



Multiscale modelling of biomolecular corona formation on metallic surfaces

Parinaz Mosaddeghi Amini*, Ian Rouse, Julia Subbotina and Vladimir Lobaskin

Full Research Paper

Open Access

Address:

School of Physics, University College Dublin, Belfield, Dublin 4, Ireland

Email:

Parinaz Mosaddeghi Amini* -
parinaz.mosaddeghiamini@ucdconnect.ie

* Corresponding author

Keywords:

all atomistic; aluminum; bionano interface; coarse grained model; lactose; milk protein; multiscale modelling; protein corona

Beilstein J. Nanotechnol. **2024**, *15*, 215–229.

<https://doi.org/10.3762/bjnano.15.21>

Received: 16 October 2023

Accepted: 16 January 2024

Published: 13 February 2024

This article is part of the thematic issue "Nanoinformatics: spanning scales, systems and solutions".

Associate Editor: K. Ariga



© 2024 Amini et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

In the realm of food industry, the choice of non-consumable materials used plays a crucial role in ensuring consumer safety and product quality. Aluminum is widely used in food packaging and food processing applications, including dairy products. However, the interaction between aluminum and milk content requires further investigation to understand its implications. In this work, we present the results of multiscale modelling of the interaction between various surfaces, that is (100), (110), and (111), of fcc aluminum with the most abundant milk proteins and lactose. Our approach combines atomistic molecular dynamics, a coarse-grained model of protein adsorption, and kinetic Monte Carlo simulations to predict the protein corona composition in the deposited milk layer on aluminum surfaces. We consider a simplified model of milk, which is composed of the six most abundant milk proteins found in natural cow milk and lactose, which is the most abundant sugar found in dairy. Through our study, we ranked selected proteins and lactose adsorption affinities based on their corresponding interaction strength with aluminum surfaces and predicted the content of the naturally forming biomolecular corona. Our comprehensive investigation sheds light on the implications of aluminum in food processing and packaging, particularly concerning its interaction with the most abundant milk proteins and lactose. By employing a multiscale modelling approach, we simulated the interaction between metallic aluminum surfaces and the proteins and lactose, considering different crystallographic orientations. The results of our study provide valuable insights into the mechanisms of lactose and protein deposition on aluminum surfaces, which can aid in the general understanding of protein corona formation.

Introduction

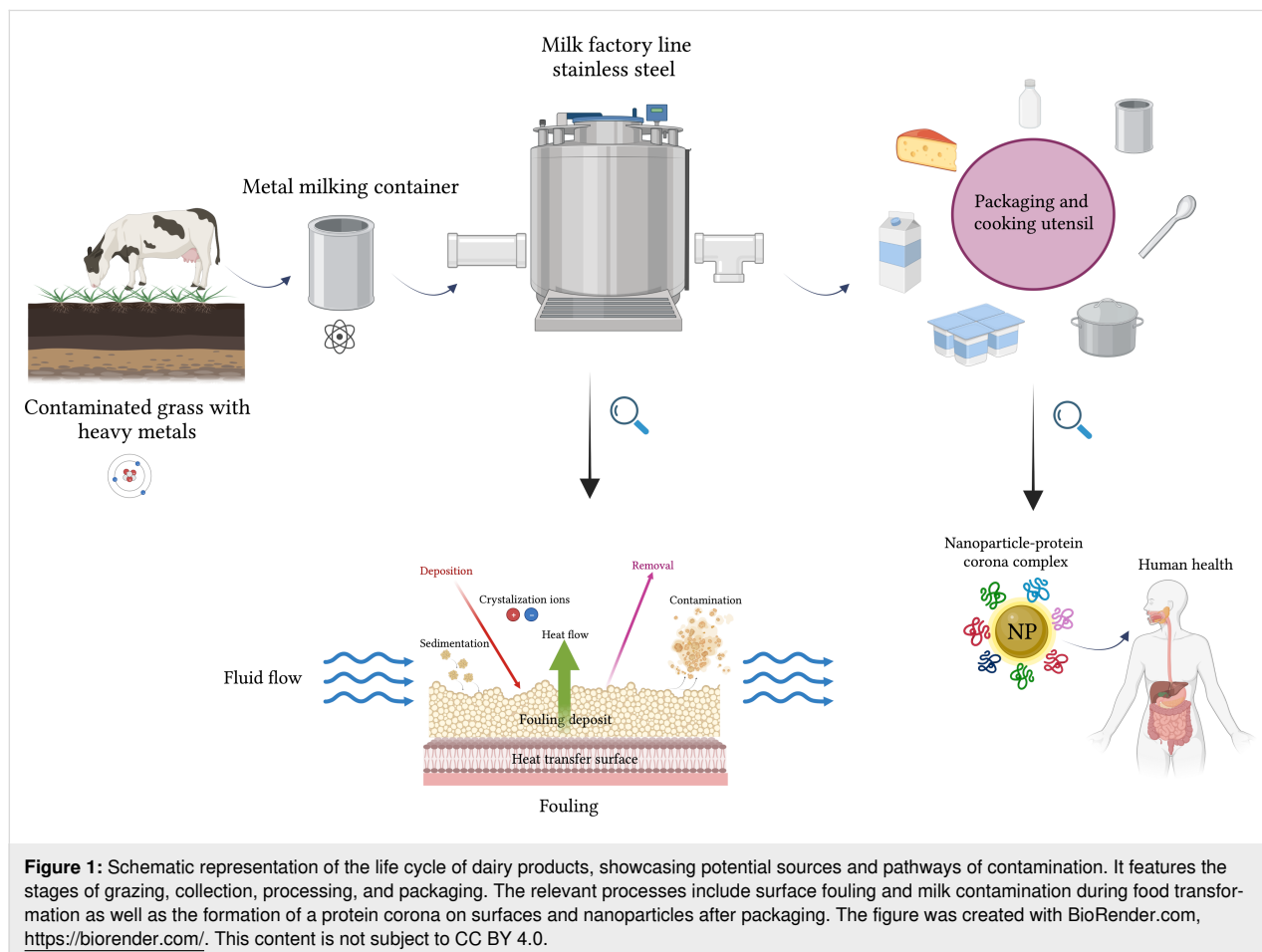
The interface between biological systems and engineered materials has gained significant attention in recent years because of its wide range of applications, spanning from food to medicine

and environmental science [1,2]. This interface plays a crucial role in ensuring the safety and quality of processed and packaged products. The selection of packaging materials and their

interaction with biological components have emerged as critical determinants impacting the preservation, shelf life, and overall acceptability of dairy products [3]. Consequently, the interface between biologically relevant molecules and nano-scale materials, such as aluminum, has become an increasingly important and intriguing area of research [4]. For long-term storage and preservation of prepared food, the choice of containers and utensils made from specific materials is essential [5]. For example, it was shown that ripened cheese and cheese spreads acquire a higher aluminum content as compared to other milk products [6]. Aside from wrapping and container packaging, aluminum has found a wide popularity in other applications, such as manufacturing of kitchen utensils, cosmetics, and components for medical and scientific equipment [7]. Figure 1 presents a schematic contamination cycle of dairy products, showcasing potential sources and pathways of aluminum pollution. It illustrates the journey of milk from a cow grazing on grass contaminated with heavy metals, highlighting the crucial role of metallic containers, metal-based equipment, and kitchen utensils in maintaining product integrity. The figure further demonstrates the potential to introduce heavy metal contamination, including iron and aluminum,

during processing and emphasizes the formation of a milk layer in form of a protein/lactose corona at the outer surface of macroscopic and micro- and nano-sized particulate after packaging. It also highlights the dynamic interactions at the bionano interface associated with potential human health hazards. Through biomolecule adsorption, change of conformation, and surface chemistry, foreign materials engage in a complex interplay of dynamic physicochemical interactions, kinetics, and thermodynamic exchanges that can lead to undesirable outcomes [1,8-10].

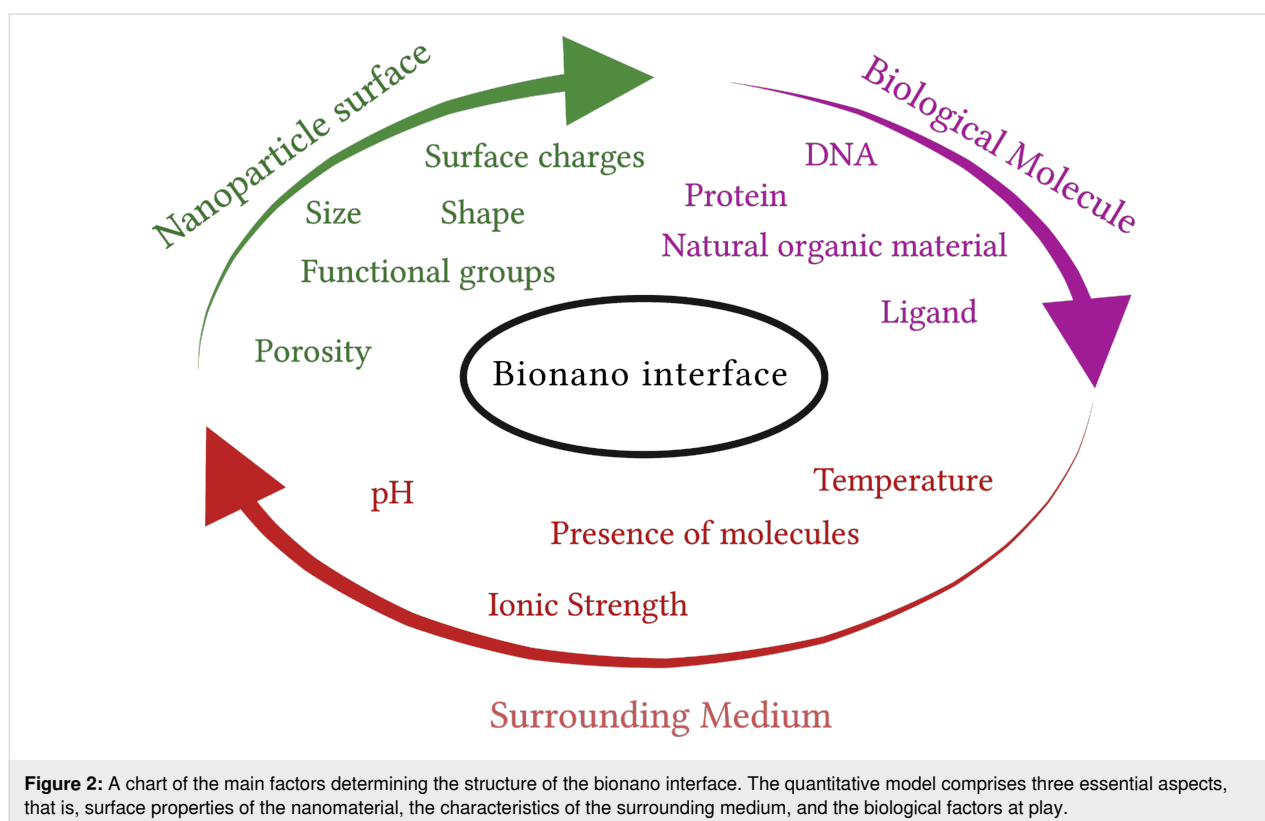
In a more general context, the importance in understanding the mechanism of bionano interactions arises from the increasing awareness and concerns regarding the safety of nanoparticles (NPs) in relation to human and animal health. The toxicity of NPs is closely linked to their chemical aggressiveness and varies with their physicochemical properties, including surface area, charge, and reactivity. Understanding the intricate interplay between these properties and the biological systems is vital for assessing and mitigating any potential adverse effects associated with exposure to NPs [11]. To advance in this field, it is crucial to comprehend the underlying forces and molecular



