Understanding adsorption behavior of α-chymotrypsin onto cation exchanger using all-atom molecular dynamics simulations

Marine Tournois, Stéphane Mathé, Isabelle André, Jeremy Esque, Maria Fernandez

Understanding adsorption behavior of α-chymotrypsin onto cation exchanger using all-atom molecular dynamics simulations

Marine Tournois, Stéphane Mathé, Isabelle André, Jérémy Esque*, María A. Fernández*

Toulouse Biotechnology Institute, TBI (ex LISBP), Université de Toulouse, CNRS, INRA, INSA, Toulouse, France

ARTICLE INFO

Article history:
Received 25 June 2019
Revised 13 November 2019
Accepted 16 November 2019
Available online 18 November 2019

Keywords:
Ion exchange chromatography
Molecular dynamics
Protein sorption
α-Chymotrypsin
Steric mass action model

ABSTRACT

The interest for a better understanding of ion-exchange mechanisms at the atomic level has strongly increased over the past decades. Indeed, molecular-level information about physico-chemical mechanisms could help optimizing chromatographic processes for protein purification, which are sub-optimized due to the lack of predictive models. A promising approach is based on the use of Molecular Dynamics (MD) simulations to study local phenomena inside adsorbents which can then be challenged against experimental results. In this work, macroscopic experimental data, consisting in the ion-exchange uptake of α-chymotrypsin onto SP Sepharose FF, have been compared to the adsorption behavior predicted by MD simulations. The chromatographic surface, represented as a uniform distribution of ligands with a counterion layer, in the presence of the protein was modeled using all-atom representation. The SMA formalism was used to describe single adsorption isotherms and to relate macroscopic observations with molecular simulations. Two SMA parameters based on physical principles, the characteristic charge n and the steric factor σ, have been estimated by both experiments and MD simulations. At pH 5 and NaCl concentration of 100 mM, our study shows a fairly good agreement between both results, especially for the characteristic charge. It is shown that the steric factor calculation is strongly dependent on the ligand density on the adsorbent surface, whose value must be carefully determined in order to obtain reliable predictions. In addition, four binding patches were identified as being involved in the adsorption, which have been confirmed through binding free energy calculations.

1. Introduction

Chromatographic processes, especially Ion Exchange Chromatography (IEC), are extensively used in protein purification. However, their industrial scale-up remains sub-optimized and is still based on empirical methods due to difficulties in developing predictive models. This might be improved, for instance, by obtaining microscopic information on physico-chemical mechanisms. Today, the process cost represents up to 80–90% of the global cost for protein production [1]. To improve cost efficiency, one major scientific challenge remains the prediction of multi-component system behavior because specific interactions – such as protein-protein interactions, protein-support competition and conformational changes – have to be taken into account. Knowledge at the atomic level inside a chromatographic support is still limited due to the complexity in adapting physical techniques. In this context, a possible way to gain knowledge of atomistic-level mechanisms is to use Molecular Dynamics (MD) simulations to study protein-surface interactions [2,3].

The protein adsorption on chromatographic media has been thoroughly investigated over the past decades, in terms of surface properties [4–7], influence of pH and/or ionic strength [8] as well as protein characterization in single and multicomponent systems [9]. Several models were developed to describe the adsorption equilibrium of macro-molecules on a porous media [10–14]. While the Langmuir isotherm, which is the most used, describes only the adsorption equilibrium phenomena, the Steric Mass Action (SMA) model [10] accounts for the counterion displacement and hindering during the ion exchange, with regard to the large size of the molecule [12]. Some other models describe both mass transport and adsorption of macro-molecules in a porous particle like the distributed pore model [15]. The multipoint nature of protein-support interactions induced by the charge distribution around the molecule surface was also studied [16]. Nevertheless, a few models integrate the electrostatic changes associated to pH. Among them, Gulat et al. [17] focused on this aspect, but molecular information about adsorption phenomena is still lacking. Microscopic experimental techniques, such as Scanning Tunneling Mi-

* Corresponding authors.
E-mail addresses: esque@insa-toulouse.fr (J. Esque), mafernan@insa-toulouse.fr (M.A. Fernández).

croscopy (STM) [18] or Confocal Laser Scanning Microscopy (CLSM) [19], were employed to better understand the protein uptake on adsorbent but are still limited notably due to light attenuation in deepest layers. Atomic Force Microscopy (AFM) could provide protein binding information [20,21] but intraparticle measurements remain difficult.

As a first approach for gaining microscopic information, Roth and Lenhoff [22] have already studied in 1993 the adsorption of lysozyme on a charged surface, by assuming the protein to be a sphere. Doing so, they showed the influence of ionic strength on protein-surface affinity. Noinville et al. [23] also worked on the molecular adsorption of α-lactalbumine and lysozyme in order to find the electrostatic patches and calculate the global charge. More recently, MD simulations were successfully used to look at the adsorption behavior of a fibrinogen γ chain fragment on different monolayer surfaces and its possible conformational change [24].

For these reasons, molecular simulation appears to be a powerful method to study protein behavior *in silico* inside an adsorbent pore, without requiring intrusive and expensive experimental techniques. Although simulations may provide information at the atomic scale, three main issues were identified by Latour et al. [2] and have to be considered: (i) using a valid and properly parametrized force-field [25]; (ii) using an accurate representation of the solvation effects and (iii) conformational sampling, which means that a sufficient number of simulations is needed to represent an average property (for example, the average binding free energy).

The protein adsorption in the particular case of ion exchange chromatography have been studied through MD simulations using both all-atom and Coarse-Grained (CG) representations. All-atom MD simulations in implicit solvent (Generalized Born) were –amongst others– used to predict the conductivity at elution time from average electrostatic energies of interaction [26]. In this study, 62 short simulations (100 ps) from different starting protein orientations were performed with the lysozyme near a SP Sepharose FF surface. The binding free energy between protein and ligands was calculated for each orientation leading to an energy map: the favorable binding patches were highlighted and appeared to be in good agreement with experiments. Although the influence of key parameters such as pH or ligands spacing was shown, this approach is limited by the small size of the system and its application to larger proteins remains difficult. CG models enable simplification of the system by reducing the number of tracked particles and thus allow for longer MD simulations to be run. Liang et al. [27,28] performed studies on the ion exchange of BSA/HSA (Bovine/Human Serum Albumine) and bHb (bovine Hemoglobin) on Q Sepharose FF combining both experimental work and MD simulations [29]. They ran long simulations (400 ns) that led to an adsorption of proteins and an estimation of parameters from the SMA model. Their experiments showed a good agreement between both approaches revealing the high potential of molecular simulations to predict macroscopic observations. Of course, computational cost of CG models is considerably reduced compared to all-atom ones, though they suffer from lower accuracy. Furthermore, in most studies, the effect of a counterion layer on the chromatographic surface is not treated.

The aim of this work is to better understand the binding mechanisms of proteins on adsorbents using MD simulations whose predictions are compared with macroscopic experiments. For this purpose, the SMA model – which accounts for the steric hindrance of the protein, its characteristic charge and the adsorption equilibrium constant – is used to describe single adsorption isotherms. The SMA parameters are based on physical principles and can be deduced from both experimental approaches and MD simulations. A well-known protein, the α-chymotrypsin (αCT), and a widely used resin, SP Sepharose FF, were chosen as model system. At first, macroscopic experiments were conducted in order to obtain adsorption isotherms with breakthrough curves measurements. Method of least squares was then used to estimate the SMA parameters. Then, all-atom MD simulations were performed on a system containing the protein, water molecules, SP Sepharose groups and a counterion layer mimicking an equilibrated resin. In this way, the counterion role in the ion-exchange mechanism can be further studied. At given pH and salt concentration, six different starting orientations of the protein were studied and for each orientation, three independent MD simulations were performed. The adsorbed states from the MD trajectories were further analyzed to determine the SMA parameters.

