Supplementary Information:
Blind prediction of toluene/water partition coefficients using COSMO-RS: Results from the SAMPL9 challenge

Thomas Nevolianis[1], Raja A. Ahmed[1], Arnim Hellweg[2], Michael Diedenhofen[2], and Kai Leonhard[1]

[1]Institute of Technical Thermodynamics, RWTH Aachen University, 52062 Aachen, Germany.
[2]BIOVIA, Dassault Systèmes Deutschland GmbH, Am Kabellager 11-15, 51063 Cologne, Germany.

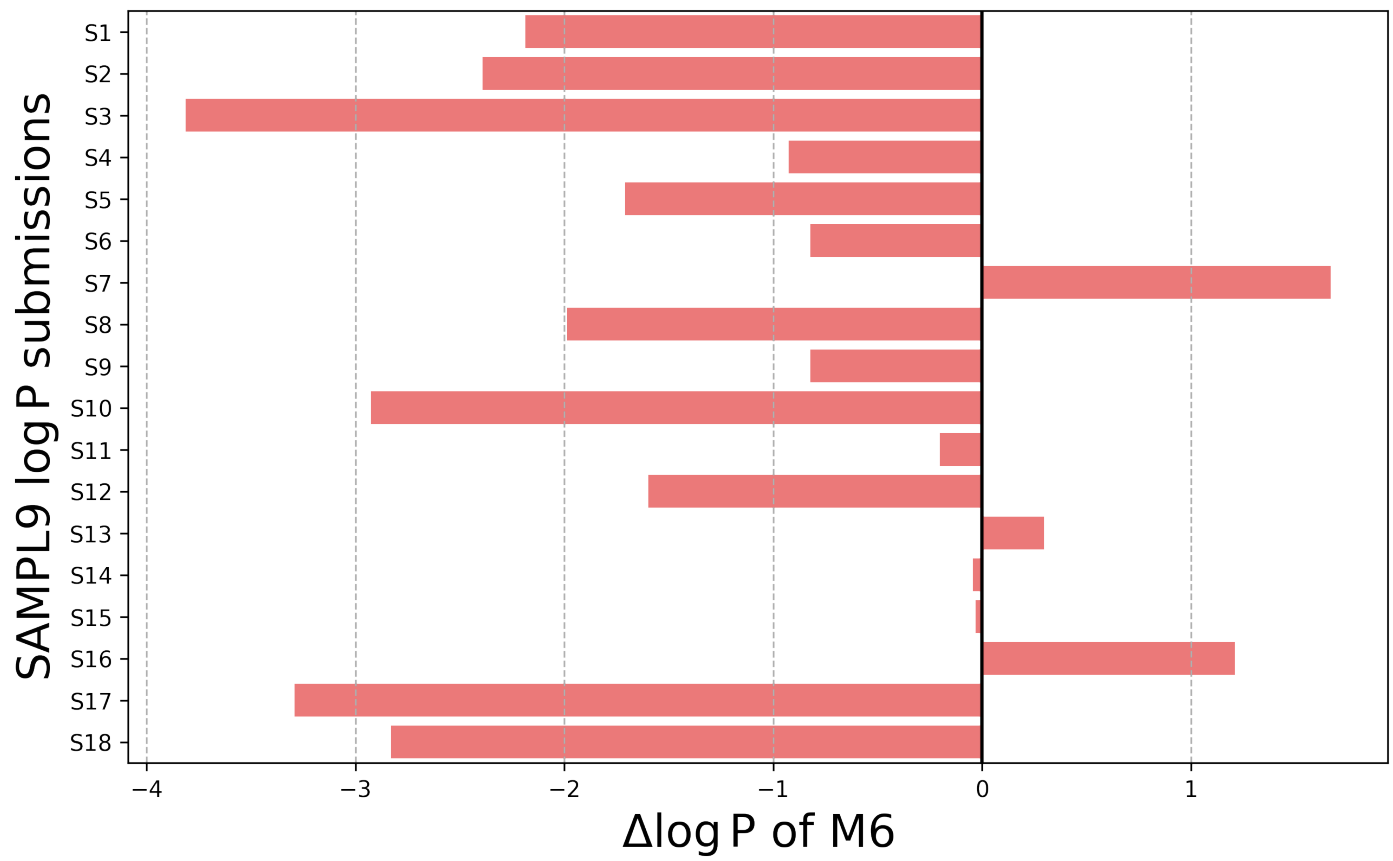## Contents

1 Performace of SAMPL9 submissions for M6

Figure S1 Distribution of $\Delta \log P$ values in SAMPL9 submissions for M6[1]. $\Delta \log P$ is the difference between each submitted prediction and experimental $\log P$ value.

2 COSMO-RS-DARE

For a comprehensive understanding of the COSMO-RS-DARE (COSMO-RS with Dimerization, Aggregation, and Reaction Extension) theory, we recommend referring to the original work by Sachsenhauser *et al.*[2]. Here, we will briefly touch upon the key aspects and motivations behind the development of COSMO-RS-DARE. The original goal of developing COSMO-RS-DARE was to accurately describe the concentration dependence of the activity coefficient of dimerizing species. This necessity arose because the traditional COSMO-RS method faced challenges in correctly handling the formation of multiple hydrogen bonds between two species, leading to inaccuracies in predicting liquid-phase properties in mixtures. COSMO-RS-DARE is an advanced method that addresses these limitations by considering dimerization effects. The method introduces a new conformer that represents a monomer in the dimerized state. The generation of this conformer necessitates additional COSMO calculations involving a dimer composed of two monomers of the dimerizing component. Post calculation, atoms and surface segments of one monomer are removed to represent a single monomer post-dimerization, creating a new surface segment that can only interact with similar segments.

