CHARACTERIZATION OF AN AFFINITY PURIFICATION SYSTEM FOR RECOMBINANT PROTEINS CONTAINING A CELLULOSE BINDING DOMAIN

by

BEATRIZ RODRIGUEZ

B.Sc., Advanced Polytechnical Institute, Cuba, 1984

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUEST FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

The Biotechnology Laboratory
Department of Chemical and Bio-Resource Engineering

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

November 1997

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Chemical and Bio-Resource Engineering

The University of British Columbia
Vancouver, Canada

Date November 12/97

DE-6 (2/88)

ABSTRACT

The selective binding of the cellulose binding domain (CBD) of the exo-β-1,4-glycanase Cex from *Cellulomonas fimi* (CBD_{Cex}) to a variety of cellulosic substrates offers the possibility of a new, cost effective, highly flexible affinity chromatography system for purification of recombinant proteins. Genetic linkage of CBD_{Cex} to a target protein results in a fusion protein which binds strongly to cellulose and retains the biological activity of the fusion partner. This thesis is concerned with development, testing, and modeling of an affinity chromatography column based on a sorbent resin of Avicel PH101, a pure semicrystalline cellulose particulate, for one-step purification of recombinant fusion proteins containing a CBD_{Cex} tag. Appropriate mobile-phase compositions for the loading, wash, and elution steps are described, as are the general requirements for column set-up. Limitations of the current column configuration and components are discussed along with suggestions for improving the system.

Frontal-loading and first-moment analyses are used to characterize basic column properties, including voidage and packing uniformity. Second moment analysis is combined with breakthrough data as a function of flow rate and concentration to determine axial diffusion coefficients, film/particle mass transfer coefficients, and solute diffusivities. The pore-size distribution and effective porosity of Avicel PH101 particles are determined by a solute exclusion technique using probe molecules of varying hydrodynamic radius. Batch adsorption-isotherm data are used to estimate binding constants for CBD_{Cex} and maximum sorbent surface capacity, which are then confirmed on column by breakthrough curve analysis.

The measured column parameters are combined with the equations governing continuity and intraparticle diffusion to develop a column model which accounts for gradients both along the column axis and in the radial direction within the sorbent particles. Flow and solute concentrations are assumed to be uniform across any cross-section of the column. Breakthrough curves predicted by the model are in good agreement with experiment over the range of realistic operating conditions.

Column performance and model predictions are verified by their application to the purification of a Protein A-CBD_{Cex} fusion and the comparison of predicted breakthrough curves with those observed experimentally.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENTS

I would like to thank to my supervisors; Dr. Charles A. Haynes for his great suggestions, for his incredible teaching abilities and for leading me into the amazing world of protein purification, which I love now, and to Dr. Douglas Kilburn for giving me the great opportunity to become a member of his group, for his advices and knowledge of life. To both, thank you for your friendship.

I would also like to thank to all of the members of the Cellulase Group and the Biotechnology Laboratory; professors, staff and students. To the Cellulase Group, for keeping such a sharing group. To the Biotechnology Laboratory, it is great to be part of such a friendly atmosphere.

I would like to express special thanks to Jürgen Koska for his very helpful work.

I would also like to thank Eduardo for his unlimited support.

To my Parents

Son las ideas ,que ascienden,
Rotas sus cárceles!

The ideas can rise
when their cage is broken!

José Martí (Cuba, 1853-1895)

INTRODUCTION

Major advances have occurred in recent years in the development of preparative and analytical techniques for the separation of peptides, proteins, and nucleic acid components (*e.g.* nucleotides). Due to this and the increasing opportunity to produce complex proteins at relatively low cost through the advent of genetic engineering techniques, the criteria for purity and biorecovery throughput in large-scale purification strategies are currently undergoing substantial reexamination.

Central to these advances is the need to develop new stratagems for the purification of specific proteins from complex fermentation broths which maximize yield and purity while minimizing costs. These objectives may be met by streamlining downstream operations into a small number of elements, each highly selective for the target protein. One such operation is the bioselective adsorption process commonly known as affinity chromatography. Affinity chromatography is based on the high degree of binding selectivity exhibited by biological molecules. For example, signaling in the human biosystem often involves highly specific and tight ($K_a \sim 10^9\ M^{-1}$) binding reactions between factor and receptor pairs, while immune responses involve binding ($K_a \sim 10^6$ to $10^9\ M^{-1}$) a specific antibody to the foreign antigen. A biospecific adsorbent can therefore be created by successfully immobilizing one member of this binding pair to a solid support without loss of activity. Due to the high affinity and exquisite selectivities of biological binding processes, affinity chromatography columns can be used to efficiently recover materials from extremely dilute and complex solutions.

One disadvantage of affinity chromatography is the cost and complexity of preparing and maintaining the affinity resin. Affinity-ligand immobilization is often expensive and difficult,

and there is no certainty that the ligand will remain active after immobilization. In operation, immobilized ligands are also susceptible to fouling and structural damage, particularly at the ligand-resin junction; this later sensitivity often leads to problems in column sterilization and capacity.

These problems can be minimized by designing affinity chromatography systems such that the target protein has selective affinity for an inexpensive and easily prepared ligand matrix. In this thesis, we develop, test, and model a generic affinity chromatography column based on production of the target protein as a genetic fusion to a polypeptide domain which has a strong and selective affinity for insoluble cellulose. Inexpensive and commercially available resins composed of semicrystalline cellulose can then be used to affinity purify any protein which has been tagged with a cellulose binding domain.

This thesis is divided into five chapters and one appendix. Chapter 1 provides a brief overview of the system and pertinent literature. Chapter 2 explains the materials and general methods used. Chapter 3 describes the chromatographic system in detail and reports data characterizing column performance, including the determination of all model parameters. The equations governing continuity and intraparticle diffusion are derived and used to develop a model capable of predicting column performance. Chapter 4 reports column and model performance for the affinity purification of a recombinant fusion protein, staphylococcal protein A-CBD_{Cex}, from an *E. coli* strain JM101 culture. Conclusions and recommendations from this work are given in Chapter 5.

CHAPTER 1

LITERATURE REVIEW

1.1 Downstream Processing in Biotechnology

Due to the continuing progress in recombinant DNA technology, the number of products obtained from biological organisms has been increasing steadily. A wide range of therapeutic proteins can now be obtained from fermentation and cell-culture processes. Depending on the microorganism and recombinant vector used, the desired substance may be excreted in the culture medium, directed in soluble form to the cytoplasm or periplasm, or targeted to inclusion bodies in a denatured form. For recombinant protein products, initial concentrations from culture are dilute, varying from several g L^{-1} to a few μg L^{-1}, as in the case of blood factors and therapeutic enzymes (Freitag and Horvath, 1995). Moreover, concentrations and number of contaminants are often high. Therefore, several isolation and purification steps are generally required for product recovery.

Downstream processing often dominates the cost of a recombinant-protein production process, in most cases representing 50-80 % of the total operating cost (Jones, 1991). In general, production costs are directly proportional to the difference between the concentration and purity of the starting material and the desired quality of the final product. The cost of purifying ethanol, which can produced to concentrations of several hundred g L^{-1} in yeast fermentations, is relatively low. Nevertheless, it is prohibitively expensive (ca. $ 0.30 per L) compared with cost of producing ethanol from natural gas (ca. $0.16 per L). Conventional downstream processing of dilute protein and peptide products, such as blood factors or therapeutic enzymes, typically requires a larger number of processing steps which significantly increase cost and decrease yield.

A typical recombinant-protein recovery process can be divided into four steps. The target protein is first separated from the producing organism and other insolubles by a solid/liquid separation step, such as centrifugation or filtration. The solution may then be concentrated by removing excess water through a second filtration, a two-phase extraction, or a precipitation induced by a chaotrophic agent such as ammonium sulfate. In many cases, substances that differ considerably from the product in their physico-chemical character are removed during this concentration process. In the third step, the target protein is separated from other substances, mainly proteins, having similar physical or functional properties. Chromatography is the primary unit operation used in this step. The final step is polishing and commercial formulation of the product and may include crystallization or lyophilization.

Chromatographic methods used by industry for recombinant protein purification include hydrophobic interaction, size exclusion and, most often, ion exchange. Purification usually requires several of these chromatographic separations in series due to the relatively non-specific physico-chemical interactions involved in the separation. Although product purities are often quite high, yields from these multi-step chromatographic processes are low due to cumulation of losses in each step. Therefore, there is a strong economic motivation for identifying purification strategies which increase yields by reducing the complexity and number of processing steps.

Affinity chromatography, based on the high degree of binding selectivity exhibited by biological molecules, is recognized as one of the most attractive approaches for simplifying purification. The replacement of a multi-step process by a single affinity step can greatly reduce production costs, provided the expense and lifetime of the affinity resin are reasonable. As a result, commercial applications of affinity chromatography are becoming more common, especially in the plasma fractionation industry (*e.g.*, antithrombin III is purified on immobilized

heparin, plasminogen is purified on lysine-sepharose), (Burnouf-Radosevich and Burnouf, 1988), and in the production of high-value research enzymes and protein therapeutics. The American Red Cross has developed an immunoaffinity chromatography method to purify human coagulation Factor IX for therapeutic treatment of hemophilia B, in which anti-Factor IX monoclonal antibody is immobilized onto activated Sepharose (Highsmith *et al.*, 1992). Hoffmann La Roche's α-interferon is produced by *E. coli* bacteria and purified after cell separation on an agarose affinity column (Staehelin, *et al.*, 1981). Du Pont uses an affinity dye, Cibracron blue, modified agarose to purify β-Interferon produced by human fibroblast cells (Knight and Fahey, 1981). Merck, Sharp and Dohme use recombinant yeast cells to produce Hepatitis B vaccine. The cells are harvested and the surface antigen is then isolated on an affinity IgG-column (McAleer, *et al.*, 1984).

1.2 Affinity Chromatography in Downstream Processing

Affinity separations are attractive because they can potentially replace difficult and expensive multi-step procedures with a single adsorption step that simultaneously purifies and concentrates the product. If an appropriate high-affinity ligand is chosen, affinity chromatography can be effective in purifying a very dilute product or in removing a very dilute contaminant. Many of the first affinity chromatography systems exploited the biospecific interaction between immobilized antibodies and protein antigens (Desai, 1990). The library of affinity ligands in use today includes sugars, synthetic dyes, chelating agents, proteins and protein fragments, and amino acids (Lowe *et al.*, 1992; Kastner and Neubert, 1991; Moks *et al.*, 1987; Buettner, *et al.*, 1996; Hopp *et al.*, 1988). Research applications of affinity chromatography are too numerous to summarize. Hundreds of proteins have been affinity

purified on the laboratory scale, and more appear each month in the literature (*e. g.*; Hughes *et al.*, 1997; Makriyannis and Clonis, 1997; Yaron *et al.*, 1996).

Like all adsorption-based separations, affinity chromatography can be divided into four basic processing steps: loading, wash, elution, and regeneration. In the loading step, the complex sample solution containing the target molecule is passed over the affinity adsorbent. The ligand specifically binds the target while all contaminating proteins, culture byproducts, and cell debris that do not interact with the adsorbent pass through unretarded. When the ligand sites are saturated, loading is stopped and a wash buffer is passed through the column until the void is cleared of contaminants. A second, harsher wash solution may follow to ensure removal of all nonspecifically (weakly) bound contaminants. The elution buffer is then introduced which allows the target to be recovered by disrupting the binding complex. When the column is free of product and contaminants, the loading buffer is reintroduced to reactivate the resin for binding.

Many issues must be addressed when selecting an appropriate affinity ligand. When immobilized, the ligand must retain activity and selectively bind the target protein with sufficient affinity to allow the impurities to be washed from the adsorbent without loss of product. It must also, however, release the protein under appropriate conditions. Very strong binding may require elution with denaturing agents which require the recovered product to be refolded, often a difficult task. Ideally, the ligands should be inexpensive, commercially available, and easily immobilized on suitable matrices. Sanitization and lifetime of the affinity media must also be considered, particularly for processes involving pharmaceutical products (Sofer, 1995; Kilburn *et al.*, 1993). Meeting all of these criteria is difficult.

Polypeptide-based ligands are relatively expensive since they too must be purified, and yields in immobilization reactions are often low (Chase, 1984). In certain cases, costs can be

reduced by using relatively simple ligands, including sugars and amino acids, and textile affinity-dyes (Lowe *et al.*, 1992; Jones, 1991) such as reactive blue. However, close attention must be paid to the synthesis and the characterization of the resin preparation since improper immobilization can reduce (or abolish) binding affinity to the point where the complex is insufficiently stable to withstand adsorption and washing operations (Marchand, 1994). There is also the tendency for these simple ligands to have affinity towards more than one biochemical compound, causing contaminants to copurify with the target product (Desai, 1990).

The selection of an appropriate solid support is also important. Beyond allowing for attachment of the ligand, the support must provide a large contact area with the mobile phase, must promote uniform flow characteristics in the column, and must not be prone to fouling or compression. A number of support materials are in common use, including beaded agarose, silica, alumina, and glass; a number of synthetic supports, particularly polyacrylamide and sepharose/sephadex beads (Hermanson *et al.*, 1992), are also in use. None of these materials is ideal. For example, agarose, possibly the most widely used support, is susceptible to microbial degradation. Moreover, many natural supports, including agarose, offer relatively few reactive sites for ligand immobilization. As a result, binding capacity is often low. The protein G-agarose column from Perstorp Biolytica, for instance, binds no more than 12 mg human IgG per ml gel, (Narayanan and Crane, 1990). Silica, alumina, and glass also suffer from nonspecific adsorption due to residual charge on the surface, and from solubility problems at alkaline pH. Polymer-based resins, such as polyacrylamide, are resistant to microbial attack and have relatively good pH stability. Their use, however, has been limited by poor mechanical stability, slow flow rates, and a tendency to shrink or swell with changes in mobile-phase composition (Marchand, 1994).

In most instances, the surface of these matrices must be chemically modified to provide functional groups for covalent coupling to the ligand (Hermanson *et* al., 1992). An example of this procedure is the introduction of cyanate esters and imidocarbonates groups on sepharose, activated by Cyanogen Bromide (CNBr), which serve as reactive sites for immobilization of the egg-white protein avidin or the related bacterial protein streptavidin (Wilchek and Bayer, 1990). Both avidin and streptavidin have a strong affinity for biotin (vitamin H). Proteins, *e.g.* antibodies, may be artificially modified with biotin through a second conjugation reaction. The biotinylated antibody is then able to bind to avidin or streptavidin through its biotin "handle" or affinity tag.

A ligand that is attached directly to a polymeric support material may not protrude far enough from the matrix surface to reach the binding site on an approaching protein molecule, resulting in a weakened interaction or no binding at all (Hermanson *et al.*, 1992). Thus, spacer arms are often introduced between the support material and the affinity ligand. These spacers usually consist of linear polar organic chains such as poly(ethylene glycol) or polyacrylamide with functionalities, *e.g.* amine groups, on both ends for easy coupling to the support and ligand.

The design and scale-up of affinity chromatography processes for the concentration and purification of biotechnological products is of prime industrial interest. Beyond those cited above, concerns raised by industry include (1) the high costs of the affinity media, matrix activation and ligand attachment, (2) the lack of appropriate ligands for all proteins of interest, (3) the difficulties of making complex affinity media operationally stable under conventional depyrogenating conditions, and (4) patent protection.

The second of these concern has led to the development of a more generic class of affinity purification systems. Most such systems are based on genetically fusing a peptide to the amino or carboxy terminus of the target protein which has specific affinity for a generic affinity resin.

1.3 Generic Affinity Chromatography Using Fusion Tails as Affinity Tags

The power of individual biospecific and biochemical interactions can be more generally applied to the concentration and purification of a wide spectrum of target proteins by genetically engineering the target protein to contain a fusion tail with specific affinity for a common affinity matrix (Ford *et al.*, 1991; Nygren *et al.*, 1994). Fusion tails have been designed to facilitate recovery of intracellular and extracellular proteins from a variety of cell types and recombinant expression systems. Specific proteolytic cleavage sites can be genetically engineered into fusion tail systems for removal of the tail from the target protein after purification (Sherwood, 1991; Nilsson *et al.*, 1992). Table 1.1 provides a partial, representative listing of affinity tag systems that have been reported in the literature (Jones *et al.*, 1995). Many of these systems have not been applied on the commercial or large scale.

Immobilized metal affinity chromatography (IMAC), based on the formation of a ternary complex centered at a divalent cation, has found widespread application on the laboratory scale and limited use on the industrial scale (Freitag and Horvath, 1995). The stationary phase contains imminodiacetate groups or similar ligands (*e.g.*, nitrolo-triacetic acid; the basis for the Hoffman-La Roche IMAC system) which are capable of chelating divalent metal ions such as Zn^{2+} and Ni^{2+}. Affinity separation is then achieved by selective interaction of a fusion tail of electron-rich amino acids, particularly polyhistidine, with the metal chelate. The extremely high affinity interaction between a 6xHis tag and the Ni^{2+} - nitrolo-triacetic acid resin permits the one-

step purification of tagged proteins from less than 1 % to greater than 95 % homogeneity (Janknecht *et al.*, 1991).

Table 1.1: Fusion Tail Systems for Affinity Purification

PURIFICATION TAG	SEPARATION METHOD	REFERENCES
1. Enzymes		
β-galactosidase	Immobilized TPEG[a]	Steers et al., 1971
Glutathione-S-Transferase (GST)	Immobilized Glutathione	Smith and Johnson, 1988
Chloramphenicol acetyltransferase (CAT)	Immobilized *p*-Aminochloramphenicol	Dykes et al., 1988
2. Polypeptide-binding proteins		
Staphylococcal protein A	Immobilized IgG	Moks et al., 1987
Streptococcal protein G	Immobilized albumin	Hammarberg et al., 1989
3. Carbohydrate binding domain		
Maltose binding protein (MBP)	Crosslinked amylose	Blondel and Bedouelle, 1990
Starch binding domain (SBD)	Starch	Ford et al., 1991
Cellulose binding domain (CBD)	Cellulose	Ong et al, 1989
4. Biotin binding domain	Immobilized avidin	Cronan, 1990
5. Antigenic epitopes		
Flag peptide[b]	Immobilized anti-Flag antibody	Hopp et al., 1988
6. Poly(His) tails	Immobilized metal affinity chromatography (IMAC)	Ljungquist et al., 1989

[a] TPEG; *p*-amino-phenyl- β-D-thiogalactoside
[b] Eight amino acid tail; Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys

Genetic-engineering technologies have been used to generate combinatorial affinity-peptide libraries (Buettner, *et al.,* 1996; Huang *et al.*, 1995) which can be screened for binding to any target molecule. However, the generation, screening, and immobilization of these libraries are often prohibitively expensive for application to a single protein product. As shown in Table 1.1, the few commercially available systems are therefore based either on high-affinity polypeptide tags such as a polypeptide-binding protein (*e.g.* Protein A) with affinity to immunoglobulin G (Moks *et al.*, 1987), an antigenic epitope with affinity for immobilized monoclonal antibodies (Field *et al.*, 1988), and the glutathione-S-transferase (GST) domain which has high affinity for glutathione (Hartman *et al.*, 1992), or on *in vivo* biotinylation by *Escherichia coli* of a specific sequence fused to the target protein which allows binding to immobilized avidin or streptavidin (Cronan, 1990).

Although they are in many ways superior to older affinity chromatography technology, these systems are not without limitations. All require either covalent or physical attachment of an often complex and expensive affinity ligand (*e.g.*, avidin) to an inert solid support matrix, which, as described above, tends to limit their applicability to large-scale separations.

1.4 Cellulose Binding Domains and Cellulose

Many bacterial and fungal enzymes that function to hydrolyze cellulose or other β-1,4-glucans contain a discrete domain, called a cellulose binding domain, that functions to bind the enzyme to the target polysaccharide (Warren, 1996)

As discussed by Tomme and coworkers, (1995a), cellulose-binding domains can be classified into 13 families. Several CBDs in families I, II, III, IV and VI have been explicitly

shown to have affinity for cellulose, while others are included solely on the basis of primary sequence comparison.

Family I CBDs are all fungal enzymes such as cellulases from *Trichoderma reesei.* They are 33-36 amino acids long, are usually C-terminal, and all contain at least three conserved aromatic residues which are believed to be important in the interaction with cellulose.

Family II includes representatives from *Streptomyces lividans, Cellulomonas fimi, Pseudomonas fluorescens* and *Thermomonospora fusca*; there are no fungal representatives. Family II CBDs are about 100 residues long, and contain three strictly conserved tryptophan residues, which are known to participate in binding (Creagh *et al.*, 1996; Din *et al.*, 1994).

Family III CBDs are prevalent in enzymes from *Bacillus* spp., *Clostridium* spp. and *Caldocellum saccharolyticum.* They contain 130-170 amino acid residues, many located internally. Family IV comprises a small group of CBDs, distantly related to family II, which have binding specificity for soluble β-1,4-glucans.

Cellulose binding domains (CBDs) bind tightly to a wide variety of cellulose-based (and chitin-based) chromatographic support matrices (Warren, 1996; Tomme *et al.*, 1995a; Tomme *et al.*, 1995b). Depending on the origin of the CBD, binding is specific for either crystalline (Creagh *et al.*, 1996) or amorphous (Tomme *et al.*, 1995b) regions of cellulose, and does not require any chemical modification of the basic insoluble cellulose matrix. Cellulose matrices for chromatographic use are available in a wide variety of forms, including low and high porosity particles (Gemeiner *et al.*, 1993), and high-throughput spiral wound membranes and fabrics which remain incompressible at high pressures (Yang *et al.*, 1992). All are inexpensive and can withstand stringent sterilization protocols.

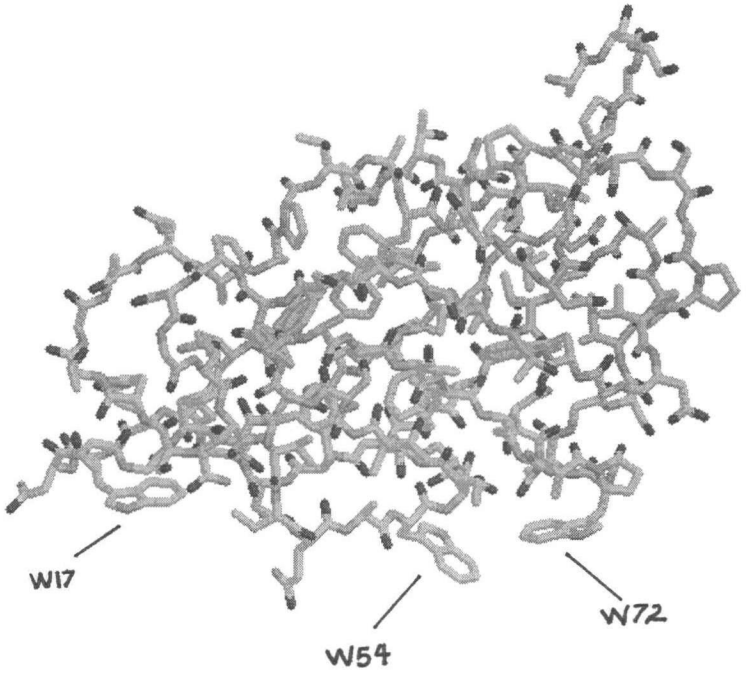


Figure 1.1: Representation of the solution structure of the C-terminal CBD_{Cex} from the exo-β-1,4-glycanase Cex of *Cellulomonas fimi* showing the three tryptophan residues (W17, W54, and W72) defining the putative binding surface.

Fusion proteins containing a CBD retain the biological activity of the fusion partner and bind tightly to cellulose (Greenwood *et al.*, 1992; Ong *et al.*, 1989). For example, CBDs have been linked to a number of heterologous proteins including *Agrobacterium* β-glucosidase (Ong *et al.*, 1989), *E. coli* alkaline phosphatase (Greenwood *et al.*, 1992), *Staphylococcus* protein A (Ramirez *et al.*, 1993), and *Streptomyces* streptavidin (Le *et al.*, 1994). CBD-tagged proteins have been expressed in a number of different recombinant hosts, including *Escherichia coli* (Le

et al., 1994), the yeast *Pichia pastoris* (Boraston, 1997), and mammalian cells (Ong *et al.*, 1995). For instance, recombinant human interleukin 2 has been produced in mammalian COS cells and affinity purified using an N-terminal CBD tag (Ong *et al.*, 1995). Of particular importance in this study is CBD_{Cex}, the C-terminal cellulose binding domain of the exo-β-1,4-glycanase Cex from the soil bacterium *Cellulomonas fimi.* CBD_{Cex} binds both crystalline and amorphous forms of insoluble cellulose. Binding of CBD_{Cex} to cellulose is mediated by a relatively hydrophobic solvent-exposed ridge (Figure 1.1) containing three linearly aligned tryptophan residues (Xu *et al.*, 1995). From calorimetry studies, specific high-affinity binding to crystalline cellulose is driven by a relatively large increase in entropy resulting from significant dehydration of the protein-cellulose contact surface (Creagh *et al.*, 1996).

Vectors for high-level expression of gene fusions containing a CBD_{Cex} have recently been constructed (Graham *et al.*, 1995; Hassenwinkle *et al.,* 1996). The gene products are produced in soluble form and directed either to the periplasm (Hassenwinkle *et al.*, 1996) or culture supernatant (Boraston, 1997), depending on the host organism and leader sequence. Soluble recombinant-protein production levels as high as 9 g L^{-1} have been attained in *E. coli*, and for more complex fusion partners such as stem-cell factor and human interleukin-2, high-level production has been achieved in the yeast *Pichia pastoris* (Boraston, 1997) and mammalian COS cells (Ong *et al.*, 1995), respectively. Once purified, the target fusion partner can be separated from the CBD tag by chemical cleavage (*e.g.*, by CNBr at a methionine residue) or sequence-specific proteolytic cleavage at a Factor Xa site (Ong *et al.*, 1995).

1.5 Modeling Affinity Chromatography

Because both the feed and ligand-bearing adsorbent are usually costly in affinity separations, it is important to be able to predict the performance of commercial-scale systems reliably based on data from laboratory-scale experiments. Chromatography models aim to describe the migration of a protein band or zone through a column and the resulting time and flow dependent protein distribution, or chromatogram, exiting the column. Figure 1.2 shows the typical features of a chromatogram for purification by affinity chromatography of a target protein from concentrated culture supernatant. The regions of the chromatogram associated with the loading, wash, and elution stages are identified.

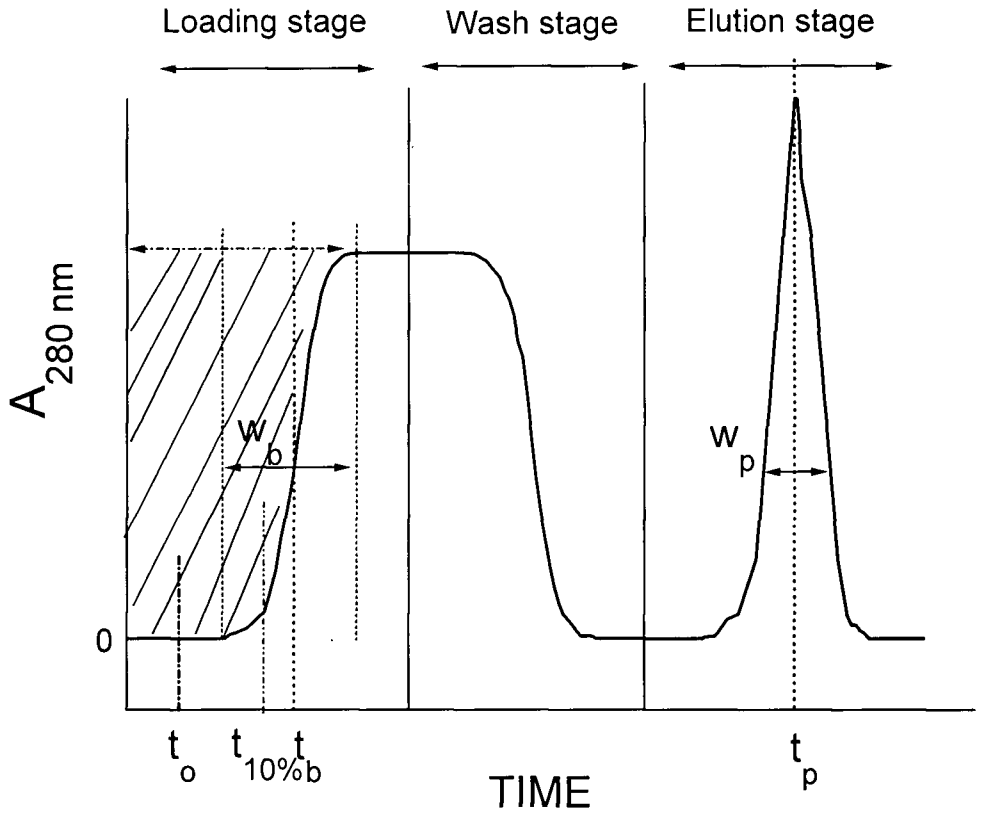


Figure 1.2: Schematic of a chromatogram showing exiting protein and contaminants concentration from an affinity column.

