Published in final edited form as:

Biotechnol J. 2015 September ; 10(9): 1400–1411. doi:10.1002/biot.201500089.

Systematic Interpolation Method Predicts Protein Chromatographic Elution from Batch Isotherm Data without a Detailed Mechanistic Isotherm Model

Arch Creasy¹, Gregory Barker², Yan Yao², and Giorgio Carta¹

¹Department of Chemical Engineering, University of Virginia, Charlottesville, VA, USA

²Biologics Process Development, Bristol-Myers Squibb, Hopewell, NJ, USA

Abstract

A method is developed to predict protein chromatographic behavior from batch isotherm using a systematic empirical interpolation (EI) scheme and without relying on a mechanistic description of the dependence of protein binding on protein and salt concentration. Coupled with a lumped kinetic model with rate parameters determined from HETP measurements for non-binding conditions, the EI scheme is used to numerically predict the column behavior. For two experimental cation exchange systems considered in this work, lysozyme on SP-Sepharose-FF and a monoclonal antibody on POROS XS, predictions based on the EI scheme are in excellent agreement with experimental elution profiles under highly overloaded conditions without using any adjustable parameters. A qualitative study of the sensitivity of predicting protein elution profiles to the precision, granularity, and extent of the batch adsorption data is conducted. Based on the results for a hypothetical system, whose properties are comparable to those found in practice for protein cation exchange chromatography the results of this show that the interpolation scheme is relatively insensitive, requiring only that the ranges of protein and salt concentrations in the experimental dataset overlap those under which the protein actually elutes from the column and along with a 10% measurement precision.

Keywords

proteins; overloaded gradient elution chromatography; predictive modeling; ion exchange

1. Introduction

The ability to predict the chromatographic behavior of proteins is desirable for process development and design and to help establish a robust operating space. Beyond a purely empirical approach, 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

Correspondence: Giorgio Carta, Department of Chemical Engineering, University of Virginia, 102 Engineers' Way, Charlottesville, Virginia 22904, USA, gc@virginia.edu.

Conflict of Interest

The authors declare no conflict of interest.

kinetics. In both cases, the model parameters are regressed to a limited number of experimental measurements. For ion exchange chromatography several isotherm models are available including Langmuir [1][2][3][4][5], steric mass action (SMA) [6][7][8], and statistical thermodynamic (ST) models [9][10][11]. Similarly, various rate models are available including models that assume the 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 [12][13][14].

Several examples of application of these models to describe protein chromatography are available in the literature. Gallant et al. [7], 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 determined the isotherm model parameters by combining 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. [4] modeled the chromatographic behavior of a polyclonal antibody on Fractogel SO_3 (65 μm diameter) and POROS HS (50 μm diameter) with a Langmuir model. In this case, the models 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. The former were used to determine the isotherm parameters affecting the linear isotherm region while the latter were used to determine the parameters affecting the non-linear region. Their model could successfully predict the salt gradient elution behavior but only for conditions that resulted in peak concentrations below 12 mg/ml.

Other authors have also used the approach of fitting experimental elution peaks to estimate protein isotherm parameters [2][5][15] [16]. However, such an approach can be limited to predicting conditions that are close to those used to fit the elution peaks if the isotherm model applied does not describe faithfully adsorption equilibrium over a broader range of conditions. For example, Osberghaus et al. [8] observed that different isotherm parameters were determined dependent on whether they were obtained by regressing the results of independent gradient elution and frontal analysis experiments or by fitting a mechanistic model directly to overloaded chromatographic data. The authors determined that although the model parameters could be adjusted to yield an accurate representation of an elution peak, the isotherm parameters were dependent on column characteristics, suggesting that these parameters were not true thermodynamic properties. In a recent review, Liang et al. [17] compared the accuracy of different isotherm models and found that multiple models could provide quantitatively similar predictions of protein chromatographic behavior. Many of these isotherm models could be used to describe empirically the dependence of the adsorption isotherm on protein concentration. However, describing simultaneously the effects of mobile phase composition over a broad range is much more challenging, especially at the high protein concentrations that are encountered during elution for high-capacity resins at high protein loads.

