

Hydration of Aromatic Heterocycles as an Adversary of π -Stacking

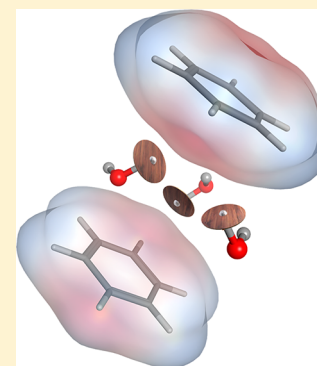
Johannes R. Loeffler,[†] Michael Schauerperl,^{*,‡} and Klaus R. Liedl[†]

[†]Institute of General, Inorganic and Theoretical Chemistry, and Center of Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innrain 80-82, A-6020 Innsbruck, Tyrol, Austria

[‡]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92039-0736, United States

Supporting Information

ABSTRACT: Hydration is one of the key players in the protein–ligand binding process. It not only influences the binding process *per se*, but also the drug's absorption, distribution, metabolism, and excretion properties. To gain insights into the hydration of aromatic cores, the solvation thermodynamics of 40 aromatic mono- and bicyclic systems, frequently occurring in medicinal chemistry, are investigated. Thermodynamics is analyzed with two different methods: grid inhomogeneous solvation theory (GIST) and thermodynamic integration (TI). Our results agree well with previously published experimental and computational solvation free energies. The influence of adding heteroatoms to aromatic systems and how the position of these heteroatoms impacts the compound's interactions with water is studied. The solvation free energies of these heteroaromatics are highly correlated to their gas phase interaction energies with benzene: compounds showing a high interaction energy also have a high solvation free energy value. Therefore, replacing a compound with one having a higher gas phase interaction energy might not result in the expected improvement in affinity. The desolvation costs counteract the higher stacking interactions, hence weakening or even inverting the expected gain in binding free energy.



INTRODUCTION

Stacking interactions between π -systems play a key role in a variety of materials and applications in chemistry, physics, and biology. These interactions effect polymers, organic synthesis, liquid crystals, DNA, RNA, and protein structures, as well as protein–ligand interactions. They include π - π ,¹ cation- π ,^{2,3} halogen- π ,^{4,5} and hydrogen bonding interactions via heteroatoms.⁶

Although aromatic stacking is omnipresent, a complete picture of its origin and effects is still elusive. A step toward a concise understanding of this phenomenon was taken by investigating stacking interactions with quantum chemical methods, performed on a selection of theory levels. However, most theoretical studies are done in the gas phase, completely neglecting the solvent and its effects.^{7–11} Studies including the solvation of aromatic compounds are sparse. Still, if solvation is in fact taken into account, usually substantial approximations are made. Recently, the group of Kim studied the solvation thermodynamics of two stacked benzene rings thoroughly, using a quantum/molecular mechanic (QM/MM) approach.¹² Stroet et al. performed a detailed study of solvation free enthalpies using thermodynamic integration.¹³ Therefore, they used a fragment library consisting of compounds frequently occurring in drug design projects. Additionally, the contribution of solvation to the binding energy between two aromatic rings might only play a minor role, if only apolar solvents are considered. This does not hold true for the most frequently occurring solvent in nature: water.

Only for very few compounds the effects of solvation and desolvation in context of aromatic π -stacking have been analyzed. QM investigations of benzene–water clusters ($n_{\text{Water}} = 1–10$) revealed that benzene can show hydrogen bond donor- and acceptor-like properties.^{14,15} Furthermore, the shape and interactions of the hydration shell of benzene and cyclohexane in water have been thoroughly explored using additive force fields (FFs).^{16,17} Stacking interactions in solvents of different polarity have presented the model of molecular balances for aromatic interactions. These studies performed and analyzed energy decomposition of aromatic stacking in different solvents and found the intramolecular component to correlate best with computational results.^{18–22} Additionally, Mobley and Guthrie made a large number of computationally and experimentally obtained solvation free energy values for a variety of compounds (including aromatics) easily accessible.²³ The consequences of hydration for aromatic interactions have, however, not been investigated in detail. Southall and Dill²⁴ found a strong correlation between hydration thermodynamics and the shape of the solute. For small, apolar solutes, the typical iceberg model holds true:²⁵ the solvent molecules are forming an ordered cage around the solute, retaining their internal hydrogen bonds. This ordering of the water molecules around the solute results in a loss of entropy.²⁴ In contrast, for larger molecules, typically water–water hydrogens bonds are broken, causing an enthalpic effect. Therefore, to fully describe

Received: May 10, 2019

Published: September 30, 2019

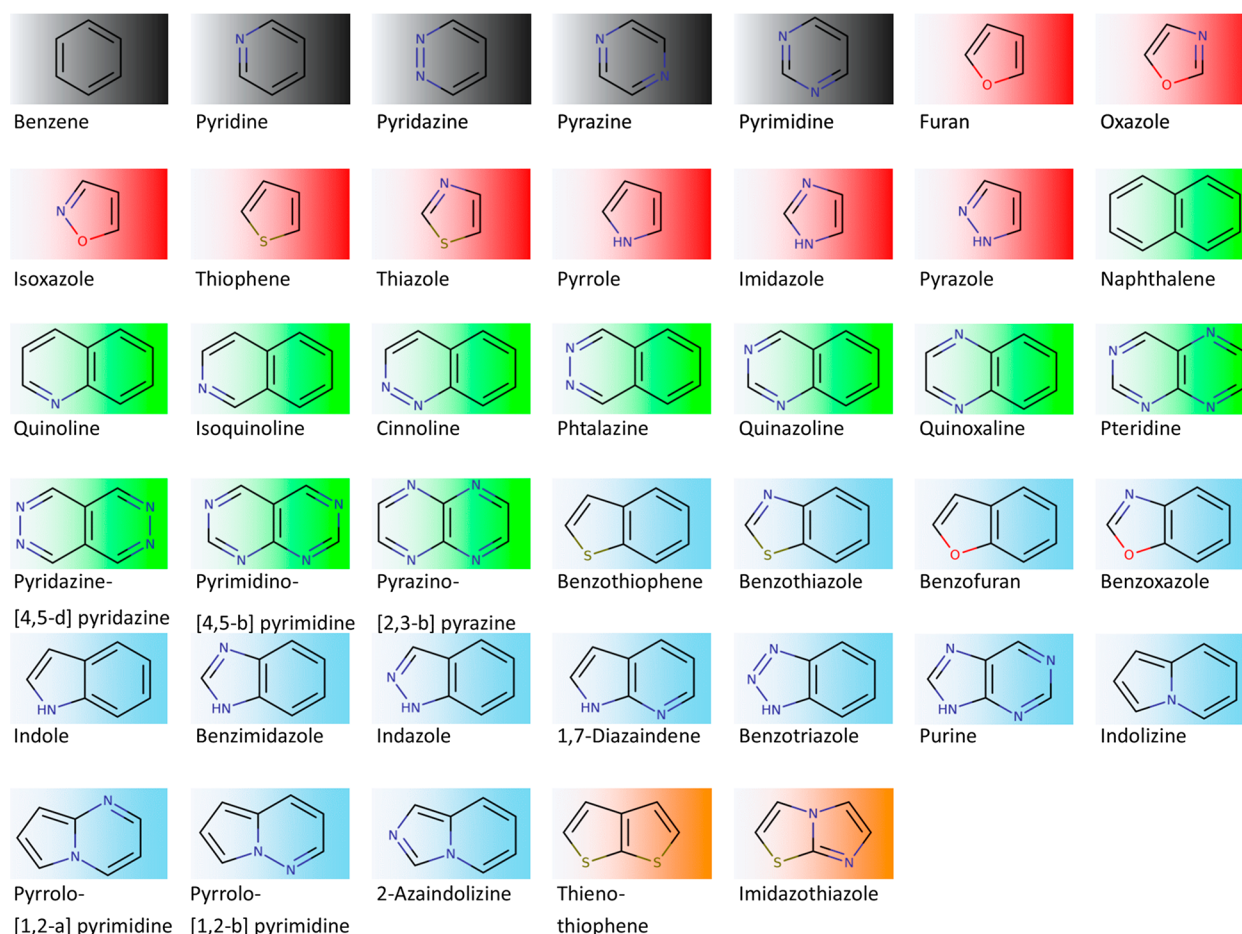


Figure 1. Investigated aromatic compounds grouped and color-coded according to their scaffold and heteroatoms. Groups 1 (black) and 2 (red) include aromatic monocycles of six- and five-membered rings, respectively. Group 3 (green) holds two fused six-membered aromatic rings. Group 4 (blue) consists of fused five- and six-membered rings. Group 5 (orange) represents two fused five-membered rings.

the solvation process, both thermodynamic contributions have to be taken into account.^{26–28}

