A two level hierarchical model of protein retention in ion exchange chromatography

Matteo Salvalaglio[a,1], Matteo Paloni[a], Bertrand Guelat[b,c], Massimo Morbidelli[b], Carlo Cavallotti[a,*]

[a] *Department Chimica, Materiali e Ingegneria Chimica, via Mancinelli 7 20131 Milano, Italy*
[b] *Institute for Chemical & Bioengineering, ETH Zurich, 8093 Zurich, Switzerland*
[c] *Novartis Pharma AG, Postfach, 4002 Basel, Switzerland*

Received 21 May 2015
Received in revised form 23 July 2015
Accepted 27 July 2015
Available online 29 July 2015

1. Introduction

Although the purification of therapeutic proteins is usually performed using at least an affinity chromatography step [1–5], it has been proposed that the level of purity required for pharmaceutical applications may be obtained also with non-affinity techniques such as Ion Exchange (IEX) Chromatography, which exhibits a large optimization margin compared to conventional processes [6–9]. Even if IEX technology is often used as a complement to traditional affinity purification steps, processes entirely based on IEX principles are rapidly becoming commercially available [6,10,11]. The rational optimization of IEX chromatography can significantly benefit from the comprehension of the fundamental phenomena occurring at the solid–liquid interface during protein adsorption [12–15]. To fulfill this goal, the analytical and numerical solutions of the Poisson–Boltzmann equation (PBE) have been often used to determine the energy of interaction of proteins with a charged surface. The use of PBE based methods for modeling interactions in IEX chromatography has been pioneered mainly by Stahlberg and co-workers [6,16–19] and has led to the development of analytical methods for the description of electrostatics-based retention [20,21].

The rapidly increasing availability of computational resources has recently attracted much interest towards the possibility of

* Corresponding author at: Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta", Politecnico di Milano, via Mancinelli 7, 20131 Milano, Italy. Tel.: +39 02 23993176; fax: +39 02 23993180.
E-mail address: carlo.cavallotti@polimi.it (C. Cavallotti).
[1] Present address: Department of Chemical Engineering, University College London, London WC1E 7JE, UK.

exploiting molecular modeling techniques to investigate complex systems at the molecular scale. Molecular dynamics simulations are in fact gradually becoming part of the arsenal of techniques applicable to address fundamental problems in protein purification [22–34]. The investigation of IEX adsorption processes at the atomistic scale has been performed with different approaches in the literature. Riccardi and Zhang focused their research on both the atomistic coarse grain simulations of charged proteins adsorption on IEX surfaces and their lateral diffusion [35] and on the molecular level modeling of the porous structure of the chromatographic supports [36–38]. Chung [39,40] and co-workers have proposed a structure-activity relationship study based on the calculation of the Poisson–Boltzmann interaction energy of a charged ligand with several mutants of the Cold Shock protein. The PBE was used to determine the influence of structural mutations on the most probable protein adsorption orientations. Similarly, Kubiak and Mulheran have investigated the dynamics of Lysozyme adsorbed on a charged surface by MD simulations to get a molecular insight on the most favorable binding orientations, finding that no tertiary structure depletion occurs when the protein adsorption takes place on a flat surface [41]. Dismer and Hubbuch have used a MD simulation protocol in order to determine the orientation of Lysozyme when adsorbed on a IEX surface, which allowed to determine an empirical correlation between the electrostatic component of the force field and the measured chromatographic retention parameters [42]. It was also found that, within typical simulation timescales, the protein does not macroscopically move its center of mass nor reorient itself with respect to the charged surface.

In principle, considering the systematic increase of computational resources, it may be conceivable to expect that within few years it will be possible to simulate protein adsorption on IEX surfaces using MD simulations in explicit solvent models. However this may not be so simple. A fully atomistic MD simulation in fact requires quite long simulation times in order to sample adequately the conformational space of the protein. In addition, using classical MD techniques, protonation states of all aminoacidic residues must be specified a priori, which may be an issue for pHs at which residues have similar probabilities of existing in the protonated and non-protonated states. Addressing properly this issue may require performing multiple MD simulations for the same system, which would increase considerably computational costs. In the present work we propose an alternative approach that combines the accuracy attainable through explicit MD simulations with the computational efficiency of models describing implicitly through the PBE the effect of the environment and, eventually, with semiempirical models. This paper is structured as follows: in Section 2 the hierarchical modeling strategy and experimental details are described in detail, while in Section 3 results are reported and discussed.

2. Method and theoretical background

The methodology here adopted to estimate protein retention coefficients is based on the calculation of free energies of interactions between the protein and a model surface. This is a problem well known to the scientific literature where many computational approaches, differing for level of complexity and accuracy, have been reported [12,19,21,24,41–45]. In the present work we investigate whether it is possible to determine the free energy of interaction of a protein with a surface using a computational approach formulated so that it retains a full molecular description of the system but it is also sufficiently simple to be applied to large and complex proteins on a routinely basis. In the following we describe the model, the adopted computational approaches, and report details of the experiments used for validation.

2.1. Free energy calculations

Two different synergic atomistic computational approaches were used to estimate the free energy of interaction of the proteins with the IEX surface, the first exploiting Umbrella Sampling (US) simulations and the second belonging to the family of the linear response approximation approaches. The two approaches differ significantly in their respective level of accuracy and computational cost. In addition, an analytical model is also used in order to test its capability to predict protein retention factors and to evaluate whether it may be used to complement and ease the computational cost of the other two models. The details of the semiempirical model are described by Guelat et al. [21], while the implementation of the MD and linear response model is described in the following two sections.

2.1.1. MD simulations

The first methodology used to determine interaction free energies between proteins and the IEX surface exploits free energy perturbation (FEP) theory. FEP was carried out through an US calculation protocol, in which the weighted histogram analysis method (WHAM) was applied to extrapolate the potential of mean force (PMF). Molecular dynamics simulations were performed in explicit water in the NPT ensemble. All simulations were performed with periodic boundary conditions and long-range interactions were computed with the Particle Mesh Ewald (PME) method. The simulation box for Lysozyme has dimensions of 145 Å × 145 Å × 100 Å and contains the IEX surface, the protein, 67,679 water molecules, and 450 Na^+ and Cl^- ions for a total of 20,8450 atoms. 162 Na^+ atoms have been added to obtain global electroneutrality. The surface was modeled by an ensemble of 169 propyl sulfonate groups organized in 13 rows of 13 elements equally spaced to give a global surface charge density of 1.57×10^{-6} mol m^{-2}. The simulations for Chymotrypsinogen A were performed using a similar system setup and parameters. The force field parameters of the IEX ligands were determined using the GAFF amber force field [45] with structure and charges calculated at the B3LYP/6-31+G(d,p) level and fitted using the RESP model. The 2VB1 (Lysozyme) [46] and 2CGA (Chymotrypsinogen A) [47] protein crystal structures were used for all the calculations. The AMBER ff03 force field [48] was used to model Lysozyme and Chymotrypsinogen A, while the TIP3P force field was used for water [49]. Throughout the simulations the IEX ligands were restrained to their position by applying a harmonic restraint of 1000 kcal mol^{-1} $Å^{-2}$ on the first and second C atoms of the propyl chain. A double restraint was applied to each ligand in order to limit its tendency to flip upside down in the course of the simulations. The cut-off for the non-bonded interactions was set to 10 Å.

The computational protocol adopted for the MD simulations was the following: first, a 2000-cycle energy minimization was carried out in order to remove bad contacts between the solute and the random-placed solvent molecules. In this step, solute molecules were restrained with a harmonic potential $k(\Delta x)^2$ with k set to 500 kcal $mol^{-1}Å^{-2}$. A second 3500-cycle energy minimization was performed for the entire system, removing the previous restraints. Then, the temperature was raised from 0 to 300 K by a simulated annealing of 20 ps at constant volume; a weak harmonic restraint (where k is now equal to 10 kcal mol^{-1} $Å^{-2}$) was applied on the solute with the purpose of avoiding wild fluctuations. This was followed by a 100 ps run at constant pressure, in order to allow the system relaxation and thus to reach the correct density. The system was finally relaxed through a 2 ns simulation performed at constant temperature and pressure. Temperature and pressure control was performed through Langevin dynamics with a collision frequency of 1 ps^{-1} and isotropic position scaling, respectively. The SHAKE algorithm was used for covalent bonds involving hydrogen atoms, allowing to adopt a time step of 2 fs. All MD simulations

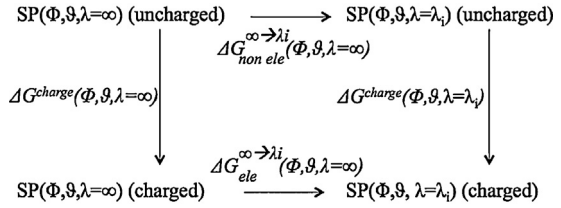

Fig. 1. Thermodynamic cycle used to calculate the free energy of interaction of globular proteins with a charged surface.

were performed using the Amber12 suite of programs [50]. The DFT calculations were performed using Gaussian 09 [51].

Following the 2 ns relaxation the protein was progressively detached from the IEX surface through a 10 ns run performed increasing constantly the distance between the center of mass (COM) of the protein and the IEX surface at steps of 0.5 Å from 22 to 39 Å for Lysozyme and from 29 to 49 Å for Chymotrypsinogen A by applying a 200 kcal mol^{-1} Å^{-2} harmonic potential. Intermediate structures were used as starting points for a series of biased US simulations (windows), which were performed imposing a harmonic restraint on the COM distance of 10 kcal mol^{-1} Å^{-2}. The simulation time for each window was 2 ns, with distances extracted every 20 ps discarding the first 0.1 ns. The unbiased statistics recovered from each window were joined using WHAM. WHAM equations were solved using our in-house code [52] until the maximum change of the PMF between successive iterations is smaller than 5×10^{-2} kJ/mol. The variance of the US trajectories for each window i from the n x_{ik} distances in the windows was computed as suggested by Zhu and Hummer [53] as:

$$var(\bar{x}_i) = \frac{1}{n(n-1)} \sum_{b=0}^{n-1} \left[\frac{1}{m} \sum_{k=bm+1}^{(b+1)m} (x_{ik} - \bar{x}_i) \right]^2 \quad (1)$$

The number of blocks b used to calculate the variance was 7. From the variance of the trajectories the cumulated variance of the PMF can be computed as:

$$var(G_i) = (K\Delta r)^2 \left[\frac{var(\bar{x}_1) + var(\bar{x}_i)}{4} + \sum_{j=2}^{i-1} var(\bar{x}_j) \right] \quad (2)$$

From var(G_i) the standard deviation, and thus the uncertainty of the PMF, was directly determined.

