Towards the systematic crystallisation of molecular ionic cocrystals

Mohamed, Sharmarke; Alwan, Ahmad A.; Friščić, Tomislav; Morris, Andrew J.; Arhangelskis, Mihails

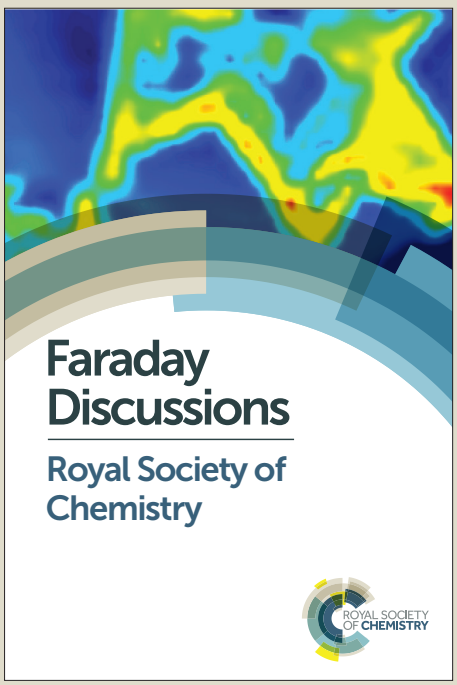
DOI:
10.1039/C8FD00036K

Document Version
Peer reviewed version

Citation for published version (Harvard):
Mohamed, S, Alwan, AA, Friščić, T, Morris, AJ & Arhangelskis, M 2018, 'Towards the systematic crystallisation of molecular ionic cocrystals: insights from computed crystal form landscapes', *Faraday Discussions*. https://doi.org/10.1039/C8FD00036K

Faraday Discussions

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: S. Mohamed, A. A. Alwan, T. Friši, A. J. Morris and M. Arhangelskis, *Faraday Discuss.*, 2018, DOI: 10.1039/C8FD00036K.

DOI: 10.1039/C8FD00036K

Towards the systematic crystallisation of molecular ionic cocrystals: insights from computed crystal form landscapes

Sharmarke Mohamed,[*a] Ahmad. A. Alwan,[a] Tomislav Friščić,[b] Andrew J. Morris[c] and Mihails Arhangelskis[b]

[a] Department of Chemistry, Khalifa University of Science and Technology, PO Box 127788, Abu Dhabi, United Arab Emirates.

[b] Department of Chemistry, McGill University, 801 Sherbrooke St. W., H3A 0B8 Montreal, Canada.

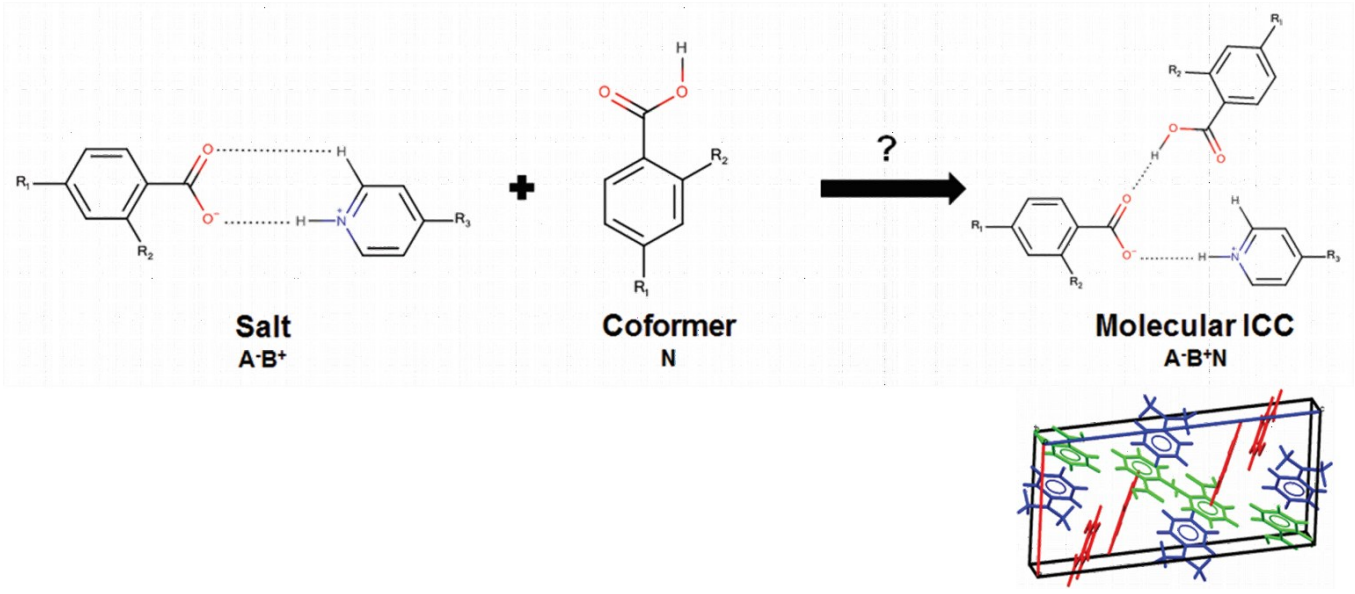
[c] School of Metallurgy and Materials, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK.

[*]Corresponding author. E-mail: sharmarke.mohamed@kustar.ac.ae

Abstract

DOI: 10.1039/C8FD00036K

The underlying molecular and crystal properties affecting the crystallisation of ionic cocrystals (ICCs) with the general formula A^-B^+N (A^-=anion, B^+=cation and N=neutral acid molecule; 1:1:1 stoichiometry) are reported for a limited set of known crystal structures determined following the cocrystallisation of either 4-aminopyridine (forms salts) or 4-dimethylaminopyridine (forms salts and ICCs) with the same set of monoprotic acids with a single hydroxy or halogen substitution at the *ortho* or *para* positions. Periodic density functional theory calculations (PBE+D2) on the energetic driving force for ICC crystallisation for a set of known crystal structures with well characterised acid, salt and ICC structures show that all but 1 of the 7 experimental ICC structures surveyed were more stable than the sum of their component salt and acid structures with 4 displaying relative stabilities (ΔE_{ICC}) ranging from 2.47-8.02 kJ mol^{-1}. The majority of molecular ICCs that are more stable with respect to their component salt and acid structures display the formation of discrete intermolecular O-H_{acid}··O_{anion} hydrogen bonds with $D_1^1(2)$ graph set between the carboxylic acid OH donor and the carboxylate oxygen acceptor of the anion. Computed crystal form landscapes for model 1:1 salts derived from acid-base pairs (involving 4-dimethylaminopyridine) known to form molecular ICCs show that on average the most stable predicted polymorphs of the 1:1 salts have efficient packing of the ions with packing coefficients in the range 65-80% and this is comparable to the packing coefficients of the most stable predicted polymorphs of 1:1 salts (involving 4-aminopyridine) that have no ICCs reported. This suggests that the cocrystallisation of equimolar amounts of the 1:1 salt and the acid to form a 1:1:1 molecular ICC is a complicated phenomenon that cannot be explained on the basis of inefficiencies in the crystal packing of the salt ions.



1. Towards the systematic crystallisation of molecular ionic cocrystals: insights from computed crystal form landscapes

View Article Online
DOI: 10.1039/C8FD00036K

1.1. Introduction

The chemical composition of crystal forms is an important intrinsic property affecting the physicochemical properties[1-3] of solid materials. For active pharmaceutical ingredients (APIs), the identity and composition of the species in the lattice is also an important regulatory[4] consideration for the classification of the solid form as a polymorph, solvate, salt or cocrystal. Each of these type of solid form offers the opportunity to enhance one or more physicochemical properties of the API[5-7]. In the domain of multicomponent crystal forms, the past decade has seen significant effort by crystal engineers to develop predictive theoretical[8] and empirical rules[9] for the synthesis of cocrystal solid forms. Cocrystals have received significant attention because of their proven utility[10-12] as vehicles for enhancing the performance characteristics and solid-state properties of a range of ionisable and non-ionisable molecules. The hierarchy of hydrogen bond synthons[13-15] observed in cocrystal solid forms suggests that theoretical models capable of predicting and accurately ranking the stabilities of different cocrystal polymorphs as a function of coformer identity and relative composition, could prove useful in experimental efforts in cocrystal selection. There is already increasing evidence[8, 16, 17] that cocrystal formation may be predicted on thermodynamic grounds given a sufficiently accurate theoretical model for calculating the relative energies of the cocrystal with respect to its component molecules.

Computational methods of crystal structure prediction (CSP) have shown to be a useful aid[18-20] in experimental efforts targeted at the discovery of new polymorphs[21-23] of organic compounds. The crystal form landscape (CFL)[18, 24] calculated from a typical CSP study is usually depicted as a scatter plot of lattice energy versus density and shows all unique lattice energy minima corresponding to hypothetical polymorphs of the system at a nominal 0 K temperature. This CFL, in principle contains enough structural information about the packing preferences and stabilities of crystal polymorphs, that an accurate algorithm for translating this CFL into the experimental conditions necessary for targeting each polymorph would provide a critical leap forward in the industrial solid-form screening strategy. The continued exponential increase in computing power is already making it possible to apply machine learning techniques to a variety of problems in science[25-27]. Although the calculation of crystal properties from calculated CFLs is still non-routine and state-of-the-art[28], recent work combining CSP with the estimation of physical properties has led to the calculation of energy-structure-function maps that have facilitated the discovery of previously unknown highly porous molecular crystals[29, 30]. The periodic blind tests[31-36] in organic crystal structure prediction organised by the Cambridge Crystallographic Data Centre (CCDC) has shown a steady improvement in the capabilities of existing CSP algorithms in predicting the crystal structures of flexible organic molecules as well as those crystal structures that could be classified as multicomponent crystal forms. Dispersion-corrected density-functional theory (DFT-D)[37-39] methods have been shown[40] to be reliable in producing final relative lattice energies for predicted structures that are often in accord with experimental stability ordering for polymorphs[41] or show that the

predicted global lattice energy minimum structure corresponds to the observed experimental structure[42]. In the domain of multicomponent crystal form synthesis, CSP methods have shown to be a useful complementary tool in experimental efforts for the discovery of cocrystals[43], hydrates[44] and salt hydrates.[24, 45]

DOI: 10.1039/C8FD00036K

The term *ionic cocrystal (ICC)*[46-48] has recently entered the crystal engineering lexicon to describe a multicomponent solid form where the crystal lattice consists of a salt and an additional chemical species which may be a neutral molecule or another salt complex. The existence of ICC solid forms has been known for a long time however, with some reports[48] suggesting that it was Kobell, who in 1843 first reported[49] an ICC consisting of sodium chloride and glucose. Despite the long history of ICC solid forms, the systematic crystallisation of ICCs using tested crystal engineering principles or theoretical methods is an underexplored area of solid state chemistry research. Here, we define a molecular ICC as any solid form with the general formula[48] A^-B^+N (where A^-=anion, B^+=cation and N=neutral molecule or salt) where the ions of the salt are derived from molecular species. In the case where N is a neutral molecule, the systematic manipulation of the chemical identity of this molecule under a range of crystallisation conditions opens up the possibility of synthesising several types of ICCs with a range of physicochemical properties[50], thereby further widening the utility of multicomponent solid forms within industrial solid-state screening laboratories. The charged intermolecular interactions between the ions are expected to be much stronger than the interactions between the neutral molecule and the salt complex. This raises interesting questions about the role of this additional neutral molecule: What is the structural role of the additional neutral molecule within the lattice? Does the inclusion of an additional neutral molecule within the lattice imply that ICCs with lattice stoichiometries such as 1:1:1 are thermodynamically more stable than the sum of the lattice energies of the 1:1 salt and the neutral acid molecule? Are ICCs more likely to be observed for acid-base combinations where the putative packing modes available to the charged ions of the hypothetical salt have low packing coefficients? All of these are interesting yet unanswered questions that will be probed in this contribution. When compared to cocrystals, salt solid forms are generally known to be more prone to crystallising in lattices with molecular stoichiometries that deviate from those expected (i.e. 1:1 or 2:1) on the basis of the number of ionisable functional groups on the acid and base, thereby leading to salt solvates[51]. The literature is replete[47, 48, 52, 53] with examples of ICCs comprised of coordination complexes of neutral and ionised forms of organic compounds bound to a metal centre. However, there remain unanswered questions about whether organic molecular ICCs are largely a serendipitous[54-56] result of the crystallisation experiment whose observation is difficult to predict *a priori* or whether there are underlying trends in the molecular properties of the constituent molecules/ions of the ICC, which make them amenable to purposeful design using computational methods of CSP.

