Molecular modeling of the interface of an egg yolk protein-based emulsion

Revised manuscript

Molecular modeling of the interface of an egg yolk protein-based emulsion

Marco Ferrari,[1] Jan-Willem Handgraaf,[2] Gianluca Boccardo,[1] Antonio Buffo,[1, a)] Marco Vanni,[1] and Daniele L. Marchisio[1]

[1)] *Department of Applied Science and Technology, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129, Torino, Italy*

[2)] *Siemens Industry Software Netherlands B.V., Galileiweg 8, 2333 BD, Leiden, The Netherlands*

(Dated: 4 January 2022)

Many food emulsions are stabilized by functional egg yolk biomolecules, which act as surfactants at the oil/water interface. Detailed experimental studies on egg yolk emulsifying properties have been largely hindered due to the difficulty in isolating individual chemical species. Therefore, this work presents a molecular model of an oil/water interfacial system where the emulsifier is one of the most surface-active proteins from the egg yolk low-density lipoproteins (LDL), the so-called Apovitellenin I. Dissipative Particle Dynamics (DPD) was here adopted in order to simulate large systems over long time-scales, when compared with full-atom molecular dynamics (MD). Instead of a manual assignment of the DPD simulation parameters, a fully-automated coarse-graining procedure was employed. The molecular interactions used in the DPD system were determined by means of a parameter calibration based on matching structural data from atomistic Molecular Dynamics (MD) simulations. Despite of the little availability of experimental data, the model was designed to test the most relevant physical properties of the protein investigated. Protein structural and dynamics properties obtained via MD and DPD were compared highlighting advantages and limits of each molecular technique. Promising results were achieved from DPD simulations of the oil/water interface. The proposed model was able to properly describe the protein surfactant behavior in terms of interfacial tension decrease at increasing protein surface concentration. Moreover, the adsorption time of a free protein molecule was estimated and, finally, an LDL-like particle adsorption mechanism was qualitatively reproduced.

[a)] Author to whom correspondence should be addressed: antonio.buffo@polito.it

I. INTRODUCTION

Food emulsions are made of a continuous water phase, a disperse phase with a high content of oil, and a surfactant that stabilizes the oil drops.[1–5] The droplet size distribution (DSD) is the most important property of the emulsion since the structure, stability, taste, and color of the final product depend on the DSD.[1–5] The DSD in turn depends on the emulsion composition, the type of process and the operating conditions under which the production process operates.[6] The production of emulsions is based on mixing the ingredients and applying a suitable mechanical energy to the emulsion for promoting droplet formation and breakage, in order to reach the desired DSD. A typical mixing process is composed by two steps: first, the ingredients (mainly egg yolk, vinegar, oil, water, salt) are mixed together in large stirred vessels at moderate rotational speed; then, this premixed emulsion is fluxed into a high-shear device, commonly a cone mill mixer, where the oil droplets undergo breakage until the final size distribution is reached.[3–5] This last step is crucial to fine-tune the DSD, in order to determine the properties of the final product.

Many food emulsions are stabilized by surface-active biopolymers that adsorb on the droplet surface and form protective coatings.[1] Some of these functional molecules are integral components of more complex food ingredients used in food products (e.g., egg yolk, milk, and flour).[1,2] Although the egg yolk is recognized as one of the most widely employed emulsifiers for both industrial and home-made food emulsion preparation,[1] many issues need to be addressed, especially the adsorption mechanism of egg yolk proteins at oil-water interface and their emulsifier behaviour.[7] Indeed, the egg yolk is a complex system with different structural levels consisting in non-soluble protein aggregates (granules) in suspension in a clear yellow fluid (plasma) that contains low-density lipoproteins (LDLs) and soluble proteins.[7] Experimental research concerning the emulsifying properties of egg yolk proteins has been hindered by the difficulties in extracting individual components from the complex matrix, therefore, they are less amenable to detailed study by being less readily available in pure form.[8–10]

During the emulsification process, the interfacial properties between disperse and continuous phases play an essential role in the formation and the stabilization of the oil droplets.[1,2] Therefore, it is important to have a fundamental understanding of the factors that influence the type, concentration, interactions, and arrangement of surface-active molecules at

interfaces.[1,2] Computer modeling techniques can greatly enhance the comprehension of the way the molecules organize themselves in a liquid.[11–14] Molecular simulations can provide valuable insight into the relationship between molecular properties and structural organization that are relevant for a better understanding of the behavior of food emulsions, including the miscibility/immiscibility of liquids, the formation of surfactant micelles, the adsorption and displacement of emulsifiers at interfaces, the transport of nonpolar molecules through aqueous phases, the conformation and flexibility of biopolymers in solution, polymer interactions, and the formation of gels.[15–24] The first step in a molecular simulation is to define the characteristics of the molecules involved (e.g., size, shape, flexibility, and polarity) and the nature of the intermolecular pair potentials that act between them, making a number of simplifying assumptions as a compromise between the model reliability and a reasonable computational time.[25] A collection of these molecules is arbitrarily distributed within a box that represents a certain region of space, and the change in the conformation and/or organization of the molecules is then monitored as they are allowed to interact with each other. Depending on the simulation technique used, one can obtain information about the evolution of the structure with time and/or about the equilibrium structure of the molecular ensemble. The most commonly used computer simulation techniques in this context are the Monte Carlo approach and Molecular Dynamics (MD). In these models the involved molecules can be described with all their atomistic details or some of them can be coarse-grained, as in Dissipative Particle Dynamics (DPD).[19,26–30]

Many molecular modeling studies of food structures were carried out employing the aforementioned approaches.[19] The adsorption of flexible proteins (β-casein[31] and a proteinlike heteropolymer[32]) at an oil-water interface was studied by means of Monte Carlo simulations. On the other hand, the majority of MD studies on protein adsorption at fluid interfaces have been on globular proteins using both all-atom and coarse-grained models, with few studies on unstructured intrinsically disordered proteins.[33–40] Few works have been carried out on protein models via coarse-grained DPD technique, although this approach allows the simulation of large systems over relatively long-time scales with respect to full-atomistic studies.[28,29,41] DPD uses simplified soft potentials and coarse-grained representations of modeled structures.[27–29] In contrast to MD, in DPD systems the intended physical properties are determined by means of parameter calibration. One of the most popular method of calibration is based on mapping onto Flory–Huggins theory.[29] Another approach is to couple DPD

with MD simulations to calibrate models by matching the structural data from the atomistic simulations.[42–44] Previous DPD studies investigated the adsorption of semi-flexible rod-like objects,[45] conformation changes[46] or the folding of small proteins.[47] However, all computer molecular techniques have been successfully employed in modeling of interfacial systems and in the calculation of the surface tension when an amphiphilic non-protein molecule act as a surfactant.[48–51] Moreover, DPD is well-suited for modeling of multi-component systems such as emulsions, and it has been used in a number of studies to look at the effect of adsorbing molecules on the stability of oil or water droplets in emulsions.[19,52–54] These have mainly been carried out on hydrocarbon oil emulsions with synthetic copolymers as the adsorbing molecules, but the methodology and the general results are relevant also for food emulsions.

The main goal of the present work is to model an oil/water interfacial system where the emulsifier is one of the most surface-active proteins from the egg yolk LDL, in order to provide new insights into physics of the food emulsion production process. Despite of the little availability of experimental data, the model was designed to test the most relevant physical properties of such a protein by means of the DPD approach in which the parameter calibration is based on MD simulations. Instead of a manual assignment, a fully *automated* coarse-graining procedure was employed to the molecules involved in the ternary system, assuming a flexible, disordered structure for the protein. Promising results were obtained in terms of both equilibrium and dynamic properties of the egg-yolk protein. Finally, the adsorption mechanism of a LDL-like particle is also qualitatively reproduced.

This paper is structured as follows: in Section II the molecular description of the studied system is presented; the molecular techniques here used are briefly introduced in Section III; the model development and calibration are explained in Section IV together with all the simulation details; Section V shows the relevant results of systems investigated and, finally, in Section VI the main conclusions are reported.

II. MOLECULAR DESCRIPTION OF THE MACROSCOPIC SYSTEM

The first step in the development of the molecular model for an egg yolk protein-based emulsion is to identify the chemical species to be simulated and to define the characteristics of the molecules involved at the interface. The basic components of the system under investigation are three: the triglyceride with three monounsaturated oleic acid residues

which stands for the oil phase, the protein Apovitellenin I coming from the egg yolk LDL and, finally, water. In this Section a general description of the macroscopic system to be modeled is provided, together with the adopted simplifications.

An example of a food emulsion where the egg yolk is widely used as an emulsifier is mayonnaise. This is a stable liquid-liquid emulsion with a high content of the dispersed oil phase. In this work a regular mayonnaise with around 70% of fat content[1] is considered and the experimental work of Dubbelboer *et al.*[3] is used as a reference to identify the ingredients of the mayonnaise, especially the molecules to play a primary role at the oil/water interface. It is important to highlight that also in this work the dispersed phase consists of the soybean oil, while the chemical species that act as surfactants are derived from the egg yolk. These two components characterize the specific type of mayonnaise studied, therefore a further description of the vegetable oil and the egg yolk used in the production of the food emulsion is presented in order to correctly select the molecules to be modeled.

Regarding the dispersed phase, a fully refined soybean oil is employed in which the triglyceride molecules are present with a concentration larger than 99%.[55] Triglycerides are tri-esters consisting of a glycerol bound to three fatty acid molecules. Based on the number of double bonds and the chain length, the fatty acids occurring in triglycerides of the soybean oil are saturated, monounsaturated and polyunsaturated with 16 or 18 carbon atoms according to an internal distribution.[55] For the sake of simplicity, here homotriglycerides are taken into account where the three fatty acids are identical (without an internal distribution). In particular, the triglyceride molecules with three monounsaturated oleic acid residues (18 carbon atoms for chain) will be modeled as the representative of the oil phase, instead of hydrocarbons as it was done in previous DPD works on similar emulsions.[52–54] It should be noted that the protein adsorption to different hydrophobic materials may cause differences in the conformation of the adsorbed molecule; in this sense our simplification may have an impact that it is difficult to quantify. That being said, it is known that the modeling of a simpler hydrocarbon–water system instead of a triglyceride–water system might not necessarily lead to realistic results,[56] therefore a triglyceride–water system was modeled in this work.

The second fundamental component in the mayonnaise production is the hen egg yolk. It is mainly composed of two fractions – plasma and granules – which are natural nano- and micro-assemblies. Plasma contains a large quantity of lipids structured as low-density

lipoproteins (LDLs), whereas granules are mainly composed of proteins aggregated in micrometric assemblies.[7] Assuming a pH equal to 3.8 for the mayonnaise,[57] plasma proteins represent about 2/3 of oil–water interface in acidic conditions (at all ionic strengths).[7] Previous works have shown that LDLs are likely to play primary roles in the formation and stabilization of egg yolk-based emulsions.[7,58–61] Consequentially, LDLs are considered to contribute mainly to yolk emulsifying properties.[7] LDLs are spherical nanoparticles (17–60 nm) with a lipid core of triglycerides and cholesterol esters in a liquid state surrounded by a monofilm of phospholipids and apoproteins.[7,62–67] The LDL adsorption mechanism at the oil-water interface was investigated by several works.[7,67–71] In fact, LDLs serve as vectors of surfactant constituents (proteins and phospholipids) that could not be soluble in water until they reach the interface. The adsorption of apoproteins and phospholipids at the interface lead to the formation of a film that stabilize the emulsion.[69] Therefore, both apoproteins and phospholipids are essential to understand the interfacial properties of egg yolk LDLs. The protein identified as Apovitellenin I is considered to be the most surface-active, among the apoproteins contained in LDL.[64,67] Due to its structure and composition, which combines amphipathic character and flexibility, Apovitellenin I shows a great capacity to adsorb at the oil–water interface in emulsions.[67] In LDL, Apovitellenin I is mostly present as a homodimer, thus containing two identical polypeptide chains of 82 amino acid residues which are linked by a single disulfide bond at the cysteine residue.[64,67] The sequence of the mature protein is available in the UniProtKB database[72] under the accession number P02659 (`www.uniprot.org/uniprot/P02659`). However, the detailed 3D structure and other physico-chemical information of Apovitellenin I are not available in the literature to the best of authors' knowledge, increasing the complexity of its modeling approach. The presence of salts, small surfactant molecules (phospholipids) or other additives is here neglected since only the emulsifying capacity of the considered egg yolk LDL protein is investigated. Furthermore, the pH of the system is kept constant and equal to 3.8. The molecular model of the oil-water interface is then described in the following sections.

