Conformational Change in Molecular Crystals: Impact of Solvate Formation and Importance of Conformational Free Energies

DOI:
10.1021/acs.cgd.1c00833

Document Version
Accepted author manuscript

Conformational Change in Molecular Crystals: Impact of solvate formation and importance of conformational free energies.

Sarah E. Wright[a], Jason C. Cole[b] and Aurora J. Cruz-Cabeza[a,c*]

a. Department of Chemical Engineering and Analytical Science, University of Manchester, Manchester, UK.
b. The Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, UK
c. Chemical Development, Pharmaceutical Technology & Development, AstraZeneca, Macclesfield Campus, SK10 2NA

*Email: aurora.cruzcabeza@manchester.ac.uk

Abstract

Changes in molecular conformation and its relationship with crystal polymorphism have been well documented in previous work. To the best of our knowledge, however, the effect of solvate formation on molecular conformation has never been explored. Using the Cambridge Structural Database (CSD) in combination with molecular modelling, we provide insights on the impact of solvate formation on the conformation adopted by a compound and whether such conformation is different to those found in its neat form(s). Typically, the more flexible a molecule is, the higher the chances that conformational change occurs upon solvate formation. There are no trends as to the relative stability of conformers and their likelihood to be observed in either the solvates or the neat forms. Typically, conformer energy differences in solvate-neat form pairs are small (< 4 kJ/mol) and when larger energy differences are observed (> 15 kJ/mol), these can be reduced significantly when both solvent as well as thermal effects are considered in the simulations. This highlights of the importance of computing thermal contributions in conformer energies as well as accounting for environmental effects. Overall, we find that conformational change in solvate-neat form pairs mirrors the behaviour of conformational change in polymorphism.

1 Introduction

A solvate is a multicomponent crystal containing at least one solvent molecule in its structure.[1] Solvate formation impacts thermodynamic and kinetic stability as well dissolution rates,

bioavailability, crystal morphology, vapour pressure and tabletability of materials.[2,3] Hydrates are the most common type of solvates given the ubiquitous nature of water.[4,5] They account for half of all solvate entries in the Cambridge Structural Database (CSD).[6] Three quarters of pharmaceutical molecules, and a third of organic molecules,[7] are thought capable of forming hydrates. Drug molecules can be exposed to many sources of water during development and manufacture, as a result hydrate formation can be problematic for the pharmaceutical industry.[8,9]

Although it is difficult to predict whether a compound will form a solvate or not from its molecular structure alone, some trends have been observed on how molecular structure, hydrogen bonding ability and crystal packing impact solvate formation.[10] For example, Infantes *et al.*[11,12] found that an increased polar surface of the molecule is correlated to increasing hydrate formation but that neither the hydrogen bond donor/acceptor ratio or molecular weight significantly had an impact. The donor/acceptor ratio of the molecule, however, was found to influence the type of water environment[5] (i.e., fewer donors favour the water donating protons, and vice versa). Similar studies focussing on solvates such as methanol and DMSO also observe that the hydrogen bond donor/acceptor ratio has little impact on solvate formation.[13,14] In terms of the thermodynamics of the process, various works agree in that solvates and hydrates form because they have lower free energies than their corresponding neat forms. Entropy and disorder can play an important role in solvate formation, especially for solvents with low melting points.[15,16]

Crystal Structure Prediction (CSP) can be used to predict solvate formation, though this may require significant computational resources. In the fifth CSP blind test, a monohydrate with multiple polymorphs was included as a target system for the first time.[17] Since it is unknown, a priori, which stoichiometry a solvate or hydrate will crystallise in, stoichiometry is a further dimension to consider in CSP studies of multicomponent crystals.[9,18,19] Sampling of the crystal structure landscapes computationally can thus become extremely challenging for solvates due to the various symmetry independent molecules present and their varying stoichiometry, beyond the usual CSP challenges of molecular flexibility.[20–22] Solvates largely remain, thus, a challenging group of solid forms to predict computationally.

Molecular conformation can influence physical properties of crystal structures. Significant conformational change, for example, is known to lead to greater differences in form stabilities and other properties (such as colour) in conformational polymorphs.[23] Whilst there is a number

of works studying conformational changes in polymorphs, to our knowledge, how the formation of solvates impacts molecular conformation has never been studied. In previous work by the authors, it was found that solvates are slightly more likely to have unusual conformations than neat forms.[24]

There are several works in the literature which discuss the impact of the crystallisation solvent on molecular conformation.[25–29] For example, the cancer drug Axitinib is known to form five anhydrous polymorphs and 66 solvates with a wealth of conformational diversity across these forms. Many of the solvates share a similar conformation to form-IV, the typical anhydrous form obtained upon desolvation with vacuum drying. Slurry experiments at high temperatures resulted in an intermediate solvate that led to the crystallisation of a most stable form-XLI, a form containing a high energy conformer[26] Another example of conformational changes involved in solvation/desolvation is Fluconazole. Multiple polymorphic forms containing difference conformers[30–35] as well as various solvates[36] and a monohydrate[37] are known for Fluconazole. Two recent studies by Basford *et al.*[38,39] investigate the hydration and dehydration mechanisms of Fluconazole revealing that a minor conformational change around a hydroxyl group initiates the dehydration of the monohydrate form. As with Axitinib, the polymorph obtained upon the dehydration of fluconazole monohydrate has a strong conformational similarity with the monohydrate.

In this study, we adopt a methodology by which we compare the conformations of molecules in solvates with those found in their corresponding neat forms. For simplicity, we would refer to solvate-neat form pairs as “SoNe” pairs. The main objective of the work is to answer several fundamental questions regarding the impact of solvate formation on crystal conformation including: a) how often is conformational change observed in SoNe pairs? b) are stable conformers more likely to crystallise as solvates than as neat forms? or c) what are typical energy differences in conformational solvates neat form pairs? An in-depth CSD analysis is presented together with computer simulations and the overall results are compared to previous data on conformational polymorphs.[23]

2 Methods

2.1 CSD searches

The Cambridge Structural Database[6] (CSD) version 5.41 (November 2019 release) was searched for two component crystal structures containing only elements commonly seen in organic molecules (H, D, C, N, O, S, P and the halides) and within the best R-factor subset (to not include redeterminations). Only crystal structures with 3D coordinates, no disorder, errors or ions and not polymeric were retrieved. One of the components had to be included in the CCDC solvent library -which consists of the 74 most common solvents in the CSD (the solvent component)- whilst the second one could be of any type (the host component). Structures containing two solvent molecules were removed. Our searches returned 20 958 crystal structures of solvates. Since searches were restricted to two component systems only, our data does not include solvates of cocrystals, salts or mixed solvates.

