**Modeling Host Cell Protein Retention in Chromatographic Polishing Operations**

Chase Evan Herman

Advisor: Professor Abraham M. Lenhoff

**Abstract**

Contemporary bioprocess design is largely empirical, despite numerous incentives to adopt model-based approaches. Mechanistic chromatography models could improve the design and optimization of preparative separations, but their utility is often limited by insufficient adsorption equilibrium information. The goal of the proposed study is to predict protein adsorption equilibrium from proteomics data. This will be applied to model the retention of host cell protein impurities in flow-through polishing operations. Isocratic and linear gradient elution experiments will be performed to generate an equilibrium data set for model proteins. Retention under these two operating conditions will be related, and colloidal theory will be applied to develop predictive models for the equilibrium data. Fractionation experiments and proteomic analysis will then be performed on bioprocess samples, and the accuracy of predictive retention models will be assessed. This will contribute to both heuristic knowledge and chromatographic theory. Preliminary results from isocratic and linear gradient experiments are presented, along with initial colloidal modeling studies.

**1. Introduction**

Biopharmaceuticals have enabled qualitatively improved medical treatments over the past half century. Well over 300 protein therapeutics are on the market today, and hundreds are estimated to be in development pipelines1,2. These therapeutics represent the fastest growing sector of the pharmaceutical industry3,4. Approximately $57 billion were invested into biopharmaceutical research and development in 20185. Coupled with this development is a need to efficiently design and optimize the downstream processes for purifying therapeutics. Chromatography is essential to this purification, but the design of preparative chromatographic separations remains largely empirical6. Heuristics and high-throughput experimentation support nearly all chromatographic process design6–8. This is in contrast to the petrochemical industry, where sufficiently accurate thermodynamic models allow process design to occur almost entirely *in silico*. The general advantages of *in silico* design are numerous: reduction of development costs and timelines, the ability to perform detailed optimization and explore large parameter spaces, and the elevation of process understanding. Such an approach is foreseeable but essentially lacking in biomanufacturing, due to the complexity of the underlying phenomena9.

* 1. **Motivation for *in silico*****BioprocessDesign**

The economic need for *in silico* design of chromatographic processes is becoming more pronounced. With the rapid development of biopharmaceuticals, time to market has become a critical determinant of product success10. Also, several biopharmaceutical patents have recently expired or are due to expire in the near future1. As this occurs, biosimilars are beginning to enter the market1. This competition is naturally creating incentive to reduce manufacturing costs, of which downstream processing represents 50-80%11. Improvements in upstream technology have dramatically increased fermentation titers to around 10 g/L, placing further pressure on optimizing downstream processes8,12.

Regulatory incentives are also promoting the implementation of bioprocess models. In response to inefficiencies in the drug approval process, the US FDA issued the current good manufacturing practice (cGMP) initiative in 200213. This, in addition to the process analytical technology (PAT)14 guidance and International Conference on Harmonization (ICH) guidelines Q8/Q9/Q1015–17, led to the FDA’s quality by design (QbD) initiative18,19. The goal of QbD is to instill quality into biopharmaceutical manufacturing by moving the industry and regulatory authorities to a more scientific, risk-based approach. It promotes a thorough understanding of the relationship between critical process parameters, the drug product’s critical quality attributes, and the product’s clinical behavior8,13. Understanding the impact of process parameters on product attributes is demonstrated to regulatory authorities through modeling. This is currently achieved with statistical and interpolatory methods20. These include response surface models, principal component analysis, and neural networks derived from high-throughput experimental data20–23.

Although high-throughput experimentation supports the development of most preparative chromatographic separations, the resulting statistical models have limited predictive capabilities. This may be attributed to their lack of a mechanistic basis. On the other hand, several mechanistic chromatography models are well-established24,25. However, their use is frequently precluded by ambiguities in selecting the appropriate model and accurately estimating the associated parameters6,26. When these issues can be overcome, mechanistic models offer the general advantage of performing well outside of their calibration ranges26. In addition to the general benefits of *in silico* design, these models can help biomanufacturers support QbD, integrate control strategies with PAT, transfer process knowledge across scales, and avoid overly robust operating windows20.

* 1. **Bioprocess Design Applications and Supporting Technologies**

Mechanistic modeling is applicable to multiple units in biopharmaceutical purification trains. An ideal purification train would contain only one chromatographic step operated in bind-and-elute mode, where the product is simultaneously concentrated and purified by adsorption. However, this certainly does not achieve requisite purities in contemporary practice. Heuristics dictate there be at least two chromatographic units. This is evident in the purification of monoclonal antibodies (mAbs), which comprise the largest class of protein therapeutics1. The downstream processing of mAbs typically consists of the following platform sequence: centrifugation, depth filtration, protein A affinity chromatography, pH viral inactivation, polishing, and multiple filtration steps10,27. Polishing includes one or more operations, which may be cation-exchange (CEX), anion-exchange (AEX), hydrophobic interaction (HIC), or multimodal (MMC) chromatography4. These steps are designed to reduce the concentration of host cell protein (HCP) impurities to acceptable levels for the drug substance, typically less than 100 ppm28. One of the polishing steps is frequently operated in flow-through mode, which is designed to promote the adsorption of impurities while the product flows through unretained. AEX is often the flow-through step in mAb purification, because the majority of secreted HCPs are more acidic than mAbs29,30.

Flow-through AEX provides an excellent application example for mechanistic modeling. Uncleared HCPs may persist through this operation due to weak adsorption, product association, or column saturation. Beyond the saturation front, the concentration of weakly adsorbing species is dilute, and linear adsorption dominates. This condition makes it more tractable to model the weakly adsorbing species than the movement of the saturation front; however, it is theoretically possible to do both with mechanistic chromatography models. Capturing product association would require complementary biophysical modeling techniques, like molecular dynamics simulations. Such a comprehensive modeling approach would enable *a priori* selection of the optimal resin and solution conditions, namely pH and ionic strength.

In this application, the mechanistic modeling alone may appear infeasible for a variety of reasons. The salient challenges include: (1) the absence of precise theory for describing protein adsorption; (2) the need for identification and quantification of all HCPs present; and (3) the computational expense. However, the latter two challenges may be addressed with recently developed technologies. Advances in mass spectrometry, namely the emergence of sequential window acquisition of all theoretical mass spectra (SWATH-MS), are enabling consistent data acquisition in high-throughput proteomics31. This allows protein identities and concentrations to be determined in complex solutions. Also, the computational expense of solving mechanistic chromatography model equations may be addressed with a recently developed software. This software, the Chromatography Analysis and Design Toolkit (CADET), is an open source modeling and simulation framework20. CADET combines modern simulation algorithms with a universal platform to efficiently solve the partial differential and algebraic equations comprising any mechanistic model.

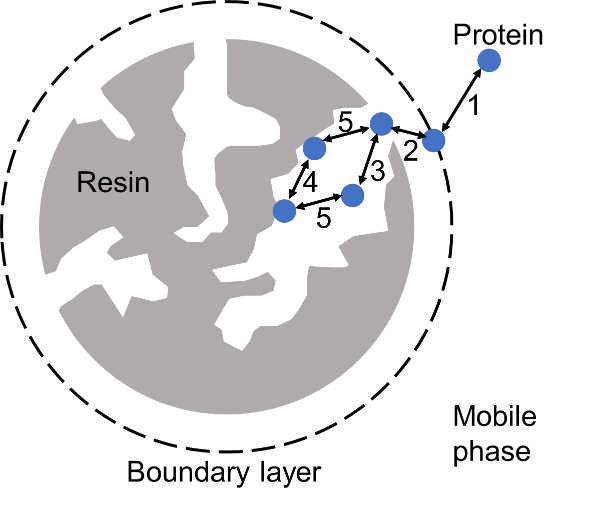
1. **Theory**
   1. **Mechanistic Chromatography Models**

In a chromatographic operation, a mobile phase containing products and impurities is passed over a stationary phase. Differences in attraction of the products and impurities to the stationary phase are exploited to achieve separation. In biomanufacturing applications, the stationary phase is typically a porous resin which is packed into a column. Mechanistic models of the chromatographic process are predicated on coupled mass balances for the mobile and stationary phases. Thermodynamic or kinetic expressions are also required to describe protein adsorption. Several mechanistic models have been developed; they differ in the extent to which the relevant transport phenomena are described. Figure 1 illustrates these phenomena25. They include: (1) convection and diffusion in the extraparticle mobile phase, (2) film diffusion through the laminar boundary layer surrounding the resin particles, (3) pore diffusion, (4) surface diffusion, and (5) adsorption and desorption. The most rigorous description, the general rate model, accounts for each of these processes. This model is given in consolidated form below20,32,33:

|  |  |
| --- | --- |
|  | (1) |
|  | (2) |

where is time, is the axial column coordinate, is the radial coordinate in a resin particle, is the resin particle radius, is the concentration of species in the extraparticle mobile phase, is the concentration of species in the intraparticle (i.e. pore) mobile phase, is the concentration of species on the resin surface, is the film mass transfer coefficient of species , is the column axial dispersion coefficient, is the pore diffusivity of species , is the surface diffusivity of species , is the extraparticle void fraction, is the intraparticle void fraction, and is the interstitial velocity of the mobile phase.

