

Application of Chromatographic Theory for Process Characterization Towards Validation of an Ion-Exchange Operation

O. Kaltenbrunner,¹ O. Giaverini,² D. Woehle,¹ J.A. Asenjo²

¹Amgen Process Development, Amgen Inc, Thousand Oaks, CA; e-mail: oliverk@amgen.com

²Centre for Biochemical Engineering and Biotechnology, Department of Chemical Engineering and Biotechnology, University Chile, Beauchef 861, Santiago, Chile

ABSTRACT: The behavior of ion-exchange chromatography is well understood with respect to changes in ionic strength, pH, resin ligand density, bed height, elution flow rate, and gradient slope. Their relative importance for any specific chromatographic situation varies. When a chromatographic operation utilized to purify a human therapeutic protein is prepared for validation before commercial production, numerous tests have to be performed to establish the relative importance of each operating parameter to define its future role and importance in the framework of in-process controls. This prioritization process is usually performed using a purely empirical approach. In this work, we demonstrate the application of a rational approach based on chromatographic theory to prioritize operating parameters. Both methodologies, empirical and rational, were performed to evaluate a specific ion-exchange chromatography operation for the preparative separation of closely related protein species. We show that the application of the rational approach has the potential to accelerate the evaluation and significantly reduce the amount of analytical testing needed.

KEYWORDS: chromatographic theory; process characterization; ion exchange chromatography; protein

Introduction

The development work that has to be invested to advance the production process of a biologic molecule to a state acceptable to regulatory agencies for commercial production is substantial. Typically, after a decision to commercialize a product candidate is reached based on early clinical results, the production process is optimized for commercial production at the anticipated production scale. While

material produced using the optimized process is tested in the pivotal clinical trial, the process is characterized in preparation for validation and regulatory submission. Figure 1 outlines an overall flow of this commercialization process. The methodology for process optimization and selection of process conditions using mathematical models has been described in a recent paper (Shene et al., 2006). One important aspect of the commercialization process is the establishment of the rationale and justification for sufficient and appropriate in-process controls for future production. Regulatory agencies expect the manufacturer to identify all critical process input and output parameters that should be monitored or controlled within predefined ranges during process validation as well as commercial production to ensure the product is of the desired quality (ICH, 2006a). Additionally, justification is expected for considering other control parameters as non-key to process performance as these will not be monitored and controlled as closely on a routine basis. For all parameters of all process operations, these justifications and acceptance ranges have to be established prior to process validation (ICH, 1999; Gardner and Smith, 2000). The methodology described in this paper is a replacement of most steps of the sub-process flow chart for screening experiments outlined in Figure 1.

The setting of appropriate validation ranges is important for future process success with respect to both production economics and regulatory compliance. The information and knowledge gained from process development, characterization studies, and manufacturing experience provide scientific understanding to support the establishment of the process design space. Working within the design space is not considered as a change while movement out of the design space is considered to be a change and would normally initiate a regulatory postapproval change process (ICH,

Overall Process Flow to Commercialization

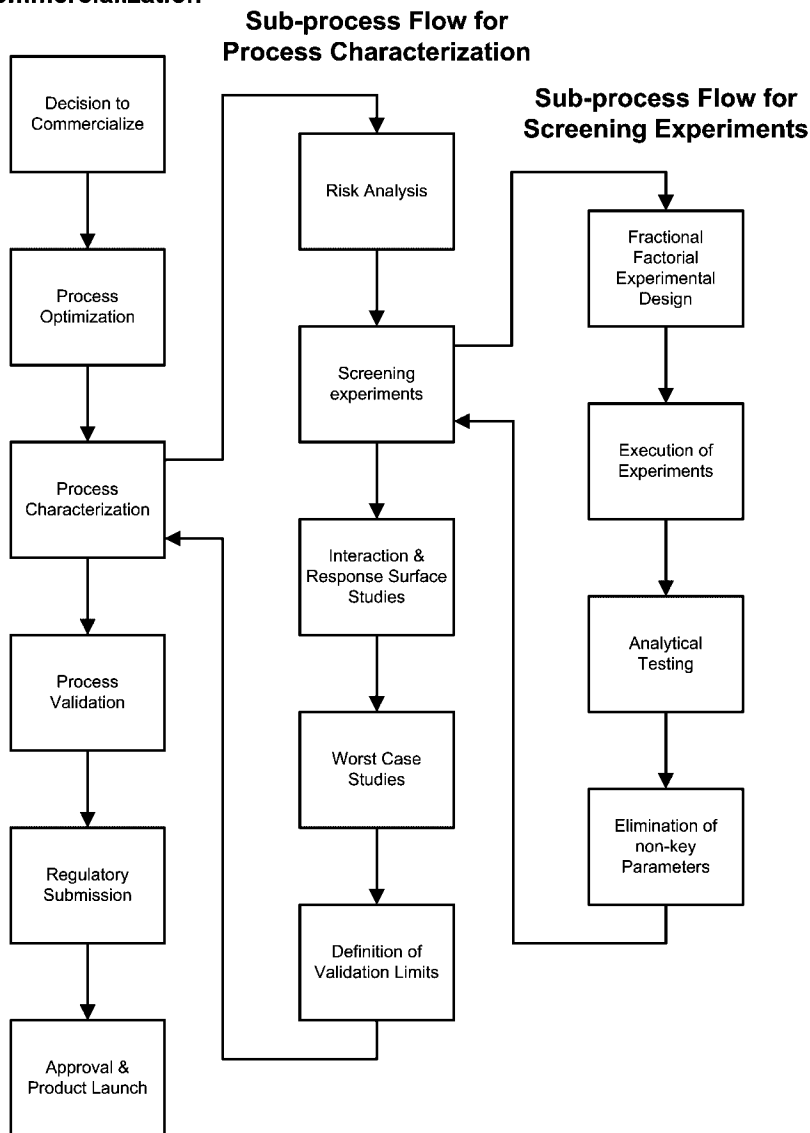


Figure 1. Process flow chart from a decision to commercialize a biologic to product launch. The sub-processes of process characterization and screening experiments are outlined.