In recent years, significant advances have been made in the determination of protein binding as a function of protein concentration and mobile phase composition over broad ranges of conditions both by using analytical tools that allow measurements with very low sample volumes as well as by implementing automated, high throughput screening (HTS) approaches [18][19][20][21]. While these approaches can generate large amounts of data, the quantitative connection between the data and column performance is not always straightforward. One approach is to use the data to regress model parameters and then predict column performance through simulation, but, obviously, this approach can only be as accurate as the model itself. Another approach, which is explored in this work, 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 can be obtained either manually or through automated HTS equipment. The proposed approach has the advantage that it is not limited by model approximations and can yield predictions whose accuracy depends only on the accuracy of the experimental measurements.

The objective of this paper is thus threefold. The first is to develop an empirical interpolation (EI) method that allows direct use of batch adsorption equilibrium data to predict protein elution behavior at high protein loads without using an isotherm model. Since mass transfer is usually limiting, a lumped kinetic model is used to describe adsorption kinetics with a rate coefficient obtained from isocratic elution experiments for non-binding conditions. The second objective is to validate the proposed EI approach using two different experimental systems, namely lysozyme on the cation exchange resin SP-Sepharose-FF, and a monoclonal antibody (mAb) on the cation exchange resin POROS XS. Gradient elution data are obtained for both systems at high protein loads and for conditions where the eluted peaks attain high protein concentrations. Predictions of protein elution profiles based on the EI approach are compared with those based on prediction using the SMA model and with experimental data. The final objective is to explore the impact of the accuracy and level of detail present in the experimental isotherm data set on the prediction of the elution profiles. While the focus of this work is on the use of batch adsorption isotherm data obtained manually, a future communication will examine the use of data obtained with fully automated HTS equipment.

2. Theoretical development

In this work, we restrict the development to the case of ion-exchange chromatography using salt gradients. However, the approach can be extended in a straightforward manner to systems involving other modes of interaction, such as hydrophobic interaction chromatography, or to separations using different types of gradients, such as pH gradients. We consider systems where protein binding depends only on two independent variables, protein and salt concentration in this case. 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 or to cases where multiple variables (e.g. salt concentration and pH) are manipulated simultaneously are possible. While details on these extensions are beyond the scope of this

work, a generalized approach is briefly discussed in the Supplementary Information (Fig. S1).

In this development, we compare predictions based on the SMA model using parameters optimally fitted to batch adsorption isotherm data, with predictions obtained by direct empirical interpolation of the same data.

2.1. Description of adsorption equilibrium

Protein adsorption equilibrium is defined in this work based on the total protein concentration held within an adsorbent particle, $\bar{q}$. 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, $\bar{q}$ is given by:

$$\bar{q} = q + \varepsilon_p K_D C \quad (1)$$

where q is the bound protein concentration in mg per ml of particle volume, C 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, $K_D \sim 1$, while for a large protein $K_D < 1$ as a result of size exclusion. In many cases, the second term in eq. 1 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 [6] 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 , to that in solution, C :

$$q = \frac{K_e [q_0 - (z + \sigma)q]^z}{(C_{\text{Na}^+})^2} C \quad (2)$$

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 C_{Na^+} is the Na^+ concentration in solution. The linear limit of the isotherm is obtained at high Na^+ concentrations when $q \ll q_0/(z + \sigma)$ and is given by the following equation, which is analogous to the so-called stoichiometric exchange model [22]:

$$q = \frac{K_e (q_0)^z}{(C_{Na^+})^z} C \quad (3)$$

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, $C_{Na^+}^E$ is related to the normalized gradient slope by the following equation [23][24][25][26]:

$$\gamma = \frac{1}{\phi} \int_{C_{Na^+}^0}^{C_{Na^+}^E} \frac{dC_{Na^+}}{A(C_{Na^+})^{-z} + \varepsilon_p (K_D - 1)} \quad (4)$$

where $A = K_e (q_0)^z$. The normalized gradient slope is given by $\gamma = \varepsilon C_{Na^+}^E / CV_G$, where ε is the column void fraction, $C_{Na^+}^E$ 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) can be obtained by regression of γ vs. $C_{Na^+}^E$ data obtained from LGE experiments conducted with different gradient slopes. 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.

The proposed EI approach does not rely on any particular equilibrium model. Instead, it uses a convenient arbitrary function to describe the relationship between q and C at each Na^+ concentration. 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 [27]:

$$q = \frac{q_m (KC)^b}{1 + (KC)^b} \quad (5)$$

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. The model parameters q_m , K , and b are treated as empirical constants obtained by independent regression at each salt concentration.

