Modeling Adsorption and Transport Behavior in Cation Exchange and Hydrophobic Resins using Numerical Column Models

A dissertation presented to the faculty of the School of Engineering and Applied Science

University of Virginia

In partial fulfillment of the requirements for

the degree of Doctor of Philosophy in Chemical Engineering

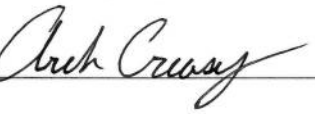
by

Arch Creasy

May 2018

APPROVAL SHEET

This Dissertation
is submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Author Signature: 

This Dissertation has been read and approved by the examining committee:

Advisor: Giorgio Carta

Committee Member: Geoffrey Geise

Committee Member: Gaurav Giri

Committee Member: Chris Highley

Committee Member: Teresa Culver

Committee Member: _____

Accepted for the School of Engineering and Applied Science:

Craig H. Benson, School of Engineering and Applied Science

May 2018

Abstract

Accurate prediction of protein chromatographic behavior is desirable for efficient and robust process development. In order to reliably predict the loading and elution behavior, an accurate description of adsorption equilibrium is required as a function of protein concentration and mobile phase composition. Traditionally, an isotherm model is used to describe the equilibrium behavior, but this approach can only be as accurate as the model itself. An alternative method is developed to predict protein chromatographic behavior from batch isotherm data that can be obtained in a high throughput process development (HTPD) mode using a systematic empirical interpolation (EI) scheme without relying on a mechanistic description of the dependence of protein binding on pH and mobile phase composition. A lumped kinetic model with rate parameters determined from HETP measurements or batch adsorption experiments can be coupled with the EI scheme to numerically predict the column elution behavior for individual or combined salt and pH gradients.

Several case studies for cation exchange chromatography are given in this work which demonstrate the EI method's general applicability to pH or salt elution, different proteins, and multi-component separations. Predictions based on the EI scheme show excellent agreement with experimental elution profiles under highly overloaded conditions for lysozyme on SP-Sepharose FF and two monoclonal antibodies (mAb) on POROS XS. Additionally, the EI method is extended to multicomponent separations and successfully predicts the separation of a monomer and dimer mAb on Nuvia HR-S.

Another major component of this dissertation is the investigation of protein retention in hydrophobic interaction chromatography. In general, an increase in kosmostropic salt concentration drives protein partitioning to the hydrophobic surface while a decrease reduces it. In some cases, however, protein retention also increases at low salt concentrations resulting in a U-shaped retention factor curve. During gradient elution the salt concentration is gradually decreased from a high value thereby reducing the retention factor and increasing the protein chromatographic velocity. For these conditions, a steep gradient can overtake the protein in the column, causing it to rebind. Two models, one based on the local equilibrium theory and the other based on the linear driving force approximation, are presented. The equilibrium behavior is described using the solvophobic

theory for cases with low protein concentrations while batch isotherm data is coupled with the EI scheme to describe cases at high protein concentrations. We show that the normalized gradient slope and protein load determines whether the protein elutes in the gradient, partially elutes, or is trapped in the column. Experimental results are presented for two different monoclonal antibodies and for lysozyme on Capto Phenyl (high sub) resin. One of the mAbs and lysozyme exhibit U-shaped retention factor curves and for each, we determine the critical gradient slope beyond where 100% recovery is no longer possible.

This dissertation demonstrates the broad applications of chromatographic modeling and how data from high-throughput automation can be properly leveraged for deeper process understanding and robust downstream process development.

Acknowledgments

I owe my deepest gratitude to my advisor, Giorgio Carta. Without your guidance, persistent help, and immense knowledge, this dissertation would not have been possible. You have inspired confidence and success that will forever be with me. I could not have imagined a better advisor and mentor for my Ph.D. study.

I would also like to thank members of my advisory committee: Prof. Geoffrey Geise, Prof. Gaurav Giri, Prof. Chris Highley, and Prof. Teresa Culver for their valuable input and discussions.

I must express a sincere thanks to my collaborators Greg Barker, Joseph Lomino, Joseph Calzada, and all the other Bristol-Myers Squibb employees who helped contribute to this work. I cannot overstate how valuable your devotion, effort, and contributions were to the success of this project. Thank you all for being so supportive of me and this research.

I would also like to thank all past and present members of Carta lab, especially Tarl, Yige, Frank, Mimi, Shaojie, Jing, Jason, Yiran, Andreas, Preston, Lucas, and Joey. Thank you all for being the best lab mates and making my time in the group exciting and fun.

To my fellow friends in and outside the department, thank you. My time in Charlottesville has been life changing. I'll miss our time playing volleyball, relaxing at the breweries, hiking the Blue Ridge Mountains, eating tacos at the farmers market, and so much more.

To Teresa, Jennifer, Vickie, Ricky, Ryan, and Jen, thank you for all your help with everything. No one would graduate without you. One of my favorite things about moving from Thornton A wing to the first floor of Chemical Engineering was getting to see you all every day.

Finally, I would like to thank my family, especially my parents, Nancy, Dave, and Cindy, and my siblings, Neala, Chris, and Matt, for their continuous support. Thank you for all your love and for always being there for me.

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbols

a	temperature dependent constant in Davies equation
A	equilibrium parameter in steric mass action model or solvophobic theory
b	isotherm parameter in Langmuir-Freundlich model or solvophobic theory
Bi	Biot number ($r_p = k_f / D_e$)
c	slope of van Deemter plot (HETP vs. u) or parameter in solvophobic theory
C	concentration in mobile phase (kg/m^3)
C_0	initial concentration in solution (kg/m^3)
C_F	feed concentration (kg/m^3)
C_M	mobile phase salt concentration (M)
$C_{M,0}$	initial mobile phase salt concentration (M)
$C_{M,f}$	final mobile phase salt concentration (M)
C_p	protein concentration in mobile phase (kg/m^3)
C_p^*	equilibrium protein concentration in mobile phase (kg/m^3)
$\overline{CV}$	mean retention volume in column volume units
CV_G	duration of the gradient in column volume units
CV_R	number of column volumes at elution
d_p	particle diameter (m)
D_0	diffusivity in free solution (m^2/s)
D_e	effective pore diffusivity (m^2/s)
I	ionic strength (M)
k	binding rate constant (s^{-1})
k'	retention factor
k'_0	retention factor of protein at zero salt concentration
k'_M	retention factor of salt
k_f	external mass transfer coefficient (m/s)
K	Henry's constant or empirical parameter in Langmuir-Freundlich model
$K_{a,n}$	thermodynamic equilibrium constant

$K'_{a,n}$ apparent equilibrium constant

K_D distribution coefficient

K_e equilibrium constant in steric mass action model

K_L equilibrium constant in multicomponent Langmuir equation

K'_w apparent ionic product of water

K_∞ isotherm parameter in quadratic Langmuir equation

N valency of charged species

N_C number of buffer components

N_{eff} effective number of plates

N_i actual number of plates

q total bound protein concentration (kg/m^3 of particle volume)

q_0 resin charge density (M)

q_m binding capacity in Langmuir equation (kg/m^3 of particle volume)

q_p total bound protein concentration (kg/m^3 of particle volume)

$\bar{q}_p$ total protein concentration held in adsorbent particle (kg/m^3 of particle volume)

r radial coordinate (m)

r_p particle radius (m)

Sh Sherwood number ($= k_f d_p / D_0$)

u superficial mobile phase velocity (m/s)

v charge of ion or mobile phase interstitial velocity (m/s)

v_c chromatographic velocity (m/s)

V_C column volume (m^3)

V_F protein feed volume (m^3)

V_r resin volume (m^3)

V_s solution volume (m^3)

t time (s)

t_G duration of gradient (s)

x column axial coordinate (m)

z effective binding charge

Greek symbols

α	empirical parameter in quadratic Langmuir equation
β	gradient slope (M/s)
δ	empirical parameter in quadratic Langmuir equation
ε	column void fraction
ε_p	intraparticle porosity
θ	empirical parameter in Langmuir equation
γ	activity coefficient or normalized gradient slope (M)
λ	empirical parameter in model from solvophobic theory
σ	steric hindrance factor
τ	empirical parameter in quadratic Langmuir equation
ϕ	column phase ratio, $(1-\varepsilon)/\varepsilon$
ω	empirical parameter in quadratic Langmuir equation

1 Introduction and Objectives

1.1 Introduction

1.1.1 Therapeutic monoclonal antibodies

Monoclonal antibodies (mAbs) have become major biotherapeutic products over the last decade and are estimated to further accelerate to a $140 billion worldwide market within the next decade[1]. Currently there are over 74 antibody-based molecules approved in either the European Union or the United States and the number of mAbs granted a first approval reached 10 for the first time in 2017 with another 12 new mAbs expected to be submitted by the end of 2018[2].

From a therapeutic standpoint, the success of mAbs is largely due to their unique antigen binding specificity which has greatly increased their use for treating a range of ailments including cardiovascular and autoimmune diseases, cancer, and inflammation[4–7]. There are many antibody based molecules such as naked IgGs, bispecific antibodies, Fc fusion proteins, antibody fragments, antibody-drug conjugates (ADC), and immunomodulatory antibodies[3]. The most exciting therapies are ADCs and immunomodulatory antibodies which both have found recent clinical success thanks to the advancements of protein engineering and better understanding of the immune system and cancer cells[8,9]. Antibody drug conjugates are cytotoxic drugs conjugated with mAbs designed to target specific antigens[10]. These drugs are intended to minimize side effects by only targeting cancer cells and disregarding healthy cells. Currently there are 87 ADCs in phase I/II or phase III clinical trials with 3 currently approved by the FDA[3].

Antibodies are also being used to indirectly target cancer by augmenting the immune response against tumor cells using immuno-oncology based mAbs[11]. Clinical studies have shown immunology drugs may be able to treat many different types of advance stage cancer since they target/activate T-cell function, a mechanism that isn't necessarily specific to any particular cancer. Currently, there are over 80 antibodies in clinical trials that modulate T-cell response[3]. This method of mAb binding to the T-cell regulatory receptors to stimulate the immune response has resulted in antitumour

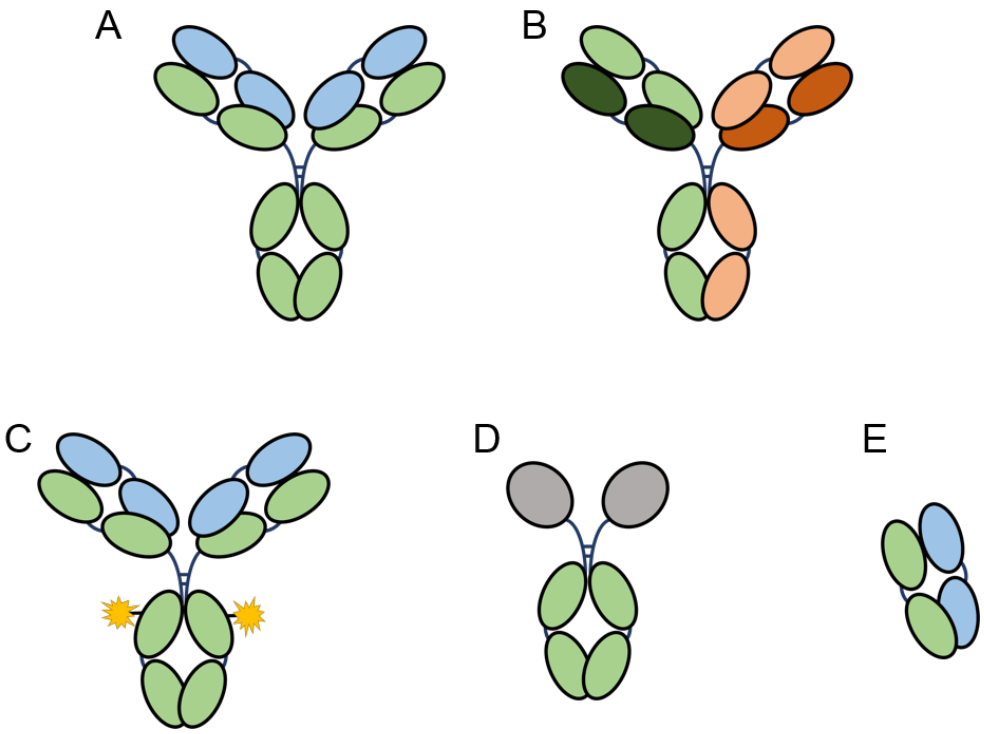


Figure 1.1: Structures of antibody-based molecules adapted from Strohl[3]. (A) IgG monoclonal antibody; (B) Heterodimeric IgG-based bivalent, bispecific antibody; (C) Antibody drug conjugate (cytotoxin denoted by yellow star); (D) Fc-protein fusion (protein denoted by gray oval); (E) FAb fragment

activity and is now approved by the FDA to treat patients with melanoma and non-small-cell lung cancer[12].

1.1.2 Downstream purification

Although modern biopharmaceutical drugs have shown to be potentially effective against many currently untreatable ailments, production costs for mAb based therapeutics are particularly high, mainly caused by the requirement to ensure the delivery of a safe and effective drug to the patient. The purity of the drug product is ultimately dictated by the downstream purification process, which often includes unit operations such as chromatographic separations to remove process impurities. Compared to other biologics, mAbs have been uniquely successful because they are ideal drug candidates for generic "platform purification" processes which rely primarily on the Protein A capture step that binds to the mAb's highly conserved Fc region with remarkable specificity[9,13,14]. The platform process often contains additional polishing steps such as cation exchange (CEX) or anion exchange (AEX) chromatography, hydrophobic interaction chromatography (HIC), or multimodal chromatography (MMC) to remove product related impurities such as charge isoforms, fragments, and aggregates.

Soluble aggregates are common impurities many mAb purification processes must remove as these impurities can be immunogenic, can have diminished therapeutic activity, and can result in reduced stability of drug products[15–17]. While large aggregates are relatively easy to remove, dimers and small oligomers can be challenging since they have molecular surface properties similar to those of the monomer. Although the molecular mass is, of course, larger than that of the monomer, molecular size is not very different since dimers have a dense structure making separations by size exclusion chromatography (SEC) or membrane filtration impractical. As a result, dimer removal is most often achieved by chromatography in a bind and elute mode. Various chromatographic modalities can be used, including ion exchange, hydrophobic interaction, and hydroxyapatite. However, since many mAbs have relatively high pI, CEX chromatography is most often used[18–20]. In CEX chromatography, separation is typically achieved with a salt gradient although pH gradient elution[21–23] and

frontal analysis[24] have also been used.

Successful removal of product and process related impurities such as misfolded isoforms, protein fragments, and aggregates has also been found using hydrophobic interaction chromatography both in the laboratory and at process scale[25–27]. The adsorption process is based on the reversible partitioning of proteins and other biopolymers between a mobile phase containing a kosmotropic salt, such as ammonium sulfate, and a mildly hydrophobic stationary phase. High kosmotropic salt concentrations promote the interaction between the immobilized hydrophobic ligands of the resin and the hydrophobic patches on the surface of the protein. Typically, the protein can be eluted with a step or linear gradient to low salt concentrations; however, in this work, we will show some proteins may exhibit retention at low salt. Understanding the adsorption equilibrium's dependence on the salt concentration is necessary to predict chromatographic behavior.

1.1.3 Predicting chromatographic behavior

The ability to predict the chromatographic behavior of proteins has become increasingly desirable in the pharmaceutical industry because it can help rationally guide process development and design and establish a robust operating space which satisfies USFDA requirements. The common approach to predicting chromatographic behavior involves employing an isotherm model to describe the effects of protein concentration and mobile phase composition on protein adsorption equilibrium and a rate model to describe the adsorption kinetics. At low protein loads, in the linear limit of the binding isotherm, equations provided by Yamamoto et al.[28–30] and by other authors[31,32] predict analytically the elution profiles with excellent accuracy. However, at higher protein loads, which is more relevant to industrial manufacturing, the isotherm is non-linear and elution profiles are substantially different from those predicted by the linear isotherm models. In this case, elution begins to occur earlier than predicted in the linear limit and the peaks become asymmetrical[33–35].

In principle, whether elution occurs in the linear region of the isotherm or under overloaded conditions, the elution profiles can be predicted on the basis of thermodynamic isotherm models. For ion exchange chromatography, several isotherm models are available including Langmuir[34,36–39],

steric mass action (SMA)[33,40,41], and statistical thermodynamic (ST) models[42–44]. In the case of HIC, retention behavior can be typically described by the solvophobic theory[45] in dilute conditions and by Langmuirian based models[46] for high protein concentrations. Similarly, various rate models are also available including models that assume that adsorption is kinetically limited, lumped rate models based on a linear driving force approximation, and the general rate model that describes the detailed mechanisms associated with diffusional mass transfer[26,47,48].

1.1.4 Modeling ion exchange chromatography

In the case of ion exchange chromatography, several examples are available in the literature showing the use of models to predict elution behavior from salt gradients at high protein loads. Gallant et al.[33], for instance, developed a model to predict the elution behavior of mixtures of α-chymotrypsinogen a, cytochrome c, and lysozyme based on the SMA isotherm model. These authors combined isotherm parameters from measurements in the linear isotherm regime with isotherm measurements obtained at high protein loads from nonlinear frontal analysis experiments. The column behavior was then simulated assuming local equilibrium, which yielded predictions that were qualitatively accurate but limited to the case of small resin bead diameter (15 μm). Müller-Späth et al.[34] modeled the chromatographic behavior of a polyclonal antibody on Fractogel SO3 (65 μm diameter) and POROS HS (50 μm diameter) with a Langmuir model. Similar to Gallant et al., the model parameters were determined by combining the results of chromatographic experiments conducted at low protein loads with those obtained from a limited number of overloaded isocratic elution experiments. Their model could successfully predict the salt gradient elution behavior but only for conditions that resulted in peak concentrations below 12 mg/ml.

Gradients in pH and combined pH-salt gradients are also sometimes used to elute proteins from ion exchange columns[21,39,49,50]. While somewhat more difficult to implement, pH gradients allow elution at low ionic strengths[23,51] and can provide greater selectivity especially when used in conjunction with multimodal resins[22,52]. Implementing and predicting protein elution with pH gradients involves two distinct modeling aspects: (a) designing buffer systems that exhibit suitably

linear pH gradients with controlled ionic strength, and (b) predicting protein binding as a function of pH.

The ability to generate predictable linear pH gradients based on modeling the mobile phase chemistry has already been demonstrated by several authors[53–56]. Kröner and Hubbuch[55], for example, using complex buffer mixtures, developed a model to predict the compositions of initial and final buffers that generate linear pH gradients with constant buffering capacity when linear mixing of the two buffers is used. Since buffering capacity rather than counterion concentration or ionic strength is kept constant, predicting protein elution becomes more challenging since the binding strength is simultaneously affected by the varying pH and counterion concentration. Moreover, having independent control of pH and ionic strength is desirable for applications where elution occurs at higher ionic strengths or when elution with simultaneous pH and salt gradients is sought to improve resolution[50,52,57,58]. A modified buffer design strategy is developed in this work to address this issue.

With regards to predicting protein binding, describing the effect of pH is typically challenging due to the multiple interacting forces at play. Some theoretical models are available but are generally subject to fairly limiting simplifying assumptions. Guélat et al.[39,59], for example, developed a model for overloaded monoclonal antibody variants on a CEX resin by treating the protein as a colloidal sphere, whose net charge is determined by pH via the deprotonation of all basic and acidic residues, which interacts with a flat surface with homogeneous surface charge density. Their model was capable of predicting the salt gradient and pH gradient elution behavior but required a substantial number of fitted parameters obtained from a large set of chromatographic data at both low and high loadings. Kluters et al.[60] developed a simpler approach based on the SMA model where pH is assumed to affect the protein effective charge through the deprotonation equilibria of a set of acidic and based residues. Fitting the number of each charged residues in this set to chromatographic data provided a functional description of pH and salt gradient elution. However, they did not validate high protein load conditions with a combined pH-salt gradient. In yet another approach, Vetter et al.[61] fitted the SMA model to isotherm data describing mAb binding to a CEX resin as a function of protein and salt concentration and then correlated the dependence of the protein binding charge

on pH using a linear empirical relationship.

Another challenge in the prediction of chromatographic behavior is that of multicomponent separations. The theory of linear gradient elution at low protein loads, in situations were separation occurs in the linear (Henry's law) region of the isotherm, is well established[29,31,32,62,63]. For these conditions, species behave essentially independently of each other, so that the separation can be predicted from the single component behavior. On the other hand, at high protein loads, competitive adsorption effects and the complex effects of salt concentration on the multicomponent isotherm become important requiring models of greater complexity[19,20,64].

The most common approach for modeling separation at high protein loads involves using a thermodynamic model to describe both competitive isotherms and the effects of salt concentration on binding coupled with a rate model to describe adsorption and desorption kinetics. Both Langmuir isotherm models and the steric mass action law (SMA) model[40] have been used extensively to describe multicomponent adsorption coupled with diffusional mass transfer kinetics[34,59,60,65]. For example, Guélat et al.[59] modeled multicomponent adsorption of antibody charge variants on cation exchange resins using a modified Langmuir equation and statistical thermodynamics expressions that treat protein adsorption as spheres interacting with a flat, charged surface. Their model could predict the resolution of four antibody charged variants at low loading. However, modeling at high protein loads required repeated refinement of the isotherm parameters by fitting model predictions to overloaded peak profiles. Examples of applications of the SMA model to describe high-load separations can be found in refs.[19,60,61,64,65]. Tao et al.[19,64] modeled the adsorption kinetics and equilibrium behavior of deamidated antibody variants on a CEX resin using the SMA model with parameters determined from single component isotherm measurements generated using purified samples of each variant. Kluters et al.[60] also used a modified SMA model to describe competitive binding with antibody monomer-aggregate mixtures, but their model parameters were determined from a combination of fitting experimental gradient elution profiles at low and high loads. A disadvantage of this approach is that obtaining high protein load elution profiles in a column requires substantial amounts of protein since fractions containing significant protein concentrations are needed to quantify monomer and aggregate content, usually by off-line analytical

SEC. Multiple such experiments are obviously needed in order to develop a robust parameter set. Moreover, the overall predictive accuracy would still be limited by the model's own simplifying assumptions. For example, previous work has also shown the SMA model could not sufficiently describe the low selectivity between a monomer and dimer antibody on a cation exchange resin at low salt concentrations[66]. At low salt, the monomer and dimer species were found to have similar affinity for the resin, but the selectivity toward the dimer increased at higher salt following a trend which is not predicted by the SMA model.

The common feature of the above modeling approaches is that the parameters of a simplified mechanistic model are fitted to experimental data at different pH and salt concentration values. In this work we test, as an alternative approach, the effectiveness of using experimental protein adsorption isotherms directly to predict column behavior without using a mechanistic model and thus removing any constraints imposed by a model's simplifying assumptions. The method involves collecting isotherm data at different salt concentrations and/or pH values, fitting the data at each salt concentration and pH with an empirical single component or multicomponent isotherm expression, and then predicting the effect of salt concentration and pH by interpolation.

Of course, a drawback of interpolation is that an extensive set of isotherm data is needed. The recent advancement of automated high throughput screening (HTS) coupled with analytical tools such as high throughput UV measurements and Ultra Performance Liquid Chromatography (UPLC) SEC and CEX analysis helps generate large data sets to interpolate the effects of protein concentration, salt concentration, and pH[67–70]. The data are systematically interpolated in a multi-dimensional experimental space and used in conjunction with a simplified mass transfer model to predict column elution with single or multiple components and with individual or combined pH and salt gradients.

In Section 2 of this work, we use two experimental cation exchange systems, lysozyme on SP-Sepharose-FF and two monoclonal antibodies on POROS XS, as models to test these ideas. A rate model is coupled with an empirical interpolation (EI) scheme to numerically predict the column elution behavior for individual or combined salt and pH gradients. Next, in Section 3 we consider the separation of a mAb monomer-dimer mixture by CEX. Monomer-dimer isotherm data are ob-

tained batchwise using analytical SEC to determine the ratio of monomer and dimer bound. The EI prediction of two-component binding are coupled with a rate model to predict the multicomponent column elution profiles for a range of protein loads.

1.1.5 Modeling hydrophobic interaction chromatography

As a final component to this work, we investigate the elution behavior of monoclonal antibodies on HIC and present a method to predict the chromatographic behavior. Normally in HIC, increasing the concentration of a kosmotropic salt increases partitioning toward the stationary phase, while reducing it increases partitioning toward the mobile phase. The effect of the kosmotropic salt concentration on protein retention in HIC is thus inverse to the well-known effect of kosmotropes on protein solubility in aqueous solutions[71].

Because of the strong dependence of protein retention on salt concentration, HIC is often operated in gradient mode starting at a high salt concentration and ending with a low salt concentration or, often, with no salt[45]. Stripping with water is even recommended[72]. Gradient operation generally improves robustness and helps prevent irreversible binding and/or denaturation as a result of strong interactions with the mobile or stationary phase or with both[45]. Of course, for relatively easy separations, HIC can also be operated in a step mode with an abrupt change from high salt conditions, used for protein binding, to no salt to facilitate elution of strongly retained species. There have also been successful attempts to use HIC resins for protein separations with low concentrations of a kosmotropic salt[73] or even without a kosmotropic salt in the mobile phase[74]. In this case, protein binding results from "hydrophobic affinity"[75], rather than salting in/out effects, which in turn, can be modulated by varying pH or supplementing the mobile phase with certain additives in order to modulate retention[74]. One difficulty is that without addition of a kosmotropic salt, protein binding is often very weak so that, in practice, as noted by Kato et al.[73] a high initial kosmotrope concentration is used.

Although protein retention in HIC is typically regarded as a monotonic function of kosmotropic salt concentration[76], in some cases, protein retention can dramatically increase at very low salt

concentrations resulting in a U-shaped retention factor curve. This behavior can become a concern in process operations since using a gradient from high to low salt concentrations can result in conditions where the protein instead of eluting becomes trapped in the column, strongly retained at low salt. For the case of low protein loads, we show that with U-shaped retention factor curves, the normalized gradient slope determines whether the protein will elute in the gradient, will partially elute, or will essentially never elute becoming trapped in the column for conditions where the retention factor increases as the kosmotrope concentration decreases. A local equilibrium analysis of this chromatographic behavior is presented along with predictions based on the numerical solution of the general rate model of chromatography. Experimental results are presented for two different monoclonal antibodies, mAb C and mAb D, and for lysozyme on Capto Phenyl (high sub) resin. mAb D and lysozyme exhibit U-shaped retention factor curves and for each of them, we determine the critical gradient slope beyond where 100% recovery is no longer possible. The remaining mAb C does not exhibit this behavior and elutes at virtually any gradient slope.

Additionally, we will investigate how U-shaped retention behavior extends to conditions with higher protein loads. Since there is a finite binding capacity in the column, it is expected peaks will breakthrough earlier and at higher ammonium sulfate concentrations as the load increases. Thus, it is expected the protein recovery will be directly affected by the column load. To understand and ultimately predict elution profiles at higher protein loads, we will measure batch isotherm data at high protein concentrations and salt concentrations which span the conditions from load to elute. In this work, we measured batch isotherms of mAb D on Capto Phenyl (high sub) at high protein concentrations and applied the EI method to predict LGE behavior for a range of column loadings up to 30% of the equilibrium binding capacity and for a range of gradient slopes. Understanding this behavior has important implications for designing gradient elutions since the protein load and slope of the gradient directly impacts the recovery of the mAb and the fraction of protein left behind which can eventually foul the column. As a final component to this work, we briefly investigate the fouling behavior on Capto Phenyl (high sub) using confocal laser scanning microscopy (CLSM) and show the effects of fouling on the LGE behavior of mAb D. This work highlights the need to develop cleaning methods that fully regenerate the column so that predictions and column performance

remain robust over many process cycles.

1.2 Research objectives

The ability to predict the chromatographic behavior of proteins is desirable for process development and to help establish a robust operating space. The classical strategy normally used for this purpose involves employing an isotherm model to describe the effects of protein concentration and mobile phase composition on protein adsorption equilibrium and a rate model to describe the adsorption kinetics. However, the accuracy of this approach is explored in this work along with another approach that does away with an isotherm model and uses batch isotherm data directly through a suitable interpolation scheme taking advantage of the typically vast matrix of experimental measurements that nowadays can be obtained through automated high-throughput screening (HTS) equipment. This work is aimed toward accurate quantitative prediction of chromatographic behavior based on simple isotherm models and interpolated HTS isotherm data to predict chromatography operations and define critical process parameters.

The specific objectives of this dissertation are to:

1. Develop an empirical interpolation (EI) method that allows direct use of high-throughput batch isotherm data to predict protein elution behavior for salt gradients on a strong cation exchanger.

    (a) Validate the EI method using two experimental systems: lysozyme on SP-Sepharose FF and a monoclonal antibody on POROS XS.

    (b) Compare the EI model predictions with those obtained using the mechanistic steric mass action model.

2. Extend EI method to a three-dimensional space that describes protein adsorption as a function of protein concentration, salt concentration, and pH using HTS experimental data. Validate this extended EI method for the prediction of pH and salt gradient elution on a strong cation exchanger. Accurate prediction of the generated pH gradient is necessary.

    (a) First develop an optimization method to calculate buffer compositions required for highly linear pH gradients.

   (b) Predict individual and combined salt and pH gradient elution behavior for a monoclonal antibody on POROS XS.

3. Extend EI method to describe multicomponent isotherm data and predict multicomponent elution profiles at high protein loads

   (a) Model separation of a monoclonal antibody monomer and dimer on Nuvia-HR S eluted with a salt gradient.

4. Predict elution behavior in hydrophobic interaction chromatography (HIC)

   (a) Describe gradient elution behavior of proteins with U-shaped retention factor curves at low protein loads.

   (b) Predict high load gradient elution behavior of a mAb on Capto Phenyl HS using the EI method.

References

(*1*) Grand View Research, Monoclonal Antibodies (mAbs) Market Size, Industry Report, 2013-2024., tech. rep., Grand View Research, 2016.

(*2*) Jarvis, L. M., (2018). FDA approvals hit a 20-year high in 2017. *Chemical & Engineering News 96*, 26–30.

(*3*) Strohl, W. R., (2018). Current progress in innovative engineered antibodies. *Protein & Cell 9*, 86–120.

(*4*) Desai, N. R., Giugliano, R. P., Zhou, J., Kohli, P., Somaratne, R., Hoffman, E., Liu, T., Scott, R., Wasserman, S. M., and Sabatine, M. S., (2014). AMG 145, a monoclonal antibody against PCSK9, facilitates achievement of national cholesterol education program-adult treatment panel III low-density lipoprotein cholesterol goals among high-risk patients. *Journal of the American College of Cardiology 63*, 430–433.

(*5*) Mizutani, N., Nabe, T., and Yoshino, S., (2015). IgE/antigen-mediated enhancement of IgE production is a mechanism underlying the exacerbation of airway inflammation and remodelling in mice. *Immunology 144*, 107–115.

(*6*) Gallagher-Colombo, S. M., Miller, J., Cengel, K. A., Putt, M. E., Vinogradov, S. A., and Busch, T. M., (2015). Erlotinib pretreatment improves photodynamic therapy of non-small cell lung carcinoma xenografts via multiple mechanisms. *Cancer Research 75*, 3118–3126.

(*7*) Savic, S., Diebold, J., Zimmermann, A.-K., Jochum, W., Baschiera, B., Grieshaber, S., Tornillo, L., Bisig, B., Kerr, K., and Bubendorf, L., (2015). Screening for ALK in non-small cell lung carcinomas: 5A4 and D5F3 antibodies perform equally well, but combined use with FISH is recommended. *Lung Cancer 89*, 104–9.

(*8*) Scott, A. M., et al. (2007). A phase I clinical trial with monoclonal antibody ch806 targeting transitional state and mutant epidermal growth factor receptors. *104*, 4071–4076.

(*9*) Scott, A. M., Wolchok, J. D., and Old, L. J., (2012). Antibody therapy of cancer. *Nature Reviews Cancer 12*, 278–287.

(*10*) Hughes, B., (2010). Antibody – drug conjugates for cancer : poised to deliver ? *Nature Reviews Drug Discovery 9*, 665–667.

(*11*) Weiner, L. M., Surana, R., and Wang, S., (2010). Monoclonal antibodies : versatile platforms for cancer immunotherapy. *Nature Reviews Immunology 10*, 317–327.

(*12*) Khalil, D. N., Smith, E. L., Brentjens, R. J., and Wolchok, J. D., (2016). The future of cancer treatment: immunomodulation, CARs and combination immunotherapy. *Nature Reviews Clinical Oncology 13*, 273–290.

(*13*) Reichert, J. M., Rosensweig, C. J., Faden, L. B., and Dewitz, M. C., (2005). Monoclonal antibody successes in the clinic. *Nature Biotechnology 23*, 1073–1078.

(*14*) Gaughan, C. L., (2016). The present state of the art in expression, production and characterization of monoclonal antibodies. *Molecular Diversity 20*, 255–270.

(*15*) Gramer, M. J., (2014). Product Quality Considerations for Mammalian Cell Culture Process Development and Manufacturing. *Advances in Biochemical Engineering-Biotechnology 139*, 123–166.

(*16*) Bessa, J., Boeckle, S., Beck, H., Buckel, T., Schlicht, S., Ebeling, M., Kiialainen, A., Koulov, A., Boll, B., Weiser, T., Singer, T., Rolink, A. G., and Iglesias, A., (2015). The immunogenicity of antibody aggregates in a novel transgenic mouse model. *Pharmaceutical Research 32*, 2344–2359.

(*17*) Haile, L. A., Puig, M., Polumuri, S. K., Ascher, J., and Verthelyi, D., (2017). In Vivo Effect of Innate Immune Response Modulating Impurities on the Skin Milieu Using a Macaque Model: Impact on Product Immunogenicity. *Journal of Pharmaceutical Sciences 106*, 751–760.

(*18*) Tao, Y., Carta, G., Ferreira, G., and Robbins, D., (2011). Adsorption of deamidated antibody variants on macroporous and dextran-grafted cation exchangers: I. Adsorption kinetics. *Journal of Chromatography A 1218*, 1519–1529.

(*19*) Tao, Y., Perez, E. X., Carta, G., Ferreira, G., and Robbins, D., (2011). Adsorption kinetics of deamidated antibody variants on macroporous and dextran-grafted cation exchangers . III . Microscopic studies. *Journal of Chromatography A 1218*, 8027–8035.

(*20*) Martin, C., Iberer, G., Ubiera, A., and Carta, G., (2005). Two-component protein adsorption kinetics in porous ion exchange media. *Journal of Chromatography A 1079*, 105–115.

(*21*) Farnan, D., and Moreno, G. T., (2009). Multiproduct high-resolution monoclonal antibody charge variant separations by pH gradient ion-exchange chromatography. *Analytical Chemistry 81*, 8846–8857.

(*22*) Holstein, M. A., Nikfetrat, A. A. M., Gage, M., Hirsh, A. G., and Cramer, S. M., (2012). Improving selectivity in multimodal chromatography using controlled pH gradient elution. *Journal of Chromatography A 1233*, 152–155.

(*23*) Talebi, M., Nordborg, A., Gaspar, A., Lacher, N. A., Wang, Q., He, X. Z., Haddad, P. R., and Hilder, E. F., (2013). Charge heterogeneity profiling of monoclonal antibodies using low ionic strength ion-exchange chromatography and well-controlled pH gradients on monolithic columns. *Journal of Chromatography A 1317*, 148–154.

(*24*) Reck, J. M., Pabst, T. M., Hunter, A. K., and Carta, G., (2017). Separation of antibody monomer-dimer mixtures by frontal analysis. *Journal of Chromatography A 1500*, 96–104.

(*25*) Shukla, A. A., and Thommes, J., (2010). Recent advances in large-scale production of monoclonal antibodies and related proteins. *Trends in Biotechnology*, 253–261.

(*26*) Carta, G., and Jungbauer, A., *Protein Chromatography: Process Development and Scale-Up*, 1st ed.; Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 2010.

(*27*) Cramer, S. M., and Holstein, M. A., (2011). Downstream bioprocessing : recent advances and future promise. *Current Opinion in Chemical Engineering 1*, 27–37.

(*28*) Yamamoto, S., Nakanishi, K., Matsuno, R., and Kamikubo, T., (1983). Ion exchange chromatography of proteins-prediction of elution curves and operating conditions. 1. Theoretical considerations. *Biotechnology and Bioengineering 25*, 1465–83.

(*29*) Yamamoto, S., Nakanishi, K., Matsuno, R., and Kamikubo, T., (1983). Ion-exchange chromatography of proteins-prediction of elution curves and operating conditions. 2. Experimental-verification. *Biotechnology and Bioengineering 25*, 1373–1391.

(*30*) Yamamoto, S., Nakanishi, K., and Matsuno, R., *Ion Exchange Chromatography of Proteins*; Marcel Dekker: New York, NY, 1988.

(*31*) Frey, D. D., (1990). Asymptotic Relations for Preparative Gradient Elution Chromatography of Biomolecules. *Biotechnology and Bioengineering 35*, 1055–1061.

(*32*) Carta, G., and Stringfield, W. B., (1992). Analytic solution for volume-overloaded chromatography. *Journal of Chromatography 605*, 151–159.

(*33*) Gallant, S., Vunnum, S., and Cramer, S. M., (1996). Optimization of preparative ion-exchange chromatography of proteins: linear gradient separations. *Journal of Chromatography A 725*, 295–314.

(*34*) Müller-Späth, T., Ströhlein, G., Aumann, L., Kornmann, H., Valax, P., Delegrange, L., Charbaut, E., Baer, G., Lamproye, A., Jöhnck, M., Schulte, M., and Morbidelli, M., (2011). Model simulation and experimental verification of a cation-exchange IgG capture step in batch and continuous chromatography. *Journal of Chromatography A 1218*, 5195–204.

(*35*) Creasy, A., Barker, G., Yao, Y., and Carta, G., (2015). Systematic interpolation method predicts protein chromatographic elution from batch isotherm data without a detailed mechanistic isotherm model. *Biotechnology Journal 10*, 1400–1411.

(*36*) Orellana, C. A., Shene, C., and Asenjo, J. A., (2009). Mathematical modeling of elution curves for a protein mixture in ion exchange chromatography applied to high protein concentration. *Biotechnology and Bioengineering 104*, 572–81.

(*37*) Forrer, N., Butté, A., and Morbidelli, M., (2008). Chromatographic behavior of a polyclonal antibody mixture on a strong cation exchanger column. Part II: Adsorption modelling. *Journal of Chromatography A 1214*, 71–80.

(*38*) Lan, Q., Bassi, A. S., Zhu, J.-X., and Margaritis, A., (2001). A modified Langmuir model for the prediction of the effects of ionic strength on the equilibrium characteristics of protein adsorption onto ion exchange/affinity adsorbents. *Chemical Engineering Journal 81*, 179–186.

(*39*) Guélat, B., Ströhlein, G., Lattuada, M., Delegrange, L., Valax, P., and Morbidelli, M., (2012). Simulation model for overloaded monoclonal antibody variants separations in ion-exchange chromatography. *Journal of Chromatography A 1253*, 32–43.

