Prediction of Solute Partition Coefficients between Polyolefins and Alcohols Using a Generalized Flory–Huggins Approach

Guillaume Gillet,†,‡ Olivier Vitrac,*,§ and Stéphane Desobry‡

Centre Énergie, Matériaux et Emballage, Laboratoire National de métrologie et d’Essais, 29 avenue Roger Hennequin, 78197 Trappes Cedex, France, LSGA-ENSAIA-INPL, Nancy Université, 2 avenue de la forêt de Haye, BP 172, 54505 Vandœuvre lès Nancy, France, and UMR 1145 Génie Industriel Alimentaire, Institut National de la Recherche Agronomique, 1 avenue des Olympiades, 91300 Massy, France

The partition coefficients of n-alkanes, n-alcohols, volatiles, and typical antioxidants between a low density polyethylene and several alcohols (methanol and ethanol) were predicted without fitting variables using an off-lattice Flory–Huggins approach. The main advantage of the proposed excluded volume constraint method was to sample at atomistic scale pairwise contact energies and the residual entropic contribution due to possible different conformers and reorientations. The positional entropy was considered both in the polymer and in the liquid as in the original Flory–Huggins theory. Possible biases due to the formulation and conformational sampling were analyzed for molecules different in size, in shape, and in stiffness. The predictions were close to the experimental uncertainty on the partition coefficients between the amorphous part of the polymer and the tested liquids. The present work confirmed that plastic additives, which are by design highly compatible with the polymer, had also a significant chemical affinity with polar liquids consisting of smaller molecules. Finally, a general predictive model of partition coefficients based on solubility coefficients was proposed.

1. Introduction

Solute partition coefficients between thermoplastics and liquids are used to predict mass diffusion1–5 in various fields including protective clothing,6 biomedical studies,7 separation techniques,8–10 and food packaging.11–13 Several recent regulations in EU14 regarding the safety of chemicals, so-called REACH directives, urge the development of predictive approaches on the migration of chemicals into the environment or through the food chain. This work was motivated by the EU regulation 1935/2004/EC15 specific for food contact materials, which enforces a safety assessment and risk management decision for all starting substances and possible degradation products coming from the material. For plastic materials, article 14 of directive 2002/72/EC16 introduces diffusion modeling as an alternative to time-consuming and costly experiments for both compliance testing and risk assessment. The principles for probabilistic modeling of packaging substance desorption into food was previously analyzed17 and applied to different situations.18–20 The main limitation in predictive approaches for both compliance testing and risk assessment is the availability of physicochemical properties for a wide range of substances and polymers under particular thermodynamic conditions (temperature, swelling). Whereas robust approaches have been developed for the prediction21–24 or overestimation25–27 of diffusion coefficients in polyolefins, no appropriate method exists to predict partition coefficients between a polyolefin and a food product or simulated food.28 A first significant attempt was made by Baner and Piringer28 using an approach mixing both the regular solution theory on the food side and Flory–Huggins approach on the polymer side. To fit experimental results, the predictions were corrected by an empirical factor, which was found to be dependent on the size and shape of the considered substance. On polymer side, Flory–Huggins interactions parameters are conventionally derived from the separation of the square root of the polymer cohesive energy29 into a dispersive contribution (short-range dispersion forces) and a long-range Coulombic contribution, which encompasses arbitrary a polar term and a term due to hydrogen bonding.30,31 Recent reviews32,33 on predictions of activity coefficients in polymer solutions showed that (i) evaluations from solubility coefficients led systematically to strong overestimations and (ii) consequently solubility coefficient decomposition should be empirically reweighted by factors less than 1. Alternative methods to predict solute activity coefficients rely on Metropolis Monte Carlo (MC)34 or molecular dynamics (MD).35,36 Conventional strategies consist in measuring the excess in free energy by molecular dynamics when the guest solute is inserted37 or removed38 into the host phase. Such methods converge rapidly for solutes smaller than accessible voids but may be prohibitive when the reordering dynamics of the host phase is required to accommodate the solute insertion/deletion. As the complete relaxation of entangled polymer segments is not tractable, the free energy difference before and after insertion/deletion can be advantageously replaced by a thermoodynamic integration (TI) along an arbitrary path39,40 possibly combined with a perturbation protocol.41,42 In complex solute mixtures, MC techniques working in the Gibbs ensemble43 may alternatively be used. Although recent developments in MC and MD sampling methods are promising for large solutes, no study has been published on a large set of bulky and hindered solutes in polymers and in liquids. In particular, all methods require ad hoc equilibration steps via molecular dynamics, which are computationally expensive for large systems involving few solutes and a much larger number of host molecules.

The current work examines a tailored MC sampling methodology based on contact energies instead of an explicit description of the binary mixture to assess, at a reasonable computational cost, excess chemical potential in polyethylene amorphous regions and in food simulants with increasing polarity (ethanol and methanol). Because reference data on excess chemical potentials are very scarce, predictions were...
compared with macroscopic partition coefficients previously measured on different classes of solutes: volatile compounds, homologous series of \( n \)-alkanes and \( n \)-alcohols, and small phenolic antioxidants. In this context and due to inherent biases in heterogeneous experimental data, our objective was to achieve predictions not far off the mark and within typical experimental errors on partitioning while taking into account the effects of crystallinity and the density of the polymer. The choice of a generalized Flory–Huggins approach was motivated by the apparent symmetry between a bulky solute dispersed among much larger segments and a bulky solute dispersed among much smaller molecules constituting the liquid simulant. In both cases, the positional entropy and the configurational entropy are expected to spread the mixed-segment and mixed-simulant interaction energies. For simple mixtures of molecules, such as \( n \)-alkanes in polyethylene, with similar chemical structures (i.e., without a significant van der Waals enthalpy) but different shapes, the Flory–Huggins theory, which calculates the positional entropy on a rigid lattice, describes remarkably their thermodynamics. The necessary condition is that the volume per mole of lattice sites is fixed. Elbro et al. demonstrated that an expression—alogous to the Flory–Huggins combinatorial term—can be derived irrespective of the considered substance shape (i.e., independent of any lattice frame of reference) by considering the volume fraction and free volumes of each substance in the mixture. A systematic derivation of the Flory–Huggins theory, which takes the free volumes into account, has been proposed by Bawendi et al. In nonideal mixtures, the heat of mixing can be approximated by the van Laar expression of first neighbor contact energies. Because the original Flory theory captures pair interaction terms only on a coarse grain scale, the corresponding mean field approximation may appear too coarse to describe the interactions with very dissimilar species or presenting different flexibilities. Similar arguments have been used to explain the large discrepancy in predictions by solubility coefficients derived from different group contributions (atomic, group based, and blob based). The main reason is that all pair configurations will not have equal probabilities at the atomic scale; moreover, the residual entropy should be included in the mean field approximation. From these considerations, both the atomistic description of interactions and the off-lattice approximation of energetic interactions taking into account the packing of atoms are expected to provide more accurate estimates of the enthalpic contribution.

The paper is organized as follows. The Theoretical Background section introduces the concepts used to derive estimates of solute partition coefficients between a semicrystalline polymer and a liquid in the framework of Flory–Huggins theory at an atomistic scale. The conditions of simulation and validation are detailed in section 3. In section 4, the simulation results are analyzed: first, by comparing the sampling biases calculated for homologous or typical solutes; second, by comparing with experimental partition coefficients between alcoholic simulants homologous or typical solutes; second, by comparing with experimental partition coefficients between alcoholic simulants. In section 4, the simulation results are analyzed: first, by comparing the sampling biases calculated for homologous or typical solutes; second, by comparing with experimental partition coefficients between alcoholic simulants.

We summarize our conclusions.

### 2. Theoretical Background

Before we introduce the concepts of the computational method used in this work, we briefly summarize the general assumptions, which support the definition of a thermodynamic equilibrium between a thermoplastic packaging, denoted \( P \), and a liquid in contact, denoted \( L \). For concision, the initial definition assumes that \( P \) is amorphous (no crystalline phase) and that the solute, denoted \( i \), is initially well dispersed in \( P \). A more general definition is subsequently proposed.

