Supplementary information

S.1 AMBER-gaff parameters

The text file AMBER-gaff_gromacs_coordinates_topologies.txt (ASCI-format) contains vacuum coordinates and topologies for all residues as used in all our AMBER-gaff calculations.

S.2 Solvent-solvent solubility

For a water-chloroform system at 293 (303) K the solubility of water in chloroform is 0.44 (0.56) mol % and chloroform in water is 0.12 (0.12) mol % [1]. For a more realistic comparison between calculated and experimental partition coefficients one should add some water to the chloroform phase and some chloroform to the water phase. As we have approximately 250 chloroform and 1000 water molecules in our simulation boxes we added 1 water (0.4 mol %) and 1 chloroform (0.1 mol %), respectively. These simulations were extended to 5 ns.

We have listed the solvation free energies (table Si) and chloroform-to-water partition coefficients (table Sii) obtained from calculations with one molecule of the other solvent added to the system. We observed a clear trend toward lower partition coefficient as a result of taking solvent-solvent solubility into account, figure S1.

Although by adding one molecule of the opposite solvent the solubility of our systems (0.4 mol % and 0.1 mol %) is in close agreement to the experimental value, our simulations cannot sample the complete ensemble of a macroscopic system. For example, we cannot exclude the possibility that in reality the dissolved solvent molecules form clusters. In addition the investigated molecule, *e.g.* the methylated nucleobases or amino acid side chain, can form a complex and drag additional molecules from one solvent into the other solvent. These possibilities cannot be taken into consideration in our simulations.

Addition of one molecule of the other solvent also creates a sampling problem. The added solvent molecule and the solute can namely be in contact or separated. To obtain a converged value, both states should be sampled with a reasonable number of transitions between them. The sampling problem manifests itself mainly in those λ-windows with (close to) full partial charges on the atoms of the investigated molecule in the simulation with chloroform + 1 water. In these λ-windows, both contact and separation between the added solvent molecule and the solute is sampled and more than one transition is observed (see figure S2). Hence, we assume the sampling to be at least sufficient for a qualitative assessment.

References

[1] Stephenson, M., *J. Chem. Eng. Data*, **1992**, 37, 80–95.

Table Si: Calculated solvation free energies (kJ mol^{-1}) of the methylated nucleobases and amino acid side chains in chloroform and water, including 1 water molecule in the chloroform phase and 1 chloroform in the water phase. Values in parentheses are uncertainties in the last significant digit.

	solvent	Ade	Gua	Cyt	Thy	Ura	Trp	Tyr	Phe
AMBER-99	TIP3P+1Chl	-59.40 (16)	-102.10 (20)	-87.05 (14)	-63.95 (14)	-62.70 (14)	-22.40 (7)	-19.81 (14)	-1.89 (9)
	Chl+1TIP3P	-54.94 (52)	-77.03 (52)	-63.44 (71)	-54.86 (24)	-51.18 (34)	-37.24 (19)	-28.56 (38)	-21.17 (6)
AMBER-gaff	TIP3P+1Chl	-65.49 (17)	-103.12 (15)	-87.88 (16)	-56.78 (16)	-59.28 (19)	-22.70 (16)	-17.69 (15)	-2.22 (12)
	Chl+1TIP3P	-54.97 (63)	-74.24 (47)	-59.84 (33)	-51.76 (41)	-50.04 (58)	-37.28 (14)	-28.44 (24)	-21.30 (9)
CHARMM-27	TIP3P+1Chl	-60.80 (16)	-100.34 (24)	-81.91 (15)	-42.60 (20)	-45.54 (14)	-16.99 (20)	-20.14 (9)	-1.12 (12)
	Chl+1TIP3P	-59.50 (46)	-78.99 (132)	-67.76 (48)	-49.06 (18)	-43.40 (43)	-38.32 (17)	-30.04 (28)	-23.50 (19)
GROMOS-45a4/53a6	SPC+1Chl	-33.55 (29)	-76.73 (29)	-66.54 (14)	-40.97 (18)	-41.42 (19)	-27.18 (21)	-27.41 (18)	-2.32 (16)
	Chl+1SPC	-47.14 (63)	-70.58 (98)	-57.51 (61)	-48.13 (14)	-45.17 (32)	-40.07 (72)	-32.20 (81)	-22.73 (9)
OPLS-AA	TIP4P+1Chl	-54.43 (23)	-91.15 (21)	-78.17 (25)	-46.28 (20)	-44.43 (11)	-17.22 (17)	-20.61 (19)	-1.77 (18)
	Chl+1TIP4P	-56.15 (37)	-76.97 (62)	-62.98 (74)	-57.04 (30)	-51.16 (21)	-36.60 (9)	-29.62 (22)	-23.36 (12)

Table Sii: Calculated water-chloroform partition coefficients ($\log(c_{water}/c_{chloroform})$) when including one molecule of the other solvent. E_{NA}, E_{tot}, E_{rel} are the average absolute error for nucleic acids only, all listed values and relative partition coefficients, respectively. An error of 1 in log P corresponds to an error of 5.74 kJ mol^{-1} in $\Delta\Delta G_{sol1-sol2}$. Values in parentheses are uncertainties in the last significant digit.