(*40*) Brooks, C. A., and Cramer, S. M., (1992). Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE Journal 38*, 1969–1978.

(*41*) Osberghaus, A., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., and Hubbuch, J., (2012). Determination of parameters for the steric mass action model–a comparison between two approaches. *Journal of Chromatography A 1233*, 54–65.

(*42*) Guélat, B., Delegrange, L., Valax, P., and Morbidelli, M., (2013). Model-based prediction of monoclonal antibody retention in ion-exchange chromatography. *Journal of Chromatography A 1298*, 17–25.

(*43*) Oberholzer, M. R., and Lenhoff, A. M., (1999). Protein adsorption isotherms through colloidal energetics. *Langmuir 15*, 3905–3914.

(*44*) Su, X., and Sun, Y., (2006). Thermodynamic model for nonlinear electrostatic adsorption equilibrium of protein. *AIChE Journal 52*, 2921–2930.

(*45*) Melander, W. R., Corradini, D., and Horvath, C., (1984). Salt-Mediated Retention of Proteins in Hydrophobic-Interaction Chromatography. *Journal of Chromatography 317*, 67–85.

(*46*) Xia, F., Nagrath, D., and Cramer, S. M., (2003). Modeling of adsorption in hydrophobic interaction chromatography systems using a preferential interaction quadratic isotherm. *Journal of Chromatography A 989*, 47–54.

(*47*) Guiochon, G., Felinger, A., Shirazi, D., and Katti, A., *Fundamentals of Preparative and Nonlinear Chromatography*, 2nd ed.; Academic Press: Boston, MA, 2006.

(*48*) LeVan, M., and Carta, G., In *Perry's Chemical Engineers' Handbook*, Green, D., Ed., 8th ed.; McGraw-Hill: New York, 2007; Chapter 16.

(*49*) Kang, X., and Frey, D. D., (2004). Chromatofocusing of peptides and proteins using linear pH gradients formed on strong ion-exchange adsorbents. *Biotechnology and Bioengineering 87*, 376–387.

(*50*) Zhou, J. X., Dermawan, S., Solamo, F., Flynn, G., Stenson, R., Tressel, T., and Guhan, S., (2007). pH-conductivity hybrid gradient cation-exchange chromatography for process-scale monoclonal antibody purification. *Journal of Chromatography A 1175*, 69–80.

(*51*) Pabst, T. M., Carta, G., Ramasubramanyan, N., Hunter, A. K., Mensah, P., and Gustafson, M. E., (2008). Separation of protein charge variants with induced pH gradients using anion exchange chromatographic columns. *Biotechnology Progress 24*, 1096–1106.

(*52*) Lee, Y. F., Schmidt, M., Graalfs, H., Hafner, M., and Frech, C., (2015). Modeling of dual gradient elution in ion exchange and mixed-mode chromatography. *Journal of Chromatography A 1417*, 64–72.

(*53*) Righetti, P. G., Fazio, M., Tonani, C., Gianazza, E., and Celentano, F. C., (1988). pH gradients generated by polyprotic buffers. II. Experimental validation. *Journal of Biochemical and Biophysical Methods 16*, 129–140.

(*54*) Giaffreda, E., Tonani, C., and Righetti, P. G., (1993). pH gradient simulator for electrophoretic techniques in a windows environment. *Journal of Chromatography 630*, 313–327.

(*55*) Kröner, F., and Hubbuch, J., (2013). Systematic generation of buffer systems for pH gradient ion exchange chromatography and their application. *Journal of Chromatography A 1285*, 78–87.

(*56*) Lingg, N., Tan, E., Hintersteiner, B., Bardor, M., and Jungbauer, A., (2013). Highly linear pH gradients for analyzing monoclonal antibody charge heterogeneity in the alkaline range. *Journal of Chromatography A 1319*, 65–71.

(*57*) Kaltenbrunner, O., Tauer, C., Brunner, J., and Jungbauer, A., (1993). Isoprotein analysis by ion-exchange chromatography using a linear pH gradient combined with a salt gradient. *Journal of Chromatography A 639*, 41–49.

(*58*) Zhang, L., Patapoff, T., Farnan, D., and Zhang, B., (2013). Improving pH gradient cation-exchange chromatography of monoclonal antibodies by controlling ionic strength. *Journal of Chromatography A 1272*, 56–64.

(*59*) Guélat, B., Khalaf, R., Lattuada, M., Costioli, M., and Morbidelli, M., (2016). Protein adsorption on ion exchange resins and monoclonal antibody charge variant modulation. *Journal of Chromatography A 1447*, 82–91.

(*60*) Kluters, S., Wittkopp, F., Matthias, J., Jöhnck, M., and Frech, C., (2016). Application of linear pH gradients for the modeling of ion exchange chromatography: Separation of monoclonal antibody monomer from aggregates. *Journal of Separation Science 39*, 663–675.

(*61*) Vetter, T., Ferreira, G., Robbins, D., and Carta, G., (2015). Resolution of protein charge variants in mixed-bed chromatography columns with step-induced pH gradients at high protein loadings. *Separation Science and Technology 50*, 2211–2219.

(*62*) Yamamoto, S., Suehisa, T., and Sano, Y., (1993). Preparative separation of proteins by gradient- and stepwise-eltuion chromatography: zone-sharpening effect. *Chemical Engineering Communications 119*, 221–230.

(*63*) Yamamoto, S., (1995). Plate height determination for gradient elution chromatography of proteins. *Biotechnology and Bioengineering 48*, 444–51.

(*64*) Tao, Y., Carta, G., Ferreira, G., and Robbins, D., (2011). Adsorption of deamidated antibody variants on macroporous and dextran-grafted cation exchangers: I. Adsorption equilibrium. *Journal of Chromatography A 1218*, 1519–29.

(*65*) Borg, N., Brodsky, Y., Moscariello, J., Vunnum, S., Vedantham, G., Westerberg, K., and Nilsson, B., (2014). Modeling and robust pooling design of a preparative cation-exchange chromatography step for purification of monoclonal antibody monomer from aggregates. *Journal of Chromatography A 1359*, 170–81.

(*66*) Reck, J. M., Pabst, T. M., Hunter, A. K., Wang, X., and Carta, G., (2015). Adsorption equilibrium and kinetics of monomer-dimer monoclonal antibody mixtures on a cation exchange resin. *Journal of Chromatography A 1402*, 46–59.

(*67*) Susanto, A., Knieps-Grünhagen, E., von Lieres, E., and Hubbuch, J., (2008). High throughput screening for the design and optimization of chromatographic processes: assessment of model parameter determination from high throughput compatible data. *Chemical Engineering & Technology 31*, 1846–1855.

(*68*) Chhatre, S., and Titchener-Hooker, N. J., (2009). Review: Microscale methods for high-throughput chromatography development in the pharmaceutical industry. *Journal of Chemical Technology & Biotechnology 84*, 927–940.

(*69*) Lacki, K. M., (2012). High-throughput process development of chromatography steps: advantages and limitations of different formats used. *Biotechnology Journal 7*, 1192–202.

(*70*) Barker, G., and Domagalski, N., (2016). A systematic approach to improve data quality in high-throughput batch adsorption experiments. *Engineering in Life Sciences 16*, 124–132.

(*71*) Melander, W. R., and Horvath, C., (1977). Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: An interpretation of the lyotropic series. *Archives of Biochemistry and Biophysics 183*, 200–215.

(*72*) GE Healthcare, Instructions 28-9955-33AB, Hydrophobic interaction media Capto Phenyl (High Sub)., 2011.

(*73*) Kato, Y., Nakamura, K., Kitamura, T., Moriyama, H., Hasegawa, M., and Sasaki, H., (2002). Separation of proteins by hydrophobic interaction chromatography at low salt concentration. *Journal of Chromatography A 971*, 143–149.

(*74*) Ghose, S., Tao, Y., Conley, L., and Cecchini, D., (2013). Purification of monoclonal antibodies by hydrophobic interaction chromatography under no-salt conditions. *mAbs 5*, 795–800.

(*75*) Jennisen, H. P., In *M.A. Vijayalakkshmi, Biochromatography: Theory and Practice*; Taylor & Francis: New York, NY, 2002, pp 55–84.

(*76*) Baca, M., Vos, J. D., Bruylants, G., Bartik, K., Liu, X., Cook, K., and Eeltink, S., (2016). A comprehensive study to protein retention in hydrophobic interaction chromatography. *Journal of Chromatography B 1032*, 182–188.

2 Salt and pH Gradient Elution on Ion Exchange

2.1 Introduction

A methodology is presented to predict protein elution behavior from an ion exchange column using both individual or combined pH and salt gradients based on high-throughput batch isotherm data. The buffer compositions are first optimized to generate linear pH gradients with defined concentrations of sodium chloride. Next, batch isotherm data are collected for lysozyme on SP-Sepharose-FF and two monoclonal antibodies on POROS XS over a range of protein concentrations, salt concentrations, and solution pH. Finally, an empirical interpolation (EI) method is used to describe protein binding as a function of the protein and salt concentration and solution pH without using an explicit isotherm model. The interpolated isotherm data are then used with a lumped kinetic model to predict the protein elution behavior. Experimental results obtained from laboratory scale columns show excellent agreement with the predicted elution curves for both individual or combined pH and salt gradients at protein loads up to 50% of the column binding capacity. Numerical studies show that the model predictions are robust as long as the isotherm data cover the range of mobile phase compositions where the protein actually elutes from the column.

2.2 Theoretical development

In this work, we first restrict the development to the case of ion-exchange chromatography using both individual and combined salt and pH gradients. However, the approach can be extended in a straightforward manner to systems involving other modes of interaction, such as hydrophobic interaction chromatography or multimodal chromatography. We consider systems where protein binding depends on the protein and salt concentration and solution pH. In its current form, the EI method is well suited for chromatographic steps where the product is bound and eluted while the impurities that are removed either flow through or are present in trace amounts only so that they do not affect the loading and elution behavior of the product in a significant way. Extensions of this method to cases where multicomponent interactions need to be taken into account will be

considered in Section 3.

The overall model has four components to predict: (1) pH gradient as a function of initial and final buffer composition; (2) protein binding equilibrium as a function of protein and salt concentration, and pH; (3) protein adsorption kinetics; and (4) column dynamics. The four model components are discussed separately below.

2.2.1 pH gradient prediction

In the ensuing development, we assume that only acid buffer components are present, which is typical for CEX resins. In general, the deprotonation of each component i is described by:

$$\left(\mathrm{H}_{N_i+1-n}\mathrm{A}_i\right)^{1-n} \Leftrightarrow \mathrm{H}^+ + \left(\mathrm{H}_{N_i-n}\mathrm{A}_i\right)^{-n}$$

where N_i is its valance and $n = 1, 2...N_i$. The apparent equilibrium constant for each deprotonation reaction, $K'_{a_i,n}$ is related to the corresponding thermodynamic value, $K_{a_i,n}$, through the following relationship:

$$K'_{a_i,n} = \frac{\left[\mathrm{H}^+\right]\left[\left(\mathrm{H}_{N_i-n}\mathrm{A}_i\right)^{-n}\right]}{\left[\left(\mathrm{H}_{N_i+1-n}\mathrm{A}_i\right)^{1-n}\right]} = K_{a_i,n} \times \frac{\gamma_{\left(\mathrm{H}_{N_i+1-n}\mathrm{A}_i\right)^{1-n}}}{\gamma_{\mathrm{H}^+}\gamma_{\left(\mathrm{H}_{N_i-n}\mathrm{A}_i\right)^{-n}}} \tag{2.1}$$

where brackets denote molarities and γ thermodynamic activity coefficients. According to the Davies equation[1], for an ion charge $\pm j$, the activity coefficient is given by:

$$\log \gamma_{\pm j} = -j^2 f(I) \tag{2.2}$$

where $f(i) = \left(\frac{a\sqrt{I}}{1+\sqrt{I}} - 0.1I\right)$, a is a temperature dependent constant equal to 0.51 at 298 K and I is the ionic strength of the solution. Combining eqs. 2.1 and 2.2 yields $K'_{a_i,n} = K_{a_i,n} \times 10^{2nf(I)}$. Considering now a mixture of N_c acidic buffer components each with valence N_i, using the apparent

deprotonation constants, the concentration of each buffer species is given by:

$$\left[\left(\mathrm{H}_{N_i+1-n}\mathrm{A}_i\right)^{1-n}\right] = \frac{C_i \prod_{m=1}^{n-1} \frac{K'_{a_i,m}}{\left[\mathrm{H}^+\right]}}{1+\sum_{n=1}^{N_i}\left(\prod_{m=1}^{n} \frac{K'_{a_i,m}}{\left[\mathrm{H}^+\right]}\right)}, n = 1, N_i+1 \tag{2.3}$$

where C_i is the total concentration of each buffer component. Combining eq. 2.3 with the following electroneutrality condition:

$$\left[\mathrm{Na}^+\right] + \left[\mathrm{H}^+\right] + \sum_{i=1}^{N_C}\left[\sum_{n=2}^{N_i+1}(1-n)\left[\left(\mathrm{H}_{N_i+1-n}\mathrm{A}_i\right)^{1-n}\right]\right] - \left[\mathrm{Cl}^-\right] - \frac{K'_w}{\left[\mathrm{H}^+\right]} = 0 \tag{2.4}$$

where K'_w is the apparent ionic product of water, provides a function of $\left[\mathrm{H}^+\right]$, which, in turn, given the desired NaCl concentration, can be used to calculate pH. To aid interpretation of these equations for specific systems, explicit forms of eqs. 2.1-2.4 are given in Section 6.2 of the Appendix for the special case of a buffer mixture containing acetate and phosphate. It should be noted that eqs. 2.1-2.4 have been written only for the mobile phase. The double layer at the charged resin surface is not considered explicitly assuming that the strong functional groups on the resin are fully deprotonated at the pH values of interest in this work. Thus, the Na^+ concentration at the resin surface is equal to the concentration of negatively charged ligands.

For the selection of buffers for pH gradient elution, the concentrations of the buffer components, C_i , are the unknowns that are optimized to define initial and final buffer compositions that will yield a pH gradient with a desired degree of linearity. An iterative computation scheme is used for this purpose. After selecting the number and type of buffer components in initial and final buffers, the ionic strength of a mixture of the two buffers mixed in a certain proportion is guessed and used to calculate the apparent deprotonation constants, the speciation of each buffer component, and the pH according to eq. 2.4. This calculation is iterated until convergence is attained for the ionic strength. The entire process is then repeated for different proportions of initial and final buffers and

the calculated pH-values are used to calculate the sum of squared residuals (SSR) according to:

$$\mathrm{SSR} = \sum_{k} \left(\mathrm{pH}_{calculated,k} - \mathrm{pH}_{linear,k} \right)^2 \tag{2.5}$$

where the summation is extended to the number of desired proportions of the two buffers and $\mathrm{pH}_{linear,k}$ is the pH for a simple linear combination of initial and final buffer pH values. The SSR is the objective function that is minimized to find optimized buffer compositions. For the calculations shown in this work, we optimized initial and final buffers comprised of mixtures of acetate, 2-(N-morpholino)ethanesulfonate (MES), and phosphate by minimizing SSR for mixtures in 5% increments from 0 to 100% using MATLAB's fmincon function. These buffer components are selected to give approximately 1 pH unit separation between pK_a values which helps maintain an approximately constant buffering capacity.

2.2.2 Protein binding equilibrium

Protein adsorption equilibrium is defined in this work based on the total protein concentration held within an adsorbent particle, $\bar{q}_p$. This quantity can be written as the sum of two terms, one accounting for the protein bound and the other for the protein simply held within the particle pores. Accordingly, $\bar{q}_p$ is given by:

$$\bar{q}_p = q_p + \varepsilon_p K_D C_p \tag{2.6}$$

where q_p is the bound protein concentration in mg per ml of particle volume, C_p is the protein concentration in solution, ε_p the intraparticle porosity, and K_D a distribution coefficient describing partitioning of non-bound protein molecules between the solution and the liquid-filled pores. For small molecules, including salt, K_D ~1, while for a large protein $K_D < 1$ as a result of size exclusion. In many cases, the second term in eq. 2.6 is of minor importance at low protein concentrations but becomes significant, for example, during elution at high protein loads for conditions where high protein concentrations are attained.

Two different methods were used to describe the bound protein concentration at equilibrium –

the SMA isotherm model, which describes the effects of both protein and counterion concentration based on a mechanistic description, and an empirical interpolation (EI) method that does not require a mechanism-based description of equilibrium. The SMA isotherm model introduced by Brooks and Cramer[2] assumes that protein binding occurs as a result of the stoichiometric exchange of protein ions and a monovalent counterion (e.g. Na^+) and is given by the following equation relating the bound protein concentration, q_p, to that in solution, C_p:

$$q_p = \frac{K_e\left[q_0 - (z+\sigma)q_p\right]^z}{\left[\mathrm{Na}^+\right]^z} C_p \tag{2.7}$$

In this equation, K_e is the equilibrium constant for the exchange of the protein with the counterion (Na^+ in our case), q_0 is the resin charge density, z is the protein effective charge, σ is the steric hindrance factor, and is the Na^+ concentration in solution. The linear limit of the isotherm is obtained at high salt concentrations and/or low protein concentrations when $q_p << q_0/(z+\sigma)$ and is given by the following equation, which is analogous to the so-called stoichiometric exchange model[3]:

$$q = \frac{K_e\left(q_0\right)^z}{\left[\mathrm{Na}^+\right]^z} C_p \tag{2.8}$$

The four model parameters, q_0, K_e, z, and σ, can be determined in different ways. A common approach is to determine q_0 from a potentiometric titration of the resin or by the frontal exchange of counterions and to determine K_e and z from linear gradient elution (LGE) at low protein loads for conditions where the protein elutes in the linear range of the isotherm. For these conditions, the Na^+ concentration at which the protein elutes, $\left[\mathrm{Na}^+\right]_{elution}$ is related to the normalized gradient slope by the following equation[4–7]:

$$\gamma = \frac{1}{\phi} \int_{\left[\mathrm{Na}^+\right]_{initial}}^{\left[\mathrm{Na}^+\right]_{elution}} \frac{d\left[\mathrm{Na}^+\right]}{A\left[\mathrm{Na}^+\right]^{-z} + \varepsilon_p\left(K_D - 1\right)} \tag{2.9}$$

where $A = K_e\left(q_0\right)^z$. The normalized gradient slope is given by $\gamma = \varepsilon \Delta C_{Na^+} / C V_G$, where ε is the

column void fraction, ΔC_{Na^+} the difference between final and initial Na^+ concentrations, CV_G the duration of the gradient in column volume units, and $\phi = (1-\varepsilon)/\varepsilon$ is the phase ratio. An analytical integral is obtained when the second term in the denominator is zero. In other cases, numerical integration is needed. In either case, however, both z and A (and, thus, K_e) are pH-dependent and can be obtained by regression of γ vs. $\left[Na^+\right]_{elution}$ data obtained from LGE experiments conducted over a range of gradient slopes at different pH values. Finally, σ can be determined from a single isotherm measurement at high loads. Alternatively, all model parameters can be regressed simultaneously to a set of batch isotherm data obtained over a range of protein and Na^+ concentrations at constant pH. Equations 2.8 and 2.9 do not explicitly account for pH effects and thus, in their current forms, are strictly limited to modeling the effect of protein and salt concentration.

The proposed EI approach does not rely on any particular equilibrium model and, for this development, will be extended to also include the effect of pH. The EI method uses a convenient arbitrary function to describe the relationship between q_p and C_p at each Na^+ concentration and pH. Examples of such functions include the SMA, Langmuir, and Langmuir-Freundlich isotherm models. In this study, we use the Langmuir-Freundlich or Sips isotherm model, which is given by[8]:

$$q_p = \frac{q_m (KC_p)^b}{1 + (KC_p)^b} \tag{2.10}$$

It should be noted that this model is used only as a convenient algebraic expression without implying that this model has any mechanistic significance in this context. This relationship was found to have sufficient flexibility to accurately fit protein batch isotherm data at each salt concentration and pH value. The model parameters q_m, K, and b are treated as empirical constants obtained by independent regression at each salt concentration and pH value.

The EI approach described in this work is obviously data-driven and assumes that a large data set of batch adsorption isotherms is available. The goal is, thus, to provide a tool to translate the batch adsorption data into the behavior of a chromatography column. Since the method does not use a mechanistic model to describe the salt concentration dependence, extrapolation beyond the experimental range is expected to be unreliable. On the other hand, the advantage is that highly

accurate predictions are possible within the experimental range without having to depend on currently available simplified mechanistic models. Since, as noted above, q_m, K, and b are fitted at each salt concentration and pH value, 3 x m parameters are used in the ensuing examples to fit each data set, where m is the number of different $\left[Na^+\right]$ and pH sets. Next, eq. 2.10 is used to calculate 50 logarithmically-spaced q_p values for each constant $\left[Na^+\right]$ and pH set. For each of these sets, the more accurate linear isotherm data from LGE experiments replace the batch isotherm data at the Na^+ concentration where the two data sets agree, typically near non-binding conditions. Finally, a piecewise cubic Hermite interpolating polynomial (PCHIP) is used to calculate q_p at intermediate $\left[Na^+\right]$ and pH values using MATLAB's 1D interpolating *pchip* function. *pchip* interpolates 50 evenly spaced values of q_p with respect to $\left[Na^+\right]$ at constant C_p and pH. Next, *pchip* interpolates 50 evenly spaced values of q_p with respect to pH at each constant C_p, and $\left[Na^+\right]$. The resulting grid of 50^3 interpolated isotherm points is then given as an input to MATLAB's *griddedInterpolant* function which uses trilinear interpolation to return any q_p value for a given input of C_p, $\left[Na^+\right]$, and pH value. If the queried isotherm point is outside the range of known $\left[Na^+\right]$ or pH, trilinear extrapolation is used. For a queried point within the known $\left[Na^+\right]$ and pH range but outside the known C_p range, the value of q_p is extrapolated using the Langmuir-Freundlich fits. Additional details regarding the specific steps taken to interpolate the experimental data are provided in Section 6.3 of the Appendix.

It should be noted that including protein adsorption data in the linear range of the isotherm is essential in order to obtain a robust prediction of the elution behavior over broad ranges of protein loads, pH, and salt concentrations. Depending upon the values of these parameters, elution can in fact occur at least in part for conditions where the isotherm is linear. We found that, for the moderate or low protein binding associated with the linear isotherm behavior, isotherm data obtained chromatographically tend to be more accurate than those obtained from batch adsorption measurements.

2.2.3 Protein adsorption kinetics

The protein adsorption kinetics is assumed to be controlled by intraparticle pore diffusion. This assumption is valid when the Biot number, $Bi = k_f r_p / D_e = Sh D_0 / 2D_e$, is larger than 10, where k_f is the external mass transfer coefficient, D_e is the protein effective pore diffusivity, r_p is the particle radius, Sh is the Sherwood number, and D_0 is the diffusivity in solution. D_e / D_0 is always much less than 1 and Sh is typically greater than 10, thus values of Bi well in excess of 10 are expected. Accordingly, the kinetics is described by the so-called linear driving force (LDF) approximation[9]:

$$\frac{\partial \bar{q}_p}{\partial t} = \frac{15 D_e}{r_p^2}\left(C_p - C_p^*\right) \tag{2.11}$$

where C_p^* is the protein concentration in equilibrium with $\bar{q}_p$. Experimentally, D_e can be obtained in several ways, two of which are explored in this work: 1) from van Deemter curves plotting the HETP obtained from pulse injections of the protein for non-binding conditions vs. the superficial mobile phase velocity, u, and 2) from batch uptake curves for conditions were the binding isotherm is highly favorable. In the first method, assuming that band-broadening is controlled by intraparticle diffusion, the van Deemter plot (HETP vs. u) is expected to be linear and the slope, c, is given by:

$$c = \frac{d\mathrm{HETP}}{du} = \frac{2}{1-\varepsilon}\left(\frac{k'}{1+k'}\right)^2 \frac{r_p^2}{15 D_e} \tag{2.12}$$

where $k' = \overline{CV} / \varepsilon - 1$ is the protein retention factor and $\overline{CV}$ is the mean retention volume in column volumes (CV units)[10]. The value of D_e can be regressed from the slope of the HETP curve.

In the second method, since external mass transfer and the solute hold-up in the pore volume are usually negligible, the batch uptake curve can be described by[11]:

$$\frac{C_0}{q_m}\frac{D_e t}{r_p^2} = f\left(\eta\right) \tag{2.13}$$

where C_0 is the initial protein concentration, q_m is the binding capacity, $\eta = \left(1 - q/q_m\right)^{1/3}$, and

$$f(\eta) = \frac{1}{3\Lambda}\ln\left(\frac{\lambda^3+\eta^3}{\lambda^3+1}\right) - \frac{1}{6\lambda\Lambda}\ln\left[\frac{\lambda^3+\eta^3}{\lambda^3+1}\left(\frac{\lambda+1}{\lambda+\eta}\right)^3\right] - \frac{1}{\lambda\Lambda\sqrt{3}}\left[\tan^{-1}\left(\frac{2\eta-\lambda}{\lambda\sqrt{3}}\right) - \tan^{-1}\left(\frac{2-\lambda}{\lambda\sqrt{3}}\right)\right] \quad (2.14)$$

where $\Lambda = V_r q_m / V C_0$ and $\lambda\left(1/\Lambda - 1\right)^{1/3}$. V_r and V are the volume of resin and solution, respectively. The validity of eq. 2.13 and its underlying assumptions can be tested by plotting experimental values vs. time and assessing the ensuing linearity. D_e can then be obtained from the slope of this plot if the binding capacity, q_m, is independently known.

2.2.4 Column dynamics

The following plug-flow model is used predict the column dynamics for protein and buffer components:

$$\varepsilon\frac{\partial C_j}{\partial t} + (1-\varepsilon)\frac{\partial \bar{q}_j}{\partial t} + u\frac{\partial C_j}{\partial x} = 0 \quad (2.15)$$

where x is the column axial coordinate and u is the superficial velocity. For the buffer components, $\bar{q}_j = \varepsilon_p C_j$ and local equilibrium is assumed, while for the protein, $\frac{\partial \bar{q}_j}{\partial t}$ is given by eq. 2.11. A numerical solution of eq. 2.15 is obtained by discretizing the axial derivative by backwards finite differences and solving the resulting set of ordinary differential equations with *ode15s* in the MATLAB library. Numerical dispersion caused by the discretization was minimized by using >80 discretization points, which caused only minimal dispersion of $\left[Na^+\right]$ and pH profiles and with no effect on spreading of the protein concentration profiles. All calculations were done in MATLAB R2013b (The Mathworks, Natick, MA, USA) on a Dell Precison T1700, Intel i7 series 3.40GHz.

2.3 Materials and methods

2.3.1 Materials

Sodium chloride, sodium acetate, sodium 2-(N-morpholino)ethanesulfonate (MES), dibasic sodium phosphate, phosphoric acid, and acetic acid used in buffer preparation were purchased from Fisher

Scientific (Fair Lawn NJ, USA) and Sigma-Aldrich (St. Louis MO, USA).

The two cation exchange resins used in the experimental study are SP-Sepharose FF (GE Healthcare, Piscataway, NJ, USA) and POROS XS (Life Technologies Corporation, Grand Island, NY, USA). The first of these two resins is agarose based while the second is based on poly(styrene-divinylbenzene). Both resins are strong cation exchangers with properties summarized in Table A1 of the Appendix.

Lysozyme (Mr ~14,300, pI ~11) was obtained from Sigma-Aldrich (St. Louis, MO, USA) and two purified monoclonal antibodies, mAb A and mAb B, were provided by Bristol-Myers Squibb (Hopewell, NJ, USA). Both mAbs have a molecular weight near 150,000 Da with slightly different pI values near 9.

The lysozyme experiments were conducted with buffers containing 10 mM Na_2HPO_4 adjusted to pH 6.5 with phosphoric acid. The mAb A experiments were conducted in buffers containing 20 mM total acetate at pH 5.5 prepared by mixing sodium acetate with acetic acid. In the two cases above, NaCl was added to these buffers to adjust the total Na^+ concentration for batch isotherm and column experiments. The mAb B experiments were conducted in buffers containing amounts of sodium acetate, sodium MES, sodium phosphate dibasic, and sodium chloride calculated using eqs. 2.1-2.5 with the pK_a values given in Beynon and Easterby[12]. The compositions of the initial and final buffers used for the pH gradient elution runs from pH 5.5 to 7 are in Table A2 of the Appendix. All experiments were performed at room temperature (22 $\pm$ 2 °C).

2.3.2 Methods

2.3.2.1 Batch adsorption isotherms

Two methods were used to collect batch adsorption isotherm data; 1) manual data collection conducted here at UVa and 2) automated data collection conducted during an internship at Bristol-Myers Squibb in Bloomsbury, NJ using a Freedom EVO® liquid handler controlled by Freedom EVOware® (Tecan US, Research Triangle Park, NC, USA). The first method was used for lysozme and mAb A to collect adsorption isotherms by equilibrating the resin in the appropriate buffer and

then removing the extraparticle liquid using 2 mL Corning Costar Spin-X microfiltration tubes (Sigma-Aldrich, St. Louis, MO, USA) with an Eppendorf Minispin bench-top centrifuge (Eppendorf North America, Hauppauge, NY, USA) operated at 5000 RPM for 15 minutes. Samples of the filtered resin (30-300 mg) were then added to either 2 or 5 mL plastic tubes and mixed with 1 to 3 ml of protein solution by slowly rotating the tubes end-over-end on a rotator for 24 h. After this time, supernatant samples were taken to determine the residual protein concentration using a spectrophotometer (Model DU640, Beckman-Coulter, Fullerton, CA, USA) at 280 nm.

mAb B adsorption isotherm data were collected using the HTS method which uses AcroPrep Advance filter plates with 0.45 μm GHP membranes and corresponding cap mats (Pall Corporation, Port Washington, NY, USA). Different resin volumes were added to each well by dispensing a slurry with known resin content and centrifuged to remove the extraparticle liquid with Hettich 460 Robotic centrifuge (Hettich America, L.P., Buford, GA, USA) at 1652 rpm for 10 minutes. 400 μL samples of protein solution were then added to the wells, which were then sealed and slowly rotated end-over-end on a wheel for 24 hours. The bulk of the supernatant was removed by centrifugation and the residual protein concentration was determined from the UV absorbance with a DropSense 96® plate reader (Trinean, Gentbrugge, Belgium).

In both methods above, the amount of protein held by the resin was then calculated by mass balance. Finally, the total protein concentration in the resin, $\bar{q}_p$, was calculated by dividing this amount by the resin sample volume. The latter was calculated from the sample mass using the density of the filtered resin determined with a pycnometer. The amount of resin added to each tube was estimated to ensure that a minimum 30% change in the protein supernatant concentration occurred between the initial and equilibrium value of the protein solution concentration in order to minimize the effects of measurement errors on the calculated adsorbed concentration.

2.3.2.2 Adsorption kinetics

HETP curves were generated for lysozyme on SP-Sepharose FF and mAb A on POROS XS. Isocratic elution experiments under non-binding conditions (1000 mM NaCl) were performed to deter-

mine the D_e values based on eq. 2.12 using 1 cm diameter x 10 cm long Tricorn columns obtained from GE Healthcare (Piscataway, NJ, USA) with an AKTA Explorer 10 system, also from GE Healthcare, at flow rates in the range 0.5 to 5 ml/min, with 50 µL injections of 5 mg/ml protein solutions, and with detection at 280 nm.

For mAb B, a batch uptake curve using HTS methods described in Section 2.3.2.1 was used to determine the binding kinetics. For this purpose, after adding the same amount of resin to each well in the filter plate and removing the equilibration buffer, samples of the protein solution were sequentially dispensed to different wells over a two hour period with constant orbital shaking using a Te-Shake (Tecan US, Research Triangle Park, NC, USA) set at 1150 rpm and 3 mm orbital radius. Immediately following the last protein solution addition, the filter plate was centrifuged and the adsorbed protein concentration was calculated by mass balance using the measured residual protein supernatant concentration in each well.

2.3.2.3 Column experiments

Linear gradient elution (LGE) experiments were done with 0.5 cm diameter x 5 cm long Tricorn columns with gradients from 100 to 500 mM NaCl for lysozyme on SP-Sepharose FF and 0 to 500 mM NaCl for mAb A on POROS XS also using an ÄKTA Explorer 10 system from GE Healthcare (Piscataway, NJ, USA). Injection volumes of 50 µl of 2 mg/ml protein with 5, 10, 15, 20, 30, and 40 CV gradients were used with a flow rate of 0.2 ml/min with detection at 280 nm in a 10 mm UV flow cell to characterize the retention behavior in the linear limit of the isotherm. Injection volumes up to 12 ml, using a 50 ml superloop from GE Healthcare (Piscataway, NJ, USA) and detection at 300 nm, were used to generate overloaded LGE data for comparison with model predictions. These runs were conducted at a flow rate of 0.5 ml/min for lysozyme and at a flow rate of 0.2 ml/min for mAb A.

For mAb B, gradient elution experiments were done with a 0.5 cm diameter x 10 cm long Tricorn column using an ÄKTA Pure system from GE Healthcare (Piscataway, NJ, USA) at a flow rate of 0.5 ml/min. For low protein loading experiments, 100 µl samples containing 5 mg/ml protein were

injected and eluted with salt gradients from 0 to 500 mM NaCl with varying gradient slopes at constant pH. Much higher sample volumes, up to 20 mL, were injected with a superloop for high protein loading experiments from 5 to 45 mg/ml of column and eluted with either salt gradients, pH gradients, or combined salt/pH gradients. Detection of the elution profile for overloaded conditions was by UV at 300 nm in a 2 mm flow cell.

2.4 Results and discussion

2.4.1 Protein binding equilibrium

2.4.1.1 Low load LGE behavior

Figure 2.1A shows the normalized gradient slope γ vs. the Na^+ concentration at elution at pH 5.5 for the LGE data at low protein loads for lysozyme on SP-Sepharose FF and for mAb A on POROS XS (see LGE chromatograms in Fig. A1 of Appendix). These data show that lysozyme elutes at higher Na^+ concentrations on SP-Sepharose FF compared to mAb A on POROS XS, which is likely due to the greater charge density of SP-Sepharose FF compared to POROS XS. The regressed values of z and A obtained by fitting these data according to eq. 2.9 are $z = 5.2$ and $A = 3.5 \times 10^{13}$ $(mM)^{5.2}$ for lysozyme on SP-Sepharose FF and $z = 10.1$ and $A = 2.9 \times 10^{23}$ $(mM)^{10.1}$ for mAb A on POROS XS. The larger effective charge for mAb A compared to lysozyme is consistent with its larger molecular size and expected higher net charge at the operating pH. Similar results were reported previously for lysozyme on SP-Sepharose FF[9] and for a different mAb on POROS 50HS, which has a structure similar to that of POROS XS[13]. A plot of the normalized gradient slope γ vs. the Na^+ concentration at elution for mAb B on POROS XS as pH 5.5, 6, 6.5 and 7 is shown in Fig. 2.1B (see LGE chromatograms in Fig. A2 of Appendix). Each constant pH data set shows that mAb B elutes at a higher Na^+ concentration with increasing gradient slope, consistent with the trend seen for mAb A on POROS XS in Fig. 2.1A. The A and z parameters regressed to the mAb B data at different pH values are given as an inset in Fig. 2.1B. As expected, the effective binding charge, z, decreases as the pH increases consistent with the lower net positive charge of a mAb at the higher pH.

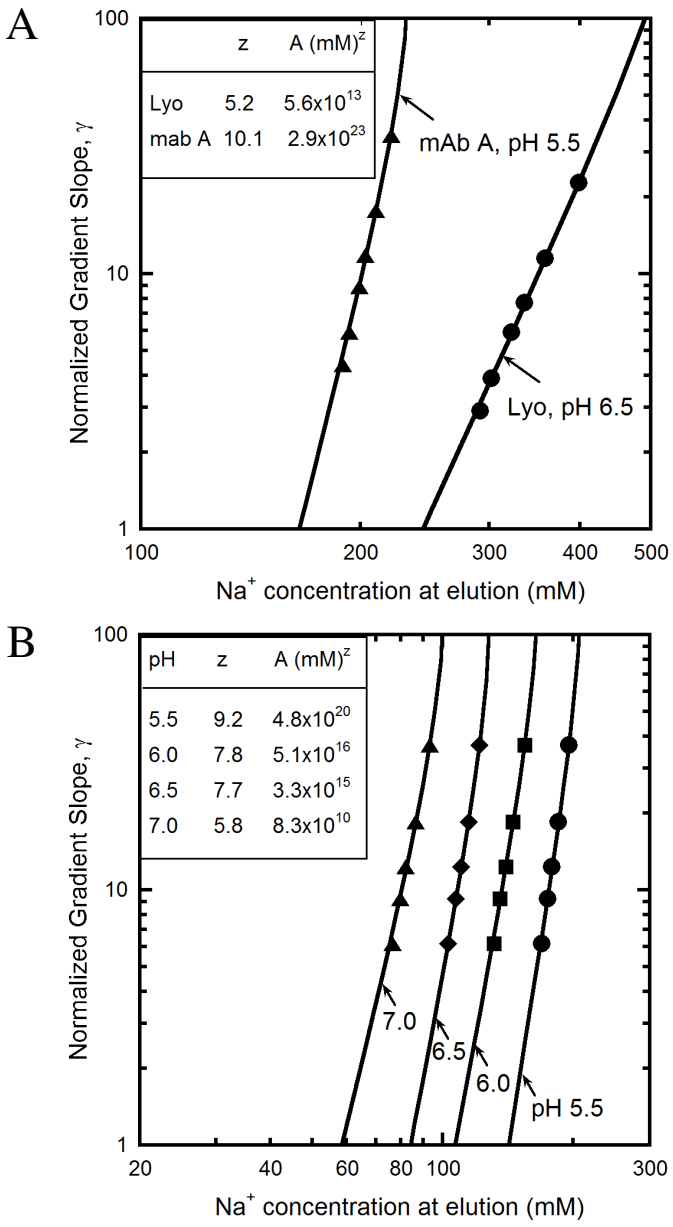


Figure 2.1: Normalized gradient slope γ vs. Na^+ concentration at elution for LGE experiments at constant pH and low protein loads for lysozyme (Lyo) on SP-Sepharose FF and mAb A on POROS XS (A) and mAb B on POROS XS (B). Lines are based on eq. 2.9 with the regressed parameters given as an inset in the figures. Experimental conditions are described in 2.3.2.3.