The importance of solvation effects for protein–ligand binding is mirrored by a steady increase of publications investigating this topic.^{29–32} This is accompanied by a steadily growing number of tools to predict the hydration of proteins and their active sites—both with and without the ligand present, e.g. SZMAP,³³ WaterMap,^{34,35} or GIST.^{36,37} In contrast to our study most of the publications focus solely on the hydration thermodynamics of the protein binding site or the protein ligand complex.^{26,31,38,39} Analyses from a ligand-based view are mostly omitted, even though the desolvation of the ligand is a crucial factor also in the binding process. The solvation properties of a biologically active compound contribute to the binding behavior of a compound to a target. In short, the compound must strip its hydration shell and compete with water molecules in the respective binding pocket. Thereby, not only is the interaction *per se* influenced by the hydration, but the drug's absorption, distribution, metabolism, and excretion (ADME) properties are also influenced by its hydrophilic/hydrophobic character.^{40,41}

Aromatic core fragments frequently occur in drugs, due to their favorable interactions with proteins, including π – π stacking, cation– π stacking, and hydrogen bonds of heteroatomics. Additionally, a large variety is synthetically accessible, and stereoselective modifications of substituents are eagerly discussed in the literature.^{9,10,42,43} In previous

studies by our group, the interactions of aromatic heterocycles with benzene were investigated.¹ The benzene molecule mimics an aromatic amino acid side chain in the protein's active site, allowing investigation of the optimal interactions between the compounds and an aromatic side chain. Nevertheless, the protein–ligand interaction is not the only thing to be optimized, in order to increase binding affinity. The desolvation of the compounds plays an equally important role in this process and is a crucial step of the binding process. Therefore, this study aims to highlight the importance to include solvation effects in the calculation of aromatic binding energy and that neglecting them can have substantial effects. While the property of interest would be the binding free energy in solution, current limitations in computational power make these detail inaccessible at the necessary level of theory. Therefore, we use a two-step approach, obtaining stacking energy from quantum chemical calculations and the desolvation energy from a force field simulation. For clarification, we use the term “stacking energy” for gas phase stacking; “desolvation energy” is used to describe the contribution between solutes and solvent; and “binding energy” includes both contributions.

To the best of our knowledge, it was never examined in detail, how changing the aromatic core changes a compound's solvation behavior and therefore influences its binding free energy. This study closes this gap by investigating 40 heterocycles, frequently occurring in drug design projects.

METHODS

Molecules for this study were selected to represent frequently used scaffolds in medicinal chemistry.⁴² In addition, compounds were selected, which allow us to draw conclusions about the influence of slightly different substitution patterns, e.g., substituting one heteroatom with another, or moving the heteroatom to a different position within the aromatic system. The investigated compounds are shown in Figure 1. Grouping, indicated by color, was chosen as follows: six-membered rings with nitrogen heteroatoms; five-membered rings containing either oxygen, and/or (a/o) sulfur, a/o nitrogen heteroatoms. Additionally, bicyclic compounds were grouped in a similar fashion: two fused six-membered rings containing nitrogen heteroatoms at different positions; fused five- and six-membered rings containing either oxygen, a/o sulfur, a/o nitrogen as heteroatoms; two fused five-membered rings.

Simulation Details. The aromatic compounds were generated from SMILES strings using MOE.⁴⁴ From the resulting files, parametrizations were performed using antechamber.⁴⁵ The van der Waals, bond-length, angle, and torsion force field parameters for the aromatic compounds were taken from the general AMBER force field (GAFF).⁴⁶ After minimization of the molecules, the partial charges were derived using the restrained electrostatic potential (RESP)⁴⁷ method based on a Gaussian09⁴⁸ ESP calculation at the HF/6-31G* theory level. AMBER topology and coordinate files were generated using AmberTools.⁴⁹ All molecules were solvated in a cubic water box of TIP3P water with a minimal wall distance of 10 Å.⁵⁰ Equilibration of the molecules followed the protocol published in ref 51. The production runs were 10 ns, NpT simulations with a time step of 2 fs. Restraints were applied to the ligand with a weight of 1000 $\frac{\text{kcal}}{\text{mol}\cdot\text{\AA}}$. The spatial configuration of the simulation system was saved every 1000 steps resulting in 5000 frames, which were further processed with GIST. To ensure NpT conditions we applied a Berendsen barostat with a pressure relaxation time of 2.0 ps and Langevin dynamics with a friction coefficient of 2.0 ps⁻¹ was used to keep the temperature constant. To ensure the combination of the used barostat and positional restraints does not alter the results significantly, additional simulations in NVT ensemble were performed for selected molecules, Table S7.

Grid Inhomogeneous Solvation Theory. GIST allows a thermodynamic analysis of water molecules based on a molecular dynamics (MD) trajectory using a grid-based approach. GIST uses a grid to replace the spatial integrals from IST with discrete sums over the voxels on the constructed grid.^{36,37} One shortcoming of this method is that the solute has to be restrained to one conformation for each simulation. However, our systems are rigid rings and therefore restraining the solute to only one conformation is no limitation. Recently published studies highlighted that GIST is a valid approach to study biomolecular systems in combination with MD simulations, allowing to differentiate between a reference state, i.e., pure water, and an investigated state. A detailed description on the calculation of entropic and enthalpic terms is given by Nguyen et al.³⁶ and our previous and further recent publications.^{36,37,52,53}

The free energy of solvation of a solute ΔG_{Solv} can be approximated according to eq 1

$$\Delta G_{\text{Solv}} = -k_{\text{b}}T \ln \int e^{-\Delta G_{\text{Solv}}(q)/k_{\text{b}}T} p(q) dq \quad (1)$$

where k_{b} is the Boltzmann constant, T is the temperature, $\Delta G_{\text{Solv}}(q)$ is the solvation free energy of a solute in a single conformation q , and $p(q)$ is the probability of finding the solute in this conformation q . $\Delta G_{\text{Solv}}(q)$, for a particular conformation q , is calculated by restraining its coordinates, making the solute retain in its initial conformation. As mentioned, the compounds investigated in this work can be considered as rigid. Therefore, the error introduced by a single conformation approach is negligible. However, if our ligands would be more flexible, we could account for that by using multiple GIST calculations and combining them using eq 1. To estimate the error, we performed all simulations three times with a different seed, starting from the same geometry but with different initial velocities, for each system and calculated the standard deviation.

In addition to the obtained grid of thermodynamic values, calculated by GIST, an in-house script was used to identify hydration hot spots. This script first searches for the grid point with the largest water density. The density of one water molecule is then subtracted from the obtained and the surrounding grid points. Next, the program finds the most abundant grid point and again the density of one water molecule is subtracted. This is done for 99% of the water density, since a 1% residual density is necessary, due to boundary effects. To assign thermodynamic values to the refined water positions, the according value (e.g., entropy, enthalpy, or free energy) of the analyzed grid points is averaged i.e., density weighted. Thereby allowing for a localization of entropy and free energy values despite the fact that both are ensemble properties. We used the cpptraj⁴⁹ implementation of GIST (v.18.01) to analyze the obtained trajectories described above with 5000 frames equally distributed over 10 ns.

Thermodynamic Integration (TI). The final frame of the equilibration protocol was used as a starting structure for the alchemical free energy calculations. For all temperatures different from the initial equilibration temperature of 300 K, another 20 ns long heating step was performed individually for each λ -window. A single-step transformation, using soft-core potentials for the van der Waals and electrostatic contributions simultaneously, was done with the pmemd implementation of AMBER as outlined by Kaus et al.⁵⁴ We used 11 λ -windows of 10 ns simulation time each and performed thermodynamic integration analysis using the trapezoid rule. These calculations were repeated at three different temperatures (280, 300, and 320 K), allowing us to calculate the enthalpy and entropy via linear regression, thus neglecting the temperature dependence of the entropy. To include errors for the solvation free energy, the calculations were performed three times for each system at each temperature. For enthalpy and entropy, errors were calculated based on errors for free energies using error propagation. Additionally, we performed 500 ns simulations of the end points to calculate the enthalpies. From the differences of the free energies and enthalpies we calculated the corresponding entropy,⁵⁵ Table S3.

