Speciation of substituted benzoic acids in solution: evaluation of spectroscopic and computational methods for the identification of associates and their role in crystallization

Agris Bērziņš[*a], Aina Semjonova[a], Andris Actiņš[a], Matteo Salvalaglio[b]*

[a] - Faculty of Chemistry, University of Latvia, Jelgavas iela 1, Riga, LV-1004, Latvia

[b] - Thomas Young Centre and Department of Chemical Engineering, University College London, Torrington Place, London, UK

ABSTRACT

Self-association of four benzoic acid derivatives 2-chloro-4-nitrobenzoic acid, 2-methyl-4-nitrobenzoic acid, 3-hydroxybenzoic acid and 2,6-dimethoxybenzoic acid in solution was investigated using spectroscopic measurements (FTIR, ^{1}H and ^{13}C NMR spectroscopy) and molecular simulation methods. Based on the formation of hydrogen bonds, solvents can be divided in two groups: apolar solvents or solvents with low hydrogen bond acceptor propensity, in which the benzoic acids form hydrogen-bonded dimers, and solvents with hydrogen bond acceptor propensity $\beta > 0.3$, interacting with the carboxylic group of benzoic acids, and thus screening its interaction in the formation of self-associates. The formation propensity and structure of self-associates stabilized by weak interactions, such as $\pi\cdots\pi$ stacking, and $CH_3\cdots\pi$ interactions, however, are determined by the substituents in the benzene ring. Despite all of the studied compounds are polymorphic, in none of the cases an unequivocal structural link between self-associates present in solution and the crystal form was observed.

1 INTRODUCTION

Crystallization from solution is the most widely used method for the large-scale separation of solid materials in the chemical, pharmaceutical and food industries. Nucleation is a crucial step in crystallization processes, as it decisively affect the polymorphic outcome, morphology and crystal size[1]. Obtaining control of the nucleation stage is thus particularly important when compounds can yield more than one crystal form. Despite the progress over the last decades towards understanding different aspects of nucleation provided by the increasing accessibility of new experimental techniques and computational methodologies [2-4], the field is still far from the possibility to predict the outcome and mechanism of nucleation from first-principles, as recently highlighted by a remark by Cruz Cabeza et al. stating that "nucleation is perhaps the least well understood step in the area of polymorphism"[5].

It is well documented in the literature that the polymorphic outcome of crystallization is dependent on multiple process variables, either affecting the relative thermodynamic stability of polymorphs or impacting their nucleation kinetics. These include temperature and pressure[6-9], clearly linked to polymorph stability, as well as solvent choice[10-13], and even the presence of impurities[14]. For these reasons, the use of additives or surface templates have been considered as potential mechanisms for polymorph control[15-17].

Solution chemistry studies have shown that the self-associates present in the liquid parent phase can act as prenucleation aggregates, and their intermolecular interaction motifs can be carried over to crystals[18-22]. The most widely studied pre-nucleation aggregates are hydrogen bonded associates[23-25], although recently attention has been drawn to the importance of aromatic and other weaker interactions[26-29], which in certain systems can play a significant role or even be the driver of the nucleation process[30-31]. Nevertheless, counterexamples exist where polymorphs precipitating from solution are not apparently related to the interactions stabilizing self-associates present in the mother phase or are governed by other factors[32-35]. Finally, in several cases the link between self-associates and polymorphic outcome of the nucleation process is rather vague, and additional details on the nucleation mechanism are necessary in order to develop a rational understanding of the impact of solute speciation on polymorph selection[27, 36-37]. The solvent can also play a key role, affecting the structure of solute self-associates, and impacting on the nucleation process[4, 38-40].

Different experimental methods have been used to investigate solute self-association, including NMR spectroscopy based on the shift of ^{1}H [18, 20, 41-42] or ^{13}C peaks[27, 37], multidimensional spectroscopy[37, 43-44], the measurement of translational diffusion coefficients or relaxation times T_1 [43-44], bands in IR[19, 21, 25] or Raman spectra[4, 22], by applying VCD spectroscopy[45-46], by measuring change of UV absorption[11, 23, 47], or by using molecular simulation methods, including *electronic structure* calculations[35, 40] and molecular dynamics simulations[22, 27, 48-50]. However, a very limited

number of studies have presented a detailed comparison of these techniques, carefully analyzing benefits and drawbacks of each method by studying several compounds.

In this study we explore the self-association of four substituted benzoic acids (2-chloro-4-nitrobenzoic acid 2C4NBA, 2-methyl-4-nitrobenzoic acid 2Me4NBA, 3-hydroxybenzoic acid 3OHBA and 2,6-dimethoxybenzoic acid 26MeOBA) in multiple solvents using information obtained from several experimental and computational methods.

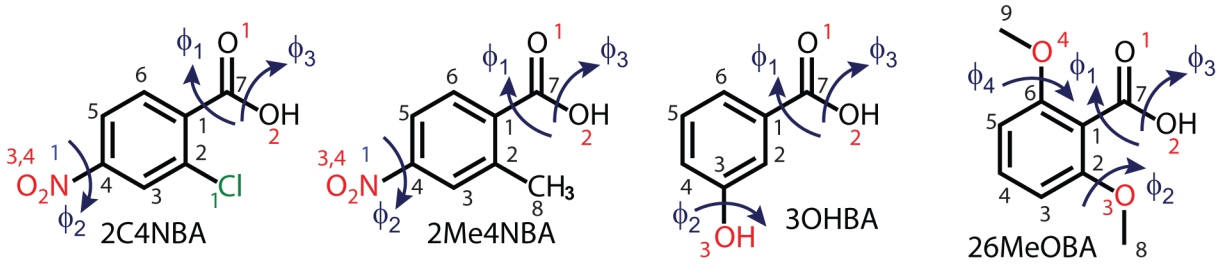


Figure 1. *Molecular structure of 2C4NBA, 2Me4NBA, 3OHBA and 26MeOBA with numbering of non-hydrogen atoms and labelling of flexible torsion angles used in the text. Labels of carbon in black, oxygen in red, nitrogen in blue, chlorine in green. Numbering of hydrogen atoms is given in Figure S1.*

2C4NBA has been identified as an active compound for the treatment of immuno-deficiency diseases, including the HIV infection, and has been reported to form 2 polymorphs[51] as well as several solvates[52]. Polymorph I is stable at room temperature and above[52], and is almost exclusively obtained from cooling crystallization, while in evaporation crystallization polymorphs I and II are concomitantly obtained[53]. Both polymorphs crystallize in $P2_1/c$ space group containing the carboxylic acid dimer motif, with Z′=1 in I and Z′=4 in II.

2Me4NBA exists in at least 5 polymorphs, with cooling crystallization producing mostly form I, while evaporation crystallization leads to the precipitation of forms I and II, as well as mixtures of polymorphs containing also forms III and V[53]. Polymorph I crystallizes in the $P2_1/c$ space group and polymorph II in the P-1 space group, both containing the carboxylic acid dimer motif.

3OHBA exists in forms I and II[54] often crystallizing concomitantly. The room temperature thermodynamically stable form I[54] crystallizes in the $P2_1/c$ space group containing the carboxylic acid dimer motif, whereas the metastable form II crystallizes in the $Pna2_1$ space group featuring intermolecular hydrogen bond chains consisting of alternate COOH and OH-groups[55]. Forms I and II are monotropically related[54], and it has been reported that the selective crystallization of either one of them can be achieved by solvent selection[54, 56-58].

Three polymorphs have been reported for 26MeOBA. The most stable, form I, displays a $P2_12_12_1$ space group featuring a catemer motif formed by COOH groups with hydrogen atom in *anti*

configuration[59-60]. Form II, obtained in water, in the presence of phenylboronic acid, crystallizes in the $P4_12_12$ space group also characterized by a non-planar conformer in the unit cell but containing carboxylic group with hydrogen atom in the *syn* configuration, involved in carboxylic acid dimer motifs[61]. It has been isolated and characterized only once – showing the hallmarks of disappearing polymorphism. Form III, crystallizing in $P2_1/c$ space group also containing carboxylic group with hydrogen atom in *syn* configuration and the carboxylic acid dimer motif has been reported recently[62-63].

The aim of this study is to combine readily available experimental techniques (FTIR, ^{1}H and ^{13}C NMR spectroscopy) and computational methods (MD simulations and DFT calculations) to identify the dominant associates of benzoic acid derivatives present in different solvents. By analyzing the case of 2C4NBA, we tuned experimental approaches and computational tools to obtain complete information about speciation in solution. Based on these results, selected experiments and calculations were carried out for the rest of the compounds. With the information about dominant associates present in solution in hand the relation between self-associates present in solution and the polymorphic outcome obtained during crystallization is critically discussed.

2 EXPERIMENTAL SECTION

2.1 Materials, crystallization experiments and crystal form identification

2C4NBA (98%, Alfa Aesar), 2Me4NBA (98%, Fluorochem), 3OHBA (99%, Acros Organics), 26MeOBA (99%, Alfa Aesar) and organic solvents of analytical grade were purchased from commercial sources and used without further purification. Deuterated solvents were purchased from Eurisotop with ≥99.9% D. Evaporation crystallization experiments were performed by preparing concentrated solutions at room temperature, filtrating and transferring the obtained solution to the selected glassware and allowing complete evaporation at ambient temperature. These experiments were performed in a fume hood at ~25 °C by evaporating 2 – 10 mL of solutions from a) Petri dish, b) 10 mL glass flasks with 14/20 joint and c) 100 mL glass flasks with 29/32 joint. For cooling crystallization concentrated solutions were prepared at elevated temperature (40–80°C, depending on the boiling point of the solvent) and the solution was filtered into a flask, which was closed and cooled without stirring by inserting in a thermostat set at –10°C. The obtained solid products were collected, air dried, and analyzed by recording Powder X-ray diffraction (PXRD) pattern.

PXRD patterns were measured at ambient temperature on a D8 Advance (Bruker) diffractometer using copper radiation (CuK_α) at the wavelength of 1.54180 Å, equipped with a LynxEye position sensitive detector. The tube voltage and current were set to 40 kV and 40 mA. The

divergence slit was set at 0.6 mm and the antiscatter slit was set at 8.0 mm. The diffraction patterns were recorded using a 0.2s/0.02° scanning speed from 3° to 35° on 2θ scale.

2.2 Study of association using FTIR spectroscopy and NMR spectroscopy

For the experimental study of association several solvents with different physicochemical properties were selected. Solutions with concentrations ranging from 1000, 500 or 250 mM down to 1 mM (for NMR) or 8 mM (for IR spectroscopy) were prepared (see details in the Supporting Information). For systems, where the desired largest concentration could not be achieved due to poor solubility, a range up to saturation was covered. For the IR spectroscopy measurements solvents without carbonyl groups were selected. Fully deuterated solvents (DMSO-*d6*, THF-*d8*, $CDCl_3$, CD_3CN, methanol-*d4*, toluene-*d8* and acetone-*d6*) were used to prepare solutions for measuring NMR spectra. The most concentrated solutions in each solvent were prepared directly, while the rest were obtained by subsequent dilutions. Analytical balance (±0.1 mg), micropipettes (±1 μL) and class A measuring pipettes and volumetric flasks were used for solution preparation. All spectra of the solutions were recorded shortly after their preparation (IR spectra and ^{1}H NMR spectra were recorded within several hours, most of the ^{13}C NMR spectra were recorded within 1 day, with exception of some low concentration solutions of 3OHBA and 26MeOBA which were recorded within 1 week).