In this thesis, I will only address modeling the loading stage, which is characterized in the chromatogram by the breakthrough curve. As shown in Figure 1.2, t_b is the characteristic time of solute breakthrough assuming equilibrium is reached instantaneously through the column at all times (*i.e.*, all mass transfer and kinetic steps are very fast). Under these ideal conditions, the breakthrough curve is a step function such that there is an instantaneous jump in effluent concentration to the feed concentration at the moment the column capacity is reached. The spreading observed in the breakthrough curve shown in Figure 1.2 is due to finite mass transfer and binding rates, and flow (channeling) and mixing (axial dispersion and dead spaces) nonidealities. The width of the breakthrough curve is denoted as w_b. The capacity of the column for a given mass flowrate of the target protein is given by the area behind the breakthrough curve (shown in cross-hatch in Figure 1.2). The amount of product lost in the effluent is given by the area under breakthrough curve.

Often it is desirable to terminate loading at some time $t < t_b$ in order to avoid loss of valuable product. For example, termination of loading at time $t_{10\%}$, where 10% indicates an absorbance reading (280 nm) which is 10% of the reading at complete breakthrough, results in a very small loss of product. Termination of loading at time t_0, where Qt_0 is the void volume V_0 of the column, results in no loss of product. In both cases, a fraction of the column remains unused (see Figure 1.3).

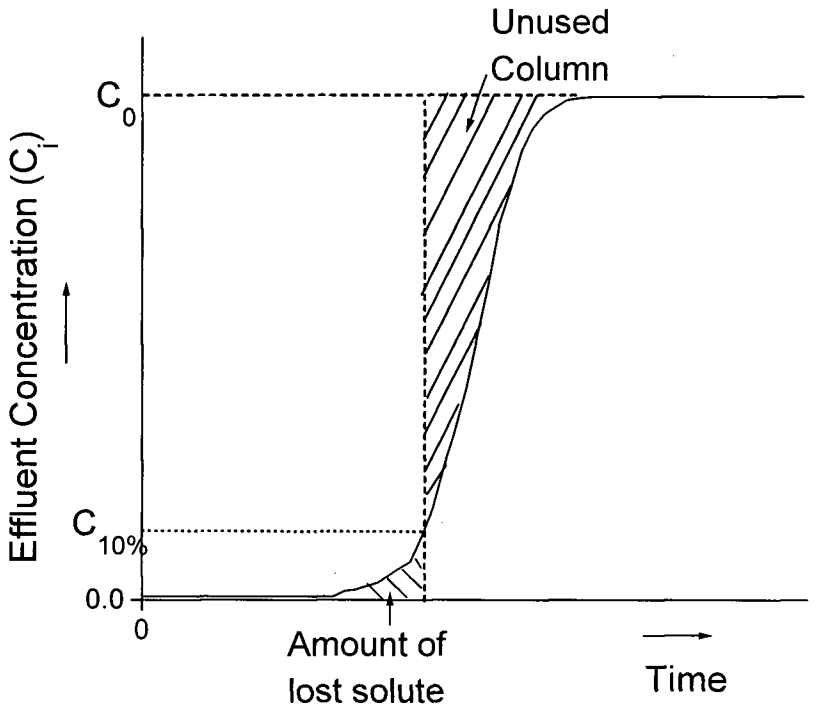


Figure 1.3: Termination of loading at 10% of the feeding concentration.

For a cylindrical chromatography column where flow is in the axial direction z, the concentration of target protein *i* at any time or position within the column is given by the continuity (mass balance) equation (see Bird, Stewart, and Lightfoot, Transport Phenomena, Wiley Press (1960)):

$$\frac{u}{\varepsilon}\nabla c_i + \frac{\partial c_i}{\partial t} + \frac{(1-\varepsilon)}{\varepsilon}\frac{\partial s}{\partial t} = D_l \nabla.\nabla c_i \qquad (1.1)$$

where u is the superficial fluid velocity (m s^{-1}), ε is the column void fraction, c_i is the concentration (mol m^{-3}) of *i* in the mobile phase, *s* is the average concentration of solute in or on the sorbent particle (mol m^{-3}), and D_l is the axial dispersion coefficient (m^2 s^{-1}). Solution of Eq. 1.1 for c_i (z, r, θ, t) requires one boundary condition in time, which is usually given by c_i (z, r, θ, t=0) = 0, and two boundary conditions in each directional coordinate. It also requires

knowledge of solute uptake $\partial s(z, r, \theta, t)/ \partial t$, and therefore an appropriate kinetic model for adsorption if binding is the rate-limiting step, or an appropriate isotherm model if the system is mass-transfer limited.

Simple affinity adsorption systems can be described by the classic Langmuir or hyperbolic-type isotherm

$$q_i = \frac{Q_{max} K_L c_i}{1 + K_L c_i} \tag{1.2}$$

where q_i is the equilibrium concentration of bound solute *i* (mol m^{-3}), Q_{max} (mol m^{-3}) is the capacity of the sorbent to bind *i*, K_L is the Langmuir-type equilibrium association constant (M^{-1}), and c_i is the free solute concentration (M). Often, however, more complex isotherms models are required to account for effects such as surface diffusion of sorbate molecules (Ma *et al.*, 1996), or non-uniform packing geometries or excluded surface effects (Jin *et al.*, 1994; Gilkes *et al.*, 1992). An objective of this thesis will therefore be to identify an appropriate binding model for binding of CBD_{cex} to crystalline cellulose.

The inverse of the separation factor α, often denoted as R_{eq} (Arnold *et al.*, 1985a), is defined as

$$R_{eq} = \frac{1}{\alpha} = \frac{1}{1 + K_L c_o} \tag{1.3}$$

where c_o is the feed concentration of the target solute. The value of R_{eq}, which varies between 1 and 0, provides an indication of the nature of the binding interaction. Linear, reversible binding processes are defined by R_{eq} values near 1. For irreversible binding processes, R_{eq} is equal to 0. The very high-affinity binding interactions required for successful affinity purification systems are therefore characterized by R_{eq} values near zero.

As shown in more detail in Chapter 2, CBD_{Cex} binding to microcrystalline cellulose is characterized by a binding constant K_L near 10^7 M^{-1}. For a CBD_{Cex} feed concentration of 20 μM (which is typical for our system), R_{eq} therefore is equal to 0.005. Adsorption is therefore irreversible, which is the primary requirement for development of a generic affinity-chromatography system based on CBD_{Cex}-fusion tails.

1.6 Thesis Objectives

The primary objective of this thesis are (1) the establishment of a generic, scalable affinity chromatography system based on fusion proteins containing a cellulose binding domain, (2) the derivation of a general model describing flow, transport and binding properties in the chromatographic column, (3) the determination of model parameters through measurement and established correlations, (4) the validation of the model over the entire range of permissible operating conditions, and (5) the application of column simulation (model) results to the purification of a recombinant fusion-protein containing a CBD tag.

CHAPTER 2

MATERIALS AND GENERAL METHODS

2.1 CBD_{cex} production and purification

A portion of CBD_{Cex} used in these studies was obtained as a gift from Repligen Sandoz Chemicals (Lexington, MA). The remainder was produced by fermentation. The gene fragment encoding for the cellulose binding domain CBD_{Cex} of the β-1,4-exoglucanase/xylanase Cex (Tomme *et al.*, 1995) from the soil bacterium *Cellulomonas fimi* was subcloned into the pTugE07 vector and expressed in *Escherichia coli* (*E. coli*) strain JM101 according to the protocol of Ong *et al.*, (1993). The pTugE07 vector contains the leader sequence from the exoglucanase Cex, which directs the recombinant protein to the periplasm of *E. coli.*

Stock and 5-mL inoculum cultures of recombinant *E. coli* JM101 were grown at 30°C and pH 7 in TYP media containing 0.05 mg mL^{-1} kanamycin to an optical density of 0.5 at 600 nm. DMSO was added to a final concentration of 10% (v/v), and 1-mL aliquots of the stock culture were stored at -70°C until required.

CBD_{Cex} was produced by growing 500-mL cultures of recombinant *E. coli* JM101 at 37°C in 2-L shake flasks containing TYP media and 0.05 mg mL^{-1} kanamycin. The culture was grown to an $\mathring{O}$.D. (600 nm) of 1.0, then induced overnight with 0.1-mM IPTG (isophenyl-thio-β-D-galactopyranoside). Previous studies have shown that significant leakage of periplasmic proteins can occur at high levels of expression (Hassenwinkle *et al.*, 1997). As shown in Figure 2.1, SDS-PAGE gels of the culture supernatant and isolated-cell lysate indicate that most of the CBD_{Cex} produced localizes in the culture supernatant.

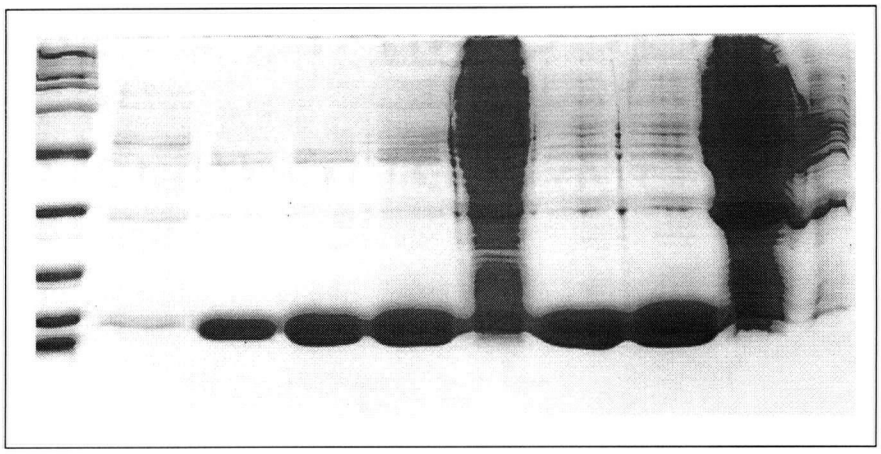


Figure 2.1. SDS-PAGE gel of broad-range molecular-weight markers (BioRad) (lane 1). Culture supernatant (10μL) (lane 2). Increasing volumes of culture supernatant (250, 500, 1000 μL, respectively) were bound (1 h) to 5 mg of Avicel. The samples were centrifuged, then washed twice with 1-mL of 1M NaCl in 50 mM phosphate buffer, centrifuged again and washed with 50 mM phosphate buffer. The supernatant was discarded. The Avicel with bound CBD was boiled in SDS sample buffer for 2 min and 10 μL were loaded into each lane (lane 3, 4, 5). Cell extract (lane 6). Increasing volumes (50 and 100 μL respectively) of cell extract mixed with Avicel (lane 7, 8). Unbound supernatant from cell extract after mixing with Avicel (lane 9).

Culture supernatant was recovered by centrifuging at 4°C and 10,000 rpm for 20 minutes the final culture, and decanting away the supernatant from cell paste. The supernatant was then mixed with an equal volume of Avicel PH101, a microcrystalline cellulose and allowed to equilibrate at 4°C for 4 h. The CBD_{Cex} -loaded Avicel was recovered by centrifugation at 15,000 rpm for 15 minutes. The Avicel pellet was resuspended in 50-mM Phosphate buffer and washed until the liquid recovered was colour-free (*ca.* 4 volumes). Sixty mL of the CBD_{Cex}-loaded Avicel was then poured into an XK-20 column (Pharmacia, Inc.; Sweden) and allowed to settle. The inlet flow adaptor was inserted and compressed to the surface of the Avicel slurry. The 60-mL column was then washed with two volumes of phosphate-buffered

saline, followed by a final wash with one volume of 50-mM phosphate buffer. Pure CBD_{Cex} was then eluted from the column with 6-M guanidinium hydrochloride.

Pure CBD_{Cex} from Repligen Sandoz Chemicals (Lexington, MA) was obtained as a 2 g L^{-1} aqueous solution and maintained in that state at 4 °C until use. CBD_{Cex} eluted from the Avicel column was renatured by dialyzing the concentrated solution against 50-mM phosphate buffer (pH 7) containing 0.01% sodium azide. A 350-mL Amicon ultrafiltration cell containing a 1-kDa molecular-weight-cutoff membrane was used to exchange the solvent. The recovered CBD_{Cex} was stored in the 50-mM phosphate buffer (0.01% NaN_3) at 4°C until use.

2.2 Production of the protein A- CBD_{Cex} fusion protein

The gene fragment encoding CBD_{cex} and the proline-threonine rich linker sequence which precedes it in Cex (O'Neil, *et al.*, 1986) were introduced into the plasmid pRIT5 (Nilsson *et al.*, 1985, obtained from Pharmacia Inc., Sweden), which contains the leader sequence and IgG binding domains of the staphylococcal protein A. The resulting pSACBD plasmid encoding the Protein A-CBD_{cex} fusion protein (molecular weight 42,000) was cloned into *Escherichia coli* strain JM101 (Yannish-Perron *et al.*, 1985; Ramirez, *et al.*, 1993). Stock cultures of the recombinant *E. coli* JM101 were prepared as described above in the presence of 0.05 mg mL^{-1} ampicillin and stored at -70°C.

Luria-broth agar plates (Difco) were streaked with a loop of stock culture and stored at 4°C. A single colony was taken from the plate and used to seed a 10-mL inoculum culture grown at pH 7 and 37°C in an optimized M9 Minimal Media (Hassenwinkle *et al.*, 1997) containing 11.76 g/L Na_2HPO_4 , 5.88 g/L KH_2PO_4 , 0.5 g/L NaCl, 1.4 g/L NH_4Cl , 0.24 g/L $MgSO_4$, supplemented with 0.2 % glucose. A trace-metal solution (0.5 mL/L) was also added which

contained the following per liter of water: 40 g $CaCl_2.H_2O$, 10 g $MnSO_4..H_2O$, 10 g $AlCl_3.6H_2O$, 4 g $CoCl_2$. $6H_2O$, 2 g $ZnSO_4.7H_2O$, 2 g $Na_2MoO_4.2H_2O$, 1 g $CuCl_2.2H_2O$, 0.5 g H_3BO_3 and 0.05 g $FeSO_4.7H_2O$. Hundred μL of a solution containing 0.5 g/L L-proline, 0.34 g/L thiamine and 100 μg ampicillin/mL (Sambrook, *et al.*, 1989), was added to the culture through a 0.22 μm sterile filter.

The inoculum culture was grown to an O.D. (600 nm) of 1 and then added to 400 mL of fresh media in 2-L shake flasks (250 rpm) at 37°C. The cultures were allowed to grow with additions of 2 g/L glucose up to 37 hours. At 23 hours culture time, proline (0.25 g/L) and ampicillin (50μg/mL) were added. Cell growth was followed by measuring the optical density of the culture at a wavelength of 600 nm. Protein A-CBD_{cex} is constitutively expressed in this construct and a large fraction is exported by leakage to the culture supernatant. The culture supernatant was separated from the cells by centrifugation at 10,000 rpm for 10 min in a Beckman Centrifuge (JA10 rotor) at 4 °C. The pH of the culture was controlled manually at 7.0 ± 0.3 by adding concentrated NH_4OH.

The concentration of Protein A-CBD_{cex} in the culture supernatant was measured by SDS-PAGE electrophoresis (Laemmli, 1970). In each gel, 0.124 mg/mL (8.8×10^{-5} μM) of purified Protein A-CBD_{cex} was used as a control. Gels were scanned by densitometry (Molecular Dynamics; Palo Alto, CA) to determine concentrations of Protein A-CBD_{cex} in the supernatant, based on a calibration curve obtained with previously purified protein. The standard deviation in densities obtained from different gels was calculated using the control sample from each gel and is shown in the calibration curve (Figure 2.2)

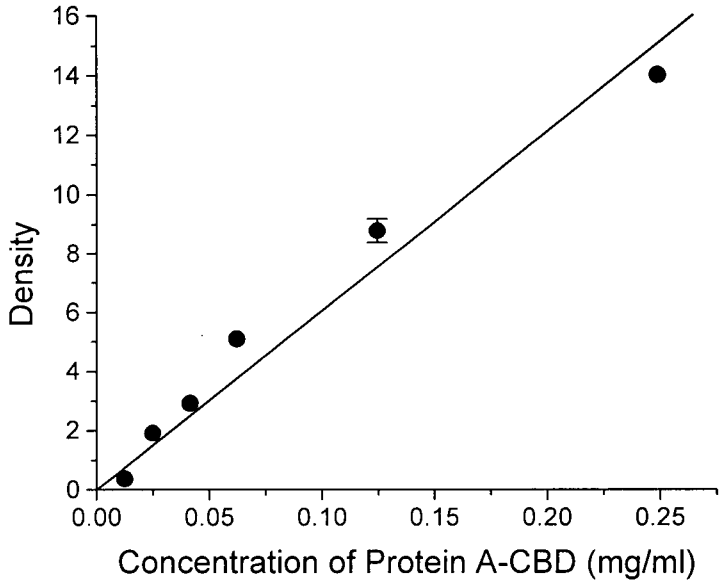


Figure 2.2. Calibration curve for Protein A-CBD_{Cex}. $y = 60.547\ x$ $(R = 0.9862)$

2.3 Cellulose matrix

CF1 particulate microcrystalline cellulose and Sigmacell Type 50 were purchased from Sigma Chemical Co. (Mississauga, Ont.). Avicel PH101, a heterogeneous particulate microcrystalline cellulose was purchased from FMC International (Little County Island, Cork, Ireland). Avicel is a semicrystalline cellulose preparation obtained from wood fibers by partial acid hydrolysis followed by spray drying of the washed slurry (Gilkes *et al.*, 1992). Prior to use, the Avicel was sieved in a set of standard sieves for 0.5 hour to obtain a set of particle size fractions based on the average particle diameter d_p (fraction 1 (350 mesh size), $d_p < 45$ μm; fraction 2, 45 μm $< d_p <$ 75 μm; fraction 3 (200 mesh size), 75 μm $< d_p$). The range of particle diameters and the average d_p of each fraction was determined by imaging 100 particles of the

sieved Avicel using an inverted microscope (Axiovert 100, Carl Zeiss, Oberkochen, Germany) equipped with a monochrome Solid State Camera (COHU Inc., San Diego, CA) connected to a real-time frame-grabber card (VISION*plus*-AT, Overlay Frame Grabber, Imaging Technology Inc., Woburn, MA) installed in an Intel 80486 personal computer. The length and width of each particle were determined from a calibrated grid overlay. Particle volumes were then calculated assuming a prolate ellipsoid geometry from which the mean particle radius of the sphere was obtained. The mean radius R_p of fraction 2 (45 μm < d_p < 75 μm), the fraction chosen for our column chromatography system, was 53 μm.

2.4 Binding isotherm and cellulose binding capacity

Static (batch) measurement of equilibrium binding isotherms followed the procedures developed by Creagh *et al.* (1996). The equilibrium binding isotherm and the capacity of sieved Avicel for Protein A-CBD_{cex} at pH 7 and 20°C were determined by binding to 3 mg of sieved Avicel increasing amounts of purified CBD_{cex} or protein A-CBD_{Cex} ($11x10^{-3}$ - $2.19x10^{-4}$ μmole/mL). Each sample of Avicel was mixed with the protein solution for at least 3 hours to ensure equilibrium had been reached. The samples were then centrifuged at 4°C for 10 min and 15,000 rpm. The concentration of unbound protein was determined in the supernatant by reading the absorbance at 280 nm; the extinction coefficients for CBD_{Cex} and protein A-CBD_{Cex} are 2.31 mL cm^{-1} mg^{-1} (Ong *et al.,* 1993*)* and 0.8 mL cm^{-1} mg^{-1} (Ramirez *et al.*, 1993), respectively. The concentration of bound protein was determined by mass balance from the difference between the initial protein concentration and that unbound. All measurements were done in triplicate. The experimental data was fit using the Langmuir isotherm equation to obtain the maximum binding capacity of the sieved Avicel. The calculated

maximum capacities, $Q_{max,}$ for CBD_{Cex} and protein A-CBD_{Cex} are 2.86 mol m^{-3} and 0.66 mole m^{-3}, respectively.

2.5 Sieved Avicel affinity chromatography

All affinity chromatography studies were executed in a Pharmacia Inc. (Uppsala, Sweden) FPLC system equipped with two P-500 reciprocating pumps, a mixing and injection valve, a custom-packed sieved Avicel column, a UV-MII flow spectrophotometer (280 nm), and a FRAC-200 fraction collector. The Avicel column was prepared as follows. 1.5 to 3 g of sieved Avicel was mixed with 50-mM phosphate buffer (pH 7) to a volume of 4.5 mL and poured into a 1-cm diameter XK-10 column (Pharmacia Inc.) with a total working volume of 8 cm^3. The column was positioned vertically to allow the cellulose to settle. Once a clear interface was formed, the outlet flow adaptor was inserted and compressed onto the cellulose to finger tightness. The column was then equilibrated by passing through a 50-mM phosphate buffer solution (pH 7, 0.05% NaN_3) at a flowrate between 0.2 and 1.0 mL min^{-1} until a stable baseline was observed.

2.6 Experimental breakthrough curves

Breakthrough curves, as described in chapter 1, were measured for feed solutions of CBD_{Cex} in 50-mM phosphate buffer (0.05% NaN_3) at 20°C. Experiments covered six different inlet protein concentrations between 17.3 to 48.2 μM, nine different volumetric flowrates ranging from 0.4 mL min^{-1} to 1.2 mL min^{-1} , and several column lengths. All breakthrough curves were measured in duplicate. The packed column was first equilibrated with a mobile phase of 50-mM phosphate buffer at a specified volumetric flowrate until a stable baseline was observed. At set

time t=0, a pure CBD_{Cex} solution (50-mM phosphate buffer, pH 7) at a specified concentration was introduced as a step change and the concentration of CBD_{Cex} in the eluent chromatogram was recorded as a function of time by continuously monitoring absorbance at a wavelength of 280 nm.

Breakthrough curves were also measured for protein A-CBD_{Cex} loaded continuously on the column from a mobile phase of culture supernatant. Up to 400 mL of culture supernatant (C_o = 5.17 μM of Protein A- CBD_{Cex}) was passed through the packed Avicel column at a superficial velocity u of 2.12×10^{-2} cm s^{-1} to obtain the breakthrough curve. Fractions (*ca.* 1 mL) from the column effluent were collected and each 30 μL sample was analyzed by the rabbit anti-β-glucosidase antibody dot-blot method described in section 2.7. Following breakthrough, the column was washed for one hour with 1-M NaCl in 50-mM phosphate buffer (pH 7), and then for another hour with 50-mM phosphate buffer pH 7. Bound and washed protein A-CBD_{Cex} was then eluted with distilled water. Pooled fractions from the elution peak were analyzed by SDS-PAGE and densitometry as described in section 2.2.

2.7 Analytical Methods

Optical densities (OD) of cultures were measured at 600 nm with a Varian DMS 200 Spectrophotometer. Glucose concentrations in culture supernatants were measured using a Beckman Glucose Analyzer II calibrated against a 1.50 g L^{-1} glucose standard. Solution eluent-chromatogram concentrations of CBD_{Cex} were calculated from a calibration curve measured by UV absorbance at 280 nm using an extinction coefficient of 2.31 mL cm^{-1} mg^{-1} (Ong *et al.*, 1993).

Concentrations of protein A-CBD_{Cex} in culture supernatants and in column flow-through were determined by an antibody-based dot-blot method. Two microliters of protein A-CBD_{Cex} containing sample, either from culture supernatant or collected fractions at the column outlet, were spotted on Whatman 1 Chr paper preblocked for 1 h with 3 % bovine serum albumin (BSA) in phosphate-buffered saline (pH 7) containing 0.05 % Tween-20. The paper was then incubated for 1 h at room temperature with rabbit anti-β-glucosidase antibody (1:500 in PBS containing 0.05% Tween and 0.5 % BSA) raised against the ß-glucosidase of *Agrobacterium* sp. (Ramirez *et al.*, 1993). The paper was washed three times with PBS (0.05 % Tween) and incubated for 1h at room temperature with goat anti-rabbit IgG coupled to horseradish peroxidase (1:7000 in PBS containing 0.05% Tween and 0.5 % BSA). The paper was washed three more times again with PBS (0.05 % Tween). The enhanced chemiluminescence kit (ECL) from Amersham was used for antibody/antigen detection. The paper was placed in a cartridge and exposed to X-ray film (Kodak).

CHAPTER 3

DEVELOPMENT, CHARACTERIZATION, AND MODELING OF THE AVICEL-BASED AFFINITY CHROMATOGRAPHY COLUMN

3.1 Introduction

This chapter focuses on the design and characterization of a novel affinity chromatography column for one-step purification of recombinant fusion proteins containing a CBD. The utility of the method is first established, along with a brief justification of the cellulose-matrix and column configuration chosen. Column performance over the realistic range of operating conditions is then determined through both pulse elution and breakthrough curve analysis. A model is then derived to describe binding, mass-transfer effects, and breakthrough-curve profiles with the aim of establishing a useful tool for scale-up and column-performance analysis. Models parameters are determined either by independent experiment or established correlations. Pulse and breakthrough curve data are combined with moments theory to establish transport properties in columns packed with semi-crystalline cellulose (Avicel) particles containing pores up to 60 Å in diameter.

3.2 Column Composition and Configuration and System Validation

Table 3.1 lists a range of cellulosic materials for which CBD_{Cex} is known to have a relatively high affinity. The association constant for CBD_{Cex} binding to any of these materials is

greater than 10^5 M^{-1} and many of them have been successfully applied to batch adsorption and purification of CBDs (Ong *et al.*, 1993; Gilkes *et al.*, 1992; Creagh *et al.*, 1996). Most of these cellulosics, however, are not suitable for chromatographic column applications. Bacterial microcrystalline cellulose and disrupted *V. ventricosa* cellulose (Jervis *et al.*, 1997), for instance, are fibrous in structure and highly compressible, resulting in column plugging at low pressure drops. Of the semicrystalline particulate celluloses investigated (CF1, Sigmacell, and Avicel), Avicel columns show the best flow characteristics and resistance to compression, while maintaining a reasonably high capacity. Particulate Avicel sieved to an average diameter of 60 μm (45 μm < d_p < 75 μm) showed no significant compression effects against back pressures up to 5 bar, making it a suitable matrix for our studies.

Table 3.1: Cellulosic Materials to which CBD_{Cex} binds (from highest to lowest degree of crystallinity)

Material	K_a (M^{-1})
Valonia (*V.ventricosa)*	10^7
BMCC (*Acetobacter xylinum*)	10^7
Avicel (FMC)	10^6
Sigmacell (Sigma)	10^6
CF1(Whatman)	-
PASC (Phosphoric Acid Swollen Cellulose)	10^6

Ong *et al.* (1993) have shown that binding of CBD_{Cex} to Avicel is characterized by an association constant near 10^6 M^{-1} which is fairly insensitive to changes in pH and ionic strength. Creagh *et al.* (1996) have also shown that binding is only slightly exothermic, indicating that binding strength is not a strong function of temperature. Thus, in principal, binding of CBD_{Cex} fusion proteins to the column can occur over a wide range of loading conditions and mobile-phase compositions.

Figure 3.1 shows an experimental chromatogram for affinity loading and elution by 6-M guanidinium hydrochloride of a 38-μM solution of CBD_{Cex} at 20°C (pH 7, 50-mM phosphate). Breakthrough is shown between points 1 and 2, column wash between points 2 and 3, and elution between points 3 and 4. The relatively steep breakthrough curve combined with the sharp elution peak indicate uniform flow and reasonably fast binding kinetics in the column, along with providing proof of concept of the affinity column design and utility.

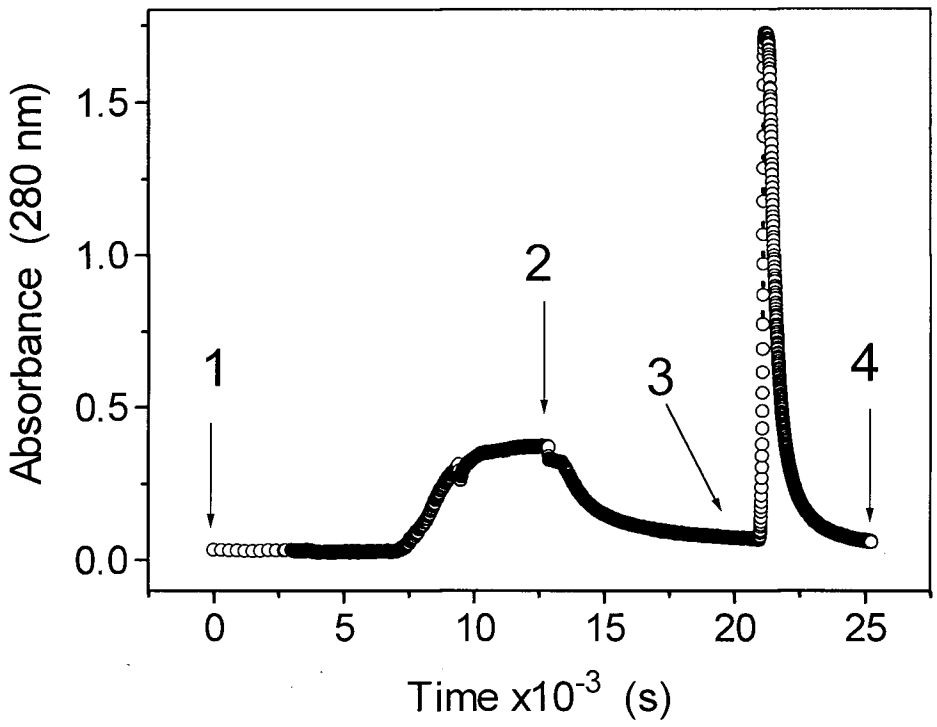


Figure 3.1 Full loading, wash and elution chromatogram of CBD_{Cex} . (1-2) Load; (2-3) Wash; (3-4) Elution. (C_o = 38 μM, u = 2.11 x10^{-4} m/s, 20 °C)

3.3 Column Performance and Efficiency

Column efficiency describes the rate at which nonbinding solute molecules in a pulse injection separate as they travel through the column. The average rate at which each solute species travels through a column depends on its partition coefficient (*i.e.*, affinity for the solid-phase matrix). However, mass transfer effects, such as nonuniform pathways (channeling) through the column packing, as well as partition forces between the solid and mobile phases will cause spreading of the peak for a given solute as it travels through the column. Eluent peak

width t_w provides a measure of column efficiency. Complete separation requires that the trailing edge of the first eluent peak A not overlap the leading edge of the next eluent peak B and so on. For complete separation of components A and B, this would mean that

$$\left(t_{RB} - \frac{1}{2}t_{wB}\right) - \left(t_{RA} + \frac{1}{2}t_{wA}\right) > 0 \qquad (3.1)$$

where t_{RA} is the retention time of peak A, and t_{wA} is the peak width at baseline. In an inefficient column, the rate of peak spreading is too fast to allow for complete peak separation. What is therefore desired is a column which provides reasonable retention of the target protein or product with minimal spreading of the peak. Thus, we want to maximize t_R/t_w or $t_R/t_{w1/2}$, the ratio of the retention time to the peak width.

