Doctoral Dissertation
Doctoral Program in Chemical Engineering (34th Cycle)

Analysis of a chromatographic purification process: an experimental and modelling combined approach

Elena Lietta

* * * * * *

Supervisors
Prof. Antonello Barresi
Prof. Marco vanni
Dr. Alessandro Pieri

Doctoral Examination Committee:
Prof. Vincenza Calabrò, Referee, Università della Calabria
Prof. Barbara Stella, Referee, Università di Torino

Politecnico di Torino
January, 2022

I hereby declare that, the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

…….....

Elena Lietta

Turin, February15, 2022

Summary

To bring a new biopharmaceutical product to market is an expensive and time-consuming process. Traditional biopharmaceutical process development is based on experimental work by DoE, trying to explore all conditions to find an optimum and gaining a partial knowledge of the process. By traditional approach an optimal and economic process must be found exploring a large number of parameters. In the downstream processing the most used purification process is chromatography. Integrating the standard procedure with a mechanistic model can increase our knowledge, speed up the process development and the scale up/down.

The aim of this study, carried out in the GSK Vaccines laboratories in Siena and in Politecnico di Torino, is to identify the main principles guiding the chromatography behaviour and describe them with a first-principle model. The test case chosen for this study is the polishing step of a recombinant protein, that consists in a Hydrophobic Interaction Chromatography (HIC).

Since the hydrophobic adsorption behaviour is still not very clear, a thermodynamic study is performed using different commercial proteins as case study with hydrophobic resins. Static methods (high-throughput experimentations) and dynamic methods (breakthrough curves) are used to obtain proteins' adsorption isotherms and the data obtained with the two methods are compared. Two different modelling approaches are then evaluated and compared.

The predictive approach consists in finding isotherms parameters from the fitting of experimental adsorption data (the experimental data are fitted with different isotherm laws) and mass transfer parameters from literature correlations and use them with an in-house code to simulate bind-elute tests. The in-house code can solve the system of Partial Differential Equations describing the component behaviour in the column. The simulated and experimental results of bind-elute tests are compared.

The estimative approach exploits a commercial software (DSPX from GoSilico) that performs the curve fitting using optimization algorithms that minimize the difference between experimental and simulated chromatograms, estimating the model parameters. To perform the parameter estimation, the bind-elute experimental chromatograms are used. The model found is validated simulating a run that is performed in operating conditions different from those used for the parameter estimation.

To investigate the impact of additional components on single-component adsorption behaviours, a high-throughput procedure is also applied to mixtures of commercial proteins. Binary mixtures of commercial proteins are tested in different conditions of salt and protein ratios to investigate their behaviours in hydrophobic interaction chromatography. The aim is to mimic a typical industrial chromatographic step where a target protein must be separated from mixtures of other proteins, like the industrial test case chosen for this study. The multi-component adsorption isotherms are determined with a full high-throughput procedure, exploiting microfluidic capillary electrophoresis in a high-throughput platform for the analysis. This method resulted quick and efficient, with an adequate accuracy considering the advantages of the high-throughput set up: very small amount of sample is needed and the time of test and analysis are very short.

Furthermore, to define a suitable model for the industrial chromatographic step chosen, DSPX is exploited to model the industrial process. Several runs were performed manipulating the process parameters that mainly affect the separation. Experiments are coupled by offline measurement exploiting Size Exclusion Chromatography (SEC) to determine concentrations and purities of the species involved in the purification process. *In silico* simulations are performed on these experiments to develop the model. DSPX is more accurate in describing the industrial process and requires less information and experimental work. The lesser need of experimental work to develop a model of an industrial process is crucial because it saves time and materials, that often are expensive and not available in large amount.

In silico models appear to be able to provide sufficiently accurate information about the process and can help the scale up/down and process development with a low workload.

Acknowledgment

I would like to acknowledge GlaxoSmithKline Biologicals SA and the Science Academy of GSK Vaccines of Siena (Italy) for supporting this PhD project. A special acknowledgement to Alessandro Pieri, Antonio Gaetano Cardillo and Elisa Innocenti for guiding and supporting me in these year from both scientific and organizational point of view. You and all the b23 team have taught me a lot, being always available to answer my questions and needs.

I acknowledge my university supervisors Antonello Barresi, Marco Vanni and Roberto Pisano for having accompanied me on this path and having participated in the drafting of the thesis, of the papers and for the support in experimental design. Thank you for being a guide even when things seemed impossible to me.

I would like to also thank Lorenzo Durbiano, whose results gained for his Master Thesis are part of this work. Working with you has been very easy and funny. I wish you the best for your future.

This work was sponsored by GlaxoSmithKline Biologicals SA. Elena Lietta is a PhD student at Politecnico di Torino and participates in a post graduate studentship program at GSK.

Contents

Chapter 1

Introduction and theoretical background

1.1 The chromatographic process

Liquid chromatography is a purification process constituted by two different phases: the stationary phase represented by a packed solid that can also be porous, and a mobile phase represented by a liquid containing the components that must be purified. During a chromatographic process the liquid phase passes through the packed bad and the component in the liquid phase interact with the solid one. The components in the liquid phase are separated based on their characteristics and their interactions with the solid. The chromatographic purification exploits different mechanisms to separate components. The Size Exclusion Chromatography bases the separation on the different size and shape of the components in the mixture. Ion Exchange (IEX) chromatography exploits the electrostatic attraction and repulsion between compounds with opposite charge and compounds with the same charge. Affinity chromatography is a highly specific chromatographic technique which aims to selectively interact with a certain component, generally used for antibody purification. Adsorption chromatography bases the separation on the different ability of the components to form weak interactions with a given stationary phase. Hydrophobic Interaction Chromatography (HIC) is a particular type of adsorption chromatography that exploits the different interaction between a hydrophobic stationary phase and the hydrophobic groups on the surface of the components.

Based on the scale and the aim of the chromatographic purification process we can distinguish between analytical and preparative chromatography. The aim of the analytical chromatography is to separate components to quantify the concentrations and purities of them. Analytical chromatography is usually in lab scale dimensions and the amount of sample injected is quite small, purity and resolution obtained with this methodology are high. For this scope, the high-performance liquid

chromatography (HPLC) is exploited, which is characterized by a high contact surface between the stationary and mobile phases, thus allowing to obtain, for the same volume of the equipment, a better separation of the different components in the liquid.

On the other hand, the preparative chromatography aims to separate a mixture to obtain a final product with the highest possible purity, considering the cost of the process. The preparative chromatography usually exploits bigger scales (pilot and production scales) and the separation efficiency is lower than the analytical one.

A chromatographic purification process is constituted by several subsequent phases. At first, the wash of the column is performed to remove the storage buffer from the column. Then the equilibration is performed, with special buffer solution, to prepare the column to interact with the component. When the column is prepared the loading phase is performed and the sample that must be purified is loaded in the column. Once the sample is loaded in the column and adsorbed to the solid, the elution is performed based on the characteristics of the sample and the elution method chosen (isocratic or gradient). At the end, different wash stages can be performed to remove any residue left in the column. Each run is then followed by a Clean In Place (CIP) run to sanitize the column and store it with the adequate solutions.

The chromatographic purification process, especially in case of preparative chromatography, can be classified also based on the retention that is exploited during the process. In particular, in bing-elute chromatography the components of interest are retained to the solid during the load phase and then removed with different method in the elution phase. In the case of breakthrough chromatography, during the load all impurities that must be removed remain retained in the column while the product of interest passes through the column.

This work focuses on the hydrophobic interaction chromatography when it is exploited to purify proteins. Hydrophobic interaction chromatography uses hydrophobic actives sites (butyl and phenyl mostly) on the solid phase to interact with hydrophobic groups on the surface of component in the mobile phase. The component interacts with the active sites with a reversible interaction and the binding or no binding of the component on the solid phase can be controlled by the ionic strength of the mobile phase. The salts used to promote the binding of the protein to the solid surface are called kosmotropic salts. This kind of salts strongly bind the water molecules, as they have high polarity, in such a way that the water does not bind to the protein and the ligand on the surface, having also a stabilizing effect on the protein (Lienqueo et al., 2007).

The mechanisms that drive the hydrophobic interaction chromatography are still not clear, different were the theories developed in the last years. The type of the salt and its concentration affect the behaviour between the component and the solid surface. The solvophobic theory was developed in 1960s by Sinanoglu and

Abdulnur (1965), describing the formation of cavities in the solvent to allow the component to enter in solvent cavity and interact with it. From the study on some proteins with different types of salt turned out that the solvophobic theory is not able to explain some differences in retention (Fausnaugh and Regnier, 1986). In the preferential interaction theory (Arakawa and Timasheff, 1984), the effect of salt concentration is related to the number of water molecules and salt ions that are moved for the protein absorption. Chen and Sun (s2003) developed a theory that describes how the salt ions create hydrophobic areas by desolvation on the protein and the ligand, followed by the interaction between the dehydrated protein and ligand (Eq. 1.1).

$$q = \frac{K_s K_p c_{salt}^{\alpha} [\Lambda - (n + \sigma) q]^n}{1 + K_s c_{salt}^{\alpha}} \tag{1.1}$$

Chen and Cramer (2007) tried to find a relationship between adsorption behavior and isotherms. They highlighted three classes of isotherm behavior on some proteins, related to the increasing salt concentration, with different hydrophobic interaction resins. *Type 1* showed a constant increase of adsorption with increasing salt concentration, corresponding to a moderate hydrophobicity of the system. *Type 2* behavior corresponds to a moderate adsorption at low salt concentration and a steep increase at higher salt concentrations; the high salt concentration can affect the protein structure, exposing more hydrophobic surface of the proteins. *Type 3* proteins have a very low adsorption at low salt concentrations and a high increase in adsorption at high salt concentrations, as if the high concentration of salt causes the unfolding of the protein, exposing its hydrophobic sites more.

Mollerup et al. (2008) developed a different theory to consider the influence of pH and solvents in the binding, that was not considered for the solvophobic cavity model (Rodler et al. 2019). The model of Mollerup and co-workers (2008) explains how the adsorption behaviour only depends on protein and salt type, independently from the properties of the solid phase. Indeed, in a previous work (Mollerup, 2006), he suggested that the intrinsic hydrophobic interaction between the protein and the resin ligand is not affected by the ionic strength.

Deitcher et al. (2010) presented a theory that considers the protein activity and the release of water molecules from hydrophobic interfaces. The model of Deitcher and co-workers (2010) does not consider the interaction of the protein with the base of the medium and the displacement of the salt ions is not included. For these

reasons, the model describes well the protein adsorption when the “salting-out” behaviour is dominant, under strongly retained conditions.

In the work of Wang et al. (2016) a model that considers the equilibrium between well-ordered and bulk-like-ordered water molecules in the hydrophobic sites is presented. The salt concentration dependency is considered in the number of hydration ions.

1.2 Biopharmaceutical products

In the last years biopharmaceutical products have assumed a high importance within the pharmaceutical industry. Biopharmaceutical products are mainly represented by therapeutic proteins like vaccines and antibodies (Baumann and Hubbuch, 2017). The process to produce a biopharmaceutical product is constituted by different phases. In the upstream phase bacteria or cell culture are exploited to produce the component of interest. In the downstream phase, the upstream material is purified with different stages to have a final content and purity of the product of interest sufficiently higher. In the downstream processing the purification of the biopharmaceutical product is articulated in different phases that mostly include preparative chromatography. Depending on the type of protein produced, and in particular if it is intracellular or extracellular the downstream phase sees different processes subsequently. If the product of interest is intracellular the cellular membrane of the cells or bacteria must be broken to obtain the protein free in mixture; this step exploits homogenization to break the cells and centrifugation to separate the cell membrane and other heavier compounds from the harvest broth that contains the protein of interest and must be purified. The material coming from the lysis, then, is subjected to different stages of filtration and chromatography to obtain the required purity of product of interest. In particular, the chromatographic purification is usually articulated in three steps: capture, intermediate and polishing. The different kinds of chromatography are combined in these steps to remove all the impurities coming from the upstream.

In this work the focus is on the last step of chromatography purification process, the polishing step. Indeed, hydrophobic interaction chromatography is particularly suitable as a polishing step, where, in addition to other impurities, the monomer must be separated from aggregates potentially formed during both upstream and downstream process. In this case the hydrophobic interaction chromatography is the best choice since it exploits the differences in hydrophobicity of the monomer and the aggregates (Queiroz et al., 2001).

Studying the adsorption dynamic of the hydrophobic interaction chromatography allows to optimize the purification and increase the separation efficiency. This is the aim of the mechanistic modelling applied to the development of a process. *In silico* simulations can be performed in order to save sample and

time to investigate how some parameters affect the system. Applying this approach to industrial process development is crucial because the amount of product available to investigate is limited and the timelines of the market are constricted (Close et al., 2014).

The limitations of time and sample for the process development have brought the high-throughput technologies of primary importance in the biopharmaceutical industry (GE Healthcare, 2009). High-throughput technologies allow investigating lot of conditions at the same time with a very small amount of sample and chemicals. The development of a purification process involves the screening of optimal process conditions in terms of buffer solutions, medium, loading etc. High-throughput technology help gaining a big amount of information in the early stage of the development of a process. In the last years, the automation increased the parallelization and miniaturization of the high-throughput technologies, also increasing the reliability of results. High-throughput methodologies range from the analytic scale (few microliters), to the filter plated (hundreds of microliters) to the robocolumn scale (hundreds of microliters to few milliliters).

1.3 Theory

Different mechanisms take place in a chromatographic column during a purification process. The component in the liquid phase move in the column subjected to different mass transfer phenomena and interact with the solid surface due to adsorption phenomena. The components are subjected to convection when in the interstitial liquid and brought to the solid surface and in the liquid inside the pores of the beads of the packed bed. The liquid in the pores of the beads is considered stagnant and the components are moved to the internal solid surface by diffusion. When the component reaches the solid surface, it interacts with the active sites based on the adsorption dynamics that characterize the process (Guiochon et al., 2006).

Different models have been developed that describes the mass transfer and adsorption phenomena, depending on the assumption that has been done and the systems to simulate.

The model considered in this study belongs to the General Rate Models (GRM) group of models (Schmidt-Traub, 2005). GRM are the most detailed models, they consider the axial dispersion and, at least, other two parameters that describe mass transfer in the liquid in the interstitial volume and in the pores. There are a lot of models reported for this category that consider adsorption kinetics, pore diffusion and surface diffusion. In most cases, to model column chromatography, some basic assumptions are made:

- the adsorbent solid is homogeneously packed, the particles are all spherical and with constant diameter;
- density and viscosity of the liquid are constant;
- the process is isothermal;
- the eluent is not interacting with the solid, its contribute is considered in the adsorption isotherm law;
- the liquid inside the pores of the beads is stagnant, there is no convection.

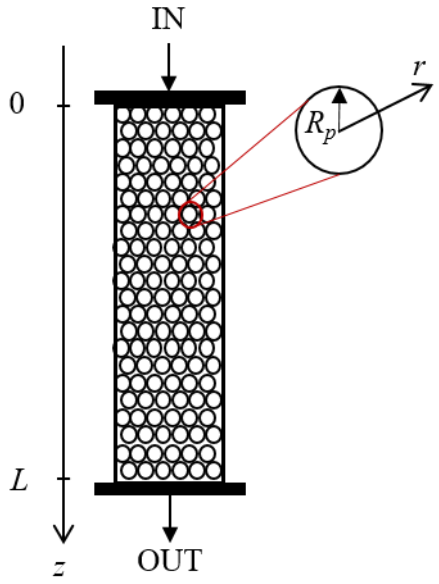


Figure 1.1: Scheme of the column.

In this study, pore diffusion and mass transfer resistance coefficient are considered besides the axial dispersion, and they are assumed as constant. The radial coordinate for the particle is considered, besides the axial coordinate along the column and time *t*. *z* represents the axial coordinate along the column, while *r* is the radial coordinate in the solid particle as highlighted in Figure 1.1.

$$\frac{\partial c_{int}}{\partial t} + v_{int}\frac{\partial c_{int}}{\partial z} = D_{ax}\frac{\partial^2 c_{int}}{\partial z^2} - \frac{3}{R_p}\frac{1-\varepsilon_c}{\varepsilon_c}k_c\left(c_{int} - c_p|_{r=R_p}\right) \tag{1.2}$$

$$\varepsilon_p\frac{\partial c_p}{\partial t} + \frac{\partial q'}{\partial t} = \varepsilon_p D_p\frac{\partial^2 c_p}{\partial z^2} - \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial c_p}{\partial r}\right) \tag{1.3}$$

Equation 1.2 describes the mass balance in the interstitial liquid volume for a single component, while Equation 1.3 is the mass balance of a single component in the liquid inside the pores of the beads. In the equations above, *q'* is the adsorbed

protein concentration referred to the beads volume, constituted by the solid phase and the stagnant liquid inside the pores of the beads.

The boundary conditions on the component concentration in the interstitial liquid are those derived by Danckwerts (1953) and reported in Eq. 1.4 and 1.5:

$$c_{feed}(t) = \left[c_{int} - \left(\frac{D_{ax}}{v_{int}}\right)\frac{\partial c_{int}}{\partial z}\right]_{z=0} \tag{1.4}$$

$$\left[\frac{\partial c_{int}}{\partial z}\right]_{z=L} = 0 \tag{1.5}$$

Regarding the concentration of the component in the pores of the beads, the boundary conditions are written in the following Equations 1.6 and 1.7.

$$\left[\frac{\partial c_p}{\partial r}\right]_{r=0} = 0 \tag{1.6}$$

$$\varepsilon_p D_p \left[\frac{\partial c_p}{\partial r}\right]_{r=R_p} = k_c \left(c_{int} - c_p|_{r=R_p}\right) \tag{1.7}$$

Values of pore diffusion D_p and mass transfer coefficient k_c can be calculated from empirical equations depending on the properties of the component and the column (Guiochon et al., 2006). The viscosity of the salt solution η_B can be found considering the mass fraction of the salt x in the solution and the viscosity of water η_W (Eq. 1.8). The molecular diffusion can be calculated with the Young correlation, derived for proteins, that consider the viscosity of the solvent η_B and the molecular mass of the protein M_A (Eq. 1.9). From the value of molecular diffusion, the pore diffusion D_p can be evaluated (Eq. 1.10). At the end, k_c can be determined with molecular diffusion value and bed properties (Eq. 1.11).

$$\eta_B = \left(1 + \frac{x}{0.3}\right)\eta_W \tag{1.8}$$

$$D_m = 8.31\times10^{-8}\frac{T}{\eta_B M_A^{1/3}} \tag{1.9}$$

$$D_p = \frac{\varepsilon_p}{(2-\varepsilon_p)^2} D_m \tag{1.10}$$

$$k_c = \frac{1.09}{\varepsilon_c}\frac{D_m}{d_p}\left(\varepsilon_c \frac{v_{int} d_p}{D_m}\right)^{0.33} \tag{1.11}$$

To describe the adsorption equilibrium two different adsorption laws are considered that are modifications of the Langmuir law. Indeed, an adsorption law that considers also salt concentration variations is needed when the elution is performed modifying the salt concentration. Thus, the exponential and power Langmuir laws are considered (Eq. 1.12 and 1.13) (Chen and Sun, 2003). Furthermore, in some cases a linear adsorption law, considering an exponential dependency from the salt concentration (Eq. 1.14), was sufficient to describe the adsorption behavior.

$$q = \frac{\lambda b e^{k c_{salt}} c_p}{1 + b e^{k c_{salt}} c_p} \tag{1.12}$$

$$q = \frac{\lambda b c_{salt}^{\alpha} c_p}{1 + b c_{salt}^{\alpha} c_p} \tag{1.13}$$

$$q = a e^{k c_{salt}} c_p \tag{1.14}$$

The adsorbed protein concentration q used in the equations above is referred to the apparent volume that is constituted by the solid and the liquid in the pores and in the interstitial volume. The relationship between q and q' is reported in Equation 1.15. The differences in the reference volume must be taken into account when performing simulations and fitting experimental data, taking into account also the compression factor *CF*, that is generally different in High Throughput and column experiments.

$$q' = \frac{q\ CF}{(1-\varepsilon_c)} \tag{1.15}$$

The modified Langmuir adsorption laws has been proposed to overcome the impossibility of the Langmuir equation of describing the salt concentration variation. The exponential Langmuir law is an empirical correlation and its

parameters do not have a physical meaning. The power Langmuir law can be considered as a reduction of the model developed by Chen and Sun (2003), when the ligand density of the solid is sufficiently low. In this last case, the model parameters have a defined physical meaning: λ is the ligand density, *b* is the lumped equilibrium constant, and it is related to the increasing of hydrophobicity due to the salt increasing and the binding of the proteins to the ligands, α is the salt coefficient that characterizes the number of salt molecules that dehydrate the protein to expose its hydrophobic surface. These two models, obviously, at constant salt concentration behave like the Langmuir isotherm: at low unbound protein concentrations the isotherm is linear, while going to higher protein concentration the isotherms reach a horizontal asymptote when the saturation of the solid is achieved. Increasing the salt concentration, the slope of the linear part of the isotherm increases and becomes steeper; the value of the asymptote, the maximum concentration, also increases when the salt concentration does.

Another adsorption law that can be used for the hydrophobic interaction chromatography is the law developed by Mollerup et al. (2008). It describes the adsorption in hydrophobic interaction chromatography as reversible interactions between a component and the hydrophobic ligands on the solid surface. Mollerup et al. considered that the ligands can be accessible or not because of steric hindrance. The adsorption equation developed by Mollerup et al. is written below (Eq. 1.16).

$$k_{kin}\frac{dq^*}{dt} = k_{eq}\left(1-\frac{q^*}{q^*_{max}}\right)^n c_p \exp\left(k_s c_{salt}+k_p c_p\right) - q^* \tag{1.16}$$

The model developed by Wang et al. (2016) (Eq. 1.17 and 1.18) is also suitable to describe hydrophobic interaction chromatography. This model considers that, when adsorption takes place, water molecules organize themselves in well-ordered structures. This phenomenon is called hydrophobic hydration and it is considered while building the adsorption law.

$$k_{kin}\frac{dq^*}{dt} = k_{eq}\left(1-\frac{q^*}{q^*_{max}}\right)^n c_p - q^{*1+n\beta} \tag{1.17}$$

$$\beta = \beta_0 \exp\left(\beta_1 c_{salt}\right) \tag{1.18}$$

We must consider that q^*, in the above equations, is the adsorbed protein concentration referred to the solid skeleton volume and it is related to q and q' as written in Equation 1.19.

$$q^* = \frac{q'}{(1-\varepsilon_p)} = \frac{q\ CF}{(1-\varepsilon_c)(1-\varepsilon_p)} \tag{1.19}$$

For more complex systems, like the in-house protein purification process also investigated in this study, and when the parameters must be found by fitting of the experimental chromatograms, the GRM could be too complex, and a simpler model can be used in order to speed up the simulations, achieving equally satisfactory results. Thus, when modelling the in-house purification process, the Lumped Rate Model (LRM) was used (Eq. 1.20).

$$\frac{\partial c_p}{\partial t} + \frac{1-\varepsilon_p}{\varepsilon_p}\frac{\partial q^*}{\partial t} = \frac{3}{R_p} \cdot \frac{k_{eff}}{\varepsilon_p}(c_{int} - c_p) \tag{1.20}$$

This model summarizes the internal and external mass transfer resistance coefficient in a unique coefficient k_{eff} (Schmidt-Traub et al., 2005). The diffusion coefficient inside the pores is assumed infinite thus no concentration distribution in the beads is considered.

Furthermore, in the industrial test case investigated a mixture of proteins is considered, where the target product must be purified from the impurities. Thus, the multi-component model of Wang and co-workers (2016) was used with the inclusion of pH dependency for the equilibrium parameter (Eq. 1.21).

$$k_{kin,i}\frac{dq_i^*}{dt} = \left(k_{eq,o,i}\exp\left(k_{eq,1,i} \cdot pH \cdot k_{eq,2,i} \cdot pH^2\right)\right)\left(1 - \sum_{j=1}^{N}\frac{q_j^*}{q_{max,j}^*}\right)^{n_i} c_{p,i} q_i^{*1+n_i\beta_i} \tag{1.21}$$

1.4 Aim and scope

This work wanted to evaluate the modelling and experimental combined approach as a tool to speed up process development. Adsorption mechanisms of the chromatographic separation were investigated to gain knowledge on separation dynamics in hydrophobic interaction chromatography.

In this work an adsorption study was performed on some commercial protein to investigate their behaviour with hydrophobic resin. Equilibrium adsorption isotherms were evaluated experimentally with both static and dynamic methods at different salt concentrations. Experimental adsorption isotherms were determined with high-throughput methodologies in 96-well plates, and with frontal analysis. Equilibrium adsorption isotherms so obtained were compared in order to determine the consistency of the two experimental methods.

The mechanistic modelling, coupled with experimental tests, was then exploited to investigate the behavior of the proteins in the column, studying the elution behavior. To do so, bind-elute tests were performed with some commercial proteins, investigating different salt concentrations and elution methods. These experimental runs were used to compare two different modeling approaches.

In the first approach, the bind-elute tests were simulated with an in-house Fortran code that solves the PDEs system and gives the outlet concentration profiles. Different adsorption models were tested that consider the dependency of model parameters from the salt concentration. Adsorption model parameters were evaluated by fitting of experimental adsorption isotherms. Simulated and experimental bind-elute tests were then compared to evaluate the robustness of the different models.

The second approach exploits a commercial software that estimates model parameters by curve fitting of experimental chromatograms. After the parameter estimation, the model was validated with an additional experimental run to assess the robustness of the model.

The two approaches were compared from different points of view: the experimental effort required, accuracy of elution forecasts in different operating conditions and process knowledge gained from the experiments and the modelling.

The modelling approach was then applied to an industrial in-development project in order to evaluate its behavior. The separation process was more complex in this case, and some assumptions were done in order to simplify the real system, that involves different species. The in-house separation process was also optimized *in silico* manipulating some process parameters to maximize the yield and purity of the target product.

Since adsorption dynamics have a very important role in the separation efficiency of a chromatographic purification of protein mixtures, multi-component adsorption isotherms were evaluated. The high-throughput methodology was exploited to determine equilibrium adsorption isotherms of binary mixtures. To do so, the microfluidic capillary electrophoresis was used to analyze samples. This study wanted also to assess the analytical method in terms of accuracy and reliability for the multi-component adsorption isotherm determination.

Chapter 2

Materials and methods

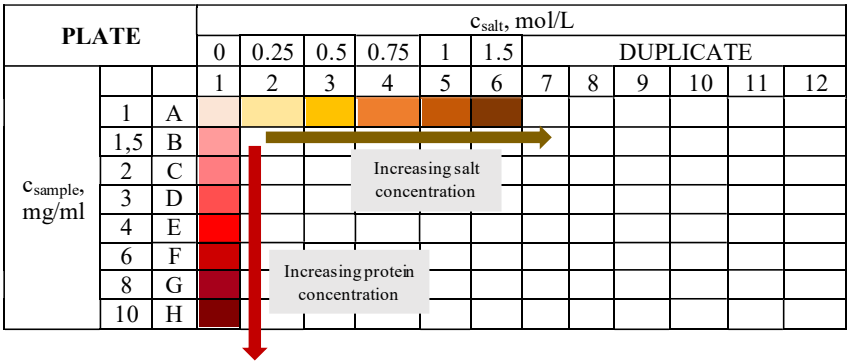
2.1 High-throughput single-component adsorption isotherms determination

The commercial proteins used for the following study are Bovine Serum Albumin (66.5 kDa) from bovine serum, Lysozyme (14.4 kDa) from chicken egg white and Cellulase from *Trichoderma reesei* (45 kDa), α-chymotrypsin from bovine pancreas (25 kDa), β-lactoglobulin from bovine milk (18.4 kDa) and Albumin from chicken egg white (44.5 kDa), that were purchased from Sigma Aldrich (St. Louis, MO, USA); CRM_{197} (58.8 kDa) was available (stored at -20°C) in GSK Laboratories (Siena, Italy). CRM_{197} is a non-toxic mutant of *diphtheria* toxin used as a carrier protein (Bröker et al., 2011).

Potassium monobasic phosphate was purchased from Sigma Aldrich (St. Louis, MO, USA) and ammonium sulfate is from Carlo Erba Reagents (Milan, Italy). Potassium hydroxide used for buffer titration was also purchased from Carlo Erba Reagents. Syringe filters used to filter protein solutions were purchased from Sartorius (Gottongen, Germany). The Butyl Sepharose HP resin bulk and al Butyl Sepharose HP columns were obtained from GE Healthcare (Uppsala, Sweden); the resin has a degree of substitution of 51 $\mu mol_{butyl\ groups}/mL_{resin}$ (data from manufacturer) and a mean bed diameter of 34 μm.

