

**Supporting Material for Conformational and Structural
Stability Calculations of the Single Molecule and Hydrogen
Bonded Clusters of Para Aminobenzoic Acid in the Gas and
Solution Phases**

Ian Rosbottom*, Dimitrios Toroz, Robert B. Hammond and Kevin J. Roberts

School of Chemical and Process Engineering, University of Leeds, Leeds, LS2 9JT, UK

Communicating author email: i.rosbottom@leeds.ac.uk

Abstract

S1: COSMO-RS Theory

S2: Free Energy of Dimerisation

S3: M06 Solution Population Calculations

S4: Tetramer and Octamer Energetics

S5: Conclusions

S1: COSMO-RS Theory

The theory behind the COSMO-RS calculations has been covered in previous publications by Klamt and co-workers (refs). However, the equations behind the statistical thermodynamics calculations in this paper are summarised below.

The screening charge densities (σ) at the surface of the clusters, calculated from the DFT optimisation, are used to calculate the interaction of the clusters with the surrounding solvent molecules, using a statistical thermodynamic ensemble¹. The screening charge densities are used to divide the cluster up into segments of different polarity, where the interactions between these segments are then used to approximate the intermolecular interactions in solution. This is used to calculate a chemical potential of the cluster in the solution, which then combined with the free energy of the cluster in the continuum (E_{cosmo}), can be used to calculate the Boltzmann population of the clusters in the solutions. The equations associated with this method are shown below.

$$E_{MF}(\sigma, \sigma') = \alpha_{eff} \frac{\alpha'}{2} (\sigma + \sigma')^2 \quad (\text{S1})$$

Equation S1 calculates the misfit electrostatic interactions between the polar segments, where σ and σ' are the surface charges of the contacting segments, α_{eff} is the contact area between the two segments and α' is an adjustable empirical parameter. If a sufficiently polar hydrogen is in contact with another oppositely polar segment then a hydrogen bonding energy E_{HB} is calculated through equation S2.

$$E_{HB}(\sigma, \sigma') = \alpha_{eff} C_{HB} \min\{0; \min(0; \sigma_{donor} + \sigma_{HB}) \max(0; \sigma_{acceptor} - \sigma_{HB})\} \quad (\text{S2})$$

In equation (2), C_{HB} is the H-bonding strength coefficient, σ_{donor} is the screening charge density of the donor segment and $\sigma_{acceptor}$ is the screening charge density of the acceptor segment. The positive and negative charges on the respective segments must be greater than the threshold parameter σ_{HB} to activate the calculation of the H-bonding energy as a product of

the excess screening charge densities, i.e. $(\sigma_{donor} + \sigma_{HB})(\sigma_{acceptor} - \sigma_{HB})$. The less directional vdW interactions are calculated from equation (S3).

$$E_{vdW}(\sigma, \sigma') = \alpha_{eff}(\tau_{vdW} + \tau'_{vdW}) \quad (S3)$$

Where τ_{vdW} and τ'_{vdW} are element specific empirical parameters.

The chemical potential of the cluster is calculated from the interactions between the polar segments. The distribution of charges around the cluster is known as the σ -profile. The σ -profile of the system is a sum of the σ -profiles of all the components of the system weighted against their mole fractions, shown in equation (S4).

$$p_s(\sigma) = \sum_{i \in S} x_i p_i(\sigma) \quad (S4)$$

The chemical potential of a surface segment is described by equation (S5).

$$\mu_s(\sigma) = -\frac{RT}{\alpha_{eff}} \ln \left[\int p_s(\sigma') \exp \left(\frac{\alpha_{eff}}{RT} (\mu_s(\sigma') - e(\sigma, \sigma')) \right) d\sigma' \right] \quad (S5)$$

Where $e(\sigma, \sigma')$ is the sum of the energies calculated from equations (S1)-(S3).

The chemical potential of compound i in system S is defined by equation (S6).

$$\mu_i^s = \mu_i^{c,s} + \int p_i(\sigma) \mu_s(\sigma) d\sigma \quad (S6)$$

The integral of the term $\mu_s(\sigma)$ encompasses the surface interactions within the liquid. The term $\mu_i^{c,s}$ is a combinatorial contribution that arises from the different sizes and shapes of the solute and solvent molecules.

