Model-based design of peptide chromatographic purification processes

David Gétaz a,1, Guido Stroehlein a,b,2, Alessandro Butté c, Massimo Morbidelli a,*

a Department of Chemistry and Applied Bioscience, Institute for Chemical and Bioengineering, ETH Zurich, CH-8093 Zurich, Switzerland
b ChromaCon AG, CH-8005 Zurich, Switzerland
c Lonza Ltd, ECM RED Tides, CH-3930 Visp, Switzerland

A R T I C L E   I N F O

Article history:
Received 27 October 2012
Received in revised form 30 January 2013
Accepted 31 January 2013
Available online 9 February 2013

Keywords:
Process
Reversed-phase
Peptide
Chromatography
Optimization
Preparative

A B S T R A C T

In this work we present a general procedure for the model-based optimization of a polypeptide crude mixture purification process through its application to a case of industrial relevance. This is done to show how much modeling can be beneficial to optimize complex chromatographic processes in the industrial environment. The target peptide elution profile was modeled with a two sites adsorption equilibrium isotherm exhibiting two inflection points. The variation of the isotherm parameters with the modifier concentration was accounted for. The adsorption isotherm parameters of the target peptide were obtained by the inverse method. The elution of the impurities was approximated by lumping them into pseudo-impurities and by regressing their adsorption isotherm parameters directly as a function of the corresponding parameters of the target peptide. After model calibration and validation by comparison with suitable experimental data, Pareto optimizations of the process were carried out so as to select the optimal batch process.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

With increasing production needs of therapeutic polypeptides, there is a significant interest in developing more efficient production processes. Since the purification step is often the bottleneck in the synthetic peptide production, it is important to develop procedures for its optimization. Among the various possible purification techniques, reversed-phase chromatography is the method of choice for peptide purification.

Usually purification processes development is done under strong time limitation and with small amount of target product. Those constraints usually prevent the use of detailed process modeling due to the time and product consuming model calibration (e.g. frontal analysis) and to the lack of methodology to build an appropriate model. However, in the recent years, strong efforts have been made to incorporate process modeling in the optimization and validation of chromatographic processes [1–9].

The chromatographic models found in the literature to describe the analyte elution can incorporate various levels of complexity. However, they all contain the same basic building blocks: a chromatographic mass balance and a multi-component adsorption isotherm. Several forms of the chromatographic mass balances can be used to describe the elution of the components (e.g. the equilibrium–dispersive model, the lumped kinetic model, the general rate model, etc.) [10]. The selection of the appropriate model depends on the level of complexity required to describe accurately the elution profiles. On the other hand, the selection of the appropriate adsorption equilibrium isotherm required to describe the distribution of the analyte between the mobile and the stationary phase is strongly dependent on the properties of the system studied (i.e. analyte properties, stationary phase type and mobile phase composition) [10]. Several methods are available for the determination of the adsorption equilibrium isotherm (e.g. frontal analysis, perturbation method, inverse method, etc.). The selection of the method used to determine the adsorption equilibrium isotherm mainly depends on the amount of product available for the model calibration and on the accuracy required to describe the analytic elution.

Typically, once the chromatographic model has been selected and calibrated, an optimization procedure is used to determine the optimal batch process based on a Pareto optimization [1–5,7]. In addition, a sensitivity analysis can be used to characterize the sensitivity of the process and to determine the critical process parameters [6–8]. These results can then be used to validate the
process and to ensure the process and product quality according to FDA regulations.

This paper presents the model-based optimization of a peptide purification process. The model calibration and the optimization study were done in a systematic manner following the step sequence in Fig. 1. It will be shown that by carefully following these steps, it is possible to develop a very robust and predictive chromatographic model, based on which reliable process optimization can be carried out. In addition, such a model could also be very useful for process sensitivity analysis and Quality-by-design.

The general procedure presented in this paper can be applied to any peptide chromatographic purification process. Of course, the choice of the adsorption isotherm will depend on the system studied. However, the methodology and the simplifications presented in this paper are valid for most of the chromatographic purification processes. In fact, the procedure was successfully applied by our research group to several peptide purification processes.

2. The framework

Often, the optimization of chromatographic processes is done empirically on the basis of heuristics and experience. This empirical approach leads in general to suboptimal process performances. In this work, an attempt was made to rationalize the process design by developing a general procedure that can be followed to design peptide purification processes using a model based approach. The procedure is sketched in Fig. 1 and it contains four main steps:

2.1. Model calibration for target component

The first step in a model based process optimization is to select the appropriate model for the system and to calibrate the model parameters. To do this, it is important to first characterize the adsorption characteristics of the pure target component and then to develop a model able to predict its elution behavior.

2.2. Impurity characterization

Once the target component adsorption has been characterized, the impurities elution has to be determined. The impurities having similar elution behavior are usually lumped into pseudo-impurities. Their selectivity is measured in diluted conditions and the remaining adsorption isotherm parameters are regressed from the target peptide isotherm parameters. The justification for doing this is that the most critical impurities are the ones eluting closer to the target component, which therefore, exhibit a chromatographic behavior very similar to that of the target component.

2.3. Model verification

Before using the model, it is important to test its prediction validity against a suitable set of experimental data.

2.4. Process optimization

Finally the model can be used to perform a multi-objective optimization. The results are represented as productivity–yield Pareto curves at constant product purity and the optimal batch process can be selected based on the optimization results and on additional economical constraints.

3. Definition of the purification process

3.1. The preparative purification

The purification of a long peptide with more than 30 amino acids was optimized using a model-based approach. The crude mixture was directly obtained from the Lonza production process, so that a real industrial multi-component mixture could be investigated. The polypeptide was synthesized by solid phase synthesis [11]. The structures of the major impurities are therefore very closely related to the target peptide structure. The starting material had a purity of about 60%. This crude mixture was purified using a 2 steps purification process. The first purification step was carried out to increase the purity of the starting material up to 94.5%. This product pool was then further purified to reach the final purity specification of 98%. This second purification step will be used in the sequel as a case study to illustrate the model-based process development procedure presented in this work.