constituents that govern the interactions between biomolecules and metals. However, traditional safety assessment methods can be costly, time-consuming, and often involve animal studies. In this regard, *in silico* modelling offers a promising alternative that can predict the interactions of NPs with living organisms. By leveraging computational approaches, *in silico* modelling provides a humane and cost-effective means of obtaining the necessary information, thus aiding in the evaluation of NP safety and reducing reliance on animal experimentation [12–14]. Data-driven methods that rely on statistical analysis are employed for this purpose, particularly when sufficient data are available. These methods leverage the power of large datasets to identify patterns, trends, and correlations between metal properties and their interactions with biomolecules [15–18]. In recent years, researchers have focused on using physics-based models to understand the mechanisms underlying the formation of NP protein corona, a complex layer of biomolecules that surrounds NPs upon their exposure to biological fluids [19,20]. It is widely recognized that composition and configuration of the protein corona play a crucial role in determining the biochemical reactivity, sensitivity of NPs, as well as their cellular uptake and systemic transfer [21]. However, in order to develop predictive models, a deeper understanding of the interactions at the bionano interface and their relationship to material and protein properties is necessary. Gathering more information on these intricate interactions will facilitate the development of accurate

predictive models, thereby advancing our ability to assess the behavior and potential implications of NPs in biological systems. The bionano interface can be broken down into three interconnected components: (i) the surface of the NP, which is influenced by its physicochemical composition, (ii) the interface between the solid NP and the surrounding liquid environment, where notable changes occur upon interaction, and (iii) the contact zone between the solid–liquid interface and biological substrates (Figure 2) [22].

In this work, we study bionano interactions involving metallic aluminum and common dairy biomolecules, namely lactose and the six most abundant milk proteins [23]. The main objective of our analysis is to computationally quantify the relative binding of these proteins on zero-valent aluminum surfaces based on their energy of adsorption and orientation. We employ a three-level multiscale method (as shown in Figure 3) to calculate the energies of adsorption and the content of the corona for these proteins on the selected surfaces. In the section “Results and Discussion”, we provide a detailed explanation of the theoretical model developed to study the interaction between protein and lactose with metals, as well as the rationale behind the parameterization scheme used. Subsequently, we discuss the simulation results and analyze the individual adsorption affinities predicted for molecules representing the biological aspect of the interface, including amino acids (AAs), milk proteins, and



carbohydrates. Additionally, we examine the preferred orientations of these molecules upon adsorption and investigate the kinetics of competitive adsorption among the proteins and lactose, aiming to understand the process of protein deposition on metallic surfaces. Finally, the key insights gained from this study are summarized, highlighting the implications and potential applications of the findings.

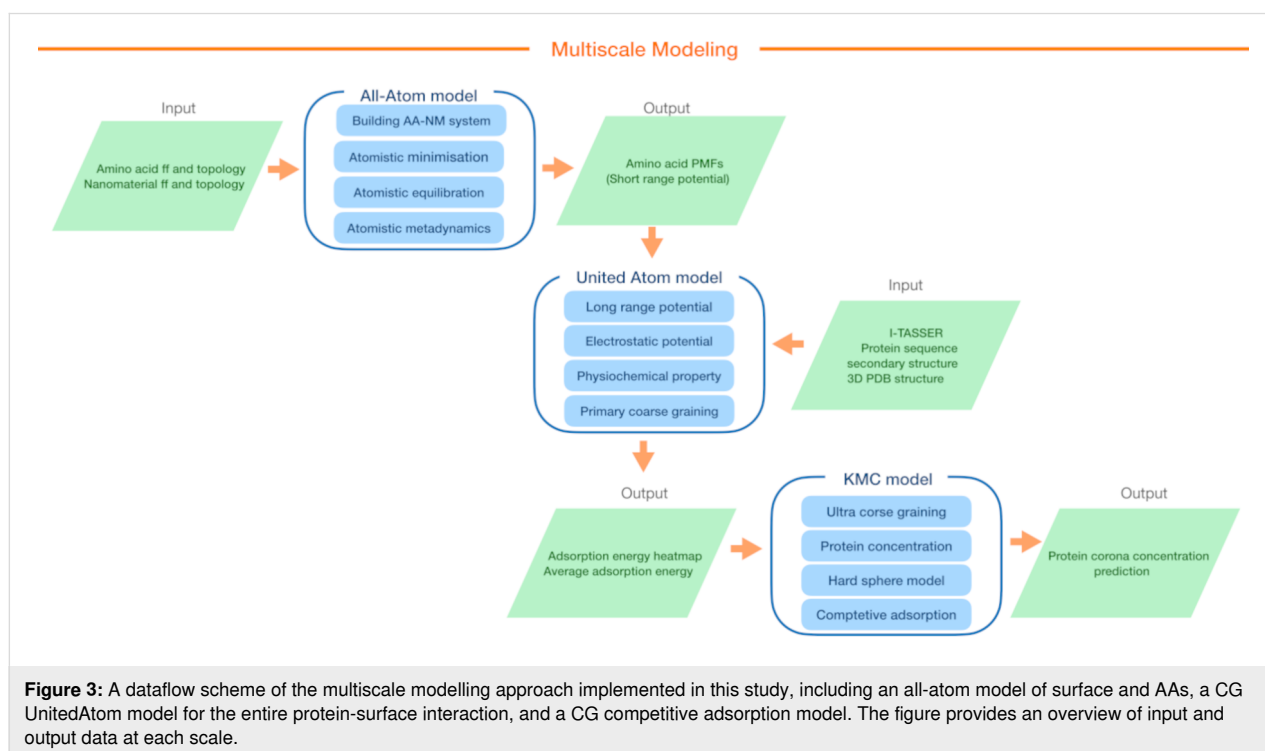
Results and Discussion

Here, we aim to predict the content of a biomolecular corona on a metallic aluminum surface. At the largest scale, our methodology employs a coarse-grained (CG) kinetic Monte Carlo (KMC) method [16] to simulate competitive adsorption of biomolecules onto the aluminum surface. To achieve this, we evaluate individual binding energies at various orientations (represented by heatmaps) for each selected protein immobilized on different fcc planes of the aluminum surface. These heatmaps for individual proteins are acquired through UnitedAtom (UA) simulations [24,25]. While the UA method has been parameterized for a range of rigid surfaces, including metals (Ag, Au, Cu, and Fe), oxides (TiO₂, SiO₂, and Fe₂O₃), carbonaceous NPs (graphene, carbon nanotubes, and carbon black), semiconductors (CdSe) [26], and polymers [27], it lacks the set of short-range potentials required for calculating milk protein-aluminum adsorption energies. Here, we compute potentials of mean force (PMF) for Al surfaces derived from explicit all-atom molecular dynamics simulations utilizing a previously established scheme [2,24,28]. These PMFs provide the input required to de-

termine the adsorption energies between milk proteins and aluminum surfaces by using multiscale UA CG model, spanning from the atomistic level of description to the complete mesoscale model of the corona. Figure 3 shows the parameterization and simulation workflow, outlining different stages and components involved in the study.

