THE ACCELERATED STUDY OF BIOPROCESS PURIFICATION SEQUENCES FOR IMPROVED BIOPROCESSING DISCOVERY

A thesis submitted to the University of London

For the degree of

ENGINEERING DOCTORATE

By

T. S. Reynolds

Department of Biochemical Engineering

University College London

Torrington Place

London WC1E 7JE

May 2004

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

In the biopharmaceutical industry being first to market is often key to product success. However process development still needs to deliver an economic, regulatory compliant and effective process. Scale-down studies are commonly used for this purpose since they are quicker and cheaper than pilot plant work. However it is necessary that scale-down methods be created which not only examine the conditions of process stages but also allow production of realistic output streams. These output streams can then be used in the development of subsequent purification operations and facilitate rapid and efficient process development while minimising early investment and risk.

Scale-down techniques were used to predict the effects of process changes in a plasma fractionation operation. These predictions were compared to pilot plant data and the full-scale process to assess their accuracy. The scale-down techniques used did not predict the performance of full-scale operations reliably despite being in close agreement with pilot plant results. This is due to the limitations of current scale-down techniques, which focus on a few core parameters and fail to examine engineering and operational details of larger processes.

Traditionally, predicting filtration operations is via a bench-top pressure filter, using constant pressure tests to examine the effect of pressure on filtrate flux rate and filter cake dewatering. Interpretation of the results into cake resistance at unit applied pressure (α) and compressibility (n) is used to predict the pressure profile required to maintain a constant, predetermined flux rate. This thesis reports on the operation of a continuous mode laboratory filter in such a way as to prepare filter cakes and filtrate similar to those produced at the industrial scale. Analysis of the filtration rate profile indicated the filter cake to have changing properties (compressibility) with time. Using the insight gained

from the new scale-down method gave predictions of the flux profile in a pilot-scale candle filter superior to those obtained from the traditional laboratory batch filter.

Very small chromatography columns can be useful tools for narrowing down the myriad of process options involved with this operation, but operational practicalities mean that traditional scale-down rules must be broken. Additionally a number of factors make their operation and subsequent interpretation of results problematic. By identifying and then quantifying the effects of these phenomena it is possible to compensate for their impact. By making several calculated adjustments to the outputs from a small column it was possible to predict the performance of much larger laboratory columns.

ACKNOWLEDGEMENTS

First of all I would like to thank Professor Mike Hoare for his continuous support, patience and optimism throughout this project. His thoughts and advice were of great assistance in writing this thesis. I would also like to thank Dr Mark Bulmer, who despite his busy schedule always found time to listen to my ideas or problems and whose technical expertise was invaluable during experimental work. Dr John More also deserves special thanks for giving me the opportunity to work in an industrial environment and for his constant encouragement and support.

There are many people both at UCL and at BPL to whom I am grateful. In particular I would like to thank Dr Michael Boychyn for helping me to settle in and get started on the project. I would like to thank Dr Dan Bracewell for the useful discussions on chromatography scale-down. I would also like to thank all the staff in lab 4 at BPL for their assistance and for making me feel at home while I was working there. I am also grateful for the friendship and support of the other Eng.D students: Andy Tait, Rhys Morris, Phillippa Gardner, Sally Lamping, James Myers and Greg Neal who could always make things more positive after a bad day in the labs!

Finally I would like to thank my family, Mum, Dad and Helena and my girlfriend Gemma, for their unwavering love, support and understanding throughout these last four years. Without them this thesis would never have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

Description of figure **Page**

LIST OF TABLES

Description of table **Page**

CHAPTER 1
INTRODUCTION

1.1 Introduction

The aging population of many countries combined with increasing lifestyle expectations across the world is driving the demand for new and improved medicines. Pharmaceutical companies are investing record amounts of money into research and development (R&D) as they search for novel treatments for ever more complex diseases. Recent advances in the fields of genomics, proteomics and bioinformatics have highlighted the huge healthcare potential of proteins and other biological molecules. This has resulted in a significant portion of R&D investment being used to identify new protein based drug candidates. This new generation of biological, macromolecular drugs could revolutionise the healthcare industry and the way disease is treated. However the significant challenge facing today's pharmaceutical and biotechnology companies is to develop the potential of these candidates into mass-produced, marketable drugs.

Over the past 25 years the process of bringing a new medicine onto the market has changed. The time and costs involved in developing a new medicine have dramatically increased due to the complexity of the science involved, a more intensive regulatory process and the need for more clinical trial data. However despite the early promise of many drug candidates the vast majority of them turn out to be unsuitable for therapeutic use in humans. This high failure rate for drug candidates makes it hard for companies to recoup all their R&D costs from the sales of their successful products and remain profitable. Therefore for companies to be successful in the future they will need to reduce R&D expenditure, speed up the drug development process and reduce the failure rate for promising drug candidates.

Companies are doing many different things to reduce costs and speed up the process of drug development. The area, which has received most attention is process development since it is one of the largest costs in drug development and has a lot of potential for improvement. Many drug products are approved with a manufacturing process that is less than optimal in order to get the product onto the market quicker, this generally reduces the profitability of the product. However, there are techniques which can speed up and improve process development so that a better manufacturing process can be designed in the time available. One such technique is scale-down, where industrial operations are mimicked at a small-scale in the laboratory. This allows a large number of process options to be investigated quickly and with minimal use of resources so that the best process can be selected.

This chapter will first of all give a brief background to the biotechnology industry and some of the challenges it faces. The drug development cycle will then be discussed with particular emphasis upon process development and the role it plays in getting a pharmaceutical onto the market. The next section will highlight the need to improve process development and ways in which this can be achieved. Finally the use of scale-down methods to improve process development will be discussed along with the benefits of such techniques and how they can best be used.

1.2 The biotechnology industry

Initially many pharmaceutical companies were reluctant to invest in biopharmaceutucal research since biological products were not thought to be of a high enough commercial value. However, with R&D costs soaring and a lack of promising therapeutic candidates companies are increasingly looking towards biotechnology in order to find new medicines. One recent industry study predicted that big pharmaceutical companies that

don't invest in large molecule drug discovery over the next 10-15 years will be seriously disadvantaged and may not survive (Dvorin, 2001).

1.2.1 Sources of competitive advantage

Pharmaceutical companies incur massive R&D costs in getting a therapeutic product to market, not only for developing the product itself but also to allow for the numerous failed drug candidates. The companies will only begin to see a return on this investment once a product is successfully put onto the market, however this is not an overnight job. The time taken from the identification of a promising candidate to getting a product on to the market is typically 10 – 15 years (Wheelwright, 1991). Once on the market the product must achieve a good market share in order to recoup the development costs as quickly as possible and fund research into potential new products.

In the manufacturing and sale of high quality biotherapeutics there are three main sources of competitive advantage: first to market, high quality (in terms of purity, efficacy or side effects) and low cost (Wheelwright, 1991). These advantages are however often inter-dependent thus improving one can only happen at the expense of the others. Product development therefore will often focus on achieving whichever of these is considered to be of most importance, usually speed to market.

1.2.1.1 Speed to market

All pharmaceutical companies strive to get their products onto the market as soon as possible so that they can start to generate revenue. However getting a product onto the market ahead of the competition is of particular importance since estimates indicate that the "first to market" product will claim 90 % of the market share (Wheelwright, 1991). A particular pharmaceutical will usually only enjoy a limited period of market domination

and the duration of this period is dependant on the patent protection for that product. To maximise the length of time a product is on the market under patent, speed to market is usually given top priority by pharmaceutical companies and substantial amounts of time and resources will be invested into ensuring delays do not occur. The cost of such delays can be enormous; one estimate states that for a drug with peak annual sales of $500 million, every day the product launch is delayed costs the company $1 million in lost income (Pisano, 1997).

1.2.1.2 High quality

The quality of a drug compared to its competitors may be of secondary importance if it is the first onto the market, but as patent life expires and competition intensifies its importance increases. A pharmaceutical product that is considered to be of a superior quality to its competitors will have a significant advantage. The higher quality of these products may be attributable to greater purity, greater activity, less side effects, different formulation or even different delivery methods. Products that come onto the market later may have a superior manufacturing process to earlier products and the first to market product may well lose a large portion of its sales. As a product reaches the end of its patent life it may become uneconomical as high quality competition arrives on the market. If the manufacturing process for the patented product is good then it will be able to adapt flexibly to market conditions and maintain some level of profitability.

1.2.1.3 Low costs

The significant advantage of being first to market often means that development and manufacturing costs are considered of secondary importance (Pisano and Wheelwright, 1995). This means that the manufacturing process may be hastily developed so as not to delay product launch, the result is that the product is not made as economically as

possible. This may be sustainable during patent protection but when this comes to an end and the threat of generic competition becomes of increasing concern an economic and well-designed process becomes a priority. In addition to this there is increasing pressure on pharmaceutical companies from large healthcare organisations to reduce the price of many drugs. Due to the effect this could have on a company's income it is becoming essential not just desirable for manufacturing processes to be as efficient and economical as possible.

1.2.2 The future for biotechnology

The number of biopharmaceuticals on the market looks likely to increase dramatically over the next few years as the development pipelines are packed with potential products. Recent advances in genomics and proteomics, most notably the completion of the human genome project will accelerate the identification of new protein based drug candidates. A number of therapeutic proteins are already on the market but by using genetic engineering protein structures can be modified to improve their usefulness as therapeutic agents. This will undoubtedly lead to many more therapeutic proteins becoming commercially available. However it is not just proteins that have potential as biopharmaceuticals, there is increasing interest in other types of organic molecules such as DNA for gene therapy and antisense oligonucleotides for their ability to prevent protein formation.

The challenge facing the pharmaceutical industry is to transform these potentially useful biomolecules into medical treatments to improve public health. This means getting promising candidates from the laboratory onto the market. In order to do this, pharmaceutical companies will have to overcome a number of new challenges since these bio-molecules are inherently different from the traditional small chemical entities

(SCE's) they are used to producing. Biological molecules are much larger than chemically manufactured molecules and have an inherent variability that is difficult to characterise. The size of biological molecules makes them particularly susceptible to damage from shear stresses and they also tend to be sensitive to heat and extremes of pH. For these reasons manufacturing processes for biotherapeutics need to be carefully designed so that the product molecule is not damaged and maintains its biological structure and function.

The process of taking a biotherapeutic from discovery through process development and clinical trials to regulatory approval and commercialisation is long, difficult and expensive. On average it will take 10 – 15 years to develop a new medicine and the development costs will be around $800 million (Pharmaceutical Industry Profile, 2003). However, since over 90% of new medicines are expected to generate sales of less than $180 million per year it is important for the future success of the biotechnology industry that R&D costs for each product are reduced significantly.

1.3 Development stages in bringing a new therapeutic to the market

This section gives an overview of the product development cycle and the various activities involved. The development cycle of a pharmaceutical product takes it from discovery through clinical trials to market launch and then on until manufacturing is stopped either due to lack of demand or for economic reasons. Within the life cycle of a product exist several distinct stages (figure 1.1), which a company will have to deal with as a product matures. The skills required for a company to negotiate successfully each stage are different and hence some companies are better equipped to tackle certain stages of the product life cycle than others. How a pharmaceutical or biotechnology company uses its skills in developing the product life cycle will ultimately determine its success.

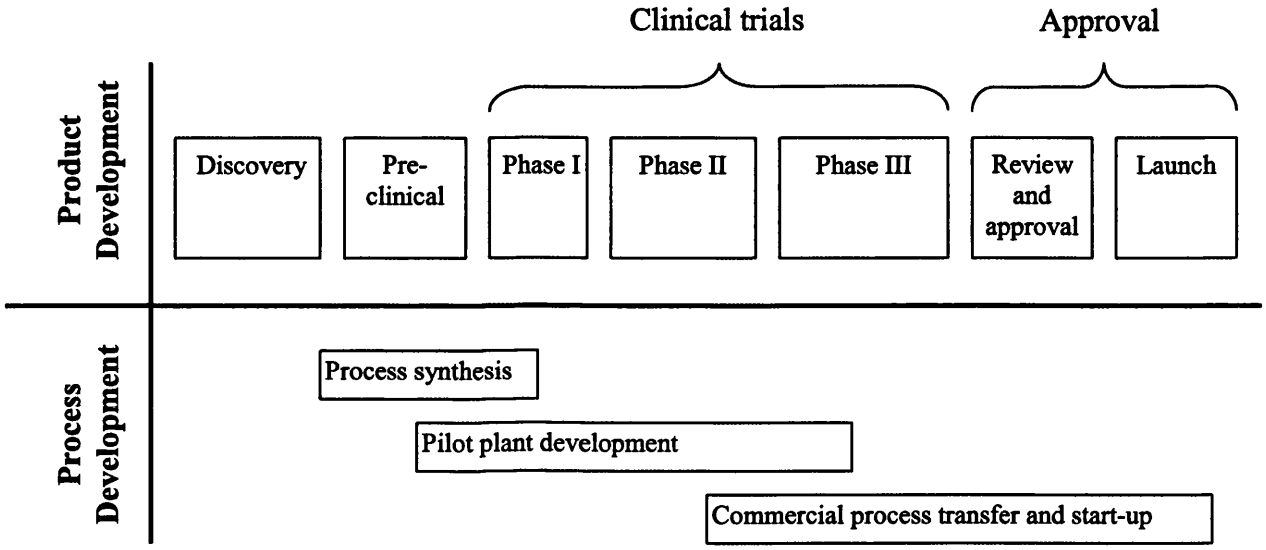

Figure 1.1: The key stages in product development and how the coincide with process development.

1.3.1 Drug discovery

The life cycle of a pharmaceutical product begins with the identification of a molecule that has some function in preventing or treating a medical condition. Understanding the biological causes of disease is key to identifying molecules that may help to control the disease. The human genome project has greatly advanced the level of understanding about the genetic causes of disease and actually identified the genes for 1400 different diseases (Wechsler, 2003). This combined with other recent advances in genomics and proteomics should accelerate the process of identifying proteins with potential therapeutic uses.

Other developments that will have an effect upon drug discovery are the emergence of combinatorial chemistry and the development of ultrahigh-throughput screening techniques. These techniques allow researchers to test rapidly and assess the therapeutic potential of a very large number of molecules. Generally simple automated tests are employed which can identify the most promising drug candidates with minimal input

from the scientist. This has enabled small research groups to increase greatly the number of molecules they can test in any given time period and thus identify many more compounds that warrant further study. The knock on effect is that development teams have many more projects on-going at any one time and the product development bottle neck shifts away from drug discovery and into the later stages of product development.

The abundance of potential therapeutic molecules and ultrahigh-throughput screening techniques should allow companies to select the most promising candidate from a range of potentially useful molecules and improve the success rates of bioprocess development. Figures typical for the industry suggest that of 5000 molecules screened only 5 will enter into clinical trials and of these only 1 will actually reach the market (Pharmaceutical Industry Profile 2002). This level of attrition is unsustainable and there is a need to streamline the drug discovery and development processes so that candidates with not only the best therapeutic properties but also the best processing and clinical properties can be identified as early as possible.

1.3.2 Pre-clinical trials

Once a potential drug candidate has been identified it enters into pre-clinical trials. The main objective of these trials is to test the toxicity of the candidate using animal models. An initial study into the pharmalogical and pharmacokinetic properties of the molecule is also performed and potential delivery routes may also be assessed. Based on this information a company will decide if the candidate has enough potential to justify proceeding onto human clinical trials. Struck (1994) estimated that the chance of a biopharmaceutical successfully completing pre-clinical trials was 54% and the chance of a biopharmaceutical making it all the way from pre-clinical trials to launch was 36%.

This is significantly better than chemical entities for which the probability of passing all the way from pre-clinical trials to launch is only 11%.

The number of molecules in pre-clinical trials will be much higher than the number of molecules in the subsequent clinical trials and the cost of development per molecule will be relatively low. Currently the most expensive stages of drug development are Phase II and Phase III clinical trials. However, due to changes in R&D practice it is predicted that by 2005 pre-clinical trials will be the most expensive stage and the number of molecules failing at this stage will increase to 75% (Arlington et al, 1998). This is because new technologies and an improved understanding of molecular biology should enable the early identification of molecules that will not make it to market, thus reducing the number of molecules in the later more expensive stages.

1.3.3 Clinical trials

Clinical trials are a regulatory requirement and often represent the most costly and resource intensive part of process development. The purpose of clinical trials is to test potential treatments in human volunteers to establish whether or not they are safe and effective. The trials consist of three sequential phases in which the number of patients and hence the amount of material and cost increases as the phases progress. However the chance of the product successfully completing each phase also increases (Struck, 1994). Careful planning, execution and analysis of clinical trials are essential to convince regulators that the drug is both safe and beneficial to the patient. The participants in the trial must also be carefully chosen to represent a diverse section of society since drugs can effect people of different age, race or gender differently. Particular care needs to be taken when dealing with a treatment that is intended for use on children and there are some specific guidelines for carrying out such trials. Table 1.1 gives typical values for

the number of patients, amount of product, duration and cost for each stage of clinical trials along with the chance of successfully completing each phase.

	No. of patients	Product required (kg)	Average duration (years)	Cost (millions $)	Chance of success (%)
Pre-clinical	na	0.01 – 0.10	2.3	66	57
Clinical Phase I	20 – 80	0.1 – 0.5	1.8	67	88
Clinical Phase II	100 - 300	0.2 – 1.0	2.2	167	86
Clinical Phase III	1000 - 5000	0.5 – 5.0	2.0	150	93
Approval	na	na	3.8	50	100
Total	≈ 2690	≈ 3.3	12.1	500	40

Table 1.1: The phases involved in testing a biopharmaceutical therapeutic product before it can be launched on the market. Data in columns 1and 2 is derived from the pharmaceutical industry profile 2003, columns 3 and 5 are from Struck (1994) and column 4 is from Arlington (1998).

In Phase I a small number of healthy human volunteers (usually between 20 and 80) are administered the drug to evaluate its safety in terms of possible side effects, study how it is metabolised and excreted and make an initial assessment of dosage. This phase usually requires only a small amount of material that can be made in a laboratory or pilot plant. Phase II clinical trials involve testing the drug on a group of humans who actually have the condition or disease to be treated. The test group usually consists of several hundred people and is used to determine the dosage and efficacy of the drug. A currently available treatment is normally used as a standard although placebos can be used if it can be justified on ethical grounds (Rados, 2003). Phase III trials are used to determine definitive safety and efficacy of the drug as well as to monitor any long term implications and less common side effects. Phase III trials will last for several years,

require several kilograms of purified product and will usually involve testing on several thousand patient volunteers. The material used in these trials must be made according to current good manufacturing practices (cGMP) and as such the pilot plant process used must be completely validated.

As a molecule progresses through the various stages of clinical trials the chances of it actually reaching the market increase (see table 1.1). The main reasons for failure during clinical trials are lack of efficacy, toxicity and poor biopharmaceutical properties. The biopharmaceutical properties that a company will look for in a molecule include chemical stability, good aqueous solubility, good cell permeation properties and a low rate of metabolism by the liver or removal by kidneys (Borchadt, 2001). These properties are sometimes called "drug like characteristics" and if these can be assessed early on in the development cycle then poor candidates can be eliminated before money is wasted on clinical trials.

1.3.4 Review and approval

If the drug candidate successfully completes phase III clinical trials then the company can submit a new drug application (NDA) to the regulatory authorities. This will consist of all the data from pre-clinical and clinical trials along with details of the manufacturing process and proof that the process will consistently produce a pure, safe and reliable drug. The regulatory authorities will then review the application and if they are satisfied with the quality and quantity of data supplied the drug will be approved for sale. Between 1983 and 1991 the typical review period for a biopharmaceutical product was 1.6 years, however there is some evidence to suggest that this review period has decreased slightly since then (Struck, 1994).

1.3.5 Process development

During process development various ways to make, capture, purify and formulate the drug are investigated. The overall goal is to make as much of the product as required at the necessary quality, as cheaply as possible and in the shortest amount of time possible. This usually involves designing a suitable fermentation to make the product and then specifying a series of steps to capture, concentrate and purify the product. Once clinical trials begin it becomes very difficult to alter the process due to regulatory restrictions. Having a high performance, robust, economic and well-understood process established early on can avoid costly delays later on in development (Bobrowicz, 1999).

1.3.5.1 The role of process development

The role of process development varies somewhat throughout product development. One of the earliest tasks is to develop assays to test for the product and contaminants, as these are the eyes and ears of the scientist. Small, purified quantities of the product are also produced in a laboratory so that initial testing and evaluation of the product can be performed. At this stage there is often little or no attention given to processing considerations since the main goal is to determine whether or not the product demonstrates suitable therapeutic and pharmacological potential to be studied further. However a well-designed manufacturing process can be a valuable source of competitive advantage, unfortunately many of the techniques used in laboratory purification are unsuitable for large-scale operation. Wheelwright (1987) developed a set of heuristics to assiast with the design of large-scale purification processes. These heuristics give useful guidance for designing a process in the lab that will be effective, economic and scalable.

The first stage in the manufacture of most bio-pharmaceuticals is to actually make the product, which usually involves some kind of fermentation. A suitable cell line needs to

be selected that combines good growth characteristics with high expression levels for the desired product. The fermentation then needs to be designed in terms of the conditions used to grow the cells in a reproducible manner. After fermentation a series of operations are needed to capture purify and concentrate the product. A wide variety of operations exist for doing this and it is the task of the process development team to select and optimise the best sequence of operations.

As soon as a product demonstrates a suitably high level of clinical promise (i.e. positive results in pre-clinical trials) the demands of process development intensify significantly. The aim is now to provide sufficient quality and quantity of material for clinical trials. Pilot plant facilities are used to do this since the quantities involved are too big to produce in a lab and regulatory requirements mean that the material must be produced to cGMP standards. The pilot plant operation will usually differ significantly from the laboratory process since many laboratory methods are not practical or economical at large scale. As a result some parts of the process may need to be redesigned and optimised in the pilot plant. However since pilot plant trials are expensive and time consuming (Zhou et al 1999) the number of process routes explored and the number of optimisation runs is usually severely restricted. In addition the process needs to be "fixed" and validated in order to produce material for Phase III clinical trials, if it isn't then the trials will have to be delayed. Once the trials do begin the focus of process development becomes scaling up and transferring the process to the full-scale facility. It is important that the selected process route is scalable because if the full-scale process performs differently from the pilot plant process then equivalency studies and revalidation will be necessary, which will delay product launch.

A company needs to find the best possible process in terms of yield, throughput, purity robustness and cost as quickly as possible. This requires substantial investment into process development, however companies often try to minimise the level of investment since the risk of product failure is still significant. It remains to be seen if improved screening techniques will reduce the risk of product failure in clinical trials and thus encourage companies to invest more in process development earlier. Currently the window for process development is fairly small since the process needs to be frozen early on in Phase I clinical trials and must be completely validated by the start of Phase III. Since clinical trial times are shortening due to more efficient design, improved data handling and streamlining of regulatory guidelines the window for process development is also shrinking. Improving the efficiency of process development can improve process productivity and avoid delays during clinical trials since the production of material for use in clinical trials is often the limiting step. Effective process development will also help to avoid scale-up problems and facilitate easier process validation.

1.3.5.2 Issues specific to bio-process development

Process development for chemically derived drugs (commonly called small chemical entities or SCE's) was traditionally left until late in product development since the chance of the molecule entering into clinical trials was extremely low. The process was often hastily designed and not optimised. However since the manufacturing costs for SCE's are often less than 10% of the products sales revenue, a low efficiency process will only have a small effect upon profits. This is not the situation with biological therapeutics however since manufacturing costs are usually in the region of 20 – 25% of total sales. This means that for basic economic reasons bioprocesses need to be efficient and well designed.

Pisano (1997) pointed to three main areas in which bioprocess development differs from more traditional process development for SCE's. Firstly the theoretical understanding of chemical processes is far more advanced than that for biological processes. Living organisms and the molecules that enable them to function such as proteins are extremely complex, much more so than the molecules encountered when chemically synthesising a drug. It is only over the past few decades that scientists have really started to discover how biological systems function at the molecular level and it is still unclear how this knowledge can be used to understand and design better bioprocesses. A consequence of this is that theory based predictive tools such as computer models are rarely used in biotechnology and experimentation is essential to overcome a lack of fundamental knowledge (Evans, 1988).

The second way bioprocess development is different is that it is difficult to characterise products and intermediates. Many proteins are chemically similar and need very specific techniques to identify them. This problem can be particularly troublesome as some proteins such as Human Immunoglobulin G (IgG) exist as a number of different subclasses. These subclasses can have different biological activities even though physically and chemically they are very similar. In addition, small conformational changes in a proteins structure can lead to loss of activity, but this change may not alter the response to analytical assays.

The third and final way in which bioprocess development is different from SCE process development is the occurrence of secondary scale-up effects. Although bioprocesses can be scaled up successfully by maintaining key parameters often secondary scaling effects can cause problems with the process. This is because biological molecules are very sensitive to the environment they are in and during scale-up there are inevitably some

small changes to the molecular environment. It is very difficult to predict the effect these changes will have on biological molecules and as a result bioprocess design still relies heavily on experimentation and to some extent on trial and error.

1.3.5.3 Process development for an existing product

The role of process development does not end once a product is on the market. Although altering a process once it has been validated is expensive as equivalency studies must be performed and revalidation may be necessary, sometimes it is beneficial. As a product nears the end of its patented lifetime the threat of cheaper, generic competition becomes a reality. If a competitor can produce a similar product cheaper or to a higher quality then they will gain the larger market share. In order for an existing process to remain economically viable the manufacturing process may need to be improved so that it is more cost effective.

A company may also find themselves in a position where they need to increase (or decrease) the production capacity of a process in order to meet market demands. If the process was well designed and is well understood then the task of scaling up the manufacturing process would be fairly straightforward. However a process that was hastily designed or is not fully understood could prove very difficult to scale up. Generally speaking good process development allows a manufacturer more flexibility with a process so that it can react better to changing market demands.

1.4 Improving process development

The practice of focusing on speed to market at the expense of process development can have significant financial consequences for a company as well as often delaying product launch (Bobrowicz, 1999). However increasing economic pressure on pharmaceutical

companies means that the need to get new medicines from the laboratory and onto the market is greater than ever. Current project failure rates are unsustainable and revenues from new product sales are not sufficient to keep R&D funding at its current level. Therefore in order for pharmaceutical research to continue and for companies to remain profitable R&D costs need to be reduced.

Since companies now have an abundance of new drug candidates to choose from the bottleneck in getting new products onto the market is likely to be process development (Nayar, 2002). The time and resources required for good process development are considerable, which means companies are reluctant to invest early due to the risk of project failure. In order to deal with these issues companies have developed a range of strategies that enable effective process development to occur while reducing the time and resources needed and while increasing the chance of project success. These strategies combine both new technological and new organisational approaches to process development. The next few sections will discuss these strategies.

1.4.1 Organisational approaches to improving process development

Many companies have attempted to streamline the drug development process by integrating drug discovery and process development at an early stage. Whereas there used to be an "over the wall" attitude between different groups in the product development cycle now many companies have interdisciplinary teams that work on a product from discovery up until launch and often beyond. This approach improves communication and facilitates the rapid exchange of information between individuals involved in the project. For example, large-scale processing limitations are considered much earlier in the project resulting in fewer problems later on. The overall result is that attrition rates decrease since scientists can study a wider range of criteria to decide

whether a molecule is worthy of further development. Also the task of process development tends to be more efficient since development scientists already have some knowledge of the product and process development begins earlier. The main disadvantage of this streamlining is that more resources are used during the early stages of a project.

Instead of streamlining product development some companies have opted to focus on their core competencies and outsource parts of product development. It is fairly common for small and medium sized pharmaceutical companies to use contract manufacturers to do their process development. In this way they can benefit from the knowledge and experience of the contract manufacturers and also avoid any capital expenditure. Of course contract manufacturers are not always an option since they can be expensive and only have limited capacity. In addition to this there can be communication problems between contractors and clients and the client often misses out on process knowledge and experience that could be invaluable in the future.

1.4.2 Technical approaches to improving process development

In the pharmaceutical industry process design has traditionally been a fairly empirical process relying heavily on expensive pilot plant trials, this is particularly true for biological products. Therefore innovative new technologies are being used to enhance and accelerate process development. The objectives are to obtain more information using less resources and in less time and to make better use of this information. Currently there are three major approaches to achieving this: computer modelling, expert systems and scale-down. These technologies can all play a role in improving process development and they are discussed in more detail below.

1.4.2.1 Computer modelling

The chemical manufacturing industry uses computers routinely in designing, evaluating and optimising manufacturing processes. However the use of computer models in biotechnology is quite rare. Since biological systems are complex, difficult to characterise and there is often a lack of reliable physical property data it is hard to construct an accurate and robust computer model. However, much work has been done to overcome these problems and computer models are gaining more widespread use in biotechnology. The role of computer models is quite varied. Some models can predict the performance of a unit operation under various conditions while others try to model a whole sequence of operations and select the best overall process route, others are used to assess process economics and optimise equipment scheduling.

The use of powerful design of experiments (DOE) packages has also led to an increase in computer usage during process development. These packages help scientists to determine quickly what the critical parameters are in a system and how these parameters interact. As with a lot of computer models there is still a need for some experimentation to obtain critical input data. The quality of this data will often determine the usefulness of the model, but it has been shown that combining some experimental data with process models can improve their accuracy significantly. If process models can be developed that reliably predict process performance then they will undoubtedly gain wide spread use since the savings in time and resources can be substantial.

1.4.2.2 Expert systems

Expert systems electronically compile and evaluate data on a certain process. They then use modelling, expert rules and heuristics to make important process decisions. A number of such systems are available for chemical processes. However until recently the

lack of experience and expert knowledge on bioprocessing has hindered their use in bioprocess development. This is starting to change as companies recognise the value of experience and seek to learn from it. Lienqueo (1996) reported on the use of an expert system to select a multi-step protein separation process for somatotropin produced in *E. coli*. Using basic physiochemical data about the product and major contaminants the expert system selected a sequence of steps very similar to an existing industrial process in around 10 minutes. This type of simulation clearly has the potential to save a lot of time and resources in process development. However they still require some experimental input and as processing problems become more difficult and novel techniques are used expert systems become inappropriate due to a lack of expertise.

1.4.2.3 Process scale-down

Process scale-down attempts to replicate the true performance of industrial operations in a laboratory. The usual strategy for scaling down a particular industrial operation is to identify the key process parameters that impact upon performance and mimic these parameters in laboratory scale experiments. All the major unit operations used in pharmaceutical manufacture including those used in bioprocessing have been scaled down with some degree of success. However laboratory scale-down mimics often perform differently to their industrial scale counterparts, which can cause problems on scale-up. This problem is compounded by the complex nature of many biological materials, which makes their behaviour and interactions with their environment difficult to predict. However, scale-down techniques are improving so that they can better predict the performance of large-scale operations. Consideration of engineering and operational issues at laboratory scale can also enable more accurate predictions and facilitate efficient process design (Wheelwright, 1987).

Good scale-down models can reduce the need for lengthy and expensive pilot plant trials by allowing much of process development to be performed in a laboratory. Scale-down experiments are generally used to assess different process options, optimise unit operations, evaluate robustness and accelerate validation. The principal advantage of scale-down is that process development can begin at an early stage in product development since only small amounts of material are required. The minimal use of resources also reduces the investment involved in the early high-risk stages of a project. Other benefits of scale-down include simple handling and operation, reduced energy usage and the ability to carry out a number of experiments very quickly. The biggest limitation of scale-down is the accuracy of the laboratory mimics and demonstrating to regulatory agencies that the mimics are capable and reliable in producing data representative of the large scale.

1.5 The use of scale-down in biotechnology

The complex nature of biological materials makes them hard to characterise and their behaviour difficult to predict. The lack of predictive tools has meant that bioprocess design has been a largely empirical process relying heavily on pilot plant trials and the experiences of the design team. Since pilot plant trials require significant amounts of time and resources small–scale laboratory experiments are commonly used to investigate process options and evaluate performance. One of the first reported uses of scale-down techniques in biotechnology was by Oosterhuis et al (1983). They found that by mimicking in small laboratory vessels the fluctuations of nutrient concentration that occur in large fermenters they were able to reproduce the performance of the large fermenter. Since this time similar theories have been applied to other unit operations used in bioprocessing to try and identify the parameters critical to performance and reproduce them in a laboratory.

The phenomenon that large-scale bioprocess operations generally display worse performance than equivalent laboratory operations complicates the scale-down process. Simply scaling all the parameters up or down on a linear basis generally will not result in the same performance and is often impossible due to fundamental differences in equipment design between laboratory and industrial scale. The critical process parameters and conditions that influence the product environment at the industrial scale must be identified. This environment must then be replicated in the laboratory by manipulating the critical parameters.

1.5.1 Developing scale-down models

Before creating a scale down model for a particular operation a thorough understanding of the industrial operation is needed. This should identify the important operating parameters as well as highlight differences between the industrial operation and commonly used laboratory equipment. After this the basis for scaling needs to be identified; what aspects of the large-scale operation need to be reproduced in the scale down mimic. Generally the simplest way to scale down an industrial operation would be to build a working scale model of the operation in the laboratory and simply scale down the important operating conditions. However this is often uneconomical and for some equipment such as high-speed centrifuges and filter presses this is simply not feasible.

There are two objectives when creating a scale-down model, the first is to identify the scale dependent aspects of a large-scale operation and the second is to keep the number of parameters to be scaled to the minimum possible. A number of strategies exist for scaling down various operations these include: fundamental methods, semi-fundamental methods, dimensional analysis, regime analysis, heuristics and trial and error (Atkinson and Mavituna, 1991). Which one of these strategies works best will depend upon the

operation being scaled down and the intended use of the scale-down model. The most thorough strategy is fundamental analysis, which is widely used in chemical engineering. This strategy relies upon completely solving fundamental equations such as mass balances, energy balances and chemical equilibria that describe the process. However the number of parameters and sheer complexity of these equations make them difficult to solve. Semi-fundamental strategies use simplifying assumptions (e.g. well mixed vessels) in order that the equations can be more easily solved; despite this the mathematical formulae involved are still quite complex, making practical use difficult. These sorts of approach have very limited use in biotechnology since the inherently complex nature of biological processes makes the application of fundamental equations very difficult.

Another strategy that has been extensively used in chemical engineering is dimensional analysis. In this strategy the process is described by a number of dimensionless groups of parameters, which represent the ratios of various process variables. The values of the dimensionless groups are kept the same at both scales so that the relative importance of the parameters and their influence on the process is maintained (Zlokarnik, 1983). Realistically it is impossible to maintain all the dimensionless groups so the most influential ones must be identified and the rest are ignored. Instead of dimensionless groups of parameters regime analysis is based upon maintaining characteristic times in the process that govern process performance. Since these characteristic times can often be determined easily, either from simple experiments or from the literature this method is particularly suitable for use in bioprocess scale-down.

Often the strategy used for creating a scale-down model is a combination of all the above methods as well as some degree of experimentation. Fundamental equations and

dimensional analysis are often useful in determining the scale dependant process parameters; then experimental techniques are used to determine the influence of these parameters on the product. Whatever the technique used the success of the scale down model will depend upon its ability to mimic the performance of the large-scale operation.

Even though determining the critical process parameters of an operation and mimicking the product environment are vital in developing a good scale-down model sometimes doing just this is not enough. The equipment used in large-scale industrial operations is fundamentally different from the experimental apparatus used in a laboratory. Industrial equipment is usually made from stainless steel and the processes are heavily automated and controlled. In contrast laboratory equipment is often made of glass or plastic and the level of control and automation is much lower. The effects of this on performance may be negligible but they should be considered when comparing a scale-down model to a larger operation. For example the location of sample points and instrumentation can affect the results they yield and biological products have a tendency to adsorb to certain materials. Thus in order to develop an accurate scale-down model it is often useful to consider not just a specific piece of equipment but also the ancillaries such as pumps, piping, valves and detectors.

1.5.2 Applications of scale-down

Scale-down models can fulfil a number of roles during process development and throughout a products lifetime. Traditionally there have been two main uses for scale down technology: these are to facilitate scale up of a new process and to evaluate modifications to an existing one. However as scale-down technology improves they are also becoming useful tools during the early stages of process development and during

process validation. The next few sections discuss the various applications of process scale down.

1.5.2.1 Scale-down use in process development

Scale-down can be a useful tool during several stages of process development. The requirements of scale-down may differ between stages and hence the accuracy and investment in a scale-down model will also vary. During the early stages of process development the aim is to determine a series of steps that will capture the product from the fermentation broth and concentrate and purify it to the required level. Since there are a wide variety of alternative operations that could be used to achieve this there is a need to rapidly screen a large number of options. Scale-down techniques can be used to assist in this task by rapidly assessing the feasibility of a number of different operations. Then comparing results and / or using heuristic rules can select the best options. Since at this stage of process development the objective is simply to determine which operations are actually feasible, the accuracy requirements are low and the investment in the scale-down model will also be low.

As a product prepares for entry into clinical trials the process development team must transfer the manufacturing process that was used in the laboratory into the pilot plant. This change of scale inevitably leads to some undesirable changes in performance due to the different environmental and operating conditions. These changes can be difficult to predict and may delay the production of clinical trial material while the process is adjusted to rectify its performance. The use of suitable scale-down models can predict these changes in performance. Therefore any surprises when scaling up are avoided and the production of clinical trial material is not delayed. Obviously the level of accuracy required from the scale down models is much greater at this stage than during process

synthesis. However, the availability of resources is also likely to be greater since the importance of getting the process right at this stage is considerable and delays to clinical trials are highly undesirable.

1.5.2.2 Scale-down of an existing bioprocess

A scaled down model of an existing bioprocess can be a very useful tool. If the model is sufficiently accurate then a company could demonstrate and validate its equivalence to the production scale process. Obviously the requirements of a scale-down model to do this are significant since the model will need to accurately simulate all aspects of the performance of the large-scale operation. The potential uses for such a model include evaluating the sensitivity of an operation to changes in feedstock, process optimisation and assessing the impact of any suggested process changes. In more general terms a scale-down process model should allow a company to develop a better understanding of its own processes. The development and validation of such a model can be difficult and resource intensive but the savings made by minimising disruption to the manufacturing process usually justify this.

1.5.2.3 Scale-down as a tool for validation

A good scale-down model of a process or operation can be a useful tool for assisting validation. The accuracy required of a scale-down model used for validation purposes is very high. A company must demonstrate equivalency in terms of performance between the model and the full-scale process. However this can remove some of the burden of validating the full-scale process. For example, part of validation involves carrying out a minimum of three verification runs on the process to determine the variability of data and show that the process is reproducible. However in many cases three runs are not sufficient due to performance differences on scale-up which mean that product

specifications or processing conditions have to be adjusted. Scale-down models can be used to set more realistic specifications so that only the minimum of three verification runs are required. This results in considerable savings in time and resources.

Scale-down models are also very useful for carrying out viral clearance studies. The removal of adventicious agents such as viruses from a biological product is critical if patient safety is to be assured. Before a biological product is approved the manufacturer must demonstrate to the regulatory authorities that a sufficient level of virus clearance is incorporated into the process (generally for mammalian cell cultures only). The most comprehensive way to measure virus clearance is using virus spiking studies, where a known virus titre is added to the feedstock for a step and the number of viruses present in the feed stream after that step are calculated. Performing this type of study in a full-scale plant would be dangerous and it would be nearly impossible to prove that all the viruses have been removed from the facility. However an accurate scale-down model of the step could be used to determine the level of virus clearance that could be expected in a much safer and controlled environment (Morfeld *et al, 1996).*

1.5.3 Scaling-down common bioprocessing operations

The vast majority of bioprocesses can be divided into two distinct areas. The first part is fermentation (sometimes referred to as up-stream) where living cells are cultivated in order to produce the desired product. The second part is down stream processing where the product is captured from the fermentation broth, concentrated and purified to the required specifications. The exact stages involved in downstream processing will depend upon the physio-chemical properties of the product and whether or not the product is excreted from the cell. Figure 1.2 shows the stages involved in a typical bioprocess along with the unit operations commonly used at each stage. A large number of operations can

be used in the downstream processing of biological products however despite the variety of options available the purification sequences for different products are often quite similar (Bonnerjea, 1986).

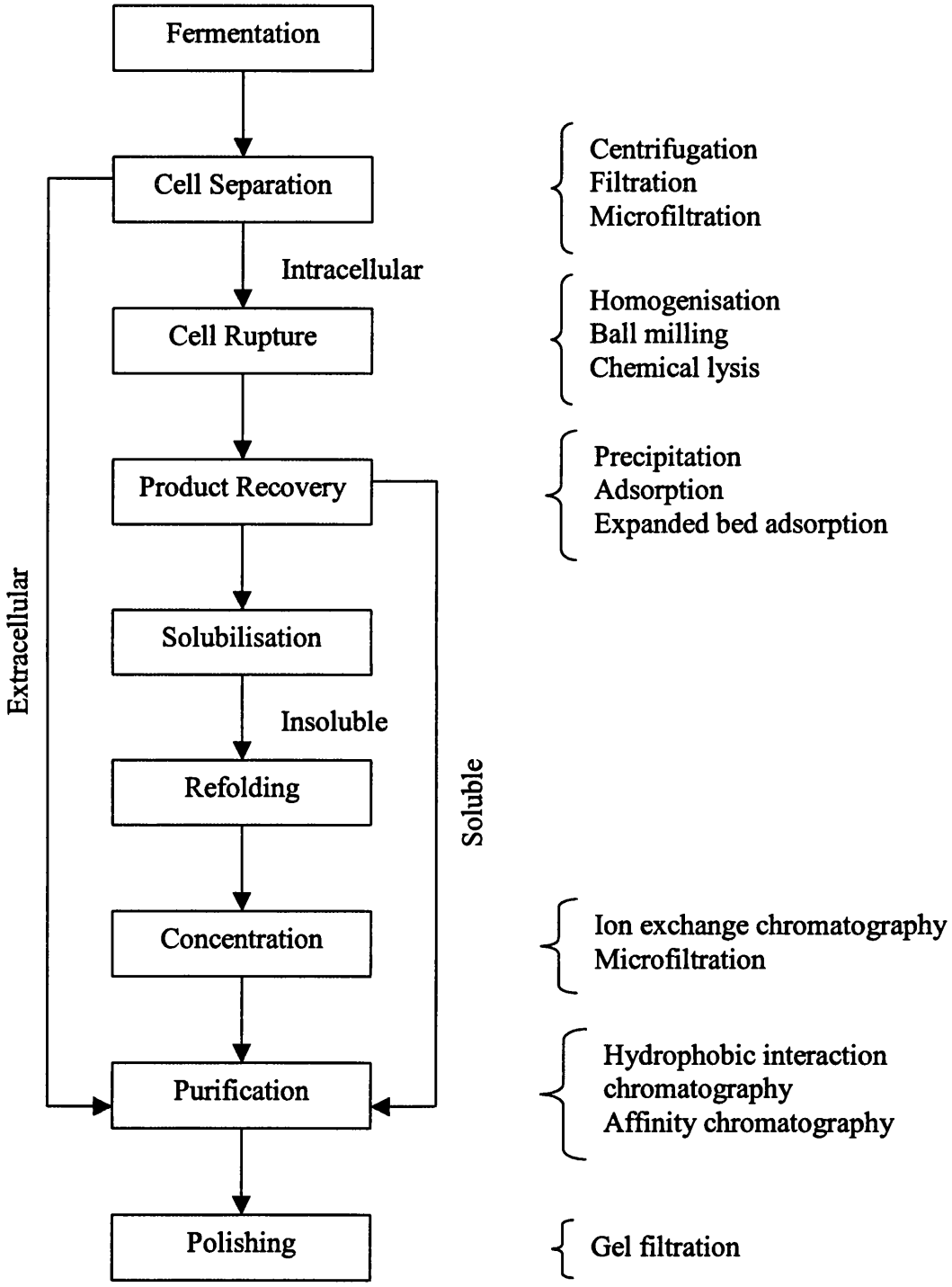

Figure 1.2: The various stages in a typical bioprocess and the unit operations commonly used at each stage. The exact nature of the process will depend upon the location and properties of the product molecule.

The downstream processing of biotherapeutics tends to rely on classical purification techniques since they can handle large volumes, are insensitive to non-protein contaminants, are scalable and are generally well understood. Although several novel technologies such as expanded bed chromatography, two phase aqueous extraction and monoliths have shown potential, there is a reluctance to use them at large scale due to uncertainties about validating them. On the other hand the validation requirements for established purification techniques are well known and basic scale-down models for these operations can be found in the literature. The next few sections will give a brief outline of these traditional scale-down models. However as economic pressure increases the requirement for high yield, low cost processes the use of novel technologies could increase. If this happens it will become necessary to develop scale-down models for each of them so that they can be assessed in the laboratory along with other operations.

1.5.3.1 Fermentation

The difference in performance between laboratory shake flasks and industrial fermenters is considerable. Growth rates in shake flasks are often lower than those of large fermenters due to problems with oxygen limitation and depleted nutrient concentrations in the smaller vessel. This has lead to the development of miniature bench-top fermenters complete with impellors, aeration devices, baffles and all the other components of larger fermenters. However these scale-down devices often perform better than their larger counterparts due to the fact that conditions in large fermenters are often non-homogenous. Factors that change on scale-up leading to non-homogenous conditions include oxygen transfer, heat transfer, quality of mixing, shear levels, superficial air velocity, time of inoculum transfer and age and stability of the culture (Humphrey, 1998). In order to produce more accurate scale-down models for fermentation these non-homogenous variations need to be mimicked (Oosterhuis, 1995). This can be done by

combining a number of ideal laboratory fermenters to mimic different conditions in an industrial fermenter. These laboratory devices can predict the performance of larger fetmenters quite well but they are often expensive and can be awkward to use. Accurately measuring and controlling the fermentation conditions at this scale can also present difficulties.

1.5.3.2 Homogenisation

Hetherington (1971) found that the degree of cell breakage in a high-pressure homogeniser is a function of pressure and the number of passes through the valve. By maintaining these two parameters constant it is possible to scale-down homogenisation operations to the laboratory scale. Siddiqi (1997) found that during homogenisation the total protein released and the particle size distribution of cell debris was the same at three different scales ranging from 40 mL to 30 L. Since homogeniser performance has been shown to be independent of both flow rate and valve geometry small-scale homogenisers can be operated in batch mode and still predict the performance of larger continuous machines. However it is important to bear in mind the exact source of the material being homogenised. Fermentation conditions can have a large effect on the strength of microbial cell walls and hence the ease of cell rupture. The cells used in the laboratory must have similar (ideally identical) properties to those grown in industrial fermenters or else homogenisation performance on scale-up may change drastically.

1.5.3.3 Precipitation

In order to scale-down an industrial precipitation operation the physical and chemical environment present in the precipitation vessel must be reproduced in the laboratory. The scale-down model will need to mimic both the changes in protein solubility and the particle size distribution produced in the industrial operation. Maintaining geometric

similarity is important for mimicking the fluid dynamics of the large vessel. This can be achieved relatively easily and a large variety of laboratory scale mixing vessels and impellors are available commercially. Other parameters that need to be maintained during scale-down are the mean velocity gradient, Camp number and the rate of reagent addition in relation to the suspension volume (Bell and Dunnill, 1982). The sensitivity of a system to these variables varies a great deal and hence realistic process material needs to be used in scale-down studies. Although many protein operations have been scaled successfully sometimes problems arise due to variations in the local environment at large scale. It is very difficult to mimic these conditions exactly in a laboratory but by considering factors such as position of reagent addition and rate of heat transfer the performance gap between scales can be minimised. Boychyn (2000) successfully scaled down protein precipitation (1000-fold) based on constant mean velocity gradient and ageing. The properties of the precipitate particles were found to have a significant impact on subsequent centrifugation and thus need to be carefully mimicked during scale-down.

