Fast model calibration of chromatographic process -Separation of protein monomers from dimers

Niklas Borg
Department of Chemical engineering
Lund University, Sweden
November 2008

Abstract

A successful setup for a calibration of a model for separation of BSA monomer and dimer is described. The BSA model is calibrated in steps using different experiments and residuals before a tuning of the model to all experiments is performed. The steps include isotherm calibration to gradient experiments, capacity calibration to load experiments, transfer resistance calibration to all experiments and pH dependency calibration to experiments with different buffers.
The advantages with this method are that iterations are not necessary and the whole process from the first experiments to the calibrated model takes only a few days.
The disadvantages are that some parameters are very hard to calibrate with the used simple method.

Keywords: BSA, Langmuir MPM, calibration, ion exchange, general rate

Introduction

In this paper a simulated model for adsorption of bovine serum albumin on an HP column is calibrated to experimental data for a general rate, Langmuir mobile phase modulator model. The calibration is performed in steps, since it is hard to calibrate 14 parameters at once.

Theory

Ion exchange chromatography (IEC) is a process where compounds are separated due to differences in ionic binding properties. A mixture of the compounds is pumped through the matrix (stationary phase) in the column. The stationary phase has immobilized charged groups onto which the compounds can form ion bindings. Depending on the differences in binding strength, the compounds are then separated, since the compound with stronger binding will take longer time to pass through the column.[i]

A way to speed up the process and to enhance the separation is to influence the binding strength during the process. The binding strength depends on the salt concentration and pH in the mobile phase, so a change in either of those will change the binding strength of the compounds. Usually the compounds are loaded at a low salt concentration to make the binding strength high so all of the compounds are adsorbed to the stationary phase. A gradient from low to high salt concentration is then applied to elute the compounds. The weakest binding will be eluted at a lower salt concentration than the stronger binding, and thus the compounds are separated.[ii]

Models

General rate model

In the general rate model (GRM) the flow only moves the volume between the particles of stationary phase. The concentrations in the mobile phase in the pores change due to diffusion instead.

The diffusion into the particles is described by the equation:

$$\frac{\partial c_p(z,r,t)}{\partial t} = \frac{D_e}{\varepsilon_p}\left(\frac{\partial^2 c_p(z,r,t)}{\partial r^2} + \frac{2}{r}\frac{\partial c_p(z,r,t)}{\partial r}\right) - r_{ads}$$

where c_p, is the concentration in the pore, r is the radius of the particle, t is the time, D_e is the effective diffusion coefficient, ε_p is the void in the stationary phase for the specific compound and r_{ads} is the adsorption term. The boundary condition in the core of the particle is a homogenous Neumann condition and at the surface the boundary condition is the Robin condition

$$\frac{\partial c_p(z,R,t)}{\partial r} = \frac{k_f}{D_e}\left(c_b(z,t) - c_p(z,R,t)\right)$$

where k_f is the outer mass transfer coefficient, c_b is the concentration in the bulk flow and R is the radius at the surface.

The flow between the particles is described by:

$$\frac{\partial c_b(z,t)}{\partial t} = -v_{int}\frac{\partial c_b(z,t)}{\partial z} + D_{ax}\frac{\partial^2 c_b(z,t)}{\partial t^2} - \frac{1-\varepsilon_c}{\varepsilon_c} * kf * \frac{3}{R}\left(c_b(z,t) - c_p(z,R,t)\right)$$

where D_{ax} is the axial dispersion coefficient and v_{int} is now calculated as

$$v_{int} = \frac{F}{R^2 * \pi} * \frac{1}{\varepsilon_c}$$

The concentration changes due to adsorption are described by:

$$r_{ads} = -\frac{1}{\varepsilon_p} * \frac{\partial q(z,t)}{\partial t}$$

where q is the concentration of adsorbed compound.[iii]

Langmuir mobile phase modulators

The mobile phase modulators model (MPM) is based on the Langmuir kinetics. The Langmuir kinetics has the equation:

$$\frac{\partial q}{\partial t} = k_{kin}\left(K_{eq} * q_{max} * c * \left(1 - \frac{q}{q_{max}}\right) - q\right)$$

where q is the concentration of adsorbed compound, k_{kin} is the reaction rate for the sorption, K_{eq} is the equilibrium constant for the sorption reaction, c is the concentration corresponding to q and q_{max} is the maximum concentration of q for the stationary phase.

