



**HAL**  
open science

# Structural and thermodynamics properties of pure phase alkanes, monoamides and alkane/monoamide mixtures with an ab initio based force-field model

Abdelmounaim Failali, Eléonor Acher, Dominique Guillaumont, Valérie Vallet, Florent Réal

## ► To cite this version:

Abdelmounaim Failali, Eléonor Acher, Dominique Guillaumont, Valérie Vallet, Florent Réal. Structural and thermodynamics properties of pure phase alkanes, monoamides and alkane/monoamide mixtures with an ab initio based force-field model. *Journal of Molecular Liquids*, 2022, 363, pp.119797. 10.1016/j.molliq.2022.119797 . hal-03481112

**HAL Id: hal-03481112**

**<https://hal.science/hal-03481112>**

Submitted on 26 Nov 2023

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# Structural and thermodynamics properties of pure phase alkanes, monoamides and alkane/monoamide mixtures with an *ab initio* based force-field model

Abdelmounaim Failali,<sup>1,2</sup> Eléonor Acher,<sup>1</sup> Dominique Guillaumont,<sup>1</sup> Valérie Vallet,<sup>2</sup> and Florent Réal<sup>2,\*</sup>

<sup>1</sup>CEA, DES, ISEC, DMRC, Univ Montpellier, Marcoule, F-30207, Bagnols-sur-Ceze Cedex, France

<sup>2</sup>Univ. Lille CNRS, UMR 8523 – PhLAM – Physique des Lasers Atomes et Molécules, F-59000, Lille, France

(Dated: Received: July 19, 2022)

A polarizable force-field (FF) model for short- and long-alkane chains and amide derivatives was constructed based solely on accurate quantum chemical (QC) calculations. First, the FF model accuracy was accessed by performing molecular dynamics (MD) simulations to calculate liquid-phase thermodynamic and structural properties for alkanes, for which experimental data are available. Second, The FF was then used to perform molecular dynamics simulations to calculate thermodynamic, structural and excess properties of monoamide/dodecane mixtures, namely DEHiBA/dodecane and DEHBA/dodecane. Aggregation phenomena appear for both types of mixtures and monoamide pure phases. A detailed structural analysis revealed, at small monoamide mole fraction the formation of dimers, while trimerization at larger monoamide concentrations and in their pure phases. Analysis of the relative orientation of the dimers have also been performed and showed a small difference for both phases.

## INTRODUCTION

Alkane and amide derivatives molecules are used and quite important in many research areas, such as biology, medicine, nuclear field and petrology. In these latter, the structural, thermodynamics and transport properties are of interest for the design and study of artificial or biological membranes and also play a major role in the recovery and refining of crude oil [1–8].

In the nuclear field, amide derivatives can be used as extracting molecules for actinide ions for nuclear fuel solutions as alternatives to the TBP molecules, with advantages highlighted in literature [9, 10]. Alkanes such as dodecane, TPH (Tetra Propylene Hydrogenated), kerosene, and so forth [11, 12] are hydrocarbons suitable as solvents to dilute TBP and/or amide derivatives and to achieve the extraction of U and Pu from the irradiated nuclear fuel.

One way to obtain insights into the properties of such solutions (alkanes + amides derivatives) at the molecular level, involves the use of atomistic molecular dynamics simulations. In this regard, several force fields have been developed over the years, such as CHARMM, GROMOS, MM4, OPLS and AMBER [13–17]. Generally, the hydrocarbons (alkanes) parameters are used to describe alkyl chains regardless of whether they are amide, acid, amine chains or lipid, peptide or some protein tails in biology applications. Most of these force-fields (FFs) have been derived focusing on individual aspects and based on the reproduction of experimental data, such as enthalpies of vaporization, vapor pressures and densities. However, the FF parameters are not always transferable to all the molecular group series, and they are reoptimized before use in most cases. As a result, different versions of AMBER, CHARMM, OPLS FFs for hydrocarbons have been developed over the years for a better description of the macroscopic properties [18–20]. For example, the first

version of OPLS FF for hydrocarbons (labeled OPLS-AA) is quite successful for short hydrocarbons but not for long [17], some studies in literature have reported the gel-phase formation at room temperature for long-chain hydrocarbons of more than eight carbons (including dodecane) using the OPLS-AA force field [21, 22]. This comes from the fact that they used experimental data measured for small molecules to develop such FF, and hence inducing a lack of physical meaning of the interactions for longer chains. Therefore, Siu *et al.* have reoptimized the OPLS-AA parameter set for long hydrocarbons, termed L-OPLS [19]. Generally, FF parameters obtained for small molecules do not always represent the real interactions between atoms, since the goal was to reproduce the reference experimental data in an average way (optimized empirically to match liquid properties). In summary, in the literature exists a zoo of FFs, the latter are mostly derived for individual molecules and specific cases, which makes the prediction of new features of newly designed molecules challenging and questionable.

In the context of nuclear fuel reprocessing, the most studied extractant are phosphorus molecules given their huge application in the field; more precisely, the TBP molecule. Over the years, many FFs have been developed (polarizable and non polarizable) to model the behavior of such molecule whether in binary mixtures with alkanes or in full phases with uranyl and nitrates ions [23–25]. As for amide derivatives (shown in Figure 1), most available FFs have been derived for small molecules and then combined with the hydrocarbons parameters to describe the alkyl chains of the long amides [26, 27]. However, as in the case of hydrocarbons, there are several FFs available in the literature but they are not always transferable and accurate enough for direct use. Hence the re-optimization of the FF parameters is often required [28–30]. For amide derivatives, it was shown that the transferability problems may arise *i)* from the fact that amides

are polar molecules and thus FFs describing this kind of molecules should include these effects explicitly [30] *ii*) and, of course, from the parameters used to describe the alkyl chains.

Nowadays, the use of *ab initio* data to parameterize FFs has become increasingly common, since a solid physical/mathematical foundation provides a better understanding of the physics and chemistry of the systems to be investigated. Moreover, most transferability problems encountered for force fields with empirical parameters are related to the reference data used. If the systems used in the parameterization process significantly differ from the ones being investigated and/or if the amount of data set used for parameterization is small, limited to certain kinds of data at restricted temperature and pressure conditions, the parameters may not be as trustworthy. The available experimental data are often limited to certain molecules at specific experimental conditions, while, quantum chemistry methods in combination with the availability of significant computational resources offer the possibility to generate a large number of quantum chemical reference data essential to the development of force fields (atomic/molecular data, dissociation curves, interaction energies, etc.).

In order to be confident that the behavior observed in molecular dynamics (MD) simulations is representative of real dynamical systems, the selection of an accurate force field is essential. Hence, our group is working on the development of a new class of polarizable *ab initio*-based force fields with the right balance between accuracy and efficiency [31–33]. Herein, we have developed a new set of FF for alkanes and amide derivatives based solely on quantum chemistry calculations. Since alkanes are known to be a non-polar molecules, polarization effects were neglected and only Coulomb, repulsion and dispersion interactions were considered. However, it should be mentioned that the alkanes do not induce polarization but that they are polarizable in the model. As for the amides, polar molecules for which polarization forces play an important role, polarization was incorporated alongside Coulomb, repulsion and dispersion interactions to ensure the transferability of the parameters for longer amides, as well as the correct description of the intermolecular interactions.

First, parameters for alkane and amide model molecules have been derived using *ab initio* quantum chemical data. Then, we have combined the two sets of parameters to describe the large amides, for which the alkyl chains have been described with the derived alkane parameters. The newly proposed parameter sets were validated on physical properties of interest, namely density, heat of vaporization as well as on the distributions of trans and gauche conformation for alkanes. The simulations on the alkane/monoamide mixtures allow us to predict the densities and excess quantities that, for the density and excess enthalpy can be compared to very re-

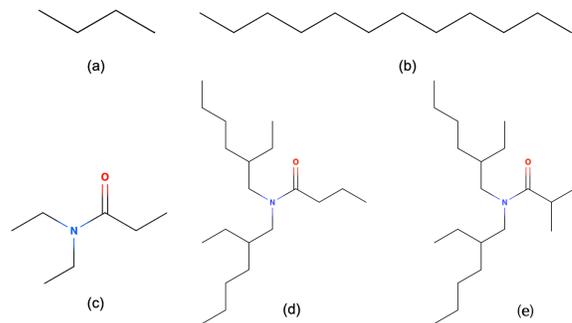


FIG. 1: Schematic representations of (a) butane, (b) dodecane, (c) N,N-Diethylpropionamide (DEPA), (d) N,N-di-2-ethylhexyl-butylamide (DEHBA) and (e) di-2-ethylhexylisobutyramide (DEHiBA) molecules.

cent experimental data [34].

## COMPUTATIONAL AND SIMULATION DETAILS

### Force-field potential model

The total potential energy  $\Delta U$  corresponding to the force field presented in this work describes the intra and intermolecular in alkanes and amide derivatives. It is a sum of six energy components:

$$\Delta U = U_{bonds} + U_{angles} + U_{dihedrals} + U_{qq'} + U_{Buck} + U_{pol}, \quad (1)$$

corresponding to bond, angle, torsion, electrostatic, Buckingham and polarization potentials, respectively.

The covalent interactions between atoms are modeled using harmonic bond stretching ( $U_{bonds}$ ) and angle bending ( $U_{angles}$ ) terms, while rotations around a bond are described by anharmonic 4-body torsional terms ( $U_{dihedrals}$ ). These three terms are developed in the following equations:

$$U_{bonds} = \sum_{bonds} \frac{1}{2} K_r (r - r_0)^2, \quad (2)$$

$$U_{angles} = \sum_{angles} \frac{1}{2} K_\theta (\theta - \theta_0)^2, \quad (3)$$

$$U_{dihedrals} = \sum_{dihedrals} \frac{1}{2} \sum_n K_n (1 + \cos(n\phi + \phi_0)). \quad (4)$$

The bond length, equilibrium bond length, and the bond stretching constants are given as  $r$ ,  $r_0$ , and  $K_r$ , respectively. The bond angle  $\theta$  has its parameters defined

similarly. The dihedral potential for every dihedral angle  $\phi$  is a sum over a number of  $n$  Fourier components. Each  $n$  value is associated with an amplitude,  $K_n$ , and a phase shift  $\phi_0$ . As for the 1-4 interactions, in our MD code, instead of using a scaling factor, they are treated with a correcting potential ( $a_{ij} \exp(-b_{ij}r_{ij})$ ), where  $a_{ij}$  and  $b_{ij}$  are adjustable parameters (parameters are available in the ESI Table S1.).

The set of bonds, angles and dihedral parameters were obtained by minimizing the difference of the relative total potential energies calculated with the force field with respect to *ab initio* calculations. The comparisons of energy profiles obtained from *ab initio* and force-field calculations are shown in Figure 3 and in the ESI Figure S1 for the longer chain n-dodecane.

The non-bonded interactions are described by a sum of *i*) Coulombic interactions between atom-centered point charges,  $U_{qq'}$ , *ii*) a Buckingham potential accounting for repulsion and dispersion interactions,  $U_{Buck}$  and lastly *iii*) a polarization term.

For a system of  $N$  atoms, the additive terms  $U_{qq'}$  and  $U_{Buck}$  are defined as:

$$U_{qq'} + U_{Buck} = \sum_{i=1}^N \sum_{j,j>i}^N \left[ \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} + A_{ij} \exp(-B_{ij}r_{ij}) - \frac{C_{ij}}{r_{ij}^6} \right], \quad (5)$$

where  $r_{ij}$  is the distance between atoms  $i$  and  $j$ ,  $q_i$  are the static point charges located on the atomic centers and obtained by QC calculations (See section ).  $A_{ij}$ ,  $B_{ij}$  and  $C_{ij}$  are adjustable parameters.

Lastly, polarization effects are incorporated with the induced dipole model, described by means of a set of induced dipole moments  $\{\mu_i\}_{i=1, N_\mu}$  located on a subset of  $N_\mu$  polarizable atomic centers. They obey

$$\mu_i = \alpha_i \left( \mathbf{E}_i^q + \sum_{j=1, j \neq i}^{N_\mu} \mathbf{T}_{ij} \mu_j \right). \quad (7)$$

Here,  $\alpha_i$  is the isotropic polarizability of the polarizable atom  $i$ , computed as described in Section -.  $\mathbf{E}_i^q$  is the electric field generated on  $i$  by the surrounding static charges  $q_j$ , and  $\mathbf{T}_{ij}$  is the dipolar tensor. They both include short-range Thole's like damping functions [35, 36] with an adjustable damping parameter  $a_{ij}$ , as described in further details the ESI.

For most of the FFs available in the literature, the latter polarization contribution is neglected for the alkanes. However, for the amide derivatives, since they are polar molecules, it should be explicitly included to ensure the transferability of the FF parameters in the series of amide derivatives. In this work, all the  $N_\mu$  non-hydrogen atoms are considered as polarizable centers, i.e., a single point polarizability is located on each non-hydrogen atomic center.

All the alkane/amide derivatives force-field parameters optimized with the procedures described in the coming section are listed in Table S1 of the ESI.

### Force-field development

The development of any new FFs consists on three main steps, besides the choice of a physically meaningful functional form for the potential energy. First, the choice and preparation of reference data, second, the adjustment of parameters to reference data and last the validation of the FF by computing macroscopic physical properties. Herein, all FF parameters have been derived based solely on QC calculations.

#### QC calculations for reference data

Before going any further, the choice of the QC level to calculate the reference data used for the development of any FF is fundamental; it has to ensure that the different interactions are accurately treated. For instance, for the alkanes, the accurate description of the interactions impacts the computed heats of vaporization that are expected to come out within chemical error,  $1 \text{ kcal} \cdot \text{mol}^{-1}$ , of the experimental data. Moreover, the dihedral potentials are crucial for reproducing the (temperature-dependent) fractions of trans and gauche isomers, as the energy difference between the two conformations is small (about  $0.6 \text{ kcal} \cdot \text{mol}^{-1}$ ) and must be accurately described for alkanes (See Figure 3).

In this work, all the structural and energetic data needed for the parametrization of our force field, were calculated using the Molpro quantum chemistry package [37] at the MP2 level (the Møller-Plesset Perturbation Theory) [38] with correlation consistent aug-cc-pVTZ basis sets by Dunning [39, 40]. Only the 1s core orbitals of C, N and O were kept frozen. This level of theory was chosen based on two arguments (1) the geometries of simple alkanes are known to be less sensitive to the size of the basis set than the energies themselves [41] (2) by performing a benchmarking of MP2 interaction energies with respect to Coupled Cluster "CCSD(T)" [42] "the gold standard" (see Supporting Information). In conclusion, the MP2 approach proved to provide a good compromise between the computational cost and the accuracy of the computed interaction energies. In the present work, electrostatic parameters and atomic polarizabilities were determined in a first stage, and the Buckingham parameters were optimized in a second stage (i.e., the electrostatic parameters were kept fixed while optimizing the Buckingham parameters).

Partial charges are calculated with CM5 method (Charge Model 5, latest update of the CMx series), a method developed by Cramer, Truhlar, and co-

workers [43]. It uses the charges obtained from a Hirshfeld population analysis (of a wave function obtained with density functional calculation) as a starting point. The charges are then varied based on some specific parameters, derived originally by fitting to gas-phase dipole moments of several molecular structures. Jorgensen *et al.*, developers of the OPLS-AA FF series [44], stated that CM5 charges yielded the best agreement with experiments in pure liquid simulations, with the extra advantage of being essentially basis set independent [43].

As for atomic polarizabilities, in literature, there are several procedures to decompose the molecular polarizability into atomic polarizabilities [45]. They differ on whether molecular polarizability was obtained from experimental refractive indices or from QC. In this work, we have opted for the method proposed by Marenich *et al.* [46] for partitioning the molecular polarizability into atomic contributions by the use of Hirshfeld population analysis [47], involving the numerical differentiation of the dipole moments computed for different values of the applied external electric field.

Lastly, the repulsion and van der Waals dispersion parameters have been derived through a systematic potential energy surface exploration, we produce a large set of reference data, interactions energies for several different dimer relative orientations and distances, starting from ethane to *n*-dodecane for alkanes and from N,N-DiEthyl-PropanAmide (DEPA) and N,N-diethyl-2-methyl-PropAmide (DEMPA), plus dimers of alkane-amide molecules (see Figure S2-S3-S4 in ESI). We have adjusted the FF parameters around the minimum and the repulsion wall up to  $15 \text{ kcal} \cdot \text{mol}^{-1}$ . A total of 57 relative orientations and conformations were considered. For the different dimer orientations, the MP2 interaction energies (*IE*) were computed using the super-molecule approach taking into account the basis set superposition error (BSSE) with the counterpoise method [48]:

$$IE = E_{AB}(AB) - E_{AB}(A) - E_{AB}(B), \quad (8)$$

where,  $E_{AB}(AB)$  is the energy of their interacting assembly (dimer).  $E_{AB}(A)$  and  $E_{AB}(B)$  denote the total energies of monomers A and B, respectively, computed with the dimer AB basis sets.