The second purification step was performed on a Kromasil 100A 10 μm C8 4.6 mm × 250 mm column obtained from EKA chemicals AB (Bohus, Sweden). The preparative buffer composition is summarized in Table 1 (i.e. buffer A1 and B1). The temperature was set to 25 °C and the flow rate was 0.5 mL/min.

The feed for the second purification step had a concentration of 3.5 g/L and the purity was 94.5%. An analytical chromatogram of the feed is shown in Fig. 2. It can be seen that the feed is a complex mixture containing more than 20 impurities. To simplify the feed characterization, the impurities were classified in 3 groups, based

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>AcOH/H2O/AcN: 0.5/97.5/2.0 (v/v/v)</td>
</tr>
<tr>
<td>B1</td>
<td>AcOH/H2O/AcN: 0.5/49.5/50.0 (v/v/v)</td>
</tr>
<tr>
<td>A2</td>
<td>200 mM KH2PO4 pH 2/ AcN: 93.5/6.5 (v/v)</td>
</tr>
<tr>
<td>B2</td>
<td>200 mM KH2PO4 pH 2/ AcN: 55/45 (v/v)</td>
</tr>
</tbody>
</table>

Fig. 2. Analytical chromatogram of the feed. Note: Three different types of impurities have been defined.
on their elution on the analytical column (A, B and C-type impurity). The final product purity specification was 98% with less than 1% of the individual impurities.

3.2. Chemicals

HPLC grade acetonitrile was purchased from Sigma–Aldrich (Buchs, Switzerland). Orthophosphoric acid 85% was purchased from Merck (Darmstadt, Germany). Acetic acid glacial was purchased from Scharlau Chemie S.A. (Barcelona, Spain). All the chemicals were used without further purification. The deionized water was purified with a Simpak2 unit (Millipore, MA, USA) before use. The synthetic polypeptide used in this study was kindly donated by Lonza Ltd (Visp, Switzerland).

3.3. Experimental setup

The experiments were carried out on an Agilent 1100 Series HPLC, equipped with an auto-sampler, a diode array detector, an online-degasser and a quaternary gradient pump. A Gilson FC 203B fraction collector (Middleton, WI, USA) was connected at the outlet of the HPLC to collect fractions during the peptide elution.

3.4. Analytical method

A Symmetry 300 C4, 3.5 μm, 300 Å, 150 mm × 4.6 mm column obtained from Waters A.G. (Baden-Dättwil, Switzerland) was used to analyze the purity of the peptide fractions collected during the chromatographic experiments. The analytical buffer composition is listed in Table 1 (i.e. buffer A2 and B2). A gradient from 57% B2 to 73.7% B2 was carried out in 50 min on the Symmetry column at a temperature of 55 °C. The UV response at 214 nm was recorded and calibrated using samples of known peptide concentration.

4. Model development

4.1. Target peptide calibration

In this manuscript, the lumped kinetic model was used to describe the chromatographic behavior of the peptide mixture under examination. The mass balance of this model for a solute i is written as follow [10]:

\[ \frac{\partial c_i}{\partial t} + \frac{1 - \varepsilon_i}{\varepsilon_i} \frac{\partial q_i}{\partial t} + u_{lin} \frac{\partial c_i}{\partial z} = D_{i,i} \frac{\partial^2 c_i}{\partial z^2} \]  

where \( c_i \) and \( q_i \) are the concentrations of component i in the mobile and in the adsorbed phase respectively; \( z \) and \( t \) are the spatial and time coordinate; \( u_{lin} \) is the linear mobile phase velocity; \( D_{i,i} \) is the axial dispersion coefficient and \( \varepsilon_i \) is the porosity accessible to the ith component. The mass transport between stationary and mobile phases is described through the solid film linear driving force model:

\[ \frac{\partial q_i}{\partial t} = k_{M,i}(q_{eq,i} - q_i) \]  

where \( k_{M,i} \) is the lumped mass transfer coefficient and \( q_{eq,i} \) is the concentration of component i adsorbed in equilibrium with the amount in the mobile phase, i.e. \( c_i \).

Several model parameters have to be determined in order to calibrate the lumped kinetic model. The determination of the lumped mass transfer coefficient \( k_{M,i} \), the axial dispersion coefficient \( D_{i,i} \) and the porosity \( \varepsilon_i \) is presented in Section 4.1.1. After that, the peptide adsorption isotherm is determined in Section 4.1.2 (for diluted conditions) and in Section 4.1.3 (for overloaded conditions).

4.1.1. Target peptide mass transfer, diffusion and porosity

The porosity accessible to the peptide is an important model parameter that has to be determined experimentally. This quantity was obtained by measuring the peptide retention time in a mobile phase composed of 100% B1 buffer. In these conditions, the peptide does not adsorb because of the high acetonitrile concentration present in the eluent [12]. The peptide retention time is therefore equivalent to the column dead time for the peptide, \( t_{0,i} \) and therefore the porosity accessible to the specific peptide can be calculated as follows:

\[ \varepsilon_i = \frac{V_{L,i}}{V_c} = \frac{Q \cdot t_{0,i}}{V_c} \]  

where \( V_c \) is the column volume, \( V_{L,i} \) is the volume accessible to the peptide and \( Q \) is the flowrate. The so obtained porosity value is reported in Table 2.

In addition to the column porosity, the hydrodynamic properties of the column have to be determined. The mass transfer coefficient \( k_{MD} \) and the axial dispersion coefficient \( D_{L,i} \) were obtained by measuring the Van Deemter curve [13], which correlates the height equivalent to a theoretical plate (HETP) to the hydrodynamic properties of the column as follows:

\[ \text{HETP}_i = 2K_{eddy,i} + 2 \left( \frac{k_i^*}{1 + k_i^*} \right) \frac{u_{lin}}{k_i^*k_i^*_{M,i}} \]  

The retention factor \( k_i^* \) was computed from the peptide retention time:

\[ k_i^* = \frac{t_{R,i} - t_{0,i}}{t_{0,i}} \]  

while \( K_{eddy,i} \) can be correlated to the axial dispersion coefficient \( D_{L,i} \) using the following relation [14]:

\[ D_{L,i} = K_{eddy,i} \cdot u_{lin} \]  

The experimental HETP of the synthetic peptide was estimated from its retention time, \( t_{R,i} \), and peak width at half height, \( w_{1/2} \), in weakly adsorbing conditions (\( k^* = 5.6 \)) as follows:

\[ \text{HETP}_i = \frac{L_c \cdot w_{1/2}^2}{5.55 \cdot t_{R,i}^2} \]  

The mass transfer coefficient \( k_{MD} \) and the eddy diffusion parameter \( K_{eddy} \) were then obtained by linear regression of the experimental HETP curve and the corresponding values are summarized in Table 2.

