

PAPER

## A multiscale model of protein adsorption on a nanoparticle surface

To cite this article: David Power *et al* 2019 *Modelling Simul. Mater. Sci. Eng.* **27** 084003

View the [article online](#) for updates and enhancements.

### You may also like

- [Protein corona formation and its constitutional changes on magnetic nanoparticles in serum featuring a polydehydroalanine coating: effects of charge and incubation conditions](#)  
Christine Gräfe, Moritz von der Lühe, Andreas Weidner *et al.*
- [Protein adsorption on nanoparticles: model development using computer simulation](#)  
Qing Shao and Carol K Hall
- [\(Invited\) A Protein Corona-Based Design Strategy for Carbon Nanotube Sensors](#)  
Rebecca Pinals, Francis Ledesma, Linda Chio *et al.*

### Recent citations

- [First principles characterisation of bio–nano interface](#)  
Ian Rouse *et al*
- [Alexander Lyubartsev and Aatto Laaksonen](#)
- [Molecular Modeling of Protein Corona Formation and Its Interactions with Nanoparticles and Cell Membranes for Nanomedicine Applications](#)  
Hwankyoo Lee





**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# A multiscale model of protein adsorption on a nanoparticle surface

David Power<sup>1</sup>, Ian Rouse<sup>1</sup>, Stefano Poggio<sup>1</sup>, Erik Brandt<sup>2</sup>,  
Hender Lopez<sup>3</sup> , Alexander Lyubartsev<sup>2</sup> and  
Vladimir Lobaskin<sup>1</sup> 

<sup>1</sup> School of Physics, University College Dublin, Belfield, Dublin 4, Ireland

<sup>2</sup> Department of Materials and Environmental Chemistry, Stockholm University, SE-10691 Stockholm, Sweden

<sup>3</sup> School of Physics and Optometric & Clinical Sciences, Technological University Dublin, Kevin Street, Dublin 8, Ireland

E-mail: [vladimir.lobaskin@ucd.ie](mailto:vladimir.lobaskin@ucd.ie)

Received 30 April 2019, revised 26 July 2019

Accepted for publication 15 August 2019

Published 3 September 2019



CrossMark

## Abstract

We present a methodology to quantify the essential interactions at the interface between inorganic solid nanoparticles (NPs) and biological molecules. Our model is based on pre-calculation of the repetitive contributions to the interaction from molecular segments, which allows us to efficiently scan a multitude of molecules and rank them by their adsorption affinity. The interaction between the biomolecular fragments and the nanomaterial are evaluated using a systematic coarse-graining scheme starting from all-atom molecular dynamics simulations. The NPs are modelled using a two-layer representation, where the outer layer is parameterized at the atomistic level and the core is treated at the continuum level using Lifshitz theory of dispersion forces. We demonstrate that the scheme reproduces the experimentally observed features of the NP protein coronas. To illustrate the use of the methodology, we compute the adsorption energies for human blood plasma proteins on gold NPs of different sizes as well as the preferred orientation of the molecules upon adsorption. The computed energies can be used for predicting the composition of the NP-protein corona for the corresponding material.

Supplementary material for this article is available [online](#)

Keywords: nanoparticle, multiscale modelling, protein, coarse-grained modelling, protein corona

(Some figures may appear in colour only in the online journal)

## 1. Introduction

Interactions of a foreign nanomaterial with biological tissues control its fate and biological activity, including eventual hazard. Quantitative study of these interactions is extremely challenging due to the presence of multiple molecule types and material chemistry—as well as structure-specific effects, and the immense system size prohibits atomistic level modelling.

It is now well accepted that the interaction of nanoparticles (NPs) with biological fluids leads to the formation of a protein layer on the surface of the NP which is known as the protein corona. It has been established that the corona plays the central role in the bioactivity of the nanomaterial. NPs of size of tens of nanometers can bind hundreds of different proteins [1–5].

The list of proteins present in the corona in a specific sample depends, on the one hand, on the NP chemistry and surface structure and reflects the immediate interactions. On the other hand, it also depends on the content of the biological fluid the NP is immersed in through the concentrations of the solutes. Due to the presence of thousands of protein types in certain biological fluids, such as blood serum, the variability of the corona content may be immense. Still, one can hope to detect the influence of the physicochemical properties of the material in the adsorbed proteins' statistics for different nanomaterials. Certain features of proteins such as charged or hydrophobic patches, aromatic residues, etc may tend to increase the propensity of molecules to adsorb on specific surfaces. This observation led researchers to an idea of construction of nanomaterial fingerprint [6]. The fingerprints appeared to be useful for prediction of the biological activity of nanomaterials, particularly NP-cell association. Although the initial statistical quantitative structure-activity relationships (QSAR) model developed in [6] were quite heavy (64 fitting parameters), it was demonstrated later that only a few of the protein characteristics are actually important [7]. Namely, the cell association of gold NPs correlates well with the sequence descriptors responsible for protein charge (such as basic, acidic, and aspartic amino acids percentage) as well as with molecular weight, and propensity of the protein to aggregation.

Despite the fact that it is generally known that the abundance of any particular protein in the corona varies for different nanomaterials (see [8] for a review), the properties responsible for this variance have not been identified. We believe that a significant improvement in the predictive power of the bioinformatics-based models can be achieved with structure descriptors characterizing size, shape, and charge distribution on the protein, i.e. the properties relevant for the interactions with NPs. For the NPs, one can expect that the possibility of quantitative predictions depends on the availability of the physicochemical characteristics of the NP surface such as dielectric constant, Hamaker constant, charge, and hydrophobicity.

In this work, we propose a multiscale methodology for modelling bio-nano interactions that directly relates the NP characteristics to protein binding affinity. Our goal is to compare and rank biomolecules by their adsorption affinity and thus form a basis for producing NP biointeraction fingerprints. We describe the modelling techniques for different levels of description, as well as their coupling in section 2. In section 3, we use these techniques to evaluate the interaction characteristics for gold NPs and common blood serum proteins. A discussion of the results and the method is presented in section 4, and conclusions in section 5.

## 2. Methodology

Once an NP comes into contact with a biological medium, a protein corona forms on its surface [9]. It has been extremely challenging to develop a model that can predict the

composition of the protein corona around an inorganic NP, as this depends on a multitude of physicochemical properties both of the protein and of the NP, such as size, shape, pH, hydrophilicity/ hydrophobicity, and charge distribution.

Computer simulations of the interactions of NPs with proteins can offer a great support to experiments because of their great speed and flexibility [10]. Full-atomistic simulations have already proven to be a valuable tool in elucidating the binding mechanisms of proteins on metallic NPs [11–13]. However, their performance is severely hindered by the inefficiency in simulating systems with large NPs due to the high number of pair interactions that need to be evaluated. To speed-up the calculations, a cut-off of the order of a nanometer is often introduced. However, this results in an underestimation of the adsorption energies of proteins on NPs due to the neglect of the core of the NP, which contributes much to their mutual attraction (especially at radii of over 10 nm) and so cannot be neglected. In this section, we describe a CG model of protein-NP interactions that overcomes most of the challenges in the inclusion of NP core in the interaction.

