TU Delft

Delft University of Technology

Model-based optimization of integrated purification sequences for biopharmaceuticals

Pirrung, Silvia M.; Berends, Carmen; Backx, Antoon H.; van Beckhoven, Ruud F.W.C.; Eppink, Michel H.M.; Ottens, Marcel

DOI
10.1016/j.cesx.2019.100025

Publication date
2019

Document Version
Final published version

Published in
Chemical Engineering Science: X

Citation (APA)
Pirrung, S. M., Berends, C., Backx, A. H., van Beckhoven, R. F. W. C., Eppink, M. H. M., & Ottens, M. (2019). Model-based optimization of integrated purification sequences for biopharmaceuticals. *Chemical Engineering Science: X*, *3*, Article 100025. https://doi.org/10.1016/j.cesx.2019.100025

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Chemical Engineering Science: X 3 (2019) 100025

Contents lists available at ScienceDirect

Chemical Engineering Science: X

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

Model-based optimization of integrated purification sequences for biopharmaceuticals

Silvia M. Pirrung [a], Carmen Berends [a], Antoon H. Backx [a], Ruud F.W.C. van Beckhoven [b], Michel H.M. Eppink [c], Marcel Ottens [a,*]

[a] Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, the Netherlands
[b] DSM Biotechnology Center, Alexander Fleminglaan 1, 2613 AX Delft, the Netherlands
[c] Synthon Biopharmaceuticals B.V., Microweg 22, 6503 GN Nijmegen, the Netherlands

ARTICLE INFO

Article history:
Received 26 September 2018
Received in revised form 9 April 2019
Accepted 27 April 2019

Keywords:
Downstream processing (DSP)
Mechanistic modelling
Ultrafiltration
Diafiltration
Buffer exchange
High-throughput process development (HTPD)

ABSTRACT

Finding the best purification process is a challenging task. Recently, mechanistic models that can accelerate the development of chromatographic unit operations, the most important purification units, became widely available. In previous work, several chromatographic models have been linked together to simulate and optimize integrated processes. However, considering only chromatographic steps may lead to a suboptimal process. Consequently, the aim of this study was to include models for ultra- and diafiltration units into the optimization approach to account for buffer exchange steps before or between chromatography units. This approach was applied to an industrial case, the purification of a monoclonal antibody, where cation exchange, hydrophobic interaction and mixed mode were the possible chromatographic separation modes. It turned out that only the duration of the total filtration step and the duration of the ultrafiltration step were crucial variables for the optimization of the ultra- and diafiltration steps. The 'best' in silico purification process was found based on the performance criteria yield and solvent usage. The purity was required to be at least 99.9%.

© 2019 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The purification of biopharmaceuticals is an area of great interest for current research in academia and industry alike. High quality assurances are generally required for pharmaceutical products, which puts pressure on the downstream process. Moreover, costs of downstream purification units such as chromatography columns typically do not benefit much from economies of scale but instead scale at least linearly. Thus, the more is produced upstream, the higher will be the proportional cost of the downstream process. Therefore, purification processes might even be the bottleneck of the whole production process (Gronemeyer et al., 2014). This increasing importance of the downstream process clearly shows the need for better development and optimization approaches.

One way to achieve that, is the use of detailed mechanistic models based on first principles. Currently, these type of models are widely used for instance for the simulation and design of chromatography units (Meyer et al., 2018; Wang et al., 2016; Sellberg et al., 2015) and allow an easy exploration of the design space for a specific unit. Prior work has dealt with linking such chromatography models together and optimizing the resulting sequences (Nfor et al., 2013; Huuk et al., 2014). In a recent study (Pirrung et al., 2017), this was even done simultaneously using a combination of detailed mechanistic models and speed-enhancing artificial neural networks (ANNs). This novel approach opened up the way for more complex optimization problems. However, a downside in all these studies was that only chromatographic separation units were considered and not ultra/diafiltration (UF/DF) units, which can be placed between chromatography columns to adjust the buffer conditions and, thus, could alter the resulting optimal process (Huuk et al., 2014; Pirrung et al., 2017).

Thus, the aim of this work is to integrate filtration units into the latter approach, which would extend the general applicability of the approach greatly. To achieve that, filtration models are implemented and applied in the optimization approach. The purification of a monoclonal antibody, IgG1, from clarified cell harvest is used as a test case. Parameters for two chromatographic resins, a cation exchange (CEX) resin and a mixed mode chromatography (MMC) resin, were already available (Pirrung et al., 2018). Therefore, additional model parameters are only determined for one extra

* Corresponding author.
E-mail address: m.ottens@tudelft.nl (M. Ottens).

https://doi.org/10.1016/j.cesx.2019.100025
2590-1400/© 2019 The Author(s). Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

hydrophobic interaction (HIC) resin as well as the UF/DF units. Subsequently, the optimization approach is applied to sequences of filtration and chromatography units to find the best purification sequence.

2. Material and methods

2.1. Modelling and optimization

2.1.1. Optimization problem definition

The optimization problem was defined and treated in a similar manner as described in a prior study (Pirrung et al., 2017). A general constrained optimization problem can be described as:

$$\max \; f(x,y) \tag{1a}$$

$$\begin{aligned} s.t. \quad & \mathrm{h}(x) = 0 \\ & g(x,y) \leq 0 \\ & \mathrm{lb} \leq x \leq \mathrm{ub}, y \in \{0,1\} \end{aligned} \tag{1b}$$

where f is the objective function. The objective function depends on two type of variables: continuous x variables, which reflect either operating conditions or design parameters for each unit, and binary y variables, which define the investigated process sequences. In this study, the objective function includes the overall process yield and the overall solvent use, since solvent use is affected by chromatography and filtration units. The equality constraints $\mathrm{h}(x)$ include for instance mass balances. Other constraints can be defined using the inequalities $g(x,y)$. Here, the final product purity was defined to be at least above 99.9%.

This overall problem was split up into smaller problems twice. First, it was divided into a master problem and subproblems. The master problem is responsible for generating all possible process sequences, which is shown by the binary y variables. These variables tell, which unit operations are in a sequence and their order. The master problem here was formulated as:

$$\max \; f(y_{m,s}) \tag{2a}$$

$$\begin{aligned} \text{s.t.} \quad & \textstyle\sum_m y_{m,s} \leq 1 \\ & \textstyle\sum_s \sum_{m=1}^{3} y_{m,s} \leq 2 \\ & \textstyle\sum_s y_{m,s} \leq 1 \quad \text{for } 1 \leq m \leq 3 \\ & \textstyle 1 - y_{4,s} + \sum_{m=1}^{3} y_{m,s+1} \geq 1 \\ & \textstyle 1 - \sum_m y_{m,2} + \sum_m y_{m,1} \geq 1 \\ & \textstyle 1 - \sum_m y_{m,3} + \sum_m y_{m,2} \geq 1 \\ & \textstyle 1 - \sum_m y_{m,4} + \sum_m y_{m,3} \geq 1 \\ & y_{m,s} \epsilon \{0,1\} \end{aligned} \tag{2b}$$

where subscript m indicates the mode used, $m\epsilon\{1,2,3,4\}$ meaning CEX, HIC, MMC and UF/DF. The number of purification steps is given by s, $s\epsilon\{1,2,3,4\}$. The first constraint defines that maximal one unit can be used per purification step. The second constraint shows that only a maximum of two chromatography units can be in a process sequence. The third constraint then defines that each chromatography mode can only be included once. Filtration units can be used more often. However, they can only occur before a chromatography unit shown in the next constraint. Finally, the last constraints mean that a unit has to be chosen for all earlier occurring steps. A scheme with all resulting process alternatives is shown in Fig. 1.

Each generated process sequence than forms its own subproblem. In each subproblem, the operating variables x for each purification unit in the specific sequence are optimised simultaneously. Since this is still a rather complex optimization, the problem was divided again into a local search and a global search problem. The global search was performed to generate good starting points for the subsequent local search increasing the chance to find a global optimum. In this study, the global search problem was defined as

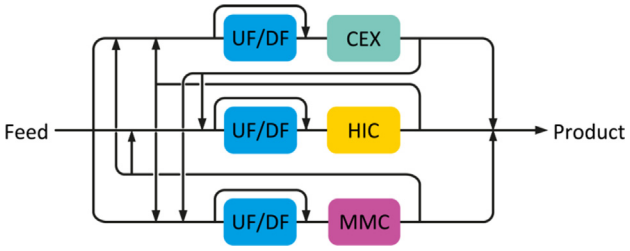

Fig. 1. Scheme with all considered process options. A maximum of two chromatographic units are allowed as defined by the second constraint in Eq. (2b).

$$\max \; f = yield + 2 * \mathrm{purity} \tag{3a}$$

$$\begin{aligned} s.t. \quad & \mathrm{h}(x) = 0 \\ & 0 \leq x \leq 1 \end{aligned} \tag{3b}$$

All variables were normalized, so that their lower boundary is zero and their upper boundary one. For an easier solution, no nonlinear constraints were included in the global search problem. Thus, the purity after the final purification unit was treated as an additional objective. Its importance over the other objectives was set higher to increase the likelihood of already feasible starting points for the local search, where purity is used as constraint. The solvent volume was not yet included in the objective, since this first optimization is primarily performed to find feasible starting points for the next optimisation. Potential yield loss in UF/DF was not taken into account, since the considered product, IgG1, cannot pass the membrane.

The global optimisation was run 40 times for each process sequence with random starting points, which were created with MATLAB's function *lhsdesign* (criterion set to correlation). As optimisation algorithm *patternsearch* was used with *searchlhs* as search method performing a complete search. The maximum number of function evaluations was set to 1000 and an output function was used, which could stop the search early, if the starting point was not promising. These and other settings are summarized in Table 1.

The x variables that were optimised for each chromatography unit were the length of the gradient elution, the length of the extra elution volume with the final eluent composition, the product pool cut points and the final salt concentration as shown in Fig. 2. The starting salt concentration as well as the pH are determined by the prior processing unit. For the filtration units, it was investigated in Section 3.1 which variables to include.