III. THEORETICAL BACKGROUND

In this Section only the main basic concepts of the standard Dissipative Particle Dynamics (DPD) method are presented, while a further detailed description of both MD and DPD

techniques can be found in the literature[25,27–29,73,74] and in the Supplementary Material.

DPD is a stochastic mesoscale particle model that it has been devised to allow the simulation of the dynamics of mesoscopic particles. Unlike classic Molecular Dynamics, each DPD particle i, called bead, represents a molecular cluster (a molecule fragment or a group of solvent molecules) rather than an individual atom. The major difference between MD and DPD, apart from the coarse-grained nature of the molecules, is the nature of the forces between them. The force acting on each bead i contains three parts: the conservative, dissipative, and stochastic (random) forces, each of which is pairwise additive. Here the conservative force felt by bead i includes: 1) contributions from repulsive interactions with surrounding beads; 2) contributions due to the springs connecting bead i to other beads in the same molecule; and 3) contributions due to angle bending interactions. The repulsive force $\mathbf{F}^r_{ij}$, which is modeled as a soft repulsion between beads i and j, is defined as follows:

$$\mathbf{F}^r_{ij} = \begin{cases} a_{ij}(1 - r_{ij}/r_c)\hat{\mathbf{r}}_{ij} & \text{if } r_{ij} \leq r_c \\ 0 & \text{if } r_{ij} > r_c \end{cases}, \tag{1}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between beads i and j at positions $\mathbf{r}_i$ and $\mathbf{r}_j$ respectively, and $\hat{\mathbf{r}}_{ij} = (\mathbf{r}_i - \mathbf{r}_j)/r_{ij}$ is the direction between the two beads. The parameters a_{ij} are the DPD interaction parameters defined for each bead pair, while r_c stands for the cutoff distance. For the system investigated in this work, their definition will be given in the Section IV B and they will be here used as fitting parameters for the calibration of the DPD model. The adjacent beads are constrained with permanent lengths and angular bonds. In this study, the bonds were modeled using harmonic spring quadratic potentials given as:

$$U^S_{ij} = k_S(r_{ij} - l_H)^2\,, \tag{2}$$

$$U^A_{ijk} = k_A(\theta_{ijk} - \theta_H)^2\,, \tag{3}$$

where l_H and θ_H are the equilibrium lengths and angles for beads i, j and k. The stiffness of the length and angular bond constraints is defined by the values of k_S and k_A.

As it is customary in DPD, the quantities here reported have to be considered reduced (dimensionless) and the scaling factors for the main properties (mass, length, time, energy) will be explained in Section IV C. Finally, it is important to point out that the coarse-graining of the molecular structures and the soft interactions allow larger systems to be modeled over significantly longer times than with (atomistic scale) molecular modeling,[41,74]

thus allowing the dynamics of mesoscopic systems to be followed over relevant time scales as well as length scales.

IV. MODELING DETAILS

In order to consider both the complex composition of the emulsion and the equilibration time required by macro-molecules to re-arrange at interfaces, the DPD approach is employed in which the parameter calibration is based on MD simulations. Next sections will present the setup of MD simulations, the DPD model development in which both the coarse-graining procedure and the calibration of parameters are explained and, finally, definitions of the main physical properties investigated here.

A. MD simulations

The purpose of all-atom MD simulations is to use their results to calibrate the DPD parameter set. Only MD simulations of one protein molecule in bulk phases (water or oil) were performed rather than the entire ternary interfacial system due to the size of the latter which would require excessive computational time. An initial guess of both protein and triglyceride structures was manually made from scratch via a molecule editor. In particular, Figure 2a shows the all-atom protein model. It can be clearly seen the disulfide bond linking two identical polypeptide chains. Furthermore, the N- and C- terminal amino acid residues and, if applicable, the functional group of side chains were protonated or deprotonated by comparing their corresponding pK_a with the pH of the solution.[75] Thus, at pH 3.8 the net charge of the protein homodimer results equal to 16 e and the protein molecular mass M is 18675.6 Da. MD simulations were performed using the OPLS-AA force field,[76,77] while water was described by the TIP3P water model.[78] A cutoff of 7.5 Å was used for long-range interactions, and both electrostatic and van der Waals interactions were handled using a smooth particle mesh Ewald summation method (SPME).[79] For the protein and the triglyceride, first 20-ps simulation in vacuum with a time step of 1 fs was performed on the single molecule to relax its initial structure. Before solvation with water or oil, the protein was centered in a rectangular box with a minimum distance of any part of the molecule defined to be at least 1 nm from box walls in order to satisfy the minimum

image convention when using periodic boundary conditions. According to the reproduced environment, the box was filled with respectively 15994 water or 325 triglyceride molecules, plus 16 Cl^- counterions to ensure the electroneutrality of the system. Thus, the resulting MD box contains a total of 50694 or 56987 atoms in the case of protein in water or oil bulk respectively. After a simple energy minimization to ensure that the system had no steric clashes or inappropriate geometry, a 0.5-ns NPT (i.e., constant number of particles, pressure, and temperature) equilibration simulation at ambient pressure (1 atm) and temperature (298 K) was performed. Pressure and temperature were fixed using the Berendsen barostat and thermostat[80] and the Verlet algorithm was used to integrate the equations of motion with an increased time step of 2 fs. To verify that the system was at the equilibrium, the fluctuations in the temperature, pressure, density, and potential energy were monitored. In particular, the average density reached during the last 0.2 ns of equilibration simulation was equal to 1059.57 and 921.85 kg/m^3 respectively for the protein in water and in oil system, both with fluctuations in the 0.1%. Finally, NVT (i.e., constant number of particles, volume, and temperature) production simulations ranging from 2 to 6 ns were performed to collect statistically averaged results by saving particle trajectories every 250 time steps.

B. Coarse-graining procedure and parameter calibration

The main steps of the DPD model development are summarized in a schematic diagram in Figure 1, in which each stage is explained in this Section.

The first step toward a realistic DPD molecular model is to obtain the coarse-grained (CG) representation of the molecules together with their full parameter set of both inter- and intra-molecular interactions. For this scope, the Automated Fragmentation and Parametrization (AFP) method is used and here a very brief introduction to this approach is provided. For a fully detailed discussion on it, the reader can refer to the work of Fraaije *et al.*[81].

Starting from their fully atomistic representations, the molecules involved in the investigated system are fragmented according to a scoring function, through a simulated annealing function that cuts through bonds; the optimal bond fission pattern is preserved and the fragments are stored. The scoring function is here defined as:

$$S = \left(1 - \frac{V}{V_0}\right)^2, \tag{4}$$

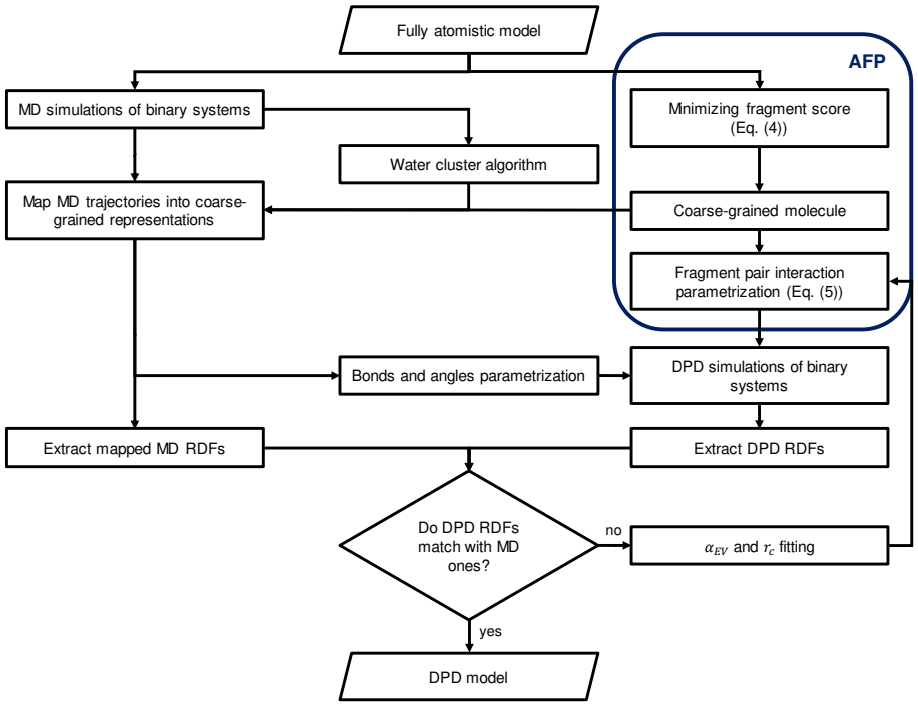


FIG. 1. Schematic diagram of the main stages followed in this work to develop the DPD model. See Section IV B for details of each step.

where V is the volume of the fragment and V_0 is the reference volume of a cluster of three water molecules in its lowest energy conformation (i.e., the reference volume used here is equal to 67.7 $Å^3$ as in the original AFP work[81]). In this approach the molecule-unique fragmentation is used in order to preserve as much as possible of the properties of the molecule. This means that the fragments are not database-unique, as is customary in coarse-grained simulations, but completely specific to a given molecule. By applying this fragmentation technique, the triglyceride molecule and the homodimer Apovitellenin I are comprised of 20 and 500 beads respectively, while each water bead corresponds to three atomistic water molecules. In particular, Figure 2 shows the all-atom (a) and the corresponding coarse-grained (b) representation of the protein molecule.

In the AFP framework, the interaction DPD parameter a_{ij} is split into two contributions,

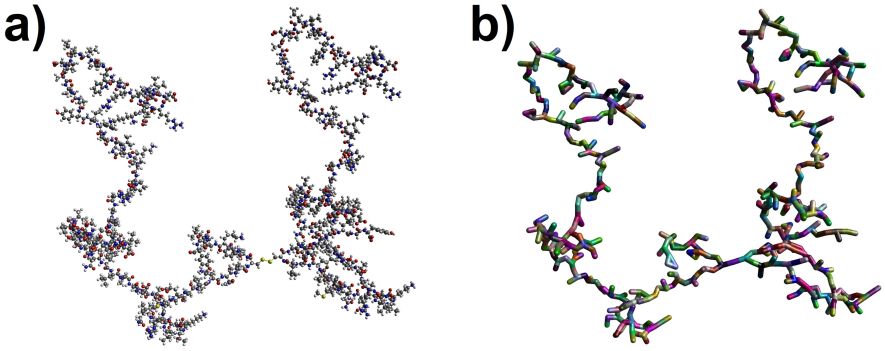


FIG. 2. All-atom (a) and corresponding coarse-grained (b) model obtained via AFP of Apovitellenin I. DPD beads are represented by colored fragments, highlighting the bond fission pattern.