2. Theory

The Steric Mass Action (SMA) model is the most widely used to describe single adsorption isotherms [10]. Unlike most models, the SMA model accounts for the steric hindrance induced by the large size of the proteins. Its main assumption is to consider the ion-exchange mechanism as a stoichiometric reaction between the protein and the complex resin/counterion. Then, the SMA model (Eq. (1)) is given as:

$$C = \left(\frac{q}{K}\right)\left(\frac{C_I}{q_R - (n+\sigma)q}\right)^n \tag{1}$$

where q and C are the protein concentrations in the solid and liquid phase respectively, C_I the counterion concentration in solution, q_R the total ionic capacity and K the equilibrium constant. n represents the characteristic charge and is defined as the number of binding ligands. σ is the steric factor and represents the number of ligands sterically hindered by the protein. In this study, the counterions for the ion-exchange phase are sodium ions and their concentration is considered as constant over time for each experiment. Indeed, the experiences are performed on columns with an open system operation, which means that the solution in the column is renewed continuously: the ion concentration in the column is equal to its initial concentration at the equilibrium state. At given pH, the SMA parameters are then estimated by curve fitting using at least three isotherms with three different salt concentrations. A reliability factor R (Eq. (2)) is introduced and minimized by only adjusting the parameters n, σ and K:

$$R = \sqrt{\frac{\sum (C_{i,exp} - C_{i,cal})^2}{\sum {C_{i,exp}}^2}} \tag{2}$$

where C_{exp} and C_{cal} are the experimental and calculated concentrations in the liquid phase, respectively. Please notice that this is not a regression coefficient and that a good fitting between experiments and the model leads to a reliability factor close to zero.

3. Materials and methods

3.1. Adsorption isotherm experiments

Adsorption equilibrium of α-chymotrypsin on SP Sepharose FF was studied through isotherm experiments at pH 5, using breakthrough curve measurements. Four different isotherms, with salt concentrations ranging from 0 to 100 mmol/L, were measured in order to estimate the SMA parameters.

3.1.1. Materials

α-chymotrypsin (C-4129, purity $\geq$ 85%) was purchased from Sigma (St. Louis, MO, USA) and used without prior purification. The salts used for buffer solutions (citric acid, trisodium citrate dihydrate and sodium chloride) were also obtained from Sigma and were of analytical grade. The adsorbent material, SP Sepharose FF,

was obtained from GE Healthcare (Uppsala, Sweden) in prepacked column HiTrap SP FF, 1 mL. The total ionic capacity q_R of the SP Sepharose FF, expressed per bed volume, was measured in triplicate by acid/base titration and estimated at 230.2 ± 3.3 mmol/L (between 180 and 240 mmol/L according to the manufacturer).

3.1.2. Determination of protein concentration

UV–vis spectroscopy

All proteins absorb ultraviolet light at 280 nm [30]. The absorbance mainly depends on the content of three amino acids: tryptophan (Trp), tyrosine (Tyr) and cysteine (Cys, disulfide bonds). Therefore, a protein concentration can be determined by direct measurement of the absorbance at 280 nm. This method is non-selective and can be used for analyzing single-protein systems. First, a calibration curve was measured by preparing protein solutions of known concentrations: a standard solution of αCT (1 mg/mL) was prepared and then diluted in different ratios (from 0.1 to 0.8 mg/L). The standard solution was prepared by dissolving αCT in the working buffer (citrate buffer pH=5, 50 mM) and then filtered through a 0.45 µm cellulose membrane. Finally, all samples were analyzed using a JASCO V630 UV–vis spectrophotometer. Each absorbance was measured in triplicate and averaged.

Size-exclusion chromatography

A common method to quantify the protein concentration is Size Exclusion Chromatography (SEC, or HPLC-SEC) in which proteins are separated according to their molecular weight (MW). This method was used to investigate the αCT dimerization. As for the UV–vis method, a calibration curve was performed. All the samples prepared for the UV–vis calibration were loaded onto a Superdex 75 Increase 10/300 GL column (GE Healthcare) and connected to an Äkta Purifier system (GE Healthcare). The column was previously equilibrated with at least 3 column volumes (CV) of working buffer. 100 µL of sample was injected and the buffer flow rate was set to 0.8 mL/min. Then, the chromatogram was analyzed: the peak area corresponding to the αCT (elution volume = 12.7 mL) was calculated and plotted against the protein concentration. Note that a prior calibration of the HPLC-SEC column was performed by preparing a multi-protein solution (6 standards proteins with known MWs), in order to correlate the molecular weight to the elution volume.

3.1.3. Breakthrough curves

Breakthrough curves were performed on the Äkta Purifier system using a 1 mL column HiTrap SP FF (GE Healthcare). Before use, the column was washed with 5 CVs of pure water and 5 CVs of equilibration buffer. Then, a protein solution was loaded continuously onto the column at 1 mL/min and the outlet protein concentration was monitored by measuring the UV absorbance. The run was conducted using the control software Äkta UNICORN (version 5.11) until the outlet concentration reached the initial concentration. The amount of q protein adsorbed in the resin, in equilibrium with a protein solution (C), can be calculated by analyzing the breakthrough curves. Each experiment leads to one point (C; q) on the isotherm: six different breakthrough curves were performed to obtain one isotherm. Finally, the column was regenerated using an elution buffer (working buffer + 1 M of sodium chloride) and equilibrated for the next experiment.

3.1.4. Enzyme activities

The enzyme activity assays of αCT were performed by following the substrate hydrolysis reaction during 5 min at room temperature (20°C). According to the protocol from Verma and Ghosh [31], 20 µL of enzyme solution were mixed with 290 µL of para-nitrophenyl acetate (pNPA, 7.3 mg in pure ethanol) and 2.69 mL of Tris–HCl buffer (100 mM, pH 7.75). The release of para-nitrophenol (pNP) was monitored by measuring the absorbance at 405 nm. For these conditions, the extinction coefficient of pNP was 14200 M.cm^{-1}. One unit (U) of enzyme activity was defined by the amount of pNP (µmol) released per minute, at pH 7.75 and T = 20°C. The concentration of αCT solutions was measured before each enzyme activity assay.

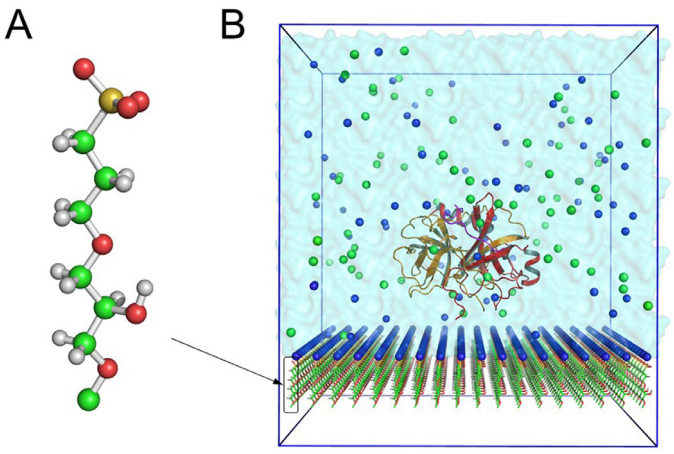

Fig. 1. A: Representation "ball-and-stick" of the SP Sepharose FF ligand. B: Simulation box representing the chromatographic surface with a counterion layer (bottom), the α-chymotrypsin (cartoon representation), chloride (green spheres) and sodium (blue spheres) ions and water. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. In silico experiments

3.2.1. System setup

Protein

The atomic coordinates of αCT were obtained from the X-ray structure (PDB-ID: 1YPH) at a resolution of 1.34 Å. Only one of two protein subunits (chains B,D and F) was extracted. Indeed, in the monomeric form, αCT (3D dimensions: 4.5 nm × 3.8 nm × 3.5 nm [32]) is composed of three protein chains, e.g. 12 amino acids (chain B), 131 amino acids (chain D) and 97 amino acids (chain F). These three fragments initially come from the same main protein chain, which was cleaved, and they are linked through disulfide bridges (two inter-chain and three intra-chain). The disulfide bridges were built using *gmx pdb2gmx* tool from GROMACS 5.0 software [33]. The protonation states of titrable residues, *i.e.* Arginine (Arg), Lysine (Lys), Aspartic acid (Asp), Glutamic acid (Glu), Histidine (His), Cysteine (Cys) and Tyrosine (Tyr) were determined using PROPKA from PDB2PQR web server [34].

Adsorbent surface

The chemical structure of the SP Sepharose FF ligand is shown in Fig. 1 [4]. In this study, only ligands were represented in the model and the agarose matrix was not included. To simulate the linkage between ligands and agarose, a carbon atom was added and only the first three atoms coordinates were restrained. The 3D ligand structure was built using the Avogadro software [35], and then parametrized using CHARMM General Force Field (CGenFF) [36,37] and automated algorithms [38,39] from CGenFF website (https://cgenff.umaryland.edu). Maximum values of 0.671 and 1.5 were obtained for charge and parameter penalties respectively. Due to correct penalty scores (< 10), no parameter and charge optimization was performed. To build the SP Sepharose matrix, this ligand was duplicated and distributed uniformly on a square surface (10 nm × 10 nm) in -x and -y direction. The minimum spacing between ligands was defined in order to have a density of 3.24 ligands per nm^2 (according to the measured ionic capacity and the specific surface area of SP Sepharose [40]).