For all solutions, except DMSO ones, infrared spectra were collected on a Frontier FTIR (PerkinElmer) spectrometer using a liquid transmission cell with KBr windows and a path length of 100 μm (except for several occasions where a 25 μm path length was used). For DMSO solutions a Universal ATR Sampling Accessory with a diamond window was used instead to avoid dissolving the KBr windows. The spectra were recorded from 650 to 4000 cm^{-1} at a 2 cm^{-1} spectral resolution with 16 scans. Spectral lines in the carbonyl group stretching region (1650 – 1850 cm^{-1}) were decomposed in individual components using the OriginPro 9.0 software. It was determined that the peak profiles in most cases are best described with a Lorentzian function, see Supporting Information. The number of individual components present were evaluated based on the obtained fit.

^{1}H and ^{13}C NMR spectra were recorded using a Bruker Fourier 300 MHz spectrometer using residual solvent as an internal standard for chemical shift referencing. The number of scans was selected to obtain an acceptable signal-to-noise ratio (8 – 256 for ^{1}H and 64 – 61440 for ^{13}C spectra). NMR spectra were recorded at ambient temperature. The spectra were processed using a MestReNova 14.1.1. Chemical shifts were allocated using the splitting information, chemical shielding values obtained in *ab initio* calculations as well as chemical shift values predicted using the MestReNova 14.1.1 software.

2.3 Ab initio calculations

Geometry optimizations, both in the gas phase and in polarizable continuum model (PCM) solvent, as well as the calculation of vibrational frequencies and NMR shielding tensors (using the gauge-independent atomic orbital (GIAO) method) were performed with *Gaussian 09*[64] at the M06-2X/6-31++G(d,p) level of theory. The molecular geometries used as initial guess in DFT calculations were obtained from experimental crystal structures (VOLZEC for 2C4NBA[51], QUPROL for 2Me4NBA[53], BIDLOP for 3OHBA[55], DMOXBA01[59] and DMOXBA03[61] for 26MeOBA). The calculated vibration frequencies were scaled by a factor 0.947[65]. ^{1}H chemical shifts were referenced to shielding values of tetramethylsilane (TMS) obtained in identical calculations. Electrostatic potential surfaces were generated from electron densities calculated in vacuum using Multiwfn 3.7 [66] and plotted using Visual Molecular Dynamics (VMD)[67].

The geometries of associates were prepared by isolating molecules containing the desired relative arrangement and intermolecular interaction motifs, such as hydrogen bonding or aromatic stacking, from the crystal structures, and modifying the relative arrangement of molecules, if necessary. The geometry of associates as prepared and that after the geometry optimization is given in the Supporting Information. Also in this case, geometry optimizations and frequency calculations were performed in *Gaussian 09*[64] at M06-2X/6-31++G(d,p) level. For calculation of association Gibbs energy (ΔG_{ass}) in solvent more accurate *in vacuo* energy and solvation energy were calculated for the optimized gas phase geometry at M06-2X/aug-cc-PVDZ level using the Solvation Model Based on Density (SMD)[68]. More details about the approach used for these calculations can be found in the literature[35, 40].

Pairwise intermolecular interaction energy calculations of crystal structures were performed in Crystal Explorer 17.5 at the B3LYP-D2/6-31G(d,p) level[69].

2.4 Molecular dynamics simulations

Unbiased molecular dynamics (MD) simulations were performed to investigate the molecule behavior in solution using an explicit representation of the solvent. Benzoic acid derivatives and solvent molecules were modelled with the General Amber Force Field (GAFF)[70]. The initial molecular geometries of the studied benzoic acid derivatives were obtained from the experimental crystal structures (see Section 2.3). The topologies of the studied benzoic acid derivatives and 2-propanol were generated using the standard GAFF[70] procedure, with AmberTools19[71]. Geometries and force field parameters for the remaining solvent molecules were obtained from the Virtual Chemistry database[72-73]. For all solvents we verified that GAFF allows to reproduce equilibrium densities consistent with experimental data (see Table S13).

The initial configuration used in MD simulations was prepared by randomly inserting 20-21 benzoic acid derivative molecules in a cubic box with box edges of 6 – 6.5 nm, and appropriately

filled with solvent molecules, resulting in a final concentration of ~0.1 – 0.15 M. Additional details are provided in the Supporting Information. MD simulations were carried out using Gromacs 5.1.4 [74]. The initial configuration was initially energy minimized with a steepest descent algorithm imposing an upper limit on the residual force of 1000 kJ mol^{-1} nm^{-1} and then equilibrated in the canonical (NVT) and isothermal-isobaric (NPT) ensembles for 100 ps. For each simulation a time step of 2 fs was used. The production runs were carried out in the NPT ensemble for 100 ns at a pressure of 1.0 bar and temperature of 300 K, using the Parrinello-Rahman barostat[75] and the Bussi-Donadio-Parrinello thermostat[76]. System coordinates were saved every 10 ps for further analysis.

The analysis of the distribution of intermolecular benzoic acid derivative distances and their relative orientations was performed using PLUMED 2.5[77-78]. The solute center of mass and the vector connecting atoms C4 to C7 (see Figure 1) were respectively used to define position and absolute orientation of solute molecules. An in-house python script[27] was used for collecting and plotting the data. The analysis of probability density distribution of hydrogen bonded dimers and hydrogen-bonded associates with solvent molecules were performed using the VMD[67] *Hbonds* plugin by counting the number of hydrogen bonds between the respective oxygen (nitrogen for acetonitrile) atoms having geometry corresponding to a conventional hydrogen bond (the distance donor(D)-acceptor(A) is less than 3.0 Å and the angle D-H-A is 180±20°). The residence time of solvent molecules hydrogen-bonded to a solute molecule was obtained postprocessing the atomistic trajectories with Plumed 2.5 [77-78], using a distance-only criterion for the identification of the bonded state, with a cutoff distance of 3 Å.

3 RESULTS AND DISCUSSION

3.1 Determination of species present in solution using FTIR spectroscopy

We recorded FTIR spectra of the benzoic acid derivative solutions of different concentration in a number of solvents (12 for 2C4NBA, 8 for 2Me4NBA, 7 for 26MeOBA, and 5 for 3OHBA). Although formation of associates involving hydrogen bond formed by the carboxylic group in general can be identified using several characteristic bands in IR spectra, including regions of carbonyl group antisymmetric stretching (1740 – 1660 cm^{-1}), carboxylic –OH···O– out-of-plane wag (960 – 875 cm^{-1}) characteristic for the carboxylic acid dimer, and the –O–H stretching (3500 – 2500 cm^{-1}),[25] the first region was the only one in which we detected bands that could be clearly interpreted as connected to association phenomena, see analysis of FTIR spectra of solid phases and solutions of 2C4NBA as well as *ab initio* predicted vibration frequencies in Supporting Information.

Overall, we observed two distinct situations within the frequency range characteristic of C=O stretching. In a first group of solvents (acetonitrile, THF, 1,4-dioxane and additionally DMSO and methanol for 2C4NBA, methyl *tert*-butyl ether (MTBE) for 2Me4NBA, DMSO for 3OHBA and DMSO and 2-propanol for 26MeOBA) we observe a single distinguishable peak maximum or two closely located overlapping peaks, see spectra of 2C4NBA in THF in Figure 2a. The separation of this region of the spectrum in individual peaks shows that this region is best described by two overlapping peaks of different intensity for 2C4NBA (see Figure 2b), 26MeOBA and 2Me4NBA, and a single peak for 3OHBA.

In case of the two overlapping peaks, the peak area and height ratio is not dependent on concentration (for an example see Figure 2c), and for 2C4NBA and 2Me4NBA the wavenumber differs by 6 – 20 cm^{-1} and for 26MeOBA by 18 cm^{-1} in acetonitrile, DMSO and nitromethane and 25 cm^{-1} in THF and 1,4-dioxane. This shows that the two peaks do not appear because of the presence of monomer/hydrogen bonded dimer equilibria[25, 45, 79] but rather because of the simultaneous presence in solution of two conformational isomers (for 2C4NBA having ϕ_1 of ~30° (conformer C1) and ~160° (conformer C2), and for 26MeOBA having ϕ_3 = ~0° (conformer C1) and 180° (conformer C2)), see Figure 1 for a definition of torsion angles), which in general is confirmed by the position difference matching well with *ab initio* calculated differences in the C=O stretching for the different conformers. Details on peak positions, width, area ratio and other characteristics as well as information on peak position and relative energy differences for different species obtained in *ab initio* calculations is available in Supporting Information.

Solutions of 2-propanol (for 2C4NBA, 2Me4NBA and 3OHBA), MTBE (for 2C4NBA), acetonitrile (for 2Me4NBA) and nitromethane (for 2C4NBA and 2Me4NBA) display behavior similar to this first group, as in these solvents the single distinguishable peak maximum in the C=O stretching region is best described with multiple overlapping peaks. The positions and other characteristics for the highest frequency peak (in case of 3OHBA) or peaks (in case of 2C4NBA and 2Me4NBA) match those observed in the solvents of the first group, and the ratio between areas of the peaks is concentration independent.

For a second group of solvents (chloroform, dichloromethane, toluene and *o*-xylene) we observe two distinguishable peak maximums, see spectra of 2C4NBA in dichloromethane in Figure 2d. Separation of this spectral region in individual peaks show that for 2C4NBA this in fact is best described with 4 (3 for chloroform) overlapping peaks of different intensity, see Figure 2e. The peak area and height ratio for peaks of the highest and lowest wavenumber (1/4) is concentration dependent (see Figure 2f) while that of the adjacent peaks 1/2 and 3/4 is not (with exception of peaks 3/4 in dichloromethane). The wavenumber of the two closest individual peaks differ by 10 – 20 cm^{-1} (for 3 and 4) and 2 – 6 cm^{-1} (for 3 and 4), while the wavenumber difference between peaks 1 and 4 (the most intense individual peaks) is 40 – 50 cm^{-1}. The peak area ratio concentration dependence and the wavenumber difference of 40 – 50 cm^{-1} show that the appearance of two peak groups 1/2 and 3/4 results from the presence of monomers and hydrogen

bonded dimers in equilibrium[25, 45, 79]. The presence of more than one peak in each group results from the presence of different conformational isomers of solute molecules either in their monomer state (peaks 3 and 4) or involved in the formation of self-associates (peaks 1 and 2) in solution. Also for 2Me4NBA and 26MeOBA in chloroform and dichloromethane we observe two distinguishable peak maxima with position differing by ~35-40 cm^{-1} and peak area ratio linearly dependent on the concentration (see Supporting Information), showing that the appearance of two peaks is resulted by presence of monomer/hydrogen bonded dimer equilibrium. For 3OHBA because of the low solubility (<10 mM) we did not record FTIR spectra in any of these solvents.