2006a). The setting of too narrow operating ranges during validation increases the likelihood of accidental deviation from, and restricts the possibility to make intentional small changes to, the specified operating range of the validated process in future commercial operation. Too wide operating ranges increase the likelihood of failing product quality acceptance criteria during validation and commercial operation. The validation range of an operating parameter has to be wider than the intended normal operating range, but within the range to meet the acceptance criteria.

Abundant information about process response to variations in operation parameters is typically available from the initial process development. However, more often than not

this information is not sufficient to justify the setting of validation criteria. Therefore, the response of a process output to variations in input parameters is investigated in some detail before process validation in a process characterization phase. To prioritize studies, characterization usually starts with a risk analysis based on the available process information and facility constraints (ICH, 2006b). As a result, insignificant operating parameters are eliminated from further analysis and the potential risk factors for the process are identified (DeSain and Sutton, 2000). After the elimination of non-important factors, the remaining factors are studied in fractional factorial experimental designs to identify the true dominating factors while

minimizing the number of experiments (Kelley, 2000). The characterization can be finalized by augmenting the experimental designs to identify interactions and response curvature for the remaining dominating parameters.

In the case of ion-exchange chromatography operations, the number of potentially important factors is large, since the potential for eliminating operating parameters during risk analysis is limited. Typically, buffer conditions like ionic strength and pH, flow rate, column load amount and quality, ligand density, pool collection criteria, bed height, and gradient slope in the case of a linear gradient elution cannot be eliminated a priori. This leads to large experimental designs and substantial analytical support requirements as each experiment may have to be analyzed with several assays to test for purity of the protein with respect to product variants, host cell contaminants, and process impurities (Winkler, 2000). On the other hand, extensive theoretical knowledge about the impact of parameters on ion-exchange behavior is available (Helfferich, 1962). The normal application of experimental screening designs is the selection of a subset of dominating effects for a regression model when little theory is available to guide the selection.

In this work, we demonstrate the applicability of chromatographic theory to reduce the experimental and analytical requirements during process characterization of a chromatographic operation. We performed both the typical screening of parameters in a fractional factorial experimental design and, independently, developed a theoretical model as described by Yamamoto et al. (1988) for a particular ion-exchange chromatography operation. While the original model is based primarily on the response of protein elution to ionic strength, flow rate, and gradient slope, extensions have been attempted to account for protein load concentrations (Jungbauer and Kaltenbrunner, 1996), pH, and resin ligand density (Watler et al., 2003). In this work, a model for linear chromatography accounting for buffer ionic strength, pH, resin ligand density, bed height, flow rate, and gradient slope is discussed. With the application of a linear model, all input parameters except column load, load quality, and pool collection criteria can be analyzed. The typical output parameters of protein purity and host cell contaminant levels cannot be described in the model, but resolution and peak position can be used instead as surrogate parameters. The results of the modeling approach versus the fractional factorial screening experiments are compared and it is shown that the model can be applied to reduce substantially the effort of parameter screening.

Theoretical Aspects

Fractional Factorial Screening

The typical approach for initial parameter screening in process characterization is the application of a fractional

factorial experimental design of resolution III, i.e. a design that will be able to identify main effects, but all main effects have aliases with two-factor interactions. These two-factor interactions must be assumed to be zero for the main effects to be meaningful (Montgomery, 2000). A resolution III design is a powerful approach as long as the assumption of sparse effects can be made. In that case, the resolution of the design can be increased for the dominant main effects by eliminating insignificant main effects. On the other hand, if all main effects in a resolution III design are significant, they cannot be distinguished from potential two-factor interactions of other significant main effects. The chance of all main factors being significant is high due to the preceding risk analysis that already eliminated all likely insignificant factors and due to the multiple responses under investigation. Factors that are insignificant for one response can significantly affect another. To get useful process information, the experimental screening designs have to be expanded to resolution IV designs; no main factor has aliases with two-factor interactions (Montgomery, 2000).

Another often-ignored quality criterion of experimental screening design results is the resolution with respect to the minimum detectable difference of a response. The desirable least significant value should be defined upfront to be able to distinguish process significance from statistical significance. However, from a basic experimental design without replicates no information is available on variability. The model is saturated and no degrees of freedom are available for error.

This fractional factorial methodology of selecting a subset of effects for a regression model is normally used when there is little theory to guide the selection. This method has been of practical use for many years despite the fact that it violates standard statistical assumptions because the resulting model has been selected rather than tested.

For the usual nine to ten potential factors such as ionic strength (1), pH (2), flow rate (3), column load amount (4) and quality (5), ligand density (6), pool collection criteria front (7) and back (8), bed height (9), and gradient slope (10), a resolution III design requires 16 experiments, while a resolution IV design requires 32 experiments. Specific factors like load quality and column bed height can drastically increase the experimental complexity for an automated chromatography system. To minimize this difficulty, the experimental matrix can be split into two or more independent subsets without increasing the experimental effort. Splitting experimental designs into subsets reduces information on two factor interactions that are not considered a main output in this early screening phase.

Ion-Exchange Model

According to the law of mass action, the equilibrium of a protein (P) and its counter ion (I) interacting with an

ion-exchange resin can be expressed as

$$z_I \cdot P + z_P \cdot \bar{I} \leftrightarrow z_I \cdot \bar{P} + z_P \cdot I \quad (1)$$

where z denotes the charge of an ion and bars denote ion-exchanger bound ions.