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, $3 \times n$ parameters are used in the ensuing examples to fit each data set, where n is the number of salt concentrations available in the data set. The data sets used as examples in this work contain $n = 6$ salt concentrations each giving a total of 18 parameters per data set. Batch adsorption data in the region where the isotherm is non-linear are used to regress these parameters, which, in turn, are used to generate a set of isotherm points calculated with eq. 5 at logarithmically-spaced protein concentration values at each salt concentration. Linear isotherm data obtained in the dilute limit from LGE experiments are then combined with the calculated isotherm points into a single data set that describes the effects of both protein and Na^+ concentrations. The two isotherm data sets are combined at the salt concentration where there is agreement between the adsorbed concentration values for the range of protein solution concentrations being studied. Isotherm points at other salt concentrations are calculated by interpolating this data set using piecewise cubic Hermite interpolating polynomials (PCHIP), which was done using MATLAB's *pchip* function. This step effectively captures the shape of the isotherm with respect to a changing salt concentration, maintains a smooth trend and conserves the monotonicity of the data. A PCHIP fit with n points requires $4 \times (n - 1)$ parameters per constant protein concentration interval. In our calculations, we used $n = 6$ salt concentrations from batch isotherm data and 50 logarithmically spaced protein concentration points calculated at each salt concentration giving a total of 1000 PCHIP parameters. A total of 50 uniformly spaced points were used to define the PCHIP curves, giving a total of 2500 isotherm points. These points are passed to MATLAB's *griddedInterpolant* function, which uses a bilinear function to interpolate the adsorbed concentration for a given pair of salt and protein concentration values. Figure 1 illustrates the bilinear interpolation method used for this task. Any queried isotherm point with a salt concentration outside the known data is determined by bilinear extrapolation. If the queried isotherm point is within the known salt range but greater than the known protein solution concentration, the Langmuir-Freundlich fits are used to extrapolate the unknown value.

2.2. Description of adsorption kinetics

The LDF model with a fluid phase concentration driving force was used to describe the kinetics of protein adsorption and desorption according to the following equation [12][28]:

$$\frac{\partial \bar{q}}{\partial t} = k(C - C^*) \quad (6)$$

where k is the LDF coefficient and C^* is the protein solution concentration at equilibrium corresponding to $\bar{q}$. The value of C^* is determined either from the isotherm model (eq. 2) or

from the EI method as a function of the local Na^+ concentration, C_{Na^+} , and the local adsorbed protein concentration value, $\bar{q}$. Because of the much faster diffusion of small ions compared to proteins, equilibrium is assumed for Na^+ between the external solution and the liquid-filled pores.

The value of the LDF coefficient is determined 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 . For these conditions, assuming that band-broadening is controlled by intraparticle diffusion, the van Deemter plot is expected to be linear and LDF rate coefficient is given by [29]:

$$k = \frac{2}{1 - \varepsilon} \left(\frac{k'}{1 + k'} \right)^2 \frac{1}{c} \quad (7)$$

where $k' = \overline{\text{CV}}/\varepsilon - 1$ is the protein retention factor, $\overline{\text{CV}}$ is the mean retention volume in column volumes (CV) units, and c is the slope of the van Deemter plot. The LDF rate coefficient, k , can also be related to the protein effective intraparticle pore diffusivity, D_e , according to the following relationship [12][14]:

$$k = \frac{60D_e}{d_p^2} \quad (8)$$

where d_p is the particle diameter. Based on this relationship, since the dominant protein transport mechanism in the resins considered in this work is expected to be pore diffusion, k is, in turn, expected to be essentially constant as any effect of the protein and Na^+ concentrations on D_e is likely to be small.

2.3. Description of column dynamics

The following plug-flow model is used to simulate the column behavior:

$$\varepsilon \frac{\partial C}{\partial t} + (1 - \varepsilon) \frac{\partial \bar{q}}{\partial t} + u \frac{\partial C}{\partial x} = 0 \quad (9)$$

where t is time and x is the column axial coordinate. This equation is discretized by backwards finite differences and the resulting system of ordinary differential equations was solved using MATLAB's multistep, variable order solver routine, *ode15s*. Axial dispersion, which, in any case is expected to have minimum impact because of the large diffusional mass transfer resistance in protein chromatography, is neglected in eq. 9 but is taken into account approximately by the numerical dispersion introduced by the discretization scheme. Calculations were done using 80 axial discretization points. For these conditions, the effects of numerical dispersion were in fact quite small compared to band broadening caused by the adsorption kinetics so that increasing the number of discretization points beyond 80 had negligible effects on the prediction of protein elution profiles. Calculations were done in

MATLAB R2013b (The Mathworks, Natick, ME, USA) on a Dell Precision T1700, Intel i7 series 3.40GHz.