2.4.1.2 High load batch isotherm behavior

Figure 2.2 shows the batch isotherm data for lysozyme on SP-Sephaose-FF with Na^+ concentrations in the range 50 to 420 mM (A) and (C) and for mAb A on POROS XS with Na^+ concentrations in the range 17 to 218 mM (B) and (D). The highest salt concentration collected was increased until adsorption values were sufficiently low, as was the case for lysozyme, or there is close agreement with the isotherm data determined from dilute linear gradient elutions, as was the case for mAb A. Although the highest salt concentration studied may be higher than the salt concentration at which the peak elutes, this last salt concentration affects the isotherm slope around the salt conditions protein does partially elute as a result of interpolation. The ranges of protein concentrations, 0-16 mg/ml for lysozyme on SP-Sepharose FF and 0-30 mg/ml for mAb A on POROS XS, cover those expected in LGE experiments at high protein loads. Lines calculated according to the SMA model (eq. 2.7) with parameters regressed simultaneously to the global set of data are shown in Figs. 2.2A and 2.2B. Attempts to use the z- and A-values determined from the LGE experiments at low protein loads and regress only q_0 and σ to the high load data resulted in unsatisfactory predictions of the trends of $\bar{q}_p$ with respect to Na^+ concentration (refer to Fig. A3 and Table A2 in the Appendix). Thus, all four SMA parameters were regressed simultaneously by minimizing the sum of squared residuals using MATLAB's optimization program *lsqnonlin*, which is based on the Levenberg-Marquardt method[14]. The regressed SMA parameter values are summarized in Table 2.1. As seen from this table, while the z-values are comparable to those obtained from the low-load LGE experiments, the A-values are very different. Moreover, as seen in Fig. 2.2A and 2.2B, while the SMA model obviously captures the qualitative trends with respect to protein and Na^+ concentration, the quantitative agreement is poor, with a mean absolute error of 43% for the lysozyme/SP-Sepharose FF data and 32% for mAb A/POROS XS data. These errors were calculated by comparing the experimental data points with the model fit at each protein and salt concentration and averaging the results. Lines calculated according to the EI scheme described in Section 2.2.2 are shown in Figs. 2.2C and 2.2D. In this case, each constant-Na^+ isotherm data set was first fitted independently with the Langmuir-Freundlich model (eq. 2.10) by non-linear regression, yielding the parameters in Table 2.2.

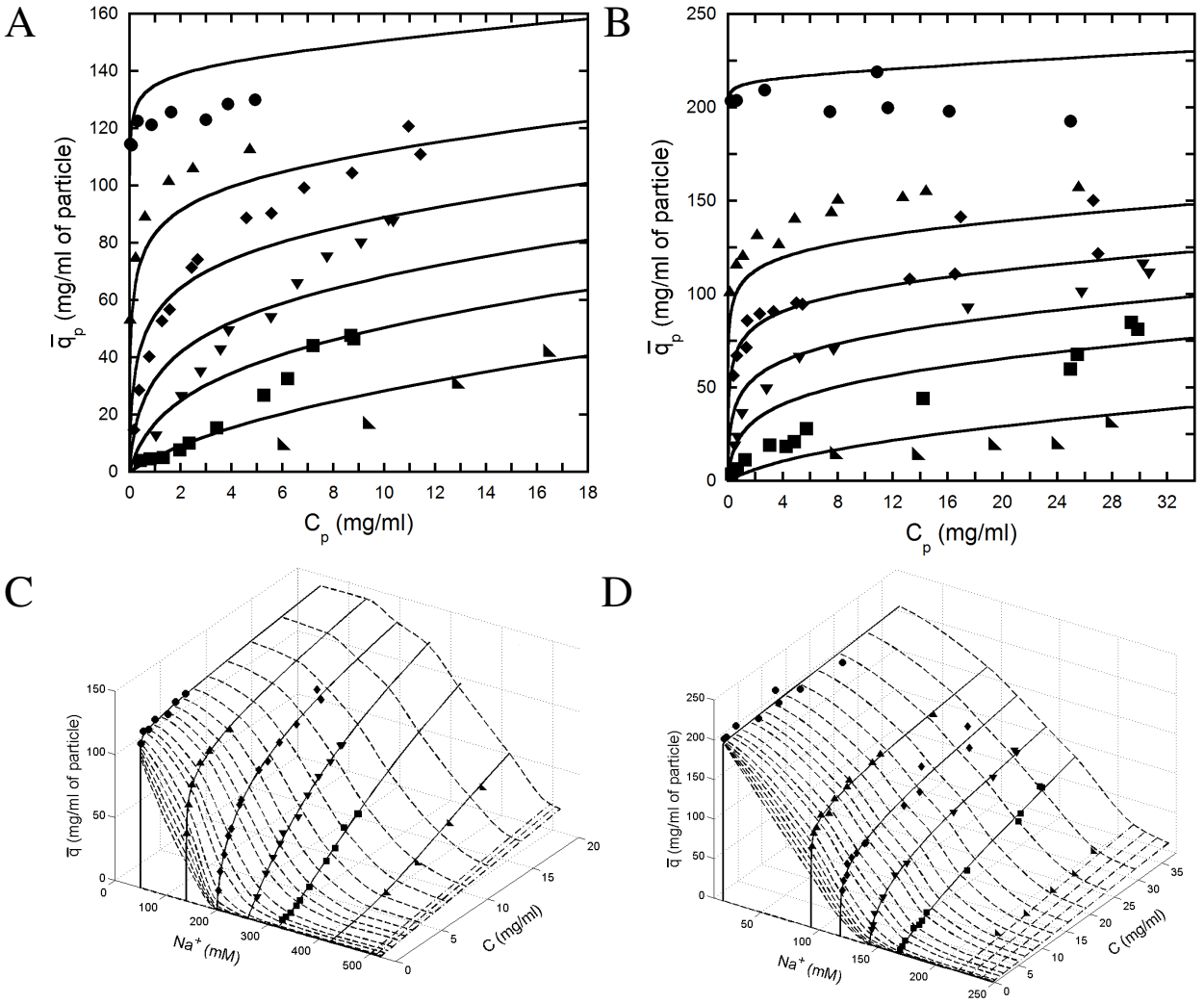


Figure 2.2: Batch adsorption isotherms for lysozyme on SP-Sepharose FF (A) and (C) and for mAb A on POROS XS (B) and (D). The solid lines in (A) and (B) are calculated with the SMA model using the global best-fit parameters given in Table 2.1. The solid lines in (C) and (D) are calculated with the Langmuir–Freundlich model with the parameters given in Table 2.2 fitted individually at each Na^+ concentration. In (A) and (C) symbols ●, ▲, ◆, ▼, ■ and ◣ represent 50, 140, 200, 260, 320, and 420 mM Na^+ concentrations, respectively. In (B) and (D) symbols ●, ▲, ◆, ▼, ■ and ◣ represent 17.6, 93, 118, 143, 168, and 218 mM Na^+ concentrations, respectively. The dash-dotted lines in (C) and (D) are the linear isotherms from the LGE experiments at low protein loads shown in Fig. 2.1A.

Table 2.1: SMA model parameters fitted to isotherm data in Figures 2.2A and 2.2B

	Lysozyme/SP-Sepharose FF	**mAb A/POROS XS**
z	6.65	12.4
A (mM)^z	3.58×10^{18}	7.66×10^{29}
q_0 (mM)	291	152
K_e	60.3	348
σ	18.3	83.1

Table 2.2: Langmuir-Freundlich parameters fitted to isotherm data in Figures 2.2C and 2.2D

Lysozyme/SP-Sepharose FF				**mAb A/POROS XS**			
[Na^+] (mM)	q_m (mg/ml)	K (ml/mg)	b	[Na^+] (mM)	q_m (mg/ml)	K (ml/mg)	b
50	158	$3.65\ 10^{3}$	0.151	17.6	200	$7.76\ 10^{9}$	1.00
140	277	$2.25\ 10^{-2}$	0.177	93.0	218	$3.49\ 10^{0}$	0.192
200	147	$3.35\ 10^{-1}$	0.721	118	414	$8.79\ 10^{-4}$	0.222
260	166	$9.02\ 10^{-2}$	1.03	143	151	$9.89\ 10^{-2}$	0.598
320	139	$6.30\ 10^{-2}$	1.42	168	499	$2.40\ 10^{-3}$	0.706
420	74.0	$5.30\ 10^{-2}$	23.2	218	181	$3.63\ 10^{-4}$	0.491

As seen in these figures, the agreement between fitted lines and experimental data is excellent with mean absolute error of less than 7% for both systems. The linear isotherm data from LGE experiments at low protein loads are represented as dash-dotted lines in Figs. 2.2C and 2.2D. For the case of lysozyme/SP-Sepharose FF in Fig. 2.2C, the linear isotherm data was added at 500 mM Na^+ near non-binding conditions since lysozyme shows non-linear isotherm behavior up to 420 mM Na^+ concentrations. In the case of mAb A/POROS XS data, the linear isotherm data was incorporated at 218 mM Na^+ since there is good agreement with the batch data at the same salt concentration up to a protein solution concentration of 28 mg/ml (Fig. 2.2D). Dashed lines show the curves generated by PCHIP as described in Section 2.2.2. These lines run perpendicular to the fitted Langmuir-Freundlich curves and linear isotherm lines and describe the effect of Na^+ concentration. Given a set of protein and Na^+ concentration values, MATLAB's function *griddedInterpolant* was used to output the corresponding interpolated values of $\bar{q}_p$. An inverse interpolating function can also be generated if the independent variables passed to the *griddedInterpolant* function are Na^+ concentration and $\bar{q}_p$ values. Thus the output would then be the corresponding protein concentration in solution, C_p^*, which is convenient for use with eq. 2.11.

Figure 2.3 shows HTS batch isotherm data for mAb B on POROS XS bound at Na^+ concentrations between 20 and 220 mM, pH between 5.5 and 7, and protein concentrations up to about 23 mg/ml. The solid lines are based on eq. 2.10 with parameters regressed to the data by minimizing the sum of squared residuals with MATLAB's function *lsqnonlin* and given in Table A4 in the Appendix. At each pH, the isotherms become linear at high Na^+ concentrations, consistent with the results of the low-loading LGE experiments. Because the scatter of the isotherm points for these weak binding conditions is substantial, the LGE results rather than the batch isotherms are used in this region. Dashed lines in each panel show the corresponding linear isotherm predictions. Thus, the effect of [Na^+] and pH on $\bar{q}_p$ is interpolated by combining the Langmuir-Freundlich fits and LGE data at constant protein solution concentration using PCHIP as discussed in Section 2.2.2. The grid of interpolated isotherm points is given as input to MATLAB's *griddedInterpolant* function which returns an interpolated $\bar{q}_p$ for a given C_p, [Na^+], and pH set. *griddedInterpolant* can also be used to generate the inverse interpolating function using [Na^+], pH and $\bar{q}_p$ values as input variables. The

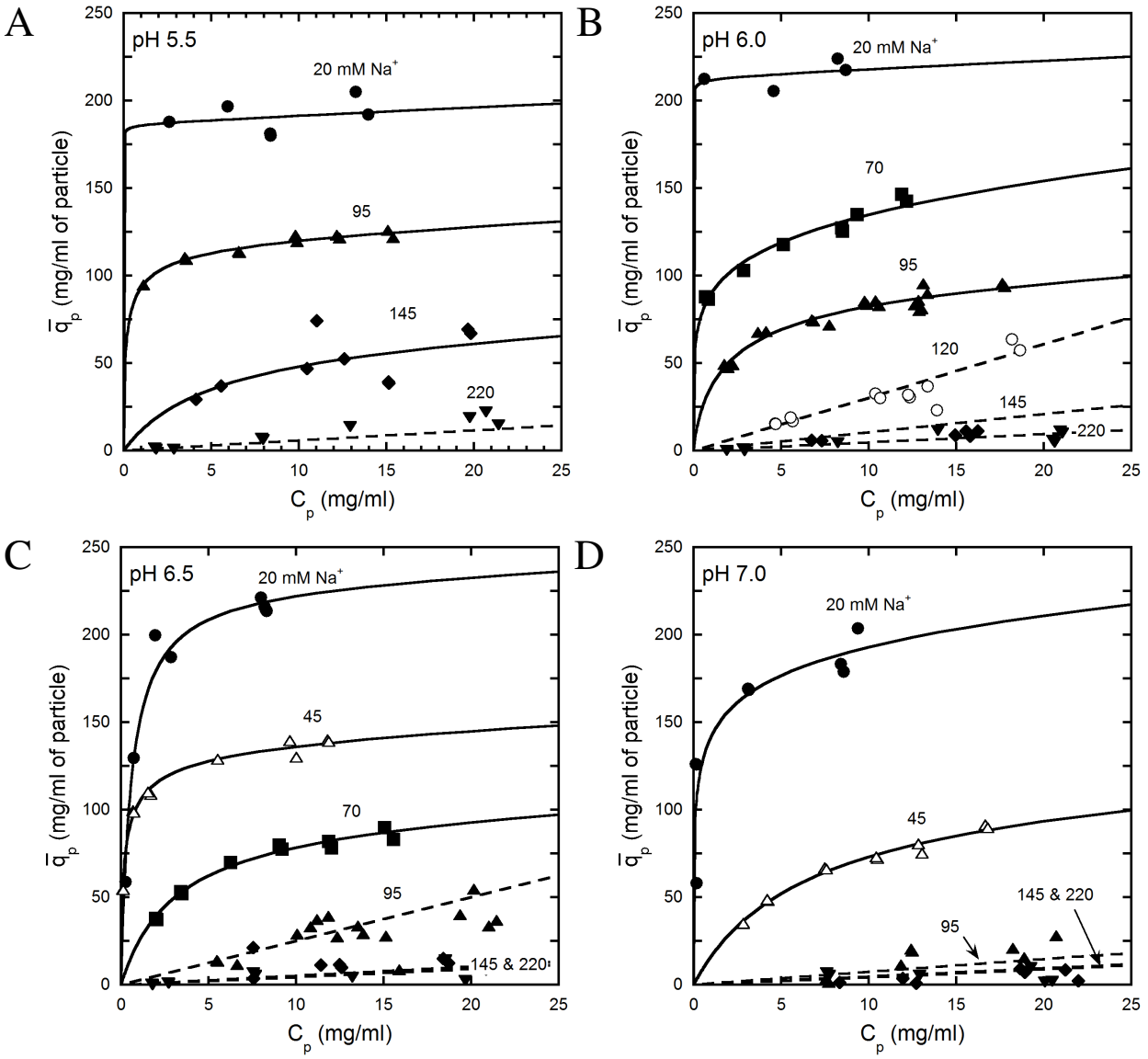


Figure 2.3: Batch adsorption isotherms of mAb B on POROS XS at constant pH values of 5.5 (A), 6.0 (B), 6.5 (C), and 7.0 (D). The solid lines are the Langmuir-Freundlich model with parameters given in Table A4 of the Appendix. Isotherms fitted individually at each Na^+ concentration and pH. The dashed lines show where the linear isotherms from LGE experiments are combined with the batch isotherm data.

inverse function returns the corresponding equilibrium protein liquid concentration C_p^* used in eq. 2.11.

2.4.2 Protein adsorption kinetics

The van Deemter curves obtained for lysozyme on SP-Sepharose FF and for mAb A on POROS XS for non-binding conditions (1000 mM Na^+) are shown in Fig. 2.4A. The data follow a linear trend of HETP vs. mobile phase velocity confirming that band broadening is mass transfer controlled. The effective diffusivities, D_e, calculated from the slope of the van Deemter curve and eq. 2.12, are $(3.09 \pm 0.18) \times 10^{-7}$ cm^2/s for lysozyme on SP-Sepharose FF and $(4.01 \pm 0.04) \times 10^{-8}$ cm^2/s for mAb A on POROS XS. Both values are substantially smaller than the corresponding free solution diffusivities of lysozyme (1.1×10^{-6} cm^2/s) and of mAb A (4.5×10^{-7} cm^2/s) suggesting that diffusion is significantly hindered in both stationary phases.

Figure 2.4B shows the batch adsorption data for mAb B on POROS XS at pH 5.5 with 20 mM Na^+ plotted in the form suggested by eq. 2.13. The plot is highly linear with an R-squared value of 0.998. The corresponding effective pore diffusivity for mAb B, obtained from the slope of this plot is $D_e = (3.6 \pm 0.1) \times 10^{-8}$ cm^2/s, about 10 times smaller than the free solution diffusivity of the mAb at room temperature ($\sim 4 \times 10^{-7}$ cm^2/s). For the buffer species, Na^+ and acetate or phosphate, the diffusional resistance is expected to be insignificant since the effective diffusivity is expected to be on the order of 10^{-5} cm^2/s.

2.4.3 Gradient elution at high protein loads

Figure 2.5 shows the experimental elution curves from salt gradients at constant pH obtained at various protein loads and gradient slopes for lysozyme on SP-Sepharose FF and for mAb A on POROS XS. The Na^+ concentration, calculated according to the effluent conductivity is also shown in each case. For both systems, the experimental curves (dashed lines) are compared with model predictions (solid lines) based either on the SMA model using the best-fit parameters summarized in Table 2.1 (Figs. 2.5A and 2.5B) or with the EI method (Figs. 2.5C, 2.5D, 2.5E, and 2.5F). In general,

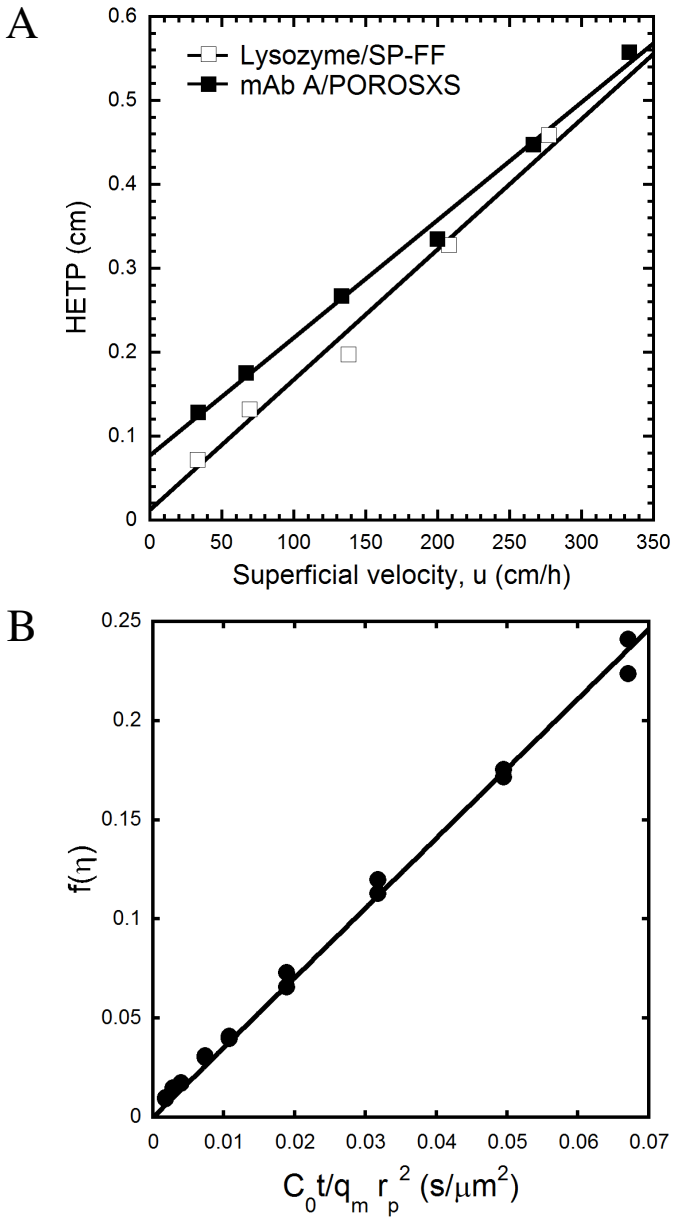


Figure 2.4: Two methods are shown to determine the effective diffusivity: HETP vs. superficial velocity obtained under non-binding conditions for lysozyme on SP-Sepharose FF and for mAb A on POROS XS (A) and HTS batch adsorption data plotted according to the analytical solution of the pore diffusion model with a rectangular isotherm for 5 mg/ml mAb B in 20 mM acetate pH 5.5 buffer on POROS XS (B). The effective pore diffusivities are regressed from the slopes of these curves according to equations given in Section 2.2.3.

there is good agreement between the experimental and calculated Na^+ concentrations (which are independent of the model used to describe protein binding) except at the highest protein load for mAb A on POROS XS (Figs. 2.5A and 2.5D). In this case, the experimental line dips below the theoretical one as the peak is eluted. We believe that this is primarily an artifact of the effect of the effluent viscosity, which, based on the data in ref.[15], reaches values substantial higher than the buffer viscosity at the approximately 25 mg/ml concentration at which the mAb is eluted. Since the Na^+ concentration is obtained directly from the electrical conductivity, which decreases as the viscosity increases, an increase in viscosity results, in turn, in an artificially low Na^+ concentration. As seen from Fig. 2.5A and 2.5B, although the SMA model captures the general trends with respect to the salt concentration, the predicted peak shapes are different from the experimental ones and show elution at salt concentrations higher than those observed experimentally. For both experimental systems, the SMA predicted profiles have a pronounced "shark fin" shape which depends on the concave shape of the SMA isotherm. It is worth noting that the SMA predictions also fail to describe the experimental behavior at low protein loads. This is the result of using global best fit parameters (Table 2.1) instead of constraining the SMA fit to isotherm data in Figs. 2.2A and 2.2B using the z- and A-values determined from the LGE experiments at low protein loads. The SMA predictions using constrained z- and A-values are shown in Fig. 2.6 and give much better agreement at low protein load, however they still significantly deviate from the data at high protein loads.

Figs. 2.5C and 2.5D show predictions based on the EI scheme and are in excellent agreement with the same experimental data. Not only are the general trends consistent with the experiments but all details of the experimental behavior are captured. For example, in the case of lysozyme on SP-Sepharose FF (Fig. 2.5C), as the protein load is increased from 0.5 to 5 mg/ml, the peak maximum shifts slightly to the right. However, with a further increase to 10, 25, and 50 mg/ml, the peak consistently shifts to the left. This behavior, captured by the EI method, is attributed to the slight S-shaped isotherm behavior for lysozyme at relatively high salt concentrations (see Fig. 2.2A). Figures 2.5E and 2.5F show additional high-protein load LGE experimental data for the lysozyme/SP-Sepharose FF system with varying gradient slope and for mAb A/POROS XS system

varying both protein load and gradient slope, respectively. Excellent agreement between predictions based on the EI method and the experimental results is seen for both systems over a broad range of conditions covering both the dilute limit as well as highly overloaded conditions.

Figure 2.7 compares the experimental protein elution curves of mAb B obtained over a fairly broad range of conditions and at high protein loads from 5 to 45 mg/ml of column (dashed lines) with predictions based on the EI method (solid lines). Since the binding capacity for mAb B on POROS XS at the load conditions (40 mM Na^+, pH 5.5 or 6) is about 160 mg/ml of particle volume and the extraparticle void fraction is 0.35, the column binding capacity is about 100 mg/ml. Thus, these protein loads are between 5% and 45% of the column binding capacity. Figure 2.7A shows the results for elution with a salt gradient at essentially constant pH with different protein loads. Similar for to the case of mAb A in Fig. 2.5, the Na^+ concentration at mAb B elution is lower at the higher protein loads compared to the dilute limit. The peak shapes remain, however, fairly symmetrical since the degree of over-loading was moderate in these runs. Figure 2.7B shows the effect of protein loading for elution with a pH gradient with a constant Na^+ concentration. As seen from this figure, the experimental pH gradients are essentially linear, consistent with predictions while the Na^+ concentration varies only minimally. Since the Na^+ concentration is obtained directly from the conductivity and the conductivity is, in turn, affected by viscosity, the variations in Na^+ concentration are associated with the higher solution viscosity caused by the high mAb concentration at elution. As protein loading increases, elution occurs at lower pH values in a manner qualitatively analogous to the trends seen with salt gradients at constant pH. As seen in Fig. 2.7B, at high protein loadings the peak shape is no longer Gaussian but becomes trapezoidal. This occurs because of the favorable nature of the isotherm for lower pH conditions corresponding to the early portion of the elution profile and the more linear nature of the isotherm for the higher pH conditions corresponding to the late portion.

Figure 2.7C shows the effect of gradient slope for elution with pH gradients designed to be linear with a constant Na^+ concentration of 100 mM and with the same protein loading. The experimental pH gradients are essentially linear further confirming the accuracy of the pH modeling approach used to predict buffer compositions. The experimental Na^+ concentration, determined again di-

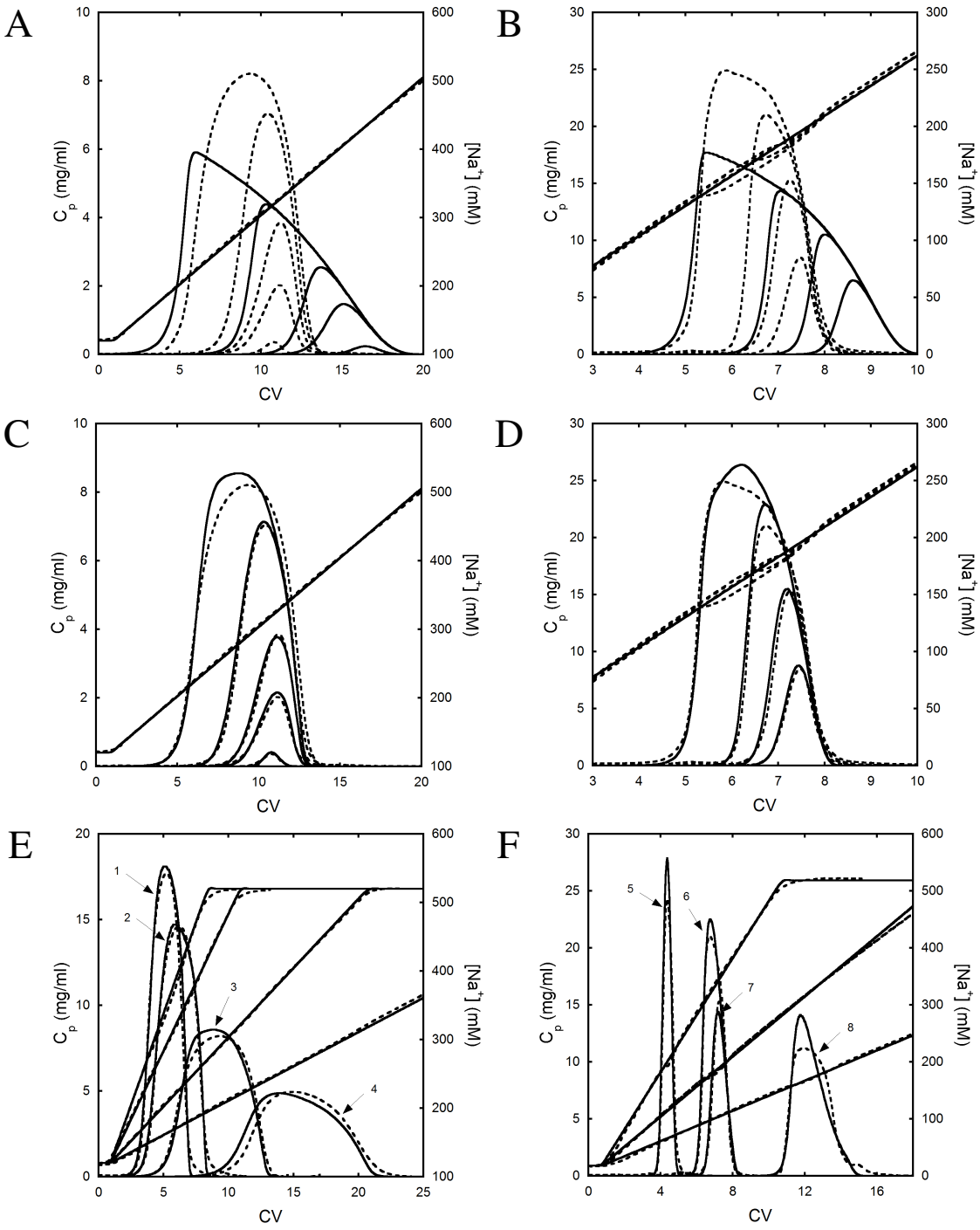


Figure 2.5: Comparison of the experimental elution profiles obtained for LGE with varying protein loads (dashed lines) with model predictions (solid lines) based on either the SMA model or the EI method for lysozyme on SP-Sepharose FF with 100-500 mM NaCl gradients in (A), (C), and (E), and for mAb A on POROS XS with 0-500 mM NaCl gradients in (B), (D), and (F). The dash-dot lines represent the Na^+ concentration. Protein loads are 0.5, 5, 10, 25, and 50 mg/ml in (A) and (C) and 6.3, 13, 27, and 53 mg/ml in (B) and (D). Gradient lengths and protein loads in (E) and (F) are: curve 1: 7.5. CV, 50 mg/ml; curve 2: 10 CV, 50 mg/ml; curve 3: 20 CV, 50 mg/ml; curve 4: 40 CV, 40 mg/ml; curve 5: 10 CV, 13 mg/ml; curve 6: 20 CV, 27 mg/ml; curve 7: 20 CV, 13 mg/ml; curve 8: 40 CV, 27 mg/ml.

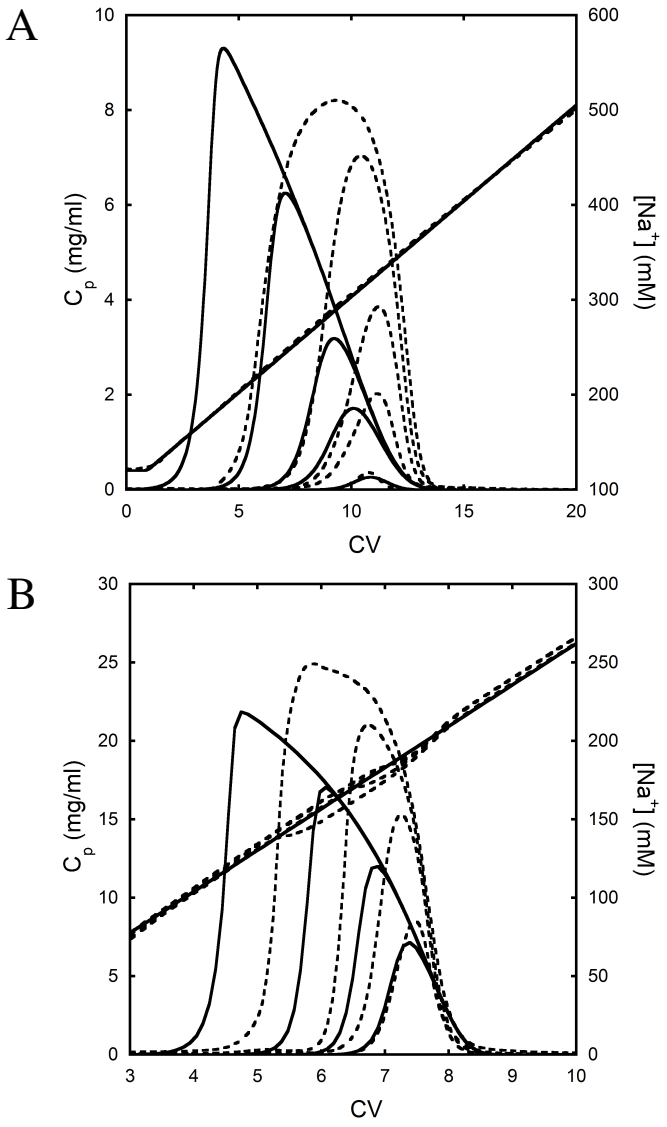


Figure 2.6: Comparison of the experimental elution profiles obtained for LGE with varying protein loads (dashed lines) with model predictions (solid lines) based on the SMA model for lysozyme (A) and for mAb A (B) using parameters in Table A2 in the Appendix. The experimental conditions are the same as Figs. 2.5A and 2.5A.

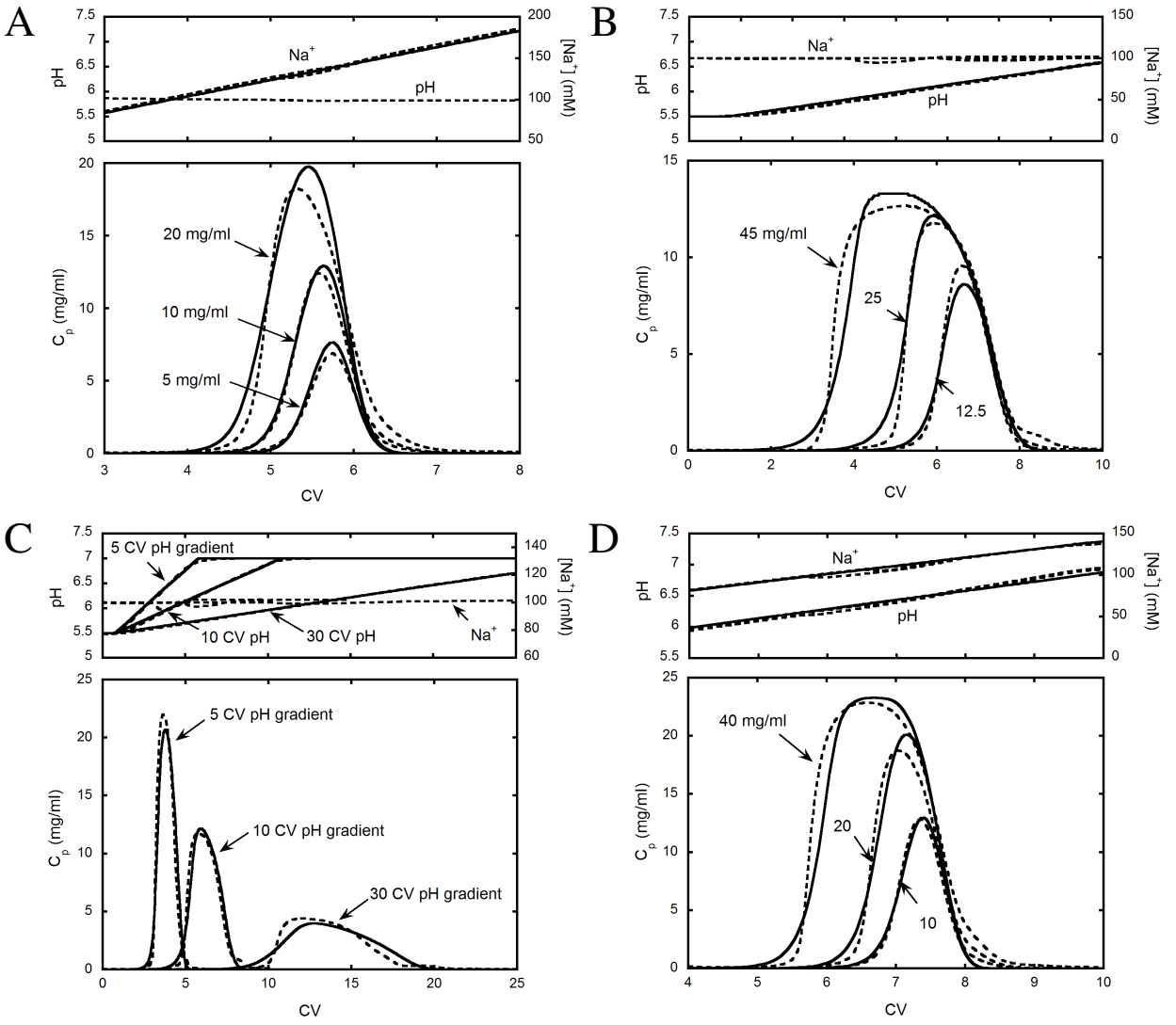


Figure 2.7: Comparison of experimental (dashed lines) and predicted (solid lines) elution profiles of mAb B on POROS XS obtained for (A) 40 to 240 mM Na^+ gradients at pH 6 in 10 CV with different protein loads; (B) pH 5.5 to 7.0 gradients at 100 mM Na^+ in 10 CV with different protein loads; (C) pH 5.5 to 7 gradients at 100 mM Na^+ with a 25 mg/ml protein load and different gradient slopes; and (D) pH 5.5 to 7.0 gradients combined with 40 to 140 mM Na^+ gradients in 10 CV with different protein loads.

rectly from the conductivity, stays constant except for the fluctuations observed during the elution of the protein peak, which are associated with higher viscosity. As expected and seen in this figure, steeper pH gradients result in sharper peaks with the peak eluting at slightly higher pH in a manner qualitatively consistent with the higher salt concentration at elution observed for steeper salt gradients at constant pH, Fig. 2.1. Finally, Fig. 2.7D shows the results obtained for elution with a simultaneous salt and pH gradient using different protein loadings. Both experimental pH and Na^+ concentration gradients are essentially linear, consistent with the buffer design approach used in this work, again with the exception of fluctuations near the peak elution associated with the high viscosity of the eluted protein. Because of the simultaneous gradients, elution and Na^+ concentration at elution are both lower compared to those of runs with pH gradients at constant Na^+ or salt gradients at constant pH.

In all four cases illustrated in Fig. 2.7, model predictions based on the EI method are in excellent agreement with the experimental profiles. The most significant deviations are seen in the front of the peak at very high protein loads and in the tailing portion of the elution profile. These relatively small deviations are likely caused by the simplified nature of the mass transfer model based on the LDF approximation and the potential presence of small amounts of aggregates that may become concentrated in the tail end of the elution peaks, thereby altering the shape of the elution profile.