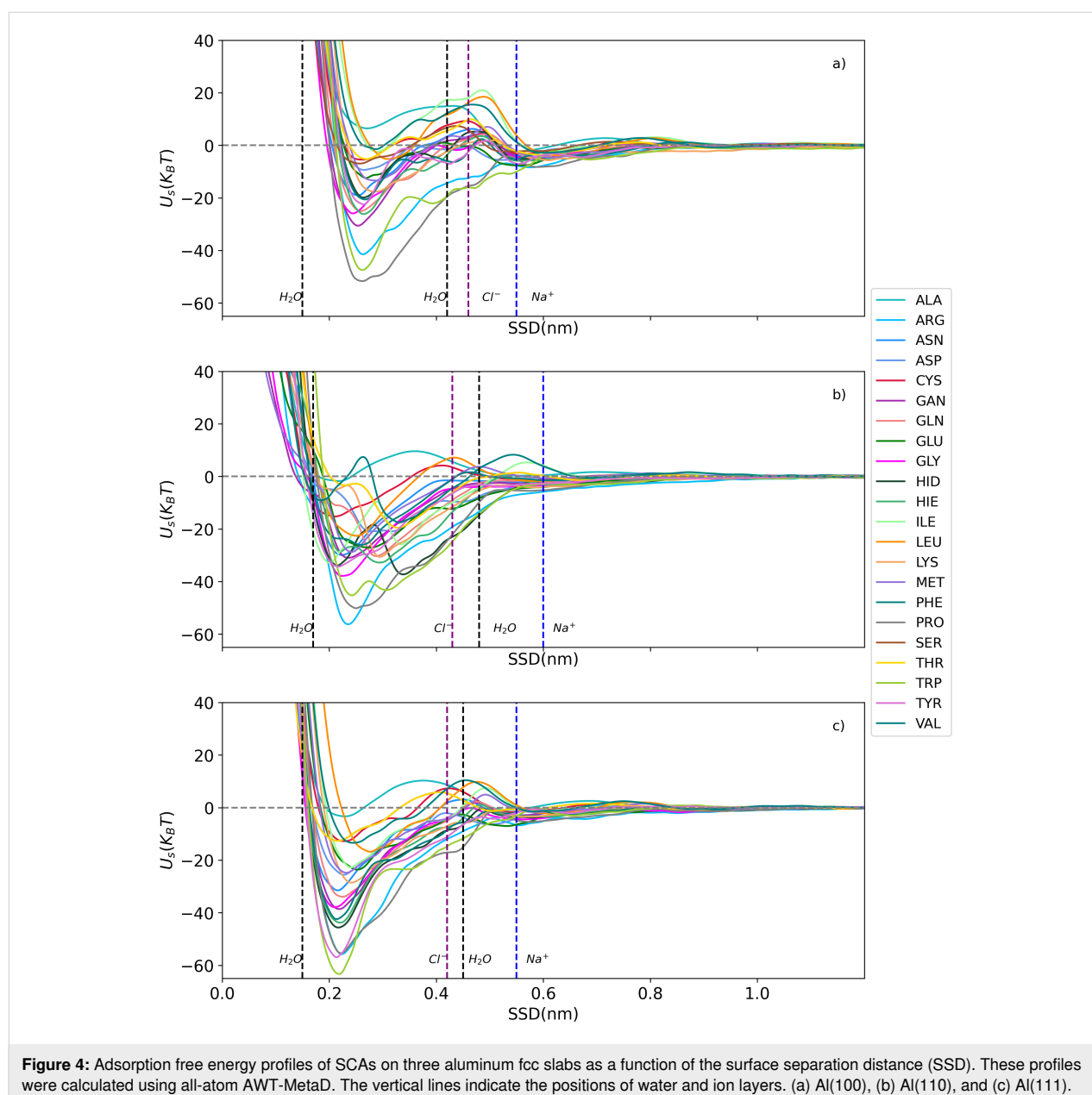
All-atoms short-range interaction modelling results

All-atom metadynamics simulations were conducted using GROMACS-2018.6 and PLUMED (PLUMED2-2.5.1.conda.5) software packages [29-31]. CHARMM-GUI/Nanomaterial Modeler was employed to construct the topology and force fields of three fcc surfaces of Al: (100), (110), and (111) [32]. The General Amber Force Field (GAFF) was utilized to model side-chains analogues (SCA) within the system [33,34]. The AMBER force field is a widely recognized and extensively validated force field that provides accurate descriptions of molecular systems [35]. We evaluated the short-range PMFs between 22 SCAs and an Al slab in a solvent environment comprising water and salt ions. The system's pH value was maintained at a neutral level, and the NaCl salt concentration was set to 150 mM, mimicking the overall ionic strength of milk and equivalent to one salt molecule per 10 nm³. The system underwent equilibration for 1.0 ns under constant pressure conditions at 1.0 bar and a temperature of 300 K, following the NPT ensemble, employing Berendsen weak coupling method [36]. Subsequently, a pre-equilibration phase was conducted for



10 nanoseconds within the NVT ensemble. For the short-range interactions, the cut-off distance was defined as 1.0 nm. In the adaptive well-tempered metadynamics (AWT-MetaD) simulations, the adsorption energy was calculated at a temperature of 300 K, a pressure of 1.0 bar, and a neutral pH within the NVT ensemble. Additionally, we measured the interaction energy as a function of surface separation distance (SSD) as a collective variable, enabling a comprehensive analysis of the AA-NP interactions. For a detailed explanation of the method used in this study, please refer to previous reports [2,24,28] where the method has been described in depth. Figure 4 and dataset [37] show the obtained free energy of adsorption in units of $k_B T$.

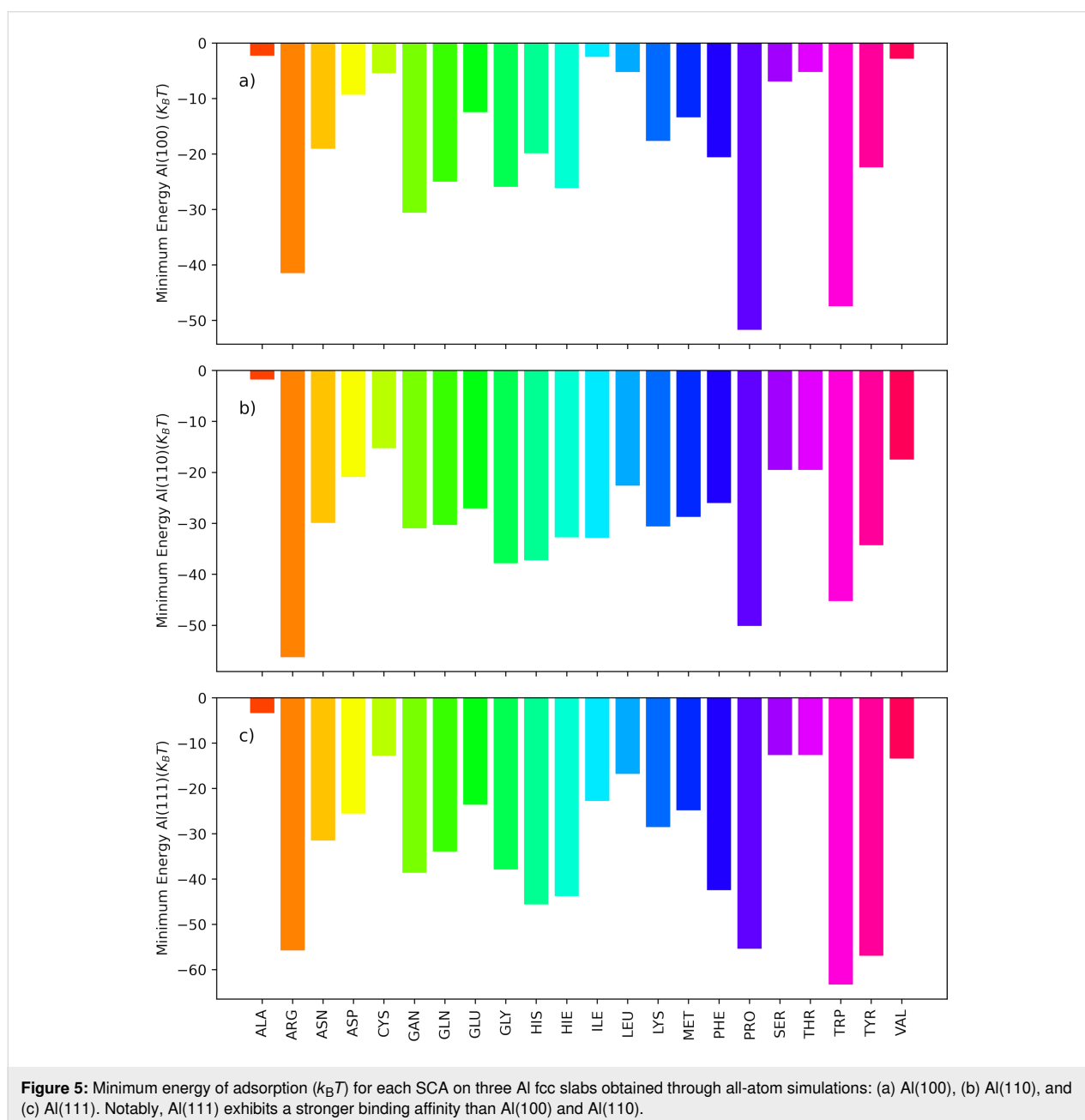
The water density profiles obtained from MD simulations for the slab–water system in the context of Al surfaces revealed characteristics that were previously observed for other simulated metallic surfaces [2,28]. The profiles exhibited two distinct regions with elevated water density located approximately 0.15–0.18 nm and 0.42–0.48 nm away from the aluminum surface. These regions corresponded to the first and second water layers adjacent to the metal surface, respectively (as depicted in Supporting Information File 1, Figure S1). Further examination of the ion density profiles indicated the presence of sodium ions within a range of 0.55–0.60 nm and chloride ions within a range of 0.42–0.46 nm from the Al surface. Notably, the positions of the chloride ions align closely with the second

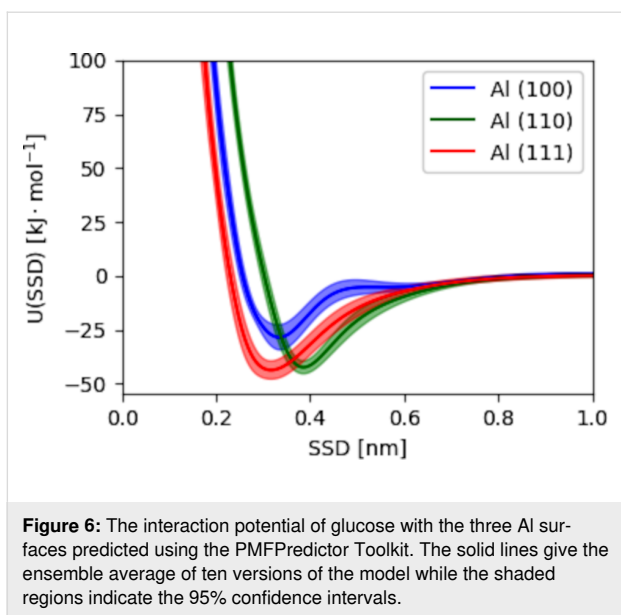


water layer, while sodium ions are located past this layer, as marked by the blue and purple vertical dashed lines in Figure 4. This alignment suggests that the chloride ions integrate into the network of water molecules comprising the second adlayer. Additionally, the analysis of the PMFs revealed a significant minimum at a distance of 0.21–0.25 nm. Figure 5 shows the minimum energy values obtained for each AA on different facets of the aluminum surface (100, 110, and 111) in a bar chart.