#### 2.1. Thermodynamic Equilibrium between the Packaging Material and the Liquid in Contact

The thermodynamic equilibrium between \( P \) and \( L \) is reached when the cumulated free energy between \( P \) and \( L \), \( G_{P,L} \), is minimal. At constant temperature and pressure, the evolution toward equilibrium is accompanied by a mass transport between \( P \) and \( L \). In this work, only the solute molecules are assumed to migrate (Figure 1a). These assumptions are particularly well satisfied when \( j \) migrates in an aqueous food or polar solvent. For an isolated system \( P + L \), \( G_{P+L} \) is:

\[
G_{P+L} = G_{i,i} + G_{i,L} = N_i \mu_i + N_L \mu_L + \sum_{j \neq i} \frac{N_j \mu_j + N_j \mu_j^L}{G_{i,j}^{L} + G_{i,j}^{P}}
\]

where \( \{N_i\}_{j=P,L} \) and \( \{N_j\}_{j=P,L} \) are the number of molecules \( j \) and the number of molecules \( i \) in \( j \), respectively. Their corresponding chemical potentials are \( \{\mu_i\}_{i=P,L} \) and \( \{\mu_j\}_{j=P,L} \), respectively. All energetic terms are further decomposed into an ideal part (id) and an excess part (excess). From a microscopic point of view, the detailed mass balance at the interface enforces that the partition coefficient at the interface between \( P \) and \( L \), noted \( K_{iLP} \), is equal to the ratio of presence probabilities from each side of the interface, \( \{p_i\}_{i=L,P} \), and to the ratio of frequencies of crossing the interface: \( k_{i,P,L} \) and \( k_{i,L,P} \). By introducing the free energy of the barrier to cross the interface, \( G^i \), the transition state theory defines \( K_{iLP} \) in the canonical ensemble as:

\[
K_{iLP} = \frac{p_i}{p_i} = \frac{k_{i,P}^{L}}{k_{i,L}^{P}} = \frac{k_{i}^{L} T}{k_{i}^{P} T} \exp \left( \frac{G_{i}^{L} - G_{i}^{P}}{k_{B} T} \right)
\]

with \( h \) Planck’s constant, \( k_{B} \) Boltzmann’s constant, and \( T \) the absolute temperature. Equation 2 can be used to estimate \( K_{iLP} \) in the Gibbs ensemble, but it requires calculating the energy of each subsystem after equilibration. An alternative relies on a macroscopic description of equilibrium (\( dG_P + dL = 0 \)) for a closed system (\( dN_L = -dN_P \)), which leads to \( \mu_L = \mu_P \). By choosing the state of pure \( i \) as reference and by expressing the activities of both nonideal mixtures from their volume fractions in \( i \), \( \{\phi_i\}_{i=P,L} \), \( K_{iLP} \) is approximated as:

\[
K_{iLP} = \frac{V_i \phi_i}{V_P \phi_P} = \frac{\mu_i^{\text{excess}} - \mu_i^{\text{excess}}}{k_{B} T} = \frac{\gamma_i^{L}}{\gamma_i^{P}}
\]

with \( \{\gamma_i^{L}\}_{i=P,L} \), the activity coefficient of \( i \) in \( j \) and \( V_i \), the partial specific volume of solute \( i \). Solutes tend to concentrate in the phase for which they have the greatest chemical affinity. From the mixing point of view and irrespective of specific molecular interactions,
solute tends to migrate in the phase where diluting the system with solutes causes the largest gain in entropy (i.e., the largest increase in the number of microconfigurations). This "attractive" effect is expected to be higher for the phase consisting of molecules smaller than the solute \( i \) \((i.e., L \text{ rather than } P)\). It is however counterbalanced by the cohesive energy associated with the creation of a cavity with a volume similar to that of the solute. Indeed, increasing the cohesion of one phase (e.g., due to hydrogen bonding) generates a resistance to the insertion of the solute and decreases consequently its frequency of occurrence. For a rubber amorphous phase, the cohesive energy of the system is expected to be lower and the partitioning to small solutes controlled mainly by the entropic contribution.

2.2. Additional Assumptions. Since the diffusion of a substance \( i \), such as a plastic additive, requires polymer cooperative motions, the substance \( i \) is assumed to diffuse only in the amorphous phase of the polymer. The partition coefficient between \( L \) and the polymer amorphous phase, \( K_{L/P} \), is therefore

\[
K_{L/P} = \frac{C_{L,eq}}{C_{P,eq}} = (1 - c) \frac{\hat{C}_{L,eq}}{a_i \hat{C}_{P,eq} - b_i C_{P,eq}}
\]

where \( C_{L,eq} \) is the concentration in \( L \) at equilibrium; \( C_{P,eq} \) is the concentration in the amorphous phase of \( P \) at equilibrium; \( \{\hat{C}_{i,eq}\}_{i=L,P} \) is the macroscopic concentration in each phase as experimentally assessed; \( C_{P,eq} \) is the initial concentration in \( P \) (i.e., before contact with \( L \)); \( c \) is the polymer crystallinity (volume fraction of crystalline phase); \( a_i \) is the extraction yield of \( i \) from \( P \); \( b_i \) is the amount of substance \( i \), which is not "well-dispersed" in \( P \). Additives with a melting point higher than the extrusion temperature (i.e., "low melting additives") or any process of hygroscopic additives (e.g., phenolic and phosphate

Figure 1. Multiscale description of interactions (b, c) between \( P \) and \( i \), and (d, e) between \( L \) and \( i \). (b, d) Mean-field approximation on a lattice; (c, e) description of interactions between first neighbors.
antioxidants) or of any nucleating substance are known to lead to incomplete dispersion in the polymer matrix. From the thermodynamic point of view, substances that are not dispersed or are poorly dispersed have a much higher chemical potential than well-dispersed substances (i.e., their activity is closer to the pure substance chemical potential, which is 1 by definition). The driving force associated with their migration is consequently higher, but desorption occurs less rapidly due to kinetic limitations (melting, trapping, and strong hydrogen bonding).

As a result, thermodynamic equilibrium described in eq 3 is assumed to be only related to the "well-dispersed" substance (at infinite dilution). From previous considerations, it is expected that the current definition of the partitioning may differ from the apparent or macroscopic partition coefficient, $K_{L/IP}$, and assessed macroscopically:

$$K_{L/IP} = \frac{\hat{C}_{i/LP}}{\hat{C}_{i/P}}$$

Definitions of partitioning derived from eqs 4 and 5 are related through the macroscopic mass balance between L and P:

$$C_{i/LP} = \frac{1}{a_i} C_{i/LP} + (1 - 1/a_i) C_{i/LP}$$

where $l$ is the dilution factor or equivalently the ratio between the volume of P and the volume of L. The relationship between $K_{L/IP}$ and $\hat{K}_{L/IP}$ is given by

$$\frac{K_{L/IP}}{\hat{K}_{L/IP}} = (1 - c) \frac{1}{1 - b_i} \frac{1}{l} \frac{L}{L_i} - (1 - c + K_{L/IP} l/b_i)$$

Since most authors assess $\hat{K}_{L/IP}$ instead of $K_{L/IP}$, comparisons between experimental and predictions from several sources must take into account the variability in the polymer material (crystallinity and density), its formulation (estimate of $b_i$), and experimental conditions ($l$ and $a_i$). When they are not available, they must be estimated from reasonable assumptions including a sensitivity analysis of these parameters. For a practical use, two extreme behaviors are worth identifying: When $b_i K_{L/IP} l \ll 1$ (e.g., $K_{L/IP} \ll 2.5$ for common experimental conditions of $l = 1/20$ and $b_i = 0.02$), $(1 - c) - K_{L/IP}$ is a constant (greater than 1), which does not depend on the true value $K_{L/IP}$. Trends for homologous molecules series and ratios obtained in similar experimental conditions can therefore be directly compared to predictions. By contrast, when $K_{L/IP} \gg L/b_i$ (e.g., $K_{L/IP} \gg 2.5$), the discrepancy between $K_{L/IP}$ and $\hat{K}_{L/IP}$ tends to be maximum. For $K_{L/IP}$ values greater than 1, $(1 - c) K_{L/IP}$ provides an underestimate of $K_{L/IP}$, whose bias $K_{L/IP} - (1 - c) K_{L/IP}$ increases as the root of $K_{L/IP}$. For these conditions, reference data are required to provide valuable comparisons with predictions. In other words, previously published partition coefficients with values much less than 1 are more valuable to validate a predictive model than larger values. Similarly, underestimating the "true" $K_{L/IP}$ may lead to a significant risk underestimation for liquid contamination.

2.3. Flory–Huggins Theory of Mixing. According to the Flory–Huggins theory, the expression of excess chemical potentials can be derived from local interactions between neighboring molecules since the conformations of both polymer segments and liquid medium molecules are not perturbed by surrounding ones beyond a critical distance. In the polymer, it is therefore assumed that excluded volume among nonconcentric monomers (belonging to the same chain or not) do not affect the overall chain conformations. A more general discussion can be found. In this framework, the entropies of mixing corresponding to mixtures $P + i$ and $L + i$ are defined as the sums of a combinatorial term and a residual term. The combinatorial term is assessed at a coarse grain level, the so-called blob size. The residual term depends on the detailed conformations of each molecule and on the dispersion of the number of contacts between pairs of molecules. When L is a noninteracting liquid, the equilibration of chemical potentials between P and L may be complicated by a secondary transport of L in P, so that the ternary mixture $P + L + i$ is involved with an additional constraint release in P. This secondary transport of L in P was neglected in this study, since only fluids (i.e., polar liquids) with high cohesive energies were considered.