RESULTS

To characterize the thermodynamic properties of the 40 chosen compounds, we used GIST and TI. In the case of GIST, values for the solvation free energy, enthalpy, and entropy were obtained by integrating over all grid points within 5 Å of the solute. This radius was chosen as a compromise to obtain most of the solvation in the first and second hydration layer but avoiding noise from including too many water

Table 1. Thermodynamic Data Obtained with GIST and TI for 40 Aromatic Compounds^a

Aromatic Compound	$\Delta G(\text{GIST})$ (kcal/mol)	$\Delta G(\text{TI})$ (kcal/mol)	$\Delta H(\text{GIST})$ (kcal/mol)	$\Delta H(\text{TI})$ (kcal/mol)	$T\Delta S(\text{GIST})$ (kcal/mol)	$T\Delta S(\text{TI})$ (kcal/mol)
benzene	-1.61 ± 0.23	-0.95 ± 0.02	-5.62 ± 0.17	-5.15 ± 0.12	-4.01 ± 0.17	-4.29 ± 0.14
pyridine	-3.68 ± 0.13	-3.20 ± 0.06	-7.37 ± 0.11	-8.52 ± 0.20	-3.71 ± 0.11	-5.44 ± 0.21
pyridazine	-6.16 ± 0.23	-6.17 ± 0.06	-10.33 ± 0.15	-11.00 ± 0.06	-4.16 ± 0.15	-4.87 ± 0.06
pyrazine	-3.82 ± 0.19	-3.43 ± 0.06	-7.06 ± 0.14	-8.48 ± 0.05	-3.24 ± 0.14	-5.12 ± 0.05
pyrimidine	-5.89 ± 0.09	-6.49 ± 0.25	-10.93 ± 0.09	-10.62 ± 1.51	-5.03 ± 0.09	-4.48 ± 1.61
furan	-2.07 ± 0.16	-1.02 ± 0.01	-4.31 ± 0.14	-4.44 ± 0.01	-2.25 ± 0.14	-3.48 ± 0.01
oxazole	-4.39 ± 0.13	-3.68 ± 0.06	-7.18 ± 0.10	-7.83 ± 0.26	-2.79 ± 0.09	-4.20 ± 0.25
isoxazole	-4.69 ± 0.07	-4.23 ± 0.03	-7.43 ± 0.11	-7.82 ± 0.02	-2.74 ± 0.11	-3.60 ± 0.02
thiophene	-1.88 ± 0.13	-0.99 ± 0.05	-4.95 ± 0.22	-5.11 ± 0.66	-3.07 ± 0.21	-4.11 ± 0.63
thiazole	-3.60 ± 0.16	-2.82 ± 0.07	-6.43 ± 0.17	-6.62 ± 0.05	-2.83 ± 0.17	-3.89 ± 0.06
pyrrole	-4.95 ± 0.16	-5.20 ± 0.08	-10.49 ± 0.06	-11.08 ± 0.64	-5.54 ± 0.06	-5.94 ± 0.65
imidazole	-7.86 ± 0.19	-8.29 ± 0.11	-12.00 ± 0.22	-15.53 ± 0.24	-4.14 ± 0.22	-7.35 ± 0.25
pyrazole	-6.41 ± 0.19	-6.82 ± 0.17	-10.77 ± 0.2	-13.09 ± 0.63	-4.35 ± 0.20	-6.43 ± 0.63
naphthalene	-2.21 ± 0.41	-2.07 ± 0.12	-6.89 ± 0.33	-8.06 ± 0.62	-4.68 ± 0.33	-6.10 ± 0.61
quinoline	-4.09 ± 0.51	-4.48 ± 0.14	-9.90 ± 0.35	-12.88 ± 0.01	-5.81 ± 0.35	-8.53 ± 9.95
isoquinoline	-4.20 ± 0.33	-4.64 ± 0.23	-9.79 ± 0.34	-12.89 ± 1.43	-5.59 ± 0.34	-8.49 ± 1.43
cinnoline	-5.08 ± 0.49	-6.16 ± 0.07	-10.31 ± 0.35	-16.02 ± 1.67	-5.23 ± 0.35	-9.96 ± 1.59
phthalazine	-7.27 ± 0.13	-8.27 ± 0.08	-13.42 ± 0.03	-14.32 ± 0.07	-6.15 ± 0.03	-6.14 ± 0.07
quinazoline	-6.37 ± 0.10	-7.15 ± 0.27	-13.73 ± 0.10	-13.89 ± 0.56	-7.37 ± 0.11	-7.01 ± 0.56
quinoxaline	-4.38 ± 0.27	-4.64 ± 0.07	-9.75 ± 0.09	-12.13 ± 0.11	-5.37 ± 0.09	-7.57 ± 0.10
pteridine	-8.62 ± 0.34	-10.73 ± 0.37	-16.12 ± 0.34	-22.09 ± 3.30	-7.49 ± 0.34	-11.57 ± 3.32
pyridazino[4,5-d]pyridazine	-9.87 ± 0.42	-11.96 ± 0.09	-16.93 ± 0.31	-20.37 ± 0.81	-7.05 ± 0.31	-8.50 ± 0.81
pyrimidino[4,5-d]pyrimidine	-11.54 ± 0.22	-16.63 ± 0.07	-20.21 ± 0.18	-31.23 ± 0.59	-8.66 ± 0.17	-14.71 ± 0.63
pyrazino[2,3-b]pyrazine	-6.65 ± 0.35	-10.19 ± 0.02	-12.83 ± 0.33	-17.84 ± 0.02	-6.58 ± 0.67	-7.56 ± 0.03
benzothiophene	-1.95 ± 0.4	-2.103 ± 0.04	-6.14 ± 0.34	-8.43 ± 0.93	-4.18 ± 0.34	-6.48 ± 0.87
benzothiazole	-4.09 ± 0.6	-4.29 ± 0.17	-9.15 ± 0.43	-10.4 ± 0.64	-5.06 ± 0.43	-6.28 ± 0.64
benzofuran	-2.73 ± 0.17	-2.28 ± 0.03	-6.82 ± 0.07	-8.69 ± 0.15	-4.08 ± 0.07	-6.51 ± 0.16
benzoxazole	-5.15 ± 0.29	-5.25 ± 0.10	-10.3 ± 0.16	-12.05 ± 0.77	-5.14 ± 0.16	-6.94 ± 0.76
indole	-5.25 ± 0.33	-5.96 ± 0.15	-11.18 ± 0.09	-13.24 ± 0.53	-5.92 ± 0.09	-7.41 ± 0.52
benzimidazole	-8.41 ± 0.27	-9.99 ± 0.17	-15.2 ± 0.12	-16.92 ± 0.03	-6.80 ± 0.11	-7.10 ± 0.03
indazole	-6.32 ± 0.39	-7.27 ± 0.21	-12.36 ± 0.37	-11.47 ± 4.36	-6.04 ± 0.37	-4.40 ± 4.36
1,7-diazaindene	-6.66 ± 0.35	-7.08 ± 0.53	-12.83 ± 0.33	-13.83 ± 0.07	-6.18 ± 0.33	-7.28 ± 0.07
benzotriazole	-7.64 ± 0.08	-8.5 ± 0.36	-12.78 ± 0.11	-15.26 ± 0.13	-5.14 ± 0.11	-6.60 ± 0.13
purine	-12.09 ± 0.21	-14.2 ± 0.17	-20.81 ± 0.17	-19.38 ± 0.50	-8.72 ± 0.17	-5.29 ± 0.51
indolizine	-3.18 ± 0.10	-3.38 ± 0.43	-8.05 ± 0.05	-9.18 ± 3.80	-4.86 ± 0.04	-6.31 ± 5.67
pyrrolo[1,2-a]pyrimidine	-6.29 ± 0.28	-6.98 ± 0.05	-12.16 ± 0.20	-12.13 ± 0.83	-5.87 ± 0.21	-5.47 ± 0.89
pyrrolo[1,2-b]pyridazine	-4.07 ± 0.26	-4.56 ± 0.17	-9.22 ± 0.25	-11.44 ± 0.69	-5.14 ± 0.25	-7.02 ± 0.67
2-azaindolizine	-6.39 ± 0.28	-7.33 ± 0.39	-11.91 ± 0.09	-15.33 ± 0.94	-5.52 ± 0.08	-8.41 ± 0.91
thienothiophene	-2.04 ± 0.41	-2.08 ± 0.04	-6.07 ± 0.12	-8.15 ± 0.19	-4.03 ± 0.12	-6.22 ± 0.18
imidazothiazole	-5.99 ± 0.18	-6.67 ± 0.53	-11.27 ± 0.12	-12.65 ± 3.64	-5.27 ± 0.12	-6.51 ± 3.64

^aThe solvation free energies, enthalpies, and the entropic terms calculated using the respective methods are given in kilocalories per mole.

molecules with already bulk-like properties. All thermodynamic values obtained with GIST are reported in Table 1: columns 1, 3, and 5. One of the main advantages of GIST is the spatial resolution of thermodynamic properties around the solute. This allows us to investigate what causes the differences in solvation thermodynamics.

We compared the calculated solvation free energies of several aromatic rings with experimental results to examine the reliability of GIST to predict solvation free energies of aromatic scaffolds (Figure 2A black line; Table S1 column 5). For monocycles, the calculated solvation free energy values from GIST calculations range from -1.6 kcal/mol for benzene (apolar) to -7.9 kcal/mol for imidazole (polar). This is not only expected from chemical intuition but is also in good agreement with experimental results. Benzene has a solvation free energy of -0.9 kcal/mol, and imidazole, one of -8.3 kcal/mol. In general, we found excellent agreement of experimental

results with the solvation free energies calculated using GIST with a Pearson Correlation of 0.96 (Figure 2A).