A comparison of the adsorption energies on aluminum and iron surfaces reveals distinct preferences for different AAs. On alu-

minum surfaces, ARG, PRO, TRP, TYR AAs show the strongest attraction ($-63.32k_B T$ to $-41.46k_B T$), followed by HIE, GLN, PHE, GAN ($-43.86k_B T$ to $-20.85k_B T$). VAL, THR, SER, CYS, ALA exhibit the weakest attraction ($-19.51k_B T$ to $-1.76k_B T$). On iron surfaces, charged and aromatic PRO, TYR, ARG, HIS AAs are strongly adsorbed ($-91.29k_B T$ to $-43.34k_B T$), while hydrophobic VAL, LEU, ALA AAs show a weaker adhesion ($-21.70k_B T$ to $2.86k_B T$) [2]. We also show the PMF for glucose with aluminum surfaces, used as the basis for a model of lactose, a sugar highly present in milk, as discussed later, computed using the PMFPredictor software in Figure 6 [38].





Protein–NP interactions

To further understand the adsorption energy and orientation of each individual protein, a primary coarse-graining step was performed. In this part, we use the UA model to predict the protein–NP binding energies. This model takes into account various factors, such as the material’s chemical composition, size, shape, surface roughness, charge, functionalization, and hydrophobicity, when constructing CG models for the bionano interface. The UA model simplifies the protein–NP interactions by representing proteins as rigid structures composed of 20 AA types, each represented by a single bead. This interaction is described through a short-range surface non-bonded potential (U_s^{nb}) (including van der Waals (vdW) repulsion and solvent effects), a long-range core vdW potential (U_1^{vdW}), and an electrostatic potential (U^{el}). Through interaction potentials for specific AAs with the NP, the overall interaction potential between the NP and the complete protein ($U_{\text{p-NP}}$) is expressed in a pairwise additive manner:

$$\begin{aligned}
 U_{\text{p-NP}} &= \sum_{i=1}^{N_{\text{AA}}} U_i(d_i(\theta, \phi)) \\
 &= \sum_{i=1}^{N_{\text{AA}}} U_i^{\text{el}}(d_i(\theta, \phi)) + \sum_{i=1}^{N_{\text{AA}}} U_i^{\text{nb}_s}(d_i(\theta, \phi)) \quad (1) \\
 &\quad + \sum_{i=1}^{N_{\text{AA}}} U_i^{\text{vdW}_1}(d_i(\theta, \phi)).
 \end{aligned}$$

The potential $U_{\text{p-NP}}$ depends on the distance d_i between the centers of mass of the NP and each AA in the protein. This distance is determined by the protein’s orientation with respect to

the NP’s surface, which is defined by two rotational angles (ϕ, θ) relative to the protein’s initial orientation. This initial orientation is set by performing a principle axis transformation such that the axis associated with the smallest moment of inertia is aligned to the z axis and the second smallest to the y axis, that is, the z axis is now typically associated with the greatest extent of the protein. Since this does not uniquely specify the orientation, further rotations of 180° are then applied if necessary such that the electric dipole moment is positive along these two axes. This produces a convenient reference state by which other orientations are defined. The specific orientation (ϕ, θ) is generated by applying a rotation of $-\phi$ around the z axis followed by a rotation of $180^\circ - \theta$ around the y axis. The short-range surface non-bonded potentials are extracted from AWT-MetaD simulations, which were described in the section “All-atoms short-range interaction modelling results”. The Hamaker technique is used to approximate the long-range term that results from the vdW forces working through the aqueous medium between the NP core and the i -th AA. The electrostatic interaction between the NP and AA is represented by the screened Coulomb potential. More comprehensive information about the theoretical aspects of the UA model can be found in our previous publications [2,25,28,39,40]. The output of the UA simulations contains a collection of rotational configurations and their corresponding $E(\theta_k, \phi_i)$ values. By employing Boltzmann averaging and weighting factors based on the potential energy as a function of distance for each angle, we calculate the average adsorption energy of these configurations. Using this approach, we evaluate the adsorption energies of the entire proteins on aluminum surfaces. To predict the three-dimensional (3D) structures of proteins, we utilize the I-TASSER (Iterative Threading ASSEMBly Refinement) 5.1 software [41], which uses the protein’s AA sequences as an input.

For this study, we have chosen six representative cow milk proteins and lactose, which constitute most of the non-fat milk solids. Table 1 displays properties of the chosen compounds. It includes their UniProt IDs, molecular weights, charges, and the number of AAs in each protein. The charge data was determined through the PROPKA method [42,43] at a pH of 7.0. We model the lactose molecule as a pair of glucose beads; it does not possess a UniProt ID or a count of AA residues. We estimated the concentration of each protein and lactose based on their weight fraction in milk and considering the fact that cow milk has 30–39 g/L of protein and 45–55 g/L of lactose in total. The molar mass of each protein was taken from AlphaFold database [44]. Following this, all proteins underwent a 50 ns equilibration in water using NVT and NPT ensembles.

The UA computations were conducted using nine different Al NPs with varying radii, namely 2, 5, 10, 20, 30, 40, 50, 80, and

Table 1: Characteristics of the selected milk proteins and lactose.

Abbreviation	UniProt ID	Compound name	MW ^a , Da	Charge, e	Res ^b	C ^c [10 ⁻⁴], mol/L	R _g ^d [Å]
AS1C	P02662	α1-casein	24528.00	-8.5	214	4	20.05
AS2C	P02663	α2-casein	26018.69	4.5	222	1	40.81
BC	P02666	β-casein	25107.33	-4.5	224	4	22.53
ALAC	P00711	α-lactalbumin	16246.61	-5	142	0.9	15.01
BLAC	P02754	β-lactoglobulin	19883.25	-6	178	2	15.50
BSA	P02769	bovine serum albumin	69293.41	-4.5	607	0.1	27.69
LAC	—	lactose	342.3	0	—	1300	4.28

^aMolecular weight, ^bNumber of residues, ^cConcentrations [mol/L] of the molecules in milk that were used in KMC calculations, ^dRadius of gyration of the biomolecules in Ångstrom.

100 nm, to investigate the influence of size and curvature on the adsorption energies. The results and detailed information on the calculation can be found in Supporting Information File 1, Figure S2 and Figure S3, which illustrate the variations in adsorption energies as a function of NP size. Within the range of 2–20 nm the binding energies of ALAC, BLAC, BC, and BSA show an initial increase on all surfaces, followed by a stabilization at larger NP sizes. In contrast, AS1C and AS2C exhibit a continuous rise in binding energy across the entire size spectrum, ranging from $-48.0k_B T$ at 2 nm to $-281.09k_B T$ at 100 nm for AS1C and $-15.26k_B T$ at 2 nm to $-275.60k_B T$ at 100 nm for AS2C, with AS2C exhibiting the most dramatic changes in binding energy as a function of size. This strong size dependence in binding energy for AS2C can be attributed to its rod-like 3D structure and the rigidity assumption in our model. As the size of the NP increases, AS2C can make more extensive contact with the surface. This increased contact area leads to enhanced binding affinity, resulting in the observed stronger binding across the size range. This is not the case for other proteins on the list as they are more compact and, therefore, reach the maximum number of contacts at relatively small NP sizes. Regarding the binding affinity rankings, for the smallest NPs (2 nm), the order from weakest to strongest is observed as AS2C, BSA, ALAC, BLAC, AS1C, and BC on Al(100), with similar rankings observed on Al(110) and Al(111) surfaces. However, for the largest (flattest) NPs (100 nm), the binding affinity ranking changes to ALAC, BLAC, BSA, BC, AS2C, and AS1C on Al(100), BC, ALAC, BLAC, BSA, AS2C, and AS1C on Al(110), and BLAC, ALAC, BC, BSA, AS2C, and AS1C on Al(111) (see Supporting Information File 1, Figure S2). In reality, protein structures are not rigid, allowing them to adapt to the surfaces upon immobilisation. This can potentially affect their binding behavior. This can be especially significant for caseins, as they belong to the group of flexible milk proteins with no tertiary structure. Globular milk proteins (lactoglobulin and lactalbumin) are expected to be less prone to this shortcoming of the UA model.

Figure 7 shows the output of the UA model for the selected milk proteins on aluminum NPs with a surface size of 80 nm with zeta potential -5 mV at pH 7.0. The heatmaps display the adsorption energies for all values of θ and ϕ . Blue areas with lower energies indicate more favorable orientations of the proteins. Each heatmap is accompanied by a 3D representation of the protein on the NP surface, with the AAs closest to the NP's surface marked. The AAs that are most likely to make contact with the metal surfaces, according to analysis, are LYS, TYR, PHE, GLU, ARG, and ASP.

