Journal of Chromatography B, 873 (2008) 102–112

Contents lists available at ScienceDirect

Journal of Chromatography B

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

Size-exclusion chromatography as a linear transfer system: Purification of human influenza virus as an example

Bernd Kalbfuss[a,*], Dietrich Flockerzi[b], Andreas Seidel-Morgenstern[c,d], Udo Reichl[a,e]

[a] Bioprocess Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany
[b] Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany
[c] Physical and Chemical Foundations of Process Engineering, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany
[d] Chair of Chemical Engineering, Otto-von-Guericke University, Magdeburg, Germany
[e] Chair of Bioprocess Engineering, Otto-von-Guericke University, Magdeburg, Germany

ARTICLE INFO

Article history:
Received 1 April 2008
Accepted 7 August 2008
Available online 12 August 2008

Keywords:
Size-exclusion chromatography
Group separation
Influenza virus
Linear time-invariant system
Maximum entropy deconvolution

ABSTRACT

Preparative size-exclusion chromatography suffers from low selectivity and productivity. Empirical optimization of operating conditions constitutes a laborious task due to many parameters. Here, a modeling framework based on linear systems theory is presented for predicting the influence of volume overloading. Impulse-responses characterizing system behavior are derived from experimental data by maximum entropy deconvolution. Theoretical derivations are validated experimentally by study of a model system and chromatography of human influenza virus. By application of the theory it is demonstrated how group separation operations can be optimized with respect to yield, purity, productivity and dilution of the product.

1. Introduction

Size-exclusion chromatography (SEC; also known as gel filtration) is a versatile technique for the separation of solutes by size. A first report is due to Porath and Flodin [1] who used cross-linked dextran for the separation of various solutes in aqueous solution. Nowadays, it plays a vital role in many biotechnological processes. Preparative SEC is commonly used for the separation of proteins from aggregates [2,3], for the desalting of proteins [4,5] and the purification of bio-particles like plasmids [6–8] or virions [9–14]. It is noteworthy that in these applications SEC does not have to be operated as a batch process but continuous operation in the form of simulated moving-bed chromatography may be used in order to overcome low productivity [15].

The separation principle of SEC is based on a molecular sieving mechanism upon diffusion of solutes into a porous stationary phase. The main advantages of SEC are low media cost (no expensive ligands) and ease of operation (isocratic elution). If operated at modest solute concentrations it further belongs to the realm of linear chromatography and is thus simple to describe and scale. The main disadvantages are inherently low productivity and selectivity. In particular low productivity usually has to be compensated by volume overloading which, however, reduces resolution. Another disadvantage is that the product is diluted and cannot be concentrated as in gradient chromatography.

First descriptions of SEC are based on the plate model of chromatography [16,17]. In these models a single source of zone spreading is assumed. Their primary purpose is the prediction of mean retention time and zone spreading for single components injected as very short pulses. Yet, they already provide a basis for predicting the effect of volume overloading [17]. A more sophisticated approach is the use of a general rate model, which includes descriptions of mass transfer kinetics (film diffusion and intraparticle diffusion) [18,19]. Validity of the suggested model was verified for a number of proteins. Contrary to plate models, it is capable of predicting zone spreading and asymmetry of elution peaks in relation to mass transfer phenomena.

A different quantitative view on SEC was provided by the stochastic modeling approach [20]. Here, SEC is dismantled into the phenomena of pore ingress and egress, which are influenced by the pore size distribution of the stationary phase. In addition, moving zone dispersion is considered as in previous models. Stochastic models are solved by means of the characteristic function from which peak moments can be immediately derived. In addition, peak shapes can be obtained by numerical inversion of characteristic functions. Validity of the stochastic model was verified using an

* Corresponding author at: Max-Planck-Institut, Sandtorstr. 1, 39106 Magdeburg, Germany. Tel.: +49 391 6110 228; fax: +49 391 6110 567.
E-mail address: kalbfuss@mpi-magdeburg.mpg.de (B. Kalbfuss).

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.08.002

extensive set of polystyrene markers [21]. Moreover, applicability to peak shape analysis has been demonstrated [22].

In this article it is shown that preparative SEC can be treated as a linear time-invariant transfer (LTI) system. For description of two- or three-component separations, this approach is clearly less powerful than the ones previously mentioned. In particular, it does not cover the effects of flow rate and column length. However, it is sufficient for describing the effect of volume overloading, which is essential for maximizing productivity.

Applicability of the theory is first demonstrated by studies with model colloids. In these studies, linearity of SEC and superposition of system response was confirmed by injection of single synthetic markers and their mixtures. Using two markers (product and impurity), a fictitious preparative SEC operation was optimized with respect to volume load subject to constraints for the yield of product and final level of impurity. Predicted operating points were validated experimentally.

Separation of influenza virus from host-cell protein was considered as an optimization problem of practical relevance. Impulse-responses for the elution of virions and total protein (viral, host-cell and protein from cell culture medium components) were determined by maximum entropy deconvolution. Global estimates were obtained from a set of three chromatograms acquired at constant flow velocity. Simulations produced by convolution of these estimates were compared to original data. Dependence of system response on flow velocity was investigated in a second set of experiments. Impulse-responses for host-cell protein were reconstructed from chromatograms of virus and total protein. Maximum purity was predicted in the limiting case of an infinitely short injection. Purity, product concentration and productivity were further predicted as a function of volume load and flow velocity. Finally, volume load was maximized under the constraint of 10% or 20% residual content of host-cell protein. Reliability and general value of the suggested method in process development are discussed. In addition, some recommendations for the precise determination of impulse-responses are given.

2. Theory

2.1. Size-exclusion chromatography as an LTI system

In an abstract manner, SEC can be regarded as an LTI transfer system. In this approach, SEC is reduced to a single-input–single-output system with respect to each solute or class of solutes. The output signal $h(t)$ (solute concentration over time at the column outlet) is considered the response $\mathbf{T}\{g\}(t)$ to perturbation by a forcing function $g(t)$ (feed concentration over time at the column inlet). For SEC to be a true LTI system it needs to obey the following equalities [23]:

$$h(t - t_0) = \mathbf{T}\{g\}(t - t_0) \tag{1}$$

$$\sum_{1}^{N} b_i h_i(t) = \sum_{1}^{N} \mathbf{T}\{b_i g_i\}(t) \tag{2}$$

Transformation of signals due to chromatography is hereby presented in the form of an abstract operator $\mathbf{T}$. The property of time-invariance is conveyed in (1), the property of linearity in (2). Time-invariance of SEC is an obvious property as chromatograms do not depend on the time of injection. They will be the same for runs under identical conditions as long as the column has been well equilibrated (i.e. the system is at rest). Linearity, in contrast, remains to be confirmed. It can be argued, however, that the principle of superposition will hold as long as inter-particle interactions remain negligible.

Response of LTI systems to perturbation by admissible forcing functions g(t) can be predicted once the response to an infinitely short pulse has been determined. This response is called the system impulse-response and its Laplace transform is referred to as the system transfer function [23]. In general, the response $h(t)$ of an LTI system at rest can be obtained by convolution of the system impulse-response $h_\delta(t)$ with the forcing function $g(t)$:

$$\begin{aligned} h(t) = \mathbf{T}\{g\}(t) &= \int_{-\infty}^{\infty} h_\delta(t-\tau) g(\tau)\,\mathrm{d}\tau = h_\delta * g \\ &= \int_{-\infty}^{\infty} g(t-\tau) h_\delta(\tau)\,\mathrm{d}\tau = g * h_\delta = \mathbf{T}'\{h_\delta\}(t) \end{aligned} \tag{3}$$

In the particular case of SEC, this translates to a convolution of the injection profile $g(t)$ with a characteristic function $h_\delta(t)$ of the system yet to be described.