2.1.2. The linear response model

The second approach used to estimate the free energy of interaction between each protein and the IEX surface is based on the assumption that there is a linear relation between the interaction energy of the solute with the environment and the binding free energy [54]. In the present work the energy of interaction of the protein with the environment is calculated through the numerical integration of the PBE for a static charge distribution [55] defined by the protein atomic coordinates and by a purposely built charge distribution representing the IEX surface. This approach allows performing computationally efficient calculations of free energy profiles as a function of the protein–surface distance λ, contextually sampling the space of all possible protein/surface mutual orientations. We choose to represent the ensemble of protein/surface configurations as a function of the azimuth (ϕ) and elevation ($\triangledown$) angles The free energy change associated with the charging process of a surface–protein system (SP) in a given point of the three dimensional space defined by ϕ, ϑ, $\triangledown$, and λ, was calculated using the thermodynamic cycle reported in Fig. 1 [56].

With λ_i, we represent a generic discrete distance between the charged surface and the closest atom of the protein, where $\Delta G(\phi, \triangledown, \lambda)$ is calculated as:

$$\begin{aligned} \Delta G(\lambda_i) &= \Delta G_{ele}^{\infty \to \lambda_i} + \Delta G_{non\,ele}^{\infty \to \lambda_i} \simeq \delta G_{ele}^{\infty \to \lambda_i} \\ &= \Delta G^{charge}(\phi, \vartheta, \lambda = \lambda_i) - \Delta G^{charge}(\phi, \vartheta, \lambda = \infty). \end{aligned} \quad (3)$$

The reference state chosen to evaluate $\Delta G(\lambda_i)$ is the protein at an infinite distance from the surface ($\lambda = \infty$), which in the calculations is assumed to be equal to three times the typical cutoff used in molecular dynamics simulations (10 Å) [57]. The discrete λ_i points considered in our calculation thus span the range between 1 and 30 Å. ΔG^{charge} was calculated as:

$$\Delta G^{charge}(\phi, \vartheta, \lambda) = \frac{1}{2} \int \rho^f(\mathbf{r}) \psi(\mathbf{r}) dv, \quad (4)$$

where $\rho^f(\mathbf{r})$ is the fixed charges distribution defined by the atomic positions of the surface/protein system, and is thus function of the ϕ, ϑ, and λ degrees of freedom. $\psi(\mathbf{r})$ is the potential obtained from the numerical solution of the Poisson–Boltzmann equation. In order to account for all possible orientations contributing to the protein–surface interaction at a given distance λ, we compute the ensemble averaged $\Delta G^{charge}(\lambda)$ over the configurational space defined by the ϕ and ϑ angles. With the aim of explicitly highlighting the contribution to $\Delta G^{charge}(\lambda)$ due to the accessibility of the whole configurational ensemble defined by rotation in ϕ and $\triangledown$ at constant λ, we explicitly write the ensemble average of $\Delta G^{charge}(\phi, \triangledown\lambda)$ considering as a reference the minimum free energy configuration $\Delta G^{charge}(\phi_{min}, \triangledown_{min}, \lambda)$. The contribution due to the rotational average is thus expressed as a scaling pre factor written as the ratio of the 2D rotational partition functions $Q^{2D}{}_{rot}$ at distances λ and ∞ and leads to the following expression for $\Delta G^{charge}(\lambda)$:

$$\begin{aligned} &\exp\left(-\frac{\Delta G^{charge}(\lambda)}{RT}\right) \\ &= \frac{Q_{rot}^{2D}(\lambda)}{Q_{rot}^{2D}(\infty)} \exp\left(-\frac{\Delta G_{min}^{charge}(\phi_{min}, \vartheta_{min}, \lambda)}{RT}\right) \\ &= \frac{A_{Rot} \int\int \exp\left(-\Delta E^{charge}(\phi, \vartheta, \lambda)/RT\right)\, d\phi\, d\vartheta}{A_{Rot} \int\int \exp\left(-\Delta E^{charge}(\phi, \vartheta, \infty)/RT\right)\, d\phi\, d\vartheta} \\ &\exp\left(-\frac{\Delta G_{min}^{charge}(\phi_{min}, \vartheta_{min}, \lambda)}{RT}\right) \end{aligned} \quad (5)$$

where ϕ_{min} and ϑ_{min} are the angles at which the rotational minimum energy is found and ΔE^{charge} is the electrostatic interaction energy, which is twice the free electrostatic energy value. A_{rot} is a constant related to the angular momenta that are conjugate to the rotational angles ϕ and $\triangledown$ and is equal to $2\pi I k_B T/h^2$, where I is the protein inertia moment, k_B and h the Boltzmann and Planck constants, and T the temperature [58]. Due to the hypothesis of a rigid protein structure, A_{rot} are equal for numerator and denominator and thus cancel out. $Q_{rot}^{2D}(\infty)$ was calculated assuming that the 2D rotation in the solvent occurs over a 2D potential energy surface (PES) having a sinusoidal periodicity with maximum energy barriers of 14 kJ/mol, which is consistent with the apparent activation energy measured for the rotational diffusion of several proteins in water [59]. This is consistent with the Eyring model for diffusion, according to which the rotational and translational diffusion of molecules in solution requires overcoming energy barriers related to the disruption of the solvation shell and thus take place over a PES whose energy barriers are comparable to the activation energy experimentally measured for diffusing solutes [60]. A sensitivity analysis showed that $Q_{rot}^{2D}(\infty)$ is slightly affected by the shape and periodicity of the PES and is mostly influenced by the height of the

energy barrier. $Q_{rot}^{2D}(\infty)$ calculated over the 2D PES is about a factor of 10 smaller than that calculated over a flat PES. The contribution of non-electrostatic interaction forces, $\Delta G_{non\,ele}^{\infty\to\lambda}$, was introduced in the model as suggested by Roth and Lenhoff [43] using the Hamaker approach [61] to calculate the interaction free energy between a sphere of radius R and an infinite plane as:

$$\Delta G_{non\,ele}^{\infty\to\lambda}(\lambda) = -\frac{A_{132}}{6}\left(\frac{R}{\lambda} + \frac{R}{2R+\lambda} + \ln\left(\frac{\lambda}{2R+\lambda}\right)\right), \quad (6)$$

where A_{132} is the Hamaker constant for the interaction of two bodies through a medium.

The $\Delta G_{non\,ele}^{\infty\to\lambda}$ contribution was computed using a Hamaker constant of 1.0×10^{-20} J ($2.4k_BT$) for both proteins. This value was used by other authors to model the interaction between proteins and IEX surfaces [62] and is in the same order of magnitude of the 1.5–2.0k_BT value recommended by Roth et al. [63] for interaction between proteins and charged surfaces and the 3.3k_BT value determined for the interaction between BSA and a negatively charged IEX surface by Bowen et al. [64].

The interaction free energy $\Delta G(\lambda,I,\mathrm{pH})$ between protein and surface at a given pH and ionic force I is finally computed as:

$$\Delta G(\lambda, I, \mathrm{pH}) = \Delta G^{charge} + \Delta G_{non\,ele}^{\infty\to\lambda}(\lambda). \quad (7)$$

The model here proposed differs from the seminal model proposed by Roth and Lenhoff [43] for the way the rotational configurations are averaged in the evaluation of the electrostatic free energy contribution in Eq. (3) and for the numerical scheme used to integrate the Poisson–Boltzmann equation.

The linearized form of the Poisson–Boltzmann equation was numerically integrated using the APBS software [65] adopting the charge distribution determined with the AMBER ff03 force field [48]. On the basis of the characteristic time analysis, it was assumed that the charge of titrable moieties is not a function of the distance between protein and IEX surface and was therefore computed as function of pH using the PropKa3 software [66–68], which allows to adequately keep into account the local protein structure. A three-grid focusing solution strategy was used to set the Dirichlet boundary conditions for the electrostatic potential calculation. In order to model the ion exchange material as the surface of a solid the accessibility and dielectric constant maps have been purposely designed dividing the calculation domain in two sub volumes: a low dielectric volume inaccessible to the ions acting as a stationary phase bulk material, and a high dielectric volume representing the buffer solution environment. The surface charge distribution was mimicked through a layer of point charges held fixed in space that are equally spaced to give the surface charge density of the IEX resin under study, that is 1.57×10^{-6} mol m^{-2}. Several schemes where the number of charged points was increased contextually decreasing the local charge to maintain the reference charge density were considered. It was found that using point charges of −1 may lead to unphysical repulsion at short distances (smaller than 3 Å) due to excessive proximity of charged groups, which would not be present in a real system because of the protein and surface internal mobility not considered in this model. It was found that using −1/3 point charges alleviates this problem. The charged surface so obtained, which is thus intermediate between a discrete and a continuous surface, was used in the simulations.

The linear response model does not account for protein conformation changes that may take place when the protein approaches the IEX surface. For this reason the PMF computed using the MD model was used to correct the protein–surface interaction free energy calculated with the linear response approach. This was performed re-scaling linearly the electrostatic component of the interaction free energy, which contributes mostly, as will be shown in the following, to the interaction free energy. The rational for this procedure is that it is reasonable to expect that structural re-organization of the protein will be mostly driven by the optimization of the electrostatic interactions between some charged groups of the protein and the IEX surface, which may get in excessive proximity when moving the protein from the bulk of the solution towards the protein surface.