Whole system

As shown in Fig. 1, the chromatographic surface was positioned at the bottom of the simulation box (10.00 nm × 10.00 nm × 11.25 nm) in the -x and -y plane. Sodium ions were first added to neutralize the chromatographic surface and positioned close to the ligands. A vacuum layer, with a thickness of 1 nm, was kept under ligands to avoid the diffusion of water molecules attributed to periodic boundary conditions (PBC).

Then, the protein was added onto the surface with a distance of 0.75 nm from the closest ligand, and centered in the -x and -y plane. Six different starting orientations were sampled in order to screen the whole surface. The protein was first rotated by 90° around the -z axis, and then around the -y axis. After the rotation step, the protein was centered and translated to the correct distance from ligands.

The box was further filled with water molecules using *gmx solvate* tool and all water molecules below sulfur group of the ligands were discarded, leading to a total molecule number of about 27,000. In addition, to balance the global positive charge of the protein and to be consistent with experiments, 100 mmol/L of sodium chloride was also added. Finally, the whole system including protein, water molecules and ions, is about 100,000 atoms.

3.2.2. Molecular dynamics simulations setup

All MD simulations were carried out using the version 5.0 of the GROMACS simulation package [33] with the all-atom CHARMM36 Force Field [41] for protein and solvent (water and ions), whereas CGenFF [36,37] was used for the parametrization of SP Sepharose ligands. TIP3P model [42] was employed to simulate water molecules. The short-range non-bonded interactions were truncated using a distance cutoff of 1.2 nm and smoothly force switched to zero from 1.0 to 1.2 nm for van der Waals interactions. Long-range electrostatic interactions were treated with Particle Mesh Ewald (PME, cubic interpolation and Fourier spacing of 0.16 nm) [43] and 3dc algorithm [44] for correcting in -xy plane. Covalent bonds involving the hydrogen atoms were constrained using LINCS algorithm and default parameters [45]. To mimic the matrix of the chromatographic media, one wall was set at the bottom of the simulation box with atom types defined as CG331. To avoid water diffusion from the top and to be compatible with pressure in -xy and PME, another wall was set with atom types defined as OT to mimic water oxygen. The wall interactions were defined using direct Lennard-Jones potential and a linear potential threshold at 0.3 nm.

For all stages of the simulation protocol, the first three atoms of SP Sepharose ligands were spatially restrained with a force constant of 2000 kJ.mol^{-1}.nm^{-2} for each. The energy of the full system was first minimized with 50,000 steps of steepest descent algorithm or a convergence maximum force lower than 1000 kJ.mol^{-1}.nm^{-1}. A heating of 100 ps in NVT ensemble was performed to reach 293K and then a short equilibration of 200 ps in NPT was performed. During heating and NPT equilibration, additional position restraints of 1000 kJ.mol^{-1}.nm^{-2} were set for the protein backbone and the other ligand atoms. Finally, the production phase was run in NPT with only position restraints on the first three atoms of SP Sepharose ligands. MD time lengths ranged from 90 to ~ 300 ns according to the adsorption time, with a performance of 95 ns/day using 144 CPUs on CALMIP mesocenter. A time-step of 2 fs was used for all MD simulations. Temperature was controlled using the modified Berendsen thermostat [46] with a time constant for coupling of 0.1 ps and groups defined as protein, ligands and solvent. Semi-isotropic pressure at 1 bar was applied using Berendsen barostat [47] with a time constant for pressure coupling of 5.0 ps and a compressibility of 4e^{-5} bar^{-1} on -z axis only.

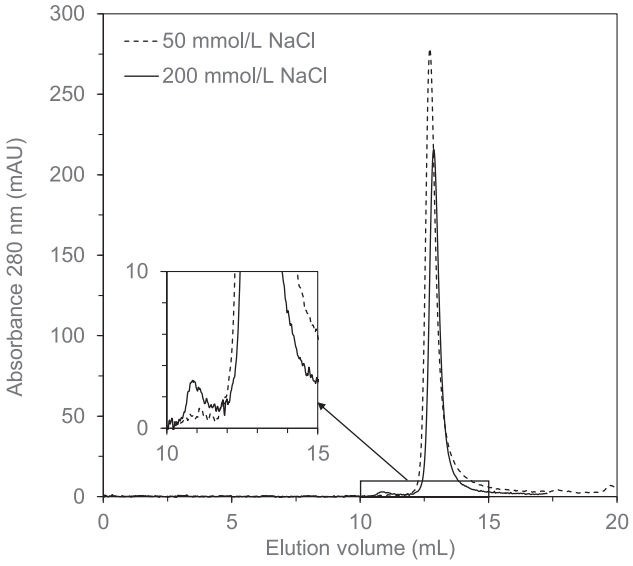

Fig. 2. HPLC-SEC chromatograms of two α-chymotrypsin solutions at pH 5 with different salt concentrations (50 and 200 mmol/L).

3.2.3. Software

Most of MD analyses were carried out using the tools provided by GROMACS software and script languages as bash, awk and python. The g_mmpbsa program, a tool developed for MM-PBSA (Molecular Mechanics Poisson–Boltzmann Surface Area) calculations using GROMACS and APBS [48,49], was used for binding free energy calculations. All 3D structure pictures were obtained using PyMol 1.7 [50]. Plots and statistical analyses were made using gnuplot and MS Excel, respectively.

4. Results and discussion

4.1. Adsorption equilibrium of α-chymotrypsin

4.1.1. Protein behavior

αCT was used to investigate a single-component adsorption equilibrium onto a cation exchange adsorbent, SP Sepharose FF. It is known that this protein can easily dimerize, depending on pH, temperature and ionic strength [51,52]. According to the literature, it appears that an increase of the ionic strength (over 100 mmol/L) and a pH around 4 favor the dimer formation. In this work, the dimerization of αCT was studied through HPLC-SEC analyses, to ensure that the monomer is the predominant form during experiments.

Fig. 2 shows the HPLC-SEC analysis of two αCT solutions with different salt concentrations (50 and 200 mmol/L) at pH 5.

The monomeric form of the protein (MW = 25 kDa) has an elution volume equal to 12.7 mL according to the column calibration; from the calibration curve and the MW of the dimer form (50 kDa), it can be deduced that this form should have an elution volume of 10.8 mL. This is confirmed in both chromatograms on Fig. 2 where a small but noticeable peak appears around 11 mL. The αCT being the only component in the solution that absorbs UV light at 280 nm, the presence of these peaks shows that protein dimerization occurs. As expected, the peak area increases with the salt concentration. However, in this case, the amount of dimers is clearly lower than the one of the monomer (less than 2%). Since the highest salt concentration used in the isotherm experiments is 100 mmol/L, the dimeric form can be neglected. Thereafter, the αCT is only considered in its monomeric form.

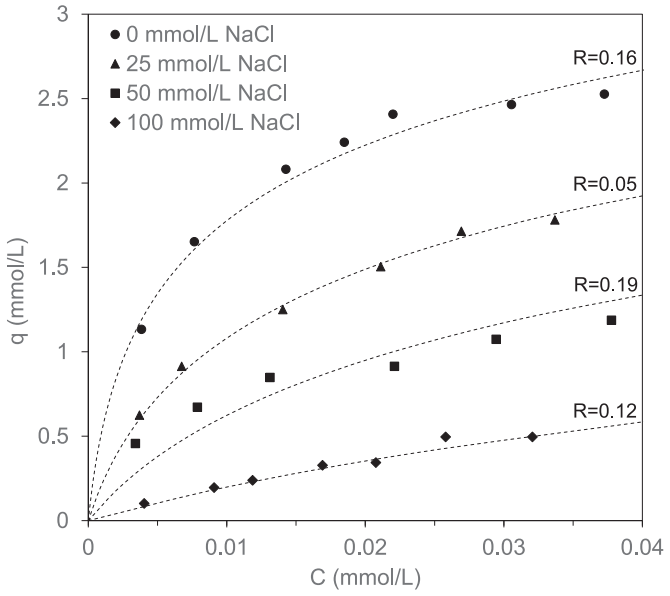

Fig. 3. Adsorption isotherms of α-chymotrypsin on SP Sepharose FF at pH 5 (50 mM sodium citrate buffer). In this figure, the protein concentration adsorbed on the SP Sepharose FF (q) is plotted against the protein concentration in solution (C). Symbols represent experimental results as a function of NaCl concentrations (circles for 0 mmol/L, triangles for 25 mmol/L, squares for 50 mmol/L and diamonds for 100 mmol/L) and dotted lines show the fitted SMA model curves.