If sample is, however, injected in overlapping cycles, as is done to enhance throughput, the assumption of a system at rest (i.e. a well-equilibrated column) is no longer valid. In preparative chromatography, it is true that $g(t) \neq 0$ only within $t_1 \le t \le T_1$ corresponding to a finite injection. If the impulse-response is assumed to be finite as well with $h_\delta(t) \neq 0$ only if $t_2 \le t \le T_2$, it is true that $h(t) \neq 0$ only within $t_1 + t_2 \le t \le T_1 + T_2$ [23]. In other words, if elution occurs within different intervals of time (referred to as windows of elution in this article), the response derived remains valid within the interval $t_1 + t_2 \le t \le T_1 + T_2$, which is ensured by the injection-elution strategy described in Fig. 8. Since we are dealing with a causal system it is further true that $t_2 \ge t_1$ and $T_2 \ge T_1$. It is noteworthy that from a mathematical point of view, $h_\delta(t)$ generally approaches zero only in the limit of $t \to \infty$. Practically, however, the assumption of a limited impulse-response is reasonable since impulse-responses in chromatography approach zero sufficiently rapidly (see Fig. 2)

2.2. Maximum entropy deconvolution

Several methods exist to determine the system impulse-response of real columns. In the simplest case, it is approximated by the response to a finite short injection. Analytical problems may, however, arise due to the resulting very low outlet concentrations. An intuitive alternative would be to acquire the response to a finite injection of sufficient extent with a known temporal concentration profile. The impulse-response could then be determined by inversion of the convolution integral in (3). This type of problem is referred to as a deconvolution problem and is related to solving a Fredholm integral equation of the first kind. Unfortunately, the latter is known to be an ill-conditioned problem in the presence of experimental noise [24,25]

In order to find meaningful solutions, regularization methods have to be applied. Variational regularization methods can be formulated as solving an optimization problem which is defined by a regularization functional $\Omega(h_\delta)$ and a data discrepancy functional $D(h_\delta, h)$ [26]. A common choice for the discrepancy functional is $D := ||\mathbf{T}' h_\delta - h||^2$ with $||\cdot||^2$ denoting the L^2 Hilbert norm. Note, that the operator $\mathbf{T}'$ in this expression is identical to the integral operator defined in (3) which uses the forcing function $g(t)$ and not the system impulse-response as kernel. Applying the penalization-based regularization approach, a reconstruction operator $R_\lambda^{\mathrm{pen}}(h)$ can be defined according to

$$R_\lambda^{\mathrm{pen}}(h) := \underset{h_\delta}{\mathrm{argmin}}\left[\left\|\mathbf{T}' h_\delta - h\right\|^2 + \lambda \Omega(h_\delta)\right] \tag{4}$$

In this approach, the regularization functional $\Omega(h_\delta)$ has the form of a penalty function containing information on the desirability of a solution. It is weighted by a regularization parameter λ.

One possible choice for the penalty function is

$$\Omega(h_\delta) = \int_0^\infty h_\delta(t) \ln h_\delta(t)\,dt \quad \text{with} \quad h_\delta \geq 0 \quad \text{and} \quad \int_0^\infty h_\delta(t)\,dt = 1 \tag{5}$$

leading to the concept of maximum entropy (MaxEnt) regularization [27]. Normalization and positivity of h_δ is thereby assumed such that it can be interpreted as a probability density function (in our case a residence time distribution). According to Rullgård et al. [26] the optimization problem in (4) can be restated as a tolerance-based reconstruction operator:

$$R_\varepsilon^{\text{tol}}(h) := \underset{\|T'h_\delta - h\| \leq \varepsilon}{\text{argmin}} \ \Omega(h_\delta) \tag{6}$$

if both optimization problems are strictly convex. In this case, there exists a correspondence between the regularization parameters λ and ε. The correspondence, however, depends on the data and is not straightforward to delineate. In the tolerance-based approach, the regularization parameter ε defines an upper bound for the data discrepancy functional D. It is equivalent to finding a solution for h_δ, which maximizes information entropy defined as $S = -\Omega$ while some deviation of the prediction from observed system behavior is allowed. Since Ω and D are both strictly convex, a single global maximum exists making the solution for h_δ unique [28].

While various alternatives for the definition of Ω exist, the principle of maximum entropy has been proven axiomatically to be a uniquely correct method for inductive inference [27]. It is the only method consistent with simple and general requirements that have been described elsewhere [27,29]. MaxEnt deconvolution has been applied successfully to a number of problems including the calculation of drug absorption rates [30], reconstruction of images [28,31], crystal structures [29], structures from EM tomography [26] and deconvolution of mass spectrometry data [24]. A systematic comparison with other deconvolution methods has been given by Madden et al. [32].

The main difficulty in applying (6) remains in choosing appropriate values for ε. In statistical regularization theory, it is assumed that D is a random function with approximately known probability distribution. If, for instance, the distribution of error is homoscedastic and variance is known a priori, χ^2-statistics can be used for deriving values of ε [28]. Regarding chromatography, however, homoscedastic distribution of error is unlikely. Experimental data is typically prone to error from baseline drift, non-linearity of detectors and deviation from ideal behavior due to particle–particle interactions. In addition, error due to discretization has to be taken into account. Sound derivation of ε is hence difficult and impractible in most cases. Consequently, an empirical approach was pursued in this article. Firstly, the minimum distance $\varepsilon_{\text{NNLS}}$ of the non-negatively constrained least-squares solution (NNLS; special case for $\lambda = 0$) was calculated. Secondly, ε was set from $2\varepsilon_{\text{NNLS}}$ to $6\varepsilon_{\text{NNLS}}$ in the following such that a meaningful solution without oscillations was obtained.

3. Experimental

3.1. Synthetic markers

Chromatography of synthetic markers was conducted on an Äkta Basic 100 chromatography system. Samples were injected using a 10-mL superloop (cat. no. 19-7585-01). Sepharose 4 FF packed into a Tricorn 10/300 housing (cat. no. 28-4064-18; all GE Healthcare Biosciences, Uppsala, Sweden) was used as column (30 cm bed height, 1 cm diameter, 23.6 mL total volume of packing). Quality of the packing was checked regularly by injection of 1% acetone ($\leq$0.02 column volume; cv) eluted at 60 cm h^{-1}. Results were reproducible without any indication of decay. At least 1200 plates and peak asymmetries close to 1 were determined. Phosphate-buffered saline was used as eluent (8 g L^{-1} NaCl, 0.2 g L^{-1} KCl, 0.20 g L^{-1} KH_2PO_4, 1.15 g L^{-1} Na_2HPO_4, pH $\sim$ 7.3). Elution of markers was monitored online at a wavelength of 220 nm.

Polymethacrylic acid (PMA) markers used in experiments were purchased from Polymer Standard Service (Mainz, Germany). 100-nm BSA-coated and 200-nm dextran-coated polystyrene latex beads were purchased from Postnova analytics (Landsberg a. Lech, Germany). Note that reported concentrations of PMA markers are only approximate since PMA supplied as powder did not dissolve completely.

0.5 mL (0.02 cv) of a 201-kDa PMA marker was injected at five concentrations (0.25, 0.5, 1, 2 and 4 mg mL^{-1}) in a first experiment. PMA markers of 1.22 and 201 kDa molecular weight (1 and 2 mg mL^{-1}, respectively) and 200-nm dextran-coated latex beads (5 μg mL^{-1}) were used as a mixture and as single components in a second experiment. Chromatograms were plotted over K_{av} = (elution volume – void volume)/(column volume – void volume) and normalized elution volume V_{e}. All runs were conducted at a superficial mobile phase velocity of 60 cm h^{-1}.

Chromatographic separation was optimized in simulation studies for a mixture of 100-nm BSA-coated latex beads (20 μg mL^{-1}) considered the product and a 201-kDa PMA marker (2 mg mL^{-1}) considered the impurity. Impulse-responses were approximated by elution profiles obtained after short injection pulses (0.1 mL or 0.005 cv) of single components acquired at seven flow velocities (7.5, 15, 30, 45, 60, 90, 120 cm h^{-1}). Permissible injection volume was then maximized in silico for each flow velocity subject to constraints for the product yield and final content of impurity (95% and 5% in this article). Chromatograms were simulated for an initial injection volume (guess value) by convolution of approximated impulse-responses with rectangular injection pulses. Next, windows of elution (overall window for product and impurity) were determined from cumulative distributions applying pre-defined threshold quantiles (10^{-3} and 0.999 in this article). Onset of elution was used to define the lower limit of the product fraction. The upper limit was set to meet the constraint for the product yield. Finally, the difference between actual content of impurity and target value was calculated and minimized iteratively by adaptation of injection volumes. Next to calculations in silico, chromatography runs with single compounds were conducted applying the predicted maximum injection volumes for experimental validation. Volumetric productivity calculated for maximized injection volumes was expressed as the volume of sample that can be processed per unit time normalized by column volume. The length of chromatography runs was set equal to the length of elution windows plus 0.2 cv safety margin (overlapping injection cycles).

Simulations were conducted in Matlab (The Mathworks, Aachen, Germany). Script files including chromatograms of synthetic markers can be downloaded from http://www.mpi-magdeburg.mpg.de/research/groups/bpt/sec-modeling.

3.2. Influenza virus

Experimental set A, comprising three chromatography runs at a constant superficial mobile phase velocity of 60 cm h^{-1} but different injection volumes (0.07, 0.14 and 0.28 cv), was taken from a previous study [34]. Experimental set B comprising three runs with constant injection volume of 0.05 cv but different flow velocities (30, 60 and 120 cm h^{-1}) was produced as described in the following.

