Contents lists available at ScienceDirect

Chemical Engineering Research and Design

journal homepage: www.elsevier.com/locate/cherd

Modelling of industrial biopharmaceutical multicomponent chromatography☆

Edward J. Close[a,b], *Jeffrey R. Salm*[c], *Daniel G. Bracewell*[b], *Eva Sorensen*[a,*]

[a] *Centre for Process Systems Engineering, Department of Chemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*

[b] *The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, Torrington Place, London WC1E 7JE, UK*

[c] *Pfizer Biopharmaceuticals, 1 Burtt Road, Andover, MA 01810, USA*

A B S T R A C T

The development and validation of a chromatography rate model for an industrial multicomponent chromatographic bioseparation is presented. The model is intended for use in a process scenario to allow specific variables critical to product quality to be studied. The chromatography provides impurity clearance whilst producing a complex product composed of six closely related variants of a dimer protein therapeutic (~30 kDa), with their monomer subunits in a specific ratio. Impurity removal is well understood, however, achieving the correct monomer subunit ratio can pose a purification challenge. We utilise a stepwise approach to develop a model for studying the effect of feed material variability on product quality. Scale down experiments are completed to quickly generate data for estimating model parameters, before an iterative procedure is employed where the industrial process is used to refine parameters in a sequential manner, until model predictions exhibit satisfactory agreement with experimental data. Final model predictions were in good agreement with experimental product quality (within 3%). The results demonstrate how good understanding of an industrial process can help facilitate model development when an exhaustive description is not required, despite considering a chromatographic bioseparation with crude feed material and challenging purification objectives.

Keywords: Process systems modelling; Quality by design; Chromatography modelling; Hydrophobic interaction chromatography; Purification process development; Therapeutic protein

1. Introduction

Advances in healthcare over the past half century have been of immense benefit to the quality of life for an increasing world population. The rapid growth in protein therapeutics has played a key role in this, and is predicted to continue with several hundred clinical candidate proteins currently estimated in company pipelines (Kelley, 2009) of which many serve significant unmet medical needs (Shukla et al., 2007). However, despite several decades of effort to improve R&D efficiency and performance, the process for bringing a new biopharmaceutical product to market remains an expensive, time-consuming, and risky proposition (Lightfoot and Moscariello, 2004).

Chromatographic separations are the workhorse of therapeutic protein purification (Kelley, 2007), but their design and operation is a challenging task. An optimal, safe and economic process must be found quickly somewhere in an extremely large parameter space which simply cannot be explored in depth using traditional experimental methodologies. Downstream process development currently depends heavily upon empirical experimentation interpreted using heuristic knowledge to arrive at an acceptable process (Lightfoot and Moscariello, 2004). The amount of material available to work with is often limited, and the work must be completed within constricted timelines to meet time to market constraints. Furthermore, US Food and Drugs Administration

* *Corresponding author*. Tel.: +44 020 7679 3802; fax: +44 020 7383 2348.
E-mail address: e.sorensen@ucl.ac.uk (E. Sorensen).

Received 2 July 2013; Received in revised form 21 October 2013; Accepted 22 October 2013
0263-8762/$ – see front matter © 2013 The Authors. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cherd.2013.10.022

Notation

C^m	mobile phase concentration (mg/ml)
C^{sp}	stationary phase concentration (mg/ml)
CF	compression factor
D_A	apparent axial dispersion coefficient (cm^2/s)
F	mobile phase flowrate (ml/min)
k_a	equilibrium constant
L	column length (cm)
N_C	number of components
N_p	number of theoretical plates
q_s	saturation capacity (mg/ml)
q	settled resin concentration (mg/ml)
t_0	retention time of an unretained small molecule (s)
t	time (s)
u	interstitial velocity (cm/s)
V_0	void volume (ml)
V_C	column volume (ml)
ϵ_T	total column porosity
Δ	peak width of an unretained molecule at half of the peak height
i	component identifier

(FDA) regulations require that the basic separation scheme is fixed prior to clinical trials, early on in the overall development process.

The FDA is now encouraging the use of quality by design (QbD) principles during process development and operation (US Food and Drug Administration, 2006). Key to a QbD approach is a thorough understanding of process inputs and their impact on performance, the relationship between the process and the products' critical quality attributes (CQA), and the association between the CQA's and a product's clinical properties (Jiang et al., 2010). The expected benefit from a QbD approach is an increase in the assurance of product quality, and in turn, the FDA will allow manufacturers greater flexibility to operate with lower regulatory burden, enabling continuous process improvement, as well as greater robustness.

Mechanistic modelling can be a useful tool for studying the impact of process parameters on process performance and product CQA's. Altering the values of process parameters may be difficult or even impossible to accomplish experimentally, e.g. feed stream composition, but is trivial in a model based approach. In addition, simulations can be completed quickly and efficiently which is valuable in an industrial scenario where time and material is often limited, and the fundamental knowledge gained by their application can be used to better understand, and reduce, processing risks.

The equations used to mathematically describe the chromatographic purification of proteins are well understood (e.g. Kaczmarski et al., 2001; Guiochon, 2002; Brooks and Cramer, 1992; Seidel-Morgenstern, 2004). The systematic development of a chromatographic model has been described for many different systems, including ion exchange (Melter et al., 2008), hydrophobic interaction (McCue et al., 2008; Nagrath et al., 2011), and protein A chromatography (Ng et al., 2012). The issue of efficient model calibration has been thoroughly addressed (Teoh et al., 2001; Persson et al., 2006; Susanto et al., 2006; Osberghaus et al., 2012a). Mechanistic models of chromatography have been successfully employed to simulate numerous case studies (Kaczmarski et al., 2002; Mollerup et al., 2007; Osberghaus et al., 2012b). In addition, there are useful examples of using chromatography models for optimisation (Degerman et al., 2006, 2007; Ng et al., 2012), scale up (Gerontas et al., 2010), design space characterisation (Degerman et al., 2009; Gétaz et al., 2013b), as well as robustness, uncertainty and sensitivity analysis (Jakobsson et al., 2005; Borg et al., 2013).

As a result of the progress in modelling chromatography that has been made over the last decade, systems with crude feed material, containing product, product-related impurities (e.g. oxidation, deamidation, acetylation, dimerisation), and process-related impurities (e.g. antifoam, DNA, protein, virus) have recently been considered (e.g. Gétaz et al., 2013a; Nfor et al., 2013). The complexity of the industrial feed material in these studies means that the model development procedures involve conducting an extensive range of detailed experiments which may not be suitable in certain scenarios. One such scenario in industry is where the majority of process development has already taken place, but there remains a desire to develop understanding of a key feature of a bioseparation. The experimental effort required to develop an exhaustive model may discourage a mechanistic modelling approach considering time and material constraints.

