Protein - surface interactions at the Nanoscale: atomistic simulations with implicit solvent models

David C. Malaspina[a], Leonor Pérez-Fuentes[b], Carlos Drummond[c], Delfi Bastos-González[b], Jordi Faraudo*[a]

[a] *Institut de Ciencia de Materials de Barcelona (ICMAB-CSIC), Campus de la UAB, E-08173 Spain*

[b] *Biocolloid and Fluid Physics Group, Department of Applied Physics, University of Granada, Av. Fuentenueva 2, E-18001 Granada (Spain)*

[c] *CNRS, Centre de Recherche Paul Pascal (CRPP), UMR 5031, Pessac, France*

Abstract

A full molecular-level understanding of protein adsorption in important situations such as the formation of protein films at solid/liquid interfaces or the formation of a protein corona over inorganic nanoparticles is still lacking. All-atomic implicit solvent molecular dynamics (MD) simulations, which are successfully employed in many related protein studies (such as protein folding for example) are emerging also as a useful tool to investigate proteins at surfaces. Implicit solvent simulations replace the detailed description of the solvent by a continuum media and effective atom-atom interactions retaining the atomistic detail in the description of the system of interest. This allows the simulation of larger systems and longer time scales as compared with full MD simulations including explicit solvent. In this brief review, we present an overview of the current state of the application of this technique to the study of problems such as the interaction of proteins with solid surfaces and the structure of protein corona over inorganic nanoparticles. Limitations of the approach and future perspectives are also outlined.

Keywords: Molecular Dynamics, Implicit solvent models, Protein surface interaction, Protein corona

Email address: jfaraudo@icmab.es (Jordi Faraudo*)

1. Introduction

Protein adsorption at interfaces has been described several times as a "common but very complicated phenomenon" [1, 2]. Different factors which are difficult to quantify such as the softness and hydrophobic or hydrophilic character of the protein or surface polarity for example are thought to play a decisive role in the properties of adsorbed protein layers and in biomaterials design [3]. In the bionanotechnology field, protein adsorption also has a prominent role. Nanoparticles (NP) in contact with biological fluids, are rapidly covered by a protein corona which determines the interactions and the biological identity of the material [4]. In fact, tailoring the protein corona is an essential step in drug delivery applications of NPs [5, 6].

Our ability to predict the behaviour and properties of proteins is growing rapidly due to the substantial increase in our knowledge of protein structure with atomistic resolution. For example, the Protein Data Bank [7, 8] contains the atomic coordinates of about 1.4×10^5 structures (and growing, at a rate of 10^4 structures per year), resolved by methods such as X-ray or NMR. In principle, these structures can be used, in combination with theoretical methods, to predict the interactions of proteins with materials and tackle the open questions related to protein adsorption, protein films and protein corona from a rational design perspective. To this end, a suitable computational tool is of course needed. A fitting candidate as a tool for performing such theoretical investigation is Molecular Dynamics (MD) simulations. Conceptually, the method is simple. It is based on the fact that at ordinary conditions (not under extreme temperatures or pressures), the motion of atoms of any molecular system can be computed implementing numerically the Newton laws of motion in a computer. The method of course needs an accurate description of the interaction forces between atoms (belonging to the same or different molecules), which can be done with good approximation using suitable, well-known force fields in many situations. The problem for the use of MD simulations for the study of protein-surface interactions is in the limitations of the method for dealing with large number of atoms and long time scales. For example, an atomistic model of a simple protein such as BSA has ≈9200 atoms (including hydrogen atoms). The protein corona of a 6 nm nanoparticle (NP) contains ≈ 10 BSA proteins [9], so the number of atoms required to model the proteins and the NP is $\approx 10^5$. The process of protein adsorption takes place always in water, so the addition of a water box large enough to have the proteins and NP in suspension in such a hypothetical simulation

will increase the number of atoms by an order of magnitude, making the simulation impossible. An additional limitation is related to time scales. MD simulations in explicit solvent are typically limited to time scales of the order of $\approx 10^2$ ns, which are too short to describe important processes such as protein diffusion or rearrangement at surfaces. For these reasons, the modelling of protein-surface interaction has been almost exclusively based on simplified models, which take into account the protein structure from a coarse-grained perspective. However, the development of implicit solvent force fields has allowed simulations of many complex systems involving proteins with full atomistic detail. In these models, water molecules are not included explicitly and the interaction of atoms from the protein are modified from classical force fields to include the effect of water in an implicit way [10, 11, 12, 13]. In this way, it is possible to reduce the number of atoms in the MD simulation to an affordable amount, with a concomitant reduction in the number of atom-atom interactions to be calculated in the simulation. It has to be noted that in implicit solvent all-atomic models, atom-atom interactions are more computationally expensive to calculate than in explicit solvent models [12, 13], but the overall balance is that simulations with implicit solvent models are (in general) substantially faster than explicit solvent simulations. Using these models, it has been possible to investigate with atomistic detail many interesting processes such as the structure of protein corona of a NP [14, 9] or how proteins adsorbed onto a surface respond to a pH change [15], among others, as we will see. Our objective in this article will be to summarize these developments. Of course, as any approach to a difficult problem, these simulations have their own difficulties and limitations, which we will also discuss.

The paper is organized as follows. In section 2, we will briefly discuss the conceptual aspects of Implicit solvent molecular dynamics simulations. In Section 3 we will discuss some basic results obtained from implicit solvent MD simulations of adsorption of a single protein onto a surface. In Section 4, we will discuss implicit solvent MD simulations involving the interaction between several proteins at a surface, such as protein films and protein corona of NPs. Finally, we will end up with Conclusions and Perspectives.