Table 1
Optimization settings with their MATLAB abbreviations in brackets.

Settings	Global optimization	Local optimization
Algorithm	*patternsearch*	*fmincon*
Search method (*SearchMethod*)	*searchlhs* with complete search	–
Initial mesh size (*Initial Mesh Size*)	0.5	–
Mesh contraction (*MeshContraction*)	0.5	–
Function tolerance (*TolFun*)	0.1	1e-6
X tolerance (*TolX*)	0.01	1e-8
Constraint tolerance (*TolCon*)	–	0.1
Maximum evaluations (*MaxFunEvals*)	1000	–
Relative step (*FinDiffRelStep*)	–	1e-3

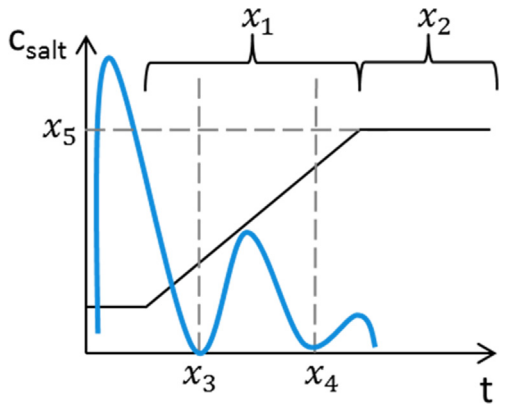

Fig. 2. Optimization variables for chromatography units.

During the global optimization, ANNs were used as surrogate models for the mechanistic models. This allows a faster evaluation of purity and yield at specific variable values. Since the ANNs sometimes find solutions that only exist due to inaccuracies in their own predictions (Pirrung et al., 2017), all starting points were checked with the mechanistic models. This is also why ANNs were only used during the global optimization. If the objective function value was below a specified value they were fed to the next optimizer.

The local search problem was then defined with the final purity as constraint and the solvent volume included in the objective:

$$max\ f = -V_{solvent} + yield \tag{4a}$$

$$\begin{aligned} s.t.\quad & h(x) = 0 \\ & purity(x) > 99.9\% \\ & 0 \le x \le 1 \end{aligned} \tag{4b}$$

The solvent volume, $V_{solvent}$, is the sum of solvent used per unit operation. Both objectives were normalized to give them a similar importance. The local search algorithm used was *fmincon*. The objective and constraint function are approximated by finite differences. Since the values of these functions are numerical solutions of ordinary and partial differential equations, the relative step of the finite difference should not be too small or too big. If it is too small, it might be for instance that there is no change in the objective function yet. If it is too big, nonlinearity might influence the finite difference. A value of 1e-3 was chosen as recommended by MATLAB for optimization problems including ordinary differential equations. All settings are summarized in Table 1.

2.1.2. Filtration modelling

The ultrafiltration/diafiltration (UF/DF) process can be described by a set of ordinary differential equations (ODE) for a variable volume diafiltration process (Foley, 2013):

$$\frac{dV}{dt} = (\alpha - 1)JA \tag{5}$$

$$\frac{dc_i}{dt} = \frac{c_i}{V}(\sigma_i - \alpha)JA \tag{6}$$

$$\frac{dV_{diluent}}{dt} = \alpha JA \tag{7}$$

where α is the ratio between diluent inflow and outflow over the membrane, $\alpha = Q/JA$ with Q the volumetric flowrate of the diluent, J the flux over the membrane and A the membrane area. V is the volume of the solution and c_i is the concentration for a compound i. The rejection coefficient σ_i defines how well compounds are being retained by the membrane. For instance for the salt, the rejection coefficient is 0, because it can freely pass the membrane. If the added diluent buffer already contains salt ($c_{s,diluent}$), Eq. (6) needs to be adjusted as follows to describe the change in salt concentration ($i = s$):

$$\frac{dc_s}{dt} = \frac{c_s}{V}(\sigma_s - \alpha)JA + \frac{c_{s,diluent}}{V}\alpha JA \tag{8}$$

The flux J can be calculated with the osmotic pressure model:

$$J = \frac{\Delta P - \Delta\pi}{\mu R_m} \tag{9}$$

where μ is the viscosity of the permeate stream, $\Delta\pi$ is the osmotic pressure and ΔP the transmembrane pressure (TMP), which is defined as:

$$\Delta P = \frac{P_{feed} + P_{retentate}}{2} - P_{permeate} \tag{10}$$

The osmotic pressure can be expressed by a virial expansion. An expansion with two virial coefficients was chosen; the third coefficient can be negligible in low concentration ranges (van Reis et al., 1997). In case of a final formulation step, where much higher concentration ranges occur, the third coefficient would need to be included. Additionally, the spatial variation in transmembrane pressure would need to be taken into account. This situation was extensively studied by Binabaji (2015).

$$\frac{\pi}{c_{i,wall}RT} = \frac{1}{M} + B_{22}c_{i,wall} + \cdots \tag{11}$$

where $c_{i,wall}$ is the concentration of protein at the wall or membrane surface. B_{22} is the second osmotic virial coefficient. In UF/DF processes, a layer forms almost immediately as a result of concentration polarization, when a solution with macromolecules is to be filtered (Baker, 2004). Assuming the instantaneous formation of this layer; $c_{i,wall}$ can be calculated by the stagnant film model (Zydney, 1997):

$$J = k \ln \frac{c_{i,wall} - c_{i,permeate}}{c_{i,bulk} - c_{i,permeate}} \tag{12}$$

where $c_{i,bulk}$ is the protein concentration in the bulk solution assuming that it is well mixed and $c_{i,permeate}$ is the concentration in the permeate. The mass transfer coefficient k is defined as the ratio of diffusivity D to the thickness of the boundary layer. It can be described depending on the Sherwood number, a commonly used dimensionless number (Da Costa et al., 1994):

$$Sh = \frac{kd_h}{D} = aRe^b Sc^c \left(\frac{d_h}{l}\right)^d \tag{13}$$

where Re is the Reynolds number defined with the density ρ and the cross membrane velocity υ as $Re = \rho\upsilon d_h/\mu$. The Schmidt number Sc can be calculated as $Sc = \mu/\rho D$. The diffusivity D was calculated using the Young correlation valid for globular proteins (Young et al., 1980).

The characteristic length of the system is shown as l. The membrane cassette used here is a flat channel with spacers enhancing mass transfer. For this module geometry, the characteristic length is the size of one mesh (Da Costa et al., 1994), which can be calculated as the sum of the mesh opening and twice the wire diameter. For the C screen type membrane cassettes, these parameters can be found in recent literature (Lutz et al., 2017). The mesh opening is thus 350 μm, the wire diameter 270 μm resulting in a total characteristic length of 890 μm. The empirical constants a, b, c and d are dependent of the specific geometry of the system as well and can be found in literature for most systems. Here, a is 0.664, b 0.5, c 0.33 and d 0.5 respectively (Da Costa et al., 1994). Additionally,

the hydraulic diameter d_h is defined as $4 * cross\ section / wetted\ perimeter$ and, thus, depends on the system used. In this case, it can be defined as (Lutz, 2015):

$$d_h = 4h\frac{\varepsilon_s}{1+\frac{2(1-\varepsilon_s)h}{r}} \tag{14}$$

where ε_s is the porosity of the spacer, h is the half-height of the channel and r is the fibre radius. Furthermore, the cross-membrane velocity can be calculated as $\upsilon = J/(a_c\varepsilon_s)$, where a_c is the ratio of feed channel area to membrane area. The spacers in the membrane cassette used were Screen C type (ε_s =0.63; h =0.026 cm; r =0.014 cm; a_c =0.0018) (Lutz, 2015).

With increasing protein concentration, the viscosity of the solution changes, which is why the dependence of the mass transfer coefficient on viscosity was incorporated. An exponential relationship between viscosity and protein concentration, c_i, can often be assumed: $\mu = \mu_0 e^{\vartheta c_i}$, where ϑ is a constant depending on the molecule and μ_0 the solvent viscosity. In this study, ϑ was found to be 0.017±0.001 L/g. It was determined by measuring the viscosity of the IgG solution at different protein concentrations using a viscometer. Subsequently, a least square fit was performed using MATLAB's *lsqcurvefit* function. The relationship between mass transfer coefficient and viscosity was derived from Eq. (13) with the module specific constants b and c:

$$k = k_0\left(\frac{\mu}{\mu_0}\right)^{-\frac{1}{6}} = k_0 e^{-\frac{1}{6}\vartheta c_i} \tag{15}$$

The initial mass transfer coefficient, k_0, was calculated using Eq. (13). Assuming that the boundary layer is sufficiently thin, the viscosity was evaluated using the bulk concentration $c_{i,bulk}$.

At initial conditions, it was assumed that the flux predicted by the osmotic pressure model is identical to the flux due to the immediately formed concentration polarization. Another assumption made was that the antibody has by far the most impact due to its higher molecular mass and concentration ($i = mAb$). Additionally, the antibody is completely retained by the membrane. Based on these assumptions, Eqs. (9) and (12) could be equated as follows:

$$k\ln\frac{c_{mAb,wall}}{c_{mAb,bulk}} - \frac{\Delta P - \Delta\pi}{\mu R_m} = 0 \tag{16}$$

This function was solved by the MATLAB function *fsolve* for finding the initial wall concentration.

The change of wall concentration over time was found by implicit differentiation of Eq. (16) as shown for a similar case in detail by Foley (2013) and incorporated in the ODE system comprised of Eqs. (5)–(8). An additional ODE described the fraction of diluent in the system at a specific time. The ODE system was solved using the ODE solver *ode15s* from MATLAB.

