



HAL
open science

Experimental and Mesoscopic Modeling Study of Water/Crude Oil Interfacial Tension

David Steinmetz, Kevin R. Arriola González, Rafael Lugo, Jan Verstraete,
Véronique Lachet, Aurélie Mouret, Benoit Creton, Carlos Nieto-Draghi

► **To cite this version:**

David Steinmetz, Kevin R. Arriola González, Rafael Lugo, Jan Verstraete, Véronique Lachet, et al..
Experimental and Mesoscopic Modeling Study of Water/Crude Oil Interfacial Tension. *Energy & Fuels*, 2021, 35 (15), pp.11858-11868. hal-03351864

HAL Id: hal-03351864

<https://ifp.hal.science/hal-03351864>

Submitted on 22 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Experimental and Mesoscopic Modeling Study of Water/Crude Oil Interfacial Tension

*David Steinmetz,¹ Kevin R. Arriola González,¹ Rafael Lugo,¹ Jan Verstraete,² Véronique Lachet,¹
Aurélie Mouret,¹ Benoit Creton,¹ Carlos Nieto-Draghi^{1,*}*

¹IFP Energies nouvelles, 1 et 4, avenue de Bois-Préau, 92852 Rueil-Malmaison, France. ²IFP
Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France.

Corresponding Author

* E-mail: carlos.nieto@ifpen.fr

Abstract

Many applications in the Oil & Gas industry require modeling physicochemical properties of complex mixtures. In this work we propose a methodology to predict the interfacial tension of water/crude oils by modeling the composition of crude oil samples using a combined approach of experimental characterization, molecular representation (surrogate) and mesoscopic simulations such as Dissipative Particle Dynamics (DPD). The methodology for molecular representation is based on the experimental analysis by separation of crude oil according to the number of carbon atoms in molecules into two fractions: C_{20-} and C_{20+} . A lumping approach was applied to the C_{20-} fraction and a stochastic reconstruction approach was employed on the C_{20+} fraction. The influence of the different variables (chemical diversity and number of molecular types in the C_{20+} fraction) of the models was analyzed to propose surrogates based on building units with different functional groups. Based on a previous work (JCTC 2018, 14, 4438–4454) a thermodynamically consistent methodology was applied to obtain the DPD interaction parameters of the different chemical building units. DPD simulation on the model crude oil provides predictive values of the interfacial tension that are in good quantitative agreement with the experimental data.

Keywords: Dissipative Particle Dynamics, coarse-grained, interfacial tension, crude oil, parametrization, stochastic reconstruction, lumping method.

1. Introduction

Crude oil is found in Earth's subsurface areas called reservoirs. Its extraction usually takes place into three phases. Primary oil recovery technique relies on the existing natural difference of pressure between the reservoir and surface. The recovery rate is low and generally around 5% OOIP (Original Oil in Place) for heavy oils and can reach 25% OOIP for light oils¹. Secondary recovery allows, on average, to increase the recovery rate to 30% OOIP by maintaining a pressure in the reservoir by injection of water or gas¹. However, a large part of the oil remains stuck in the pores of the rock due to capillary forces². Tertiary recovery methods called Enhanced Oil Recovery (EOR) have been developed in order to increase significantly the quantities of oil that can be extracted from a reservoir. EOR methods aim to affect the properties of mobility and wettability of the oil to facilitate its displacement to production wells.

Several EOR processes exist to increase the mobility of oil by modifying either (i) the temperature of the reservoir by the injection of hot water or steam, (ii) the miscibility of the oil with water by the addition of a phase of hydrocarbons or (iii) the interfacial tension between oil and water by the introduction of an ASP (Alkaline/Surfactant/Polymer) formulation. This latter process is called chemical EOR (cEOR). This approach consists of determining the best ASP combinations to mobilize the oil trapped in the reservoir by acting on capillary forces. To reduce capillary forces, the viscosity of the fluid can be increased and/or the interfacial tension (IFT) between oil and water reduced. The ASP formulation aims at reaching an ultralow IFT for the brine/surfactant/crude oil systems. However, the difficulty of the formulation design stands in the fact that each oil reservoir is unique and they differ from each other by their composition, salinity, pressure, and temperature condition; so a specific S/SP/ASP formulation for a given reservoir is necessary.

Identification and selection of relevant surfactants are challenging due to complex involved phenomena, and it requires a large number of trial-and-error tests. This procedure could in principle be improved if a reliable method to predict the physicochemical properties of crude oils based on their composition is available. This hypothesis is based on the assumption that crude oil properties can be correlated with its composition. Modeling tools such as molecular simulation techniques are suitable for this task since they provide information about phenomena occurring at the molecular level of interfaces.³ Performing molecular simulation requires to establish a molecular representation of the system and select relevant force fields. However, petroleum is a mixture of thousands of molecules. Even now, the precise molecular description of the petroleum with available analytical techniques is a real challenge.

Analytical techniques commonly used in the oil industry provide only global and average information such as elemental analysis, average density or distillation curves of petroleum and its fractions⁴. The most common analytical techniques (bidimensional chromatography, RMN spectroscopy, mass spectrometry, etc.) provide statistical data on the distribution of atoms or groups of atoms in petroleum fractions^{5,6}. The characterization of crude oil, from a molecular point of view, has been the subject of many works since the 1970's⁷⁻⁹. Recent developments on advanced analytical techniques, such as reversed-phase high-temperature comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry (GC × GC-TOFMS)¹⁰ or Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS)¹¹ allow the identification of a large number of chemical species present in complex fluids such as crude oils. The use of detailed compositions for physical property predictions requires two aspects: 1) having access to 3D detailed structural information of the molecules, and 2) an adapted number of chemical species (surrogates) that can be handled by any kind of predictive model or method.

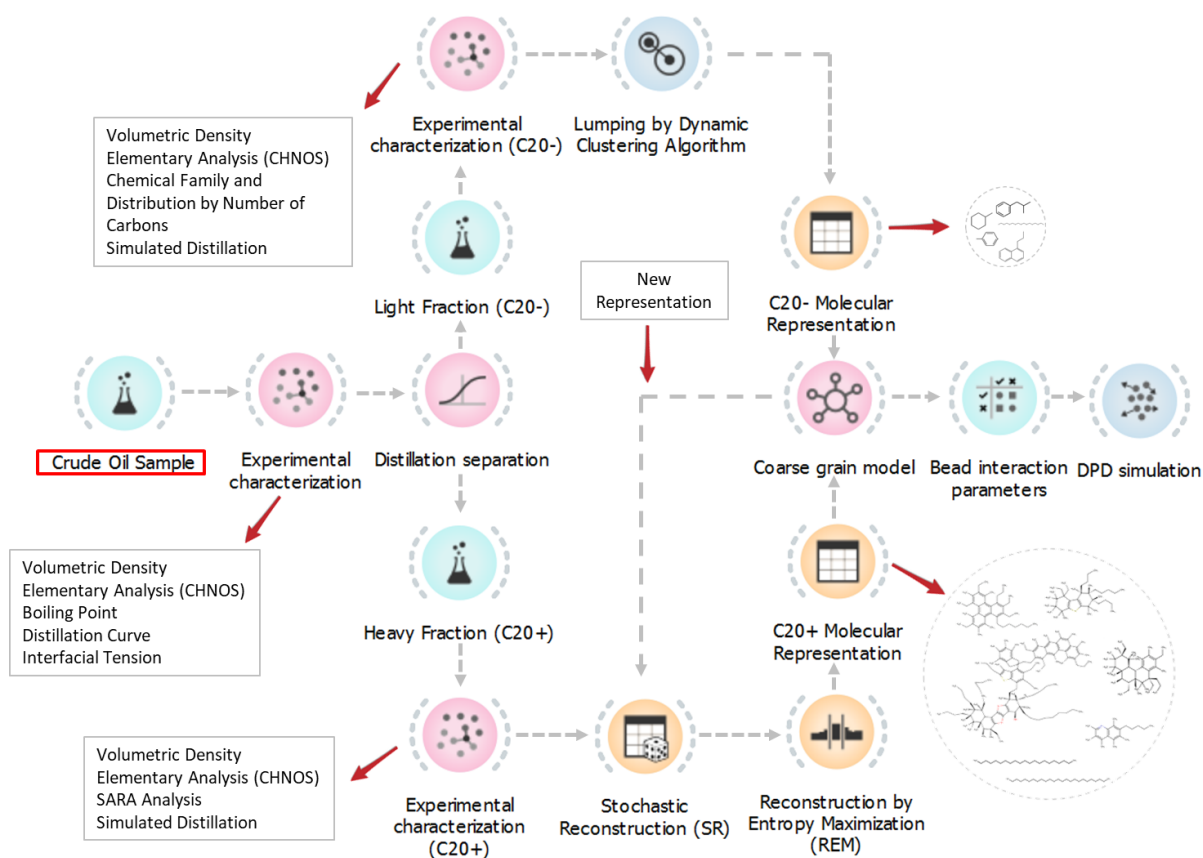
Several works have been devoted in the literature to obtain surrogate models of different petroleum cuts such as gasoline,¹² diesels,¹³ jet fuels¹⁴ and coal liquid fuels,¹⁵ but only few works attempt to propose representative specific molecular structures to model the behavior of real complex fluids. We can mention the work of Ungerer et al.¹⁶ who performed Monte Carlo (MC) simulations to determine the vapor-liquid equilibrium (VLE) with hydrocarbons ranging from C₁₀ to C₃₆, validating with experimental data the saturation pressure and the liquid density. Then, they applied this approach to represent nine crude oils, ranging from light (density < 870 kg/m³) to heavy oils (920 - 1000 kg/m³). Experimental methods as SARA analysis (saturates, aromatics, resins and asphaltenes) and HTGC (High-Temperature Gas Chromatography) were used to describe the molecular structure of each crude oil. Guan et. al¹⁷ proposed a methodology based on the structural unit (SU) and Dissipative Particle Dynamics (DPD) to reproduce the self-assembly behavior of heavy petroleum systems. They performed a SARA analysis to determine the molecular component distributions. Then, they generated molecular structures to feed molecular simulations based on different experimental methods (microscopic atomic force, residue composition analysis).

