Calibration and Optimization of Hydrophobic Interaction Chromatography

Alex Olsson

Department of Chemical Engineering, Lund University, Sweden

2012-11-01

Abstract

One of the most commonly used method of choice when separating molecules are ion-exchange chromatography and this has resulted in more extensive studies on this method. The aim of this article is to calibrate and optimize the separation of insulin variants using hydrophobic interaction chromatography by mechanistic models. This thesis focuses on calibration and optimization of HIC experimental data. For the optimization part, two objection functions were minimized. These functions are yield and productivity by varying the loading time together with the salt gradient. Further studies were also made to the optimization step with different types of salt gradients. These gradients consist of an isocratic, step and a linear gradient and then to find the most optimal way to separate insulin on HIC.

Keywords: Hydrophobic Interaction Chromatography, Optimization, Calibration

Introduction

The regulatory agencies such as the US Food and Drug Administration, demand high quality production of pharmaceuticals and at the same time low cost drugs, which puts a pressure on a well-defined mathematical model used in the simulation of the separation. HIC is based on separation by the different hydrophobicity of molecules and it needs a high salt concentration in the mobile phase to prevent self-association and to adsorb the molecules to the column. The column is packed with small particle which interact with the target protein. To elute the protein, the salt concentration needs to be lowered by either a step or a linear gradient.

The simulation consists of a system of partial differential equation which cannot be solved analytically and needs to be changed to ordinary differential equation by using finite volume methods which makes this kind of procedure very heavy to calculate on a computer.

In this master thesis the study is focused on calibrating and optimizing the separation of insulin variants by HIC. The data was received from Novo Nordisk and the calibration together with the optimization was simulated with MATLAB. The data received was run with an isocratic salt gradient and the parameters that were going to be calibrated were the hydrophobic characteristics, the adsorption rate coefficient and Henry's constants.

Theory

The hydrophobic interaction chromatography (HIC) is a preparative chromatography process which means that it is used to purify substances, such as pharmaceutical drugs. HIC is usually used to separate peptides, proteins and other biological molecules and are based on the

hydrophobicity of a molecule [1]. To get the molecules to interact with the ligands, a high salt concentration is needed to prevent the proteins to self-associate. To elute the protein, the most common case is to lower the salt concentration by either using a step or a linear salt gradient. There are other methods that can be used [1], which are addition of organic solvents that changes the polarity of the solvent, the use of chaotropic salts or adding detergent to disrupt the hydrophobic interaction.

Models

The column model is a reaction dispersive model and is described as the following partial differential equation of a component i [2]:

$$\frac{\delta c_i}{dt} = D_{ax} \cdot \frac{\delta^2 c_i}{dx^2} - \nu_{int} \cdot \frac{\delta c_i}{dx} - \frac{1-e_c}{e_c} \cdot \frac{\delta q_i}{dt} \qquad (1)$$

where D_{ax} is the dispersion coefficient (m^2/s), ν_{int} is the interstitial velocity (m/s), ε_c is the void fraction in the packed bed (-), c_i is the concentration of component i in the mobile phase (mol/m^3), q_i is the concentration of component i in the stationary phase (mol/m^3), x is the axial coordinate from the inlet of the column along the axis (m) and t is the time (s).

The reaction dispersive model requires boundary condition and for the inlet, the boundary condition is set to a Robin condition [2]:

$$\frac{\delta c_i}{dx} = \frac{\nu_{int}}{D_{ax}}(c_i - c_{inlet}) \text{ at } x = 0 \qquad (2)$$

where c_{inlet} is the inlet concentration. The boundary condition for the outlet is considered a Neumann condition [2]:

$$\frac{\delta c_i}{dx} = 0 \text{ at } x = L \qquad (3)$$

where L is the length of the column (m)

The isotherm is set to a Langmuir MPM kinetic model [2]:

$$\frac{\delta q_i}{dt} = k_{kin,i}\left(H_i \cdot c_{x,i}\left(1 - \sum_{j=1}^{n} \frac{q_j}{q_{max,j}}\right) - q_i\right) \qquad (4)$$

where $k_{kin,i}$ is the alternative adsorption rate coefficient, q_{max} is the maximum concentration in the stationary phase, H_i is henry's constant and are described as the following equation [3]:

$$H_i = H_{i,0} \cdot e^{\gamma_i \cdot S} \qquad (5)$$

γ_i is the hydrophobic characteristics for each component and S is the salt concentration.

