Process Control for Flexible Pooling in Chromatographic Purification

Karolina Johansson

Department of Chemical Engineering, Lund University, Sweden

June 2011

Abstract

As a master's thesis in chemical engineering, process control for flexible pooling in a chromatographic process has been studied. The process is an RPC step for removal of primarily HCP from a therapeutic protein produced by Pfizer Health AB. From experimental data, a model was developed and used for sensitivity analyses. The parameters studied were column load, feed concentration and HCP to target protein mass ratio in the feed. A sensitivity analysis, giving the optimal pooling limits for maximum yield without violation of the purity limits, was performed. The results showed that front pooling should be performed at the base line but for safety, a low fixed limit was chosen. With this, another sensitivity analysis was performed and functions of one or two of the parameters were fitted to the results. Finally, a control algorithm for back pooling based on column load was chosen. With the control algorithm, the minimum yield was increased by 15%. Since the experimental data used was not optimal for calibration, there are model uncertainties which propagate to the control algorithm. Consequently, further studies are required.

Keywords: Preparative chromatography, Mechanistic modeling, Calibration, Simulation, Sensitivity analysis, Process control, Quality by design

Introduction

As stated in the *ICH Harmonised Tripartite Guideline* on pharmaceutical development (Q8), the concept of quality by design (QbD) is recommended for development and improvement of processes for production of pharmaceuticals. The main objective of QbD is a pro-active approach by which a high quality product is guaranteed, as opposed to the traditional quality by testing (QbT) where the product is discarded if the quality demands are not met in the final tests.

QbD requires good knowledge of the process in question and an insight into the underlying mechanisms. It is also important to investigate the influence of different process parameters, including possible interaction effects, on the product quality. Based on such knowledge, measures can be taken to avoid unfavorable running conditions. One method is to implement some kind of process control. [1]

As a small step towards QbD, a control algorithm for flexible pooling in a chromatographic purification step at Pfizer Health AB has been developed.

Based on experimental data provided by Pfizer Health AB, a mechanistic model of the purification step was developed. A number of sensitivity analyses were performed using simulation data and starting from the results, a pooling control algorithm, including the parameters found relevant, was developed.

Theory

In this section, a brief theoretical background on preparative chromatography and sensitivity analysis is given.

Preparative Chromatography

Chromatography can be used for one of two main purposes; analysis or separation. In the latter case it is generally referred to as preparative chromatography. Separation is achieved when the components in the mobile phase, the solutes, are more or less strongly adsorbed on the stationary phase. In preparative chromatography, the mobile phase is a liquid solvent and the stationary phase consists of either particles or a monolith, with functional groups giving the desired adsorption properties. [2]

Chromatographic processes are often classified by the type of interactions causing adsorption. In reversed phase chromatography (RPC), the stationary phase is non-polar and the mobile phase is polar. Adsorption occurs when hydrophobic patches on the solutes are bound to the stationary phase. The adsorption equilibrium is often described by an isotherm such as the Langmuir isotherm (eq. 1). [3]

$$q = \frac{q_{max} \cdot K \cdot c}{1 + K \cdot c} \tag{1}$$

q is the concentration of adsorbates in the stationary phase, q_{max} is the maximum adsorbate concentration, c is the solute concentration and K is the adsorption equilibrium constant. [3]

To achieve separation, the components must be desorbed and leave the column, i.e. be eluted. This generally occurs naturally in time during isocratic elution, but measures are often taken to shorten the run time. [2] During gradient elution, the mobile phase properties are altered towards those of the stationary phase, promoting desorption. This is accomplished by changing the concentration of the modifier which in RPC is a non-polar solvent. The change can be made as a step or gradually, in the latter case often linearly. [3]

During elution, the liquid leaving the column is collected in pools, divided by limits in units of for example volume or absorbance. Pooling should give one pool with each target component and pools with impurities which are discarded. In industrial applications, peaks commonly overlap, making pooling a compromise between yield and purity. [2]

