How to Model Inter- and Intramolecular Hydrogen Bond Strengths with Quantum Chemistry
Supporting Information

Christoph A. Bauer

Contents

1 Description of the Data Sets

The inter- and intramolecular data sets are described here. All the 3D structures are optimized at the dispersion-corrected DFT level of theory (TPSS-D3(BJ)/def2-TZVP).

1.1 pK_{AHY}29 Data Set

The pK_{AHY}29 data set consists of 59 structures in total. The first structure is the reference molecule NMP, followed by 29 hydrogen bond donor molecules, followed by 29 complexes. Figure S1 shows two structures of such 1:1 complexes. As mentioned in the main manuscript, the conformers do not necessarily align, since the methodology followed is a single-conformer ansatz.

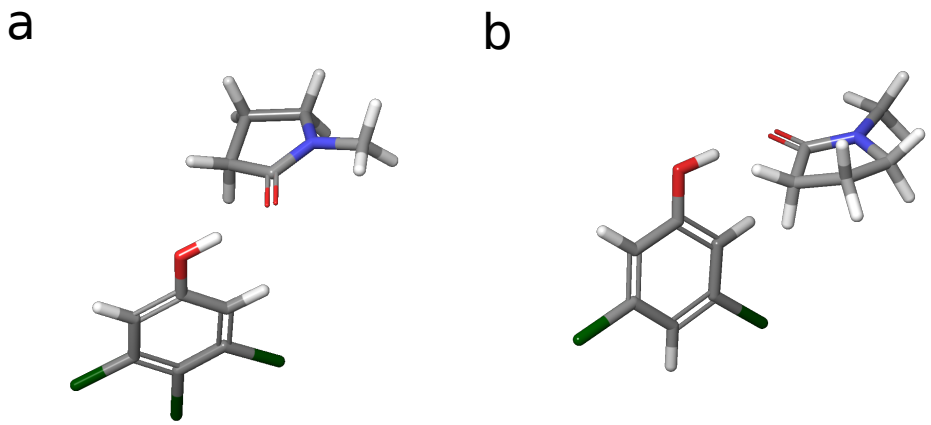

Figure S1: Optimized D structures of the complexes of (a) 3,4,5,-trichlorophenol and (b) 3,5-dichlorophenol with NMP as contained in the pK_{AHY}29 data set.

1.2 IMHB16

The IMHB data set consists of 32 structures in total. There are 16 IMHB structures and 16 open structures where there is no IMHB formation. As an illustration, Figure S2 shows the IMHB and open structures of OBC-11.

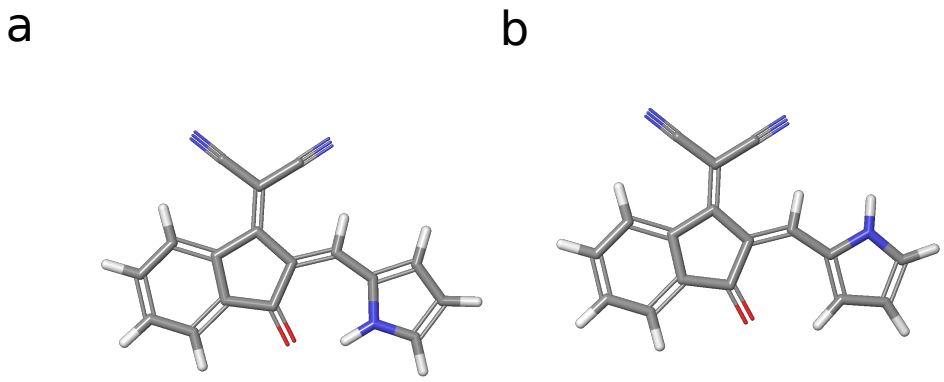

Figure S2: Optimized 3D structures of OBC-11 (a) in its IMHB structure (b) in its non-IMHB (open form) as contained in the IMHB16 data set.

For flexible molecules like OBC-15 and OBC-16, where the aliphatic rings can adopt multiple conformations, it is important to compare conformers that are maximally aligned. Figure S3 shows that it is indeed the case for the molecules after for OBC-16, a manual check of four conformers has been conducted.