3. Experimental

3.1. Materials

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 1.

Lysozyme ($M_r \sim 14,300$, pI ~ 11) was obtained from Sigma-Aldrich (St. Louis, MO, USA) and a purified monoclonal antibody ($M_r \sim 150,000$, pI ~ 9) was provided by Bristol-Myers Squibb (Hopewell, NJ, USA). Other chemicals were purchased from Fisher Scientific (Waltham MA, USA) and Sigma-Aldrich (St. Louis MO, USA).

The lysozyme experiments were conducted with buffers containing 10 mM Na_2HPO_4 adjusted to pH 6.5 with phosphoric acid while the mAb experiments were conducted in buffers containing 20 mM total acetate at pH 5.5 prepared by mixing a suitable amount of sodium acetate with acetic acid. In both cases, NaCl was added to these buffers to adjust the total Na^+ concentration for batch isotherm and column experiments. All experiments were performed at room temperature ($22 \pm 2^\circ\text{C}$).

3.2. Methods

3.2.1 Batch adsorption isotherms—Batch adsorption isotherm data at varying protein and Na^+ 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 (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. 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. 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.

3.2.2. Column experiments—Isocratic elution experiments under non-binding conditions (1000 mM NaCl) were performed to determine the LDF rate coefficient based on eq. 7 using 1 cm diameter $\times$ 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 μ L injections of 5 mg/ml protein solutions, and with detection at 280 nm.

Linear gradient elution (LGE) experiments were done with 0.5 cm diameter $\times$ 5 cm long Tricorn columns with gradients from 100 to 500 mM NaCl for lysozyme on SP-Sepharose and 0 to 500 mM NaCl for the mAb on POROS XS also using an AKTA Explorer 10 system. Injection volumes of 50 μ L 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 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 310 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 the mAb.

4. Results and discussion

4.1. Adsorption equilibrium

A plot of the normalized gradient slope γ vs. the Na^+ concentration at elution, $C_{\text{Na}^+}^E$ for the LGE data at low protein loads for lysozyme on SP-Sepharose-FF and for the mAb on POROS XS is shown in Fig. S2 in the Supporting Information. These data show that lysozyme elutes at higher Na^+ concentrations on SP-Sepharose-FF compared to the mAb 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. 4 are $z = 5.17$ and $A = 8.59 \times 10^{13} \text{ (mM)}^{5.17}$ for lysozyme on SP-Sepharose-FF and $z = 10.1$ and $A = 2.92 \times 10^{23} \text{ (mM)}^{10.1}$ for the mAb on POROS XS. The larger effective charge for the mAb 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 [29] and for a different mAb on POROS 50HS, which has a structure similar to that of POROS XS [30].

Figure 2 shows the batch isotherm data for lysozyme on SP-Sepharose-FF with Na^+ concentrations in the range 50 to 420 mM (a) and (c) and for the mAb 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 the mAb. 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 of SP-Sepharose-FF and 0–30 mg/ml for the mAb on POROS XS, cover those expected in LGE experiments at high protein loads. Lines calculated according to the SMA model (eq. 2) with parameters regressed simultaneously to the global set of data are shown in Figs. 2a and 2c. Attempts to use the z - and A -values determined from the LGE experiments at low protein loads and regress only K_e and σ to the high load data resulted in

unsatisfactory predictions of the trends of q with respect to Na^+ concentration (not shown in this work for brevity). 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 [31].

The regressed parameter values are summarized in Table 2. 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. 2a and 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 the mAb/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 2.1 are shown in Figs. 2c and 2d. In this case, each constant- Na^+ isotherm data set was first fitted independently with the Langmuir-Freundlich model (eq. 5) by non-linear regression, yielding the parameters in Table 2. 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. 2c and 2d. For the case of lysozyme/SP-Sepharose-FF in Fig. 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 the mAb/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. 2d). Dashed lines show the curves generated by PCHIP as described in Section 2.1. 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 $\bar{q}$. An inverse interpolating function can also be generated if the independent variables passed to the *griddedInterpolant* function are Na^+ concentration and $\bar{q}$ values. Thus the output would then be the corresponding protein concentration in solution, C^* , which is convenient for use with eq. 6.