Chromatography was conducted on an Äkta Explorer 100 chromatography system (GE Healthcare Biosciences, Uppsala, Sweden).

Sample was injected using a 1.5-mL capillary loop (partial injection). The same column as for chromatography of synthetic markers and experimental set A was used. HEPES-buffered saline (200 mM NaCl, 20 mM HEPES, pH 7.3) was chosen as eluent. Elution of solutes was monitored online at a wavelength of 280 nm and by a static light scattering detector (Dawn EOS, Wyatt, Dernbach, Germany). Eluates were fractionated (0.6 mL or 0.025 cv per fraction) into deep well microtiter plates (cat. no. 732-0585, VWR, Darmstadt, Germany) using a Frac-950 collector (GE Healthcare Biosciences, Uppsala, Sweden). The fraction collector and UV detector had been aligned before by injection of a phenol red solution (~5 mg mL^{-1}) through a bypass capillary. Collected fractions (0.2 mL per fraction into analytical microtiter plates) were scanned with a plate photometer (Rainbow Spectra, Tecan Instruments, Crailsheim, Germany) and compared to the online signal. The delay volume was chosen to minimize mismatch between profiles. Delay between the UV and light scattering detectors was determined by matching the first peak in the UV trace with the single peak detected by light scattering. Collected fractions were analyzed for virus content (HA activity) and total protein (Bradford assay) as described before [34,35]. However, the ratio of extinction at wavelengths 470 and 495 nm (instead of extinction at 495 nm alone) was used in order to enhance sensitivity of the protein assay.

For the characterization of extra-column contributions to band broadening the column was replaced by a union (while capillaries were maintained). Injection profiles were acquired by injection of 1% acetone at the same volume and flow velocity as in runs with virus sample.

A 20-fold concentrate of cell culture supernatant was used for chromatography. Influenza virus A/PR/8/34 (H1N1) was propagated in mammalian cell culture using Madin-Darby Canine Kidney cells as host. Clarification, inactivation and concentration was performed as previously described [35]. The concentrate was stored at −70 °C.

Online signals (UV absorbance and light scattering) were normalized to represent true density distributions. Both signals were re-sampled to a uniform grid by linear interpolation (200 data points per column volume). Offline data (HA activity and total protein) were re-sampled to the same grid by the following procedure: data were first summed up and normalized to represent true cumulative distributions. Cumulative distributions were then interpolated using weighted splines (Matlab Spline Toolbox, The Mathworks, Aachen, Germany) and differentiated to yield density distributions. Distributions of host-cell protein were reconstructed by subtracting re-scaled distributions of HA activity from distributions of total protein. The scaling factor was determined under the assumption of pure virus in early fractions (set A: 0.34–0.42 cv; set B: 0.33–0.38 cv elution volume).

Impulse-responses for all types of signals were estimated by MaxEnt deconvolution of normalized chromatograms using experimentally determined injection profiles (see section on MaxEnt deconvolution above). Simulation of chromatography was conducted similarly to synthetic markers. Limits of the product fraction were determined from cumulative distributions of HA activity based on 0.01 and 0.91 quantiles. Windows of elution were determined from cumulative distributions of the UV signal (10^{-3} and 0.999 quantiles). A safety margin of 0.2 cv was added for calculation of the total run length. The final concentration of product, content of host-cell protein and volumetric productivity were predicted as functions of injection volume and flow velocity. Concerning productivity, a second mode of operation was considered. In this mode, flow velocity was increased to 250 cm h^{-1} after elution of product in order to wash out impurities as fast as possible (see Fig. 8). Deconvolution of chromatograms and simulations were conducted in Matlab (The Mathworks, Aachen, Germany). Script files including chromatograms can be downloaded from http://www.mpi-magdeburg.mpg.de/research/groups/bpt/sec-modeling.

4. Results

4.1. Synthetic markers

Linearity of SEC was confirmed experimentally by chromatography of model colloids. In a first experiment 0.5 mL (0.02 cv) of a 201-kDa PMA marker was injected at concentrations ranging from 0.25 to 4 mg mL^{-1} (not shown). Peak position and shape of eluting peaks were almost identical for all injections (width at half height $\Delta w_{0.5} = 2.55 \pm 0.006$, peak elution volume $V_p = 0.453 \pm 0.014$ cv; mean ± S.D.). However, an increasing shift towards higher elution volumes was observed with increasing concentration. The height of peaks scaled proportionally with marker concentration. Comparison of peak area was not feasible due to impurities in the sample appearing as small peaks before and after the main peak.

In a second experiment 0.075, 0.15 and 0.30 cv of a mixture of 200-nm dextran-coated polystyrene beads (5 μg mL^{-1}) and PMA markers of 1.22 and 201 kDa (1 and 2 mg mL^{-1}, respectively) were injected. For comparison, each colloid was also injected individually. Chromatograms from individual injections with the same injection volume were summed up and compared to chromatograms of the mixture (Fig. 1). Only slight deviation was observed which was strongest at the highest load of 0.30 cv.

Before any attempts were made to predict virus elution, a third experiment with model colloids was conducted to ensure that SEC can indeed be predicted by means of linear systems theory. In this experiment, 100 nm BSA-coated polystyrene beads (20 μg mL^{-1}) were used to mimic influenza virions or totally excluded colloids in general. The 201 kDa PMA marker (2 mg mL^{-1}), which had been used before, was used as model impurity. Impulse-responses for both colloids were approximated from chromatograms obtained after short injections (0.005 cv of sample). Superficial mobile phase velocity was varied between 7.5 and 120 cm h^{-1} (0.25 and 4 cv h^{-1}). Chromatograms of polystyrene beads were almost independent of flow velocity in contrast to chromatograms of PMA marker. The normalized chromatogram of polystyrene beads at 60 cm h^{-1} and chromatograms of PMA marker at four out of seven flow velocities are illustrated in Fig. 2.

Operating conditions of a fictitious preparative SEC operation were found by maximizing injection volume and hence produc-

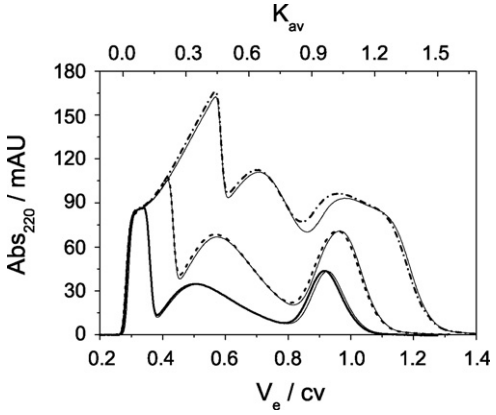

Fig. 1. 200-nm dextran-coated polystyrene beads (5 μg mL^{-1}) and two PMA markers of 1.22 and 201 kDa (1 and 2 mg mL^{-1}, respectively) injected individually and as a mixture. Chromatography was conducted by injection of 0.075 (—), 0.15 (– – –) and 0.30 cv (-·-). Chromatograms after injection of single components (for a certain injection volume) were summed up (thin lines) and compared to chromatograms of the mixture. Concentrations are reported in arbitrary absorbance units (mAU).

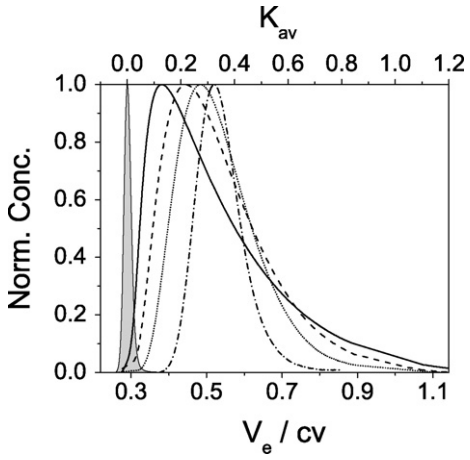

Fig. 2. Chromatograms of model colloids eluted at different superficial mobile phase velocities. Chromatogram of BSA-coated polystyrene beads (20 μg mL^{-1}) at 60 cm h^{-1} (shaded curve) and chromatograms of 201 kDa PMA marker (2 mg mL^{-1}) at 120 (—), 60 (– – –), 30 (···) and 7.5 cm h^{-1} (-·-). Concentrations were normalized to their maximum value.

tivity under the constraint of 95% product yield and 5% residual impurity (Fig. 3). Volume overloading was simulated by convolution of impulse-responses with rectangular injection pulses of desired length. Separate optimization was conducted for all seven flow velocities. Volumetric productivity (here defined as the normalized volume of feed that can be processed per unit time) was calculated from maximized injection volumes and corresponding fractionation limits. Similarly, the final product concentration was calculated under the assumption of full recovery over the whole run.