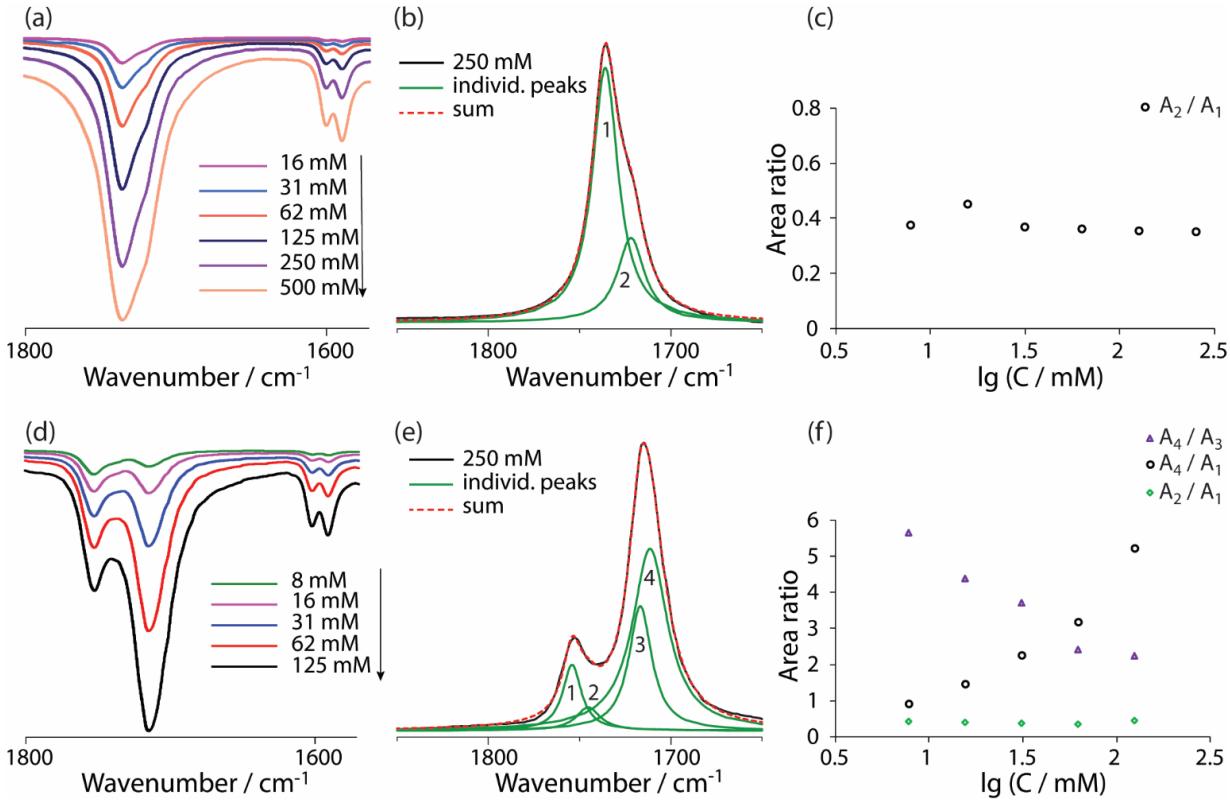


Figure 2. *IR spectra C=O stretching region of 2C4NBA solution of different concentration in THF (a) and dichloromethane (d). Separation of this spectral region in individual peaks for 250 mM solution in THF (b) and dichloromethane (e). Area ratio of the individual peaks in C=O stretching region for THF (c) and dichloromethane (f).*

Based on the analysis of the IR spectra, we can conclude that in chloroform, dichloromethane, toluene and *o*-xylene (Group 2 solvents) 2C4NBA, 2Me4NBA and 26MeOBA form hydrogen bonded dimers, that are in equilibrium with monomers. Moreover, for 2C4NBA each of these two states can be characterized by up to 2 different conformers. In contrast, in the remaining solvents (Group 1 solvents) no hydrogen bonded dimers can be detected for none of the studied benzoic acid derivatives. Instead, to some extent, association with solvent tends to shift the C=O

stretching peak to lower frequencies, with the most notable redshift observed in DMSO and methanol, while the smallest one detected in acetonitrile and nitromethane. In these solvents 2C4NBA and 26MeOBA exist in two of its conformers. The summary of the peak positions and determined identity of species present in solution of the studied benzoic acid derivatives is given in Figure 3.

We note that the additional peak in 2-propanol appears at frequency ~20 – 30 cm^{-1} lower than peaks present in other Group 1 solvents, suggesting that the lower frequency peak could correspond to associates where a solvent molecule acts as hydrogen bond donor, which would result in larger redshift of C=O stretching peak compared to formation of associates where the solvent is the hydrogen bond acceptor. In contrast, the appearance of an additional low frequency peak in nitromethane for 2C4NBA and 2Me4NBA, acetonitrile for 2Me4NBA and MTBE for 2C4NBA is not yet completely clear.

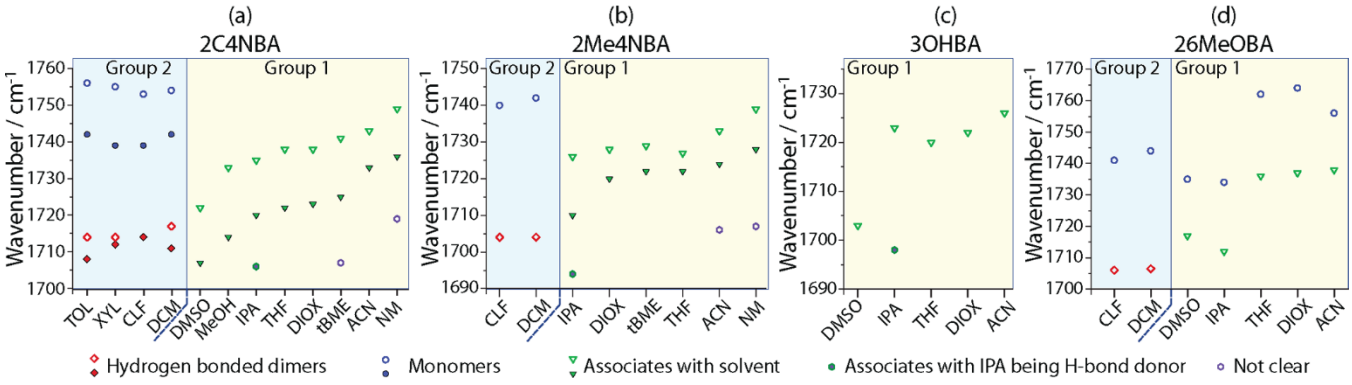


Figure 3. *Summary of individual peak positions in C=O stretching region in solution of 2C4NBA (a), 2Me4NBA (b), 3OHBA (c) and 26MeOBA (d) in different solvents, with designation of each individual peak to the corresponding species present in the solution.*

3.2 Determination of species present in solution using NMR spectroscopy

In part of the solvents in which FTIR measurements were carried out we recorded ^{1}H NMR and ^{13}C NMR spectra of 2C4NBA, 26MeOBA and 3OHBA at different concentration. These included fully deuterated DMSO, THF, chloroform, DCM, acetonitrile, methanol and toluene, as well as acetone. All except for the DCM were used for recording ^{1}H NMR spectra of 2C4NBA, while part for recording ^{13}C NMR spectra of 2C4NBA and ^{1}H and ^{13}C NMR spectra of 26MeOBA and 3OHBA. Based on the concentration dependence of ^{13}C and also ^{1}H chemical shifts these solvents can again be classified in two groups.

In the Group 2 solvents increase of the concentration results in significant increase of the ^{13}C chemical shift of C7 (for an example see data in $CDCl_3$ for 2C4NBA in Figure 4) and also gradual change of the ^{1}H chemical shift for hydrogen atoms in the benzene ring (in $CDCl_3$ measured for 2C4NBA and 26MeOBA, in DCM-*d2* for 26MeOBA and, although not clearly

systematic for all the data, also in toluene-*d8* for 2C4NBA, as shown in the Supporting Information). This is fully consistent with the formation of hydrogen bonded dimers in these solvents (*ab initio* calculated difference of C7 chemical shift of 2C4NBA in hydrogen bonded dimer and monomer is 4-6 ppm depending on the particular conformer).

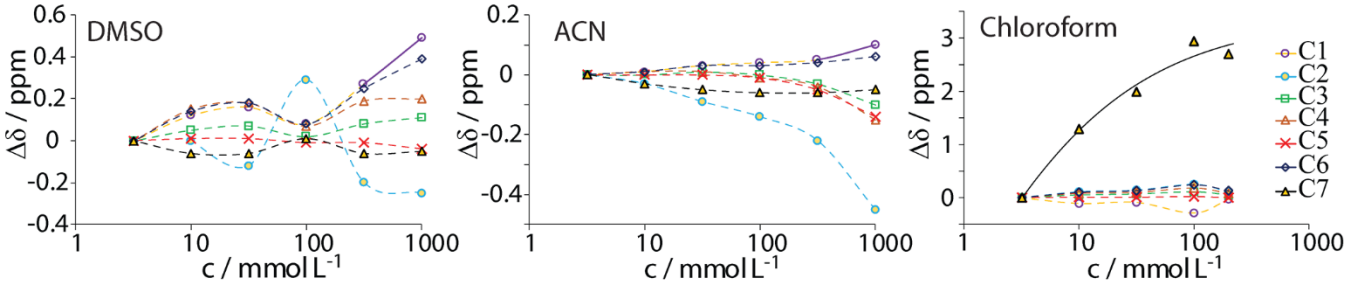


Figure 4. *Concentration dependence of ^{13}C chemical shift of 2C4NBA carbon atoms in $CDCl_3$, DMSO-d6 and acetonitrile-d3. The dashed lines are guide for the eye, whereas the solid line for C7 in $CDCl_3$ is calculated with theoretical dimerization model (details on the mathematical model used and the fitting procedure are provided in the Supporting Information).*

In the Group 1 solvents some change of chemical shift was observed in ^{1}H and ^{13}C spectra recorded in DMSO-*d6*, THF-*d8* and methanol-*d4* for 2C4NBA, 26MeOBA and 3OHBA, acetonitrile-*d3* for 2C4NBA and 3OHBA and acetone-*d6* for 2C4NBA (see Supporting Information), which indicated on some association. However, in none of these solvents this was characteristic for formation of hydrogen bonded dimers (most importantly – there was no notable chemical shift concentration dependence for ^{13}C chemical shift of C7 peak). For 2C4NBA increasing the solution concentration results in a decrease of the ^{1}H chemical shift for hydrogen atoms in the benzene ring, as shown Figure 5. In most of the solvents, however, the change actually occurs in two stages – an initial increase of the concentration (1 – 32 mM) results in increase of the chemical shift, while further increase of the concentration (32 – 1000 mM) results in decrease of the chemical shift. This is most clearly visible in the most polar solvents DMSO-*d6* and methanol-*d4*. Additionally, this effect appears most pronounced for H3 (see Figure 5 and Figure S1 for numbering of hydrogen atoms). Although our investigation and analysis were unable to provide a solid explanation for the observed chemical shift concentration dependence, most likely, the decrease of chemical shift by increasing the concentration is resulted by the formation of π···π stacked dimers and trimers. We, however, note that the experimental procedure used for collection of the ^{1}H chemical shift concentration dependence was not especially optimized which did not allow obtaining such accuracy in these highly procedure sensitive data to make the above presented conclusions irrefutable.

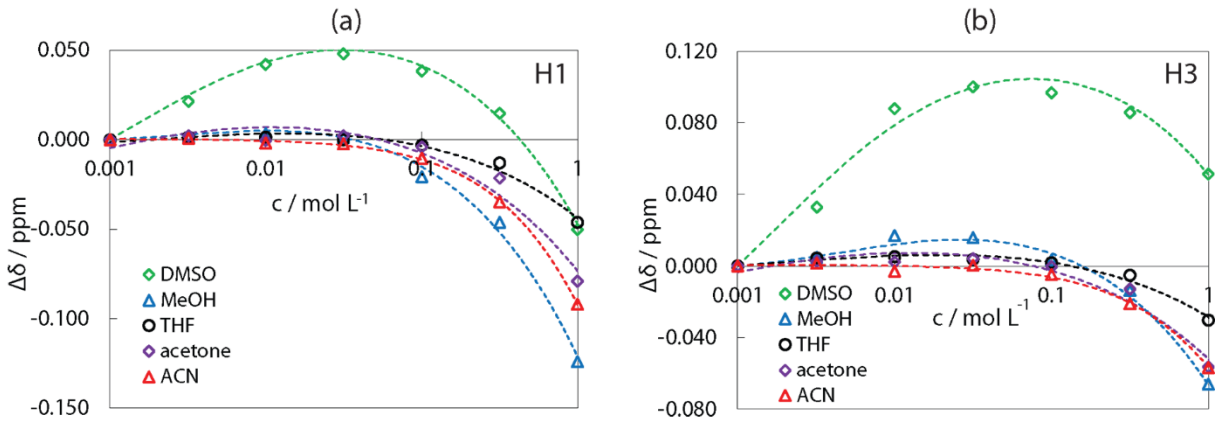


Figure 5. *Concentration dependence of ^{1}H chemical shift of 2C4NBA hydrogen H1 (a) and H3 (b) in various fully deuterated solvents. The dashed lines are calculated with cubic equation and should be interpreted as guide for the eye (see details in the Supporting Information).*

The chemical shift changes for 3OHBA in all the fully deuterated solvents and 26MeOBA in DMSO-*d6* did not indicate on formation of π···π stacked associates, while the decrease of chemical shift by increasing the concentration for 26MeOBA in THF-*d8* and methanol-*d4* suggested that formation of some associates involving aromatic interactions, like CH_3···π or π···π stacking, could take place.

