ELSEVIER

Journal of Chromatography A, 937 (2001) 1–11

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

Graphical method for the calculation of chromatographic performance in representing the trade-off between purity and recovery

S.H. Ngiam, Y.H. Zhou, M.K. Turner, N.J. Titchener-Hooker*

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

Received 23 February 2001; received in revised form 11 September 2001; accepted 19 September 2001

Abstract

A simple engineering framework that enables the rapid representation of the performance of liquid chromatographic separations is provided in this paper. The fractionation diagram and its associated maximum purification factor versus product yield, and contamination index versus product yield diagrams, may be derived directly from chromatographic data. The fractionation diagram plots the relative change in the cumulative fractional mass of product eluted with the corresponding fractional total mass eluted, while the maximum purification factor versus yield diagram shows the degree of trade-off between the levels of purity and recovery achieved in the chromatographic step. The minimum contamination index versus yield plot is especially suitable for cases where the product and impurity are expressed in different units and shows how the extent of contaminant removal changes relative to product yield. These diagrams are more straightforward and easily interpretable compared to the basic conventional chromatograms and enable investigation of the degree of trade-off between purity and recovery for any set of operating conditions to be made. The approach is demonstrated for two different chromatographic systems. In the first, a set of simulation results from a verified size exclusion model is used to demonstrate the approach for product recovery. In the second, a set of experimental results for the removal of endotoxin from DNA is used. This demonstrates a problem where the product and impurity content are measured by different assay techniques and are expressed in different units, and also where the quality of process information is limited by the small number of fractions collected. The studies show how such an approach can help to identify the optimal operating conditions, in terms of acceptable yield and desired level of contaminant removal, and to redefine the location of product fractions needed to achieve these specifications.

Keywords: Fractionation diagram; Purity; Recovery; Purification factor; Contamination factor; Product yield

1. Introduction

Due to its excellent resolving power, chromatographic separation is a pre-eminent downstream processing operation for the isolation of biological products and is almost indispensable in the manufacturing of biopharmaceuticals destined for human diagnostic and therapeutic purposes [1]. Indeed in order to ensure the stringent and exact purification specifications are met, regulatory authorities require

*Corresponding author. Tel.: +44-207-679-3796; fax: +44-207-383-2348.

E-mail address: nigelth@ucl.ac.uk (N.J. Titchener-Hooker).

0021-9673/01/$ – see front matter
PII: S0021-9673(01)01299-7

at least one chromatographic step to be employed in the manufacture of a protein-based pharmaceutical [2]. Such requirements place an increasing demand for more efficient and reliable preparative and large-scale chromatographic techniques and for methods to determine how best to operate them in order to achieve desired levels of performance, usually expressed in terms of yield and/or purity.

Prediction of chromatography performance is currently rather limited and time consuming given the complex nature of the mechanisms governing separation. Chromatograms are often the final output of modelling predictions [3–5], but they are not very straightforward as a means of quantifying the consequences for performance and the sensitivity of the chromatographic separation quality to changes in operating conditions. This is especially laborious for systems where both the product and impurity content are expressed in different units, and where the quality of process information is limited by the small number of fractions collected. The need for more straightforward and faster approaches for the prediction and visualisation of chromatographic performance forms the focus of the research work reported in this paper which aims to provide a robust yet straightforward representation of chromatographic performance which is of use in determining operating strategies for such separations.

The methodology proposed serves as a useful tool for identifying the process trade-offs during a chromatographic separation and indicating the impact on further processing of the cut-point decisions that are made. In the manufacture of pharmaceuticals, the chromatographic step is typically a validated process operation where the operating variables and conditions have to be maintained as consistent as possible. We therefore envisage the fractionation maximum purification–yield and minimum contamination index–yield diagrams being developed on the basis of an initial separation and then applied to all subsequent separations as a reference for cut-point decision.

Richardson and co-workers [6,7] developed the fractionation diagram approach for the study of fractional protein precipitation and have successfully illustrated the use of such an approach for the optimisation of the precipitation conditions to use in order to realise a desired objective function expressed in terms of the process yield and the precipitate product purity. They used the fractionation diagram and its associated maximum purification factor against yield diagram to display how process parameters, such as the pH at precipitation, affected the relationship between product yield, purity and the maximum purification that could be achieved. Similar diagrams have also been employed to represent the performance of membrane filtration separation [8], where a methodology was developed for optimisation purposes.

In preparative and process-scale chromatographic separations, purity and yield are often the two major process outputs whose trade-off have to be considered in determining the best operating conditions. The direct analogy between the separation sought in protein precipitation and that achieved in chromatography was investigated in this paper and specifically the utility of the fractionation diagram and maximum purification factor against yield diagram approaches for visualising the impact of processing conditions on the chromatographic output. The approach is demonstrated for two different chromatographic separation systems: a set of simulation results from a verified size exclusion model separating three globular proteins (i.e. γ-globulin, ovalbumin and Ribonuclease A), and a set of experimental results for the removal of endotoxin from DNA by ion-exchange chromatography. The paper details the theoretical basis behind the fractionation, purification factor–yield and contamination index–yield diagrams and the mathematical algorithms used to connect chromatographic data with these before describing the experimental systems used to demonstrate the approach.

