

# Protein-surface interactions at the nanoscale: Atomistic simulations with implicit solvent models

David Malaspina, Leonor Pérez-Fuentes, Carlos Drummond, Delfi Bastos-González, Jordi Faraudo

### ▶ To cite this version:

David Malaspina, Leonor Pérez-Fuentes, Carlos Drummond, Delfi Bastos-González, Jordi Faraudo. Protein-surface interactions at the nanoscale: Atomistic simulations with implicit solvent models. Current Opinion in Colloid & Interface Science, 2019, 41, pp.40-49. 10.1016/j.cocis.2018.11.005 . hal-01962389

## HAL Id: hal-01962389 https://hal.science/hal-01962389

Submitted on 25 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

David C. Malaspina<sup>a</sup>, Leonor Pérez-Fuentes<sup>b</sup>, Carlos Drummond<sup>c</sup>, Delfi Bastos-González<sup>b</sup>, Jordi Faraudo<sup>\*a</sup>

<sup>a</sup>Institut de Ciencia de Materials de Barcelona (ICMAB-CSIC), Campus de la UAB, E-08173 Spain

<sup>b</sup>Biocolloid and Fluid Physics Group, Department of Applied Physics, University of Granada, Av. Fuentenueva 2, E-18001 Granada (Spain) <sup>c</sup>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 an 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 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

Preprint submitted to Curr. Op. Col. Interf. Sci.

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

#### 1 1. Introduction

Protein adsorption at interfaces has been described several times as a 2 "common but very complicated phenomenon" [1, 2]. Different factors which 3 are difficult to quantify such as the softness and hydrophobic or hydrophilic 4 character of the protein or surface polarity for example are thought to play a 5 decisive role in the properties of adsorbed protein layers and in biomaterials 6 design [3]. In the bionanotechnology field, protein adsorption also has a 7 prominent role. Nanoparticles (NP) in contact with biological fluids, are rapidly covered by a protein corona which determines the interactions and 9 the biological identity of the material [4]. In fact, tailoring the protein corona 10 is an essential step in drug delivery applications of NPs [5, 6]. 11

Our ability to predict the behaviour and properties of proteins is growing 12 rapidly due to the substantial increase in our knowledge of protein struc-13 ture with atomistic resolution. For example, the Protein Data Bank [7, 8] 14 contains the atomic coordinates of about  $1.4 \times 10^5$  structures (and growing, 15 at a rate of  $10^4$  structures per year), resolved by methods such as X-ray or 16 NMR. In principle, these structures can be used, in combination with the-17 oretical methods, to predict the interactions of proteins with materials and 18 tackle the open questions related to protein adsorption, protein films and 19 protein corona from a rational design perspective. To this end, a suitable 20 computational tool is of course needed. A fitting candidate as a tool for 21 performing such theoretical investigation is Molecular Dynamics simulations 22 (MD). Conceptually, the method is simple. It is based on the fact that at or-23 dinary conditions (not under extreme temperatures or pressures), the motion 24 of atoms of any molecular system can be computed implementing numeri-25 cally the Newton laws of motion in a computer. The method of course needs 26 an accurate description of the interaction forces between atoms (belonging to 27 the same or different molecules), which can be done with good approxima-28 tion using suitable, well-known force fields in many situations. The problem 29 for the use of MD simulations for the study of protein-surface interactions 30 is in the limitations of the method for dealing with large number of atoms 31 and long time scales. For example, an atomistic model of a simple protein 32 such as BSA has  $\approx 9200$  atoms (including hydrogen atoms). The protein 33 corona of a 6 nm nanoparticle (NP) contains  $\approx 10$  BSA proteins[9], so the 34 number of atoms required to model the proteins and the NP is  $\approx 10^5$ . The 35 process of protein adsorption takes place always in water, so the addition of 36 a water box large enough to have the proteins and NP in suspension in such 37