2.2. Evaluation of the retention factor from computational results

The adsorption equilibrium constant in chromatography is expressed through the Henry constant, defined as the ratio between the adsorbed and bulk protein concentrations in the limit of highly diluted systems, i.e. linear adsorption. As reported in the literature [21], it is difficult to define the adsorbed volume concentration since adsorption takes place at the solid/liquid interface. Thus it is useful to define the Henry constant as the excess concentration at the solid/liquid interface determined by the protein–surface attraction. The excess concentration can be defined as the difference between the concentration profile at a given I and pH and the same concentration profile in non-adsorbing conditions:

$$q^V(\lambda, I, \mathrm{pH}, I_{ref}, \mathrm{pH}_{ref}) = c(\lambda, I, \mathrm{pH}) - (\lambda, I_{ref}, \mathrm{pH}_{ref}) \quad (8)$$

where λ represents the plane-protein distance and can be considered as the reaction coordinate for the calculation of the concentration profiles of the protein at equilibrium as:

$$c(\lambda, I, \mathrm{pH}) = c_{bulk}\,\exp\left(\frac{-\Delta G(\lambda, I, \mathrm{pH})}{RT}\right) \quad (9)$$

The adsorption Henry coefficient can thus be conveniently expressed in function of the surface excess concentration q^S, which can be obtained by integration of the volume excess concentration q^V on the protein–surface distance degree of freedom λ:

$$H(I, \mathrm{pH}, I_{ref}, \mathrm{pH}_{ref}) = \frac{q^S(I, \mathrm{pH}, I_{ref}, \mathrm{PH}_{ref})}{c_{bulk}} = \int_0^\infty \frac{q^V(\lambda, I, \mathrm{pH}, I_{ref}, \mathrm{pH}_{ref})}{c_{bulk}} d\lambda \quad (10)$$

The Henry coefficient expressed as a function of $\Delta G(\lambda,I,pH)$ yields:

$$\begin{aligned} &H(I, \mathrm{pH}, I_{ref}, \mathrm{pH}_{ref}) \\ &= \int_0^\infty \frac{q^V(\lambda, I, pH)}{c_{bulk}} d\lambda - \int_0^\infty \frac{q^V(\lambda, I_{ref}, \mathrm{pH}_{ref})}{c_{bulk}} d\lambda \\ &= \int_0^\infty \exp\left(\frac{-\Delta G(\lambda, I, pH)}{RT}\right) d\lambda \\ &- \int_0^\infty \exp\left(\frac{-\Delta G(\lambda, I_{ref}, \mathrm{pH}_{ref})}{RT}\right) d\lambda \end{aligned} \quad (11)$$

The I_{ref} and pH_{ref} values, defining non-adsorbing conditions, are 0.8 and 6, respectively. In such conditions no adsorption is experimentally observed for the two proteins considered in this work, thus the second integral is equal to one.The retention factor k^I, in diluted conditions can now be calculated as:

$$k^I(I, \mathrm{pH}) = \phi H(I, \mathrm{pH}) \quad (12)$$

where ϕ is the ratio between the accessible surface area of the stationary phase and the volume of the mobile phase.

To compute the retention factor from molecular models we are assuming that the distribution of protein configurations in the space of the φ and θ angles has sufficient time to relax to the Boltzmann probability distribution and that both the surface and the protein possess time-independent surface charges. To justify these hypotheses we have carried out an analysis of the characteristic

times associated with protonation/deprotonation reactions, relaxation of the ionic atmosphere around a charged group exchanging protons with the environment, and diffusion processes occurring at the Debye length scale. The derivation of the characteristic times for the processes of rotational and translational diffusion and for proton transfer reactions is reported in the supporting information (SI), while the relative importance of the different processes is analyzed in Section 3.1.

2.3. Experimental determination of the retention factor

The retention factor was determined experimentally for Lysozyme and Chymotrypsinogen A as a function of pH and ionic strength. Following the procedure described in detail by Guelat et al. [21], the retention factor $k'(pH, I)$ was computed as:

$$k'(\mathrm{pH}, I) = \frac{t_R(\mathrm{pH}, I) - t_{ref}}{t_{ref}} \quad (13)$$

where $t_R(\mathrm{pH}, I)$is the retention time measured at assigned pH and ionic strength, and t_{ref} is the retention time measured in non-adsorbing conditions.

In order to favor a fair comparison between the retention factor calculated from molecular modeling and its experimental counterpart, experiments were performed with a non-porous strong cation exchanger column (commercial name: YMC-BioPro SP-F). The average particle size of the stationary phase is 5 mm and sulphopropyl groups are used as charged ligands. The retention time measurements were performed in isocratic conditions at pH between 4 and 6 in 20 mM buffers. A phosphate buffer was used in experiments performed at pH 6, while the other experiments were performed using an acetate buffer. The phosphate buffer was prepared mixing sodium dihydrogen phosphate (Fluka, Switzerland), disodium hydrogen phosphate (Lancaster, England) and sodium chloride (J.T. Baker, USA). The acetate buffer was prepared mixing sodium acetate (Merck, Germany), acetic acid (Carlo Erba reagents, Italy), and sodium chloride (J.T. Baker, USA). Water was filtered through a Millipore Synergy system before use. Experiments in non-adsorbing conditions were performed using a 20 mM phosphate buffer at pH 6. All experiments were performed using an Agilent 1100 Series HPLC, equipped with a quaternary gradient pump with a degasser, an autosampler, a column oven, and a diode-array detector. The retention times were determined from the maximum peak of the UV signal at 280 nm.

3. Results and discussion

In this section the results of model and experiments are reported and compared. The results of the MD simulations are discussed in Section 3.1, a characteristic times analysis, reported in detail in the SI and that is relevant in order to support some of the hypotheses on which the linear response model rests, is described in Section 3.2, the free energy of interaction of the considered proteins with the surface calculated with the linear response model is reported and discussed in Section 3.3, while experimental and calculated protein retention factors are compared in Section 3.4. The section is concluded by a comparison of the predictions of the coupled MD – linear response model with those of the semiempirical model of Guelat et al. [21].

3.1. Adsorption free energy from Molecular Dynamics simulations

Molecular dynamics simulations were used to investigate the conformational evolution of the protein on the surface and to sample the protein-IEX free energy surface (FES) as a function of the distance between the centers of mass of protein and IEX surface. The purpose of these simulations was to determine the interaction

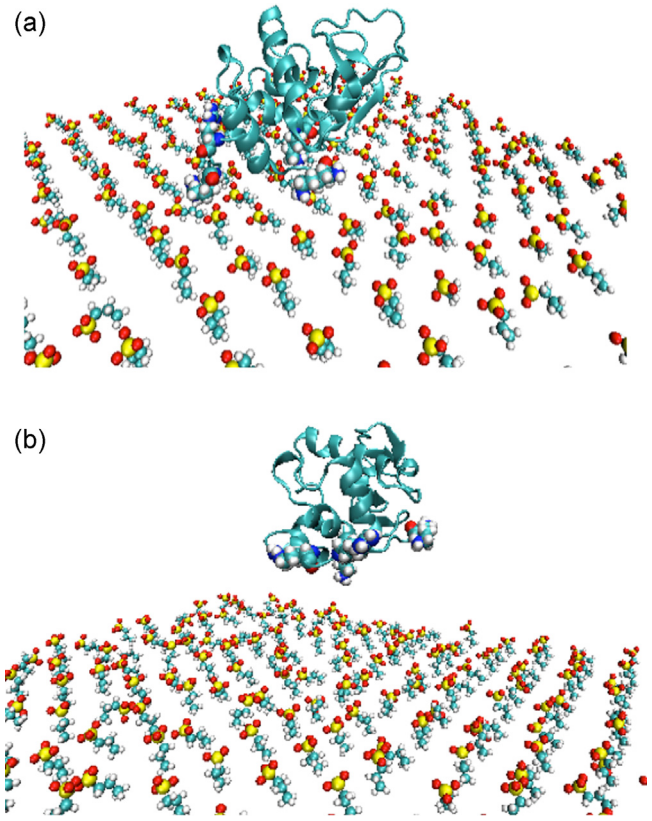

Fig. 2. Structure of Lysozyme adsorbed on the IEX surface (a) and in its desorbed conformation (b) determined using MD simulations.

energy between Lysozyme and Chymotrypsinogen A and the IEX surface accounting explicitly for the conformational change that is known to take place following the adsorption process [69]. Given the computational cost of these calculations (about 40 ns of simulation and 250,000 atoms), the MD simulations were performed only at a single pH and ionic strength (pH 4.0 and a ionic strength of 0.4 mol/L for Lysozyme; pH 6.0 and ionic strength 0.2 mol/L for Chymotrypsinogen A). The pHs at which simulations were performed were chosen so that the protonation state of each residue can unambiguously by assigned for each protein (assignment performed on the basis of the predictions of the propKa3 software).

Two snapshots of the simulations showing the minimum energy structure found for Lysozyme adsorbed onto the IEX surface and the structure it assumes at the end of the umbrella sampling simulation are reported in Fig. 2.

Simulations for Lysozyme were started from the minimum energy conformation determined through the rotational analysis performed using the linear response model (see Section 3.3) and were protracted for 2 ns to let the protein relax its conformation up to the structure shown in Fig. 2a. A successive 10 ns run was performed increasing constantly the distance between the COM of the protein and the IEX surface at steps of 0.5 Å from 22 to 39 Å. This allowed the protein to progressively relax its conformation up to the structure of Fig. 2b. Intermediate structures were used as starting points for the umbrella sampling simulations. A similar protocol was used for Chymotrypsinogen A. Post-processing the trajectories using the code described by Salvalaglio et al. [52] allowed calculating an adsorption free energy 27 kJ/mol for Lysozime and of 24 kJ/mol for Chymotrypsinogen A. The PMF profiles as a function of the protein-IEX surface distance are reported and discussed in Section 3.3.

3.2. Characteristic times analysis

The calculation of the adsorption free energy with the linear response model is based on the hypothesis of a stationary charges distribution. The aim of the present analysis is to assess the validity of this hypothesis through an order-of-magnitude analysis of the characteristic times of relevant processes taking place during adsorption. The characteristic times considered pertain to the following processes: ionic atmosphere relaxation (τ_{relax}, Section S1.1), proton exchange (τ_p, Section S1.2), translation diffusion (τ_D, Section S1.3), and rotational diffusion (τ_R, Section S1.4).

Since the relaxation of the ionic atmosphere occurs in the time span separating two successive proton exchange events, to ensure pseudo-stationary conditions the relaxation characteristic time has to be compared with the characteristic proton exchange frequency calculated when the maximum number of proton exchange events is observed, i.e. when pH = pKa, since in these conditions the characteristic time for protonation τ_p is equal to the characteristic deprotonation time τ_d. The ratio between the relaxation time τ_{relax} and the proton exchange reaction time τ_p has been calculated for carboxylic functional groups (pKa = 3.9) such as the Glutamic and Aspartic acid side chains solvated in a buffer with symmetric and monovalent ions (see Section S2 for details) and found to be at least two orders of magnitude faster than the proton exchange event. It means that the composition of the solution that surrounds the protein can be considered in a pseudo stationary state with respect to the change of the local surface charge of the protein.

The proton exchange events affect the protein electrostatic potential, inducing local surface charge fluctuations that have an impact on the interactions with the IEX surface. The ratio between the characteristic times of diffusion and proton exchange can be expressed as:

$$\beta = \frac{\tau_D}{\tau_p} = \frac{3\pi\varepsilon_0\varepsilon_r\eta r \cdot k_p 10^{-pka}}{N_A e^2 I} \tag{14}$$

The β parameter is function of the protein structure through its sphere-equivalent radius r, the ionic strength I, and the proton exchange rate (k_p).