The rankings of protein adsorption on each aluminum surface are shown in Table 2, highlighting the variations in adsorption energies ($E_{\text{ads}}/k_B T$) and the particular protein–surface interactions (θ and ϕ in degrees). Moreover, the minimum distance (r_{min} in nm) indicates the closest approach of the protein to the aluminum surface during the adsorption process.

The ranking of adsorption energies highlights the distinct adsorption behaviors of various proteins on different metal fcc surfaces. We can see that AS1C exhibits the highest adsorption energy on Al(100) and Al(111) surfaces, while on Al(110), AS1C, and AS2C show similar adsorption energies. In contrast, on metallic iron, AS1C consistently demonstrates the highest adsorption energy on Fe(100), Fe(110), and Fe(111) surfaces. This result reflects the size and shape of the AS1C protein, which allows it to make the largest number of contacts with the metal as compared to the other proteins. Regarding the most weakly bound proteins, on aluminum surfaces, ALAC consistently exhibits the lowest adsorption energy across all three surfaces, while BLAC shows slightly higher adsorption energies. In contrast, on iron surfaces, ALAC and BLAC demonstrate comparable adsorption energies, with ALAC exhibiting slightly lower energies on Fe(110) and Fe(111) surfaces [2]. We note that generally the binding of proteins to aluminum is weaker than to iron, which may be caused by the smaller lattice constant of fcc iron and higher density of surface atoms.

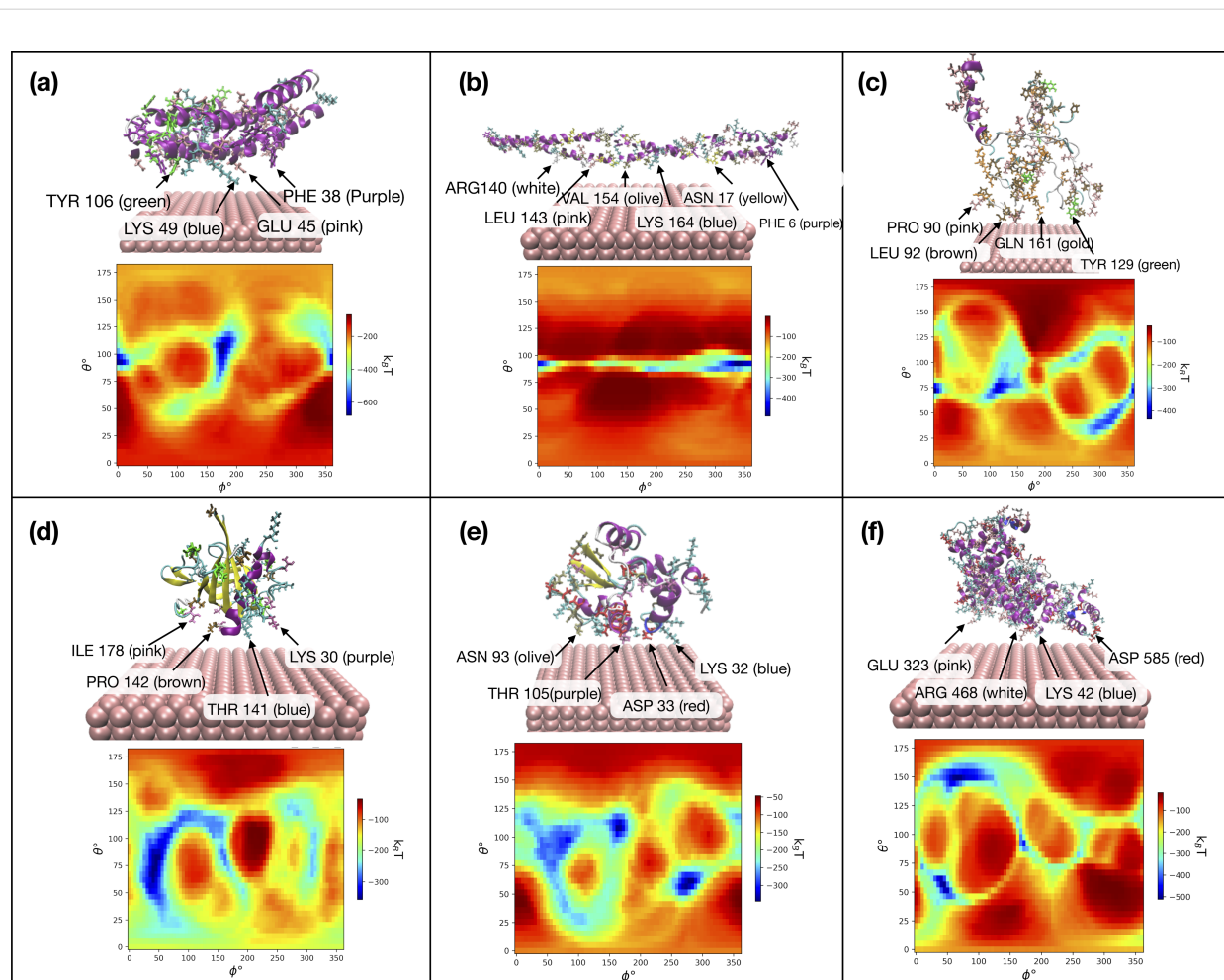


Figure 7: Adsorption energy heatmaps obtained from the UnitedAtom model and corresponding 3D representations of the interactions of (a) AS1C, (b) AS2C, (c) BC, (d) BLAC, (e) ALAC, and (f) BSA with Al(110) in the preferred orientations. The figure highlights the closest AAs to the surface of the material.

Supporting Information File 2, Table S2 reports the preferred orientations of all 820 milk proteins based on the lowest energy from the UnitedAtom output. In our investigation of these proteins, we focused on identifying the most strongly adsorbing proteins when exposed to Fe and Al. These proteins, including P19660, A6QP30, G3X745, F1MMI6, E1BBY7, A6QLY7, and Q9N2I2, demonstrated remarkable similarity in their binding behavior towards Fe(100) and Al(100) surfaces, E1BGJ4, A5D7M6, F1MMI6, A6QP30, G3X745, and F1N1C7 on Fe(110) and Al(110) surfaces, and F1MMI6 and E1B748 and A6QP30 on Fe(111) and Al(111) surfaces.

In the subsequent step, we predicted the composition of the milk protein layer at the aluminum surfaces. For this analysis, we consider the Al surface as a spherical NP with the protein layer uniformly adsorbed on its entire surface, forming the protein corona.

Competitive adsorption and biomolecular corona

Kinetic Monte Carlo (KMC) simulations as implemented in the CoronaKMC tool [26] were employed to investigate competitive adsorption and to determine the composition of the protein corona. This method models adsorbates as hard spheres, which adsorb and desorb to the surface of the NPs, with different orientations of each protein treated as different potential adsorbates to allow for a more physically realistic model of corona formation for anisotropic proteins. In brief, a standard kinetic Monte Carlo routine is used to advance the simulation from one event, collision of an incoming adsorbate with the NP or desorption of an adsorbed species, to the next, with events occurring with a probability proportional to their rate. In the initial form of the model, adsorption is assumed to occur with unit probability if the incoming species does not overlap with any currently adsorbed species and fails to take place otherwise. We

Table 2: Comparison of milk proteins' binding affinities and orientations on Al(100), Al(110), and Al(111) with NP radius of 80 nm, derived from the UnitedAtom model and ordered by the binding strength on each surface.

Individual protein adsorption description on Al(100)				
Protein,	$E_{\text{ads}}/k_{\text{B}}T$	$\phi, ^\circ$	$\theta, ^\circ$	$r_{\text{min}}, \text{nm}$
AS1C	−145.65	175	100	0.19
BC	−108.13	305	40	0.13
AS2C	−96.12	315	95	0.05
BSA	−91.11	45	60	0.11
BLAC	−67.35	65	90	0.19
ALAC	−49.12	125	35	0.20
Individual protein adsorption description on Al(110)				
Protein,	$E_{\text{ads}}/k_{\text{B}}T$	$\phi, ^\circ$	$\theta, ^\circ$	$r_{\text{min}}, \text{nm}$
AS1C	−278.37	175	100	0.32
AS2C	−224.01	345	90	0.10
BSA	−173.77	40	60	0.23
BLAC	−157.70	50	95	0.28
ALAC	−155.17	70	90	0.29
BC	−132.52	0	70	0.20
Individual protein adsorption description on Al(111)				
Protein,	$E_{\text{ads}}/k_{\text{B}}T$	$\phi, ^\circ$	$\theta, ^\circ$	$r_{\text{min}}, \text{nm}$
AS1C	−242.93	175	100	0.15
AS2C	−181.65	330	90	0.11
BSA	−137.46	45	60	0.13
BC	−131.93	140	110	0.15
ALAC	−125.76	75	90	0.17
BLAC	−113.39	45	75	0.20

parameterize this model using adsorption and desorption rate constants extracted from UnitedAtom results as described previously [16,45]. In brief, each potential adsorbate (e.g., a small molecule or a particular orientation of a protein) is projected onto the surface of the NP and a convex hull procedure used to estimate the area of the NP occupied by that adsorbate, A_i . The adsorbate is then assigned an effective radius R_i such that a sphere projected onto the NP would produce the same radius [16]. The per-site adsorption rates are calculated using kinetic theory for the rate of collisions between two spheres in solution, normalized by the number of binding sites for that protein,

$$k_a = \frac{A_i}{4\pi R_{\text{NP}}^2} [4\pi D N_A (R_{\text{NP}} + R_i)], \quad (2)$$

where R_{NP} is the radius of the NP, N_A is Avogadro's number, R_A is the effective adsorbate radius, D is the pair diffusion coefficient given by

$$D = \frac{k_{\text{B}}T}{6\eta} (R_{\text{NP}}^{-1} + R_A^{-1}), \quad (3)$$

taking the viscosity $\eta = 8.9 \times 10^{-4}$ Pa·s. We employ SI units in the above calculation, noting that k_a must then be multiplied by 1000 to convert from units $\text{m}^3 \cdot \text{mol}^{-1}$ to $\text{L} \cdot \text{mol}^{-1}$. Desorption rates are found by requiring that $k_d = k_a \times 1 \frac{\text{mol}}{\text{L}} \times e^{E_{\text{ads}}/k_{\text{B}}T}$, where E_{ads} is the value obtained for that orientation using UnitedAtom [45]. A concentration is then assigned to the adsorbate based on the bulk concentration of that adsorbate, weighted by the relative abundance of that orientation of the adsorbate if necessary. This means that for protein i with a bulk concentration of C_i and a set of orientations θ_k , an orientation θ_j is assigned a concentration

$$C_{i,j} = C_i \frac{\sin \theta_j}{\sum_k \sin \theta_k} \quad (4)$$

to ensure that orientations are correctly weighted and the total concentration summed over orientations is correctly reproduced. Scripts to automate this parameterization based on UA output and adsorbate structure files are available as part of the UnitedAtom repository [26].