1.5.3.4 Centrifugation

A common method for comparing different centrifuges is the Sigma concept, which is defined as the area of gravity settling tank required to perform the same separation (Ambler, 1959). Correction factors have to be introduced to account for areas of non-ideal flow, which then allow centrifuges of different geometries to be compared. Scale-down of centrifuges can be achieved by simply maintaining the ratio of flow rate to sigma. Since the vast majority of small-scale centrifuges operate in batch mode while large-scale ones are continuous several other factors need to be considered. Firstly the relatively short spin times used with laboratory centrifuges mean that acceleration and deceleration times can be significant and sigma values need to be adjusted accordingly (Maybury, 2000). Secondly the sediment dewatering in laboratory centrifuges is lower

than that occurring in industrial machines since scaling using sigma theory is based upon liquid residence time (flow rate) not solids residence time. This can be overcome by removing the liquid from a batch centrifuge after spinning and then spinning the sediment again to match the solids residence time in an industrial machine. The final consideration is the high shear stresses that materials are exposed to in the feed zone of large-scale centrifuges. These shear stresses alter the particle size distribution by breaking up larger particles, thus making their removal more difficult. Boychyn (2001) overcame this problem by exposing slurries to high shear stresses in a purpose built laboratory shear cell prior to centrifugation. This technique allowed the performance of an industrial centrifuge to be predicted in a laboratory.

1.5.3.5 Filtration

Filtration scale-down has traditionally been based upon maintaining the ratio of slurry volume to filter area and maintaining the applied pressure. Experimental studies using realistic process material are vital for filtration scale-up due to the limitations of filtration theory and the evolving nature of the filter cake (Wolthius and Dichiaria, 1997). Simple laboratory batch filtrations are performed using constant pressure tests to examine the effects of pressure on the filtrate flux rate and filter cake dewatering. The results from these tests can be used to predict the flux profile obtained from a known pressure profile in an industrial filter. However the complex interactions between process parameters and subtle differences in cake formation can cause problems on scale-up. Chapter 3 of this thesis discuses filtration scale-down in more depth and describes the use of a new scale-down model to predict the performance of large-scale filtration operations.

1.5.3.6 Membrane filtration

Accurate scale-down of membrane filtration operations can only be achieved by maintaining fluid dynamic parameters that are independent of scale (van Reis, 1997). The fluid dynamics of a membrane system are dependent on the operating parameters, contact materials and system geometry of that particular system. Traditional scale-down techniques have focused on maintaining the ratio of filtrate volume to membrane surface area, membrane material and pore size, channel height and flow path type and filtrate and retentate pressures constant. The width and number of channels are reduced in order to reduce the overall membrane filtration area. However even if all the above parameters are scaled down accurately problems can still arise on scale up. One common cause of scale-up problems is that the channel length is increased. This can lead to increased pressure drops, reduced retentate flow rates and vortex formation which all affect the overall performance of the operation (van Reis, 1997). Other areas that require consideration when developing a scale-down model include flow distribution, channel height compression and entrance/exit effects. Since many industrial membrane filtration operations are scaled-down by simply reducing the number of membrane cassettes involved it is important to consider the manner in which the cassettes are set up. For example the cassettes are commonly run in parallel to increase filtration area but often feed ports run in series thus the pressure and fluid dynamic properties of the feed vary between cassettes. This means that as with most unit operations it is important to consider the equipment set-up and mode of operation used when developing a scale-down model.

1.5.3.7 Chromatography

Chromatography is a complex process with a large number of inter-linked variables that make purely theoretical predictions of column performance extremely difficult (Rathore,

2003). The traditional approach to chromatography scale-down is to maintain the same bed height, linear flow rate, pressure drop and ratio of feed volume to matrix volume (i.e. protein loaded per unit of matrix). The diameter of the column and the flow rate are reduced on scale down to enable the other properties to stay the same. It is also essential when scaling down chromatography operations that the actual chromatography system used remains the same. This means that the matrix used (particle size, degree of cross linkage, functional group, etc.), buffers used (pH, conductivity, etc), feed material (pH, conductivity, protein concentration, etc.) and elution gradients must all be identical to the large scale. Adhering to the criteria above should result in the same peak shape and resolution at both scales. However in practice large-scale column performance is often inferior to that of laboratory scale-down models due to poor feed distribution and non-homogenous packing, which can both cause a decrease in resolution (Rathore, 2003). Other scale-up problems are related to column hydrodynamics and column packing. As column diameter increases wall support for the matrix decreases which results in greater compression, higher pressure drops and bed instability, which may result in maldistribution of flow across the column. Stickel et al (2001) addressed this problem by developing an empirical model to correlate pressure drop with the aspect ratio of the packed bed and the superficial velocity. This and other issues related to chromatography scale-down are discussed in more detail in chapter 4.

1.5.4 Scaling-down whole bioprocesses

The production and purification of biological substances usually consists of a sequence of unit operations. These operations will run in a specific order and the product stream from one operation becomes the feed stream for the next operation, thus the performance of one step will have a direct impact upon subsequent steps. Although interactions between unit operations are often significant they are rarely studied during process

synthesis or optimisation. Current approaches to scale-down usually see one operation studied and optimised at a time, this type of approach means that overall process performance is likely to be sub-optimal (Groep, 2000). If all the steps in a bioprocess are scaled-down simultaneously and run in sequence then the interactions between steps can be studied. This would allow the performance of the whole process to be optimised and would also enable easier identification of critical process parameters and thorough robustness testing.

In order for whole bioprocesses to be scaled down current scale-down models will need to be improved. The interactions between unit operations mean that if the performance at one stage changes then the feed stream to the next stage changes and hence the performance of all subsequent operations and the overall process will be affected. Therefore whole bioprocess modelling will require very accurate scale-down models that are capable of producing material representative of the large-scale. As well as improving accuracy, the scope of scale-down models will need to be increased to incorporate all aspects of a particular stage. Most scale-down models only deal with the core elements of a unit operation and are not suitable for studying secondary operations such as filter cake washing or drying. Industrial processes may also have intermediate holding or transfer stages between key operations, which may involve a number of pumps and valves and a network of pipes. Since these factors can have an effect upon the product, for example shear damage in pumps or protease activity in a holding vessel, they need to be accounted for in scale-down models.

1.6 Conclusions

The length, complexity and costs of pharmaceutical research and development have increased steadily over the last couple of decades. The vast majority of drug candidates

never reach the market and of those that do only a third generate enough revenue to compensate for the huge R&D costs. Much effort has gone into speeding up drug discovery and considerable innovation has lead to the development of combinatorial libraries, high throughput screening techniques and the new disciplines of proteomics and genomics. However in order to avoid a bottleneck at the process development stage similar levels of effort and innovation are required. The concept of process scale-down provides an opportunity for this since more information can be generated using less material and in less time. The critical challenge is to determine how scale-down can be used to produce accurate, quantitative and reliable process data.

1.7 Synopsis of thesis

Chapter 2 of this thesis will present a case study in process optimisation. Small-scale experiments and pilot-plant trials are used to predict the performance of a large-scale operation under different operating conditions. Process data from the large-scale operation is then used to assess the accuracy of the predictions and explain any deviations. The chapter aims to highlight the usefulness of scale-down models and also establish some of the basic factors that can make scale-down problematic.

In chapter 3 the operation of depth filtration will be studied. Fundamental differences between the operation of industrial filters and laboratory filters have limited the accuracy of traditional scale –down techniques for filtration. In this chapter a new approach to scale-down will be presented which attempts to overcome these differences so that the model is more representative of large-scale filters. The resulting scale-down mimic enables accurate predictions to be made about the performance of large-scale filters and enables realistic process material to be produced in the laboratory.

In chapter 4 scale-down of ion exchange chromatography will be investigated. Since this is the most widely used operation in the purification of biological pharmaceuticals there are huge potential savings to be made by developing accurate scale-down models. Traditional scale-down techniques work but only for columns above a certain size since the operation of very small columns is problematic. If chromatography is to be scaled-down to the same level as other operations (tens of millilitres) then these problems need to be addressed. Chapter 4 looks at how very small chromatography columns can be used to generate useful process data and predict the performance of larger columns.

CHAPTER 2
A CASE STUDY IN PROCESS OPTIMISATION

2.1 Introduction

This chapter deals with a case study in process optimisation where pilot scale trials, small-scale experiments and process modelling were used to assess the impact of changing process conditions on the yield and purity of Human Immunoglobulin G (IgG). Full-scale production batches were then used to verify the results of the trial batches and model predictions. The outcomes of this study will also be used to highlight the potential uses for ultra scale-down techniques and establish what are the important considerations when developing a scaled-down process.

2.1.1 The market for human IgG

The demand for purified Human Normal IgG (HNIG) from blood plasma has risen significantly over the past two decades and demands worldwide currently exceed supply (Lebing et al, 2000). This growth in the market has come about as a direct result of increased usage in healthcare procedures. In addition to the broad range of immunodeficiency disorders that IgG has traditionally been used to treat it is now used widely in transplant procedures, in the treatment of AIDS patients and in overcoming resistant strains of bacteria. The current trends in polyclonal IgG usage are likely to continue, giving the IgG market a strong outlook for years to come (Lebing et al, 2000).

In addition to the shortfall in availability of IgG, the availability of processable human plasma is also extremely limited. This means that in spite of the market potential it is virtually impossible for new plasma fractionators to set up, or for existing fractionators to increase their overall production capacity due to a lack of raw materials. However the importance of IgG as a healthcare product and its increasing use create a significant driving force to increase IgG output. To take advantage of this opportunity IgG yields

from existing processes need to be improved. There may be several possibilities for improving process yields since in a multi-step process losses occur at each stage.

2.1.2 Process optimisation

The ideal purification process for a biological product would capture 100 % of the product and remove all contaminants, producing a completely pure substance. In reality this is virtually impossible since the inherent complexity and variability of biological process streams makes separating one species from all the others extremely difficult. Since a perfect process is unachievable a company will try to find the best possible process and the optimum conditions to operate at. The chosen process will meet or exceed all the product requirements such as purity and potency, while maximising yield and minimising cost. However the various process outputs are closely interlinked and an improved performance for one criterion will usually mean a drop in performance for one or more of the others. For this reason improving process performance commonly involves a delicate balancing act between the product yield, product purity and production costs.

Once a process is given regulatory approval the conditions of that process can only be altered within the ranges specified in the product license; otherwise re-approval is required. However despite this obvious barrier to process optimisation after approval, it is a common occurrence since often the speed to market outweighs the need for a fully optimised process. This is not to say that process optimisation does not occur before licensing; it does but the degree of optimisation is often limited by time or resources and then the process is frozen to allow the production of clinical trial material. Once the process is licensed it may be optimised further within the scope of its license to improve the overall profitability.

2.2 Large-scale production of human IgG

The industrial scale production of human immunoglobulins has been taking place since the late 1940's. The product has an excellent track record in terms of safety and efficacy, which goes some way to explaining why the technology used in production has only begun to change significantly over the last two decades (More and Harvey, 1991). Recently several modern techniques such as chromatography have been successfully developed for producing high purity IgG from plasma (Lebing et al, 2000; Karnik et al, 2001). A number of other methods have combined chromatography with the traditional method of cold ethanol precipitation. This approach combines the safety, and robustness of the traditional process with the high product purities achievable using chromatography.

2.2.1 Cold ethanol precipitation

In the 1940's a method for precipitating proteins from human plasma by varying the pH and adding ethanol was developed, this method became known as the Cohn fractionation process (Cohn et al, 1946). The method is used to purify a number of proteins from human plasma including albumin, immunoglobulins and the various prothrombin complexes. The method of cold ethanol precipitation is robust, well characterised, cost effective and has an excellent safety record hence it has been used more widely than any other technique (Foster, 1994). The original method has undergone many changes since its initial development but the main principles of the process remain the same and are still being used by plasma fractionators today.

The exact way in which cold ethanol precipitation works is rather more complex than was originally thought. Alcohol water mixtures tend to have a lower dielectric constant than water alone, which increases the force of molecular interactions. However it was

demonstrated that the low temperatures of the Cohn process cause the dielectric constant of the medium to remain largely unchanged by alcohol addition and the real driving force to be the dehydration of proteins by alcohol (van Oss, 1989). The dehydrated proteins then become strongly attracted to each other and begin to form agglomerates. Another parameter that has an influence on the solubility of the proteins is pH. At pH values above or below the isoelectric point the net charge of the protein moves away from zero making the protein more soluble.

The use of cold ethanol precipitation to fractionate blood plasma is based upon the varying solubilities of the different plasma proteins. By manipulating the five key variables; ethanol concentration, pH, temperature, ionic strength and protein concentration, selected proteins can be made to precipitate thus enabling their separation from the others by either filtration or centrifugation. The purity and yield of each precipitated protein is a function of all the variables above. The complex interactions involved and the interdependence of the variables makes mathematical modelling of the process problematic. For this reason the operating conditions of large scale fractionation processes tend to be carefully selected based on a combination of experimentation, knowledge and experience (Stryker, 1985). Scale-down experiments can be very useful for this purpose provided that they can deal with the complex nature of the process material.

2.2.2 The fractionation process at BPL

The production process at BPL is based upon cold ethanol fractionation. The exact conditions used are a combination of those from Cohn (Method 6) and those developed by Kitsler and Nitchman (1962). Albumin and IgG are largely separated by the first precipitation step (referred to as the Fraction A+1 stage) and then undergo a series of

further precipitations to remove impurities before final purification is achieved using ion exchange chromatography. The type of chromatography used at BPL is a contaminant capture, batch adsorption step on DEAE-Sephadex media. This step was added to the original process to improve robustness and add an extra degree of assurance of contaminant removal. After chromatography the process stream is practically pure IgG (>99%), this is processed further to remove adventitious agents such as viruses and is then adjusted to the desired concentration, formulated and sterile filled ready for distribution.

Figure 2.1 shows the major steps and conditions up to Fraction II (F. II) precipitation in the IgG production process at BPL. The same stages and conditions in the Laboral production process are given in figure 2.2. The changes between the two processes mainly occur at the precipitation steps and are highlighted in figure 2.2. It is at these steps that the greatest product losses occur, mainly due to denaturation and co-precipitation. Understanding how all the process variables interact and bring about precipitation is complex but slight changes can alter the ratio of precipitating proteins and thus affect process yields and product purity.

After F. II precipitation the two processes follow the same path. F. II paste is removed from the centrifuge bowls and dissolved in water ready for application to the DEAE column. The pH and conductivity of the solution entering the column need to be adjusted to ensure adequate separation is achieved on the column. The Laboral process stream must have similar composition, conductivity and pH to the BPL process stream at this stage in order to minimise pre-column treatment. Major additions to the feed stream, particularly high dilutions will have severe effects on later downstream operations and could make processing to final product unfeasible.

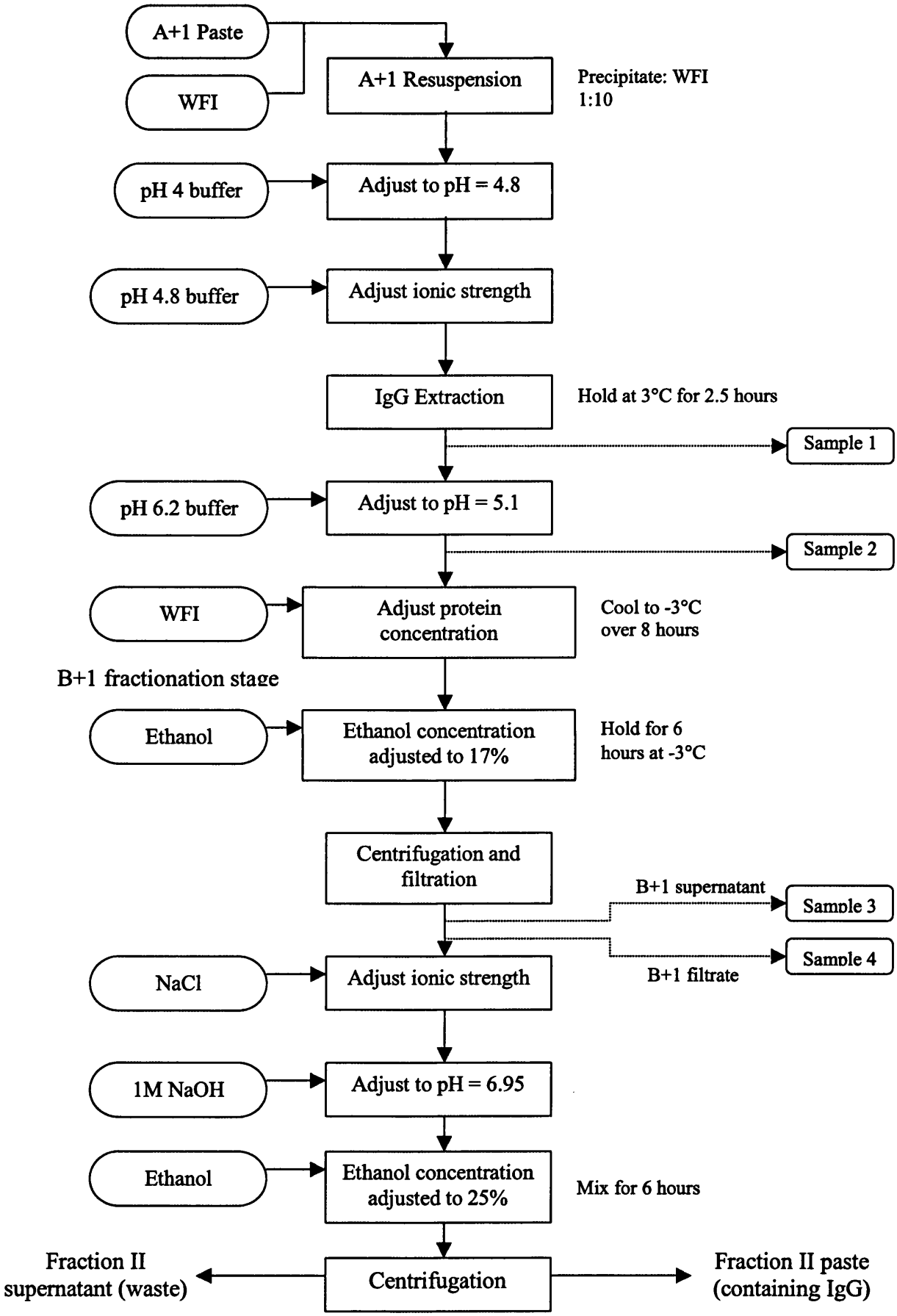

Figure 2.1: Process flow sheet for BPL's full scale IgG purification process.

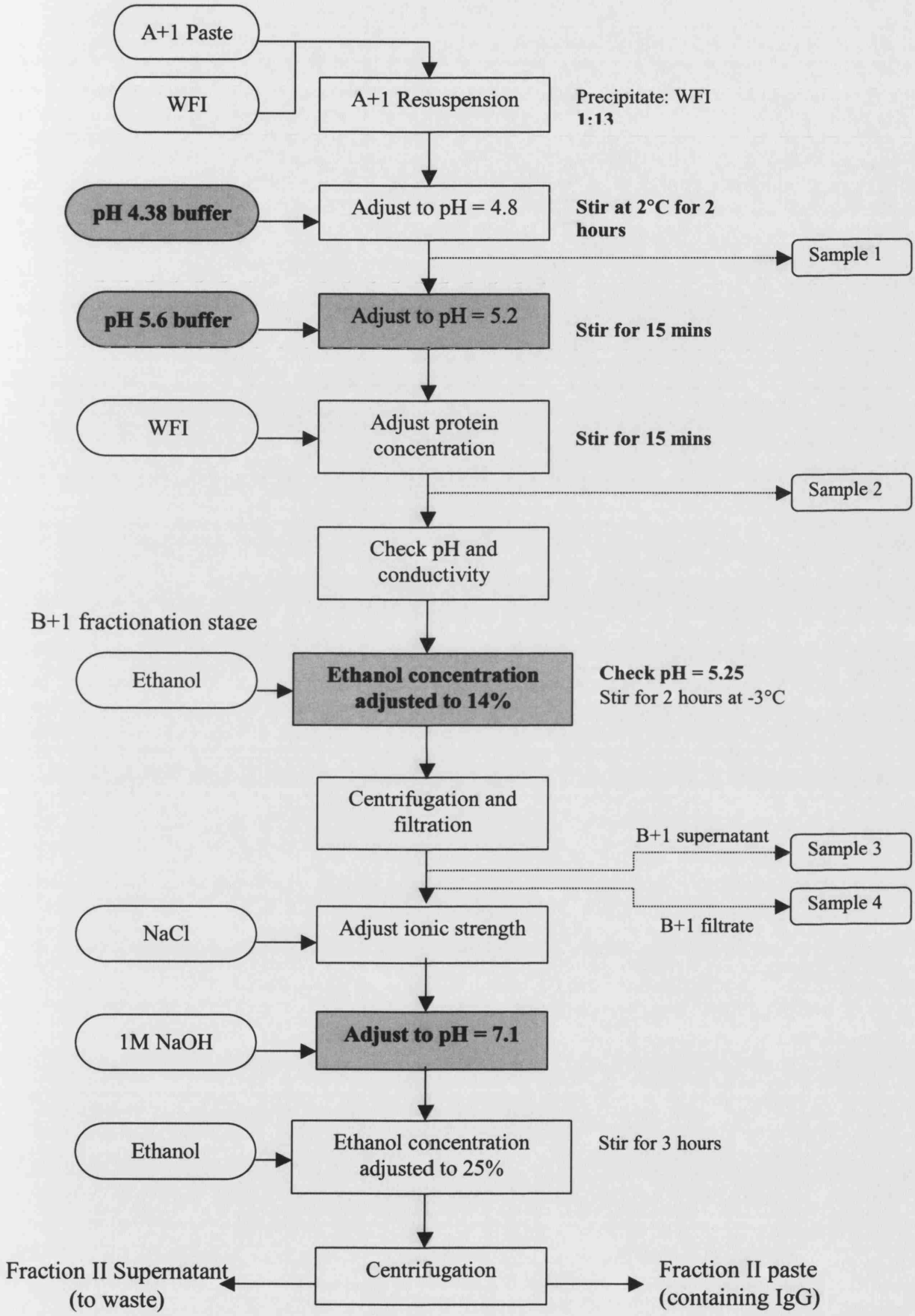

Figure 2.2: Process flow sheet for Laboral's IgG production process to improve IgG yields. Highlighted sections show the major changes from the BPL process.

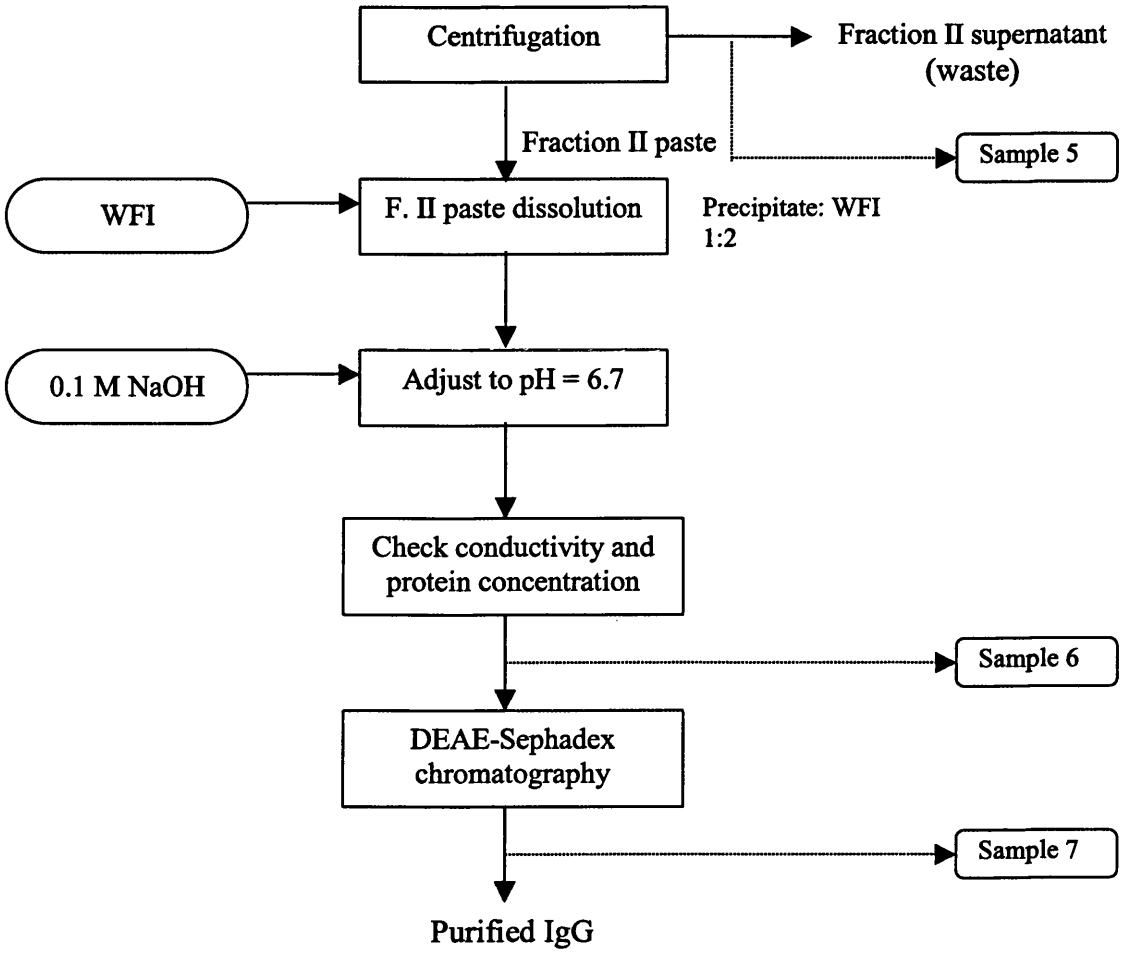

Figure 2.3: Process flow sheet after Fraction II precipitation, showing the position of samples 5, 6 and 7. This stage of the process is identical for Laboral and BPL processes.

A number of sample points are indicated on the above figures. The sample points are at key stages throughout the process so that a measure of the process performance can be made. All the samples taken were assayed and used to calculate yields and losses across the various steps so that the BPL and Laboral processes could be compared. For more information about how the samples were assayed see section 2.4.2.5 and for more information about what the data from the samples was used for see section 2.5.2

2.2.3 Optimising the process

The aim of this study was to increase the yield of IgG in BPL's production facility, which is currently approximately 3.74 grams of IgG per kilogram of plasma paste

processed. The greatest losses in IgG occur at the F. A+1 extraction and F. B+1 precipitation stages so efforts were focussed around these steps. While the yield of IgG was to be increased it was important that the amount and identity of contaminating proteins was maintained or reduced so that there was minimal impact on the rest of the downstream purification sequence. The safety and efficacy of the product are of primary concern and therefore equivalency between the "new" product and previously demonstrated safe, old product must be shown.

The yield could have been improved by changing to a chromatographic processing method, but this would be very expensive and would have moved away from the core skills and experience at BPL. Therefore process optimisation was used in order to increases IgG yields without having to purchase any major new equipment or significantly change the process. This was achieved by altering critical process parameters such as pH, temperature, mixing times and ethanol concentration. When doing this care was taken to remain within the process constraints outlined in BPL's product licence.

A team of external consultants known collectively as Laboral were employed to advise on what parameters to change and how to optimise the process. The suggested changes were initially reviewed by BPL and then tested by performing a number of pilot and full-scale trial batches. Each pilot trial consisted of one batch processed under normal BPL conditions and another processed under the conditions suggested by Laboral. The two batches were sampled at key points and the samples were assayed for product and contaminants so that the Laboral process could be assessed in comparison to the normal process. Figures 2.1, 2.2 and 2.3 show process flow diagrams for the BPL and Laboral processes and the sample points along each.

2.3 Tools to aid process optimisation

The yield and purity obtained in a particular process can often be improved via the use of process optimisation. Achieving this usually requires some of the key process parameters or methods used to be altered. However, the question of what to change and by how much is critical. Methods are needed to test the impact of different process changes with minimal use of time and resources.

2.3.1 Computer modelling

Performing optimisation studies at full or pilot plant scale is expensive as the volume of materials used will be high and there is a risk of wasting valuable product. In addition investigating all the process parameters, particularly in a multi-variable system is often prohibitively expensive and time consuming. A number of tools exist to speed up and reduce the risks during process optimisation. Computer modelling is one such tool, which has great potential to perform process optimisation studies very quickly and with minimum experimentation. Computer models can reduce costs significantly, minimise product or raw materials waste and can predict the effects of changing process parameters. However, constructing accurate and reliable models is often difficult and requires simplifying assumptions to be made, hence models are often used as a guide in process optimisation but always require experimental verification.

Process models can be constructed on a purely theoretical basis, based upon experimental data (empirical models) or a combination of the two. Theoretical models minimise the use of materials since no experimentation is necessary but tend to be very complex due to the large number of factors influencing a real process. Empirical models are often much simpler but are only valid over the range of conditions under which

experimentation was carried out. Usually a combination is used where experiments are employed to find out certain parameters in the underlying theoretical equations.

2.3.2 Process scale-down

Another tool, which can be used to improve process optimisation, is scale-down. The fundamental advantage of scale-down studies is that they require much less time and materials than pilot plant studies and thus are potentially more efficient in terms of obtaining useful process data. A large number of variables can be studied using minimal resources and the results from these experiments can be used to select the best operating conditions for a process and optimise its performance. Obviously the degree of optimisation possible relies on the quality of the scale-down procedures used. However, in almost all situations scale-down experiments can be used to find specific process parameters that assist with process optimisation. Obtaining these values by experimental means is often easier than by purely theoretical means since the level of understanding required is likely to be lower, in some cases experimentation is the only way to find a particular parameter.

Limitations in the accuracy and reliability of scale-down models have restricted their use as tools for process optimisation and pilot or full-scale verification is always necessary. However scale-down is still useful as it can narrow the range of process variables and highlight particular aspects of a process that could be improved. As the accuracy of scale-down techniques improves it may be possible to eliminate large-scale optimisation studies and simply verify the performance of the process designed at laboratory scale.

2.3.3 Process modelling and scale-down in plasma fractionation

Very little work seems to have been done on modelling plasma fractionation, since companies tend to trust and rely upon their own knowledge and experience. There is also very little published evidence of scale-down work on plasma fractionation. However, the unit operations used in the fractionation process are standard biotechnology purification operations and as such general scale-down models exist for these steps. The principle operation used in plasma fractionation is precipitation and this can be scaled down by maintaining geometric similarity between vessels, mean velocity gradient, Camp number and rate of reagent addition (Bell and Dunnill, 1982). The use of realistic process conditions such as choice of precipitant and pH is also critical.

The differences between the Laboral and BPL processes occur in the precipitation steps. Hence scale-down work will focus on predicting protein solubility at the various conditions. Filtration, centrifugation and chromatography conditions remain unchanged between the two processes so these operations will not be modelled. Obviously, changes in the precipitation steps will have an impact upon the other steps, (Bell et al, 1983; Gosele, 1993) but since these steps have a high level of robustness to accommodate the natural variability of plasma, it is assumed that the changes will have minimum effect. The sampling and assay plan will aim to validate this assumption.

2.3.3.1 Predicting protein solubility

The various methods of protein precipitation and their mechanisms have been well studied in the past (Clarkson, 1995; Rothstein, 1994; Bell et al, 1983). In addition to these studies a number of models have been devised for predicting the solubility of proteins in the presence of various concentrations of precipitants. These models vary from highly mathematical inter-molecular force balances (Debye and Huckle, 1923) to

simple empirical equations (Niktari et al, 1990). However all these models deal with salt precipitation of proteins and very little work has been performed on producing models to predict the solubility of plasma proteins during cold ethanol precipitation. However the dehydration mechanism of ethanol precipitation suggested by van Oss (1989) is very similar to the mechanism of precipitation by salting out, suggesting that the general form of the models may still be applicable.

Cohn adopted a simple, semi-empirical approach to protein solubility. It was suggested that the solubility of a protein at varying salt concentrations could be expressed using a simple linear equation of the form:

$$\text{Log } S = \beta - K_S C_S \qquad (2.1)$$

Where S is the protein solubility, C_S is the salt concentration, β and K_S are constants for a particular system. The Cohn equation has been applied to blood plasma fractionation, however its application is limited since it only describes the linear portion of the solubility curve.

Niktari et al (1990) used an empirical equation to describe the solubility of proteins, which was later successfully used in modelling the solubility and purification of yeast alcohol dehydrogenase (ADH) (Clarkson et al, 1992). The equation was of the form:

$$F = \frac{1}{1 + (C_s / \alpha)^m} \qquad (2.2)$$

Where F is the fraction of protein remaining soluble, C_s is the salt concentration expressed as percentage saturation and α and m are constants. The advantage of the Niktari equation is that it can describe the entire solubility curve and so is of greater practical use in protein fractionation since this usually involves precipitation outside the linear regions.

2.3.4 Developing a scale-down model for use in this study

Scale-down experiments will be used in this study to predict the solubilities of the main plasma proteins under various processing conditions. Laboratory experiments will be performed with IgG and albumin in order to determine the effect of ethanol concentration and pH on protein solubility. The experimental data will then be used to determine an empirical relationship that can predict the amount of protein precipitating under various processing conditions.

Ideally these solubility experiments would be performed with real process material and not pure proteins since interactions between proteins may affect solubility (mainly by co-precipitation). However, for this study pure proteins will be used since this allows much simpler assay techniques to be used (e.g.: OD 280). In addition to this the rate of addition of precipitant and the maturation time of the precipitates in the laboratory vessel should be representative of the large scale. However, scaling these factors down properly would make the duration of each experiment impractically long, hence it will be assumed that the effect of these factors on the model is negligible.

2.4 Materials and methods

2.4.1 Raw materials

All chemicals used were obtained from Sigma-Aldrich chemical Co. (Dorset, UK) and were of analytical grade. Chemicals used in the fractionation process were tested to ensure they met all criteria stated in the raw materials compliance. All water used in the two processes was purified water for injections (WFI).

2.4.2 Large / pilot scale trial batches

2.4.2.1 Starting material preparation

The starting material for each batch was F.A+1 paste, which was obtained from BPL's production process. The F.A+1 paste was weighed out into two equal lots, one to be processed under Laboral conditions and one to be processed under BPL conditions. In each trial the F. A+1 paste used was derived from the same plasma pool to ensure that the start material for Laboral and BPL batches was identical.

The trial batches took place at two different scales, pilot and full. The pilot scale batches required 15.75 kg of F.A+1 paste for the initial resuspension and had a maximum working volume of 800 litres at the Fraction II supernatant stage. The large-scale batches each required 235 kg of F.A+1 paste for the initial resuspension and the maximum working volume was 11500 litres.

2.4.2.2 Process equipment

Buffer preparation was carried out in stainless steel mixing vessels (100-500 L) in accordance with the relevant standard operating procedures (SOP) at BPL.

Precipitation and ageing was carried out in stainless steel Cohn fractionation vessels (750 L). These are essentially large, baffled, mixing vessels cooled by means of a glycol filled jacket. The suspensions were mixed by impellors at the bottom of the vessel and all additions were made through the top of the vessel.

At pilot scale centrifugation was performed in Westfalia BKA 6 multi-chamber bowl centrifuges (Odele, Germany). At full scale centrifugation was performed in Westfalia BKB 45 multi-chamber bowl centrifuges. These centrifuges both have two chambers and

are geometrically similar. The main bowl is cooled via a spray nozzle and the centripetal pump feeding the centrifuge is also cooled. When centrifugation is complete any liquid still inside the centrifuge bowl is removed using a siphon. The centrifuge is then dismantled and the paste is manually removed from the two chambers. Properties and operating conditions of the two centrifuges are given in table 2.2.

	Centrifuge	Rotational speed	Flow rate (m^3/s)	Sigma value (m^2)	$Q/\Sigma \times 10^{-8}$ (m/s)
Full Scale					
Fraction B+1	BKB 45	5500	1.25×10^{-4}	4061	3.08
Fraction II	BKB 45	5500	1.67×10^{-4}	4061	4.11
Pilot Scale					
Fraction B+1	BKA 6	4200	3.33×10^{-5}	1424	2.34
Fraction II	BKA 6	4200	2.00×10^{-5}	1424	1.40

Table 2.2: Centrifuge properties and conditions for the various centrifugation steps within the IgG production process. The conditions were identical in Laboral and BPL batches.

Fraction B+1 filtration was performed in 20 litre lenticular (Cuno) filters. The filter media used was made from glass fibres and the filter was placed in a cold room at 3°C during operation.

2.4.2.3 Operating schedule

Each trial run takes three days to complete up to the recovery of Fraction II (F.II) paste. The schedule for each batch is shown in table 2.1. Equipment scheduling allowed the second batch to begin on day 2 enabling a Laboral and BPL batch to be processed in a working week. The F.II paste can be frozen and re-dissolved at a later date lending much greater flexibility to the DEAE-Sephadex stage than other processing steps.

Day	Time	Operation
Day 1	Morning	Fraction A+1 Resuspension
Day 1	Afternoon	Fraction B+1 Precipitation
Day 1	Overnight	Fraction B+1 Maturation
Day 2	Morning	Centrifugation
Day 2	Afternoon	Filtration
Day 2	Evening	Fraction II Precipitation
Day 2	Overnight	Fraction II Maturation
Day 3	Morning	Centrifugation
Day 3	Afternoon	Fraction II Paste Recovery

Table 2.1: The schedule for each trial batch. A Laboral batch and a BPL batch were run in parallel one day apart to give time for equipment to be cleaned (so the second batch was started on day 2).

2.4.2.4 Sampling plan

Samples of the production stream were taken at key points along the process for analysis. Following centrifugation and filtration the product stream was collected into a vessel and mixed before sampling so that a representative sample could be collected. Liquid samples generally had a volume of 250 mL. The position within the process of each sample is described in Table 2.3 and illustrated in figures 2.1, 2.2 and 2.3. In addition to samples from the liquid stream samples of centrifuge paste were also collected for dry weight analysis. The samples were taken perpendicular to the centrifuge bowl so that they covered the whole depth of the centrifuge paste. One sample was taken from each chamber of each centrifuge bowl.

2.4.2.5 Sample analysis

2.4.2.5.1 General sample preparation

Sample analysis was performed in the R&D department. Each sample was assayed using a Turbitime unit (TT) and RID plates (see sections 2.4.2.5.2 and 2.4.2.5.3); the

absorbance at 280 nm was also measured. Some samples were diluted in 0.9 % saline prior to assaying. The appropriate dilutions and storage temperatures are given along with the description of each sample in table 2.3. All samples were assayed as quickly as possible, which was usually within an hour. Samples taken outside of normal working hours were samples at the first available opportunity and always within 12 hours of being taken.

Sample No	Sample Description	IgG Dilutions TT	IgG Dilutions RID	Albumin Dilutions RID	A280 Dilutions
1*	A+1 extract (pre pH 5.6 buffer)	1/10	1/5	1/5	1/25
2*	A+1 extract (pre ethanol addition)	1/5	1/3	1/2	1/10
3**	B+1 Supernatant	1/3	1/2	neat	1/3
4**	B+1 Filtrate	1/3	1/2	neat	1/3
5**	F II Supernatant	-	LC	neat	neat
6*	Pre Dex	1/100	1/50	1/25	1/100
7*	Post Dex	1/100	1/50	neat	1/100

*Store at 4°C
** Store at 0°C

Table 2.3: Description and dilution of process samples taken during trial batches. Samples 1 and 2 were centrifuged prior to dilution in a Biofuge Stratos (Heraeus Instruments) at 5000 rpm, 4°C for 10 minutes and then passed through a 0.2 μm syringe filter (Whatman, USA). Other samples were filtered only if the liquid appeared to contain particulates.

2.4.2.5.2 Turbitime unit

The amount of IgG present in samples was measured using a Turbitime system (Behring, Marburg, Germany). This system is based upon the principles of kinetic turbidimetry and measures both the maximum reaction velocity and the time taken to reach this velocity when sample is mixed with precalibrated antisera. A 20 μL sample is pipetted into a plastic cuvette and placed into the machine. Next 500 μL of IgG antisera is added to the cuvette, a small magnetic stirrer in the bottom of the cuvette mixes the liquid

automatically, then the absorbance is determined and the concentration of IgG is calculated and printed out. This method was also used with different antisera to assay some impurities; namely IgA, IgM and Albumin.

2.4.2.5.3 RID plates

Three different types of radial immunodiffusion (RID) plates (Behring, Marburg, Germany) were used, one for Albumin and two for IgG, one of which was a specialised low concentration plate (LC) for detecting very low levels of IgG. The plates consist of an agarose gel, which contains specific antibodies. The antibodies form precipitate complexes with the target proteins and become visible as opaque rings in the gel. The diameter of the ring is directly proportional to the concentration of the target protein in the sample.

Each plate consists of twelve circular wells cut into the agarose gel, 0.5 μL of sample is dispensed into each well using a positive displacement pipette and the plate is left to stand for 2-3 days at room temperature (15-25 °C). The first three wells are filled with standard solutions of different concentrations so that a calibration curve can be constructed. After incubation the size of the product ring around each of the wells is measured using a magnifying eyepiece and a ruler. The calibration curve is constructed and the concentration of the analyte in each sample is calculated from the curve.

2.4.2.5.4 UV absorbance

The UV absorbance at 280 nm of each sample was measured using a Perkin-Elmer Lambda 20 spectrophotometer (Perkin-Elmer Instruments).

2.4.2.5.5 Moisture content

The moisture content of centrifuge paste samples was measured using a Halogen Moisture analyser HG53 (Mettler-Toledo Instruments) set to 170°C and a sensitivity of 3. The dry solids content of liquid samples was also measured by the same means.

2.4.3 Constructing a scale-down model

The scale-down methods described here are not novel techniques and were used as a way of exploring some of the problems commonly encountered in scale-down. The model was based upon precipitation experiments performed with purified IgG and albumin to determine the solubility of these two proteins under different conditions. The empirical relationships determined from these experiments were then used within a computer spreadsheet in Excel, which performed a mass balance across the system to determine how much of each protein remained in solution. From the mass balance the yield and purity of IgG were estimated.

2.4.3.1 Precipitation experiments

The main differences between the Laboral and BPL processes are the pH and ethanol concentration at the fractionation steps (F.B+1 and F.II). The pH values at these steps are 5.1 and 6.95 respectively for the BPL process, and 5.2 and 7.1 respectively for the Laboral process. Precipitation experiments were performed at all these pH values (± 0.01) by varying the ethanol concentration from 0 – 40% and determining the solubility of the proteins at various concentrations. A solubility curve for each protein at each pH was then plotted.

Purified IgG and albumin were obtained from BPL's production process for use in these precipitation experiments. The concentration of the proteins was adjusted to

approximately 5 g/L using deionised (DI) water and the conductivity was checked to ensure it was below 5 mS/cm. Precipitation was carried out in a 200 mL jacketed glass reaction vessel (Fisher Scientific, UK) with a four-baffle insert placed inside. 100 mL of protein solution was poured into the vessel and mixed at 350 rpm via an overhead mixer (Ika, Eurostar Digital) fitted with a 4-blade turbine. The laboratory vessel used for precipitation had a similar geometry to the large scale Cohn vessels. The number of baffles and the type of impellor were also the same as at large and pilot scale. The ethanol was fed from the top of the vessel into the impellor region to ensure good mixing and minimise areas of high ethanol concentration, which can denature proteins. The vessel was cooled to -3°C by the flow of glycol through the jacket from a Haake C35 chiller unit (Haake, Denmark).

Ethanol (96%) was pumped into the vessel using a Watson Marlow 101U peristaltic pump (Watson Marlow). The pump was calibrated prior to use and the flow rate was set to 33 mL/hour so that the ethanol concentration was increased from 0 to 40% over 2 hours. Samples were taken from the bulk mixture at various different ethanol concentrations, centrifuged for 10 mins at 5000 rpm, -3°C and then the absorbance at 280 nm was recorded as a measure of how much protein was still in solution.

2.5 Results and discussion

2.5.1 Protein solubility experiments

The small-scale solubility experiments were used to determine a relationship between ethanol concentration and the fraction of soluble protein at various pH values. The experimental data was plotted in several different forms and different equations were fitted to the data in order to determine which equation gave the best fit. The Cohn equation fitted the linear portion of the solubility curve well, but as already discussed it is not applicable to the non-linear regions, which is where this process was operating. The Niktari equation also did not match the experimental data well. The general shape of the curve was sigmoidal like the experimental data, but it was much shallower and deviated significantly around the transitional regions of the curve. Neither of these equations was deemed to have a good enough fit of the experimental data to be of use in predicting the performance of the large-scale operations so different relationships were tried out in order to find one that was a good fit of the scale-down data.

Since the experimental data displayed a sigmoidal type of curve other equations with this shape were fitted to the data. The equation that had the best fit to all the solubility curves was the Boltzman equation, which had a chi-squared value of below 0.0008 in all cases. The general form of the Boltzman equation is:

$$F = \frac{P_1 - P_2}{1 + e^{\frac{(E - P_3)}{P_4}}} + P_2 \qquad (2.3)$$

Where F is the fraction of protein remaining soluble, E is the percentage concentration of ethanol and P_1, P_2, P_3 and P_4 are constants. The experimental data and Boltzman fits for all solubility curves are shown in figures 2.4 and 2.5 along with the standard deviation

for multiple runs. The size of the error bars show that the reproducibility of the solubility experiments was high and that the variation in solubility between experiments was low.

As previously mentioned the solubility of the proteins is not the only critical factor, the size and density of the particles formed will determine how well they can be separated from the liquid phase during centrifugation. In this study the scale-down precipitation vessel has the same mean velocity gradient as the larger vessels but the precipitates are not matured for as long since it is impractical. This is likely to cause the production of weaker particles that may be damaged by shear forces and broken into small particles that are more difficult to separate from the liquid phase. However the effect of ageing and mixing conditions upon particle properties is something that can be investigated separately once the ideal conditions for precipitation have been determined and so will not be discussed here.