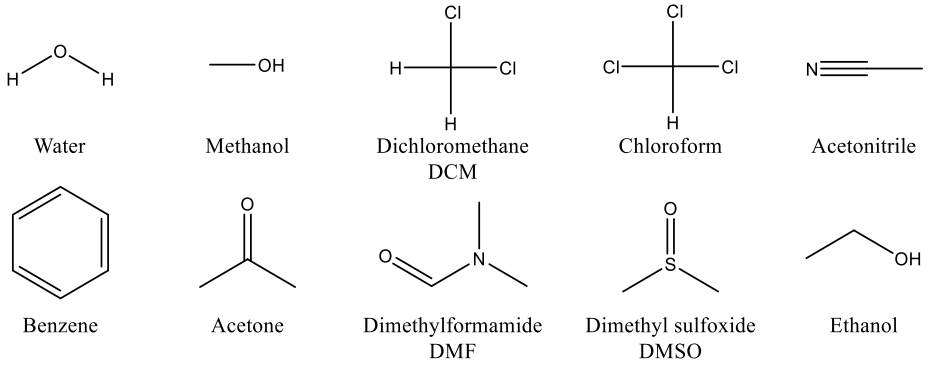


Figure 1: Structure of solvents investigated in the analysis of Solvates in this study.

Solvates containing the top ten solvents (accounting for 84% of the solvate structures) were then analysed further. The ten most common solvents in our solvates dataset were water, methanol, dichloromethane (DCM), chloroform, acetonitrile, benzene, acetone, dimethylformamide (DMF), dimethyl sulfoxide (DMSO) and ethanol (Figure 1). Using the CSD Python API,[6] the host component in each of the solvate crystal structure was then searched for its corresponding neat crystal forms available in the CSD. For these searches, InChI keys were used in order to account for the correct stereochemistry of the host compound. Redeterminations of the same crystal structures were removed. This analysis identified 1518 crystal structures of solvates having at least one neat crystal form of the host compound (Figure

2). For simplicity, we will refer to these solvate-neat form pairs to as “SoNe pairs”. If multiple forms are available for the same two components, a pair will be generated for each unique solvate structure (thus different polymorphs or different stoichiometries of the same two components were accounted for separately). If multiple neat forms were existing, those would be considered separately as the analysis was done per unique Solvate structure. Since only a small proportion of solvates and neat forms would show polymorphism, the overall data statistics would remain very similar if the approach was to cluster SoNe pairs of neat forms as a unique pair.

Once a SoNe pair was identified, the host compound geometry as found in the solvate structure and in the neat form were compared and overlayed using the CSD Python API. An RMSD of atomic positions was calculated for each of those overlays. If the RMSD was greater than 0.375 Å, conformational change was identified within the SoNe pair and the pair was referred to as a Conformational SoNe pair. This RMSD cut-off value was chosen from recommendations of previous work on Conformational Change.[23]

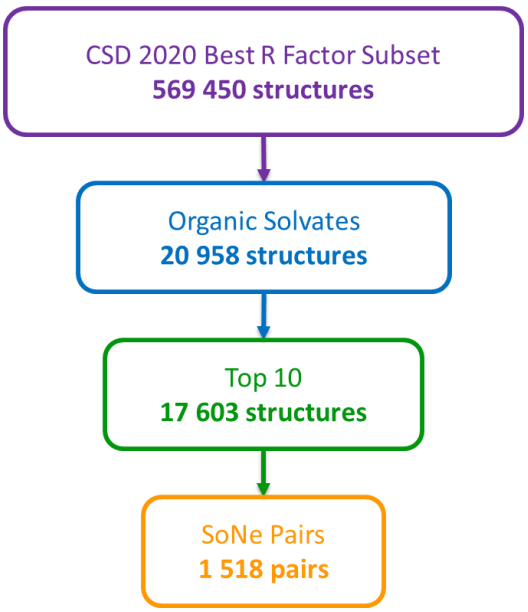


Figure 2: Methodology for retrieving SoNe pairs from the CSD.

2.2 Calculation of DOFlex

We calculate the number of degrees of flexibility (DOFlex) for the host compounds using the CSD Python API. This was calculated as the sum of the number of acyclic R-bonds (not accounting for terminal methyl groups) and the number of aliphatic cyclic rings (with six or

more bonds) minus the number of triple bonds. This DOFlex property for the host compound was calculated similarly to the previous work on Conformational Polymorphs so that data could be compared.[23]

2.3 Conformer energies in Conformational SoNe pairs

Relative conformer energies of the SoNe pairs were computed with DFT. In order to inform a choice of DFT method, conformer energies of the SoNe pairs were computed using several DFT methods and basis sets for two compounds: N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine and o-acetamidobenzamide. Because of its balance in terms of computational time and accuracy, the M06/6-31+G(d,p) level of theory was then taken for the calculation of all other systems in this study. All calculations were performed using Gaussian09.[40]

Conformer geometries were retrieved from the solvate as well as the neat form crystal structures. Only pairs differing by conformational change (as per the RMSD test, Conformational SoNe pairs) were used in these computations. For each calculation, the molecular geometry of the host was optimised without constraints in the gas-phase followed by a single point energy calculation in various SMD[41] solvation models (a number of solvents were explored). The energy difference between the conformer of the host compound as found in the neat form and the solvate form were calculated as in Equation 1.

$$\Delta E_{Change}^{Neat \rightarrow Solv} = E_{Conformer}^{Solvate} - E_{Conformer}^{Neat} \qquad (1)$$

After geometry optimisation, if the SoNe pair conformer energies were small (< 2.5 kJ mol^{-1}), the optimised molecular geometries were overlayed again using Mercury to ensure that they had not optimised to the same conformer. If the RMSD of the overlayed pairs was 0.00 Å (to 2 decimal places), the conformations had converged to the same conformer geometry and thus were reclassified as non-conformational SoNe pairs and were removed from the analysis.

