

Enhancing Solubility Predictions of Plastic Additives in Depolymerization Solvents Using Quantum Chemical Conformer Analysis and COSMO-RS

Amin Shariatmadar Tehrani, Ruben Goeminne, Pieter Cnudde, Veronique Van Speybroeck, Kathy Elst, Karolien Vanbroekhoven, Elias Feghali, Kevin M. Van Geem, and Steven De Meester*



Cite This: *ACS Omega* 2026, 11, 59039–59053



Read Online

ACCESS |



Metrics & More

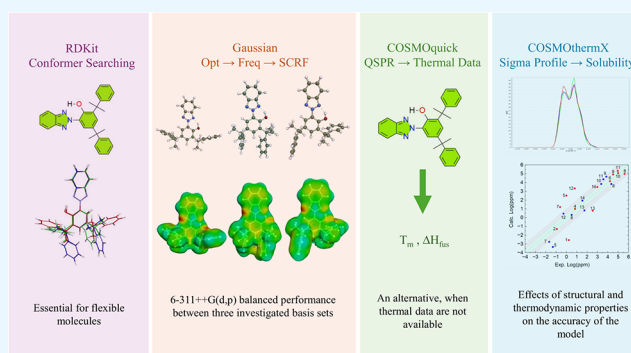


Article Recommendations



Supporting Information

ABSTRACT: The success of recycling plastics by means of depolymerization requires the selective purification of regenerated monomers from complex mixtures. This can be assisted by accurate prediction of the phase behavior of large, flexible, and multifunctional additive molecules, which remains a challenge for current thermodynamic models. In this study, the solubility of 11 representative additives and three PET monomers was measured in two common depolymerization solvents, water and ethylene glycol, covering a wide range from practically insoluble (<1 ppm) to 41.6 wt %. To benchmark predictive capabilities, a quantum-chemistry workflow was investigated, integrating RDKit conformer generation, Gaussian DFT optimizations, and COSMO-RS thermodynamic modeling. Systematic conformer sampling proved critical for bulky, high-molecular-weight additives such as antioxidants. A comparison of three Pople-type basis sets revealed a trade-off between computational cost and accuracy. The 6–311++G(d,p) basis set provided the best balance (RMSE = 1.31 log units), while 3–21+G(d) reduced computational cost at lower accuracy, and 6–311++G(3df,3pd) yielded only minor improvements for select cases at significantly higher cost. Prediction accuracy was highest for more soluble compounds but decreased with increasing molecular weight, conformational flexibility, and Gibbs free energy of fusion. To improve solid–liquid equilibrium predictions, melting points and fusion enthalpies were measured using DSC and compared with QSPR estimates. This combined computational-experimental approach supports the development of energy-efficient purification strategies in chemical recycling of condensation polymers such as PET.



1. INTRODUCTION

Plastic pollution has emerged as a critical global challenge, making innovations in this field more urgent than ever to reduce its environmental impact. Next to reduction, reusing, and eco-design, plastic recycling, encompassing both mechanical and chemical methods, serves as the cornerstone of circular and more sustainable plastics. For many condensation polymers, such as PET, depolymerization is a promising approach to achieve closed-loop recycling by converting the polymer into its building blocks through processes like glycolysis, hydrolysis, or methanolysis.¹ Scientific literature has focused a lot on the depolymerization reaction itself.² Yet, when developing an industrial process, one of the most energy-intensive stages in depolymerization, the purification of its reaction products, has been understudied. This challenge arises from the inherent complexity of the initial waste feedstock, which contains not only polymers but also various (Non) Intentionally Added Substances ((N)IAS) such as additives,³ catalysts from the polymerization process, plastic degradation products, and contaminants introduced during consumer use.⁴

These components may exist as molecules, long-chain oligomers, or ionic species, and they can be organic or mineral, which complicates the downstream processing of depolymerization reactions.^{5,6}

The selection and design of purification processes typically depend on properties such as solubility, vapor pressure, and partition coefficients that can be estimated based on thermodynamic parameters, such as activity coefficients.^{7,8} Such property estimation methods are known or have been studied for petrochemical molecules for decades.⁹ Yet, for many (N)IAS present in plastic depolymerization media, such as additives, these thermodynamic properties have not been

Received: April 16, 2026

Revised: August 2, 2026

Accepted: August 4, 2026

Published: September 21, 2026



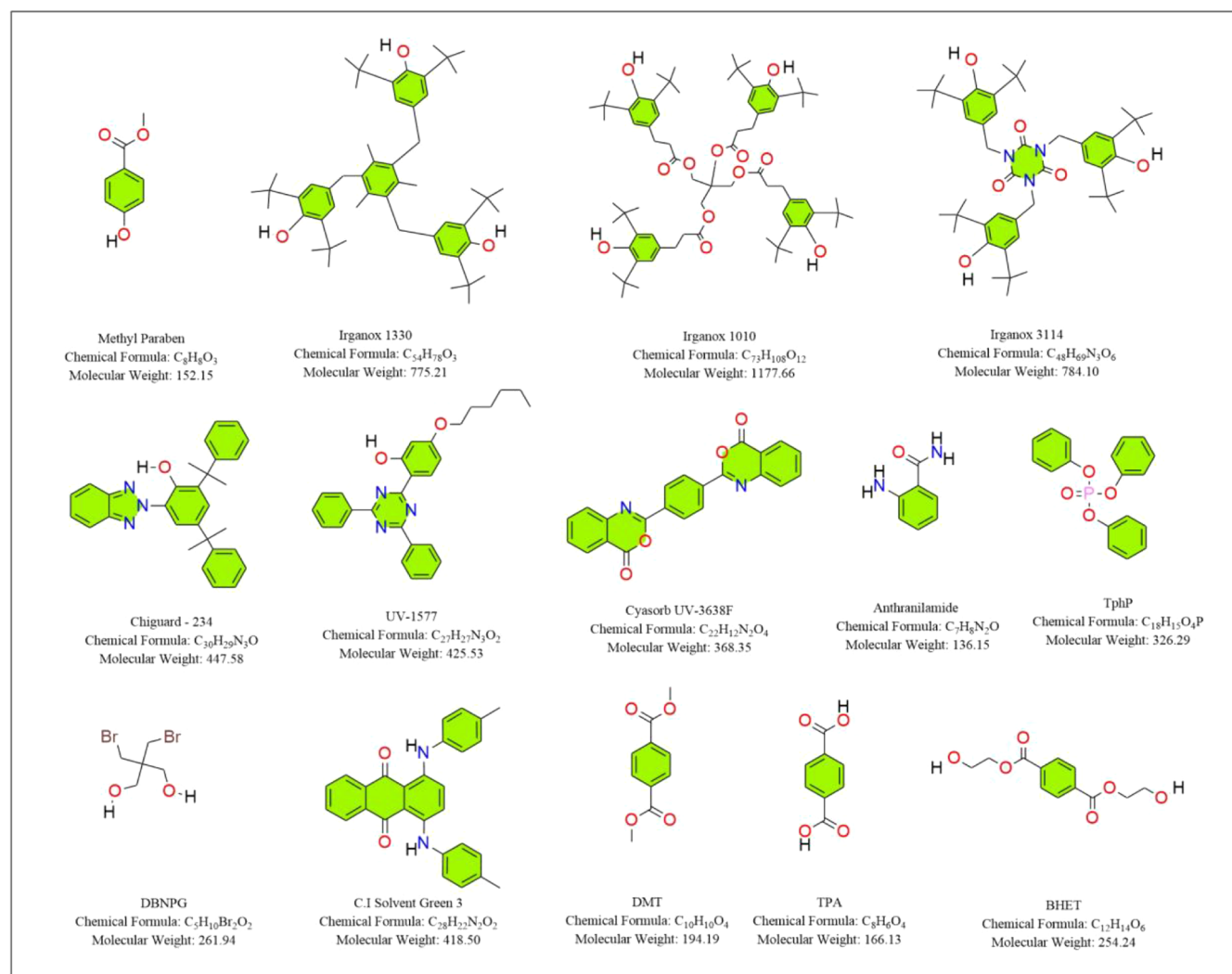


Figure 1. 2D structures of the studied components.

described.^{10–12} Traditional thermodynamic models, such as Excess Gibbs Energy models (e.g., NRTL, UNIQUAC), rely heavily on experimental equilibrium data to regress binary interaction parameters for all the component pairs in the system, which can be very expensive, time-consuming, or even impossible.¹³ In this regard, the COSMO-RS (Conductor-like Screening Model for Real Solvents) thermodynamic model has gained a lot of attention as it allows a deeper understanding of a solution's nonideality and intermolecular forces.^{14,15} This model has become particularly popular because it can require minimal to no experimental input, relying primarily on the results of Quantum Chemistry (QC) calculations, which provide information about the molecular structure, volume, and area, as well as the sigma profile. Such a sigma profile represents the charge density distribution on the molecular surface, serving as the basic input, along with molecular volume and area, for predicting thermodynamic properties and ultimately phase equilibria.^{16,17} Such calculations require a stepwise approach, starting with conformer searching and geometry optimization, and progressing to calculating thermodynamic properties and solvation. This process, especially for relatively large and complex molecules such as additives, involves numerous methodological choices.

In this context, this study aims to evaluate a computational workflow for predicting the solubility of 11 additives and three PET-related monomers in water and ethylene glycol, two typical solvents used in hydrolysis and glycolysis depolymerization, respectively. Furthermore, the solubility measurements reported in Section 4.2 serve as the validation reference for the workflow throughout this study. The selected molecules, representative of additive categories commonly expected in depolymerization processes, are depicted in Figure 1. Knowing their solubility is important as residual additives need to be separated to achieve a pure monomer, allowing repolymerization to high-quality PET. The compounds were selected for their structural diversity and relevance to PET, a key polymer in the circular economy of plastics. By comparing three different basis sets, the trade-offs between computational cost and accuracy are assessed to identify an optimal approach and its limitations for use in depolymerization system modeling.

2. THEORETICAL BASIS AND COMPUTATIONAL FRAMEWORK

In this paper, the solubility of selected monomers and additives is estimated following a workflow shown in Figure 2. The process begins with an automated search of molecular conformations, utilizing RDKit (V.2023.09.3) to generate a