4.2. Adsorption kinetics

The van Deemter curves obtained for lysozyme on SP-Sepharose-FF and for the mAb on POROS XS for non-binding conditions (1000 mM Na^+) are shown in Fig. S3 of the Supporting Information. The data follow a linear trend of HETP vs. mobile phase velocity confirming that band broadening is mass transfer controlled. The values of the LDF rate coefficient, k , calculated based on these data according to eq. 7 are $0.185 \pm 0.011 \text{ s}^{-1}$ for lysozyme on SP-Sepharose-FF and $0.096 \pm 0.001 \text{ s}^{-1}$ for the mAb on POROS XS. The corresponding effective diffusivities, D_e , calculated from eq. 8 are $(3.09 \pm 0.18) \times 10^{-7} \text{ cm}^2/\text{s}$ for lysozyme on SP-Sepharose-FF and $(4.01 \pm 0.04) \times 10^{-8} \text{ cm}^2/\text{s}$ for the mAb on POROS XS. Both values are substantially smaller than the corresponding free solution diffusivities of lysozyme ($1.1 \times 10^{-6} \text{ cm}^2/\text{s}$) and of the mAb ($4.5 \times 10^{-7} \text{ cm}^2/\text{s}$) suggesting that diffusion is

significantly hindered in both stationary phases. 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 $10^{-5} \text{ cm}^2/\text{s}$.

4.3. Overloaded linear gradient elution

Figure 3 shows the experimental elution curves obtained at various protein loads and gradient slopes for lysozyme on SP-Sepharose-FF and for the mAb 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 (Figs. 3a and 3b) or with the EI method (Figs. 3c, 3d, 3e, and 3f). In general, 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 the mAb on POROS XS (Figs. 3b and 3d). 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. [32], 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. 3a and 3b, 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. As seen in Figs. 3c and 3d predictions based on the EI scheme 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. 3c), 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. 2a).

Figures 3e and 3f show additional high-protein load LGE experimental data for the lysozyme/SP-Sepharose-FF system with varying gradient slope and for the mAb/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.

4.4. Sensitivity analysis

An important question to be addressed regards the level of detail in the experimental data set needed to obtain reliable predictions. Obviously a data set with greater granularity will provide more faithful predictions, but this has to be balanced against the cost of acquiring

the data, which can be significant even when HTS tools are employed. Additionally, the often-limited supply of protein samples can impose limits on the number of measurements that can be performed. In order to test this effect, a hypothetical isotherm dataset, $\bar{q}$ as a function of protein concentration and different Na^+ concentrations, was generated using the SMA model (eq. 2) with values of $q_0 = 150 \text{ mM}$, $z = 14$, $K_e = 300$, $\sigma = 80$, $\epsilon_p = 0.6$, and $K_D = 0.7$ for a hypothetical 150 kDa protein. Calculated isotherm data points with simulated random error were generated using a normal distribution centered around each $\bar{q}$ -value defined with a standard deviation of 5% such that 95% of the distribution was within 10% of the mean. Other approaches, including Monte Carlo simulations can also be used to determine the effects of experimental variability. However, the simpler approach used in this work was deemed sufficient for an initial assessment. The isotherm data points calculated in this manner are shown in Fig. 4a along with the Langmuir-Freundlich isotherm fits for each salt concentration. The mean absolute error of the fit was 4.5%. Elution profiles were then calculated either using the SMA model with the same parameters used to generate the isotherm data set or with the EI method using the Langmuir-Freundlich model to fit each constant- Na^+ isotherm and bilinear interpolation to calculate $\bar{q}$ at intermediate values of C and C_{Na^+} . Other parameters employed in these calculations are the same as those used for the mAb/POROS XS experimental system described in the previous sections. Figures 4b through 4e shows the elution profiles predicted by the EI method (solid lines) vs. the exact predictions using the SMA model (dashed lines) at varying protein loads. The base case scenario, shown in Fig. 4b, includes 12 points for each isotherm and 8 different Na^+ concentrations of 15, 50, 85, 120, 155, 190, 225, and 260 mM. Experimentally, such a data set would fit on a standard 96 well plate for HTS measurements. Excellent agreement with a mean absolute error of 9.4% is seen in this case between the exact SMA model predictions and those obtained with the EI method. Figure 4c shows the results obtained using only 4 isotherms at 15, 85, 155, and 225 mM Na^+ concentrations each comprising 12 different protein concentrations. This isotherm fit had a mean absolute error of 4.3%. As seen in this figure, the EI method predictions are now much less accurate and deviate from the exact SMA model prediction with a mean absolute error of 16%. Note that in this case isotherm points with a salt concentration higher than 225 mM Na^+ were determined by bilinear extrapolation but since the peak eluted by 235 mM Na^+ , this likely has a small effect on the prediction shown in Fig. 4c. This lower accuracy is due in part to the lack of isotherm information in the range of Na^+ concentrations where the protein actually elutes. Since the peaks elute between 120 and 225 mM Na^+ concentrations, the isotherm information needs to be well defined in this range. Figure 4d shows the results using eight Na^+ concentrations (same as Fig. 4b) and only the first five protein concentration points up to 10 mg/ml. In this case, while the isotherm fit has a mean absolute error of only 3.3%, poor predictions of the elution profile are obtained resulting in a mean absolute error of 37%. Figure 4e shows the predictions based on four isotherms only at 15, 85, 155, and 225 mM Na^+ with just six points per isotherm and a random error of 10% of each $\bar{q}$ -value. The mean absolute error of the EI method prediction is now 50%. These results suggest that the quality and granularity of the isotherm data and the ranges of C and C_{Na^+} selected play important roles in predicting chromatography behavior.