For comparison, we also included hydration thermodynamic values obtained by TI (Figure 2A red line; Table 1 columns 2, 4, and 6), which is a well-established method to predict solvation thermodynamics of small molecules. The entropies and enthalpies from TI (Table 1: columns 4 and 6, respectively) were obtained by performing calculations at three different temperatures (280, 300, and 320 K) and using a linear regression of the Gibbs eq (eq 2). The assumption was made that the enthalpy and the entropy are temperature independent in this temperature range.

$$\Delta G = \Delta H - T\Delta S \quad (2)$$

GIST calculations as well as TI simulations showed excellent agreement with the experimentally observed trend (cf. Figure 2A black and red line, respectively) showing an RMSE of 1.09

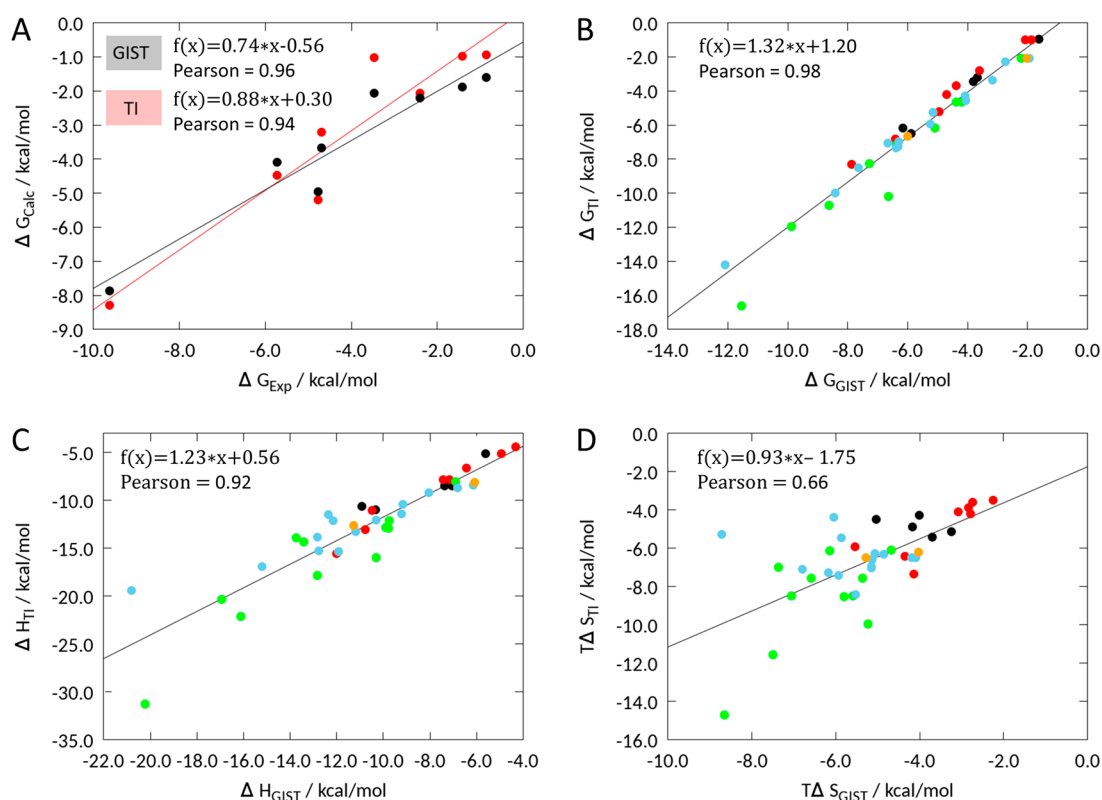


Figure 2. Pearson correlations of (A) experimental, compared to calculated solvation free energies using GIST (black) and TI (red) (Table S1: column 5). (B) Pearson correlations of solvation free energies calculated with GIST and TI for six- (black) and five-membered monocycles (red), fused six-six- (green), five-six- (blue), and five-five-rings (orange). (C) Respective enthalpy correlation. (D) Respective Pearson correlation for entropy calculations.

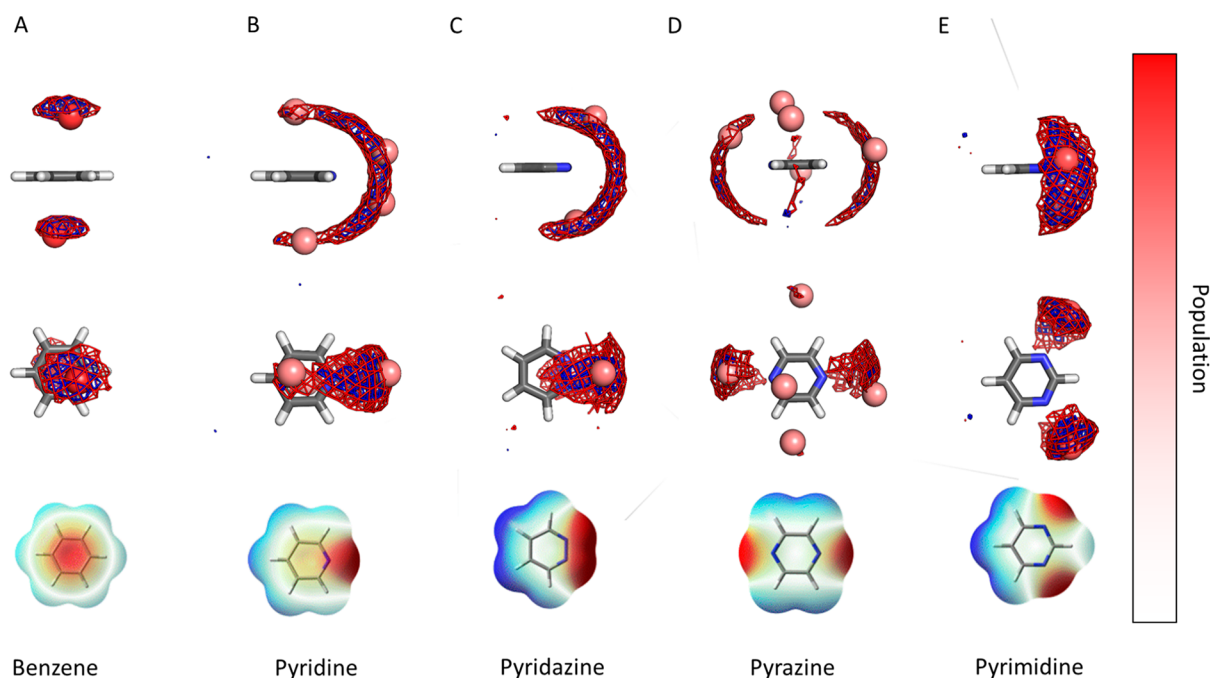


Figure 3. GIST maps revealing the most occupied hydration positions for six-membered monocycles from a side (top row) and top view (middle row) given as large spheres in shades of red according to population. The more intense the red of the sphere, the more located the water molecules are at one position during the simulation. White spheres equal the value for bulk water. Red mesh: favorable solute–water enthalpy (cutoff $-0.1 \text{ kcal}/(\text{mol}\cdot\text{\AA}^3)$). Blue mesh: low water entropy (cutoff $-0.08 \text{ kcal}/(\text{mol}\cdot\text{\AA}^3)$). Electrostatic potential calculated with DFT at $\omega\text{B97XD}^{56}/\text{cc-pVDZ}^{57}$ level using Gaussian09 mapped on the electron density (bottom row).

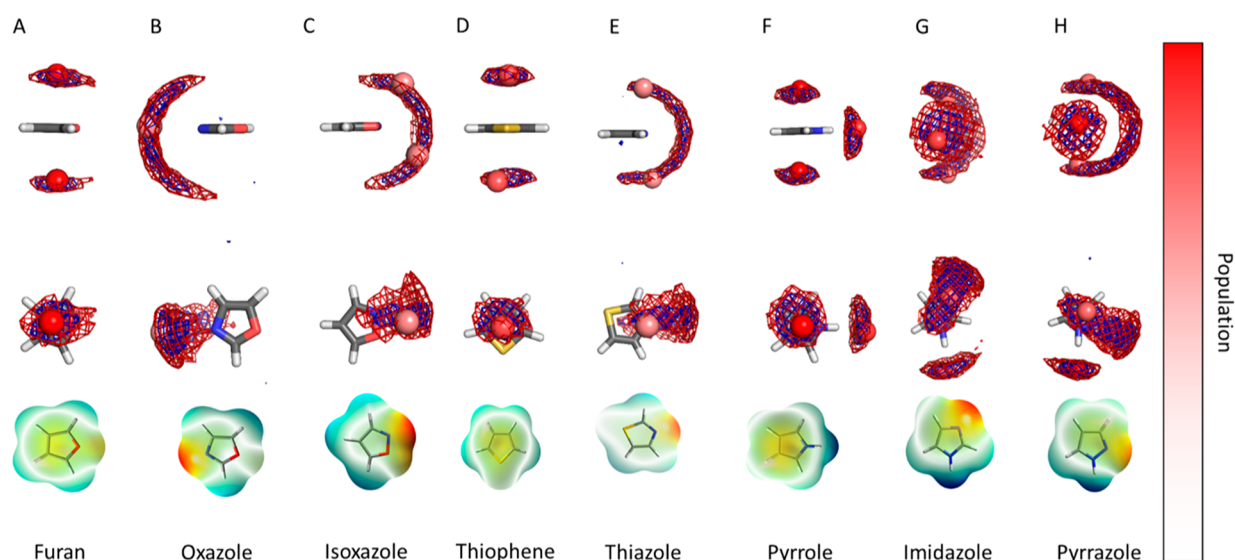


Figure 4. GIST maps revealing the most likely hydration hot spots for five-membered heterocycles from a side (top row) and top view (middle row). Red mesh: favorable water solute enthalpy (cutoff -0.1 kcal/(mol $\cdot\text{\AA}^3$)). Blue mesh: low water entropy (cutoff -0.08 kcal/(mol $\cdot\text{\AA}^3$)). More intensely red spheres show more located water molecules. Electron densities calculated with DFT at the ω B97XD/cc-pVDZ level using Gaussian09 mapped on the electrostatic potential also calculated using Gaussian09 at the same level of theory (bottom row).

kcal/mol for GIST and 1.23 kcal/mol for TI respectively, (Table S5). In general, the GIST results show slightly better correlation with experimental results. If the most negative data point is excluded, the correlation for GIST is still strong (0.88), whereas TI shows a slightly lower correlation (0.85) for the residual data points.