The equivalent representation of interactions between P and i on one hand and between L and i on the other hand are illustrated in Figure 1. The reference mesh size or blob size corresponds either to the partial specific volume of the solute, $V_r$, or to a volume which exposes a similar contact surface. $P$ is represented by many interconnected lattice sites (Figure 1b,c), whereas a single site represents many L molecules (Figure 1d,e). Because the end-chain effects are neglected, P or L sites are supposed randomly dispersed on the lattice without taking into account their real connectivity. We can then notice that the same argument and mathematical expressions can be used in P as well as in F. On simple lattices, the number of neighbors does not depend on the blob size. Following the arguments in ref 61, two different values of $V_r$ were compared in this work: the volume of van der Waals, denoted $V^w$, calculated as the volume enclosed within the Connolly surface for a spherical probe of 0.1 Å with a grid step of 0.15 Å, and the molar volume $V^M$. For molecules j with a volume fraction $\phi_j$ and consisting of $r_j$ blobs of volume $V_r$, one gets at infinite dilution of i:

$$\frac{(\mu_{i,j} - \mu_{i,j}^0)}{k_B T} = \ln \chi_{i,j} - \ln (1 - \frac{r_j}{V^w}) - \ln \frac{\gamma_{i,j}}{\gamma_{i,j}^w} = \left(1 - \frac{1}{r_j}\right) + \chi_{i,j}$$

where $\chi_{i,j} = H_{i,j}^\text{excess}/k_B T$ is the heat of mixing. The first term represents the effect of the configurational entropy associated with the increase of microstates due to the distribution of j around i. The absolute value of the first term is expected to be small in polymers ($r_p \gg 1$), while it is expected to be significant in liquids consisting in molecules much smaller than i. The term $r_j - 1$ is equivalent to the cavity dimensionless size (relative to the size of j) required to insert one molecule i. Since the size of P is large compared to i ($r_p \gg 1$), then the simplification $1/r_p \rightarrow 0$ is reasonable. Equations 3 and 8 lead to the following model for the partition coefficient:

$$\ln K_{L/IP} = \ln \gamma_{i,j}^L - \ln \gamma_{i,j}^P = \frac{r_j - 1}{r_j} + \chi_{i,j} - \chi_{i,j}^L$$

It is worth noting that the exact meaning of $r_j - 1$ is difficult to establish as it may vary according to the nature of the host L. The lower bound is necessarily the number (possibly fractional) of L molecules to be displaced to insert a volume corresponding to the van der Waals volume of i. In interacting liquids, the upper bond depends on the pressure—volume work done by the bulk liquid when it loses one or several hydrogen bonds at the interface with the solute. As an example in water, a study showed that the reordering of L was dependent on the solute and led to low-density ordered water in the vicinity of the solute.
2.4. Estimation of $\chi_{\text{AA}}$ from Pairwise Contact Energies.

Our off-lattice method to calculate excess chemical potentials in $P$ and $L$ is derived from the methodology described by Fan et al.\textsuperscript{65} Its main advantage is to take into consideration the excluded volume in pairwise interactions and in the coordination number. Indeed, the blob size and the coordination number cannot be modified in original lattice methods, so the lattice approximation is less accurate for estimating the interactions between structures dissimilar in size and shape.\textsuperscript{61} A continuous representation of the binary interaction parameter was used instead:

$$2k_BT\chi_{\text{AA}}(T)_{ij}=g_{ij}+g_{ji}-g_{ij}-g_{ji}=(\zeta_{ij}/\varepsilon_{ij})T \frac{\zeta_{ij}/\varepsilon_{ij}}{\zeta_{ij}/\varepsilon_{ij}}-\left(\zeta_{ij}/\varepsilon_{ij}\right)\frac{\zeta_{ij}/\varepsilon_{ij}}{\zeta_{ij}/\varepsilon_{ij}} \frac{T}{10}$$

For any pair of molecules $A$ (denoted seed molecule) and $B$ (denoted contact molecule), possibly equivalent but with non-symmetric roles, $g_{A+B} = \langle \zeta_{A+B} \varepsilon_{A+B} \rangle_T$ represents the ensemble-averaged potential energy associated with a set of molecules $B$ in contact with a single molecule $A$. $z_{A+B}$ is the coordination number of the arrangement when $A$ is surrounded by molecules $B$ (i.e., number of neighbors $B$), and $\varepsilon_{A+B}$ is the contact energy associated with the contact(s) between a molecule $A$ and a molecule $B$. $\langle \rangle$ and $\langle \rangle_T$ denote averaging on the sampled configurational space and Boltzmann-weighted ensemble averaging, respectively. By introducing the Boltzmann weight, $\exp(-\varepsilon/2k_BT)$, the expectation of the distribution of contact energies $p_{A+B}(\varepsilon)$ becomes

$$\langle \varepsilon_{A+B} \rangle_T = \int_{\varepsilon_{\min}}^{\varepsilon_{\max}} p_{A+B}(\varepsilon) e^{-\varepsilon/2k_BT} d\varepsilon$$

The sampling of $p_{A+B}(\varepsilon)$ was based on a large set of conformers set $A$ and $B$ representative of their condensed state at temperature $T$ and based on all possible contacts of their van der Waals envelopes with elliptical symmetry probability around their centers of mass. The coordination number was determined similarly on a large number of packed configurations, where van der Waals envelopes were put in contact without overlapping. Polymers based on few monomers were idealized by discarding all configurations where head and tail atoms were in contact with any surface. Because the radii of all exposed atoms were chosen equal to their van der Waals radii at their fundamental states (incompressible mixtures), the higher packing related to hydrogen bonding was locally underestimated.

3. Simulation Details

3.1. MC Sampling. The sampling of conformers and pairwise contact energies was performed using the client/server Materials Studio environment version 4.1 (Accelrys, San Diego, CA), its scripting features, and the atom-based COMPASS force field. All van der Waals and Coulombic interactions were calculated without any cutoff since the computational cost of pairwise interactions was not limiting. The commercial implementation of eqs 10 and 11 in Materials Studio via the package Blends was not privileged as the sampling method is mainly optimized for small and spherical solutes. The excluded volume constraint method was instead implemented using an importance sampling optimized for aspherical and bulky solutes. The algorithm to retrieve configurations representative of the condensed state of the mixture $A + B$ is summarized as follows.

1. Two configurations are randomly picked for $A$ (seed molecule) and $B$ (contact molecule).

2. $A$ is oriented along the referential $R_A$ formed by its principal axes; the principal axes of $B$ are also oriented along $R_A$.

3. A random position, $O$, was assigned with a normal radial distribution centered on the center of mass of $A$ and with a standard deviation equal to the gyration radius of $A$.

4. A random direction starting from the center of mass of $A$, $D_A$, is chosen. The angular distribution was chosen to match the differential cross section per unit solid angle of the ellipsoid enclosing $A$. A random direction $D_B$ is similarly chosen from the ellipsoid enclosing $B$.

5. The main axis of $B$ is oriented along $D_B$ while a random rotation around $D_B$ is applied.

6. $B$ is translated from $O$ along $D_B$ until the van der Waals surfaces are in contact (multiple contacts are permitted) while minimizing the overlapping volume. Nonfeasible displacements or configurations are discarded.

7. $\varepsilon_{A+B}$ is calculated.

For conformers presenting an internal cavity and small contact molecules, step 6 was modified to minimize the risk of oversampling of internal cavities. The initial position of $B$ was chosen along the straight line defined by $O$ and $D_A$ either outside $A$ or in $O$. The displacement of $B$ proceeded subsequently by a contraction (displacement toward $O$) or by an expansion (displacement from $O$). The modifications introduced made it possible to sample efficiently the configuration space associated with two conformers of $A$ and $B$ in contact. Usually $10^2$ to $10^3$ samples were sufficient to achieve a convergence for two particular conformers. Because the whole procedure was repeated for a random pairs of $10^2$ to $10^6$ of conformers, $\langle \varepsilon_{A+B} \rangle_T$ was calculated from the distribution of $10^6$ to $10^9$ estimates of contact energies.

To generate conformers realistic at the desired temperature $T$ while minimizing the computational cost of their generation, $10^3$ to $10^4$ conformers were extracted as a first guess from 20 ns MD simulations in a vacuum. The first 2 ns were discarded. For flexible chains much shorter than their persistence lengths, this protocol was in particular assumed to be compatible with the real configuration of tested molecules in a theta solvent. All molecular dynamics were simulated with Discover (Accelrys, USA) and the same COMPASS force field. The temperature $T$ was set to 298 or 313 K in eq 11 and MD simulations to match experimental tests recommended by the EU regulation of plastics materials in contact with food.

