

Supplementary information of:

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

*D. Steinmetz¹, K. R. Arriola González¹, Rafael Lugo¹, Jan J. Verstraete², Véronique Lachet¹,
Aurélie Mouret¹, Benoît Creton¹, Carlos Nieto-Draghi¹*

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

INDEX

1. EXPERIMENTAL DATA OF CRUDE OIL AND ITS FRACTIONS	5
1.1 EXPERIMENTAL METHODS.....	5
1.1.1 ELEMENTAL ANALYSIS	5
1.1.2 BOILING TEMPERATURE AND DISTILLATION CURVES	5
1.1.3 VOLUMETRIC MASS DENSITY	8
1.1.4 INTERFACIAL TENSION (IFT)	8
1.2 EXPERIMENTAL DATA.....	9
1.2.1 EXPERIMENTAL DATA OF CRUDE OIL A1	9
1.2.2 EXPERIMENTAL DATA OF THE LIGHT FRACTION FOR CRUDE A1 ($T < 344^{\circ}\text{C}$)	10
1.2.3 EXPERIMENTAL DATA OF THE HEAVY FRACTION FOR CRUDE A1 ($T > 344^{\circ}\text{C}$)	11
2. THERMODYNAMIC MODELS SELECTION FOR PREDICTION OF LIQUID-LIQUID EQUILIBRIA AND PARAMETRIZATION OF DPD INTERACTIONS.	13
2.1 PREDICTION OF THE COMPOSITION OF LIQUID-LIQUID EQUILIBRIA	13
2.1.1 PREDICTION OF DPD INTERACTION PARAMETERS USING A THERMODYNAMIC MODEL.....	18
2.1.2 TRANSFERABILITY OF INTERACTION PARAMETERS FOR ALKANES.....	23

3.	<u>CALCULATION OF THERMODYNAMIC PROPERTIES FOR CRUDE OIL</u>	26
3.1	LUMPING METHOD	26
3.2	MOLECULAR REPRESENTATION	28
3.3	DPD SIMULATIONS OF CRUDE OIL	33
4.	<u>REFERENCES</u>	68

LIST OF TABLES

Table 1.	Elemental analysis and density for crude oil A1.	9
Table 2.	Elemental analysis and density for the light fraction of crude oil A1.	10
Table 3.	Chemical family of the light fraction of crude oil A1.	10
Table 4.	Weight fraction of the light fraction of crude oil A1.	10
Table 5.	Simulated distillation for the light fraction of crude oil A1.	11
Table 6.	Elemental analysis and density for the heavy fraction of crude oil A1.	11
Table 7.	SARA analysis of the heavy fraction of crude oil A1.	11
Table 8.	Simulated distillation for the heavy fraction of crude oil A1.	12
Table 9.	Comparison of thermodynamic models for prediction of liquid-liquid equilibrium for the water/benzene/1,4-dioxane system at 298.15K.	15
Table 10.	Interaction parameters calculated for binary systems using the SRK-MHV2-UNIFAC thermodynamic model for water/solute and organic compound/solute binary systems. (Compositions: 1 mol% of solute and 99 mol% of solvent). Parameters are compared with those obtained in CG simulations (using approach #2 of ref. ³) except for the water/benzene/1,4-dioxane system.	22
Table 11.	Interaction parameters calculated for <i>n</i> -alkane/solute binary systems using the thermodynamic model SRK-MHV2-UNIFAC. (Composition used is 1 mol% in solute and 99 mol% in hydrocarbon).	24
Table 12.	Conversion of the simulated distillation curve to an ASTM D86 curve, and then to a TBP curve for crude oil A1.	26
Table 13.	Calculation of molecule and pseudo-component properties.	27
Table 14.	Methods for the properties of molecules and mixtures used in the SR approach.	29
Table 15.	Definition of the structural attributes used in the SR. Extracted from the work of de Oliveira <i>et al.</i> ⁴⁴	30
Table 16.	Parameters of the genetic algorithm for the heavy fraction reconstruction. Extracted from the work of de Oliveira <i>et al.</i> ⁴⁴	31
Table 17.	Hildebrand solubility parameters extracted from the DIPPR ⁴⁶ .	33
Table 18.	Composition of the simulation box for the crude oil/water system A1 (Representation 1).	34
Table 19.	Composition of the simulation box for the crude oil/water system A1 (Representation 2).	34
Table 20.	Composition of the simulation box for the crude oil/water system A1 (Representation 3).	36

Table 21. Composition of the simulation box for the crude oil/water system A1 (Representation 4).....	36
Table 22. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 5).....	37
Table 23. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 6).....	38
Table 24. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 7).....	39
Table 25. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 8).....	39
Table 26. Chemical structures and coarse-grained models of Representation 1 (R1) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.	41
Table 27. Chemical structures and coarse-grained models of Representation 2 (R2) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.	43
Table 28. Chemical structures and coarse-grained models of Representation 3 (R3) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.	45
Table 29. Chemical structures and coarse-grained models of Representation 4 (R4) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.	47
Table 30. Chemical structures and coarse-grained models of Representation 5 (R5) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.	49
Table 31. Chemical structures and coarse-grained models of Representation 6 (R6) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.	53
Table 32. Chemical structures and coarse-grained models of Representation 7 (R7) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.	56
Table 33. Chemical structures and coarse-grained models of Representation 8 (R8) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.	60

LIST OF FIGURES

Figure 1. Principle of the IFT measurement using the Wilhelmy plate method.	9
Figure 2. Prediction of liquid-liquid equilibrium ternary diagrams for the water/benzene/1,4-dioxane system at 298.15K using the SRK-MHV2-UNIFAC model. Experimental data are plotted in black (solid lines for the conodal lines) and the results from the thermodynamic model are in red (dashed lines for the conodal lines).....	18
Figure 3. (a) Chemical potential of the 1,4-dioxane in the aqueous phase and (b) in the organic phase for the water/benzene/1,4-dioxane system.	20

Figure 4. (a) Interaction parameters of the 1,4-dioxane with water for the water/benzene/1,4-dioxane system. (b) Interaction parameters of the 1,4-dioxane with benzene for the water/benzene/1,4-dioxane system.....	20
Figure 5. Representation of interactions between a solute and linear alkanes such as <i>n</i> -hexane, <i>n</i> -nonane and <i>n</i> -dodecane.....	24
Figure 6. Comparison between the initial fluid (20 pseudo-components) and the final fluid resulting from the Lumped method (5 pseudo-components). The calculated phase envelopes and the calculated distillation curves are shown for crude oil A1.....	28
Figure 7. Density profile of crude oil/water A1 system (Representation 1). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	64
Figure 8. Density profile of crude oil/water A1 system (Representation 2). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	64
Figure 9. Density profile of crude oil/water A1 system (Representation 3). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	65
Figure 10. Density profile of crude oil/water A1 system (Representation 4). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	65
Figure 11. Density profile of crude oil/water A1 system with 20 molecules (Representation 5). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	66
Figure 12. Density profile of crude oil/water A1 system with 20 molecules (Representation 6). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	66
Figure 13. Density profile of crude oil/water A1 system with 20 molecules (Representation 7). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	67
Figure 14. Density profile of crude oil/water A1 system with 20 molecules (Representation 8). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.	67

1. Experimental data of crude oil and its fractions

1.1 Experimental methods

1.1.1 Elemental analysis

Elemental analysis provides percentages of carbon (C), hydrogen (H), heteroatoms (S, N, O) contained in crude oil or in oil fractions. The elemental analysis were performed with a Microvario Cube elemental analyzer (Elementar Analysensysteme), using a methodology adapted from the ASTM 5291 standard for CHNS analysis and by infrared detection for oxygen analysis. From the elemental analysis, the atomic ratios of the various elements to carbon (i.e. H/C, N/C, O/C, and S/C) are determined and provide indications of the overall character of the feedstock. According to Speight ¹, proportions of the elements in petroleum vary only slightly over narrow limits.

1.1.2 Boiling temperature and distillation curves

Distillation techniques consist in separating the different constituents of a mixture according to their boiling point. The boiling point may be presented by a curve of boiling temperature versus volume fraction (vol %) or mass fraction (wt %). The boiling point of the lightest component in a mixture is called Initial Boiling Point (IBP) and the boiling point of the heaviest compound is called the Final Boiling Point (FBP). For crude oil, FBP is above 550°C, however, it can happen that some heavier compounds cannot vaporize. Therefore, the FBP value measured is not accurate and does not correspond to the boiling of heaviest compound present in the mixture. There are three main types of distillation curves: the distillation D86 (ASTM D86), the True

Boiling Point (TBP) (ASTM D2892) and the Simulated Distillation (SD) curve (ASTM D2887).

In this work, we followed the SD method, which will be explained below.

Simulated Distillation by Gas Chromatography

A distillation curve produced by a Gas chromatography (GC) is called Simulated Distillation (SD) and the method is described in ASTM D2887 test method. Simulated Distillation method is known to be simple, consistent, and reproducible. This method is applicable to petroleum fractions with a FBP up to 538°C (even 700°C) and a boiling range of greater than 55°C. Samples are analyzed on a non-polar chromatographic column that separates hydrocarbons according to their boiling points. Each component has a certain retention time depending on the structure of compound, type of column and stationary phase, flow rate of mobile phase, length, and temperature of column. The retention time is the amount of time required for a given component to cross the column. More volatile compounds with lower boiling points have lower retention times. Distillation curves by SD are presented in terms of boiling point versus wt% of mixture vaporized. SD curves are very close to actual boiling points shown by TBP curves. But these two types of distillation data are not identical and conversion methods should be used to convert SD to TBP curves.

SARA analysis

SARA (Saturates, Aromatics, Resins, Asphaltenes) analysis is based on a solvent separation approach; it means that components in crude oil are divided according to their solubility in a particular solvent. Solubility of compounds in solvents depends on their polarity. When two compounds have the same polarity (polar-polar or apolar-apolar), they tend to be miscible, while

compounds with different polarities are generally insoluble. Asphaltenes, which are the most polar compounds in crude oil, are not soluble in apolar paraffinic solvent.

In this work, SARA content was measured following modified ASTM 2007 procedure as described below. Results are expressed in wt %.

Distillation to separate heavy fraction and light fraction of the crude oil

For the light cut C20-: preparatory distillation according to the ASTM D-2892-1 standard allowing to obtain the fraction with a boiling point below 344°C.

For the heavy cut C20+: a vacuum distillation is added according to the ASTM D-5236-13 standard allowing to obtain the fraction with a boiling point above 344°C.

After these distillation steps, the weight percentages of the C20- and C20+ cuts are obtained.

Carburane for the light fraction < 344°C

A gas chromatography is performed on the distillation cut corresponding to boiling point below 344°C. The sample is eluted with helium (vector gas) and components are detected with a flame-ionization detector. The weight fraction of saturated and aromatics on the C20- cut is deduced.

SARA on the heavy fraction > 344°C

The distillation cut with a boiling point above 344°C is analyzed in order to separate the four SARA oil fractions according to their solubility. The method is derived from the standards ASTM (ASTM 2007) and AFNOR (NF T01-005 and NF T01-042).

The first step is the asphaltenes separation by precipitation with n-heptane, filtration and quantification of the asphaltenes by comparing the filter weight before and after.

The filtrate collected at the end of the filtration is then diluted in n-heptane and injected in a chromatography column with alumina and activated silica.

The SAR family separation is made successively from the less polar one to the most polar:

- S fraction eluted by the column by n-heptane

- A fraction eluted by a mixture of n-heptane/toluene
- R fraction eluted by a mixture of dichloromethane/toluene/methanol

The SAR contents for the heavy fraction (C20+) are then determined by weight after solvent evaporation.

1.1.3 Volumetric mass density

The volumetric mass density measurement on crude oil A1, as well as on its light and heavy fractions, was carried out with a Metrohm DMA 4500M oscillating tube density meter.

1.1.4 Interfacial tension (IFT)

Interfacial tension was measured between the crude oil A1 and a reference brine (NaCl – 5g/L) using a Krüss K100 tensiometer based on the Wilhelmy plate method, at room temperature². In this method, a vertically suspended thin plate exerts a force F at the crude oil / aqueous phase interface. This force is measured using a microbalance and allows to calculate the interfacial tension (γ) in mN/m using the Wilhelmy equation:

$$\gamma = \frac{F}{L \cos \theta} \quad (1)$$

where L (in m) is the wetted length of the plate (equal to its perimeter when its thickness is negligible), F (in mN) the force measured by a microbalance and θ ($^{\circ}$) the contact angle between the liquid and the plate. The plate is in platinum because it is chemically inert and can also be optimally wetted. Thus, the contact angle θ can be approximated to 0° with most liquids ($\cos\theta = 1$) Figure 1 shows the principle of the measurement.

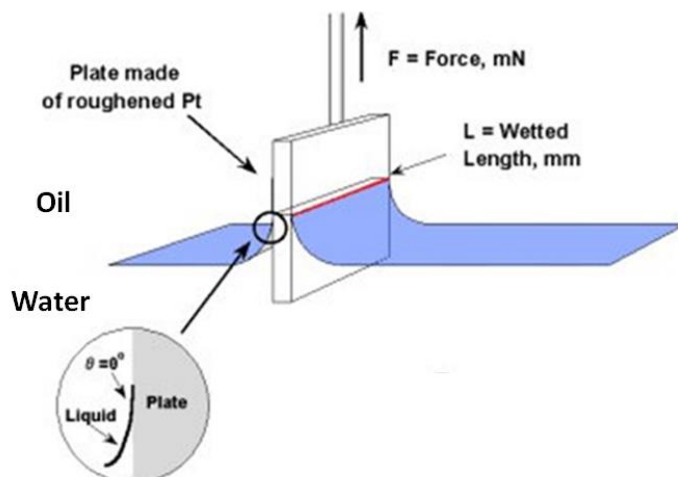


Figure 1. Principle of the IFT measurement using the Wilhelmy plate method.

1.2 Experimental data

Different analyses were carried out on the crude oil A1, as well as on its light and heavy fractions. The corresponding experimental data are detailed in the following sections.

1.2.1 Experimental data of crude oil A1

For crude oil A1, the experimental data available are the density and the elemental analysis (C,H,O,N,S).

Table 1. Elemental analysis and density for crude oil A1.

Elemental analysis (% w/w):		Volumetric mass density:	
Carbon content	83.6	at 15 °C (g/cm ³)	0.8745
Hydrogen content	12.47	at 20 °C (g/cm ³)	0.871
Nitrogen content	0.14		
Oxygen content	0.76		
Sulfur content	2.24		

1.2.2 Experimental data of the light fraction for crude A1 ($T < 344^{\circ}\text{C}$)

For the light fraction of the crude oil A1, the experimental data available are elemental analysis (C,H,O,N,S), density, chemical family, distribution by number of and SD curve.

Table 2. Elemental analysis and density for the light fraction of crude oil A1.

Elemental analysis (% w/w):		Volumetric mass density:	
Carbon content	85.1	at 15 °C (g/cm^3)	0.7975
Hydrogen content	13.69	at 20 °C (g/cm^3)	0.8012
Nitrogen content	0.09		
Oxygen content	0		
Sulfur content	0.53		

Table 3. Chemical family of the light fraction of crude oil A1.

Chemical family (% w/w):	
<i>n</i> -Paraffins	13.8
<i>i</i> -Paraffins	10
Naphthenes	7.8
Aromatics	5.4
Saturates C ₁₅₊	5.6
Aromatics C ₁₅₊	3.6
Unknown C ₂₀₋	0

Table 4. Weight fraction of the light fraction of crude oil A1.

Distribution by number of C (Wt %):			
C ₂	0.01	C ₁₂	2.99
C ₃	0.1	C ₁₃	3.57
C ₄	0.46	C ₁₄	2.64
C ₅	1.27	C ₁₅	2.91
C ₆	2.54	C ₁₆	2.45
C ₇	3.6	C ₁₇	2.14
C ₈	3.7	C ₁₈	2.84
C ₉	4.36	C ₁₉	0.89
C ₁₀	4.69	C ₂₀	1.51
C ₁₁	3.46		

Table 5. Simulated distillation for the light fraction of crude oil A1.

Simulated distillation:							
Weight	T (°C)	Weight	T (°C)	Weight	T (°C)	Weight	T (°C)
1.00%	23.3	13.00%	141.9	25.00%	216	37.00%	287
2.00%	55.1	14.00%	151	26.00%	219.5	38.00%	294.7
3.00%	69	15.00%	151	27.00%	228.3	39.00%	302
4.00%	74.6	16.00%	160.4	28.00%	235	40.00%	306.4
5.00%	86.7	17.00%	165.7	29.00%	238.2	41.00%	313.8
6.00%	98	18.00%	174	30.00%	247.2	42.00%	317.9
7.00%	98	19.00%	174.7	31.00%	254	43.00%	325.2
8.00%	111.3	20.00%	182.5	32.00%	257.3	44.00%	330.4
9.00%	116.7	21.00%	189.4	33.00%	265.1	45.00%	338
10.00%	126	22.00%	196	34.00%	271	46.00%	344
11.00%	126	23.00%	201.8	35.00%	275.8		
12.00%	136.8	24.00%	209.1	36.00%	282.2		

1.2.3 Experimental data of the heavy fraction for crude A1 ($T > 344^{\circ}\text{C}$)

For the heavy fraction of the crude oil A1, the experimental data available are elemental analysis (C,H,O,N,S), density, SARA analysis, and SD curve.

Table 6. Elemental analysis and density for the heavy fraction of crude oil A1.

Elemental analysis (% m/m):		Volumetric mass density:	
Carbon content	84.2	at 15 °C (g/cm ³)	0.9685
Hydrogen content	11.39	at 20 °C (g/cm ³)	0.9651
Nitrogen content	0.19	at 70 °C (g/cm ³)	0.9308
Oxygen content	0.25		
Sulfur content	3.59		

Table 7. SARA analysis of the heavy fraction of crude oil A1.