A comparison of GIST and TI results of all 40 molecules allows us to evaluate the differences and errors of the methods, as the errors arising from the use of the force field are the same for both methods. The correlation between the free energies obtained with the two methods is strong, with a Pearson correlation coefficient of 0.98 (Figure 2B). The slope of 1.32 likely results from truncating the GIST analysis at 5 \AA and therefore missing some solvent effects further apart from the solute. The enthalpies obtained from GIST and TI show a similarly strong resemblance (Figure 2C). The high interception between the GIST and the TI results is, on first sight, surprising. We assume it arises from the fact that GIST underestimates the enthalpic interactions for enthalpically very favorable binders. Strong enthalpic effects may have a larger prolongation than 5 \AA and therefore are not entirely covered within the used cutoff. However, the correlation for the enthalpy values obtained through GIST and TI is still strong but drops significantly for the entropy, with a Pearson correlation coefficient of only 0.66 (Figure 2D). Nevertheless, these results are in good agreement with previously published studies on amino acids.⁵² The enthalpies and entropies calculated with the end point method⁵⁵ are similar and do not improve the correlation, Table S3. The errors cannot be significantly reduced by increasing GIST sampling time and/or number of snapshots. From the GIST analyses we can identify “hot spots”, where water molecules are permanently or frequently located. We visualize these hot spots in Figure 3 with large spheres in different shades of red in the first two rows. The more intense the red, the more located a water molecule is throughout the simulation. The positions of the spheres were calculated from the population obtained through GIST analyses. To depict the solvation thermodynamics in the proximity of the solute, favorable solute–water enthalpy is

shown as a red mesh (cutoff -0.1 kcal/(mol $\cdot\text{\AA}^3$)), and low water entropy areas is given as a blue mesh (cutoff -0.08 kcal/(mol $\cdot\text{\AA}^3$)). From the GIST simulation of benzene, the interaction of its aromatic π -cloud with water can be identified (Figure 3 column A). Intuitively, for benzene, this is the only region where strong solute–solvent interactions can be identified.

In comparison to benzene, compounds with heteroatoms show additional, strong interactions with water molecules. Furthermore, these heteroatoms also influence the electrostatic properties of the aromatic systems (π -cloud) and, therefore, the interaction potential with water molecules. Introducing a nitrogen atom, as i.e., in pyridine, offers an additional interaction with the solvent via an H-bond accepting atom. However, the favorable interaction of water with the aromatic π -cloud is still present (Figure 3B).

The point of substitution considerably influences the effect of adding a second heteroatom. The relative position of the nitrogen atoms has substantial influence, not only on the electrostatic properties of an aromatic monocycle but also on the solvation free energy. In monocycles containing two heteroatoms, the dipole moment is the strongest when the heteroatoms are positioned next to each other. In pyridazine, with the highest dipole moment of 4.01 D calculated from the QM calculation (ω B97XD/cc-pVTZ), the hydration free energy is hence the lowest with -6.16 kcal/mol. Pyrimidine, with a dipole moment of 2.20 D, shows a solvation free energy of -5.89 kcal/mol. Pyrazine, which has no dipole moment, has a substantially higher solvation free energy with -3.8 kcal/mol (Table 1 columns 1 and 2). The influence of the relative positions of two heteroatoms is also evident in the assignment of point charges during the parametrization process. Interestingly, the compound with the highest solvation free energy has the lowest partial charges—which might be counterintuitive (Figure S1). With respect to the influence in solvation free energy and the point charge distribution, we already see that introducing a nitrogen atom to the aromatic ring can have different effects, nonadditive behavior must be considered.

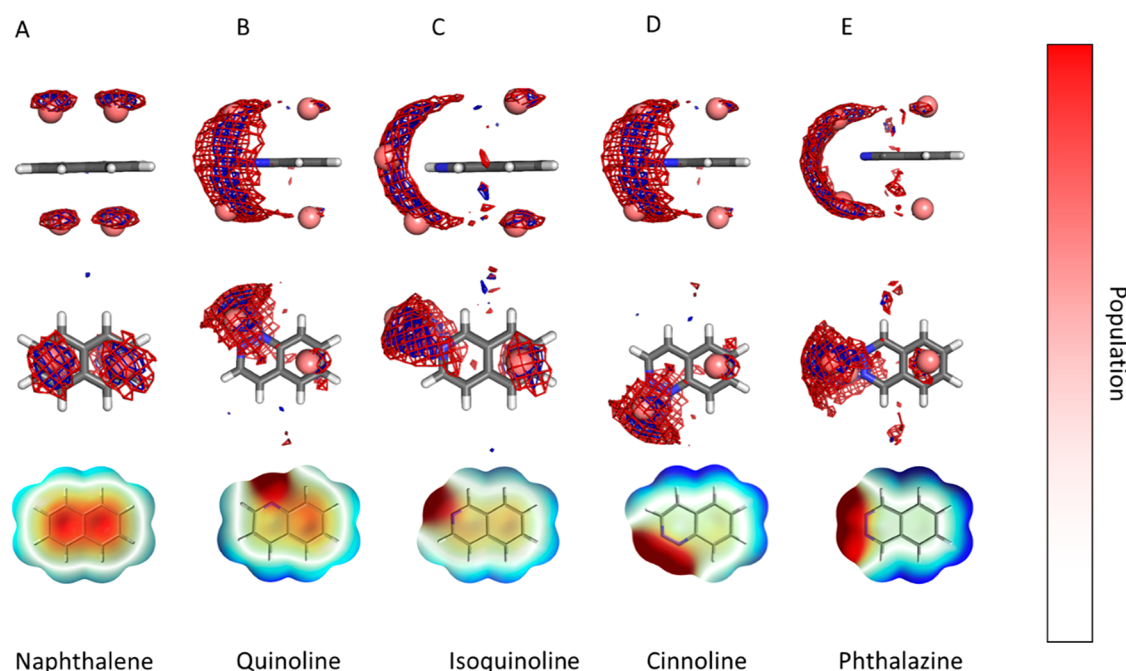


Figure 5. GIST maps of most likely hydration hot spots for fused six-six-rings from a side (top row) and top view (middle row). Red grid: favorable water–solute enthalpy (cutoff -0.1 kcal/(mol·Å³)). Blue grid: low water entropy (cutoff -0.08 kcal/(mol·Å³)). Positions of spheres were calculated from the population in GIST. More intense red represents more located water, and white equals bulk water. Electron densities calculated with DFT at the ω B97XD/cc-pVDZ level using Gaussian09 (bottom row) are mapped on the electrostatic potential also calculated using Gaussian09 at the same level of theory.

In additive force field based approaches, the electrostatic properties are treated via relatively simple point charges (Figure S1). The modeling of the charge distribution for heteroaromatics is a possible weak spot of additive FFs without off-site charges or atomic multipoles.⁵⁸ However, to confirm that the obtained results are not only due to the simplistic representation of the electrostatic potential (the RESP charges for pyrimidine are high), we compared the obtained GIST results to quantum calculations at DFT ω B97XD/cc-pVDZ level (calculated with Gaussian09). The correlation between the QM electrostatic potential (ESP) and the thermodynamic hydration grids, respectively the hydration hot spots predicted from our GIST simulations is nearly perfect (Figure 3). We can see a clear increase of the electrostatic potential above the center of the aromatic ring. Whereas the exact explanation might be different, as pointed out by Wheeler and Bloom,⁵⁹ in a simplified picture we can interpret this as a decrease in electron density at the π -cloud when heteroatoms are added to the aromatic ring.