Here, we compare the potential of two relatively simple organic bases (**1-2**) to form salts and molecular ICCs using the results of previously reported cocrystallisation experiments involving the combination of **1** or **2** with three aromatic monoprotic acids (**3-5**) with a single substitution at the *ortho* or *para* positions (Table 1). These molecules (**1-5**) were carefully chosen to provide a contrast in the behaviour of 4-aminopyridine (**1,** 4-AP) which forms salts with the acids **3-5** and 4-dimethylaminopyridine (**2,** 4-DMAP) which forms both salt

(1:1) and molecular ICC structures (1:1:1). Table 1 summarises the numbering system for the 1:1 salts (denoted **I-VI**) that are expected following the cocrystallisation of **1** or **2** with the acids **3-5** as well as the previously reported multicomponent solid forms reported in the Cambridge Structural Database (CSD)[57] for each combination of acid and base. The objectives of this work are to probe the underlying molecular and crystal properties that give rise to ICC crystallisation. Given that ICC solid forms could be synthesised via the cocrystallisation of stoichiometric amounts of a salt with an acid[50, 58, 59], we can consider ICC formation to be a type of "cocrystallisation" experiment between a salt and an acid coformer. Computed CFLs of model 1:1 salts are therefore of interest in understanding whether ICC crystallisation could be inferred on the basis of the packing modes available to the salt ions. Calculated DFT-D relative lattice energies of a series of experimentally characterised 1:1:1 ICCs and their corresponding 1:1 salt and acid crystal structures were used to quantify the thermodynamic driving force for ICC crystallisation. The calculated packing efficiencies of the ions in the predicted polymorphs of the 1:1 salts were used to quantify the degree to which ICC crystallisation is driven by packing frustration of the salt ions.

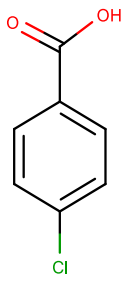
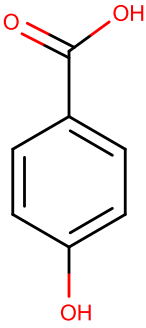
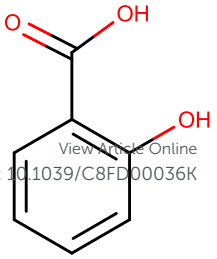
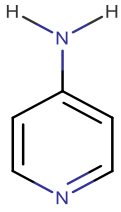
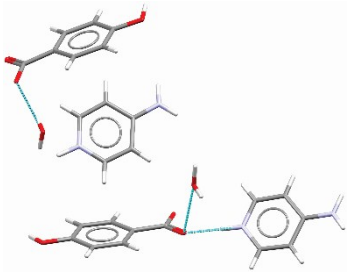
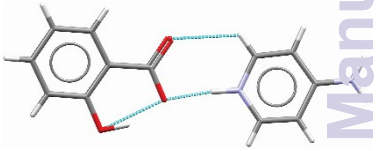
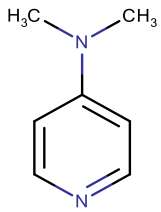
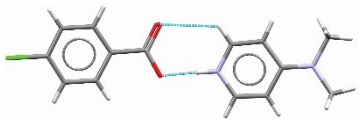
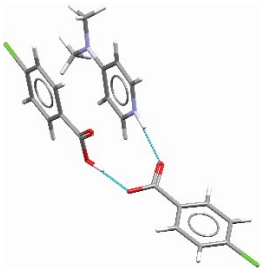
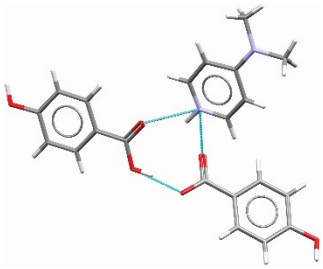
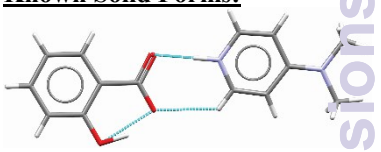
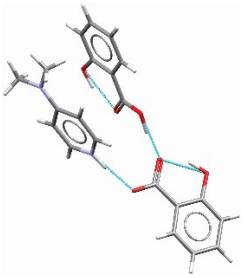
	 3 4-Chlorobenzoic Acid	 4 4-Hydroxybenzoic Acid	 5 2-Hydroxybenzoic acid
 1 4-Aminopyridine	**I**=4-Aminopyridinium 4-chlorobenzoate (1:1) **NO SALT/ICC KNOWN**	**II**=4-Aminopyridinium 4-hydroxybenzoate (1:1) **Known Solid Forms:**  **MOYQOH**[60] 4-Aminopyridinium 4-hydroxybenzoate monohydrate	**III**=4-Aminopyridinium 2-hydroxybenzoate (1:1) **Known Solid Forms:**  **DUSYOG**[61] 4-Aminopyridinium 2-hydroxybenzoate
 2 4-Dimethylaminopyridine	**IV**=4-Dimethylaminopyridinium 4-chlorobenzoate (1:1) **Known Solid Forms:**  **CUKNED**[62] 4-Dimethylaminopyridinium 4-chlorobenzoate  **CUKNON**[62] 4-Dimethylaminopyridinium 4-chlorobenzoate 4-chlorobenzoic acid	**V**=4-Dimethylaminopyridinium 4-hydroxybenzoate (1:1) **Known Solid Forms:**  **SOLGUX**[63] 4-Dimethylaminopyridinium 4-hydroxybenzoate  **CUKNUT**[62] 4-Dimethylaminopyridinium 4-hydroxybenzoate 4-hydroxybenzoic acid	**VI**=4-Dimethylaminopyridinium 2-hydroxybenzoate (1:1) **Known Solid Forms:**  **KUJDEA**[58] 4-Dimethylaminopyridinium 2-hydroxybenzoate  **KUJDIE**[58] 4-Dimethylaminopyridinium 2-hydroxybenzoate 2-hydroxybenzoic acid

Table 1: Grid depicting the numbering system (**I-VI**) for the 1:1 salts resulting from the combination of **1** or **2** with the acids **3-5**. The asymmetric units of the known experimental solid forms following cocrystallisation of the bases **1** or **2** with the acid coformers **3-5** under varying conditions are shown. Cambridge Structural Database (CSD)[57] reference codes for the known crystal structures are shown in bold capital letters. For all acid-base combinations, the calculated[64] aqueous ΔpK_a was in the range 4.2-6.2, illustrating high probability of ionisation to form the salt[65-67].

1.2. Methodology

1.2.1. Equilibrium Solubility Measurements

The saturation temperatures (T_s) corresponding to saturated solutions of **1-5** were determined in methanol solvent assuming a total solvent volume of 1 ml. Turbidity measurements as implemented in the Crystal16 parallel crystallisation unit (Technobis Crystallization Systems) were used to determine the solution saturation temperatures. Four heating and cooling cycles were set up for each compound **1-5** at a range of solution concentrations starting with a saturated solution of the compound at room temperature. Solution concentrations differed by approximately 5-10 mg/ml for each system. Each heating and cooling cycle had a minimum temperature of -15°C and a maximum temperature of 50°C. A constant heating/cooling rate of 1°C/min was used for all experiments. Nitrogen purge gas was supplied to the reactors at a constant pressure of 0.4 bar. The contents of each vial were stirred at a rate of 500rpm using a single magnetic flea stirrer bar using the bottom-stirring option of the Crystal16 unit. The resulting solubility data were analysed according to the measured clear and cloud points. The Van't Hoff equation was used to estimate the mole fractions (x_m) of each solute at a reference saturation temperature (T_s) of 298 K.

1.2.2. Crystal Structure Prediction

The crystal form landscapes (CFLs) for the 1:1 salt (**I-VI**) complexes (Table 1), were calculated using the Polymorph Predictor module of *BIOVIA Materials Studio 8.0*[68], which uses a Monte Carlo simulated annealing method[69, 70] to generate trial crystal structures. For the salt complexes, we assumed a 1:1 ratio of the ions in the asymmetric unit. Hypothetical crystal structures were generated in the following space groups: $P\bar{1}, P2_1, P2_1/c, P2_12_12_1, Pbca, C2/c$. Motion groups were defined for each molecule/ion in the asymmetric unit and these motion groups were kept rigid throughout the cell refinement stage. For all systems except **II** and **V**, the assumed input geometry for the ions in the asymmetric unit corresponded to the calculated global minimum conformation obtained from gas phase geometry optimisation at the MP2/6-31G(d,p) level of theory using *GAUSSIAN09*[71]. For **II** and **V**, several input conformations of the 4-hydroxybenzoate anion (see Section 1.1.1 in the SI) were used in separate rigid body searches with the corresponding 4-aminopyridinium (**II**) and 4-dimethlyaminopyridinium (**V**) cation geometries fixed to their gas phase global minimum conformations. For all CSP searches, torsion angles were not optimised during the cell refinement stage. For each of the above space groups used to generate trial crystal structures, the search space for the Monte Carlo simulation was limited to 7000 steps. The minimum temperature that the simulation could reach was 300 K and the maximum temperature was 1×10^5 K. Initial estimates of the lattice energies of predicted crystal structures were made within the Polymorph Predictor module of *BIOVIA Materials Studio 8.0* using the Dreiding[72] force field with atomic charges derived from fitting to the molecular electrostatic potential of the *ab initio* wavefunction of the optimised conformation for the molecule/ion. The ChelpG[73] scheme as implemented in *GAUSSIAN09* was used to fit the atomic charges to the molecular electrostatic potential. Clustering of the predicted crystal structures was performed within *BIOVIA Materials Studio 8.0* in order to

remove duplicates of the predicted lattice energy minima using a tolerance of 0.13, a cut-off of 7 Å for the distance-distribution analysis using 140 bins and a maximum cluster size of 500.

View Article Online
DOI: 10.1039/C8FD00036K

1.2.3. Final energy rankings of predicted structures using empirical force fields

For all systems except **II** and **V**, the 1000 most stable structures produced following clustering were passed to DMACRYS[74] for lattice energy minimisation using a distributed multipole model[75, 76] for the electrostatic contribution towards the lattice energy. For systems **II** and **V**, since several rigid body searches were performed as a result of different input conformations for the 4-hydroxybenzoate anion, only the most stable 200 structures from each rigid body search were passed to DMACRYS. Multipoles were calculated up to rank 4 (hexadecapole) for all atoms by performing a distributed multipole analysis of the *ab initio* charge density using *GDMA2.2*[77]. The dispersion-repulsion contributions towards the lattice energy were estimated using a Buckingham *exp-6* function using the FIT potential parameters for C, H_C (hydrogen attached to carbon), N, O and Cl from the work of Williams[78-80] as well as the parameters for H_N[81] (hydrogen attached to nitrogen) and H_O[82] (hydrogen attached to oxygen) that were subsequently determined by fitting to crystal structures containing the N-H···O=C interactions and carboxylic acid structures respectively. For each search, a final clustering step was performed to remove all duplicate structures with lattice energies within 0.2 kJ mol^{-1}, cell volumes within 1.0 Å^3 and a powder pattern similarity index[83] of at least 0.97. The Crystal Packing Similarity module[84] of Mercury 3.9[83] was used to match predicted lattice energy minima with the observed salt structure retrieved from the CSD. This was done by estimating the root-mean squared deviation for matching at least 15 molecules ($RMSD_{15}$) in the coordination spheres of the experimental and predicted structures using a tolerance of 20% for the distances and 20° for the angles. In order to facilitate the crystal packing similarities, the lattice energies of the experimental structures were minimised using the same conformation and energy model used in the CSP. The Kitaigorodsky Packing Index (KPI)[85] of the predicted polymorphs on the CFLs of **I-VI** where calculated using PLATON[86] and the distribution of the KPI for the predicted polymorphs within 20 kJ mol^{-1} of the global minimum structure were used to rationalise the packing efficiencies of the ions. Computationally demanding DFT structural optimisations (atomic positions and unit cell parameters) were performed for a limited set of systems using the most stable structures (up to rank 20) following DMACRYS lattice energy minimisation. This was done in order to investigate the extent of re-ranking in the relative stability of the predicted polymorphs as a function of the DFT method used. Further details about the methodology for these calculations and a discussion of the results can be found in Section 1.4 of the SI.