In this work, a chromatography model for predicting product quality in an industrial multicomponent bioseparation is developed and validated. The model is intended for use in a process scenario to allow specific variables critical to product quality to be studied. The chromatography utilises a hydrophobic interaction retention mechanism to purify a multicomponent product from a complex mixture of impurities. Process parameters were predefined prior to this work. Impurity removal is well understood and therefore a model description of this feature of the chromatography is not required. However, the step must also deliver the multicomponent product composed of six closely related variants of a dimer protein therapeutic (~30 kDa) with their monomer subunits in a specific ratio. Variability in the feed material poses a purification challenge, and consequently, there is a risk that the products' monomer subunit ratio will not meet product quality specifications incurring significant losses. Therefore, a model which can study product quality as a function of the load material concentration and composition is developed and validated in this work. A systematic procedure is used to determine key model parameter values, first using targeted experimental studies to quickly generate experimental data for estimation of model parameters, before employing an iterative procedure where laboratory scale column runs of the industrial process using real feed material are used to refine parameters in a sequential manner until model predictions exhibit satisfactory agreement with experimental data. We demonstrate how good understanding of an industrial process can facilitate model development, despite considering a chromatographic bioseparation with crude feed material and challenging purification objectives.

2. Problem description

The hydrophobic interaction chromatography (HIC) considered in this work is a complex separation with challenging purification objectives. The product of interest is a disulphide linked dimer protein molecule (MW = 30 kDa), comprised of two monomer subunits. Three variations of the monomer

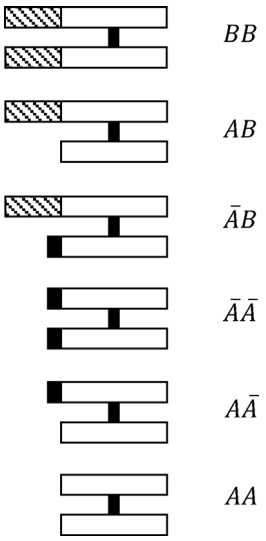

Fig. 1 – The product is a disulphide linked dimer protein therapeutic (MW ≈ 30 kDa), comprised of two monomer subunits. Three variations of the monomer subunit exist due to slight variations in the amino acid sequence, here denoted *A*, $\bar{A}$ and *B*, resulting in six different forms of the product. All six closely related variants of the dimer protein therapeutic must be present in the final product, with their monomer subunits in a specific ratio.

subunit exist due to slight variations in the amino acid sequence, here denoted A, $\bar{A}$ and B. This results in six possible isoforms of the dimer (AA, $\bar{A}\bar{A}$, $A\bar{A}$, $\bar{A}B$, AB and BB) as shown in Fig. 1. The corresponding analytical chromatogram is shown in Fig. 2A. Each form is an active component of the final product which must contain a specific ratio of the monomer subunits, $(A + \bar{A}) : B$, i.e. not just one product form at a given total amount is required, but six closely related dimer variants, with a specific ratio of their monomer subunits. Specifically, subunit B must account for between 25% and 45% of all monomer subunits in the product, i.e. $0.25 < B < 0.45$. In addition to the product, approximately 30% of the HIC feed material was product related impurities including the individual monomer subunits (A, $\bar{A}$ and B), incorrectly formed product species (I1, I2, I3, I4), and host cell related contaminants consisting of mainly host cell protein. The corresponding analytical chromatogram is shown in Fig. 2B.

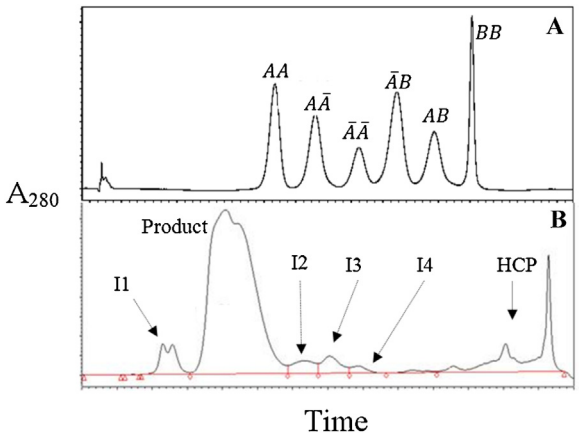

Fig. 2 – Analytical chromatogram of (A) the product and (B) the feed material. (Axis values removed for confidentiality purposes).

3. Systematic model development

3.1. Determining modelling approach

Minimising the time and amount of material required to develop, validate, solve and determine model solutions is extremely important, as process development timelines are typically very constricted (Steinmeyer and McCormick, 2008). Especially so, as many simulations may be required to map the impact of process parameters on process performance and product CQA's (Degerman et al., 2006; Karlson et al., 2004). Thus, we aimed to minimise the necessary experimental and computational effort required to model the HIC and our ultimate goal was not to develop the most *exact* model description, but to develop a model with *sufficient accuracy* for use in industry.

3.2. Mathematical model

An equilibrium dispersive model was used to simulate the HIC (Guiochon et al., 1994). This model assumes that the mass transfer kinetics between the mobile phase moving through the column bed and the particles is infinitely fast. Thus the concentration of component i in the mobile phase is equal to the average concentration of component i in the intra-particular mobile phase, and the axial dispersion coefficient is replaced with an apparent axial dispersion coefficient which includes the contribution from the mass transfer kinetics. The model has the following additional assumptions; (i) the column is one-dimensional (radially homogeneous), (ii) the chromatographic separation is isothermal and adiabatic, (iii) the compressibility of the mobile phase is negligible and thus the velocity profile is flat, and (iv) the mass transfer parameters are independent of component concentration.

With the assumption that the column is radially homogeneous, the differential mass balance in the bulk mobile phase is described by (Guiochon et al., 1994):

$$\frac{\partial C_i^m}{\partial t} + \frac{(1-\epsilon_T)}{\epsilon_T} \cdot \frac{\partial C_i^{sp}}{\partial t} + u \cdot \frac{\partial C_i^m}{\partial z} = D_A \cdot \frac{\partial^2 C_i^m}{\partial z^2} \quad \forall i = 1, 2, \ldots, N_C \quad \text{and} \quad z \in (0, L) \tag{1}$$

where C_i^m is the concentration of component i in the mobile phase, t is the time, ϵ_T is the total column porosity, C_i^{sp} is the concentration of component i in the stationary phase, u is the interstitial velocity, z is the axial coordinate, D_A is the apparent axial dispersion coefficient, N_C is the number of components in the system, and L is the column length. $\partial C_i^m/\partial t$ is the rate per unit volume of accumulation of component i in the mobile phase, $((1-\epsilon_T)/\epsilon_T) \cdot (\partial C_i^{sp}/\partial t)$ is the rate per unit volume of accumulation of component i in the stationary phase, $u \cdot (dC/dz) \cdot (\partial C_i^m/\partial z)$ is the rate per unit volume of mass transfer by convection down the column, and $D_A \cdot (\partial^2 C_i^m/\partial z^2)$ is the rate per unit volume of mass transfer by dispersion and particle mass transfer kinetics lumped into one term. The apparent axial dispersion coefficient can be estimated from the number of theoretical plates of the column, N_p (Guiochon et al., 1994):

$$D_A = \frac{uL}{2N_p} \tag{2}$$

The number of theoretical plates is determined directly from the chromatogram of an unretained small molecule

according to the following equation (Synder and Kirkland, 2008):

$$N_p = 5.54\left(\frac{t_0}{\Delta}\right)^2 \quad (3)$$

where t_0 is the retention time of an unretained small molecule and Δ is the peak width of the unretained molecule at half of the height of the chromatogram peak (both determined experimentally). The total column porosity, ϵ_T, is defined as the ratio between the void volume, V_0, and the column volume, V_C:

$$\epsilon_T = \frac{V_0}{V_C} \quad (4)$$

The void volume, V_0, can be determined from the product of the retention time of an unretained small molecule (determined experimentally), t_0, and the mobile phase flowrate, F:

$$V_0 = t_0 F \quad (5)$$

The boundary conditions for Eq. (1) are the following (Guiochon et al., 1994):