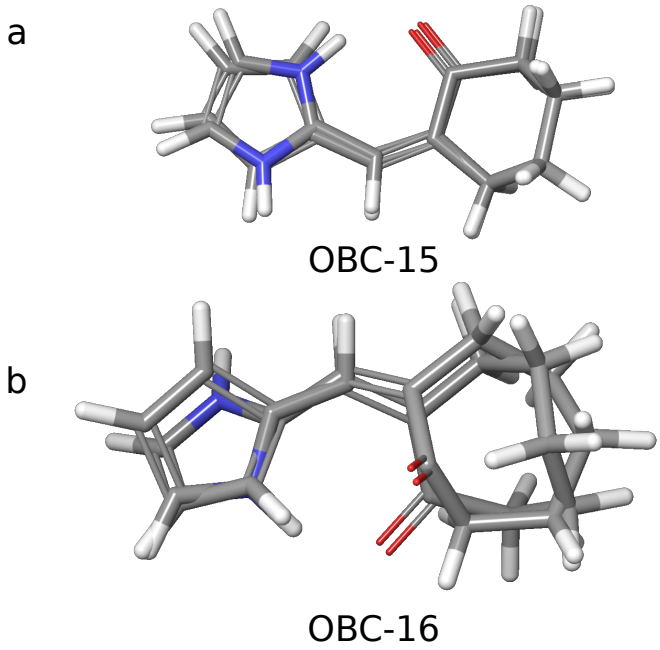

Figure S3: Aligned 3D structures of the IMHB and open forms of OBC-15 (a) and OBC-16 (b) as contained in the IMHB16 data set.

2 Molecular Properties

All properties contained by the molecules in the SDF files are described in this section. The energies given in units of E_h must always be used at the given precision. The energies in kJ mol^{-1} are reported at one or two decimal places.

2.1 Mol_ID

This is the unique identifier of each molecule. For pK_{AHY}29, the identifiers are of the forms 'NMP' (the reference acceptor molecule), 'pKAHY-XX' (the hydrogen bond donor molecules), and 'pKAHY-XX_c001' (the 1:1 complex structures with NMP). For IMHB16, the identifiers are of the form 'intraHB-OBC-XX' and 'openform-OBC-XX' for the intramolecularly hydrogen mbonded and non-intramolecularly hydrogen bonded forms, respectively.

2.2 Delta G_exp (kJ/mol)

This is the experimental target value in units of kJ mol^{-1}.

2.3 E_el (TPSS)

This is the total energy E_{tot} as calculated by quantum chemistry at the TPSS-D3(BJ)/def2-TZVP level of theory. This quantity is given in units of E_{h}

2.4 E_el (PW6B95)

This is the total energy E_{tot} as calculated by quantum chemistry at the PW6B95-D3(BJ)/def2-TZVP level of theory. This quantity is given in units of E_{h}

2.5 G_HO

These are the thermal corrections to each structure for the computation of ΔG_{HO}. This quantity is given in units of E_{h}

2.6 E_el (BP86)

This is the total energy E_{tot} as calculated by quantum chemistry at the BP86/def2-TZVP level of theory. This number is used as the gas phase energy for the computation of δG_{solv}. This quantity is given in units of E_{h}

2.7 E_el (SMD)

This is the total energy E_{tot} as calculated by quantum chemistry at the SMD(BP-86/def2-TZVP) level of theory. This number is used as the solution phase energy for the computation of δG_{solv}. For pK_{AHY}29, the solvent is carbon tetrachloride. For IMHB16, the solvent is chloroform. This quantity is given in units of E_{h}

2.8 dG_solv (kJ/mol)

This is the solvation free energy δG_{solv} in kJ/mol, computed at the SMD level of theory. Note that for the pK_{AHY}29 data set, a further shift of -22.6 kJ mol^{-1} to correct for standard conditions (1.0 M, CCl_4 solution) has to be added manually, see the main manuscript.

2.9 Name of HBD molecule

Only pK_{AHY}29 data set: The name of the hydrogen bond donor molecule.