3.2. Biases Associated with Configurational Space Sampling. In lattice methods, the coordination number depends only on the lattice and not on the considered seed molecule so that $z_{A+B}$ and $\varepsilon_{A+B}$ remain uncorrelated. In the off-lattice formulation proposed in eq 10, approximating the enthalpic terms $g_{A+B} = \langle \zeta_{A+B} \varepsilon_{A+B} \rangle_T$ by the product of their averages reduced drastically the overall computational method cost but introduced a bias, denoted $\beta_{A+B}$ which is equal to the covariance between $z_{A+B}$ and $\varepsilon_{A+B}$:

$$\beta_{A+B} = \langle \zeta_{A+B} \varepsilon_{A+B} \rangle_T - \langle \zeta_{A+B} \rangle \langle \varepsilon_{A+B} \rangle_T$$

In the presence of biases, the potential energy of $i$ in $P$ might depend on the length of the polymer segment chosen to represent $i$. The optimal length would therefore correspond to a minimal bias. Since it was too expensive to calculate $g_{A+B} = \langle \zeta_{A+B} \varepsilon_{A+B} \rangle_T$ for all possible pairs $A$ and $B$ by MD simulations, biases were estimated from the neighborhood-based local covariance for a set of homologous series of molecules:

$$\beta_{A+B} = 2s_{A+B} \frac{\Delta_{A+B}}{s_{A+B}} \frac{s_{A+B} \Delta_{A+B} + s_{A+B}}{2} \leq s_{A+B} \frac{\Delta_{A+B}}{s_{A+B}}$$
where $s_X^2$ is the variance estimate of $X$ within the considered sample of homologous molecules and $\Delta_{A+B}$ is referred as the derivative of $z_{A+B}$ with $\epsilon_A+\epsilon_B$. $\Delta_{A+B}$ was approximated by the gradient assessed when either $A$ or $B$ is replaced by a similar molecule:

$$\Delta_{A+B} = \frac{\partial z_{A+B}}{\partial \epsilon_{A+B}} = \frac{1}{2} \left( \frac{\partial z_{A+B}}{\partial \epsilon_A \text{homologous } B} + \frac{\partial z_{A+B}}{\partial \epsilon_B \text{homologous } A} \right)$$

Estimating $\beta_{A+B}$ on homologous molecules instead on conformers made it possible to keep similar ensemble averages on a large number of conformers.

According to eqs 10 and 12, the estimation of $\langle z_{i,j} \rangle_{p,L}$ becomes

$$2k_BT \langle z_{i,j} \rangle_{p,L} = \langle z_{i,j} \rangle_{p} \langle \epsilon_{i,j} \rangle_{p} + \langle z_{i,j} \rangle_{x} \langle \epsilon_{i,j} \rangle_{x} - \langle z_{i,j} \rangle_{x} \langle \epsilon_{i,j} \rangle_{p} - \langle z_{i,j} \rangle_{p} \langle \epsilon_{i,j} \rangle_{x}$$

$$+ \langle z_{i,j} \rangle_{p} \langle \epsilon_{i,j} \rangle_{x} - \langle z_{i,j} \rangle_{x} \langle \epsilon_{i,j} \rangle_{p} = \langle z_{i,j} \rangle_{p} \langle \epsilon_{i,j} \rangle_{p}$$

According to eqs 2.4 and 15, the estimate of the partition coefficient between $L$ and $P$ becomes

$$2k_BT \ln K_{i,j,p,L} = 2k_BT r_k^{-1} + \langle z_{i,j} \rangle_{p} \langle \epsilon_{i,j} \rangle_{p} - \langle z_{i,j} \rangle_{x} \langle \epsilon_{i,j} \rangle_{p}$$

$$+ \langle z_{i,j} \rangle_{p} \langle \epsilon_{i,j} \rangle_{x} - \langle z_{i,j} \rangle_{x} \langle \epsilon_{i,j} \rangle_{p}$$

It is worth noting that the final result did not depend on the reference chemical potential $\mu_{i,j}$. Besides, some symmetric biases ($i + P$ and $P + i$, $i + L$ and $L + i$) were expected to balance each other due to a possible exchange of roles between the seed and the contact molecule.

### 3.3. Reference Partition Coefficients

A set of experimental partition coefficients around our targeted temperature (313 K) were collected from the literature. The corresponding values of $K_{i,j,p}$ are gathered in Table 1 along with estimates of $V_{i,j}^d$, $V^m$, melting ranges, and $b_i$. It is worth noting that a significant uncertainty exists on $K_{i,j,p}$ values derived from apparent partitioning $K_{i,j,p}$ (via eq 7), since the data were obtained on different polyethylenes with different formulations and for temperatures ranging between 298 and 313 K. When data on polymer crystallinity and density were missing, likely values were applied.

### 4. Results and Discussion

In eq 2.4, $\chi_{i,j} - \chi_{L,i}$ encompasses both the enthalpic contribution (four pairwise terms) and the residual entropy. For convenience, the usual terminology of enthalpic contribution will be used for $\chi_{i,j} - \chi_{L,i}$. Similarly, $\tau_{n-1}$ will be related to the entropic contribution, whereas it accounted only for the positional entropy. In the perspective to develop a tailored computational method, we detail, first, the important steps in calculation of terms of eq 16 and analyze possible sampling biases according to statistics and mechanical arguments. The results obtained in polyethylene with our methodology are, second, compared with classical results of Flory–Huggins.
theory as a first validation step. As the partial specific volume of a given solute \( i \) in an alcoholic solution \( L \) in relation to its size, shape, and polarity is not known a priori, the choice of a simple approximation will be discussed by comparing predictions of eq 16 with experimental values. Finally, the best approximation in \( L \) will be applied without fitting to predict partitioning between \( P \) and \( L \) for all tested solutes.

4.1. Sampling of Contact Energies. For a given solute \( i \) in \( P \) and in \( L \), the default strategy consisted in applying the same sample of conformers of \( i \) to the calculation of all pairs \( \{i+j\}_{j=P,L} \). Each conformer was subsequently weighted via the Boltzmann factor according to \( \varepsilon_{i+j} \) values (see eq 11).

By contrast, real configurations are weighted by the packing energy of each structure \( \{i+j\}_{j=P,L} \), formally \( z_{i+j} \varepsilon_{i} + z_{j+i} \varepsilon_{j} \).

Since we started as a first guess from a set of conformers \( i \) representative of their configurations in a gas phase rather than in their respective condensed phase, the effects of the contact method together with the set of conformers were particularly investigated.

Figure 2 presents some typical pair configurations with very low interaction energies and representative of the different pairs tested in this work. One molecule is outlined by its trace embedded inside its van der Waals envelope, while the other is plotted as a calotte model with atoms radii matching their van der Waals ones. Both representations led to very similar results, and the determination of contacts was determined for efficiency from the van der Waals radii of atoms. The depicted configurations with minimal potential energy are known to be not necessarily privileged at high temperature and required an appropriate weighting. They illustrate however two significant trends respectively for large molecules and for a mixture of H donor and H acceptor molecules.

For flexible molecules without strong attractive interactions, the configurations of minimal energies correspond to a crossing over of molecules on their concave surfaces (Figure 2a,b). This crossover configuration is very likely for isolated molecules but is not very realistic in packed systems since all curved molecules expose both an intrados and an extrados to the contact with...
their neighbors. For large molecules such as polymers, the configuration depicted in Figure 2b is very similar to the one envisioned for an entanglement. In the case of polyethylene, the persistence length imposes an entanglement every about 30 monomers.\(^6^3\) As a result, the depicted situation has a relative weight about 7/30.

When hydrogen bonding exists, the importance of large molecule configuration is much less important and the proximity between the H donor and the H acceptor is favored irrespective of the length of the molecule (Figure 2c,d). The frequency of such a contact depends however on the size and shape of the molecule and on the strategy used to sample the contacts. Since the configurational space was weighted according to pair interactions instead of packing energies, the default method could not describe a possible alteration of the configuration of the gross side chain due to a long-lived hydrogen bonding (see Figure 2c as an example). The probability to sample a contact with the oxygen atom of decanol (denoted \(A\)) depicted in Figure 2c was in turn proportional to its solid angle when it was observed from the center of mass of the contact molecule (methanol, denoted \(B\)). Since the solid angle was inversely proportional to the square of the distance between \(A\) and \(B\), \(d_{AB}\), or equivalently inversely proportional to the gyration radius of \(A\), the probability to put in contact \(B\) with a terminal atom of a linear molecule \(A\) (here an oxygen), denoted \(p_{A+B}^O\), depended on its length, \(n_A\), and on its considered configuration, according to the following scaling relationship:

\[
\frac{p_{A+B\text{ rigid}}^O}{p_{A+B\text{ random}}^O} = \left(\frac{d_{AB\text{ random}}}{d_{AB\text{ rigid}}}\right)^2 \propto \left(\frac{n_A^{0.5}}{n_A}\right)^2 = \frac{1}{n_A} \tag{17}
\]

where the subscripts “rigid” and “random” indicate a rigid configuration (as in good solvent) and a random configuration close to a Gaussian (as in a theta solvent).