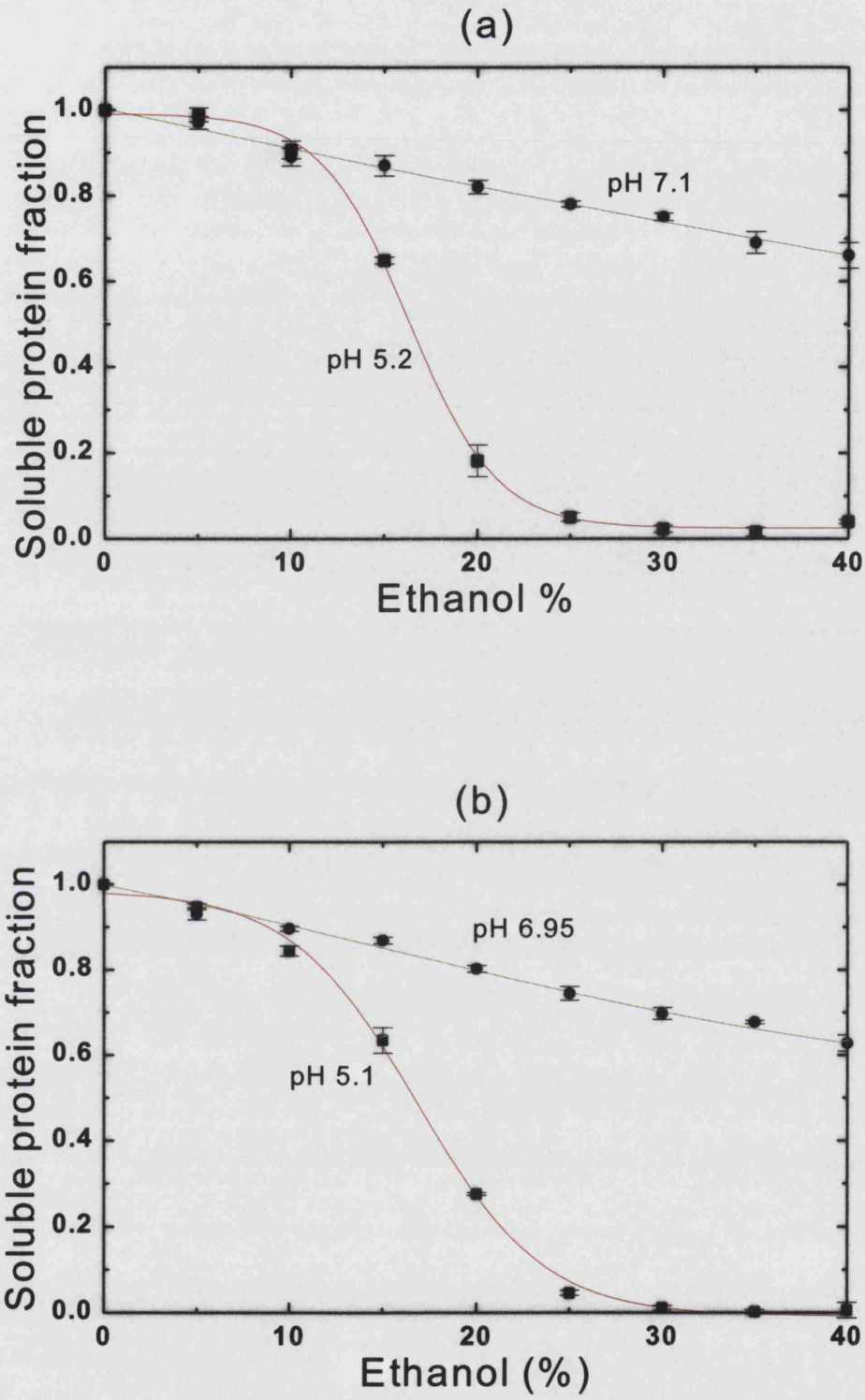

Figure 2.4: Solubility curves for albumin under Laboral (a) and BPL (b) fractionation conditions. Lines of best fit are generated using the Boltzman equation for each set of conditions. Error bars show the standard deviation from three experiments at each pH value.

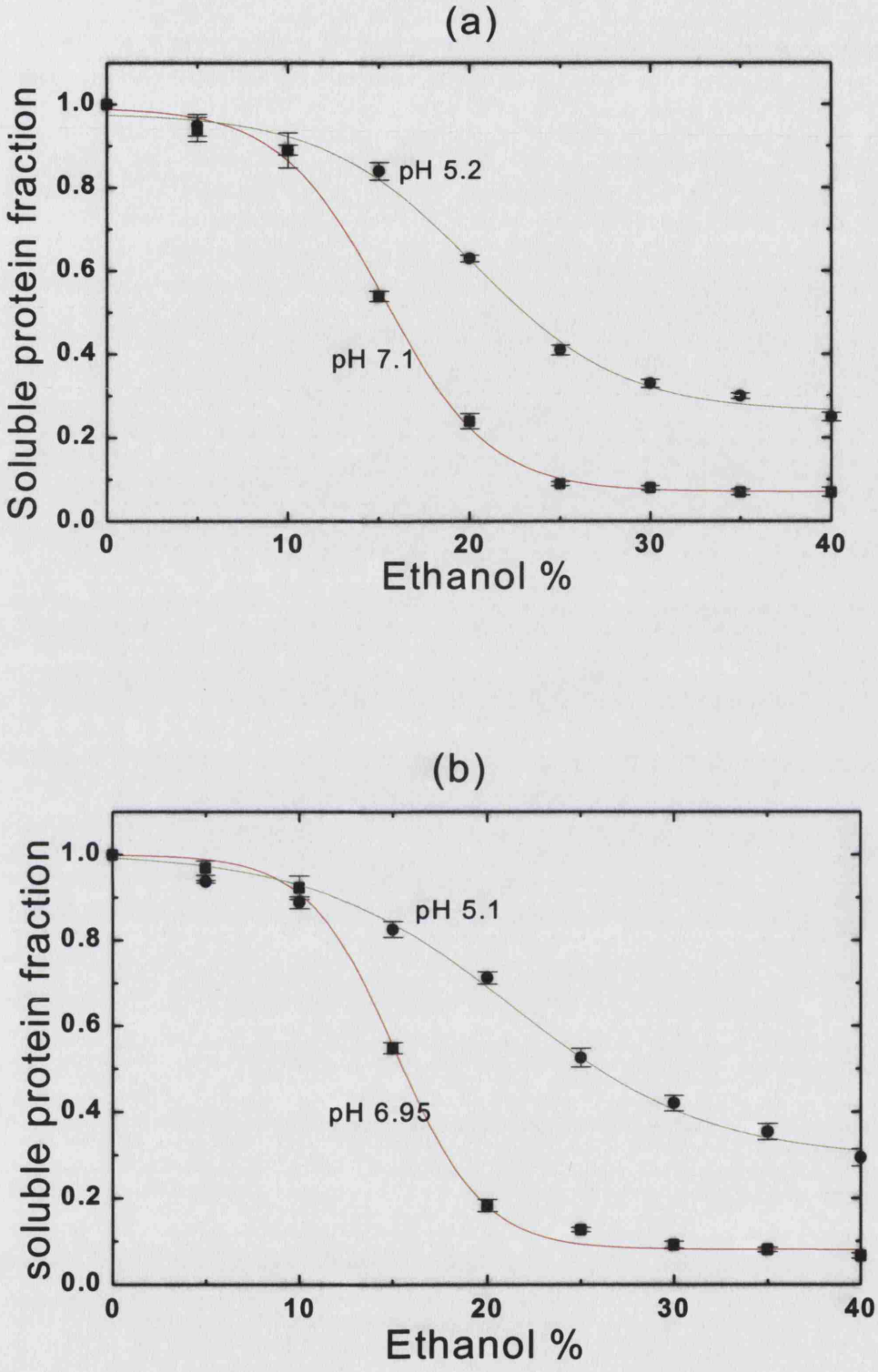

Figure 2.5: Solubility curves for IgG under Laboral (a) and BPL (b) fractionation conditions. Lines of best fit are generated using the Boltzman equation for each set of conditions. Error bars show the standard deviation from three experiments at each pH value.

The Boltzman equation was fitted to all the solubility curves and the constants P_1-P_4 were determined for each pH value. The scale-down experiments were designed to mimic the physical and chemical characteristics of the large-scale precipitation operation. However there is no theoretical basis underlying the use of the Boltzman equation thus the model is completely empirical and is only applicable to this system and under the conditions investigated. To make use of the scale-down data the starting properties of the F A+1 suspension such as volume, mass of IgG and mass of albumin must be known. These can be easily determined by analysing the start material. Then the pH of the precipitation step is used to select the appropriate curves for IgG and albumin. Next using the relevant equation and the percentage of ethanol added the fraction of each protein remaining soluble is calculated. It is then assumed that 100% of the precipitate phase is removed during centrifugation and filtration and that no losses occur to the liquid phase before the next stage. This procedure was followed using typical values for both the BPL and Laboral processes and the results are shown in table 2.4.

The results from the scale-down experiments can be used to perform a basic mass balance across each precipitation step (table 2.4) and the IgG yield and the amount of albumin removed can be calculated. For the BPL process the yield of IgG in the Fraction II paste is 71% and this paste contains 11% of the albumin that was in the initial F.A+1 suspension. The Laboral process has an IgG yield of 76% and contains 15% of the starting albumin. This suggests that the Laboral process will have a higher yield of IgG at the F.II stage but that the paste will contain more impurities. The increased levels of albumin should not be a problem since the DEAE-Sephadex step should remove the majority of it. Therefore on this basis the Laboral process would appear to be an improvement over the BPL process. However, this prediction depends upon the accuracy and reliability of the empirical model and the scale-down experiments it is based upon.

					Liquid Phase		Solid Phase	
		pH	Ethanol (%)	Volume (L)	IgG (g)	Albumin (g)	IgG (g)	Albumin (g)
BPL	F A+1	4.80	0 %	250	1700	350	0	0
	F B+1	5.20	17 %	535	1332.5	152.9	367.5	197.1
	F II	7.10	25 %	605	124.6	114.2	1207.9	38.7
Laboral	F A+1	4.80	0 %	226	1700	350	0	0
	F B+1	5.10	14 %	670	1439.8	239.9	260.2	110.1
	F II	6.95	25 %	750	146.8	187.0	1293.4	52.9

Table 2.4: Predictions based on the solubility diagrams in figures 3 and 4 of the IgG yield and albumin removal at the Fraction B+1 and Fraction II precipitation stages in both the BPL and Laboral pilot-scale fractionation processes. The mass of IgG and albumin, and the volume at the F.A+1 stage are based upon typical values for pilot scale batches. The shaded cells represent material that is disposed of as waste (F.B+1 precipitate and F.II supernatant)

2.5.2 Pilot scale batches

Initially three pilot scale batches were run under each set of processing conditions and the results from these trials were compared to determine if and by how much the Laboral conditions improved the IgG yield. However after these batches had been processed there were still some concerns over the contaminant profile in the Laboral product and over the robustness of the process so two more batches were run under Laboral conditions. The following section discusses the results obtained from all the pilot scale trials under both sets of conditions.

2.5.2.1 Product yield

Process samples were taken at various points in order to give an indication of what was happening to the IgG and where losses were occurring throughout the process. Three different assays were used to calculate the amount of IgG present in each sample,

Turbitime, RID plates and the absorbance at 280 nm. The variation between these three assays made determining the true IgG concentration difficult, especially during the earlier stages of the process when the purity of IgG is low. The most sensitive and reliable technique is A280, which has a coefficient of variation of less than 1% between samples. The main limitation of this technique is it cannot be used with a mixture of proteins since each protein has a different extinction coefficient and the concentrations are not known. The RID plates also only worked well with fairly pure IgG but were not as accurate as A280 since considerable human error arises in measuring the diameter of the sample ring; the coefficient of variation was about 5% between assays but more between different plates. In contrast the turbitime unit can be used with a complex protein mixture since the reagent reacts specifically with IgG, however this technique also had a coefficient of variation of 5 %.

The assay results for all pilot-scale trials are shown in appendix A along with the calculated amount of IgG at each stage. Due to the limitations of each of the three assay techniques different results were used to calculate the total IgG at each stage. Samples 3, 4, 6 and 7 used A280 since this is considered the most accurate method and any protein left after the F B+1 stage is predominantly IgG (i.e.: the IgG is pure). Samples 1 and 2 used the turbitime result since with protein mixtures it is considered most reliable and sample 5 used the RID result since the low IgG concentration meant the other techniques were out of range.

		BPL		Laboral		
	A+1 ppt Plasma equivalent (Kg)	Post DEX IgG (Kg)	IgG per kg Plasma (G/kg)	Post DEX IgG (Kg)	IgG per kg Plasma (G/kg)	Laboral yield improvement (%)
1st trial	321.4	0.89	2.77	1.03	3.21	15.9
2nd trial	311.3	0.99	3.17	1.14	3.67	15.8
3rd trial	301.8	1.07	3.55	1.15	3.80	7.0
4th trial	300.8	-	-	1.21	4.02	13.2*
5th trial	300.8	-	-	1.10	3.65	2.8*

* Based on the highest BPL yield achieved (3.55 g/kg)

Table 2.5: The mass of IgG present at the after the DEAE-Sephadex stage and IgG yield per kg of plasma for all pilot-scale trial runs under both sets of conditions. The concentration of IgG at the post DEAE-Sephadex stage is calculated from the absorbance at 280 nm assuming that all protein is IgG and the extinction coefficient of IgG is 1.4. Measuring the amount of protein in the F.A+1 paste and converting this to an equivalent amount of plasma calculates the yield per kg of plasma.

Table 2.5 provides a summary of the IgG yield achieved at the post DEAE-Sephadex stage in each of the pilot plant trials under both Laboral and BPL conditions. The change in yield as a percentage of the BPL yield is also given. Looking at the table it is evident that despite the similar amount of IgG in solution at the F.A+1 stage there is more IgG left in the F.II supernatant and filtrate of the Laboral process. From this it can be concluded that less IgG is lost at the F.B+1 precipitation stage of the Laboral process, which is in good agreement with the predictions made from the scale down experiments. Decreased product loss over the F.B+1 stage is significant since this was the stage at which the greatest losses occurred. The reasons behind the improved yield at this stage are most likely due to the lower ethanol concentration and reduced mixing times employed in the Laboral process. Lower ethanol concentrations avoid over precipitation and discourage co-precipitation of IgG with other F.B+1 constituents. Long maturation

times increase the likelihood of co-precipitation of IgG and so need to be minimised. Chang (1988) reported on a continuous fractionation process that involved very low maturation times that resulted in less co-precipitation and greater product yields. From this it can be assumed that extending the length of the maturation period longer than necessary risks product losses.

The results also show that the final (post chromatography) yield of IgG is higher in the Laboral process indicating that IgG losses across the entire process are reduced. This was the case in all trial runs, however in order to make comparisons between batches the results must be normalised to reflect the variable protein concentration of the F.A+1 paste. This is done by calculating the mass of plasma equivalent to the mass of F.A+1 paste used in each batch and then converting the mass of IgG at the post DEAE-Sephadex stage into a yield of IgG per kg of plasma equivalent (PEq). Expressing the yield in this manner also enables the productivity and maximum plant capacity to be calculated for both processes (see section 2.5.5)

In the first three trial runs the yield of IgG per kilogram of plasma processed was greater in the Laboral batch than in the corresponding BPL batch and the final two Laboral batches had yields higher than the best BPL yield. An average yield increase of 10.9 % was achieved across all five batches. On the basis of this yield improvement it would seem beneficial to switch from the normal processing conditions at BPL to the Laboral conditions on a permanent basis. However, when the yields obtained from these pilot plant trial batches were compared to those from full-scale production batches it became evident that the small-scale BPL trials all had abnormally low yields. Production batches routinely obtain yields of 3.74 g/kg after DEAE-Sephadex chromatography, which is only slightly below the best Laboral yield. Since the time and effort involved in changing

from one set of process conditions to another would be considerable the increase in yield needs to be significant to justify the change.

2.5.2.2 Product purity

The amount of IgG recovered from the process is not the only factor that warrants consideration when assessing the performance of the Laboral batches. The contaminant profile in the product was also investigated to prove that the quality of the product has not been affected. The final IgG product has written acceptable limits for all the major contaminants that must be adhered to if the product is to be sold. The DEAE-Sephadex column acts as a final purification step to produce virtually pure IgG and contaminant levels after this stage must be within the acceptable range. During the earlier stages the contaminant profile is not as important since it is assumed that the IgG will be pure after the chromatography stage. However if the load on the DEAE column is increased then breakthrough of some contaminants could occur into the product. In addition to this a change in the contaminant profile entering the column may have an effect upon column cleaning and matrix lifetime.

A summary of the contaminants found before and after the DEAE-Sephadex stage for the first three trial batches is given in table 2.6. On the whole the difference in contaminants between the two processes is small and all the results met with product purity criteria after the DEAE stage. The second Laboral batch gives cause for concern since it has higher contamination levels for five of the impurities tested after the DEAE column. Another cause for concern is the much higher IgA levels in all three of the Laboral batches prior to chromatography. IgA levels need to be minimised to prevent shock in some patients. The DEAE column successfully removes the IgA but there is a risk that the column will become fouled causing the level of IgA in the product to reach

unacceptably high concentrations. If this were to happen the batch would be ruined at considerable financial cost. The contaminant profile in the other two Laboral batches was very similar to that of the first three. IgA levels were slightly higher than in the BPL batches and in one batch plasmin and plasminogen levels were also a little bit higher. However as with the first three batches all purity criteria were met after chromatography.

		1st trial		2nd trial		3rd trial	
		Pre Column	Post Column	Pre Column	Post Column	Pre Column	Post Column
Albumin (g/L)	BPL	2.6	ND	1.16	0.01	6.6	ND
	Laboral	0.5	ND	1.87	0.01	0.54	ND
IgA (g/L)	BPL	0.17	0.01	0.33	0.01	0.22	0.03
	Laboral	0.62	ND	0.56	0.04	0.44	0.01
Transferin (g/L)	BPL	0.46	<0.02	0.35	ND	0.72	ND
	Laboral	0.09	<0.02	0.15	ND	0.13	ND
Haemoplexin (g/L)	BPL	0.12	ND	0.10	ND	0.11	0.02
	Laboral	0.10	ND	0.19	ND	0.19	0.09
IgM (g/L)	BPL	<0.01	<0.01	0.06	<0.01	0.28	0.09
	Laboral	<0.01	<0.01	ND	ND	0.12	ND
α-2-macro (g/L)	BPL	0.02	ND	0.02	ND	ND	ND
	Laboral	0.02	ND	0.08	0.01	0.63	ND
PKA (U/mL)	BPL	28.8	1.5	47.3	8.6	895.0	34.5
	Laboral	54.1	1.5	158.0	14.1	174.0	14.0
Plasmin (iu/mL)	BPL	0.07	0.05	0.06	0.04	0.07	0.04
	Laboral	0.06	0.05	0.08	0.06	0.05	0.02
Plasminogen (iu/mL)	BPL	0.30	0.20	0.45	0.21	0.08	0.04
	Laboral	0.04	0.01	0.51	0.32	0.14	0.06

Table 2.6: Concentrations of the key contaminants (albumin, immunoglobulin A, transferin, haemopexin, immunoglobulin M, alpha-2-macroglobulin, pre-kalikrien activator, plasmin and plasminogen) in Laboral and BPL batches; before and after DEAE-Sephadex chromatography. All measurements were by validated QC assay procedures performed in house at BPL.

2.5.2.3 Laboral vs. BPL

The decision of whether or not to implement the changes suggested by Laboral was a difficult one. The Laboral yields from all the trial runs were superior to those from the BPL pilot scale batches, but were not a significant improvement on the IgG yields

achieved in production. The most likely cause of this is a low protein content in the initial F.A+1 paste used for the trial runs and if this is the case a further yield improvement should be expected if paste with a more typical protein content is used. It is also possible that the Laboral yield would increase on scale-up (from pilot scale to production) just like the BPL process seems to do.

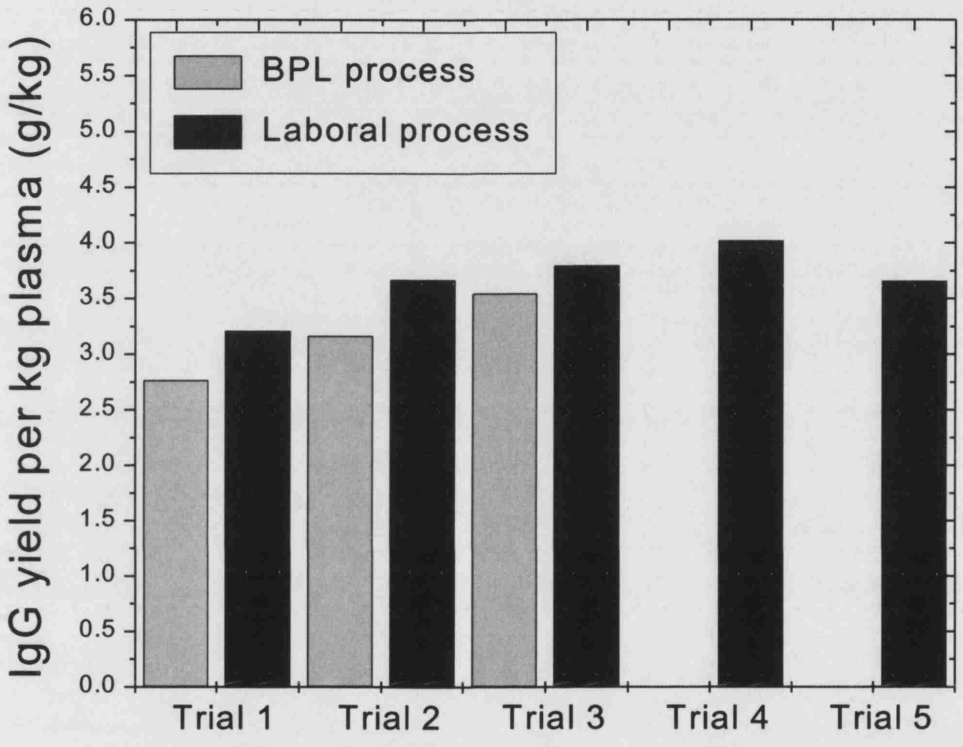

Figure 2.6: A comparison of the BPL and Laboral process yields after the DEAE-Sephadex stage in pilot-plant trials. Yields are expressed as grams of IgG per Kg of plasma processed.

The final (post DEAE-Sephadex) yield of IgG for all the pilot scale batches are shown in figure 2.6. It is easily noticeable that the Laboral batches have higher IgG yields than the corresponding BPL batches. The last four Laboral batches also show a high level of consistency and reproducibility, indicating that the process is robust. The overall average yield for the Laboral process is 3.67 g IgG/kg plasma, which is roughly 0.1 g/kg, better than the best BPL yield at the same scale. On this basis it was decided to proceed with

full-scale Laboral batches despite the fact that BPL yields at full scale were superior to Laboral. The exact reasons for the low yields are unclear but testing did indicate that the F.A+1 paste used had lower than normal protein content. Since the Laboral process performed in a superior manner at pilot scale similar improvements were expected over the current process at full scale.

2.5.2.4 Centrifugation performance

Alterations however small, to precipitation conditions may have an effect on the physical properties of the protein precipitates (e.g. size distribution, density, etc). These physical properties are critical in determining how easily the precipitates can be separated from the liquid phase. Low-density particles are harder to separate than high-density particles and are also more susceptible to shear damage. High shear stresses such as those found in centrifuge feed zones break the precipitates up into smaller particles, which are difficult to separate. To avoid having weak particles with a wide size distribution the mixing conditions such as time and impellor speed need to be carefully selected. The precipitate ageing time is much lower in the Laboral process than in the BPL process but impellor speed is the same, this could affect the performance of the centrifuges.

In order to asses whether the Laboral process changes had an effect upon precipitate recovery dry weight tests were performed on centrifuge pastes and liquid samples. This information can be used to assess if the current centrifugation conditions will be suitable for the Laboral process or if they need to be modified to improve the separation. A summary of the dry weight analysis for Laboral and BPL samples is given in table 2.7 along with the actual mass of paste recovered from the centrifuges at each stage.

	Wet Paste Mass (kg)			Mean dry weight as % of initial weight			
				Fraction A+1	Fraction B+1		Fraction II	
	F A+1	F B+1	F II		Inner Bowl	Outer Bowl	Inner Bowl	Outer Bowl
BPL 1	15.75	12.89	3.65	41.55 ± 1.8	34.53 ± 1.0	31.23 ± 1.8	29.35 ± 3.8	-
LABORAL 1	15.75	11.30	3.40	40.05 ± 1.2	34.47 ± 4.8	30.35 ± 1.8	31.48 ± 3.9	-
BPL 2	15.75	12.52	3.94	37.86 ± 1.4	31.22 ± 1.5	31.41 ± 3.1	32.76 ± 4.9	19.79 ± 2.8
LABORAL 2	15.75	11.29	5.01	39.86 ± 1.3	38.64 ± 2.8	31.14 ± 1.4	29.38 ± 6.0	23.78 ± 2.2
BPL 3	15.75	10.74	4.56	44.62 ± 6.2	36.35 ± 2.4	30.82 ± 3.4	29.42 ± 1.0	17.15 ± 1.7
LABORAL 3	15.75	10.10	4.17	39.95 ± 4.1	36.86 ± 4.1	35.41 ± 5.6	29.50 ± 3.8	21.07 ± 2.3
LABORAL 4	15.75	11.01	4.58	38.69 ± 2.5	35.42 ± 3.9	28.92 ± 7.2	31.48 ± 0.5	25.39 ± 1.2
LABORAL 5	15.75	10.90	4.49	41.36 ± 1.7	35.86 ± 2.7	34.31 ± 1.9	31.10 ± 1.4	22.48 ± 1.9

Table 2.7: The mass and moisture content of centrifuge pastes from BPL and Laboral pilot scale batches. Mean values are quoted for between 3 and 5 readings; error values indicate the range of the readings.

	Mean dry weight as % of initial weight				
	1	2	3	4	5
BPL 1	2.71 ± 0.2	1.63 ± 0.2	0.51 ± 0.4	0.71 ± 0.2	0.66 ± 0.3
LABORAL 1	2.67 ± 0.3	1.24 ± 0.1	0.33 ± 0.2	0.30 ± 0.0	0.51 ± 0.3
BPL 2	2.67 ± 0.3	1.51 ± 0.2	0.40 ± 0.1	0.29 ± 0.1	0.51 ± 0.1
LABORAL 2	3.05 ± 0.3	1.48 ± 0.2	1.02 ± 0.4	0.41 ± 0.1	0.44 ± 0.2
BPL 3	2.49 ± 0.1	1.35 ± 0.1	0.38 ± 0.0	0.37 ± 0.0	0.38 ± 0.5
LABORAL 3	3.17 ± 0.1	1.49 ± 0.3	0.35 ± 0.3	0.43 ± 0.2	0.39 ± 0.1
LABORAL 4	2.70 ± 0.1	1.18 ± 0.1	0.39 ± 0.1	0.37 ± 0.1	0.41 ± 0.1
LABORAL 5	2.60 ± 0.3	1.23 ± 0.1	0.49 ± 0.1	0.39 ± 0.1	0.44 ± 0.1

Table 2.8: Dry weights of liquid samples from Laboral and BPL batches. Mean values are quoted for 3 readings; error values indicate the range of the readings

Analysis of the centrifuge pastes did not show any significant differences between the two processes (Table 2.7). The normal BPL centrifugation conditions (flow rate, rotational speed, etc) seem to be suitable for use with the Laboral process also. There was no breakthrough of solids into the supernatant and the maximum bowl capacity was not exceeded at either the F B+1 or F II stage. The mass of paste collected at the F B+1 stage was generally lower in the Laboral process; this may be due to less IgG co-precipitating with the albumin and other contaminants. At the F II stage the mass of paste collected in each process was approximately the same. However the paste from the Laboral batches appeared whiter than the BPL paste, which had a yellowish tinge. The difference in colour suggests that the Laboral paste has a higher IgG content and lower albumin content since albumin is the most likely cause of this discolouration.

At both the F B+1 and F II stages of both processes the paste in the inner centrifuge chamber had lower moisture content than the outer chamber. This is surprising since the higher centrifugal force in the outer chamber should facilitate greater dewatering of the paste. However this effect is negated by the fact that the precipitate particles in the outer chamber are smaller than those in the inner bowl and have a large amount of water loosely bound. Since smaller particles have a greater surface area to volume ratio than large particles they tend to bind more water per unit volume of solids. This loosely bound water is difficult to remove from the precipitates and results in a high moisture content paste.

2.5.3 Comparison between scale-down and pilot plant

The scale-down experiments and empirical model predicted that the yield of IgG would be approximately 5% higher in the Laboral process. This prediction is calculated for the F.II stage but should be the same at the post DEAE stage since IgG losses on the column have been shown to be negligible. In practice the Laboral process achieved yields of up to 15% better than the BPL process and so exceeded the predicted improvement. Figure 2.7 shows the predicted and experimental yields at the F.II precipitation stage in all the pilot-scale batches. The predictions use the actual assay results to define the start conditions and then use the empirical models to calculate how much albumin and IgG precipitates at each stage. The model also predicted that the F.II paste from the Laboral process would be less pure than the BPL paste (see table 2.4). This last prediction does not appear to be correct since the Laboral paste appeared whiter (cleaner) than the BPL paste and analysis showed that it actually contained less albumin prior to chromatography.

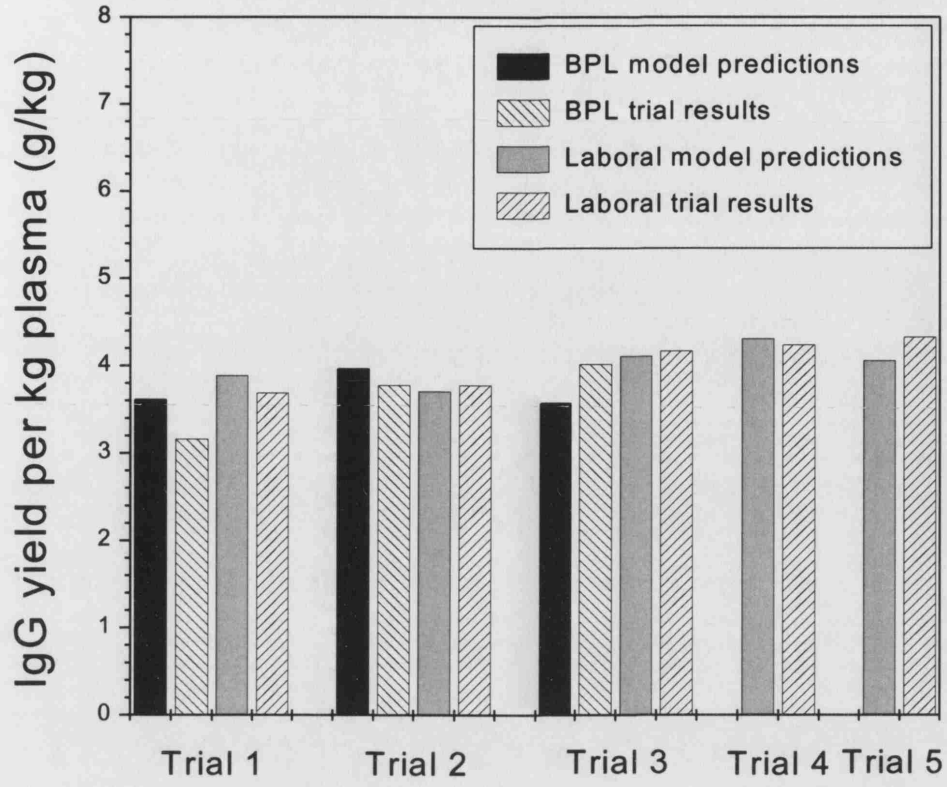

Figure 2.7: A comparison of the predicted and actual yields obtained for Laboral and BPL pilot scale batches at the Fraction II precipitation stage. Predicted yields are calculated using the actual protein concentrations and volumes measured in the F.A+1 suspension at the start of each batch.

The model predictions for the Laboral process support the experimental data and are all within 6% of the actual yields. The predictions for the BPL process however, are not as reliable and vary by as much as 15% from the experimental data. This could be because the scale-down experiments were more representative of the Laboral process; hence are capable of predicting its performance more accurately. One reason for this could be because the mixing and maturation times in the laboratory experiments were closer to those used in the Laboral process. Generally these parameters control the size and density of the precipitates produced but they can exhibit some influence over the amount and type of protein precipitating. The longer maturation periods in the

BPL process could have lead to a greater degree of protein interactions and co-precipitation, which did not occur at the laboratory scale and thus the effects of these were not predicted.

The inaccuracies in the model demonstrate how difficult, but likewise how important it is to scale-down every aspect of the process carefully. Using pure protein solutions for the solubility experiments made analysis easier and was good enough to give reasonably accurate predictions about each process. However if real process material had been used then a better insight into the contaminant profile of each process could have been gained as well as better yield predictions. The fact that the model was more accurate at predicting the performance of the Laboral batches than the BPL ones suggests that there were some aspects of the process that were not scaled down adequately. If the performance had depended solely on the pH, ethanol concentration and conductivity of the solutions then predictions should have been accurate for both sets of trials. However certain factors, that could be considered insignificant were not maintained properly upon scale-down and thus in some ways the scale-down model did not reflect the differences between the two processes. For example, the different process volumes and consequently protein concentrations were not accounted for in the laboratory and this could have affected both product purity and yield. Other apparently minor differences such as the composition and rate of addition of the precipitating agent and buffers could also have influenced process performance.

2.5.4 Full-scale batches

Two full-scale batches were processed under Laboral conditions. The full-scale batches used approximately 18 times the amount of F.A+1 paste used in the pilot

scale trials (283 kg). This amount is 10% lower than the normal start pool used for standard BPL batches. The reason for the smaller start pool was that the increased volumes in the Laboral process meant that the volumetric capacity of the precipitation vessels would be exceeded if a normal size start pool were used. The decreased start pool size has implications for the output capacity of the plant because if less material is processed then less product will be produced and existing customer demands may not be met. This means that the yield improvements must be sufficiently high to offset the smaller start pool since a decrease in the annual production capacity would be unacceptable.

The assay results for both the full-scale Laboral batches are shown in appendix A. The mass of IgG at the post DEAE Sephadex stage works out to a yield of 3.26 g IgG/kg plasma. This yield is unacceptable since it is considerably lower than the average yield for the BPL process at full scale, which is 3.74 g/kg. A second full scale Laboral batch gave a slightly improved yield of 3.44 g/kg at the post DEAE Sephadex stage, but this is still inferior to the BPL yields. The yields achieved in all full-scale batches along with the model predictions for these batches are given in table 2.9.

	Actual Yield (g/kg)	**Model Yield (g/kg)**	**% Difference**
Laboral Trial 6	3.26	3.85	18.10
Laboral Trial 7	3.44	3.63	5.52
Average	3.35	3.74	11.64

Table 2.9: A summary of the yields obtained in the full-scale Laboral trials and the predicted yields for these trials. The yields are expressed in terms of the mass of IgG (in grams) obtained post DEAE-Sephadex chromatography per mass of plasma equivalent used (in kilograms).

The IgG yield was considerably lower than that predicted by the model, which is unexpected since the model predicted the outcome of the pilot-scale batches well. This demonstrates that even the relatively simple task of scaling from pilot plant to full-scale can be complicated by unforeseen factors and careful attention needs to be paid to both the equipment used and the procedures that are followed in the process.

The model predictions for the full-scale Laboral batches were 3.85 and 3.63 g/kg, which work out at an average of 12% higher than actually obtained. The reasons for the reduced yield are difficult to pinpoint since this is a complex multistage process. The model predicted that around 15 % of the total IgG would be lost over the F B+1 precipitation and a further 9 % over the F II precipitation. However, in the trial batches the average loss across each stage was 12 %. This means that the product losses are throughout the process and hence further process optimisation studies could focus on either or both of the precipitation steps.

The model predicted that the Laboral conditions would enable IgG yields to be improved; however despite the good level of agreement with pilot-scale trials this was never achieved. In contrast to the Laboral process the actual yields achieved in full scale BPL batches are better than those predicted by the model. All this points to the fact that parameters such as mixing speeds and times, rates of buffer addition and volume (hence concentration) can be just as important in a precipitation operation as the chemical environment bringing about precipitation. Simple scale-down models can be a useful guideline in process design and optimisation but to truly represent a process at laboratory scale the influence of all aspects of an operation must be mimicked or somehow compensated for.

2.5.5 Plant productivity

In order to meet current customer demands for IgG the plant must produce a certain amount of product each year. The amount of IgG produced per year will depend upon the size and composition of the A+1 start pool, the yield of IgG per batch and the number of batches per year. Since the Laboral process has both a different size of start pool and a different yield to the BPL process the plant capacity will change if the Laboral conditions are adopted.

The amount of IgG produced per annum in the plant can be calculated from the following equation:

$$Productivity = \begin{matrix} mass\ of\ start \\ pool \end{matrix} X \begin{matrix} yield\ per \\ batch \end{matrix} X \begin{matrix} Batches\ per \\ year \end{matrix} \quad (2.4)$$

Under normal operating conditions the plant will process 100 batches a year. Since the average yield for a BPL batch is 3.74 g/kg and the start pool size is 314 kg, the overall plant production capacity for IgG is 117 kg/yr. Under Laboral conditions the start pool size is reduced to 283 kg and the average yield (from full scale batches) was found to be 3.35 g/kg. On the basis of 100 batches a year this gives an annual IgG production capacity of 95 kg/yr, which is approximately 19 % lower than the current BPL capacity.

2.6 Conclusions

The Laboral process did not achieve the yield improvements that were expected and at full scale the yield was approximately 10% lower than that of a standard BPL batch. The purity of the F II paste was a slight improvement over the BPL process however this improvement in purity was accompanied by an increase in conductivity

and pH, which led to complications further downstream. Furthermore there were concerns over the robustness of the Laboral process since the level of some contaminants (such as PKA) in certain batches at the post DEAE Sephadex stage were very close to the acceptable limits.

The lower IgG yields of the Laboral process coupled with the smaller start pool mean that a change from BPL conditions to Laboral conditions on a permanent basis would not be feasible. In addition to this the Laboral batches had a tendency towards high conductivity, which made processing on the DEAE-Sephadex column awkward. Since the Laboral process has a lower annual IgG production than the BPL process it would be difficult to meet current demands let alone increase sales. This means that there would be no economic advantage in switching to the Laboral conditions.

The model predicted the yield of the Laboral process at pilot-scale quite accurately but over-predicted the full-scale yield considerably. It also under-predicted the BPL yield at full scale. This demonstrates the limits of simple process scale-down since it became clear that factors, which had not been adequately considered when scaling down, had a significant effect on the process. BPL's experience with the process over many years had allowed them to adjust the operating conditions to improve yields. In order for a computer or scale-down model to work accurately it must be able to reproduce the effects of these changes. Ultra-scale down aims to reproduce every aspect of an operation in the laboratory so that highly reliable scale-up data can be obtained and the impact of any process changes can be investigated.

A major limitation of this study was that problems that arose in one trial had to be corrected for the next trial based purely upon speculation of what the root cause was. This was because there was not time between batches to do a thorough study into the reasons behind any unusual or undesirable results. Since there were a large number of interacting effects it was difficult to establish exactly what was going on. In addition to this the cost involved in doing extra batches was prohibitively expensive. In this sort of situation good ultra scale-down models would be extremely useful. The short experimental times and low volumes of material required allow a large numbers of experiments to be carried out quickly so that a wide range of variables can be investigated. The risk involved in carrying out such experiments is also much lower than that associated with pilot scale trials since the investment it is not as great.

A number of other process optimisation tools could have been used in this study such as experimental design and robustness testing. These methods could have been employed to investigate what are the critical parameters in the process and which parameters should be altered to give the biggest improvement in yield. However these techniques would still require good scale-down procedures to be followed if the experiments were to have any bearing on the full-scale process.

The discussions above have highlighted some of the potential pitfalls and limitations that are present if scale-down techniques are not carefully designed and operated. The next 2 chapters look at different unit operations and ways in which the reliability of scale-down methods can be improved so that they can be used with greater confidence to predict large-scale performance.

CHAPTER 3
ULTRA-SCALE-DOWN OF CAKE FILTRATION

3.1 Introduction

This chapter will examine cake filtration and how the operation can be scaled down from industrial size filters to the laboratory. Traditionally there have been some fundamental differences between the ways filtration is conducted in a laboratory compared to how it is conducted at large scale. These differences can affect filter performance and may lead to unreliable scale-up data. This work outlines a new approach to laboratory filtration so that a realistic mimic of large-scale filtration is achieved.

In this chapter the effectiveness of filtration as a method of solid-liquid separation will be studied. The degree of separation can be assessed by measuring either the solids still present in the liquid phase (clarification) or the liquid trapped in the solid phase (dewatering). Separation is never absolute and the relative importance of having a dry solid or a particle free liquid will depend on the process (Svarovsky, 1979). These criteria will be used to compare filtration performance to that of laboratory and industrial-scale centrifugation.

3.2 Solid-liquid separation in biotechnology

Solid-liquid separation of some type is a necessity within the majority of processes in the biotechnology industry. There are a large number of operations available for this task (Svarovsky, 1979), most of which are based upon either filtration or sedimentation. However, when separating biological solids from a liquid stream a number of unique problems may arise due to the nature of the materials. Many techniques have been developed to overcome these problems and some of these will be discussed in this chapter.

3.2.1 Centrifugation in biotechnology

Since the density differences between solid and liquid phases in many biological processes are small, sedimentation as a separation method would be extremely slow. Centrifuges speed up the sedimentation of solids by effectively increasing the force of the gravitational field up to several thousand times. This enables biological separations to be performed in a reasonable timescale.

Various types of centrifuge are used in the biotechnology industry including; disc stack, tubular bowl, multichamber bowl, basket and scroll. Although centrifuges have been used successfully in many applications there are some aspects of their operation that are less than ideal. The high shear rates, particularly in the feed zone can damage shear sensitive materials (Boychyn et al, 2001). Also the heat produced by bowl friction means sufficient cooling must be incorporated to avoid denaturing proteins and other materials. The risk of damaging products along with the high maintenance and running costs of centrifuges has led some manufacturers to investigate alternative methods (de Jonge et al, 1993).

3.2.2 Filtration in biotechnology

Filtration is commonly used in the biotechnology industry, either on its own or in combination with centrifugation. A variety of equipment designs are available (see section 3.3.3), but the basic principle remains the same. Essentially the feed stream is passed through a porous media that retains solids but allows the liquid to pass through. Filtration has several advantages over centrifugation according to de Jonge (1993), including; faster rate of separation and thus a high capacity per unit of floor space, improved compliance with current GMP requirements (e.g. clean-in-place), easy maintenance, less noise disturbance and reduced energy costs.

The biggest problem with the use of filtration is the fact that the compressible nature of most biological solids means that cake porosity can be very low, resulting in low filtrate flux rates (Rees, 1990). In order to increase flux rates and improve cake permeability filter aids are used. These are inert, rigid powders with tiny pores to trap solids while encouraging the formation of open flow channels for the liquid. However, there are several problems associated with the use of filter aids such as leaching of hazardous substances (eg: heavy metals), adsorption of products and abrasion in pumps, valves and other equipment. Recently, improvements in filter aid technology have addressed some of these issues making filtration more attractive as an alternative to centrifugation (Sulpizio, 1999).

3.2.3 Solid-liquid separation in plasma fractionation

The various precipitation steps in the Cohn fractionation process necessitate several solid-liquid separation steps. Currently centrifugation is the most common method employed by plasma fractionators, although filtration has recently gained a more active role since it has several advantages (More and Harvey, 1991). Since the Cohn fractionation operates at sub-zero temperatures the cooling requirement of centrifuges is significant due to the heat they generate; this can account for a large portion of running costs. The bowls also have limited solids holding capacity meaning that several bowl changes may be necessary for each batch of plasma processed (even with multiple centrifuges). Maintenance requirements for centrifuges are also high resulting in substantial down-times (Johnston, 1997).

Hao (1985) carried out a pilot plant scale preparation of human serum albumin (HSA) and concluded that depth filtration was superior to centrifugation in terms of time and labour savings, though a small percentage of albumin was not recoverable due to

entrapment in the filter bed. This problem is exacerbated by filter-aids, which adsorb albumin (Johnston, 1997). Friedli et al. (1976) claimed that the adsorptive capacity of diatomaceous earth caused unacceptably low recoveries in purer plasma fractions, but others suggested that in-situ washing sufficiently alleviates this problem (More and Harvey 1991). If the product losses due to entrapment in the filter cake and adsorption to filter aid can be overcome then filtration would seem to be preferable to centrifugation for a number of reasons. These are a greater separation capacity, lower maintenance costs and improved containment and GMP compliance.

Some work has been conducted in order to improve filterability of plasma precipitates by altering the precipitation conditions. Suspensions with a more uniform particle size display improved filtration characteristics. De Jonge et al. (1993) investigated ways of tightening precipitate distributions by avoiding localised supersaturated areas of precipitant, which favour the formation of large numbers of nuclei instead of orthokinetic growth, and avoiding excessively vigorous mixing leading to shear break-up of particles. They were successful in reducing the amount of filter aid required for most fractions by decreasing the rate of ethanol addition, adding the precipitant in the fluid just above the tip of the mixer and generating only a mildly turbulent flow regime. The filtration properties of Fraction IV, which consists primarily of albumin, could not be bettered by the above methods, most probably due to the fine particle size (de Jonge, 1993).

3.3 Filtration basics

Filtration equipment has been developed from practical experience with limited attention paid to theory (Wolthuis and Dichiaria, 1997). This is in part due to the limitations of filtration theory as a result of the complex interactions between the different process parameters; these vary during the filtration process itself on account of the evolving

mechanical properties of the cake. It is vitally important that filtration experiments are performed with actual process material since accurately determining parameters such as cake permeability (or filtration resistance) or porosity depends heavily on the properties of the slurry (Tiller, 1966, Mayer, 2000). Usually a number of laboratory experiments are performed with the slurry in order to gain sufficient information for process-scale predictions.

3.3.1 Scale-up and down

The scaling of filtration is traditionally based on the same applied pressure profile, pore size, and ratio of filtrate volume to surface area. Batch laboratory tests are first performed on a filter of simple geometry to narrow the myriad process options (filter cloth, type of filter aid, precoat, body feed, etc.) using a single pressure. Next, the average flux and the cake resistance are determined as functions of pressure so that the type and size of the large-scale filter can be selected. The laboratory data are then used to calculate the pressure profile required for continuous, constant-rate operation of the industrial filter. This marks the major difference between laboratory-scale and manufacturing-scale filtration; the former is a batch operation usually with compressed air as the driving force, whereas at the manufacturing scale material is constantly fed to the filter via a pump. Discrepancies often arise between laboratory predictions and industrial performance; these are commonly due to differing geometries between laboratory-scale and process scale filters and the interaction of process material with ancillary equipment at the larger scale (eg: shear effects). To surmount these difficulties an ultra scale-down (USD) version of a rotating vertical leaf filter has been developed by modifying a standard Nutsche filter to mimic the manner of operation of large-scale filters. Previously the formation of a filter cake in laboratory filtrations differed from that of larger scale operations since with batch operations the average residence time of the

slurry in the filter is increased allowing particles to settle out under gravitational forces. If the mechanisms of cake formation are different then the structure and properties of the cake are also likely to change. Use of the USD filter should enable accurate performance predictions to be made but should also enable the production of material representative of large-scale operations (cake and filtrate) and give a useful insight into the structure and properties of the cake itself.

3.3.2 Filter performance

As with all other solid/liquid separation operations, filtration performance is assessed by the mass fraction of the solids recovered, often referred to as the separation efficiency or degree of clarification, and the moisture content in the recovered solids. Several steps in the entire filtration process contribute to its performance, these being the application of precoat, actual filtration of the slurry, washing, cake dewatering, removal, and cleaning. Clarity of the filtrate is mostly a function of the precoat type, particle size distribution, porosity and thickness. Washing is concerned with removing the soluble material from the liquid remaining in the cake and is conducted in two ways: on the filter itself or re-slurry washing where the cake is re-suspended with clean wash liquor and re-filtered, the latter method being considerably more expensive (Purchas and Wakeman, 1986). Dewatering refers to the process of purging the liquid, which is trapped in the pores of the filter cake. This can be achieved by either blowing large volumes of air through the cake or by mechanical compression of the cake. Effective cake dewatering and washing are essential to avoid low yields in processes involving multiple separation stages such as plasma fractionation (Pinheiro and Cabral, 1993).