2.1.3. Filtration optimization

Different objectives can be defined in the optimization of UF/DF units such as processing time (Foley, 1999; Paulen et al., 2011), economic factors (Fikar et al., 2010), minimum amount of diluent (Paulen et al., 2011), product loss, if the product is not fully rejected by the membrane, or a combination of all of these factors (Fikar, 2014). For the optimization of only the filtration step, the minimum amount of diluent was chosen as single objective, since the results can be easily evaluated experimentally. Therefore, the optimization problem was formulated as:

$$max\ f = \min_{\alpha(t)} \int_{t_0}^{t_{end}} \alpha(t) A J(t) dt \tag{17a}$$

$$\begin{aligned} s.t.\quad & V(t_0) = V_0; c_i(t_0) = c_{i,0}; c_i(t_{end}) = c_{i,end} \\ & 0 \le \alpha(t) \le 1 \\ & \alpha(t) = \textstyle\sum_{k=1}^{N} \alpha_k \end{aligned} \tag{17b}$$

where $\alpha(t)$ is described as a piecewise constant for N time intervals. α_k is then the value α assumes during a specific step k. The duration of each step k, the value of α_k during the step and the overall duration of the ultrafiltration/diafiltration process were used as variables. To define constraints for an easier comparison, a simple base case was simulated with only one step and a constant α(α = 0.75). The final concentrations experienced during the base case were used as constraints,$c_{i,end}$. In this base case, the diafiltration was performed for 20 min, the concentration of the antibody was increased from 1.0 to 5.6 g/L and the salt concentration was decreased from 58.4 to 0.4 g/L. In total, 123 mL of diluent were consumed.

The optimizations were performed with MATLAB's function *fmincon*. The same settings as described above for the local optimization were used. Each optimization was performed with 20 different starting points, which were generated with the function *lhsdesign* and the criterion set to correlation.

2.1.4. Chromatography modelling

The equilibrium transport dispersive model can be applied to model chromatography columns (Guiochon et al., 2006). The following mass balance is used to describe the mobile phase (Eq. (18)):

$$\frac{\partial c_i}{\partial t} + \frac{1-\varepsilon_b}{\varepsilon_b}\frac{\partial q_i}{\partial t} = -\upsilon\frac{\partial c_i}{\partial x} + D_{L,i}\frac{\partial^2 c_i}{\partial x^2} \tag{18}$$

where c_i is the concentration of protein i in the bulk phase, ε_b is the bed porosity and $D_{L,i}$ is the axial dispersion coefficient. The interstitial velocity of the mobile phase υ is defined as $\upsilon = u/\varepsilon_b$ with u, the superficial velocity of the mobile phase.

The change in q_i, the concentration in the stationary phase, over time was approximated with the linear driving force approach for the mass transfer in the liquid phase (Eq. (19)).

$$\frac{\partial q_i}{\partial t} = k_{ov,i}(c_i - c_i^*) \tag{19}$$

where $k_{ov,i}$ is the overall mass transfer coefficient. An appropriate adsorption isotherm needs to be used to calculate c_i^*, the protein concentration inside the particle pores. One possible option is the mixed-mode isotherm developed based on thermodynamic principles by Nfor et al. (2010):

$$\frac{q_i}{c_i^*} = A_i\left(1 - \sum_{j=1}^{m}\frac{q_j}{q_j^{max}}\right)^{\nu_i+n_i} \tag{20}$$

Its validity range spans mixed-mode chromatography, ion-exchange chromatography and hydrophobic interaction. m stands for the number of proteins, q^{max} for the maximum binding capacity and j for the protein species. The stoichiometric coefficient in hydrophobic interaction chromatography is n_i and in ion exchange chromatography ν_i respectively. ν_i can be calculated as z_p/z_s with z_p, the effective binding charge of the protein, and z_s, the charge on the salt counter ion.

The initial slope of the isotherm A_i is defined as:

$$A_i = K_{eq,i}\Lambda^{(\nu_i+n_i)}(z_s c_s)^{-\nu_i} c_v^{\ -n_i}\gamma_i \tag{21}$$

where K_{eq} is the thermodynamic equilibrium constant, c_s the salt concentration, c_v the molarity of the solution in the pore volume and Λ the ligand density. The activity coefficient can be expressed as

$$\gamma_i = e^{2B_{22}c_i} \tag{22}$$

Assuming that protein-protein interactions are only of importance between one protein species and that these interactions mostly occur between two molecules (Pirrung et al., 2018). Additionally, it is assumed that salt protein interactions are negligible, which is valid when salts with a low salting-out effect are used such as chlorides (Mollerup et al., 2008). For hydrophobic interaction chromatography, however, sulphates are commonly used, which makes this assumption invalid. Therefore, the salting-out constant K_s and the salt concentration were included in the definition of the activity coefficient in case sulphates are used:

$$\gamma_i = e^{K_{s,i}c_s + 2B_{22}c_i} \tag{23}$$

The mechanistic model was solved as described in (Pirrung et al., 2017) employing the mass transfer correlations as shown in Pirrung et al. (2018).

2.1.4.1. Artificial neural networks. Artificial neural networks (ANNs) were generated and trained with the Neural Network Toolbox™ by MATLAB as described previously (Pirrung et al., 2017). Here, seven variables were used: the gradient length, the length of the extra elution volume after the end of the gradient, both product pool cut points, the injection volume, the starting and final salt concentration as well as the pH. Different ANNs were trained to predict the protein concentration for all proteins in the product pool, the volume of the product pool, the solvent volume as well as the salt concentration of the product pool. The neural networks were trained for different number of sample points (1000 or 2000), hidden layer sizes (10 to 40) as well as different amounts of hidden layers (1 or 2); networks with the best R^2 were chosen further. Typically an R^2 above 0.8 indicates sufficient predictive power (Forrester et al., 2008).

2.1.5. Computations

All computations were performed on an Intel® Xeon® Processor E5-1620 v2 with 3.7 GHz. MATLAB's Parallel Computing Toolbox™ was used to compute in parallel on four cores whenever possible.

2.2. Materials

2.2.1. Sample preparation

A clarified Chinese hamster ovary (CHO) cell culture supernatant containing a monoclonal immunoglobulin G (IgG1) was used as sample with a concentration of 1.3 mg/mL of IgG1 and a pH pf 7.72. The pI of IgG1 was determined to be 8.6 by capillary isoelectric focusing. Additionally, a sample of the same antibody after Protein A purification was used.

2.2.1.1. Buffers and chromatography resins. The buffers used in this study are shown in Table 2. In the chromatographic steps, the acetic acid buffer was mainly used for the cation exchange resin Poros 50 HS (Thermo Fisher Scientific Breda, Breda, The Netherlands), the MOPS buffer for the mixed mode resin Capto MMC (GE Healthcare, Eindhoven, The Netherlands) and the Tris-HCl buffer for the hydrophobic interaction resin Cellufine Phenyl (AMS Biotechnology, Abingdon, United Kingdom). All buffers were purchased in buffer grade.

2.3. Experimental techniques

2.3.1. Filtration

Fig. 3 shows a scheme of the experimental set-up used for all UF/DF experiments. The sample solution was added to the feed tank, from where it was continuously circulated over an 88 cm^2 Pellicon 3 Ultracel 30 kDa membrane cassette placed in a cassette holder (Merck Millipore, Amsterdam, The Netherlands) by a peristaltic Masterflex L/S pump (Metrohm Netherlands B.V., Barendrecht, The Netherlands). If desired, diluent solution was added with a LC-8A HPLC pump (Shimadzu, s'Hertogenbosch, The Netherlands). The mass of permeate and retentate was constantly recorded with the help of PG 3001-S scales (Mettler Toledo, Tiel, The Netherlands). Signals from all other sensors were sent as an analogous signal to the data acquisition device DAQ USB 6009 (National Instruments Netherlands BV, Woerden, The Netherlands), which converts them to digital signals. The program SignalExpress (National Instruments Netherlands BV, Woerden, The Netherlands) was then used to process the data.

2.3.1.1. Initial membrane resistance. The flux of a pure water stream J was measured at a flowrate of 20 mL/min. Subsequently, the initial membrane resistance was calculated to be $7.92 \pm 0.44 * 10^{12}$ 1/m using the following equation and the viscosity of water:

$$R_m = \frac{\Delta P}{\mu J} \tag{24}$$

2.3.1.2. Rejection coefficient. A relation for the rejection coefficient, which describes how well proteins are retained by the membrane, was found based on model molecules as listed in Table 3. Model proteins were used instead of dextrans, which are linear polymers, since not only molecular weight but also molecule structure influences retention behavior (Baker, 2004). For each experiment, the protein concentration was 0.2 g/L.

2.3.1.3. Second virial coefficient. The second virial osmotic coefficient was determined and fitted for the Tris-HCl buffer as described in detail elsewhere using self-interaction chromatography (Pirrung et al, 2018). In that study, data for IgG1 in the acetate and MOPS buffer can also be found as shown in Table 4.

A second order polynomial function was fitted to the determined B_{22} values using MATLAB's *fit* function:

$$B_{22} = b_1 + b_2 pH + b_3 c_s + b_4 pH c_s + b_5 pH^2 + b_6 c_s^2 \tag{25}$$

The salt concentration is shown as c_s. The resulting B_{22} was in the units (mol mL)/g^2 as required for Eq. (11). To use the determined B_{22} in the chromatographic mechanistic model as shown in Eqs. (22) and (23), the units need to be changed to L/mol with the squared molecular weight.

The pH during the diafiltration was assumed to be a sum of the pH of each buffer multiplied with its fraction of the total solution. This assumption was proven to be applicable through mixing experiments (data not shown).

2.3.1.4. Validation experiments. Experiments were performed at least in duplicates. The TMP was kept constant during the experi-

Table 2
Buffer specifications.