2. Theoretical development

2.1. Fractionation diagram

The basis of the fractionation diagram approach explored in this paper is shown in Fig. 1. A fractionation diagram is constructed based upon the concentration profiles of the different components being separated, obtained either directly from the elution chromatogram, if a specific on-line assay exists for the product and the total material, or from

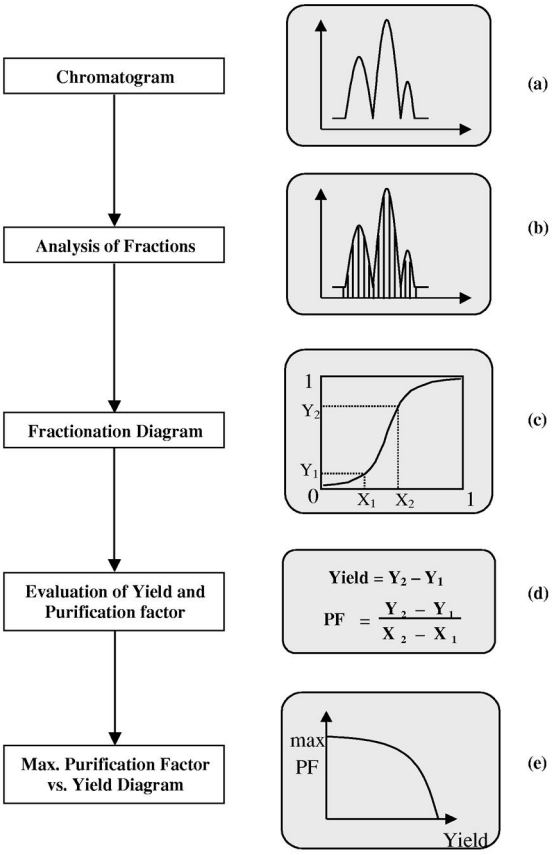

Fig. 1. Schematic illustration of the procedure to generate the fractionation diagram and the corresponding maximum purification factor–yield diagram from a elution chromatogram obtained either by experiments or simulation. *X* and *Y*, which are the *x*- and *y*-axis of the fractionation diagram, represent the cumulative fraction of total material and target product, respectively, eluted at any time. Subscripts 1 and 2 denote the start and end points of product collection on the fractionation curve. *PF* denotes the purification factor between any two points as defined by Eq. (6).

the corresponding off-line data (Fig. 1a). The problem is first simplified into a three component separation, the product and all the other contaminants being treated as two pseudo-components, one which elutes before the product, the other after. This recognises that for the purposes of processing the aim is to isolate the product peak from all the other species. The chromatogram is then fractionated into *N* steps (Fig. 1b), such that the whole chromatographic profile is divided into *N* elements with equal width (time or volume intervals).

Hypothetically, if a system comprises three components (Fig. 1a), the product (P) and the two pseudo-impurities (A and B), the amounts of product (M_P), impurity A (M_A) and impurity B (M_B) in each interval can be calculated. For the *i*-th interval they would be denoted as $M_{P,i}$, $M_{A,\,i}$ and $M_{B,\,i}$. Hence the total amount of material in a particular fraction (M_T) is the sum of all these all quantities:

$$M_{T,\,i} = M_{P,\,i} + M_{A,\,i} + M_{B,\,i} \tag{1}$$

For each interval, we define an arithmetic mean time (t_m) as defined between the upper and lower limit of the interval, for example, for the first interval:

$$t_m = \frac{1}{2}(t_I + t_{II}) \tag{2}$$

The fractionation diagram plots the changes in the cumulative fractional mass of product eluted with the corresponding fractional total mass eluted. Hence the axes are defined as:

$$\textit{Fractional mass of material eluted, } X = \frac{\textit{Cumulative mass of material eluted at time } t}{\textit{Total mass eluted at } t = \infty}$$

$$\textit{Fractional mass of product eluted, } Y = \frac{\textit{Cumulative mass of product eluted at time } t}{\textit{Total mass of product eluted at } t = \infty}$$

Since *X* and *Y* are fractions, the values fall in the range between 0 and 1. A theoretical fractionation diagram can then be generated as shown in Fig. 1c.

2.2. Purification factor

Having defined the fractionation diagram it is now possible to calculate operating performance parameters such as the purification factor (*PF*), defined as the ratio between the final purity of the product after purification to the starting purity of a load sample:

$$PF = \frac{\textit{Final purity}}{\textit{Initial purity}} = \left[\frac{M_P^{(2)} - M_P^{(1)}}{M_T^{(2)} - M_T^{(1)}}\right] \Big/ \left[\frac{M_0}{M_S}\right] \tag{3}$$

where M_S and M_0 are the total amount of product

and impurity in the sample load and the amount of product in the sample load, respectively. Rearranging them:

$$PF = \left[\frac{M_P^{(2)}}{M_0} - \frac{M_P^{(1)}}{M_0} \right] \Big/ \left[\frac{M_T^{(2)}}{M_S} - \frac{M_T^{(1)}}{M_S} \right] \quad (4)$$

where P and S represent the amount of product component and sample, respectively, subscript 0 represents the initial condition, and superscripts (1) and (2) are the points of the starting and end collection times, respectively. The terms (M_T/M_S) and (M_P/M_0) define the x-axis (X) and y-axis (Y), respectively, in the fractionation diagram. Therefore Eq. (4) is actually the gradient of the fractionation diagram between any two points corresponding to the start and end of product collection in the chromatogram.