one from the excluded volume and the second from the residual interactions:

$$a_{ij} = \alpha_{EV} v_i v_j + \alpha_{res} \sqrt{v_i v_j} \beta \Delta G_{res,ij} \,, \tag{5}$$

where $v_i = V_i/V_0$ is the scaled molecular volume of fragment i, $\beta = 1/k_bT$, α_{EV} and α_{res} represent two global adjustable parameters and $\Delta G_{res,ij}$ is the residual Gibbs energy of mixing of a *hypothetical* equimolar mixture of fragments i and j. The Gibbs energy of mixing was calculated through COSMO-RS calculations,[82,83] using the charge envelope of the fragments (the so-called sigma profiles). The COSMO charge envelope is here computed via a modified version of AM1,[84–86] using atomic partial charges derived from the charge equilibration (QEq) method.[87] By definition the residual Gibbs energy of mixing between identical fragments is zero, i.e., $\Delta G_{res,ii} = 0$, thus it follows trivially that a_{ii} is reduced only to the excluded volume contribution and, in particular, for water bead self-interaction $a_{ww} = \alpha_{EV}$. It is also important to point out here that the bead-size effect is taken into account in the definition of DPD a_{ij} parameter given in Eq. (5) by considering the fragment volume scaled with respect to the reference volume, V_0, of a cluster of three water molecules. This allows to consider a constant DPD base unit of length, h, for all fragments irrespective of size or composition. As in the original AFP work,[81] here the value of h is assumed equal to 7.65 Å as the yardstick for length in DPD approach. This value corresponds to five 3-mer water clusters per cell of size h^3, or, in terms of the DPD dimensionless unit system, this corresponds to a density of 5 for water under ambient conditions. The soft-core repulsion

potential employed here is devoid of the short-range Lennard-Jones divergence. Also, the typical long-range electrostatic Coulomb term is avoided completely, through using the close-contact electrostatic interaction of the COSMO model. Both interactions are therefore replaced by a soft repulsive potential that is local, with a length scale limited to the cutoff, r_c. Hence, in AFP approach the fragment-specific chemical information is condensed in only one parameter: the DPD a parameter. The magnitude of the repulsion (not the spatial extension) is modified depending on the volume of the underlying molecular fragment, and residual interactions. In order to map the characteristics of the atomistic models into the DPD system, MD simulations of protein in water and oil bulks were used to extract molecular characteristics such as radial distribution functions as well as the distributions of lengths and angles for molecules bonded with length and angular bonds. To make MD and DPD models physically comparable, it is necessary to map atomistically detailed trajectories into their corresponding coarse-grained representations considering a length scale factor, h, to convert atomistic coordinates and MD box dimensions into a CG model. When dealing with the triglyceride and the protein in which their fragmentation information has been already well-defined through the AFP approach, the mapped MD trajectories of such molecules are easily determined by replacing the fully atomistic coordinates with the centre-of-mass positions of provided molecular fragments. However, in the case of atomistic water models, where the water particles move independently, their CG representation has to be dynamically identified. Therefore, a clustering method is required to enable the mapping of multiple water molecules into a single CG bead. Here, the water molecules clustering algorithm proposed by Pieczywek, Płaziński, and Zdunek[88] was employed, which is based on a step-wise iterative nearest neighbour search algorithm. The number of water molecules per bead in all clusters is kept constant and equal to the degree of coarse-graining employed here, i.e., a 3 to 1 CG ratio, corresponding to the number of clustering steps performed for each simulation time frame. This represents the major advantage compared to other approach where, instead, the total number of beads in the system have to be provided,[89] leading to some issues converging with the desired number of equally sized clusters. Very briefly, as the algorithm initialization, a grid of fixed-size cubes was superimposed onto the MD simulation box and initial positions of bead centers were generated by randomly choosing coordinates of water molecules from the first time frame. For each step of the algorithm, an iterative search for the unique nearest water molecule was carried out in the area adjacent to the unit

cell in which the coarse-grained bead is located. The unique nearest water molecule was defined by means of the Euclidean distance from the center-of-mass of a CG bead. When all of the CG beads had the same number of molecules assigned to them (equal to the CG ratio), the algorithm finished and the positions of the beads were updated by calculating the center-of-mass of the molecular clusters. Hence, for each MD simulation time frame, the water molecules were divided into equally sized groups based on their proximity.

The mapped MD trajectories were used to extract radial distribution functions (RDFs) of coarse-grained molecules. Thus, using the AFP method as a basis, a further DPD parameter calibration was carried out by using the MD RDFs as reference curves to be compared with those extracted from DPD simulations. Since the RDF is solely determined by the conservative force,[90] the repulsion force coefficients was adjusted to match MD and DPD RDFs. As the specific fragment pair interactions were defined in Eq. (5), the global adjustable parameters which serve to define the mutual repulsive interaction between *all* the beads belonging to a single type of molecule can be used to calibrate the DPD model. In particular, α_{EV} and the cutoff distance, r_c, were used as fitting parameters, while for all the fragment pairs the DPD-sigma parameter was set to the standard value of 3.0[29] and α_{res} was kept equal to 6.1 as in the original AFP work.[81] Therefore, from both MD and DPD simulations of protein in water and in oil bulk, only RDFs referring to all beads belonging to water, oil, and protein were extracted and the results of the calibration are presented and discussed in Section V. Obviously, from simulations of the binary systems only water-water, oil-oil, water-protein, and oil-protein interactions can be exactly calibrated. However, the remaining interactions, i.e., oil-water and protein-protein, must be determined to build the DPD model of the ternary system. In particular, the oil-water α_{EV} value was obtained by simply fitting the experimental interfacial tension between purified soybean oil and water,[51] found to be equal to 31-32 mN/m and independent on the presence of salt.[91] For the protein-protein repulsive interaction, the same α_{EV} value of water-protein was arbitrarily chosen as a first guess. This value could be of paramount importance since the self-protein interaction may effect the structural configuration of the protein as well as equilibrium and dynamics properties of the ternary system. The study of protein-protein interactions needs therefore a deeper insight, which could be the scope of future works.

The parametrization of intra-molecular interactions (bonds and angles) of CG molecules was also based on MD simulations. The basic concept is to construct the distribution

function of each of these quantities from atomistic model simulations. By using again the molecular fragment information obtained via AFP within the atomistic MD trajectories, the distribution functions of bond lengths and bending angles were calculated based on the center of the coarse-grained fragments. Then, a robust and fast approach when dealing with hundreds of bond and angle interaction types generated from the automated coarse-graining procedure employed in this work (AFP) is to derive parameters from distributions directly,[43,92,93] instead of fitting each bond-stretching and bending angle potential obtained from Boltzmann inversion with a harmonic approximation.[94] When assuming a harmonic bond potential (Eq. (2)), the resulting distribution is a Gaussian that can be equated with the distribution of the bonds. It follows that the equilibrium bond length, l_H, is simply the average of the distribution and the bond constant, k_S, can be expressed in terms of the standard deviation of that distribution.[43,92,93] For angles, the same would hold for harmonic potentials (Eq. (3)), except that the angle is bounded between 0° and 180°. This means that the distribution for a purely harmonic potential will not be a Gaussian, but rather a Gaussian that is cut off at 180°. However, a reasonable procedure is to simply take the angle where the distribution is maximal and treating that as if it were the average, equating it to the equilibrium angle, θ_H. Taking the standard deviation to calculate the angle potential strength, k_A, also is reasonable.[43] It is important to point out that this procedure is not able to capture multiple maxima and/or minima in bond and angle distributions from atomistic MD simulations.[43] Without a further modification, bonded interaction parameters directly derived from MD distributions can be used in DPD simulations by using a shorter time step than that typically used in DPD works (i.e., $\Delta t = \mathcal{O}(0.01)$[29]). In fact, the exact replication of the MD structures required the strength of bonds to become too large for relatively long time step, resulting in unstable simulations.[41] Therefore, in order to preserve the distance and angular bond characteristics, a dimensionless time step of $\Delta t = 0.001$ was used to integrate the DPD equations of motion.[88]

C. DPD simulation parameters

To avoid using excessively large or small numbers and to simplify the calculations, DPD systems were usually scaled by arbitrarily chosen base units. As it was already discussed in the previous subsection, the conversion factor $h = 7.65$ Å was here employed as base

unit of length. The mass of one water bead consisting of three water molecules equal to 8.974×10^{-26} kg, was used as the base mass unit. Both MD and DPD simulations were performed at ambient temperature (298 K), giving $k_bT = 4.11 \times 10^{-21}$ J used as the base unit for energy, where k_b is the Boltzmann constant. The base time unit τ was estimated by evaluating the diffusion coefficient. This is computed from both MD and DPD simulations by using the standard mean-squared displacement (MSD) method through the well-known Einstein relation.[25] By defining the scaling factor $S = D_{W,Exp}/D_{W,\mathrm{DPD}} = 7.63 \times 10^{-9}$ m^2/s, where $D_{W,Exp}$ and $D_{W,\mathrm{DPD}}$ are respectively the experimental water self-diffusion coefficient at ambient conditions and the simulated one via DPD, the base unit used to convert the reduced DPD time into real unit reads as follows:

$$\tau = \frac{h^2}{S} \approx 77 \text{ ps}\,. \tag{6}$$

Therefore, the real protein diffusion coefficient computed from DPD simulations was simply determined by multiplying the simulated value for the scaling factor, S.[95] Since no experimental measurement is available in the literature, the protein diffusion D computed via MD and DPD were compared with three correlations proposed for the prediction of protein diffusion coefficients in free solution, based on the molecular weight M (Eq. (7a)[96]), on the radius of gyration R_g (Eq. (7b)[97]), and on both the molecular weight and the radius of gyration of the protein (Eq. (7c)[98]), respectively:

$$D = 8.34 \times 10^{-8} \left(\frac{T}{\eta M^{1/3}} \right) , \tag{7a}$$

$$D = 5.78 \times 10^{-8} \left(\frac{T}{\eta R_g} \right) , \tag{7b}$$

$$D = 6.85 \times 10^{-8} \left(\frac{T}{\eta \sqrt{M^{1/3} R_g}} \right) , \tag{7c}$$

where η is the solvent viscosity, i.e., 0.894 and 50 cP at 25 °C for water[99] and for soybean oil,[100] respectively.

Several DPD simulation configurations were investigated in this work. In order to match the coarse-grained characteristics from MD simulations, the binary systems were reproduced using DPD. The MD box was scaled according to the length conversion factor h and one CG protein molecule was located at its center. According to the binary environment, the box was then filled with water beads or oil CG molecules to obtain the overall DPD density $\rho = 5$. The DPD simulations were performed with an equilibration period of 10^5 steps, then

followed by a production phase of 10^6 steps, saving particle trajectories every 250 steps. Once DPD parameters have been calibrated as explained in the previous subsection, two DPD configurations of the interfacial system were carried out in order to study the equilibrium properties at increasing protein interface concentration c_i and the protein adsorption at the oil/water interface. Both initial configurations consisted of a central water phase segregated by two oil phases, thus forming two planar interfaces in equidistant yz-planes. The 50/50 oil-to-water bead ratio was kept constant for all DPD simulations and both the number of water beads and oil CG molecules was adjusted to keep the same overall DPD density of 5 when the protein molecules were also added in the DPD box. The equilibrium simulations were conducted with increasing protein interface concentration c_i, which is simply calculated by multiplying the number of the protein molecules at each interface for the protein molecular mass M, divided for the constant interface yz-area expressed in real units. The protein molecules were initially located at the oil–water interface to make sure that both interfaces contain the same number at equilibrium in order to perform averages on both interfaces. For equilibrium DPD simulations, the box was an orthorhombic cell of reduced size $L_x \times L_y \times L_z$, where $L_y = L_z = 32$ and L_x was properly adjusted up to 52 based on the protein molecule number to allow both interfaces to be independent. Simulations were run for 2.5×10^5 equilibration steps and for a production period of 10^6 steps, saving time frame data for post-processing every 500 steps. Here the interfacial tension, σ_{DPD}, was computed by integrating the difference between normal and tangential stress across the interface separating the segregated components.[101] Thus, if the normal to the interface lies along the x-direction, the interfacial tension is deduced from the local components of the pressure tensor:

$$\sigma_{\mathrm{DPD}} = \frac{1}{2}\int \left(p^*_{\mathrm{N}} - p^*_{\mathrm{T}}\right) dx = \frac{1}{2}\int \left(p^*_{xx} - \frac{1}{2}\left(p^*_{yy} + p^*_{zz}\right)\right) dx\,, \tag{8}$$

where p^*_{N} and p^*_{T} are the normal and tangential components of the pressure tensor profile in reduced DPD units. The factor 1/2 before the integral sign is due to the presence of two symmetric interfaces in the DPD simulation box when using periodic boundary conditions. Since the oil droplets of a food emulsion have a diameter of the order of microns,[3] it is reasonable to neglect the curvature effect when modeling the interfacial system at the nano-scale, thus allowing to use the above formula, valid for planar geometry only.[101] The conversion of σ_{DPD} to real units operates as follows: $\sigma_{\mathrm{calc}} = \frac{k_b T}{h^2}\sigma_{\mathrm{DPD}}$. The quantity σ_{calc} can be directly

compared with experimentally measured interfacial tension. The free protein adsorption at the oil/water interface was also studied by locating one protein molecule in the center of an orthorhombic DPD box $L_x \times L_y \times L_z$, where $L_y = L_z = 20$ and L_x was ranged from 40 to 56 in order to properly increase the mutual initial distance between the protein center and the interface. In addition, the adsorption at the oil/water interface was tested for an LDL-like particle configuration by initially creating a small droplet of 15 oil CG molecules surrounded by one protein molecule. These latter DPD simulations were performed with 2×10^5 equilibration steps and a production period of up to 4×10^6 steps, saving simulation time frames every 500 steps to check if the protein adsorption has taken place.