Single component adsorption isotherms were determined experimentally, for the proteins mentioned above, with a high-throughput methodology using AcroPrep 96 Filter Plates with PTFE membrane from Pall Corporation (New York, USA). The stock solution with 10 mg/mL protein concentration was prepared per each protein and filtered with a Minisart Syringe Filter with a 0.22 μm cut-off. Eight protein concentrations were prepared by dilution from the stock, from 1 to 10 mg/mL of protein. The protein solutions were prepared in a 0.05 mol/L potassium monobasic phosphate buffer with six different ammonium sulfate concentrations, from 0 to 1.5 mol/L (layout in Figure 2.1 a); 1.2 and 1.4 mol/L conditions were investigated in a subsequent series of tests (layout in Figure 2.1 b) to integrate the previous series of data. These two last conditions were investigated using the same resin bulk batch to minimize variability. 6 mol/L potassium hydroxide is used for buffer titration to reach neutral pH.

a)



b)

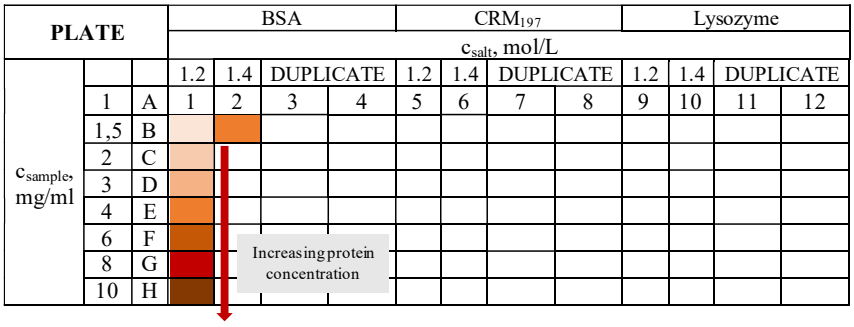


Figure 2.1: Operating conditions layout for filter plates tests. a) describes the layout of tests performed with the range of salt concentration from 0 mol/L to 1.5 mol/L; b) describe the layout of tests performed subsequently with 1.2 and 1.4 mol/L of salt concentration.

The resin bulk is available as slurry of 70% of solid and 30% of ethanol 20%$_{v/v}$. To obtain a slurry that is 50% of apparent volume of the solid, about 20 mL of the resin slurry is poured into a 50 mL graduated cylinder, where it settles by gravity; after the slurry settles down a volume of water equal to the apparent volume of the solid is added. The apparent volume is here considered as the volume occupied by the solid settled for gravity and the liquid present in the interstitial volume of it. The resin slurry is then kept under stirring in a becker in order to make it homogeneous before loading in the plate.

The supernatant is removed from the plate by vacuum using a Vacuum Manifold system (Pall Corporation). The Vacuum Manifold system is a housing for filter plates that can be connected to a vacuum line (or pump) and allows the filtration. The system has a lid with gasket, the collecting plate is placed under the lid, where the vacuum line is connected, while the filter plate with the resin and solutions is placed above the lid. At the bottom of the filter plates there is a membrane and the liquid supernatant in the filter plate is collected in the collection plate under the lid, by vacuum filtration.

Steps of the experimental procedure are the following:

- each well of the filter plate is filled with 50 µL of resin slurry;
- the resin slurry is washed twice pipetting 300 µL of water and removing the supernatant;
- the resin is equilibrated twice with 300 µL of the buffer with the defined salt concentration, then the supernatant is removed again;
- in each well 120 µL of loading solution (the salt concentration in the loading solution must be the same as the equilibration buffer for each well) are pipetted;
- the plate is kept under stirring for 6 hours at 1300 rpm and 25°C to be sure that the equilibrium is reached in each well homogeneously;
- the supernatant is removed from the wells and collected in the collection plates by filtration.

Tests were also performed in the same conditions but with different times of stirring. In particular, 4, 8 and 15 hours of stirring were used. In the case of 4, 6 and 8 hours of stirring no appreciable differences in adsorption were found. When 15 hours of stirring were used, several wells resulted clogged and it was not possible to filter; this may happen because of protein aggregation due to too long incubation time. In order to compare data with literature, the 6 hours incubation time results were considered for the study.

Test were performed also at a lower velocity, 150 rpm, with 6 hours of incubation. No differences were noticed, since the stirring time was sufficiently long to achieve the equilibrium, even with a lower stirring velocity.

The loading and flowthrough solutions are analysed with a Multiscan Sky Spectrophotometer from Thermo Scientific (Waltham, MA, USA) at 280 nm to obtain absorbance values, using Greiner UV-Star 96 well plates from Merck (Darmstad, Germany). It was verified that buffer solutions do not give interference at this wavelength. Each condition of salt and protein concentration, is repeated four times and the average values of free and loaded protein are considered, excluding outliers that were identified by Tukey criteria (Tuckey, 1949).

The total mass balance (Eq. 2.1) (GE Healthcare, 2009) is applied to each well to calculate adsorbed protein concentration from loading and flowthrough concentration values.

$$V_{medium}\, q = V_{liquid}(c_0 - c_{unbound}) \tag{2.1}$$

$$c_0 = \frac{c_{sample}\, V_{sample}}{V_{sample}+V_r} \tag{2.2}$$

$$V_{liquid} = V_{sample} + V_r \tag{2.3}$$

$$V_r = V_{medium}\, 0.6 + 6\, \mu L \tag{2.4}$$

V_{medium} is the apparent volume of the solid, V_{sample} is the volume of loaded protein solution and V_{liquid} is the total liquid phase volume in each well. We must consider the amount of liquid that is retained from the solid phase and the membrane at the bottom of the filter plates. Indeed, the protein solution is diluted when loaded into the well and this must be considered in the mass balance. V_{liquid} includes the retained volume V_r, that has been determined experimentally in a previous work (GE Healthcare, 2009) and is equal to the volume retained by the adsorbent phase (60% of the adsorbent phase volume) plus the liquid volume retained in the filter membrane (6 µL) (Eq. 2.2 to 2.4). c_{sample} is the loaded protein concentration while c_0 is the corrected protein concentration in the well. q is the adsorbed protein concentration in equilibrium with $c_{unbound}$, the unbound protein concentration.

The variability of the bed properties when a different resins batches are used is not well known. It is not clear how it affects the separation efficiency of the process, and rarely its influence is evaluated. To determine this variability, the adsorption behaviour of some commercial proteins interacting with two resin batches with different ligand density is investigated. To do so, equilibrium adsorption isotherms are determined with the procedure described above. The protein used for this study are the α-chymotrypsin, β-lactoglobulin, Albumin and BSA. In this case, the salt concentrations investigated were 0 mol/L, 0.5 mol/L, 1mol/L and 1.4 mol/L ant the plate layout is shown in Figure 2.2. The UV-Vis analysis of the load and supernatant, in this case, were performed with a Spark Multimode Microplate Reader (Tecan, AG, Switzerland).

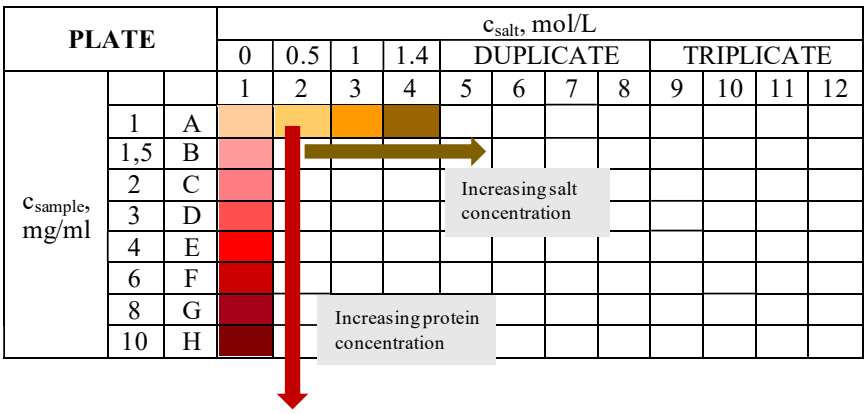


Figure 2.2: Layout of the filter plates carried out for the resin batches variability study

Two different batches of Butyl Sepharose HP resin were used for this study. From the manufacturer handbook of the resin bulk, the ligand density is 40 µmol/mL$_{resin}$. This value is given by the manufacturer as a reference value of ligand density, but each resin batch has a different value. The specific ligand density of the two batches available for this study were 51 (the same resin slurry used for the isotherm determination described before) and 55 µmol/mL$_{resin}$ (data given by manufacturer).

The high-throughput isotherm determination was carried out with both resin batches and, to minimize the experimental variability, per each protein the same load solution was used for the two batches. Adsorption isotherms so obtained are then compared.

2.2 Multi-component adsorption isotherms and assessment of electrophoretic method

Binary mixtures adsorption isotherms are investigated. A batch high-throughput method was used for isotherm determination. Proteins involved in this study were BSA, Lysozyme and CRM_{197}. Single protein stock solutions were prepared at 10 mg/mL. Four total protein concentrations were investigated: 2, 5, 7, and 10 mg/mL obtained by mixing and then diluting single protein stock solutions. For each binary mixture, three different protein ratios were considered: 70%-30%, 50%-50%, and 30%-70%. In all cases, the solutions were prepared at three different ammonium sulfate concentrations: 0, 0.5, 1 mol/L in a 0.05 mol/L of potassium monobasic phosphate buffer. In this case also, a 6 mol/L potassium hydroxide solution was used for titration to reach the neutral pH of the solutions. The experimental procedure was the same as described in section 2.1. In multi-component isotherm determination, the supernatant was removed via centrifugation to minimize contamination between wells.

Microfluidic capillary electrophoresis was exploited to analyze samples and determine protein concentrations in binary mixtures. The concentration analysis was performed with the LabChip GXII Touch protein characterization system from Perkin Elmer (Waltham, MA, USA), that is a high-throughput mode platform. The LabChip system performs determination of the concentration and sizing (via calibration curves) and purity of the sample analyzed. 2 μL of sample are needed for the analysis that must be prepared according to Protein Express Assay Quick Guide (Perkin Elmer, 2019). Filter tubes, ladder and buffer tubes were provided together with the reagent kit from the manufacturer. NuPAGE Sample Reducing Agent from Thermo Scientific was used to reduce samples before analysis. HardShell PCR Plate with 96 wells from Perkin Elmer were used to prepare and analyse samples with the LabChip System. The analysis can be performed in two ways depending on the amount of protein in the sample. With the standard sensitivity mode, 2 μL of protein solution are denatured at 100°C for five minutes with 7 μL of reducing solution and, after denaturation, 35 μL of water are added per each well to stop the reaction. Using the high sensitivity mode, the volume of the protein solution in the sample is higher (7 μL) but the total volume in the well is the same as the standard sensitivity mode. This second analysis modality allows the analysis of samples that have concentrations that would be below the limit of detection when prepared with the standard sensitivity

mode. The choice of analysis mode can be made knowing that the limit of detection of the instruments is 5 μg/mL. We used the standard sensitivity mode since pure proteins were used with known concentrations. Furthermore, Lysozyme has high adsorption in fluorescence; if the sample is too concentrated the peak in the electropherogram is too high and it goes towards saturation of the sensor failing the analysis. Adsorbed protein concentration is calculated with the mass balance (Eq. 2.1) applied to the single protein concentration values on loadings and flowthrough. Calculation made in case of binary mixture are the same made for single-component isotherms determination (Eq. 2.1 to 2.4).

Furthermore, to verify the performances of the microfluidic capillary electrophoresis, the LabChip was tested from different point of view. At first, the separation efficiency of the instrument was verified. The Protein Express Assay LabChip separates proteins by their molecular weight. A 90%CRM_{197}-10%BSA and 10%CRM_{197}-90%BSA samples were analysed in order to verify if the separation occurred correctly. BSA and CRM_{197} were used for the test because they have similar molecular weights (66.4 kDa and 58.8 kDa, respectively).

Additionally, in the case of LabChip, the measurement error was determined, excluding the variability resulting from the preparation of the samples, to assess the precision of the instrument only, and to determine how much the preparation of samples affects the measurement. The test was carried out by preparing a larger amount of sample, according to the procedure described above and dispensing it into ten wells. This procedure should remove some of the variability introduced with the preparation and denaturation of ten samples in ten different wells.

Since the capillary electrophoresis is not such an established method for determination of multi-component adsorption isotherms as absorbance analysis with UV-Vis spectrophotometer is for single-component isotherms determination, an accuracy validation was also performed for the LabChip. A BSA standard sample (BSA standard at 2 +/-0.03 mg/mL calibrated by direct comparison to purified BSA from the National Institute of Standard and Technology, purchased from Thermo Scientific) was diluted and analysed five times with the LabChip and the concentration value so obtained was compared with the value declared from the manufacturer. The BSA sample was diluted 1:1 to reach the concentration of 1 mg/mL since the linearity in fluorescence analysis between concentration and peak area is, for most of the proteins, guaranteed under a concentration of 2 mg/mL. The error made for the dilution of the sample was considered negligible.

2.3 High-throughput method precision investigation

High-throughput methodologies allow investigation of different operating conditions in parallel, wasting a very low amount of sample and chemicals. On the other end, the experimental procedure requires a high experimental effort and

several steps are needed. For this reason, the error affecting the results can be quite high and must be evaluated to assess the precision of the method. Mostly, as the experimental errors depend on the operator precision and conditions that cannot be controlled, it is very difficult to calculate these errors a priori.

In both binary mixture and single-component adsorption isotherms the uncertainty relative to the unbound and adsorbed protein concentration values was calculated in different ways. The unbound protein concentration was directly analysed from the spectrophotometer to obtain absorbance for single-component isotherms and from the LabChip in the case of multi-component isotherms to obtain electropherograms. In both cases, after the analysis, values obtained (absorbance and electropherogram peak area) were converted in concentration data by using a calibration line built with standard solutions. Thus, the error that affects unbound concentration values is the sum of the variability of the measurement equipment and the calibration line error.

The error that comes from the variability of the measurement was obtained analysing the same sample several times and calculating the standard deviation and correlation coefficient of the data so obtained. In the case of absorbance analysis, it is performed on 16 samples of BSA at 1 mg/mL. Standard deviation and %RSDs were calculated for absorbance values. The same study was performed analysing 1 mg/mL BSA samples with the LabChip. In this case the standard deviation and %RSD were calculated for peak area data.

The error that is due to the use of a calibration line was calculated as the average of the relative error of each point of the curve. The calibration curve error was calculated for LabChip and UV-Vis analysis. The final error bars applied to $c_{unbound}$ concentration values (x-axis) consider the sum of equipment measurement variability and calibration curve error.

The adsorbed protein concentration was calculated using the mass balance (Eq. 2.1) from the values of loading and unbound protein concentrations, assuming specific amount of resin and solutions. For this reason, adsorbed concentration values are affected by all experimental uncertainties of the experimental procedure. Evaluating all these uncertainties is complex and sometimes not even feasible (e.g., homogeneity of the slurry, pipetting of the slurry and solutions, retention of the liquid in the plate). Experimental evaluation of the precision of the experimental procedure was performed. Experimental tests were performed in filter plates with the same procedure described in section 2.1, investigating representative conditions. Operating condition considered in this study were 1mg/mL of protein with no salt in the buffer, 10 mg/mL of protein with no salt in the buffer and 10 mg/mL of protein with 1 mol/L of salt in the buffer. Each condition is investigated in 32 wells. In the condition without salt in the buffer, no (or very low) adsorption occurred in the wells for the three proteins, so this condition was chosen to evaluate the experimental error resulting only from manipulation, excluding adsorption. In the case of 1 mol/L of ammonium sulfate in the buffer, the adsorption of the protein to the solid is higher (Chen and Cramer, 2007), hence, variability in the adsorption

behaviour of the protein per each well can be evaluated along with the experimental procedure.

In the case of protein mixtures, the condition investigated for the reproducibility study is 50%BSA-50%CRM_{197} using a total protein concentration of 5 mg/mL with 0.5 mol/L of ammonium sulfate in the buffer. The condition here described is investigated in 30 wells of the filter plate. This combination of lower protein concentration with lower salt concentration, with respect to the single component tests, results in basically the same effect, from the adsorption point of view, of higher salt and protein concentrations. Salt concentration was chosen in such a way that the protein is almost equally distributed between the solid and liquid phases, and the latter has a concentration within the linear range of fluorescence analysis (0-2 mg/mL). This choice was made to avoid diluting the samples that must be analysed, an operation that would increase the manipulations on the samples and therefore the variability. We must consider that protein concentration used for protein mixture reproducibility investigation was four times lower than the single component test. For this reason, the salt concentration used in this case was lower (0.5 mol/L); if a higher concentration were used, the adsorption would be too high, and no peaks would be detected in the flowthrough. The supernatant of the filter plate was analysed with a UV-Vis spectrophotometer, in the case of a single-component system, and with the LabChip, in the case of binary mixtures.

The standard deviation for the so obtained values of absorbance, in UV-Vis analysis and peak area, in the case of the LabChip analysis, was calculated together with the %RSD and referred to the average value. In both cases, the concentration values were obtained with calibration curves from absorbance and peak area values; standard deviation and %RSD were also calculated for the concentration data that was obtained. These last values of %RSD, and standard deviation, also include the error committed by using calibration curve to obtain concentration values. The error bars applied to adsorbed protein concentration values (y-axis) are represented by the variation coefficient obtained from the tests described before, considering the condition with adsorption in the wells.

2.4 Column and plant parameter determination

Mass transfer parameters and bed properties are necessary for interpretation of frontal analysis and bind-elute tests. Acetone and blue dextran are used to determine experimentally the bed and plant parameters (Hahn et al., 2016). Acetone and Blue Dextran used for column parameter determination were purchased from Carlo Erba Reagents and Sigma Aldrich, respectively. Acetone is a small molecule (58.08 g/mol) that can penetrate pores of the beads in the column. Pulse injections with a $10\%_{v/v}$ acetone solution are performed; loop injection is carried out and the volume of the loop is equal to 1% of the column volume (thus 10 µL for HiTrap and 50 µL for HiScreen columns). Pulse injections of acetone solution are performed without the column to determine the dead volume V_d from the injection to the UV sensor.

When the injection is performed in the column, the total void volume V_f of the bed is determined from the peak retention volume V_{RetAc} (Eq. 2.5), i.e., the volume of fed liquid from the beginning of the operation to the detection of the maximum of the peak at the outlet sensor. The operation is repeated with blue dextran, which is a big molecule (2000 kDa) that cannot penetrate into the pores. Thus, pulse injections performed with blue dextran evaluate the interstitial volume of the column V_{int} from the dextran peak retention volume V_{RetDex} (Eq. 2.6). Blue dextran pulse injections are performed with a 1% of the column volume loop with a 2 mg/mL blue dextran solution.

The total porosity ε_t, the column porosity ε_c and the particle porosity ε_p, are then calculated (Eq. 2.7 to 2.9). In the following equations, V_c is the geometric volume of the column.

$$V_f = V_{RetAc} - V_d \tag{2.5}$$

$$V_{int} = V_{RetDex} - V_d \tag{2.6}$$

$$\varepsilon_t = \frac{V_f}{V_c} = 1 - (1 - \varepsilon_p)(1 - \varepsilon_c) \tag{2.7}$$

$$\varepsilon_c = \frac{V_{int}}{V_c} \tag{2.8}$$

$$\varepsilon_p = \frac{V_f - V_{int}}{V_c - V_{int}} \tag{2.9}$$

The peak resulting from the dextran pulse injection also allows us to calculate the axial dispersion in the column. Axial dispersion D_{ax} can be calculated from σ_{Dex}, the standard deviation of dextran peak obtained from pulse injection (Eq. 2.10), considering the column length L, the interstitial velocity and volume v_{int} and V_{int}. The standard deviation of the dextran peak is directly calculated from the chromatograph software Unicorn (GE Healthcare).

$$D_{ax} = \frac{1}{2} L v_{int} \left(\frac{\sigma_{Dex}}{V_{int}}\right)^2 \tag{2.10}$$

Each pulse injection is performed three times and the resulting average peak retention and standard deviation are considered.

2.5 Frontal analysis

Breakthrough tests were performed with BSA, Cellulase, CRM_{197} and Lysozyme on a ÄKTA avant 25 (GE Healthcare), using HiTrap Butyl HP columns (0.7 cm diameter and 2.5 cm bed height, that results in a column volume of 1 mL). Since breakthrough experiments are sample wasteful a 1 mL column was chosen for this study. The HiTrap column format with a 1 mL of column volume represents the smallest format of this type of column that can be connected to the chromatographic system used. Protein concentrations investigated were 3.5, 5, 7.5 and 10 mg/mL, prepared as described in section 2.1. Salt concentrations investigated were 0, 1 and 1.5 mol/L of ammonium sulfate for BSA, Lysozyme and Cellulase, and 0.25, 0.75 and 1.2 mol/L of ammonium sulfate for CRM_{197}. The different conditions chosen for CRM_{197}, were due to different behaviour of the protein especially at 0 and 1.5 mol/L of ammonium sulfate (discussion in the following chapter).

Breakthrough runs consist of different phases whose length is measured in CVs (column volume):

- 2 CVs of wash with water to remove storage solution from the column;
- 5 CVs of equilibration with the buffer solution at the desired salt concentration;
- 30 to 60 CVs of loading, the length of this phase depend on the adsorption in the condition investigated; if the adsorption is higher the saturation plateau will occur later and the sample amount necessary to reach the plateau will be higher;
- 5 CVs of rinse with the same buffer of equilibration, this step is made to remove the protein from the interstitial liquid in the column (in particular in the case of lysozyme this step needs to be longer (10 CVs) to remove all the protein from the liquid phase);
- 8 CVs of stripping with the buffer solution without salt to allow the protein to detach from the solid surface;
- 5 CVs of stripping with water to remove any protein residue left in the column.

Each breakthrough test was performed three times, and, before the run, bypass breakthrough test was performed without the column, thus without any adsorption. The bypass test was made to have an accurate reference value of the plateau concentration that needs to be reached. Indeed, the breakthrough plateau is reached when the solid is saturated and no more protein can be adsorbed, thus the inlet concentration and the outlet concentration are the same.

The obtained breakthrough curve data were used to calculate the concentration of protein absorbed to the solid using the mass balance with numeric integration of the curve (Eq. 2.11) (Schmidt-Traub, 2005):

$$V_d(c_f - c_{in}) + V_c\left\{\left(\varepsilon_t(c_f - c_{in})\right) + (1-\varepsilon_t)\left[q^*(c_f) - q^*(c_{in})\right]\right\} = \dot{V}\int_0^{t_{des}}\left[c_f - c(t)\right]dt \quad (2.11)$$

In the mass balance above, c_{in} is the concentration of the protein in the liquid in equilibrium with $q(c_{in})$ adsorbed to the solid surface, in the initial state. In the initial state there is no protein in the column and these variables are equal to zero. The same variables in the final condition, when the plateau is reached, are c_f and $q(c_f)$, c_f is the breakthrough concentration. V_d is the dead volume of the column and equipment, determined from the injection of the sample to the UV sensor, ε_t is the total porosity of the column. V_d and ε_t were determined experimentally (see section 2.5).

When breakthrough isotherms are compared with high-throughput data, the difference in solid volume between the gravity settled resin (in high-throughput filter plates) and the packed resin (in pre-packed column) must be considered. The compression factor *CF* is the ration between the height of the bed settled for gravity and the packed bed height. Using this factor is necessary when a comparison between the static and dynamic methodologies is done. V_c must be multiplied for the Compression Factor *CF* equal to 1.15 (manufacturer data).

The Lambert-Beer law was used to convert the absorbance data at 280 nm in concentration (Eq. 2.12):

$$A = \varepsilon_\lambda \cdot l \cdot M \quad (2.12)$$

In the equation above A is the absorbance obtained with the UV sensor of the chromatographic system. The UV cell path length l depends on the equipment, in this case it is 0.2 cm. M is the molar concentration of the solution passing through the UV-cell and ε_λ is the molar extinction coefficient for a determined wavelength that depends on the protein.

Since, per each condition, three tests were performed, the average values of the adsorbed and free concentration obtained from the three tests were considered.

After each run, CIP (Clean In Place) and sanitization of the column are performed.

The breakthrough curves performed with Cellulase gave abnormal results. To clarify this behaviour, that is explained in section 4.1.1, an SDS-PAGE analysis was performed on some eluted fractions of the Cellulase breakthrough runs.

The SDS-PAGE analysis is performed using NativePAGE Running Buffer (20X) and 10-wells Native PAGE Gels. The sample is prepared with NativePAGE Sample Buffer (4X) and reduced with NuPAGE Sample Reducing Agent (10X) for five minutes at 100°C. The samples and the iBright Prestained Protein Ladder are loaded in the wells of the gel and then the gel is run in a Mini Gel Tank with the

running buffer. All the reagents and equipment mentioned were purchased from Thermo Scientific (Waltham, MA, USA). The gel is then acquired with a GS-900 Calibrated Densitometer (Bio-Rad, Hercules, CA, USA).

2.6 Bind-elute experiments

Bind-elute experiments are performed with BSA, Lysozyme and CRM_{197} in HiScreen Butyl HP columns on a ÄKTA avant 25 (GE Healthcare). For protein loading a 5 mL loop is used, the protein concentration varies from 2 to 5 mg of protein per mL of resin. In the case of Lysozyme, since this protein adsorbs more than others and at high concentrations saturation of the sensor is incurred, lower protein concentrations are used, from less that 1 up to 2 mg of protein per mL of resin. For these tests, salt concentration varies from 1 to 1.4 mol/L for equilibration and loading phases. The salt concentration range investigated is narrow. Indeed, when concentration lower than 1 mol/L is used, the adsorption is low and the protein moves to the outlet before the elution phase. When a salt concentration higher than 1.4 mol/L is used the chaotropic effect could occur and the solubility and hydrophobicity of the protein may be affected. In Table 2.1 below the experimental design is resumed.

Per each protein isocratic and gradient elution are performed. Different elution conditions are performed considering high-salt buffers and no salt-buffers. Some bind-elute tests are used for the validation of the model.

The method used for bind-elute tests is the following:

- 3 CVs of wash with water to remove storage solution from the column;
- 5 CVs of equilibration in 0.05 mol/L potassium monobasic phosphate at the required salt concentration;
- Loading of 5 mL of the protein in the same buffer solution of equilibration;
- 2 CVs of rinse with same buffer of equilibration to remove the protein remained in the interstitial liquid;
- 10 or 5 CVs of elution in step or gradient mode, depending on the run, to elute the protein from the column;
- 3 CVs of stripping with water to remove eventual residues of protein on the resin in the column.
    - In this case also, CIP and sanitization of the column are performed after each run.
    - The amount of loaded protein is known, but a double check is made on the elution peaks. The chromatograph software Unicorn directly calculates the amount of protein based on the selected peak area, given the extinction coefficient of the protein at the working wavelength.
    - Experimental bind-elute runs are compared to the simulated runs (see section 4.2)

Table 2.1: Experimental design for bind-elute tests.

Protein	Elution Method	Elution CVs	Salt concentration, mol/L	Load protein concentration, mg/mL
BSA	Isocratic	10	1.2	4.66
	Gradient	10	1	4.60
	Gradient	5	1.4	4.17
CRM_{197}	Isocratic	10	1.2	3.21
	Gradient	10	1	2.33
	Gradient	5	1.4	2.86
Lysozyme	I	10	1.2	0.86
	Gradient	10	1	2.98
	Gradient	5	1.4	2.18

2.7 Modelling and optimization of the polishing step of in-house protein

The industrial in-development purification process chosen as case study of this work is the polishing step of a recombinant protein purification. In this step the hydrophobic interaction chromatography is exploited to purify the final product, the monomer (82 kDa) from the impurities, that were not removed from the previous steps. The load of the polishing step is, indeed, the elution product of the intermediate step and can be considered constituted of three groups of pseudo-components: the target product or monomer, the high molecular weight compounds (assumed around 160 kDa) and the low molecular weight compounds (assumed around 50 kDa). In this process, the monomer is present in four charge variants; the elution profile is constituted by two main peaks corresponding to the monomer elution. The four charge variants of the monomer elute by couple, two charge variants elute together in the first peak, and the other two variants elute in the second peak. The low molecular weight compounds are constituted by truncations of the monomer, while the high molecular weight compounds are aggregates of the monomer.

To produce the in-house protein, the upstream cell paste was purified in larger scale. Homogenization by lysis was performed, subsequently the harvest broth was centrifuged and the supernatant was collected and filtered. The capture and intermediate purification step were performed with cation exchange and mixed

mode chromatography respectively. The intermediate elution product is the starting material for the polishing step, and it was produced from a unique cell paste, in such a way that all the runs were performed with the same starting material. Thus, the system was not affected from the variability that comes from the upstream process.

The equilibration buffer consists in 0.2 mol/L of potassium monobasic phosphate and 1 mol/L of ammonium sulfate, while elution buffer is 0.05 mol/L of potassium monobasic phosphate; 6 mol/L of potassium hydroxide was used for buffer pH titration. Experiments are performed in a HiScreen Butyl HP column (77x100 mm) using a ÄKTA avant 25 supported by Unicorn (GE Healthcare). The flowrate was changed according to the opening of the pump and handled for the equilibration buffer and the elution buffer from two different pumps of the ÄKTA that will be called, respectively, A pump and B pump.