It is emphasized that the bias identified for linear molecules with eq 17 was not associated with accessibility effects but with our isotropic sampling of pair interactions. To minimize this effect for oblong molecules, the initial expansion/contraction method was modified as follows. The space around the molecule was randomly cut into two half-spaces around its center of mass; the center of mass of atoms belonging to one of the half-spaces was chosen as the new center of expansion/contraction. Previous considerations suggested that \(n\)-alcohols represented a particular worst case to validate our whole strategy of sampling of pair contact energies.

The distributions of pair contact energies, \(\varepsilon_{A+B}\), corresponding to pairs similar to those depicted in Figure 2 are plotted in Figure 3. In the absence of significant polar interactions, the raw
sampled distributions had a negative skew, which was characteristic of dispersive van der Waals forces (induced dipole–dipole forces). The associated contact energies were negative since all atoms were separated by a distance equal to or greater than their

Figure 5. Typical packing configurations to sample the number of first neighbors in contact with a seed molecule: (a) BHT + polyethylene segment consisting of seven monomers, denoted PE5; (b) PE5 + PE5; (c) BHT + ethanol; (d) ethanol + ethanol.

Figure 6. Estimation of biases due to the sampling procedure of excess contact energies with direct neighbors in $P$: (a, b) number of neighbors when PE segments idealized by 1 (denoted 1) to 10 monomers (denoted 10) are in contact with n-alcohols (from decanol, denoted C10, to octacosanol, denoted C28) versus the corresponding sampled thermalized pair contact energy; (c) overall bias calculated with eqs 12–14; (d) effects of the number of monomers to idealize PE segments on the estimation of $\chi_{i,P}$. 
van der Waals radii. By contrast, pair molecules with strong dipolar interactions created almost symmetric distributions with a high kurtosis including a sharp maximum and fat tails. The pair potential could be significantly positive due to possible repulsions between close partial charges with the same signs.

It is worth emphasizing that partial overlapping of atoms in the case of significant attractive Coulombic interactions was not allowed by the sampling algorithm.

Weighting previous distributions with the Boltzmann factor (see eq 11) determined the relative probability of each sampled
pair configuration as if the system was in thermodynamic equilibrium with a thermostat at the desired temperature (Figure 3a). Further details on the distinction between so-called global and local equilibrium states can be found. In the absence of dipolar interactions, the thermalization process decreased the average interaction energy by a few dozen percent without changing the overall distribution shape. By contrast, the distributions including dipolar interactions were drastically modified after thermalization. A sharp secondary mode associated with hydrogen bonding was noticed for values less than $-2 \text{ kcal} \cdot \text{mol}^{-1}$. Although the potential associated with a hydrogen bond is about 10 times the potential of London forces, its importance decreases rapidly with the size of linear alcohols. For small molecules, such as methanol, distributions presented several maxima. For nonconvex molecules, there was a risk to oversample internal cavities larger than the contact molecule. The situation is depicted in Figure 4a,b for heptadecanol in contact with methanol. Such configurations promoted the creation of a hydrogen bond. As a result, the expansion method starting from a position close to the center of the cavity yielded probabilities of hydrogen bonding much higher than those observed for shorter molecules such as decanol, which was not realistic. By contrast, the contraction method minimized the contribution of hydrogen bonding in long n-alcohols. The behavior of both methods was systematically studied for all n-alcohols used as seed molecules, A, in contact with small alcohols, B, with increasing size. The criterion to create an internal cavity was defined as

Figure 8. Flory–Huggins parameters versus number of carbon atoms in i: (a) $L$ = methanol; (b) $L$ = ethanol; (c) $P$ = polyethylene segment with optimal length. CAM = camphor; DPM = diphenylmethane; DPO = diphenyl oxide; EUG = eugenol; IAA = isoamyl acetate; INX = Irganox 1076; LIM = l-limonene; LNA = linalyl acetate; MEN = DL-menthol; PEA = phenylethyl alcohol.

Figure 9. Ratio of cavity sizes to insert a solute i in L defined in eq 20: (a) partition between methanol and PE; (b) partition between ethanol and PE. Vertical bars represent the dispersion index of the ratio due to experimental errors in experimental $K_{i,L}$ values.
where $V_A^H$ is the hydrodynamic volume occupied by the molecule $A$; $V_A^{vdW}$ and $V_B^{vdW}$ are the van der Waals volume of molecules $A$ and $B$, respectively. The values of the criterion averaged over more than 50 conformers are plotted for all tested $n$-alcohols and liquids in Figure 4c. The critical size to create a cavity for tested $n$-alcohols appeared between dodecanol and tetradecanol. The corresponding contribution of hydrogen bonding was assessed by calculating the cumulative probability to sample, at temperature $T$, a contact energy less than $-2$ kcal mol$^{-1}$. Beyond the critical size, both contacting methods led to dramatically dissimilar results and only the contraction method gave the expected trend: the contribution of hydrogen bonding should decrease linearly with the number of carbons in the chain.

Additional error due to the number of conformers is illustrated in Table 2 while keeping the same number of MC trials. For the same reasons as discussed previously, it was significant for flexible molecules due to the large extent of the conformational phase space. To minimize this error, up to $10^4$ conformers were sampled to achieve an appropriate convergence.

4.2. Sampling of the Number of First Neighbors. The number of neighbors in contact with a seed molecule was assigned only on the basis of packing considerations as illustrated in Figure 5. The convergence was faster for molecules with similar sizes and shapes. To ensure that the configuration space of each molecule was conveniently sampled, at least $10^4$ packing samples were performed. Infinitely long molecules such as polymers were simulated by preventing the head and tail atoms from being in contact with any other molecule (Figure 5a,b). In the case of associating molecules, multiple hydrogen bonding could be achieved (Figure 5d), but since no thermalization process was applied this configuration was not favored in the sample. Thus, the packing of associating liquids such as methanol or ethanol are expected to differ from their real structures as they could be assessed using molecular dynamics. As we were interested in the average number of neighbors and not in the structure factor of the liquid, this approximation was hence not limiting. Nevertheless, it is worth noting that since packaging and enthalpic considerations were treated separately, it was not possible to account for possible multiple hydrogen bonds involving more than two molecules. Thus, in ethanol at 298 K, about 79% of molecules have two hydrogen bonds versus 15% that have a single hydrogen bond. The approximation introduced by the Flory–Huggins theory is however partly compensated by the highest lifetime (or equivalently the lowest free energy) of simple hydrogen bonds or double hydrogen bonds involving only two molecules.

The effect of the number of sampled conformers on the contact neighbors is presented in Table 2. The effect was much lower than for pair interaction energies.

4.3. First Attempt of the Validation of the Flory–Huggins Approximation. The mean-field theory of Flory–Huggins relies on two important assumptions, which should be verified by our simulation results. (i) The interactions in $P$ do not depend on the length of polymer segments used to represent the polymer. (ii) The pair contact energies and the number of neighbors are not correlated together, and consequently the average of the product can be approximated efficiently by the product of the averages. According to Figure 5a, the first assumption seemed unrealistic in our simulations when the size of polymer segments was much smaller than the solute size. Indeed, the noncontact condition of head and tail atoms in $P$ tended to underestimate the real number of contacts and neighbors. The violation of the second assumption was more obvious as all calculations demonstrated that $z_{A+B}$ was always highly correlated to $z_A$ and $z_B$ (i.e., their variations are opposite). This condition was almost sufficient to ensure a minimum covariance between the random variable $z_{A+B}$ and the random variable $z_A + z_B$. It is underline that the conventional Flory–Huggins description on a lattice is by construction symmetric (Figure 1), because the size of the blob increases with the volume of $i$ while the number of neighbors remains imposed by the lattice coordination number.
Table 4. Calculated Flory–Huggins Interaction Parameters