Buffer type	Supplier	pH	Buffer [mM]	Salt type	Salt range [mM]
Acetic acid	Sigma Aldrich, Zwijndrecht, The Netherlands	4.5	25	Sodium chloride	0–1000
MOPS	Applichem GmbH Darmstadt, Germany	6.75	25	Sodium chloride	0–1000
Tris-HCl	Sigma Aldrich, Zwijndrecht, The Netherlands	7.5	25	Ammonium sulphate	0–750

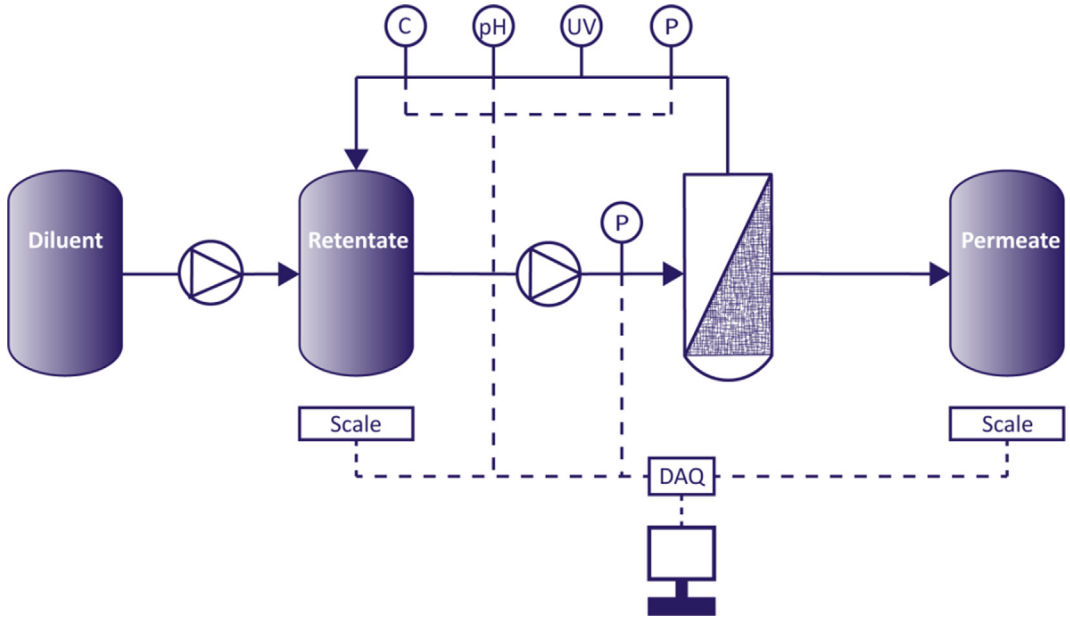

Fig. 3. Scheme of the experimental setup with the ultrafiltration/diafiltration membrane.

Table 3
Model proteins used to determine the rejection coefficient.

Name	Molecular weight (kDa)	Stokes radius (nm)
Lysozyme	14.3	1.86
α-Chymotrypsin	25.0	2.09
Albumin from hen egg white	44.3	2.93
Albumin from bovine serum	66.5	3.48

ments. The starting concentration of IgG1 was 1 g/L. Modelling results were compared with experimental data with help of the coefficient of determination (R^2), which was calculated from the correlation coefficients given by MATLAB's *corrcoef* function.

2.3.1.5. Cleaning. After each use, the membrane was cleaned according to the manufacturer's instructions.

2.3.2. Chromatography

2.3.2.1. Ion exchange and mixed mode chromatography. Model input parameters for the cation exchange resin, POROS 50 HS, and the mixed mode resin, Capto MMC, were taken from literature (Pirrung et al., 2018).

2.3.2.2. Hydrophobic interaction chromatography. Parameters for the hydrophobic interaction resin, Cellufine Phenyl, were determined by a 3D liquid chromatography method as described in Hanke et al. (2016), Tsintavi (2015). The first dimension, which is mainly to reduce the complexity of the mixture and to allow a focus on critical impurities, is identical to the one shown in Pirrung et al. (2018). Here, the mixture was prefractioned twice, once on a cation exchange resin and once on an anion exchange resin. However, more fractions were analysed further, since the adsorption behavior on cation or anion exchange resins is quite different from the adsorption behavior due to hydrophobic interaction. To avoid correlation errors during the regression, several parameters, which can be assumed as constant in the investigated range, were combined with the equilibrium constant to form one parameter ($K_{comb,i} = K_{eq,i}(\Lambda/c_v)^n$) (Hanke et al., 2016). The isotherm parameters K_{comb} and K_s for proteins contained in these fractions were determined on RoboColumns with a volume of 200 µL prefilled with the resin of interest, Cellufine Phenyl, by Repligen (Weingarten, Germany). For that, three types of gradient elution experiments (12, 24 and 36 CV gradient length) were performed for each fraction. The salt concentration in the beginning of the gradient, $c_{s,in}$, was 1 M ammonium sulphate and the salt concentration in the end, $c_{s,f}$, 0 M respectively. The following equation as derived in Hanke et al. (2016) was used to fit the resulting data:

$$V_{R,g,i} = \frac{V_G}{-K_{s,i}(c_{s,f} - c_{s,in})} \times \ln\left(1 + V_{column}(1-\varepsilon_b)\varepsilon_p K_{D,i}\frac{-K_{s,i}(c_{s,f} - c_{s,in})}{V_G} K_{comb,i} e^{K_{s,i}c_{s,in}}\right) \quad (26)$$

As determined in Hanke et al. (2016), the bed porosity of the RoboColumns filled with Cellufine Phenyl was 0.3 and the particle porosity 0.93. The other isotherm parameters q_{max} and n were determined in batch uptake experiments as described in Pirrung et al. (2018). Different to the method described there, the salt concentration to perform these experiments was set to 0.75 M of ammonium sulphate, since binding capacities are typically higher in HIC at higher salt concentrations.

2.3.2.3. Validation experiments. Validation experiments were performed on OPUS® ValiChrom 11.3/100 columns prepacked with the resin Cellufine Phenyl by Repligen (Weingarten, Germany) on

Table 4
Parameters to calculate the B_{22} for IgG1 in the acetate and MOPS buffer taken from Pirrung et al. (2018). They need to be multiplied with 10^{-4}.

	b_1	b_2	b_3	b_4	b_5	b_6
Acetate buffer	6.791	−2.794	1.249	−0.575	0.237	1.474
MOPS buffer	2.119	−0.807	−0.199	0.222	0.046	−1.013

an Äkta Avant 25 (GE Healthcare, Uppsala, Sweden). The bed porosity was determined with the help of the Blake-Kozeny equation that correlates pressure drop and applied flowrate. The flowrate was 182 cm/h. Linear salt gradients of 12 CV were used during the elution. 20% Ethanol was used as storage solution. Absorption was recorded at 210, 230 and 280 nm.

2.3.3. Quantification

If not defined otherwise, protein concentrations were determined in an ultra-high performance liquid chromatography (UHPLC+) (Thermo Fisher Scientific, Waltham, USA) system as described by Hanke et al. (2016).

2.3.4. Protein identification

A mass spectrometric (MS) analysis was performed for the fractions of the cation exchange prefractionation. The samples were desalted and concentrated prior to a proteolytic digestion, which was based on literature (Wang et al., 2016). They were then analysed with liquid chromatography followed by mass spectrometry (LC-MS). Protein names of all proteins and their intensity in each fraction were determined with the help of UniProt. After that, they were matched with the tracked proteins. Additionally, the isoelectric point for the antibody was measured externally with capillary isoelectric focusing (CIEF). Isoelectric points for the impurities were estimated using ExPASy (the Expert Protein Analysis System, a bioinformatics resource portal by the SIB Swiss Institute of Bioinformatics).

3. Results and discussion

The first section focusses on UF/DF units. Here, parameter determination, modelling and subsequent validation of the model is shown. Additionally, a short section on the optimization of UF/DF units is included. The second section is about chromatography units. It starts by describing the chromatographic model input parameters and then shows the mechanistic model together with its validation. This section finishes with a short discussion on the training of ANNs to predict the chromatography units. The last section shows the results of the process optimisation including the best found process option.

3.1. Filtration

3.1.1. Model parameters

3.1.1.1. Rejection coefficient. To understand how well differently sized proteins are retained by the membrane, sieving experiments were performed. The rejection coefficient was approximated with a logistic function and the molecular weight, MW, in kDa as:

$$\sigma_i = \frac{1}{1 + (a_1 * MW_i)^{a_2}} \tag{27}$$

The experimental results together with the fitted curve can be found in Fig. 4. As expected, smaller proteins up to 15 kDa can easily pass the 30 kDa membrane. Above 50 kDa, proteins are almost fully retained. The fitted parameter a_1 was found to be 0.037±0.001 and a_2 −5.9±1.1. For an improved description of the retention of impurities, the dependence of the rejection coefficient on the flux should be taken into account additionally.

3.1.1.2. Second virial coefficient. All B_{22} results for the Tris-HCl buffer are summarized in Fig. 5. Fig. 5a shows the second order polynomial function. It was fitted to the experimentally determined B_{22} values resulting in the following constants $\times 10^{-4}$: b_1=−7.016; b_2=2.145; b_3=2.431; b_4=−0.226; b_5=−0.187 and b_6=−1.407. Ammonium sulfate concentrations greater than 0.75M were not investigated due to stability issues such as aggregation. Two general trends can be observed from the results: First, the B_{22} values decrease with increasing pH. Second, salt concentration only has a strong influence on protein-protein interactions at high pH values. Both trends can well be explained by the fact that the IgG is getting closer to its pI (around 8.6) and therefore is less charged. With a decrease in protein charge, repulsive interactions between IgGs decrease as shown by the decrease in B_{22}.

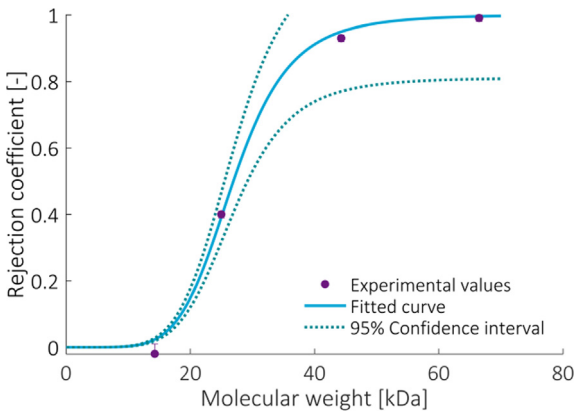

Fig. 4. Rejection coefficient.