Apart from the water cluster algorithm, which was performed in the MATLAB environment,[88] all MD and DPD simulation setup, runs, and post-processing analyzes were conducted within the CULGI software package,[102] together with all other tools and algorithms employed in this work.

V. RESULTS AND DISCUSSION

The results of the DPD model calibration explained in Section IV B are shown in Figure 3, where the distance is expressed in real units, and in Table I. Using the MD RDFs as references, the DPD RDFs were adjusted in order to best match curve heights and shapes by calibrating both α_{EV} and r_c of molecule bead pairs. These two terms define both the magnitude (via Eq. (5)) and the spatial extension of the repulsive force (Eq. (1)). Typically, in standard DPD the cutoff value also represents the base unit of length and, therefore, is often set equal to 1 in dimensionless unit.[29] In contrast, here the dimensionless value of r_c resulting from fitting the first peaks of RDF curves shown in Figure 3 was found to be equal to 0.7. Hence, the cutoff, r_c, and the length factor, h, were decoupled in order to assure both the constant DPD number density of 5 and the repulsive force calibration. The results of α_{EV} fitting are summarized in Table I. Although the oil-water α_{EV} turned out to be substantially smaller than all the others in Table I, the overall repulsion between water and oil beads was properly reproduced due to the two contributions in Eq. (5) and a cutoff, r_c, equal to 1 in this specific case, in which a sophisticated calibration was not needed.

The molecular model is tested and the main findings are presented here, paying a particular attention to verify the emulsifying behaviour of Apovitellenin I at the oil/water interface.

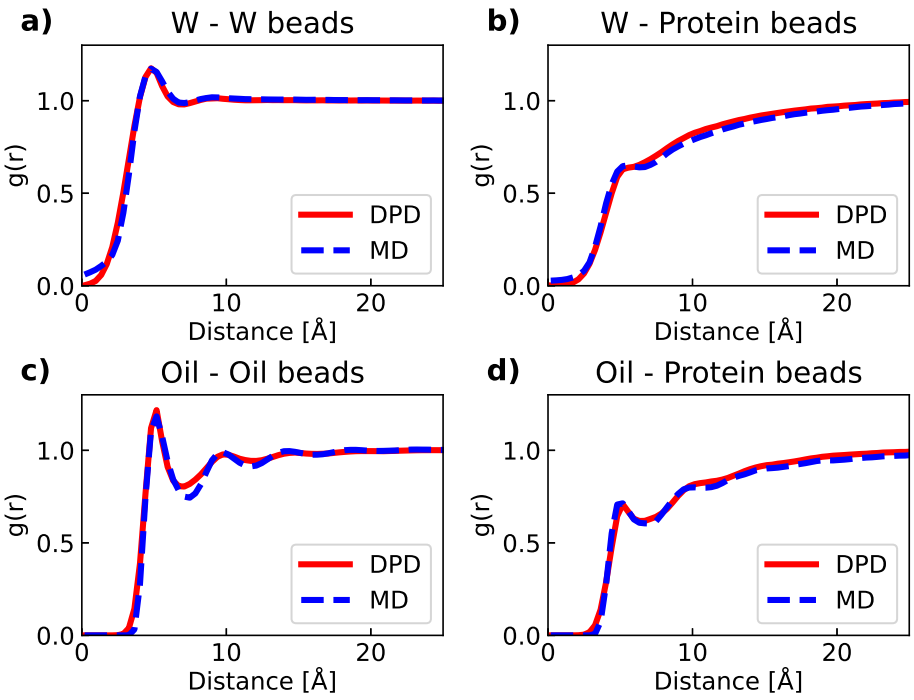


FIG. 3. Results of the DPD parameter calibration of water-water (a), water-protein (b), oil-oil (c), and oil-protein (d) interactions based on matching RDFs of the mapped MD reference model (dashed blue line) with corresponding RDFs extracted from DPD simulations (solid red line).

First, preliminary structural and dynamic quantities of the protein are estimated by performing both MD and DPD simulations of one protein molecule in bulk phases. Then, the DPD simulation results of the ternary system are discussed in terms of both equilibrium and dynamic aspects.

Table II reports End-to-End distance and radius of gyration mean values and standard deviations of Apovitellenin I in water and oil bulks computed via MD and DPD simulations. The MD values were averaged over the simulation time, meanwhile 10 independent DPD simulations with the same initial configuration were carried out from which the reported values are extrapolated by computing their respective arithmetically averaged frequency distributions. It is important to recall that Apovitellenin I is modeled here as a homodimer,

TABLE I. Values of the global parameter α_{EV} used in Eq. (5) to define the mutual repulsion between all the beads belonging to water, oil, and protein in the DPD model of this work. The cutoff distance, r_c, is equal to 0.7 unless otherwise specified.

$\alpha_{\mathbf{EV}}$	W	Oil beads	Protein beads
W	25[a]	-	-
Oil beads	8.5[b]	100	-
Protein beads	40	100	40[c]

[a] Exactly corresponding to a_{ww}.

[b] Value obtained by fitting experimental interfacial tension between soybean oil and water,[91] with a cutoff distance, r_c, equal to 1.

[c] Arbitrarily chosen equal to the water-protein value.

so the two polypeptide chains are labeled as 1 and 2 in Table II where the End-to-End distance is that between the N-terminal and the C-terminal of each chain, while the protein radius of gyration refers to the homodimer itself. By looking at mean values reported in the Table II, it can be noticed that a good accordance between the two molecular technique is achieved. The largest differences are only related to the chain 1 End-to-End distance and the radius of gyration of the protein in water environment. The MD radius of gyration data suggest that the protein is more compact in water than in oil environment, while an opposite trend is detected via DPD. Another considerable dissimilarity regards the standard deviation values calculated with the two techniques. Both MD and DPD were able to identify a smaller error of the respective quantity in oil than in water bulk meaning a less flexible protein structure in the former environment than in the latter. However, all the DPD standard deviations are significantly higher than those obtained via MD. This might be due to two main reasons. First, combining distributions from independent DPD simulations into a single arithmetically averaged distribution involves that the variance of the averaged one is always at least as large as the minimum of the variances of input distributions.[103] Secondly, the soft potential applied in the DPD force field can provide less steric hindrance compared to the Lennard-Jones potential used in MD. Moreover, the higher variation in DPD than MD may be related to the lack of additional bond constraints for intra-protein molecular interaction[46,104] in the present DPD framework, thus assuming a completely flexible nature

TABLE II. End-to-End distance and radius of gyration mean values and standard deviations of Apovitellenin I in water and oil bulk phases computed via MD and DPD simulations.

			MD	**DPD**[a]
Apovitellenin I in Water	**End-to-End distance** [Å]	Chain 1	50.46 ± 2.93	62.06 ± 18.84
		Chain 2	69.84 ± 2.82	65.87 ± 18.37
	Radius of gyration [Å]		24.98 ± 0.50	35.67 ± 5.26
Apovitellenin I in Oil	**End-to-End distance** [Å]	Chain 1	57.22 ± 0.96	58.38 ± 14.59
		Chain 2	64.49 ± 0.49	63.39 ± 14.20
	Radius of gyration [Å]		27.04 ± 0.13	29.39 ± 2.84

[a] The reported values are extrapolated from respective frequency distributions arithmetically averaged over 10 independent simulations.

of Apovitellenin I without a specific secondary structure. This latter explanation can be also given to the opposite trend of the mean value of the protein radius of gyration reported by means of MD and DPD in the two bulk phases.

Table III shows the comparison of diffusion coefficient values, D, of Apovitellenin I in water and oil bulk calculated by means of three correlations found in the literature (Eq. (7)[96–98]) and computed from MD and DPD simulations. MD protein radius of gyration in the respective solution reported in Table II are used in expressions based on such a property (Eqs. (7b) and (7c)). Table III also reports the diffusion errors in terms of ranges of variability. In particular, the accuracy of correlation results was taken from the corresponding previous works,[96–98] meanwhile MD and DPD uncertainties were directly estimated from simulations. As it can be seen, both correlation and simulation results show a difference in the protein diffusion coefficient of at least one order of magnitude between the water and oil solution. The larger diffusion coefficient in water than in oil is mostly likely due to the larger oil viscosity than the water one that can be responsible of the limited mobility of Apovitellenin I in oil phase. By comparing the results for water environment, MD and DPD give a remarkable agreement between them although all the correlations indicate a slightly higher value. On the other hand, the accordance on simulation results is relatively lost when dealing with oil bulk, but the DPD value is noticeably close to those predicted via empirical

TABLE III. Comparison of diffusion coefficient values of Apovitellenin I in water and oil bulk as predicted by three correlations (Eq. (7)) and as computed from MD and DPD simulations.

$D \times 10^{-12}$ [m^2/s]	**Correlation results**			**MD**	**DPD**[a]
	Eq. (7a)[96]	Eq. (7b)[97]	Eq. (7c)[98]		
Apovitellenin I in Water	82.3 – 127.2	65.7 – 89.0	80.6 – 97.0	22.7 – 24.0	20.9 – 26.1
Apovitellenin I in Oil	1.47 – 2.27	1.10 – 1.45	1.40 – 1.65	0.296 – 0.297	1.97 – 2.92

[a] Averaged on 10 independent simulations.

correlations. It is also important to highlight here that the diffusion coefficient of proteins in solution computed by molecular simulation techniques tends to be underestimated when compared to the true value.[105] That being said, although it is really hard to validate the data reported in Tables II and III without experimental evidence, it is possible to affirm that molecular modeling techniques lead to very reasonable results.

Let us move now on the discussion of the ternary system made by oil, water and protein via DPD simulations. In order to study the equilibrium properties of such a system, the starting configuration of the DPD box consists of two symmetrical interfaces due to the periodic boundary conditions applied in the three directions. Figure 4 shows the equilibrated DPD boxes representing the oil-water interface where Apovitellenin I acts as the surfactant at increasing protein surface concentrations and by highlighting the planar interfaces. Figure 5 reports profiles of the number density of oil, water and protein (i) and stress profiles (difference between normal and tangential pressures, $p^*_{\mathrm{N}} - p^*_{\mathrm{T}}$) (ii) along the normalized x-direction normal to the interfaces at increasing protein interface concentrations corresponding to those of Figure 4 (a, b, and c). The dashed lines represent the interface position in the initial DPD configuration. It points out the initial phase separation and the resulting mutual interpenetration of each component at equilibrium. The profile plots show the symmetry of the equilibrated ternary system and define the interfacial region that contains the protein layer and the bulk region that lies between the interfaces. As it can be seen in Figures 5 a.i), b.i) and c.i), the most interesting result is that the protein molecules penetrate the water bulk to a much larger extent than the oil bulk, especially at higher interface protein concentrations. As expected by looking at Table I, this is mostly likely due to the higher overall repulsion between protein and oil than that between protein and wa-

ter. By looking at Figures 5 a.ii), b.ii) and c.ii), the mechanical equilibrium of the system is reached in both oil and water phases since the stress profiles fluctuate with small oscillations around zero in the bulk regions. As a consequence, the local contribution to the interfacial tension is located only at the interfaces, with an increase in the stress in the protein region. Therefore, the accuracy of the interfacial tension calculation is achieved. In order to avoid size effects along x-axis and allow both interfaces to be independent, the bulk phases must be large enough to reach the mechanical equilibrium by increasing the L_x dimension as the number of protein molecules increases keeping the interface yz-area constant.