4.1.2. Adsorption isotherms

The adsorption isotherms of αCT are presented on Fig. 3. As expected for single-component equilibrium, an increase of salinity leads to a weaker adsorption of the protein on the chromatographic surface [8,53]. Indeed, adding salt increases the number of counterions in competition with the protein on the charged ligands.

The four isotherms were then used to estimate the SMA parameters – the characteristic charge n, the steric factor σ and the equilibrium constant K – by curve fitting. For the four isotherms, the best fitting, $R = 0.14$ (see *Theory* section for further details) led to the following values: $n = 5.1$, $\sigma = 28.3$ and $K = 11.0$. The reliability factor was then calculated for each isotherm from the fitted parameters and the values are indicated on Fig. 3. Regarding the isotherms shape, it can be concluded that αCT has a good affinity with SP Sepharose FF at pH 5, with a relatively high characteristic charge. Indeed, in comparison with the steric factor which is dependent on the protein size, αCT seems to have a strong interaction with the chromatographic surface. These results are consistent with the fact that the working pH was lower than the protein isoelectric point *pI* (~ 8.3) [54]. The isotherms do not reach a plateau, which means that the protein does not cover the entire surface and is probably adsorbed in a monolayer. Indeed, a multilayer adsorption, an aggregation or precipitation of the protein at the surface would lead to an isotherm in the shape of a stepwise curve. However, it is not impossible that protein dimerization may occur at the chromatographic surface, but the amount of dimer would be insufficient to impact the isotherms shape, considering the low dimerization in solution (Fig. 2).

In addition, a breakthrough curve experiment was performed to verify whether the quantity of eluted protein corresponds to the q calculated. In this experiment, the calculated q was equal to 64.3 mg for 1 mL of bed volume. Then, the elution was performed using a 1 M NaCl solution (with working buffer) in isocratic mode, and the amount of eluted protein was determined to be around 70 mg. The protein quantity in eluate is slightly higher than the calculated q, but this can be explained by the residual protein solution in the pipes, which was eluted as well. Nevertheless, this result demonstrates that all the proteins retained in the column can be eluted.

4.1.3. Enzyme activities

Enzyme activity is expressed in moles of converted substrate per time unit and describes the quantity of active enzyme in a solution. A loss of activity may indicate for example a conformational change of the protein. To ensure that the ion exchange mechanism does not impact the active fold of the αCT, enzyme activity assays were performed before and after ion exchange experiments. The αCT activity of the initial solution was 93.5 mU/mg$_{\alpha CT}$. After the experiment, the protein was eluted from the column and the eluate was kept and analyzed. Its activity was 96.0 mU/mg$_{\alpha CT}$. The relative difference between those values (less than 3%) is mainly due to measurement uncertainties, the activity after ion-exchange has no physical reason to be higher than before. This result suggests that the αCT does not undergo conformational structure alteration due to the ion-exchange mechanism.

4.2. MD simulations

4.2.1. Protein behavior and adsorption stability

All MD simulations were carried out until a stable adsorption state was reached. The adsorption time was defined by monitoring two metrics during the simulation: (i) the distance between the protein center of mass (COM) and the ligands in the -z axis (*Min dist COM-LIG*); and (ii) the minimum distance between any atom of the protein and any atom of the ligands (*Min dist PROT-LIG*). All results are given in Supplementary Information (Figs. S1 and S2). The protein is considered adsorbed when the minimum distance between the Protein and the support is less than 0.2 nm, and the minimum distance between COM and the ligands does not fluctuate significantly (< 3-3.25 nm). These two conditions should be respected during at least 10 ns continuously to define stable adsorption and adsorption time.

As shown on Table 1, most of simulations led to stable adsorption before 100 ns (average adsorption time = 54 ns). It appeared that none of the six initial orientations led to unfavorable adsorption of the protein on the chromatographic surface. Two specific cases can be highlighted: (i) one MD has required a longer simulation time (up to 300 ns, see MD3 - Orientation 5 in Table 1), (ii) another one converged slowly for the minimal distance (Min dist COM-LIG) (see MD2 - Orientation 6 in Table 1 and Figs. S1 and S2). Fig. 4 illustrates this last case (MD2-O6). As shown, the protein changes its position and rotates in the simulation box during the trajectory. Sometimes, a desorption event is observed as shown at 60 ns, but at the end, the adsorption is stable with a constant value at least for the minimum distance (Prot-LIG), around 0.17 nm. The evolution of *COM-LIG* suggests more that the protein needed to reorient before and after adsorption to find its appropriate orientation on the surface. This explains why the evolution of *COM-dist* and *min-dist* are not simultaneously stable and *COM-dist* seems to converge slower. Notice that this behavior is valid for all other MD simulations (Figs. S1 and S2 in SI). However, it must be noticed that it is easier to set the threshold for the minimum distance between the protein and the ligands (Prot-LIG), compared to the minimum distance between COM and the ligands (COM-LIG). Moreover, this latter could not be lower than the gyration radius of the protein (~ 1.7 nm) and depends on the adsorption geometry. Thus, steady values for these two parameters indicate that the protein has a stable interaction with the chromatographic surface.

According to these results, the last 10 ns of each MD simulation seem representative of the adsorption state and were used for further analyses (binding patches and SMA parameter calculations).

Table 1
Total simulation time and adsorption time for each MD simulations (3 MD × 6 starting orientations). The adsorption time was defined by monitoring *COM-dist* (< 3.0 nm) and *min-dist* (< 0.2 nm) during at least 10 ns continuously of trajectory. *Only the adsorption time of this MD simulation was obtained with a threshold of 3.25 nm for *COM-dist*.

Orientation	MD Repeats	Total simulation time (ns)	Adsorption time (ns)
1	1	130	70
	2	90	25
	3	100	70
2	1	90	15
	2	100	48
	3	100	34
3	1	110	90
	2	95	54
	3	100	29
4	1	95	36
	2	100	11
	3	150	106
5	1	90	16
	2	100	24
	3	300	227
6	1	100	8
	2	120	91*
	3	100	16

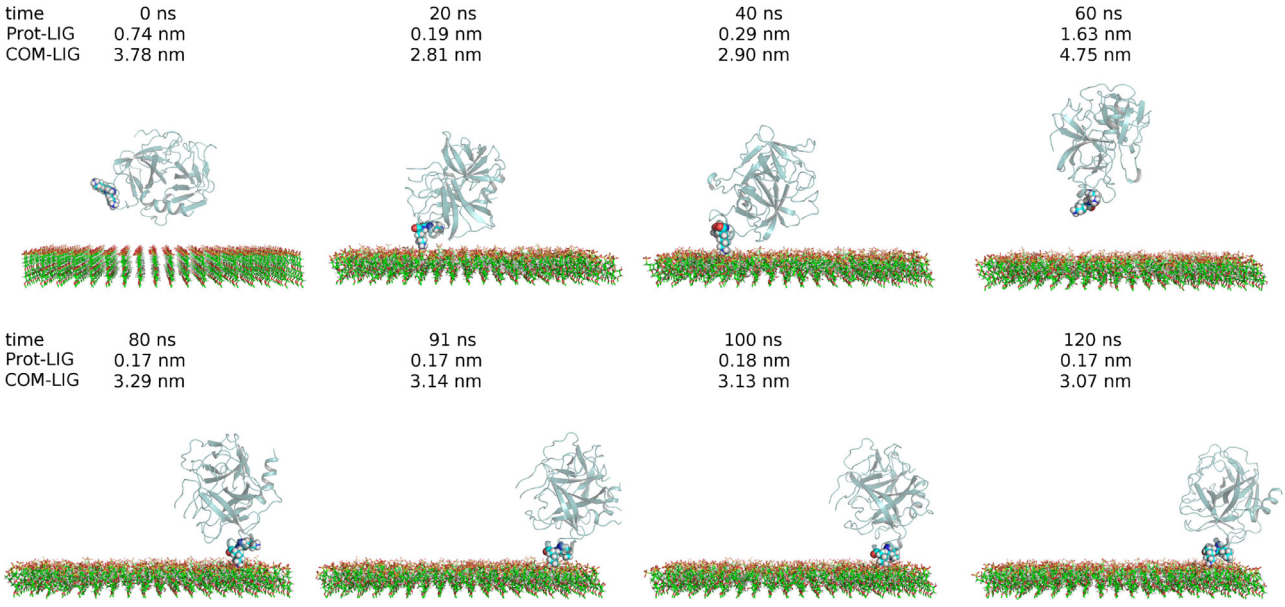

Fig. 4. Visualization of the diffusion and adsorption from MD2 - Orient6 simulation. For each snapshot, the MD time is given in nanosecond (ns). The minimum distance between COM and Ligand (COM-LIG) and the minimum distance between the αCT and the ligands are given in nanometer (nm). The protein is shown in lightteal cartoon whereas the chromatography surface (SP Sepharose) is shown in sticks. Two residues (Lys169 and Lys170) are shown in spheres and correspond to main residues involved in the adsorption.