The ratio between translational and rotational diffusion characteristic times determines if a protein can reorient itself while approaching the surface at the Debye length scale. It can be expressed as:

$$\alpha = \frac{\tau_D}{\tau_\theta} = \frac{3\varepsilon_0\varepsilon_r}{N_A e^2 I} \cdot \frac{k_b T}{8r^2(2\pi)^2} \tag{15}$$

This parameter accounts for the number of full rotations made by a protein of radius r while traveling along distances of the Debye length order of magnitude. The parameter α decreases both with the inverse of the square of r and with the inverse of I.

A rationalization of all the possible scenarios for a single protein approaching the IEX surface can be formulated as a function of α and β as follows (see Fig. 3): (i) proton exchange is faster than diffusion, while rotation is slower ($\alpha < 1$, $\beta > 1$); it means that the solution of the PBE has to be averaged on the equal protonation states, but not on all the possible orientations since the protein does not have enough time to visit a significant number of orientation configurations while approaching the IEX surface; (ii) both proton exchange and rotation are faster than diffusion ($\alpha > 1$, $\beta > 1$); it means that the electrostatic potential must be averaged on equal protonation states and on all the possible orientations; (iii) both proton exchange and rotation are slower than diffusion ($\alpha < 1$, $\beta < 1$); it means that each protein approaches the surface with a certain protonation configuration and also with a fixed orientation; *iv*) rotation is faster than diffusion, while proton exchange is slower ($\alpha > 1$, $\beta < 1$); it means that the electrostatic potential produced by the protein must be averaged on all the possible orientations but not on all the equal protonation states, since the protein does not have enough time to visit a representative number of protonation states while approaching the IEX surface. In Fig. 3 it can be observed how large globular proteins, with radii ranging from 2.5 nm to 10 nm, are characterized by α values of about one and β values larger than one, which corresponds to a scenario that is intermediate between (i) and (ii).

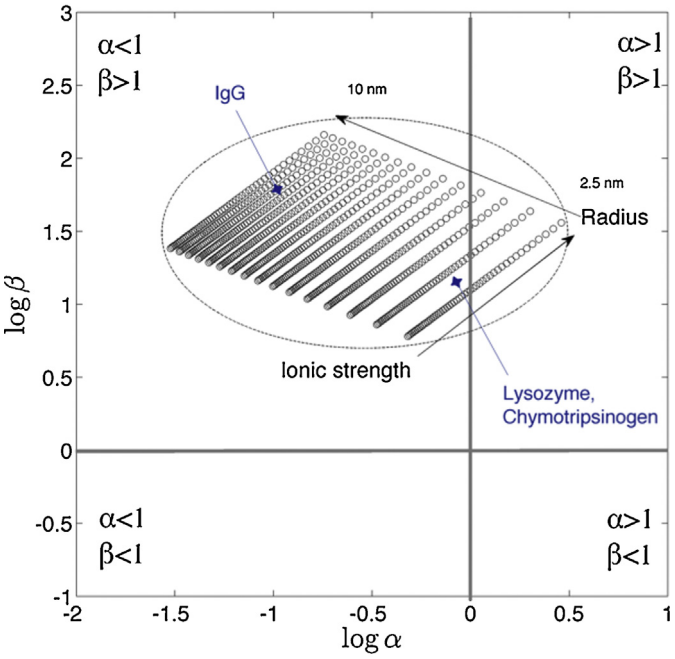

Fig. 3. Characteristic time ratios α and β calculated for different radii and ionic strengths.

From the characteristic times analysis the following considerations can thus be drawn on Lysozyme and Chymotrypsinogen A:

(I) the ionic atmosphere relaxation is always at least one order of magnitude faster than the fastest proton exchange event. The electrostatic potential around the protein can thus be considered always in equilibrium, since the reorganization transient does not affect it. Therefore it is reasonable to use of an implicit approach based on the PBE to determine its local value.

(II) Proton exchange events are faster than diffusion at the Debye length scale. This means that the protein interaction with the surface can be described with an electrostatic potential averaged over equivalent protonation states.

(III) Rotation is comparable to diffusion, so that a complete rotation of the protein while approaching the surface will be improbable. This means that the relaxation to the thermodynamic limit of the adsorbed proteins distributions in the space of all possible orientations may require a complex mechanism involving several adsorption and reorientation steps rather than a contextual roto-translation taking place while the protein is approaching the surface.

Therefore, in order to determine the characteristic time of evolution of the orientation of an adsorbed protein a simple kinetic model, reported and described in detail in the supporting information, was developed. Although this model was obtained with an order-of-magnitude approach based on strong simplifying hypotheses, its results show that kinetic effects of reorientation of the adsorbed proteins do not affect the orientation distribution. We find that such distribution relaxes to the thermodynamic limit

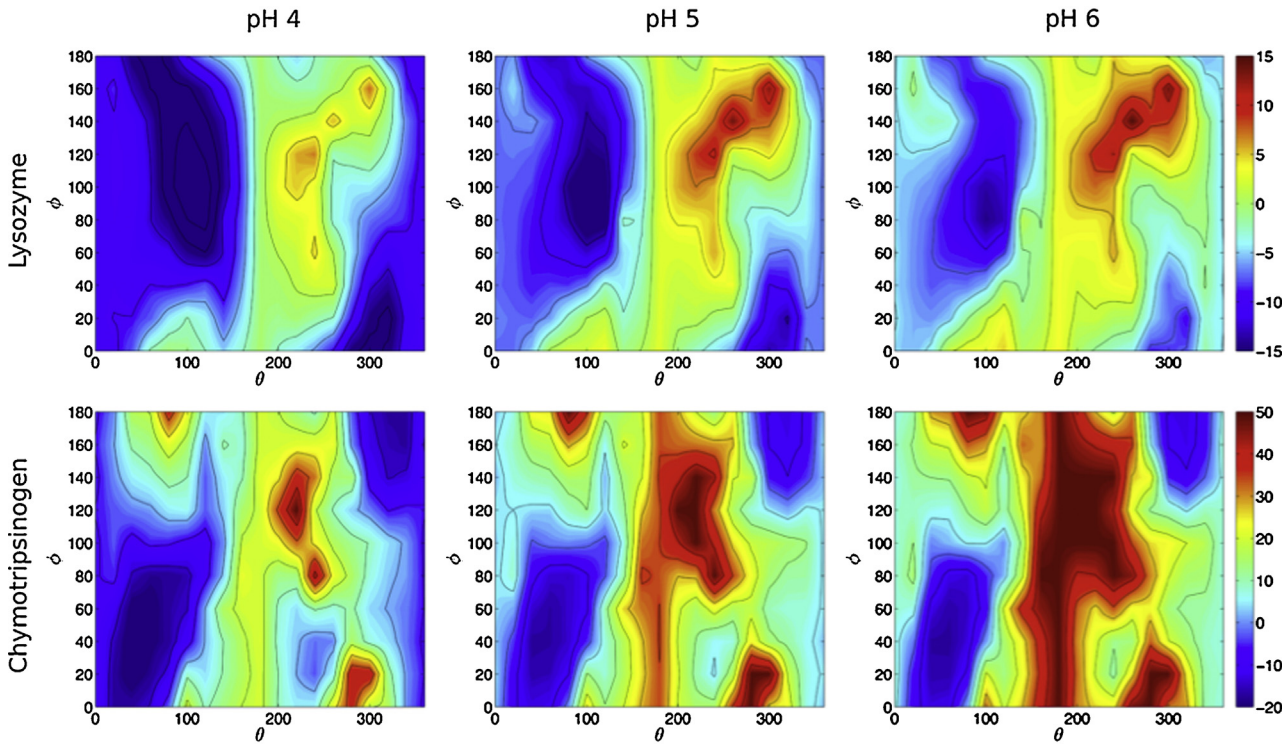

Fig. 4. Energy (kJ/mol) maps calculated at the plane – protein distance where the interaction energy is maximum for Lysozyme and Chymotrypsinogen A interacting with a distribution of charges at three different pH values, reported as a function of the angles defining the relative orientation between protein and surface. The ionic strength is 0.25 mol/L for the Lysozyme maps and 0.35 mol/L for the Chymotrypsinogen A maps.

within times of the order of magnitude of 10^{-3} s, much smaller than the typical elution time.

An approach based on the calculation of equilibrium-averaged adsorption free energies can thus be safely adopted to estimate the affinity of globular proteins with charged surfaces.

3.3. *Adsorption free energy from linear response simulations*

The free energy of interaction of protein and surface was determined as a function of three parameters: two angles, defining the relative orientation of the protein with respect to the surface, and the distance between the center of mass of the protein and the surface. Calculations were performed for pH values of 4, 4.5, 5, and 6 at different ionic strengths. The results calculated at pH 4, 5, and 6 at the distance at which the electrostatic interaction free energy ΔG^{charge} is minimum for a selected ionic strength are sketched in Fig. 4. The present results do not incorporate the contribution of van der Waals interaction free energies, which were computed as described in the method section and are about 4 kJ/mol at the distance at which ΔG^{charge} assumes its minimum value. Except for a few configurations in which the interaction energy is positive, the free energy surfaces computed for Lysozyme (Fig. 4, top) highlight how, generally, the vast majority of the configurations exhibit a negative free energy change associated to the protein interaction with the surface. This finding reflects the widespread positive surface charge that characterizes Lysozyme, which has several ARG and LYS vicinal residues on its surface. The number of orientations characterized by negative interaction energy diminishes with the increase of pH, following the decreasing trend of the free energy. The change of the interaction energy is the consequence of the dependence of the protein charge on pH, as the positively charged amino groups and the neutral carboxylic acid groups are deprotonated with increasing pH and become respectively neutral and negatively charged.