SARA analysis:		
Saturates (% w/w)	29.7	± 1.1
Aromatics (% w/w)	44.4	± 1.1
Resins (% w/w)	21.6	± 1.1
Asphaltenes (% w/w)	1.7	± 0.2
Waste	2.5	

Table 8. Simulated distillation for the heavy fraction of crude oil A1.

Simulated distillation:							
Weight	T (°C)	Weight	T (°C)	Weight	T (°C)	Weight	T (°C)
IBP (0.50 %)	357	23.00%	443	46.00%	516.2	69.00%	609.7
1.00%	366.2	24.00%	446.1	47.00%	519.7	70.00%	614.2
2.00%	374.1	25.00%	449.1	48.00%	523.6	71.00%	619
3.00%	379.6	26.00%	452.1	49.00%	527.9	72.00%	624.2
4.00%	384	27.00%	455.2	50.00%	532.4	73.00%	629.6
5.00%	388.1	28.00%	458.1	51.00%	535.9	74.00%	634.7
6.00%	391.4	29.00%	461.3	52.00%	539	75.00%	640.2
7.00%	395.1	30.00%	464.5	53.00%	542.1	76.00%	646.1
8.00%	398.6	31.00%	467.4	54.00%	545.4	77.00%	652.5
9.00%	401.7	32.00%	470.5	55.00%	550	78.00%	658.1
10.00%	405	33.00%	473.4	56.00%	554.7	79.00%	665.3
11.00%	408.2	34.00%	476.5	57.00%	558.8	80.00%	672.4
12.00%	411.4	35.00%	479.6	58.00%	562.7	81.00%	679.5
13.00%	414.2	36.00%	482.7	59.00%	566.6	82.00%	686.8
14.00%	417.2	37.00%	486.1	60.00%	570.1	83.00%	693.5
15.00%	420.1	38.00%	489.4	61.00%	573.7	84.00%	700.7
16.00%	422.7	39.00%	492.9	62.00%	577.9	85.00%	707.2
17.00%	425.6	40.00%	496.2	63.00%	582.3	86.00%	714.2
18.00%	428.6	41.00%	499.5	64.00%	586.6	87.00%	721.2
19.00%	431.3	42.00%	502.6	65.00%	590.8	88.00%	727.4
20.00%	434.4	43.00%	505.9	66.00%	595.3	89.00%	734
21.00%	437.2	44.00%	509.2	67.00%	599.9	90.00%	740
22.00%	440.1	45.00%	512.7	68.00%	604.9	91.00%	749.1

2. Thermodynamic models selection for prediction of liquid-liquid equilibria and parametrization of DPD interactions.

2.1 Prediction of the composition of liquid-liquid equilibria

In previous work,³ it has been shown that interaction parameters for CG simulations can be determined using compositional data of liquid-liquid equilibria (LLE). Whenever the experimental data are not available, thermodynamic based methods exist, such as activity coefficient models and Equations of State (EoS). These thermodynamic models allow calculating thermophysical properties and phase equilibrium for a mixture at a given temperature (T), pressure (P), and composition. An Equation of State is a PVT relation, with V the volume, that is applicable over wide ranges of temperatures, pressures and compositions, and it can provide satisfactory results for hydrocarbons, gases, non-polar compounds and, even polar compounds for some specific EoS. Activity coefficient models are more reliable for predicting the behavior of non-ideal liquids, especially for polar mixtures. It can be noted that activity coefficient models are limited to liquid systems while EoS can be used for liquid and gaseous systems.

During phase equilibrium calculations, the chemical potential μ_i of a molecule i is defined as:

$$\mu_i = \mu_i^{ref}(T_0, P_0) + RT \times \ln \left(\frac{f_i}{f_i^{ref}} \right) \quad (2)$$

where μ_i^{ref} is the chemical potential of the molecule i at the reference state, f_i the fugacity of the molecule i and f_i^{ref} the fugacity of the molecule i at the reference state. The reference state is

generally taken either as the pure liquid solvent or as the fluid mixture in the ideal gas state at the same pressure and temperature. In the first case, the equation (2) becomes:

$$\mu_i = \mu_i^*(T, P) + RT \times \ln(x_i \gamma_i) \quad (3)$$

where γ_i is the activity coefficient, which can be calculated using a suitable activity coefficient model. The reference state (indicated by a “*”) is generally taken, for neutral molecules, at its vapor pressure. The drawback of this approach is that the pressure is not taken into account since the activity coefficient models are generally pressure independent. In the second case, equation (2) becomes:

$$\mu_i = \mu_i^\#(T, P, x) + RT \times \ln(\varphi_i) \quad (4)$$

where φ_i is the fugacity coefficient that can be calculated using an EoS. The reference state (indicated by “#”) is then the fluid mixture taken as an ideal gas at the same temperature T and pressure P as the fluid mixture, and x compositions. The relationship between the two approaches can be obtained using the definition of activity coefficients:

$$\gamma = \frac{\varphi_i}{\varphi_i^*} \quad (5)$$

where φ_i^* is the fugacity coefficient of molecule i at the reference state.

There are a large number of EoS and activity coefficient models. For example, Peng-Robinson EoS⁴ (PR) and the Soave-Redlich-Kwong⁵ (SRK) EoS belongs to the cubic EoS. Other forms of EoS are based on statistical thermodynamic concepts such as the Statistical Associating Fluid Theory (SAFT) EoS, which was proposed by Chapman *et al.*^{6,7} on the basis of Wertheim’s perturbation theory^{8,9}. Among the activity coefficient models, the most commonly known are

Wilson¹⁰, NRTL (Non-random two-liquid model)¹¹, UNIQUAC (UNIversal QUAsiChemical)¹² and UNIFAC (UNIversal Functional Activity Coefficient)¹³. Moreover, many extensions have been developed for each of these thermodynamic models. Description of all thermodynamic models and their extensions are out of the scope of this work.

In order to extend the validity of the parametrization approach proposed in this work, a comparative study has been conducted between several thermodynamic models in order to select the most appropriate to predict liquid-liquid compositional data. The final goal is to use LLE to parametrize the DPD repulsion parameters using the methodology developed in our previous work³. For simplicity, LLE are calculated for the same experimental compositions and for the six systems studied in this previous work, including partially miscible systems including hydrocarbons and polar molecules able to form hydrogen bonds with water. As an example, results for the water/benzene/1,4-dioxane system are presented in Table 9 (other systems not shown). Thermodynamic models compared in Table 9 are those available in the Simulis Thermodynamics software¹⁴.

Table 9. Comparison of thermodynamic models for prediction of liquid-liquid equilibrium for the water/benzene/1,4-dioxane system at 298.15K.

Thermodynamic models		$R_{solubility}^2$ ^[a]
EoS models	PR	No solute distribution between bulk phases
	SRK	No solute distribution between bulk phases
	GC-PPC-SAFT	0.962

	NRTL-SAC	0.983
Activity models	UNIFAC original	0.958
	UNIFAC LL	0.961
	SRK-MHV2-UNIFAC	0.998

$^{[a]} R_{solubility}^2 = 1 - \left[\frac{\sum_{i=1}^{2 \times N} (x_i^{exp} - x_i^{calc})^2}{\sum_{i=1}^{2 \times N} (x_i^{exp} - \bar{x})^2} \right]$ with $\bar{x} = \frac{1}{2 \times N} \sum_{i=1}^{2 \times N} x_i^{exp}$ where x_i is the molar fraction of solute in the aqueous phase and in the organic phase for the i^{th} composition. N is the number of compositions and a factor 2 is added because the molar fraction of the solute is taken into account for both bulk phases.

The criterion for comparing the thermodynamic models is the regression function, $R_{solubility}^2$, which corresponds to the error on the solute distribution between the aqueous and the organic phases compared to the experimental data. Note that this criterion was used in ref ³ to compare the parameterization approaches of CG simulations to reproduce LLE.

PR and SRK do not reproduce the LLE of water/benzene/1,4-dioxane system, as shown in Table 9. These results are not surprising since these cubic EoS are not suitable for polar molecules. Group-Contribution Polar Perturbed-Chain SAFT (GC-PPC-SAFT) model is an extension of the SAFT model where a specific term (PPC) is added to deal with polar systems¹⁵. This model relies on a group contribution (GC) method to compute equation's settings for a large number of chemical families. GC method is a technique to estimate and predict a property value only knowing the molecular structures of the studied fluids. The predictive capability of the GC-PPC-SAFT EoS for liquid-liquid equilibria of water/benzene/1,4-dioxane system is good considering the agreement between predicted and experimental data, as shown in Table 9 ($R_{solubility}^2 = 0.962$). However, for other systems not shown here, LLE are not well reproduced using GC-PPC-SAFT EoS.

Among existing activity coefficient models, we propose to compare NRTL-sac and UNIFAC model. The NRTL-SAC (SAC = Segment Activity Coefficient) model¹⁶⁻¹⁸ is based on the original NRTL¹¹ and polymer NRTL¹⁹, and its particularity stands in the characterization of molecules in terms of predefined conceptual segments. As shown in Table 9, NRTL-SAC

provides LLE for the water/benzene/1,4-dioxane system in good agreement with experimental data ($R_{solubility}^2 = 0.983$). The UNIFAC model, which is based on a GC method, also provides LLE predictions in agreement with experimental data for water/benzene/1,4-dioxane system ($R_{solubility}^2 = 0.958$). In addition, UNIFAC LL which is a specific model for liquid-liquid equilibria^{20,21}, improves the accuracy of LLE reproduction with $R_{solubility}^2 > 0.900$ in most cases studied in this work.

The thermodynamic model that provides the most accurate predictions of LLE for the six systems studied here is the SRK-MHV2-UNIFAC model. Developed in 1978 by Vidal^{22,23}, SRK-MHV2-UNIFAC is based on a combination of EoS (SRK) with G^E mixing rules (MHV2). This kind of approach allows using cubic EoS for polar compounds (water, alcohol, etc.). G^E mixing rules obtain the interaction information from excess Gibbs energy G_V^E models, originally developed for the prediction of liquid activity coefficients. In the case of SRK-MHV2-UNIFAC, interaction information is given by the UNIFAC model. Figure 2 presents the prediction of the liquid-liquid equilibrium ternary diagram for the water/benzene/1,4-dioxane system using the SRK-MHV2-UNIFAC model and the results are in very good agreement with available experimental data ($R_{solubility}^2 = 0.998$).

It has been shown that thermodynamic models can predict compositional data for LLE in good agreement with experimental data. Therefore, it is possible to parameterize DPD simulations using the parameterization approaches developed in ref³ for liquid-liquid equilibrium systems when experimental data is available or not.

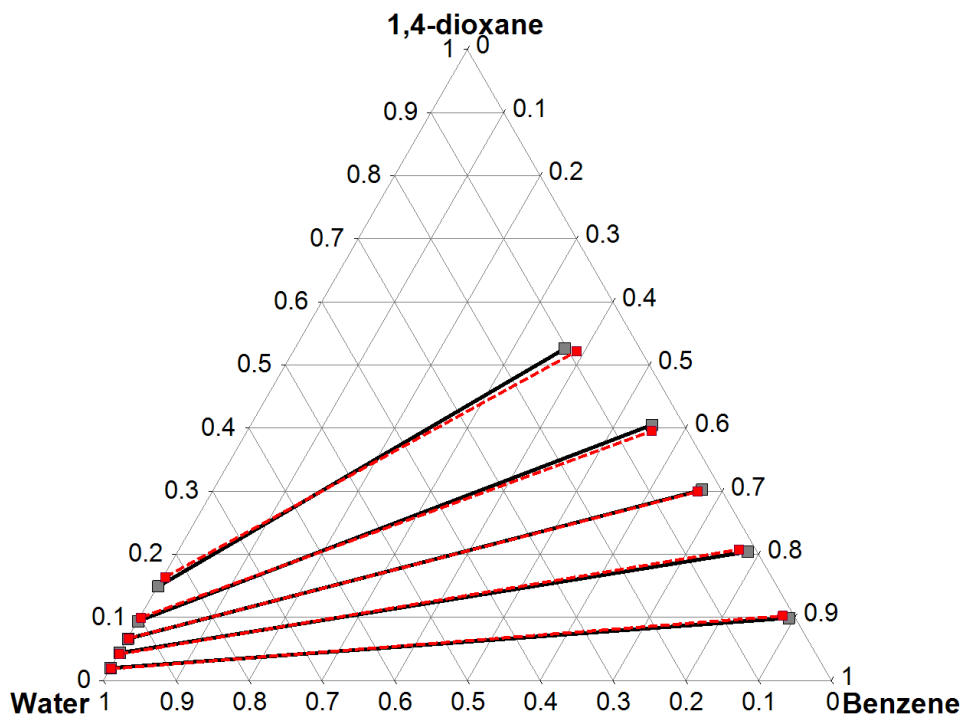


Figure 2. Prediction of liquid-liquid equilibrium ternary diagrams for the water/benzene/1,4-dioxane system at 298.15K using the SRK-MHV2-UNIFAC model. Experimental data are plotted in black (solid lines for the conodal lines) and the results from the thermodynamic model are in red (dashed lines for the conodal lines).

2.1.1 Prediction of DPD interaction parameters using a thermodynamic model

This section aims at defining the parametrization approach where the interaction parameters are obtained directly from chemical potential or activity coefficients calculated using a thermodynamic model. The phase composition data and the activity coefficient obtained with the thermodynamic model are used to obtain the Flory-Huggins parameter by means of the equation (13) of the main manuscript. Similar approaches have been proposed in the literature where activity coefficients at infinite dilution have been related to the interaction parameters^{24,25}.

The choice of the thermodynamic model is an important step for predicting the activity coefficient and, therefore, appropriate interaction parameters. Alasiri and Chapman²⁴ have proposed to estimate the activity coefficients at infinite dilution using COSMO-RS model²⁶. However, Vishnyakov *et al.*²⁵ have shown that calculations of parameters using COSMO-RS “did not always lead to plausible results”. Indeed, they obtained results which deviate from reference experimental data for a system where the polar head group of a surfactant is more soluble in the organic phase than in the aqueous phase. In this section, a comparative study is carried out in order to determine the best thermodynamic model to predict activity coefficients and chemical potentials of solute in a solvent. A set of validated interaction parameters allowing reproducing the liquid-liquid equilibrium and the variation of the interfacial tension have been obtained for the water/benzene/1,4-dioxane (other systems such as water/chloroform/acetone and water/benzene/acetic acid were also explored). These parameters are used as a reference to choose the most appropriate thermodynamic model.

In a first step, the chemical potential of the solute in the aqueous phase and in the organic phase are calculated with several thermodynamic models and values are reported for each bulk phases (Figure 3 (a) for the aqueous phase and (b) for the organic phase). Chemical potential used in CG simulations were derived from interaction parameters (using $a_{ij}=3.5*\chi_{ij}+25$ and equation 13 of the main manuscript). Interaction parameters feeding CG simulations can be compared to those obtained from thermodynamic models (Figure 4 (a) for the aqueous phase and (b) for the organic phase).

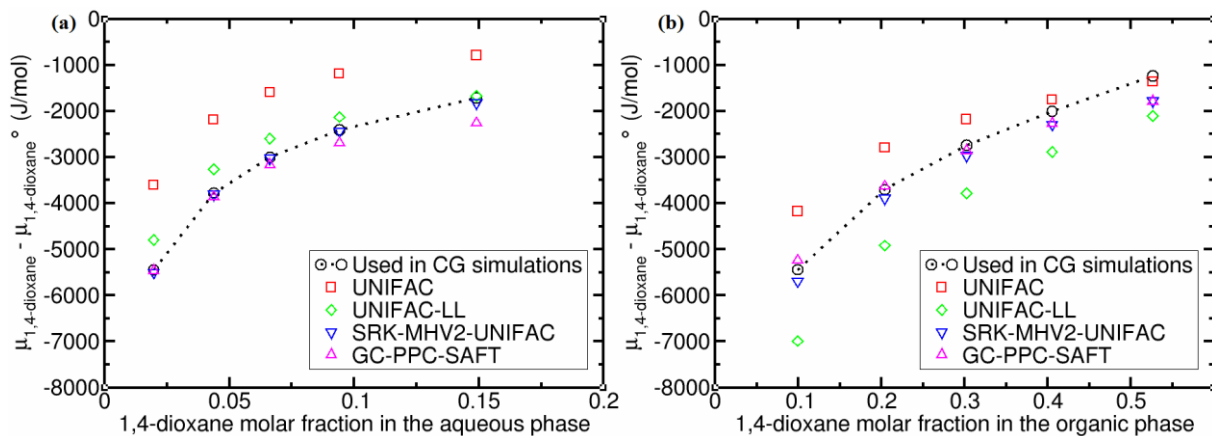


Figure 3. (a) Chemical potential of the 1,4-dioxane in the aqueous phase and (b) in the organic phase for the water/benzene/1,4-dioxane system.