The standard purification process for the polishing step involved, at first, the usual equilibration of the column, the loading of the protein and the rinse of the column with the same buffer of the equilibration. Subsequently, a first step was made with the 10% of the B pump for 3 CV, then the elution step was performed at 23% of the B pump for 10 CV. After the isocratic step, the gradient elution was carried out, for others 10 CV, from 23% to 50% of the B pump. At the end, a stripping with water for 5 CV was performed to remove any residues left in the column. In the standard conditions, the buffers were set to the neutral pH and the loading of the column was 2 $mg_{monomer}/mL_{resin}$. The loading volume was then calculated based on the protein loading, depending on the column used. The starting material, that comes from the previous elution step, needs to be adjusted in terms of conductivity to reach in the loading solution the same conductivity of the equilibration buffer, to allow the protein to attach to the column. The conductivity adjustment is made with a 0.4 mol/L potassium monobasic phosphate and 2 mol/L ammonium sulfate buffer. A pH adjustment is also performed with 6 mol/L of potassium hydroxide or 1 mol/L of phosphoric acid to reach pH 7.

To estimate model parameters using DSPX, several experiments were performed on this process. The model chosen was then validated with a run in different operating conditions from those used for the model parameters estimation. At the end, using the model found, the process has been optimized maximizing the yield and purity of the target product among the other species.

Several experiments were performed manipulating those parameters that affect the adsorption behaviour: elution method, monomer loading and pH. Furthermore, the elution method has been simplified to be modelled more easily: the wash step is not performed and the elution step has been reduced to a first step of 20 CV at 23% of B pump followed by a 3 CV step at 100% of B pump. Using this method, runs were performed changing the pH and the monomer loading from the standard conditions, the operating conditions of the runs performed is resumed in Table 2.2.

Table 2.2: Operating conditions of runs performed for the modelling of the industrial process

Experiment	Elution method	Monomer loading, ($mg_{monomer}/mL_{resin}$)	pH
Ref	Simplified standard elution method	2	7
Grad	20 CV Gradient from 0% to 50% of B pump	2	7
pH6	Simplified standard elution method	2	6
pH8	Simplified standard elution method	2	8
Load1	Simplified standard elution method	1	7
Load3	Simplified standard elution method	3	7

The eluted product was collected in 4.7 mL fractions and stored; each fraction is equal to a column volume. As written before, the usual elution profile consists in two peaks in which the monomer is mainly eluting and a third peak of the stripping phase in which the high molecular weight compounds are manly eluting. Analytical pools of each peak were collected for every run. Size Exclusion Chromatography (SEC) analysis is performed on each analytical pool and loading to evaluate concentration and purity of the monomer with respect to high and low molecular weight components. The SEC analysis were performed from the GSK Vaccines laboratories using a ACQUITY UPLC Protein BEH SEC Column (200 Å, 1.7 µm, 4.6 mm X 300 mm) on a ACQUITY UPLC system, both were purchased from Waters (Milford, MA, USA).

To use the estimative approach, equilibrium adsorption isotherms were evaluated with high-throughput methodology for the monomer. Since the four charge variants elute in couples, two variants in the first peak and the other two in the second peak, they are supposed to have, by couple the same adsorption behaviour. For this reason, when performing the high-throughput experiments, two monomer species were considered, Monomer 1 that is the mixture of the two charge variants eluting in the first peak, and Monomer 2, the mixture of the two charge variants eluting in the second peak. The eluted material produced in the experiments described before was collected per each peak. Amicon Ultra Centrifugal Filters (Merck, Darmstad, Germany) were used to concentrate and exchange the buffer of the eluted material to have the protein at the desired protein concentrations in a 0.2 mol/L potassium monobasic phosphate at three different ammonium sulfate concentrations, 0.75, 1 and 1.2 mol/L. The buffer and salt concentrations respect the concentrations used in the process that promotes the binding of the protein to the solid, while the ammonium sulfate concentrations were chosen around the 1 mol/L value used in the process. The material available was limited and the stability

of the target protein at high concentration and with high salt concentration is still not well known. For these reasons, the protein concentrations investigated varies between 0.4 to 1.5 mg/mL for Monomer 1 and from 0.5 to 2.0 mg/mL for Monomer 2. The monomer concentrations used allowed to perform sufficient high-throughput tests to determine the equilibrium adsorption isotherms, considering that the amount of protein was limited. By the way, it must be also considered that the ranges of protein concentrations investigated are coherent with the operating load condition which does not exceed 1 $mg_{monomer}$/mL.

The high-throughput methodology used for the determination of equilibrium adsorption isotherms followed the procedure described in section 2.1. The sample, in this case cannot be analysed with the UV-Vis analysis since the protein investigated is not pure as the commercial proteins. The purity reached for the product is around 93% with the presence of the low and high molecular weight compounds. Analysing the samples with UV-Vis would give misleading results due to the different adsorption behaviour of the different species. The microfluidic capillary electrophoresis was also tested in this case to analyse load and flowthrough samples. Unfortunately, this kind of analysis is not suitable for the product of interest (see section 4.1).

To avoid the use of the microfluidic capillary electrophoresis, the SEC analysis was performed on the samples. Despite this kind of analysis is not suitable for a high-throughput methodology, since it requires long time to run the samples and integrate the peaks, it was the only option that allow to determine the monomers concentrations and purities. With the load and flowthrough concentration values obtained with the SEC, the mass balance was then used to calculate the adsorbed concentrations as described before (section 2.1)

Chapter 3

Numerical simulations

Numerical simulations were performed to investigate the behaviour of protein in hydrophobic interaction chromatography. To do so two different approaches were used. The estimative approach exploits a commercial software DSPX (GoSilico, Karlsruhe, Germany). The second approach uses an in-house Fortran code. These two approaches were used to simulate bind-elute runs and the simulated and experimental results were compared to evaluate the robustness and applicability of each approach. The purification process of the in-house protein was also modelled, validated and optimized with DSPX.

3.1 DSPX

The commercial software DSPX from GoSilico (Karlsruhe, Germany) was used to perform parameter estimation. DSPX is a chromatography modelling software in which it is possible to estimate model parameters from experimental chromatograms and data. The estimation in DSPX is performed as an optimization of model parameters minimizing the difference between simulated and experimental chromatograms.

DSPX utilizes a finite element scheme for the discretization of the spatial variables of the system. For the time discretization, different methods are implemented in DSPX: Explicit Euler, Implicit Euler, Crank-Nicolson, Fractional step, BDF and IDAS. For both commercial and in-house protein, the IDAS method was used. For the spatial and the time coordinates the step is variable, 10 seconds was set as time step, while for the spatial coordinates 30 axial cells and 10 radial cells were used. These values are set by default from the software and can be changed if necessary, but in the systems investigated they were suitable, in terms of complexity of the model and quickness of simulations.

The estimation can be intended as an optimization of model parameters minimizing the difference between simulated and experimental chromatograms. To do so, different algorithms are implemented in DSPX that can be grouped in five different classes based on the systematic way the algorithm determine new parameter set-ups. The five classes (with the respective algorithms) available are: Levenberg-Marquardt (CMinPack, CERES), Trust Region (MKL), Interior Point

(IPOPT), Simulated Annealing (ASA) and Genetic (GALib, OpenBeagle, OpenBeagleMultiObjective). To perform an estimation and develop the model, the heuristic algorithm (ASA) is used at first to find a first approximation of parameters with an algorithm that works faster ranging in the whole space and do not stop in local minimum, then the model is refined with a deterministic algorithm (CERES) that is slower but more precise. Different error norms are also implemented. The error norm selected defines how the algorithm calculates the difference between the experimental and simulated chromatograms. The error norms available are: individual norms, NRMSE (Normal Root Mean Square Error), NMRSE Integral, Normalized Dynamic Time Warping, Normalized L1, Normalized L1 Integral, Log Likehood. For the estimation performed in this study the NRMSE norm was use, in particular this is the best choice when some very small peaks are present in the chromatograms and they can be neglected, like the flowthrough of the in-house purification process.

Online data from the experimental tests are exported from the chromatograph management software (e.g., Unicorn) and imported to DSPX. When data from the experimental run are imported in DSPX, all information regarding the sensor and the method are uploaded. Before starting the parameter estimation, it is necessary to set up the experiment characteristics like type of column and adsorber, porosities, buffers, load concentrations. In DSPX it is possible to choose the models to describe the interstitial liquid mass transfer, the pore liquid mass transfer and the adsorption kinetics. The Transport Dispersive model (Eq. 1.2) is considered to describe the mass transfer in the interstitial liquid, while the General Rate model (Eq. 1.3) is used for the mass transfer in the liquid inside the pores of the beads, when modelling the commercial proteins. As written in section 1.3, two models are implemented in the software to describe adsorption kinetics (Eq. 1.16-1.18) in hydrophobic interaction chromatography. Both were tested to perform the estimation on bind-elute tests. In the case of the in-house protein purification process, since it is more complex, the Lumped Rate model was used to describe mass transfer phenomena in the liquid inside the pores of the beads (Eq. 1.20).

To simulate the bind-elute tests the experimental chromatogram were imported in DSPX and the operating conditions together with the characteristics of the system were set in the software. Everything was fixed to respect experimental conditions (described in section 2.7), salt and protein concentration, porosities and dead volume, method.

Additionally, a simple run was simulated with DSPX in order to be compared with the respective with the Fortran code, to prove the consistency of the two different codes. This run considered an injection of 5 mL of sample, constituted by BSA at 4.6 mg/mL. The system was considered to respect the Langmuir adsorption law with the adsorption equilibrium assumption. No dependency from the salt concentration was considered. The Transport Dispersive model was used to

describe the mass transfer in the interstitial liquid, while the General Rate model describes the mass transfer in liquid inside the pore of the beads.

3.2 In-house Fortran code

An in-house Fortran code was used to simulate bind-elute tests. The code allows to solve the partial differential equations system that consists in the mass balance in the interstitial liquid and in the pore of the bead, and the adsorption equilibrium at the solid surface. The code discretizes the equations in the two spatial coordinates (z and r) by collocation of the orthogonal polynomials of Jacobi for z and Lobatto for r on the zeros. For the time (t), the integration exploits the *lsode* routine of *odepack*. The system is moderately stiff; the resolution is performed with a Backword Differentiation Formula (BDF) method of changeable order and step, with error control.

The models that describe the mass transfer were the same used in DSPX (Eq. 1.2 and 1.3). With this code, bind-elute tests were predicted testing different isotherm laws that consider the variation of salt concentration (Eq. 1.12-1.14). Loading concentration and volume, dimension and characteristics of the column, mass transfer parameters and elution method were given as input information to the code. The code solves the PDEs system giving profiles of the outlet concentration of the component, the concentration in the interstitial liquid and the concentration in the liquid inside the pore of the beads.

In this case also, a simple run was simulated in the same conditions as described in previous section, to be compared with the simulation performed with DSPX.

3.3 Matlab curve fitting

Experimental isotherm data are used to find isotherm parameters. With the adsorption laws considered in this work, the dependency from salt concentration is considered, thus the whole bundle of experimental isotherms is fitted to find parameter values. The Levenberg-Marquardt method is used to fit the bundle of experimental isotherms and find isotherm law parameter values with Matlab R2020b (MathWorks, USA).

Chapter 4

Results and discussion

In this chapter the results for the experimental tests and simulated runs are shown. Results are grouped for the kind of test performed, but also comparison and connection are made through the chapter.

4.1 Equilibrium adsorption isotherms of proteins

4.1.1 High-throughput and frontal analysis single-component isotherms determination

The four proteins aforementioned were investigated as explained in section 2.1 to determine equilibrium adsorption isotherms. In Figure 4.1 equilibrium adsorption isotherm are reported. It can be noticed that different proteins have different adsorption behaviour (Chen and Cramer, 2007). In any of the cases here investigated, the typical isotherm plateau was not reached; the ratio between the protein loaded and the solid phase is not enough to reach the saturation of the resin. In particular, it can be noticed that Lysozyme isotherms remained in the linear part of the curve. This behaviour can be due, partially, to the small dimensions of the protein (14 kDa): the adsorption was less affected by the steric hindrance between the molecules.

It can be noticed also that the adsorption isotherm in the case of 1.5 mol/L of ammonium sulfate (green dots in Figure 4.1) were sometimes lower than the isotherms at 1.2 and 1.4 mol/L. This could be due to the chaotropic effect of the high salt concentration in the buffer, even though the salt used is anti-chaotropic. Indeed, it is known that a very high salt concentration can have a chaotropic effect, weakening the hydrophobic effect of proteins (Lienqueo et al., 2007).

Experimental adsorption isotherms were first compared with isotherms obtained with frontal analysis to compare the methods and evaluate the robustness of the experimental data. Later, the experimental adsorption data were fitted with different models to find adsorption parameters.

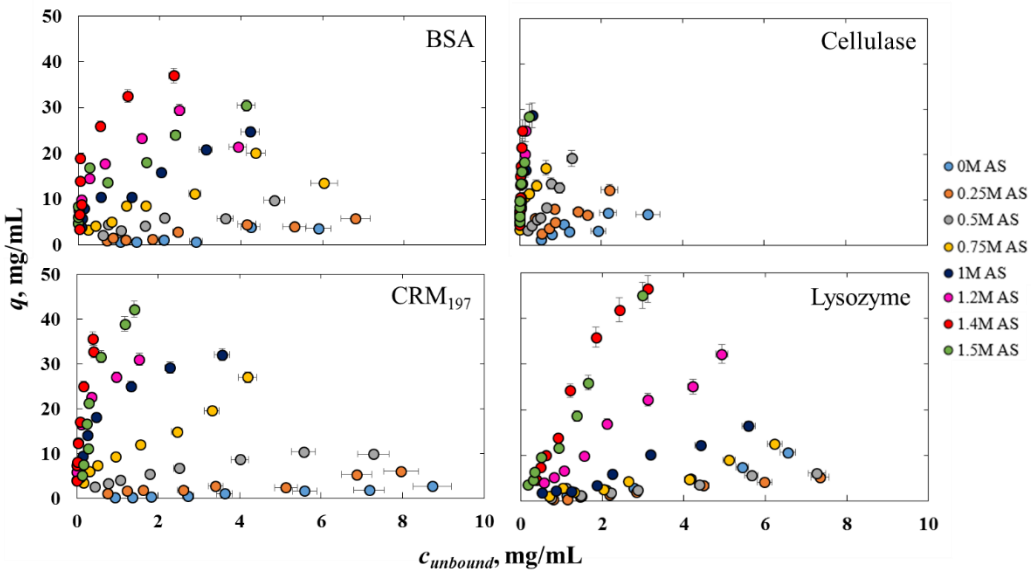


Figure 4.1: Equilibrium adsorption isotherms of BSA, CRM_{197}, Cellulase and Lysozyme for different concentration of ammonium sulfate (AS); determined by HT experiments.

Only four concentrations of protein at three salt concentrations of the buffer were investigated to build adsorption isotherms with frontal analysis. Indeed, frontal analysis is a very accurate method, but it needs a large amount of sample. For this reason, the HiTrap column format was used: to exploit the smallest format possible with the chromatographic system in use and use a smaller amount of sample and chemicals.

Breakthrough experiments were performed as described in section 2.5 and, from the breakthrough curve, the amount of protein adsorbed to the solid $q(c_f)$ was calculated with Eq. 2.11. As written before, in order to compare data between frontal analysis and high-throughput method, the relative value of the compression factor *CF* must be considered: this factor for the column used is equal to 1.15 (data provided from manufacturer), while 1 is assumed for the batch. Values of dead volume and total porosity for the columns were evaluated experimentally as described in section 2.4. The dead volume resulted equal to 0.47 mL while the total and column porosity resulted equal to 0.96 and 0.37, respectively. The comparison between high-throughput and frontal analysis experimental data is reported in Figure 4.2, in which the adsorbed protein concentration data obtained from frontal analysis were corrected with the packing factor and porosity to be consistent with high-throughput data. The experimental data obtained from the frontal analysis resulted consistent with the high-throughput experimental data. We can notice that the unbound protein concentrations investigated were higher and the isotherms obtained from the breakthrough curves follow the trend started by the high-throughput results at lower concentration.

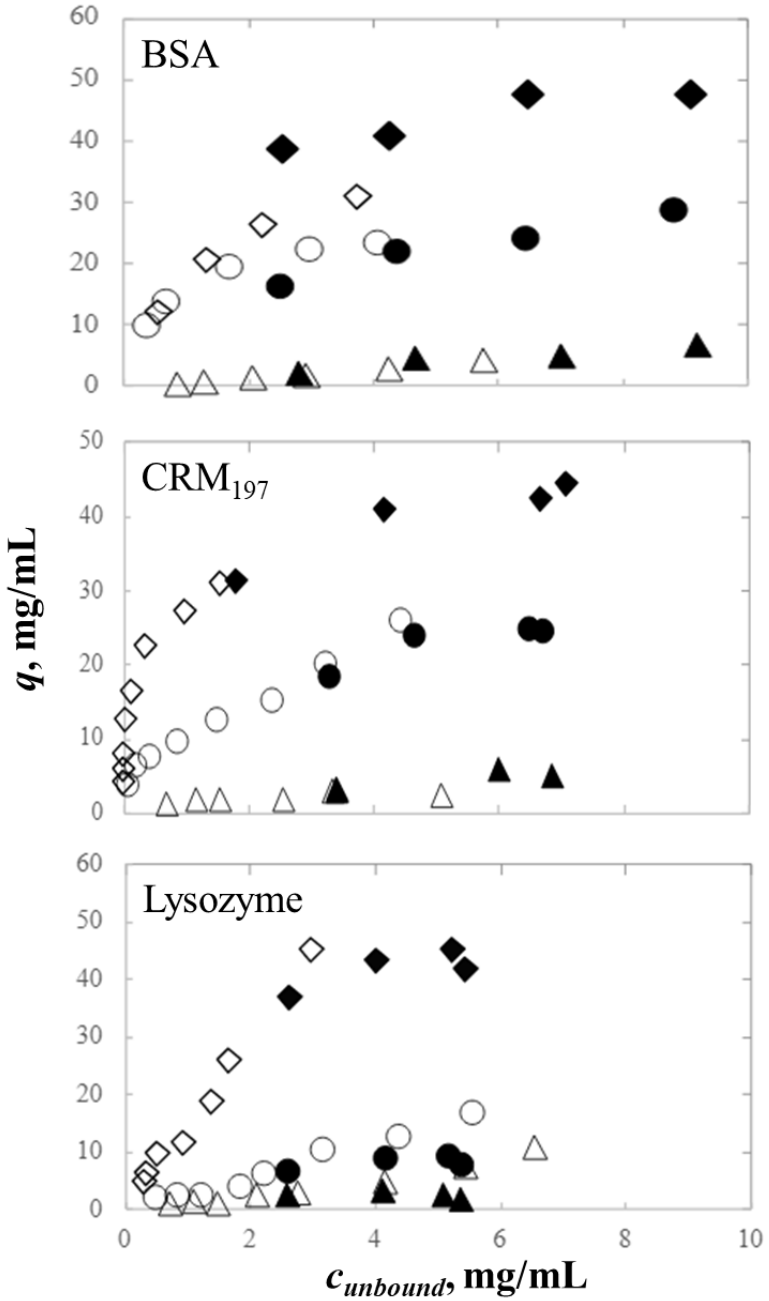


Figure 4.2: Comparison of high-throughput (empty symbols) and breakthrough (filled symbols) experimental isotherms for BSA, CRM_{197} and Lysozyme. Breakthrough data were corrected to be consistent with high-throughput data (dividing the estimated *q* value by column *CF*). In the case of BSA and Lysozyme triangles, circles and diamonds represent, respectively 0, 1 and 1.5 mol/L of ammonium sulfate in the buffer. In the case of CRM_{197}, the salt concentrations investigated and compared are 0.25, 0.75 and 1.2 mol/L, represented by, respectively, triangles, circles and diamonds.

The experimental data of the adsorption isotherms obtained with the two different methods, the static and the dynamic method, resulted consistent. The two methods are well known in literature. While the static method is favourable because it requires a very small amount of sample and chemicals, the dynamic method needs a high amount of sample to load the column until the saturation. On the other hand, the high-throughput methodologies require a high experimental effort, and this

introduces an important variability on the experimental data that must be taken into account (and will be evaluated in the next section). The frontal analysis method, instead, gives more accurate results and the execution is simple and does not require a long series of manipulations on the samples that may affect the accuracy of the results.

Breakthrough tests were also performed with Cellulase, but it resulted in two different plateaus (Figure 4.3). It seems that the Cellulase is constituted by two kinds of proteins with different hydrophobicity. This was not noticed in the static method because the UV analysis is not able to distinguish the different species. To further investigate this point, an SDS-Page analysis has been performed on seven fractions of the breakthrough: four fractions belong to the first lower plateau, two fractions correspond to the step between the two plateaus, and the last fraction is taken at the end of the higher plateau.

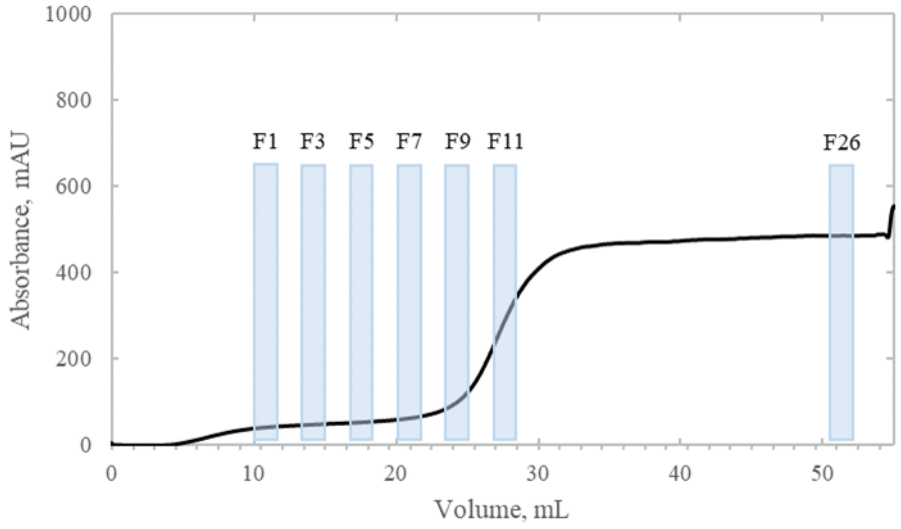


Figure 4.3: Breakthrough curve performed with 2.5 mg/mL of Cellulase in 1 mol/L ammonium sulfate buffer. It is possible to notice two different plateaus in the curve, a lower one that occurs first and a higher one that occurs later. Fraction analyzed with the SDS-Page are highlighted in light blue.

Two SDS-Page are performed (Figure 4.4); per each gel in the first lane the marker was loaded while the second lane refers to the load of the breakthrough run, to have a comparison with the starting material loaded in the column. Samples coming from each fraction were loaded in the gel both as it is and diluted, since the concentration was not known a priori. The correspondence between fraction and lanes is reported in Table 4.1.

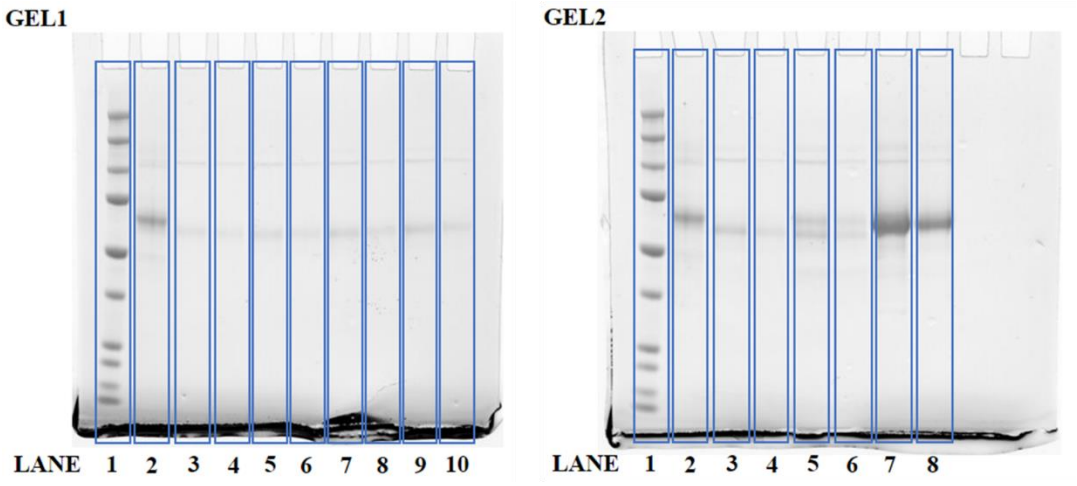


Figure 4.4: SDS-Page of the fractions collected for the Cellulase breakthrough run.

Table 4.1: Scheme of the lane of the SDS-Page in Figure 4.4

Gel 1		Gel 2	
Lane	**Fraction**	**Lane**	**Fraction**
1	Marker	**1**	Marker
2	Load	**2**	Load
3	F1	**3**	F9
4	F1 diluted	**4**	F9 diluted
5	F3	**5**	F11
6	F3 diluted	**6**	F11 diluted
7	F5	**7**	F26
8	F5 diluted	**8**	F26 diluted
9	F7	**9**	-
10	F7 diluted	**10**	-

Two species can be noticed: especially in lane 5 of Gel 2 it is possible to see two different bands. This lane is representative of the fraction 11, collected in the transition between the lower and the higher plateau. These two species may be the responsible of the two different plateaus, they have slightly different molecular weight and different hydrophobicity, that could be the reason of the double plateau. Since this behavior was not clear, we decided to not investigate more this protein and focus on the others.

The abnormal behavior of the Cellulase highlights the limits of the high-throughput methodologies. The Cellulase used for this study seems to be constituted by two species with a slightly different molecular weight and hydrophobicity. The presence of these two species can be noticed when a run performed in column is made because the two species have different behavior during the breakthrough test. This abnormal behavior cannot be noticed with the static method because the UV-Vis analysis performed on the load and supernatant does not allow to discriminate between different proteins. Thus, when performing a high-throughput tests, the sample investigated must be pure. On the other hand, the frontal analysis allows to notice this kind of abnormal behavior but wasting a large amount of sample and time. A similar but even more complex problem will be faced investigating the behavior of the in-house protein as will be discussed in section 4.1.4.

The high-throughput experimental adsorption isotherms were chosen for the determination of the adsorption isotherms in the following sections. Data obtained with the static method were more complete in terms of protein and salt concentrations investigated.

4.1.2 Multi-component adsorption isotherms and assessment of electrophoretic method[1]

Binary mixtures adsorption behavior was investigated using an electrophoretic analytical method. In order to prove the consistency of the results obtained, the analytical method was studied to asses its precision in determining multi-component adsorption isotherms. To do so, single-component adsorption isotherms were evaluated as a first step with both the analytical methods and compared.

Single-component isotherms obtained with UV-Vis analysis resulted consistent with those obtained with LabChip analysis (Figure 4.5). It can be seen how BSA, CRM_{197} and Lysozyme adsorption isotherms obtained with the two different methods were basically the same, considering also that preparation of the samples in the case of LabChip is more complex (see section 2.2) than UV-Vis analysis.

Mixtures were prepared according to the procedure in section 2.2 and isotherms were obtained applying mass balance (Eq. 2.1) to the single protein concentrations.

[1] The results presented in this section were published in Lietta et al. (2021).