The analysis of the orientation dependence of the adsorption free energy for Chymotrypsinogen A highlights a fundamental difference with respect to Lysozyme. The dipole-like nature of the surface charge distribution for this protein, due to the presence of both positively and negatively charged residues organized in patches, reflects on the ratio between favored and unfavored interacting orientations, which is close to 50%. The distribution of the charged groups on the protein surface determines the concentration of the populated configurations in restricted areas rather than in wide regions of the ϕ–θ space. The most favored orientations for Lysozyme and Chymotrypsinogen A, shown in Fig. S4, are characterized by the maximum possible exposure of positively charged groups to the negatively charged surface. For Lysozyme the most favored orientation is characterized by simultaneous exposure of Lys 1, Arg 5, Lys 33, Arg 125, and Arg 128 to the charged surface. These residues form in fact a positively charged patch that is capable of the strongest interactions with the surface. Chymotrypsinogen A is characterized by a smaller net charge compared to Lysozyme and by the aforementioned dipolar surface charge distribution. This aspect leads to the scattering of the most favored configurations over the ϕ–θ space. The charged residues that constitute the main positively charged patch are Lys 36, Lys 90, Lys 93, Lys 84 Lys 87, Lys 107, and Lys 82. The interaction free energy as a function of the plane-protein distance and ionic strength is reported for Chymotrypsinogen A in Fig. 5.

The PMFs of Lysozyme and Chymotrypsinogen A computed using the linear response approach without performing the rotational conformational average are compared with the PMFs determined using the WHAM protocol in Fig. 6a andb. The reason why the contribution of the conformational average on the free energy estimation was not included in the present comparison is that the limited simulation time we used in the MD simulations does not allow to explore efficiently the translation and rotational conformational space. A similar situation was encountered by Bouvier [70].

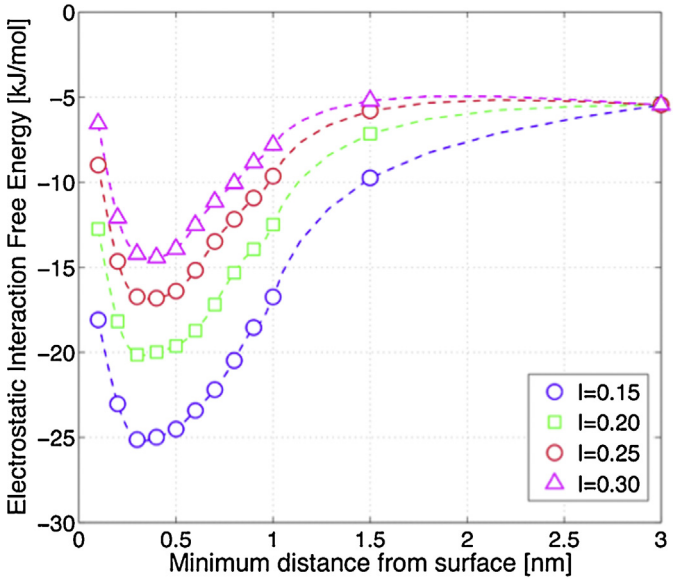

Fig. 5. Electrostatic free energy of interaction as a function of separation and ionic strength calculated at pH 4 between Chymotrypsinogen A and the IEX surface using the linear response model. The free energies have been calculated using Eq. (7) and averaged over all orientations through Eq. (5).

Luckily, this is not an issue in the present study as the conformational average is properly included in the linear response calculations, so that the uncertainties introduced by the limited simulation time used in the MD simulations are compensated by the extensive rotational conformational sampling performed with the linear response model. In perspective, increasing the simulation time for each umbrella sampling window, thus allowing for an ergodic sampling of all possible protein/surface orientations, will enable a direct comparison of the MD results with the conformationally averaged free energies of interaction obtained using the linear response model.

The analysis of the PMFs shown in Fig. 6 evidences that the linear response simulations underestimate the interaction free energy computed with the WHAM approach for Lysozyme and slightly overestimate it for Chymotrypsinogen A. For Lysozyme, this is determined by the mobility of its positively charged residues (most notably Lys 1, Arg 5, Lys 33, Arg 125, and Arg 128) that, in proximity of the charged surface, are able to move sufficiently to optimize the interaction with the IEX surface. The difference between the free energies computed at the two levels of theory for Chymotrypsinogen A, a factor of 0.9, is within the uncertainty of the present calculations and suggests that the conformational reorganization of the protein is probably not energetically relevant for this system. In order to correct the linear response simulations with the US/WHAM estimations of the protein – IEX interaction free energy the electrostatic component of the linear response PMF was rescaled by a factor of 1.2 for Lysozyme and 0.9 for Chymotrypsinogen A. The standard deviation of the WHAM PMF, calculated as suggested by Zhu and Hummer [53], is about 4 kJ/mol for both proteins, indicating that this is the level of uncertainty of the present simulations for what concerns the WHAM protocol. Finally, it is interesting to observe that the difference in the PMF shapes computed using WHAM and the linear response approach for Chymotrypsinogen, most notably the plateau of the WHAM PMF between 38 and 45 Å, is determined by a macroscopic rotation of the protein that optimizes its interaction energy as the distance between the COM is forced to increase.

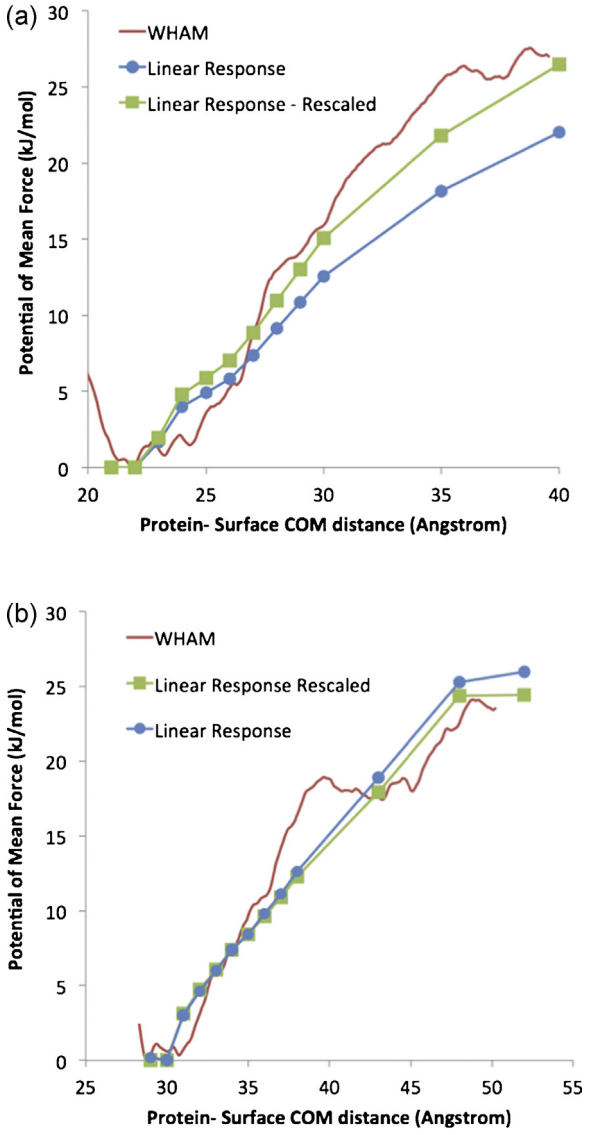

Fig. 6. Potential of mean force calculated for Lysozyme (a) and Chymotrypsinogen A (b): using the WHAM analysis of the MD trajectories (line), the linear response approach without accounting for the rotational average (line with circles), and rescaling the electrostatic component of the linear response PMF using a factor of 1.2 for Lysozyme and 0.9 for Chymotrypsinogen A to fit the WHAM profile (line with squares).

3.4. Retention factor: prediction and comparison with experiments

The retention factor k^I can be calculated from the free energy $\Delta G(\mathrm{pH},I,\lambda)$ as a function of pH and ionic strength through Eq. (12). The k^I factors determined for Lysozyme and Chymotrypsinogen A are reported and compared with experimental data in Figs. 7 and 8, respectively. As it can be observed, the model predictions, obtained using the linear response model and rescaled over the WHAM PMF, are in good qualitative and quantitative agreement with the experimental retention factors determined as a function of pH and ionic strength for both proteins, in particular considering that no fitting parameter was used. One noteworthy aspect is that the linear response free energies of all the simulations where rescaled using the same correction factors, thus indicating that using as correction factor for the protein conformational reorganization on the IEX surface a constant fraction of the minimum interaction free energy is a reasonable approximation. These results

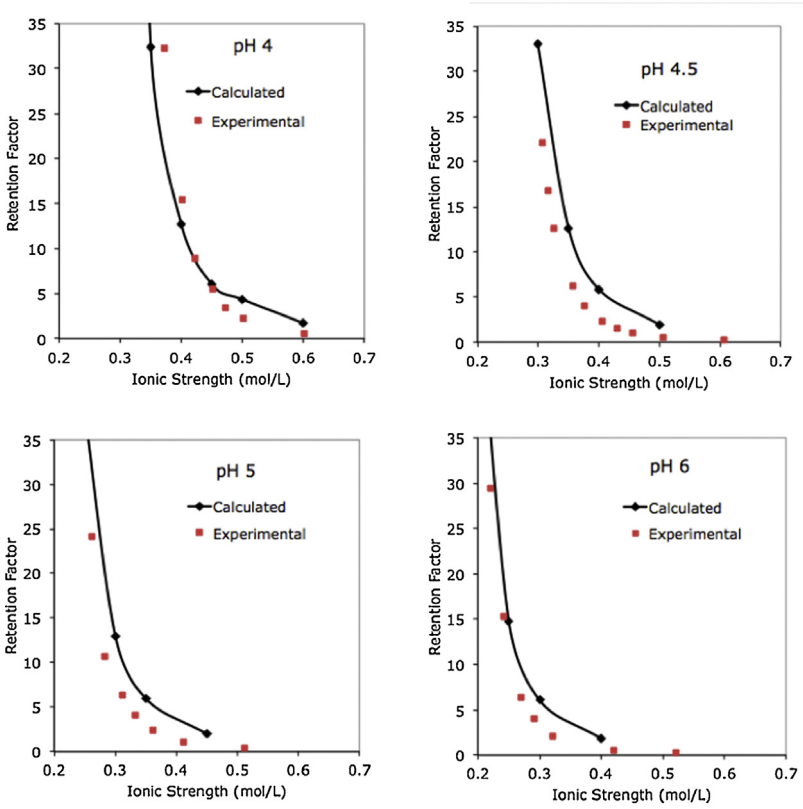

Fig. 7. k^l retention factor calculated (continuous line) and measured experimentally (points) for Lysozyme at pH 4, 4.5, 5, 5.5, and 6.