<table>
<thead>
<tr>
<th>First estimate</th>
<th>Fine tuning</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_{R,\text{pep}} ) [−]</td>
<td>0.37</td>
</tr>
<tr>
<td>( k_{MD,\text{pep}} ) [min]</td>
<td>61.7</td>
</tr>
<tr>
<td>( K_{eddy,\text{pep}} ) [cm]</td>
<td>2.8 × 10−3</td>
</tr>
<tr>
<td>( \alpha_{\text{pep}} ) [−]</td>
<td>2.89 × 103</td>
</tr>
<tr>
<td>( \alpha_{\text{pep}} ) [−]</td>
<td>0.533</td>
</tr>
<tr>
<td>( \alpha_{\text{pep}} ) [−]</td>
<td>0.507</td>
</tr>
<tr>
<td>( \alpha_{\text{pep}} ) [g/L]</td>
<td>203.5</td>
</tr>
<tr>
<td>( \alpha_{\text{pep}} ) [g/L]</td>
<td>318.0</td>
</tr>
<tr>
<td>( q_{\text{satmax,pep}} ) [g/L]</td>
<td>−0.984</td>
</tr>
<tr>
<td>( q_{\text{satmax,pep}} ) [g/L]</td>
<td>0.84</td>
</tr>
<tr>
<td>( q_{\text{satmax,pep}} ) [g/L]</td>
<td>269.2</td>
</tr>
</tbody>
</table>
4.1.2. Target peptide elution in diluted conditions

The next step in the model calibration is to characterize the target peptide chromatographic behavior in diluted conditions. In these conditions, the adsorption isotherm can be assumed to be linear and the initial slope of the adsorption isotherm (i.e. Henry coefficient $H_i$) can be obtained from the measured retention time:

$$t_{R,i}(c) = t_{0,i} \left(1 + \frac{1 - \varepsilon_i}{\varepsilon_i H_i} c_i\right)$$

The measured values of the target peptide Henry coefficient at different acetonitrile concentration are shown in Fig. 3. The obtained results were fitted to the stoichiometric displacement model suggested by Geng and Reignier for the retention of proteins and peptides in reversed-phase chromatography [15]:

$$H_i = \alpha_{1,i}\left(c_{ACN}\right)^{-\alpha_{2,i}}$$

where $c_{ACN}$ is the acetonitrile volume fraction, $\alpha_{1,i}$ is a fitting parameter related to the phase ratio and to the number of stationary phase ligand involved in the peptide-stationary phase interaction and $\alpha_{2,i}$ is a parameter proportional to the amount of acetonitrile required for the peptide displacement. All the corresponding fitting parameters values are summarized in Table 2.

4.1.3. Target peptide elution in overloaded conditions

Now that the peptide elution in diluted conditions has been characterized, it is important to study the peptide elution in overloaded conditions. In industrial purification processes, in fact, the column is usually overloaded to increase the process productivity and therefore it is important to have good model predictions in overloaded conditions.

The peptide elution profile for different loadings in isocratic conditions is shown in Fig. 4. It can be seen that at relatively low loading (Loading <0.24 g peptide/L column) the peak has a shock in the front and a dispersed rear. This kind of peak shape is typically due to a Langmuir adsorption isotherm [16]. On the other hand, when the column is further overloaded (Loading >0.24 g peptide/L column), a shock is created in the rear and the front becomes dispersed. This peak shape is usually encountered when adsorbate–adsorbate interactions are present. This behavior can be described using the Moreau isotherm [17]. Accordingly, in order to fully describe the peptide adsorption, it was assumed that two different adsorption sites were present on the stationary phase surface. The first site exhibits a Langmuir type of adsorption isotherm and it is predominant at low peptide concentrations. The second one exhibits, instead, a Moreau type of adsorption isotherm. This site becomes predominant at higher peptide concentration. Thus the overall adsorption equilibrium model is given by:

$$q_{eq,i} = \frac{H_{1,i}c_i}{1 + \left(\frac{H_{1,i}}{q_{sat,1,i}}\right)c_i} \left(1 + \frac{H_{2,i}c_i(1 + (l_{2,i}/(2 \cdot q_{sat,2,i}))c_i))}{1 + \left(H_{2,i}/q_{sat,2,i}\right)c_i + l_{2,i}(H_{2,i}/(2 \cdot q_{sat,2,i}))^2}\right)^2$$

(10)

where, $l_{2,i}$ is the adsorbate–adsorbate interaction parameter for the interaction between two components $i$. $H_{1,i}$, $H_{2,i}$, $q_{sat,1,i}$ and $q_{sat,2,i}$ are the Henry coefficient and saturation capacity of component $i$ on the first and second site, respectively. By definition, the Henry coefficient and the saturation capacity of the two sites are related to the overall Henry coefficient $H_i$ and $q_{sat,i}$:

$$H_i = H_{1,i} + H_{2,i}$$

(11)

$$q_{sat,i} = q_{sat,1,i} + q_{sat,2,i}$$

(12)

Eqs. (10)–(12) form a system of three equations with eight unknown variables (i.e. $q_{sat,1,i}$, $H_{1,i}$, $H_{2,i}$, $q_{sat,1,i}$, $q_{sat,2,i}$ and $l_{2,i}$) for a given target peptide, $i$. Since the overall Henry coefficient $H_i$ was already measured in Section 4.1.2, only four adsorption isotherm parameters (i.e. $H_{1,i}$, $q_{sat,1,i}$, $q_{sat,2,i}$ and $l_{2,i}$) have to be determined to fully characterize the adsorption isotherm.