2. Implicit solvent models for all-atomic Molecular Dynamics simulations of proteins

In implicit solvent models, the solvent is replaced by a continuum dielectric medium and the expressions employed to compute the interactions between atoms are modified (from those employed with explicit water) to include the solvent in an effective way [10, 11, 13]. The first implicit solvent force field now known as Generalized Born Implicit Solvent model (GBIS, also sometimes abbreviated as GBSA to emphasize the inclusion of the effects due to the solvent accessible surface) was originally proposed in the 90s [10] to describe the interaction energy between molecules in solution for use in molecular dynamics simulations without the need to describe explicitly the solvent molecules. The main concept in the GBIS force field is to combine the concepts of classical force fields for explicit solvent MD simulations with the basic concepts of the Born theory of molecular solvation and the basic theory of the hydrophobic effect. In its original formulation [10], MD simulations employing the GBIS force field described with a very good approximation the solvation free energy of a wide range of small molecules, from inorganic ions to organic molecules. However, the theory produced inaccurate results in the case of molecules with large interior regions with many atoms buried inside, as in the case of proteins [16]. Over the years, refinements in the calculation of the dielectric screening [17] and electrostatics near the surfaces of the atoms [11] provided the necessary improvements and good results were obtained when used with MD of proteins. The accuracy of the current versions of the GBIS force field for MD simulations of biomolecules is discussed in many reviews (see for example [18, 19]). Nowadays, implementations of GBIS force fields for implicit solvent MD simulations are available in standard MD codes such as AMBER [11], NAMD [12, 13] or GROMACS [20].

The GBIS implicit solvent force field, as implemented in the MD codes mentioned above, describe the atoms with the same atomic partial charges and Lennard-Jones atomic parameters as employed in the explicit solvent force field for the calculation of electrostatic and van der Waals interactions. However, two substantial modifications are made in the energy and force equations of the force field. The first modification is that electrostatic interactions are calculated using a modified expression (instead of a direct Coulomb's law) which takes into account the degree of exposure of each atom to the solvent [10, 11]. And the second modification is that hydrophobic interactions are included as an additional term in the force field which

is proportional to the exposed hydrophobic surface and an effective surface tension of the molecule-solvent interface [10, 13].

The modified expression for the electrostatic interaction has a more complex spatial dependence than Coulomb's law, and it contains as a fundamental quantity determining the length scale of the interactions, the so-called Born radii. Each atom in a molecule is represented by a sphere filled uniformly with material of dielectric constant ϵ_r=1 [11]. The exterior of the atom is filled with a continuum medium with an effective dielectric constant of the solvent which also depends on the implicit ion concentration [13]. The size of the atom is given by the Born radius which describes the exposure of a given atom to the solvent (thus determining the degree of screening of electrostatic interaction by the solvent). The Born radius depends on the distance to neighbouring atoms. For an isolated atom, its Born radius is equal to its van der Waals radius [21], while for a deeply buried atom, its Born radius is much larger than its van der Waals radius. Depending on the specific implementation of the GBIS force field, the Born radius is calculated by the Onufriev-Bashford-Case (OBC) method [11] or by the Hawkins-Cramer-Truhlar method [17].

We recall here that the generalized Born implicit solvent GBIS model described above gives an accurate description of the polar contribution to the energy of solvation. The second modification of the force field for implicit solvent accounts for the nonpolar (i.e. hydrophobic) contribution to the energy of solvation. As we said, the model assumes that the nonpolar, hydrophobic solvation energy is proportional to the exposed hydrophobic surface, with a surface tension ≈ 0.005 kcal/mol Å^2 [13] which is a typical value for hydrocarbons in water. This is calculated by computing the surface exposed by each atom weighted by an atom-type weight.

Using an implicit solvent force field, one reduces greatly the number of atoms in the simulation and more importantly the number of atom pair interactions to be computed. But the price to pay is an increased computational cost of the calculations. The increased computational time has two sources. First, in the GBIS model, it is not possible to use the Particle Mesh Ewald (PME) method to speed up electrostatic calculations [12, 13], as it is usually done with Coulomb interactions in explicit solvent MD simulations. The only option in GBIS force fields is to compute long-range screened electrostatics with a large cut off. The second source of increased computational cost is related to the complexity of the GBIS equations (which require also the calculation of the Born radius which depends on the distribution of atoms).

Typical estimations suggest that electrostatic GBIS calculations are typically 7 times more expensive than PME electrostatic calculations [13].

The actual source of speed up of implicit solvent MD simulations is due to two factors: a) the absence of viscosity (which induces faster speed in relaxation processes) and b) absence of slow processes involving explicit water molecules. Therefore, the comparison between the actual gain in implicit solvent MD depends on the amount of water required in the simulation box to perform the explicit solvent simulation of the same problem. This balance is nicely illustrated by the calculations reported in Ref. [22]. For example, an implicit solvent MD simulation of a 25,100-atom model of a nucleosome is nominally 1.6 times slower than the explicit-solvent PME simulation with a small water solvent box extending 1 nm from the solute, whereas it is 1.6 times faster compared to a PME simulation in a larger 3.6 nm water solvent box.