The population of a conformer (in this study the conformer refers to a cluster) j in a solvent system S containing k different conformers, π_j^s , is calculated from Equation S7, in which E_{COSMO} is the quantum-chemical total energy in the COSMO conductor, μ^s is the chemical potential of a conformer in solvent system S, w is an assigned multiplicity, k the Boltzmann constant and T the absolute temperature

$$\pi_j^S = \frac{w_j \exp\left\{-\frac{E_{\text{COSMO}}^j + \mu_j^S}{kT}\right\}}{\sum_k \omega_k \exp\left\{-\frac{E_{\text{COSMO}}^k + \mu_k^S}{kT}\right\}} \quad (\text{S7})$$

Since the chemical potential of a component depends on the population of its conformers, Equation 7 is iterated to self-consistency from an initial guess of the population based on the component's chemical potential being equated to zero.

These comparisons were carried out in a pairwise fashion hence all combinations of two dimers selected from the total of seven dimers (irrespective of order) were chosen and compared directly. A similar procedure was adopted for the tetramers and octamers. When, for example, two dimers were compared these were treated as two molecular conformations in the context of COSMOthermX.

S2: Free Energy of Dimerisation

The relative solution populations calculations were complimented by calculating the free energy of reaction of two monomers of PABA becoming each of the four identified dimers. This is shown for aqueous solution in Figure S1.

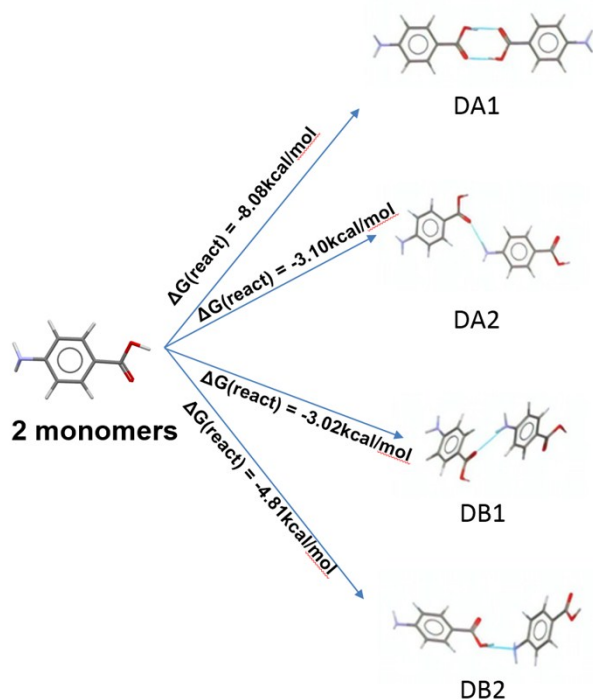


Figure S1: Free energy of formation of the DA1, DA2, DB1 and DB2 dimers in aqueous solution of 0.05 mole fraction of PABA monomers, calculated using the COSMO-RS approach

The free energy of formation of all the dimers was negative, suggesting that these are all plausible building blocks for the two forms. In line with the solution population calculations from the main text, the DA1 dimer was found to have the most favourable free energy of formation.

S3: M06 Solution Population Calculations

Since it has been highlighted that DFT calculations of organic materials can often poorly estimate the effects of dispersion interactions^{2,3}, a comparison of the results using the B3LYP and dispersion corrected M06-2X functional. Figure S2 shows the solution populations of the dimers optimised using the M06-2X functional.