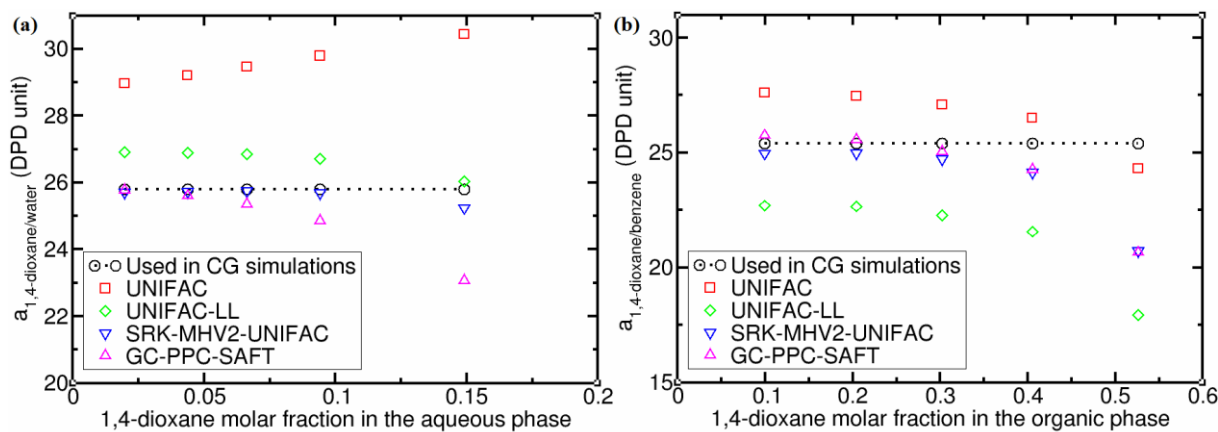


Figure 4. (a) Interaction parameters of the 1,4-dioxane with water for the water/benzene/1,4-dioxane system. (b) Interaction parameters of the 1,4-dioxane with benzene for the water/benzene/1,4-dioxane system.

For the water/benzene/1,4-dioxane system presented in Figure 3 (a) and (b), variation of the chemical potential with the composition is in a very good agreement with the values of chemical potential in DPD simulations using the SRK-MHV2-UNIFAC and GC-PPC-SAFT models. It can be noted that GC-PPC-SAFT does not give a good prediction of the composition

of the LLE for these two systems, therefore, predicted chemical potential values are not relevant. Among the thermodynamic models compared, it seems that SRK-MHV2-UNIFAC model is the most suitable for parameterization of CG simulations.

In Figure 4 (a) and (b), the interaction parameters of 1,4-dioxane with water and with benzene are plotted as a function of the composition in the aqueous and organic phase, respectively. Since there is a direct relationship between interaction parameters and chemical potential, SRK-MHV2-UNIFAC still provides the best interaction parameters compared to those used in DPD simulations. However, it can be noticed that interaction parameters obtained from thermodynamic models are sensitive to the composition. The composition dependence of interaction parameters has already been studied in ref ³. These results obtained from the thermodynamic models confirm that the use of constant parameters over the whole composition range is an important approximation in the DPD simulations and its use should be verified for the specific system studied.

Finally, interaction parameters are calculated using the SRK-MHV2-UNIFAC model for water/solute and organic compound/solute binary systems and results are presented in Table 10. Parameters are calculated for a given composition since the equation (13) of the main manuscript between the activity coefficient and the Flory-Huggins parameter involves the composition of the system (volume fractions). The chosen composition corresponds to a relatively diluted system with 1 mol% of solute and 99 mol% of the solvent. This method consists, to a certain extent, in calculating the activity coefficient at infinite dilution activity with a validated thermodynamic model.

Table 10. Interaction parameters calculated for binary systems using the SRK-MHV2-UNIFAC thermodynamic model for water/solute and organic compound/solute binary systems. (Compositions: 1 mol% of solute and 99 mol% of solvent). Parameters are compared with those obtained in CG simulations (using approach #2 of ref. ³) except for the water/benzene/1,4-dioxane system.

Systems	$a_{water/solute}$		$a_{organic\ compound/solute}$	
	Used in CG simulations	Calculated for binary systems	Used in CG simulations	Calculated for binary systems
water/benzene/1,4-dioxane ($N_m = 5$)	25.80	25.74	25.41	25.10
water/chloroform/acetone ($N_m = 4$)	28.20	26.91	22.07	20.91
water/benzene/acetic acid ($N_m = 4$)	21.38	23.97	31.12	30.45
water/benzene/2-propanol ($N_m = 5$)	12.46	27.83	18.96	34.08
water/hexane/acetone ($N_m = 4$)	23.58	26.91	27.44	33.83
water/hexane/2-propanol ($N_m = 4$)	23.60	28.58	29.01	38.68

Table 10 shows that $a_{water/solute}$ and $a_{organic\ compound/solute}$ interaction parameters are in good agreement with those used in CG simulations for the water/benzene/1,4-dioxane, (also for water/chloroform/acetone and water/benzene/acetic acid systems). These systems have been well parameterized to represent liquid-liquid equilibrium and interfacial tension variation in CG

simulations. In this section, different thermodynamic models were used to calculate interaction parameters of a solute in a solvent, and results were compared to those obtained from parametrization approaches based on experimental data. It has been shown that SRK-MHV2-UNIFAC model provides interaction parameters of the same order of magnitude as those used in DPD simulations using experimental data. It can be noted that these interaction parameters are not precise enough to allow an accurate reproduction of the complete LLE diagram. However, thermodynamic models allow directly calculating interaction parameters for solute/solvent beads whenever a certain degree of polarity/hydrogen bond is present, such as the heteroatoms in our system. In addition, the use of one composition at infinite dilutions to determine the interaction parameters is therefore faster compared to the other approaches based on LLE compositional data since it does not need to represent the complete liquid-liquid equilibrium of ternary systems (see reference ³ for details). Complex mixtures such as crude oil require to estimate a large amount of interaction parameters. Therefore, this approach based on the SRK-MHV2-UNIFAC model is suitable to evaluate interaction parameters for heteroatom molecules present in the crude oil.

2.1.2 Transferability of interaction parameters for alkanes

Representing a complex system such as a crude oil should result in using a higher number of bead types to represent its chemical complexity, and therefore, to a large number of interactions need to be parameterized. Transferability of interaction parameters could simplify the parameterization step. In this section, we present a study carried out to evaluate the transferability of interaction parameters to representing a linear alkane with a solute. Due to the chosen coarse-grained level, *n*-hexane molecule is represented with two beads, *n*-nonane molecules with three beads, and *n*-dodecane molecules with four beads (see Figure 5). The interactions of a solute

(1,4-dioxane, acetone, acetic acid and 2-propanol) with the three alkanes were calculated using the SRK-MHV2-UNIFAC thermodynamic model and the results are presented in Table 11. The results presented in Table 11 show that variation of the interaction parameters between the solute beads and the alkane beads as function of the alkane length is low. Therefore, it seems reasonable in the case of systems containing alkanes of different lengths to use the same solute/alkane interaction parameters. A similar hypothesis is used for other types of hydrocarbon beads.

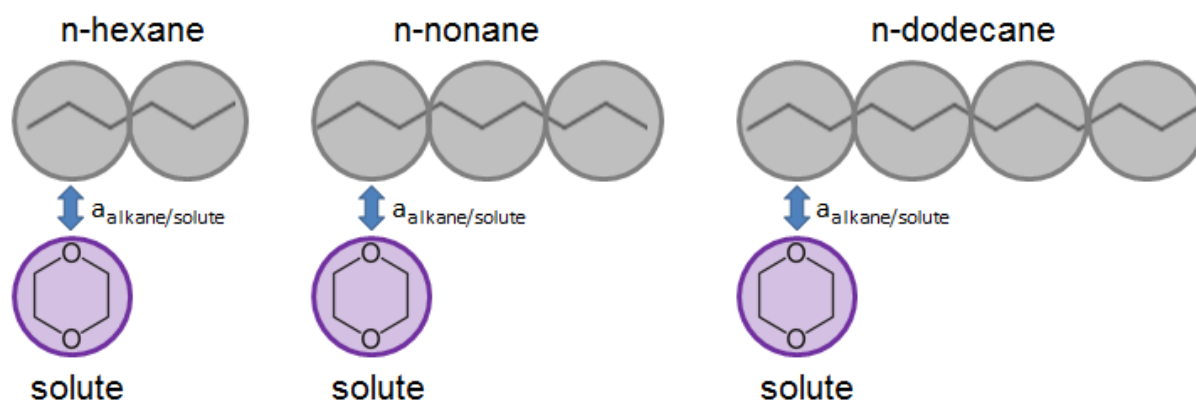


Figure 5. Representation of interactions between a solute and linear alkanes such as *n*-hexane, *n*-nonane and *n*-dodecane.

Table 11. Interaction parameters calculated for *n*-alkane/solute binary systems using the thermodynamic model SRK-MHV2-UNIFAC. (Composition used is 1 mol% in solute and 99 mol% in hydrocarbon).

Systems	χ_{ij}	a_{ij}	Systems	χ_{ij}	a_{ij}
n-hexane/1,4-dioxane	2.23	32.82	n-hexane/acetone	2.53	33.86
n-nonane/1,4-dioxane	2.39	33.38	n-nonane/acetone	2.66	34.31
n-dodecane/1,4-dioxane	2.34	33.19	n-dodecane/acetone	2.52	33.83
Systems	χ_{ij}	a_{ij}	Systems	χ_{ij}	a_{ij}

n-hexane/acetic acid	3.10	35.87	n-hexane/2-propanol	3.87	38.54
n-nonane/acetic acid	3.35	36.71	n-nonane/2-propanol	4.09	39.30
n-dodecane/acetic acid	3.32	36.62	n-dodecane/2-propanol	4.01	39.03

3. Calculation of thermodynamic properties for crude oil

3.1 Lumping method

A “breakdown” procedure, developed by Eckert and Vanek²⁷ was applied to define a list of pseudo-components. The density was estimated using the Watson approach²⁸. Other thermodynamic properties of the pseudo-components (pseudo-formula, critical properties and acentric factor) were obtained using correlations from Twu²⁹ and it should be noted that the breakdown procedure can only be performed with a TBP (True Boiling Point) distillation curve if the ReFGen³⁰ software is used. To convert the simulated distillation curve into a TBP curve, we chose the procedures developed by Daubert and presented in the sixth edition of API-TDB³¹. The simulated distillation curve (ASTM D2887) is converted to an ASTM D86 distillation curve. Then, the resulting ASTM D86 curve is converted to a TBP distillation curve as shown in Table 12. Finally, the Lumping procedure proposed by Montel et al.^{32,33}, described in our previous works of Aquing et al.³⁴ and Lugo et al.³⁰ on light fuels, is used here to model the light fraction of the crude oil. We provide here the required input data to perform the selection of pseudo-components.

Table 12. Conversion of the simulated distillation curve to an ASTM D86 curve, and then to a TBP curve for crude oil A1.

% Volume	Distillation curves	
	ASTM D86	TBP
	T (°C)	T (°C)
10	106.23	83.54
30	151.88	147.62
50	198.01	205.58
70	257.38	270.33

90	308.71	321.83
100	328.88	347.37

Table 13. Calculation of molecule and pseudo-component properties.

Properties of pseudo-components derived from the distillation curve		Properties of pseudo-components derived from the Lumping method	
Molecular weight (g/mol)	Riazi ³⁵ (Procedure API B2B.1 ³¹)	Molecular weight (g/mol)	Weighted average value on initial compounds
Chemical formula (C _x H _y)	Riazi ³⁵	Chemical formula (C _x H _y)	Riazi ³⁵
Critical temperature (T _C)	Twu ²⁹	Critical temperature (T _C)	Weighted average value on initial compounds
Critical pressure (P _C)	Twu ²⁹	Critical pressure (P _C)	Weighted average value on initial compounds
Acentic factor (ω)	Edmister ³⁶	Acentic factor (ω)	Weighted average value on initial compounds
		Enthalpy of formation (kJ/mol)	Lower heating value ³⁰
		Molar volume (mol/L)	Peng et Robinson EoS ⁴
		Normal boiling temperature (K)	Riazi (Procedure API 2B1.1 ³¹)

To assign a real molecule to each pseudo-components the following objective function was minimized numerically.

$$F = \sum_{i=1}^{N_P} (w_i^{\text{pseudo}} - w_i^{\text{real}})^2 \quad (6)$$

Where w_i^{pseudo} is one of the i^{th} physical property of the pseudo-component described in Table 13 and w_i^{real} corresponds to the same i^{th} physical property of one of the molecules included in the Durand et al. database.³⁷ Comparison of an initial (containing 20 molecules) and final simplified

lumped fluid (containing only 5 molecules used in Table 4 of the main manuscript) phase envelopes are shown in Figure 6.

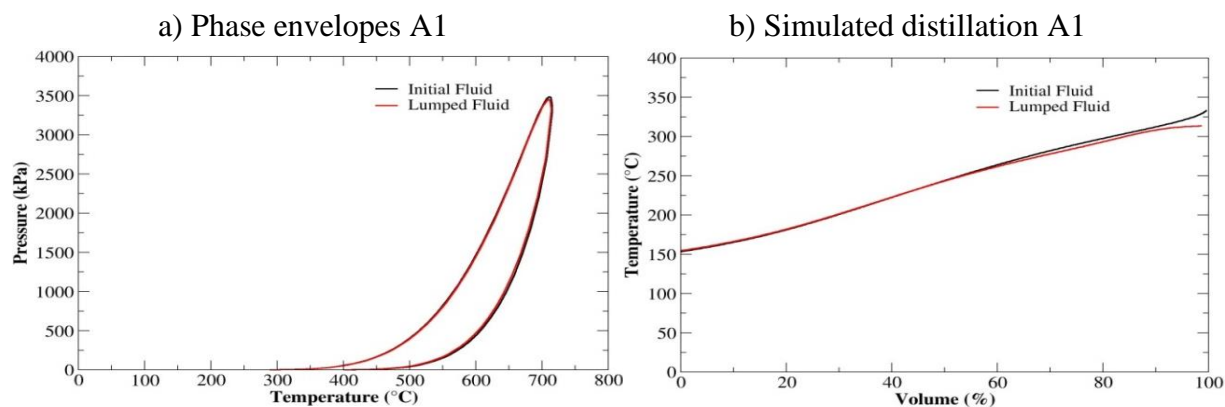


Figure 6. Comparison between the initial fluid (20 pseudo-components) and the final fluid resulting from the Lumped method (5 pseudo-components). The calculated phase envelopes and the calculated distillation curves are shown for crude oil A1.

3.2 Molecular representation

Stochastic Reconstruction (SR)

The physicochemical and thermodynamic properties of molecules set by the SR method are estimated either by direct inspection of the structure of the molecule (e.g., chemical formula, molar mass) or from correlations and group contribution methods (e.g., boiling point, density). The properties of the mixture formed by the molecules are calculated from the properties of the molecules weighted by their molar fractions. The properties of the mixture are then compared with the experimental data. The objective function F (equation (6)) is minimized by modifying

parameters of the probability distribution functions of the molecular attributes so that properties of the mixture are as close as possible to reference experimental data.

$$F = \frac{1}{N_p} \sum_{i=1}^{N_p} W_i \delta_i \quad (7)$$

where N_p is the number of properties considered in the objective function, W_i represents weights associated with δ_i the relative difference between the calculated and experimental values of the property i following equation 8.

$$\delta_i = \frac{1}{N_{M,i}} \sum_{j=1}^{N_{M,i}} \frac{|X_{j,i}^{exp} - X_{j,i}^{calc}|}{X_{j,i}^{exp}} \quad (8)$$

where $X_{j,i}^{exp}$ is the reference experimental datapoint value for the measure j of the property i , $X_{j,i}^{calc}$ is the calculated value for the measure j of the property i and $N_{M,i}$ the number of measures of property i . We selected a genetic algorithm to minimize the objective function and determine the overall minimum of the function³⁸.

Table 14. Methods for the properties of molecules and mixtures used in the SR approach.

Properties of molecules	Methods
SARA family	Wiehe ^{39,40}
Boiling temperature	Hudebine et Wahl ⁴¹
Density	Hudebine et Wahl ⁴¹
Critical temperature	Joback ⁴²
Critical pressure	Joback ⁴²
Critical molar volume	Joback ⁴²
Standard enthalpy of formation	Joback ⁴²
Cohesive energy	Fedors ⁴³

Table 15. Definition of the structural attributes used in the SR. Extracted from the work of de Oliveira *et al.*⁴⁴

	Structural attribute	Values	Distribution	Number of parameters
1	Type of molecule ^[a]	0, 1, 2 or 3	Histogram	3
2	Number of cores	$1 < X < 5$	Exponential	1
3	Type of heterocycle ^[b]	0, 1, 2 or 3	Histogram	3
4	Number of benzene rings per core	$1 < X < 5$	Exponential	1
5	Total number of rings per core	$0.5 < X < 2$	Gamma	1
6	Number of thiophenes per core	0, 1 or 2	Histogram	2
7	Number of pyridines per core	0, 1 or 2	Histogram	2
8	Number of pyrroles per core	0, 1 or 2	Histogram	2
9	Number of furans per core	0, 1 or 2	Histogram	2
10	Acceptance probability for a peripheral carbon	0 or 1	Histogram	1
11	Length of the paraffinic chains	$1 < X < 10$	Gamma	1
12	Length of an alkyl chain (lateral and intercore)	$1 < X < 10$	Exponential	1
13	Probability of sulfur substitution for aliphatic CH ₃ or CH ₂	0 or 1	Histogram	1
14	Substitution probability of a carbon atom by a heteroatom	0 or 1	Histogram	1
15	Type of heteroatom substitution ^[c]	0 or 1	Histogram	1
16	Type of oxygen group ^[d]	0 or 1	Histogram	1

^[a]Type of molecule: 0 – paraffin, 1 – naphthene, 2 – aromatic monocore, 3 – aromatic multicore.

^[b]Type of heterocycle: 0 – thiophene, 1 – pyridine, 2 – pyrrole, 3 – furan.

^[c]Type of heteroatom: 0 – nitrogen, 1 – oxygen.

^[d]Type of oxygen group: 0 – ether function, 1 – carbonyl function

Table 16. Parameters of the genetic algorithm for the heavy fraction reconstruction. Extracted from the work of de Oliveira *et al*⁴⁴

Parameters	Values
Number of molecules per individual	5000
Number of iterations of the algorithm	100
Initial number of individual	2048
Maximum percentage of refused molecules	50
Average mutation percentage per individual	25
Number of crossover points	1
Population percentage kept for the next generation	50
Minimum number of individuals	128
Reduction rate of the number of individuals	1
Frequency of reevaluation of the parents	1

Reconstruction by Entropy Maximization (REM)

The term of Shannon entropy S (which is not the thermodynamic definition) in equation (9) must be maximized in order to obtain the optimal result:

$$S = - \sum_{i=1}^N x_i \times \ln x_i \quad (9)$$

with

$$\sum_{i=1}^N x_i = 1 \quad (10)$$

where N is the total number of molecules in the library and x_i is the molar fraction of the molecule i .