It is possible to operate at any positions along the fractionation curve that satisfy the following criterion: the vertical distance between any two points gives the product yield; the slope of the tangent between these two points is the value of purification factor corresponding to that yield. By varying the position of the collection points a plot of purification factor against yield can then be produced. The plot obtained represents the set of all the possible values of purification factors achievable for any combination of two points on the graph (Fig. 2) and will range from low yields and high purification factors to the opposite of high yields and low purification factors. Of particular interest to the engineer is the ability to select conditions that maximise the purification factor for a given yield or to determine the best yield that can be achieved for a specific maximum purification factor.

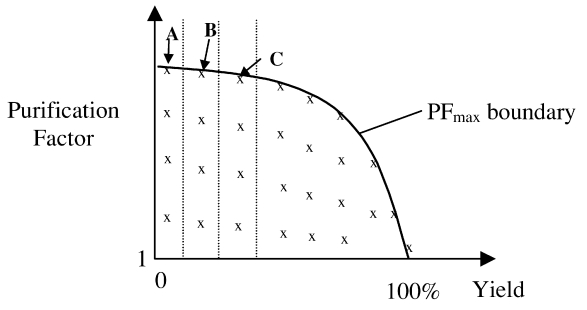

Fig. 2. Theoretical plot of purification factor against product yield. For any value of product yield, there exist a number of possible purification factors. However, it is often only the maximum purification factor (PF_{max}) that is of interest. This exists at the upper bound of the purification factor. Point A is the maximum value in the first interval, point B in the second and so on.

To generate systematically the PF_{max} vs. yield relationship a searching-type computer algorithm was written. The algorithm was utilised to search through all the purification factors achievable for each yield, to compare the magnitudes of those and to select the maximum purification factor value corresponding to that particular yield. This maximises the tie-line gradient for a given yield enabling a plot of optimum purification factor versus yield to be constructed. For example, in Fig. 2, point A is the maximum purification factor for the first value of yield, point B for the second, C for the third and so on. From such an analysis a plot of the maximum purification factor for a corresponding set of yields can then be produced.

In an ideal separation, the highest possible product purity that can be achieved is 100% where the collection of product is facilitated such that no peak overlapping happens (i.e. baseline separation). Therefore, the ideal value of purification factor for a given system with initial product purity of x% is:

$$\textit{Maximum purification factor}, PF_{max} = \frac{100}{x} \quad (5)$$

For instance, if the initial purity of the sample loaded is 20%, the maximum purification factor in any selected fraction would be 5. As mentioned earlier, this global maximum value can only be realised in a baseline separation where no overlapping contaminant and component peaks exist. This situation rarely happens in real operations.

The final use of the yield vs. purification factor diagram is for identifying the retention times or retention volumes for sample collection. Having determined a desired specific value of product yield and/or purification factor, the time (t_1 and t_2) or retention volume (V_{R1} and V_{R2}) for the cut (sample collection) can be determined by retrieving the data points which correspond to this value of yield or purification factor. These data are generated and stored during the simulation of the purification factor versus yield plot as depicted in Fig. 2.

2.3. Contamination index

In reality it is not uncommon for the product and impurity of a chromatographic separation to be expressed in different units and where there may be no mass equivalence. For example, in the separation of an enzyme from total protein, the enzyme level is often expressed in terms of enzymatic activity, whilst total protein is expressed in normal concentration units, i.e. mg ml^{-1}; in the purification of plasmid DNA for gene therapy, the DNA level is often expressed in mg ml^{-1} whereas that of a major contaminant, endotoxin, is expressed in EU (endotoxin unit). Conversion to consistent units is often not straightforward. For such systems the fractionation diagram is modified to a plot where the relative change in the cumulative fractional amount of impurity with the corresponding fractional mass of product eluted is shown.

Having defined the modified fractionation diagram it is now possible to determine the contamination index (*CI*) that is defined as the amount of impurity remaining in a unit mass of product:

$$CI = \text{Gradient of fractionation diagram} \times \frac{\text{Total amount of impurity}}{\text{Total amount of product}} \quad (6)$$

It is possible to operate at any positions along the fractionation curve that satisfy the following criterion: the horizontal distance between any two points gives the yield; the slope of the tangent between these two points leads to the calculation of the contamination index corresponding to that yield. By varying the position of the collection points, a plot of contamination index against yield can then be produced. Again, similar to the *PF* vs. yield diagram illustrated previously, the plot obtained in this case represents the set of all the possible values of the contamination index achievable for any yield (Fig. 3). For the case of impurity removal, the objective is to minimise the contamination index for a specified product yield. In order to obtain the minimum contamination index (CI_{min}), the searching-type computer algorithm mentioned above was used again. However in this case the algorithm was employed to search through all the contamination index values achievable for each product yield, and to select the minimum value. This enables a plot of minimum contamination index versus product yield to be generated. For example, in Fig. 3, point A is the minimum contamination index for the first value of yield, point B for the second, C for the third and so on. Having specified a desired value of product yield and/or contamination index, the retention time or volume for the sample collection can be determined by retrieving the corresponding data points generated and stored during the simulation of the contamination index versus yield plot.