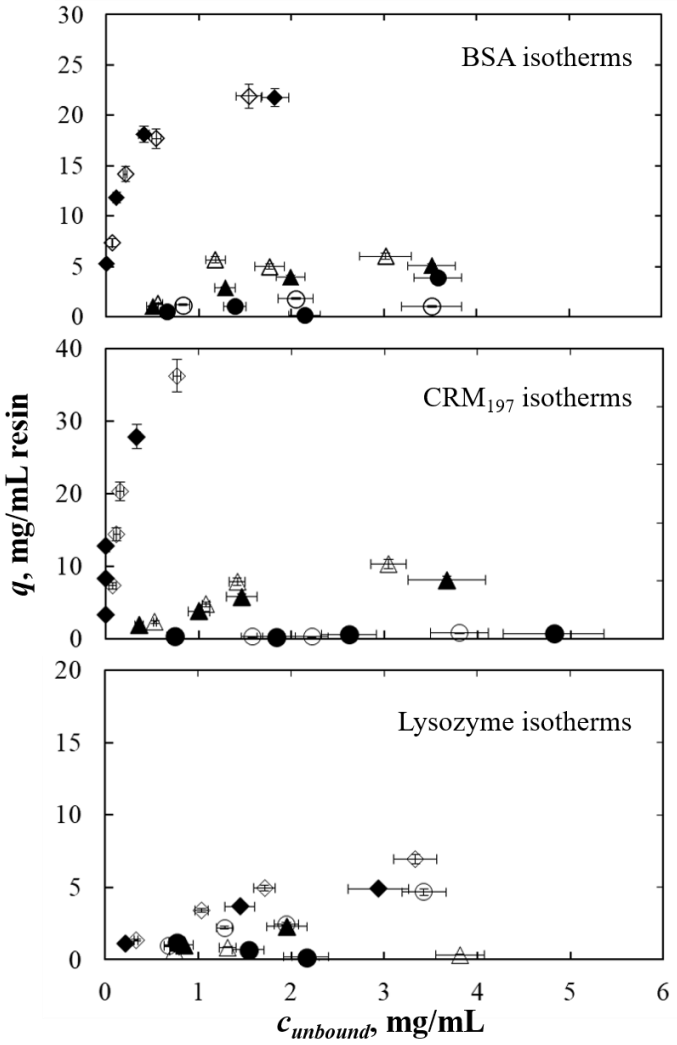


Figure 4.5: Comparison of adsorption isotherms obtained with UV-Vis spectrophotometer and LabChip for the BSA, CRM_{197} and Lysozyme when investigated as single-component at different salt concentrations. For each protein filled symbols refer to LabChip analysis, while empty symbols refer to UV-Vis spectrophotometer analysis. Circles, triangles and diamonds refer to isotherms obtained at 0, 0.5 and 1 mol/L of ammonium sulfate respectively.

In Figures 4.6 a), the comparison of single and multi-component isotherm is shown for BSA and CRM_{197}. With this couple of proteins, we can notice that in the case of 0 and 0.5 mol/L of ammonium sulfate in the buffer, the mutual adsorption behavior did not change when the protein is in mixture. When the concentration of the salt was 1 mol/L, BSA adsorption isotherm, when in mixture with CRM_{197}, resulted lower than the single-component isotherm. CRM_{197} could affect BSA adsorption reducing its binding capacity.

In Figures 4.6 b) isotherms for the CRM_{197}-Lysozyme mixture are shown. In 0 and 0.5 mol/L of ammonium sulfate in the buffer, multi-component isotherms followed the single-component isotherms behavior; in these conditions, CRM_{197}

and Lysozyme showed low adsorption. With 1 mol/L of salt in the buffer, the adsorption for both proteins was higher and, also in this case, both followed their single-component behavior.

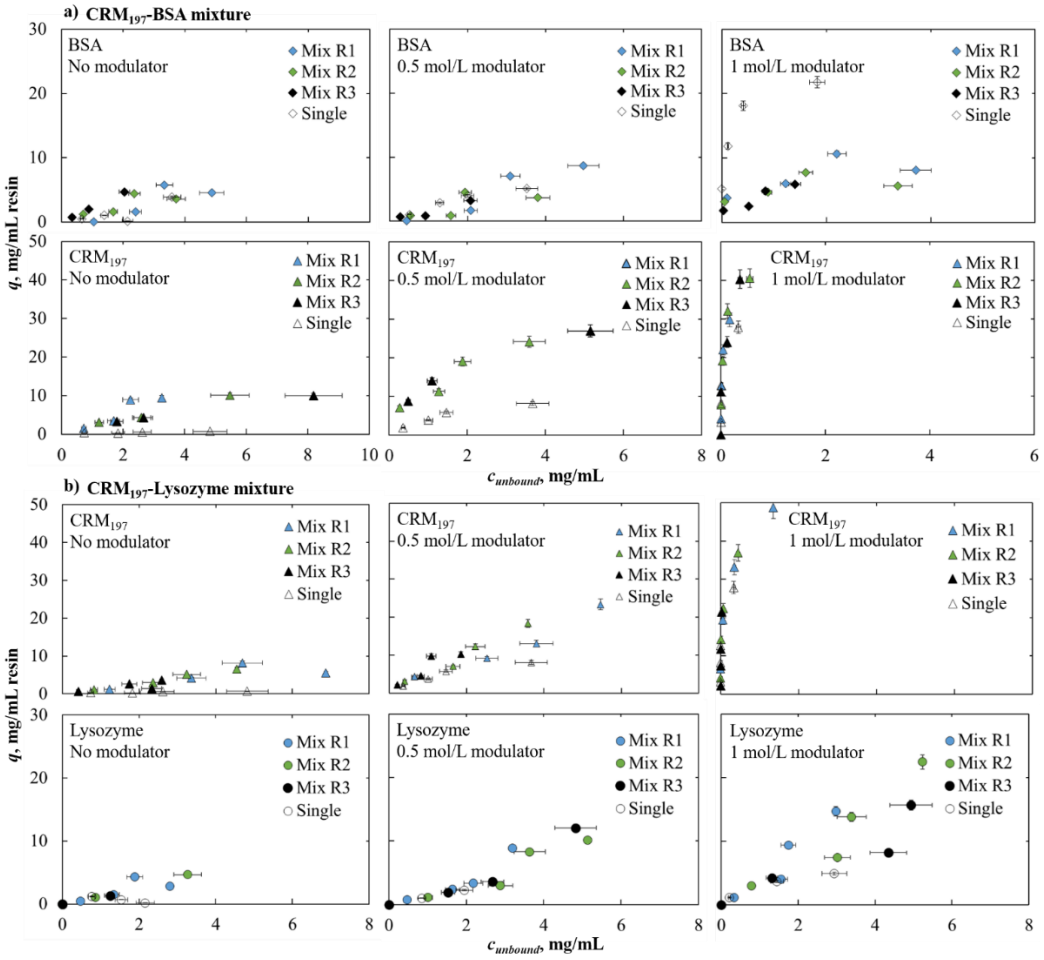


Figure 4.6: Adsorption isotherm comparison for binary mixtures: section a) shows the mixture BSA-CRM_{197}, section b) shows the mixture CRM_{197}-Lysozyme. Each graph presents data for a fixed ammonium sulfate concentration (0, 0.5 and 1 mol/L from left to the right). In each graph the empty symbols named labelled as "Single" represent the adsorption isotherm of the protein in single-component case, "Mix R1" is the protein ratio 70%-30%, "Mix R2" is the protein ratio 50%-%50%, "Mix R3" is the protein ratio 30%-70% considering that, in all the ratios, the first percentage is relative to the protein shown in the upper graphs of the section and the second percentage is relative to the protein in the lower graphs of the section.

The binary mixture of BSA and Lysozyme was also investigated, but in this case the tests gave abnormal results. In particular, the plate with BSA-lysozyme mixture was tested three times and concentration values obtained from the analysis were variable; the values were in some case anomalously high or low and in any case they were inconsistent. Furthermore, the single component adsorption isotherms obtained (on the same plate) with the LabChip were not consistent with the adsorption isotherms obtained with the UV-Vis analysis. This did not happen in the case of the other two mixtures, as previously shown. This problem could be due

to some sort of interference of this mixture with the gel electrophoresis in presence of ammonium sulfate, or to the difference in surface hydrophobicity and intramolecular forces of the two proteins (Moon et al., 2000). Further studies are required to clarify this behavior.

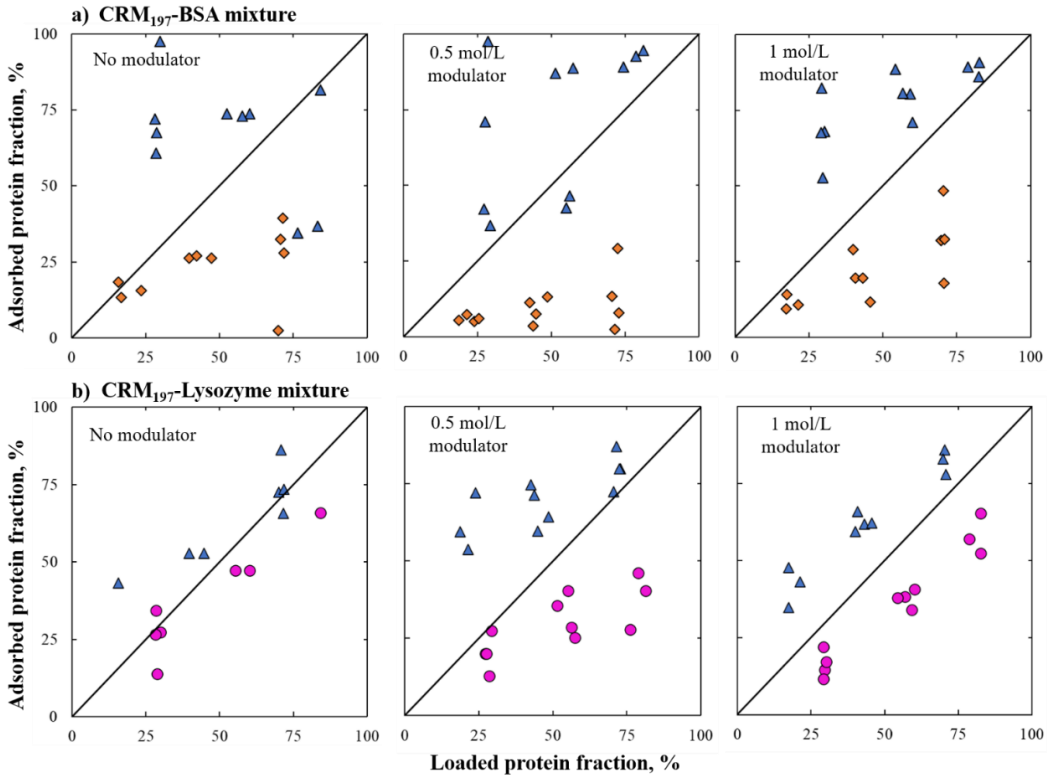


Figure 4.7: Parity plots describe the relationship between adsorbed and loaded protein fractions at different salt concentrations; blue triangles represent CRM_{197} fractions, pink circles represent Lysozyme fractions and orange diamonds represent BSA fractions. Section a) BSA-CRM_{197} and section b) CRM_{197}-Lysozyme mixture.

The different adsorption behavior of proteins could also be evidenced when reporting adsorption data on parity plots (Figure 4.7). The experimental protein ratios were reported in parity plots of loaded protein fractions (x-axis) and adsorbed protein fractions (y-axis). From the parity plots, it can be noted that protein fractions are always close to the bisector in the case of the CRM_{197}-Lysozyme mixture, and for this reason, the parity plots confirmed that protein adsorption behavior did not change when they were in a mixture compared to what was observed in a single-component system. In the case of the BSA-CRM_{197} mixture, it can be noticed that, especially in the case of 1 mol/L of ammonium sulfate, for BSA the adsorbed fraction was generally lower than the loaded fraction. Thus, the opposite happened to the CRM_{197}, for which the adsorbed fraction was higher than the loaded fraction. In these graphs, it can be noticed that the data are stratified and grouped according to the different protein ratios used. Indeed, the graphs show the correspondence

between the data series and the protein ratios used. The higher the loaded protein concentration, the higher the protein fraction adsorbed to the solid. Additionally, Figure 4.7 highlights the impact of salt concentration on the mixture separation efficiency.

To evaluate the performances of the microfluidic capillary electrophoresis as analysis method to determine adsorption isotherms, five samples of standard BSA at 1 mg/mL were analyzed with the LabChip as described in section 2.2. Concentration values of the samples were obtained using the calibration curve and standard deviation and %RSD were calculated on concentration values. As previously explained, the standard sample of BSA 2 mg/mL was diluted 1:1 in order to remain in the linear range of fluorescence analysis. The average value of the concentration of the samples was 1.070 mg/mL, close to the theoretic concentration value, the semidispersion of the five concentration values resulted 0.071 mg/mL. The accuracy, calculated with Equation 4.1 (c_{av} and c_{th} are respectively the average and theoretical concentration), was 7.04%; this value also considers the variability given by the manipulation on each sample that is the main cause of error in the procedure. Moreover, in the work of Field et al. (2017), the accuracy of the PLS method resulted within 5% that is not far from the accuracy that is determined for this application of the capillary electrophoresis. The accuracy study was done in a unique point of the calibration curve in the middle of the working range of concentrations (0-2 mg/mL). We assume that the results so obtained are valid for all the concentrations belonging to the calibration curve.

$$Accuracy = \frac{c_{av} - c_{th}}{c_{th}} \cdot 100 \tag{4.1}$$

To determine how the preparation of the samples affects variability in microfluidic electrophoresis, a larger amount of sample (BSA 1 mg/mL) was prepared as described in section 2.2 and dispensed in ten wells to exclude samples preparation variability. The %RSD in this case was 1.09%, which is lower than the previous case. This difference shows how the manipulation, necessary to prepare samples, affects the measurement error, more than equipment variability.

When verifying the separation efficiency, electropherograms obtained from the analysis of 90%CRM_{197}-10%BSA (Sample 1) and 10%CRM_{197}–90%BSA (Sample 2) mixtures are shown in Figure 4.8. As written above, BSA and CRM_{197} have similar molecular weights, nevertheless, the peak resolution was high, the two peaks turned out to be well separated. In both cases, the protein present in smaller quantities was adequately separated. Values of protein fractions were obtained from LabChip analysis (Table 4.2). The experimental protein fractions were reasonably close to the theoretic values, considering errors that could occur in diluting a single protein solution to reach the desired concentration and mixing them in the desired ratio, in addition to the sample preparation required for analysis. Indeed, as previously demonstrated, manipulation of the sample mainly affects the

measurement. Thus, we can say that no considerable amount of protein was entrained with each other during separation, and the peaks were representative of the individual separated proteins. The difference between theoretical and experimental fraction values is assumed to be due to sample handling.

Table 4.2: Protein fraction obtained with LabChip analysis on BSA-CRM_{197} mixture.

Sample	CRM_{197} fraction, %	BSA fraction, %
Sample 1	92.88	6.88
Sample 2	5.75	94.25

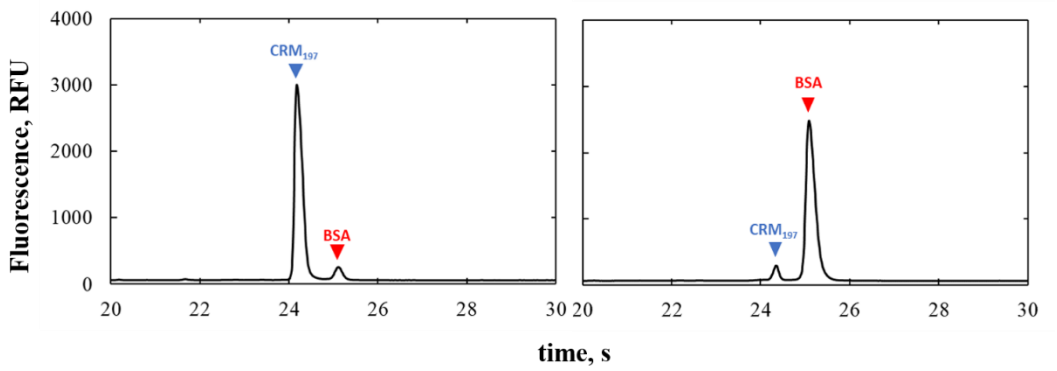


Figure 4.8: Electropherograms of binary mixtures of 90%CRM_{197}-10%BSA (Sample 1), on the left, and 10%CRM_{197}-90%BSA (Sample 2), on the right.

4.1.3 Experimental variability of the high-throughput method

For the three proteins in a single-component system, the reproducibility of the high-throughput method was evaluated to assess the robustness of the results.

The UV-Vis spectrophotometer measurement variability was investigated and the %RSD turned out to be 4.06% with a standard deviation of 0.0018 AU.

The relative error was calculated for all the curve building points. The final calibration line error considered was the average value of all the relative errors of the line building points. The average relative errors and correlation coefficients of the UV-Vis calibration curves of BSA, CRM_{197} and Lysozyme are reported in Table 4.3. The sum of the UV-Vis spectrophotometer measurement error and calibration line error was applied to $c_{unbound}$ values (x-axis) for single-component adsorption isotherms.

Table 4.3. Correlation coefficients and average relative errors of the UV-Vis calibration curves uses for quantification of BSA, CRM_{197} and Lysozyme.

	Calibration line correlation coefficient R^2	**Average relative error, %**
BSA	0.9989	5.15
CRM_{197}	0.9997	2.54
Lysozyme	0.9998	2.80

Table 4.4. Correlation coefficients and average relative errors of the capillary electrophoresis calibration curves uses for quantification of BSA, CRM_{197} and Lysozyme.

	Calibration line correlation coefficient R^2	**Average relative error, %**
BSA	0.9996	1.74
CRM_{197}	0.9957	4.98
Lysozyme	0.9997	4.72

In the case of microfluidic capillary electrophoresis, the measurement variability was calculated on the peak area obtained from the electropherograms of the samples. In this case the %RSD was 6.35%.

The calibration curve error was calculated for LabChip analysis with the same procedure used for UV-Vis analysis. Calibration curve average relative error calculated for BSA, CRM_{197} and Lysozyme are reported in Table 4.4. The sum of %RSD and calibration curve absolute error was applied to the $c_{unbound}$ concentration values (x-axis) in multi-component isotherms.

Results of the experiments performed with filter plates are shown in Tables 4.5 and 4.6. The standard deviation of no/low-adsorption absorbance and concentration values resulted generally slightly lower than the case of 1 mol/L of salt in the buffer. While calculating the standard deviation, outliers were identified with the Tukey criteria (Tukey, 1949) and excluded. The difference between standard deviation and %RSD values of absorbance and concentration in the two different adsorption conditions, shows how the adsorption in plates is affected by random variations that cannot be directly quantifiable. The semi-dispersion, calculated for the UV-Vis set of data, resulted in the range between 1% and 15% approximately. Semi-dispersion calculated values resulted generally slightly higher as the adsorbed concentration

increased. This behavior is consistent with what was found previously: adsorption increase variability in the procedure.

Table 4.5: Average value, standard deviation and %RSD for the proteins when analyzed with UV-Vis spectrophotometer in low adsorption condition (without ammonium sulfate).

		BSA	**CRM_{197}**	**Lysozyme**
Absorbance, AU	Average	0.274	1.850	0.964
	Standard Deviation	0.011	0.050	0.044
Absorbance %RSD		3.83	2.69	4.57
Concentration, mg/mL	Average	6.83	9.09	7.25
	Standard Deviation	0.29	0.25	0.35
Concentration **%RSD**		4.22	2.76	4.81

Table 4.6: Average value, standard deviation and %RSD for the proteins when analyzed with UV-Vis spectrophotometer in adsorption condition (1 mol/L ammonium sulfate).

		BSA	**CRM_{197}**	**Lysozyme**
Absorbance, AU	Average	0.458	0.756	2.48
	Standard Deviation	0.022	0.045	0.11
Absorbance %RSD		3.83	4.74	5.89
Concentration, mg/mL	Average	4.61	3.56	7.19
	Standard Deviation	0.25	0.23	0.31
Concentration %RSD		4.22	5.42	6.29

In the case of protein mixtures, the results of the test in mild adsorption conditions are shown in Table 4.7. In this case, the variables investigated were the peak area of the resulting electropherogram and the related concentration obtained with the calibration curve.

The value of %RSD calculated for the samples was considered and applied to *q* concentration values (*y*-axis) for both UV-Vis and microfluidic capillary electrophoresis analysis. In the case of protein mixtures, semidispersion calculated for experimental data is between 2 and 19% in the mixtures experimentally investigated BSA-CRM_{197} and CRM_{197}-Lysozyme, while the mixture BSA-

Lysozyme gave abnormal results that are discussed previously. For this reason, and since semidispersion values resulted coherent between the different mixtures, the Lysozyme variability was not experimentally investigated, but assumed as the average of BSA and CRM_{197} %RSDs. %RSD, and then precision, has been reported as approximately 1.6-2.1% in a previous work (Osberghaus et al., 2012), but in that case a robotic platform was used. Furthermore, precision of adsorbed protein concentration turned out to be less than 3.6% in another work (Field et al., 2017) where multivariate analysis of spectra has been exploited for analysis. In this last case also, a robotic liquid handler is used. Results here reported were higher, but it must be considered that all the operations described were performed manually by the operator. Moreover, has been demonstrated in Osberghaus (2012) that precision of the procedure and analysis increases with the training of the experimenter on the device. The manipulation factor mainly affects the results.

Table 4.7: Average value, standard deviation and %RSD for BSA and CRM_{197} when analyzed with LabChip in binary mixture and adsorption condition (0.5 mol/L ammonium sulfate).

		BSA	**CRM_{197}**
Peak area, RFU · s	Average	2300	1140
	Standard Deviation	100	60
Peak area %RSD		4.43	5.44
Concentration, mg/mL	Average	1.774	0.961
	Standard Deviation	0.074	0.057
Concentration %RSD		4.18	5.92

4.1.4 In-house protein equilibrium adsorption isotherms

The predictive approach was also tested on the in-house purification process besides the estimative modelling approach. This approach is not suitable for complex systems because of the high experimental effort required (see section 4.1.2). Nevertheless, equilibrium adsorption isotherms were determined with high-throughput methodologies on the target monomer. Equilibrium adsorption isotherms were investigated as written in section 2.7.

In this case, when the microfluidic capillary electrophoresis was performed to analyse samples, if the traditional procedure is performed, the proteins need to be denatured, which makes impossible to distinguish between Monomer 1 and Monomer 2 and aggregates. When the samples are analysed without denaturation, the electropherograms gave abnormal results. The behaviour of the target product using the capillary electrophoresis is still not clear and requires further studies. For

this reason, a more traditional analysis was performed, the Size Exclusion Chromatography, as written in section 2.7.

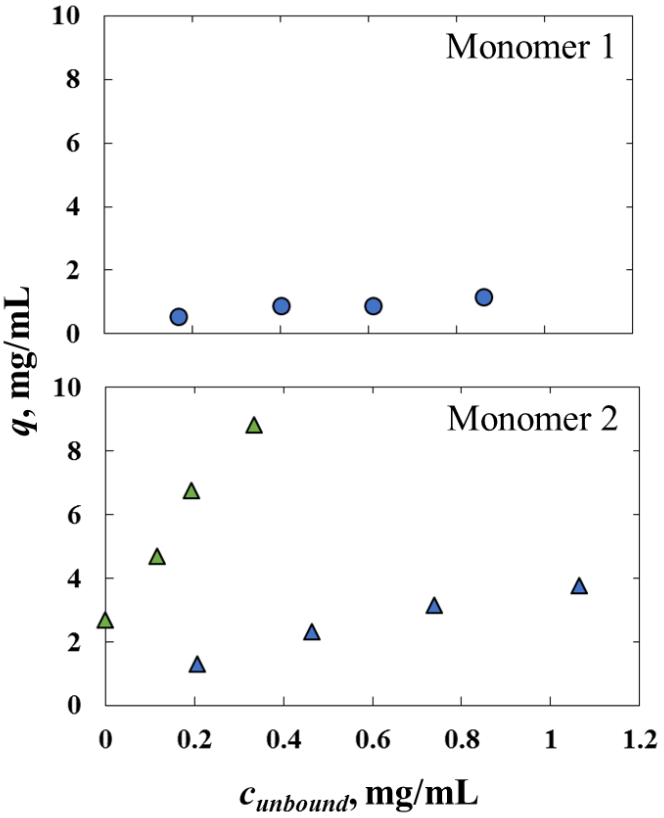


Figure 4.9: Equilibrium adsorption isotherm for the industrial protein. Monomer 1 at 0.75 mol/L of ammonium sulfate (blue circles). Monomer 2 at 0.75 mol/L of ammonium sulfate (blue triangles) and 1 mol/L of ammonium sulfate (green triangles).

For both the species, from the SEC analysis performed on the supernatant when the ammonium sulfate concentration is 1.2 mol/L, no monomer was detected. There was no unbounded protein in the liquid, at these salt and protein concentrations the adsorption is high, and all the protein is bounded to the solid. The same effect resulted with the supernatant at 1 mol/L of ammonium sulfate for the Monomer 1. Indeed, because of the low availability of the starting material, the load concentrations of Monomer 1 were lower, and all of the monomer bounded to the solid. Monomer 2, for which the availability of the starting material was higher, resulted in two isotherms, but we can notice how the isotherm at 1 mol/L of ammonium sulfate (process concentration) is steep and the adsorption is high (Figure 4.9). From the comparison between the isotherms of Monomer 1 and Monomer 2 at 0.75 mol/L of ammonium sulfate, we can recognize that the isotherm of Monomer 1 is lower. This behavior is coherent with their behavior in the column:

Monomer 1 that elute earlier has a lower hydrophobicity and then a lower adsorption. The opposite is true for the Monomer 2 that elutes later and has an isotherm that reaches higher adsorbed protein concentrations since his hydrophobicity is higher.

Because of the scarcity of isotherm data, the estimation approach was not performed for the in-house protein purification process. Furthermore, this demonstrates that the predictive approach is not suitable for such a complex system.

4.1.5 Experimental isotherms fitting

As a preliminary study equilibrium adsorption isotherm data from literature were selected between those reported by Chen and Cramer (2007). Experimental isotherms of HSA, Lysozyme, Cellulase and Catalase were considered. The proteins investigated by Chen and Cramer turned out to have different behavior with the increase of the salt concentration. As written in section 1.1, the proteins investigated were classified in three classes. In order to evaluate different behaviors, from the adsorption point of view, proteins used for this study were also selected from the different classes identified by Chen and Cramer (2007).

In particular, HSA and Lysozyme belong to the first class, while Cellulase and Catalase belong to the second class. The first class is characterized by a constant adsorption increase with the salt concentration increase. The second class of proteins has a low adsorption at low salt concentration and a quick adsorption increase when a certain value of salt concentration is reached.

Furthermore, combining the proteins belonging to different classes, can be representative of a mixture, in which the components have different behaviors and interact differently with the active site on the solid, as was shown in section 4.1.2.

This study aims to investigate the behavior of different proteins when interacting with hydrophobic resin and evaluate if the Langmuir adsorption law can describe that behavior. As written before, as a first step, the literature data were fitted with the Langmuir adsorption law at the different salt concentration (Eq. 4.2). Furthermore, the values of Langmuir parameters found were related to the salt concentration to evaluate if the exponential and power Langmuir law were suitable to fit experimental data.

$$q = \frac{ac_p}{1 + bc_p} \tag{4.2}$$

Figure 4.10 shows the Chen and Cramer (2007) experimental isotherms in Butyl Sepharose HP resin, fitted with the Langmuir adsorption law. In Figure 4.11, the values of parameters *a* and *b* are related with the respective salt concentration (values of Langmuir parameters are reported in the Appendix, Table A1). The

dependence of Langmuir parameters from the salt concentration suggests an exponential or power trend. It is possible to notice that the Langmuir adsorption law, even though it is very simple, fits the experimental isotherm data quite well (Fig. 4.10). Some anomalous values can be highlighted, but these correspond to a set of experimental data that refer to a non-monotonic behavior. In most cases the Langmuir isotherm is able to fit both the linear part of the curve, at low protein concentrations, and the asymptote at higher salt concentrations.

In the work of Chen and Cramer (2007) different hydrophobic resins have been also tested with the proteins mentioned before. Two hydrophobic resins have been used in addition to Butyl Sepharose HP already discussed: Phenyl Sepharose High Sub and Phenyl Sepharose Low Sub. These two resins have phenyl ligands with, respectively, a high and low substitution degree. In Figure 4.12 the adsorption isotherms of HSA are shown with the three different hydrophobic resins (Butyl Sepharose HP is shown again for comparison). Values of Langmuir parameters for the three resins are reported in Appendix in Table A2. The Langmuir law fits well the adsorption data also in this case, confirming that it is suitable to describe, as a first approximation, different hydrophobic systems.

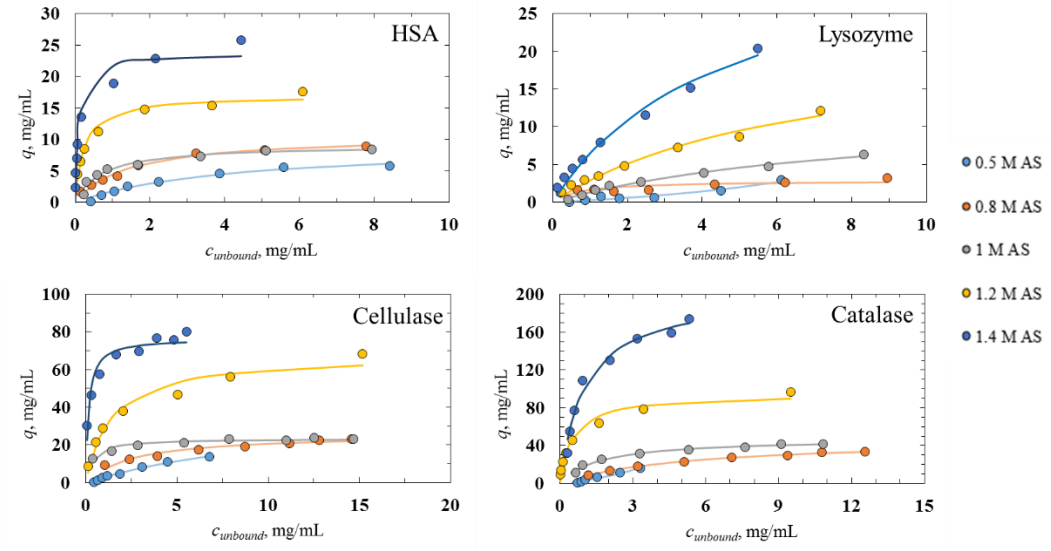


Figure 4.10: Equilibrium adsorption isotherms of HSA, Lysozyme, Cellulase and Catalase with Butyl Sepharose HP resin at different ammonium sulfate concentrations from literature data (Chen and Cramer, 2007), fitted with Langmuir adsorption law. Data from HT experiments.