At the inlet of the column, i.e. at $z=0$, the mobile phase concentration, C_i^m, depends on convection and dispersion:

$$\left[uC_i^m - D_A\frac{\partial C_i^m}{\partial z}\right]|_{z=0} = uC_{i,0}^m \quad \forall i = 1, 2, \ldots, N_C \quad (6)$$

where $C_{i,0}^m$ is the inlet concentration.

At the outlet of the column, only convective transport is considered:

$$\frac{\partial C_i^m}{\partial z}|_{z=L} = 0 \quad \forall i = 1, 2, \ldots, N_C \quad (7)$$

An initial condition is also required to solve Eq. (1) which states that the rate per unit volume of accumulation in the mobile phase of component i at $t=0$ is zero at all points interior to the column:

$$\frac{\partial C_i^m}{\partial t} = 0 \quad 0 < z < L \quad \forall i = 1, 2, \ldots, N_C \quad (8)$$

In this work, protein adsorption onto the stationary phase was modelling using a competitive Langmuir adsorption isotherm (Seidel-Morgenstern, 2004):

$$q_i = \frac{q_s \cdot k_{a,i} \cdot C_i^m}{1 + \sum k_{a,i} \cdot C_i^m} \quad \forall i = 1, 2, \ldots, N_C \quad z \in (0, L) \quad (9)$$

where q_s is the resin saturation capacity, and $k_{a,i}$ is the equilibrium constant of component i, and q_i represents the amount of protein adsorbed per unit volume of settled resin. When the adsorption isotherm (Eq. (9)) is linked with the differential mass balance in the bulk mobile phase (Eq. (1)), the amount of protein adsorbed per unit volume of settled resin, q_i, is converted to the amount of protein adsorbed per unit volume of stationary phase in the packed bed, C_i^{sp}:

$$C_i^{sp} = \frac{CF \cdot q_i}{(1-\epsilon_T)} \quad \forall i = 1, 2, \ldots, N_C \quad z \in (0, L) \quad (10)$$

where dividing q_i by $(1-\epsilon_T)$ accounts for the phase ratio (Mollerup, 2008), and multiplying q_i by a compression factor, CF, defined as the ratio between settled bed volume and packed bed volume, accounts for bed compression (Gerontas et al., 2010). All model equations are implemented and solved using the dynamic simulation tool gPROMS™ (Process Systems Enterprise, 2013). Discretisation of the column in the axial coordinate is done using the built-in orthogonal collocation on finite element method (OCFEM).

3.3. *Parameter estimation and model validation*

We used a systematic approach to model calibration, as illustrated in Fig. 4. In the procedure, targeted micro well experimentation is utilised to estimate the adsorption isotherm parameters, q_s and $k_{a,i}$. Scale down column studies are used to determine the total column porosity, ϵ_T, and give an initial estimate of the mass transfer, D_A, parameters of the system (Fig. 4 Steps 1 and 2). Then an iterative procedure is employed, where laboratory scale column runs (7 ml CV, 7.4 cm bed height) of the industrial process using untreated feed material are used to refine parameters in a sequential manner until model predictions exhibit satisfactory agreement with experimental data (illustrated in Fig. 4 Steps 2 and 3, results for the HIC described in detail in Section 4). Such a procedure is required when model predictions after the initial calibration effort (Fig. 4 Steps 1 and 2) are not satisfactory, because generating more pure material for further model development is normally prohibitively costly, and using data from the actual industrial process is a superior method from our industrial perspective, as will become clear in the following section. Experimental results were used to estimate the values of model parameters using the 'parameter estimation' entity in gPROMS based on the SRQPD sequential quadratic programming code. Parameter estimation was based on the maximum likelihood formulation, which determines values for the uncertain physical and variance model parameters that maximise the probability that the model will predict the measurement values obtained from the experiments. The statistical variance model of constant variance was used in this case (Process Systems Enterprise, 2013).

4. Model development of HIC chromatography

In the following section we describe the application of the procedure outlined above to the industrial multicomponent hydrophobic interaction chromatography considered in this work.

4.1. *Generation of purified materials for model calibration experiments*

One of the key challenges of developing a mechanistic model of industrial chromatographic processes is the limited availability of purified material, which is required for many tasks in the overall effort to bring a protein therapeutic to market e.g. drug trials, stability studies, toxicology studies, etc. We had to generate our own protein solutions for model development experiments from bulk feed material. Following initial purification by pseudo affinity capture, the material contained the 6 product forms of interest, as well as a range of product related impurities and host cell proteins (HCP's). The product forms were further purified and isolated from impurities over multiple runs on the hydrophobic interaction chromatography considered in this work. Multiple runs were required as it

was particularly challenging to separate the product isoform *BB* from closely related product related impurities. Fractions were taken every column volume (CV) and analysed by CEX HPLC in order to determine the isoform distribution of the sample. Multiple samples with a range of isoform distributions were generated in this way, and later pooled in order to generate material with desired isoform distributions for development experiments. Confirmation of removal of product related impurities and HCP's was determined by phenyl RP HPLC. We were unable to generate pure samples of each product isoform which prevented determining traditional single component isotherms. We therefore required an approach to model development which used multicomponent mixtures of the six product forms. This involved (1) measuring isotherms competitively and (2) using an advanced parameter estimation facility to fit the experimental data to a competitive isotherm model.

4.2. *Assumptions*

The similar amino acid sequence of two of the monomer subunits (*A* and $\bar{A}$) results in similar separation properties of the product isoforms *AA*, $\bar{A}\bar{A}$, $A\bar{A}$ and the product isoforms $\bar{A}B$, *AB*. In order to simplify the modelling problem, the six product isotherms were reduced in the model to three components: *AA*, *AB* and *BB*. The chromatography cycle is divided into different steps (see Appendix for detailed description). Firstly, the product is applied to the column (load step), then buffer without any product is applied (wash step), before bound protein is collected by applying an elution buffer (elution step). We assumed that all product isoforms that remain bound to the column after the load and wash steps are subsequently collected in the elution step. This assumption was confirmed experimentally. Our approach also assumed that the product related impurities and HCP's in the feed stream had a negligible impact on the separation of the product of interest, as the impurities are observed to flow through during the load phase of the chromatographic cycle (Fig. 3B). To confirm this we also compared the product form distributions in fractions collected every CV during HIC runs with and without impurities in the feed material. Runs were identical in all other aspects e.g. load challenge, product form concentrations and wash length. We found that the impurities had no effect on the product distributions (Fig. 3A). In addition, by comparing the UV traces in Fig. 3A, one can clearly see where the impurities are flowing through during the load step, before the two UV traces merge and are in exact agreement.