Molecular simulation techniques can provide information not easily accessible by experimental techniques such as the local composition of mixtures at interfaces. Several works have been devoted to analyze the influence of complex molecules, such as asphaltenes, on the interfacial tension of water/hydrocarbons.¹⁸⁻²¹ However, to the best of our knowledge, there is no information in literature about a quantitative prediction of the water/crude oil interfacial tension (IFT) based on a realistic molecular representation of a crude oil. So, for this work, we propose to combine existing approaches to establish a simplified molecular representation for a crude oil sample. The present article is organized as follows: section 2 is devoted to the description of the methods used

to characterize the crude oil samples and the workflow strategy to obtain molecular structure representation of crude oils followed by the parameterization of the interactions required for DPD simulations. In section 3 we provide the computational details used to obtain the IFT of the different systems by means of DPD simulations. The analysis of the results and comparison between experimental and simulation data is provided in section 4 followed by conclusions in section 5.

2. Workflow strategy and methods

The general strategy developed in our work is presented in Scheme 1. Detailed description of the different methods is presented in the following sub-sections.



Scheme 1. Workflow of the steps involved in the mesoscopic modeling of a crude-oil.

2.1. Experimental characterization of the crude oil: In a first step, one crude oil sample is separated into two fractions according to the number of carbon atoms by distillation: Light fraction with molecules with less than 20 carbon atoms (C_{20-} , using the ASTM D-2892-1 norm) and a heavy fraction (C_{20+} , using the ASTM D-5236-13 norm). The C_{20-} fraction corresponds to compounds whose boiling point is lower than 344 °C while the C_{20+} fraction contains compounds whose boiling point is higher than 344 °C. According to the mass repartition the light fraction of our sample represents 51.6 % (w/w) whereas the heavy fraction represents 48.4% (w/w) of the whole fluid.

The elementary analysis (EA, ASTM D5291) SARA analysis, boiling point (BP), distillation curves (DC), simulated distillation (SD, ASTM D2887) and volumetric mass density (VMD) have been performed on our sample where a detailed description is provided in Supporting Information. The EA and the SARA analysis are reported in Table 1 and Table 2 respectively. As boiling points of oil fractions increase, amounts of the C/H ratio, sulfur, nitrogen and metallic content increase. This experimental characterization is required as an input data for the molecular representation of a crude oil proposed in this work as described below.

Table 1. Elementary analysis (in % w/w according to the ASTM D5291) of the crude oil sample used in this work.

Atom	C₂₀₋	C₂₀₊
C	85.1	84.2
H	13.69	11.39
N	0.09	0.19
O	0	0.25
S	0.53	3.59

Undefined	0.67	0.38
-----------	------	------

Table 2. SARA analysis (derived from the standards ASTM D2007) and AFNOR (NF T01-005 and NF T01-042) of the C₂₀₊ fraction of the crude oil sample used in this work.

Family	% (w/w)
Saturates	29.7
Aromatics	44.4
Resins	21.6
Asphaltenes	1.7
Undefined	2.6

Interfacial tension was measured between the crude oil sample and a reference brine using a commercial device based on the Wilhelmy plate method, at room temperature. Details are provided in the Supporting Information. It is important to note that for crude oil sample, the experimental IFT was measured with a salt concentration (NaCl) of 5 g/L in the aqueous phase. We consider that this value is low enough to assume negligible effects of the salt on measured IFT which is 29.3 ± 0.2 mN/m when the system is at equilibrium (i.e. stable IFT values over time).

2.2. Molecular representation of crude oil fractions: In a recent work, Alvarez-Majmutov et al. proposed a method for generating a computational mixture of representative hydrocarbon molecules that mimics the properties of heavy vacuum gas oil samples.²² In their approach a combined series of experimental characterization with statistical sampling (Monte Carlo) was used to create 3D structures of molecules that can be fitted in a two-dimensional gas chromatography with a flame ionization detector (GC × GC–FID) and a sulfur chemiluminescence detector (GC × GC–SCD), respectively. We follow a similar approach but considering a full crude oil sample.

Both light (C₂₀₋) and heavy (C₂₀₊) fractions were analyzed to measure fluid's physical and chemical properties required for the molecular representation. To represent the light fraction, (C₂₀₋), we used the dynamic clustering algorithm to group components according to their physico-chemical and thermodynamic properties (lumping method).¹² This approach has already been used to represent petroleum cuts (gasoline²³ and diesel²⁴) and it allows to limit the number of representative compounds. To represent the heavy fraction, (C₂₀₊), two methods were sequentially used: the Stochastic Reconstruction (SR), followed by the Reconstruction by Entropy Maximization (REM). These methods were developed to represent heavy oil fractions (from diesel to vacuum residues²⁵⁻³¹). Then, the molecular representations of light and heavy fractions were converted in a coarse-grained (CG) model using different chemical structures as building units (see details in the Methods section).

2.2.1 Light fraction representation: We propose to represent the light fraction of the crude oil with lumping methods; which allow reducing the composition of a fluid to a few representative compounds. The algorithm used in this work is known as “dynamic clustering algorithm”^{24,32,33}. The simulated distillation curve for the C₂₀₋ fraction of our crude oil sample is available in the Supporting Information.

This two-step methodology provides slightly better results than a direct conversion of a simulated distillation curve to a True-Boiling-Point (TBP) curve³⁴. Conversion from the experimental simulated distillation curve into an ASTM D86 norm and then to a TBP curve can be found in the Supporting Information. Once calculated, the TBP distillation curve is divided into 20 equivalent temperature intervals. Pseudo-components derived from the distillation curve are used as initial fluid in the lumping method. The objective of the lumping method is to reduce the number of pseudo-components to only 5, but still preserving the capability of reproducing the

phase behavior of the original real fluid. The last step consists in assigning to each pseudo-component a representative molecule extracted from a gasoline database containing about 250 molecules³⁵ (see details of the complete procedure in the Supporting Information).

2.2.2 Heavy fraction representation: Lumping methods are effective for representing hydrocarbon mixtures with low molecular weight and low heteroatoms content; however, the presence of these heteroatoms cannot be ignored for the case of resins and asphaltenes due to their polarity and the possibility of forming hydrogen bonds. So, for the heavy fraction (C₂₀₊), the SR and the REM methods are needed. These methods are briefly explained hereafter.

Stochastic Reconstruction (SR): SR is a method for building molecules to form a mixture whose properties are in agreement with experimental data. It is based on the assumption that a petroleum fraction can be entirely characterized by a set of probability distribution functions of molecular structural attributes. Any molecule in the petroleum feedstock can be considered to be an assembly of molecular attributes (for example the type of a molecule, the number of aromatic rings in a molecule, the number of aliphatic chains in a molecule, etc.). The occurrence frequency of an attribute is given by a probability distribution function. By arranging basic molecular building unit structures next to each other, it becomes possible to form complex molecules. Molecular attributes must be chosen on the basis of the chemical characteristics of the petroleum fraction to be represented. Based on the work of Schnongs *et al.*³¹, de Oliveira *et al.*^{29,36} proposed to describe the fraction of vacuum residues with a total of 16 structural attributes. The list of attributes, their possible values, their type of distribution function and the number of parameters for the distribution functions used for the genetic algorithm are presented in section 3.2 of the Supporting Information.

The parameters of the genetic algorithm have been fixed at those used by de Oliveira *et al.*²⁹ for the heavy fraction. The result of the SR is a list of 10 to 20 molecules that should represent the heavy fraction of crude oils (C₂₀₊), though their proportion (molar fraction) is not necessarily optimal. In fact, additional information is required to improve the molar fraction of the heavy fraction. This optimization is performed by the REM method described below.

Reconstruction by Entropy Maximization (REM): The REM method was developed at IFP Energies nouvelles by Hudebine and Verstraete^{26,30} to improve the concordance between the properties of a mixture with the analytical data. The molar fractions of a library of molecules are adjusted based on the information entropy criterion. Originally formulated in the context of Shannon's information theory³⁷, this entropic criterion can be adapted for the study of oil composition. Additional technical details concerning the application of the REM method are provided in the Supporting Information.