At the production scale a very important parameter is the rate of filtration (i.e. the volume of filtrate collected per unit time) since this has a large effect on the processing

time and costs. An emphasis upon a high initial rate of filtration often proves detrimental to performance if the solids are non-homogeneous. If the starting pressure is high, this will force colloidal particles into interstices of the filter cloth, essentially plugging it and decreasing the rate of filtration. It is advisable to slowly ramp up the pressure so that the filter cake can develop a more open structure. This results in a higher overall filtration rate and easier removal of the cake from the filter medium. However, in filters with large cross-sectional areas it is more difficult to obtain a perfectly uniform cake structure due to regional pressure fluctuations caused by an uneven flow distribution (i.e. turbulent eddies). This often leads to poorer than expected performance of the large-scale filtration device.

3.3.3 Types of filters

Types of filters used in the isolation of protein precipitates are the Nutsche (laboratory-scale only), candle, plate and frame, tubular, leaf filters (see figure 3.1 for a photo of a rotating vertical leaf filter at BPL's manufacturing facility) and microfiltration systems. The initial stages of filtration specification usually consist of small-scale laboratory experiments; these are most commonly performed in a Nutsche filter. A Nutsche filter (Figure 3.2) is a simple batch filter composed of a pressure vessel with a perforated bottom, which supports a filter medium. The slurry is fed into the tank and then the solid-liquid separation takes place as a result of gravitational forces, pressure, vacuum or a combination of these forces (Rushton et al. 1996). Operation is generally at constant pressure such that the flux decreases with time due to the increasing depth and resistance of the filter cake. This is different from larger filters which generally operate in a variable pressure, constant rate mode.

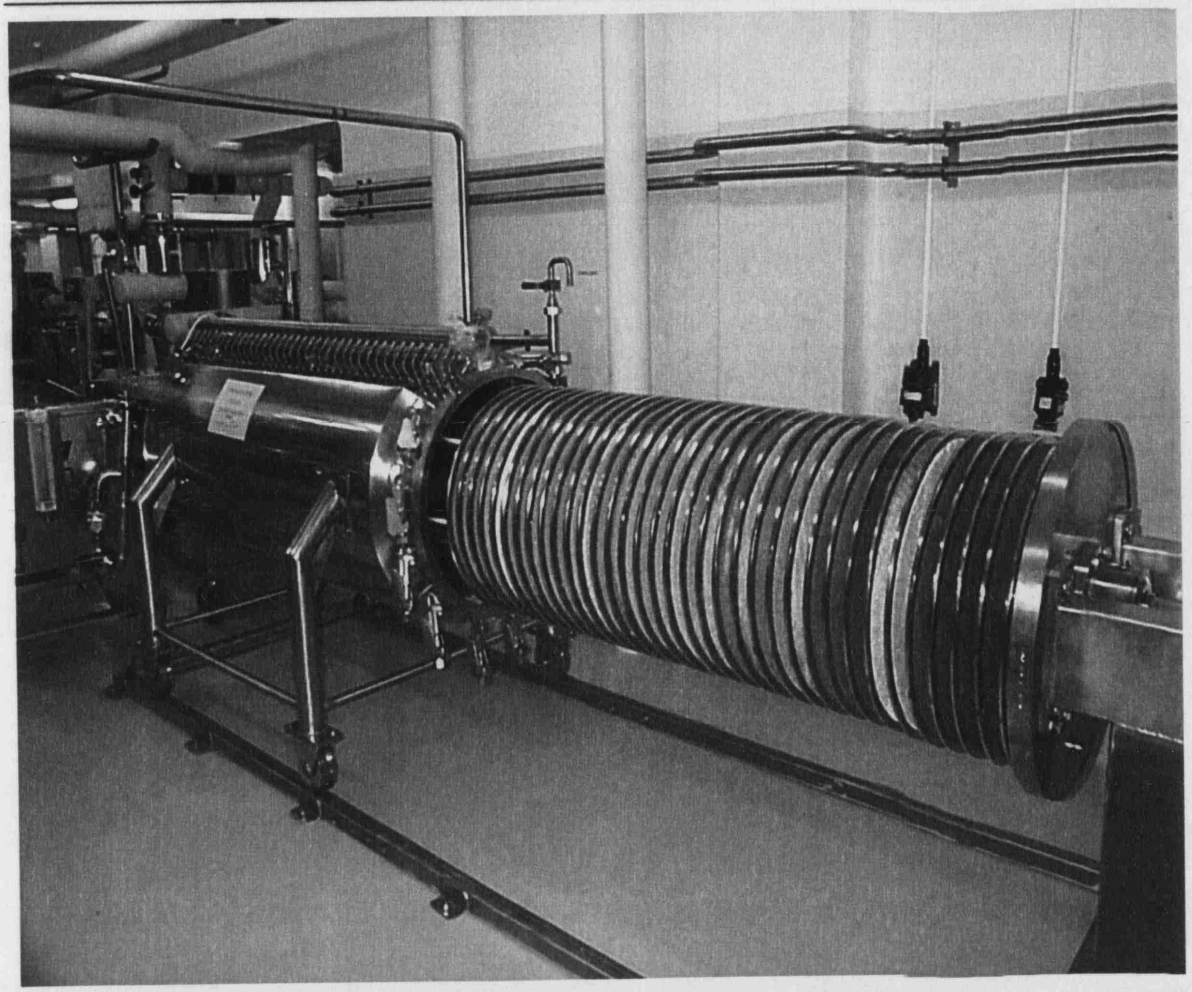

Figure 3.1: This is a photograph of the 24m^2 rotating vertical leaf filter at BPL's manufacturing facility. The vertical plates that can be seen on the rigt slide inside the main housing on the left and rotate during operation. Feed is supplied to the filter from the series of pipes that can be seen on top of the housing and filtrate leaves the filter via a pipe through the center of the filter.

Leaf filters are constructed with horizontal or vertical leaves in a vertical or horizontal vessel. Horizontal leaves are used in intermittent operations such as in a polishing filter where the solids loading is low, since solids can only accumulate on the upper leaf faces. They are also used when cake adherence to the filter cloth is poor. Vertical leaves are designed for ease of cake removal and when solids fraction is higher since particles may accumulate on two surfaces. A photograph of an industrial rotating vertical leaf filter is shown in figure 3.1, the rotation of the filter is essential to ensure an even cake build up on the filter leaves. Pressure-leaf filters are generally considered to be cleaner and more

reliable than other pressure filters. However, careful attention must be paid to filter cake properties (Rushton et al, 1996). In the case of a vertical leaf filter the stability of the cake on the filter element is crucial because a loss of pressure during the filter cycle could result in some of the cake falling away from the element. Furthermore, competition occurs between gravitational and stabilising forces caused by fluid flow and friction, and any change in fluid velocity through the cake may create local areas of instability. To promote fluid flow and thus uniform cake formation, most circular leaf units with vertical elements are rotated slowly (3 rpm). Bridging between leaves must be avoided by considering plate spacing relative to the volume of cake filtered per batch. Bridging can lead to mechanical damage to the end leaves due to excessive pressure.

3.4 Theoretical considerations

Filtration theory is well reviewed elsewhere (Purchas et al, 1986, Rushton et al, 1996). The following section provides a summary of some of the basic equations used in filtration as a point of reference for further development in this chapter.

3.4.1 Permeability

Darcy stated that the pressure loss ΔP, of a fluid passing through a porous bed of solid spheres, is directly proportional to the flow rate dV/dt:

$$\frac{\Delta P}{L} = \frac{\mu}{kA}\frac{dV}{dt} \tag{3.1}$$

Where L is the depth of the bed of solids, μ is the liquid viscosity, A is the filtration area, k is the bed permeability and V is the volume of fluid flowing in time t. Darcy's law assumes that the pressure loss is only due to friction between the solid and liquid (i.e. skin friction or form drag) under streamline flow conditions and that filter medium resistance is negligible (Rushton et al. 1996).

Filter design is basically an empirical exercise due to the inability to calculate accurately cake permeability, particularly for complex mixtures. Kozeny derived the best-known equation for permeability:

$$k = \frac{\varepsilon^3}{K(1-\varepsilon)^2 S_v^2} \tag{3.2}$$

Where ε is the porosity of the bed of solid particles, S_v is the surface area per unit volume of the particles and K is the Kozeny constant. Substituting Equation 3.2 into Darcy's law gives the Kozeny-Carman equation:

$$\frac{\Delta P}{L} = \frac{\mu}{A}\left(\frac{K(1-\varepsilon)^2 S_v^2}{\varepsilon^3}\right)\frac{dV}{dt} \tag{3.3}$$

The Kozeny-Carman expression is used extensively, but it is only valid for rigid particles in a fixed geometry and in point contact with each other. The only forces considered are liquid drag and pressure; these are the predominant forces for particles larger than 10 μm. The equation is a function of porosity, particle size, shape, distribution and packing, rate of cake formation, concentration of the slurry, etc. (Rushton et al. 1996). With such a large number of variables it is often more convenient to measure the permeability experimentally.

3.4.2 Constant-pressure filtration

Filtration can be operated in three different modes: constant-pressure filtration; constant-rate filtration; and variable-pressure, variable-rate filtration. In all three cases Darcy's law provides a basic description of the process. Integration of this equation with respect to time and volume of filtrate enables predictions to be made about the filtration process. For constant-pressure filtration, the linearised parabolic rate law is commonly used:

$$\frac{t}{V} = \frac{\mu c \alpha_{av}}{2A^2 \Delta P_c} V + \frac{\mu R_m}{A \Delta P_m} \quad (3.4)$$

Where t is the time, V is the volume of filtrate collected, A is the filter area, c is the mass of dry solids deposited per unit volume of filtrate, α_{av} is the average specific cake resistance, and R_m is the medium resistance (ie: the resistance of the filter cloth without any cake). The total pressure differential (ΔP) is defined as that over the cake (ΔP_c) and the medium (ΔP_m). Essentially c accounts for the increasing cake depth throughout filtration. A *specific* resistance is used because the cake resistance will rise with cake depth. The mass of dry solids deposited per unit volume of filtrate (c) is determined from the following equation:

$$c = \frac{s\rho}{1 - sm} \quad (3.5)$$

Where s is the mass fraction of solids in the feed, ρ is the liquid density and m is the mass ratio of wet cake to dry. The determination of m can be difficult with biological materials since there is often a high level of associated water. However, assuming m to be unity often results in considerable error (Holdich et al, 1993).

Some modifications are made to Equation 3.4 when using a filter aid for difficult clarifications; R_m becomes the resistance of both the medium and the precoat; α_{av} is the combined specific resistance of the filter aid body feed and solids intrinsic to the suspension; c includes not only the dry solids mass of the original suspension, but also the body feed dosage. In practice, it is the mass of filtrate M, which is recorded so that M/ρ replaces V. The medium resistance (R_m) is usually negligible compared to that of the cake so long as blinding does not occur (Mayer, 2000), so Equation 3.4 becomes:

$$\frac{t}{M} = \frac{\mu c \alpha_{av}}{2A^2 \Delta P \rho^2} M \quad (3.6)$$

Plotting *t/M* against *M* theoretically yields a straight line; the cake resistance can be determined from the slope ($\mu c \alpha_{av}/2A^2\Delta P\rho^2$). In practice however, some deviations from linearity occur. When filtering a *uniform* suspension under constant pressure, a filter cake of approximately constant concentration results; hence, the concept of specific resistance. Nevertheless, the filtration of mixtures containing particles of considerably different properties (e.g. filter aid consists of large, porous particles unlike the much smaller precipitates) means the larger ones will settle faster than the smaller ones. The slurry concentration can be maintained uniform if adequate mixing is provided, but this is not usually the case in laboratory, batch filtration vessels such as the Nutsche. Here the feed slurry is mixed with filter aid, decanted into the chamber, the latter is sealed and pressure applied. This leads to the formation of a cake, which varies in composition from bottom to top, as observed by a changing (usually increasing) resistance during the filtration, i.e. a non-linear plot of *t/M* vs. *M*. Therefore, an *average* cake concentration and specific resistance is defined (Rushton et al. 1996).

Several forms of the parabolic rate equation (3.4) exist. Plotting instantaneous rate of filtration against volume should also produce a straight line, but this requires graphical differentiation of the filtration volume-time curve, which is difficult to perform accurately. Another form of the law involves plotting $(t\text{-}t_1)/(M\text{-}M_1)$ against *M*, where t_1 and M_1 are some arbitrary datum values; this is especially useful when analysing filtration data in which step changes in pressure were introduced (Rushton et al. 1996).

In the case of compressible cakes constant pressure filtration is generally used at laboratory scale to determine the dependence of specific cake resistance on pressure over the cake. The usual empirical relationship between these two variables is:

$$\alpha_{av} = \alpha_o(1-n)(\Delta P)^n \qquad (3.7)$$

Where α_o is a constant determined largely by the size and morphology of the particles in the cake, and n is the cake compressibility coefficient. Constant pressure experiments at laboratory scale are used to define this relationship.

3.4.3 Constant-rate filtration

Large-scale filtration operations are usually performed under constant rate conditions as opposed to constant pressure. Knowledge of the compressibility allows calculation of the pressure increase over time needed to maintain constant flow:

$$\Delta P = \mu c \alpha_0 (1-n) \Delta P_c^{\,n} \left(\frac{Q}{A}\right)^2 t + \mu R_m \left(\frac{Q}{A}\right) \tag{3.8}$$

The expression is simplified in the case of an incompressible cake (i.e. $n = 0$). R_m can be subject to substantial variation but is often negligible compared to the cake resistance so that Equation 3.8 reduces to:

$$\Delta P = \left(\mu c \alpha_0 (1-n) \left(\frac{Q}{A}\right)^2 t \right)^{1/1-n} \tag{3.9}$$

In reality, it is extremely difficult to maintain the flux entirely constant (Cain, 1990); there is always some delay between sensors detecting a decrease in flow rate and delivery pressure by the feed pump (usually a centrifugal or diaphragm pump). This occurs as a consequence of increasing total resistance to filtration due to the increase in cake depth. The result of this is that a filtration process may be operated under conditions of varying pressure and rate. This is a complicated situation with many variables. One way to deal with this situation is to combine the independent variables, pressure and flow-rate, this way the slope of the line becomes independent of pressure.

$$\frac{\Delta P^{1-n}}{Q^2} = \frac{(1-n)\alpha_0 \mu c}{A^2} t \tag{3.10}$$

Calculating the compressibility coefficient (n) and α_0 in this situation can be difficult since the specific cake resistance varies with flow rate and pressure and thus needs to be calculated for each different set of values. Since the instantaneous flux at each pressure cannot be determined easily calculating the specific cake resistance is hard and usually requires some type of iterative process. This concept will be discussed more in the results and discussion section.

A true description of the filtration process involving a compressible cake is extremely difficult and would involve modelling the mechanical properties of the filter cake. This is reliant upon detailed knowledge of the formation and structure of the filter cake, knowledge that is very rarely available. Exact numerical analysis of the formation and build up of filter cake has been shown to give slightly better predictions than conventional theory (Tien et al, 2003), however these methods are mathematically complex. The methodologies described in this report attempt to overcome this problem, using a combination of new scale-down and modelling methods to produce and identify the properties of filter cakes at a laboratory scale representative of large-scale operations.

3.5 Material and methods

3.5.1 Chemicals

All chemicals, unless specified otherwise, were obtained from Sigma-Aldrich Co. Ltd. (Dorset, UK) and were of analytical grade.

3.5.2 Description of equipment

Filtration was performed in a stainless steel Fundabac Nutsche filter (Dr Muller AG, Switzerland) of 35.5 mm internal diameter (i.e. active area of 10^{-3} m^2) with *Celpure*^{TM}

1000 filter aid (World Minerals Inc, Santa Barbara, CA, USA) supported on a GF/B glass microfibre filter cloth (Whatman International Ltd, Maidstone, UK). During batch operation the driving force for filtration was provided by compressed air controlled via a regulator. In continuous operation the feed was supplied to the filter by means of an NF1.30KNDCG Micro-Diaphragm Liquid Pump (KNF Neuberger AG, Balterswil, Switzerland). The filtrate was collected in a beaker (with funnel to prevent evaporation of ethanol) placed on a balance connected to a computer with a software program recording the filtrate mass every 5 seconds. A candle filter (Dr Muller, AG, Switzerland) with a filtration area of 0.012 m^2 was also used for filtration. It was operated in a continuous manner with the suspension fed to the filter from a stirred, jacketed vessel. Filtrate mass was logged every 5 seconds.

3.5.3 Preparation of feed suspension

The feed suspension was Fraction IV (F:IV; 3-5% w/v protein, 40% v/v ethanol, pH 5.85, -5°C) of the Cohn (method 6) fractionation process used in the manufacture of Human Serum Albumin. This suspension was selected due to its small particle size (mean diameter of approximately 1 μm), which presents a substantial separation challenge. One litre of F:IV was obtained from Bio Products Laboratory's (BPL) Production Department from standard manufacturing batches, placed in a cooled (-5°C), baffled, glass vessel and agitated at 2.5 s^{-1}.

3.5.4 Filtration

The Nutsche filter was operated in two modes. First, it was used vertically as a standard, laboratory batch filter in which flow is driven by compressed air (figure 3.2). The

second involved modifying it into an ultra scale-down rotating vertical leaf filter (RVLF), as described below (figure 3.2).

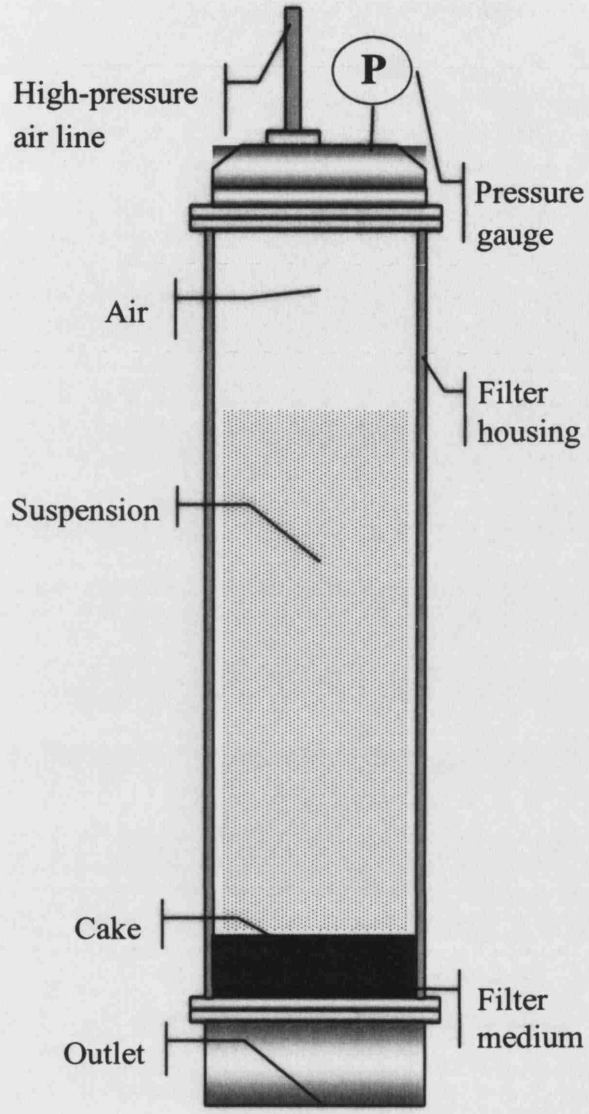

Figure 3.2: A cross sectional view of a standard laboratory Nutsche filter driven by compressed air.

3.5.4.1 Laboratory batch

Celpure^{TM} 1000 (0.5 g) was suspended in 100 mL of RO water, mixed vigorously and poured into the filter housing to give a precoat of 0.5 kg m^{-2} which was formed under 1.0 bar of pressure. The pressure was maintained for 30 s after the entire buffer had been filtered. For body feed, *Celpure*^{TM} 1000 (1.5 g) was then added to 98.5 g of F:IV to give

a concentration of 15 g L^{-1} (1.5 %). The suspension was mixed carefully to minimise foaming and poured into the filter housing on top of the precoat. Constant pressure experiments were performed to determine the cake resistance over the range of 0.5-4.0 bar, as described by Holdich *et al.* (1993). The filtrate was collected in a beaker (no cooling), placed on a balance. In all experiments, the final pressure was maintained until the same filtrate mass was observed three times consecutively, signalling the end of filtration.

3.5.4.2 Ultra scale-down of filtration

The goal was to create a laboratory mimic of a production-scale rotating vertical leaf filter (RVLF), which has a filtration area of 24 m^2. A major difference between the RVLF and a laboratory filter is that the latter is normally operated batchwise, whereas the RVLF is continuously fed suspension via a pump. This difference is overcome using a plunger inserted into the Nutsche filter housing designed to enable continuous flow of feed from a diaphragm pump. Another crucial variable is the gap between the plunger and the filter cloth. The spacing between filter plates in the production scale RVLF is 40 mm, hence the plunger is positioned in the Nutsche to give a filtration chamber with a depth of 20 mm, so as to mimic one plate from the production scale. The plunger is fitted with a double O-ring that prevents leakage from the filtration chamber. This allowed the ultra scale-down filter to be operated at pressures up to 4.0 bar while maintaining the integrity of the seal. Since many industrial filters have a vertical configuration to make best use of space, the Nutsche was turned on its side so that the filtration surface was vertical. A diagram of the ultra scale-down (USD) filter is shown in figure 3.3.

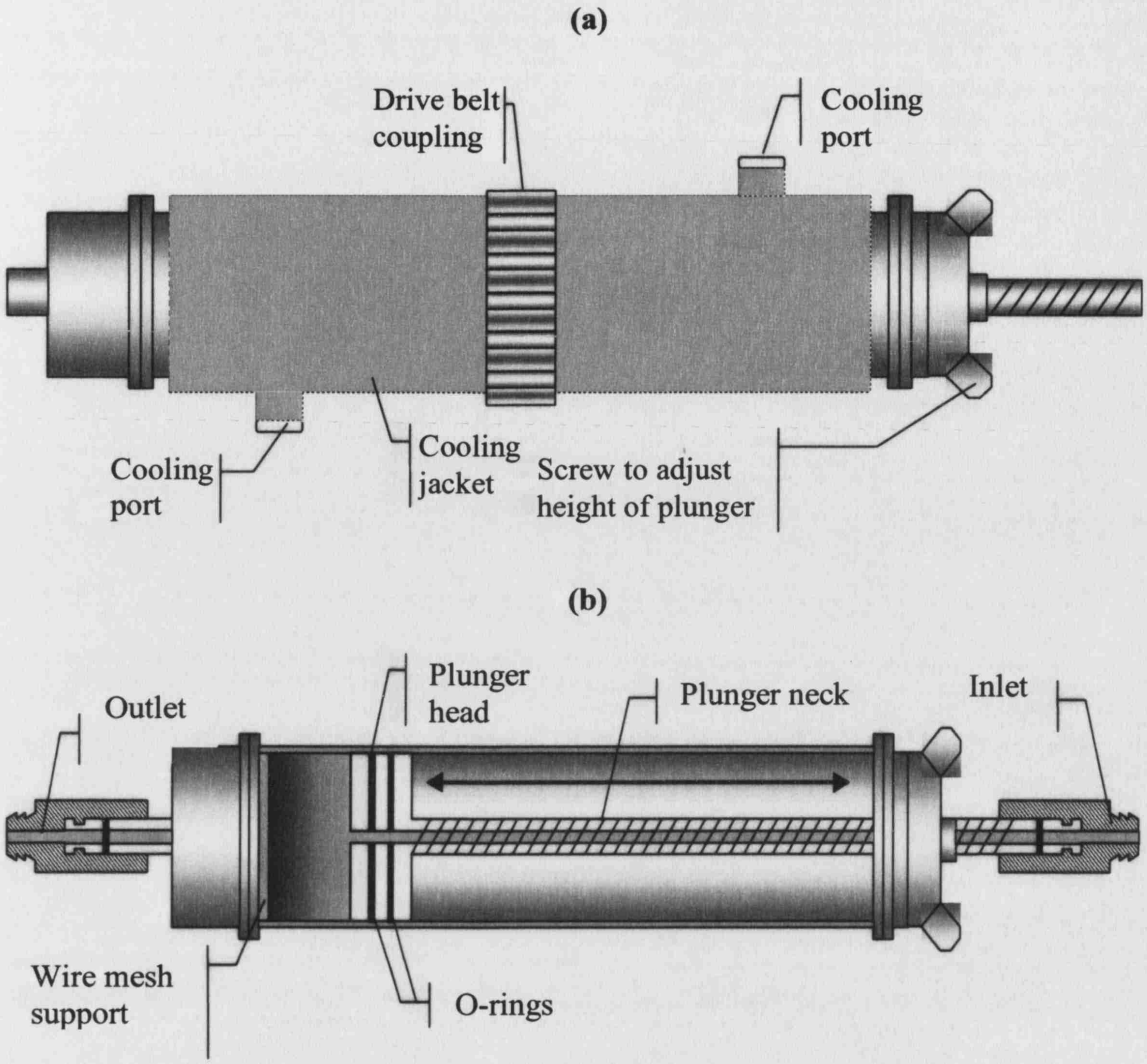

Figure 3.3: Schematic of the ultra scale-down (USD) laboratory filter: (a) view of filter's exterior and (b) its cross section. The ultra scale-down filter consists of a laboratory Nutsche placed horizontally, with a plunger insert so as to adjust chamber volume (i.e. distance between the filter cloth and the chamber inlet).

The USD filter required rotating to mimic the motion of the production scale RVLF and prevent settling of the bodyfeed. This was accomplished by mounting the Nutsche to quick-connect couplings to allow the filter case to rotate without requiring the feed pipe to rotate. The filter was rotated by means of a toothed belt drive connected to a motor, which enabled the filter to rotate at 10 ± 3 rpm. The pressure head exerted by the pump was monitored with a P200 XH Manometer (Digitron, Hertfordshire, UK) and the pumping rate controlled by a DC power supply, (0 to 24 volts) connected to the pump.

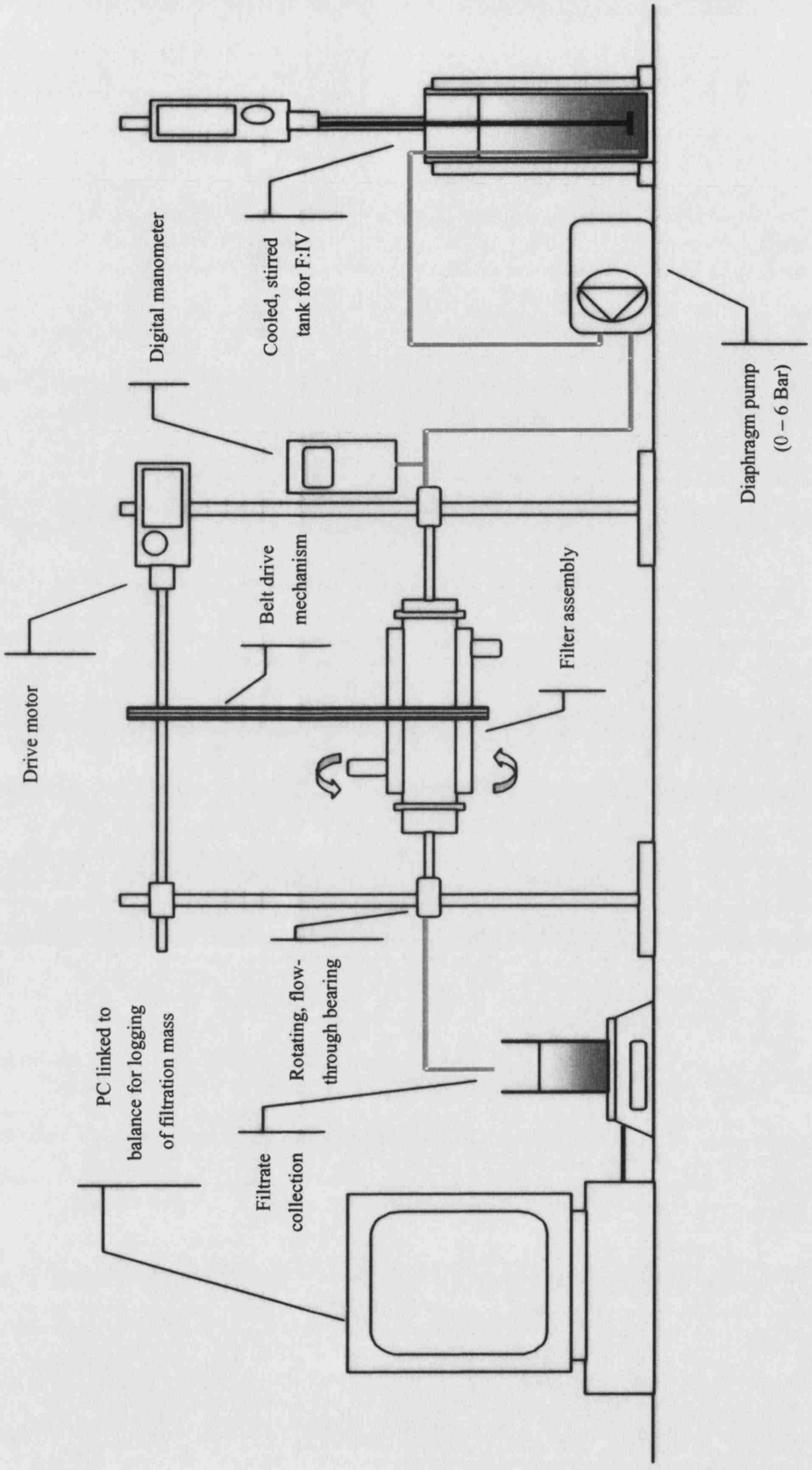

Figure 3.4: Illustration of the ultra scale-down experimental apparatus. The filter assembly is rotated by means of a belt-drive mechanism and motor, and a universal coupling attached to the inlet and outlet of the housing. Suspension is fed to the filter via a diaphragm pump.

When operating the USD filter the pre-coat was mixed with 100 mL of water in a stirred beaker and fed to the filter at a constant pump setting. As soon as all pre-coat had been added the pipe was immediately switched to the mixed vessel containing the F:IV and body feed. Care was taken to avoid air entering the system, which could cause cracking of the pre-coat. The feed rate was adjusted to the desired level and as the feed was seen to enter the filter the mass recording program was started. The pressure was continuously monitored and recorded every 30 s. The complete experimental set-up for the USD filter is shown in figure 3.4.

The USD filter was operated in a manner designed to mimic the operation of an industrial filter. This involved operating at a constant feed rate until a predetermined maximum pressure was reached after which the rate was decreased to maintain this pressure. The maximum pressure selected was 1.5 bar, since this pressure was easily reached before all the feed material had been filtered.

3.5.4.3 Candle filter

Large-scale continuous filtration was performed in a candle filter with a filtration area of 0.012 m^2 (Dr Muller AG, Mannedorf, Switzerland). The support medium for the filter was a PTFE cloth perforated with 10 x 30 μm openings. A pre-coat of 0.5% *Celpure*^{TM} 1000 filter aid was applied to the filter medium and then the feed suspension was immediately fed to the filter via a pneumatically driven Yamada NDP-5FST pump, capable of flow rates up to 8 L/min. The pump was set to a maximum pressure of 1.5 bar and automatically reduced its delivery rate to maintain this pressure. The feed suspension was held in a stirred, jacketed vessel and contained 1.5% *Celpure*^{TM} 1000 filter aid. The filtrate was collected in a large beaker placed on a balance and the filtrate mass was logged every 5 seconds.

3.5.5 Analyses

3.5.5.1 Turbidity

The turbidity of each stream was measured in a Digital Direct-Reading Turbidity Meter (Orbeco-Hellige, UK). A sample was filtered through a 0.2 μm filter and the turbidity of this sample was used as a reference point for 100% clarification.

3.5.5.2 Moisture content

The moisture content of filter cakes was measured with a HG53 Halogen Moisture Analyser (Mettler Toledo, Greifensee, Switzerland) set to 175°C and a sensitivity level of 3 until a constant mass was achieved.

3.5.5.3 Physical properties

The dynamic viscosity of the F:IV suspension was determined to be 6.5 mPas (at -5 °C), using a Contraves Rheomat 225 viscometer with concentric cylinders (Contraves Industrial Products, Ruislip, UK) and the density was measured as 967 kg m^{-3} with a specific gravity bottle.

3.6 Results and discussion

3.6.1 Filter aid dosage

Initially the use of filter aid was investigated. The application of a pre-coat of *Celpure*^{TM} 1000 was found to greatly affect the clarity of the resulting filtrate and to lower the specific cake resistance at a given pressure. Early filtration runs carried out at 1.0 bar with no filter aid (pre-coat or body feed) had an average clarification of 96% and a specific cake resistance of 1.9×10^{13} m kg^{-1}. These filtrations did not go to completion, most of the liquid remained trapped above the filter medium. The addition of 0.5 kg/m^2

Celpure™ 1000 as a pre-coat decreased the resistance to $1.2x10^{13}$ m kg^{-1} and improved the clarification to at least 99%. However, the filtrations still did not go to completion, so the use of filter aid body feed was investigated. The dosage of body feed is critical since too low a dosage leads to unacceptably low flux rates and too high a dosage causes solids bridging which can damage filter leaves. After some preliminary experiments a body feed dosage of 15 g/L was selected (data not shown here). With the addition of body feed and pre-coat the specific cake resistance was lowered to an average of $3.91x10^{11}$ m kg^{-1} at 1.0 bar and the filtration went to completion within an hour.

3.6.2 Traditional batch filtration

Batch filtration of F.IV was performed at various pressures from 0.5-4.0 bar. The cumulative mass of filtrate collected was recorded against time and filtration curves for a range of pressures in the batch Nutsche are shown in figure 3.5. These curves show how the mass of filtrate builds up over time. It can be seen that higher pressures yield higher flux rates enabling a fixed amount of filtrate to be collected in a shorter time. The filtration curves all exhibit the same initial steep slope gradually levelling out into a plateau. This shape is a result of the increasing resistance of the filter cake as more solids are deposited and the depth of the cake grows making it more difficult for the liquid to flow through. This build up of cake usually has the effect of improving the clarification with time since the cake captures a greater proportion of the solids.

At around 90 g of filtrate the curves all plateau and just after the plateau there is a sudden jump in filtrate mass, this occurs as the head of liquid above the cake disappears and the cake is exposed directly to the pressurised air. This signifies the end of the filtration; any liquid left within the filter is trapped within the filter cake and will be more difficult to

recover. Generally this liquid is displaced by a wash buffer to reduce product losses and the cake is dewatered either by air or mechanical compression.

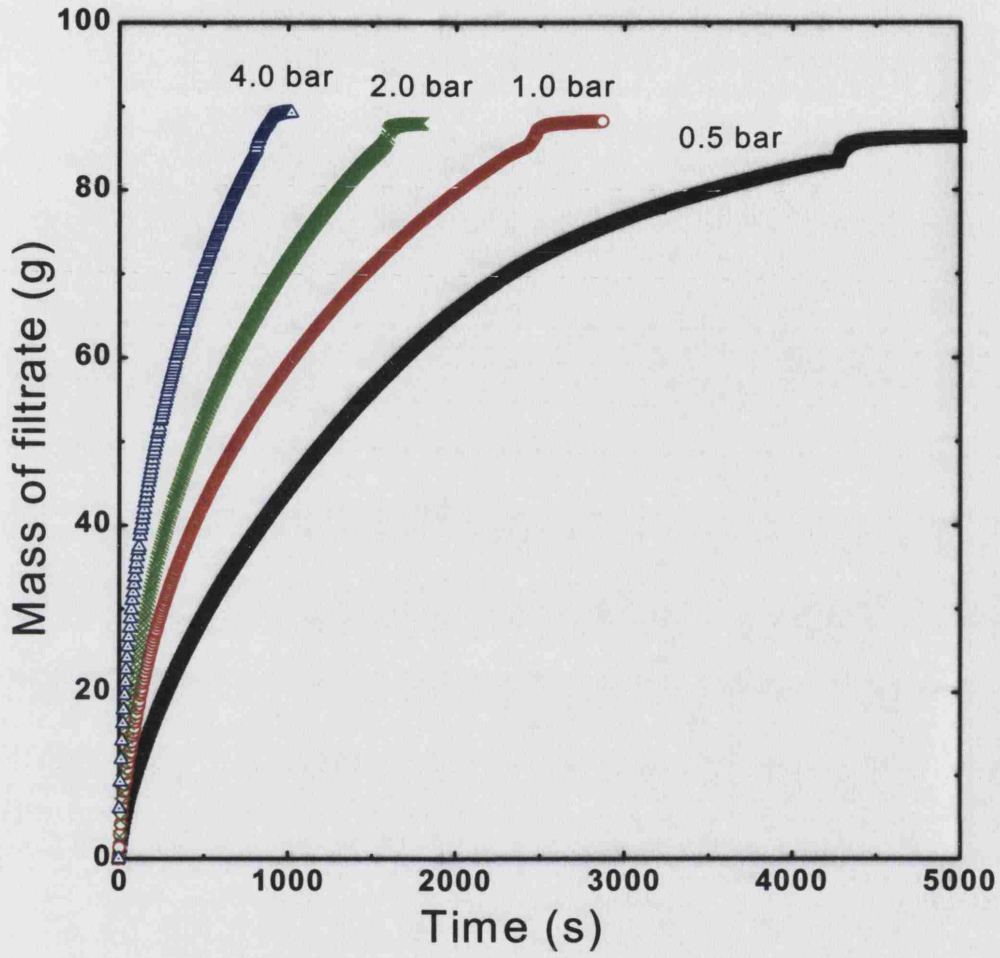

Figure 3.5: Filtration curves for constant pressures of 0.5, 1.0, 2.0 and 4.0 bar in laboratory batch filtration of Fraction IV plasma precipitate suspension. The filter aid dosage in the suspension was 0.5 kg m^{-2} pre-coat and 15 g/L bodyfeed, using *Celpure*™ 1000. The suspensions were filtered at pH 5.8 and –5 °C. For all pressures the general shape of the curve is the same, mass increases up to about 90 g and then plateaus. The time taken to reach this constant mass decreases as pressure increases. (note: all pressures are gauge)

The filtration curves can be linearised using equation 3.6 and the average specific cake resistance can be found from the gradient of this curve. Figure 3.6 shows linearised plots of the filtration curves for different pressures. There are some deviations from linearity, predominantly at the end of filtration. These deviations arise since all the liquid above

the cake has filtered through, signalling the end of filtration and the start of the dewatering process. There is also a small deviation from linearity at the start of filtration; this arises as a result of the time delay between pouring the sample into the filter and securing the top on the filter and applying the pressure.

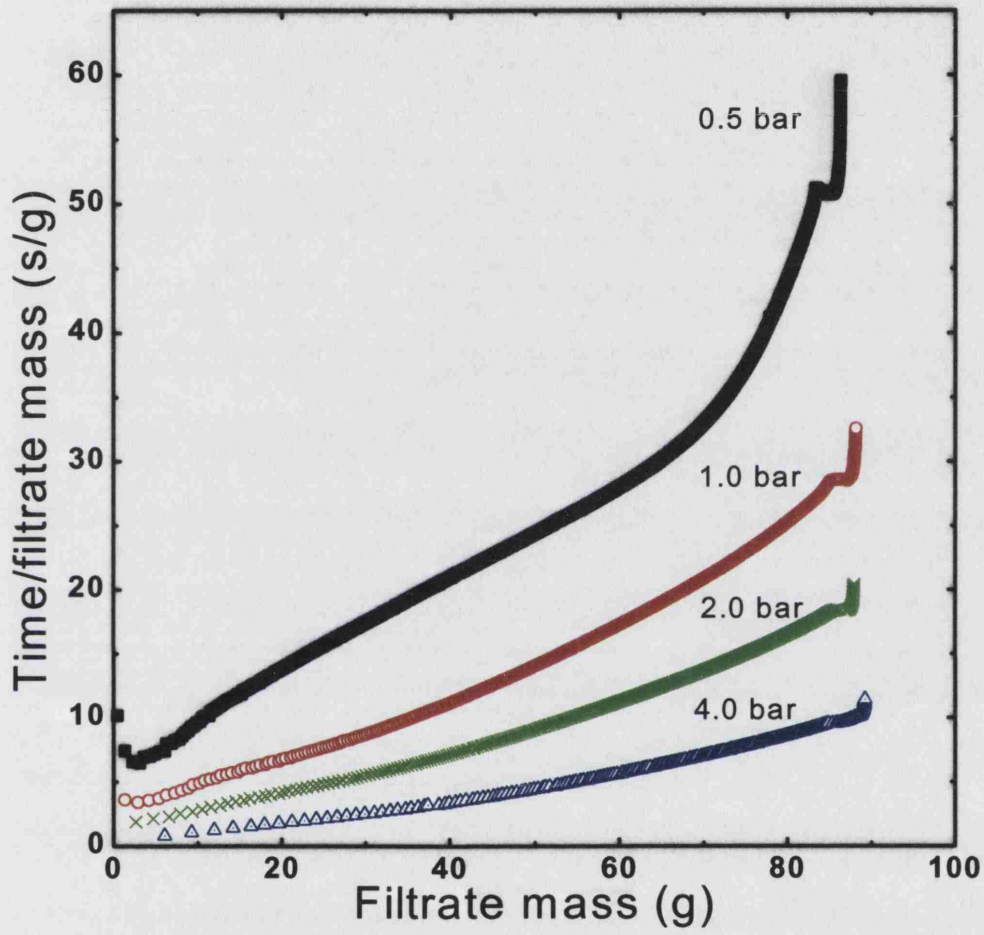

Figure 3.6: Linearised plots of the filtration curves displayed in figure 4 showing the dependence of the ratio of time to mass of filtrate on mass of filtrate. Conditions are as per figure 4. Beginning and end of the curves are non-linear due to pressure build up and dewatering respectively. The slopes for the linear region of each curve (20-80 g) were calculated to be 0.51, 0.32, 0.21 and 0.13 s/g^2, in order of increasing pressure. The specific cake resistances (α_{av}) for the pressures 0.5, 1.0, 2.0 and 4.0 bar were calculated to be 3.02×10^{11}, 3.82×10^{11}, 5.11×10^{11} and 6.34×10^{11} m/kg respectively. (note: all pressures are gauge)

The slope of the plot can be used to calculate the average specific cake resistance (see equation 3.6), when doing so the final dewatering region is ignored. Since the actual

resistance of the cake varies with depth an average cake resistance is determined. The average specific cake resistances for 0.5, 1.0, 2.0 and 4.0 barg were calculated to be $3.02x10^{11}$, $3.82x10^{11}$, $5.11x10^{11}$ and $6.34x10^{11}$ m kg^{-1} respectively. As the pressure increases the specific cake resistance also increases since higher pressures lead to a more compact cake, which has greater resistance to flow. Up to a limit higher pressures will enable higher flux rates. However once the cake becomes too compact to allow liquid to flow through filtration will stop completely and the equipment may become damaged.

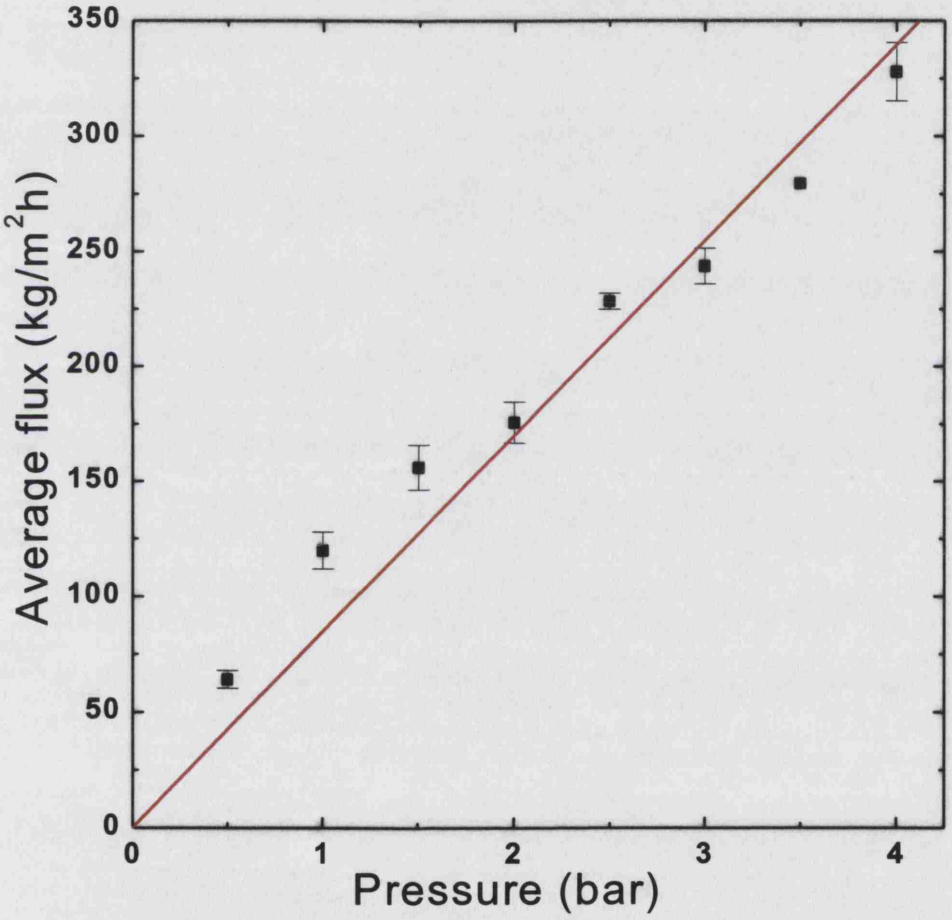

Figure 3.7: The average flux at various constant pressures in the batch filtration of fraction F.IV (precoat of 0.5 kg m^{-2} and body feed of 1.5 kg m^{-3} using Celpure 1000, pH 5.8, -5°C). The line of best fit was determined using linear regression with the origin as a fixed point. Error bars indicate 95% confidence interval for multiple runs (≥ 3).

The average flux rate during filtration is shown for several pressures in the batch Nutsche in figure 3.7. The average flux was determined by dividing the total amount of filtrate collected by the total time taken for filtration to go to completion, ignoring the final dewatering region. An approximately linear relationship between pressure and flux is obtained; hence doubling the pressure should halve the filtration time. However, high pressure filters are more expensive to run, difficult to contain and are more prone to media blinding than lower pressure filters. For these reasons industrial filters rarely operate at pressures above a few bar.