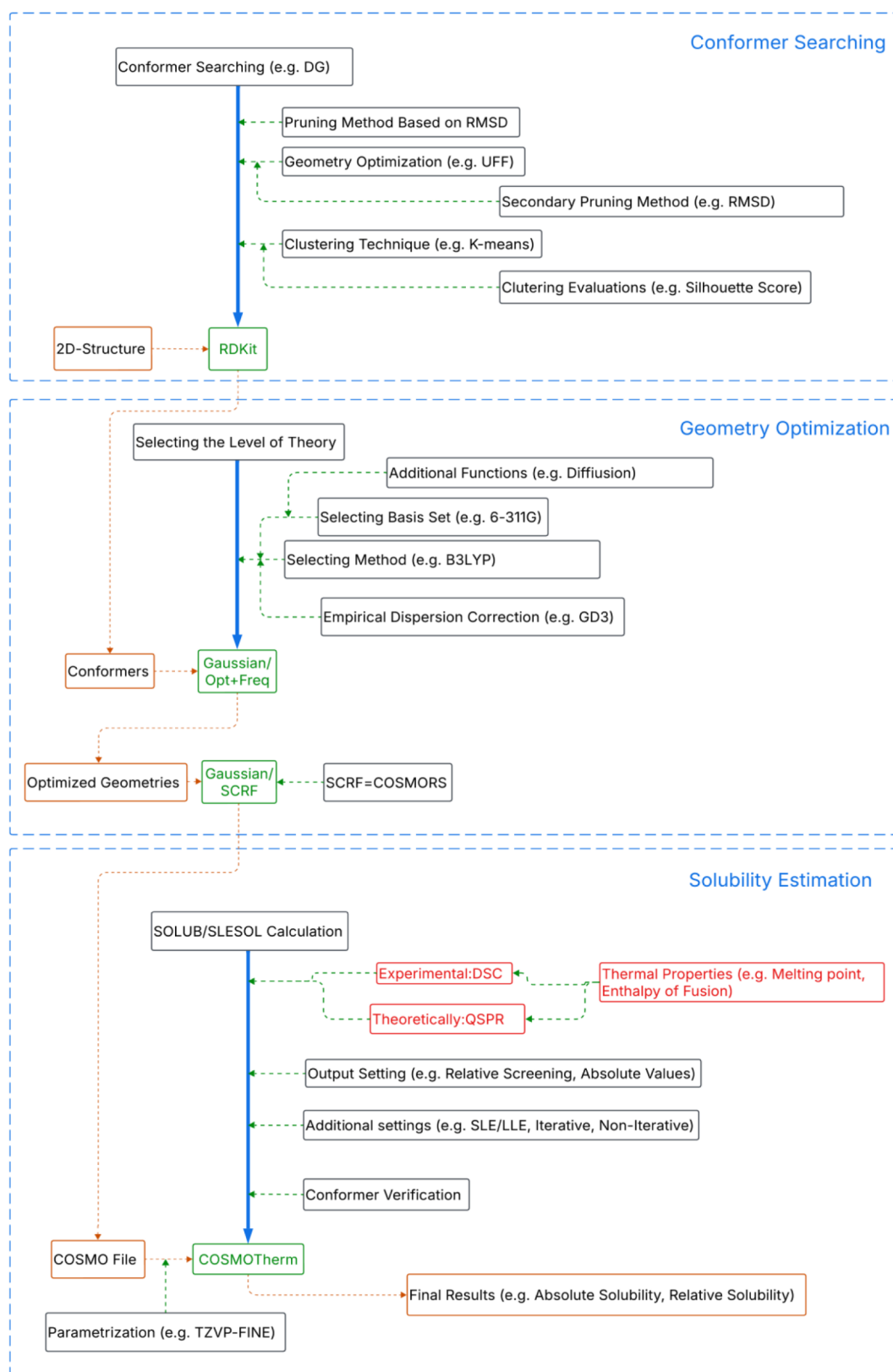


Figure 2. QC-driven workflow for thermodynamic property estimation in COSMO-RS (COSMOtherm).

library of conformers. Subsequently, the resulting geometries are optimized using static calculations with the Gaussian 16 (revision C.01) software, and COSMO input files are generated. Finally, these inputs are collected and analyzed in COSMOtherm 20.0.0 (revision 5273M) to predict the relevant

thermodynamic properties and to estimate the solubilities of the selected components in two solvents, water and ethylene glycol. Below, the execution of each step is discussed in detail.

Performing a proper conformer screening to identify local or global minimum energy structures is an essential step. For

small molecules, exhibiting a unique, well-defined conformation, using the lowest-energy conformer is typically sufficient.^{18,19} However, (N)IAS such as additives are often relatively large (350–1000 g/mol or higher) and flexible molecules, for which the potential energy surface is quite complex, exhibiting multiple and sometimes nearly isoenergetic local minima. Here, the choice of the initial structure impacts the outcome of the optimization algorithm. To improve the likelihood of finding the global minimum, a robust conformer search should be implemented to ensure all diverse and relevant conformers are considered during optimization.²⁰ Different software solutions are available to automate the search for various conformations, such as RDKit, which is a versatile open-source toolkit widely used in cheminformatics and computational chemistry, as it can start from formats like SMILES (Simplified Molecular Input Line Entry System) and InChI (International Chemical Identifier), making it particularly valuable when no initial 3D structure is available.^{21,22}

It should be noted that the accuracy of COSMO-based models, such as COSMO-RS, is sensitive to the chosen quantum-chemical (QC) level of theory. Furthermore, COSMO-RS accuracy declines in aqueous systems and halogenated compounds due to simplified hydrogen-bond treatment and challenges in accounting for dispersion and polarization,²³ which is particularly relevant for halogenated additives in polymers. Finally, conformer selection also plays a critical role as partition coefficients are shown to be even more sensitive to conformational diversity than solubility. In particular, flexible molecules, especially ones capable of intramolecular hydrogen bonding, require that all relevant low-energy conformers be taken into account for accurate predictions.^{24,25} Consequently, in many cases, validation of QC-derived structures is essential and can be achieved by comparing calculated results with experimental reference data, such as vapor pressure or activity coefficients for liquids, and solubility for solids.^{23,24}

2.1. Conformer Searching

The first step consists of the identification of various possible conformers. In this preoptimization step, RDKit's package was used in a Python environment, and the "EmbedMultipleConfs" function from RDKit's "rdDistGeom" module was employed to generate 1,000 conformers for each molecule from their 2D mol files. This function uses the basic DG (Distance Geometry) algorithm to compute a distance matrix based on molecular topology and employs random coordinate embedding to create initial conformations. The parameters "useRandomCoords = True" and "pruneRmsThresh = 1" were applied to ensure the generation of a diverse set of conformers while filtering out those that were too similar. However, while DG (Distance Geometry) explores a broader and less constrained conformational space, unlike ETKDG (Experimental Torsion Knowledge Distance Geometry), DG does not account for torsion angle preferences, the planarity of aromatic systems, or steric clash avoidance. To address this, all generated conformers are optimized using the UFF (Universal Force Field) optimization tool.²⁶ Following optimization, an additional Root-mean-square deviation (RMSD) based pruning (threshold of 1 Å) was performed to eliminate conformers that converged into more similar structures.^{26–28} The output of conformer generation is usually a library of conformers, which are organized based on a structural similarity descriptor. This descriptor is calculated based on the RMSD of atoms

coordinated between two conformers. A higher RMSD indicates greater differences in the atomic configurations between the conformers. For an output of *N* conformers, an *N* × *N* symmetric matrix is constructed and is then further processed for clustering.²⁹

For structures lacking structural diversity due to a simple, rigid backbone or few rotatable bonds (i.e., number of conformers <4), all of the conformers were geometrically optimized. For molecules with more than four remaining conformers, K-means clustering (using the scikit-learn implementation with K-means++ initialization)³⁰ was applied to select representative conformer structures based on RMSD.³¹ The number of rotatable bonds was calculated using "CalcNumRotatableBonds" from "rdMolDescriptors". The optimal number of clusters was determined by evaluating three validation metrics: the Silhouette Score (SiS), the Calinski-Harabasz Index (CHI), and the Davies-Bouldin Index (DBI).³² While no single metric is universally superior, combining them provides complementary insights to ensure compact and meaningful clustering.^{29,32}

Silhouette Score (SiS) shown by eq 1 is a clustering metric that measures how well each data point fits within its assigned cluster compared to other data points. It calculates the average distance to points within the same cluster and those in the nearest neighboring cluster. Scores range from −1 to 1, where values near 1 indicate well-defined clusters, values near 0 indicate overlapping clusters, and values near −1 suggest misclassification. However, the score may be less reliable when clusters differ significantly in size or density.³³

$$\begin{aligned} \text{SiS} &= \frac{1}{NC} \sum_x \frac{b_i - a_i}{\max(b_i, a_i)}, \quad a_i \\ &= \frac{1}{n-1} \sum_{y \in C, y \neq x} d(x, y), \quad b_i \\ &= \min_{j, j \neq i} \left[\frac{1}{n_j} \sum_{y \in C_j} d(x, y) \right] \end{aligned} \quad (1)$$

The Calinski-Harabasz Index (CHI)³⁴ shown by eq 2 evaluates clustering by comparing between-cluster dispersion to within-cluster compactness. It uses the ratio of between-cluster to within-cluster sum of squares. Higher CHI suggests better-defined clusters; however, the index has no upper bound and may favor clusters of similar size, thereby limiting reliability for clusters with varied shapes or densities.^{32,34}

$$\text{CHI} = \frac{\sum_i \frac{n_i \|c_i - c\|^2}{NC-1}}{\sum_i \sum_{x \in C_i} \frac{\|x - c_i\|^2}{n - NC}} \quad (2)$$

The Davies-Bouldin Index (DBI)³⁵ shown by eq 3, scores clustering by averaging, for each cluster, the highest ratio of intracluster dispersion to intercluster separation. Dispersion is the average distance within a cluster; separation is the centroid distance to the most similar cluster. A lower DBI (ideally 0) indicates compact, well-separated clusters. DBI tolerates size imbalance but may favor more clusters by reducing both dispersion and centroid distances.^{35,36}

$$DBI = \frac{1}{NC} \sum_{j \neq i} \max \left(\frac{\frac{1}{n_i} \sum_{x \in C_i} d(x, c_i) + \frac{1}{n_j} \sum_{x \in C_j} d(x, c_j)}{d(c_i, c_j)} \right) \quad (3)$$

In the equations above, the commonly used variables are NC, which denotes the number of clusters, c_i , the centroid of cluster i , and n_i , the number of data points assigned to that cluster. The symbols x and y represent individual points, while $d(x,y)$ denotes the Euclidean distance between them. For the Silhouette Score (SiS), a_i refers to the mean intracluster distance of point x , and b_i corresponds to the mean nearest-cluster distance.

Clustering metrics were computed using Python functions from the sklearn.metrics library. The optimal number of clusters was analyzed from two upward until limited by the total number of conformers or computational resources. The optimal cluster count was qualitatively chosen when at least two metrics aligned, aiming to strike a balance between the identification of distinct geometries and computational efficiency.^{28,32}

2.2. Geometry Optimization

Subsequently, DFT calculations to perform geometry optimizations on the RDKit output structures were carried out using the Gaussian software. The hybrid B3LYP functional was chosen as the level of theory, as it is known to give reliable geometry predictions.²³ This method was supplemented with additional Grimme D3 dispersion corrections to properly account for dispersion (van der Waals) interactions.³⁷

This study employs three Pople-style basis sets of increasing complexity: 3-21+G(d), 6-311++G(d,p), and 6-311++G(3df,3pd). The 3-21+G(d) offers low computational cost, and its accuracy is evaluated relative to larger split-valence sets.³⁸ The 6-311++G(d,p) incorporates additional diffuse and polarization functions that enhance the description of non-covalent interactions. The 6-311++G(3df,3pd) further extends polarization on both heavy atoms and hydrogens, enabling higher accuracy for systems requiring greater angular flexibility.³⁹ A vibrational analysis is carried out on the optimized geometries to ensure a stable (local) minimum is found with no imaginary frequencies.

The SCRF module with COSMO-RS settings is then used to generate the COSMO file for the COSMO-RS model. This file is used as input for COSMOtherm to estimate thermodynamic properties and calculate the total corrected energies.

2.3. Solubility Estimation

The final step in this computational workflow is the use of COSMOtherm, a collection of modules based on the COSMO-RS thermodynamic model.^{40–42} In this model, the chemical potential of compound “ i ” in solvent “ S ” is calculated by integrating $\mu_S(\sigma)$ over the sigma surface by eq 4:

$$\mu_S^i = \mu_{C,S}^i + \int p^i(\sigma) \times \mu_S(\sigma) d\sigma \quad (4)$$

In this equation, $\mu_{C,S}^i$, which depends on the area and volume of the compounds in the mixture, is the combinatorial term accounting for the size and shape differences of the molecules. The chemical potential of surface segments is calculated using eq 5:

$$\mu_S(\sigma) = -\frac{RT}{a_{eff}} \ln \left[\int p_S(\sigma') \exp \left(\frac{a_{eff}}{RT} (\mu_S(\sigma') - E_{misfit}(\sigma, \sigma') - E_{HB}(\sigma, \sigma')) \right) d\sigma' \right] \quad (5)$$

This equation uses the effective contact area a_{eff} , donor sigma potential $p_S(\sigma')$, misfit Energy E_{misfit} , and hydrogen bonding energy E_{HB} , which is a function of hydrogen bond strength and threshold for hydrogen bonding.