Results, similar to those for six-membered rings, can be found for the smaller five-membered rings (Figure 4). We distinguish between aromatic five-rings containing heteroatoms, solely serving as H-bond acceptors, and aromatic five-rings containing a secondary amine: First, GIST identifies the higher H-bond acceptor strength of nitrogen over oxygen, and nitrogen over sulfur, respectively (Figure 4B and E). In contrast to furan and thiophene, pyrrole offers an additional, very favorable interaction site with water, resulting from the secondary amine, as shown in the GIST mesh representation in Figure 4E. This additional strong interaction site naturally affects the solvation free energy, resulting in a solvation free energy of -2.1 kcal/mol for furan, -1.9 kcal/mol for thiophene, and -5.0 kcal/mol for pyrrole from the GIST calculations. This is again in good agreement with experimental

results for furan, with a solvation free energy of -3.4 kcal/mol, thiophene, with -1.4 kcal/mol, and pyrrole, with -4.7 kcal/mol, respectively (Table S1 column 4).

In five-membered rings, adding a second heteroatom to the aromatic ring generally results in a more negative solvation free energy. Investigating these derivatives, we found that they offer a strong hydrogen bond accepting site, clearly visible in the GIST meshes (Figure 4). In all three cases, the solvation free energy is lowered by approximately 2 kcal/mol as listed in Table 1 column 2. While adding a nitrogen to pyrrole also results in a lower solvation free energy, the point of substitution has a strong influence. In contradiction to the six-membered rings, separating the nitrogen atoms shows a stronger effect on the solvation free energy, than if they are positioned next to each other. The increase from pyrrole with -5.0 kcal/mol to pyrazole with -6.4 kcal/mol is smaller than to imidazole with -7.9 kcal/mol. Nitrogen is acting as the strongest H-bond acceptor in the heteroaromatics in group 2, which is in good agreement with the corresponding H-bond strength. When focusing on the point charges, pyrazole shows a substantially less negative charge on the nitrogen atom of the secondary amine. This is due to the close vicinity of the second aromatic nitrogen atom in the aromatic ring. However, in terms of entropic and enthalpic effects, we see that these are mainly influenced by the positive charge of the secondary amine's hydrogen (Figure S2).

Fusing two benzene rings to a naphthalene, two hydration hot spots at the respective π -clouds are identified (Figure 5A). Not only does the resulting GIST map compare to the map of two benzene rings, the solvation free energy is approximately double the solvation free energy of one benzene. Joining benzene and pyridine, depending on the orientation of the added monocycles, resulting in quinoline and isoquinoline, once again shows additivity, when it comes to identifying the

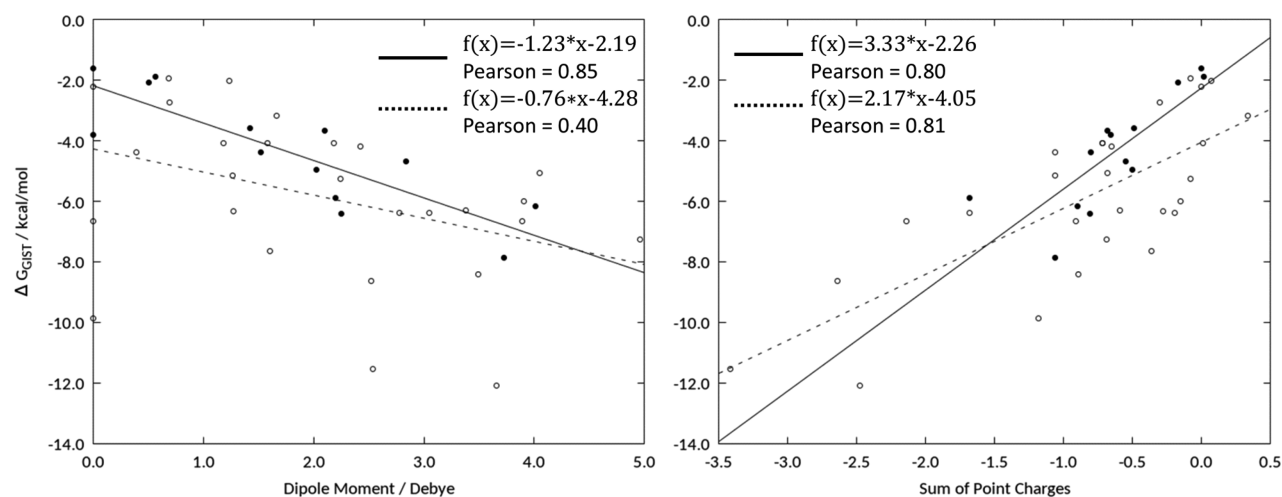


Figure 6. (A) Correlation between the molecular dipole moments and the solvation free energies: filled circles monocycles; hollow circles bicycles; solid line linear regression of monocycles; dashed line correlation for bicycles. (B) Correlation of the sums of point charges in the heteroatoms calculated using RESP with the solvation free energies. For secondary amines, the partial charge of the corresponding hydrogen atom was included in the calculation.

hydration hot spots in the GIST maps (Figure 5B and C). Adding more nitrogen atoms to the studied molecules, cinnoline and phthalazine show a slight divergence in the similarity of the GIST maps. Especially the probability of water molecules at the benzene π -cloud is reduced, when comparing the GIST maps and hydration hot spots of the unsubstituted ring of cinnoline and phthalazine to the respective GIST maps of benzene (cf. Figures 5D and E and 3A). Comparing the dipole moment of this series (given in Table S1 column 2) further emphasizes the strong influence of electrostatic properties on the solvation free energy. This increase in dipole moment from 2.18 and 2.42 D for quinoline and isoquinoline, respectively, compared to 4.05 and 4.96 D, respectively, for cinnoline and phthalazine correlates with the decreasing solvation free energy of -4.09 kcal/mol for quinoline with the lowest dipole moment to -7.27 kcal/mol for phthalazine.

The strong correlation for the solvation free energy of bicycles and their respective monocycles holds up throughout the data set. We obtain an overall Pearson Correlation of 0.89 (Figure S3). This correlation allows to further predict solvation free energies of bicyclic aromatic compounds based on their respective monomers. Outliers from this strong correlation are fused five- and six-membered heteroaromatics, where a secondary amine is a bridging atom. The deviation is most likely due to the change of a secondary amine to an aromatic nitrogen, thereby losing the amine hydrogen, one major contributor to the solvation free energy of pyrrole, Table S6.

DISCUSSION

In this study, we investigate the solvation free energy of aromatic compounds using two different, well-established methodologies: GIST and TI. Results from these calculations agree well in terms of free energy and enthalpy, however, as previously reported, the entropic contributions differ quite substantially. Despite the disagreement in entropy, with our setup, GIST is more efficient in predicting thermodynamic properties, as only a single simulation is needed. In contrast, for TI, multiple simulations are needed to describe the different, nonphysical states of the alchemical process, making it an average of ten times slower (for the used simulation setup). In addition to get reliable values for enthalpy and

entropy either longer runs of the end points⁵⁵ or simulations at different temperatures are necessary.

In previous studies, a strong correlation was identified between the molecular dipole moment and the π -stacking interaction energy for aromatic monocycles.¹ We further observe a strong inverse correlation between the solvation free energy and the dipole moment (Figure 6A), pointing toward the necessity of including solvation free energies for predicting the overall binding energy: When a compound with a higher stacking energy is chosen, the compound's desolvation penalty is also higher, thereby reducing or even annihilating the expected gain in binding free energy.

QM calculations (details are given in the Supporting Information) were used to describe the interaction of a single water molecule with benzene. The most favorable water position (-4.6 kcal/mol) is above the center of the solute (Table S4). This value is already significantly higher than the value of -3.6 kcal/mol for benzene–benzene interactions obtained by Huber et al.,¹ highlighting that the contribution from desolvation can overcompensate the effect introduced by the change in stacking energy.

For the bicyclic heteroaromatics, the correlation of a dipole moment cannot be confirmed—neither for the solvation free energy nor for the stacking interaction (Figures 6A and 7). This is mainly stemming from symmetric substitution patterns, which can lead to the reduction or complete annihilation of the molecular dipole moment (Table S1 column 2). Nevertheless, the electrostatic potential of these compounds is not uniform.

As most additive force fields describe electrostatic properties of atoms with point charges, we used the sum of absolute heteroatomic point charges as a descriptor instead and compared it to the solvation free energy. The point charges were obtained using RESP fitting, as mentioned in the Methods section. If the molecule contained a secondary amine, we added the corresponding hydrogen atom to account for charge compensation between the two linked atoms. The charge descriptor shows a strong correlation with the solvation free energy, highlighting how the charges and the electrostatic potential are responsible for the hydration behavior of a molecule.