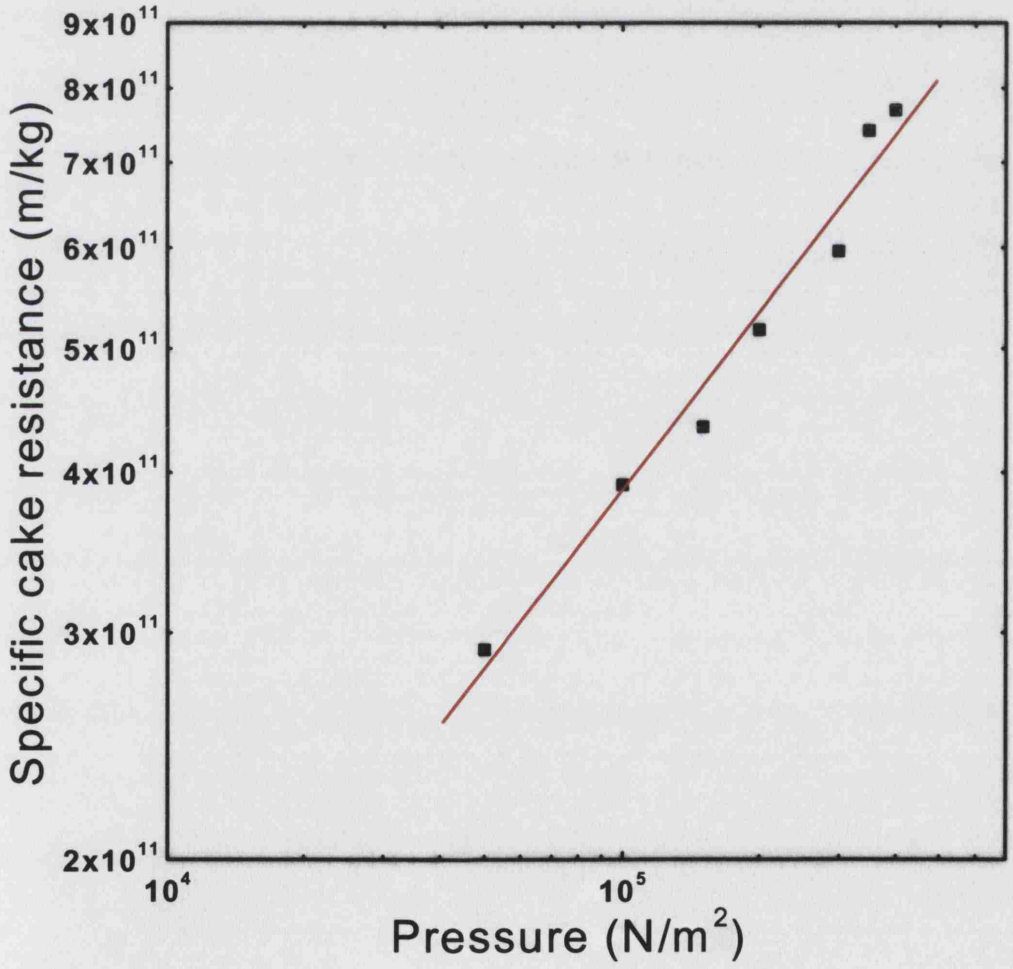

Figure 3.8: The average specific cake resistance plotted against pressure on logarithmic axes. The pressures used in the filter varied from 0.5 to 4.0 bar. The slope of the line gives the cake compressibility coefficient, which was found to be 0.46 and the intercept gives α_0, which was found to be 3.45×10^9 m/kg.

The compressibility curve for F.IV batch filtration is shown in figure 3.8. This curve is a summary of the average filter cake resistance encountered at all the pressures studied (equation 3.7 plotted on logarithmic axis). The data shows strong linear relationships between log pressure and log specific cake resistance. From the slope of this curve the compressibility coefficient for this system was found to be 0.46 and from the intercept the value of α_0 was found to be 3.45×10^9 m kg^{-1}. These parameters are assumed to be constant values for a given system and can be used in combination with a given pressure profile to predict the flux profile obtained in continuous filtration using equation 3.9.

3.6.3 Continuous laboratory filtration

The modified Nutsche filter (ultra scale-down filter) was operated in a variable pressure variable rate mode, where pressure was increased to a maximum and then held at a constant value of 1.5 bar. The resulting filtrate flux and pressure profiles obtained in the ultra scale-down filter are shown in figure 3.9 for two different pump settings. An initial high flux was obtained which then rapidly decreased; this decrease occurred before the pressure had reached its maximum value and become effectively constant. The main cause of this is thought to be the accumulation and compression of material within the filter. As the suspended solids accumulate in the filter cake, layers of solids near to the filter medium are compressed and increase resistance to flow. This causes a concomitant increase in the backpressure to the pump. The increase in backpressure then causes the feed rate and hence flux rate to decrease. It may be possible to achieve a more constant flux by using a different pump with a flatter delivery head. However achieving truly constant flux is very uncommon in industrial filtration operations.

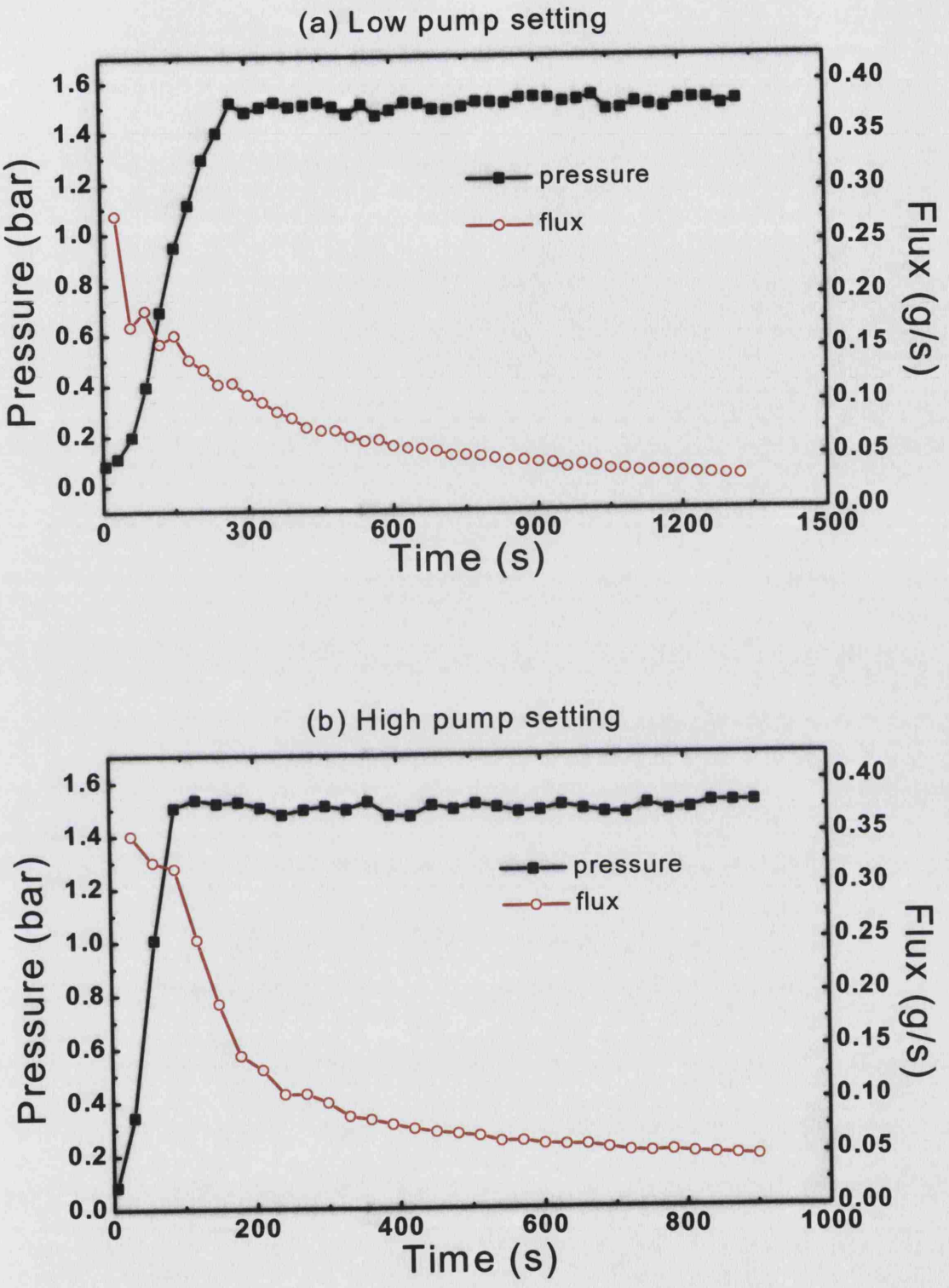

Figure 3.9: The pressure and flux profiles from the ultra scale-down filter at two different pump settings. In both cases the pump was initially set to a constant rate and pressure allowed to build up, then once the pressure reached 1.5 bar the pump rate was reduced to maintain the pressure at this level. Graph (a) is for a low pump setting and corresponds to an initial flowrate of 0.28 g/s. Graph (b) is for a higher pump setting and corresponds to an initial flowrate of 0.35 g/s. (note: all pressures are gauge)

Using equation 3.9 it is possible to predict the flux profile in the ultra scale-down filter from the applied pressure profile. However in order to do this the values of α_0 and n must be known for the slurry being filtered. Figure 3.10 shows these predictions using α_0 and n values determined from the batch Nutsche experiments. The author recognises that using one scale-down technique to predict the performance of another is unusual and of little commercial significance but it is useful in illustrating the importance of the differences between the two methods. It can be seen that the predictions do not match the experimental values closely. In the early stages of the filtration the predicted flux is much lower than the experimental flux but as the process continues the experimental and predicted fluxes become very similar. The discrepancies between the predicted and actual fluxes are most likely caused by the evolving properties of the cake. In batch filtration it is assumed that the cake properties are established very quickly and are approximately constant throughout filtration. However figure 3.10 suggests that this may not be the case and that during the initial stages of cake formation properties such as compressibility and hence porosity may differ from the average value.

The structure of the cake formed during filtration is extremely important in determining both the flux rate and the degree of clarification achieved. An open or more porous cake structure allows liquid to pass through relatively unhindered leading to greater filtrate flux rates. However the same cake will have poorer clarification than a less porous cake since particulates are less likely to be removed. For these reasons the differences in cake formation between batch and continuous operation will affect performance and make it difficult to use one form of filtration to predict the outcome of the other. Furthermore the properties of the cake change with time.

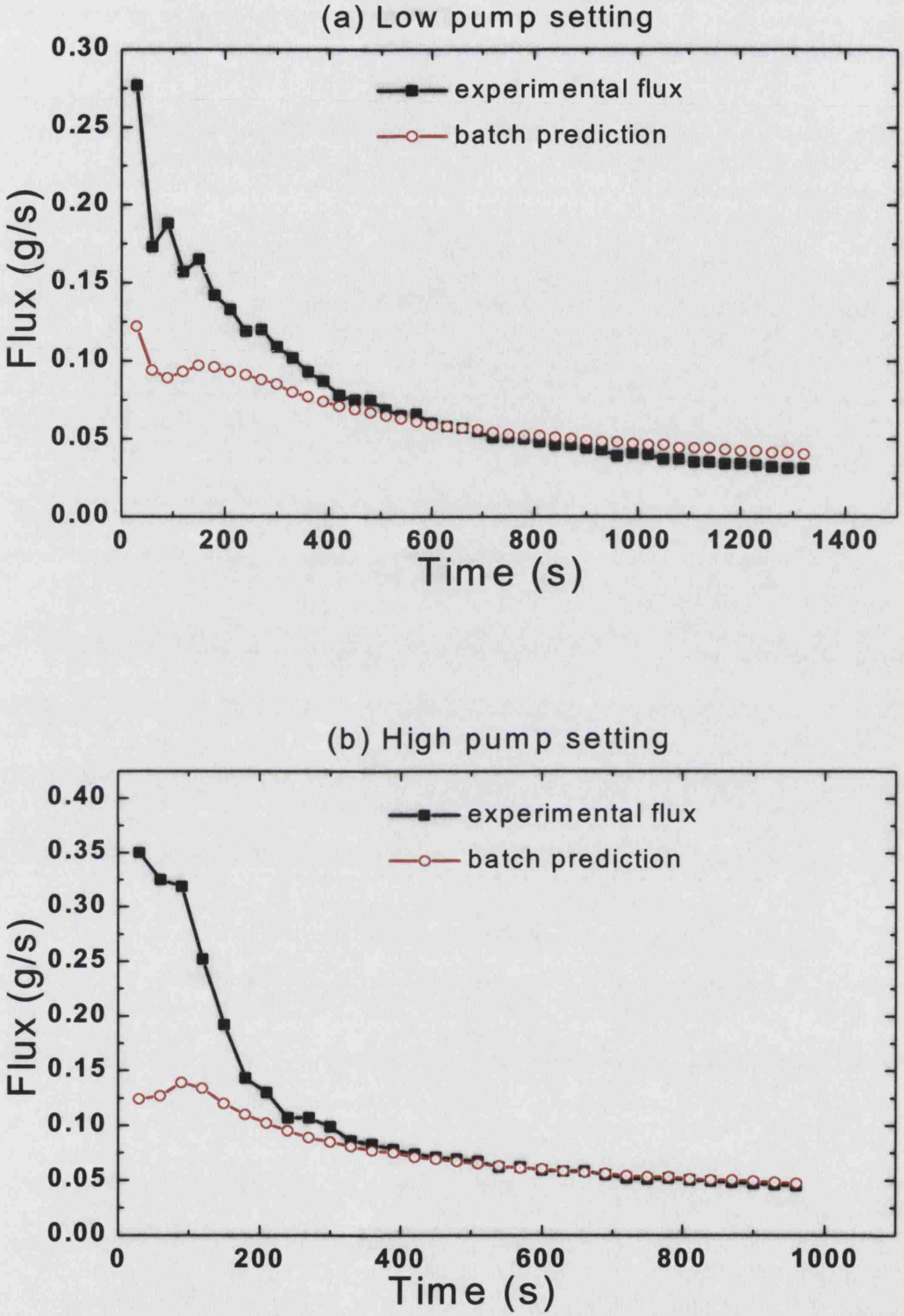

Figure3.10: Experimental flux in the ultra scale-down filter and predictions of the flux from batch Nutsche data. The predictions from the batch Nutsche are significantly lower than the experimental flux at the beginning of filtration but improve with time. This suggests that the batch Nutsche does not predict what happens during the initial stages of cake formation, although there could be other explanations.

Figure 3.11 demonstrates the differences in cake structure that arise as a result of the way filtration is carried out. In batch filtration the filter aid body-feed which is much denser than the protein precipitates settles out very quickly forming a layer on top of the precoat. This layer is fairly porous since it is mostly filter aid and allows the liquid to flow through easily. As the liquid flows through the precipitate particles are blocked by the filter aid and begin to form another layer in the cake, which consists almost solely of protein. Since the precipitate particles have a wide size distribution and are compressible the porosity of this section of the cake will be low and the resistance to liquid flow is high. The formation of this low porosity protein layer causes the flux to decrease more rapidly than expected but improves the clarification since very few particulates pass through.

In continuous filtration (sometimes called dynamic filtration) the body feed filter aid is mixed with the slurry and pumped into the filter throughout the filtration process. Providing that the mixing in the feed vessel is good the filter aid will stay in suspension then as the mixture is pumped into the filter the cake builds up. The cake formed will have a more open structure than that formed in batch operations since the filter aid is evenly distributed throughout the cake. Evidence for this was obtained by drying filter cakes from the USD and batch Nutsche filters then cutting them in half for inspection. The cakes from the batch Nutsche had a clearly stratified structure whereas those from the USD filter had an even appearance. The formation of this homogenous cake gives improved flux properties and generally allows more material to be filtered through a given area before the flux falls to unacceptably low rates. The use of filter aid is particularly important for biological solids since their compressible nature results in filter cakes of very low porosity that can completely stop filtrate flow (as seen in this study, see section 3.6.1).

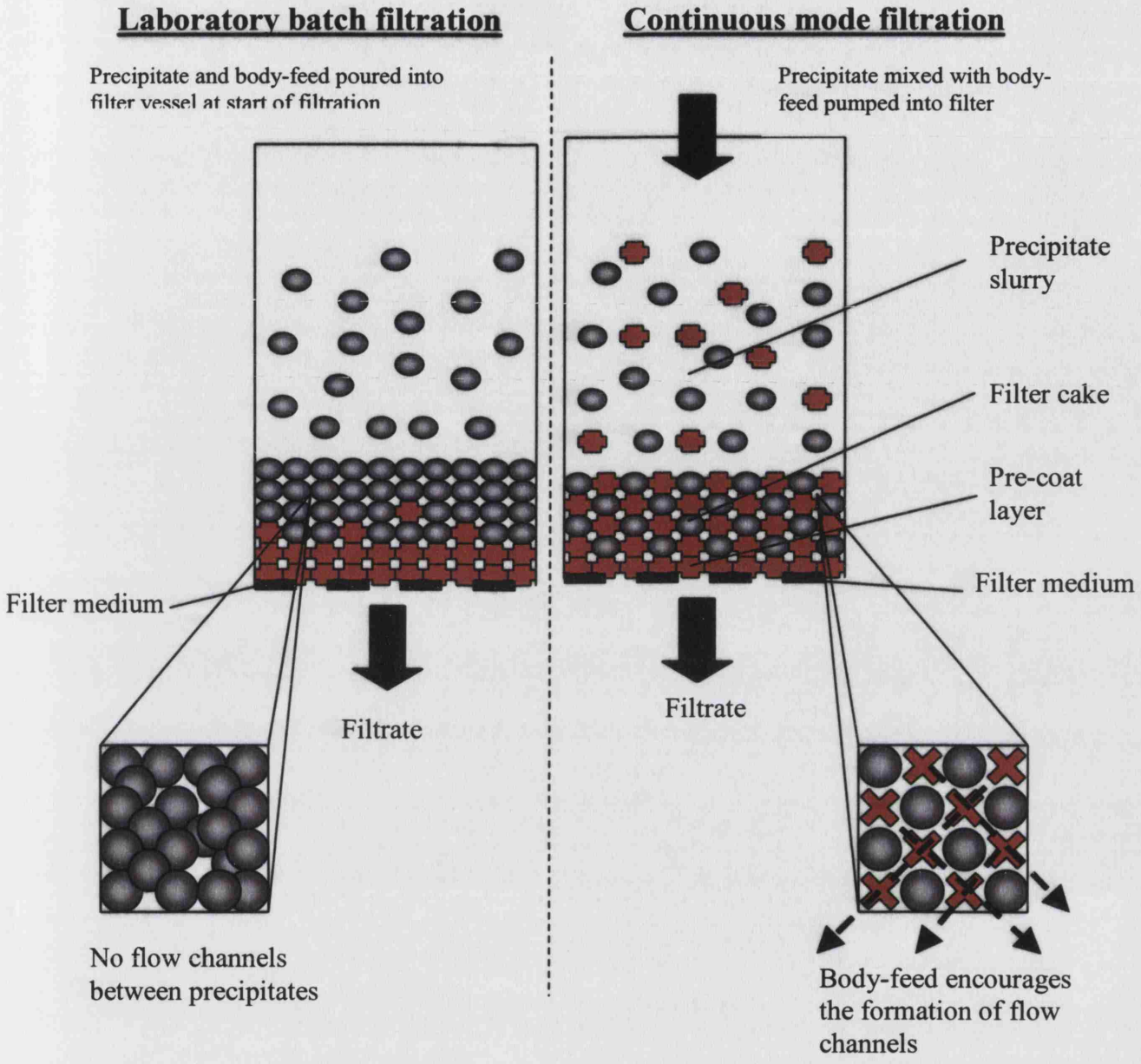

Figure 3.11: The different cake structures that arise from batch and continuous filtration. The even distribution of filter aid within the cake encourages the formation of flow channels which improve flux rates.

Equation 3.10 assumes that the cake properties are quickly established in comparison with the overall time of filtration. However with highly compressible biological materials such as protein precipitates this may not necessarily be the case. The value of α_0 is likely to be a constant since it can be defined as the resistance of the material at unit pressure and hence is only dependant on the material properties. However the

compressibility coefficient of the cake may vary during the filtration. The value of the compressibility coefficient is likely to start off at a low value (highly compressible cake) and then consolidate to a higher value. This would be a time-dependant process and a first order style relationship is assumed:

$$n = n_\infty - e^{-bt}(n_\infty - n_0) \qquad (3.11)$$

Where b is a constant, t is time and n_0 and n_∞ are the upper and lower limits of the compressibility coefficient.

To solve equations 3.10 and 3.11 without any prior knowledge of the cake an iterative process must be used. Values are assumed for n_0, n_∞ and b and then n is calculated for the filtration and inserted into equation 3.10. Plotting equation 3.10 allows α_0 to be calculated which is in turn used to find how n varies with time (by matching flux predictions to experimental data). An equation is fitted to the n verses time data and from this the values of the constants n_0, n_∞ and b are found. New values for n are then predicted and the process is repeated again until the vales of α_0, n_0, n_∞ and b converge. If the parameters in equation 3.11 are known then it will predict the relationship between n and time in any continuous filtration operation, providing the feed is not altered. This in turn could be used in equation 3.9 to predict the flux. Finding α_0 and n from the USD filter requires an iterative process. By combining the independent variables pressure and flux as shown in equation 3.10 it is possible to find α_0. Start values are assumed for the constants in equation 3.11 and a value for α_0 is then found which can be used to find a new equation for n. This process is repeated until the values consolidate.

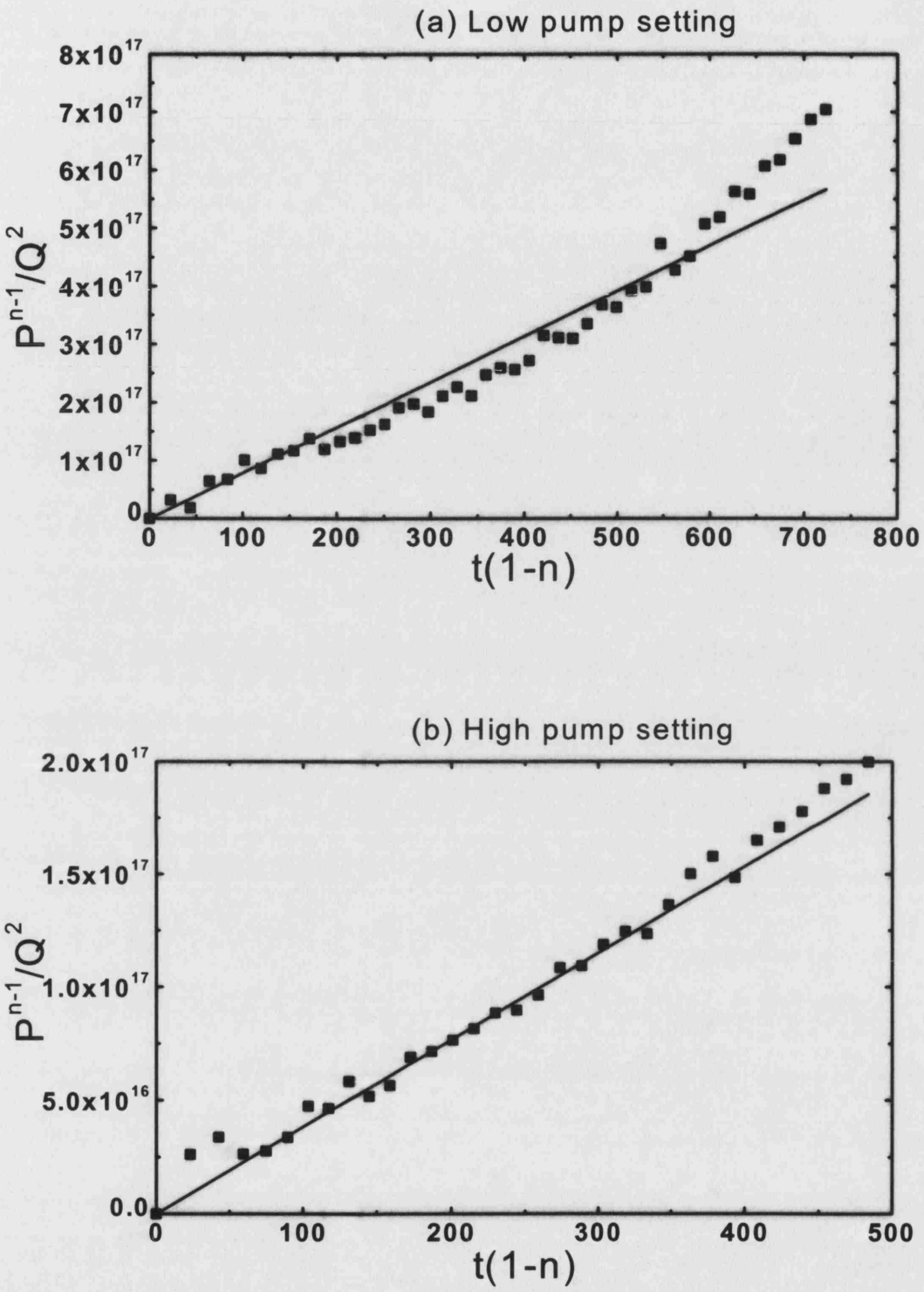

Figure 3.12: Linearised filtration plots for the variable-pressure variable-rate operation of the ultra scale-down filter (see equation 3.10). The gradient of the curve can be used to find the value of α_0, which was found to be 2.459×10^9 m/kg at the higher pump setting and 5.021×10^9 m/kg at the lower pump setting.

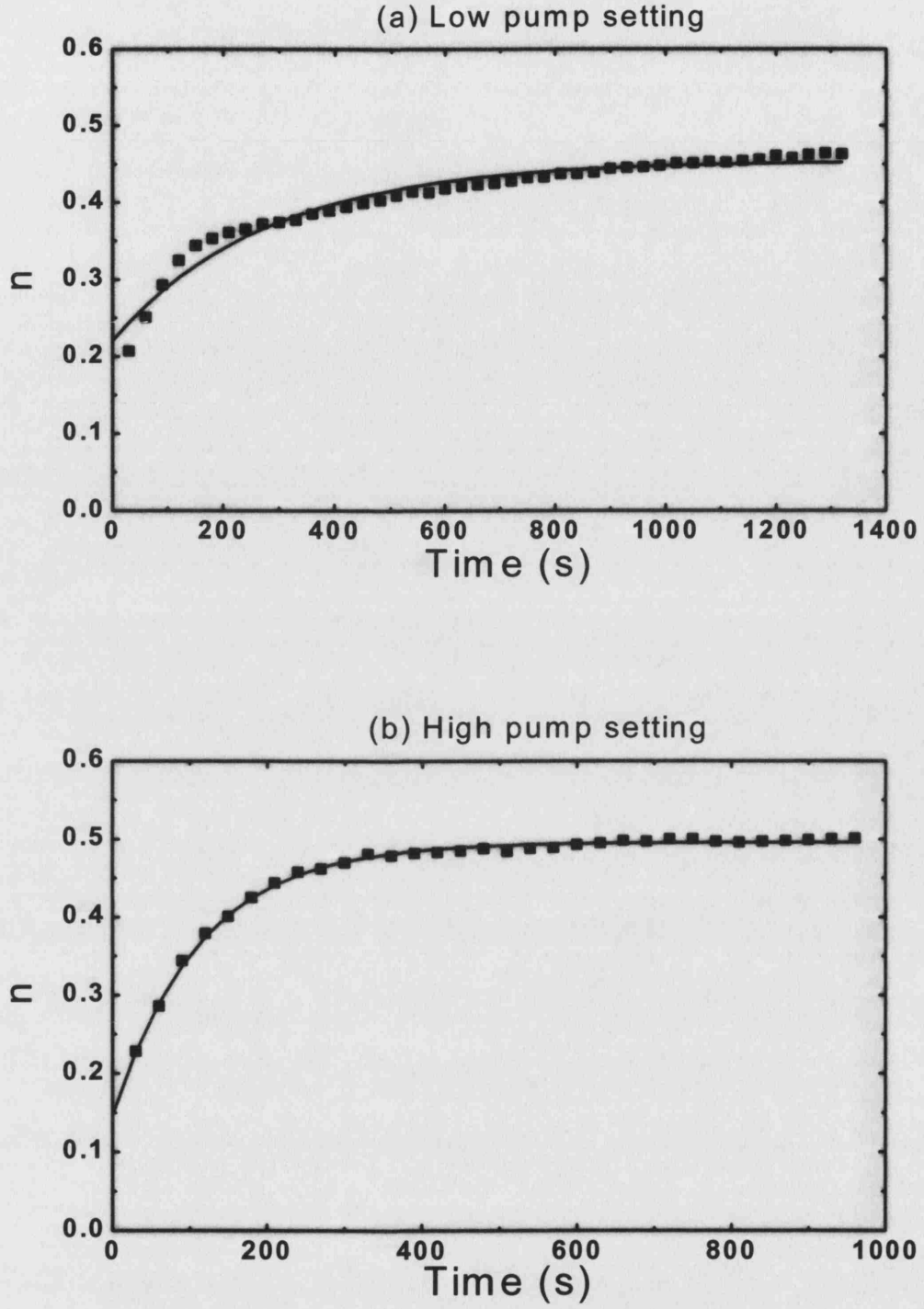

Figure 3.13: A graph showing how n varies with time in the continuous ultra scale-down filter (higher pump setting). The value of n increases in an exponential manner from an initial value before reaching a plateau. The equation for n is of the form:

$$n = n_{\infty} - e^{-bt}(n_{\infty} - n_0)$$

From the above plot the constants n_0, n_{∞} and b were found to be 0.10, 0.365 and 0.009 respectively at the high pump setting and 0.22, 0.23 and 0.004 at the low pump setting. This relationship for n can be combined with the value of α_0 and used to predict the flux in a large-scale filter using the same feedstock.

Figure 3.12 shows the final values for α_0, which were found to be 2.46×10^9 m kg^{-1} at the high pump setting and 5.02×10^9 m kg^{-1} at the lower pump setting. Figure 3.13 shows the final relationships between n and time, which were:

Low pump setting: $n = 0.454 - e^{-0.004t}(0.454-0.221)$

High pump setting: $n = 0.496 - e^{-0.009t}(0.496-0.147)$

It is difficult to understand the physical significance of a cake compressibility that changes with time. Some analogies can be drawn between this concept and the phenomenon of creep whereby materials deform over time when a constant stress is applied. It may be that the apparent changing compressibility of the cake is due to the time taken for particles to deform and find a stable shape and position within the filter cake.

3.6.4 Pilot-scale predictions

To test the above theories equation 3.9 was used to make predictions about the flux in a pilot scale candle filter. This was to see whether the ultra scale-down filter would predict the performance of a larger continuous filter better than the batch Nutsche. Figure 3.14 shows all the predictions along with the experimental flux in the candle filter. As observed with the ultra scale-down filter, predictions from the batch Nutsche were significantly lower than the experimental flux at the beginning of filtration, but improved as filtration progressed. In contrast the predictions from the ultra scale-down filter matched the experimental flux in the candle filter closely throughout the filtration. This supports the theory of changing cake compressibility and highlights the advantages of continuous laboratory filtration.

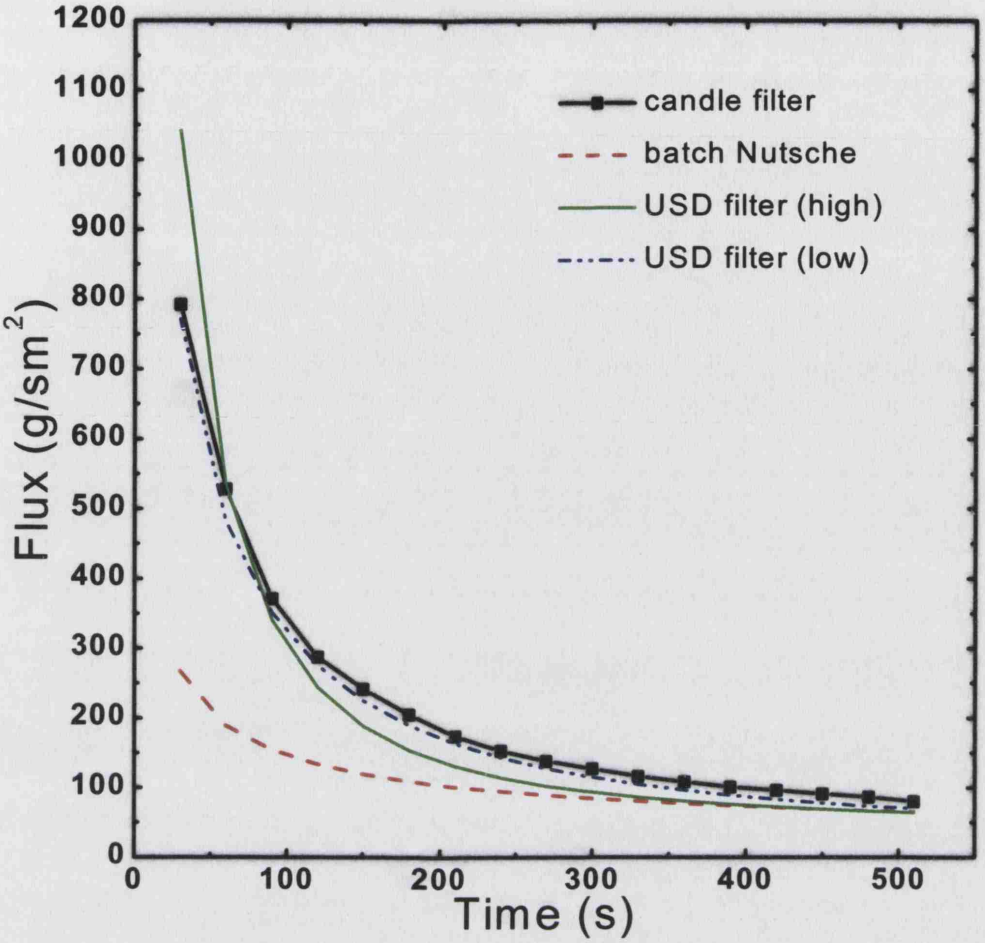

Figure 3.14: A plot to show the actual and predicted flux profiles in the candle filter. The predictions from the USD filter closely represent the flux from the candle filter. The predictions from the batch Nutsche differ significantly to the candle filter at the start of filtration but after about 500 s they match quite closely. This demonstrates that the initial period of cake formation in continuous operation is different to that of batch operation. By generating a relationship between time and cake compressibility the USD filter can be used to predict the flux of a continuous filter accurately, whereas the batch predictions only become reliable after the initial stages of cake formation.

3.6.5 Comparison of filtration and centrifugation

Table 3.1 provides a comparison between batch and USD laboratory filtration and pilot-scale filtration in terms of clarification and sediment dewatering. Clarification in the filter was found to be mainly dependent on the application of a pre-coat, and practically independent of pressure above 1.0 bar. The USD filter performed slightly worse than the batch Nutsche, but had very similar clarification to the candle filter. This poorer

clarification is likely to be caused by the initial period of low-pressure filtration in continuous operation, where there has been very little cake build up; hence fine particles can pass through. At industrial scale this can be avoided by recycling the early filtrate until sufficient cake has built up.

Pressure (barg)	Sediment dryness (%)	Solids carryover (%)
0.5	44.6 ± 1.5	2.7 ± 2.1
1.0	47.9 ± 1.5	1.0 ± 0.5
1.5	50.0 ± 2.0	0.5 ± 0.2
2.0	51.4 ± 1.8	0.8 ± 0.2
2.5	54.8	1.0
3.0	52.4 ± 1.3	1.2 ± 0.1
3.5	57.4	1.0
4.0	60.2 ± 5.9	1.1 ± 0.2
0.5-1.5-2.5-3.5	61.1 ± 0.5	0.9 ± 0.5
USD filter (Low pump setting)	31.9 ± 1.0	1.8 ± 0.7
USD filter (High pump setting)	30.9 ± 1.4	1.4 ± 0.5
Candle filter	48.2 ± 2.3	1.8 ± 0.3
Reference centrifuge (RCF_{mzx} = 8010, t = 0.25 h; $V_{lab}/t_{lab}C_{lab}\Sigma_{lab}$ = 1.7 x10^{-9} m/s)	35.0 ± 2.7	0.9 ± 0.2
Laboratory centrifuge (RCF_{mzx} = 8010, t = 0.25 h; $V_{lab}/t_{lab}C_{lab}\Sigma_{lab}$ = 15 x10^{-9} m/s)	44.1 ± 1.1	23.5 ± 3.4
Production centrifuge (RCF_{mzx} = 8140, t = 8.0 h; $Q_{mc}/C_{mc}\Sigma_{mc}$ = 14 x10^{-9} m/s)	40.6 ± 2.4	54.4 ± 2.6

Table 3.1: A comparison of the sediment dryness and clarification achieved in the centrifugation and filtration of Cohn Fraction IV (from human blood plasma) at various conditions.

Much lower levels of dewatering were achieved in the ultra scale-down filter (30.9% and 31.9% dry weight) than in the batch Nutsche (50.0%), since there was no air-drying. The candle filter was operated with and without air-drying. However, without air-drying the cake in the candle filter would come away from the filter medium and form sludge in the bottom of the filtration chamber. Cake dewatering in the candle filter was dramatically improved by air-drying (48.2%) and this would doubtless be the case in the USD filter.

3.7 `Conclusions

Using the results from the USD filter to predict how the cake compressibility changes with time allows accurate predictions to be made about the flux profile in larger continuous filters. These predictions are superior to those made using a traditional batch filtration approach as they also predict the flux in the high rate initial period of filtration. In addition to this the USD filter allows the production of materials (cake and filtrate) representative of a large-scale continuous filter. Which enables the laboratory production of material for further studies. This could be of vital importance when studying subsequent operations to filtration, such as chromatography. The material used in chromatography studies needs to be as close as possible to actual process material in order that the complex interactions of feed and matrix can be monitored and the operation can be optimised.

The USD filter provided an insight into the structure and properties of highly compressible biological filter cakes. Such cakes have extremely low permeabilities if filter aid body feed is not used resulting in a rapid decrease of filtrate flux. The structure of the cake can have a major influence on all aspects of filter performance and thus it is important in scale-down to be able to mimic cakes formed in industrial filters. The even distribution of filter aid throughout the cake is essential for high flux rates and minimises the effects of compressible solids on cake permeability.

CHAPTER 4
ULTRA SCALE-DOWN OF CHROMATOGRAPHY

4.1 Introduction

This chapter presents a study into the scale-down of ion exchange chromatography and explores some of the problems associated with very small columns. Traditionally chromatography operations are scaled up several times from the laboratory, through pilot-scale to full scale in order to avoid any undesirable surprises. This process may well involve some degree of parameter adjustment at each scale in order to maximise performance. However this very cautious approach is wasteful of material and slows down the overall process development cycle.

Traditional chromatography scale-down is based around maintaining the column bed height and linear velocity of the sample. This method has been shown to work well for columns with diameters in a certain range but for very large or small columns linear scaling becomes more difficult. Changing column bed height enables more practical small-scale columns to be used and can provide greater flexibility at pilot or production scale (Hansen, 2003). However there is a reluctance to move away from the traditional approach to scaling chromatography operations. This chapter looks into how corrections can be made if these rules cannot be followed. The intended result is a very small, ultra scale-down column and the method by which it can be used in process development.

Ultra-scale down techniques for chromatography that could predict the performance of much larger columns would be an extremely useful tool for rapid process development. However, even when following typical chromatography scale-down rules the performance of very small columns can differ significantly from larger columns due to a number of effects. The biggest obstacle to ultra scale-down chromatography is

identifying these effects and developing ways to prevent them or adjust results to correct for them. If this can be achieved it may be possible to scale-up columns across several orders of magnitude quickly and easily and optimise processing conditions at an early stage, thus reducing process development times.

4.1.1 The industrial importance of chromatography

Chromatography has for many years been used as an analytical tool in the chemical industries. Over the last few decades it has also become firmly established at the processing scale as a method of isolating pure substances for the pharmaceutical and biotechnology industries (Sofer, 1989; Rathore, 2003). The wide varieties of solid (stationary) and liquid (mobile) phases make it a very flexible operation, which can be used in the purification of a great number of substances ranging from complex organic molecules (e.g. antibiotics) to inorganic chemicals (e.g. industrial dyes). In many cases the only way to achieve the high purities necessary for these substances is via the use of chromatography.

4.1.2 Chromatography and biotechnology

Chromatography is one of the most important and widely used operations in the biotechnology industry. It can fulfil a variety of roles within the downstream processing sequence such as primary capture, contaminant removal or final product polishing. Chromatography has become an integral part of biotechnology and is used in the purification of virtually every biological product produced commercially (Bonnerjea,1986). This widespread usage is likely to continue as the biotechnology market expands and new groups of products such as DNA vaccines are commercialised.

The high levels of purity required by regulatory bodies in the biotechnology industry are almost exclusively achieved via the use of chromatographic separations since few other operations can achieve such purity at an industrial scale (Heuer, 1998). In addition to the high purity that can be achieved chromatography also has the advantages of being a low temperature and low shear operation. These factors are critical in retaining the function of sensitive bio-molecules. In fact chromatography is so important within the industry that regulatory agencies such as the US Food and Drug Administration (FDA) consider at least one chromatography step to be essential in the purification of therapeutics (Sofer et al, 1989).

Although the process is well established, technological advances continue to be made within the field of chromatography. Increased process automation and greater use of instrumentation can make the process simpler, more efficient and more reproducible. Sanitary design of columns and ancillary equipment has eased the regulatory burden and new chromatography media enable higher binding capacities and throughputs (Boschetti, 1994). In addition, new techniques such as simulated moving bed (SMB) chromatography and expanded bed chromatography (EB) have further broadened the range of uses for chromatographic separations.

4.1.3 Chromatography in plasma fractionation

The excellent safety record and ability to handle large process volumes has made cold ethanol precipitation the most widely used technique for plasma fractionation (More and Harvey, 1991). However the low operating temperatures of the process result in very high running costs, which are often only viable in very large, well established fractionation facilities. This has encouraged many newer and smaller fractionaters to look at chromatography as an alternative method. Yap et al (1993) developed a

chromatographic process for the preparation of Human Serum Albumin (HSA) from pooled plasma. The resulting process consistently had higher yields and produced a product with greater purity and monomer content than that of traditional fractionation processes.

As a processing technique chromatography has several advantages over the traditional cold ethanol method. The principal advantages being that very high recoveries and purities are possible, running costs are greatly reduced and a greater degree of process automation is possible (Curling, 1977; Yap *et al*, 1993). In addition, the mild processing conditions of chromatography ensure that protein denaturation and damage is minimised.

Plasma fractionaters around the world have employed a variety of chromatographic techniques for protein purification. Some processes use chromatography as the main method of protein separation whereas others have integrated chromatography into more traditional fractionation methods. The most commonly used type of chromatography and the first method developed involved the use of ion exchange resins. Since then highly specific affinity resins have been successfully tested for some applications, but the cost of theses resins generally makes them less appealing. Many fractionators (including BPL) have investigated chromatographic methods as a way of improving yields of IgG, albumin and other plasma proteins and in the future chromatography may eventually replace cold ethanol precipitation as the predominant method.

4.2 Development of a chromatographic process for the manufacture of high purity IgG from human plasma

All experimental work in this chapter focuses on a chromatographic process that is being developed at Bio Products Laboratory (BPL). The process is being developed as a

potential alternative to cold ethanol fractionation for the recovery and purification of polyclonal IgG from human blood plasma. Prior to chromatography Fraction A+1 paste (from the Cohn fractionation process) is suspended in deionised water (1 part paste to 10 parts water). The suspension pH is adjusted and then it is mixed to extract and dissolve all the IgG present in the paste. After mixing, the suspension is filtered to remove all the insoluble material. The filtrate is then diluted and the pH and conductivity are adjusted ready for column loading. Some experimental work has already been performed by BPL to determine the best resin to use (resin screening) and optimal loading conditions. This work will not be shown here but the outcomes from the previous work will be used as a basis for the chromatography in this thesis. A process flowsheet covering load preparation and chromatography is shown in figure 4.1.

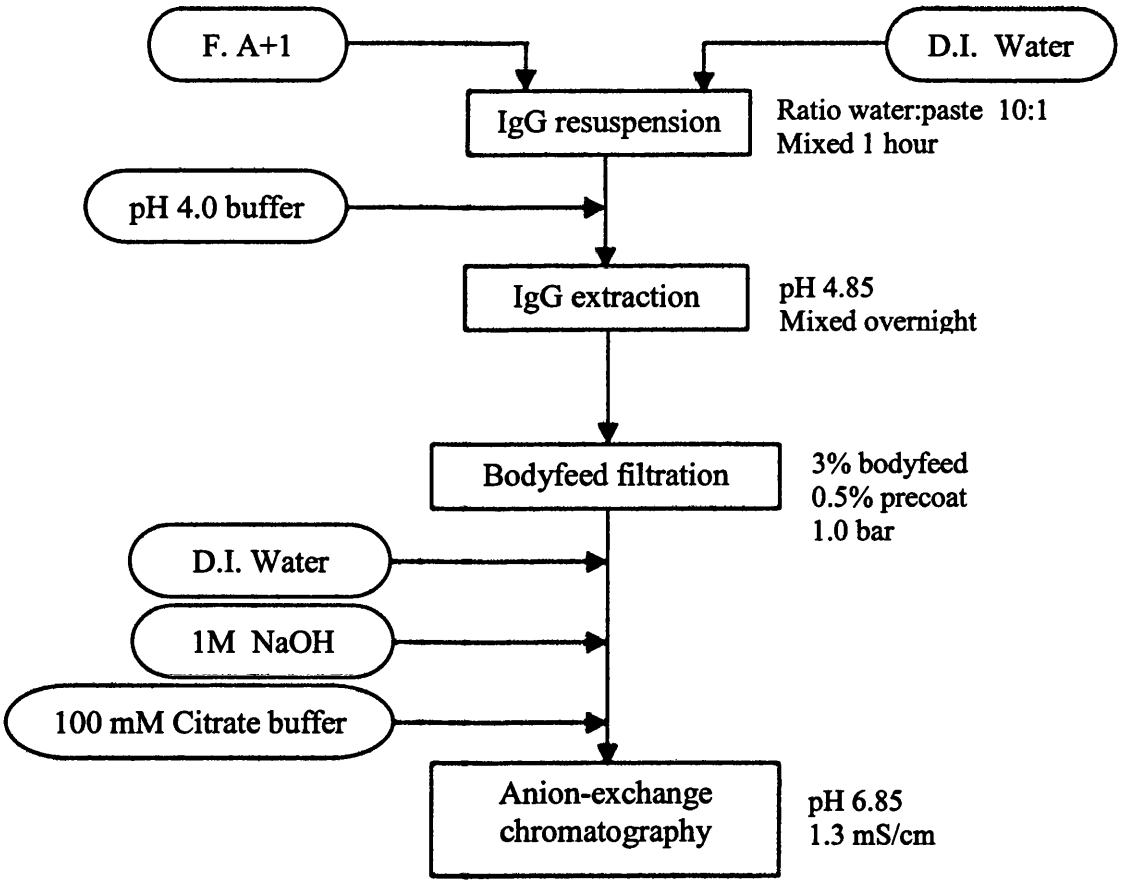

Figure 4.1: A diagram to show the major steps and conditions involved in the purification of IgG from human blood plasma. This diagram is based upon the laboratory scale process used during the early stages of process development.

The chromatography step itself is to be a contaminant capture step, so IgG will pass straight through the column while the matrix will capture other proteins. The reasons for using contaminant capture at this stage (as opposed to product capture) are to reduce the column size required and to overcome patent restrictions. The bound material will then be eluted in a single step and disregarded, while the flowthrough material will be further purified most likely with more chromatography steps. Early development work has identified the most suitable matrix and binding conditions for this step (pH, conductivity, loading buffer, linear velocity, etc). These conditions will be fixed so that the performance of the chromatography step can be determined. This step will then be scaled down to the millilitre scale and the influence of operating scale on performance will be studied.

4.3 Fundamental concepts in chromatography

Chromatography is essentially a separation process in which target molecules are selectively bound based on some physiochemical property. A chromatography system will consist of two phases, one mobile and one stationary. Generally a multi-component mixture (the mobile phase) is loaded onto a column containing the chromatography matrix (the stationary phase). Since the components within the mixture have different properties they are differentially separated, so that components with a high affinity for the solid phase (matrix) are removed from the liquid phase first and then components with progressively lower affinities. Components with no affinity for the matrix remain in the liquid phase and pass straight through the column.