1.2.4. Solid-state periodic DFT-D calculations for estimating the thermodynamic driving force for ICC crystallisation

Experimental crystal structures corresponding to molecular ICCs with 1:1:1 stoichiometry and their corresponding 1:1 salt and acid crystal structures were retrieved from the CSD and optimized using CASTEP 8.0[37] as implemented in *BIOVIA Materials Studio 8.0* using the Perdew-Burke-Ernzerhof (PBE) generalised

gradient approximation (GGA) exchange-correlation functional[87] and ultrasoft pseudopotentials[88]. The D2 (G06) semi-empirical dispersion correction of Grimme[89] was used in all cases. Brillouin zone integrations were performed on a symmetrised Monkhorst-Pack *k*-point grid[90] generated using the highest quality setting on *Materials Studio 8.0* corresponding to a *k*-point separation of 0.07 $Å^{-1}$. The plane-wave basis set cut-off was set at 340 eV and the self-consistent field (SCF) convergence on the total energy was set at 1×10^{-6} eV/atom. The BFGS[91] algorithm was used for geometry optimisation of all atomic positions and unit cell parameters, within the constraints of the space group symmetry. The structural optimisation was considered complete when the following convergence criteria were satisfied: maximum energy change of 1×10^{-5} eV/atom, maximum force of 3×10^{-2} eV/Å, maximum stress of 5×10^{-2} GPa and maximum displacement of 1×10^{-3} Å. In some cases, the BFGS algorithm did not converge after reaching 100 iterations and geometry optimisation was performed stepwise by first optimising only the atomic positions (fixed unit cell) followed by a full geometry optimisation on all atomic positions and the unit cell. Such a step-wise approach has been shown[16, 92] to aid convergence in the optimisation of organic crystal structures when using DFT methods. Where a system displayed crystal polymorphism, the most stable reported structure (as determined using thermal methods) was used to carry out the DFT-D lattice energy calculations.

1.3. Results and Discussion

1.3.1. Molecular and crystal properties of systems known to form ICCs: Insights from the CSD

A search of the Cambridge Structural Database (CSD)[57] for all three-component (Z''=3)[93] crystal structures consisting of the general formula A^-B^+N (Figure 1) where the cation is a pyridinium moiety, the anion is an aromatic carboxylate and N is any neutral molecule has revealed that there are a total of 185 crystal structures in the database that match the search query. Approximately 52% of these structures were salt hydrates where the additional chemical species N is a molecule of water. Molecular ICCs where the additional molecule N is the same acid used to form the anion (A^-B^+A) account for 27%, whilst ICCs where the additional molecule N is the base (A^-B^+B) account for only 2.7%. The remaining 18% of structures are ternary ICCs where N is a distinct molecular species to the acid or base used to form the salt.

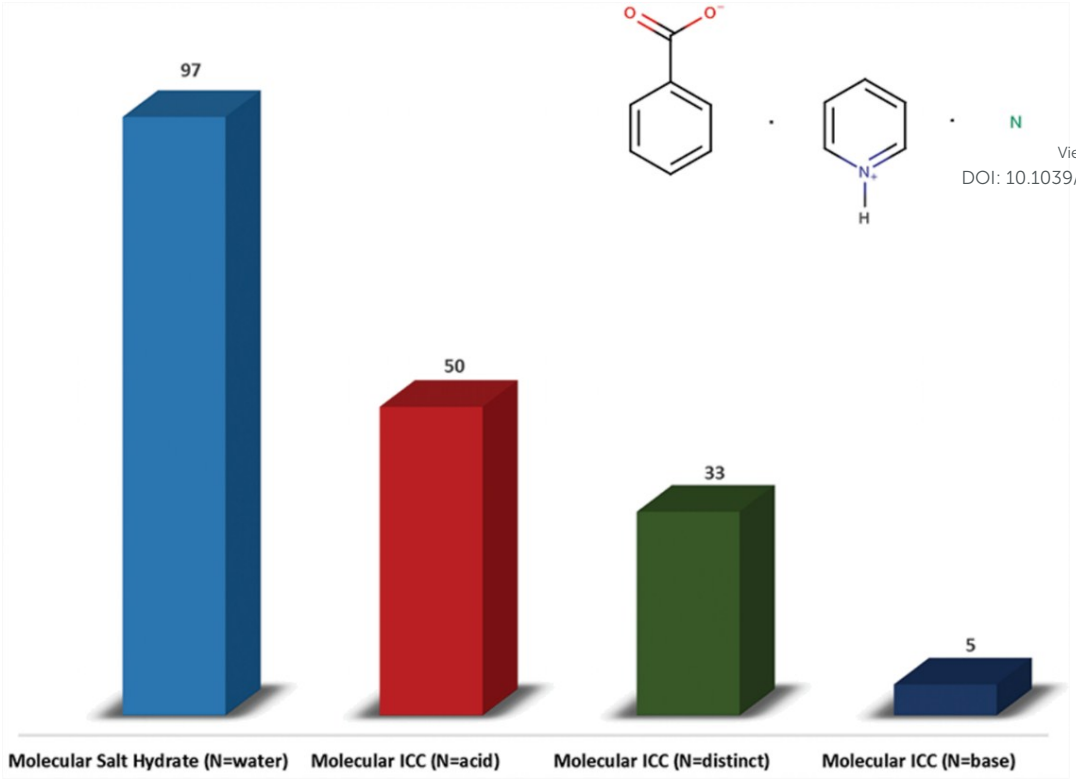


Figure 1: Bar graph depicting the frequencies of different types of organic molecular ICCs retrieved from the CSD. All categories of structures depicted above consist of 1:1:1 stoichiometry of anion:cation:neutral molecule (N) in the lattice, where N is a molecule of water, acid, base or a distinct molecular species to the acid and base. The inset on the top right-hand corner shows the search query used in the CSD search. The Cambridge Structural Database Version 5.37 (November 2015 + 3 updates up to and including the May 2016 release) was searched using Conquest V1.18 to retrieve the crystal structures of all pyridinium carboxylate salts containing 3 residues. The following filters were used in the search: 3D coordinates determined, not disordered, no errors, not polymeric, no powder structures and only organics.

With the exception of salt hydrates, given that molecular ICCs with 1:1:1 stoichiometry are predominantly those where the additional molecule N is the same acid that forms the anion (A^-B^+A), we conducted a careful analysis of the molecular properties of the acid-base pairs used to crystallise the ICCs with the view of gaining a better understanding of the underlying molecular properties that favour ICC crystallisation. All acids that form ICCs have molecular structures with one or more hydrogen bond acceptor groups and in 41% of the acids surveyed, there were more acceptor groups than donor groups. By contrast, all the basic molecules used to form ICCs have more acceptors than donors in line with expectation. The polar surface area (PSA) of the acid molecules, which is a good measure of the potential of a molecule to engage in intermolecular hydrogen bonding, is on average 79 $Å^2$ across the sample space compared to just 19 $Å^2$ across all the basic molecules known to form ICCs (Figure 2). This suggests that there is a significant mismatch in the available molecular surface area that could be used to form intermolecular hydrogen bond heterosynthons in acid-base pairs that are prone to crystallising as ICCs.

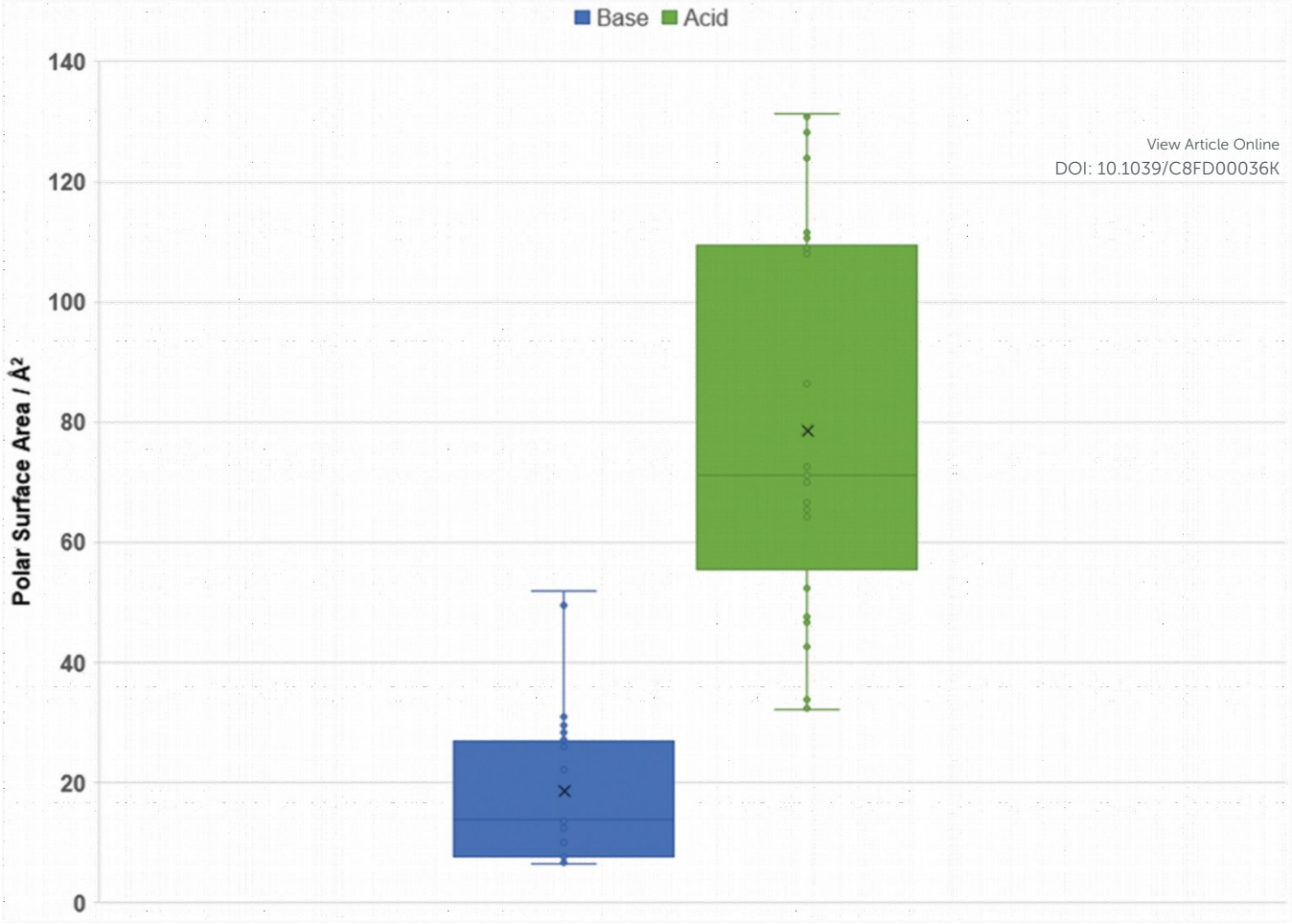


Figure 2: Box-plots depicting the distribution in the calculated polar surface areas (PSA) of the acid (green) and base (blue) molecules that form ICCs according to the data retrieved from the CSD. The PSA for each acid or base was estimated using Spartan'16[94] via the summation of the area (in Å^2) occupied by nitrogen, oxygen and hydrogens attached to either nitrogen or oxygen. The Corey-Pauling-Koltun (CPK)[95] model was used to represent the atoms. The mean value for each distribution is shown as a cross within the box-plot for each category. See Table S2 in the SI for more details.

The hydrogen bond synthons observed in molecular ICCs of the type A^-B^+A (A^-=anion, B^+=cation and A=acid; 1:1:1) retrieved from the CSD display two common types of interactions, which are denoted as Type 1 or Type 2 (Figure 3). In both cases, the acid engages in discrete $O\text{-}H_{acid}\cdot\cdot O_{anion}$ [graph set: $D_1^1(2)$] hydrogen bonding interactions with the oxygen of the anion. The key difference between the motifs is that the Type 1 motif occurs in ICCs where the OH donor of the acid and the N^+-H donor of the cation are both engaged in hydrogen bonding interactions with the same oxygen of the carboxylate group, whilst in Type 2 interactions, these donor groups interact with different hydrogen bond acceptors on the anion. The Type 2 motif is found in 39% of all ICC structures surveyed whilst the Type 1 motif occurs in just 18% of structures. Approximately 63% of all ICCs that display a Type 1 motif were found to have cation-anion heterosynthons with the $R_2^2(7)$ graph set, whereas structures with Type 2 motifs all displayed discrete $D_1^1(2)$ heterosynthons between the cation and anion.