There are several methods available to obtain the adsorption isotherm parameters (e.g. frontal analysis, perturbation method, inverse method, etc.). In this study, the isotherm parameters were obtained by the inverse method (i.e. peak fitting) because of the limited amount of target peptide available for the model development. A series of overloaded injections were carried out at three different acetonitrile concentrations. The isotherm parameters were fitted to the elution profiles using a Matlab gateway routine to Lester Ingber’s Adaptive Simulated Annealing [18]. A set of four adsorption isotherm parameters (i.e. $H_{2,i}$, $q_{sat,1,i}$, $q_{sat,2,i}$ and $l_{2,i}$) were therefore obtained for each acetonitrile concentration. The experimental and fitted elution profiles are compared in Fig. 5. The value of the fitted isotherm parameters are summarized in Table 3 together with the overall Henry coefficient, $H_{sep}$, for the target peptide. It is seen that all regressed isotherm parameters show a strong dependence on the acetonitrile concentration. Instead of correlating the second site Henry coefficient $H_{2,i}$ to the acetonitrile concentration this was correlated to the overall Henry coefficient $H_i$, using a power function [2]:

$$H_{2,i} = \alpha_{3,i}\left(H_i^{\nu_{2,i}}\right)$$

(13)

It is important to note that this correlation implies that the Henry coefficient of the second adsorption site follows the retention model proposed by Geng and Reignier for the elution of peptides in RP-chromatography [15]. The fitting of Eq. (13) to the regressed adsorption isotherm parameters is shown in Fig. 6.

Few studies have been reported regarding the variation of the saturation capacity with the modifier concentration [2,19–21]. Jacobson et al. [19] suggested that the saturation capacity remains almost constant over the useful range of variation of the modifier concentration. This assumption was experimentally confirmed by El Fallah and Guichon for both the elution of small molecule [20] and proteins [21] in reversed-phase chromatography. The range
Table 3
Adsorption isotherm parameters of Eq. (10) determined at different acetonitrile concentration. Note: The corresponding peptide overall Henry coefficient values $H_{pp}$ was calculated from Eq. (9).

<table>
<thead>
<tr>
<th>Acetonitrile concentration [v/v]%</th>
<th>$H_{pp}$ [-]</th>
<th>$q_{sat,pp}$ [g/L]</th>
<th>$H_{2,pp}$ [-]</th>
<th>$q_{sat2,pp}$ [g/L]</th>
<th>$l_{pp}$ [g/L]</th>
</tr>
</thead>
<tbody>
<tr>
<td>25.58</td>
<td>16.4</td>
<td>0.81</td>
<td>2.31</td>
<td>164</td>
<td>136</td>
</tr>
<tr>
<td>25.98</td>
<td>10.9</td>
<td>0.81</td>
<td>1.64</td>
<td>144</td>
<td>209</td>
</tr>
<tr>
<td>26.52</td>
<td>6.3</td>
<td>0.91</td>
<td>1.41</td>
<td>149</td>
<td>217</td>
</tr>
</tbody>
</table>

of acetonitrile concentration used in those studies was however limited. This limited range is actually not representative of the strong variation of acetonitrile concentration encountered during the industrial gradient purification of peptides. Aumann et al. [2] have described the effect of acetonitrile on the saturation capacity of Calcitonin by correlating its saturation capacity to its Henry coefficient using a power function. The drawback of this correlation is that it lacks of physical meaning. In fact, at very low acetonitrile concentration...
concentration, the Henry coefficient tends toward very high values and the saturation capacities obtained with this correlation are strongly over-evaluated. In this work, a saturation-type function was chosen to correlate the saturation capacities of the two adsorption sites, $q_{sat1,i}$ and $q_{sat2,i}$, with their corresponding Henry coefficient, $H_{1,i}$ and $H_{2,i}$ [14]:

$$q_{sat1,i} = \frac{\alpha_{5,i} H_{1,i}}{1 + (\alpha_{5,i}/q_{sat1,max,i}) H_{1,i}}$$

$$q_{sat2,i} = \frac{\alpha_{6,i} H_{2,i}}{1 + (\alpha_{6,i}/q_{sat2,max,i}) H_{2,i}}$$

(14) (15)

This correlation implies that at very low acetonitrile concentration, when the adsorption strength is very high, the saturation capacity reaches a maximum value, i.e., $q_{sat1,max,i}$ and $q_{sat2,max,i}$. This saturation capacity corresponds to the close packing of peptides on the stationary phase surface. When the acetonitrile concentration is increased, the adsorption strength decreases and the saturation capacity decreases due to the repulsion between the adsorbed peptide molecules [22]. The fitting of Eqs. (14) and (15) to the regressed adsorption isotherm parameters is shown in Fig. 6.

Finally, the variation of the adsorbate–adsorbate interaction parameter $I_{ii}$ with the modifier concentration has been seldom reported in the literature [23]. In this study, it was correlated to the Henry coefficient of the second adsorption site $H_{2,i}$, using a power function:

$$I_{ii} = \alpha_{7,i} H_{2,i}^2$$

(16)

The fitting of Eq. (16) to the regressed adsorption isotherm parameters is shown in Fig. 6.

The parameter values of Eqs. (13)–(16) are summarized in the first column of Table 2. Now that the effect of the acetonitrile concentration on the peptide adsorption has been characterized, it is possible to simulate gradient elutions. A series of overloaded gradient elutions were performed. The gradient elution profiles obtained were fitted together with the isocratic elution profiles shown in Fig. 5 to perform a fine-tuning of the parameter values, $\alpha_{3,i}$, $\alpha_{4,i}$, $\alpha_{5,i}$, $\alpha_{6,i}$, $\alpha_{7,i}$, $\alpha_{8,i}$, $q_{sat1,max}$, and $q_{sat2,max}$. The fitting of the gradient experiments is shown in Fig. 7. The values of the adjustable parameters of Eqs. (13)–(16) are shown in the second column of Table 2. The correlation between the different adsorption isotherm parameters is shown in Fig. 6. It can be seen that the isotherm parameters obtained by fitting of the gradient and isocratic experiments together are very similar to the parameters obtained by fitting only the isocratic experiments.