For this equilibrium, the equilibrium constant $K_{P,I}^e$ can be expressed based on activities (a) as

$$K_{P,I}^e = \frac{\bar{a}_P^{|z_I|} \cdot a_I^{|z_P|}}{a_P^{|z_I|} \cdot \bar{a}_I^{|z_P|}} \quad (2)$$

For a monovalent counter-ion and the introduction of an activity coefficient (γ) as the ratio of the activity (a) and the fluid phase concentration (c) and the absorbed phase concentration (n), Equation (2) can be expressed as

$$K_{P,I}^e = \frac{n_P}{c_P} \left(\frac{c_I}{n_I} \right)^{|z_P|} \cdot \frac{\bar{\gamma}_P}{\gamma_P} \left(\frac{\gamma_I}{\bar{\gamma}_I} \right)^{|z_P|} \quad (3)$$

From this the distribution coefficient (K) of the protein P can be expressed as

$$K = \frac{n_P}{c_P} = K_{P,I}^e \cdot \Gamma \cdot \left(\frac{n_I}{c_I} \right)^{|z_P|} \quad (4)$$

where the lumped activity coefficient term Γ is

$$\Gamma = \frac{\bar{\gamma}_P}{\gamma_P} \left(\frac{\bar{\gamma}_I}{\gamma_I} \right)^{|z_P|} \quad (5)$$

For electroneutrality on the resin, the concentration of available stationary ionic groups (n^s) has to be balanced by protein and counter ion as

$$n^s = |z_P| \cdot n_P + n_I \quad (6)$$

For linear chromatographic conditions $|z_P| \cdot n_P \ll n_I$ and Equation (4) and Equation (6) can be combined to

$$K = \frac{n_P}{c_P} = K_{P,I}^e \cdot \Gamma \cdot \left(\frac{n^s}{c_I} \right)^{|z_P|} \quad (7)$$

This derivation was outlined, expanded and simplified by Yamamoto et al. (1988). The nomenclature of symbols is aligned with the comprehensive review of ion-exchange by LeVan et al. (1997). An empirical relation of the distribution coefficient K and ionic strength c_I was described by Yamamoto et al. as

$$K(I) = A \cdot I^{-B} + K_{crt} \quad (8)$$

where K_{crt} takes the sieving effect under non-binding conditions into account, B represents the number of

interacting charges of the protein $|z_P|$, I represents liquid phase counter ion concentration c_I , and A is a lumped parameter including equilibrium constant, activities of species, and resin ligand density.

$$A = K_{P,I}^e \cdot \Gamma \cdot (n^s)^{|z_P|} \quad (9)$$

Equation (9) allows the expansion of Equation (8) to account for changes in resin ligand density in the form of

$$K(I, n^s) = A' \cdot \left(\frac{n^s}{I} \right)^B + K_{crt} \quad (10)$$

A' is the combined parameter of the equilibrium constant ($K_{P,I}^e$) and the activity coefficient term (Γ).

To include the effect of pH in the model, the effect of pH on n^s , $|z_P|$, and A' has to be considered. For the case of a separation on a strong ion-exchange resin, the operating pH will be more than two pH units apart from the pK of the ionogenic ligand. In that case, the apparent resin ligand density (n^s) will be minimally affected by pH changes and pH changes will mainly change the number of binding sites of the protein $|z_P|$ interacting with the resin. This dependency of the number of binding sites on solution pH can be found in the literature for several standard proteins (Ståhlberg and Jönsson, 1999; Yamamoto and Ishihara, 1999). While the underlying theory is complex and rigorously developed (Ståhlberg and Jönsson, 1999; Shen and Frey, 2004), for the purpose of the evaluation of chromatographic behavior within the narrow operational pH range of a commercial protein chromatographic operation the dependence of $|z_P|$ on pH can be approximated by a quadratic polynomial.

The effects of pH on A' are complex. From Equation (5) it is obvious that Γ will change with pH as the number of interacting charges changes. $K_{P,I}^e$ can be derived from the Donnan potential as a function of swelling pressure (π) of the ion-exchanger and partial molar volumes (v) of the solutes as

$$K_{P,I}^e = \exp \left[-\frac{\pi}{RT} \cdot (v_P - v_I \cdot |z_P|) \right] \quad (11)$$

This indicates that $K_{P,I}^e$ is also a function of pH as swelling pressure changes with pH (Michaeli and Katchalski, 1957). However, this change is orders of magnitude smaller than the change of the other terms with pH. Consequently, Equation (10) can be expanded to include dependence on pH as

$$K(I, n^s, pH) = p \cdot \left(q \cdot \frac{n^s}{I} \right)^{B(pH)} + K_{crt} \quad (12)$$

where

$$p = K_{p,I}^e \cdot \frac{\gamma_P}{\bar{\gamma}_P}, \quad (13)$$

$$q = \frac{\bar{\gamma}_I}{\gamma_I}, \quad (14)$$

and

$$B(pH) = r + s \cdot pH + t \cdot pH^2 \quad (15)$$

In the theory described by Yamamoto et al. the parameters A and B can be determined from linear gradient data according to

$$GH = \frac{I_R^{B+1}}{A \cdot (B+1)} \quad (16)$$

where GH is the normalized gradient slope of a gradient of length V_G from I_0 to I_{end}

$$GH = \frac{(I_{end} - I_0) \cdot (V_t - V_0)}{V_G} \quad (17)$$

and I_R is the ionic strength at the peak maximum during gradient elution. Using Equation (16) to determine parameters p , q , r , s , and t for multiple species (i), the equation can be expanded to

$$GH = \frac{I_R^{B(pH,i)+1}}{A(pH, n^s, i) \cdot (B(pH, i) + 1)} \quad (18)$$

with

$$A(pH, n^s, i) = p_i \cdot (q \cdot n^s)^{B(pH,i)} \quad (19)$$

$$B(pH, i) = r_i + s_i \cdot pH + t_i \cdot pH^2 \quad (20)$$

Furthermore, the ionic strength at the peak maximum can be calculated for any combination of gradient slope (GH), pH , and ligand density (n^s) as

$$I_R(GH, pH, n^s) = (A(pH, n^s, i) \cdot (B(pH) + 1) \cdot GH)^{-\frac{1}{B(pH)+1}} \quad (21)$$