While the number of atom–atom pairs in this problem is extremely large, the overall energy includes numerous instances of the same contributions, like the pair interaction of a given amino acid (AA) with a unit volume of the nanomaterial. One can dramatically reduce the amount of calculations by pre-computing those interactions for the specific material. For a protein, one would need to pre-compute about  $2 \times 20$  potentials, having in mind that different potentials will be needed for the interaction with the NP surface and the core. Assuming these interactions do not depend on the position of the AA inside the protein and are additive (the two most significant approximations of our model), one can quickly scan multiple proteins once the potentials are known.

We calculate the coarse-grained adsorption energy for a protein molecule as a sum of energies of non-bonded (van der Waals + excluded volume)  $U_{\text{nb}}$  and electrostatic  $U_{\text{el}}$  interactions between the AA and segments of the NP:

$$U = U_{\text{nb}} + U_{\text{el}}, \quad (1)$$

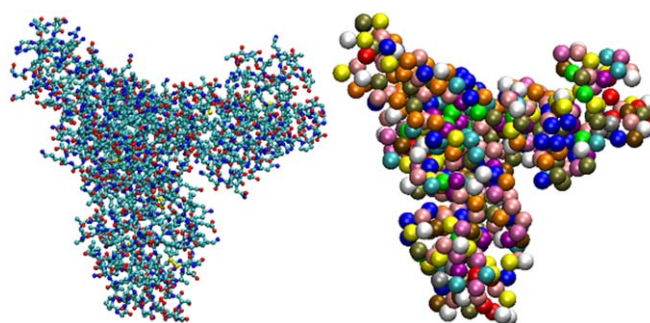
$$U_{\text{nb}}(d, \theta, \phi) = \sum_{i=1}^{N_{\text{AA}}} U_i(h_i(d, \theta, \phi)), \quad (2)$$

$$U_{\text{el}}(d, \theta, \phi) = \sum_{i=1}^{N_{\text{AA}}} \frac{\psi_s q_i R}{R + h_i(d, \theta, \phi)} e^{-\kappa h_i(d, \theta, \phi)}. \quad (3)$$

Here,  $R$  is the NP radius,  $h_i(d, \theta, \phi)$  is the distance between the AA centre and the NP surface, which depends on the distance  $d$  from the protein centre of mass (COM) to the surface and orientation angles  $(\theta, \phi)$ ,  $\kappa = \sqrt{4\pi l_B I_0}$  is the inverse Debye length where  $l_B = \frac{e^2}{4\pi\epsilon\epsilon_0 k_B T}$  is the Bjerrum length and  $I_0$  is the solution ionic strength,  $\psi_s$  is the electrostatic surface potential of the NP,  $q_i$  is the charge of the AA, and  $N_{\text{AA}}$  is the number of AAs. The non-bonded interaction energy includes contributions from the surface and the core of the NP

$$U_i(h_i(d, \theta, \phi)) = (U_s)_i + (U_c)_i, \quad (4)$$

where  $U_s$ ,  $U_c$  are the potential energies of interaction between the  $i$ th AA and surface or core parts of the NP, respectively. The surface part is calculated using atomistic molecular dynamics (MD) simulation with a fixed cut-off and thus includes occasional hydration effects or surface geometry, while the core part covers the long-range contributions to the interaction with the rest of the NP, beyond the MD cut-off distances.



**Figure 1.** All-atomistic (left) and united atom—UA (right) models of lung surfactant protein D (SP-D).

In the remainder of this section, we describe the systematic procedure of evaluation of the required potentials of mean force. In the final section, we test the model by simulating the adsorption of some common blood serum proteins on gold NPs.

### 2.1. Coarse-grain protein model

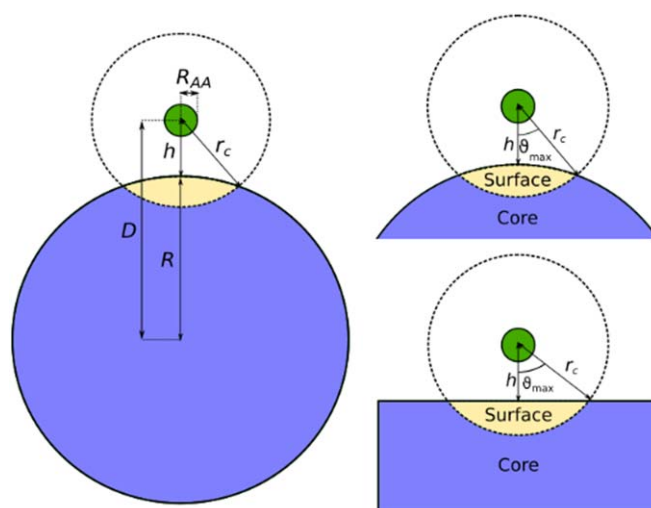
In contrast to nanomaterials, which populate a whole universe of chemistries and structures, the description of biomolecules can be significantly compressed due to their chemical uniformity, e.g. the same AAs present in all proteins or same nucleic acids in all DNA. The AA sequence alone, however, does not provide exhaustive information as the shape and functionality of the protein depends on its 3D structure. The 3D structure of the molecule, where possible, can be retrieved from the Protein Data Bank (PDB). For our calculations, we consider the proteins as rigid 3D globules with fixed conformations corresponding to their crystal structures as presented in the PDB. The model thus preserves two main structural features that guide the binding mechanisms, i.e. the overall shape of the protein and the charge distribution.

We previously proposed a one-bead-per-AA (united atom—UA) model of globular proteins, which is suitable for estimating the adsorption energy [14]. Our UA protein model, in which every AA in the protein is substituted by a single bead whose centre is placed at the position of the  $\alpha$ -carbon atom, is illustrated by figure 1.

### 2.2. Coarse-grain nanoparticle

The protein model described above allows us to reduce the number of components in treating the protein. NP size, however, also plays a fundamental role in the formation of the corona and in the interactions at the bio-nano interface. The number of atoms needed to represent an NP is again a severe limitation to all atomistic calculations. Simulation of NPs of size greater than 10 nm in biological media is an unfeasible task even for modern computers and CG model for describing NPs is therefore highly needed.

