Chemical Science

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: L. Wittmann, C. E. Selzer and S. Grimme, *Chem. Sci.*, 2025, DOI: 10.1039/D5SC06406F.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/chemical-science

View Article Online
DOI: 10.1039/D5SC06406F

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

A Diverse and Chemically Relevant Solvation Model Benchmark Set with Flexible Molecules and Conformer Ensembles†

Lukas Wittmann,[a‡] Christian Erik Selzer,[a‡] and Stefan Grimme*[a]

Received Date
Accepted Date

DOI: 00.0000/xxxxxxxxxx

We introduce *FlexiSol* – a flexible solvation benchmark set with molecule ensembles. *FlexiSol* is the first of its kind to combine structurally and functionally complex, highly flexible solutes with exhaustive conformational sampling for systematic testing of solvation models. The dataset contains 824 experimental solvation energy and partition ratio data points (1551 unique molecule-solvent pairs) at standard-state conditions, focusing on drug-like, medium-to-large flexible molecules (up to 141 atoms), with over 25000 theoretical conformer/tautomer geometries across all phases. The set is publicly available and data points were selected in order to have minimal overlap with existing sets. Using this benchmark, we evaluate a broad spectrum of popular implicit solvation approaches, including physics-based (quantum-chemical and semiempirical) and data-driven models. We find that partition ratios are generally computed more accurately compared to solvation energies, likely due to partial error cancellation, yet most models still systematically underestimate strongly stabilizing interactions while overestimating weaker ones in both solvation energies and partition ratios. Additionally, we investigate the impact of three key ingredients: conformational ensemble, geometry choice (phase-specific *vs.* single-phase), and underlying electronic energy method. We find that full Boltzmann-weighted ensembles or just the lowest-energy conformers yield very similar accuracy – still both require conformational sampling – whereas random single-conformer selection degrades performance, especially for larger and flexible systems. Geometry relaxation and the level of electronic structure theory both influence results; however, the magnitude and sometimes direction of these effects can vary by method, as fortuitous error cancellation sometimes masks underlying deficiencies present in the models. As a complement to existing data sets, *FlexiSol* will enable more systematic development and evaluation of solvation models.

1 Introduction

Everyday observations like dissolving sugar in coffee or the low solubility of creatine in water[1] illustrate the ubiquitous nature of solvation – which refers to any stabilizing interaction between a solute and its solvent[2] and affects all almost all areas of chemistry, biology and materials science.[3,4] Over the past decades, theoretical chemistry has become an indispensable partner to experiment, offering both conceptual frameworks for interpreting chemical phenomena and predictive tools that guide new investigations.[5,6] Especially in the field of Green Chemistry,[7] theory helps to identify and optimize reaction pathways, predict solvent and catalyst effects *in silico*, and guide the design of sustainable processes with reduced waste and energy use.[8–10] Besides sustainability aspects, the behavior of emerging environmental pollutants, such as per- and polyfluorinated alkyl substances (PFAS),[11–13] can be modeled under climate change, predicting their compartmental distribution, degradation, and bioavailability.[14–17] A major challenge, however, is the scarcity of experimental data,[18,19] making predictive calculations not just valuable but essential for assessing their behavior in solution – a need recognized in guidelines for more efficient risk assessment by the Organisation for Economic Co-operation and Development (OECD).[12,20–22] Beyond macroscopic equilibria, solvation also shapes molecular properties measured by spectroscopic methods. Solvent-induced shifts in infra-red (IR) absorption and nuclear magnetic resonance

[a] *Mulliken Center for Theoretical Chemistry, Beringstraße 4, D-53115 Bonn, Germany*
* *grimme@uni-bonn.de*

† Supplementary Information available: The Supporting Information includes a PDF with all supplementary figures (systematic-error trends, solvent-specific performance), extended tables, and additional discussion. Further, the full set of conformer geometries and all computed energies; experimental reference values with full bibliographic citations; and computed solvation energies and partition ratios for each model. See DOI: 00.0000/00000000.

‡ These authors contributed equally to this work

View Article Online
DOI: 10.1039/D5SC06406F

(NMR) chemical shifts[23–26] can be large. This often results from a significant, non-negligible change in molecular geometry upon solvation,[27–29] that can change the preferred tautomeric or conformational state, *e.g.*, amino acids are neutral in the gas phase but zwitterionic in water, and capsaicin adopts an open form in aqueous solution versus a folded form in methanol.[30–32]

Such descriptions can be obtained using three main approaches to model solute-solvent interactions. The most direct approach is explicit solvation, where solvent molecules are included in the calculation. Commonly using free-energy perturbation[33,34] or thermodynamic integration[35,36] with dynamic simulations, these methods require exhaustive sampling of the configurational space and are thus computationally demanding, typically only feasible with classical methods like force fields. Together with integral-equation theories like RISM and its extensions (EC-RISM, RISM-SCF-cSED),[37,38] they are often grouped as statistical methods. The so-called implicit approach can be seen as a simplification in which the solvent is approximated as some form of continuum, drastically improving the computational cost and enhancing applicability.[39,40] Most implicit solvation models are based on a quantum-mechanical (QM) treatment of the solute, while the continuum is often treated classically.[41] The Poisson (or Poisson-Boltzmann) equation solved around a molecular cavity is often used as the basis in so called polarizable continuum models (PCM).[42] More sophisticated solvation models describe additional solvent-solute interactions, *e.g.*, hydrogen bonding, dispersion interactions and cavitation energy. Popular approaches are the usage of a surface-area-dependent term,[43] or using empirical statistical mechanical frameworks.[44,45] Hybrid strategies, such as cluster-continuum (*microsolvation*) approaches, combine a few explicit solvent molecules with an implicit bulk description.[46–48] While often successful, they are difficult to apply, computationally demanding, and usually require expert intervention,[49–51] although several automated microsolvation workflows are available.[52,53] The third and most empirical approach is the descriptor-based approach. These models rely on correlations between known descriptors and target properties, replacing explicit physical components with mappings learned from large experimental or quantum-mechanical databases. These ones are, *e.g.*, quantitative structure-activity and structure-property relationship (QSAR, QSPR) models like UNIFAC[54] and OPERA.[55] These approaches are complemented by machine-learning (ML) models. Recent ML models include QM-GNNIS for solvation energies,[56] QupKake for pK_a,[57] MF-LOGP for octanol–water partition ratios,[58] and FASTSOLV for solubility prediction.[59]

Generally, with increasing model empiricism, more data is required for parameterization and testing. This means that ML-based approaches, and even empirical QM models, depend on particularly extensive datasets. However, not only the quantity of the data is of importance – equally the quality and diversity, since models trained on bad or incomplete datasets can inherit these shortcomings.[60] For solution phase properties (*e.g.*, solvation energies) there is no practicable *ab-initio* method or protocol to obtain accurate, theoretical reference values – unlike for general electronic structure theory.[61–63] Because of this, experimental values have to be taken as reference. Popular data sets include the Minnesota Solvation Database (MNSOL)[64], which contains around 3000 data points spanning 92 solvents with roughly 800 unique molecules, or the FreeSolv database,[65] containing around 650 molecules in aqueous solution, however, only 250 of those are not already contained in MNSOL.[66] There are other databases that only contain the name or atom connectivity of the molecules (*e.g.*, via SMILES), but lacking geometric information required by QC methods. These are, for example, the SOLV@TUM database or the large collection of acid dissociation constants of IUPAC.[67,68]

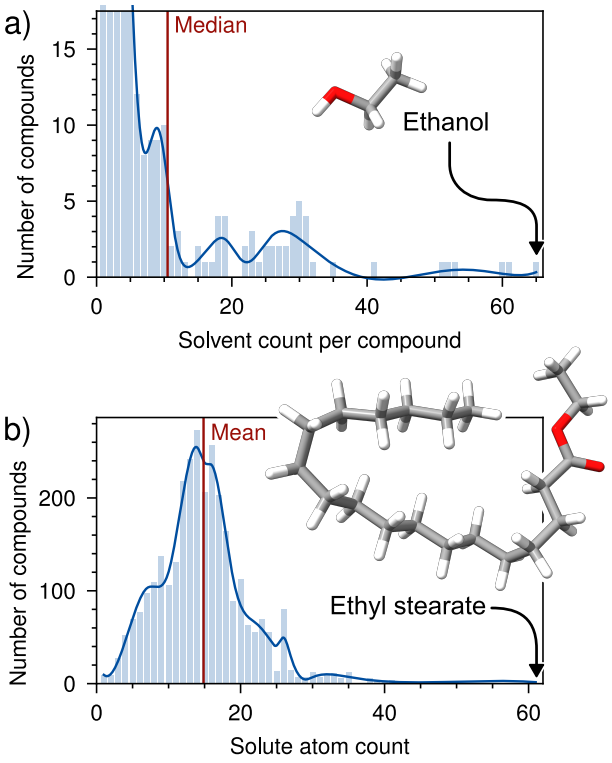

Fig. 1 Distributions of solvent occurrence per solute and solute sizes in the MNSOL dataset. (a) Histogram of the number of unique solvents for each unique solute, with the red line marking the median. Ethanol (shown) is the most common solute, available in 65 different solvents (right tail). (b) Histogram of the number of atoms per solute molecule, with the red line indicating the mean. Ethyl stearate (shown) is the largest solute with 62 atoms.

Despite their great utility, the present sets have limitations. They predominantly feature small molecules, lacking diversity in larger, more complex molecular structures and motifs (see Fig. 1b), and often contain lower amounts of unique molecules compared to their number of data points. For example, half of the around 3000 data points in MNSOL originate from just 54 unique molecules in a large number of different solvents (Fig. 1a), making it chemically relatively homogeneous. Additionally, the present QM-ready sets most often only provide a single gas phase structure per unique molecule (*e.g.*, MNSOL). Without proper minimum-energy geometries for each phase, the solvent-induced geometric and conformational changes are not accounted for and can thus introduce systematic biases. Moreover, as many models are parameterized on these databases, the limited availability of

View Article Online
DOI: 10.1039/D5SC06406F

independent sets leaves only a small fraction of data for testing, which may reduce the ability to fully assess model robustness or identify potential overfitting.[69,70]

To address these limitations, we have compiled a diverse data set of 824 experimental solvation energies and partition ratios (1551 unique molecule-solvent pairs) for drug-like, medium-to-large flexible molecules (mainly in the 30-80 atom range, up to 141), including their conformational ensembles with over 25000 conformers and tautomers in total. The complete benchmark, along with all coordinates, computed energies, and reference values, is freely available for use and extension. This set not only extends beyond the size and structural diversity of existing collections, but also allows us to systematically dissect three critical factors in their influence on model accuracy: conformational sampling, geometry choice (phase-specific versus single-phase), and underlying electronic structure method. By doing so, we aim to provide a rigorous foundation for testing modern methods and guiding the development of more robust, efficient solvation-prediction protocols.

2 Background

To place our work in context, we briefly review the theoretical foundations of solvation modeling, the involved quantities, the common approach and introduce the most popular solvation models, especially focusing on the implicit approach.

2.1 Thermodynamic Quantities

Gibbs energies are the direct bridge between molecular calculations and real-world observables like solubilities, partition ratios, and equilibrium constants, which dictate concentrations, yields, and distribution of compounds under realistic conditions. For a given experimental equilibrium reaction, the Gibbs–Helmholtz equation

$$\Delta G = -RT \ln K, \tag{1}$$

states the relationship between the equilibrium constant K and the difference in Gibbs energy ΔG for the respective process at temperature T, with R being the ideal gas constant. In principle, any solvation-dependent experimental quantity can be used to evaluate a model's performance, however, solvation energies and partition ratios depend predominantly on the solvation description, making them especially suitable for testing theoretical approaches. The partition ratio describes the equilibrium of a substance A between two phases α and β

$$A_{(\alpha)} \overset{K_{\alpha/\beta}}{\rightleftharpoons} A_{(\beta)},$$

which can be expressed as the ratio of concentrations in the respective phases as

$$K = \frac{[A]^{(\alpha)}}{[A]^{(\beta)}}. \tag{2}$$

The phase transfer Gibbs energy $\Delta_{\mathrm{tr}}G$ can be obtained by the difference of Gibbs energies of the compound in each phase.

$$\left(\log K_{\alpha/\beta}\right)_A = \frac{\Delta_{\mathrm{tr}} G_A^{\alpha\to\beta}}{RT \ln 10} \quad \text{with } \Delta_{\mathrm{tr}} G_A^{\alpha\to\beta} = G_A^{(\beta)} - G_A^{(\alpha)} \tag{3}$$

Partition ratios (often also called partition constant) give the partitioning of a substance between two immiscible solution phases, whereas Henry's Law constants (HLC) denotes the partitioning of a substance between a solution phase and the gas phase.[71] Both, the partition ratios and HLCs can be used to obtain the respective Gibbs energy difference. In this work, we will describe the air-solvent partitioning via their respective solvation energy ($\Delta_{\mathrm{solv}}G$) and the solvent-solvent partitioning via the partition ratio ($\log K_{\alpha/\beta}$). Although it is recommended to include temperature and solvent in the notation, we will simplify this since only standard conditions (298.15 K and 1 atm) at infinite dilution are used.

Solvent-solvent partition ratios are most commonly measured by the shake-flask method, in which a solute is equilibrated between two immiscible solvents and concentrations in each phase are determined – often by UV-Vis spectroscopy or ^{1}H-NMR.[22,72] For highly hydrophobic compounds that form can stable emulsions, slow-stirring techniques can offer improved reliability.[73] Additional measurement approaches such as reversed-phase HPLC and generator-column methods have been employed to expand coverage across diverse solute-solvent systems.[74,75] Henry's law constants, are often obtained by either static equilibrium experiments, where gas and liquid phase concentrations are measured in closed cells once equilibrium is reached, or by dynamic *gas-purge* approaches,[76,77] in which a purge gas or bubble column achieves thermodynamic equilibrium and the decay of the gas phase solute concentration is monitored over time.[78,79] Static methods provide direct equilibrium measurements, while dynamic methods can be more suitable for volatile or low-solubility compounds. The accuracy of such methods is discussed in more detail in Sec. 4.1.

2.2 Ensemble Averaged Gibbs Energies

To calculate Henry's Law constants or partition ratios (via Eqs. 1 and 3), the Gibbs energy of the respective substance of interest in the respective phases α and β has to be known. The Gibbs energy of a unique structure i of substance A in a solvent is obtained from

$$G_{A,i}^{(\alpha)} = E_{\mathrm{el},A,i} + G_{\mathrm{trv},A,i}^{(\alpha)} + \Delta_{\mathrm{solv}} G_{A,i}^{(\alpha)}, \tag{4}$$

where $E_{\mathrm{el},A,i}$ is the gas phase total electronic energy as obtained by, *e.g.*, by Density Functional Theory (DFT). The ro-vibrational Gibbs energy contribution at finite temperature $G_{\mathrm{trv},A,i}$, which includes temperature-dependent entropic as well as enthalpic terms, both arising from vibrational, rotational, and translational contributions.[80] Due to the associated computational cost and the fact that the change of the ro-vibrational Gibbs energy contribution is generally absorbed into the solvation model itself or rather in its parametrization (see Sec. 2.3 and Ref. 43,81), we will not compute this term explicitly. Finally, a solvation energy component $\Delta_{\mathrm{solv}} G_{A,i}^{(\alpha)}$ is needed, which will be detailed in Sec. 2.3.