<table>
<thead>
<tr>
<th>Solute</th>
<th>$X_{c(n)}$-methylene</th>
<th>$X_{c(n)}$-ethanol</th>
<th>$X_{c(n)}$-isopropanol</th>
<th>$X_P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>decane</td>
<td>5.89 ± 0.01</td>
<td>4.40 ± 0.01</td>
<td>3.28 ± 0.003</td>
<td>0.3 ± 0.05</td>
</tr>
<tr>
<td>undecane</td>
<td>6.17 ± 0.0005</td>
<td>4.65 ± 0.003</td>
<td>3.50 ± 0.004</td>
<td>0.30 ± 0.001</td>
</tr>
<tr>
<td>dodecane</td>
<td>6.62 ± 0.001</td>
<td>5.06 ± 0.003</td>
<td>3.87 ± 0.001</td>
<td>0.31 ± 0.002</td>
</tr>
<tr>
<td>tridecane</td>
<td>6.90 ± 0.005</td>
<td>5.31 ± 0.003</td>
<td>4.10 ± 0.01</td>
<td>0.30 ± 0.01</td>
</tr>
<tr>
<td>tetradecane</td>
<td>7.26 ± 0.01</td>
<td>5.64 ± 0.01</td>
<td>4.38 ± 0.02</td>
<td>0.23 ± 0.003</td>
</tr>
<tr>
<td>pentadecane</td>
<td>7.84 ± 0.02</td>
<td>6.16 ± 0.01</td>
<td>4.85 ± 0.01</td>
<td>0.24 ± 0.01</td>
</tr>
<tr>
<td>hexadecane</td>
<td>8.03 ± 0.05</td>
<td>6.35 ± 0.05</td>
<td>5.03 ± 0.04</td>
<td>0.37 ± 0.04</td>
</tr>
<tr>
<td>heptadecane</td>
<td>8.44 ± 0.05</td>
<td>6.72 ± 0.05</td>
<td>5.36 ± 0.04</td>
<td>0.28 ± 0.04</td>
</tr>
<tr>
<td>octadeacne</td>
<td>8.76 ± 0.05</td>
<td>7.01 ± 0.05</td>
<td>5.62 ± 0.04</td>
<td>0.44 ± 0.04</td>
</tr>
<tr>
<td>nonadecane</td>
<td>8.82 ± 0.05</td>
<td>7.09 ± 0.05</td>
<td>5.66 ± 0.04</td>
<td>0.55 ± 0.04</td>
</tr>
<tr>
<td>eicosane</td>
<td>9.15 ± 0.02</td>
<td>7.38 ± 0.02</td>
<td>5.93 ± 0.01</td>
<td>0.68 ± 0.01</td>
</tr>
<tr>
<td>docosane</td>
<td>9.71 ± 0.02</td>
<td>7.91 ± 0.01</td>
<td>6.40 ± 0.01</td>
<td>0.11 ± 0.01</td>
</tr>
<tr>
<td>tetracosane</td>
<td>9.96 ± 0.06</td>
<td>8.21 ± 0.06</td>
<td>6.68 ± 0.05</td>
<td>0.39 ± 0.05</td>
</tr>
<tr>
<td>octacosane</td>
<td>10.87 ± 0.7</td>
<td>9.08 ± 0.6</td>
<td>7.48 ± 0.5</td>
<td>0.35 ± 0.05</td>
</tr>
<tr>
<td>decanol</td>
<td>3.77 ± 0.05</td>
<td>2.49 ± 0.04</td>
<td>1.86 ± 0.04</td>
<td>0.79 ± 0.03</td>
</tr>
<tr>
<td>undecanol</td>
<td>4.43 ± 0.05</td>
<td>3.03 ± 0.04</td>
<td>2.09 ± 0.04</td>
<td>0.77 ± 0.04</td>
</tr>
<tr>
<td>dodecanol</td>
<td>5.29 ± 0.01</td>
<td>3.71 ± 0.03</td>
<td>2.47 ± 0.01</td>
<td>0.69 ± 0.002</td>
</tr>
<tr>
<td>tridecanol</td>
<td>5.19 ± 0.04</td>
<td>3.77 ± 0.03</td>
<td>2.27 ± 0.03</td>
<td>0.68 ± 0.01</td>
</tr>
<tr>
<td>tetradecanol</td>
<td>5.81 ± 0.03</td>
<td>4.35 ± 0.03</td>
<td>2.49 ± 0.03</td>
<td>0.40 ± 0.003</td>
</tr>
<tr>
<td>pentadecanol</td>
<td>6.21 ± 0.02</td>
<td>4.80 ± 0.02</td>
<td>2.52 ± 0.01</td>
<td>0.46 ± 0.01</td>
</tr>
<tr>
<td>hexadecanol</td>
<td>6.18 ± 0.06</td>
<td>4.92 ± 0.06</td>
<td>2.37 ± 0.05</td>
<td>0.54 ± 0.05</td>
</tr>
<tr>
<td>heptadecanol</td>
<td>6.89 ± 0.06</td>
<td>5.62 ± 0.06</td>
<td>2.70 ± 0.05</td>
<td>0.47 ± 0.04</td>
</tr>
<tr>
<td>octadecanol</td>
<td>6.76 ± 0.03</td>
<td>5.71 ± 0.03</td>
<td>2.34 ± 0.02</td>
<td>0.70 ± 0.02</td>
</tr>
<tr>
<td>nonadecanol</td>
<td>7.22 ± 0.06</td>
<td>6.04 ± 0.06</td>
<td>2.49 ± 0.05</td>
<td>0.72 ± 0.05</td>
</tr>
<tr>
<td>eicosanol</td>
<td>7.95 ± 0.02</td>
<td>6.72 ± 0.02</td>
<td>2.94 ± 0.01</td>
<td>0.76 ± 0.002</td>
</tr>
<tr>
<td>camphor</td>
<td>-0.54 ± 0.005</td>
<td>-0.81 ± 0.005</td>
<td>-1.02 ± 0.005</td>
<td>0.67 ± 0.002</td>
</tr>
<tr>
<td>diphenyl oxide</td>
<td>-1.05 ± 0.005</td>
<td>-0.78 ± 0.005</td>
<td>-0.55 ± 0.005</td>
<td>0.42 ± 0.002</td>
</tr>
<tr>
<td>diphenylmethane</td>
<td>3.37 ± 0.02</td>
<td>2.73 ± 0.01</td>
<td>2.12 ± 0.01</td>
<td>-0.05 ± 0.01</td>
</tr>
<tr>
<td>n-limonene</td>
<td>5.12 ± 0.03</td>
<td>3.80 ± 0.01</td>
<td>2.73 ± 0.01</td>
<td>0.04 ± 0.01</td>
</tr>
<tr>
<td>DL-menthol</td>
<td>2.02 ± 0.03</td>
<td>1.55 ± 0.03</td>
<td>1.09 ± 0.03</td>
<td>0.56 ± 0.03</td>
</tr>
<tr>
<td>eugenol</td>
<td>3.07 ± 0.02</td>
<td>2.59 ± 0.01</td>
<td>2.02 ± 0.04</td>
<td>0.43 ± 0.03</td>
</tr>
<tr>
<td>isoamyl acetate</td>
<td>1.36 ± 0.01</td>
<td>0.71 ± 0.01</td>
<td>0.21 ± 0.01</td>
<td>0.49 ± 0.01</td>
</tr>
<tr>
<td>linalyl acetate</td>
<td>2.40 ± 0.03</td>
<td>1.77 ± 0.02</td>
<td>1.19 ± 0.02</td>
<td>0.20 ± 0.02</td>
</tr>
<tr>
<td>phenylethyl alcohol</td>
<td>1.62 ± 0.02</td>
<td>1.30 ± 0.01</td>
<td>0.93 ± 0.01</td>
<td>0.62 ± 0.01</td>
</tr>
<tr>
<td>BHT</td>
<td>3.85 ± 0.003</td>
<td>3.12 ± 0.61</td>
<td>2.30 ± 0.18</td>
<td>0.51 ± 0.06</td>
</tr>
<tr>
<td>Irganon 1076</td>
<td>8.65 ± 0.13</td>
<td>7.14 ± 0.48</td>
<td>6.03 ± 0.18</td>
<td>0.15 ± 0.16</td>
</tr>
</tbody>
</table>

Values at 298 K. b Values at 313 K.