The Shannon entropic term measures the homogeneity of the molar fractions of the molecules. Maximizing this term ensures that the distribution of molecules is as uniform as possible. However, if the value of the Shannon entropic term is maximized without setting any constraints, the distribution of the molecules will be equimolar (as it is by default by the SR method described before). Constraints correspond to the analytical data and are added to the Shannon entropic term by means of Lagrange parameters as shown in equation (11):

$$H = - \sum_{i=1}^N x_i \cdot \ln x_i + \mu \cdot \left(1 - \sum_{i=1}^N x_i \right) + \sum_{j=1}^J \lambda_j \cdot \left(f_j - \sum_{i=1}^N x_i \cdot f_{i,j} \right) \quad (11)$$

where H represents the constrained Shannon entropy criterion, f_j is the value of the constraint j , $f_{i,j}$ represents the property or coefficient of molecule i for constraint j , μ represents the Lagrange multiplier associated with the mass balance constraint, λ_j represents the Lagrange multiplier associated with constraint j , and J is the total number of equality constraints.

3.3 DPD simulations of crude oil

Considerations about the system size for interfacial tension calculation:

We have performed simulations with several system sizes and determined that $L_x=L_y>10$ is the most appropriate limit for the interfacial tension calculation (this fact was also observed by Ghoufi et al for alkanes⁴⁵). To keep a security range, we decided to choose an optimum interface area of $L_x=L_y=30$ to allow larger molecules to easily accommodate at the interface, but not excessively large to prevent the formation of a curved interface. Three other criteria should be taken into consideration to select the appropriate system size: 1) an aqueous phase sufficiently large to avoid any interference between both water/oil interfaces (at least 15 DPD units), 2) the possibility to include enough molecules of the different species with the lowest molar fraction considered (at least 10 molecules per specie) and, 3) left enough space to develop the interface zone (~5 or 10 DPD units for each interface) with the possibility to recover the concentrations of the bulk phase.

Table 17. Hildebrand solubility parameters extracted from the DIPPR⁴⁶.

Bead types	Hildebrand solubility parameters, δ (in $(\text{J}/\text{cm}^3)^{1/2}$)
Water	47.86
<i>n</i> -paraffin (<i>n</i> -hexane)	14.88
<i>i</i> -paraffin (2-butane)	12.56
Benzene	18.73
p-xylene	17.94
Methylcyclohexane	16.09
Furan	18.50
Phenol	24.90
Thiophene	20.15
Thiol (methanethiol)	19.37
Pyridine	21.60
Ethylamine	19.10
Benzaldehyde	21.63
Acetaldehyde	20.19

Table 18. Composition of the simulation box for the crude oil/water system A1 (Representation 1).

	Molecule name or ID	Molar fraction/fraction (% mol)	Molar fraction (% mol)	Number of molecules
	Water			54005
Light fraction (C ₂₀ -)	Methylcyclohexane	13.1	9.59	2299
	p-xylene	18.4	13.44	3222
	Isobutylbenzene	22.7	16.60	3979
	n-heptadecane	20.5	14.96	3584
	1-propyl- naphthalene	25.3	18.51	4436
Heavy fraction (C ₂₀ +)	(1) Aromatic	7.19	1.93	463
	(2) Aromatic	9.08	2.44	585
	(3) Aromatic	39.03	10.5	2515
	(4) Saturate	6.43	1.73	414
	(5) Saturate	17.1	4.61	1105
	(6) Saturate	8.8	2.37	567
	(7) Resin	5.42	1.46	350
	(8) Resin	6.80	1.85	444

Table 19. Composition of the simulation box for the crude oil/water system A1 (Representation 2).

	Molecule name or ID	Molar fraction/fraction (% mol)	Molar fraction (% mol)	Number of molecules
	Water			53985
Light fraction (C ₂₀ -)	Methylcyclohexane	13.1	9.59	2373
	p-xylene	18.4	13.44	3326
	Isobutylbenzene	22.7	16.60	4107

	n-heptadecane	20.5	14.96	3700
	1-propyl-	25.3	18.51	4579
	naphthalene			
<hr/>				
	(1) / Aromatic /	7.95	2.14	529
	(2) / Saturated /	6.33	1.70	422
	(3) / Aromatic /	24.9	6.70	1658
Heavy fraction	(4) / Resin /	10.6	2.85	705
(C ₂₀₊)	(5) / Saturate /	33.98	9.14	2261
	(6) / Aromatic /	13.3	3.58	883
	(7) / Saturate /	2.72	0.73	181
	(8) / Asphaltene /	0.24	0.06	16
<hr/>				

Table 20. Composition of the simulation box for the crude oil/water system A1 (Representation 3).

	Molecule name or ID	Molar fraction/fraction (% mol)	Molar fraction (% mol)	Number of molecules
	Water			54016
Light fraction (C ₂₀₋)	Methylcyclohexane	13.1	0.10	1951
	p-xylene	18.4	0.13	2735
	Isobutylbenzene	22.7	0.17	3377
	n-heptadecane	20.5	0.15	3042
	1-propyl- naphthalene	25.3	0.19	3765
Heavy fraction (C ₂₀₊)	(1) Aromatic	15.45	4.16	845
	(2) Aromatic	16.10	4.33	881
	(3) Resin	3.95	1.06	216
	(4) Saturate	11.94	3.21	654
	(5) Resin	8.72	2.35	477
	(6) Resin	3.70	1.00	202
	(7) Aromatic	9.70	2.61	531
	(8) Saturated	6.02	1.62	329
	(9) Saturate	24.41	6.57	1336

Table 21. Composition of the simulation box for the crude oil/water system A1 (Representation 4).

	Molecule name or ID	Molar fraction/fraction (% mol)	Molar fraction (% mol)	Number of molecules
	Water			54002
Light fraction (C ₂₀₋)	Methylcyclohexane	13.1	9.59	2200
	p-xylene	18.4	13.44	3084
	Isobutylbenzene	22.7	16.60	3808

	n-heptadecane	20.5	14.96	3431
	1-propyl-naphthalene	25.3	18.51	4246
Heavy fraction (C ₂₀₊)	(1) Saturate	12.72	3.42	785
	(2) Resin	9.66	2.60	596
	(3) Aromatic	7.72	2.08	476
	(4) Saturate	4.67	1.26	288
	(5) Aromatic	51.81	13.94	3197
	(6) Saturate	8.41	2.26	519
	(7) Aromatic	5.02	1.35	310

Table 22. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 5).

	Molecule name or ID	Molar fraction/fraction (% mol)	Molar fraction (% mol)	Number of molecules
	Water			53984
Light fraction (C ₂₀₋)	Methylcyclohexane	13.12	9.59	2299
	p-xylene	18.39	13.44	3222
	Isobutylbenzene	22.71	16.60	3979
	n-heptadecane	20.46	14.96	3584
	1-propyl-naphthalene	25.32	18.51	4436
Heavy fraction (C ₂₀₊)	(1) Saturate	7.27	1.96	437
	(2) Aromatic	5.87	1.58	353
	(3) Resin	6.42	1.73	385
	(4) Saturate	10.63	2.86	639
	(5) Aromatic	0.98	0.26	59
	(6) Aromatic	8.92	2.40	536
	(7) Aromatic	2.06	0.55	124
	(8) Aromatic	3.89	1.05	233
	(9) Aromatic	4.55	1.22	273

(10) Aromatic	9.26	2.49	556
(11) Saturate	16.79	4.52	1008
(12) Aromatic	18.40	4.95	1105
(13) Resin	4.92	1.32	296

Table 23. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 6).

	Molecule name or ID	Molar fraction/fraction (% mol)	Molar fraction (% mol)	Number of molecules
	Water			53987
Light fraction (C ₂₀ -)	Methylcyclohexane	13.12	9.59	2392
	p-xylene	18.39	13.44	3353
	Isobutylbenzene	22.71	16.60	4141
	n-heptadecane	20.46	14.96	3731
	1-propyl- naphthalene	25.32	18.51	4617
Heavy fraction (C ₂₀ +)	(1) Aromatic	12.11	3.26	813
	(2) Saturate	1.86	0.50	125
	(3) Aromatic	23.95	6.44	1607
	(4) Saturate	8.34	2.24	560
	(5) Aromatic	4.52	1.22	303
	(6) Saturate	0.54	0.15	36
	(7) Saturate	1.86	0.50	125
	(8) Saturate	0.54	0.15	36
	(9) Saturate	8.40	2.26	564
	(10) Resin	4.03	1.08	270
	(11) Aromatic	13.68	3.68	918
	(12) Resin	8.38	2.25	562
	(13) Resin	3.13	0.84	210
	(14) Saturate	6.52	1.75	438
	(15) Aromatic	2.13	0.57	143

Table 24. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 7).

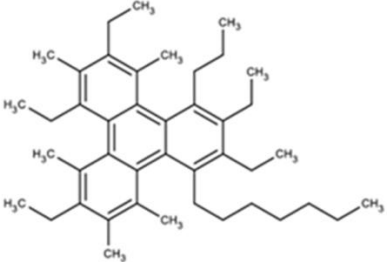
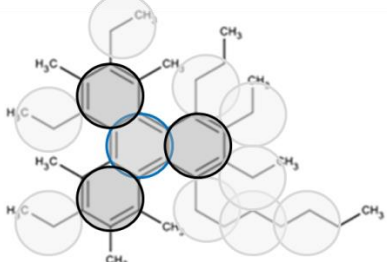
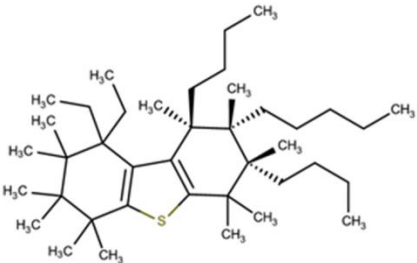
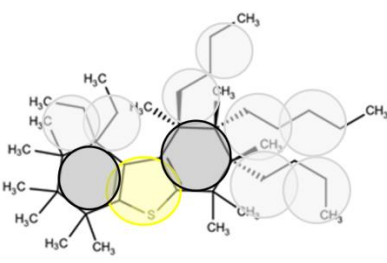
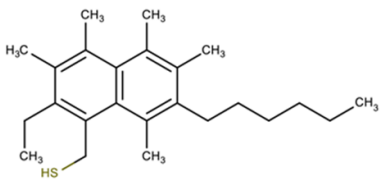
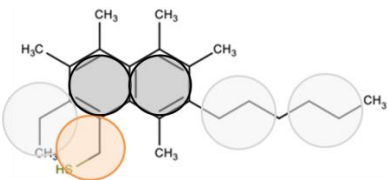


	Molecule name or ID	Molar fraction/fraction (%) mol)	Molar fraction (% mol)	Number of molecules
	Water			53994
Light fraction (C ₂₀₋)	Methylcyclohexane	13.12	9.59	1907
	p-xylene	18.39	13.44	2673
	Isobutylbenzene	22.71	16.60	3301
	n-heptadecane	20.46	14.96	2974
	1-propyl-naphthalene	25.32	18.51	3681
Heavy fraction (C ₂₀₊)	(1) Aromatic	1.00	0.27	53
	(2) Aromatic	13.57	3.65	726
	(3) Aromatic	7.92	2.13	424
	(4) Saturate	0.87	0.23	47
	(5) Aromatic	3.52	0.95	188
	(6) Aromatic	5.04	1.36	270
	(7) Saturate	11.35	3.05	607
	(8) Saturate	13.85	3.73	741
	(9) Saturate	0.87	0.23	47
	(10) Saturate	17.26	4.64	923
	(11) Aromatic	10.84	2.92	580
	(12) Resin	1.94	0.52	104
	(13) Aromatic	4.90	1.32	262
	(14) Resin	7.06	1.90	377

Table 25. Composition of the simulation box for the crude oil/water system A1 with 20 molecules (Representation 8).

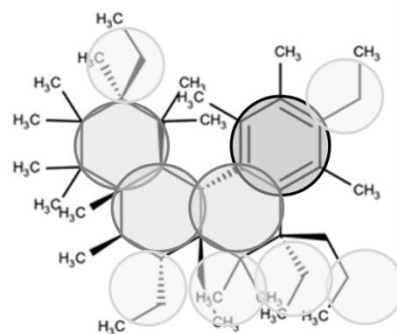
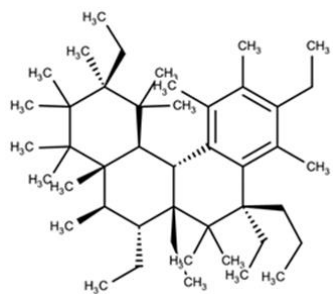
	Molecule name or ID	Molar fraction/fraction (%) mol)	Molar fraction (% mol)	Number of molecules
	Water			53953
Light fraction (C ₂₀₋)	Methylcyclohexane	13.12	9.59	1967
	p-xylene	18.39	13.44	2757

	Isobutylbenzene	22.71	16.60	3405
	n-heptadecane	20.46	14.96	3067
	1-propyl-naphthalene	25.32	18.51	3796
<hr/>				
	(1) Saturate	14.73	3.96	813
	(2) Aromatic	0.23	0.06	13
	(3) Resin	2.46	0.66	136
Heavy fraction	(4) Saturate	0.23	0.06	13
(C ₂₀₊)	(5) Aromatic	11.11	2.99	613
	(6) Aromatic	19.38	5.21	1069
	(7) Aromatic	7.14	1.92	394
	(8) Aromatic	10.03	2.70	553
	(9) Aromatic	8.71	2.34	481
	(10) Aromatic	7.47	2.01	412
	(11) Saturate	11.42	3.07	630
	(12) Aromatic	2.86	0.77	158
	(13) Resin	0.65	0.18	36

Table 26. Chemical structures and coarse-grained models of Representation 1 (R1) of the heavy fraction for crude oil A1¹ with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.

Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mol
		(1) / Aromatic / 7.19 %
		(2) / Aromatic / 9.08%
		(3) / Aromatic / 39.04%
		(4) / Saturated / 6.43%

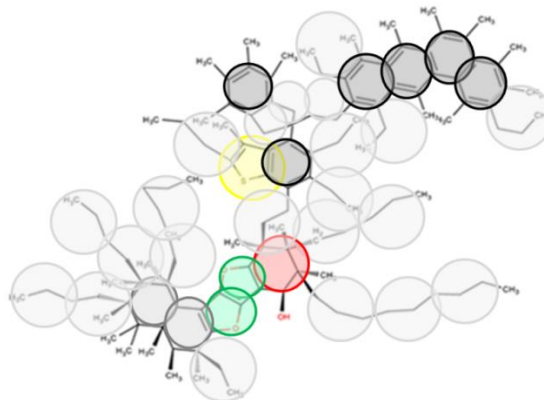
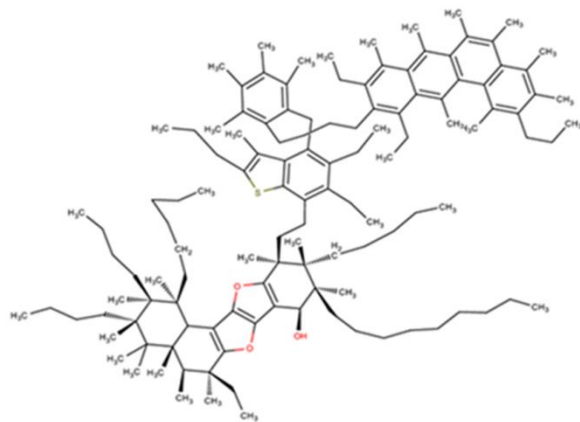
¹ For reasons of readability, an identification number (ID) is given to each molecule. The chemical family according to the SARA separation and the molar percentage of each molecule is given in the third column. The degree of coarse-graining is $N_m = 4$ (the volume of a bead corresponds to 4 water molecules).



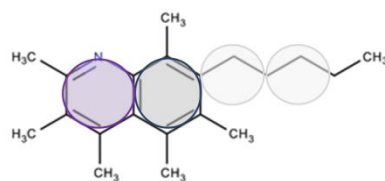
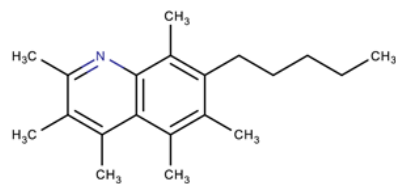
(5) / Saturate / 17.14%



(6) / Saturate / 8.8%

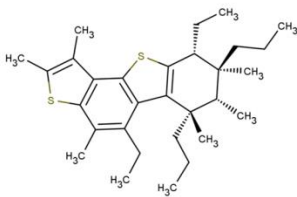

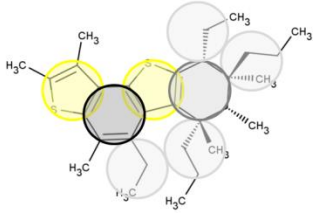

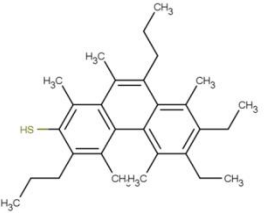
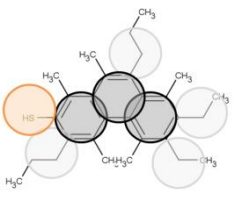
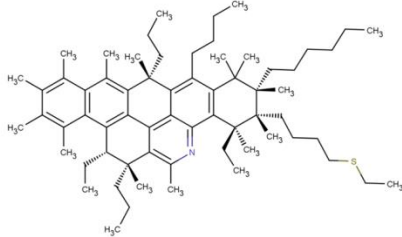
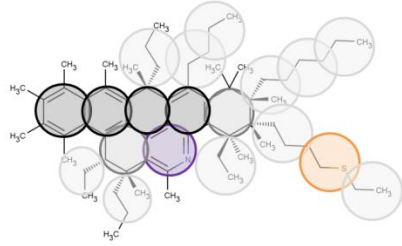


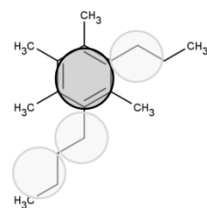
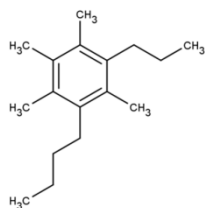
(7) / Resin / 5.42%



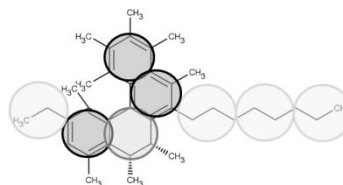
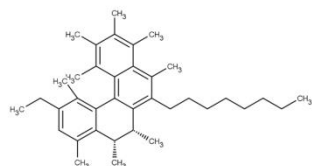
(8) / Resin / 6.88%

Table 27. Chemical structures and coarse-grained models of Representation 2 (R2) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.