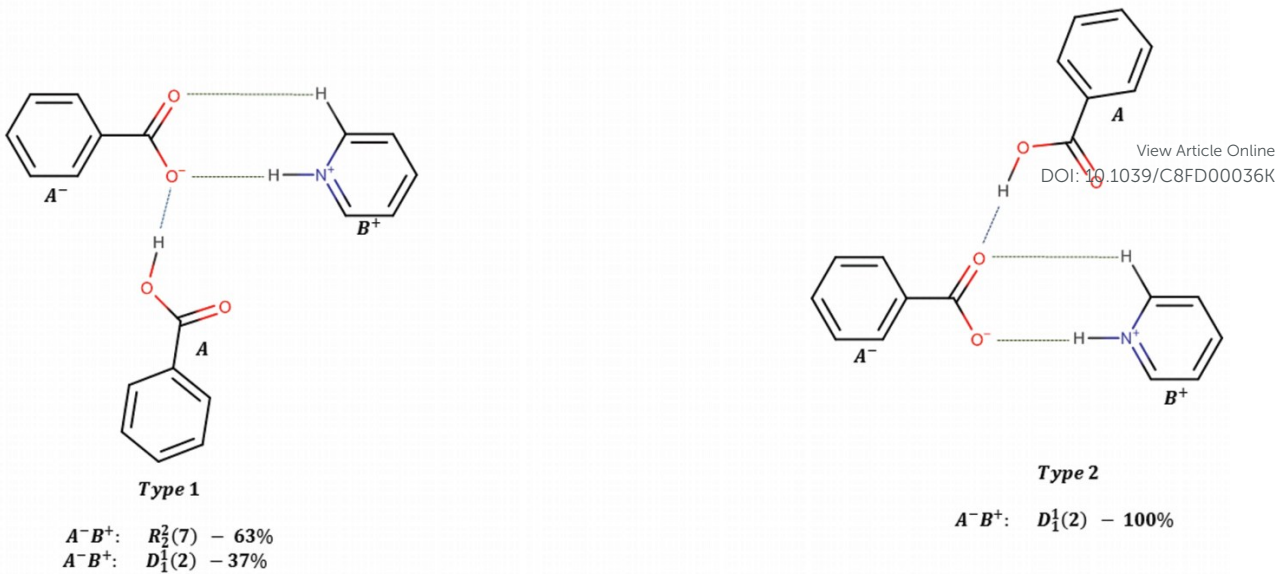


Figure 3: Illustration of the common Type 1 and Type 2 hydrogen-bond heterosynthons found in the crystal structures of molecular ICCs with the formula A^-B^+A (A^-=anion, B^+=cation and A=acid) derived from the cocrystallisation of substituted carboxylic acid and pyridine molecules. The hydrogen bond heterosynthons are depicted using graph-set notation[96].

1.3.2. Estimating the thermodynamic driving force for ICC crystallisation

Dispersion-corrected density functional theory methods (DFT-D) were used to estimate the relative lattice energies of molecular ICCs with 1:1:1 stoichiometry (A^-B^+A) with respect to the lattice energies of the corresponding 1:1 salt (A^-B^+) and acid (A) structures known for each system. These relative lattice energies provide an indication of the thermodynamic driving force for ICC crystallisation given that the cocrystallisation of equimolar amounts of a salt and acid is known[50, 58, 59] to lead to the formation of molecular ICCs. Only 7 (of the total 50) well-characterised molecular ICC crystal structures retrieved from the CSD were found to have known 1:1 salt and acid crystal structures and hence only these systems were used to carry out the relative stability estimates. In each case, the thermodynamic driving force for ICC crystallisation (ΔE_{ICC}) was estimated according to the equation

$$\Delta E_{ICC}(A^-:B^+:A) = E_{ICC}(A^-:B^+:A) - [E_{SALT}(A^-:B^+) + E_{ACID}(A)]$$

where $E_{ICC}(A^-:B^+:A)$, $E_{SALT}(A^-:B^+)$ and $E_{ACID}(A)$ are the total lattice energies of the 1:1:1 ICC, 1:1 salt and acid crystal structures respectively expressed as per formula units.

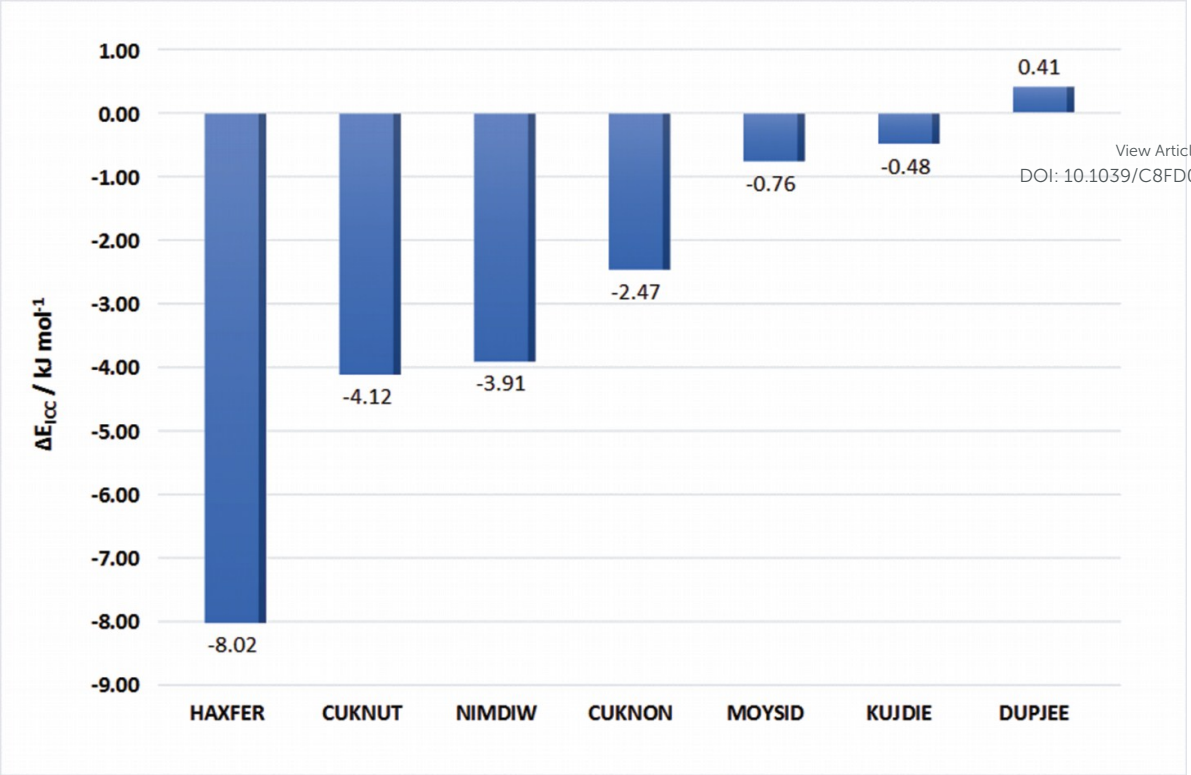


Figure 4: Relative stabilities (ΔE_{ICC}) of the 1:1:1 experimental ICCs (labelled according to CSD refcodes) with respect to the stabilities of the known experimental 1:1 salt and acid crystal structure. All relative stabilities were calculated using PBE+D2 lattice energies of the experimental ICC, salt and acid crystal structures obtained following full structural optimisation with CASTEP.

The relative energy estimates (Figure 4) revealed that for 6 out of the 7 systems surveyed, the 1:1:1 ICC crystal structure was more stable than the sum of the energies of its corresponding 1:1 salt and acid crystal structures. In the case of **HAXFER**[97], a molecular ICC formed from 3,5-dimethylpyridine and 3,5-dinitrobenzoic acid, the ICC is approximately 8 kJ mol^{-1} more stable than the sum of the lattice energies of the 1:1 salt and acid structures. The ICCs **CUKNUT**[62] and **NIMDIW**[98], which were both crystallised using mono-substituted hydroxybenzoic acids also have clear thermodynamic stabilities of approximately 4 kJ mol^{-1} with respect to their component salt and acid structures. In all three cases (**HAXFER**, **CUKNUT** and **NIMDIW**), careful inspection of the hydrogen bond synthons in the molecular ICCs when compared to those observed in the salt structures shows that the additional acid molecule in the ICC is engaged in discrete hydrogen bonding interactions with the carboxylate of the anion. In all cases, there is a significant perturbation to the crystal packing of the salt ions upon inclusion of the acid in the lattice leading to the ICC structure (Figure 5). For example, in **HAXFER** which displays the highest relative stability with respect to its constituent salt and acid structures, the $R_2^2(7)$ dimer motif between the ions in the salt (**PUHRIT**[99]) is lost in the ICC, potentially due to the significant change in the relative packing of the ions as a result of acid inclusion into the lattice. Instead, discrete $D_1^1(2)$ acid-anion hydrogen bonds are formed. By comparison, for the **DUPJEE**[100] structure which has a lattice energy that is comparable in value to the sum of the lattice energies of the salt and acid crystal structures, the formation of the ICC does not lead to a change in the hydrogen bonded synthons involving the ions. This is evidenced by the persistence of the $R_2^2(7)$ dimer motif between the salt ions, which is a constant feature (Figure 5) in both the salt (**BOVTEM**[101]) and ICC (**DUPJEE**) crystal

packings. For three of the observed molecular ICCs (**MOYSID**[60], **KUJDIE**[58] and **DUPJEE**[100]), the calculated energy differences are too small (< 1 kJ mol^{-1}) to be indicative of a clear energetic driving force for the formation of a molecular ICC with 1:1:1 stoichiometry. The crystallisation of **MOYSID**[60] may be an interesting case study in the kinetics of ICC crystallisation as the authors report[60] the appearance of single crystals of **MOYSID** (3-aminopyridinium 3,5-dihydroxybenzoate 3,5-dihydroxybenzoic acid; 1:1:1) from a methanol solution of the acid and base following solvent evaporation at low temperatures inside a domestic refrigerator. If a solution with the same composition is allowed to undergo slow solvent evaporation under ambient conditions, single crystals of the 1:1 molecular salt 3-aminopyridinium 3,5-dihydroxybenzoate (**MUCMUT**[60]) are observed instead. Both the ICC and salt are reported to crystallise from the same methanol solution prepared via the cocrystallisation of equimolar amounts of 3-aminopyridine and 3,5-dihydroxybenzoic acid. The small relative energy ($\Delta E_{ICC} = -0.76\ kJ\ mol^{-1}$) of **MOYSID** with respect to the known salt and acid solid forms, may therefore be indicative of a kinetic ICC crystallisation product. The lattice energies of a wider set of experimental ICCs and their corresponding salt and acid structures as well as the experimental conditions used to target the observed solid forms must be surveyed before drawing any general conclusions about the thermodynamic driving force for ICC crystallisation. But the evidence from this limited set of structures is that DFT-D relative energies can be diagnostic enough to guide the selection of acid-base pairs with potential to form stable molecular ICCs when coupled with CSP or X-ray diffraction techniques for determining the crystal structures.

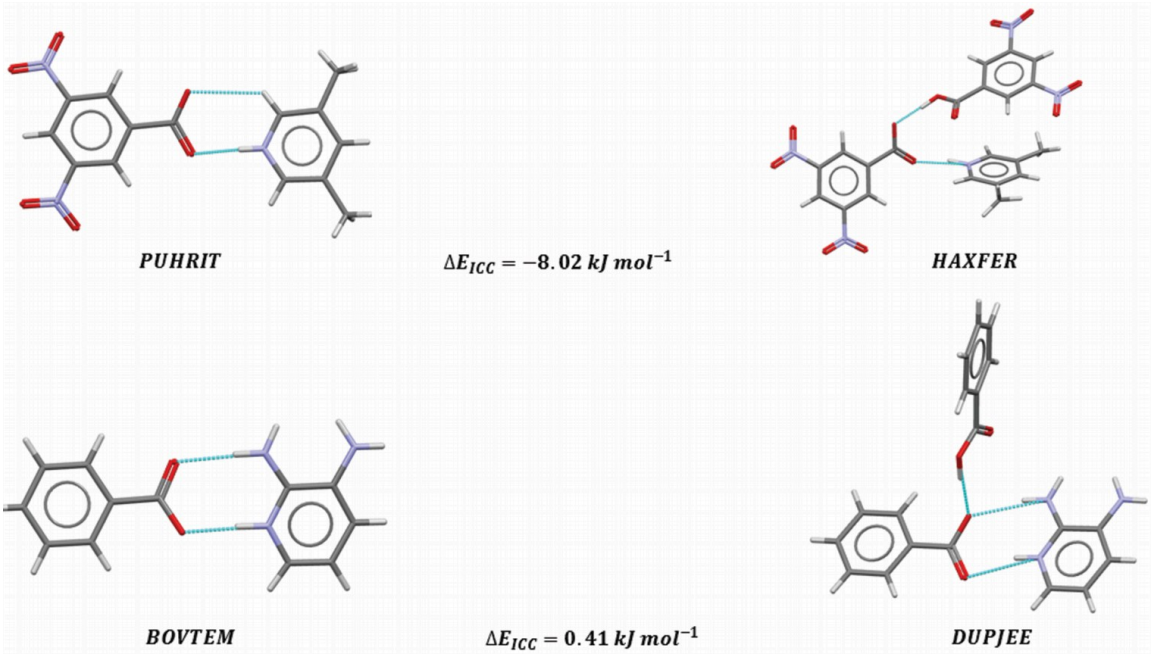


Figure 5: Comparison of the differences induced in the hydrogen-bonded heterosynthons of the ions upon switching from the 1:1 salts **PUHRIT** and **BOVTEM** to three-component ICCs **HAXFER** and **DUPJEE**, respectively. The quoted ΔE_{ICC} values correspond to the calculated DFT-D relative lattice energies of these ICCs with respect to the known salt and acid structure for each system as reported in Figure 4.