4.3. *Initial model calibration experiments and determination of parameter values*

The apparent axial dispersion coefficient, D_A, and total column porosity, ϵ_T, were determined from pulse injections of a small unretained molecule using Eq. (2) to Eq. (5). The determined total column porosity value, 0.9, was in agreement with previous literature estimations for the same resin (McCue et al., 2007). The number of theoretical plates, N_p, was determined as 309 with a plate height of 0.0239 cm, resulting in an apparent axial dispersion coefficient, D_A, value of 0.0001 cm^2/s.

We made special effort to determine accurate isotherm parameters at the start of model development, as the accuracy of the equilibrium isotherm is the most important part of the chromatographic model (Mollerup et al., 2009). Competition

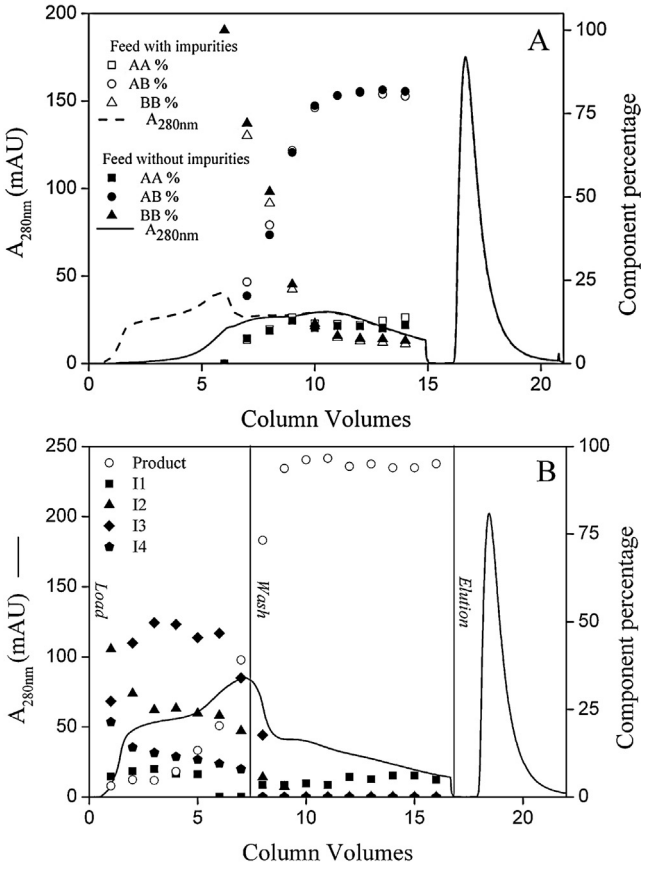

Fig. 3 – (A) Experimental comparison between column runs using feed material with and without impurities (7 ml CV, 7.4 cm bed height, 4.2 CV/h, inlet concentration = 0.34 mg/ml, load challenge = 2 mg/ml). Similar product form percentages and overlapping $A_{280\,nm}$ trace during wash and elution indicates that impurities have minimal impact on separation of product forms and can be neglected in the model. (B) Chromatogram showing the $A_{280\,nm}$ trace and the percentage of product related impurities and product in samples taken every CV during a standard HIC run, determined by phenyl RP HPLC. The figure shows that the majority of impurities in the feed material elute from the column during the load phase, product forms begin to elute from the column at the end of the load phase and continue throughout the wash.

between product forms was a significant part of the separation, as the multiple product forms were closely related and had similar adsorption properties. It was important to capture these competitive effects in the adsorption isotherm model for simulation accuracy. We used batch adsorption experimental studies on a microwell plate to generate competitive adsorption data (i.e. the amount of component *i* adsorbed per unit volume of settled resin, q_i (mg/ml), as a function of the concentration of all components in the liquid (mobile phase), C_i^m, (mg/ml)), which was fitted with a competitive Langmuir isotherm model (Eq. (9), using gPROMS to estimate the saturation capacity, q_s, and equilibrium constants, $k_{a,i}$.

Batch adsorption in a microwell plate is tedious requiring labour intensive experiments, and can give inaccurate results (Seidel-Morgenstern, 2004). However, it can be automated using robotic liquid handling to reduce experimental burden, and is simpler than most alternative methodologies (e.g. perturbation method, dispersed front analysis, peak fitting) which are difficult when applied to multicomponent

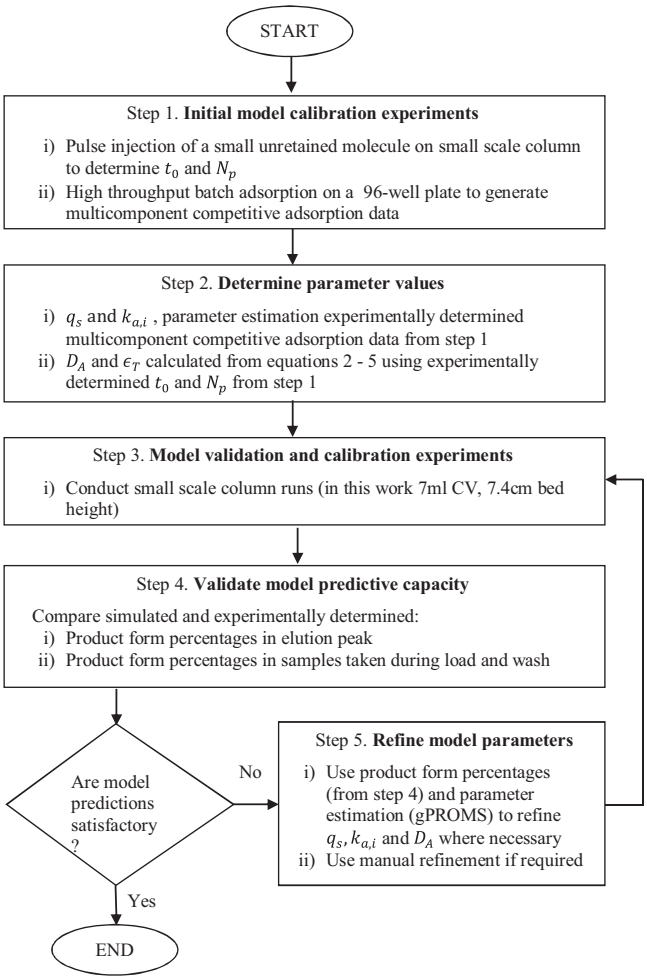

Fig. 4 – Flow diagram showing the model development procedure used in this work. We utilised a stepwise approach, first using targeted scale down experimental studies to quickly generate data for estimation of model parameters (Steps 1 and 2). We then employed an iterative procedure where laboratory scale column runs of the industrial process were used to refine parameters in a sequential manner until model predictions were in agreement with experimental data (Steps 3, 4 and 5).

separations. Only frontal analysis is relatively straightforward, but was unfeasible in this work due to high costs and limited availability of required material. Our exact methodology for the batch adsorption studies was based upon previous work on high throughput screening of chromatographic separations (Coffman et al., 2008), and is described in the Appendix. All batch adsorption experiments were repeated in triplicate, and averages are shown in this paper.