a hypothetical simulation will increase the number of atoms by an order of 38 magnitude, making the simulation impossible. An additional limitation is 39 related to time scales. MD simulations in explicit solvent are typically lim-40 ited to time scales of the order of  $\approx 10^2$  ns, which are too short to describe 41 important processes such as protein diffusion or rearrangement at surfaces. 42 For these reasons, the modelling of protein-surface interaction has been al-43 most exclusively based on simplified models, which take into account the 44 protein structure from a coarse-grained perspective. However, the develop-45 ment of implicit solvent force fields has allowed simulations of many complex 46 systems involving proteins with full atomistic detail. In these models, water 47 molecules are not included explicitly and the interaction of atoms from the 48 protein are modified from classical force fields to include the effect of wa-49 ter in an implicit way [10, 11, 12, 13]. In this way, it is possible to reduce 50 the number of atoms in the MD simulation to an affordable amount, with a 51 concomitant reduction in the number of atom-atom interactions to be calcu-52 lated in the simulation. It has to be noted that in implicit solvent all-atomic 53 models, atom-atom interactions are more computationally expensive to cal-54 culate than in explicit solvent models [12, 13], but the overall balance is that 55 simulations with implicit solvent models are (in general) substantially faster 56 than explicit solvent simulations. Using these models, it has been possible 57 to investigate with atomistic detail many interesting processes such as the 58 structure of protein corona of a NP [14, 9] or how proteins adsorbed onto 59 a surface respond to a pH change [15], among others, as we will see. Our 60 objective in this article will be to summarize these developments. Of course, 61 as any approach to a difficult problem, these simulations have their own 62 difficulties and limitations, which we will also discuss. 63

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.

# Implicit solvent models for all-atomic Molecular Dynamics sim ulations of proteins

In implicit solvent models, the solvent is replaced by a continuum di-73 electric medium and the expressions employed to compute the interactions 74 between atoms are modified (from those employed with explicit water) to 75 include the solvent in an effective way [10, 11, 13]. The first implicit solvent 76 force field now known as Generalized Born Implicit Solvent model (GBIS, 77 also sometimes abbreviated as GBSA to emphasize the inclusion of the ef-78 fects due to the solvent accessible surface) was originally proposed in the 90s 79 [10] to describe the interaction energy between molecules in solution for use 80 in molecular dynamics simulations without the need to describe explicitly the 81 solvent molecules. The main concept in the GBIS force field is to combine the 82 concepts of classical force fields for explicit solvent MD simulations with the 83 basic concepts of the Born theory of molecular solvation and the basic theory 84 of the hydrophobic effect. In its original formulation [10], MD simulations 85 employing the GBIS force field described with a very good approximation 86 the solvation free energy of a wide range of small molecules, from inorganic 87 ions to organic molecules. However, the theory produced inaccurate results 88 in the case of molecules with large interior regions with many atoms buried 89 inside, as in the case of proteins [16]. Over the years, refinements in the cal-90 culation of the dielectric screening [17] and electrostatics near the surfaces of 91 the atoms [11] provided the necessary improvements and good results were 92 obtained when used with MD of proteins. The accuracy of the current ver-93 sions of the GBIS forcefield for MD simulations of biomolecules is discussed 94 in many reviews (see for example [18, 19]). Nowadays, implementations of 95 GBIS force-fields for implicit solvent MD simulations are available in stan-96 dard MD codes such as AMBER [11], NAMD [12, 13] or GROMACS [20]. 97

The GBIS implicit solvent force field, as implemented in the MD codes 98 mentioned above, describe the atoms with the same atomic partial charges 99 and Lennard-Jones atomic parameters as employed in the explicit solvent 100 force field for the calculation of electrostatic and Van der Waals interac-101 tions. However, two substantial modifications are made in the energy and 102 force equations of the force field. The first modification is that electrostatic 103 interactions are calculated using a modified expression (instead of a direct 104 Coulomb's law) which takes into account the degree of exposure of each 105 atom to the solvent [10, 11]. And the second modification is that hydropho-106 bic interactions are included as an additional term in the force field which 107