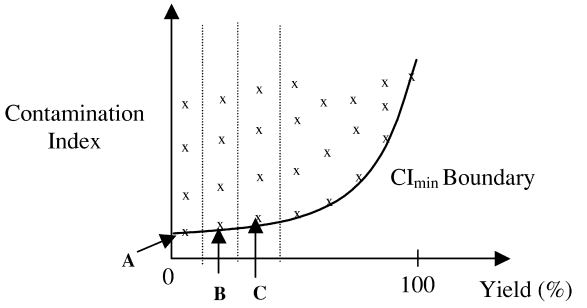

Fig. 3. Theoretical plot of contamination index against product yield. For any value of product yield, there exist a number of possible values of contamination index. The minimum contamination index (CI_{min}) is usually of most interest. This exists at the lower bound of the contamination index. Point A is the minimum value in the first interval, point B in the second and so on.

3. Materials and methods

3.1. Gel filtration model

The fractionation diagram approach was initially established using a computational model to simulate the chromatographic separations required to test the method. Size exclusion chromatography (SEC) was selected due to its simplicity. The model of Boyer and Hsu [9] based upon the general rate model for linear chromatography was employed and a Fast Fourier Transform (FFT) [10–13] numerical technique was used to solve the set of equations. The model was simulated in MATLAB (The Mathworks, Natick, MA, USA) programming environment [14].

The separation of three globular type proteins, γ-globulin, ovalbumin and Ribonuclease-A, was simulated on a column of 49.3 cm bed length and 1.6

Table 1
Operating conditions and process parameters used in the simulation of the gel filtration model

Parameter/operating condition	Value
No. of components	3
Radius of matrix particle, R_P (μm)	4.4×10^{-3}
Length of bed, L (cm)	49.3
Internal diameter of the column, D_i (cm)	1.6
Operating temperature, T (°C)	20
Mobile phase flow-rate, Q (ml min^{-1})	0.1, 0.3 and 0.5
Void fraction of the bed, ϵ	0.332
Density of fluid, ρ (kg m^{-3})	1004.1
Polymer concentration in particle, c_f (g cm^{-3})	0.06
Number of time steps, N	512
Period for FFT algorithm (s)	60 000

cm internal column diameter, packed with Pharmacia® Sepharose 6B (mean matrix particle diameter, 88 μm). Ovalbumin, which has an intermediate molecular weight, was assumed to be the product of interest. The initial concentration of the protein mixture was 50% ovalbumin, 25% γ-globulin and 25% Ribonuclease A. Therefore the initial purity of the sample with respect to ovalbumin was 50%. Separation of 1% by volume load at mobile phase flow-rates of 0.1, 0.3 and 0.5 ml min^{-1} was simulated. Tables 1 and 2 summarise the operating parameters of the simulations and the values of physical properties and calculated transport parameters for the three proteins used in the simulations,

Table 2
Values of physical parameters and calculated transport parameters for the three globular type proteins used in the gel filtration simulations

Protein	γ-Globulin	Ovalbumin	Ribonuclease A
Physical parameters			
Molecular weight of the protein, M_W (Da)	156 000	45 000	13 700
Stokes radius of the particle, r_s (cm)	5.35	2.93	1.83
Intraparticle inclusion porosity, ϵ_p	0.58	0.69	0.81
Initial concentration, C_0 (mg ml^{-1})	1.5	3.0	1.5
Calculated parameters			
Solute diffusivity in unbounded solution, D_m (cm^2 s^{-1})	3.59×10^{-7}	6.55×10^{-7}	1.05×10^{-6}
Effective intraparticle diffusivity, D_e (cm^2 s^{-1})	4.30×10^{-8}	1.41×10^{-7}	3.27×10^{-7}
Schmidt number, *Sc*	3.10×10^{4}	1.70×10^{4}	1.06×10^{4}
Reynolds number, *Re*			
@ 0.1 ml min^{-1}	6.55×10^{-4}	6.55×10^{-4}	6.55×10^{-4}
@ 0.3 ml min^{-1}	0.002	0.002	0.002
@ 0.5 ml min^{-1}	0.0033	0.0033	0.0033
Fluid phase mass transfer coefficient, k_f (cm s^{-1})			
@ 0.1 ml min^{-1}	3.51×10^{-5}	3.96×10^{-4}	5.76×10^{-4}
@ 0.3 ml min^{-1}	3.38×10^{-4}	5.32×10^{-4}	7.63×10^{-4}
@ 0.5 ml min^{-1}	3.96×10^{-4}	6.19×10^{-4}	8.82×10^{-4}
Convective axial dispersion coefficient, D_L (cm^2 s^{-1})			
@ 0.1 ml min^{-1}	3.51×10^{-5}	3.35×10^{-5}	3.23×10^{-5}
@ 0.3 ml min^{-1}	1.15×10^{-4}	1.10×10^{-4}	1.06×10^{-4}
@ 0.5 ml min^{-1}	1.99×10^{-4}	1.90×10^{-4}	1.83×10^{-4}

respectively. The equations and correlations employed to evaluate the transport parameters involved in solving the general rate model are provided in the literature [9,12,13].