Figure 6 reports the trend of the protein layer thickness (a), the protein mean radius of gyration, $\langle R_{g,Protein} \rangle$ (b), and, finally, the interfacial tension (c) as a function of the interface concentration of Apovitellenin I. Three independent DPD runs were carried out and the averaged values are shown together with the corresponding standard deviations. Error bars are generally smaller than symbols indicating a high reproducibility of the current DPD model. The most remarkable result is the interfacial tension decrease as the protein interface concentration increases. This trend clearly evidences the capability of Apovitellenin I to behave as a surfactant. As expected, the minimum value of the interfacial tension is reached at the saturation of the interface, which does no longer allow direct interactions between oil and water. As shown in Figure 6c, the saturation is obtained at the protein interface concentration equal to 3.0-3.5 mg/m^2, where the interfacial tension ranges between 8 and 10 mN/m. The maximum protein coverage (about 3.0 mg/m^2) of the present system is in line with that observed in an experimental work where the oil-in-water emulsion stabilized by flexible proteins (caseins) was studied.[106] Moreover, Dauphas *et al.*[69] reported that the equilibrium interfacial tension for the oil-water interface with adsorbed LDL film at pH 3 is 9.5 mN/m, which is markedly consistent with our result. It is also important to highlight that, when no protein molecules are added, the interfacial tension between water and oil phase modeled as homotriglycerides is accurately reproduced in agree with the experimental value.[1,91] $\langle R_{g,Protein} \rangle$ (Figure 6b) is computed from the mean value of the protein R_g distribution, further averaged over 3 DPD simulations. Therefore, $\langle R_{g,Protein} \rangle$ provides information about the conformation and packing of protein molecules at the interface. At low concentration, the protein radius of gyration is higher than its corresponding DPD value in both bulk situations (see Table II). This can indicate that, when very few protein molecules are absorbed at the oil-water interface, they assume a more elongated conformation than that in water or

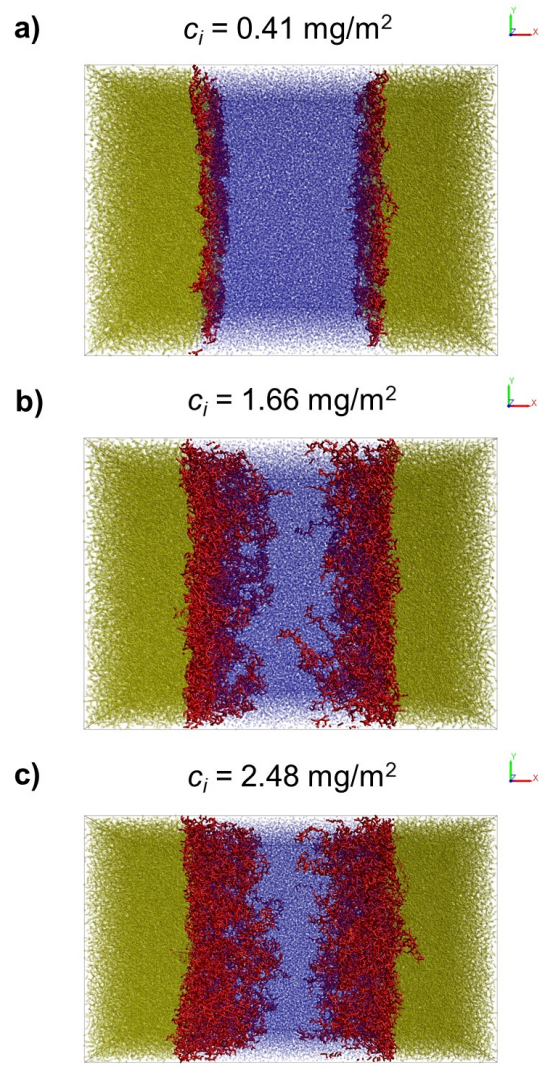


FIG. 4. Snapshots of equilibrated DPD boxes of the interface between oil (yellow) and water (blue) where Apovitellenin I (red) acts as the surfactant at increasing protein interface concentration, c_i (a, b, and c).

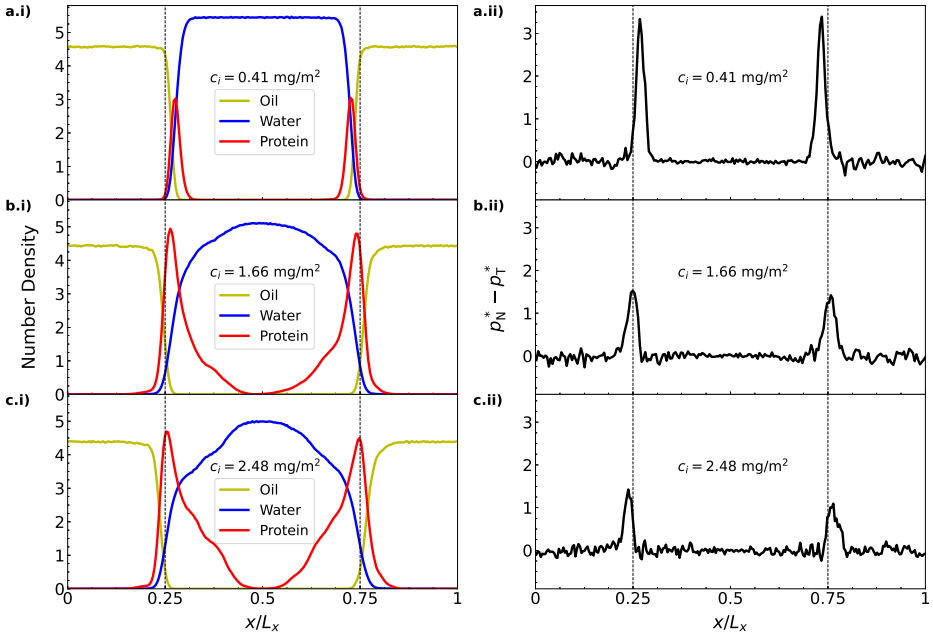


FIG. 5. Profiles of the number density of oil, water and protein (i) and of the difference between normal and tangential pressures, $p_N^* - p_T^*$, (ii) along the normalized x-direction normal to the interfaces at increasing protein interface concentrations (a, b, and c).

oil solution. Meanwhile, at increasing protein concentration, the mean radius of gyration of Apovitellenin I at the interface decreases to a stable value and becomes comparable to that in free solution. Thus, the packing mode of protein molecules at interface can be considered similar to that observed in bulk phases, when the protein interface concentration is high. Regarding the thickness of the protein layer (Figure 6a), it is directly derived from the width of the protein density profile along the x-direction normal to the interface surface (see Figures 5i for reference). As expected, the protein layer thickness increase from 2 to 13 nm as the protein interface concentration increases until the saturation of the interface where the maximum and stable value for the thickness is reached. Fang and Dalgleish [106] reported that the adsorbed layer of casein molecules at the maximum coverage of the oil-water interface was about 10 nm thick so that the protein molecules protrude further into the solution, as also shown in this work (Figures 4 and 5i). Moreover, previous works[107,108] found that the

interfacial layer surrounding oil droplets in mayonnaise have an average thickness of around 14 nm, which is comprised of surface-active proteins and lecithin-protein granules from egg yolk. Those findings are reasonably in accordance with our results. It is also straightforward to point out here that the emulsifier behaviour of only one LDL apoprotein is tested since it is identified as one of the most surface-active. LDL phospholipids may also have an effect on the interfacial tension of LDL-based emulsion by a further decrease of its saturation value.

In order to study the adsorption of Apovitellenin I at the oil-water interface, DPD simulations of a box containing two equidistant interface and one free protein molecule initially located in the center of the water phase were carried out. So, the protein diffusion from the aqueous environment towards the oil-water interface is investigated as represented in Figure 7, where an illustrative example shows the three main steps of the protein adsorption mechanism . First, the protein moves to the interface (a), then a portion of the molecule initiates the protein adsorption (b) and, after a certain time, Apovitellenin I is totally adsorbed at the oil-water interface (c). Apparently, there is no specific reason for the protein to be preferably adsorbed at the right rather than at the left interface as the two sides are symmetrical. Moreover, the protein desorption has not been observed meaning that the adsorption process is most likely irreversible as also reported in previous experimental works.[7,67] To estimate the time required by a protein molecule to be fully absorbed as a function of its distance from the oil-water interface, multiple DPD simulations were performed by increasing the box size in the x-direction normal to the interfaces and the results are summarized in Figure 8. Since the oil-to-water bead ratio is kept constant and the protein molecule is placed in the center of the water phase at the beginning of the simulation (see Figure 7 for reference), the abscissa of Figure 8 represents the initial distance between the geometric center of the protein molecule and the oil-water interface. The y-coordinate of Figure 8 expresses the time elapsed from the start of the simulation to the moment in which the protein molecule is totally absorbed at one of the interfaces and it is estimated by visual inspection of simulation time frames. As also done in Figure 6, for each point three independent DPD simulations were carried out from which the mean value and the standard deviation were extracted. Although the error bars are relatively large, a linear trend passing through the origin of the axes can be identified in the range of investigated distances. The slope of 0.978 ns/Å can be considered as an estimation of the required time of a liberated Apovitellenin I molecule to be totally adsorbed at a free interface as a function of their mutual distance.

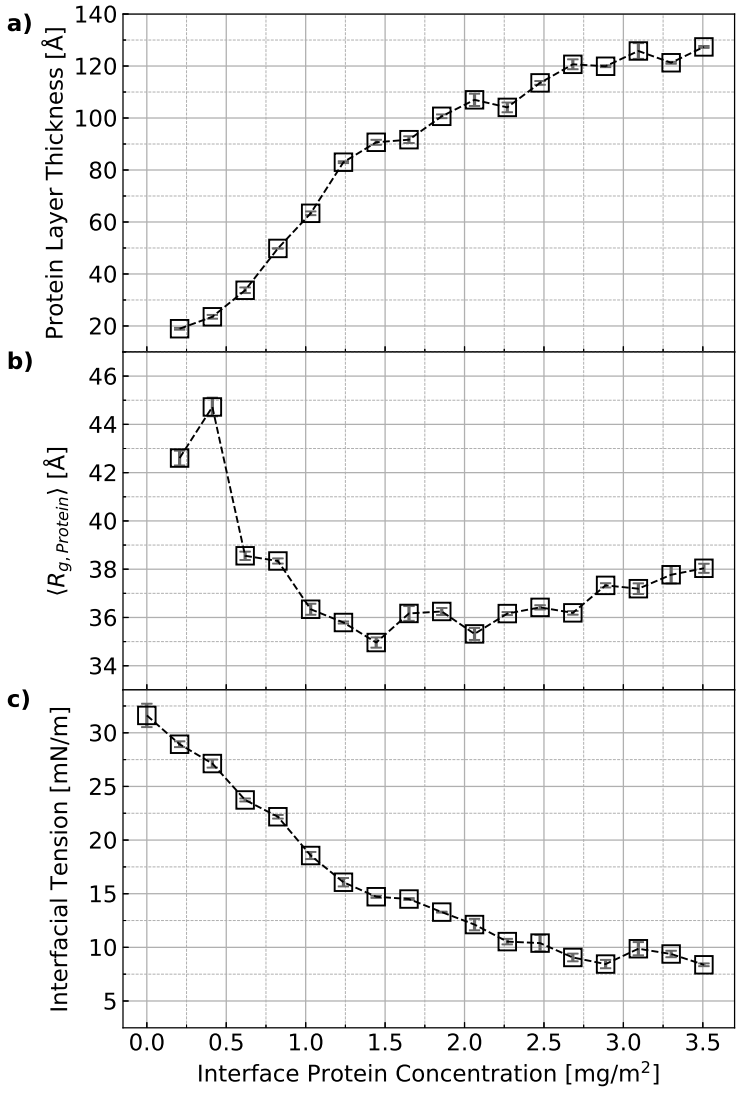


FIG. 6. Protein layer thickness (a), protein mean radius of gyration, $\langle R_{g,Protein} \rangle$ (b), and interfacial tension (c) as a function of the interface concentration of Apovitellenin I. Error bars are estimated from three independent DPD simulations.