In addition to these main contributions, non-rigid molecules often have multiple relevant conformers contributing to the total Gibbs energy of a given chemical species.[82] This firstly requires the exploration of the respective potential energy surface for finding relevant conformers, and secondly, the calculation of all afore-

View Article Online
DOI: 10.1039/D5SC06406F

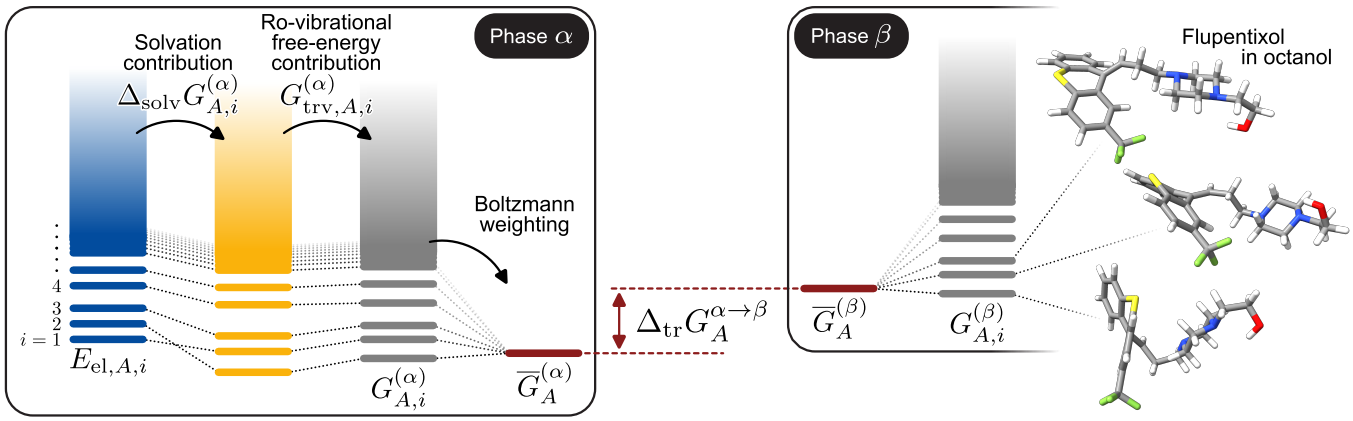

Fig. 2 Scheme illustrating how conformational ensembles are used to compute the solute transfer Gibbs energy between two phases. All conformers (and tautomers) within a defined energy window are considered, and their per-conformer Gibbs energies $G_{A,i}$ are Boltzmann weighted to yield the ensemble average Gibbs energy $\overline{G}_A$ for each phase α and β. The phase-transfer Gibbs energy $\Delta_{\mathrm{tr}}G^{\alpha\to\beta}_A$ is obtained as the difference between these ensemble averages. Shown is the example of three conformers of flupentixol in octanol.

mentioned contributions for each conformer. To obtain the Gibbs energy of the substance, the Gibbs energies of all ensemble members i need to be Boltzmann-weighted via

$$G_A \equiv \overline{G}_A = \sum_i^N p_{A,i} G_{A,i}, \quad \text{with } p_{A,i} = \frac{\exp\left(-\beta G_{A,i}\right)}{\sum_i^N \exp\left(-\beta G_{A,i}\right)}, \tag{5}$$

where $\beta = (k_B T)^{-1}$ with k_B being the Boltzmann constant. This is shown schematically in Fig. 2. The ensemble-averaged Gibbs energy $\overline{G}_A$ of A will be denoted as G_A for the rest of this work.

Depending on the functional groups of a molecule, multiple tautomeric states are possible.[83] They are included by extending the Boltzmann sum to cover all relevant tautomers, treating them as additional ensemble members i together with their associated G_i.

In principle, the Gibbs energy of a conformer ensemble additionally includes a conformational Gibbs energy part G_{conv} that stems from the conformational entropy ($-TS_{\mathrm{conv}}$), as a result of mixing multiple populated conformers.[84] For the gas phase, this is already a challenging problem due to the required very extensive exploration of the potential energy surface. For the solution phase, the limitations of the implicit solvation model make accurate determination of the conformational entropy computationally not feasible.[84,85] Therefore, we omit determining these contributions in our study.

2.3 Accounting for Solvation Effects

To compute the Gibbs energy of each conformer, the solvation energy $\Delta_{\mathrm{solv}}G$ is needed, which refers to the change in Gibbs energy when an ion or molecule is transferred from a vacuum (or the gas phase) to a solvent α.[86] One common decomposition of contributions in the implicit-solvation framework is given in Eq. 6.

$$\Delta_{\mathrm{solv}}G^{(\alpha)}_A = G^{(\alpha)}_{\mathrm{ES}} + G^{(\alpha)}_{\mathrm{NES}} + G^{(\alpha)}_{\mathrm{N}} + G^{\circ\to *}_{\mathrm{corr}} \tag{6}$$

Here, G_{ES} is the electrostatic term, G_{NES} the non-electrostatic term, G_{N} the geometry-relaxation contribution, and $G^{\circ\to *}_{\mathrm{corr}}$ is the standard-state correction term. In the following paragraphs, these components are explained in more detail.

The electrostatic contribution G_{ES} (often referred to as polarization) captures the stabilization arising the mutual polarization of the solute's electronic density and the surrounding dielectric continuum. In implicit models, this term effectively represents the electrostatic solute-solvent interactions that would be described explicitly in a molecular simulation by Coulomb interactions. G_{ES} is obtained by solving the Poisson or Poisson-Boltzmann equation using methods such as IEF-PCM,[87] COSMO or CPCM.[88,89]

The non-electrostatic term G_{NES} accounts for, *e.g.*, cavity formation around the solute, solvent restructuring, and London dispersion forces. Especially in less polar solvents, an electrostatics-only description is not sufficient, making the inclusion of non-electrostatic effects necessary. Popular models include the surface-area-based SM*x* family of solvation methods, with SM12 or SMD,[43,90] or the variants of the Miertus-Scrocco-Tomasi (MST) approach.[42,91–94] Another popular method is COSMO-RS,[44,45,95] which uses an empirical statistical-thermodynamic framework. Noteworthy is also the less empirical composite method for implicit representation of solvent CMIRS,[96–99] and newer approaches, like the spherically averaged liquid susceptibility ansatz SaLSA[100] or charge-asymmetric non-locally determined local-electric solvation model CANDLE.[101] The growing interest in larger structures has also led to the need for more efficient solvation models. To this category, the ALPB model,[81] CPCM-X,[102] and the easy solvation estimation ESE models[103] belong to. Current implicit models do not explicitly take into account the entropy penalty upon solvation, resulting from the loss of translational, rotational, and often also low-frequency vibrational degrees of freedom, while rigorous explicit approaches (*e.g.*, via MD) are prohibitively expensive. This omission is in fact why most often a separate thermostatistical correction for solution phase properties is left out: the entropy change is already fitted into the empirical parameterization of the solvation models, although several schemes have been proposed in the literature that would enable an approximate explicit calculation.[84,104,105]

View Article Online
DOI: 10.1039/D5SC06406F

The geometry (or nuclear relaxation) contribution is independent of the solvation model itself, but depends on the used procedure. This contribution is given by

$$G_{\mathrm{N}}^{(\alpha)} = E_{\mathrm{el}}^{(\alpha)} - E_{\mathrm{el}}^{(\mathrm{gas})}, \tag{7}$$

and describes the change in electronic energy when a molecule distorts from its gas phase to its solution equilibrium geometry. A common, but crude, approximation in many solvation models or approaches[43,81] is setting $G_{\mathrm{N}}^{(\alpha)} = 0$ as it simplifies the calculation procedure (*cf.* 4.2.2). This constitutes a systematic approximation whose associated error is typically absorbed or compensated for during the parameterization of the respective solvation model against experimental references.

To obtain the standard-state corrected solvation energy $\Delta_{\mathrm{solv}}G^{\circ}$, denoted as $\Delta_{\mathrm{solv}}G$ from now on, a standard state correction factor $G_{\mathrm{corr}}^{\circ\to *}$ is needed.[106] It accounts for the standard state mismatch between an ideal gas at 1 atm (gas phase) and an ideal solution at 1 mol·L^{-1} (solution phase), ensuring both phases are referenced to 1 mol·L^{-1} concentration. It is given by

$$G_{\mathrm{corr}}^{\circ\to *} = RT \ln\left(\frac{RTc^{\ominus}}{p^{\ominus}}\right), \tag{8}$$

where $c^{\ominus} = 1\,\mathrm{mol{\cdot}L^{-1}}$ is the standard concentration, and $p^{\ominus} = 1\,\mathrm{atm}$ is standard pressure and amounts to 1.89 kcal·mol^{-1} assuming ideal gas behavior.[106]

Unlike physics-based continuum models, ML approaches bypass the decomposition into physical contributions and instead directly predict the overall solvation energy (*i.e.*, both, $G_{\mathrm{ES}}^{(\alpha)}$ and $G_{\mathrm{NES}}^{(\alpha)}$) from molecular representations.

3 Methods

In this section, the workflow for the generation of the conformer ensembles, the computational methods, the tested solvation models and the data curation is outlined.

3.1 Ensemble Generation

The procedure described in the following serves to generate phase-specific conformational ensembles and is shown in Fig. 3. The computational details are detailed in Sec. 3.3. Initial geometries were created by converting respective SMILES identifiers to three dimensional geometries.[107] These were subsequently optimized using an efficient tight-binding method. Tautomeric states were automatically screened using alternating protonation and deprotonation cycles, as implemented in the CREST program package.[108,109] The obtained ensemble of tautomers was subsequently cleaned using MolBar,[110] removing not only redundant conformers but also artifacts generated during tautomerization, which often correspond to entirely different compounds. In a second clean-up step, high-lying (*i.e.*, $> 12.0\,\mathrm{kcal{\cdot}mol^{-1}}$) tautomers are sorted out using a low-cost DFT method (PBE-D4/def2-SV(P)). For each of the remaining tautomers, a conformer search was carried out using GOAT[111] (energy window of 6.0 kcal·mol^{-1}). The last step of the conformer generation process is the screening and subsequent optimization, where the ensemble is further refined by ensemble geometry optimization as

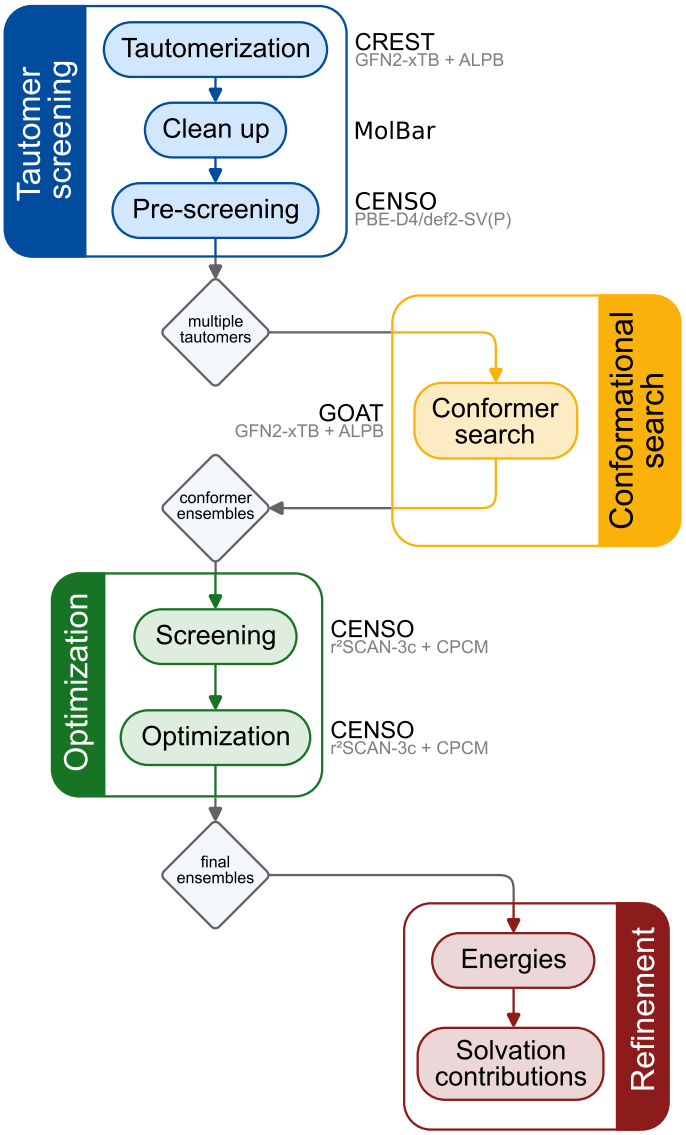

Fig. 3 Illustration of the workflow used to generate the structure ensemble for a given substance in a given phase. The used software/tool (black) and level of theory (gray) is given as text next to the individual steps.

implemented in CENSO[82] using ORCA (final energy window of 4.0 kcal·mol^{-1}).[112]

For the conformer searches and geometry optimizations in solution, the ALPB solvation model[81] was employed for xTB/CREST calculations, while for any DFT calculations, the CPCM model[89,113] was used. We also tested the more sophisticated SMD model. However, owing to possible convergence difficulties[114] and our observation that geometries from CPCM and SMD differ only marginally (see Supplemental Material, Sec. C.1), we employ CPCM throughout this work.

3.2 Solvation Models

We grouped the models into three categories according to their underlying electronic structure method: a) quantum mechanical models requiring DFT (or similar level of QM), b) semiempirical models based on efficient approximate electronic structure meth-

View Article Online
DOI: 10.1039/D5SC06406F

ods, and c) machine-learning models predicting $\Delta_{\text{solv}}G$, directly from molecular representations. Table 1 summarizes the tested methods, their included computed contributions, and required molecule input type. A more detailed description can be found in the Supplemental Material, Sec. A.1. In the main manuscript, we present only the 10 most widely used and representative solvation models. The results for all other methods can be found in Supplemental Material, Sec. E and F. All $\Delta_{\text{solv}}G$ and $\log K_{\alpha/\beta}$ values correspond to standard-state conditions (298.15 K, 1 atm) at infinite dilution.

For the self-consistently calculated solvation energies (*i.e.*, with CPCM, SMD, D-COSMO-RS, ALPB and ddCOSMO), the solvation energy is obtained via

$$\Delta_{\text{solv}}G^{(\alpha)}_{A,i} = G^{(\alpha)}_{A,i} - E_{\text{el},A,i}, \tag{9}$$

where the $G^{(\alpha)}_{A,i}$ is the Gibbs energy of the conformer i of molecule A in phase α and $E_{\text{el},A,i}$ is the respective gas phase energy evaluated on the same geometry.

Only solvation models capable of predicting $\Delta_{\text{solv}}G$ for all solvents in our benchmark were included; models limited to specific solvents (*e.g.*, water-only models) were excluded. Although we intended to evaluate additional ML- and QSAR-based tools, many proved to be property-specific (*e.g.*, QupKake for pK_a, IFSQSAR[128] and Vega[129] for other physicochemical properties), lacked publicly accessible or maintained code, or were insufficiently documented to support a reproducible integration.