Another important point, in order to compare implicit solvent and explicit solvent MD simulations is that nominal simulation time (this is, the time step in real units multiplied by the number of steps) has different meanings. In an explicit solvent MD simulation, the nominal simulation time is expected to correspond to the clock time elapsed in the real system in an equivalent experimental situation. But in the case of implicit solvent models, the absence of explicit water molecules and the absence of viscosity implies that the exploration of the configuration space takes place at a higher speed than in a real situation. In other words, for each ns of simulation, a protein will explore more possible configurations in an implicit solvent MD simulation than in an explicit solvent MD simulation. Therefore, the quantification of the speed up obtained in implicit solvent models is not a trivial task. This question has been investigated in detail in Ref [22] by comparing the speed of five different conformational changes in implicit solvent and explicit solvent MD simulations (a dihedral angle flip, nucleosome histone tail collapse, DNA unwrapping from the nucleosome histone tail collapse, DNA unwrapping from the nucleosome histone core and miniprotein folding). The authors compare not only the differences in nominal simulation times for explicit and implicit water but also estimate the time that will correspond to a explicit water simulation to explore the configurations obtained in the trajectories of the implicit model MD simulations. Taking into account all these factors, the speed up of the conformational change in implicit solvent was shown to be strongly depended on the particular problem. In the case of a dihedral angle flip, the speed up is only by a factor of 1.6 whereas that in the case of

miniprotein folding the speed up is by a factor of ≈54.

Concerning the accuracy of implicit solvent models, comparison with explicit solvent models [11, 12, 13, 22] reveals that implicit solvent MD simulations predict correctly protein configuration in bulk conditions, as compared with explicit solvent simulations and experimental data. Now the question is the performance of the implicit solvent models in the case of the interaction of proteins with surfaces.

3. MD simulations of the adsorption of a single protein

Once the validity of the implicit solvent MD simulations of proteins has been established (see Section 2), the next step is to consider the adsorption of a single protein onto a surface. Albeit simple, there are many interesting questions that have been studied from such simulations, as illustrated in Figure 1. Some of these questions are the identification of the specific amino acids involved in the adsorption of a given protein at a given surface (and thus the driving force for adsorption), structural changes of proteins after adsorption, the effect of changing conditions (such as pH) after protein adsorption or the effect of curvature in adsorption (adsorption onto a planar surface compared with adsorption onto small nanoparticles).

We will discuss now some representative results related to these questions, by comparing the results obtained for different proteins and different substrates. In Figure 1, we summarize the results obtained in the case of four different systems: (a) the adsorption of β-Lactoglobulin onto a planar metal surface [23], (b) the adsorption of β-casein onto a generic model of a planar hydrophobic surface, (c) the adsorption of ubiquitin onto a 10 nm diameter Ag nanoparticle and (d) the adsorption of a BSA protein onto a 6 nm diameter nanoparticle.

In all these cases, the implicit solvent simulations allowed to identify the specific amino acids or the specific protein domain involved in adsorption, as shown in Figure 1. One interesting result obtained in some of these simulations is that the residues involved in protein adsorption -for a given protein- are essentially the same in different situations (such as for charged or neutral surfaces or at different pH). For example, in the case of bovine β-Lactoglobulin [23] adsorbed onto a gold surface, it was found that the residues involved in the adsorption were the same for a neutral surface or for surfaces with different values of positive charge. The protein residues adsorbing at the neutral Au surface are the hydrophilic Thr 125, Thr 18, Lys 100

and Gln 13 residues and the Pro 50 hydrophobic residue (see Figure 1a). In the simulations, the effect of charging the surface with positive charge density was also investigated. When charging the surface, some residues from 125 to 135 (which are close to the surface in the neutral surface case) also change its charge, leading to stronger adsorption by attractive electrostatic residue-surface interactions.

Analogous results were obtained in [15] for adsorption of β-casein onto model hydrophobic surfaces of different charges (Figure 1b). At neutral pH, the negatively charged protein (with a charge of -6e) adsorbs onto a neutral hydrophobic surface by contact with several, mostly hydrophobic residues but also by some polar residues situated near these hydrophobic ones. The residues involved in the contact between the protein and the neutral surface were six hydrophobic aminoacids (Pro 76, 78, 194, 196, 201 and Phe 205), and three polar aminoacids (Tyr 75, Val 193 and Ile 202), all indicated in Figure 1b). No significant change in the adsorption of the protein (same adsorbed residues and same area of contact) was observed when changing the charge of the surface from neutral to negative (-0.62 e/nm^2), in spite of the electrostatic repulsion (recall that both the surface and the protein were negatively charged). Finally, the charge of the adsorbed protein was changed from -6e to +8e, corresponding to a pH change from 7 to 4 (the charge of the surface was maintained unchanged at -0.62 e/nm^2). In this case, the adsorbed protein increases the contact with the surface, by adsorbing not only with the same residues as in the case of neutral pH but also with other protonated, positively charged residues located near the previously adsorbed domain.

In addition to these studies over planar surfaces, there are also a few simulations [9, 14] that study similar questions in the case of adsorption of proteins over NPs. These studies are more difficult than simulations over a planar surface (because they involve the simulation of the full NP) and are typically restricted to small NPs, with sizes of 5-10 nm.

In Ref. [14], the authors studied the adsorption of ubiquitin over a 10 nm Ag NP. They found that the protein adsorbs by binding of the Glu 18 residue to the NP (see Figure 1c) and also to some extend with an interaction of the Gln 8 residue with the NP. No significant change in the secondary structure of the protein was found due to adsorption. These simulation results are in agreement with NMR studies [24] which found that the binding domain of human ubiquitin to Ag NP was located in the residues Gln 2-Ile 3 and Leu 15-Glu 18 . Concerning the kinetics of adsorption, they obtained that

protein re-orientation was the rate-limiting step in protein adsorption.