2.3 Dissipative Particle Dynamics (DPD): The size and complexity of molecular structures in a crude oil (particularly the heavy fraction), and the time and length scales involved in the interfacial phenomena (e.g., IFT) makes impossible the use of molecular simulation at atomic level. Therefore, we decided to consider DPD technique developed by Hoogerbrugge and Koelman^{38,39} to handle the crude-oil/water system. DPD simulations consist of pairwise interaction between “beads” (particles representing groups of atoms or molecules). The total force \mathbf{f}_i exerted on a bead i by another bead j is defined as the sum of the conservative (\mathbf{F}_{ij}^C), dissipative (\mathbf{F}_{ij}^D), random (\mathbf{F}_{ij}^R) forces and intramolecular force (\mathbf{F}_{ij}^{intra}) as shown in Eqs. (1) to (10):

$$\mathbf{f}_i = \sum_{j \neq i} (\mathbf{F}_{ij}^C + \mathbf{F}_{ij}^D + \mathbf{F}_{ij}^R + \mathbf{F}_{ij}^{intra}) \quad (1)$$

$$\mathbf{F}_{ij}^C = a_{ij} \left(1 - \frac{r_{ij}}{r_c}\right) \hat{\mathbf{r}}_{ij} \quad (2)$$

$$\mathbf{F}_{ij}^D = -\gamma_{ij} \omega_D(r_{ij}) (\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij}) \hat{\mathbf{r}}_{ij} \quad (3)$$

$$\mathbf{F}_{ij}^R = \sigma_{ij} \omega_R(r_{ij}) \theta_{ij} \hat{\mathbf{r}}_{ij} \quad (4)$$

$$\mathbf{F}_{ij}^{\text{intra}} = -K(r_{ij} - r_0) \hat{\mathbf{r}}_{ij} \quad (5)$$

where a_{ij} is the interaction parameter representing the maximum repulsive magnitude between beads i and j , r_c is the cutoff radius, $\hat{\mathbf{r}}_{ij}$ is the position unit vector defined by beads i and j . $\omega_D(r_{ij})$ is the dissipative weight function, γ_{ij} the friction coefficient, \mathbf{v}_{ij} is the relative velocity between bead i and bead j , $\omega_R(r_{ij})$ is the random weight function, σ_{ij} is the random force amplitude for the bead i and bead j , θ_{ij} is a random number with zero mean and unity variance when averaged over time, K is the spring constant and r_0 is the equilibrium spring distance.

According to Español and Warren⁴⁰, the dissipative forces (\mathbf{F}_{ij}^D) and the random forces (\mathbf{F}_{ij}^R) must be coupled with a fluctuation-dissipation relation in order to define the temperature of the system. They derived the Fokker-Plank equation for the DPD model proposed by Hoogerbrugge and Koelman, relating the friction coefficient γ_{ij} between particles i and j with the random force amplitude σ in Eq. (6) and the weight functions $\omega^D(r_{ij})$ and $\omega^R(r_{ij})$ in Eq. (7).

$$\sigma^2 = 2\gamma_{ij} k_B T, \quad (6)$$

$$\omega^D(r_{ij}) = [\omega^R(r_{ij})]^2 = \begin{cases} \left(1 - \frac{r_{ij}}{r_c}\right), & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases} \quad (7)$$

where k_B is the Boltzmann constant and T is the temperature. According to these assumptions, the system is ensured to maintain the temperature equilibrium and follow the constant number of particles, volume and temperature (canonical NVT) ensemble.

To perform DPD simulations, the molecular representations of light fraction (C_{20-}) and heavy fraction (C_{20+}) are converted into a coarse-grained model. In this work, the coarse-grained model is built according to three criteria:

- a) Represent as closest as possible the same molecular volume for all DPD beads used to model the IFT^{41,42}.
- b) Using a coarse-grained level (N_m) that better represents all the functional groups used to represent the molecules.
- c) Limit the CG representation to only a few types of beads, so the number of interactions that must be parameterized remains limited.

In this work, we have selected a set of 12 different chemical structures (or functional groups) as building units or DPD beads for the mesoscopic model as can be seen in Figure 1. In order to preserve the chemical diversity, six building blocks are used to represent the different functional groups containing heteroatoms: thiol, thiophene, pyridine, ethylamine, phenol and furan. The parameterization of interactions between DPD beads were carried out in order to reproduce crude oil/water equilibrium. The methodology proposed here was based on a previous work⁴³, where water/hydrocarbon/solute ternary systems were parameterized. Beads used in the coarse-grained representation of the crude oil were classified into three categories: water beads that represent the aqueous phase, hydrocarbon beads that represent the organic phase and compounds with a heteroatom that represent chemical functions present in some crude oil molecules.

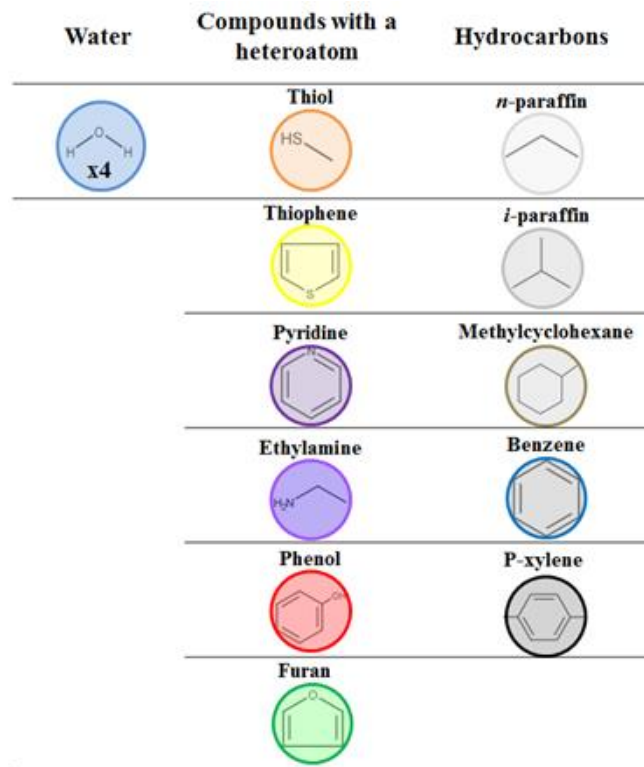


Figure 1. List of beads used to represent the crude oils. Water beads contain four molecules of water ($N_m = 4$). For simplicity, sulfide functional groups (R-S-R') are considered similar to thiol (R-SH) functional groups.

Interaction parameters between DPD beads were estimated using different approaches: (i) Hildebrand solubility parameters or (ii) thermodynamic model SRK-MHV2-UNIFAC calculations (accordingly to the type of bead interactions). This methodology has been applied and validated for water/hydrocarbon/solute ternary systems⁴³. The parameters for a crude oil/water system were obtained as follows:

- Interactions between **like beads**. The corresponding interaction parameters are calculated from the isothermal compressibility of water as proposed by Groot and Warren⁴⁴(here, $a_{ii} = a_{jj} = 25$).

b) **Functional group/water** and **functional group/hydrocarbon interactions**. The interaction between molecules with heteroatoms and water have been obtained by means of the Flory-Huggins parameter (χ) using the following relation,

$$\chi_{ij} = \frac{[\ln(x_i \cdot \gamma_{ij}) - \ln(1 - \phi_j)]}{\phi_j^2} \quad (8)$$

where x_i is the molar fraction of i and ϕ_j is the volume fraction of j , γ_{ij} is the activity coefficient of a solute molecule i in a solvent j . We propose to use a thermodynamic model to predict the values of γ_{ij} . We have tested a certain number of thermodynamic models available in the literature (see Supporting Information for details). Their quality has been evaluated by direct comparison of the model predictions with reference experimental data on liquid-liquid equilibria for several ternary systems containing two immiscible solvents (water and a hydrocarbon) and a partially miscible organic molecule containing an oxygen heteroatom (forming a hydrogen bond with water) used in our previous work.⁴³ After our analysis, we have selected the SRK-MHV2-UNIFAC⁴⁵ thermodynamic model to estimate the functional groups/water activity coefficient (details of the models and the evaluation tests are provided in the Supporting Information). Then the interaction parameter can be estimated as suggested by Groot and Warren⁴⁴ as $a_{ij}=3.5*\chi_{ij}+25$.

c) **Functional group/functional group** and **water/hydrocarbon** interactions. These interaction parameters are calculated using Hildebrand solubility parameters (δ). We assume that the overall concentration of heteroatoms is low, and therefore, the number of functional group/functional group interactions is too small to influence the result of the simulations and can be fairly well estimated by means of Hildebrand solubility parameter. The χ parameter can be obtained as,

$$\chi_{ij} = \frac{v_b}{k_B T} (\delta_i - \delta_j)^2 \quad (9)$$

where v_b is the mean volume of a bead. Hildebrand solubility parameters are extracted from the DIPPR⁴⁶ and are given in the Supporting Information.

The complete matrix of interaction parameters for the crude oil/water system is presented in Table 3. Some of the water/functional group interactions presented in Table 3 reflect the presence of hydrogen bonds with negative values of the χ parameter (e.g., those between water/thiol and water/ethylamine with $\chi_{W/Thiol} = -1.63$ and $\chi_{W/Ethylamine} = -3.46$, respectively) and with repulsive parameters which are less than 25 (i.e., less than the water-water interaction).

Table 3. Interaction parameters* calculated for the crude oil/water system (in DPD unit) for the different bead types defined in Figure 1.