3.3 Computational Details

Quantum chemical calculations were performed with xTB 6.6.1, Turbomole 7.7.1,[130–132] and ORCA 6.0.1.[112,133,134] If not stated otherwise, default settings were applied for all calculations. LibXC was used for some of the functionals.[135] All quantum chemical calculations use matching def/def2 effective small core potentials (ECPs) for heavy elements with Z > 36.[136,137] For ORCA related calculations, matching general-purpose auxiliary basis sets are constructed on the fly using *AutoAux*[138] and the *RI-JCOSX*[139–141] approximation was used. Turbomole calculations were done using RI-J.[142,143]

For the workflows, Open Babel 3.1.0,[144] CREST 3.0.2,[109,145] GOAT,[111] smi2xyz,[107] a development version of CENSO 2.0,[82] and MolBar 1.1.0 were used.[110] The tautomerization protocol,[108] as well as the conformational sampling was done using GFN2-xTB with ALPB.[81,146] For pre-screening, PBE-D4/def2-SV(P)[147–151] is used. Final geometries were obtained using the r^{2}SCAN-3c composite method,[152,153] combined with CPCM for any non-gas phase structures.[89] Calculations with the new general extended tight-binding method g-xTB use the publicly available development version 1.0.0.[127] Hybrid DFT calculations utilize ωB97M-V[154] with the def2-TZVPPD (aTZ) basis with ORCA.[136,151]

COSMO-RS 16.01 is calculated using Turbomole 7.7.1 with COSMOtherm C30-1601.[155,156] COSMO-RS uses per default BP86/def-TZVP.[157–159] D-COSMO-RS solvation contributions were calculated via Turbomole 7.7.1 using BP86/def-TZVP. Solvation contributions of the CPCM and SMD[43] models are obtained using r^{2}SCAN-3c, and openCOSMO-RS[115] uses (per default) BP86/def2-TZVPD in ORCA. Solvation contributions of the GBSA, ALPB,[81] and CPCM-X[102] models were calculated using GFN2-xTB within xTB.[160] CPCM-X was calcualted in xTB 6.7.1 due to availability. ddCOSMO was calculated using ddX[161] combined with GFN2-xTB within tblite. Solvation contributions of uESE 1.2,[103,162], ESE-PM7 1.2, and Solv 1.0 (with 500/150 grid setting for electrostatics and non-electrostatics, respectively),[121] are calculated using Mulliken charges obtained with g-xTB. Solvation contributions from the ESE-GB-DNN model (version: September 2023)[163] and the ESE-EE-DNN model (version: June 2024)[125] were obtained using their respective published implementations. Solvation contributions of DirectML 0.0.3 and CIGIN (version: August 2020) were obtained using the SMILES strings of the solvent and solute.[123]

3.4 Data Curation

Curating a truly novel solvation test set required painstaking effort to identify high-quality experimental data that had not already been used by MNSol, FreeSolv, or SOLV@TUM. We used the following criteria to select suitable candidates: First, we focus only on the elements HCNO, F, Cl, Br, I, S, P, and Si. On the one hand, some of the tested solvation models are only parametrized for these elements, and on the other hand, the available number of experimental data points for, *e.g.*, (transition) metals and heavy elements is very small. Secondly, the focus was set on drug-like, medium- to large-sized (30-80 atoms) molecules that contain, for example, heteroatom-rich scaffolds, zwitterionic moieties, macrocycles, and halogens. The inclusion of molecules that contain networks of hydrogen-bond donors and acceptors, and nonstandard ring sizes further increase the complexity and diversity. We prioritized compounds of biological, pharmaceutical, or environmental importance, such as per-and polyfluoroalkyl substances, active pharmaceutical ingredients, and natural products to guarantee that *FlexiSol* addresses a highly relevant chemical space. Thirdly, we deliberately included conformationally flexible molecules or molecules with multiple protonation/tautomeric states, to investigate the importance of including the corresponding chemical space. We limited our set to solvents with abundant, high-quality data and broad practical relevance: primarily water and 1-octanol, which together account for the majority of published solvation energies and partition ratios for medium-to-large organic molecules. Sources were curated through an extensive literature study, where the primary literature was consulted when possible. Care was taken to ensure that every data point was explicitly labeled as experimental. All experimental solvation energies in this work are given at standard-state conditions (298.15 K, 1 atm) in the molar reference state, meaning that $\Delta_{\text{solv}}G$ is the Gibbs energy for transferring a solute molecule from the ideal gas phase at 1 mol·L^{-1} into an ideal solution at the same solute concentration.

3.5 Analysis

For the analysis of the mean or average errors, a 3σ criterion is applied. Hence, values (now considered an outlier) of each

View Article Online
DOI: 10.1039/D5SC06406F

Table 1 Overview of all tested solvation models, showing the required input (3D coordinates or SMILES) and explicitly computed energy terms (electrostatic, ES; non-electrostatic, NES), with brief descriptions and references. Details can be found in the Supplemental Material, Sec. A.1.

Type	Model	Input	ES	NES	Description	Ref.
Quantum mechanical (QM)	CPCM	3D	✓	✗	Conductor-like PCM (electrostatics only)	89
	SMD	3D	✓	✓	SM*x* model with empirical cavity-dispersion (CDS) term	43
	openCOSMO-RS	3D	✓	✓	Open-source variant of COSMO-RS	115
	COSMO-RS	3D	✓	✓	COSMO-based screening model for real solvents	44,45
	D-COSMO-RS[1]	3D	✓	✓	Direct, self-consistent variation of COSMO-RS	116–118
Semiempirical (SQM)	ALPB	3D	✓	✓	Analytically linearized Poisson-Boltzmann model	81
	ddCOSMO	3D	✓	✗	Domain-decomposition formulation of COSMO/CPCM	102,119,120
	CPCM-X	3D	✓	✓	ddCOSMO with post-processing (similar to COSMO-RS & SMD)	102
	Solv[2]	3D	✓	✓	Non-iterative COSMO-like electrostatics with NES	121
	uESE[2]	3D	✓	✓	Non-iterative COSMO-like electrostatics with NES	103
	ESE-PM7[2]	3D	✓	✓	Cheaper, PM7 charges-based variant of uESE	122
Machine Learning (ML)	DirectML	SMILES		—	Directed message-passing neural network (NN)	123
	CIGIN	SMILES		—	Chemically interpretable graph interaction network	124
	ESE-EE-DNN	3D	✗	✗	Dense NN with empirical charges, electrostatics, and NES	125
	ESE-GB-DNN	3D	✗	✗	Simplified electrostatics version of ESE-EE-DNN	126

[1] D-COSMO-RS results do not include the datapoints in Hexadecane due to no available parameterization.
[2] Input charges for Solv, uESE, and ESE-PM7 are computed using g-xTB.[127]

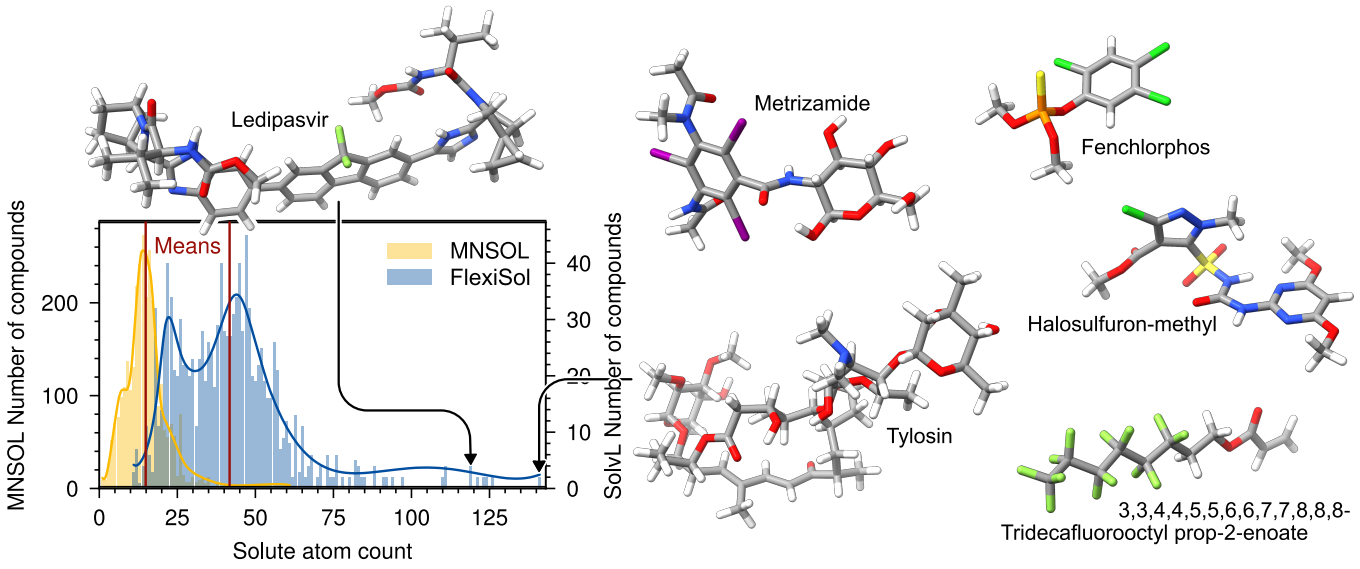

Fig. 4 Histogram of the solute molecule sizes in the *FlexiSol* and MNSOL datasets, with vertical lines indicating the respective means. The largest compound of *FlexiSol*, tylosin (141 atoms), and five other representative molecules from the set are shown.

model that deviate more than three times the standard deviation from the mean are removed from the statistics. This is done to mitigate the impact of outliers that would otherwise dominate the statistics and make the interpretation difficult. Cases where models (or their underlying method) failed to converge or delivered obviously faulty results were also removed from the statistics of the respective model. In the final set, this occurred only for openCOSMO-RS and exclusively for three specific conformers of three different compounds, detailed in Supplemental Material, Sec. C.7.

The error of all respective methods will be analyzed using the mean absolute error (MAE), standard deviation (SD) and root mean squared error (RMSE). For definitions of all used statistical measures in this work, see Supplemental Material, Sec. B. The solvation energies will be discussed in kcal·mol^{-1}, while the partition ratios are given in log units, where one log unit is equivalent to 1.36 kcal·mol^{-1} according to Eq. 1.

To discuss the influence of different factors on the computed results (Sec. 4.2.1-4.2.3), we will use relative errors, *i.e.*, the *change* in root mean square error (RMSE) relative to a baseline approach that will be introduced in Sec. 4.2. This is done by varying one variable at a time (*ceteris paribus*). An improvement in error is

View Article Online
DOI: 10.1039/D5SC06406F

given as a negative change in RMSE, while a worsening is indicated by a positive change. The RMSE is used as a measure as it concludes the statement by MAE and SD. For the discussion of the relative errors, the 3σ criterion is not applied in order to avoid inconsistencies.

Due to the large range of experimental reference values in *FlexiSol*, absolute errors alone can be misleading when comparing methods across the entire set. Therefore, relative errors with respect to the experimental reference (in %) are additionally reported in the Supplemental Material, Sec. C.5. We also performed an analysis in which deviations that are smaller than the reported experimental uncertainties were counted as zero error – explicitly taking into account the experimental uncertainty. These can be found in Supplemental Material, Sec. C.4. Both the relative error analysis and the explicit treatment of the experimental uncertainties yielded trends essentially identical to those from the absolute errors. As these additional analyses did not alter the overall conclusions, we proceed with the main analysis based solely on absolute errors.

4 Results and Discussion

4.1 Database

Absolute solvation energy data (or the respective Henry's law constants) was taken from Refs. 18,68,123,164–173. Partition ratio reference data was collected from Refs. 18,45,123,167,172,174–179 The Supplemental Material provides a detailed list of reference values and their sources (Supplemental Material, Sec. G), together with the curated experimental data, optimized geometries, and raw energies for every structure and method.

FlexiSol contains 824 experimental values: 530 solvation energies and 294 partition ratios, covering 734 unique molecules in 10 solvents (1551 molecule-solvent pairs). Including all conformer/tautomer ensembles, the set contains over 25000 geometries. A chord diagram of solvent distribution is shown in Fig. 5. Molecules range from 11 to 141 atoms (mean 42) with up to 25 rotatable bonds (Supplemental Material, Fig. D1a); for comparison, MNSOL averages 15 atoms. Of the heavy atoms, 30% are non-carbon, mainly nitrogen or oxygen (21%) and halogens (8%), with the remainder being sulfur, phosphorus, and silicon (2%). The solvation energies cover a range of 32.5 kcal·mol^{-1} (from -27.7 to 4.9 kcal·mol^{-1}), with a mean absolute reference energy of 11.1 kcal·mol^{-1} (MNSOL, for comparison, has 4.5 kcal·mol^{-1}). The partition ratios cover a range of 14.2 log units (from -4.3 to 9.9 log units), with a mean absolute reference of 3.3 log units.

Because most often no experimental uncertainty is stated for the data points used, we will take an educated guess for the uncertainty. MNSOL states an experimental uncertainty of around 0.2 kcal·mol^{-1},[64] and FreeSolv state, where available, an uncertainty per data point, ranging from 0.0 to 1.9 kcal·mol^{-1}, with a default value of around 0.5 kcal·mol^{-1} if no literature value is given.[65] This is in good agreement with the literature, where often an uncertainty of around 0.3 to 0.5 kcal·mol^{-1} is stated.[180–182] For partition ratios, experimental repeatability is around ± 0.3 log units according to OECD,[183] with inter-method differences of up to ± 0.5 log units.[22] However, as the experimental determination of larger (and strongly polar or nonpolar) molecules is more difficult, *e.g.*, due to poor solubility.[184–187] For the mentioned reasons, we estimate the uncertainty for solvation energies to be around ± 0.6 kcal·mol^{-1} and for the partition ratios to be around ± 0.5 log units.

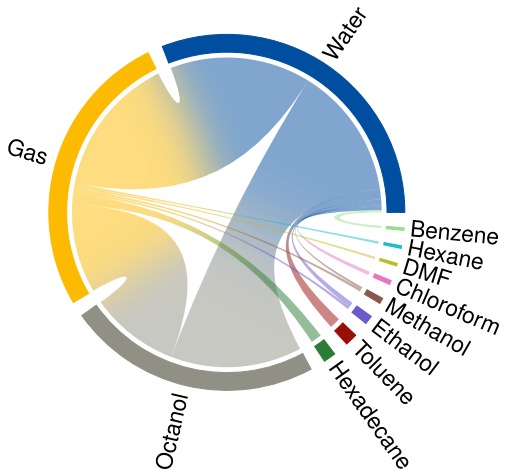

Fig. 5 Chord diagram showing solvent frequency and the composition of data points in *FlexiSol*. Gas-solvent ribbons represent solvation energies ($\Delta_{\text{solv}}G$) and solvent-solvent ribbons represent partition ratios ($\log K_{\alpha/\beta}$). The arc size reflects how often each solvent appears in the dataset, while ribbon thickness indicates the number of data points for each combination.

While this benchmark set was constructed to reflect chemically relevant and challenging solutes, some inherent biases and limitations remain: i.), the majority of reference data refers to water and octanol, which dominate both the solvation and partition ratios (see Fig. 5). ii.), the chemical space is skewed toward molecules composed mainly of C, H, N, O, halogens, and to a lesser extent S, P, and Si. iii.) the set contains only neutral solutes. While there are works presenting partition coefficients[188] or solvation energies[64,66,189–191] for ionic substances that could in principle be used as reference values, the accuracy of the values is uncertain due to the larger experimental errors associated. iv.) the focus on drug-like, medium-sized molecules means that small molecules and inorganic/organometallic systems are not present. v.) the set only contains data points at standard conditions. vi.) the set contains only data for infinite dilution only and thus no concentration/activity effects. Accordingly, our conclusions are restricted to neutral, flexible and polyfunctional drug-like organic molecules in mainly water or 1-octanol at standard conditions.