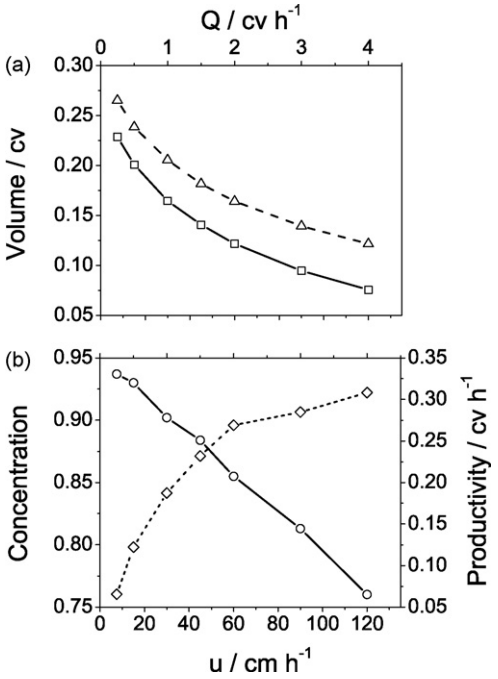

Fig. 3. Optimization of a fictitious preparative SEC operation. 100-nm BSA-coated polystyrene beads were used to mimic influenza virions (product). A 201-kDa PMA marker was used as model impurity. Impulse-responses at seven flow velocities (7.5–120 cm h^{-1}) approximated from chromatograms after short injections (0.5% cv) were used for simulation. Injection volumes of preparative runs were maximized under the constraint of 95% product yield and 5% residual impurity. Maximum injection volumes (□) and corresponding upper fractionation limits (△) were determined iteratively by numerical optimization (a). Lower fractionation limits were set to allow a loss of 0.1% of product. Final normalized product concentration (○) and productivity (◊) were calculated from maximized injection volume and corresponding fractionation limits (b). Volumetric flow rate Q is provided in addition to superficial mobile phase velocity *u*.

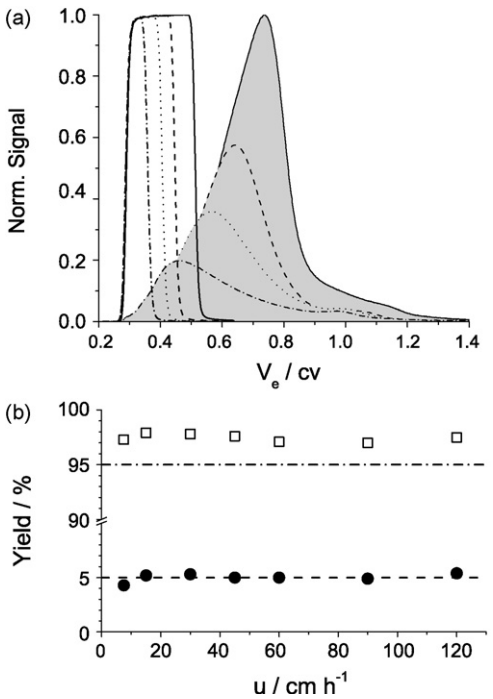

Fig. 4. Experimental validation of predictions for a fictitious SEC operation. 100-nm BSA-coated polystyrene beads (product) and a 201-kDa PMA marker (impurity) were injected at volumes predicted to be optimal for a certain flow velocity. Four out of seven chromatograms at superficial mobile phase velocities of 7.5 cm h^{-1} (-·-), 30 cm h^{-1} (···), 60 cm h^{-1} (- - -) and 120 cm h^{-1} (—) are plotted over normalized elution volume (a). Chromatograms of PMA marker are shaded in gray. Yields were determined by integration of UV signals. An average yield of 97% was determined for the product (95% targeted), while the residual content of impurity was 5% (5% targeted) (b).

Predictions were validated experimentally by applying previously determined (i.e. maximized) injection volumes. Chromatograms for four out of seven flow velocities (7.5–120 cm h^{-1}) are illustrated in Fig. 4a. Each marker was injected separately such that yields could be determined by integration of UV signals. An average product yield (beads) of 97.5 ± 0.3% (mean value ± S.D.) close to the constraint of 95% was obtained (Fig. 4b). Experimentally determined yields for the impurity (PMA marker) were 5.0 ± 0.4% on average thus matching the constraint.

4.2. *Influenza virus*

Separation of influenza virus from host-cell protein (HCP) by preparative SEC was considered as an optimization problem of practical relevance. Chromatography was conducted on the same column as used for separation of synthetic markers. Two lots of 20-fold concentrated virus-containing cell culture supernatant stored at −70 °C were used as sample.

In contrast to synthetic markers, no online signals were available for the specific detection of virions and protein but elution profiles had to be determined by offline analysis. Consequently, impulse-responses could not be approximated from chromatograms after short injections. On the one hand, feed solutions used in experiments were fairly dilute such that the limit of quantitation would have been undercut for total protein. On the other hand, offline

measurements would not have resolved the narrow elution peak of virions sufficiently.

Instead, impulse-responses were determined by maximum entropy (MaxEnt) deconvolution of chromatograms obtained after injection of extended pulses (at least 0.05 cv). Eluates were fractionated at high resolution and analyzed offline for virus content (HA activity) and total protein concentration (Bradford assay). Based on the assumption of pure virus in early fractions an attempt was made to reconstruct the impulse-response for HCP (see materials and methods for details). In addition to offline signals, UV absorbance was acquired as an unspecific but precise online signal in order to validate results of the deconvolution algorithm. All chromatograms (including impulse-responses) were normalized to represent true density distributions of elution volume.

Global impulse-responses for a set of three chromatograms (0.07, 0.014 and 0.028 cv injection volume eluted at 60 cm h^{-1}; set A) were determined for each signal (UV absorbance, HA activity and total protein). Values between $2\varepsilon_{NNLS}$ and $6\varepsilon_{NNLS}$ were set as the limit for misfit in MaxEnt deconvolution. An overlay of all three responses and original data is provided in Fig. 5 (bold lines). For clarity, the response to an injection of 0.014 cv was omitted from the plot of total protein.

UV absorbance showed two distinct peaks near the total exclusion and inclusion limits. Absorbance was low but not zero between these peaks. The impulse-response for UV absorbance was almost identical to the chromatogram obtained after the shortest injection (0.07 cv), but both peaks were narrower. Virions reflected by HA activity eluted in a single tailing peak with maximum around 0.3 cv (total exclusion). Similar to UV absorbance, the impulse-response, was close to the chromatogram obtained after the shortest injection. Elution of total protein, in contrast, was spread over a wide range from 0.3 to 1.2 cv. A definite peak coinciding with the peak of eluting virions indicated that viral protein was detected in these fractions.

Impulse-responses for UV absorbance, HA activity and total protein were used to simulate original data, i.e. chromatograms were predicted for injections of 0.07, 0.14 and 0.28 cv, by convolution with experimentally determined injection profiles. Results of simulations are provided in Fig. 5 (thin lines). A perfect match was obtained for UV absorbance—the signal of highest quality. Slight deviation was observed for HA activity but original data was well resembled. Significant mismatch was, in contrast, observed for total protein but the principle features of original data were still reflected. This mismatch was particularly evident at the highest load.

In a second experiment, the effect of superficial mobile phase velocity was investigated (set B). Due to a limitation in starting material, a different lot of virus concentrate had to be used. Moreover, the eluent employed was of lower ionic strength (only 0.2 M NaCl compared to 0.65 M). The high ionic strength of the eluent used in set A had only been necessary for pre-conditioning to a subsequent anion-exchange chromatography step [34]. Three chromatography runs were conducted by injection and elution of 0.05 cv at different flow velocities (30, 60 and 120 cm h^{-1}). In addition to UV absorbance, scattering of red laser light at an angle of 90° was acquired online. A uniform value of $5\varepsilon_{NNLS}$ was used as the limit for misfit in MaxEnt deconvolution. Impulse-responses at different flow velocities are illustrated in Fig. 6. Results for HA activity were almost indistinguishable. If at all, peak tailing was less pronounced at higher flow velocities. Significant differences were, in contrast, observed for total protein and hence HCP. In the case of total protein, a second peak at about 0.8 cv could be observed at the lowest flow velocity of 30 cm h^{-1}. At 60 cm h^{-1} this peak was less pronounced and shifted towards lower elution volume (about 0.6 cv). At the highest flow velocity of 120 cm h^{-1} the peak had merged almost completely with the peak of virus protein, being merely visible as a shoulder. This trend could be followed even more clearly in reconstructions for HCP. The peak previously located