	solvent-solvent	Ade	Gua	Cyt	Thy	Ura	Trp	Tyr	Phe
AMBER-99	TIP3P+1Chl-Chl+1TIP3P	0.78 (5)	4.39 (5)	4.14 (9)	1.59 (1)	2.02 (2)	-2.60 (1)	-1.53 (3)	-3.38 (1)
AMBER-gaff	TIP3P+1Chl-Chl+1TIP3P	1.84 (7)	5.06 (4)	4.91 (2)	0.88 (3)	1.62 (6)	-2.55 (1)	-1.88 (1)	-3.34 (1)
CHARMM-27	TIP3P+1Chl-Chl+1TIP3P	0.23 (4)	3.74 (31)	2.48 (4)	-1.13 (1)	0.37 (3)	-3.74 (1)	-1.73 (1)	-3.92 (1)
GROMOS-45a4/53a6	SPC+1Chl-Chl+1SPC	-2.38 (7)	1.08 (17)	1.58 (7)	-1.25 (1)	-0.66 (2)	-2.26 (9)	-0.84 (12)	-3.57 (1)
OPLS-AA	TIP4P+1Chl-Chl+1TIP4P	-0.30 (3)	2.48 (7)	2.66 (10)	-1.88 (2)	-1.18 (1)	-3.39 (1)	-1.58 (1)	-3.78 (1)
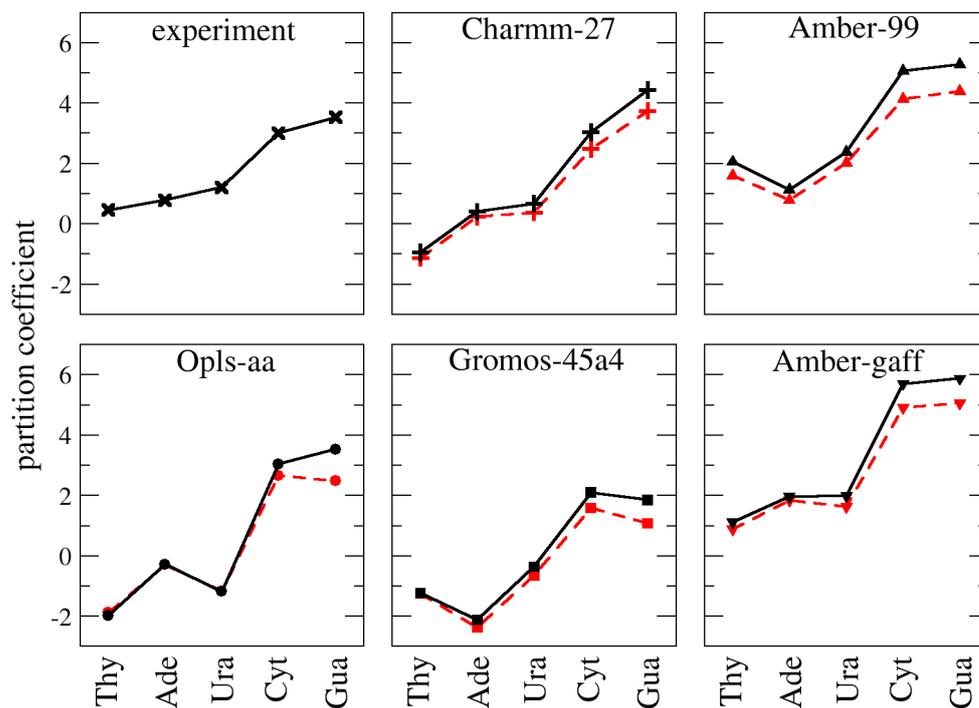
Figure S1: Nucleobase hydrophobicity scale. On the x-axis, the nucleobases are ordered according to their experimental hydrophobicity. A (partly) declining line means that the respective hydrophobicity scale deviates from the experimental. Both pure solvent (solid black) and including solvent-solvent solubility (dashed red) are shown. For each force field only the results for the native solvent model are shown, the other tested solvent models yielded very similar graphs.
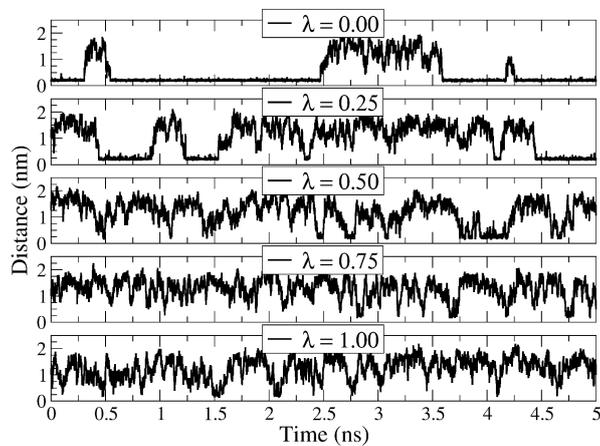
Figure S2: Sampling problem for one water and one 9-methylguanine molecule in chloroform. For various λ-values for charge disappearance the minimum distance between the water and the methylated guanine base are plotted. In all windows both contact (less than 0.25 nm) and separation (more than 0.25 nm) is sampled. Trajectories correspond to CHARMM-27 simulations, which resulted in the largest statistical error.