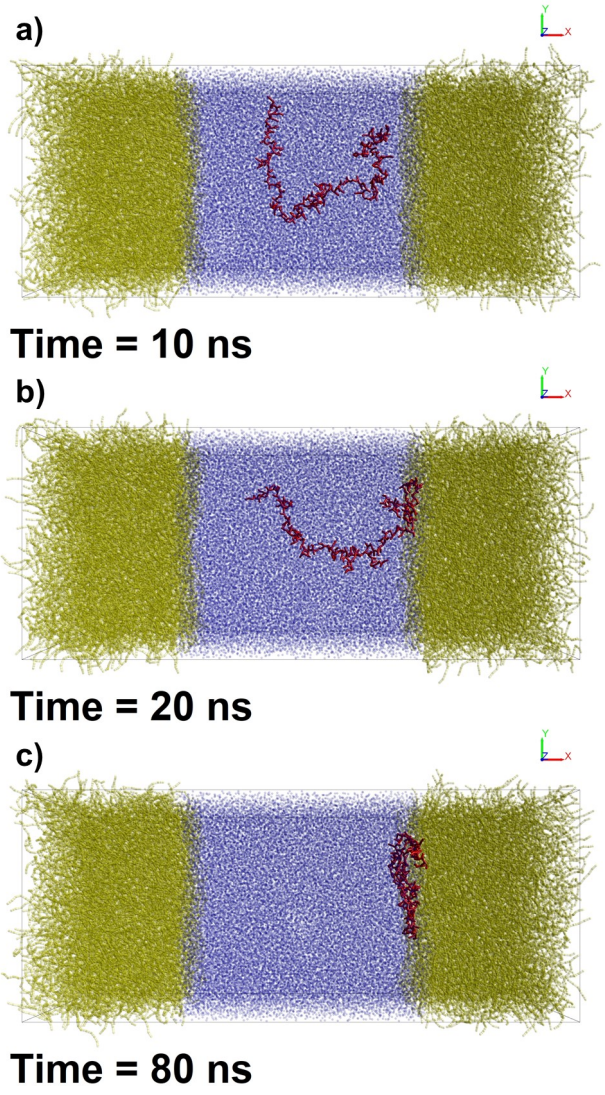


FIG. 7. Snapshots of the DPD simulation showing an illustrative example of the adsorption process of Apovitellenin I (one free molecule in red) at the interface between oil (yellow) and water (blue). The most significant steps of the adsorption mechanism are successively represented in a, b, and c.

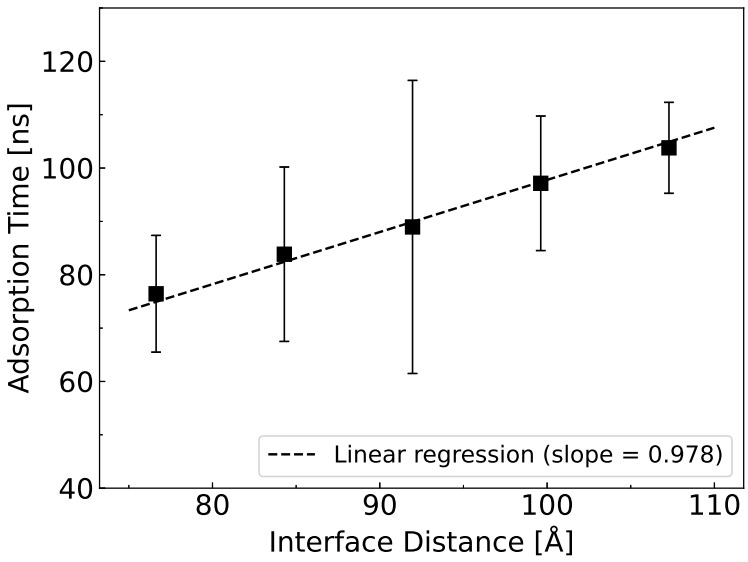


FIG. 8. Trend of the time required by one free molecule of Apovitellenin I to be fully adsorbed at the oil-water interface as a function of the initial distance between the protein geometric center and the oil-water interface. Error bars are estimated from three independent DPD simulations.

As already stated, LDL particles act as vectors of surfactant constituents (e.g., Apovitellenin I) that could not be soluble in water until they reach the interface. Therefore, a DPD simulation of a LDL-like particle with a lipid core surrounded by one molecule of Apovitellenin I was performed and the adsorption mechanism at the oil-water interface was tested. Although it is clear that this structure is far from being a realistic representation of a LDL particle, surprisingly the adsorption process proposed by Anton[7] is qualitatively reproduced as it can be seen in Figure 9 (Multimedia view). Indeed, first the LDL-like particle diffuses in the water bulk (a) until the protein situated on the particle surface comes into contact with the interface causing the unfolding of the LDL-like particle (b). Thus, the protein molecule initiates the LDL-like particle disruption by its anchorage at the oil-water interface. Then, the neutral lipids are released from the particle core and merge with the oil phase, while the protein molecule adsorbs at the interface (c). Since the system dimensions of Figure 9 (Multimedia view) are the same of those represented in Figure 7, a general comparison can be made between two configurations, namely the liberated protein and the LDL-like

particle. In particular, the adsorption time of the LDL-like particle is significantly higher than that of the free protein. This can be intended as a greater stability of Apovitellenin I when surrounding the LDL-like particle rather than as a free molecule, also confirming that the liberated protein is supposed to be almost insoluble in water. Finally, it is important to remark that the representation of the LDL-like particle here presented must be considered qualitative, since both LDL size and its specific composition, namely including also the lipid distribution of the LDL core and all surfactant components situated on the LDL surface (e.g., phospholipids and other apoproteins), were not considered in the analysis.

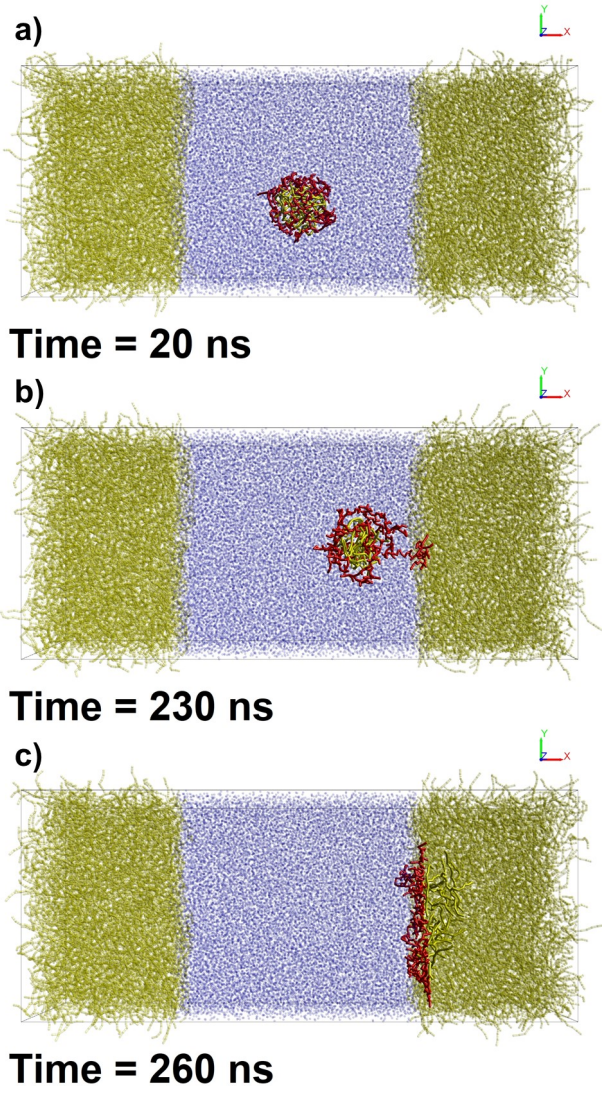


FIG. 9. Snapshots of the DPD simulation showing the adsorption process of a LDL-like particle with a lipid core (bright yellow) surrounded by one molecule of Apovitellenin I (red) at the interface between oil (yellow) and water (blue). The most significant steps of the adsorption mechanism are successively represented in a, b, and c (Multimedia view).

VI. CONCLUSIONS

Although egg yolk is widely used as an emulsifier in many food emulsion preparations, little experimental research on emulsifying properties of its individual components has been carried out since their extraction and isolation from the egg yolk complex matrix turned out to be difficult. Hence, this work focuses on the molecular model of an oil/water interface stabilized by one of the most surface-active protein of egg yolk LDLs, called Apovitellenin I. In order to take into account the system size, composition and the equilibration time needed by macro-molecules to re-arrange at interfaces, the molecular modeling technique here proposed is the Dissipative Particle Dynamics approach. Once the chemical species were determined, especially the biomolecule that should act as a surfactant at the oil/water interface, an automated coarse-graining procedure was carried out on the molecules involved in the ternary system. In DPD systems the intended physical properties are determined by means of a parameter calibration, which was here based on coupling DPD with all-atom Molecular Dynamics simulations of a single protein molecule in two different solvents, water and oil. Thus, both inter- and intra-molecular interactions employed in the DPD system are solely determined by matching the structural data from the atomistic simulations. The model was designed to test the most relevant physical properties of the protein studied, especially its emulsifier behavior. The results of MD and DPD simulations are compared in terms of protein structural and dynamics properties (radius of gyration, end-to-end distance, and diffusion coefficient), showing a good agreement between the two molecular techniques. Then, the oil-water interface system was simulated via the DPD technique. In particular, the present molecular modeling approach was able to properly describe the protein surfactant behavior by interfacial tension decrease at increasing protein surface concentration. The protein density profile, layer thickness, and adsorption time at the oil-water interface were also investigated, giving reasonable results in line with experimental evidence of similar protein systems. In addition, the adsorption mechanism of an LDL-like particle is qualitatively reproduced. The modeling method here presented shows how computer molecular simulations can greatly help in the comprehension of food emulsion behavior and, in general, offer the advantage of estimating properties that are difficult to measure experimentally.

These results are encouraging and could be a starting point to explore the role of other surfactant molecules from egg yolk with an analogous molecular modeling method. More-

over, the main findings of this work together with non-equilibrium studies at the meso-scale will pave the way for a better understanding of the breakage and coalescence events of the oil droplets occurring in the food emulsion preparation. This information can be eventually transferred to a computational fluid dynamics study coupled to a population balance model thus achieving a complete, general, and multi-scale digital twin of the food emulsion production process.

SUPPLEMENTARY MATERIAL

See Supplementary Material for a further description of MD and DPD techniques used in this work.

ACKNOWLEDGMENTS

This work was carried out in the context of the VIMMP project (`www.vimmp.eu`), where the entire workflow will contribute to populate a marketplace for generic multiscale and multiphysics simulations. The VIMMP project has received funding from the European Union's Horizon 2020 Research Innovation Programme under Grant Agreement n. 760907. We thank Dr. Piotr Pieczywek (Institute of Agrophysics, Polish Academy of Sciences, Doświadczalna 4, 20-270, Lublin, Poland) for sharing with us the MATLAB code for the water cluster algorithm employed in this work. We greatly appreciate the suggestion of a reviewer to include Figure 1.

CONFLICT OF INTEREST

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are openly available in Zenodo at `http://doi.org/10.5281/zenodo.5703247`, reference number 109.

REFERENCES

[1]D. McClements, *Food Emulsions: Principles, Practice, and Techniques* (CRC Press, Boca Raton, FL, 2005).

[2]S. Friberg, J. Sjoblom, and K. Larsson, *Food Emulsions* (CRC Press, Boca Raton, FL, 2003).

[3]A. Dubbelboer, J. Janssen, H. Hoogland, E. Zondervan, and J. Meuldijk, "Pilot-scale production process for high internal phase emulsions: Experimentation and modeling," Chemical Engineering Science **148**, 32–43 (2016).

[4]A. Dubbelboer, *Towards optimization of emulsified consumer products : modeling and optimization of sensory and physicochemical aspects*, PhD dissertation, Technische Universiteit Eindhoven, Department of Chemical Engineering and Chemistry (2016).

[5]S. Maindarkar, A. Dubbelboer, J. Meuldijk, H. Hoogland, and M. Henson, "Prediction of emulsion drop size distributions in colloid mills," Chemical Engineering Science **118**, 114–125 (2014).

[6]P. Walstra, "Principles of emulsion formation," Chemical Engineering Science **48**, 333–349 (1993).

[7]M. Anton, "Egg yolk: structures, functionalities and processes," Journal of the Science of Food and Agriculture **93**, 2871–2880 (2013).

[8]D. G. Dalgleish, "Food emulsions," in *Emulsions and emulsion stability*, edited by J. Sjoblom (Marcel Dekker Inc., 1996) p. 287–325.

[9]R. W. Burley, "Isolation and properties of a low-molecular-weight protein (apovitellenin I) from the high-lipid lipoprotein of emu egg yolk," Biochemistry **12**, 1464–1470 (1973).