2.4.4 Sensitivity analysis

Obviously, the accuracy of predictions based on the EI method is expected to increase with the amount of available isotherm data. However there is a trade-off between accuracy and the required time and cost of obtaining large amounts of data. To assess the effect of the "granularity" of the isotherm data, predicted elution curves were generated for the combined pH-salt gradient experiment shown in Fig. 2.7D using different subsets of the isotherm data. The results are compared in Fig. 2.8. Predictions in Fig. 2.8A use only the isotherm data at 20, 95, 145, and 220 mM Na^+. Limited to only these four Na^+ concentrations lowers the level of data granularity at higher pH values where the transition from binding to non-binding conditions narrows to a lower Na^+ concentration

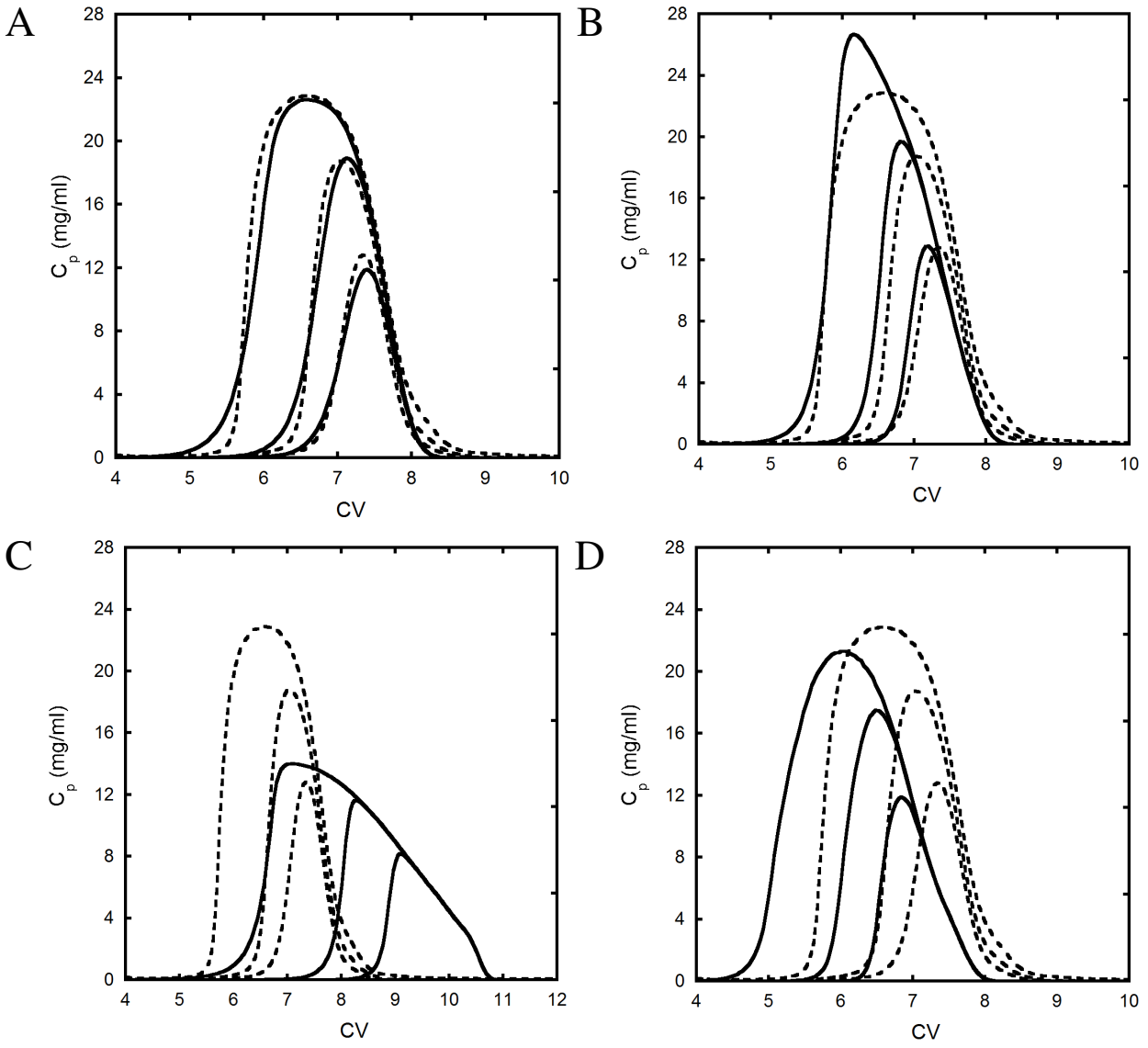


Figure 2.8: Comparison of experimental (dashed lines) and predicted (solid lines) elution profiles using experimental batch isotherm data sets from Fig. 2.3 with varying levels of granularity. (A) predictions using isotherms data only at 20, 95, 145, 220 mM Na^+; (B) predictions using isotherm data only at protein concentrations less than 10 mg/ml; (D) predictions using isotherms data only at pH 5.5 and 7.0; and (D) predictions using isotherm data only at 20 and 220 mM Na^+. Experimental elution data are the same as those in Fig. 2.7D.

range. Although more isotherm space is being interpolated, this decrease in the amount of data did not have a large impact on predictions. Prediction in Fig. 2.8B use only the isotherm data points at protein concentrations less than 10 mg/ml. The predictions become much less accurate than with the full data set (Fig. 2.7D), especially at the higher protein loadings, which, upon elution result in protein concentrations well above 10 mg/ml. Predictions in Fig. 2.8C use only the isotherm data at pH 5.5 and 7. Between two pH values, in the pH dimension of the experimental isotherm space, the effect of pH is linearly interpolated resulting in large discrepancies between predicted and experimental curves. Finally, predictions in Fig. 2.8D use only the isotherm data at 20 and 220 mM Na^+. As in the previous case, large deviations are again seen which results from inaccuracy of linearly-interpolating the effect of the Na^+ concentration to accurately predict the elution behavior. Based on these results, for the case at hand, satisfactory predictions require isotherm data within the region of protein concentrations experienced during elution, data at pH values in 0.5 pH unit interval over the range of elution pHs and salt concentrations covering the region of the isotherm space where the transition between strongly bound and weakly bound conditions occur.

References

(*1*) Money, R. W., and Davies, C. W., (1932). The extent of dissociation of salts in water. IV. Bi-bivalent salts. *Transactions of the Faraday Society 28*, 609.

(*2*) Brooks, C. A., and Cramer, S. M., (1992). Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE Journal 38*, 1969–1978.

(*3*) Velayudhan, A., and Horvath, C., (1986). On the stoichiometric model of electrostatic interaction chromatography for biopolymers. *Journal of Chromatography 367*, 160–162.

(*4*) Yamamoto, S., Nakanishi, K., and Matsuno, R., *Ion Exchange Chromatography of Proteins*; Marcel Dekker: New York, NY, 1988.

(*5*) Yamamoto, S., Nakanishi, K., Matsuno, R., and Kamikubo, T., (1983). Ion exchange chromatography of proteins-prediction of elution curves and operating conditions. 1. Theoretical considerations. *Biotechnology and Bioengineering 25*, 1465–83.

(*6*) Yamamoto, S., Nakanishi, K., Matsuno, R., and Kamikubo, T., (1983). Ion-exchange chromatography of proteins-prediction of elution curves and operating conditions. 2. Experimental-verification. *Biotechnology and Bioengineering 25*, 1373–1391.

(*7*) Zhu, M., and Carta, G., (2014). Adsorption of polyethylene-glycolated bovine serum albumin on macroporous and polymer-grafted anion exchangers. *Journal of Chromatography A 1326*, 29–38.

(*8*) Sips, R., (1948). On the structure of a catalyst surface. *16*, 490–495.

(*9*) Glueckauf, E., (1955). Theory of chromatography. *Transactions of the Faraday Society 51*, 1540.

(*10*) Carta, G., Ubiera, A. R., and Pabst, T. M., (2005). Protein Mass Transfer Kinetics in Ion Exchange Media: Measurements and Interpretations. *Chemical Engineering & Technology 28*, 1252–1264.

(*11*) Teo, W. K., and Ruthven, D. M., (1986). Adsorption of water from aqueous ethanol using 3-A molecular sieves. *Industrial & Engineering Chemistry Process Design and Development 25*, 17–21.

(*12*) Beynon, R. J., and Easterby, J. S., *Buffer Solutions-The Basics*; Oxford University Press: Oxford, 1996.

(*13*) Pabst, T. M., Suda, E. J., Thomas, K. E., Mensah, P., Ramasubramanyan, N., Gustafson, M. E., and Hunter, A. K., (2009). Binding and elution behavior of proteins on strong cation exchangers. *Journal of Chromatography A 1216*, 7950–6.

(*14*) Marquardt, D. W., (1963). An algorithm for least-squares estimation of non-linear parameters. *Journal of the Society for Industrial and Applied Mathematics 11*, 431–441.

(*15*) Monkos, K., and Turczynski, B., (1999). A comparative study on viscosity of human, bovine and pig IgG immunoglobulins in aqueous solutions. *International Journal of Biological Macromolecules 26*, 155–159.

3 Multicomponent Separations on Ion Exchange

3.1 Introduction

The empirical interpolation (EI) method is extended to predict highly overloaded multicomponent elution behavior on a cation exchange (CEX) column based on batch isotherm data. Instead of a fully mechanistic model, the EI method employs an empirically modified multicomponent Langmuir equation to correlate two-component adsorption isotherm data at different salt concentrations. Piecewise cubic interpolating polynomials are then used to predict competitive binding at intermediate salt concentrations. The approach is tested for the separation of monoclonal antibody monomer and dimer mixtures by gradient elution on the cation exchange resin Nuvia HR-S. Adsorption isotherms are obtained over a range of salt concentrations with varying monomer and dimer concentrations. Coupled with a lumped kinetic model, the interpolated isotherms predict the column behavior for highly overloaded conditions. Predictions based on the EI method showed good agreement with experimental elution curves for protein loads up to 40 mg/mL column or about 50% of the column binding capacity. The approach can be extended to other chromatographic modalities and to more than two components.

3.2 Theoretical development

The method is developed specifically for the case of two component ion-exchange chromatography using salt gradient elution. However, the method can to be extended to more than two components since each component's isotherm behavior is described separately as a function of all protein liquid concentrations and the salt concentration. The model can also be extended to pH gradients using the same approach for the single component case presented in Chapter 2.

In this work, we compare the predictions from the best SMA model fit to the predictions from the EI method. The latter has three components: (1) protein binding equilibrium as a function of protein and salt concentrations; (2) protein binding kinetics; and (3) column dynamics.

3.2.1 Protein binding equilibrium

The total concentration of each component within the chromatographic particles at equilibrium, $\bar{q}_i$, is the sum of the bound protein and the protein held within the particle pores. Accordingly,

$$\bar{q}_i = q_i + \varepsilon_p K_{D,i} C_i \tag{3.1}$$

where q_i is the bound protein concentration, C_i is the protein concentration in solution, ε_p is the intraparticle porosity, and $K_{D,i}$ is the partition coefficient accounting for the steric exclusion of each protein in the resin pores. The intraparticle porosity ε_p is measured from the chromatographic retention of a small molecule (e.g. salt) and $K_{D,i}$ from the chromatographic retention of each protein under non-binding conditions. Note that both $\bar{q}_i$ and q_i are expressed in units of mg per mL of resin bead volume.

Two methods are used here to describe the bound concentration: the mechanistic SMA model and the EI method that does not rely on a mechanistic equation. The SMA model assumes a stoichiometric displacement of counterions, Na^+ in our case, by charged protein molecules. For two components, M (monomer) and D (dimer), the SMA equation can be written as[1,2]:

$$q_i = \frac{K_{e,i}\left[q_0 - (z_M + \sigma_M)q_M - (z_D + \sigma_D)q_D\right]^{z_i}}{\left[\mathrm{Na}^+\right]^{z_i}} C_i \tag{3.2}$$

where $K_{e,i}$ is the equilibrium constant, q_0 is the resin charge density, z_i is the protein effective charge, and σ_i is the steric hindrance factor that represents shielding of ligands by bound protein. For conditions where $(z_M + \sigma_M)q_M + (z_D + \sigma_D)q_D << q_0$ (i.e. at high salt concentrations and/or very low protein concentrations), eq. 3.2 reduces to:

$$q = \frac{K_{e,i}\left(q_0\right)^{z_i}}{\left[\mathrm{Na}^+\right]^{z_i}} C_i \tag{3.3}$$

The SMA parameters can be determined using different methods. In one approach, used here, the

parameters $K_{e,i}(q_0)^{z_i}$ and z_i are determined from linear salt gradient elution experiments at low protein loads as shown in ref.[3]. The charge density q_0 and the σ_i values are then determined by regressing eq. 3.2 to isotherm data at high protein loads over a range of salt concentrations. This approach has the advantage of preserving the Henry's law limit behavior, which is critically important in gradient elution. The EI method uses an arbitrary function to fit two-component adsorption data at each salt concentration. The following equations proposed by Gu et al.[4] for systems with unequal saturation capacities are used in this work:

$$q_M = \frac{K_{L,M}C_M\left[\left(1+K_{L,D}C_D\right)q_{m,M} - \theta_{M,D}K_{L,D}C_D q_{m,D}\right]}{1+K_{L,M}C_M+K_{L,D}C_D+\left(1-\theta_{M,D}\right)K_{L,M}K_{L,D}C_M C_D} \tag{3.4a}$$

$$q_D = \frac{K_{L,D}C_D\left[\left(1+K_{L,M}C_M\right)q_{m,D} - K_{L,M}C_M q_{m,M}\right]}{1+K_{L,M}C_M+K_{L,D}C_D+\left(1-\theta_{M,D}\right)K_{L,M}K_{L,D}C_M C_D} \tag{3.4b}$$

The five parameters $q_{m,M}$, $q_{m,D}$, $K_{L,M}$, $K_{L,D}$, and $\theta_{M,D}$ are regressed independently at each of the [Na^+] values used in the experimental data set. In this work, eqs. 3.4a and 3.4b were fitted to experimental isotherm data sets at three different Na^+ concentrations which results in five isotherm parameters for each of the three [Na^+] values to give a total of 15 model parameters to describe both q_M and q_D as a function of C_M and C_D at constant [Na^+].

The procedure used to calculate the bound protein concentrations q_M^* and q_D^* at given values of C_M^*, C_D^*, and [Na^+] according to the EI method is illustrated schematically in Fig. 3.1. In the first step, eqs. 3.4a and 3.4b are used to generate q_M' and q_D' values at C_M^*, C_D^* for each experimental [Na^+] value. In the second step, the q_M' and q_D' values are used to construct piecewise cubic interpolating polynomials (PCHIP), shown as solid lines in Fig. 3.1. In the final step, q_M^* and q_D^* are calculated from PCHIP at [Na^+]*. To increase the speed of the column simulation, the interpolation process is repeated for a combination of one hundred values of C_M^*, one hundred values of C_D^*, and one hundred values of [Na^+]* to generate 100^3 interpolated q_M^* and q_D^* values that are used as a look up table. Each grid of points corresponding q_M^* and q_D^* is given as a separate input to MATLAB's *griddedInterpolant* function which uses trilinear interpolation to quickly return q_M^* and q_D^* values

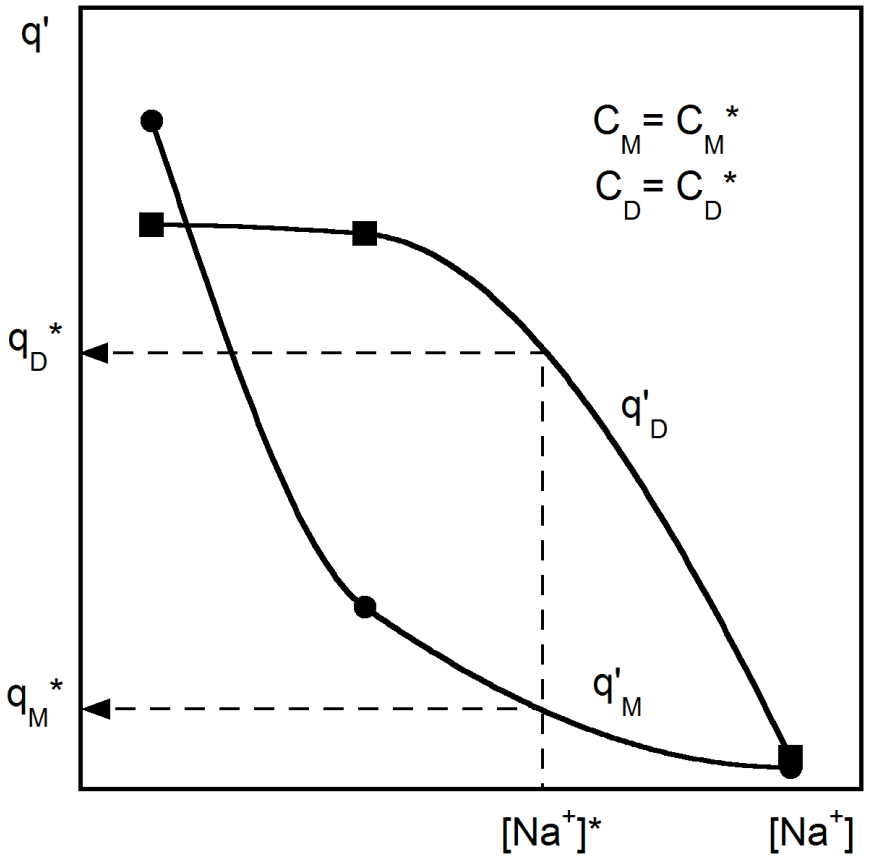


Figure 3.1: Illustration of the PCHIP interpolation method to calculate unknown q_M^* and q_D^* at given values of C_M^*, C_D^*, and [Na$^+$]*. The q'-values evaluated for the monomer and dimer are circles and squares, respectively, calculated at $C_M = C_M^*$ and $C_D = C_D^*$ using eqs. 3.4a and 3.4b at each experimental [Na$^+$]. The two solid lines are the PCHIP curves constructed independently using the q'-values and evaluated at [Na$^+$]*.

for a given input of C_M^*, C_D^*, and [Na^+]*. If the queried isotherm point is outside the range of known [Na^+] values, trilinear extrapolation is used. For a queried point within the known [Na^+] range but outside the known C_M^*, C_D^* range, the values of q_M^* and q_D^* are extrapolated using eqs. 3.4a and 3.4b.

3.2.2 Protein adsorption kinetics

The protein binding kinetics is described according to the linear driving force (LDF) approximation based on the liquid phase concentration driving force:

$$\frac{\partial \bar{q}_i}{\partial t} = \frac{60 D_{e,i}}{d_p^2}\left(C_i - C_i^*\right) \tag{3.5}$$

where d_p is the particle diameter, $D_{e,i}$ is the effective diffusivity of species i, and C_i^* is the concentration of species i in equilibrium with $\bar{q}_i$. In this work, we use the slope of van Deemter curves which is the height equivalent to a theoretical plate (HETP) vs. the superficial velocity to experimentally measure the value of $D_{e,i}$, although several other methods can be used as well[5]. The data in previous work using the same antibody monomer and dimer and resin are used for this purpose in this work[3].

3.2.3 Column dynamics

The column dynamics for the protein mixture and buffer components are modeled using the equation:

$$\varepsilon \frac{\partial C_i}{\partial t} + (1-\varepsilon)\frac{\partial \bar{q}_i}{\partial t} + u\frac{\partial C_i}{\partial x} = 0 \tag{3.6}$$

where x is the column axial coordinate and u is the superficial velocity. The term $\partial \bar{q}_i / \partial t$ is given by eq. 3.5. Equation 3.6 is solved numerically by discretizing the axial derivative by backwards finite differences over the length of the column, L, and solving the resulting set of ordinary differential equations with MATLAB's *ode15s* built-in routine. Numerical dispersion caused by the discretization was minimized by using at least 80 discretization points which gave only minimal dispersion

of [Na^+] profiles and no significant effect on the band broadening of the protein concentration profiles. All calculations were done in MATLAB R2016a (The Mathworks, Natick, ME, USA) on a Dell Precison T1700, Intel i7 series 3.40GHz.

3.3 Materials and methods

3.3.1 Materials

The monoclonal antibody (M_r ~150,000, pI ~9) used in this work was obtained from a stock solution containing 28% dimer according to SEC analysis, and was provided by MedImmune (Gaithersburg, MD, USA). Samples that were highly enriched in either monomer or dimer were obtained from this stock solution using preparative SEC with a Superdex 200 Increase 10/300 GL column from GE Healthcare (Marlborough, MA, USA). For this purpose, sequential injections of 100 µl of 38 mg/mL protein were eluted at 0.5 mL/min in 10 mM phosphate containing 140 mM NaCl at pH 7.4. Fractions of enriched dimer and monomer were then buffer exchanged using PD-10 desalting columns from GE Healthcare (Marlborough, MA, USA) and stored at 4 °C.

The CEX resin used in this work is Nuvia HR-S, from Bio-Rad Laboratories (Hercules, CA, USA). The resin is a macroporous strong cation exchanger with a mean particle radius $d_p = 52$ µm and an intraparticle porosity $\varepsilon_p = 0.77$, based on the chromatographic retention of sodium chloride. The non-binding partition coefficients K_D are 0.52 and 0.47 for the mAb monomer and dimer, respectively. All buffers were made with dibasic sodium phosphate titrated with phosphoric acid to pH 7. Sodium chloride, dibasic sodium phosphate, and phosphoric acid used in buffer preparations were purchased from Fisher Scientific (Fair Lawn NJ, USA).

3.3.2 Methods

3.3.2.1 Analytical SEC

Analytical SEC was done using a Waters ACQUITY UPLC H-Class system (Milford, MA, USA) with a Waters ACQUITY BEH SEC column (200Å, 1.7 µm, 4.6 mm X 150 mm) in a running buffer

consisting of 100 mM Na_2HPO_4, 300 mM NaCl , pH 6.8 at 0.3 mL/min. The monomer and dimer were assumed to have the same extinction coefficient of 1.33 mL mg^{-1} cm^{-1}, determined from a protein assay, to convert UV absorbance at 280 nm to mass concentrations in mg/mL. Since the SEC peaks exhibited significant tailing, especially for the dimer, quantitation was accomplished by fitting the peak profiles with exponentially-modified-Gaussian functions (EMG) to determine the ratio of monomer and dimer peak areas in a mixture as shown in Fig. A4 of the Appendix.

3.3.2.2 Batch adsorption isotherms

Batch adsorption isotherm data were obtained at pH 7 while varying monomer, dimer, and Na^+ concentrations. The resin was first equilibrated in the desired buffer followed by removal of the extraparticle liquid using 2 mL Corning Costar Spin-X microfiltration tubes (Sigma-Aldrich, St. Louis, MO, USA) with an Eppendorf Minispin bench-top centrifuge operated at 5000 rpm for 15 min. Samples of the filtered resin were added to 2 mL plastic tubes and mixed with 0.5 to 3 mL of protein solution with varying monomer/dimer ratios for 24 h. The final concentrations were measured with a Nanodrop™ 2000c (Thermo Fisher Scientific, Waltham, MA, USA) and the final monomer/dimer ratios were measured by UPLC SEC as described in Section 3.3.2.1. The mass of protein taken up by the resin was calculated using a mass balance for both the monomer and dimer and the bound concentrations were calculated by dividing their respective masses by the resin volume. The resin volume was determined by converting the hydrated resin mass using the density of the resin, 1.076 g/mL, determined from pycnometer measurements.

3.3.2.3 Column experiments

Linear gradient elution (LGE) experiments were done with a 0.5 cm diameter × 5 cm long Tricorn column (GE Healthcare, Marlborough, MA) packed with the Nuvia HR-S resin with gradients from 0 to 300 mM NaCl in 25 CV using an AKTA Pure 25 system from GE Healthcare (Marlborough, MA, USA) at a flow rate of 0.5 mL/min (~4 min residence time). The column extraparticle porosity was determined to be $\varepsilon = 0.39$ based on pressure drop data using the Blake-Kozeny equation with a numerical constant of 150[6]. Injection volumes between 2.4 and 9.7 mL were loaded using a 50

mL superloop corresponding to total protein loadings ranging from 10 to 40 mg/mL of column. The elution profiles were collected in 0.25 mL fractions and analyzed by UPLC SEC as described in Section 3.3.2.1.

3.4 Results and discussion

3.4.1 Protein binding equilibrium

Figure 3.2 shows the batch isotherm data at Na^+ concentrations of 20, 80, and 200 mM for monomer-dimer mixtures spanning concentration ranges of 0-6 mg/mL for the monomer and 0-3 mg/mL for the dimer. The total monomer concentration in the resin, $\bar{q}_M$, is plotted in panels A and C while that of the dimer, $\bar{q}_D$, is plotted in panels B and D. At 20 mM Na^+, the isotherms are favorable for both the pure monomer and the pure dimer with binding capacities in the range of 130-140 mg/mL for both. Competitive binding is evident for these conditions, becoming more pronounced at 80 mM Na^+. At higher salt concentrations, the binding capacity becomes much smaller and the isotherms become linear and independent of each other.

The surfaces that are shown accompanying the data points in Fig. 3.2 are based on the SMA model (eq. 3.2) for panels A and B and on eqs. 3.4a and 3.4b for panels C and D. As shown in ref.[3], at Na^+ concentrations ≥ 170 mM for the monomer and ≥ 200 mM for the dimer, the isotherms are linear and consistent with eq. 3.3 with $z_M = 10.2$ and $K_{e,M}(q_0)^{z_M} = 2.01\times10^{23}$ for the monomer, and $z_D = 14.8$ and $K_{e,D}(q_0)^{z_D} = 3.21\times10^{34}$ for the dimer, both with q_0 in mmol/L units. For the full SMA model description, the parameters σ_M , σ_D, and q_0 were regressed simultaneously to the global dataset using MATLAB's built-in solver, *lsqnonlin*, and are listed in Table 3.1. While the model obviously fits the high Na^+ data, since the parameters were constrained to ensure such agreement, large deviations are seen in Fig. 3.2A and 3.2B at lower Na^+ concentrations. The average deviation of the model from the data was 49.8% at 20 mM Na^+ and 30.5% at 80 mM Na^+. The reason why the SMA model fails to represent the data in this case can be found in the trend of selectivity vs. Na^+ concentration. According to the SMA model, the selectivity toward the protein with higher charge is predicted to always increase as the Na^+ concentration decreases[3]. On the other hand, according

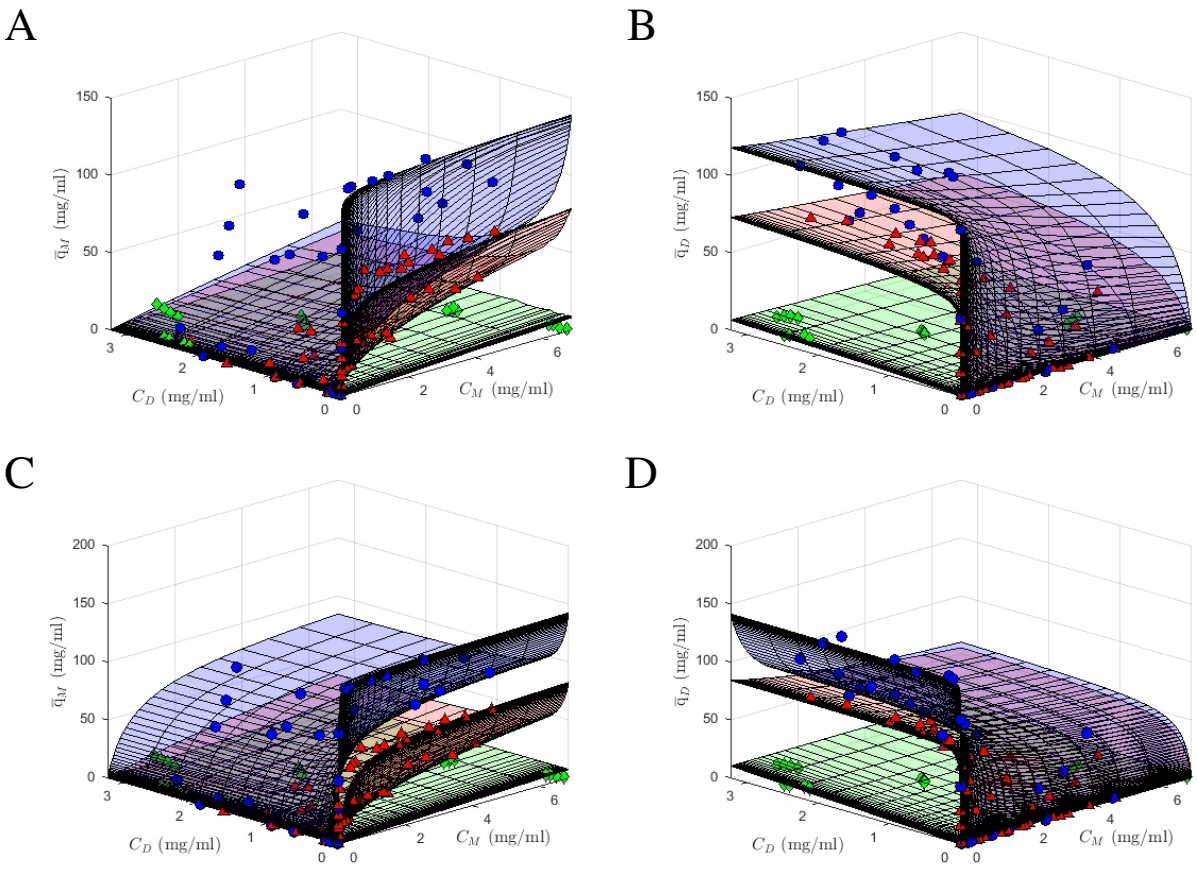


Figure 3.2: Batch adsorption isotherms for the monomer in panel A and C and for the dimer in panels B and D. The symbols ● (blue), ▲ (red), and ◆ (green) represent 20, 80, and 200 mM Na^+ concentrations, respectively. The surface lines in A and B are calculated with the SMA model using the global best-fit parameters given in Table 3.1. In C and D, the blue and red surface lines for 20 and 80 mM [Na^+], respectively, are calculated with eqs. 3.4a and 3.4b with the parameters given in Table 3.2. The green surfaces for 200 mM [Na^+] in C and D are the linear isotherms from ref.[3].

to the experimental data, at 20 mM Na^+ binding is favorable for both monomer and dimer and the selectivity has a relatively low average value of 2.9. At 80 mM Na^+, however, while binding is still favorable for both individual components, the selectivity increases dramatically reaching an average value of 23. Finally, at 200 mM Na^+, when the isotherms become linear, the selectivity becomes 3.0, continuing to decrease as the Na^+ concentration is increased further. The inability of the SMA model to accommodate the experimental selectivity trends results in a poor average description of the data when the model parameters are fitted to the global dataset.

As seen in panels C and D of Fig. 3.2, eqs. 3.4a and 3.4b provide a much better agreement with the experimental data compared to the SMA model with an average deviation of the model from the data 18.3% at 20 mM Na^+ and 22.1% at 80 mM Na^+. In this case, the parameters were regressed separately at each salt using the MATLAB function *lsqnonlin*. The best-fit parameters determined at 20 and 80 mM Na^+ are listed in Table 3.2 and the corresponding surfaces are shown in panels C and D of Fig. 3.2. At 200 mM Na^+ the surfaces are calculated according to the linear isotherm model (eq. 3.3) with the parameters from Reck et al.[3] and are the same as those shown for the SMA model.

Table 3.1: SMA model parameters fitted to isotherm data in Figures 3.2A and 3.2B

	Monomer	**Dimer**
z_i	10.2	14.8
q_0 (mM)	200	200
$K_{e,i}$	0.68	2.83
σ_i	180	440

Table 3.2: Langmuir parameters regressed at each Na^+ concentration in Figures 3.2C and 3.2D

$[Na^+]$ (mM)	20		80	
	Monomer	**Dimer**	**Monomer**	**Dimer**
q_m (mg/ml)	140	140	83	83
$K_{L,i}$ (ml/mg)	130	780	3.8	110
$\theta_{M,D}$	0.99		0.98	

To speed up the EI-based computation of chromatographic elution, separate grids of 100 × 100 × 100 values of q_M and q_D are generated for 100^3 combinations of C_M, C_D, and [Na^+] values. These q_M and q_D values are given as a separate inputs to MATLAB's *griddedInterpolant* function which uses trilinear interpolation to quickly return q_M and q_D values for a given input of C_M, C_D, and [Na^+]. The same approach can be used to return corresponding equilibrium protein liquid concentrations, C_M and C_D, by giving *griddedInterpolant* q_M, q_D, and [Na^+] values as inputs. This work uses the inverse to return equilibrium C_M and C_D values which are more convenient with eq. 3.5.

3.4.2 Protein adsorption kinetics

The effective pore diffusivity for the monomer and dimer were determined from prior work based on the slope of van Deemter curves under non-binding conditions[3]. The effective pore diffusivities $D_{e,M} = 8.0 \times 10^{-8}$ cm^2/s and $D_{e,D} = 4.1 \times 10^{-8}$ cm^2/s for the monomer and dimer, respectively, are about 10 times smaller than their free solution diffusivities at room temperature. For the salt, D_e was assumed to be 1.0×10^{-5} cm^2/s. Because of this high value, mass transfer resistance is predicted to be negligible for salt resulting in little broadening of the gradient profile in agreement with experimental measurements.

3.4.3 Gradient elution at high protein loads

Figure 3.3 shows the experimental elution curves for the feed mixture containing 28% dimer with total protein loads of 10, 20 and 40 mg per mL of column. Based on the batch isotherm data and the column extraparticle porosity of 0.39, the binding capacity is around 80 mg/mL of column at 20 mM Na^+. Thus, these loads correspond to up to about 50% of the column capacity. The experimental monomer and dimer concentrations are shown in this figure as solid and hollow points, respectively, while the Na^+ concentration, derived from the effluent conductivity, is shown as a dotted line. Model predictions are shown as solid lines for the monomer and salt and as dashed lines for the dimer based either on the SMA model using the parameters in Table 3.1 in panels A-C

or the EI method using the parameters in Table 3.2 in panels D-F. As seen in this figure, the SMA model predicts earlier elution for the monomer and dimer and, in the case of the 40 mg/mL load, predicts shark fin-shaped elution peaks that are not seen experimentally. Conversely, the EI method predictions are in excellent agreement with the experimental data at all protein loads tested. As the load increases, the EI method accurately predicts the monomer and dimer eluting at higher peak concentrations, at lower Na^+ concentrations, and with broader peaks compared to the low loading results.

A final consideration is to what extent the peak profiles can be predicted using exclusively the linear limit of the isotherm, eq. 3.3. According to Yamamoto[7], for these conditions the peak profile is calculated from the following equations:

$$\frac{C_i}{C_{i,F}} = \frac{V_F}{V_C \varepsilon \left(1 + k'_{R,i}\right)} \left[\frac{N_{eff,i}}{2\pi \left(t_i^*\right)^3}\right]^{\frac{1}{2}} \exp\left[-\frac{N_{eff,i}}{2} \frac{\left(t_i^* - 1\right)^2}{t_i^*}\right] \tag{3.7}$$

where V_F is the feed volume, V_C is the column volume, $k'_{R,i} = \left(1-\varepsilon\right) K_{e,i} \left(q_0 / [Na^+]_{R,i}\right)^{z_i} / \varepsilon$, and $t_i^* = t / \left[\left(\varepsilon L / u\right)\left(1 + k'_{R,i}\right)\right]$. In these equations, $[Na^+]_{R,i}$ is the Na^+ concentration at which the peak is predicted to elute and is given by[7]:

$$[Na^+]_{R,i} = \left\{ \left(1-\varepsilon\right) K_{e,i} q_0^{z_i} \left(z_i + 1\right) \frac{\Delta\left[Na^+\right]}{C V_G} \right\}^{\frac{1}{z_i+1}} \tag{3.8}$$

where $\Delta\left[Na^+\right]$ is the difference between final and initial Na^+ concentrations and CV_G is the duration of the gradient in column volume units. Finally, $N_{eff,i}$ is the effective number of plates for the protein which is obtained by dividing the actual number of plates for the protein, N_i, by a correction factor given by eq. 11 in ref.[7] to account for peak compression in the gradient. In our work, according to ref.[5], we calculated $N_i = 30(1-\varepsilon) D_{e,i} L / u d_p^2$, using the same $D_{e,i}$ values used for the numerical predictions. Figure 3.4 shows the peak profiles calculated according to eq. 3.7, shifted in time to match the $[Na^+]_{R,i}$ calculated from eq. 3.8[7], in comparison with the numerical predictions using the EI method for a 40 mg/mL total protein load, for the conditions of Fig. 3.3F. As seen in this figure, the EI method predicts, which is in agreement with the experimental data, earlier elution and

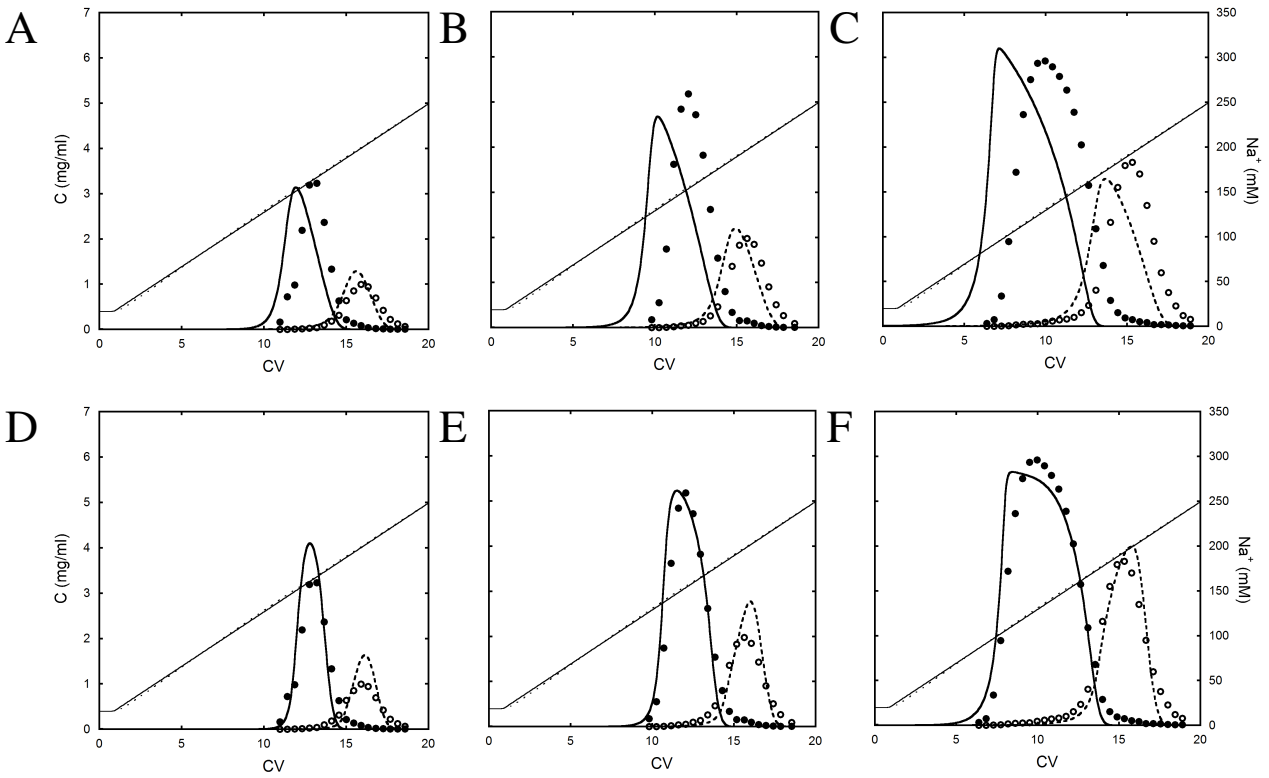


Figure 3.3: Comparison of experimental data (points) and model predictions (lines) based on the SMA model (panels A, B, and C) and the EI method (panels D, E, and F) with 20 to 320 mM Na^+ gradients at pH 7 in 25 CV. Protein loads are 10 (A and D); 20 (B and E); and 40 mg/mL column (C and F). Monomer and dimer are solid and hollow points, respectively. Predicted concentrations of the monomer and dimer are solid and dashed lines, respectively.

broader peaks than those predicted by eqs. 3.7-3.8 as a result of the overloaded conditions used. Reducing the protein load leads to closer agreement with eqs. 3.7-3.8, but quantitative agreement between numerical predictions and eqs. 3.7-3.8 is seen only at protein loads less than about 5-10% of the column binding capacity or about 4-8 mg/mL for the case at hand.