3.3 Determination of species present in solution using MD simulations

We also studied the behavior of the benzoic acid derivatives in solution using MD simulations. For the simulations discussed here we used a simulation box representing solutions with a concentration falling within the range investigated in experiments (0.1 – 0.16 M), while identical conclusions on species present in all the solvents for 2C4NBA was obtained also from simulations representing solutions with a concentration of ~0.30 M, see Supporting information. Preliminary MD simulations of 26MeOBA showed that an ad-hoc tuning of the forcefield parameters is required to correctly capture the relative energy of conformers C1 and C2, thus MD simulations for 26MeOBA solutions were not included in this study.

First of all, we analyzed the MD trajectories by identifying the extent of self-association of benzoic acid derivatives in solution. Although in most of the simulations the most abundant species are solute monomers, dimers are also frequently detected. Similarly to the results emerged from FTIR and NMR, we observe that also MD simulations indicate that solvents can be divided in two groups. In the Group 2 solvents (chloroform and toluene) we can clearly observe the formation of hydrogen bonded dimers. For 2C4NBA and 2Me4NBA the cluster size distribution shows that in these simulations most of the molecules exist in dimers, with larger clusters corresponding to dimers and trimers in close proximity also contributing to the population of self-associates. The presence of self-associates stabilized by hydrogen bonds was

further confirmed by enumerating the solute-solute hydrogen bonds using VMD, as reported in Figure 6 for 2C4NBA.

For 3OHBA in chloroform the concentration used in the simulation was notably larger than the solubility of the 3OHBA in this solvent (<10 mM), resulting in formation of large clusters during the simulation. Nevertheless, at the beginning of the simulation hydrogen bonded dimers are clearly detected, confirming this to be the main intermolecular interaction motif in chloroform.

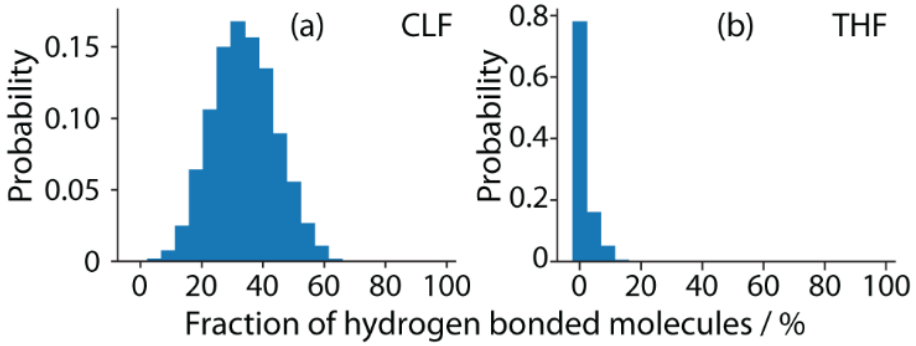


Figure 6. *Probability of 2C4NBA being involved in formation of hydrogen bond O2-H···O1 obtained from trajectories of MD simulation of ~0.1 M 2C4NBA solution in chloroform (a) and THF (b).*

In the Group 1 solvents (DMSO, 2-propanol, THF and acetonitrile) for 2C4NBA and 2Me4NBA π···π stacked dimers are the only notable, well-defined, self-association intermolecular motifs present, with adducts stabilized by other weak interactions also being occasionally observed. Among these solvents, only in THF we observed the occasional formation of a small number of hydrogen-bonded dimers. The distribution of cluster size shows that most of the molecules exist in monomeric form or are associated in dimers, while on some occasions also larger clusters can appear (see Supporting Information).

To quantify the relative abundance of different interaction motifs, for each of the systems investigated we computed the joint distribution of intermolecular solute distances and of their relative orientations. For 2C4NBA and 2Me4NBA in the Group 1 solvents the most likely distance between molecule centers is slightly below 4 Å. Moreover, interacting molecules do not display a well-defined relative orientation, indicating that the statistically dominating motif is π···π stacking (see Figure 7 (a) and (d)). In contrast, in the Group 2 solvents the most likely distance between molecule centers is slightly above 9 Å, with a distribution of relative orientations peaked at ~180°. These features are characteristic of hydrogen bonded dimers (see Figure 7 (b) and (e)). In this case, also a smaller second peak characteristic for π···π stacked molecules can be observed. A visual inspection of the trajectories reveals that such π···π stacked molecules are also typically involved in hydrogen bonding, forming tetramers or trimers through

a combination of interactions A and B as shown in Figure 7 (c) and (f) and corresponding to the adducts recently observed in benzoic acid solution[26].

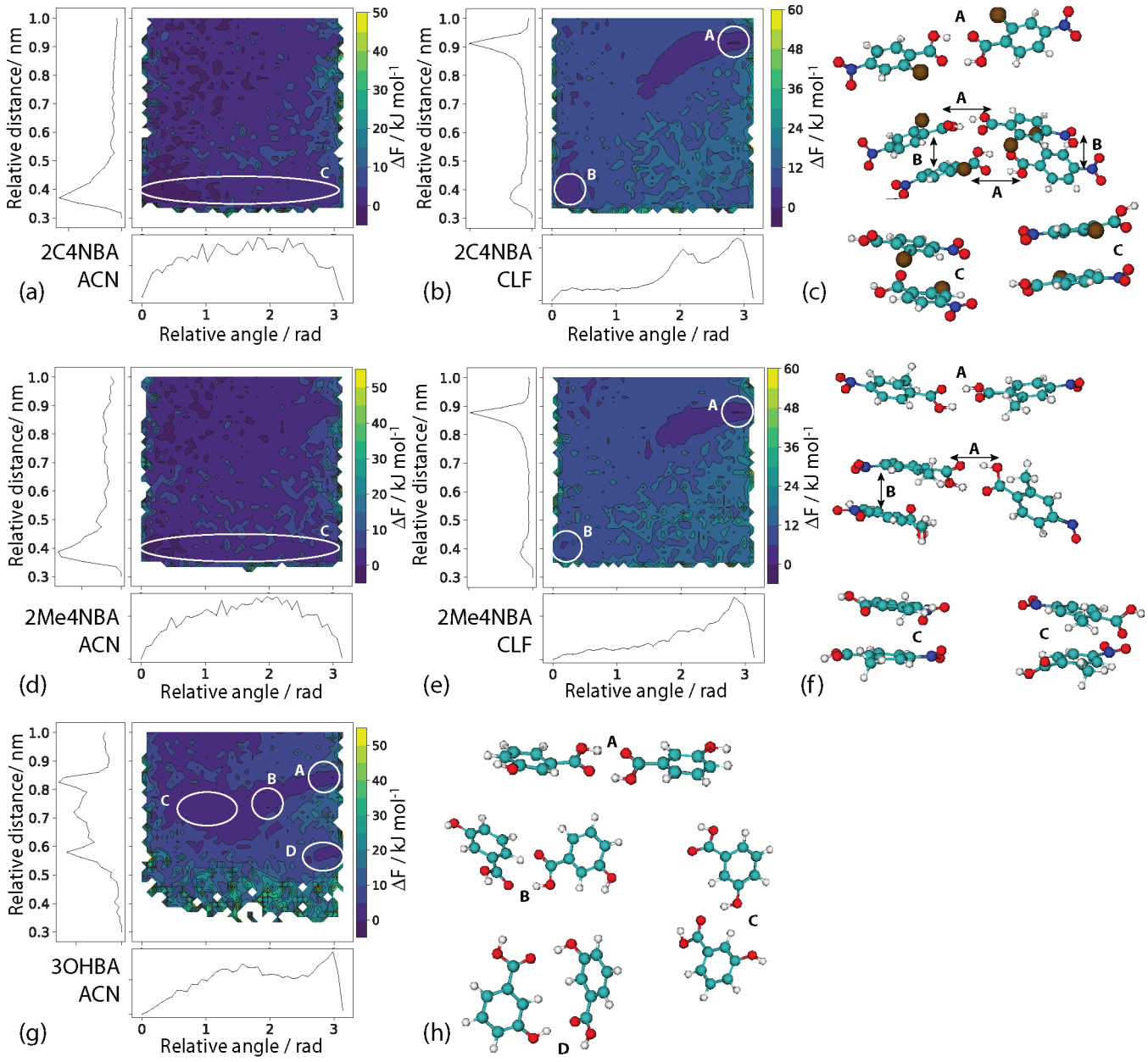


Figure 7. *Distribution of intermolecular benzoic acid derivative distances and relative orientations defined by the center of mass and direction from C4 to carboxylic group obtained from trajectories of MD simulation of ~0.1-0.15 M solution in acetonitrile (a) and chloroform (b) for 2C4NBA, acetonitrile (d) and chloroform (e) for 2Me4NBA, and THF for 3OHBA (g). Characteristic molecule pairs detectable in these distribution diagrams is given in (c), (f) and (h).*

MD simulations of 3OHBA solutions showed that in THF part of the molecules form carboxyl-carboxyl hydrogen bonded dimers (A), as well as dimers hydrogen bonded by phenolic hydroxyl group O3-H···O1 (C), and dimers stabilized by two O3-H···O1 interactions (D), as represented in the distribution of intermolecular 3OHBA distances and relative orientations, reported in Figure 7 (g) and (h). In acetonitrile and 2-propanol all these associates formed by hydrogen bonds appear notably less frequent, while they are absent in DMSO. In none of the solvents the formation of π···π stacked associates is detected.

Besides the formation of solute self-associates, in order to characterize the dominant motifs in solution it is key to assess the propensity of solvent molecules to selectively bind solutes by forming hydrogen-bonds[27-28]. We note that all solvents from the Group 1 can hydrogen-bond the benzoic acid derivative molecules, thus screening the carboxylic moieties and inhibiting the formation of hydrogen bonded dimers. In these systems the probability of solute molecules to be involved in hydrogen bonds of the type O2-H···O_{Solv}/N_{Solv} is always significant, however it varies from solvent to solvent. For instance, we find that DMSO is involved in hydrogen bonds with more than ~60% of 2C4NBA molecules, while acetonitrile is involved in hydrogen bonds with no more than ~30% of 2C4NBA molecules as reported in Figure 8.