To account for the "missing" interaction that arises due to the inability of COSMO-RS to handle multiple hydrogen bonds formation correctly, an interaction free energy is introduced and fitted to experimental data. This leads to the introduction of two new adjustable parameters, cH and cS, representing the enthalpic and entropic contributions to the dimerization interaction free energy, respectively. These parameters are essential in calculating the interaction free energy between the dimerization segments as,

$$\bar{E}_{\text{dim}}(\bar{d}_k, \bar{d}_l) = -2\Delta\bar{E}_{\text{conf}} + c_{\text{H}} - Tc_{\text{S}}, \tag{1}$$

where $\bar{E}_{\text{dim}}(\bar{d}_k, \bar{d}_l)$ is the interaction energy between two dimerization segments and $\Delta\bar{E}_{\text{conf}}$ is the COSMO-energy difference between the conformers the interacting segments originate from. The two interaction parameters are determined through fitting properties computed with COSMO-RS-DARE to experimental data, and are dependent on the number of conformers used for a dimerizing component. COSMO-RS-DARE views dimerization as a concentration-dependent reaction. Importantly, it should be noted that dimerization occurs at finite concentrations only. At higher mole fractions of a dimerizing component in a mixture, the likelihood of forming a dimer increases, as there are more dimerization segments available to interact.

The obtained interaction parameters values for dimerized M6 fall within the same range as those calculated for various organic molecules in studies reported by Cysewski *et al.*[3] and Sachsenhauser *et al.*[2].

3 Conformers of M6 dimers

(a) Conformer 1

(b) Conformer 2

(c) Conformer 3
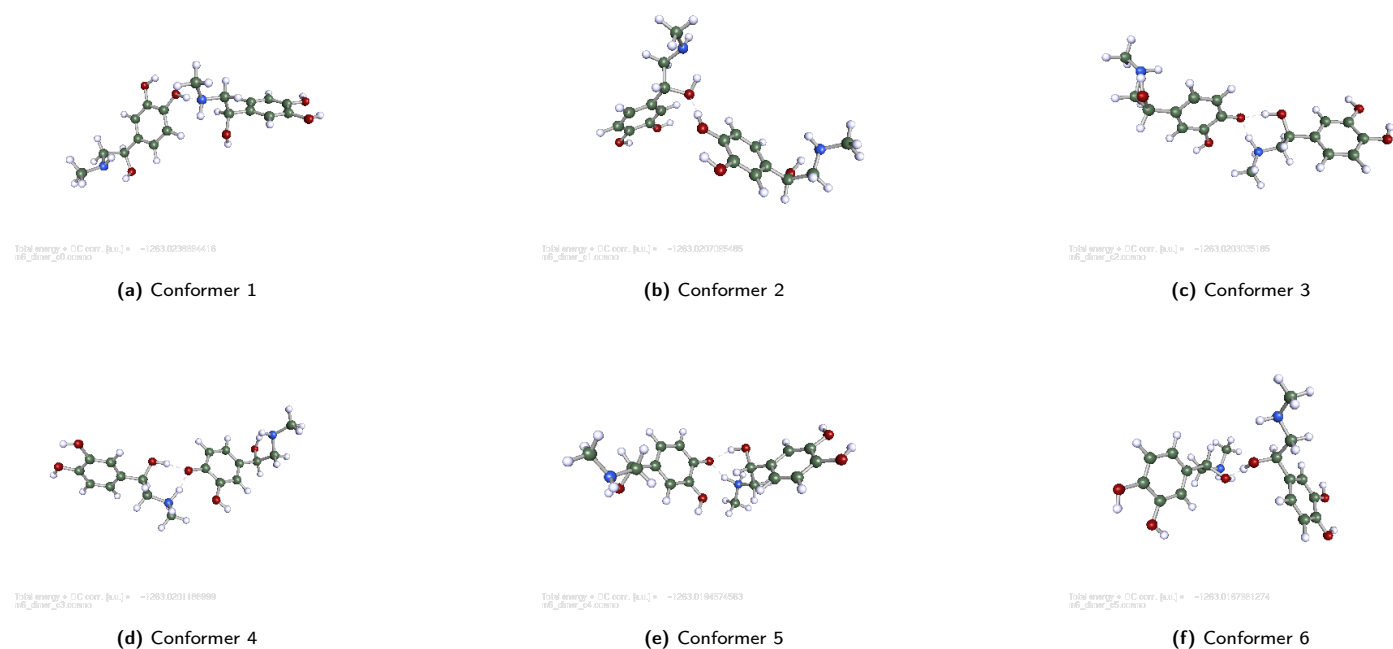
(d) Conformer 4

(e) Conformer 5

(f) Conformer 6

Figure S2 All conformers of M6 dimers in COSMO phase generated by COSMOconf at BP86/TZVPD parametrization with FINE COSMO cavity.

Name	Cosmo energy (kcal/mol)	Gas phase energy (kcal/mol)
Conformer 1	-5.1	-8.5
Conformer 2	-3.1	-8.3
Conformer 3	-2.8	-4.7
Conformer 4	-2.7	-4.6
Conformer 5	-2.3	-3.8
Conformer 6	-0.6	-2.8

Table S1 Cosmo and gas phase dimerization energies of all conformers of M6. The dimerization energies are the difference between the energy of the dimer and the energies of the monomers.

References

[1] M. Amezcua, D. L. Mobley and T. D. Bergazin, *samplchallenges/SAMPL9: 0.8*, 2023, `https://zenodo.org/record/7644720`.

[2] T. Sachsenhauser, S. Rehfeldt, A. Klamt, F. Eckert and H. Klein, *Fluid Phase Equilibria*, 2014, **382**, 89–99.

[3] P. Cysewski, T. Jeliński and M. Przybyłek, *Molecules*, 2022, **27**, year.