In Ref. [9], the authors studied experimentally and by MD implicit solvent simulations the adsorption of BSA onto a 6 nm iron oxide NP covered by citrate. The simulations show that as the protein contacts the NP, the BSA protein changes its conformation and the protein domain in contact with the NP partially spreads over the surface of the NP (see Figure 1d). In fact, the surface of the protein increased from 325 nm^2 in solution to 345 nm^2 at the NP surface. The change in the secondary structure of the protein was small; the α-helix content changed from 72% in solution to 66% for a BSA adsorbed onto a NP. However, a few residues changed significantly its environment. It was particularly interesting the case of the Trp213 and Trp134 residues which contribute to the UV-vis spectra of the BSA protein. Simulations also indicate a change of their environment from being buried into the protein to a more solvent exposed location, a result in agreement with UV-vis spectroscopy measurements [9].

All the simulation results discussed so far of MD simulations with implicit solvent indicate a small change in the conformation of a protein after adsorption, with the changes localized at a few particular residues. Probably the most significant exception is the case of adsorption of proteins at carbon surfaces (graphite or graphene) in which several works using implicit solvent MD simulations report substantial unfolding of proteins or protein fragments over these carbon surfaces. Examples include unfolding of albumin and fibronectin fragments after adsorption onto graphite and carbon nanotubes [25], unfolding of lysozyme on graphite [26] and unfolding of a BMP-2 protein onto graphite [27]. Interestingly, the results for unfolding of BMP-2 protein onto graphite [28] were further confirmed by MD simulations in explicit solvent using accelerated MD simulations. Other explicit solvent MD simulations of proteins and peptides onto carbon surfaces indicate similar conclusions. Explicit solvent MD simulations of two "de novo" designed α-helical peptides [29] show that they unfold and assemble into an amorphous dimer at a graphene surface, in agreement with circular dichroism spectroscopy and scanning tunneling microscopy measurements. It seems that the specific role of the carbon surface is important in these results. For example, extensive explicit solvent simulations of lysozyme adsorption/desorption on polymeric hydrophobic surfaces [30, 31] only show small secondary structure changes in particular residues (typically those located near the adsorption region), but they do not show unfolding of the protein. Interestingly, in these works it was found that there is an energy barrier for adsorption mainly arising

from protein and surface hydration in the first hydration shell [30, 31]. The protein cannot be adsorbed even though the substrate surface exhibits attraction, whenever the protein-surface interaction energy is not large enough to overcome the barrier of breaking the hydration shells, which is strongly residue and surface dependent.

The question of the specific role of carbon surfaces in protein unfolding and the validity of its modelling by using implicit solvent MD simulations was analysed in detail in Ref. [32]. First, the authors repeated and extended previous studies [33] which reported unfolding of BSA after adsorption onto graphene. They found large unfolding of BSA upon adsorption onto a graphene surface, with an α-helix content of only 35% after free adsorption onto graphene. The authors also reported extensive all-atomic MD simulations of BSA adsorption onto graphene with explicit solvent and in this case the α-helix content was about 60% after free adsorption. Therefore, in this case the unfolding of BSA was observed only in the case of implicit solvent MD simulations. This discrepancy probably indicates that the employed force field for carbon surfaces greatly exaggerates the strength of the interaction of BSA with graphene, inducing a BSA unfolding as a simulation artefact. We think that better, thoroughly validated force fields for implicit solvent simulations of carbon surfaces are needed for implicit solvent simulations of protein adsorption. As a general consequence, we can say that more attention has to be paid to implicit solvent models of surfaces for use in implicit solvent simulations involving biomolecular adsorption.

Another interesting question to discuss, in light of simulation results, is the reversibility or irreversibility of protein adsorption. In all the simulations discussed so far, only adsorption (with no desorption) was reported, indicating that protein adsorption is essentially irreversible at the time scales probed by atomistic MD simulations. This question has been discussed in a recent all-atomic MD study (with explicit solvent) of the energy landscape for BSA adsorption onto silica [34]. The results indicate that the time scales for protein desorption are of orders of hours in this case, whereas time scales for protein diffusion are in the 100 ns range. Therefore, protein desorption is far beyond the scales accessible to MD (even with implicit solvent models) whereas protein diffusion should be observable, if present, by all-atomic MD simulations.

4. MD Simulations of adsorption of many proteins: protein films and protein corona

Most of the simulation results published in the literature correspond to the situation considered in Section 3, this is, the interaction of a single protein with a planar surface or a nanoparticle, which correspond to studies of the protein-surface interaction.

Of course, as the coverage of adsorbed protein increases, protein-protein interactions became more and more important, but these situations are computationally expensive because they require the consideration of many proteins at the surface. The cost of the simulations increases not only because the number of atoms in the simulation increases but also because the dynamics of the problem itself is much slower. This makes the simulation of high coverages prohibitively expensive in many cases. For this reason, only recently simulations considered atomistic simulations of protein films or protein corona in which protein-protein interactions are essential. In spite of these difficulties, a few of the studies discussed in the previous section consider not only the case of a single protein adsorption but also the formation of an adsorbed protein layer with atomistic detail, which we will now discuss.

In addition to the study of single β-casein protein adsorption (discussed in Section 3, see Figure 1b), this work also considered simulations with two and three adsorbed proteins onto a surface of 45.5 nm^2. These simulations with two and three proteins correspond to a mass coverage of 1.74 mg/m^2, and 2.6 mg/m^2 which are about 47% and 70% of the maximum experimental coverage for a β-casein monolayer [15]. Comparing the simulation results at the different coverages (see Figure 2a), it was observed that the thickness of the film decreases from $\approx$ 4 nm at the lowest coverage to $\approx$ 3 nm at the highest coverage, indicating that at higher coverages the proteins tend to be adsorbed in a more compact configuration. As shown in the snapshot in Figure 2b, the adsorbed proteins are in contact, forming a compact structure. Hence, the results indicate a substantial protein-protein attraction in spite of the electrostatic repulsion between proteins (each protein has a charge of -8e in the simulation, corresponding to pH=7).