All simulations were run on a Hewlett-Packard Vectra-Pentium II 400 MHz, 64 MB RAM PC (Hewlett-Packard, London, UK) to provide output files tabulating the concentration versus time profiles of the three individual components.

3.2. Removal of endotoxin from plasmid DNA

The fractionation diagram approach was further verified by using experimental data for the removal of endotoxin from a DNA product by ion-exchange chromatography (provided by Biopharmaceutical Product Development, GlaxoSmithKline, Kent, UK). The size of the endotoxin clearance column is ~3.5 l. Four sets of chromatographic data were provided corresponding to different fermentation batches where each had been harvested and then purified by the same purification process. In these particular chromatographic runs the target was to achieve a minimum specification in terms of the clearance of endotoxin from DNA product. The precise value for endotoxin contamination is related to the dose regime envisaged for the final product. As an illustration, a level of less than 1 EU/mg DNA was selected as being typical of a medical medication requiring a repeated dose.

3.3. Experimental data treatment

The fractionation diagram generated from the raw chromatogram data is actually a set of data points describing the relative change in the cumulative fraction of impurity eluted from a column with respect to the cumulative fraction of product eluted. In the case of the gel filtration model data, the data points generated were sufficiently numerous and closely spread as to form a smooth fractionation curve. Whilst the interval between each data point can be made to be very small in the case of simulation data the same is not true for real experimental systems where the density of data points is fixed by the size of the fractions taken. For the endotoxin and the DNA the small number of data points available were fitted with a smoothed line obtained by applying a polynomial curve fitting function (to the degree of one) in MATLAB (The Mathworks, Natick, MA, USA) [14]. The purpose of such fitting was to produce a continuous function for subsequent estimation of the tie-line gradients which form the basis of the contamination versus yield plots. After curve-fitting the minimum operating tie-line gradient for a particular yield was derived using these relationships to describe the fractionation diagram.

4. Results and discussions

4.1. Simulation results

The gel filtration model was run to simulate the individual component peak profiles and the resultant chromatogram shape as a function of flow-rate of the mobile phase. The results were used to generate fractionation diagrams and maximum purification factor versus yield diagrams. The sensitivity and variation of the diagram towards the changes in the operating flow-rates were investigated for mobile phase flow-rates of 0.1, 0.3 and 0.5 ml min^{-1}.

A detailed examination of the individual component profiles for the 0.3 ml min^{-1} data is given in Fig. 4. With respect to the product, ovalbumin, peak overlapping with γ-globulin happens between 235 and 265 min and peak overlapping with Ribonu-

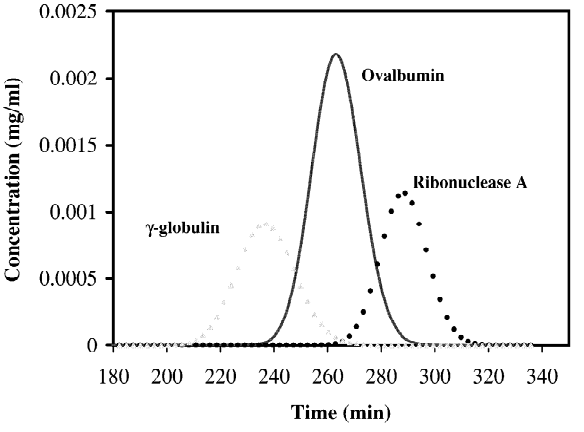

Fig. 4. Simulated SEC chromatogram for elution of a three component model system at a mobile phase flow-rate of 3 ml min^{-1}.

clease A exists between 260 and 300 min. Hence if fraction collection is between 260 and 265 min, a high purity of sample can be achieved but at the expense of product yield which will be very low as much product has to be sacrificed to realise a sample comprised of ovalbumin with only traces of impurities.

Analysis of the simulated chromatograms was carried out to calculate the changes in cumulative fractions of total protein and product component, purification factor and yield obtained during the elution period. Figs. 5 and 6 are the resultant fractionation curve and maximum purification factor against product yield diagram, respectively. All of the fractionation diagrams display a common ''S'' shaped curve. The fractionation diagram for the lowest flow-rate, 0.1 ml min^{-1}, had the steepest curve, while 0.5 ml min^{-1} is the flattest. The regions of greatest variance are at the start and end of the elution step (low and high values of X). Here it can be seen that at the highest flow-rate the product starts to elute ($Y>0$) at $X=0.15$ compared to $X=0.23$ for the lowest flow-rate. Product elution is complete at $X=0.75$ in the case of the lowest flow-rate but continues to $X=0.95$ for the highest flow-rate. These data show how the degree of product and impurity overlap increases with flow-rate. As seen in the corresponding purification factor versus yield plot even apparently small differences in the fractionation diagram lead to significant changes in the process output performance (Fig. 6). Since purification factor is defined as the gradient of the tie line on the fractionation curve between any two points, it can be deduced that the separation attained at the lowest mobile phase flow-rate has the highest value of purification factor. The simulated purification factor vs. yield results demonstrate this result well (Fig. 6) and show clearly how the separation performance improves with a reduction in flow-rate.