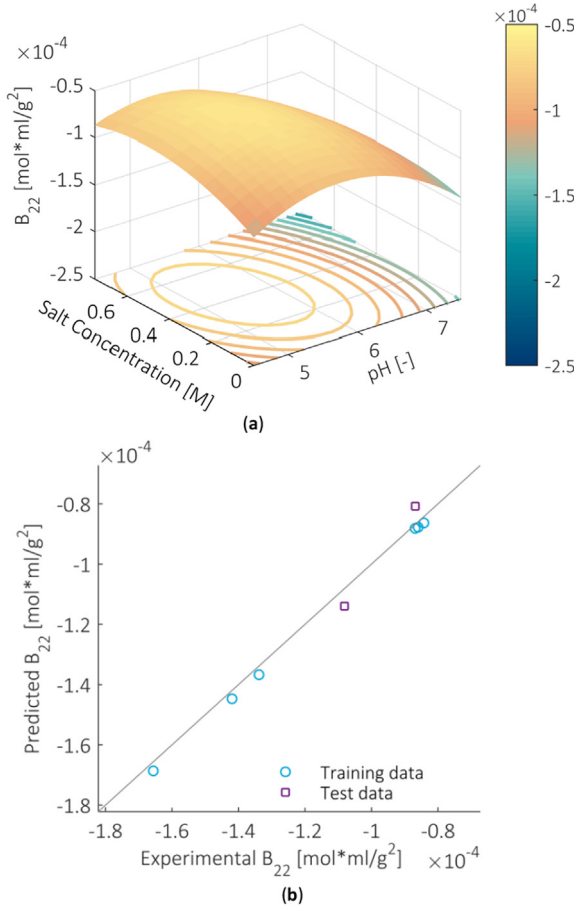

Fig. 5. B_{22} values of IgG1 as a function of salt concentration and pH for the Tris-HCl buffer. a: Second order polynomial function that was fitted on experimental data; b: Comparison of experimentally obtained B_{22} values with values given by the polynomial function.

The correlation between B_{22} values predicted by the polynomial and the experimentally determined ones is plotted in Fig. 5b. A good correlation was also found for two test data points, that were not included in the data set used for fitting.

3.1.2. Modelling and validation

Fig. 6 compares model predictions for UF/DF processes with experimental data. An excellent agreement (R^2 of at least 0.99) was found in all cases supporting the validity of the model approach employed. In the ultrafiltration experiment in Fig. 6a, only a small deviation between experimental values and model predictions was found indicating that the third virial coefficient as well as fouling of the membrane are indeed negligible in the investigated concentration range of up to around 15 g/L. This was not clear beforehand considering the high wall concentrations, which were predicted to be around 5 g/L at the start of the experiment and around 85 g/L at the end.

Fig. 6b and c show different dilution strategies; even the results for the most complex dilution strategy was well predicted as is shown in Fig. 6b. Here a random trajectory with ten different steps was chosen. In conclusion, the UF/DF model can well be used in the investigated concentration range.

These validation experiments were performed with a purified solution of IgG1. To see, if other proteins have an additional impact on the flux, experiments with the clarified cell harvest were performed at different salt concentrations, buffer compositions and pH values. Fluxes predicted fell in the range of experimental error ($1.3\ 10^{-6}$ m/s) of experimentally determined fluxes, which is why the prediction with only osmotic data of IgG1 was declared sufficient.

Experiments, where buffer was exchanged from initial clarified cell harvest, resulted in unexpected aggregation. Therefore, the final protein concentration for the initial diafiltration step was constrained to stay below 5 g/L. For filtration units that were placed between chromatography units it was constrained to stay below 10 g/L. If higher protein concentrations are being used, as would be needed for instance for a final formulation step, the influence of fouling on the ultra/diafiltration needs to be investigated in more detail.

3.1.3. Optimization

The purpose of the optimization as described in Eq. (17) was the reduction of diluent volume. Typically, an ultrafiltration/diafiltration process is performed as follows: First step, the volume is reduced until the product concentration reaches the desired value ($\alpha = 0$); second step, the diafiltration is performed until the desired salt concentration or pH is achieved ($\alpha = 1$). It was investigated if a higher flexibility in α, meaning more steps with variable α and variable duration, would lead to a further reduction in diluent volume. Consequently, optimizations were performed with two, five, ten and forty steps.

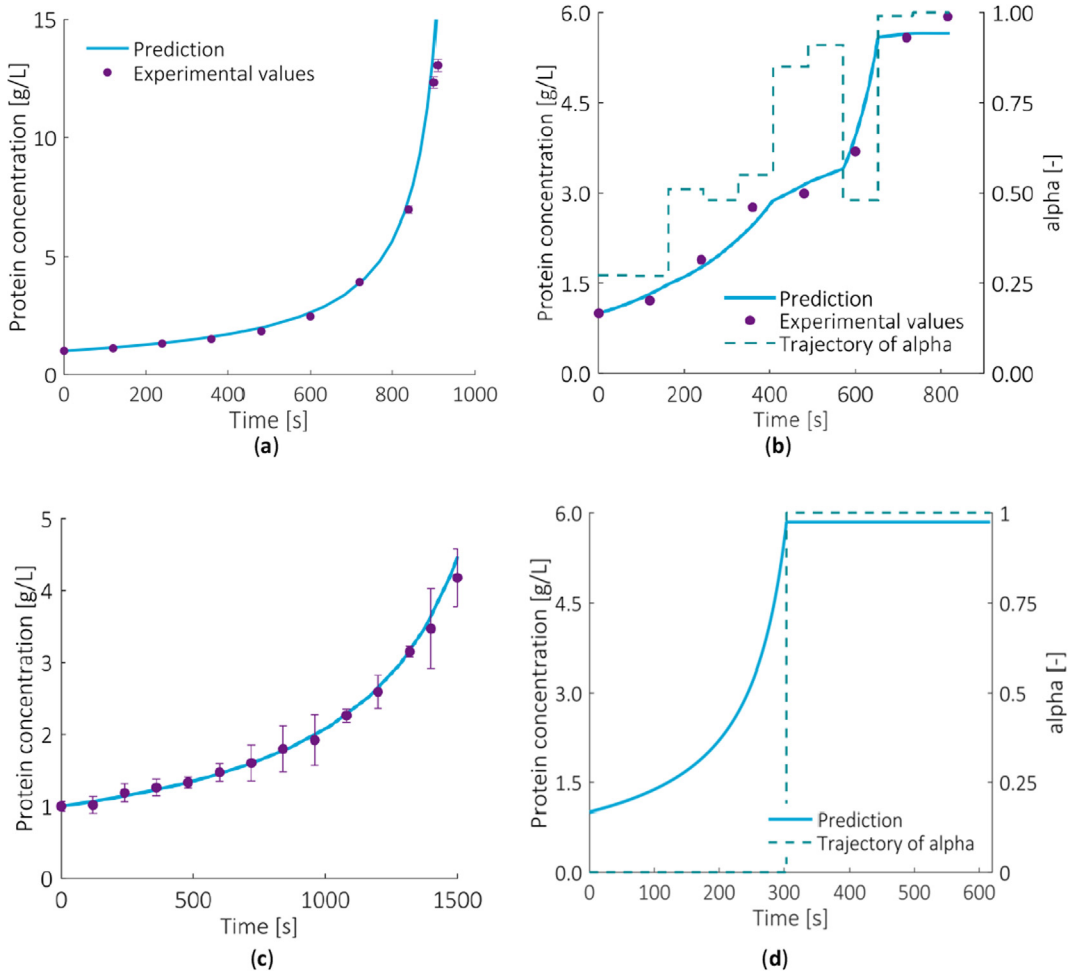

Fig. 6. Model predictions for comparison with ultrafiltration/diafiltration experiments (a), (b) and (c) and the optimization result (d). a: Compared with experimental values for an ultrafiltration experiment. The TMP was kept constant at 93 kPa. b: For a diafiltration experiment with alpha as a piece wise constant following the shown trajectory and a constant TMP of 93 kPa; c: For a diafiltration experiment with constant alpha of 0.75. The TMP was constant at 123 kPa; d: Regardless of the number of steps.

Table 5
Isotherm parameters regressed from retention volume curves determined in RoboColumns with their standard deviation as taken from Pirrung et al. (2018) for POROS 50 HS and Capto MMC or determined in this study for Cellufine Phenyl.

Resin	Protein ID		rh [nm]	Keq or Kcomb [–]	ν [–]	Ks [L/mol]
	AEX	CEX				
POROS 50 HS	1		2.4	12.6±0.54	2.9±0.5	–
POROS 50 HS	2	1	4.2	34.6±1.7	9.8±1.3	–
POROS 50 HS	3		2.7	2.2±0.2	7.4±0.7	–
POROS 50 HS	4	2	2.2	177.1±16.1	5.4±1.1	–
POROS 50 HS	6	4	2.2	0.9±0.8	7.0±0.3	–
POROS 50 HS	7	3	4.2	2.0±0.2	2.5±0.2	–
POROS 50 HS	8		2.4	0.2±0.1	16.9±6.4	–
Capto MMC	2	1	4.2	51.5±2.1	3.6±0.4	–
Capto MMC	4	2	2.8	16.6±5.8	4.7±1.6	–
Cellufine Phenyl	2	1	3.5	13.9±6.4	–	4.7±2.0
Cellufine Phenyl	4	2	2.2	6.4±3.2	–	3.3±1.0
Cellufine Phenyl	6	4	2.2	11.3±3.9	–	6.4±6.2
Cellufine Phenyl	7	3	3.5	17.0 ± 2.4	–	3.8 ± 0.5

MW: molecular weight; Keq or Kcomb: equilibrium constant; ν: stoichiometric coefficient; Ks: Salting out constant.

In all cases, the same optimal result was found, which is shown in Fig. 6d. The minimal diluent volume needed was 43 mL which is a great reduction of 65% compared to the base case. This result shows that the traditional way of operating is indeed optimal in regards to diluent usage confirming the finding by Paulen et al. (2011). However, the same group showed in a later study that more complex control strategies for α can be beneficial in certain cases (Paulen et al., 2012). Since no benefits were found by using more complex profiles for α in this case, the typical two step profile was chosen as described above in all further studies. Thus, only the total length of the ultrafiltration/diafiltration process and the duration of the ultrafiltration step were used as variables further on, which greatly reduces the complexity of the optimization problem. The duration of the diafiltration step was calculated as difference between total length of the filtration step and the ultrafiltration step.