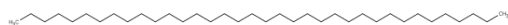
Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mol
 	 	(1) / Aromatic / 7.95 % (2) / Saturated / 6.33 %
		(3) / Aromatic / 24.9 %
		(4) / Resin / 10.6 %



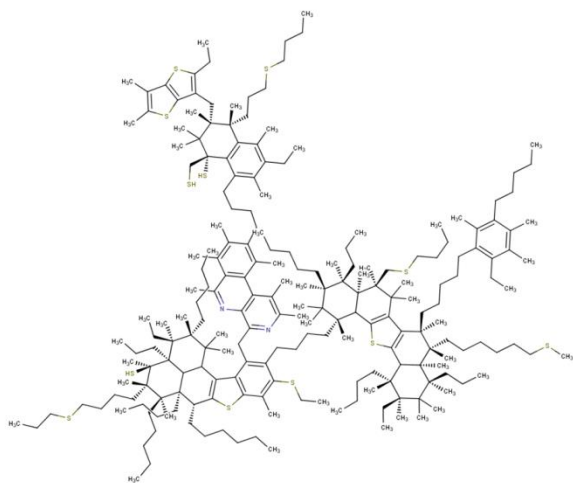
(5) / Saturate / 33.98%



(6) / Aromatic / 13.3%

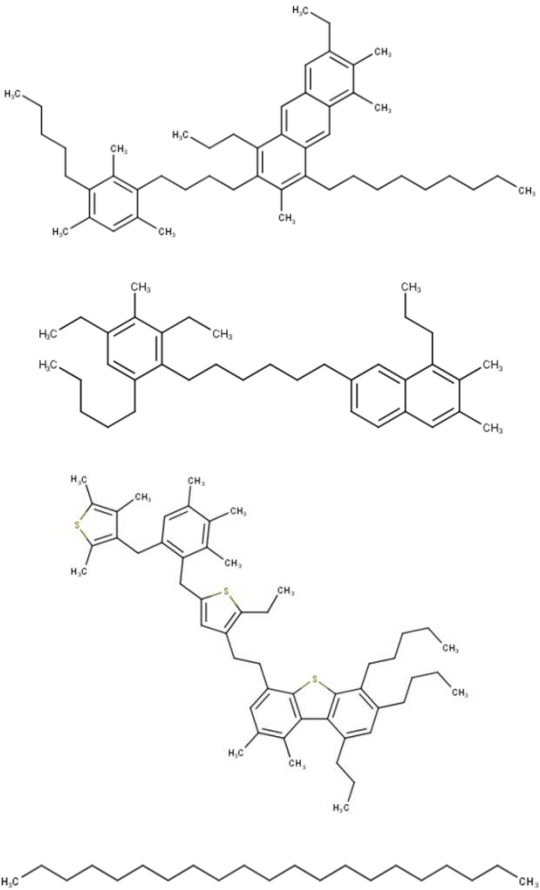
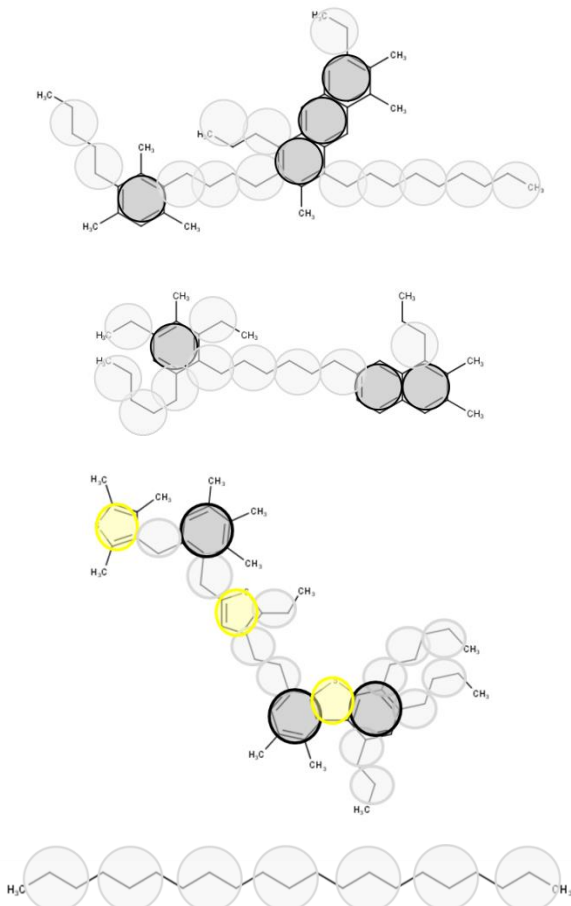


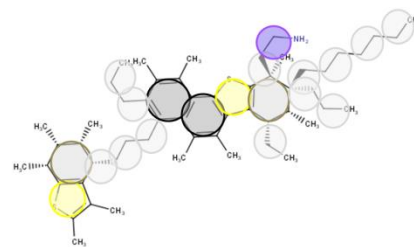
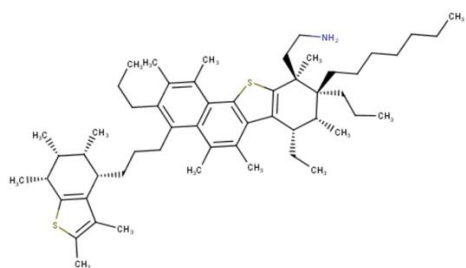
(7) / Saturate / 2.72%



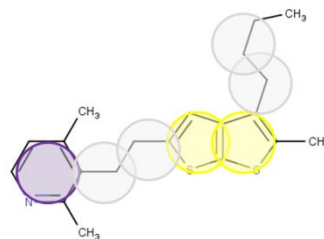
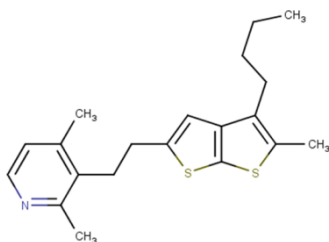
(8) / Asphaltene / 0.24%

Table 28. Chemical structures and coarse-grained models of Representation 3 (R3) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.

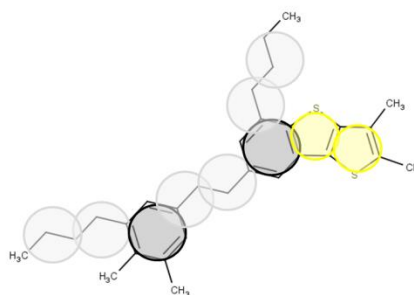
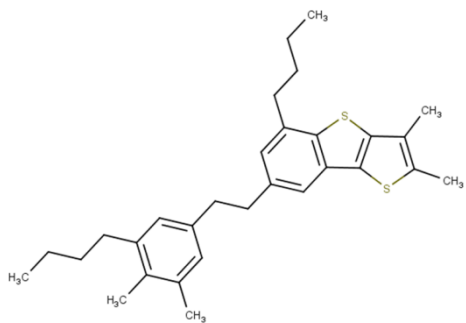
Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mole
		<p>(1) / Aromatic / 15.45 %</p> <p>(2) / Aromatic / 16.10 %</p> <p>(3) / Resin / 3.94 %</p> <p>(4) / Saturate / 11.94 %</p>



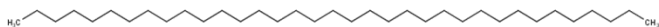
(5) / Resin / 8.72%



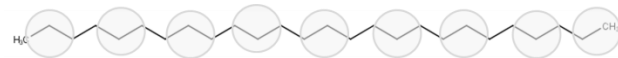
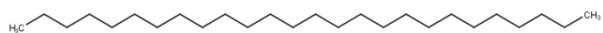
(6) / Resin / 3.7%



(7) / Aromatic / 9.7%

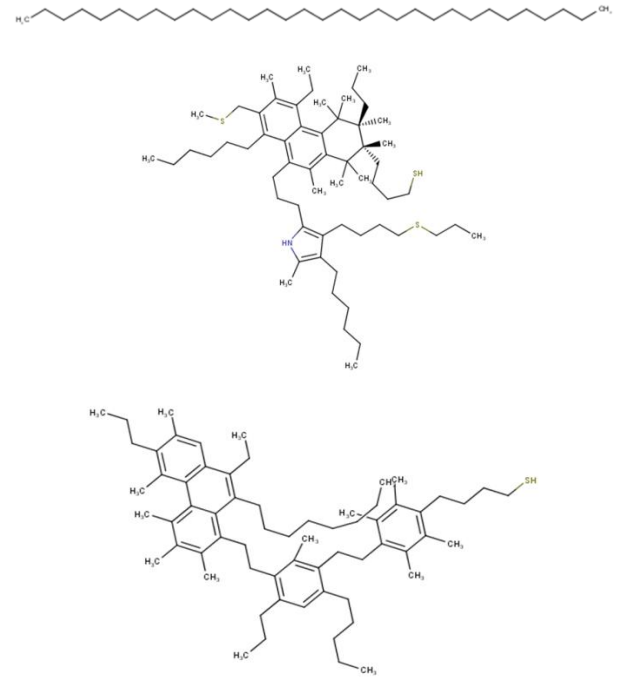
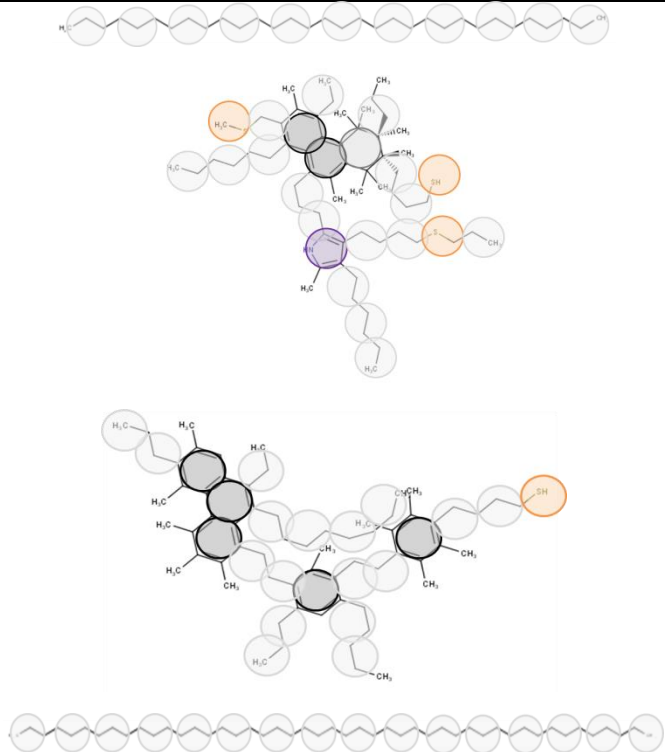


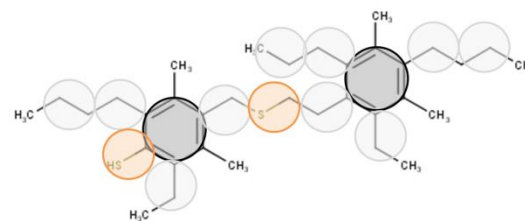
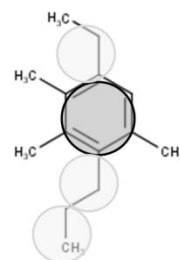
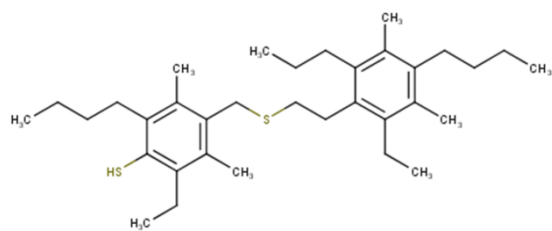
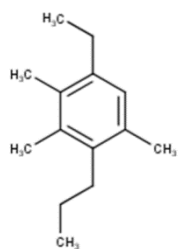
(8) / Saturated / 6.02%



(9) / Saturate / 24.41%

Table 29. Chemical structures and coarse-grained models of Representation 4 (R4) of the heavy fraction for crude oil A1 with maximum 10 molecules. Mole fractions are given with respect to the heavy fraction.

Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mole
		<p>(1) / Saturate / 12.72 %</p> <p>(2) / Resin / 9.66 %</p> <p>(3) / Aromatic / 7.71 %</p> <p>(4) / Saturate / 4.67 %</p>

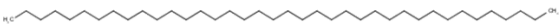

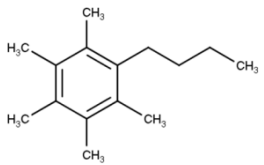
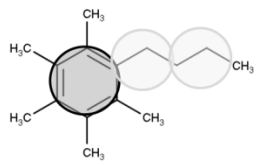
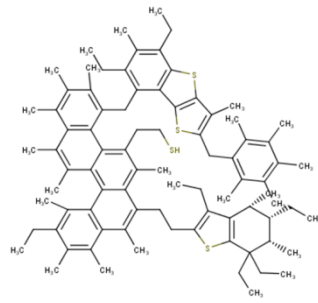
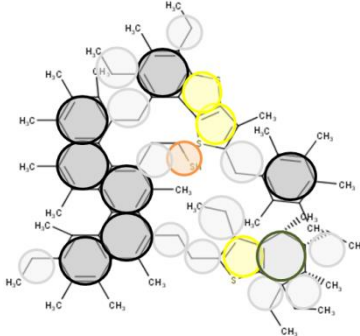
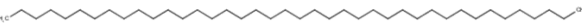

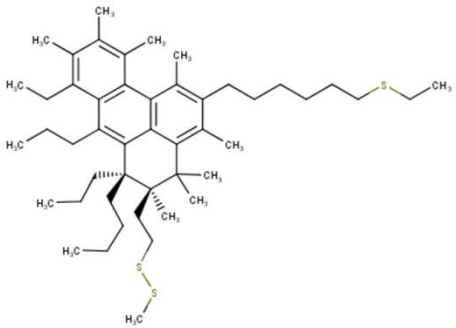
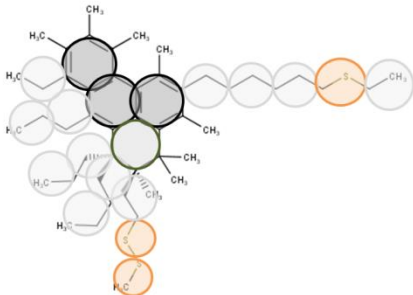


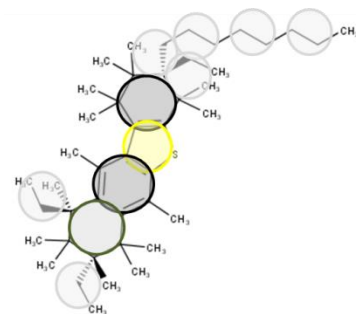
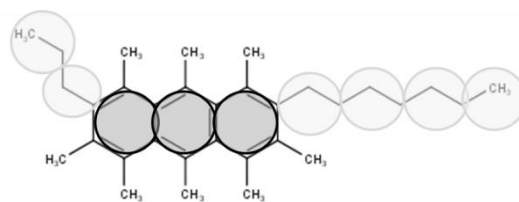
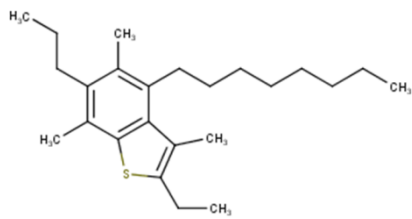
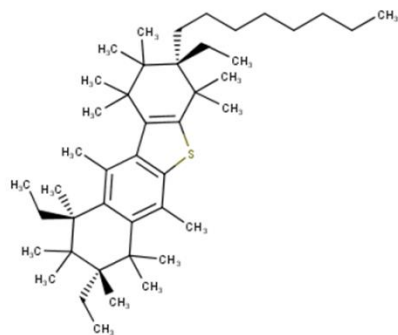
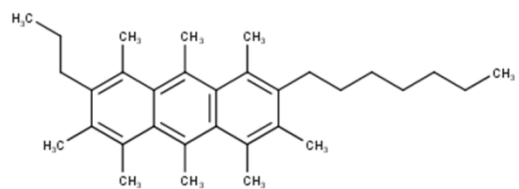
(5) / Aromatic / 51.81 %

(6) / Saturate / 8.4%

(7) / Aromatic / 5.02%

Table 30. Chemical structures and coarse-grained models of Representation 5 (R5) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.