4.2 Benchmark Results

In this section, we assess how well computed values of the models agree with the experimental references on the *FlexiSol* set. We start with our *baseline* approach – phase-specific conformational ensembles, Boltzmann weighting, and r^{2}SCAN-3c electronic method – and use this as the reference point for all further comparisons. This approach has proven robustness with moderate cost in many prior works.[82,192,193] From there, we examine the impact of i.), ensemble sampling, Sec. 4.2.1, ii.), geometry choice, Sec. 4.2.2, and iii.), the electronic structure method

View Article Online
DOI: 10.1039/D5SC06406F

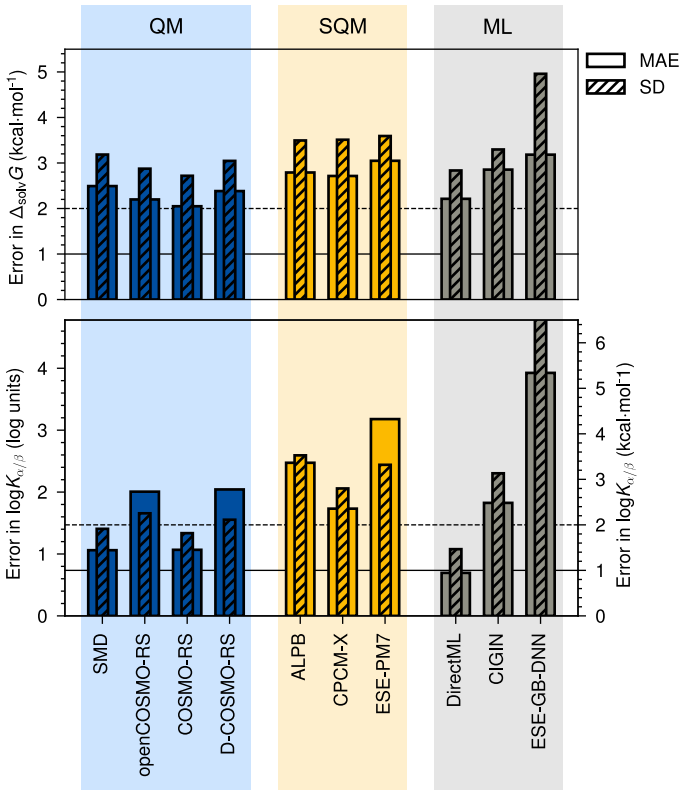

Fig. 6 Mean absolute error (solid bars) and standard deviation (hatched bars) are shown for solvation energies (top, $\Delta_{solv}G$ in kcal·mol^{-1}) and partition ratios (bottom, log $K_{\alpha/\beta}$ in log units, with the secondary right axis in kcal·mol^{-1}). Results are given for the most popular solvation models, additional models can be found in the Supplemental Material.

for the description of the nuclear relaxation, Sec. 4.2.3. In this subsection, we start with a detailed investigation of the baseline approach. The mean absolute error (MAE) and standard deviation (SD) for the whole set is shown in Fig. 6 for each solvation method.

We find that the computed solvation energies show larger errors compared to experiment on our benchmark set compared to those reported in prior studies, as seen on, *e.g.*, MNSOL. We attribute this to the generally larger, more challenging solutes in our set, the fact that many models (*e.g.*, SMD) were parameterized on MNSOL, and *FlexiSol*'s substantially wider range of $\Delta_{solv}G$ and log $K_{\alpha/\beta}$ values. For solvation energies, we find the DFT-based methods to deliver overall the best results, with COSMO-RS leading the category with a mean absolute error of 2.0 kcal·mol^{-1}. The SQM-based models yield a slightly worse result with CPCM-X being the best performer with an MAE of 2.7 kcal·mol^{-1}. The machine-learning models yield results of similar accuracy compared to the DFT-based models, demonstrated by DirectML's MAE of 2.2 kcal·mol^{-1}, but generally show a greater difference between the individual models.

Calculated partition ratios generally agree better with experiment compared to solvation energies, likely because solvent-independent, substance-specific errors partially cancel out in the ratio. An example is decachlorobiphenyl, where the computed free solvation energies in octanol and water both show larger errors (error of QM-based methods around 4.0 kcal·mol^{-1}) compared to the respective partition ratio (error of 1.0 log units).[194] For the DFT-based methods, we find SMD and COSMO-RS to both provide a mean absolute error of 1.0 log units, followed by the SQM-based models with CPCM-X and an MAE of 1.8 log units. The ML-based models provide a more diverse picture, with ESE-GB-DNN showing the worst result and DirectML providing overall the best performance with an MAE of 0.7 log units.

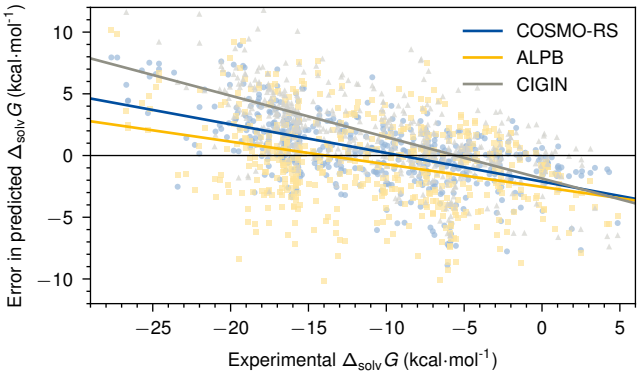

Fig. 7 Error in computed solvation energy ($\Delta_{solv}G$) versus experimental $\Delta_{solv}G$ for three representative models: COSMO-RS (QM, blue circles), ALPB (SQM, yellow squares), and CIGIN (ML, gray triangles). Solid lines show linear regression lines of the errors, indicating each model's systematic underestimation of large (positive) and overestimation of small (negative) values. All values are given in kcal·mol^{-1}.

Most models are found to suffer from a systematic error in their computed solvation energies, overestimating small ones, and underestimating larger ones, as shown in Fig. 7 (see Supplemental Material, Sec. D.2.1 for all methods). This trend is very similar for most methods with a slope between −0.1 and −0.2 kcal·mol^{-1} per kcal·mol^{-1} of reference $\Delta_{solv}G$. Partition ratios show the same systematic error: The models underestimate the affinity for the favored phase, *i.e.*, negative errors for very positive partition ratios and positive errors for very negative partition ratios (Supplemental Material, Sec. D.2.2). Systematic errors can also be found for some specific functional groups or structural motifs (Supplemental Material, Sec. D.3). Systems containing primary and secondary heteroatoms are systematically overestimated in terms of their solvation energy, whereas the opposite holds for tertiary heteroatoms. For heteroatom-heteroatom bonds, partition ratios are found to be generally underestimated, favoring the polar phase. This highlights the difficulty in describing strong solvent-solute interactions like hydrogen-bonding – often a major difficulty for implicit solvation models.[39,41] Because our dataset is dominated by water and octanol, and contains far fewer data in other solvents, broad conclusions about solvent-dependence are inherently limited. Nonetheless, we observe that the MAE values are worse for octanol compared to water, by about 0.3 kcal·mol^{-1} for most methods, with SMD and CIGIN both worsening most by over 0.6 kcal·mol^{-1}. In the next paragraphs, the three model categories will be discussed in more detail.

Among the DFT-based models (SMD, COSMO-RS and openCOSMO-RS), absolute solvation energies are calculated best by COSMO-RS (MAE 2.0 kcal·mol^{-1}), closely followed by its open

View Article Online
DOI: 10.1039/D5SC06406F

source variant openCOSMO-RS (MAE 2.2 kcal·mol^{-1}) and SMD (MAE 2.5 kcal·mol^{-1}). For partition ratios, COSMO-RS and SMD perform very similarly (MAE 1.0 log units), with openCOSMO-RS showing a larger systematic shift (*i.e.*, SD<RMSE) with an mean error of 1.8 log units, thus yielding a worse agreement (MAE 2.0 log units) with experiment. One of the most overestimated solvation energies (6.0 kcal·mol^{-1} on average) is found for polyhydroxy compounds, more specifically sugar-type alcohols like sorbitol, adonitol, mannitol, or galacticol. Especially COSMO-RS overestimates the mentioned values by more than 10.0 kcal·mol^{-1}, SMD, for example, only by 4.5 kcal·mol^{-1}. SMD, however, struggles more with solvation energies of very lipophilic substances, underestimating molecules such as decafluorobiphenyl by 7.0 kcal·mol^{-1}. Organic-aqueous partition ratios for heteroatom-dominant (especially nitrogen-containing) are underestimated by both COSMO-RS and SMD, *e.g.*, for substances, like cytidine diphosphate (COSMO-RS, error of −5.6 log units) and azimsulfuron (SMD, error of −5.4 log units). Such systematic errors have been noted already in, *e.g.*, Refs. 195–198

The SQM-based approaches deliver robust results, with MAEs of around 2.7 kcal·mol^{-1} for $\Delta_{\text{solv}}G$ and 1.8 log units for $\log K_{\alpha/\beta}$, with the best performer being CPCM-X. Most of the SQM-based methods are found to struggle with halogen-dominated substances like chlorothalonil (error of −13.0 kcal·mol^{-1}) or very oxygen and nitrogen rich interactions in polar solvents like benzo-18-crown-6 in water (error of 10.0 kcal·mol^{-1}) on *FlexiSol*. For partition ratios, the SQM-based models seem to overestimate the affinity of the organic (octanol) phase over the aqueous one for large, polycyclic and heteroatom-rich substances like tacrolimus (with an average error of 10.0 log units). ALPB underestimates the octanol/water partitioning for polyfluorinated substances like perfluorooctanesulfonic acid (with an error of −9.0 log units).

The ML approaches deliver overall good results. Besides the low mean errors, however, the number of outliers according to the 3σ criterion, which assumes a Gaussian error distribution, is generally slightly larger compared to the other model classes. As ML models do not always follow a normal distribution of errors,[199] the application of this criterion results in the removal of a larger number of outliers (around 5 more compared to the QM/SQM-based models, see Supplemental Material, Sec. F). CIGIN and DirectML are found to show larger errors (*i.e.*, ±10 kcal·mol^{-1}) for the solvation energies of cyclic heteroatom-containing xenobiotics, like flumioxazin, milbemycin A3, or trimethoprim. The excellent performance for DirectML cannot be matched by any other tested ML-based model.

4.2.1 Ensemble

Because the sampling of the conformational space to obtain the ensembles is computationally demanding (*cf.* Sec. 4.3), we analyze the influence of the conformational ensemble on the results by testing two simplified approaches: (a), using only the lowest-energy conformer for each phase, thereby avoiding higher-energy conformers and requiring only one final single-point and solvation model evaluation, and (b), a single random conformer in each phase, simulating the absence of conformational sampling by mimicking the outcome if one initial geometry is used for the optimization in both phases. Both approaches are illustrated schematically in Fig. 8, and their impact relative to the baseline is shown in Tab. 2.

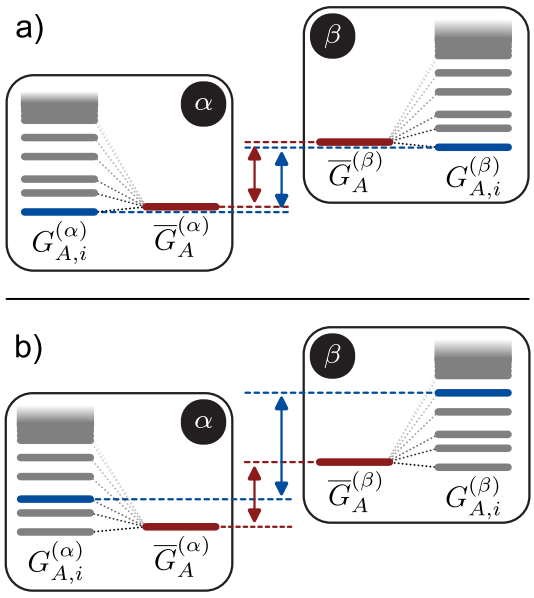

Fig. 8 Schematic of the two approximate ways to calculate solvation energies with respect to the selection of the used conformers. The baseline approach uses all conformers (and tautomers) within the chosen energy window. Per-conformer Gibbs energies $G_{A,i}$ are Boltzmann-weighted to yield the ensemble average for each phase (shown in red, and in Fig. 2). (a) Lowest-energy conformer: Only the single lowest-energy structure is used in both phases. (b) Random conformer: A single conformer i_{rand} is chosen at random and used for the ΔG. This aims to simulate the error introduced by the absence of conformational sampling.

We find that using only the lowest-energy conformer instead of the full Boltzmann-weighted ensemble yields overall very similar results for both solvation energies and partition ratios. The largest deviations are about ±0.7 kcal·mol^{-1} for solvation energies, as seen for example in simvastatin (octanol, improves) or 15-crown-5 (water, worsens). In cases where there is a noticeable change, many conformers lie close in energy, so their combined Boltzmann weights differ significantly from that of the single lowest conformer. For instance, simvastatin has 41 conformers in the gas phase within 4.0 kcal·mol^{-1}, five of which have Boltzmann population above 5%, while the lowest conformer accounts for only 17%. Since the typical model errors on the full set (MAE ∼ 2.0 kcal·mol^{-1}) exceed the magnitude of the changes, such improvements are within statistical error of the models and likely reflect error-compensation effects. Therefore, the use of a single conformer in scenarios with many conformers that are energetically close together can lead to small errors (∼ 0.7 kcal·mol^{-1}), while the results remain unchanged without (almost) degenerate conformers.

For the random conformer approach, the accuracy deteriorates notably – on average by more than 0.2 kcal·mol^{-1} for solvation energies and 0.4 log units for partition ratios – compared to the baseline. It is important to emphasize that our test still favors lower-energy conformers: The random conformer was drawn

View Article Online
DOI: 10.1039/D5SC06406F

Table 2 Change in RMSE (relative to the baseline) for solvation energies ($\mathrm{kcal\cdot mol^{-1}}$) and partition ratios (log units) upon using different approaches to obtain the final Gibbs energy in each phase. A negative change in RMSE indicates an improvement, a positive change a worsening. *Low.* denotes using only the lowest-energy conformer per phase; *Rand.* denotes using a single random conformer per phase.

Method	$\Delta_{\mathrm{solv}}G$ ($\mathrm{kcal\cdot mol^{-1}}$)		$\log K_{\alpha/\beta}$ (log units)	
	Low.	Rand.	Low.	Rand.
SMD	0.0	0.3	0.0	0.3
openCOSMO-RS	−0.1	0.2	0.0	0.2
COSMO-RS	0.0	0.3	0.0	0.4
D-COSMO-RS	−0.1	0.2	0.0	0.4
ALPB	0.0	0.3	0.0	0.4
CPCM-X	0.0	0.2	0.0	0.5
ESE-PM7	−0.1	0.1	0.0	0.3
DirectML	0.0	0.3	0.0	1.1
CIGIN	0.0	0.7	0.0	0.5
ESE-GB-DNN	0.0	0.4	0.0	0.2

from the already optimized ensemble, which only contains conformers within the defined 4.0 $\mathrm{kcal\cdot mol^{-1}}$ window. That also restricts the random conformer selection to that window, and means that the degradation observed here is a lower bound to "real world" calculation without conformer sampling. In a true one-geometry workflow, for example, optimizing a single arbitrary starting structure (*e.g.*, output of a 1D → 3D conversion) in each phase without prior conformer screening, larger deviations can potentially occur, particularly for large and flexible molecules or systems with multiple low-energy conformers. These findings reinforce that a tautomer and conformational analysis step is recommended for accurate solvation and partition calculations, regardless of the underlying solvation model. In a recent study, lysergic acid diethylamide (LSD) was shown to exhibit several tautomeric forms differing by up to −12 $\mathrm{kcal\cdot mol^{-1}}$ relative to the commonly depicted tautomer (*e.g.*, the SMILES-derived form); only a comprehensive QM tautomer search identified the dominant species and thus yielded accurate solvation free energies.[200]

4.2.2 Geometry

In addition to examining conformational effects, we explore how nuclear relaxation influences solvation energy computations and their agreement with experiment. Before discussing the respective approaches and their influence on the computed values, we will investigate the geometry change and general magnitude of the nuclear relaxation energy on the *FlexiSol* set when transferring a solute from gas to solution phase.