In general, this is a sensible approach. First the volume is reduced until the desired product concentration is reached leading to a smaller amount of diluent that is needed to exchange the buffer. However, it should be kept in mind that this way of processing means that the buffer is exchanged at the highest protein concentration. This might cause stability problems, which should be investigated first.

3.1.3.1. Speed. The time needed to evaluate the filtration model is very fast compared to the chromatography model. It only takes in between 0.01 and 0.05 s, when the relative tolerance is set to 10^{-4} and the absolute tolerance to 10^{-6}. Thus, the speed increase due to artificial neural networks (around 0.005 s per evaluation) would not compensate for the initial time needed to train them and the general loss of accuracy. The complete optimization for a two-step process with five variables took with 20 different starting points around 200 s for a parallel computation on four cores. In 14 of these optimizations, the constraints could be fulfilled and the same minimal diluent was found.

3.2. Chromatography

3.2.1. Model parameters

The parameters regressed from all linear gradient experiments with the HIC resin Cellufine Phenyl are summarized in Table 5 together with the results taken from previous work for the other two resins (Pirrung et al., 2018). There it was shown that proteins can be given an ID and tracked through different experiments by knowing, which proteins are in each fraction and what their hydrodynamic radius is. In the case of Cellufine Phenyl, two prefractionations were performed with an anion exchange and a cation exchange resin. Many fractions were taken during the prefractionations, so that all impurities with a high abundancy could be characterised by their hydrophobicity.

Determining hydrophobic interaction parameters for fractions of both prefractionations allowed the matching of protein ID's from the different prefractionations to each other. Since the hydrodynamic radius of proteins, which were known to be identical, varied in the different experiments on Cellufine Phenyl, the peak volume as well as the elution behaviour were used in addition for this protein matching. Nonetheless, the varying hydrodynamic radii made the ID matching less conclusive. Therefore, the worst case scenario, which means the most challenging separation, was considered: if a protein showed near identical behaviour on HIC even though the hydrodynamic radius was different, it was assumed to be the same protein.

Even so, it was difficult to assess if any protein of the AEX prefractionation is equivalent to the protein with ID 2 from the CEX prefractionation (ID2 CEX). It was assigned to be the same protein as ID4 AEX but it might also be equivalent to ID3 AEX. To eliminate any ambiguity in future studies, it would be recommended to perform all parameter determination studies with fractions of both prefractionations. Another option would be to analyse all fractions of the prefractionations by MS.

How the model parameters for the hydrophobic interaction resin were determined is shown in Fig. 7 taking the product as example. The product IgG1 was assigned ID 2 during the anion exchange prefractionation (ID2 AEX) and ID 1 during the cation exchange prefractionation (ID1 CEX). In Fig. 7(a), experimental retention volumes for IgG1 (ID2 AEX & ID1 CEX) at different gradient lengths are shown with markers; the curves resulted from the parameter fitting using Eq. (26). Fraction 2 and 3 were taken during the AEX prefractionation, while fraction 4 and 5 were taken during the CEX prefractionation. In all other fractions, the concentrations of IgG1 were too small to be analysed. The retention volumes found for IgG1 in the different fractions were quite similar to each other, showing that the prefractionation method does not change the hydrophobicity of the protein.

The maximal capacity of the resin for IgG1 was determined with batch uptake experiments as shown in Fig. 7b. It was found to be 31.8±4.7 g/L. This capacity is in the range reported for other types of proteins on the same resin (JNC-Corporation, 2017). The hydrophobic stoichiometric coefficient was regressed to be 2.8±0.9. The particle size average (85 μm) was taken from the technical data sheet supplied by the manufacturer (JNC-Corporation, 2017). The pore size was estimated to be 45 nm based on a published pore size accessibility curve (Hanke et al., 2016).

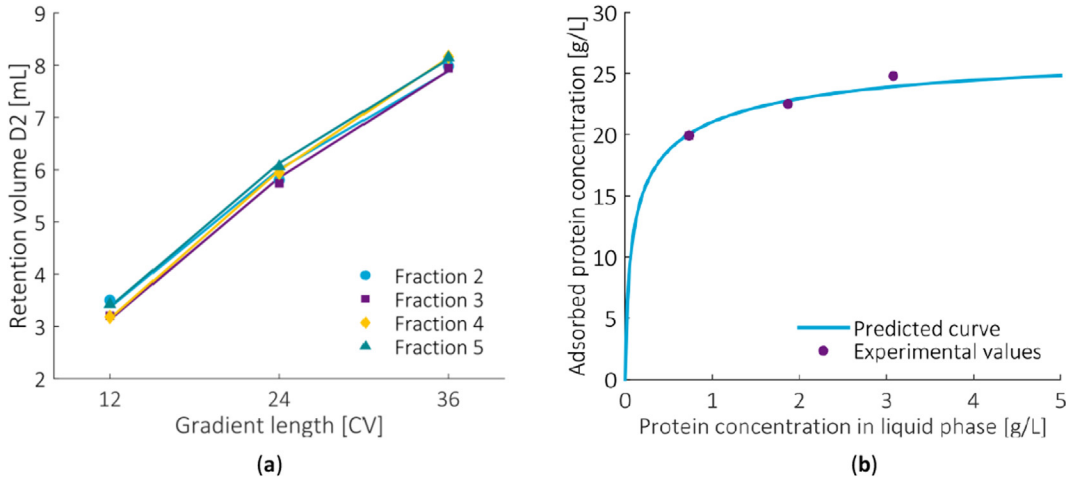

Fig. 7. Results of the linear gradient RoboColumn experiments (a) and the batch uptake experiments (b) for IgG1 on the hydrophobic interaction resin.

3.2.1.1. pH dependence of the stoichiometric ion exchange coefficient. Hydrophobic interactions are almost not influenced by pH changes. However, electrostatic interactions are highly depending on pH. The variance of the stoichiometric ion exchange coefficient, which is in this case identical to the effective binding charge, with pH was assumed to follow this relationship (Mollerup, 2008):

$$\nu_i = g_{1,i} + g_{2,i} \ln pH \tag{28}$$

However, this trend is only valid at pH values that are not too close to the pKa of the ligand (Nfor et al., 2010). Since the pKa of Capto MMC is 3.3 (Pinto et al., 2015) and the pKa of the functional group of POROS 50 HS is 1.2, this is given in the investigated pH range. Additionally, the isoelectric points (pI) for the proteins were used assuming that net charge and binding charge are identical at that point. The pI's for the impurities were estimated with ExPASy, while the pI for IgG1 was determined experimentally. This partly explains the much smaller standard deviation found for the fitting parameters of the monoclonal antibody as shown in Table 6.

Generally protein charge does not follow a strict relationship as given in Eq. (28), as can be seen for instance in published protein net charge data (Lehermayr et al., 2011). Therefore, Eq. (28) and the fitted parameters shown in Table 6 should be seen as estimates. The influence of the stoichiometric coefficient on the equilibrium constant was not taken into account.

3.2.2. Modelling and validation

Validation of the mechanistic chromatography model with experimental data was shown previously for POROS 50 HS and Capto MMC (Pirrung et al., 2018). Therefore, validation experiments were only performed with the hydrophobic interaction resin, Cellufine Phenyl. Unfortunately, IgG1 (ID1 CEX) did not show the expected behaviour in experiments on the lab scale column packed with this resin. It did not elute with ammonium sulphate as an eluent even though the elution was predicted by a simulation, which is depicted in Fig. 8a. A potential reason for this might be varying ligand density caused by a different resin lot used in the lab scale columns than in the RoboColumns. This difference in ligand density might be up to 20% (Susanto et al., 2008), which could explain this observation in part. As investigated by Deitcher et al. (2010), a change in the ligand density of 7.5% could already lead up to a change in the retention factor of 50–100%. Therefore, simulations were performed at varying ligand densities, which is shown for IgG1 in Fig. 8b. As can be seen, even a 20% difference in ligand density would lead to the elution of IgG1 with the salt buffer at the end of the gradient.

In our experimental approach, the ligand density was lumped together with another parameter. Since it appears to be an important parameter, however, this would not be recommended for future work. An experimental determination of the ligand density for each used resin batch is highly recommendable (Huuk et al., 2016).

Generally, the purification capability of this resin is quite poor for IgG1, since IgG1 as well as critical impurities elute in the very end or even after the ammonium sulphate gradient. To avoid this problem, a less hydrophobic resin such as for instance one with a butyl group as functional group should be investigated instead. Another option to improve this step would be the use of a different salt with a smaller salting-out capacity such as sodium chloride. Nevertheless, Cellufine Phenyl was used during the process optimisation later; it can still show how the approach works.

Table 6
pH dependence of the stoichiometric coefficient for ion exchange.

Protein ID AEX	pI	g_1	g_2
1	7.1[b]	12.4[c]	−6.3[c]
2	8.6[a]	32.6±0.2	−15.1±0.11
3	5.6[b]	58.3[c]	−33.8[c]
4	8.7[b]	17.5±8.6	−7.6±4.6
6	7.6[b]	26.9[c]	−13.2[c]
7	4.7[b]	89.0[c]	−57.5[c]
8	7.5[b]	66.4[c]	−32.9[c]

[a] Experimentally determined with capillary isoelectric focusing.
[b] Estimated with ExPASy.
[c] Not enough data points available to calculate the standard deviation.

3.2.3. Artificial neural networks

Since more variables were included compared with a previous study (Pirrung et al., 2017), the number of hidden layers, the amount of neurons in them and the needed starting points were reevaluated. It was found that two hidden layers increased the predictive quality of the ANNs considerably. The amount of neurons were varied from 10 to 20 in each layer during training; the networks with the best R^2 were used further. For the cation exchange and hydrophobic interaction resins, 1000 sample points were sufficient, while for the mixed mode resin the sample points had to be increased to 2000 to obtain ANNs with an R^2 above 0.8 for yield.