MPM is named after the modulators added to the Langmuir kinetics:

$$\frac{\partial q}{\partial t} = k_{kin} * s^{\beta}\left(K_{eq} * q_{max} * s^{-\beta} * e^{-\gamma s} * e^{-\delta(pH - pH_{ref})} * c * \left(1 - \frac{q}{q_{max}}\right) - q\right)$$

where β describes how much faster q is desorbed, γ how hydrophobic the binding is, depending on the concentration of salt and δ the pH dependency.[iv]

Experimental design

The objective in the experimental design for bovine serum albumin (BSA) was to calibrate a model for simulator of purification of monomers of BSA (mBSA) from dimers of BSA (dBSA). The column used was GE Healthcare HiTrap Q HP 1 ml at the flow rate 1 ml/minute.

Previous experience has given an experimental design to start with.[v] The sample buffer (100 ml) was composed of 50 mM TRIS-buffer, 0 mM with a pH of 8.8 and 1 g/l BSA. The buffers used in the experiments were composed of 50 mM TRIS-buffer, 0 mM NaCl (or 1000 mM NaCl for the elution buffer) and a pH of either 8.8, 8.6, 8.4, 8.2, 8.0, 7.8 or 7.6. The pH was varied first, to see what pH would be good to use. The buffers were pumped through the loop, both during the loading and the elution. The volume pumped through for loading was 3 ml in all experiments except the last experiment with pH 8.0, when 6 ml was pumped trough. The gradient length was varied to give experiments to calibrate the adsorption isotherm with. The load volume was varied to study the capacity of the column. Both load and gradient experiments could then be used to calibrate the transfer resistances. See table 1 for a list of the experiments.

pH	Load volume	Gradient length
8.8	2 ml	40 ml
8.6	2 ml	40 ml
8.4	2 ml	40 ml
8.2	2 ml	40 ml
8.0	2 ml	40 ml
7.8	2 ml	40 ml
7.6	2 ml	40 ml
8.0	2 ml	20 ml
8.0	2 ml	30 ml
8.0	5 ml	40 ml
8.4	2 ml	20 ml
8.4	2 ml	30 ml
8.4	2 ml	60 ml
8.4	0.1 ml	40 ml
8.4	0.5 ml	40 ml
8.4	5 ml	40 ml

Table 1. Experiments with BSA

The sample was filtered before each experiment. Time spent to do these 16 experiments was approximately 30 hours. This includes the experiments at pH 8.0 which was performed while it was realised that pH 8.4

gave better separation. They are now used for validating the model instead.

Calibration

When calibrating a model to experimental results a large number of simulations have to be done. The better the initial guess is, the faster the calibration will be. With a bad initial guess the calibration might even not be possible. The parameters that matters the most in a simulation of IEC is K_{eq}, q_{max} and β. When these are in the correct range k_{kin}, D_e and k_f can be approximated. These can then be combined to a good initial guess.

The approach used was to look at the difference in position for the maximum of the experimental and simulated peaks and at the difference in width between the simulated and experimental peaks.

The pH was calibrated last of all, since this pH can be used as pH reference, and will thus not change while calibrating the pH dependency. This was done by calibrating Henry's constant, H, for each pH and look at the trend. The reason this was not calibrated directly with simulations was that there were errors in the experimental setup. These were systematic though, so the trend should still be the same, just shifted in pH.

First the positions for mBSA and dBSA were calibrated from an initial guess that was taken from literature. Simulations were performed on a homogenous model with rough grid, 50 points and a three point difference discretization. The value to minimize was the sum of the squares of the distances between experimental UV peak positions and simulated concentration peak positions for gradient experiments. That the concentration and not UV curve was used as the simulated curve was because the dBSA-peak is much harder to detect from a UV curve. This calibration gave rough idea about where H (Henry's constant) and β should be. If it had been unknown compounds, any parameters that give concentration peaks within the gradient should work start from, since this optimization is relatively fast and easy for the optimizer fminsearch.

To separate H into K_{eq} and q_{max} the sum of the squares of the distances between experimental UV peak positions and simulated concentration peak positions for loading experiments were minimized. The same model was used, and β was allowed to change, to adapt to the new behaviour. The initial guess for q_{max} was 100 kg/m^3, and K_{eq} calculated from $H=K_{eq}*q_{max}$. This optimization with fminsearch was as fast as the step above.