We further analyze the results for adsorption of milk components obtained from KMC simulations, specifically focusing on the mean absolute and relative abundance of proteins (10^{-3} nm^2) adsorbed on Al surfaces per unit area (nm^2). Table 3 shows the abundances of proteins and lactose on Al surfaces.

The simulations were performed using NPs with a radius of 80 nm, and the results are collected in Table 3. It presents the number concentration and mass abundance of proteins adsorbed on three different Al surfaces, namely Al(100), Al(110), and Al(111). Each protein's adsorption behavior is quantified in terms of its number concentration (expressed in units of 10^{-3} nm^{-2}) and mass abundance (represented as a percentage of the total adsorbed mass). These calculations were performed utilizing the most recent KMC method modifications, including an alternative mode in which the acceptance–rejection criteria for incoming adsorbates are altered to allow replacement of pre-existing adsorbates. We should note that Al(111) has the lowest energy of all three surfaces, according to the Materials Project data, so we expect the adsorption profile in real systems to be similar to that predicted for Al(111).

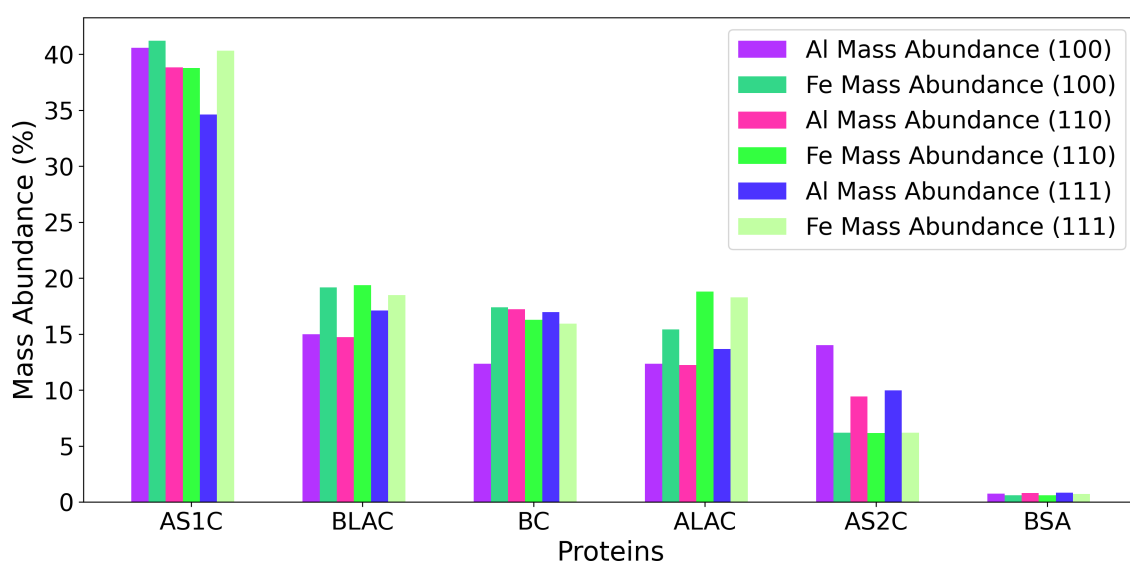
Table 3: Mean amounts of proteins adsorbed on Al surfaces per unit area: number concentration (per nm²) and mass abundance obtained from KMC simulations with NPs of radius 80 nm. These calculations have been done using the KMC method with displacements.

Protein	Al(100)		Al(110)		Al(111)	
	$N_{\text{ads}} [10^{-3}, \text{nm}^{-2}]$	$M_{\text{ab}}, \%$	$N_{\text{ads}} [10^{-3}, \text{nm}^{-2}]$	$M_{\text{ab}}, \%$	$N_{\text{ads}} [10^{-3}, \text{nm}^{-2}]$	$M_{\text{ab}}, \%$
AS1C	12.26	57.16	16.70	67.82	27.21	83.19
BC	4.45	21.24	3.38	14.07	1.91	5.84
BLAC	2.91	10.99	2.97	9.79	1.00	2.43
LAC	96.59	6.28	89.13	5.05	84.50	3.62
ALAC	1.14	3.51	1.13	3.05	1.84	3.60
AS2C	0.11	0.55	0.04	0.16	3.00	1.09
BSA	0.02	0.25	0.00	0.05	0.02	0.21

We also compared the protein composition in the corona on aluminum and iron [2], obtained in our previous work using the original KMC approach without molecular displacements. This comparison is shown in Figure 8. AS1C exhibited the highest abundance on both iron and aluminum among the studied proteins, indicating a strong affinity for both metals with both KMC methods as well as its high number concentration in solution. The following AS1C, BC, BLAC, and ALAC also showed fairly equal abundances on the surfaces of iron and aluminum. In contrast, BSA displayed the lowest abundance on both metals because of its larger size and the relatively low molar fraction in milk as compared with other proteins. Figure 8 shows the mass abundance of each protein on both aluminum (Al(100), Al(110), and Al(111)) and iron (Fe(100), Fe(110), and Fe(111)) surfaces.

We can also observe that AS1C, BLAC, and ALAC display significantly enhanced presence on Fe surfaces in contrast to Al. Conversely, AS2C shows greater adsorption on Al surfaces as compared to Fe. Overall, we expect a somewhat different corona formed on these metallic surfaces.

Real-life organic media do not consist only of proteins, but they also include many other molecules, for example, sugars and other organic compounds that may bind to NPs along with proteins. It can reasonably be assumed that these molecules may alter both the kinetics and equilibrium state of the corona and, moreover, may play a role in biological outcomes. Thus, it is of interest to include these small molecules in the corona simulation to not only gain further insight into this particular case of

**Figure 8:** Mass abundance of proteins on Al and Fe surfaces (100, 110, and 111) using the original KMC approach without molecular displacements and a NP radius of 80 nm.

aluminum in milk, but also to establish a methodology by which more general molecules can be included in these simulations. We choose lactose as a prototypical example of a small molecule capable of binding to NPs, since it is present at a high concentration in milk. We model the lactose molecule as a pair of glucose beads separated by a distance determined by the equilibrium structure of lactose. Although this is not completely rigorous, it demonstrates how the UnitedAtom software can be adapted to model larger molecules other than proteins using the same fragment-based approach. To avoid the need to run a time-consuming parameterization protocol based on metadynamics simulations, we produce PMFs for the glucose bead using a machine-learning technique (PMFPredictor) trained on previous metadynamics results [38]. For the lactose molecule, each constituent glucose bead is assigned a charge of 0, and the Hamaker term is neglected because of the small size of these beads. Following this parameterization, the coarse-grained lactose molecule is processed identically to proteins using the same automated pipeline, that is, UnitedAtom is run to produce a table of orientation-specific binding energies. These are mapped to rate constants for adsorption and desorption. We stress that this procedure is sufficiently generic that essentially arbitrary organic molecules can be included in the simulation by performing a fragment-based decomposition, generating PMFs via traditional or machine-learning approaches, and constructing a coarse-grained representation for input to UA. To simplify this procedure for more complex molecules, we have developed a Python script (MolToFragments.py) employing RDKit [46] to automate splitting larger molecules into suitable fragments and producing coarse-grained input files suitable for UnitedAtom and included it in this repository [26].

The addition of lactose (or other small molecules) to the corona simulation poses a challenge for the form of the CoronaKMC algorithm previously employed because of the high concentration and very small binding area of this small molecule relative to proteins [16,45]. As a consequence of these factors, the original form of the algorithm results in rapid coverage of the NPs with a very large quantity of lactose. This greatly increases the required computational time, which scales as $\mathcal{O}(N^2)$ for N adsorbed particles. Moreover, in this original form of the model, a single adsorbed lactose molecule inhibits the adsorption of a large protein, no matter how strongly the protein may adsorb. To counteract these issues, the following features were added to the new version of the CoronaKMC software. First, we implemented a method to accelerate the simulation by adjusting rate constants for quasi-equilibrated processes (e.g., the adsorption of lactose) according to the methodology of Dybeck and co-workers [47]. Second, we added an optional mode in which the acceptance–rejection criteria for an incoming adsorbate are modified such that an incoming adsorbate is no longer immedi-

ately rejected if it overlaps with a pre-existing adsorbate. Instead, the incoming adsorbate is accepted with a probability p given by,

$$p(\Delta E) = \frac{\exp[-\Delta E/k_{\text{B}}T]}{1 + \exp[-\Delta E/k_{\text{B}}T]}, \quad (5)$$

where ΔE is the difference in energy between the two states,

$$\Delta E = E_{\text{ads}} - \sum_j E_j, \quad (6)$$

where j is the set of all adsorbed particles that would overlap with this particle, taking $\Delta E = E_{\text{ads}}$ if no overlaps are found. If the adsorbate is accepted, then all the overlapping particles are removed from the NP. We note that this breaks the principle of detailed balance in that it allows for the replacement of a set of adsorbates by a single molecule, but does not allow for the converse in which a set of incoming molecules can displace an adsorbate. We justify this neglect on the basis that the required event of multiple simultaneous collisions on a single target would occur so rarely that it would essentially not be sampled in the course of a simulation. The probabilistic acceptance to regions of the NP without explicit adsorbates present effectively multiplies the adsorption rate by a factor of $p(E_{\text{ads}})$. Thus, to maintain the same equilibrium constant, we must multiply the desorption rate by this same factor, noting that this correction is only significant for very weakly adsorbing particles with $E_{\text{ads}} \gtrsim -3k_{\text{B}}T$. This methodology does not treat adsorption of water to the NP explicitly. Instead, it is assumed that all binding energies are defined relative to the adsorption of water, which is assigned an affinity $E_{\text{ads}} = 0k_{\text{B}}T$, and that the concentration of water is sufficiently high such that any region of the NP without an explicit adsorbate can be assumed to be covered in water.