Modeling and Simulation of Chromatography

There are various models of chromatography and the choice of model depends on the resources available and the accuracy required. For this project, a lumped rate model called the reaction dispersive model was chosen (eq. 2). It includes convective and dispersive transport of the solutes and adsorption kinetics but mass transfer resistance is neglected. The stationary phase particles are assumed to be uniform and spherical with pores which all solutes can access fully. Radial dispersion is neglected leaving one spatial variable, due to symmetry. [2]

$$\frac{\partial c_i}{\partial t} = -v \cdot \frac{\partial c_i}{\partial x} + D_{ax} \cdot \frac{\partial^2 c_i}{\partial x^2} - \frac{(1-\varepsilon_c)}{\varepsilon_c} \cdot \psi \tag{2}$$

c_i is the concentration of solute i, t is time, v is the mobile phase linear velocity, x is the axial position in the column, D_{ax} is the axial dispersion coefficient, ε_c is the column porosity and ψ is the adsorption rate. [2] In this case a more advanced adsorption rate formula, the Langmuir MPM (mobile phase modulators) model including the effect of the modifier, was used. The version for RPC and the reaction dispersive model is given by eq. 3. [4]

$$\psi = k_{kin,i} \cdot \left(H_i \cdot e^{\gamma_i \cdot (s - s_{ref,i})} \cdot c_i \cdot \left(1 - \sum_j \frac{q_j}{q_{max,j}}\right) - q_i \right) \tag{3}$$

k_{kin} is the adsorption rate constant, H is Henry's constant, γ is the hydrophobicity and s is the modifier concentration while s_{ref} is a reference concentration of the same. [4]

Since the model is given by a number of partial differential equations (PDEs), discretization is required to enable simulation. The column is divided into grid points along the axis and the PDEs are transformed into ordinary differential equations (ODEs). Discretization of the function *f* with respect to the variable *x* is described by eq. 4.

$$\frac{\partial f}{\partial x} = \frac{f_{x+\Delta x} - f_x}{\Delta x} \quad (4)$$

A finer grid increases the simulation accuracy but requires a longer computational time. [2]

All simulations for this project have been performed with the MATLAB based simulation tool *pcs* (Preparative Chromatography Simulator), developed at the Department of Chemical Engineering. [4]

Sensitivity Analysis

The main purpose of a sensitivity analysis is to investigate the effect of changes in the input parameters on the process outcome. In some cases, the relative importance of the parameters is studied. There are many different methods for sensitivity analysis, varying in features and complexity. A common classification is based on how the parameter values are chosen and divide the methods into mathematical, using chosen parameter values, and statistical, using random sampling. [5]

For the sensitivity analyses in this study, the Monte Carlo method has been used. It is statistical and involves three steps. Firstly, intervals of allowed values for the parameters are chosen. Secondly, sets of parameter values within the intervals are sampled according to the chosen probability distribution. The sampling can be completely random or more controlled, for example using Latin hypercube sampling (LHS). LHS follows the distribution applied but ensures unique and more evenly distributed values. Finally, simulations are performed for each set of parameter values. [6]

Methods

The process studied is an RPC step for purification of a therapeutic protein and removal of primarily HCP (host cell proteins). Due to confidentiality issues, the target protein is called component A and the product related impurities are referred to as components B-D. Additionally, values of variables such as column load and pooling limits have been re-scaled while some plots lack scale.

The Experimental Data

All experimental data originates from a DoE study previously performed at Pfizer Health AB. The experiments were performed in lab-scale but at conditions corresponding to the full-scale process. The stationary phase consists of non-polar particles and the modifier is a mix of two organic solvents. Each cycle has four steps; regeneration, load, wash and elution. During load and wash, the mobile phase is water with a negligible addition of buffer. Elution is achieved with a linear gradient and pooling is performed according to fixed absorbance limits.

In the DoE study, column load (mg comp A/ml resin), temperature and pooling limits have been varied. Except the set point values, one lower and one higher value given by an equal decrease or increase for either parameter has been tested (Table 1).

Table 1. The variations in column load, temperature, front and back pooling given as percent relative to the set point values.

Parameter	**Load**	**Temp**	**Front**	**Back**
Δ (%)	31.4	27.3	38.5	17.6

The samples for the DoE study already fulfilled the demands except for the HCP content which was nearly 34 times too high.