The calibration of the target peptide adsorption isotherm has been developed in this section. The chromatographic model developed so far is able to predict the overloaded elution of the target peptide in gradient and isocratic conditions. The model can accommodate a total of twelve parameters. Two parameters (i.e., $k_{id}$ and $k_{id},i$) are required to describe the column hydrodynamics. Two parameters (i.e., $\alpha_{1,i}$, $\alpha_{2,i}$) are required to describe the peptide elution in diluted conditions and eight parameters (i.e., $\alpha_{3,i}$, $\alpha_{4,i}$, $\alpha_{5,i}$, $\alpha_{6,i}$, $\alpha_{7,i}$, $\alpha_{8,i}$, $q_{sat1,max}$, and $q_{sat2,max}$) are required to describe the non-linear range of the peptide adsorption isotherm. Suitable experimental procedures have been described for the evaluation of each one of these parameters. The next step in the chromatographic model development is to characterize the chromatographic behavior of the impurities.

4.2. Peptide impurity calibration

4.2.1. Impurity characterization

The crude poly peptide mixture that has to be purified contains many different impurities. In fact, it can be seen from the analytical chromatogram of the feed shown in Fig. 2 that there are over 20 different impurities. Since so many impurities are present in the feed, it is not possible to isolate all of them and determine their adsorption isotherm. For this reason, a series of simplification have to be made to reduce the complexity of the system.

First of all, the number of impurities simulated was reduced by grouping them together based on their adsorption behavior on the preparative stationary phase. An overloaded injection of the feed on the preparative column is shown in Fig. 8. It can be seen that the impurities can be lumped in 3 main groups, referred to in the sequel as pseudo-impurities. The pseudo-impurity L1 has a very low selectivity and it is completely spread below the target peptide peak. The pseudo-impurity L2 is eluting in front of the target peptide peak and the pseudo-impurity H1 is eluting in the rear of the main peak.

The elution of the 3 pseudo-impurities in diluted conditions is shown in Fig. 9. The pseudo-impurity L1 exhibits a very low
selectivity with respect to the target peptide. Its elution was very different from the other light impurities. Therefore it had to be independently simulated. On the other hand, impurity L2 is a pseudo-impurity representing the elution of all the impurities having selectivity lower than that of the pseudo-impurity L1. Finally, the impurity H1 is a pseudo-impurity representing all the late eluting impurities (i.e. all the impurities having a selectivity with respect to the target peptide higher than one).

It is worth pointing out that the pseudo-impurities introduced above represent groups of impurity exhibiting similar elution behavior on the preparative column. However, their elution on the analytical column might differ and this complicates their quantification. For example, not all the early eluting impurities (A and C-type impurity) shown on the analytical chromatogram of Fig. 2 elute before the target peptide on the preparative material. The selectivity on the analytical and on the preparative column might in fact be different.

In order to better characterize the pseudo-impurities, the composition of the fractions collected during the overloaded injection shown in Fig. 8 was studied in more detail. The fractions containing mainly impurity L2 (i.e. fractions eluting between 22 and 45 min) were pooled together and reanalyzed. The same was done with the fractions containing mainly impurity H1 (i.e. fractions eluting between 45 and 60 min). Based on the analytical results, it was found that the pseudo-impurity L2 is composed of 62% of B-type impurity (i.e. impurity eluting before the target peptide on the analytical columns) and 38% of C-type impurity (i.e. impurity eluting after the target peptide on the analytical columns). Similarly, it was found that the pseudo-impurity H1 is composed of 26% of B-type impurity and 74% of C-type impurity. Finally, it was found that the pseudo-impurity L1 is exclusively composed of A-type impurity. Now that the composition of the pseudo-impurities L1, L2 and H1 has been characterized in terms of impurities detectable through the adopted analytical procedure, i.e. A–C, we can quantitatively represent the feed crude composition in terms of pseudo-impurities.

4.2.2. Impurity mass transfer, diffusion and porosity

The polypeptide used in this study was synthesized by solid phase synthesis [11]. The structures of the impurities created during the solid phase synthesis are usually very closely related to the target peptide structure. The mass transfer $k_{M,i}$, the Eddy diffusion parameter $K_{E,i}$ and the porosity accessible can therefore be assumed to be the same for all the components in the system.

4.2.3. Impurity selectivity in diluted conditions

The elution of the 3 pseudo-impurities in diluted conditions is shown in Fig. 9. The pseudo-impurity L1 elutes with a very low selectivity. On the other hand, impurity L2 is a pseudo-impurity representing the elution of all the impurities having selectivity lower than the selectivity of the pseudo-impurity L1. Of course, every single impurity constituting the pseudo-impurity L2 has a slightly different retention time. However, it is possible to approximate their elution by assuming that all the impurities constituting the pseudo-impurity L2 have a retention equal to the closest eluting impurity of the group. Other approximation, such as taking the mean retention time, are in principle possible. However, this approach was favored because it is the most conservative one, since it reduces the chances of process performances overestimation due to overestimated impurity selectivity. The same approach was taken to determine the selectivity of the pseudo-impurity H1.