The relation in Equation (21) can only hold for the assumption that protein elution is completely dominated by the gradient and the absorbed species is virtually immobilized on the stationary phase at the buffer condition of load, wash, and gradient initiation. If this assumption cannot be satisfied, the peak profile has to be generated iteratively by numerical procedures as described by Yamamoto et al. (1988).

Resolution between two species in a chromatographic separation is defined as

$$R_S = \frac{t_{R,2} - t_{R,1}}{2 \cdot (\sigma_1 + \sigma_2)} \quad (22)$$

Here t_R and σ are retention time and variance of the two species under consideration. Yamamoto et al. (1988) have shown that, for ion-exchange chromatography, in the case of a linear gradient elution of slope G , this definition of resolution can be expressed for two species eluting at ionic strength $I_{R,1}$ and $I_{R,2}$ as

$$R_S = \frac{I_{R,2} - I_{R,1}}{2 \cdot G \cdot (\sigma_{\theta,1} + \sigma_{\theta,2})} \quad (23)$$

Here the variance of the peaks is expressed in dimensionless time ($\theta = t \cdot u / L$) if L is column bed height and u is linear flow velocity. Finally, by replacing variances with the appropriate equations for band broadening in linear gradient elution, resolution can be expressed as (Yamamoto et al., 1988):

$$R_S(G, pH, n^s, u, L) = \frac{(I_{R,1} - I_{R,2}) \cdot \sqrt{L \cdot \left(-\frac{dK}{dI}\right)_{I=I_{R,1}} \cdot H}}{(1 + H \cdot K_{I_{R,1}}) \cdot \sqrt{8 \cdot G \cdot \left(\frac{2 \cdot D_L}{u} + \frac{d_p^2 \cdot H \cdot K_{I_{R,1}}^2}{30 \cdot (1 + H \cdot K_{I_{R,1}})^2 \cdot D_{crt,1} \cdot K_{crt,1}}\right)}} \quad (24)$$

$D_{crt,1}$ is the gel-phase diffusion coefficient of protein 1 as obtained by the moment method under non-binding conditions. D_L is the effective longitudinal dispersion coefficient.

Equation (24) can be utilized to assess the relative importance of changes in ionic strength, pH , resin ligand density, flow velocity, and bed height on the separation.

Materials And Methods

Fractional Factorial Screening

In the traditional parameter screening approach, 11 parameters were screened, and of these, 7 parameters were relevant for the comparison to the modeling approach. These parameters were resin lot (1), load concentration (2), load amount (3), equilibration ionic strength (4), gradient slope of elution conductivity (5), buffer pH (6), and flow rate (7). The experiments were performed in two independent matrices of eight experiments each for subsets of factors. All main effects are identifiable in this design, but two-factor interactions have aliases. For the purpose of this comparison, the two matrices were combined as blocks of experiments and analyzed for the factors of the original

designs. This allows having all factors of interest in one analysis. As responses to load concentration and load amount cannot be described with the simple model used in this study, these responses are not presented. These experimental results are presented only in the context of the comparison to the modeling data.

Load Material for Model Definition Experiments

The load material for the gradient experiments was enriched in impurities to allow the detection of pre-peak and post-peak impurities in the UV profile directly without the need of analytical testing. The enriched load was prepared by combining pre and post collection fractions of a large-scale separation of the same operation. This combined material was concentrated and buffer exchanged to create the final load material. The column load amount was chosen such that pre and post peak impurities were detectable in the UV profile of the gradient experiments as separate peaks. In a typical elution profile for a preparative separation, the pre-peak and post-peak impurities merge with the main peak (Fig. 2A). In the elution profiles created by the enriched load with low loading, the species are distinguishable from the UV profile (Fig. 2B). This allows the direct retention measurement of the main product species and pre-peak and post-peak impurities without the need for analytical testing.

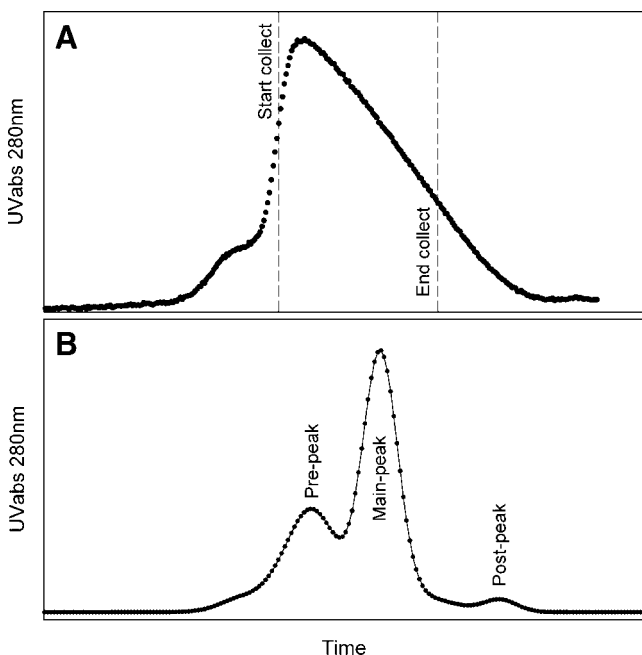


Figure 2. Comparison of chromatograms. (A) Typical elution profile for preparative column load. Start and end collection points are indicated as dashed lines. Pre-, main, and post-peak are merged. (B) Representative elution profile with the enriched column load material and low loading amount used for the model definition experiments. Pre-, main, and post-peak are indicated on the UV profile.