For select examples, geometry optimisations and frequency calculations at 25^{o}C were further computed, free of constraints, in various SMD solvation models (eight in total, see ESI) for all different conformers found in experimental crystal structures of neat forms as well as solvates and cocrystals. The frequency analysis provided us with the thermal free energy correction (G_{corr}[T]) which can be added to the electronic energy (E_e) to provide the Gibbs free energy

(G), as shown by Equation 2.[42] This analysis was then repeated at 60°C, however, little change was noted thus only the energetics at 25°C are presented in the results section.

$$G[T] = E_e + G_{corr}[T] \quad (2)$$

The energies (E) and free energies (G) of these specific systems were then calculated relative to the most stable experimental conformer (Equation 3).

$$\Delta E_{change} = E_{conformer} - E_{stable\ conformer} \quad (3)$$

3 Results

3.1 CSD Statistics

3.1.1 Overall Solvates Statistics

The CSD was searched for two component solvates containing only one host component (HC) and one solvent component. This resulted in a total of 20 958 solvate crystal structures. With regards to the number of solvates and their nature, we note that the CSD has changed significantly in the last two decades.[43] Our data is presented together with the seminal solvate analysis by Görbitz and Hersleth[43] in Table 1 for the 10 most common solvates. Our dataset sees toluene and tetrahydrofuran (THF) solvates move out of the top 10 list whilst DMF and DMSO move into it. The ten top solvate types were then analysed further. These were solvates of water, methanol, DCM, chloroform, acetonitrile, benzene, acetone, DMF, DMSO and ethanol. The molecular structure of the HC of these solvates was then taken for further searches of the CSD in order to retrieve, if any, all crystal structures of its neat forms. SoNe pairs were then identified (Table 1).

On average, only 9% of solvate crystal structures have a neat form counterpart for the HC. The breakdown across the different solvate types can be seen in Table 1, where DMSO and DCM are the most and the least likely solvates to have SoNe pairs (17% and 4% respectively). We performed a χ^2 test on the proportion of solvates with a SoNe pair (Table 1), which suggests that there is a statistically significant relationship between solvate type and the existence of a neat form ($\chi^2 = 173$, $p < 0.001$), therefore the two variables, solvate type and the existence of a neat form, are dependent on each other. Whilst hydrates make up the largest portion of our

data, its omission has little effect on the significance of the relationship. We divided our data into solvent groups: polar protic (water, methanol and ethanol), polar aprotic (acetonitrile, DMF, DMSO and acetone) and chlorinated (chloroform and DCM) to investigate whether individual solvent types in a given class influences the proportion of SoNe pairs. χ^2 tests within the polar protic group ($\chi^2 = 1.84$, $p > 0.3$) and the chlorinated group ($\chi^2 = 3.4$, $p > 0.05$) indicate that the individual solvents within the respective groups are not associated with the proportion of SoNe pairs. However, χ^2 tests within the polar aprotic group ($\chi^2 = 15.3$, $p < 0.01$) indicate that the individual solvents within the group are related to the proportion of SoNe pairs.

Why are so many of these HCs observed as solvates but not as neat forms? In answer to this question, a study by Görbitz and Hersleth and two more recent studies by Werner and Swift are relevant. Görbitz and Hersleth[43] observed that the increasing complexity of molecules being synthesised relates to an increased proportion of solvates; since these more complex molecules, they are more likely to include cavities and channels within their packing arrangement thus including solvent molecules in their structure. The authors also suggested that the use of solvent mixtures during crystallisations may prove useful for obtaining crystals of molecules with a high molecular weight. A more recent study by Werner and Swift[44] found that the solvents of crystallisation used have become less diverse, but the use of solvent mixtures has led to a disproportionate increase in reports of heterosolvates. The authors also noted an increased use of DMSO and DMF as solvents, and a significant decrease in the use of Benzene, agreeing with the trends observed in Table 1.

Werner and Swift made further observations on the stoichiometric trends and crystal systems of organic hydrates and solvates.[44,45] Organic solvates adopt a much narrower range of solvent:HC stoichiometries compared to organic hydrates. Hydrates with and without neat forms show almost identical trends in their water stoichiometries, with smaller stoichiometries being preferred overall. Within the hydrate-neat pairs subset it was observed that hydrates would crystallise in lattices with lower symmetry relative to their neat forms. Comparison of solvates and neat forms indicated that solvates tended to favour triclinic lattices (in particular heterosolvates) and also noted an increasing fraction of trigonal lattices for solvates which had a known neat form, however they note that due to the small number of pairs that assigning significance to this would be premature. Variations of space-group preferences of solvates have been also reviewed in other works.[46]

Another recent study by Cole *et al.*,[47] evaluated the prior likelihood (PL) of a solvent to form a solvate, which has implications for predicting solvate formation and also as a guide for avoiding solvation during crystallisation. Some key takeaways are that aromatic solvents with electron-withdrawing substituents have high PLs. Alcohols tend to have lower average PLs and aliphatic hydrocarbons, which form very weak interactions, have very low PLs. The authors also investigate the space-group preferences of solvates. Occupancy of $P\bar{1}$ is always higher for solvates, similarly for C2/c with a few exceptions. Non-solvates (structures not containing the solvent of interest) generally have a higher occupancy for $P2_1/c$, $P2_12_12_1$, $P2_1$ and *P*bca. Solvents which can be centred on 2-fold axes or inversion centres increases the chances of the space group incorporating those types of symmetries. Another relevant study by Cruz-Cabeza *et al.*[15] found some correlation between the likelihood of a solvent to make solvates with their corresponding entropy cost. Thus, solvents having a low entropy penalty for solvate formation (i.e. DMSO) forming more solvates than those with a high entropy penalty (i.e. ethyl acetate).

Table 1: The most frequent solvates in the CSD in 2000 and in 2020 (CSD 5.41) with the proportion of them (in %) having SoNe pairs.