Firstly, we generally investigate the structural changes associated with bringing a solute in solution. On our set, we find final structure root mean square deviations (RMSD) between the solution and gas phase geometry from 0.0 to 4.9 Å, using the lowest-energy conformer in each phase as ranked by r^{2}SCAN-3c with the SMD solvation model. This RMSD linearly correlates with the number of atoms and the number of rotatable bonds in the molecule (Supplemental Material, Fig. D1a and b). The nuclear relaxation contribution associated with the change in geometry upon solvation is on average around 0.6 $\mathrm{kcal\cdot mol^{-1}}$ on *FlexiSol*. Of those nuclear relaxation contributions, around 300 are larger than 1.0 $\mathrm{kcal\cdot mol^{-1}}$ and around 120 larger than 2.0 $\mathrm{kcal\cdot mol^{-1}}$ (Supplemental Material, Fig. D1c). One example of a very high nuclear relaxation contribution with 13.0 $\mathrm{kcal\cdot mol^{-1}}$ is cytidine diphosphate (CDP), shown in Fig. 9. This large contribution results from the opening of intramolecular hydrogen bonds. This means that, for flexible and very functionalized molecules, explicitly including the nuclear relaxation contribution is essential for a good description of solvation energies.

Table 3 Change in RMSE (relative to the baseline) for solvation energies ($\mathrm{kcal\cdot mol^{-1}}$) and partition ratios (log units) upon using different geometry approaches. A negative change in RMSE indicates an improvement, a positive change a worsening. *Gas ph.* denotes using only the gas phase geometry; *Sol. ph.* denotes using the solution phase geometry.

Method	$\Delta_{\mathrm{solv}}G$ ($\mathrm{kcal\cdot mol^{-1}}$)		$\log K_{\alpha/\beta}$ (log units)
	Gas ph.	Sol. ph.	Gas ph.
SMD	0.0	0.1	−0.1
openCOSMO-RS	0.0	−0.1	0.4
COSMO-RS	0.0	0.0	0.1
D-COSMO-RS	0.0	0.1	0.4
ALPB	−0.1	0.2	0.0
CPCM-X	−0.1	0.1	−0.1
ESE-PM7	0.1	0.0	−0.1
DirectML	0.2	0.2	−0.3
CIGIN	−0.3	−0.3	−0.3
ESE-GB-DNN	−0.1	0.0	−0.5

Additionally, we investigated the geometry effects by comparing two different approaches to our baseline (phase-specific geometries): (a), the gas phase geometry approach, where we relied exclusively on gas phase optimized structures for both phases, effectively assuming that the molecular geometry remains unchanged upon transfer to solution, and (b), the solution phase approach, where we only employ the solution phase geometry, which equally does not capture the respective nuclear relaxation component. By systematically comparing these three protocols, we assess the sensitivity of the results to nuclear relaxation effects. Importantly, the models's solvation energies depend explicitly on the molecular geometry supplied. Thus, evaluating the same solvation model on a gas-phase versus a solution-phase optimized structure does not simply omit the nuclear relaxation contribution but also changes the calculated solvation term itself, making the total result non-linear with respect to G_{N}. The resulting change in RMSE compared to the baseline is given in Tab. 3.

Compared to the prior section, no general trends are visible; only methods within a category show a similar tendency: QM-based methods benefit only partly from using solvent-specific geometries, and SQM- and ML-based methods give more a inhomogeneous picture, often improving noticeably ($< -0.3\,\mathrm{kcal\cdot mol^{-1}}$ and < -0.3 log units) when gas-phase structures are used for computing solvation energies and partition ratios.

For the QM-based methods, no overall change is observed for solvation energies, while partition ratios worsen on average by

View Article Online
DOI: 10.1039/D5SC06406F

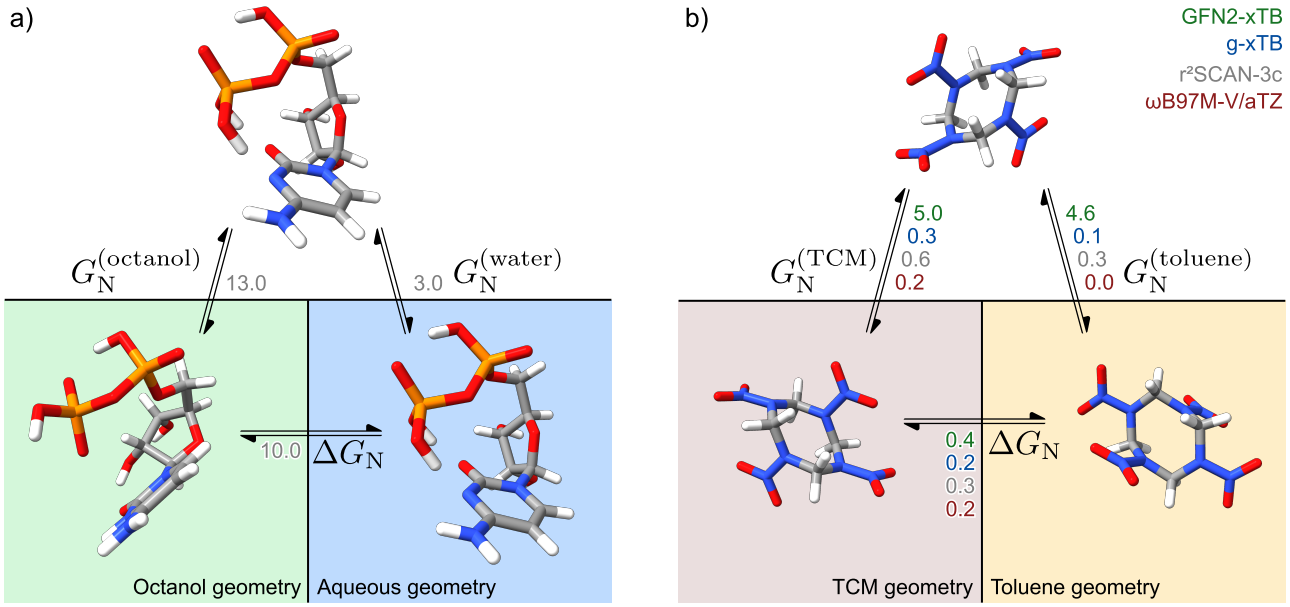

Fig. 9 The top structure is the gas phase geometry; the lower (shaded background) ones are the solution phase geometries. Vertical arrows give method-specific G_N values (kcal·mol^{-1}). The horizontal arrow between solution phase structures shows ΔG_N. (a) Nuclear relaxation for cytidine diphosphate (CDP) in octanol and water. (b) Error cancellation in ΔG_N for partition-ratio calculations of cyclotetramethylene tetranitramine in trichloromethane (TCM) and toluene.

about 0.3 log:units when using gas phase geometries only. This effect is most pronounced for openCOSMO-RS and D-COSMO-RS, which rely on phase-specific geometries (worsening by $\sim$ 0.4 log:units), whereas SMD and COSMO-RS are only marginally affected ($\pm$0.1 kcal·mol^{-1} and 0.1 log:units). The largest deteriorations occur for larger, highly functionalized molecules, where nuclear relaxation contributions are substantial. For example, the solvation energy of penoxsulam in water worsens by 4.8 kcal·mol^{-1} across the QM models. Of this, 2.3 kcal·mol^{-1} arises from the neglected nuclear relaxation, and the remaining error stems from using an inappropriate geometry, which causes the solvation term itself to be poorly described.

For the SQM-based models, the trends are more heterogeneous, especially for the solvation energies which are improved by using gas phase geometries, and worsened upon using the solution phase one. Solvation energies and partition ratios are improved by around 0.1 kcal·mol^{-1} and log units upon using the gas phase geometry; solvation energies are worsened by 0.1 kcal·mol^{-1} when using solution phase geometries. Due to the generally larger errors in these models, the influence of the geometry used is not very large overall, even though the computed values change noticeably (on average by 0.3 kcal·mol^{-1} and 0.3 log units for $\Delta_{solv}G$ and $\log K_{\alpha/\beta}$, respectively). For the machine learning models, the respective change is very method dependent: Computed solvation energies with DirectML worsen upon using either only gas or solution phase geometries, while CIGIN improves. Partition ratios, however, are improved in all cases upon using the gas phase geometry. For the partition ratios, we find generally a noticeable improvement throughout both SQM and ML-bases methods of around 0.3 log units.

This is noteworthy as those trends of SQM- and ML-based models are opposite of the QM-based ones which is attributed to the parameterization strategy and model design: Especially machine-learning models are meant to predict the total solvation energy – including the nuclear contribution.[123] This means that using phase-specific geometries for these models (*i.e.*, including the geometry relaxation additionally) will double-count this effect, thus resulting in worse results. Similarly, this is also the reason for the semiempirical models, which were for the most part also parameterized on gas phase geometries and thus implicitly account for this effect already.[81,102,126] It is noteworthy, that SMD is also parameterized on gas phase structures only (MNSOL database), which can explain the reason SMD does not significantly profit from using phase-specific geometries.[43]

In this context, we would also like to note, that in order to obtain phase-specific solution phase geometries efficiently (*i.e.*, perform geometry optimizations), the analytic nuclear gradient for a method has to be available and implemented. This is not given for most of the tested methods: Only CPCM, SMD, D-COSMO-RS,

4.2.3 Electronic Energy

From the previous section, we saw that the nuclear contribution G_N is often non-negligible, especially for larger, more complex solutes. Here, we test whether evaluating the underlying electronic energies at a lower- or higher-level method affects the results. Importantly, we used the same gas- and solution-phase geometries as in the previous sections – no new optimizations were performed. Only the single-point electronic energies were recomputed at different levels of theory to assess their impact on G_N. Besides the baseline r^{2}SCAN-3c, we test: (a) the efficient GFN2-xTB method, (b) its much improved successor g-xTB, and (c) a high-level hybrid DFT functional, ωB97M-V, combined with a large triple-ζ basis set, as our highest-accuracy method. We selected these methods to cover a reasonable accuracy-cost

View Article Online
DOI: 10.1039/D5SC06406F

Table 4 Change in RMSE (relative to the baseline) for solvation energies (kcal·mol^{-1}) and partition ratios (log units) upon using different levels of theory for the underlying electronic energy to calculate G_N. A negative change in RMSE indicates an improvement, a positive change a worsening.

Method	$\Delta_{solv}G$ (kcal·mol^{-1})			$\log K_{\alpha/\beta}$ (log units)		
	GFN2-xTB	g-xTB	ωB97M-V	GFN2-xTB	g-xTB	ωB97M-V
SMD	0.0	0.0	−0.1	0.1	0.0	0.1
openCOSMO-RS	0.0	0.0	0.0	0.1	0.1	0.0
COSMO-RS	0.3	0.2	0.0	0.0	0.0	0.0
D-COSMO-RS	0.0	0.1	0.0	0.1	0.0	0.0
ALPB	0.0	0.0	−0.1	0.1	0.0	−0.1
CPCM-X	−0.1	−0.1	−0.1	0.1	0.0	0.0
ESE-PM7	0.1	0.1	0.1	0.1	0.1	0.0
DirectML	−0.1	−0.1	−0.1	0.1	0.1	0.1
CIGIN	0.3	0.3	−0.1	0.0	0.0	−0.1
ESE-GB-DNN	0.1	0.0	−0.1	−0.5	−0.3	0.0

range for G_N. GFN2-xTB and g-xTB provide very low-cost electronic energies, routinely used for screening and sampling purposes.[127] r^{2}SCAN-3c delivers accurate relative conformer energies and geometries for main-group organic systems at a moderate cost, with consistent performance across large benchmarks, making it a robust baseline for our set.[152,201–203] As our high-accuracy method, ωB97M-V with a large augmented triple-ζ basis represents a high-accuracy range-separated hybrid, chosen because it has been shown to perform exceptionally well, often rivaling double-hybrid functionals.[154,204] Tab. 4 shows the change in error relative to the baseline when substituting its electronic energy with these alternatives.

We generally find only small changes in the RMSE for solvation energies and partition ratios (≤ 0.1 kcal·mol^{-1}, ≤ 0.1 log units), with partition ratios being less affected. QM-based methods tend to slightly worsen at lower levels of theory, while some ML models show a more erratic result and small apparent improvements – likely due to error-cancellation effects. The smaller changes observed for partition ratios arise from cancellation of errors in G_N between the two individual phases. This is observed for, *e.g.*, cyclotetramethylene tetranitramine (shown in Fig. 9b), where the method specific error is present for both phases (chloroform and toluene), canceling perfectly for the ΔG_N between the phases and thus the calculation of the partition ratio.

Moving from r^{2}SCAN-3c to the much more expensive ωB97M-V produces negligible improvements. In most cases the changes in computed values are well below 0.5 kcal·mol^{-1}, with only 15 calculated values changing by more than 1 kcal·mol^{-1}. The most notable improvements are found for the resulting solvation energy of difethialone in water or that of bensulide in octanol, which both improve on average by around 1.5 kcal·mol^{-1} across the tested models. Other well-performing hybrids such as PBE0[147,205] or B3LYP,[206,207] combined with D4 or MBD[208,209] perform similarly well on various benchmark sets[61,210] and are therefore expected to yield results comparable to r^{2}SCAN-3c and ωB97M-V on this set. While GFN2-xTB yields slightly higher overall errors, this degradation is driven by only a small number of structures. Notable cases include fructose, penoxsulam, and 15-crown-5 (shown in Supplemental Material, Fig. C3), which worsen by about 3.0, 5.0, and 8.0 kcal·mol^{-1}, respectively, compared to r^{2}SCAN-3c. Such trends mirror errors observed in relative conformational-energy benchmarks[127,146] like the GLUCOSE205[211] or UPU46[212] sets. Using the much improved g-xTB method already eliminates most issues seen with GFN2-xTB and shows good performance for the computation of G_N.

4.3 Timings

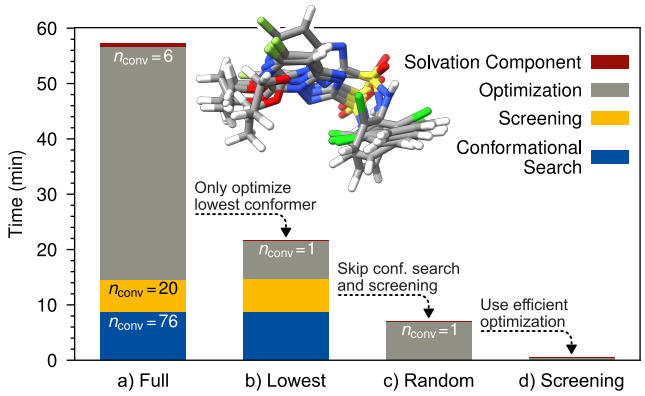

Fig. 10 Timings (wall-clock) in minutes to obtain a single Gibbs energy for diclosulam in water on 48 cores of an Intel Xeon Platinum 8468 CPU. Bars decompose total cost into conformer search (blue), screening (yellow), geometry optimization (gray) and final solvation evaluation (red); panels a-d compare the full ensemble, lowest-conformer, random-conformer and screening-only protocols, respectively (annotated n_{conv} values show retained conformers). The inset shows the six final optimized conformers of (a).

To give practical context for the investigated approaches above, we report wall-clock timings for the computational cost of those respective approaches (see Sec. 4.2.1. The Gibbs energy of diclosulam in water was chosen as a representative test case: diclosulam's molecular size, conformational flexibility, and the measured timings for the steps closely match the medians across *FlexiSol* (distributions shown in Supplemental Material, Sec. C.6.1). The timings are shown in Fig. 10. Computations were done on 48 cores of an Intel Xeon Platinum 8468 (Sapphire Rapids) CPU.

View Article Online
DOI: 10.1039/D5SC06406F

To obtain a final solvation energy, the same workflow must be run twice – in this case additionally with the gas phase, however, without the solvation model part. For completeness, timings of the solvation models themselves with their corresponding underlying method can be found in Supplemental Material, Sec. C.6.2.