#### FF parameterization procedure

The parameterization strategy used to derive the FF parameters for both the bonded and non-bonded potentials for alkanes and amides, is based solely on quantum chemistry calculations, often labeled "Bottom-Up", meaning that the parameters are derived to match the computed atomic/molecular-scale data (atomic/molecular data, dissociation curves, interaction energies, etc.).

TABLE I: Atom types and their definitions in the present work

No.	Atom Type	Description
1	CT3	Alkane methyl C $-\text{C}-\text{CH}_3$
2	CT2	Alkane methylene C $-\text{C}-\text{CH}_2-\text{C}$
3	CT1	Alkane branched C $-\text{C}-\text{CH}-\text{C}$
4	CT2-N	Methylene C bonded to amide N
5	HN	H atom bonded to CT2-N
6	HA	Aliphatic H atom (for alkane and amides)
7	N	Amide N atom
8	O	Amide carbonyl O atom
9	C	Amide carbonyl C atom

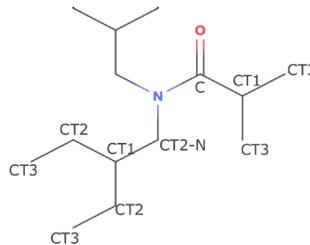


FIG. 2: Illustration of atom type used in the present work

Table I summarizes the atom types used in this work (See also Figure 2). We define HA as the hydrogens in the aliphatic compounds and HN the hydrogen atoms bonded to the specific carbon atom CT2-N. For carbons, we introduce four types: methyl carbon (CT3), methylene carbon (CT2), alkane branched carbon (CT1) and (CT2-N) for methylene carbon bonded to amide nitrogen atom. The oxygen of the carbonyl group is denoted O, the carbon atom double-bonded to an oxygen atom (C), the nitrogen atom of the amide function (N).

In most available FFs in the literature, bond stretching and angle bending are described using harmonic potentials that contain the bond and angle force constants and equilibrium bond lengths and angles. These equilibria are taken from the optimized geometries of the molecules, and the motion force constants are derived by performing scans over the bond/angle of interest, considering how the energy of a bond/angle changes with its length/angle. In this work, we have compared our parameters with that proposed by the CHARMM and Amber FFs [20, 49]. It turned out that for alkanes, the CHARMM FF parameters matched very well the QC scans, and could be taken without further refinements (see ESI).

As for repulsion and van der Waals parameters, our methodology consists in, first, computing interactions energies using the super molecule approach (Eq. 8) for the different dimer relative orientations and distances (See Figures S2-S3-S4-S5 in ESI for more details). Second, we have used the PEST optimization code [50] to refine the parameters, by adjusting the interaction energies com-

puted with the FF model to the MP2 QC values.

In order to later speed up the MD simulations and ease the parametrization process, Coulomb, polarization, repulsion and dispersion contributions were only considered for interactions between carbon, nitrogen, oxygen atoms. As for all interactions that involve hydrogen atoms, only the Coulomb and repulsion terms were kept. The optimized parameters for used in this work are reported in the ESI (Table S1).

The dihedral parameters are chronologically the last ones to be derived, since torsion angle motions embed contributions from both the non-bonded (van der Waals, electrostatic and 1-4 interactions) terms, as well as angle bendings. The torsional parameters are therefore intimately coupled to the non-bonded and bonded parameters. In this work, complete QC scans of the dihedral torsions were performed as a sequence of constrained optimization in which the torsion angle is varied in steps of  $10^\circ$ . For alkanes, all possible torsions CTx-CT2-CT2-CTx with  $x = 2, 3$  were scanned and resulted in quite similar torsion *ab initio* energy profiles (difference between maxima  $\leq 0.2 \text{ kcal} \cdot \text{mol}^{-1}$  as shown in Figure 3). For most available FFs in literature, the same torsional parameters are used for both torsion involving terminal methyl and middle torsion. In this work, since we differentiate the non-bonded parameters of CT3 from those of CT2, and since the dihedral parameters do dependent on the non-bonded ones, we have derived parameters specific to each dihedral torsion (CT3-CT2-CT2-CT2, CT3-CT2-CT2-CT3 and CT2-CT2-CT2-CT2).

Figure 3 illustrates the almost perfect superposition of the classical and QC torsion curves, making us confident to explore further the different trans-gauche populations of the various species in the bulk phase. The same approach was adopted to derive torsion parameters for amide derivatives.

Lastly, in order to assess the reliability of the proposed classical model one needs to confront macroscopic data based on MD simulations to experimental condensed phase properties, when available. In particular, for the pure system the density and heat of vaporization will be compared.

### Molecular dynamics simulation details and liquid-phase properties

+

All MD simulations were carried out with the Polaris-MD code [51], and considering periodic boundary conditions; first, in the isothermal-isobaric ensemble (NPT) to equilibrate the systems, and then the production part in the constant-volume ensemble (NVT). The temperature was maintained by a Nose-Hoover thermostat and the system pressure isotropically by an Andersen's barostat.

Verlet leapfrog scheme was employed with a 1 fs integration time step of the dynamic equations of motion. The equilibration period of our systems varies between 5 ns for pure phases (alkanes, monoamides) and 8 ns for alkanes/monoamides mixtures, with a temperature scaling interval each 10 steps, the equilibration period was fixed based on analysis of the time evolution of density, energy and structural analysis (see Figure S6 in the ESI). The production run is about 20 ns, which was enough considering the size of the systems that we simulated. The simulation box size impact on ligands properties have also been analyzed and was shown to be minor (see ESI Table S4). More details about the systems are available in Table II. The C-H structural parameters were constrained to their gas-phase equilibrium values thanks to the RATTLE algorithm (the convergence criteria was set to  $1 \times 10^{-6} \text{ \AA}$ ) [52]. The fast multipole method was used for computing Coulomb electrostatic forces and polarization interactions [53].

As mentioned previously, the density and heat of vaporization were computed for each pure systems of interest. The average bulk density is computed by performing NPT simulations using the formula:

$$\langle \rho \rangle = \frac{m}{\langle V \rangle} = \frac{N_{mol}M}{\langle V \rangle N_A}, \quad (9)$$

where  $\langle V \rangle$  is the average volume of the simulation box,  $N_{mol}$  is the number of molecules in the simulation box,  $M$  is the molar mass of the molecule and  $N_A$  is the Avogadro constant.

The enthalpy of vaporization is computed by performing NVT simulations, using the formula:

$$\Delta H_{vap}(T) = U_g^{pot} - U_l^{pot} + k_B T, \quad (10)$$

where,  $U_l^{pot} = \frac{U_T^{pot}}{N_{mol}}$  is the potential energy of a molecule in liquid phase and  $U_T^{pot}$  is the total potential energy of the system, and  $U_g^{pot}$  the potential energy of one molecule in the gas phase (in vacuum).

The binary monoamide/dodecane mixture are solutions that may not behave ideally. The deviations from ideal behavior, consequence of new interactions between the monoamides and dodecane within the mixture, can be characterized by excess quantities. The excess volume  $\Delta V_{ex}$  has been calculated by subtracting the pure component molar volume from the mixture molar volume, using the following equation:

$$\Delta V_{ex} = \frac{x_1 M_1 + x_2 M_2}{\rho} - \frac{x_1 M_1}{\rho_1} - \frac{x_2 M_2}{\rho_2}, \quad (11)$$

where  $x_1$  and  $x_2$  are mole fractions of monoamide and dodecane, respectively,  $M_1$  and  $M_2$  the respective molar masses, and  $\rho_1$  and  $\rho_2$  the densities of the respective pure components;  $\rho$  is the density of the mixture.

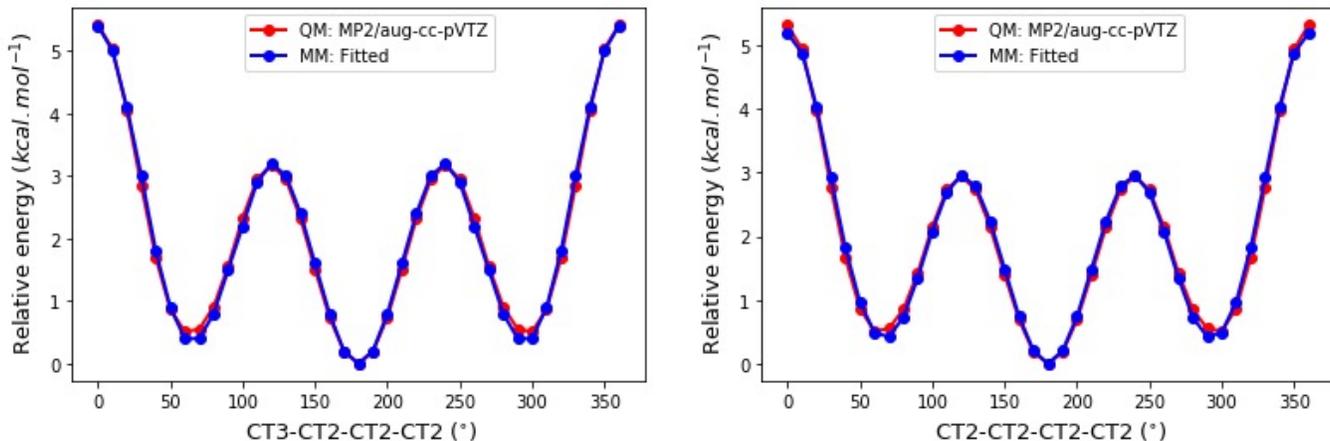


FIG. 3: Relative energy profiles of the dihedral angle scans in heptane (CT2–CT2–CT2–CT2 middle torsion on the left and CT3–CT2–CT2–CT2 terminal torsion on the right). The red lines represent the MP2 QC curves, while the blue ones the fitted FFs.

The excess enthalpy  $\Delta H_{ex}$  of the binary mixture corresponds to the thermodynamic activity of each component in the mixture. It is computed as:

$$\Delta H_{ex} = H_{mix} - x_1 H_1 - x_2 H_2, \quad (12)$$

where  $H_{mix}$  is the enthalpy of the monoamide/dodecane mixture,  $x_1$  and  $x_2$  being defined above, and  $H_1$  and  $H_2$  are the enthalpies of pure monoamide and pure dodecane solvents.

### PHYSICAL PROPERTIES OF PURE PHASES: FF VALIDATION

All the figures labeled “S” below are provided as the supplementary material.

#### Density and heat of vaporization

In Table II, we report the density and the enthalpy of vaporization for the alkane series (ethane, propane, butane, heptane, dodecane), the two monoamides DEHiBA and DEHBA (displayed in Figure 1) and the primary amide N,N-Dimethylacetamide (DMA).

We can clearly notice that the bulk density values agree well with the experimental ones, as the largest deviation is 5% for dodecane and 6% for DEHiBA. This confirms the correct description of the phase equilibria, connected the accurate positioning of the minima of potential wells by our FFs. For the heat of vaporization, we can solely consider the alkanes for the comparison to experimental data. The largest deviation is less than  $1 \text{ kcal} \cdot \text{mol}^{-1}$  for dodecane. For the simulated monoamides no experimental heat of vaporization is available in the literature, but since the same parametrization approach was used

for both systems, we can safely trust our predicted values. To our knowledge, we are the first group to report the value of the heat of vaporization of DEHiBA and DEHBA via computational methods. The heats of vaporization of DEHiBA and DEHBA come out quite close for both molecules ( $\simeq 23.5 \text{ kcal} \cdot \text{mol}^{-1}$  for DEHiBA and  $\simeq 25 \text{ kcal} \cdot \text{mol}^{-1}$  for DEHBA), this was expected since both molecules are quite similar as they only differ by the branching of the alkyl chain attached to the carbonyl group (butyl vs. iso-butyl). Overall, the computed properties yield a great agreement with experiment, and the errors are within the error bars of alkanes traditional FF developed based on density and heat of vaporization, about 5% for the densities and up to  $1 \text{ kcal} \cdot \text{mol}^{-1}$  for the heats of vaporization [16, 17, 19, 56].

Some FFs reported in the literature [13–17, 19], may give very accurate densities and heats of vaporization for some carbon chain lengths. This is not surprising as the latter properties were used in their parameterization process. However, as already mentioned, these FFs are not always transferable for all molecular groups. For instance, the dodecane heat of vaporization obtained with OPLS-AA FF is about  $\simeq 22.3 \text{ kcal} \cdot \text{mol}^{-1}$  with a deviation of  $\simeq 7.6 \text{ kcal} \cdot \text{mol}^{-1}$  from the experiment, and a density of  $839 \text{ kg} \cdot \text{m}^{-3}$  overestimating the experimental value of  $745 \text{ kg} \cdot \text{m}^{-3}$  [19]. Moreover, some studies in literature have reported the gel-phase formation at room temperature for long-chain hydrocarbons of more than eight carbons (including dodecane) using the OPLS-AA force field [21, 22]. These deviations pair with the fact that the OPLS-AA FF was adjusted to reproduce liquid densities and enthalpies of vaporization of short alkanes (ethane, propane, and butane). As a result, Siu *et al.* have re-optimized these parameters for longer alkanes (labeled L-OPLS [19]), following the same parameterization approach of the original OPLS (based on experi-

TABLE II: Densities  $\rho$  (in  $\text{kg} \cdot \text{m}^{-3}$ ) and heats of vaporization  $\Delta H_{vap}$  (in  $\text{kcal} \cdot \text{mol}^{-1}$ ) of alkanes and monoamides (DEHiBA and DEHBA) using the parameters derived in this work. <sup>a</sup> Experimental values are taken from Haynes *et al.* [54]. <sup>b</sup> OPLS-AA values taken from Refs. 55, and 17. <sup>c</sup> OPLS-L values taken from Ref. 19. Error estimates were obtained by block averaging, considering the oscillations of the energy and temperature during the last 100 ps of the production trajectory. The temperatures were selected based on the available experimental data.

Name	T(K)	$\rho(\text{kg} \cdot \text{m}^{-3})$			$\Delta H_{vap}$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )		
		This work	Expt. <sup>a</sup>	OPLS	This work	Expt. <sup>a</sup>	OPLS
Ethane	185	566	544	538 <sup>b</sup>	$3.65 \pm 0.06$	3.60	3.44 <sup>b</sup>
Propane	225	604	587	580 <sup>b</sup>	$3.57 \pm 0.07$	4.40	4.50 <sup>b</sup>
Butane	273	596	601	589 <sup>b</sup>	$4.66 \pm 0.10$	5.36	5.00 <sup>b</sup>
Heptane	298	700	677	679 <sup>b</sup>	$7.69 \pm 0.10$	8.60	9.58 <sup>b</sup>
	371	637	616		$7.84 \pm 0.05$	7.60	
Decane	298	761	730	727 <sup>b</sup> , 726 <sup>c</sup>	$12.90 \pm 0.10$	12.30	13.35 <sup>b</sup>
Dodecane	298	784	745	750 <sup>b</sup> , 744 <sup>c</sup>	$15.43 \pm 0.19$	14.70	22.3 <sup>c</sup>
	490	604	590		$11.51 \pm 0.20$	10.54	
Tridecane	300	781	756	840 <sup>b</sup>	$14.90 \pm 0.10$	15.70	12.9 <sup>b</sup>
Isopentane	301	648	617		$6.32 \pm 0.04$	5.88	
DMA	298	937	936	911 <sup>b</sup>	$10.81 \pm 0.11$	11.75	11.99
DEHiBA	298	898	865		$23.34 \pm 0.15$		
	308	892	858		$23.13 \pm 0.15$		
	318	883	851		$21.46 \pm 0.15$		
DEHBA	298	916	861		$24.93 \pm 0.21$		
	308	905	854		$24.44 \pm 0.21$		
	318	895	847		$22.56 \pm 0.21$		

mental densities and vaporization heats). Still, this may lead to the same disadvantages for much longer molecules than the ones used in the parameterization. It should also be mentioned that, even though cis-9-octadecene molecule was used in the parameterization process of L-OPLS, a deviation of  $\approx 2 \text{ kcal} \cdot \text{mol}^{-1}$  from reference data was accepted which is more than the maximum deviation we report in this work ( $1 \text{ kcal} \cdot \text{mol}^{-1}$ ). All of this affirm the strength and reliability of our parametrization approach and the developed FFs in simulating/predicting thermodynamic properties of alkanes and monoamides.