Subsequently, the activity coefficient and solubility of each solute are calculated using eqs 6 and 7, respectively:

$$\gamma_S^i = \exp \left(\frac{\mu_S^i - \mu_i^i}{RT} \right) \quad (6)$$

$$\log x_j = \log \left(\exp \left(\frac{\mu_j^{pure} - \mu_j^{solvent} - \Delta G_{f,j}^{fusion}}{RT} \right) \right) \quad (7)$$

For the process of dissolution, the total Gibbs free energy can be written as the sum of two components as eq 8:⁴³

$$\Delta G_{i,dissolution} = \Delta G_{i,fusion} + \Delta G_{i,mixing} \quad (8)$$

This equation shows that the Gibbs free energy of dissolution for solids has two contributors: the first term accounts for the energy required to transition from the solid or crystalline phase to the liquid phase, while the second term arises from the interactions and mixing between the solute in its supercooled liquid form and the solvent molecules, which include effects related to molecular size and shape (combinatorial), electrical forces, hydrogen bonding, dispersion forces, and polarization.^{44,45}

COSMO-RS was initially developed for liquid-phase equilibria, making accurate activity coefficient predictions in solid–liquid equilibrium (SLE) particularly challenging for large, flexible molecules.^{24,46} Estimating solid solubility with COSMO-RS also requires thermal corrections using melting point and enthalpy of fusion, obtained experimentally (e.g., DSC) or predicted via Group Contribution (GC) theory⁴⁷ or Quantitative Structure–Property Relationship (QSPR)⁴⁸ which relates the molecular descriptors to thermodynamic properties. However, advancements in QSPR methodologies have improved the accuracy of these thermal property estimations over classic GC theories.⁴⁹

For the enthalpy of fusion, COSMOquick v.2020 employs six molecular descriptors in a multivariate linear regression model trained on a data set of 2,332 organic molecules. In contrast, melting point prediction is more challenging due to its nonlinear nature.⁵⁰ Therefore, a random forest regression was trained on 12,000 compounds with 20 molecular descriptors (all descriptors are listed in Table S.1).⁴⁸ In this paper, the predictive performance of these QSPR models for melting point and enthalpy of fusion was evaluated by comparing them to DSC results. The DSC procedure is described in Section 3.3.

The Gibbs free energy of fusion (ΔG_{fusion}) is a key parameter that determines whether the melting process is spontaneous at a specific temperature. At the melting point (T_m), the solid and liquid phases are in equilibrium, and $\Delta G_{fusion} = 0$. Below T_m , ΔG_{fusion} is positive, making fusion nonspontaneous. The general expression for the Gibbs free energy of fusion is given by eq 9:⁴⁸

Table 1. List of Additives, Monomers, and Solvents Investigated in This Study

Category	Chemical name	Chemical Formula	Trade name/Abb.	Chemical Family	Cas no	Company	Purity/Quality
Anti-Microbial	Methyl 4-hydroxybenzoate	C ₈ H ₈ O ₃	Methyl Paraben	Aromatic ester—Parabens	99-76-3	Sigma-Aldrich	CRM
Antioxidant	1,3,5-Trimethyl-2,4,6-tris(3,5-ditert-butyl-4-hydroxybenzyl)benzene	C ₅₄ H ₇₈ O ₃	Irganox 1330	Primary phenolic	1709-70-2	Sigma-Aldrich	99%
	Pentaerythritol tetrakis(3,5-ditert-butyl-4-hydroxyhydrocinnamate)	C ₇₃ H ₁₀₈ O ₁₂	Irganox 1010	Primary phenolic	6683-19-8	Sigma-Aldrich	+95%
	Tris(3,5-ditert-butyl-4-hydroxybenzyl) isocyanurate	C ₄₈ H ₆₉ N ₃ O ₆	Irganox 3114	Primary phenolic	27676-62-6	Supelco	CRM/MQ300
UV Stabilizer	2-(2H-Benzotriazol-2-yl)-4,6-bis(1-methyl-1-phenylethyl)phenol	C ₃₀ H ₂₉ N ₃ O	Chiguard-234	Hydroxyphenyl benzotriazole	70321-86-7	Sigma-Aldrich	MQ100
	2-(4,6-Diphenyl-1,3,5-triazin-2-yl)-5-[(hexyl)oxy]-phenol	C ₂₇ H ₂₇ N ₃ O ₂	UV-1577	Hydroxyphenyl triazine	147315-50-2	Sigma-Aldrich	MQ100
	2,2'-(1,4-Phenylene)bis(4H-3,1-benzoxazin-4-one)	C ₂₂ H ₁₂ N ₂ O ₄	Cyasorb UV-3638F	Benzoxazinone	18600-59-4	TCI	+98%
Acetaldehyde Scavenger	Anthranilamide	C ₇ H ₈ N ₂ O	—	Amide	88-68-6	Sigma-Aldrich	+98%
Flame Retardant	Triphenyl phosphate	C ₁₈ H ₁₅ O ₄ P	TPhP	Aromatic phosphate ester	115-86-6	Sigma-Aldrich	+99%
	Dibromoneopentyl Glycol	C ₅ H ₁₀ Br ₂ O ₂	FR-522 (DBNPG)	Hydroxyl—Brominated	3296-90-0	Sigma-Aldrich	98%
Pigment	1,4-bis(p-tolylamino)anthraquinone	C ₂₈ H ₂₂ N ₂ O ₂	C.I. Solvent Green 3	Anthraquinones	128-80-3	Sigma-Aldrich	95%/MQ200
Monomers	Bis(2-Hydroxyethyl) terephthalate	C ₁₂ H ₁₄ O ₆	BHET	Terephthalate diol ester	959-26-2	Sigma-Aldrich	MQ100
	Terephthalic acid	C ₈ H ₆ O ₄	TPA	Terephthalate	100-21-0	Sigma-Aldrich	+98%
	Dimethyl terephthalate	C ₁₀ H ₁₀ O ₄	DMT	Terephthalate	120-61-6	Sigma-Aldrich	+99.0%
Solvents	Ethylene glycol	C ₂ H ₆ O ₂	EG	diol	107-21-1	Sigma-Aldrich	+99%
	Water	H ₂ O	—	—	7732-18-5	Sigma-Aldrich	HPLC

$$\Delta G_{\text{fusion}} = \Delta H_{\text{fusion}} - T \Delta S_{\text{fusion}} - \Delta C_{p,\text{fusion}} \left[(T_m - T) + T \ln \left(\frac{T}{T_m} \right) \right] \quad (9)$$

where ΔH_{fusion} is the enthalpy of fusion, ΔS_{fusion} is the entropy of fusion, and $\Delta C_{p,\text{fusion}}$ is the change in heat capacity during the phase transition. To simplify this equation, the change in heat capacity is assumed to be approximately zero, $C_{p,\text{fusion}} \approx 0$. This assumption is valid for temperatures close to the melting point.⁵¹ Considering this condition, the equation reduces to eq 10:

$$\Delta G_{\text{fusion}} = \Delta H_{\text{fusion}} - T \Delta S_{\text{fusion}} \quad (10)$$

At the melting point ($T = T_m$), the system is at equilibrium, so $\Delta G_{\text{fusion}} = 0$. From this condition, the entropy change during fusion can be calculated as eq 11:

$$\Delta S_{\text{fusion}} = \frac{\Delta H_{\text{fusion}}}{T_m} \quad (11)$$

Substituting eq 11 for ΔS_{fusion} into eq 10, it will be simplified to eq 12:

$$\Delta G_{\text{fusion}} = \Delta H_{\text{fusion}} \left(1 - \frac{T}{T_m} \right) \quad (12)$$

Furthermore, the solvents, water, and ethylene glycol, were selected from the COSMOtherm v.20 database, and COSMO files were imported with the TZVPD_FINE parametrization.

The T_m and ΔH_{fus} values were then incorporated into the compounds' properties.

When available, comparing estimated properties, such as boiling point or vapor pressure, with experimental data (self-association) helps identify anomalous conformers. For solids, solubility estimates enable comparison within a conformer set or with experimental data.^{23,25} In the absence of experimental data, nonparametric outlier detection methods can be used to detect outlier conformers, as conformer solubilities do not follow a normal distribution. Models such as the modified Z-score, presented as eq 13, can then be used to identify conformers that poorly align with others.⁵²

$$Z_i = \frac{0.6745 \times (x_i - \tilde{x}_i)}{\text{MAD}} \quad (13)$$

In which, $\tilde{x}_i$ and MAD are the median of data points, and the median absolute deviation of the data points, respectively.

Finally, to predict the absolute solubility, conformers were combined into a single set. The conformers are weighted internally using the Boltzmann distribution, considering their energy, chemical potential, and dielectric energy correction with a weight prefactor that accounts for degeneracy and symmetry.¹⁸ Finally, the SLE/LLE (Liquid–Liquid Equilibrium) option is selected, and to specify ΔG_{fus} , the experimental option was chosen to incorporate the DSC data.

The deviations from the experimental solubilities are calculated using the RMSE of log units in ppm over the relevant data by eq 14:

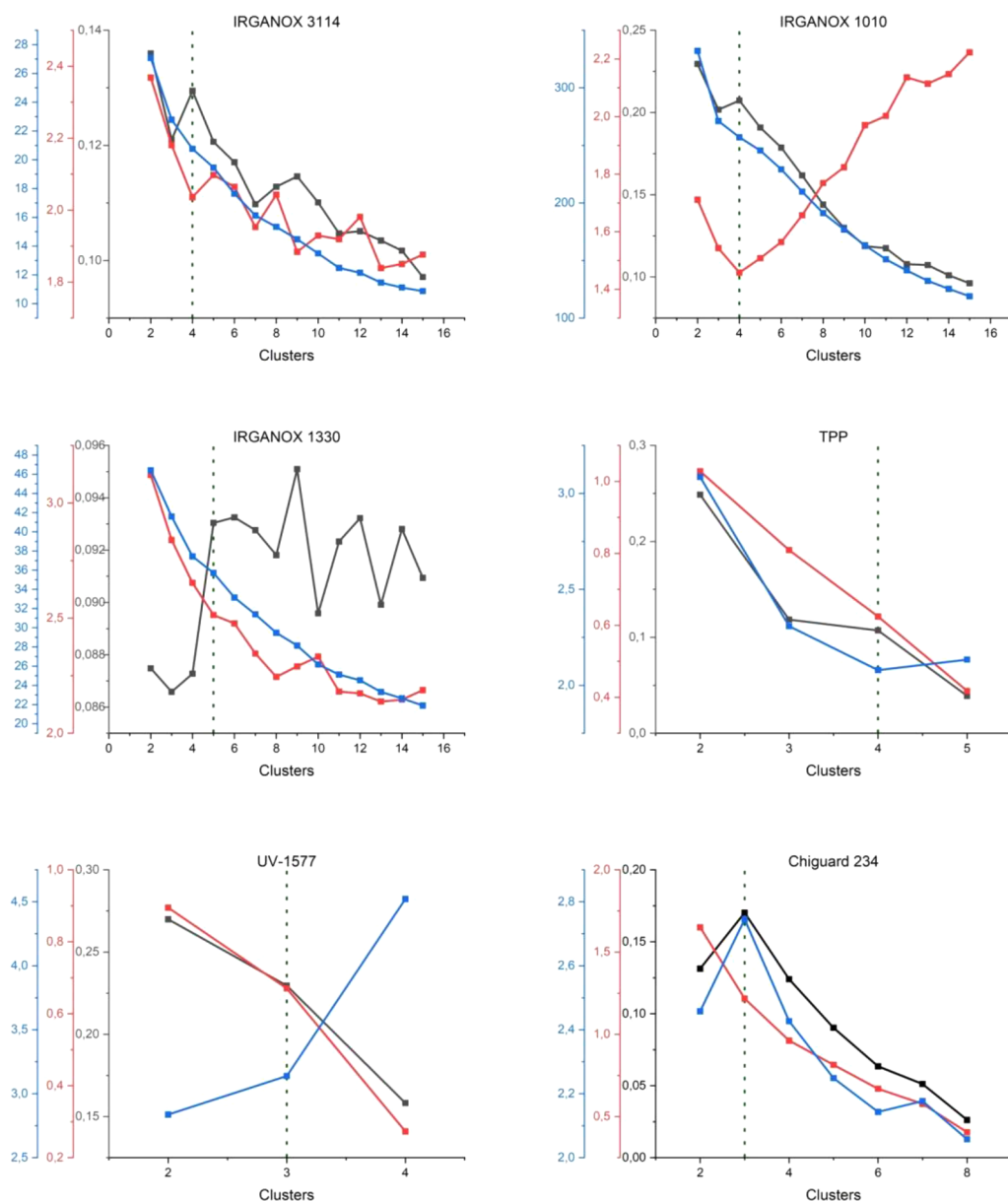


Figure 3. Evaluation of clustering performance using three metrics: SiS (Black), DBI (Red), CHI (Blue), and qualitatively selected cluster number (dotted line, Green).