Our model starts by considering the contributions that different atoms in the NP give to the binding interaction. Based on the nature of these contributions, we partition the NP into a core segment and a surface segment. The outer layer on the NP surface is directly in contact with the solvent and the pair interaction with the protein residues must include solvent effects as well as the chemical composition, charge, and hydrophilicity/ hydrophobicity of the NP surface. Therefore, the interaction of each residue with the nearest part of the surface must be



**Figure 2.** Left: A schematic of an amino acid bead of radius  $R_{AA}$  interacting with a nanoparticle of radius  $R$ , where the distance between the centres of the two particles is given by  $D$  and the distance between the surface of the NP and the centre of the bead is given by  $h$ . The nanoparticle is divided into a surface region (yellow) containing the volume within a distance  $r_c$  of the centre of the bead, and a core region (blue) containing the volume outside this cutoff range. Right: Schematics of the surface region for a spherical NP (top) and planar slab (bottom), indicating the angle  $\vartheta_{\max}$  used for the calculation of the surface correction factor.

parameterized to reflect these details, using full-atomistic simulations. The size of the surface segment is thus determined by the cutoff  $r_c$  used in the full-atomistic simulation (typically, 1–2 nm). Geometrically, the surface segment is a lens formed by an intersection of a sphere of radius  $r_c$ , centred on the AA bead, and a sphere of radius of the NP,  $R$ , centred on the NP itself (figure 2). The core comprises majority of the atoms, but these only interact with the protein via long-range forces, for which we assume that a continuum-level description is sufficient. The core of the NP is then modelled as a single bead of the shape of a sphere of radius  $R$  with a cut-out surface lens. The potential between the core and the AA beads in our model is calculated using the Lifshitz theory [15] for interaction between two macroscopic bodies. In the next two subsections, we describe how the potentials are parameterized for the AA-NP interaction.

### 2.3. Generation of surface potentials

Adsorption free energy profiles (or potentials of mean force—PMFs) can be calculated using adaptive well-tempered metadynamics (AWT-MetaD) [16–20]. This method uses a time-dependent bias term which is evolved according to

$$\dot{V}(z, t) = \omega e^{-V(z(t), t)/\Delta T} \exp\left[\frac{-(z - \bar{z}(t))^2}{2\sigma^2(t)}\right], \quad (5)$$

with  $\omega$  being the initial filling rate and  $\Delta T$  being a temperature boost factor that determines how large part of the free energy space that is reached. The positions and widths of the added Gaussians are changed on-the-fly as exponentially weighted averages

$$\bar{z}(t) = \frac{1}{\tau_D} \int_0^t dt' z(t') e^{-\frac{(t-t')}{\tau_D}}, \quad (6)$$

and

$$\sigma^2(t) = \frac{1}{\tau_D} \int_0^t dt' (z(t') - \bar{z}(t'))^2 e^{-\frac{(t-t')}{\tau_D}}. \quad (7)$$

The single parameter  $\tau_D$  determines the time window used to estimate fluctuations in  $z$ . The free energy is recovered in a single simulation from the expression [20]

$$F_\Omega(z) = - \lim [V(z, t) + k_B T \ln \Omega(z, t) \text{ const.} \times t]_{t \rightarrow \infty}, \quad (8)$$

where  $\Omega(z, t)$  is the accumulated histogram of the reaction coordinate  $z$  up to time  $t$ . The constant on the right-hand side is independent of  $z$  and  $t$ . The logarithmic divergence as  $t \rightarrow \infty$  is irrelevant as long as we are interested in free energy differences only along  $z$ -axis. The AWT-MetaD simulations need to be run for ca. 200 ns to allow the system to visit all states many times for each AA or lipid segment.

We further used the free energies  $F_\Omega(z)$  for each AA as estimates of the surface terms  $U_s(z)$ ,  $U_s(z) = F_\Omega(z)$  with the integration constant set to zero. Upon evaluating the PMFs, we also calculated the mean adsorption free energies for each AA as:

$$E_{AA} = -k_B T \ln \left[ \frac{1}{\delta} \int_0^\delta \exp \left( -\frac{U_s(z) - U_s(\delta)}{k_B T} \right) dz \right]. \quad (9)$$

Generally, the interaction with a convex surface of a NP of finite radius is less than that for the flat slab due to the lesser number of atoms of the NP within the interaction cut-off distance. To account for this reduction, we correct the PMFs for the flat surface by a distance-dependent multiplicative function  $f(h)$  that reflects also the cut-off radius ( $r_c$ ) used in the calculations as well as the radius  $R$  of the NP.

$$U_s(h, R) = U_s(h, \infty) f(h). \quad (10)$$

Here,  $U_s(h, R)$ ,  $U_s(h, \infty)$  are the PMFs for the curved and flat surfaces, respectively, and  $h$  is the minimum distance between the AA-bead centre and the NP surface. By taking the appropriate limits for  $R$ , we can calculate a correction factor for any geometry. A diagram showing how an AA bead interacts with a CG NP is shown in figure 2.

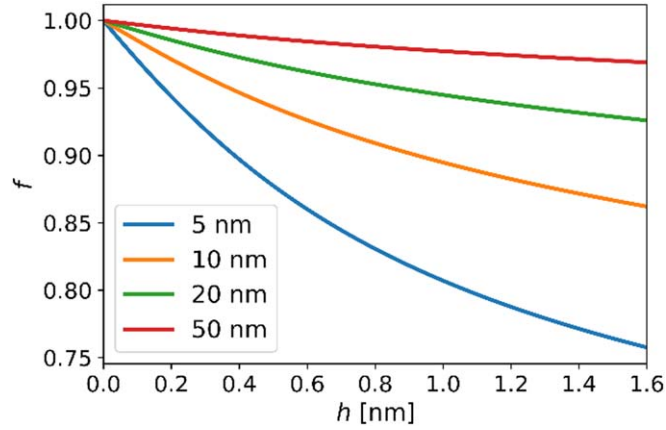
We assume that the relevant point–point interaction is dominated by dispersion forces that scale as  $r^{-6}$ , where  $r$  is the distance between the interaction centres (e.g. atoms). We then assume that AA are much smaller than the NPs. In this case, the attraction energy for a small particle of volume  $V$  to a large sphere of radius  $R$ , such that  $R \gg V^{1/3}$ , given the finite cut-off  $r_c$ , can be calculated as:

$$U_s(h, R) = \varepsilon V \int_h^{r_c} \int_0^{\vartheta_{\max}} \int_0^{2\pi} \frac{r^2 \sin \vartheta}{r^6} d\varphi d\vartheta dr, \quad (11)$$

$$\vartheta_{\max} = \cos^{-1} \left( \frac{r^2 - R^2 + (R + h)^2}{2r(R + h)} \right)_{\max}, \quad (12)$$

$$U_s(h, R) = -\frac{\pi \varepsilon V}{h + R} \left( \frac{h - 2R}{12h^3} + \frac{-6r_c^2 + 8r_c(h + R) - 3h(h + 2R)}{12r_c^4} \right). \quad (13)$$

Here,  $\varepsilon$  is the interaction energy per unit volume. For  $R \rightarrow \infty$ , equation (13) reduces to that of a flat surface:



**Figure 3.** Correction factor  $f(h)$  versus distance from the surface  $h$ , equation (15), for a range of NP radii.

$$U_s(h, \infty) = \pi\epsilon V \left( \frac{1}{6h^3} - \frac{2}{3r_c^3} + \frac{h}{2r_c^4} \right) \quad (14)$$

and the correction factor is calculated as:

$$f = \frac{U_s(h, R)}{U_s(h, \infty)} = -\frac{r_c^2(h - 2R) + 2r_ch(h - 2R) - 3h^2(h + 2R)}{2(r_c^2 + 2r_ch + 3h^2)(h + R)}. \quad (15)$$

Figure 3 shows how the volume correction factor changes with the distance from the AA to the surface for a set of NP radii. One can see that, as the radius increases approaching the flat surface limit,  $f \rightarrow 1$  for all values of  $h$ .