4.3.1 Structure of the stationary phase

The stationary phase in chromatography is commonly referred to as the matrix or media. It consists of tiny porous beads, which are roughly spherical in shape. The matrix is

poured into the column as a slurry then packed down by pumping a suitable buffer through the bed of particles. The size of the particles in the matrix varies from 5 to 200 microns, with most applications using particles of between 30 and 90 microns. Smaller particles give rise to higher-pressure drops across the bed while large particles have low surface area to volume ratios, thus reducing column capacity. Usually the beads contain pores, which greatly increase the surface area available for binding so long as molecules are small enough to penetrate these pores. Other beads are composed of smaller particles fused together (macro-reticular), which also confer a very large surface area. The surface of the beads are covered to various degrees with functional groups referred to as ligands, these interact with the target molecules and bring about binding.

The ideal chromatography matrix would be highly porous to increase surface area, hydrophilic to enable thorough wetting, chemically inert and physically strong enough to resist compression caused by high liquid velocities. The majority of matrices are made from polysaccharide based materials such as cellulose, agarose and dextran. These materials have an abundance of free hydroxyl groups, which allow a variety of different ligands to be attached making them suitable for many applications. Although polysaccharide based matrices are generally hydrophilic and inert they are also mechanically soft and compressible, thus at high flow rates the particles deform. Other materials such as silica and glass are much more rigid but not widely used due to problems with non-specific binding and protein denaturation (Nakamura, 1998).

4.3.2 Modes of chromatography operation

There are several ways in which a chromatography column may be operated. The three main ways are elution, displacement and frontal chromatography (Garcia et al, 2002). All

three of these modes have uses in an analytical laboratory but only elution chromatography has gained widespread use at process or production scale.

In elution chromatography a sample of the mixture to be separated is loaded onto of the column then a mobile phase with a lower affinity for the matrix is loaded after it. The components of the sample migrate through the column at different rates depending on their affinity for the stationary phase and pass out the other end of the column in concentrated bands. There are two types of elution possible based on the strength of the eluting solvent (eluent). Isocratic elution is where the strength of the solvent is constant throughout the chromatographic run and gradient elution is where the strength of the solvent changes continuously.

An additional type of chromatography that could be classed as a simplified version of elution chromatography is batch adsorption. In this technique the sample is loaded onto the column as normal or mixed in a vessel along with the stationary phase. Then an isocratic solvent is passed down the column to remove all the bound protein, however all the protein migrates at a similar rate so only one band, or fraction of protein needs to be collected. This method differs from normal elution chromatography since separation is brought about by adsorption or non-adsorption and not by different rates of migration through the column.

4.3.3 Modes of interaction in chromatography

A range of chromatography operations have been successfully employed at the process scale for separating biomolecules. These operations can be classified according to which property they exploit to bring about separation. The main techniques used in the biotechnology industry are presented here.

4.3.3.1 Ion exchange chromatography

Most biological molecules, including proteins have charged molecular groups in them. These charged groups combine to give the molecule an overall net charge, the size and sign of which is determined by the molecules immediate environment. By manipulating the pH and conductivity of a particular solution the charge on a molecule can be changed. Ion exchange chromatography is based upon the interaction between charged groups on molecules in the mobile phase and oppositely charged groups on the chromatographic matrix. The three dimensional configuration of the molecule will determine which groups interact with the matrix.

There are two types of ion exchange matrix: cation exchangers (negatively charged matrix) and anion exchangers (positively charged matrix). Sample constituents bind to the matrix at a low ionic strength and are selectively eluted by increasing the ionic strength or changing the pH. The strength and selectivity of binding are determined by the initial pH and conductivity of the solution. For best results it is important to find the optimum combination of mobile phase conditions and matrix. Matrices generally vary in the type of ligand used and the density of binding sites. To find the best combination for a particular application many matrices will be tested at a variety of pH and conductivities, this scouting process can be time consuming and costly so small-scale laboratory columns are regularly used for this purpose.

Ion exchange chromatography is the most widely used type of chromatography (Bonnerjea, 1986). It is not as selective as affinity chromatography but can still be used to separate dozens of proteins when combined with gradient elution. The matrices are often cheaper than other types of matrix and the process is well suited to large-scale operation because of its high capacity and concentrating effects.

4.3.3.2 Hydrophobic interaction chromatography

Most proteins have some hydrophobic regions on their surface and although they are not as abundant as hydrophilic ones they can sometimes be used as a basis for separation. Hydrophobic interaction chromatography (HIC) is based upon the varying strengths of the interactions between bio-molecules and an uncharged media containing hydrophobic groups. In aqueous salt solutions hydrophobic groups are attracted to each other and the strength of this attraction is dependant on the hydrophobicities of the two surfaces. High salt concentrations prevent non-specific ionic interactions with the media. The bound proteins can then be selectively eluted by reducing the salt concentration or by the addition of organic solvents.

Hydrophobic interaction chromatography has the advantage that it can be used to separate proteins of a similar size and charge based on differences in hydrophobicity. The media are also generally resistant to strong acids and bases, detergents and organic solvents making them fairly simple to clean and regenerate. The main limitation of HIC is that the high salt concentrations involved may cause some proteins to denature.

4.3.3.3 Reverse phase chromatography

Reverse phase chromatography is used predominantly for analytical purposes in high performance liquid chromatography (HPLC) but has found some preparative applications. It uses hydrophobic adsorbents similar to those used in HIC the main difference being that the non-polar solvents used in elution cause proteins to denature. The method is however very sensitive and enables a high degree of resolution which is why it is such a powerful analytical tool.

4.3.3.4 Affinity chromatography

This is a highly specific technique, which involves the reversible formation of a complex between the target molecule and an immobilised ligand. The nature of the complex formed and the mechanism of binding vary as the method exploits a particular affinity of the target molecule. The exact nature of the ligand also varies greatly between applications. Some ligands have a very high specificity for a particular bio-molecule such as protein A or monoclonal antibodies, while others exhibit a high specificity for whole groups of proteins such as metal ions. Generally any molecule that can be immobilised on a suitable matrix can be utilised in affinity chromatography and as the variety of ligands increases so do the number of applications.

The extremely high selectivity of affinity chromatography makes it a powerful purification tool. Very large volumes of dilute product solution can be processed on a single column and purification factors of 1000 or more have been reported (Bonnerjea et al, 1986). Furthermore the capacity of the matrix is not reduced by non-specific binding, as is the case with ion exchange chromatography and HIC. However the cost of affinity matrices is much higher than other types of matrix due to the complexity of immobilising many of the ligands. Affinity matrices also tend to have a lower capacity than other types of matrix making very large-scale operations extremely expensive.

4.3.3.5 Gel filtration chromatography

In gel filtration chromatography separation of a mixture occurs on the basis of molecular size. The mixture is loaded onto a column packed with beads very similar to other chromatographic matrices except that theses beads have no ligands attached and so are chemically inert. The beads contain a network of pores into which the mobile phase passes and depending on their size other molecules also pass. Larger molecules can only

pass a small way into the pores or are too big to enter the pores at all. Since the majority of the bed volume is made up of the porous beads the residence time of a molecule in the column will be increased if it is small enough to penetrate the pores. Thus, completely excluded molecules pass through the column quickly while smaller molecules pass through it progressively slower.

Gel filtration is an important operation in industry, which is generally used as a polishing step towards the end of the purification sequence. It is particularly useful for removing polymers and aggregates since these are difficult to separate on a chemical basis. It is a simple operation, which requires much less method development than other chromatographic techniques since performance depends only on the matrix selected. The biggest problem with gel filtration arises from matrices, which are not truly inert and both ionic and hydrophobic interactions have been reported. Other problems are caused by the compressible nature of the matrices since separation ability is affected as the particles deform and bed porosity is reduced (Danilov et al, 1997).

4.3.4 Measuring chromatographic performance

Predicting the performance of a chromatographic process is a complex task due to the large number of variables involved and the dynamic nature of the process (Sorensen, 2002). The fundamental measures of chromatographic performance are the yield and purity of the product molecule. However in order to specify an operation that will satisfy the necessary product requirements the performance of a column needs to be correlated somehow to the system parameters. In order to describe the performance of a chromatography operation several concepts and parameters have been developed which can describe the size and shape of the chromatographic peak and the separating power of the process. A number of models and equations exist which link these performance

characteristics to the operating conditions and properties of the chromatography column. The two main categories of chromatographic model are plate models and rate models. Using these methods it is possible to predict the performance of a particular chromatographic operation or make comparisons between different columns. The following sections will review some of the main concepts used to describe chromatographic performance.

4.3.4.1 The concept of the theoretical plate

The concept of the theoretical plate is derived from the analyses of distillation column performance. Distillation columns are divided into a number of equally spaced, individual sections (called plates). In each section vapour rises, cools and eventually condenses to liquid on the plate surfaces. The vapour and liquid phases within each individual section are in equilibrium and the solute is divided accordingly between the two phases. Martyn and Synge (1941) realised that vapour liquid equilibria in distillation columns were analogous to adsorption isotherms in chromatography, which describe the partition of a solute between the stationary and mobile phases. Thus a chromatography column can be considered as a large number of separate sections in which the solute concentrations between solid and liquid phases are in equilibrium. Solutes pass down the column as equilibrated mobile phase passes from one section to the next. The number of these theoretical plates in a column gives a measure of column efficiency, with a greater number of plates enabling greater separation.

For real chromatography systems the height equivalent to one theoretical plate (HETP) can be found by examining an eluted chromatographic peak. The HETP can be expressed by the relationship:

$$HETP = \frac{\sigma_L^2}{Rut_R} = \frac{\sigma_L^2}{L} \tag{4.1}$$

where R is the retardation factor, t_R is the retention time of the component, u is the interstitial fluid velocity, σ_L is the standard deviation of the eluted peak measured in column length units and L is the length of the bed (Giddings, 1965). A low HETP value means that there is only a small amount of band broadening per unit length of column and thus the separation will be efficient. Sometimes it may be more convenient to express the standard deviation of the eluted peak in either time or volume units.

$$\sigma_L = \sigma_t Ru = \frac{\sigma_V Ru}{Q} \tag{4.2}$$

where Q is the volumetric flowrate of the mobile phase. These can be substituted into equation 4.1 to find the HETP.

From the HETP value the number of theoretical plates, N, in a particular column can be calculated either by dividing into the bed length or from:

$$N = \left(\frac{L}{\sigma_L}\right)^2 = \left(\frac{t_R}{\sigma_t}\right)^2 = \left(\frac{V_R}{\sigma_V}\right)^2 \tag{4.3}$$

The derivation of these equations is based upon statistical theory and the method of moments (see Garcia et al, 2003 for a full derivation). In practice N is usually calculated directly from a sample peak:

$$N = 16\left(\frac{t_R}{W}\right)^2 \tag{4.4}$$

where W is the peak width measured in the same units as t_R the baseline ($W = 4\sigma_t$) and can be measured directly from an experimental peak.

A high number of theoretical plates (low HETP) indicates an efficient column, which will produce narrow solute elution peaks and good separation. This concept is widely used as a measure of column efficiency since it is simple and easy to measure, however there are some limitations that should be borne in mind. Firstly the exact value of N will actually vary for different components on a column due to varying diffusion coefficients thus separation may not be as good as expected. Also, plate theory assumes that equilibrium is reached infinitely fast which is unrealistic especially for large bio-molecules. If equilibration is very slow then separation performance will be greatly affected.

4.3.4.2 The rate theory of chromatography

The actual shape of a chromatographic peak is influenced by a number of factors. As a sample travels along a chromatography column the molecules within that sample become increasingly spread out due to dispersion processes. This causes what is known as band broadening and decreases the efficiency of separation. Van Deemeter (1956) derived an equation based on the rate theory for gas chromatography, which related the separation efficiency of a packed column to the characteristics and operation of that column. Three factors were identified which contribute to overall band broadening and therefore reduce separation, these are: molecular diffusion along the column, resistance to mass transfer and flow disturbances. Since each of these factors is random and independent of the others the total band broadening effect is the sum of the individual contributions. Van Deemeter described the effects of band broadening in terms of plate heights:

$$H = H_{flow} + H_{diffusion} + H_{masstransfer} \qquad (4.5)$$

Thus the total plate height is the sum of the individual contributions. Each of these contributions will be briefly explained in the next section.

The flow of the mobile phase through and around the matrix particles is random and tortuous. Different solute molecules will take different paths and thus some molecules will take a relatively long path and fall behind the bulk while other molecules find particularly short paths and move ahead of the bulk. This dispersive process is called "eddy diffusion" and is dependant upon the size of the matrix particles and the non-uniformity of packing; it is essentially independent of mobile phase flowrate.

The molecules within the mobile phase will diffuse randomly in all directions until there are no concentration gradients present in the system. This ideal is never achieved in a chromatography column since the mobile phase continually moves in one direction and not all the molecules remain within the mobile phase for very long. The degree of diffusion that occurs and hence the contribution to band broadening depends on the mobile phase velocity, the interstitial tortuosity and the diffusivity of the solute molecule. If the velocity is low then the column residence time will be high and the effects of diffusion will be increased. However in the case of large biological molecules the band broadening effects due to diffusion are often insignificant since their molecular diffusivities are very low.

Equilibration between the stationary and mobile phases does not happen instantaneously since it takes time for a molecule to move to the phase boundary and then into the other phase. During this time the mobile phase is still moving and so the molecules also move a certain distance down the column, this hinders the approach to equilibrium and causes the molecules to spread out due to differing affinities for the stationary phase. If the mobile phase velocity is high, then the band broadening effect is worsened as molecules are spread further apart.

The van Deemter equation relates the plate height to the three independent dispersion processes and is normally written in the form:

$$H = A + \frac{B}{u} + Cu \tag{4.6}$$

where u is the mobile phase velocity and A, B and C are constants for eddy diffusion, longitudinal diffusion and resistance to mass transfer respectively. The van Deemter equation is extremely useful since it relates the properties of the mobile and stationary phases to column efficiency and HETP.

4.3.4.3 Extra column band broadening

In addition to the band broadening that occurs in the chromatography column the pipe work, instrumentation and other equipment associated with the column will contribute to the total band broadening. This contribution is negligibly small for most columns, but it can be significant for very small columns where the extra column space is similar to or exceeds the actual bed volume. The factors that contribute to extra column band broadening are dispersion in pipe work, system dead volumes, finite detector volumes and the finite response rate of electronic sensors and recorders (Kaltenbrunner, 1997). Dispersion in tubing and the finite sensing volume of detectors both introduce a symmetrical Gaussian type (σ) of band broadening while dead volumes and electronic response rates give rise to exponential (τ) band broadening. The total contribution of extra column band broadening is additive and can be expressed as:

$$\sigma_{ex}^2 = \sigma_{pipe}^2 + \sigma_{\det}^2 + \tau_{dead}^2 + \tau_{el}^2 \tag{4.7}$$

where the subscripts refer to pipe work, detector volume, dead volume and electronic response. The total contribution of extra column effects to band broadening is the sum of those that occur before and after the column itself. Usually the contribution of dispersion in pipe work is substantially greater than all the others since detector and dead volumes

are small in relation to the sample volume and electronic response times are very fast. The main consequence of extra column band broadening is that the apparent efficiency of the column is reduced.

4.3.4.4 Column packing and bed compression

The matrix in a chromatography column has to be packed down to form a tightly packed bed prior to use. If this is not done then the bed may compress during operation leaving a gap between the top of the bed and the adapter. Compression is caused by the flow of liquid through the bed, which gives rise to drag forces on the matrix beads. Since most matrices are polysaccharide based the beads are not completely rigid and deform under stress causing the bed to shrink. The degree of compression will depend upon the linear velocity of the liquid and the extent of wall support given to the matrix. Bed compression (λ) is defined by the equation:

$$\lambda = \frac{L_0 - L}{L_0} \tag{4.8}$$

where L_0 is the gravity settled bed height and L is the packed bed height. Bed compression leads to a decrease in bed porosity (ε) and reduced permeability as defined by the relationship:

$$\varepsilon = \frac{\varepsilon_0 - \lambda}{1 - \lambda} \tag{4.9}$$

where ε_0 is the gravity settled bed porosity. Since the hydrodynamic behaviour of the column is dependent on the bed porosity and packing structure it is important to have a well-packed column for good separation. However as compression increases so does the pressure drop across the column. Eventually the point is reached where any further increases in pump speed do not result in a higher linear velocity but a very rapid pressure increase occurs. This point is termed the critical velocity since it is the highest steady

state velocity possible through the column; its value depends on the matrix type and column geometry (Soriano, 1997). Generally column packing is performed at between 80-90% of this value and column loading is performed at a lower velocity, usually determined from the van Deemeter equation to give maximum HETP.

4.4 Scale-up / down of chromatography

When scaling a chromatography operation there are both chromatographic and non-chromatographic factors that require consideration. In process development it is common for the scale of operation to be increased several times in the laboratory before chromatography is scaled up to pilot-plant and full production scale. The degree of scaling in a single step is usually less than 150-fold since different phenomena control column performance at different scales. If the effects of these various scale-dependant phenomena can be predicted then greater degrees of scaling would be possible and ultimately a laboratory column with only a few millilitres of matrix could be scaled up to a full-scale column in a single step and vice versa. Currently our understanding of the phenomena that bias the performance of small-scale columns is limited and thus if better scaling is to become possible these peripheral effects need to be overcome (Lode, 1998).

4.4.1 Chromatographic factors

The chemical environment present in a chromatography column determines the rate and extent to which species are retained on the matrix. Since the interactions between different components in the feed material and the matrix are complex and difficult to predict it is essential that both the mobile and stationary phases be kept the same regardless of scale. Any buffers used in the process for washing or elution also need to be identical since they will impact upon column performance. Traditionally as well as keeping the same process materials the column bed height and linear flow rate are also

maintained. For scale-down the diameter of the column is reduced and hence the overall bed volume decreases. The volumetric flowrate is also decreased but column residence time remains the same since bed height and velocity don't change. The ratio of feed to matrix volume must also be maintained to avoid under or over-loading the column. Table 4.1 lists the major considerations when scaling down a chromatographic operation including what elements need to be maintained or decreased and the most common causes of scaling problems.

Maintain	Matrix Bed height Linear velocity / Residence time Load material (concentration, pH, conductivity, etc) Ratio of load:matrix volume Elution gradient / bed volume Sample injection time
Decrease	Column diameter Sample volume to cross sectional area Volumetric flowrate Gradient volume to cross sectional area
Factors that may change with scale	Wall support / pressure drop Sample distribution System dead volumes and hold up

Table 4.1: General guidelines for scaling down a chromatography operation.

A number of phenomena may influence column performance depending on the scale of operation and these factors must be considered when scaling up or down. Large industrial columns have much lower surface area to volume ratios than laboratory columns and as a result the matrix is not supported as much by the walls of the column.

In order to avoid excessively high pressure drops caused by the loss of wall support the flow rate in large columns may need to be reduced. This needs to be borne in mind when using scale-down models to specify operating conditions for larger columns since linear velocities used in the lab may not be feasible at large-scale.

The distribution system used in any column needs to spread the sample over the whole cross section of the bed so that maximum use is made of the matrix. Distribution in laboratory columns is usually very good but with larger columns it can be uneven leading to dead volumes and stream lining in the liquid phase. Careful design of distribution systems can avoid these problems. However for practical engineering reasons the design and materials of construction used for a laboratory column distributor will usually differ from those of a larger distributor. Therefore it is important to ensure that process materials do not bind to or react with the distributor in any way.

The volume of pipe work associated with a chromatography column is virtually impossible to scale up or down since it is difficult for a laboratory scientist or process engineer to control this value. As a general rule, the length of pipe work should be kept to a minimum to avoid extra-column band broadening and large system hold up volumes. As the scale of operation decreases the volume of pipe work increases relative to the actual bed volume so that at very small scales the effects of diffusion and back mixing become significant before the feed even reaches the column. This phenomenon makes it very difficult to interpret chromatograms from very small columns since it is difficult to separate extra column effects from those that occur on the column itself.

4.4.2 Non-chromatographic factors

Generally speaking if the chemical and physical properties of the two phases are kept the same then performance will be the same, however this is not always the case since performance is influenced by some aspects of operation and system design. The materials of construction for example can interact with the process stream. All wetted materials are required to be non-adsorptive, chemically inert, resistant to operating and cleaning conditions, non-leaching and stable at the temperatures and pressures employed. Large-scale columns along with connecting pipe work and any distribution system tend to be made from stainless steel, the column itself is usually electro-polished to give a smooth surface and reduce fouling. Small-scale columns on the other hand tend to be made of glass or plastic with plastic tubing, connections and distributor. It is important to ensure when scaling down that changing the materials for any wetted parts does not impact upon the process.

The design of the overall chromatography system also needs to be considered when scaling down. Some aspects are bound to change such as length of pipe work, number and style of valves, type and position of sensors and type of pump. The important thing is to ensure that any system changes on scale-down do not affect the performance. Good practices include keeping pipe work volume to a minimum, avoiding any stagnant regions, using simple easy to clean valves (e.g. diaphragm) and carefully positioning any monitors, meters or sensors. The type of pump used should also be low shear and able to provide pulsation free flow over a wide range. At large scale diaphragm or gear pumps are used but peristaltic pumps are more suitable for small-scale use.

As with most unit operations, considering chromatography in isolation when scaling down may not give a true representation of the process. Upstream operations can have a

significant effect upon chromatography performance and the consequences of even small process changes can be dramatic. For example if chromatography is preceded by a filtration step in which the type of filter aid used is changed the resulting filtrate may contain higher concentrations of a species that was previously removed (e.g. lipids). The higher concentrations of this species could block the column or adsorb to the matrix thus reducing its capacity. Any holding stages prior to chromatography can also have an impact; these stages are very rarely necessary in a laboratory but are common at production scale. Whilst the material is being held precipitation may occur or product may be destroyed by the action of proteolytic enzymes. This alters the composition of feed to the column and therefore affects performance.

4.4.3 Ultra scale-down of chromatography

Typical chromatography scale-down to laboratory levels still uses columns with matrix volumes as large as a few hundred millilitres. The volume of feed and buffers required for experimental work on these sorts of columns is often in the order of litres and experimental times can run into days for a complete cycle. In the early stages of product development when process synthesis is not complete the availability of product material is likely to be extremely limited. Therefore it is highly desirable to reduce the quantity of material needed to carry out meaningful chromatographic studies. Current practice makes use of very small pre-packed columns (e.g. 1 mL Hi-trap® columns from Pharmacia) to do method scouting where the most suitable matrix and binding conditions are selected. However the problems surrounding the use of very small columns has limited their applications to these simple comparative tests in which very little real process data is obtained. If these problems can be overcome then columns of this size would gain a much greater role in process development and provide valuable data about process performance and optimisation.

There is a practical limit to how far a chromatographic step can be scaled down. Theoretically the smallest diameter column would be the same as that of a single bead of matrix. However manufacturing a column this small and ensuring the internal diameter did not vary too much along its length would be extremely difficult and prohibitively expensive. The smallest readily available columns have internal diameters of around 5 mm, which is a sensible size to work with since there will be a reasonable number of beads in any given cross section to facilitate random packing similar to industrial columns. Figure 4.2 is a schematic of an industrial column along with a typical scale-down column and three possible geometries for an ultra scale-down (USD) column with a volume of 1 mL. The amount of material required to load the ultra scale-down columns is nearly 40 times less than that required for the traditional scale-down column, thus enabling many more experiments to be performed with a limited amount of material. If useful information can be derived from these experiments then a large-scale chromatography operation could be designed and to some extent optimised reliably in the lab.

The first USD column in figure 4.2 would be impractical to use since it is narrow and fragile. Additionally the large surface area to volume ratio would make it difficult to pack and the effects of wall friction would hinder achieving plug flow. The second and third USD columns would be readily available to buy and simple to use. The reduced surface area to volume ratio would facilitate packing and the influence of wall effects would be lessened. Essentially there are two extremes at which this unorthodox scale-down column may be operated. The first is to keep the residence time of the feed in the column constant by reducing the linear velocity and the second is to maintain the linear velocity so that residence time is reduced. Maintaining residence time may be very

important to enable binding and separation, but the low flow rates required to do this will encourage linear diffusion and promote band broadening. It is difficult to predict which mode of operation, constant residence time or constant velocity, would give a better representation of large-scale operation. The answer would most likely depend on the system being studied and the relative diffusion rates and binding kinetics of the species involved. It may also be possible to generate useful process data from both modes of operation and use the combined data to make process decisions.

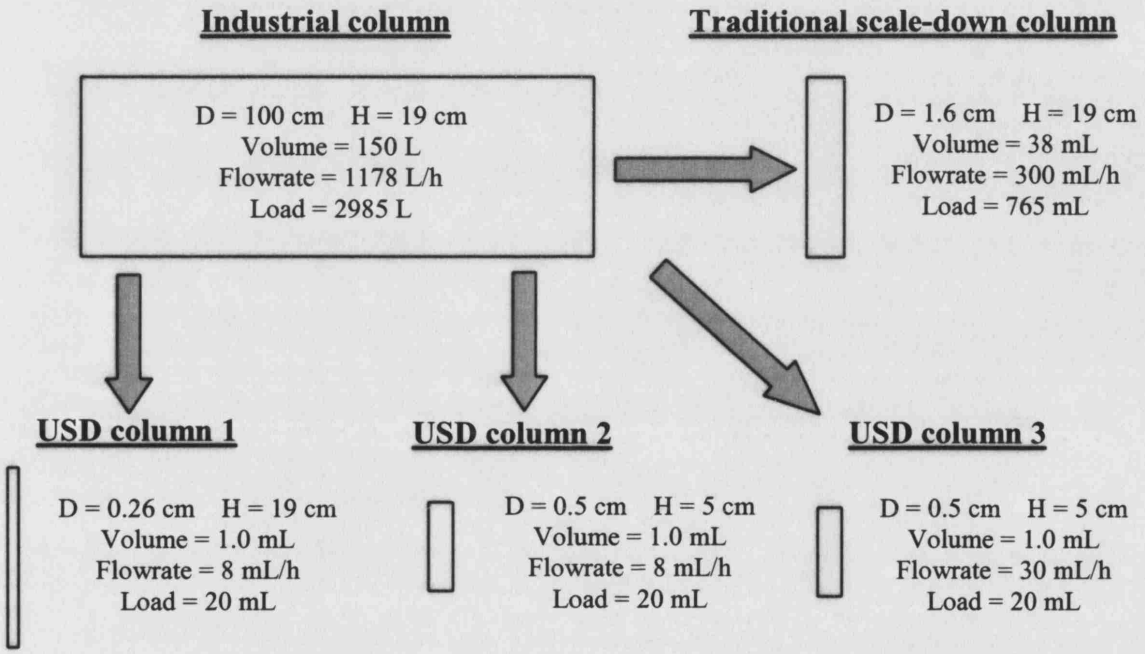

Figure 4.2: The dimensions (diameter and height), bed volume, flowrate required to give a linear velocity of 150 cm/h and volume of material required to load 20 column volumes from an industrial column and three various scale-down columns. The traditional scale-down column has the same bed height, residence time and linear velocity as the industrial column. The first ultra scale-down column follows traditional rules by maintaining bed height but has an impractically small diameter. The second and third ultra scale-down columns have a larger, more practical diameter but bed height is reduced; for this geometry only linear velocity or residence time can be maintained not both.

In this study the purification of IgG by contaminant removal chromatography is investigated using both a traditional scale-down column and a USD column. The USD

column is operated in both modes (constant residence time / velocity) and the resulting performance is compared to that of the traditional column, which is assumed to perform very similarly to an industrial column.

4.5 Materials and methods

4.5.1 Description of equipment

Chromatography was performed in four different sized columns; XK 16, XK 26, XK 50 and HR-Hiprep columns all obtained from Amersham Biosciences (Uppsala, Sweeden). The diameter of these columns is 16 mm, 26 mm, 50 mm and 5 mm respectively. All columns were packed with Q Sepharose XL media, a strong anion exchange matrix from Amersham Biosciences. Material was supplied to the columns via a positive displacement pump, either a P-500 or a P-50 depending on the required flowrate. The column outlet was connected to a single path UV monitor and control unit (Amersham Biosciences) that measured the absorbance at 280 nm of the column effluent. Outputs from the pump and UV monitor (in the range 0 – 50 mV) were recorded every 5 seconds using a 1000 series Squirrel data logger (Grant Instruments Ltd, Cambridge, UK). These outputs were later converted into absorbance units and pressure drop across the column. Column fractions were collected in plastic tubes using a FRAC-100 fraction collector (Amersham Biosciences).

4.5.2 Chromatography buffers

All buffers were prepared using chemicals from Fischer Scientific and deionised water. Buffers were stored at room temperature and were replaced after one month or when necessary. The composition of all the buffers used during chromatography is given in table 4.2.

Buffer	Composition	pH	Conductivity (mS/cm)
Pre-equilibration	5 mM Citrate, 80 mM Glycine	5.45 ± 0.05	1.3 ± 0.03
Equilibration	5 mM Citrate, 80 mM Glycine	6.85 ± 0.05	1.3 ± 0.03
Elution	5 mM Citrate, 80 mM Glycine, 1 M Sodium Chloride	6.85 ± 0.05	-
Cleaning	1 M Sodium Hydroxide	-	-
Storage	20 % Ethanol	-	-

Table 4.2: Buffers used during QXL chromatography to remove contaminants from filtered F. A+1 suspension. The pH and conductivity were not specified for the cleaning and storage buffers.

4.5.3 Preparation of load material

The protein solution used in all chromatography runs was derived from filtered F. A+1 paste (see section 2.2.2 and figure 2.1 for more information). Once prepared load material was kept refrigerated at 4°C and used within 3 days of preparation.

4.5.3.1 Fraction A+1 paste resuspension

Fraction A+1 paste from the Cohn fractionation process was obtained from BPL's production department. The paste was divided into aliquots of approximately 100g and frozen at -40°C. One day prior to resuspension, 100 g of paste was removed from the freezer then placed in a fridge at 0°C and left overnight (12-24 hours) to defrost.

Once defrosted approximately 90g of F.A+1 paste was weighed out into a stainless steel pot and mixed with 3 parts of chilled deionised water (ratio paste: water is 1:3). The mixture was then homogenised using a Silverson mixer (Silverson Machines LTD, Chesham, UK) with a 44mm disintegrating head at 50 s^{-1} for 5 minutes. The mixture was

checked for lumps and homogenised for a further 1 minute if necessary. The homogenisation process was performed in a 4°C cold room to minimise protein denaturation caused by temperature increases.

Once the paste was completely suspended the mixture was poured into a jacketed glass vessel cooled to 3°C. A further 7 parts of chilled deionised water were added to increase the dilution to 1:10 and the suspension was mixed using an overhead mixer (Janke and Kunkle) at 5.8 s^{-1} for 1 hour. After this time 10 mL of the suspension was removed, diluted 1/5 with 0.16 M saline and titrated to pH 4.85 ±0.05 using 50mM citrate buffer pH 4.0. The amount of buffer required to adjust the remaining suspension was calculated and added slowly to the glass vessel. The suspension was then left mixing overnight to allow the IgG in the paste to dissolve.

4.5.3.2 Filtration

The suspension was filtered in a stainless steel laboratory Nutsche filter (Dr Muller AG, Switzerland) of 127 mm internal diameter (filtration area 0.013 m^2), with a glass fibre prefilter support (Sartorius, Gottingen, Germany). One litre of the suspension was removed from the stirred vessel and combined with 30g of *Celpure C100* (World Minerals Inc, Santa Barbara, CA, USA) to give a body-feed filter aid concentration of 3% w/v. A further 6.3g of *Celpure C100* was mixed with 1 litre of 50mM citrate buffer to form a pre-coat mixture (0.5 kg/m^2). The pre-coat mixture was poured into the filter housing and a pressure of 1.0 bar was applied. The compressed air supplying the pressure was turned off once the flow of liquid had stopped. With the pre-coat in place the protein suspension was mixed briefly to suspend the filter aid and then poured into the filter housing carefully so as to cause minimal disruption to the pre-coat. A pressure of 1.0 bar was applied and filtrate was collected in a beaker placed on a balance. The

balance was connected to a computer which logged the filtrate mass every 5 seconds. Once air started to come through the filter outlet pipe the pressure was turned off and the filtration ended.

4.5.3.3 Filtrate treatment prior to chromatography

Before being loading onto a column the filtrate / supernatant was diluted with an equal volume of deionised water. Glycine was added to a final concentration of 80 mM to stabilise the IgG. The pH was then adjusted to 6.85 ±0.05 using a 5M sodium hydroxide solution and the conductivity was adjusted to 1.35 ±0.03 mS/cm using 100mM citrate buffer (pH 6.85). The adjusted filtrate was then left to sit for a few hours in a fridge at 4°C after which time it was filtered through a 0.2-micron syringe filter (Whatman International Ltd, Maidstone, UK).

4.5.4 Operation of chromatography columns

4.5.4.1 Column packing

The media used was Q Sepharose XL (Amersham Biosciences), a strong anion exchanger. The matrix was stored in 20% ethanol. Before packing the matrix was suspended in the ethanol by vigorously shaking the bottle and was then poured into a measuring cylinder. The ethanol was siphoned off and the resin was resuspended in approximately double this volume of loading buffer before being poured into the columns. Matrix was allowed to settle under gravity to the appropriate bed height (either 19 or 5 cm). The top column adapter was then lowered into the buffer carefully to avoid air entrapment. Equilibration buffer was then passed through the column at a predetermined velocity to pack the media tightly. Once the bed height had stabilised at this flow rate the adapter was lowered to the top of the bed and pushed approximately

2mm into the bed to prevent the formation of a gap between the top of the bed and the adapter.

4.5.4.2 Column loading

Once packed and equilibrated the columns were loaded with protein suspension. The load pump (P-50), stopwatch, data logger and fraction collector were started simultaneously. Once the desired volume of protein suspension was loaded the feed inlet was switched to equilibration buffer and the column was washed to remove entrapped protein. After washing the inlet was switched to elution buffer and the bound proteins were eluted. Once the absorbance of the column effluent fell below 0.2 mV the process was stopped and fractions were assayed for IgG and total protein.

4.5.5 Analyses

4.5.5.1 Filtration clarification

The turbidities of the F.A+1 resuspension and filtrate were estimated by measuring the absorbance at 670 nm using a Lambda 20 digital spectrophotometer (Perkin-Elmer Instruments). A sample was filtered through a 0.2 μm syringe filter and the turbidity of this sample was used as a reference point for 100% clarification.

4.5.5.2 Filter cake dry weight

The moisture content of the solids phase was measured with a HG53 Halogen Moisture Analyser (Mettler Toledo, Greifensee, Switzerland) set to 175°C and a sensitivity level of 3 until a constant mass was achieved.

4.5.5.3 IgG measurement

The amount of IgG present in samples was measured using a Turbitime system (Behring, Marburg, Germany). This system is based upon the principles of kinetic turbidimetry and measures both the maximum reaction velocity and the time taken to reach this velocity when sample is mixed with precalibrated antisera. A 20 µL sample is pipetted into a plastic cuvette and placed into the machine. Next 500 µL of IgG antisera is added to the cuvette, a small magnetic stirrer in the bottom of the cuvette mixes the liquid automatically, then the absorbance is determined and the concentration of IgG is calculated and printed out.

4.5.5.4 Total protein measurement

The total amount of protein in samples (including IgG) was measured using a bicinchoninic acid (BCA) assay kit (Pierce, Rockford, USA). This method is based upon the reduction of Cu^{+2} to Cu^{+1} by protein in an alkaline environment (the biuret reaction), which causes a colour change from green to purple. This purple reaction product exhibits a strong absorbance at 562 nm that is linear with increasing protein concentrations. A standard curve covering the range 0.05 – 1.15 mg/mL was made using purified IgG. Each unknown sample was diluted assayed and the protein concentration was determined from the standard curve. The assay was performed according to the manufacturers instructions.

4.6 Results and discussion

4.6.1 Fraction A+1 resuspension and filtration

The process of resuspending and filtering the F.A+1 paste proved to be quite reproducible in terms of the material that was produced for chromatography. The initial pH of the suspension was always within a narrow range (5.40 – 5.60) and the quantity of

buffer required to adjust the pH for IgG extraction was equally invariant. The average flux rate during filtration was 7.8 L/m^{2}h and the clarification achieved was extremely high with an average of 98.8 ± 0.9 %. The recovery of IgG over this step was not investigated but for 1 litre of filtered suspension approximately 930 mL of filtrate was recovered and the concentration of IgG in the filtrate was in the range 6.85 – 7.21 mg/mL (average 6.98 mg/mL). The average mass of filter cake was 118 ± 5 g and the dry weight was on average 37 ± 2%. If it is assumed that the wet fraction of the cake has the same composition as the filtrate then the loss of IgG is calculated to be 0.5 g. This is a very approximate estimate but it should represent the worst-case scenario. Recovering this IgG should be possible by washing the cake with a small volume (3-4 times the cake volume) of buffer.

When preparing the filtrate for column loading it was found that adjusting the pH to 6.85 resulted in the precipitation of some material. The fibrous appearance of the precipitate suggested it was fibrinogen, a major contaminant in the process. In order to prevent excessive column fouling the load material was passed through a 0.2 μm filter directly before loading. This removed the precipitate. However after storing the load material for several days in the fridge it reappeared. To minimise the impact of this precipitation the filtrate was always adjusted and filtered on the same day as chromatography was performed.

4.6.2 Column packing

Before any actual process material was run through the columns it was necessary to determine appropriate packing velocities. Packing at the same linear velocity in both columns resulted in much higher levels of compression in the XK16 column due to the reduced wall support. This approach to packing was unsuitable since if the matrix in the

USD column is to be representative of that in the traditional scale-down column then it must be packed such that the bed compression and hence porosity and permeability are identical. Models exist to predict bed compression in different columns using forces balances, but application of these models is difficult since they require intricate knowledge of the frictional properties of the column wall and matrix particles (Stickle, 2001). Therefore for this study experimental techniques were used to determine the dependence of pressure drop and compressibility on linear velocity. Figure 4.3 shows plots of pressure drop and bed height verses linear velocity for the XK16 and USD columns. Comparing the two plots it becomes evident that for any given linear velocity the pressure drop in the USD column is lower due to the shorter bed height and greater wall support. It is also apparent that the critical linear velocity is not reached for either column; this is due to limitations with the pump used. Column packing normally occurs at a velocity that is 80-90 % of the critical velocity, however lower packing velocities can be used so long as they are higher than the operating velocity. Since the important issue for this study is that both columns are packed to the same degree the packing velocities were chosen based upon the amount of bed compression they produced. The compression in each column was calculated at different linear flow rates using equation 4.7 and the results are plotted in figure 4.4. From this graph it was possible to select packing velocities for each column that would result in the same level of compression.

Ideally the packing in a chromatography column should be uniform in nay cross sectional area. However it has been demonstrated that in reality this is not the case since the packing density varies in the radial direction as well as the axial. The primary reason for this is friction between the column walls and the beads, which result in higher compression stresses close to the walls. Guichon *et al* (1997) claimed that the mobile phase velocity was up to 8% higher in the column centre and that the HETP near the

walls might be up to 150% higher than the centre. These cross column variations are likely to affect column performance in some way, although exactly how remains to be determined. However, it is almost certain that these so called "wall effects" will be more pronounced in smaller columns due to the increased ratio of wall area to bed volume.

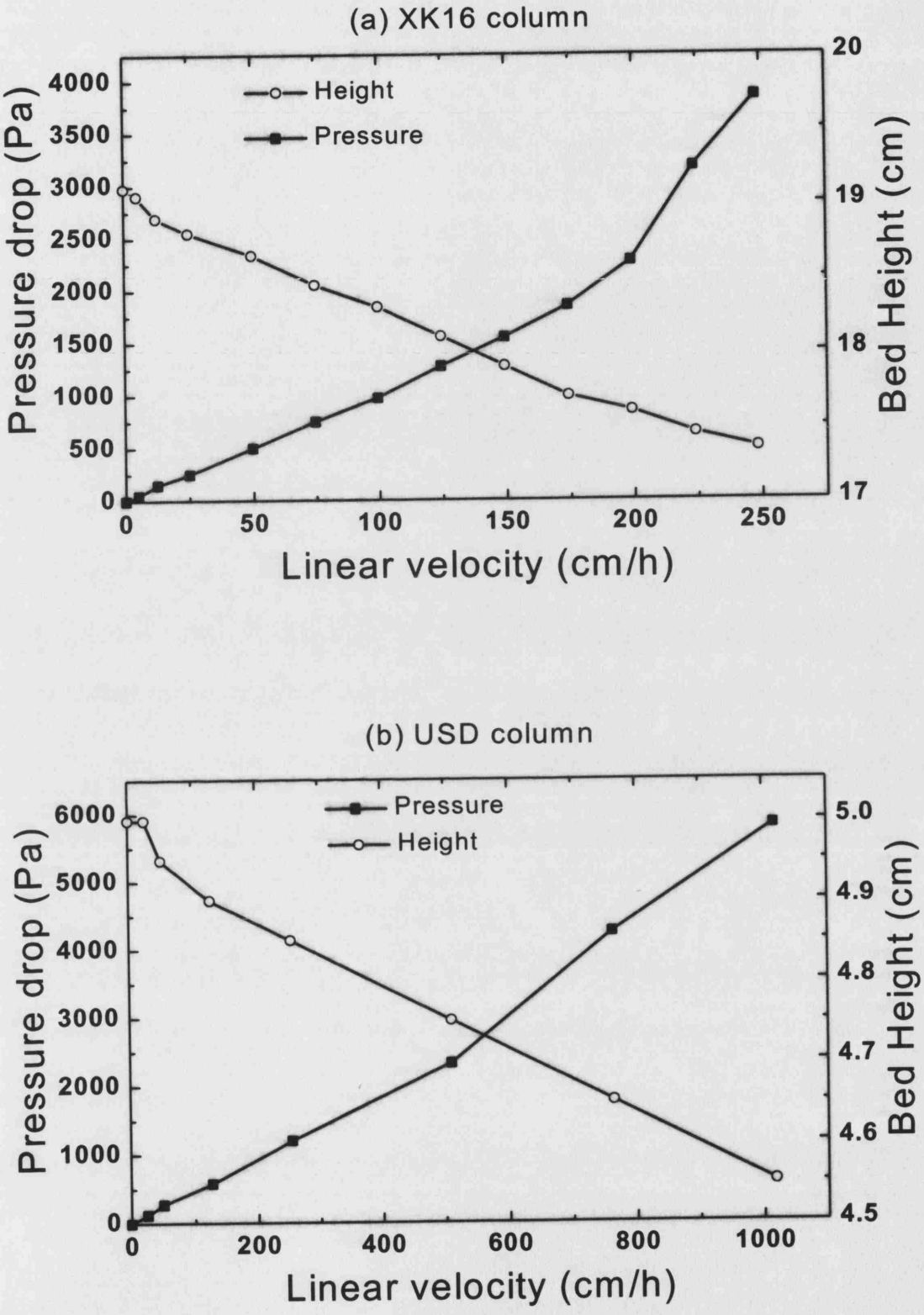

Figure 4.3: Pressure-flow curves from the XK16 column and ultra scale-down column. The decrease in bed height with increasing linear flow rate is also plotted. Lines simply connect the data points and are not lines of best fit.

The packing velocity selected for the XK16 column was 170 cm/h, which translates to a flowrate of 340 mL/h. For the USD column a packing velocity of 750 cm/h or 150 mL/h was selected to give the same bed compression. At these flowrates the bed compression is 0.07 in both columns (from equation 4.8). For all further experiments the columns were packed using equilibration buffer at these flowrates. The gravity settled and packed bed heights were measured to ensure that the same degree of compression was achieved each time.

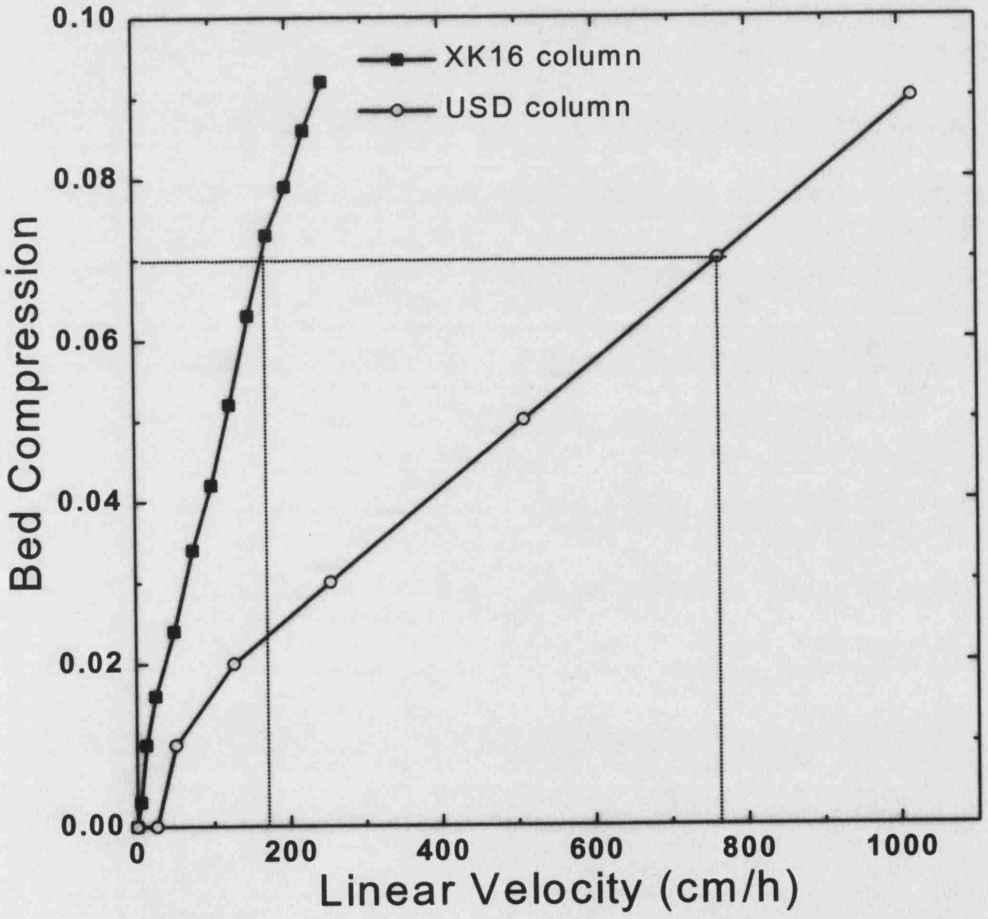

Figure 4.4: Bed compression in the XK 16 and USD columns at various flowrates from 0-500 mL/h. Packing flowrates for each of the columns were selected to give the same level of compression in both beds. The packing flowrates also needed to be higher than operating flowrates to avoid the formation of a gap between the top of the bed and the adapter. The packing flowrates chosen were 150 mL/h for the USD column and 340 mL/h for the XK16 column, at these flowrates the bed compression is 0.07.

4.6.3 HETP calculations

The efficiency of packing in each column was assessed by measuring the height equivalent to a theoretical plate (HETP). The length of tubing between the sample loop and column inlet was deliberately very small in order to minimise the impact of extra column band broadening. The average HETP value for the XK16 column was found to be 0.0295 ± 0.001 cm and for the USD column it was 0.0304 ± 0.001 cm (using equation 4.4). The similarity of these values indicates that the efficiency of the columns should be roughly the same. However, due to the different bed heights the number of theoretical plates in the columns will differ significantly. It was calculated that the XK16 column had an average of 644 theoretical plates while the USD column contained an average of only 165 theoretical plates. The lower number of plates in the USD column would have a significant effect on the separation power of the column, particularly for displacement or frontal chromatography. However since the operation under study here is essentially batch adsorption and no separation of the bound phase is required (because it is all waste) the performance implications of the low plate number should be minimal. The overall binding capacity of the matrix should still be the same and therefore it should be possible to mimic the performance of the XK16 column using the USD column.