1.3.3. Solubility differences of acid-base pairs known to form molecular ICCs

Differences in the solubilities of acid-base pairs are known to affect the stoichiometry of cocrystal solid forms[102]. The measured solubility differences of acid-base pairs are therefore critical parameters to consider when conducting multicomponent crystallisation screens. The Crystal16 platform was used to measure the

saturation temperatures (T_s) as a function of the solution concentrations for **1-5** (Table 1) in methanol solvent. Methanol was chosen as the solvent of choice given that all the ICCs (**CUKNON, CUKNUT** and **KUJDIE**) of interest in this study (see Table 1) were crystallised from this solvent. The intention behind the solubility screen was to obtain a quantitative initial assessment of the likely solution concentrations of the acid-base pairs in methanol solvent at a given saturation temperature. This was deemed as a necessary first step in conducting more detailed studies on the effects of component solution concentrations on the stoichiometry of ICCs. To the best of our knowledge, the use of ternary phase diagrams for exploring the range of solution concentrations for targeting ionic cocrystals and the correlation of such concentrations with the stoichiometry of the resulting solid form are not typically reported in the literature. For a narrow temperature range, the Van't Hoff equation can be used to describe the variation in the mole fraction (x_m) of a given solute (**1-5**) as a function of the experimental saturation temperature, (T_s), according to the equation

$$ln\, x_m = -\frac{\Delta H_{fus}}{R}\left(\frac{1}{T_s} - \frac{1}{T_0}\right)$$

where ΔH_{fus} is the enthalpy of fusion, R is the gas constant and T_0 is the melting temperature. For a series of experimental x_m values corresponding to different solution concentrations of **1-5**, the saturation temperatures of these solutions were measured using turbidity sensors. Linear fitting of the experimental solubilities and saturation temperatures from the Crystal16 platform were used to calculate the mole fraction (x_m) corresponding to a saturated solution of **1-5** at a reference saturation temperature of $T_s = 298\ K$. Figure 6 shows the x_m values at $T_s = 298\ K$ for **1-5** in methanol solvent. Comparison of the relative solubilities of the acid-base pairs in methanol shows that 4-chlorobenzoic acid (**3**) had the lowest solubility whilst 4-hydroxybenzoic acid (**4**) and 2-hydroxybenzoic acid (**5**) have comparable solubilities in methanol. The solubility of 4-dimethylaminopyridine (**2**) was at least a factor of 3 larger than the solubilities of **4** and **5**. By contrast 4-aminopyridine (**1**) had a marginally higher solubility than **4** and **5**. For this limited set of systems studied, acid-base pairs known to form ICCs (combinations involving **2** and the acids **3-5**) therefore have greater differences in relative solubilities at a given saturation temperature when compared to acid-base pairs that only form 1:1 salts or hydrates of 1:1 salts (combinations involving **1** and the acids **4** and **5**) from solution crystallisation experiments. The extent to which a comprehensive experimental screen for ICCs of **1** with the acid coformers **3-5** was performed and how exhaustive[103] such screens were in being confident that ICC crystallisation is not favourable, remains a clear caveat in this analysis. However, given that chemists often rely on specific molar stoichiometric ratios of acid-base pairs when performing cocrystallisation experiments rather than relative concentrations of the acid or base at a given saturation temperature, the significant differences in the solubilities of **2** and each of the acids **3-5** show that it is important to consider the solution concentrations of the acid-base pair as this will determine the shape of the ternary phase diagram[104] and hence the composition of the resulting multicomponent solid form under a given set of experimental conditions.

Several molecular ICCs[100, 105-107] with 1:1:1 stoichiometry of the type A^-B^+A retrieved from the CSD were synthesised from solution crystallisation experiments involving 1:1 molar ratios of the acid and base. It is therefore conceivable that a 1:1 molar ratio of the acid:base translates to a slightly supersaturated solution of the acid than is assumed in the molar ratios, especially in acid-base pairs (like those involving **2** and the acids **3-5**) where the differences in the solubilities are significant. This is worthy of further investigation by looking at a wider set of relative solubilities for acid-base pairs known to form molecular ICCs.

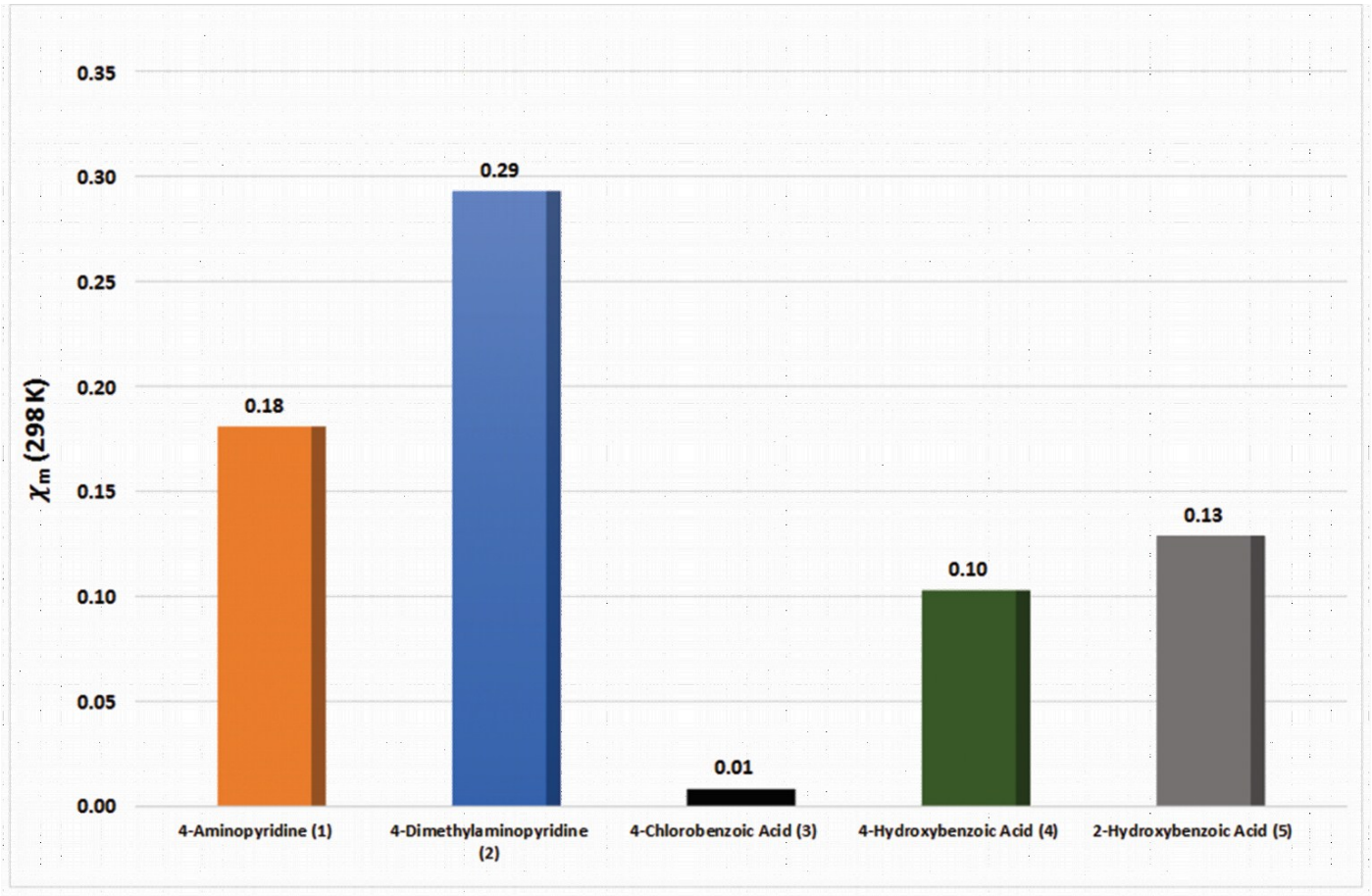


Figure 6: Equilibrium solubilities expressed as mole fractions, x_m, for the molecules **1-5** in methanol solvent. The mole fractions, x_m, were derived for all molecules **1-5** in methanol solvent at a reference saturation temperature of 298 K. All mole fractions, x_m, shown above were derived from solubility data in methanol solvent using the Van't Hoff equation at a defined saturation temperature of 298 K.

1.3.4. Model crystal form landscapes of 1:1 salts: packing efficiencies of ions and inferences about ICC crystallisation

Model crystal form landscapes (CFLs) of the 1:1 salts **I-VI** (Table 1) were calculated in an attempt to see if the packing efficiencies of the salt ions could provide a strong indication as to the potential of the salt system to form a molecular ICC. The hypothesis being tested here is that molecular salts with inefficient packing of the ions are likely to form 1:1:1 ICCs (A^-B^+A) via the inclusion of a neutral acid molecule (A) into the salt lattice following the cocrystallisation of equimolar amounts of A and the 1:1 salt (A^-B^+), provided there is a thermodynamic driving force (Figure 4) for the crystallisation of the ICC. This is worthy of investigation

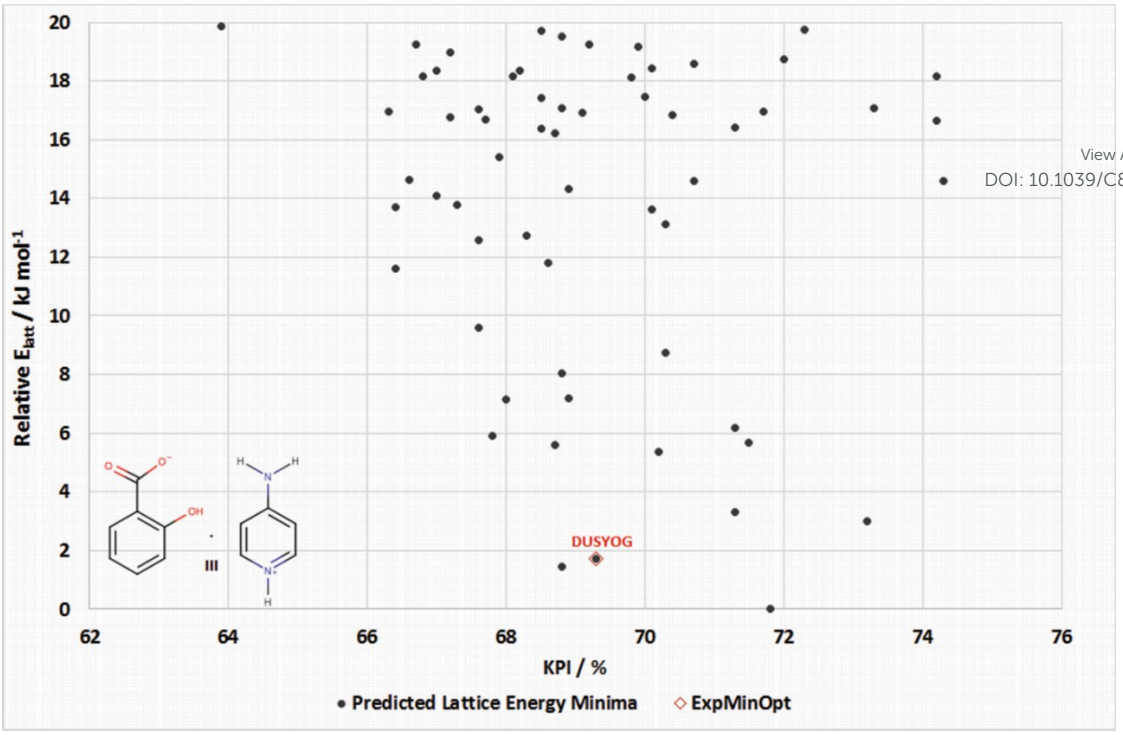
since the CSD survey (Figure 1) revealed a number of ICC crystal structures that were described as "acid solvates"[100, 106-108] according to their entry in the CSD or in one case, the title of the structure report itself[100]. The implication here, is that the neutral acid molecule occupies the empty space in the ICC lattice created by the inefficient packing of the salt ions, in the same way that solvent molecules such as water may be said to occupy the voids in the crystal lattice of salt hydrates[45, 109]. The 4-dimethylaminopyridinium 2-hydroxybenzoate 2-hydroxybenzoic acid ICC (**KUJDIE**) with 1:1:1 stoichiometry can be prepared from the solution crystallisation of equimolar amounts of 2-hydroxybenzoic acid (**5**) and 4-dimethylaminopyridinium 2-hydroxybenzoate (1:1) salt (**VI, KUJDEA**) in methanol solvent[58]. By contrast, the cocrystallisation of equimolar amounts of **5** with 4-aminopyridine (**1**) leads to a 1:1 salt of 4-aminopyridinium 2-hydroxybenzoate (**III, DUSYOG**) with no ICC of the salt and **5** reported in the CSD (Table 1). Of the 1:1 salt systems **I-III** (Table 1) that could potentially be formed via the cocrystallisation of equimolar amounts of **1** and each of the acids **3-5**, only the combination **1** and **5** has a known anhydrate 1:1 salt structure (**III, DUSYOG**). The combination of **1** and **3** does not have a known salt or ICC structure in the CSD and cocrystallisation of **1** and **4** leads to a salt monohydrate (**MOYQOH**). Figure 7 contrasts the CFLs of the salt systems **III** (Figure 7a) and **VI** (Figure 7b), which are plotted for each predicted lattice energy minimum in terms of the Kitaigorodsky Packing Index (KPI, %) as a function of the relative lattice energy using the predicted global minimum structure on each CFL as the reference point.