To ensure that the experimental data used to fit the competitive Langmuir model contained competitive information, we varied the product form distribution in the load material used in the batch adsorption experiments. Each graph in Fig. 5 shows adsorption data from microwell experiments conducted at a different load material product form distribution, shown in the ratio AA%:AB%:BB% in the top left hand corner of each graph. Note that although the graphs show the bound concentration of the product form as a function of its mobile phase concentration, the mobile phase concentration of the other two product forms are also affecting the bound concentration. The effects of competition for binding sites is clear when the graphs in Fig. 5 are compared. In graph C, the

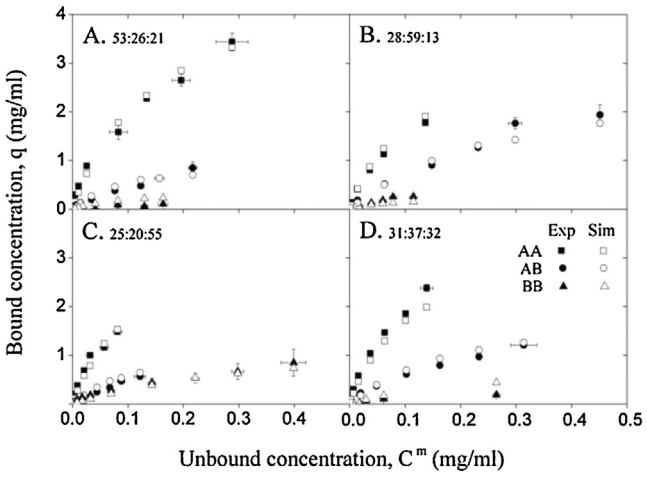

Fig. 5 – Experimental and simulated multicomponent competitive adsorption isotherms at a range of load material product distributions, as shown on the graphs in the order AA%:AB%:BB%. The experimental data is from micro well plate batch adsorption followed by CEX HPLC analysis. All experimental points were repeated in triplicate and standard error is shown on graphs.

BB stationary phase concentrations are significantly higher, especially compared to Graph B. This is due to the favourable product distribution of the load material resulting in fewer competing components, allowing more BB to bind (graph C load material 20% AA:25% AB:55% BB, graph B load material 28% AA:59% AB:13% BB.

The estimated isotherm parameter values are shown in Table 1. The standard deviations of the estimated parameters are approximately ten percent, indicating there is still some uncertainty around the parameter values. The coefficient of determination, r^2, for the model fit to experimental data was 0.96, which was found to be sufficient for satisfactory agreement between model predictions and experimental data as shown in Fig. 5, given the inherent uncertainties of the batch adsorption studies.

4.4. *Model validation and calibration experiments*

The product form distribution in the elution peak, and in samples taken from the column outlet every column volume (CV) during the load and the wash, was measured during experimental small-scale column runs using CEX HPLC. The ability of the model to predict these product form distributions was an important and industrially relevant test to validate model accuracy, since achieving a specific percentage distribution of the product isoforms in the elution pool is a key objective of the

Table 1 – Experimentally determined model parameter values after initial model calibration.

Parameter name	Parameter symbol	Value	StDev
AA equilibrium constant	$k_{a,1}$	5.31	0.58
AB equilibrium constant	$k_{a,2}$	1.49	0.16
BB equilibrium constant	$k_{a,3}$	0.52	0.07
Saturation capacity	q_s	6.45	0.36
Total column porosity	ϵ_T	0.9	–
Apparent axial dispersion coefficient	D_A	0.0001	–
Compression Factor	CF	1.25	–

Table 2 – Model validation runs: product percentage in load, load concentration, wash length, and load challenge.

Run identifier	Load challenge (mg/ml resin)	Load concentration (mg/ml)	Wash length (CV)	Percentage AA	Percentage AB	Percentage BB
A	1.5	0.26	10	35	35	30
B	2.4	0.44	10	14	38	48
C	2.2	0.35	10	40	44	16
D	1	0.4	3	38	20	42
E	1	0.11	3.2	38	20	42

chromatography. This data was also suitable for model calibration should the first iteration of the model give unsatisfactory predictions.

4.5. *Validation of model predictive capacity*

The first iteration of the model was unable to give satisfactory predictions of the chromatographic process after the initial model calibration, as shown in Fig. 6A. This was not unexpected, as although we had spent extra effort ensuring that the isotherm parameters were accurate, we used a lumped mass transfer coefficient and had determined its value using the residence time of a unretained molecule (NaCl) that was significantly smaller than the protein, and thus would be expected to experience faster mass transfer. We could have considered using a more complex model (Kaczmarski et al., 2001), and/or used alternative experimental approaches to determining mass transfer parameters more accurately, e.g. conducting multiple pulse injections of the product at non-binding conditions, or using van Deemter plot methodology, etc. (Muller-Spath et al., 2011; Ng et al., 2012). However, applying these techniques to this industrial separation was problematic. For example, we could not find any non-binding conditions suitable for pulse injections of the product which did not significantly alter the system, and the lack of any pure component material, the highly competitive system and closely related isotherms of the product forms meant that the van Deemter plot method was not practical. In addition, we were conscious that there may also have been unidentified phenomenon occurring which could potentially affect these approaches, for example proteins unfolding on the HIC surface. Therefore, we took an alternative approach as discussed in the following section.

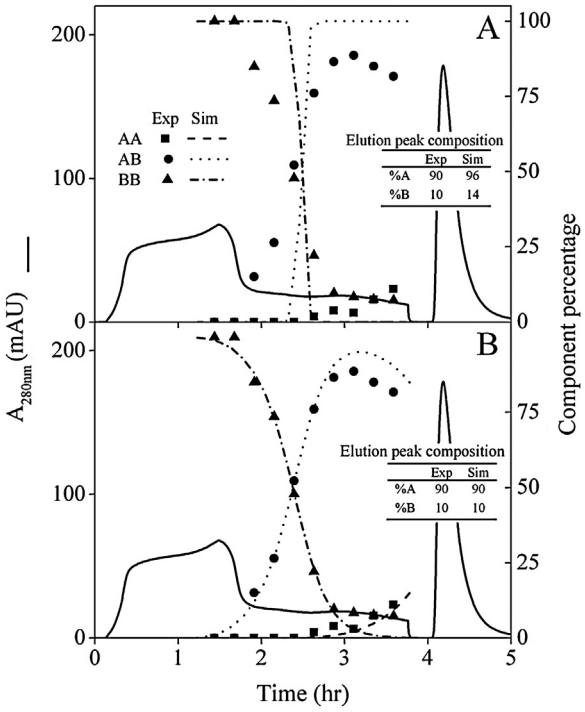

Fig. 6 – Experimental and simulated product form distributions during load, wash and in final elution peak. (A) Before model refinement. (B) After model refinement. The apparent axial dispersion coefficient and the AA adsorption constant were modified from 0.0001 cm^2/s to 0.003 cm^2/s, and 5.31 to 3.5, respectively.