$$\sigma = \sqrt{\frac{\sum (\log(w_{ppm,exp.}) - \log(w_{ppm,est.}))^2}{N}} \quad (14)$$

3. EXPERIMENTAL PROCEDURES

3.1. Sample Preparation

To determine the maximum solubility, samples were prepared by filling 40 mL screw-capped, clear glass N24 flat-bottom vials with either water or ethylene glycol. Solutes, as listed in Table 1, were gradually added until visible precipitation occurred, indicating solubility limits. The vials were then placed in a thermal bath equipped with a magnetic stirrer, using an IKA-MAG-HS7 heater-stirrer plate. The solution temperature was controlled at 25 ± 0.1 °C using an IKA-ETS-D5 temperature probe suspended in the bath. Each mixture was stirred continuously for 7 days to ensure equilibrium, as indicated by the presence of visible precipitations, and to allow sufficient time to reach SLE in case the kinetics were slow. To prevent

degradation from light or oxygen exposure, the vials were filled to the top and stored in a dark environment.

After equilibration, a 5 mL aliquot was withdrawn from each vial and centrifuged at 7000 rpm for 10 min using a UNIVERSAL-320R centrifuge. The temperature was controlled by an AD07R-20 refrigeration/heating bath (462–0226). This step removed suspended particles without disturbing the SLE. The resulting supernatant was then filtered through 0.1 μ m PTFE syringe filters to remove any remaining fine particulates, yielding a clear solution suitable for further analysis. All postprocessing steps were performed in triplicate on independent samples.

3.2. HPLC Analysis

To analyze the maximum solubility, an Agilent 1260 Infinity II HPLC system (G7111B) equipped with an autosampler (G7129A) and a photodiode array (PDA) detector (1260 DAD HS) was used. Column temperature was maintained isothermally at 25 ± 0.5 °C using a 1260-MCT column oven throughout the analysis. Method conditions, including gradient parameters, PDA detection wavelengths, analytical column specifications, and corresponding retention

times, are provided in Table S.2. The limits of detection (LOD) and quantification (LOQ) were calculated based on the standard deviation of the response and the slope of the calibration curve, as described in eqs 15 and 16, respectively.⁵³

$$LOD = \frac{3.3 \times \sigma}{S} \quad (15)$$

$$LOQ = \frac{10 \times \sigma}{S} \quad (16)$$

where σ is the standard deviation of calibration points and S is the slope of the calibration curve.

3.3. DSC Analysis

Differential Scanning Calorimetry (DSC) was performed to obtain calorimetric thermograms. A measured amount of the sample was weighed using a KERN ABJ-NM/ABS-N analytical balance and press-sealed in aluminum crucibles (Concavus, Netzsch). Thermal analysis was performed using a NETZSCH 204 F1 DSC at a heating rate of 10 K/min. The enthalpy of fusion (ΔH_{fus}) was determined from the area under the melting peak, and the peak temperature of the melting transition was recorded as the melting point (T_m) of the compound.

4. RESULTS AND DISCUSSION

4.1. Conformer Searching

Conformational sampling of plastic additives and monomers (Table S.3) highlights the variability in molecular flexibility. Rigid molecules such as methyl paraben, TPA, and DMT yielded only a single conformer. Moderately flexible compounds such as TPhP, UV-1577, and Chiguard-234 produced between 5 and 9 conformers with average RMSDs between 2.6 and 3.1 Å, while more flexible compounds like Irganox 1010 (24 rotatable bonds) and Irganox 1330 (12 rotatable bonds) generated a higher number of conformers (457 and 1000 structures, respectively) with average RMSDs of 6.5 Å and 4.7 Å.

As explained in the Section 2.1, the selection of the optimal number of clusters (Figure 3), was performed by evaluating the Silhouette Score (SiS), Davies-Bouldin Index (DBI), and Calinski-Harabasz Index (CHI). Clustering evaluation was initiated from two clusters upward to reduce the number of conformers while retaining those with potentially different interactions with solvents. A given cluster number was selected when at least two of the three metrics showed consistent behavior (i.e., local extrema or favorable trends), after which representative conformers were extracted.

Using these criteria, four clusters were selected for Irganox 3114, where both SiS and DBI exhibited local extrema. For Irganox 1010, four clusters were chosen because the SiS reached a local maximum and the DBI reached a global minimum, indicating compact and well-separated clusters. For Irganox 1330, five clusters were selected based on a pronounced increase in SiS and a concurrent decrease in DBI, while the declining CHI suggested shorter distances between cluster centroids. In the case of TPhP, four clusters were deemed appropriate, as the DBI continued to decrease from three to four clusters while the SiS remained stable. For UV-1577, three clusters were selected based on alignment between DBI and CHI. Finally, Chiguard-234 showed a global maximum in both SiS and CHI at three clusters, accompanied by a decrease in DBI, supporting this choice.^{29,32}

The observation that some larger or more structurally complex molecules yield fewer effective clusters can be attributed to steric constraints. Highly branched structures restrict the diversity of accessible conformations, leading to

fewer distinct clusterable geometries. This clustering approach offers advantages, including computational efficiency and explicit consideration of molecular geometry rather than relying solely on conformer energies. However, it also has limitations, as subtle differences in functional-group orientation that may influence solubility can be overlooked. Such methods, however, are not fully optimized and should be regarded as semiquantitative, serving primarily as a practical strategy for reducing conformer space while preserving major geometrical variations relevant to solvation.

4.2. Comparing Experimental and Model Solubility

For the number of clusters selected, the solubility of each representative conformer per cluster and their set was calculated with three different basis sets and compared with the experimental values (Table 2) and the estimated log-solubilities (above 1×10^{-3} ppm) are shown in Figure 4. As a general trend, higher concentrations/solubilities show better agreement between predicted and experimental values in both water and ethylene glycol. For example, using the 6-311++G(d,p) basis set, DBNPG shows moderate deviations ($|\Delta \log(\text{ppm})| = 0.51$ in water and 0.64 in ethylene glycol), while anthranilamide and BHET display larger deviations in water (0.79 and 0.81, respectively) but smaller deviations in ethylene glycol (0.36 and 0.41, respectively). In contrast, methylparaben and TPhP showed the lowest deviations in ethylene glycol (0.07 and 0.06, respectively) compared to the studied compounds. COSMO-RS has been widely used to predict solids' solubility, particularly for APIs, with typical RMSE values less than ± 1.5 log units, commonly between ± 0.5 and ± 1 .^{24,54-57} Deviations exceeding ± 1.5 log units mainly occurred for compounds with solubilities below 100 ppm. This is valid for all 3 basis sets. For depicted values in Figure 4, the 3-21+G(d) yielded a high RMSE of 2.12 log units due to poor treatment of heavy atoms and even failed to optimize brominated compounds (e.g., FR-522). Instead, the 6-311++G(d,p) basis set required more optimization steps and job restarts, particularly for large molecules, but this effort improved the accuracy with an RMSE of 1.31 log units. However, deviations remained for some of the poorly soluble compounds (e.g., UV-1577, C.I. Solvent Green 03). The more flexible and computationally intensive option, 6-311++G(3df,3pd), showed similar trends. This basis set, achieved a slightly better RMSE of 1.28 log units. However, it still struggled with underestimation for additives containing heavier elements (phosphorus in TPhP and bromine in FR-522 (DBNPG)). Furthermore, while geometry optimizations converged efficiently, frequency calculations for large molecules were typically prone to restarts, such that for the Irganox 1010, did not reach a checkpoint and therefore could not be completed.

Outliers were identified using modified Z-scores calculated across all solvent-basis set combinations (Table S.4). Although a conventional threshold of 3.5 is typically used, a more conservative threshold of 10^3 was applied to account for the inherently wide variability in solubility predictions among conformers, particularly for flexible, high-molecular-weight molecules.^{25,58} The discrepancy observed for Irganox 3114 (Conf. 3) at lower-level basis sets ($Z = 2.11 \times 10^3$) was resolved with higher-level methods, whereas Irganox 3114 (Conf. 1) and Irganox 1330 (Conf. 5) remained extreme outliers ($Z > 1 \times 10^6$) and were excluded from the set, likely due to suboptimal initial geometries or inaccurate electronic structures. The SLE algorithm failed for some Irganox 1330 conformers and all Irganox 1010 conformers, yielding "N/A" results due to COSMOtherm limitations at extremely low solubility levels. Irganox 1010 also triggered frequent errors, likely due to its very low solubility and complex electronic structure associated with its large molecular size or the fact that the parametrization in COSMO-RS has been limited to small and medium-sized molecules.^{41,46}

Although a greater number of rotatable bonds generally corresponds to a more complex conformational space, Figure 5 shows that molecules with as few as two rotatable bonds can exhibit inconsistent solubility predictions. This suggests that the size of

Table 2. Experimental Solubility of Studied Compounds in Water and Ethylene Glycol, and Estimated Values Using Three Basis Sets for Individual Conformers and Their Sets at 25 °C^{a,b,c}