Eqs. 1 and 2 are one-dimensional mass balances for the extraparticle mobile and stationary phases, respectively. The stationary phase is defined here as both the solid resin and the contained intraparticle fluid phase. Danckwerts boundary conditions are usually applied to Eq. 1; a film diffusion expression and finite radial symmetry are applied as boundary conditions for Eq. 220. Implicit in the use of the general rate model are the following assumptions: (1) the resin particles are much smaller than the column dimensions and may be treated as homogenously distributed, and (2) radial transport may be neglected in the macroscopic column. As its name suggests, this model is the most general description of transport in biochromatography. Several other mechanistic models exist, including the equilibrium, dispersive, transport, transport-dispersive, reactive, and reactive-dispersive models25. These may all be effectively obtained through simplifications of the general rate model.



*Figure 1.* Transport *phenomena in protein chromatography.*

* 1. **Adsorption Isotherms**

Key to the implementation of any mechanistic model is the relationship between pore and surface concentration. This relationship may be described as a kinetic approach to equilibrium. However, the adsorption and desorption steps are relatively fast in protein chromatography, and local equilibrium is frequently assumed25. In either case, adsorption equilibrium is described by an isotherm. The isotherm is a function of pH, ionic strength, resin, and solute34. The classic isotherm example is the Langmuir model, which was originally developed for gas adsorption35. This model has frequently been applied to biochromatography on an empirical basis24. Similarly, numerous isotherm models exist and have been applied with various degrees of empiricism. This is because protein adsorption cannot be currently treated with precise theory. The reason involves several complicating factors, including: the anisotropy of protein shape, charge, and hydrophobicity; local and global protein unfolding; aggregation; non-specific binding; the absence of accurate molecular descriptions of resins; the disparate length scales and ionic strengths encountered; poorly-understood solvation effects; and challenges in extending single-component adsorption information to multi-component mixtures24.

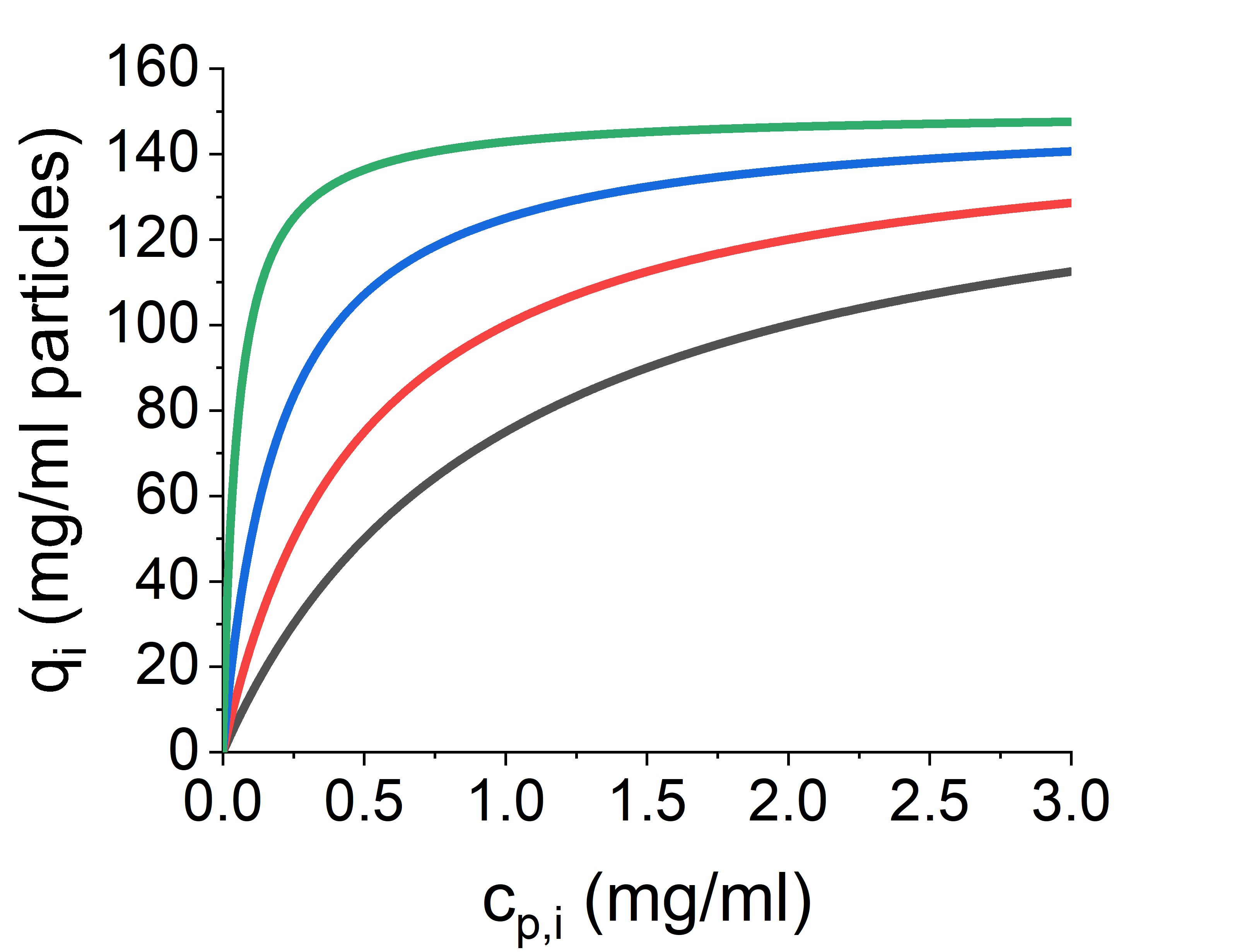


Figure 2. Example of Langmuir adsorption isotherms with different equilibrium constants.

Despite their diversity, all isotherm models share qualitatively similar characteristics. In the infinitely dilute solution limit, pore and surface concentrations are linearly related by . is the dilute solution adsorption equilibrium constant, known in the chromatography literature as the Henry coefficient25,36. Also, all isotherms asymptotically approach a saturation capacity. Figure 2 provides an illustrative example of Langmuir isotherms with equivalent saturation capacities and different equilibrium constants. Isotherm parameters are almost always determined from experimental data. The equilibrium constant may be quantified independently from other parameters through isocratic pulse tests. Such experiments provide measurements of the chromatographic retention factor, . This is a dimensionless quantity characterizing the retention time of a solute relative to its retention time under nonadsorbing conditions. The retention factor is related to the equilibrium constant by , where the phase ratio is defined as ()/. Once obtained, the equilibrium constant provides sufficient isotherm information to mechanistically model dilute solutions. Several static and dynamic methods exist for measuring isotherm parameters related to nonlinear adsorption25.

* 1. **Isotherm Parameter Prediction**

Due to their modeling utility, the prediction of adsorption isotherm parameters has been a long-standing goal37. This goal has been notably pursued for the equilibrium constant in ion exchange chromatography, where the mechanism of action is relatively well-understood. Yamamoto et al. provided one of the first theoretical descriptions of the equilibrium constant’s dependence on ionic strength38–41. Their work, termed the stoichiometric displacement model (SDM), used the law of mass action to describe protein adsorption as a stochiometric ion exchange process. This was combined with plate theory, where a chromatographic column is treated as a sequence of equilibrium stages, to relate retention under isocratic and linear gradient elution (LGE) conditions. Recently, an analogous expression has been derived by Pfister et al. from the more realistic, albeit highly simplified, equilibrium model42. Their derivation was based on the method of characteristics and perturbation analysis. They also derived an allegedly more accurate version of the isocratic-LGE relationship.