Overall, the full baseline (a) approach takes around 1 h. The initial conformer screening takes around 9 min ($\sim$ 15%) and yields 76 initial conformers. The following screening takes 6 min ($\sim$ 10%) and reduced the number of relevant conformers to 20. The most expensive step, the optimization, takes around 42 min ($\sim$ 74%) and yields 6 final conformers in the specified 4 kcal·mol^{-1} window. The final solvation model evaluation (COSMO-RS) takes less than one minute ($<$ 1%), making it negligible relative contribution to the workflow. Using only the lowest-energy conformer per phase dramatically reduces optimization cost, because only a single geometry optimization per phase is required. The actual speed-up depends on how many conformers survive the screening step (and thus how many optimizations would otherwise be needed); for diclosulam this delivers roughly a 5$\times$ reduction in wall time for the solvation energy, while for very flexible cases (*e.g.*, ledipasvir, $\sim$ 935 initial conformers) the savings can reach orders of magnitude. Accelerated variants of the CENSO workflow have been proposed for such reduced-cost protocols and can be adopted where appropriate.[213,214] Skipping ensemble generation entirely (c), thereby simulating the common one-geometry practice by selecting a single random conformer, reduces the computation time moderately. While this eliminates the ensemble-generation and screening steps, it still requires at least one full geometry optimization per phase, so wall-time savings are limited compared to (b). For rapid, low-cost screening, it could therefore make sense to combine the random-conformer approach with a cheaper optimization method such as GFN2-xTB or g-xTB – forming a possible efficient screening protocol (d). This strategy could in principle reduce the computational cost of solvation energies and partition ratios by orders of magnitude compared to the full ensemble approach (a). However, we emphasize that we did not systematically test or benchmark such an SQM/ML-based screening workflow.

5 Conclusion and Outlook

In this work, we present *FlexiSol*, a comprehensive solvation model benchmark set with experimental reference data, consisting of 824 data points and 1551 unique molecule-solvent pairs. With the inclusion of most relevant conformers and tautomers for each respective phase, this totals over 25000 geometries. This set focuses on drug-like molecules between 30 to 80 atoms, with the mean being 42 atoms, and the largest molecule having 141 atoms – far surpassing common sets in that regard. Additionally, we focused on relevant and more difficult structures that are underrepresented in existing benchmarks: larger, flexible, and polyfunctional molecules. The benchmark, including all geometries and energies, is publicly available for future testing and development. In addition to creating the benchmark itself, we systematically tested and evaluated a wide range of popular solvation models and approaches, trying to isolate the effects of ensemble sampling, geometry choice, and underlying electronic energy.

Overall findings. Errors on *FlexiSol* exceed those previously seen on MNSOL or FreeSolv, reflecting the larger and more complex nature of the solutes and the wider range of $\Delta_{\text{solv}}G$ and $\log K_{\alpha/\beta}$ values. The QM-based models (COSMO-RS, SMD) show most consistent results with the best performing MAE of 2.0 kcal·mol^{-1} for $\Delta_{\text{solv}}G$, and 1.0 log units for $\log K_{\alpha/\beta}$. SQM-based methods perform slightly worse with an MAE of 2.8 kcal·mol^{-1} and 1.7 log units, respectively. The best performing ML method (DirectML) nearly matches the QM methods for solvation energies (MAE of 2.2 kcal·mol^{-1}) and yields the best result for partition ratios with an MAE of 0.7 log units. Most models systematically overestimate weak stabilization and underestimate strong stabilization, particularly for heteroatom-rich/H-bonding motifs, and errors increase in less polar solvents (octanol vs. water by around 0.3 kcal·mol^{-1} on average).

Ensemble. Using the phase-specific lowest-energy conformers reproduces the result using the full Boltzmann ensemble with minor deviations, with the largest changes being $\sim \pm 0.7$ kcal·mol^{-1} for solvation energies. Choosing a single random conformer per phase significantly degrades accuracy by > 0.3 kcal·mol^{-1} and ($\Delta_{\text{solv}}G$) and > 0.3 log units ($\log K_{\alpha/\beta}$) on average, showing that the conformational analysis step is essential. A simplified protocol using only the lowest conformer in each phase can reduce costs by over an order of magnitude at minimal accuracy loss.

Geometry (nuclear relaxation). Nuclear relaxation contributions G_{N} average ~ 0.6 kcal·mol^{-1} on our set (with $>$ 2.0 kcal·mol^{-1} in flexible/H-bonding cases). However, this does not directly translate into shifts in $\Delta_{\text{solv}}G$, because of solvation-model uncertainties and different model parameterization choices. For some QM-based methods (openCOSMO-RS and D-COSMO-RS), using non-phase-specific geometries worsens the result consistently, *e.g.*, for partition ratios by ≈ 0.3 log units – whereas SMD and CORMO-RS are hardly affected. SQM/ML-based models show an improvement using only gas phase geometries because G_{N} is partly absorbed into the empirical parameterization. Omitting the nuclear relaxation contributions by only using gas phase structures approximately halves the computational cost.

Electronic energy. Changing the level of the underlying electronic structure method for the description of the nuclear relaxation changes results only slightly (typically < 1.0 kcal·mol^{-1}); partition ratios are even less sensitive due to possible cancellation of G_N across phases. Upgrading from r^{2}SCAN-3c to the much more costly ωB97M-V/aTZ method produces little practical gain, while fast tight-binding approaches (notably g-xTB) typically provide an adequate and computationally efficient description of G_{N}.

Recommended protocols. In line with our findings, we recommend two practical protocols: The physics-based, high-rigor approach: Employ a QM-based solvation model such as COSMO-RS or SMD together with conformer/tautomer sampling, ensuring phase-specific geometries. Whether the full Boltzmann-weighted ensemble or just the lowest-energy conformer per phase is used makes little difference to the final accuracy, but conformer sampling is essential. Such a pro-

View Article Online
DOI: 10.1039/D5SC06406F

tocol, combined with r^{2}SCAN-3c single-point energies for the electronic component, yields robust accuracy on *FlexiSol* (MAE 2.0 kcal·mol^{-1} for $\Delta_{\text{solv}}G$; 1.0 log units for $\log K_{\alpha/\beta}$) and is preferred for heteroatom-rich/H-bonding or foreseen more complicated cases. Fast semiempirical or ML-based screening: Apply a modern ML- or SQM-based solvation model (DirectML or CPCM-X) together with a single gas phase geometry. This yields errors in the range of (MAE 2.2 kcal·mol^{-1} to 2.8 kcal·mol^{-1} for $\Delta_{\text{solv}}G$; 0.7 log units to 1.8 log units for $\log K_{\alpha/\beta}$), at a fraction of the cost – especially when resorting to a lower-cost method for the optimization.

Outlook and future directions. Finally, we want to outline options for extending scope and utility of future solvation benchmark sets. An important step is the extension beyond neutral drug-like molecules to inorganic and transition metal compounds,[165,215] proteins and macromolecular fragments,[216] and ions.[188,217,218] Complementing this chemical space expansion, a broadened solvent coverage and inclusion of additional properties such as acid dissociation constants,[68] solubilities,[219] and vapor pressures would be very beneficial.[220] Consistency-focused curation, *i.e.*, thermodynamic-cycle checks, cross-source reconciliation, and uncertainty propagation can improve reliability of experimental references.[221–223] Additionally to improvements to the set itself, an extension to explicit solvation approaches alongside implicit models using classical force fields like GAFF[224] and GFN-FF[225,226] or ML interatomic potentials like UMA,[227] SO3LR,[228] or PaiNN could give additional insights into computational modeling of solvation.[229]

Author Contributions

Conceptualization: LW, SG; Data curation: LW, CES; Formal Analysis: LW, CES; Funding acquisition: LW, SG; Investigation: LW, CES; Methodology: LW, CES; Project administration: LW, SG; Resources: SG; Software: LW; Supervision: LW, SG; Validation: LW, CES; Visualization: LW, CES; Writing (original draft): LW, CES; Writing (review) & editing: LW, CES, SG

Conflict of Interest

There are no conflicts of interest to disclose.

Data Availability Statement

All raw data for *FlexiSol* – including geometries (`xyz` files), all computed energies for all conformers/tautomers, computed solvation and partition ratios for each model and approach (`csv` files), experimental reference values and their bibliographic sources (`csv` file) – are provided as Supporting Information with this work. This data, including a tool for easy evaluation of methods on *FlexiSol* can additionally also be found on GitHub: https://github.com/grimme-lab/flexisol

Supporting Information Available

The Supporting Information for *FlexiSol* can be found attached to this paper and on GitHub (https://github.com/grimme-lab/flexisol) includes the following files:

- `flexisol-cli` (*FlexiSol* Analyzer)
  - Command-line tool to reproduce the data, compute statistics, and evaluate new methods on *FlexiSol*.
  - Repository and install instructions: `https://github.com/grimme-lab/flexisol` (see `README.md`).
  - Python-based, well-documented CLI; supports Boltzmann/minimum weighting, geometry approaches (gas/solv/full), sigma-clipping, per-datapoint and summarized statistics, batch evaluation, and plug-in of new methods.
- Additional Figures, Tables and Discussion

  PDF compile of all supplementary figures (systematic-error trends, solvent-specific performance) and extended tables (full model-coverage matrix), including additional discussion and comments.
- Geometries and Conformers
  - `flexisol`: XYZ coordinates for all conformers of each solute.
- Raw structures and Energies
  - `structures.csv`: Relative paths of all conformers with properties.
  - `energies.csv`: All raw computed energies for all conformers, methods and models in Hartree. `el_*` denotes the electronic energies; all other energies of the models are the obtained solvation contribution.
- Experimental reference values
  - `dgsolv-references.csv`: Experimental reference solvation energies $\Delta_{\text{solv}}G$ in kcal·mol^{-1}.
  - `logkab-references.csv`: Experimental reference partition rations $\log K_{\alpha/\beta}$ in log units.
- Computed results and errors for the discussed approaches. The regular results are given with the respective model name, the associated errors begin with `err_*`. The computed values and errors are kcal·mol^{-1} for $\Delta_{\text{solv}}G$ and in log units for $\log K_{\alpha/\beta}$.
  - `r2scan-3c-full-boltzmann-results.csv`: Baseline approach, r^{2}SCAN-3c using phase-spefific geometries and full Boltzmann weighting.
  - `r2scan-3c-full-minimum-results.csv`: r^{2}SCAN-3c using phase-spefific geometries and lowest conformer.
  - `r2scan-3c-full-random-results.csv`: r^{2}SCAN-3c using phase-spefific geometries and random conformer.
  - `r2scan-3c-gas-boltzmann-results.csv`: r^{2}SCAN-3c using gas phase geometries and full Boltzmann weighting.
  - `r2scan-3c-solv-boltzmann-results.csv`: r^{2}SCAN-3c using solution phase geometries and full Boltzmann weighting.

View Article Online
DOI: 10.1039/D5SC06406F

- `gfn2-full-boltzmann-results.csv`: GFN2-xTB using phase-spefific geometries and full Boltzmann weighting.
- `gxtb-full-boltzmann-results.csv`: g-xTB using phase-spefific geometries and full Boltzmann weighting.
- `wb97mv-def2tzvppd-full-boltzmann-results.csv`: Results for ωB97M-V/def2-TZVPPD with full Boltzmann weighting.

Acknowledgements

L. W. greatly acknowledges the support of the Stiftung Stipendien-Fonds des Verbandes der Chemischen Industrie e.V. through its Kekulé Fellowship.
The authors gratefully acknowledge the granted access to the Marvin cluster hosted by the University of Bonn and the help and support of the HPC/A Lab.
The authors acknowledge Dr. Uwe Huniar for providing parameter files for D-COSMO-RS.
The authors thank Thomas Gasevic, Tim K. Schramm, and Christoph Plett for fruitful discussions. Thomas Gasevic and Tim K. Schramm are additionally acknowledged for proofreading our manuscript.

Notes and references

1 R.-M. Dannenfelser and S. H. Yalkowsky, *Science of The Total Environment*, 1991, **109-110**, 625–628.
2 P. Muller, *Pure and Applied Chemistry*, 1994, **66**, 1077–1184.
3 G. Dyrda, E. Boniewska-Bernacka, D. Man, K. Barchiewicz and R. Słota, *Mol Biol Rep*, 2019, **46**, 3225–3232.
4 T. Gholami, H. Seifi, E. A. Dawi, M. Pirsaheb, S. Seifi, A. M. Aljeboree, A.-H. M. Hamoody, U. S. Altimari, M. Ahmed Abass and M. Salavati-Niasari, *Materials Science and Engineering: B*, 2024, **304**, 117370.
5 W. Thiel, *Angewandte Chemie International Edition*, 2011, **50**, 9216–9217.
6 M. Barbatti, *Pure and Applied Chemistry*, 2025.
7 M. Poliakoff, J. M. Fitzpatrick, T. R. Farren and P. T. Anastas, *Science*, 2002, **297**, 807–810.
8 M. G. Quesne, F. Silveri, N. H. De Leeuw and C. R. A. Catlow, *Front. Chem.*, 2019, **7**, 182.
9 G. N. Simm, A. C. Vaucher and M. Reiher, *The Journal of Physical Chemistry A*, 2019, **123**, 385–399.
10 O. Engkvist, P.-O. Norrby, N. Selmi, Y.-h. Lam, Z. Peng, E. C. Sherer, W. Amberg, T. Erhard and L. A. Smyth, *Drug Discovery Today*, 2018, **23**, 1203–1218.
11 C. M. A. Eichler and J. C. Little, *Environmental Science: Processes & Impacts*, 2020, **22**, 500–511.
12 B. B. de Souza and J. Meegoda, *Science of The Total Environment*, 2024, **926**, 171738.
13 E. Panieri, K. Baralic, D. Djukic-Cosic, A. Buha Djordjevic and L. Saso, *Toxics*, 2022, **10**, 44.
14 F. Haque and C. Fan, *iScience*, 2023, **26**, 107649.
15 D. S. Aga, F. Samara, L. Dronjak, S. Kanan, M. M. Mortula and L. Vahapoglu, *ACS ES&T Water*, 2024, **4**, 2785–2788.
16 V. Hatje, M. Sarin, S. G. Sander, D. Omanović, P. Ramachandran, C. Völker, R. O. Barra and A. Tagliabue, *Front. Mar. Sci.*, 2022, **9**, 936109.
17 T. Salthammer, J. Zhao, A. Schieweck, E. Uhde, T. Hussein, F. Antretter, H. Künzel, M. Pazold, J. Radon and W. Birmili, *Indoor Air*, 2022, **32**, e13039.
18 S. Endo, J. Hammer and S. Matsuzawa, *Environmental Science & Technology*, 2023, **57**, 8406–8413.
19 I. Abusallout, C. Holton, J. Wang and D. Hanigan, *Journal of Hazardous Materials Letters*, 2022, **3**, 100070.
20 M. Mudlaff, A. Sosnowska, L. Gorb, N. Bulawska, K. Jagiello and T. Puzyn, *Environment International*, 2024, **185**, 108568.
21 Q. Xiang, G. Shan, W. Wu, H. Jin and L. Zhu, *Environmental Pollution*, 2018, **242**, 1283–1290.
22 *Test No. 117: Partition Coefficient (n-Octanol/Water), HPLC Method | OECD*, https://www.oecd.org/en/publications/test-no-117-partition-coefficient-n-octanol-water-hplc-method_9789264069824-en.html.
23 Y. Li, H. Zhang and Q. Liu, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2012, **86**, 51–55.
24 A. Allerhand and P. v. R. Schleyer, *Journal of the American Chemical Society*, 1963, **85**, 371–380.
25 A. D. Buckingham, T. Schaefer and W. G. Schneider, *The Journal of Chemical Physics*, 1960, **32**, 1227–1233.
26 P. Laszlo, *Progress in Nuclear Magnetic Resonance Spectroscopy*, 1967, **3**, 231–402.
27 E. Jonas, S. Kuhn and N. Schlörer, *Magnetic Resonance in Chemistry*, 2022, **60**, 1021–1031.
28 A. García Alejo, N. De Silva, Y. Liu, T. L. Windus and M. Pérez García, *Solvent Extraction and Ion Exchange*, 2023, **41**, 241–251.
29 Y. Y. Rusakov, Y. A. Nikurashina and I. L. Rusakova, *The Journal of Chemical Physics*, 2024, **160**, 084109.
30 E. Antoniou, C. F. Buitrago, M. Tsianou and P. Alexandridis, *Carbohydrate Polymers*, 2010, **79**, 380–390.
31 K.-J. Liu and J. L. Parsons, *Macromolecules*, 1969, **2**, 529–533.
32 N. D. Kambaine, D. M. Shadrack and S. A. Vuai, *Journal of Molecular Liquids*, 2022, **345**, 117794.
33 S. Wan, R. H. Stote and M. Karplus, *The Journal of Chemical Physics*, 2004, **121**, 9539–9548.
34 W. L. Jorgensen, J. F. Blake and J. Buckner, *Chemical Physics*, 1989, **129**, 193–200.
35 T. P. Straatsma and H. J. C. Berendsen, *The Journal of Chemical Physics*, 1988, **89**, 5876–5886.
36 G. Duarte Ramos Matos, D. Y. Kyu, H. H. Loeffler, J. D. Chodera, M. R. Shirts and D. L. Mobley, *Journal of Chemical & Engineering Data*, 2017, **62**, 1559–1569.
37 T. Kloss, J. Heil and S. M. Kast, *The Journal of Physical Chemistry B*, 2008, **112**, 4337–4343.
38 K. Imamura, D. Yokogawa and H. Sato, *The Journal of Chemical Physics*, 2024, **160**, 050901.