The broadening of the peaks was estimated with a heterogeneous model with 5 radial grid points and 200 axial grid points. The width of the peak was calculated as the distance between the peak position and the position where the height was half of the peak. First a Latin hypercube (lhs) sampling of k_f, k_{kin} and d_e was performed. The value to minimize was the sum of the squares of the differences between experimental UV widths and simulated concentration widths for three experiments. The three experiments used was the one with shortest gradient, the one with longest gradient and the one with highest loading. The best result was used as an initial guess in an fminsearch optimization. The value to minimize was the sum of the squares of the differences between experimental UV widths and simulated UV widths and the differences between experimental UV peak position and experimental peak positions for six experiments. That the UV curve was used for the simulated values here depends on that the dBSA peak is slightly influenced by the tail of the mBSA. This effect would not be taken into account if only concentrations were compared to the simulated UV curve.

The scan for the width moved the peak positions; the final step was to optimize both the positions and the width together with the advanced model with fminsearch.

All of these steps were made in one script in Matlab. When one step was completed, the resulting values were saved. The steps were run individually, by commenting out the old steps when a new was written, but theoretically this could be run as one script, without any input from the user. The most important parameter here is β, since it will not be able to change very much after the first two steps while using fminsearch.

The pH dependency was calculated as a relative change in Henrys constant. The

relation was approximated to be linear between the logarithm of the H^+ concentration and the logarithm of the Henry's constant.

The calibrated values for k_f, D_e, k_{kin}, Keq, q_{max}, δ and β are presented in table 2.

Compound	mBSA	dBSA
k_f	3.156e-4	6.7e-3
D_e	4.46e-12	3.2e-12
k_{kin}	17.8059	0.1396
δ	-0.91	-0.80
Keq	5.346e8	1.936e9
q_{max}	1.3876	0.9565
β	3.2754	3.2911

Table 2. Final calibration of model.

Validation

The experiments used to calibrate, marked as solid lines in figure 1, can be predicted by simulation of the model, marked as dashed lines in the figure.

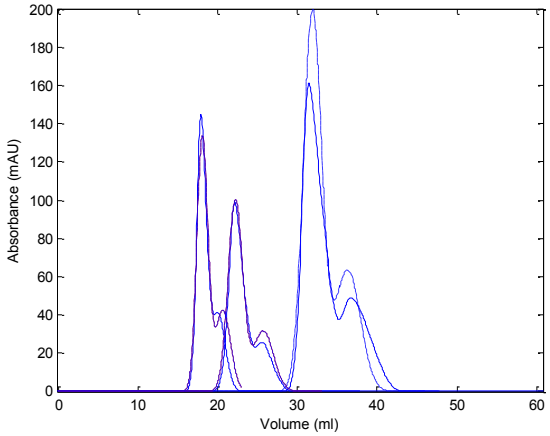

Figure 1. Validation of model

Discussion

The experiments are not quite accurate, since they were performed in quite a haste to be presentable to a meeting. If an industrial application should be designed, a new parameter estimation from more careful experiments is advised. A known flaw in the experiments used here were that the pumps and the tubes from pump to mixing tank was not flushed when the buffer was changed.

Fminsearch could probably not find very good estimations, since it is hard to change all parameters to get a better guess. What it did was more likely to find a good balance close to the initial guess. The calibration is not bad, but it could be better.

Conclusion

Advantages for this method:

- Fast calibration: can be calibrated within 48 hours.
- Low amount of target compound used: 100 mg.
- Good enough calibration: position and width of experiments coincide with simulations.

Disadvantages for this method:

- Not very good shape of the peaks.

With a more advanced target function for the method to minimize, the shape of the peaks could be taken into account. A clever way to calculate this and weigh it to fit the size of the position and width residuals is left as future work.

References

[i] Guichon, G. Golshan-Shirazi, S. and Katti, A.M., 1994 Fundamentals of preparative and nonlinear chromatography, Academic Press Limited, London, 726-727.

[ii] Guichon, G. Golshan-Shirazi, S. and Katti, A.M., 1994 Fundamentals of preparative and nonlinear chromatography, Academic Press Limited, London, 699-700.

[iii] Guichon, G. Golshan-Shirazi, S. and Katti, A.M., 1994 Fundamentals of preparative and nonlinear chromatography, Academic Press Limited, London, 36-39.

[iv] Melander, W.R., El Rasse, S. and Horváth, C., 1989. Interplay of hydrophobic and electrostactic interactions in biopolymer chromatography, *Journal of chromatography,* **469**, 3-27

[v] Jakobsson, N., 2006, A Model-Based Approach to Robustness Analysis and Optimisation of Chromatography Processes, *Ph.D. Thesis.*