The Yamamoto et al.38–41 model allowed chromatographers to compute ion exchange equilibrium constants from readily-obtained LGE data; however, no *a priori* predictions could be made. This challenging task was instead approached with colloidal theory. Roth and Lenhoff furnished one of the first predictions of protein adsorption based on detailed protein structural information and continuum electrostatics computations43. Namely, they solved the linearized Poisson-Boltzmann equation numerically; they also included van der Waals contributions to the free energy of adsorption. Their work captured the ionic strength dependence of the equilibrium constant semi-quantitatively44. However, later developments demonstrated the approach’s ability to discriminate retention differences caused by slight protein structural variations45. Detailed calculations involving complete protein structures were also compared with reductionist approximations43,44. These approximations treated proteins as spheres with or without a dipole moment. Interestingly, the detailed description and the homogeneously charged sphere approximation were found to produce very similar predictions for lysozyme adsorption onto quartz. Incorporation of lysozyme’s dipole moment scaled the monopole predictions almost linearly. However, for chymotrypsinogen A, the dipole moment scaled the predictions nonlinearly.

The surprisingly tolerable performance of spherical approximations suggests that computationally inexpensive colloidal models may suffice for isotherm parameter estimation. This is particularly desirable for implementation in mechanistic chromatography models, where the intended use in optimization algorithms necessitates multiple calculations. Recently, an analytical solution of the nonlinear Poisson-Boltzmann equation was published for two interacting spheres46. This relatively inexpensive result was implemented by Guélat et al. as a descriptive model for the equilibrium constant’s dependence on both pH and ionic strength47. Based on prior work36, theoretical titration curves were later incorporated to approximate protein net charge48. This left only one adjustable parameter in the equilibrium constant model: the minimum protein-resin separation distance which bounds van der Waals contributions. Although empirical, resin-specific correlations were later developed for this parameter49. The model was also extended to nonlinear adsorption conditions49,50, but this introduced more adjustable parameters.

1. **Goals and Objectives**

The goal of the proposed study is to mechanistically model HCP behavior in chromatographic polishing columns. This will involve a combination of experimental and computational work. The number of HCP species present in polishing, which is on the order of 100, prohibits the experimental determination of individual isotherm parameters51. However, HCP sequences may be ascertained with SWATH-MS proteomics. Using this information, predictive approaches will be implemented for isotherm parameters. Emphasis will be placed on the equilibrium constant’s dependence on pH and ionic strength for any combination of resin and protein. This represents the most tractable goal for predictive modeling. This is also immediately relevant to flow-through AEX columns, for which there is a paucity of modeling studies in the literature. Industrial collaborators, who are specifically interested in this polishing step, will provide the necessary bioprocess samples. Whenever possible, experiments will be designed to improve both heuristic understanding and theoretical techniques. The following briefly enumerates specific objectives.

**Aim 1:** Generate dilute solution isocratic pulse data on ion-exchange resins.

**Aim 2:** Generate analogous LGE data, and asses the accuracy of models relating retention under isocratic and LGE conditions.

**Aim 3:** Develop and implement predictive models for the ion-exchange equilibrium constant.

**Aim 4:** Fractionate protein A eluate samples, analyze the fractions with SWATH-MS proteomics, and test the methods developed in aims 2 and 3 for predicting linear adsorption behavior.

**Aim 5:** Extend predictive models to nonlinear isotherm parameter estimation.

The essence of aim 1 is to gather direct equilibrium adsorption measurements. Ionic strength will be varied over several levels to provide retention factor data spanning two to three orders of magnitude. This will be performed at a few levels of pH for a variety of model proteins and resins. At least a few AEX resins of interest will be employed to identify general resin retentivity trends. Literature data for other model proteins may also be compiled. The essence of aim 2 is to assess the accuracy of models and techniques for indirectly measuring equilibrium constants from LGE experiments. Although the Yamamoto et al.38–41 method for relating retention under isocratic and LGE conditions is the most frequently used, very few comparisons of isocratic retention and the associated LGE predictions are found in the literature. The few comparisons which do exist were made at relatively low retention factors40. This fails to provide a complete assessment of model validity, because model inaccuracies are expected to become more pronounced as the retention factor increases. Both the Yamamoto et al.38–41 and Pfister et al.42 models will be implemented. These models require minimal experimental data, namely retention time and gradient slope, to predict the equilibrium constant. Inverse fitting parameter estimation techniques, which rely on complete LGE chromatogram information, will also be implemented. The accuracy of each method will be assessed through comparison to the aim 1 experimental data.

Aim 3 will seek to develop *a priori* predictions for the equilibrium constant. The Guélat et al.47 model was compared to experimental data for mAbs only48–50. Instead, the prediction accuracy will be assessed here for a variety of proteins. Previous colloidal isotherm theory52–54 will also be applied, and equilibrium constant predictions will be compared to the aim 1 experimental data. If justified, further computational and theoretical study will be made to improve prediction accuracy. The models found to be most suitable in aims 2 and 3 will then be applied in aim 4. Protein A eluate samples will be obtained, and these will be fractionated under linear adsorption conditions (i.e. a dilute solution relative to the column size will be employed). Two or more AEX LGE experiments will be performed with different gradient slopes. Following an approach similar to Levy et al.4 and Nfor et al.55, the flow-through will be collected, along with multiple eluate fractions. The fractions will then be concentrated and subjected to 1D SDS-PAGE. Densitometry analysis will be used to determine the relative abundance of HCPs in each fraction.

The parent fractions will be pooled as necessary for proteomics analysis, and simultaneous HCP identification and quantification will be performed with SWATH-MS. Using the densitometry and proteomics results, chromatograms will be constructed for the most abundant HCPs. The corresponding equilibrium constants will be predicted with the aim 2 results, and these will be compared to aim 3 *a priori* predictions based solely on HCP sequences. Even semi-quantitative *a priori* predictions would be useful in directing bioprocess design efforts. The utility of these predictions will then be demonstrated with analogous experiments under isocratic conditions. Here, the flow-through will be fractionated, analyzed, and the identities of unretained HCPs will be compared to predictions from sequence information. Finally, aim 5 will attempt to extend model predictions to nonlinear isotherm parameters based on previous colloidal isotherm theory52–54. These predictions will be compared to isotherm measurements obtained from inverse fitting experimental breakthrough curves56. Similar fractionation experiments will then be performed under overloaded conditions, and the results compared to mechanistic model predictions. Preliminary results for aims 1—3 are described below.

1. **Materials and Methods**
   1. **Materials and Solutions**

Sodium hydroxide, hydrochloric acid, sodium chloride, monobasic sodium phosphate, dibasic sodium phosphate, acetic acid, and sodium acetate were purchased from Fisher (Fair Lawn, NJ). Ethanolamine, lithium nitrate, and 2000 kDa blue dextran were purchased from Sigma (St. Louis, MO). Bovine liver catalase, bovine milk β-lactoglobulin, hen egg white lysozyme, hen egg white ovalbumin, and *S. cerevisiae* alcohol dehydrogenase were also purchased from Sigma (St. Louis, MO). GE Healthcare Capto Q, Applied Biosystems POROS 50 HQ, and Applied Biosystems POROS XQ were kindly provided by Bristol-Myers Squibb (Devens, MA). SP Sepharose FF was purchased from GE Healthcare (Upsala, Sweden). The physicochemical properties of these resins are given in Table 1.

Table 1. Resin properties as provided by manufacturers.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Resin | Mode | Base matrix | Mean  dp (µm)a | Ligand | Ionic capacity (µmol/ml)b | Dynamic Capacity (mg/ml)b |
| POROS XQ | AEX | Poly[styrene divinylbenzene] | 50 | Quaternary amine | 80 - 130 | ≥ 140c |
| POROS 50 HQ | AEX | Poly[styrene divinylbenzene] | 50 | Quaternary PEI | Not given | 65 - 90c |
| Capto Q | AEX | Agarose | 90 | Quaternary amine | 160 - 220 | ≥ 100d |
| SP Sepharose FF | CEX | Agarose | 90 | Sulfonate | 180 - 250 | 70e |

a Mean particle diameter; b Per unit packed bed volume; c 5% breakthrough of BSA; d 10% breakthrough of BSA; e 10% breakthrough of ribonuclease.