Methods

The data that was given was run isocraticly at different buffer concentrations. The different non-salt buffer concentration was 60, 65, 70, 75, 80, 85, 90 and 100%. The sample that was used contained three types of insulin where the focus was to separate the second peak (figure 1).

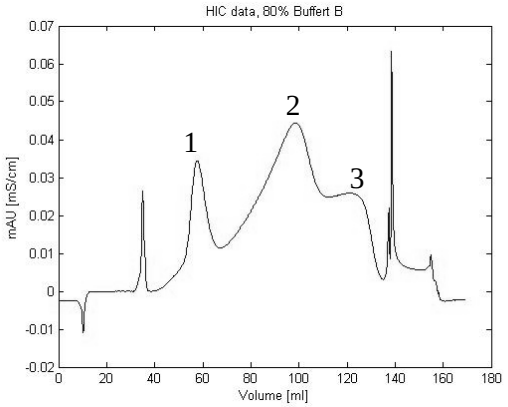

Figure 1. One of the chromatogram of the experiment at a non-salt buffer concentration of 80%. The target protein is the second peak and the impurities are at peak 1 and 3.

The other data was a salt breakthrough curve; this is to calculate the dead volume of the system. The breakthrough curve is shown in figure 2:

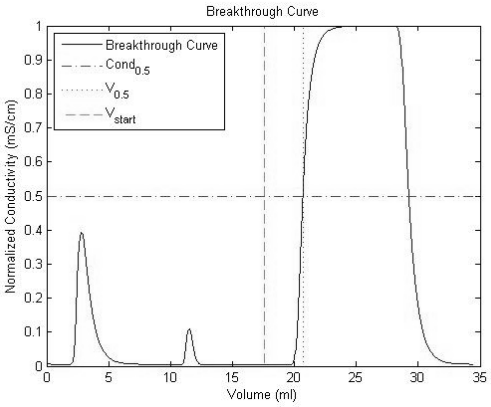

Figure 2. Breakthrough curve of the system at a flow of 0.7 ml/min. This curve is required to calculate the dead volume of the system.

The parameters that are being calculated are hydrophobic characteristics, the adsorption rate coefficient and Henry's constants. This is done by using the following equation [4]:

$$V_R = V_{Dead} + V_{NA} + f \cdot H_0 \cdot e^{\gamma \cdot S} \quad (6)$$

where V_R is the retention volume of every component at the different salt concentration and V_{Dead} is the dead volume. V_{NA} is the non-adsorbed volume which is calculated as follows:

$$V_{NA} = V_{col} \cdot \left(\varepsilon_c + (1 - \varepsilon_c) \cdot \varepsilon_p\right) \quad (7)$$

Equation 6 is only calibrating γ and H_0 but not $k_{kin,i}$. $k_{kin,i}$ is calibrated by adjust the parameters for every component to fit the experimental chromatogram.

Equation 1 cannot be solved analytical since it is a partial differential equation (PDE) and has to be simulated by using finite volume methods, discretization. This method makes the PDE to act as an ordinary differential equation by generating grid points to form a mesh which is used to approximate the equations values.

For the optimization of the process, MATLAB was used together with a simplex method, *fminsearch.* The simplex method minimizes a given objective function by inserting an initial guess value. From this initial value, this method generates two other points and forms a triangle [5]. This triangle will continue to move towards the local minimum and finally enclose and find the minimum value. The chosen objective functions are productivity and yield because they work opposite each other [6].

$$Yield = \frac{\int_{cp1}^{cp2} C_{ut} dt}{C_L \cdot V_L} \quad (8)$$

$$Productivity = \frac{C_L \cdot V_L \cdot Yield}{t_{cycle}} \quad (9)$$

where C_L and the V_L is the loading concentration respective volume of the sample. The $cp1$ and $cp2$ are calculated with a fractionizer called *simplefrac*. These cutpoints are time interval in which a set purity of 95% is given to the target protein. In this case t_{cycle} is set to $cp2$.