The ratio $t_R/t_{w1/2}$ for the product peak in the eluent is a central parameter in the theory of chromatography. For a chromatogram in which the eluent peaks are Gaussian, the number of theoretical plates N_{th} in the column is proportional to the square of $t_R/t_{w1/2}$,

$$N_{th} = 5.54\left(\frac{t_R}{t_{w1/2}}\right)^2 \qquad (3.2)$$

A large value for N_{th} is therefore the best indicator of an efficient column (Giddings, 1965).

Determination of N_{th} for our affinity column requires measuring elution peaks from pulse injections of the target protein under nonbinding (but native protein) conditions. CBD_{Cex}, however, binds Avicel under essentially all native protein conditions. To overcome this problem, we selected a nonbinding protein analogue to CBD_{Cex}. Figure 3.2 shows eluent peaks from our packed Avicel column as a function of superficial velocity u for 200 μL pulses (15.8 μM) of

horse-heart myoglobin (Sigma Chemical Co., M-1882). Myoglobin (MW 17 kDa) is similar to CBD_{Cex} in both size and net charge (pI 7), but shows no binding affinity for the Avicel column. Elution peaks at low u are near Gaussian, allowing determination of theoretical plates in the column using Equation (3.2). N_{th} is equal to *ca.* 30 for the 4.5 cm column used in these studies, and increases to greater than 150 for a 25-cm column of the same diameter.

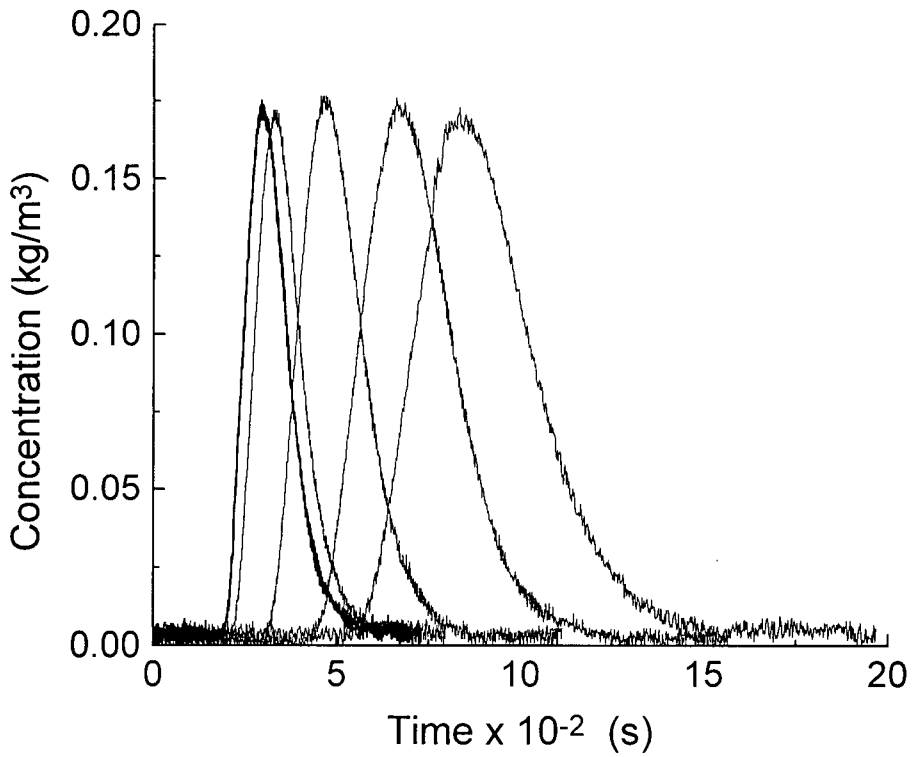


Figure 3.2. Eluent peaks of myoglobin from the Avicel column as a function of superficial velocity. From right to left: u = 8.5 $x10^{-3}$, 1.1 $x10^{-2}$, 1.5 $x10^{-2}$, 2.1 $x10^{-2}$, 2.3 $x10^{-2}$ cm/s.

At superficial velocities in excess of 0.02 cm s^{-1}, eluent peak profiles show more of a Poison-type distribution, indicating increased flow nonidealities. For the column used, this

equates to a volumetric flowrate of 1 mL min^{-1}, which should therefore be viewed as an upper limit for mobile-phase flow. At all flowrates, including those above 1 mL min^{-1}, no compression of the column or associated anomalous increase in back pressure was observed.

3.4 Binding Equilibrium Constant and Sorbent Capacity

Batch adsorption-isotherm experiments were performed to determine the association constant K_a (M^{-1}) and maximum sorbent capacity Q_{max} (μmol g^{-1} Avicel) for binding of CBD_{Cex} to sieved Avicel at 20°C and pH 7. Figure 3.3 shows a typical adsorption isotherm. The near-infinite initial slope of the isotherm indicates the very high affinity of CBD_{Cex} for Avicel. The adsorption plateau of 2.9 mol m^{-3} yields a value for Q_{max} of 3.1 μmol g^{-1} Avicel, where the conversion from specific surface concentration to volumetric concentration was made using the measured density of the Avicel particles (930 ± 50 kg m^{-3}). Thus, provided all of this surface area remains available after packing, our affinity Avicel column is capable of binding ca. 30 g of CBD_{Cex} per liter of column, which compares favorably with most commercial affinity columns. The 4 mL column used in most these studies is capable of binding 120 mg of CBD_{Cex}.

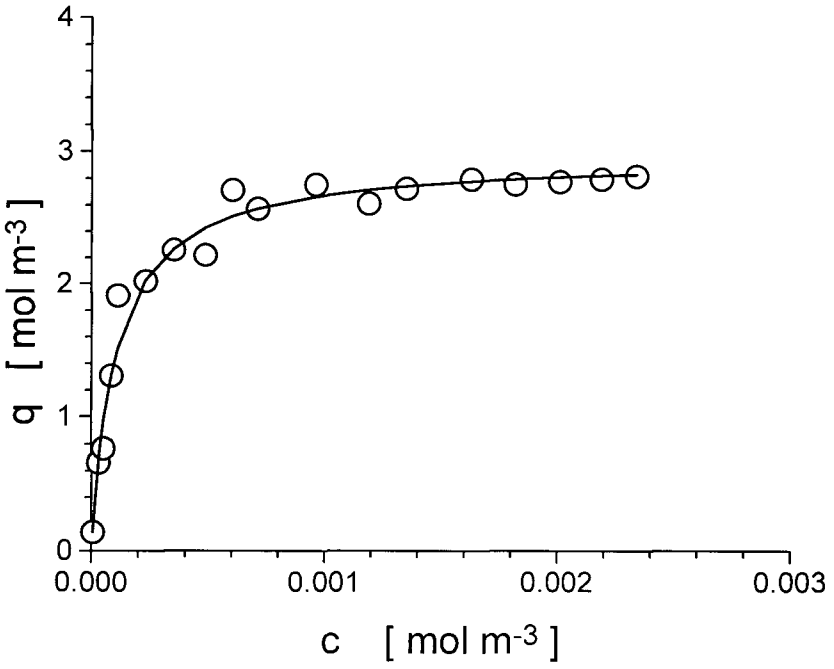


Figure 3.3. Adsorption isotherm for CBD_{Cex} on Avicel PH101 (unsieved), 25°C, pH 7 (50-mM phosphate Buffer). (0) Experiment; (solid line) Langmuir Equation fit

Regression of the binding constant K_a from the isotherm in Figure 3.3 requires an appropriate adsorption-isotherm model. To establish this model, we consider a binary solution (phase o) of solvent w and protein p in equilibrium contact with a sorbent surface (phase s) containing N moles of independent identical binding sites. At equilibrium, all S surface binding sites will be occupied such that r_w sites are filled with adsorbed water and r_p sites are filled with a specifically adsorbed protein (CBD_{Cex}) molecule. At constant temperature and pressure, the Gibbs energy change for the exchange of molecules is

$$dG = \mu_{wo} dn_w + \mu_{ws} dr_w + \mu_{po} dn_p + \mu_{ps} dr_p \qquad (3.3)$$

where n_w and n_p are the moles of water and protein in the bulk solution, r_w and r_p are the moles of each on the sorbent surface, and μ is the chemical potential. The following independent mass

balance constraints apply: conservation of w, conservation of p, and conservation of total binding sites. That is,

$$n_w + r_w = N_w \qquad (3.4a)$$

$$n_p + r_p = N_p \qquad (3.4b)$$

$$r_w + r_p = N \qquad (3.4c)$$

Therefore,

$$dn_p = - dr_p = dr_w = - dn_w \qquad (3.5)$$

At equilibrium, dG is equal to zero. Substitution of Equation (3.5) into Equation (3.3) and imposing the condition of equilibrium then gives,

$$\frac{dG}{dn_p} = \mu_{po} - \mu_{ps} - (\mu_{wo} - \mu_{ws}) = 0 \qquad (3.6)$$

It is convenient to define each chemical potential μ_i in Equation (3.6) in terms of activity a_i

$$(\mu^o_{po} + RT \ln a_{po}) - (\mu^o_{ps} + RT \ln a_{ps}) - [(\mu^o_{wo} + RT \ln a_{wo}) - (\mu^o_{ws} + RT \ln a_{ws})] = 0 \qquad (3.7)$$

As is generally done in gas-phase adsorption equilibria theory, we express each activity in terms of mole fraction. For the bulk phase, x_p is the mole fraction protein and $x_w = (1 - x_p)$ is the mole fraction water. For the adsorbed phase, the mole fraction of binding sites filled by protein is $\theta_p = r_p / N$, and $\theta_w = (1 - \theta_p)$ is the mole fraction of total binding sites occupied by solvent. Equation (3.7) then becomes

$$(\mu^o_{po} + RT \ln x_p) - (\mu^o_{ps} + RT \ln \theta_p) - [(\mu^o_{wo} + RT \ln x_w) - (\mu^o_{ws} + RT \ln \theta_w)] = 0 \qquad (3.8)$$

The dimensionless equilibrium association constant K_a' for the binding reaction is given by (Smith and van Ness, 1986)

$$- RT \ln K_a' = \Sigma \nu_i \mu_i^o = \mu^o_{po} - \mu^o_{ps} - (\mu^o_{wo} - \mu^o_{ws}) \qquad (3.9)$$

so that from Equation (3.8)

$$K_a' = \frac{\theta_p x_w}{\theta_w x_p} = \frac{\theta_p(1-x_p)}{(1-\theta_p)x_p} \qquad (3.10)$$

Solving for θ_p, the surface fraction of adsorbed protein, gives

$$\theta_p = \frac{K_a' x_p}{(1-x_p)+K_a' x_p} \qquad (3.11)$$

In Equation (3.11), K_a' is the equilibrium binding constant of protein p to sorbent s, but also includes the energetics associated with release of water from the sorbent surface, which thereby distinguishes the result from classic Langmuir theory (for a derivation of Langmuir theory, see A. W. Adamson, *Physical Chemistry of Surfaces,* 5th Ed., Wiley-Interscience, New York (1990)). From the form of Equation (3.11), however, it is clear that binding of solute in solution is closely analogous to gas adsorption on solid surfaces.

For $K_a' >> 1$, Equation (3.11) simplifies to

$$\theta_p = \frac{K_a' x_p}{1+K_a' x_p} \qquad (3.12)$$

which is identical in form to the Langmuir equation. Equation (3.12) can be converted to molar concentrations without additional assumptions

$$\Gamma = \frac{\Gamma^{max} K_a[P]}{1+K_a[P]} \qquad (3.13)$$

where K_a now carries units of M^{-1}, Γ^{max} is the monolayer capacity of the sorbent surface, and Γ is the surface concentration of protein.

Figure 3.3 shows the results of a nonlinear least-squares regression of Equation (3.13) to the adsorption isotherm data for CBD_{Cex} binding to Avicel in aqueous solution at 20°C. The

regressed K_a is 9.5 (± 1.3) x 10^6 M^{-1}, which satisfies the assumption (K_a>>1) made in deriving Equation (3.12) and indicates very high affinity of CBD_{Cex} for Avicel in aqueous solution.

3.5 Affinity Column Model

Breakthrough and elution profiles in chromatographic columns may depend on any number of kinetic and mass-transfer subprocesses (*e.g.*, film mass-transfer resistance, pore diffusion, binding rate). The fundamental mechanisms underlying many of these subprocesses have been identified and used to derive appropriate continuity and constitutive equations for quantifying time and spatial variations in protein concentration within the column (Chase (1984); McCormick (1988); Liapis *et al.*, (1989); Ruthven (1984)). These results have been used in combination with an appropriate binding rate equation to derive an unsteady-state differential model describing time-dependent CBD-fusion-protein migration and uptake within our Avicel affinity chromatography column.

3.5.1 Continuity and solute uptake equations

The model we have developed to describe our Avicel affinity chromatography column is depicted in Figure 3.4. The affinity column of length L, cross-sectional area A and void fraction ε is treated as a uniform packed bed of monodisperse porous sorbent spheres of radius R_p, porosity ß and density ρ_p. The mobile phase, at volumetric flowrate Q, has a solute concentration c(z,t) which is assumed to be independent of radial position within the column cross-section. Transport through the column void can occur by convection or axial dispersion (which includes eddy and molecular diffusion). Changes in c(z,t) can also occur through solute uptake by the sorbent spheres, where the solute concentration in the particles is s(r,z,t). The

sorbate concentration in the particle is $q_i(r,z,t)$, and the solute concentration in the pores is $c_i(r,z,t)$, such that

$$s(r,z,t) = ßc_i(r,z,t) + \rho_p\, q_i(r,z,t) \qquad (3.14)$$

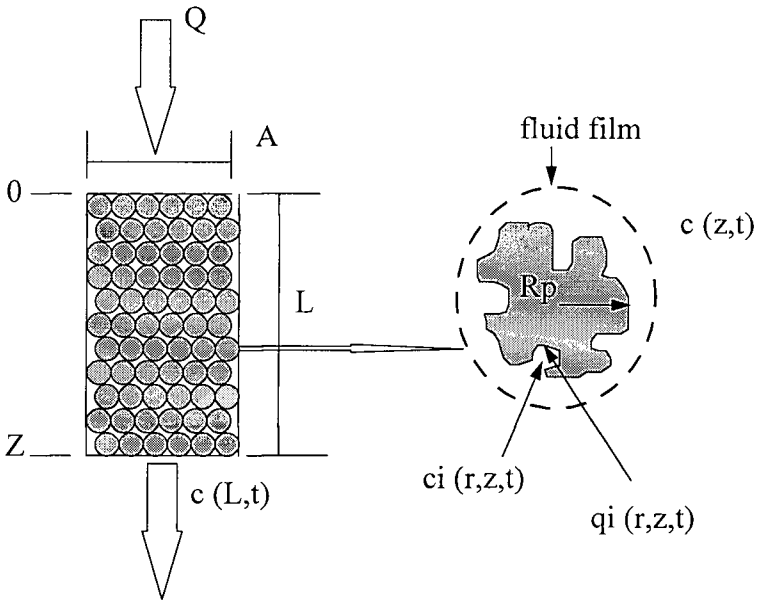


Figure 3.4. Description of the affinity chromatography column

Changes in solute concentration with time along the column length are described by the differential mass balance over a section of the column

$$\frac{\partial c}{\partial t} = D_L \frac{\partial^2 c}{\partial z^2} - \frac{u}{\varepsilon}\frac{\partial c}{\partial z} - \frac{(1-\varepsilon)}{\varepsilon}\frac{\partial s}{\partial t} \qquad (3.15)$$

where D_L is the axial diffusion coefficient ($m^2\ s^{-1}$), and u is the superficial liquid velocity ($m\ s^{-1}$). The terms on the right-hand side of Equation (3.15) account for axial dispersion, convective transport and intraparticle uptake of solute, respectively.

The rate of change of the average solute concentration in the particles depends on the flux of solute N_o into the particles:

$$\frac{(1-\varepsilon)}{\varepsilon}\frac{\partial s}{\partial t} = \frac{3(1-\varepsilon)}{R_p \quad \varepsilon} N_o = \frac{3(1-\varepsilon)}{R_p \quad \varepsilon} D_p \left|\frac{\partial c_i}{\partial r}\right|_{r=R_p} \qquad (3.16)$$

where $3(1-\varepsilon) / R_p$ is the surface area per unit bed volume of spherical particles, and D_p is the solute diffusion coefficient within the pore liquid ($m^2\ s^{-1}$) based on the entire particle volume. The solute concentration in the bulk liquid c(z,t) is coupled to $c_i|_{r=R_p}$ by the rate of mass transfer through the stagnant fluid film surrounding each particle, such that

$$D_p \left|\frac{\partial c_i}{\partial r}\right|_{r=R_p} = k_f\left(c - c_i|_{r=R_p}\right) \qquad (3.17)$$

where k_f is the film mass-transfer coefficient.

Finally, gradients in free solute concentration within the pores of the sorbent particles caused by intrapore diffusion and adsorption of sorbate are described by an intraparticle mass balance

$$\frac{\partial c_i}{\partial t} = D_p\left(\frac{\partial^2 c_i}{\partial r^2} + \frac{2}{r}\frac{\partial c_i}{\partial r}\right) - \frac{\rho_p}{\beta}\frac{\partial q_i}{\partial t} \qquad (3.18)$$

where r is the radial position within the particle and varies from 0 to R_p. Convection is not considered within the particle volume, and we assume that intraparticle solute concentrations are angularly (θ and ϕ) independent at a given radial position.

Together, Equations 3.15 to 3.18 model transport and adsorption of solute in the column as a series of convection, diffusion and reaction steps. The solute must first be transported by convection to the bulk fluid outside the particle, then diffuse through the stagnant liquid film

surrounding the particle, and finally diffuse through the pores of the particle to find an available binding site. The binding reaction must then occur until the free solute within the neighboring pore liquid and the sorbate are in equilibrium.

3.5.2 Binding kinetics/equilibrium model

Solutions of Equations 3.14 to 3.18 requires a final equation relating the local sorbate concentration $q_i(r,z,t)$ to the local solute concentration in the pore liquid $c_i(r,z,t)$. If the rate of formation of the affinity complex on the pore wall is rate limiting, this relation takes the form of a binding rate equation of the form

$$\frac{\partial q_i}{\partial t} = f\left(c_i, q_i\right) \qquad (3.19)$$

Often, largely for simplicity, the rate equation is assumed second order in the forward direction and first order in the reverse direction, so that it obeys Langmuir-type adsorption theory (see Equation (3.13)) at equilibrium:

$$\frac{\partial q_i}{\partial t} = \mu_f c_i\left(Q_{max} - q_i\right) - \mu_r q_i \qquad (3.20)$$

where the forward and reverse rate constants μ_f and μ_r, respectively, are related to the Langmuir-type association constant (K_L) by $K_L = \mu_f / \mu_r$, and Q_{max} is the total concentration of available binding sites. Equation (3.20), and the associated Langmuir adsorption theory, often provide good quantitative correlation of high-affinity protein binding data (Liapis *et al.,* 1989; McCoy and Liapis, 1991; Arnold *et al.*, 1986). However, the derivation of Langmuir adsorption theory involves a number of rather severe assumptions, including the absence of lateral sorbate interaction, and the equal availability (both energetically and sterically) of all unoccupied binding

sites on the sorbent surface. Recently, a series of binding-rate models have been derived which relax one or more of these assumptions (Jin *et* al., 1994; Mao *et al.*, 1991; Ma *et al.*, 1996).

The fundamental repeat unit of cellulose is cellobiose. Gilkes *et al.*, (1992) have estimated that bound CBD_{Cex} covers approximately 30 cellobiose units of solvent exposed crystalline cellulose. Thus, although CBD_{Cex} may specifically bind to only a few cellobiose units, it is more appropriate, from a modeling perspective, to define the size of one binding lattice unit as that occupied by *ca.* 30 cellobiose units. Jervis *et al.* (1997) have recently shown that bound CBD_{Cex} diffuses in two dimensions on crystalline cellulose. In principal, this would mean that bound CBD_{Cex} can arrange itself in closest possible packing on the cellulose surface such that all available lattice sites are filled at saturation. However, depending on surface concentration and sorbent preparation, between 30% and 40% of the bound CBD_{Cex} remains immobile. Thus, there exists the potential for steric exclusion of binding lattice sites on the cellulose surface such that the CBD covers or overlaps potential binding sites, and thus not all sites are filled at plateau adsorption levels (*i.e.*, the surface jamming limit is less than the total number of solvent-exposed binding sites). As noted by Gilkes *et al.*, (1992) this results in an observed non-linearity in Scatchard plot of the adsorption isotherm data.

Jin *et al.*, (1994) have derived a binding rate model which incorporates surface exclusion effects into Langmuir theory. Surface concentrations are written in the form of surface lattice fractions so that

$$\frac{\partial \theta^*}{\partial t} = \mu_f \, c_i \left(1-\theta^*\right)\left(1 - B_1(\alpha)\theta^* - B_2(\alpha)\theta^{*2}\right)^2 - \mu_r \, \theta^* \qquad (3.21)$$

where θ^* is the fraction of filled surface lattice sites, α is the average number of overlapping sites covered by a bound CBD_{Cex} molecule, and B_1 and B_2 are constants which depend on α the following ways

$$B_1(\alpha) = \frac{0.7126\alpha + 1.404\alpha^{1.5}}{1 + 3.4363\alpha + 2.4653\alpha^{1.5}} \qquad (3.22)$$

$$B_2(\alpha) = \frac{0.07362\alpha + 0.1204\alpha^{1.5}}{1 + 0.5443\alpha + 0.2725\alpha^{1.5}} \qquad (3.23)$$

Based on the results of Gilkes *et al.*, (1992) and Jervis *et al.*, (1997), Equation (3.21) would appear to offer a more realistic model representation of CBD_{Cex} adsorption to crystalline cellulose. However, incorporation of this rate equation into our differential column model significantly complicates solution of Equation (3.14) to (3.18).

Under column-loading conditions where the rate of adsorption is fast compared with transport to the intraparticle binding site, the sorbate concentration q_i will be in equilibrium with the local concentration of solute in the pore liquid c_i, so that

$$\frac{dq_i}{dt} = 0 \qquad \text{(diffusion limited process only)} \qquad (3.24)$$

Under these conditions, Equation (3.20) reduces to the classic Langmuir adsorption isotherm equation

$$q_i = \frac{Q_{max}(\mu_f / \mu_r)c_i}{1 + (\mu_f / \mu_r)c_i} = \frac{Q_{max}K_L c_i}{1 + K_L c_i} \qquad (3.25)$$

where K_L is the Langmuir equilibrium association constant (M^{-1}). At local equilibrium conditions, the Jin *et al.* (1994) model reduces to

$$\theta^* = \frac{K_a c_i\left(1 - B_1(\alpha)\theta^* - B_2(\alpha)\theta^{*2}\right)^2}{1 + K_a c_i\left(1 - B_1(\alpha)\theta^* - B_2(\alpha)\theta^{*2}\right)^2} \qquad (3.26)$$

where K_a is the equilibrium association constant (M^{-1}).

3.5.3 Boundary conditions

When axial dispersion and mixing at the column exit are considered, the continuity equation (Equation 3.15) for solute i in the column requires two boundary conditions in z and one initial condition at time t = 0. We assume constant velocity u of the fluid stream (*i.e.*, u is independent of the space variable z) since the mobile phases entering the column are very dilute. Appropriate initial condition and boundary conditions for Equation (3.15) are then:

$$c = 0 \qquad t = 0 \qquad 0 \le z \le L \tag{3.27a}$$

$$\varepsilon D_L \frac{\partial c_{inlet}}{\partial z} = u\left(c_{inlet} - c\right) \qquad z = 0 \qquad t \ge 0 \tag{3.27b}$$

$$\frac{\partial c_{out}}{\partial z} = 0 \qquad z = L \qquad t \ge 0 \tag{3.27c}$$

This set of boundary conditions was first proposed by Dankwerts (1953). In certain cases, axial dispersion can be neglected, and a single boundary conditions of the form

$$c = c_{inlet} \qquad z = 0 \qquad t \ge 0 \tag{3.28}$$

is required along with Equation (3.27a) to solve Equation (3.15).

The complementary set of initial and boundary conditions for Equation (3.18), describing solute uptake within the particle volume, are

$$c_i = 0 \qquad t = 0 \qquad 0 \le r \le R_p \tag{3.29a}$$

$$\left.\frac{\partial c_i}{\partial r}\right|_{r=0} = 0 \qquad r = 0 \qquad t > 0 \tag{3.29b}$$

Equation (3.17) provides the final radial boundary condition by specifying that the concentrations c_i and c in the particle volume at $r = R_p$ and the bulk liquid surrounding the particle, respectively, are coupled by the rate of mass transfer through the fluid film.

3.5.4 Solution algorithm

Under certain limiting conditions, Equation (3.14) to (3.18) can be solved analytically. Chase (1984), for instance, obtained an analytical result by lumping into the forward and reverse rate constants for binding all mass transfer effects characterizing solute transport. The resulting model provides a reasonable description of breakthrough curve patterns generated by frontal loading when binding follows Langmuir theory, but does not allow one to determine the magnitudes of the various mass transfer resistances. Moreover, as discussed by Arnold and Blanch (1985a), Chase's definition of time is ambiguous since it depends on the size of the nonadsorbing species.

Vermeulen *et al.* (1973) obtained an analytical solution to Equations (3.14) to (3.18) which is valid for a plug-flow column exhibiting constant pattern behavior in which the adsorption is irreversible and the rate of adsorption is very fast.

An analytical solution to Equation (3.14) to (3.18) is not available when axial dispersion effects are considered and Dankwerts' boundary conditions are applied. The model equations, described in sections 3.5.1 to 3.5.3 were solved numerically for our Avicel affinity column, when either the binding rate equation of Langmuir (Equation (3.20)) or of Jin *et al.* (1994) (Equation (3.21)) was applied, using the iteration scheme and program of Koska and Haynes (1997) which solves the coupled transport and reaction equations by finite difference methods. The spatial derivatives in all equations were discretized using a control-volume method (Patankar, 1980).

The discretized forms of the diffusion-reaction equations (Equations 3.15 and 3.18) yield a set of nonlinear equations that were solved by a Newton-Raphson root finding method. A second-order Adams-Bashforth predictor -corrector method (Finlayson, 1980) was used to perform the time integration.

Figures 3.5 to 3.8 show the dependence of calculated breakthrough curves for a 10-cm model column on the value of the axial dispersion coefficient D_L, the film mass transfer coefficient k_f, the intrapore diffusivity D_p, and the forward adsorption rate constant μ_f (based on Equation (3.20)). The breakthrough curve represented by the solid line was obtained using the calculated or experimentally measured coefficients. To analyze the effect of each coefficient on the dynamic behavior of the breakthrough curve, values close to those measured or calculated were used (*e.g.* ± 20 %).

