Article

Solvation Thermodynamics in Different Solvents: Water–Chloroform Partition Coefficients from Grid Inhomogeneous Solvation Theory

Johannes Kraml, Florian Hofer, Anna S. Kamenik, Franz Waibl, Ursula Kahler, Michael Schauperl, and Klaus R. Liedl*

Cite This: *J. Chem. Inf. Model.* 2020, 60, 3843−3853

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Reliable information on partition coefficients plays a key role in drug development, as solubility decisively affects bioavailability. In a physicochemical context, the partition coefficient of a solute between two different solvents can be described as a function of solvation free energies. Hence, substantial scientific efforts have been made toward accurate predictions of solvation free energies in various solvents. The grid inhomogeneous solvation theory (GIST) facilitates the calculation of solvation free energies. In this study, we introduce an extended version of the GIST algorithm, which enables the calculation for chloroform in addition to water. Furthermore, GIST allows localization of enthalpic and entropic contributions. We test our approach by calculating partition coefficients between water and chloroform for a set of eight small molecules. We report a Pearson correlation coefficient of 0.96 between experimentally determined and calculated partition coefficients. The capability to reliably predict partition coefficients between water and chloroform and the possibility to localize their contributions allow the optimization of a compound's partition coefficient. Therefore, we presume that this methodology will be of great benefit for the efficient development of pharmaceuticals.
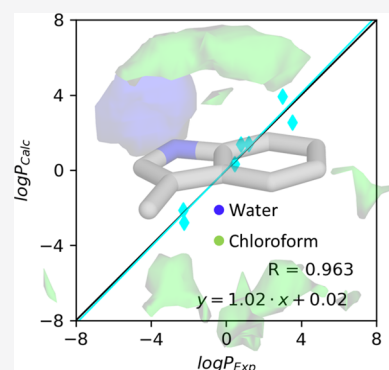
■ INTRODUCTION

For a sufficiently high bioavailability, a druglike compound generally needs to follow a set of four basic rules, also known as Lipinski's rule of five.[1,2] These simplified guidelines describe the complex interplay of physicochemical properties, which are required for a molecule to act as a drug, including solubility in environments of different polarities. On the one hand, molecules aiming at intracellular targets have to pass through the cell membrane, which is usually facilitated through a certain degree of lipophilicity.[3,4] On the other hand, it is decisive that the molecule features a sufficient number of hydrophilic groups to ensure solubility in the polar intra- and extracellular environments. Furthermore, it is well known that rather lipophilic, that is, "fatty" drugs, such as tetrahydrocannabinol, actually do show some extent of solubility in polar solvents.[5,6] However, these drugs can accumulate in fatty tissue, which often is followed by an extended uncontrolled release that causes several adverse effects.[7−10] Hence, a balanced equilibrium between solvent preferences, which can be described by the partition coefficient, is of utmost importance in drug design.[11]

One of Lipinski's rules states that the partition coefficient between water and octanol of a bioavailable compound, log P_{OW}, may not exceed 5. This rule highlights the importance of a reliable prediction of partition coefficients for drug discovery and design, as this allows preselection of compounds with

sufficient bioavailability.[12,13] Therefore, it is not surprising that in various countries, log P_{OW} must either be measured or calculated before a chemical compound can be licensed for commercial use.[14] Yet, the experimental approaches to measure log P suffer from various shortcomings, and recently, steps have been taken to remedy this situation.[15]

Several approaches have been proposed to predict partition coefficients for different solvents, both knowledge-based[16−26] and physics-based.[27−31] One of the standard methods to calculate the partition coefficient is the cLog P method.[25,32] This method splits the query molecule into a set of fragments for which the log P is known. The final value is computed by summing up the log P values for the fragments multiplied by their respective occurrences in the query molecule. Furthermore, steric, hydrogen bonding and electronic interactions are considered for the final cLog P value. Another approach to the cLog P's fragment-based method is to use the atomic contributions to calculate the log P. There are multiple

approaches that make use of this concept, for example, Alog P,[21,33,34] Xlog P,[26] and Mlog P.[23]

Most of the physics-based methods rely on the calculation of the differences in solvation free energies between solvents. One important approach to calculate the solvation free energy is the three-dimensional (3D) reference interaction site model (3DRISM).[35,36] This approach has recently been generalized for solvents other than water. Octanol in particular was the focus of Huang et al.[28] who studied the calculation of partition coefficients using the 3DRISM approach. The major advantage of 3DRISM over grid inhomogeneous solvation theory (GIST) is its lower calculation cost. The theoretical framework of 3DRISM inherently relies on an approximation when calculating the entropy, called the closure relation. The fundamental assumptions within the theory of inhomogeneous solvation theory (IST)-based methods, on the other hand, do not rely on this estimate. Nevertheless, because of practical limitations, GIST also introduces an approximation in the entropy calculation, which is usually stopped after the two-body term for the entropy. The entropy for IST-based approaches is directly calculated from the phase space of water. Additionally, the GIST method yields molecular distribution functions, whereas 3DRISM provides separate atomic distribution functions, which are more difficult to interpret.

Here, we propose a method based on GIST.[37−39] GIST calculates thermodynamic properties of a solvent around a solute on a grid. Recently, we introduced a new version of this algorithm, implemented on the GPU (GIGIST).[40] One of the key features of this algorithm is the tremendous speedup compared to the standard CPU implementation present in the program cpptraj[37−39,41] of AmberTools.[42] In addition to speeding up the calculations, we implemented a more generalized version of GIST enabling the use of chloroform as a solvent.

GIST calculations of localized thermodynamic quantities have already facilitated a broad range of applications, including the analysis of water structures in streptavidin,[43] calculation of hydrophobicity,[40,44] integration of solvation thermodynamics into structure−affinity relationships,[45] improvement of docking scoring functions,[46] and correlation of desolvation of aromatic moieties with binding free energies.[47] In all these approaches, the localized thermodynamic properties of water around a solute were found to be a key aspect in understanding the studied phenomena. Here, we aim toward extending this powerful tool to a different solvent, chloroform. Chloroform is particularly interesting, as it has been used as a simplistic model to approximate the polarity inside of proteins as well as membranes.[31] It is assumed to be a better suited model system than many other apolar solvents, as its relative permittivity of 4.3 is close to that typically expected for the interior of membranes.[48,49] Additionally, its structure shows only minimal flexibility, allowing it to be modeled as a fixed conformation. Therefore, chloroform can be treated similarly as water molecules in GIST.

Differences in solvation free energies in water and chloroform have previously been studied by Wolf and Groenhof[50] using TI simulations. The focus of their study was to benchmark the accuracy of varying force fields and water models in predicting solvation free energies. Additionally, they combine the calculated solvation free energies in water and chloroform to compare computationally and experimentally determined partition coefficients between

these two solvents. Their results indicate that there is a much higher impact of the given force field than of the different water models.

Here, we use the model systems studied by Wolf and Groenhof to test the reliability of GIST solvation free energies in estimating the partition coefficients between water and chloroform. We present a more generalized version of GIST that can calculate thermodynamic properties on a grid for chloroform in addition to water. From these values, we calculate solvation free energies, which are referenced to the solute in vacuum. Furthermore, we calculate partition coefficients between water and chloroform, which do not include the vacuum as a reference state anymore. We compare our results to earlier calculations of both, solvation free energies in water and chloroform, as well as to the experimental partition coefficients between these solvents. Furthermore, we calculate localized differences in the solvation free energies and their contributions.

■ METHODS

Compounds. We used the same eight compounds (Figure 1) as Wolf and Groenhof[50] to benchmark our approach
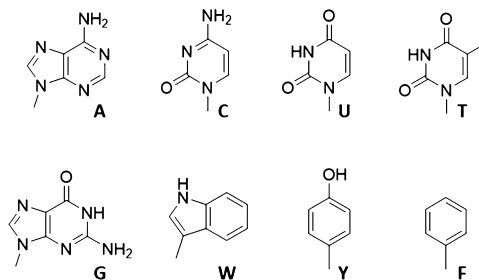

Figure 1. Eight compounds from Wolf and Groenhof[50] were used as a benchmark data set for our GIST-based approach. (**A**) 9-methyl-adenine, (**C**) 1-methyl-cytosine, (**U**) 1-methyl-uracil, (**T**) 1-methyl-thymine, (**G**) 9-methyl-guanine, (**W**) 3-methyl-indol, (**Y**) p-cresol, and (**F**) toluene. The abbreviation in parenthesis corresponds to the respective nucleobase or amino acid mimicked by the compound.

because these compounds are common motifs in biomolecules. Therefore, we will refer to these compounds by the single letter code of these biomolecules; that is, we will call the nucleobases as **A**, **G**, **C**, **T**, and **U** and 3-methylindole will be referred to as **W**, p-cresol as **Y**, and toluene as **F** in the following.