$m$ was increasing. The best suitable number of monomers, $m$, was chosen to minimize the criterion $C_{P_i}^{(m)}$.

$$C_{P_i}^{(m)} = 1 - \frac{\partial g_{P_i}^{(m)}}{\partial g_{P_i}^{(m)}} - \frac{\partial g_{P_i}^{(m)}}{\partial g_{P_i}^{(m)}}$$

Because the configuration sampling of large polymer systems was computationally more expensive, the optimal $m$ value was expected to occur for an intermediate value of $m$ as the best trade-off between convergence costs and sampling biases. Figure 6 plots the correlations between $z_{m+P}$, $Z_{P+i}$, and $e_{P+i}$, the associated cumulated biases, $\beta_{P+i}$, $\beta_{P+i}$, and $\beta_{P+i}$, as defined in eq 16, and their consequences on the estimate of $\chi_{IP}$ when $m$ was increasing. The antisymmetry between the role of the polymer used as seed molecule or as contact molecule was best illustrated in Figure 6a, when it was in contact with aliphatic molecules, which resembled it. When $P$ was the contact molecule (Figure 6a), the number of $P$ segments increased around the solute $i$ while the contact energy was increasing (i.e., the number of contacts per molecule of $P$ decreased). Linear solutes behaved symmetrically when they were placed around a single segment of $P$ (Figure 6b). Corresponding biases are listed in Table 3 for all tested molecules. Despite significant biases introduced by the separation of the calculation of $z_{A+B}$ and $Z_{P+i}$, and $e_{A+B}$, a minimum global bias was finally achieved by considering the sum $z_{A+B} + z_{P+i}$, and $e_{A+B}$, in particular minimal when the roles of seed and contact molecule were
exactly interchangeable, that is, when their sizes were similar (Figure 6c). When the solute did not resemble the polymer, the bias was higher. The consequence was that $\chi_{ij}^{(p)}$ calculated from eq 15 admitted a minimum value close to the value that was responsible for the minimal overall bias (Figure 6d). For too small $m$ values, head and tail effects dominated leading to overestimations of $\chi_{ij}$. For large $m$ values, end effects vanished and the dependence to $m$ was lower, as expected in the Flory–Huggins theory. The accuracy of results was more sensitive to the quality of sampling: all biases increased with $m$ and in particular the contribution of $\beta_{p-p}$. It is worth noting that the asymptotic bias estimated in our simulations (between $0.14k_B T$ and $0.31k_B T$) was close to the theoretical correction of $0.35k_B T$ conventionally applied to fit the Flory–Huggins theory to the solubility of linear $n$-alkanes in polyethylene.$^{33}$ The minimum value of $\chi_{ij}^{(p)}$ over a likely range of $m$ values coincided with the minimum of the criterion and was chosen to estimate $K_{ij,LP}$ in eq 2.4. For an arbitrary solute, the examined range of $m$ values was chosen to mimic a polymer segment with a hydrodynamic volume ranging between half and twice the hydrodynamic volume of $i$.

Previous biases associated with our off-lattice approach of polymer–solute interactions were observed also in $L$ (Figure 7a,b). By contrast with $P$, there was however no degree of freedom (i.e., $m$) available to control the overall bias. As $L$ molecules were much smaller than solute ones, the cumulated biases appeared mainly negative. It is nevertheless underlined that it was almost constant for linear solutes (Figure 7c,d).

$\chi_{ij}$, and $\chi_{ij}$ values calculated with our off-lattice Flory–Huggins approach are gathered in Figure 8 versus the number of solute carbons and are compared along with conventional values derived from van Krevelen solubility coefficients.$^{70}$ The interesting feature was that both $\chi_{ij}$ curves were parallel but with values $9–11$ times lower when they were calculated with the our off-lattice approach. This result confirmed indirectly the bad solubility coefficient weighting for the prediction of activity coefficients in interacting liquids.$^{32}$ Since almost all tested molecules appeared to be highly compatible with the polymer, with $\chi_{ij}$ values close to or less than 1, errors and biases on $\chi_{ij}$, impacted a few on $\chi_{ij} - \chi_{ij}$ differences (see eqs 15 and 16) and only for molecules with a good or a moderate affinity for $L$, that is, mainly for volatiles and phenolic compounds.

4.4. Combinatorial Entropic Contribution. The volume of the cavity required to insert a molecule $i$ in the medium $L$, also called the partial specific volume, affected significantly the prediction of partitioning (see eq 8). In particular, increasing the volume of the cavity in $L$ increased the affinity for $L$ due to a higher entropic contribution. In polar media, the introduction of a hydrophobic substance $i$ is expected to modify the network of hydrogen bonding so that the orientation of $L$ molecules may not be random around $i$ and may increase the size of the effective cavity. On the contrary, the possibility of hydrogen bonding between $i$ and $L$ may lead to a contraction of the cavity. Because our objective was to keep within the simplified framework of the Flory–Huggins theory, we argued that the expected size of the cavity was comprised between the van der Waals volume of $i$ (hard core volume) and its molar volume. This assumption has been studied in a companion work to be published by isobaric molecular dynamics for different fractions in solutes. A dimensionless size of the cavity, $r_{L}^{-1}$ (as used in eq 2.4), is obtained by dividing the previous volume with the molar volume of pure $L$ at the considered temperature. The choice of the molar volume to normalize the typical blob size was consistent with the assumption that no modification of bulk properties occurs at infinite solute dilution. Conversely, the choice of the molar volume for pure solutes $i$ that are solid at the considered temperature (see Table 1) might be questionable since they did not correspond to a disordered state as $i$ was dispersed among $L$ molecules. To check which dimensionless cavity size matched best the experimental partition coefficients between $L$ and the amorphous phase of the polymer, the following ratio, $R$, was calculated at $T = 298$ K and $T = 313$ K:

$$ R = \frac{r_{L}^{-1}(T)}{(\ln K_{ij,L,p}^{(0)} - \chi_{ij}^{(p)} - \chi_{ij}^{(d)})} \quad (20) $$

$R$ values for both assumptions are compared for all tested solutes and liquids in Figure 9. A ratio less than 1 indicated an underestimated cavity volume. In agreement with conventional approaches based on cohesive energy densities, solute molar volumes tended to lead more accurate results and in particular for all linear molecules. The significant discrepancy among experimental results, however, made it not possible to confirm the assumption for all tested molecules. As a result, both definitions were used to estimate $K_{ij,LP}$ with eq 2.4 and a likely value was proposed as the geometric mean of possible bonds.

Besides, calculated $R$ values confirmed unambiguously the predominant role of the positional entropic contribution in $L$ for the prediction of the partitioning. This effect was neglected in the pioneer work of Baner and Piringer.$^{28}$

4.5. Predictions of Partitioning. The comparisons between predicted and experimental values for all tested molecules are presented in Figure 10. Experimental values corresponded to those listed in Table 1. Predicted values were calculated with eq 16 using $e_{L-p} + z_{L-p}$ values obtained from our MC sampling. The corresponding $\{\chi_{ij,ij},\} = p-L$ and $\beta_{p-p}$ values are gathered in Tables 4 and 3, respectively. The vertical error bars are related to the choice of $V_{p}^{MOL}$ or $V_{M}$ to estimate $r_{L}^{-1}$, whereas the horizontal bars include the uncertainty on polymer crystallinity and on solute dispersion in the matrix, as described in eq 7. The predictions were in very satisfactory agreement with measurements for a wide range of molecules highly different in size, shape, stiffness, and polarity. The higher discrepancy observed for linear aliphatic chains in contact with small polar liquids (methanol) was related to the deviation of their conformations in a poor solvent from the one sampled in a gas phase.

Since our description of partitioning was consistent with previously published values, we tried to modify the solubility coefficient based approach of Baner and Piringer$^{28}$ to match our findings and in particular to fit the values of $\chi_{ij,ij}$ and $\chi_{ij}$ estimated by our Monte Carlo approach. The first modification consisted in introducing a weighting sum of the solubility coefficients as described by Lindew et al.$^{32}$

$$ \chi_{i,j} = p-L = \frac{V_{p}^{MOL}}{RT} \left[ \frac{1}{\alpha} \left( \delta_{i}^{2} - \delta_{i}^{2} \right)^{2} + 0.25(\delta_{i}^{2} - \delta_{i}^{2})^{2} + 0.25(\delta_{i}^{2} - \delta_{i}^{2})^{2} \right] $$

$$ \quad (21) $$

where $\{\delta_{i}^{2}\} = p-L$, $\{\delta_{i}^{2}\} = p-L$, and $\{\delta_{i}^{2}\} = p-L$ are respectively the dispersive, the polar, and the hydrogen contributions to the solubility coefficients. The main advantage of such a model is that it uses coefficients that have been extensively tabulated by group contribution for a wide range of substances.$^{31,33}$ As suggested, we used the solubility coefficients of van Krevelen$^{70}$ compared with results for low density polyethylene.$^{28}$ The scaling coefficient $\alpha$ was set to 0.35 to match our $\chi_{ij}$ values for $n$-alkanes. This choice was consistent with the suggested values between 0.25 and 0.3 assessed for mixtures including polymer and hydrogen bonding solvents.$^{32}$ The second modi-
the general trends. The same scaling relationship was also
optimized on quantum mechanics results. Leaving the mean-
are calculated at the atomistic scale using a generic force field
representation of the whole polymer. The molecular interactions
neighbors and consequently does not require an explicit
approach, it takes into account interactions only with nearest
proposed to predict the partitioning of solutes between polymer
financial support.