Bead type	Symbol	W	n-par	i-par	Benz	P-xyl	Cycl	Fur	Phe	Thp	Thl	Pyr	Ethl	Acid
Water	W	25.0	136.0	152.1	111.6	116.3	128.0	113.0	32.3	30.6	19.3	26.4	12.9	25.2
n-paraffin	n-par	136.0	25.0	25.5	26.5	26.0	25.1	26.3	38.5	29.3	29.5	33.5	30.7	32.9
i-paraffin	i-par	152.1	25.5	25.0	28.9	28.0	26.3	28.6	37.7	28.3	27.5	33.0	29.3	30.8
Benzene	Benz	111.6	26.5	28.9	25.0	25.1	25.7	25.0	29.8	25.2	25.5	25.6	26.4	26.2
p-xylene	P-xyl	116.3	26.0	28.0	25.1	25.0	25.3	25.0	30.8	24.8	24.7	26.7	27.5	26.7
Methylcyclohexane	Cycl	128.0	25.1	26.3	25.7	25.3	25.0	25.6	36.3	25.9	27.6	30.3	29.8	31.2
Furan	Fur	113.0	26.3	28.6	25.0	25.0	25.6	25.0	29.2	25.3	25.1	26.0	25.0	25.3
Phenol	Phe	32.3	38.5	37.7	29.8	30.8	36.3	29.2	25.0	27.3	28.1	26.1	28.4	27.3
Thiophene	Thp	30.6	29.3	28.3	25.2	24.8	25.9	25.3	27.3	25.0	25.1	25.2	25.1	25.0
Thiol	Thl	19.3	29.5	27.5	25.5	24.7	27.6	25.1	28.1	25.1	25.0	25.5	25.0	25.1
Pyridine	Pyr	26.4	33.5	33.0	25.6	26.7	30.3	26.0	26.1	25.2	25.5	25.0	25.6	25.2
Ethylamine	Ethl	12.9	30.7	29.3	26.4	27.5	29.8	25.0	28.4	25.1	25.0	25.6	25.0	25.1
Acetaldehyde	Acid	25.2	32.9	30.8	26.2	26.7	31.2	25.3	27.3	25.0	25.1	25.2	25.1	25.0

* Parameters obtained from the isothermal compressibility of water , parameters calculated using a thermodynamic SRK-MHV2-UNIFAC model , and parameters calculated using the Hildebrand solubility parameters .

3. Computational details

All simulations of crude oil/water systems were performed in the native NVT ensemble of DPD using the NEWTON Molecular Dynamics simulation package.⁴⁷ Initial simulation boxes

containing about 162000 beads were built using the PACKMOL software package^{48,49}. All initial configurations have been constructed in such a way to have equal composition at each side of the interfaces. The detailed molecular composition of the simulations boxes is given in the Supporting Information. Boxes dimensions were set to $L_z = 60$, $L_x = L_y = 30$ (in DPD units). Two planar interfaces were created normal to the z-axis. In all simulations, periodic boundary conditions were imposed in all directions.

For DPD simulations, reduced units were used. The mass for all beads is fixed to unity. The bead volume, v_b , is defined as $v_b = N_m \times (\text{volume of a water molecule})$, where N_m is the degree of coarse-graining. For this work $N_m = 4$, the bead volume is equal to $v_b = 120 \text{ \AA}^3$. The cut-off radius is given from the volume of the beads with $r_c = (\bar{\rho} \times v_b)^{1/3}$. With an overall DPD number density $\bar{\rho}$ set to 3, the cut-off radius is equal to $r_c = 7.11 \text{ \AA}$. For all flexible molecules the spring constant $K = 100.0$ and equilibrium distance r_0 is $0.7 * r_c$ (both values in DPD units). These values are commonly used in the literature for bond forces.⁵⁰⁻⁵² The dissipative force γ and random force σ were set to 4.5 and 3, respectively, to keep the temperature fixed at $k_B T = 1$; thus satisfying the fluctuation-dissipation relation.

A modified version of the velocity-Verlet algorithm⁴⁴ was applied in this work, and the time step was fixed at $\delta t = 0.01$ in DPD units. Simulations have been carried out using 1.5×10^6 steps of equilibration and 3×10^5 steps of production. DPD simulations of each representation of a crude oil/water system were repeated four times using completely independent initial configurations (position and initial velocities of the molecules were randomly generated). The interfacial tension (IFT) for each configuration was calculated in a similar manner as in our previous work,⁴³ according to the local method proposed by Irving and Kirkwood.⁵³

Another relevant issue when performing simulations with explicit interfaces is the eventual migration of molecules from bulks to interfaces. This is the case of molecules containing heteroatoms behaving as surface-active species. This migration of molecules will modify the composition of the bulk phase. We recently showed that a way to solve this problem would be to carry out an equilibration step in the osmotic ensemble (μ NPT) using Monte Carlo (MC) simulations.⁴³ This ensemble would correct the composition of the bulk phase to compensate the variation of the composition mentioned before. The chemical potential of each specie of the oil could be estimated by Widom test insertion of a separate independent bulk phase (without the explicit interface). However, the main problem with this approach would be the efficiency of the MC method to insert large complex molecules, which would make this simulation impractical. Alternatively, we correct the composition of the bulk phases by using an iterative approach where we survey the composition of each molecule in the oil phase after 1×10^5 steps (phase composition is estimated by integrals of the local density profiles). If a significant deviation from the initial composition of all the species is detected ($>2\%$), we corrected the number of molecules of the concerned specie (respecting as much as possible the symmetry for the number of molecules in the bulk phases), and we restart the simulation (always verifying the global mass balance). This procedure allows converging to the desired bulk initial composition in only few iterations (between three and four).

4. Results and discussion

We generated one representation for the light fraction and eight different representations for the heavy fraction. The choice of one representation for the light fraction can be justified because the C₂₀- fraction is expected to have a lower diversity in terms of molecular complexity as compared to the heavy fraction, and also by the experience acquired in the past using lumping method to

represent fuel gasolines²³ and diesels²⁴ which can be assumed similar to the C₂₀-. The representation of the C₂₀₊ fraction is more complicated since the molecular diversity in terms of chemical nature and size is much larger. Heavy compounds have a strong impact on transport or interfacial properties, and the representation of the heavy fraction should be proceeded with care. Thus, representations of the C₂₀₊ fraction were generated by means of SR and REM methods. We have investigated the influence of the number of molecular types to represent the heavy fraction, and the initial SR step is used to produce mixtures containing maximum 10 molecules (representations R1 to R4) or maximum 20 molecules (representations R5 to R8). Finally, molar fractions of molecules in the heavy fraction are optimized by the REM to better fit the experimental constraints (or even completely remove some molecules provided by the SR step). We can then produce different mixtures of molecules (with different structures and number of building blocks) still having the same overall mean physical properties. Each crude oil is represented by eight mixtures (named R1 to R8) containing the same light fraction solely varying the heavy fraction. The objective is to compare the interfacial tension of these mixtures to evaluate the sensitivity of the proposed approach. Additionally, information about the influence of the number of compounds considered for the representation of the heavy fraction on the final value of the ITF should be derived.

The 5 molecules generated by the lumping method to represent the C₂₀- fraction are presented in Table 4. The set of molecules obtained for the C₂₀- fraction of crude oil sample contains no heteroatom in agreement with elementary analysis provided in Table 1. This representation includes hydrocarbon families such as n-paraffin, naphthene, mono- and poly-aromatics.

The R1 (representation 1) for the heavy fraction of the crude oil sample is proposed in Table 5, whereas other representations (R2 to R8) are given in the Supporting Information. As expected the

heavy fraction contains heteroatoms which are embedded in molecules that can be classified according to the SARA analysis (see Table 2 for details). Since our crude oil contains only 1.7% of asphaltenes, such molecules are not systematically present in the representation of the C₂₀₊.

Table 4. Chemical structures of the representative molecules of the light fraction (C₂₀₋) of the crude oil sample.

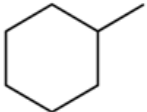


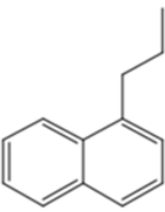
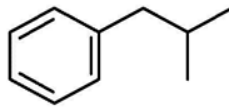
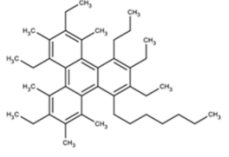
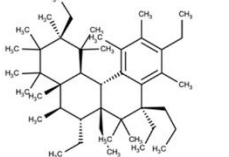
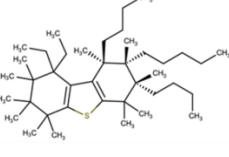
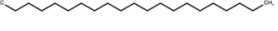
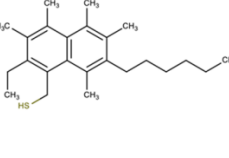
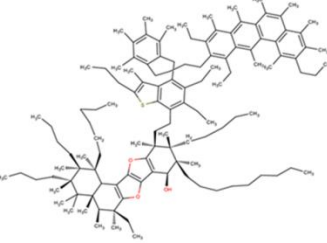
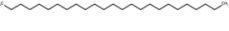
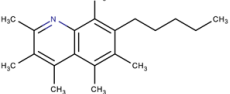
Light fraction (C ₂₀₋)	
Methylcyclohexane	<i>n</i> -heptadecane
	
p-Xylene	1-propyl-naphthalene
	
Isobutylbenzene	
	

Table 5. Chemical structures of the representative molecules (R1) of the heavy fraction (C₂₀₊) for crude oil sample. For reasons of readability, an identification (ID) number is given to each molecule and the chemical family according to the SARA separation.

Molecular representation	(ID) / Chemical family	Molecular representation	(ID) / Chemical family

	(1) / Aromatic		(5) / Aromatic
	(2) / Aromatic		(6) / Saturate
	(3) / Aromatic		(7) / Resin
	(4) / Saturate		(8) / Resin

For our crude oil, the IFT values predicted by the DPD simulations using different representations are shown in Table 6. R1 to R4 initially contain 10 molecules whereas R5 to R8 initially contains 20 molecular types to represent the C_{20+} fraction (obtained with the SR method), this number of molecular types has been reduced after the molar fraction refinement performed by the REM method. We obtained similar IFT values from a statistical point of view using both sets of representations (with standard deviations of 0.88% for R1 to R4 and 2.3% for R5 to R8). We can mention, however, that we obtained higher values of the IFT using representations having 13 or more molecules than the ones using 7 or 8 molecules. In view of the uncertainties obtained, we can conclude that the representation of 10 molecules generated by the SR method is enough to model the heavy fraction of the crude oil.