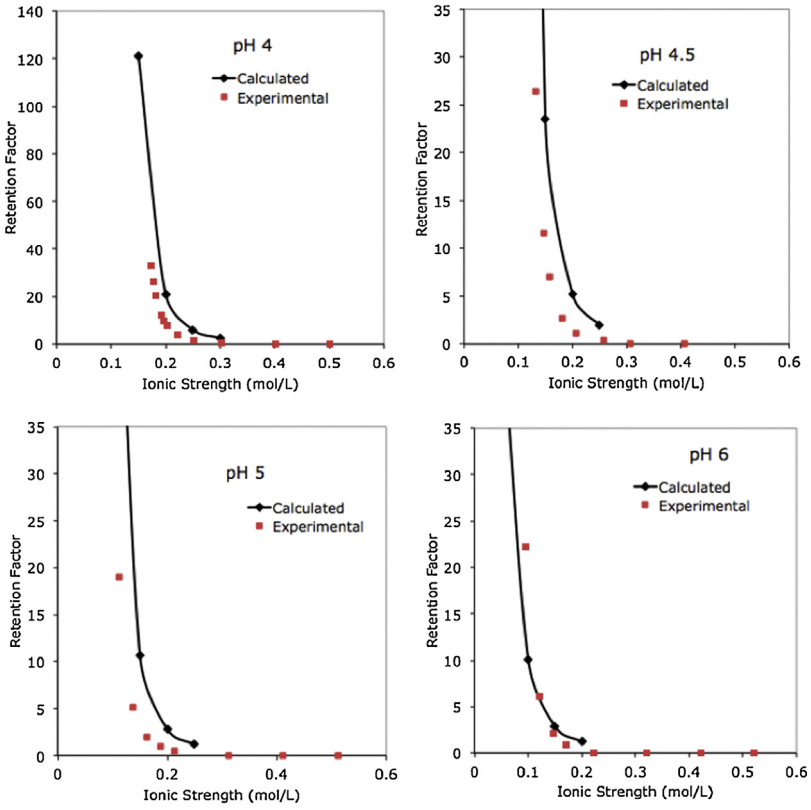

Fig. 8. k^l retention factor calculated (continuous line) and measured experimentally (points) for Chymotrypsinogen A at pH 4, 4.5, 5, and 6.

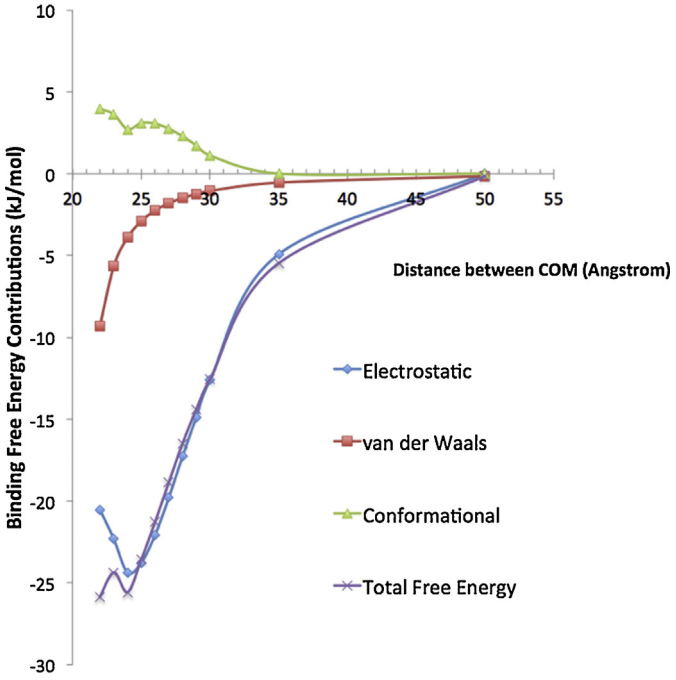

Fig. 9. Breakdown of the contributions to the interaction free energy between Lysozyme and the IEX surface computed at pH4 for a ionic strength of 0.4 mol/L as a function of the distance between the protein-IEX surface centers of mass.

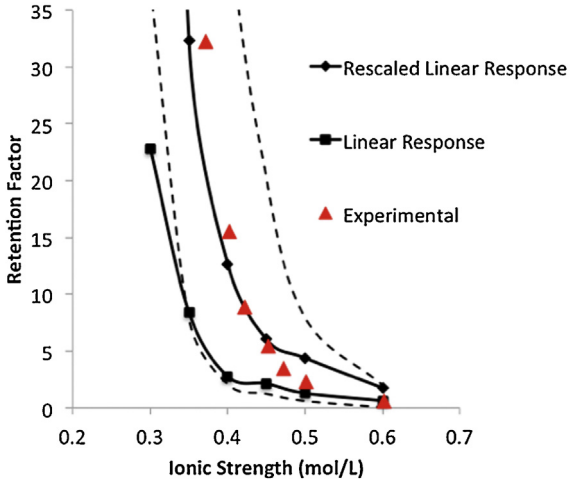

Fig. 10. Upper and lower limits (dashed lines) for the k^l retention factor calculated for Lysozyme at pH 4 increasing and decreasing the minimum interaction free energy by the standard deviation of the WHAM simulations (4 kJ kcal mol).

can be interpreted considering the breakdown of the total interaction free energy $\Delta G(\mathrm{pH},I,\lambda)$ between protein and surface into three different contributions: electrostatic at the minimum energy conformation ($\Delta G^{\mathrm{charge}}(\phi_{\mathrm{min}}, \nabla_{\mathrm{min}}\ \lambda)$), van der Waals, and conformational ($-RT\ \ln Q_{\mathrm{rot}}^{2D}(\lambda)/Q_{\mathrm{rot}}^{2D}(\infty)$). The three contributions to $\Delta G(\mathrm{pH},I,\lambda)$ calculated for Lysozyme at a pH of 4 and for a salt concentration of 0.5 mol/L are compared in Fig. 9 as a function of the distance between the IEX and protein COMs. Qualitatively similar results are obtained for Chymotrypsinogen A and for different pHs and ionic strengths.

The data of Fig. 9 show that, as expected, the interaction free energy is dominated by the electrostatic component, though van der Waals and conformational contributions can be quite important, especially at short distances between protein and surface.

The conformational free energy, which can also be interpreted as rotational entropy, contributes positively to the interaction free energy, thus decreasing the extent of the interaction, as in the adsorbed state the mobility of the protein is significantly limited with respect to the free state, where the rotational motion is unhindered. In order to account for the effect that the uncertainty in the estimation of the free energy discussed in Section 3.3 has on the calculated retention factor the retention factors for Lysozyme at pH 4 are computed increasing and decreasing the global interaction free energy by 4 kJ/mol, which is a reasonable uncertainty factor for the present calculations. The retention factor calculated without rescaling the interaction free energy of Lysozyme is reported as well to highlight the importance of accounting properly for the protein structural reorganization (Fig. 10).

3.5. Comparison between molecular and semiempirical models

In order to compare the predictive capabilities of semiempirical models with those of the two molecular models here used the experimental retention factors were fitted employing the analytical model recently proposed by Guelat et al. [21] This analytical model is based on the fundamental assumption of considering the protein as a spherical colloidal particle interacting with a charged plane. However representing a protein that possesses an inherently heterogeneous charges distribution with a homogeneously charged spherical colloid introduces a level of simplification that requires further verification.

The model implementation requires the estimation of several parameters, some of which can be determined from known physical properties while two, the density of positively and negatively charged protein groups, must be fitted either over experimental data or over the calculated retention factors. The main model parameters, their values, and their origin are summarized in Table S1 while the values of the fitted parameters and the extrapolated number of charged residues are reported in Tables S2 and S3, respectively. Due to the heterogeneous distribution of charged groups on the protein surface, their surface densities cannot be predicted directly from the protein amino acid sequences. Therefore, these parameters are fitted on the experimental retention factors following the same procedure described in Guelat et al. [21]. The experimental and fitted retention factors for Lysozyme and Chymotrypsinogen A on YMC BioPro SP-F are shown in Fig. 11. For both proteins, the fitted parameters (Table S3) are only in a semiquantitative agreement with those determined from the amino acid sequence. Even though the absolute numbers may be influenced by the many simplifications of the model (homogenous charge densities, spherical protein, etc.), the relative amount of each type of groups suggests the existence of a preferred interaction orientation for both proteins, similarly to what found in the molecular computations. An indirect indication that this may be the case comes from the observation that, in the case of Lysozyme, the ratio between carboxyl and amino groups in the amino acid sequence is 0.53, roughly corresponding to the global balance between negative and positive patches on the whole protein surface. This ratio is however only 0.38 for the semiempirical model. The underestimation of the ratio between negatively and positively charged groups indicates that the semiempirical model, which is based on the hypothesis that charges are homogeneously distributed on the protein surface, needs to fictitiously increase the positive charge of the protein in order to fit adequately the protein-IEX interaction energy as it is unable to account for the fact that, in reality, charged groups are distributed in patches. For Chymotrypsinogen A, the ratio between carboxyl and amino groups computed using the amino acid sequence is 0.71. However, for the semiempirical model, this ratio decreases to 0.37, which is very similar to Lysozyme. Also in this case a similar explanation to what found for Lysozyme applies.

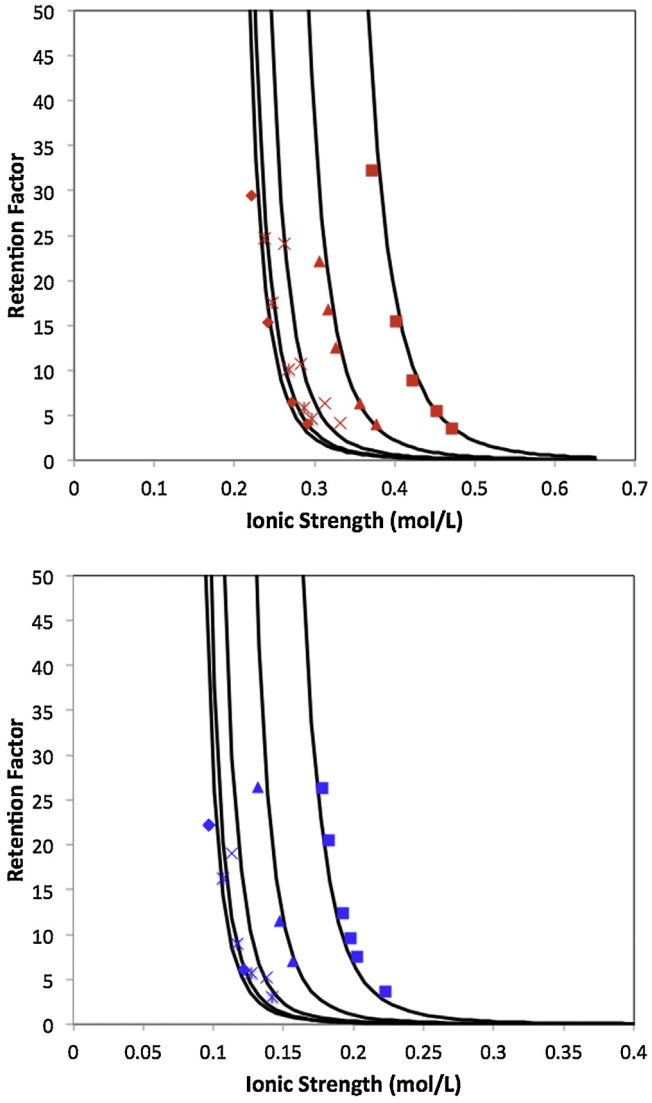

Fig. 11. Retention factor of Lysozyme (upper panel) and Chymotrypsinogen A (lower panel) on YMC BioPro SP-F as a function of the ionic strength of the buffer at pH 4, 4.5, 5, 5.5 and 6 (from right to left). Markers: experimental data; solid lines: fitted data.