Model Development

Most of the process data required for the modeling was supplied by Pfizer Health AB but the porosities and the adsorption

parameters had to be determined by calculations and calibration, respectively, using the experimental data. Since the samples contained a small amount of salt, the conductivity of the eluted liquid changed abruptly at the end of the load step. The delay compared to when the wash was initiated was used to determine the total porosity. Any dead volume was neglected but the data available was still insufficient for calculation of both column and particle porosity. Based on guidelines, a column porosity of 35% was assumed and the other one could be calculated. Calibration was performed with a least-squares method and for each experiment, the UV curve and the pool concentrations for either component were included.

As seen in eq. 3, the absorbance coefficients to determine were H, γ, s_{ref}, k_{kin} and q_{max}. With the amount of experimental data available, it was unreasonable to calibrate five parameters for each of the five components. Therefore, H was set to 1 and q_{max} was assumed to be equal for all components. Individual values of γ and s_{ref} were calibrated while k_{kin} was assumed to be equal for components A-D. Since the UV absorbance was unknown, qualified guessing was applied.

Initially, attempts were made to use data from all the experiments for calibration. Since the results were unsatisfactory, calibration for solely the set point temperature was tried and resulted in an acceptable fit.

Sensitivity Analyses

By request from Pfizer Health AB, the effects of the following three parameters were investigated in the sensitivity analyses:

- Column load (mg comp A/ml resin)
- Feed concentration (mg comp A/ml feed)
- HCP content in the feed (ng HCP/mg comp A)

200 combinations of parameter values were generated for subsequent simulation. LHS with a normal distribution based on the re-scaled historical values in Table 2 was applied.

Table 2. Mean and standard deviation for each parameter tested in the sensitivity analyses.

Parameter	Mean	Std. Dev.
Column Load	1	0.120
Concentration	1	0.103
HCP Mass Ratio	1	0.291

Two different approaches were tried; one giving the product composition and the yield with fixed pooling limits and one giving the optimal pooling limits for maximum yield without violating the purity limits. For the approach with fixed pooling limits, both the original limits and new ones suggested in the DoE study were tested. The suggested pooling limits are 51.1% lower at the front and 15.6% higher at the back, compared to the original ones. The same sets of parameter values were used for all analyses.

Development of a Control Algorithm

According to the sensitivity analyses, the front pooling limit should be fixed at the base line since the HCP peak is eluted last. As the composition of the HCP is unknown, it has been modeled as one single component while actually it is a mix of proteins of varying size and character. Thus, in reality there might be several HCP peaks and for safety, the front pooling limit was set to the value suggested in the DoE study. With the front pooling fixed, a new sensitivity analysis was performed to find the optimal back pooling limit maximizing yield while keeping the impurity levels within the limits. In this analysis, a 30% safety margin was added to the purity limit for the HCP. To include combinations of more extreme parameter values, a rectangular distribution given by the minimum and maximum values from the former analyses was applied (Table 3).

Table 3. The minimum and maximum values for the rectangular distribution.

Parameter	Min	Max
Column Load	0.663	1.386
Concentration	0.670	1.414
HCP mass ratio	0.116	1.860

With the rectangular distribution, 200 sets of parameter values were generated by LHS and simulations were performed for each of them. Different functions of one or two of the parameters were then fitted to the results for the back pooling by optimization with a least-squares method. A constraint was added to ensure that the back pooling limit given by the function would give the required purity or better for all parameter values. Each control algorithm was evaluated by a sensitivity analysis using the parameter value sets based on the normal distribution. Front pooling was performed at the suggested limit and back pooling according to the control algorithm.

Results and Discussion

In this section, the relevant results from the model development, via the sensitivity analyses, to the final control algorithm are summarized.

Model Development

The porosities were determined to 66.3% for the particles and 78.1% in total with the column porosity set to 35.0%.

Regarding the calibration, the resulting values for the adsorption coefficients at the set point temperature are given in Table 4.

Table 4. Final calibrated values for the adsorption coefficients at the set point temperature. The units are m^3/mol for γ, mol/m^3 for s_{ref} and s^{-1} for k_{kin}.