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 com-110 plex spatial dependence than Coulomb's law, and it contains as a fundamen-111 tal quantity determining the length scale of the interactions, the so-called 112 Born radii. Each atom in a molecule is represented by an sphere filled uni-113 formly with material of dielectric constant  $\epsilon_r = 1$  [11]. The exterior of the 114 atom is filled with a continuum medium with an effective dielectric constant 115 of the solvent which also depends on the implicit ion concentration [13]. The 116 size of the atom is given by the Born radius which describes the exposure 117 of a given atom to the solvent (thus determining the degree of screening of 118 electrostatic interaction by the solvent). The Born radius depends on the dis-119 tance to neighbouring atoms. For an isolated atom, its Born radius is equal 120 to its van der Waals radius [21], while for a deeply buried atom, its Born 121 radius is much larger than its Van der Waals radius. Depending on the spe-122 cific implementation of the GBIS force field, the Born radius is calculated by 123 the Onufriev-Bashford-Case (OBC) method [11] or by the Hawkins-Cramer-124 Truhlar method [17]. 125

We recall here that the generalized Born implicit solvent GBIS model de-126 scribed above gives an accurate description of the polar contribution to the 127 energy of solvation. The second modification of the force field for implicit sol-128 vent accounts for the nonpolar (i.e. hydrophobic) contribution to the energy 129 of solvation. As we said, the model assumes that the nonpolar, hydrophobic 130 solvation energy is proportional to the exposed hydrophobic surface, with a 131 surface tension  $\approx 0.005$  kcal/mol Å<sup>2</sup> [13] which is a typical value for hydro-132 carbons in water. This is calculated by computing the surface exposed by 133 each atom weighted by an atom-type weight. 134

Using an implicit solvent force field, one reduces greatly the number of 135 atoms in the simulation and more importantly the number of atom pair inter-136 actions to be computed. But the price to pay is an increased computational 137 cost of the calculations. The increased computational time has two sources. 138 First, in the GBIS model, it is not possible to use the Particle Mesh Ewald 139 (PME) method to speed up electrostatic calculations [12, 13], as it is usually 140 done with Coulomb interactions in explicit solvent MD simulations. The only 141 option in GBIS force fields is to compute long-range screened electrostatics 142 with a large cut off. The second source of increased computational cost 143 is related to the complexity of the GBIS equations (which require also the 144 calculation of the Born radius which depends on the distribution of atoms). 145

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 148 to two factors: a) the absence of viscosity (which induces faster speed in 149 relaxation processes) and b) absence of slow processes involving explicit water 150 molecules. Therefore, the comparison between the actual gain in implicit 151 solvent MD depends on the amount of water required in the simulation box 152 to perform the explicit solvent simulation of the same problem. This balance 153 is nicely illustrated by the calculations reported in Ref [22]. For example, 154 an implicit solvent MD simulation of a 25,100-atom model of a nucleosome 155 is nominally 1.6 times slower than the explicit-solvent PME simulation with 156 a small water solvent box extending 1 nm from the solute, whereas it is 1.6 157 times faster compared to a PME simulation in a larger 3.6 nm water solvent 158 box. 159

Another important point, in order to compare implicit solvent and ex-160 plicit solvent MD simulations is that nominal simulation time (this is, the 161 time step in real units multiplied by the number of steps) has different mean-162 ings. In an explicit solvent MD simulation, the nominal simulation time is 163 expected to correspond to the clock time elapsed in the real system in an 164 equivalent experimental situation. But in the case of implicit solvent models, 165 the absence of explicit water molecules and the absence of viscosity implies 166 that the exploration of the configuration space takes place at a higher speed 167 than in a real situation. In other words, for each ns of simulation, a protein 168 will explore more possible configurations in an implicit solvent MD simula-169 tion than in an explicit solvent MD simulation. Therefore, the quantification 170 of the speed up obtained in implicit solvent models is not a trivial task. This 171 question has been investigated in detail in Ref [22] by comparing the speed 172 of five different conformational changes in implicit solvent and explicit sol-173 vent MD simulations (a dihedral angle flip, nucleosome histone tail collapse, 174 DNA unwrapping from the nucleosome histone tail collapse, DNA unwrap-175 ping from the nucleosome histone core and miniprotein folding). The authors 176 compare not only the differences in nominal simulation times for explicit and 177 implicit water but also estimate the time that will correspond to a explicit 178 water simulation to explore the configurations obtained in the trajectories of 179 the implicit model MD simulations. Taking into account all these factors, 180 the speed up of the conformational change in implicit solvent was shown to 181 be strongly depended on the particular problem. In the case of a dihedral 182 angle flip, the speed up is only by a factor of 1.6 whereas that in the case of 183

<sup>184</sup> miniprotein folding the speed up is by a factor of  $\approx 54$ .