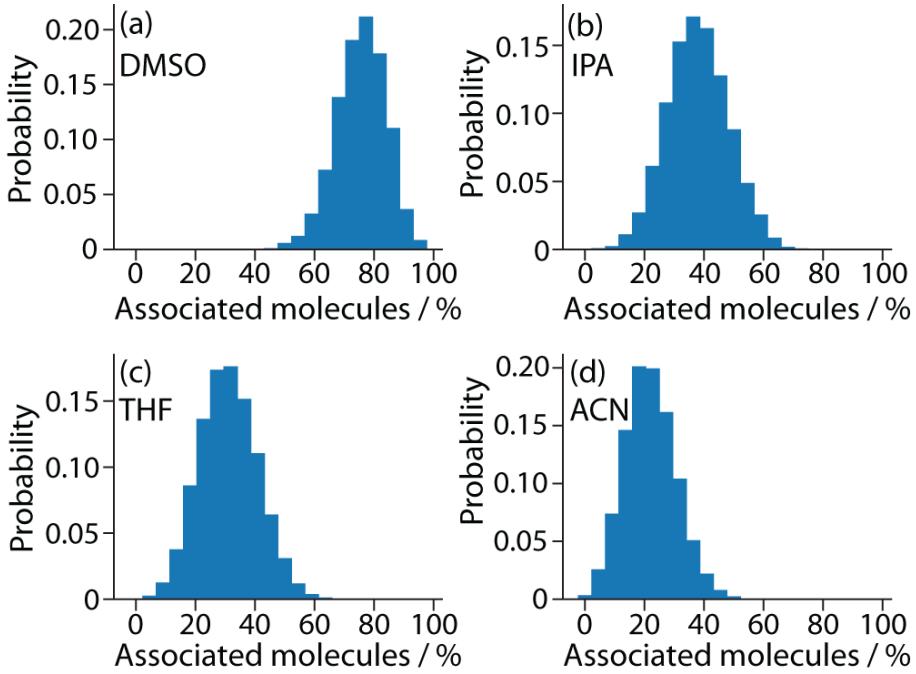


Figure 8. *Probability that the given relative amount of 2C4NBA molecules will be involved in classical hydrogen bonding O2-H···O_{Solv}/N_{Solv} as determined by HBonds plugin in VMD from trajectories of MD simulation of ~0.1 M 2C4NBA solution in DMSO (a), 2-propanol (b), THF (c) and acetonitrile (d).*

As 2-propanol is also able to get involved in the hydrogen bonds as donor by forming interaction O_{Solv}-H···O1, the occurrence of this interaction was also analyzed. Although such interaction is less frequent than O2-H···O_{Solv}, it is still highly likely, as reported in the Supporting Information. As expected, for 3OHBA also the phenolic hydroxyl group is involved in hydrogen bonding with solvent O3-H···O_{Solv}/N_{Solv}.

3.4 Stability of solute adducts via DFT calculations

To further investigate the stability of different interaction motifs in different environments, we carry out DFT calculations of the association Gibbs energy in solvent (ΔG^{solv}_{ass}) using *Gaussian 09* to evaluate the stability of various dimeric (hydrogen bonded dimers, various π···π stacked dimers and other possible dimers) as well as also some trimeric and tetrameric 2C4NBA species and thus assess their relative stability in numerous solvents. We calculated ΔG_{ass} in vacuum and in 11 different solvents, modelled using a SMD continuum model (but see the discussion in Section 3.6 for shortcomings of this approach). The geometry and all ΔG^{solv}_{ass} values for the 2C4NBA species considered are available in the Supporting Information.

Overall, these calculations show that among the dimers analyzed in all solvents except for alcohols hydrogen bonded dimers (particularly HBD_C consisting of conformers C2) should be the most stable species, see data in selected solvents in Table 1 (geometry is provided in Figure S45). Among the considered π···π stacked dimers πA_F, πS_D, $\pi 2S_B$ (geometry is provided in Figure S46) are the most stable ones, but even they are predicted to be stable only in alcohols, particularly in methanol. Although the considered trimers are predicted to be stable relative to the monomers, particularly in 1,4-dioxane and toluene, their formation is not favorable when compared to the highly favorable hydrogen bonded dimers and monomers. In contrast, formation of tetramers, particularly Tt_B, is calculated to be favorable both against the monomers as well as the dimers in all solvents.

Table 1. *Association Gibbs energy ΔG_{ass} in vacuum and several selected solvent continuum (in kJ mol^{-1}) with respect to corresponding monomers (and also dimers for the tetramer Tt_B) for selected 2C4NBA dimers and tetramer.*

	vacuum	ACN	DCM	DIOX	DMSO	IPA	MeOH	THF	TOL
HBD_C	-26.08	-7.64	-9.44	-21.89	-10.81	1.18	1.31	-14.30	-21.62
πA_F	5.71	6.98	9.19	4.25	3.85	-0.36	-3.58	7.87	9.43
πS_D	2.04	7.06	8.53	2.30	4.48	-0.01	-3.68	7.48	7.29
$\pi 2S_B$	7.09	6.40		3.83		-0.07	-2.46		9.73
Tt_B [a]	-65.13	-26.12	-25.25	-56.84	-37.66	-23.03	-31.63	-35.93	-46.23
Tt_B [b]	-12.97	-8.28	-3.65	-10.15	-13.58	-23.02	-32.00	-4.64	0.00

[a] – with respect to monomers, [b] – with respect to dimers.

3.5 Comment on formation of π···π stacked dimers

Differences in the formation of classical π···π stacked dimers for 2C4NBA and 2Me4NBA compared to 3OHBA and 26MeOBA can be rationalized based on the electrostatic potential (ESP) maps of these molecules, mirrored in the packing characteristics of their crystal structures. In the ESP map of 2C4NBA and 2Me4NBA, the area above and below the benzene ring is more positive than for 3OHBA and 26MeOBA, as shown in Figure 9. The surface extrema values above and below the benzene ring determined using Multiwfn 3.7 [66] were +8 to +13 kcal mol^{-1} for 2C4NBA, +3 to +6 kcal mol^{-1} for 2Me4NBA, –10 to –7 kcal mol^{-1} for 3OHBA and –10 to –5 kcal mol^{-1} for 26MeOBA C1 and –14 to –12 kcal mol^{-1} for C2. We thus infer that parallel stacked aromatic interactions are favored for 2C4NBA and 2Me4NBA[80-81], while CH_3···π are favored for 26MeOBA.

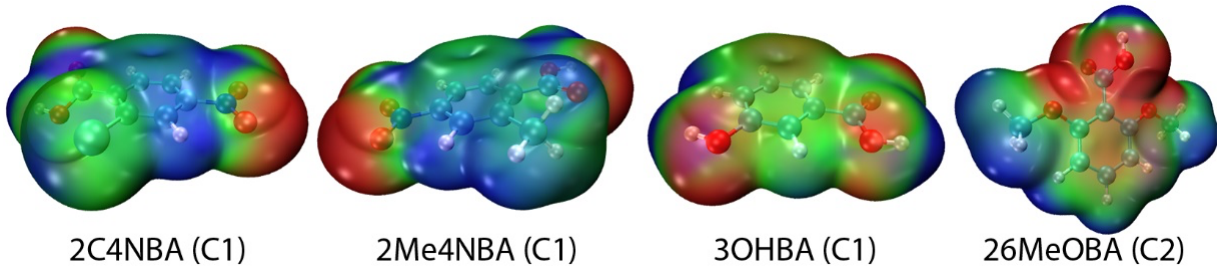


Figure 9. *Electrostatic potential maps of one of the conformers of the studied molecules. The maps for the other conformers are nearly identical for 2C4NBA, 2Me4NBA and 3OHBA and highly similar for benzene ring region for 26MeOBA, see Figure S100.*

This tendency also clearly emerges when analyzing packing features in the crystal phase of polymorphs of the studied compounds. In polymorphs I and II of 2Me4NBA and II of 2C4NBA classical parallelly stacked molecules can be found, with total interaction energies ranging from –22 to –40 kJ mol^{-1} and dispersion components from –32 to –49 kJ mol^{-1}. There are no classical parallelly stacked molecules in 2C4NBA polymorph I, but this structure is rather an exception as it is shown to be the only such structure for 2C4BA[53]. However, the only structure with classical parallelly stacked molecules for 3OHBA and 26MeOBA is 3OHBA polymorph II, but note that the total interaction energy (–19 kJ mol^{-1}) and the dispersion component (–32 kJ mol^{-1}) is the least negative among the classical parallelly stacked molecules, as given in Table S25.

3.6 Summary of species present in solution

It can be concluded that in apolar solvents or solvents with hydrogen bond acceptor propensity β < 0.2[82] (chloroform, dichloromethane, toluene and *o*-xylene) the analyzed benzoic acid derivatives form conventional hydrogen bonded dimers. In contrast, in polar solvents, or solvents able to act as hydrogen bond acceptors characterized by β > 0.3 (acetonitrile, DMSO, THF, 1,4-dioxane, methanol, 2-propanol, nitromethane, MTBE and acetone) the formation of hydrogen

bonded dimers is inhibited, in favor of π···π stacked associates for 2C4NBA and 2Me4NBA. The formation of hydrogen bonded dimers can be reliably detected with FTIR and ^{13}C spectroscopy measurements, while the formation of hydrogen bonds with the solvent can be captured quite clearly in FTIR spectra. Although ^{1}H spectra capture detailed information on a self-association, a clear identification of the species affecting changes in the chemical shifts is challenging. For 2C4NBA the decrease in the chemical shifts observed by increasing solute concentration is likely due to the formation of π···π stacked associates, however such assumption is difficult to prove unambiguously. Furthermore, all of the processes occurring in solution could be captured by analyzing MD simulations.

However, ΔG^{solv}_{ass} calculated using DFT method cannot be used to assess which species are likely to be present in solution, as spectroscopic methods and MD simulations show that 2C4NBA hydrogen bonded dimers, and, based on the MD simulations, to some extent also tetramers formed from two hydrogen bonded dimer pairs, are present only in toluene and dichloromethane (from solvents considered in *ab initio* calculations), while formation of π···π stacked associates is observed in numerous solvents. Most likely this is because these calculations cannot account for the conformational entropy associated with the large structural fluctuations of dimers and with explicit solvent. Moreover, directional interactions with the solvent cannot be captured with a continuum model.

3.7 Detailed characterization of the benzoic acid derivative associates formed with solvent

All of the studied molecules clearly form hydrogen bonds with the solvent, as observed in MD simulations and also in IR spectra by analyzing the shift in the position of C=O stretching band. In IR spectra, the most pronounced correlation between the wavenumber of C=O stretching peaks and solvent hydrogen bond acceptor propensity β is observed for 2C4NBA as shown in Figure 10, while for other molecules the correlation is not that linear and some of the solvents do not fit in this trend (as can be assessed from Figure 3). This can be associated with additional interactions present between the solvent and the benzoic acid derivative or by the modification of the self-association behavior of the benzoic acid derivative by the solvent.

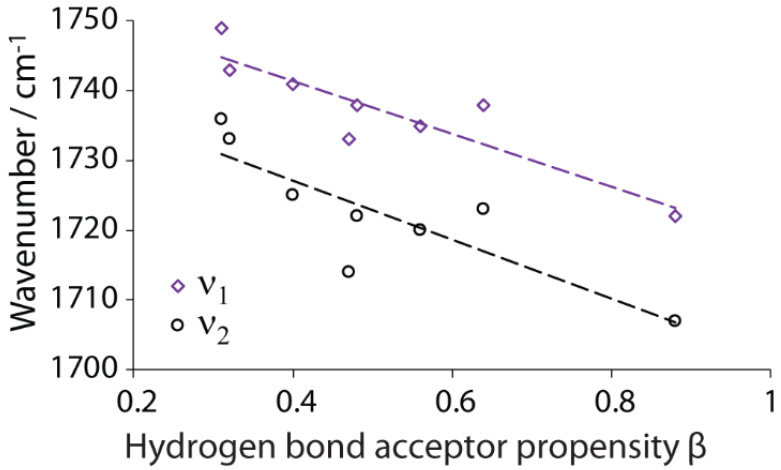


Figure 10. *Correlation between the C=O stretching peak position and the solvent hydrogen bond acceptor propensity β for 2C4NBA in the Group 1 solvents.*

Hydrogen bonds between benzoic acid derivatives and the Group 1 solvents O2-H···O_{Solv}/N_{Solv} are also clearly detected by analyzing the obtained MD trajectories. Number of 2C4NBA, 2Me4NBA and 3OHBA molecules hydrogen bonded to the solvent correlate with the hydrogen bond acceptor propensity β of solvent: DMSO forms hydrogen bonds with the largest number of benzoic acid derivative molecule, while acetonitrile – the lowest number. The most molecules are solvated for 2C4NBA (see Figure 11 where the probability maximum P_{max} in the distribution of relative amount of benzoic acid derivative molecules involved in hydrogen bond with solvent is depicted), while for 2Me4NBA and 3OHBA this number is similar and notably lower.