[10]R. W. Burley, "Studies on the Apoproteins of the Major Lipoprotein of the Yolk of Hen's Eggs I. Isolation and Properties of the Low-molecular-weight Apoproteins." Australian Journal of Biological Sciences **28**, 121–132 (1975).

[11]W. Norde, *Colloids and Interfaces in Life Sciences and Bionanotechnology* (CRC Press, Boca Raton, FL, 2011).

[12]C.-A. Palma, M. Cecchini, and P. Samorì, "Predicting self-assembly: from empirism to determinism," Chemical Society Reviews **41**, 3713–3730 (2012).

[13]A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, U.K., 2013).

[14]A. D. Lavino, M. Ferrari, A. A. Barresi, and D. Marchisio, "Effect of different good solvents in flash nano-precipitation via multi-scale population balance modeling-CFD coupling approach," Chemical Engineering Science **245**, 116833 (2021).

[15]S. R. Euston, "Computer simulation of proteins: adsorption, gelation and self-association," Current Opinion in Colloid & Interface Science **9**, 321–327 (2004).

[16]L. A. Pugnaloni, E. Dickinson, R. Ettelaie, A. R. Mackie, and P. J. Wilde, "Competitive adsorption of proteins and low-molecular-weight surfactants: computer simulation and microscopic imaging," Advances in Colloid and Interface Science **107**, 27–49 (2004).

[17]D. Zahn, "On the role of the solvent in biosystems: atomistic insights from computer simulations," Frontiers in Bioscience-Landmark **14**, 3586–3593 (2009).

[18]E. Dickinson, "Structure and rheology of colloidal particle gels: Insight from computer simulation," Advances in Colloid and Interface Science **199-200**, 114–127 (2013).

[19]S. Euston, "14 - Modelling and computer simulation of food structures," in *Food Microstructures*, Woodhead Publishing Series in Food Science, Technology and Nutrition, edited by V. Morris and K. Groves (Woodhead Publishing, 2013) pp. 336–385.

[20]A. Jusufi, "Molecular simulations of self-assembly processes of amphiphiles in dilute solutions: the challenge for quantitative modelling," Molecular Physics **111**, 3182–3192 (2013).

[21]J.-W. Handgraaf and F. Zerbetto, "Molecular dynamics study of onset of water gelation around the collagen triple helix," Proteins: Structure, Function, and Bioinformatics **64**, 711–718 (2006).

[22]A. D. Lavino, N. Di Pasquale, P. Carbone, and D. L. Marchisio, "A novel multiscale model for the simulation of polymer flash nano-precipitation," Chemical Engineering Science **171**, 485–494 (2017).

[23]A. D. Lavino, P. Carbone, and D. Marchisio, "MARTINI coarse-grained model for poly-ϵ-caprolactone in acetone-water mixtures," The Canadian Journal of Chemical Engineering **98**, 1868–1879 (2020).

[24]A. D. Lavino, L. Banetta, P. Carbone, and D. L. Marchisio, "Extended Charge-On-Particle Optimized Potentials for Liquid Simulation Acetone Model: The Case of Acetone–Water Mixtures," The Journal of Physical Chemistry B **122**, 5234–5241 (2018).

[25]D. Frenkel and B. Smit, *Understanding Molecular Simulation*, 2nd ed. (Academic Press, Inc., USA, 2001).

[26]C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2nd ed. (Wiley, New York, 2004).

[27]P. J. Hoogerbrugge and J. M. V. A. Koelman, "Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics," Europhysics Letters (EPL) **19**, 155–160 (1992).

[28]P. Español and P. Warren, "Statistical Mechanics of Dissipative Particle Dynamics," Europhysics Letters (EPL) **30**, 191–196 (1995).

[29]R. D. Groot and P. B. Warren, "Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation," The Journal of Chemical Physics **107**, 4423–4435 (1997).

[30]N. Lauriello, J. Kondracki, A. Buffo, G. Boccardo, M. Bouaifi, M. Lisal, and D. Marchisio, "Simulation of high Schmidt number fluids with dissipative particle dynamics: Parameter identification and robust viscosity evaluation," Physics of Fluids **33**, 073106 (2021).

[31]E. Dickinson and S. R. Euston, "Monte Carlo simulation of colloidal systems," Advances in Colloid and Interface Science **42**, 89–148 (1992).

[32]R. E. Anderson, V. S. Pande, and C. J. Radke, "Dynamic lattice Monte Carlo simulation of a model protein at an oil/water interface," The Journal of Chemical Physics **112**, 9167–9185 (2000).

[33]G. Dalkas and S. R. Euston, "Molecular simulation of protein adsorption and conformation at gas-liquid, liquid–liquid and solid–liquid interfaces," Current Opinion in Colloid & Interface Science **41**, 1–10 (2019).

[34]D. Zare, K. M. McGrath, and J. R. Allison, "Deciphering β-Lactoglobulin Interactions at an Oil–Water Interface: A Molecular Dynamics Study," Biomacromolecules **16**, 1855–1861 (2015).

[35]D. Zare, J. R. Allison, and K. M. McGrath, "Molecular Dynamics Simulation of β-Lactoglobulin at Different Oil/Water Interfaces," Biomacromolecules **17**, 1572–1581 (2016).

[36]D. L. Cheung, "Adsorption and conformations of lysozyme and α-lactalbumin at a water-octane interface," The Journal of Chemical Physics **147**, 195101 (2017).

[37]D. L. Cheung, "Conformations of Myoglobin-Derived Peptides at the Air–Water Interface," Langmuir **32**, 4405–4414 (2016).

[38]S. R. Euston, "Molecular Dynamics Simulation of Protein Adsorption at Fluid Interfaces: A Comparison of All-Atom and Coarse-Grained Models," Biomacromolecules **11**, 2781–2787 (2010).

[39]S. R. Euston, P. Hughes, M. A. Naser, and R. E. Westacott, "Comparison of the Adsorbed Conformation of Barley Lipid Transfer Protein at the Decane-Water and Vacuum-Water Interface: A Molecular Dynamics Simulation," Biomacromolecules **9**, 1443–1453 (2008).

[40]S. R. Euston, P. Hughes, M. A. Naser, and R. E. Westacott, "Molecular Dynamics Simulation of the Cooperative Adsorption of Barley Lipid Transfer Protein and cis-Isocohumulone at the Vacuum-Water Interface," Biomacromolecules **9**, 3024–3032 (2008).

[41]F. Sepehr and S. J. Paddison, "Dissipative Particle Dynamics interaction parameters from *ab initio* calculations," Chemical Physics Letters **645**, 20–26 (2016).

[42]H. Lei, B. Caswell, and G. E. Karniadakis, "Direct construction of mesoscopic models from microscopic simulations," Physical Review E **81**, 026704 (2010).

[43]W. Tschöp, K. Kremer, J. Batoulis, T. Bürger, and O. Hahn, "Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates," Acta Polymerica **49**, 61–74 (1998).

[44]A. Vishnyakov and A. V. Neimark, "Self-Assembly in Nafion Membranes upon Hydration: Water Mobility and Adsorption Isotherms," The Journal of Physical Chemistry B **118**, 11353–11364 (2014).

[45]K. Patterson, M. Lisal, and C. M. Colina, "Adsorption behavior of model proteins on surfaces," Fluid Phase Equilibria **302**, 48–54 (2011).

[46]A. Vishnyakov, D. S. Talaga, and A. V. Neimark, "DPD Simulation of Protein Conformations: From α-Helices to β-Structures," The Journal of Physical Chemistry Letters **3**, 3081–3087 (2012).

[47]K. Okuwaki, H. Doi, K. Fukuzawa, and Y. Mochizuki, "Folding simulation of small proteins by dissipative particle dynamics (DPD) with non-empirical interaction parameters based on fragment molecular orbital calculations," Applied Physics Express **13**, 017002 (2019).

[48]M. Ndao, F. Goujon, A. Ghoufi, and P. Malfreyt, "Coarse-grained modeling of the oil–water–surfactant interface through the local definition of the pressure tensor and interfacial tension," Theoretical Chemistry Accounts **136**, 21 (2017).

[49]A. Ghoufi, P. Malfreyt, and D. J. Tildesley, "Computer modelling of the surface tension of the gas–liquid and liquid–liquid interface," Chemical Society Reviews **45**, 1387–1409 (2016).

[50]A. Khedr and A. Striolo, "DPD Parameters Estimation for Simultaneously Simulating Water–Oil Interfaces and Aqueous Nonionic Surfactants," Journal of Chemical Theory and Computation **14**, 6460–6471 (2018).

[51]A. Maiti and S. McGrother, "Bead–bead interaction parameters in dissipative particle dynamics: Relation to bead-size, solubility parameter, and surface tension," The Journal of Chemical Physics **120**, 1594–1601 (2004).

[52]S.-l. Lin, M.-y. Xu, and Z.-r. Yang, "Dissipative particle dynamics study on the mesostructures of n-octadecane/water emulsion with alternating styrene–maleic acid copolymers as emulsifier," Soft Matter **8**, 375–384 (2012).

[53]F. Alvarez, E. A. Flores, L. V. Castro, J. G. Hernández, A. López, and F. Vázquez, "Dissipative Particle Dynamics (DPD) Study of Crude Oil-Water Emulsions in the Presence of a Functionalized Co-polymer," Energy & Fuels **25**, 562–567 (2011).

[54]L. Rekvig, B. Hafskjold, and B. Smit, "Molecular Simulations of Surface Forces and Film Rupture in Oil/Water/Surfactant Systems," Langmuir **20**, 11583–11593 (2004).

[55]E. G. Perkins, "Chapter 2 - composition of soybeans and soybean products," in *Practical Handbook of Soybean Processing and Utilization*, edited by D. R. Erickson (AOCS Press, 1995) pp. 9–28.

[56]J. Leaver and D. G. Dalgleish, "Variations in the binding of β-casein to oil-water interfaces detected by trypsin-catalysed hydrolysis," Journal of Colloid and Interface Science **149**, 49–55 (1992).

[57]R. Xiong, G. Xie, and A. Edmondson, "Modelling the pH of mayonnaise by the ratio of egg to vinegar," Food Control **11**, 49–56 (2000).

[58]M. Anton and G. Gandemer, "Composition, Solubility and Emulsifying Properties of Granules and Plasma of Egg Yolk," Journal of Food Science **62**, 484–487 (1997).

[59]J. N. Dyer-Hurdon and I. A. Nnanna, "Cholesterol Content and Functionality of Plasma and Granules Fractionated from Egg Yolk," Journal of Food Science **58**, 1277–1281 (1993).

[60]M. Le Denmat, M. Anton, and V. Beaumal, "Characterisation of emulsion properties and of interface composition in O/W emulsions prepared with hen egg yolk, plasma and granules," Food Hydrocolloids **14**, 539–549 (2000).

[61]V. Martinet, V. Beaumal, M. Dalgalarrondo, and M. Anton, "Emulsifying properties and adsorption behavior of egg yolk lipoproteins (LDL and HDL) in oil-in-water emulsions," in *Food emulsions and dispersions*, edited by M. Anton (Research Signpost, Trivandrum, 2002) p. 103–116.

[62]W. Cook and W. Martin, "Egg lipoproteins," in *Structural and Functional Aspects of Lipoproteins in Living Systems*, edited by E. Tria and A. Scanu (Academic Press, London, 1969) p. 579–615.

[63]R. W. Burley and W. H. Cook, "Isolation and composition of avian egg yolk granules and their constituent α- and β-lipovitellins," Canadian Journal of Biochemistry and Physiology **39**, 1295–1307 (1961).

[64]P. Jolivet, C. Boulard, V. Beaumal, T. Chardot, and M. Anton, "Protein components of low-density lipoproteins purified from hen egg yolk," Journal of Agricultural and Food Chemistry **54**, 4424–4429 (2006).

[65]P. Jolivet, C. Boulard, T. Chardot, and M. Anton, "New Insights into the Structure of Apolipoprotein B from Low-Density Lipoproteins and Identification of a Novel YGP-like Protein in Hen Egg Yolk," Journal of Agricultural and Food Chemistry **56**, 5871–5879 (2008).

[66]R. J. Evans, D. H. Bauer, S. L. Bandemer, S. B. Vaghefi, and C. J. Flegal, "Structure of egg yolk very low density lipoprotein. polydispersity of the very low density lipoprotein and the role of lipovitellenin in the structure," Archives of Biochemistry and Biophysics **154**, 493–500 (1973).