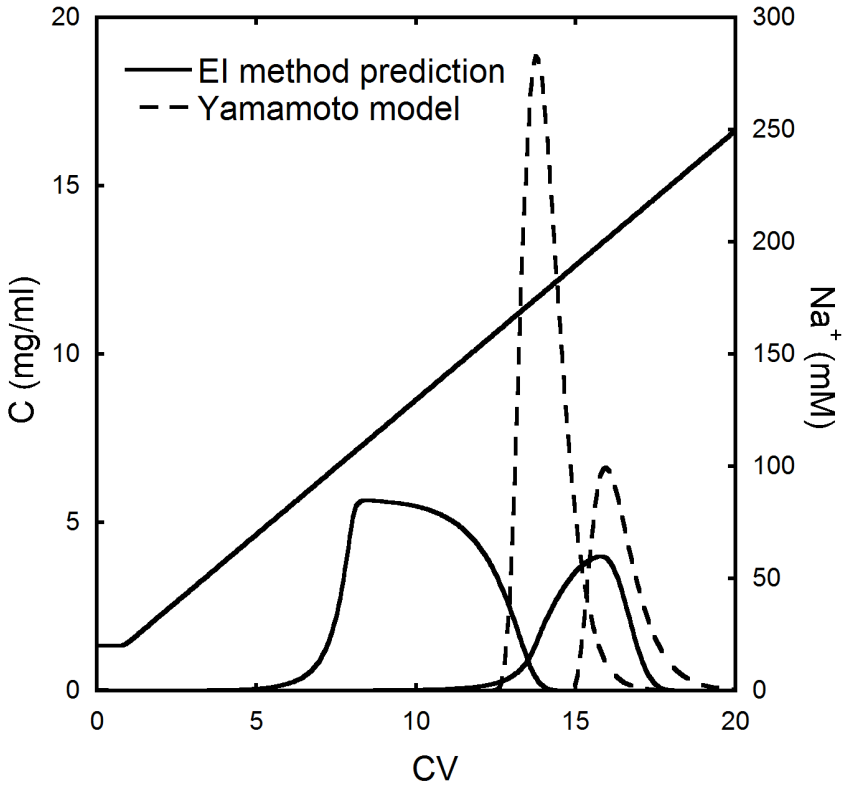


Figure 3.4: Comparison of peaks predicted from the model of Yamamoto[7] (eqs. 8-9) (dashed lines) with the EI method predictions (solid lines) for the conditions of Fig. 3.3F with a protein load of 40 mg/mL.

References

(*1*) Reck, J. M., Pabst, T. M., Hunter, A. K., and Carta, G., (2017). Separation of antibody monomer-dimer mixtures by frontal analysis. *Journal of Chromatography A 1500*, 96–104.

(*2*) Brooks, C. A., and Cramer, S. M., (1992). Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE Journal 38*, 1969–1978.

(*3*) Reck, J. M., Pabst, T. M., Hunter, A. K., Wang, X., and Carta, G., (2015). Adsorption equilibrium and kinetics of monomer-dimer monoclonal antibody mixtures on a cation exchange resin. *Journal of Chromatography A 1402*, 46–59.

(*4*) Gu, T., Tsai, G.-J., and Tsao, G. T., (1991). Multicomponent adsorption and chromatography with uneven saturation capacities. *AIChE Journal 37*, 1333–1340.

(*5*) Carta, G., Ubiera, A. R., and Pabst, T. M., (2005). Protein Mass Transfer Kinetics in Ion Exchange Media: Measurements and Interpretations. *Chemical Engineering & Technology 28*, 1252–1264.

(*6*) Bird, R., Stewart, W., and Lightfood, E., *Transport Phenomena*; John Wiley & Sons: New York, 1960, p 199.

(*7*) Yamamoto, S., (1995). Plate height determination for gradient elution chromatography of proteins. *Biotechnology and Bioengineering 48*, 444–51.

4 Hydrophobic Interaction Chromatography

4.1 Introduction

Protein retention in hydrophobic interaction chromatography is described by the solvophobic theory as a function of the kosmostropic salt concentration. In general, an increase in salt concentration drives protein partitioning to the hydrophobic surface while a decrease reduces it. In some cases, however, protein retention also increases at low salt concentrations resulting in a U-shaped retention factor curve or isotherm. During gradient elution the salt concentration is gradually decreased from a high value thereby reducing the retention factor and increasing the protein chromatographic velocity. For these conditions, a steep gradient can overtake the protein in the column, causing it to rebind. Two rate models, one based on the local equilibrium theory and the other based on the linear driving force (LDF) approximation, are presented. We show that the normalized gradient slope determines whether the protein elutes in the gradient, partially elutes, or is trapped in the column. Experimental results are presented for two different monoclonal antibodies, mAb C and mAb D, and for lysozyme on Capto Phenyl (high sub) resin. mAb D and lysozyme exhibit U-shaped retention factor curves at low loads and for each, we determine the critical gradient slope beyond where 100% recovery is no longer possible. Elution with a reverse gradient from low to high salt is also demonstrated at low protein lows for the proteins with U-shaped retention factor curves.

Another experimental study with mAb D on Capto Phenyl (high sub) is conducted at high protein loads. Batch isotherm data is collected over a range of ammonium sulfate concentrations and described using the empirical interpolation (EI) method. Predictions of protein elution profiles based on the EI approach are compared to experimental LGE data for a range of protein loads from 2% to 30% of the column binding capacity. Understanding this behavior has implications in the design of gradient elution since the protein load and gradient duration impact the recovery of the mAb and the fraction of protein left behind which can eventually foul the column.

As a final component to this work, we will briefly consider the effects of fouling on Capto Phenyl (high sub). After cycling the resin with large loads of protein and multiple cleaning steps, the

adsorption behavior can significantly change as binding sites are blocked. Confocal laser scanning microscopy (CLSM) and a dye designed to conjugate with residual protein shows significant fouling can occur in this resin. Finally, we show resin fouling can lead to significant changes in LGE behavior over time emphasizing the need to develop cleaning methods that fully regenerate the column.

4.2 Theoretical development

4.2.1 Protein binding equilibrium

Melander and Horvath[1] and later Melander et al.[2] showed that the physico-chemical basis of both hydrophobic affinity chromatography at low salt concentrations and HIC at high salt concentrations is explained by the so-called solvophobic theory. Accordingly, the logarithm of the isocratic retention factor at low protein concentrations is expressed as a summation of Gibbs free energy differences between mobile and stationary phases that are associated with cavity formation, electrostatic effects, and van der Waals interactions. Following the treatment of Melander and co-workers, protein retention can be expressed as a function of salt concentration, C_M, by the following equation:

$$\ln(k' - k'_M) = \ln k'_0 - \frac{b\sqrt{C_M}}{1 + c\sqrt{C_M}} + \lambda C_M \tag{4.1}$$

where

$$k' = \phi(\varepsilon_p + K) \tag{4.2}$$

is the retention factor of the protein, $\phi = (1-\varepsilon)/\varepsilon$ is the phase ratio, ε is the extraparticle or external porosity, ε_p is the intraparticle or internal porosity, K is the protein Henry's constant describing protein binding in the dilute limit, and k'_M is the retention factor of the salt. If the salt is not bound, then $k'_M = \phi\varepsilon_p$. Although, as shown by Melander et al.[2], the parameters k'_0, b, c and λ are theoretically related to thermodynamic functions, in practice, they can be treated as empirical constants whose values are determined by data fitting. It should be noted, that the solvophobic theory expresses the salt concentration as a molality. In practice, however, if the salt concentrations

are relatively low, replacing molalities with molarities results in a relatively small error (refer to Fig. A5 in the Appendix).

At high values of C_M, the last term on the right hand side of eq. 4.1 becomes dominant and this equation reduces to:

$$\ln(k' - k'_M) = \ln k'_0 - \frac{b}{c} + \lambda C_M = A + \lambda C_M \tag{4.3}$$

which gives the exponential dependence of protein retention on salt concentration normally observed in HIC. At low C_M values, the third term in eq. 4.1 becomes less important and k' values that decrease as C_M increases are predicted resulting in a U-shaped curve with a minimum k' that depends on the specific values of $\ln k'_0$, b, c, and λ. This possibility was experimentally observed for several proteins by Melander et al.[2] using $NaClO_4$ as the mobile phase modulator in HIC and by Machold et al.[3] for lysozyme on a number of HIC resins using ammonium sulfate as the modulator. In addition to the solvophobic theory, other adsorptive interactions can also lead to U-shaped retention factor curves in HIC. For example, hydrophobic binding and protein unfolding on the resin surface can lead to strong retention at low salt concentrations, weaker interactions at intermediate kosmotrope concentrations and strong binding again at high kosmotrope concentration. Additionally, U-shaped retention factor curves can be a result of the interplay of hydrophobic and electrostatic interactions which has been observed for protein binding to multimodal chromatography media by Melander at al.[4], Nfor et al.[5], Kallberg et al.[6], and Lee et al.[7].

The solvophobic model in eq. 4.1 is limited to the dilute limit of the isotherm where the adsorption behavior is linear with respect to the protein concentration. In order to model adsorption at higher protein concentrations near surface saturation, it is necessary to account for nonlinear adsorption. Langmuirian based models are typically used to describe nonlinear retention in HIC such as classic Langmuir[8], exponentially modified Langmuir[9], and polynomial Langmuir[10,11]. Xia et al. used the quadratic form of the polynomial Langmuir equation coupled with the preferential interaction model to predict adsorption behavior under both linear and nonlinear conditions over a wide range of salt concentrations[12]. Good agreement between theory and experimental data was achieved for lysozyme and lectin on TOYOPEARL Phenyl-650M.

In this work, the EI method is extended to predict nonlinear adsorption in HIC using an arbitrary function to fit single component adsorption data at each ammonium sulfate concentration. The function used in this work is the quadratic Langmuir equation since it has sufficient flexibility in the nonlinear range and can be coupled with the solvophobic model to constrain the function's dilute limit behavior. The resulting model is given as the following:

$$\bar{q} = \frac{K_\infty\left(\alpha C + \delta C^2\right)}{1 + K_\infty\left(\omega C + \tau C^2\right)} + \varepsilon_p C \tag{4.4}$$

where C is the protein concentration in the mobile phase, $K_\infty = \left(k' - k'_M\right)\big/\phi\alpha$ and α, δ, ω, and τ are empirical constants. At low values of C, eq. 4.4 reduces to $\bar{q} = k'C\big/\phi$ which is the dilute limit described by eq. 4.1. Constraining k' in eq. 4.4 to the value calculated in eq. 4.1 is essential in order to obtain a robust prediction of the elution behavior over broad ranges of protein loads and salt concentrations.

To interpolate the effect of the salt concentration, piecewise cubic Hermite interpolating polynomials (PCHIP) are used according to the method given in Section 2.2.2 with additional details in Section 6.3 of the Appendix.

4.2.2 Protein retention in gradient elution

4.2.2.1 Local equilibrium model

Protein retention in gradient elution can be predicted from the isocratic retention factor using the local equilibrium theory for conditions where the binding isotherm is linear and the mass transfer kinetics is fast. The theory is well known[13–16] and only a brief outline is presented here. With a linear gradient, the salt concentration at a particular time, t, and distance x from the column entrance is given by:

$$C_M = C_{M,0} + \beta\left[t - \frac{x}{\upsilon}\left(1 + k'_M\right)\right] \tag{4.5}$$

where $C_{M,0}$ is the initial salt concentration, β is the gradient slope, and υ is the mobile phase interstitial velocity. The gradient slope is given by $\beta = \left(C_{M,f} - C_{M,0}\right)\big/t_G$ where $C_{M,f}$ is the salt

concentration at the end of the gradient and t_G is the duration of the gradient in time units. Taking the differential of eq. 4.5, we obtain:

$$\frac{dC_M}{dx} = \frac{\partial C_M}{\partial t}\frac{dt}{dx} + \frac{\partial C_M}{\partial x} = \beta\frac{dt}{dx} - \frac{\beta}{\upsilon}\left(1 + k'_M\right) \tag{4.6}$$

Movement of the protein through the column is described by the chromatographic velocity

$$\upsilon_c = \frac{dx}{dt} = \frac{\upsilon}{1 + k'\left(C_M\right)} \tag{4.7}$$

where $k'(C_M)$ is the protein retention factor at C_M. Combining eqs. 4.6 and 4.7 yields:

$$\frac{dC_M}{dx} = \frac{\beta}{\upsilon}\left(1 + k'\right) - \frac{\beta}{\upsilon}\left(1 + k'_M\right) = \frac{\beta}{\upsilon}\left(k' - k'_M\right) \tag{4.8}$$

In turn, this equation can be written as:

$$\frac{dC_M}{d\gamma} = k' - k'_M \tag{4.9}$$

where $\gamma = \beta x/\upsilon$ is a normalized gradient slope. The latter can also be expressed in the terms of the duration of the gradient in column volumes units, CV_G, as $\gamma = \varepsilon\left(C_{M,f} - C_{M,0}\right)/CV_G$. Equation 4.9 can be integrated with initial condition $C_M\big|_{\gamma=0} = C_{M,0}$ to determine the salt concentration, $C_{M,R}$ at which the protein elutes. The corresponding elution time is found from eq. 4.5 as:

$$t_R = \frac{C_{M,R} - C_{M,0}}{\beta} + \frac{x}{\upsilon}\left(1 + k'_M\right) \tag{4.10}$$

This equation can be rewritten as:

$$CV_R = \varepsilon\left[\frac{C_{M,R} - C_{M,0}}{\gamma} + \left(1 + k'_M\right)\right] \tag{4.11}$$

where $CV_R = t\varepsilon\upsilon/x$ is the number of column volumes at elution. An analytical solution of eq. 4.9

is easily obtained if k' is described by eq. 4.3 and is given by[14,17]:

$$C_M = -\frac{1}{\lambda}\ln\left[e^{-\lambda C_{M,0}} - A\lambda\gamma\right] \tag{4.12}$$

where $A = k'_0 e^{-b/c}$. Numerical integration of eq. 4.9 is however required in the general case of eq. 4.1.

4.2.2.2 Rate model

The theory described in Section 4.2.2.1 can only describe situations where band broadening is small, which are encountered when small resin particles are used, when the intraparticle diffusivity is high, or when the residence time is high, and the isotherm is linear near dilute conditions. In practice, however, when dealing with larger proteins, band broadening is substantial and the peak elutes over a range of salt concentration rather than at the single value predicted by the local equilibrium model. Nevertheless, even for those conditions the local equilibrium model is expected to describe the general trends with respect to the peak mean retention time provided the isotherm is linear—i.e., in the dilute limit. A model taking into account band broadening is needed however for a more precise description of the elution behavior, especially when conditions are such that only a portion of the injected protein elutes in the gradient.

Two models were used in this work to describe these scenarios – the general rate model, taking into account pore diffusion, and the linear driving force approximation (LDF) model, using a liquid phase concentration driving force. In both cases, axial dispersion and film mass transfer resistance are neglected since intraparticle diffusion is expected to be controlling. The model equations are as follows:

Column differential material balance

$$\varepsilon\frac{\partial C_i}{\partial t} + (1-\varepsilon)\frac{\partial \bar{q}_i}{\partial t} + \varepsilon v\frac{\partial C_i}{\partial x} = 0 \tag{4.13a}$$

$$x = 0 \rightarrow C_i = C_{i,F} \tag{4.13b}$$

Pore diffusion model

$$\frac{\partial q_i}{\partial t} = \frac{D_{e,i}}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial c_i}{\partial r}\right) \tag{4.14a}$$

$$r = 0 \rightarrow \partial c_i \big/ \partial r = 0 \tag{4.14b}$$

$$r = r_p \rightarrow c_i = C_i \tag{4.14c}$$

$$\frac{\partial \bar{q}_i}{\partial t} = -\frac{3D_{e,i}}{r_p}\frac{\partial c_i}{\partial r}\bigg|_{r=r_p} \tag{4.14d}$$

LDF model

$$\frac{\partial \bar{q}_i}{\partial t} = \frac{15D_{e,i}}{r_p^2}\left(C_i - C_i^*\right) \tag{4.15}$$

In these equations, q_i is the local concentration of species i at each point in the particle, $\bar{q}_i$ is the value of q_i averaged over the particle volume, C_i^* is the mobile phase protein concentration in equilibrium with $\bar{q}_i$, $C_{i,F}$ is the feed concentration, $D_{e,i}$ is the effective pore diffusivity, and r_p is the particle radius. For low protein concentrations in the dilute limit, eq. 4.1 is used to calculate C_i^* in eq. 4.15. To model cases with high protein concentrations, the EI method with eq. 4.4 was used to calculate C_i^* in eq. 4.15.

The analytical solution of both rate models in ref.[18] was used to describe the low load isocratic elution behavior yielding nearly indistinguishable results. The LDF model was thus used to describe the LGE runs. In this case, eqs. 4.13 and 4.15 were solved numerically as described in Section 2.2.4. The analytical solution of the LDF model for LGE in ref.[17] was used to corroborate the validity of the numerical solution for cases where complete elution of the peak was predicted and low protein loads were used.

4.3 Materials and methods

4.3.1 Materials

The resin used in this work is Capto Phenyl (high sub), which was obtained from GE Healthcare (Uppsala, Sweden). The particle diameter of the sample used (volume average $d_p = 78$ μm) was determined from a particle size distribution measurement from microphotographs (see Fig. A6 in the Appendix). The intraparticle porosity ($\varepsilon_p = 0.91$) and the mean pore radius ($r_{pore} = 30$ nm) were obtained by inverse size chromatography (iSEC) using dextran standards according to the method in ref.[19] as described in ref.[20] (results shown in Fig. A7).

The proteins used in this work are lysozyme (pI ~11, MW ~15 kDa), obtained from MilliporeSigma (St. Louis, MO, USA), and two monoclonal antibodies, mAb C (pI ~8.2, MW ~150 kDa), and mAb D (pI ~8, MW ~150 kDa). Both mAbs were highly pure with undetectable aggregate levels. Chemicals used in buffer preparation (trisodium phosphate, phosphoric acid, and ammonium sulfate) were obtained from MilliporeSigma (St. Louis, MO, USA) and Fisher Scientific (Fair Lawn, NJ, USA).

All buffers were prepared by mixing trisodium phosphate and ammonium sulfate with distilled-deionized water to yield 30 mM Na^+ and the desired ammonium sulfate molarity. The pH was then adjusted to 7.2 by stirring in drop-wise concentrated phosphoric acid. Lysozyme solutions were prepared by adding weighed amounts of the protein to these buffers. mAb solutions were prepared by size exclusion chromatography with a HiPrep 26/10 desalting column from GE Healthcare (Piscataway, NJ, USA).

4.3.2 Methods

4.3.2.1 Isocratic elution experiments

Isocratic elution experiments were conducted by slurry packing the resin into a 0.5x5 cm Tricorn column from GE Healthcare according to the resin manufacturer packing instructions. Packing quality was verified using pulse injections of 0.8 M NaCl in water with 0.4 M NaCl in water as eluent. The extraparticle or external porosity of this column was determined to be $\varepsilon = 0.43$ based

on the elution of 2,000 kDa dextran. Protein samples (20 μL, containing 11 g protein/L for the mAbs and 5 g protein/L for lysozyme) were injected and eluted isocratically at a flow rate of 0.5 mL/min (153 cm/h) using an Acquity H-Class UPLC system from Waters (Milford, MA, USA) with UV monitoring at 280 nm. This system has low dead volumes providing more accurate results for the smaller column used in these experiments. The ensuing chromatograms were used to determine the protein retention factor from the first moment of the eluted peak after adjustment for the extracolumn volume.

4.3.2.2 Batch adsorption isotherms

Batch adsorption isotherm data at varying protein and ammonium sulfate concentrations were obtained by first equilibrating resin in the appropriate buffer and then removing the extraparticle liquid using 2 mL Corning Costar Spin-X microfiltration tubes (Sigma-Aldrich, St. Louis, MO, USA) with an Eppendorf Minispin bench-top centrifuge (Eppendorf North America, Hauppauge, NY, USA) operated at 5000 RPM for 15 minutes. Samples of the filtered resin (15-300 mg) were then added to 2 mL plastic tubes and mixed with 1 ml of protein solution by slowly rotating the tubes end-over-end on a wheel for 24 h. After this time, supernatant samples were taken to determine the residual protein concentration using a NanoDrop 2000c UV–vis spectrophotometer (Thermo Scientific, Wilmington, DE, USA) at 280 nm. The amount of protein held by the resin was then calculated by mass balance. Finally, the total protein concentration in the resin, $\bar{q}$, was calculated by dividing this amount by the resin sample volume. The latter was calculated from the sample mass using the density of the filtered resin determined with a pycnometer.

4.3.2.3 Gradient elution experiments

All LGE experiments were performed using an AKTA Pure 25 system at a flow rate of 0.5 mL/min (153 cm/h). UV280 and conductivity were recorded and converted to the corresponding protein and ammonium sulfate concentrations using independently obtained calibration curves. Appropriate shifts of UV and conductivity signals were made to take into account the dead volumes between columns and detectors. Because of the larger columns used for these experiments, the larger dead

volumes in this system compared to the UPLC system did not compromise the accuracy of the results. All experiments were conducted at room temperature, 21±1 °C.

Low protein loads

Low load LGE experiments were carried out with a 0.5x20 cm Tricorn column packed to a column bed height of 19.3 cm. The extraparticle porosity of the column $\varepsilon = 0.41$ based on the elution of 2,000 kDa dextran.

Protein samples (100 μL, containing 11 g protein /L for the mAbs and 5 g protein/L for lysozyme) were injected at either 1 M ammonium sulfate for lysozyme or 0.2 M ammonium sulfate for the mAbs and eluted with a linear gradient to 0 M ammonium sulfate followed by a 5 CV hold step for the mAbs and a 3 CV hold step for lysozyme at 0 M ammonium sulfate. The ammonium sulfate concentration at elution was determined from the first moment of the eluted peaks. A clean-in-place step with 0.1 M sodium hydroxide was used to regenerate the column after each low load LGE run.

High protein loads

The high load experiments used two 0.5x20 cm Tricorn columns: virgin resin packed to a height of 18.8 cm and cycled resin packed to a height of 19 cm. The virgin column was newly packed and was only exposed to protein used for the experiments listed below. The cycled resin for this work had been used in multiple LGE experiments and exposed to more than 30 mg/ml of mAb D per ml of column before its use in the experimental runs shown in this work.

High load LGE experiments with mAb D used injection volumes between 1.1 and 5 mL loaded using a 50 mL superloop corresponding to total protein loadings ranging from 0.3 to 5.3 mg/mL of column. mAb D was loaded in 0.2 M ammonium sulfate and eluted with a linear gradient to 0 M ammonium sulfate followed by a 5 CV hold step at 0 M ammonium sulfate.

The virgin column was stripped with 70% isopropyl alcohol over 2 CVs followed by a clean-in-place step with 0.1 M sodium hydroxide to regenerate the column after each high load LGE run. The cycled column was only cleaned with 0.1 M sodium hydroxide after each high load LGE run.

4.3.2.4 Confocal laser scanning microscopy

Confocal laser scanning microscopy (CLSM) was used to test for the presence of residual protein in cycled resin. The cycled resin for this experiment was obtained from Bristol-Myers Squibb (Hopewell, NJ) after it was saturated with mAb D at 450 mM ammonium sulfate, eluted with 0 mM ammonium sulfate step gradient, cleaned with water, 0.5 M NaOH, and 70% isopropyl alcohol. Both virgin and cycled resin were conjugated with Dylight 488 Lightning-Link from Innova Biosciences (Babraham, Cambridge, UK) in Na_2HPO_4 pH 7 according to supplier protocol. Resin was then washed and re-suspended in 10 mM Na_2HPO_4 pH 7 buffer. Imaging was performed with a Zeiss LSM 510 microscope with a Plan-Apochromat 63×/1.4 NA oil objective (Carl Zeiss MicroImaging, LLC, Thornwood, NY, USA). Confocal images were taken using an excitation wavelength of 488 nm and measuring emission at 524 nm.

4.4 Results and discussion

4.4.1 Isocratic elution behavior

Figure 4.1 shows the isocratic elution behavior of the three proteins used in this work. As seen in Fig. 4.1A, mAb C could be eluted isocratically at different ammonium sulfate concentrations from 0.2 M to 0, with essentially 100% recovery in each case. This protein exhibited the normally expected behavior of decreasing retention with decreasing ammonium sulfate concentrations. On the other hand, both lysozyme and mAb D exhibited a minimum retention at an intermediate ammonium sulfate concentration and neither protein could be eluted at 0 M ammonium sulfate. The lysozyme peaks are considerably sharper than those obtained from mAb C as a result of the smaller size and larger diffusivity. mAb C is more strongly retained than lysozyme at the higher ammonium sulfate concentration suggesting a greater hydrophobic character.

Figure 4.2 shows the retention factors calculated from the isocratic elution runs of Fig. 4.1 as a function of the ammonium sulfate concentration. While k' for mAb C increases monotonically with C_M, both lysozyme and mAb D exhibit U-shaped retention factor curves. This behavior is

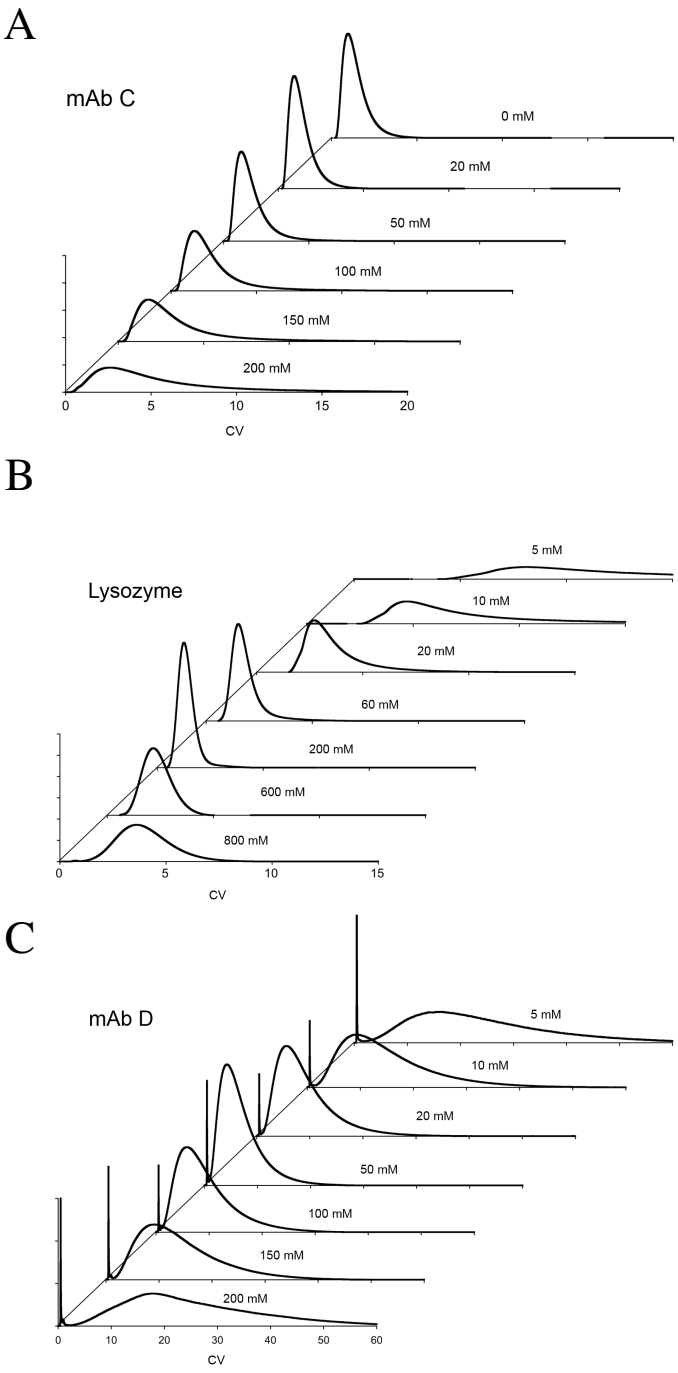


Figure 4.1: Isocratic elution peaks obtained for (A) mAb C, (B) lysozyme, and (C) mAb D at the ammonium sulfate concentrations indicated for each curve.

especially pronounced for mAb D, which exhibits a strong minimum at about 0.06 M ammonium sulfate. The lines in this figure are based on the solvophobic model, eq. 4.1, with the parameters $\ln k'_0$, b, c, and λ obtained by least squares fit using KaleidaGraph v. 4.5 (Synergy Software) and summarized in Table 4.1. k'_M was set equal to 1.21, based on the iSEC data.

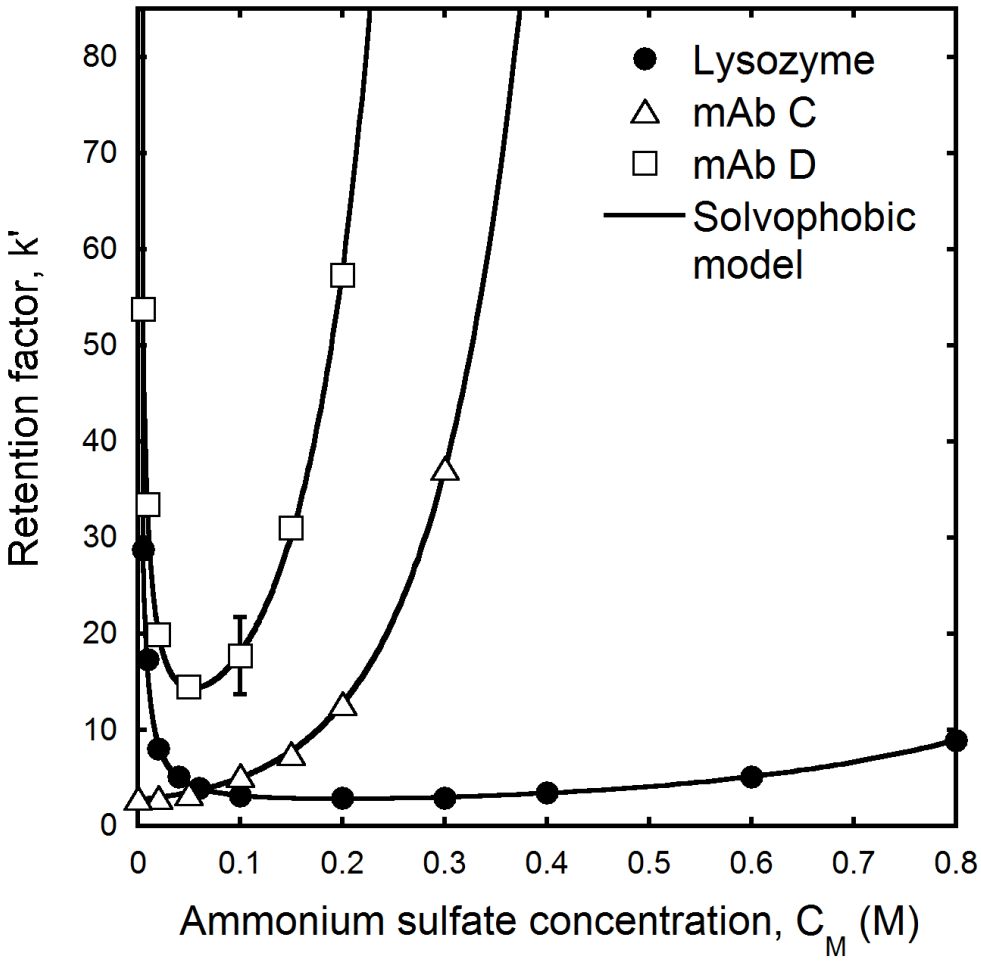


Figure 4.2: Retention factors obtained from the data in Fig. 4.1 for mAb C, Lysozyme, and mAb D as a function of the ammonium sulfate concentration. All solutions contained 10 mM trisodium phosphate adjusted to pH 7.2 with phosphoric acid. The lines are based on the solvophobic model, eq. 4.1, with parameters in Table 4.1.

Table 4.1: Parameters of the solvophobic model (eq. 4.1) obtained from the isocratic retention factor for the proteins used in this work with C_M in mol/L and $\varepsilon = 0.43$.

Parameter	**mAb C**	**Lysozyme**	**mAb D**
$\ln\left(k_0'\right)$	0.566 ± 0.118	5.53 ± 0.26	6.17 ± 0.24
b	1.95 ± 1.98	42.0 ± 7.14	45.7 ± 7.4
c	1.35 ± 5.30	4.74 ± 1.12	5.74 ± 1.20
λ	12.1 ± 1.6	4.64 ± 0.59	17.9 ± 1.0

4.4.2 High load batch isotherm behavior

Figure 4.3 shows the batch isotherm data for mAB D on Capto Phenyl HS with ammonium sulfate concentrations C_M in the range 200 to 0 mM and protein concentrations C in the range 0 to 4 mg/ml which cover the expected conditions in LGE experiments at high protein loads. The solid lines in 4.3A and 4.3B are calculated with the quadratic Langmuir model eq. 4.4 with the parameters given in Table 4.2 fitted individually at each ammonium sulfate concentration C_M. The dashed lines in Fig. 4.3B show the curves generated by PCHIP in the EI method as described in Section 2.2.2. These lines run perpendicular to the fitted quadratic Langmuir curves and describe the effect of the ammonium sulfate concentration.

The isotherm capacity is approximately 30 mg/ml of particle at 200 mM C_M and decreases as the salt concentration decreases to 50 mM. At 0 mM C_M, the binding capacity dramatically increases again analogous to the trend seen in the retention factor curve at low protein loads (Fig. 4.2). This behavior results in a U-shaped isotherm most clearly seen in Fig. 4.3B. The agreement between fitted lines used in the EI method and experimental data is excellent with mean absolute error of less than 3%.

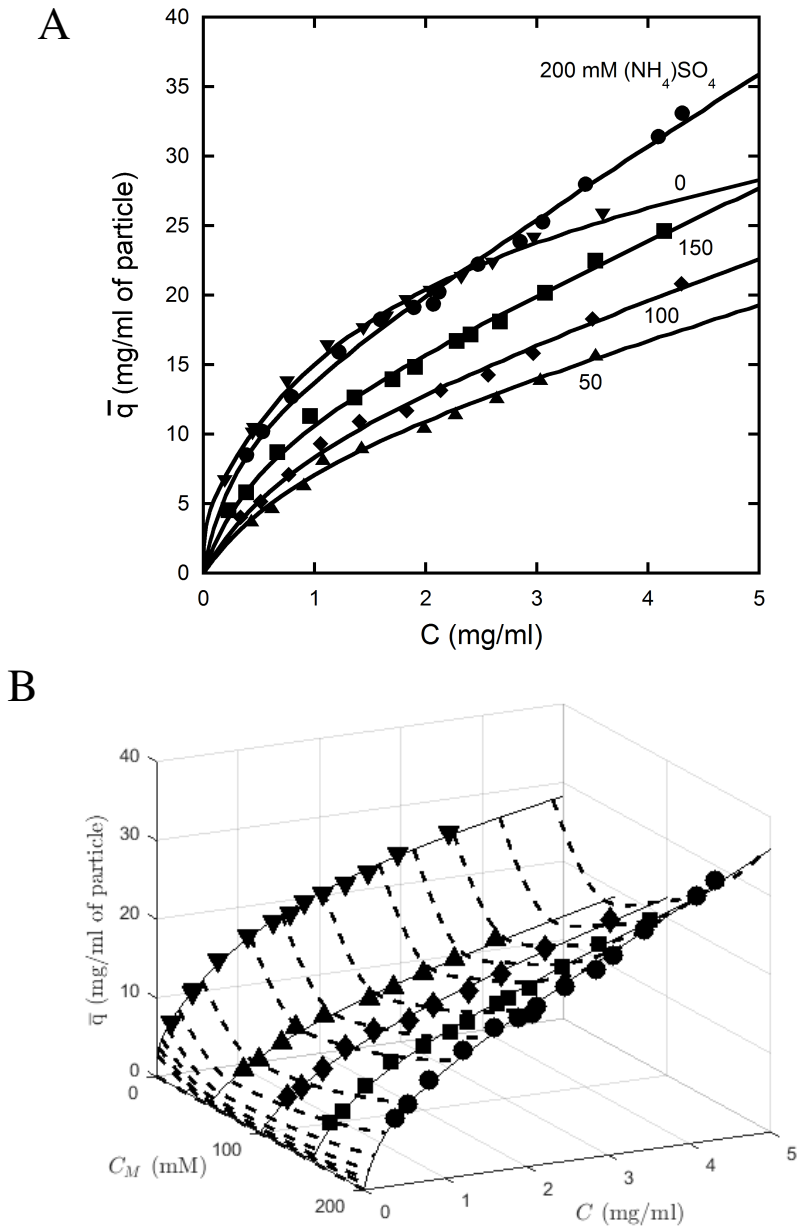


Figure 4.3: Batch adsorption isotherms for mAb D on Capto Phenyl HS. Panel A shows the 2-dimensional view of the isotherms and panel B shows the 3-dimensional view with C_M as a third axis. The solid lines are calculated with the quadratic Langmuir model with the parameters given in Table 4.2 fitted individually at each ammonium sulfate concentration C_M. The dotted lines are the PCHIP curves used to interpolate the effect of C_M. In (A) and (B) symbols ●, ■, ◆, ▲, and ▼ represent 200, 150, 100, 50, and 0 mM ammonium sulfate concentrations, respectively.

Table 4.2: Quadratic Langmuir equation parameters fitted to isotherm data in Figure 4.3

C_M	α	ω	τ	δ
200	0.551	0.044	0.000	0.185
150	0.806	0.064	0.000	0.170
100	1.343	0.103	0.000	0.178
50	1.548	0.129	0.000	0.153
0	0.352	0.081	0.051	1.518

4.4.3 Gradient elution behavior at low loads

Figures 4.4-4.6 show the LGE results for the three proteins used in this work. As seen in these figures different behaviors are exhibited by the three different proteins as a function of the gradient slope. For mAb C (Fig. 4.4) elution is obtained with essentially 100% recovery even with very sharp gradients. The elution peak becomes sharper and elutes at lower ammonium sulfate concentrations as the gradient slope increases (or the gradient volume, CV_G, decreases). Even with a step change to 0 M ammonium sulfate, essentially complete elution is obtained with the elution peak emerging from the column immediately after the ammonium sulfate step.

For lysozyme (Fig. 4.5), the elution behavior is analogous to that of mAb C for the shallower gradient slopes. However, a different result is obtained with a step change to 0 M ammonium sulfate, with the peak eluting isocratically only many CVs after the ammonium sulfate step. The elution peak is very shallow in this case as a result of the long retention in the column. Finally, for mAb D (Fig. 4.6), complete elution occurs only with relatively shallow gradients. Gradient with duration of 10 CV or less resulted in incomplete elution, with much of the protein remaining trapped in the column at the end of the gradient and even during the ensuing hold step. Almost no elution is obtained with a 2 CV gradient and no elution at all was seen even after 16 CVs using a step change to 0 M ammonium sulfate.