We consider these eight compounds as an ideal set to evaluate the reliability of our GIST calculations because for all but one (**Y**) of these compounds, reliable and coherent experimental data are available on their partition coefficients between water and chloroform. This is particularly beneficial as data on experimental partition coefficients between water and chloroform are significantly sparser than those between water and octanol, for example. Second, all compounds in this selected set are very rigid and essentially confined to the same conformation in water and chloroform. This eliminates the need for conformational sampling prior to the GIST simulations, which would be necessary for flexible solutes to obtain reasonable ensemble averages in both solvents. Because the grid-based approach requires restrained molecules to be analyzed, the GIST simulations would need to be repeated for multiple adequate ensemble representatives.

Grid Inhomogeneous Solvation Theory. A short introduction to GIST will be given in the following section.

For a more complete introduction into GIST,[37−39,51] and the underlying IST by Lazaridis,[52,53] the reader is referred to the original publications.

$$\Delta A(r_k) = \Delta E_{total}(r_k) - T\Delta S_{uv}^{total}(r_k) \tag{1}$$

$$\Delta E_{total}(r_k) = \Delta E_{vv}(r_k) + \Delta E_{uv}(r_k) \tag{2}$$

$$\Delta S_{uv}^{total}(r_k) = \Delta S_{uv}^{trans}(r_k) + \Delta S_{uv}^{orient}(r_k) \tag{3}$$

Generally, GIST can be used to calculate the free energy of solvation localized on a grid. This can be achieved by splitting the free energy (ΔA) into its two contributions, enthalpy (ΔE_{total}) and entropy (ΔS_{uv}^{total}), where u denotes the solute and v denotes the solvent (eq 1). The energetic contributions and the entropic contributions to the free energy are calculated on the grid, indicated by the r_k. The energetic contributions can be split into the solvent−solvent contributions (ΔE_{vv}) and the solute−solvent contributions (ΔE_{uv}), eq 2. The force field is used to calculate the different contributions to the energy. These values are then stored at the appropriate grid positions. The grid positions are assigned using a "central" atom, that is, oxygen in water and the carbon atom in chloroform.

The entropic contribution can be split into two parts, the translational entropy (ΔS_{uv}^{trans}) and the orientational entropy (ΔS_{uv}^{orient}) (eq 3). Here, the calculation is truncated after the two-body term of the entropy.

Both two-body entropy terms are calculated using a nearest-neighbor estimate.[39,54] The calculation of the nearest neighbor for the translational entropy is straightforward, as this can be calculated directly by using the Euclidean distance between the molecules. The nearest neighbor for the orientational entropy is calculated using quaternion distances $(\Delta\omega)$,[55] which are defined in eq 4.

$$\Delta\omega = 2 \cdot \arccos(|\langle q_1, q_2 \rangle|) \tag{4}$$

where $\langle q_1, q_2 \rangle$ denotes the inner product between the two quaternions (q_1, q_2) and $|.|$ denotes the absolute value. The quaternions are constructed for each water molecule using the vectors from the oxygen to the first hydrogen and from the oxygen to the second hydrogen. This has the effect that the different hydrogen atoms can be distinguished from each other, which has the disadvantage that more sampling is required until convergence.

Instead of splitting the entropy in the orientational and translational part, they can be estimated together, directly approximating the integral of sixth order. To achieve this, a distance (d) in the orientational and translational space is defined, which is a simple l^2 norm (eq 5). Using this distance, again the nearest-neighbor entropy estimator is used.

$$d = \sqrt{\Delta\omega^2 + d_{euclid}^2} \tag{5}$$

$\Delta\omega$ again refers to the orientational distance, calculated via eq 4, and d_{euclid} refers to a simple Euclidean distance in space.

Our implementation of GIGIST uses cpptraj[37−39,41] of AmberTools[42] as a base, benefitting from its various functionalities and making it straightforward to use. Similar to water, we also defined a central atom for chloroform in the GIST calculations. For chloroform, we used the carbon atom as a center for the binning calculation. Furthermore, we used the C−H bond and the bond between the carbon and the first chlorine atom for the calculation of the quaternions. As in the

water implementation for the hydrogen atoms, the chlorine atoms are distinguishable.

System and Simulation Setup. The structures of the solute molecules were generated with a molecular operating environment (MOE).[56] Hereafter, the parameters of the molecules were derived with the *antechamber* module of AmberTools19[42] and the GAFF/GAFF2[57] force fields. Partial charges were derived with Gaussian 16[58] and the RESP[59] procedure using HF/6-31G*.

With LEaP, the molecules were placed in TIP3P,[60] TIP4P,[60] TIP5P,[61] OPC3,[62] OPC,[63] SPC/E,[64] and CHCl$_3$[65] solvent boxes with a minimum wall distance of 20 Å. Restrained simulations for GIST were performed with the GPU implementation of the pmemd[66] module of AMBER18[42] on our in-house GPU cluster. A Langevin thermostat[67] with a collision frequency of 2 ps^{-1} was used to maintain a constant temperature of 300 K, and a Berendsen barostat[68] with a pressure relaxation time of 2 ps was used to maintain the atmospheric pressure. Long-range electrostatics were treated with the particle-mesh Ewald method,[69] and a van der Waals cutoff of 8 Å was used. Because a time step of 1 fs was necessary for the CHCl$_3$ box, the same time step was used for all simulations for consistency. All nonsolvent atom positions were restrained with a harmonic potential with a weight of 1000 kcal/mol·Å^2. All simulations were run for 100 ns, collecting frames every 10 ps, resulting in 10,000 frames per simulation.

Calculation of Solvation Free Energies. For the calculation of the solvation free energy, GIST analyses were performed on the simulations' trajectories. For the GIST calculations, the reference density and the solvent−solvent reference energy of the AMBER manual[42] were used, where they were available. For the systems for which no reference values were reported in the manual, a simple solvent box with a box size of 40 × 40 × 40 Å^3 was simulated. From this calculation, both the solvent−solvent reference energy and the reference density were calculated (see Supporting Information Table S1).

For the solute simulations, the solvation free energy was obtained with an in-house python script. The reference energy was subtracted from the solvent−solvent interaction energy. The energy density was recalculated from the normalized, referenced energy and the population on the given voxel. Then, the free energy was calculated as shown in eq 1. For the entropy contribution, the approximation of the integral of sixth order was used. The free energy density in each voxel was integrated around the molecule within 9 Å for the water models and 11 Å for chloroform to yield the solvation free energy. As a center for the cutoff, the heavy atoms of the compounds were used.

The cut-off distance was chosen based on an radial distribution function (RDF) analysis of solute simulations (see Supporting Information Figure S1). For both solvents, that is, water and chloroform, the cut-off distances were chosen to include the first two solvation shells of the molecules.