4.2.2. Binding patches

The pH of the solution was set to 5, which is lower than the αCT isoelectric point *pI* (~ 8.3) [54] and leads to a positive global charge (+7) as shown in Fig. 5. This figure represents the charge distribution from APBS calculations on the protein surface. At pH 5, the protein is mainly positively charged (blue colored) despite some local neutral or negatively charged regions. It also highlights a strongly positively charged patch, from Lys84 to Lys177.

Protein residues were considered as being bound to the surface on the basis of their (positive) charge and their distance from the ligands (less than an arbitrary threshold value of 0.2 nm). In addition to Lysines and Arginines, which are positively charged at pH 5, Alanine-149 was also involved in some adsorptions, which can be explained by the protein structure: α-chymotrypsin is the digested form of α-chymotrypsinogen and is composed of three chains linked by disulfide bridges. Ala149 is the N-terminal of the third chain and is charged positively at working pH. Other residues (threonine, serine) were located close to the ligands during the adsorption, due to their spatial proximity with involved residues (example: Ser92 and Lys93), but also they can form hydrogen bonds with their hydroxyl group and the oxygen atoms of the sulfonic group of the ligands. These latter were not considered as binding residues.

Fig. 6 represents the presence-rate of binding residues near the ligands (time fraction during which the residue-ligand distance is lower than 0.2 nm) during the last 10 ns of each MD simulation.

This presence-rate allows accounting for residues that happen to move near ligands but end-up not remaining in interaction with them. This analysis highlights the presence of four recurring patches, which are represented on Fig. 7. These four Patches (P)

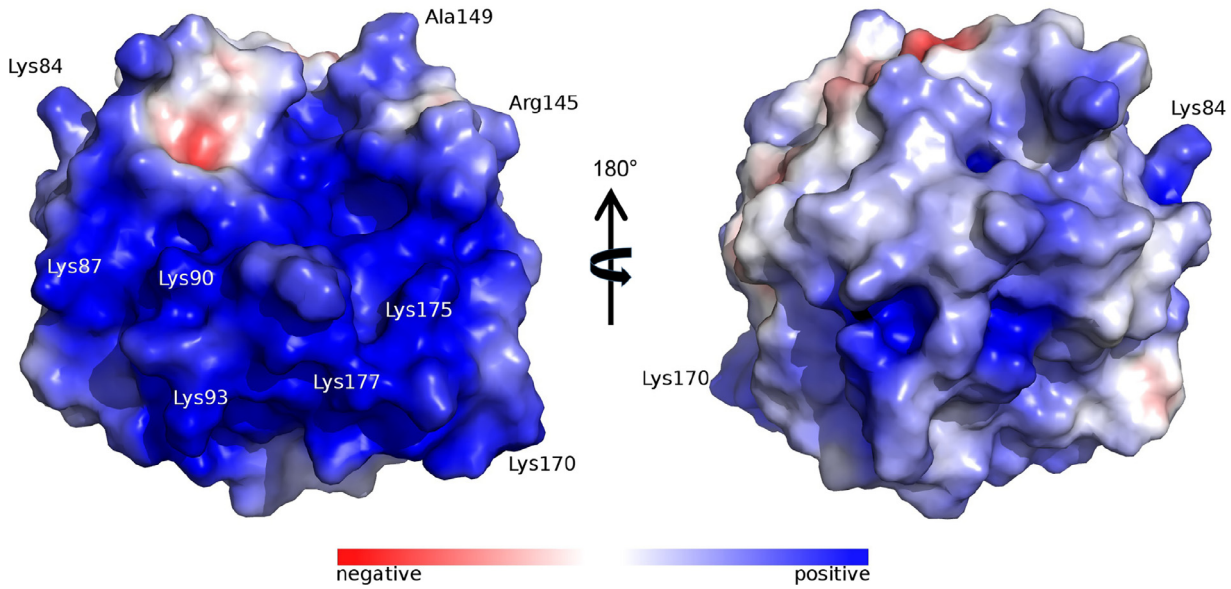

Fig. 5. α-chymotrypsin surface representation colored according to electrostatic potential (−5 kEV, red; +5 kEV, blue), at pH = 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

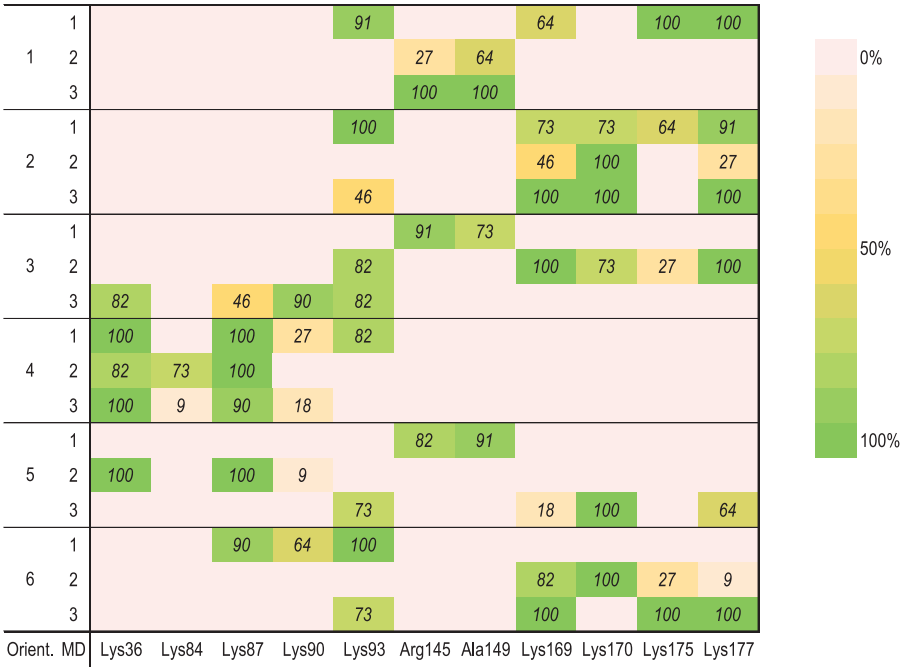

Fig. 6. Presence-rate of positively charged binding residues over the last 10 ns of each MD simulations.

can be defined as; Arg145-Ala149 (P1), Lys36-Lys84-Lys87-Lys90 (P2), Lys93 (P3) and Lys169-Lys170-Lys175-Lys177 (P4). Moreover, most residues from these patches belong to the "blue positively charged region" identified on Fig. 5. It must be noticed that several patches could participate to the final adsorption, which is mostly driven by the protein topology. For example, Lys93 is located between P2 and P4, and can be part of both (for example, MD1 from orientation 1 has a binding patch that includes P3 and P4; MD3 from orientation 3 includes P2 and P3). It is therefore geometrically impossible for the protein to have an interaction involving P2 and P4 at the same time. Furthermore, P1 is not located near the lysines and has a particular form: because these residues are located at chains termini, they can penetrate into the ligands layer and have a strong interaction.