Solid masses were measured with an Ohaus Scout® Pro 4000 g and a Denver Instrument 110 g balance. Solutions were prepared using deionized water further purified with an EMD Millipore Synergy® UV system. Solution pH and conductivity were measured with a Cole-Parmer PC200 pH/conductivity meter, and pH adjustments were made with the addition of concentrated sodium hydroxide or hydrochloric acid. Solutions were filtered using Fisher 0.2 µm aPES membranes (Pittsburgh, PA). With the exception of lysozyme at pH 7, protein solutions were purified to remove lyophilized excipients with three buffer exchange cycles using EMD Millipore Amicon® ultracentrifugal filter units (Cork, Ireland) and a Cole-Parmer MP centrifuge. All protein solutions were filtered using Thermo Scientific 0.2 µm PVDF membranes (Rockwood, TN). Protein concentrations were measured with a Thermo Scientific NanoDrop Lite Spectrophotometer. Experiments were conducted with lysozyme at pH 5, 7, and 9 in 75 mM sodium acetate, 25 mM sodium phosphate, and 25 mM ethanolamine buffers, respectively. Experiments conducted with all other proteins were performed at pH 7 in 20 mM sodium phosphate. Buffer solutions were prepared with 0 and 1 M NaCl.

* 1. **Experimental Methods**

Resin storage solutions were exchanged three times with the appropriate packing buffer using an Eppendorf MiniSpin® Plus centrifuge. SP Sepharose FF was gravity packed into a 0.5 x 10 cm Waters® AP chromatography column to a final bed volume of 2.0 ml. All other resins were flow packed at 600 cm/h using an Äkta Explorer liquid chromatography system. POROS XQ and POROS 50 HQ were packed into 0.5 x 5 cm Waters® AP columns to final bed volumes of 0.8 ml. Capto Q was packed into a 1 x 20 cm GE C column to a final bed volume of 2.9 ml. The SP Sepharose FF column was validated at pH 7 by equilibrating with 5 column volumes (CVs) of 0 M NaCl, injecting a 100 µl pulse of 1 M NaCl, and washing the column with 0 M NaCl. All other columns were validated in the same manner, except a 0.3 M NaCl solution was used during equilibration and washing. Columns with conductivity trace asymmetries between 0.8 and 1.6 were accepted. All tests were performed at 300 cm/h.

For both isocratic and LGE experiments, at least a 5 CV equilibration period in the appropriate buffer preceded a 100 µl sample injection. Protein solutions of at most 10 mg/ml were used. Absorbance was measured at 215 and 280 nm. Each test was followed by a 5 CV wash of 1 M NaCl. Columns were regenerated at least once every ten experiments with 500 mM NaOH. Requisite NaCl concentrations were achieved for isocratic tests through inline mixing of 0 and 1 M NaCl buffers. Gradients for LGE tests began immediately after sample injection and proceeded from 0 to 1 M NaCl. The Äkta P-960 pump was used to load samples in breakthrough experiments.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Table 2. Model protein properties.* | | | | |
| Protein | pI | Predominant State at pH 7 | Mass (kDa) a | Coordination Groups a |
| Ovalbumin | 4.7 57 | Monomer 58,59 | 43 | None |
| β-lactoglobulin | 5.2 60 | Dimer 60 | 37 | None |
| Catalase | 5.4 b | Tetramer 61 | 240 | 4 Heme 61 |
| Alcohol dehydrogenase | 5.6 b | Tetramer 62 | 147 | 4 Zn2+  63 |
| Lysozyme | 11.4 64 | Monomer 65 | 14 | None |
| a Corresponding to the predominant multimer; b Provided by the manufacturer. | | | | |

1. **Results and Discussion**

Isocratic pulse tests were performed with the three AEX resins listed in Table 1. These were selected based on their utility in flow-through operations and their incomplete characterization in the literature. Ovalbumin, β-lactoglobulin, catalase, and alcohol dehydrogenase were selected as model proteins due to their size range and acidic isoelectric points. This information is listed in Table 2, along with relevant properties used in modeling.

In each isocratic test, a chromatogram in absorbance was obtained. Baseline drift was observed; this was corrected by linearly adjusting the absorbance measurement to offset the drift. Retention time was computed from the first moment of the adjusted chromatogram. The nonadsorbing reference time was measured for each protein-resin combination by performing duplicate isocratic tests at 1 M NaCl. These data were used to compute retention factors. Figure 3 gives the results of isocratic AEX tests performed to date. As assumed in the SDM, retention appeared to follow a power-law in ionic strength. The monotonic decrease in retention is attributable to the greater strength of electrostatic protein-surface interactions than protein-protein interactions66. These data provide a relative retentivity ordering of the three AEX resins. The power-law slope is nearly equivalent for the two POROS resins, with POROS XQ consistently being more retentive by a factor of ~3. The relative retentivity of Capto Q is variable. This may be due to its surface morphology. Of the three resins, Capto Q is the only one with dextran surface extenders. These tentacles provide multiple interaction sites and may be implicated in protein size effects67.