Solvent	Görbitz and Hersleth (2000) *	Structures CSD 5.41 (2020)	% SoNe pairs (2020)	No. SoNe pairs† (2020)
Hydrate	-	8724	8%	713
Methanol	762	1918	7%	139
DCM	457	1439	4%	63
Chloroform	395	1334	6%	79
Acetonitrile	389	794	10%	83
Benzene	444	693	13%	93
Acetone	332	683	12%	82
DMF	114	703	14%	97
DMSO	129	698	17%	119
Ethanol	398	617	8%	50
Total	-	17603	9%	1518

* We note that the filtering used in the Görbitz study is different to in this study. The Görbitz study does not investigate hydrates.

† Number of solvate structures with at least one anhydrous form.

3.1.2 Conformational SoNe Pairs Statistics

We define here a conformational SoNe pair as a pair of crystal structures of a solvate and the neat form whereby the host component in the neat form and the solvate has experienced conformational change. For all solvents investigated, on average, the incidence of conformational change in SoNe pairs is 46%. The solvent with the most conformational SoNe pairs is ethanol with conformational change occurring in 64% of the SoNe pairs, and the least conformational is DMSO with only 35%. A χ^2 test was performed, on the data in Table 2, which suggested that there is a likely relationship between solvate type and conformational change in SoNe pairs ($\chi^2 = 27.5$, $p < 0.01$). Again, the omission of hydrates from our data has little effect with the χ^2 value still being significant.

In comparison, the work by Cruz-Cabeza and Bernstein[23] on polymorphs found that ~36% of polymorphic compounds would display conformational polymorphism. The statistical differences between the SoNe pairs and the polymorph observations may be related to the nature of compounds most likely found to form solvates. For example, drugs crystallising as hydrates and solvates generally have a higher number of rotatable bonds than those crystallising as pure neat forms.[48] It has been suggested that this trend may occur as a result of inefficient crystal packing becoming more likely as molecular size and flexibility increases. To mitigate for this, molecules may crystallise in frameworks with voids or channels which get further stabilised through the inclusion of solvent. These, on average, larger more flexible molecules crystallising as solvates may also be more susceptible to conformational change.

Table 2: Proportion of SoNe pairs for which conformational change is observed.

Solvate	Solvent Dielectric Constant ε	No. of SoNe Pairs	% Conformational SoNe Pairs (No.)	% Non-Conformational SoNe Pairs (No.)
Hydrate	78	713	43% (304)	57% (409)
Methanol	32	139	50% (70)	50% (69)
DCM	9	67	54% (36)	46% (31)
Chloroform	5	79	38% (30)	62% (49)
Acetonitrile	36	83	57% (47)	43% (36)
Benzene	2	93	48% (45)	52% (48)
Acetone	20	82	55% (45)	45% (37)
DMF	37	97	47% (46)	53% (51)
DMSO	47	119	34% (41)	66% (78)
Ethanol	25	50	64% (32)	36% (18)
Total All	-	1518	46% (696)	54% (826)

3.1.3 Capability for conformational change

In Figure 3 we plot the mean number of DOFlex for SoNe pairs that are conformational (blue) versus those that are non-conformational (orange) classified by solvent type. We observe that,

on average, the HC in conformational SoNe pairs is more flexible than in non-conformational SoNe pairs.

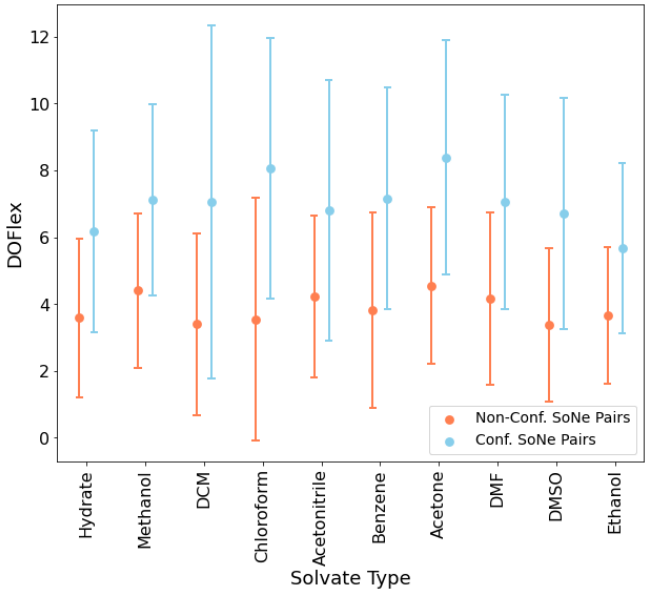


Figure 3: Average DOFlex of the HC found in conformational (blue) versus non-conformational (orange) SoNe pairs. Error bars represent the range of DOFlex in each of the subset whilst the data point is the mean value.

A more in-depth analysis of this data, through histogram plots in Figure 4, allows us to compare the effect of increasing DOFlex of the HC on the resulting incidence of conformational change in the SoNe pairs. For illustration purposes, the data for the top 5 solvate datasets is compared with that of conformational change in polymorphs – as taken from the Cruz-Cabeza and Bernstein[23] analysis. Here, as with polymorphs, we observe a gradual increase in the proportion of conformational SoNe pairs with increasing DOFlex. When trends slightly deviate, we note that the number of experimental observations become very small so the data should be interpreted with care. The proportion of conformational change observed in SoNe pairs analysed per DOFlex is extremely similar to the data on conformational change in polymorphs. This suggests that, for a compound with a given DOFlex, conformational change due to the formation of a new crystal forms is as likely to occur through polymorphism or solvate formation. The data also shows that the more flexible the compound crystallising either as polymorphs or solvates is, the more likely the compound is to crystallise in a different conformer. For example, a compound with 8 DOFlex crystallising as a SoNe pair or a polymorphic pair has >60% probability to do so with a new conformer whilst as compound with 3 DOFlex, only has ~ 30% of probability.