Comp. ID	Components	Conf. ID	3-21+G(d)				6-311++G(d,p)				6-311++G(3df,3pd)				Experimental ^a	
			Water		EG		Water		EG		Water		EG		Water	EG
			Conformer	Set	Conformer	Set	Conformer	Set	Conformer	Set	Conformer	Set	Conformer	Set	ppm	ppm
1	Irganox 3114															
		1**	1.29E-9	4.38E-4	2.10E+4	1.03E+3	5.81E-9	2.15E-10	1.50E+4	2.73E-3	4.35E-9	1.49E-10	9.78E+3	2.80E-3	6.09E-1	1.72E+0
		2	2.81E-7	1.20E-1	1.20E-1	3.27E-10	3.27E-10	1.69E-10	1.93E-3	1.74E-10	1.33E-10	3.07E-11	4.42E-3	1.31E-3		
		3	4.38E-4	1.03E+3	1.03E+3	2.26E-2	2.84E-11	4.27E-10	5.15E-2	5.00E-2	8.96E-11	8.96E-11	5.47E-4	3.00E-1	7.02E-1	1.17E-1
		4	6.98E-8	5.14E-9	3.53E+4	3.53E+4	N/A	N/A	N/A	N/A	3.02E-11	3.02E-11	2.52E-3			
2	Irganox 1330															
		1	6.09E-9	1.42E-8	2.53E-2	1.39E-2	1.84E-11	6.70E-1	1.19E-3	8.57E-12	8.57E-12	1.43E-2	7.36E-4			
		2	7.60E-10	N/A	N/A	N/A	N/A	N/A	N/A	3.23E-13	3.23E-13	2.14E+4	1.43E-2			
		3**	N/A	0	Error	0	N/A	N/A	N/A	0	N/A	0	N/A	0	1.44E-1	4.54E-2
3	Irganox 1010															
		1	Error***	Error	Error	Error	Error	Error	Error	Error	Error	Error	Error	Error		
		2	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
		3	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
		4	Error	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A		
4	Chignard-234															
		1	1.85E-4	9.60E-5	3.81E+1	3.33E+1	1.15E-5	5.96E-6	1.40E+1	9.86E+0	7.27E-6	4.64E-6	1.08E+1	8.01E+0	4.46E-2	6.82E+0
		2	9.46E-5	3.32E+1	3.32E+1	5.07E-6	5.07E-6	1.75E-5	9.40	4.08E-6	4.08E-6	7.63	1.12E+1			
		3	4.31E-4	5.26E+1	5.26E+1	7.06E+2	8.36E-5	4.48E-4	1.64E+2	3.17E+2	4.41E-5	2.28E-4	1.32E+2	2.66E+2	4.85E-2	1.01E+0
5	UV1577															
		1	1.91E-3	4.57E-2	3.73E+2	7.06E+2	4.51E-4	1.07E-4	3.18E+2	2.29E-4	2.29E-4	5.80E-5	1.73E+2	2.67E+2		
		2	4.57E-2	7.06E+2	7.06E+2	4.16E+2	1.07E-4	2.03E+2	2.03E+2	1.47E+4	1.45	1.25E+0	1.09E+4	9.89E+3	3.22E+0	1.63E+4
		3	2.73E-3	4.10E+4	4.10E+4	3.56E+4	2.08	2.03E+0	1.47E+4	1.43E+4	1.45	1.25E+0	1.09E+4	9.89E+3	3.22E+0	1.63E+4
6	TPHP															
		1	1.52E+1	1.51E+1	4.10E+4	3.56E+4	2.08	2.03E+0	1.47E+4	1.43E+4	1.45	1.25E+0	1.09E+4	9.89E+3	3.22E+0	1.63E+4
		2	1.47E+1	2.56E+4	2.56E+4	3.11E+4	1.50	2.08	1.46E+4	1.46E+4	1.34	1.08E+4	1.08E+4	1.08E+4		
		3	1.99E+1	3.11E+4	3.11E+4	4.09E+4	2.04	2.08	8.51E+3	6.20E-1	6.20E-1	6.13E+3	6.13E+3	6.13E+3		
		4	1.52E+1	4.09E+4	4.09E+4	4.09E+4	2.04	2.08	1.46E+4	1.46E+4	1.44	1.09E+4	1.09E+4	1.09E+4		
7	Cyasorb UV3638-F															
		1	5.30E-2	5.43E-2	7.26E+1	7.14E+1	1.64E-3	1.74E-3	1.57E+1	1.59E+1	1.01E-3	1.13E-3	1.17E+1	1.19E+1	2.15E-2	2.38E-1
		2	5.64E-2	7.01E+1	7.01E+1	1.86E-3	1.86E-3	1.63E+1	1.63E+1	1.27E-3	1.27E-3	1.22E+1	1.22E+1	1.22E+1		
8	FR-522 (DBNPG)															
		1	N/A	N/A	N/A	N/A	3.88E+4	6.79E+3	2.07E+5	9.48E+4	2.56E+4	4.18E+3	2.83E+5	1.07E+5	2.22E+4	4.16E+5
		2	N/A	N/A	N/A	N/A	6.41E+3	9.61E+4	9.61E+4	3.83E+3	3.83E+3	3.12E+3	1.06E+5	1.06E+5		
		3	N/A	N/A	N/A	N/A	5.04E+3	7.41E+4	7.41E+4	3.12E+3	3.12E+3	1.09E+4	7.06E+4	6.00E+4	7.27E+3	3.32E+4
9	BHET															
		1	7.68E+5	7.68E+5	4.49E+5	4.54E+5	3.50E+4	4.70E+4	9.34E+4	8.46E+4	1.01E+4	1.09E+4	6.78E+4	6.00E+4	7.27E+3	3.32E+4
		2	7.58E+5	8.84E+4	4.58E+5	1.17E+5	5.21E+4	6.82E+3	7.67E+4	1.19E+5	1.11E+4	3.90E+3	5.54E+4	1.18E+5	2.31E+3	1.01E+5
10	Methyl Paraben															
		1	4.30E+4	4.30E+4	2.58E+5	2.58E+5	2.31E+4	2.31E+4	2.16E+5	2.16E+5	1.41E+4	1.41E+4	2.00E+5	2.00E+5	3.72E+3	9.37E+4
11	Anthraniamide															
		1	1.43E+3	1.43E+3	1.75E+4	1.75E+4	2.83E+0	2.83E+0	2.08E+3	2.08E+3	6.30E-1	6.30E-1	1.34E+3	1.34E+3	5.20E-1	6.33E+0
12	C.I. Solvent Green 3															
		1	4.09E+2	4.09E+2	6.05E+0	6.05E+0	6.61E+0	6.61E+0	6.14E+0	6.14E+0	4.40E+0	4.40E+0	5.52E+0	5.52E+0	5.46E+1	3.45E+2
13	TPA															
		1	1.32E+3	1.32E+3	2.09E+4	2.09E+4	8.54E+1	8.54E+1	3.08E+3	3.08E+3	5.68E+1	5.68E+1	2.55E+3	2.55E+3	3.54E+1	9.81E+2
14	DMT															
		1	2.12	2.12	1.31	1.31	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.28

^aLOD and LOQ values are reported in Table S2, and stddev is reported in Table S5. ^bExcluded conformers from the conformer set. ^cError caused by extreme charge-density values in the sigma surface.

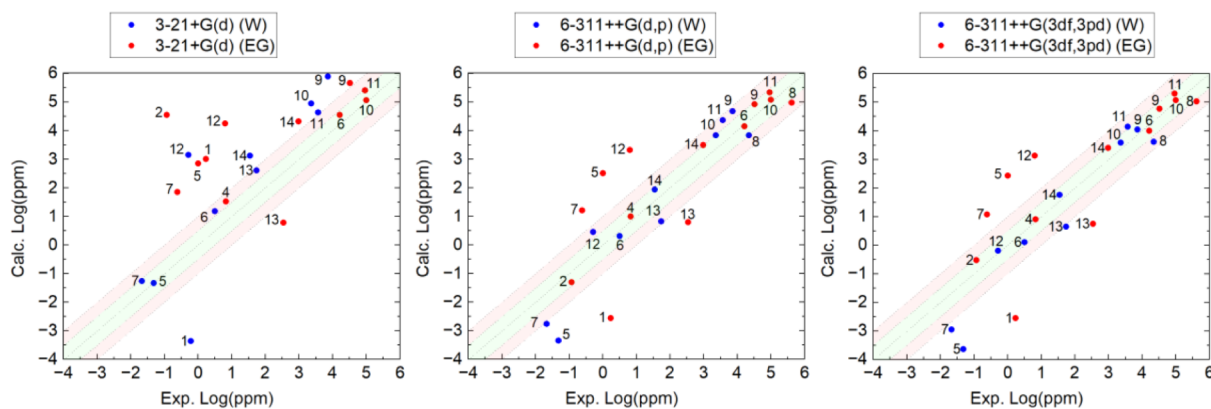


Figure 4. Log of solubilities (ppm) for studied compounds predicted using the 3–21+G(d), 6–311++G(d,p), and 6–311++G(3df,3pd) basis sets paired with B3LYP versus experimental solubility at 25 °C for water (W) and ethylene glycol (EG). ± 0.5 and ± 1 log range indicated by green and red bands, respectively. Data labels correspond to comp. ID in Table 2.

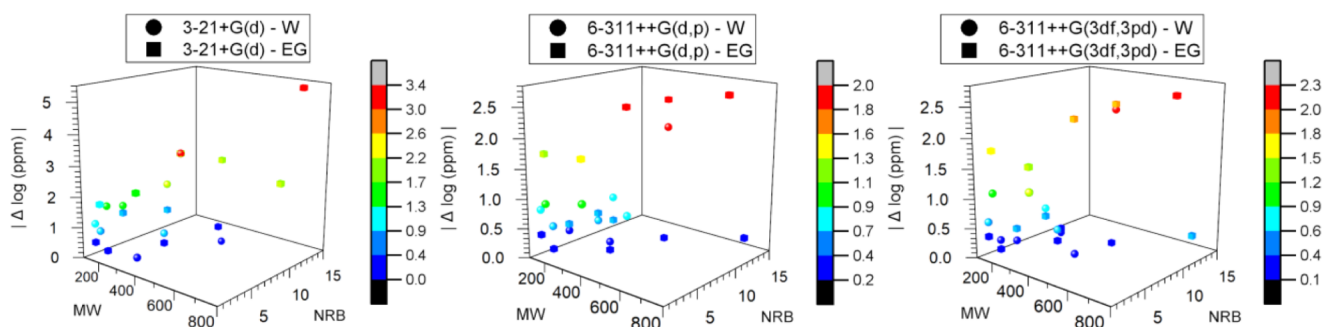


Figure 5. Comparing the accuracy of solubility estimations by COSMO-RS at 25 °C compared to the respective solutes' molecular weight and number of rotatable bonds (NRB) across all three basis sets.

fragments attached to rotatable bonds and the nature and number of functional groups, such as hindered phenols or esters capable of intramolecular hydrogen bonding, may have a larger influence.^{25,46} Improving conformer sampling by expanding the initial conformer library, relaxing the RMSD pruning threshold, or employing clustering algorithms that favor more clusters can enhance predictions for solutes above 400 g/mol, though computational cost becomes limiting.

In general, as can be seen in Figure 5, the lower-level basis set 3–21+G(d) was more sensitive to molecular size and flexibility, whereas higher-level sets (6–311++G(d,p) and 6–311++G(3df,3pd)) provided more consistent results, although occasional errors persisted for the larger molecules.⁵⁹ Traditionally, the minimum-energy conformer has been the primary focus in conformer selection.⁶⁰ COSMO-derived conformer energies (Table S.6) decrease on average by 0.56% from 3–21+G(d) to 6–311++G(d,p), with a negligible 0.007% change when optimized at 6–311++G(3df,3pd). However, as illustrated in Figure S.1, no consistent correlation exists between gas-phase conformer energy and solubility prediction accuracy, within each basis set. For larger molecules capable of intramolecular hydrogen bonding, such as Irganox 3114 or Chiguard-234, higher-energy conformers often yield more accurate results. Even small molecules like DBNPG show improved predictions from their highest-energy conformers in water and ethylene glycol.⁵⁷ This behavior reflects solvent-induced stabilization in polar solvents; functional groups can reorient to optimize hydrogen bonding and dipole interactions, resulting in conformations that are higher in energy in the gas phase but better suited to the solvent environment.