Comp	**A**	**B**	**C**	**D**	**HCP**
$\boldsymbol{\gamma \cdot 10^4}$	-5.89	-7.92	-5.25	-6.66	-12.36
$\boldsymbol{s_{ref} \cdot 10^{-3}}$	7.02	5.48	7.18	6.32	7.24
$\boldsymbol{k_{kin} \cdot 10^3}$				3.34	3.69

As seen, q_{max} is not specified in Table 4. Since the initial calibrations gave unreasonably large variations in this value, a sensitivity analysis for q_{max} was performed. With no change in the peaks at values above 10^3 mol/m^3, q_{max} was fixed at that value.

UV curves based on experimental and simulation data, respectively, are plotted together in Figure 1 for comparison.

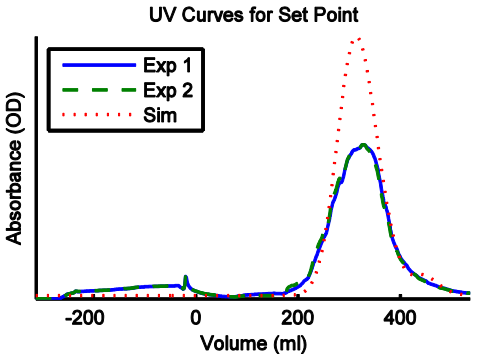

Figure 1. Experimental and simulation data for the set point conditions.

The peak based on simulation data is obviously too high but since the pooling is performed within the area where the experimental and simulated peaks overlap, the height is unimportant for this application.

Sensitivity Analyses

The two sensitivity analyses using fixed pooling limits gave similar results with negligible variations in the mass ratios for all components except HCP. In Figure 2, the dependence of HCP mass ratio and yield on either parameter is shown. These results apply for the original fixed pooling limits.

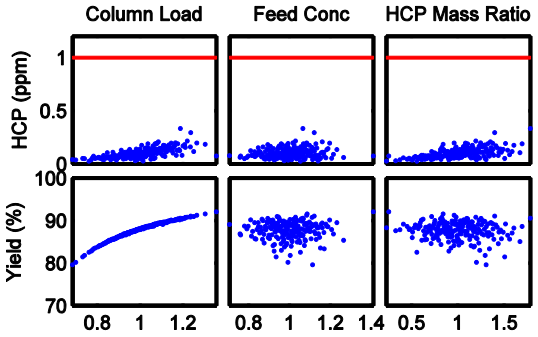

Figure 2. Results from the sensitivity analysis with the original fixed pooling limits.

For the suggested limits, the HCP content is lower, the yield is higher and the ranges of both are narrower. With a larger portion of the front of the peak included, this is expected. The analysis with flexible pooling limits gave an optimal front pooling limit at the base line for all parameter values. Dependence both on the column load and on the HCP mass ratio, but not on the feed concentration, was found for the back pooling limit. In general, the

results were very similar to those in Figure 3. A comparison of the intervals of the yield obtained for each sensitivity analysis is found in Table 5.

Table 5. Statistics for the yield obtained in the sensitivity analysis for either set of pooling limits.

Limits	Yield (%)		
	Min	Mean	Max
Original	79.6	87.7	92.0
Suggested	88.1	92.3	94.8
Flexible	97.9	98.5	99.4

The results with the flexible limits are impressive but this kind of pooling is not feasible in reality. Firstly, there are no safety margins and secondly, it would require a perfect correlation between the parameter values and the optimal pooling limit. Regarding the results for fixed pooling, the improvement achieved with the suggested limits is clear but an additional increase might be possible with a control algorithm.

Development of a Control Algorithm

Figure 3 shows the results from the sensitivity analysis using front pooling at the suggested fixed limit and flexible back pooling. As expected, the results showed that the HCP mass ratio was at or just below the purity limit.