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.

### <sup>191</sup> 3. MD simulations of the adsorption of a single protein

Once the validity of the implicit solvent MD simulations of proteins has 192 been established (see Section 2), the next step is to consider the adsorption 193 of a single protein onto a surface. Albeit simple, there are many interest-194 ing questions that have been studied from such simulations, as illustrated 195 in Figure 1. Some of these questions are the identification of the specific 196 amino acids involved in the adsorption of a given protein at a given surface 197 (and thus the driving force for adsorption), structural changes of proteins 198 after adsorption, the effect of changing conditions (such as pH) after protein 199 adsorption or the effect of curvature in adsorption (adsorption onto a planar 200 surface compared with adsorption onto small nanoparticles). 201

We will discuss now some representative results related to these ques-202 tions, by comparing the results obtained for different proteins and different 203 substrates. In Figure 1, we summarize the results obtained in the case of 204 four different systems: (a) the adsorption of  $\beta$ -Lactoglobulin onto a planar 205 metal surface [23], (b) the adsorption of  $\beta$ -case onto a generic model of 206 a planar hydrophobic surface, (c) the adsorption of ubiquitin onto a 10 nm 207 diameter Ag nanoparticle and (d) the adsorption of a BSA protein onto a 6 208 nm diameter nanoparticle. 209

In all these cases, the implicit solvent simulations allowed to identify 210 the specific amino acids or the specific protein domain involved in adsorp-211 tion, as shown in Figure 1. One interesting result obtained in some of these 212 simulations is that the residues involved in protein adsorption -for a given 213 protein- are essentially the same in different situations (such as for charged 214 or neutral surfaces or at different pH). For example, in the case of bovine 215  $\beta$ -Lactoglobulin [23] adsorbed onto a gold surface, it was found that the 216 residues involved in the adsorption were the same for a neutral surface or for 217 surfaces with different values of positive charge. The protein residues adsorb-218 ing at the neutral Au surface are the hydrophilic Thr 125, Thr 18, Lys 100 219

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 residuesurface interactions.

Analogous results were obtained in [15] for adsorption of  $\beta$ -casein onto 226 model hydrophobic surfaces of different charges (Figure 1b). At neutral pH, 227 the negatively charged protein (with a charge of -6e) adsorbs onto a neutral 228 hydrophobic surface by contact with several, mostly hydrophobic residues 229 but also by some polar residues situated near these hydrophobic ones. The 230 residues involved in the contact between the protein and the neutral surface 231 were six hydrophobic aminoacids (Pro 76, 78, 194, 196, 201 and Phe 205), 232 and three polar aminoacids (Tyr 75, Val 193 and Ile 202), all indicated in 233 Figure 1b). No significant change in the adsorption of the protein (same 234 adsorbed residues and same area of contact) was observed when changing 235 the charge of the surface from neutral to negative  $(-0.62 \text{ e/nm}^2)$ , in spite of 236 the electrostatic repulsion (recall that both the surface and the protein were 237 negatively charged). Finally, the charge of the adsorbed protein was changed 238 from -6e to +8e, corresponding to a pH change from 7 to 4 (the charge of 239 the surface was maintained unchanged at  $-0.62 \text{ e/nm}^2$ ). In this case, the 240 adsorbed protein increases the contact with the surface, by adsorbing not 241 only with the same residues as in the case of neutral pH but also with other 242 protonated, positively charged residues located near the previously adsorbed 243 domain. 244

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 250 Ag NP. They found that the protein adsorbs by binding of the Asp 18 residue 251 to the NP (see Figure 1c) and also to some extend with an interaction of the 252 Gln 8 residue with the NP. No significant change in the secondary structure 253 of the protein was found due to adsorption. These simulation results are 254 in agreement with NMR studies [24] which found that the binding domain 255 of human ubiquitin to Ag NP was located in the residues Gln 2-Ile 3 and 256 Leu 15-Asp 18. Concerning the kinetics of adsorption, they obtained that 257

<sup>258</sup> protein re-orientation was the rate-limiting step in protein adsorption.