### Structural properties: Chain conformation (trans-gauche populations)

The chain conformation of alkane is analyzed in the light of the population of the gauche and trans states in the alkane series (heptane, decane, dodecane and tridecane), which is characterized by the **D** ratio defined as  $\frac{\text{population of trans}}{\text{population (gauche+trans)}}$  for each specific torsion angle labelled as in Figure 4. The gauche and trans conformations were classified based on the torsional barriers at  $120^\circ$  and  $240^\circ$  between the two states (Figure 3). This kind of analysis serves to determine the preferred conformations in the liquid phases, and among the fundamental structural units in biology and chemistry in general [57–59].

Table III compiles the trans-gauche population ratio as a function of the dihedral angle along the carbon chain

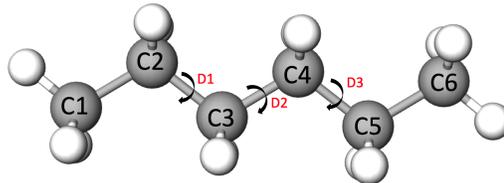


FIG. 4: Representation of dihedral angle along the carbon chain for the hexane molecule.

obtained from our FF and some other FFs (OPLS-AA and L-OPLS), the only available experimental values in literature are those of tridecane. We can clearly observe the good agreement of our work with the experimental data for tridecane, a difference of 6 % for the terminal torsion populations (D1 and D9) and 2 % for the torsion in the middle of the chain (Dx with  $x \in [3..7]$ ) population was noted. The largest deviation was observed for heptane, about 7 % for the terminal torsion population and 1 % for the middle torsion ones.

It is worth mentioning here that the results obtained with the OPLS-AA FF for tridecane deviate by about 34 % from the experimental ones for the chain terminal populations and 25 % for the trans populations of the interior of the chain. This again confirms the inability of OPLS-AA FF to correctly describe the trans-gauche populations for alkanes. The more recent L-OPLS FF results are in quite good agreement with the experimen-

tal data reported by Casal *et al.* [60], with an average deviation of about 3 %, and quite close with the prediction of our FF (average deviation 4 %) for the simulated alkanes.

### Radial distribution functions

The structural organization of the pure alkanes solvents was investigated by calculating the radial distribution functions (RDFs,  $g(r)$ ), using VMD package [61], as they reveal the distribution of neighboring molecules and the long-range solvent organization. The  $g(r)$  for carbon atoms in alkanes are plotted in Figure 5.

We have drawn the C–C RDFs starting at 2 Å, hence, the first C–C intramolecular interaction at 1.53 Å is not visible. The peak at 2.55 Å corresponding to the intramolecular C<sub>1</sub>–C<sub>3</sub> second-neighbor interaction is present for all alkanes longer than ethane. The third peak at  $\approx 3.1$  Å corresponds to the gauche C<sub>1</sub>–C<sub>4</sub> molecular segments. At the distance of  $\approx 3.88$  Å appears a peak that can be attributed to the trans C<sub>1</sub>–C<sub>4</sub> molecular segments. The peak positions reported in Table IV are quite consistent with the experimental X-ray diffraction results of Habenschuss *et al.* [62]. We have also calculated the structure factor for heptane and the results are in good agreement with experiment (see ESI Figure S7).

We can also notice the influence of temperature upon the RDFs, for dodecane at two different temperature (298 K and 490 K). It is readily apparent in Figure 5 (c), that the intensity of the 3rd peak increases while the 4th peak simultaneously decreases, implying that the gauche population increases with temperature and the reverse for the trans population. This was confirmed by the calculation of the populations as presented in Table III. Indeed as the temperature rises, it becomes more likely to populate conformations with higher energies (gauche). This observation further support the strength and reliability of our FF in describing the structural organization of the pure alkanes.

### MOLECULAR PROPERTIES OF DODECANE/MONOAMIDES MIXTURES

After having validated our FFs on pure alkane solvent properties, herein we present the results obtained for DEHBA and DEHiBA mixtures with dodecane. This study was performed to gain insight and to visualize molecular-level behavior of these extraction molecules (alkane-monoamides mixtures), for instance, to understand the impact of the monoamide structure, in particular the branching of the alkyl chain on the molecular organization of alkane-monoamides mixtures. At the end, this may give insights about the sharp difference

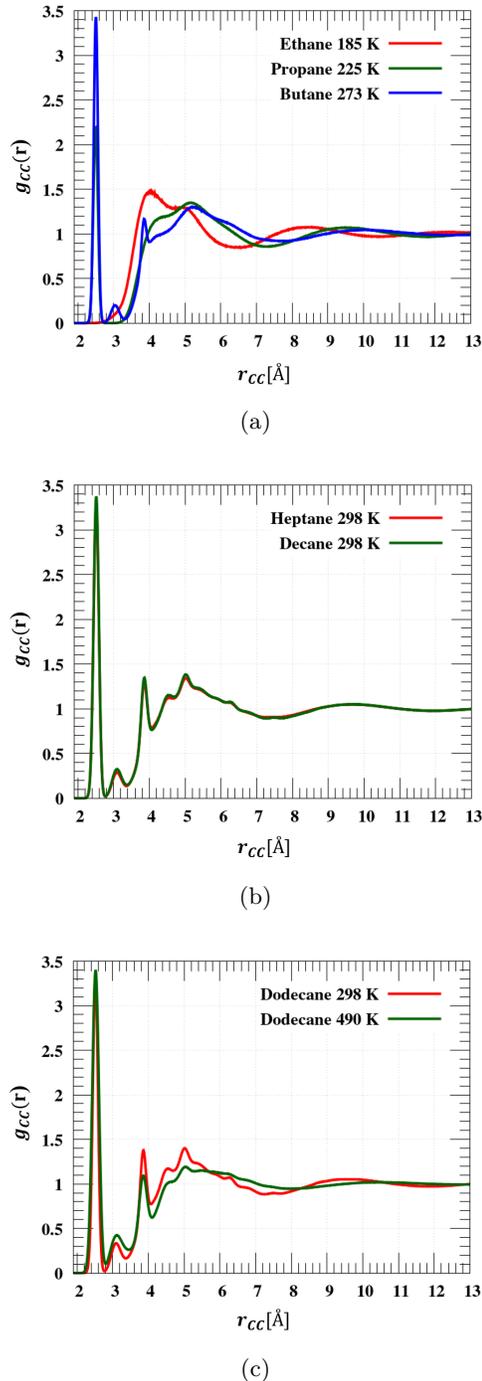


FIG. 5: Radial distribution functions of carbon atoms for selected alkanes: ethane, propane, butane (a), and heptane, decane (b) and dodecane (c).

measured for Pu(IV) extraction between DEHiBA and DEHBA [9, 10].

In order to approach as much as possible experimental conditions, the simulation boxes were constructed based on available experimental data for DEHiBA (den-

TABLE III: Trans populations in % in heptane, decane, dodecane and tridecane solvents as a function of dihedral angle noted Dn along the carbon chain. <sup>a</sup> values from Monte Carlo (MC) simulations [55] with the original OPLS-AA FF [17]; <sup>b</sup> values from MD simulations with the L-OPLS FF [19]; <sup>c</sup> Experimental values for the first and the sixth carbon torsion of tridecane from Casal *et al.* [60]

Molecule	Dihedral angle	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
heptane (298 K)	This work	65	68	68	65						
	OPLS-AA <sup>a</sup>	73	82	78	76						
decane (300 K)	This work	65	69	66	66	66	69	65			
	OPLS-AA <sup>a</sup>	78	81	84	82	81	81	79			
dodecane (298 K)	This work	64	68	65	65	65	65	68	64		
	OPLS-AA <sup>a</sup>	80	83	79	80	78	80	81	83	81	
dodecane (490 K)	This work	56	61	57	57	57	57	61	56		
tridecane (298 K)	L-OPLS <sup>b</sup>	55	66	64	64	64	64	64	66	55	
	OPLS-AA <sup>a</sup>	92	97	98	98	98	98	98	97	92	
	This work	64	67	64	65	65	65	65	68	65	64
	Expt. <sup>c</sup>	58	-	-	-	-	67	-	-	-	58

TABLE IV: Comparison of the positions in Å of the local maxima of the C–C RDFs computed in this work to X-ray diffraction data [62].

	C <sub>1</sub> –C <sub>2</sub>		C <sub>1</sub> –C <sub>3</sub>		gC <sub>1</sub> –C <sub>4</sub>		tC <sub>1</sub> –C <sub>4</sub>	
	MD	Expt.	MD	Expt.	MD	Expt.	MD	Expt.
Butane 273 K	1.54	1.55	2.55	2.56	3.10	3.07	3.88	3.93
Heptane 298 K	1.53	1.55	2.54	2.56	3.10	3.14	3.87	3.92
Decane 298 K	1.53	1.54	2.55	2.57	3.10	3.15	3.88	3.95
Dodecane 298 K	1.53		2.55		3.10		3.88	

sities) at room temperature and atmospheric pressure (298 K, 1 atm). Since the experimental data were only available for DEHiBA [34], the same simulation conditions were considered for the DEHBA/dodecane mixtures to be able to compare the solvent mixtures, as a function of the nature of the monoamide. Molecular compositions for the simulations are given in ESI Table S5. The extractant mole fraction  $x_{mono}$  was varied from 10 to 62 %.

### Density

The computed densities of different DEHiBA/dodecane and DEHBA/dodecane mixtures are displayed in Figure 6. The calculated values for the former compare favorably to experimental data, with an average percent error of 5 % across the range of solvents ( $33 \text{ kg} \cdot \text{m}^{-3}$ ). The densities of the DEHBA/dodecane mixtures appear to be close to the DEHiBA/dodecane ones; this was expected since the structures of the two ligands are quite alike and also from the fact that density of both monoamides in pure phase are nearly the same, within  $4 \text{ kg} \cdot \text{m}^{-3}$ .

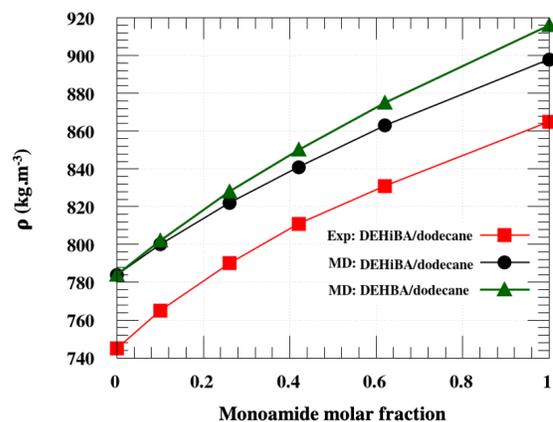


FIG. 6: Mass density of the DEHiBA/dodecane and DEHBA/dodecane mixtures as a function of the mole fraction. Dark blue circles and green diamonds correspond to the simulated densities for the DEHiBA/dodecane and DEHBA/dodecane, respectively. Red squares correspond to experimental measurements by Coquil *et al.* [34].

### Radial distribution functions and aggregation numbers

To gain insights into the structures of DEHiBA/DEHBA in the liquid phase as well as the organization of DEHiBA/DEHBA within the dodecane solvent, the RDFs between nitrogen and/or oxygen belonging to different monoamide were computed using a bin width of  $0.10 \text{ Å}$  and a cutoff distance of  $25 \text{ Å}$  (about half of the simulation box length), averaged over the 20 ns of the MD trajectories. The O–O RDFs of the monoamide/dodecane mixtures at different concentrations are reported in Figure 7. For the pure

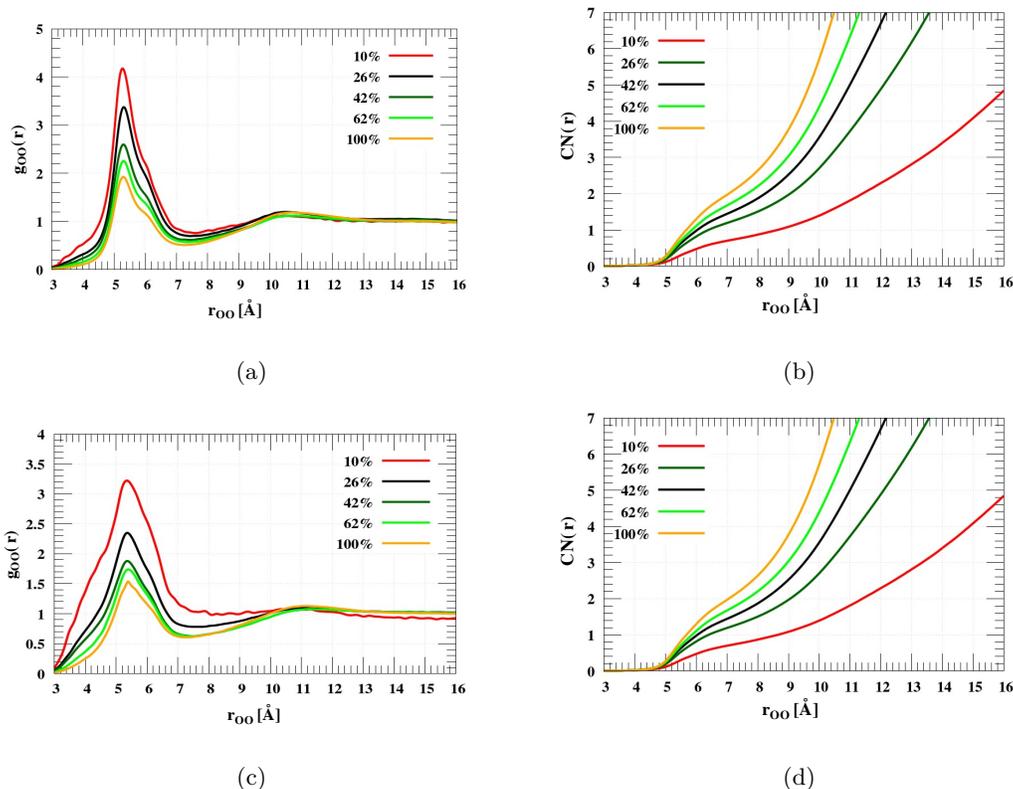


FIG. 7: Radial distribution functions and the coordination number (CN) of oxygen atoms for DEHiBA/dodecane (a-b) and DEHBA/dodecane mixtures (c-d) at different mole fractions.

monoamide phases, the RDF profiles are equivalent, suggesting that the molecular organizations for these two alike monoamides are quite similar (displayed in Figure 1). However, the peak positions are shifted by about  $+0.1$  Å for DEHBA as compared to DEHiBA. The relative steric hindrance increases by the branching of the alkyl group bonded to carbonyl site is probably responsible for the shift (Figures S10-S11-S12-S13 of the ESI).

Using Figure 7, one can discuss the evolution of the O–O RDFs of DEHiBA/DEHBA for various monoamide/dodecane ratios. In both sets of mixtures, the peak heights of the RDFs gradually decrease as the concentration of monoamide increases. However, the peak positions are insensitive to the concentration indicating that the dilution of the extractant in dodecane has no significant impact on the molecular organization. These observations also apply to the O–N and N–N RDFs showed in Figures S10 and S11 of the ESI.

The (O–O) and (N–N) RDFs suggest a "first coordination shell" ranging out to  $7.2$  Å (Figure 7 and 8), which corresponds to self-assembly of monoamides. The average coordination numbers were estimated by integrating (O–O) and (N–N) RDFs (see Figure 7). For  $x_{mono}=10\%$ , it amounts to about 0.6 and about 1.2 for  $x_{mono}=26\%$ , for both DEHiBA and DEHBA, which suggests the possibility of monoamides dimer formations in

these phases. The coordination number steadily increases as the monoamide concentration grows, in a similar way for both monoamides; it reaches 1.5 for  $x_{mono}=42\%$ , about 1.9 for  $x_{mono}=62\%$ , and tops at 2.4 for the pure phases.

The snapshots showed in Figure S8 of the ESI, exemplify the formation of aggregates in dodecane. We suspect that aggregation occurs because the monoamide molecules are polar, and prefer interacting with each other rather than with alkanes. In general, alkanes are known to be "insoluble" in polar solvent (such as water). They are soluble only in non-polar and slightly polar solvents. In our case, DEHiBA and DEHBA are polar molecules with a dipole moment<sup>[63]</sup> of  $\simeq 3.64$  debye (water  $\simeq 1.85$  debye), yet, the fact that they are long molecules with lipophilic tails makes them soluble in alkanes.