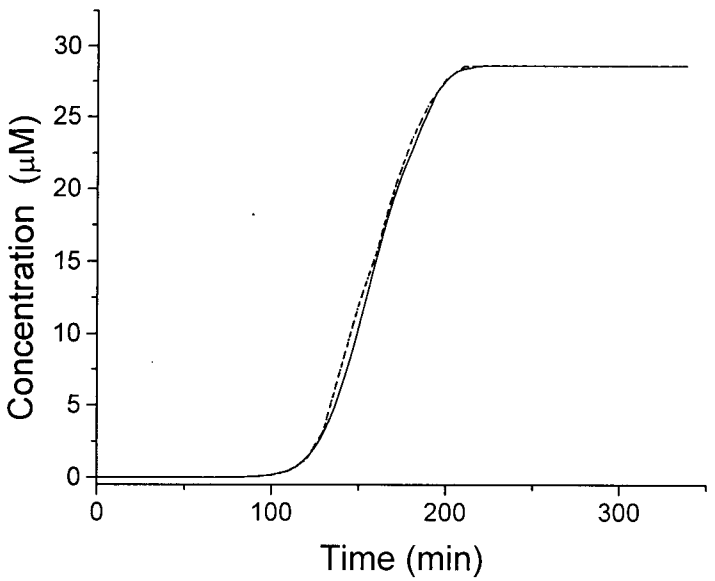


Figure 3.5. Dependence of calculated breakthrough curves on the value of the axial dispersion coefficient, D_L. Initial concentration C_o = 28.6 μM, u = 2.1 x10^{-2} m/s. (. .) 1.2D_L, (---) 0.8D_L, (–) D_L

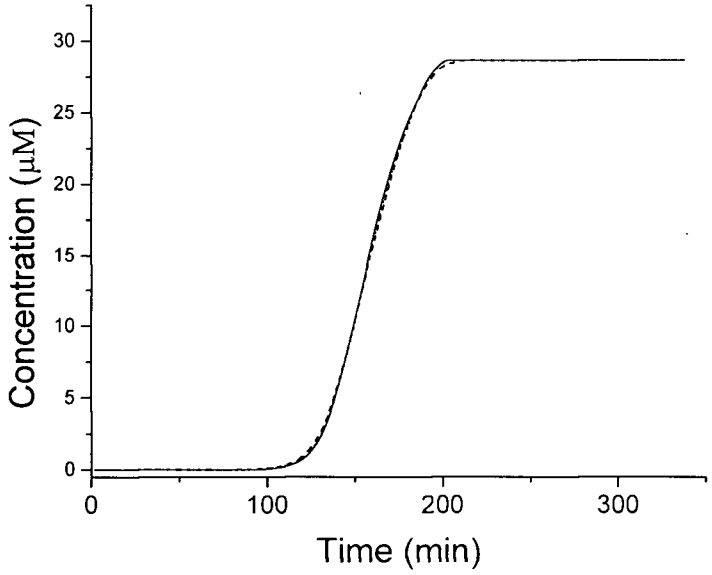


Figure 3.6. Dependence of calculated breakthrough curves on the value of the film mass transfer coefficient, k_f. Initial concentration C_o = 28.6 μM, u = 2.1 x10^{-2} m/s. (. .) 1.2k_f, (---) 0.8k_f , (-) k_f

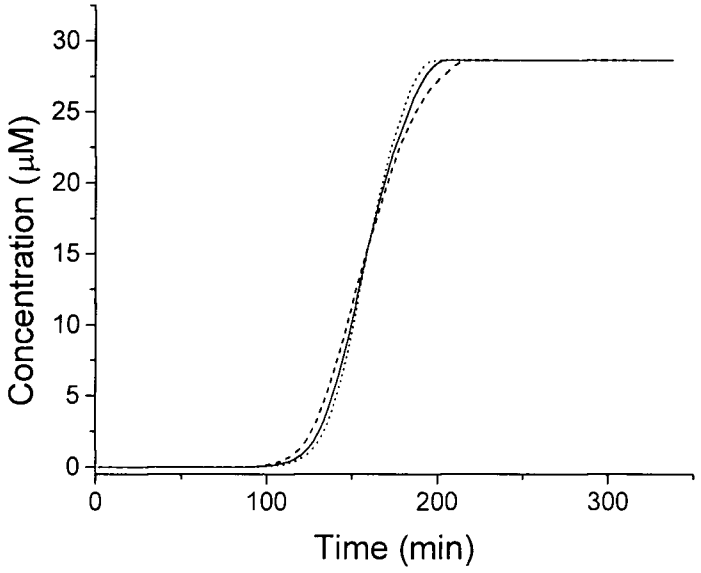


Figure 3.7. Dependence of calculated breakthrough curves on the value of the intrapore diffusivity, D_p. Initial concentration C_o = 28.6 μM, u = 2.1 x10^{-2} m/s. (. .) 1.2 D_p, (---) 0.8 D_p, (-) D_p

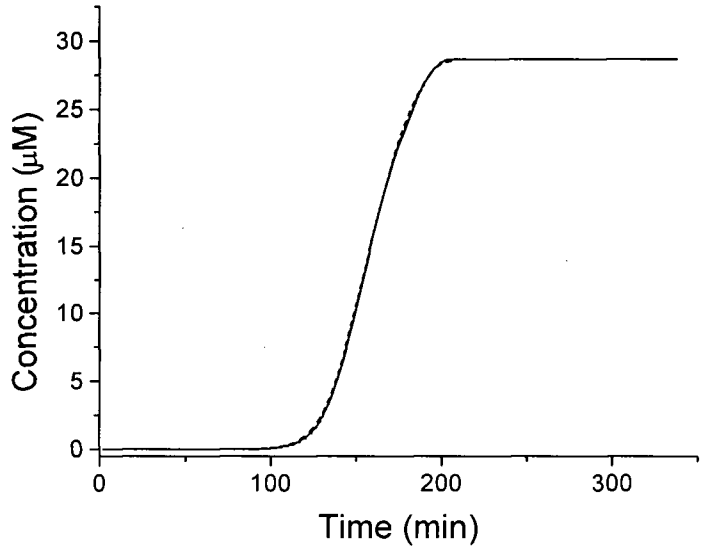


Figure 3.8. Dependence of calculated breakthrough curves on the value of the forward adsorption rate constant, μ_f. C_o = 28.6 μM, u = 2.1 x10^{-2} m/s. (. .) 1.2μ_f, (---) 0.8μ_f, (−)μ_f

Increasing D_l, k_f or μ_f or decreasing in 20 % did not show to affect the calculated breakthrough curves. On the other hand, the calculated breakthrough curve using 20% increasing or decreasing of the D_p value resulted in a difference between the breakthrough curves indicating the dependence of the calculated breakthrough curves on the value of the intrapore diffusivity.

Figure 3.9 compares breakthrough curves calculated when the binding equilibrium model is given by Langmuir theory (Equation (3.25)) or by Equation (3.26) based on the site-exclusion model of Jin *et al.* (1994). To apply the model of Jin *et al.* (1994), a value for α, the overlap factor, must be chosen. Based on the projected surface area of CBD_{Cex} relative to a single cellobiose unit, the fundamental repeating unit of crystalline cellulose, α should be *ca.* 30 (Gilkes *et* al. (1992). Breakthrough curves calculated with Equation (3.26) represent the experimental data more closely than that obtained with Equation (3.25). Both curves compare favorably with

that obtained from the experimental data up to approximately 50% of the curve, after that the symmetric curve obtained from Equation (3.25) deviates from the experimental values.

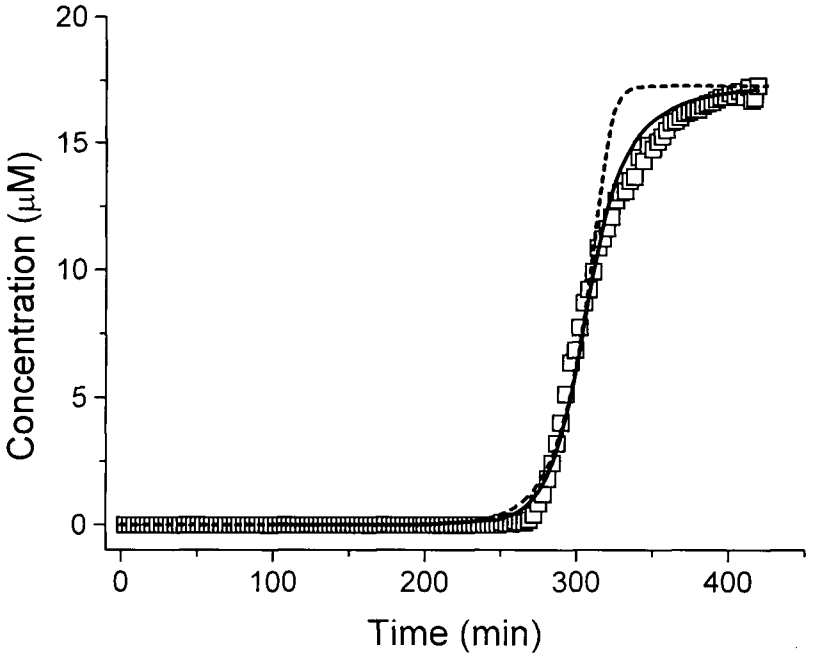


Figure 3.9. Breakthrough curves calculated when the binding equilibrium model is given by Langmuir (----), or by Equation (3.26) (___). Experiment (□); with Initial concentration C_o = 17.3 μM, u = 2.1 x10^{-4} m/s

3.6 Determination of Model Parameters

Application of the numerical model described in section 3.5 to our Avicel-based affinity chromatography column requires values for the model parameters. These parameters include the column void fraction ε, the particle porosity β, the axial dispersion coefficient D_L, the film mass

transfer coefficient k_f, the intrapore diffusivity D_p, and the forward adsorption rate constant μ_f (μ_r is then given by K/μ_f).

3.6.1 Avicel particle porosity (solute-exclusion technique)

Solute exclusion is a well-proven method for determining porosities of fully hydrated chromatographic resins (Lin *et al.*, 1987; Gama *et al.*, 1994). The size-exclusion method is argued to be superior to the more common mercury porosimetry technique because it estimates porosity of the particle under actual column conditions, instead of the dry powder form (Grethlein, 1985). This is particularly true for compressible materials such as cellulose, which tend to exhibit pore-wall collapse during drying such that the dry sample has much less surface area than the wet sample (Stone and Scallan, 1968).

The solute-exclusion technique is based on the bulk-phase concentration increase which occurs when a solute is too large to penetrate the pores of a particulate solid; the pores then contain only solvent. For particles with a distribution of pore sizes, the measured concentration increase, which is inversely related to the porosity, will depend on the size of the solute. As a result, the solute-exclusion experiment is often carried out over a range of solute or 'probe' sizes, with near-monodisperse poly(ethylene glycol) (PEG) standards serving as the probe.

Seven different PEG standards, ranging in size from 200 to 20,000 g mol^{-1}, were used to measure the porosity of sieved Avicel PH101 particles. Light scattering studies have shown that PEG chains in this size range configure in solution as hydrated excluded volume spheroids with molecular diameters of 13 Å to 130 Å, respectively (Haynes *et al.*, 1993; Nelson and Oliver, 1971). In addition, PEG standards are produced in tight molecular-weight distribution and do not

preferentially adsorb to cellulose (Lin *et al.*, 1987), making them a reasonably good choice for exclusion studies.

Each PEG standard was used to make a set of 8 to 10 aqueous probe solutions ranging in concentration from 2 to 10 g L^{-1}. Calibration curves (see Appendix 1) were constructed from refractive index readings (Lin *et al.*, 1987). Dry sieved Avicel PH101 (1 g) was then added to 11 mL of each probe solution and mixed in an orbital shaker for 5 hours. Controls were prepared by mixing the Avicel with distilled water. After equilibration, the suspensions were transferred to test tubes and allowed to settle for 1 hour. Supernatants were collected and centrifuged for 10 min at 5000 rpm (Gama *et al.*, 1994) for further clarification. The final probe concentration C_f, and the associated concentration increase due to solute exclusion, were determined by refractive index (RI) on a Waters HPLC. RI data for three independent samples, with five replicate RI readings per sample, were used to obtain an average C_f for each probe.

Consider then a sample containing p grams of Avicel and q grams of water. The final solute concentration in the equilibrium sample is given by:

$$C_f = \frac{M_p}{W + q - S} \qquad (3.30)$$

where M_p is the grams of PEG standard added, W is the initial volume (L) of solvent added, and q is the volume (L) of water associated with the 'dry' Avicel before mixing. The moisture content in the dry Avicel was determined to be 10.75(± 0.38)% by dry-weight measurements after extensive lyophilization of four independent solid samples. The pore volume accessible to solvent but inaccessible to probe molecules is S (L), the value of which allows the determination of ß in the following way. First, the inaccessible pore volume per unit weight of solid sorbent, I, is calculated

$$I = \frac{S}{p} = \frac{W + q - \frac{M_p}{C_f}}{p} \qquad (3.31)$$

As the diameter of the probe molecule increases, I will reach its maximum value I_{max}. Figure 3.10 plots I as a function of probe diameter. For probe diameters greater than 50 Å, I is at its maximum value of I_{max} = 0.87 (± 0.13) mL g^{-1}. The pore volume V_a accessible to a given probe is calculated by

$$V_a = I_{max} - I \qquad (3.32)$$

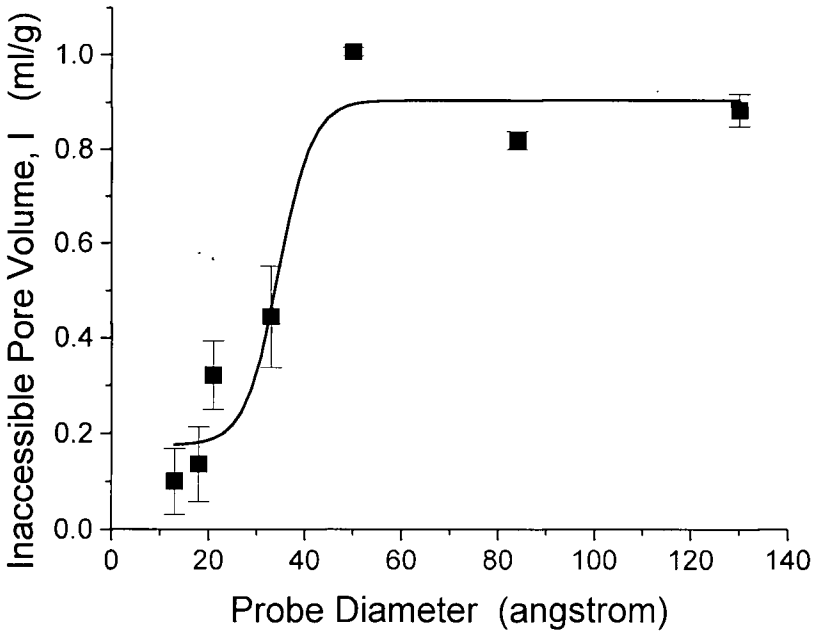


Figure 3.10. The inaccessible pore volume, I, as a function of probe diameter

Finally, the particle porosity ß corresponding to a particular probe diameter is given by V_a multiplied by the density of the particle ρ_p.

The particle density ρ_p is related to the bulk density ρ_b through the fraction of external voids ε in a packed bed (Perry (1984))

$$\rho_p = \frac{\rho_b}{1-\varepsilon} \qquad (3.33)$$

where

$$\rho_b = \frac{p - \rho_w q}{V} \qquad (3.34)$$

ρ_w is the solvent density and V is the bed volume (L).

Figure 3.11 plots porosity as a function of probe diameter for the sieved Avicel PH101 particles used in our affinity column. The porosity decreases sharply from its maximum of *ca.* 0.7 when the probe diameter exceeds 10 Å. The particles are nonporous for probes or proteins with diameters greater than 60 Å, which is in agreement with limited data reported by Gilkes *et al.* (1992). Figure 3.11 indicates that the bead porosity accessible to CBD_{Cex}, which has a mean diameter of 28.5 Å, is 0.45.

To facilitate determination of ß for model calculations, the data in Figure 3.11 was fit to a Boltzmann-type function of the form

$$\beta = \frac{\alpha}{1 + exp(\gamma X - \delta)} \qquad (3.35)$$

where α, γ, and δ are fitted parameters and X is the molecular probe diameter in Å. The solid curve shown in Figure 3.11 represents the best fit, for which α = 0.73, γ = 0.1423, and δ = 4.5861.

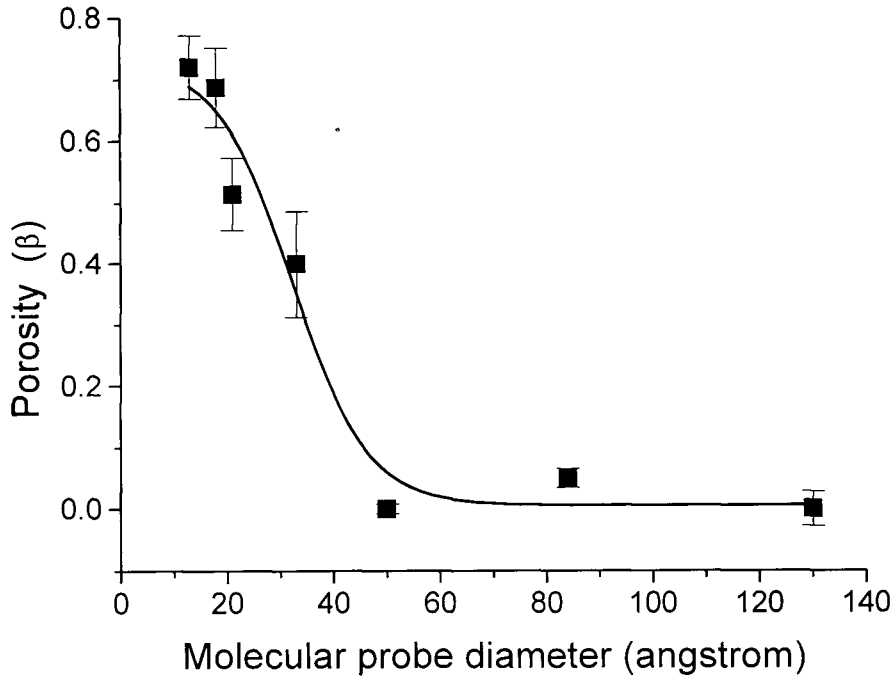


Figure 3.11. Porosity of sieved Avicel as a function of probe diameter. (■) experimental values; (solid line) fitted curve.

3.6.2 Column void fraction (first moment analysis)

First moment analysis provides a convenient method for determining the fraction of external voids ε in a column which generates Gaussian or near Gaussian eluent peak profiles under nonbinding conditions (Villermaux, 1987; Kucera, 1965). Kucera (1965) attempted to solve the chromatography equations listed in section 3.5 for a pulse input of the form

$$c(0,t) = \begin{cases} c_o & 0 \le t \le t_o \\ 0 & t > t_o \end{cases} \qquad (3.36)$$

where t_o is the injection duration and assuming that the binding isotherm is linear or that binding does not occur. The model was transformed into the Laplace domain, but could not be inverted to yield an analytical solution. However, the transformed solution was used to obtain analytical expressions for the first and second peak moments. The first moment (μ_1) of a peak is given by

$$\mu_1 = \frac{\int_0^\infty c(L,t)\, t\, dt}{\int_0^\infty c(L,t)\, dt} \qquad (3.37)$$

Equation (3.37) can be solved under nonbinding conditions for any eluent peak shape. Since ε represents the void of the column but not of the particles, a nonbinding probe molecule must be chosen with sufficiently large diameter to ensure zero penetration into the particle volume. For our analysis, 0.52 μM ferritin, which has a molecular weight of 500 kDa and an average diameter of 84 Å, was used.

For a uniformly packed column operating under plug-flow conditions, Kucera (1965) and Haynes and Sarma (1973) have shown for pulse-injection loading conditions that the retention time t_R, which corresponds to the solute peak maximum of a Gaussian eluent peak, of a nonbinding nonpenetrating solute is equal to the first moment μ_1. Under such conditions, where the width of the injected sample is infinitesimally small compared to L, solution of Equation (3.37) yields an equation for μ_1 which is inversely proportional to u, the superficial velocity,

$$\mu_1 = t_R = \frac{L}{u}\left(\varepsilon + (1-\varepsilon)\beta\right) \qquad (3.38)$$

where L is the column length. A plot of μ_1 versus $L\, u^{-1}$ should therefore yield a line with a slope $[\varepsilon+(1-\varepsilon)ß]$, from which the void fraction ε can be computed.

Retention times were measured for ferritin pulses injected into our packed Avicel column as a function u over the range 6.37×10^{-3} cm s^{-1} to 2.54×10^{-2} cm s^{-1}. Figure 3.12 plots the data according to Equation (3.38). The slope of the resulting line is 0.60 (± 0.04), from which a value for ε of 0.60 (± 0.04) was determined since ß is equal to zero.

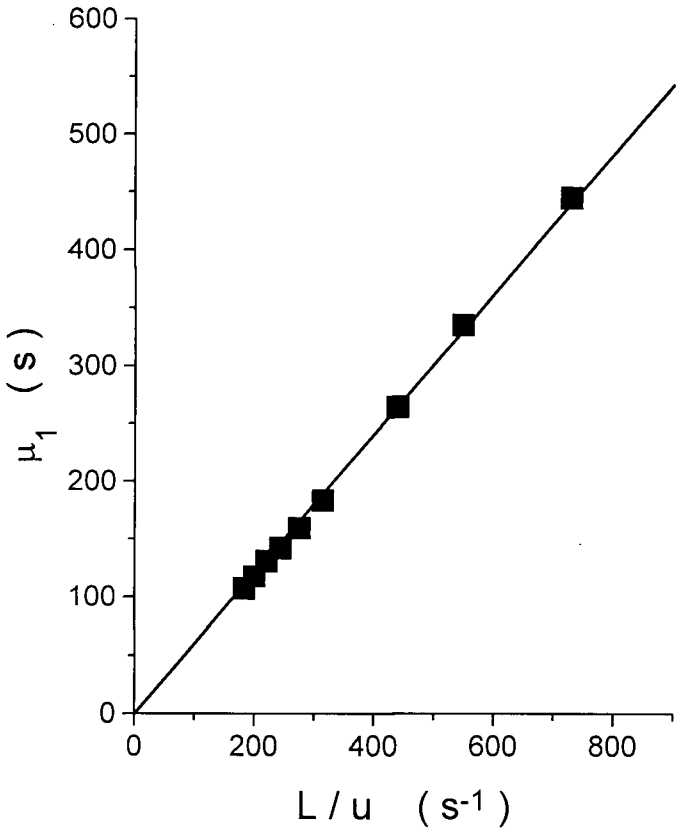


Figure 3.12. Retention times measured for ferritin pulses. Range of u from 6.37×10^{-3} to 2.54×10^{-2} cm s^{-1} .

3.6.3 On column equilibrium binding constant and sorbent capacity

The equilibrium constant K_a and, in particular, the sorbent capacity Q_{max} characterizing solute binding within a chromatographic column can differ from values measured in batch suspension due to, for instance, steric blockage of potential binding surface in the tightly packed column environment. First moment analysis, when combined with frontal-loading data for the target solute, provides a convenient method for determining equilibrium binding properties within the column.

In the frontal-loading (or breakthrough) experiment, a mobile phase containing a fixed solute (*e.g.*, CBD_{Cex}) concentration is continuously fed through the column so that the average elution volume V_e corresponds to an elution time of t_b, which is equal to the first moment μ_1 calculated from the first derivative of the breakthrough curve. In affinity chromatography, the partitioning of a target protein between the mobile and solid phases is governed by the strength of binding of the solute to the matrix. Larger elution volumes are therefore observed for strong binding solutes. To account for this effect, Dunn and Chaiken (1974) separated V_e into a sum of two terms.

$$V_e = V_o + \frac{s}{C_o} V_s \qquad (3.39)$$

where V_o is the column void volume, V_s is volume of sorbent particles in the column, C_o is the concentration of solute in the feed, and s is the total solute concentration in and on the sorbent particles. The first term on the right hand side of Equation (3.39) accounts for the volume of fluid $V\varepsilon$ which must be displaced from the column void before breakthrough of a nonbinding, nonpenetrating solute would be observed. The second term accounts for the additional volume

of mobile phase which must pass through the column to elute a solute which can partition into the solid phase and bind specifically to the sorbent surface.

We now specify a flowrate for the feed solution which is slow enough to allow attainment of equilibrium everywhere behind the solute front. The concentration of solute within the pores of the sorbent particles will then be equal to C_o. Application of Equations (3.14) and (3.19) then allows us to transform Equation (3.39) into an equation which explicitly relates V_e to the feed concentration C_o, the binding constant K_a (or K_L) and the sorbent capacity Q_{max}. For instance, if we assume binding is well described by Langmuir theory (Equation (3.25)), Equation (3.39) becomes

$$\frac{V}{V_e - [\varepsilon + (1-\varepsilon)\beta]\, V} = \frac{1}{(1-\varepsilon)\rho_p\, Q_{max}}\, c + \frac{1}{(1-\varepsilon)\rho_p\, Q_{max}\, K_L} \qquad (3.40)$$

where V is the total volume of the column.

Figure 3.13 plots elution volume data for CBD_{Cex} as a function of feed concentration according to form of Equation (3.40). The slope of the least-squares linear fit yields a value for Q_{max} of 3.1 mole m^{-3}, and the intercept a value for K_L of 9470 m^3 mol^{-1}. These results compare favorably with the Q_{max} and K_L values of 2.9 mol m^{-3} and 9500 m^3 mol^{-1}, respectively, determined from batch adsorption-isotherm experiments (see section 3.4). This indicates that flow patterns and packing in the packed Avicel column are sufficiently uniform to allow the CBD_{Cex} solute to access all available sorbent surface area.

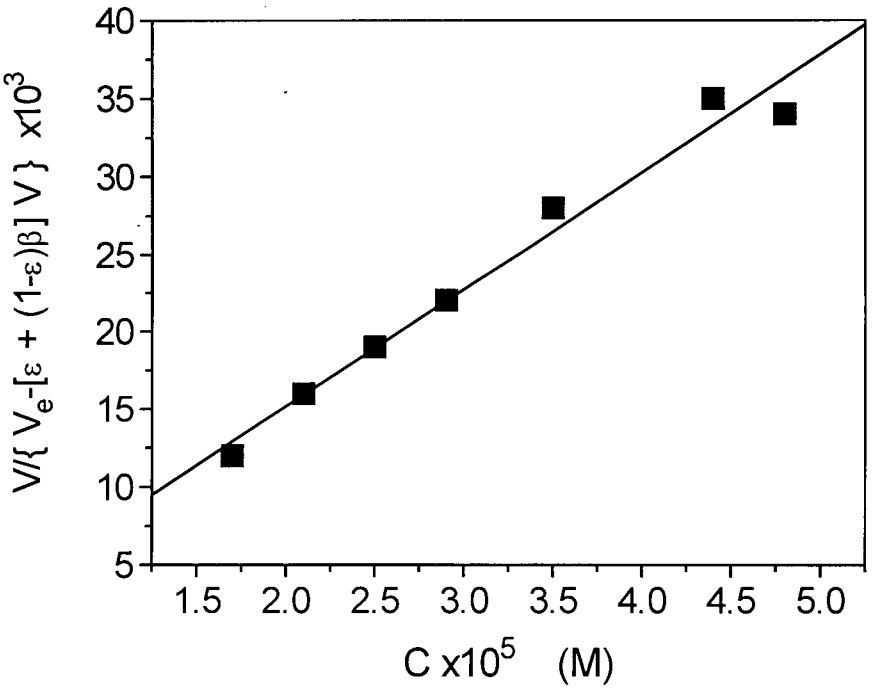


Figure 3.13. Elution volume data for CBD_{Cex} as a function of feed concentration. (■) Experimental values; (line) linear fit. y = 7.556 x + 0.08 (R = 0.987)

3.6.4 Solute mass-transfer effects (second moment analysis)

The second central moment μ_2 of an eluent peak is given by (Kucera, 1965)

$$\mu_2 = \frac{\int_0^\infty c(L,t)\,(t-\mu)^2\,dt}{\int_0^\infty c(L,t)\,dt} \qquad (3.41)$$

Under nonbinding solute conditions, Ruthven (1984) and Haynes and Sarma (1973) have solved the chromatographic model equations shown in section 3.5 for μ_2

$$\mu_2 = \frac{2L\varepsilon}{u}\left\{\frac{D_L}{u^2}\left(\varepsilon+(1-\varepsilon)\,\beta\right)^2+\left(\frac{1-\varepsilon}{\varepsilon}\right)\beta^2\,\frac{R_p^{\,2}}{15}\left(\frac{1}{\beta D_p}+\frac{5}{k_f R_p}\right)\right\} \qquad (3.42)$$

where R_p is the average particle radius. Thus, the second central moment μ_2 (a measure of peak spreading) depends on all mass-transfer resistances in the column, including axial dispersion, film mass transfer, and intrapore diffusion.

For Gaussian-shaped peak profiles, μ_2 is equal to the standard peak variance σ^2, which is given by the peak width at $e^{-1/2}$ times the peak height. Villermaux (1987) has shown that the equality of μ_2 and σ^2 also holds with reasonable accuracy for slightly non-Gaussian peak shapes. Under such conditions, moments analysis can be used to calculate the height equivalent to a theoretical plate (HETP), a useful gauge of column performance, according to the classic chromatography theory of van Deemter *et al.*, (1956). According to van Deemter *et al.* (1956), one HETP is equal to

$$\mathrm{HETP} = \frac{\sigma^2 L}{t_r^{\,2}} = \frac{2\varepsilon D_L}{u} + 2u\left[\frac{R_p^{\,2}}{15}\left(\frac{1}{\beta D_p} + \frac{5}{k_f R_p}\right)\frac{(1-\varepsilon)\beta^2}{\left(\varepsilon + (1-\varepsilon)\beta\right)^2}\right] \qquad (3.43)$$

$$= \frac{2D_L}{u_o} + 2u_o\left(\frac{1}{K_m}\right)\left(\frac{\varepsilon}{1-\varepsilon}\right)\left(1 + \frac{\varepsilon}{(1-\varepsilon)\beta}\right)^{-2}$$

where K_m is the overall mass transfer coefficient (s^{-1}), given by

$$\frac{1}{K_m} = \frac{R_p^{\,2}}{15\beta D_p} + \frac{R_p}{3k_f} \qquad (3.44)$$

and u_o (= u/ε) is the interstitial fluid velocity in the column.