The behaviors exhibited by the three proteins appear to be qualitatively consistent with the corresponding k' trends observed isocratically. The normal elution behavior of mAb C is a consequence of the monotonically increasing k' values. In this case, lowering the ammonium sulfate concentration always results in reduced retention and, thus, increased chromatographic velocity. As a result, the peak always elutes within the gradient. For the other two proteins, but especially for mAb D, because of the U-shaped retention factor plot, the chromatographic velocity increases during the initial portion of the gradient allowing the protein to move down the column but decreases as the protein is exposed to the final portion of the salt gradient. Whether the protein reaches the column outlet or not depends on the gradient slope. If the gradient slope is too high (10 CVs or less in the case of mAb D), lower ammonium sulfate concentrations overtake the protein in the column

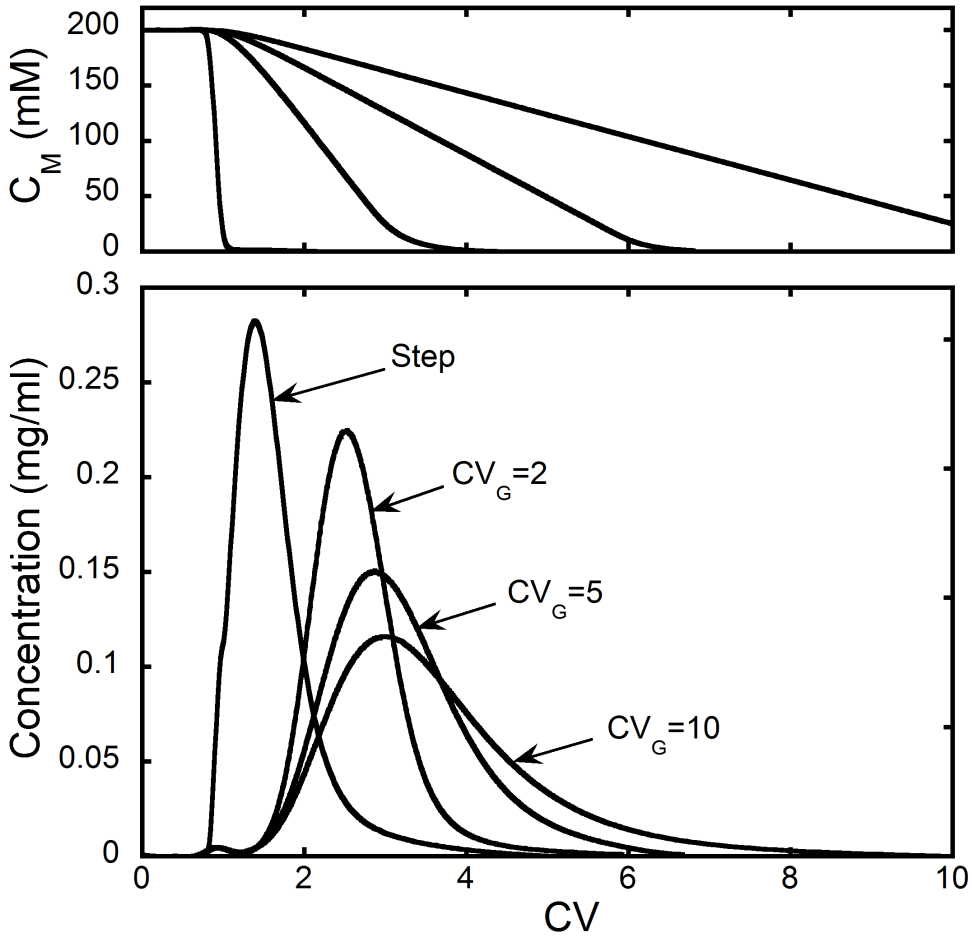


Figure 4.4: LGE results for mAb C with linear gradients from 0.2 to 0 M ammonium sulfate. CV_G is the duration of the gradient in column volumes units. The gradient profiles shown in the top panel are obtained from the conductivity signal using a calibration curve.

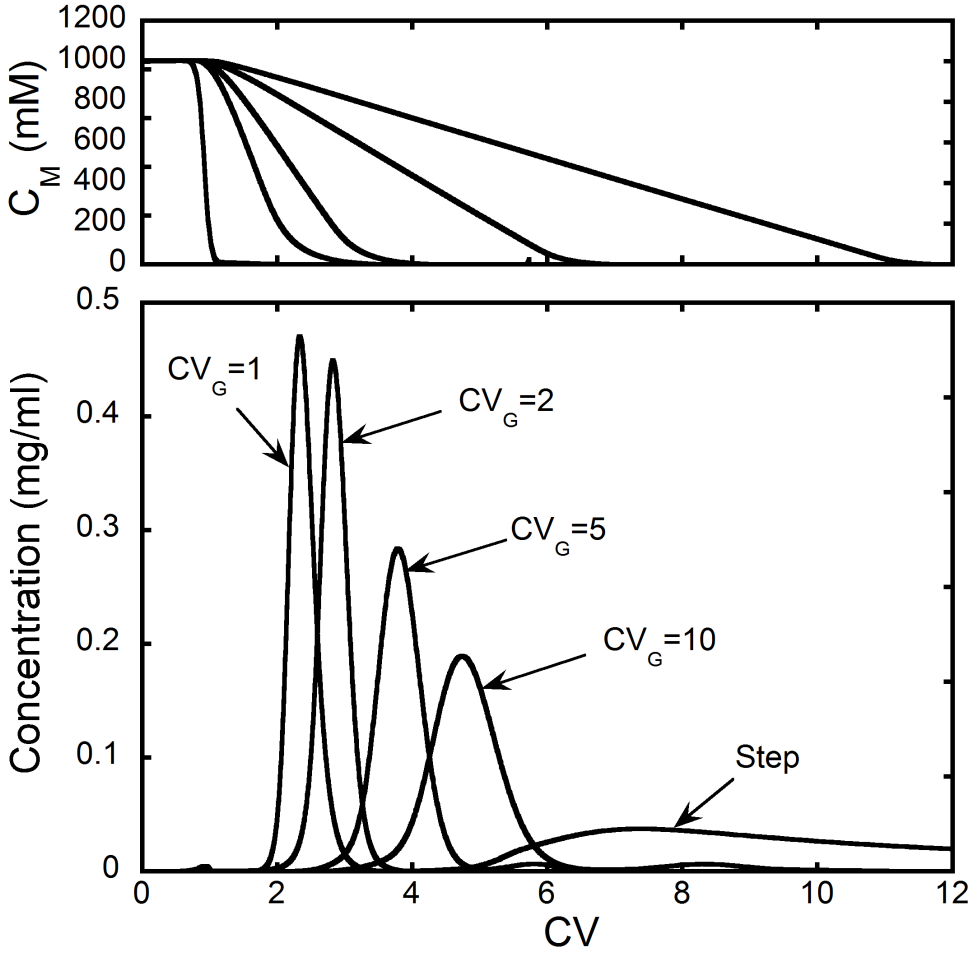


Figure 4.5: LGE results for lysozyme with linear gradients from 1 to 0 M ammonium sulfate. CV_G is the duration of the gradient in column volumes units. The gradient profiles shown in the top panel are obtained from the conductivity signal using a calibration curve.

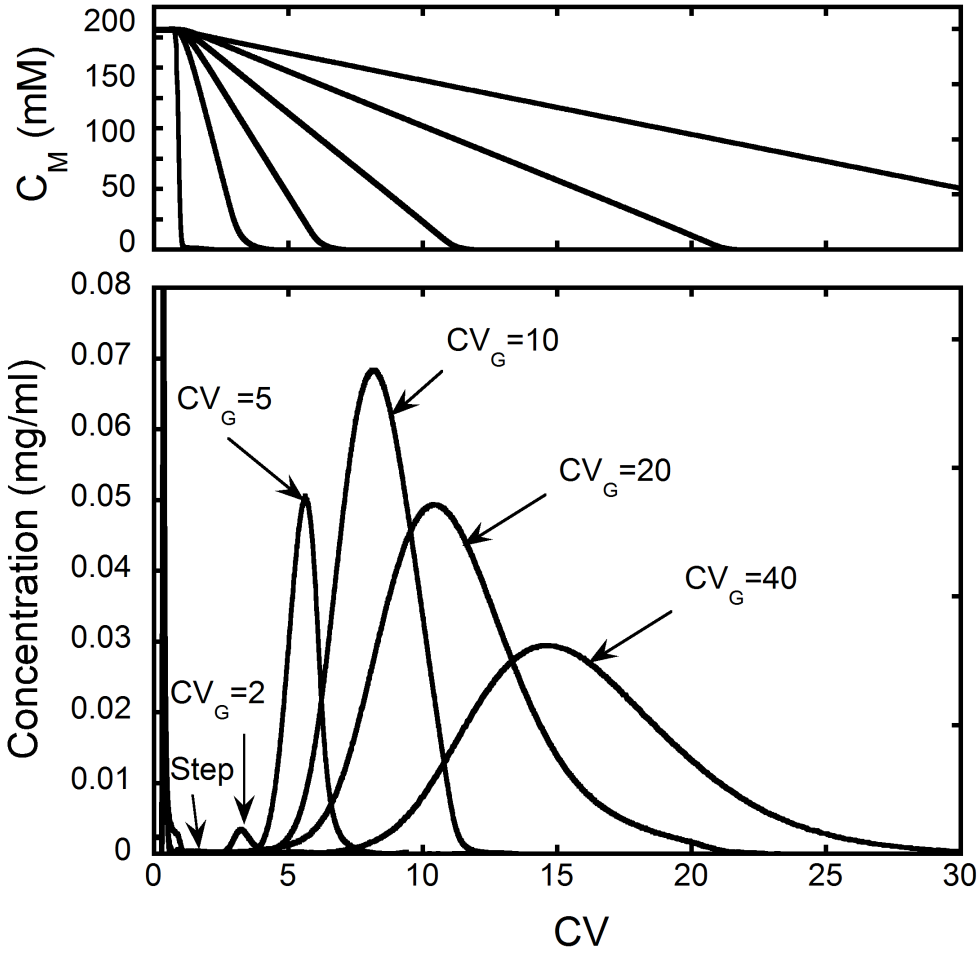


Figure 4.6: LGE results for mAb D with linear gradients from 0.2 to 0 M ammonium sulfate. CV_G is the duration of the gradient in column volumes units. The gradient profiles shown in the top panel are obtained from the conductivity signal using a calibration curve.

increasing retention and causing the chromatographic velocity to decrease. At this point the protein becomes trapped in the column in the down-sloping region of the retention factor plot resulting in low or no recovery even for long CVs. The behavior is less striking for lysozyme since the retention factor plot is much broader with a less pronounced minimum. Even so, the extremely sharp gradient generated by a step change to 0 M ammonium sulfate at the column entrance results in a strong retention of the protein in the column.

4.4.4 Prediction of low load LGE behavior from isocratic data

4.4.4.1 Local equilibrium model results

Figure 4.7 shows the ammonium sulfate concentrations at peak elution, $C_{M,R}$, vs. the normalized gradient slope, $-\gamma$, for the three different proteins calculated according to eqs. 4.1 and 4.9-4.11 using the parameters in Table 4.1. The corresponding protein chromatographic velocity normalized by the modulator chromatographic velocity, $v_c/v_{c,M} = \left(1+k'_M\right)/\left(1+k'\right)$, is also shown in these figures. v_c is the rate at which the protein band moves through the column at each $C_{M,R}$-value. Qualitatively similar behaviors are observed for all three proteins at low values of $-\gamma$ (i.e. for shallow gradients) and experimental values of $C_{M,R}$ (shown only for the LGE runs that gave essentially 100% recovery) are in good agreement with the model predictions. On the other hand, distinctly different behaviors are predicted for the three proteins at high values of $-\gamma$ (i.e. for steep gradients). For mAb C, as $-\gamma$ increases the protein chromatographic velocity increases monotonically, reaching a high value close to the chromatographic velocity of the ammonium sulfate modulator as $C_{M,R}$ approaches zero. Thus, the model predicts that mAb C will elute within the gradient or shortly after the gradient regardless of gradient slope. On the other hand, for both lysozyme (Fig. 4.7B) and mAb D (Fig. 4.7C), the chromatographic velocity decreases sharply at high values of $-\gamma$ as a result of the U-shaped retention factor curve. In the case of mAb D, in particular, the chromatographic velocity is close to zero at $C_{M,R} = 0$ suggesting that, for a given initial modulator concentration, a critical gradient slope exists beyond which this protein becomes trapped in the column and does not elute over reasonable time scales. From Fig. 4.7C, the critical value of $-\gamma$ is about 0.009 M, which

corresponds to a gradient duration CV_G ~9. Indeed, as seen in Fig. 4.6, gradient durations smaller than 10 CVs resulted in decreasing recovery with little or no elution for very steep gradients.

4.4.4.2 Rate model results

Detailed predictions based on the rate model were made only for mAb D, which exhibited the greatest deviations from the normally expected behavior. This model requires the protein effective diffusivity, D_e, which was obtained by matching the experimental isocratic elution curves to the analytical solution. A value of $D_e = (1.0 \pm 0.2) \times 10^{-7}$ cm^2/s provided a good fit for all isocratic elution runs. Figure 4.8 shows representative examples at 0.05 and 0.2 M ammonium sulfate. As seen from this figure there is little difference between the pore diffusion model and the LDF model.

Figure 4.9 shows numerical predictions of the elution curves for the experiments of Fig. 4.6 using the LDF model. Comparing Figs. 4.6 and 4.9 shows that model predictions are in reasonable agreement with the experimental data. In particular, the numerical model predicts incomplete elution with gradient lengths of 10 CVs or less and virtually no elution for a 2 CV gradient or for a step change to 0 M ammonium sulfate. Note that for 5 and 10 CV gradients, the numerical model predicts extremely slow elution of the protein trapped in the column at the end of the gradient, which is not observed in the experiment. This small discrepancy is likely due to the difficulty of describing retention at 0 M ammonium sulfate. Since, experimentally, the protein did not elute isocratically at 0 M ammonium sulfate, the retention factor at $C_M = 0$ is extrapolated from the fit of the solvophobic model and is thus affected by considerable uncertainty. A higher value of $\ln k'_0$ around 6.9 would eliminate this discrepancy.

Figure 4.10 compares predicted and experimental protein recoveries as a function of gradient duration. In both cases, the percentage recovery was calculated from the eluted peak area. As seen in this figure, the rate model is in excellent agreement with the data predicting incomplete elution at $CV_G \leq 10$ and essentially no recovery at $CV_G \leq 2$.

4.4.4.3 Reverse gradients

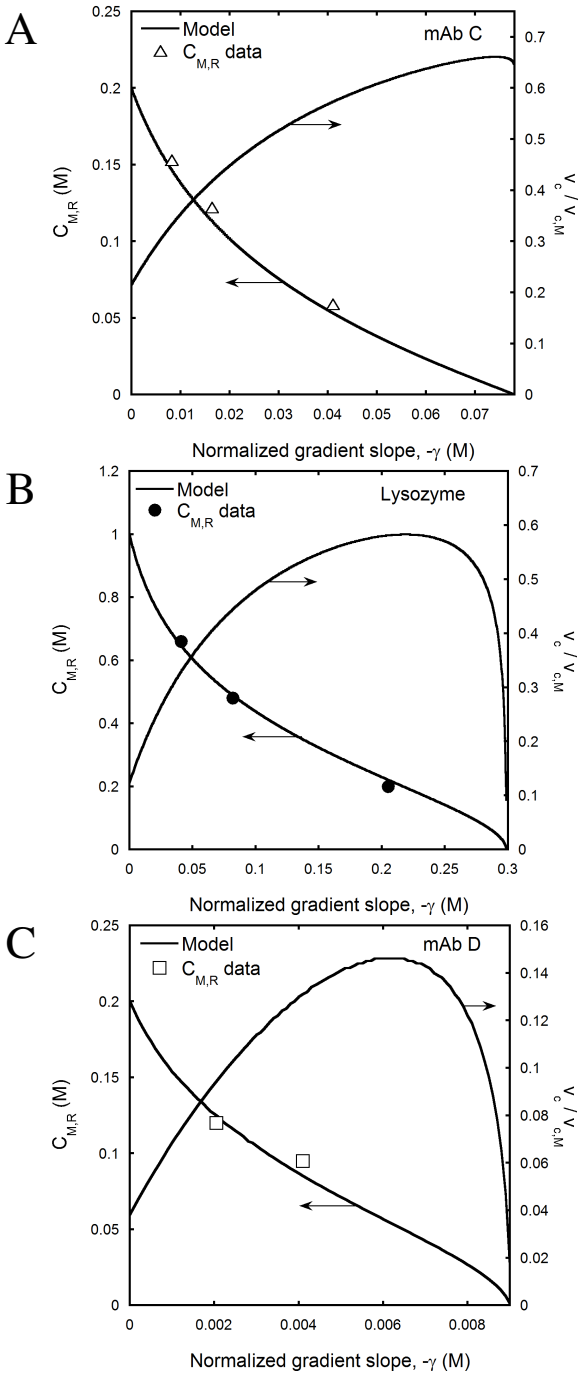


Figure 4.7: Plots of the ammonium sulfate concentration at elution, $C_{M,R}$, vs. the normalized gradient slope, $-\gamma$, for (A) mAb C, (B) lysozyme, and (C) mAb D based on the local equilibrium model. The vertical right-hand side axis gives the protein chromatographic velocity v_c divided by the chromatographic velocity of the modulator, $v_{c,M}$. Conditions are the same as those in Figs. 4.4-4.6. Experimental $C_{M,R}$ values based on the ammonium sulfate concentration at the first moment of the protein peak are shown only for conditions where protein recovery was essentially 100%.

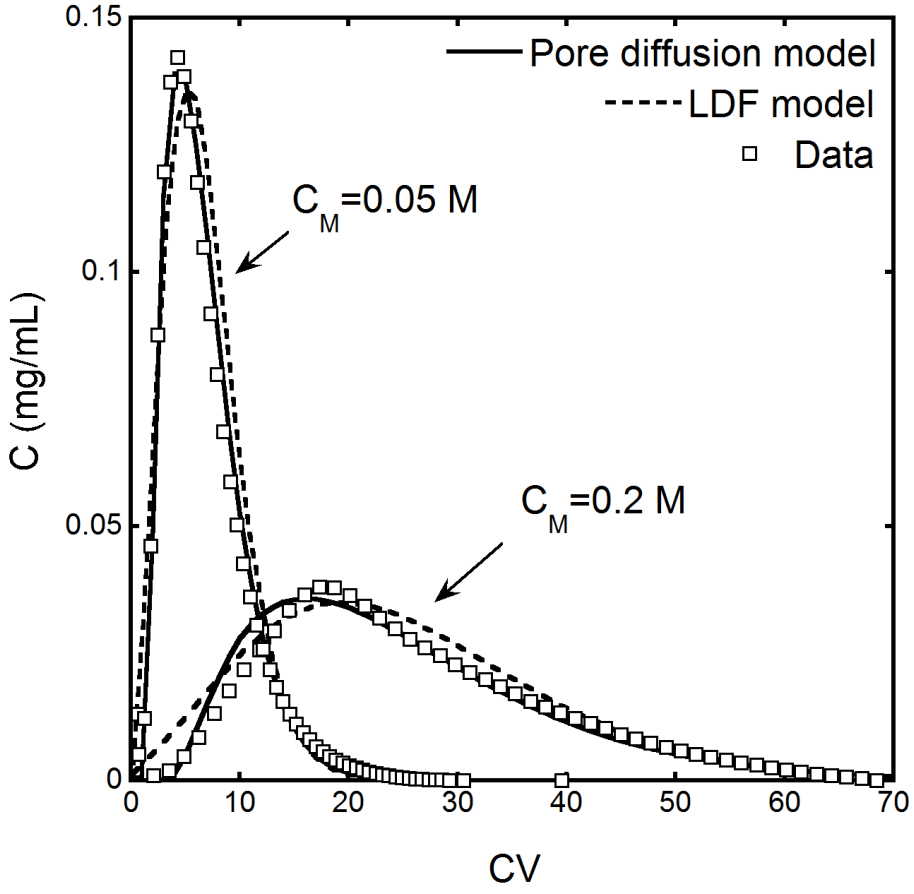


Figure 4.8: Comparison of experimental isocratic elution curves of mAb D at 0.05 and 0.2 M ammonium sulfate concentrations with curves calculated according to both the pore diffusion model and the LDF model using a fitted value of $D_e = (1.0 \pm 0.2) \times 10^{-7}$ cm^2/s. The curves are calculated with the analytical solution of eqs. 4.13-4.15 in ref.[18]. Experimental conditions are the same as in Fig. 4.1C.

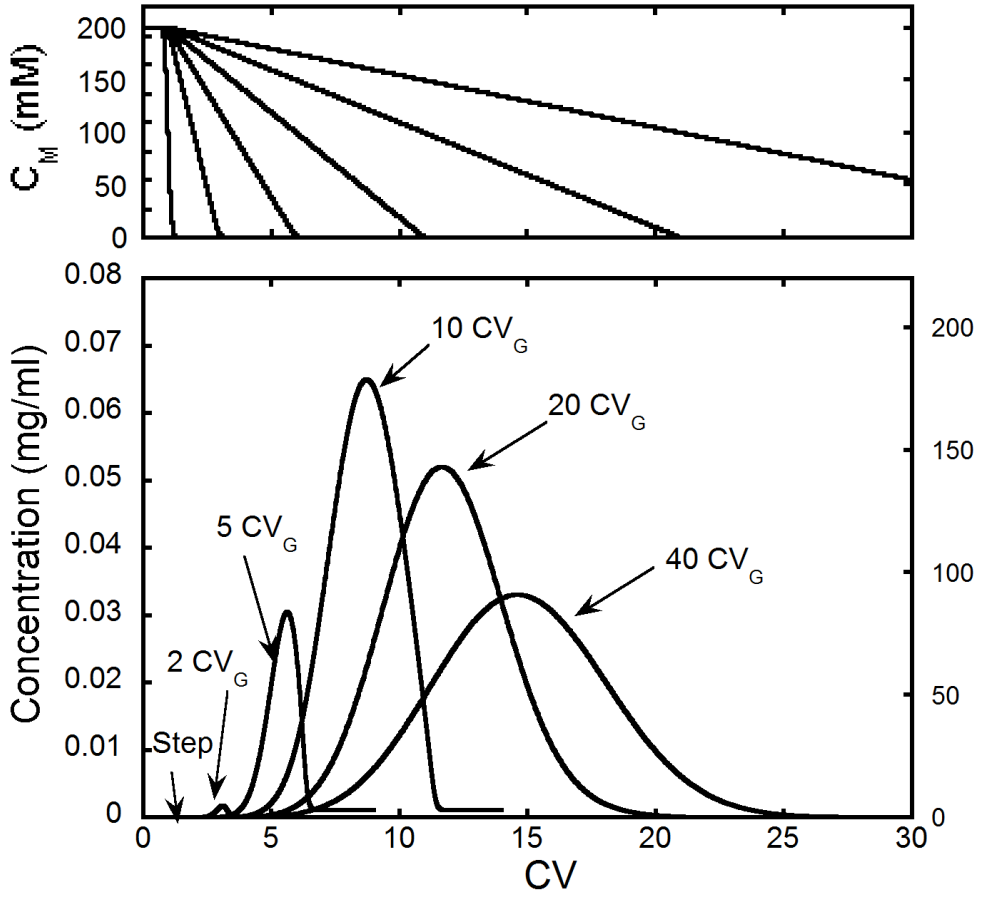


Figure 4.9: Predicted elution profiles for mAb D with linear gradients from 0.2 to 0 M ammonium sulfate using the LDF rate model with a D_e value of 1×10^{-7} cm^2/s for the protein and a retention factor described by eq. 4.1 with parameters in Table 4.1 (bottom panel). CV_G is the duration of the gradient in column volumes units. The top panel shows the gradient profiles at the column outlet predicted using a D_e value of 1×10^{-5} cm^2/s for ammonium sulfate. Conditions simulated are the same as those in Fig. 4.6.

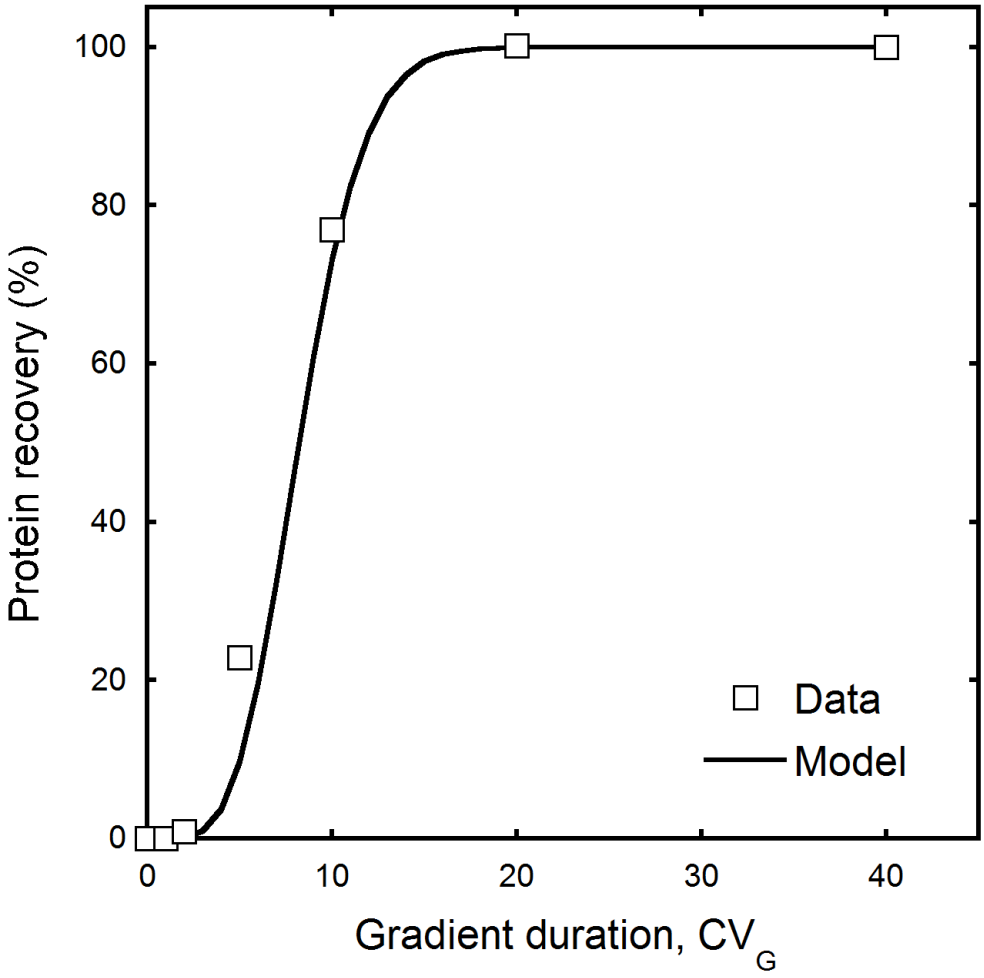


Figure 4.10: Comparison of experimental and predicted protein recovery for mAb D with linear gradients from 0.2 to 0 M ammonium sulfate. Model and experimental conditions are the same as those in Figs. 4.6 and 4.9.

It is worth considering whether, given the U-shape retention plots, lysozyme and mAb D can be eluted with a "reverse gradient" (i.e. positive gradient) after loading in 0 M ammonium sulfate. Figure 4.11A and 4.11B show the results for lysozyme and mAb D, respectively, the former using either gradient or a step from 0 to 1 M ammonium sulfate and the latter using a gradient or step from 0 to 0.2 M ammonium sulfate. In both cases, the protein elutes essentially completely during the gradient but elutes isocratically only many CVs after the step change in ammonium sulfate concentration has passed through the column. The results are consistent with the U-shaped retention plots. For lysozyme with a positive 2 CV gradient from 0 to 1 M ammonium sulfate, eq. 4.1 and 4.9 predict $C_{M,R} = 0.45$ M vs. the experimental value of $C_{M,R} = 0.33$ M, while for mAb D with a positive 20 CV gradient from 0 to 0.2 M ammonium sulfate, these equations predict $C_{M,R} = 0.075$ M vs. the experimental value of $C_{M,R} = 0.068$ M. In both cases, the local equilibrium model predicts late elution for the extremely steep gradient ensuing from the corresponding step changes.

4.4.5 Prediction of high load LGE behavior from batch isotherm data

Figure 4.12 compares the experimental protein elution curves of mAb D obtained on the virgin Capto Phenyl (high sub) column over a range of protein loads from 0.3 to 5.3 mg/ml of column (dashed lines) with predictions based on the EI method (solid lines) coupled with the LDF rate model and using $D_e = (1.0 \pm 0.2) \times 10^{-7}$ cm^2/s determined from isocratic elution curve fits. The binding capacity of mAb D at 200 mM ammounium sulfate is about 30 mg/ml and the extraparticle porosity is 0.41, the column binding capacity is about 18 mg/ml. Thus, these protein loads are between 1% and 30% of the column binding capacity. It is clear that as the protein load increases, the peak begins to elute at higher salt concentrations compared to the low loading case allowing for more protein to be recovered. At the highest load of 5.3 mg/ml elution begins near the start of the salt gradient as a result of mass overload and significant dispersion from mass transfer resistance. The asymmetrical shape of these peaks is a result of the nonlinear adsorption behavior seen in Fig. 4.3. Figure 4.14A shows the recovery for each load compared to the predicted values using the EI method. Overall there is good agreement between the data and model showing an increase in recovery as the protein load increases.

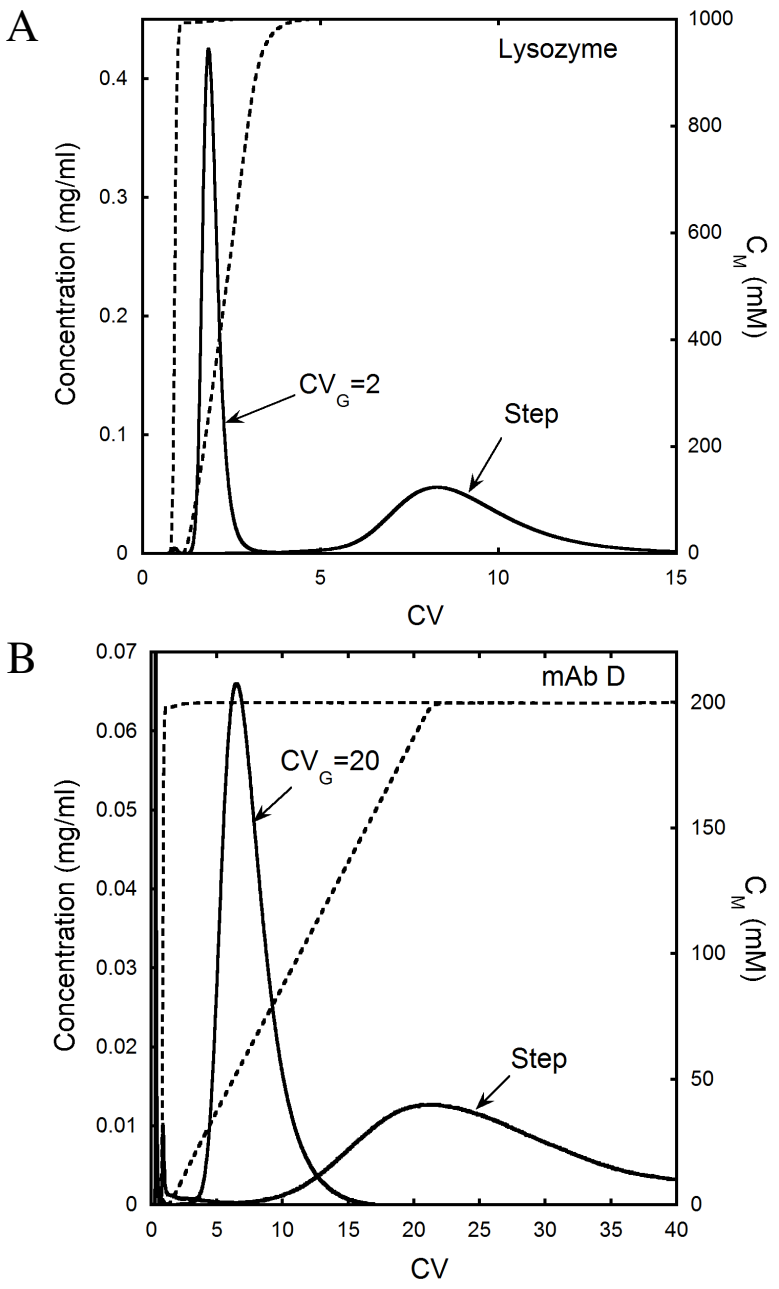


Figure 4.11: Elution results for (A) lysozyme eluted with a 2 CV gradient or a step gradient from 0 to 1 M ammonium sulfate and (B) mAb D eluted with a 20 CV gradient or a step gradient from 0 to 0.2 M ammonium sulfate, both following loading in 0 M ammonium sulfate. Solid lines show the protein concentration based on the UV280 signal and dashed lines the ammonium sulfate concentration based on the conductivity signal. Essentially 100% protein recovery is obtained with the 2 CV gradient for lysozyme and with the 20 CV gradient for mAb D.

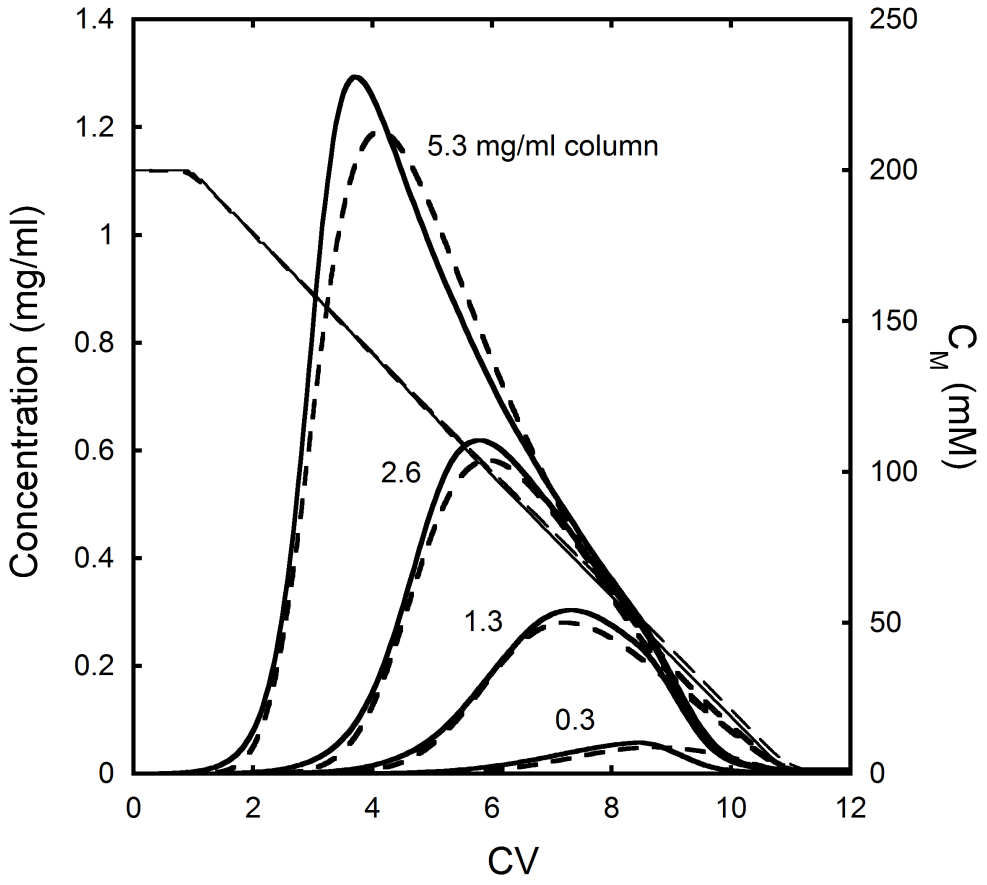


Figure 4.12: Comparison of the experimental elution profiles obtained for LGE with varying protein loads (dashed lines) with model predictions (solid lines) based on the EI method for mAb D on Capto Phenyl HS with a linear gradient from 0.2 to 0 M ammonium sulfate in 10 CV. The gradient profile is obtained from the conductivity signal using a calibration curve.

Figure 4.13 shows the effects of varying the gradient slope with a 5.3 mg/ml load over the range from a 20 CV_G to a step gradient. Excellent agreement between predictions based on the EI method and the experimental results is seen. As the gradient becomes steeper from 20 CV_G to 5 CV_G, the protein peak becomes shaper and elutes at lower salt concentrations in the gradient. However, when a step gradient is applied the peak elutes isocratically over many CVs. Even for this condition, the EI method accurately predicts the elution profile. Figure 4.14B shows the recovery obtained for each gradient length with a constant load of 5.3 mg/ml. Similar to the case illustrated in Fig. 4.10, the recovery decreases as gradient duration decreases. However, in the case of high protein loads, a relatively high recovery >80% can still be achieved with steep gradients since the protein breaks through much earlier compared to the low protein load case. The asymmetrical shape of the peaks at high loads and gradient durations >0 CV_G results in a large fraction of protein never exposed to low salt concentrations where there is strong binding.

In all the cases illustrated in Figs. 4.12 and 4.13, model predictions based on the EI method are in excellent agreement with the experimental profiles. The most significant deviations are seen in the tailing of the peaks at 50 mM C_M. This is likely a result of the insufficiently fine granularity of the batch isotherm data used by the EI method to interpolate from 50 mM to 0 mM C_M. Fig. 4.3B shows PCHIP interpolates a sharp transition in binding capacity from 50 mM to 0 mM. Collecting isotherm data between 0 mM and 50 mM would likely improve the model's description of this region.

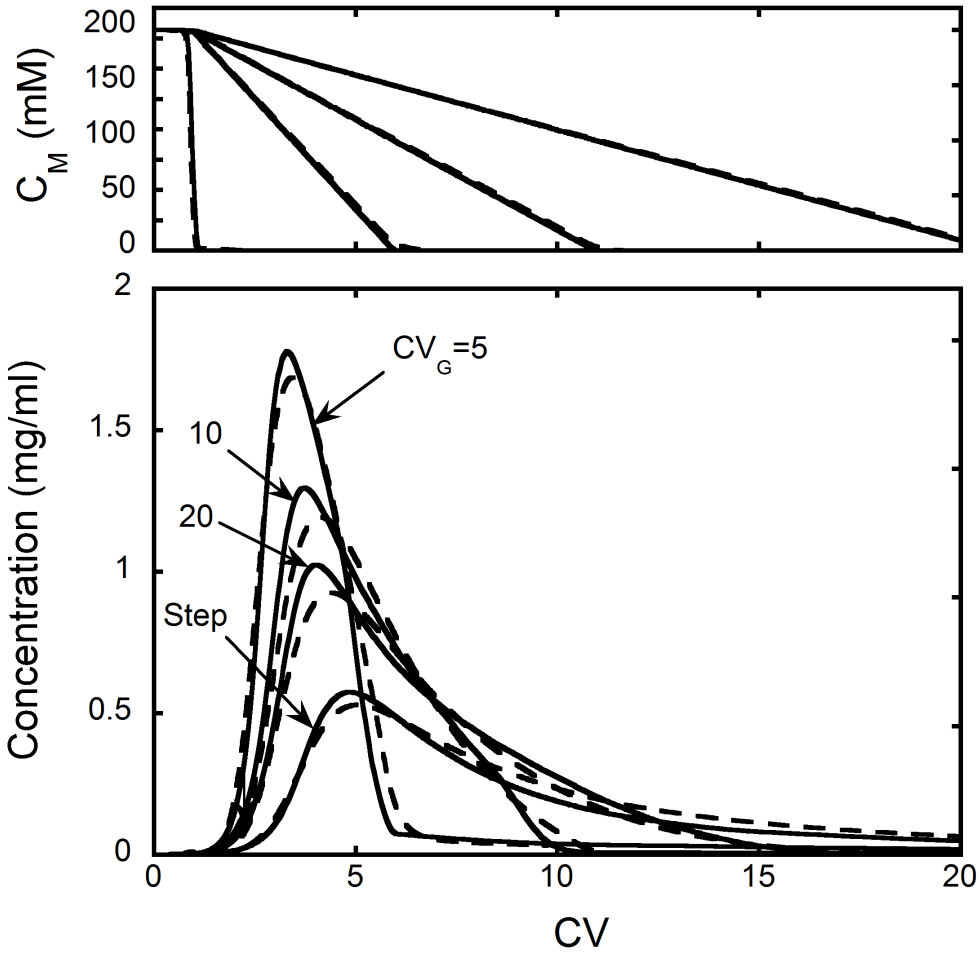


Figure 4.13: Comparison of the experimental elution profiles obtained for LGE with varying gradient durations (dashed lines) with model predictions (solid lines) based on the EI method for mAb D on Capto Phenyl HS with loads of 5.3 mg/ml and linear gradients from 0.2 to 0 M ammonium sulfate. CV_G is the duration of the gradient in column volumes units. The top panel shows the experimental gradient profiles (dashed) and predicted profiles (solid) using a D_e value of 1×10^{-5} cm^2/s for ammonium sulfate.