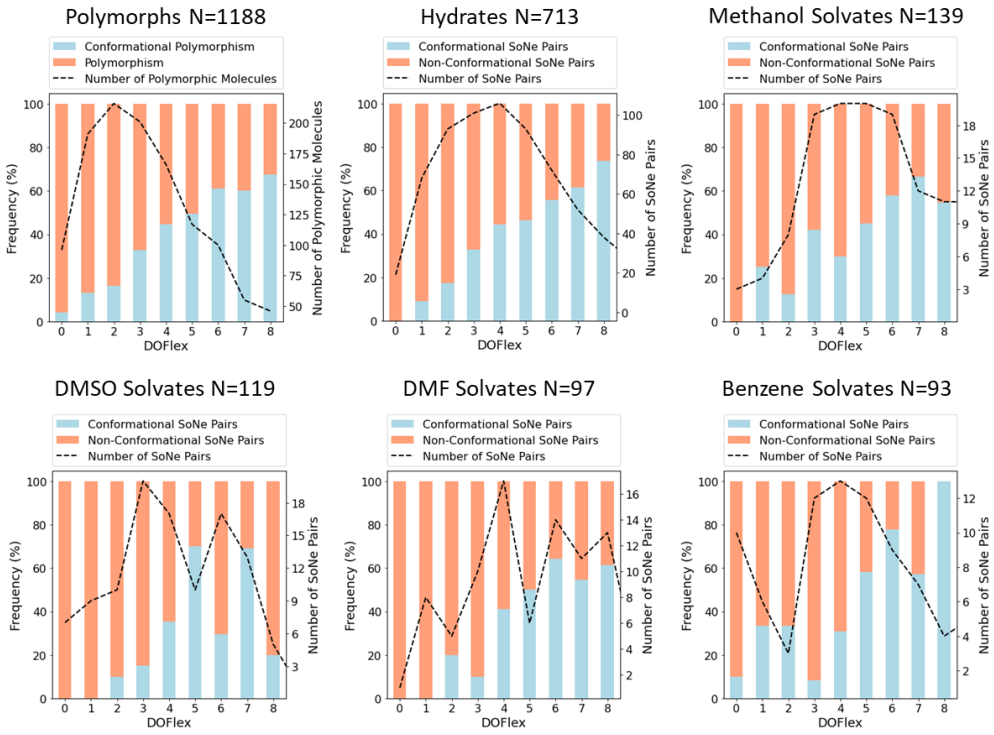


Figure 4: Incidence of conformational (blue) and non-conformational (orange) SoNe pairs as a function of the HC's degree of flexibility (DOFlex), for the five most common solvates with SoNe pairs. The data is compared with the incidences for conformational and non-conformational polymorphism [23].

3.2 Conformational energies

In this section we study several systems more in detail with computational methods. For simplicity, we refer to the experimental crystal conformers as "N", "S" or "C" to refer whether they are found in a neat form structure, solvate structure or co-crystal structure respectively. If a conformer is observed in both a neat and a solvate structure it would be referred to as "N/S". The conformers are then further numbered based on their relative gas phase stabilities.

3.2.1 Choice of computational method for conformer energies

The calculation of accurate and reliable conformational energies proves very challenging. Different methods can account for the various intramolecular interactions in the molecule differently. For example, intramolecular van der Waals or hydrogen bonding can be challenging to compute.[49] Other electronic effects such as electron correlation in delocalised systems can be also difficult to account for.[50,51] For example, recent works by Greenwell *et al.* [50,51] show that standard DFT methods can over stabilise highly conjugated systems and that

these needs for molecular energies to be computed at the MP2-D level of theory. Nyman *et al.*[52] have also shown the difficulties of computing molecular energies for the well-known polymorphic compound ROY.

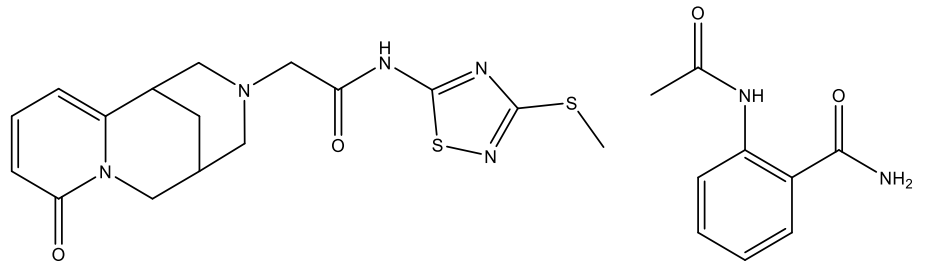


Figure 5: Molecular structures of N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine and o-acetamidobenzamide.

Here we compare the relative conformer energies obtained by several computational methods Table 3) on two systems, N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine and o-acetamidobenzamide (Figure 5, the conformers of each can be found in Figure 10 and Figure 12 respectively).

First we explore B3LYP[53–56] since it is the most widely used DFT method.[57] However, B3LYP functional is known to underperform when weak interactions are important for conformer energies (i.e. intramolecular dispersion).[58] Then, we use the M06[59] functional (a highly parameterised hybrid meta functional by the Truhlar group) since this functional is able to model weak interactions[59,60] and thus has proven to be a good method for simulating conformationally flexible molecules.[61] Third, we use the B97D[62] (a pure functional based of Becke's 97 functional[63]) with its hybrid ωB97xD[64] version since these both include Grimme's empirical dispersion corrections and are known to model well covalent systems[65] and non-covalent intramolecular interactions.[64] Fourth and last, the MP2[66–69] method, known to model well the electron correlation, was also used.

All conformers were taken from crystal structures in the CSD,[6] and first optimised with the M06/6-31+G(d,p) level of theory, followed by a further optimisation with the def2TZVpp basis set and finally a single point with the def2QZVpp basis set. The optimised conformers with the M06/def2TZVpp were then used as the starting point for the other methods.

Table 3: Comparison of different DFT methods for computing relative conformational energies of N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine and o-acetamidobenzamide. Calculations are started from the optimised structure obtained using the M06/6-31+G(d,p) method, see text for more

details. Conformers are referred to as N (originating from neat form) or S (originating from solvate form). See section 3.2.