Despite the shortcomings discussed above, the semiempirical model is able to fit well the experimental data reported in Fig. 11. This indicates that, after a selected set of protein retention factors has been computed using the more physically consistent models previously discussed and used to fit the semiempirical model parameters, retention factors can be predicted using the semiempirical model in a wider range of operative conditions at essentially no computational cost retaining a high level of accuracy.

As discussed at length, electrostatic interactions are dominating the adsorption energetics. In order to further compare the insight obtained from molecular modeling with the available experimental information a Stoichiometric Displacement Model (SDM) has been used to infer the effective number of protein charged groups interacting with the stationary phase. Details of the SDM have been reported in paragraph S6 of the SI. In the case of Lysozyme the SDM effective charge ranges from 7.25 (pH 6) to 9.25 (pH 4), while for Chymotrypsinogen A it adopts values in the interval between 5.65 (pH 6) and 8.82 (pH 4). For both proteins the effective charge is lower than the net charge of the protein in the same pH range. This observation is consistent with the existence of patches on the protein surface that are responsible for the interaction with the stationary phase. It is also interesting to notice that the number of charged groups belonging to those patches, identified through the analysis of the energy contour maps reported in Fig. 4, is of the same order of magnitude of the charges computed via SDM. In the case of Lysozyme, five residues (Lys 1, Arg 5, Lys 33, Arg 125, and Arg 128) have been identified as the constituents of a positively charged patch conferring a very localized nature to the protein–surface interaction. The comparison with the SDM results suggests that such localized patch is responsible for more than half of the effective charge, while the remaining half is probably due to smaller neighboring charged patches, characterizing the surface of the protein (Fig. S4a). In the case Chymotrypsinogen A, as shown in Fig. 4, the interaction is much less localized. In this case the largest interacting patch is spread on the protein surface (Fig. S4b) and is constituted by seven residues (Lys 36, Lys 90, Lys 93, Lys 84 Lys 87, Lys 107, and Lys 82), approximately accounting for the whole effective charge fitted through the SDM.

4. Conclusions

The purpose of the present work was to develop a molecular modeling approach that could be used to perform a priori estimates of protein retention factors for IEX materials in a wide range of conditions. It was found that the combination of MD simulations with a linear response model is able to describe well the dependence of the protein retention factor as a function of pH and ionic strength for the two proteins considered in this study: Chymotrypsinogen A and Lysozyme. The proposed model rests on two pillars. The first is the accurate determination of the interaction free energies using explicit MD simulations and the WHAM protocol at a specific pH and ionic strength. The second consists in the estimation of the change of the interaction free energy over pH and ionic strength treating the solvent implicitly and integrating the PBE using atomic charges for the proteins averaged over the possible protonation states. Notably, the conformational sampling and the study of the effect of the environment are performed using the implicit model, which allows saving considerable computational time. Also, it was found that the model computational efficiency may be increased if a semiempirical model is added to the computational protocol, provided that sufficient data, theoretical or experimental, are available to determine the model parameters.

The results of the molecular simulations provide some interesting insights into the investigated systems. It was in fact found that the interaction energy strongly depends on the orientation of the proteins relatively to the IEX surface. This is due to the fact that the protein surface charge density, which is heterogeneously spread on the surface and organized in patches, plays a major role in determining the interaction energy. Also, it was found that including in the model the effect of the structural relaxation of the protein as it approaches the surface is in one case important (Lysozyme), and in the other almost negligible (Chymotrypsinogen A). The different behavior may be due to the fact that Lysozyme exhibits a significantly higher density of charged residues on its surface than Chymotrypsinogen A. Also, the present study allows probably for the first time to quantify the effect that the consideration of the conformational rotational sampling has on the estimation of the interaction free energy between protein and surface. It was thus found that the loss of the rotational mobility of the protein in the adsorbed state decreases the interaction free energy by about 5 kJ/mol.

Finally, it should be observed that the IEX surface investigated in the present study is quite simple. While this was instrumental in the present work, which was focused on testing the reliability of the proposed computational approach, it would be interesting in a successive study to investigate whether the proposed approach may be suitable to study more complex IEX surfaces.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version.

References

[1] S. Hober, K. Nord, M. Linhult, A. Protein, chromatography for antibody purification, J. Chromatogr. B 848 (2007) 40–47.

[2] A.C.A. Roque, C.S.O. Silva, M.A. Taipa, Affinity-based methodologies and ligands for antibody purification: advances and perspectives, J. Chromatogr. A 1160 (2007) 44–55.

[3] A.C.A. Roque, C.R. Lowe, M.A. Taipa, Antibodies and genetically engineered related molecules: Production and purification, Biotechnol. Prog. 20 (2004) 639–654.

[4] A.C.A. Roque, C.R. Lowe, Advances and applications of de novo designed affinity ligands in proteomics, Biotechnol. Adv. 24 (2006) 17–26.

[5] A.R. Newcombe, C. Cresswell, S. Davies, K. Watson, G. Harris, K. O'Donovan, R. Francis, Optimised affinity purification of polyclonal antibodies from hyper immunised ovine serum using a synthetic Protein A adsorbent MAbsorbent (R) A2P, J. Chromatogr. B 814 (2005) 209–215.

[6] D. Low, R. O'Leary, N.S. Pujar, Future of antibody purification, J. Chromatogr. B 848 (2007) 48–63.

[7] A. Ljunglof, K.M. Lacki, J. Mueller, C. Harinarayan, R. van Reis, R. Fahrner, J.M. Van Alstine, Ion exchange chromatography of antibody fragments, Biotechnol. Bioeng. 96 (2007) 515–524.

[8] N. Forrer, A. Butte, M. Morbidelli, Chromatographic behavior of a polyclonal antibody mixture on a strong cation exchanger column. Part I: adsorption characterization, J. Chromatogr. A 1214 (2008) 59–70.

[9] N. Forrer, A. Butte, M. Morbidelli, Chromatographic behavior of a polyclonal antibody mixture on a strong cation exchanger column. Part II: Adsorption modelling, J. Chromatogr. A 1214 (2008) 71–80.

[10] T. Muller-Spath, L. Aumann, G. Strohlein, H. Kornmann, P. Valax, L. Delegrange, E. Charbaut, G. Baer, A. Lamproye, M. Johnck, M. Schulte, M. Morbidelli, Two step capture and purification of IgG2 using multicolumn countercur-rent solvent gradient purification (MCSGP), Biotechnol. Bioeng. 107 (2010) 974–984.

[11] D.K. Follman, R.L. Fahrner, Factorial screening of antibody purification processes using three chromatography steps without protein A, J. Chromatogr. A 1024 (2004) 79–85.

[12] C.M. Roth, J.E. Sader, A.M. Lenhoff, Electrostatic contribution to the energy and entropy of protein adsorption, J. Colloid Interface Sci. 203 (1998) 218–221.

[13] D. Asthagiri, A.M. Lenhoff, Influence of structural details in modeling electrostatically driven protein adsorption, Langmuir 13 (1997) 6761–6768.

[14] Y. Yao, A.M. Lenhoff, Electrostatic contributions to protein retention in ion-exchange chromatography 1. Cytochrome c variants, Anal. Chem. 76 (2004) 6743–6752.

[15] Y. Yao, A.M. Lenhoff, Electrostatic contributions to protein retention in ion-exchange chromatography 2. Proteins with various degrees of structural differences, Anal. Chem. 77 (2005) 2157–2165.

[16] J. Stahlberg, B. Jonsson, C. Horvath, Combined effect of coulombic and van der Waals interactions in the chromatography of proteins, Anal. Chem. 64 (1992) 3118–3124.

[17] J. Stahlberg, B. Jonsson, C. Horvath, Theory for electrostatic interaction chromatography of proteins, Anal. Chem. 63 (1991) 1867–1874.

[18] J. Stahlberg, B. Jonsson, Influence of charge regulation in electrostatic interaction chromatography of proteins, Anal. Chem. 68 (1996) 1536–1544.

[19] B. Jonsson, J. Stahlberg, The electrostatic interaction between a charged sphere and an oppositely charged planar surface and its application to protein adsorption, Colloids Surf. B 14 (1999) 67–75.

[20] W.R. Bowen, L.C. Pan, A.O. Sharif, Predicting equilibrium constants for ion exchange of proteins – a colloid science approach, Colloids Surf. A 143 (1998) 117–131.

[21] B. Guelat, G. Strohlein, 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.

[22] G.W. Slater, C. Holm, M.V. Chubynsky, H.W. de Haan, A. Dube, K. Grass, O.A. Hickey, C. Kingsburry, D. Sean, T.N. Shendruk, L.X. Nhan, Modeling the separation of macromolecules: a review of current computer simulation methods, Electrophoresis 30 (2009) 792–818.

[23] L. Zamolo, M. Salvalaglio, C. Cavallotti, B. Galarza, C. Sadler, S. Williams, S. Hofer, J. Horak, W. Lindner, Experimental and theoretical investigation of effect of spacer arm and support matrix of synthetic affinity chromatographic materials for the purification of monoclonal antibodies, J. Phys. Chem. B 114 (2010) 9367–9380.

[24] M. Salvalaglio, L. Zamolo, V. Busini, D. Moscatelli, C. Cavallotti, Molecular modeling of Protein A affinity chromatography, J. Chromatogr. A 1216 (2009) 8678–8686.