5. Conclusions

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. For the case of protein ion exchange chromatography using salt as the mobile phase modulator, the approach developed in this work uses an empirical isotherm equation to describe the dependence of protein binding on protein concentration at constant salt and a purely empirical interpolation scheme to describe the dependence on salt concentration. For the two experimental systems considered in this work, lysozyme on SP-Sepharose-FF and a mAb on POROS XS, the empirical interpolation scheme yields predictions of the overloaded gradient elution behavior that are in excellent agreement with experimental data obtained in laboratory scale columns. 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 two example cases considered, 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 protein and salt concentrations at which elution actually takes place. A limited number of measurements appears sufficient, however, provided that the experimental precision is 10% or better. Modern HTS-based approaches facilitate the task of obtaining large sets of batch experimental data. The approach developed in this work permits the rapid and effective translation to the column behavior.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by Bristol-Myers Squibb and by the NIH Biotechnology Training Program at the University of Virginia. We are grateful to Xuankuo Xu, Steve Traylor, Nate Domagalski, Siegfried Rieble, and Sibylle Herzer for providing useful comments and materials used in this research.

Abbreviations

EI	empirical interpolation
HETP	height equivalent to a theoretical plate

HTS	high throughput screening
LGE	linear gradient elution
SMA	steric mass action

References

- Orellana CA, Shene C, Asenjo JA. Mathematical modeling of elution curves for a protein mixture in ion exchange chromatography applied to high protein concentration. *Biotechnol. Bioeng.* 2009; 104:572–581. [PubMed: 19593757]
- Forrer N, Butté A, Morbidelli M. Chromatographic behavior of a polyclonal antibody mixture on a strong cation exchanger column. Part II: Adsorption modelling. *J. Chromatogr. A.* 2008; 1214:71–80. [PubMed: 18962650]
- Lan Q, Bassi AS, Zhu J-X, Margaritis A. 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. *Chem. Eng. J.* 2001; 81:179–186.
- Müller-Späth T, Ströhlein G, Aumann L, Kornmann H, Valax P, et al. Model simulation and experimental verification of a cation-exchange IgG capture step in batch and continuous chromatography. *J. Chromatogr. A.* 2011; 1218:5195–51204. [PubMed: 21696747]
- Guélat B, Ströhlein G, Lattuada M, Delegrange L, Valax P, et al. Simulation model for overloaded monoclonal antibody variants separations in ion-exchange chromatography. *J. Chromatogr. A.* 2012; 1253:32–43. [PubMed: 22795935]
- Brooks CA, Cramer SM. Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE J.* 1992; 38:1969–1978.
- Gallant SR, Vunnum S, Cramer SM. Optimization of preparative ion-exchange chromatography of proteins: linear gradient separations. *J. Chromatogr. A.* 1996; 725:295–314.
- Osberghaus A, Hepbaldikler S, Nath S, Haindl M, von Lieres E, et al. Determination of parameters for the steric mass action model—a comparison between two approaches. *J. Chromatogr. A.* 2012; 1233:54–65. [PubMed: 22386131]
- Guélat A, Delegrange L, Valax P, Morbidelli M. Model-based prediction of monoclonal antibody retention in ion-exchange chromatography. *J. Chromatogr. A.* 2013; 1298:17–25. [PubMed: 23759301]
- Oberholzer MR, Lenhoff AM. Protein adsorption isotherms through colloidal energetics. *Langmuir.* 1999; 17:3905–3914.
- Su X, Sun Y. Thermodynamic model for nonlinear electrostatic adsorption equilibrium of protein. *AIChE J.* 2006; 52:2921–2930.
- Guiochon, G., Felinger, A., Shirazi, D., Katti, A. Fundamentals of preparative and nonlinear chromatography. 2. Academic Press; Boston, MA: 2006.
- LeVan, MD., Carta, G. Adsorption and Ion Exchange. In: Green, DW., Perry, RH., editors. *Perry's Chemical Engineers' Handbook*. 8. McGraw-Hill; New York, NY: 2007.
- Carta, G., Jungbauer, A. Protein chromatography: Process development and scale-up. Wiley-VCH Verlag; Weinheim, Germany: 2010.
- Melter L, Butté A, Morbidelli M. Preparative weak cation-exchange chromatography of monoclonal antibody variants I. Single-component adsorption. *J. Chromatogr. A.* 2008; 1200:156–165. [PubMed: 18565531]
- Chan S, Titchener-Hooker N, Bracewell DG, Sørensen E. A systematic approach for modeling chromatographic processes - Application to protein purification. *AIChE J.* 2008; 54:965–977.
- Liang J, Fieg G, Shi Q-H, Sun Y. Single and binary adsorption of proteins on ion-exchange adsorbent: the effectiveness of isothermal models. *J. Sep. Sci.* 2012; 35:2162–2173. [PubMed: 22888059]
- Lacki KM. High-throughput process development of chromatography steps: advantages and limitations of different formats used. *Biotechnol. J.* 2012; 7:1192–1202. [PubMed: 22745056]