DFT Model	N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine					o-acetamidobenzamide		
	Relative Conformer Energy (kJ/mol)				Av. CPU Time (min)	Relative Conformer Energy (kJ/mol)		Av. CPU Time (min)
Conformer	N1/S1	N2/S2	S3	S4		N1/S1	N2	
B3LYP/6-31+G(d,p)	0.00	7.60	13.30	20.99	568	0.00	42.83	115
B3LYP/def2TZVpp	0.00	6.73	10.75	17.86	2083	0.00	42.56	584
B3LYP/def2QZVpp	0.00	6.69	10.55	17.34	2609	0.00	42.48	436
M06/6-31+G(d,p)	0.00	7.23	21.61	28.66	1155	0.00	37.38	239
M06/def2TZVpp	0.00	6.35	18.23	24.04	3079	0.00	37.07	824
M06/def2QZVpp	0.00	6.05	17.98	23.90	3075	0.00	36.71	426
B97-D/def2TZVpp	0.00	4.90	22.54	29.04	3123	0.00	34.80	608
B97-D/def2QZVpp	0.00	4.73	22.07	28.28	2134	0.00	34.51	973
ωB97xD/def2TZVpp	0.00	5.12	23.81	31.56	5359	0.00	36.03	961
ωB97xD/def2QZVpp	0.00	4.93	23.41	30.88	4578	0.00	35.77	1276
MP2/def2TZVpp	-	-	-	-	-	0.00	37.84	13 521
MP2/def2QZVpp	-	-	-	-	-	0.00	37.79	2 602

For both systems investigated, all methods provided the same relative ranking of the conformers. The B3LYP, however, provided relative conformer energies which are significantly different to the other methods used. Whilst B3LYP is still very popular,[57] it is known to perform less well where weak interactions play a major role.[58,70] In 2011, Goerigk and Grimme[71] discouraged its use as a standard method without closer inspection of the results.

The geometry optimisations with MP2 were found to be very expensive, using significantly more CPU time than any of the other DFT methods and failing to converge in the optimisation of our larger molecule (see Table 3). The MP2 method, however, could have been used with a single point calculation on geometries optimised with other DFT models to reduce on computational time. Despite M06 methods providing the initial optimisation for the subsequent methods, the CPU times are still sufficiently smaller than the other methods, with exception of B3LYP. Larger basis sets are naturally more computationally expensive. However, we found that the relative energies obtained from larger basis sets are comparable to those with smaller basis sets so the use of larger more expensive basis sets was deemed unnecessary for this study.

Overall, the ranking for these conformers with the M06/6-31+G(d,p) method is reasonable within a few kJ/mol of the other compared but needing substantially shorter CPU times. Therefore, we adopted this method for the computation of relative conformational energies of a large number of SoNe pairs in the following sections.

3.2.2 Conformer Energies: Conformational Polymorphs vs. SoNe pairs

The relative conformer energies of conformational SoNe conformer pairs were calculated in this section for 148, 36, 21 and 18 SoNe pairs of water, methanol, DMF and benzene solvates respectively. This was achieved by retrieving molecular conformations of the HCs from the relevant crystal structures and performing relaxed gas-phase optimisations followed by a single point energy calculation with an implicit solvation model for the solvent in question (following the procedure outlined in Section 2.3). If the SoNe pairs differed in conformer energies by less than 2.5 kJ/mol (~RT), they were classified as having similar energies; if they differed by more than 2.5 kJ/mol they were classified as either pairs where the solvate conformer was more stable ($\Delta E_{Change}^{Neat \rightarrow Solv} < -2.5\ kJ/mol$) or pairs where the neat form conformer was more stable ($\Delta E_{Change}^{Neat \rightarrow Solv} > 2.5\ kJ/mol$). The data is analysed by solvent type in Figure 6.

Of the conformational SoNe pairs studied, we observe that in the majority of cases the conformers observed in the neat forms have similar energies to the conformers observed in the solvates (within 2.5 kJ/mol). Where the conformers were energetically different, it was equally likely that the most stable conformer would be found in either the solvate or the neat crystal structure. This is a surprising observation. We would have expected for solvates to have a greater potential to stabilise higher energy conformers through an increase in the number of possible intermolecular interactions (such as hydrogen bonds) because of the presence of the solvent. However, the data suggests that this is not the case, and that the incorporation of solvent in a structure does not always stabilise a higher energy conformer.

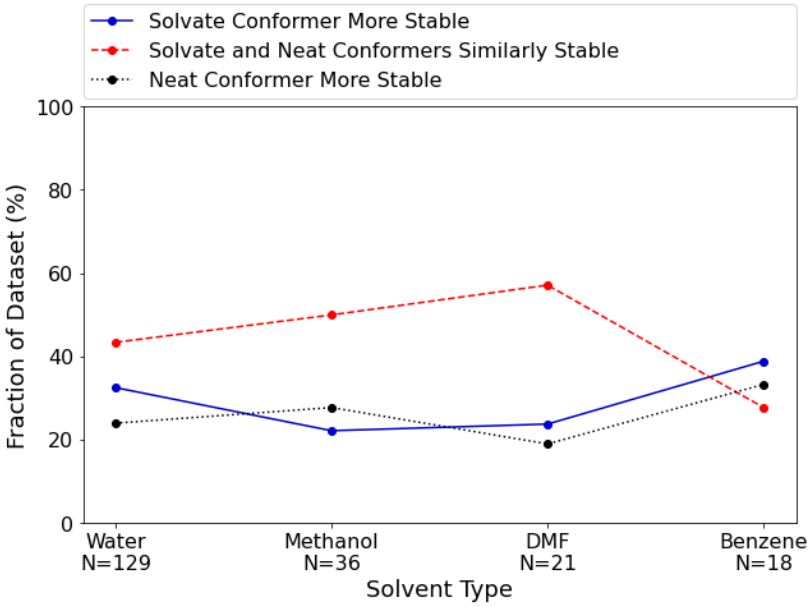


Figure 6: Comparison of the relative conformer stabilities of the SoNe conformer pairs for different solvates. Conformer pairs are classified as having similar stability when they differ in conformer energies by less than 2.5 kJ/mol. SMD solvation models for water, methanol, DMF and benzene were used for the water, methanol, DMF and benzene solvates respectively. N is the number of SoNe conformer pairs studied in each solvate category.