Results

Linear gradient experiments were performed at different gradient slope and different buffer pH. As shown by Yamamoto et al. (1988), the elution data of varying gradient slope for a specific buffer condition can be used to determine the dependence of the distribution coefficient on salt concentration and, ultimately, to predict the peak profiles. An overlay of the gradient data in this study and simulation results are shown in Figure 3. To assess the dependence of the elution behavior on pH, experiments were performed at four different pH values around the target pH. At each buffer pH, five to nine experiments at different gradient slopes were performed to build a data set of 29 gradient experiments. For each experiment, the ionic strength at peak maximum was determined for the pre, main, and post-peak in the elution profile. These experimental data were fitted to Equation (18) for all protein species (i) simultaneously to determine parameters p_i , q , r_i , s_i , and t_i of Equations (12) and (15). This resulted in the fit of 13 parameters simultaneously to the 74 available data points. Figure 4 shows a representative fit of gradient data for the subset of main peak data. Equivalent data fits are performed simultaneously for pre and post peaks. All fitted parameters are given in Table I. Utilizing Equations (19) and (20), the dependence of parameters A and B of Equation (8) on pH can be visualized (Fig. 5).

Similarly, ionic strength at peak maximum (I_R) can be calculated for any gradient slope and pH using Equation (21). However, if the gradient condition is such that the protein elutes very early in the gradient, the moving protein zone has not yet reached its maximum velocity. In this case, Equation (21) does not hold and retention can only be predicted by numerical procedures. The deviation between the numerical procedure and Equation (21) increases with increasing gradient slope (see Fig. 6) and increasing wash volume between load and elution. In general, we can assume

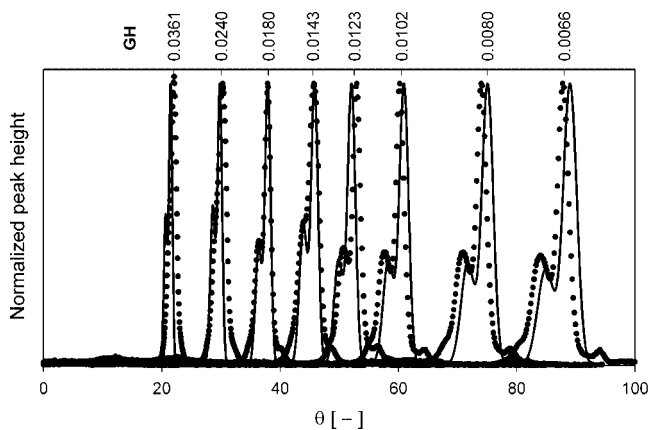


Figure 3. Comparison of small-scale experimental data (symbols) and simulation results (lines) for eight gradient slopes (GH). Retention is expressed in normalized time, $\theta = t \cdot u/L$.

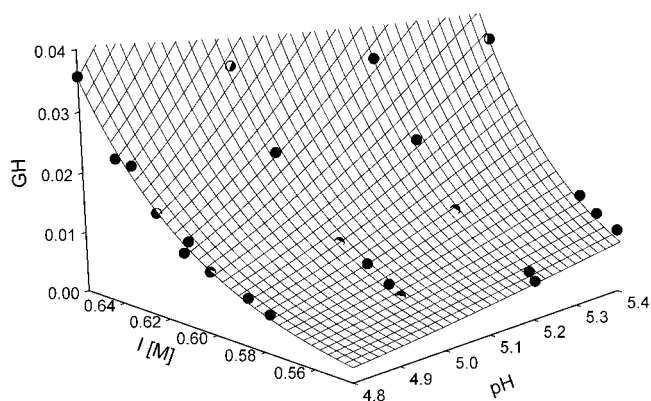


Figure 4. Surface plot of gradient data of gradient slope (GH) vs. pH and ionic strength at elution peak maximum. Data are for the main peak species. The graph represents one surface of the simultaneous fit with Equations (18)–(20). Fit results are shown in Table 1.

that Equation (21) can be applied if the distribution coefficient of a species is at least 200 at the beginning of the gradient (K_{I0}). Additionally, this sensitivity emphasizes the importance of the duration of the isocratic wash phase before gradient initiation in case of insufficient retention at the initial buffer condition.

Within the ranges of input parameters (Table II), all values of K_{I0} are above 200 with the exception of conditions combining low ligand density (n^s) with high initial salt concentration (I_0). While this exception immediately identifies the condition as inappropriate and unstable for the process, for the purpose of this study all ionic strengths at peak maximum (I_R) were calculated numerically to eliminate the deviation caused by the simplified approach using equation 21.

Parameter Prioritization by Modeling

Resolution between the species and retention of all species were calculated for all combinations of parameters from Table I at high, low, and center point level. This leads to a full factorial set of 2^5 data points that were created from numerical procedures and equation 24. These 32 data points can now be analyzed to rank the individual contribution of single factors and their interactions to the resulting resolution between the species and to retention of the

Table I. Model parameters for the three protein species of interest.

Species	$p (\times 10^6)$	r	s	t
Pre-peak	0.49	-18.5	9.9	-0.77
Main-peak	5.48	-26.9	13.8	-1.15
Post-peak	2.21	-24.7	12.9	-1.10

Experimental data for three proteins are fit simultaneously with Equations (18)–(20). Parameter $q = c_I/n_I = 1.098$, $n^s = 0.17$ mol/L resin, $R^2 = 0.974$, $N = 74$.