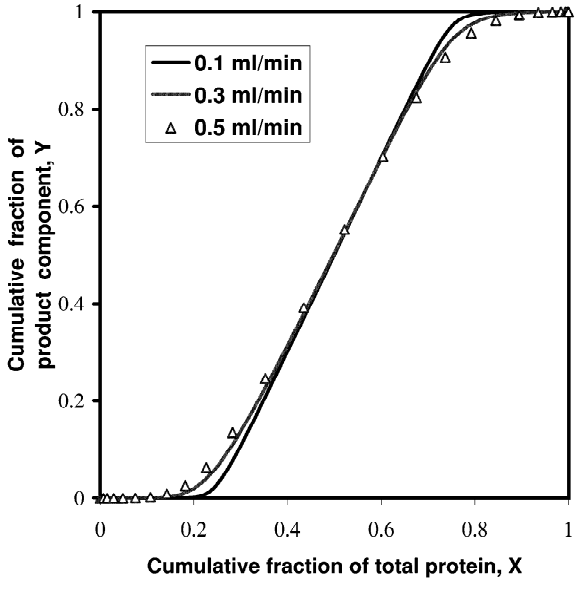

Fig. 5. Fractionation curves for the simulated gel filtration separation of ovalbumin from a mixture of Ribonuclease A and γ-globulin as a function of mobile phase flow-rates (solid line, 0.1 ml min^{-1}; broken line, 0.3 ml min^{-1}; Δ, 0.5 ml min^{-1}).

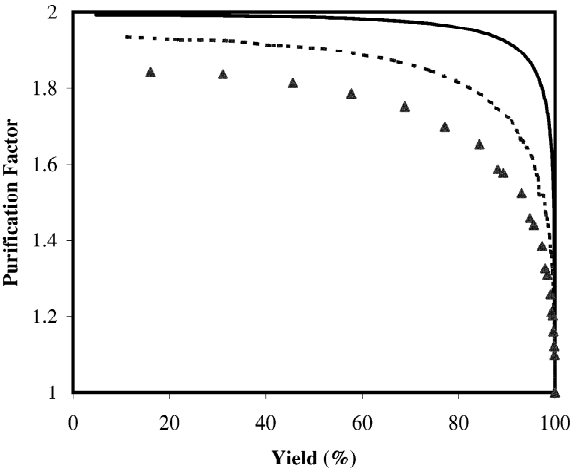

Fig. 6. Purification factor versus yield diagram corresponding to Fig. 5 for the simulated gel filtration separation of ovalbumin from a mixture of γ-globulin and Ribonuclease A at different flow-rates (solid line, 0.1 ml min^{-1}; broken line, 0.3 ml min^{-1}; Δ, 0.5 ml min^{-1}).

Since the initial purity of the sample load for all simulations was 50%, the maximum purification factor of this separation is 2. However, only by operating at 0.1 ml min^{-1} does the simulated maximum purification factor approach this theoretical limit. This is mainly because band-overlapping is more significant as the flow-rate increases and this is reflected in the separation becoming more difficult. For a flow-rate of 0.1 ml min^{-1}, the degree of peak overlapping is insignificant and there is almost a baseline separation for the three components between the times of 770 and 810 min. Therefore, the

purification factor–yield curve achieves the maximum purification factor of the ideal system in this period. However, when considering this result it must also be borne in mind that it is often not economical to operate a separation under such a low mobile phase velocity due to the longer elution time. Selected operating strategies would probably reflect this trade-off.

To demonstrate how the data generated could be used to select appropriate sample collection times corresponding to different values of purification factors, we took the data evaluated for a fixed mobile flow-rate of 0.3 ml min^{-1}. Table 3 tabulates typical combinations of performance data for the system. With the aid of such a table the operator can select at which point to commence product collection to achieve the desired value of purification factor or product yield. Process trade-offs are unambiguously presented and very precise control of the chromatogram output can be realised.

4.2. Experimental results

Data for the concentration of endotoxin and plasmid DNA in each of the chromatograms generated from the ion-exchange endotoxin clearance column are presented in Fig. 7 for the four different fermentation batches. The mass of endotoxin (in terms of endotoxin unit, EU) and plasmid DNA (in terms of mg) corresponding to each fraction collected were determined, respectively. These data were processed to produce a modified fractionation diagram (Fig. 8) that shows the relative change in mass of endotoxin and plasmid DNA. In Fig. 8 then

Table 3
Calculated sample collection times (t_1 and t_2) corresponding to different values of yield and purification factor for the gel filtration system, simulated at a buffer flow-rate of 0.3 ml min^{-1}

Purification factor	Yield	t_1 (min)	t_2 (min)
1.95	11.1	261.7	264.3
1.90	51.7	256.5	269.5
1.85	74.2	251.3	272.1
1.80	79.4	251.3	274.7
1.60	95.2	243.5	279.9
1.40	98.9	238.3	285.2
1.20	99.8	230.5	290.3
1.00	100.0	196.6	334.6