The case of the maximum possible coverage was considered in the case of the formation of a BSA protein corona onto a 6 nm diameter nanoparticle studied in Ref. [9] (see Figure 2c). In that case, the simulations were employed to determine the number of proteins in the corona, their organization and to identify possible secondary structure differences between proteins in

the corona or free proteins. It should be noted also that the BSA protein is commonly employed in experiments as a cheaper alternative to experimentation with the human serum albumina (HSA) protein, which is more relevant for biomedical applications. For this reason, we have also repeated the analysis in [9] using HSA instead of BSA [35] and we have also considered two different particle sizes (6 nm and 3 nm of diameter). Starting from the simulation with a single adsorbed protein, the simulations of the protein corona were done by systematically adding more proteins in solution to a previous simulation (see [9]) with a smaller number of proteins and performing long simulation runs in order to allow for structural relaxations and adsorption of additional proteins [9]. It should be noted that the total number of atoms in these simulations with many proteins is $\approx 10^5$ atoms (for a 10 protein simulation) while the same system in explicit solvent would have more than 1 million of atoms in the simulation box. The simulations indicate that the maximum number of BSA or HSA proteins that can be accommodated onto the 6 nm NP is of 10 proteins (see Figure 2b), a result which is in agreement with experiments [9] (in which the number of BSA at the corona was estimated from the change in size of the particle before and after protein corona formation). It is also interesting to recall that the maximum number of proteins considered in the simulations was 12 and that in these simulations with excess protein, 2 of the proteins remain in bulk without adsorbing and without interacting significantly with the layer of adsorbed proteins [9]. This result also suggests that these proteins form only a monolayer over the NP, since a soft corona (a second layer of protein adsorbed onto the protein corona) was not observed in our simulations.

The size of the NP has a deep impact in the size and organization of the protein corona, as can be seen in Figure 2c. In the case of a 3 nm NP, we obtain a protein corona of only 3 HSA proteins which are clearly separated (i.e., they are not in contact) due to the curvature of the surface. This has to be compared with the compact structure made by the 10 proteins adsorbed onto a 6 nm NP (also note here that a decrease of the surface by a factor of 4 involved a decrease in the number of proteins by a factor of 3.3).

In Figure 2, we also show structural details that can be obtain from data analysis of the simulations. In figure 2d, we compare the α-carbon radial density data from Ref. [9] for BSA or HSA protein corona [35] respectively of a 6 nm NP. The results were very similar in both cases. There is a high density peak at the NP surface and a constant density region (the compact corona) that extends up to a distance of 6.5 nm of the centre of the NP. This

corresponds to a diameter of the NP+protein corona entity of about 13 nm, in agreement with DLS measurements with BSA [9].

As we discussed before, implicit solvent models can provide accurate representation of protein secondary structure, so we can discuss the secondary structure of the proteins in the protein corona. Figure 2c shows the α-helix content for each of the HSA or BSA proteins adsorbed on the protein corona, compared with a protein in bulk solution. The results show that the changes in secondary structure are small in all cases, but there is also a tendency indicating that these changes are smaller for the last adsorbed proteins. The explanation for this effect is that at low coverages, there is more space available for the spreading of the protein over the surface (recall here previous section and Figure 1b). On the contrary, near saturation, the last adsorbed protein has a small surface available for adsorption and a small contact with the NP, without possibility of spreading over the surface.

5. Conclusions and outlook

Implicit solvent models for protein simulation are an interesting alternative for simulations of proteins retaining all-atomic details with affordable resources. Their use in the study of problems involving protein adsorption onto surfaces, such as protein films or the protein corona of NPs is a young and promising approach that has delivered several interesting insights but still has drawbacks that need to be tackled. We can summarize the accomplishments and shortcomings of MD atomistic simulations of proteins at surfaces using implicit solvent GB models as follows:

- The method can be used for atomistic simulation of protein films over planar surfaces under different conditions (e.g. different pH) even in the case of protein and surface charged with charges of the same sign.

- The method can be used for atomistic simulation of protein corona over small NPs (< 10 nm). Predictions of the size of protein corona are in agreement with experimental results. The method is currently the only feasible option to study, with atomistic detail, the structure of a protein corona and its interactions with their environment.

- Predictions of secondary structure changes after adsorption are in general in agreement with experimental data.

- The model of the surface must be consistent with the implicit solvent model employed for the description of the protein. This is in general a nontrivial question, so some kind of validation of the results (by comparing with reference simulations with explicit solvent or with experimental results) is advisable.

- The role of ions is considered implicitly, at a Poisson-Boltzmann level of description. This is clearly not enough for describing the complex specific effects of ions with proteins, which are known to play a substantial role in protein films [36].

- Charge regulation effects taking place during protein adsorption (protonation or deprotonation of charged groups at the surface and/or protonation or deprotonation of protein residues near the surface) are important, but not considered explicitly in MD simulations. Currently, these effects are considered implicitly by performing simulations at different protonation states to mimic the effect of pH.