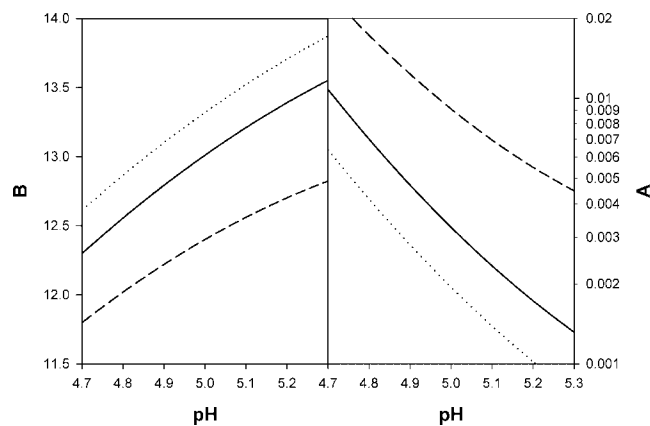


Figure 5. Dependence of parameters A and B of Equation (12) on pH. Dotted line: pre-peak, solid line: main-peak, dashed line: post-peak. $n^s = 0.17$ mol/L resin.

individual peaks. To compare the size of effects, a pareto plot of orthogonalized estimates can be used. From the pareto plot ranking of 31 factors and interactions (not shown), the combined effects of ligand density, I_0 , pH, flow velocity, and gradient slope accounted for more than 60% of the total variability due to main factors. Table III ranks these main factors according to the strength of their relative effect on the resolution of pre-peak from main peak, the resolution of main peak from post-peak, and the retention of main peak. This analysis identifies the three parameters I_0 , pH, and

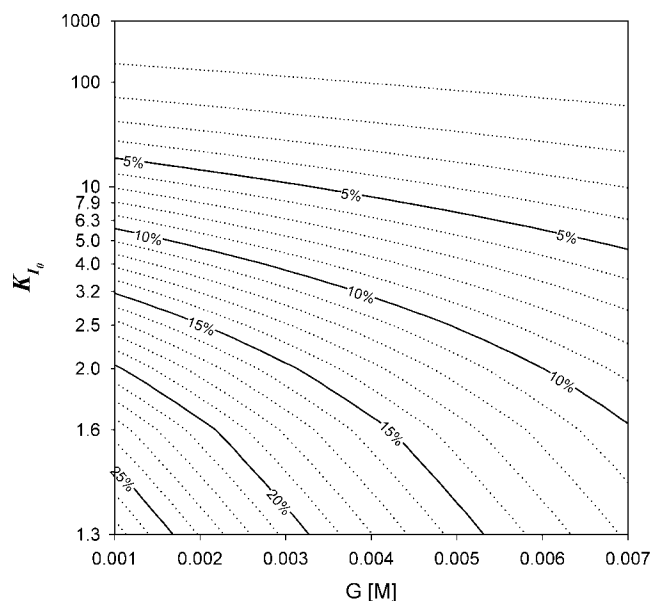


Figure 6. Deviations between I_R calculated by Equation (21), and by numeric iteration for three proteins as a function of distribution coefficient and gradient slope G . Increasing gradient slope and decreasing K value at initial ionic strength (I_0) decrease the accuracy of simplified Equation (21) compared to the numerical, iterative calculation.

Table II. Input parameters and ranges for the full factorial model of resolution and retention.

Factor	Low	High	Units
u	90	110	cm/h
I_0	0.405	0.495	M
G	0.00495	0.00605	M/CV
pH	4.8	5.2	—
Ligand density	0.15	0.20	M/L resin

Table III. Relative contribution (expressed as %) of factors to the outcome of resolution and peak shift expressed as shift in I_R .

Factor	Resolution		ΔI_R (M)
	Pre–Main	Main–Post	
Ligand density	54	33	73
I_0	34	36	9
pH	4	20	15
u	6	7	0
G	2	4	3

Contributions are from linear regression analysis of modeling results using a full factorial matrix of input data over the ranges given in Table II.

ligand density as the dominating factors for resolution and peak position, while gradient slope (G) and flow velocity (u) have only minor influence.

Fractional Factorial Parameter Screening

Two independent factorial designs of eight experiments each were performed. The purpose of these experimental designs was to identify significant main effects. The step recoveries and analytical results for product purity were analyzed within their respective experimental designs. While the statistical power of the two experimental data sets was different, for the purpose of this analysis, the sizes of the main effects were combined and normalized by the average response of their respective data set. Table IV ranks these main factors according to the strength of their relative effect on step recovery and key product quality indicators. This analysis identifies the three parameters I_0 , pH, and ligand density as the dominating factors for resolution and peak

Table IV. Relative contribution (expressed as %) of factors to changes in step recovery and key elution pool quality indicators.

Factor	Quality attributes			Step recovery
	Pre-peak	Main-peak	Post-peak	
I_0	58.6	63.5	50.3	51.1
pH	22.8	7.5	16.6	38.0
Ligand density	9.0	14.0	13.8	6.2
G	4.3	8.5	10.1	2.2
u	5.3	6.6	9.2	2.5

Contributions are given relative to the average output value. Values are from linear regression analysis of two separate fractional factorial experimental data sets for the input ranges given in Table II.

position, while gradient slope (G) and flow velocity (u) have only minor influence.

Discussion

Parameter screening by chromatographic modeling and by fractional factorial experimental design both rank the ionic strength at the beginning of the elution gradient, pH, and resin ligand density as the parameters that have the highest effect on separation behavior (Fig. 7.) Similarly, both methods identify gradient slope and flow rate—within the range considered in this study—as parameters that have a lesser effect on the separation behavior. This comparison indicates that the modeling approach outlined by Yamamoto et al. (1988) can be applied for the initial screening of operational parameters during process characterization.