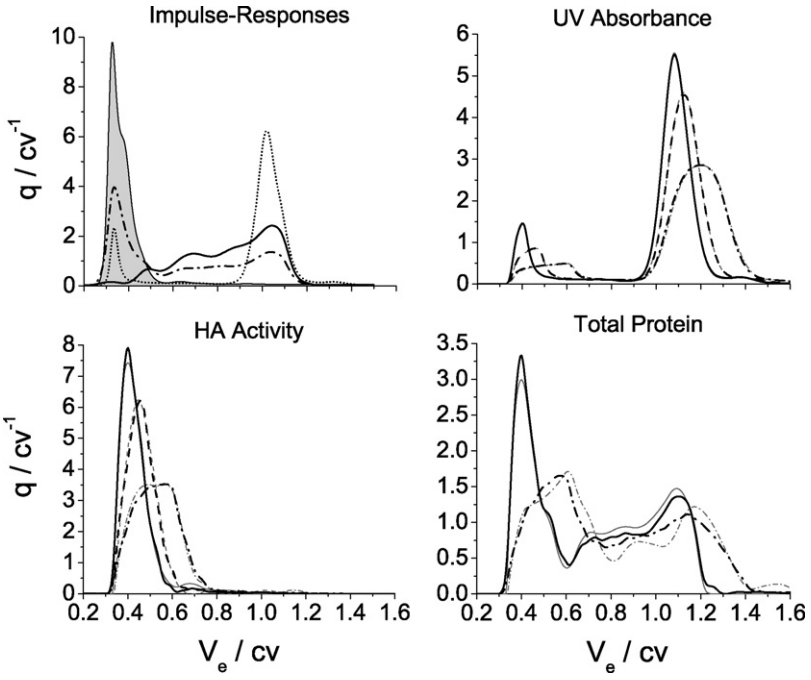

Fig. 5. Impulse-responses for UV absorbance (···), HA activity (shaded curve), total (-·-) and host-cell protein (—) at a constant superficial mobile phase velocity of 60 cm h^{-1} (upper left corner). Impulse-responses were determined from a set of three chromatograms by maximum entropy deconvolution. Distributions of UV absorbance were acquired online, distributions of HA activity and total protein were determined offline. The impulse-response for host-cell protein was reconstructed from distributions of HA activity and total protein under the assumption of pure virus in early fractions. Results were used to simulate original data. Overlays of original data (bold lines) and simulations (thin lines) for injections of 0.07 (—), 0.14 (- - -) and 0.28 cv (-·-) are provided in remaining graphs. For clarity, results for 0.14 cv were omitted from the plot of total protein.

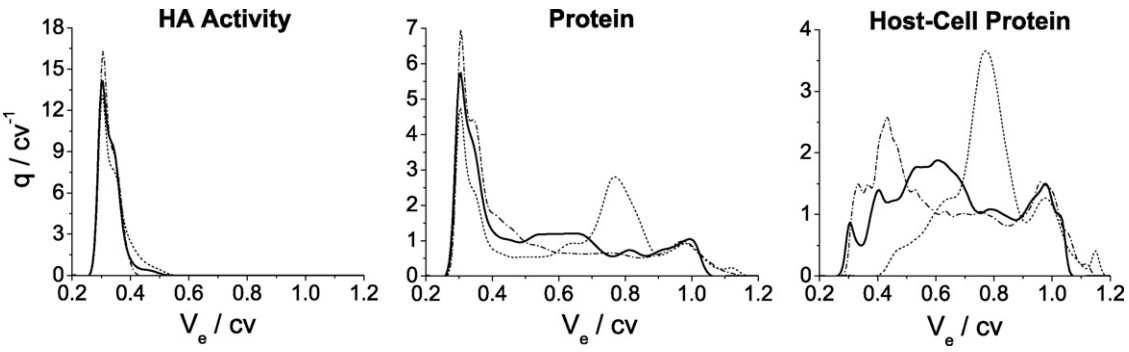

Fig. 6. Impulse-responses for HA activity, total and host-cell protein at different superficial mobile phase velocities. Injections of 0.05 cv were eluted at 30 (···), 60 (—) and 120 cm h^{-1} (- - -). Impulse-responses were determined by maximum entropy deconvolution. The impulse-response for host-cell protein was reconstructed from distributions of HA activity and total protein. Results for HA activity were almost indistinguishable. Impulse-responses for total protein and particularly host-cell protein, in contrast, showed strong dependence on flow velocity.

at 0.8 cv became wider and shifted towards lower elution volume with increasing flow velocity similar to the 201 kDa PMA marker characterized before (Fig. 2).

Experimentally determined impulse-responses (both sets) were used to simulate the outcome of a preparative SEC operation for the separation of virions from HCP. Firstly, maximum purity was estimated from cumulative distributions of elution volume in the limiting case of an infinitely short injection (i.e. from integrals of impulse-responses). The lower limit of the product fraction was set to 0 cv, the upper limit was chosen such that 90% of eluting product (i.e. HA activity) would be collected in the product fraction. Results for all signals are summarized in Table 1. The content of light scattering area (only set B), UV area, total protein and HCP content were calculated accordingly. Good correlation of light scattering area and HA activity was observed in all cases (86–90%) indicating a high specificity of light scattering for virions. The minimum content of total protein increased from 30% to 40% with increasing flow velocity for set B. Similarly the estimated content of HCP increased from 0.2% to 9.3%. Prediction of total protein for set A was slightly higher with 44% (compared to set B at 60 cm h^{-1}) but content of HCP was similar (5.9% and 6.4%).

Secondly, finite injections with volumes ranging from 0.01 to 0.5 cv were considered (Fig. 7). Impulse-responses determined from experimental set B were used for simulation. Rectangular injection profiles were assumed. Lower limits of the product fraction were chosen flexibly to permit a loss of 1% of product but were almost constant with values around 0.3 cv (not shown). Upper limits of product fractions were predicted as functions of injection volume such that 90% of product would be recovered (bold lines). Correlations were almost identical for all flow velocities due to similar impulse-responses for HA activity.

The residual content of HCP (shaded areas) was predicted as a function of injection volume and upper fraction limit (not necessarily the limit corresponding to 90% product yield). Numbers in shaded areas denote the maximum content of HCP to be expected for a given value pair. Evidently, the degree of impurity increased with increasing injection volume and flow velocity. The final content of HCP for the special case of 90% product yield can be determined along bold lines.

Additionally, the final concentration of product normalized by the initial concentration and volumetric productivities are provided for the special case of 90% product yield. Again, predicted concentrations were almost identical at different flow velocities due to similar impulse-responses. Concerning productivity two strategies for elution were considered: in the first scenario, sample is injected and eluted at a constant flow velocity. In the second scenario, flow velocity is increased to 250 cm h^{-1} (recommended maximum for Sepharose 4 FF) after virus has eluted (referred to as two-step elution). A scheme of the second scenario is depicted in Fig. 8. In both scenarios overlapping injection cycles were assumed. Naturally, productivity increased with injection volume and flow velocity. A significant gain in productivity was achieved by the second elution strategy—in particular at the lowest flow velocity. For higher flow velocities the difference was less pronounced.

Crosshairs were included in the first diagram to give an example of how to interpret graphs. In this example, injection of 0.3 cv requires an upper fractionation limit of 0.62 cv for 90% recovery of product. The normalized concentrations of product after chromatography (0.9) and volumetric productivities (0.42 and 0.21 cv h^{-1}) are obtained at intersections of the vertical line with corresponding graphs (not shown). The maximum content of host-cell protein (5%) is indicated by the shaded area underneath.

Table 1
Maximum purity at a product yield of 90% based on HA activity in the limiting case of an infinitely short injection

Set	u[a] (cm h^{-1})	V_r[b] (cv)	Product		Impurity		
			HA[c] (%)	LS[d] (%)	UV[e] (%)	Protein (%)	HCP[f] (%)
B	30	0.41	90	90	8.5	30	0.2
A	60	0.49	90	–	12.5	44	5.9
B	60	0.38	90	86	9.2	38	6.4
B	120	0.37	90	86	9.8	40	9.3

[a] Superficial flow velocity.
[b] Upper limit of product fraction.
[c] HA activity.
[d] Light scattering area.
[e] UV area.
[f] Host-cell protein.