[25] D. Moiani, M. Salvalaglio, C. Cavallotti, A. Bujacz, I. Redzynia, G. Bujacz, F. Dinon, P. Pengo, G. Fassina, Structural characterization of a Protein A mimetic peptide dendrimer bound to human IgG, J. Phys. Chem. B 113 (2009) 16268–16275.

[26] C. Boi, V. Busini, M. Salvalaglio, C. Cavallotti, G.C. Sarti, Understanding ligand-protein interactions in affinity membrane chromatography for antibody purification, J. Chromatogr. A 1216 (2009) 8687–8696.

[27] C.J. Morrison, R. Godawat, S.A. McCallum, S. Garde, S.M. Cramer, Mechanistic studies of displacer-protein binding in chemically selective displacement systems using NMR and MD simulations, Biotechnol. Bioeng. 102 (2009) 1428–1437.

[28] Z.Q. Hu, J.W. Jiang, Separation of amino acids in glucose isomerase crystal: insight from molecular dynamics simulations, J. Chromatogr. A 1216 (2009) 5122–5129.

[29] R.B. Kasat, N.H.L. Wang, E.I. Franses, Experimental probing and modeling of key sorbent-solute interactions of norephedrine enantiomers with polysaccharide-based chiral stationary phases, J. Chromatogr. A 1190 (2008) 110–119.

[30] F.F. Liu, X.Y. Dong, T. Wang, Y. Sun, Rational design of peptide ligand for affinity chromatography of tissue-type plasminogen activator by the combination of docking and molecular dynamics simulations, J. Chromatogr. A 1175 (2007) 249–258.

[31] L. Zhang, D.N. Lu, Z. Liu, Dynamic control of protein conformation transition in chromatographic separation based on hydrophobic interactions. Molecular dynamics simulation, J. Chromatogr. A 1216 (2009) 2483–2490.

[32] C.F. Zhao, N.M. Cann, The docking of chiral epoxides on the Whelk-O1 stationary phase: A molecular dynamics study, J. Chromatogr. A 1149 (2007) 197–218.

[33] V. Busini, D. Moiani, D. Moscatelli, L. Zamolo, C. Cavallotti, Investigation of the influence of spacer arm on the structural evolution of affinity ligands supported on agarose, J. Phys. Chem. B 110 (2006) 23564–23577.

[34] L. Zamolo, V. Busini, D. Moiani, D. Moscatelli, C. Cavallotti, Molecular dynamic investigation of the interaction of supported affinity ligands with monoclonal antibodies, Biotechnol. Prog. 24 (2008) 527–539.

[35] X. Zhang, J.C. Wang, K.M. Lacki, A.I. Liapis, Molecular dynamics simulation studies of the conformation and lateral mobility of a charged adsorbate biomolecule: Implications for estimating the critical value of the radius of a pore in porous media, J. Colloid Interface Sci. 290 (2005) 373–382.

[36] X. Zhang, J.C. Wang, K.M. Lacki, A.I. Liapis, Construction by molecular dynamics modeling and simulations of the porous structures formed by dextran polymer chains attached on the surface of the pores of a base matrix: characterization of porous structures, J. Phys. Chem. B 109 (2005) 21028–21039.

[37] E. Riccardi, J.C. Wang, A.I. Liapis, A molecular dynamics study on the transport of a charged biomolecule in a polymeric adsorbent medium and its adsorption onto a charged ligand, J. Chem. Phys. 133 (2010) 084904.

[38] E. Riccardi, J.C. Wang, A.I. Liapis, Porous polymer adsorbent media constructed by molecular dynamics modeling and simulations: the immobilization of charged ligands and their effect on pore structure and local nonelectroneutrality, J. Phys. Chem. B 113 (2009) 2317–2327.

[39] W.K. Chung, S.T. Evans, A.S. Freed, J.J. Keba, Z.C. Baer, K. Rege, S.M. Cramer, Utilization of lysozyme charge ladders to examine the effects of protein surface charge distribution on binding affinity in ion exchange systems, Langmuir 26 (2010) 759–768.

[40] W.K. Chung, Y. Hou, A. Freed, M. Holstein, G.I. Makhatadze, S.N. Cramer, Investigation of protein binding affinity and preferred orientations in ion exchange systems using a homologous protein library, Biotechnol. Bioeng. 102 (2009) 869–881.

[41] K. Kubiak, P.A. Mulheran, Molecular dynamics simulations of hen egg white lysozyme adsorption at a charged solid surface, J. Phys. Chem. B 113 (2009) 12189–12200.

[42] F. Dismer, J. Hubbuch, 3D structure-based protein retention prediction for ion-exchange chromatography, J. Chromatogr. A 1217 (2010) 1343–1353.

[43] C.M. Roth, A.M. Lenhoff, Electrostatic and van der Waals contributions to protein adsorption – computation of equilibrium-constants, Langmuir 9 (1993) 962–972.

[44] P.M. Biesheuvel, M. van der Veen, W. Norde, A modified Poisson–Boltzmann model including charge regulation for the adsorption of ionizable polyelectrolytes to charged interfaces, applied to lysozyme adsorption on silica, J. Phys. Chem. B 109 (2005) 4172–4180.

[45] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174.

[46] J.W. Wang, M. Dauter, R. Alkire, A. Joachimiak, Z. Dauter, Triclinic lysozyme at 0.65 angstrom resolution, Acta Crystallogr. Sect. D: Biol. Crystallogr. 63 (2007) 1254–1268.

[47] D.C. Wang, W. Bode, R. Huber, Bovine Chymotrypsinogen-A X-ray crystal-structure analysis and refinement of a new crystal form at 1.8 a resolution, J. Mol. Biol. 185 (3) (1985) 595–624.

[48] Y. Duan, C. Wu, S. Chowdhury, M.C. Lee, G.M. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J.M. Wang, P. Kollman, A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations, J. Comput. Chem. 24 (2003) 1999–2012.

[49] W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, J. Chem. Phys. 79 (1983) 926.

[50] D.A. Case, T.A. Darden, T.E. Cheatham III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C.

Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, K.P.A. ollman, AMBER 12, University of California, San Francisco, 2012.
[51] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J.A. Peralta, J.E. Ogliaro, F. Bearpark, M. Heyd, J.J. Brothers, E. Kudin, K.N. Staroverov, V.N. Kobayashi, R. Normand, J. Raghavachari, K. Rendell, A. Burant, J.C. Iyengar, S.S. Tomasi, J. Cossi, M. Rega, N. Millam, N.J. Klene, M. Knox, J.E. Cross, J.B. Bakken, V. Adamo, C. Jaramillo, J. Gomperts, R. Stratmann, R.E. Yazyev, O. Austin, A.J. Cammi, R. Pomelli, C. Ochterski, J.W. Martin, R.L. Morokuma, K. Zakrzewski, V.G. Voth, G.A. Salvador, P. Dannenberg, J.J. Dapprich, S. Daniels, A.D. Farkas, Ö. Foresman, J.B. Ortiz, J.V. Cioslowski, J.D.J. Fox, Gaussian 09, Revision A.1, Gaussian, Inc., Wallingford CT, 2009.
[52] M. Salvalaglio, I. Muscionico, C. Cavallotti, Determination of energies and sites of binding of PFOA and PFOS to human serum albumin, J. Phys. Chem. B 114 (2010) 14860–14874.
[53] F. Zhu, G. Hummer, Convergence and error estimation in free energy calculations using the weighted histogram analysis method, J. Comput. Chem. 33 (2012) 453–465.
[54] Y.Y. Sham, Z.T. Chu, H. Tao, A. Warshel, Examining methods for calculations of binding free energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA calculations of ligands binding to an HIV protease, Proteins: Struct. Funct. Genet. 39 (2000) 393–407.
[55] J.M. Wang, P. Morin, W. Wang, P.A. Kollman, Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA, J. Am. Chem. Soc. 123 (2001) 5221–5230.
[56] F. Fogolari, A. Brigo, H. Molinari, The Poisson–Boltzmann equation for biomolecular electrostatics: a tool for structural biology, J. Mol. Recognit. 15 (2002) 377–392.
[57] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, R.J. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688.
[58] R.G. Gilbert, S.C. Smith, Theory of Unimolecular and Recombination Reactions, Blackwell Scientific Publications, Oxford, 1990.
[59] Y.E. Shapiro, E. Meirovitch, Evidence for domain motion in proteins affecting global diffusion properties: a nuclear magnetic resonance study, J. Phys. Chem. B 113 (2009) 7003–7011.
[60] J.F. Kincaid, H. Eyring, A.E. Stearn, The theory of absolute reaction rates and its application to viscosity and diffusion in the liquid state, Chem. Rev. 28 (1941) 301–365.
[61] H.C. Hamaker, The London-van der Waals attraction between spherical particles, Physica 4 (10) (1937) 1058–1072.
[62] M.E. Thrash Jr., N.G. Pinto, Incorporating water-release and lateral protein interactions in modeling equilibrium adsorption for ion-exchange chromatography, J. Chromatogr. A 1126 (2006) 304–310.
[63] C.M. Roth, B.L. Neal, A.M. Lenhoff, Van der Waals interactions involving proteins, Biophys. J. 70 (1996) 977–987.
[64] W.R. Bowen, L.-C. Pan, A.O. Sharif, Predicting equilibrium constants for ion exchange of proteins – a colloid science approach, Colloids Surf. A 143 (1998) 117–131.
[65] N. Baker, D. Sept, S. Joseph, M. Holst, J. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. U.S.A. 98 (2001) 10037–10041.
[66] H. Li, A.D. Robertson, J.H. Jensen, Very fast empirical prediction and rationalization of protein pK(a) values, Proteins: Struct. Funct. Bioinform. 61 (4) (2005) 704–721.
[67] D.C. Bas, D.M. Rogers, J.H. Jensen, Very fast prediction and rationalization of pK(a) values for protein-ligand complexes, Proteins: Struct. Funct. Bioinform. 73 (3) (2008) 765–783.
[68] M.H.M. Olsson, C.R. Sondergaard, M. Rostkowski, J.H. Jensen, PROPKA3: consistent treatment of internal and surface residues in empirical pK(a) predictions, J. Chem. Theory Comput. 7 (2) (2011) 525–537.
[69] A. Steudle, J. Pleiss, Modelling of lysozyme binding to a cation exchange surface at atomic detail: the role of flexibility, Biophys. J. 100 (12) (2011) 3016–3024.
[70] B. Bouvier, Decoding the patterns of ubiquitin recognition by ubiquitin-associated domains from free energy simulations, Phys. Chem. Chem. Phys. 16 (2014) 48–60.