For the typical parameter classification by fractional factorial experimental design, numerous samples have to be analyzed for each experiment to establish the response to the variations in the inputs. In contrast, in the case of parameter classification by modeling, no analytical testing of samples had to be performed. This elimination of analytical testing could be achieved by preparing column load material enriched in impurities and under-loading the column. Consequently, the resolution achieved during elution was sufficient to identify the dominating impurities eluting just prior and post the main species directly from the UV trace of the elution. Therefore, although 29 gradient experiments were performed to define the model, versus 16 fractional factorial experiments to screen for main effects, the overall time and resource savings was much greater by using the model.

In the modeling approach, only dependence on pH, initial ionic strength (I_0), and gradient slope (G) are investigated explicitly. The remaining model parameters are addressed based on the theoretical dependencies. The dependence of the separation behavior on resin ligand density is calculated in Equation (21). The experiments are only performed for one resin lot and the dependence of residence time and resolution is assessed by extrapolation. This is especially useful as the availability of resin samples over the entire specification range accepted by the resin vendor is typically limited. The dependence on flow rate is addressed while calculating resolution by Equation (24). It should be noted that the dependence on column bed height could be assessed in a similar fashion during calculation of resolution. In our work, column bed height was not considered in the fractional factorial experimental design, as it was considered insignificant in the initial risk analysis. When added to the parameter classification by modeling, the effect of bed height variability $\pm 10\%$ of target was of similar magnitude as the effects of gradient slope and flow rate.

In addition to the correct classification of parameters and the elimination of costly analytical testing, the modeling approach has other advantages. In the model, the test matrix can easily be expanded with a third parameter level to

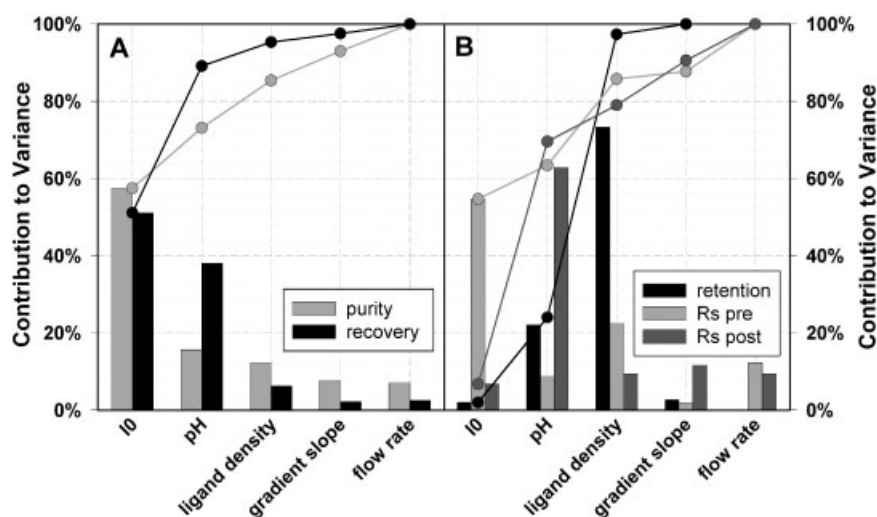


Figure 7. Comparison of results from (A) traditional screening approach and (B) modeling approach. Bars indicate individual contributions of variance and lines and symbols indicate cumulative contributions to variance of the input parameters.

investigate the curvature of the responses. This can be helpful guidance during experimental design for the final phases of process characterization, when interactions of factors and curvature of responses are investigated.

Some limitations of the modeling approach should be kept in mind. As the model does not account for nonlinear effects of column overloading, no information of the effect of loading can be included in this assessment. If a particular separation were dominated by nonlinear effects, this analysis would be less realistic in its predictions. However, even in this case, the prioritization of other parameters should hold and be useful for eliminating non-significant factors from more detailed further studies.

The model parameters are not identical to the parameters typically monitored and analyzed during development and characterization of a separation operation. While during development quality indicators based on analytical testing and product recovery are usually studied, in the case of our non-representative column loading with respect to both load amount and quality these parameters would be virtually meaningless. As demonstrated in Figure 7, in this analysis we investigated the effects of input parameters on surrogate responses, specifically retention and resolution. The effects of input parameters on individual species retention time and resolution between the species of interest and early and late eluting impurities should be a good approximation of the separation behavior under normal operating conditions. Any operational parameter that significantly affects the relative elution position of a species, under these model conditions, will very likely also have an effect on the elution behavior under normal operating conditions. Any operational parameter that has no significant effects on relative elution positions of a species will very likely also have no effect on the elution behavior under normal operating conditions. As differences in elution behavior between the

target protein and host cell impurities are typically much larger than the differences between the target protein and the pre-peak and post-peak impurities analyzed in this study, this general assumption should also be applicable to the case of host cell or process impurities.

Since the modeling approach uses lumped parameters obtained by fitting of a certain number of experimental data, the physical significance of the individual model parameters is relatively limited and wider interpretations and applications of an established model for a specific operation have to be approached carefully. While Figure 5 demonstrates that the values for model parameters *A* and *B* are of a realistic magnitude they are not expected to reflect the true physical value of the underlying fundamental parameters.