Calculation of log *P* and Δ*G* Difference. The difference in solvation free energy between water and chloroform can be understood in terms of the thermodynamic cycle. The solvation free energy corresponds to the transfer of the solute from vacuum to the solvent $(\Delta G_{water}$ and $\Delta G_{chloroform})$. The difference in these solvation free energies closes the thermodynamic cycle[70] and is related to the partition coefficient between the two solvents via eq 6.[50]

Table 1. Solvation Free Energies Calculated Using GAFF2 and GAFF Parameters of AMBER in Combination with Various Water Models[a]

	Solvent	A	G	C	T	U	W	Y	F
GAFF2	TIP3P	−12.2	−18.4	−16.6	−10.2	−11.3	−3.4	−1.8	−0.2
	TIP4P	−8.8	−16.6	−15.4	−9.4	−10.7	−1.9	−0.9	0.9
	TIP5P	−7.5	−13.8	−11.9	−6.8	−8.4	−3.5	−1.7	−0.3
	OPC	−5.8	−11.9	−11.0	−5.7	−6.8	0.8	2.9	3.4
	OPC3	−5.2	−11.7	−9.5	−4.4	−5.3	2.7	3.9	4.9
	SPCE	−9.5	−17.3	−15.2	−9.1	−10.9	−2.0	−0.9	0.0
	CHCl$_3$	−10.3	−14.9	−11.2	−9.8	−9.4	−7.2	−4.2	−3.1
GAFF	TIP3P	−13.6	−20.8	−18.1	−13.1	−13.9	−5.0	−3.7	−2.0
	CHCl$_3$	−11.3	−15.7	−11.6	−10.2	−10.1	−6.2	−4.3	−2.3
WOLF[50]	TIP3P	−15.5	−24.5	−21.0	−13.6	−14.1	−5.3	−4.2	−0.5
	CHCl$_3$	−12.8	−16.5	−13.2	−12.1	−11.4	−8.8	−6.6	−5.0
EXP	water[72,73]	−13.6			−(9−13)[b]		−5.9	−6.1	−0.8
	CHCl$_3$[50]	−12.5[c]					−9.0[c]		−3.9[c]

[a]The values calculated by Wolf and Groenhof[50] are provided for comparison. The experimental values are reported below. All energies are given in kcal/mol. [b]The value reported in ref 71 is an approximated range. [c]These values are calculated from the experimentally determined solvation free energy in water and the partition coefficient between water and chloroform.
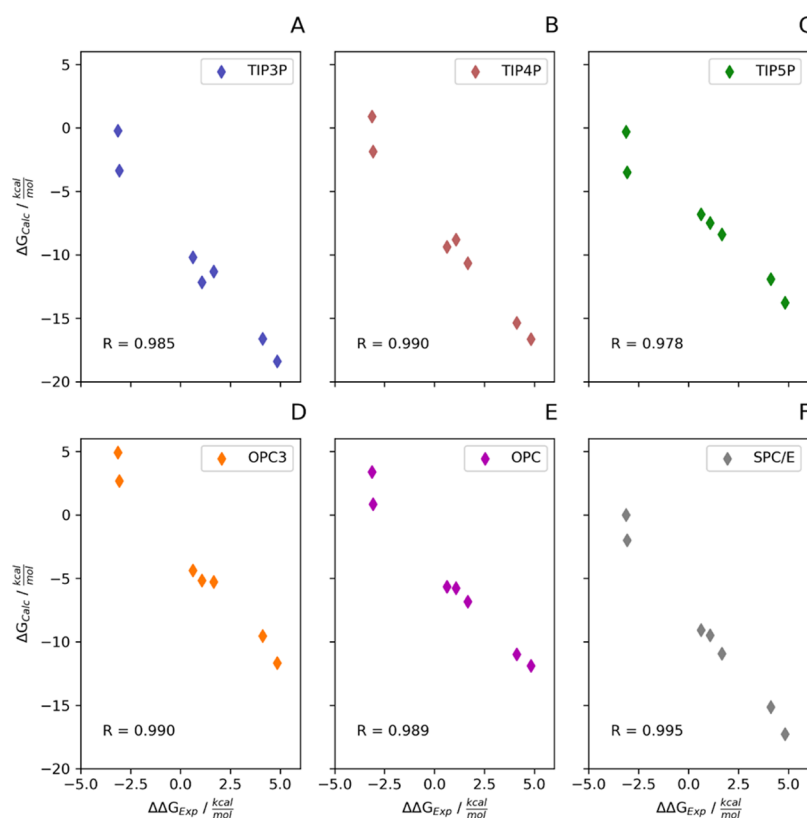
Figure 2. Solvation free energies of the compounds in the different water models, compared with the experimentally determined difference in solvation free energy between water and chloroform. (A) TIP3P, (B) TIP4P, (C) TIP5P, (D) OPC3, (E) OPC, and (F) SPC/E. The absolute Pearson correlation coefficient (R) is given for each solvent within the graph.

$$\log_{10}\left(\frac{c_{\text{water}}}{c_{\text{chloroform}}}\right) = -\frac{1}{2.303 \cdot RT}(\Delta G_{\text{water}} - \Delta G_{\text{chloroform}})$$

(6)

where c_{water} is the concentration of the solute in water and $c_{\text{chloroform}}$ is the concentration of the compound in chloroform. ΔG_{water} and $\Delta G_{\text{chloroform}}$ are the different solvation free energies, respectively, and RT is the ideal gas constant multiplied with the temperature.

The difference in solvation free energy is readily calculated, as mentioned above. The experimental difference in solvation free energy was calculated by reorganizing eq 6 to yield eq 7 and using the experimentally determined value for the water−chloroform partition coefficient.

$$\Delta G_{\text{chloroform}} - \Delta G_{\text{water}} = \log_{10}\left(\frac{c_{\text{water}}}{c_{\text{chloroform}}}\right) \times 2.303 \cdot RT$$

(7)

Table 2. Partition Coefficients between Water and Chloroform for All Eight Compounds Used in This Study with Experimental Values Taken from Wolf and Groenhof[50a]

	Model	A	G	C	T	U	W	Y	F	rmse
GAFF2	TIP3P	1.4	2.6	4.0	0.4	1.5	−2.7	−1.7	−2.1	0.6
	TIP4P	−1.0	1.3	3.1	−0.2	1.0	−3.8	−2.4	−2.9	1.3
	TIP5P	−2.0	−0.8	0.6	−2.1	−0.7	−2.6	−1.8	−2.0	2.5
	OPC	−3.2	−2.2	−0.1	−2.9	−1.8	−5.8	−5.1	−4.7	3.8
	OPC3	−3.7	−2.3	−1.2	−3.9	−3.0	−7.1	−5.9	−5.8	4.6
	SPC/E	−0.5	1.8	2.9	−0.5	1.2	−3.7	−2.4	−2.2	1.1
GAFF	TIP3P	1.7	3.8	4.7	2.2	2.8	−0.8	−0.3	−0.2	1.5
WOLF[50]	TIP3P	2.0	5.9	5.7	1.1	2.0	−2.6	−1.8	−3.4	1.6
EXP[74,75]		0.8	3.5	3.0	0.5	1.2	−2.2		−2.3	0

[a]The last column lists the rmse between the partition coefficient of chloroform and the studied water model. Because there is no experimental value for p-cresol (Tyr), we ignored it for the calculation of the rmse.

Grid Visualization. The values for the solvation free energy, the referenced energy, and the entropy were calculated on grids. To visualize the origins of the different contributions to the solvation free energy, a factor of 0.5 in the solvent–solvent interaction energy was retained to avoid double counting.[38] We calculated the differences in the two grids by simply subtracting the chloroform grid from the water grid. The resulting grid was visualized with cutoffs of ±0.05 kcal/(mol·A³). For free energy and reference energy, the negative cutoff (blue surface) shows highly favorable sites of water, and the positive cutoff (green surface) shows highly favorable sites for chloroform. For the entropic contributions, we visualized $-T\Delta S$: blue regions are entropically favorable sites for water and green regions entropically favor chloroform. All molecules and grids were visualized using PyMOL.[71]

■ RESULTS

Solvation Free Energies. The calculations of the free energies of solvation in water and chloroform were performed as described in the Methods section. In Table 1, the results for the two solvents are summarized and compared to the experimental solvation free energies and to the TI-based calculations by Wolf and Groenhof.[50] In general, our values are approximately 2 kcal/mol lower than the experimental values. Furthermore, the results calculated with the GAFF2 parameters are consistently lower than the results obtained by Wolf and Groenhof using the GAFF parameters. For better comparability, we calculated TIP3P and CHCl₃ solvation free energies using GAFF. Interestingly, even when using these parameters, GIST results are lower, by about 2 kcal/mol compared to TI.

The other water models seem to be less accurate than the TIP3P calculations. Comparison with experimental data shows that there is a root-mean-square error (rmse) of about 2.5 kcal/mol between the experimental values and the calculated values for TIP3P. For the other water models, the deviation is even larger and can be as high as 8.3 kcal/mol for OPC3. However, despite the offset in absolute values for OPC3, the Pearson correlation for this model is similar to the correlations found for TIP3P and TIP4P (all 0.99). While we find the worst correlation between the experimental partition coefficients and solvation free energies estimated from chloroform alone, the respective Pearson correlation coefficient of 0.91 is still very high (Supporting Information Figure S2).