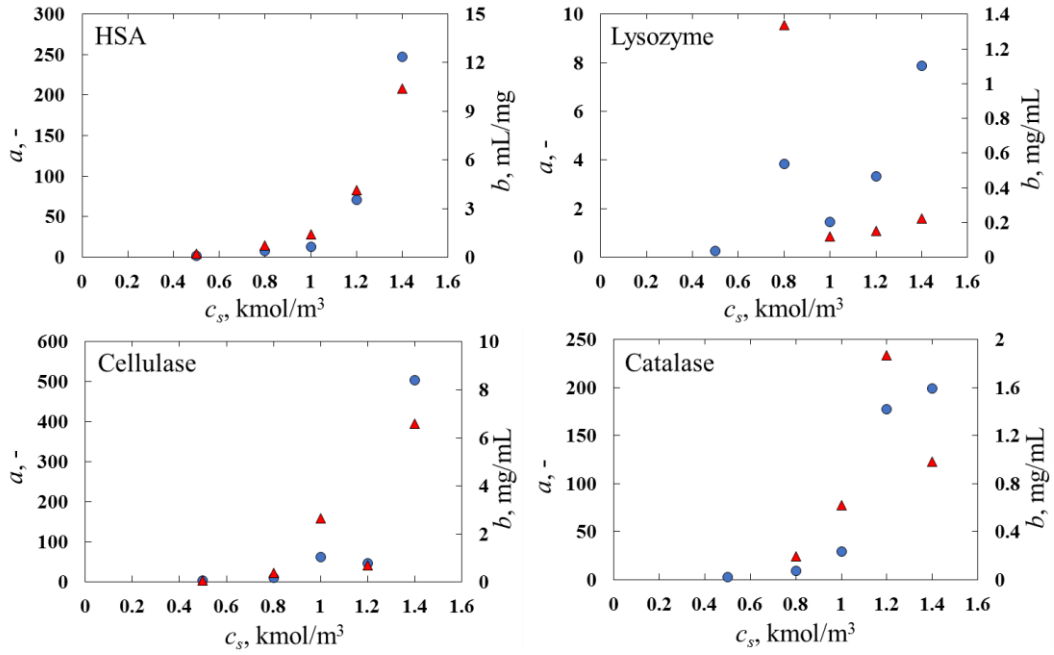


Figure 4.11: Langmuir adsorption parameters related to salt concentrations for HSA, Lysozyme, Cellulase and Catalase (fitting curves shown in Figure 4.10). Blue circles refer to *a* values and red triangles to *b* values.

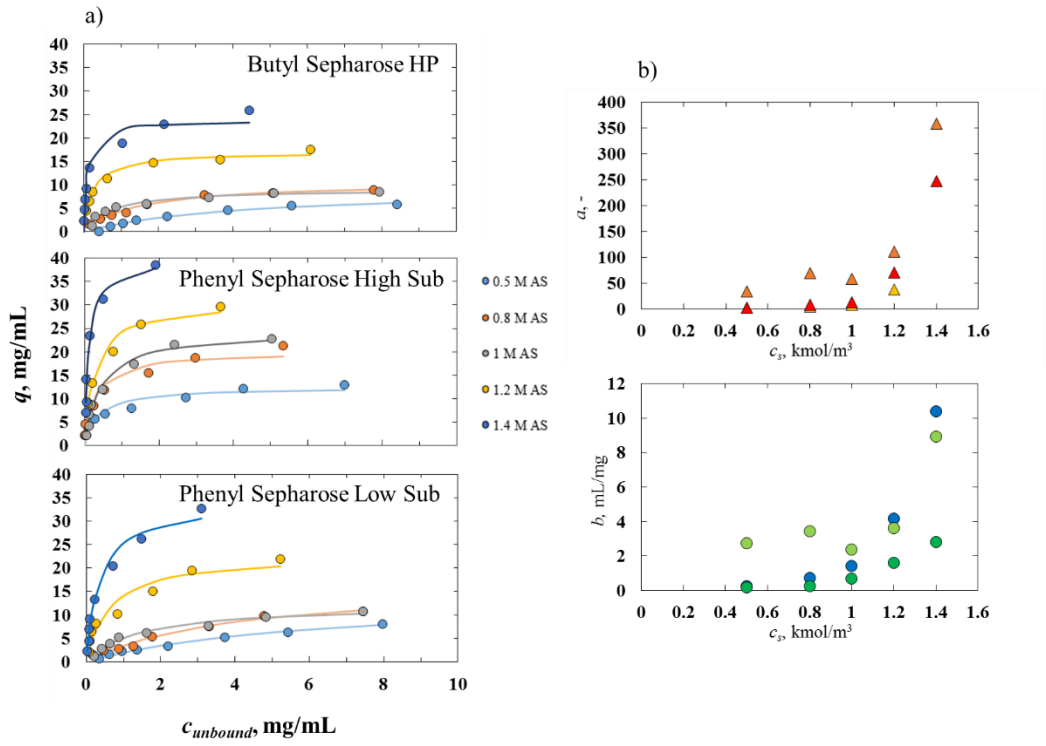


Figure 4.12: a) HSA equilibrium adsorption isotherms in Butyl Sepharose HP, Phenyl Sepharose High Sub, Phenyl Sepharose Low Sub (Chen and Cramer, 2007). b) Langmuir parameters related to the salt concentration. In the upper graph, *a* parameter where red, orange and yellow triangles are respectively the values of *a* in Butyl Sepharose HP, Phenyl Sepharose High Sub and Phenyl Sepharose Low Sub. In the lower graph, *b* parameter where blue, light green and dark green are respectively the values of *b* in Butyl Sepharose HP, Phenyl Sepharose High Sub and Phenyl Sepharose Low Sub. Data from HT experiments.

Successively, the experimental adsorption isotherms obtained with high-throughput experiments were fitted to obtain isotherm parameters. The fitting was performed on the entire bundle of isotherms to also consider the dependence from the salt concentration. The adsorption laws tested for this study were described in section 1.3. In particular, for the predictive approach, the modified Langmuir laws were used. It is known from literature (Chen and Sun, 2003) that the exponential and the power Langmuir laws generally works fine in a narrow range of salt concentration, while the salt range investigated in this work is quite wide. Furthermore, Figure 4.11 has confirmed that the trend of the Langmuir parameters can be interpreted with an exponential or power law.

When this study was performed, the experimental adsorption isotherms at 1.5 mol/L of ammonium sulfate in the buffer were excluded from the curve fitting because of the abnormal behavior, probably due to a sort of salting out effect as written in section 4.1.1 (see figure 4.1). Furthermore, in this case the adsorbed protein concentrations values were converted to *q'* for the column configuration (Eq. 1.15), which is the adsorbed protein concentration used in the in-house code, multiplying HT experimental q values for the column relative CF and considering column porosity, beside using kmol/m^3 instead of mg/mL.

In Figure 4.13 we can notice that the exponential modified Langmuir law fits quite well CRM_{197} and Lysozyme experimental adsorption isotherms. In the case of BSA, the model does not fit well at high salt concentration, in particular the relationship between the slope of the curve and the asymptote is not well interpreted by the model used. The bad fitting can be noticed also at low salt concentration for Lysozyme. Parameters obtained from the fitting are reported in Table 4.8.

Lysozyme experimental adsorption isotherms have a different shape with respect to the other proteins. As written before, Lysozyme adsorption isotherms remain in the linear part of the curve also at high salt and protein concentrations. For this reason, the linear adsorption law was also tested, with an exponential dependency from the salt concentration (Eq 1.14). The adsorption isotherm fitting with the linear model is quite similar to the exponential Langmuir law (Figure 4.14) and parameters values (Table 4.9) also are similar with those of the previous model. Indeed, Lysozyme adsorption behavior is reflected, obviously, on λ and b parameters that are very different from the other two proteins. Instead, parameters values for BSA and CRM_{197} are similar in that they have similar behavior. At low salt concentrations, the Lysozyme isotherms fitting does not work well as in the previous case.

Table 4.8: Exponential modified Langmuir law (Eq. 1.12) parameters obtained from the fitting of experimental data for the proteins investigated (corrected for column *CF*).

Model parameter	BSA	CRM_{197}	Lysozyme
λ, -	0.0004	0.0007	0.0204
b, $m^3/kmol$	389.7640	385.9780	2.8800
k, $(kmol/m^3)^{-1}$	5.5602	5.0126	4.1996

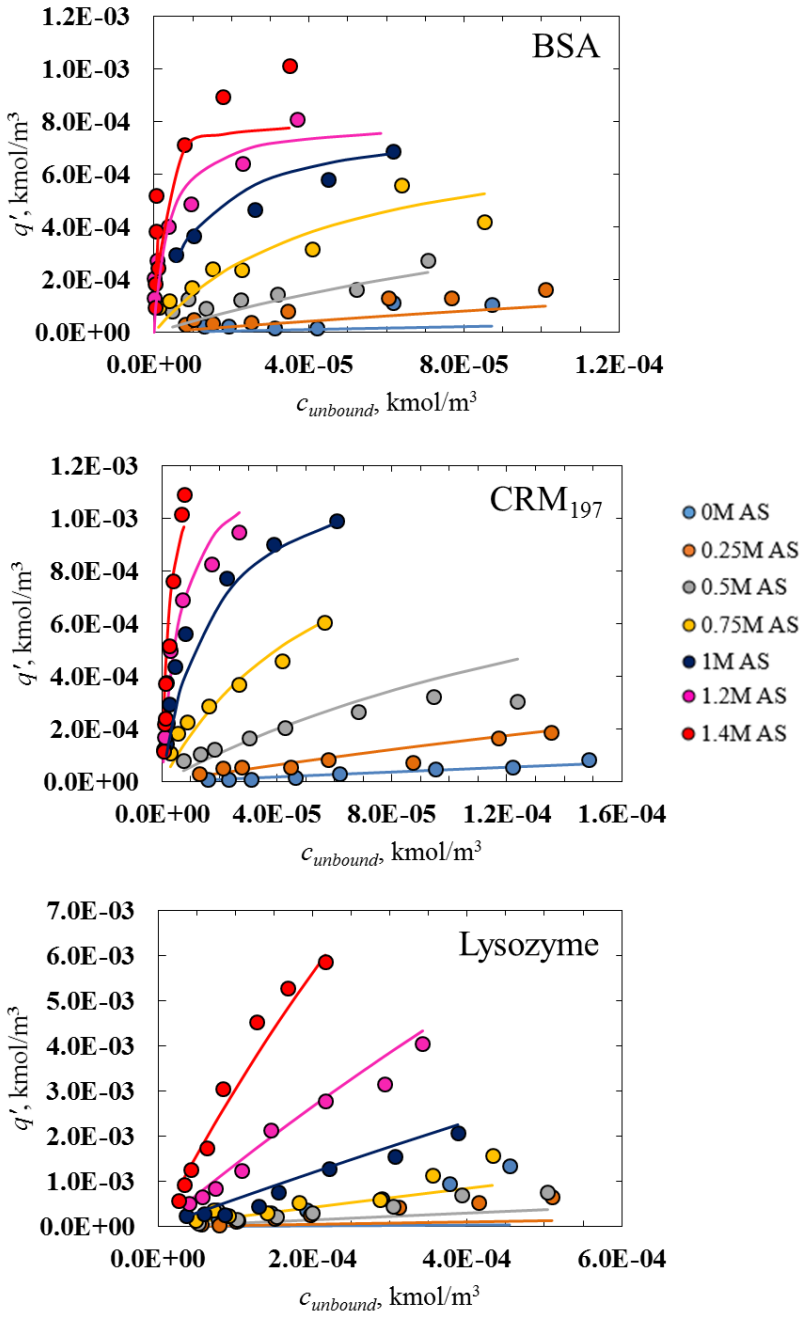


Figure 4.13: Adsorption isotherms fitting with the exponential modified Langmuir law. Experimental data corrected for column *CF*, to refer to column conditions.

Table 4.9: Linear adsorption fitting parameters (Eq. 1.14) for Lysozyme adsorption isotherms (corrected for column *CF*)..

Model parameter	Lysozyme
***a*, -**	0.0280
***k*, (kmol/m^3)$^{-1}$**	4.4588

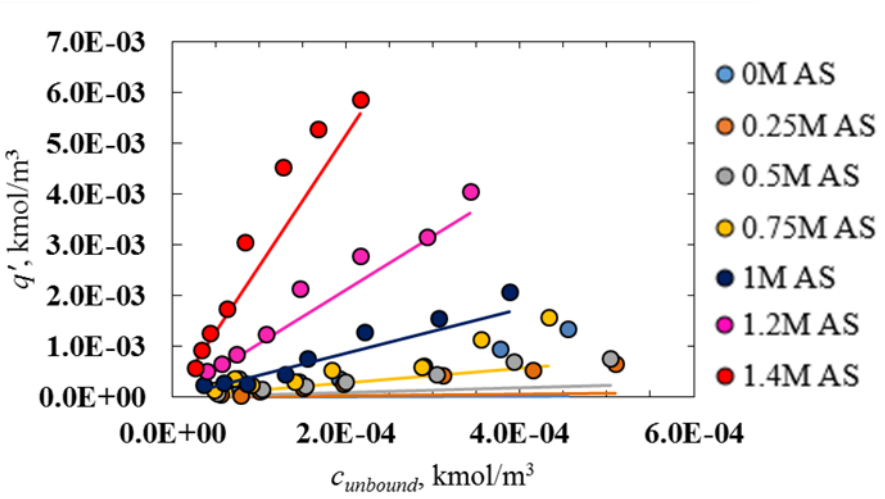


Figure 4.14: Lysozyme adsorption isotherms fitting with the exponential linear law (corrected for column *CF*).

The power modified Langmuir law was also tested to fit experimental data. This model has some constraints, besides the narrow salt concentration range in which it works well, it is not able to describe adsorption at 0 mol/L of salt since if c_{salt} is 0, q will be 0 as well (Eq. 1.13). BSA and CRM_{197} adsorption isotherms were fitted with the power Langmuir law (Fig. 4.15), while for the Lysozyme, that has linear adsorption isotherms the linear law was used (Figure 4.14).

The power Langmuir model works worse than the previous one for BSA and CRM_{197} (Figure 4.15, values in Table 4.10). It is possible to notice that, especially for BSA, the fitting is not able to describe neither the slope nor the asymptote trends.

The parameter values and models here described, determined by fitting experimental data, were used to simulate bind-elute tests with the same operating conditions of the experimental tests. Results of simulations are reported in the following section.

Table 4.10: Power modified Langmuir law parameters (Eq. 1.13) obtained from the fitting of experimental data for the proteins investigated.

Model parameter	BSA	CRM_{197}
λ, -	$4.89 \cdot 10^{-4}$	$4.92 \cdot 10^{-4}$
b, $m^3/kmol$	$1.73 \cdot 10^{5}$	$1.90 \cdot 10^{5}$
α, -	4.7375	4.4557

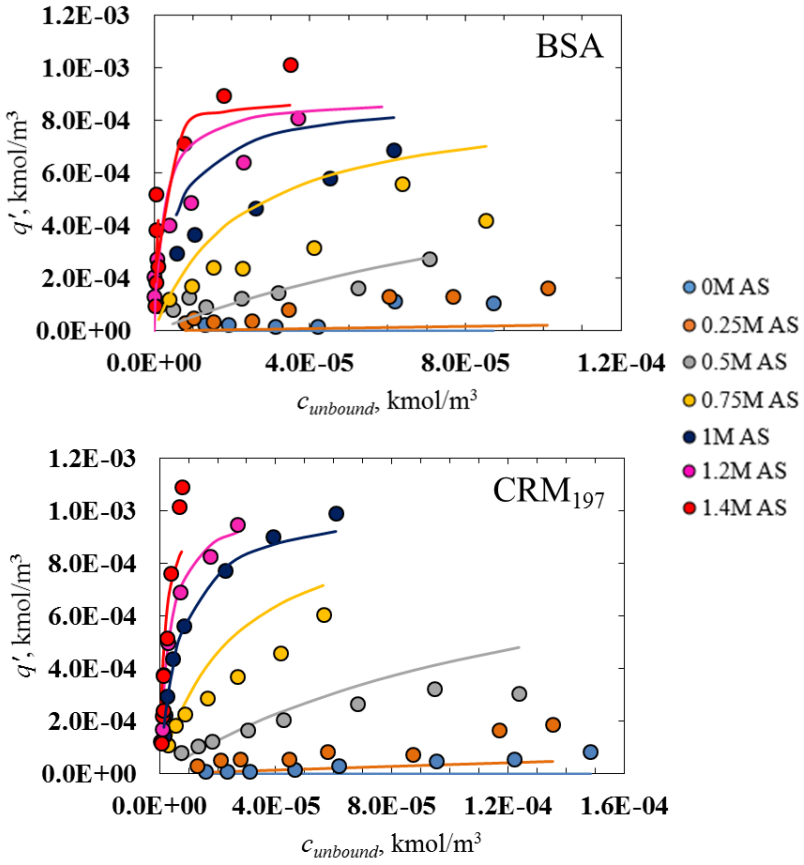


Figure 4.15: adsorption isotherms fitting with the power modified Langmuir law for BSA and CRM_{197}. Experimental data corrected for column *CF*, to refer to column conditions.

The experimental adsorption data were also compared with the predictions of the models by Mollerup et al. (2008) and Wang et al. (2016) using the parameters founds from DSPX curve fitting of experimental bind-elute runs to make a more coherent comparison between the two modelling approaches described in section 4.2. In fact, when using the estimation approach, two different adsorption models can be used in DSPX to describe hydrophobic interaction chromatography, the model developed by Wang et al. (2016) and the model of Mollerup et al. (2008)

(Eq. 1.16-1.18). As both models consider the kinetic equilibrium dynamics, the models must be modified to fit experimental isotherms that were obtained from batch experiments in which the adsorption equilibrium was reached. The right term of Eq. 1.16 and 1.17 is set equal to zero to consider adsorption equilibrium and fit the experimental adsorption isotherms. While fitting the experimental isotherm we must consider that the model parameters found with DSPX come from an estimation based on test performed in the column, while the experimental isotherms were determined in the batch high-throughput mode. The packing of the solid affects k_{eq} and q^*_{max} values and, thus, they were corrected to take into account the differences in packing and called $k_{eq,batch}$ and $q_{max,batch}$. The experimental isotherms data were fitted as a bundle of curve using parameters found by DSPX and estimating the values of $k_{eq,batch}$ and $q_{max,batch}$. Furthermore, in this case also, the values of adsorbed and unbound protein concentrations were converted in kmol/m^3 to be fitted more easily, since model parameters are expressed in kmol/m^3.

In the case of the model of Mollerup et al. (2008), as we can notice from Eq. 1.16, the isotherm equation is not explicit neither in q nor in c_p. For this reason, the isotherm of Mollerup and co-workers (2008) was used in the linear version (Mollerup, 2006) considering the experimental protein concentrations in the linear part of the isotherm only. Considering that the protein loading is quite low in the bind-elute tests performed (maximum 2 mg/mL$_{resin}$ while the maximum saturation capacity is around 40 mg/mL$_{resin}$, as it was calculated in the frontal analysis), we can assume that the amount of protein loaded is much less that the maximum saturation capacity (Eq. 4.4).

$$q^* \ll q^*_{max} \tag{4.4}$$

With this assumption, the linear Mollerup model (2006) is written (Eq. 4.5):

$$q^* = k_{eq} c_p exp(k_s c_{salt} + k_p c_p) \tag{4.5}$$

The values of $k_{eq,batch}$ found for BSA, CRM$_{197}$ and Lysozyme referred to q_{batch} are respectively 0.4745, 0.3294 and 0.0138 M^{-1} (values found from the DSPX curve fitting are reported in Table 4.17, section 4.2).

When interpreting experimental isotherms with the Mollerup linearized model (2006), for BSA and CRM$_{197}$, only the data at low protein concentration were considered, in order to fit only the linear part of the curve (Figure 4.16). In the case of Lysozyme, which experimental isotherms are linear, the whole dataset was fitted with the linearized Mollerup model (2006). It can be seen that the linear Mollerup model (2006) describes quite well the isotherms in the linear part of the curves

(Figure 4.16) for the three proteins. In particular, for Lysozyme the models is able to describe well the entire experimental dataset.

In the case of the fitting with the model of Wang et al. (2016), since Eq. 1.17 is not explicit for q, the fitting was performed on c_p from the experimental values of q. As written before, the entire bundle of isotherms was fitted to estimate the values of $k_{eq,batch}$ and $q_{max,batch}$,which are reported in Table 4.11 (values found from the DSPX curve fitting are reported in Table 4.18, section 4.2).

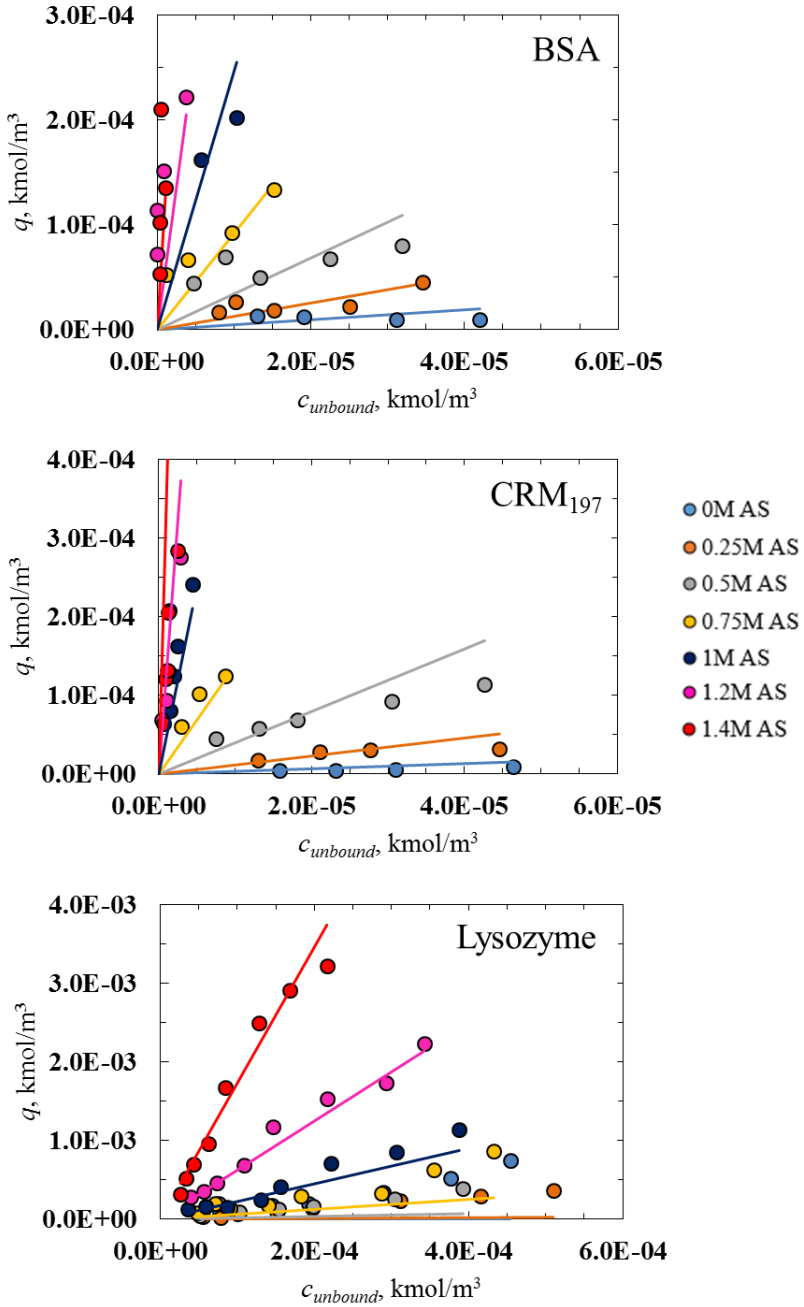


Figure 4.16: Experimental adsorption isotherms of BSA, CRM197 and Lysozyme fitted with the linear model of Mollerup (2006). Data from HT experiments.

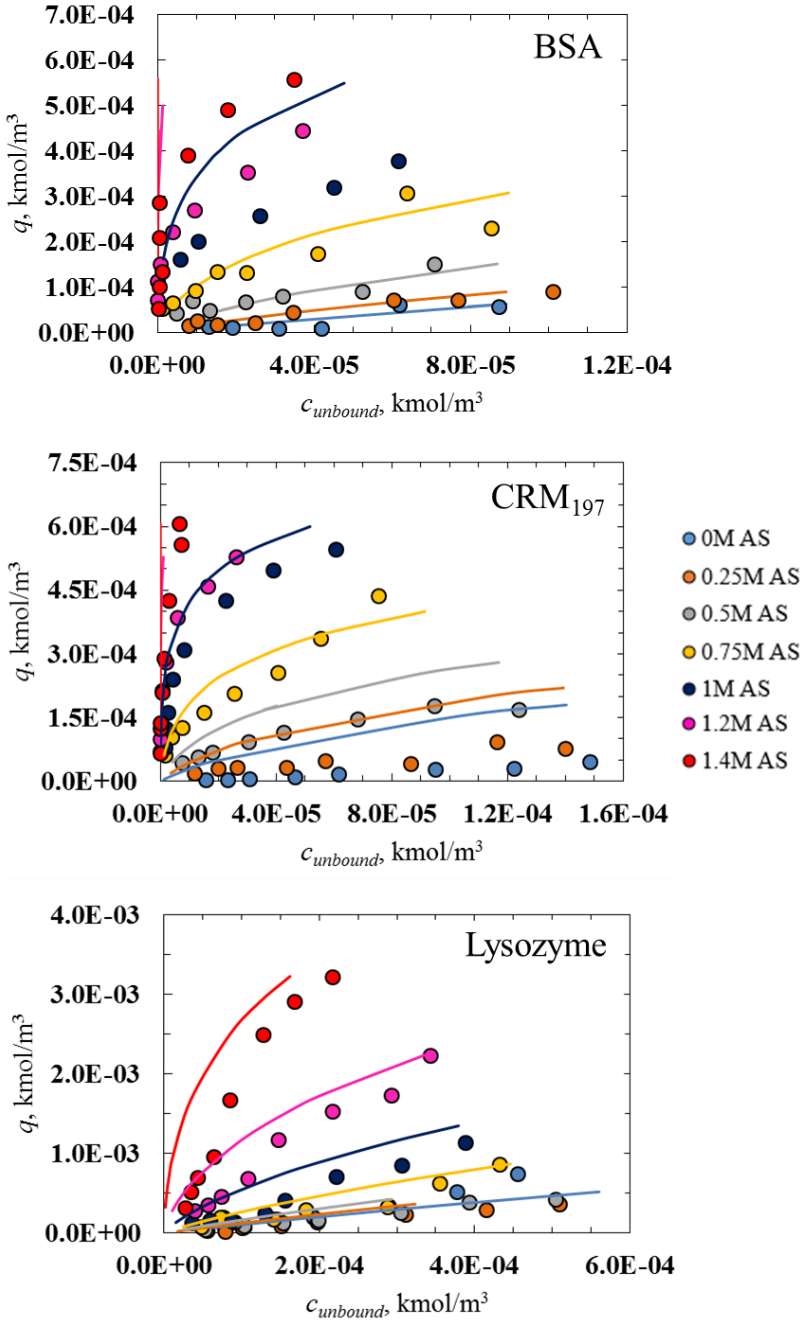


Figure 4.17: Experimental adsorption isotherms of BSA, CRM197 and Lysozyme fitted with the Wang et al. (2016) model at the equilibrium. Data from HT experiments.

The comparison with the predictions of the model developed by Wang et al. (2016) at the equilibrium is shown in Figure 4.17. The agreement is sometimes poor but the quality of agreement is comparable with those obtained in the work of Wang *et al.* (2016). The model works better in case of Lysozyme, while for the other two proteins the fitting works worse, especially at high salt concentrations.

Table 4.11: Values of model parameter adjusted for batch experiments with the Wang et al. (2016) model.