In general, the pseudo-impurities selectivity $S_i$ in diluted conditions is calculated from data obtained in isocratic conditions using the following relation [2]:

$$S_i = \frac{H_i}{H_{main}}$$  \hspace{1cm} (17)

where $H_i$ and $H_{main}$ are the pseudo-impurity and target peptide Henry coefficient respectively. However, in this case, the impurities peaks were not visible in isocratic conditions because of their very low concentration. The pseudo-impurities selectivity was therefore evaluated from gradient elution experiments using the Yamamoto method [24]. A series of gradient elution experiments with varying gradient slope were carried out. The Henry coefficient parameters
\( \alpha_{1,j} \) and \( \alpha_{2,j} \) (see Eq. (9)) were then evaluated using the following relation:

\[
G_\phi = \frac{(c_{\text{AcN elution},i})^{\alpha_{2,j}+1}}{[\alpha_{1,j}([\alpha_{2,j}]+1)]}
\]  

(18)

where \( c_{\text{AcN elution},i} \) is the acetonitrile concentration at the elution of component \( i \) and \( G_\phi \) is the normalized gradient slope:

\[
G_\phi = \frac{S \cdot \epsilon_i \cdot V_c}{Q}
\]  

(19)

with \( S \) being the gradient slope. A log–log plot of Eq. (18) is shown in Fig. 10. The linear regression of these data leads to the estimation of the Henry coefficient parameters \( \alpha_{1,j} \) and \( \alpha_{2,j} \) for all the impurities considered. The corresponding parameter values are summarized in Table 4 and the pseudo-impurities selectivities are shown in Fig. 11 as a function of the acetonitrile concentration. It can be seen that the pseudo-impurities selectivities are strongly dependent on the acetonitrile concentration. In particular, the pseudo-impurities L1 and H1 are converging toward the target peptide when the acetonitrile concentration decreases, whereas the opposite behavior is observed for the pseudo-impurity L2.

![Graph showing \( G_\phi \) vs. \( c_{\text{AcN elution},i} \) for the impurities and the target component.](image)

**Fig. 10.** \( G_\phi \) vs. \( c_{\text{AcN elution},i} \) for the impurities and the target component. (solid line and triangle) target peptide; (dashed-dotted line and circle) impurity L2; (dashed line and square) impurity L1; (dotted line and diamond) impurity H1.

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adsorption isotherm parameters ( \alpha_{1,j} ) and ( \alpha_{2,j} ) for the key impurities determined by the Yamamoto method.</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>( \alpha_{1,j} ) [-]</td>
</tr>
<tr>
<td>L1</td>
</tr>
<tr>
<td>L2</td>
</tr>
<tr>
<td>H1</td>
</tr>
</tbody>
</table>

4.2.4. Impurity elution in overloaded conditions

Due to the similar chemical structures of the impurities and the target peptide, it is reasonable to assume that they exhibit a very similar adsorption behavior. Accordingly, also the adsorption behavior of the impurities has been described with a two sites competitive adsorption isotherm:

\[
H_{2,j}\epsilon_j \left(1 + \sum_{i=1}^{n} K_{ij}(H_{2,i}/(2 \cdot q_{\text{sat},i}) \epsilon_i) + \sum_{i=1}^{n} \sum_{k=1}^{n} I_{ik}(H_{2,i}/(2 \cdot q_{\text{sat},i}))\epsilon_i(q_{\text{sat},i}/(2 \cdot q_{\text{sat},k})\epsilon_k)ight)
\]  

(20)

coupled with Eq. (13)–(16) to account for the effect of the acetonitrile on the values of the isotherm parameters. Due to the limited amount of impurities available for model calibration, it was not possible to determine the parameter values for each impurity independently. A certain number of simplifications were therefore made in order to reduce the number of fitting parameters.

First of all, it was assumed that the abundance of the adsorption sites of type 1 and 2 was the same for every component in the system [2]. Therefore, the values of the parameters \( \alpha_{1,j} \) and \( \alpha_{2,j} \) in Eq. (13) are the same for the impurities and for the target peptide. Moreover, it was assumed that the impurity saturation capacity was equal to the one of the target peptide [2]. This assumption implies that the parameters \( \alpha_{1,j} \), \( \alpha_{2,j} \), \( q_{\text{sat},i} \), and \( q_{\text{sat},i} \) are the same for all the components in the system. This is indeed a reasonable assumption since the saturation capacity is mainly dependent on the peptide size and the available surface area [14].

The only unknown parameters remaining are then \( \alpha_{2,j} \) and \( \alpha_{2,j} \) in Eq. (16). These parameters characterize the peptide–peptide, peptide–impurity and impurity–impurity interactions in the adsorbed state. It can be expected that these parameters do indeed depend on the chemical structure of the involved chemical species and are therefore specific for each binary-interaction. Since there are 4 components in the system (i.e. the target peptide and three key impurities), we have \( 4! = 24 \) binary interactions and a total of 48 parameters to be estimated. Again we need some reasonable assumptions in order to reduce the complexity of the system. First of all, it was assumed that the effect of the adsorption strength on the binary interactions was the same for all interactions. Parameter \( \alpha_{2,j} \) is therefore the same for every binary-interaction. The second key assumption is that the impurity–impurity interactions are negligible. Since the impurities are present in very small amounts their interaction is in fact going to be negligible. Therefore only the interactions between the target peptide and the impurities are accounted for, thus reducing the number of unknown parameters to 3. In particular, the parameter \( \alpha_{2,j} \) describes the interaction between the pseudo-impurity L1 and the target peptide, the parameter \( \alpha_{2,j} \) the interaction between the pseudo-impurity L2 and the target peptide and the parameter \( \alpha_{2,j} \) the interaction between the pseudo-impurity H1 and the target peptide.

Considering the complexity of the system and the large number of assumptions, it is not really needed to go through a rigorous fitting procedure to evaluate the parameters \( \alpha_{2,j} \), \( \alpha_{2,j} \), and \( \alpha_{2,j} \). First of all, these were set to the same value as the corresponding ones of the target peptide, i.e. \( \alpha_{2,11,j} = \alpha_{2,12,j} = \alpha_{2,3,j} = 392 \) (see Table 2).

Then, an overloaded injection of the feed in isocratic conditions was considered. The corresponding experimental elution profile is compared to the model prediction in Fig. 12. It is seen that the model prediction is not in good agreement with the experimental elution profile. The parameters \( \alpha_{2,11,j}, \alpha_{2,12,j} \), and \( \alpha_{2,3,j} \) were then adjusted manually until getting a reasonable agreement with the experimental values. It was found that the parameters \( \alpha_{2,11,j} \) and \( \alpha_{2,12,j} \) have to be decreased down to 350 and 300 respectively, while parameter \( \alpha_{2,3,j} \) has to be increased up to 480. The comparison between the experimental elution profile and the model prediction with the so adjusted parameter values is shown in Fig. 13.