The values calculated for chloroform also differ in our approach compared to the one by Wolf and Groenhof. The values we calculated with the GAFF2 parameters and our

GIST approach are more than 2 kcal/mol higher, that is, more positive, than the values calculated with TI.

Figure 2 shows the solvation free energy for the different water models (ΔG_{calc}). It is clearly visible that all the water models perform very similarly, when analyzed individually. All of them show a very high correlation with the difference in solvation free energies between water and chloroform ($\Delta\Delta G_{exp}$). It should be noted that we compare the hydration free energy to a difference in solvation free energies in Figure 2, and therefore, we do not expect the absolute numbers to match but rather focus on the correlation between these values. Intriguingly, the SPC/E water model has the highest correlation of 0.995, albeit only by a very small margin. The lowest correlation coefficient can be found for TIP5P, with 0.978, which is still very close to the highest 0.995 of SPC/E.

Partition Coefficients. From the difference of solvation free energies between each individual water model and chloroform, we then calculated the partition coefficient between the two solvents using eq 6. The results of the calculation for the water models are shown in Table 2. It is worth noting that we always use the same chloroform model. The water models show differences in terms of correlation and rmse with the experimental values. Most intriguingly, the TIP3P water model in combination with GAFF2 parameters does not only show a striking correlation with the experimental values but indeed almost matches them numerically. We find a Pearson correlation coefficient of 0.96 and an rmse in log P of 0.6 (rmse in $\Delta\Delta G$ of 0.89 kcal/mol). The correlation and rmse are worse for the TIP4P water model, which shows a Pearson correlation coefficient of 0.94 and an rmse of 1.3. Of the TIP water models, TIP5P clearly performs worst in this study, as it only shows a correlation coefficient of 0.81 with the experimental results and a significantly higher rmse of 2.5. While both OPC-type water models show a moderate Pearson correlation (0.89 for OPC and 0.95 for OPC3), the rmse values (3.8 for OPC and 4.6 for OPC3) are notably worse, compared to the results we find for the TIP family.

Compared to the previous study by Wolf and Groenhof (rmse of 1.6), in our approach, TIP3P, TIP4P, and SPC/E clearly perform better at reproducing the absolute ΔG values, with rmses of 0.6, 1.3, and 1.1, respectively. We also compared TIP3P with GAFF parameters to the calculations of Wolf and Groenhof. Although we find similar rmse and correlation values for our calculations (rmse 1.5 compared to 1.6; correlation coefficient 0.97 compared to 0.99), our approach does not reproduce their values exactly.

In Figure 3, the computed partition coefficients between water and chloroform are plotted against the experimentally
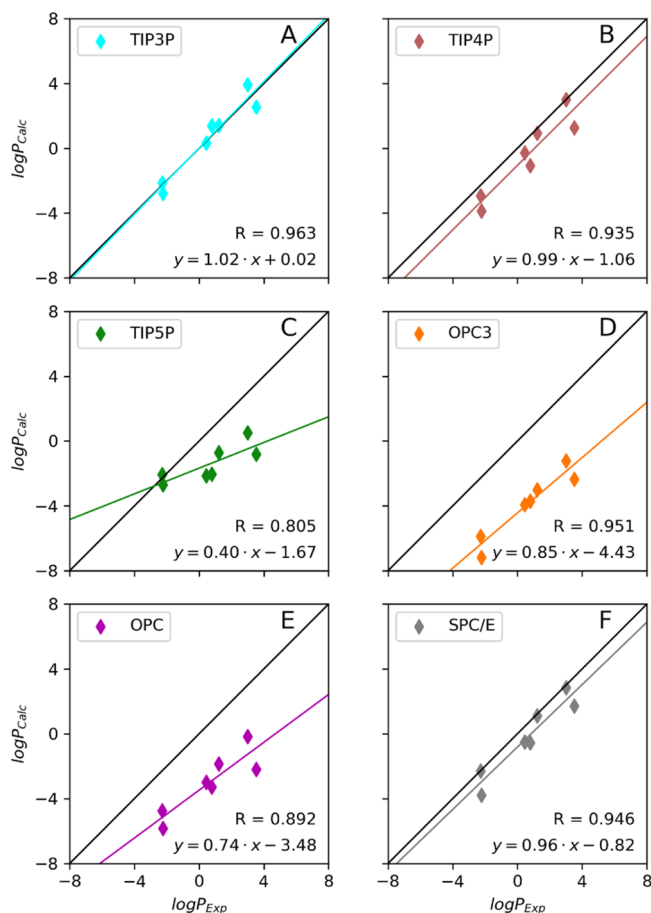

Figure 3. Difference of the solvation free energy of chloroform for the different water models: (A) TIP3P (cyan), (B) TIP4P (red), (C) TIP5P (green), (D) OPC3 (orange), (E) OPC (magenta), and (F) SPC/E (gray). The identity line between experimental and calculated values is depicted as a black line in all subplots.

determined ones. Therefore, the ideal values would be arranged along the diagonal; this is most true for TIP3P, which shows a slope of the linear fit of 1.02 and an axis intercept at 0.02. For the TIP4P and SPC/E water models, the slope is also very close to 1 (TIP4P: 0.99; SPC/E: 0.96), but the axis intercept is significantly farther away (TIP4P: −1.06; SPC/E: −0.82). Interestingly, in the linear fit for the OPC-like water models, the slope is still within an acceptable range (OPC3: 0.85; OPC: 0.74). However, the axis intercept differs stronger from zero than for the water models mentioned above (OPC3: −4.43; OPC: −3.48). Finally, the worst fit is shown by TIP5P, which has a slope of 0.40 and an axis intercept of −1.67. Interestingly, the linear fit on the TI data of Wolf and Groenhof is much steeper than the calculations in this study with a slope of 1.6 and an axis intercept at 0.51 (Supporting Information Figure S4).

Localized Thermodynamic Properties. An insightful feature of GIST is the ability to localize the thermodynamic properties and show where chloroform has more favorable interactions than water and vice versa. These localized data can be visualized as a grid around the molecules (free energy in Figure 4). Here, atoms that would generally be considered

hydrophilic are shown to prefer water as a solvent, whereas hydrophobic atoms prefer the chloroform phase more strongly.
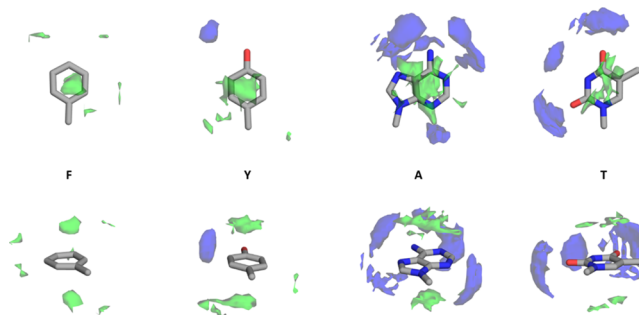

Figure 4. Difference in solvation free energy between water and chloroform. Blue regions are favorable water interaction sites (<-0.05 kcal/mol/Å^3) and green regions are favorable chloroform interaction sites (>0.05 kcal/mol/Å^3). (**F**) toluene, (**Y**) p-cresol, (**A**) 9-methyl-adenine, and (**T**) 1-methyl-thymine.

For **F**, chloroform is preferred at all positions around the molecule. Especially at the pi cloud above the ring, chloroform is much more favorable than water. **Y** shows a very similar behavior. The pi cloud is again strongly favored by chloroform. However, a strong interaction with water can be found at the hydrogen bond position of the hydroxyl group. Note that the molecule is constrained to a single conformation for the GIST calculation; thus, the strong interaction with the water molecules is only present at one side of the oxygen atom. Interestingly, this interaction is so strong that it almost outweighs the two interactions with the pi cloud, which heavily favor chloroform. However, despite its strength, the hydrogen bond is limited to a singular interaction site, while the interaction with the pi cloud is possible above and below the ring. Together with the remaining (rather apolar) parts of the molecule, these comparably weaker interactions in sum lead to favorization of chloroform over water.

The other two compounds in Figure 4, that is, **A** and **T**, favor water over chloroform. Here again, the solvation free energy at the pi cloud positions is less favorable than for chloroform. Interestingly, the regions in the ring plane show a very high preference for water. For these two molecules, this is mostly due to the multiple hydrophilic groups and possible hydrogen bonding sites. For **A**, the most favorable interactions occur around the nitrogen atoms, and for **T**, the strongest interaction is observed in the space surrounding the nitrogen atom and encapsulated by the two oxygen atoms. Additionally, the two oxygen atoms also show strongly favorable interactions with water.