19. Nfor BK, Noverraz M, Chilamkurthi S, Verhaert PDEM, van der Wielen LAM, et al. High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. *J. Chromatogr. A*. 2010; 1217:6829–6850. [PubMed: 20880532]
20. 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. *Chem. Eng. Technol.* 2008; 31:1846–1855.
21. Chhatre S, Titchener-Hooker NJ. Microscale methods for high-throughput chromatography development in the pharmaceutical industry. *J. Chem. Technol. Biotechnol.* 2009; 84:927–940.
22. Velayudhan A, Horváth Cs. On the stoichiometric model of electrostatic interaction chromatography for biopolymers. *J. Chromatogr. A*. 1986; 367:160–162.
23. Yamamoto, S., Nakanishi, K., Mastuno, R. Ion exchange chromatography of proteins. Dekker; New York, NY: 1988.
24. Yamamoto S, Nakanishi K, Matsuno R, Kamikubo T. Ion-exchange chromatography of proteins - prediction of elution curves and operating-conditions. 1. Theoretical considerations. *Biotechnol. Bioeng.* 1983; 25:1465–1483. [PubMed: 18551435]
25. S Yamamoto S, Nakanishi K, Matsuno R, Kamikubo T. Ion-exchange chromatography of proteins - prediction of elution curves and operating-conditions. 2. Experimental-verification. *Biotechnol. Bioeng.* 1983; 25:1373–1391. [PubMed: 18548766]
26. Zhu M, Carta G. Adsorption of polyethylene-glycolated bovine serum albumin on macroporous and polymer-grafted anion exchangers. *J. Chromatogr. A*. 2014; 1326:29–38. [PubMed: 24398224]
27. Sips R. On the structure of a catalyst surface. *J. Chem. Phys.* 1948; 16:490–495.
28. Antos D, Kaczmarzski K, Wojciech P, Seidel-Morgenstern A. Concentration dependence of lumped mass transfer coefficients linear versus non-linear chromatography and isocratic versus gradient operation. *J. Chromatogr. A*. 2003; 1006:61–76. [PubMed: 12938876]
29. Carta G, Ubiera AR, Pabst TM. Protein Mass Transfer Kinetics in Ion Exchange Media: Measurements and Interpretations. *Chem. Eng. Technol.* 2005; 28:1252–1264.
30. Pabst TM, Suda EJ, Thomas KE, Mensah P, Ramasubramanyan N, et al. Binding and elution behavior of proteins on strong cation exchangers. *J. Chromatogr. A*. 2009; 1216:7950–7956. [PubMed: 19815222]
31. Marquardt DW. An algorithm for least-squares estimation of non-linear parameters. *J. Soc. Ind. Appl. Math.* 1963; 11:431–441.
32. Monkos K, Turczynski B. A comparative study on viscosity of human, bovine and pig IgG immunoglobulins in aqueous solutions. *Int. J. Biol. Macromol.* 1999; 26:155–159. [PubMed: 10517523]

Figure 1.