4.6.4 Column performance

Chromatograms for the XK16 column and USD column under both modes of operation are shown in figure 4.5. The x-axis is expressed in column volumes to enable simple comparison between the different sized columns. The y-axis is normalised by dividing the absorbance of the column effluent by the absorbance of the feed, this accounts for any feed variations between experiments. Studying the chromatograms is a quick and simple way to assess column performance. It is evident immediately that the performance of the different columns is not exactly the same.

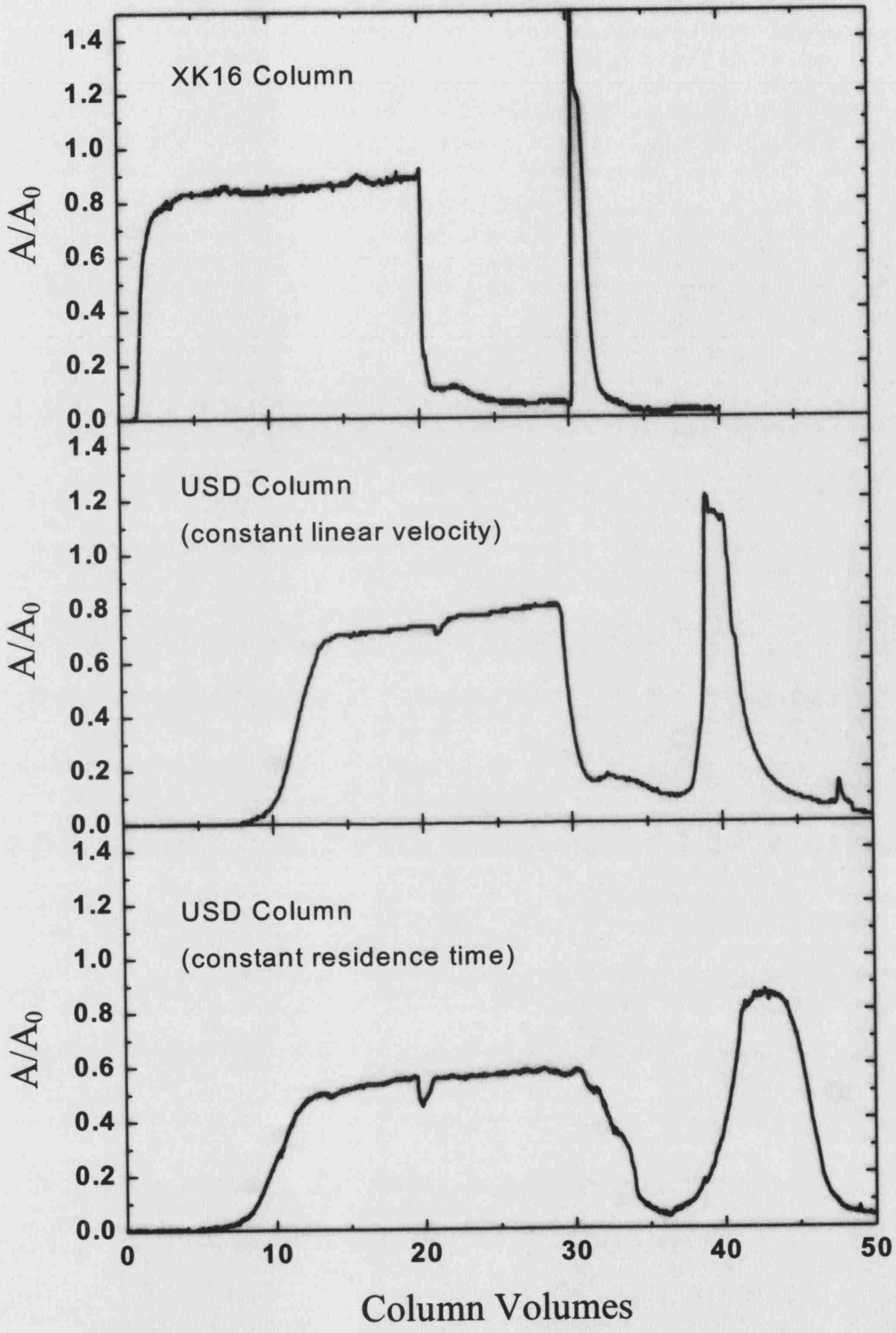

Figure 4.5: Chromatograms for the XK16 column and the USD column in constant residence time and constant linear velocity mode (constant refers to kept constant during scale-down, hence is the same as XK16 column). The columns were loaded with 20 column volumes of feed material, then washed with 10 column volumes of equilibration buffer and then eluted by a step change in salt concentration.

The chromatograms immediately highlight some ways in which the performance of the USD column is different from that of the larger XK16 column. Firstly there is a significant hold up volume associated with the USD column, which when combined with the low operating flowrates means that a significant amount of time passes before the load material reaches the column inlet. Secondly the chromatograms for the USD column show evidence of band broadening, most likely caused by molecular diffusion, this can be seen by the more gradual transition between stages (flowthrough, wash and eluate). Finally it appears that more material is bound on the USD column since the size of the elution peak relative to the size of the flowthrough is larger than that from the XK16 column. This extra "bound" material may indeed be bound to the matrix or could be adsorbing to some other component of the system such as the column walls. In order to further assess the performance of the columns the yield and purity of IgG in the flowthrough were calculated and are shown in table 4.3.

		1	2	3	4	5	Average	STD
XK16 column	**Yield (%)**	95.02	95.50	95.60	95.91	96.21	95.65	0.45
	Purification factor	NA	1.12	1.12	1.18	1.20	1.15	0.04
USD (linear velocity)	**Yield (%)**	92.88	92.84	93.23	92.27	91.91	92.63	0.48
	Purification factor	NA	1.21	1.15	1.20	1.09	1.16	0.01
USD (residence time)	**Yield (%)**	89.90	90.31	87.13	84.47	87.88	87.94	2.36
	Purification factor	NA	1.18	1.09	1.13	1.10	1.13	0.04

NA = not applicable since total protein not measured

Table 4.3: The yield and purity of IgG recovered from the various chromatography columns used. The yield was calculated by combining the IgG in the flowthrough and wash stages. Purification factor was calculated by dividing the final purity by the initial purity. The standard deviation of these values over the 5 runs is also shown.

It can be seen from table 4.3 that the IgG yield in the USD column is worse than that of the XK16 column and that operating the USD column in constant residence time mode produces a lower yield than operating in constant linear velocity mode. For this particular operation a lower yield indicates that product is binding to the column when it should be passing through unaffected. The different amounts of binding between different operating modes suggest that it is not part of the column that is binding the protein. Since the physio-chemical environmental conditions such as pH, conductivity and protein concentration are identical in each case the apparent increased capacity of the USD column must be somehow related to the hydrodynamic conditions within the column. If the USD column is operated in constant residence time mode then the time the sample spends within the column is the same as the XK16 column, but the sample contact time with any individual section of matrix is actually higher due to the lower linear velocity. The effect of this could be a greater degree of non-specific binding which would decrease the yield. However, if this is the case one would not expect to see a lower yield in the USD column when operated in constant linear velocity mode. Since a lower yield is observed the cause must be attributed at least partially to something else. It may be somehow related to wall effects, which have been reported to dominate the performance of small-scale columns (Shalliker, 2000).

The different modes of operating the USD column did have an impact upon chromatographic performance. However it was decided at this point that operating the column at the same linear velocity as the XK16 column was more useful as a scale-down tool than operating at constant residence time. The main reason for this was that the performance when operating at constant residence time had a very low reproducibility, as could be seen by the higher standard deviations for this column in table 4.3. Additionally the chromatograms produced varied substantially in shape and defining where to make

the cut between regions was very difficult since they often overlapped. The cause of this is thought to be the very low flowrate (0.13 mL/min), which results in diffusion processes becoming increasingly dominant. For these reasons all further work focussed on operating the USD column in constant linear velocity mode. This mode has a further advantage over constant residence time mode in that the experimental time required is much less. This is significant since the whole aim of scale-down is to gain as much information as possible in as short a time and using as little material.

4.6.5 Correcting the chromatogram for small-scale effects

The results from the USD column show that there are some factors that have a significant effect upon the performance of very small columns. Avoiding these effects is difficult since they are inherent to small-scale chromatography systems; therefore if useful information is to be gained from experiments at this scale there must be some way to quantify the impact of these effects and compensate for them.

Studying the results from the USD column there would appear to be three aspects of small-scale operation that have an impact on column performance. The first of these is the system hold-up volume, which may be several times larger than the volume of the bed itself. This encourages back mixing and diffusion and would also affect fraction collection since processing times would be different between scales. The second factor is extra column band broadening as a result of molecular diffusion in the tubing before and after the column. This happens at large scale too, but the relatively high hold-up volumes and low flow rates exaggerate the effects at small-scale (Kaltenbrunner, 1997). The final aspect is the apparent extra binding capacity of the USD column. It is assumed that this phenomenon is somehow related to wall effects, which can be highly influential in smaller columns (Shalliker, 2000). The next section discusses how these factors can be

compensated for and shows that it is possible to quantify these effects and adjust the performance of the small column accordingly.

4.6.5.1 Hold up volume

Measuring the hold-up volume of the system is a relatively simple task and can be done by simply pushing the column adapters together (so that bed volume is zero) and passing a pulse of some adsorbing material (e.g. acetone) through the system. The absorbance at the column outlet is monitored and then from the system residence time and the flowrate the hold-up volume can be calculated. It is important to include any tubing between the outlet and the detector in the calculation since band broadening occurs post-column as well as pre-column and this volume contributes to the delayed response. The total system hold-up for the USD column was found to be 9.1 mL, which is equivalent to 9.1 column volumes of material. The chromatogram for the USD column is adjusted by setting 9.1 on the x-axis as zero, so everything is moved 9.1 units left along the x-axis. This effect of this is shown in figure 4.6, graph 1.

4.6.5.2 Extra column band broadening

Extra column band broadening effects are attributable to dispersion in tubing, dead volumes, finite detector volumes and the dynamic behaviour of transducers and electronics (Kaltenbrunner et al, 1997). For a chromatography system it can be assumed that the contribution to band broadening of detector volumes and electronics is negligibly small. Any dead volumes in this system are also likely to be very small compared to the hold volume since this is several times the actual column volume; therefore the major contribution to band broadening comes from dispersion in tubing. Taylor (1953) demonstrated that diffusion in tubing results in a symmetrical gaussian type broadening of a sample. So the first stage in compensating for this broadening was to quantify it by

fitting gaussian peaks to the flowthrough and eluate sections of the USD chromatogram (see figure 4.6 graph 2). Next it was necessary to estimate how much band broadening occurred on the column and how much could be defined as extra column. This is not easy to measure but Kaltenbrunner (1997) calculated that for a 1 mL column the extra column band broadening would account for approximately 70% of the total. The exact ratio between extra and intra-column band broadening is likely to depend on the ratio of total column volume to extra-column volume. Since the extra-column volume is the same for both columns studied (XK16 and USD) the relative amounts of extra-column broadening on each column should be very similar those found in Kaltenbrunner's study despite the flowrate and extra-column volumes being different. Hence if the variance (σ) of the fitted gaussian curves is reduced by 70% the resulting curves should represent how the chromatogram would look if only intra-column band broadening occurs. The chromatogram can then be modified by incorporating the new gaussian curves in the appropriate place. However, simply replacing the experimental eluate profile with the new gaussian curve will result in a change of area. This cannot happen since the physical significance of the area under the curve is the mass of protein and the effects of band broadening do not alter this, they simply spread it out. Increasing the height of the eluate peak enables the area under the curve to be maintained so that no protein is lost. It is also important that the area of the flowthrough be maintained by careful positioning of the new curves at either end. The final result of these adjustments is shown in figure 4.6, graph 3.

4.6.5.3 Column wall effects

In order to assess the impact of wall effects on column performance further experiments were done using the XK16 column, but the bed height was reduced and hence the ratio of wall area to bed volume increased. With the bed height of the XK16 column reduced to 5

cm the yield of IgG was also reduced slightly. It was assumed that increased surface area reduced the yield due to the formation of an almost stagnant region near the column walls, which greatly slowed down the passage through the column of any material within it. Material binding to the column walls was ruled out as a possibility since both columns were made of the same material and mass balances proved that all the protein was removed from the columns. In order to try and quantify the effects of this wall region a set of simultaneous equations were derived based upon assay data from the USD column and the reduced bed height XK16 column. The loss of IgG on each column was considered to be a result of two different mechanisms. The first was non-specific binding of IgG to the matrix or other bound proteins and the second was IgG becoming stuck in the wall region and taking a very long time to pass through the column. The relevant simultaneous equations are given below:

$$V_{USD}I_B + A_{USD}I_W = \text{IgG loss in USD column} \quad (4.10)$$

$$V_{XK}I_B + A_{XK}I_W = \text{IgG loss in XK16 column} \quad (4.11)$$

where V is the bed volume, I_B is the concentration of bound IgG, A is the internal surface area of the column wall and I_W is the concentration of IgG in the wall region. The eluate fraction from each column was assayed to determine the IgG loss. Mass balances showed that all the IgG loaded onto the columns was recovered indicating that irreversible binding does not occur. The mass balances for IgG and total protein on both of the columns used are shown in Appendix B. Substituting the appropriate values into equations 4.9 and 4.10 gives:

$$\text{USD column: } (0.98 \times I_B) + (7.86 \times I_W) = 4.73 \quad (4.12)$$

$$\text{XK16 column: } (10.01 \times I_B) + (25.14 \times I_W) = 27.34 \quad (4.13)$$

By solving the above equations it was found that $I_B = 1.75$ mg/mL and $I_W = 0.38$ mg/mL.

Using the values of I_B and I_W the IgG loss in any column can be estimated from the column dimensions (assuming other factors such as load volume and protein concentration do not change). To test this the chromatography step was repeated in a number of columns with different aspect ratios and the yield of IgG was calculated for each one. The predicted IgG loss was then converted to a yield based on the amount loaded onto the column. Table 4.4 lists the predicted and experimental yields for various columns along with the column dimensions. When calculating the predicted yields, to simplify things the wall region is assumed to have a thickness of unity. Since the thickness of the wall region should be independent of scale this assumption is valid, as the volume of the stagnant region will depend solely on the internal surface area of the column wall.

Dimensions H x D (cm)	Bed volume (mL)	Wall area (cm^2)	IgG loaded (mg)	IgG loss (mg)	Predicted yield (%)	Experimental yield (%)
5 x 0.5	1.0	7.86	63.6	4.70	92.60	92.57
5 x 1.6	10.1	25.14	575.5	27.14	95.28	95.15
10 x 1.6	20.1	50.27	1280.0	54.29	95.75	95.96
19 x 1.6	38.2	95.52	2563.1	103.16	95.97	95.83
5 x 2.6	26.5	40.85	1689.4	61.98	96.33	96.13
5 x 5	98.2	78.55	5995.1	201.68	96.63	96.74

Table 4.4: Predicted and experimental yields for a variety of laboratory scale columns. The first two columns were used to calculate the concentration of bound IgG (I_B) and the concentration of IgG stuck in the wall region (I_W). These values were then use to calculate yields in the other columns based upon the volume of matrix and the surface area of the column walls.

The chromatograms for all the columns except for the 0.5 cm diameter column were all virtually identical. This indicates that above a certain column diameter small scale effects including wall effects become insignificant. This would make scaling between these columns a relatively simple task since it would just be a case of maintaining important

process parameters and no corrections would need to be made to the chromatograms. However, there is some evidence to suggest that in very large diameter columns wall effects again become significant since without the support of the walls the matrix can collapse leading to excessive pressure drops.

A similar set of equations to the ones above can be derived for the total amount of protein removed by the columns assuming that the same two mechanisms of protein removal apply. These simultaneous equations are based upon the total protein assays and are expressed as:

$V_{USD}P_B + A_{USD}P_W$ = Total protein loss in USD column (4.14)

$V_{XK}P_B + A_{XK}P_W$ = Total protein loss in XK16 column (4.15)

Where P_B is the concentration of all the bound protein and P_W is the concentration of all the protein in the wall region. From the mass balances for each experiment it is possible to calculate the amount of protein removed by the columns (see Appendix B for mass balances). The average amount of protein removed by the USD column was 25.58 mg and the average amount removed by the 5 cm bed height XK16 column was 160.82 mg. Substituting these values and those for column properties into the equations gives:

USD column: $(0.98 \times P_B) + (7.86 \times P_W) = 25.58$ (4.16)

XK16 column: $(10.01 \times P_B) + (25.14 \times P_W) = 160.82$ (4.17)

Solving these equations shows that $P_B = 11.23$ mg/mL and $P_W = 1.78$ mg/mL

The values of P_B and P_W enable the prediction of the amount of protein removed in any sized column. Using these values the predicted protein removal for the traditional laboratory column (XK16, 19 cm bed height) was estimated to be 599.7 mg. This compares very well with the experimental average over four runs, which is 626.6 mg. By comparing the relative amounts of protein removed on the USD column and the

predicted amount on the XK16 column it is possible to derive a correction factor to compensate for the increased wall effects in the smaller column. This correction factor can then be applied to the elution peak of the USD column to adjust its size. The correction factor (C) can be determined by the equation:

$$\frac{\left[\frac{protein\ \ removed}{protein\ \ loaded}\right]_{USD}}{\left[\frac{protein\ \ removed}{protein\ \ loaded}\right]_{XK}} = C \qquad (4.18)$$

In order to adjust the USD chromatogram all the data points used to plot the elution peak were divided by C, which was found to be 1.63. This reduces the area of the elution peak proportionately. The difference in area between the old and new peaks represents the extra mass of protein that was present in the eluate from the USD column. If this area is discarded then it will be impossible to complete a mass balance across the column, so the extra material is added to the flowthrough and the area is increased in accordance. This represents the final adjustment to the USD chromatogram and is shown in figure 4.6 graph 4. This correction factor should be applicable to any sized column so long as the loading conditions are the same, its value however will vary depending on the relative wall area and volume of the column.

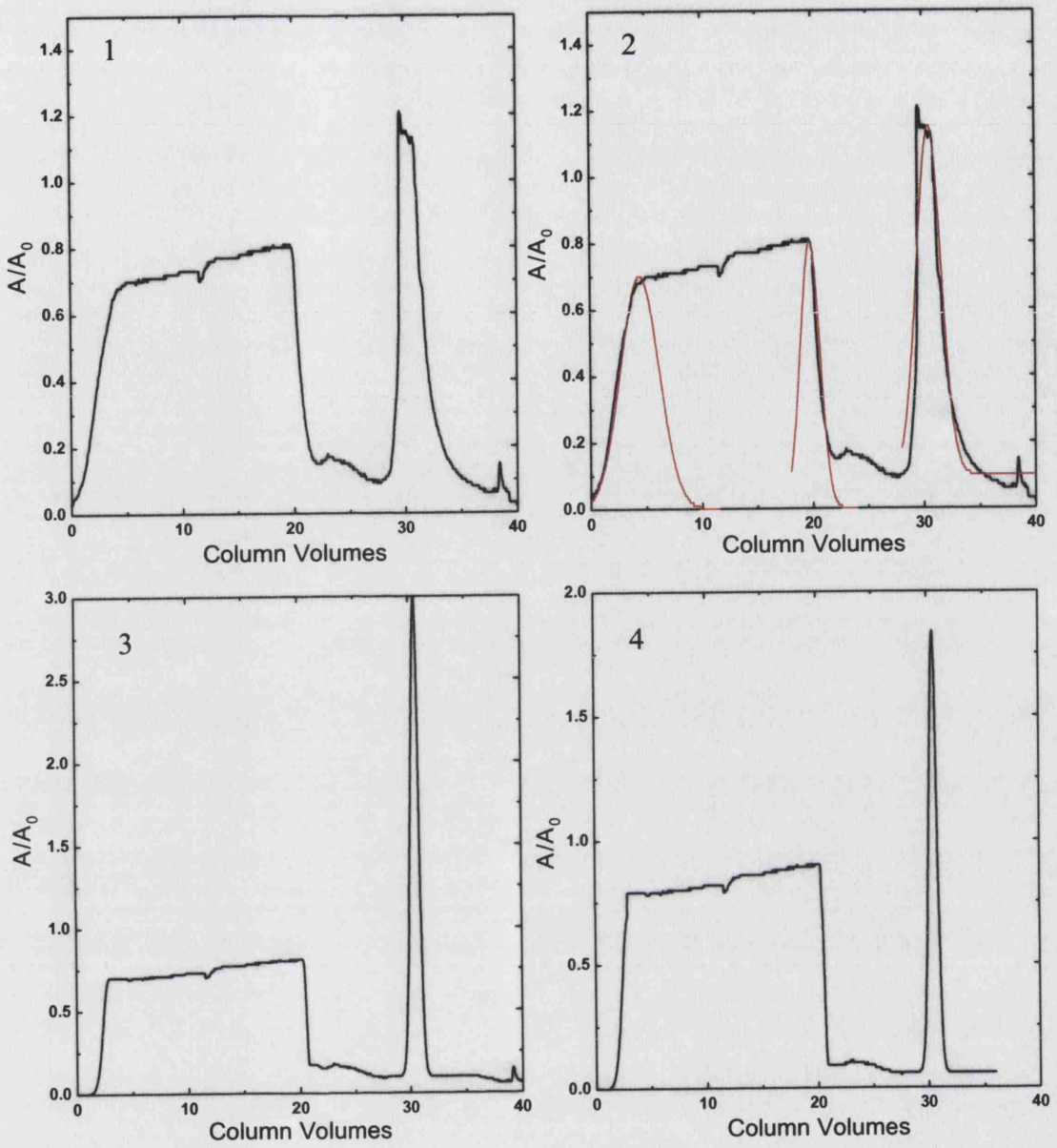

Figure 4.6: The stages in adjusting the USD chromatogram to account for small-scale column effects. In graph 1 the x-axis is adjusted to account for the system hold-up volume. In graph 2 gaussian curves are fitted to either end of the flowthrough section and to the elution peak. In graph 3 the fitted gaussian curves are adjusted to compensate for extra column band broadening and replace the corresponding sections in the chromatogram; care is taken to maintain the area of both the flowthrough and eluate sections (note change in y-axis scale). In graph 4 the area of the wash and eluate sections are adjusted to account for IgG loss as a result of wall effects. The degree of adjustment required is calculated from experimental assay data and any protein removed from the wash and eluate sections is evenly distributed across the flowthrough section.

4.6.5.4 Comparison of corrected and uncorrected USD chromatograms

In figure 4.7 the chromatogram from the traditional laboratory column is compared to the chromatogram from the USD column before and after adjustment. It is evident that after adjustment the USD chromatogram resembles that of the XK16 column much more closely. More importantly the areas under the various sections of the curve (flowthrough, wash and eluate) are very similar indicating that a similar yield and purification would be achieved. Some discrepancies still exist most notably at the beginning of the flowthrough and in the shape of the eluate peak. One important reason for this is that the fitted gaussian curves were not perfect matches for the experimental data. The eluate peak in this process is made up of many different proteins which will eluate at slightly different times giving it an irregular shape, the peak also exhibits substantial tailing. Since finding an equation to describe fully the shape of this peak is not practical the closest fitting gaussian curve was used. Careful attention is paid to maintaining the area of this peak but it is impossible to convert the gaussian curve back to the same shape as the original peak after the size has been adjusted like in figure 4.6. In a product capture process the eluate peak is likely to have a more gaussian shape and this problem shouldn't arise.

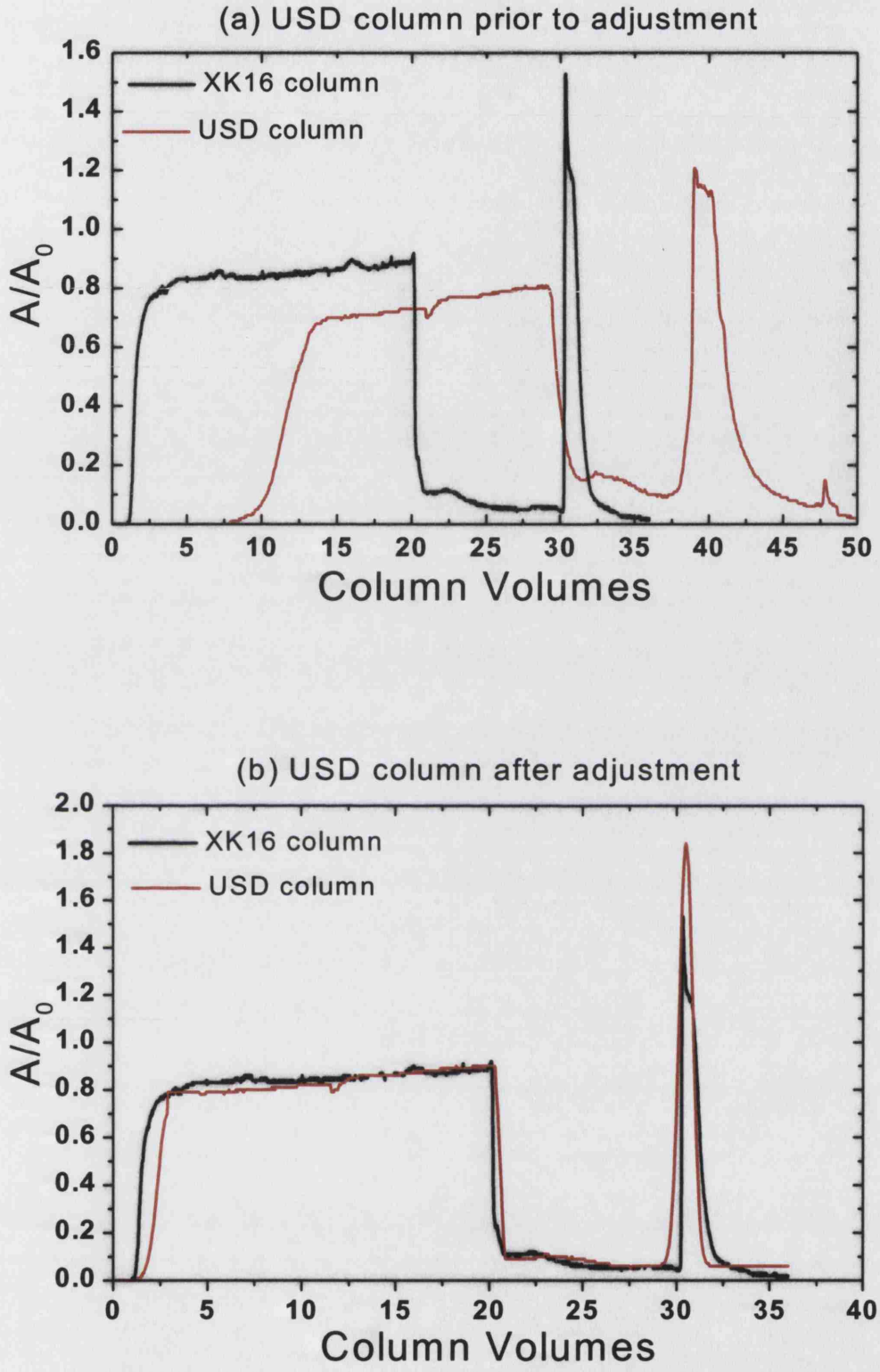

Figure 4.7: A comparison between the chromatogram generated by the XK16 column and that of the USD column operated in constant linear velocity mode; (a) shows the experimentally determined USD chromatogram and (b) shows the USD chromatogram after it has been modified to correct for small-scale effects such as hold-up volume, extra column band broadening and wall effects.

4.7 Use of the USD column as a predictive tool

It has been shown that it is possible to correct the chromatogram produced from the USD column to compensate for the effects of small scale phenomena such as hold-up volumes, extra column band broadening and wall effects. In order to calculate the correction factor required to account for wall effects it is necessary to perform experiments in another column with different dimensions. However, this column can also be small and thus time and resources are still conserved.

Analyses of the modified USD chromatogram can yield useful process information about the chromatography step. Quantifying protein and product removal in terms of matrix capacity and wall effects allows predictions to be made about the yield and purity achievable in a larger column. Integrating the chromatogram ensures that material is not lost during the correction process and enables mass balances to be modified to reflect the corrections. Although the chromatography process studied here is somewhat unusual in that product binding is not the objective and the bound material is actually waste, the theory employed should still apply to other chromatographic processes. Standard product capture chromatography would be subject to the same small-scale effects if the USD column was used and the same correction procedures would be applicable. Some parts of the procedure may even be simplified since the elution peak for a singe protein is more likely to be gaussian in shape and thus will be easier to fit an equation to. However, peak separation, particularly with gradient elution is likely to be more difficult to reproduce using the USD column due to the lower plate number and dominant wall effects. In this situation it may be necessary to reduce the operating velocity of the USD column to compensate for these factors. Though as observed in this study very low flowrates result in a lot of molecular diffusion, which can make interpretation of results difficult.

The USD chromatography column and methodologies described here would be a useful tool during process development. The usual scale-down benefits of reduced time and material usage are achieved enabling an early start to process development and the rapid procurement of vast amounts of process data. The ability to gain quality process data quickly is particularly useful for chromatographic operations since the complexity of the interactions involved restricts the use of modelling and other predictive tools, making experimentation essential. Furthermore, the large number of process variables makes optimisation difficult. Using traditional scale-down techniques to develop a chromatography step often results in some variables being arbitrarily set or ignored since there is not enough time to study their full impact on process performance. For example alternative loading buffers are commonly overlooked and the final loading conditions selected may not be ideal. The USD column would enable a wide range of process variables and combinations of variables to be studied resulting in a well understood and thoroughly developed process.

Chromatography operations are very sensitive to changes in feed material, thus it is worth investigating a number of options for the stages preceding chromatography. Subtle differences between solid-liquid separation steps can lead to drastic changes in column performance, for example lipids which are removed in a batch operated laboratory centrifuge may not be so in an industrial centrifuge or in a filter. It would be extremely useful to investigate the impact of pre-purification upon chromatography at the laboratory scale during process development because this will help to avoid serious problems later on at larger scale. The USD column would be ideal for this purpose but in order to make this possible good scale-down mimics would be required for the other operations. In many cases these mimics are already available and should facilitate faster more efficient process development.

4.8 Conclusions

The use of very small columns such as the USD column used in this study for preparative and design purposes is feasible. A number of phenomena make operation and interpretation of results difficult but it is possible to identify, quantify and compensate for these effects so that useful process data is obtained. The extra experimentation and calculations required to make the necessary adjustments are straightforward and quick to perform so that rapid data acquisition is possible. Overall the use of an USD column in process development would enable chromatographic operations to be studied quickly, thoroughly and with minimal use of material. It should be possible to predict yields and purification factors achievable at large scale and determine the optimum loading levels for the column. The impact of early separation steps upon chromatography could also be investigated so long as there was realistic process material available to load the column with.

Although the USD column has several potential applications it is not the perfect scale-down model as there are still some significant limitations. This study did not look at peak resolution since it was not relevant to this particular process, but this is a critical factor in many chromatographic operations and it remains to be seen how the USD column could be operated to mimic this in a larger column. Another limitation is that the USD column does not produce material that is representative of the large scale due to the column wall effects. This makes studying the next step difficult since realistic feed material is not available.

CHAPTER 5
VALIDATION

5.1 Introduction

Before a manufacturer can bring a new biological product to market, regulatory approval of the process and product must be obtained. Regulatory requirements and industry standards have increased dramatically over the past three decades which has lead manufacturers to increase their spending in order to achieve a compliant process. One important requirement of all the major regulatory bodies is that the manufacturing process must be validated. Process validation can be defined as establishing documented evidence that a process consistently produces a product that meets predetermined specifications and quality requirements. This is a vital part of ensuring the quality, safety and efficacy of the product (Rosendale 2002).

Validation can be a lengthy and expensive process but it is essential to ensure product safety and quality. Consideration of validation issues should begin as early as possible, usually at the laboratory development stage since the process will need to be validated before clinical trial material can be manufactured. Even when a process has been operating successfully for many years validation will still be ongoing to ensure that equipment is still operating within specification (Rosendale 2002, Hoare *et al* 2001). To make validation as simple as possible a company should have well-established programmes for quality control, process development and general GMP compliance. Since technology is always changing so are approaches to validation and it is important that a company keeps up to date with the latest trends (e.g risk based approach to equipment operation).

5.2 The validation process

Process validation involves several stages. The first stage is creating the Validation Master Plan (VMP) which states what is to be validated, how, when and by whom. Each stage must be completed, written up and signed off before the next stage can begin. Figure 1 outlines the main components of the validation process.

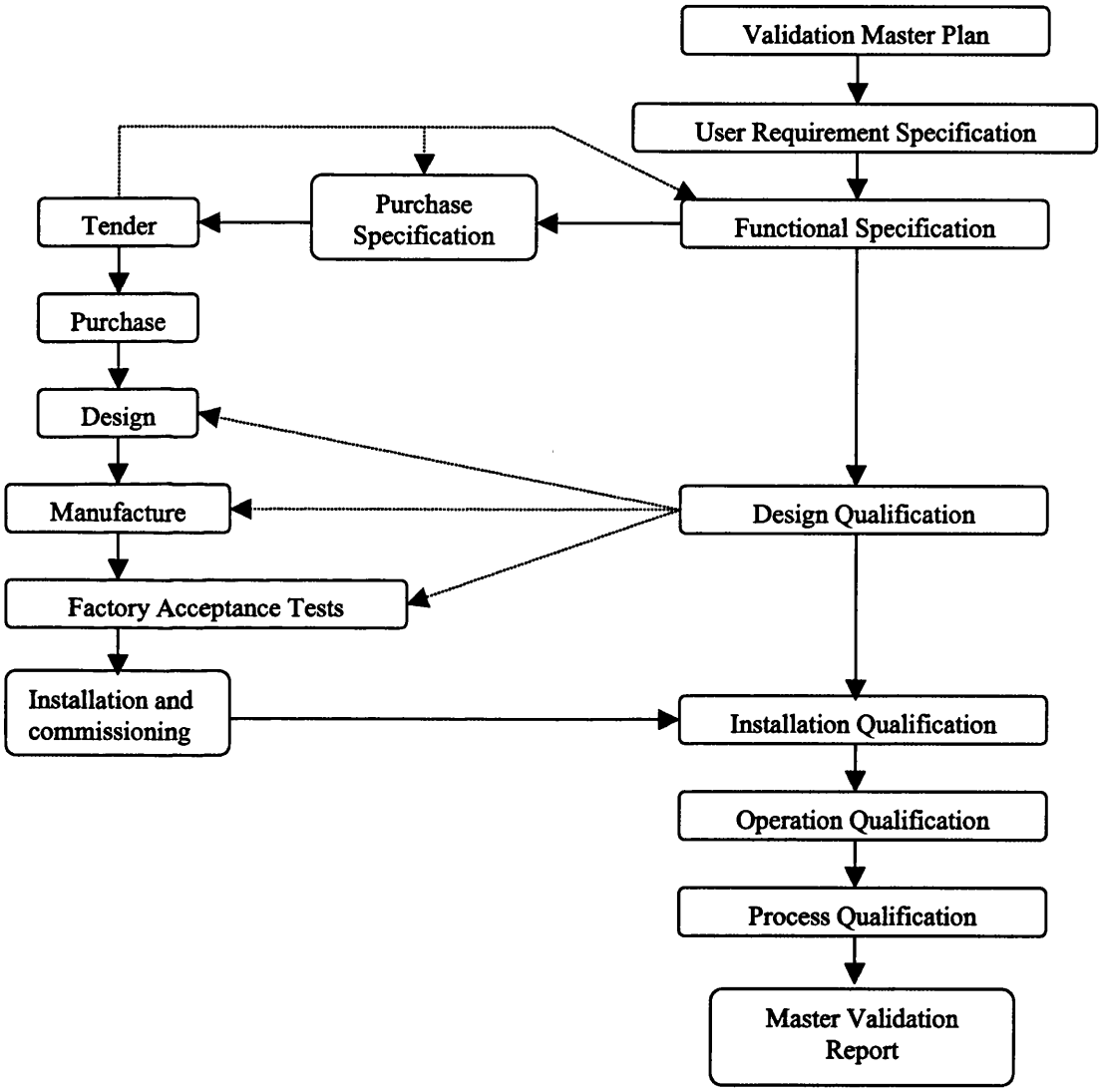

Figure 5.1. Stages in the validation of a typical process.

The validation master plan (VMP) is prepared by a multidisciplinary team and outlines the overall validation strategy and the various responsibilities. The user requirement specification (URS) then outlines exactly what is required of the equipment in terms of process, environmental, safety and regulatory requirements. This then forms the basis for the functional and purchase specifications, which are often drawn up in conjunction with the equipment supplier. Once these have been finalised, design of the equipment will commence. Both supplier and customer will approve the final design before manufacture begins. The design qualification (DQ) ensures that the equipment design conforms to operational and regulatory expectations and that all the user requirements have been satisfied. This will normally incorporate electrical schematics, process and instrumentation diagrams and descriptions of all the major components. Once the equipment is manufactured factory acceptance tests (FAT's) ensure the final design has been followed and URS is met. If the equipment passes factory acceptance tests it can be delivered to the end user for installation. However, this may not be the end of the supplier's role in the validation process as they may be required to assist in installation and IQ of the equipment.

The validation process now involves a series of tests and procedures that aim to show the system is installed, configured and operating to the written standards that have been set. The main testing areas can be defined as follows. The installation qualification (IQ) aims to verify that the equipment is properly installed, meets specifications and is in an environment suitable for its use. IQ activities will usually include instrument calibration and functionality testing of pumps, valves etc. Once IQ is complete the operational qualification (OQ) is used to demonstrate that the equipment operates as intended across all anticipated ranges. The OQ will involve verification of proper operation of controllers, recorders, indicators and alarms and also ensures the system can operate as

an integrated whole. Up to this point only water is used in the system but at the next stage, process qualification (PQ), real process materials are used. PQ essentially aims to show that the process can produce a product that meets predetermined acceptance criteria; this will involve the integration of equipment, materials, utilities, personnel and procedures. The final stage of the PQ will consist of at least 3 full-scale batches, which are analysed in detail to show that all acceptance criteria have been met and the product is suitable for use. Normally before the PQ batches are run a manufacturer will perform several shakedown runs to familiarise operators with equipment and fine-tune the operational process. This reduces the risk of PQ runs being ruined by operator errors.

In addition to equipment validation the process and the product itself need to be validated. During the early stages of process development, often while the process is still in the lab the product is defined in terms of its physical and biochemical characteristics so that acceptance criteria can be set for the product. The acceptance criteria will specify the product identity, purity, concentration and acceptable levels of impurities. Defining the product allows the critical process parameters to be identified (ones that affect product quality) so that values and tolerances can be set for these parameters. Once the acceptance criteria have been set the aim becomes to demonstrate that the process operates within preset ranges so that a product of defined purity, potency and identity is made.

As well as the above procedures another important aspect of validation is documentation (Chapman 1984). Documentation should provide an overview of the process and highlight the location of important validation information. Protocols and SOP's are required for operating, cleaning, maintaining and storing of equipment and staff must be trained to perform these tasks. Batch and training records are used to show that the

proper procedures were followed and that staff are up to date on their training. A senior member of Quality Assurance (QA) staff will usually sign off batch records to ensure all protocols and SOPs were correctly adhered to.

5.3 A validation example: Chromatography

Validating a chromatography operation follows the same principles as validating any other unit operation. The cleaning and reuse of chromatography media poses particular difficulties since it is hard to quantify the degree of fouling on a column or prove that all contaminants have been removed. Some manufacturers will only use chromatography resin once before discarding it, however with the right approach it is possible to validate the use of resin for many process cycles, often significantly reducing process costs.

5.3.1 Process conditions

Chromatography is virtually always used to remove contaminants so a well-characterised feed stream is necessary to define what happens on the column (Lutsch 2002). The identity and amount of bound materials will need to be established so that contaminant removal levels and product yield can be calculated. It is usual for the process flow rate, fraction collection times, pH and conductivity to be controlled. Other parameters such as pressure drop and temperature are monitored to demonstrate that they remain within predetermined limits. The precise loading and elution conditions will also need to be specified since these will greatly affect process performance. Monitoring the absorbance of the column effluent can be used to determine when to stop loading and to separate different fractions.

To determine what happens to various species in the feed when loaded onto the column specific assays will be needed. The assays themselves must be validated in order to prove

they are reliable and determine the range, accuracy and limit of detection. A wide range of assays may be required to test for all the various components including species that are not present in the feed such as product aggregates and ligands that have leaked from the matrix. These assays are usually performed offline and are retrospective.

5.3.2 Raw materials

The raw materials used in a chromatography operation must be of a suitably high standard and will need to meet certain acceptance criteria stated in the raw materials compliance before they can be used. Raw materials suppliers will usually be audited by the manufacturer and warehouse or other storage conditions will be tested for suitability. Raw materials that should be tested will include water and chemicals used in buffer preparation and the chromatography media itself. In house testing should be performed in addition to the manufacturer's own analysis to confirm the identity and compliance of the raw materials. Unntill test results have been verified the material should be quarantined and not released for use in production. When commercial cleaning products are used on or in the system the exact composition of these products must be known so that no unknown species can be introduced into the system, which may affect product quality.

5.3.3 Hardware and software

The complete chromatographic system should be described in the DQ. This should include the column itself along with distributors; all pipe work, pumps and valves and all sensors and detectors. The OQ should demonstrate that each individual piece of equipment (pumps, sensors, valves, etc) is working within acceptable ranges and tolerances. The OQ should also demonstrate the readiness of a packed chromatography column for use. This will involve measuring the HETP and peak asymmetry on the

column to show that column packing procedures are suitable and that the bed is free from air bubbles and other anomalies. This must be repeated each time the column is packed. PQ will demonstrate that the chromatography operation can achieve the desired separation and produce a product that meets all acceptance criteria.

Any computer software or hardware used on the column will also need to be validated. Validating computer-controlled processes will normally involve demonstrating that GAMP (Good Automated Manufacturing Practice) standards have been followed (Birkin 2002). These are designed to ensure that any automated process or equipment operates correctly when brought into use and that suitable fail-safe and backups are in place. This is quite a specialised area of validation and often requires individuals with a programming background. If specialised software (not off the shelf) is used then a copy of the source code and programming language used will be required. With all software the date of installation, serial numbers, minimum specifications and version number should be included in the IQ and OQ. The exact function of the software will also need to be specified in a software functional specification.

5.3.4 Cleaning and storage

Since many chromatographic matrices will be used more than once, cleaning and storage conditions are of utmost importance. A thorough cleaning and regeneration regime must be applied to the column after every use in order to reduce the risk of contamination and to prolong the column lifetime (Lutsch 2002). The cleaning and regeneration cycle will need to be defined in terms of reagents used, concentrations, contact times, temperature and order of use. Usually a combination of small-scale studies and full-scale runs are employed in cleaning validation. Small-scale studies are used to develop the best cleaning regime and test it by challenging the column with a series of dirty streams to

represent worst-case scenarios. Large-scale runs then demonstrate that the cleaning regime is effective in a real process situation which involves extensive testing to ensure that cleaning has been effective. This will usually involve swab testing, monitoring rinse water properties (conductivity, TOC) as well as visual inspection and testing for specific contaminants (Gariepy et al 1998). Testing (e.g. HETP) is performed between operations to ensure that cleaning has been effective and verify the matrix is still suitable for use. However before the column is reused it is also important to show that all the cleaning reagents have been removed from the system. Criteria need to be set to indicate when a column has been successfully cleaned, if the criteria are not met then the matrix must be replaced and the column will need to be repacked.

If there is a sizable time delay (typically more than 12 hours) between one column use and the next then the column must be stored between uses. A suitable storage buffer will need to be found that maintains the functionality of the matrix, is compatible with all wetted parts of the system and restricts microbial growth. The effects of the storage buffer will need to be studied with time to determine how long a column can be safely stored. When the column is revived from storage the storage buffer should be rinsed out of the system and the matrix re-equilibrated ready for use. It is common to perform a HETP test prior to use to verify column integrity.

5.3.5 Column lifetime

The chromatographic media will inevitably deteriorate with time so there needs to be some system in place to indicate when the column requires a re-pack. A column lifetime study will indicate how many cycles the matrix can undergo before it needs replacing. Often chromatography columns will have a validated lifetime, which is a measure of how many cycles the matrix can undergo before it requires replacing. There are two

approaches to validating the lifetime of a chromatography resin. The first is a concurrent approach where the resin is used at full scale to manufacture actual product for market. The second approach is prospective and uses small-scale laboratory mimics. The latter method is often beneficial since it reduces the level of analysis required after each step and eliminates the need to put product in quarantine while the resin is assessed (O'Leary, 2001).

5.4 Validating scale-down mimics

Good scale-down methods can make validation of a large-scale process easier by enabling greater process knowledge and understanding. Since a lot of data that would be too expensive to generate at large scale can be generated using scale-down techniques a thorough study of the process responses is possible. This makes it easier to define critical process parameters and set operating limits and acceptance criteria for the process.

5.4.1 Uses for scale-down mimics

The vast majority of process validation work is carried out at full or pilot plant scale. However, in some situations it is advantageous either in terms of time, cost or safety to carry out validation studies at small-scale (Sofer 1996). The ultra scale-down (USD) chromatography column described in this thesis could be validated if the user could demonstrate that it was a reliable mimic of an industrial column. Once validated the USD column could be used to investigate a number of process variables in the lab using minimal time and resources. Several (minimum of 3) full-scale runs would still be required for process validation as part of PQ. However, good scale-down techniques would enable some aspects of the process to be tested and fine-tuned in the laboratory thus minimising the risk of expensive, unforeseen problems at full-scale. Ultimately the

process knowledge and experience gained from scale-down work could keep the number of validation runs to a minimum and facilitate faster regulatory approval.

Viral clearance studies are virtually always performed on scale-down mimics since achieving realistic virus titres at large scale is not possible (Sofer 1996, Cameron et al 1997). The hazardous natures of viruses mean large-scale studies pose a significant risk both to workers and the environment. Scale-down mimics improve safety, are easier to control, reduce time and costs and allow higher more realistic virus titres. Other studies that may be performed using laboratory mimics include host cell protein clearance, nucleic acid removal, matrix regeneration and cleaning and column lifetime studies. O'Leary (2001) demonstrated that small-scale prospective validation studies could accurately predict the performance of resins at manufacturing-scale.

5.4.2 Design considerations

Performing validation studies at small-scale requires laboratory mimics that accurately predict the performance of industrial scale operations. These laboratory mimics must represent the full-scale operation in every way possible and perform in the same manner. In order to achieve this every detail of the scale-down mimic must be considered from the geometry of equipment to the types of material used.

5.4.2.1 Materials and apparatus

An accurate scale-down model will include all pipe work and auxiliary equipment associated with the operation. A mimic of a chromatography operation will consist of the column, matrix, distributors, pipe work, detectors, valves and pumps. It is important that all these components are represented as any changes to the system could result in a change in performance. It is also preferable to keep materials of construction for any wetted parts the same at both scales. This is because different materials may adsorb

proteins or leach contaminating molecules into the process streams. Obviously it is not always possible to use the same materials of construction at the two scales but great care should be exercised to ensure that whatever material is used does not interfere with the process.