Table 6. Average values of IFT obtained from DPD simulations for the crude oil sample using 5 molecules to model the light fraction and different numbers of molecules to model the heavy

fraction. R1 to R4 contains 10 molecules obtained by the SR, whereas systems R5 to R8 contain 20 molecules. Size of C_{20+} gives the final number of molecular types in the C_{20+} fraction obtained after the REM method.

Representation/system	Size of C_{20+}	$\overline{IFT}_{\text{averaged}}$ (DPD units)	IFT_{averaged} (mN/m)	IFT_{overall} (mN/m)	$IFT_{\text{experimental}}$ (mN/m)
R1. Crude oil/water	8	3.71 ± 0.01	30.20 ± 0.08		
R2. Crude oil/water	8	3.91 ± 0.03	31.86 ± 0.24		
R3. Crude oil/water	9	3.76 ± 0.013	30.59 ± 0.11		
* C_{20-} /water	-	4.71 ± 0.021	38.32 ± 0.17	30.62 ± 0.88	
* C_{20+} /water	9	3.28 ± 0.013	26.69 ± 0.11		29.3 ± 0.2
R4. Crude oil/water	7	3.66 ± 0.03	29.84 ± 0.24		
R5. Crude oil/water	13	4.02 ± 0.03	32.76 ± 0.24		
R6. Crude oil/water	15	4.3 ± 0.02	35.03 ± 0.16		
R7. Crude oil/water	14	3.62 ± 0.03	29.48 ± 0.24	32.69 ± 2.34	
R8. Crude oil/water	13	4.11 ± 0.05	33.49 ± 0.41		

$\overline{IFT}_{\text{averaged}}$ corresponds to the mean value of each crude oil representation computed using four simulation runs with different initial conditions (conformations and velocities). A factor of 8.143 [mN/m] is used for the conversion between $\overline{IFT}_{\text{averaged}}$ (in DPD units) and IFT_{averaged} (in mN/m) according to the relation $\overline{IFT} = IFT \left(\frac{r_c^2}{k_B T} \right)$. IFT_{overall} (in mN/m) is the mean value of the four IFT_{averaged} obtained for each representation of the crude oil, the error is the standard deviation of the mean values R1, R2, R3 and R4. C_{20-} /water and C_{20+} /water are the IFT obtained using the light and heavy fraction composition respectively in contact with water. *Additional simulations where IFT are obtained using either the light (C_{20-}) or heavy (C_{20+}) fractions of R3, these two tests do not contribute to the IFT_{overall}

If we focus on R1 to R4, we can observe that the relative error deviation between the experimental IFT and the average IFT value obtained by our DPD simulations is only 4.5%. We can consider this a remarkable result in view of the complexity in terms of composition of a crude oil. To better understand the role played by light and heavy fractions on the IFT of the complete crude oil, we report in Table 6 the IFT obtained separately between the C_{20-} and water and the one obtained by

the C_{20+} and water (both fractions used in the R3). Our results show a higher value of the IFT ($\sim 43\%$) obtained by the light fraction than the heavy fraction. At the origin of this difference is the presence of heteroatoms in the molecules of the heavy fraction that act as surface-active molecules (this point will be discussed later on the analysis of the composition profiles). The light fraction represents 51.6% w/w ($w_{C_{20-}}$) of the crude oil (see section 2.1). Though there is no reason to expect that the IFT of a mixture of hydrocarbons follows a linear mixing of its constituents, we perform a rough estimation of the crude oil IFT by using a mass fraction weighting average (estimated as $IFT_{c-o} = w_{C_{20-}} \cdot IFT_{C_{20-}} + (1 - w_{C_{20-}}) \cdot IFT_{C_{20+}}$) of the IFT of the light and heavy fractions obtained for the R3. We obtained a value of 32.7 mN/m which is higher than the value obtained in simulations with the complete crude-oil/water system (30.59 mN/m). The reason for the difference is probably the migration of surface-active molecules modifying the IFT of the complete system.

The density profiles computed in the cases of R1 and R5 of the crude oil/water systems are shown in Figure 2. For reasons of readability, only the density of the molecules of the heavy fraction and water are displayed. Molecules are indicated by their ID number, given in Table 5. It is important to mention that the interface zone (considering the two interfaces) spans between 70 Å to 140 Å (~ 10 to 20 DPD unit length), which represents between 16% to 32% of the total simulation box size in the direction perpendicular to the interface (60 DPD units or 427 Å). The size of the interface with respect to the simulation box, in addition to the statistical treatment of the different samples, clearly justify the use of a mesoscopic approach employed in this work (in comparison with an atomistic one).

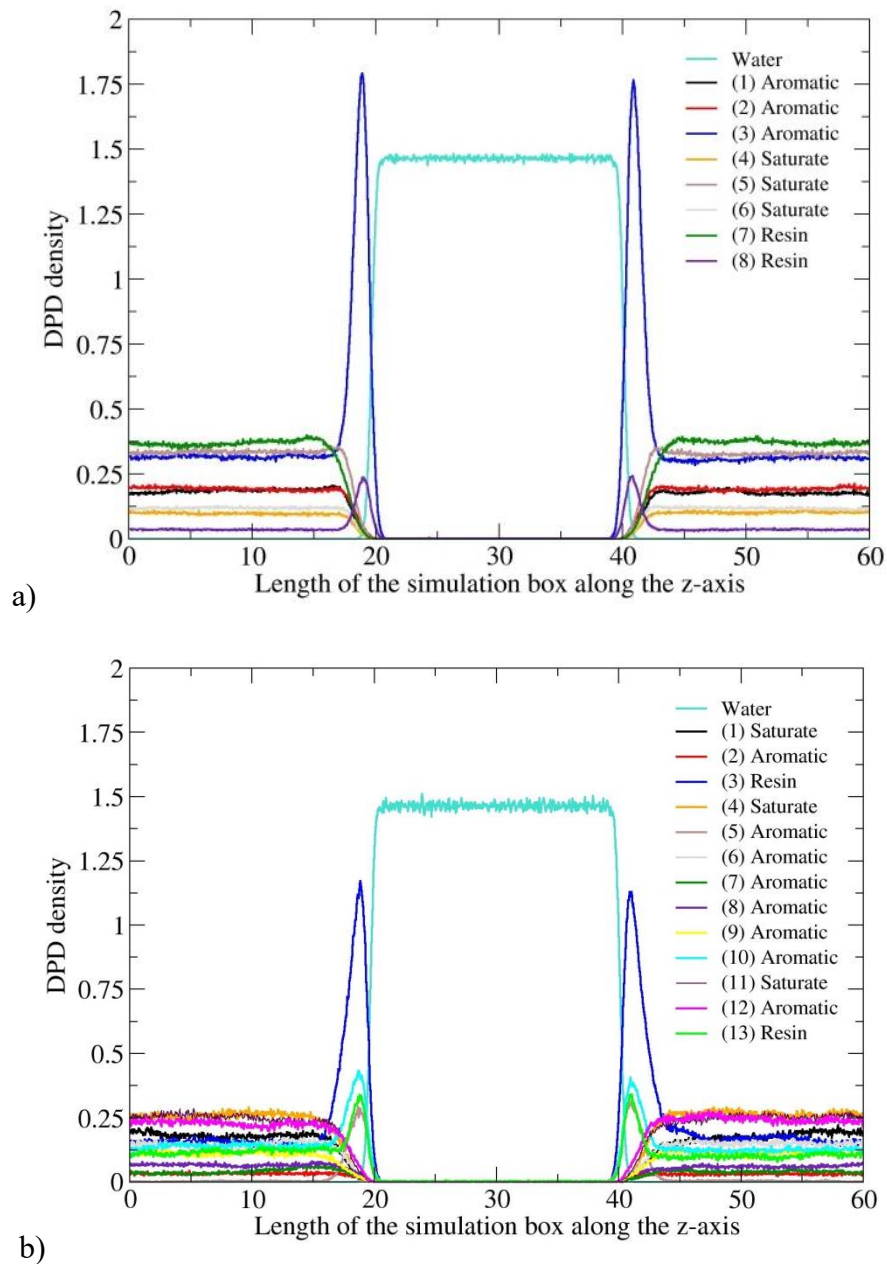


Figure 2. Density profile of the different species of the crude oil/water system using a) 8 and b) 13 molecular types to represent the heavy fraction (systems R1 and R5 respectively). Molecules of the light fraction are omitted for simplicity. For reasons of readability, molecules are indicated by their ID number provided in Table 5 and by the chemical family according to the SARA composition analysis (details on R5 composition are given in the Supporting Information). Note

that the DPD density of water has been divided by 2 on diagrams. The length of the simulation box is in DPD units.