#### Relative orientations in monoamide dimers

RDFs solely provide a 1D picture of the molecular organization for the DEHBA/DEHiBA pure and dodecane mixed solutions. To further characterize it and have a clearer vision in 3D, we have analysed the relative orientations of the carbonyl groups belonging to two clustered

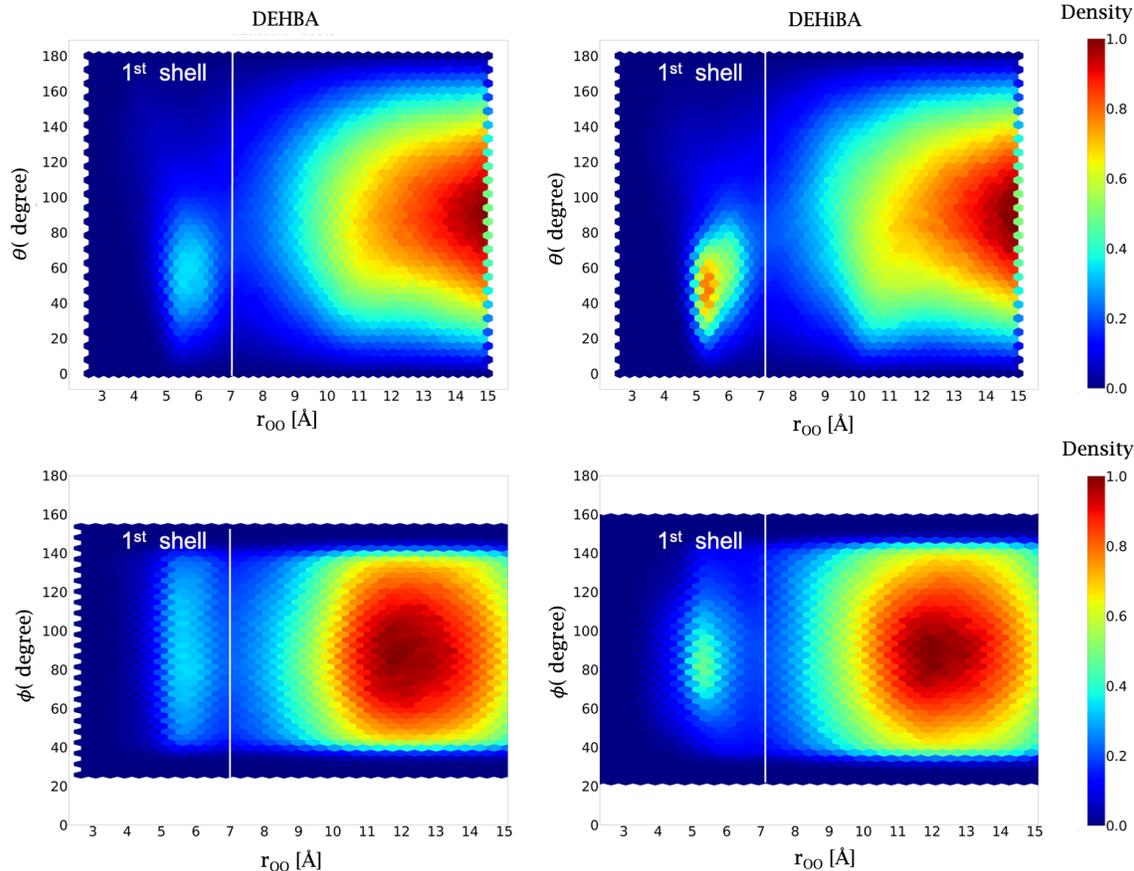


FIG. 8: Analysis of the carbonyl group C=O orientation,  $\theta(r_{OO})$ , and relative orientations of the NCO planes,  $\phi(r_{OO})$ , as defined in Figure 9), as a function of the distance between the oxygen of the carbonyl group for the pure DEHBA and DEHiBA solvents. The color bar indicates the normalized count density.

monoamide molecules, defined by the angle  $\theta$  as well as the angle  $\phi$  between the two NCO planes (See Figure 9), as a function of the mole fraction. Figure 8 shows the angles  $\phi$  and  $\theta$  as a function of the distance between the oxygen of the carbonyl group for both extractants for pure systems (DEHBA and DEHiBA); the results for  $x=10, 26, 42, 62$  and 100% mole fractions are available in ESI (see Figure S14) but do follow exactly the same trend that will be discussed below, as well as a time evolution of the relative orientations as a function of time (see Figure S15) that indicate that the property is well converged after 8 ns.

For DEHBA, the angular analysis of  $\phi$  shows that, at  $d_{OO} = 5.5 \text{ \AA}$ , a vast range of  $\phi$  angles (43 to 135°) is preferred in mixtures and pure phases. Contrary to DEHBA, DEHiBA seems to orient perpendicularly (70 to 100°) to another DEHiBA molecule while making dimers. As for long  $d_{OO}$  distances ( $d_{OO} = 9$  to 16 Å) corresponding to the "second coordination sphere", one can notice that both monoamides molecules prefer perpendicular conformations.

As for the analysis of the carbonyl group orientation,

for both systems (DEHBA and DEHiBA) diluted in dodecane or in pure phase, the top panels of Figure 8 suggest that the monoamides preferably interact with each other in a conformation where the angle  $\theta$  between carbonyl groups is in the range of 20 to 80°, see Figure 10.

### Excess volume and excess enthalpy

The excess volume of mixing and excess enthalpy of mixture are a result of complex combination of molecular properties between two molecular species such size, shape, as well as the inherent molecular interaction. These properties have been extensively calculated for TBP/n-alkanes mixtures in the literature [23, 24]. It was used as a measure of the force-field accuracy for such solutions and also to give insight into extractant-solvent interactions. Such data have been recently measured experimentally by Coquil *et al.* [34], but it remains of great interest to predict them with MD simulations. Our predictions are drawn in Figure 11.

The excess volume of mixing is found to be positive for

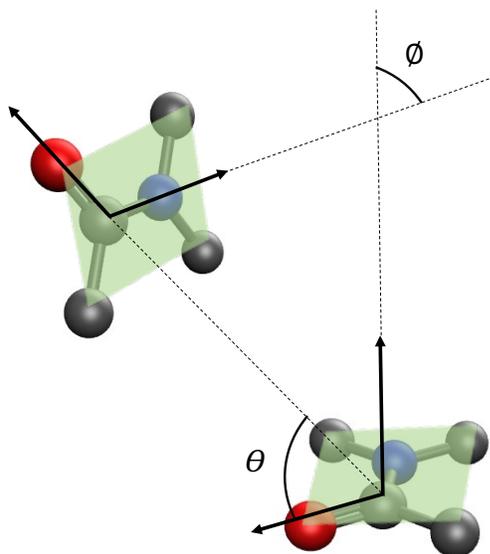


FIG. 9: Illustration of the orientation convention used to analyse the molecular organization for monoamide/dodecane mixtures. The carbonyl group C=O orientation ( $\theta$ ) and the orientation of the flat surface of the amide function "NCO" ( $\phi$ ).

the entire range of mole fraction, which indicates looser packing in the mixtures compared to the pure phases. The excess enthalpy of mixing appears to be endothermic for both the DEHBA/dodecane and DEHiBA/dodecane mixtures and reaches a maximum for the 42 % mole fraction, hence dodecane shows an unfavorable mixing for all mole fractions and for both extractant ligands. Comparing to experimental data, our values are bit higher than expected, but the tendency and the endothermic behavior is well captured with a maximum around the 42 % mole fraction.

The uncertainties for both properties were also estimated (see details in the ESI) and the average errors are about  $0.1 \text{ kcal} \cdot \text{mol}^{-1}$  for the excess enthalpies and about  $0.01 \text{ cm}^3 \cdot \text{mol}^{-1}$  for the excess volumes. Only the uncertainties for  $\Delta H_{ex}$  are displayed in Figure 11.

Recently, Coquil *et al.* [34] have measured the enthalpy of mixing for DEHiBA with dodecane. A comparison to these experimental data with our values seem to overestimate this properties, however this does not affect the accuracy of our FF since, in the work of Servis *et al.* [23] and also the one of Arya Das *et al.* [64] have showed that it is important to get the correct sign for the excess enthalpy of mixing which proves the accuracy of the force field.

The comparison of the properties of monoamide/dodecane mixtures to that reported by Servis *et al.* [23] for phosphoric extractants, such as tributyl phosphate (TBP), triamyl phosphate (TAP), dibutyl butyl phosphonate (DBBP) and diamyl amyl

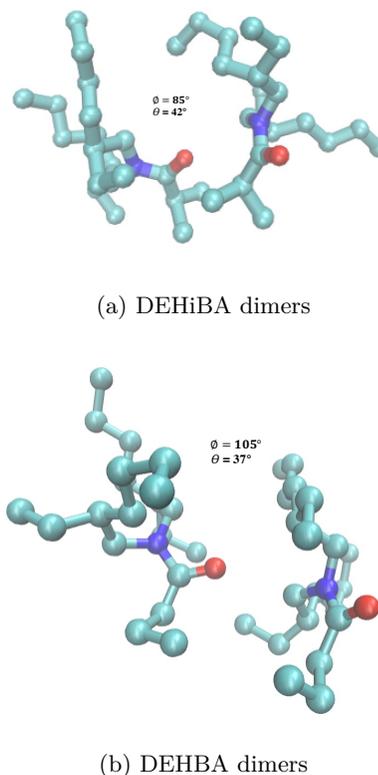


FIG. 10: The preferred DEHiBA (a) and DEHBA (b) dimer structures according to the C=O and N-C=O relative orientation analysis.

phosphonate (DAAP), shows that the enthalpy of mixing is overall larger,  $0.8 \text{ kcal} \cdot \text{mol}^{-1}$ , for monoamides than for the organophosphorus extractant structures, e.g.  $0.35 \text{ kcal} \cdot \text{mol}^{-1}$  for TBP.

## CONCLUSION

In this work, *ab initio* based force-field model for short- and long-alkane chains and amide derivatives was successfully constructed by solely considering quantum chemical calculations (at the MP2 level of theory) and taking explicitly into account polarization effects. The different intra-molecular and inter-molecular parameters were derived and validated by performing MD simulations to calculate liquid-phase thermodynamic and structural properties. The reported simulation results for the alkanes are in great agreement with experimental data. These high-quality results for both molecular families suggest that this purely *ab initio* parametrization methodology is promising and can be applied to other molecules of interest.

The developed FFs were also used to simulate and calculate properties of monoamide/dodecane mixtures,

## ACKNOWLEDGEMENTS

We first would like to thank M. Masella and P. Guilbaud for fruitful discussions. We acknowledge support by the French government through the Program "Investissement d'avenir" (Grants LABEX CaPPA/ANR- 11-LABX-0005-01 and I-SITE ULNE/ANR16-IDEX-0004 ULNE), as well as by the Ministry of Higher Education and Research, Hauts de France council and European Regional Development Fund (ERDF) through the Contrat de Projets État Region (CPER CLIMIBIO). We further acknowledge the "Groupement de recherche" GDR 2035 SolvATE. In addition, this work was granted access to the HPC resources of CINES/IDRIS/TGCC under the allocation 2019–2020 (A0070801859) made by GENCI. We also acknowledge the CEA for the Ph.D. grant given to A.F.

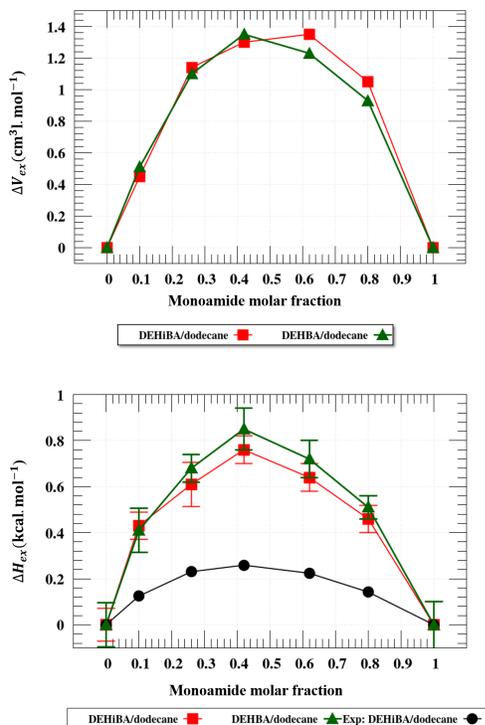


FIG. 11: Excess Volume ( $\Delta V_{ex}$ ) and Excess Enthalpy ( $\Delta H_{ex}$ ) of mixing of the DEHiBA/dodecane and DEHBA/dodecane binary mixtures.

namely DEHiBA/dodecane and DEHBA/dodecane. The structural analysis revealed that for both mixtures, amide ligands tend to self-assemble (aggregate) in the organic solution. However, The RDFs calculations showed that the molecular organization for both systems is quite similar. The relative orientation analysis for the C=O carbonyl group and the amide function for both systems showed a small difference on the molecular organization for both monoamides. Excess volume and excess enthalpy have been also calculated and seemed to be quite similar for both systems. All the results and predictions are the first step towards a more realistic description of more complex phenomena in which these mixtures could interact with strongly ionic complexes, such as plutonium nitrate ions, and be in contact with other types of solvent, namely water. In particular, mimicking liquid-liquid extraction remains a challenging task in term of FF development but also computational resources, since way longer simulations as well as larger systems will be required.

## CONFLICTS

There are no conflicts to declare.

- \* [florent.real@univ-lille.fr](mailto:florent.real@univ-lille.fr)
- [1] K. D. Papavasileiou, L. D. Peristeras, A. Bick, and I. G. Economou, Molecular dynamics simulation of pure n-alkanes and their mixtures at elevated temperatures using atomistic and coarse-grained force fields, *J. Phys. Chem. B* **123**, 6229 (2019).
  - [2] F. Goodsaid-Zalduondo and D. Engelman, Conformation of liquid n-alkanes, *Biophys. J.* **35**, 587 (1981).
  - [3] P. L. Yeagle, *The structure of biological membranes* (CRC press, 2004).
  - [4] J. Murgich, Molecular simulation and the aggregation of the heavy fractions in crude oils, *Mol. Simulat.* **29**, 451 (2003).
  - [5] C. Bin, G. Jinsen, and X. Chunming, The applications of molecular simulation technology in the fields of petroleum, *Prog. Chem.* **16**, 3 (2004).
  - [6] P. Ungerer, V. Lachet, and B. Tavitian, Applications of molecular simulation in oil and gas production and processing, *Oil. Gas. Sci. Technol.* **61**, 387 (2006).
  - [7] X. Chen, L. Hou, X. Wei, and D. Bedrov, Transport properties of waxy crude oil: A molecular dynamics simulation study, *ACS omega* **5**, 18557 (2020).
  - [8] A. A. Petrov, *Petroleum hydrocarbons* (Springer Science & Business Media, 2012).
  - [9] D. Prabhu, G. Mahajan, and G. Nair, Di (2-ethyl hexyl) butyramide and di (2-ethyl hexyl) isobutyramide as extractants for uranium (vi) and plutonium (iv), *J. Radioanal. Nucl. Chem.* **224**, 113 (1997).
  - [10] R. Taylor, ed., *Reprocessing and Recycling of Spent Nuclear Fuel*, Woodhead Publishing Series in Energy (Woodhead Publishing, Oxford, 2015) pp. xiii – xv.
  - [11] W. W. Schulz and J. D. Navratil, *Science and technology of tributyl phosphate* (CRC Press, Inc, United States, 1987).
  - [12] Y. Ban, S. Hotoku, N. Tsutsui, A. Suzuki, Y. Tsubata, and T. Matsumura, Uranium and plutonium extraction by n,n-dialkylamides using multistage mixer-settler extractors, *Procedia Chem.* **21**, 156 (2016).
  - [13] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger, and W. F. van Gunsteren, The GRO-MOS biomolecular simulation program package, *J. Phys. Chem. A* **103**, 3596 (1999).
  - [14] N. L. Allinger, Y. H. Yuh, and J. H. Lii, Molecular mechanics. the mm3 force field for hydrocarbons. 1, *J. Am. Chem. Soc.* **111**, 8551 (1989).
  - [15] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caffisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, CHARMM: The biomolecular simulation program, *J. Comput. Chem.* **30**, 1545 (2009).
  - [16] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, The amber biomolecular simulation programs, *J. Comput. Chem.* **26**, 1668 (2005).
  - [17] W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, *J. Am. Chem. Soc.* **118**, 11225 (1996).
  - [18] I. V. Vorobyov, V. M. Anisimov, and A. D. MacKerell, Polarizable empirical force field for alkanes based on the classical drude oscillator model, *J. Phys. Chem. B* **109**, 18988 (2005).
  - [19] S. W. I. Siu, K. Pluhackova, and R. A. Böckmann, Optimization of the opls-aa force field for long hydrocarbons, *J. Chem. Theory Comput.* **8**, 1459 (2012).
  - [20] A. M. Nikitin, Y. V. Milchevskiy, and A. P. Lyubartsev, A new AMBER-compatible force field parameter set for alkanes, *J. Mol. Model.* **20**, 2143 (2014).
  - [21] Q. N. Vo, C. A. Hawkins, L. X. Dang, M. Nilsson, and H. D. Nguyen, Computational study of molecular structure and self-association of tri-n-butyl phosphates in n-dodecane, *J. Phys. Chem. B* **119**, 1588 (2015).
  - [22] X. Ye, S. Cui, V. F. de Almeida, and B. Khomami, Effect of varying the 1–4 intramolecular scaling factor in atomistic simulations of long-chain n-alkanes with the opls-aa model, *J. Mol. Model.* **19**, 1251 (2013).
  - [23] M. J. Servis, D. T. Wu, and J. C. Shafer, The role of solvent and neutral organophosphorus extractant structure in their organization and association, *J. Mol. Liq.* **253**, 314 (2018).
  - [24] S. Cui, V. F. de Almeida, and B. Khomami, Molecular dynamics simulations of tri-n-butyl-phosphate/n-dodecane mixture: Thermophysical properties and molecular structure, *J. Phys. Chem. B* **118**, 10750 (2014).
  - [25] J. Mu, R. Motokawa, C. D. Williams, K. Akutsu, S. Nishitsuji, and A. J. Masters, Comparative molecular dynamics study on tri-n-butyl phosphate in organic and aqueous environments and its relevance to nuclear extraction processes, *J. Phys. Chem. B* **120**, 5183 (2016).
  - [26] A. Paquet, O. Diat, L. Berthon, and P. Guilbaud, Aggregation in organic phases after solvent extraction of uranyl nitrate: X-ray scattering and molecular dynamic simulations, *J. Mol. Liq.* **277**, 22 (2019).
  - [27] B. Qiao, G. Ferru, M. Olvera de la Cruz, and R. J. Ellis, Molecular origins of mesoscale ordering in a metalloamphiphile phase, *ACS Cent. Sci.* **1**, 493 (2015).
  - [28] A. Jagielska, Y. A. Arnautova, and H. A. Scheraga, Derivation of a new force field for crystal-structure prediction using global optimization: nonbonded potential parameters for amines, imidazoles, amides, and carboxylic acids, *J. Phys. Chem. B* **108**, 12181 (2004).
  - [29] J.-H. Lii and N. L. Allinger, The mm3 force field for amides, polypeptides and proteins, *J. Comput. Chem.* **12**, 186 (1991).
  - [30] A. Hagler, E. Huler, and S. Lifson, Energy functions for peptides and proteins. i. derivation of a consistent force field including the hydrogen bond from amide crystals, *J. Am. Chem. Soc.* **96**, 5319 (1974).
  - [31] F. Réal, V. Vallet, J.-P. Flament, and M. Masella, Revisiting a many-body model for water based on a single polarizable site: From gas phase clusters to liquid and air/liquid water systems, *J. Chem. Phys.* **139**, 114502 (2013).
  - [32] F. Réal, M. Trumm, B. Schimmelpfennig, M. Masella, and V. Vallet, Further insights in the ability of classical nonadditive potentials to model actinide ion–water interactions, *J. Comput. Chem.* **34**, 707 (2012).