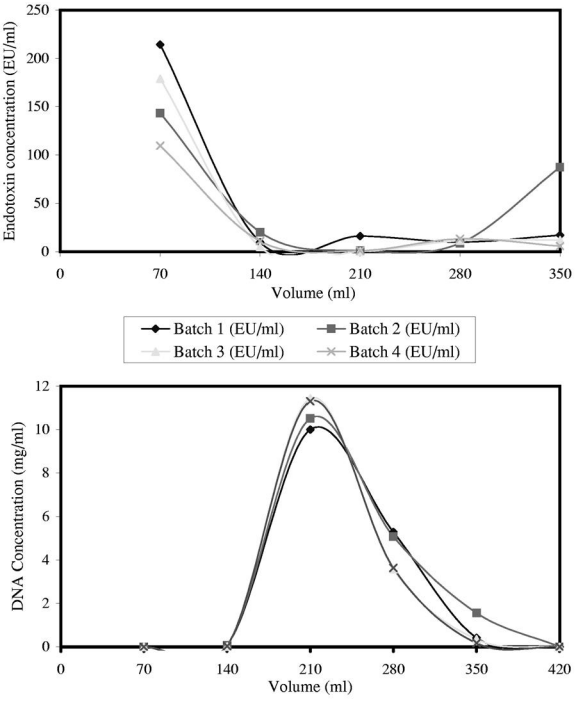

Fig. 7. Fraction data for the amount of (top panel) endotoxin and (bottom panel) plasmid DNA collected from an ion-exchange chromatographic separation of endotoxin from DNA product. The batch numbers refer to separate fermentation batches where the product has been processed by a common purification sequence. Data from Biopharmaceutical Product Development, Glaxo-SmithKline, Kent, UK.

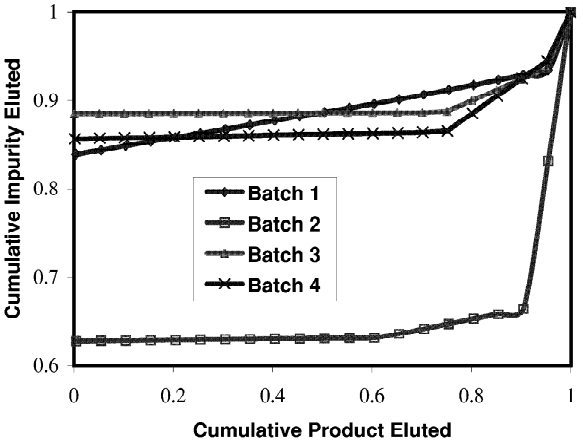

Fig. 8. Modified fractionation diagrams for the endotoxin clearance column data, showing the effect of the different fermentation batches.

the axes plot the cumulative fractional mass of plasmid DNA and the corresponding fractional amount of endotoxin eluted, respectively. For each set of data, a fractionation line was obtained by employing linear-line-fitting to generate a continuous function. The fractionation lines do not pass through the origin, since only endotoxin was eluted in a significant amount in the initial portion of the chromatogram. So for batch 2, ~63% of the total endotoxin collected appears in the first fraction, whilst for the other batches ~85±3% elutes in the first fraction. For all but batch 1, the fractionation plots are fairly flat between 0 and 0.6 cumulative DNA. This indicates that after the first fraction no significant amount of endotoxin was eluted while the amount of DNA collected increased from 0 to 60%. For batch 1, a gradually increasing tangent shows the constant co-elution of endotoxin and DNA. The large amount of endotoxin found in the last fractions for all batches gives rise to the abrupt increase in slope of the fractionation curves at ~0.9 on the cumulative DNA axis.

From a regulatory viewpoint it is crucial to be able to determine a suitable operational trade-off between achieving a desired minimum level of contaminant removal whilst realising an acceptable product yield. This feature is explored in Fig. 9 which shows the minimum contamination index against product yield. A typical specification for a DNA product might be related to the endotoxin level. Whilst the precise amount that may be tolerated is dose-specific, a limit of 1 EU mg^{-1} DNA was taken as the target for the study. Batch 1 gives an endotoxin level higher than the specification and much greater than the other three batches. Batches 2, 3 and 4 display constant impurity levels for product yields less than 60% with batch 3 achieving nearly an order of magnitude greater contaminant removal in this interval. For the principal region of interest, i.e. product yields which are greater than 90%, the changes of contamination index are abrupt. Of the three batches which satisfy the <1 EU mg^{-1} criteria in this region, batches 3 and 4 out-perform batch 2 by realising an in-specification product at 95% yield. Data for batch 2 show that this will be out of specification at a yield of $>92\%$ and that further steps to reduce the endotoxin level will be needed if higher yields are sought. The figure therefore provides an analysis of the process