Equation (3.43) can be linearized with respect to u_o^2

$$u_o \times \mathrm{HETP} = \frac{u_o \sigma^2 L}{t_R^{\,2}} = 2D_L + 2u_o^{\,2}\left(\frac{1}{K_m}\right)\left(\frac{\varepsilon}{1-\varepsilon}\right)\left(1 + \frac{\varepsilon}{(1-\varepsilon)\beta}\right)^{-2} \qquad (3.45)$$

to allow for determination of the axial dispersion coefficient D_L and the overall mass transfer coefficient K_m from the intercept and slope, respectively, under nonbinding sorbent-penetrating

conditions. To apply Equation (3.45) to the determination of the parameters characterizing mass-transfer resistances encountered by CBD_{Cex} and CBD_{Cex}-fusions, we must identify nonbinding probe molecules with structural properties similar to, for instance, CBD_{Cex}.

Horse heart myoglobin (MW 17,800 Da) is a globular protein similar in size and shape to CBD_{Cex} which shows no binding affinity for particulate Avicel (Jervis *et al.*, 1997). Figure 3.14 shows chromatogram peaks as a function of u for 200-μL 15.8-μM pulses of myoglobin loaded onto a uniformly packed 14.2 cm^3 column of sieved Avicel PH101. Both the mobile phase and the myoglobin pulse contained 50-mM phosphate buffer (pH 7). The column was operated at 20°C. At low flowrates, the eluent peaks are Gaussian, and they remain near-Gaussian at high superficial velocities.

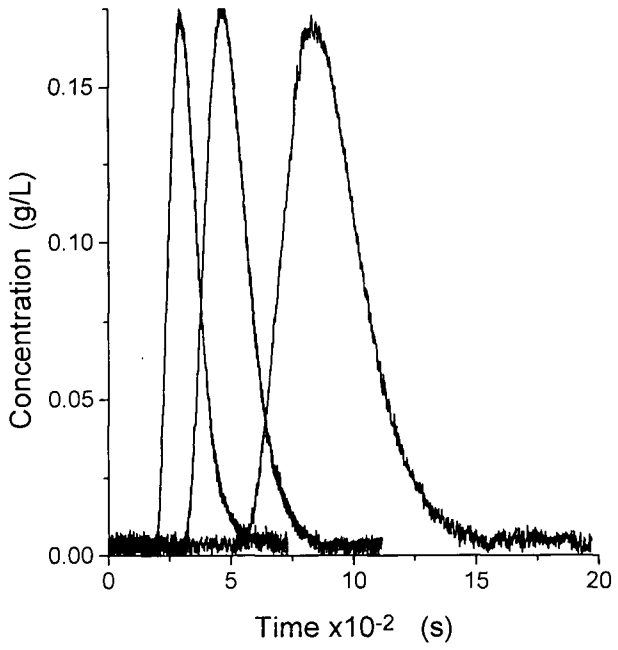


Figure 3.14. Chromatogram peaks of myoglobin pulses. Range of u from 6.37×10^{-3} to 2.54×10^{-2} cm s^{-1}.

Figure 3.15 plots first and second moments data regressed from pulse-injection chromatographic peaks for myoglobin in a form consistent with Equation (3.45). The resulting linear function confirms the applicability of van Deemter HETP theory and allows for regression of D_L and K_m from the intercept and slope, respectively. Linear least-squares regression (R=0.9918) gives values of $D_L = 8.0 \times 10^{-8}$ m^2 s^{-1} and $K_m = 4.62 \times 10^{-2}$ s^{-1}.

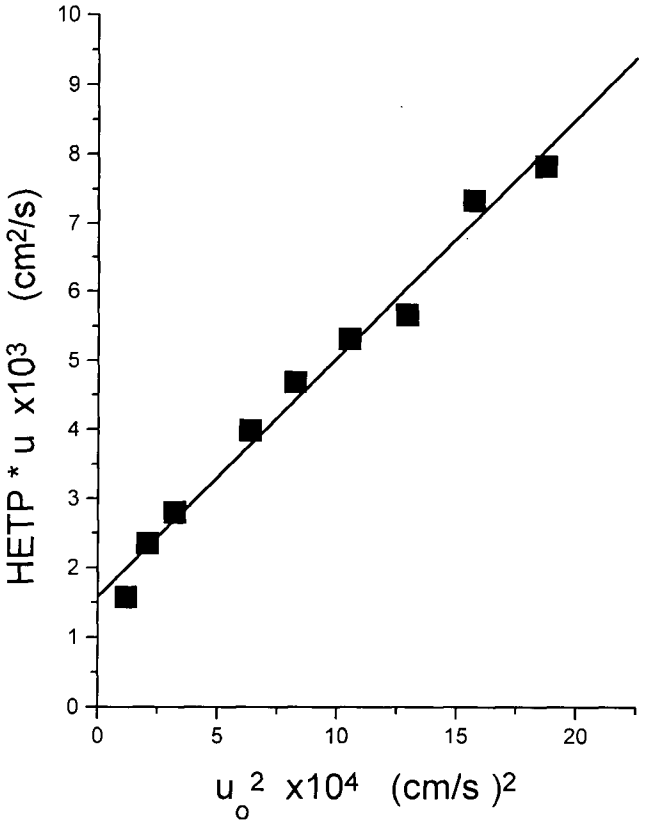


Figure 3.15. First and second moments data regressed from pulse-injection chromatographic peaks for myoglobin

As shown in Equation (3.44), the overall mass transfer coefficient K_m contains contributions from film and intrapore mass transfer resistances. Segregation of K_m into its film and intrapore mass transfer components requires an appropriate correlation for determining the value of either k_f or D_p. For spherical particles, good correlations for determining k_f are available (Wakao *et al.,* 1958; Foo and Rice, 1975). Foo and Rice (1975), for instance, derived and verified the following dimensionless correlation for low Reynolds number flow:

$$Sh = 2 + 1.45\, Re^{0.5}\, Sc^{0.33} \qquad (3.46)$$

where *Sh* is the Sherwood number, given by

$$Sh = \frac{k_f 2R_p}{D_m} \qquad (3.47)$$

Re is the Reynolds number, given by

$$Re = \frac{2R_p u}{\nu} \qquad (3.48)$$

and *Sc* is the Schmidt number, given by

$$Sc = \frac{\nu}{D_m} \qquad (3.49)$$

Solution of Equation (3.46) for k_f requires a value for the diffusion coefficient D_m of CBD_{Cex} in free solution. The Polson equation (Polson, 1950; Tyn and Gusek, 1990) provides simple method for accurately estimating free-solution diffusion coefficients of globular proteins of molecular weight greater than 1 kDa

$$D_m = \frac{A}{M^{\frac{1}{3}}} \qquad (3.50)$$

where A is a constant of value 2.85 x 10^{-5} cm^2 s^{-1} $g^{1/3}$ $mol^{1/3}$, and M is the molecular weight of the protein (g mol^{-1}). The solution diffusivity D_m for CBD_{Cex} in water is estimated from Equation (3.50) to be 12.8 x 10^{-11} m^2 s^{-1}. The correlation of Foo and Rice (Equation (3.46)) then gives an estimated value for k_f of 7.8 x 10^{-6} m s^{-1} and, from Equation (3.44), a value for D_p of 2.2 x 10^{-11} m^2 s^{-1}. This estimated value for D_p is consistent with a number of intrapore diffusivities reported in the literature for similar protein, particle and pore sizes (Carlsson *et al.*, 1994; Arnold *et al.*, 1985b).

3.6.5 Forward binding rate constant

Mass transfer resistances associated with solute diffusion to the particle and through its pores preclude independent measurement of the rate constant μ_f for binding of CBD_{Cex} to the surface of Avicel. Film and intrapore diffusion are generally (in fact almost always) rate-limiting in chromatographic processes involving nanoporous, micron or larger sized sorbent particles (Jungbauer, 1996). As a result, binding to the pore wall is rapid relative to diffusion through the film and into the pore, and local solute-sorbate equilibrium is established everywhere within the pore volume as described by Equation (3.25) or (3.26). Binding therefore depends on the equilibrium association constant but not on μ_f. Hall *et al.* (1976) have shown that the assumption of local solute-sorbate equilibria is exact when $C_O > 1/K_a$, which is true in all cases reported here since $K_a \approx 10^7 >> 1$. The local equilibria assumption is therefore made of all model calculations reported in section 3.7 and in Chapter 4.

Although it is not required for model calculations, an approximate value for μ_f can be obtained by fitting the model derived in section 3.5 with binding kinetics described by Equation (3.21) to an experimental breakthrough curve for CBD_{Cex}. As described in section 3.6.1 to 3.6.4, all remaining model parameters have been independently determined either by experiment or correlation. Figure 3.16 shows the best fit of the model when only μ_f is adjusted to breakthrough-curve data for a 43.8 x 10^{-3} mol m^3 feed solution of CBD_{Cex} on our 14.2 cm^3 column of Avicel. The resulting estimate for μ_f is 0.21 m^3 mol^{-1} s^{-1}, which is consistent with earlier studies of Gilkes *et al.* (1992) which indicate that CBD_{Cex} binds rapidly to crystalline cellulose.

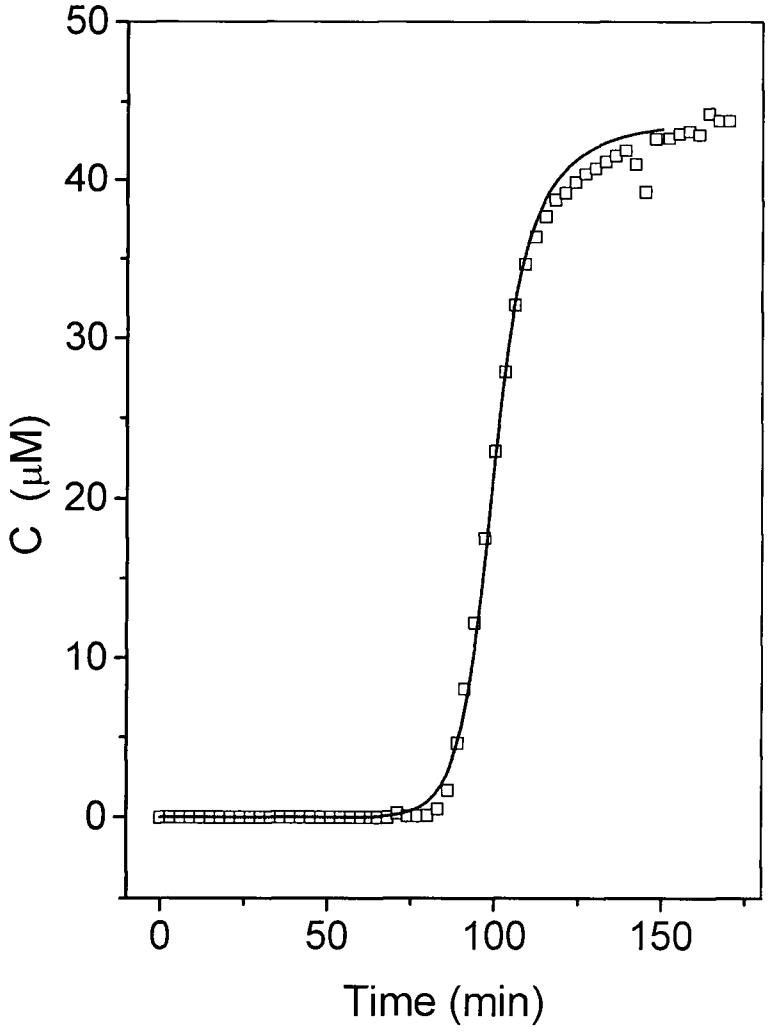


Figure 3.16. Best fit of the model. Estimated $\mu_f = 0.21$ m^3 mol^{-1} s^{-1} (□) Experimental data,; (solid line) Model regression. ($C_o = 43.8x10^{-3}$ mole/m^3).

3.7 Model Results and Validation

Table 3.2 provides a summary of all measured or calculated column properties and parameters required for solution of the model described in section 3.5 for loading of CBD_{Cex} onto our Avicel-packed column. For completeness, the estimated value for the association rate constant μ_f has been included but is not strictly required since local equilibrium is assumed. All remaining parameters listed in Table 3.2 were obtained from independent experiment (*i.e.*, not from breakthrough curve data for CBD_{Cex}). Validation of the predictive capabilities of the model, and any potential simplifying assumptions which can be made to ease computation, can therefore proceed by comparison with breakthrough data for CBD_{Cex} over a range of realistic operating conditions.

Table 3.2: Measured or calculated properties and parameters describing the packed Avicel column and the binding of CBD_{Cex} to the column.

Column length, L	variable
Empty bed volume, V	3.54 ($\pm$ 0.24) x10^{-6} m^3
Particle density, ρ_p	0.93 ($\pm$ 0.05) x10^3 Kg/m^3
Bead radius, R_p	5.34 ($\pm$ 2) x10^{-5} m
Void fraction , ε	0.60
Bead porosity, ß	0.45
Langmuir equilibrium binding constant, K_L	
Frontal analysis	9470 m^3/mole
Binding Isotherms	9500 m^3/mole
Maximum capacity, Q_{max}	
Frontal analysis	3.14 mole/m^3
Binding Isotherms	2.86 mole/m^3
Axial dispersion coefficient, D_L	8.0 x10^{-8} m^2/s
Film mass transfer coefficient, k_f	7.83 x10^{-6} m/s
Diffusion coefficient in bulk liquid, D_m	12.78 x10^{-11} m^2/s
Diffusion coefficient in pore liquid, D_p	2.2 x 10^{-11} m^2/s
Forward rate constant, μ_f	0.2136 m^3/moles

Before making this comparison, it is instructive to differentiate between the rate of solute mass transfer through the stagnant fluid film and that through the pores of the sorbent particles. In column chromatography, the Biot number (B_i)

$$B_i = \frac{k_f R_p}{3D_p} \qquad (3.51)$$

defines the dimensionless ratio of the rate of film mass transfer to intraparticle diffusion. A Biot number of the order of 100 or greater indicates that intraparticle diffusion limits the binding process; a Biot number of the order of 0.01 or less indicates that the rate of film mass transfer limits the process. Intermediate values indicate that both effects are important. The sieved Avicel-packed affinity chromatography column used in these studies is characterized by a Biot number of *ca.* 6, indicating that the rates are similar and both mass transfer resistances are important when ß is reasonably large (*i.e.* when the solute size is small). As the solute size increases, ß becomes vanishingly small (see Figure 3.11) and the percentage of solute bound within the sorbent pores relative to that bound on the particle exterior becomes negligible. Under these conditions, the effect of film mass transfer on the breakthrough curve becomes dominant.

As shown in Figure 3.11, the total voidage of the particulate cellulose used is 0.73, which is consistent with previous literature results (Gamma *et al.*, 1994). The particle voidage available to a solute of a given size can be determined from the fractional area beneath the porosity versus solute diameter plot shown in Figure 3.11. The derivative of the Boltzmann fitting function (Equation 3.35) for the particle porosity data is given by :

$$\frac{d\beta}{dX} = -\frac{\alpha\gamma e^{(\gamma X-\delta)}}{\left(1+e^{\gamma X-\delta}\right)^2} \qquad (3.52)$$

Equation (3.52), which is plotted in Figure 3.17, gives the pore size distribution of the sieved Avicel particles. It has a fairly broad pore size distribution which is centered around an average pore diameter of *ca.* 30 Å.

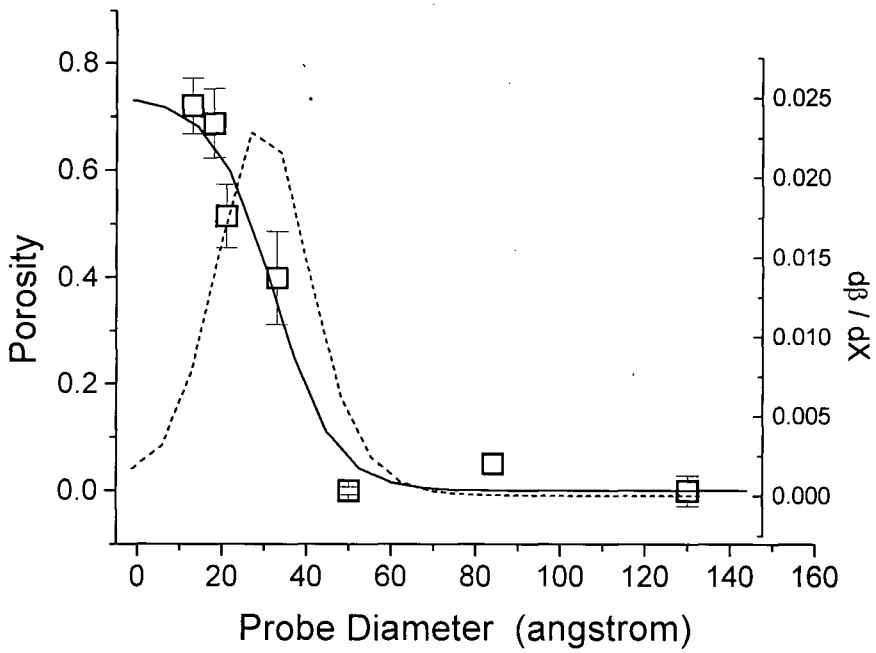


Figure 3.17. Pore size distribution of the sieved Avicel particles. The doted line represents the derivative of the Boltzmann fitting function for the particle porosity data.

CBD_{Cex} has an average diameter of 29 Å, which is near the average pore diameter of the semicrystalline Avicel particles used. As a result, over 40% of the pore volume is inaccessible to CBD_{Cex} and thereby severely limits the intraparticle surface area available for binding. The small pore sizes of the Avicel particles therefore suggest that most of the CBD_{Cex} binds to or near the exterior surface of the Avicel particles. For fusion proteins of average diameter greater than *ca.* 60 Å, essentially all of the binding will occur on the particle exterior.

3.7.1 Comparison of Binding Equilibrium Models

In section 3.5.2, we proposed two possible models for describing equilibrium binding isotherms of CBD_{Cex} and CBD_{Cex}-fusions on particulate Avicel: (1) the Langmuir binding isotherm (Equation (3.25)) and the random sequential adsorption (RSA) isotherm model (Equation (3.26)) of Jin *et al.*, (1994). It is well known that application of the Langmuir equation leads to prediction of symmetric breakthrough curves (see for instance Jungbauer (1996)). The effect of the RSA model on breakthrough curve analysis is the prediction of band broadening near column saturation conditions. Figure 3.18 shows an experimental breakthrough curve for a 43.8 x 10^{-3} mol m^{-3} feed of CBD_{Cex} flowing through a 14.2 cm^3 column of Avicel at 20°C, pH 7 and a flowrate of 1 mL min^{-1}. This breakthrough curve is representative of all of the observed data for the Avicel affinity column. It shows a modest but significant asymmetry such that the leading edge of the breakthrough curve is sharper than the approach to saturation. Such asymmetry cannot be modeled with the Langmuir equation.

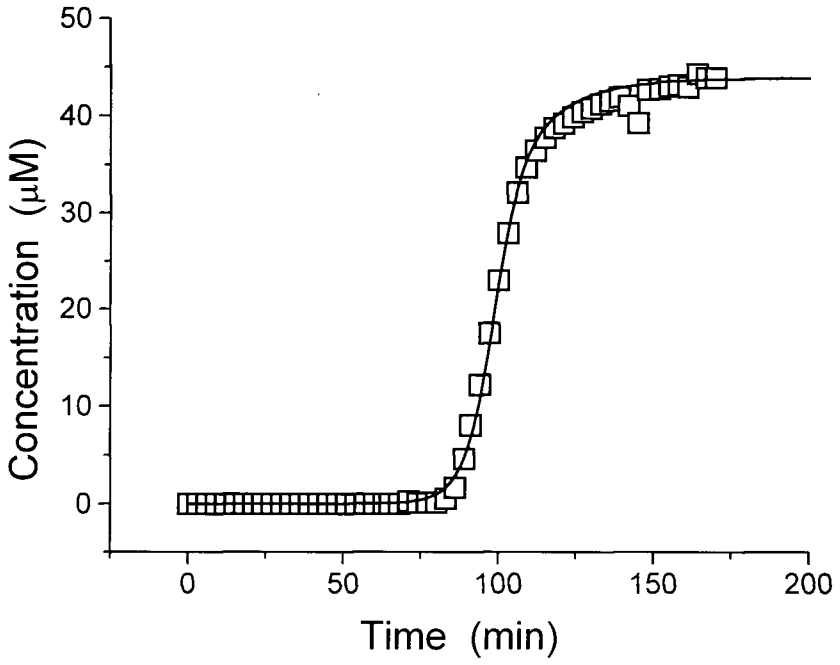


Figure 3.18. Experimental breakthrough curve for a 43.8 μM feed of CBD_{Cex}, flowrate 1 mL min^{-1}, pH 7, 20°C. Solid line; model prediction

As shown in Figure 3.18, when incorporated into the column model derived in section 3.5, the RSA model does a reasonably good job of describing the asymmetric breakthrough curve measured for frontal loading of CBD_{Cex}. In the model results shown, the overlap parameter α which specifies the average number of overlapping sites covered by a bound CBD_{Cex} molecule was set to 10 in an effort to best represent the shape of the breakthrough curve. The equilibrium form of the RSA model (Equation (3.26)) was used in all column simulations.

The sigmoidal shape of the breakthrough curve is the result of a complex interplay of equilibrium and non-equilibrium processes. Steep breakthrough curves are often observed in chromatography of small solutes, where binding kinetics and rates of diffusion are fast. Broader breakthrough curves are usually observed in protein affinity chromatography due to slower

binding and diffusion. All of these effects are captured in the model equations described in section 3.5. Distortion of the breakthrough curve symmetry however, is only accounted for in the RSA binding model. In reality, this distortion can be due to a number of effects, such as nonuniform mixing in some regions of the bed, non-specific adsorption, and conformational changes in the adsorbed protein (Johnson *et al.*, 1990). Thus, it may be more accurate to view α, the overlap factor, as an adjustable parameter which corrects for breakthrough curve distortion.

3.7.2 Predicted breakthrough curves for changing feed concentration

Changes in feed concentration of the adsorbing solute provide a fairly stringent test of our model since gradients in mobile-phase concentrations and across the fluid film dictate the driving force toward adsorption equilibrium. Figure 3.19 shows measured breakthrough curves for frontal loading of CBD_{Cex} as a function of feed concentration. Increasing the feed concentration decreases elution time and increases the sharpness of the breakthrough curve as would be expected for a self-sharpening function where kinetic and mass transfer events are balanced. For protein affinity chromatography columns, the breakthrough curves shown in Figure 3.19 are relatively sharp, indicating uniform flow and binding characteristics (Johnston *et al.,* 1990).

Calculated breakthrough curves are also shown in Figure 3.19. The model accurately predicts the average time of breakthrough and closely mirrors the shape of the breakthrough curve. There is a small departure from experiment when saturation is approached, indicating that in addition to binding-site overlap, non-equilibrium effects such as nonuniform flow contribute to breakthrough curve asymmetry. Previous models for affinity chromatography tend to show much larger deviations from the experimental data (Liapis *et al.*, 1989; Mao *et al.*, 1991). In

general, they underpredict the time of column breakthrough and saturation, a phenomena which is commonly attributed to steric hindrance effects which are accounted for in the present model.

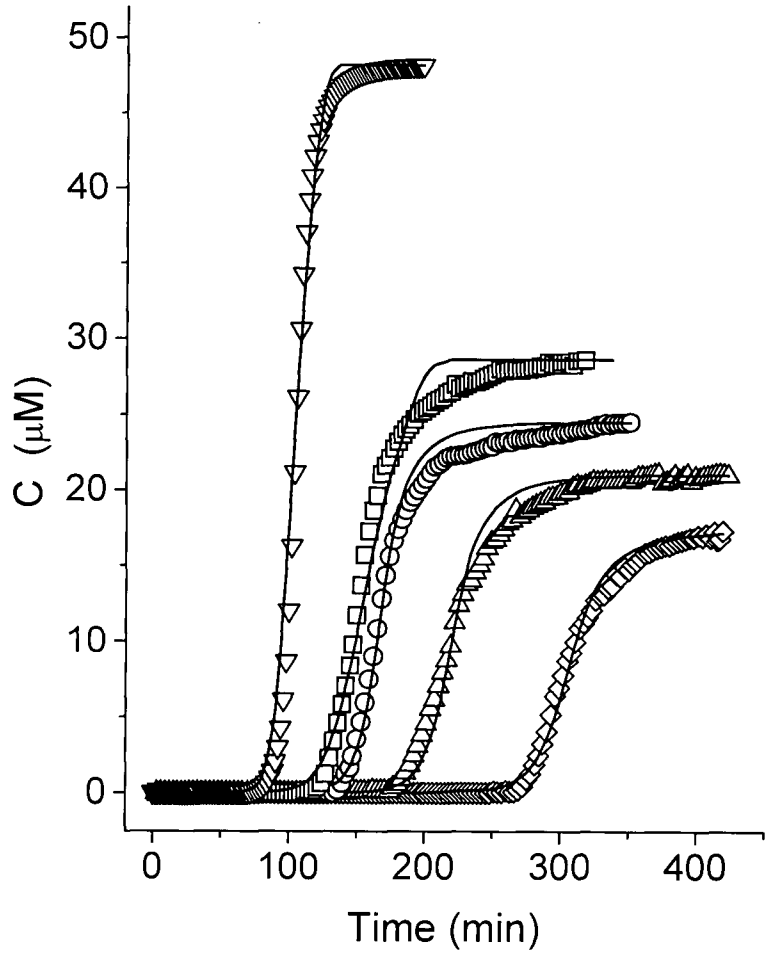


Figure 3.19. Measured breakthrough curves for frontal loading of CBD_{Cex} as a function of feed concentration. (◊) 17.3 $x10^{-3}$ mole/m^3, (Δ) 21 $x10^{-3}$ mole/m^3, (O) 24.5 $x10^{-3}$ mole/m^3, (□) 28.6 $x10^{-3}$ mole/m^3, (▽) 48.2 $x10^{-3}$ mole/m^3; (solid lines) Model predictions.

Although complete agreement between the model and experiment is desired, one is generally most interested in the accurate prediction of the onset of breakthrough since that will dictate the time at which the sample loading stage should be terminated. The loading process is usually stopped before or at the onset of breakthrough to avoid excessive loss of product.

3.7.3 Predicted breakthrough curves for changing operating conditions

Changes in process parameters such as flowrate and column length/volume can significantly affect column performance, particularly when packing or flow nonuniformities are present. Avoiding and identifying the introduction of column irregularities is a critical element of successful column scale-up. By comparing experimental breakthrough curves for new column conditions with model predictions, it is possible to analyze the degree to which these changes have affected column performance (McCoy and Liapis, 1991; Liapis, 1990).

Regrettably, the time investment required to produce large quantities of CBD_{Cex} limited our ability to test the column over a wide range of operating conditions and scales. Based on our on-column binding isotherm data shown in Table 3.2, a 1 L sieved Avicel affinity column will bind *ca.* 34 g of CBD_{Cex}. This is a prohibitively large amount of protein since production of CBD_{Cex} was not the focus of this thesis. However, sufficient protein was produced to allow us to evaluate column performance over a modest range of operating conditions and scales.

Figure 3.20 shows experimental breakthrough curves for a feed concentration of 28.6 mol m^{-3} CBD_{Cex} at two different superficial velocities: 2.11 x 10^{-4} m s^{-1} (1 mL min^{-1}) and 1.06 x 10^{-4} m s^{-1} (0.5 mL min^{-1}). Slower mobile-phase velocities result in a slight broadening of the breakthrough curve near saturation conditions which is indicative of flow nonuniformities

leading to small convective dead-zones in the column. Saturation of such dead zones can only occur through diffusion.

A decrease in interstitial velocity can result in an increase in the thickness of the fluid-film boundary layer around each particle, and thus a decrease in k_f. The change in k_f accompanying a change in column hydrodynamics can be determined from Equation (3.46), and the resulting value used in model calculations for the specified flow conditions. Figure 3.20 shows predicted breakthrough curves as a function of u for a CBD_{Cex} feed concentration of 28.6 mol m^{-3}. The range of simulated velocities covers the realistic range of flowrates (1.2 mL min^{-1} to 0.4 mL min^{-1}) for this systems. Two of the simulation conditions, u=2.11 x 10^{-4} m s^{-1} and 1.06 x 10^{-4} m s^{-1}, match those used in the two experimental breakthrough curves. At both conditions, the predicted and experimental breakthrough curves are reasonably well matched, particularly with respect to the time of initial breakthrough. It should be noted that the two breakthrough curve measurements were performed on two independently packed columns. The more pronounced broadening of the breakthrough curve measured at u= 1.06 x 10^{-4} m s^{-1} may therefore be due to packing nonidealities. Nevertheless, excess broadening near saturation is also observed in the high-flow column relative to that predicted by the model, indicating that the current method of preparing the Avicel column tends to result in unwanted flow and packing nonidealities.