[67]M. Anton, V. Martinet, M. Dalgalarrondo, V. Beaumal, E. David-Briand, and H. Rabesona, "Chemical and structural characterisation of low-density lipoproteins purified from hen egg yolk," Food Chemistry **83**, 175–183 (2003).

[68]V. Martinet, P. Saulnier, V. Beaumal, J.-L. Courthaudon, and M. Anton, "Surface properties of hen egg yolk low-density lipoproteins spread at the air–water interface," Colloids and Surfaces B: Biointerfaces **31**, 185–194 (2003).

[69]S. Dauphas, V. Beaumal, A. Riaublanc, and M. Anton, "Hen egg yolk low-density lipoproteins film spreading at the air-water and oil-water interfaces," Journal of Agricultural and Food Chemistry **54**, 3733–3737 (2006).

[70]S. Dauphas, V. Beaumal, P. Gunning, A. Mackie, P. Wilde, V. Vié, A. Riaublanc, and M. Anton, "Structures and rheological properties of hen egg yolk low density lipopro-

tein layers spread at the air–water interface at pH 3 and 7," Colloids and Surfaces B: Biointerfaces **57**, 124–133 (2007).

[71]S. Dauphas, V. Beaumal, P. Gunning, A. Mackie, P. Wilde, V. Vié, A. Riaublanc, and M. Anton, "Structure modification in hen egg yolk low density lipoproteins layers between 30 and 45mN/m observed by AFM," Colloids and Surfaces B: Biointerfaces **54**, 241–248 (2007).

[72]The UniProt Consortium, "Uniprot: the universal protein knowledgebase in 2021," Nucleic Acids Research **49**, D480–D489 (2020).

[73]M. Allen and D. Tildesley, *Computer Simulation of Liquids*, 2nd ed. (Oxford University Press, Oxford, U.K., 2017).

[74]E. Moeendarbary, T. Y. Ng, and M. Zangeneh, "Dissipative particle dynamics: Introduction, methodology and complex fluid applications — A review," International Journal of Applied Mechanics **01**, 737–763 (2009).

[75]D. R. Lide, ed., *CRC Handbook of Chemistry and Physics*, 85th ed. (CRC press, Boca Raton, FL, 2005).

[76]W. L. Jorgensen and J. Tirado-Rives, "The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin," Journal of the American Chemical Society **110**, 1657–1666 (1988).

[77]W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, "Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids," Journal of the American Chemical Society **118**, 11225–11236 (1996).

[78]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," The Journal of Chemical Physics **79**, 926–935 (1983).

[79]U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, "A smooth particle mesh Ewald method," The Journal of Chemical Physics **103**, 8577–8593 (1995).

[80]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," The Journal of Chemical Physics **81**, 3684–3690 (1984).

[81]J. G. E. M. Fraaije, J. van Male, P. Becherer, and R. Serral Gracià, "Coarse-grained models for automated fragmentation and parametrization of molecular databases," Journal of

Chemical Information and Modeling **56**, 2361–2377 (2016).

[82]M. Diedenhofen and A. Klamt, "COSMO-RS as a tool for property prediction of IL mixtures—A review," Fluid Phase Equilibria **294**, 31–38 (2010).

[83]A. Klamt, *COSMO-RS: From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design* (Elsevier, Amsterdam, 2005).

[84]A. Jakalian, B. L. Bush, D. B. Jack, and C. I. Bayly, "Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method," Journal of Computational Chemistry **21**, 132–146 (2000).

[85]A. Jakalian, D. B. Jack, and C. I. Bayly, "Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation," Journal of Computational Chemistry **23**, 1623–1641 (2002).

[86]P. C. Petris, P. Becherer, and J. G. E. M. Fraaije, "Alkane/Water Partition Coefficient Calculation Based on the Modified AM1 Method and Internal Hydrogen Bonding Sampling Using COSMO-RS," Journal of Chemical Information and Modeling **61**, 3453–3462 (2021).

[87]A. K. Rappe and W. A. Goddard, "Charge equilibration for molecular dynamics simulations," The Journal of Physical Chemistry **95**, 3358–3363 (1991).

[88]P. M. Pieczywek, W. Płaziński, and A. Zdunek, "Dissipative particle dynamics model of homogalacturonan based on molecular dynamics simulations," Scientific Reports **10**, 14691 (2020).

[89]K. R. Hadley and C. McCabe, "On the Investigation of Coarse-Grained Models for Water: Balancing Computational Efficiency and the Retention of Structural Properties," The Journal of Physical Chemistry B **114**, 4590–4599 (2010).

[90]E. Zohravi, E. Shirani, and A. Pishevar, "Influence of the conservative force on transport coefficients in the DPD method," Molecular Simulation **44**, 254–261 (2018).

[91]A. G. Gaonkar, "Effects of salt, temperature, and surfactants on the interfacial tension behavior of a vegetable oil/water system," Journal of Colloid and Interface Science **149**, 256–260 (1992).

[92]D. Reith, H. Meyer, and F. Müller-Plathe, "Mapping Atomistic to Coarse-Grained Polymer Models Using Automatic Simplex Optimization To Fit Structural Properties," Macromolecules **34**, 2335–2345 (2001).

[93]Y. Li, B. C. Abberton, M. Kröger, and W. K. Liu, "Challenges in Multiscale Modeling of Polymer Dynamics," Polymers **5**, 751–832 (2013).

[94]V. Agrawal, G. Arya, and J. Oswald, "Simultaneous Iterative Boltzmann Inversion for Coarse-Graining of Polyurea," Macromolecules **47**, 3378–3389 (2014).

[95]J. G. E. M. Fraaije, J. van Male, P. Becherer, and R. Serral Gracià, "Calculation of Diffusion Coefficients through Coarse-Grained Simulations Using the Automated-Fragmentation-Parametrization Method and the Recovery of Wilke–Chang Statistical Correlation," Journal of Chemical Theory and Computation **14**, 479–485 (2018).

[96]M. E. Young, P. A. Carroad, and R. L. Bell, "Estimation of diffusion coefficients of proteins," Biotechnology and Bioengineering **22**, 947–955 (1980).

[97]M. T. Tyn and T. W. Gusek, "Prediction of diffusion coefficients of proteins," Biotechnology and Bioengineering **35**, 327–338 (1990).

[98]L. He and B. Niemeyer, "A Novel Correlation for Protein Diffusion Coefficients Based on Molecular Weight and Radius of Gyration," Biotechnology Progress **19**, 544–548 (2003).

[99]J. Kestin, M. Sokolov, and W. A. Wakeham, "Viscosity of liquid water in the range -8°C to 150°C," Journal of Physical and Chemical Reference Data **7**, 941–948 (1978).

[100]F. C. Magne and E. L. Skau, "Viscosities and Densities of Solvent-Vegetable Oil Mixtures," Industrial & Engineering Chemistry **37**, 1097–1101 (1945).

[101]J. H. Irving and J. G. Kirkwood, "The statistical mechanical theory of transport processes. iv. the equations of hydrodynamics," The Journal of Chemical Physics **18**, 817–829 (1950).

[102]Culgi B.V., The Netherlands, "The Chemistry Unified Language Interface (CULGI)," `www.culgi.com` (2020), version 13.0.0.

[103]T. P. Hill and J. Miller, "How to combine independent data sets for the same quantity," Chaos: An Interdisciplinary Journal of Nonlinear Science **21**, 033102 (2011).

[104]E. K. Peter, K. Lykov, and I. V. Pivkin, "A polarizable coarse-grained protein model for dissipative particle dynamics," Physical Chemistry Chemical Physics **17**, 24452–24461 (2015).

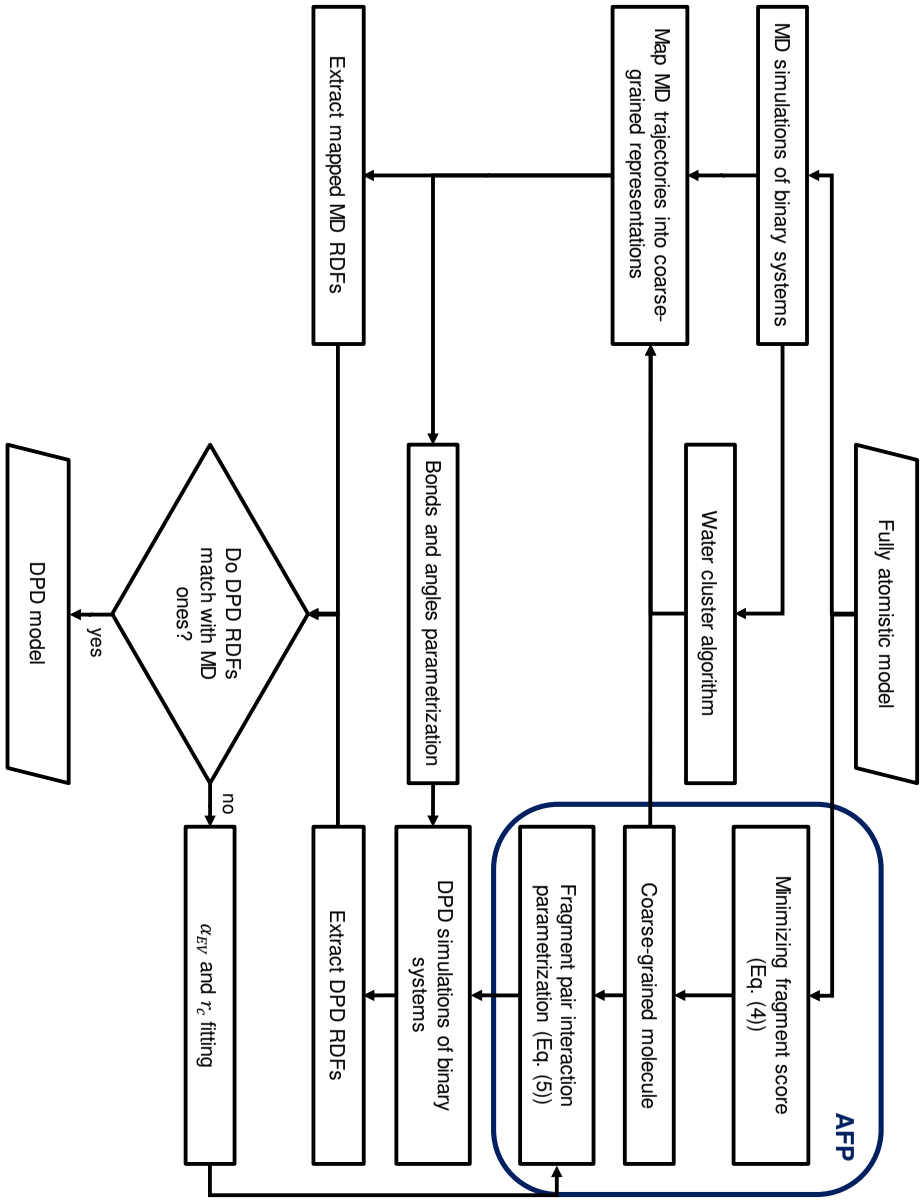
[105]J. Wang and T. Hou, "Application of molecular dynamics simulations in molecular property prediction II: Diffusion coefficient," Journal of Computational Chemistry **32**, 3505–3519 (2011).

[106]Y. Fang and D. G. Dalgleish, "Dimensions of the Adsorbed Layers in Oil-in-Water Emulsions Stabilized by Caseins," Journal of Colloid and Interface Science **156**, 329–334 (1993).

[107]L. Ford, R. Borwankar, R. Martin, and D. Holcomb, "Dressings and sauces," in *Food Emulsions*, edited by S. Friberg and K. Larsson (Marcel Dekker, New York, 1997) 3rd ed., pp. 361–412.

[108]M. Langton, E. Jordansson, A. Altskär, C. Sørensen, and A.-M. Hermansson, "Microstructure and image analysis of mayonnaises," Food Hydrocolloids **13**, 113–125 (1999).

[109]M. Ferrari, J.-W. Handgraaf, G. Boccardo, A. Buffo, M. Vanni, and D. L. Marchisio, "Dataset for "Molecular modeling of the interface of an egg yolk protein-based emulsion"," Zenodo (2021), Dataset, https://doi.org/10.5281/zenodo.5703247.



PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0079883