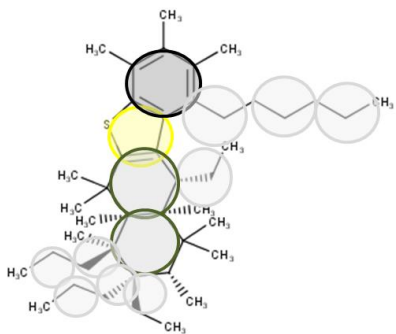
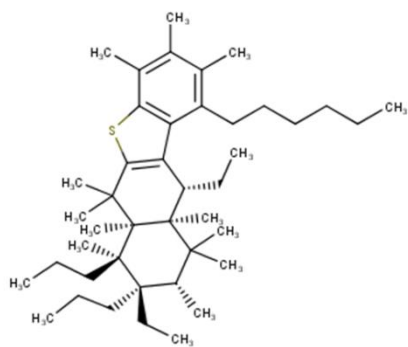
Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mole.
		(1) / Saturated / 7.27 %
		(2) / Aromatic / 5.87 %
		(3) / Resin / 6.42 %
		(4) / Saturated / 10.63 %
		(5) / Aromatic / 0.98 %



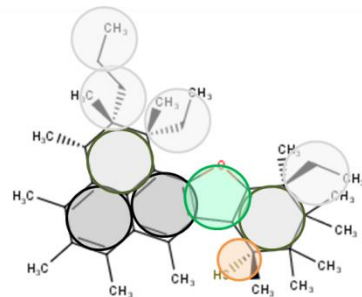
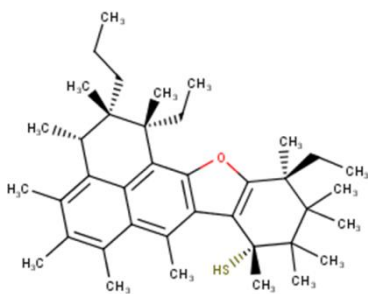
(6) / Aromatic / 8.92 %

(7) / Aromatic / 2.06 %

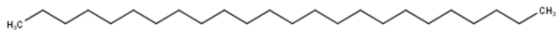
(8) / Aromatic / 3.89 %



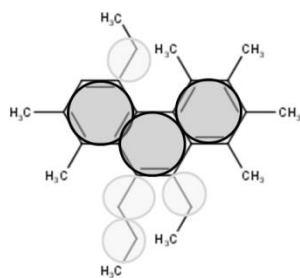
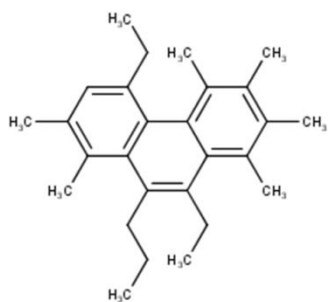
(9) / Aromatic/ 4.55 %



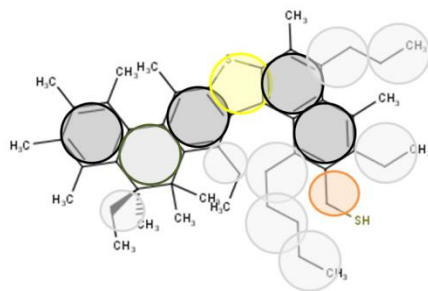
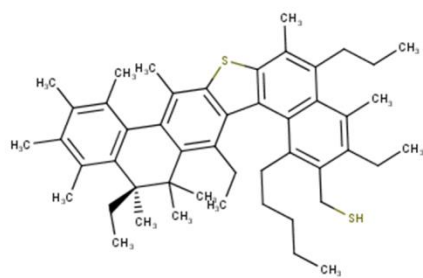
(10) / Aromatic/ 9.26 %



(11) / Saturated / 16.79 %

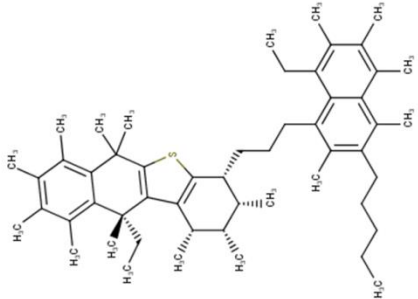

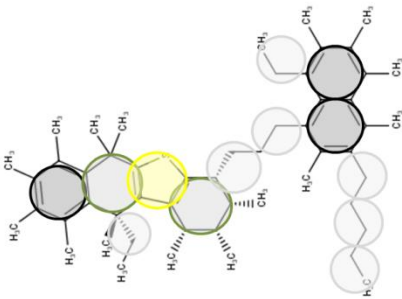

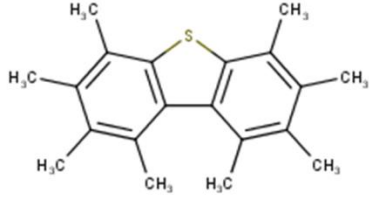
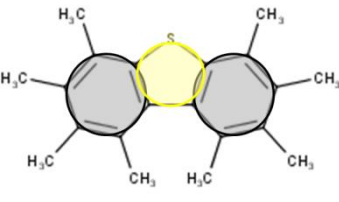
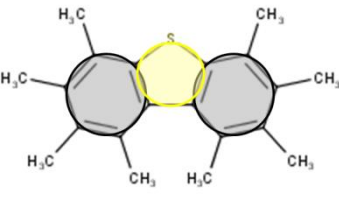

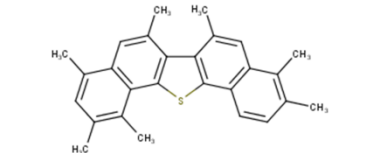
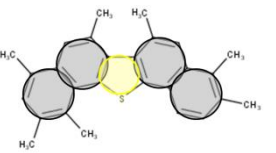
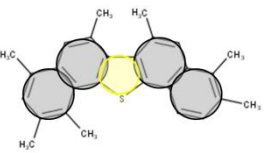

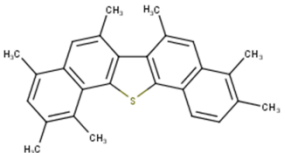
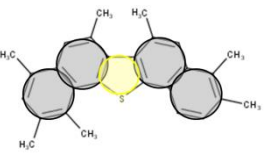
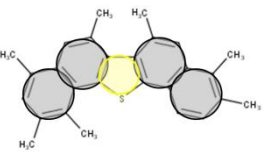



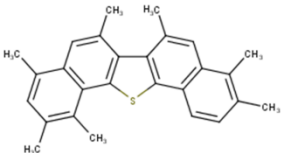
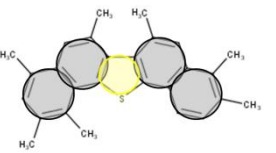
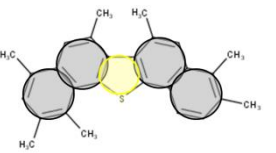









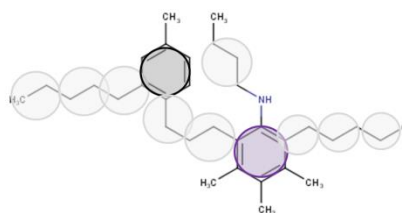
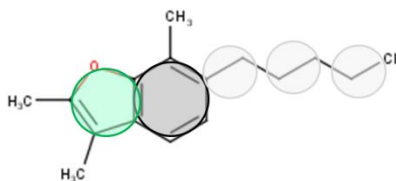
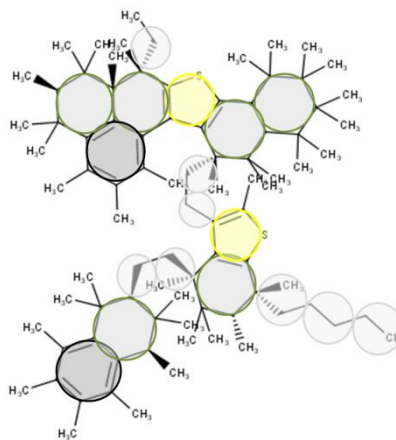
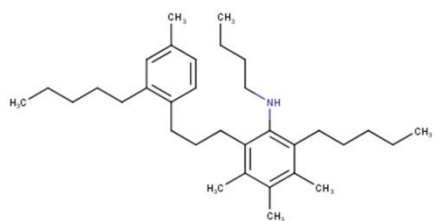
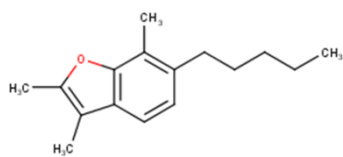
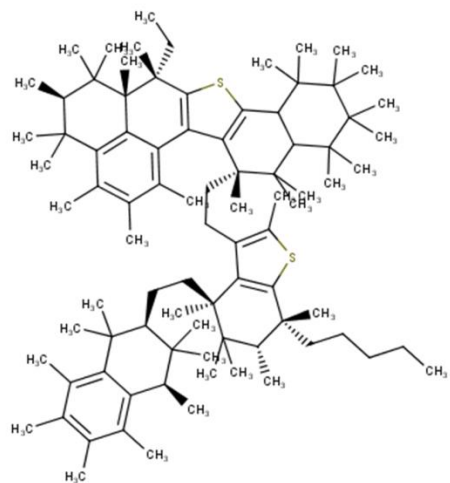
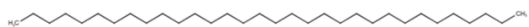
(12) / Aromatic / 18.40%



(13) / Resin / 4.92 %

Table 31. Chemical structures and coarse-grained models of Representation 6 (R6) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.

Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mole
 	 	(1) / Aromatic / 12.11 %
 	 	(2) / Saturated / 1.86 %
 	 	(3) / Aromatic / 23.95 %
 	   	(4) / Saturated / 8.34 %
 	 	(5) / Aromatic / 4.52 %
		(6) / Saturated / 0.54 %
		(7) / Saturated / 1.86 %
		(8) / Saturated / 0.54 %

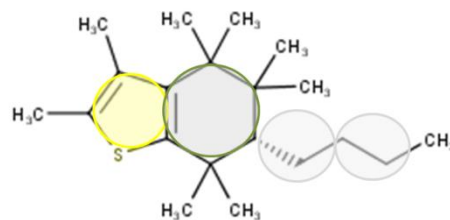
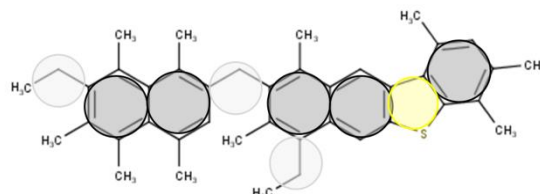
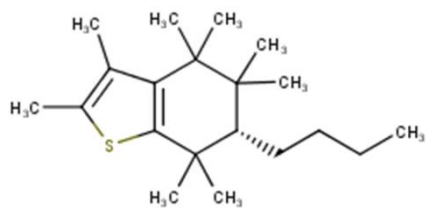
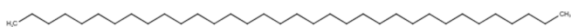
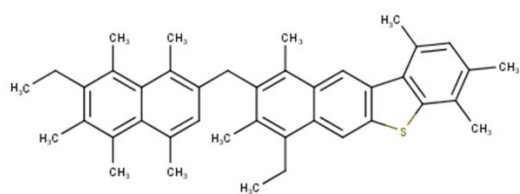


(9) / Saturate/ 8.40 %

(10) / Resin/ 4.03 %

(11) / Aromatic / 13.68 %

(12) / Resin / 8.38%

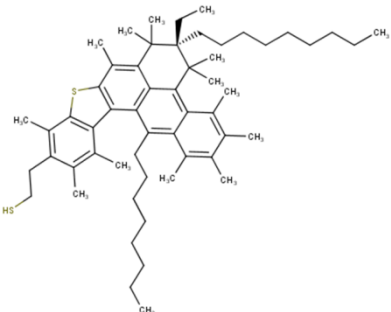
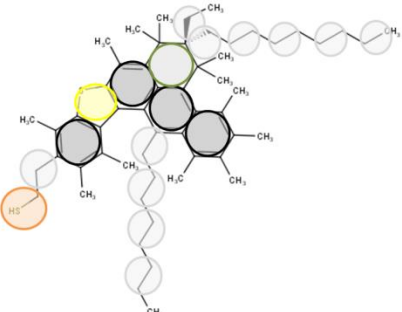
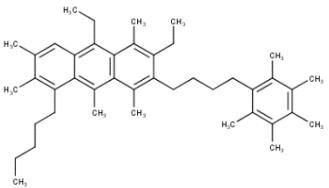
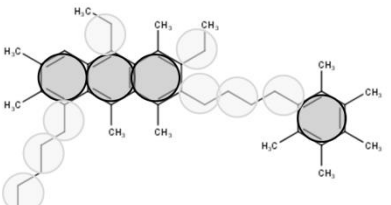
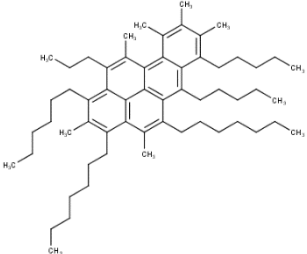
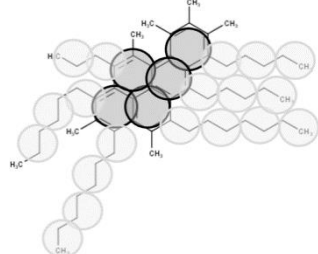
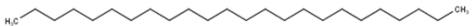



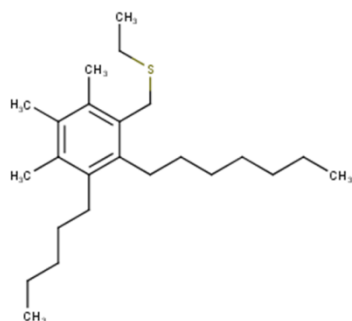
(13) / Resin / 3.13 %

(14) / Saturate / 6.52%

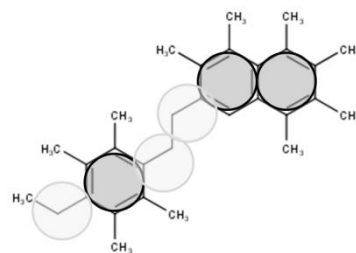
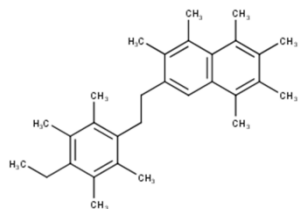
(15) / Aromatic / 2.13 %

Table 32. Chemical structures and coarse-grained models of Representation 7 (R7) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.

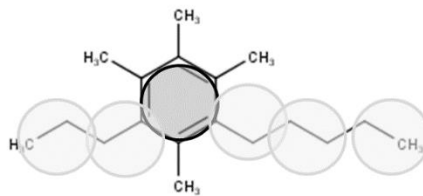
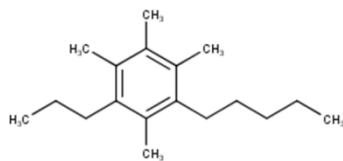
Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mole
		(1) / Aromatic / 0.99 %
		(2) / Aromatic / 13.57 %
		(3) / Aromatic / 7.92 %
		(4) / Saturate / 0.87 %



(5) / Aromatic/ 3.52 %



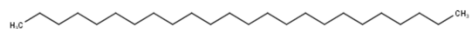
(6) / Aromatic/ 5.04 %



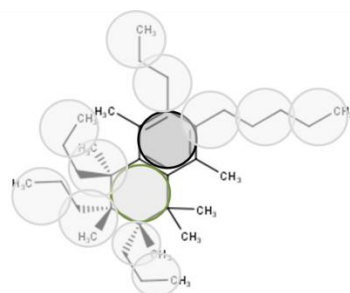
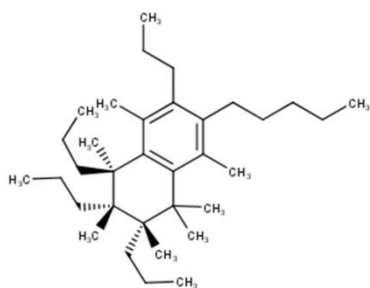
(7) / Saturate / 11.35 %



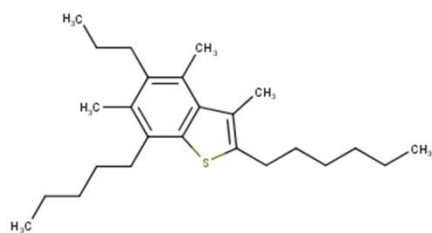
(8) / Saturate / 13.85 %



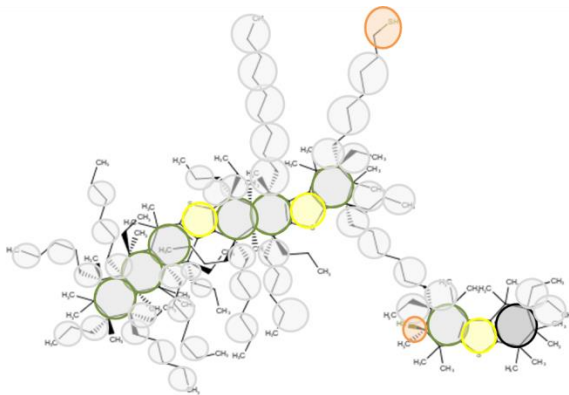
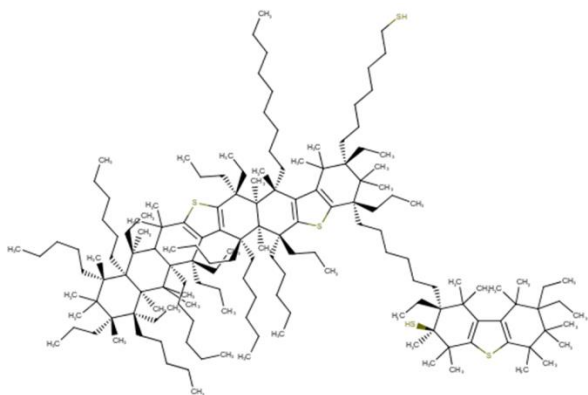
(9) / Saturate/ 0.87 %