![Graph showing selectivity for AcN vs. AcN concentration.](image)

**Fig. 11.** Selectivity of the pseudo-impurities. (dashed-dotted line) impurity L2; (dashed line) impurity L1; (dotted line) impurity H1.
Fig. 12. Overloaded injection of the crude peptide in isocratic conditions. Load = 7.5 g/L; c\textsubscript{0} = 26.52 (v/v); comparison between the experimental data (symbols) and the model prediction (lines): dashed line and square target peptide; (dotted line and diamond) impurity L2; (dotted line and triangle) impurity L1.

Fig. 13. Overloaded injection of the crude peptide in isocratic conditions. Load = 7.5 g/L; c\textsubscript{0} = 26.52 (v/v); Comparison between the experimental data (symbols) and the model prediction (lines): (thick solid line) calibrated UV signal; (thin solid line and triangle) target peptide; (dashed-dotted line and circle) impurity L2; (dashed line and square) impurity L1; (dotted line and diamond) impurity H1.

5. Model validation against experimental data

Two complementary approaches can be used to check the reliability of a chromatographic model. The first one is to verify the reliability of the adsorption equilibrium model against equilibrium data measured with alternative method, such as breakthrough experiments and batch uptake experiments. In Section 5.1, a method is proposed to evaluate the total saturation capacity of the column while using a minimum amount of peptide, which can then be compared with the model prediction. In the second approach, described in Section 5.2, the purity–yield Pareto curves predicted by the model are compared to the corresponding experimental data.

5.1. Adsorption equilibria

The verification of the entire adsorption equilibrium model determined requires not only a significant experimental effort but also a significant amount of target peptide which is typically not available in the early stages of the process development. It is therefore advisable to directly check some selected model parameters. In this section, a method is proposed to evaluate the total saturation capacity of the column using a minimum amount of target peptide.

As discussed above, we assumed that the peptide saturation capacity is maximum at very large Henry coefficient values (i.e. when the acetonitrile concentration is very low). According to Eqs. (14) and (15), the total saturation capacity in 100% buffer A1 is in fact equal to $q_{\text{sat,max}} = q_{\text{sat1,max}} + q_{\text{sat2,max}} = 264 \text{g/L}$. This value could be easily checked by performing a breakthrough experiment in 100% buffer A1, but this would have required a very large amount of peptide (~400 mg), if the experiment was to be carried out on a standard 4.6 mm × 250 mm column. Therefore, this particular breakthrough experiment was carried out on a 2.1 mm × 50 mm column, thus limiting the amount of peptide to about 17 mg. Thus, from a breakthrough experiment using the pure peptide in 100% buffer A1 on the 2.1 mm × 50 mm column at a flowrate of 0.2 ml/min, the saturation capacity was estimated from the breakthrough time $t_{\text{BTC,i}}$, as follows:

$$q_{\text{sat.max,i}} = \frac{Q \cdot t_{\text{BTC,i}} - V_e \cdot (1 - \epsilon_i) \cdot c_1}{V_c \cdot (1 - \epsilon_i)}$$

(21)

The so obtained saturation value is 254 g/L, which saturation value is only about 3.8% from the model prediction.

It is worth mentioning that the breakthrough experiment was repeated with the crude peptide instead of the pure one. In this case, the total concentration of peptide (i.e. target peptide + impurities) was considered in the calculations since both the target peptide and the impurities adsorb on the stationary phase. A 2% variation was obtained with respect to the case with the pure peptide. This shows that this model validation test can be run also in the very early stage of the process development when only the crude peptide is available.

5.2. Chromatographic column performance

Before using the model for process development, it is important to assess its prediction reliability. In particular, we have to make sure that the model provides reliable estimate of the process yield and productivity for different operating conditions. In this section, an overloaded gradient elution typical of industrial operation was carried out. During the experiment, fractions were collected every 0.75 min and in each one of them the target peptide and the impurity concentrations were measured. The yield and purity were then calculated for all possible fraction pooling and a purity–yield Pareto curve was determined. The results are compared to the model predictions in Fig. 14. It can be seen that the model predictions are in very good agreement with the experimental data. The chromatographic model developed in this work can therefore be used to describe the target peptide purification process and then to perform its optimization.
6. Model based process optimization

In a preparative chromatographic process, there are often several competing objective functions, such as purity, yield, productivity, product pool concentration and buffer consumption, that need to be optimized [25]. Of course, it is unlikely that the same values of the design parameters, such as loading, acetonitrile gradient, can simultaneously lead to the optimal values for all the objective functions. It is therefore important to determine the trade-off between all of them by performing a multi-objective optimization.

In this work, the multi-objective optimization was done using an Elitist Non-dominated Sorting Genetic Algorithm routine (NSGA-II) shared by Seshadri on the Matlab Central [26]. The population was set to 100 individuals and the number of generations created was between 30 and 40 depending on the number of design variables.

The objective functions used are the productivity

\[ P = \frac{m_{p,\text{out}}}{\tau_{\text{exp}} \cdot V_C} = \frac{m_{p,\text{in}} \cdot \gamma}{(t_{\text{load}} + t_{\text{grad}} + t_{\text{reg}} + t_{\text{req}}) \cdot V_C} \]  

and the process yield \( \gamma \)

\[ \gamma = \frac{m_{p,\text{out}}}{m_{p,\text{in}}} \]  

where \( m_{p,\text{in}} \) and \( m_{p,\text{out}} \) are the mass of target peptide loaded and recovered from the chromatographic column, respectively. \( \tau_{\text{exp}} \) is the total process time, \( t_{\text{load}} \) the loading time, \( t_{\text{grad}} \) the gradient duration, \( t_{\text{reg}} \) the time required for the column regeneration at high acetonitrile concentration and \( t_{\text{req}} \) the time required for the column reequilibration between two consecutive runs. The regeneration time \( t_{\text{req}} \) was assumed to be 10 min as well as the reequilibration time \( t_{\text{reg}} \). Since the target peptide concentration in the feed \( c_{p,\text{feed}} \) was fixed at 3.5 g/L. The loading time \( t_{\text{load}} \) depends only on the amount of feed loaded in one chromatographic batch \( L \) expressed in mg/mL column volume:

\[ t_{\text{load}} = \frac{L \cdot V_C}{c_{p,\text{feed}} \cdot Q} \]  

In order to complete the formulation of the optimization problem, we have to consider the constraints that the process has to satisfy. The most important ones refer to the target peptide purity which has to be larger than a specified value and to the maximum amount tolerable for some of the critical impurities. Next, we need to impose limits on the operating variables due to limitation on the specific hardware. Typical examples are the limits on the flowrate values imposed by the characteristics of the installed pumps.