3.3. Optimization of process sequences

Finally, an overall optimization of all process sequences was performed. The scheme to generate process sequences, which

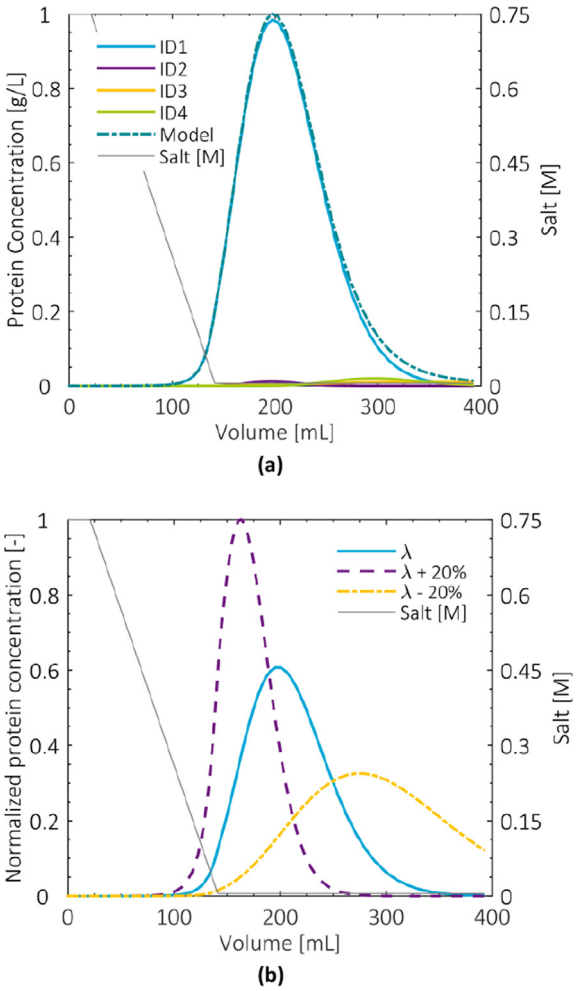

Fig. 8. Modelling of a chromatographic separation in a column filled with the hydrophobic interaction resin Cellufine Phenyl. a: The behaviour of all proteins of interest was simulated. The ID's are from the cation exchange prefractionation; b: The behaviour of IgG1 was simulated at varying ligand densities.

was previously described in Eq. (2), employs y variables that tell which unit is in a sequence and at which position. Having the duration of the UF/DF process as one of the x variables, greatly simplifies this scheme. With that, the filtration units can be included as fixed units before each chromatography unit without having y variables attached to it, because the optimizer can simply change the duration of the UF/DF to zero if a filtration unit was not needed in that position. Thus, only the chromatographic units needed to be taken into account for generating all process alternatives. The amount of process sequence variants was therefore reduced from 30 to 9 while still taking all possibilities into account:

$$\max \; f(y_{m,s}) \tag{29a}$$

$$\begin{aligned} \text{s.t.} \quad & \textstyle\sum_m y_{m,s} \le 1 \\ & \textstyle\sum_s y_{m,s} \le 1 \\ & 1 - \textstyle\sum_m y_{m,2} + \textstyle\sum_m y_{m,1} \ge 1 \\ & y_{m,s} \epsilon \{0,1\}; m \epsilon \{1,2,3\}; s \epsilon \{1,2\} \end{aligned} \tag{29b}$$

The optimization of all resulting process sequences took about 9.5 h with the local optimization taking around 80% of the total time.

Table 7 summarizes the best results for each process sequence; the nonlinear constraint of at least 99.9% purity was fulfilled for all shown options. The solvent use includes the solvent used in filtration and chromatography units alike. As can be seen, all sequences including the hydrophobic interaction resin show a much higher solvent use, which is due to the very late elution of the product. Thus, the resulting peaks are very broad and the elution is more time and solvent consuming. Moreover, this complicates the separation of product and impurities resulting in a trade-off between solvent use and yield. It was the only resin that would not be expected to purify the solution when applied as single chromatography column in the process sequence.

The best process option found, which consists of a UF/DF step and a subsequent cation exchange column, is shown in Fig. 9. During the first step, the bigger molecules such as the product (ID2 AEX) and the biggest impurity (ID7 AEX) are being concentrated.

Smaller molecules, which are able to pass the membrane such as the salt and the smallest impurity (ID4 AEX), are being diluted. Additionally, this step adjusts the pH for the following chromatographic step. The cation exchange chromatography is able to separate the product from the impurities reaching the wanted purity while still having a generally high yield.

To test the robustness of this process, a simulation was performed with four times spiked impurity levels. Under these conditions, a purity of only 99.6% could be reached, which is below the defined constraint of 99.9%. This could be counteracted by moving the product pool cut points accordingly leading to a great loss of yield. Another option would be to add more processing units such as suggested in the process sequence CEX-MMC. This would lead to a more robust process that could easily remove higher concentrations of impurities.

Even though some of the influences of pH were already included in this study, it would be very interesting to see them implemented in more detail. Chromatography units could for instance use pH or dual gradients (a combination of salt and pH) (Lee et al., 2015) as eluents, which could definitely change the result of the optimization.

4. Conclusion

This paper presented an improved approach for the development and optimization of process sequences to purify a monoclonal antibody from clarified cell harvest. The main improvement is the addition of ultra/diafiltration units. To add new unit operations into the optimization approach, the following steps needed to be carried out: First, mechanistic models describing the considered unit operations were developed. Next, the

Table 7
Best results for each possible process sequence fulfilling the constraint of 99.9% purity.

Sequence	CEX	HIC	MMC	CEX - HIC	CEX - MMC	HIC - CEX	HIC - MMC	MMC - CEX	MMC-HIC
Filtration steps	**1**	–	1	2	2	2	2	2	2
Solvent use [mL]	**63.6**	–	106.9	297.2	133.2	259.3	257.7	196.0	260.3
Yield [–]	**96.8**	–	95.7	94.7	95.5	84.7	89.8	95.8	75.1

The best process option is highlighted in bold. UF/DF units were added by the optimizer before each chromatography unit in every shown process sequence; CEX: cation exchange chromatography; HIC: hydrophobic interaction chromatography; MMC: mixed-mode chromatography.

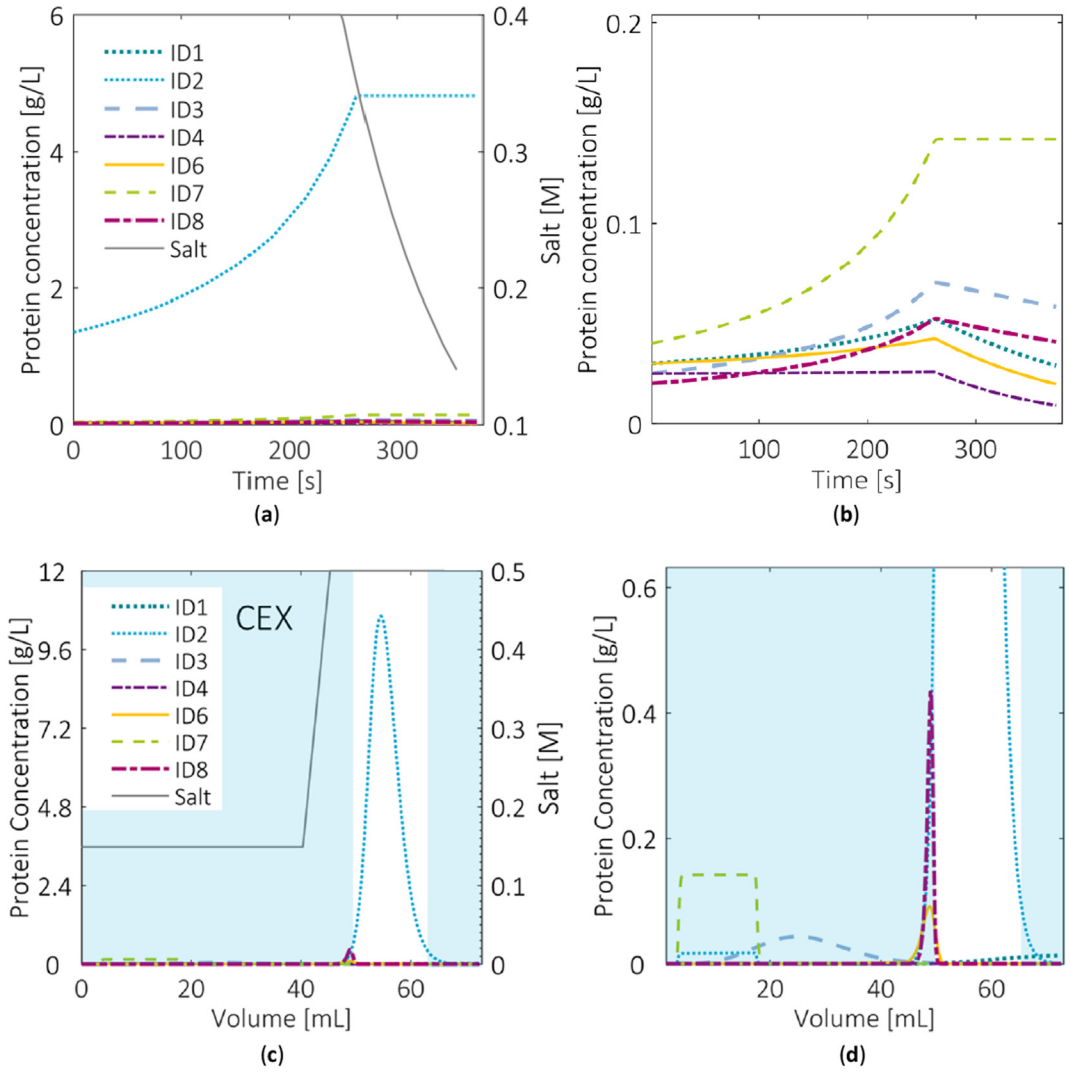

Fig. 9. Best process option. a and b: First step is a UF/DF unit to adjust salt level and pH shown in full (a) and as zoom-in (b). 50 mL of clarified cell harvest was used. The TMP was set to 123 kPa. c and d: Second and final step is a cation exchange column at pH 5.6 with the full chromatogram (c) and the zoom-in (d). The white area is the product pool. The ID's are the ones according to the anion exchange prefractionation.

needed model input parameters were determined. After that, optimizations of only the respective unit were performed to decide, which variables need to be included in further optimizations. For the ultra/diafiltration unit, it was found that only two of the investigated variables are crucial: The duration of the total filtration step and the duration of the ultrafiltration step.