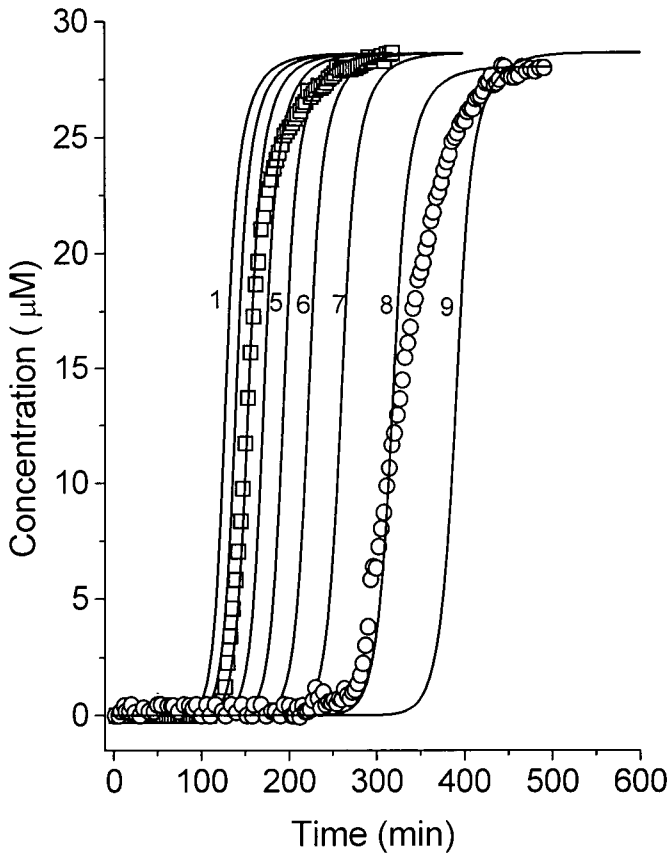


Figure 3.20. Model prediction and experimental curves at different superficial velocities. (□) Experimental data (C_o = 28.6 x10^{-3} moles/m^3) ; (lines) Model predictions. (1) 2.54 x10^{-4} ; (2) 2.33 x10^{-4}; (3) 2.11 x10^{-4}; (4) 1.91 x10^{-4}; (5) 1.69 x10^{-4}; (6) 1.48 x10^{-4}, (7) 1.27x10^{-4}; (8) 1.06 x10^{-4}, (○) Experimental data (C_o = 28 x10^{-3} moles/m^3); (9) 8.48 x10^{-5} m/s.

Nonuniform packing effects, and the resulting disturbances in flow distribution, are known to be a strong function of bed length (Mao *et al.*, 1991). In Figure 3.21, experimental breakthrough curves at two different column lengths are compared directly by normalizing them against their first moment μ_1. Differences in the shapes of the breakthrough curves are small, indicating that flow and packing nonidealities are relatively small.

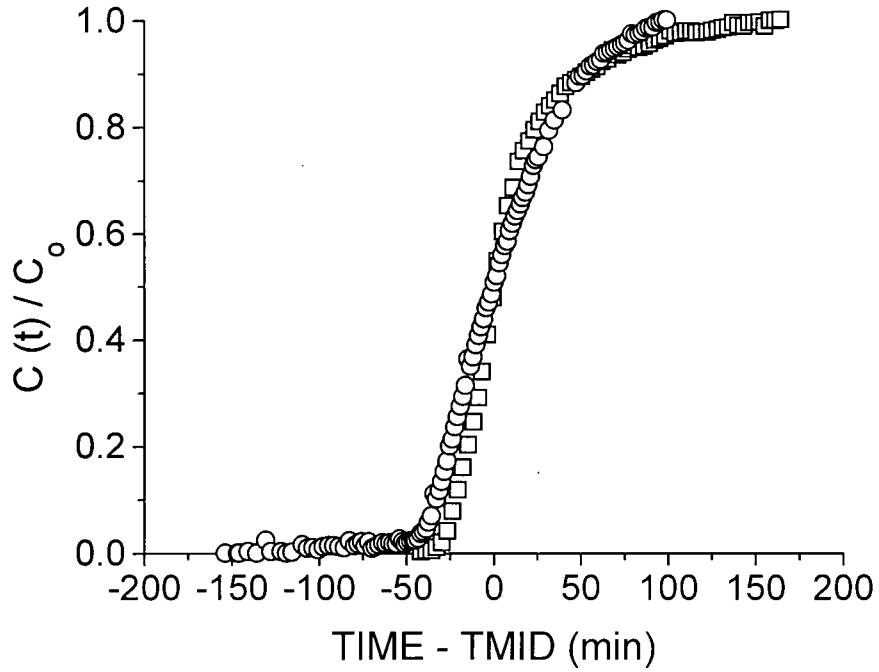


Figure 3.21. Experimental breakthrough curves obtained from different bed length. TMID corresponds to the time at 50% breakthrough, the midpoint, after subtracting the column void volume. (□) 4.7 x10^{-2} m; (O) 6.7 x10^{-2} m.

Figure 3.22 compares predicted breakthrough curves as a function of column length with experiment. Once again, the model calculations agree well with experiment, indicating that the model provides an effective method for column simulation and scale-up. For example, the model can be used to predict the dependence of loading time on column size and geometry. A 4.4-fold increase in column length results in a 6-fold increase in the volume of feed which can be loaded before breakthrough.

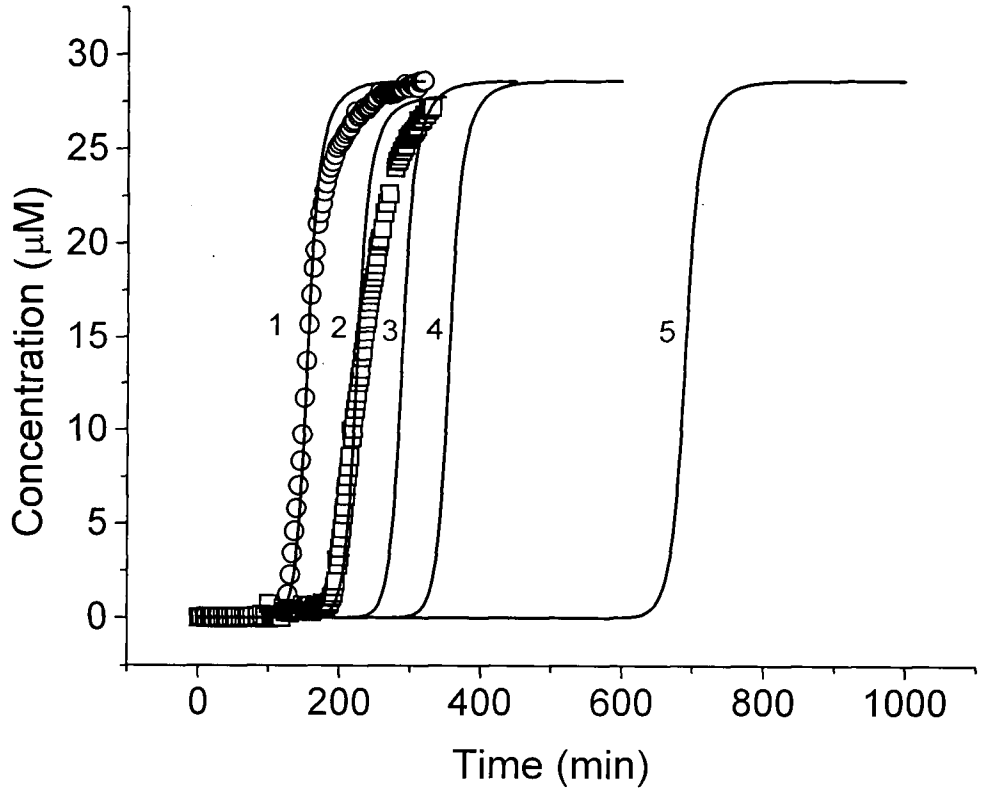


Figure 3.22. Model prediction and experimental curves using different column length. (points) Experimental data, $C_o = 28.6 \times 10^{-3}$ mole/s; (lines) Model predictions. (1) 4.7×10^{-2} m; (2) 6.7×10^{-2} m, $C_o = 28 \times 10^{-3}$ mole/m^3; (3) 8.7×10^{-2} m; (4) 10.7×10^{-2} m; (5) 20.7×10^{-2} m.

CHAPTER 4

APPLICATION OF THE SIEVED AVICEL COLUMN TO THE AFFINITY PURIFICATION OF A STAPHYLOCOCCAL PROTEIN A - CBD_{CEX} FUSION

4.1 Introduction

As discussed in Chapter 1, one of the goals of this thesis was to develop an inexpensive generic cellulose-based affinity chromatography column for one-step purification of fusion proteins containing a cellulose-binding-domain tag from culture broths. In this chapter, the sieved-Avicel packed column developed and characterized in Chapter 3 is applied to the purification of a staphylococcal protein A - CBD_{Cex} (proA- CBD_{Cex}) fusion protein produced recombinantly in *E. coli*. The proA- CBD_{Cex} fusion was chosen because of its potential applications in affinity purification and immobilization of antibodies. Immobilization of antibodies through binding to a proA- CBD_{Cex} fusion can, for instance, be applied to the development of a range of antigen diagnostic systems utilizing cellulosic materials.

The column model derived and characterized in Chapter 3 is generalized and evaluated as a means of predicting breakthrough of proA- CBD_{Cex} when it is frontal loaded as a component of a complex culture supernatant. In contrast to the breakthrough curves we measured using pure CBD_{Cex} to validate our predictive model, the breakthrough of the target solute (*e.g.*, ProA-CBD_{Cex}) in a complex culture supernatant can not be observed directly because it will be masked by the co-elution of contaminating proteins. On-line determination of column saturation is therefore problematic. Under such conditions, column simulation can provide a valuable estimate of breakthrough and a means of controlling the various stages of operation.

4.2 Model extension

The column model derived in section 3.5 strictly applies to affinity adsorption of a single solute from a binary aqueous feed solution. The potential effect of contaminating solutes on the binding and breakthrough behavior of the target protein is not considered. Although extension of the model to specifically include a mixture of non-binding contaminants is theoretical possible, our objective is to establish a useful, simple algorithm for estimating product breakthrough. To achieve this goal, we have assumed that the contaminating solutes are sufficiently dilute that their presence does not influence the transport and binding properties of the target protein within the column. The feed solution can then be regarded as a pseudo-binary mixture of the fusion protein in the solvent, and the model equations presented in section 3.5 can be applied.

As noted in Chapter 3, the model requires a number of parameters (see Table 3.1), several of which depend on the properties, particularly the size, of the target protein (the CBD_{Cex}-ProA fusion has a molecular weight of 42 KDa). These include the porosity ß, the binding constant K_a, the maximum capacity Q_{max}, the axial dispersion coefficient D_L, the film mass transfer coefficient k_f, and the intrapore diffusivity D_p. Often one does not have sufficient pure protein available to independently measure these parameters according to the methods described in Chapter 2 and 3. In such cases, reasonable methods are required for estimating or eliminating these model parameters.

To simplify matters, we will make the following assumptions: (1) since affinity binding is specific for CBD_{Cex}, K_a for the fusion protein can be set to the value measured for isolated CBD_{Cex}, (2) axial dispersion makes only a minor contribution to breakthrough behavior, so that D_L for the fusion can be set equal to the measured D_L for CBD_{Cex} without significant error, and (3) sufficient protein is available to make a single point batch measurement of Q_{max} as described

in Chapter 3. These assumptions reduce the number of parameters which must be estimated to three: k_f, ß and D_p.

4.2.1 Film mass transfer coefficient

The dependence of k_f on column hydrodynamics (*i.e.*, the superficial velocity u) and the bulk-phase diffusivity D_m of the target protein is provided by the semi-empirical correlation of Foo and Rice (1975) discussed in Chapter 3

$$Sh = \frac{2R_p k_f}{D_m} = 2 + 1.45 R_e^{0.5} S_c^{0.33} = 2 + 1.45 \left(\frac{2R_p u}{\nu}\right)^{0.5} \left(\frac{\nu}{D_m}\right)^{0.33} \qquad (4.1)$$

The bulk-solution solute diffusivity D_m is the only parameter in Equation (4.1) that depends on the molecular weight of the target protein. The Polson (1950) equation, proposed in 1950 based on a limited set of polymer diffusivity data, appears to provide a reasonable estimate of the molecular weight dependence of D_m (Tyn and Gusek, 1990). The form of the Polson equation is given by Equation (3.50), where A is a constant which depends on temperature and probe geometry.

Table 4.1 is a compilation from the literature of diffusivities at 20°C of a number of globular proteins in aqueous solution. Protein molecular weights range from 11 to 240 kDa. Figure 4.1 plots experimental D_m data in the form of the Polson equation. The proportionality constant A is then given by the slope of the resulting line. From the figure, A is equal to 26.81×10^{-6} cm^2 s^{-1} g$^{1/3}$ mol$^{-1/3}$ and the resulting correlation for D_m in m^2 s^{-1} at 20°C is

$$D_m = \frac{26.81 x 10^{-10}}{MW^{1/3}} \qquad (4.2)$$

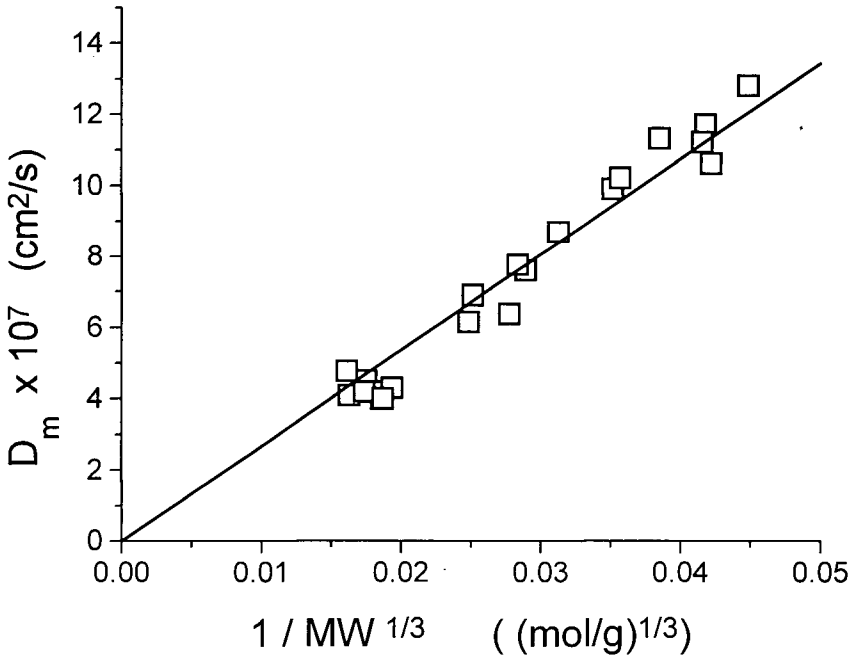


Figure 4.1. Experimental bulk-solution diffusivities (D_m) for protein molecular weights range from 11 to 240 kg mol^{-1}. Proportionality constant; $A = 26.81 \times 10^{-6}$ cm^2 s^{-1} $g^{1/3}$ $mol^{-1/3}$ (R=0.9845)

Together, Equation (4.1) and (4.2) provide a general method for estimating the diffusivities of globular proteins in aqueous solution at 20°C. It should not be applied to fibrous or random-coil protein structures, or at operating temperatures far removed from ambient.

TABLE 4.1: Dependence of globular protein diffusivities (D_m) and sizes on molecular weight

Protein	MW (g/mole)	$D_m \times 10^7$ (cm^2/s)	Size (Å)	Reference
CBD_{cex}	11,081	12.78	45x25x25	1, 2
Ferricytochrome c (horse heart)	11,980	12.85	30x34x34	3
α-Lactalbumin (bovine milk)	13,300	10.6	-	4
Ribonuclease (bovine pancreas)	13,700	11.7	38x28x22	4, 5
Lysozyme (chicken egg white)	13,930	11.2	30x30x45	4, 6
Prophospholipase A_2 (porcine pancreatic)	14,752	12.0	25x28x35	7
Flavodoxin (*Clostridium pasteurianum*)	15,000	11.93	36x42x28	8
Myoglobin (horse heart)	17,500	11.3	-	4
Chymotrypsin B (bovine pancreas)	23,000	9.9	-	4
Chymotrypsinogen α (bovine pancreas)	25,000	10.2	50x40x50	9, 4
Pepsin (swine stomach mucosa)	32,700	8.7	-	4
Carboxypeptidase B (bovine pancreatic)	34,606	9.05	50x40x38	10
Acid protease (*Rhizopus chinensis*)	35,000	9.02	37x46x61	11

Protein	MW (g/mole)	$D_m \times 10^7$ (cm^2/s)	Size (Å)	Reference
Acid proteinase (fungus *Endothia parasitica)*	36,000	8.93	60x40x35	12
ß-Lactoglobulin (bovine milk)	41,400	7.6	-	4
Ovalbumin	44,000	7.76	70x45x50	4, 13
Enolase (yeast)	46,650	8.2	60x50x45	14
Phosphoglycerate-kinase (yeast)	46,800	6.38	-	4
Phosphoglycerate kinase (horse muscle)	48,000	8.12	70x45x35	15
Hemoglobin	63,000	6.9	-	4
Bovine serum albumin	65,400	6.15	-	4
Alcohol dehydrogenase (horse liver)	80,000	6.61	45x60x110	16
Phosphoglucose isomerase (pig muscle)	120,000	6.0	70x70x100	17
Lysine-tRNA ligase	138,000	4.3	-	4
γG-Immunoglobulin (IgG)	156,000	4.0	-	4
Glycogen phosphorylase (rabbit muscle)	163,000	4.2	-	4
Phosphorylase b (rabbit muscle)	185,000	5.2	110x65x55	18
Malate synthese (yeast)	187,000	4.5	-	4
Pyruvate kinase (*S. carlsbergenis)*	190,800	4.2	-	4

Protein	MW (g/mole)	$D_m \times 10^7$ (cm^2/s)	Size (Å)	Reference
Catalase (bovine liver)	232,000	4.1	-	4
Pyruvate kinase (rabbit muscle)	240,000	4.78	75x95x125	17

(1) Ong *et al.* (1993)
(2) Xu *et al.* (1995)
(3) Dickerson *et al.* (1971)
(4) Tyn and Gusek (1990)
(5) Kartha *et al.* (1967)
(6) Artymiuk *et al.* (1979)
(7) Drenth *et al.* (1976)
(8) Ludwig *et al.* (1971)
(9) Kraut *et al.* (1962)
(10) Schimd and Herriott (1976)
(11) Subramanian *et al.* (1977)
(12) Jenkins *et al.* (1975)
(13) Stein *et al.* (1991)
(14) Stec and Lebioda (1990)
(15) Blake *et al.* (1972)
(16) Klund *et al.* (1976)
(17) Blake *et al.* (1975)
(18) Fischer *et al.* (1970)

4.2.2 Particle porosity and pore diffusivity

Figure 3.11 shows the dependence of the porosity ß of the Avicel particles on the hydrodynamic diameter of the probe molecule. Thus, if the hydrodynamic radius of the fusion protein of interest was known, Figure 3.11 and the associated Equation (3.35) could be used to determine ß. However, no such data are available for fusion proteins containing a CBD tag. In general, only the molecular weight of the fusion protein is known. Thus, we require a method to estimate ß for a given fusion-protein molecular weight.

To establish a useful correlation for the dependence of ß on fusion-protein molecular weight MW_p, we assume that the volume V_p and general shape of the fusion protein are well described as an ellipsoid of semi-axes r_a, r_b, and r_c (Spiegel, 1968) so that

$$V_p = \frac{4}{3}\pi\, r_a r_b r_c = \frac{1}{6}\pi\, d_h^{\,3} \qquad (4.3)$$

The second equality in Equation (4.3) defines d_h, the effective hydrodynamic diameter of the protein. Table 4.1 gives the ellipsoidal dimensions (diameters), determined from crystal structures, for a number of globular proteins of known molecular weight. Application of Equation (4.3) to determine d_h for each of these proteins allows us to (crudely) convert the x-axis of Figure 3.11 into a globular protein molecular weight scale, as shown in Figure 4.2.

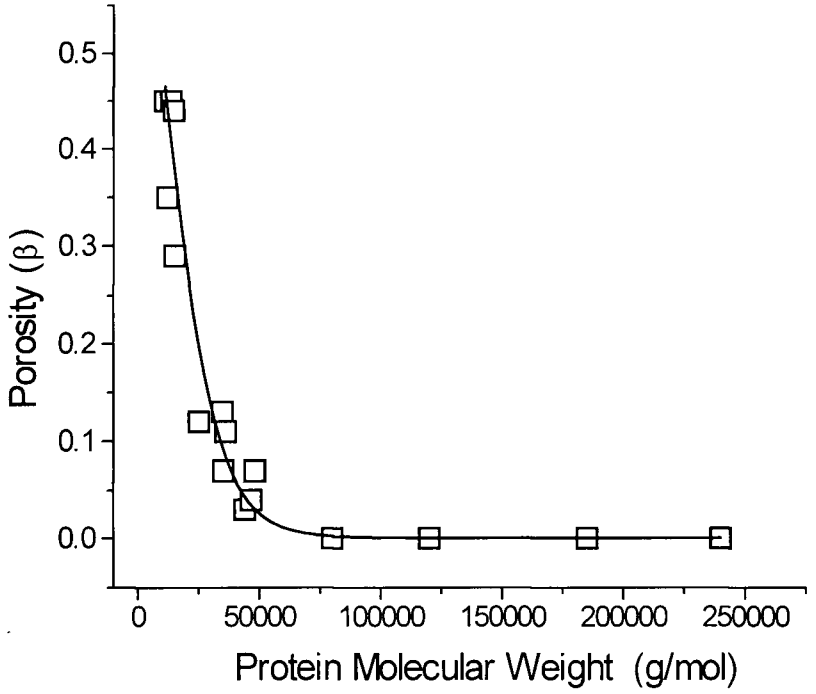


Figure 4.2. Correlation between porosity (β) available to different proteins as a function of their molecular weights; (□) calculated, (solid line) fitted

Figure 4.2 predicts that a fusion protein of molecular weight greater than *ca.* 40 kDa will not penetrate the internal volume of the Avicel particles. Since CBD_{Cex} has a molecular weight of 11 kDa, any fusion partner of molecular weight greater than *ca.* 29 kDa will be excluded from the particle interior. Thus, all adsorption will occur on the external surface of the particles and intrapore diffusion can be ignored in the column model. For fusion proteins of molecular weight less than 40 kDa, the value of ß can be determined from the regression function

$$\beta = \frac{\sigma}{1 + e^{(\xi MW - \psi)}} \qquad (4.4)$$

where $\sigma = 1.027$, $\xi = 8.9 \times 10^{-5}$, and $\psi = 0.7978$.

What remains then is a method for estimating D_p for those fusion proteins which can penetrate the pores of the sorbent particles. Epstein (1989) has shown that the intrapore and bulk solute diffusivities are related by

$$D_p = \frac{\beta D_m}{\kappa} \qquad (4.5)$$

where κ is the tortuosity factor. An estimated value for κ can be obtained from the D_p value for CBD_{Cex} which was independently measured by second-central-moments analysis in Chapter 3. From Table 3.2, D_p and D_m for CBD_{Cex} are 2.2×10^{-11} m^2 s^{-1} and 12.8×10^{-11} m^2 s^{-1}, respectively.

Application of Equation (4.5) with β=0.45 gives an estimated value for κ of 2.5. For porous solids, κ usually lies between 2 and 6 (Epstein, 1989; Miyabe and Suzuki, 1992). Combination of Equation (4.2) and Equation (4.5) therefore provides a method for estimating D_p as a function of fusion-protein molecular weight.

It must be noted that none of the above estimation methods (with the exception of the Foo and Rice correlation k_f) have been tested to any meaningful extent. Thus, more work is required before these correlations can be used with confidence. Nevertheless, as shown below,

incorporation of these estimated parameters into our column model leads to an excellent prediction of the breakthrough behavior of proA-CBD_{Cex}, suggesting that the estimations are reasonably accurate.

Table 4.2 shows all estimated and measured model parameters for proA-CBD_{Cex} which differ from those values shown in Table 3.2 for isolated CBD_{Cex}. The proA-CBD_{Cex} construct has a molecular weight of 42 kDa. Application of Figure 4.2 yields a sorbent porosity of less than 0.05, indicating that the protein is effectively excluded from the particle interior. Intraparticle diffusion effects are therefore negligible in this system, and film mass transfer becomes the rate-limiting step to adsorption. The estimated value for k_f is consistent with experimental results reported in the literature for proteins of similar size. For instance, Carlsson *et al.* (1994) report a k_f value of 6.9×10^{-6} m s^{-1} for lysozyme uptake at similar superficial velocities.

TABLE 4.2: Measured or estimated model parameters for Protein A-CBD_{cex}

Film mass transfer coefficient, k_f (estimated using Eq. 4.1)	5.1×10^{-6} m/s
Diffusion coefficient in pore liquid, D_p (estimated using Eq. 4.5)	1.6×10^{-12} m^2/s
Maximum capacity, Q_{max} (measured experimentally)	0.66 $mole/m^3$
Bead porosity, ß (taken from Figure 4.2)	< 0.05

4.3 Model Simplification for proA-CBD_{Cex} System

As shown in section 4.2.2, diffusion of proA-CBD_{Cex} into the pores of the Avicel particles is negligible. Intraparticle diffusion effects can therefore be neglected and the continuity equation derived in section 3.5 simplified to

$$\frac{\partial c}{\partial t} = D_L \frac{\partial^2 c}{\partial z^2} - \frac{u}{\varepsilon}\frac{\partial c}{\partial z} - \frac{(1-\varepsilon)}{\varepsilon}\rho_p \frac{\partial q}{\partial t} \qquad (4.6)$$

where q is the sorbate concentration on the exterior surface of the particle. Intraparticle diffusion, including Equation (3.18), is ignored and all adsorption is assumed to occur on the particle exterior.

The rate of mass transfer in the fluid film is then given by

$$\frac{\partial c_f}{\partial t} = \frac{3}{R_p}\frac{1-\varepsilon}{\varepsilon} k_f (c - c_f) \qquad (4.7)$$

where c_f is the concentration of the solute in the fluid film. The proportionality constant, $3/R_p\,(1\text{-}\varepsilon)/\varepsilon$, is the interfacial area per interstitial void volume of the packed bed.

In the absence of intrapore diffusion, the mass transfer and surface binding events occur in series and the mass balance between the solute and sorbate can be written as

$$\frac{\partial c_f}{\partial t} = \frac{1-\varepsilon}{\varepsilon}\frac{\partial q}{\partial t} \qquad (4.8)$$

The dependence of q on c_f is then given by Equation (3.21) or Equation (3.26) with $\theta^* = q/Q_{max}$ and $c_i = c_f$.

Equations (4.6) to (4.8) can be solved using the initial and boundary conditions given in Equation (3.27) to yield an appropriate column model when intrapore diffusion cannot occur due to the size of the fusion protein.

4.4 Results and Discussion

4.4.1 Production of proA-CBD_{Cex}

Protein production and final cell densities in batch fermentations of recombinant *E. coli* are often inhibited by the high levels of organic acids (particularly acetate) produced when the culture is subjected to a high initial concentration of glucose (Hassenwinkle *et al.*, 1997; Li and Taylor, 1994; Robbins and Taylor, 1989; Yamane and Shimizu, 1984; Yee and Blanch, 1992). Fed-batch fermentations, where essential nutrients such as glucose are periodically fed at noninhibiting levels, generally result in significantly higher cell densities and recombinant protein yields (Hassenwinkle *et al.*, 1997).

Figure 4.3 shows typical growth, initially with 0.2 % of glucose, and proA-CBD_{Cex}-production curves for a fed-batch fermentation of recombinant *E. coli* JM101 harbouring the pSACBD plasmid as described in section 2.2. The arrows in Figure 4.3 indicate where glucose additions to a final concentration of 2 g L^{-1} were made. Final cell densities between 6 and 7 optical density (OD_{600nm}) units were generally achieved, resulting in proA-CBD_{Cex} supernatant concentrations between 200 and 250 mg L^{-1}. Glucose concentrations remained low throughout the culture, and a direct correlation between cell density and protein production was observed.

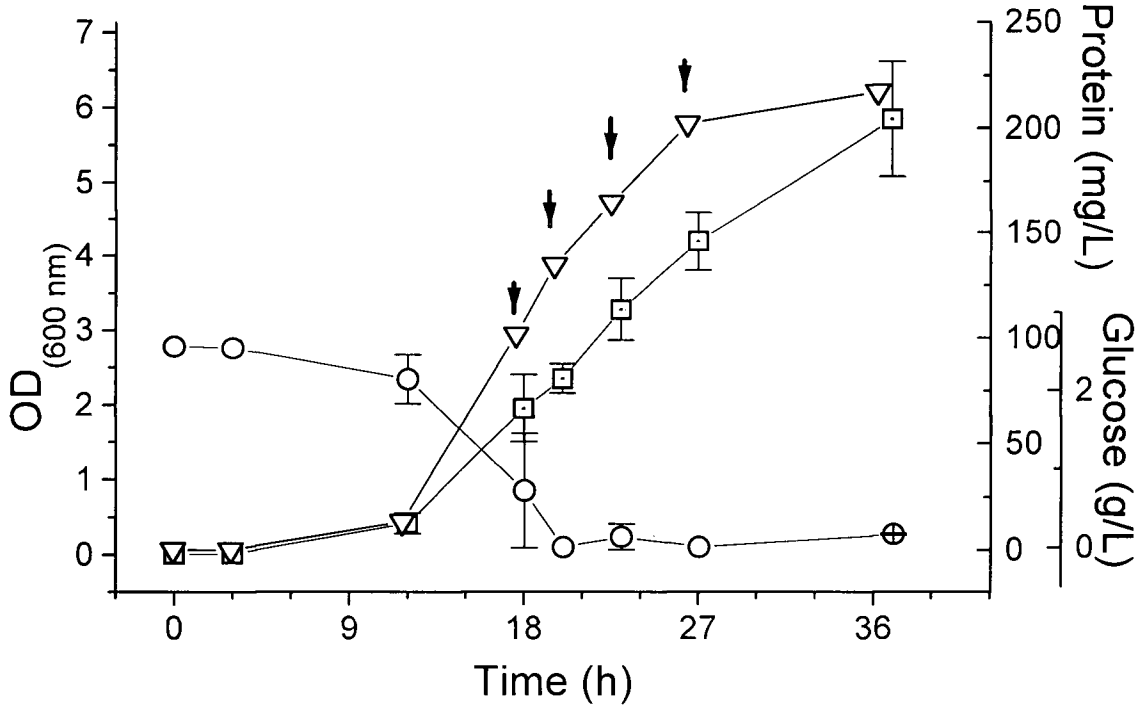


Figure 4.3. Cell growth, protein and glucose concentrations during growth of *E. coli* on M9 media with additions of 2 g/L glucose: (O) glucose concentration (g/L); (□) optical density, (A_{600}); (▽) Protein A-CBD_{cex} concentration (mg/L); (↓) 2 g/L glucose additions.