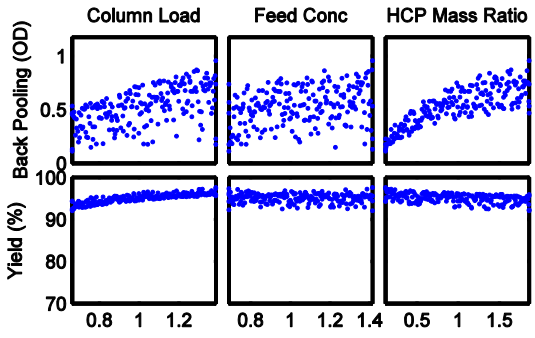

Figure 3. Results from the sensitivity analysis for control algorithm development.

According to Figure 3, column load and HCP mass ratio affect the back pooling but since the latter cannot be measured on-line, a control algorithm including only the former was developed (eq. 5).

$$OD_{back}(l) = 0.060 \cdot l + 0.247 \quad (5)$$

OD_{back} is the pooling limit and l is the column load. In Figure 4, the line given by eq. 5 is plotted together with the data points on which it is based.

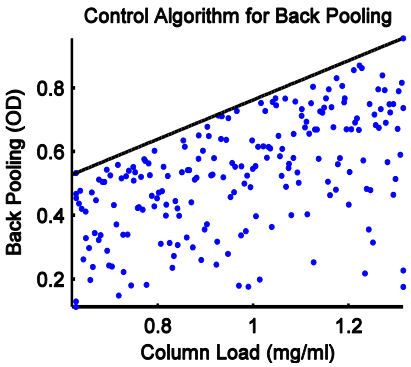

Figure 4. The control algorithm for back pooling, based on column load, plotted together with the data points from the sensitivity analysis.

As seen in Figure 4, the fit for the higher pooling limits is good but the distance between many of the data points and the line is large. The sensitivity analysis performed with this control algorithm showed a gap between the HCP purity limit and the data points, accompanied by a less than optimal yield. With a view to a better fit, control algorithms including either of the two other parameters were developed and evaluated. A slightly better fit was achieved but neither that improvement nor the increase in yield, presented in Table 6, was enough to motivate a two-parameter control algorithm. Thus, the control algorithm described by eq. 5 was found to be the best result achievable under the circumstances.

Table 6. Statistics for the yield obtained in the sensitivity analysis for either control algorithm.

Parameters Included	Yield (%)		
	Min	Mean	Max
Load	92.1	94.1	95.3
Load and HCP	92.2	94.7	96.6
Load and Conc	92.0	94.3	95.4

Conclusions

With a control algorithm based on column load, the yield was increased and the process robustness was improved. Compared to pooling with fixed limits, the minimum yield was increased by 16% relative to the original limits and by 5% relative to the suggested ones. It could also be concluded that inclusion of a second parameter in the control algorithm would only give a negligible improvement.

Due to the uncertainties in the calibration results, caused by the low amount of experimental data available, further work is required to produce a reliable pooling control algorithm. With data from experiments performed at the set point temperature but for varying column load, feed concentration, HCP mass ratio and gradient, a better calibration would be possible. With the improved calibration, a new and better control algorithm can be developed.

References

1. ICH. *ICH Harmonized Tripartite Guideline: Pharmaceutical Development Q8 (R2)*. 2009.
2. Schmidt-Traub, H., *Preparative Chromatography of Fine Chemicals and Pharmaceutical Agents*. 2005, Weinheim, Germany: Wiley-VCH.
3. Carta, G., Jungbauer, A., *Protein Chromatography: Process Development and Scale-Up*. 2010, Weinheim, Germany: Wiley-VCH Verlag Gmbh.
4. Borg, N., Andersson, N., Max-Hansen, M., Westerberg, K., Nilsson, B., *Preparative Chromatography Simulation Platform*, in *SPICA 2010*, 2010: Stockholm, Sweden.
5. Frey, H.C., Patil, S. R., *Identification and Review of Sensitivity Analysis Methods*. Risk Anal., 2002. **22**(3): p. 553-578.
6. Sin, G., Gernaey, K. V., Eliasson Lantz, A., *Good modeling practice for PAT applications: Propagation of input uncertainty and sensitivity analysis*. Biotechnol. Progr., 2009. **25**(4): p. 1043-1053.