Parameter	BSA	CRM_{197}	Lysozyme
$q_{max,batch}$, $kmol/m^3$	0.0031	0.4931	0.0159
$k_{eq,batch}$, $(kmol/m^3)^{-1}$	0.0022	1.8040	0.8859

4.1.6 Evaluation of the adsorption variability due to different resin batches

To evaluate the impact of the resin ligand density variability on separation efficiency, equilibrium adsorption isotherms were evaluated for Albumin, α-chymotrypsin and β-lactoglobulin: the procedure was also repeated with BSA. The adsorption isotherms of the proteins were determined for the two resin batches available with different ligand densities. Equilibrium adsorption isotherms are determined with the already mentioned high-throughput procedure.

In Figure 4.18 the adsorption isotherms are shown comparing the two resin batches. In particular, in the picture, the variation is more evident for adsorption isotherms with 1 and 1.4 mol/L. The isotherms here reported are fitted with the exponential Langmuir law and the isotherm parameters found for the proteins in the two conditions are reported in the Appendix in Table A3.

From the different graphs of adsorption isotherms at higher salt concentrations (Figure 4.18), it can be seen that by varying the batch of resin used in the process, the adsorption isotherms of albumin and BSA varied. While α-chymotrypsin and β-lactoglobulin did not appear to be significantly influenced by the batch of resin used in the experimental phase. From the Figure 4.18, it can be seen that for albumin and BSA the use of a batch of resin with a higher density of ligands results in adsorption isotherms with a greater slope of the linear part of the curve and a higher asymptote. This behaviour was expected since the higher the ligand density, the higher the adsorption.

To evaluate how the resin ligand density affects the adsorption it is also useful to observe the behavior of the partition coefficient K_d with respect to the variation in the density of the resin ligands. The partition coefficient describes how the protein is divided between the liquid and solid phases. This coefficient can be estimated by evaluating the slope of the isotherms at the origin of the axes (Chen and Sun, 2003). Therefore, by plotting the logarithm of the partition coefficient as a function of the logarithm of the resin ligand density, it is possible to observe the relationship between the slope of the linear part of the isotherms and the ligand density of the resins studied.

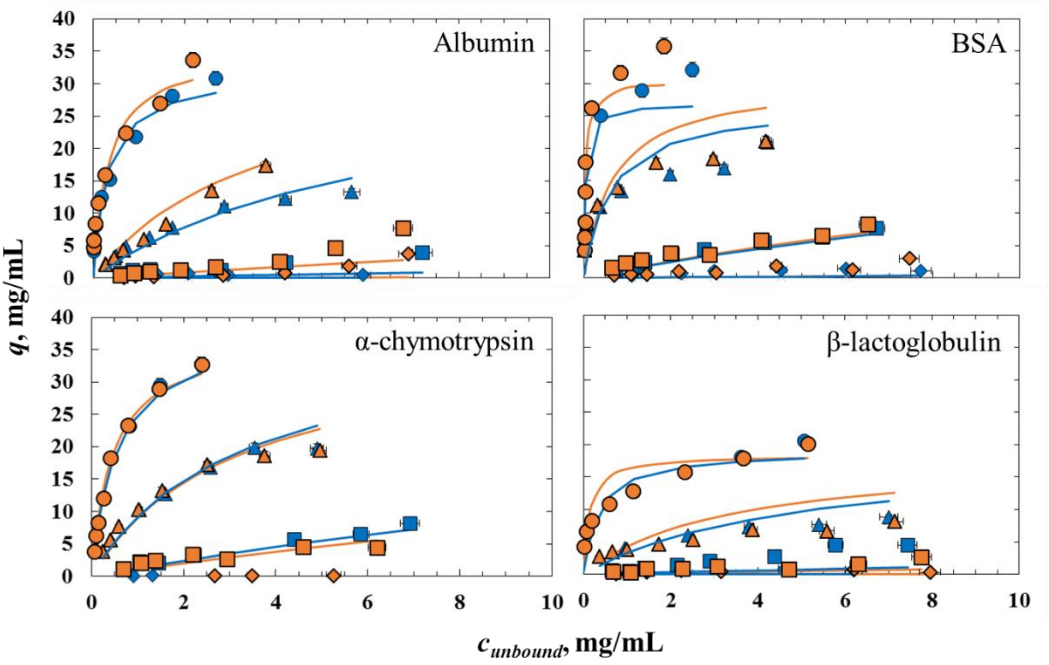


Figure 4.18: Comparison of adsorption isotherms when two different resin batches are compared. Blue symbols refer to the resin with 51 µmol/mL and orange symbols refers to resin batch with 55 µmol/mL of ligand density. Diamonds refer to isotherms at 0 mol/L of ammonium sulfate, squares refer to isotherms at o.5 mol/L of ammonium sulfate, triangles refer to isotherm at 1 mol/L of salt and circles refer to isotherms at 1.4 mol/L of salt.

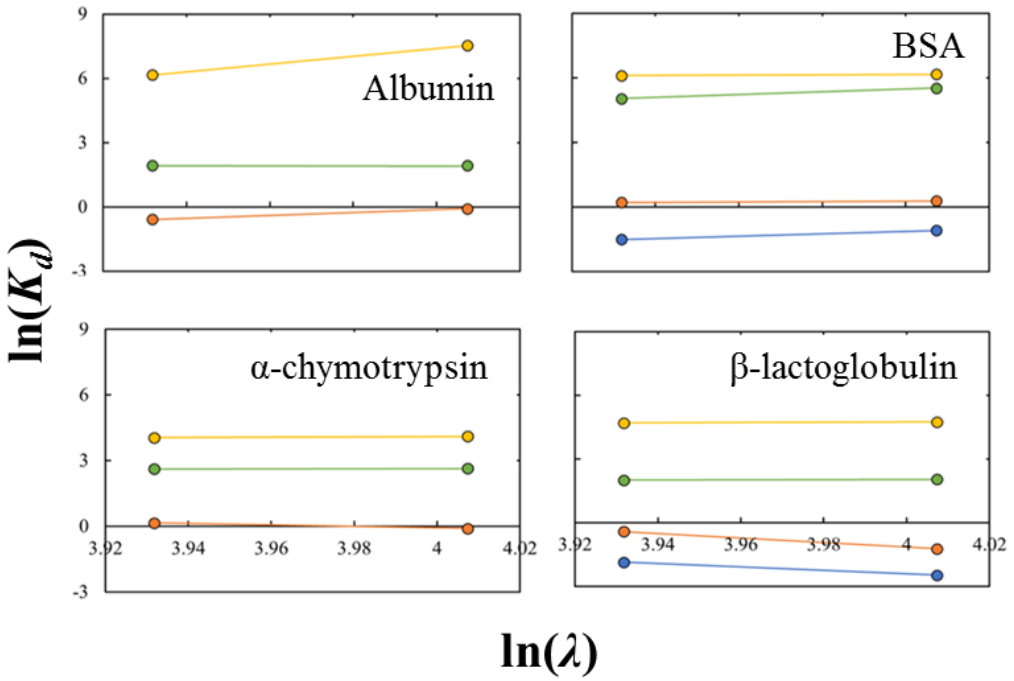


Figure 4.19: Logarithm variation of the partition coefficient as a function of the logarithm of the resin ligand density for the four proteins studied. The density of ligands is expressed in µmol/ml. Each graph contains four different sets of data in reference to the concentration of ammonium sulfate present in the solution: in blue the data relating to the no salt case; in orange the data relating to a concentration of 0.5 mol/L; in green the data relating to a 1 mol/L concentration; in yellow the data relating to a 1.4 mol/L concentration.

As can be seen in Figure 4.19, for albumin and BSA it resulted that an increase in the logarithm of the density of ligands corresponds to a slight increase in the logarithm of the partition coefficient, i.e. the slope of the isotherms in their linear section. For α-chymotrypsin and β-lactoglobulin, however, no dependence was observed.

However, it must be considered that the latter analysis is particularly approximate as it involves the determination of the slope of the isotherms in the linear part of the curve, where only few points were available. Furthermore, the data relating to a 0 mol/L concentration of ammonium sulfate are not to be considered relevant for this type of evaluation as they are characterized by a considerable dispersion, since the adsorption in this case is particularly low.

To determine how the variability of the ligand density affects the process some simulations were performed. Indeed, the parameter λ of the exponential Langmuir law can be considered related to the ligand density when some assumptions are done (Chen and Sun, 2003). We can see from Table A3 (Appendix) that the variation of the λ parameters is small when using the different resin batches. For this reason, a bind-elute simulation with isocratic elution is performed, varying the value of λ of the 50%. This variation is high and not representative of real batch variability, but it was chosen to highlight that even though the variation is high, it does not affect noticeably the elution, as we can see in Figure 4.20. The elution peak is slightly anticipated when λ is lower; this case can be representative of using a hydrophobic resin with lower ligand density in which the adsorption is lower and then, the retention is lower. The behavior is the opposite when the λ was assumed higher. It can be assumed that the variability in the resin batches, for the conditions here investigated does not affect the system in a way that may affects the separation efficiency significantly.

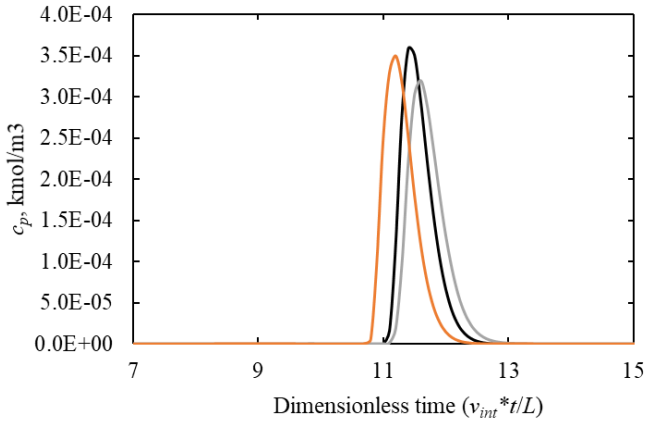


Figure 4.20: Elution peaks simulated for a bind-elute test performed with BSA changing the values of the λ parameter. The black, orange and grey curves refer to elution peaks obtained using, respectively, default λ value, λ+50% and λ-50% values.

4.2 Predictive and estimation modelling approach: simulated and experimental results

In this section, two different modeling approaches were compared. From one side, simulations were performed with the in-house Fortran code (see section 3.2) that solves the system of partial differential equations and gives concentration profiles of the components at the column outlet with the adsorption model chosen. On the other hand, a commercial software DSPX was used to estimate model parameters by curve fitting of experimental chromatograms.

The two codes were first compared to investigate the consistency between them. An *in silico* run was performed considering an ideal system. The simulated run consists in an injection of BSA in the column that respect the Langmuir adsorption law in equilibrium assumption and without dependency from the salt concentration, that is considered constant during the run. The eluted peaks obtained with DSPX and the in-house code resulted consistent, thus the modeling tools can be compared (Figure 4.21).

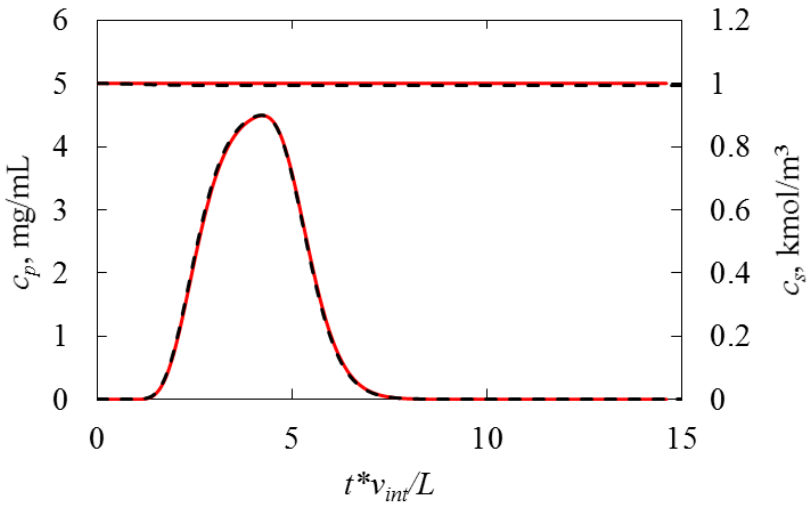


Figure 4.21: Comparison between the elution peak simulated with DSPX (red line) and the in-house code (black dashed line).

The isotherm parameters found from the fitting of experimental adsorption isotherm data (see the previous section), were used in the Fortran code to predict the elution peaks for the bind-elute tests. Simulated and experimental bind-elute tests were compared to evaluate the model chosen. This kind of modelling approach is predictive since isotherm and other model parameters were evaluated *a priori* and then simulation were performed based on operating conditions of experimental tests (the bind-elute tests in this case reported in Tab. 2.1, but it can be applied also to breakthrough or flowthrough mode).

Values of mass transfer and adsorption parameters (Tables from 4.12 to 4.16) were given as inputs to the code together with the elution method and components concentration.

Table 4.12: Column parameters evaluated with procedure described in section 2.2

V_d, mL	0.39
ε_t	0.958
ε_c	0.367
ε_p	0.934
D_{ax}, m²/s	$9.13 \cdot 10^{-7}$

Table 4.13: Ammonium sulfate mass transfer parameters calculated with Equations from 1.9 to 1.11.

Ammonium sulfate parameters	
D_m, m²/s	$4.86 \cdot 10^{-10}$
D_p, m²/s	$4.00 \cdot 10^{-10}$
k_c, m/s	$1.29 \cdot 10^{-4}$

Table 4.14: BSA mass transfer parameters calculated in different salt conditions.

Parameter	BSA		
	1 mol/L salt	1.2 mol/L salt	1.4 mol/L salt
D_m, m2/s	$4.24 \cdot 10^{-11}$	$4.00 \cdot 10^{-11}$	$3.94 \cdot 10^{-11}$
D_p, m2/s	$3.49 \cdot 10^{-11}$	$3.29 \cdot 10^{-11}$	$3.24 \cdot 10^{-11}$
k_c, m/s	$2.52 \cdot 10^{-5}$	$2.42 \cdot 10^{-5}$	$2.40 \cdot 10^{-5}$

Table 4.15: CRM_{197} mass transfer parameters calculated in different salt conditions.

Parameter	CRM197		
	1 mol/L salt	1.2 mol/L salt	1.4 mol/L salt
D_m, m2/s	$4.42 \cdot 10^{-11}$	$4.17 \cdot 10^{-11}$	$4.11 \cdot 10^{-11}$
D_p, m2/s	$3.63 \cdot 10^{-11}$	$3.42 \cdot 10^{-11}$	$3.38 \cdot 10^{-11}$
k_c, m/s	$2.59 \cdot 10^{-5}$	$2.49 \cdot 10^{-5}$	$2.47 \cdot 10^{-5}$

Table 4.16: Lysozyme mass transfer parameters calculated in different salt conditions.

Parameter	**Lysozyme**		
	1 mol/L salt	1.2 mol/L salt	1.4 mol/L salt
D_m, **m2/s**	$7.07 \cdot 10^{-11}$	$6.66 \cdot 10^{-11}$	$6.57 \cdot 10^{-11}$
D_p, **m2/s**	$5.81 \cdot 10^{-11}$	$5.47 \cdot 10^{-11}$	$5.40 \cdot 10^{-11}$
k_c, **m/s**	$3.55 \cdot 10^{-5}$	$3.41 \cdot 10^{-5}$	$3.38 \cdot 10^{-5}$

In Figure 4.22 the experimental conductivity signal is shown, together with the outlet protein concentration, for the gradient and isocratic case.

In Figure 4.23, experimental and simulated peaks with both exponential and power Langmuir law are compared. The gradient and isocratic elution were performed, respectively, with a 1 mol/L and 1.2 mol/L ammonium sulfate in the buffer. When the exponential and power Langmuir law (Eq. 1.12 and 1.13) were used to simulate the bind-elute tests, the differences in the elution peaks were small (Figure 4.23).

First, we must consider that, when the elution occurs, the chromatographic system turns off the salt buffer line pump and turns on the no-salt buffer line pump in the case of isocratic elution; in the case of gradient elution the two pumps are working together to decrease gradually the salt buffer flowrate and, at the same time, increase the no-salt buffer flow rate. These mechanisms are affected by several phenomena that cannot be easily determined, as back-mixing in the lines or dead volume between the buffer pumps, the column, and the sensors.

It is possible to notice, from Figure 4.22, that the experimental conductivity signal has a roundish profile, while theoretically the gradient should be straight. This may be due to the imperfect and not immediate mixing of the salt and no-salt buffers during the gradient elution, or to an effect of the ionic strength on the conductivity signal. Experimental data are available in literature (Emerson, 2010) that show the dependency of the conductance from the solute concentration in water. With the ammonium sulfate concentrations used for the bind elute test (from 1.2 mol/L to 0 mol/L) the trend of the conductance resulted quite linear. The roundish shape of the conductivity signal could therefore be due to a combination of these two effects that are not easy to determine. In Figure 4.23 (and later in Figure 4.24), the experimental conductivity was diagrammed using the same scale as the salt, in order to compare more easily the two profiles.

The adsorption laws used for this study only consider the dependency of model parameters from the salt concentration and, for this reason, the effect of the salt concentration on the specific conductivity was considered negligible in the

interpretation of the results. The salt elution profile in both isocratic and gradient elution was simulated in order to coincide with the experimental conductivity profile, in such a way that the simulation was the more faithful to the real system.

We can notice, from Figure 4.23, that the gradient elution was worse predicted for all the three proteins. In general, the exponential Langmuir law gave narrower peaks with both elution methods. The simulated peaks are sharper and more symmetric than the experimental peaks. The real system, indeed, is different from the ideal one summarized in section 1.3. Different factors can affect the symmetry and the width of the peaks, like the goodness of the packing of the bed, homogeneity of beads shape and dimension, but also obstacles in the pipelines that affect the axial dispersion in the dead volume. The three proteins investigated in this study showed different behavior during elution (Figure 4.23), consistently with their equilibrium adsorption isotherms behavior. Indeed, we can see that BSA and CRM_{197} have shown similar elution peaks during elution and the adsorption isotherm of the two proteins have similar behavior. Lysozyme elution behavior is different. In particular, we can notice that, in the gradient elution, the experimental elution peak occurs before the conductivity decreases. Indeed, the gradient bind-elute test is performed with a 1 mol/L of ammonium sulfate in the equilibration and load buffer, and the adsorption in this condition is quite low for the Lysozyme, how we can infer from the equilibrium adsorption isotherm of the protein (Figure 4.1). In these conditions the salt concentration is not high enough to let the Lysozyme attach to the solid and it elute before the conductivity decreases.

Generally, the exponential modified Langmuir law seemed to better describe the system. Lysozyme is an exception in this case also: the linear model with the exponential dependency from salt concentration worked better than the exponential modified Langmuir law, besides being a simpler model. Indeed, the fitting with the linear law better described the experimental adsorption isotherm data, especially at low protein concentrations, where the conditions are more similar with those of the bind-elute tests. When the exponential modified Langmuir law was used (Figure 4.13), it is possible to notice that the fitting isotherms are overestimating the adsorption in the initial part of the curve. This overestimation of the adsorption reflected on the simulated bind-elute test (Figure 4.23): the simulated elution peak occurs slightly later than the experimental peak.

It must be considered that the eluted peak, for all the runs below is the one in correspondence of the salt concentration decrease (Figure 4.22). The other peaks that can be noticed in the chromatograms are relative to the water stripping phase, that the model was not able to consider, performed at the end of the run. However, these peaks were quite narrow and, thus, relative to a very small amount of protein that can be considered negligible compared with the main peak.

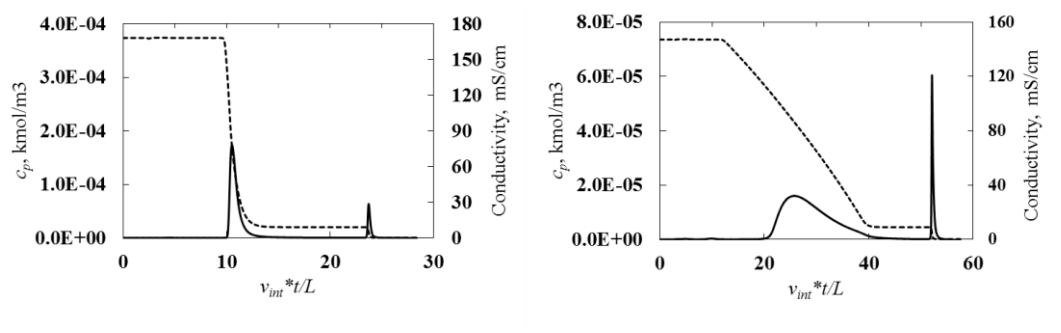


Figure 4.22: Example of bind-elute test performed with BSA, on the left the isocratic elution and on the right the gradient elution. The continuous line is the outlet protein concentration, while the dashed line is the experimental conductivity profile.

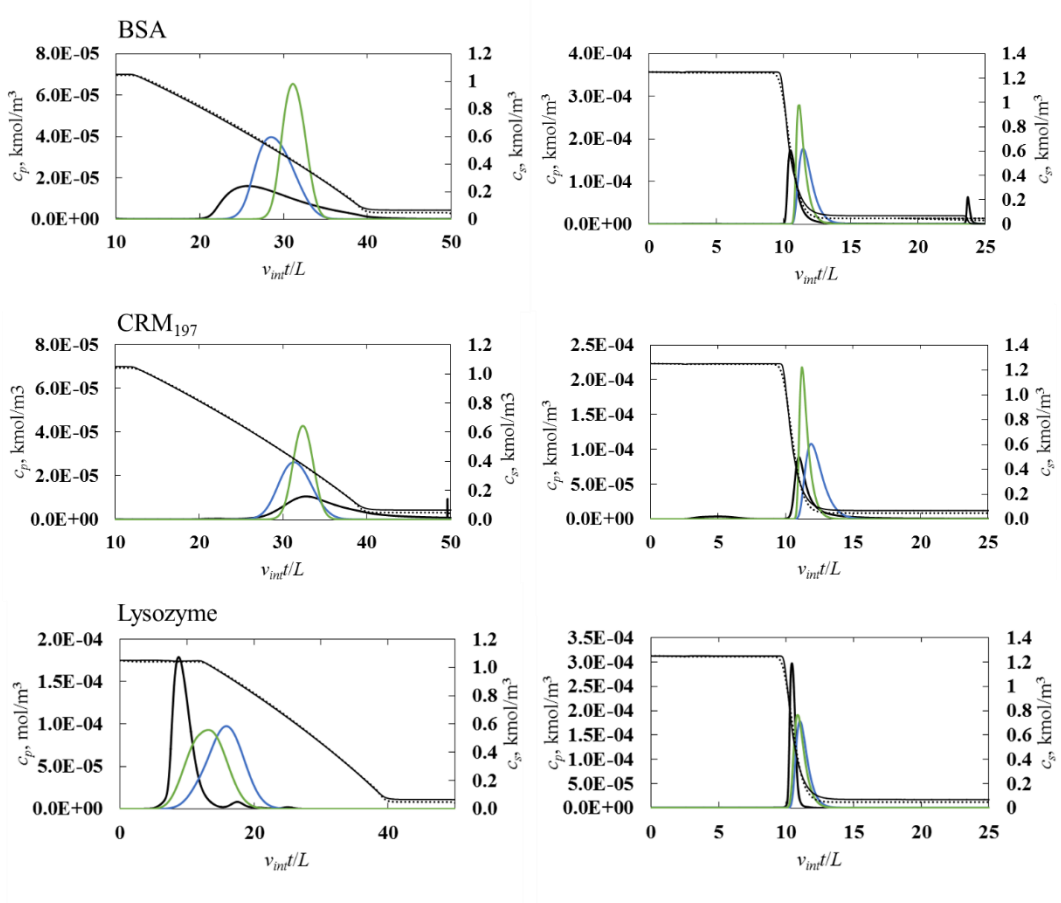


Figure 4.23: comparison of experimental and simulated bind-elute tests on BSA, CRM_{197} and Lysozyme. On the left the gradient elution, on the right the isocratic elution. The thinner black line is the scaled conductivity signal, the pointed line is the simulated outlet salt concentration. The black peak is relative to the experimental protein concentration, the blue line the simulated protein concentration when exponential Langmuir is used and the green line is the simulated outlet protein concentration when the power Langmuir law is used. In the case of Lysozyme, the green line refers to the eluted peak when the linear model is used.

The same experimental bind-elute tests, both isocratic and gradient elution, were used also for the estimation of model parameters with DSPX. Chromatograms resulted from the experimental tests performed were imported into the software and

the estimation of model parameters was run. The same values of axial dispersion and mass transfer resistance coefficient previously used in the predictive approach were employed, data were already reported in Tables 4.12 to 4.16. Performing estimation with DSPX, both the Mollerup et al. (2008) and Wang et al. (2016) models were used and compared. Results of the simulation are reported in Figure 4.24; parameters values for each protein relative to the estimation are reported in Tables 4.17 and 4.18.

It must be considered that in this case, the outlet salt concentration resulted slightly different from the experimental normalized conductivity. With DSPX, indeed, it is not possible to impose the salt concentration profile, as was done in previous case, but the profile is calculated by the software from the elution method given and the characteristics of the system set.

Table 4.17: Values of model parameter obtained with estimation on DSPX with the Mollerup et al. (2008) model.

Parameter	BSA	CRM_{197}	Lysozyme
k_{kin}, s	99.9927	63.3131	1.5013
k_{eq}, -	13.0609	9.0676	0.0878
k_s, $(kmol/m^3)^{-1}$	3.9516	4.9833	5.0967
k_p, $(kmol/m^3)^{-1}$	$1.00 \cdot 10^{-6}$	$1.40 \cdot 10^{-6}$	$1.37 \cdot 10^{-6}$
n, -	9.8474	1.7633	8.9567
q^*_{max}, $kmol/m^3$	0.0376	0.5377	0.5744

Table 4.18: Values of model parameter obtained with estimation on DSPX with the Wang et al. (2016) model.

Parameter	BSA	CRM_{197}	Lysozyme
k_{kin}, s	21.3236	73.8006	0.4162
k_{eq}, -	12.5240	20.0277	4.8991
β_0, -	0.0059	0.0059	0.0059
β_1, $(kmol/m^3)^{-1}$	2.4386	2.4386	2.4386
n, -	9.9954	9.9877	3.8541
q^*_{max}, $kmol/m^3$	0.1008	0.1000	0.6822

The two models tested with DSPX turned out to describe quite well the bind-elute chromatograms used for the curve fitting. The peak retention of the simulated elution peaks in most cases coincides with that of the experimental elution peaks. The adsorption parameters found seem to reflect the protein adsorption behavior:

BSA and CRM_{197} has a similar behavior, while Lysozyme is different (Tables 4.17 and 4.18).

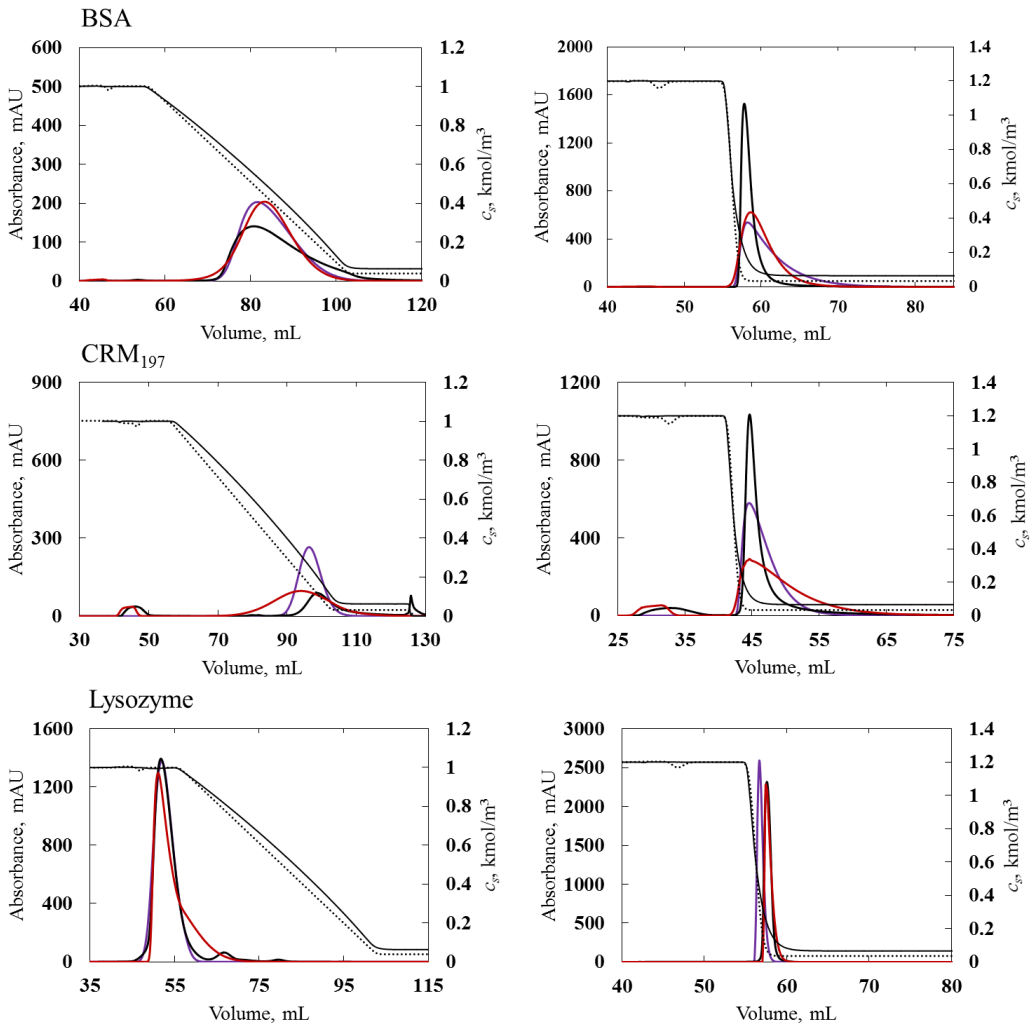


Figure 4.24: Comparison between experimental and simulated bind-elute tests chromatograms of BSA, CRM_{197} and Lysozyme after the model parameter estimation. To the left the gradient elution, to the right the isocratic elution. The black, purple and red lines are, respectively, relative to the experimental peak, the peak simulated with the model of Mollerup et al. (2008) and the peak simulated with the model of Wang et al. (2016). The thinner black line is the experimental scaled conductivity while the dashed line is the simulated outlet salt concentration.