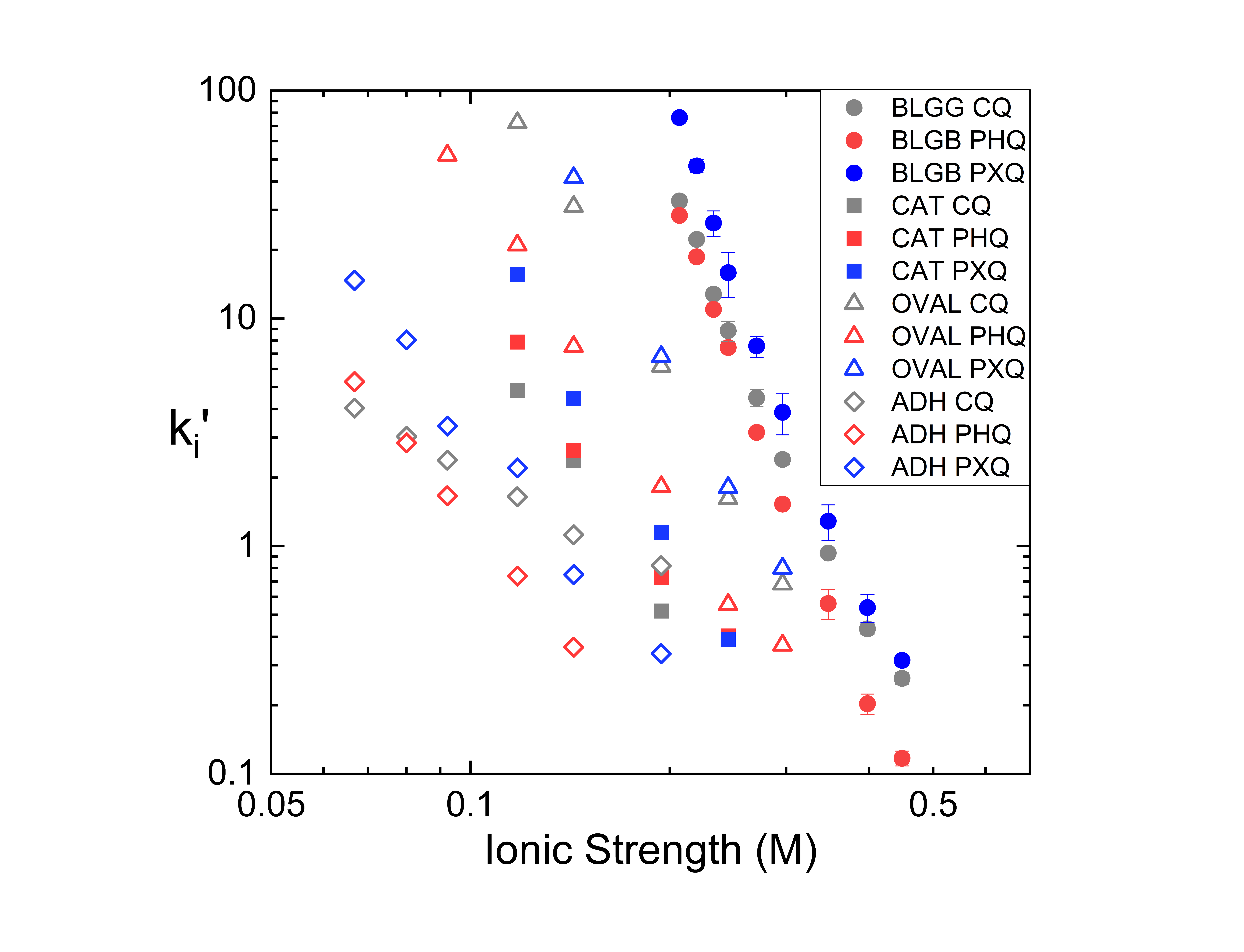
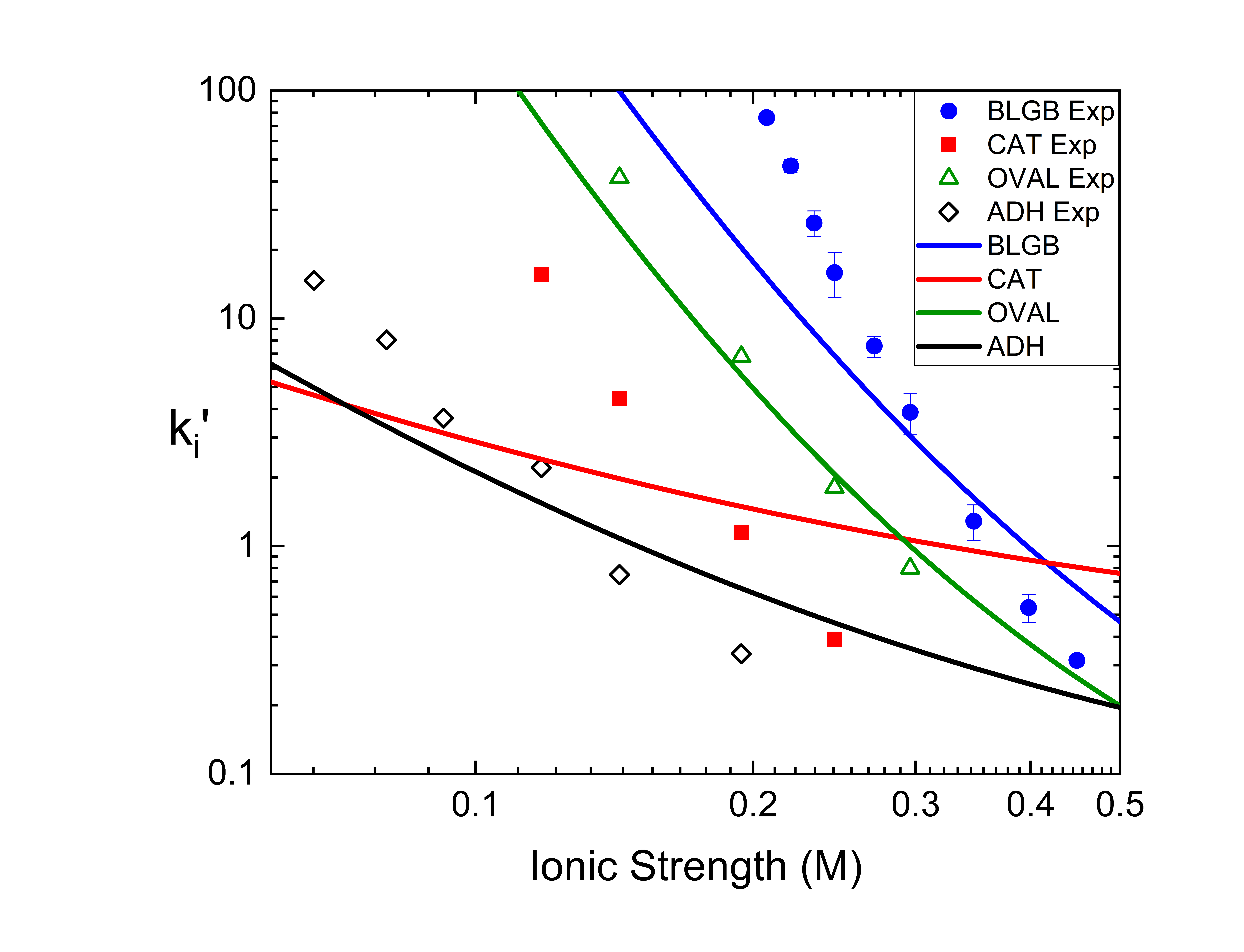
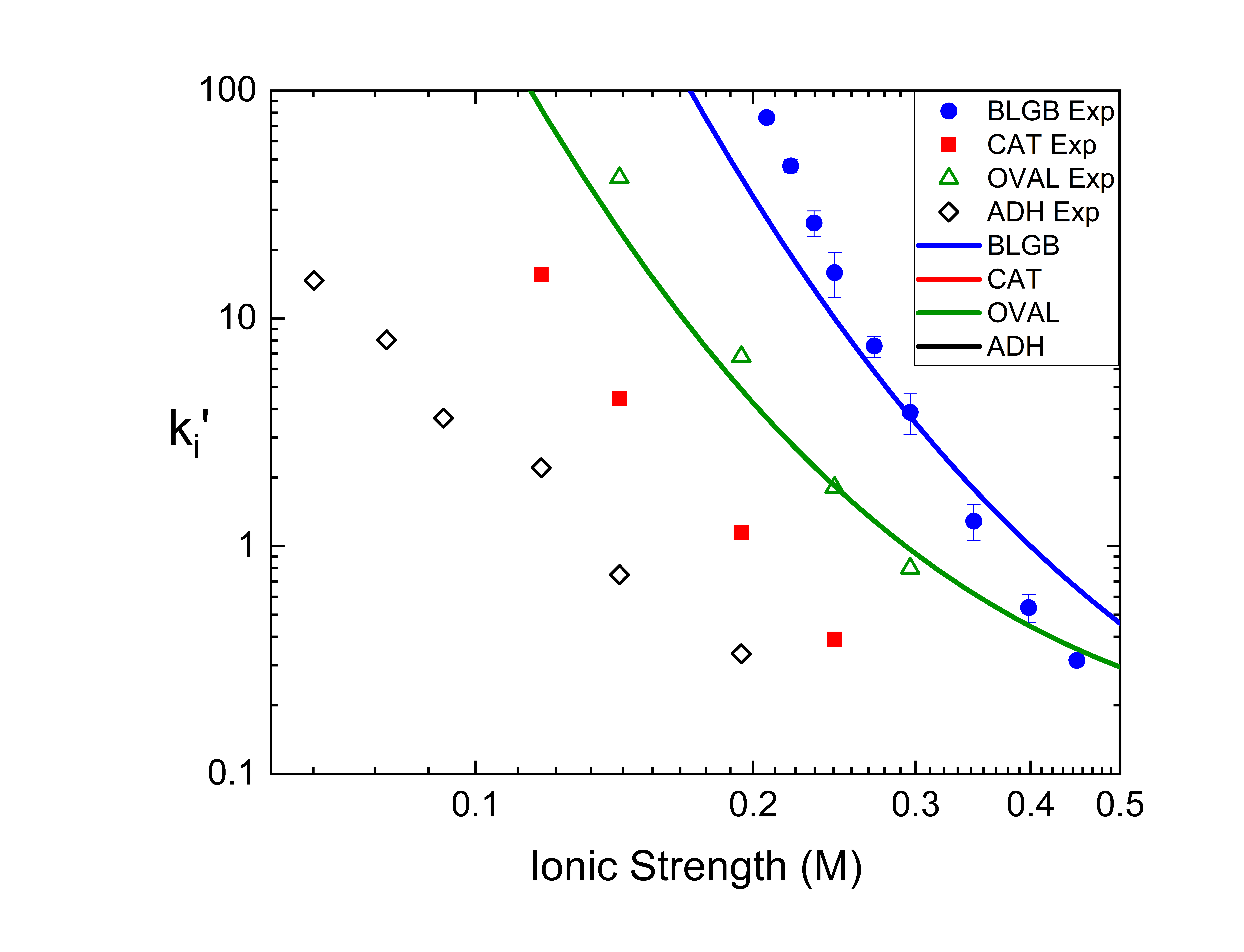
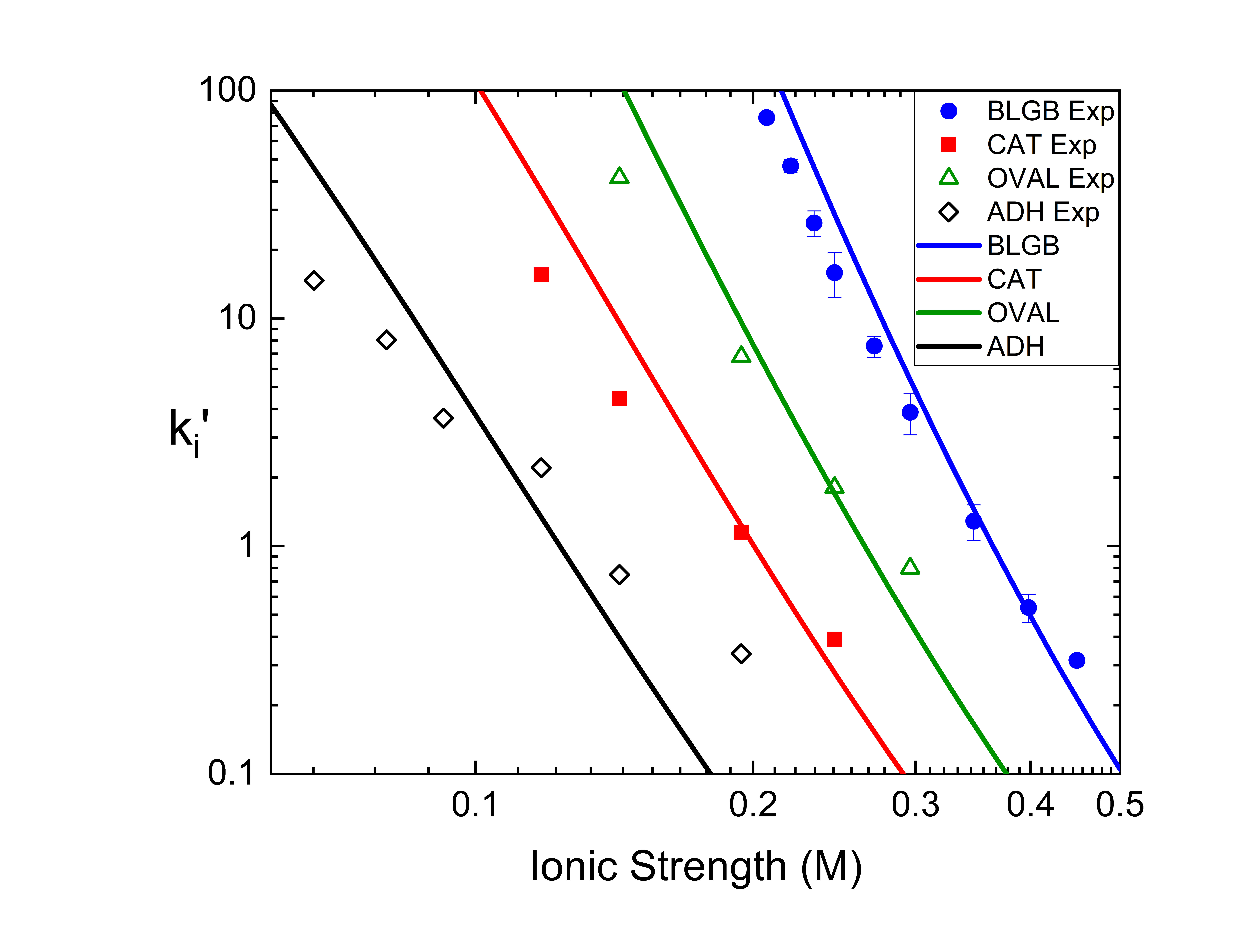
Similar isocratic tests were performed with lysozyme on SP Sepharose FF at three levels of pH. Corresponding LGE experiments were also performed on this system with five linear gradients ranging incrementally from 10 to 50 CV in duration. The models of Yamamoto et al.38–41 and Pfister et al.42 were followed to regress the LGE data, resulting in SDM parameters. These were used to predict isocratic retention; the results are shown in Figure 4. Although it should theoretically be more accurate, the model from Pfister et al. evidently underpredicts the power-law slope for this system. Further study is required to inversely fit the LGE chromatograms with colloidal isotherm parameters, which may provide more accurate predictions.