As we have discussed in this review, the direction of the field in the last years has been clearly focused on the implementation of the method in the software usually employed in MD simulations and to speed up the implementations of the method (by adding features such as the use of GPU). We are sure that these implementations of the method will fuel exciting new uses of the method to study more complex problems. But before more applications of greater complexity can be considered, it will be of great interest to advance and improve the theory. We think that, in view of the points listed above, the most pressing questions to be tackled are the development of more accurate implicit solvent models for nanostructured surfaces of interest and also the development of consistent models for the explicit inclusion of ions (and their specific effects) and the inclusion of charge regulation in implicit solvent MD simulations.

Acknowledgements

We acknowledge financial support from the Spanish Government grants MAT2015-64442-R, FIS2016- 80087-C2-1-P and SEV-2015-0496 and from the Junta de Andalucia Grant CTS-6270. D.C.M. is supported by the European Union's Horizon 2020 research and innovation programme under Marie Sklodowska-Curie grant agreement No 6655919.

References

[1] K. Nakanishi, T. Sakiyama, K. Imamura, On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon, Journal of Bioscience and Bioengineering 91 (2001) 233–244.

[2] M. Rabe, D. Verdes, S. Seeger, Understanding protein adsorption phenomena at solid surfaces, Advances in Colloid and Interface Science 162 (2011) 87–106.

[3] M. M. Ouberai, K. Xu, M. E. Welland, Effect of the interplay between protein and surface on the properties of adsorbed protein layers, Biomaterials 35 (2014) 6157–6163.

[4] M. P. Monopoli, C. Åberg, A. Salvati, K. A. Dawson, Biomolecular coronas provide the biological identity of nanosized materials, Nature Nanotechnology 7 (2012) 779–786.

[5] J. Mariam, S. Sivakami, P. M. Dongre, Albumin corona on nanoparticles a strategic approach in drug delivery, Drug Delivery (2015) 1–9.

[6] V. H. Nguyen, B.-J. Lee, Protein corona: a new approach for nanomedicine design, International Journal of Nanomedicine Volume 12 (2017) 3137–3151.

[7] H. Berman, K. Henrick, H. Nakamura, Announcing the worldwide Protein Data Bank, Nature Structural Biology 10 (2003) 980.

[8] H. Berman, K. Henrick, H. Nakamura, J. L. Markley, The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data, Nucleic Acids Research 35 (2007) D301–D303.

[9] S. Yu, A. Perálvarez-Marín, C. Minelli, J. Faraudo, A. Roig, A. Laromaine, Albumin-coated SPIONs: an experimental and theoretical evaluation of protein conformation, binding affinity and competition with serum proteins, Nanoscale 8 (2016) 14393–14405. [** This paper provides direct comparison between experimental data in protein corona and implicit solvent MD simulations.].

[10] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, Semianalytical treatment of solvation for molecular mechanics and dynamics, Journal

of the American Chemical Society 112 (1990) 6127–6129. [** Classical article founder of the implicit solvent method].

[11] A. Onufriev, D. Bashford, D. A. Case, Exploring protein native states and large-scale conformational changes with a modified generalized born model, Proteins: Structure, Function, and Bioinformatics 55 (2004) 383–394. [* Improvement in the implementation of Generalized Born algorithm and test on AMBER force field].

[12] D. E. Tanner, K.-Y. Chan, J. C. Phillips, K. Schulten, Parallel Generalized Born Implicit Solvent Calculations with NAMD, Journal of Chemical Theory and Computation 7 (2011) 3635–3642.

[13] D. E. Tanner, J. C. Phillips, K. Schulten, GPU/CPU Algorithm for Generalized Born/Solvent-Accessible Surface Area Implicit Solvent Calculations, J. Chem. Theory Comput. 8 (2012) 2521–2530. [* Implementation and test of fast algorithm for implicit solvent molecular dynamics simulations].

[14] F. Ding, S. Radic, R. Chen, P. Chen, N. K. Geitner, J. M. Brown, P. C. Ke, Direct observation of a single nanoparticle-ubiquitin corona formation, Nanoscale 5 (2013) 9162–9169.

[15] L. Pérez-Fuentes, C. Drummond, J. Faraudo, D. Bastos-González, Adsorption of milk proteins (beta-casein and beta-lactoglobulin) and BSA onto hydrophobic surfaces, Materials 10 (2017) 893. [* Includes an implicit solvent simulation study of effect of pH and surface charge in protein adsorption].

[16] J. Srinivasan, M. W. Trevathan, P. Beroza, D. A. Case, Application of a pairwise generalized Born model to proteins and nucleic acids: inclusion of salt effects, Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta) 101 (1999) 426–434.

[17] G. D. Hawkins, C. J. Cramer, D. G. Truhlar, Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium, The Journal of Physical Chemistry 100 (1996) 19824–19839.

[18] J. Chen, C. L. Brooks, J. Khandogin, Recent advances in implicit solvent-based methods for biomolecular simulations, Current Opinion in Structural Biology 18 (2008) 140–148.

[19] J. Kleinjung, F. Fraternali, Design and application of implicit solvent models in biomolecular simulations, Current Opinion in Structural Biology 25 (2014) 126–134.

[20] P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess, E. Lindahl, Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models, Journal of Chemical Theory and Computation 6 (2010) 459–466.

[21] A. Bondi, van der Waals Volumes and Radii, The Journal of Physical Chemistry 68 (1964) 441–451.

[22] R. Anandakrishnan, A. Drozdetski, R. C. Walker, A. V. Onufriev, Speed of Conformational Change: Comparing Explicit and Implicit Solvent Molecular Dynamics Simulations, Biophysical Journal 108 (2015) 1153–1164. [** Detailed study on the interpretation of time scales in implicit solvent MD simulations.].