- [33] C. Houriez, F. Réal, V. Vallet, M. Mautner, and M. Masella (Foreign), Ion hydration free energies and water surface potential in water nano drops: The cluster pair approximation and the proton hydration gibbs free energy in solution, *J. Chem. Phys.* **151**, 174504 (2019).
- [34] M. Coquil, N. Boubals, M. Duvaill, M.-C. Charbonnel, and J.-F. Dufrière, On interactions in binary mixtures used in solvent extraction: Insights from combined isothermal titration calorimetry experiments and molecular dynamics simulations, *J. Mol. Liq.*, **116985** (2021).
- [35] M. Masella and P. Cuniasse, A many-body model to study proteins. i. applications to ml complexes,  $M^{m+} = Li^+, Na^+, K^+, Mg^{2+}, Ca^{2+},$  and  $Zn^{2+}$ ,  $l = H_2O, CH_3OH, HCONH_2$ ,  $n = 1-6$ , and to small hydrogen bonded systems, *J. Chem. Phys.* **119**, 1866 (2003).
- [36] B. T. Thole, Molecular polarizabilities calculated with a modified dipole interaction, *Chem. Phys.* **59**, 341 (1981).
- [37] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, Molpro: a general-purpose quantum chemistry program package, *WIREs Comput. Mol. Sci.* **2**, 242 (2012).
- [38] M. J. Frisch, M. Head-Gordon, and J. A. Pople, A direct mp2 gradient method, *Chem. Phys. Lett.* **166**, 275 (1990).
- [39] T. H. Dunning Jr, Gaussian basis sets for use in correlated molecular calculations. i. the atoms boron through neon and hydrogen, *J. Chem. Phys.* **90**, 1007 (1989).
- [40] R. A. Kendall, T. H. Dunning, and R. J. Harrison, Electron affinities of the first-row atoms revisited. systematic basis sets and wave functions, *J. Chem. Phys.* **96**, 6796 (1992).
- [41] J. B. Klauda, B. R. Brooks, A. D. MacKerell, R. M. Venable, and R. W. Pastor, An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and a dppc bilayer, *J. Phys. Chem. B* **109**, 5300 (2005).
- [42] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, A fifth-order perturbation comparison of electron correlation theories, *Chem. Phys. Lett.* **157**, 479 (1989).
- [43] A. V. Marenich, S. V. Jerome, C. J. Cramer, and D. G. Truhlar, Charge model 5: An extension of hirshfeld population analysis for the accurate description of molecular interactions in gaseous and condensed phases, *J. Chem. Theory Comput.* **8**, 527 (2012).
- [44] L. S. Dodda, J. Z. Vilseck, K. J. Cutrona, and W. L. Jorgensen, Evaluation of CM5 charges for nonaqueous condensed-phase modeling, *J. Chem. Theory Comput.* **11**, 4273 (2015).
- [45] D. Bedrov, J.-P. Piquemal, O. Borodin, A. D. MacKerell Jr, B. Roux, and C. Schröder, Molecular dynamics simulations of ionic liquids and electrolytes using polarizable force fields, *Chem. Rev.* **119**, 7940 (2019).
- [46] A. V. Marenich, C. J. Cramer, and D. G. Truhlar, Reduced and quenched polarizabilities of interior atoms in molecules, *Chem. Sci.* **4**, 2349 (2013).
- [47] F. L. Hirshfeld, Bonded-atom fragments for describing molecular charge densities, *Theoret. Chim. Acta* **44**, 129 (1977).
- [48] S. F. Boys and F. Bernardi, The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced error, *Mol. Phys.* **19**, 553 (1970).
- [49] J. R. Maple, M.-J. Hwang, T. P. Stockfisch, U. Dinur, M. Waldman, C. S. Ewig, and A. T. Hagler, Derivation of class ii force fields. i. methodology and quantum force field for the alkyl functional group and alkane molecules, *J. Comput. Chem.* **15**, 162 (1994).
- [50] J. Doherty, Pest: Model-independent parameter estimation and uncertainty analysis (2010), see <http://www.pesthomepage.org>.
- [51] The polaris-md code, <http://biodev.cea.fr/polaris/>, accessed: 2020-10-30.
- [52] S. Miyamoto and P. A. Kollman, Settle: An analytical version of the shake and rattle algorithm for rigid water models, *J. Comput. Chem.* **13**, 952 (1992).
- [53] J. P. Coles and M. Masella, The fast multipole method and point dipole moment polarizable force fields, *J. Chem. Phys.* **142**, 024109 (2015).
- [54] W. M. Haynes, ed., *CRC handbook of chemistry and physics*, 95th ed. (CRC press, 2014).
- [55] L. L. Thomas, T. J. Christakis, and W. L. Jorgensen, Conformation of alkanes in the gas phase and pure liquids, *J. Phys. Chem. B* **110**, 21198 (2006).
- [56] J. Wang and T. Hou, Application of molecular dynamics simulations in molecular property prediction. 1. density and heat of vaporization, *J. Chem. Theory Comput.* **7**, 2151 (2011).
- [57] M. Dettenmaier, Conformation of n-alkane molecules in the melt and in cyclohexane solution studied by small-angle neutron scattering, *J. Chem. Phys.* **68**, 2319 (1978).
- [58] L. S. Bartell and D. A. Kohl, Structure and rotational isomerization of free hydrocarbon chains, *J. Chem. Phys.* **39**, 3097 (1963).
- [59] F. M. Menger and L. L. D'Angelo, Does the conformation of hydrocarbon chains depend on solvation?, *J. Am. Chem. Soc.* **110**, 8241 (1988).
- [60] H. L. Casal and H. H. Mantsch, Positional dependence of solvent effects on the conformation of liquid n-alkanes. an infrared spectroscopic study, *J. Mol. Struct.* **192**, 41 (1989).
- [61] W. Humphrey, A. Dalke, and K. Schulten, Vmd: visual molecular dynamics, *J. Mol. Graphics* **14**, 33 (1996).
- [62] A. Habenschuss and A. H. Narten, X-ray diffraction study of some liquid alkanes, *J. Chem. Phys.* **92**, 5692 (1990).
- [63] Obtained with quantum calculations at the MP2 level and aug-cc-pVTZ basis set.
- [64] A. Das and S. M. Ali, Molecular dynamics simulation for the test of calibrated OPLS-AA force field for binary liquid mixture of tri-iso-amyl phosphate and n-dodecane, *J. Chem. Phys.* **148**, 074502 (2018).

**Supporting Information:**

**Structural and thermodynamics properties of  
pure phase monoamides and alkane/monoamide  
mixtures with an ab initio based force-field model**

Abdelmounaim Failali,<sup>†</sup> Eléonor Acher,<sup>†</sup> Dominique Guillaumont,<sup>†</sup> Valérie  
Vallet,<sup>‡</sup> and Florent Réal<sup>\*,‡</sup>

<sup>†</sup>*CEA, DES, ISEC, DMRC, Univ Montpellier, Marcoule, F-30207 Bagnols-sur-Ceze Cedex,  
France*

<sup>‡</sup>*Université Lille, CNRS, UMR 8523-PhLAM-Physique des Lasers, Atomes et Molécules,  
F-59000 Lille, France*

E-mail: florent.real@univ-lille.fr

# Force-field parameters

Table S1: The Buckingham, 1-4 interactions parameters derived and used in this work.

i	j	$A_{ij}$	$B_{ij}$	damping	$C_{ij}$	correc tor (1-4)	
CT3	CT3	3.86	320 643	0.3	1850	4.82	1 000 000
CT3	CT2	3.80	320 000	0.3	1500	4.82	1 000 000
CT3	NCT3	4.45	320 000	0.3	1500	5.30	1 000 000
CT3	C	3.33	71 347	0.3	1550	4.70	1 000 000
CT3	O	4.25	300 000	0.3	0	4.16	1 000 000
CT3	HA	3.84	5440	0.5	0	7.50	75 000
CT3	N	4.65	600 000	0.3	700	4.90	1 000 000
CT2	CT2	4.55	320 000	0.3	1400	4.82	1 000 000
CT2	NCT3	4.50	1 004 868	0.3	500	4.82	1 000 000
CT2	C	3.65	72 000	0.3	1500	4.82	1 000 000
CT2	O	4.40	1 000 000	0.3	500	4.16	1 000 000
CT2	HA	3.70	5700	0.5	0	7.50	75 000
CT2	N	4.15	304 000	0.3	700	4.82	1 000 000
NCT3	C	3.43	101 347	0.3	700	4.82	1 000 000
NCT3	O	4.25	500 000	0.3	700	4.16	1 000 000
NCT3	HA	3.23	1476	0.5	0	7.50	75 000
NCT3	N	4.75	500 000	0.3	700	5.30	1 000 000
C	C	4.50	304 000	0.3	500	4.82	1 000 000
C	O	4.20	500 000	0.4	1000	5.30	1 000 000
C	HA	3.23	1476	0.5	0	7.50	75 000
C	N	4.30	1 000 000	0.3	700	5.30	800 000
O	O	5.10	900 000	0.3	0	4.80	1 000 000
O	HA	7.00	75 000	0.5	0	8.00	50 000
O	N	5.20	500 000	0.5	1200	5.10	1 000 000
HA	HA	3.65	1000	0.5	0	8.00	50 000
HA	N	5.50	350 000	0.5	0	8.00	50 000
N	N	4.70	9 000 000	0.3	700	5.60	500 000

Table S2: Partial CM5 charges ( $q(\text{CM5})$  a.u.) and atomic polarizabilities ( $\alpha$  in  $\text{\AA}^3$ ) used for alkanes and amides derivatives.

Atom Type	$q(\text{CM5})$	$\alpha(\text{\AA}^3)$
CT3	-0.21	2.0
CT2	-0.14	2.0
CT1	-0.07	2.0
CT2-N	-0.02	2.0
HN	0.10	0.0
HA	+0.07	0.0
N	-0.34	1.3
O	-0.42	1.3
C	0.32	1.0

Table S3: Fitted torsion coefficients for alkanes and alkyls groups.

Dihedral torsion	$K_1$	$K_2$	$K_3$	$\Phi_0$
CT3-CT2-CT2-CT3	2	0.61	2.42	0
CT3-CT2-CT2-CT2	1.5	0.45	2.16	0
CT2-CT2-CT2-CT2	1.3	0.36	1.97	0
CT2-CT1-CT2-CT2	1.15	0.36	2	0
HA-CT-CT-HA	0.25	0	0	0

## Polarization interaction term

The corresponding polarization energy term is defined as

$$U_{pol} = \frac{1}{2} \sum_{i=1}^{N_\mu} \frac{\boldsymbol{\mu}_i^2}{\alpha_i} - \sum_{i=1}^{N_\mu} \boldsymbol{\mu}_i \cdot \mathbf{E}_i^q - \frac{1}{2} \sum_{i=1}^{N_\mu} \sum_{j=1, j \neq i}^{N_\mu} \boldsymbol{\mu}_i \mathbf{T}_{ij} \boldsymbol{\mu}_j, \quad (1)$$

with the dipolar tensor defined as:

$$\mathbf{T}_{ij} = \frac{1}{4\pi\epsilon_0} \left( \frac{f_5(r_{ij})}{r_{ij}^5} \begin{bmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{bmatrix} - \frac{f_3(r_{ij})}{r_{ij}^3} I_3 \right), \quad (2)$$

Here,  $I_3$  is the identity matrix. The  $f_5$  and  $f_3$  functions are introduced to account for short-range damping effects to prevent the so-called "polarization catastrophe", proposed by Thole<sup>S1</sup> as:

$$f_3(r_{ij}) = 1 - \exp(-a_{ij}r_{ij}^3), \quad (3)$$

$$f_5(r_{ij}) = 1 - (1 + a_{ij}r_{ij}^3)\exp(-a_{ij}r_{ij}^3), \quad (4)$$

where  $a_{ij}$  is an adjustable damping factor, a parameter which depends on the nature of atoms  $i$  and  $j$ . In the present study, the damping factor was fixed at 0.3 at the beginning for all polarizable sites and fitted when needed for a better reproduction of the long-range interaction (See Table S1), the 0.3 value was chosen arbitrary.

# Intramolecular potential and Interaction energy curves

Herein, we reported the energy profiles of the dihedral angle scans in the dodecane and just some of the dimer structures with the corresponding interaction energy curves as a function of distance between the two molecules. In total, we fitted over 40 other structures and overall the MM energy curves fits well all the QM reference data.

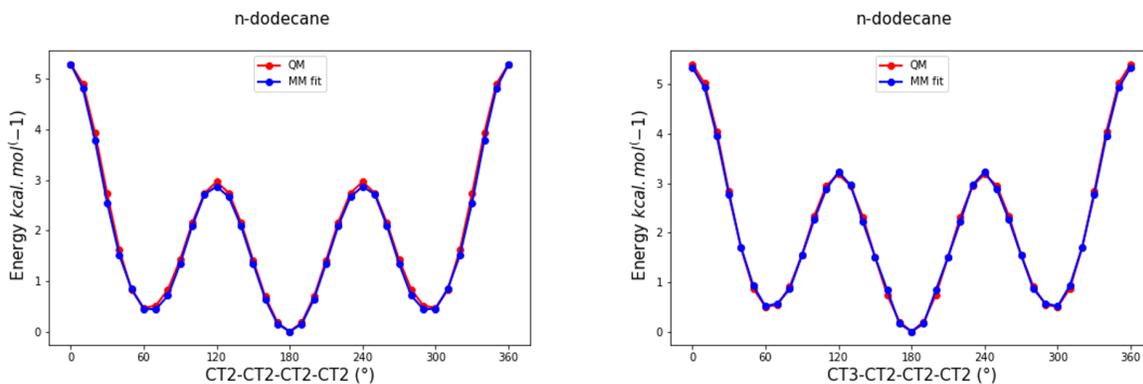


Figure S1: Relative energy profiles of the dihedral angle scans in dodecane (CT2-CT2-CT2-CT2 middle torsion on the left and CT3-CT2-CT2-CT2 terminal torsion on the right). The red lines represent the MP2 QC curves, while the blue ones the fitted FFs.