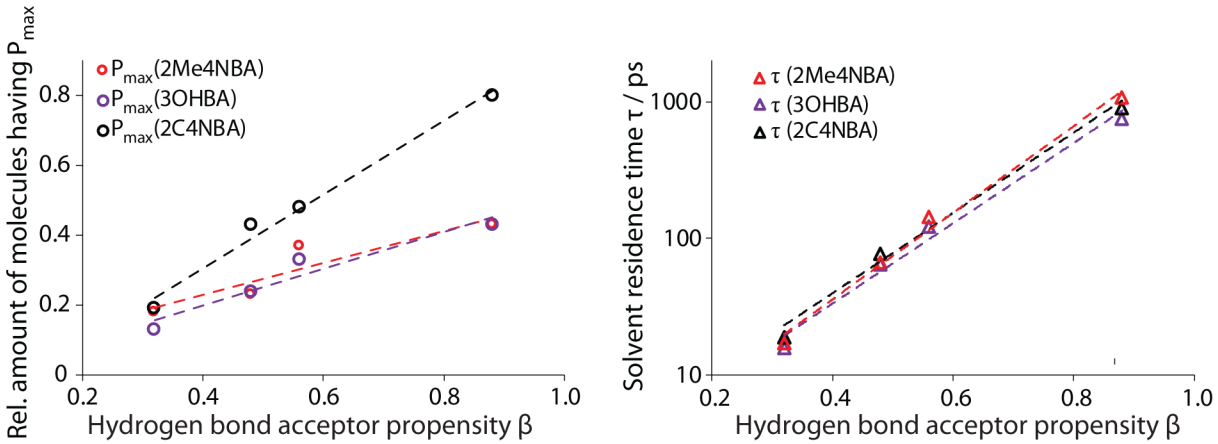


Figure 11. *Correlation of solvent hydrogen bond acceptor propensity β with probability maximum P_{max} corresponding to relative amount of benzoic acid derivative molecules involved in hydrogen bond with solvent (the lines correspond to a linear model) and with solvent residence time τ in hydrogen bonded state to benzoic acid derivative molecules (in logarithmic scale, the lines correspond to an exponential model).*

It can also be added that for 3OHBA also the phenolic hydroxyl group forms hydrogen bonds with solvent O3-H···O_{Solv}/N_{Solv}, and the number of 3OHBA molecules involved in such interaction is in fact higher than the number of molecules involved in the interaction O2-H···O_{Solv}/N_{Solv}.

Additionally, the only hydrogen bond donor solvent, 2-propanol, also forms hydrogen bonds O_{Solv} -H···O1 and with 3OHBA also O_{Solv} -H···O3. The number of such hydrogen bonds for 2C4NBA and 2Me4NBA is lower, while for 3OHBA higher than the corresponding hydrogen bonds where solvent is the acceptor.

For an additional characterization of the hydrogen bonds between the benzoic acid derivatives and the solvent molecules for all unique solute-solvent pairs investigated in simulation we determined the lifetime distribution, reported in Figure 12. The mean lifetime of solute-solvent associates τ was obtained by fitting the escape time distribution to decaying exponential. The obtained order of solvent residence time parameter τ is in agreement with the probability for observing solute molecules hydrogen bonded to the solvent, and therefore to the solvent hydrogen bond acceptor propensity β (see Figure 11). For 2C4NBA the highest lifetime is 906 ps, recorded in DMSO, and the lowest is 18.9 ps, in acetonitrile. Solvent residence time for 2-propanol acting as hydrogen bond donor is shorter (46 ps for 2C4NBA) than the lifetime of adducts in which 2-propanol is the hydrogen bond acceptor (143 ps).

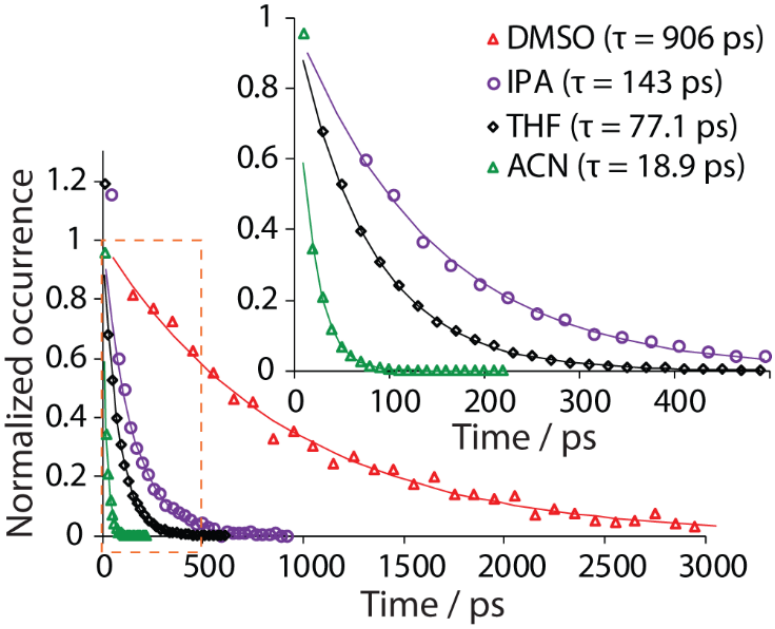


Figure 12. *Distribution of lifetime of hydrogen bonded 2C4NBA-solvent pairs (detected based on only by the distance between donor and acceptor atoms being lower than 3.0 Å) obtained from trajectories of MD simulation of ~0.1 M 2C4NBA solution in DMSO, 2-propanol, THF and acetonitrile. The lines are calculated by exponentially decaying function. The occurrence is normalized by height constant K of the theoretical equation.*

The mean lifetime of solute-solvent associates τ for 2Me4NBA are similar to those of 2C4NBA, while for 3OHBA the lifetimes of associates formed by carboxyl group are slightly lower and those formed by phenolic hydroxyl group even lower, except for 2-propanol, for which τ is 186 ps, see details in Supporting Information.

3.8 Crystal form obtained in crystallization experiments in context with the species present in solution

We also performed crystallization of all the studied benzoic acid derivatives from various solvents also used in the spectroscopic and MD studies reported in this work. Several separate (3-5) cooling crystallizations without stirring were performed at –10 °C, while evaporation crystallizations were performed at ~25 °C from 3 different type of glassware, on several occasions performing up to 3 parallel crystallizations. Summary of the crystal forms obtained in these experiments is given in Table 2. Although we cannot exclude some effect on the polymorphic outcome of the crystallization by the presence of impurities in the chemicals used, in general the solvent and selection of the crystallization method have more notable effect.

Table 2. *Summary of the crystal form obtained in the crystallization of benzoic acid derivatives in cooling and evaporation crystallization.*

	2C4NBA		2Me4NBA		3OHBA		26MeOBA	
	Cool.	Evap.	Cool.	Evap.	Cool.	Evap.	Cool.	Evap.
Group 1								
MeOH	I	I / II	I	II	I (+ II)	I + II	I (+ III)	I / I + III
2-propanol	I	I + II	I	I / II / I+II	I	I + II	I	I / I + III
Acetone	I	I (+ II) [a]	I	I + II	II (+ I)	II (+ I)	I	I / I + III
THF	I	I	I	I (+ II)	I (+ II)	I + II	I	I + III
1,4-dioxane	S_{DIOX} [b]	–	I	I / II	I (+ II)	I + II	I (+ III)	I + III / III
Acetonitrile	I	I / II	V / II	V/I/II	II (+ I)	I + II	I	I / I (+ III)
Nitromethane	I	I (+ II)	I	I/II/V	I (+ II)	II	I	I
Group 2								
DCM	I	I (+ II)	I	I / II / I+II	N/A [c]	II	I (+ III)	I (+ III)
Chloroform	I	I + II	I / S_{CLF} [b]	–	II	II	I	I (+ III)
Toluene	I	I / I + II	I	I	N/A [c]	I + II	I	I + III

[a] – in parenthesis phase present in minor quantity is given. [b] – solvate is obtained in crystallization experiments, no evaporation experiments have been performed. [c] – solubility was very low and not enough product for phase identification was obtained in the experiments.

Despite the two clear self-association modes present for 2C4NBA and 2Me4NBA in solution, the self-associates seem to have no effect on the crystal form obtained during the crystallization. However, this could be explained by the fact that crystal structures of polymorphs I and II for both compounds contain only carboxyl acid dimers (see Figure S101). Thus, apparently other criteria are responsible for selection of the crystal form. In cooling crystallization, the thermodynamic stability plays an important role, as in nearly all cases thermodynamically most

stable form is obtained (for both compounds polymorph I). In evaporation crystallization instead, the formation of crystal forms is notably less selective and both polymorphs can be obtained. Additional evaluation of the effect of evaporation rate (provided by use of different opening size of the used glassware) and container material did not lead to any conclusions about the control mechanism of the formation of these polymorphs.

The results of association obtained for 3OHBA suggest that the formation of polymorph I, containing hydrogen bonded dimers (see Figure S101), should be facilitated in chloroform, dichloromethane and toluene, which is not the case. In fact, in chloroform and dichloromethane only polymorph II was obtained, while in other solvents there was no clear selection criteria for the formation of these polymorphs. Thus, the associates present in solution do not seem to affect the polymorph obtained in the crystallization, as also proposed earlier[35]. The associates present in solution do not determine the obtained polymorph also for 26MeOBA, as in evaporation often mixture of I and III (with only the latter containing hydrogen bonded dimers, see Figure S101) is obtained, while in cooling crystallization the mixture is obtained from solvents with different properties providing different association of 26MeOBA in the solution.

4 CONCLUSIONS

All the compounds analyzed form hydrogen bonded dimers only in apolar solvents, or in solvents with hydrogen bond acceptor propensity $\beta < 0.2$. In contrast, in solvents able to act as hydrogen bond acceptors ($\beta > 0.3$) there is no notable amount of hydrogen bonded dimers, while the solute forms hydrogen bonds with solvent molecules. It is likely that these hydrogen bonds are in fact inhibiting the formation of hydrogen bonded dimers, even when established with rather weak hydrogen bond acceptors such as acetonitrile and nitromethane. Furthermore, we note that the functional groups attached to the benzene ring of the solute govern the propensity to form self-associates stabilized by weaker interactions altering the electrostatic potential of the benzene ring. 2C4NBA and 2Me4NBA tend to form $\pi\cdots\pi$ stacked associates, 26MeOBA can form $CH_3\cdots\pi$ associates, while 3OHBA does not form any associates stabilized by aromatic interactions in any of the solvents considered in this study.

The formation of hydrogen bonded dimers can be detected using FTIR and ^{13}C NMR spectroscopy measurements, while formation of hydrogen bonds with the solvent can be seen clearly in FTIR spectra. Although ^{1}H spectra seem to capture detailed information of the association, a clear identification of the species resulting the chemical shift change is challenging. The self-association processes occurring in solution can be captured by analyzing the performed MD simulations, which confirm the experimental finding and provide additional details on the structural arrangement of associates and on their lifetime. On the contrary ΔG^{solv}_{ass} calculated using DFT with implicit solvation methods does not reliably predict which species are likely to be present in solution, as shown for 2C4NBA.

Self-association in the solution for the studied compounds does not show any clear correlation with the crystal forms obtained in the crystallization, showing that the associates present in these solutions are not likely to act as the growth units, and do not directly determine the polymorphic outcome of the crystallization process.