To evaluate the strength of the interaction, a standard binding free energy calculation was performed using the MM-PBSA method. One may note that this method has been widely used to calculate binding free energy between a protein and another bio-molecule (protein or ligand). In the simulations, there are 324 charged ligands which could lead to over-estimated and non realistic free energy values, mostly on the electrostatic terms. To balance this effect, a counterion layer was kept over the ligands surface for the MM-PBSA calculations. It must be noticed that only the last frame of each simulation was analyzed and the energetic information was used in a qualitative and comparative way. The MM-PBSA results are presented in Table 2 and allow identifying the most energetically favorable patches. As expected, P1 (Arg145-Ala149) has the highest binding free energy while the lowest are

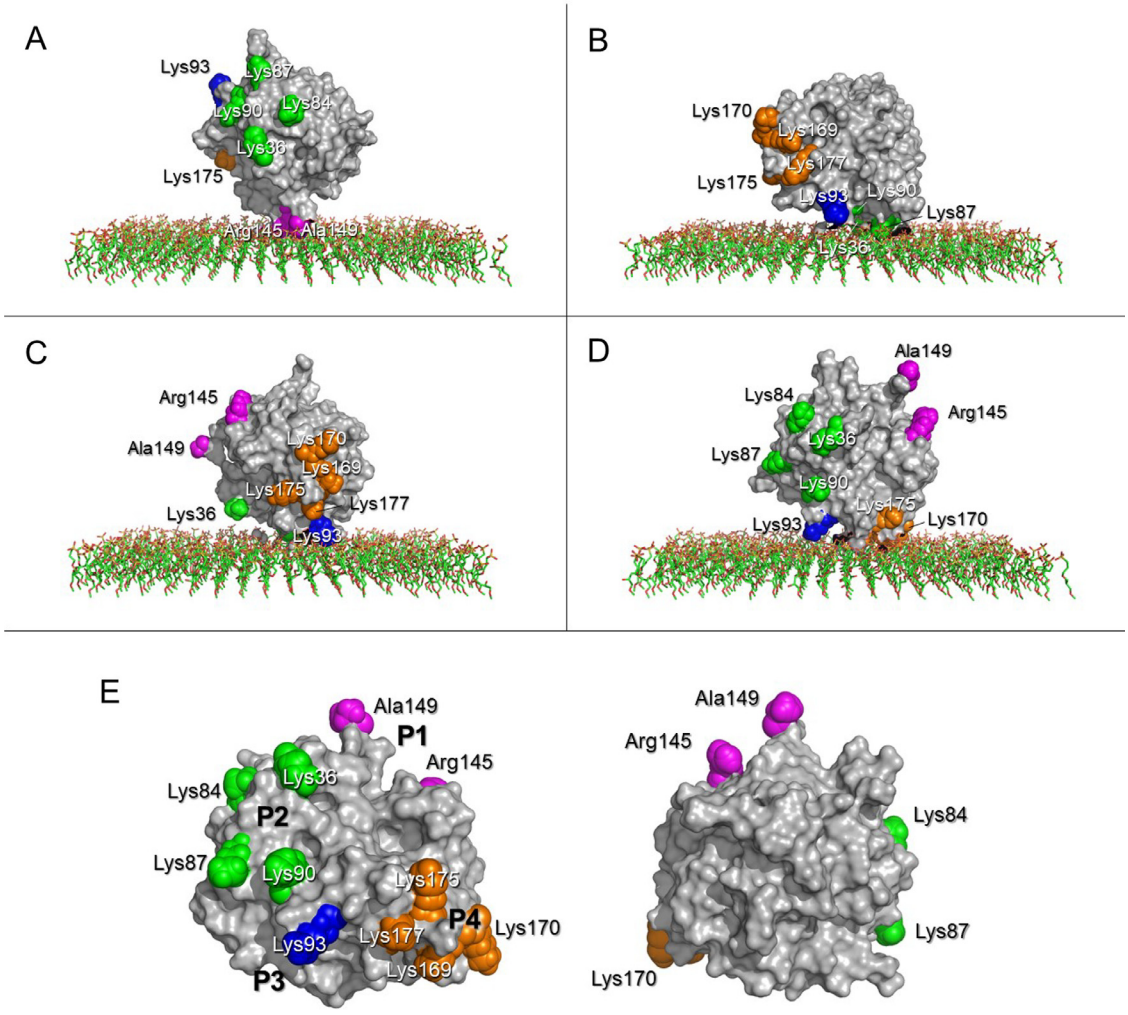

Fig. 7. Representation of the protein adsorbed state at the end (last frame) of four MD simulations, representing the four binding patches identified (Figures A from MD3-O1; B from MD3-O4; C from MD1-O6; D from MD3-O2). Patches are highlighted in color: P1 (magenta), P2 (green), P3 (blue) and P4 (orange). Figure E shows the protein and binding patches with the same view as in Fig. 5. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Binding residues during the last 10 ns of each MD simulations and the total binding energy. The energy was calculated using the MM-PBSA method and was performed only on the last frame.

Orientation	MD Repeats	Binding sites (Patches)	Total binding energy (kJ/mol)
1	1	Lys93 - Lys169 - Lys175 - Lys177 (P3-P4)	−2 913
	2	Arg145 - Ala149 (P1)	−2 225
	3	Arg145 - Ala149 (P1)	−2 589
2	1	Lys93 - Lys169 - Lys170 - Lys175 - Lys177 (P3-P4)	−4 430
	2	Lys169 - Lys170 - Lys177 (P4)	−2 764
	3	Lys93 - Lys169 - Lys170 - Lys177 (P3-P4)	−3 290
3	1	Arg145 - Ala149 (P1)	−2 128
	2	Lys93 - Lys169 - Lys170 - Lys175 - Lys177 (P3-P4)	−4 830
	3	Lys36 - Lys87 - Lys90 - Lys93 (P2-P3)	−3 564
4	1	Lys36 - Lys87 - Lys93 (P2-P3)	−1 184
	2	Lys36 - Lys84 - Lys87 (P2)	−2 477
	3	Lys36 - Lys87 (P2)	−2 554
5	1	Arg145 - Ala149 (P1)	−1 540
	2	Lys36 - Lys87 (P2)	−1 914
	3	Lys93 - Lys170 - Lys177 (P4)	−2 559
6	1	Lys87 - Lys90 - Lys93 (P2-P3)	−4 921
	2	Lys169 - Lys170 (P4)	−3 100
	3	Lys93 - Lys169 - Lys175 - Lys177 (P3-P4)	−2 830

Table 3
Characteristic charges (n) and steric factors (σ) calculated from MD simulations. Two averages are calculated (non-weighted and an energy-weighted) with their standard deviations and compared to the values from experiments.

Orientation	MD Repeats	Charge n	Steric factor σ
1	1	6.8	42.8
	2	1.0	56.6
	3	4.6	52.8
2	1	6.5	45.5
	2	3.3	44.0
	3	5.9	45.1
3	1	2.6	52.7
	2	6.4	44.3
	3	5.6	51.5
4	1	6.0	49.4
	2	4.2	52.5
	3	4.2	53.5
5	1	2.6	55.3
	2	4.4	50.7
	3	4.4	46.9
6	1	4.6	51.4
	2	4.1	47.7
	3	7.0	41.2
Average values from	Non-weighted	4.7 ± 1.6	49.1 ± 4.6
MD simulations	Energy-weighted	4.9 ± 1.5	48.5 ± 4.3
Values from experiments		5.1	28.3

mostly patches that include at least 4 or 5 lysine residues such as P2 and P4. The major part of the total energy is due to the electrostatic effects, mainly caused by the high number of charged ligands. Finally, it is reasonable to assume that patches having the lowest binding free energies (the most negative values) are likely to be the predominant protein orientation during ion-exchange experiments.

4.3. Comparison between experiments and MD simulations

The two SMA parameters n and σ, estimated both from experiments and MD simulations, were compared thanks to the physical meaning of these parameters which is consistent both at molecular and macroscopic scales.

First, the characteristic charge n experimentally corresponds to the number of binding sites and is related to the strength of protein adsorption. Based on this definition, this parameter n was computationally determined from a maximum distance of 0.2 nm between any atom of ligands and the protein. To be consistent with the SMA model and ion-exchange mechanism, only the ligands in interaction with positively charged residues were counted. The last 10 ns of each simulation were analyzed and results were averaged over this time range. Two metrics were compared, both a non-weighted average and weighted average based on the free energy value. For both experiments and MD simulations, results are given in Table 3. According to the isotherm experiments, the characteristic charge n is about 5.1. The non-weighted average from MD simulations is in very good agreement with a value of 4.7 and a standard deviation of 1.6. Interestingly, the energy-weighted average also tends to approach the experimental n, with a low relative difference of 3.9%. As expected, the energetically favorable patches might be the predominant protein orientation.