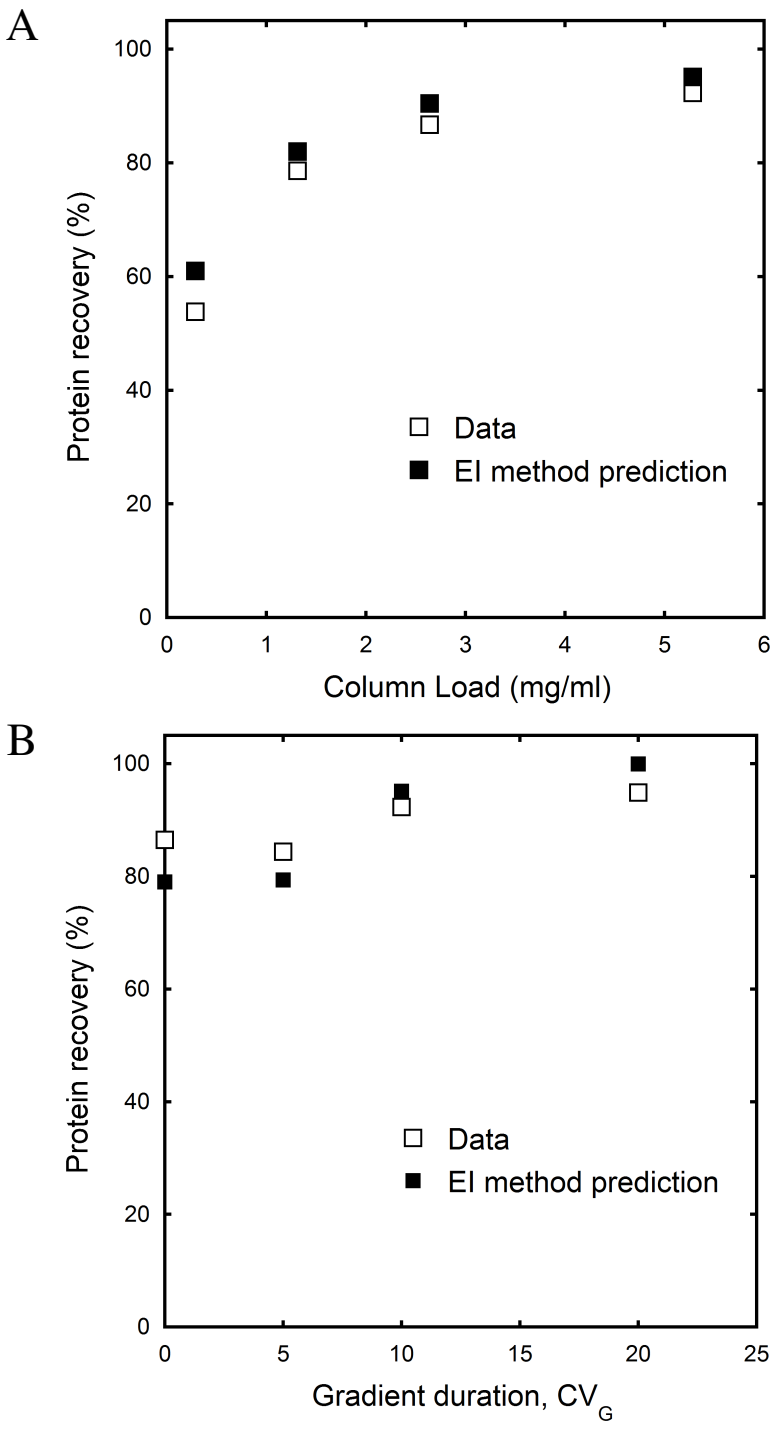


Figure 4.14: Comparison of experimental and predicted protein recovery for mAb D with varying loads (A) and varying gradient durations (B). Model and experimental conditions are the same as those in Figs. 4.12 and 4.13.

4.4.6 Fouling behavior in cycled resin

As a final component to this work, we consider the possibility of fouling that may occur when protein is left behind in the column and becomes irreversibly bound and resistant to removal by cleaning steps. For this study, we used cycled Capto Phenyl (high sub) resin obtained from Bristol-Myers Squibb after the resin had been cycled with mAb D according to the method given in Section 4.3.2.4. To test for fouling in this batch of cycled Capto Phenyl, a reactive fluorescence dye, Dylight 488, was used to label any residual mAb D present in the resin bead. Figure 4.15 shows representative CLSM images of virgin resin (A) and cycled resin (B) both conjugated with Dylight 488. The fluorescence is much greater in the cycled resin compared to the virgin resin suggesting there is residual protein present. This is likely mAb D which was trapped in the column and retained at low ammonium sulfate concentrations where binding is favorable (Figs. 4.2 and 4.3). This result also suggests the cleaning steps used between experimental runs were insufficient to regenerate the column.

Based on the results of the CLSM data in Fig. 4.15, it is worth considering how the LGE behavior may change as the resin is cycled with multiple LGE elutions of mAb D. For this study, a column packed with Capto Phenyl HS had been cycled with multiple LGE experiments using mAb D and exposed to a cumulative loading greater than 30 mg/ml of protein per ml of column. Figure 4.16 shows the comparison of elution profiles obtained on the cycled column (solid lines) vs. the virgin column (dashed lines). Compared to the virgin column, the protein peaks from the cycled column elute earlier in the gradient at higher salt concentrations indicating the binding is weaker on the cycled column.

The results in Fig. 4.16 strongly suggest cycling Capto Phenyl HS with mAb D changes the adsorption behavior. The CLSM data in Fig. 4.15 indict this could be a result of protein irreversibly binding to the resin. Capto Phenyl HS may have a heterogeneous ligand surface and the high protein loads saturate the strongest binding sites that are responsible for the stronger retention at low ammonium sulfate. Cleaning with NaOH could hydrolyze the protein retained on the surface making it bind more strongly. Regardless of the mechanism, it is clear the adsorption behavior changes

after cycling the resin and the elution profiles deviate significantly from the predictions in Fig. 4.12 which were generated using batch isotherm data with virgin Capto Phenyl HS. These results highlight the need to develop effective elution and cleaning steps that properly regenerate the column so that predictions and process performance remain robust over many cycles.

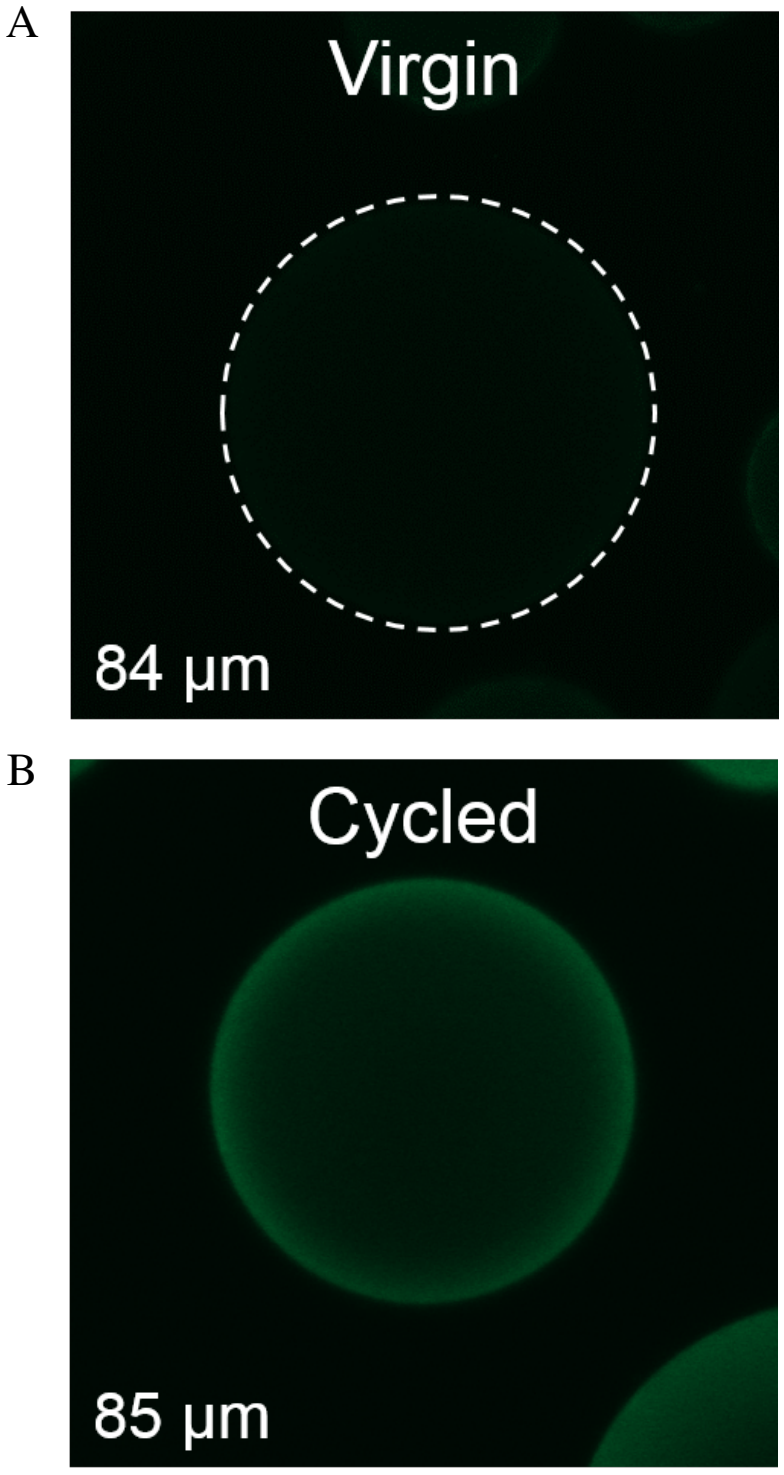


Figure 4.15: Representative CLSM images of virgin (A) and cycled (B) Capto Phenyl (high sub) resin conjugated with Dylight 488 on similarly sized resin particles.

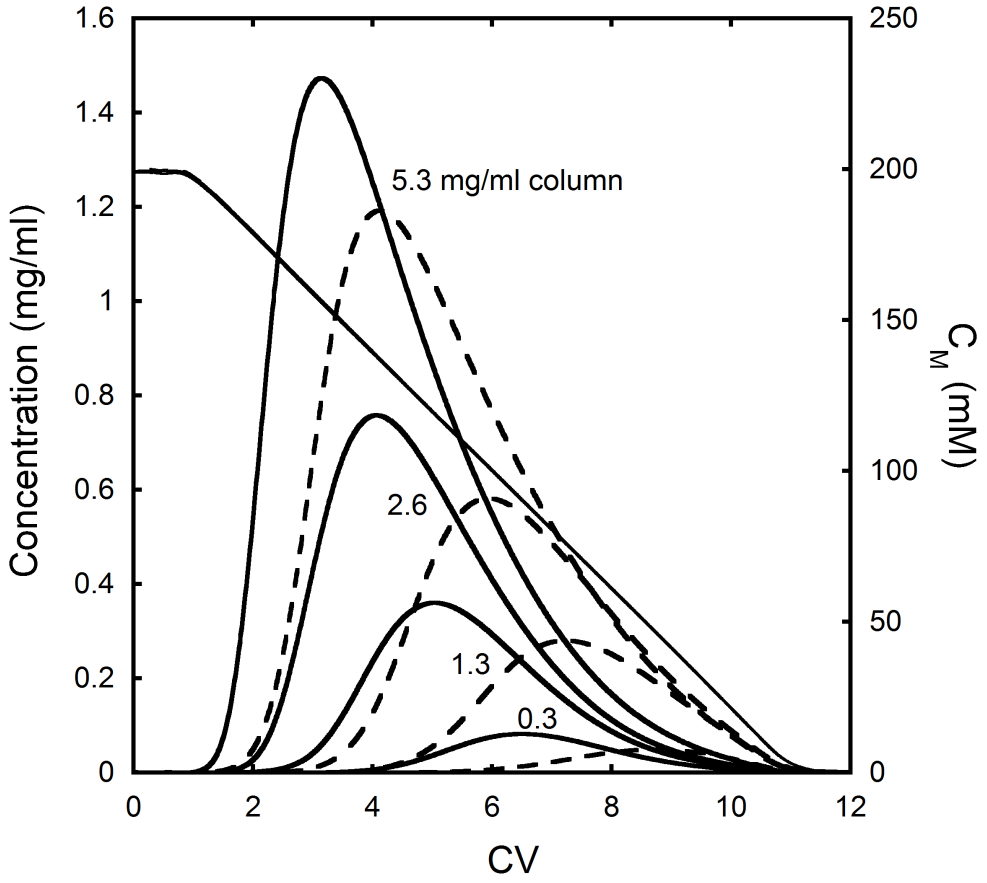


Figure 4.16: Comparison of mAb D LGE behavior on virgin Capto Phenyl (dashed lines) and cycled Capto Phenyl (solid lines) with varying protein loads. The linear gradient is from 0.2 to 0 M ammonium sulfate in 10 CV. Experimental data from the virgin column are the same as those in Fig. 4.12.

References

(*1*) Melander, W. R., and Horvath, C., (1977). Salt effects on hydrophobic interactions in precipitation and chromatography of proteins: An interpretation of the lyotropic series. *Archives of Biochemistry and Biophysics 183*, 200–215.

(*2*) Melander, W. R., Corradini, D., and Horvath, C., (1984). Salt-Mediated Retention of Proteins in Hydrophobic-Interaction Chromatography. *Journal of Chromatography 317*, 67–85.

(*3*) Machold, C., Deinhofer, K., Hahn, R., and Jungbauer, A., (2002). Hydrophobic interaction chromatography of proteins I . Comparison of selectivity. *Journal of Chromatography A 972*, 3–19.

(*4*) Melander, W. R., Rassi, Z. E., and Horvath, C., (1989). Interplay of Hydrophobic and Electrostatic Interactions in Biopolymer Chromatography: Effect of Salts on the Retention of Proteins. *Journal of Chromatography 469*, 3–27.

(*5*) Nfor, B. K., Noverraz, M., Chilamkurthi, S., Verhaert, P. D. E. M., van der Wielen, L. A. M., and Ottens, M., (2010). High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. *Journal of Chromatography A 1217*, 6829–50.

(*6*) Kallberg, K., Johansson, H.-O., and Bulow, L., (2012). Multimodal chromatography : An efficient tool in downstream processing of proteins. *Biotechnology Journal 7*, 1485–1495.

(*7*) Lee, Y. F., Graalfs, H., and Frech, C., (2016). Thermodynamic modeling of protein retention in mixed-mode chromatography: An extended model for isocratic and dual gradient elution chromatography. *Journal of Chromatography A 1464*, 87–101.

(*8*) Meng, Q., Wang, J., Ma, G., and Su, Z., (2013). Isotherm Type Shift of Hydrophobic Interaction Adsorption and its Effect on Chromatographic Behavior. *Journal of Chromatographic Science 51*, 173–180.

(*9*) Antia, F. D., and Horvath, C., (1989). Gradient elution in non-linear preparative liquid chromatography. *Journal of Chromatography 484*, 1–27.

(*10*) Guiochon, G., Golshan-Shirazi, S., and Katti, A., *Fundamentals of Preparative and Nonlinear Chromatography*; Academic Press: Boston, MA, 1994.

(*11*) Mollerup, J. M., (2008). A Review of the Thermodynamics of Protein Association to Ligands, Protein Adsorption, and Adsorption Isotherms. *Chemical Engineering & Technology 31*, 864–874.

(*12*) Xia, F., Nagrath, D., and Cramer, S. M., (2003). Modeling of adsorption in hydrophobic interaction chromatography systems using a preferential interaction quadratic isotherm. *Journal of Chromatography A 989*, 47–54.

(*13*) Carta, G., and Jungbauer, A., *Protein Chromatography: Process Development and Scale-Up*, 1st ed.; Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 2010.

(*14*) Snyder, L. R., *Gradient Elution*; Horvath, C., Ed.; Academic Press: New York, 1980; Vol. 1, pp 207–316.

(*15*) Yamamoto, S., Nakanishi, K., Matsuno, R., and Kamikubo, T., (1983). Ion exchange chromatography of proteins-prediction of elution curves and operating conditions. 1. Theoretical considerations. *Biotechnology and Bioengineering 25*, 1465–83.

(*16*) Yamamoto, S., Nomura, M., and Sano, Y., (1987). Resolution of proteins in linear gradient elution ion-exchange and hydrophobic interaction chromatography. *Journal of Chromatography A 409*, 101–110.

(*17*) Carta, G., and Stringfield, W. B., (1992). Analytic solution for volume-overloaded chromatography. *Journal of Chromatography 605*, 151–159.

(*18*) Carta, G., (1988). Exact analytic solution of a mathematical model for chromatographic operations. *Chemical Engineering Science 43*, 2877–2883.

(*19*) Hagel, L., Ostberg, M., and Andersson, T., (1996). Apparent pore size distribution of chromatography media. *Journal of Chromatography A 743*, 33–42.

(*20*) Stone, M. C., and Carta, G., (2007). Protein adsorption and transport in agarose and dextran-grafted agarose media for ion exchange chromatography. *Journal of Chromatography A 1146*, 202–215.

5 Concluding Remarks and Recommendations

5.1 Conclusions

This dissertation presents a methodology to predict protein elution behaviors for a wide range of chromatography modes and elution methods using column models coupled with high throughput screening (HTS) data. This work does away with mechanistic isotherm models and explores a universal approach that uses batch isotherm data directly through a suitable interpolation scheme. The interpolated isotherm data are coupled with a lumped kinetic model using rate parameters determined from adsorption kinetic measurements to numerically predict the column behavior for both individual or combined pH and salt gradients. Experimental results obtained from laboratory scale columns show excellent agreement with the predicted elution curves for several proteins on ion exchange and hydrophobic interaction media. The EI method could predict elution behavior for both single and multi-component cases and protein loads up to 50% of the column capacity. This method provides a practical tool for rational process development and defining a robust operating space. More specific conclusions and recommendations for future research are addressed below.

5.1.1 Salt and pH gradient elution on ion exchange

The goal of this work was to predict the chromatographic elution behavior of proteins for highly overloaded conditions based on independent measurements of protein adsorption isotherm without relying on a mechanistic model of adsorption equilibrium. The work in this chapter successfully demonstrates the EI method to predict protein elution using either salt gradients at constant pH, pH gradients at constant Na^+ concentration, or simultaneous pH and salt gradients. The method accomplishes such predictions by (a) predicting buffer compositions that will yield linear pH gradients; (b) predicting protein binding by interpolating batch isotherm data with varying protein concentration, Na^+ concentration, and pH; (c) predicting protein binding kinetics based on a linear driving force approximation combined with adsorption kinetic measurements; and (d) predicting the evolution of concentration profiles in the column with a plug flow model.

For the systems presented in Section 2, lysozyme on SP-Sepharose-FF and two mAbs on POROS XS, the empirical interpolation scheme yields predictions of the overloaded gradient elution behavior that are in excellent agreement with experimental data for salt gradient elution, pH gradient elution, and combined pH-salt gradient elution. In the case of salt gradient elution, the agreement between experiments and predictions with this method is far superior to that obtained using the SMA model as a mechanistic description of protein binding. For the lysozyme and mAb A systems, the SMA model was able to describe qualitatively the trends of protein binding with respect to protein and salt concentration with only 4 parameters. However, large deviations between experimental and predicted elution profiles were seen with this model. The EI approach requires a much larger number of parameters, but the additional degrees of freedom remove limitations from model approximations and allow predictions whose accuracy is based entirely on the accuracy of the experimental batch data set.

The results of a qualitative study on the level of detail needed for the experimental isotherm data set show that the interpolation model predictions are relatively insensitive to the granularity of the adsorption equilibrium data. As expected, the data need to cover the ranges of pH and protein and salt concentrations at which elution actually takes place. In particular, the batch isotherm data must cover the range of salt concentrations and pH where conditions change from strong binding, with non-linear binding isotherm behavior, to weak binding with a linear isotherm. The potential of the data-driven prediction method is best achieved when coupled with the power of HTS and liquid handling robots.

Recommendation for future work would include applying the EI method to multimodal chromatography (MMC). Multimodal ligands combine aliphatic or aromatic groups with amino, carboxyl or sulfonic groups. As a result, protein binding can occur through simultaneous hydrophobic, electrostatic, and hydrogen bonding interactions mimicking biological affinity. Thus, it is often required to simultaneously change the salt concentration and pH to elute protein hence making the EI method ideally suited for simulating MMC behavior.

5.1.2 Multicomponent separations on ion exchange

This work shows the EI method can be extended to multicomponent systems to successfully predict highly loaded elutions of an antibody monomer and dimer mixture using a salt gradient. The lumped kinetic model using effective pore diffusivities determined from van Deemter curves can accurately describe the adsorption kinetics for both components. The SMA model does not appear to have the flexibility needed to describe the isotherm behavior and is thus unable to quantitatively predict the multicomponent column behavior. Conversely, the EI method using an empirically modified Langmuir equation, with parameters optimized at each salt concentrations can accurately describe the batch isotherm data and predict multicomponent column behavior using interpolated parameter values at intermediate salt concentrations. Since a mechanistic description of the effects of the mobile phase composition is not needed, the EI method can be easily extended to other chromatographic modalities and to more than two components provided that an empirical model can be found to describe the competitive binding. Regardless of which isotherm model is used, the model does not need to include the effect of the salt concentration, or alternatively pH, since these variables are described by PCHIP. As more dimensions are interpolated, obviously, more data are required to describe the effect of mobile phase composition and competitive binding from loading to elution during column operation. The large data requirement for predicting multicomponent column behavior can be aided by modern high-throughput technology and analytical methods to screen systems with more than two components and varying pH.

5.1.3 Hydrophobic interaction chromatography

U-shaped retention factor curves observed for proteins in HIC with a highly hydrophobic stationary phase are shown to result in quantitatively different elution behaviors dependent on the gradient slope, when the initial and final modulator concentration bracket the retention factor minimum. Starting with a high ammonium sulfate concentration, the protein can completely elute in the gradient, be partially eluted, or the elution peak can disappear altogether as the gradient slope is increased. Normal elution peaks are obtained for shallow gradients while distorted peak with

reduced protein recovery are obtained when the gradient slope is increased above a certain critical value. The interplay of protein and modulator chromatographic velocities and the temporal variation of the modulator concentration at the column entrance is responsible for this behavior. A local equilibrium model can be used to explain the qualitative trends at low protein loads and based on measurements of the isocratic retention factor while a rate model accounting for band broadening is shown to be capable of quantitative predictions under a broad range of conditions. Furthermore, we have shown that proteins that exhibit a U-shaped retention factor plot can be loaded at low ammonium sulfate concentrations and eluted at a higher ammonium sulfate concentration provided that a sufficiently shallow gradient is applied to the column. We have also shown higher protein loads can improve recovery as more protein elutes earlier in the gradient as a result of asymmetrical peaks from nonlinear adsorption. The EI method using batch isotherm data at high protein concentrations quantitatively predicts the elution profiles under nonlinear adsorption conditions for a range of loads and gradient lengths. From a practical perspective, while U-shaped retention factor curves and U-shaped isotherms have been observed previously, our theoretical development provides the means to understand their impact of the dynamic of the chromatography column and helps explain the phenomenon of disappearing elution peaks observed in HIC with steep gradients. The related models can be used to select gradients that avoid undesirable behaviors by manipulating either the gradient slope or the final modulator concentration or both.

Since the focus of this work is on the column dynamics that arise from U-shaped retention behavior, the molecular basis for retention in the absence of ammonium sulfate was not directly addressed. Nevertheless, we can advance some hypotheses. One possibility is that this phenomenon is correlated with the effects of ammonium sulfate on protein solubility. Although the Hofmeister series is typically used to classify salts with respect to their ability to desolvate proteins and promote their binding in HIC, the actual salting-out properties can vary dependent on the salt concentration range. For example, as shown by Green[1], while increasing ammonium sulfate concentrations above 0.5 M reduce the solubility of carboxyhemoglobin, increasing the salt concentration in a range below about 0.5 M actually increase its solubility. It is possible that the reversed salting-out properties of ammonium sulfate in the low concentration range facilitate elution of our proteins with an in-

creasing ammonium sulfate gradient. Another, related possibility is that the protein unfolds on the hydrophobic surface in absence of the stabilizing effects of ammonium sulfate, resulting in strong retention. In this case, adding ammonium sulfate in low concentration could allow the protein to refold causing it to elute. Further exploration of mAb D's biophysical properties could provide some additional insight into a mechanistic explanation of it's hydrophobic retention behavior. Future work should address the effects of pH, salt type and concentration on protein solubility and the retention on resins with different hydrophobicity.

References

(*1*) Green, A. A., (1932). Studies in the physical chemistry of the proteins. X. The solubility of hemoglobin in solutions of chlorides and sulfates of varying concentration. *Journal of Biological Chemistry 95*, 47–66.

6 Appendix

6.1 Supporting tables and figures

Table A1: Summary of resin properties

Property	SP Sepharose FF	POROS XS
Mean particle diameter (µm)[(a)]	100	50
Column extraparticle porosity, ε [(b)]	0.29	0.35
Intraparticle porosity, ε_p [(c)]	0.87	0.61
Lysosyme distribution coefficient, K_D [(d)]	0.83	-
mAb A distribution coefficient, K_D [(d)]	-	0.62
mAb B distribution coefficient, K_D [(d)]	-	0.71
Hydrated particle density	1.02	1.03

(a) Nominal particle diameter based on resin manufacturer data
(b) Averaged value based on pressure drop with Karman-Cozeny equation
(c) Averaged value based on retention of NaCl
(d) Based on protein retention under non-binding conditions

Table A2: SMA model parameters fitted to isotherm data in Figures A3A and A3B using the z- and A-values determined from the LGE experiments at low protein loads

	Lysozyme/SP-Sepharose FF	**mAb A/POROS XS**
z	5.17	10.1
A (mM)^z	3.51×10^{13}	2.92×10^{23}
q_0 (mM)	270	217
K_e	9.53	0.736
σ	12.9	115

Table A3: Buffer compositions for individual and combined pH and salt gradients calculated from pH model

pH	Acetate (mM)	MES (mM)	Phosphate (mM)	[NaCl] (mM)	[Na^+] (mM)
6.0	6.1	8.4	27.8	0	40
6.0	27.3	20.9	3.6	200	240
5.5	13.0	11.3	25.3	60.0	100
7.0	0.0	0.0	25.1	60.0	100
5.5	41.8	19.6	0.0	0.0	40
7.0	0.0	27.2	9.8	100	140

Table A4: Langmuir-Freundlich isotherm parameters regressed at each pH and Na^+ concentration for mAb B on POROS XS

pH	[Na^+] (mM)	q_m (mg/ml)	K (ml/mg)	b
5.5	20	189	7.00×10^5	0.3
5.5	95	127	5.9	0.6
5.5	145	64.8	0.20	1.0
6.0	20	216	1.06×10^6	0.3
6.0	70	632	0.0	0.2
6.0	95	101	0.5	0.8
6.5	20	230	1.7	1.0
6.5	45	146	4.5	0.6
6.5	70	96.4	0.3	1.0
7.0	20	312	0.5	0.3
7.0	45	110	0.2	1.0

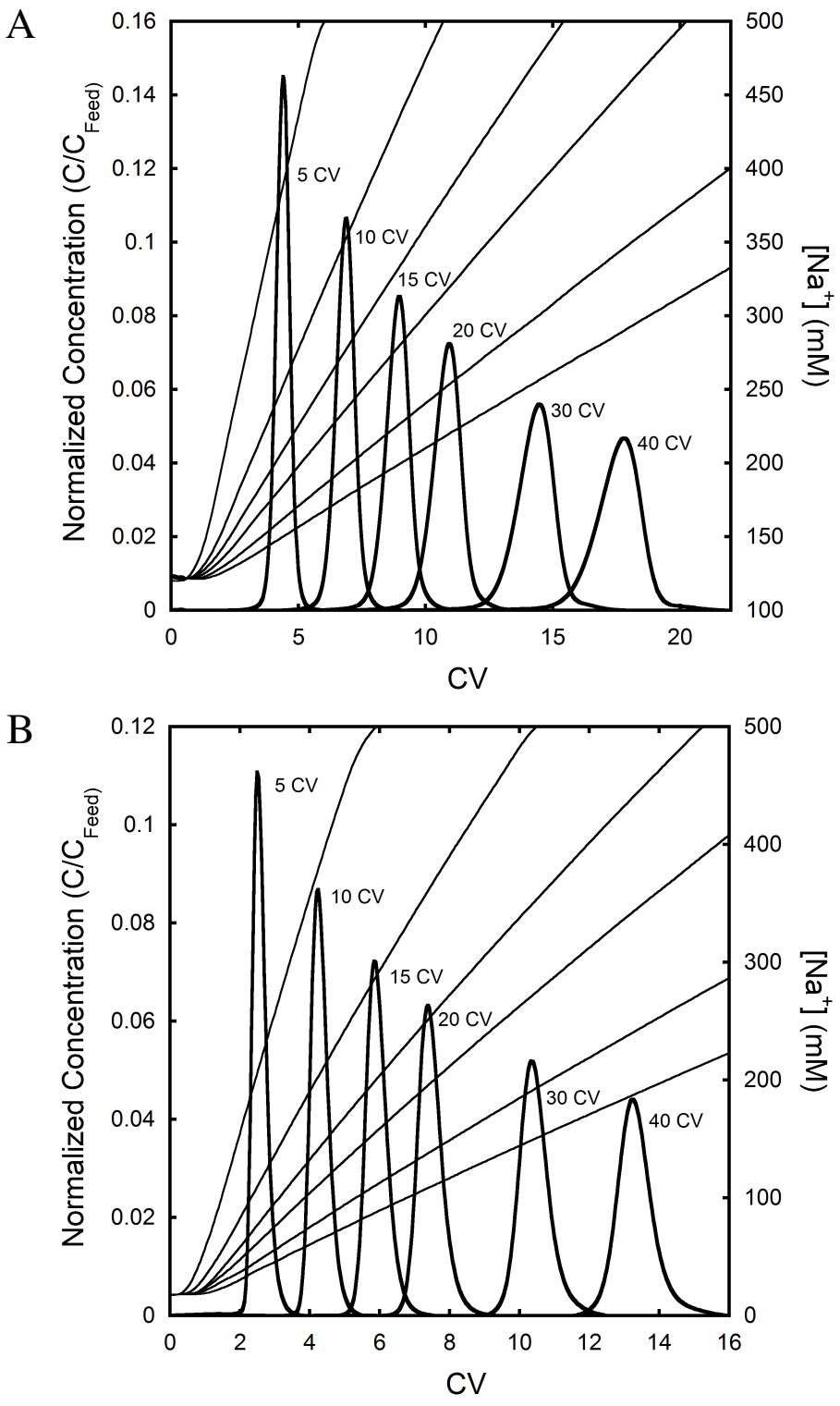


Figure A1: LGE chromatograms for lysozyme on SP-Sepharose FF (A) and for mAb A on POROS XS (B) with gradients from 0 to 500 mM NaCl in 5, 10, 15, 20, 30 CV.

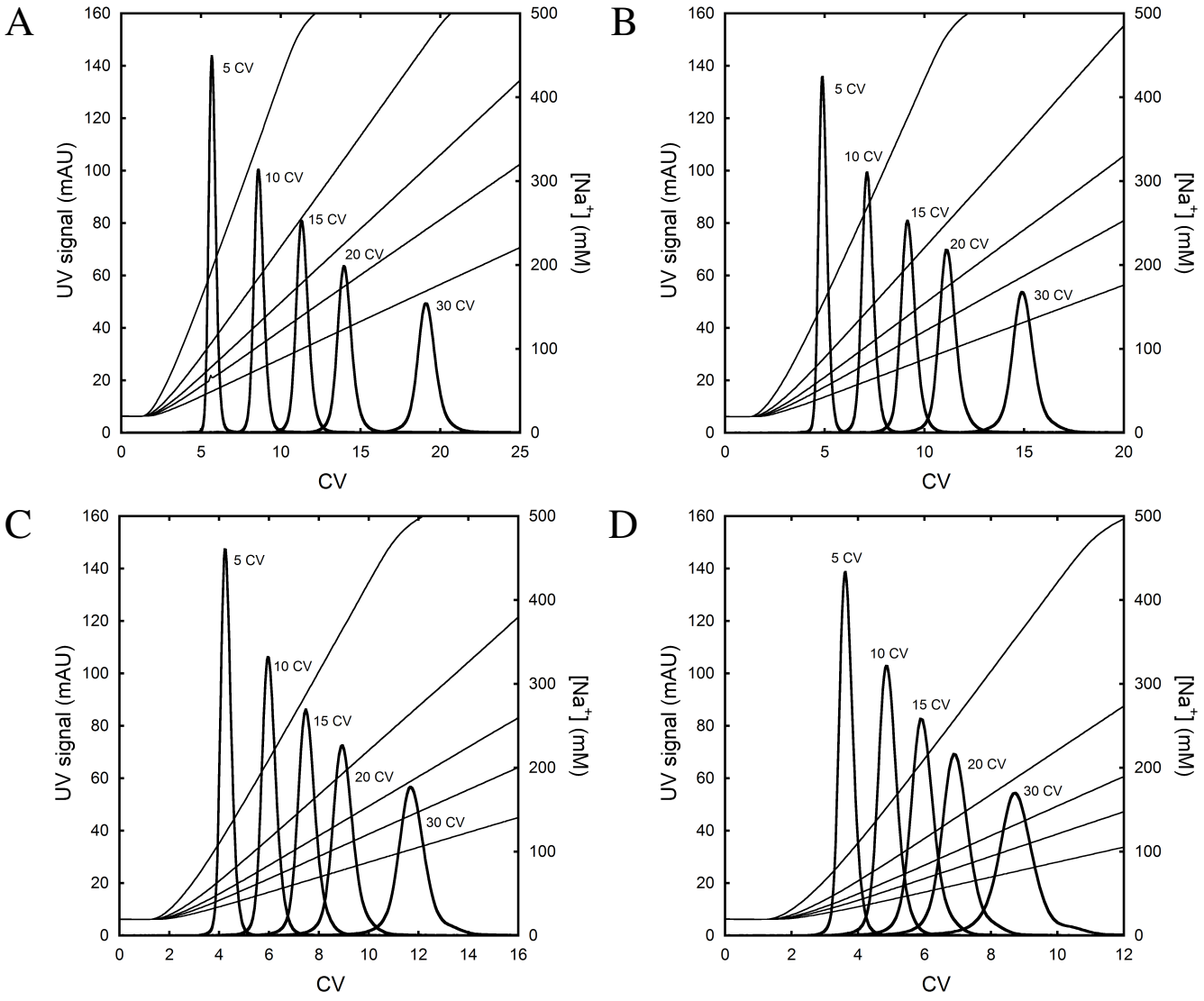


Figure A2: LGE chromatograms for mAb B on POROS XS at pH 5.5 (A), 6 (B), 6.5 (C), and 7 (D) with gradients from 0 to 500 mM NaCl in 5, 10, 15, 20, 30 CV.

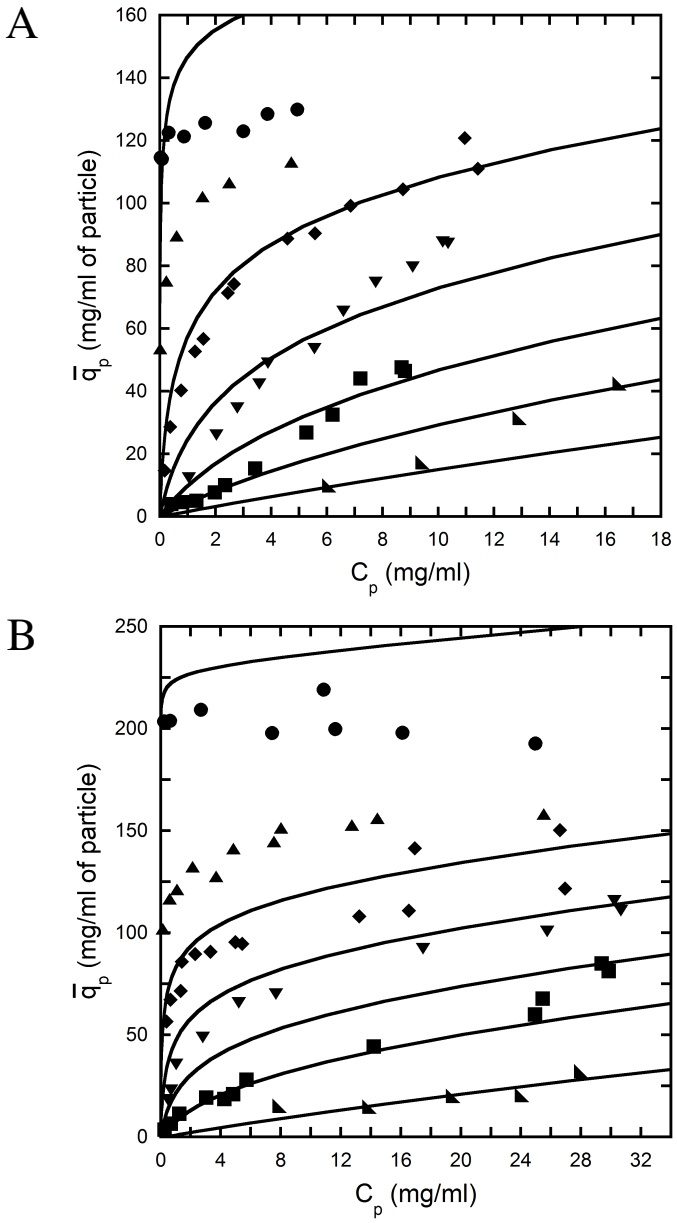


Figure A3: Batch adsorption isotherms for lysozyme on SP-Sepharose FF (A) and for mAb A on POROS XS (B). The solid lines in (A) and (B) are calculated with the SMA model using the *z*- and *A*-values determined from the LGE experiments at low protein loads and best-fit parameters q_0 and σ given in Table A2. In (A) symbols ●, ▲, ◆, ▼, ■ and ◣ represent 50, 140, 200, 260, 320, and 420 mM Na^+ concentrations, respectively. In (B) symbols ●, ▲, ◆, ▼, ■ and ◣ represent 17.6, 93, 118, 143, 168, and 218 mM Na^+ concentrations, respectively.