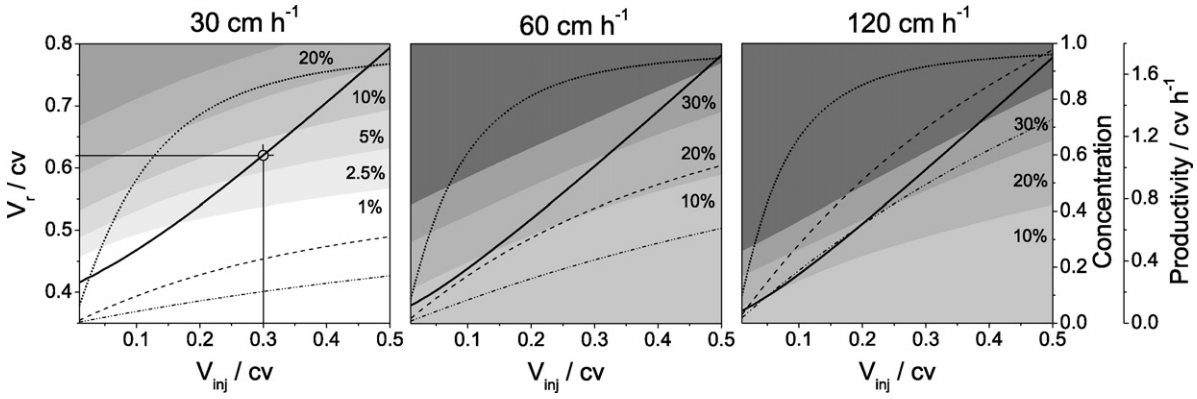

Fig. 7. Model-based design of a preparative SEC operation for the separation of influenza virus from host-cell protein. Upper limits of the product fraction V_r (—) were predicted as a function of injection volume V_{inj} (90% product yield). Corresponding normalized concentration of the product (···) and volumetric productivities for a constant superficial mobile phase velocity (-···-) and two-step elution strategy (- - -) are provided in addition. The maximum residual content of host-cell protein is indicated by shaded areas underneath.

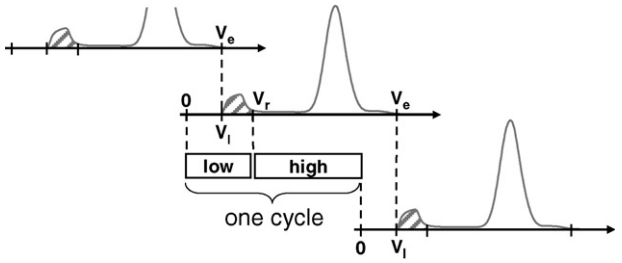

Fig. 8. Two-step elution strategy for maximizing productivity in size-exclusion chromatography. Sample is injected and eluted in overlapping cycles. Chromatograms of three cycles (UV trace over elution volume) are illustrated as cascaded graphs. Product (hatched area) elutes first and is to be separated from impurities (open area). Sample is injected and product eluted at low speed (respective phase marked by box "low"). The left and right limits of the product fraction are indicated by V_l and V_r. After elution of the product, resolution is no longer important and flow velocity is increased to the maximum rate compatible with the chromatography medium (phase marked by box "high"). The endpoint of elution is denoted by V_e. The next cycle is started such that V_l and V_e match and there is no contamination of the product by residual impurities in the column.

Analogous to predictions for synthetic markers, injection volume was eliminated as a free parameter by setting limits for the final content of HCP. Consequently, injection volume, the final concentration of product and productivity reduced to functions of flow velocity alone. Simulation results for the two scenarios of 10% and 20% final HCP content are illustrated in Fig. 9. As before, two strategies were considered for the calculation of productivity: a constant flow velocity and two-step elution.

Remarkably, the trend in productivity was inverse for both scenarios. For the lower limit of 10% HCP, productivity decreased with increasing flow velocity. In contrast, for the higher limit of 20% a significant gain in productivity was achieved. Productivity of the two-step elution strategy was always higher than for elution at a constant flow velocity. Since higher flow velocities required the reduction of injection volume due to stronger band broadening, concentration of the product was likewise reduced. Higher flow velocities also led to slight shortening (up to 0.27 cv) of injection-elution cycles due to a shift of HCP towards lower elution volumes (not shown).

Finally, impulse-responses from both sets were used to predict outcome of a preparative SEC operation that had been established before [34]. In these experiments, a column of the same length but wider inner diameter was used. This column was loaded with 0.15 cv of virus concentrate and injected solutes were eluted at a constant flow velocity of 60 cm h^{-1}. Product fractions were collected between 0.30 and 0.58 cv and analyzed for HA activity and total protein. Product yields based on HA activity and the final content of total protein (predicted and experimental) are summarized in Table 2. Predictions for flow velocities of 30 and 120 cm h^{-1} are

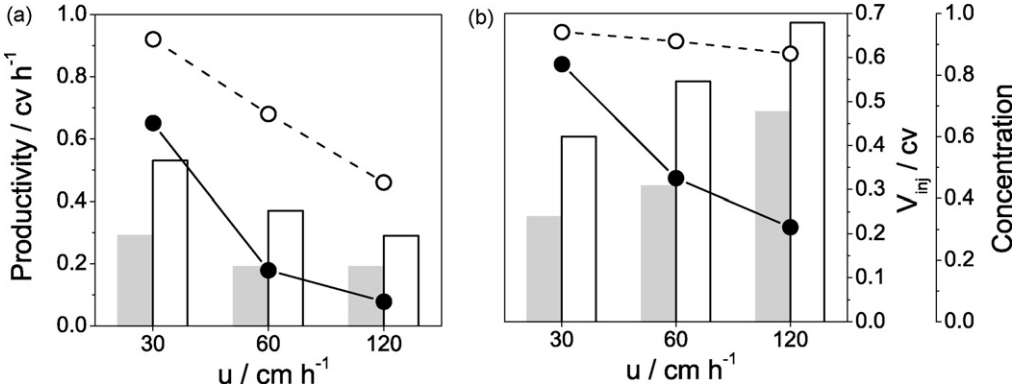

Fig. 9. Performance of a preparative SEC operation for the purification of influenza virus. Limits of the product fraction were set to result in 90% recovery of product. Injection volume V_{inj} (●) was maximized under the constraint of 10% (a) or 20% (b) final content of host-cell protein. The normalized concentration of product (○) and productivity were predicted as functions of superficial mobile phase velocity. Productivity was calculated assuming a constant flow velocity (shaded bars) or a two-step elution strategy designed to maximize productivity (non-shaded bars). Trends of productivity with respect to flow velocity were inverse for 10% and 20% final content of host-cell protein. The two-step elution strategy yielded significant gains in productivity.

Table 2
Prediction of yield and purity for a preparative SEC operation

Set	u[a] (cm h^{-1})	V_{inj}[b] (cv)	Product		Impurities			Productivity	
			HA[c] (%)	LS[d] (%)	UV[e] (%)	Protein (%)	HCP[f] (%)	Constant[g] (cv h^{-1})	Two-step[h] (cv h^{-1})
B	30	0.15	97	95	9	35	4.1	0.11	0.25
A	60	0.15	90	–	13	45	7.1	0.21	0.37
B	60	0.15	98	97	11	52	25	0.22	0.44
B	120	0.15	98	98	14	60	37	0.49	0.72
Exp.[i]	60	0.15	95.8 ± 0.4	–	15.3 ± 1.4	50.2 ± 3.7	–		

[a] Superficial mobile phase velocity.
[b] Normalized injection volume.
[c] HA activity.
[d] Light scattering area.
[e] UV area.
[f] Host-cell protein.
[g] Elution at constant flow velocity.
[h] Two-step elution strategy.
[i] Experimental data from four preparative runs (mean value ± S.D.).

provided for completeness. The yield of UV area was included as a precise but abstract measure. The content of HCP has not been determined in these experiments due to lack of an appropriate assay. Light scattering measurements were neither available since the detector saturated for large injection volumes. Supplementary to data on yield and purity productivities were included for both strategies: elution at a constant flow velocity and the two-step strategy. Please note, that the slightly different productivities for set A and B at a flow velocity of 60 cm h^{-1} result from differences in corresponding UV traces, which were used to determine the length of elution (see Section 3 for details).

Average recovery of virions (95.8%) was close to the predicted value based on set B (98%). The same was true for the final content of total protein (50.2% compared to 52%). In general, both quantities were slightly overestimated by predictions based on set B and underestimated by predictions based on set A. Recovery of UV area, in contrast, was slightly underestimated by both sets (15.3% compared to 13% and 11%). Most striking was the difference in predicted HCP content (7.1% and 25%). Unfortunately, experimental data for comparison was not available here.

5. Discussion

5.1. Size-exclusion chromatography as an LTI system

Despite the fact that SEC is one of the best-characterized principles in chromatography, development of an efficient SEC operation still constitutes a laborious task. A two-dimensional space is spanned by the parameters flow velocity and injection volume alone already requiring a large number of experiments. If column length was to be altered in addition, a three-dimensional parameter space would be obtained. Including further parameters like the choice of resin or eluent, a vast number of combinations arises that can be barely covered experimentally. Viable approaches are therefore needed to reduce the problem to manageable extend.