The results of simulations obtained with the updated CoronaKMC (i.e., including the molecule displacement) are shown in Table 3, and they suggest a notable variation in the abundances of proteins and lactose among different Al crystallographic orientations. Notably, on all surfaces studied, AS1C and BC consistently exhibited the highest protein abundances, while BLAC, LAC, and ALAC demonstrated moderate adsorption levels. In contrast, AS2C and BSA consistently displayed the lowest adsorption among the proteins considered in our simulations. Furthermore, when considering different Al facets, it is evident that the (110) surface consistently exhibited the weakest average adsorption across all proteins. When the displacement is allowed, AS1C gains much more space in the corona by replacing other proteins, mostly BLAC, ALAC, and AS2C.

Figure 9 presents a comparison between the protein abundances in the corona on Al and Fe obtained using the enhanced version of the KMC algorithm with molecular displacements. As discussed earlier, this improved algorithm addresses computational efficiency concerns and more accurately represents long-term scenarios during protein corona formation. As shown in the Figure, these algorithmic improvements have a profound impact on the mass concentration of milk proteins on metallic surfaces, particularly on iron. In the original algorithm (Figure 8), proteins showed comparable mass abundances on both metals. However, the enhanced algorithm reveals a distinct change in the adsorption behavior of the AS1C protein on Fe and Al surfaces, characterized by a substantial increase in mass concentration compared to other proteins. The data in Table 3 show that in terms of mass abundance lactose ranks fourth among the corona components (see Supporting Information File 1, Figure S3). As compared to the algorithm without displacement [2], the protein abundance ranking on iron (NP radius 80 nm) surfaces changes to AS1C \gg BC \geq BLAC \geq ALAC $>$ AS2C \approx BSA. A comparable affinity ranking is also now observed for aluminum surfaces (80 nm) studied in current work: AS1C \gg BC \geq BLAC \geq ALAC $>$ AS2C \approx BSA.

Conclusion

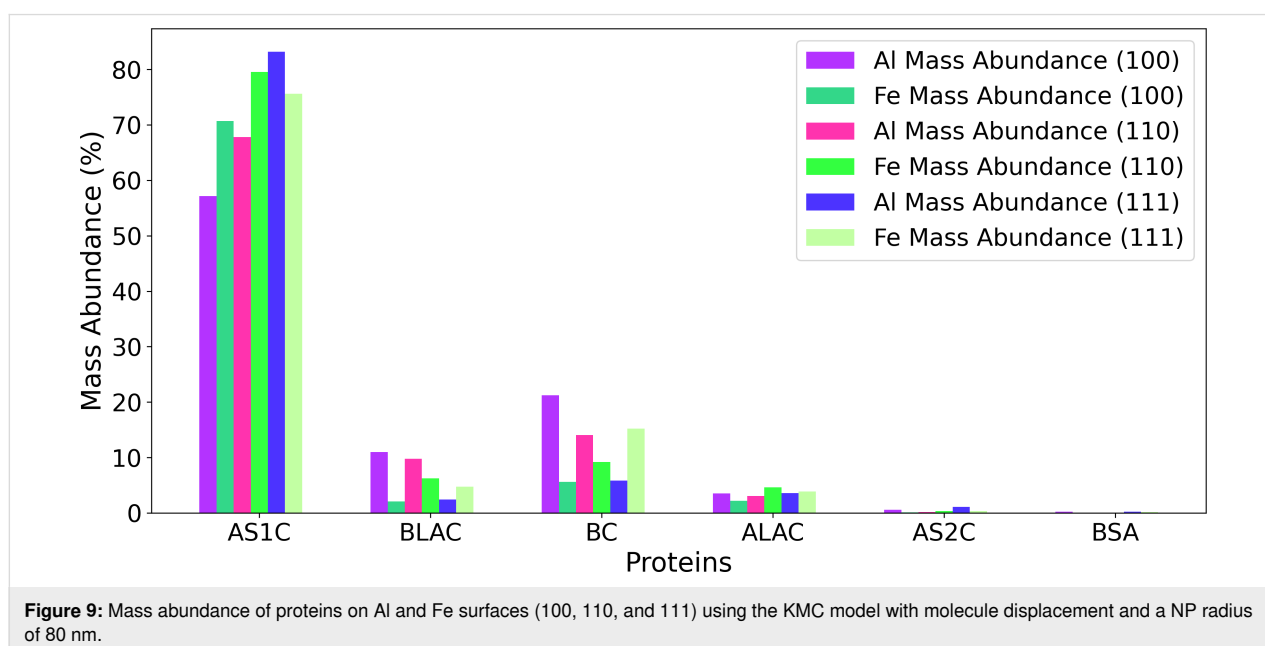
In this work, we applied a multiscale computational model to study the adsorption of milk solids on the metallic surfaces of aluminum, widely used in food processing/packaging. The milk model contained the six most common milk proteins and lactose. To account for the size differences of selected milk constituents, we used an improved competitive adsorption algo-

rithm that can potentially achieve a realistic description of biocorona formation processes with diverse adsorbates (e.g., for predicting an eco-corona).

Our computational model predicts strong binding of milk proteins to pure aluminum surfaces, which is in agreement with our previous observations for metallic iron surfaces [2]. For aluminum, we also found that AS1C and AS2C exhibited the strongest binding to the metal, followed by BSA, BC, BLAC, and ALAC, which displayed weaker adsorption. We also found similar protein abundances in the corona for the two metals demonstrated by KMC simulation results. AS1C dominates the adsorption as the most abundant protein on aluminum surfaces, with BSA being the least abundant. We found a small difference in the predicted corona content between the two metals: BC and BLAC prefer Al(100) and Al(110) to iron, while AS1C prefers Fe(100) and Fe(110) over aluminum.

Although the adsorption energy regulates the interaction strength between proteins and surfaces, the mass concentration of proteins in the solution has a major effect on the amount of protein adsorbed onto the surface. Expanding the milk model by adding lactose into the mix did not alter the ranking of protein abundance in the corona. Despite the high concentration in the milk, lactose does not exceed the mass abundance of specific proteins such as AS1C due to its small size. In our model, it essentially forms a thin monolayer on the surface.

Overall, our freely accessible multiscale computational model [26] allows us to make predictions of the binding strength, preferred orientations, and relative abundance of the specified



molecules on the specified material surfaces or NPs and, thus, gives an insight into the mechanisms of bionano interaction. We can compare different materials in terms of the protein binding affinity and corona content and optimize the processes in food and chemical industry. The presented methodology can be easily extended to other molecules, materials, and contexts involving the bionano interface such as environmental safety, health, medical devices, or toxicology.

Supporting Information

Table S1: Adsorption free energies for each SCA on Al surfaces; Figure S1: Water density profiles for aluminum slabs: (a) Al(100), (b) Al(110), (c) Al(111), Figure S2: Influence of the NP size on the adsorption energies; Figure S3: Milk molecules ranking based on mass abundance in the corona, Figure S4: Example of AlNP size-dependent interaction of ALAC: (a) 2 nm, (b) 5 nm, (c) 10 nm, (d) 20 nm, (e) 40 nm, (f) 50 nm, (g) 80 nm, (d) 100 nm; Figures S5–S10: Comparison of interaction of AS1C, AS2C, β -casein, ALAC, BLAC, BSA, with different Al fcc surfaces: (a) Al(100), (b) Al(110), (c) Al(111); Table S2: Description of 820 milk proteins interaction with Al (100, 110, 111) based on the lowest energy values of the adsorption heatmaps.

Supporting Information File 1

Supporting material.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-15-21-S1.pdf>]

Supporting Information File 2

820-Milk-protein-table.

[<https://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-15-21-S2.pdf>]

Acknowledgements

The authors extend their gratitude to the Irish Centre for High-End Computing (ICHEC) and the UCD Sonic High-Performance Computing Centre for providing the necessary computational resources for this research. The graphical abstract of this article was created with <https://biorender.com/> and its content is not subject to CC BY 4.0.

Funding

The research presented in this study was supported by funding from Science Foundation Ireland under grant number 16/IA/4506, the European Union Horizon 2020 Programme grant 814572 (NanoSolveIT), and European Union Europe Programme grant 101092741 (nanoPASS).