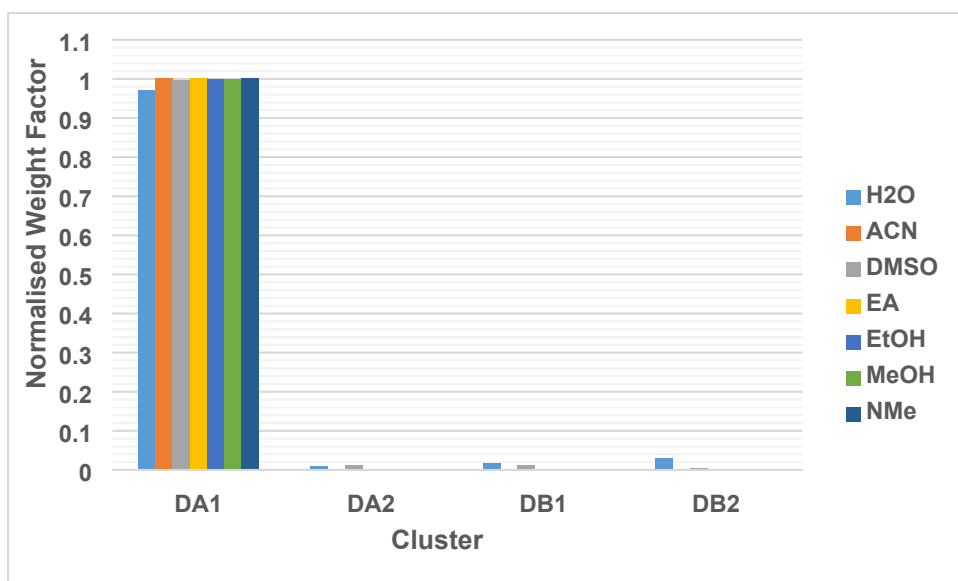
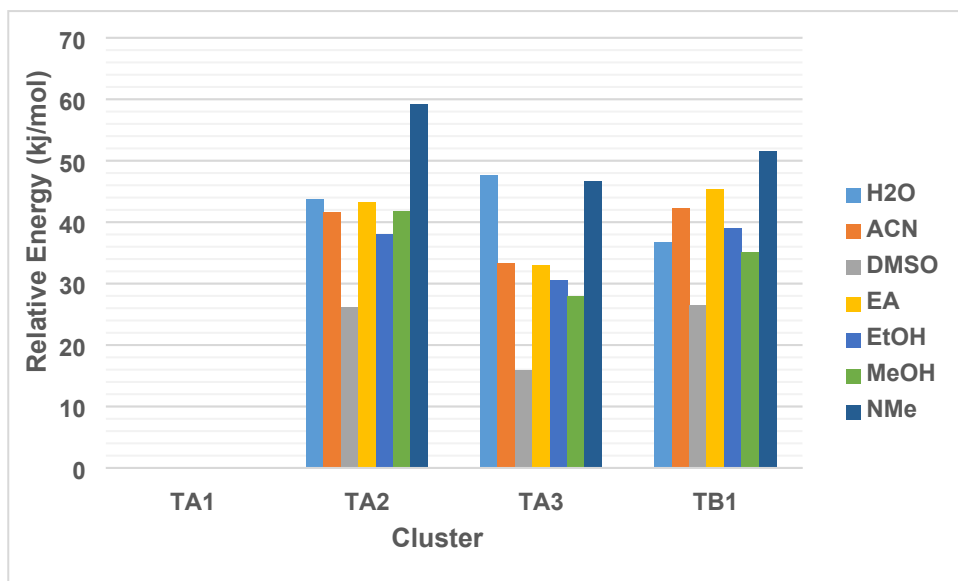


Figure S2: Normalised weight factors of the DA1, DA2, DB3 and DB4 clusters, optimised using the M06-2X/6-31G* approach