[23] T. Hagiwara, T. Sakiyama, H. Watanabe, Molecular Simulation of Bovine beta-Lactoglobulin Adsorbed onto a Positively Charged Solid Surface, Langmuir 25 (2009) 226–234.

[24] L. Calzolai, F. Franchini, D. Gilliland, F. Rossi, ProteinNanoparticle Interaction: Identification of the UbiquitinGold Nanoparticle Interaction Site, Nano Letters 10 (2010) 3101–3105.

[25] G. Raffaini, F. Ganazzoli, Understanding the Performance of Biomaterials through Molecular Modeling: Crossing the Bridge between their Intrinsic Properties and the Surface Adsorption of Proteins, Macromolecular Bioscience 7 (2007) 552–566.

[26] G. Raffaini, F. Ganazzoli, Protein Adsorption on a Hydrophobic Surface: A Molecular Dynamics Study of Lysozyme on Graphite, Langmuir 26 (2010) 5679–5689.

[27] C. Mücksch, H. M. Urbassek, Adsorption of BMP-2 on a hydrophobic graphite surface: A molecular dynamics study, Chemical Physics Letters 510 (2011) 252–256.

[28] C. Mücksch, H. M. Urbassek, Enhancing Protein Adsorption Simulations by Using Accelerated Molecular Dynamics, PLoS ONE 8 (2013) e64883+.

[29] L. Ou, Y. Luo, G. Wei, Atomic-Level Study of Adsorption, Conformational Change, and Dimerization of an -Helical Peptide at Graphene Surface, The Journal of Physical Chemistry B 115 (2011) 9813–9822.

[30] T. Wei, M. A. Carignano, I. Szleifer, Lysozyme Adsorption on Polyethylene Surfaces: Why Are Long Simulations Needed?, Langmuir 27 (2011) 12074–12081.

[31] T. Wei, M. A. Carignano, I. Szleifer, Molecular Dynamics Simulation of Lysozyme Adsorption/Desorption on Hydrophobic Surfaces, The Journal of Physical Chemistry B 116 (2012) 10189–10194.

[32] J. G. Vilhena, P. Rubio-Pereda, P. Vellosillo, P. A. Serena, R. Pérez, Albumin (BSA) Adsorption over Graphene in Aqueous Environment: Influence of Orientation, Adsorption Protocol, and Solvent Treatment, Langmuir 32 (2016) 1742–1755. [* Comparison of implicit and explicit solvent molecular dynamics simulations].

[33] C. Mücksch, H. M. Urbassek, Molecular Dynamics Simulation of Free and Forced BSA Adsorption on a Hydrophobic Graphite Surface, Langmuir 27 (2011) 12938–12943.

[34] K. Tokarczyk, K. Kubiak-Ossowska, B. Jachimska, P. A. Mulheran, Energy Landscape of Negatively Charged BSA Adsorbed on a Negatively Charged Silica Surface, The Journal of Physical Chemistry B 122 (2018) 3744–3753.

[35] D. C. Malaspina, J. Faraudo, Molecular dynamics study of human serum albumin protein corona in an inorganic nanoparticle (in preparation).

[36] L. Pérez-Fuentes, C. Drummond, J. Faraudo, D. Bastos-González, Interaction of organic ions with proteins, Soft Matter 13 (2017) 1120–1131.