Results and Discussion

The calibration was made using the equation 6 by rewriting and by taking the logarithm of it to give it a linear relation:

$$\ln(V_R - V_{Dead} - V_{NA}) = \ln(f \cdot H_0) + \gamma \cdot S \quad (10)$$

V_{Dead} was calculated by using the breakthrough curve (figure 2) with this equation:

$$V_{dead} = V_{0.5} - V_{start} \quad (11)$$

The retention volume, non-adsorbed volume, dead volume, $\ln(V_R - V_{Dead} - V_{NA})$, is plotted against the salt concentration and the following graph is plotted:

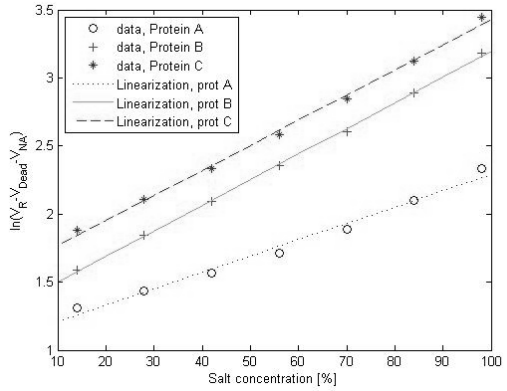

Figure 3. The experimental points for each of the protein A, B and C is plotted against the salt concentration together with the linearization of them.

As seen in figure 3, the experimental values are a good fit to the linear regression and the average deviation for both protein B and C are 0.042 respective 0.086. However, protein A has a larger deviation of 0.11 and has a slightly other character than a straight line. One issue can be that a slightly wrong retention volume has been extracted from the experimental data.

The linear regression from figure 3 gives the following values:

Table 1. The values from the linear regression $y = kx + m$ for every protein A, B and C.

Protein	k	m
A	1.2050	1.0884
B	1.8820	1.3121
C	1.8454	1.5830

By using the linear relation in equation 10, one can calculate the values for γ and H_0. The calculated parameters are presented in the table below.

Table 2. The hydrophobic characteristics and Henry's constant

Protein	γ	H_0
A	1.2050	1.9798
B	1.8820	2.4759
C	1.8454	3.2465

The final parameters that was calibrated was $k_{kin,i}$. This was done manually where the parameters were adjusted to fit the experimental chromatogram. These values for each of the rate constants for protein A, B and C are 0.1, 0.007 and 0.05 s^{-1}.

Next step is to optimize the separation by minimize the objective function 8 and 9. The parameters that are going to change to find the optimum values are loading volume, salt concentration (isocratic, step salt gradient) and end time (linear salt gradient).

The first gradient to optimize is the isocratic and the surface plot it is simulated with a salt concentration ranging from 5-50% and a loading time of 100 to 2500 seconds.

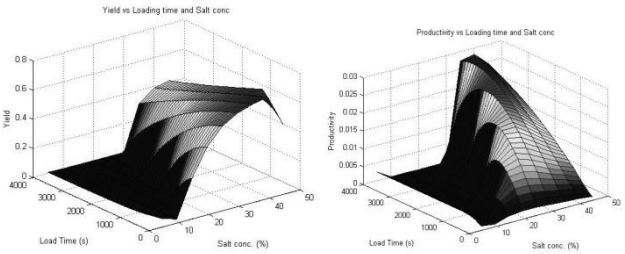

Figure 4. Surf plots of the yield (left) and productivity (right) as a function of loading time and salt concentration.

As seen in figure 4 the optimum value for yield is at a low loading time together with a high salt concentration. The productivity favours a high loading time and a high Salt concentration.