A validation test was carried out to verify if the model, with the fitted parameters, can predict the system in different operating condition. For this reason, for each protein, a further bind-elute test was performed. For all the proteins investigated a 1.4 mol/L of ammonium sulfate buffer is used for equilibration and load phases; the elution was performed with a gradient, but the gradient was performed in five column volume instead of ten (Figure 4.25). Keeping the parameters found during estimation, the validation test was imported in DSPX and the run was simulated.

In this case also only the peaks relative to the salt concentration decrease must be considered, the other peaks are relative to the water stripping phase and, in the case of CRM_{197}, the stripping with sodium hydroxide was also performed at the end of the run.

In Figure 4.26 results of the validation runs are shown. It can be noticed that both Mollerup et al. (2008) and Wang et al. (2016) models work very well with the validation experiments. In particular, in the case of BSA, the simulated peaks are overlapping. The case of the CRM_{197} is slightly less consistent. This behavior could be due to the not correct fitting of the pre-peak present before the elution peak (that is an impurity present in the purified CRM_{197} and cannot be removed, present in very low quantities, and considered negligible). The model of Mollerup et al. (2008) was not able to fit this behavior and the elution peak has a bigger area. However, the retention time of the peak is well predicted in both cases.

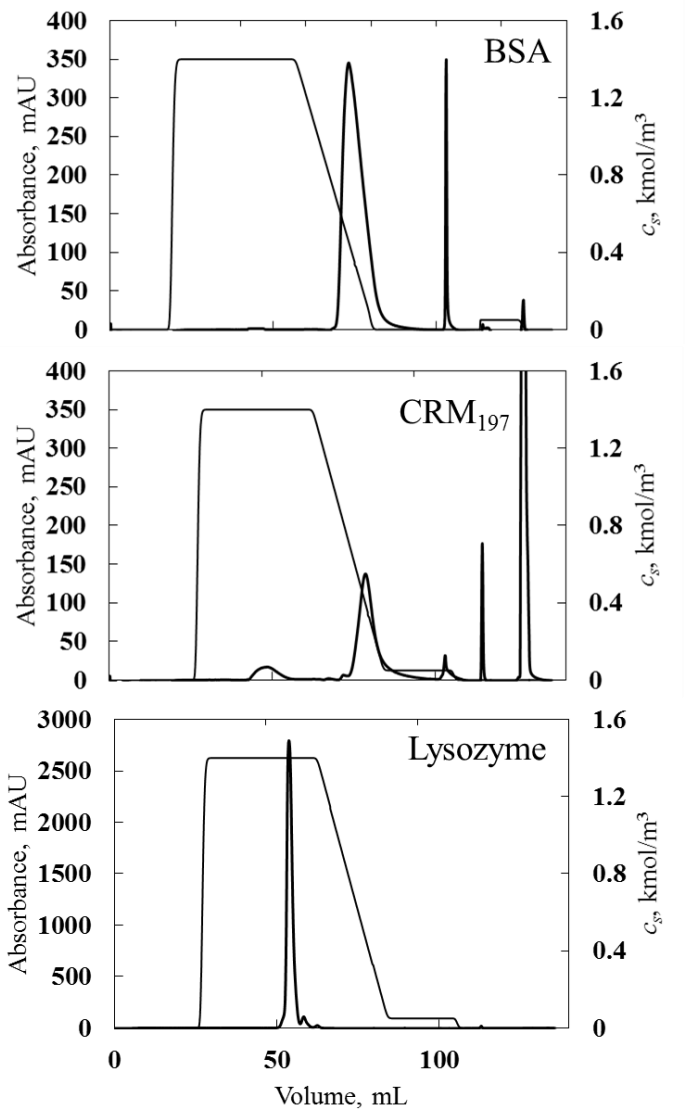


Figure 4.25: Chromatograms of the validation runs for BSA, CRM_{197} and Lysozyme to test the DSPX model, the thicker line is relative to the protein absorbance, while the thinner line refers to the outlet salt concentration.

Both models are considered robust and valid and can be used for other operating condition with the components used.

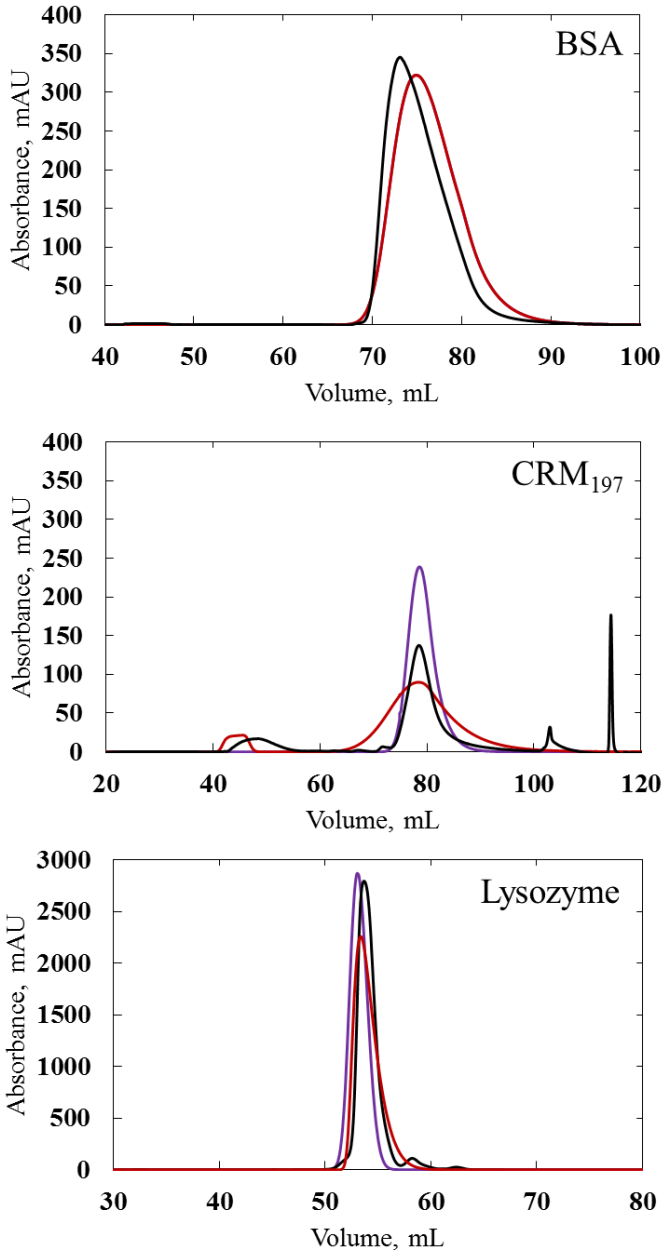


Figure 4.26: Validation runs for BSA, CRM_{197} and Lysozyme to test the DSPX model, the violet line represents the absorbance profile when the model Mollerup et al. (2008) is tested, the red line is the absorbance profile when the model of Wang et al. (2016) is tested, the black line is the experimental absorbance profile. BSA simulated peaks are overlapped.

This kind of approach is effective, the estimation with DSPX allowed to find model parameters that resulted in elution peaks quite faithful to the experimental one. On the other hand, to perform the estimation no information about the adsorption dynamics were needed, adsorption parameters are estimated only on the base of the elution peaks. This approach, thus, does not allow us to gain knowledge on the adsorption mechanisms.

In the predictive approach we investigated the adsorption behavior to obtain model parameters that were used for the simulation. In this case the knowledge gained was huge but the experimental effort was high. Furthermore, this kind of approach can be applied to protein with a high purity and low cost, otherwise the effort is not sustainable.

4.3 Modelling of the in-house protein purification process

The industrial in-house process chosen for this study, as written in section 2.7, consists in the purification of a target product (the monomer), from impurities consisting of high and low molecular weight species. The estimative modelling procedure was applied to this industrial process, since the process is quite complex and consists in a purification with different species. In order to estimate model parameters, several experiments were performed on this process. The model chosen was then validated with a run in operating conditions different from those used for the model parameters estimation. At the end, using the set of parameters found, the process has been optimized maximizing the yield and purity of the target product among the other species.

In DSPX it is possible to choose the model that better represent the system, in terms of mass transfer and adsorption dynamics. For the industrial process, since it is quite complex and involve different species, the Transport Dispersive model (Eq. 1.2) was maintained for the mass transfer in the interstitial column volume, while the Lumped Rate model was chosen for the mass transfer in the pore of the beads (Eq. 1.20). Furthermore, the adsorption law chosen for this study is the multi-component Wang et al. model (2016) with a modification for the pH dependency (Eq. 1.21).

The chromatogram resulting from the reference run (with the standard values of pH and monomer loading, but the simplified elution method) is shown in Figure 4.27. It can be noticed, from the chromatogram, that there are two peaks in the first step of elution and a third peak in the second elution step. With this method (23% of B pump in the first step and 100% in the second step) the eluting material in correspondence of the water stripping is negligible, because everything is removed in the last step (100% of B pump).

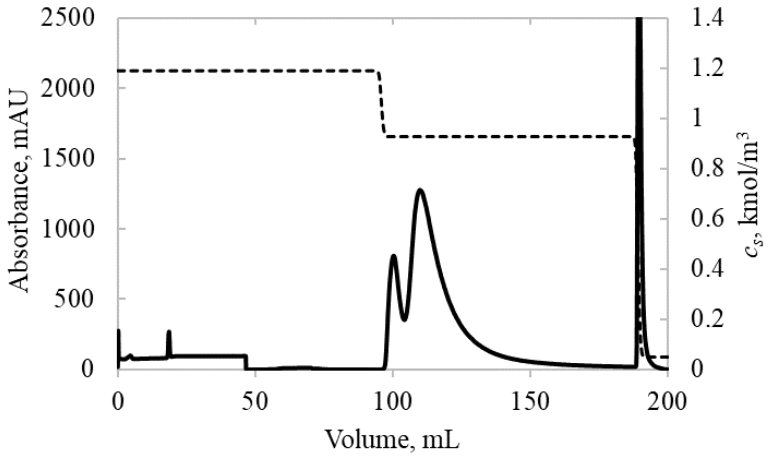


Figure 4.27: Chromatogram of the simplified standard process.

As reported before, the target product is a monomer that is present in four charge variants. In particular, the four charge variants elute in pairs: two variants elute in the first peak and the other two monomer variants elute in the second peak. As described in section 2.7, from the chromatogram we can infer that in the first peak, since it is quite sharp and symmetric, the two variants have the same hydrophobicity. The second peak has a tail that makes it asymmetric, and this behavior could be due to a small difference in the hydrophobicity of the two charge variants that elute here: the first one representative of the main peak, and the second one, eluting slightly later, responsible of the tailing of the peak. A small amount of monomer is also eluting in the third peak, probably carried away by the impurities during the elution. In order to interpret the behavior of the monomer during the purification, some assumptions were done. In the first peak the two variants were considered as a unique specie, since they elute together, and can be considered having the same adsorption behavior: this specie was called Monomer 1. In the second peak, since the tail of the peaks seems to suggest the presence of two variants with different behavior, two species were considered eluting here, called Monomer 2a and Monomer 2b. Furthermore, since it is not perfectly clear why in the third peak there is still the monomer, the monomer eluting in this peak was considered as another species called Monomer 3.

High and low molecular weight compounds (called HMWC and LMWC) were considered as two different species eluting together with the monomer during the whole purification.

The SEC analysis allows to know the monomer concentration and purity per each peak. It is not possible to know the real concentration of HMWC and LMWC since they are a set of aggregates and truncations of the monomer and do not represent real components, thus it is not possible to build a calibration curve to convert the peak area from the SEC in concentration values. For this reason, another

approach was adopted: concentration values of the HMWC and LMWC are obtained from the values of purity and concentration of the monomer (Eq. 4.8 to 4.10, where P_i indicates the purities and c_i the concentration of the species), assuming for the HMWC and LMWC pseudo-components during the SEC separation the same intensity of the signal of the monomer. Indeed, using the Eq. 4.8 to 4.10 to calculate concentrations of HMWC and LMWC, means that the peak area referred to these species is converted into concentration using the calibration line relative to the monomer and, thus, a monomer-intensity equivalent concentration is considered.

$$c_{tot} = \frac{c_{mon} \cdot 100}{P_{mon}} \tag{4.8}$$

$$c_{HMWC} = \frac{c_{tot} \cdot P_{HMW}}{100} \tag{4.9}$$

$$c_{LMWC} = \frac{c_{tot} \cdot P_{LMW}}{100} \tag{4.10}$$

This assumption causes an error in the evaluation of the unknown compounds' concentration, but this error is considered acceptable since there are no methods available to determine HMWC and LMWC concentration values.

Experimental online (chromatograms) and offline (SEC analysis results) data were imported in DSPX and the model parameter estimation was performed. All the runs reported in Table 2.2 were used for the curve fitting. Results of the parameters estimation are shown in Figure 4.28 to 4.33, and the parameters values are reported in Table 4.19 (see Eq. 1.21 for reference).

In the figures below (left side), the black dashed line refers to the experimental absorbance, the pink line is the simulated absorbance. In the graphs in the right side, the red line refer to the Monomer 1 concentration, dark blue line to the Monomer 2a concentration, green line to Monomer 2b concentration and purple line to Monomer 3 concentration. Light blue and yellow line refer to, respectively, LMWC and HMWC concentration.

From the eluted profiles of the different components, it is possible to notice that the retention time of each component is well predicted and respects the assumption done on the components and the relative elution behavior. The HMWC and LMWC are worse predicted, but these pseudo-components are the more affected by the assumptions done, as they are classified in only two species and the concentration values are approximately estimated. These assumptions reflect on the predicted

peaks of HMWC and LMWC: the LMWC elute in the flowthrough in most of the runs simulated and the LMWC, especially in the run at pH 8, are eluting mostly at the end of the second peak, which is not consistent with the real system.

Table 4.19: Wang et al. (2016) parameter values obtained with estimation on DSPX, Eq. 1.21.

Parameter	Monomer 1	Monomer 2a	Monomer 2b	Monomer 3	HMWC	LMWC
β_0, -	$2.96 \cdot 10^{-3}$	$2.96 \cdot 10^{-3}$	$2.96 \cdot 10^{-3}$	$2.96 \cdot 10^{-3}$	$2.96 \cdot 10^{-3}$	$2.96 \cdot 10^{-3}$
β_1, kmol/m^3	$2.25 \cdot 10^{5}$	$2.25 \cdot 10^{5}$	$2.25 \cdot 10^{5}$	$2.25 \cdot 10^{5}$	$2.25 \cdot 10^{5}$	$2.25 \cdot 10^{5}$
$k_{eq,1}$, -	$3.68 \cdot 10^{-1}$	$3.11 \cdot 10^{-1}$	$3.97 \cdot 10^{-1}$	$-9.80 \cdot 10^{-1}$	$5.46 \cdot 10^{-1}$	$5.80 \cdot 10^{-1}$
$k_{eq,2}$, -	0.00	0.00	0.00	0.00	0.00	0.00
k_{eq}, -	$3.21 \cdot 10^{5}$	$1.66 \cdot 10^{5}$	$2.23 \cdot 10^{5}$	$3.34 \cdot 10^{5}$	$1.18 \cdot 10^{5}$	$2.51 \cdot 10^{-1}$
k_f, m/s	$9.70 \cdot 10^{-3}$	$9.70 \cdot 10^{-3}$	$9.70 \cdot 10^{-3}$	$9.70 \cdot 10^{-3}$	$3.62 \cdot 10^{-2}$	$9.75 \cdot 10^{-2}$
k_{kin}, s	$1.02 \cdot 10^{5}$	$1.05 \cdot 10^{5}$	$1.11 \cdot 10^{5}$	$1.79 \cdot 10^{5}$	$1.02 \cdot 10^{5}$	$1.59 \cdot 10^{5}$
q^*_{max}, kmol/m^3	$6.98 \cdot 10^{-2}$	$8.28 \cdot 10^{-2}$	$5.48 \cdot 10^{-2}$	$1.65 \cdot 10^{-2}$	$4.39 \cdot 10^{-2}$	$9.22 \cdot 10^{-2}$
n	$1.28 \cdot 10^{5}$	$1.24 \cdot 10^{5}$	$9.75 \cdot 10^{5}$	$4.53 \cdot 10^{5}$	$2.82 \cdot 10^{4}$	$3.47 \cdot 10^{5}$

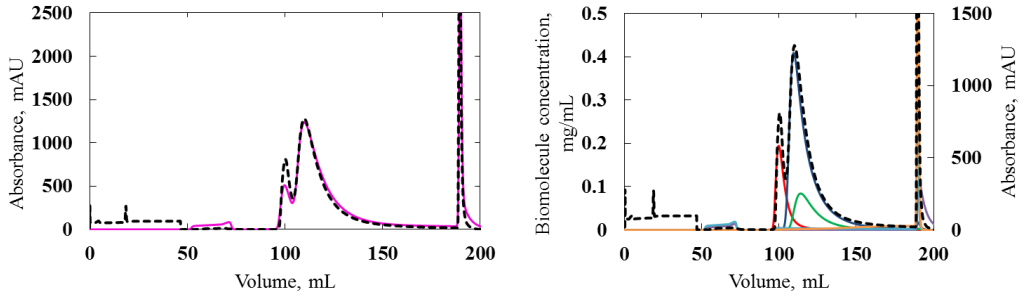


Figure 4.28: Chromatograms resulted from model parameters estimation performed with DSPX for the standard run (pH 7, 2 $mg_{monomer}/mL_{resin}$).

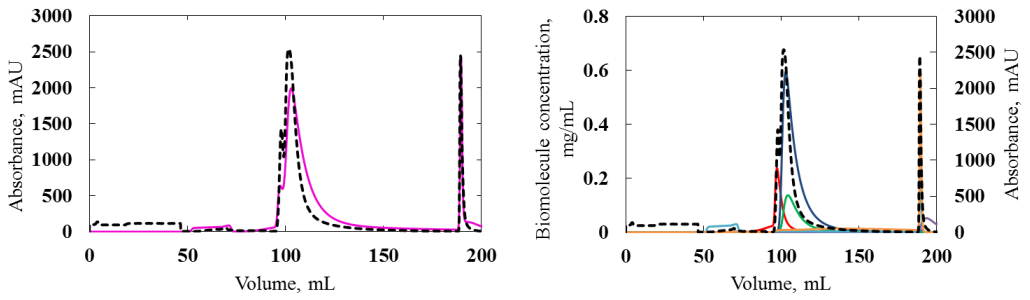


Figure 4.29: Chromatograms resulted from model parameters estimation performed with DSPX for the run with pH 6.

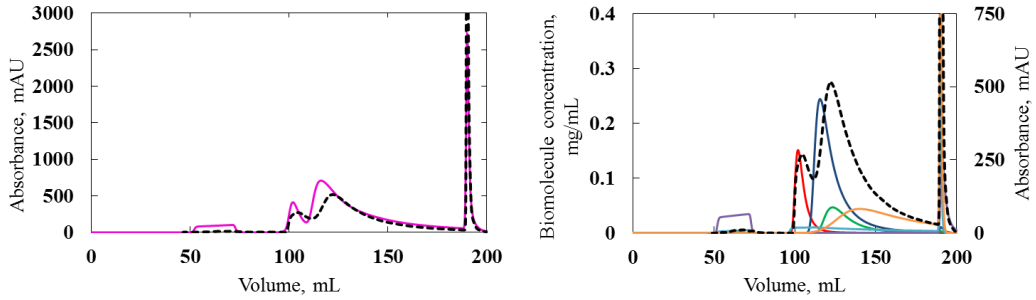


Figure 4.30: Chromatograms resulted from model parameters estimation performed with DSPX for the run with pH 8.

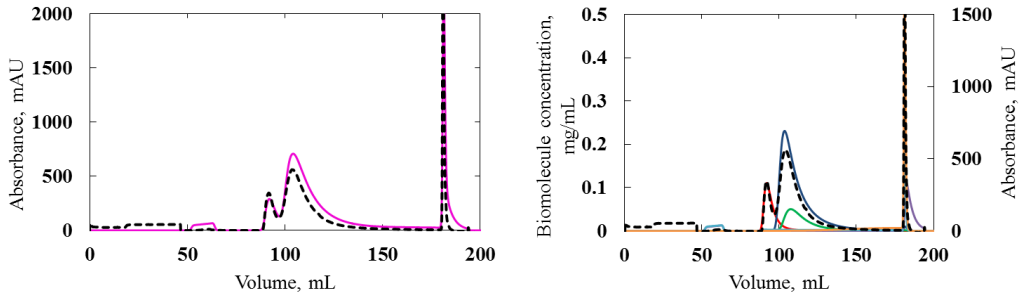


Figure 4.31: Chromatograms resulted from model parameters estimation performed with DSPX for the run with load 1 mg/mL.

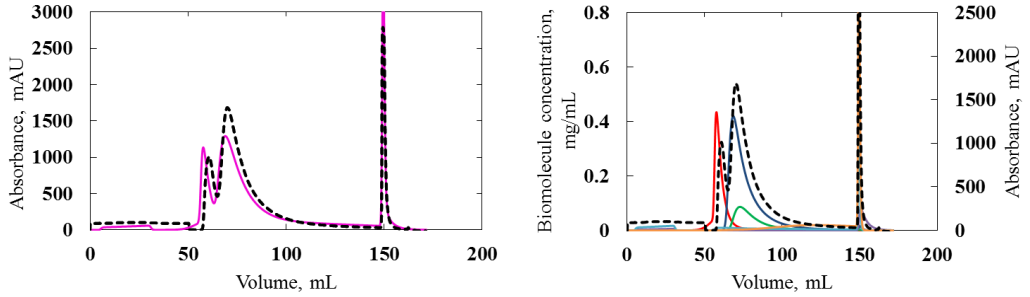


Figure 4.32: Chromatograms resulted from model parameters estimation performed with DSPX for the run with load 3 mg/mL

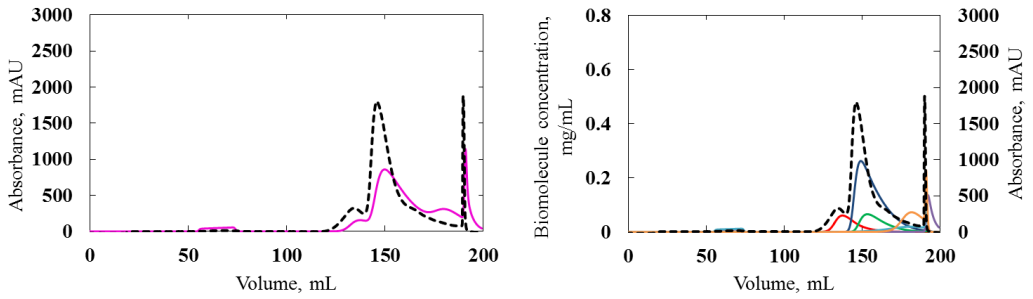


Figure 4.33: Chromatograms resulted from model parameters estimation performed with DSPX for the run with the gradient elution.

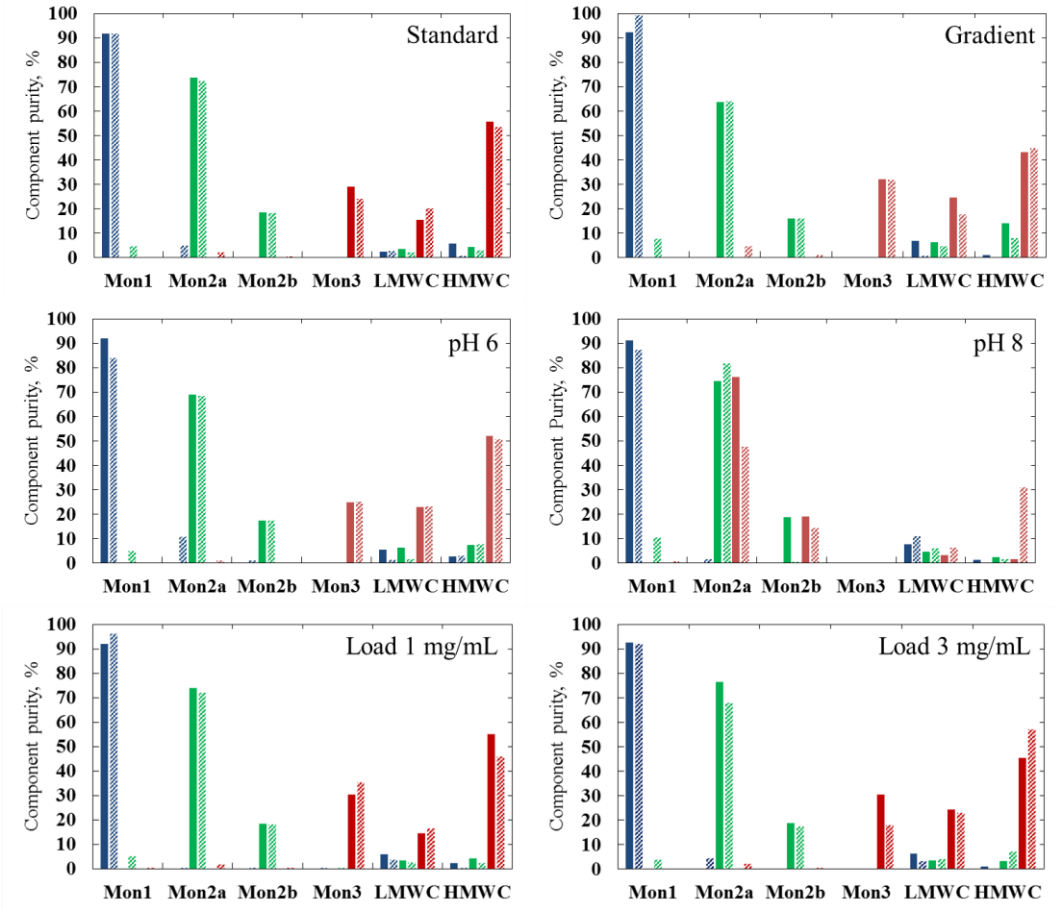


Figure 4.34: Comparison of experimental (full bars) and simulated (dashed bars) purities of the different components in the three elution peaks of the runs performed. Blue, green and red bars refer respectively to Peak 1, Peak 2 and Peak 3.

For the runs used for the estimation of parameters, experimental and simulated values of purities in the different peaks were compared to evaluate the goodness of

the prediction. Results of the comparison are shown in Figure 4.34. Purity values are reported in Appendix in Tables A4 to A9. In Figure 4.34, we can see that in most cases the monomers species are well predicted in terms of purity and eluting peak. These data give us another confirmation of what was written before, HMWC and LMWC are the species for which the assumption done were stronger and, for this reason, it is more difficult to predict their behavior faithfully.

In this case also, the model is tested in order to proof the robustness of it. A validation run is performed in different operating conditions; the monomer loading concentration and pH are the same as the standard reference run, but the isocratic elution has a lower ionic strength (18% of B pump instead of 23%).

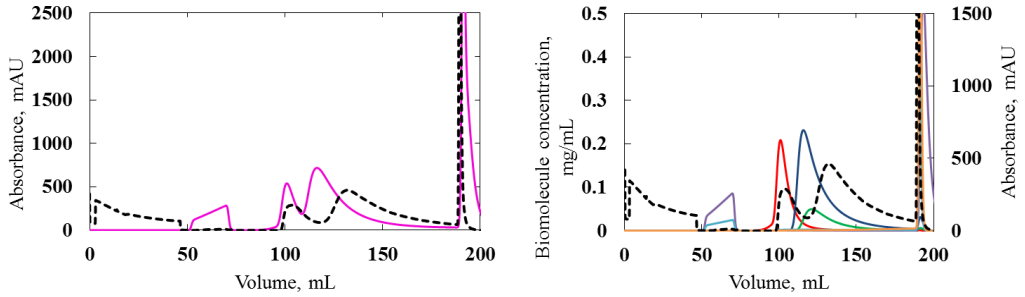


Figure 4.35: Results of the validation run performed for the industrial in-development process.