Second, the steric factor σ represents the number of ligands which are sterically hindered by the protein that are neither participating in the interaction nor available for another adsorption. To determine this σ parameter from MD simulations, a geometrical approach was developed in this work. As previously, the last 10 ns were analyzed. For each frame of the MD trajectory, the convex hull of the x and y coordinates of all protein atoms and its surface area were computed (Figure S3 in Supplementary Information) using the Python class called *scipy.spatial.Convexhull*. Then, knowing the ligand density, the total number of ligands inside the convex hull $N_{convhull}$ was calculated. Finally, the number of hindered ligands σ was determined by subtracting the characteristic charge n to $N_{convhull}$. The main advantage of this method is its independence from the ligand distribution, but it depends on two other parameters which are the total ionic capacity q_R and the specific surface S. Both were used to estimate the number of ligands in the simulation box (324 ligands for 100 nm^2). q_R was determined experimentally in triplicate with a low uncertainty (see Section 3.1.1), whereas S was found in the literature [40]. Table 3 gives the results for both experimental and computational approaches. Compared to the characteristic charge n, the difference between both approaches is higher, *i.e.* 28.3 from isotherms against 49.1 ± 4.6 (non-weighted average) from MD simulations. In spite of this discrepancy, the results are in the same range and seem to be in a fairly good agreement.

It should be noticed that uncertainties on experimental value of σ can be quite higher, mainly because of the determination of the specific surface area S of the chromatographic medium. These uncertainties will directly impact σ. For instance, an uncertainty of +20% on S would decrease the σ from 48.5 (energy-weighted average) down to 40.7. Indeed, experimental methods used to estimate this value, such as the Brunauer–Emmett–Teller (BET) adsorption method, are usually unsuccessfully applied to the SP Sepharose FF (and other sugar-based solid) because of possible matrix alterations, especially during drying. For this reason, DePhillips and Lenhoff [5] have suggested inverse size-exclusion chromatography in order to experimentally characterize internal surfaces of ion exchange resins. Nevertheless, it is difficult to conclude about the relevancy of these S values to estimate ligand density, as the surface accessible for ligands during the resin manufacturing process could be higher than that accessible in a packed column. According to experimental results, this analysis suggests that the ligand density determined from total capacity and specific surface area might be overestimated.

Both for n and σ, the other source of uncertainty comes from the least square method used to fit the SMA model on experimental data, which does not allow the estimation of standard deviations. This is a numerical method that leads to the best fitting, however other SMA parameters could lead to an accurate fitting regarding the measurement uncertainties. An other approach, such as Monte Carlo method, would be more appropriate to estimate the SMA parameters along with their standard deviations, taking into account the deviation observed on repeated experiments. This method was unsuccessfully applied to the experimental results because of the non-reversibility of the SMA equation. However, this method has to be studied further in order to be applied to the experimental data. Obviously, MD simulations results are also dependent on the used Force Field, the conformational sampling and the environment consideration. Taking into account all these limitations, our results open a door in the rational design of chromatographic processes.

5. Conclusion

In this paper, molecular simulations were used to better understand α-chymotrypsin behavior during ion-exchange on SP Sepharose FF, at chosen pH and ionic strength. Six starting orientations of the protein were studied with three MD simulation runs for each. Thus, the analyses were performed over the 18 simulations and averaged. These results were compared to macroscopic experiments through the Steric Mass Action formalism. Indeed, the SMA parameters, estimated from fitting the SMA law with the experimental data, have a physical basis and can be calculated from MD simulations at local scale.

Four charged amino acid patches (binding residues) were identified to be the predominant form of adsorption, which was confirmed through qualitative MM-PBSA binding free energy calculations. The results were in good agreement with the experiments, especially for predicting the characteristic charge. The steric factor calculation needs to be improved, and more specifically the ligands density determination. Despite this, the steric factor value is relatively close to the anticipated one.

This work allowed validating the relevance of atomic scale simulations to predict the protein adsorption behavior in adsorbents, and offers new perspectives to limit long and costly experiences in chromatography field for optimization of protein purification processes.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Marine Tournois: Investigation, Methodology, Formal analysis, Software, Validation, Visualization, Writing - original draft. **Stéphane Mathé:** Conceptualization, Supervision, Writing - review & editing. **Isabelle André:** Conceptualization, Resources, Supervision, Writing - review & editing. **Jérémy Esque:** Conceptualization, Methodology, Software, Supervision, Visualization, Writing - review & editing. **María A. Fernández:** Conceptualization, Project administration, Resources, Supervision, Writing - review & editing.

Acknowledgments

This work was supported by the French Ministry of Higher Education, Research and Innovation. We are grateful for its financial support. We wish to thank Dr. Patrick Fuchs for interesting discussions about protein adsorption using MD simulations. This work was granted access to the HPC resources on the TGCC-Occigen supercomputer and the Computing mesocenter of Occitanie Region (CALMIP, Toulouse, France).

References

[1] H.B. Costa, P.M. Fernandes, W. Romão, J.A. Ventura, A new procedure based on column chromatography to purify bromelain by ion exchange plus gel filtration chromatographies, Ind. Crops Prod. 59 (2014) 163–168, doi:10.1016/j.indcrop.2014.04.042.

[2] R.A. Latour, Molecular simulation of protein-surface interactions: benefits, problems, solutions, and future directions (review), Biointerphases 3 (3) (2008) FC2–FC12, doi:10.1116/1.2965132.

[3] L. Zhang, Y. Sun, Molecular simulation of adsorption and its implications to protein chromatography: a review, Biochem. Eng. J. 48 (3) (2010) 408–415, doi:10.1016/j.bej.2009.12.003.

[4] B.-L. Johansson, M. Andersson, J. Lausmaa, P. Sjvall, Chemical characterisation of different separation media based on agarose by static time-of-flight secondary ion mass spectrometry, J. Chromatogr. A 1023 (1) (2004) 49–56, doi:10.1016/j.chroma.2003.10.008.

[5] P. DePhillips, A.M. Lenhoff, Pore size distributions of cation-exchange adsorbents determined by inverse size-exclusion chromatography, J. Chromatogr. A 883 (12) (2000) 39–54, doi:10.1016/S0021-9673(00)00420-9.

[6] A.M. Lenhoff, Protein adsorption and transport in polymer-functionalized ion-exchangers, J. Chromatogr. A 1218 (49) (2011) 8748–8759, doi:10.1016/j.chroma.2011.06.061.

[7] A.M. Lenhoff, Ion-exchange chromatography of proteins: the inside story, Mater. Today 3 (10) (2016) 3559–3567, doi:10.1016/j.matpr.2016.10.038.

[8] S.R. Dziennik, E.B. Belcher, G.A. Barker, A.M. Lenhoff, Effects of ionic strength on lysozyme uptake rates in cation exchangers. i: uptake in SP sepharose FF, Biotechnol. Bioeng. 91 (2) (2005) 139–153, doi:10.1002/bit.20503.

[9] J. Liang, G. Fieg, Q.-H. Shi, Y. Sun, Single and binary adsorption of proteins on ion-exchange adsorbent: the effectiveness of isothermal models: liquid chromatography, J. Separat. Sci. 35 (17) (2012) 2162–2173, doi:10.1002/jssc.201200101.

[10] C.A. Brooks, S.M. Cramer, Steric mass-action ion exchange: displacement profiles and induced salt gradients, AIChE J. 38 (12) (1992) 1969–1978, doi:10.1002/aic.690381212.

[11] P. Raje, N.G. Pinto, Combination of the steric mass action and non-ideal surface solution models for overload protein ion-exchange chromatography, J. Chromatogr. A 760 (1) (1997) 89–103, doi:10.1016/S0021-9673(96)00812-6.

[12] J.C. Bosma, J.A. Wesselingh, Available area isotherm, AIChE J. 50 (4) (2004) 848–853, doi:10.1002/aic.10080.

[13] X.-L. Su, Y. Sun, Thermodynamic model for nonlinear electrostatic adsorption equilibrium of protein, AIChE J. 52 (8) (2006) 2921–2930, doi:10.1002/aic.10900.

[14] A. Creasy, G. Barker, Y. Yao, G. Carta, Systematic interpolation method predicts protein chromatographic elution from batch isotherm data without a detailed mechanistic isotherm model, Biotechnol. J. 10 (9) (2015) 1400–1411, doi:10.1002/biot.201500089.

[15] B. Coquebert de Neuville, A. Tarafder, M. Morbidelli, Distributed pore model for bio-molecule chromatography, J. Chromatogr. A 1298 (2013) 26–34, doi:10.1016/j.chroma.2013.04.074.

[16] Y. Huang, J. Bi, L. Zhao, G. Ma, Z. Su, Regulation of protein multipoint adsorption on ion-exchange adsorbent and its application to the purification of macromolecules, Protein Express. Purific. 74 (2) (2010) 257–263, doi:10.1016/j.pep.2010.07.002.