The CFL of **III** shows that the experimental structure is the third ranked structure in lattice energy (Figure 7a) using the empirical FIT potential coupled with a distributed multipole electrostatic model. The predicted rank 3 structure corresponding to the experimental salt structure of **III** (**DUSYOG**) displays a KPI of 69.3%. By contrast the CFL of **VI** shows that the experimental salt structure (**KUJDEA**) is the global minimum in lattice energy. The CFL of **VI** displays fewer energetically competitive packings of the salt ions, with the rank 2 structure approximately 4 kJ mol^{-1} higher in energy when compared to the predicted global minimum structure. Structural optimisation (atomic positions and unit cell parameters) of the most stable predicted polymorphs (up to rank 20) for **III** and **VI** using the PBE functional[87] combined with a plane-wave implementation of the Grimme D2[89] dispersion correction (PBE+D2), has shown (see Section 1.4 of the SI for more details) that there is significant re-ranking in the stabilities of the predicted polymorphs when the results of the empirical force field calculations are compared with those obtained from the PBE+D2 method. For **III**, the experimental rank 3 structure from the empirical force field calculations is the global minimum in lattice energy using the PBE+D2 method, with the second most stable structure (originally the global minimum using the empirical force field) approximately 1.89 kJ mol^{-1} higher in energy. Optimisations performed with the PBE functional coupled with the TS dispersion correction (PBE+TS)[110] as well as optimisations performed using the uncorrected LDA functional, reveal that in both cases the experimental structure of **III** is ranked as the second most stable structure. For **VI**, the re-ranking observed is much more pronounced and sensitive to the DFT method used. PBE+D2 and PBE+TS optimisations reveal that the experimental structure of **VI** changes from rank 1 using the empirical force

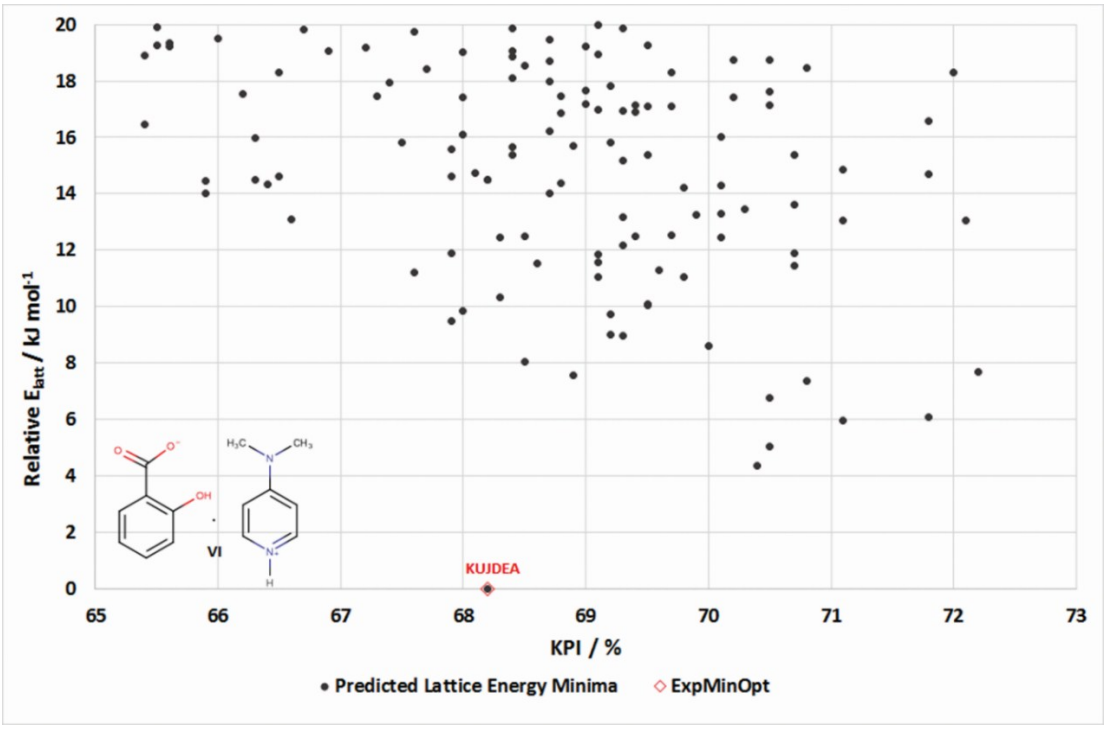
field to rank 4 with relative lattice energies of 3.61 kJ mol^{-1} and 2.26 kJ mol^{-1} respectively. The results of the DFT optimisations show that the relative energies of predicted polymorphs for molecular organic salts are sensitive to the type of DFT method used. This contrasts with recently reported results on the relative energies of metal pentazolate[111] framework structures, where the rankings of the structures were consistent across a range of dispersion-correction schemes used.

The KPI of the predicted global minimum structure of **VI** is 68.2% and is comparable in value to the predicted rank 3 structure on the CFL of **III** which has a KPI of 69.3%. The comparable KPI in the predicted lattice energy minima of **III** and **VI** that correspond to the experimental structures **DUSYOG** and **KUJDEA** respectively, suggest that the observation of an ICC in **VI** is a complicated phenomenon that has more to do with the finely balanced competing intermolecular interactions in both systems rather than arguments based on packing efficiency alone. This is supported by a closer inspection (Figure 8) in the distribution of the calculated KPIs of all predicted polymorphs for salts **I-III** (which do not form ICCs) and salts **IV-VI** (which do form ICCs). Within a relative lattice energy range of 20 kJ mol^{-1} with respect to the global minimum structure, the majority of the predicted polymorphs of all salts have KPIs that are in the range 65-80%, which is within the range we would expect for stable molecular crystals[112]. Moreover, the most stable polymorphs on each CFL have KPIs (see SI: Tables S3-S8) in the range 65-70%, which suggests no packing frustration of the salt ions. However, approximately 20% of the predicted salt polymorphs for **I** and **II** within a relative lattice energy range of 20 kJ mol^{-1} with respect to the global minimum display low KPIs (60-65%) when compared to the KPIs of the lattice energy minimum corresponding to the experimental salts of **III** and **VI**. This may be significant given that anhydrate salt structures for **I** and **II** have yet to be reported.

Careful contrast of the hydrogen bond motifs adopted by the predicted low energy structures in the CFLs of **III** and **VI** show that all predicted polymorphs of **III** within 5 kJ mol^{-1} of the global minimum structure display N-H··O hydrogen bonds between the 4-amino group of the cation and the oxygen of the carboxylate/hydroxyl group on the anion (Figure 9). The crystal form landscapes of the most stable salt polymorphs of **I** and **II** also display similar N-H··O hydrogen bonds between the ions. These N-H··O hydrogen bonds in **III** are analogous to the Type 2 O-H_{acid}··O_{anion} hydrogen bonds (Figure 3) found in some of the ICCs reported in Table 1.



(a)



(b)

Figure 7: Comparison of the calculated crystal form landscapes (CFLs) of the 1:1 salts **III** (a) and **VI** (b). The CFLs are plotted as relative lattice energies with respect to the calculated global minimum structure as a function of the calculated Kitaigorodsky Packing Index (KPI, %) for each lattice energy minimum. On each CFL, the predicted polymorphs are shown in dark-grey circles whilst the polymorph corresponding to the experimental structure (ExpMinOpt) is indicated in an open red diamond shape. See Tables S3-S8 in the SI for a detailed listing of the most stable predicted lattice energy minima for all the 1:1 salt complexes **I-VI**.

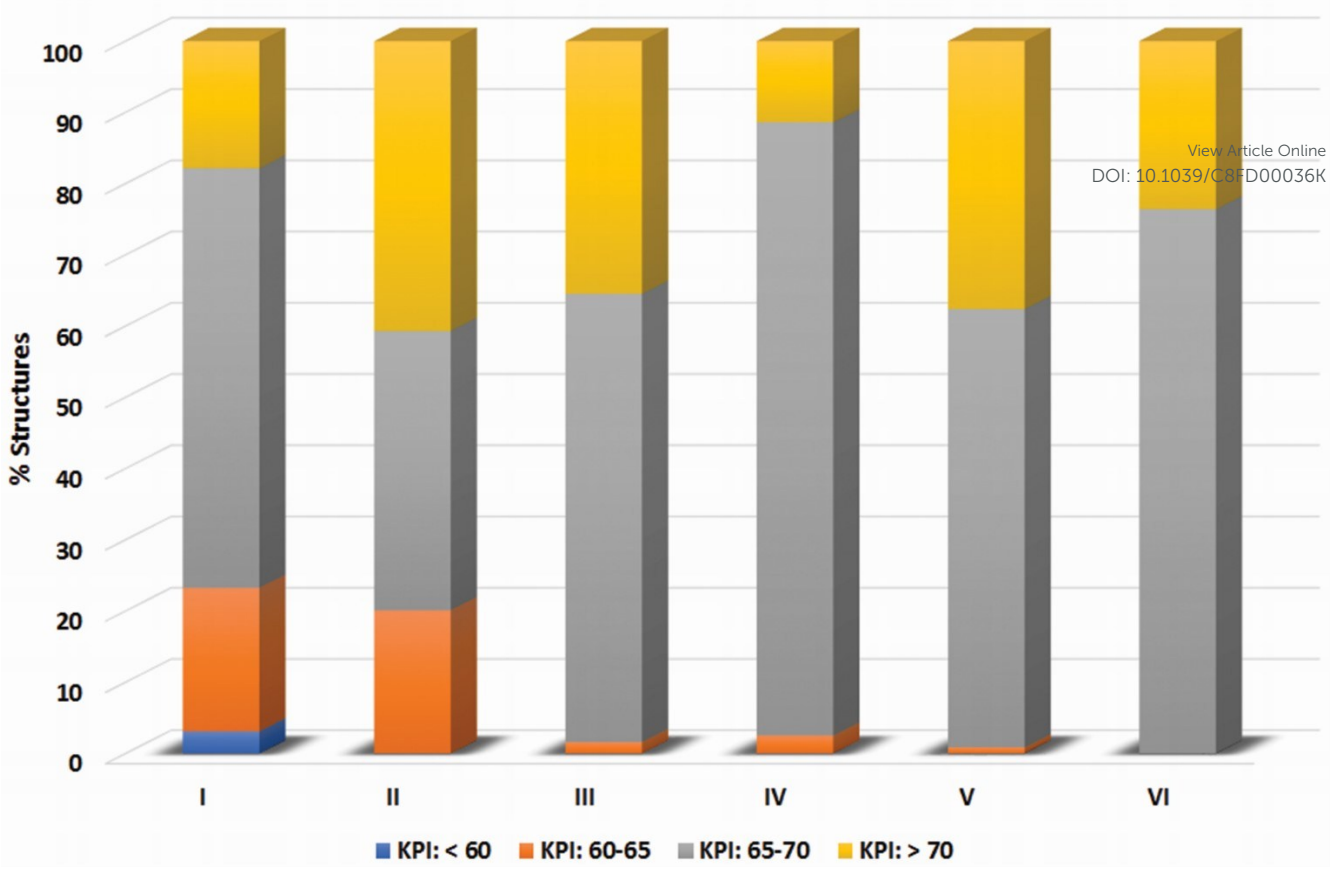


Figure 8: Distribution in the calculated Kitaigorodsky Packing Index (KPI) for the 1:1 molecular salts **I-VI** (Table 1) for the predicted lattice energy minima within 20 kJ mol^{-1} of the predicted global minimum structure for each system. The y-axis shows the percentage of structures with a particular KPI range (see legend). For all 1:1 salt systems, the majority of the predicted lattice energy minima within 20 kJ mol^{-1} have packing coefficients with KPI=65-70 or KPI > 70.