We can see that the retention volume of the peaks was worse predicted (Figure 4.35). The retention of simulated Peak 1 was quite consistent with the experimental one, the simulated Peak 2 was slightly anticipated. We can notice that the model, in this case also, considered the LMWC eluting in the flowthrough. HMWC and LMWC were the most difficult species to model. The comparison between experimental and simulated purities of the different species of the monomer is performed again and the results are shown in Figure 4.36. Values of components purities are reported in the Appendix in Table A10. It must be specified in this case that, since the second peak is slightly anticipated with respect to the experimental chromatograms, the purities values simulated are shifted because the peaks volume is considered the same of the experimental chromatograms.

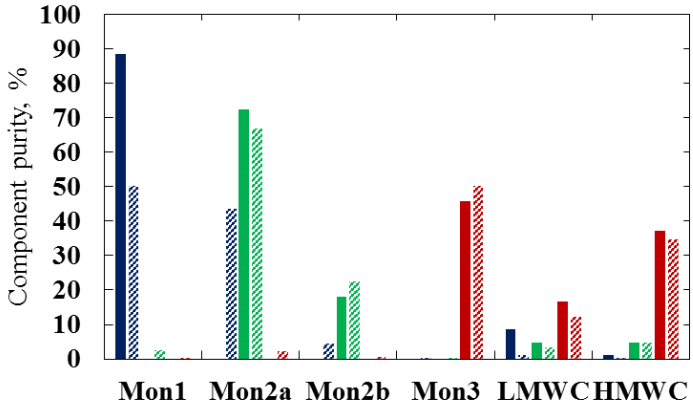


Figure 4.36: Comparison of experimental (full bars) and simulated (dashed bars) purities of the different components in the three elution peaks of the validation run. Blue, green and red bars refer respectively to Peak 1, Peak 2 and Peak 3.

In this case the experimental purities of the species in the mixture are less consistent (Figure 4.36). In particular, for Peak 1 and Peak 3 the data are quite faithful to the experimental results, while in the case of Peak 2 the values are different, especially for the Monomer 2a that is the main responsible for this peak. These inconsistencies were slightly expected since this run is not used for the parameter estimation and then the model did not use these operating conditions to estimate parameters.

At the end, the *in silico* optimization was performed. The yield and purity of all the Monomer species was maximized. The process parameters changed in order to maximize the yield and purity were the pH and percentage of B pump open, thus the ionic strength, during elution. The pH range of variation was set from 4.5 to 9.5, while the B pump percentage was varied from 15% to 35%. The pH range was chosen larger than what was considered in previous tests, but is a suitable one, with no denaturation of proteins. The ionic strength range was chosen in such a way that is not too low to lose the protein in the flowthrough, but also not too high that the proteins cannot desorb from the solid or they desorb too late. The optimization, performed *in silico* with DSPX, gave as best result a collective purity of the monomer species (all of them) of 93.67%, and a yield of 81.86%, calculated as amount of monomer species in the elution with respect to the amount of monomer species loaded in the column. The result of the optimization is shown in Figure 4.37. These values of yield and purity were reached *in silico* when a pH 5 is used with a 33% of B pump during the isocratic elution.

Since the yield and purity of all the monomers species were optimized, we can see that the optimized run has a unique peak. This solution can be valid if the four charge variants can be merged and collected all together. At the moment, since this

process is still in-development, we don’t know if this is the best solution. Anyway, constraints can be set to keep the two peaks divided.

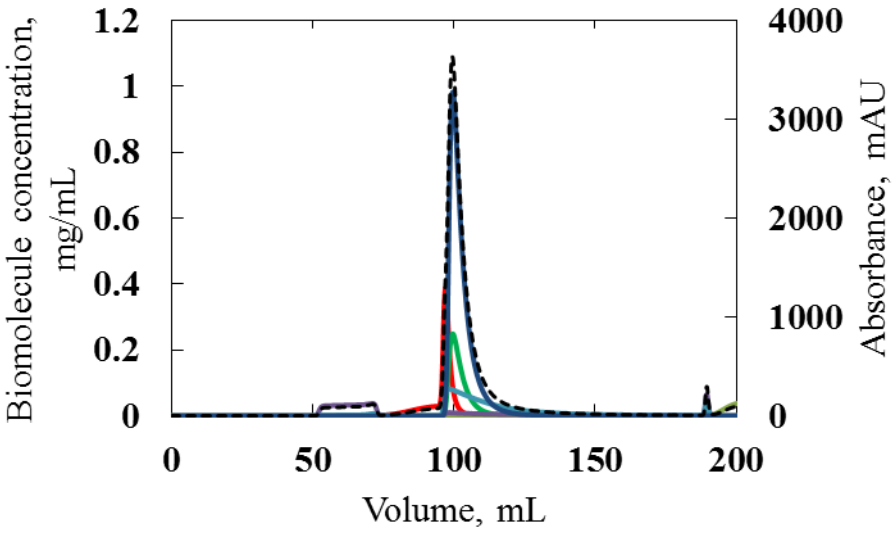


Figure 4.37: Results of the *in silico* optimization performed for the industrial in-development process.

Chapter 5

Conclusions

The chromatography process is a powerful tool exploited to purify biopharmaceutical products. Chromatography purification is a complex process affected by several parameters in which different dynamics take place to separate components in a mixture.

In this work a study is performed focusing hydrophobic interaction chromatography. Hydrophobic interaction chromatography is used when removing aggregates from a target product because it exploits the difference in hydrophobicity to separate components. While the mass transfer in hydrophobic interaction chromatography is common to other type of chromatographic purifications, the adsorption behaviour is still not clear, and different theories, with corresponding different adsorption laws, were proposed to explain the adsorption behaviour in Hydrophobic interaction chromatography.

An adsorption study that exploits experimental work and modelling is performed on some commercial protein as a proof of concepts. When determining equilibrium adsorption isotherms experimentally, a comparison is done between static and dynamic methods.

High-throughput methodology used for equilibrium adsorption isotherms determination turned out to be a very powerful tool in both single and multi-component cases. A very small amount of sample and chemical are needed and several conditions can be investigated at a time. On the other hand, since high-throughput experimentations require a high experimental effort, the experimental error that can occur during manipulation and analysis must be evaluated. Precision and accuracy of the method and equipment were evaluated in order to assess this methodology for isotherm determination. The precision of the high-throughput procedure was investigated by performing tests in low and mild adsorption. The %RSD resulted quite low and shows how high-throughput methodologies are reliable, even though the experimental procedure and analysis are performed manually.

Single-component adsorption isotherms were also determined with frontal analysis. Breakthrough experiments were performed with the commercial proteins in operating conditions comparable to those used in the static method. Frontal analysis is a more sample wasteful method. Indeed, a large amount of protein and

chemicals are required to reach the concentration plateau, even in a 1 mL column (the smaller format that can be used on a AKTA avant 25). On the other hand, the procedure is very easy and more accurate results can be obtained. Furthermore, the frontal analysis is a more powerful tool to identify anomalies. We demonstrated how the Cellulase was actually constituted by at least two different species with different molecular weight and hydrophobicity. This behaviour was highlighted by the two plateaux of the breakthrough curve, but it was not recognizable in the high-throughput isotherm determination, as a static method.

The comparison between the static and the dynamic method to determine single-component adsorption isotherms turned out in consistent isotherm data with all the protein investigated. The experimental isotherms obtained with the high-throughput methodology and the ones obtained with frontal analysis resulted to follow the same isotherm shape. The two methods have different advantages and drawback that make them suitable depending on the system investigated. It must be taken into account that, with the high-throughput method, it is not possible to recognize any anomalies.

In the case of multi-component isotherms, determined with high-throughput technologies, microfluidic capillary electrophoresis in a high-throughput platform is exploited for the analysis. Results of precision study were comparable with those made for single-component adsorption isotherms and showed that precision of the high-throughput method is reliable compared with other works. The limit in multi-component isotherm determination is represented by the analysis of the samples. Indeed, mixtures of proteins needs to be separated to quantify all the component in it. Using traditional analytical methods, like HPLC or UPLC, is not suitable because they would take too long, losing the high-throughput advantages, despite having high precision. The microfluidic capillary electrophoresis resulted a fast and easy method to quantify protein mixtures in high-throughput mode for adsorption isotherm determination. Furthermore, this method is more accurate than most of colorimetric assays (e.g., BCA or Lowry protein assays) used as routine analysis in process development of a biopharmaceutical.

A dynamic study is then performed exploiting mechanistic modelling. Two different approaches were compared to model simple systems with the commercial proteins and hydrophobic interaction chromatography columns.

The first approach is predictive, it uses fitting of experimental isotherms to obtain isotherms parameters. The isotherms parameters are then used to run simulation with an in-house Fortran code that solves the PDEs system and simulate the behaviour of a component in a column. Bind-elute tests are simulated with the code and experimental and simulated outlet concentration profiles are compared.

The second approach exploits a commercial software, DSPX from GoSilco, in which the chromatograms relatives to the bind-elute tests are imported and the operating conditions are set. DSPX uses optimization algorithms to minimize the

difference between simulated and experimental isotherms to estimate the model parameters. A validation bind-elute test is then performed in different operating conditions to validate the model

Using both approaches the comparison between experimental and simulated concentration profiles turned out to be good, the simulated peaks resulted similar to the experimental peaks, especially the retention times of the elution peaks. In the case of the predictive approach the validation run resulted successful and validated the model found. The predictive approach requires a high experimental effort to find equilibrium adsorption isotherms and fit them and is not sustainable when investigating complex systems. On the other hand, the knowledge on process dynamics gained from using this approach is huge. Indeed, with this approach it was possible to notice that the simulated peaks are more symmetric than the experimental ones. This behaviour can be due to the heterogeneity of the packing of the column. This effect is lumped in the other parameters of the model in the estimative approach. The estimative approach is more effective and requires less experimental work, just a few bind-elute tests were used. However, the model parameters are estimated from the software and no information about the adsorption dynamic are given.

The estimative approach is also applied to the industrial purification process. The process of interest is quite complex and involves different species: the monomer, its aggregates, and its truncations. Several experiments were performed manipulating the parameters that affect more the efficiency of the process (pH, load concentration, elution method). The experimental tests were used to estimate model parameters with DSPX. Assumptions on the composition of the load were done to simplify such a complex system. In this case also, a validation run was performed to confirm the model. The approach and assumption resulted effective for the industrial process and the validation run confirmed the reliability of the model. In particular in all the runs the retention time of the peaks and the elution of each component was consistent with the assumptions done.

Instead of experiment complete DoEs (Design of Experiments) designed to investigate and optimize a purification process, the modelling can reduce the material requirement and speed up the development of the process. Modelling allows to perform in silico experiment and predict how the system reacts to different operating conditions without wasting time and material. A modelling approach increases the knowledge on the dynamics of the process and speeds up the process development and optimization, with low experimental effort. It is a powerful tool when facing new and complex systems that requires short times to be established.

List of symbols

A	absorbance, AU
a	isotherm parameter, -
b	isotherm parameter, mL/mg
c_0	load protein concentration, mg/mL
c_{av}	average concentration value, mg/mL
c_f	protein concentration in the outlet when the plateau is reached, mg/mL
c_{feed}	feed concentration, mg/mL
c_{HMWC}	HMWC concentration, mg/mL
c_{in}	protein concentration in the feed in the initial state, mg/mL
c_{int}	protein concentration in the interstitial liquid, mg/mL
c_{LMWC}	LMWC concentration, mg/mL
c_{mon}	monomer concentration, mg/mL
c_p	protein concentration in the pore liquid, mg/mL
c_{salt}	salt concentration, mol/L
c_{sample}	sample protein concentration, mg/mL
c_{th}	theoretical concentration value, mg/mL
c_{tot}	total protein concentration, mg/mL
$c_{unbound}$	unbound protein concentration, mg/mL
CF	compression factor
D_{ax}	axial dispersion, m^2/s
D_m	molecular diffusion, m^2/s
D_p	pore diffusion, m^2/s
d_p	particle diameter, m
k	adsorption parameter of linear and exponential Langmuir law, $m^3/kmol$
k_c	film mass transfer coefficient, m/s
k_{eff}	effective film mass transfer coefficient, m/s
k_{eq}	equilibrium constant, -
$k_{eq,batch}$	equilibrium constant for batch experiments, -

$k_{eq,0}$	equilibrium constant of pH modified Wang model, -
$k_{eq,1}$	equilibrium constant of pH modified Wang model, -
$k_{eq,2}$	equilibrium constant of pH modified Wang model, -
k_{kin}	kinetic constant, s
K_d	partition coefficient, -
K_p	adsorption equilibrium constant of Chen and Sun model (2003), -
K_s	protein dehydration equilibrium constant of Chen and Sun model (2003), -
k_p	protein coefficient of Mollerup model, M^{-1}
k_s	salt coefficient of Mollerup model, M^{-1}
L	column length, m
l	optical path length, cm
M	molarity, -
M_A	protein molecular mass, g/mol
n	stoichiometric number of ligands, -
P_{HMW}	HMWC purity, %
P_{LMW}	LMWC purity, %
P_{mon}	monomer purity, %
q	adsorbed protein concentration, $mg/mL_{apparent\ solid\ volume}$
q_{max}	maximum concentration, $mg/mL_{apparent\ solid\ volume}$
q'	adsorbed protein concentration, $mg/mL_{particles}$
q^*	adsorbed protein concentration, $mg/mL_{solid\ skeleton}$
q^*_{max}	maximum concentration, $mg/mL_{solid\ skeleton}$
$q^*_{max,batch}$	maximum concentration in batch conditions, $mg/mL_{solid\ skeleton}$
R_p	particle radius, m
r	radial particle coordinate
T	temperature, K
t	time
t_{des}	end time of breakthrough curve
$\tilde{V}$	volume flowrate, m^3/sec
V_c	column volume, mL
V_d	dead volume, mL
V_f	total liquid volume, mL
V_{int}	interstitial volume, mL

V_{liquid}	liquid volume in the wells, mL
V_{medium}	apparent solid volume, mL
V_r	retained volume, mL
V_{RetAc}	retention time of Acetone injection peak, mL
V_{RetDev}	retention time of Dextran injection peak, mL
V_{sample}	sample liquid volume, mL
v_{int}	interstitial velocity, m/s
x	salt mass fraction, -
z	axial column coordinate

Greek letters

α	salt concentration exponent, -
β	salt dependency parameter, -
β_0	isotherm parameter, -
β_1	isotherm parameter, -
ε_c	column porosity, -
ε_p	particle porosity, -
ε_t	total porosity, -
ε_λ	molar extinction coefficient, AU*cm^{-1}*(mg/mL)$^{-1}$
η_B	salt solution viscosity, cP
η_W	water viscosity, cP
Λ	hydrophobic ligand density, mol/mL
λ	isotherm parameter related to ligand density, -
σ	steric factor, -
σ_{Dex}	standard deviation of the Dextran injection peak, mL

References

Arakawa, T. and Timasheff, S. N. (1984), Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding, *Biochemistry*, 23(25), 5912–5923. doi: 10.1021/bi00320a004

Baumann, P. and Hubbuch, J. (2017), Downstream process development strategies for effective bioprocesses: Trends, progress, and combinatorial approaches, *Engineering in Life Sciences*, 17, 1142–1158. doi: 10.1002/elsc.201600033.

Bröker, M., Costantino, P., DeTora, L., McIntosh, D. E. and Rappu-oli, R. (2011) Biochemical and biological characteristics of cross-reacting material 197 (CRM197), a non-toxic mutant of diphtheria toxin: Use as a conjugation protein in vaccines and other potential clinical applications, *Biologicals*, 39(4), 195–204. doi: 10.1016/j.biologicals.2011.05.004.

Chen, J. and Cramer, S. M. (2007), Protein adsorption isotherm behavior in hydrophobic interaction chromatography, *Journal of Chromatography A*, 1165(1–2), 67–77. doi: 10.1016/j.chroma.2007.07.038.

Chen, J. and Sun, Y. (2003), Modeling of the salt effects on hydrophobic adsorption equilibrium of protein, *Journal of Chromatography A*, 992(1–2), 29–40. doi: 10.1016/S0021-9673(03)00277-2.

Close, E. J., Salm, J. R., Bracewell, D. G. and Sorensen, E. (2014), Modelling of industrial biopharmaceutical multicomponent chromatography, *Chemical Engineering Research and Design*, 92(7), 1304–1314. doi: 10.1016/j.cherd.2013.10.022.

Dahlmann, J., Budakowski, W. R. and Luckas, B. (2003), Liquid chromatography-electrospray ionisation-mass spectrometry based method for the simultaneous determination of algal and cyanobacterial toxins in phytoplankton from marine waters and lakes followed by tentative structural elucidation of microcystins, *Journal of Chromatography A*, 994(1–2), 45–57. doi: 10.1016/S0021-9673(03)00485-0.

Danckwerts, P. V. (1953), Continuous flow systems, *Chemical Engineering Science*, 2(1), 1–13. doi: 10.1016/0009-2509(53)80001-1.

Deitcher, R.W., Rome, J.E., Gildea, P.A., O'Connel, J.P. and Fernandez, E.J. (2010), A new thermodynamic model describes the effects of ligand density and type, salt concentration and protein species in hydrophobic interaction

chromatography, *Journal of Chromatography A*, 1217(2), 199–208. doi: 10.1016/j.chroma.2009.07.068.

Emerson and Rosemount Analytical 2010, Conductance data for commonly used chemicals. 44-6039/rev. B.

Fausnaugh, J. L. and Regnier, F. E. (1986), Solute and mobile phase contributions to retention in hydrophobic interaction chromatography of proteins, *Journal of Chromatography A*, 359, 131–146. doi: 10.1016/0021-9673(86)80068-1.

Field, N., Konstantinidis, S. and Velayudhan, A. (2017), High-throughput investigation of single and binary protein adsorption isotherms in anion exchange chromatography employing multivariate analysis, *Journal of Chromatography A*, 1510, 13–24. doi: 10.1016/j.chroma.2017.06.012.

GE Healthcare Adsorption equilibrium isotherm studies using a high-throughput method, Application note 28-9403-62 AA, 2009. Available online: https://cdn.cytivalifesciences.com/dmm3bwsv3/AssetStream.aspx?mediaformatid=10061&destinationid=10016&assetid=14249 (accessed on 19 January 2021).

Guiochon, G., Felinger, A., Shirazi, D. G. and Katti, A.M., *Fundamentals of Preparative and Nonlinear Chromatography (Second edition)*, Elsevier (San Diego, CA, USA) (2006), 975 pp. doi: 10.1016/s0009-5893(07)82075-3.

Hahn, T., Baumann, P., Huuk, T., Heuveline, V. and Huubuch, J. (2016), UV absorption-based inverse modeling of protein chromatography, *Engineering in Life Sciences*, 16, 99–106. doi: 10.1002/elsc.201400247.

Lienqueo, M. E., Mahn, A., Salgado, C. J. and Asenjo, J. A. (2007), Current insights on protein behavior in hydrophobic interaction chromatography, *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences*, 849(1–2), 53–68. doi: 10.1016/j.jchromb.2006.11.019.

Lietta, E., Pieri, A., Innocenti, E., Pisano, R., Vanni, M. and Barresi, A. A. (2021), Use of microfluidic capillary electrophoresis for the determination of multi-component protein adsorption isotherms: application to high-throughput analysis for hydrophobic interaction chromatography, Pharmaceutics, 13(12), 2135. doi: 10.3390/pharmaceutics13122135.

Mollerup, J. M. (2006), Applied thermodynamics: A new frontier for biotechnology, *Fluid Phase Equilibria*, 241(1–2), 205–215. doi: 10.1016/j.fluid.2005.12.037.

Mollerup, J. M., Budde Hansen, T., Kidal, S. and Staby, A. (2008), Quality by design-Thermodynamic modelling of chromatographic separation of proteins,

Journal of Chromatography A, 1177(2), 200–206. doi: 10.1016/j.chroma.2007.08.059.

Moon, Y. U., Curtis, R. A., Anderson, C. O., Blanch, H. W. and Prausnitz, J. M. (2000), Protein-protein interactions in aqueous ammonium sulfate solutions. Lysozyme and bovine serum albumin (BSA), *Journal of Solution Chemistry*, 29, 699–718. doi: 10.1023/a:1005112927213.

Osberghaus, A., Baumann, P., Hepbildikler, S., Nath, S., Haindl, M., von Lieres, E. and Hubbuch, J. (2012), Detection, quantification, and propagation of uncertainty in high-throughput experimentation by Monte Carlo methods, *Chemical Engineering and Technology*, 35, 1456–1464. doi: 10.1002/ceat.201100610.

Perkin Elmer. Protein Express Assay Quick Guide LabChip® GXII Touch and LabChip® GX II, PN CLS140169 Rev. F, 2019. Available online: https://resources.perkinelmer.com/lab-solutions/resources/docs/GDE-protein-express-quick-guide.pdf (accessed on 19 January 2021). doi: 10.1016/S0168-1656(01)00237–1.

Queiroz, J.A., Tomaz, C.T., and Cabral, J.M.S. (2001), Hydrophobic interaction chromatography of proteins, *Journal of Biotechnology*, 87(2), 143–159. doi: 10.1016/s0168-1656(01)00237-1.

Rodler, A., Ueberbacher, R., Beyer, B. and Jungbauer, A. (2019), Calorimetry for studying the adsorption of proteins in hydrophobic interaction chromatography, *Preparative Biochemistry & Biotechnology*, 49(1), 1-20. doi: 10.1080/10826068.2018.1487852.

Schmidt-Traub, H., Schulte, M. and Seidel-Morgenstern, A. (2005), *Preparative Chromatography: of Fine Chemicals and Pharmaceutical Agents*, Wiley-VCH Verlag GmbH & Co. KGaA (Weinheim, Germany), 458 pp.

Sinanoglu, O. and Abdulnur, S. (1965), Effect of water and other solvents on the structure of biopolymers, *Federation Proceedings*, 24, 12–23. PMID: 14314560.

Smith, M. T., Zhang, S., Adams, T., DiPaolo, B. and Dally, J. (2017), Establishment and validation of a microfluidic capillary gel electrophoresis platform method for purity analysis of therapeutic monoclonal antibodies, *Electrophoresis*, 38(9–10), 1353–1365. doi: 10.1002/elps.201600519.

Tukey, J. W. (1949), Comparing individual means in the analysis of variance, *Biometrics*, 5(2), 99–114. doi: 10.2307/3001913.

Wang, G., Hahn, T. and Hubbuch, J. (2016), Water on hydrophobic surfaces: mechanistic modeling of hydrophobic interaction chromatography, *Journal of Chromatography A*, 1465, 71–78. doi: 10.1016/j.chroma.2016.07.085.

Appendix

Table A1: Langmuir parameters *a* and *b* for HSA, Lysozyme, Cellulase and Catalase.

	HSA		Lysozyme		Cellulase		Catalase	
	a, -	**b, mg/mL**	**a, -**	**b, mg/mL**	**a, -**	**b, mg/mL**	**a, -**	**b, mg/mL**
0.5 M AS	2.20	0.24	0.26	-0.07	3.15	0.15	2.88	-0.13
0.8 M AS	7.58	0.71	3.84	1.34	10.05	0.75	9.08	0.19
1 M AS	12.84	1.40	1.46	0.12	61.69	5.10	29.40	0.62
1.2 M AS	70.60	4.15	3.33	0.15	47.70	1.35	177.25	1.87
1.4 M AS	247.47	10.42	7.88	0.22	503.62	12.74	198.74	0.98

Table A2: Langmuir parameters *a* and *b* for HSA with Butyl Sepharose HP, Phenyl Sepharose High Sub and Phenyl Sepharose Low Sub resins.

	Butyl Sepharose HP		Phenyl Sepharose High Sub		Phenyl Sepharose Low Sub	
	a, -	**b, mg/mL**	**a, -**	**b, mg/mL**	**a, -**	**b, mg/mL**
0.5 M AS	2.20	0.24	33.99	2.74	2.25	0.16
0.8 M AS	7.58	0.71	68.68	3.43	4.26	0.25
1 M AS	12.84	1.40	58.02	2.38	8.54	0.70
1.2 M AS	70.60	4.15	110.80	3.62	36.88	1.62
1.4 M AS	247.47	10.42	358.48	8.95	95.28	2.79

Table A3: adsorption isotherms parameters of the proteins when investigated in the resin of 51 μmol/mL (RES51) and 55 μmol/mL (RES55).

Parameter	**Albumin**		**BSA**		**α-chymotrypsin**		**β-lactoglobulin**	
	RES51	RES55	RES51	RES55	RES51	RES55	RES51	RES55

λ, -	31.9	35.1	26.8	30.4	38.4	37.2	19.1	18.4
b, $m^3/kmol$	4.45	26.7	99.6	92.96	87.5	62.5	5.52	1.84
k, M^{-1}	7.41	6.11	7.03	6.98	4.5	4.85	6.54	8.02

Table A4: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the standard reference run.

	Standard Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	91.90	91.87	0.00	4.53	0.00	0.03
Mon2a	0.00	4.76	73.68	72.29	0.00	2.01
Mon2b	0.00	0.02	18.42	18.22	0.00	0.36
Mon3	0.00	0.01	0.00	0.01	29.00	24.09
LMWC	2.50	2.72	3.40	2.11	15.40	20.12
HMWC	5.60	0.62	4.40	2.84	55.60	53.40

Table A5: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the run with the lower monomer load concentration of 1 $mg_{monomer}/mL_{resin}$.

	Load1 Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	92.00	96.10	0.00	5.00	0.00	0.02
Mon2a	0.00	0.06	73.84	72.12	0.00	1.72
Mon2b	0.00	0.00	18.46	18.09	0.00	0.29
Mon3	0.00	0.03	0.00	0.02	30.30	35.43
LMWC	5.80	3.70	3.50	2.52	14.50	16.51
HMWC	2.20	0.11	4.30	2.25	55.10	46.04

Table A6: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the run with the lower monomer load concentration of 3 $mg_{monomer}/mL_{resin}$.

	Load3 Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	92.50	92.18	0.00	3.72	0.00	0.02
Mon2a	0.00	4.31	76.64	67.87	0.00	1.98
Mon2b	0.00	0.03	18.66	17.28	0.00	0.35
Mon3	0.00	0.00	0.00	0.00	30.40	17.84
LMWC	6.20	3.21	3.50	3.95	24.40	22.84
HMWC	1.00	0.27	3.20	7.18	45.30	56.96

Table A7: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the run with the gradient elution.

	Gradient Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	92.40	99.14	0.00	7.55	0.00	0.02
Mon2a	0.00	0.07	63.76	63.93	0.00	4.63
Mon2b	0.00	0.00	15.94	15.97	0.00	1.13
Mon3	0.00	0.00	0.00	0.06	32.00	31.83
LMWC	6.80	0.75	6.20	4.63	24.70	17.53
HMWC	0.90	0.04	14.10	7.87	43.30	44.85

Table A8: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the run with the lower pH 6.

	pH6 Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	92.10	84.03	0.00	4.94	0.00	0.02
Mon2a	0.00	10.70	68.96	68.41	0.00	1.02
Mon2b	0.00	1.06	17.24	17.33	0.00	0.10
Mon3	0.00	0.00	0.00	0.00	24.90	25.08
LMWC	5.30	1.27	6.40	1.61	22.90	23.07
HMWC	2.60	2.94	7.50	7.72	52.10	50.72

Table A9: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the run with the higher pH 8.

	pH8 Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	91.10	87.25	0.00	10.35	0.00	0.83
Mon2a	0.00	1.53	74.48	81.71	76.32	47.73
Mon2b	0.00	0.00	18.62	0.41	19.08	14.24
Mon3	0.00	0.08	0.00	0.05	0.00	0.07
LMWC	7.70	10.87	4.60	6.01	3.10	6.18
HMWC	1.20	0.27	2.30	1.48	1.50	30.95

Table A10: comparison between experimental and simulated purities of the monomers, HMWC and LMWC for the validation run.

	Validation Run					
	Peak 1		Peak 2		Peak 3	
Component	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %	Experimental purity, %	Simulated purity, %
Mon1	88.50	50.23	0.00	2.45	0.00	0.05
Mon2a	0.00	43.65	72.48	67.02	0.00	2.11
Mon2b	0.00	4.54	18.12	22.42	0.00	0.43
Mon3	0.00	0.08	0.00	0.26	45.80	50.29
LMWC	8.70	1.20	4.80	3.23	16.80	12.31
HMWC	1.00	0.31	4.60	4.63	37.30	34.80