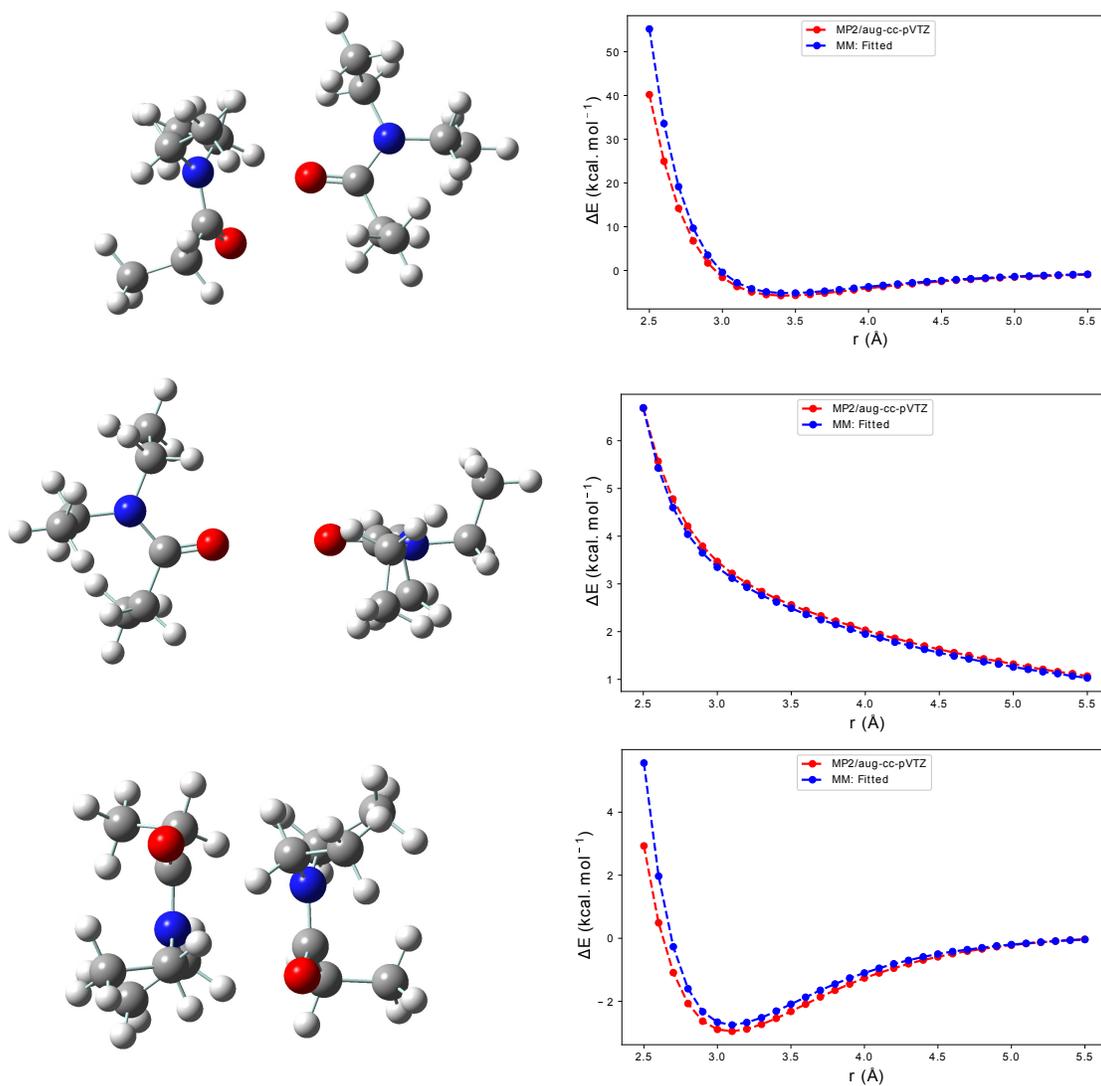


Figure S2: Selected DEPA dimer structures on the left with the corresponding interaction energy curves as a function of distance between the two molecules. MP2 values with aug-cc-pVTZ basis sets are in red and the fitted FF in blue.

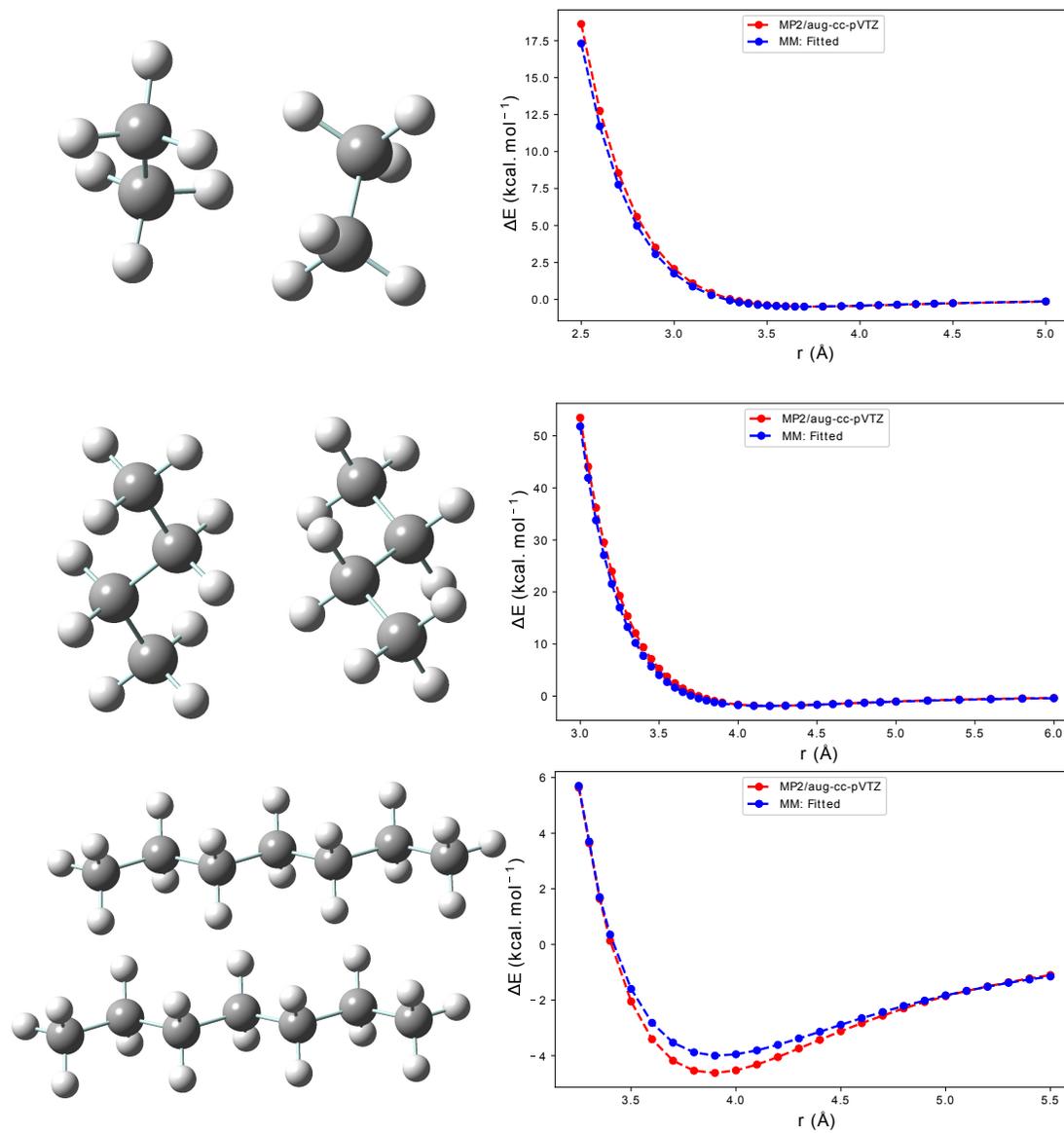


Figure S3: Selected dimer structures (for ethane, butane, and heptane) opposite to their interaction energy curves as a function of distance between the two molecules. MP2 values with aug-cc-pVTZ basis sets are in red and the fitted FF in blue.

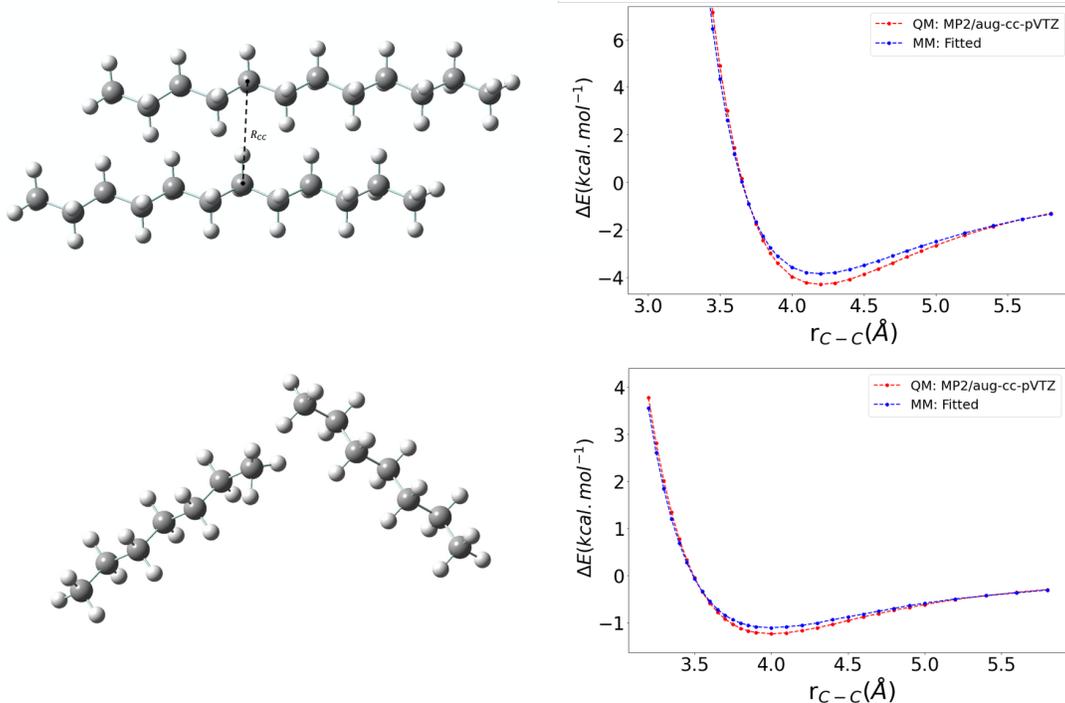


Figure S4: Selected dimer structures for dodecane opposite to their interaction energy curves as a function of distance between the two molecules. MP2 values with aug-cc-pVTZ basis sets are in red and the fitted FF in blue.

## Error estimation

The approach used to estimate the errors in this work is based entirely on the statistical nature of the results. Assuming that a MD simulation is performed for a total period of  $t_{tot}$  (production time). We select the last 20% of the production (noted  $t_{st}$ , we can divide  $t_{st}$  into  $N$  segments with the end point of each segment being  $t_i = i\Delta t$  (with  $i = 1, 2, \dots, N$ ) where  $\Delta t = t_{st}/N$ . Any time-averaged property can be calculated for each of the time intervals  $\Delta t_i = t_i - t_{i-1} = \Delta t$ , and as a result, each MD simulation will produce  $N$  values of the property  $P$ . If we denote each estimate of  $P$  to be  $P_i$  (with  $i = 1, 2, \dots, N$ ), the best estimate of the property can be calculated as :

$$\bar{P} = \frac{\sum_i P_i}{N} \tag{5}$$

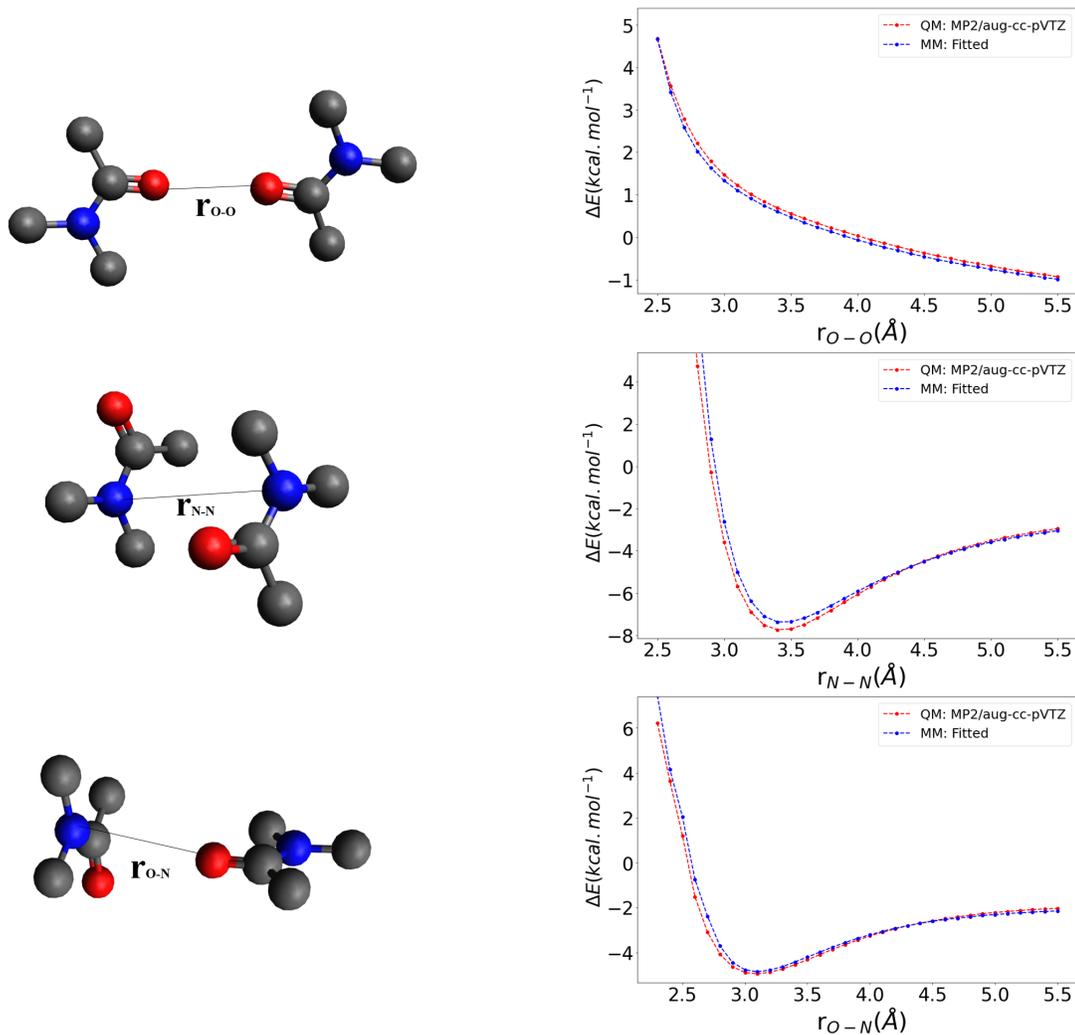


Figure S5: Selected dimer structures for DEHBA opposite to their interaction energy curves as a function of distance between the two molecules. MP2 values with aug-cc-pVTZ basis sets are in red and the fitted FF in blue.