Conclusion

The potential for applying a model describing the behavior of an ion-exchange separation has been demonstrated. While this work is limited to a linear gradient elution situation, the same model was applied successfully to describe elution by a stepwise change in ionic-strength (Ishihara and Yamamoto, 2005) and isocratic elution (Watler et al., 2003). This is an indication that this modeling method would prove useful for process characterization for all types of ion-exchange operations. Such a modeling approach can be used to reduce a screening matrix of 10 parameters to a matrix of 4. The remaining parameters are load amount and quality, and pooling criteria on both front and back of the elution profile. Using this work as an example, the effect of load amount and quality can quickly be tested in separate experiments. This screening can then be followed by rigorous characterization of the parameters identified as influential while analyzing pseudo-pools of

different pooling criteria for each elution condition to include the characterization of pooling criteria that are critical for both step recovery and product quality.

The study of mock loads enriched in the significant impurities to levels detectable on standard chromatography equipment proved very useful. In this case, the separation behavior could be tested without incurring the usual analytical expenses and time requirements. Such an approach can be applied for the identification and elimination from further consideration of insignificant process input parameters. The overall impact of parameters found to be process significant and their potential interactions would then be studied in a more comprehensive subsequent study in combination with all other factors found significant. This subsequent phase of process characterization also allows verification of the parameter ranking provided by this methodology with respect to all output parameters monitored.

Notations

a	activity
A, B	empirical parameters in Yamamoto model
c	fluid phase concentration
D_{crt}	longitudinal dispersion coefficient at non-binding conditions
D_L	longitudinal dispersion coefficient
d_p	resin particle diameter
G	gradient slope normalized to column void volume; $(I_{end} - I_0) \cdot V_0 / V_G$
H	void fraction; $(1 - \varepsilon) / \varepsilon = (V_t - V_0) / V_0$
I	ionic strength in Yamamoto model (c_I)
I_0	ionic strength at the start of the linear gradient
I_{end}	ionic strength at the end of the linear gradient
I_R	ionic strength at peak maximum during gradient elution
K	distribution coefficient
K_{crt}	distribution coefficient at non-binding conditions
K^e	equilibrium constant
L	column bed height
n	adsorbed phase concentration
n^s	resin ligand density, i.e. ion-exchange capacity
P	protein
p, q, r, s, t	fit parameters
R_s	resolution between adjacent peaks
t_R	retention time
u	linear flow velocity
V_0	column void volume
V_G	gradient volume
V_t	total column volume
z	charge of protein or counter-ion
γ	activity coefficient
θ	dimensionless time; $t \cdot u / L$
Γ	lumped activity coefficient term as defined in Equation (5)
ν	partial molar volume
π	swelling pressure
σ	variance around retention time

Subscripts

I	counter-ion
P	protein

References

- DeSain C, Sutton C. 2000. Process hazard analysis and critical point Identification. *BioPharm* 13(10):36–40.
- Gardner AR, Smith TM. 2000. Identification and establishment of operating ranges of critical process variables. In: Sofer G, Zabriskie DW, editors. *Biopharmaceutical Process Validation*. New York: Marcel Dekker. p 61–76.
- Helfferich F. 1962. *Ion exchange*. New York: McGraw-Hill.
- ICH. 1999. Specifications: Test Procedures and Acceptance Criteria for Biotechnological/Biological Products. ICH Harmonized Tripartite Guideline Q6B.
- ICH. 2006a. Pharmaceutical Development. ICH Harmonized Tripartite Guideline Q8.
- ICH. 2006b. Quality Risk Management. ICH Harmonized Tripartite Guideline Q8.
- Ishihara T, Yamamoto S. 2005. Optimization of monoclonal antibody purification by ion-exchange chromatography: Application of simple methods with linear gradient elution experimental data. *J Chromatogr A* 1069:99–106.
- Jungbauer A, Kaltenbrunner O. 1996. Fundamental questions in optimizing ion-exchange chromatography of proteins using computer-aided process design. *Biotechnol Bioeng* 52:223–239.
- Kelley BD. 2000. Establishing process robustness using designed experiments. In: Sofer G, Zabriskie DW, editors. *Biopharmaceutical Process Validation*. New York: Marcel Dekker. p 29–59.
- LeVan MD, Carta G, Yon CM. 1997. Adsorption and ion-exchange. In: Perry RH, Green DW, Maloney JO, editors. *Perry's Chemical Engineers' Handbook*. 7 ed: McGraw-Hill.
- Michaeli I, Katchalski A. 1957. Potentiometric titration of polyelectric gels. *J Polym Sci* 23:683–696.
- Montgomery DC. 2000. *Design and Analysis of Experiments*. New York: John Wiley & Sons.
- Shen H, Frey DD. 2004. Charge regulation in protein ion-exchange chromatography: Development and experimental evaluation of a theory based on hydrogen ion Donnan equilibrium. *J Chromatogr A* 1034:55–68.
- Shene C, Lucero A, Andrews BA, Asenjo JA. 2006. Mathematical modelling of elution curves for a protein mixture in ion exchange chromatography and for the optimal selection of operational conditions. *Biotechnol Bioeng* 95:704–713.
- Ståhlberg J, Jönsson B. 1999. Influence of charge regulation in electrostatic interaction chromatography of proteins. *Analyt Chem* 68:1536–1544.
- Watler P, Kaltenbrunner O, Feng D, Yamamoto S. 2003. Engineering aspects of ion-exchange chromatography. In: Rathore AS, Velayudhan A, editors. *Scale-Up and Optimization on Preparative Chromatography*. New York: Marcel Dekker, Inc.
- Winkler ME. 2000. Purification issues. In: Sofer G, Zabriskie DW, editors. *Biopharmaceutical Process Validation*. New York: Marcel Dekker. p 143–155.
- Yamamoto S, Ishihara T. 1999. Ion-exchange chromatography of proteins near the isoelectric points. *J Chromatogr, A* 852(1):31–36.
- Yamamoto S, Nakanishi K, Matsuno R. 1988. In: Cazes J, editor. *Ion-Exchange Chromatography of Proteins*. New York: Marcel Dekker p 401.