In addition, GIST allows the visualization of contributions from enthalpy (Figure 5 top) and entropy (Figure 5 bottom) to the free energy. Interestingly, all compounds in this study exhibit an energetic preference for water but show a strong entropic penalty (Supporting Information Tables S2, S3), which is much less pronounced in chloroform. Indicative for this are the large blue areas around each molecule in the top of Figure 5. In contrast, chloroform shows fewer regions that are entropically unfavorable, which is indicated by the larger green areas in the bottom of Figure 5. The entropically favored regions and the enthalpically favored regions are almost always complementary to each other, highlighting the effect of enthalpy−entropy compensation.
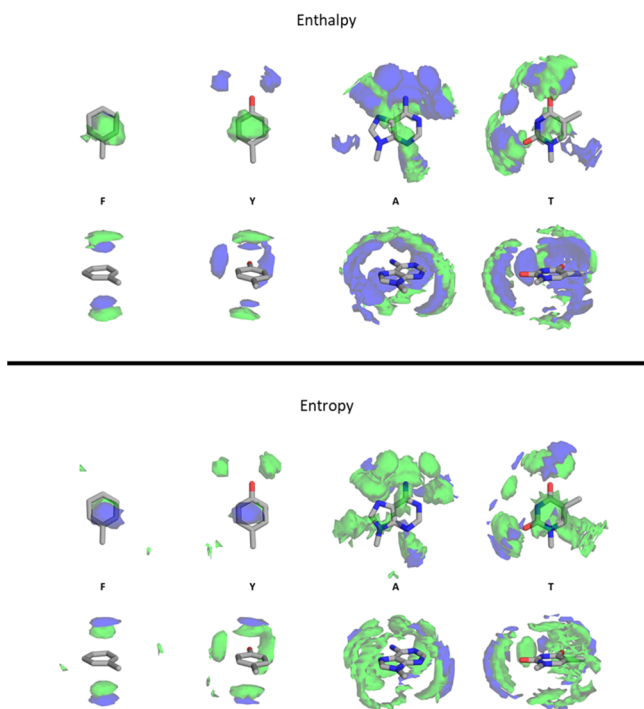
Figure 5. Difference in the enthalpic (top) and entropic (bottom) contributions to the solvation free energy between water and chloroform, visualized as a grid. Note that for the entropic contributions, $-T\Delta S$ is visualized. Blue regions are favorable water interaction sites (-0.05 kcal/mol/Å^3), and green regions are favorable chloroform interaction sites (0.05 kcal/mol/Å^3). (**F**) toluene, (**Y**) p-cresol, (**A**) 9-methyl-adenine, and (**T**) 1-methyl-thymine.

For **F**, both solvents show enthalpically favored regions on top and bottom of the ring. For chloroform, this region is farther away than for the water molecules, which is in concordance with the size of chloroform and the stronger electrostatic interactions with water. However, from the free energy plot, we see that water is never significantly more favorable than chloroform, indicating that either its entropy is lower or the enthalpic contributions are higher.

For **Y**, a similar picture can be observed, as there are again two favorable areas around the pi interaction sites. Again, these locations are farther away for chloroform and closer for water. However, in the free energy plot (Figure 4), these interactions are more favorable in chloroform. Also, at the hydrogen bonding site, strong enthalpic interactions are visible in water (top Figure 5), which come at the price of a significantly higher entropic loss (bottom Figure 5). These entropic penalties are not as significant as the enthalpic interactions in these spots, resulting in a more favorable free energy in water at the hydrogen bonding sites.

In **A**, we again see the larger distance for chloroform than for water. For water, the enthalpic contribution is more favorably close to the molecule, whereas chloroform has its interaction hot spots a little farther away from **A**. In general, the enthalpic interactions are pronounced very strongly close to the plane of the ring, where the interactions with the nitrogen atoms and the possible hydrogen bonds are visible. The most significant contribution in chloroform seems to be above the ring, where again interactions with the pi cloud are visible. In general, the surface for the favorable water interactions is much clearer and more pronounced, when compared to the more rugged surface in chloroform. The stronger enthalpic interactions persist even

into the free energy, making **A** more hydrophilic than the previous molecules.

Last, **T** shows three very strong enthalpic interaction positions. Two are located close to the oxygen atom and one is close to the exposed nitrogen atom. Like **A**, the favorable contributions in chloroform are farther away and the surface is again more uneven, indicating the higher enthalpic preference for water. Despite the unfavorable entropic interactions, resulting from the increased ordering of the water molecules around the hydrophilic atoms, the free energy stays favorable for water in almost all regions. Again, the position just above the ring remains the only exception.

■ DISCUSSION

A reliable prediction of log P values between water and chloroform is of high interest especially for biological systems, where it finds application in the estimation of membrane permeability. In a multitude of experimental and computational studies, the relation of membrane permeability and a compound's hydrophobicity has been thoroughly investigated.[4,76] While several of these studies, particularly computational investigations, already consider partition coefficients between water and chloroform, the typical experimental approach to determine a compound's hydrophobicity is its log P_{OW}, the partition coefficient between water and octanol. We have previously shown that a solute's hydrophobicity relates to GIST solvation free energies.[40,44] These water-based approaches to GIST however approximate the hydrophobicity of a compound as the difference in solvation free energies and vacuum. Here, we extend this description as the difference between solvation free energies in water and chloroform.

In a first step, we calculated the hydration free energy using different water models and the GAFF2 parameters. For all the investigated water models, that is, TIP3P, TIP4P, TIP5P, OPC, OPC3, and SPC/E, we find an extremely high correlation with the experimental partition coefficient, in concordance with our previous studies.[40,44] The outstanding correlation with the experimental partition coefficient between water and chloroform shows that for similar sized compounds, the hydration free energy already correlates remarkably well with the partition coefficient. Furthermore, a lowest correlation of 0.978 is observed for TIP5P. However, this correlation is still remarkably high. These findings are reassuring as they indicate that our previously postulated relation of hydrophobicity and GIST solvation free energies in proteins can also be extended to small molecule systems. We want to emphasize that as a proof of principle, we chose rather small molecules, which do not differ much in size and functionalities, for this study, as these compounds are ideally suited for the GIST calculations. However, we surmise that molecules with a similar chloroform/water partition coefficient, which differ strongly in size, would yield very different hydration free energies. This should simply occur because a larger molecule has more interactions with the water around it. However, the same reasoning would also apply to the solvation free energy in chloroform. Therefore, we hypothesize that the contributions of the apolar solvent are necessary to predict the partition coefficients of molecules of varying size and polarity.

When compared to the few available experimental hydration free energies, we can observe that TIP3P closely reproduces the experimental data, with an rmse of about 2.5 kcal/mol. The other water models perform worse in this regard, with rmses ranging up to almost 9 kcal/mol. This is interesting, as most

other water models are considered better than TIP3P in many aspects.[77] Additionally, the correlation coefficient of the other water models, with TIP5P being the lone exception, is higher than that of TIP3P. We surmise that one possible reason for this is that the GAFF parameters were optimized in the context of the TIP3P water model. The more advanced water models are parameterized to model pure water properties correctly, which does not necessarily increase the reliability of hydration free energy calculations.

The trend of solvation free energy in chloroform is rather similar to the trend found by Wolf and Groenhof[50] when considering the relative differences. Interestingly, all absolute values are about 2 kcal/mol off. This could be due to an incompletely converged nearest-neighbor estimate of the entropy, whose influence is enhanced by the additional voxels we used due to the large cutoff. However, we deemed it necessary to include the second solvation shell in this study, as it is very pronounced in chloroform. Another source for the difference in the solvation free energies between GIST and TI calculations could be the approximation when calculating the entropy term. In the current GIST implementations, only the two-body term of the entropy is considered. We also performed TI calculations on the TIP3P water model and chloroform, which are gathered in the Supporting Information. Our results show an almost constant offset in the solvation free energies. However, smaller fluctuations can still be observed, even when shifting the values by this offset.