Two different systems are compared in Figure 2, using 8 (Figure 2.a, using R1) or 13 (Figure 2.b, using R5) molecules to represent the heavy fraction of our crude oil sample. We first analyze Figure 2.a, where one clearly see that two molecules of the heavy fraction have migrated to the interface: the aromatic molecule (ID 3 in Table 5) with a thiol functional group in its structure (10.5% mol/mol of the crude oil) and the resin (ID 8 in Table 5) with a pyridine group (1.85% mol/mol of the crude oil). The migration of these molecules towards the interface can be explained by a lower functional group/water interaction parameter (i.e. $\chi_{w/Thiol}=-1.63$ and $\chi_{w/Pyridine}=0.4$) than functional groups/hydrocarbons interaction parameters (see Table 3). In other words, the presence of functional groups such as the thiol or pyridine in molecules of the heavy fraction confers them a “natural surfactant” character then contributing in diminishing the crude oil/water IFT value. Not all the molecules having a heteroatom behave like surface-active molecules, for instance the aromatic molecule (ID 2 in Table 5), having a thiophene group (and $\chi_{w/thiophene}=1.6$), and the resin molecule (ID 7 in Table 5), having two furan groups and a phenol group ($\chi_{w/phenol}=2.0$ and $\chi_{w/furan}=25$), do not migrate to the interface. There are probably two reasons for this, firstly the fact that the interaction parameters of these functional groups with water are not as favorable as the ones of thiols and pyridine, and secondly these functional groups are sterically hindered by other (less favorable) functional groups in their respective molecules (see details in Table 5).

The analysis of Figure 2.b (for R5) shows a similar behavior for the density profiles as the one observed in Figure 2.a, where molecules having heteroatoms tend to migrate to the interface (see Supporting Information for details on composition of this system). In this case four molecules having functional groups are surface active: two resins with thiol and thiophene groups (ID 3 with

1.73% mol/mol and ID 13 with 1.32% mol/mol), an aromatic with furan and thiol groups (ID 10 and 2.49% mol/mol) and an aromatic with sulfide and thiol groups (ID 5 and 0.26% mol/mol).

The information provided by the density profiles allows to qualitatively identify surface-active molecules. To provide a quantitative analysis of the impact of surface-active molecules on the IFT values, we can evaluate their affinity to the interface (I) by means of the Gibbs free energy of adsorption (ΔG_{ads}). ΔG_{ads} can be estimated using the density profiles obtained in our DPD simulations through the bulk-interface partition coefficients. Partition coefficients are generally used to represent the distribution of a solute between two bulk phases and can be expressed with a distribution constant, K_D . For our systems, the distribution constant of the solute between the aqueous and organic phase is expressed with $K_D^{aq \rightarrow org}$. Considering the interface as a separate phase in our simulations, two other distribution constants can be established: a distribution constant between the aqueous phase and the interface, $K_D^{aq \rightarrow I}$, and a distribution constant between the organic phase and the interface, $K_D^{org \rightarrow I}$. Assuming ideal mixtures, these distribution constants are expressed using solute local density as shown in the following equations.

$$\frac{\rho_{solute}^{org.}}{\rho_{solute}^{aq.}} = K_D^{aq \rightarrow org.} = \exp\left(\frac{-\Delta G_{transfer}^{aq \rightarrow org.}}{RT}\right), \quad (10)$$


$$\frac{\rho_{solute}^I}{\rho_{solute}^{aq.}} = K_D^{aq \rightarrow I} = \exp\left(\frac{-\Delta G_{ads}^{aq \rightarrow I}}{RT}\right), \quad (11)$$

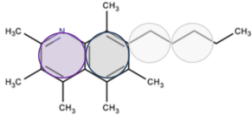
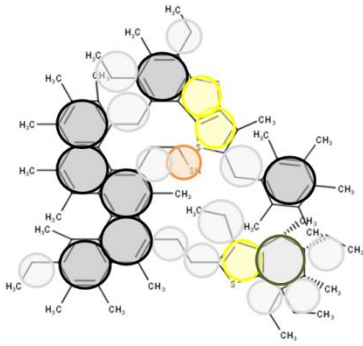
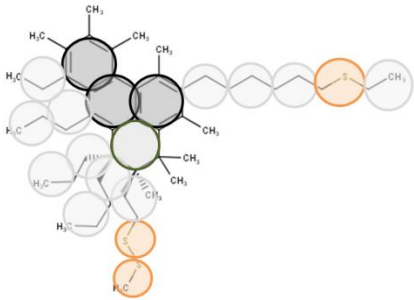
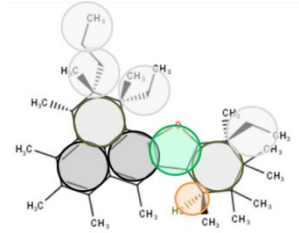
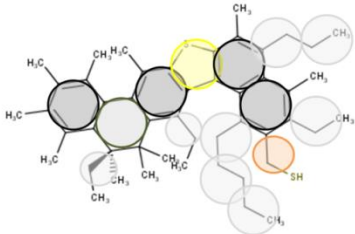
$$\frac{\rho_{solute}^I}{\rho_{solute}^{org.}} = K_D^{org \rightarrow I} = \exp\left(\frac{-\Delta G_{ads}^{org \rightarrow I}}{RT}\right). \quad (11)$$

The solute density at the interface, ρ_{solute}^I , and the solute density at the bulk, $\rho_{solute}^{org.}$ for the different surface-active molecules are directly obtained from the density profiles shown in Figure 2 as the values of the picks at the water/crude-oil interface and the values of the plateau at the bulk

respectively. The Gibbs free energy of transfer of solute molecules from the aqueous phase to the organic phase, $\Delta G_{transfer}^{aq.\rightarrow org.}$, corresponds to the energy required for one mole of solute to cross the interface from the aqueous phase to the organic phase. In other words, the Gibbs free energy of transfer is the addition of the Gibbs free energy of adsorption of solute at the interface from the aqueous phase, $\Delta G_{ads}^{aq.\rightarrow I}$, with the Gibbs free energy of desorption of solute from the interface to the organic phase, $\Delta G_{ads}^{org.\rightarrow I}$. Since $\rho_{solute}^{aq.}$ is zero in our simulations, the only accessible quantity in our simulations is $\Delta G_{ads}^{org.\rightarrow I}$ which has been obtained through equation 12 and the results are presented in Table 7 for the surface-active molecules of R1 and R5 systems (the list of molecules for all systems studied is provided in the Supporting Information).

Table 7. Surface active molecules obtained in our simulations for the crude oil sample using different representations. Hydrophilic/Lipophilic balance of the DPD model molecule (HLB). Bulk organic phase to interface (*I*) distribution constant $K_D^{org.\rightarrow I}$. Gibbs free energy of adsorption of surface-active molecules from the bulk organic phase to the interface $\Delta G_{ads}^{org.\rightarrow I}$. The molecular structure contains the coarse graining scheme using the functional groups with color code presented in Figure 1.

ID (Number of molecular types in C ₂₀₊)	Structure	HLB ^[b]	$K_D^{org.\rightarrow I}$	$\Delta G_{ads}^{org.\rightarrow I}$ [kJ/mol]
3 (8)		3.3	4.8	-3.9

8 (8)		5.0	8.0	-5.2
3 (13)		3.3	7.4	-5.0
5 (13)		3.2	2.6	-2.4
10 (13)		4.0	3.3	-2.9
13 (13)		2.7	2.8	-2.6

^bThe HLB for the DPD representation of the molecule have been estimated as $HLB=20 \cdot n_H/n_{tot}$ where n_H and n_{tot} are the number of beads bearing an heteroatom and the total number of beads of the molecule respectively.

It is interesting to see the variety in topology for the surface-active molecules identified for systems R1 and R5. R1 contains molecules much smaller than the ones observed for system R5,

but their affinity to the interface represented by the $\Delta G_{ads}^{org \rightarrow I}$ are quite similar with values of the same order of magnitude ranging from -2.4 kJ/mol to -5.2 kJ/mol. We also introduce a crude approximation of the hydrophilic/lipophilic balance (HLB) of the DPD molecules listed in Table 7, counting the ratio of beads bearing a heteroatom with respect to the total number of beads. Molecules with higher HLB values (more hydrophilic) tend to have the lower Gibbs free energy of adsorption at the interface.

To end our analysis we compare the variation of the crude oil/water IFT with the overall interfacial concentration (C_{surf}) of surface-active molecules for the different systems studied in Figure 3. As expected, we observe a tendency where the IFT tends to decrease with the increase of C_{surf} . This figure reveals the impact of the concentration of surface-active molecules (having heteroatoms) in the heavy fraction on the behavior of the crude oil/water IFT.

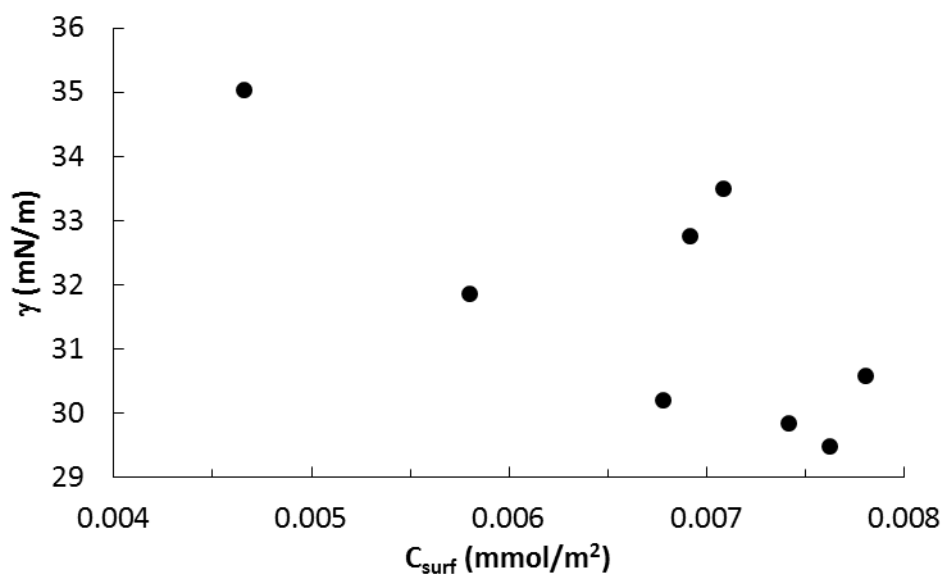


Figure 3. Variation of the interfacial tension obtained by the different systems studied in this work (R1 to R8) in function of the surface concentration of “active” molecules present at the interface.