In Ref. [9], the authors studied experimentally and by MD implicit solvent 259 simulations the adsorption of BSA onto a 6 nm iron oxide NP covered by 260 citrate. The simulations show that as the protein contacts the NP, the BSA 261 protein changes its conformation and the protein domain in contact with the 262 NP partially spreads over the surface of the NP (see Figure 1d). In fact, 263 the surface of the protein increased from  $325 \text{ nm}^2$  in solution to  $345 \text{ nm}^2$ 264 at the NP surface. The change in the secondary structure of the protein 265 was small; the  $\alpha$ -helix content changed from 72% in solution to 66% for 266 a BSA adsorbed onto a NP. However, a few residues changed significantly 267 its environment. It was particularly interesting the case of the Trp213 and 268 Trp134 residues which contribute to the UV-vis spectra of the BSA protein. 269 Simulations also indicate a change of their environment from being buried 270 into the protein to a more solvent exposed location, a result in agreement 271 with UV-vis spectroscopy measurements [9]. 272

All the simulation results discussed so far of MD simulations with implicit 273 solvent indicate a small change in the conformation of a protein after adsorp-274 tion, with the changes localized at a few particular residues. Probably the 275 most significant exception is the case of adsorption of proteins at carbon sur-276 faces (graphite or graphene) in which several works using implicit solvent MD 277 simulations report substantial unfolding of proteins or protein fragments over 278 these carbon surfaces. Examples include unfolding of albumin and fibronectin 270 fragments after adsorption onto graphite and carbon nanotubes [25], unfold-280 ing of lysozyme on graphite [26] and unfolding of a BMP-2 protein onto 281 graphite [27]. Interestingly, the results for unfolding of BMP-2 protein onto 282 graphite<sup>[28]</sup> were further confirmed by MD simulations in explicit solvent 283 using accelerated MD simulations. Other explicit solvent MD simulations 284 of proteins and peptides onto carbon surfaces indicate similar conclusions. 285 Explicit solvent MD simulations of two "de novo" designed  $\alpha$ -helical pep-286 tides [29] show that they unfold and assemble into an amorphous dimer at 287 a graphene surface, in agreement with circular dichroism spectroscopy and 288 scanning tunneling microscopy measurements. It seems that the specific role 289 of the carbon surface is important in these results. For example, extensive 290 explicit solvent simulations of lysozyme adsorption/desorption on polymeric 291 hydrophobic surfaces [30, 31] only show small secondary structure changes 292 in particular residues (typically those located near the adsorption region), 293 but they do not show unfolding of the protein. Interestingly, in these works 294 it was found that there is an energy barrier for adsorption mainly arising 295

from protein and surface hydration in the first hydration shell. The protein cannot be adsorbed even though the substrate surface exhibits attraction, whenever the proteinsurface 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 301 and the validity of its modelling by using implicit solvent MD simulations 302 was analysed in detail in Ref. [32]. First, the authors repeated and ex-303 tended previous studies [33] which reported unfolding of BSA after adsorp-304 tion onto graphene. They found large unfolding of BSA upon adsorption onto 305 a graphene surface, with an  $\alpha$ -helix content of only 35% after free adsorption 306 onto graphene. The authors also reported extensive all-atomic MD simu-307 lations of BSA adsorption onto graphene with explicit solvent and in this 308 case the  $\alpha$ -helix content was about 60% after free adsorption. Therefore, 300 in this case the unfolding of BSA was observed only in the case of implicit 310 solvent MD simulations. This discrepancy probably indicates that the em-311 ployed force field for carbon surfaces greatly exaggerates the strength of the 312 interaction of BSA with graphene, inducing a BSA unfolding as a simulation 313 artefact. We think that better, thoroughly validated force fields for implicit 314 solvent simulations of carbon surfaces are needed for implicit solvent sim-315 ulations of protein adsorption. As a general consequence, we can say that 316 more attention has to be paid to implicit solvent models of surfaces for use 317 in implicit solvent simulations involving biomolecular adsorption. 318