Histograms of conformer energy differences (ΔE_{change}) for these datasets are presented in Figure 7, with the methanol, DMF and benzene solvates combined as one plot due to the fewer number of datapoints available. We compare the conformational change histograms in the hydrates SoNe pairs and the solvates SoNe pairs with the conformational change histograms in conformational polymorphs.[23] The histograms for hydrates, solvates (methanol, DMF and benzene) and Conformational Polymorphs are all very similar, with many of the conformers being within 2.5 kJ/mol of each other (RT at room temperature), but with some high energy conformers being possible. 10% of both hydrate and solvate conformational SoNe pairs have relative conformer energies of over 20 kJ/mol.

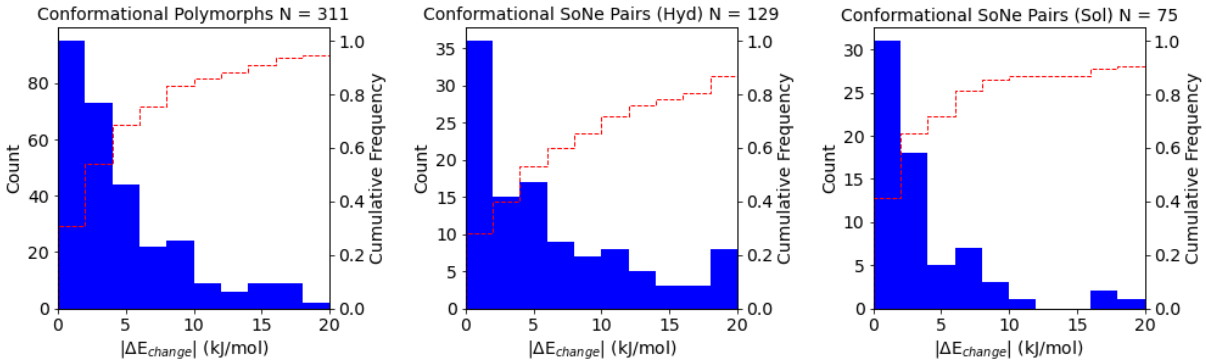


Figure 7: Histograms for the absolute gas phase conformer energies (M06/6-31+G(d,p)) between conformational SoNe Pairs (separated into hydrates (b) and solvates (c) combining data of methanol,

DMF and benzene solvates) compared to Conformational Polymorphs (a) adapted from Cruz-Cabeza and Bernstein[23] (B97-D/cc-pVTZ). The blue bars show the number of conformer pairs with relative energies within each bin (bin size of 2 kJ/mol), and the red dashed line shows the cumulative frequency.

3.3 Case Studies

A series of case studies have been chosen here for further study. We have focused on SoNe pairs with conformers differing more significantly in their stabilities. Energetics of the conformers are presented in three different solvents of increasing polarity: benzene, acetone and water. Calculations were done in five additional solvents (ethyl acetate, DMSO, 2-propanol, ethanol and methanol) but are only presented in the ESI to aid the readability of results in this section. Several of these systems have multiple solvate types, with different conformers present.

CASE 1: A tetra-triazolium macrocycle

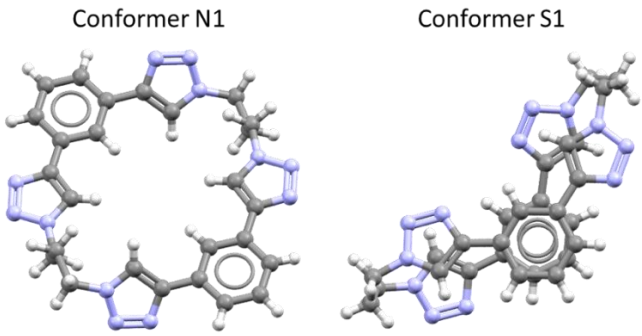


Figure 8: The two conformers of this tetra-triazolium macrocycle, left: the structure of the anhydrous form, right: the structure of the trihydrate with a folded conformation.

This tetra-triazolium macrocycle (3,4,5,9,10,11,19,20,21,25,26,27-dodecaazaheptacyclo[27.3.1.12,5.19,12.113,17.118,21.125,28]octatriaconta-1(33),2(38),3,10,12(37),13,15,17(36),18(35),19,26,28(34),29,31-tetradecaene) and its trihydrate occupy two very different conformers[72] (Figure 8) which differ in energy by 61 kJ/mol in the gas phase, the largest energy difference observed in all of the SoNe pairs studied. Conformer S1 is folded with aromatic interactions between the overlapping phenyl groups, and additionally stabilised in the crystal structure with hydrogen bonds to the water molecules, whereas conformer N1 is open, which forms pi-pi stacking interactions in the crystal structure, and has higher energy than conformer S1, as seen in Table 4. Accounting for the thermal free energy contributions to the conformational free energy in the calculations significantly reduces

the relative energies of the conformers by 17 kJ/mol (gas-phase). Further accounting for solvent implicitly also reduced the relative stabilities of the conformers. The two effects combined reduce the relative energy differences of these two conformers by over 30 kJ/mol from 61 kJ/mol (electronic energy only in the gas-phase) to around 30 kJ/mol (free energy in acetone or water).

Table 4: Conformer* Energies of the two conformers of the Tetra-triazolium Macrocycle in the gas phase, benzene, acetone and water.

	Conformer	Gas Phase	Apolar Benzene	Polar Aprotic Acetone	Polar Protic Water
ΔE_{change} (kJ/mol)	S1	0.00	0.00	0.00	0.00
	N1	61.08	50.40	43.32	45.42
ΔG_{change} (kJ/mol)	S1	0.00	0.00	0.00	0.00
	N1	43.20	37.63	27.27	30.72

*Conformer N1: anhydrous (LAYPAE), Conformer S1: hydrate (LAYPEI).