#### 2.4. Generation of the core potential

The NP core plays a crucial role in the protein adsorption as it contains most of the nano-material. A serious limitation of the all-atom models is the difficulty of a correct account for the attraction by the core atoms. This problem is mainly due to the short-range cut-off employed in simulations leading to a considerable underestimation of the adsorption energies. The latter, however, can be easily calculated in the continuum approximation, which is commonly used in colloid science. The correction we propose in this report is to evaluate the contribution of the core of the NP at distances  $r > r_c$  by treating the remote part of the NP as a single sphere less the part within the cut-off distance from the specific AA. The interaction energy between an AA and the core can be computed using the Hamaker method for dispersion forces. We take into account only the part of the NP that is beyond the reach of the PMF cut-off distance  $r_c$ . Then, for two spheres or radii  $R, R_{AA}$  at a distance  $D$  between their centres, such that  $D > R + r_c$  the interaction energy is the Hamaker potential:

$$U_{c,\text{long}}(D) = \frac{-A_{123}}{12} \left( \frac{4RR_{AA}}{D^2 - (R + R_{AA})^2} + \frac{4RR_{AA}}{D^2 - (R - R_{AA})^2} + 2 \ln \left( \frac{D^2 - (R - R_{AA})^2}{D^2 - (R + R_{AA})^2} \right) \right). \quad (16)$$



At the shorter distances, the Hamaker potential must be corrected for the interaction with the lens cut out of the NP by a sphere of radius  $r_c$  centred at the AA centre of mass, as it is covered by the PMF

$$U_c(D) = U_{c,\text{long}}(D) + U_{c,\text{lens}} \approx U_{c,\text{long}}(D) + \frac{A_{123}}{12} \left( \frac{4\pi^2 R_{AA}^3}{3D} \left( \frac{D-3R}{(D-R)^3} + \frac{-6r_c^2 + 8r_c D - 3h(D+R)}{r_c^4} \right) \right) \quad (17)$$

In the above equations,  $A_{123}$  is the Hamaker constant for interaction between material 1 (e.g. protein) with material 2 (e.g. the NP) through material 3 (e.g. water), the only material dependent term in the equation. The Hamaker constant for the AA (phase 1)-material (phase 2) interactions through a medium (phase 3) can be obtained as follows:

$$A_{123} = \frac{3}{4} k_B T \frac{(\varepsilon_1 - \varepsilon_3)(\varepsilon_2 - \varepsilon_3)}{(\varepsilon_1 + \varepsilon_3)(\varepsilon_2 + \varepsilon_3)} + \frac{3h\nu_e}{8\sqrt{2}} \frac{(n_1^2 - n_3^2)(n_2^2 - n_3^2)}{(n_1^2 + n_3^2)^{\frac{1}{2}}(n_2^2 + n_3^2)^{\frac{1}{2}}((n_1^2 + n_3^2)^{\frac{1}{2}} + (n_2^2 + n_3^2)^{\frac{1}{2}})}, \quad (18)$$

where  $n_i$  are the refractive indices of the materials in the visible region,  $\nu_e$  is the main electronic absorption frequency in the UV (typically  $3 \times 10^{15} \text{ s}^{-1}$ ) for material 2 and  $\varepsilon_i$  are the dielectric permittivities of the materials, which are equal to  $n^2$  in the visible part of the spectrum. In the case one of the materials is a conductor, equation (18) must be modified to take into account the high values of the polarisability and therefore of the dielectric constant. The equation for a dielectric-conductor interaction in a medium is given by [15]:

$$A_{123} = \frac{3}{8\sqrt{2}} \left( \frac{n_1^2 - n_3^2}{n_1^2 + n_3^2} \right) \frac{h\sqrt{\nu_1\nu_3} \cdot \nu_2}{\sqrt{\nu_1\nu_3} + \frac{\nu_2}{\sqrt{n_1^2 - n_3^2}}}, \quad (19)$$

where  $\nu_i$  are the frequencies of maximum absorption for the material in the UV region, for metals this corresponds to the plasma frequency.

## 2.5. Evaluation of the adsorption free energy

To evaluate the average adsorption energy for a protein globule on a NP we scan the configurational space (i.e. all possible orientation in which a protein can adsorb on the surface of the NP) by a systematic rotation of the protein and calculate the Boltzmann-averaged energy. There are three degrees of freedom (DOF) that must be scanned [14]. The orientation of the protein can be described by a vector from the centre of mass (COM) to an arbitrary point of the molecule. It is characterized by two angles:  $\phi$  and  $\theta$  and by rotating the molecule an angle  $-\phi$  about the  $z$ -direction and then by an angle  $-\theta + 180^\circ$  about the  $y$ -axis will make the position vector point towards the surface (along the negative  $z$ -axis). The third DOF is the distance from the COM to the closest point of the surface,  $d_{\text{COM}}$ . Instead of obtaining the actual adsorption free energy by calculating the PMF for all orientations and conformations of the protein molecule, we calculate a composite energy  $U$ , which is the sum of all the pairwise interactions between the surface and the AAs. In the energy  $U$ , some of the interactions (van der Waals) are represented by the potential energy while the other ones (electrostatic and surface PMFs) include ensemble averages over positions of water molecules, ions, and segment orientations, and in this sense represent free energy of adsorption. For each

configuration  $(\phi_i, \theta_j)$ , the total energy is calculated as a function of distance of the COM,  $U(D, \phi_i, \theta_j)$ , to the surface for the case of a slab or to the centre of the NP for the case of a NP. Following a similar approach as in Kokh *et al* [21], and denoting the reaction coordinate  $D = z$ , the mean interaction energy between a protein with orientation  $(\phi_i, \theta_j)$  and a spherical NP is given by:

$$E(\phi_i, \theta_j) = -k_B T \ln \left[ \frac{3}{(R+a)^3 - R^3} \int_R^{R+a} D^2 \exp\left(-\frac{U(D, \phi_i, \theta_j)}{k_B T}\right) dD \right], \quad (20)$$

where  $a = a(\phi_i, \theta_j)$  is the maximum interaction distance from the centre of mass of the protein to the surface for the given orientation. The mean adsorption free energy is obtained by averaging this interaction energy over the distribution of orientations

$$E_{\text{ad}} = \frac{\sum_i \sum_j P_{ij} E(\phi_i, \theta_j)}{\sum_i \sum_j P_{ij}}. \quad (21)$$

The weighting factor  $P_{ij}$  in equation (21) is given by:

$$P_{ij} = \sin \theta_j \exp\left(-\frac{E(\phi_i, \theta_j)}{k_B T}\right). \quad (22)$$

## 2.6. Simulation settings

We have used AWT-MetaD to compute PMFs of AA residues at a gold surface. There are 20 naturally occurring AAs. We have also considered two protonated versions of histidine (HID and HIE) so there are 22 biomolecules in total. The backbone fragments of amino acids were terminated by neutral  $\text{NH}_2$  and  $\text{COH}$  groups in order to mimic their behaviour in a long peptide chain. AAs were described by the AMBER03 force field [1]. Considering gold, we note that the face-centred (FCC) elementary metals are not dominated by Coulomb interactions, and the atomistic interactions can be well described by the Lennard-Jones potential alone. Thus, these simulations are performed for an uncharged surface, noting that the effects of an overall charge of the NP can be included in equation (1) if required. Here we used parameters from the force field by Heinz *et al* [2] which are  $\sigma_{\text{AU}} = 2.629 \text{ \AA}$  and  $\varepsilon_{\text{AU}} = 22.149 \text{ kJ mol}^{-1}$ . We used Lorentz–Berthelot combination rules to describe interaction of gold with water and AA. For water, TIP3P model was used.