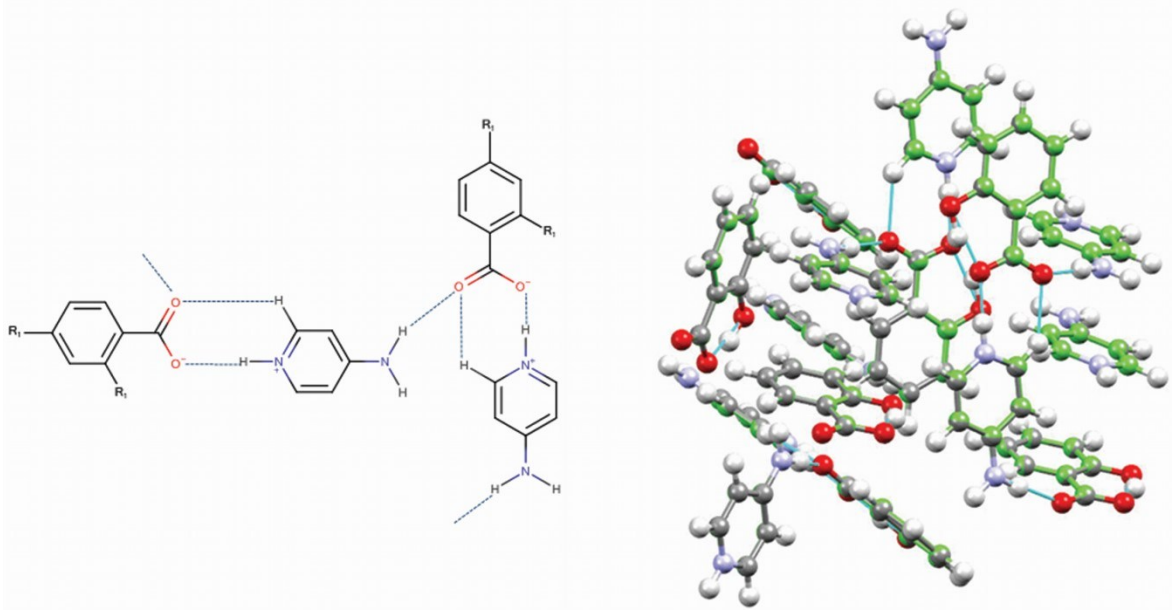


Figure 9: Hydrogen bond motif found in the experimental (left) structure of **III** (**DUSYOG**) and an overlay (right) of the predicted rank 3 structure from the CFL of **III** with the experimental structure (**DUSYOG**) of **III** whose lattice energy was minimised using the same energy model used in the CSP (denoted ExpMinOpt). The calculated ExpMinOpt is coloured by elements and the predicted rank 3 structure is coloured green. All 15/15 molecules in the coordination spheres of the two structures matched with a powder X-ray diffraction similarity index of 1.

1.4. Conclusions

The cocrystallisation of an acid and base can lead to a variety of multicomponent solid forms such as a salt, cocrystal or salt hydrate. A number of qualitative descriptors have been proposed to account for the factors

that affect the synthesis of cocrystal solid forms. Some of the proposed molecular descriptors/properties affecting the crystallisation of cocrystals include the complementarity of hydrogen bond synthons, the relative acidity (ΔpK_a) and the solubility regime of the acid-base pair on the ternary phase diagram. Whilst there are many reported examples of molecular ionic cocrystals (ICCs) in the literature, to date there has been no systematic investigation as to the factors that control the crystallisation of molecular ICCs, let alone the proposal of empirical or theoretical rules for predicting the crystallisation of such solid forms. The empirical evidence from the crystal structures of a small set of 50 molecular ICCs reported in the CSD as well as a limited grid of salt and molecular ICC solid forms resulting from the cocrystallisation of pyridines with a range of substituted monoprotic acids, was used to probe the underlying molecular properties of acid-base pairs that may be important drivers for ICC crystallisation. Acid-base pairs known to form ICCs have statistically significant differences in the polar surface areas available for intermolecular hydrogen bonding. Solubility differences between acid-base pairs known to form ICCs were also more pronounced than those known to form only salts. Solid-state DFT calculations show that comparisons of the relative lattice energies of 1:1:1 ICCs with the lattice energies of the known 1:1 salt and acid crystal structures reveal a thermodynamic preference for ICC crystallisation ranging between 2.47-8.02 kJ mol^{-1} for 4 of out of the 7 experimental ICC systems surveyed. However, the energy differences were less than 1 kJ mol^{-1} for the remaining 3 systems. This illustrates the challenges of relying on just the relative lattice energies when attempting to draw definitive conclusions about the crystallisation of multicomponent solid forms, which display many more finely balanced intermolecular interactions than single component crystal forms. Finally, the hypothesis that ICCs are favoured when the 1:1 salt ions do not have favourable modes of packing with themselves was disproven by the results of computed crystal form landscapes, which show that the predicted polymorphs of 1:1 salts known to form ICCs are just as likely to have a range of poor, reasonable and good packing coefficients when compared to anhydrous 1:1 salts that do not have reported ICCs. The description of ICCs as "acid solvates" is discouraged due to the observation of discrete intermolecular O-H$_{acid}$··O$_{anion}$ hydrogen bonds between the carboxylic acid OH donor and the carboxylate oxygen acceptor of the anion in a significant number of the ICC crystal structures surveyed, suggesting that the additional acid molecule is integral to the ICC structure rather than just occupying empty space within the crystal. Future work will address some of the above unanswered questions from this study. These include the role of solution concentrations of acid-base pairs in affecting the stoichiometry of molecular ICCs. We will also conduct a wider study looking at the functional groups responsible for directing the finely balanced set of halogen and hydrogen bonds that are evident in the crystal structures of molecular ICCs. This will pave the way for the discovery of useful supramolecular design motifs that may be used by crystal engineers to facilitate the systematic crystallisation of molecular ICCs with useful solid-state properties.

Conflicts of Interest

The authors confirm that there are no conflicts of interest to declare.

Acknowledgement

We are grateful for the support of the research computing department at Khalifa University and we acknowledge the use of the high-performance computing clusters at the Masdar Institute and at the Nuclear Engineering Department of Khalifa University. We acknowledge the support of the National Science and Engineering Council (NSERC) Discovery Grant (RGPIN-2017-06467) and E. W. R. Steacie Memorial Fellowship (SMFSU- 507347-17).

Bibliography

1. R. K. Khankari and D. J. W. Grant, *Thermochimica Acta*, 1995, **248**, 61-79.
2. L. B. Ta, H. M. Hobgood, A. Rohatgi and R. N. Thomas, *Journal of Applied Physics*, 1982, **53**, 5771-5775.
3. J. F. Remenar, S. L. Morissette, M. L. Peterson, B. Moulton, J. M. MacPhee, H. R. Guzmán and Ö. Almarsson, *Journal of the American Chemical Society*, 2003, **125**, 8456-8457.
4. A. V. Trask, *Molecular pharmaceutics*, 2007, **4**, 301-309.
5. W. Curatolo, *Pharmaceutical Science & Technology Today*, 1998, **1**, 387-393.
6. L.-F. Huang and W.-Q. Tong, *Advanced Drug Delivery Reviews*, 2004, **56**, 321-334.

7. R. J. Bastin, M. J. Bowker and B. J. Slater, *Organic Process Research & Development*, 2000, **4**, 427-435.
8. N. Issa, P. G. Karamertzanis, G. W. A. Welch and S. L. Price, *Crystal Growth & Design*, 2009, **9**, 442-453.
9. C. C. Seaton, K. Chadwick, G. Sadiq, K. Guo and R. J. Davey, *Crystal Growth & Design*, 2009, **10**, 726-733.
10. S. Hiendrawan, B. Veriansyah, E. Widjojokusumo, S. N. Soewandhi, S. Wikarsa and R. R. Tjandrawinata, *International journal of pharmaceutics*, 2016, **497**, 106-113.
11. J. Li, L. Wang, X. Fu, Z. Deng, Y. Q. Ye, Q. Ren and H. Zhang, *Eur J Pharm Sci*, 2016, **85**, 47-52.
12. N. Shan, M. L. Perry, D. R. Weyna and M. J. Zaworotko, *Expert Opin. Drug Metab. Toxicol.*, 2014, **10**, 1255-1271.
13. G. R. Desiraju and T. Steiner, *The weak hydrogen bond : in structural chemistry and biology*, Oxford University Press, Oxford, 2001.
14. C. B. Aakeroy and D. J. Salmon, *CrystEngComm*, 2005, **7**, 439-448.
15. T. R. Shattock, K. K. Arora, P. Vishweshwar and M. J. Zaworotko, *Crystal Growth & Design*, 2008, **8**, 4533-4545.
16. C. R. Taylor and G. M. Day, *Crystal Growth & Design*, 2017, DOI: 10.1021/acs.cgd.7b01375.
17. H. S. Chan, J. Kendrick, M. A. Neumann and F. J. Leusen, *CrystEngComm*, 2013, **15**, 3799-3807.
18. S. L. Price, D. E. Braun and S. M. Reutzel-Edens, *Chemical Communications*, 2016, **52**, 7065-7077.
19. S. L. Price and S. M. Reutzel-Edens, *Drug Discovery Today*, 2016, **21**, 912-923.
20. M. A. Neumann, J. van de Streek, F. P. A. Fabbiani, P. Hidber and O. Grassmann, *Nat Commun*, 2015, **6**.
21. A. T. Hulme, S. L. Price and D. A. Tocher, *Journal of the American Chemical Society*, 2005, **127**, 1116-1117.
22. J.-B. Arlin, L. S. Price, S. L. Price and A. J. Florence, *Chemical Communications*, 2011, **47**, 7074-7076.
23. A. G. Shtukenberg, Q. Zhu, D. J. Carter, L. Vogt, J. Hoja, E. Schneider, H. Song, B. Pokroy, I. Polishchuk, A. Tkatchenko, A. R. Oganov, A. L. Rohl, M. E. Tuckerman and B. Kahr, *Chemical Science*, 2017, **8**, 4926-4940.
24. S. Mohamed, *Acta Crystallographica Section E*, 2016, **72**, 1348-1352.
25. Y. Liu, T. Zhao, W. Ju and S. Shi, *Journal of Materiomics*, 2017, **3**, 159-177.
26. F. Musil, S. De, J. Yang, J. E. Campbell, G. M. Day and M. Ceriotti, *Chemical Science*, 2018, DOI: 10.1039/C7SC04665K.
27. R. Burbidge, M. Trotter, B. Buxton and S. Holden, *Computers & Chemistry*, 2001, **26**, 5-14.
28. E. L. Willighagen, R. Wehrens, W. Melssen, R. de Gelder and L. M. C. Buydens, *Crystal Growth & Design*, 2007, **7**, 1738-1745.
29. A. Pulido, L. Chen, T. Kaczorowski, D. Holden, M. A. Little, S. Y. Chong, B. J. Slater, D. P. McMahon, B. Bonillo, C. J. Stackhouse, A. Stephenson, C. M. Kane, R. Clowes, T. Hasell, A. I. Cooper and G. M. Day, *Nature*, 2017, **543**, 657.
30. G. M. Day and A. I. Cooper, *Advanced Materials*, DOI: 10.1002/adma.201704944, 1704944-n/a.
31. A. M. Reilly, R. I. Cooper, C. S. Adjiman, S. Bhattacharya, A. D. Boese, J. G. Brandenburg, P. J. Bygrave, R. Bylsma, J. E. Campbell, R. Car, D. H. Case, R. Chadha, J. C. Cole, K. Cosburn, H. M. Cuppen, F. Curtis, G. M. Day, R. A. DiStasio Jr, A. Dzyabchenko, B. P. van Eijck, D. M. Elking, J. A. van den Ende, J. C. Facelli, M. B. Ferraro, L. Fusti-Molnar, C.-A. Gatsiou, T. S. Gee, R. de Gelder, L. M. Ghiringhelli, H. Goto, S. Grimme, R. Guo, D. W. M. Hofmann, J. Hoja, R. K. Hylton, L. Iuzzolino, W. Jankiewicz, D. T. de Jong, J. Kendrick, N. J. J. de Klerk, H.-Y. Ko, L. N. Kuleshova, X. Li, S. Lohani, F. J. J. Leusen, A. M. Lund, J. Lv, Y. Ma, N. Marom, A. E. Masunov, P. McCabe, D. P. McMahon, H. Meekes, M. P. Metz, A. J. Misquitta, S. Mohamed, B. Monserrat, R. J. Needs, M. A. Neumann, J. Nyman, S. Obata, H. Oberhofer, A. R. Oganov, A. M. Orendt, G. I. Pagola, C. C. Pantelides, C. J. Pickard, R. Podeszwa, L. S. Price, S. L. Price, A. Pulido, M. G. Read, K. Reuter, E. Schneider, C. Schober, G. P. Shields, P. Singh, I. J. Sugden, K. Szalewicz, C. R. Taylor, A. Tkatchenko, M. E. Tuckerman, F. Vacarro, M. Vasileiadis, A. Vazquez-Mayagoitia, L. Vogt, Y. Wang, R. E. Watson, G. A. de Wijs, J. Yang, Q. Zhu and C. R. Groom, *Acta Crystallographica Section B*, 2016, **72**, 439-459.