ASSOCIATED CONTENT

Supporting Information. The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/
Detailed and additional results from FTIR and NMR spectroscopy measurements and molecular simulation methods, detailed results obtained in decomposition of C=O stretching region into individual peaks, results from *ab initio* calculations for interpretation of FTIR and NMR spectra, information on the mathematical models used for fitting the concentration dependence of ^{13}C and ^{1}H chemical shifts and fitting procedure, analysis of molecular conformation of 2C4NBA in MD simulations, additional ESP surfaces and results from pairwise interaction energy calculation. (PDF)

AUTHOR INFORMATION

Corresponding Author

*Telephone: +(371)-67033903. E-mail: agris.berzins@lu.lv.

Funding Sources

The research was funded by the European Regional Development Fund, within the Project activity 1.1.1.2 “Post-doctoral Research Aid” (project no. 1.1.1.2/VIAA/1/16/195).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We thank Rihards Klūga for recording NMR spectra and Ilva Kresse and Sabīne Blaua for performing part of the association studies and solid form screening experiments as part of their research projects.

REFERENCES

1. Chen, J.; Sarma, B.; Evans, J. M. B.; Myerson, A. S., Pharmaceutical Crystallization. *Cryst. Growth Des.* **2011,** *11* (4), 887-895.
2. Myerson, A. S.; Trout, B. L., Nucleation from Solution. *Science* **2013,** *341* (6148), 855-856.
3. Davey, R. J.; Schroeder, S. L. M.; ter Horst, J. H., Nucleation of Organic Crystals—A Molecular Perspective. *Angew. Chem. Int. Ed.* **2013,** *52* (8), 2166-2179.
4. Khamar, D.; Zeglinski, J.; Mealey, D.; Rasmuson, Å. C., Investigating the Role of Solvent–Solute Interaction in Crystal Nucleation of Salicylic Acid from Organic Solvents. *J. Am. Chem. Soc.* **2014,** *136* (33), 11664-11673.
5. Cruz-Cabeza, A. J.; Feeder, N.; Davey, R. J., Open questions in organic crystal polymorphism. *Commun. Chem* **2020,** *3* (1), 142.
6. Aaltonen, J.; Allesø, M.; Mirza, S.; Koradia, V.; Gordon, K. C.; Rantanen, J., Solid form screening – A review. *Eur. J. Pharm. Biopharm.* **2009,** *71* (1), 23-37.
7. Kitamura, M., Polymorphism in the crystallization of L-glutamic acid. *J. Cryst. Growth* **1989,** *96* (3), 541-546.
8. Neumann, M. A.; van de Streek, J.; Fabbiani, F. P. A.; Hidber, P.; Grassmann, O., Combined crystal structure prediction and high-pressure crystallization in rational pharmaceutical polymorph screening. *Nat. Commun* **2015,** *6* (1), 7793.
9. Oswald, I. D. H.; Chataigner, I.; Elphick, S.; Fabbiani, F. P. A.; Lennie, A. R.; Maddaluno, J.; Marshall, W. G.; Prior, T. J.; Pulham, C. R.; Smith, R. I., Putting pressure on elusive polymorphs and solvates. *CrystEngComm* **2009,** *11* (2), 359-366.
10. Davey, R. J.; Allen, K.; Blagden, N.; Cross, W. I.; Lieberman, H. F.; Quayle, M. J.; Righini, S.; Seton, L.; Tiddy, G. J. T., Crystal engineering – nucleation, the key step. *CrystEngComm* **2002,** *4* (47), 257-264.
11. Davey, R. J.; Blagden, N.; Righini, S.; Alison, H.; Quayle, M. J.; Fuller, S., Crystal Polymorphism as a Probe for Molecular Self-Assembly during Nucleation from Solutions: The Case of 2,6-Dihydroxybenzoic Acid. *Cryst. Growth Des.* **2001,** *1* (1), 59-65.
12. Svärd, M.; Nordström, F. L.; Jasnobulka, T.; Rasmuson, Å. C., Thermodynamics and Nucleation Kinetics of m-Aminobenzoic Acid Polymorphs. *Cryst. Growth Des.* **2010,** *10* (1), 195-204.
13. Martínez-Ohárriz, M. C.; Martín, C.; Goñi, M. M.; Rodríguez-Espinosa, C.; Tros De Ilarduya-Apaolaza, M. C.; Sánchez, M., Polymorphism of Diflunisal: Isolation and Solid-State Characteristics of a New Crystal Form. *J. Pharm. Sci.* **1994,** *83* (2), 174-177.
14. Mukuta, T.; Lee, A. Y.; Kawakami, T.; Myerson, A. S., Influence of Impurities on the Solution-Mediated Phase Transformation of an Active Pharmaceutical Ingredient. *Cryst. Growth Des.* **2005,** *5* (4), 1429-1436.
15. Banerjee, M.; Brettmann, B., Combining Surface Templating and Confinement for Controlling Pharmaceutical Crystallization. *Pharmaceutics* **2020,** *12* (10), 995.
16. Black, J. F. B.; Cruz-Cabeza, A. J.; Davey, R. J.; Willacy, R. D.; Yeoh, A., The Kinetic Story of Tailor-made Additives in Polymorphic Systems: New Data and Molecular Insights for p-Aminobenzoic Acid. *Cryst. Growth Des.* **2018,** *18* (12), 7518-7525.
17. Warzecha, M.; Guo, R.; M. Bhardwaj, R.; Reutzel-Edens, S. M.; Price, S. L.; Lamprou, D. A.; Florence, A. J., Direct Observation of Templated Two-Step Nucleation Mechanism during Olanzapine Hydrate Formation. *Cryst. Growth Des.* **2017,** *17* (12), 6382-6393.

18. Hunter, C. A.; McCabe, J. F.; Spitaleri, A., Solvent effects of the structures of prenucleation aggregates of carbamazepine. *CrystEngComm* **2012,** *14* (21), 7115-7117.
19. Parveen, S.; Davey, R. J.; Dent, G.; Pritchard, R. G., Linking solution chemistry to crystal nucleation: the case of tetrolic acid. *Chem. Commun.* **2005,** (12), 1531-1533.
20. Spitaleri, A.; Hunter, C. A.; McCabe, J. F.; Packer, M. J.; Cockroft, S. L., A 1H NMR study of crystal nucleation in solution. *CrystEngComm* **2004,** *6* (80), 490-493.
21. Edkins, R. M.; Hayden, E.; Steed, J. W.; Fucke, K., Conserved hydrogen bonding in tetrahydrocarbazolone derivatives: influence of solution-state assembly on crystal form nucleation. *Chem. Commun.* **2015,** *51* (25), 5314-5317.
22. Warzecha, M.; Verma, L.; Johnston, B. F.; Palmer, J. C.; Florence, A. J.; Vekilov, P. G., Olanzapine crystal symmetry originates in preformed centrosymmetric solute dimers. *Nature Chemistry* **2020,** *12* (10), 914-920.
23. Mattei, A.; Li, T., Polymorph Formation and Nucleation Mechanism of Tolfenamic Acid in Solution: An Investigation of Pre-nucleation Solute Association. *Pharm. Res.* **2012,** *29* (2), 460-470.
24. Kulkarni, S. A.; McGarrity, E. S.; Meekes, H.; ter Horst, J. H., Isonicotinamide self-association: the link between solvent and polymorph nucleation. *Chem. Commun.* **2012,** *48* (41), 4983-4985.
25. Davey, R. J.; Dent, G.; Mughal, R. K.; Parveen, S., Concerning the Relationship between Structural and Growth Synthons in Crystal Nucleation: Solution and Crystal Chemistry of Carboxylic Acids As Revealed through IR Spectroscopy. *Cryst. Growth Des.* **2006,** *6* (8), 1788-1796.
26. Tang, W.; Zhang, M.; Mo, H.; Gong, J.; Wang, J.; Li, T., Higher-Order Self-Assembly of Benzoic Acid in Solution. *Cryst. Growth Des.* **2017,** *17* (10), 5049-5053.
27. Bobrovs, R.; Drunka, L.; Auzins, A. A.; Jaudzems, K.; Salvalaglio, M., Polymorph-Selective Role of Hydrogen Bonding and π - π Stacking in p-Aminobenzoic Acid Solutions. *Cryst. Growth Des.* **2021,** *21* (1), 436–448.
28. Tang, W.; Mo, H.; Zhang, M.; Parkin, S.; Gong, J.; Wang, J.; Li, T., Persistent Self-Association of Solute Molecules in Solution. *J. Phys. Chem. B* **2017,** *121* (43), 10118-10124.
29. Jones, C. D.; Walker, M.; Xiao, Y.; Edkins, K., Pre-nucleation aggregation based on solvent microheterogeneity. *Chem. Commun.* **2019,** *55* (33), 4865-4868.
30. Cruz-Cabeza, A. J.; Davey, R. J.; Sachithananthan, S. S.; Smith, R.; Tang, S. K.; Vetter, T.; Xiao, Y., Aromatic stacking – a key step in nucleation. *Chem. Commun.* **2017,** *53* (56), 7905-7908.
31. Liu, Y.; Xu, S.; Zhang, X.; Tang, W.; Gong, J., Unveiling the Critical Roles of Aromatic Interactions in the Crystal Nucleation Pathway of Flufenamic Acid. *Cryst. Growth Des.* **2019,** *19* (12), 7175-7184.
32. Du, W.; Cruz-Cabeza, A. J.; Woutersen, S.; Davey, R. J.; Yin, Q., Can the study of self-assembly in solution lead to a good model for the nucleation pathway? The case of tolfenamic acid. *Chem. Sci.* **2015,** *6* (6), 3515-3524.
33. Chiarella, R. A.; Gillon, A. L.; Burton, R. C.; Davey, R. J.; Sadiq, G.; Auffret, A.; Cioffi, M.; Hunter, C. A., The nucleation of inosine: the impact of solution chemistry on the appearance of polymorphic and hydrated crystal forms. *Faraday Discuss.* **2007,** *136* (0), 179-193.
34. Hylton, R. K.; Tizzard, G. J.; Threlfall, T. L.; Ellis, A. L.; Coles, S. J.; Seaton, C. C.; Schulze, E.; Lorenz, H.; Seidel-Morgenstern, A.; Stein, M.; Price, S. L., Are the Crystal