Then, the simulation time was evaluated to see, if it is necessary to train fast artificial neural networks, which can be used instead of the more detailed mechanistic models during optimization. Since the simulation time of the UF/DF only takes in between 0.01 and 0.05 s, this was deemed unnecessary. For unit operation models that take a longer time to simulate such as the chromatography units, however, this step needed to be performed. Finally, the unit operations were included into the process sequences for an overall optimization.

By adding new unit operations and objectives, this paper showed the flexibility of the optimization approach. The approach could easily be extended to include more unit operations. Also the objectives and constraints can be changed as long as they can be simulated by the used mechanistic models. Additionally, the number of variables used as inputs for ANNs was doubled, which showed that more complex problems can be tackled with only few adjustments. By looking at other recent publications, a further surge of new applications including ANNs or other surrogate models for chromatography is expected due to their ease of use and great applicability.

Furthermore, it should be kept in mind that only high molecular weight contaminants were considered as impurities in this study. The removal of other impurities such as viruses, DNA or endotoxins was not yet investigated with this method. However, this could be a very interesting addition in the future making the results more realistic.

In conclusion, the presented approach can greatly influence and improve the way purification processes are being developed. It is especially useful in early process development stages when little is known about the sample material and availability is limited.

Acknowledgment

The authors want to thank Eleni Tsintavi and Carme Pons Royo for performing some of the experiments. Additionally, we would like to thank Song Yi for his help in the laboratory. This work was financially supported under grant F2.003 by the Ministry of Economic Affairs of the Netherlands and BE-Basic partner organizations

(www.be-basic.org) through BE-Basic, a public private NWO-ACTS program.

Declaration of interest

The authors declare no conflict of interest.

References

Baker, R.W., 2004. Ultrafiltration. Membrane Technology and Applications. John Wiley & Sons, Ltd, Chichester, UK, pp. 237–274.

Binabaji, E., 2015. Ultrafiltration of highly concentrated monoclonal antibody solutions [Doctoral Dissertation]. The Pennsylvania State University, Ann Arbor.

JNC-Corporation. Hydrophobic Interaction Chromatography Media Cellufine® Butyl Cellufine® Phenyl. http://www.jnc-corp.co.jp/fine/en/cellufine/guide/pdf/hydrophobic/TD_HIC_DF013015_V3_1_EN_20140711.pdf. Accessed 19.12.2017.

Da Costa, A.R., Fane, A.G., Wiley, D.E., 1994. Spacer characterization and pressure drop modelling in spacer-filled channels for ultrafiltration. J. Membr. Sci. 87 (1–2), 79–98.

Deitcher, R.W., Rome, J.E., Gildea, P.A., O'Connell, J.P., Fernandez, E.J., 2010. A new thermodynamic model describes the effects of ligand density and type, salt concentration and protein species in hydrophobic interaction chromatography. J Chromatogr A. 1217 (2), 199–208.

Fikar, M., Kovacs, Z., Czermak, P., 2010. Dynamic optimization of batch diafiltration processes. J. Membr. Sci. 355 (1–2), 168–174.

Fikar, M., 2014. Modelling, control, and optimisation of membrane processes. In: Paper Presented at: Control Conference (ICCC), 2014 15th International Carpathian; 28-30 May 2014.

Foley, G., 1999. Minimisation of process time in ultrafiltration and continuous diafiltration: the effect of incomplete macrosolute rejection. J. Membr. Sci. 163 (2), 349–355.

Foley, G., 2013. Membrane Filtration: A Problem Solving Approach with MATLAB®. Cambridge University Press, Cambridge, UK.

Forrester, A., Sobester, A., Keane, A., 2008. Engineering Design Via Surrogate Modelling: A Practical Guide. John Wiley & Sons, Chichester, UK.

Gronemeyer, P., Ditz, R., Strube, J., 2014. Trends in upstream and downstream process development for antibody manufacturing. Bioengineering 1 (4), 188–212.

Guiochon, G., Felinger, A., Shirazi, D.G., Katti, A.M., 2006. Fundamentals of Preparative and Nonlinear Chromatography. Elsevier Inc..

Hanke, A.T., Tsintavi, E., Ramirez Vazquez, M.D., et al., 2016. 3D-liquid chromatography as a complex mixture characterization tool for knowledge-based downstream process development. Biotechnol. Prog. 32 (5), 1283–1291.

Huuk, T.C., Hahn, T., Osberghaus, A., Hubbuch, J., 2014. Model-based integrated optimization and evaluation of a multi-step ion exchange chromatography. Sep. Purif. Technol. 136, 207–222.

Huuk, T.C., Briskot, T., Hahn, T., Hubbuch, J., 2016. A versatile noninvasive method for adsorber quantification in batch and column chromatography based on the ionic capacity. Biotechnol. Prog. 32 (3), 666–677.

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

Lehermayr, C., Mahler, H.C., Mader, K., Fischer, S., 2011. Assessment of net charge and protein-protein interactions of different monoclonal antibodies. J. Pharm. Sci. 100 (7), 2551–2562.

Lutz, H., 2015. Ultrafiltration for Bioprocessing. Woodhead Publishing, Oxford, UK.

Lutz, H., Arias, J., Zou, Y., 2017. High concentration biotherapeutic formulation and ultrafiltration: Part 1 pressure limits. Biotechnol. Prog. 33 (1), 113–124.

Meyer, K., Huusom, J.K., Abildskov, J., 2018. High-order approximation of chromatographic models using a nodal discontinuous Galerkin approach. Comput. Chem. Eng. 109 (Supplement C), 68–76.

Mollerup, J.M., 2008. A review of the thermodynamics of protein association to ligands, protein adsorption, and adsorption isotherms. Chem. Eng. Technol. 31 (6), 864–874.

Mollerup, J.M., Hansen, T.B., Kidal, S., Staby, A., 2008. Quality by design–thermodynamic modelling of chromatographic separation of proteins. J. Chromatogr. A 1177 (2), 200–206.

Nfor, B.K., Noverraz, M., Chilamkurthi, S., Verhaert, P.D., van der Wielen, L.A., Ottens, M., 2010. High-throughput isotherm determination and thermodynamic modeling of protein adsorption on mixed mode adsorbents. J. Chromatogr. A 1217 (44), 6829–6850.

Nfor, B.K., Ahamed, T., van Dedem, G.W.K., et al., 2013. Model-based rational methodology for protein purification process synthesis. Chem. Eng. Sci. 89, 185–195.

Paulen, R., Fikar, M., Kovacs, Z., Czermak, P., 2011. Process optimization of diafiltration with time-dependent water adding for albumin production. Chem. Eng. Process. 50 (8), 815–821.

Paulen, R., Foley, G., Fikar, M., Kovacs, Z., Czermak, P., 2011. Minimizing the process time for ultrafiltration/diafiltration under gel polarization conditions. J. Membr. Sci. 380 (1–2), 148–154.

Paulen, R., Fikar, M., Foley, G., Kovacs, Z., Czermak, P., 2012. Optimal feeding strategy of diafiltration buffer in batch membrane processes. J. Membr. Sci. 411, 160–172.

Pinto, I.F., Aires-Barros, M.R., Azevedo, A.M., 2015. Multimodal chromatography: debottlenecking the downstream processing of monoclonal antibodies. Pharm. Bioprocess. 3 (3), 263–279.

Pirrung, S.M., van der Wielen, L.A.M., van Beckhoven, R., van de Sandt, E., Eppink, M. H.M., Ottens, M., 2017. Optimization of biopharmaceutical downstream processes supported by mechanistic models and artificial neural networks. Biotechnol. Prog. 33 (3), 696–707.

Pirrung, S.M., Parruca da Cruz, D., Hanke, A.T., et al., 2018. Chromatographic parameter determination for complex biological feedstocks. Biotechnol. Prog. 34 (4), 1006–1018.

Sellberg, A., Ojala, F., Nilsson, B., 2015. Model-based comparison of antibody dimerization in continuous and batch-wise downstream processing. Antibodies 4 (3), 157–169.

Susanto, A., Knieps-Grünhagen, E., von Lieres, E., Hubbuch, J., 2008. High throughput screening for the design and optimization of chromatographic processes: assessment of model parameter determination from high throughput compatible data. Chem. Eng. Technol. 31 (12), 1846–1855.

Tsintavi, E., 2015. Multidimensional Crude Feedstock Profiling. TU Delft report.

van Reis, R., Goodrich, E.M., Yson, C.L., Frautschy, L.N., Whiteley, R., Zydney, A.L., 1997. Constant Cwall ultrafiltration process control. J. Membr. Sci. 130 (1), 123–140.

Wang, Y., Li, X., Liu, Y.-H., et al., 2016. Simultaneous monitoring of oxidation, deamidation, isomerization, and glycosylation of monoclonal antibodies by liquid chromatography-mass spectrometry method with ultrafast tryptic digestion. Paper presented at: mAbs2016.

Wang, G., Hahn, T., Hubbuch, J., 2016. Water on hydrophobic surfaces: Mechanistic modeling of hydrophobic interaction chromatography. J. Chromatogr. A 1465, 71–78.

Young, M., Carroad, P., Bell, R., 1980. Estimation of diffusion coefficients of proteins. Biotechnol. Bioeng. 22 (5), 947–955.

Zydney, A.L., 1997. Stagnant film model for concentration polarization in membrane systems. J. Membr. Sci. 130 (1–2), 275–281.