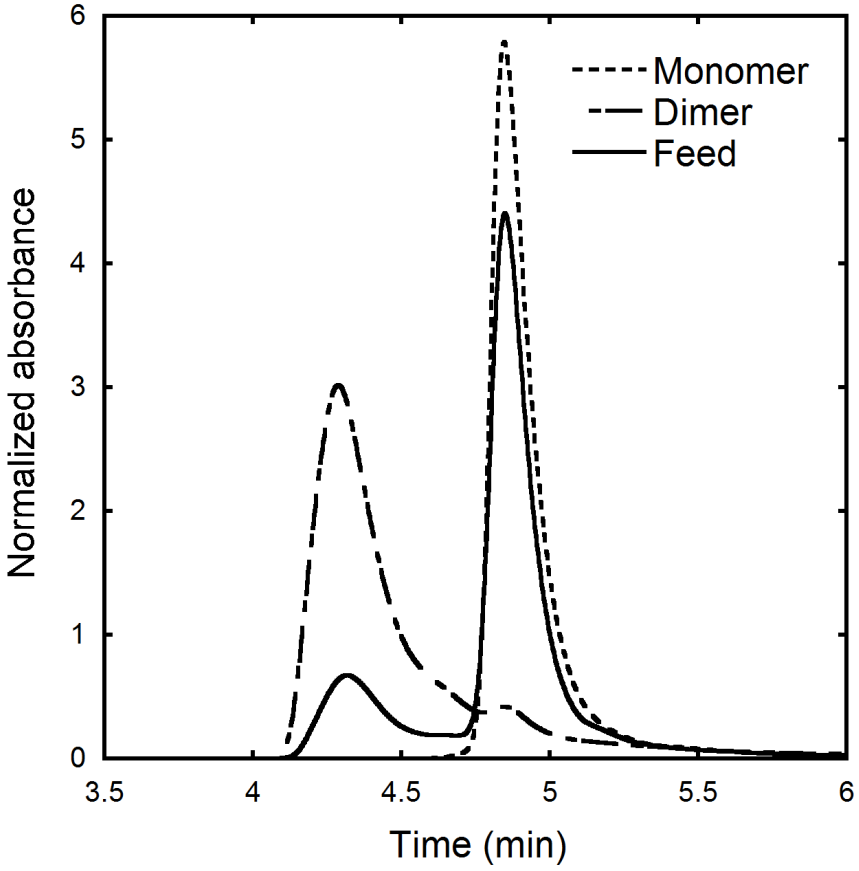


Figure A4: SEC peaks obtained for the monomer-dimer feedstock containing 28% dimer, the enriched monomer sample containing 99% monomer, and the enriched dimer sample containing 98% dimer. The experimental method is described in the main paper.

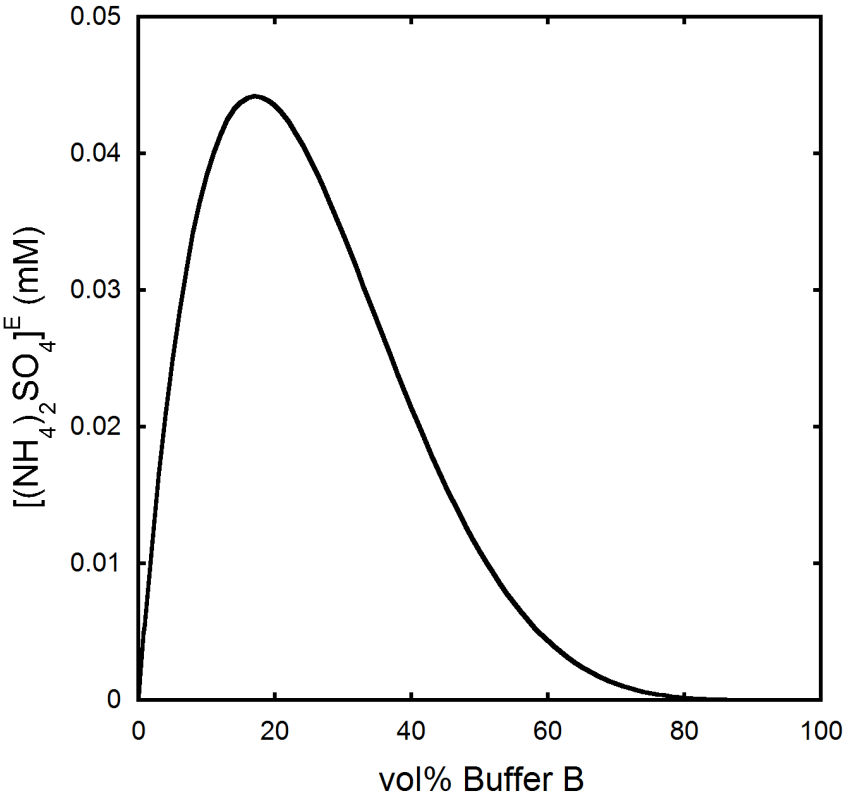


Figure A5: Calculated excess molarity when a buffer with 500 mM ammonium sulfate (Buffer A) is mixed with 0 mM ammonium sulfate (Buffer B). Calculations were made using solution density data from Perry's Chemical Engineers' Handbook.

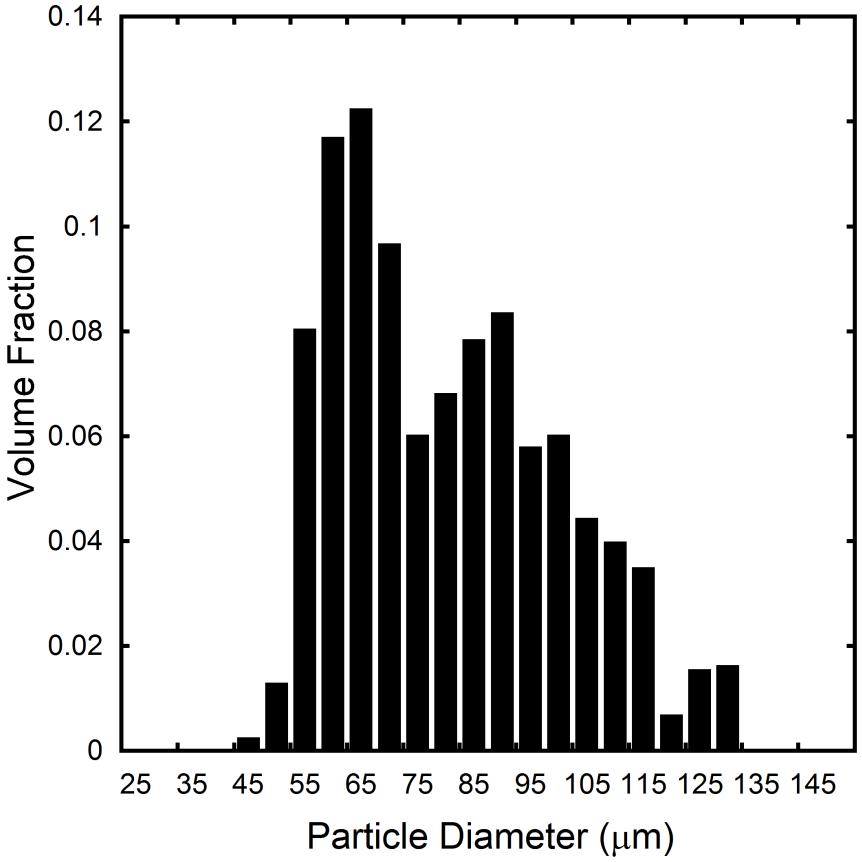


Figure A6: Particle size distribution of Capto Phenyl HS obtained from micrographs.

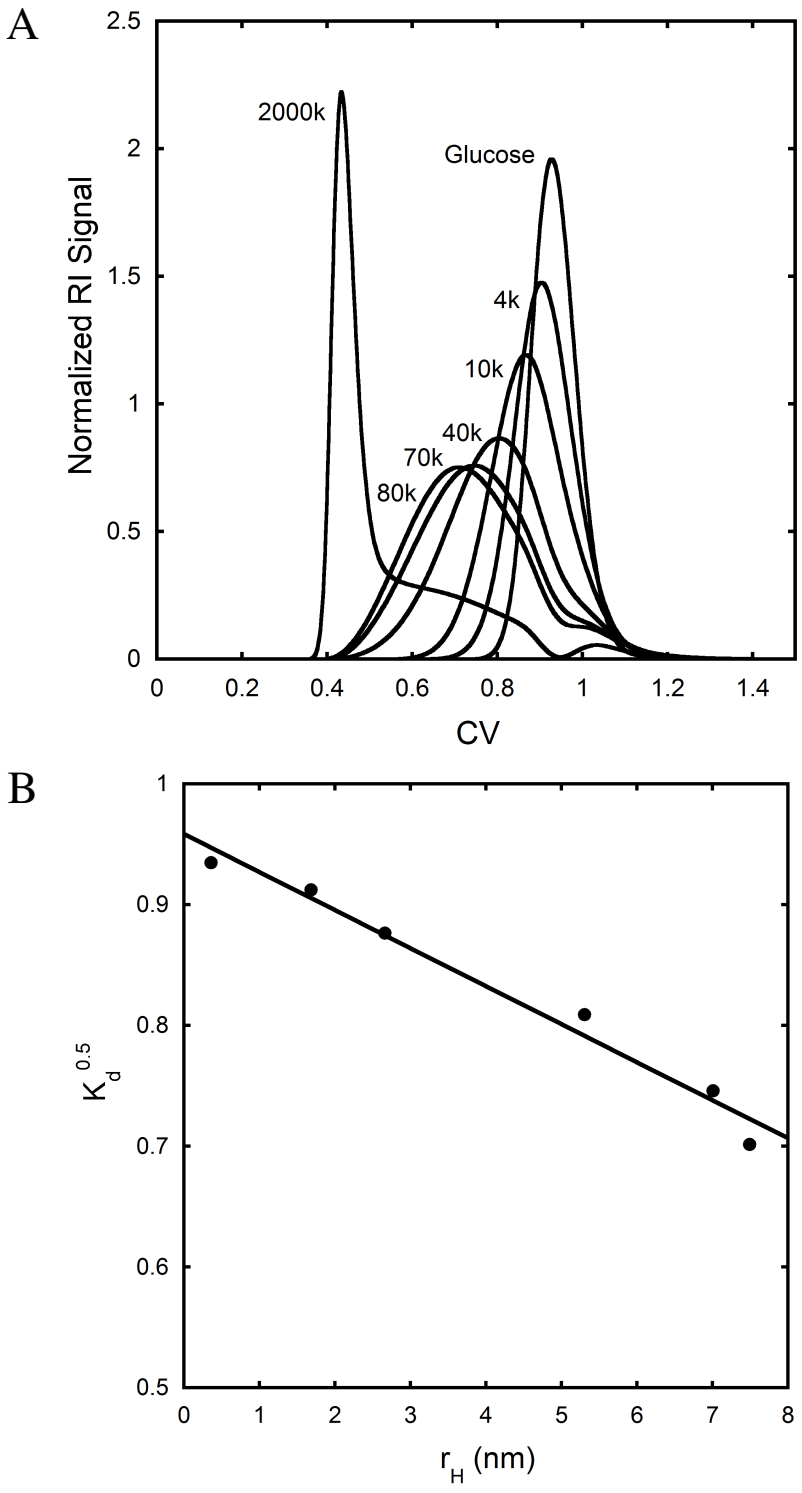


Figure A7: Inverse size exclusion chromatography peaks for glucose and dextran probes on Capto Phenyl HS (A) and square root of partition coefficient vs. hydrodynamic radius of probe (B).

6.2 Buffer calculations

Example of application of eqs. 2.1-2.4 in main paper.

Explicit equations derived from eqs. 2.1-2.4 in the main paper to predict the pH of mixtures containing acetate and phosphate.

Acetate, $\mathbf{A_i = CH_3COO}$, $N_i = 1$, $n = 1$

$$\left(HCH_3COO\right) \Leftrightarrow \left(CH_3COO\right)^{-1} + H^+, K'_{a,1} = K_{a,1} \times 10^{2f(I)}$$

$$\left[\left(HCH_3COO\right)\right] = \frac{C_{\text{Acetate}}}{1+K'_{a,1}/\left[H^+\right]}$$

$$\left[\left(CH_3COO\right)^{-1}\right] = \frac{C_{\text{Acetate}} K'_{a,1}/\left[H^+\right]}{1+K'_{a,1}/\left[H^+\right]}$$

Phosphate, $\mathbf{A_i = PO_4}$, $N_i = 3$, $n = 1,2,3$

$$\left(H_3PO_4\right) \Leftrightarrow \left(H_2PO_4\right)^{-1} + H^+, K'_{a,1} = K_{a,1} \times 10^{2f(I)}$$

$$\left(H_2PO_4\right)^{-1} \Leftrightarrow \left(HPO_4\right)^{-2} + H^+, K'_{a,2} = K_{a,2} \times 10^{4f(I)}$$

$$\left(HPO_4\right)^{-2} \Leftrightarrow \left(PO_4\right)^{-3} + H^+, K'_{a,3} = K_{a,3} \times 10^{6f(I)}$$

$$\left[\left(H_3PO_4\right)\right] = \frac{C_{\text{Phosphate}}}{1+\frac{K'_{a,1}}{\left[H^+\right]}+\frac{K'_{a,1}K'_{a,2}}{\left[H^+\right]^2}+\frac{K'_{a,1}K'_{a,2}K'_{a,3}}{\left[H^+\right]^3}}$$

$$\left[\left(H_2PO_4\right)^{-1}\right] = \frac{C_{\text{Phosphate}}\frac{K'_{a,1}}{\left[H^+\right]}}{1+\frac{K'_{a,1}}{\left[H^+\right]}+\frac{K'_{a,1}K'_{a,2}}{\left[H^+\right]^2}+\frac{K'_{a,1}K'_{a,2}K'_{a,3}}{\left[H^+\right]^3}}$$

$$\left[\left(HPO_4\right)^{-2}\right] = \frac{C_{\text{Phosphate}}\frac{K'_{a,1}K'_{a,2}}{\left[H^+\right]^2}}{1+\frac{K'_{a,1}}{\left[H^+\right]}+\frac{K'_{a,1}K'_{a,2}}{\left[H^+\right]^2}+\frac{K'_{a,1}K'_{a,2}K'_{a,3}}{\left[H^+\right]^3}}$$

$$\left[\left(PO_4\right)^{-3}\right] = \frac{C_{\text{Phosphate}}\frac{K'_{a,1}K'_{a,2}K'_{a,3}}{\left[H^+\right]^3}}{1+\frac{K'_{a,1}}{\left[H^+\right]}+\frac{K'_{a,1}K'_{a,2}}{\left[H^+\right]^2}+\frac{K'_{a,1}K'_{a,2}K'_{a,3}}{\left[H^+\right]^3}}$$

Electroneutrality condition

$$\left[\mathrm{Na^+}\right]+\left[\mathrm{H^+}\right]-\left[\left(\mathrm{CH_3COO}\right)^{-1}\right]-\left[\left(\mathrm{H_2PO_4}\right)^{-1}\right]-2\left[\left(\mathrm{HPO_4}\right)^{-2}\right]-3\left[\left(\mathrm{PO_4}\right)^{-3}\right]-\left[\mathrm{Cl^-}\right]-\frac{K'_w}{\left[\mathrm{H^+}\right]}=0$$

6.3 Details of the empirical interpolation method

The method used to find the value of q_p at set values of C_p^*, $[Na^+]^*$, and pH^* by interpolation is illustrated in Fig. A8 using a hypothetical data set as an example. It is assumed that experimental q_p-values are available at different C_p, $[Na^+]$, and pH conditions as given in Fig. A8A. The first step is to fit the isotherms at each pH and $[Na^+]$ value. The fitted lines are shown in Fig. A8A. The q_p-values (q_{p1} through q_{p9} in the hypothetical example at hand) are then calculated at $C_p = C_p^*$ using the fitted lines at each experimental pH and $[Na^+]$. These values are shown by the open symbols in Fig. A8B.

The second step uses PCHIP to construct interpolating lines at each pH as a function of $[Na^+]$. q_p' -values (q_{p1}', q_{p2}', and q_{p3}' in the hypothetical example at hand) are then calculated from these lines at $[Na^+]=[Na^+]^*$. These values are shown by the open symbols in Fig. A8C.

The third step is to use PCHIP again to construct an interpolating line as a function of pH at $C_p = C_p^*$ and $[Na^+]=[Na^+]^*$, which is finally used to calculate q_p at the given pH^* (see Fig. A8C).

The final step, used to speed the column calculations is to repeat the three steps above to generate a tri-dimensional matrix of q_p-values calculated over broad ranges of C_p, $[Na^+]$, and pH, which is used to rapidly generate q_p-values using trilinear interpolation for any input set of C_p, $[Na^+]$, and pH. In our case, we used MATLAB's *griddedinterpolant* function to perform this last task.

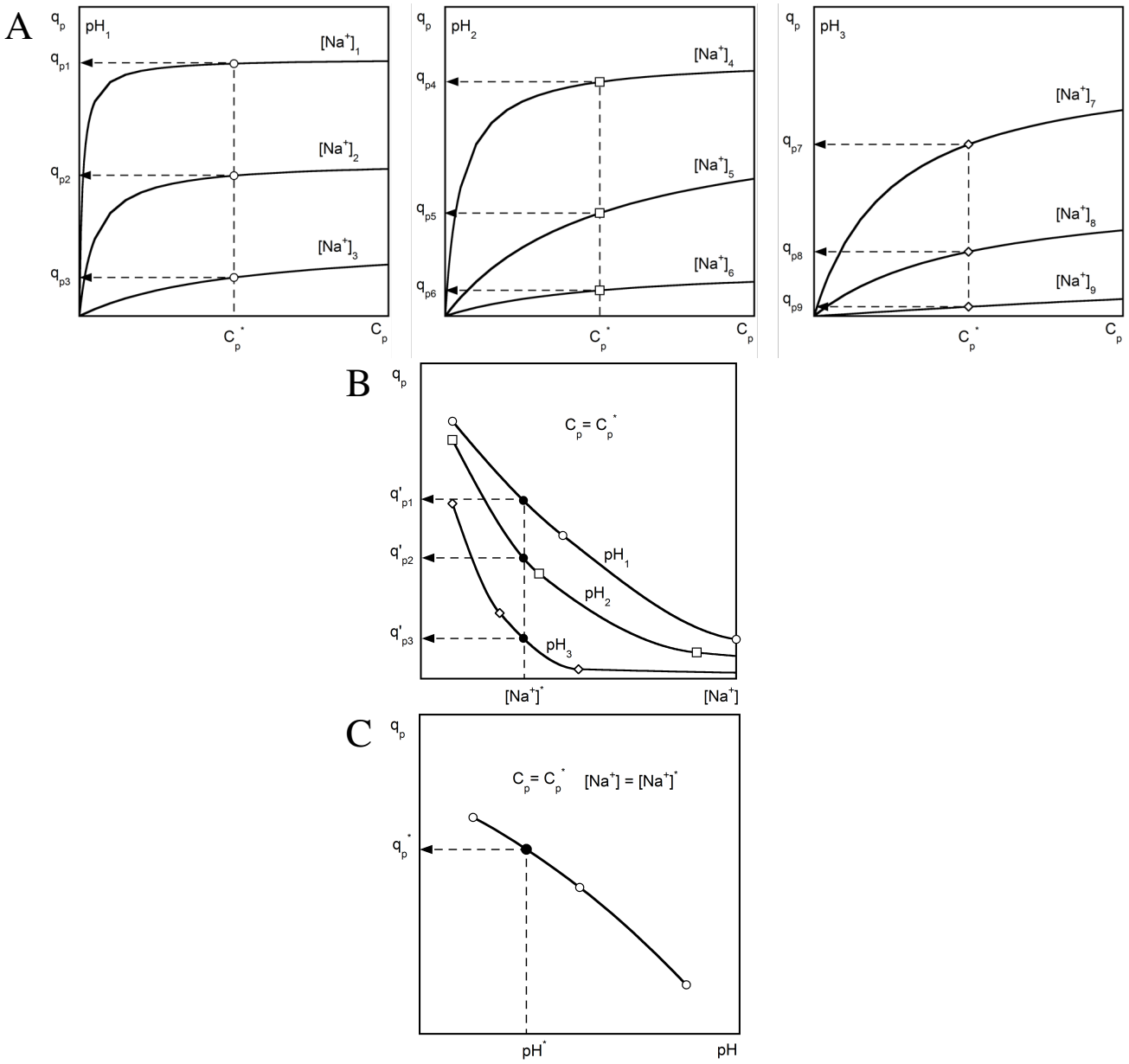


Figure A8: Schematic of the empirical interpolation method shown for a hypothetical example with adsorption isotherms obtained at three pH values and nine $[Na^+]$ values.

6.4 Surface induced two-peak elution behavior

This section specifically highlights the modeling contributions made to another body of work by Jing Guo et al.[1].

6.4.1 Introduction

Ion exchange chromatography plays an important role in the purification of therapeutic proteins. Cation exchange chromatography (CEX), in particular, is commonly used to remove product related impurities from monoclonal antibodies, such as aggregates and charge variants. In most cases, the bound protein is eluted from CEX columns in a single peak at a salt concentration that depends on pH and gradient slope[2–4]. Recent studies, however, have shown that multiple peak elution can also occur on CEX columns as a result of different binding conformations of the protein. Voitl et al.[5], for example, described a two-peak elution behavior for human serum albumin on Fractogel SO3 CEX columns and explained the experimental behavior by assuming that the protein was bound in two different conformations with different binding strength and kinetics[6].

In work by Jing et al.[1], a monoclonal antibody was found to exhibit a two- or three-peak elution behavior when loaded on the CEX resin POROS XS in a sodium acetate buffer at pH 5 and eluted with a salt gradient following a hold step in the load buffer during which the protein remains bound to the column. Two peaks are observed without a hold step while a third more strongly retained peak becomes noticeable with a hold time as low as 10 min. Dynamic light scattering analysis shows that the third peak contains significant levels of aggregates formed in the column while the two early eluting peaks regardless of hold time are shown to comprise exclusively monomeric species.

This section focuses on the two-peak elution behavior obtained for POROS XS when the protein is eluted immediately after loading, i.e. with zero hold time. Mechanistic modeling is used to support the hypothesis that these two early eluting peaks form as a result of the presence of weak and strong binding sites on the resin having, respectively, fast and slow binding kinetics.

6.4.2 Theoretical development

While developing a model to describe the unfolding/aggregation behavior is beyond the scope of this work, a mechanistic model was developed for the two-peak behavior with the intent of helping to elucidate the underlying causes. The model assumes that POROS XS contains two independent binding sites: weak binding sites (1) assumed to have fast on/off kinetics and strong binding sites (2) assumed to have slow on/off kinetics. The physical nature of the two assumed binding sites is not certain, but POROS resins are designed with a bi-modal pore size distribution[7,8] including very large pores, where protein molecules can presumably bind quickly but more weakly, and small pores whether presumably protein molecules can bind more strongly but also more slowly. Accordingly, for each of the two types of binding sites $i = 1$ and $i = 2$, protein binding equilibrium is described by the mass action law model, which can be written as follows[2,4]:

$$q_i = A_i \left[\mathrm{Na}^+\right]^{z_i} C_p \tag{6.1}$$

where q_i is the bound protein concentration, A_i is an equilibrium constant, z_i is the effective protein binding charge, and c is the protein concentration in the particle pores. The corresponding binding kinetics is described for each site type by the following equation:

$$\frac{\partial q_i}{\partial t} = k_i \left(c - \frac{\left[\mathrm{Na}^+\right]^{z_i}}{A_i} q_i \right) \tag{6.2}$$

where k_i is a binding rate constant. Equations 6.1 and 6.2 are coupled with the following equations and boundary conditions:

$$\varepsilon_p \frac{\partial c}{\partial t} + (1 - \varepsilon_p) \sum_i \frac{\partial q_i}{\partial t} = \frac{D_e}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial c}{\partial r} \right) \tag{6.3a}$$

$$t = 0 \rightarrow c = 0, q_i = 0 \tag{6.3b}$$

$$r = 0 \rightarrow \partial c \big/ \partial r = 0 \tag{6.3c}$$

$$r = r_p \rightarrow c = C \tag{6.3d}$$

$$\varepsilon \frac{\partial C}{\partial t} - \frac{3(1-\varepsilon)D_e}{r_p}\left(\frac{\partial c}{\partial r}\right)_{r=r_p} + u\frac{\partial C}{\partial x} = 0 \tag{6.4a}$$

$$t = 0 \rightarrow C = 0 \tag{6.4b}$$

$$x = 0 \rightarrow C = C_F \tag{6.4c}$$

which describe, respectively, mass transfer in the particles and convective transport along the column length, x. In these equations, C is the protein concentration in the mobile phase outside the particles, ε_p is the intraparticle porosity, D_e is the effective pore diffusivity, and r_p is the particle radius. The Na^+ concentration as a function of time and position in the column during the gradient was simulated neglecting any dispersion effects. The resulting set of partial differential equations was solved numerically by finite differences using 30 and 100 discretization points in the radial and axial direction, respectively, which were sufficient to eliminate any numerical dispersion effects. The resulting system of ordinary differential equation was integrated using MATLAB's variable order solver routine *ode15s*.

6.4.3 Results and discussion

Figure A9 shows the two-peak elution behavior of POROS XS by illustrating the effects of load and elution flow rates (Fig. A9A and A9B, respectively). As shown in this figure, the load flow rate had no effect, while the elution flow rate significantly affected the relative magnitude of the two peaks. In the latter case, while the breadth of the peaks decreased as the elution flow rate was increased, as expected because of the increased residence time and, thus, increased plate number, a much greater portion of the protein eluted in the second peak. This result suggests that the distribution of the protein between the two peaks is kinetically controlled over time scales that are comparable to the elution times.

The parameters appearing in these equations were estimated as follows[1]. $\varepsilon = 0.35$ was obtained using the Carman-Kozeny equation from the column pressure drop. $\varepsilon_p = 0.58$ and $D_e = 7.0 \times 10^{-8}$ cm^2/s were obtained from pulse injections of the mAb under non-binding conditions (1 M NaCl) at

different flow rates of the mobile phase as described in ref.[9]. Next we assumed that the weak binding sites have fast kinetics and used $k_1 = 100\ s^{-1}$, which is sufficiently large to ensure that the simulated binding kinetics does not influence retention of the first peak. In this case, $A_1 = 5.37\times10^{28}$ $(mM)^{12.3}$ and $z_1 = 12.3$ for the weak binding sites could be determined with the method outlined in ref.[9] from the retention of the first peak as a function of gradient slope obtained in separate experimental linear gradient elution runs[1]. The remaining parameters $k_2 = 0.013\ s^{-1}$, $A_2 = 2.56\times10^4 9$ $(mM)^2 0.6$, and $z_2 = 20.6$ for the strong binding sites were estimated by regressing the data at varying salt gradients using MATLAB's nonlinear least squares function, *lsqnonlin*. Figure A10 shows the modeling results illustrating the predicted dependence of the elution peaks for different values of the rate constant for adsorption on the strong binding sites (Fig. A10A) and the predicted dependence on the flow rate, using $k_2 = 0.013\ s^{-1}$ (Fig. A10B) for the experimental conditions of Fig. A9B. As seen in Fig. A10A, a value of $k_2 = 0$ obviously leads to a single early-eluting peak and a value of $k_2 = 0.1\ s^{-1}$ leads to a single late-eluting peak. Intermediate values around $0.01\ s^{-1}$ obviously yield two peaks. As seen in Fig. A10B using the regressed values of $k_2 = 0.013\ s^{-1}$, the model predicts a two-peak elution profiles that vary with flow rate in a manner consistent with the experimental results shown in Fig. A9B. In this case, as shown by the model at 1 ml/min, a majority of the protein elutes early as the elution time is too short to permit the protein's full interaction with the strong, but slow binding sites. At 0.25 ml/min, more time is available for interaction with the strong binding sites, which end up dominating the elution process.

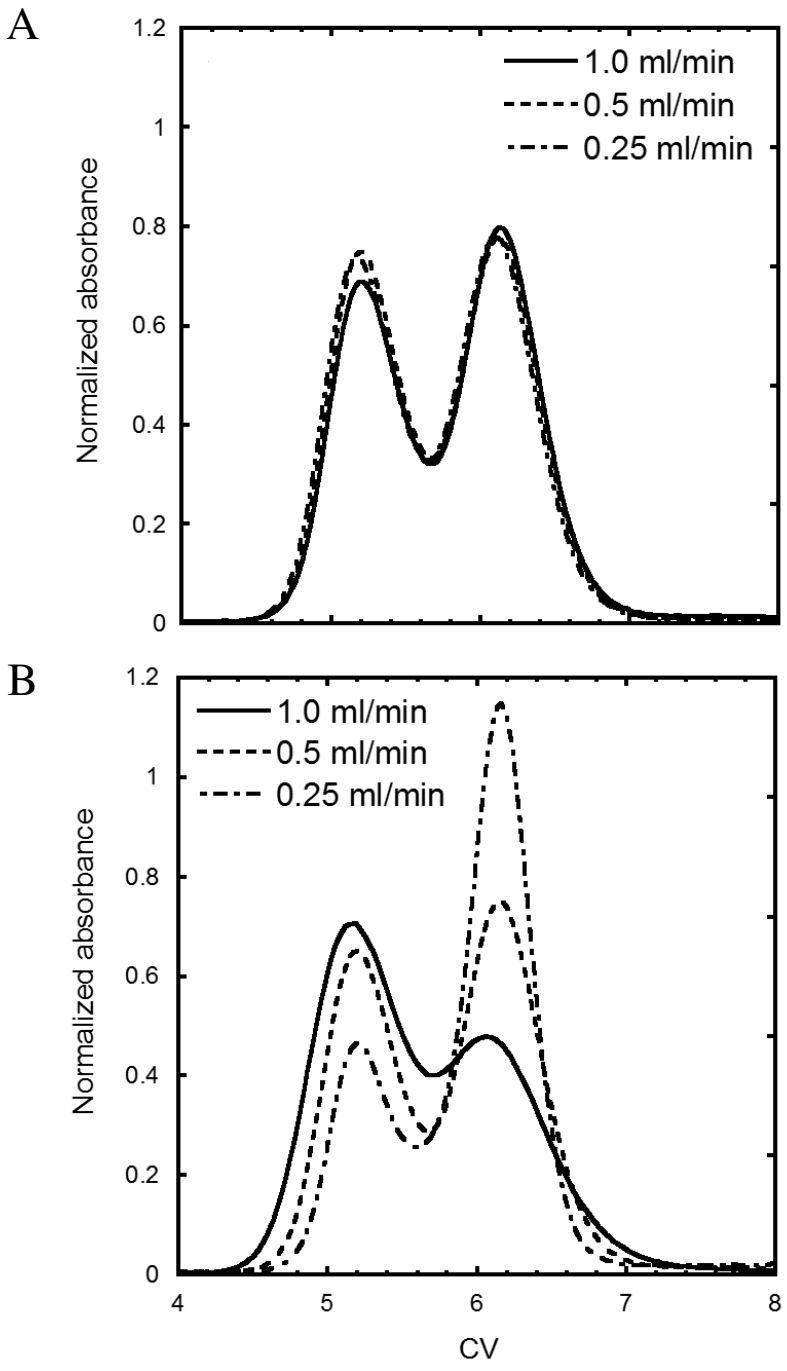


Figure A9: Effects of (A) load flow rate and (B) elution flow rate on elution behavior of the POROS XS column with 0 min hold time followed by a 20 CV 0-1 M NaCl gradient in 40 mM sodium acetate at pH 5.

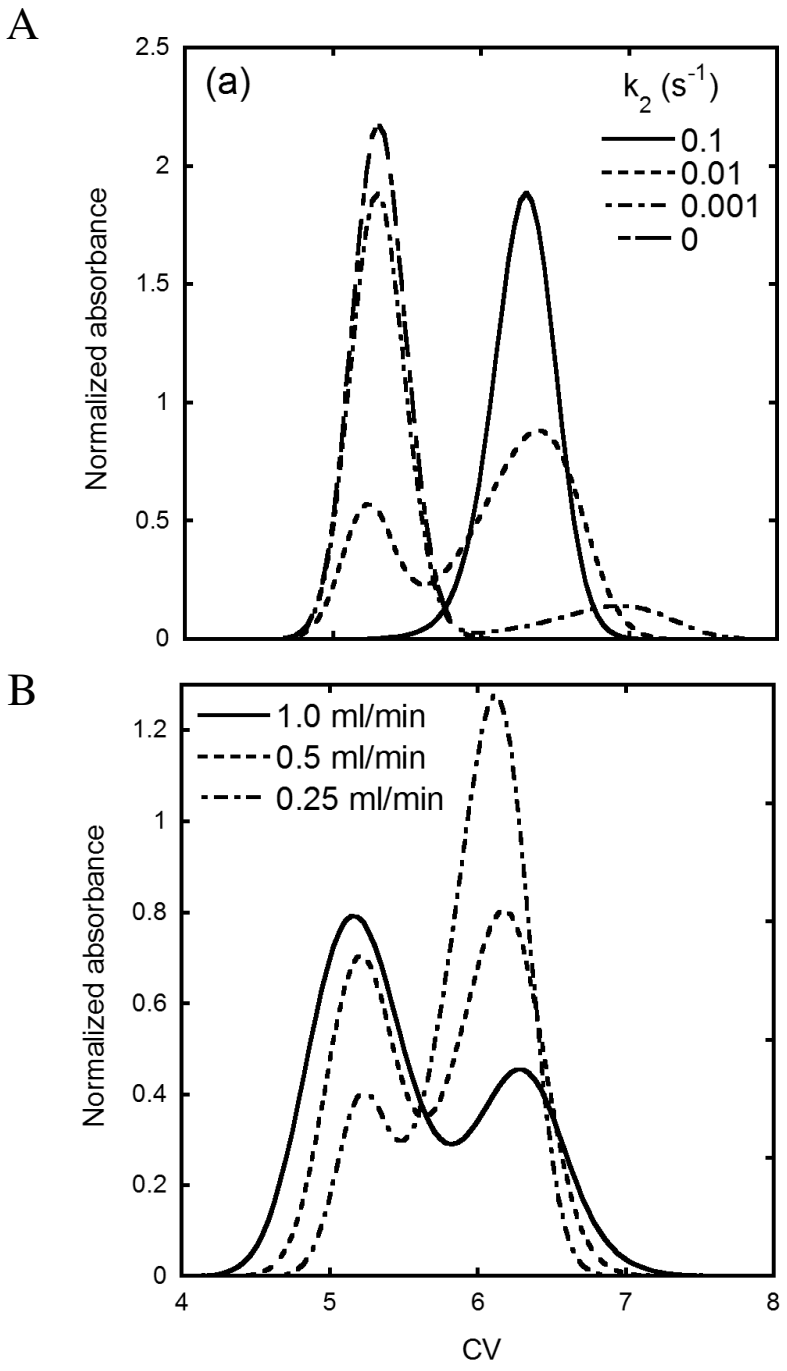


Figure A10: Elution profiles predicted by the model described in Section 6.4.2 for POROS XS with a 20 CV 0-1 M NaCl gradient. (A) Effect of varying k_2 while keeping $k_1 = 100$ s^{-1} at 0.25 ml/min. (B) Effect of elution flow rate for the conditions of Fig. A9B predicted with $k_1 = 100$ s^{-1} and $k_2 = 0.013$ s^{-1}. Other model parameters are given in Section 6.4.3.

6.5 Publications

1. Creasy, A., Reck, J., Carta, G., “Systematic interpolation method predicts two component chromatographic elution”, *Biotechnology Journal*, submitted (2018).

2. Creasy, A., Lomino, J., Barker, G., Carta, G., “Gradient elution behavior of proteins in hydrophobic interaction chromatography with U-shaped retention factor curves”, *Journal of Chromatography A*, 1547, 53-61, (2018).

3. Creasy, A., Barker, G., Carta, G., “Systematic interpolation method predicts protein chromatographic elution with salt gradients, pH gradients and combined salt/pH gradients”, *Biotechnology Journal*, 12, 1600636 (2017).

4. Guo, J., Creasy, A., Barker, G., Carta, G., “Surface induced three-peak elution behavior of a monoclonal antibody during cation exchange chromatography, *Journal of Chromatography A*, 1474, 85-94 (2016).

5. Creasy, A., Barker, G., Yao, Y., Carta, G., “Systematic interpolation method predicts protein chromatographic elution from batch isotherm data without a detailed mechanistic isotherm model”, *Biotechnology Journal*, 10, 1400-1411 (2015).

References

(*1*) Guo, J., Creasy, A., Barker, G., and Carta, G., (2016). Surface induced three-peak elution behavior of a monoclonal antibody during cation exchange chromatography. *Journal of Chromatography A 1474*, 85–94.

(*2*) Yamamoto, S., Nakanishi, K., and Matsuno, R., *Ion Exchange Chromatography of Proteins*; Marcel Dekker: New York, NY, 1988.

(*3*) Yamamoto, S., Nakanishi, K., Matsuno, R., and Kamikubo, T., (1983). Ion exchange chromatography of proteins-prediction of elution curves and operating conditions. 1. Theoretical considerations. *Biotechnology and Bioengineering 25*, 1465–83.

(*4*) Carta, G., and Jungbauer, A., *Protein Chromatography: Process Development and Scale-Up*, 1st ed.; Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 2010.

(*5*) Voitl, A., Butté, A., and Morbidelli, M., (2010). Behavior of human serum albumin on strong cation exchange resins: I. Experimental analysis. *Journal of Chromatography A 1217*, 5484–5491.

(*6*) Voitl, A., Butté, A., and Morbidelli, M., (2010). Behavior of human serum albumin on strong cation exchange resins: II. Model analysis. *Journal of Chromatography A 1217*, 5492–5500.

(*7*) Afeyan, N. B., Gordon, N. F., Mazsaroff, I., Varady, L., Fulton, S. P., Yang, Y. B., and Regnier, F. E., (1990). Flow-through particles for the high-performance liquid chromatographic separation of biomolecules : perfusion chromatography. *Journal of Chromatography A 519*, 1–29.

(*8*) Wu, Y., Simons, J., Hooson, S., Abraham, D., and Carta, G., (2013). Protein and virus-like particle adsorption on perfusion chromatography media. *Journal of Chromatography A 1297*, 96–105.

(*9*) Carta, G., Ubiera, A. R., and Pabst, T. M., (2005). Protein Mass Transfer Kinetics in Ion Exchange Media: Measurements and Interpretations. *Chemical Engineering & Technology 28*, 1252–1264.