4.6. *Refine model parameters*

We used data from the previously conducted model validation run of the industrial process to refine the model parameters. Experimentally determined product distributions in fractions taken every CV were used to estimate the new value of the apparent axial dispersion coefficient, D_A, using the 'parameter estimation' entity in gPROMS. The previously determined value, estimated from the number of theoretical plates of the column, N_p, was used as an initial guess. Adsorption isotherm parameters, q_s and $k_{a,i}$, were refined manually, guided by the previously determined values, standard deviations and experimental product distributions. This method made optimal use of validation data, and meant no further experiments were required other than for further model validation. Runs of the industrial process are straightforward to conduct, can be left to run unsupervised with minimal preparation time, and can use feed material that needs no pre treatment and is readily available. This part of our approach can be easily integrated with traditional process development where column runs of the industrial process are conducted regularly. The drawback of such a procedure is that it can result in sub optimal parameter values. Estimating model parameters directly from the industrial process under normal operation was not possible without prior knowledge of suitable initial values, due to the complexity of the feed material and separation considered in this work.

We found that model predictions were in good agreement with experimental results after one iteration of refining the model parameters using model validation run data, shown in Figs. 6 and 7. The figures show the good agreement between the predicted and experimentally determined product form distributions in fractions taken every CV, and in the elution peak. The apparent axial dispersion coefficient and the AA equilibrium constant were modified from 0.0001 cm^2/s to 0.003 cm^2/s, and 5.31 to 3.5, respectively.

Model validation studies (including the model refinement run) were conducted using qualified scale down columns to provide a rigorous test of model accuracy. The flowrate, bed height and mobile phase conditions were kept constant throughout. The composition of the load material, total load

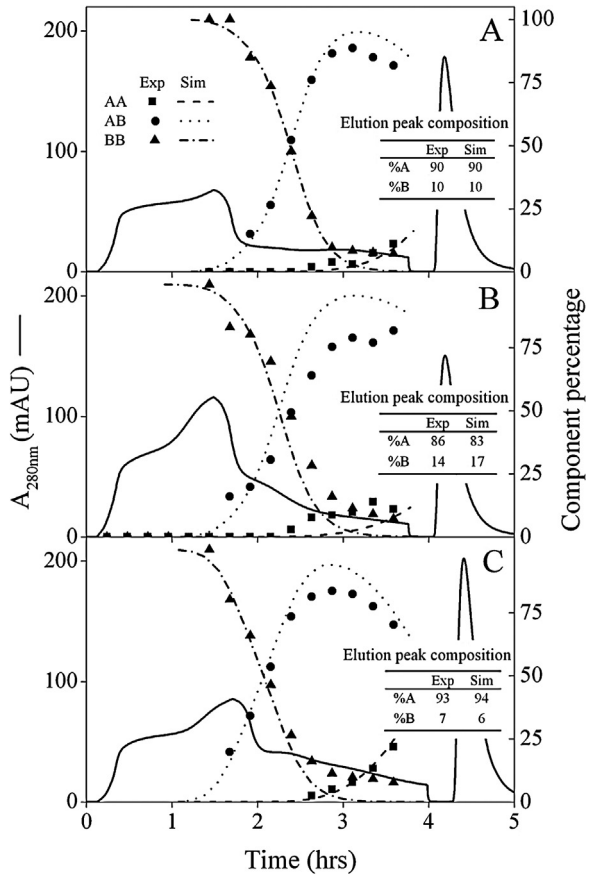

Fig. 7 – Experimental and simulated product form distributions during load, wash and in final elution peak in model validation runs. (7 ml CV, 7.4 cm bed height, 4.2 CV/h, load details shown in Table 2).

concentration, load challenge and wash length were varied as described in Table 2. Results are presented in Table 3. Validation run A was chosen to give a good representation of the isoform distributions experienced in normal load material and the load challenges used during day-to-day operation of the chromatography (Fig. 7A). The model gave very accurate predictions of the product form percentages in samples taken during the load and the wash phase, and was able to accurately predict the final products' monomer subunit ratio in the elution peak. The isoform distributions chosen for runs B and C were chosen to provide a challenge for the model, and would rarely be experienced during day-to-day operation. It was important to test these artificially created, rare load conditions in order to understand the limitations of the model. The model gave good predictions across the extended range of conditions, especially for the case study C (Fig. 7C). The difference between model predictions and experimental data for validation run B (Fig. 7B) was due to the very challenging isoform distribution in the load material, in particular that product isoform *BB* occupied 48% of the load, combined with a high load concentration and load challenge. The model slightly overestimated the percentage of *AB* during the wash length, which was attributed to tailing of *BB* that was not captured by the model. This was not observed during studies with fresh feed material, and thus it is likely that a small amount of product was damaged during the multiple applications, buffer exchanges and concentration steps used to generate the feed material for validation studies. Runs D and E were conducted to test model predictions at shorter wash lengths more applicable to producing a product of the desired quality, and thus no samples were taken during the wash. We found that the model predictions were again in good agreement with experimental data (predicted component percentages within 3%).

Table 3 – Model validation: experimental and predicted product quality.

Run identifier	Experimental %B	Simulation %B
A	10	10
B	14	17
C	7	6
D	44	46
E	34	37

5. Concluding remarks

The development and validation of a predictive mechanistic model of industrial biopharmaceutical multi-component chromatography has been described. The equilibrium dispersive model with a competitive Langmuir isotherm was able to successfully predict product quality for an extended range of inlet concentrations, load challenges and inlet product distributions. We used targeted small scale experimentation for initial model calibration, spending extra effort on micro well batch adsorption experiments for the estimation of the adsorption isotherm parameters. Then an iterative procedure was employed where laboratory scale column runs of the industrial process were used to refine parameters in a sequential manner until model predictions were in satisfactory agreement with the experimental data. Final model predictions were within 3% of the final products' monomer subunit ratio found in the chromatography elution peaks during validation studies. The model was also able to accurately predict the product form distribution in samples taken during the wash phase of the chromatographic cycle in validation runs A, B and C (samples were not taken in runs D and E). The model can now be used in subsequent studies to explore the effect of load material on product quality. The results demonstrate how good understanding of an industrial process can facilitate simpler model development when an exhaustive description is not required, despite considering a chromatographic bioseparation with crude feed material and challenging purification objectives.

Acknowledgements

The support of Pfizer and the contributions of Jenna Davison, Andrew Wood and Victoria Brook are gratefully acknowledged. The support of the UK Engineering and Physical Sciences Research Council (EPSRC) for the Innovative Manufacturing Research Centre (IMRC) in Bioprocessing and the EPSRC Centre for Innovative Manufacturing in Emergent Macromolecular Therapies is acknowledged gratefully. The IMRC and the EPSRC Centre are each part of The Advanced Centre for Biochemical Engineering, Department of Biochemical Engineering, University College London, with collaboration from a range of academic partners, biopharmaceutical and biotechnology companies.