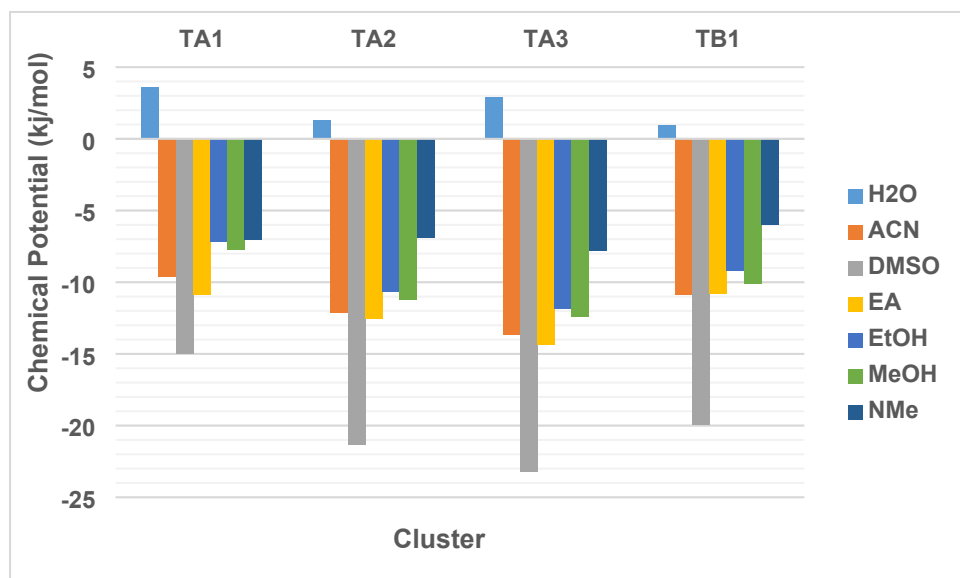
Figure S2 shows that the DA1 cluster was found to dominate the solution populations when optimised using the M06-2X functional. Though the DB1 and DB2 clusters have slightly higher populations in water, in comparison to the clusters optimised using the B3LYP functional, it is not enough to suggest that the use of the M06-2X functional makes a significant difference to the results. The solution populations for the tetramers and octamers calculated using the M06-2X functional were also found to be similar to those calculated using the B3LYP functional.

S4: Tetramer and Octamer Energetics

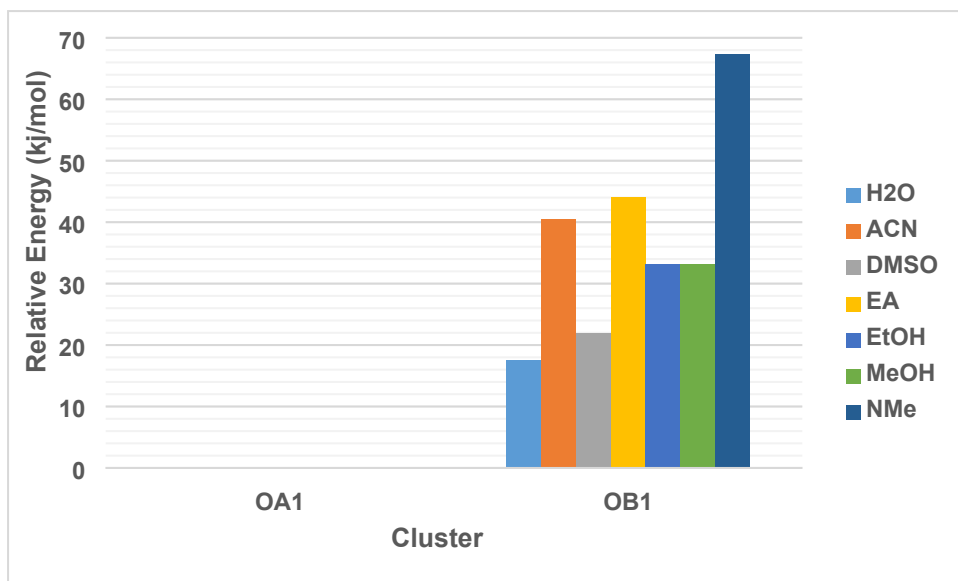
Figure S3 shows the energetics calculated for the tetramers and octamers from α - and β -PABA.



(a)



(b)



(c)