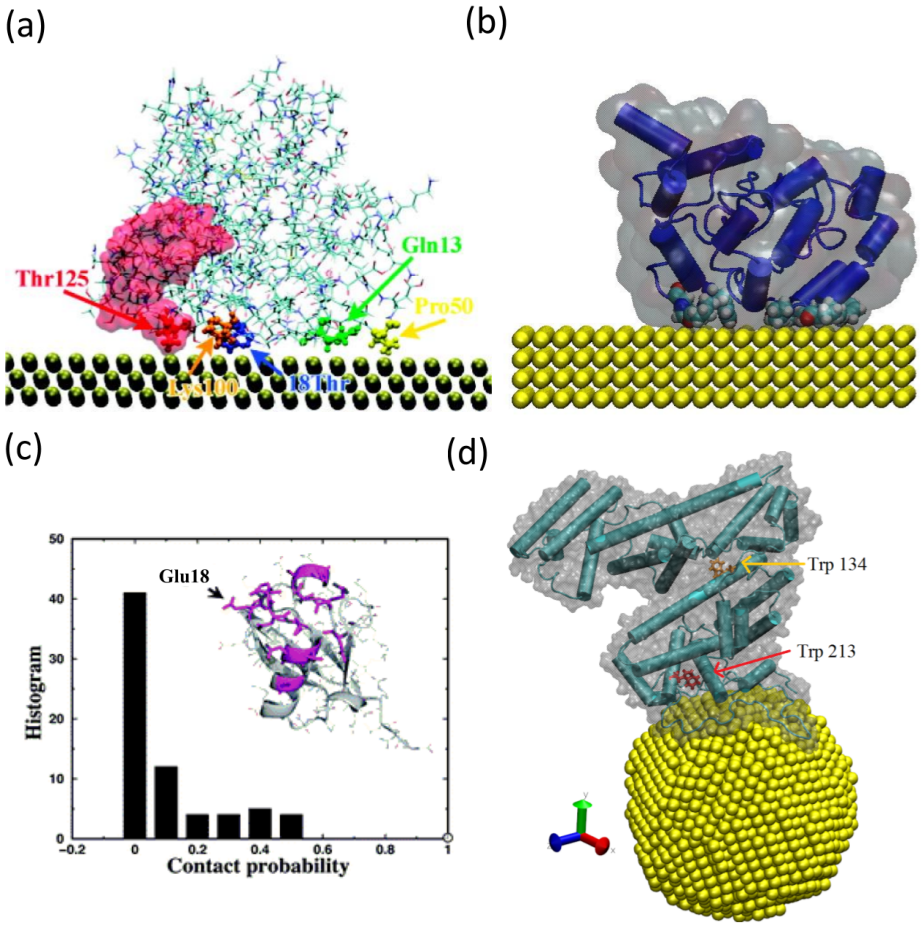

Figure 1: Examples of results from implicit solvent all-atomic MD simulations of protein adsorption (a) Adsorption of β-Lactoglobulin onto an Au(100) surface, with indication of the residues involved in protein adsorption (reproduced from [23] with permission), (b) Adsorption of β-casein onto a generic model of a planar hydrophobic surface (redrawn with data from [15]). The protein is shown in cartoon representation with indication of its size by a glassy surface. The residues involved in adsorption are shown in van der Waals representation. (c) Simulation of adsorption of human ubiquitin over a 10 nm Ag NP. The histogram indicates the number of residues with a given adsorption probability (adapted from [14] with permission). The inset shows the structure of ubiquitin with indication of the residue with the highest adsorption probability. (d) Snapshot from a simulation of adsorption of a BSA protein over a 6 nm NP, with indication of the residues with the most significant conformational change due to adsorption. Note the spread of the protein over the NP (reproduced with permission from [9]).

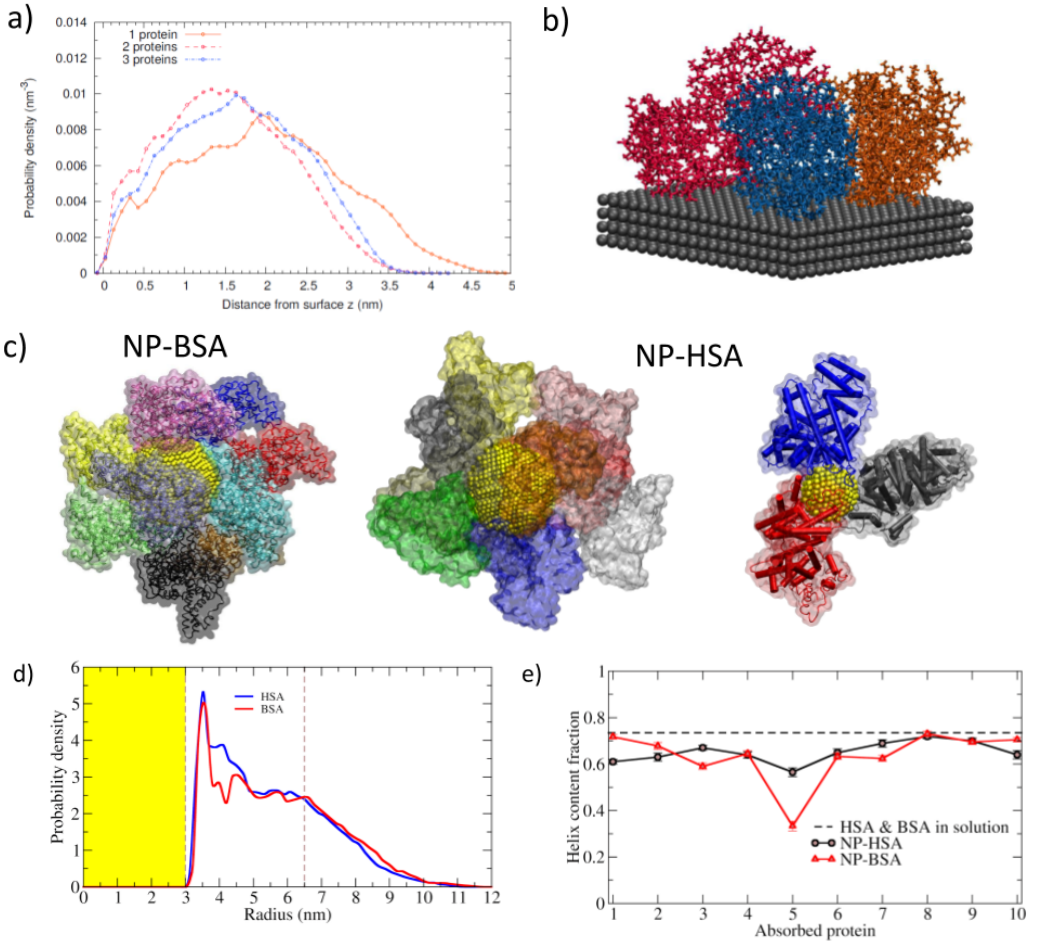

Figure 2: Examples of results from MD simulations with implicit solvent of systems with many adsorbed proteins: a) and b) correspond to a simulation of a β-casein film from [15] and (c)-(e) correspond to simulations of NP protein corona [9, 35]. (a) Probability distribution of β-casein atoms as a function of the distance from the adsorbing surface obtained in MD simulations with one, two or three proteins adsorbed onto a 45.5 nm^2 surface (b) Snapshot of the film formed by three β-casein proteins with each protein in a different colour. c) Snapshots comparing different protein corona obtained in simulations: a 6 nm NP covered with 10 BSA proteins (left), the same NP covered with 10 HSA proteins (center) and a 3 nm NP covered with HSA (right). Each colour correspond to a different protein. d) Atom density distribution (atoms/nm^3) of carbon atoms from HSA or BSA proteins in the corona of a 6 nm NP. e) α-helix content for each of the 10 proteins (HSA or BSA) in the corona of a 6 nm NP compared with its value in bulk solution.