All computations have been carried out in a rectangular periodic cell of the size  $24.4 \times 24.4 \times 80 \text{ \AA}^3$ . A charge-neutral gold slab composed of  $5 \times 5 \times 5$  unit cells of the gold crystal FCC structure, periodic in  $xy$  directions and cut in  $z$ -direction along (100) surface, was inserted into the simulation cell. The interaction cut-off was set at  $r_c = 1.6 \text{ nm}$ . The rest of the system was filled by water (about 800 molecules) and the adsorbing molecule. In the case of charged AAs, a counterion ( $\text{Na}^+$  or  $\text{Cl}^-$ ) was added to the system. The starting configuration was first energy minimized for 10 000 steps using the steepest gradient method. Then the system was equilibrated in  $NVT$ -ensemble simulations for 20 ps and in anisotropic  $NPT$  ensemble for 200 ps with time step 1 fs. The temperature was set to 300 K and pressure to 1 bar. Production simulations with metadynamics were carried out in  $NVT$  conditions, with established in preliminary  $NPT$  simulations volume, for 500 ns. A velocity-rescale thermostat with relaxation time 1 ps was used to ensure correct ensemble fluctuations. The Particle-Mesh Ewald method was used to treat long-range electrostatic interactions. Lennard-Jones interactions were truncated at  $10 \text{ \AA}$ . The motion of the centres of mass of material and solvent were removed separately to avoid artificial flow of the system through periodic boundaries. Gromacs v.5.1 [3] with PLUMED plugin v 2.3 [4] were used in all simulations.

In the metadynamics setup, the initial height of Gaussians was set to  $1 \text{ kJ mol}^{-1}$  with the deposition each 500 steps. The nearest distance between amino acid centre of mass and gold atoms was used as a collective variable. The adaptive well-tempered version of metadynamics (AWT-MetaD) was used in which width of the Gaussians is dynamically tuned to the local free energy landscape. The bias factor of the well-tempered metadynamics was set to 0. In order to reduce time spent by the adsorbate in the bulk solvent far from the surface, visiting of such states was prevented by a soft wall potential:

$$U_{\text{wall}}(z) = k(z - a)^4 \quad (23)$$

with the force constant  $k = 40 \text{ kJ mol}^{-1} \text{ \AA}^4$  and  $a = 1.8 \text{ nm}$ . Statistical analysis carried out over 100 ns pieces of the trajectories showed that the uncertainty of the resulting potentials of mean force was always within  $1 \text{ kJ mol}^{-1}$ .

The adsorption energy calculations described in section 2.5 are implemented in a C++ program UnitedAtom which handles the summation of the interaction potential over all AA beads and the numerical integration of equation (20) and will be described in more detail elsewhere. Briefly, the potential as a function of distance is tabulated for each bead, and this data used together with the known structure of the protein to tabulate the total potential as a function of the distance from the centre of mass of the protein to the NP. The calculations for this work have been performed for a set of blood serum proteins on gold NPs of radii ranging from 5 to 250 nm. We sampled protein orientations  $\phi$  from  $0^\circ$  to  $355^\circ$  in steps of  $5^\circ$  and  $\theta$  from  $0^\circ$  to  $175^\circ$  in steps of  $5^\circ$ . Energies were thus evaluated for 2592 different orientations. A more detailed account of the rotation procedure is given in [14, 22]. To improve the sampling of orientations, the energy at each nominal orientation is calculated as the average of 16 randomly-sampled orientations uniformly distributed in the interval  $[\phi, \phi + 5^\circ]$  and  $[\theta, \theta + 5^\circ]$ ; note that for the evaluation of equation (21) this requires offsetting the value of  $\theta$  by  $2.5^\circ$  to ensure it correctly represents the average value of  $\theta$  sampled.

### 3. Results

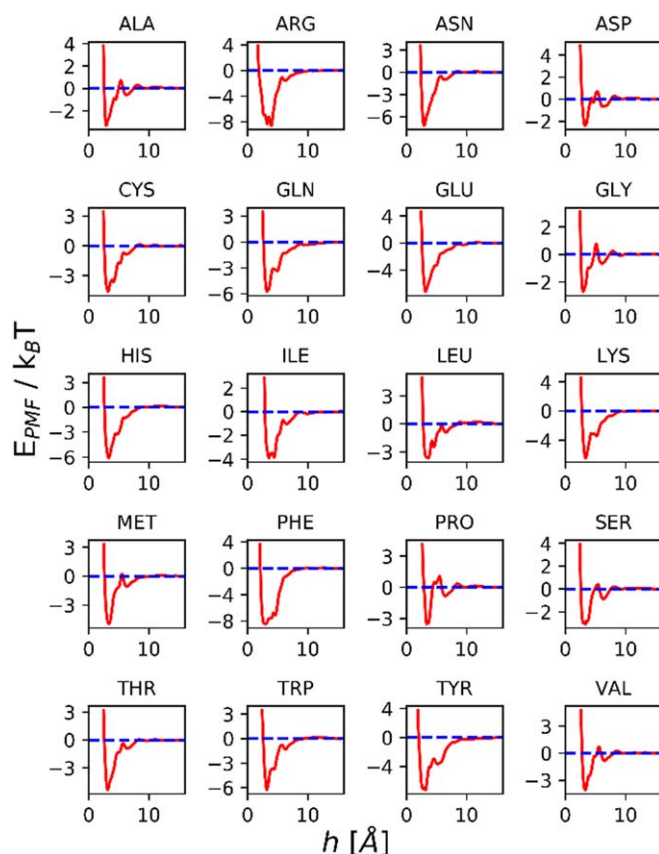
#### 3.1. Evaluation of the interaction parameters and preparation of the configurations

The model has been validated for a set of common globular proteins on Au NPs. Table 1 summarizes the Hamaker constants for AA-Au interaction in water calculated from equation (13), the UV absorption frequencies for AA are taken from [23] and the refractive indices from the procedure described in [24].