5. Conclusions

A workflow for the representation of a crude oil composition with a limited number of molecular types has been developed based on experimental data. The proposed strategy consists of the separation and analysis of a crude oil in two main cuts: a light fraction, C_{20-} (with molecules having boiling temperatures less than 617.15 K) and a heavy fraction, C_{20+} (with boiling temperatures greater than 617.15 K), by means of a distillation approach. Specific battery of analysis has been performed to each cut to characterize their physicochemical properties.

Using this information, a lumping method was applied to the light fraction C_{20-} using the ReFGen tool. From a distillation curve and a molecular database of species typically observed in gasoline, a representation with only 5 constitutive molecules has been established to represent the light fraction. The heavy fraction C_{20+} was modeled by using the Stochastic Reconstruction (SR) method with a subsequent refinement of the molar composition by the Entropy Maximization Method (REM) to produce a set of molecules in agreement with the experimental data. The heavy fraction of the crude oil has been modeled by several systems having different numbers of molecules (between 7 and 15) and compositions. Molecules representing the crude oil (including the light and heavy fractions) have been coarse-grained and the crude oil modeled with Dissipative Particle Dynamics simulations in order to predict the crude oil/water interfacial tension (IFT). Our simulation predictions show a good agreement on the IFT when compared to the experimental reference data with a relative deviation of less than 5%. To the best of our knowledge, this successful quantitative comparison between simulations and experiment on crude oil/water IFT has never been published before. The explanation of this achievement is twofold: 1) a consistent methodology to provide a reliable representation of molecules of a crude oil with special attention on the representation of complex surface-active molecules containing heteroatoms and 2) a

thermodynamically consistent parameterization strategy of the fluid/fluid interactions applicable for DPD simulations.

The strategy presented in this work can be used to model quantitatively crude oil/water IFT for EOR and water treatment applications. This can be useful for a better understanding of the complex interactions of the natural surface-active molecules present in crude oils with additives such as surfactants, ions and polymers. Work is in progress to thoroughly test this methodology with other crude oil samples with a large diversity on composition, including asphaltene molecules and the specific consideration of acid/basic equilibrium reactions to properly describe the interfacial composition of such complex fluids.

ASSOCIATED CONTENT

The Supporting Information is available free of charge on the ACS Publication website at DOI: 10.1021/acs.jctc.XXXX

Further information is available on the experimental characterization of the crude oil sample in addition to the light (C20-) and heavy (C20+) fractions, the lumping procedure to represent the light fraction, the stochastic reconstruction and the entropy maximization to represent the heavy fraction, the choice of the thermodynamic model to parameterize the beads interaction, the detailed composition and model molecules of the different representation of the heavy fraction of the crude oil and additional simulation results (PDF).

ACTUAL ADDRESS

K. R. Arriola González actual position: (kr.arriolagonzalez@ugto.mx) Division de Ciencias e Ingenierías, Campus Leon, Universidad de Guanajuato, Loma del Bosque 103, Lomas del Campestre 37150 León, Guanajuato, México.

DEDICATION

The present work is dedicated to the memory of Dr Rafael Lugo.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGEMENTS

K. R. Arriola González would like to thank CONACyT for scholarship grant number 468141

REFERENCES

- (1) Thomas, S. Enhanced Oil Recovery - An Overview. *Oil & Gas Science and Technology - Rev. IFP*, 2008, 9–19.
- (2) Avendano, J. Viscoélasticité et récupération améliorée du pétrole, Université Paris-Est, Ph.D thesis, 2012.
- (3) Creton, B.; Nieto-Draghi, C.; Pannacci, N. Prediction of Surfactants' Properties using Multiscale Molecular Modeling Tools:A Review. *Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles* **2012**, 67 (6), 969–982. DOI: 10.2516/ogst/2012040.
- (4) Riazi, M. R. *Characterization and properties of petroleum fractions*, 1st ed.; ASTM International, 2005.
- (5) Hudebine, D. Reconstruction moléculaire de coupes pétrolières, Ecole Normale Supérieure de Lyon, Ph.D thesis, 2003.
- (6) Chacón-Patiño, M. L.; Smith, D. F.; Hendrickson, C. L.; Marshall, A. G.; Rodgers, R. P. Advances in Asphaltene Petroleomics. Part 4. Compositional Trends of Solubility Subfractions Reveal that Polyfunctional Oxygen-Containing Compounds Drive Asphaltene Chemistry. *Energy Fuels* **2020**, 34 (3), 3013–3030. DOI: 10.1021/acs.energyfuels.9b04288.
- (7) Hirsch, E.; Altgelt, K. H. Integrated structural analysis. Method for the determination of average structural parameters of petroleum heavy ends. *Anal. Chem.* **1970**, 42 (12), 1330–1339. DOI: 10.1021/ac60294a005.
- (8) Speight, J. A structural investigation of the constituents of Athabasca bitumen by proton magnetic resonance spectroscopy. *Fuel* **1970**, 49 (1), 76–90. DOI: 10.1016/0016-2361(70)90010-4.
- (9) Boduszynski, M. M. Composition of heavy petroleums. 2. Molecular characterization. *Energy Fuels* **1988**, 2 (5), 597–613. DOI: 10.1021/ef00011a001.
- (10) Jennerwein, M. K.; Eschner, M. S.; Wilharm, T.; Zimmermann, R.; Gröger, T. M. Proof of Concept of High-Temperature Comprehensive Two-Dimensional Gas Chromatography Time-of-

Flight Mass Spectrometry for Two-Dimensional Simulated Distillation of Crude Oils. *Energy Fuels* **2017**, *31* (11), 11651–11659. DOI: 10.1021/acs.energyfuels.7b01799.

(11) Palacio Lozano, D. C.; Gavard, R.; Arenas-Diaz, J. P.; Thomas, M. J.; Stranz, D. D.; Mejía-Ospino, E.; Guzman, A.; Spencer, S. E. F.; Rossell, D.; Barrow, M. P. Pushing the analytical limits: new insights into complex mixtures using mass spectra segments of constant ultrahigh resolving power. *Chem. Sci.* **2019**, *10* (29), 6966–6978. DOI: 10.1039/c9sc02903f. Published Online: Jul. 5, 2019.

(12) Montel, F.; Gouel, P. L. A New Lumping Scheme of Analytical Data for Compositional Studies. In *SPE Annual Technical Conference and Exhibition*, 1984. DOI: 10.2118/13119-MS.

(13) Mueller, C. J.; Cannella, W. J.; Bruno, T. J.; Bunting, B.; Dettman, H. D.; Franz, J. A.; Huber, M. L.; Natarajan, M.; Pitz, W. J.; Ratcliff, M. A.; Wright, K. Methodology for Formulating Diesel Surrogate Fuels with Accurate Compositional, Ignition-Quality, and Volatility Characteristics. *Energy Fuels* **2012**, *26* (6), 3284–3303. DOI: 10.1021/ef300303e.

(14) Huber, M. L.; Smith, B. L.; Ott, L. S.; Bruno, T. J. Surrogate Mixture Model for the Thermophysical Properties of Synthetic Aviation Fuel S-8: Explicit Application of the Advanced Distillation Curve. *Energy Fuels* **2008**, *22* (2), 1104–1114. DOI: 10.1021/ef700562c.

(15) Huber, M. L.; Lemmon, E. W.; Diky, V.; Smith, B. L.; Bruno, T. J. Chemically Authentic Surrogate Mixture Model for the Thermophysical Properties of a Coal-Derived Liquid Fuel. *Energy Fuels* **2008**, *22* (5), 3249–3257. DOI: 10.1021/ef800314b.

(16) Ungerer, P.; Yiannourakou, M.; Mavromaras, A.; Collell, J. Compositional Modeling of Crude Oils Using C 10 –C 36 Properties Generated by Molecular Simulation. *Energy Fuels* **2019**, *33* (4), 2967–2980. DOI: 10.1021/acs.energyfuels.8b04403.

(17) Guan, D.; Feng, S.; Zhang, L.; Shi, Q.; Zhao, S.; Xu, C. Mesoscale Simulation for Heavy Petroleum System Using Structural Unit and Dissipative Particle Dynamics (SU–DPD) Frameworks. *Energy Fuels* **2019**, *33* (2), 1049–1060. DOI: 10.1021/acs.energyfuels.8b04082.

(18) Rezaei, H.; Amjad-Iranagh, S.; Modarress, H. Self-Accumulation of Uncharged Polyaromatic Surfactants at Crude Oil–Water Interface: A Mesoscopic DPD Study. *Energy Fuels* **2016**, *30* (8), 6626–6639. DOI: 10.1021/acs.energyfuels.6b00254.

(19) Ervik, Å.; Lysgaard, M. O.; Herdes, C.; Jiménez-Serratos, G.; Müller, E. A.; Munkejord, S. T.; Müller, B. A multiscale method for simulating fluid interfaces covered with large molecules such as asphaltenes. *J. Comput. Phys.* **2016**, *327*, 576–611. DOI: 10.1016/j.jcp.2016.09.039.