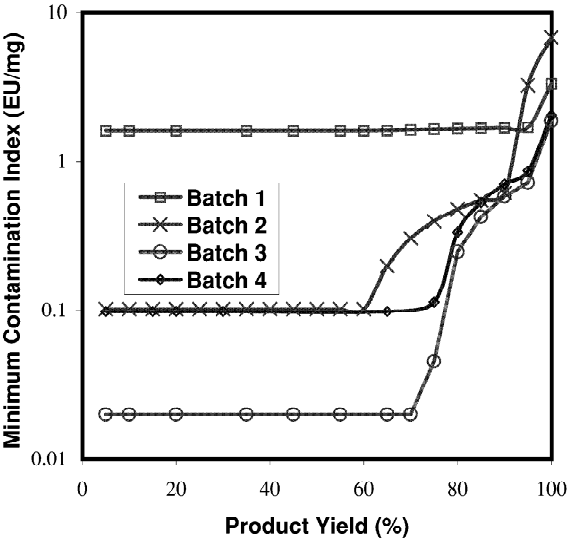

Fig. 9. Minimum contamination index versus product yield for endotoxin clearance column data given in Figs. 7 and 8 showing the effect of different fermentation batches on the separation performance.

trade-offs and an indication of the impact on further processing of the cut-point decisions that are made.

Finally we selected to report the different sample collection retention volumes corresponding to the required product yield or contamination index for batch 3 product (Table 4). For a known value of fraction volume and flow-rate of the chromatographic run, the sampling points can be re-defined on the chromatogram. The results show how the volume of material to be collected increases with the desired yield but at the expense of the degree of contamination.

Table 4
Calculated sample collection points (Cut 1 and Cut 2) corresponding to different values of yield and contamination index achieved in an ion-exchange endotoxin clearance column

Yield (%)	Contamination index (EU/mg DNA)	Cut 1 (ml)	Cut 2 (ml)
40	0.020	172	210
60	0.020	154	212
80	0.262	138	227
90	0.593	138	257
95	0.733	138	272
99	1.138	138	322

Data are for batch 3 fermentation product.

5. Conclusion

A simple engineering framework that enables the quick and straightforward quantification of the performance of liquid chromatographic separations has been proposed in this paper. The combination of fractionation diagrams, to capture the elution chromatographic data, and the resultant maximum purification factor against yield and minimum contamination index versus yield curves have been shown to represent the performance changes occurring during chromatographic separation. Between them the plots facilitate an examination of the desired degree of trade-off between purity and yield corresponding to any set of operating conditions. The approach has been applied to two sets of chromatographic data, and shown to provide an easily-interpretable framework for process decision-making and control. With the help of such an approach, the optimal operating conditions, in terms of acceptable yield, desired purity and level of contamination, can be chosen easily and the product fractions required to achieve these calculated.

6. Nomenclature

C	Concentration of solute in the mobile phase (mg ml^{-1})
i	Number of time steps from $t=0$
M	Amount of material (mg)
N	Number of time steps
PF	Purification factor
t	Time (s)
x	Initial purity of the sample
X	Fractional mass of material eluted
Y	Fractional mass of product eluted
z	Axial position along the column (cm)

Subscripts

A	Impurity component A
B	Impurity component B
max	Maximum
O	Overall product
P	Product component (in each fraction)
S	Overall sample
T	Sample (in each fraction)
(1)	Point when sample collection starts
(2)	Point when sample collection ends

Acknowledgements

Financial support from both the University of London Academic Trust and UK Overseas Research Scheme for S.H. Ngiam are gratefully acknowledged, as is the research funding provided by the Biotechnology and Biological Science Research Council for The Advanced Centre for Biochemical Engineering.

References

[1] G. Walsh, H. Dennis, in: Protein Biotechnology, Wiley, Chichester, UK, 1996, p. 99.
[2] G. Walsh, Biopharmaceuticals: Biochemistry and Biotechnology, Wiley, Chichester, UK, 1998.
[3] S. Golshan-Shirazi, G. Guiochon, J. Chromatogr. 603 (1992) 1.
[4] G. Guiochon, S. Golshan-Shirazi, A. Katti, Fundamental of Preparative and Non-Linear Chromatography, Academic Press, Boston, MA, USA, 1994.
[5] T. Gu, Mathematical Modeling and Scale-Up of Liquid Chromatography, Springer, Berlin, 1995.
[6] P. Richardson, M. Hoare, P. Dunnill, Biotechnol. Bioeng. 36 (Vol. 4) (1990) 354.
[7] P. Richardson, M. Hoare, P. Dunnill, Chem. Eng. Res. Des. 67 (1989) 273.
[8] R. van Reis, S. Saksena, J. Membr. Sci. 129 (1997) 19.
[9] P.M. Boyer, J.T. Hsu, AIChE J. 38 (Suppl.) (1992) 259.
[10] T.L. Chen, J.T. Hsu, AIChE J. 35 (Vol. 2) (1989) 332.
[11] J.T. Hsu, J.S. Dranoff, Comput. Chem. Eng. 11 (Vol. 2) (1987) 101.
[12] J.T. Hsu, U.P. Ernst, Chem. Eng. Sci. 45 (Suppl. 4) (1990) 1017.
[13] G.A. Soriano, Ph.D. Thesis, University College London, London, UK, 1995.
[14] MATLAB, version 6, release 12, The MathWorks Inc., Natick, MA, USA, 2000.