ORCID® iDs

Parinaz Mosaddeghi Amini - <https://orcid.org/0009-0003-6003-6830>

Ian Rouse - <https://orcid.org/0000-0002-3686-7701>

Julia Subbotina - <https://orcid.org/0000-0002-2227-0787>

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

Preprint

A non-peer-reviewed version of this article has been previously published as a preprint: <https://doi.org/10.3762/bxiv.2023.45.v1>

References

- Pulido-Reyes, G.; Leganes, F.; Fernández-Piñas, F.; Rosal, R. *Environ. Toxicol. Chem.* **2017**, *36*, 3181–3193. doi:10.1002/etc.3924
- Mosaddeghi Amini, P.; Subbotina, J.; Lobaskin, V. *Nanomaterials* **2023**, *13*, 1857. doi:10.3390/nano13121857
- Karaman, A. D.; Özer, B.; Pascall, M. A.; Alvarez, V. *Food Rev. Int.* **2015**, *31*, 295–318. doi:10.1080/87559129.2015.1015138
- Lamberti, M.; Escher, F. *Food Rev. Int.* **2007**, *23*, 407–433. doi:10.1080/87559120701593830
- Das, S.; Panda, S. H.; Bole, M.; Pal, N.; Samantaray, B. R.; Thatoi, H. Facets of nanotechnology in food processing, packaging, and safety: an emerald insight. In *Bio-Nano Interface: Applications in Food, Healthcare and Sustainability*; Arakha, M.; Pradhan, A. K.; Jha, S., Eds.; Springer: Singapore, 2022; pp 75–92. doi:10.1007/978-981-16-2516-9_5
- Nabrzyski, M.; Gajewska, R.; Czuprynska-Rzepko, A.; Sandak-Bosak, K. *Rocz. Panstw. Zakl. Hig.* **1994**, *45*, 1181–1190. <https://yadda.icm.edu.pl/yadda/element/bwmeta1.element.agro-article-e854232b-cdd5-4e6d-b974-4bf212e50601>
- Barabasz, W.; Albińska, D.; Jaśkowska, M.; Lipiec, J. *Pol. J. Environ. Stud.* **2002**, *11*, 199–204.
- Tang, L.; Thevenot, P.; Hu, W. *Curr. Top. Med. Chem.* **2008**, *8*, 270–280. doi:10.2174/156802608783790901
- Williams, D. F. *Biomaterials* **2008**, *29*, 2941–2953. doi:10.1016/j.biomaterials.2008.04.023
- Dobrovolskaia, M. A.; Germolec, D. R.; Weaver, J. L. *Nat. Nanotechnol.* **2009**, *4*, 411–414. doi:10.1038/nnano.2009.175
- Landsiedel, R.; Ma-Hock, L.; Kroll, A.; Hahn, D.; Schnekenburger, J.; Wiench, K.; Wohlleben, W. *Adv. Mater. (Weinheim, Ger.)* **2010**, *22*, 2601–2627. doi:10.1002/adma.200902658
- Winkler, D. A.; Burden, F. R.; Yan, B.; Weissleder, R.; Tassa, C.; Shaw, S.; Epa, V. C. *SAR QSAR Environ. Res.* **2014**, *25*, 161–172. doi:10.1080/1062936x.2013.874367
- Darabi Sahneh, F.; Scoglio, C.; Riviere, J. *PLoS One* **2013**, *8*, e64690. doi:10.1371/journal.pone.0064690
- Shao, Q.; Hall, C. K. *J. Phys.: Condens. Matter* **2016**, *28*, 414019. doi:10.1088/0953-8984/28/41/414019
- Angioletti-Uberti, S.; Ballauff, M.; Dzubiella, J. *Mol. Phys.* **2018**, *116*, 3154–3163. doi:10.1080/00268976.2018.1467056
- Rouse, I.; Lobaskin, V. *Biophys. J.* **2021**, *120*, 4457–4471. doi:10.1016/j.bpj.2021.09.002
- Rouse, I.; Power, D.; Brandt, E. G.; Schneemilch, M.; Kotsis, K.; Quirke, N.; Lyubartsev, A. P.; Lobaskin, V. *Phys. Chem. Chem. Phys.* **2021**, *23*, 13473–13482. doi:10.1039/d1cp01116b

18. Wyrzykowska, E.; Mikolajczyk, A.; Lynch, I.; Jeliakovska, N.; Kochev, N.; Sarimveis, H.; Doganis, P.; Karatzas, P.; Afantitis, A.; Melagraki, G.; Serra, A.; Greco, D.; Subbotina, J.; Lobaskin, V.; Bañares, M. A.; Valsami-Jones, E.; Jagiello, K.; Puzyn, T. *Nat. Nanotechnol.* **2022**, *17*, 924–932. doi:10.1038/s41565-022-01173-6
19. Monopoli, M. P.; Walczyk, D.; Campbell, A.; Elia, G.; Lynch, I.; Baldelli Bombelli, F.; Dawson, K. A. *J. Am. Chem. Soc.* **2011**, *133*, 2525–2534. doi:10.1021/ja107583h
20. Vilanova, O.; Mittag, J. J.; Kelly, P. M.; Milani, S.; Dawson, K. A.; Rädler, J. O.; Franzese, G. *ACS Nano* **2016**, *10*, 10842–10850. doi:10.1021/acsnano.6b04858
21. Walkey, C. D.; Olsen, J. B.; Song, F.; Liu, R.; Guo, H.; Olsen, D. W. H.; Cohen, Y.; Emili, A.; Chan, W. C. W. *ACS Nano* **2014**, *8*, 2439–2455. doi:10.1021/nn406018q
22. Nel, A. E.; Mädler, L.; Velegol, D.; Xia, T.; Hoek, E. M. V.; Somasundaran, P.; Klaessig, F.; Castranova, V.; Thompson, M. *Nat. Mater.* **2009**, *8*, 543–557. doi:10.1038/nmat2442
23. Eskin, N. A. M.; Shahidi, F. *Biochemistry of Foods*, 3rd ed.; Academic Press, 2013. doi:10.1016/b978-0-08-091809-9.00018-2
24. Brandt, E. G.; Lyubartsev, A. P. *J. Phys. Chem. C* **2015**, *119*, 18126–18139. doi:10.1021/acs.jpcc.5b02670
25. Power, D.; Rouse, I.; Poggio, S.; Brandt, E.; Lopez, H.; Lyubartsev, A.; Lobaskin, V. *Modell. Simul. Mater. Sci. Eng.* **2019**, *27*, 084003. doi:10.1088/1361-651x/ab3b6e
26. Physical-chemical aspects of protein corona: relevance to in vitro and in vivo biological impacts of nanoparticles. <https://bitbucket.org/softmattergroup/unitedatom/>.
27. Subbotina, J.; Rouse, I.; Lobaskin, V. *Nanoscale* **2023**, *15*, 13371–13383. doi:10.1039/d3nr03264g
28. Subbotina, J.; Lobaskin, V. *J. Phys. Chem. B* **2022**, *126*, 1301–1314. doi:10.1021/acs.jpcc.1c09525
29. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435–447. doi:10.1021/ct700301q
30. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.; Lindahl, E. *Bioinformatics* **2013**, *29*, 845–854. doi:10.1093/bioinformatics/btt055
31. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Comput. Phys. Commun.* **2014**, *185*, 604–613. doi:10.1016/j.cpc.2013.09.018
32. Jo, S.; Kim, T.; Iyer, V. G.; Im, W. J. *Comput. Chem.* **2008**, *29*, 1859–1865. doi:10.1002/jcc.20945
33. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Mod.* **2006**, *25*, 247–260. doi:10.1016/j.jmglm.2005.12.005
34. Sousa da Silva, A. W.; Vranken, W. F. *BMC Res. Notes* **2012**, *5*, 367. doi:10.1186/1756-0500-5-367
35. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174. doi:10.1002/jcc.20035
36. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690. doi:10.1063/1.448118
37. Potential of Mean Force (PMFs) for Zerovalent Aluminium (100-110-111) NanoParticles.. <https://doi.org/10.5281/zenodo.8334053> (accessed Sept 11, 2023). doi:10.5281/zenodo.8334053
38. Rouse, I.; Lobaskin, V. *Faraday Discuss.* **2023**, *244*, 306–335. doi:10.1039/d2fd00155a
39. Lopez, H.; Lobaskin, V. *J. Chem. Phys.* **2015**, *143*, 243138. doi:10.1063/1.4936908
40. Rouse, I.; Power, D.; Brandt, E. G.; Schneemilch, M.; Kotsis, K.; Quirke, N.; Lyubartsev, A. P.; Lobaskin, V. *Phys. Chem. Chem. Phys.* **2021**, *23*, 13473–13482. doi:10.1039/d1cp01116b
41. Roy, A.; Kucukural, A.; Zhang, Y. *Nat. Protoc.* **2010**, *5*, 725–738. doi:10.1038/nprot.2010.5
42. Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 525–537. doi:10.1021/ct100578z
43. Søndergaard, C. R.; Olsson, M. H. M.; Rostkowski, M.; Jensen, J. H. *J. Chem. Theory Comput.* **2011**, *7*, 2284–2295. doi:10.1021/ct200133y
44. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D. *Nature* **2021**, *596*, 583–589. doi:10.1038/s41586-021-03819-2
45. Hasenkopf, I.; Mills-Goodlet, R.; Johnson, L.; Rouse, I.; Geppert, M.; Duschl, A.; Maier, D.; Lobaskin, V.; Lynch, I.; Himly, M. *Nano Today* **2022**, *46*, 101561. doi:10.1016/j.nantod.2022.101561
46. RDKit. <https://www.rdkit.org/> (accessed Sept 1, 2023).
47. Dybeck, E. C.; Plaisance, C. P.; Neurock, M. *J. Chem. Theory Comput.* **2017**, *13*, 1525–1538. doi:10.1021/acs.jctc.6b00859

License and Terms

This is an open access article licensed under the terms of the Beilstein-Institut Open Access License Agreement (<https://www.beilstein-journals.org/bjnano/terms>), which is identical to the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0>). The reuse of material under this license requires that the author(s), source and license are credited. Third-party material in this article could be subject to other licenses (typically indicated in the credit line), and in this case, users are required to obtain permission from the license holder to reuse the material.

The definitive version of this article is the electronic one which can be found at: <https://doi.org/10.3762/bjnano.15.21>