The second gradient is a linear gradient and the 3D plot is simulated with a loading time between 1 to 8000 seconds and an end time ranging from 8500 to 20000 seconds

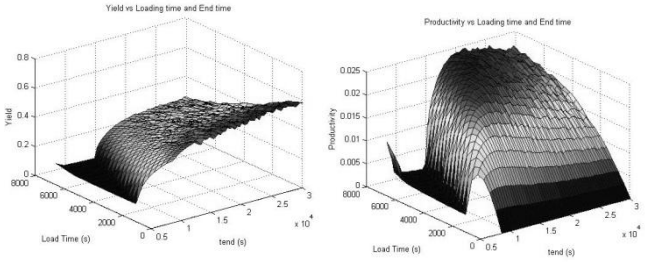

Figure 5. 3D plots of the yield (left) and the productivity (right) as a function of loading time and end time.

As seen in the figure 5, a maximum yield is achieved by a low loading time and a high end time. The productivity is required a high loading volume and a short end time.

The third and last gradient to optimize is the step salt gradient which is simulated with a salt concentration between 1 – 40% and a loading time from 1 – 6500 seconds.

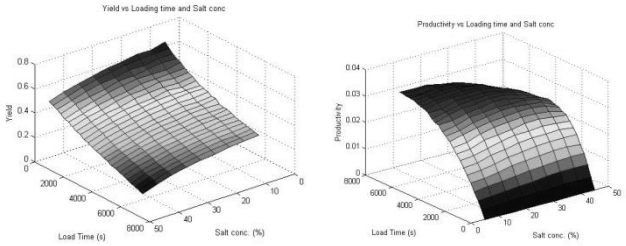

Figure 6. Surf plots of the yield (left) and the productivity (right) as a function of loading time and salt concentration.

As seen in figure 6, a high yield needs a low loading volume and a low salt step while the maximum productivity needs a high loading time together with a small salt step.

Table 3. Comparison between the different salt gradients.

	Isocratic		**Linear**		**Step**	
	Salt conc.	Load time	End time	Load time	Salt conc.	Load time
Productivity	40%	4999 s	19826 s	7066 s	10%	5728 s
value $[kg/(m^3 \cdot s)]$	0.0291		0.0248		0.0340	
Yield	42.8%	3.3 s	56697 s	479 s	1.7%	24.6 s
value [-]	0.698		0.612		0.637	
Yield·Productivity	40%	2670 s	18964 s	4786 s	10%	3964
value $[kg/(m^3 \cdot s)]$	0.0117		0.0079		0.0110	

Above is a comparison between all de different gradients and since yield and productivity work opposite each other, the yield·productivity values are also listed in the table. As seen in table 3, the optimum productivity is achieved by using a step salt gradient. The yield is almost the same size, but the highest yield is by using an isocratic salt gradient and the combined yield·productivity is highest at an isocratic salt gradient.

Conclusion

The calibration is based on mathematical models and there is always a deviation that occurs and some estimation of the parameters, as well as reading the experimental data can give rise to deviations as well.

The highest yield is when using an isocratic salt gradient while a high productivity is preferred to run at a step gradient. A combination of yield and productivity, yield·productivity, has its maximum at an isocratic gradient.

References

[1] McCue J.T., 2009, *Methods in Enzymology: Chapter 25 Theory and Use of Hydrophobic Interaction Chromatography in Protein Purification Applications*, 2nd Edition, Guide to Protein Purification, p 405-414, Vol: 463.

[2] Jakobsson N., Degerman M., Nilsson B., 2005, *Optimisation and robustness analysis of a hydrophobic interaction chromatography step*, Journal of Chromatography.

[3] Degerman M., Jakobsson N., Nilsson B., 2007, *Modeling and optimization of preparative reversed-phase liquid chromatography for insulin purification*, Journal of Cromatography.

[4] Guiochon G., 1994, *Fundamentals of Preparative and Nonlinear Chromatography*, Academic Press Limited
[5] Beers K.J., 2006, *Numerical Methods for Chemical Engineering: Applications in MATLAB*, Massachusetts Institute of Technology, p 213-215
[6] Degerman M., 2004, *Optimization of chromatography process – Methods and robustness*, Lund University