Various strategies have been suggested in the past (plate, general rate or stochastic models) for the prediction of chromatographic outcome. What is common to all of them is that the composition of the feed stream has to be known a priori and individual components need to be characterized in order to come up with valid predictions. Prediction of the effect of volume overloading is furthermore not straight forward although usually feasible.

The potential of treating SEC as an LTI system is that it does not rely on any such information but is an inverse method. The effect of volume overloading can be described after a single experiment used to determine the system impulse-response. It can further deal with abstract signals like UV absorbance or total protein concentration without having to decompose them into their individual contributions. It has to be noted, however, that the effect of flow velocity or column length cannot be predicted as opposed to other methods but has to be determined experimentally. Still, by treating SEC as an LTI system, the parameter space spanned by injection volume, flow velocity and column length is efficiently reduced by one dimension.

5.2. Synthetic markers

Chromatography of a single marker at various concentrations and comparison of elution of a mixture to its components indicated that equalities defined in (1) and (2) do hold for the system investigated. Experimental validation of predictions for the separation of 100-nm polystyrene beads from a 201-kDa PMA marker provided even stronger evidence. Impulse-responses of synthetic markers (Fig. 2) seemed to be well approximated by chromatograms after short injections indicated by the close match between prediction and experiment (Fig. 4). Slight bias in predictions for the product yield (97.5% instead of 95%) was noticed, however. In general, approximation of impulse-responses by chromatograms obtained after short injections leads to estimates that are slightly broader. The same is true for simulations based on these. Collection of product according to simulations consequently leads to higher recovery than predicted since product peaks in experiments are narrower than assumed. In contrast, for the PMA marker no bias was observed. A fact which is likely related to the significantly wider elution peaks such that error due to approximation of impulse-responses is negligible.

Chromatograms for the PMA marker (representative for a partially excluded solute) further illustrate the effect of limiting mass transfer at high flow velocities on zone broadening and peak asymmetry (Fig. 2). Accordingly, injection volume and hence productivity and the final concentration of product were strong functions of flow velocity (Fig. 3). Chromatograms of polystyrene beads (representative for totally excluded solutes), in contrast, were almost indistinguishable due to the absence of mass transfer (not shown). Flattening of productivity towards higher flow velocities suggests that a maximum exists. Whether this maximum is approached asymptotically or inverse behavior can be observed at even higher flow velocities depends on which effect is dominant: shortening of injection-elution cycles due to higher flow velocity or decrease in injection volume due to increased zone broadening.

In principle, the strategy presented serves as a general protocol for optimizing group separations (i.e. an excluded product is to be separated from partially included impurities). Firstly, impulse-

responses are to be determined for a number of flow velocities. Secondly, free parameters like the injection volume or limits of fractionation are fixed indirectly by specifying objectives for the yield and final content of impurity. Thirdly, chromatography is simulated subject to the chosen constraints. Finally, benchmarks like productivity or final product concentration are calculated from simulations as functions of flow velocity. An adequate operating point can now be chosen according to process requirements.

5.3. Influenza virus

Successful prediction of chromatograms from three runs by means of a single impulse-response (one for each type of signal) underlines applicability of linear systems theory in the case of complex feed streams (Fig. 5). Prediction was particularly accurate for UV absorbance—the signal of highest quality. A good match between simulation and experiment was also obtained for HA activity. The poor match for total protein can be related to low concentrations in some fractions close to the limit of quantitation (about 5 $\mu g\,mL^{-1}$). In addition, matrix interference has been observed for this particular assay due to media components leading to unclosed material balances (typically around 60%). Results for total protein therefore do not invalidate the general approach but only reflect the low quality of the signal.

Chromatography at different flow velocities produced uniform impulse-responses for HA activity (Fig. 6), a result which confirms that virions are totally excluded from the resin. In the absence of mass transfer to the stationary phase (film and pore diffusion) solutes are only subject to moving zone dispersion. Moving zone dispersion, again, is characterized by the Peclet number which is only a weak function of flow velocity [18] explaining the results.

Reconstructed impulse-responses for HCP, on the contrary, strongly depended on flow velocity. Increasing broadening and a shift of the distribution towards lower elution volume was observed at high flow velocities (120 and 60 $cm\,h^{-1}$ compared to 30 $cm\,h^{-1}$). Similar to the 201-kDa PMA marker (Fig. 2) these changes can be interpreted as the result of limiting mass transfer. As derived by Li et al. [18], broad and skewed peaks are the result of low η number, which again is caused by high flow velocities. If the accessible pore fraction ε_a and effective diffusivity D_p are further low (which is the case for almost excluded solutes) very low numbers for η are the result and broad asymmetric peaks are to be expected. In terms of the stochastic model [20], limitation in mass transfer is represented by low values of α—the relative magnitude of the size-exclusion effect for the egress process m_p with respect to the total size-exclusion effect (ingress and egress). Low α means domination of the ingress process which occurs due to lower residence time in the column at high flow velocities and the hence reduced mean values of the ingress step number $\bar{n}_p$. According to theory, low α leads to low effective plate heights (i.e. broad) and skewed peaks. Again, the effect is predicted to be strongest for solutes close to the exclusion limit ($K_{SEC} \to 0$).

Interestingly, impulse-responses for HCP derived from experimental sets A and B deviate significantly. A much higher content of HCP eluting at low elution volume was seen in the response derived from set B. To some extend this may be related to the fact that HCP was reconstructed and not measured. But qualitative differences in reconstructed distributions (Figs. 5 and 6) suggest that real differences exist. Whether these have to be attributed to batch-to-batch variability, inappropriate storage of sample or different eluents cannot be decided at this point. More experiments would have to be conducted using different lots of virus in order to collect statistical data.

Predicted minimum content of UV area, total protein and HCP were consistent in the sense that values increased with increasing flow velocity due to stronger band broadening. Besides, values for the flow velocity of 60 $cm\,h^{-1}$ derived from both sets compared well. Obviously, the low value of 0.2% for HCP at a flow velocity of 30 $cm\,h^{-1}$ should not be overstressed. Reconstructions for HCP are certainly only of modest accuracy and predicted proportions will be particularly inaccurate for values close to zero. Nevertheless, it can be concluded that although complete separation from HCP does not seem to be feasible, a final content below about 10% can be achieved.

Comparison of a previously established SEC operation [34] to simulations resulted in good agreement for the yield of product based on HA activity (Table 2). As expected, yields of light scattering area resembled that of HA activity quite well confirming that light scattering can be used as an online signal for the quantitation of virions [36]. It is noteworthy that simulations and experimental results were both close to a yield of 95%—the value originally targeted when designing the operation.

A good match was also observed for the final content of UV area and total protein. The value of such a comparison is, however, questionable since neither of these signals is particular specific for the product or impurities. Changes in product yield or purity are therefore less reflected by changes in material balances of these signals. This becomes immediately evident by looking at predictions for the final content of HCP. Despite similar predictions for the content of total protein (45% and 52%), a large discrepancy occurred with respect to HCP (7.1% and 25%)—a result that owes to the clear difference between estimated impulse-responses.

Productivity of the operation assuming non-overlapping injection-elution cycles (injection volume: 0.15 cv, flow velocity: 60 $cm\,h^{-1}$, run length: 1.91 cv) was calculated previously to be 0.15 $cv\,h^{-1}$ [34]. By admitting overlapping injection-elution cycles (including a 0.2 cv safety margin) a significant gain to 0.21 or 0.22 $cv\,h^{-1}$ (set A or B, respectively) can be achieved (Table 2). Applying the two-step elution strategy described in Fig. 8 productivity even rises to 0.37 or 0.44 $cv\,h^{-1}$, respectively. Since, however, the high content of HCP predicted by set B is less desirable, operation at a flow velocity lower than 60 $cm\,h^{-1}$ should be considered. The predicted content of HCP at 30 $cm\,h^{-1}$ was again below 10% but higher purity would be achieved at the cost of reduced productivity. A compromise could be to reduce the target product yield slightly to about 90% (instead of 95%), which would allow for injection of at least 0.3 cv at 30 $cm\,h^{-1}$ according to predictions in Fig. 7. In doing so, productivity would be approximately restored to previous values while the final content of HCP should remain below 10%.