Appendix. Experimental materials and methods

A.1. Chromatography resin and equipment

Butyl Sepharose 4B fast flow hydrophobic interaction resin was obtained from GE Healthcare (Uppsala, Sweden). All preparative scale laboratory experiments were carried out using an ÄKTA FPLC chromatography system from GE Healthcare (Uppsala, Sweden). Laboratory columns were 1.1 cm in diameter and 7.4 cm in bed height. Tosoh Bioscience TSKgel Phenyl-5PW RP and GE Healthcare Mono S column (5.0 mm × 50 mm) high performance liquid chromatography (HPLC) columns were used for in assays.

A.2. Cation exchange HPLC assay

The cation exchange (CEX) HPLC assay utilises a Mono S column and a gradient of sodium acetate, acetonitrile and sodium chloride at pH 5.0 in order to determine the relative percentages of the six dimer isoforms of the product in the sample. After equilibrating the column for 30 min, 100 μl samples at 0.5 mg/ml are injected onto a column at a flowrate of 1 ml/min. Over the course of the gradient, separation of the isoforms is accomplished based upon competitive ionic exchange of the sample ions with a counter ion in the mobile phase, for fixed cationic functional groups on the column resin. Absorbance at 280 nm is measured at the column exit. Integration of the resulting chromatogram and analysis of the relative percentage area of each peak indicates the percentage of each isoform in the sample. The total time to run each sample is 30 min.

A.3. Phenyl reverse phase HPLC assay

The phenyl reversed phase (RP) HPLC assay utilises a TSK-Phenyl reversed phase column and a water/acetonitrile/trifluoroacetic acid gradient system to determine the relative amount of product and product related impurities in samples. After equilibrating the column for 30 min, 100 μl samples at 1 mg/ml are injected onto a column equilibrated with a low percentage of acetonitrile mobile phase at a flowrate of 1 ml/min. As the organic modifier (acetonitrile) is increased over the course of the gradient, separation of the product related species and impurities is accomplished. Absorbance at 214 nm is measured at the column exit. Integration of the resulting chromatogram and analysis of the relative percentage area of each peak indicates the percentage of each species in the sample. The total time to run each sample is 80 min.

A.4. Hydrophobic interaction chromatography

During all runs the columns were first equilibrated with 50 mM Tris, 1.0 M NaC1, 0.50 M Arg-HCl, pH 7.00 equilibration buffer. Isocratic experiments were then conducted. The elution peak from a preceding pseudo affinity capture chromatography unit operation was brought to the correct NaCl concentration and applied to the column at 0.49 ml/min followed by a 10 column volume (CV) wash step using the equilibration buffer. Elution buffer consisting of 20% propylene glycol, 50 mM Tris, 0.50 M Arg-HCl, pH 7.00 was then applied and the product peak collected. Any remaining bound protein was removed using 0.1 M sodium acetate, pH 4.00 sanitisation buffer, and the column was stored in storage buffer when not in use. All experiments were conducted between 4 and 8 °C.

A.5. High throughput batch adsorption

Batch binding studies were conducted in a 96-well filter plate. The filter plates used throughout the experiments were round-well 800 μl plates with 0.45-μm pore-size polypropylene membrane. 25 μl of resin was taken from a bulk reservoir and dispensed by the robotic liquid handler into the individual wells as 25% (v/v) slurry in the appropriate equilibration buffer. The plate was then centrifuged to evacuate excess liquid and leave damp resin. Subsequently, other solutions composed of pure product, having various total protein concentrations (0.5–1 mg/ml) and isoform distributions (each component was varied between 20% and 60%) were added into wells containing the resin. The initial concentration and component distribution for each filter plate well were fashioned by mixing together protein from a bulk solution of known component distribution and concentration, with the appropriate amount of equilibration buffer from a bulk solution in order that the total volume of liquid dispensed into each well was 275 μl (V_{tot}). The resin and solutions were then agitated on a platform shaker for 120 min. Separate batch uptake studies indicated that equilibrium was reached in less than 30 min, and therefore that this incubation time was suitable. Foil adhesive tape was used on the underside of the filter-plate to prevent liquid loss during shaking. After incubation, a centrifuge evacuated the supernatant into a UV-transparent 96 well microplate which was stacked beneath the filter plate for analysis. The supernatant was then analysed by a 96-well UV spectrophotometer (SpectraMax 250, Molecular Devices, Sunnyvale, CA) to determine the concentration of protein in the supernatant, C_i^m. CEX HPLC was used to determine the percentage of each component in the supernatant, P_i^m. The concentration of the protein in the mobile phase is then calculated from Eq. (A1).

$$C_i^m = \frac{C_{equil} \cdot P_i^m}{100} \tag{A1}$$

where C_i^m is the concentration of component i in the mobile phase in mg/ml, C_{equil} is the measured concentration in the supernatant of the micro well, determined by UV spectroscopy, and P_i^m is the percentage of component i in the mobile phase as determined by CEX HPLC. An elution cycle was then conducted following the same methodology as the load cycle, where 275 μl of elution buffer was added to each well, the plate agitated on a platform shaker for 120 min and the supernatant subsequently collected as described previously and analysed using the spectrophotometer and CEX HPLC. The total amount of protein added to each micro well was then determined by Eq. (A2).

$$M_t = \frac{C_{elution}}{V_{elution}} + \frac{C_{equil}}{V_{equil}} \tag{A2}$$

where M_t is the total amount of protein added to the micro well (mg), $C_{elution}$ is the concentration of the elution supernatant (mg/ml), $V_{elution}$ is the volume of the elution supernatant (ml), C_{equil} is the concentration of the equil supernatent (mg/ml), and V_{equil} is the volume of the equil supernatant. The amount

of protein adsorbed per unit volume settled resin, q_i, is calculated using Eq. (A3).

$$q_i = \frac{((M_t \cdot P_i^{load}/100) - (C_i^m \cdot V_{equil}/1000))}{V_{resin}} \quad \text{(A3)}$$

References

Borg, N., Westerberg, K., Andersson, N., von Lieres, E., Nilsson, N., 2013. Effects of uncertainties in experimental conditions on the estimation of adsorption model parameters in preparative chromatography. Comput. Chem. Eng. 55, 148–157.

Brooks, C.A., Cramer, S.M., 1992. Steric mass-action ion exchange: displacement profiles and induced salt gradients. AIChE J. 38, 1969–1978.

Coffman, J.L., Kramarczyk, J.F., Kelley, B.D., 2008. High-throughput screening of chromatographic seperations: 1. Method development and column modeling. Biotechnol. Bioeng. 100, 605–618.

Degerman, M., Jakobsson, N., Nilsson, B., 2006. Constrained optimization of a preperative ion-exchange step for antibody purification. J. Chromatogr. A 1113, 92–100.

Degerman, M., Jakobsson, N., Nilsson, B., 2007. Modeling and optimization of preperative reversed-phase liquid chromatography for insulin purification. J. Chromatogr. A 1162, 41–49.

Degerman, M., Westerberg, K., Nilsson, B., 2009. A model-based approach to determine the design space of preparative chromatography. Chem. Eng. Technol. 32, 1195–1202.