4.3. Influence of Molecule Properties on Its Solubility

Overall, ethylene glycol provided higher solubility for plastic additives and monomers compared to water. The solutes display nonpolar regions, as indicated by peaks near $\sigma = 0$ in

their σ -profiles ($-0.01 < \sigma < 0.01$)⁶¹ (Figures S.2 and S.3), which align more closely with the more nonpolar character of ethylene glycol compared to water (Figure S.4). Ethylene glycol can simultaneously engage polar functional groups through hydrogen bonding and nonpolar moieties via van der Waals interactions. Its bulkier, more flexible structure produces a less rigid hydrogen-bonding network compared to water (Figure S.5), facilitating the incorporation of solutes into the solvent matrix.

In ethylene glycol, dibromoneopentyl glycol (DBNPG) exhibits the highest solubility (41.6% w/w), consistent with its σ -profile, which shows strong polar regions from hydroxyl (–OH) groups and nonpolar regions enhanced by brominated carbons (C–Br). Hydrogen bonding with EG's (–OH) groups and increased London dispersion forces contribute to its high solubility, whereas interactions between DBNPG's quaternary carbon and EG's methylene (–CH₂–) units remain weak. Methyl paraben (10.1% w/w), anthranilamide (9.37% w/w), and BHET (3.32% w/w) display moderate solubility, driven by hydrogen bonding between their polar functional groups, ester carbonyl (C=O) and phenolic hydroxyl (–OH) in methyl paraben, amide carbonyl (C=O) and ortho-amino (–NH₂) in anthranilamide, and ester carbonyl (C=O) and hydroxyl (–OH) in BHET, and EG's hydroxyls, along with van der Waals interactions between their (partially) aromatic rings and EG's ethylene backbone. Triphenyl phosphate (TPhP, 1.63% w/w) is less soluble; its phosphoryl oxygen (P=O) forms hydrogen bonds with EG, but the three nonpolar phenyl rings impede the solvation by ethylene glycol.

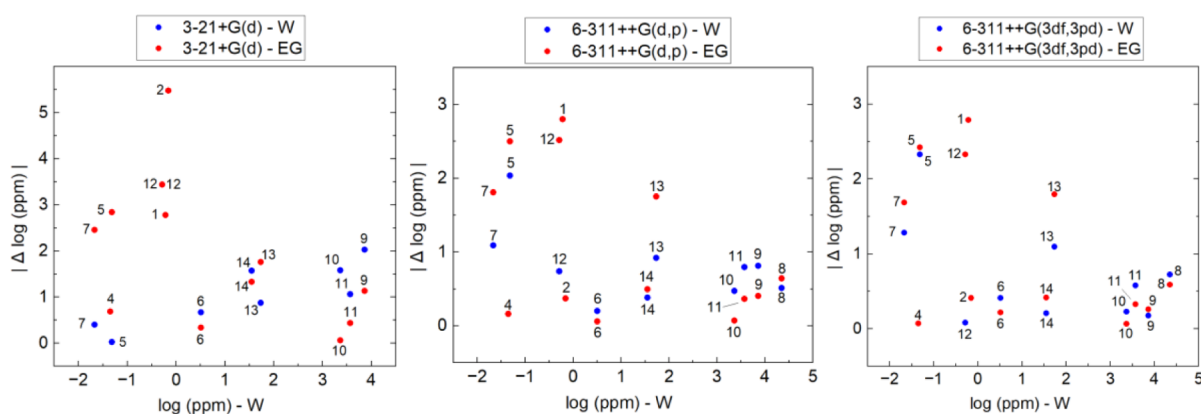


Figure 6. Comparing the accuracy of solubility estimations by COSMO-RS at 25 °C compared to the respective experimental water solubility (hydrophilicity) across all three basis sets. Data labels correspond to comp. ID in Table 2.

Table 3. Comparison of Melting Points and Enthalpies of Fusion for Studied Compounds Obtained through DSC Analysis and the QSPR Method Predictions, along with Respective Absolute Deviations and RMSEs

Component	Melting point Exp. (°C)	Melting Point (QSPR) (°C)	ΔT (°C)	Enthalpy of Fusion Exp. (kJ/mol)	NIST ref. (kJ/mol)	Enthalpy of Fusion (QSPR) (kJ/mol)	Δ (ΔHfus) (kJ/mol)
Methyl Paraben	130.0 ± 0.7	151.2	21.3	26.3 ± 1.7	24.3	20.3	6.0
Irganox 1330	246.7 ± 0.5	225.5	21.2	56.3 ± 5.5	—	56.2	0.1
Irganox 1010	117.4 ± 0.7	146.3	28.9	70.0 ± 1.6	65.9	96.1	26.1
Irganox 3114	221.9 ± 0.5	214.4	7.5	60.4 ± 1.9	—	58.5	1.9
Chiguard-234	143.9 ± 1.1	N/A ^a	N/A ^a	37.5 ± 1.9	—	N/A ^a	—
UV-1577	154.0 ± 1.7	223.5	69.5	39.5 ± 2.8	—	53.5	14.0
Cyasorb UV-3638F	317.1 ± 0.6	N/A ^a	N/A ^a	51.4 ± 3.8	—	N/A ^a	—
Anthranilamide	114.9 ± 0.2	164.3	49.6	21.4 ± 0.4	21.9	23.3	1.9
TPhP	54.4 ± 1.2	116.5	60.0	30.3 ± 1.0	29.6	33.2	2.9
FR-522 (DBNPG)	113.9 ± 3.0	88.2	25.7	26.5 ± 7.1	30.1	27.3	0.8
C.I. Solvent Green 3	224.7 ± 0.3	236.6	11.9	28.9 ± 1.9	—	42.6	13.8
BHET	109.6 ± 0.5	116.1	6.5	34.0 ± 1.2	—	35.2	1.2
TPA	427.0	312.6	114.4	63.4	—	25.3	38.1
DMT	146.7 ± 0.9	104.5	42.2	31.5 ± 0.9	31.6	22.9	8.6
Overall RMSE ^b			48.6				14.9

^aInvalid structures. ^bRMSE = $\sqrt{\frac{1}{n} \sum_i (y_{\text{exp}} - y_{\text{calc}})^2}$

In water, dibromoneopentyl glycol (DBNPG) shows the highest solubility (2.2% w/w), owing to its two hydroxyl groups that form strong hydrogen bonds and its small, compact structure, which minimizes the hydrophobic effect of the bromomethyl groups. BHET (0.73% w/w), methyl paraben (0.23% w/w), and anthranilamide (0.37% w/w) exhibit relatively low solubility despite having polar functional groups capable of hydrogen bonding. Their nonpolar aromatic rings favor the formation of a stable crystal lattice, which water cannot readily disrupt. The coexistence of polar and nonpolar regions in these molecules results in poor solubility in water compared to more compatible solvents, such as ethylene glycol, where flexible methylene groups facilitate solute incorporation.⁶² In general, the other additives and monomers exhibited low solubility in water and ethylene glycol, which can be attributed to the absence of hydrogen-bonding functional groups or, more importantly, the steric shielding of such groups by bulky nonpolar substituents.

COSMO-RS generally overestimates absolute solubilities for compounds above 100 ppm in both water and ethylene glycol. Furthermore, it has been demonstrated that COSMO-RS

overestimates the solubility of nonpolar compounds in polar solvents by neglecting hydrophobic hydration entropy and the disruption of hydrogen-bond networks (Figure 6).^{63,64} In contrast, solubility is often underestimated for compounds containing postsecond-row atoms, likely due to limitations in quantum chemical calculations, which can underestimate polarity, and the use of generalized, nonelement-specific dispersion models that do not fully account for the enhanced polarizability of these heavier atoms.⁴¹ For example, in the case of triphenyl phosphate (TPhP), this effect is less pronounced: the experimental solubility in ethylene glycol is 1.63% w/w, while the predicted value is 1.43% w/w (using the 6-311++G(d,p) basis set). In contrast, the discrepancy is much higher for DBNPG, which contains two heavy halogenated (C-Br) bonds. Its experimental solubility in ethylene glycol is 41.6% w/w, whereas COSMO-RS predicts only 9.48% w/w (using the 6-311++G(d,p) basis set). Similar trends have been observed in other studies.^{23,65} The highest absolute overestimation occurred with anthranilamide, which has an experimental solubility of 9.37% w/w, but is predicted to be 21.6% w/w (6-311++G(d,p) basis set). This is likely due to

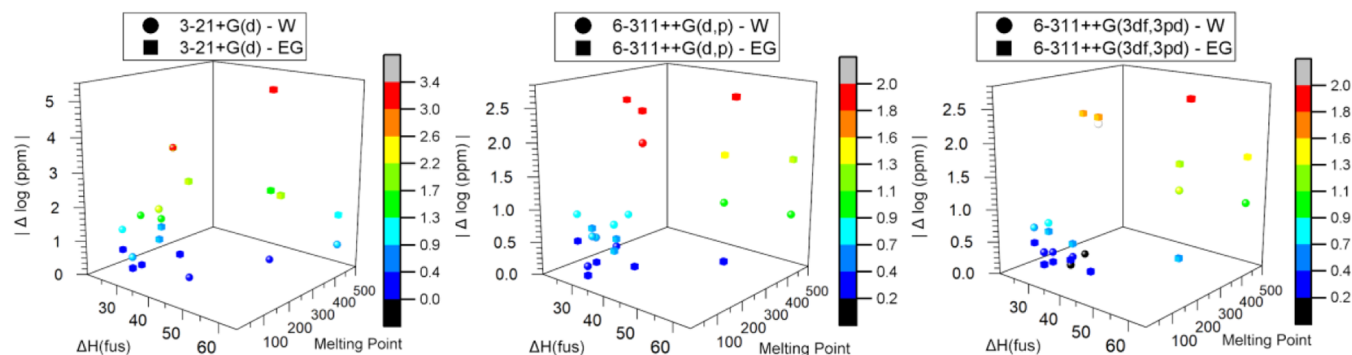


Figure 7. Comparing the accuracy of solubility estimations by COSMO-RS at 25 °C compared to the respective solutes' melting point (°C) and heat of fusion (kJ/mol).

an inaccurate hydrogen bond geometry between the ortho-amino group ($-\text{NH}_2$) and the carbonyl oxygen ($\text{C}=\text{O}$) in the input structure. Such an error can lead to an overestimation of polarity. Additionally, this hydrogen bond site may be less accessible to ethylene glycol due to steric hindrance, further contributing to the discrepancy.

4.4. Need for Experimental Enthalpy of Fusion

Melting points and enthalpies of fusion obtained from DSC and the QSPR method are summarized in Table 3. The predictions from the QSPR method show an overall observed RMSE of 48.6 °C compared to the experimental data for the melting point and 14.9 kJ/mol for the enthalpy of fusion (Relative deviation provided in Table S.8), which is in a similar order of magnitude compared to previous studies.^{48,49} However, according to a previous study, RMSEs of 22.43 and 35.26 °C in melting point lead to RMSEs of 1.20 and 2.13 in the natural log unit of molar solubility, respectively.⁴⁹ Predictions based on experimental thermal data consistently outperform those using modeled data across all basis sets. For example, the 3-21+G(d) basis set yields an RMSE of 2.42, representing a 0.30 log-unit increase relative to predictions using experimental data. Over the available data, higher-level basis sets, 6-311++G(d,p) and 6-311++G(3df,3pd), still show an increase in RMSE of 0.23 and 0.13 log units, respectively (Table S.7).

Therefore, COSMOquick's QSPR predictions can serve as an order-of-magnitude first estimation, but might not be preferred for accurately predicting absolute solubility. These models are typically trained on specific data sets with defined molecular descriptors (values for the studied components are provided in Table S.9), which makes predictions less reliable when descriptor values fall near the edge, or outside, the model's applicability domain. In some cases, such as Chiguard-234 and Cyasorb UV3638-F, predictions failed, probably due to issues with molecular fragmentation and the inability to calculate descriptors for their fused-ring systems. Regarding terephthalic acid, a PET monomer produced from hydrolysis, which sublimates at approximately 402 °C, values from the ASPEN Plus database differed from QSPR predictions by 114.4 °C for T_m , and 38.1 kJ/mol for ΔH_{fus} .