#### 3.2. Potentials of mean force and the mean adsorption free energies for AA side chains

In figure 4, we show the PMFs for AA side chains as obtained from atomistic simulations. Although all of them are attractive, they exhibit different structures, which can be grouped into several major categories. PMFs for ALA, ASP, GLY, MET, PRO, SER, VAL contain a well-expressed maximum after the primary minimum at  $3 \text{ \AA}$ , and a further minimum at  $6\text{--}7 \text{ \AA}$ . The maximum corresponds to a preference to have a water monolayer in between. Another group of potentials is purely attractive or has only a tiny maximum: these include ARG, ASN, CYS, GLU, GLN, HIS, ILE, LEU, LYS, PHE, THR, TYR, TRP.

The mean adsorption energies for the AA side chains based on these PMFs i.e., excluding the core potential, are collected in table 2. The figures range from  $-2.96k_B T$  for ASP to  $-17.75k_B T$  for PHE. The strongest attraction is predicted by our model for charged ARG, GLU and hydrophobic PHE, TYR, and ASN. These figures are roughly in agreement with



**Figure 4.** PMFs for interaction of AA side chain with a flat 100 FCC Au surface as obtained from atomistic simulations.

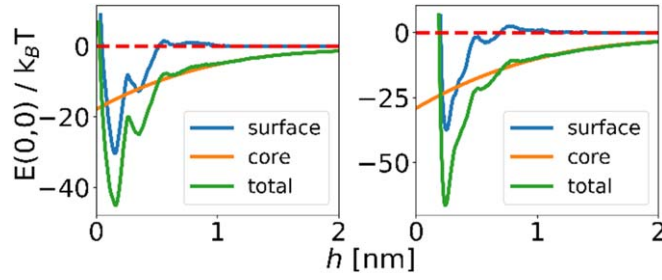
**Table 1.** Hamaker constants for AA-Au interaction through water according to equation (14). All values are reported in units of  $10^{-20}$  J.

|       |       |       |       |       |       |       |       |
|-------|-------|-------|-------|-------|-------|-------|-------|
| ARG   | HIS   | LYS   | ASP   | GLU   | SER   | THR   | ASN   |
| 29.70 | 30.57 | 28.78 | 30.10 | 29.49 | 29.40 | 29.10 | 29.88 |
| CYS   | GLY   | PRO   | ALA   | VAL   | ILE   | LEU   | MET   |
| 30.08 | 29.44 | 28.73 | 28.84 | 28.51 | 28.50 | 28.41 | 29.43 |
| GLN   | PHE   | TYR   | TRP   | HIP   | HID   | HIE   |       |
| 29.94 | 29.96 | 29.92 | 30.21 | 30.08 | 30.08 | 30.08 |       |

those reported in [25], where  $E_{ad}$  of  $-15 \text{ kJ mol}^{-1}$  to  $-50 \text{ kJ mol}^{-1}$  (ca.  $-6k_B T$  to  $-20k_B T$ ) have been reported for gold 111 surface. The highest attraction energies reported in [25] were found for TRP, PHE, TYR, and MET.

### 3.3. Mean adsorption free energies for selected proteins on NPs of different sizes and surface charges

Next, we wish to calculate the adsorption energy of proteins on the nanoparticle, including the contributions from the surface interaction described by the PMFs, the surface charge of the



**Figure 5.** Interaction energy  $E$  of the  $\phi = 0, \theta = 0$  orientation of HSA on  $R = 5$  nm (left) and  $R = 50$  nm (right) Au NP as a function of the distance of closest approach  $h$ , showing the total potential and the contributions from the surface and core.

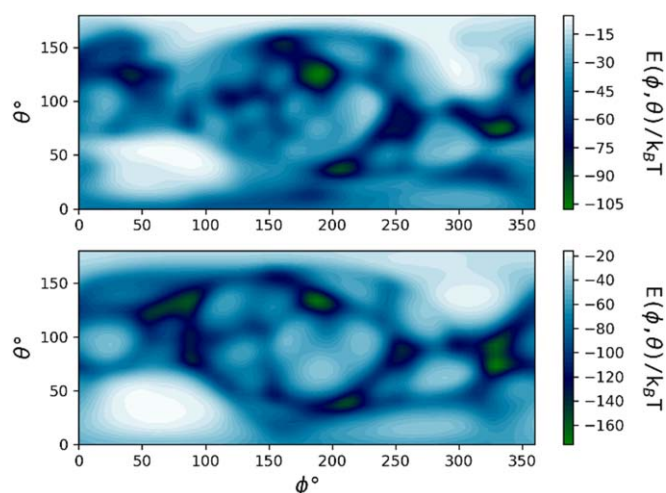
**Table 2.** Mean adsorption free energies for AA side chains on a flat gold surface at  $T = 300$  K, calculated from the surface PMFs. All values are reported in units of  $k_B T$ .

|        |        |        |        |        |       |       |
|--------|--------|--------|--------|--------|-------|-------|
| ARG    | HIS    | LYS    | ASP    | GLU    | SER   | THR   |
| -17.39 | -11.20 | -11.69 | -2.29  | -13.63 | -4.34 | -9.26 |
| CYS    | GLY    | PRO    | ALA    | VAL    | ILE   | LEU   |
| -7.62  | -2.98  | -5.02  | -4.23  | -6.17  | -6.54 | -5.69 |
| GLN    | PHE    | TYR    | TRP    | ASN    | MET   |       |
| -10.43 | -17.75 | -14.32 | -11.09 | -13.72 | -8.55 |       |

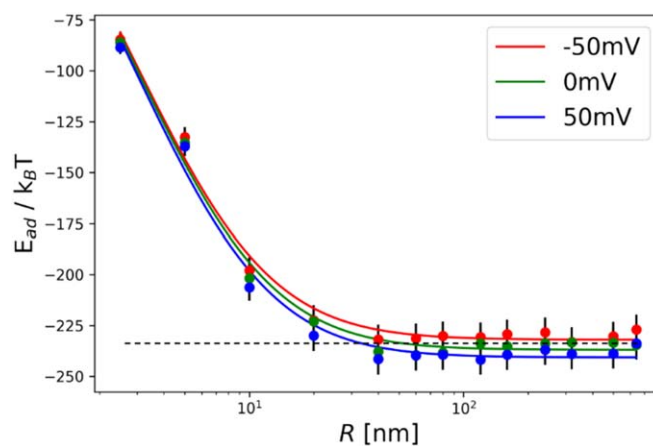
NP and the core potential. The procedure to do so is outlined in section 2.5 and the parameters for these calculations in section 2.6. In brief, given the orientation of the protein ( $\phi_i, \theta_j$ ) and the distance of the protein from the NP, the total interaction potential is evaluated using equation (1) by summing over the contribution from each bead in the protein. Figure 5 below shows the total potential for one such orientation of human serum albumin protein (HSA) on 5 and 50 nm gold NP with  $q_{\text{NP}} = 0$ . The contributions from the core and the surface have been separated to illustrate the difference in their behaviour. The contribution from the NP core is evident in both NP sizes, however it drastically increases as NP radius is increased and eventually becomes the dominant contribution. For the smaller NP we observe a highly structured short-range part of the potential, which includes a contact minimum of about  $-40k_B T$ , then a maximum that corresponds to a hydration layer between the protein and the NP surface and a much shallower secondary minimum. These effects are present but less visible in the larger NP due to the increased contribution from the core potential, which also leads to an increased depth of the contact minimum to  $-70k_B T$ .