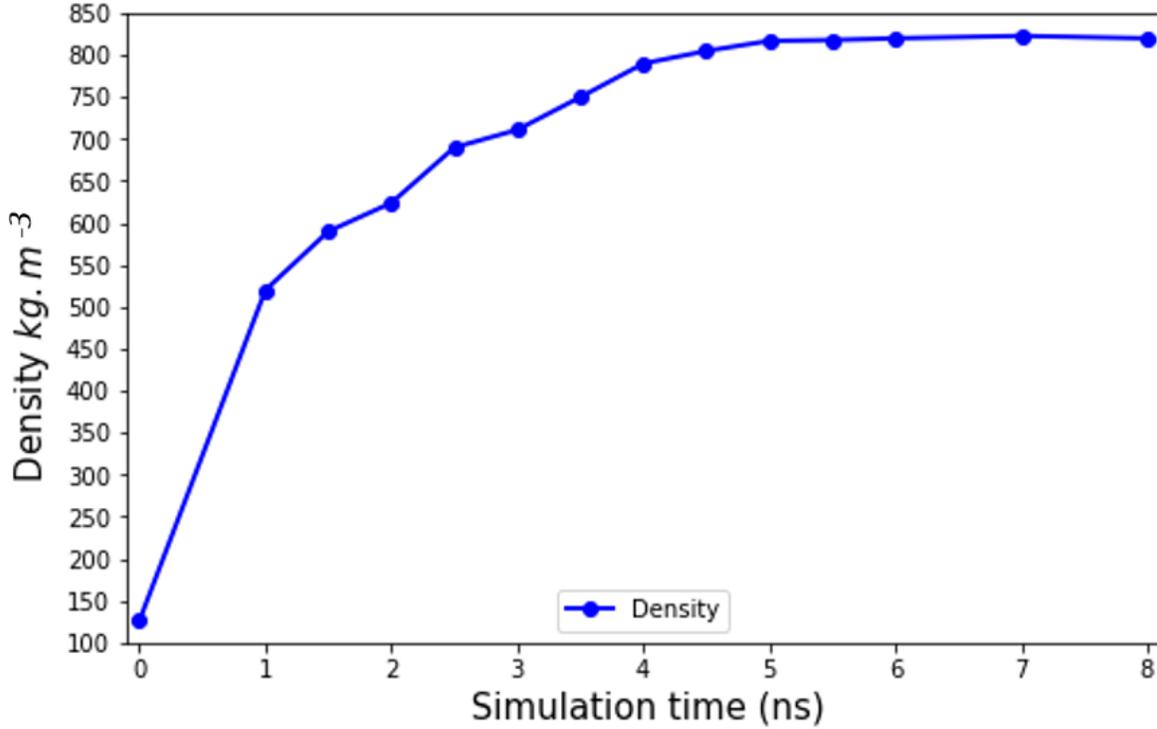


Figure S6: Time evolution of the density for 26% DEHBA in dodecane during the equilibration phase up to 8 ns.

The uncertainty of the samples  $P_i$  was quantified "in our work" by the sample standard deviation defined as :

$$\sigma = \sqrt{\frac{\sum_i (P_i - \bar{P})^2}{N - 1}}. \quad (6)$$

The estimation of  $P$  also associated with an uncertainty  $\Delta P = \sigma$  is written as :

$$P = \bar{P} \pm \Delta P. \quad (7)$$

In our case, the uncertainties for the density ( $\rho$ ), heat of vaporization ( $\Delta H_{vap}$ ), the excess volume ( $\Delta V_{ex}$ ) and excess enthalpy ( $\Delta H_{ex}$ ) were evaluated by considering the uncertainty on volumes and energies. The two latter being weighted sums of several variables  $x, y, z$ , the

errors can be estimated by :

$$\sigma_f = \sqrt{\left(\frac{\partial f}{\partial x} \cdot \sigma_x\right)^2 + \left(\frac{\partial f}{\partial y} \cdot \sigma_y\right)^2 + \left(\frac{\partial f}{\partial z} \cdot \sigma_z\right)^2}, \quad (8)$$

where  $\sigma_f$ ,  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_z$  are the absolute errors associated to the observable  $f, x, y$  and  $z$ , respectively.

Table S4: Impact of the simulation box size on ligands properties, namely the density ( $\rho$ ) and heat of vaporization ( $\Delta H_{vap}$ ). Experimental values are taken from Haynes *et al.*<sup>S2</sup> Relative deviations  $\delta$  (in %) and absolute deviation  $\Delta$  (in  $\text{kcal} \cdot \text{mol}^{-1}$ ) with respect to experimental values are reported.

Ligand	$\rho$ ( $\text{kg} \cdot \text{m}^{-3}$ )				$\Delta H_{vap}$ ( $\text{kcal} \cdot \text{mol}^{-1}$ )			
	# Molecules				# Molecules			
	343	729	Expt.	$\delta(\%)$	343	729	Expt.	$\Delta(\text{kcal} \cdot \text{mol}^{-1})$
DMA	937	925	900	4-3	10.2	10.4	10.9	0.7
Butane	596	592	601	1-2	3.6	3.7	4.4	0.8
Heptane	700	690	677	3-2	7.7	7.8	8.6	0.9
DEHiBA	898	892	865	4-3	23.3	23.3		

The analysis of the data in Table S4 reveals that the maximum deviation between the two box sizes is 3% for density and  $0.2 \text{ kcal} \cdot \text{mol}^{-1}$  for the heat of vaporization, hence we conclude that the initial simulation boxes were adequately chosen, and that the box size effect is in our case is minor.

Table S5: Composition of DEHiBA/dodecane and DEHBA/dodecane mixtures, associated simulation boxes and densities  $\rho$  ( $\text{kg} \cdot \text{m}^{-3}$ ). L denotes the extractant ligand. Experimental values are taken from Ref. S3.

$C_{mono}$ ( $\text{mol} \cdot \text{L}^{-1}$ )		$\rho$ ( $\text{kg} \cdot \text{m}^{-3}$ )				# Molecules				
DEHiBA		DEHBA		DEHiBA		DEHBA				
Expt.	MD	Expt.	MD	Expt.	MD	Expt.	MD	L	dodecane	% L
0.40	0.42	-	0.43	765	-	800	802	34	316	10
1.00	1.04	-	1.07	790	-	822	828	92	258	26
1.50	1.44	-	1.47	811	-	841	850	149	201	42
2.00	2.07	-	2.03	832	-	863	875	217	133	62

Table S6: Coordination numbers derived from the radial distribution function of oxygen and nitrogen atoms, for the DEHiBA/dodecane and DEHBA/dodecane mixtures as a function of the  $x_{mono}$  mole fraction of monoamide.

$x_{mono}$	DEHiBA/dodecane		DEHBA/dodecane	
	(O–O)RDF	(N–N)RDF	(O–O)RDF	(N–N)RDF
0.10	0.67	0.65	0.63	0.63
0.26	1.23	1.22	1.17	1.20
0.42	1.54	1.55	1.51	1.50
0.62	1.97	1.93	1.92	1.91
1.00	2.40	2.43	2.40	2.40

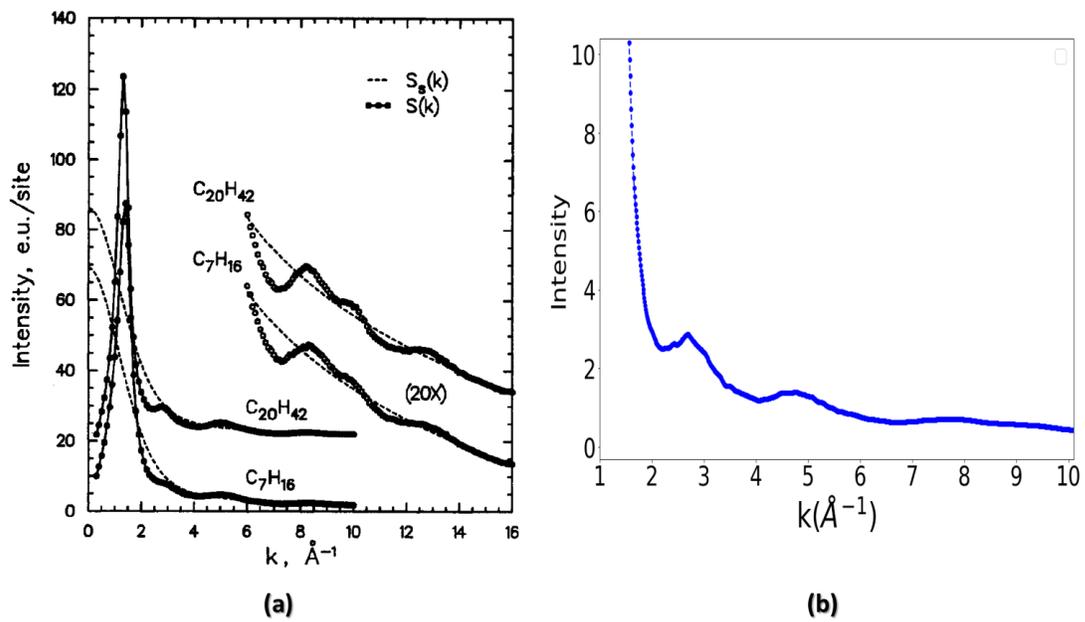
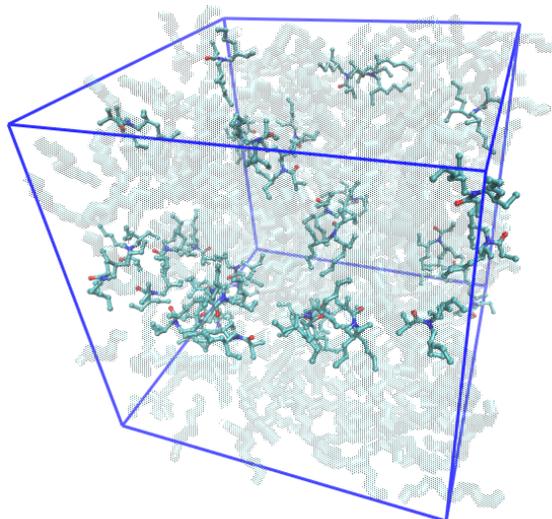
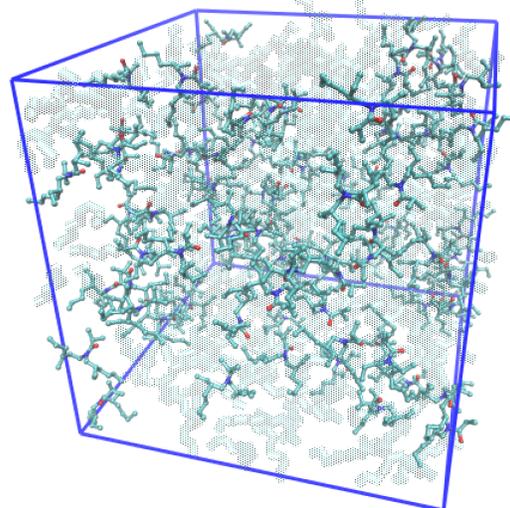


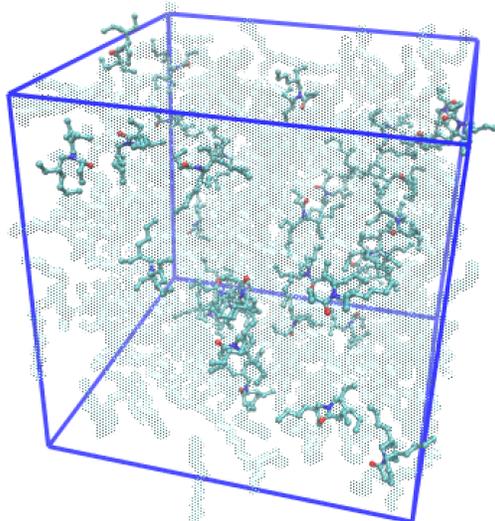
Figure S7: Structure factors for heptane, (a) Experimental data from Habenschuss *et al.*,<sup>S4</sup> (b) simulations using Debye code.<sup>S5</sup>



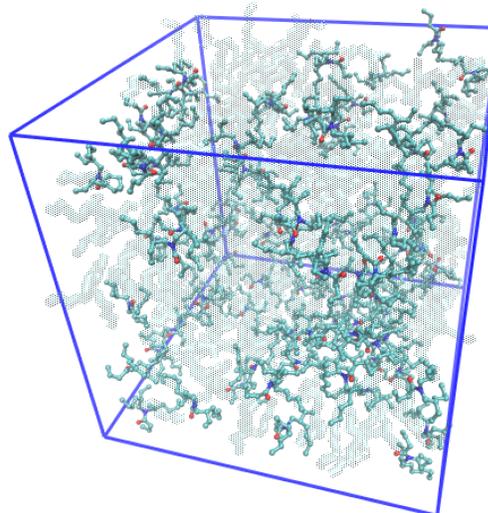
(a) 10% DEHiBA



(b) 26% DEHiBA

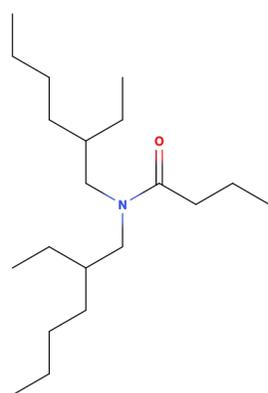


(c) 10% DEHBA

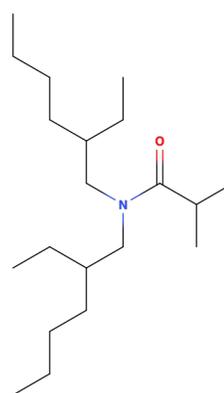


(d) 26% DEHBA

Figure S8: Perspective views illustrating self-assembly of the monoamides in DE-HiBA/dodecane (a–b) and DEHBA/dodecane (c–d) mixtures, with 10–26 % of monoamide concentrations. Dodecane is greyed for clarity.

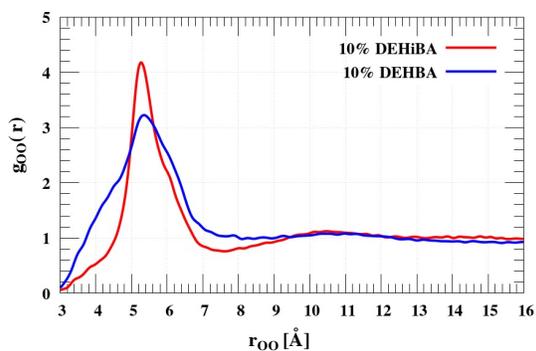


(a)

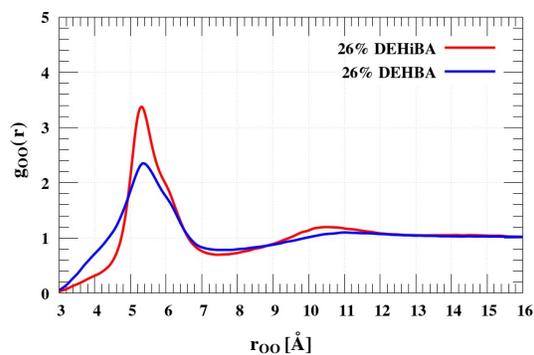


(b)

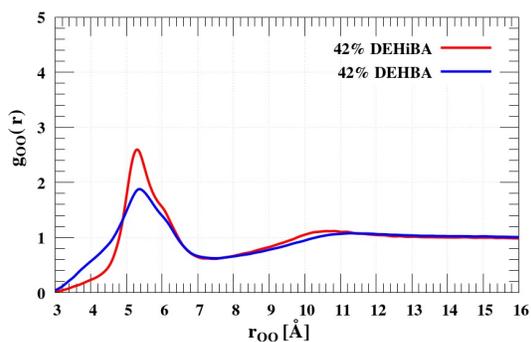
Figure S9: Structure of (a) di-2-ethylhexyl-butylamide (DEHBA) and (b) di-2-ethylhexylisobutyramide (DEHiBA).



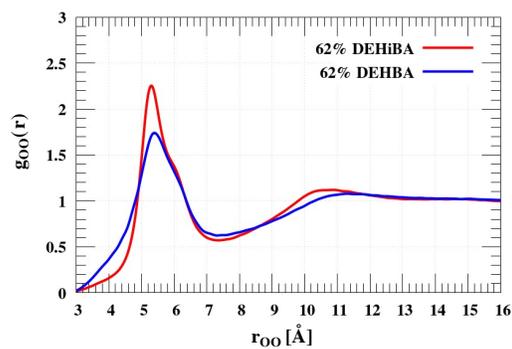
(a)  $x_{mono} = 10\%$



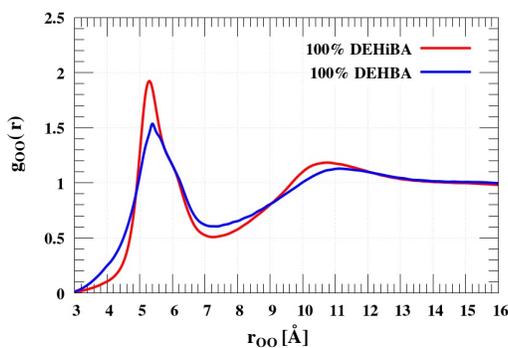
(b)  $x_{mono} = 26\%$



(c)  $x_{mono} = 42\%$

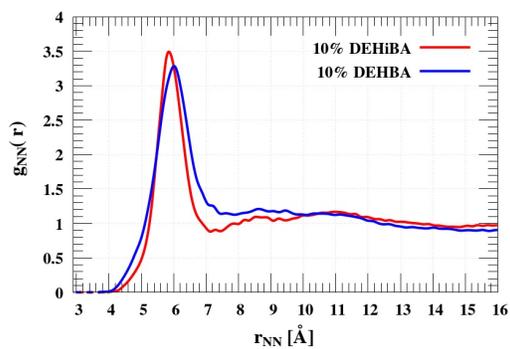


(d)  $x_{mono} = 62\%$

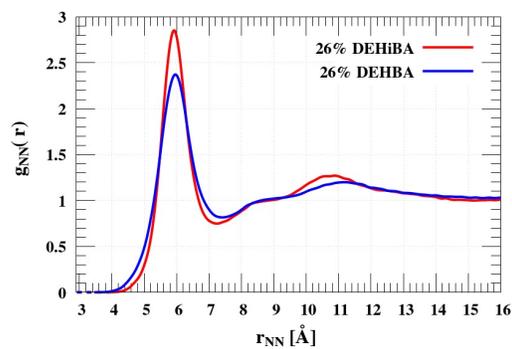


(e)  $x_{mono} = 100\%$

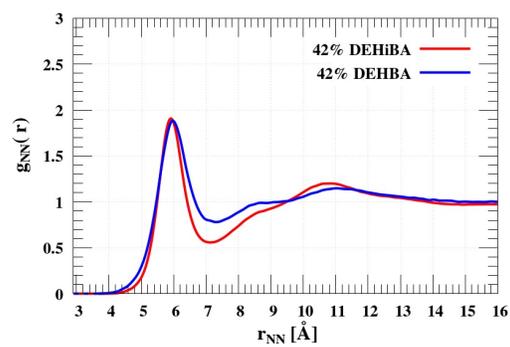
Figure S10: Radial distribution functions of oxygen atoms for DEHiBA/dodecane and DEHBA/dodecane mixtures at different monoamide  $x_{mono}$  mole fractions.