5.4. Determination of impulse-responses

Deconvolution of experimental (i.e. noisy) data is a challenging task. First attempts to deconvolve chromatograms by polynomial long-term division (i.e. direct inversion) were unsuccessful yielding negative values for the concentration and wild oscillations in reconstructions (not shown). Attempts to deconvolve chromatograms by NNLS somewhat improved the result but still lead to noisy impulse-responses (neither shown). Although valid predictions could be derived from these, two major disadvantages remained. On the one hand, noisy reconstructions were difficult to interpret suggesting structure in impulse-responses that was not real. On the other hand, cumulative distributions calculated from impulse-responses (e.g. for the calculation of maximum purity) contained steps due to oscillations in particular for narrow peaks.

Smooth solutions were, in contrast, obtained by MaxEnt deconvolution (up to equal distributions if the admitted degree of misfit was chosen inappropriately). Evidently, the success of MaxEnt deconvolution very much depends on the way data is integrated.

An elegant approach for the integration of data was suggested by Skilling and Bryan [28] based on χ^2 statistics. In practice, however, accuracy of analytical methods is often unknown and sometimes difficult to determine. For this reason, a general cross-validation strategy was suggested by Amato and Hughes [33] to estimate accuracy directly from the data. Although χ^2 statistics appears to be a useful approach, some limitations exist since experimental data may be biased due to miscalibration (e.g. as a result of detector drift), interpolation of offline data or matrix interference which frequently occurs in biological assays. Deconvolution is further strongly affected by the assumption of an inaccurate injection profile. Such sources of error cannot be covered adequately by any χ^2 statistics.

In this article, misfit was therefore still controlled empirically by limiting ε to multiples of $\varepsilon_{\mathrm{NNLS}}$ in an arbitrary manner. While this approach is certainly unsatisfactory and decision should rather be based on sound statistics, it was sufficient to demonstrate that the problem of estimating impulse-responses can be successfully attacked by MaxEnt deconvolution. Further improvement of the algorithm shall be left to experts in the field.

Independently of the algorithm used, limited resolution of offline measurements poses a serious restriction. Some improvement can be achieved by (weighted) spline interpolation. But in general, some if not most of the fine structure is lost and quality of predictions reduced. Moreover, offline measurements often pose an experimental bottleneck. Unfortunately, conventional online detectors like UV or RI are of limited value in biochromatography of complex feed streams as they are usually not specific for the product or impurities. New strategies have thus to be pursued for the generation of appropriate online signals. With respect to influenza virus online light scattering has been suggested as an approximate measure of virion quantity [36]. In the case of total protein, after-column derivatization combined with fluorometric detection can be suggested. Importantly, availability of online signals would not only improve resolution but would simultaneously speed up experiments allowing for additional investigations, e.g. studies on batch-to-batch variability.

6. Conclusion

It was demonstrated that SEC can be treated as an LTI system. Based on linear systems theory, the effect of volume overloading was predicted for various solutes by convolution of impulse-responses with injection profiles. In contrast to other methods, no a priori knowledge was required about the composition of feed streams by the inverse approach pursued. MaxEnt deconvolution hereby proved as the key for obtaining impulse-responses from experimental data.

With respect to chromatography of biological feed streams, offline analytics can be identified as the major bottleneck. While, for example, online light scattering appears to be an acceptable replacement for HA activity in the case of influenza virus, no such alternative exists for total protein or HCP. But only after such online signals have become available determination of impulse-responses at many flow velocities or studies on batch-to-batch variability of feed streams can be conducted efficiently. Development of appropriate online detectors is therefore mandatory before the suggested method will become truly useful.

Concerning the method itself, a combination of linear systems theory with mechanistic models should be considered in the future in order to incorporate the effects of flow velocity and column length. Ideally, inverse methods would be developed that allowed to gain sufficient information on the composition of feed streams and properties of the principal components in a limited number of experiments. Impulse-responses could then be predicted by a rate or statistical model and volume overloading subsequently simulated by convolution with the desired injection profile.

References

[1] J. Porath, P. Flodin, Nature 183 (1959) 1657.
[2] A. Buchacher, G. Iberer, A. Jungbauer, H. Schwinn, D. Josic, Biotechnol. Progr. 17 (2001) 140.
[3] S. Gibert, N. Bakalara, X. Santarelli, J. Chromatogr. B 737 (2000) 143.
[4] K. Hashimoto, S. Adachi, Y. Shirai, Agric. Biol. Chem. 52 (1988) 2161.
[5] K. Reissner, A. Prior, J. Wolfgang, H.J. Bart, C.H. Byers, J. Chromatogr. A 763 (1997) 49.
[6] L.Z. Li, Y. Liu, M.S. Sun, Y.M. Shao, J. Chromatogr. A 1139 (2007) 228.
[7] G. Paredes, S. Makart, J. Stadler, M. Mazotti, Chem. Eng. Technol. 28 (2005) 1335.
[8] J. Urthaler, W. Buchinger, R. Necina, Acta Biochim. Polonica 52 (2005) 703.
[9] H.H. Gschwender, W. Haller, P.H. Hofschneider, Biochim. Biophys. Acta 190 (1969) 460.
[10] J.T. Heyward, R.A. Klimas, M.D. Stapp, J.F. Obijeski, Arch. Virol. 55 (1977) 107.
[11] C.C. Loa, T.L. Lin, C.C. Wu, T.A. Bryan, H.L. Thacker, T. Hooper, D. Schrader, J. Virol. Methods 104 (2002) 187.
[12] D.P. Nayak, S. Lehmann, U. Reichl, J. Chromatogr. B 823 (2005) 75.
[13] M.D. Segura, A. Kamen, P. Trudel, A. Garnier, Biotechnol. Bioeng. 90 (2005) 391.
[14] J. Transfiguracion, D.E. Jaalouk, K. Ghani, J. Galipeau, A. Kamen, Hum. Gene Ther. 14 (2003) 1139.
[15] Y. Xie, S. Mun, J. Kim, N.H.L. Wang, Biotechnol. Progr. 18 (2002) 1332.
[16] J.C. Giddings, K.L. Mallik, Anal. Chem. 38 (1966) 997.
[17] S. Yamamoto, M. Nomura, Y. Sano, J. Chromatogr. 512 (1990) 77.
[18] Z.G. Li, Y.S. Gu, T.Y. Gu, Biochem. Eng. J. 2 (1998) 145.
[19] B. Zelic, B. Nesek, Eng. Life Sci. 6 (2006) 163.
[20] F. Dondi, A. Cavazzini, M. Remelli, A. Felinger, M. Martin, J. Chromatogr. A 943 (2002) 185.
[21] L. Pasti, F. Dondi, M. van Hulst, P.J. Schoenmakers, M. Martin, A. Felinger, Chromatographia 57 (2003) S171.
[22] A. Felinger, L. Pasti, F. Dondi, M. van Hulst, P.J. Schoenmakers, M. Martin, Anal. Chem. 77 (2005) 3138.
[23] Z. Gajic, Linear Dynamic Systems and Signals, Pearson Education, Upper Saddle River, NJ, 2003.
[24] A. Mohammad-Djafari, J.-F. Giovannelli, G. Demoment, J. Idier, Int. J. Mass Spectrometry 215 (2002) 176.
[25] S. Bott, in: Dahneke (Ed.), Measurement of Suspended Particles by Quasi-Elastic Light Scattering, Wiley, New York, 1983.
[26] H. Rullgard, O. Öktem, U. Skoglund, Inverse Problems 23 (2007) 2121.
[27] J.E. Shore, R.W. Johnson, IEEE Trans. Inf. Theory IT-26 (1980) 26.
[28] J. Skilling, R.K. Bryan, Month. Notices Royal Astr. Soc. 211 (1984) 111.
[29] A.K. Livesey, J. Skilling, Acta Crystallographica A41 (1985) 113.
[30] M.K. Charter, S.F. Gull, J. Pharmacokinet. Biopharmaceut. 15 (1987) 645.
[31] S.F. Gull, G.J. Daniell, Nature 272 (1978) 686.
[32] F.N. Madden, K.R. Godfrey, M.J. Chappell, R. Hovorka, R.A. Bates, J. Pharmacokinet. Biopharmaceut. 24 (1996) 283.
[33] U. Amato, W. Hughes, Inv. Prob. 7 (1991) 793.
[34] B. Kalbfuss, M. Wolff, R. Morenweiser, U. Reichl, Biotechnol. Bioeng. 96 (2007) 932.
[35] B. Kalbfuss, Y. Genzel, M. Wolff, A. Zimmermann, R. Morenweiser, U. Reichl, Biotechnol. Bioeng. 97 (2007) 73.
[36] B. Kalbfuss, M. Wolff, L. Geisler, A. Tappe, R. Wickramasinghe, V. Thom, U. Reichl, J. Membr. Sci. 299 (2007) 251.