6.1. Linear gradient optimization

Let us consider a first process optimization example where we assume to use a linear gradient of the acetonitrile concentration. Four design parameters were selected in this optimization: the loading \( L \), the gradient duration \( t_{\text{grad}} \) and the initial and final acetonitrile concentration values \( c_{\text{AcN,in1}} \) and \( c_{\text{AcN,final}} \). The productivity–yield Pareto obtained from the multi-objective optimization is shown in Fig. 15. These results have been obtained using as a constraint that the purity of the target Pareto should not be smaller than 98%. One of the operating points along the Pareto curve was then selected and reproduced experimentally. In particular, the calculated point indicated by the red open triangle was selected and its experimental verification was performed by using the same operating conditions used in the simulation, leading to the experimental performance indicated by the red closed triangle. It can be seen in Fig. 15 that experimental and model results are in good agreement.

6.2. Bi-linear gradient optimization

A bi-linear gradient of the acetonitrile concentration can be used instead of the linear one to increase the number of design variables and therefore potentially improve the separation. In a bi-linear gradient, six design variables can be optimized (instead of four for the linear gradient): the loading \( L \), the duration of gradient 1 and 2, \( t_{\text{grad1}} \) and \( t_{\text{grad2}} \), the initial acetonitrile concentration of gradient 1 and 2, \( c_{\text{AcN,in1}} \) and \( c_{\text{AcN,in2}} \) and the final acetonitrile concentration of gradient 2, \( c_{\text{AcN,final2}} \). In Fig. 15, the productivity–yield Pareto curve obtained for the bi-linear gradient optimization is superimposed to the one obtained for the linear gradient elution as discussed in the previous section. One of the operating points along the Pareto curve was selected and also in this reproduced experimentally. The experimental verification is again in good agreement with the model prediction. In conclusion, the results reported above indicate the good reliability of the model developed in this work despite the complex adsorption behavior of the peptide under consideration.

6.3. Optimal batch purification selection

Based on the results of the multi-objective optimization of the purification process shown in Fig. 15, it is possible to identify the most convenient process operating conditions. Of course in the case under examination it can be concluded that the introduction of a bi-linear gradient allows further improvement of the process performance only in the high productivity region and to a modest extent. The two Pareto curves do not in fact differ significantly. Of course, the optimal batch purification process selected for a large scale industrial application does not only depend on the yield and productivity trade-off. Additional constraints have to be accounted for in the process selection, besides the product purity. These include reproducibility, sensitivity to process disturbances (robustness) and ease of implementation. For example, in this particular case most likely the linear gradient operation would be selected over the bi-linear gradient because of its higher simplicity and comparable performance.
7. Conclusion

A general procedure for the model-based optimization of the chromatographic purification process of a polypeptide crude mixture is presented in this work and applied to an industrial crude as an illustrative example. First the transport properties of the packed bed were determined by measuring a Van-deemter curve. Secondly, a procedure for evaluating the adsorption equilibria was developed. It was shown that the peptide under examination exhibits very strong peptide–peptide interactions in the adsorbed phase and that its adsorption equilibria could be modeled with a two sites multi-component equilibrium model. The first adsorption site is described by a Langmuir isotherm and dominates at low peptide concentrations. The second one is instead described by a Moreau adsorption isotherm and prevails at higher peptide concentrations. The variation of the adsorption equilibrium isotherm parameters with the modifier concentration was accounted for. In particular, it was found that the Henry coefficients of the two sites were strongly decreasing with the modifier concentration. This decrease was modeled using the stoichiometric displacement model. It was shown that the saturation capacity of the first site was very low and not affected by the modifier concentration. On the other hand, the saturation capacity of the second site was strongly decreasing with the modifier concentration. Finally it was found that the parameter representing peptide–peptide interactions in the adsorbed state increases with the modifier concentration following a power function.

The chromatographic behavior of the impurities was described by lumping them into pseudo-impurities and by estimating their adsorption equilibrium isotherm parameters directly from the corresponding ones of the target peptide. The selectivity of the key-impurities was regressed from gradient experiments using the Yamamoto method. About the overloaded conditions, it was assumed that the impurity saturation capacity was equal to that of the target peptide. This is due to the fact that all impurities have a very similar structure to the target peptide. As a final step for the development of the impurity equilibrium model, the peptide–impurity interaction parameter was tuned to match the impurity elution profiles in a preparative elution chromatogram.

The so obtained model was validated against experimental data. First the total saturation capacity value obtained from the model was confirmed by performing breakthrough experiments at very low acetonitrile concentration. Then the reliability of the prediction of the chromatographic model was assessed by comparing experimental and modeled yield–purity Pareto curve of an overloaded gradient elution typical in industrial operations.

At this point the developed model can be used to perform Pareto optimizations of the process. This is a very useful tool for process optimization which has been used in the particular case under examination to investigate operating conditions that optimize yield and productivity of the target peptide while maintaining its purity within specifications. With illustrative purposes the option of using a linear or a bi-linear gradient of the modifier concentration has been quantitatively investigated.

The developed procedure is sufficiently general to be adapted to the purification of most peptide crudes, leading to significant performance improvements, particularly on the industrial scale.

Acknowledgment

This work was financially supported by Lonza AG, Visp, Switzerland.

References