Evaluating equation (20) using this interaction potential produces the binding energy for a specific orientation of the protein, denoted  $E(\phi, \theta)$ . Figure 6 shows  $E(\phi, \theta)$  calculated for HSA on 5 and 50 nm Gold NPs over a range of orientations obtained by systematically scanning the angles ( $\phi, \theta$ ). We see that the binding energy changes in magnitude with the NP size as the large NP produces deeper minima. Moreover, the map for the 50 nm particle contains more features: in addition to the minima at  $(340^\circ, 70^\circ)$  and  $(200^\circ, 140^\circ)$  for the 5 nm NP, we find a preferred orientation at  $(70^\circ, 120^\circ)$  and another minimum at  $(340^\circ, 80^\circ)$ .

The orientation-specific binding energies are then ensemble-averaged using equation (21) as discussed in section 2.5 to produce the mean adsorption energy  $E_{\text{ad}}$ . The result for HSA as



**Figure 6.** Orientation-specific binding energy  $E$  of HSA on  $R = 5$  nm (top) and  $R = 50$  nm (right) Au NP.



**Figure 7.** Orientation-averaged binding energy calculated for protein HSA on a gold nanoparticle as a function of the nanoparticle radius  $R$  for 3 different surface potentials ( $\psi_s = -50, 0$  and  $50$  mV). The solid lines are guide to the eye and the dashed line represents the binding energy for a neutral planar surface.

a function of  $R$  is shown in figure 7. Note that the binding becomes stronger as the size of the nanoparticle increases until it converges to the value obtained for binding to a planar surface.

### 3.4. Ranking proteins by adsorption affinity

To illustrate the power of our approach, we performed calculation of mean adsorption energies for several globular proteins, most of them present in blood serum (see table 3). The results presented here are for a neutral gold surface, but we note that varying the charge of the NP, so that the surface potential is  $\pm 50$  mV, produces only minor changes in the results as

**Table 3.** Mean adsorption free energies for selected globular proteins on neutral gold NPs at  $T = 300$  K. All values are reported in units of  $k_B T$ . The proteins are ranked by ascending adsorption energy on the 50 nm NP.

| Protein                    | PDB ID | $E_{ad}/k_B T$ | $E_{ad}/k_B T$ |
|----------------------------|--------|----------------|----------------|
|                            |        | $R = 5$ nm     | $R = 50$ nm    |
| Surfactant protein, SP-D   | 3DBZ   | -87.60         | -148.67        |
| Ubiquitin                  | 1UBQ   | -123.50        | -149.15        |
| Haemoglobin                | 5NI1   | -132.32        | -195.92        |
| Myoglobin                  | 3RGK   | -165.10        | -205.91        |
| Immunoglobulin gamma, IgG1 | 5JII   | -172.90        | -210.85        |
| Immunoglobulin gamma, IgA1 | 3H09   | -134.85        | -221.28        |
| Transferrin                | 2HAV   | -113.88        | -230.73        |
| Macroglobulin alpha-2, A2M | 4ACQ   | -120.29        | -233.76        |
| Immunoglobulin gamma, IgE  | 5MOL   | -132.42        | -236.76        |
| Human serum albumin, HSA   | 1N5U   | -136.04        | -239.26        |
| Alpha-1-antitrypsin, A1A   | 5IO1   | -180.41        | -263.91        |

can be seen in figure 7. The obtained energies are very high, corresponding to essentially irreversible adsorption. These figures are, however, in agreement with simulations and experiments predicting the adsorption energies on gold in the range of  $10^2 k_B T$  [11, 21]. The strongest binding is predicted for a hydrophobic surfactant protein Alpha-1-antitrypsin, A1A. The adsorption to large NPs (50 nm) is visibly stronger as compared to small NPs due to the large contribution of the long-range van der Waals attraction. We note that the ranking is different for different NP sizes, the most significant changes seen for Macroglobulin Alpha-2, A2M: it is ranked 4/11 for the large NP but 10/11 for the small NP. For Ubiquitin, the rank is 10/11 for the large NP but 8/11 for the small NP. Ubiquitin has the smallest increase in  $E_{ad}$  between the 5 nm NP and the 50 nm NP of the 11 tested proteins. Similar change in ranking for Ubiquitin has been reported in [6], where it was found more prevalent in the corona of small gold NPs compared with the corona of large NPs. These differences are related to the number of contacts the protein globules can make with the NPs of different curvatures. We note also that the ranking is sensitive to the NP surface charge. The latter has also been reported in [6].

#### 4. Discussion

The methodology presented here opens far-reaching prospects for materials modelling and characterization. It allows one to evaluate adsorption energies for multiple proteins on NPs of different sizes and surface charges in a systematic multiscale approach where the coarser representation is parameterized using detailed atomistic simulations and materials constants. A number of protein adsorption energies and rankings, as well as their dependence on the NP size are reported here for the first time. The numbers are in agreement with the data reported in more detailed and demanding calculations [11, 21, 25].

The most computationally intensive part of our multiscale scheme is the calculation of the PMFs for AAs, during which the effects of water and AA orientation are integrated out. It very much depends on the availability of accurate force fields, which may require additional optimization taking into account the surface properties such as hydration enthalpy.

At the same time, the adsorption energies calculated within our model may have significant systematic errors due to strong approximations made in the model. We suppose that the main source of errors is the additivity approximation combined with the averaging over the orientations of the AA side chains. In reality, only one specific conformation and orientation with respect to the surface may be preferable for each AA, depending on its position within the folded protein. This assumption might affect the surface part of the energy  $(U_s)_i$ , where a higher energy (weaker adsorption) should be generally expected on average, as not all AA would be able to assume the most favourable orientation due to restrictions from the neighbouring AAs inside the protein globule. The next possible source of errors is in the charge regulation, which is neglected in the model. We expect that the assumption of constant charge on the AAs and the use of linearised screened Coulomb model for the electrostatics would generally overestimate the charge effects in the real systems. Finally, we neglect the possible changes in the protein conformation due to the interaction with the NP. This, on the contrary, should lead to underestimation of the adsorption energies for all proteins as in reality, the proteins would be able to reach a more favourable conformation at the surface by flexing themselves. Other minor artefacts are introduced by fixing the protein in its crystal structure, because of which some NP sizes may be singled out as they match the curvature of the rigid protein globule. Some additional challenge is posed by the evaluation of van der Waals forces, where we assume that the AA and the core of the NP are separated only by the solvent. While we see that van der Waals forces may play a major role in the adsorption, it is challenging to improve the accuracy of this contribution in particular because the Hamaker constants are not always easily accessible, and literature may give quite a variety of values for the same material.