Structures of Enantiopure and Racemic Mandelic Acids Determined by Kinetics or Thermodynamics? *J. Am. Chem. Soc.* **2015,** *137* (34), 11095-11104.
35. Di Tommaso, D., The molecular self-association of carboxylic acids in solution: testing the validity of the link hypothesis using a quantum mechanical continuum solvation approach. *CrystEngComm* **2013,** *15* (33), 6564-6577.
36. Huang, J.; Stringfellow, T. C.; Yu, L., Glycine Exists Mainly as Monomers, Not Dimers, in Supersaturated Aqueous Solutions: Implications for Understanding Its Crystallization and Polymorphism. *J. Am. Chem. Soc.* **2008,** *130* (42), 13973-13980.
37. Tang, W.; Mo, H.; Zhang, M.; Gong, J.; Wang, J.; Li, T., Glycine's pH-Dependent Polymorphism: A Perspective from Self-Association in Solution. *Cryst. Growth Des.* **2017,** *17* (10), 5028-5033.
38. Mealey, D.; Zeglinski, J.; Khamar, D.; Rasmuson, Å. C., Influence of solvent on crystal nucleation of risperidone. *Faraday Discuss.* **2015,** *179* (0), 309-328.
39. Sullivan, R. A.; Davey, R. J.; Sadiq, G.; Dent, G.; Back, K. R.; ter Horst, J. H.; Toroz, D.; Hammond, R. B., Revealing the Roles of Desolvation and Molecular Self-Assembly in Crystal Nucleation from Solution: Benzoic and p-Aminobenzoic Acids. *Cryst. Growth Des.* **2014,** *14* (5), 2689-2696.
40. Gaines, E.; Maisuria, K.; Di Tommaso, D., The role of solvent in the self-assembly of m-aminobenzoic acid: a density functional theory and molecular dynamics study. *CrystEngComm* **2016,** *18* (16), 2937-2948.
41. Bobrovs, R.; Seton, L.; Dempster, N., The reluctant polymorph: investigation into the effect of self-association on the solvent mediated phase transformation and nucleation of theophylline. *CrystEngComm* **2015,** *17* (28), 5237-5251.
42. Chadwick, K.; Davey, R. J.; Dent, G.; Pritchard, R. G.; Hunter, C. A.; Musumeci, D., Cocrystallization: A Solution Chemistry Perspective and the Case of Benzophenone and Diphenylamine. *Cryst. Growth Des.* **2009,** *9* (4), 1990-1999.
43. Mukherjee, A.; Dixit, K.; Sarma, S. P.; Desiraju, G. R., Aniline-phenol recognition: from solution through supramolecular synthons to cocrystals. *IUCrJ* **2014,** *1* (4), 228-239.
44. Mattei, A.; Mei, X.; Miller, A.-F.; Li, T., Two Major Pre-Nucleation Species that are Conformationally Distinct and in Equilibrium of Self-Association. *Cryst. Growth Des.* **2013,** *13* (8), 3303-3307.
45. Kuppens, T.; Herrebout, W.; van der Veken, B.; Bultinck, P., Intermolecular Association of Tetrahydrofuran-2-carboxylic Acid in Solution: A Vibrational Circular Dichroism Study. *J. Phys. Chem. A* **2006,** *110* (34), 10191-10200.
46. Bünnemann, K.; Pollok, C. H.; Merten, C., Explicit Solvation of Carboxylic Acids for Vibrational Circular Dichroism Studies: Limiting the Computational Efforts without Losing Accuracy. *J. Phys. Chem. B* **2018,** *122* (33), 8056-8064.
47. Loya, J. D.; Qiu, J.; Unruh, D. K.; Cozzolino, A. F.; Hutchins, K. M., Co-Crystallization of the Anti-Cholesterol Drug Bezafibrate: Molecular Recognition of a Pharmaceutical Contaminant in the Solid State and Solution via Hydrogen Bonding. *Cryst. Growth Des.* **2018,** *18* (9), 4838-4843.
48. Hamad, S.; Moon, C.; Catlow, C. R. A.; Hulme, A. T.; Price, S. L., Kinetic Insights into the Role of the Solvent in the Polymorphism of 5-Fluorouracil from Molecular Dynamics Simulations. *J. Phys. Chem. B* **2006,** *110* (7), 3323-3329.

49. Chen, J.; Trout, B. L., Computational Study of Solvent Effects on the Molecular Self-Assembly of Tetrolic Acid in Solution and Implications for the Polymorph Formed from Crystallization. *J. Phys. Chem. B* **2008,** *112* (26), 7794-7802.
50. Gavezzotti, A., Molecular Aggregation of Acetic Acid in a Carbon Tetrachloride Solution: A Molecular Dynamics Study with a View to Crystal Nucleation. *Chem. Eur. J.* **1999,** *5* (2), 567-576.
51. Barsky, I.; Bernstein, J.; Stephens, P. W.; Stone, K. H., The study of the polymorphic system of 2-chloro-4-nitrobenzoic acid. *New J. Chem.* **2008,** *32* (10), 1747-1753.
52. Aitipamula, S.; Chow, P. S.; Tan, R. B. H., Solvates and polymorphic phase transformations of 2-chloro-4-nitrobenzoic acid. *CrystEngComm* **2011,** *13* (3), 1037-1045.
53. Bērziņš, A.; Kons, A.; Saršu¯ns, K.; Belyakov, S.; Actiņš, A., On the Rationalization of Formation of Solvates: Experimental and Computational Study of Solid Forms of Several Nitrobenzoic Acid Derivatives. *Cryst. Growth Des.* **2020,** *20* (9), 5767-5784.
54. Nordström, F. L.; Rasmuson, Å. C., Polymorphism and thermodynamics of m-hydroxybenzoic acid. *Eur. J. Pharm. Sci.* **2006,** *28* (5), 377-384.
55. Gridunova, G.; Furmanova, N.; Struchkov, Y. T.; Ezhkova, Z.; Grigoryeva, L.; Chayanov, B., Crystal-Structures of 2 Polymorphic Modifications of M-Oxybenzoic Acid. *Kristallografiya* **1982,** *27* (2), 267-272.
56. Svärd, M.; Rasmuson, Å. C., m-Hydroxybenzoic Acid: Quantifying Thermodynamic Stability and Influence of Solvent on the Nucleation of a Polymorphic System. *Cryst. Growth Des.* **2013,** *13* (3), 1140-1152.
57. He, G.; Wong, A. B. H.; Chow, P. S.; Tan, R. B. H., Effects of the rate of supersaturation generation on polymorphic crystallization of m-hydroxybenzoic acid and o-aminobenzoic acid. *J. Cryst. Growth* **2011,** *314* (1), 220-226.
58. Nordström, F. L.; Svärd, M.; Malmberg, B.; Rasmuson, Å. C., Influence of Solution Thermal and Structural History on the Nucleation of m-Hydroxybenzoic Acid Polymorphs. *Cryst. Growth Des.* **2012,** *12* (9), 4340-4348.
59. Bryan, R. F.; White, D. H., 2,6-Dimethoxybenzoic acid. *Acta Crystallogr., Sect. B* **1982,** *38* (3), 1014-1016.
60. Portalone, G., Redetermination of 2,6-dimethoxybenzoic acid. *Acta Crystallogr., Sect. E* **2009,** *65* (2), o327-o328.
61. Portalone, G., A new polymorph of 2,6-dimethoxybenzoic acid. *Acta Crystallogr., Sect. E* **2011,** *67* (12), o3394-o3395.
62. Pal, R.; Jelsch, C.; Malaspina, L. A.; Edwards, A. J.; Murshed, M. M.; Grabowsky, S., syn and anti polymorphs of 2,6-dimethoxy benzoic acid and its molecular and ionic cocrystals: Structural analysis and energetic perspective. *J. Mol. Struct.* **2020,** *1221*, 128721.
63. Portalone, G., Crystal structure and Hirshfeld surface analysis of a third polymorph of 2,6-dimethoxybenzoic acid. *Acta Crystallogr., Sect. E* **2020,** *76* (12), 1823-1826.
64. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A. J.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R.

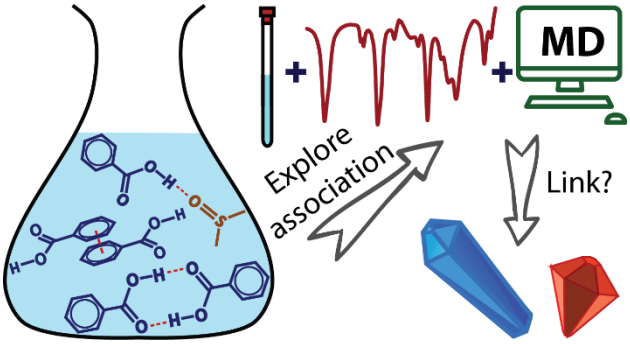
E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision D.01*; Gaussian Inc.: Wallingford, CT, 2009.
65. Kashinski, D. O.; Chase, G. M.; Nelson, R. G.; Di Nallo, O. E.; Scales, A. N.; VanderLey, D. L.; Byrd, E. F. C., Harmonic Vibrational Frequencies: Approximate Global Scaling Factors for TPSS, M06, and M11 Functional Families Using Several Common Basis Sets. *J. Phys. Chem. A* **2017,** *121* (11), 2265-2273.
66. Lu, T.; Chen, F., Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012,** *33* (5), 580-592.
67. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *J. Mol. Graph.* **1996,** *14* (1), 33-38.
68. Marenich, A. V.; Cramer, C. J.; Truhlar, D. G., Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009,** *113* (18), 6378-6396.
69. Mackenzie, C. F.; Spackman, P. R.; Jayatilaka, D.; Spackman, M. A., CrystalExplorer model energies and energy frameworks: extension to metal coordination compounds, organic salts, solvates and open-shell systems. *IUCrJ* **2017,** *4* (5), 575-587.
70. 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.
71. Case, D. A. Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Giambasu, G.; Giese, T.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Huang, Y.; Izadi, S.; Kovalenko, A.; et al. *AMBER 2019*; University of California: San Francisco, 2019.
72. Caleman, C.; van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; van der Spoel, D., Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **2012,** *8* (1), 61-74.
73. van der Spoel, D.; van Maaren, P. J.; Caleman, C., GROMACS molecule & liquid database. *Bioinformatics* **2012,** *28* (5), 752-753.
74. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.
75. Parrinello, M.; Rahman, A., Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981,** *52* (12), 7182-7190.
76. Bussi, G.; Donadio, D.; Parrinello, M., Canonical sampling through velocity rescaling. *J. Chem. Phys* **2007,** *126* (1), 014101.
77. Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banáš, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; Capelli, R.; Carloni, P.; Ceriotti, M.; Cesari, A.; Chen, H.; Chen, W.; Colizzi, F.; De, S.; De La Pierre, M.; Donadio, D.; Drobot, V.; Ensing, B.; Ferguson, A. L.; Filizola, M.; Fraser, J. S.; Fu, H.; Gasparotto, P.; Gervasio, F. L.; Giberti, F.; Gil-Ley, A.; Giorgino, T.; Heller, G. T.; Hocky, G. M.; Iannuzzi, M.; Invernizzi, M.; Jelfs, K. E.; Jussupow, A.; Kirilin, E.; Laio, A.; Limongelli, V.; Lindorff-Larsen, K.; Löhr, T.; Marinelli, F.; Martin-Samos, L.; Masetti, M.; Meyer, R.; Michaelides, A.; Molteni, C.; Morishita, T.; Nava, M.; Paissoni, C.; Papaleo, E.; Parrinello, M.; Pfaendtner, J.; Piaggi, P.; Piccini, G.; Pietropaolo, A.; Pietrucci, F.; Pipolo, S.; Provasi, D.; Quigley, D.; Raiteri, P.; Raniolo, S.; Rydzewski, J.; Salvalaglio, M.; Sosso, G. C.; Spiwok, V.; Šponer, J.; Swenson, D. W. H.; Tiwary, P.; Valsson,

O.; Vendruscolo, M.; Voth, G. A.; White, A.; The, P. c., Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019,** *16* (8), 670-673.
78. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014,** *185* (2), 604-613.
79. Fujii, Y.; Yamada, H.; Mizuta, M., Self-association of acetic acid in some organic solvents. *J. Phys. Chem.* **1988,** *92* (23), 6768-6772.
80. Hunter, C. A.; Sanders, J. K. M., The nature of .pi.-.pi. interactions. *J. Am. Chem. Soc.* **1990,** *112* (14), 5525-5534.
81. Chen, T.; Li, M.; Liu, J., π–π Stacking Interaction: A Nondestructive and Facile Means in Material Engineering for Bioapplications. *Cryst. Growth Des.* **2018,** *18* (5), 2765-2783.
82. Gu, C.-H.; Li, H.; Gandhi, R. B.; Raghavan, K., Grouping solvents by statistical analysis of solvent property parameters: implication to polymorph screening. *Int. J. Pharm.* **2004,** *283* (1–2), 117-125.

For Table of Contents Use Only

Speciation of substituted benzoic acids in solution: evaluation of spectroscopic and computational methods for the identification of associates and their role in crystallization

Agris Bērziņš, Aina Semjonova, Andris Actiņš, Matteo Salvalaglio



Self-association of four benzoic acid derivatives in solvent was investigated using spectroscopic measurements and molecular simulation methods. In apolar solvents or solvents with low hydrogen bond acceptor propensity β the studied compounds form hydrogen-bonded dimers, while in solvents with $\beta > 0.3$ the studied compounds form hydrogen bonds with solvent and associates formed by aromatic interactions.