Another interesting question to discuss, in light of simulation results, is 319 the reversibility or irreversibility of protein adsorption. In all the simulations 320 discussed so far, only adsorption (with no desorption) was reported, indicat-321 ing that protein adsorption is essentially irreversible at the time scales probed 322 by atomistic MD simulations. This question has been discussed in a recent 323 all-atomic MD study (with explicit solvent) of the energy landscape for BSA 324 adsorption onto silica [34]. The results indicate that the time scales for pro-325 tein desorption are of orders of hours in this case, whereas time scales for 326 protein diffusion are in the 100 ns range. Therefore, protein desorption is 327 far beyond the scales accessible to MD (even with implicit solvent models) 328 whereas protein diffusion should be observable, if present, by all-atomic MD 329 simulations. 330

# 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 337 interactions became more and more important, but these situations are com-338 putationally expensive because they require the consideration of many pro-339 teins at the surface. The cost of the simulations increases not only because 340 the number of atoms in the simulation increases but also because the dy-341 namics of the problem itself is much slower. This makes the simulation of 342 high coverages prohibitively expensive in many cases. For this reason, only 343 recently simulations considered atomistic simulations of protein films or pro-344 tein corona in which protein-protein interactions are essential. In spite of 345 these difficulties, a few of the studies discussed in the previous section con-346 sider not only the case of a single protein adsorption but also the formation 347 of an adsorbed protein layer with atomistic detail, which we will now discuss. 348 In addition to the study of single  $\beta$ -case protein adsorption (discussed) 349 in Section 3, see Figure 1b), this work also considered simulations with two 350 and three adsorbed proteins onto a surface of  $45.5 \text{ nm}^2$ . These simulations 351 with two and three proteins correspond to a mass coverage of  $1.74 \text{ mg/m}^2$ . 352 and 2.6 mg/m<sup>2</sup> which are about 47% and 70% of the maximum experimental 353 coverage for a  $\beta$ -case monolayer [15]. Comparing the simulation results at 354 the different coverages (see Figure 2a), it was observed that the thickness 355 of the film decreases from  $\approx 4$  nm at the lowest coverage to  $\approx 3$  nm at the 356 highest coverage, indicating that at higher coverages the proteins tend to 357 be adsorbed in a more compact configuration. As shown in the snapshot in 358 Figure 2b, the adsorbed proteins are in contact, forming a compact structure. 359 Hence, the results indicate a substantial protein-protein attraction in spite 360 of the electrostatic repulsion between proteins (each protein has a charge of 361 -8e in the simulation, corresponding to pH=7). 362

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 1c). 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 368 commonly employed in experiments as a cheaper alternative to experimenta-369 tion with the human serum albumina (HSA) protein, which is more relevant 370 for biomedical applications. For this reason, we have also repeated the anal-371 ysis in [9] using HSA instead of BSA [35] and we have also considered two 372 different particle sizes (6 nm and 3 nm of diameter). Starting from the sim-373 ulation with a single adsorbed protein, the simulations of the protein corona 374 were done by systematically adding more proteins in solution to a previous 375 simulation (see [9]) with a smaller number of proteins and performing long 376 simulation runs in order to allow for structural relaxations and adsorption of 377 additional proteins [9]. It should be noted that the total number of atoms 378 in these simulations with many proteins is  $\approx 10^5$  atoms (for a 10 protein 379 simulation) while the same system in explicit solvent would have more than 380 1 million of atoms in the simulation box. The simulations indicate that the 381 maximum number of BSA or HSA proteins that can be accommodated onto 382 the 6 nm NP is of 10 proteins (see Figure 2b), a result which is in agree-383 ment with experiments [9] (in which the number of BSA at the corona was 384 estimated from the change in size of the particle before and after protein 385 corona formation). It is also interesting to recall that the maximum number 386 of proteins considered in the simulations was 12 and that in these simula-387 tions with excess protein, 2 of the proteins remain in bulk without adsorbing 388 and without interacting significantly with the layer of adsorbed proteins [9]. 389 This result also suggests that these proteins form only a monolayer over the 390 NP, since a soft corona (a second layer of protein adsorbed onto the protein 391 corona) was not observed in our simulations. 392