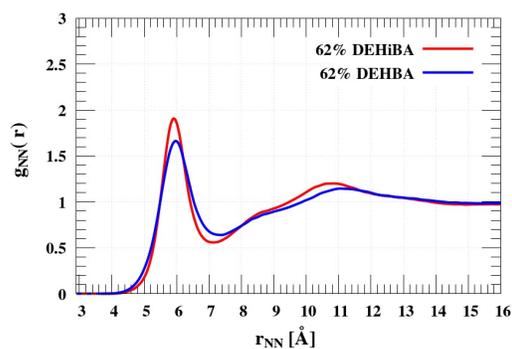
(a)  $x_{mono} = 10\%$



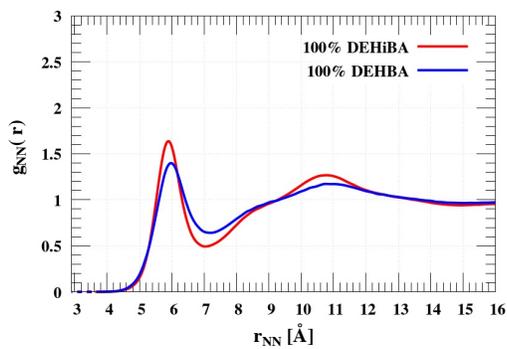
(b)  $x_{mono} = 26\%$



(c)  $x_{mono} = 42\%$

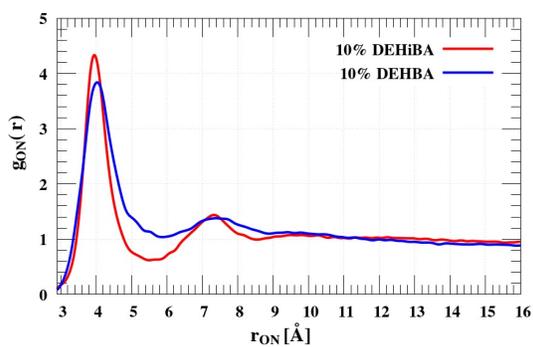


(d)  $x_{mono} = 62\%$

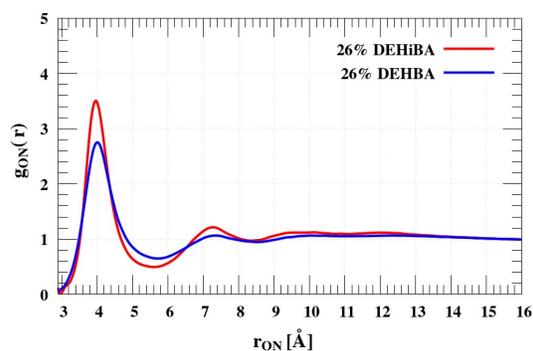


(e)  $x_{mono} = 100\%$

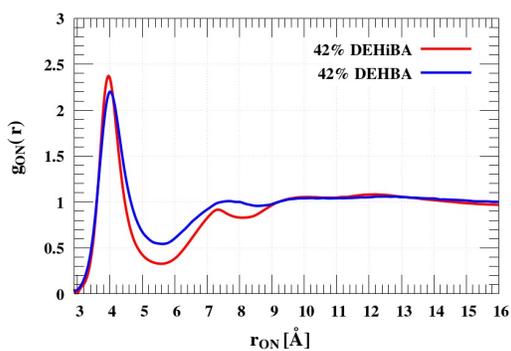
Figure S11: Radial distribution functions of nitrogen atoms for DEHiBA/dodecane and DEHBA/dodecane mixtures at different monoamide  $x_{mono}$  mole fractions.



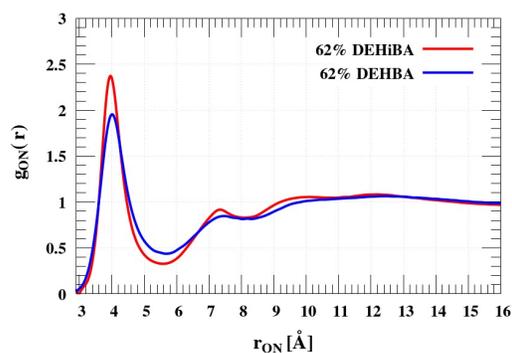
(a)  $x_{mono} = 10\%$



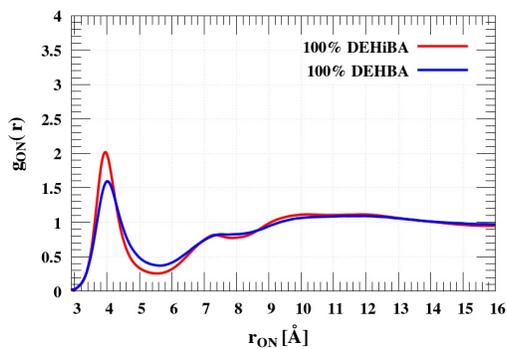
(b)  $x_{mono} = 26\%$



(c)  $x_{mono} = 42\%$

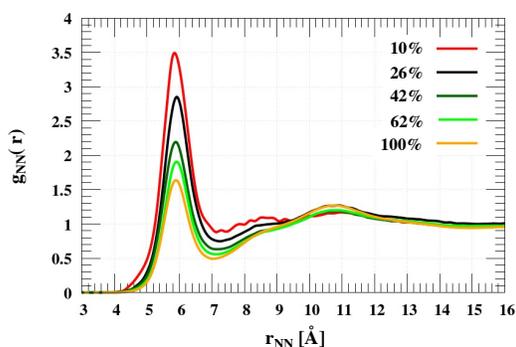


(d)  $x_{mono} = 62\%$

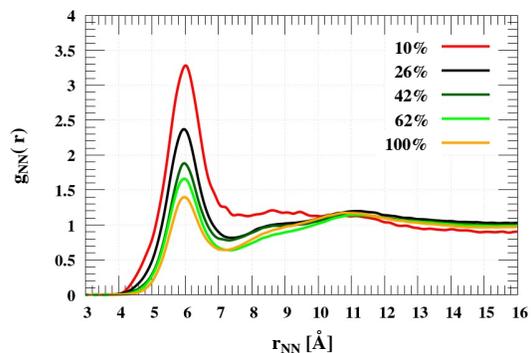


(e)  $x_{mono} = 100\%$

Figure S12: Radial distribution functions between oxygen and nitrogen atoms for DE-HiBA/dodecane and DEHBA/dodecane mixtures at different monoamide  $x_{mono}$  mole fractions.

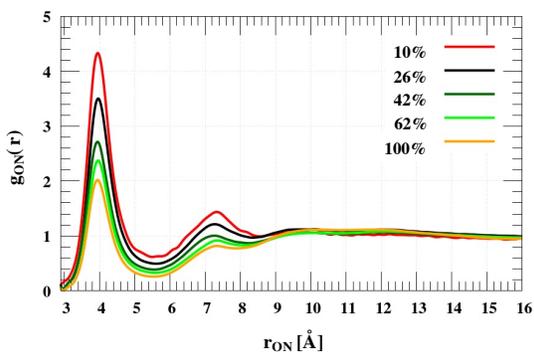


(a) DEHiBA/dodecane mixtures

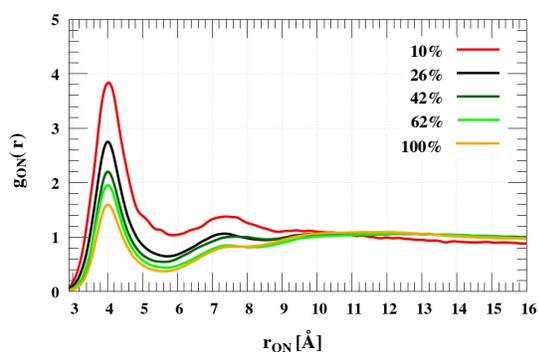


(b) DEHBA/dodecane mixtures

Figure S13: Radial distribution functions of nitrogen atoms for DEHiBA/dodecane and DEHBA/dodecane mixtures at at different monoamide mole fractions ( $x_{mono}=10, 26, 42, 62$  and 100 %).



(a) DEHiBA/dodecane mixtures



(b) DEHBA/dodecane mixtures

Figure S14: Radial distribution functions between oxygen and nitrogen atoms for DEHiBA/dodecane and DEHBA/dodecane mixtures at at different monoamide mole fractions ( $x_{mono}=10, 26, 42, 62$  and 100 %).

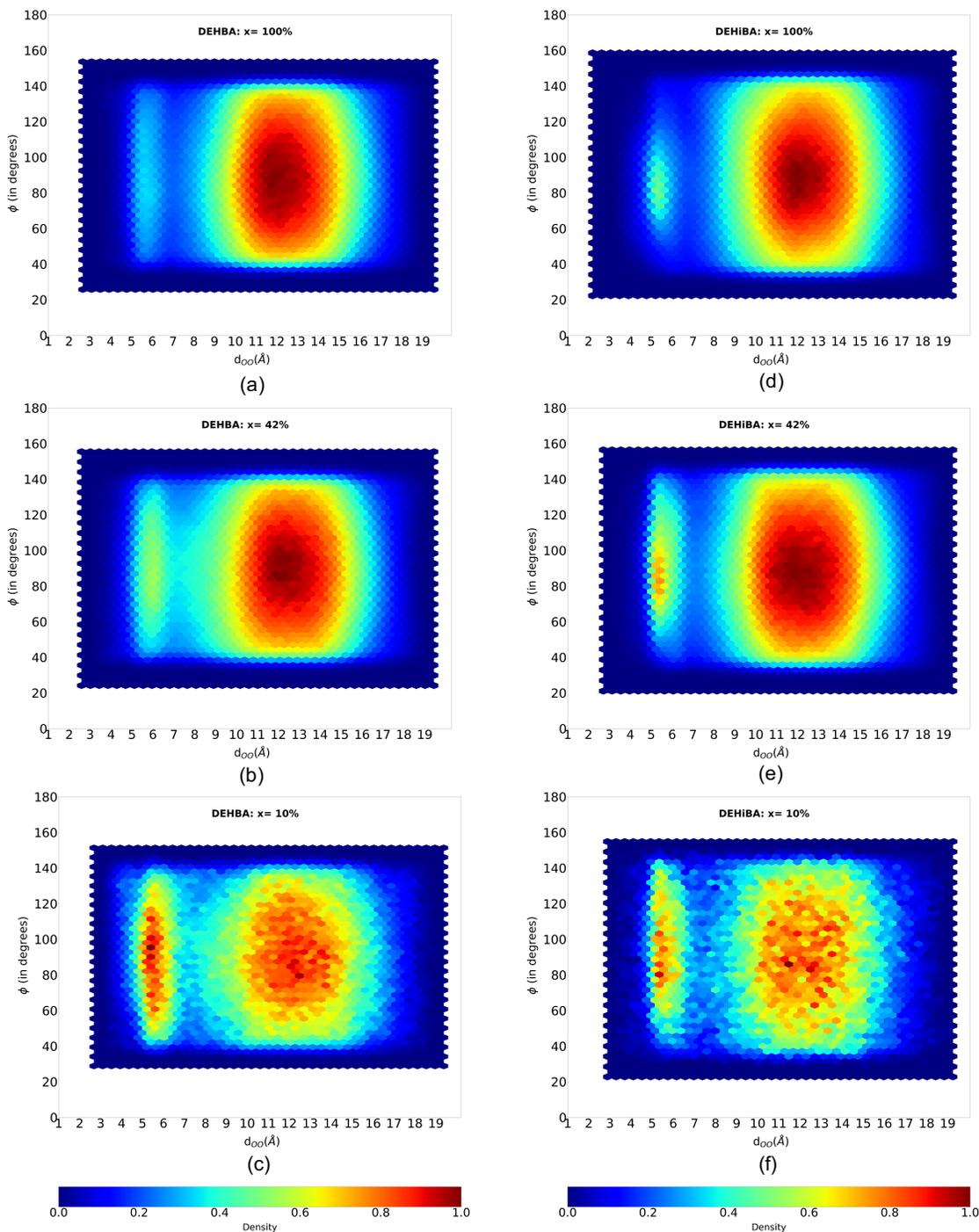


Figure S15: Analysis of the relative orientations of the NCO planes,  $\phi(r_{OO})$ , as a function of the distance between the oxygen of the carbonyl group for DEHBA/dodecane (left column) and DEHiBA/dodecane mixtures (right column). The color bar indicates the normalized count density. 100% DEHBA, (b) 42% DEHBA, (c) 10% DEHBA, (d) 100% DEHiBA, (e) 42% DEHiBA and (f) 10% DEHiBA. The color bar indicates the normalized count density.

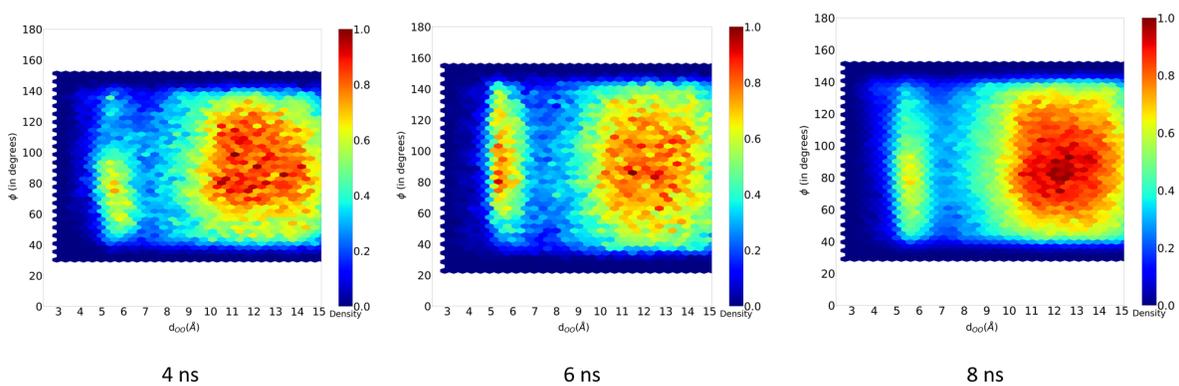


Figure S16: Analysis of the relative orientations of the NCO planes,  $\phi(r_{OO})$ , as a function of the distance between the oxygen of the carbonyl group for 26% DEHBA in dodecane averaged for 4 ns, 6 ns and 8 ns of the equilibration phase. The color bar indicates the normalized count density. The color bar indicates the normalized count density.

## References

- (S1) Thole, B. T. Molecular polarizabilities calculated with a modified dipole interaction. *Chem. Phys.* **1981**, *59*, 341–350, DOI: 10.1016/0301-0104(81)85176-2.
- (S2) Haynes, W. M., Ed. *CRC handbook of chemistry and physics*, 95th ed.; CRC press, 2014.
- (S3) Coquil, M.; Boubals, N.; Duvail, M.; Charbonnel, M.-C.; Dufrêche, J.-F. On interactions in binary mixtures used in solvent extraction: Insights from combined Isothermal Titration Calorimetry experiments and Molecular Dynamics simulations. *J. Mol. Liq.* **2021**, 116985, DOI: 10.1016/j.molliq.2021.116985.
- (S4) Habenschuss, A.; Narten, A. H. X-ray diffraction study of some liquid alkanes. *J. Chem. Phys.* **1990**, *92*, 5692–5699, DOI: 10.1063/1.458500.
- (S5) Debyer. <https://debyer.readthedocs.io/en/latest/>.