Figure 3. Retention factors from isocratic pulse experiments on AEX resins at pH 7. CQ = Capto Q, PHQ = POROS 50 HQ, and PXQ = POROS XQ. BGLB = β-lactoglobulin, CAT = catalase, OVAL = ovalbumin, and ADH = alcohol dehydrogenase. N = 2 for β-lactoglobulin experiments; N = 1 otherwise.

Figure 4. Comparison of isocratic pulse data (symbols, N = 2) to predictions from LGE experiments (lines, 5 gradient slopes with N = 1) for lysozyme on SP Sepharose FF at three pH levels. Y = the Yamamoto et al.38–41 model and P = the Pfister et al.42 model for relating LGE to isocratic data. Error bars on isocratic data for pH 5 and 9 are smaller than the symbol size.



**(a)**

**(b)**

**(c)**

Figure 5. Comparison of POROS XQ experimental data to predictions made with (a) colloidal isotherm theory, (b) the Guélat et al.47 model, and (c) the modified Guélat et al. model.

In order to model retention, it is necessary to know the pore size distribution (PSD) and surface charge density of the resins. Resin PSD may be measured with a series of dextran probes68; these measurements have previously been completed for POROS 50 HQ and SP Sepharose FF69,70. It is reasonable to assume the same PSD for both POROS resins, because they have the same base matrix. However, the PSD of SP Sepharose FF likely differs from Capto Q due to its dextran tentacles. The Capto Q PSD is yet to be characterized.

The resin surface charge density may be estimated from ionic capacity. This was lacking for POROS 50 HQ, but it can be determined with a titration experiment. Following the approach of Gadam et al.71, LiNO3 was loaded onto the POROS 50 HQ column until breakthrough. Absorbance was measured at 310 nm, and a calibration curve relating LiNO3 concentration to absorbance was also constructed. A mass balance25 then allowed for computation of the number of adsorbed nitrate groups. Triplicate experiments yielded an ionic capacity measurement of 72 ± 1 µmol/ml of packed bed. Under the assumption of equivalent PSDs, this is consistent with the retention ordering observed in Figure 3 between the two POROS resins.

This information was used with colloidal isotherm theory to develop a retention model. Specifically, theoretical titration curves were constructed from protein sequences48 and the information shown in Table 2. Equivalent sphere radii were computed from molecular masses and an assumed protein density of 1.35 g/ml72. This was used to estimate homogeneous surface charge densities for proteins. Expressions developed by Sader54, based on asymptotic matching solutions for the spherical nonlinear Poisson-Boltzmann equation, were then used to relate surface charge density and surface potential. The Grahame equation73 was used to relate surface charge density and potential for the resin, where the charge density was estimated from ionic capacity. Once obtained, the two surface potentials were used with previous colloidal isotherm theory53 to estimate equilibrium constants. These were related to chromatographic retention factors with phase ratios computed from protein radii and resin PSDs. Van der Waals contributions to adsorption were also included. A minimum protein-resin separation distance was thereby introduced. This represents the only adjustable parameter; decreasing it almost linearly increases retention predictions. Future work is required to modify the inclusion of van der Waals contributions to eliminate this adjustable parameter.

The results of this preliminary modeling effort are shown in Figure 5a. Power law slopes were clearly underpredicted for all proteins other than ovalbumin. This issue is indicative of a need for nonlinear scaling, reminiscent of that previously observed with dipole moment incorporation44. Future work may address this issue by replacing the homogenously charged sphere approximation with a collection of point charges, following an approach similar to Kirkwood74. Also, the Guélat et al.47 model was applied, and the results are shown in Figure 5b. This model failed to predict the retention of catalase and alcohol dehydrogenase at pH 7, both of which are close to their theoretical isoelectric points. Investigation revealed this was due to the sign of the first term in the model’s expression for the electrostatic free energy of adsorption. Strictly for exploratory purposes, the Guélat et al. model was modified *ad hoc* by reversing the sign of this term; the resulting predictions are shown in Figure 5c. Although entirely empirical, this modification enabled the prediction of catalase and alcohol dehydrogenase retention while simultaneously improving the β-lactoglobulin power-law slope. Comparable trends were also observed with the POROS 50 HQ and SP Sepharose FF predictions (not shown). Future study is required to assess whether this modification can be theoretically justified.

1. **Conclusions**

Economic and regulatory incentives are promoting the use of mechanistic biochromatography models. Such models could provide a basis for the rational design and optimization of preparative separations. However, their use is largely limited by a need for HCP adsorption equilibrium information. The experimental determination of adsorption equilibria is currently prohibited by the number of species involved. Colloidal models and proteomics may offer a solution, at least for the prediction of dilute solution equilibrium constants which are relevant to flow-through AEX operations. Even semi-quantitative *a priori* predictions would help direct experimental efforts to develop this polishing step. The goal of the proposed study is to develop such predictions and simultaneously improve heuristic understanding. Preliminary results include direct equilibrium measurements for three AEX resins and a few model proteins. These allowed for resin retentivity ordering. Similar measurements were completed for lysozyme on SP Sepharose FF at three levels of pH, along with corresponding LGE experiments. The accuracy of models relating LGE data to isocratic retention was assessed. The ionic capacity of POROS 50 HQ was also determined, and colloidal models for retention were implemented. Immediate future studies will expand the acquired data set and seek to develop truly predictive retention models.