Another potential source of deviation in COSMO-RS estimations is related to the treatment of ΔC_p . In the present work, the approximation $\Delta C_p \approx 0$ was adopted, which discards the last term in eq 12. This simplification is commonly applied near the melting point, but the solubility measurements considered here were performed at 25 °C, which sits substantially below T_m for most compounds studied. Consequently, the validity of this approximation decreases as the gap between the melting point and the solubility temperature grows.^{66–68} The absolute magnitude of the discarded term is set by the solute alone, scaling with ΔS_{fus} , and becomes larger for compounds with high melting points and large fusion enthalpies. Its impact on predicted solubility, however, depends on the balance between ΔG_{fus} and ΔG_{mixing} in eq 8. In strongly favorable solvents, where ΔG_{mixing} is very negative and may dominate $\Delta G_{\text{dissolution}}$, the

dominant source of error is the COSMO-RS estimate of solute–solvent interactions. However, when ΔG_{fus} represents a substantial fraction of $\Delta G_{\text{dissolution}}$, ignoring the ΔC_p contribution may shift predicted solubilities appreciably. As illustrated in Figure 7, deviations from experimental solubility increase with higher melting points and fusion enthalpies, highlighting limitations of the simplified ΔG_{fus} expression, particularly when T_m greatly exceeds the solubility temperature.⁶⁹ For compounds with $\Delta H_{\text{fus}} > 40$ kJ/mol and $T_m > 150$ °C, the accuracy of thermal data therefore becomes particularly important for reliable room-temperature solubility predictions.

Based on the present data set, the workflow performs reliably ($\text{RMSE} \leq 1$ log unit) for compounds with molecular weight below ~600 g/mol, fewer than ~10 rotatable bonds, predicted solubility above ~100 ppm, and T_m within ~150 K of the solubility temperature. Outside these ranges, deviations grow systematically along four directions: (i) at very low predicted solubilities (<100 ppm), where the COSMO-RS SLE algorithm approaches numerical limits and weak long-range solute–solvent interactions gain weight; (ii) for bulky, flexible molecules ($\text{MW} > 600$ g/mol, $\text{NRB} > 10$), where the gas-phase minimum-energy conformer is no longer sufficient and broader conformer sampling with explicit treatment of intramolecular hydrogen bonding becomes necessary; (iii) for high-melting compounds, where the $\Delta C_p \approx 0$ simplification of ΔG_{fus} contributes additional deviation; and (iv) in the thermal inputs themselves (T_m , ΔH_{fus}), which are independent of COSMO-RS and are governed by the auxiliary fusion equations. These regions are best understood as targets for future improvement of the underlying thermodynamic model rather than as hard operational limits of the present workflow.

5. CONCLUSION

This study demonstrates the effectiveness of a combined computational and experimental framework for predicting the solubility of plastic additives and PET monomers in depolymerization solvents. Accurate solubility prediction is crucial for guiding efficient separation and purification processes, particularly in light of the challenge of removing diverse additives and contaminants. Solubility measurements ranged from less than 1 ppm in water to 2.2 wt % and up to 41.6 wt % in ethylene glycol, providing a benchmark for assessing quantum chemistry-based predictions. Systematic conformer sampling emerged as a key factor in reliable solubility predictions, with molecular flexibility, reflected in the number and diversity of conformers, strongly influencing results. Considering a broad set of conformers, combined with outlier detection to remove unrealistic geometries, ensures inclusion of relevant structures, particularly for complex and high-molecular-weight additives such as antioxidants. The basis set choice in quantum chemical optimization was also critical: 6-311++G(d,p) offered the best balance between accuracy

and computational cost, while 3–21+G(d) minimized computational expense, and 6–311++G(3df,3pd) provided only marginal gains in selected cases. COSMO-RS thermodynamic modeling performed well for most compounds, especially in ethylene glycol, but predictions were less reliable for low-solubility, large, or structurally complex molecules.

For high-melting, stable crystals with large Gibbs free energies of fusion, incorporating experimental thermal properties, including melting points and fusion enthalpies measured by DSC, is recommended over QSPR estimates. In addition, accounting for heat capacity differences (ΔC_p) may further improve solid–liquid equilibrium predictions. Overall, integrating conformer sampling, quantum chemical methods, thermodynamic modeling, and targeted experiments provides an efficient and accurate framework to predict the order of magnitude solubility of additives in depolymerization media, thereby supporting the design of energy-efficient downstream processes that advance more environmentally sustainable chemical recycling of plastics.

A ± 1 log unit error in solubility corresponds to approximately an order-of-magnitude uncertainty in equilibrium concentration. In process terms, this implies a similar uncertainty in the solvent required to dissolve a given mass of solute, or in the amount of material predicted to remain dissolved during crystallization. The practical impact depends on the operating regime. Quantifying these effects for specific process configurations will require dedicated process-simulation studies.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsomega.6c04278>.

detailed materials and methods, additional computational outputs from RDKit, Gaussian, COSMOtherm, and COSMOquick, as well as lists of abbreviations and symbols (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Steven De Meester – *Laboratory for Circular Process Engineering (LCPE), Department of Green Chemistry and Technology, Ghent University, Kortrijk 8500, Belgium;*
orcid.org/0000-0002-5246-3918;
Email: Steven.DeMeester@UGent.be

Authors

Amin Shariatmadar Tehrani – *Laboratory for Circular Process Engineering (LCPE), Department of Green Chemistry and Technology, Ghent University, Kortrijk 8500, Belgium;* orcid.org/0000-0002-2333-5300

Ruben Goeminne – *Center for Molecular Modeling (CMM), Department of Applied Physics, Ghent University, Zwijnaarde 9052, Belgium;* orcid.org/0000-0001-8920-3362

Pieter Cnudde – *Center for Molecular Modeling (CMM), Department of Applied Physics, Ghent University, Zwijnaarde 9052, Belgium;* orcid.org/0000-0002-6735-0078

Veronique Van Speybroeck – *Center for Molecular Modeling (CMM), Department of Applied Physics, Ghent University, Zwijnaarde 9052, Belgium;* orcid.org/0000-0003-2206-178X

Kathy Elst – *Sustainable Polymer Technologies (SPOT) Team, Flemish Institute for Technological Research (VITO), Mol 2400, Belgium;* orcid.org/0000-0002-9873-6379

Karolien Vanbroekhoven – *Sustainable Polymer Technologies (SPOT) Team, Flemish Institute for Technological Research (VITO), Mol 2400, Belgium*

Elias Feghali – *Sustainable Polymer Technologies (SPOT) Team, Flemish Institute for Technological Research (VITO), Mol 2400, Belgium; Chemical Engineering Program, Notre Dame University-Louaize, Zouk Mosbeh 1211, Lebanon;* orcid.org/0000-0003-1561-3259

Kevin M. Van Geem – *Laboratory for Chemical Technology (LCT), Department of Materials, Textiles and Chemical Engineering, Ghent University, Zwijnaarde 9052, Belgium;* orcid.org/0000-0003-4191-4960

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acsomega.6c04278>

Author Contributions

A.S.T.: Conceptualization, Methodology, Software, Data curation, Formal analysis, Investigation, Writing—Original Draft, Visualization. R.G.: Methodology, Software, Validation. P.C.: Validation, Writing—Review and Editing. V.V.S.: Supervision, Writing—Review and Editing. K.E.: Writing—Review and Editing. K.V.: Writing—Review and Editing. E.F.: Writing—Review and Editing. K.M.V.G.: Writing—Review and Editing. S.D.M.: Conceptualization, Supervision, Project administration, Funding acquisition, Writing—Review and Editing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support from the Flemish Government and Flanders Innovation & Entrepreneurship (VLAIO) through the Moonshot project (CYCLOPS HBC.2021.0584). The computational resources (Stevin Supercomputer Infrastructure) and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by Ghent University, FWO, and the Flemish Government – Department EWI.

■ REFERENCES

- (1) Cao, F.; Wang, L.; Zheng, R.; Guo, L.; Chen, Y.; Qian, X. Research and progress of chemical depolymerization of waste PET and high-value application of its depolymerization products. *RSC Adv.* **2022**, *12* (49), 31564–31576.
- (2) Barnard, E.; Arias, J. J. R.; Thielemans, W. Chemolytic depolymerisation of PET: a review. *Green Chem.* **2021**, *23* (11), 3765–3789.
- (3) McGill, C.; Shaw, K.; Knauer, K.; Seeley, M.; Lynch, J. Plastic additives in the ocean: Use of a comprehensive dataset for meta-analysis and method development. *Chemosphere* **2024**, *358*, 142172.
- (4) Xin, J.; Zhang, Q.; Huang, J.; Huang, R.; Jaffery, Q. Z.; Yan, D.; et al. Progress in the catalytic glycolysis of polyethylene terephthalate. *J. Environ. Manage.* **2021**, *296*, 113267.
- (5) Godiya, C. B.; Gabrielli, S.; Materazzi, S.; Pianesi, M. S.; Stefanini, N.; Marcantoni, E. Depolymerization of waste poly(methyl methacrylate) scraps and purification of depolymerized products. *J. Environ. Manage.* **2019**, *231*, 1012–1020.
- (6) McNeeley, A.; Liu, Y. A. Assessment of PET Depolymerization Processes for Circular Economy. 1. Thermodynamics, Chemistry,