Gerontas, S., Asplund, M., Hjorth, R., Bracewell, D.G., 2010. Integration of scale-down experimentation and general rate modelling to predict manufacturing scale chromatographic seperations. J. Chromatogr. A 1217, 6917–6926.

Gétaz, D., Stroehleina, G., Butté, A., Morbidelli, 2013a. Model-based design of peptide chromatographic purification processes. J. Chromatogr. A 1284, 69–79.

Gétaz, D., Butté, A., Morbidelli, M., 2013b. Model-based design space determination of peptide chromatographicpurification processes. J. Chromatogr. A 1284, 80–87.

Guiochon, G., Goldshan-Shirazi, S., Katti, A.M., 1994. Fundamentals of Preparative and Non-linear Chromatography. Academic Press, Boston.

Guiochon, G., 2002. Review: preparative liquid chromatography. J. Chromatogr. A 965, 129–161.

Jakobsson, N., Degerman, M., Nilsson, B., 2005. Optimization and robustness analysis of a hydrophobic interaction chromatography step. J. Chromatogr. A 1063, 99–109.

Jiang, C., Flansburg, L., Ghose, S., Jorjorian, P., Shukla, A., 2010. Defining process design space for a hydrophobic interaction chromatography (HIC) purification step: application of quality by design (QbD) principles. Biotechnol. Bioeng. 107, 985–997.

Karlson, D., Jakobsson, N., Axelsson, A., Nilsson, B., 2004. Model-based optimization of a preperative ion-exchange step for antibody purification. J. Chromatogr. A 1055, 29–39.

Kaczmarski, K., Antos, D., Sajonz, H., Sajonz, P., Guiochon, G., 2001. Comparitive modeling of breakthrough curves of bovine serum albumin in anion-exchange chromatography. J. Chromatogr. A 925, 1–17.

Kaczmarski, K., Cavazzini, A., Szabelski, P., Zhou, D., Liu, X., Guiochon, G., 2002. Application of the general rate model and the generalized Maxwell–Stefan equation to the study of the mass transfer kinetics of a pair of enantiomers. J. Chromatogr. A 962, 57–67.

Kelley, B.D., 2007. Review: very large scale monoclonal antibody purification: the case for conventional unit operations. Biotechnol. Prog. 23, 995–1008.

Kelley, B.D., 2009. Industrialization of mAb production technology. mAbs 1, 443–452.

Lightfoot, E.N., Moscariello, J.S., 2004. Bioseparations. Biotechnol. Bioeng. 87, 259–273.

McCue, J.T., Engel, P., Ng, A., Macniven, R., Thömmes, J., 2008. Modeling of protein/aggregate purification and separation using hydrophobic interaction chromatography. Bioprocess Biosyst. Eng. 31, 261–275.

McCue, J.T., Cecchini, D., Chu, C., Liu, W., Spann, A., 2007. Application of a two-dimensional model for predicting the pressure-flow and compression properties during column packing scale-up. J. Chromatogr. A 1145, 89–101.

Melter, L., Butte, A., Morbidelli, M., 2008. Preperative weak cation-exchange chromatography of monoclonal antibody variants: 1. Single component adsorption. J. Chromatogr. A 1200, 156–165.

Mollerup, J.M., Hansen, T.B., Kidal, S., Sejergaard, L., Hansen, E., Staby, A., 2007. Development, modelling, optimisation and scale-up of chromatographic purification of a therapeutic protein. Fluid Phase Equilibr. 261, 133–139.

Mollerup, J.M., 2008. A review of the thermodynamics of protein association to ligands, protein adsorption, and adsorption isotherms. Chem. Eng. Technol. 31, 864–874.

Mollerup, J.M., Hansen, T.B., Kidal, S., Sejergaard, L., Hansen, E., Staby, A., 2009. Use of quality by the design for the modelling of chromatographic seperations. J. Liq. Chromatogr. 32, 1577–1597.

Muller-Spath, T., Strolein, G., Aumann, L., Kornmann, H., Valax, P., Delegrange, L., Charbaut, E., Baer, G., Lamproye, A., Johnck, M., Schulte, M., Morbidelli, M., 2011. Model simulation and experimental validation of a cation exchange IgG capture step in batch and continuous chromatography. J. Chromatogr. A 1218, 5195–5204.

Nagrath, D., Xia, F., Cramer, S.M., 2011. Characterization and modeling of nonlinear hydrophobic interaction chromatographic systems. J. Chromatogr. A 1218, 1219–1226.

Nfor, B.K., Ahamed, T., van Dedem, G.W.K., Verhaert, P.D.E.M., van der Wielen, L.A.M., Eppink, M.H.M., van de Sandt, E.J.A.X., Ottens, M., 2013. Model-based rational methodology for protein purification process synthesis. Chem. Eng. Sci. 89, 185–195.

Ng, C.K.S., Osuna-Sanchez, H., Valery, E., Sørensen, E., Bracewell, D.G., 2012. Design of high productivity antibody capture by protein A chromatography using an integrated experimental and modeling approach. J. Chromatogr. B 899, 116–126.

Osberghaus, A., Drechsel, K., Hansen, S., Hepbildikler, S.K., Nath, S., Haindl, M., vonLieres, E., Hubbuch, J., 2012a. Model-integrated process development demonstrated on the optimization of a robotic cation exchange step. Chem. Eng. Sci. 76, 129–139.

Osberghaus, A., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E., Hubbuch, J., 2012b. Optimizing a chromatographic three component separation: a comparison of mechanistic and empiric modeling approaches. J. Chromatogr. A 1237, 86–95.

Persson, P., Gustavsson, P.E., Zacchi, G., Nilsson, B., 2006. Aspects of estimating dependencies in a detailed chromatography model based on frontal experiments. Process Biochem. 41, 1812–1821.

Process Systems Enterprise, 2013. gPROMS, www.psenterprise.com/gproms

Seidel-Morgenstern, A., 2004. Review: experimental determination of single solute and competitive adsorption isotherms. J. Chromatogr. A 1037, 255–272.

Shukla, A.A., Hubbard, B., Tressel, T., Guhan, S., Low, D., 2007. Downstream processing of monoclonal antibodies – application of platform approaches. J. Chromatogr. B 848, 28–39.

Susanto, A., Wekenborg, K., Hubbuch, J., Schmidt-Traub, H., 2006. Developing a chromatographic column model for bovine serum albumin on strong anion-exchange Source30Q using data from confocal laser scanning microscopy. J. Chromatogr. A 1137, 63–75.

Steinmeyer, D.E., McCormick, E.L., 2008. The art of antibody process development. Drug Discov. Today 13, 613–618.

Synder, L.R., Kirkland, J.J., 2008. Introduction to Modern Liquid Chromatography. Wiley, New York.
Teoh, H., Turner, M., Titchener-Hooker, N., Sorensen, E., 2001. Experimental verification and optimisation of a detailed dynamic high performance liquid chromatography column model. Comput. Chem. Eng. 25, 893–903.
US Food and Drug Administration, 2006. Guidance for industry; Q8 pharmaceutical development.