**References**

1. Walsh, G. Biopharmaceutical benchmarks 2018. *Nature Biotechnology* **36**, 1136-1145 (2018).
2. Liu, H. F., Ma, J., Winter, C. & Bayer, R. Recovery and purification process development for monoclonal antibody production. *MAbs* **2**, 480–499 (2010).
3. Ecker, D. M., Jones, S. D. & Levine, H. L. The therapeutic monoclonal antibody market. *MAbs* **7**, 9–14 (2015).
4. Levy, N. E., Valente, K. N., Lee, K. H. & Lenhoff, A. M. Host cell protein impurities in chromatographic polishing steps for monoclonal antibody purification. *Biotechnology and Bioengineering* **113**, 1260–1272 (2016).
5. Morrison, C. & Lähteenmäki, R. Public biotech 2018—the numbers. *Nature Biotechnology* **30**, 714–721 (2019).
6. Lenhoff, A. M. Ion-exchange chromatography of proteins: the inside story. *Materials Today: Proceedings* **3**, 3559–3567 (2016).
7. Close, E. J., Salm, J. R., Bracewell, D. G. & Sorensen, E. Modelling of industrial biopharmaceutical multicomponent chromatography. *Chemical Engineering Research and Design* **92**, 1304–1314 (2014).
8. Nfor, B. K., Verhaert, P. D. E. M., van der Wielen, L. A. M., Hubbuch, J. & Ottens, M. Rational and systematic protein purification process development: the next generation. *Trends in Biotechnology* **27**, 673–679 (2009).
9. Lenhoff, A. M. A natural interaction: Chemical engineering and molecular biophysics. *AIChE Journal* **49**, 806–812 (2003).
10. Shukla, A. A., Wolfe, L. S., Mostafa, S. S. & Norman, C. Evolving trends in mAb production processes. *Bioengineering and Translational Medicine* **2**, 58–69 (2017).
11. Lowe, C. R., Lowe, A. R. & Gupta, G. New developments in affinity chromatography with potential application in the production of biopharmaceuticals. *Journal of Biochemical and Biophysical Methods* **49**, 561–574 (2001).
12. Singh, N., Pizzelli, K., Romero, J. K., Chrostowski, J., Evangelist, G., Hamzik, J., Soice, N. & Cheng, K. S*.* Clarification of recombinant proteins from high cell density mammalian cell culture systems using new improved depth filters. *Biotechnology and Bioengineering* **110**, 1964–1972 (2013).
13. Rathore, A. S. & Winkle, H. Quality by design for biopharmaceuticals. *Nature Biotechnology* **27**, 26–34 (2009).
14. US Food and Drug Administration. PAT Guidance for industry PAT—a framework for innovative pharmaceutical development, manufacuring, and quality assurance. (US Department of Health and Human Services, Rockville, MD, 2004).
15. US Food and Drug Administration. Guidance for industry: Q8(R2) pharmaceutical development. (US Department of Health and Human Services, Rockville, MD, 2009).
16. US Food and Drug Administration. Guidance for industry: Q9 quality risk management. (US Department of Health and Human Services, Rockville, MD, 2006).
17. US Food and Drug Administration. Guidance for industry: Q10 pharmaceutical quality system. (US Department of Health and Human Services, Rockville, MD, 2008).
18. Rathore, A. S. Roadmap for implementation of quality by design (QbD) for biotechnology products. *Trends in Biotechnology.* **27**, 546–553 (2009).
19. US Food and Drug Administration. Pharmaceutical quality for the 21st century: a risk-based approach progress report. (US Department of Health and Human Services, Rockville, MD, 2007).
20. Leweke, S. & von Lieres, E. Chromatography analysis and design toolkit (CADET). *Computers and Chemical Engineering* **113**, 274–294 (2018).
21. Susanto, A., Knieps-Grünhagen, E., von Lieres, E. & Hubbuch, J. 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 (2008).
22. Bergander, T., Nilsson-Välimaa, K., Öberg, K. & Lacki, K. M. High-throughput process development: determination of dynamic binding capacity using microtiter filter plates filled with chromatography resin. *Biotechnology Progress* **24**, 632–639 (2008).
23. Rathore, A. S. & Bhambure, R. High-throughput process development: process chromatography. In *Protein Downstream Processing: Design, Development and Application of High and Low-Resolution Methods*, 29–37 (Humana Press, New York, NY, 2014).
24. Carta, G. & Jungbauer, A. *Protein Chromatography: Process Development and Scale-Up*. (Wiley-VCH, Weinheim, Germany, 2010).
25. Schmidt-Traub, H., Schulte, M. & Seidel-Morgenstern, A. *Preparative Chromatography*. (Wiley-VCH, Weinheim, Germany, 2012).
26. Hanke, A. T. & Ottens, M. Purifying biopharmaceuticals: knowledge-based chromatographic process development. *Trends in Biotechnology* **32**, 210–220 (2014).
27. Allen, L. The evolution of platform technologies for the downstream processing of antibodies. In *Process Scale Purification of Antibodies*, 365–389 (John Wiley & Sons, hoboken, NJ, 2017).
28. Valente, K. N., Levy, N. E., Lee, K. H. & Lenhoff, A. M. Applications of proteomic methods for CHO host cell protein characterization in biopharmaceutical manufacturing. *Current Opinion in Biotechnology* **53**, 144–150 (2018).
29. Jin, M., Szapiel, N., Zhang, J., Hickey, J. & Ghose, S. Profiling of host cell proteins by two-dimensional difference gel electrophoresis (2D-DIGE): implications for downstream process development. *Biotechnology and Bioengineering* **105**, 306–316 (2010).
30. Kelley, B. D., Tobler, S. A., Brown, P., Coffman, J. L., Godavarti, R., Iskra, T., Switzer, M. & Vunnum, S. Weak partitioning chromatography for anion exchange purification of monoclonal antibodies. *Biotechnology and Bioengineering* **101**, 553–566 (2008).
31. Ludwig, C., Gillet, L., Rosenberger, G., Amon, S., Collins, B. C. & Aebersold, R. Data‐independent acquisition‐based SWATH‐MS for quantitative proteomics: a tutorial. *Molecular Systems Biology* **14**, 1-23 (2018).
32. Lenhoff, A. M. Significance and estimation of chromatographic parameters. *Journal of Chromatography* **384**, 285–299 (1987).
33. Qamar, S., Akram, N. & Seidel-Morgenstern, A. Analysis of general rate model of linear chromatography considering finite rates of the adsorption and desorption steps. *Chemical Engineering Research and Design* **111**, 13–23 (2016).
34. Xu, X. & Lenhoff, A. M. A predictive approach to correlating protein adsorption isotherms on ion-exchange media. *The Journal of Physical Chemistry B* **112**, 1028–1040 (2008).
35. Langmuir, I. The adsorption of gases on plane surfaces of glass, mica and platinum. *Journal of the American Chemical Society* **40**, 1361–1403 (1918).
36. Guélat, B., Ströhlein, G., Lattuada, M. & Morbidelli, M. Electrostatic model for protein adsorption in ion-exchange chromatography and application to monoclonal antibodies, lysozyme and chymotrypsinogen A. *Journal of Chromatogaphy A* **1217**, 5610–5621 (2010).
37. Ladiwala, A., Rege, K., Breneman, C. M. & Cramer, S. M. A priori prediction of adsorption isotherm parameters and chromatographic behavior in ion-exchange systems. *Proceedings of the National Academy of Sciences* **102**, 11710–11715 (2005).
38. Yamamoto, S., Nakanishi, K., Matsuno, R. & Kamikubo, T. Ion exchange chromatography of proteins—prediction of elution curves and operating conditions: I. theoretical considerations. *Biotechnology and Bioengineering* **XXV**, 1465–1483 (1983).
39. Yamamoto, S., Nakanishi, K., Matsuno, R. & Kamikubo, T. Ion exchange chromatography of proteins—prediction of elution curves and operating conditions: II. experimental verification. *Biotechnology and Bioengineering* **XXV**, 1373–1391 (1983).
40. Yamamoto, S., Nomura, M. & Sano, Y. Adsorption chromatography of proteins: determination of optimum conditions. *AIChE Journal* **33**, 1426–1434 (1987).
41. Yamamoto, S., Nakanishi, K. & Matsuno, R. *Ion-Exchange Chromatography of Proteins*. (Marcel Dekker, New York, NY, 1988).
42. Pfister, D., Steinebach, F. & Morbidelli, M. Linear isotherm determination from linear gradient elution experiments. *Journal of Chromatography A* **1375**, 33–41 (2015).
43. Roth, C. M. & Lenhoff, A. M. Electrostatic and van der Waals contributions to protein adsorption: computation of equilibrium constants. *Langmuir* **9**, 962–972 (1993).
44. Roth, C. M. & Lenhoff, A. M. Electrostatic and van der Waals contributions to protein adsorption: comparison of theory and experiment. *Langmuir* **11**, 3500–3509 (1995).
45. Yao, Y. & Lenhoff, A. M. Electrostatic contributions to protein retention in ion-exchange chromatography. 1. cytochrome c variants. *Analytical Chemistry* **76**, 6743–6752 (2004).
46. Liu, B. T. & Hsu, J. P. Analytic expressions for electrical energy and electrical force of two spheres. *The Journal of Chemical Physics* **130**, 1–8 (2009).
47. Guélat, B., Ströhlein, G., Lattuada, M., Delegrange, L., Valax, P. & Morbidelli, M. Simulation model for overloaded monoclonal antibody variants separations in ion-exchange chromatography. *Journal of Chromatography A* **1253**, 32–43 (2012).
48. Guélat, B., Delegrange, L., Valax, P. & Morbidelli, M. Model-based prediction of monoclonal antibody retention in ion-exchange chromatography. *Journal of Chromatography A* **1298**, 17–25 (2013).
49. Khalaf, R., heymann, J., LeSaout, X., Monard, F., Costioli, M. & Morbidelli, M. Model-based high-throughput design of ion exchange protein chromatography. *Journal of Chromatography A* **1459**, 67–77 (2016).
50. Guélat, B., Khalaf, R., Lattuada, M., Costioli, M. & Morbidelli, M. Protein adsorption on ion exchange resins and monoclonal antibody charge variant modulation. *Journal of Chromatography A* **1447**, 82–91 (2016).
51. Zhang, Q., Goetze, A. M., Cui, H., Wylie, J., Trimble, S., Hewig, A. & Flynn, G. C*.* Comprehensive tracking of host cell proteins during monoclonal antibody purifications using mass spectrometry. *MAbs* **6**, 659–670 (2014).
52. Oberholzer, M. R. & Lenhoff, A. M. Protein adsorption isotherms through colloidal energetics. *Langmuir* **15**, 3905–3914 (1999).
53. Yuan, Y., Oberholzer, M. R. & Lenhoff, A. M. Size does matter: electrostatically determined surface coverage trends in protein and colloid adsorption. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **165**, 125–141 (2000).
54. Sader, J. E. Accurate analytic formulae for the far field effective potential and surface charge density of a uniformly charged sphere. *Journal of Colloid and Interface Science* **188**, 508–510 (1997).
55. Nfor, B. K., Ahamed, T., Pinske, M. W. H., van der Wielen, A. M., Verhaert, P. D. E. M., van Dedem, G. W. K., Eppink, M. H. M., van de Sandt, E. J. A. X. & Ottens, M. Multi-dimensional fractionation and characterization of crude protein mixtures: toward establishment of a database of protein purification process development parameters. *Biotechnology and Bioengineering* **109**, 3070–3083 (2012).
56. Khanal, O., Kumar, V., Westerberg, K., Schlegel, F. & Lenhoff, A. M. Multi-column displacement chromatography for separation of charge variants of monoclonal antibodies. *Journal of Chromatography A* **1586**, 40–51 (2018).
57. Egelandsdal, B. Heat-induced gelling in solutions of ovalbumin. *Journal of Food Science* **45**, 570–581 (1980).
58. Strixner, T. & Kulozik, U. Egg proteins. In *Handbook of Food Proteins*, 150–209 (Woodhead Publishing Limited, Cambridge, UK, 2011).
59. Demoruelle, K., Guo, B., Kao, S. McDonald, H. M., Nikic, D. B., Holman, S. C. & Wilson, W. W. Correlation between the osmotic second virial coefficient and solubility for equine serum albumin and ovalbumin. *Acta Crystallographica* **58**, 1544–1548 (2002).
60. Kelly, M. J. & Reithel, F. J. A Thermodynamic analysis of the monomer-dimer association of β-lactoglobulin A at the isoelectric point. *Biochemistry* **10**, 2639–2644 (1971).
61. Reid III, T. J., Murthy, M. R. N., Sicignano, A., Tanaka, N., Musick, W. D. L. & Rossmann, M. G. Structure and heme environment of beef liver catalase at 2.5 Å resolution. *Proceedings of the National Academy of Sciences* **78**, 4767–4771 (1981).
62. Jörnvall, H., Persson, B. & Jeffery, J. Characteristics of alcohol/polyol dehydrogenases. *European Journal of Biochemistry* **167**, 195–201 (1987).
63. Veillon, Claude and Sytkowski, A. J. The intrinsic zinc atoms of yeast alcohol dehydrogenase. *Biochemical and Biophysical Research Communications* **67**, 1494–1500 (1975).
64. Tanford, C. & Roxby, R. Interpretation of protein titration curves: application to lysozyme. *Biochemistry* **11**, 2192–2198 (1972).
65. Sophianopoulos, A. J. & van Holde, K. E. Physical studies of muramidase (lysozyme): II. pH-dependent dimerization. *The Journal of Biological Chemistry* **239**, 2516–2524 (1964).
66. Oberholzer, M. R., Wagner, N. J. & Lenhoff, A. M. Grand canonical Brownian dynamics simulation of colloidal adsorption. *The* *Journal of Chemical Physics* **107**, 9157–9167 (1997).
67. Lenhoff, A. M. Protein adsorption and transport in polymer-functionalized ion-exchangers. *Journal of Chromatography A* **1218**, 8748–8759 (2011).
68. DePhillips, P. & Lenhoff, A. M. Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography. *Journal of Chromatography A* **883**, 39–54 (2000).
69. Yao, Y. & Lenhoff, A. M. Pore size distributions of ion exchangers and relation to protein binding capacity. *Journal of Chromatography A* **1126**, 107–119 (2006).
70. DePhillips, P. & Lenhoff, A. M. Determinants of protein retention characteristics on cation-exchange adsorbents. *Journal of Chromatography A* **933**, 57–72 (2001).
71. Gadam, S. D., Jayaraman, G. & Cramer, S. M. Characterization of non-linear adsorption properties of dextran-based polyelectrolyte displacers in ion-exchange systems. *Journal of Chromatography A* **630**, 37–52 (1993).
72. Creighton, T. E. *Proteins: Structures and Molecular Properties* (W. H. Freeman and Company, New York, NY, 1992).
73. Gray, C. G. & Stiles, P. J. Nonlinear electrostatics: The Poisson-Boltzmann equation. *European Journal of Physics* **39**, 1–27 (2018).
74. Kirkwood, J. G. Theory of solutions of molecules containing widely separated charges with special application to zwitterions. *The Journal of Chemical Physics* **2**, 351–361 (1934).