- Purification, and Process Design. *Ind. Eng. Chem. Res.* **2024**, 63 (8), 3355–3399.
- (7) Demirel, Y. Thermodynamic Analysis of Separation Systems. *Sep. Sci. Technol.* **2004**, 39 (16), 3897–3942.
- (8) Abrams, D. S.; Prausnitz, J. M. Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *AIChE J.* **1975**, 21 (1), 116–128.
- (9) Richard, L.; Helgeson, H. C. Calculation of the thermodynamic properties at elevated temperatures and pressures of saturated and aromatic high molecular weight solid and liquid hydrocarbons in kerogen, bitumen, petroleum, and other organic matter of biogeochemical interest. *Geochim. Cosmochim. Acta* **1998**, 62 (23), 3591–3636.
- (10) Gillet, G.; Vitrac, O.; Desobry, S. Prediction of Partition Coefficients of Plastic Additives between Packaging Materials and Food Simulants. *Ind. Eng. Chem. Res.* **2010**, 49 (16), 7263–7280.
- (11) Loschen, C.; Klamt, A. Prediction of Solubilities and Partition Coefficients in Polymers Using COSMO-RS. *Ind. Eng. Chem. Res.* **2014**, 53 (28), 11478–11487.
- (12) Zhou, P.; Yu, J.; Sánchez-Rivera, K. L.; Huber, G. W.; Lehn, R. C. V. Large-scale computational polymer solubility predictions and applications to dissolution-based plastic recycling. *Green Chem.* **2023**, 25 (11), 4402–4414.
- (13) Renon, H.; Prausnitz, J. M. Estimation of Parameters for the NRTL Equation for Excess Gibbs Energies of Strongly Nonideal Liquid Mixtures. *Ind. Eng. Chem. Process Des. Dev.* **1969**, 8 (3), 413–419.
- (14) Pye, C. C.; Ziegler, T.; van Lenthe, E.; Louwen, J. N. An implementation of the conductor-like screening model of solvation within the Amsterdam density functional package—Part II. COSMO for real solvents. *Can. J. Chem.* **2009**, 87 (7), 790–797.
- (15) Esteban, J.; Vorholt, A. J.; Leitner, W. An overview of the biphasic dehydration of sugars to 5-hydroxymethylfurfural and furfural: a rational selection of solvents using COSMO-RS and selection guides. *Green Chem.* **2020**, 22 (7), 2097–2128.
- (16) Klamt, A. The COSMO and COSMO-RS solvation models. *WIREs Comput. Mol. Sci.* **2011**, 1 (5), 699–709.
- (17) Mullins, E.; Oldland, R.; Liu, Y. A.; Wang, S.; Sandler, S. I.; Chen, C.-C.; et al. Sigma-Profile Database for Using COSMO-Based Thermodynamic Methods. *Ind. Eng. Chem. Res.* **2006**, 45 (12), 4389–4415.
- (18) Cordova, I. W.; Teixeira, G.; Ribeiro-Claro, P. J. A.; Abranches, D. O.; Pinho, S. P.; Ferreira, O.; et al. Using Molecular Conformers in COSMO-RS to Predict Drug Solubility in Mixed Solvents. *Ind. Eng. Chem. Res.* **2024**, 63 (21), 9565–9575.
- (19) Cavin, A. T.; Hillisch, A.; Uellendahl, F.; Schnecken, S.; Göller, A. H. Reliable and Performant Identification of Low-Energy Conformers in the Gas Phase and Water. *J. Chem. Inf. Model.* **2018**, 58 (5), 1005–1020.
- (20) Crippen, G. M.; Havel, T. F. *Distance Geometry and Molecular Conformation*; Research Studies Press, 1988.
- (21) Landrum, G.; Tosco, P.; Kelley, B.; Ric, C.; D, S., et al. *rdkit/rdkit: 2023_09_3 (Q3 2023) Release*, 2023. doi: <https://doi.org/10.26434/chemrxiv-2023-09-3>
- (22) RDKit: open-source cheminformatics. <https://www.rdkit.org>.
- (23) Mu, T.; Rarey, J.; Gmehling, J. Performance of COSMO-RS with Sigma Profiles from Different Model Chemistries. *Ind. Eng. Chem. Res.* **2007**, 46 (20), 6612–6629.
- (24) Klajmon, M. Purely Predicting the Pharmaceutical Solubility: What to Expect from PC-SAFT and COSMO-RS? *Mol. Pharm.* **2022**, 19 (11), 4212–4232.
- (25) Hyttinen, N.; Prisle, N. L. Improving Solubility and Activity Estimates of Multifunctional Atmospheric Organics by Selecting Conformers in COSMOtherm. *J. Phys. Chem. A* **2020**, 124 (23), 4801–4812.
- (26) Riniker, S.; Landrum, G. A. Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. *J. Chem. Inf. Model.* **2015**, 55 (12), 2562–2574.
- (27) Ebejer, J.-P.; Morris, G. M.; Deane, C. M. Freely Available Conformer Generation Methods: How Good Are They? *J. Chem. Inf. Model.* **2012**, 52 (5), 1146–1158.
- (28) Friedrich, N.-O.; de Bruyn Kops, C.; Flachsenberg, F.; Sommer, K.; Rarey, M.; Kirchmair, J. Benchmarking Commercial Conformer Ensemble Generators. *J. Chem. Inf. Model.* **2017**, 57 (11), 2719–2728.
- (29) Kim, H.; Jang, C.; Yadav, D. K.; Kim, M.-H. The comparison of automated clustering algorithms for resampling representative conformer ensembles with RMSD matrix. *J. Cheminf.* **2017**, 9 (1), 21.
- (30) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, 12 (85), 2825–2830.
- (31) MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*; University of California Press, 1967; pp 281–298.
- (32) Chen, L.; Roe, D. R.; Kochert, M.; Simmerling, C.; Miranda-Quintana, R. A. K.-M. N. An Improved Clustering Algorithm for Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2024**, 20 (13), 5583–5597.
- (33) Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, 20, 53–65.
- (34) Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat.* **1974**, 3 (1), 1–27.
- (35) Davies, D. L.; Bouldin, D. W. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **1979**, PAMI-1224–227.
- (36) Arbelaitz, O.; Gurrutxaga, I.; Muguerza, J.; Pérez, J. M.; Perona, I. An extensive comparative study of cluster validity indices. *Pattern Recognit.* **2013**, 46 (1), 243–256.
- (37) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, 32 (7), 1456–1465.
- (38) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. Efficient diffuse function-augmented basis sets for anion calculations. III. The 3–21+G basis set for first-row elements, Li–F. *J. Comput. Chem.* **1983**, 4 (3), 294–301.
- (39) Andersson, M. P.; Uvdal, P. New Scale Factors for Harmonic Vibrational Frequencies Using the B3LYP Density Functional Method with the Triple- ζ Basis Set 6–311+G(d,p). *J. Phys. Chem. A* **2005**, 109 (12), 2937–2941.
- (40) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *J. Phys. Chem.* **1995**, 99 (7), 2224–2235.
- (41) Klamt, A.; Jonas, V.; Bürger, T.; Lohrenz, J. C. W. Refinement and Parametrization of COSMO-RS. *J. Phys. Chem. A* **1998**, 102 (26), 5074–5085.
- (42) Eckert, F.; Klamt, A. Fast solvent screening via quantum chemistry: COSMO-RS approach. *AIChE J.* **2002**, 48 (2), 369–385.
- (43) Skyner, R. E.; McDonagh, J. L.; Groom, C. R.; Mourikvan, T.; Mitchell, J. B. O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* **2015**, 17 (9), 6174–6191.
- (44) Müller, S.; Nevolianis, T.; Garcia-Ratés, M.; Riplinger, C.; Leonhard, K.; Smirnova, I. Predicting solvation free energies for neutral molecules in any solvent with openCOSMO-RS. *Fluid Phase Equilib.* **2025**, 589, 114250.
- (45) Klamt, A.; Diedenhofen, M. Calculation of Solvation Free Energies with DCOSMO-RS. *J. Phys. Chem. A* **2015**, 119 (21), 5439–5445.
- (46) Kolář, M.; Fanfrlík, J.; Lepšík, M.; Forti, F.; Luque, F. J.; Hobza, P. Assessing the Accuracy and Performance of Implicit Solvent Models for Drug Molecules: Conformational Ensemble Approaches. *J. Phys. Chem. B* **2013**, 117 (19), 5950–5962.
- (47) Yan, F.; Zhang, Y.; Niu, H.; Feng, X.; Xiong, J.; Jia, Q.; et al. QSPR models for enthalpy and entropy of organic compounds based on a set of norm indices. *Fluid Phase Equilib.* **2023**, 573, 113869.

- (48) Loschen, C.; Klamt, A. New Developments in Prediction of Solid-State Solubility and Cocrystallization Using COSMO-RS Theory. In *Computational Pharmaceutical Solid State Chemistry*; Abramov, Y. A., ed.; Wiley, 2016; pp. 211–233.
- (49) Wyttenbach, N.; Niederquell, A.; Kuentz, M. Machine Estimation of Drug Melting Properties and Influence on Solubility Prediction. *Mol. Pharm.* **2020**, *17* (7), 2660–2671.
- (50) Hughes, L. D.; Palmer, D. S.; Nigsch, F.; Mitchell, J. B. O. Why Are Some Properties More Difficult To Predict than Others? A Study of QSPR Models of Solubility, Melting Point, and Log P. *J. Chem. Inf. Model.* **2008**, *48* (1), 220–232.
- (51) Song, Z.; Wang, J.; Sundmacher, K. Evaluation of COSMO-RS for solid–liquid equilibria prediction of binary eutectic solvent systems. *Green Energy Environ.* **2021**, *6* (3), 371–379.
- (52) Iglewicz, B.; Hoaglin, D. C. *Volume 16: How To Detect And Handle Outliers*; Quality Press, 1993.
- (53) Center for Biologics Evaluation and Research Center for Drug Evaluation and Research Q2(R1) Validation of Analytical Procedures: Text and Methodology Guidance for Industry. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/q2r1-validation-analytical-procedures-text-and-methodology-guidance-industry>.
- (54) Klamt, A.; Eckert, F.; Hornig, M.; Beck, M. E.; Bürger, T. Prediction of aqueous solubility of drugs and pesticides with COSMO-RS. *J. Comput. Chem.* **2002**, *23* (2), 275–281.
- (55) Schröder, B.; Santos, L. M. N. B. F.; Marrucho, I. M.; Coutinho, J. A. P. Prediction of aqueous solubilities of solid carboxylic acids with COSMO-RS. *Fluid Phase Equilib.* **2010**, *289* (2), 140–147.
- (56) Verma, N. R.; Gopal, G.; Anantharaj, R.; Banerjee, T. (Solid + liquid) equilibria predictions of ionic liquid containing systems using COSMO-RS. *J. Chem. Thermodyn.* **2012**, *48*, 246–253.
- (57) Mullins, E.; Liu, Y. A.; Ghaderi, A.; Fast, S. D. Sigma Profile Database for Predicting Solid Solubility in Pure and Mixed Solvent Mixtures for Organic Pharmacological Compounds with COSMO-Based Thermodynamic Methods. *Ind. Eng. Chem. Res.* **2008**, *47* (5), 1707–1725.
- (58) Pung, A.; Leito, I. Predicting Relative Stability of Conformers in Solution with COSMO-RS. *J. Phys. Chem. A* **2017**, *121* (36), 6823–6829.
- (59) Vauthier, E. C.; Cossé-Barbi, A.; Blain, M.; Fliszár, S. Low-level DFT calculations II: molecules containing fluorine and second-row atoms. *J. Mol. Struct.: THEOCHEM* **1999**, *492* (1), 113–121.
- (60) Bolton, E. E.; Kim, S.; Bryant, S. H. PubChem3D: Conformer generation. *J. Cheminf.* **2011**, *3* (1), 4.
- (61) Mohan, M.; Keasling, D.; Simmons, B. A.; Singh, B. S. In silico COSMO-RS predictive screening of ionic liquids for the dissolution of plastic. *Green Chem.* **2022**, *24* (10), 4140–4152.
- (62) Israelachvili, J. N. *Intermolecular and Surface Forces*; Academic Press, 2011.
- (63) Klamt, A. Prediction of the mutual solubilities of hydrocarbons and water with COSMO-RS. *Fluid Phase Equilib.* **2003**, *206* (1), 223–235.
- (64) Klamt, A. COSMO-RS for aqueous solvation and interfaces. *Fluid Phase Equilib.* **2016**, *407*, 152–158.
- (65) AlSaleem, S. S.; Zahid, W. M.; AlNashef, I. M.; Hadj-Kali, M. K. Solubility of Halogenated Hydrocarbons in Hydrophobic Ionic Liquids: Experimental Study and COSMO-RS Prediction. *J. Chem. Eng. Data* **2015**, *60* (10), 2926–2936.
- (66) Nordström, F. L.; Rasmuson, C. Determination of the activity of a molecular solute in saturated solution. *J. Chem. Thermodyn.* **2008**, *40* (12), 1684–1692.
- (67) Tsivintzelis, I.; Economou, I. G.; Kontogeorgis, G. M. Modeling the solid–liquid equilibrium in pharmaceutical-solvent mixtures: Systems with complex hydrogen bonding behavior. *AIChE J.* **2009**, *55* (3), 756–770.
- (68) Nordström, F. L.; Rasmuson, C. Prediction of solubility curves and melting properties of organic and pharmaceutical compounds. *Eur. J. Pharm. Sci.* **2009**, *36* (2), 330–344.
- (69) Neau, S. H.; Bhandarkar, S. V.; Hellmuth, E. W. Differential Molar Heat Capacities to Test Ideal Solubility Estimations. *Pharm. Res.* **1997**, *14* (5), 601–605.