Schematic of the bilinear interpolation method used to estimate an unknown value $\bar{q}_{x,y}$ (black dot) from four known values (hollow dots). The points $\bar{q}_{x,y+1}$ and $\bar{q}_{x,y-1}$ (gray dots) are first determined from linear interpolation at a constant protein concentration (dashed lines). The value $\bar{q}_{x,y}$ is then determined from linear interpolation using the values $\bar{q}_{x,y+1}$ and $\bar{q}_{x,y-1}$ at constant Na^+ concentration.

Figure 2.

Batch adsorption isotherms for lysozyme on SP-Sepharose-FF (a) and (c) and for the mAb 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. The solid lines in (c) and (d) are calculated with the Langmuir-Freundlich model with the parameters given in Table 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 determined from the LGE experiments at low protein loads.

Figure 3.

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 the mAb on POROS with 0–500 mM NaCl gradients in (b), (d), and (f). The dash-dot lines represent the Na^+ concentration based on the effluent conductivity. Protein loads given as mg of protein loaded per ml of column volume 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 length and protein load, given as mg of protein loaded per ml of column volume, 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.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4.

(a) Adsorption isotherms generated with the SMA model with $q_0 = 150$ mM, $z = 14$, $K_e = 300$, $\sigma = 80$, $\varepsilon_p = 0.6$, and $K_D = 0.7$. The lines represent the SMA model at Na^+ concentrations of 15, 50, 85, 120, 155, 190, 225, and 260 mM. The dots are 12 calculated data points including $\pm 10\%$ random error at each Na^+ concentration. (b)–(e): Comparison of the LGE profiles for varying protein loads predicted with the SMA model (dashed lines) with predictions based on the EI method (solid lines) using data sets with different detail using: (b) isotherms at 15, 50, 85, 120, 155, 190, 225, and 260 mM Na^+ with 12 equally spaced points each; (c) isotherms at 15, 85, 155, and 225 mM Na^+ and 12 points each; (d) isotherms

at 15, 50, 85, 120, 155, 190, 225, and 260 mM Na⁺ and 5 points each up to 10 mg/ml; and (e) isotherms at 15, 85, 155, and 225 mM Na⁺ and 6 points each. Gradient: 0–500 mM NaCl in 20 CV. Protein loads are the same as those used for the mAb/POROS XS system in Figs. 3b and 3d. The Na⁺ concentration scale shown at right is 50–300 mM Na⁺.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Summary of resin properties

Property	SP-Sepharose-FF	POROS XS
Mean particle diameter (μm) ^(a)	100	50
Column extraparticle porosity, ϵ ^(b)	0.29	0.34
Intraparticle porosity, ϵ_p ^(c)	0.87	0.58
Lysozyme distribution coefficient, K_D ^(d)	0.83	ND
mAb distribution coefficient, K_D ^(d)	ND	0.62
Hydrated particle density	1.02	1.03

^(a) Nominal particle diameter based on resin manufacturer data

^(b) Based on retention of blue dextran and pressure drop for 0.5×5 cm columns

^(c) Based on retention of NaCl

^(d) Based on protein retention under non-binding conditions

Table 2

Adsorption equilibrium parameters

SMA model parameters fitted to entire isotherm data set				
	Lysozyme/SP-Sepharose-FF		mAb/POROS XS	
z	6.65		12.4	
A (mM) ^z	3.58×10^{18}		7.66×10^{29}	
q_0 (mM)	291		152	
K_c	60.3		348	
σ	18.3		83.1	

Langmuir-Freundlich parameters fitted to isotherm data at each Na ⁺ concentration						
	Lysozyme/SP-Sepharose-FF			mAb/POROS XS		
C_{Na^+} (mM)	q_m (mg/ml)	K (ml/mg)	b	C_{Na^+} (mM)	q_m (mg/ml)	b
50	158	3.65×10^3	0.151	17.6	200	7.76×10^9
140	277	2.25×10^{-2}	0.177	93.0	218	3.49×10^0
200	147	3.35×10^{-1}	0.721	118	414	8.79×10^{-4}
260	166	9.02×10^{-2}	1.03	143	151	9.89×10^{-2}
320	139	6.30×10^{-2}	1.42	168	499	2.40×10^{-3}
420	74.0	5.30×10^{-2}	2.32	218	181	3.63×10^{-4}