View Article Online
DOI: 10.1039/D5SC06406F

39 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *Journal of Chemical Theory and Computation*, 2008, **4**, 877–887.
40 S. Decherchi, M. Masetti, I. Vyalov and W. Rocchia, *European journal of medicinal chemistry*, 2015, **0**, 27–42.
41 J. M. Herbert, *WIREs Comput Mol Sci*, 2021, **11**, e1519.
42 S. Miertuš, E. Scrocco and J. Tomasi, *Chemical Physics*, 1981, **55**, 117–129.
43 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *The Journal of Physical Chemistry B*, 2009, **113**, 6378–6396.
44 A. Klamt, *J. Phys. Chem.*, 1995, **99**, 2224–2235.
45 A. Klamt, V. Jonas, T. Bürger and J. C. W. Lohrenz, *The Journal of Physical Chemistry A*, 1998, **102**, 5074–5085.
46 J. R. Pliego, *J. Phys. Chem. A*, 2024, **128**, 6440–6449.
47 S. C. L. Kamerlin, M. Haranczyk and A. Warshel, *ChemPhysChem*, 2009, **10**, 1125–1134.
48 J. R. Pliego and J. M. Riveros, *J. Phys. Chem. A*, 2001, **105**, 7241–7247.
49 S. A. Katsyuba, S. Spicher, T. P. Gerasimova and S. Grimme, *J. Phys. Chem. B*, 2020, **124**, 6664–6670.
50 S. N. Steinmann, P. Sautet and C. Michel, *Phys. Chem. Chem. Phys.*, 2016, **18**, 31850–31861.
51 E. Tang, D. Di Tommaso and N. H. De Leeuw, *Phys. Chem. Chem. Phys.*, 2010, **12**, 13804.
52 M. Steiner, T. Holzknecht, M. Schauperl and M. Podewitz, *Molecules*, 2021, **26**, 1793.
53 P. L. Türtscher and M. Reiher, *J. Chem. Theory Comput.*, 2025, **21**, 5571–5587.
54 A. Fredenslund, R. L. Jones and J. M. Prausnitz, *AIChE Journal*, 1975, **21**, 1086–1099.
55 K. Mansouri, C. M. Grulke, R. S. Judson and A. J. Williams, *Journal of Cheminformatics*, 2018, **10**, 10.
56 P. Katzberger, F. Pultar and S. Riniker, *J. Chem. Theory Comput.*, 2025, **21**, 7450–7459.
57 O. D. Abarbanel and G. R. Hutchison, *Journal of Chemical Theory and Computation*, 2024, **20**, 6946–6956.
58 D. H. Kenney, R. C. Paffenroth, M. T. Timko and A. R. Teixeira, *J Cheminform*, 2023, **15**, 9.
59 L. Attia, J. W. Burns, P. S. Doyle and W. H. Green, *Nature Communications*, 2025, **16**, 7497.
60 K. Mansouri, C. M. Grulke, A. M. Richard, R. S. Judson and A. J. Williams, *SAR and QSAR in Environmental Research*, 2016, **27**, 911–937.
61 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.
62 J. Řezáč, *J. Chem. Theory Comput.*, 2020, **16**, 6305–6316.
63 N. Mardirossian and M. Head-Gordon, *Molecular Physics*, 2017, **115**, 2315–2372.
64 A. V. Marenich, C. P. Kelly, J. D. Thompson, G. D. Hawkins, C. C. Chambers, D. J. Giesen, P. Winget, C. J. Cramer and D. G. Truhlar, *Minnesota Solvation Database (MNSOL) Version 2012*, 2020.
65 D. L. Mobley and J. P. Guthrie, *J Comput Aided Mol Des*, 2014, **28**, 711–720.
66 C. Plett, M. Stahn, M. Bursch, J.-M. Mewes and S. Grimme, *J. Phys. Chem. Lett.*, 2024, **15**, 2462–2469.
67 C. . Hille, S. . Ringe, M. . Deimel, C. . Kunkel, W. E. . Acree, K. . Reuter and H. . Oberhofer, *Solv@TUM v 1.0*, 2018, `https://mediatum.ub.tum.de/1452571?v=1`.
68 J. Zheng and L.-J. Olivier, *IUPAC/Dissociation-Constants: v2.3b*, 2025, `https://zenodo.org/records/15375522`.
69 X. Ying, *J. Phys.: Conf. Ser.*, 2019, **1168**, 022022.
70 D. M. Hawkins, *J. Chem. Inf. Comput. Sci.*, 2004, **44**, 1–12.
71 R. Sander, W. E. Acree, A. De Visscher, S. E. Schwartz and T. J. Wallington, *Pure and Applied Chemistry*, 2022, **94**, 71–85.
72 H. Cumming and C. Rücker, *ACS Omega*, 2017, **2**, 6244–6249.
73 J. De Bruijn, F. Busser, W. Seinen and J. Hermens, *Environmental Toxicology and Chemistry*, 1989, **8**, 499–512.
74 B. McDuffie, *Chemosphere*, 1981, **10**, 73–83.
75 H. Devoe, M. Miller and S. Wasik, *Journal of Research of the National Bureau of Standards*, 1981, **86**, 361.
76 J. Staudinger and P. V. Roberts, *Critical Reviews in Environmental Science and Technology*, 1996, **26**, 205–297.
77 S.-H. Lee, S. Mukherjee, B. Brewer, R. Ryan, H. Yu and M. Gangoda, *J. Chem. Educ.*, 2013, **90**, 495–499.
78 J. Kames, S. Schweighoefer and U. Schurath, *Journal of Atmospheric Chemistry*, 1991, **12**, 169–180.
79 H. S. S. Ip, X. H. H. Huang and J. Z. Yu, *Geophysical Research Letters*, 2009, **36**, 2008GL036212.
80 M. Bursch, J.-M. Mewes, A. Hansen and S. Grimme, *Angew Chem Int Ed*, 2022, **61**, e202205735.
81 S. Ehlert, M. Stahn, S. Spicher and S. Grimme, *Journal of Chemical Theory and Computation*, 2021, **17**, 4250–4261.
82 S. Grimme, F. Bohle, A. Hansen, P. Pracht, S. Spicher and M. Stahn, *The Journal of Physical Chemistry A*, 2021, **125**, 4039–4054.
83 Y. C. Martin, *Journal of Computer-Aided Molecular Design*, 2009, **23**, 693–704.
84 D. Suárez and N. Díaz, *WIREs Comput Mol Sci*, 2015, **5**, 1–26.
85 J. Gorges, S. Grimme, A. Hansen and P. Pracht, *Phys. Chem. Chem. Phys.*, 2022, **24**, 12249–12259.
86 *IUPAC - Solvation Energy (ST07102)*, https://goldbook.iupac.org/terms/view/ST07102.
87 E. Cancès, B. Mennucci and J. Tomasi, *The Journal of Chemical Physics*, 1997, **107**, 3032–3041.
88 A. Klamt and G. Schüürmann, *Journal of the Chemical Society, Perkin Transactions 2*, 1993, **0**, 799–805.
89 V. Barone and M. Cossi, *J. Phys. Chem. A*, 1998, **102**, 1995–2001.
90 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *Journal of Chemical Theory and Computation*, 2013, **9**, 609–620.
91 S. Miertuš and J. Tomasi, *Chemical Physics*, 1982, **65**, 239–245.
92 J. Tomasi and M. Persico, *Chemical Reviews*, 1994, **94**, 2027–

View Article Online
DOI: 10.1039/D5SC06406F

2094.
93 C. Colominas, F. J. Luque, J. Teixidó and M. Orozco, *Chemical Physics*, 1999, **240**, 253–264.
94 R. D. Cunha, S. Romero-Téllez, F. Lipparini, F. J. Luque and C. Curutchet, *Journal of Computational Chemistry*, 2025, **46**, e70027.
95 A. Klamt, *WIREs Comput Mol Sci*, 2018, **8**, e1338.
96 A. Pomogaeva, D. W. Thompson and D. M. Chipman, *Chemical Physics Letters*, 2011, **511**, 161–165.
97 A. Pomogaeva and D. M. Chipman, *Journal of Chemical Theory and Computation*, 2011, **7**, 3952–3960.
98 A. Pomogaeva and D. M. Chipman, *The Journal of Physical Chemistry A*, 2013, **117**, 5812–5820.
99 Z.-Q. You and J. M. Herbert, *Journal of Chemical Theory and Computation*, 2016, **12**, 4338–4346.
100 R. Sundararaman, K. A. Schwarz, K. Letchworth-Weaver and T. A. Arias, *The Journal of Chemical Physics*, 2015, **142**, 054102.
101 R. Sundararaman and W. A. Goddard, III, *The Journal of Chemical Physics*, 2015, **142**, 064107.
102 M. Stahn, S. Ehlert and S. Grimme, *The Journal of Physical Chemistry A*, 2023, **127**, 7036–7043.
103 S. F. Vyboishchikov and A. A. Voityuk, *Journal of Computational Chemistry*, 2021, **42**, 1184–1194.
104 O. J. Conquest, T. Roman, A. Marianov, A. Kochubei, Y. Jiang and C. Stampfl, *J. Chem. Theory Comput.*, 2021, **17**, 7753–7771.
105 D. L. Mobley, K. A. Dill and J. D. Chodera, *J. Phys. Chem. B*, 2008, **112**, 938–946.
106 V. Radtke, D. Stoica, I. Leito, F. Camões, I. Krossing, B. Anes, M. Roziková, L. Deleebeeck, S. Veltzé, T. Näykki, F. Bastkowski, A. Heering, N. Dániel, R. Quendera, L. Liv, E. Uysal and N. Lawrence, *Pure and Applied Chemistry*, 2021, **93**, 1049–1060.
107 *GitHub - hoelzerC/Smi2xyz: Conversion of SMILES to Xyz*, https://github.com/hoelzerC/smi2xyz.
108 P. Pracht, C. A. Bauer and S. Grimme, *Journal of Computational Chemistry*, 2017, **38**, 2618–2631.
109 P. Pracht, S. Grimme, C. Bannwarth, F. Bohle, S. Ehlert, G. Feldmann, J. Gorges, M. Müller, T. Neudecker, C. Plett, S. Spicher, P. Steinbach, P. A. Wesołowski and F. Zeller, *The Journal of Chemical Physics*, 2024, **160**, 114110.
110 N. Van Staalduinen and C. Bannwarth, *Digital Discovery*, 2024, **3**, 2298–2319.
111 B. De Souza, *Angew Chem Int Ed*, 2025, **64**, e202500393.
112 F. Neese, *WIREs Comput Mol Sci*, 2025, **15**, e70019.
113 M. Garcia-Ratés and F. Neese, *Journal of Computational Chemistry*, 2020, **41**, 922–939.
114 A. Kovács, *Symmetry*, 2024, **16**, 1668.
115 T. Gerlach, S. Müller, A. G. De Castilla and I. Smirnova, *Fluid Phase Equilibria*, 2022, **560**, 113472.
116 S. Sinnecker, A. Rajendran, A. Klamt, M. Diedenhofen and F. Neese, *The Journal of Physical Chemistry A*, 2006, **110**, 2235–2245.
117 M. Renz, M. Kess, M. Diedenhofen, A. Klamt and M. Kaupp, *J. Chem. Theory Comput.*, 2012, **8**, 4189–4203.
118 A. Klamt and M. Diedenhofen, *J. Phys. Chem. A*, 2015, **119**, 5439–5445.
119 M. Nottoli, R. Nifosì, B. Mennucci and F. Lipparini, *Journal of Chemical Theory and Computation*, 2021, **17**, 5661–5672.
120 F. Lipparini, G. Scalmani, L. Lagardère, B. Stamm, E. Cancès, Y. Maday, J.-P. Piquemal, M. J. Frisch and B. Mennucci, *The Journal of Chemical Physics*, 2014, **141**, 184108.
121 Y. Minenkov, *Journal of Chemical Theory and Computation*, 2023, **19**, 5221–5230.
122 S. F. Vyboishchikov and A. A. Voityuk, *Journal of Chemical Information and Modeling*, 2021, **61**, 4544–4553.
123 Y. Chung, F. H. Vermeire, H. Wu, P. J. Walker, M. H. Abraham and W. H. Green, *Journal of Chemical Information and Modeling*, 2022, **62**, 433–446.
124 Y. Pathak, S. Mehta and U. D. Priyakumar, *Journal of Chemical Information and Modeling*, 2021, **61**, 689–698.
125 S. F. Vyboishchikov, *Journal of Chemical Information and Modeling*, 2023, **63**, 6283–6292.
126 S. F. Vyboishchikov, *Journal of Computational Chemistry*, 2025, **46**, e70104.
127 T. Froitzheim, M. Müller, A. Hansen and S. Grimme, *G-xTB: A General-Purpose Extended Tight-Binding Electronic Structure Method For the Elements H to Lr (Z=1–103)*, 2025.
128 *GitHub - Tnbrowncontam/Ifsqsar*, https://github.com/tnbrowncontam/ifsqsar.
129 A. Pedretti, A. Mazzolari, S. Gervasoni, L. Fumagalli and G. Vistoli, *Bioinformatics*, 2021, **37**, 1174–1175.
130 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chemical Physics Letters*, 1989, **162**, 165–169.
131 S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. Van Wüllen, V. K. Voora, F. Weigend, A. Wodyński and J. M. Yu, *The Journal of Chemical Physics*, 2020, **152**, 184107.
132 Y. J. Franzke, C. Holzer, J. H. Andersen, T. Begušić, F. Bruder, S. Coriani, F. Della Sala, E. Fabiano, D. A. Fedotov, S. Fürst, S. Gillhuber, R. Grotjahn, M. Kaupp, M. Kehry, M. Krstić, F. Mack, S. Majumdar, B. D. Nguyen, S. M. Parker, F. Pauly, A. Pausch, E. Perlt, G. S. Phun, A. Rajabi, D. Rappoport, B. Samal, T. Schrader, M. Sharma, E. Tapavicza, R. S. Treß, V. Voora, A. Wodyński, J. M. Yu, B. Zerulla, F. Furche, C. Hättig, M. Sierka, D. P. Tew and F. Weigend, *J. Chem. Theory Comput.*, 2023, **19**, 6859–6890.
133 F. Neese, *WIREs Comput Mol Sci*, 2012, **2**, 73–78.
134 F. Neese, F. Wennmohs, U. Becker and C. Riplinger, *The Journal of Chemical Physics*, 2020, **152**, 224108.
135 S. Lehtola, C. Steigemann, M. J. Oliveira and M. A. Marques, *SoftwareX*, 2018, **7**, 1–5.