In general, we observe the poorest correlation between solvation free energy and partition coefficient in chloroform. However, as this correlation is still above 0.9, we surmise that also for chloroform, the employed methodology works very well. A possible explanation for the decreased correlation of chloroform could be that the applied force field parameters describing the small molecules were derived with the aim to represent the molecule in water. The used RESP charges and GAFF/GAFF2 parameter combination are developed to account for self-polarization in water. However, self-polarization in chloroform is significantly smaller than in water, and therefore, the parameters are overestimating the interaction with chloroform. To get similar accuracy for chloroform, it might be necessary to refit parameters to account for the change in polarity.[78,79] Additionally, we tested various water models but used only one chloroform model. We want to note that other chloroform models might improve the results even further.

The calculation of the partition coefficients between water and chloroform not only preserved the high correlation (Figure 3) but also reproduced the experimental values almost numerically. This is prominent for the TIP3P water model, where an rmse of only 0.6 and a correlation of 0.96 were found. Furthermore, the calculated values align almost perfectly with the experimentally determined ones, which is visible in Figure 3 from the parameters of the linear fit. The slope of this linear fit is almost exactly 1, and the axis intercept is approximately 0. Other water models have similarly low rmse values, while also maintaining a high correlation, that is, TIP4P (rmse: 1.3, R: 0.94) and SPC/E (rmse: 1.1, R: 0.95). Additionally, the linear fit in Figure 3 still shows a slope of almost 1 for these water models, but they show a parallel displacement from the ideal diagonal. It is intriguing that the highest correlation, lowest rmse, and best linear fit are found for the TIP3P model, as TIP3P is nowadays generally considered to be a subpar water model.[77] All other models overestimate the likelihood to find

the compound in chloroform. As described above, this might be caused by an overpolarization of the solute in chloroform with GAFF2/RESP parameters. Hence, this systematic overpolarization in the chloroform phase is likely the reason for the parallel displacement of the differences in solvation free energies, which is visible from Figure 3. Another possible explanation to the effect could be favorable error compensation between the different solvent models, that is, chloroform and TIP3P. In addition, GIGIST and GIST treat the electrostatic interactions via the nearest image convention, which could lead to larger offsets in better water models.

Interestingly, the TI calculations did not reproduce the experimental reference values of the partition coefficient as closely as the GIST calculations. Because the TI calculations produce the exact solvation free energy for a given force field, the reason for this discrepancy might be due to a favorable error compensation between the approximation of GIST, using only the two body terms for the entropy, and the force field. This could hint toward an inaccurate description of the higher-order entropy terms of the solvent phase in the used force fields.

GIST allows for the localization and analysis of thermodynamic properties around the molecule of interest, which has been shown to be a valuable tool in analyzing water properties around biomolecules.[38,40,43,51] Here, we analyze differences in the local thermodynamic properties of the different compounds solvated in water and chloroform. One of the most interesting features found in this study is the preference of the pi cloud, above and below an aromatic moiety, for chloroform and not water. Water is preferred over chloroform on atoms, which are traditionally considered as hydrophilic, such as nitrogen and oxygen.

Furthermore, our method allows for the analysis of entropic and enthalpic contributions to the free energy. In all investigated compounds, we were able to clearly see that an enthalpically favorable interaction site is accompanied by a sizeable entropic penalty. This highlights the concept of enthalpy−entropy compensation, which states that a very favorable enthalpic interaction is accompanied by an unfavorable entropic penalty.[80] To achieve a favorable free energy, the enthalpic interaction must overcome the entropic cost. In our examples, this is the case for the hydrophilic atoms, such as nitrogen or oxygen. However, interactions that are also accompanied by a large entropic cost and that are not favorable are often more preferred in chloroform. One example is the pi cloud of the aromatic rings, where a weak interaction is always visible between water and the ring. The restricted space for the water molecules leads to a high entropic penalty, which is much smaller in chloroform, and therefore, chloroform is favored above the pi cloud. This is again visible for **F**, which is the most hydrophobic molecule in this set. It shows no significant favorable interactions in water, and interestingly, the area of the pi cloud is heavily favored in chloroform. This is in line with a previous work from our group,[47] which showed that the solvation of an aromatic moiety counteracts its ability to exhibit pi-stacking interactions, which is much more pronounced in heterocycles with multiple hydrophilic atoms.

Interestingly, in terms of enthalpy, all compounds would favor water over chloroform. This is somewhat surprising, as on the one hand, water has stronger interactions with the compound and therefore a higher energetic contribution of the solute−solvent interaction energy. On the other hand, a water molecule also has high interactions in the bulk, in the form of

hydrogen bonds, to other water molecules in its vicinity. However, entropically, all compounds favor chloroform over water. This can easily be explained by entropy—enthalpy compensation. While water is bound more strongly to hydrophilic groups, an entropic penalty for this strong interaction occurs because of the introduced order. This leads to generally more favorable entropic contributions for chloroform, compared to water. However, even for the solvation in pure chloroform, an entropic penalty occurs, which is just weaker than the one in water.

The calculation speedup, which is provided by the GPU-accelerated version of GIST, facilitates highly efficient analysis of trajectories, even for large molecular systems. We already showed the speedup of these calculations for water boxes of increasing size.[40] The bottleneck for the calculations of this study is the simulation preceding GIST analysis, which—depending on the size of the solute and the solvent box—can take multiple hours to complete. The analysis itself can be performed within minutes for a single molecule. Hence, while GIST is probably not the fastest approach to estimate partition coefficients of small molecules, it represents a reliable method which can also be applied to large biomolecular systems.

CONCLUSIONS

We showed that our new GIGIST implementation can effectively compute thermodynamic properties of water and chloroform as input solvents. Additionally, we evaluate the reliability of this new feature in GIGIST by calculating the differences in solvation free energies and thus partition coefficients between water and chloroform.

In summary, we were able to calculate the partition coefficients between chloroform and water with a surprising accuracy, even more so as the TI calculations do not reproduce these values closely. We surmise that a favorable error compensation between the two-body approximation and the force field plays a vital role. Interestingly, the TIP3P water model performs best in this analysis. This might be due to the design of force fields, which are usually developed in the context of the TIP3P water model.

Furthermore, the presented algorithm can localize the major contributors to the difference in solvation free energies. We use this advantage to highlight positions around the molecules where chloroform is largely favored over water, which is in all cases at the position of the pi clouds. Additionally, we found that the favorable positions for water are mostly close to hydrophilic groups. Not surprisingly, we found that enthalpically favored positions around the molecule come at an entropic cost. In general, water shows stronger enthalpic interactions with the solutes but also a higher entropic penalty.

The GIGIST implementation of GIST allows for localized and global calculations of solvation free energies in water and chloroform. We surmise that the presented algorithm and method provide a robust framework for the development of advanced tools to predict membrane permeabilities or biomolecular hydrophobicities.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c00289.

Information on the reference data that were not available in the AMBER manual, entropic and enthalpic

contribution to the solvation free energy, RDFs for all compounds and water models and chloroform, and grid visualizations for the remaining four compounds (PDF)

AUTHOR INFORMATION

Corresponding Author

Klaus R. Liedl — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria;* ◉ orcid.org/0000-0002-0985-2299; Email: klaus.liedl@uibk.ac.at

Authors

Johannes Kraml — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria*

Florian Hofer — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria*

Anna S. Kamenik — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria*

Franz Waibl — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria*

Ursula Kahler — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria*

Michael Schauperl — *Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, A-6020 Innsbruck, Austria*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.0c00289

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.
GIGIST is available free of charge from the github page of the Liedl Lab (https://github.com/liedllab/gigist.git).

ABBREVIATIONS

GIST, grid inhomogeneous solvation theory; IST, inhomogeneous solvation theory

REFERENCES

(1) Lipinski, C. A. Lead- and Drug-Like Compounds: The Rule-of-Five Revolution. *Drug Discovery Today: Technol.* **2004**, *1*, 337−341.

(2) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and Computational Approaches to Estimate Solubility and Permeability in Drug Discovery and Development Settings. *Adv. Drug Delivery Rev.* **1997**, *23*, 3−25.

(3) Sugano, K.; Kansy, M.; Artursson, P.; Avdeef, A.; Bendels, S.; Di, L.; Ecker, G. F.; Faller, B.; Fischer, H.; Gerebtzoff, G.; Lennernaes, H.; Senner, F. Coexistence of Passive and Carrier-Mediated Processes in Drug Transport. *Nat. Rev. Drug Discovery* **2010**, *9*, 597−614.