[17] B. Guélat, G. Ströhlein, M. Lattuada, M. Morbidelli, Electrostatic model for protein adsorption in ion-exchange chromatography and application to monoclonal antibodies, lysozyme and chymotrypsinogen a, J. Chromatogr. A 1217 (2010) 5610–5621, doi:10.1016/j.chroma.2010.06.064.

[18] L. Haggerty, A.M. Lenhoff, STM and AFM in biotechnology, Biotechnol. Progress 9 (1) (1993) 1–11, doi:10.1021/bp00019a001.

[19] K. Yang, Q.-H. Shi, Y. Sun, Modeling and simulation of protein uptake in cation exchanger visualized by confocal laser scanning microscopy, J. Chromatogr. A 1136 (1) (2006) 19–28, doi:10.1016/j.chroma.2006.09.036.

[20] C.A. Johnson, Y. Yuan, A.M. Lenhoff, Adsorbed layers of ferritin at solid and fluid interfaces studied by atomic force microscopy, J. Colloid Interface Sci. 223 (2) (2000) 261–272, doi:10.1006/jcis.1999.6680.

[21] S. Demanèche, J.-P. Chapel, L.J. Monrozier, H. Quiquampoix, Dissimilar ph-dependent adsorption features of bovine serum albumin and α-chymotrypsin on mica probed by AFM, Colloids Surf. B 70 (2) (2009) 226–231, doi:10.1016/j.colsurfb.2008.12.036.

[22] C.M. Roth, A.M. Lenhoff, Electrostatic and Van der Waals contributions to protein adsorption: computation of equilibrium constants, Langmuir 9 (4) (1993) 962–972, doi:10.1021/la00028a015.

[23] V. Noinville, C. Vidal-Madjar, B. Sebille, Modeling of protein adsorption on polymer surfaces. computation of adsorption potential, J. Phys. Chem. 99 (5) (1995) 1516–1522, doi:10.1021/j100005a023.

[24] M. Agashe, V. Raut, S.J. Stuart, R.A. Latour, Molecular simulation to characterize the adsorption behavior of a fibrinogen gamma-chain fragment, Langmuir 21 (3) (2005) 1103–1117, doi:10.1021/la0478346.

[25] M. Ozboyaci, D.B. Kokh, S. Corni, R.C. Wade, Modeling and simulation of proteinsurface interactions: achievements and challenges, Q. Rev. Biophys. 49 (2016), doi:10.1017/S0033583515000256.

[26] F. Dismer, J. Hubbuch, 3D structure-based protein retention prediction for ion-exchange chromatography, J. Chromatogr. A 1217 (8) (2010) 1343–1353, doi:10.1016/j.chroma.2009.12.061.

[27] J. Liang, G. Fieg, F.J. Keil, S. Jakobtorweihen, Adsorption of proteins onto ion-exchange chromatographic media: a molecular dynamics study, Ind. Eng. Chem. Res. 51 (49) (2012) 16049–16058, doi:10.1021/ie301407b.

[28] J. Liang, G. Fieg, S. Jakobtorweihen, Molecular dynamics simulations of a binary protein mixture adsorption onto ion-exchange adsorbent, Ind. Eng. Chem. Res. 54 (10) (2015) 2794–2802, doi:10.1021/ie504374x.

[29] J. Liang, G. Fieg, S. Jakobtorweihen, Ion-exchange adsorption of proteins: experiments and molecular dynamics simulations, Chemie Ingenieur Technik 87 (7) (2015) 903–909, doi:10.1002/cite.201400095.

[30] C.N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, How to measure and predict the molar absorption coefficient of a protein, Protein Sci. 4 (11) (1995) 2411–2423, doi:10.1002/pro.5560041120.

[31] S.K. Verma, K.K. Ghosh, Activity, stability and kinetic parameters for α-chymotrypsin catalysed reactions in aot/isooctane reverse micelles with non-ionic and zwitterionic mixed surfactants, J. Chem. Sci. 125 (4) (2013) 875–882, doi:10.1007/s12039-013-0434-6.

[32] B.W. Matthews, P.B. Sigler, R. Henderson, D.M. Blow, Three-dimensional structure of tosyl-α-chymotrypsin, Nature 214 (5089) (1967) 652–656, doi:10.1038/214652a0.

[33] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1–2 (2015) 19–25, doi:10.1016/j.softx.2015.06.001.

[34] T.J. Dolinsky, J.E. Nielsen, J.A. McCammon, N.A. Baker, PDB2Pqr: an automated pipeline for the setup of poisson-Boltzmann electrostatics calculations, Nucl. Acids Res. 32 (2004) W665–7, doi:10.1093/nar/gkh381.

[35] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, J. Cheminform. 4 (2012) 17, doi:10.1186/1758-2946-4-17.

[36] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A.D.M. Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force field, J. Comput. Chem. 31 (2010) 671–690, doi:10.1002/jcc.21367.
[37] W. Yu, X. He, K. Vanommeslaeghe, A.D.M. Jr., Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations, J. Comput. Chem. 33 (2012) 2451–2468, doi:10.1002/jcc.23067.
[38] K. Vanommeslaeghe, A.D.M. Jr., Automation of the CHARMM general force field (CGenff) i: bond perception and atom typing, J. Chem. Inf. Model. 52 (2012) 3144–3154, doi:10.1021/ci300363c.
[39] K. Vanommeslaeghe, E.P. Raman, A.D.M. Jr., Automation of the CHARMM general force field (CGenff) II: assignment of bonded parameters and partial atomic charges, J. Chem. Inf. Model. 52 (2012) 3155–3168, doi:10.1021/ci3003649.
[40] C. Chang, A.M. Lenhoff, Comparison of protein adsorption isotherms and uptake rates in preparative cation-exchange materials, J. Chromatogr. A 827 (2) (1998) 281–293, doi:10.1016/s0021-9673(98)00796-1.
[41] J. Huang, A.D. MacKerell, CHARMM36 Allatom additive protein force field: validation based on comparison to nmr data, J. Comput. Chem. 34 (2013) 2135–2145, doi:10.1002/jcc.23354.
[42] S.R. Durell, B.R. Brooks, A. Ben-Naim, Solvent-induced forces between two hydrophilic groups, J. Phys. Chem. 98 (1994) 2198–2202, doi:10.1021/j100059a038.
[43] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, A smooth particle mesh ewald method, J. Chem. Phys. 103 (1995) 8577–8593, doi:10.1063/1.470117.
[44] I.-C. Yeh, M.L. Berkowitz, Ewald summation for systems with slab geometry, J. Chem. Phys. 111 (1999) 3155–3162, doi:10.1063/1.479595.
[45] B. Hess, H. Bekker, H.J.C. Berendsen, J.G.E.M. Fraaige, LINCS: A linear constraint solver for molecular simulations, J. Comput. Chem. 18 (1997) 1463–1472.
[46] G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling, J. Chem. Phys. 126 (2007) 14101, doi:10.1063/1.2408420.
[47] H.J.C. Berendsen, J.P.M. Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, J. Chem. Phys. 81 (1984) 3684–3690, doi:10.1063/1.448118.
[48] R. Kumari, R. Kumar, A. Lynn, G_mmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. 54 (7) (2014) 1951–1962, doi:10.1021/ci500020m.
[49] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. 98 (18) (2001) 10037–10041, doi:10.1073/pnas.181342398.
[50] Schrödinger, LLC, The PyMOL Molecular Graphics System, Version 1.7, 2015.
[51] K.C. Aune, S.N. Timasheff, Dimerization of alpha-chymotrypsin. i. ph dependence in the acid region, Biochemistry 10 (9) (1971) 1609–1617, doi:10.1021/bi00785a017.
[52] K.C. Aune, L.C. Goldsmith, S.N. Timasheff, Dimerization of alpha-chymotrypsin. II. ionic strength and temperature dependence, Biochemistry 10 (9) (1971) 1617–1622, doi:10.1021/bi00785a018.
[53] M.A. Hashim, K.-H. Chu, P.-S. Tsan, Effects of ionic strength and ph on the adsorption equilibria of lysozyme on ion exchangers, J. Chem. Technol. Biotechnol. 62 (3) (1995) 253–260, doi:10.1002/jctb.280620307.
[54] P. Hudky, G. Kaslik, I. Venekei, L. Grf, The differential specificity of chymotrypsin a and b is determined by amino acid 226, Eur. J. Biochem. 259 (1–2) (1999) 528–533, doi:10.1046/j.1432-1327.1999.00075.x.