View Article Online
DOI: 10.1039/D5SC06406F

136 D. Andrae, U. Häußermann, M. Dolg, H. Stoll and H. Preuß, *Theoretica chimica acta*, 1990, **77**, 123–141.
137 K. A. Peterson, D. Figgen, E. Goll, H. Stoll and M. Dolg, *The Journal of Chemical Physics*, 2003, **119**, 11113–11123.
138 G. L. Stoychev, A. A. Auer and F. Neese, *Journal of Chemical Theory and Computation*, 2017, **13**, 554–562.
139 F. Neese, F. Wennmohs, A. Hansen and U. Becker, *Chemical Physics*, 2009, **356**, 98–109.
140 R. Izsák and F. Neese, *The Journal of Chemical Physics*, 2011, **135**, 144105.
141 B. Helmich-Paris, B. De Souza, F. Neese and R. Izsák, *The Journal of Chemical Physics*, 2021, **155**, 104109.
142 K. Eichkorn, O. Treutler, H. Öhm, M. Häser and R. Ahlrichs, *Chemical Physics Letters*, 1995, **240**, 283–290.
143 F. Weigend, *Phys. Chem. Chem. Phys.*, 2006, **8**, 1057.
144 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *Journal of Cheminformatics*, 2011, **3**, 33.
145 P. Pracht, F. Bohle and S. Grimme, *Physical Chemistry Chemical Physics*, 2020, **22**, 7169–7192.
146 C. Bannwarth, S. Ehlert and S. Grimme, *Journal of Chemical Theory and Computation*, 2019, **15**, 1652–1671.
147 K. Burke, M. Ernzerhof and J. P. Perdew, *Chemical Physics Letters*, 1997, **265**, 115–120.
148 E. Caldeweyher, C. Bannwarth and S. Grimme, *The Journal of Chemical Physics*, 2017, **147**, 034112.
149 E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, *The Journal of Chemical Physics*, 2019, **150**, 154122.
150 L. Wittmann, I. Gordiy, M. Friede, B. Helmich-Paris, S. Grimme, A. Hansen and M. Bursch, *Phys. Chem. Chem. Phys.*, 2024, **26**, 21379–21394.
151 F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297.
152 S. Grimme, A. Hansen, S. Ehlert and J. M. Mewes, *J. Chem. Phys.*, 2021, **154**, 64103.
153 T. Gasevic, J. B. Stückrath, S. Grimme and M. Bursch, *J. Phys. Chem. A*, 2022, **126**, 3826–3838.
154 N. Mardirossian and M. Head-Gordon, *The Journal of Chemical Physics*, 2016, **144**, 214110.
155 F. Eckert and A. Klamt, *AIChE Journal*, 2002, **48**, 369–385.
156 C. G. . C. KG, *COSMOtherm, Version C3.0, Release 17.01*, 2017.
157 J. P. Perdew and W. Yue, *Physical Review B*, 1986, **33**, 8800–8802.
158 A. Schäfer, H. Horn and R. Ahlrichs, *The Journal of Chemical Physics*, 1992, **97**, 2571–2577.
159 K. Eichkorn, F. Weigend, O. Treutler and R. Ahlrichs, *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 1997, **97**, 119–124.
160 A. Katbashev, M. Stahn, T. Rose, V. Alizadeh, M. Friede, C. Plett, P. Steinbach and S. Ehlert, *The Journal of Physical Chemistry A*, 2025, **129**, 2667–2682.
161 M. Nottoli, M. F. Herbst, A. Mikhalev, A. Jha, F. Lipparini and B. Stamm, *WIREs Computational Molecular Science*, 2024, **14**, e1726.
162 A. A. Voityuk and S. F. Vyboishchikov, *Physical Chemistry Chemical Physics*, 2020, **22**, 14591–14598.
163 S. F. Vyboishchikov, *Journal of Chemical Theory and Computation*, 2023, **19**, 8340–8350.
164 S. Baskaran, Y. D. Lei and F. Wania, *Journal of Physical and Chemical Reference Data*, 2021, **50**, 043101.
165 R. Sander, *Atmos. Chem. Phys.*, 2023, **23**, 10901–12440.
166 F. H. Vermeire and W. H. Green, *Chemical Engineering Journal*, 2021, **418**, 129307.
167 L. M. Grubbs, M. Saifullah, N. E. De La Rosa, S. Ye, S. S. Achi, W. E. Acree and M. H. Abraham, *Fluid Phase Equilibria*, 2010, **298**, 48–53.
168 G. Bronner, K. Fenner and K.-U. Goss, *Fluid Phase Equilibria*, 2010, **299**, 207–215.
169 S. Lee, K.-H. Cho, C. J. Lee, G. E. Kim, C. H. Na, Y. In and K. T. No, *Journal of Chemical Information and Modeling*, 2011, **51**, 105–114.
170 S. H. Hilal, S. N. Ayyampalayam and L. A. Carreira, *Environmental Science & Technology*, 2008, **42**, 9231–9236.
171 R.-U. Ebert, R. Kühne and G. Schüürmann, *Environmental Science & Technology*, 2023, **57**, 976–984.
172 R. Naef and W. E. Acree, *Liquids*, 2024, **4**, 231–260.
173 S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang and E. E. Bolton, *Nucleic Acids Research*, 2024, **53**, D1516–D1525.
174 C. F. Poole, *Journal of Chromatography A*, 2023, **1706**, 464213.
175 M. Nendza, V. Kosfeld and C. Schlechtriem, *Environ Sci Eur*, 2025, **37**, 44.
176 C. Knox, M. Wilson, C. M. Klinger, M. Franklin, E. Oler, A. Wilson, A. Pon, J. Cox, N. E. L. Chin, S. A. Strawbridge, M. Garcia-Patino, R. Kruger, A. Sivakumaran, S. Sanford, R. Doshi, N. Khetarpal, O. Fatokun, D. Doucet, A. Zubkowski, D. Y. Rayat, H. Jackson, K. Harford, A. Anjum, M. Zakir, F. Wang, S. Tian, B. Lee, J. Liigand, H. Peters, R. Q. R. Wang, T. Nguyen, D. So, M. Sharp, R. da Silva, C. Gabriel, J. Scantlebury, M. Jasinski, D. Ackerman, T. Jewison, T. Sajed, V. Gautam and D. S. Wishart, *Nucleic Acids Research*, 2024, **52**, D1265–D1275.
177 D. S. Wishart, A. Guo, E. Oler, F. Wang, A. Anjum, H. Peters, R. Dizon, Z. Sayeeda, S. Tian, B. L. Lee, M. Berjanskii, R. Mah, M. Yamamoto, J. Jovel, C. Torres-Calzada, M. Hiebert-Giesbrecht, V. W. Lui, D. Varshavi, D. Varshavi, D. Allen, D. Arndt, N. Khetarpal, A. Sivakumaran, K. Harford, S. Sanford, K. Yee, X. Cao, Z. Budinski, J. Liigand, L. Zhang, J. Zheng, R. Mandal, N. Karu, M. Dambrova, H. B. Schiöth, R. Greiner and V. Gautam, *Nucleic Acids Research*, 2022, **50**, D622–D631.
178 Y. Liang, D. T. Kuo, H. E. Allen and D. M. Di Toro, *Chemosphere*, 2016, **161**, 429–437.
179 W. J. Zamora, A. Viayna, S. Pinheiro, C. Curutchet, L. Bis-

View Article Online
DOI: 10.1039/D5SC06406F

bal, R. Ruiz, C. Ràfols and F. J. Luque, *Physical Chemistry Chemical Physics*, 2023, **25**, 17952–17965.

180 R. Wolfenden, L. Andersson, P. M. Cullis and C. C. Southgate, *Biochemistry*, 1981, **20**, 849–855.

181 J. Hine and P. K. Mookerjee, *The Journal of Organic Chemistry*, 1975, **40**, 292–298.

182 J. P. Guthrie, *The Journal of Physical Chemistry. B*, 2009, **113**, 4501–4507.

183 *Test No. 107: Partition Coefficient (n-octanol/water): Shake Flask Method*, 1995, https://www.oecd.org/en/publications/test-no-107-partition-coefficient-n-octanol-water-shake-flask-method_9789264069626-en.html.

184 G. Ermondi, M. Vallaro, G. Goetz, M. Shalaeva and G. Caron, *Future Drug Discovery*, 2019, **1**, FDD10.

185 E. Fritschka and G. Sadowski, *Molecular Pharmaceutics*, 2025.

186 M. Işık, D. Levorse, D. L. Mobley, T. Rhodes and J. D. Chodera, *Journal of computer-aided molecular design*, 2020, **34**, 405–420.

187 G. Koch, A. Engstrom, J. Taechalertpaisarn, J. Faris, S. Ono, M. R. Naylor and R. S. Lokey, *Journal of Medicinal Chemistry*, 2024, **67**, 19612–19622.

188 Y. H. Zhao and M. H. Abraham, *J. Org. Chem.*, 2005, **70**, 2633–2640.

189 B. Case and R. Parsons, *Transactions of the Faraday Society*, 1967, **63**, 1224.

190 C. P. Kelly, C. J. Cramer and D. G. Truhlar, *The Journal of Physical Chemistry B*, 2006, **110**, 16066–16081.

191 T. Nevolianis, J. W. Zheng, S. Müller, M. Baumann, S. Tshepelevitsh, I. Kaljurand, I. Leito, I. Smirnova, W. H. Green and K. Leonhard, *Journal of the American Chemical Society*, 2025.

192 M. Stahn, S. Grimme, T. Salthammer, U. Hohm and W.-U. Palm, *Environmental Science: Processes & Impacts*, 2022, **24**, 2153–2166.

193 P. Pracht and S. Grimme, *The Journal of Physical Chemistry A*, 2021, **125**, 5681–5692.

194 S. Xu and B. Kropscott, *Analytical Chemistry*, 2012, **84**, 1948–1955.

195 A. Klamt, F. Eckert, M. Diedenhofen and M. E. Beck, *The Journal of Physical Chemistry A*, 2003, **107**, 9380–9386.

196 O. Andreussi, N. G. Hörmann, F. Nattino, G. Fisicaro, S. Goedecker and N. Marzari, *J. Chem. Theory Comput.*, 2019, **15**, 1996–2009.

197 J. Zhang, H. Zhang, T. Wu, Q. Wang and D. van der Spoel, *Journal of Chemical Theory and Computation*, 2017, **13**, 1034–1043.

198 F. Šebesta, Ž. Sovová and J. V. Burda, *The Journal of Physical Chemistry B*, 2024, **128**, 1627–1637.

199 P. Pernot, B. Huang and A. Savin, *Machine Learning: Science and Technology*, 2020, **1**, 035011.

200 L. Wittmann, T. Salthammer and U. Hohm, *Environmental Science: Processes & Impacts*, 2025.

201 C. Plett, S. Grimme and A. Hansen, *J Comput Chem*, 2024, **45**, 419–429.

202 S. Ehlert, S. Grimme and A. Hansen, *J. Phys. Chem. A*, 2022, **126**, 3521–3535.

203 C. Plett, S. Grimme and A. Hansen, *J. Chem. Theory Comput.*, 2024, acs.jctc.4c00801.

204 L. Wittmann, H. Neugebauer, S. Grimme and M. Bursch, *The Journal of Chemical Physics*, 2023, **159**, 224103.

205 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

206 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.

207 C. Lee, W. Yang and R. G. Parr, **37**, 785–789.

208 A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler, *Phys. Rev. Lett.*, 2012, **108**, 236402.

209 A. Ambrosetti, A. M. Reilly, R. A. DiStasio, Jr. and A. Tkatchenko, *J. Chem. Phys.*, 2014, **140**, 18A508.

210 M. Puleva, L. Medrano Sandonas, B. D. Lőrincz, J. Charry, D. M. Rogers, P. R. Nagy and A. Tkatchenko, *Nature Communications*, 2025, **16**, 8583.

211 M. Marianski, A. Supady, T. Ingram, M. Schneider and C. Baldauf, *Journal of Chemical Theory and Computation*, 2016, **12**, 6157–6168.

212 H. Kruse, A. Mladek, K. Gkionis, A. Hansen, S. Grimme and J. Sponer, *Journal of Chemical Theory and Computation*, 2015, **11**, 4972–4991.

213 B. B. Mészáros, K. Kubicskó, D. D. Németh and J. Daru, *J. Chem. Theory Comput.*, 2024, **20**, 7385–7392.

214 H. Mun, W. Lorpaiboon and J. Ho, *J. Phys. Chem. A*, 2024, **128**, 4391–4400.

215 J. R. Pliego, *Journal of Molecular Liquids*, 2022, **359**, 119368.

216 N. Ancona, A. Bastola and E. Alexov, *Journal of Computational Biophysics and Chemistry*, 2023, **22**, 515–524.

217 T. Nevolianis, M. Baumann, N. Viswanathan, W. A. Kopp and K. Leonhard, *Fluid Phase Equilibria*, 2023, **571**, 113801.

218 H. Zhang, V. Juraskova and F. Duarte, *Nature Communications*, 2024, **15**, 6114.

219 *Solubility Data Series - IUPAC | International Union of Pure and Applied Chemistry*, https://iupac.org/what-we-do/databases/solubility-data-series/.

220 H. Marques and S. Müller, *Fluid Phase Equilibria*, 2025, **592**, 114335.

221 J. Ferraz-Caetano, F. Teixeira and M. N. D. S. Cordeiro, *J. Chem. Inf. Model.*, 2024, **64**, 2250–2262.

222 A. A. Toropov, A. P. Toropova, A. Roncaglioni, E. Benfenati, D. Leszczynska and J. Leszczynski, *Molecules*, 2023, **28**, 7231.

223 D. Zhang, S. Xia and Y. Zhang, *J. Chem. Inf. Model.*, 2022, **62**, 1840–1848.

224 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of Computational Chemistry*, 2004, **25**, 1157–1174.

225 S. Spicher and S. Grimme, *Angewandte Chemie*, 2020, **132**, 15795–15803.

226 S. Grimme and T. Rose, *Zeitschrift für Naturforschung B*,

View Article Online
DOI: 10.1039/D5SC06406F

2024, **79**, 191–200.

227 B. M. Wood, M. Dzamba, X. Fu, M. Gao, M. Shuaibi, L. Barroso-Luque, K. Abdelmaqsoud, V. Gharakhanyan, J. R. Kitchin, D. S. Levine, K. Michel, A. Sriram, T. Cohen, A. Das, A. Rizvi, S. J. Sahoo, Z. W. Ulissi and C. L. Zitnick, *UMA: A Family of Universal Models for Atoms*, 2025.

228 A. Kabylda, J. T. Frank, S. Suárez-Dou, A. Khabibrakhmanov, L. Medrano Sandonas, O. T. Unke, S. Chmiela, K.-R. Müller and A. Tkatchenko, *Journal of the American Chemical Society*, 2025, **147**, 33723–33734.

229 K. T. Schütt, O. T. Unke and M. Gastegger, *Equivariant message passing for the prediction of tensorial properties and molecular spectra*, 2021, `http://arxiv.org/abs/2102.03150`, arXiv:2102.03150 [cs].

View Article Online
DOI: 10.1039/D5SC06406F

Data Availability Statement

Lukas Wittmann, Christian Erik Selzer, Stefan Grimme

October 9, 2025

All raw data for *FlexiSol* – including geometries (`xyz` files), all computed energies for all conformers/tautomers, computed solvation and partition ratios for each model and approach (`csv` files), experimental reference values and their bibliographic sources (`csv` file) – are provided as Supporting Information with this work. This data, including a tool for easy evaluation of methods on *FlexiSol* can additionally also be found on GitHub: https://github.com/grimme-lab/flexisol