The final culture supernatant containing proA- CBD_{Cex} was separated form the cells by centrifugation at 10,000 rpm for 10 min. Lane 2 of the 15% SDS-PAGE gel in Figure 4.4 shows that the resulting supernatant is rich in proA-CBD_{Cex} and contains a relatively small number of contaminating proteins. The proA-CBD_{Cex} fusion accounts for *ca.* 45% of the total protein content in the supernatant. Lane 3 in Figure 4.4 identifies those proteins in the culture supernatant which bind to Avicel in a batch suspension. A 42-kDa protein, which corresponds to

the molecular weight of proA-CBD_{Cex}, is the only protein in the culture supernatant which binds to particulate Avicel, showing that the chromatographic media is highly specific for proA-CBD_{Cex} and shows little or no nonspecific "background" binding.

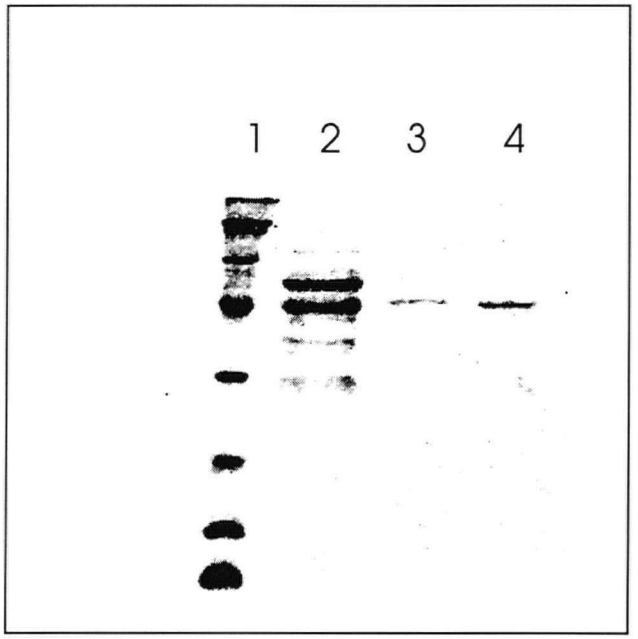


Figure 4.4. Protein A-CBD_{cex} in culture supernatant and bound to cellulose. Lane 1: molecular mass standards (kg/mole); Myosin 200; βGalactosidase 116.25; Phosphorylase B 97.4; BSA 66.2; Ovalbumin 45; Carbonic Anhydrase 31; Trypsin inhibitor 21.5; Lysozyme 14.4; Aprotinin 6.5. Lane 2: culture supernatant. Lane 3: culture supernatant mixed with Avicel for 3 h, the Avicel was washed with 1M NaCl in phosphate buffer, centrifuged, washed with phosphate buffer, centrifuged, then boiled in SDS. Lane 4: sample obtained from the pooled fractions after the chromatographic column was eluted with water

4.4.2 One-step affinity purification of proA-CBD_{Cex} from culture supernatant

Figure 4.5 shows a typical chromatogram for the (1) loading, (2) high-salt wash (1 M NaCl), (3) low-salt wash (50-mM phosphate buffer, pH 7), (4) elution (pure water), and (5) regeneration (6-M guanidinium hydrochloride (Gu-HCl)) stages of the affinity purification of proA-CBD_{Cex} from culture supernatant. The feed solution (pH 7, 20°C) contained 217 mg L^{-1} proA-CBD_{Cex} based on the calibration procedure described in section 2.2. 400 mL of culture supernatant were loaded onto the column. Initial breakthrough of nonbinding contaminating proteins occurs in approximately one void volume.

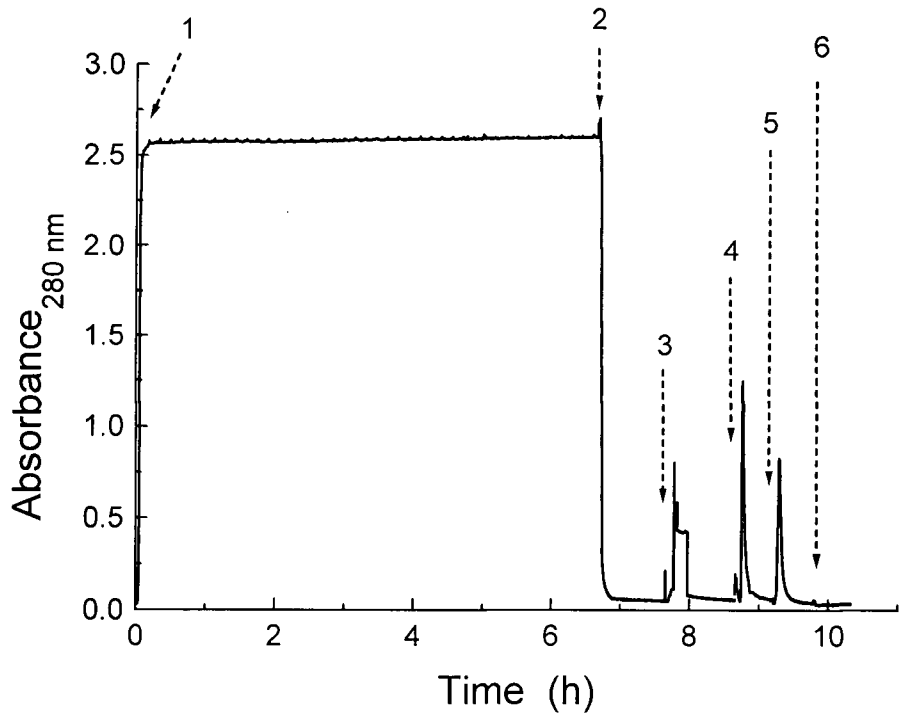


Figure 4.5. Chromatogram for the stages of the affinity purification of proA-CBD_{Cex}. (1-2) Loading, (2-3) high salt wash, (3-4) low-salt wash, (4-5) elution with water, (5-6) regeneration, (6) storage buffer.

For model analysis, breakthrough of proA-CBD_{Cex} was monitored by the antibody-marker dot-blot assay described in section 2.6. Since this assay requires binding of a rabbit anti-ß-glucosidase antibody to protein A, it provides further evidence that the 42-kDa protein identified in Figure 4.4 is intact proA-CBD_{Cex}. Figure 4.6 shows some of the dot-blot assay results obtained for breakthrough of proA-CBD_{Cex} through the Avicel affinity column. Blot 1A corresponds to proA-CBD_{Cex} in the unprocessed culture supernatant. Breakthrough of proA-CBD_{Cex} in the 3.7 mL sieved Avicel column occurs at around 210 min when the superficial velocity is 2.12×10^{-2} cm s^{-1}. Saturation is reached at *ca.* 270 min. Very similar proA-CBD_{Cex} breakthrough results were found when eluent fractions were run on a 15% SDS-PAGE (see Figure 4.4) and the resulting proA-CBD_{Cex} band (42 kDa) was read by densitometry. The complete breakthrough curve data for proA-CBD_{Cex} are shown in Figure 4.7.

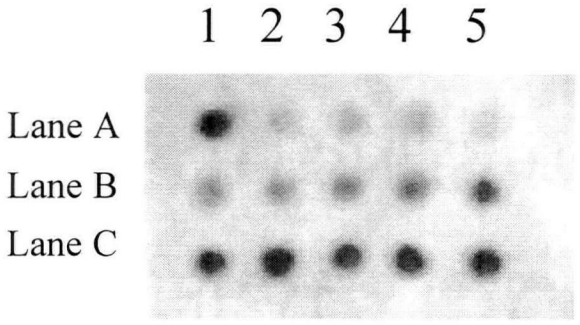


Figure 4.6. Dot blot assay results obtained for breakthrough of proA-CBD_{Cex}
Dot blots using 2µl of fractions collected from the breakthrough curve. Lane A; 1A: supernatant loaded into the column; 2A: fraction collected at column outlet after 5 min of initial loading; 3A: after 15 min; 4A: 85 min; 5A: 100 min. Lane B; 1B: 145 min; 2B: 155 min; 3B: 165 min; 4B: 210 min; 5B: 235 min. Lane C; 1C: 250 min; 2C: 270 min; 3C: 290 min; 4C: 310 min; 5C: 325 min.

Bound proA-CBD_{Cex} was eluted from the washed column with pure water as shown by the first elution peak in Figure 4.5. A recent isothermal titration calorimetry study showed CBD_{Cex} binding to crystalline cellulose is primarily driven by the entropy increase accompanying dehydration of the contacting protein and sorbent surfaces (Creagh *et al.*, 1996). Dehydration effects are minimized in pure water due to the excess concentration of free water (*i.e.*, water not involved in ion solvation). As a result, many (but not all) CBD_{Cex} fusion proteins which have been produced can be eluted with pure water (Kilburn *et al.* (1993)). In general, pure water is not an ideal eluent solution for proteins primarily because of the loss of buffering power. In this case, it also suffers from its lack of generality and the tendency for it to not elute all bound protein. Both site-directed mutagenesis (B. McLean, Ph.D. student) and random mutagenesis (O. Shoseyov, visiting scientist) studies of CBD_{Cex} are therefore in progress in the laboratory to engineer quantitative elution by, for instance, a change in pH.

As shown in Lane 4 of Figure 4.4, the proA-CBD_{Cex} recovered by the water elution step is relatively pure. No major contaminant protein bands are observed. The recovered proA-CBD_{Cex} is also fully bi-functional: able to bind cellulose and rabbit anti-ß-glucosidase antibody, as shown in Figure 4.6. Regrettably, not all of the bound proA-CBD_{Cex} is eluted with pure water. A fraction of presumably more tightly bound or entrapped protein must be eluted with 6-M GuHCl (see Figure 4.5). The yield of active proA-CBD_{Cex} recovered in the pure water elution is 57%. The remaining 43% of bound proA-CBD_{Cex} is eluted in the 6-M GuHCl regeneration step. Through the mutagenesis research on-going in the laboratory, we hope to improve yields by developing a more efficient elution strategy.

4.4.3 Model predictions

The estimated parameters shown in Table 4.2 for proA-CBD_{Cex} were combined with the column model shown in section 4.3 to predict breakthrough behavior of proA-CBD_{Cex} when loaded as a component of a complex culture supernatant. In applying this simplified model, we have assumed that proA-CBD_{Cex} does not diffuse into the internal pores of the Avicel particles. Thus, all binding is to the exterior of the particle. All remaining model parameters were taken from Table 3.2.

Figure 4.7 compares the predicted proA-CBD_{Cex} breakthrough curve with that measured at the same conditions using the rabbit anti-ß-glucosidase antibody dot-blot assay. Local binding equilibrium is assumed in the calculation such that the sorbate concentration q is defined by c_f according to Equation (3.26). Model predictions are in excellent agreement with experiment. Although the model must be tested more extensively to fully verify its utility, these results suggest that at least for low-density cultures the pseudo-binary solution assumption (whereby we ignore the effects of contaminating proteins on the binding and transport properties of proA-CBD_{Cex}) is reasonable. It also suggests that the parameter estimation methods described in section 4.2 are reasonably accurate.

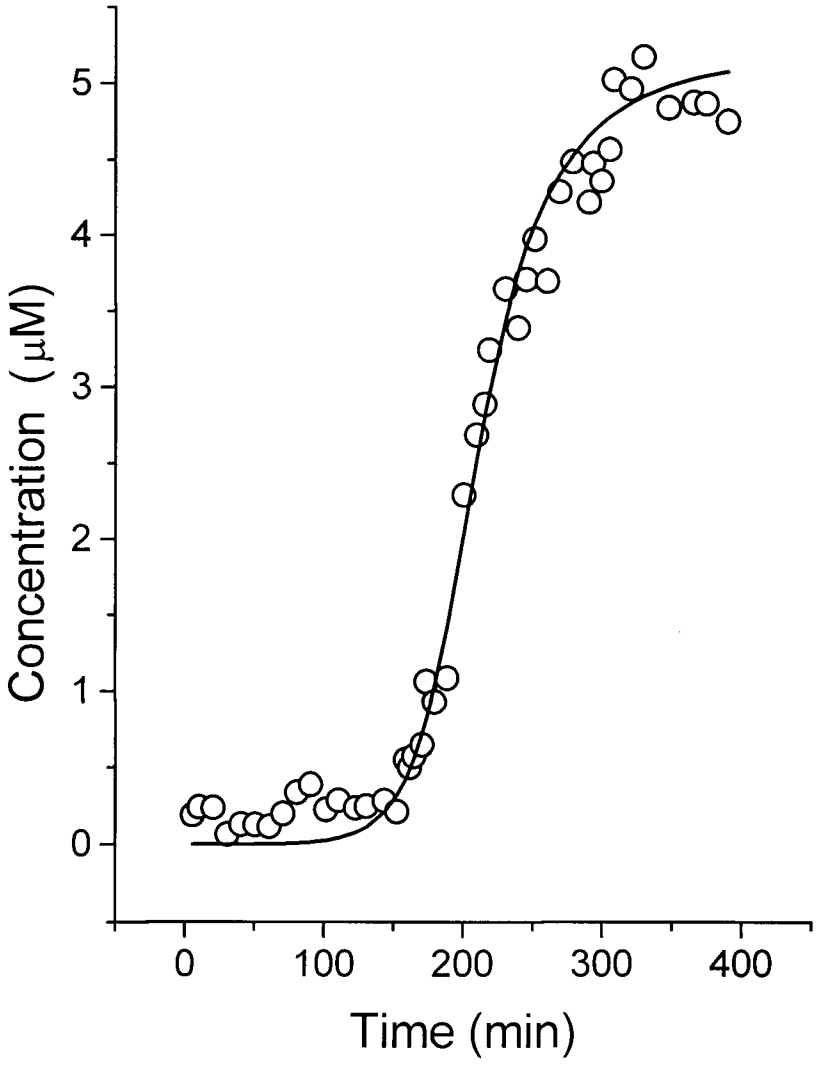


Figure 4.7. Comparison between experimental breakthrough curve and model prediction for Protein A-CBD_{cex}. (O) Experimental values; (line) Model prediction.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

With the emergence of off-patent drug manufacturers and shrinking Federal health budgets, production efficiency and economics are becoming critical issues in Biotechnology. As a result, companies are now seeking new methods for streamlining downstream processing of products made by fermentation or cell culture. Chromatography remains the primary method for purifying proteins, nucleic acids, and secondary metabolites. To reduce the number of required chromatographic steps, industry is increasingly turning to affinity chromatography. Limitations of current commercial affinity chromatography matrices include complex immobilization chemistries, resin heterogeneities, expense, and poor matrix stability against nonspecific adsorption and during sterilization.

This thesis addresses one possible solution to these limitations: the application of genetic fusion to a cellulose binding domain as a generic means of purifying recombinant proteins on an inexpensive and stable semi-crystalline-cellulose-based affinity chromatography column. The sieved Avicel affinity column developed in this thesis was successfully applied to the one-step purification of a staphylococcal proA-CBD_{Cex} fusion protein from the supernatant of a recombinant *E. coli* culture. Protein yield from the purification was 57 %, and could be improved by the development of a more effective elution protocol. Chromatograms from the Avicel affinity column are characterized by steep breakthrough curves, indicating fast binding kinetics and uniform flow distribution, and tight elution peaks, allowing for the recovery of a concentrated protein product. Modest scale-up of the column did not alter column performance.

A detailed differential column model, which takes into account all mass-transfer and binding kinetics/thermodynamics, was developed and solved to evaluate column performance, the rate-limiting step(s) to adsorption, and the dependence of breakthrough on column size and operating conditions. Complex moments analyses combined with breakthrough-curves data and solute-exclusion data were used to determine all essential model parameters. Model analysis indicated that the rates of solute diffusion through the stagnant fluid film surrounding the sorbent particles and within the pores of the sorbent were similar and rate limiting. Both of these limitations could be at least partially reduced by new matrix geometries which minimize diffusional path lengths within the sorbent phase. Although no data are reported here, we are currently testing a new column geometry based on a woven cellulose fabric which greatly reduces the momentum boundary layer by allowing for convective transport through the woven fibers.

The study of breakthrough curves was used to obtain detailed information about the behavior of the chromatographic system. The dependence of curve shape on column length, superficial velocity, and feed concentration, provided estimates for the acceptable range of column operating conditions. Comparison of calculated and experimental breakthrough curves was then used to validate the proposed column model and the associated protein adsorption-desorption behavior within the column. Accurate model predictions can be used to design and optimize columns during process scale-up so that the number of bench-scale and pilot-plant experiments, which are tedious and expensive, can be reduced. However, further model testing is required before such predictive calculations can be performed with confidence.

The model is able to predict protein breakthrough as a function of operating conditions, including protein feed concentration and superficial velocity. Under an appropriate set of

simplifying assumptions, the model is also able to predict target protein breakthrough and column performance for the separation of recombinant proA-CBD_{Cex} from a complex culture supernatant. A set of useful correlations were established to obtained critical model parameters for fusion proteins with a CBD_{Cex} affinity tag. Although further testing is required, the model could be used to optimize operating conditions and feed cycles for the affinity purification of any protein genetically fused to the CBD_{Cex} affinity tag (*e.g.* ß-glucosidase- CBD_{Cex} ~ 68 kg/mole (Ong *et al.*, 1989)).

NOMENCLATURE

a_i	activity (Equation 3.7)	
A	constant of value 2.85×10^{-5} cm^2 s^{-1} $g^{1/3}$ $mol^{1/3}$ (Eq. 3.50)	
B_i	Biot number ($B_i = \frac{k_f R_p}{3D_p}$)	
c	solute concentration in bulk liquid	(mol m^{-3})
c_i	solute concentration in pore liquid	(mol m^{-3})
c_f	concentration of the solute in the fluid film	(mol m^{-3})
C_f	final probe concentration (Eq. 3.30)	(g / L)
C_o	initial concentration of solute in bulk liquid	(μM)
dG	Gibbs energy change	
d_h	effective hydrodynamic diameter of the protein (Eq. 4.3)	(Å)
d_p	particle diameter	(μm)
D_L	axial dispersion coefficient	(m^2 s^{-1})
D_m	diffusion coefficient in bulk liquid	(m^2 s^{-1})
D_p	intrapore diffusivity	(m^2 s^{-1})
I	inaccessible pore volume (Eq. 3.31)	(ml / g)
I_{max}	maximum inaccessible pore volume (Eq. 3.32)	(ml / g)
k_f	fluid film mass transfer coefficient	(m s^{-1})
K_a	binding equilibrium association constant	(M^{-1})
K_L	Langmuir equilibrium binding constant	(M^{-1})
K_m	overall mass transfer coefficient	(s^{-1})

L	column length	(m)
M	molecular weight of the protein (Eq. 3.50)	(g mol^{-1}).
M_p	amount of poly(ethylene glycol) standard, (Eq. 3.30)	(g)
MW	protein molecular weight	(kg $mole^{-1}$)
MW_p	fusion protein molecular weight	(kg $mole^{-1}$)
n_p	moles of protein in the bulk solution (Eq. 3.3 to 3.5)	(moles)
n_w	moles of water in the bulk solution	(moles)
N	moles of independent identical binding sites	(moles)
N_o	flux of solute into the particles	
N_{th}	Number of theoretical plates in the column	
q	sorbate concentration on the exterior surface of the particle	(mol m^{-3})
q_i	equilibrium concentration of bound solute *i*	(mol m^{-3})
q	amount of water added (Eq. 3.30)	(g)
p	amount of Avicel added (Eq. 3.30)	(g)
Q	volumetric flow rate	(m^3 s^{-1})
Q_{max}	maximum capacity of the sorbent to bind	(mol m^{-3})
r	radial coordinate	
r_a, r_b, and r_c	ellipsoid semi-axes (Eq. 4.3)	(Å)
R_{eq}	inverse of the separation factor (α), (Equation 1.3)	
Re	Reynolds number ($Re = \frac{2R_p u}{v}$)	
R_p	particle radius	(m^{-2})
r_w	sites filled with adsorbed water (Eq. 3.3 to 3.5)	(moles)

r_p	sites filled with a specifically adsorbed protein	(moles)
S	surface binding sites	
S	pore volume inaccessible to probe molecules (Eq. 3.30, 3.31)	(L)
Sc	Schmidt number ($Sc = \frac{\nu}{D_m}$)	
Sh	Sherwood number ($Sh = \frac{k_f 2R_p}{D_m}$)	
s	average concentration of solute in or on the sorbent particle	(mol m^{-3})
t	time	(s)
t_b	time of solute breakthrough	(s)
t_p, t_R	retention time of eluent peak; (Fig. 1.2), (Eq. 3.1)	(s)
t_w	peak width at baseline	(s)
$t_{w1/2}$	half of peak width at baseline	(s)
u	superficial liquid velocity	(m s^{-1})
V	bed volume	(L)
V_a	pore volume accessible to a given probe (Eq. 3.32)	(ml / g)
V_e	average elution volume	(L)
V_p	volume of the fusion protein	
V_o	column void volume	(L)
V_s	volume of sorbent particles in the column	(L)
W	the initial volume of solvent added (Eq. 3.30)	(L)
w_b	width of the breakthrough curve	(s)
w_p	width of the elution peak	(s)
X	molecular probe diameter (Eq.3.3 and 3.52)	(Å)

x_p	the mole fraction protein	
x_w	the mole fraction water	
z	axial coordinate	

Greek letters

α	separation factor (Equation 1.3), average number of overlapping sites covered by a CBD_{Cex} molecule (Equation 3.26), fitted parameter (Eq. 3.35 and 3.52)	
ß	bead porosity	
δ	fitted parameter (Eq. 3.35 and 3.52)	
θ^*	fraction of filled surface lattice sites	
θ_p	the mole fraction of binding sites filled by protein	
θ_w	the mole fraction of total binding sites occupied by solvent	
γ	fitted parameter (Eq. 3.35 and 3.52)	
ε	column void fraction	
σ	fitted parameter (Eq. 4.4)	
ξ	fitted parameter (Eq. 4.4)	
ψ	fitted parameter (Eq. 4.4)	
κ	tortuosity factor	
μ	chemical potential (Equation 3.3)	
μ_f	forward adsorption rate constant	(m^3 $mole^{-1}$ s^{-1})
μ_r	reverse adsorption rate constant	(s^{-1})
μ_1	first central moment of an eluent peak	(s^{-1})

μ_2	second central moment of an eluent peak	(s^{-1})
ρ_p	particle density	($kg\ m^{-3}$)
ρ_b	bulk density (Eq. 3.33)	($kg\ m^{-3}$)
ρ_w	solvent density (Eq. 3.34)	($kg\ m^{-3}$)
Γ	the surface concentration of protein	($\mu mol\ g^{-1}$)
Γ^{max}	maximum sorbent capacity	($\mu mol\ g^{-1}$)
ν	liquid kinematic viscosity	($m^2\ s^{-1}$)
ν_i	reaction coefficient (Eq. 3.9)	

BIBLIOGRAPHY

Arnold, F. H., Blanch, H. W., and Wilke, C. R. "Analysis of affinity separations. I : Predicting the performance of affinity adsorbers". Chem. Eng. J., **30**, B9-B23 (1985a)

Arnold, F. H., Blanch, H. W., and Wilke, C. R. "Analysis of affinity separations. II: The characterization of affinity columns by pulse technique". Chem. Eng. J., **30**, B25-B36 (1985b)

Arnold, F. H., Schofield, S. A., and Blanch, H. W. "Analytical affinity chromatography. I. Local equilibrium theory and the measurement of association and inhibition constants". J. Chromatogr., **355**, 1-12 (1986)

Artymiuk, P. J., Blake, C. C., Grace, D. E., Oatley, S. J., Phillips, D. C., and Sternberg, M. J. "Crystallographic studies of the dynamic properties of lysozyme". Nature, **280**, 563-568 (1979)

Blake, C. C. "X-Ray studies of glycolytic enzymes". Essays in Biochem., **11**, 37-79 (1975)

Blake, C. C., Evans, P. R., and Scopes, R. K. "Structure of horse-muscle phosphoglycerate kinase at 6 Å resolution". Nature New Biol., **235**, 195-198 (1972)

Blondel, A., and Bedouelle, H. "Export and purification of a cytoplasmic dimeric protein by fusion to the maltose-binding protein of *Escherichia coli*. Eur. J. Biochem., **193**, 325-330 (1990)

Boraston, A. Personal communication, 1997

Buettner, J. A., Dadd, C. A., Baumbach, G. A., Masecar, B. L., and Hammond, D. J. "Chemically derived peptide libraries: A new resin and methodology for lead identification". Int. J. Peptide Protein Res., **47**, 70-83 (1996)

Burnouf-Radosevich, M., and Burnouf, T. French Patent application No 8807530 (1988)

Carlsson, F., Axelsson, A., and Zacchi, G. "Mathematical modelling and parametric studies of affinity chromatography". Computers Chem. Engng., **18**, S657-S661 (1994)

Chase, H. A. "Affinity separations utilising immobilised monoclonal antibodies- A new tool for the biochemical engineer". Chem. Eng. Science., **39**, 1099-1125 (1984)

Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G., and Haynes, C. A. "Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven". Proc. Natl. Acad. Sci. USA, **93**, 12229-12234 (1996)

Cronan, J. "Biotination of proteins *in vivo*; a posttranslational modification to label, purify, and study proteins". J. Biol. Chem., **265**, 10327-10331 (1990)

Dankwerts, P. V. "Continuous flow systems". Chem. Eng. Sci., **2**, 1-13 (1953)

Desai, M. "Immunoaffinity adsorption: Process-scale isolation of therapeutic-grade biochemicals". J. Chem. Techn. Biotechnol., **48**, 105-126 (1990)

Dickerson, R. E., Takano, T., Eisenberg, D., Kallai, O. B., Samson, L., Cooper, A., and Margoliash, E. Ferricytochrome c. I. "General features of the horse and bonito proteins at 2.8 Å resolution". J. Biol. Chem., **246**, 1511-1535 (1971)

Din, N., Forsythe, I. J., Burtnick, L. D., Gilkes, N. R., Miller, R. C. Jr., Warren, R. A. J., and Kilburn, D. G. "The cellulose-binding domain of endoglucanase A (Cen A) from *Cellulomonas fimi*: evidence for the involvement of tryptophan residues in binding". Mol. Microbiol., **11**, 747-755 (1994)

Drenth, J., Enzing, C. M., Kalk, K. H., and Vessies, J. C. "Structure of porcine pancreatic prephospholipase A_2". Nature, **264**, 373-377 (1976)

Dunn, B. M., and Chaiken, I. M. "Quantitative affinity chromatography. Dertermination of binding constants by elution with competitive inhibitors". Proc. Nat. Acad. Sci. USA., **71**, 2382-2385 (1974)

Dykes, C. W., Bookless, A. B., Coomber, B. A., Noble, S. A., Humber, D. C., and Hobden, A. N. "Expression of atrial natriuretic factor as a cleavable fusion protein with chloramphenicol acetyltransferase in *Escherichia coli.*" Eur. J. Biochem., **174**, 411-416 (1988)

Epstein, N. "On tortuosity and the tortuosity factor in flow and diffusion through porous media". Chem. Engng. Sci., **44**, 777-779 (1989)

Field, J., Nikawa, J., Broek, D., MacDonald, B., Rodgers, L., Wilson, I., Lerner, R., and Wigler, M. "Purification of a RAS-responsive adenylyl cyclase complex from *Saccharomyces cerevisiae* by use of an epitope addition method". Mol. Cell. Biol., **8**, 2159-2162 (1988)

Finlayson, B. A. Nonlinear Analysis in Chemical Engineering, McGraw-Hill INC., New York (1980)

Fischer, E. H., Pocker, A., and Saari, J. C. "The structure, function and control of glycogen phosphorylase". Essays in Biochem., **6**, 23-68 (1970)

Foo, S. C., and Rice, R. G. "On the prediction of ultimate separation on parametric pumps". AIChE J., **21**, 1149-1158 (1975)

Ford, C. F., Suominen, I., and Glatz, C. E. "Fusion tails for the recovery and purification of recombinant proteins". Protein Expression Purif., **2**, 95-107 (1991)

Freitag, R. and Horvath, C. "Chromatography in the downstream processing of biotechnological products". Adv. Biochem. Engng., **53**, 17-59 (1995)