Yet, we expect that all these factors play a minor role due to the characteristic scales of the problem. First, the strong adhesion makes the protein adsorption process effectively irreversible, and thus the question of conformational changes is irrelevant. The molecules, once stuck to the NP surface, would not detach in reasonable time. This is particularly valid for materials like gold, which produce strong van der Waals attractions. Next, variability and diversity of the protein content of the biological fluids (e.g. blood serum) makes it impossible to guarantee the presence of any specific protein in the NP corona, so that only adsorption statistics can be reliably predicted. This, however, is handled well by our multiscale model that includes all the major factors systematically evaluated for the series of proteins and NPs and produces correct ranking of proteins by adsorption affinity. We expect that it is the ranking and concentration in solution and not the absolute values of the adsorption energies that determine the relative abundance of specific molecules in the corona.

We should also stress the significance of systematic evaluation of material characteristics relevant for bionano interactions that our model provides. Here, we introduced several dozens of novel material descriptors (Hamaker constants, AA adsorption energies, protein adsorption energies) that can be used to develop predictive QSAR models for biological activity of NPs.

## 5. Summary

We presented a bottom-up multiscale modelling methodology for quantitative characterization of bio-nano interface. In particular, it can be used to describe NP corona by estimating the adsorption energy and ranking the proteins by adsorption affinity. The calculations can be applied in a uniform fashion to a large number of biomolecules and NP sizes and shapes in a short time, at the cost of making the approximations that pairwise additivity holds and that charge regulation and changes in the protein conformation can be neglected. The possibility



to effectively scan over multiple biomolecular segments is especially useful for the development of bio- and nanoinformatics models, required to predict NP functionality or toxicity base on their physicochemical properties. In addition to this, our method delivers preferred orientations for proteins on a specific NP, thus systematically taking into account the NP shape and size. We also demonstrated how the effect of NP core can be included in the calculation, which makes the size-dependence of the interaction more credible.

## Acknowledgments

This research was funded by EU Horizon2020 Grant No. 686098 (SmartNanoTox project) and by Science Foundation Ireland Grant No. 16/IA/4506 (Bio-Interface project).

## ORCID iDs

Hender Lopez  <https://orcid.org/0000-0003-1083-6234>

Vladimir Lobaskin  <https://orcid.org/0000-0002-5231-0639>

## References

- [1] Cedervall T, Lynch I, Foy M, Berggard T, Donnelly S, Cagney G, Linse S and Dawson K 2007 Detailed identification of plasma proteins adsorbed on copolymer nanoparticles *Angew. Chem. Int. Ed.* **46** 5754–6
- [2] Landsiedel R R, Ma-Hock L, Kroll A, Hahn D, Schnekenburger J, Wiench K and Wohlleben W 2010 Testing metal-oxide nanomaterials for human safety *Adv. Mater.* **22** 2601–27
- [3] Lundqvist M, Sethson I and Jonsson B H 2004 Protein adsorption onto silica nanoparticles: conformational changes depend on the particles curvature and the protein stability *Langmuir* **20** 10639–47
- [4] Lundqvist M, Sethson J and Jonsson B H 2005 Transient interaction with nanoparticles ‘freezes’ a protein in an ensemble of metastable near-native conformations *Biochemistry* **44** 10093–9
- [5] Karajanagi S S, Vertegel A A, Kane R S and Dordick J S 2004 Structure and function of enzymes adsorbed onto single-walled carbon nanotubes *Langmuir* **20** 11594–9
- [6] Walkey C D, Olsen J B, Song F, Liu R, Guo H, Olsen D W H, Cohen Y, Emili A and Chan W C W 2014 Protein corona fingerprinting predicts the cellular interaction of gold and silver nanoparticles *ACS Nano* **8** 2439–55
- [7] Kamath P, Fernandez A, Giralt F and Rallo R 2015 Predicting cell association of surface-modified nanoparticles using protein corona structure—activity relationships (PCSAR) *Curr. Top. Med. Chem.* **15** 1930
- [8] Saptarshi S R, Duschl A and Lopata A L 2013 Interaction of nanoparticles with proteins: relation to bio-reactivity of the nanoparticle *J. Nanobiotechnol.* **11** 26
- [9] Rahman M, Laurent S, Tawil N, Yahia L and Mahmoudi M 2013 *Protein-Nanoparticle Interactions (Springer Series in Biophysics)* vol 15 (Berlin: Springer)
- [10] An D, Su J, Li C and Li J 2015 Computational studies on the interactions of nanomaterials with proteins and their impact *Chin. Phys. B* **24** 120504
- [11] Brancolini G, Kokh D B, Calzolari L, Wade R and Corni S 2012 Docking of ubiquitin to gold nanoparticles *ACS Nano* **6** 9863–78
- [12] Ding F, Radic S, Chen R, Chen P, Geitner N, Brown J and Ke P 2013 Direct observation of a single nanoparticle–ubiquitin corona formation *Nanoscale* **5** 9162–9
- [13] Khan S, Gupta A and Nandi C 2013 Controlling the fate of protein corona by tuning surface properties of nanoparticles *J. Phys. Chem. Lett.* **4** 3747–52
- [14] Lopez H and Lobaskin V 2015 Coarse-grained model of adsorption of blood plasma proteins onto nanoparticles *J. Chem. Phys.* **143** 243138
- [15] Israelachvili J N 2011 *Intermolecular and Surface Forces* 3rd edn (Israel: Academic)

- [16] Reith D, Pütz M and Müller-Plathe F 2003 Deriving effective mesoscale potentials from atomistic simulations *J. Comput. Chem.* **24** 1624–36
- [17] Izvekov S and Voth G A 2005 A multiscale coarse-graining method for biomolecular systems *J. Phys. Chem. B* **109** 2469–73
- [18] Noid W G, Chu J W, Ayton G S, Krishna V, Izvekov S, Voth G A, Das A and Andersen H C 2008 The multiscale coarse-graining method: I. A rigorous bridge between atomistic and coarse-grained models *J. Chem. Phys.* **128** 244114
- [19] Duan Y *et al* 2003 A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations *J. Comput. Chem.* **24** 1999–2012
- [20] Brandt E and Lyubartsev A P 2015 Molecular dynamics simulations of adsorption of amino acid side chain analogues and a titanium binding peptide on the TiO<sub>2</sub> (100) surface *J. Phys. Chem. C* **119** 18126–39
- [21] Kokh D, Corni S, Winn P, Hoefling M, Gottschalk K and Wade R 2010 ProMetCS: an atomistic force field for modelling protein-metal surface interactions in a continuum aqueous solvent *J. Chem. Theory Comput.* **6** 1753–68
- [22] Sun Y, Welsh W and Latour R 2005 Prediction of the orientations of adsorbed protein using an empirical energy function with implicit solvation *Langmuir* **21** 5616–26
- [23] Glasel J A 1995 Validity of nucleic acid purities monitored by a260/a280 absorbance ratios *Biotechniques* **18** 62–3
- [24] Zhao H, Brown P H and Schuck P 2011 On the distribution of protein refractive index increments *Biophys. J.* **100** 2309–17
- [25] Hoefling M, Iori F, Corni S and Gottschalk K-E 2010 Interaction of amino acids with the Au(111) surface: adsorption free energies from molecular dynamics simulations *Langmuir* **26** 8347–51