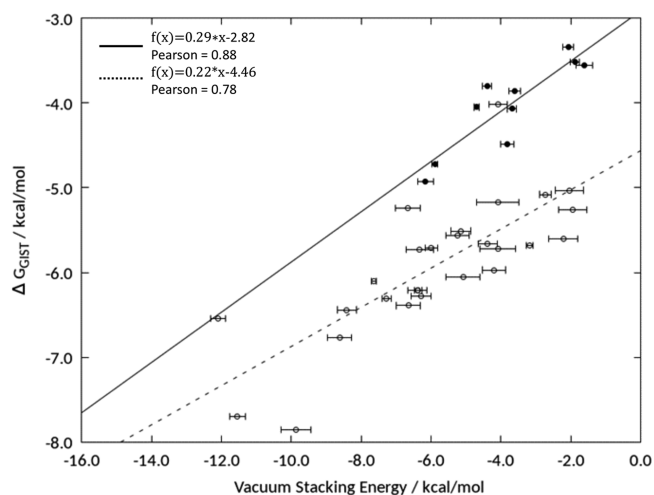


Figure 7. Correlation of the solvation free energies calculated using GIST with QM based vacuum interaction energies. Pearson correlation coefficients are 0.88 for monocycles, and 0.78 for bicycles, respectively. Overall correlation results in a Pearson coefficient of 0.76: filled circles monocycles; hollow circles bicycles.

Correlating the solvation free energies with calculated vacuum interaction energies for π - π stacking of monocycles, published by Huber et al.,¹ reveals an almost perfect agreement, with a Pearson correlation coefficient of 0.88 (Figure 7). For bicycles, our group also performed extensive scans to identify low energy stacking conformations of bicycles with benzene, using the same calculation setup used for the monocycles.⁶⁰ A good correlation, with a Pearson correlation coefficient of 0.78, between the bicycle interaction energies with benzene and solvation free energies presented in this work could also be observed (Figure 7).

In general, the vacuum stacking interaction energies for the monocycles are higher (less favorable) than for the bicyclic compounds, due to their smaller size. The solvation free energy, on the other hand, is not necessarily higher for larger bicycles. For instance, while furan and naphthalene are within the respective error range in terms of solvation free energy, with a mean difference of 0.14 kcal/mol (furan -2.07 ± 0.16 and naphthalene -2.21 ± 0.41 kcal/mol), the difference in vacuum stacking interaction is substantial, with -2.26 kcal/mol. Similarly, thiazole and indolizine are quite similar in terms of solvation free energy (thiazole -3.60 ± 0.16 kcal/mol and indolizine -3.18 ± 0.10 kcal/mol) but significantly differ in stacking interaction energy by 1.8 kcal/mol.

For polar bicyclic heteroaromatics, such as purine, and the tetra-aza-substituted naphthalene derivatives, such as pteridine, we observed a generally strong vacuum stacking interaction energy. Still, the hydration free energies are strongly favorable, resulting in a high desolvation penalty.

We can also identify compounds, which show a very similar vacuum interaction energy, but strongly differ in solvation free energy. For instance, focusing on the series of fused five- and six-membered rings with a nitrogen in a bridge position, we find that, while the vacuum stacking interaction energy for indole and indazole solely differs by 0.12 kcal/mol, the solvation free energy diverges by 1.2 kcal/mol. Indazole and indolizine show an even closer similarity of 0.10 kcal/mol in the calculated vacuum stacking interaction energy, while the solvation penalty for indazole is 3.1 kcal/mol higher than for indolizine.

CONCLUSION

In this study, we calculated the solvation free energy of 40 heteroaromatic compounds using GIST and TI. We found that the correlation of TI and GIST is excellent; both our results agree well with previously published experimental and computational data. However, the calculation with GIST allows for much more efficient generation of data. Additionally, GIST gives information on hydration hot spots, which can be of great interest in drug design projects.

Identifying pairs with similar interaction energy and/or solvation free energy will be of great interest for scaffold hopping. First, identifying compounds with similar properties in hydration and binding free energies allows to escape patented chemical space, without changing key properties. Second, including information on solvation free energy might help to explain why sometimes the expected gain in affinity from vacuum calculations cannot be observed in experiments. Furthermore, in the hit-to-lead stage of drug development, when optimizing an aromatic core by scaffold hopping,⁶¹ combining the knowledge of vacuum interaction energies with information on physicochemical and thermodynamic properties of a candidate can be used to optimize the binding affinity, without losing sight of solvation properties. Additionally, the resulting GIST map can be used as a general approach to screen for complementarity in the binding pocket, as it can easily be visualized where potential strong hydrogen bond donors and acceptors are located.

Foremost, our study reveals the strong inverse correlation between the solvation free energy and stacking interaction energies calculated in vacuum, showing the necessity to include solvation properties for predicting binding free energies and affinities.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.9b00395.

Data necessary to generate the plots in the manuscript and further correlations with additional databases, information on the calculation setup for the water benzene interaction, and a table containing the pairs used to predict the additivity of monocycles (PDF)

AUTHOR INFORMATION

Corresponding Author

*Email: mschauperl@ucsd.edu.

ORCID

Johannes R. Loeffler: 0000-0002-5724-655X

Michael Schauerl: 0000-0001-5648-8170

Klaus R. Liedl: 0000-0002-0985-2299

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Austrian Science Fund (FWF), P 30565, and an Erwin Schrödinger fellowship J-4150 (granted to MS). Computational results have been achieved using the high-performance computing infrastructure LEO of the University of Innsbruck. We thank Dr. Julian Fuchs for his support in the initial phase of this project. We would also like

to express our gratitude to Dr. Teresa Ortner for her support in preparing the manuscript.

■ ABBREVIATIONS

ADME, absorption, distribution, metabolism, excretion; DFT, density functional theory; GAFF, general Amber force field; GIST, grid inhomogeneous solvation theory; RESP, restrained electrostatic potential; SMILES, simplified molecular input line entry system, TI, thermodynamic integration