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  $\alpha$ -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 rep-408 resentation of protein secondary structure, so we can discuss the secondary 409 structure of the proteins in the protein corona. Figure 2c shows the  $\alpha$ -helix 410 content for each of the HSA or BSA proteins adsorbed on the protein corona, 411 compared with a protein in bulk solution. The results show that the changes 412 in secondary structure are small in all cases, but there is also a tendency 413 indicating that these changes are smaller for the last adsorbed proteins. The 414 explanation for this effect is that at low coverages, there is more space avail-415 able for the spreading of the protein over the surface (recall here previous 416 section and Figure 1b). On the contrary, near saturation, the last adsorbed 417 protein has a small surface available for adsorption and a small contact with 418 the NP, without possibility of spreading over the surface. 419

#### 420 5. Conclusions and outlook

Implicit solvent models for protein simulation are an interesting alterna-421 tive for simulations of proteins retaining all-atomic details with affordable 422 resources. Their use in the study of problems involving protein adsorption 423 onto surfaces, such as protein films or the protein corona of NPs is a young 424 and promising approach that has delivered several interesting insights but 425 still has drawbacks that need to be tackled. We can summarize the ac-426 complishments and shortcomings of MD atomistic simulations of proteins at 427 surfaces using implicit solvent GB models as follows: 428

429 430 431 • 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 (i10 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 gen eral 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 in a implicit way, at a Poisson-Boltzmann
  level of description. This is clearly not enough for describing the complex specific effects of ions with proteins, which are know 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 in MD simulations.

As we have discussed in this review, the direction of the field in the last years 452 has been clearly focused on the implementation of the method in the software 453 usually employed in MD simulations and to speed up the implementations 454 of the method (by adding features such as the use of GPU). We are sure 455 that these implementations of the method will fuel exciting new uses of the 456 method to study more complex problems. But before more applications of 457 greater complexity can be considered, it will be of a great interest to advance 458 and improve the theory. We think that, in view of the points listed above, the 459 most pressing questions to be tackled are the development of more accurate 460 implicit solvent models for nanostructured surfaces of interest and also the 461 development of consistent models for the explicit inclusion of ions (and their 462 specific effects) and the inclusion of charge regulation in implicit solvent MD 463 simulations. 464

### 465 Acknowledgements

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

### 471 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+.
- 484 [5] J. Mariam, S. Sivakami, P. M. Dongre, Albumin corona on nanoparticles
  485 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 pro tein 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.
- [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.

- [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.
- [12] D. E. Tanner, K.-Y. Chan, J. C. Phillips, K. Schulten, Parallel general ized 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.
- <sup>511</sup> [14] F. Ding, S. Radic, R. Chen, P. Chen, N. K. Geitner, J. M. Brown,
  <sup>512</sup> P. C. Ke, Direct observation of a single nanoparticleubiquitin corona
  <sup>513</sup> formation, Nanoscale 5 (2013) 9162+.
- 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+.
- [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.
- I9] 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.
- 542 [23] T. Hagiwara, T. Sakiyama, H. Watanabe, Molecular simulation of
  543 bovine beta-Lactoglobulin adsorbed onto a positively charged solid sur544 face, Langmuir 25 (2009) 226–234.
- 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.
- <sup>558</sup> [28] C. Mücksch, H. M. Urbassek, Enhancing protein adsorption simulations <sup>559</sup> 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.
- [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.
- <sup>576</sup> [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)
  <sup>579</sup> 3744–3753.
- <sup>580</sup> [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, Inter action 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  $\beta$ -Lactoglobulin onto an Au(100) surface, with indication of the residues involved in protein adsorption (reproduced from [23] with permission), (b) Adsorption of  $\beta$ -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 10nm Ag NP. The histogram indicates the number of residues with a given adsorption probability (reproduced from [14] with permission). The inset shows the structure of ubiquitin with indication of adsorption of a BSA protein over a 6nm 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  $\beta$ -casein film from [15] and (c)-(e) correspond to simulations of NP protein corona [9, 35]. (a) Probability distribution of  $\beta$ -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<sup>2</sup> surface (b) Snapshot of the film formed by three  $\beta$ -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<sup>3</sup>) of carbon atoms from HSA or BSA proteins in the corona of a 6 nm NP compared with its value in bulk solution.