(4) Liu, X.; Testa, B.; Fahr, A. Lipophilicity and Its Relationship with Passive Drug Permeation. *Pharm. Res.* **2011**, *28*, 962−977.

(5) Rawitch, A. B.; Rohrer, R.; Vardaris, R. M. Delta-9-Tetrahydrocannabinol Uptake by Adipose Tissue: Preferential Accumulation in Gonadal Fat Organs. *Gen. Pharmacol.* **1979**, *10*, 525−529.

(6) Garrett, E. R.; Hunt, C. A. Physicochemical Properties, Solubility, and Protein Binding of Δ9 -Tetrahydrocannabinol. *J. Pharm. Sci.* **1974**, *63*, 1056−1064.

(7) Gunasekaran, N.; Long, L.; Dawson, B.; Hansen, G.; Richardson, D.; Li, K.; Arnold, J.; McGregor, I. Reintoxication: the release of fat-stored Δ9-tetrahydrocannabinol (THC) into blood is enhanced by food deprivation or ACTH exposure. *Br. J. Pharmacol.* **2009**, *158*, 1330−1337.

(8) Young, E. Triggering a Second 'High'. *New Sci.* **2009**, *203*, 13.

(9) Goto, A.; Moriya, Y.; Mandai, T.; Wakabayashi, T.; Tsukamoto, T.; Tagawa, Y.; Kondo, T.; Asahi, S. Disposition of the Highly Fat Distributed Compound 1-(4-Methoxyphenyl)-4-(2,2,4,6,7-Pentamethyl -2,3-Dihydro-1-Benzofuran-5-yl)Piperazine (TAK-357) in Rats and Dogs. *Drug Res.* **2016**, *67*, 38−45.

(10) Jandacek, R. J.; Tso, P. Factors Affecting the Storage and Excretion of Toxic Lipophilic Xenobiotics. *Lipids* **2001**, *36*, 1289−1305.

(11) Hansch, C.; Björkroth, J. P.; Leo, A. Hydrophobicity and Central Nervous System Agents: On the Principle of Minimal Hydrophobicity in Drug Design. *J. Pharm. Sci.* **1987**, *76*, 663−687.

(12) Leo, A.; Hansch, C.; Elkins, D. Partition Coefficients and Their Uses. *Chem. Rev.* **1971**, *71*, 525−616.

(13) Işık, M.; Levorse, D.; Mobley, D. L.; Rhodes, T.; Chodera, J. D. Octanol−Water Partition Coefficient Measurements for the SAMPL6 Blind Prediction Challenge. *J. Comput.-Aided Mol. Des.* **2020**, *34*, 405−420.

(14) Renner, R. The KOW Controversy. *Environ. Sci. Technol.* **2002**, *36*, 410A−413A.

(15) Low, Y. W.; Blasco, F.; Vachaspati, P. Optimised Method to Estimate Octanol Water Distribution Coefficient (logD) in a High Throughput Format. *Eur. J. Pharm. Sci.* **2016**, *92*, 110−116.

(16) Riniker, S. Molecular Dynamics Fingerprints (MDFP): Machine Learning from MD Data To Predict Free-Energy Differences. *J. Chem. Inf. Model.* **2017**, *57*, 726−741.

(17) Safder, U.; Nam, K.; Kim, D.; Shahlaei, M.; Yoo, C. Quantitative Structure-Property Relationship (QSPR) Models for Predicting the Physicochemical Properties of Polychlorinated Biphenyls (PCBs) Using Deep Belief Network. *Ecotoxicol. Environ. Saf.* **2018**, *162*, 17−28.

(18) Kim, T.; Park, H. Computational prediction of octanol-water partition coefficient based on the extended solvent-contact model. *J. Mol. Graphics Modell.* **2015**, *60*, 108−117.

(19) Toropov, A. A.; Toropova, A. P.; Cappelli, C. I.; Benfenati, E. CORAL: Model for Octanol/Water Partition Coefficient. *Fluid Phase Equilib.* **2015**, *397*, 44−49.

(20) Liao, Q.; Yao, J.; Yuan, S. SVM Approach for Predicting LogP. *Mol. Diversity* **2006**, *10*, 301−309.

(21) Ghose, A. K.; Crippen, G. M. Atomic Physicochemical Parameters for Three-Dimensional Structure-Directed Quantitative Structure-Activity Relationships I. Partition Coefficients as a Measure of Hydrophobicity. *J. Comput. Chem.* **1986**, *7*, 565−577.

(22) Cheng, T.; Zhao, Y.; Li, X.; Lin, F.; Xu, Y.; Zhang, X.; Li, Y.; Wang, R.; Lai, L. Computation of Octanol−Water Partition Coefficients by Guiding an Additive Model with Knowledge. *J. Chem. Inf. Model.* **2007**, *47*, 2140−2148.

(23) Moriguchi, I.; Hirono, S.; Liu, Q.; Nakagome, I.; Matsushita, Y. Simple Method of Calculating Octanol/Water Partition Coefficient. *Chem. Pharm. Bull.* **1992**, *40*, 127−130.

(24) Molnár, L.; Keserű, G. M.; Papp, Á.; Gulyás, Z.; Darvas, F. A Neural Network Based Prediction of Octanol−Water Partition Coefficients Using Atomic5 Fragmental Descriptors. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 851−853.

(25) Leo, A. J. [25] Hydrophobic Parameter: Measurement and Calculation. *Methods in Enzymology*; Academic Press, 1991; Vol. *202*, pp 544−591.

(26) Wang, R.; Fu, Y.; Lai, L. A New Atom-Additive Method for Calculating Partition Coefficients. *J. Chem. Inf. Comput. Sci.* **1997**, *37*, 615−621.

(27) Genheden, S. Predicting Partition Coefficients with a Simple All-Atom/Coarse-Grained Hybrid Model. *J. Chem. Theory Comput.* **2016**, *12*, 297−304.

(28) Huang, W.; Blinov, N.; Kovalenko, A. Octanol−Water Partition Coefficient from 3D-RISM-KH Molecular Theory of Solvation with Partial Molar Volume Correction. *J. Phys. Chem. B* **2015**, *119*, 5588−5597.

(29) Michalík, M.; Poliak, P.; Klein, E.; Lukeš, V. On the Toxicity of Para-Substituted Phenols and Their Quinone Metabolites: Quantum Chemical Study. *Chem. Phys. Lett.* **2018**, *709*, 71−76.

(30) Klopman, G.; Iroff, L. D. Calculation of Partition Coefficients by the Charge Density Method. *J. Comput. Chem.* **1981**, *2*, 157−160.

(31) Ono, S.; Naylor, M. R.; Townsend, C. E.; Okumura, C.; Okada, O.; Lokey, R. S. Conformation and Permeability: Cyclic Hexapeptide Diastereomers. *J. Chem. Inf. Model.* **2019**, *59*, 2952−2963.

(32) Leo, A. J. Calculating log Poct from Structures. *Chem. Rev.* **1993**, *93*, 1281−1306.

(33) Ghose, A. K.; Crippen, G. M. Atomic Physicochemical Parameters for Three-Dimensional-Structure-Directed Quantitative Structure-Activity Relationships. 2. Modeling Dispersive and Hydrophobic Interactions. *J. Chem. Inf. Comput. Sci.* **1987**, *27*, 21−35.

(34) Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. Atomic Physicochemical Parameters for Three Dimensional Structure Directed Quantitative Structure-Activity Relationships. 4. Additional Parameters for Hydrophobic and Dispersive Interactions and Their Application for an Automated Superposition of Certain Naturally Occurring Nucleoside Antibiotics. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163−172.

(35) Kovalenko, A.; Hirata, F. Self-Consistent Description of a Metal−Water Interface by the Kohn−Sham Density Functional Theory and the Three-Dimensional Reference Interaction Site Model. *J. Chem. Phys.* **1999**, *110*, 10095−10112.

(36) Kovalenko, A.; Hirata, F. Three-Dimensional Density Profiles of Water in Contact with a Solute of Arbitrary Shape: A RISM Approach. *Chem. Phys. Lett.* **1998**, *290*, 237−244.

(37) Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T. Thermodynamics of Water in an Enzyme Active Site: Grid-Based Hydration Analysis of Coagulation Factor Xa. *J. Chem. Theory Comput.* **2014**, *10*, 2769−2780.