■ REFERENCES

- (1) Huber, R. G.; Margreiter, M. A.; Fuchs, J. E.; von Grafenstein, S.; Tautermann, C. S.; Liedl, K. R.; Fox, T. Heteroaromatic π -Stacking Energy Landscapes. *J. Chem. Inf. Model.* **2014**, *54*, 1371–1379.
- (2) Gallivan, J. P.; Dougherty, D. A. Cation- π Interactions in Structural Biology. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9459–9464.
- (3) Dougherty, D. A. The Cation- π Interaction. *Acc. Chem. Res.* **2013**, *46*, 885–893.
- (4) Wallnofer, H. G.; Fox, T.; Liedl, K. R.; Tautermann, C. S. Dispersion Dominated Halogen- π Interactions: Energies and Locations of Minima. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14941–14949.
- (5) Wilcken, R.; Zimmermann, M. O.; Lange, A.; Joerger, A. C.; Boeckler, F. M. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *J. Med. Chem.* **2013**, *56*, 1363–1388.
- (6) Graton, J.; Le Questel, J.-Y.; Maxwell, P.; Popelier, P. Hydrogen-Bond Accepting Properties of New Heteroaromatic Ring Chemical Motifs: A Theoretical Study. *J. Chem. Inf. Model.* **2016**, *56*, 322–334.
- (7) Wheeler, S. E.; Bloom, J. W. G. Toward a More Complete Understanding of Noncovalent Interactions Involving Aromatic Rings. *J. Phys. Chem. A* **2014**, *118*, 6133–6147.
- (8) McGaughey, G. B.; Gagné, M.; Rappé, A. K. Pi-Stacking Interactions. Alive and Well in Proteins. *J. Biol. Chem.* **1998**, *273*, 15458–15463.
- (9) Wheeler, S. E. Understanding Substituent Effects in Noncovalent Interactions Involving Aromatic Rings. *Acc. Chem. Res.* **2013**, *46*, 1029–1038.
- (10) Harder, M.; Kuhn, B.; Diederich, F. Efficient Stacking on Protein Amide Fragments. *ChemMedChem* **2013**, *8*, 397–404.
- (11) Janiak, C. A Critical Account on π - π Stacking in Metal Complexes with Aromatic Nitrogen-Containing Ligands. *J. Chem. Soc., Dalton Trans.* **2000**, 3885–3896.
- (12) Lee, H.; Dehez, F.; Chipot, C.; Lim, H. K.; Kim, H. Enthalpy-Entropy Interplay in π -Stacking Interaction of Benzene Dimer in Water. *J. Chem. Theory Comput.* **2019**, *15*, 1538.
- (13) Stroet, M.; Caron, B.; Visscher, K. M.; Geerke, D. P.; Malde, A. K.; Mark, A. E. Automated Topology Builder Version 3.0: Prediction of Solvation Free Enthalpies in Water and Hexane. *J. Chem. Theory Comput.* **2018**, *14*, 5834–5845.
- (14) Slipchenko, L. V.; Gordon, M. S. Water-Benzene Interactions: An Effective Fragment Potential and Correlated Quantum Chemistry Study \ddagger . *J. Phys. Chem. A* **2009**, *113*, 2092–2102.
- (15) Prakash, M.; Samy, K. G.; Subramanian, V. Benzene-Water (BZW_n ($n = 1 - 10$)) Clusters. *J. Phys. Chem. A* **2009**, *113*, 13845–13852.
- (16) Schravendijk, P.; van der Vegt, N. F. A. From Hydrophobic to Hydrophilic Solvation: An Application to Hydration of Benzene. *J. Chem. Theory Comput.* **2005**, *1*, 643–652.
- (17) Raschke, T. M.; Levitt, M. Nonpolar Solutes Enhance Water Structure within Hydration Shells While Reducing Interactions between Them. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6777–6782.
- (18) Cockroft, S. L.; Hunter, C. A. Desolvation Tips the Balance: Solvent Effects on Aromatic Interactions. *Chem. Commun.* **2006**, 3806–3808.
- (19) Mati, I. K.; Adam, C.; Cockroft, S. L. Seeing through Solvent Effects Using Molecular Balances. *Chem. Sci.* **2013**, *4*, 3965–3972.
- (20) Yang, L.; Brazier, J. B.; Hubbard, T. A.; Rogers, D. M.; Cockroft, S. L. Can Dispersion Forces Govern Aromatic Stacking in an Organic Solvent? *Angew. Chem., Int. Ed.* **2016**, *55*, 912–916.
- (21) Yang, L.; Adam, C.; Nichol, G. S.; Cockroft, S. L. How Much Do van Der Waals Dispersion Forces Contribute to Molecular Recognition in Solution? *Nat. Chem.* **2013**, *5*, 1006–1010.
- (22) Yang, L.; Adam, C.; Cockroft, S. L. Quantifying Solvophobic Effects in Nonpolar Cohesive Interactions. *J. Am. Chem. Soc.* **2015**, *137*, 10084–10087.
- (23) Mobley, D. L.; Guthrie, J. P. FreeSolv: A Database of Experimental and Calculated Hydration Free Energies, with Input Files. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 711–720.
- (24) Southall, N. T.; Dill, K. A. The Mechanism of Hydrophobic Solvation Depends on Solute Radius. *J. Phys. Chem. B* **2000**, *104*, 1326–1331.
- (25) Frank, H. S.; Evans, M. W. Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *J. Chem. Phys.* **1945**, *13*, 507–532.
- (26) Schauerl, M.; Czodrowski, P.; Fuchs, J. E.; Huber, R. G.; Waldner, B. J.; Podewitz, M.; Kramer, C.; Liedl, K. R. Binding Pose Flip Explained via Enthalpic and Entropic Contributions. *J. Chem. Inf. Model.* **2017**, *57*, 345–354.
- (27) Wienen-Schmidt, B.; Jonker, H. R. A.; Wulsdorf, T.; Gerber, H.-D.; Saxena, K.; Kudlinzki, D.; Sreeramulu, S.; Parigi, G.; Luchinat, C.; Heine, A.; et al. Paradoxically, Most Flexible Ligand Binds Most Solvation-Favored: Intriguing Impact of Ligand Flexibility and Solvation on Drug-Kinase Binding. *J. Med. Chem.* **2018**, *61*, S922–S933.
- (28) Verteramo, M. L.; Stenström, O.; Ignjatović, M. M.; Caldararu, O.; Olsson, M. A.; Manzoni, F.; Leffler, H.; Oksanen, E.; Logan, D. T.; Nilsson, U. J.; et al. Interplay between Conformational Entropy and Solvation Entropy in Protein-Ligand Binding. *J. Am. Chem. Soc.* **2019**, *141*, 2012–2026.
- (29) Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. Contribution of Explicit Solvent Effects to the Binding Affinity of Small-Molecule Inhibitors in Blood Coagulation Factor Serine Proteases. *ChemMedChem* **2011**, *6*, 1049–1066.
- (30) Bodnarchuk, M. S.; Viner, R.; Michel, J.; Essex, J. W. Strategies to Calculate Water Binding Free Energies in Protein-Ligand Complexes. *J. Chem. Inf. Model.* **2014**, *54*, 1623–1633.
- (31) Uehara, S.; Tanaka, S. AutoDock-GIST: Incorporating Thermodynamics of Active-Site Water into Scoring Function for Accurate Protein-Ligand Docking. *Molecules* **2016**, *21*, 1604.
- (32) Balius, T. E.; Fischer, M.; Stein, R. M.; Adler, T. B.; Nguyen, C. N.; Cruz, A.; Gilson, M. K.; Kurtzman, T.; Shoichet, B. K. Testing Inhomogeneous Solvation Theory in Structure-Based Ligand Discovery. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E6839–E6846.
- (33) SZMAP 1.2.1.4; OpenEye Scientific Software, Santa Fe, NM, 2015.
- (34) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the Active-Site Solvent in the Thermodynamics of Factor Xa Ligand Binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817–2831.
- (35) 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*, 808–813.
- (36) 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* (7), 2769–2780.
- (37) 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.
- (38) Nguyen, C.; Yamazaki, T.; Kovalenko, A.; Case, D. A.; Gilson, M. K.; Kurtzman, T.; Luchko, T. A Molecular Reconstruction Approach to Site-Based 3D-RISM and Comparison to GIST

Hydration Thermodynamic Maps in an Enzyme Active Site. *PLoS One* **2019**, *14* (7), No. e0219473.

(39) Fuchs, J. E.; Huber, R. G.; Waldner, B. J.; Kahler, U.; von Grafenstein, S.; Kramer, C.; Liedl, K. R. Dynamics Govern Specificity of a Protein-Protein Interface: Substrate Recognition by Thrombin. *PLoS One* **2015**, *10*, No. e0140713.

(40) Munos, B. Lessons from 60 Years of Pharmaceutical Innovation. *Nat. Rev. Drug Discovery* **2009**, *8*, 959–968.

(41) Swinney, D. C.; Anthony, J. How Were New Medicines Discovered? *Nat. Rev. Drug Discovery* **2011**, *10*, 507–519.

(42) Ertl, P.; Jelfs, S.; Mühlbacher, J.; Schuffenhauer, A.; Selzer, P. Quest for the Rings. In Silico Exploration of Ring Universe To Identify Novel Bioactive Heteroaromatic Scaffolds. *J. Med. Chem.* **2006**, *49* (15), 4568–4573.

(43) Schuffenhauer, A.; Ertl, P.; Roggo, S.; Wetzel, S.; Koch, M. A.; Waldmann, H. The Scaffold Tree - Visualization of the Scaffold Universe by Hierarchical Scaffold Classification. *J. Chem. Inf. Model.* **2007**, *47* (1), 47–58.

(44) *Molecular Operating Environment (MOE)*; Chemical Computing Group Inc.: Montreal, QC, Canada, 2017.

(45) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graphics Modell.* **2006**, *25* (2), 247–260.

(46) 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* (9), 1157–1174.

(47) 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* (40), 10269–10280.

(48) Frisch, M. J. et al. *Gaussian 09*; Gaussian Inc.: Wallingford CT, 2009.

(49) Case, D. A.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Greene, D.; Homeyer, N. et al. *AMBER 2017*; University of California, San Francisco, 2017.

(50) 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* (2), 926–935.

(51) Wallnoefer, H. G.; Handschuh, S.; Liedl, K. R.; Fox, T. Stabilizing of a Globular Protein by a Highly Complex Water Network: A Molecular Dynamics Simulation Study on Factor Xa. *J. Phys. Chem. B* **2010**, *114* (21), 7405–7412.

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

(53) Koehl, A.; Hu, H.; Maeda, S.; Zhang, Y.; Qu, Q.; Paggi, J. M.; Latorraca, N. R.; Hilger, D.; Dawson, R.; Matile, H.; et al. Structure of the M-Opioid Receptor-G i Protein Complex. *Nature* **2018**, *558* (7711), 547.

(54) Kaus, J. W.; Pierce, L. T.; Walker, R. C.; McCammon, J. A. Improving the Efficiency of Free Energy Calculations in the Amber Molecular Dynamics Package. *J. Chem. Theory Comput.* **2013**, *9* (9), 4131.

(55) Peter, C.; Oostenbrink, C.; van Dorp, A.; van Gunsteren, W. F. Estimating Entropies from Molecular Dynamics Simulations. *J. Chem. Phys.* **2004**, *120*, 2652–2661.

(56) Chai, J.-D.; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.

(57) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90* (2), 1007–1023.

(58) Kramer, C.; Spinn, A.; Liedl, K. R. Charge Anisotropy: Where Atomic Multipoles Matter Most. *J. Chem. Theory Comput.* **2014**, *10* (10), 4488–4496.

(59) Wheeler, S. E.; Bloom, J. W. G. Anion- π Interactions and Positive Electrostatic Potentials of N-Heterocycles Arise from the

Positions of the Nuclei, Not Changes in the π -Electron Distribution. *Chem. Commun.* **2014**, *50*, 11118–11121.

(60) Loeffler, J. R.; Tautermann, C. S.; Fox, T.; Bergner, A.; Liedl, K. R.; Fuchs, J. E. Pi-Stacking Interactions of Bicyclic Heteroaromatics, *manuscript in preparation*.

(61) Wang, L.; Deng, Y.; Wu, Y.; Kim, B.; LeBard, D. N.; Wandschneider, D.; Beachy, M.; Friesner, R. A.; Abel, R. Accurate Modeling of Scaffold Hopping Transformations in Drug Discovery. *J. Chem. Theory Comput.* **2017**, *13* (1), 42–54.