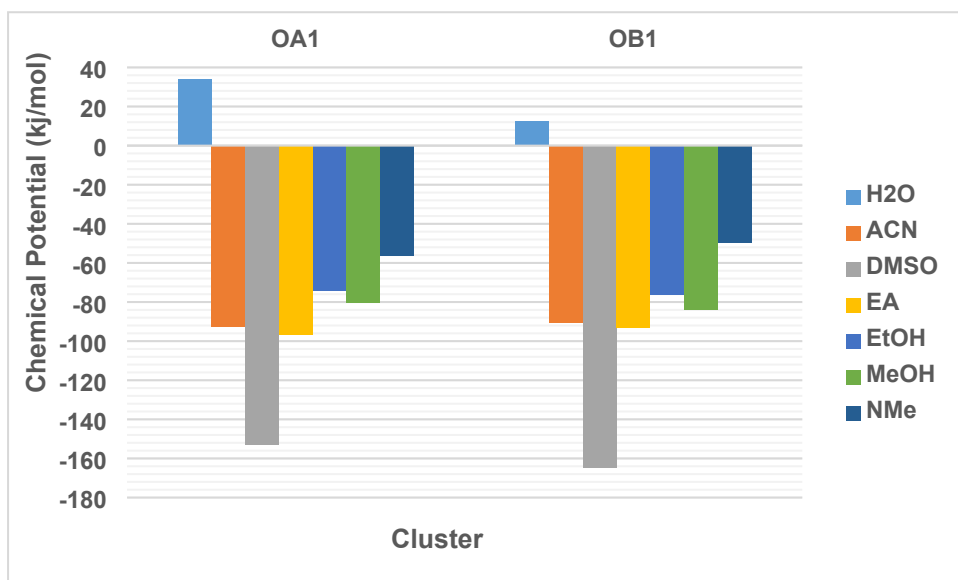


Figure S3: (a) Energy of the tetramers in the solvent continuum; (b) chemical potential of the tetramers in solution; (c) energy of the octamers in the solvent continuum; (d) chemical potential of the octamers in solution

Figure S3 shows that the domination of the solution populations by the tetramers and octamers derived from the α -form structure (Figure 10, main text) is due to the stability of such clusters

(d)

in the solvent continuum. Interestingly, the energy difference between the octamers of α - and β -PABA is less than between the tetramers of α - and β -PABA. This is found to be especially the case for water, where the OB1 energy in the continuum is less than 20kJ/mol less stable than OA1. Combined with the unfavourable chemical potential of the α -PABA tetramers and octamers results in the stabilities of the clusters of the two forms moving slightly closer. It would be interesting to simulate the stability of clusters of significantly higher sizes to see if the more extended network of β -PABA becomes stabilised at larger cluster sizes.

S5: Basis Set Superposition Errors

Table S1. Basis set super position error (BSSE) for the dimers R1 to R6 in six different solvents.

	DB1	DB2	DA1	DA2
H ₂ O	3.817	0.052	5.886	5.754
DMSO	3.717	0.086	5.870	5.586
Ethylacetate	5.994	0.041	6.124	6.347
Ethanol	4.014	0.051	5.808	5.678
Methanol	3.859	0.053	5.899	5.634
Nitromethane	3.809	0.054	5.897	5.614
Water	3.580	0.087	5.901	5.686
Gas-Phase	17.347	8.084	7.991	9.214

Table S1 shows the basis set superposition errors calculated for the seven dimers in the solvents considered. For some dimers the BSSE error is negligible, e.g. for cluster (DA3), whereas for some clusters the BSSE was found to be significant (around 6 kJ/mol) although not so large as to affect the results drastically. However, in the gas phase much larger BSSE values were found and this should be considered when performing calculations on molecular clusters in vacuum.

S5: Conclusions

The calculations presented in the supporting material suggest that, at least for a small molecule such as PABA, the different electronic structure theory approaches give very similar outcomes in terms of conformer stability and solution populations. Future studies on larger, complex molecules, along with clusters more dominated by dispersive interactions would be of interest to see if the same trends hold.

References

- (1) Klamt, A., Conductor-like screening model for real solvents - a new approach to the quantitative calculation of solvation phenomena. *Journal of Physical Chemistry* **1995**, 99, (7), 2224-2235.
- (2) von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D., Optimization of effective atom centered potentials for London dispersion forces in density functional theory. *Physical Review Letters* **2004**, 93, (15).
- (3) Thakkar, A. J., Small clusters of formic acid: Tests and applications of density functional theory with dispersion-correcting potentials. *Chemical Physics Letters* **2013**, 560, 71-74.
- (4) Coriani, S.; Marchesan, D.; Gauss, J.; Hättig, C.; Helgaker, T.; Jørgensen, P., The accuracy of ab initio molecular geometries for systems containing second-row atoms. *The Journal of Chemical Physics* **2005**, 123, (18), 184107.
- (5) Head-Gordon, M.; Pople, J. A.; Frisch, M. J., MP2 energy evaluation by direct methods. *Chemical Physics Letters* **1988**, 153, (6), 503-506.