Gama, F. M., Teixeira, J. A., and Mota, M. "Cellulose morphology and enzymatic reactivity : A modified Solute Exclusion Technique". Biotechnol. Bioeng., **43,** 381-387 (1994)

Gemeiner, P., Stefuca, V., and Bales, V. "Biochemical engineering of biocatalysts immobilized on cellulosic materials". Enzyme Microb. Technol., **15**, 551-566 (1993)

Giddings, J. C. "Dynamics of Chromatography". New York, M Dekker (1965)

Gilkes, N. R., Jervis, E., Henrissat, B., Tekant, B., Miller, R. C. Jr., Warren, R. A. J., and Kilburn, D. G. "The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose". J. Biolog. Chem., **267**, 6743-6749 (1992)

Graham, R. W., Greenwood, J. M., Warren, R. A. J., Kilburn, D. G., and Trimbur, D. E. "The pTugA and pTugAS vectors for the high-level expression of cloned genes in *Escherichia coli*". *Gene*, **158**, 51 (1995)

Greenwood, J. M., Ong, E., Gilkes, N. R., Warren, R. A. J., Miller, R. C., Jr. and Kilburn, D. G. "Cellulose-binding domains: potential for purification of complex proteins". Protein Engineering, **5**, 361-365 (1992)

Grethlein, H. E. "The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulosic substrates". Bio/Technol., **2**, 155-160 (1985)

Hall, K. R., Eagleton, L. C., Acrivos, A., and Vermeulen, T. Ind. Eng. Chem. Fund., **212**, 213 (1976)

Hammarberg, B., Nygren, P.-A., Holmgren, E., Elmblad, A., Tally, M., Hellman, U., Moks, T., and Uhlen, M. "Dual affinity fusion approach and its use to express recombinant human insulin-like growth factor II". Proc. Natl. Acad. Sci. USA, **86**, 4367-4371 (1989)

Hartman, J., Daram, P., Frizzell, R. A., Rado, T., Benos, D. J., and Sorscher, E. J. "Affinity purification of insoluble recombinant fusion proteins containing Glutathione-S-Transferase". Biotechnol. Bioeng., **39**, 828-832 (1992)

Hassenwinkle, D., Jervis, E., Kops, O., lin, C., Lesnicki, G., Haynes, C. A., and Kilburn, D. G. "Very high-level production and export in *Escherichia coli* of a cellulose binding domain for use in a generic secretion-affinity fusion system". Biotechnol. Bioeng., **6**, 854-863 (1997)

Haynes, C. A., Benitez, F. J., Blanch, H. W., and Prausnitz, J. M. "Application of Integral-Equation Theory to aqueous two-phase partitioning systems" AIChE. J., **39**, 1539-1557 (1993)

Haynes, H. W., and Sarma P. N. AIChEJ., **19**, 1043 (1973)

Hermanson, G. T., Mallia, A. K. and Smith, P. K. "Immobilized affinity ligand techniques", Academic Press,INC. San Diego, California (1992)

Highsmith, F., Regan, T., Clark, D., Drohan, W., and Tharakan, J. "Evaluation of CNBr, FMP and hydrazide resins for immunoaffinity purification of Factor IX". BioChromatography, **12**, 418-423 (1992)

Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerretti, D. P., Urdal, D. L., and Conlon, P. J. "A short polypeptide marker sequence useful for recombinant protein identification and purification" Bio/Technol., **6**, 1204-1210 (1988)

Huang, P. Y., and Carbonell, R. G. "Affinity purification of proteins using ligand derived from peptides libraries". Biotechnol. Bioeng., **47**, 288-297 (1995)

Hughes, J., Lakey, J. H., and Hughes, M. A. "Production and characterization of a plant α-Hydroxynitrile lyase in *Escherichia coli*". Biotechnol. Bioeng., **35**, 332-338 (1997)

Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A., Nordheim, A., and Stunnenberg, H. G. "Rapid and efficient purification of native histidine-tagged protein expressed by recombinant vaccinia virus". Proc. Natl. Acad. Sci. USA, **88**, 8972-8976 (1991)

Jenkins, J. A., Blundell, T. L., Tickle, I. J., and Ungaretti, L. "The low resolution structure analysis of an acid proteinase from *Endothia parasitica*". J. Mol. Biol., **99**, 583-590 (1975)

Jervis, E. J., Haynes, C. A., and Kilburn, D. G. "Surface diffusion of cellulases and their isolated binding domains on cellulose". J. Biological Chemistry (accepted) (1997)

Jin, X., Talbot, J., and Wang, N. -H. L "Analysis of steric hindrance effects on adsorption kinetics and equilibria". AIChE J., **40**, 1685-1696 (1994)

Johnston, A., and Hearn, M. T. W. "High-performance liquid chromatography of amino acids, peptides and proteins. CIII. Mass transfer resistances in ion-exchange and dye-affinity chromatography of proteins". J. Chromatogr., **512**, 101-114 (1990)

Jones, C., Patel, A., Griffin, S., Martin, J., Young, P., O'Donnell, K., Silverman, C., Porter, T., and Chaiken, I. "Current trends in molecular recognition and bioseparation". J. Chromatogr. *A*, **707**, 3-22 (1995)

Jones, K. " A review of biotechnology and large scale affinity chromatography" Chromatographia, **32**, 469-480 (1991)

Jungbauer, A. "Insights into the chromatography of proteins provided by mathematical modeling" Curr. Opin. Biotechnol., **7**, 210-218 (1996)

Kartha, G., Bello, J., and Harker, D. "Tertiary structure of ribonuclease". Nature, **213**, 862-865 (1967)

Kastner, M., and Neubert, D. "High-performance metal chelate affinity chromatography of cytochromes P-450 using Chelating Superose". J. Chromotogr., **587**, 43-54 (1991)

Kilburn, D. G., Turner, R. F. B., Coutinho, J. B., Din, N., Gilkes, N. R., Greenwood, J. M., Hobbs, J. B., Miller, R. C., Jr., Ong, E., Phelps, M. R., Ramirez, C., Warren, R. A. J. Cellulose Binding Domains: Applications in Biotechnology, pp. 45-54. In: Kennedy, J. F., Phillips, G. O., Wiltham, P. A. Ellis Horwood Ltd, Simon and Schuster International Group (eds.), Cellulosics: Materials for selective separations and other technologies. West Sussex England (1993)

Klund, H. E., Nordstrom, B., Zeppezauer, E., Soderlund, G., Ohlsson, I., Boiwe, T., Soderberg, B., Tapia, O., Branden, C., and Akeson, A. "Three-dimensional structure of horse liver alcohol dehydrogenase at 2.4 Å resolution". J. Mol. Biol., **102**, 27-59 (1976)

Knight, E., and Fahey, D., Jr., "Human fibroblast interferon" J. Biological Chemistry., **256**, 3609-3611 (1981)

Koska, J., and Haynes, C. A. Personal communication (1997)

Kraut, J., Sieker, L. C., High, D. F., and Freer, S. T. "Chymotrypsinogen: a three-dimensional fourier synthesis at 5 Å resolution". Proc. Nat. Acad. Sci., Wash., **48**, 1417-1421 (1962)

Kucera, E. "Contribution to the theory of chromatography. Linear non-equilibrium elution chromatography". J. Chromatgr., **19**, 237-248 (1965)

Laemmli, U. K. "Cleavage of structural proteins during the assembly of the head of bacteriophage T4". Nature, **227**, 680-685 (1970)

Le, K. D., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr., Saddler, J. N., and Warren, R. A. "A streptavidin-cellulose-binding domain fusion protein that binds biotinylated proteins to cellulose". Enzyme Microb. Technol., **16**, 496-500 (1994)

Li, X., and Taylor, K. B. "Effect of glucose on the expression parameters of recombinant protein in *Escherichia coli* during batch growth in complex medium". Biotechnol. Prog., **10**, 160-164 (1994)

Liapis, A. I. "Modelling affinity chromatography". Separation and Purification Methods, **19**, 133-210 (1990)

Liapis, A. I., Anspach, B., Findley, M. E., Davies, J., Hearn, M. T. W., and Unger, K. K. "Bioespecific adsorption of lysozyme onto monoclonal antibody ligand immobilized on nonporous silica particles". Biotechnol. Bioeng., **34**, 467-477 (1989)

Lin, J. K., Ladisch, M. R., Patterson, J. A., and Noller, C. H. "Determining pore size distribution in wet cellulose by measuring solute exclusion using a differential refractometer". Biotechnol. Bioeng., **29**, 976-981 (1987)

Ljungquist, C., Breitholtz, A., Bruik-Nilsson, H., Moks, T., Uhlen, M., and Nilsson, B. "Immobilization and affinity purification of recombinant proteins using histidine peptide fusions". Eur. J. Biochem., **186**, 563-569 (1989)

Lowe, C. R., Burton, S. J., Burton, N. P., Alderton, W. K., Pitts, J. M. and Thomas, J. A. "Designer dyes: "biomimetic" ligands for the purification of pharmaceutical proteins by affinity chromatography". TIBTECH, **10**, 442-448 (1992)

Ludwig, M. L., Andersen, R. D., Apgar, P. A., Burnett, R. M., LeQuesne, M. E., and Mayhew, S. G. "The structure of a clostridial flavodoxin, an electron-transferring flavoprotein. III. An

interpretation of an electron-density map at a nominal resolution of 3.25 Å". Cold Springs Harbor Symp. Quant. Biol., **36**, 369-380 (1971)

Ma, Z., Whitley, R. D., and Wang, N.-H. L. "Pore and surface diffusion in multicomponent adsorption and liquid chromatography systems" AIChE J., **42**, 1244-1262 (1996)

Makriyannis, T., and Clonis, D. "Design and study of peptide-ligand affinity chromatography adsorbents: application to the case of trypsin purification from bovine pancreas". Biotechnol. Bioeng., **53**, 49-57 (1997)

Mao, Q. M., Johnston, A., Prince, I. G., and Hearn, M. T. W. "High -performance liquid chromatography of amino acids, peptides and proteins. CXIII. Predicting the performance of non-porous particles in affinity chromatography of proteins". J. Chromatogr., **548**, 147-163 (1991)

Marchand, D. H. "Chromatography in analytical biotechnology". Curr. Opinion Biotechnol., **5**, 72-76 (1994)

McAleer, W. J., Buynak, E. B., Maigetter, R. Z., Wampler, D. E., Miller, W. J., and Hilleman, M. R. "Human hepatitis B vaccine from recombinant yeast". Nature, **307**, 178-180 (1984)

McCormick, D. "Chromatography and affinity separations". Bio/Technology, **6**, 158-165 (1988)

McCoy, M. A., and Liapis, A. I. "Evaluation of kinetic models for biospecific adsorption and its implications for finite bath and column performance". J. Chromatogr., **548**, 25-60 (1991)

Miyabe, K., and Suzuki, M. "Chromatography of liquid-phase adsorption on octadecylsilyl-silica gel". AIChE J., **38**, 901-910 (1992)

Moks, T., Abrahmsen, L., Holmgren, E., Bilich, M., Olsson, A., Uhlen, M., Pohl, G., Sterky, C., Hultberg, H., Josephson, S., Holmgren, A., Jornvall, H., and Nilsson, B. "Expression of human insulin-like growth factor I in bacteria: Use of optimized gene fusion vectors to facilitate protein purification." Biochemistry, **26**, 5239 (1987)

Narayanan, S. R. and Crane, L. J. "Affinity chromatography supports: a look at performance requirements". TIBTECH, **8**, 12-16 (1990)

Nelson R., and Oliver, D. W. "Study of cellulose structure and its relation to reactivity". J. Polymer Sci.: Part C, **36**, 305-320 (1971)

Nilsson, B., Abrahmsen, L., and Uhlen, M. "Immobilization and purification of enzymes with staphylococcal protein A gene fusion vectors". EMBO J., **4**, 1075-1080 (1985)

Nilsson, B., Forsberg, G., Moks, T., Hartmanis, M., and Uhlen, M. "Fusion proteins in biotechnology and structural biology". Curr. Opin. Struct. Biology, **2**, 569-575 (1992)

Nygren, P., Stahl, S., and Ulhen, M. "Engineering proteins to facilitate bioprocessing". TIBTECH, **12**, 184-188 (1994)

O'Neill, G. P., Goh, S. H., Warren, R. A. J., Kilburn, D. G., and Miller, R. C., Jr. "Structure of the gene encoding the exoglucanase of *Cellulomonas fimi*". Gene, **44**, 325-330 (1986)

Ong, E., Alimonti, J. B., Greenwood, J. M., Miller, R. C., Jr., Warren, R. A., and Kilburn, D. "Purification of human interleukin-2 using the cellulose-binding domain of a prokaryotic cellulase." Bioseparation, **5**, 95-104 (1995)

Ong, E., Gilkes, N. R., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. "The Cellulose-Binding Domain (CBDcex) of an exoglucanase from *Cellulomonas fimi*: Production in *Escherichia coli* and characterization of the polypeptide". Biotechnol. Bioeng., **42,** 401-409 (1993)

Ong, E., Greenwood, J. M., Gilkes, N. R., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. J. "The cellulose-binding domains of cellulases: tools for biotechnology". Trends Biotechnol.,**7**, 239-243 (1989)

Patankar, S. V. Numerical Heat Transfer and Fluid Flow, Hemisphere Publ. Co., New York (1980)

Perry's Chemical Engineering Handbook, 6th Ed., Don W. Green, Ed., McGraw-Hill, New York (1984)

Polson, A. "Some aspects of diffusion in solution and a definition of a colloidal particle". J. Phys. Colloid Chem., **54**, 649-652 (1950)

Ramirez, C., Fung, J., Miller, R. C., Jr., Warren, R. A. J., and Kilburn, D. G. "A bifunctional affinity linker to couple antibodies to cellulose". Bio/Technology, **11**, 1570-1573 (1993)

Robbins, J. W., Jr., and Taylor, K. B. "Optimization of *Escherichia coli* growth by controlled addition of glucose". Biotechnol. Bioeng., **34**, 1289-1294 (1989)

Ruthven, D. *Principles of adsorption and adsorption processes*. John Wiley and Sons, New York (1984)

Sambrook, J., Fritsch, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual, 2nd. ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. (1989)

Schimd, M. F., and Herriott, J. R. "Structure of carboxypeptidase B at 2.8 Å resolution". J. Mol. Biol., **103**, 175-190 (1976)

Sherwood, R. "Protein fusions: Bioseparation and application". Trends Biotechnol., **9**, 1-3 (1991)

Smith, D. B., and Johnson, K. S. "One- step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione S-transferase". Gene, **67**, 31-40 (1988)

Smith, J. M., and van Ness, H. C. "Introduction to Chemical Engineering Thermodynamics". Fourth Ed., McGraw Hill, New York (1986)

Sofer, G. "Preparative chromatographic separations in pharmaceutical, diagnostic, and biotechnology industries: current and future trends". J. Chromatogr. A, **707**, 23-28 (1995)

Spiegel, M. Mathematical Handbook of Formulas and Tables, Schaum's Outline Series, McGraw-Hill, New York (1968)

Staehelin, T., Hobbs, D. S., Kung H., Lai, C. Y., and Pestka, S. "Purification and characterization of recombinant human leukocyte interferon (IFLrA) with monoclonal antibodies". J. Biolog. Chemistry, **256**, 9750-9754 (1981)

Stec, B., and Lebioda, L. "Refined structure of yeast apo-enolase at 2.25 Å resolution". J. Mol. Biol., **211**, 235-248 (1990)

Steers, E., Jr., Cuatrecasas, P., and Pollard, H. B. "The purification of β-galactosidase from *Escherichia coli* by affinity chromatography". J. Biol. Chem., **246**, 196-200 (1971)

Stein, P. E., Leslie, A. G., Finch, J. T., and Carrell, R. W. "Crystal structure of uncleaved ovalbumin at 1.95 A resolution". J. Mol. Biol., **221**, 941-959 (1991)

Stone , J. F., and Scallan, A.M. "A structural model for the cell wall of water-swollen wood pulp fiber based on their accessibility to macromolecules". Cellulose Chem. and Tech., **2,** 343-358 (1968)

Subramanian, E., Swan, I. D., Lin, M., Davies, D. R., Jenkins, J. A., Tickle, I. J., and Blundell, T. L. "Homology among acid proteases: Comparison of crystal structures at 3 Å resolution of acid proteases from *Rhizopus chinensis* and *Endothia parasitica*". Proc. Nat. Acad. Sci., USA, **74**, 556-559 (1977)

Tomme, P., Warren, R. A. J., and Gilkes, N. R. "Cellulose hydrolysis by bacteria and fungi" in Advances in Microbial Physiology (Robert K. Poole, Ed), pp 2-66, Volume 37, Academic Press, London (1995a)

Tomme, P., Warren, R. A. J., Miller, R. C., Jr., Kilburn, D. G., and Gilkes, N. R. "Cellulose-binding domains - Classification and properties" in Enzymatic degradation of insoluble polysaccharides (Saddler, J. N. and Penner, M., Eds) pp 142-163, American Chemical Society (1995b)

Tyn, M. T., and Gusek, T. W. "Prediction of diffusion coefficients of proteins". Biotechnol. Bioeng., **35**, 327-338 (1990)

van Deemter, J. J., Zuiderweg, F. J., and Klinkenberg, A. "Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography". Chem. Eng. Sci., **5**, 271-289 (1956)

Vermeulen, T., Klein, G., and Hiester, N. K., Chemical Engineer's Handbook, R. H. Perry and C. H. Chilton (Eds.), McGraw-Hill, New York (1973)

Villermaux, J. "Chemical engineering approach to dynamic modelling of linear chromatography. A flexible method for representing complex phenomena from simple concepts". J. Chromatogr., **406**, 11-26 (1987)

Wakao, N., Oshima, T., and Yagi, S. "Mass transfer from packed beds of particles to a fluid". Chem. Eng. Jap., **22**, 780-788 (1958)

Warren, R. A. J. "Microbial hydrolisis of polysaccharides". Annu. Rev. Microbiol., **50**, 183-212 (1996)

Wilchek, M., and Bayer, E. A. Meth Enzymol., **184**, (1990)

Xu, G., Ong, E., Gilkes, N. R., Kilburn, D. G., Muhandiram, D. R., Harris-Brandts, M, Carver, J. P., Kay, L. E., and Harvey, T. S. "Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by Nuclear Magnetic Resonance Spectroscopy". Biochemistry. **34**, 6993-7009 (1995)

Yamane, T., and Shimizu, S. "Fed-batch techniques in microbial processes". Adv. Biochem. Eng., **30**, 147-194 (1984)

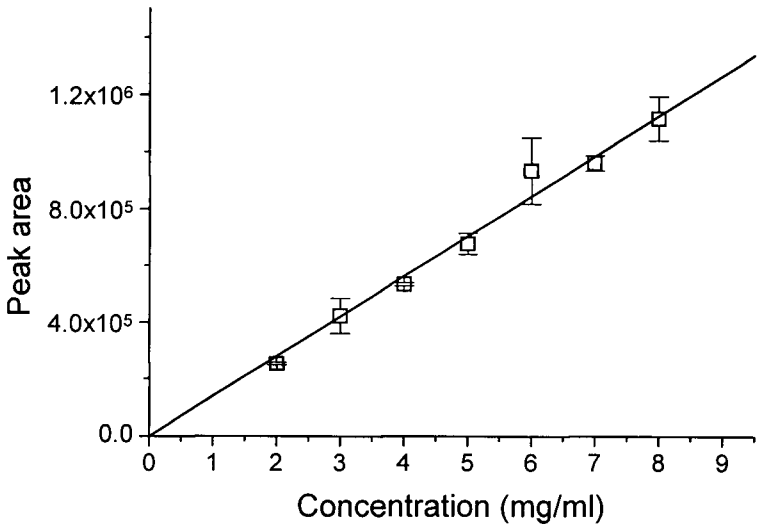
Yang, Y., Velayudhan, A., Ladisch, C. M., and Ladisch, M. R. "Protein chromatography using a continuous stationary phase". J. Chromatogr., **598**, 169-180 (1992)

Yannish-Perron, C., Vieira, J., Messing, J. "Improved m13 phage cloning vectors and host strains: nucleotide sequence of the m13mp18 and pUC19 vectors". Gene, **33**, 103-119 (1985)

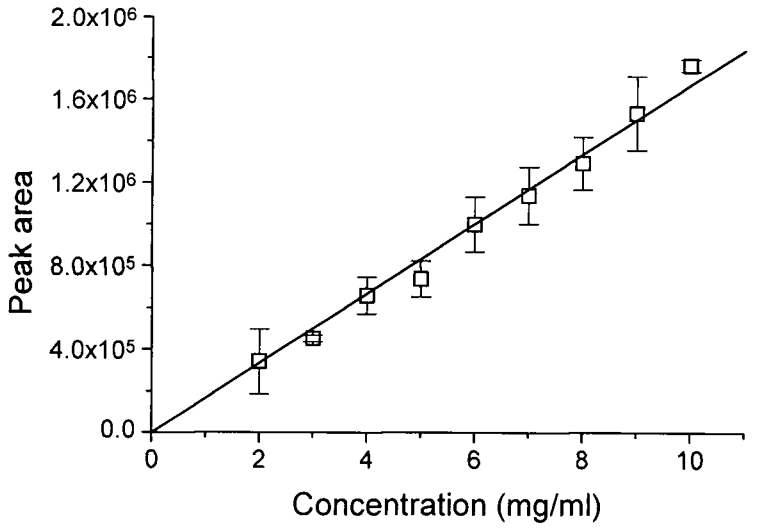
Yaron, S., Shimon, L. J., Frolow, F., Lamed, R., Morag, E., Shoham, Y., and Bayer, E. A. "Expression, purification and crystallization of a cohesin domain from the cellulosome of *Clostridium thermocellum*". L. Biothechnol., **51**, 243-249 (1996)

Yee, L., and Blanch, H. W. "Recombinant protein expression in high cell density fed-batch cultures of *Escherichia coli*." Bio/Technology, **10**, 1550-1556 (1992)

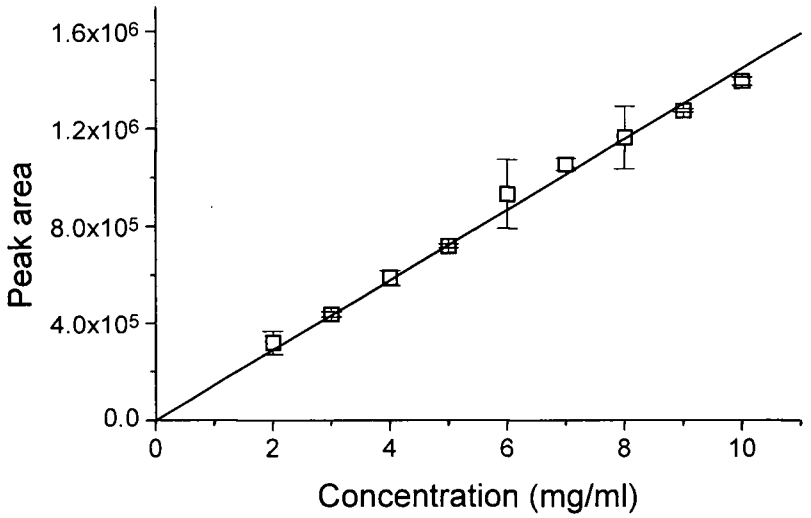
APPENDIX 1 CALIBRATION CURVES FOR THE SOLUTE EXCLUSION TECHNIQUE



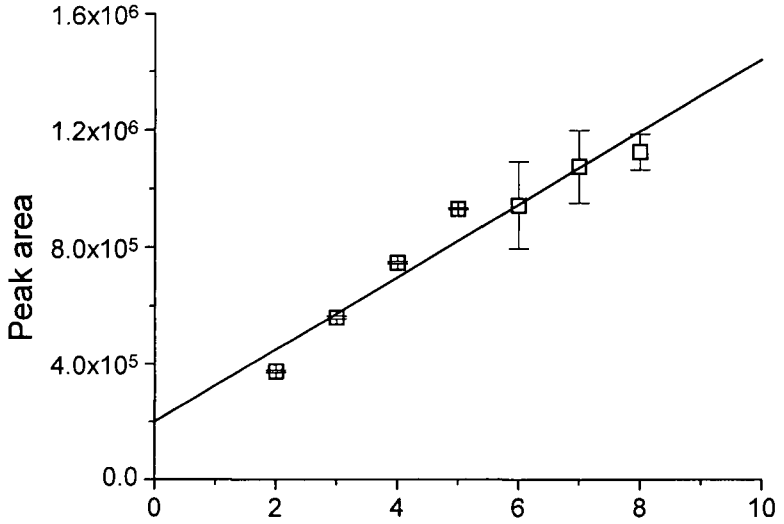
Calibration curve for PEG-200
y = 140373 x (R = 0.991)



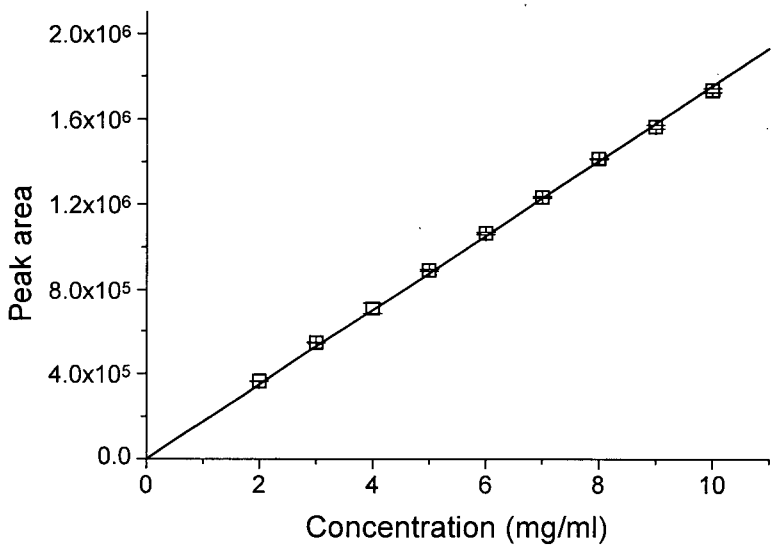
Calibration curve for PEG-400
y=166853 x (R=0.995)



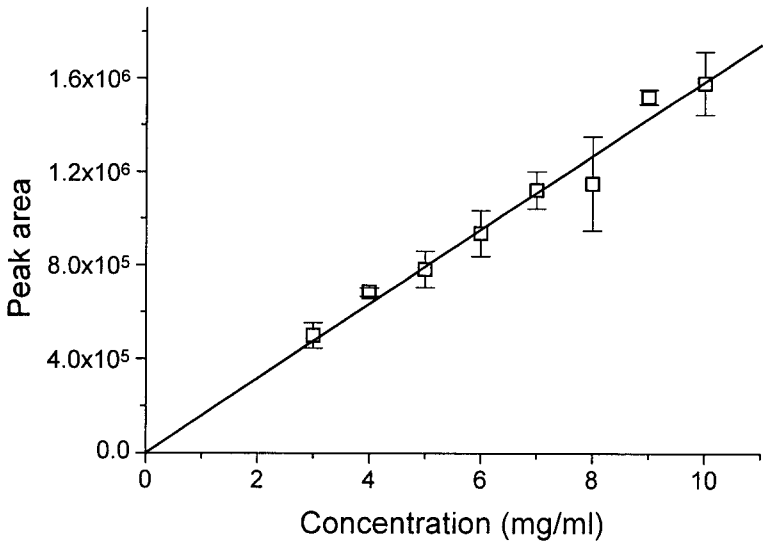
Calibration curve for PEG-600
y = 144379 x (R = 0.997)



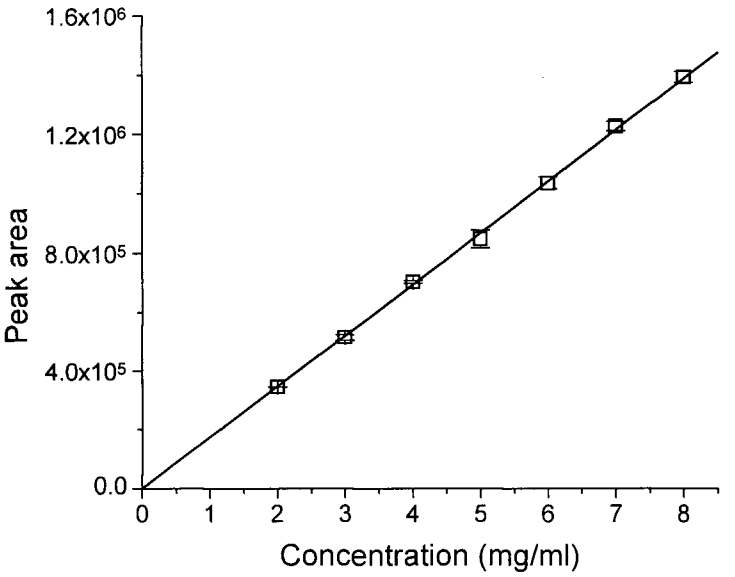
Calibration curve for PEG-1450
y=124035 x + 201549 (R=0.972)



Calibration curve for PEG-3350
y = 175274 x (R = 0.999)



Calibration curve for PEG-8000
y = 158493 x (R = 0.987)



Calibration curve for PEG-20 000

$y = 173113\ x$ $(R = 0.999)$