**Appendix: Future Work**

Section 3 delineates the goals and objectives of the proposed study. An estimated timeline for the completion of these objectives is elaborated below. Figure 5 provides a corresponding Gantt chart.

Figure 5. Estimated Timeline.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Aim | Year 1 | | | | Year 2 | | | | Year 3 | | | | Year 4 | | | |
| Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**One-year Outlook:**

Replicate experiments will be performed for all unrepeated isocratic tests with the three AEX resins. The isocratic AEX data set will also be expanded to include two more levels of pH. This will indicate whether the observed resin retentivity trend is consistent, and it will conclude aim 1. Analogous LGE tests will then be performed for at least one protein on the AEX resins. This will enable assessment of whether the LGE-isocratic relationships exhibit the same behavior across different protein-resin systems. Inverse chromatogram fitting will also be implemented. The equilibrium constant model will be developed concurrently with these experiments. Modifications will be explored to eliminate the previously described adjustable parameter, making the model truly predictive. Point charges will also be introduced into the model.

**Further Outlook:**

LGE data acquisition will be completed, and a method will be developed to accurately relate retention under isocratic and LGE conditions. This will conclude aim 2. The equilibrium constant model will be developed until sufficiently accurate predictions can be made with model protein data. Protein A eluate samples will then be obtained, and the fractionation experiments described in section 3 will be performed. If successful, this will conclude aims 3 and 4. Concurrently with the aim 4 experimental work, model extensions to nonlinear isotherm parameter estimation will be explored. The development of such models in conjunction with related experimental studies constitutes aim 5.