(20) Ruiz-Morales, Y.; Mullins, O. C. Coarse-Grained Molecular Simulations to Investigate Asphaltenes at the Oil–Water Interface. *Energy Fuels* **2015**, *29* (3), 1597–1609. DOI: 10.1021/ef502766v.

(21) Gao, F.; Xu, Z.; Liu, G.; Yuan, S. Molecular Dynamics Simulation: The Behavior of Asphaltene in Crude Oil and at the Oil/Water Interface. *Energy Fuels* **2014**, *28* (12), 7368–7376. DOI: 10.1021/ef5020428.

(22) Alvarez-Majmutov, A.; Gieleciak, R.; Chen, J. Deriving the Molecular Composition of Vacuum Distillates by Integrating Statistical Modeling and Detailed Hydrocarbon Characterization. *Energy Fuels* **2015**, *29* (12), 7931–7940. DOI: 10.1021/acs.energyfuels.5b02082.

(23) Nieto-Draghi, C.; Bocahut, A.; Creton, B.; Have, P.; Ghoufi, A.; Wender, A.; Boutin, A.; Rousseau, B.; Normand, L. Optimisation of the dynamical behaviour of the anisotropic united atom model of branched alkanes: application to the molecular simulation of fuel gasoline. *Mol. Sim.* **2008**, *34* (2), 211–230. DOI: 10.1080/08927020801993370.

- (24) Aquing, M.; Ciotta, F.; Creton, B.; Féjean, C.; Pina, A.; Dartiguelongue, C.; Trusler, J. P. Martin; Vignais, R.; Lugo, R.; Ungerer, P.; Nieto-Draghi, C. Composition Analysis and Viscosity Prediction of Complex Fuel Mixtures Using a Molecular-Based Approach. *Energy Fuels* **2012**, *26* (4), 2220–2230. DOI: 10.1021/ef300106z.
- (25) de Oliveira, Luís Pereira; Verstraete, J. J.; Kolb, M. Simulating vacuum residue hydroconversion by means of Monte-Carlo techniques. *Catal. Today* **2014**, *220-222*, 208–220. DOI: 10.1016/j.cattod.2013.08.011.
- (26) Hudebine, D.; Verstraete, J. J. Reconstruction of Petroleum Feedstocks by Entropy Maximization. Application to FCC Gasolines. *Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles* **2011**, *66* (3), 437–460. DOI: 10.2516/ogst/2011110.
- (27) Van Geem, Kevin M.; Hudebine, D.; Reyniers, M. F.; Wahl, F.; Verstraete, J. J.; Marin, G. B. Molecular reconstruction of naphtha steam cracking feedstocks based on commercial indices. *Comput. Chem. Eng.* **2007**, *31* (9), 1020–1034. DOI: 10.1016/j.compchemeng.2006.09.001.
- (28) Hudebine, D.; Verstraete, J.; Chapus, T. Statistical Reconstruction of Gas Oil Cuts. *Oil Gas Sci. Technol. – Rev. IFP Energies nouvelles* **2011**, *66* (3), 461–477. DOI: 10.2516/ogst/2009047.
- (29) de Oliveira, L. Pereira; Vazquez, A. T.; Verstraete, J. J.; Kolb, M. Molecular Reconstruction of Petroleum Fractions: Application to Vacuum Residues from Different Origins. *Energy Fuels* **2013**, *27* (7), 3622–3641. DOI: 10.1021/ef300768u.
- (30) Hudebine, D.; Verstraete, J. J. Molecular reconstruction of LCO gasoils from overall petroleum analyses. *Chem. Eng. Sci.* **2004**, *59* (22-23), 4755–4763. DOI: 10.1016/j.ces.2004.09.019.
- (31) Verstraete, J. J.; Schnongs, P.; Dulot, H.; Hudebine, D. Molecular reconstruction of heavy petroleum residue fractions. *Chem. Eng. Sci.* **2010**, *65* (1), 304–312. DOI: 10.1016/j.ces.2009.08.033.
- (32) Montel, F.; Gouel, P. L. A New Lumping Scheme of Analytical Data for Compositional Studies. *59th annual Technical Conference and Exhibition-Houston*, 1984.
- (33) Lugo, R.; Ebrahimian, V.; Lefebvre, C.; Habchi, C.; Hemptinne, J.-C. de. A Compositional Representative Fuel Model for Biofuels - Application to Diesel Engine Modelling. *SAE 2010 Powertrains Fuels & Lubricants Meeting*, 2010. DOI: 10.4271/2010-01-2183.
- (34) Daubert, T. E.; Danner, R. P. *API Technical Data Book - Petroleum Refining*, 6th ed.; American Petroleum Institute (API), 1997.
- (35) Durand, J. P.; Fafet, A.; Barreau, A. Direct and automatic capillary GC analysis for molecular weight determination and distribution in crude oils and condensates up to C20. *J. High Resol. Chromatogr.* **1989**, *12* (4), 203. DOI: 10.1002/jhrc.1240120408.
- (36) Pereira de Oliveira, Luís; Verstraete, J. J.; Kolb, M. A Monte Carlo modeling methodology for the simulation of hydrotreating processes. *Chem. Eng. J.* **2012**, *207-208*, 94–102. DOI: 10.1016/j.cej.2012.05.039.
- (37) Shannon C.E. A Mathematical Theory of Communication: The Bell System Technical. *The Bell Syst. Tech. J.* **1948**, *27* (1), 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.
- (38) Hoogerbrugge, P. J.; Koelman, J. M. V. A. Simulating Microscopic Hydrodynamic Phenomena with Dissipative Particle Dynamics. *Europhys. Lett.* **1992**, *19* (3), 155–160. DOI: 10.1209/0295-5075/19/3/001.
- (39) Koelman, J. M. V. A.; Hoogerbrugge, P. J. Dynamic Simulations of Hard-Sphere Suspensions Under Steady Shear. *Europhys. Lett.* **1993**, *21* (3), 363–368. DOI: 10.1209/0295-5075/21/3/018.
- (40) Español, P.; Warren, P. Statistical Mechanics of Dissipative Particle Dynamics. *Europhys. Lett.* **1995**, *30* (4), 191–196. DOI: 10.1209/0295-5075/30/4/001.

- (41) Goel, H.; Chandran, P. R.; Mitra, K.; Majumdar, S.; Ray, P. Estimation of interfacial tension for immiscible and partially miscible liquid systems by Dissipative Particle Dynamics. *Chem. Phys. Lett.* **2014**, *600*, 62–67. DOI: 10.1016/j.cplett.2014.03.061.
- (42) Rezaei, H.; Modarress, H. Dissipative particle dynamics (DPD) study of hydrocarbon–water interfacial tension (IFT). *Chem. Phys. Lett.* **2015**, *620*, 114–122. DOI: 10.1016/j.cplett.2014.12.033.
- (43) Steinmetz, D.; Creton, B.; Lachet, V.; Rousseau, B.; Nieto-Draghi, C. Simulations of Interfacial Tension of Liquid-Liquid Ternary Mixtures Using Optimized Parametrization for Coarse-Grained Models. *J. Chem. Theory Comput.* **2018**, *14* (8), 4438–4454. DOI: 10.1021/acs.jctc.8b00357.
- (44) Groot, R. D.; Warren, P. B. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* **1997**, *107* (11), 4423. DOI: 10.1063/1.474784.
- (45) Dahl, S.; Michelsen, M. L. High-pressure vapor-liquid equilibrium with a UNIFAC-based equation of state. *AIChE J.* **1990**, *36* (12), 1829–1836. DOI: 10.1002/aic.690361207.
- (46) Rowley, R. L.; Wilding, W. V.; Oscarson, J. L.; Zundel, N. A. *DIPPR 801*; Design Institute for Physical Property Research/AIChE, 2014.
- (47) Nguyen, T. V.-O.; Houriez, C.; Rousseau, B. Viscosity of the 1-ethyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide ionic liquid from equilibrium and nonequilibrium molecular dynamics. *Phys. Chem. Chem. Phys.* **2010**, *12* (4), 930–936. DOI: 10.1039/b918191a.
- (48) Martínez, J. M.; Martínez, L. Packing optimization for automated generation of complex system's initial configurations for molecular dynamics and docking. *J. Comput. Chem.* **2003**, *24* (7), 819–825. DOI: 10.1002/jcc.10216.
- (49) Martínez, L.; Andrade, R.; Birgin, E. G.; Martínez, J. M. PACKMOL: a package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* **2009**, *30* (13), 2157–2164. DOI: 10.1002/jcc.21224.
- (50) Rekvig, L.; Kranenburg, M.; Vreede, J.; Hafskjold, B.; Smit, B. Investigation of Surfactant Efficiency Using Dissipative Particle Dynamics. *Langmuir* **2003**, *19* (20), 8195–8205. DOI: 10.1021/la0346346.
- (51) Rekvig, L.; Hafskjold, B.; Smit, B. Simulating the effect of surfactant structure on bending moduli of monolayers. *J. Chem. Phys.* **2004**, *120* (10), 4897–4905. DOI: 10.1063/1.1645509.
- (52) Rekvig, L.; Hafskjold, B.; Smit, B. Molecular simulations of surface forces and film rupture in oil/water/surfactant systems. *Langmuir* **2004**, *20* (26), 11583–11593. DOI: 10.1021/la048071p.
- (53) Irving, J. H.; Kirkwood, J. G. The Statistical Mechanical Theory of Transport Processes. IV. The Equations of Hydrodynamics. *J. Chem. Phys.* **1950**, *18* (6), 817–829. DOI: 10.1063/1.1747782.
-

TOC Graphic