32. D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, *Acta Crystallographica Section B*, 2011, **67**, 535-551.
33. G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, *Acta Crystallographica Section B*, 2009, **65**, 107-125.
34. G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, *Acta Crystallographica Section B*, 2005, **61**, 511-527.
35. W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallographica Section B*, 2002, **58**, 647-661.
36. J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, *Acta Crystallographica Section B*, 2000, **56**, 697-714.
37. S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson and M. C. Payne, *Z. Kristall.*, 2005, **220**, 567-570.
38. M. A. Neumann and M.-A. Perrin, *The Journal of Physical Chemistry B*, 2005, **109**, 15531-15541.
39. M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias and J. D. Joannopoulos, *Reviews of Modern Physics*, 1992, **64**, 1045-1097.
40. M. A. Neumann, F. J. Leusen and J. Kendrick, *Angewandte Chemie International Edition*, 2008, **47**, 2427-2430.
41. H. C. S. Chan, J. Kendrick and F. J. J. Leusen, *Angewandte Chemie International Edition*, 2011, **50**, 2979-2981.
42. A. Asmadi, M. A. Neumann, J. Kendrick, P. Girard, M.-A. Perrin and F. J. J. Leusen, *The Journal of Physical Chemistry B*, 2009, **113**, 16303-16313.
43. M. H. D. B. Alsirawan, V. R. Vangala, J. Kendrick, F. J. J. Leusen and A. Paradkar, *Crystal Growth & Design*, 2016, **16**, 3072-3075.
44. D. E. Braun, P. G. Karamertzanis and S. L. Price, *Chemical Communications*, 2011, **47**, 5443-5445.
45. S. Mohamed, D. P. Karothu and P. Naumov, *Acta Crystallographica Section B*, 2016, **72**, 551-561.
46. A. J. Smith, S.-H. Kim, N. K. Duggirala, J. Jin, L. Wojtas, J. Ehrhart, B. Giunta, J. Tan, M. J. Zaworotko and R. D. Shytle, *Molecular Pharmaceutics*, 2013, **10**, 4728-4738.
47. A. R. Buist and A. R. Kennedy, *Crystal Growth & Design*, 2014, **14**, 6508-6513.
48. N. K. Duggirala, A. J. Smith, Ł. Wojtas, R. D. Shytle and M. J. Zaworotko, *Crystal Growth & Design*, 2014, **14**, 6135-6142.
49. F. v. Kobell, *Journal für Praktische Chemie*, 1843, **28**, 489-491.
50. S. L. Childs, L. J. Chyall, J. T. Dunlap, V. N. Smolenskaya, B. C. Stahly and G. P. Stahly, *Journal of the American Chemical Society*, 2004, **126**, 13335-13342.
51. C. B. Aakeröy, M. E. Fasulo and J. Desper, *Molecular Pharmaceutics*, 2007, **4**, 317-322.
52. D. P. Karothu, I. Jahovic, G. Jovanovski, B. Kaitner and P. Naumov, *CrystEngComm*, 2017, **19**, 4338-4344.
53. K. Honer, E. Kalfaoglu, C. Pico, J. McCann and J. Baltrusaitis, *ACS Sustainable Chemistry & Engineering*, 2017, **5**, 8546-8550.
54. M. Rafilovich and J. Bernstein, *Journal of the American Chemical Society*, 2006, **128**, 12185-12191.

55. K. Shankland, C. K. Leech, S. Mohamed, S. A. Barnett and D. A. Tocher, *Acta Crystallographica Section E*, 2007, **63**, o3574.
56. D.-K. Bučar, R. W. Lancaster and J. Bernstein, *Angewandte Chemie International Edition*, 2015, **54**, 6972-6993.
57. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallographica Section B*, 2016, **72**, 171-179.
58. B. Lou, S. R. Perumalla and C. C. Sun, *Crystal Growth & Design*, 2015, **15**, 24-28.
59. A. Sikorski and D. Trzybiński, *Journal of Molecular Structure*, 2013, **1049**, 90-98.
60. B. Sarma, N. K. Nath, B. R. Bhogala and A. Nangia, *Crystal Growth & Design*, 2009, **9**, 1546-1557.
61. H.-K. Fun, M. Hemamalini and V. Rajakannan, *Acta Crystallographica Section E*, 2010, **66**, o2010-o2011.
62. B. Lou, S. R. Perumalla and C. C. Sun, *Journal of Molecular Structure*, 2015, **1099**, 516-522.
63. A. Mukherjee and G. R. Desiraju, *Crystal Growth & Design*, 2014, **14**, 1375-1385.
64. Calculator Plugins were used for structure property prediction and calculation, *MarvinSketch*, Version 16.3.7.0, ChemAxon (http://www.chemaxon.com), 2016.
65. A. J. Cruz-Cabeza, *CrystEngComm*, 2012, **14**, 6362-6365.
66. S. Mohamed, D. A. Tocher, M. Vickers, P. G. Karamertzanis and S. L. Price, *Crystal Growth & Design*, 2009, **9**, 2881-2889.
67. S. L. Childs, G. P. Stahly and A. Park, *Molecular Pharmaceutics*, 2007, **4**, 323-338.
68. Dassault Systèmes BIOVIA, *BIOVIA Materials Studio*, Version 8.0.100.21, San Diego: Dassault Systèmes, 2014.
69. H. R. Karfunkel and R. J. Gdanitz, *Journal of Computational Chemistry*, 1992, **13**, 1171-1183.
70. R. L. Akkermans, N. A. Spenley and S. H. Robertson, *Molecular Simulation*, 2013, **39**, 1153-1164.
71. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09* Revision D.01, Wallingford, CT, USA, 2009.
72. S. L. Mayo, B. D. Olafson and W. A. Goddard, *Journal of Physical Chemistry*, 1990, **94**, 8897-8909.
73. C. M. Breneman and K. B. Wiberg, *Journal of Computational Chemistry*, 1990, **11**, 361-373.
74. S. L. Price, M. Leslie, G. W. A. Welch, M. Habgood, L. S. Price, P. G. Karamertzanis and G. M. Day, *Physical Chemistry Chemical Physics*, 2010, **12**, 8478-8490.
75. A. J. Stone, *Chemical Physics Letters*, 1981, **83**, 233-239.
76. A. J. Stone and M. Alderton, *Molecular Physics*, 1985, **56**, 1047-1064.
77. A. J. Stone, *Journal of Chemical Theory and Computation*, 2005, **1**, 1128-1132.
78. L.-Y. Hsu and D. E. Williams, *Acta Crystallographica Section A*, 1980, **36**, 277-281.
79. S. R. Cox, L.-Y. Hsu and D. E. Williams, *Acta Crystallographica Section A*, 1981, **37**, 293-301.
80. D. E. Williams and S. R. Cox, *Acta Crystallographica Section B*, 1984, **40**, 404-417.
81. D. S. Coombes, S. L. Price, D. J. Willock and M. Leslie, *The Journal of Physical Chemistry*, 1996, **100**, 7352-7360.
82. T. Beyer and S. L. Price, *The Journal of Physical Chemistry B*, 2000, **104**, 2647-2655.
83. C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. v. Streek and P. A. Wood, *Journal of Applied Crystallography*, 2008, **41**, 466-470.
84. J. A. Chisholm and S. Motherwell, *Journal of Applied Crystallography*, 2005, **38**, 228-231.
85. A. Kitaigorodsky, *Molecular crystals and molecules*, Elsevier, 2012.
86. A. Spek, *Acta Crystallographica Section D*, 2009, **65**, 148-155.
87. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.

88. D. Vanderbilt, *Physical Review B*, 1990, **41**, 7892-7895.
89. S. Grimme, *Journal of Computational Chemistry*, 2006, **27**, 1787-1799.
90. H. J. Monkhorst and J. D. Pack, *Physical Review B*, 1976, **13**, 5188-5192.
91. B. G. Pfrommer, M. Côté, S. G. Louie and M. L. Cohen, *Journal of Computational Physics*, 1997, **131**, 233-240.
92. J. van de Streek and M. A. Neumann, *Acta Crystallographica Section B*, 2010, **66**, 544-558.
93. B. P. van Eijck and J. Kroon, *Acta Crystallographica Section B*, 2000, **56**, 535-542.
94. Wavefunction Inc. , *Spartan'16*, Version 2.0.7, 18401 Von Karman Ave., Suite 370, Irvine CA 92612, 2017.
95. W. L. Koltun, *Biopolymers*, 1965, **3**, 665-679.
96. M. C. Etter, J. C. MacDonald and J. Bernstein, *Acta Crystallographica Section B*, 1990, **46**, 256-262.
97. L. B. Jerzykiewicz, T. Lis, Z. Malarski and E. Grech, *Journal of Crystallographic and Spectroscopic Research*, 1993, **23**, 805-812.
98. J. R. Jebamony and P. Thomas Muthiah, *Acta Crystallographica Section C*, 1998, **54**, 539-540.
99. L. B. Jerzykiewicz, Z. Malarski, L. Sobczyk, T. Lis and E. Grech, *Journal of Molecular Structure*, 1998, **440**, 175-185.
100. M. Hemamalini and H.-K. Fun, *Acta Crystallographica Section E*, 2010, **66**, o479-o480.
101. K. Balasubramani and H.-K. Fun, *Acta Crystallographica Section E*, 2009, **65**, o1519.
102. A. Jayasankar, L. S. Reddy, S. J. Bethune and N. Rodríguez-Hornedo, *Crystal Growth & Design*, 2009, **9**, 889-897.
103. D. E. Braun, P. G. Karamertzanis, J.-B. Arlin, A. J. Florence, V. Kahlenberg, D. A. Tocher, U. J. Griesser and S. L. Price, *Crystal Growth & Design*, 2011, **11**, 210-220.
104. R. A. Chiarella, R. J. Davey and M. L. Peterson, *Crystal Growth & Design*, 2007, **7**, 1223-1226.
105. G. Smith, U. D. Wermuth, P. C. Healy and J. M. White, *Australian Journal of Chemistry*, 2003, **56**, 707-713.
106. T. Wu, X.-P. Zhang and X.-Z. You, *RSC Advances*, 2013, **3**, 26047-26051.
107. D. Arman Hadi and R. T. Tiekink Edward, *Journal*, 2013, **228**, 289.
108. M. Hemamalini and H.-K. Fun, *Acta Crystallographica Section E*, 2010, **66**, o2535.
109. D. E. Braun, T. Gelbrich, V. Kahlenberg and U. J. Griesser, *Molecular Pharmaceutics*, 2014, **11**, 3145-3163.
110. A. Tkatchenko and M. Scheffler, *Phys. Rev. Lett.*, 2009, **102**, 073005.
111. M. Arhangelskis, A. D. Katsenis, A. J. Morris and T. Friscic, *Chemical Science*, 2018, DOI: 10.1039/C7SC05020H.
112. C. P. Brock and J. D. Dunitz, *Chemistry of Materials*, 1994, **6**, 1118-1127.