(38) Nguyen, C. N.; Kurtzman Young, T.; Gilson, M. K. Grid Inhomogeneous Solvation Theory: Hydration Structure and Thermodynamics of the Miniature Receptor Cucurbit[7]uril. *J. Chem. Phys.* **2012**, *137*, 044101.

(39) Ramsey, S.; Nguyen, C.; Salomon-Ferrer, R.; Walker, R. C.; Gilson, M. K.; Kurtzman, T. Solvation Thermodynamic Mapping of Molecular Surfaces in AmberTools: GIST. *J. Comput. Chem.* **2016**, *37*, 2029−2037.

(40) Kraml, J.; Kamenik, A. S.; Waibl, F.; Schauperl, M.; Liedl, K. R. Solvation Free Energy as a Measure of Hydrophobicity: Application to Serine Protease Binding Interfaces. *J. Chem. Theory Comput.* **2019**, *15*, 5872−5882.

(41) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.

(42) Case, D.; Ben-Shalom, I.; Brozell, S.; Cerutti, D.; Cheatham, T., III; Cruzeiro, V.; Darden, T.; Duke, R.; Ghoreishi, D.; Gilson, M. *AMBER 2018*, 2018.

(43) Young, T.; Abel, R.; Kim, B.; Berne, B. J.; Friesner, R. A. Motifs for Molecular Recognition Exploiting Hydrophobic Enclosure in Protein−Ligand Binding. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 808.

(44) Schauperl, M.; Podewitz, M.; Waldner, B. J.; Liedl, K. R. Enthalpic and Entropic Contributions to Hydrophobicity. *J. Chem. Theory Comput.* **2016**, *12*, 4600−4610.

(45) Hüfner-Wulsdorf, T.; Klebe, G. Protein-Ligand Complex Solvation Thermodynamics: Development, Parameterization and Testing of GIST-based Solvent Functionals. *J. Chem. Inf. Model.* **2020**, *60*, 1409−1423.

(46) Uehara, S.; Tanaka, S. AutoDock-GIST: Incorporating Thermodynamics of Active-Site Water into Scoring Function for Accurate Protein-Ligand Docking. *Molecules* **2016**, *21*, 1604.

(47) Loeffler, J. R.; Schauperl, M.; Liedl, K. R. Hydration of Aromatic Heterocycles as an Adversary of π-Stacking. *J. Chem. Inf. Model.* **2019**, *59*, 4209−4219.

(48) Witek, J.; Keller, B. G.; Blatter, M.; Meissner, A.; Wagner, T.; Riniker, S. Kinetic Models of Cyclosporin A in Polar and Apolar Environments Reveal Multiple Congruent Conformational States. *J. Chem. Inf. Model.* **2016**, *56*, 1547−1562.

(49) Bockus, A. T.; Lexa, K. W.; Pye, C. R.; Kalgutkar, A. S.; Gardner, J. W.; Hund, K. C. R.; Hewitt, W. M.; Schwochert, J. A.; Glassey, E.; Price, D. A.; Mathiowetz, A. M.; Liras, S.; Jacobson, M. P.; Lokey, R. S. Probing the Physicochemical Boundaries of Cell Permeability and Oral Bioavailability in Lipophilic Macrocycles Inspired by Natural Products. *J. Med. Chem.* **2015**, *58*, 4581−4589.

(50) Wolf, M. G.; Groenhof, G. Evaluating Nonpolarizable Nucleic Acid Force Fields: A Systematic Comparison of the Nucleobases Hydration Free Energies and Chloroform-to-Water Partition Coefficients. *J. Comput. Chem.* **2012**, *33*, 2225−2232.

(51) Nguyen, C.; Gilson, M. K.; Young, T. Structure and Thermodynamics of Molecular Hydration via Grid Inhomogeneous Solvation Theory. **2011**, arXiv:1108.4876. arXiv preprint.

(52) Lazaridis, T. Inhomogeneous Fluid Approach to Solvation Thermodynamics. 1. Theory. *J. Phys. Chem. B* **1998**, *102*, 3531−3541.

(53) Lazaridis, T. Inhomogeneous Fluid Approach to Solvation Thermodynamics. 2. Applications to Simple Fluids. *J. Phys. Chem. B* **1998**, *102*, 3542−3550.

(54) Singh, H.; Misra, N.; Hnizdo, V.; Fedorowicz, A.; Demchuk, E. Nearest Neighbor Estimates of Entropy. *Am. J. Math. Manage. Sci.* **2003**, *23*, 301−321.

(55) Huynh, D. Q. Metrics for 3D Rotations: Comparison and Analysis. *J. Math. Imaging Vis.* **2009**, *35*, 155−164.

(56) *MOE (Molecular Operating Environment), 2019.1*; Chemical Computing Group (CCG), 2019.

(57) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General AMBER Force Field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(58) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Petersson, G.; Nakatsuji, H. *Gaussian 16*; Gaussian, Inc.: Wallingford, CT, 2016.

(59) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. A Well-Behaved Electrostatic Potential Based Method using Charge Restraints for Deriving Atomic Charges: the RESP Model. *J. Phys. Chem.* **1993**, *97*, 10269−10280.

(60) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.

(61) Mahoney, M. W.; Jorgensen, W. L. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112*, 8910−8922.

(62) Izadi, S.; Onufriev, A. V. Accuracy Limit of Rigid 3-Point Water Models. *J. Chem. Phys.* **2016**, *145*, 074501.

(63) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863−3871.

(64) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269−6271.

(65) Fox, T.; Kollman, P. A. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B* **1998**, *102*, 8070−8079.

(66) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.

(67) Adelman, S. A.; Doll, J. D. Generalized Langevin equation approach for atom/solid-surface scattering: General formulation for classical scattering off harmonic solids. *J. Chem. Phys.* **1976**, *64*, 2375−2388.

(68) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(69) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(70) Bannan, C. C.; Calabró, G.; Kyu, D. Y.; Mobley, D. L. Calculating Partition Coefficients of Small Molecules in Octanol/Water and Cyclohexane/Water. *J. Chem. Theory Comput.* **2016**, *12*, 4015−4024.

(71) DeLano, W. L. *PyMOL*; Schrödinger, Inc., 2002.

(72) Ferguson, D. M.; Pearlman, D. A.; Swope, W. C.; Kollman, P. A. Free Energy Perturbation Calculations Involving Potential Function Changes. *J. Comput. Chem.* **1992**, *13*, 362−370.

(73) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. Affinities of Amino Acid Side Chains for Solvent Water. *Biochemistry* **1981**, *20*, 849−855.

(74) Cullis, P. M.; Wolfenden, R. Affinities of Nucleic Acid Bases for Solvent Water. *Biochemistry* **1981**, *20*, 3024−3028.

(75) Radzicka, A.; Wolfenden, R. Comparing the Polarities of the Amino Acids: Side-Chain Distribution Coefficients between the Vapor Phase, Cyclohexane, 1-Octanol, and Neutral Aqueous Solution. *Biochemistry* **1988**, *27*, 1664−1670.

(76) Parisio, G.; Stocchero, M.; Ferrarini, A. Passive Membrane Permeability: Beyond the Standard Solubility-Diffusion Model. *J. Chem. Theory Comput.* **2013**, *9*, 5236−5246.

(77) Vega, C.; Abascal, J. L. F. Simulating Water with Rigid Non-Polarizable Models: A General Perspective. *Phys. Chem. Chem. Phys.* **2011**, *13*, 19663−19688.

(78) Schauperl, M.; Nerenberg, P. S.; Jang, H.; Wang, L.-P.; Bayly, C. I.; Mobley, D. L.; Gilson, M. K. Non-bonded force field model with advanced restrained electrostatic potential charges (RESP2). *Commun. Chem.* **2020**, *3*, 44.

(79) Zhou, A.; Schauperl, M.; Nerenberg, P. S. Benchmarking Electronic Structure Methods for Accurate Fixed-Charge Electrostatic Models. *J. Chem. Inf. Model.* **2020**, *60*, 249−258.

(80) Chodera, J. D.; Mobley, D. L. Entropy-Enthalpy Compensation: Role and Ramifications in Biomolecular Ligand Recognition and Design. *Annu. Rev. Biophys.* **2013**, *42*, 121−142.