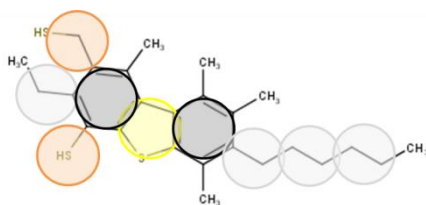
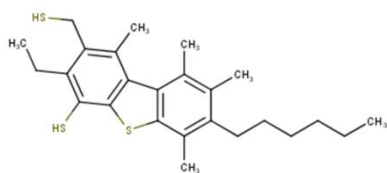
(10) / Saturate / 17.26 %



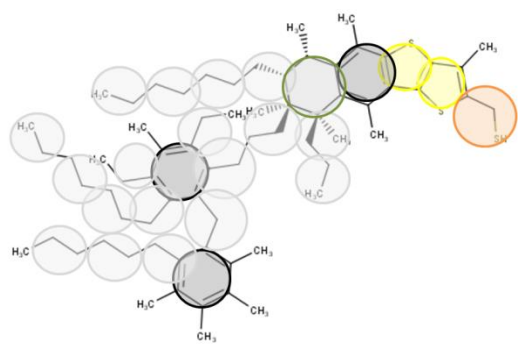
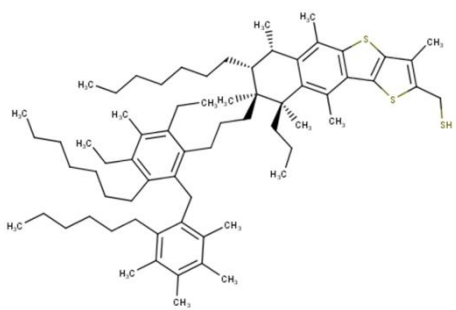
(11) / Aromatic / 10.84 %



(12) / Resin / 1.94%

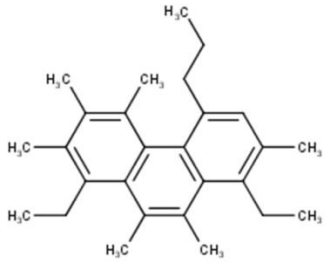
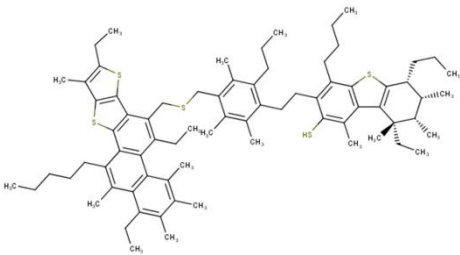
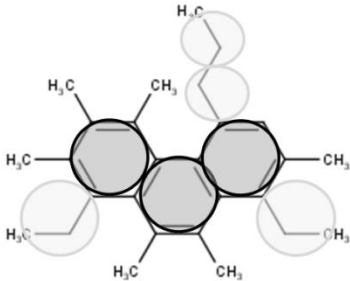
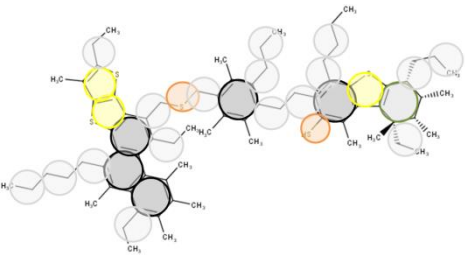


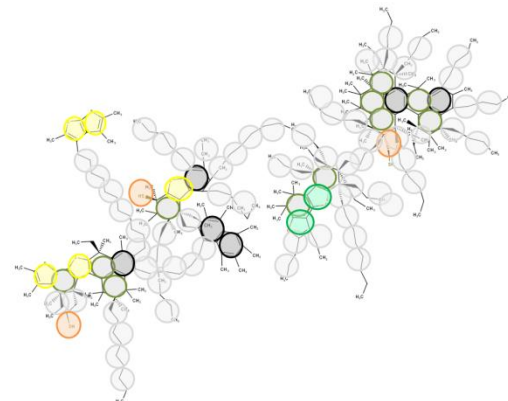
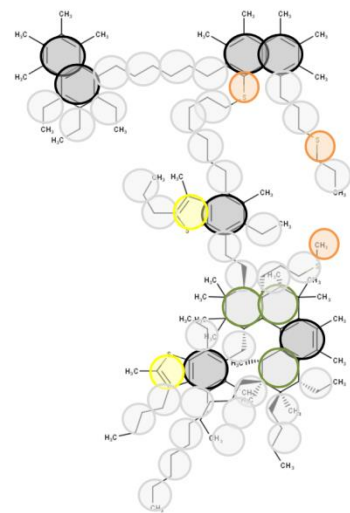
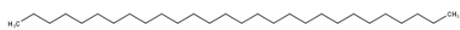
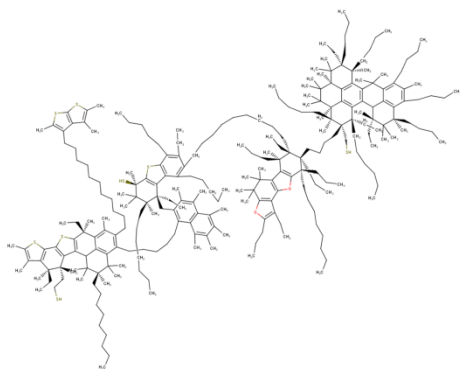
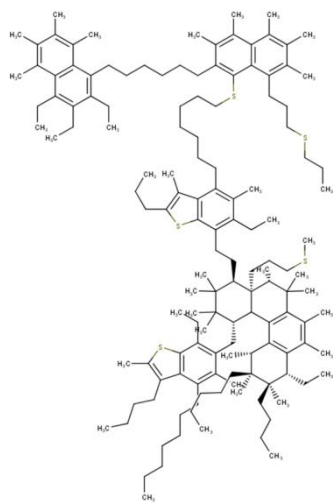
(13) / Aromatic / 4.9 %



(14) / Resin / 7.05%

Table 33. Chemical structures and coarse-grained models of Representation 8 (R8) of the heavy fraction for crude oil A1 with maximum 20 molecules. Mole fractions are given with respect to the heavy fraction.

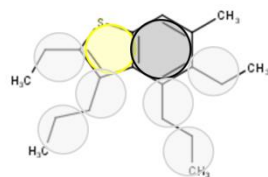
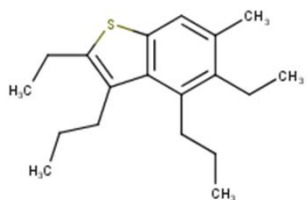
Molecular representation	Coarse-grained representation	(ID) / Chemical family / % mole
 	 	<p>(1)/ Aromatic/ 14.73 %</p> <p>(2) / Resin/ 0.23 %</p>



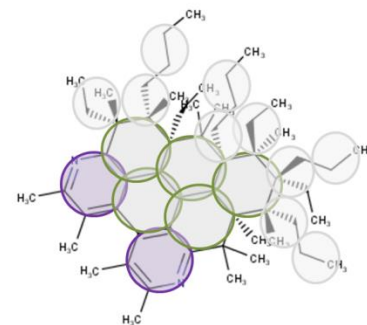
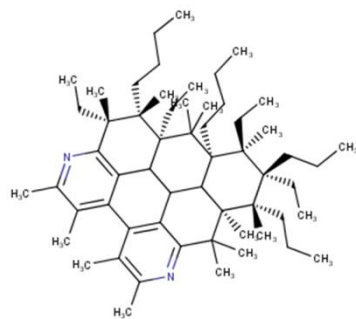
(3) / Resin / 2.46 %

(4) / Asphaltene / 0.23 %

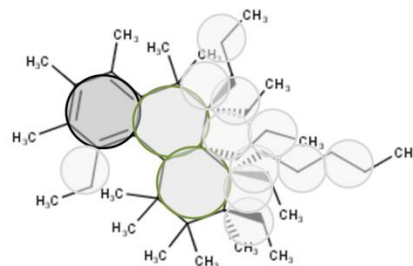
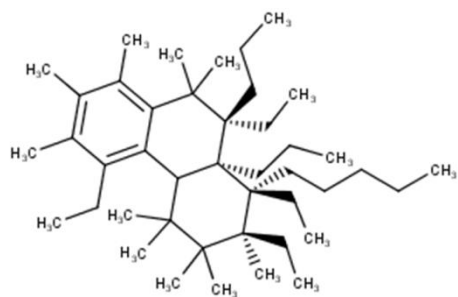
(5) / Saturate/ 11.1 %



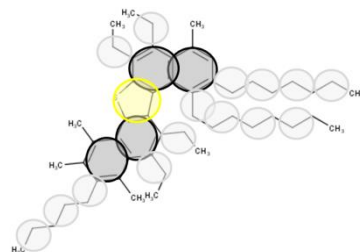
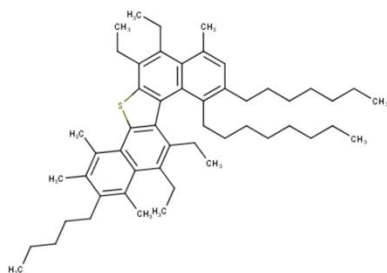
(6) / Aromatic/ 19.4 %



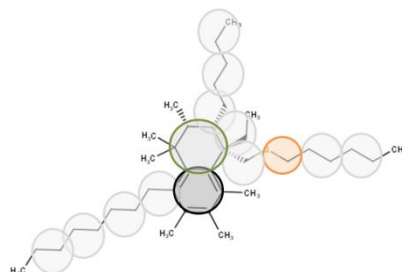
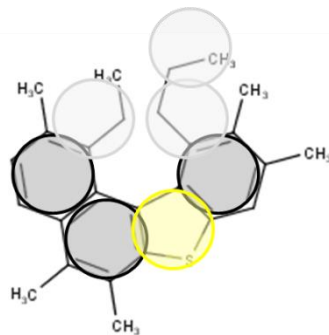
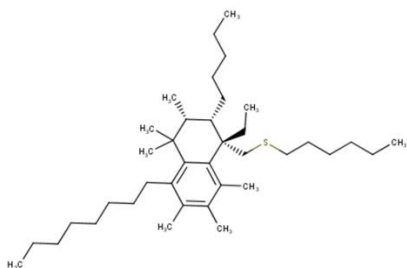
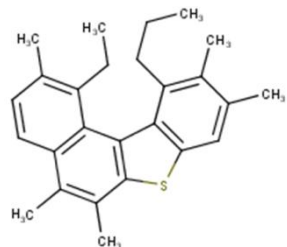
(7) / Resin / 7.14 %



(8) / Saturate / 10.02 %



(9) / Aromatic/ 8.71 %



(10) / Saturate / 7.47 %

(11) / Aromatic / 11.42 %

(12) / Saturate / 2.86 %

(13) / Saturate / 0.65 %

(14) / Aromatic / 3.55 %

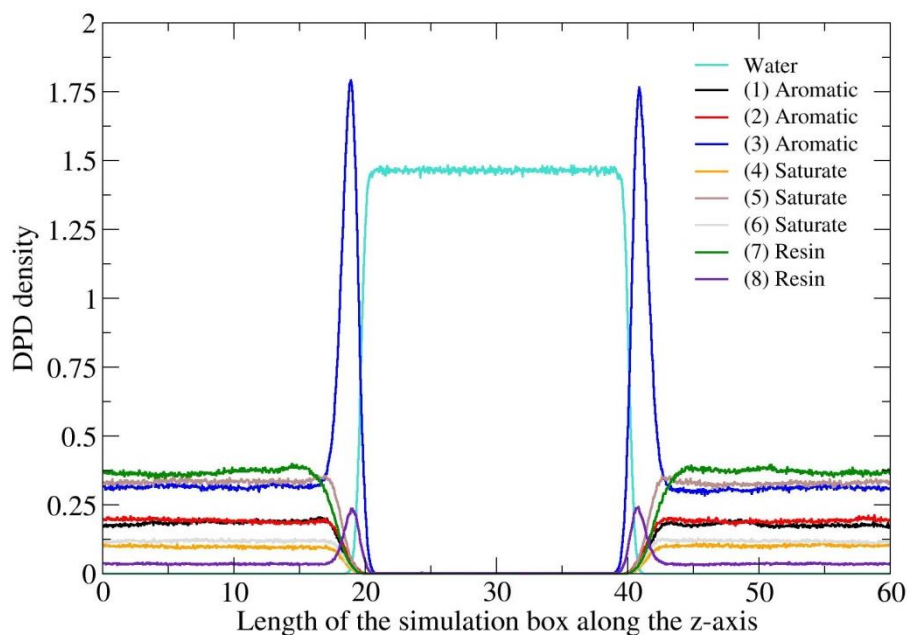


Figure 7. Density profile of crude oil/water A1 system (Representation 1). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

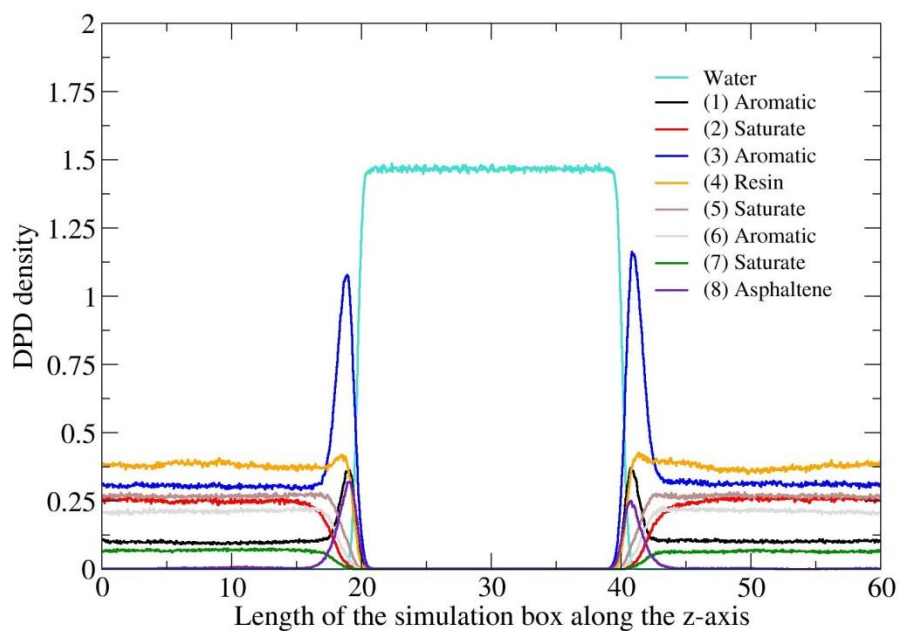


Figure 8. Density profile of crude oil/water A1 system (Representation 2). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

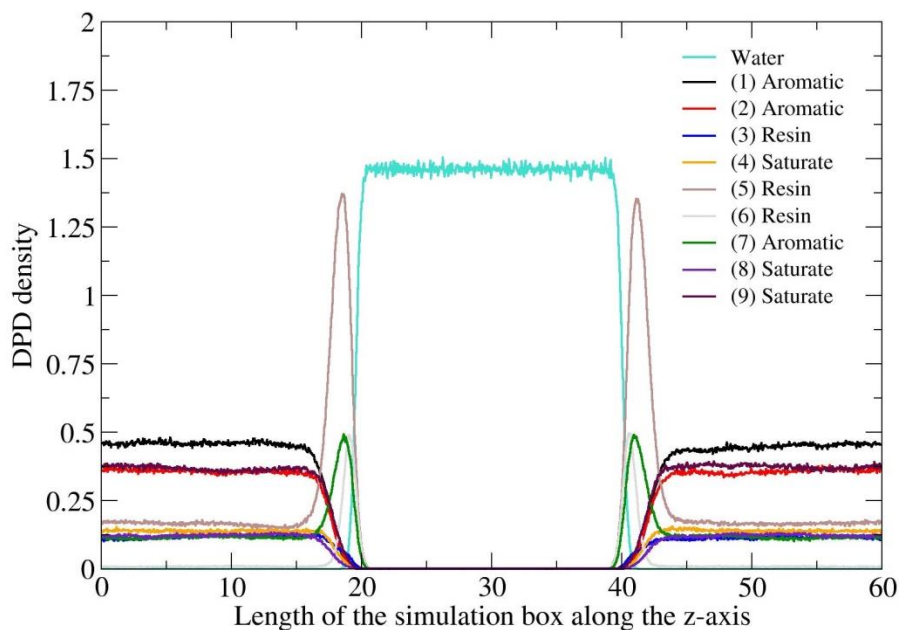


Figure 9. Density profile of crude oil/water A1 system (Representation 3). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

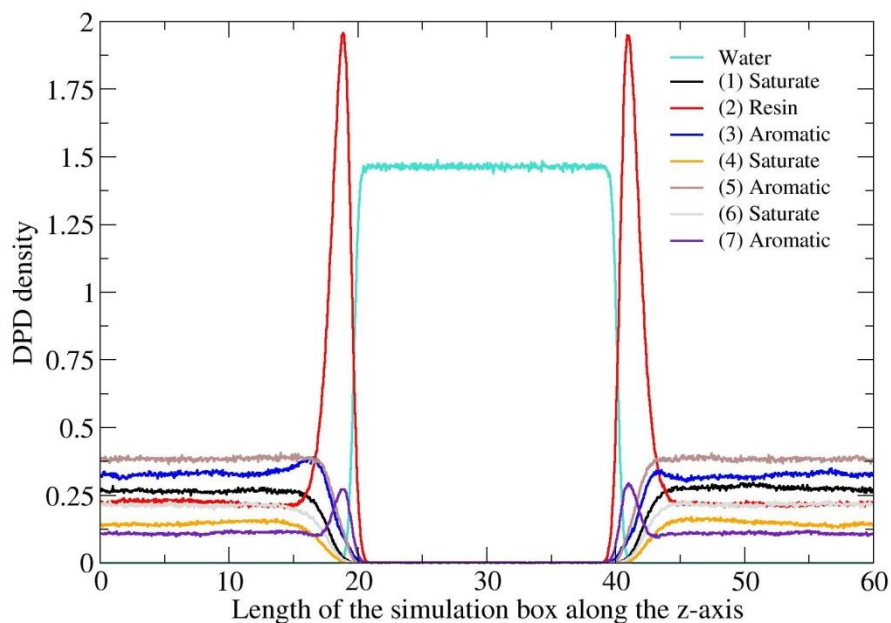


Figure 10. Density profile of crude oil/water A1 system (Representation 4). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