5.4.2.2 Column dimensions and design

The design of the chromatography column itself will have a substantial effect on the performance of the laboratory mimic. Typically scale down is performed by maintaining the column bed height and linear flow rate and reducing the column diameter and volumetric flow rate. However, it is still possible to create a reliable laboratory mimic if the bed height is not maintained, although this may be more difficult. The most important factor is that the laboratory mimic behaves in the same way as the larger column. If a laboratory column, such as the USD column can mimic the yield and product purity of the industrial scale column then it can almost certainly be validated. Such a column could be used to study changes in process conditions, cleaning regimes and column lifetime but may not be suitable for virus removal studies. This is because the column bed height may have a significant effect on the level of virus removal.

The distributor is of critical importance in any chromatography column. Ideally when scaling down the distributor used in the small-scale column should be the same as that used in the industrial column. In practice this is often quite difficult since industrial columns use large stainless steel distributors, which are very difficult and expensive to mimic at smaller scales. It is often easier to use a different distribution system and demonstrate that the alternative system distributes the feed to the same degree as the larger distributor.

5.4.3 Operational considerations

There are a great many process parameters that need to be considered when scaling down. The volumetric flow rate should be suitably reduced but either the linear velocity or residence time (or both) should be maintained to replicate the contact times at industrial scale. These may have an effect upon binding, so if they are changed it will be necessary to show that the change has negligible effect on performance.

5.4.3.1 Temperature

The temperature of the operation also needs to be considered. Temperature can affect the viscosity of solutions, alter the rate of binding and disrupt the conformation of a protein. It is therefore important that the temperature of the process is controlled at both scales. This applies not just to the column itself but to all elements of the chromatography system, so if buffers are stored at 4°C on the industrial scale then the laboratory mimic must reflect this.

5.4.3.2 pH and conductivity

In order for a laboratory column to mimic an industrial column the pH and conductivity of the feed and buffers used must be maintained since they often have a direct impact upon binding. Adjustments to pH or conductivity must be performed using the same concentration and type of solution, and proper mixing is required to avoid localised extreme conditions. Since the pH may also make a significant contribution to the level of virus clearance it is important to maintain the contact times at specific pH's and to determine whether clearance is being achieved by removal or inactivation.

5.4.3.3 Feedstock

The concentration of salt, protein, product and other components in the feedstock must be identical at both scales. Ideally for validation studies actual feedstock from the industrial process should be used. However at small scale the addition of material to the feed, such as a viral spike, can have an effect on concentration, ionic strength and pH. The consequences of any changes are variable and may be hard to detect. To avoid any undesirable effects the pH and/or ionic strength should be adjusted to the appropriate values. If it is not possible to get the conditions exactly as they are in the commercial process then a design of experiments approach is used to determine the sensitivity of the process to the parameters. If the process is insensitive to the parameter then it may not be essential to maintain that parameter upon scale-down.

5.4.4 Chromatography buffers and media

Every company will have its own, well-established protocols for large-scale buffer preparation. In order to validate a laboratory mimic the same protocols should be followed at lab scale to ensure all the properties of the buffer are retained. It is also important that the ingredients used in buffer preparation are of the same standard as the commercial scale. So if WFI is used to make buffers at the industrial scale it must also be used for the laboratory mimic.

The media used in the scale-down mimic must be exactly the same as the media used in the commercial process. It is not sufficient to use media with the same functional group but different base matrix since this can affect porosity, permeability and non-specific binding. Any media used must also come from a lot that has been shown to meet the required user specifications. The manufacturer may perform tests on the media but if any

additional testing is carried out before commercial use, the media for the laboratory mimic must also pass these tests.

The media used in the laboratory mimic could be media that has already been used in the industrial scale process. This is because used media is more representative of material used at large scale and hence will behave more like the commercial process. Used media from manufacturing can also be used to see if the ability of the media to remove a particular contaminant has changed with usage. However the used media must not have exceeded the maximum number of cycles determined from the lifetime study.

5.4.5 Ultra-scale-down compared to traditional scale-down

Both USD and traditional scale down methods are quick, simple and inexpensive to perform. However, USD techniques have several advantages over their traditional counterparts. The major limitation of the traditional techniques is that they tend to focus on a few important parameters and rely solely upon calculating how these parameters should be scaled down. As a result of this, traditional scale-down methods often represent an ideal situation that cannot be achieved at large scale since important process considerations are ignored. The effect of this is that the scale-down methods over predict the performance of large-scale equipment.

Ultra-scale-down techniques take into account all aspects of large-scale operation including scale-down parameters, mode of operation and ancillary equipment. Factors such as shear in centrifugation (Boychyn et al 2001), cake structure and formation in filtration (Reynolds et al 2003) and diffusion effects in chromatography may impact upon the performance of an operation and so need to be considered. USD techniques incorporate these factors either by mimicking them directly (e.g. USD filter) or by

establishing a correction factor (e.g. diffusion in chromatography). Overall, USD techniques through a greater understanding of large-scale processes allow the production of representative process material, improve the quality of process data and enable better process understanding. All this should contribute to a more streamlined validation process with fewer unforeseen process of operational issues.

5.5 Conclusions

Scale-down laboratory mimics of large-scale chromatography operations can be useful validation tools. It is often advantageous to use small-scale mimics in validation studies since they reduce the time and resources required, leading to a reduction in costs. However it is absolutely essential to demonstrate that the scale-down mimic duplicates the performance of the large-scale column. Since the molecular interactions that occur during chromatography are very complex it is important to employ sound scientific principles when scaling down. The laboratory mimic should imitate the industrial process in every way possible and where discrepancies occur it is important to demonstrate that they do not affect performance.

The USD column could probably be used in some validation studies such as lifetime testing or assessing storage buffers when operated at the same linear velocity as the larger column. Assessing cleaning regimes and virus removal may also be possible although very careful assessment of the contribution of contact time would be necessary and it is likely that the USD column would have to be operated at the same residence time as the larger column. Since the different modes of operation can be used for different validation studies it will be necessary to prove equivalency in both modes.

CHAPTER 6
CONCLUSIONS

Accurately mimicking the performance of large-scale unit operations requires that the environments to which the process material is exposed be reproduced in the laboratory. A number of factors can influence this environment, the most obvious being physical parameters such as temperature and chemical parameters such as pH and conductivity. These parameters are widely considered as the most important process variables and thus have been the focus of most traditional scale-down work. However, for scale-down to provide a true representation of the industrial scale there are other factors that must be considered these include but are not limited to; quality of materials used, equipment set-up, time dependencies, modes of operation (batch verses continuous) and sample analysis. Elements of the large-scale operation that are not mimicked or accounted for somehow in the scale-down model may have an unexpected impact upon process performance and thus affect the accuracy of scale-down data. In chapter 2 it was found that process changes, which appeared to have a favourable impact on product yields in, the lab did not translate to full-scale operation, despite good agreement with pilot plant data. This highlights the need to consider every aspect of an operation that could affect performance, whether it be engineering, biochemical or operational. Ultra scale-down (USD) aims to do this and in doing so make better use of scale-down studies.

Making process changes in one unit operation will have an effect on subsequent operations and may alter overall performance. When studying unit operations in isolation the full impact of any changes further downstream can be very difficult to predict. The consequences of this could be that performance is negatively affected, operational limits exceeded (e.g tank volumes) or regulatory commitments violated. To avoid this, ideally

the entire purification sequence should be studied, but in order for this to happen the scale-down mimics must produce material representative of the large-scale.

The development of the USD filter in chapter 3 aimed to address some of the issues of filtration scale-down and enable better scale-up predictions. Using data from the USD filter it was possible to predict the flux in a larger continuous operation filter with the same applied pressure profile. These predictions were superior to those made using data from batch laboratory tests as they successfully accounted for the initial low-pressure high rate period of the operation. The reason for this was found to be that cake compressibility actually varies throughout filtration. In batch tests a single value (the average) for cake resistance is found while the USD filter can be used to determine how compressibility varies with time. The result was that compressibility increased in an exponential manner between a lower and upper limiting value.

The benefits of the USD filter over traditional laboratory batch filtration arise from the fact that the continuous mode of operation facilitates the formation of a homogenous filter cake, with precipitate and filter aid roughly equally distributed. This type of filter cake is more representative of those formed in large-scale filters (due to their mode of operation) and results in the production of a more representative filtrate. The clarity of this filtrate is actually worse than that produced by laboratory batch filtration, but this is explained by the more open pore structure bestowed upon the homogenous cake by the even distribution of filter aid. Since the USD filter can produce "realistic" process material it should enable further process steps to be studied at laboratory-scale or form part of a whole process model.

The use of very small columns such as those used in chapter 4, to enable rapid process development is an exciting possibility since chromatography is so widely used in biotechnology and can be quite resource intensive even at laboratory-scale. However several phenomenon have been identified that can make operation and interpretation of results difficult. These phenomena are caused by three main factors diffusion in pipework, the extra-column hold up volume and wall effects. It was found that with some extra experimentation and careful analyses of results it is possible to compensate for these effects and adjust the resulting chromatogram and process outputs (yield and purity). This enabled the USD column to predict the performance of larger laboratory columns, which have been shown to be good mimics of industrial-scale columns.

It was discovered that operating the USD column in a manner that mimicked the residence time of larger columns was extremely problematic. Due to the very low flow rates involved the effects of molecular diffusion, mainly in the pre-column pipework, became very significant and made it difficult to distinguish between different stages of operation in the chromatogram. This also resulted in a loss of product since peak resolution was so poor. It was much easier to operate the USD column in a manner based on scaling down at a constant linear velocity. This mode of operation gave some useful information about process yields and purities for the system studied. However it may be difficult to apply to a true chromatography step that involved peak separation since this is heavily dependant on residence time and column length.

CHAPTER 7
FUTURE WORK

7.1 Introduction

This thesis has looked at scale-down as a tool to aid rapid and efficient process development. It has been demonstrated that careful design and operation of scale-down mimics to overcome the inherent differences between laboratory and industrial-scale operations can yield reliable process data to aid in design and optimisation. However current techniques still have room for improvement and there are several other potential uses for scale-down technology within process development that remain largely unexplored. This chapter discusses some of the directions scale-down research could take in the future.

7.2 Further development of ultra scale-down mimics

7.2.1 Depth filtration

Since the completion of this work an improved version of the USD filter has been constructed and tested. The new design is more compact since rotation is achieved by placing the main filter chamber on a bearing, which is connected to a small motor, this eliminates the need for the belt drive and makes operation less awkward. The new design also allows the filter to be swapped easily between a horizontal and vertical position and is mounted on a solid base so that it is more stable and convenient to transport. The set-up could be improved further if automatic data recording was used, particularly for pressure, pump rate, flux and even filtrate clarity. It would also be useful to have a control program that could set the pump rate based upon flux or pressure measurements. This would enable industrial operating conditions to be mimicked very easily in the laboratory and reduce the laboriousness of experiments.

In this work it was not possible to air-dry the filter cakes in the USD filter due to its design. However air-drying and cake washing are both important parts of any filtration operation since they can impact upon process yields. Further work could investigate the volumes of air or wash liquid required for optimum results, this type of work is difficult to perform with traditional batch experiments since they generally require dismantling for additions to be made. In contrast with the correct set-up the USD filter could be used to study the whole filtration operation from pre-coat application right through to washing, drying and possibly even clean in place (CIP) operations.

7.2.2 Chromatography

Further research needs to be done with the USD column. The chromatographic system studied in this thesis was a contaminant capture step and therefore did not involve any peak separation or elution gradients. However gradient elution is very common in industrial operations since it enables excellent separation. It would be useful to investigate how the small-scale phenomena observed in chapter 4 affect peak separation. The resolution of peaks depends on several factors including the slope of the gradient, the linear velocity and bed height; therefore altering these parameters will affect separation. It may be that a similar approach to that taken in this thesis can compensate for any differences in performance; although laboratory operating conditions would have to be carefully selected to reproduce the changing conditions across the column as the gradient progresses.

Understanding of chromatographic operations has advanced considerably over the last few decades but it is still not possible to predict the separation of complex biological feedstocks. As a result, small-scale experimentation is likely to play a significant role in the development of chromatographic operations for the foreseeable future. In this thesis

the problems of small-scale operation were identified and compensated for. However ideally these problems should be overcome by good design and engineering of scale-down models so that they can be operated to directly mimic large-scale performance. In order for this to happen a better understanding of the microscopic structure and properties of packed beds is needed. This will lead to a better understanding of packing processes, bed compression and column wall effects. Ultimately if very accurate scale-down models can be designed then it will be possible to study in depth the topic of column fouling and the effects of pre-purification, which would be invaluable when developing and optimising a purification process.

7.2.3 Other unit operations

The accuracy of data from scale-down studies is of vital importance if they are to be employed to make decisions regarding process development. Many currently employed scale-down techniques focus on the core aspect of a unit operation (e.g. clarification in centrifugation) but in order for fully informed development decisions to be made other aspects of operation often need to be considered. Ultra scale-down techniques can provide more reliable scale-down data than traditional methods by incorporating all the important aspects of large-scale operation. However, if extensive process development is to take place at the laboratory scale then USD techniques will be required for all major processing operations. This will allow different combinations and sequences of process steps to be investigated and compared. Although much work has already been performed on scaling down the commonly used unit operations more work is required to account for the variety of industrial equipment and for newer technologies (e.g. expanded bed chromatography). The accuracy of scale-down techniques can also be improved further through better design and consideration of all performance affecting factors.

7.3 Scaling-down entire processes

Laboratory scale-down tends to focus on a single operation at a time, which is far from ideal since it becomes difficult to assess the impact that variations in the performance of one operation have further downstream. In addition, for convenient scheduling and increased flexibility large-scale operations commonly have intermediate freezing or holding stages that are not taken into account during scale-down studies. This means the product is exposed to certain conditions for much greater time periods at large-scale which could result in denaturation, aggregation or transformation. If small-scale studies mimicked the entire process then these and other unexpected effects would become evident at an early stage and costly large-scale losses could be avoided. Scaling-down an entire process would also make it easier to determine what the real critical process parameters were and therefore optimisation would be more straightforward. However in order to do this accurate scale-down mimics for all the operations involved are required and the material passed from one step to the next must be representative of the large-scale. The biggest challenge to performing this kind of scale-down study will most likely be mimicking the transfers between stages and linking all the operations together.

7.4 Integration of scale-down and process modelling

Computer models are widely used in process development throughout the chemical and pharmaceutical industries, for a variety of purposes. However the application of these models to biological systems is difficult since process materials are less well defined, there is a lack of historical data and complex interactions can occur between process components. The integration of scale-down and computer modelling has the potential to overcome these problems and make bioprocess modelling a realistic development tool. Scale-down experiments could be used to determine critical process parameters or material properties that are otherwise difficult to predict (e.g. specific filter cake

resistance). This information could then be fed into a computer model along with appropriate process variables to predict the performance. The key to this sort of approach is identifying the key determinants of process performance and obtaining accurate scale-down data to input into the model.

Computer modelling can be used to study many different aspects of a process, not just the overall performance. The economics of a process have a huge influence upon the commercial viability of a product and can be used as a source of competitive advantage. However estimating costs, both capital and running costs for a new process is not easy. Economic models require inputs relating to material usage, timing, size of equipment and energy consumption, but this sort of information is not usually available until late in the process development stage forcing scientists to give a best guess in order to asses costs. This problem could be overcome using scale-down models to give a more realistic idea of the process requirements (e.g. estimating volumes of wash buffers required or giving an idea of the turnaround time required to clean a column). The integration of scale-down and economic modelling would be a powerful tool for assessing the business impact of process decisions and developing a cost effective process. Again the challenge remains to develop scale down techniques that accurately mimic the operation and performance of industrial operations.

APPENDIX A
ASSAY RESULTS FOR BPL AND LABORAL IgG PROCESSES

BPL trial results

Total IgG is calculated from A280 result for samples 3, 4, 6 and 7; turbitime result for samples 1 and 2; and RID for sample 5. For a description of each sample see table 2.2. *All protein is assumed to be IgG and an extinction coefficient of 1.4 is used to calculate the IgG concentration from the absorbance at 280 nm.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	250	6.6	7.45	1.37	19.40	-	1.65
2	480	3.4	3.55	0.61	8.50	-	1.63
3	535	1.9	2.10	0.25	3.48	2.49	1.33
4	535	1.7	1.78	0.23	2.99	2.14	1.14
5	610	-	0.10	0.14	0.30	-	0.06
6	12.6	73.1	-	-	112.9	80.64	1.02
7	13.3	67.1	-	-	94	67.14	0.89

Table A1: Assay results for the first pilot-scale BPL trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	250	7.0	8.03	1.55	23.80	-	1.75
2	464	3.5	3.72	0.76	8.65	-	1.62
3	539	2.4	2.48	0.37	3.80	2.71	1.46
4	539	2.3	2.74	0.30	3.42	2.44	1.32
5	604	-	0.07	0.21	0.29	-	0.04
6	14.3	NT	-	-	116.90	83.50	1.19
7	14.1	NT	-	-	98.07	70.05	0.99

Table A2: Assay results for the second pilot-scale BPL trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	250	6.09	7.48	1.45	26.61	-	1.52
2	464	3.79	4.22	0.71	10.49	-	1.76
3	536	2.1	2.86	0.30	3.66	2.61	1.40
4	536	2.1	2.64	0.28	3.62	2.59	1.39
5	605	-	0.03	0.12	0.29	-	0.02
6	14.55	83.4	-	-	123.6	88.3	1.28
7	14.34	64.9	-	-	104.4	74.6	1.07

Table A3: Assay results for the third pilot-scale BPL trial.

Laboral trial results

Total IgG is calculated from A280 result for samples 3, 4, 6 and 7; turbitime result for samples 1 and 2; and RID for sample 5. For a description of each sample see table 2.2.
*All protein is assumed to be IgG and an extinction coefficient of 1.4 is used to calculate the IgG concentration from the absorbance at 280 nm.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	226	7.28	8.91	1.59	19.30	-	1.69
2	600	2.65	3.17	0.59	6.70	-	1.59
3	672	1.50	2.22	0.3	3.00	2.33	1.57
4	672	1.50	2.2	0.26	2.67	2.26	1.52
5	788	-	0.02	0.26	0.27	-	0.02
6	14.4	82.3	-	-	113.8	81.28	1.17
7	14.4	65.4	-	-	100.0	71.42	1.03

Table A4: Assay results for the first pilot-scale Laboral trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	226	6.6	6.71	1.50	12.9	-	1.49
2	544	4.1	3.44	0.64	8.1	-	2.23
3	626	2.3	2.39	0.33	3.9	2.79	1.75
4	626	2.19	2.29	0.30	3.4	2.43	1.52
5	734	-	0.04	0.26	0.2	-	0.03
6	16.8	NT	-	-	105.5	75.36	1.26
7	16.3	NT	-	-	98.3	70.21	1.14

Table A5: Assay results for the second pilot-scale Laboral trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	227	7.2	9.6	1.81	22.6	-	1.63
2	576	3.09	3.29	0.64	7.78	-	1.78
3	667	1.74	2.25	0.34	3.26	2.33	1.55
4	667	1.65	2.25	0.31	3.17	2.26	1.51
5	777	-	0.01	0.29	0.28	-	0.01
6	24.12	51.9	-	-	73.10	52.21	1.26
7	24.63	39.8	-	-	65.16	46.54	1.15

Table A6: Assay results for the third pilot-scale Laboral trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	228	6.58	8.20	1.75	19.6	-	1.50
2	568	2.53	3.35	0.65	7.3	-	1.44
3	647	1.97	2.30	0.50	3.64	2.60	1.68
4	647	1.75	2.00	0.40	3.27	2.34	1.51
5	756	-	0.01	-	0.32	-	0.01
6	65	NT	-	-	27.46	19.61	1.27
7	65	NT	-	-	26.06	18.61	1.21

Table A7: Assay results for the fourth pilot-scale Laboral trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	226	7.1	6.7	1.35	18.8	-	1.60
2	520	3.01	2.8	0.50	7.6	-	1.57
3	574	1.82	1.5	0.30	3.26	2.33	1.34
4	574	1.75	1.4	0.30	3.11	2.22	1.27
5	664	-	NT	0.20	0.33	-	-
6	16.59	NT	-	-	109.9	78.50	1.30
7	21.62	NT	-	-	71.2	50.86	1.10

Table A8: Assay results for the fifth pilot-scale Laboral trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	4158	7.06	7.47	1.32	17.00	-	29.36
2	9142	2.79	2.41	0.55	7.40	-	25.51
3	10232	1.70	1.16	0.29	3.01	2.15	22.00
4	10232	1.75	1.45	0.28	3.10	2.21	22.61
5	11821	0	0	0.21	0.32	-	2.48
6	243	64.6	64.0	0.84	108.64	77.60	18.89
7	238	76.9	64.9	0	102.76	73.40	17.47

Table A9: Assay results for the first full-scale Laboral trial.

Sample	Volume (L)	IgG TT (g/L)	IgG RID (g/L)	Alb RID (g/L)	A280 (nm)	IgG* by A280 (g/L)	Total IgG (kg)
1	3997	7.06	7.10	1.53	19.6	-	28.21
2	9317	3.05	3.11	0.52	7.02	-	28.42
3	10136	2.02	2.12	0.33	3.64	2.60	20.47
4	10136	1.99	2.12	0.31	3.24	2.31	20.17
5	11398	0.02	NT	0.26	0.32	-	-
6	260	61.3	-	-	110.6	79.00	20.54
7	250	58.4	-	-	106.1	75.79	18.95

Table A10: Assay results for the second full-scale Laboral trial.

APPENDIX B
MASS BALANCES

Mass balances were performed for all experiments. IgG was measured using a Behring turbitime unit and total protein was measured by a BCA assay. The flowthrough and wash fractions were combined to calculate the yield and final purity of IgG.

	1	2	3	4	5		
Load protein	-	4768.0	5235.1	5093.4	5062.4		
Load IgG	2685.8	2267.1	2424.4	2591.6	2591.6		
FT protein	-	3799.1	4332.5	3831.5	3909.1		
FT IgG	2552.0	2065.5	2323.4	2405.8	2488.0		
Wash protein	-	130.9	201.5	261.1	222.3		
Wash IgG	0.0*	21.0	21.0	43.0	48.0		
Eluate protein	-	640.2	610.5	600.9	654.6		
Eluate IgG	111.7	98.3	107.8	104.4	100.0		
Total protein	-	4570.3	5144.5	4693.5	4786.0		
Total IgG	2663.7	2184.8	2452.2	2553.1	2636.0		
% Difference	-	4.15	1.73	7.85	5.46		
% Difference	0.82	3.63	-1.14	1.49	-1.71	**Mean**	**STD**
Initial purity	-	0.48	0.46	0.51	0.51	0.49	0.03
Final purity	-	0.53	0.52	0.6	0.61	0.58	0.05
Purification factor	-	1.12	1.12	1.18	1.20	1.15	0.04
Yield	95.02	95.50	95.60	95.91	96.21	95.65	0.45

*wash included in flowthrough
% Difference is the difference between the amount of protein in the load and the total amount of protein in the individual fractions.

Table B1: Mass balance for the XK16 column.

	1	2	3	4	5		
Load protein	-	138.2	133.2	136.9	134.3		
Load IgG	62.8	61.8	62.2	64.8	65.2		
FT protein	-	88.9	82.8	87.4	89.6		
FT IgG	52.9	50.2	49.2	51.9	52.9		
Wash protein	-	16.7	23.7	14.5	17.3		
Wash IgG	2.5	5.6	5.0	2.9	4.4		
Eluate protein	-	25.7	30.4	33.5	28.7		
Eluate IgG	7.0	5.4	9.1	8.0	4.5		
Total protein	-	131.3	136.9	135.4	135.6		
Total IgG	62.4	61.2	63.2	62.7	61.8		
% Difference	-	4.94	-2.83	1.09	-0.98		
% Difference	0.72	0.98	-1.68	3.23	5.23	**Mean**	**STD**
Initial purity	-	0.45	0.47	0.47	0.49	0.47	0.02
Final purity	-	0.53	0.51	0.54	0.54	0.53	0.01
Purification factor	-	1.18	1.09	1.13	1.10	1.13	0.04
Yield	89.9	90.13	87.13	84.47	87.88	87.94	2.36

% Difference is the difference between the amount of protein in the load and the total amount of protein in the individual fractions.

Table B2: Mass balance for the USD Column scaled down based on constant residence time.

	1	2	3	4	5		
Load protein	-	131.7	135.0	136.6	138.4		
Load IgG	62.6	63.46	62.6	60.5	68.0		
FT protein	-	94.3	93.7	97.5	101.9		
FT IgG	55.4	57.9	56.6	55.6	57.9		
Wash protein	-	8.5	10.2	10.8	11.6		
Wash IgG	3.0	2.3	2.0	1.8	2.9		
Eluate protein	-	25.9	24.1	25.1	26.1		
Eluate IgG	4.5	4.6	4.3	4.8	5.4		
Total protein	-	128.7	128.0	133.3	139.6		
Total IgG	62.8	64.8	62.8	62.2	66.2		
% Difference	-	2.23	5.20	2.37	-0.87		
% Difference	0.72	-2.17	-0.27	-2.87	2.27	**Mean**	**STD**
Initial purity	-	0.48	0.46	0.44	0.49	0.46	0.02
Final purity	-	0.59	0.56	0.53	0.54	0.56	0.03
Purification factor	-	1.21	1.15	1.20	1.09	1.16	0.01
Yield	92.88	92.84	93.23	92.27	91.91	92.63	0.48

% Difference is the difference between the amount of protein in the load and the total amount of protein in the individual fractions.

Table B3: Mass balance for the USD Column scaled down based on constant linear velocity.

NOMENCLATURE

Letters	Description	Units
A	filtration area	m^2
A_{USD}	wall area of ultra scale-down column	cm^2
A_{XK}	wall area of XK16 column	cm^2
b	constant in equation 3.11	
c	mass of solids deposited per unit volume of filtrate	kgm^{-3}
C_{lab}	correction factor for laboratory centrifuge	-
C_{mc}	correction factor for multi-chamber bowl centrifuge	-
C_S	salt concentration	% saturation
E	concentration of ethanol	% v/v
F	fraction of protein remaining soluble	-
H_{flow}	contribution to HETP by hydrodynamic factors	cm
$H_{diffusion}$	contribution to HETP by diffusion	cm
$H_{mass\ transfer}$	contribution to HETP by mass transfer resistance	cm
HETP	height equivalent to a theoretical plate	cm
I_B	concentration of bound IgG	$mgmL^{-1}$
I_W	concentration of IgG in wall region	$mgmL^{-1}$
k	permeability of solids bed	
K	Kozeny constant	
K_S	constant in Cohn equation	
L	depth of solids bed	cm
L	length of chromatography bed	cm
m	constant in Niktari equation	
m	mass ratio of wet cake to dry cake	-
M	mass of filtrate	g
n	compressibility coefficient	-

n_0	lower limit of compressibility coefficient	-
n_∞	upper limit of compressibility coefficient	-
N	number of theoretical plates	-
P_B	concentration of bound protein	$mgmL^{-1}$
P_W	concentration of protein in wall region	$mgmL^{-1}$
P_{1-4}	constants in Boltzman equation	-
Q	filtrate flow rate	gs^{-1}
Q_{mc}	flowrate through multi-chamber bowl centrifuge	Lh^{-1}
R	retardation factor in equation 4.1	
R_m	resistance of filter medium	
S	protein solubility	
s	mass fraction of solids in suspension	
S_V	surface area per unit volume of particles	m^{-1}
t	time	s
t_{lab}	centrifugation time	s
t_R	retention time of component	s
u	interstitial fluid velocity	cmh^{-1}
V	volume of filtrate	mL
V_{lab}	volume of laboratory centrifuge sample	mL
V_R	retention volume of component	mL
V_{USD}	bed volume of ultra scale-down column	mL
V_{XK}	bed volume of XK16 column	mL
W	eluted peak width	cm

Greek letters	Description	Units
α	constant in Niktari equation	
α_{av}	average specific cake resistance	mkg^{-1}
α_0	cake resistance at unit applied pressure	mkg^{-1}
β	constant in Cohn equation	
ΔP	pressure drop	Nm^{-2}
ΔP_c	pressure drop across filter cake	Nm^{-2}
ΔP_m	pressure drop across filter medium	Nm^{-2}
ε	porosity of particulate bed	-
ε_0	gravity settled bed porosity	-
λ	chromatography bed compression	-
Σ_{lab}	equivalent settling area for laboratory centrifuge	m^2
Σ_{mc}	equivalent settling area for multi-chamber bowl centrifuge	m^2
ρ	density	kgm^{-3}
σ_L	standard deviation of eluted peak (length units)	cm
σ_V	standard deviation of eluted peak (volume units)	mL
σ_t	standard deviation of eluted peak (time units)	s
σ_{ex}	total extra column band broadening	
σ_{pipe}	contibution of pipe work to extra column band broadening	
σ_{det}	contibution of detector volume to extra column band broadening	
τ_{dead}	contibution of dead volumes to extra column band broadening	
τ_{el}	contibution of electronic response times to extra column band broadening	
μ	viscosity	

Abbreviations	Description
A280	Absorbance at 280 nm
α-2-macro	Alpha-2-macroglobulin
BPL	Bio Products Laboratory
cGMP	Current good manufacturing practices
DEAE	Diethylaminoethyl
DI	Deionised water
DOE	Design of experiments
F.A+1	Fraction A+1 of the Cohn fractionation process
F.B+1	Fraction B+1 of the Cohn fractionation process
F.II	Fraction II of the Cohn fractionation process
F.IV	Fraction IV of the Cohn fractionation process
IgA	Immunoglobulin A
IgG	Immunoglobulin G
NDA	New drug application
OD	Optical density
Peq	Plasma equivalent
PKA	Pre-kalikrien activator
RID	Radial immunodiffusivity assay
RVLF	Rotating vertical leaf filter
SCE	Small chemical entity
SOP	Standard operating procedure
TT	Turbitime unit
USD	Ultra scale-down
WFI	Water for injections

REFERENCES

Ambler, C. M. (1959). The theory of scaling up laboratory data for the sedimentation type centrifuge. *Journal of Biochemistry and Microbiology*, **1**: 185-205.

Arlington, S., Hughes, S., McKiernan, J., Hamilton, J., Palo, J. and Solano, P. (1998). Pharma 2005: An Industrial Revolution in R & D. Price Waterhouse Coopers.

Atkinson, B. and Mavituna, F. (1991). Scale-up of processes. Biochemical Engineering and Biotechnology Handbook, 2nd edition. Macmillan, Basingstoke, UK.

Author unknown. (2003). Increased length and complexity of the research and development process. In Pharmaceutical Industry Profile 2003. PhRMA.

Bell, D. J. and Dunnill, P. (1982). Shear disruption of Soya protein precipitate particles and the effect of ageing in a stirred tank. *Biotechnology and Bioengineering*, **24**: 1271-1285.

Bell, D. J., Hoare, M. and Dunnill, P. (1983). The formation of protein precipitates and their centrifugal recovery. *Advances in Biochemical Engineering*, **26**: 1-72

Birkin, S. (2002). GAMP Validation and Membrane Filtration Plant Specification. *Biopharm Europe*, March 2002, 58-64.

Bobrowicz, G. (1999). The compliance cost of hasty process development. *Biopharm*, August, 35-38.

Bonnerjea, J., Oh, S., Hoare, M. and Dunnill, P. (1986). Protein purification: The right step at the right time. *Bio/Technology*, **4**: 954-958.

Borchardt, R. T. (2001). Integrating drug discovery and development. *The Scientist*, **15**: 1-5.

Boschetti E. (1994). Advanced sorbents for preparative protein separation purposes. *Journal of Chromatography A*, **658**: 207-236.

Boychyn, M., Doyle, W., Bulmer, M., More, J. and Hoare, M. (2000). Laboratory scale-down of protein purification processes involving fractional precipitation and centrifugal recovery. *Biotechnology and Bioengineering*, **69:** 1-10.

Boychyn, M., Yim S.S., Ayazi-Shamlou, P., Bulmer, M., More, J. and Hoare, M. (2001). Characterization of flow intensity in continuous centrifuges for the development of laboratory mimics. *Chemical Engineering Science*, **56:** 4759-4770.

Byrom, D. (2000). Role and timing of process development for biopharmaceutical manufacture. *Pharmaceutical Technology Europe*, **12**: 52-56.

Cain, C. W. (1990). Putting the principles to work: Filter cake filtration. *Chemical Engineering*, **97**: 72-75.

Cameron, R., Davies, J., Adcock, W., MacGregor, A., Barford, J. P., Cossart, Y. and Harbour, C. (1997). The removal of model viruses, Poliovirus type 1 and Canine Parvovirus, during the purification of human albumin using ion-exchange chromatographic procedures. *Biologicals*, **25**: 391-401.

Chang, C. E. (1988). Continuous fractionation of human plasma proteins by precipitation from the suspension of the recycling stream. *Biotechnology and Bioengineering*, **31**: 841-846.

Chapman, K. G. (1984). The PAR approach to process validation. *Pharmaceutical Technology*, **8**: 22-36.

Clarkson, A. I., Bogle, I. D. and Titchner-Hooker, N. J. (1992). Modelling and simulation of fractional precipitation – comparison with pilot plant data. *Computers and Chemical Engineering*, **10**: 441-448.

Cohn, E. J., Strong, L. E., Hughes, W. L., Mulford, D. J., Ashworth, J. N., Melin, M. and Taylor, H. L. (1946). Preparation and properties of serum and plasma proteins IV. A system for the separation into fractions of the proteins and

lipoprotein components of biological tissues and fluids. *Journal of the American Chemical Society*, **68**: 459-475.

Curling J. M. (1983). Current practice and future possibilities in plasma protein fractionation. Uppsala: Pharmacia Fine Chemicals AB.

Danilov, A. V., Vagenina, I. V., Mustaeva, L. G., Moshnikov, S. A., Gorbunova, E. V., Cherskii, V. V. and Baru, M. B. (1997). Liquid chromatography on soft packing material, under axial compression: Size exclusion chromatography of polypeptides. *Journal of Chromatography A*, **773**: 103-114.

Datar, R. (1986). Economics of primary separation steps in relation to fermentation and genetic engineering. *Process Biochemistry*, **21**: 19-26.

De Jonge, E., Van Leeuwen, M. A. W., Radema, H., Linssen P. H. J. M. and Over, J. (1993). Filtration processes in the Cohn fractionation process. *Biotechnology of Blood Proteins*, **227**: 49-54.

Dvorin, J. (2001). Large Molecules: Too Late For Big Pharma? *In Vivo: The Business and Medicine Report*, December 2001, 53-58.

Evans, L. B. (1988). Bioprocess simulation: A new tool for process development. *Bio/Technology*, **6**: 200-203.

Foster, P. R. (1994). Protein precipitation. Engineering processes for Bioseparations. Butterworth-Heinmann, Oxford.

Friedli, H., Mauerhoffer, M., Faes, A. and Kistler, P. (1976). Studies on new procedures in plasma fractionation on an industrial scale: Removal of precipitates through a horizontal leaf filter with centrifugal cleaning as an alternative to tubular-bowl centrifuges. *Vox Sang*, **31**: 289-295.

Friedli, H., Fournier, E., Volk, T., Ingold, W., Meier, R and Kistler, P. (1976). Studies on new procedures in plasma fractionation on an industrial scale: Determination

of the pH value of ethanol containing protein solutions at low temperatures. *Vox Sang*, **31**: 277-282.

Garcia, F. A. P., Prazeres, D. M. F. and Cabral, J. M. S. (2002). Industrial biochromatography: engineering aspects. Biochromatography Theory and Practice. Taylor and Francis, London.

Gariepy, P. A., Moore, S. R. and Sofer, G. (1998). Multiuse chromatography system: A clean in place study. *BioPharm*, September 1998, 32-36.

Gebauer, K. H., Luo, X. L., Barton, N. G. and Stokes, A. N. (2003). Efficiency of preparative and process column distribution systems. *Journal of Chromatography A,* **1006**: 45-60.

Giddings, J. C. (1965). Dynamics of chromatography. Part I: Principles and theory. Chromatographic Science Series. Marcel Dekker, New York.

Gosele, W. (1993). Precipitation to obtain easily separated solids. *Chemical Engineering Science*, **48**:295-303.

Grandgeorge, M., Vernon, J. L. and Gueille, G. (1989). Separation of protein precipitates by tangential flow microfiltration during purification of human albumin. *Biotechnology of Plasma Proteins: Fractionation and Application*, **175**: 25-32.

Groep, M. E., Gregory, M. E., Kershenbaum, L. S. and. Bogle, I. D. L. (2000). Performance modelling and simulation of biochemical process sequences with interacting unit operations. *Biotechnology and Bioengineering*, **67**: 30-311.

Guiochon, G., Farkas, T., Guan-Sajonz. H., Koh, J., Sarker, M., Stanley, B. and Yun, T. (1997). Consolidation of particle beds and packing of chromatographic columns. *Journal of Chromatography A*, **762**: 83-88.

Hansen, E (2003). Chromatographic scale-up on a volume basis. Presented at the EPSRC sponsored IMRC meeting, University College London, 8th April 2003.

Hao, Y. (1985). Pilot-plant scale preparation of human serum albumin. *Vox Sang*, **49**: 1-8.

Hetherington, P. J., Follows, M., Dunnill, P. and Lilly, M. D. (1971). Release of protein from bakers yeast (*Saccharomyces Cerevisiae*) by disruption in an industrial homogeniser. *Transactions of the Institute of Chemical Engineers*, **49**: 142-148.

Heuer, C., Kniep, H., Falk, T. and Seidel-Morgenstern, A. (1998). Comparison of various process engineering concepts of preparative chromatography. *Chemical Engineering Technology*, **21**: 469-477.

Hoare, M. and Titchener-Hooker, N. (2001). MBI lecture notes, University College London.

Holdich, R. G., Tarleton, E. S. and Shaw, F. J. (1993). A new experimental technique for the analysis of cake filtration. Filtech 93 Conference, Karlsruhe, Germany.

Humphrey, A. (1998). Shake flask to fermenter: What have we learned? *Biotechnology Progress*, **14**: 3-7.

Johnston, A., Davies, J. R., Turner, P. J. and Wilkie, B. J. (1997). Filtration of plasma mixtures using cellulose-based filter aids. International patent number: WO 97/32654.

Kaltenbrunner, O., Jungbauer, A. and Yamamoto, S. (1997). Prediction of the preparative chromatography performance with a very small column. *Journal of Chromatography A*, **760**: 41-53.

Karnik, S. S., Kamath, M. V. and Ray, V. (2001). Manufacturing of safe blood products using biotechnology. Available from: http://www.exicom.org/bio_med/novdec96/blood.htm

Kistler, P. and Nitschmann, H. S. (1962). Large-scale production of human plasma fractions. *Vox Sang*, **7**: 414-424.

Lebing, W. R., Lee, D. C., Davies, A.O., Nixon, C. E. and Paul, H. I. (1999). Development and scale-up of a production-scale chromatographic process for the production of human IgG. *Downstream,* **31**: 19.

Lienqueo, M. E., Leser, E. W. and Asenjo, J. A. (1996). An expert system for the selection and synthesis of multistep protein separation processes. *Computers and Chemical Engineering*, **20**: 189-194.

Lode, F. G., Rosenfeld, A., Yuan Q. S., Root, T. W. and Lightfoot, E. N. (1998). Refining the scale-up of chromatographic separations. *Journal of Chromatography A,* **796**: 3-14.

Lutsch, C. (2002). Biochromatography validation. Biochromatography Theory and Practice. Taylor and Francis, London.

Martin, A. J. P. and Synge, R. L. H. (1941). TITLE??? *Biochemical Journal*, **35**: 1358-1368.

Maybury, J. P., Hoare, M. and Dunnill, P. (2000). The use of laboratory centrifugation studies to predict performance of industrial machines: Studies of shear-sensitive and shear-insensitive materials. *Biotechnology and Bioengineering*, **43**: 32-34.

Mayer, E. (2000). Cake filtration theory and practice. *Chemical Engineering Journal*, **80**: 233-236.

Mitra, G. and Lundblad, J. (1978). Continuous fractionation of human plasma. *Biotechnology and Bioengineering*, **20**: 1037-1044.

More, J. E. and Harvey, M. J. (1991). Purification technologies for human plasma albumin. Blood Separation and Plasma Fractionation. Wiley-Liss, Inc.

Morfeld, F., Schutz, R., Josie, D. and Schwinn, H. (1996). Manufacturing and down scaling process for a virus-inactivated pooled human plasma. *Pharmaceutical Industry*, **58**: 433-435.

Moscariello, J., Purdom, G., Coffman, J., Root, T. W. and Lightfoot, E. N. (2001). Characterizing the performance of industrial-scale columns. *Journal of Chromatography A*, **908**: 131-141.

Nakamura, K. and Matsumoto, K. (1998). Adsorption behaviour of BSA in microfiltration with porous glass membrane. *Journal of Membrane Science*, **145**: 119-128.

Nayar, R. and Manning, M. C. (2002). High-throughput biopharmaceutical drug development: Meeting the coming challenge. *Biopharm*, February 2002, 20-28.

Niktari, M., Chard, S., Richardson, P. and Hoare, M. (1990). The monitoring and control of protein purification and recovery processes. Separations for Biotechnology 2. Elsevier Applied Science, 622-634.

Oja, M. and Nystrom, L. (1995). Evaluation of filtration tests with gradual pressure increase and sedimentation. Filtech 95 Conference, Karlsruhe, Germany.

O'Leary, R. M., Feuerhelm, D., Peers, D., Xu, Y. and Blank, G. S. (2001). Determining the useful lifetime of chromatography resins. BioPharm, September 2001, 10-18.

Oosterhuis, N. M. G., Kossen, N. W. F., Olivier, A. P. C. and Shenk, E. S. (1995). Scale-down and optimisation studies of the gluconic acid fermentation by *Gluconabacter Oxydans. Biotechnology and Bioengineering*, **27**: 711-720.

Petrides, D. P., Koulouris, A. and Siletti, C. A. (2002). Throughput Analysis and Debottlenecking of Biomanufacturing Facilities. *Biopharm*, August 2002, 28-34.

Pinheiro, H. and Cabral, J. M. S. (1993). Chapter 4: Filtration. Recovery Processes for Biological Materials. John Wiley & Sons Ltd.

Pisano, G. P. (1997). The development factory. Harvard Business School Press, Boston, USA.

Pisano, G. and Wheelwright, S. C. (1995). The new logic of high-tech R&D. *Harvard Business Review*. Sept-Oct.

Purchas, D. B. and Wakeman, R. J. (1986). Filtration fundamentals. Solid/Liquid Separation Equipment Scale-up. London, Uplands press.

Rathore, A. and Velayudhan, A. (2003). Guidelines for optimisation and scale-up in preparative chromatography. *BioPharm International,* January 2003, 34-42.

Rees, R. H. (1990). Let diatomite enhance your filtration. *Chemical Engineering*, **97**: 76-79.

Reynolds, T., Boychyn, M., Sanderson. T., Bulmer, M., More, J. and Hoare, M. (2003). Scale-down of continuous filtration for rapid bioprocess design: Recovery and dewatering of protein precipitate suspensions. *Biotechnology and Bioengineering*, **83**: 454-464.

Rosendale, D. M. (2002). Practical techniques for successful validation of pharmaceutical process equipment. Available from: http://www.vectorcorporation.com/downloads/val-interphex.pdf

Rothstein, F. (1994). Differential precipitation of proteins. Protein Purification Process Engineering. Marcel-Dekker, New York.

Rushton, A., Ward, A. S. and Holdich, R. G. (1996). Filtration fundamentals. Solid-Liquid Filtration and Separation Technology. VCH, Weinheim, Germany.

Shalliker, R. A., Broyles, B. S. and Guiochon, G. (2000). Physical evidence of two wall effects in liquid chromatography. *Journal of Chromatography A*, **888**: 1-12.

Siddiqi, S. F., Titchener-Hooker, N. J. and Shamlou, P. A. (1997). High pressure disruption of yeast cells: The use of scale down operations for the prediction of protein release and cell debris size distribution. *Biotechnology and Bioengineering*, **55**: 642-649.

Sofer, G. (1996). Validation: Ensuring the accuracy of scaled-down chromatography models. *BioPharm*, October 1996, 51-54.

Sofer, G. and Hagel, L. (1997). Handbook of process chromatography. Academic press limited, London.

Sofer G. K. and Nystrom L. E. (1989). Process chromatography. A practical guide 1st edition. Academic press Ltd. London.

Sorensen, E., Teoh, K. and Titchener-Hooker, N. (2002). Modelling and optimisation of chromatographic processes. *European Pharmaceutical Review*, 13-16.

Soriano, G. A., Titchner-Hooker, N. J. and Shamlou, P. A. (1997). The effects of processing scale on the pressure drop of compressible gel supports in liquid chromatographic columns. *Bioprocess Engineering,* **17**: 115-119.

Stickel, J. J. and Fotopoulos, A. (2001). Pressure-flow relationships for packed beds of compressible chromatography media at laboratory and production scale. *Biotechnology Progress*, **17**: 744-751.

Struck, M. M. (1994). Biopharmaceutical R&D success rates and development times. *Bio/Technology*, **12**: 674-677.

Stryker, M. H., Bertolini, M. J. and Hao, Y. L. (1985). Blood fractionation: proteins. *Advances in Biotechnological Processes*, **4**: 275-336.

Sulpizio, T. E. (1999). Advances in filter aid and precoat filtration technology. Presentation at the American Filtration and Separtions Society, Annual Technical Conference, Boston.

Svarovsky, L. (1979). Advances in solid liquid separation-I: Filtration and allied operations. Chemical Engineering, July 1979, 62-76.

Taylor, G. (1953). Dispersion of soluble matter in solvent flowing slowly through a tube. *Proceedings of the Royal Society of London, Series A*, **219**: 186-203.

Tien, C. and Bai, R. (2003). An assessment of the conventional cake filtration theory. *Chemical Engineering Science*, **58**: 1323-1336.

Tiller, F. M. (1966). Solid liquid separation operations. *Chemical Engineering*, **20**: 151-162.

Van Deemter, J. J., Zuiderberg, F. J. and Klinkenberg, A. K. (1956). Longitudinal diffusion and resistance to mass transfer as causes of non-ideality in chromatography. *Chemical Engineering Science*, **5**: 271-289.

van Oss, C. J. (1989). On the mechanism of the cold ethanol precipitation method of plasma protein fractionation. *Journal of Protein Chemistry*, **8**: 661-668.

van Reis, R., Goodrich, E. M., Yson, C. L., Frautschy, L. N., Dzengeleski, S. and Lutz, H. (1997). Linear scale ultrafiltration. *Biotechnology and Bioengineering*, **55**: 737-746.

Watt, J. G. (1970). Automatically controlled continuous recovery of plasma protein fractions for clinical use. A preliminary report, *Vox Sang*, **23**: 268-271.

Wechsler, J. (2003). Reshaping biotech drug development. *Biopharm International*, June 2003, 14-16.

Wheelwright, S. M. (1987). Designing downstream processes for large-scale protein purification. *Bio/Technology*, **5**:790-793.

Wolthuis, R. and Dichiaria, V. C. (1997). Conventional filtration. Handbook of Downstream Processing. Blackie academic and professional, London.

Yap. H. B., Young, I. F., Micucci, V., Herrington, R. W., Turner, P. J. and Davies J. R. (1993). Development of a process for the preparation of human serum albumin using chromatographic methods. *Biotechnology of Blood Proteins,* **227**: 143-149.

Zhou, Y. H. and Titchener-Hooker, N. J. (1999). Visualizing integrated bioprocess designs through "Windows of operation". *Biotechnology and Bioengineering*, **65**: 550-557.

Zlokarnik, M. (1983). Model scale-up in chemical engineering. *Chemical Engineering Technology*. **55**: 362-372.