Acknowledgment

theta solvents.

conditions where diffusion in
possibly in water should be revised, in particular under
related to the significant contribution of the positional entropy
molecules. The high affinity of bulky solutes for liquids was
significant chemical affinity for liquids consisting of small
additives have a good solubility in polymers, they have also a
previously mentioned in Figure 8a,b, the overestimation of the
interactions. The overall correction 0.35
nonweighted sum of solubility coefficients would be mainly
required for \( \chi_{i,P} \) since the nonweighted sum overestimated the
values obtained by our Monte Carlo method.

5. Conclusions

An off-lattice generalized Flory—Huggins approach has been
proposed to predict the partitioning of solutes between polymer
and liquid amorphous phase. As in the original Flory—Huggins
approach, it takes into account interactions only with nearest
neighbors and consequently does not require an explicit
representation of the whole polymer. The molecular interactions
are calculated at the atomic scale using a generic force field
optimized on quantum mechanics results. Leaving the mean-
field approximation of polymer interactions used in conventional
Flory—Huggins approach made it possible to sample directly
the entropic contribution due to conformers and reorientations.
However, the separation of the calculations corresponding to
the number of neighbors and the pair interaction energies
introduced systematic biases in the sampling, which has also
been shown in the original version. They could be minimized
in the polymer by choosing a segment length that matched the
hydrodynamic solute volume. Additionally, it was demonstrated
that isotropic sampling of pair interactions might induce
additional biases for nonspherical molecules: either linear or
including a cavity. By both performing a cautious sampling and
taking into account the positional entropy in the polymer and
in the liquid, it has been demonstrated that partition coefficients
could be predicted without any fitting. These results were used
to adapt a predictive model of partition coefficients based on
solubility coefficients and on the weighted sum.32

Additionally, the simulations demonstrated that even if plastic
additives have a good solubility in polymers, they have also a
significant chemical affinity for liquids consisting of small
molecules. The high affinity of bulky solutes for liquids was
related to the significant contribution of the positional entropy
in L. This effect was either neglected or underestimated in
previous studies. As a result, the migration assessment of large
additives such as hindered phenolic antioxidants in alcohols and
possibly in water should be revised, in particular under
conditions where diffusion in P is not limiting (e.g., at high
temperature). A companion work is dedicated to generalizing
the proposed approach of partitioning to water—ethanol mix-
tures, while providing a reliable estimate of \( r_F^{-1} \) in poor and
theta solvents.

Acknowledgment

The authors would like to thank the Association de Coordina-
tion Technique pour l’Industrie Agroalimentaire and the As-
sociation Nationale pour la Recherche Technique (ANRT) for
financial support.

List of Symbols

\( A \) = any molecule with a seed role
\( B \) = any molecule with a contact role
\( \{C_{i_{\text{eq}}}\}_{i=L,P} \) = macroscopic concentration at equilibrium of solute
\( i \) in \( j = L, P \) as experimentally assessed (kg m\(^{-3}\))
\( C_{i_{\text{eq}}} \) = concentration of solute \( i \) in liquid \( L \) at equilibrium (kg m\(^{-3}\))
\( C_{i_{\text{eq}}} \) = concentration of solute \( i \) in the amorphous phase of polymer
\( P \) at equilibrium (kg m\(^{-3}\))
\( C_{i_{\text{eq}}}^{\text{init}} \) = initial concentration of solute \( i \) in polymer \( P \) (i.e., before
contact with \( L \)) (kg m\(^{-3}\))
\( C_{i_{\text{eq}}}^{\text{ct}} \) = criterion defined in eq 19
\( G_i \) = free energy to cross the interface (J)
\( G_{i,k} \) = free energy between \( j = L, P \), \( i \), \( k = L, P \), \( i, j \neq k \) (J)
\( H_{i,j} \) = enthalpy between solute \( i \) and \( j = L, P \) (J)
\( k_{i,L,P} \) = partition coefficient of solute \( i \) between \( L \) and the amorphous
region of \( P \)
\( \hat{k}_{i,L,P} \) = apparent partition coefficient of solute \( i \) between \( L \) and a
semicrystalline polymer \( P \)
\( L \) = liquid
\( N_i \) = number of molecules \( j = L, P \)
\( N_{i,j} \) = number of solute \( i \) in \( j = L, P \)
\( P \) = polymer
\( R \) = cavity size criterion defined in eq 20
\( T \) = absolute temperature (K)
\( V_i \) = hydrodynamic volume occupied by the molecule \( A \) (m\(^3\))
\( V_i \) = partial specific volume solute \( i \) (m\(^3\))
\( V_i^{\text{aw}} \) = van der Waals volume of molecule \( A = L, i \) (m\(^3\))
\( V_i^{\text{mol}} \) = molar volume of the solute \( i \) (m\(^3\))
\( a_i \) = extraction yield of solute \( i \) from polymer \( P \)
\( b_i \) = amount of solute \( i \), which is not “well-dispersed” in polymer \( P \)
\( c \) = polymer crystallinity (volume fraction of crystalline phase)
\( g_{a+b} \) = ensemble-averaged potential energy associated with a set
of molecules \( B \) in contact with a single molecule \( A \)
\( i \) = solute index
\( h \) = Planck’s constant \((6.63 \times 10^{-34} \text{ J s}^{-1})\)
\( k_B \) = Boltzmann’s constant \((1.38 \times 10^{-23} \text{ J K}^{-1})\)
\( k_{i,j,k} \) = frequency of \( i \) crossing the interface from \( j \) \( P \), \( k \) = \( L \), \( P \), \( L \) with \( j \neq k \) (s\(^{-1}\))
\( l \) = dilution factor, or equivalently the ratio between the volume
of polymer \( P \) and the volume of liquid \( L \)
\( m \) = number of monomers
\( n_i \) = length of molecule \( A = L, P \)
\( r_{a+b}^{\text{dist}} \) = distribution of contact energies between \( A \) and \( B \)
\( \hat{p}_{i,j} \) = probability to put in contact a molecule \( B \) with a terminal
atom of a linear molecule \( A \)
\( \hat{p}_{i_{\text{eq}}} \) = presence probability at equilibrium of solute \( i \) from side of
\( j = L, P \) at the interface
\( \hat{r}_j \) = number of blobs of volume \( V_i \) for molecule \( j = L, P \)
\( \xi_{a+b} \) = coordination number of the arrangement when \( A \) is
surrounded by molecules \( B \)

Greek Symbols

\( \beta_{A-B} \) = covariance between \( \xi_{a+b} \) and \( \varepsilon_{a+b} \)
\( \gamma_i \) = activity coefficient of solute \( i \) in \( j = L, P \)
\( \delta_i^d \) = dispersive contribution to the solubility coefficients for \( j \)
\( L \), \( P \), \( i \) (J mol\(^{-1}\))
\( \delta_i^p \) = polar contribution to the solubility coefficients for \( j \)
\( L \), \( P \), \( i \) (J mol\(^{-1}\))
\( \delta_i^h \) = hydrogen contribution to the solubility coefficients for \( j \)
\( L \), \( P \), \( i \) (J mol\(^{-1}\))
\( \Delta_{A-B} \) = derivative of \( \xi_{a+b} \) with \( \varepsilon_{a+b} \) (see eq 14)
\( \varepsilon_{i,k} \) = contact energy associated with the contact(s) between a
molecule \( j = A, B, L, P \), \( i \) and a molecule \( k = A, B, L, P \), \( i \neq k \)

List of Symbols
\( \mu_j = \text{chemical potential of } j = P, L \)
\( \mu_{ij} = \text{chemical potential of molecules } i \text{ in } j = L, P \)
\( \phi_j = \text{volume fraction of } j = L, P \)
\( \phi_{ij} = \text{volume fraction of solute } i \text{ in } j = L, P \)
\( \{x_{ij}\}_{j,f} = \text{Fory-Huggins parameter} \)

**Subscripts/Superscripts**

\( \text{eq} = \text{value at equilibrium} \)
\( \text{excess} = \text{excess part} \)
\( \text{id} = \text{ideal part} \)

**Mathematical Operators**

\( \text{pr}(X;x) = \text{probability that variable } X \text{ was less than or equal to} \)
\( \text{value } x \)
\( \langle \rangle_X = \text{variance estimate of } X \)
\( \{ \rangle_X = \text{averaging on the sampled configurational space} \)
\( \langle \rangle_T = \text{Boltzmann-weighted ensemble averaging on the sampled configurational space} \)

**Literature Cited**


Received for review July 24, 2008
Revised manuscript received February 15, 2009
Accepted March 20, 2009

IE801141H