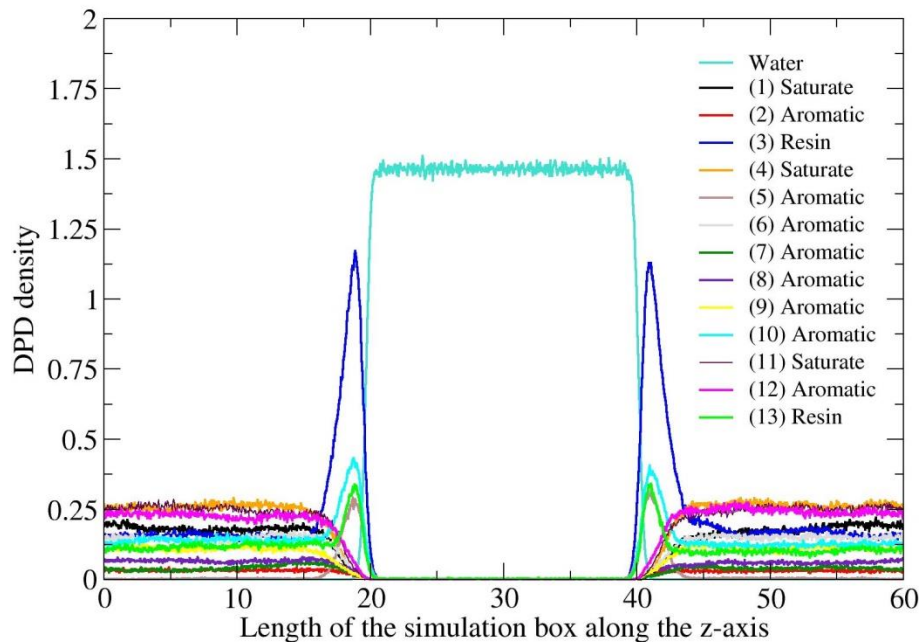


Figure 11. Density profile of crude oil/water A1 system with 20 molecules (Representation 5). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

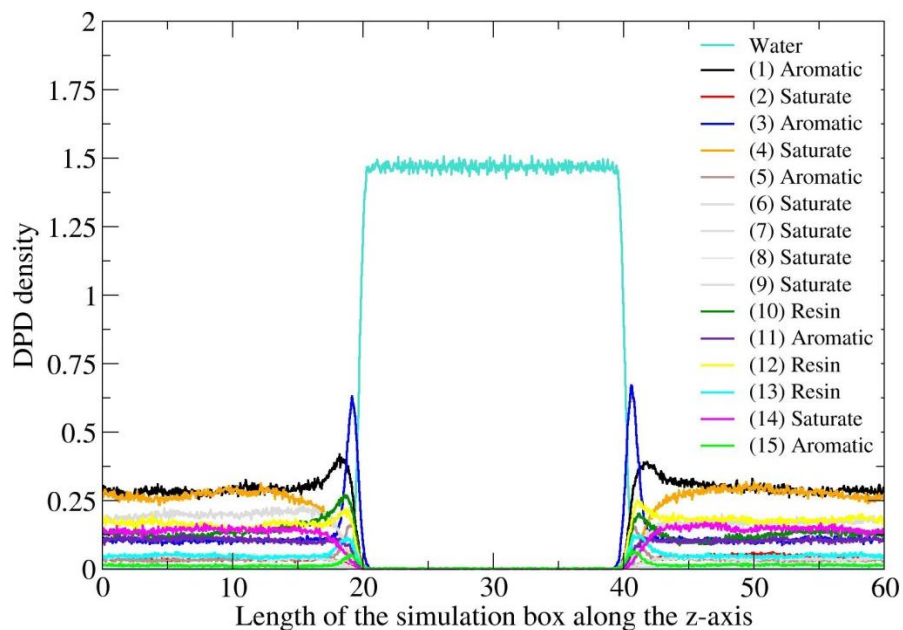


Figure 12. Density profile of crude oil/water A1 system with 20 molecules (Representation 6). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

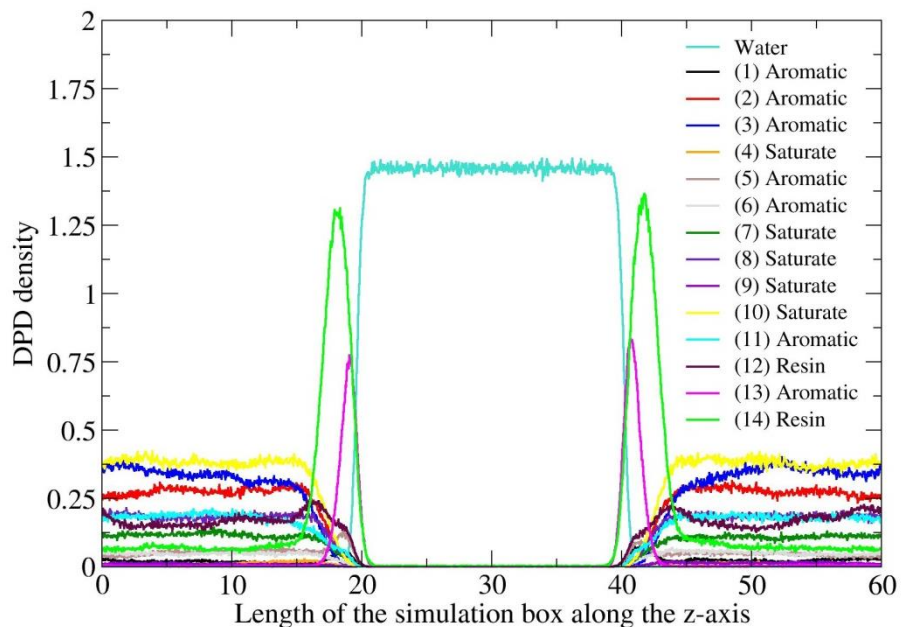


Figure 13. Density profile of crude oil/water A1 system with 20 molecules (Representation 7). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

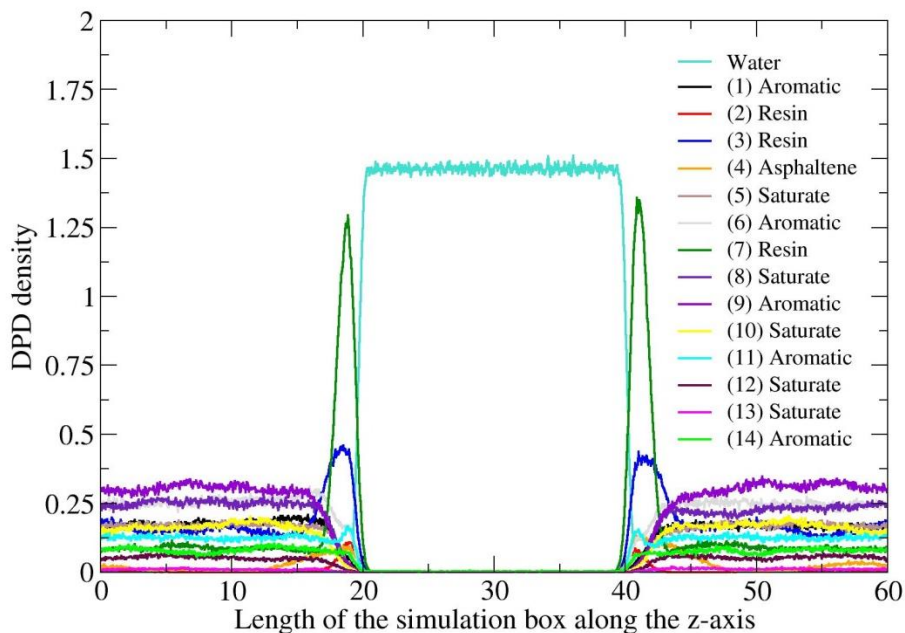


Figure 14. Density profile of crude oil/water A1 system with 20 molecules (Representation 8). Only molecules of heavy fraction are given. Molecules are indicated by ID number and the chemical family according to SARA analysis. Density of water has been halved.

4. References

- (1) James G. Speight. *The Chemistry and Technology of Petroleum*, 4th ed.; Fourth Edition, CRC Press/Taylor & Francis, 2006.
- (2) Drelich, J.; Fang, C.; White, C. L. *Measurement of Interfacial Tension in Fluid-Fluid Systems*; Encyclopedia of Surface and Colloid Science; Marcel Dekker, Inc., 2002.
- (3) 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. *Journal of chemical theory and computation* **2018**, *14* (8), 4438–4454. DOI: 10.1021/acs.jctc.8b00357.
- (4) Peng, D.-Y.; Robinson, D. B. A New Two-Constant Equation of State. *Ind. Eng. Chem. Fund.* **1976**, *15* (1), 59–64. DOI: 10.1021/i160057a011.
- (5) Soave, G. Equilibrium constants from a modified Redlich-Kwong equation of state. *Chemical Engineering Science* **1972**, *27* (6), 1197–1203. DOI: 10.1016/0009-2509(72)80096-4.
- (6) Chapman, W. G.; Gubbins, K. E.; Jackson, G.; Radosz, M. SAFT:Equation-of-state solution model for associating fluids. *Fluid Phase Equilibria* **1989**, *52*, 31–38. DOI: 10.1016/0378-3812(89)80308-5.
- (7) Chapman, W. G.; Gubbins, K. E.; Jackson, G.; Radosz, M. New reference equation of state for associating liquids. *Ind. Eng. Chem. Res.* **1990**, *29* (8), 1709–1721. DOI: 10.1021/ie00104a021.
- (8) Wertheim, M. S. Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations. *J Stat Phys* **1984**, *35* (1-2), 35–47. DOI: 10.1007/BF01017363.
- (9) Wertheim, M. S. Fluids of dimerizing hard spheres, and fluid mixtures of hard spheres and dispheres. *The Journal of Chemical Physics* **1986**, *85* (5), 2929–2936. DOI: 10.1063/1.451002.
- (10) Wilson, G. M. Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing. *J. Am. Chem. Soc.* **1964**, *86* (2), 127–130. DOI: 10.1021/ja01056a002.
- (11) Renon, H.; Prausnitz, J. M. Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE J.* **1968**, *14* (1), 135–144. DOI: 10.1002/aic.690140124.
- (12) Abrams, D. S.; Prausnitz, J. M. Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE J.* **1975**, *21* (1), 116–128. DOI: 10.1002/aic.690210115.
- (13) Fredenslund, A.; Gmehling, J.; Michelsen, M. L.; Rasmussen, P.; Prausnitz, J. M. Computerized Design of Multicomponent Distillation Columns Using the UNIFAC Group Contribution Method for Calculation of Activity Coefficients. *Ind. Eng. Chem. Proc. Des. Dev.* **1977**, *16* (4), 450–462. DOI: 10.1021/i260064a004.
- (14) ProSim. *Simulis Thermodynamics, Version 2.0.16*.
- (15) Nguyen, T.-B.; Hemptinne, J.-C. de; Creton, B.; Kontogeorgis, G. M. GC-PPC-SAFT Equation of State for VLE and LLE of Hydrocarbons and Oxygenated Compounds. Sensitivity Analysis. *Ind. Eng. Chem. Res.* **2013**, *52* (21), 7014–7029. DOI: 10.1021/ie3028069.
- (16) Chen, C.-C.; Song, Y. Solubility Modeling with a Nonrandom Two-Liquid Segment Activity Coefficient Model. *Ind. Eng. Chem. Res.* **2004**, *43* (26), 8354–8362. DOI: 10.1021/ie049463u.

- (17) Chen, C.-C.; Crafts, P. A. Correlation and Prediction of Drug Molecule Solubility in Mixed Solvent Systems with the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL-SAC) Model. *Ind. Eng. Chem. Res.* **2006**, *45* (13), 4816–4824. DOI: 10.1021/ie051326p.
- (18) Tung, H.-H.; Tabora, J.; Variankaval, N.; Bakken, D.; Chen, C.-C. Prediction of pharmaceutical solubility Via NRTL-SAC and COSMO-SAC. *Journal of Pharmaceutical Sciences* **2008**, *97* (5), 1813–1820. DOI: 10.1002/jps.21032.
- (19) Chen, C.-C. A segment-based local composition model for the gibbs energy of polymer solutions. *Fluid Phase Equilibria* **1993**, *83*, 301–312. DOI: 10.1016/0378-3812(93)87033-W.
- (20) Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. Group-contribution estimation of activity coefficients in nonideal liquid mixtures. *AIChE J.* **1975**, *21* (6), 1086–1099. DOI: 10.1002/aic.690210607.
- (21) Magnussen, T.; Rasmussen, P.; Fredenslund, A. UNIFAC parameter table for prediction of liquid-liquid equilibria. *Ind. Eng. Chem. Proc. Des. Dev.* **1981**, *20* (2), 331–339. DOI: 10.1021/i200013a024.
- (22) Dahl, S.; Fredenslund, A.; Rasmussen, P. The MHV2 model: A UNIFAC-based equation of state model for prediction of gas solubility and vapor-liquid equilibria at low and high pressures. *Ind. Eng. Chem. Res.* **1991**, *30* (8), 1936–1945. DOI: 10.1021/ie00056a041.
- (23) Dahl, S.; Macedo, E. A. The MHV2 model: A UNIFAC-based equation of state model for vapor-liquid and liquid-liquid equilibria of mixtures with strong electrolytes. *Ind. Eng. Chem. Res.* **1992**, *31* (4), 1195–1201. DOI: 10.1021/ie00004a033.
- (24) Alasiri, H.; Chapman, W. G. Dissipative particle dynamics (DPD) study of the interfacial tension for alkane/water systems by using COSMO-RS to calculate interaction parameters. *Journal of Molecular Liquids* **2017**, *246*, 131–139. DOI: 10.1016/j.molliq.2017.09.056.
- (25) Vishnyakov, A.; Lee, M.-T.; Neimark, A. V. Prediction of the Critical Micelle Concentration of Nonionic Surfactants by Dissipative Particle Dynamics Simulations. *The journal of physical chemistry letters* **2013**, *4* (5), 797–802. DOI: 10.1021/jz400066k.
- (26) Klamt, A.; Eckert, F. COSMO-RS: A novel and efficient method for the a priori prediction of thermophysical data of liquids. *Fluid Phase Equilibria* **2000**, *172* (1), 43–72. DOI: 10.1016/S0378-3812(00)00357-5.
- (27) Eckert, E.; Vaněk, T. New approach to the characterisation of petroleum mixtures used in the modelling of separation processes. *Computers & Chemical Engineering* **2005**, *30* (2), 343–356. DOI: 10.1016/j.compchemeng.2005.10.005.
- (28) Wauquier, J.-P. *Petroleum Refining 1*, Editions Technip; Editions TECHNIP, 1995.
- (29) Twu, C. H. An internally consistent correlation for predicting the critical properties and molecular weights of petroleum and coal-tar liquids. *Fluid Phase Equilibria*, 1984, 137–150.
- (30) 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.
- (31) Daubert, T. E.; Danner, R. P. *API Technical Data Book - Petroleum Refining*, 6th ed.; American Petroleum Institute (API), 1997.
- (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) 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.
- (34) Aquino, 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.
- (35) Riazi, M. R. *Characterization and properties of petroleum fractions*, 1st ed.; ASTM International, 2005.
- (36) Edmister, W. C. *Applied hydrocarbon thermodynamics, part 4: Compressibility Factors and Equations of state*, Petroleum Refiner; Petroleum Refiner, 1958.
- (37) 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.
- (38) Pereira de Oliveira, Luís; Verstraete, J. J.; Kolb, M. A Monte Carlo modeling methodology for the simulation of hydrotreating processes. *Chemical Engineering Journal* **2012**, 207-208, 94–102. DOI: 10.1016/j.ccej.2012.05.039.
- (39) Wiehe, I. A. The Pendant-Core Building Block Model of Petroleum Residua. *Energy Fuels* **1994**, 8 (3), 536–544. DOI: 10.1021/ef00045a003.
- (40) Wiehe, I. A. *Process chemistry of petroleum macromolecules*; Chemical industries, Vol. 121; CRC Press, 2008.
- (41) Hudebine, D. Reconstruction moléculaire de coupes pétrolières, Ecole Normale Supérieure de Lyon, Ph.D thesis, 2003.
- (42) Poling, B.E., Prausnitz, J.M., O’Connell, J.P. *The properties of gases and liquids*, 5th ed.; McGraw-Hill, 2000.
- (43) Fedors, R. F. A method for estimating both the solubility parameters and molar volumes of liquids. *Polym. Eng. Sci.* **1974**, 14 (2), 147–154. DOI: 10.1002/pen.760140211.
- (44) 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.
- (45) Goujon, F.; Dequidt, A.; Ghoufi, A.; Malfreyt, P. How Does the Surface Tension Depend on the Surface Area with Coarse-Grained Models? *Journal of chemical theory and computation* **2018**, 14 (5), 2644–2651. DOI: 10.1021/acs.jctc.8b00158. Published Online: Apr. 13, 2018.
- (46) Rowley, R. L.; Wilding, W. V.; Oscarson, J. L.; Zundel, N. A. *DIPPR 801*; Design Institute for Physical Property Research/AIChE, 2014.