CASE 2: 2,5-bis(pyridin-2-yl)octahydropentalene-2,5-diol

2,5-Bis(pyridin-2-yl)octahydropentalene-2,5-diol exists as two polymorphs, two hydrates and a benzene solvate. In these crystal structures four different conformers are observed (Figure 9).[73] conformer N1/S1 is more stable than the other conformers, due to the intramolecular hydrogen bond which forms between the two hydroxyl groups. The other three conformers have a different ring conformation and no intramolecular hydrogen bond, conformer S2 and conformer N2 differ by the rotation of a pyridine group, and the least stable conformer is conformer S3 (Table 5).

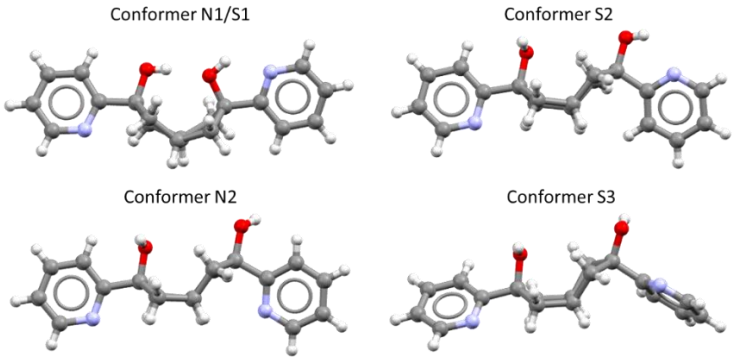


Figure 9: Conformers of 2,5-bis(pyridin-2-yl)octahydropentalene-2,5-diol.

The solvent type has a great effect on the relative stability of the conformers of this compound. Polar protic solvents appear to stabilise all conformers the most, with all conformers within 10 kJ/mol of the most stable. Thermal free energy corrections have a large effect on this system, in particular when combined with solvent effects. In water and accounting for thermal corrections, conformer S2 becomes slightly more stable than conformer N1/S1, and conformer S3 becomes much more accessible (only 3.85 kJ/mol compared to 30.75 kJ/mol conformer energy difference in the gas-phase).

Table 5: Conformer* energies of the four conformers of 2,5-bis(pyridin-2-yl)octahydropentalene-2,5-diol in the gas phase, benzene, acetone and water.

	Conformer	Gas Phase	Apolar Benzene	Polar Aprotic Acetone	Polar Protic Water
ΔE_{change} (kJ/mol)	N1/S1	0.00	0.00	0.00	0.00
	S2	16.45	15.05	10.52	1.13
	N2	21.85	19.97	15.64	6.55
	S3	30.75	27.13	18.87	8.00
ΔG_{change} (kJ/mol)	N1/S1	0.00	0.00	0.00	0.00
	S2	13.35	13.88	8.55	-0.48
	N2	17.49	17.94	11.59	3.94
	S3	23.60	23.18	13.72	3.85

* Conformer N1/S1: anhydrous (PEKJEW) and benzene solvate (PEKJOG), Conformer S2: hydrate (PEKJUM), Conformer N2: anhydrous (PEKJEW01), Conformer S3: hydrate (PEKJUM01).

CASE 3: N-((3-(Methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine

N-((3-(Methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine has three polymorphs, a hydrate and four solvates (methanol, benzene, chloroform and dioxane) with four different

conformers observed in total (Figure 10).[74,75] The folded conformers in conformer N1/S1 and conformer N2/S2 are both more stable than the more linear conformers S3 and S4.

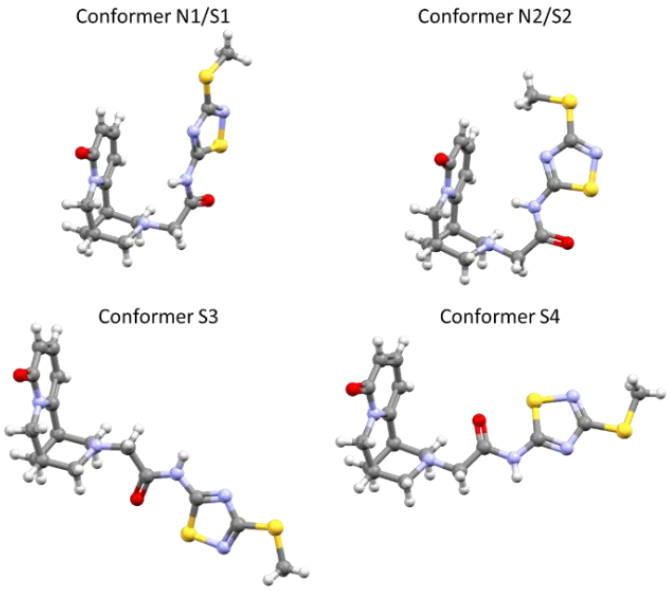


Figure 10: Conformers of N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine.

All conformers are stabilised significantly in polar solvents relative to the gas-phase (Table 6). For this system, however, the inclusion of thermal free energy corrections does not seem to have a significant re-ranking impact for these conformers.

Table 6: Conformer* energies of the four conformers of N-((3-(methythio)-1,2,4-thiadiazol-5-yl)aminocarbonylmethyl)cytisine in the gas phase, benzene, acetone and water.

	Conformer	Gas Phase	Apolar Benzene	Polar Aprotic Acetone	Polar Protic Water
ΔE_{change} (kJ/mol)	N1/S1	0.00	0.00	0.00	0.00
	N2/S2	7.23	5.72	2.53	0.36
	S3	21.61	21.47	18.86	9.45
	S4	28.66	25.82	17.86	6.43
ΔG_{change} (kJ/mol)	N1/S1	0.00	0.00	0.00	0.00
	N2/S2	8.24	5.60	6.05	7.14
	S3	16.38	21.97**	14.43	10.13
	S4	21.90	24.97	15.61	6.80

* Conformer N1/S1: anhydrous (IFILAL02), dioxane (LAJCUW), benzene (LAJDAD) and pyridine (IFILUF), Conformer N2/S2: anhydrous (IFILAL and IFILAL01) and chloroform (IFILEP), Conformer S3: benzene (IFILUF) and pyridine (LAJDAD), Conformer S4: hydrate (IFILIT) and methanol (IFILOZ). ** Even though Conformer S3 optimised to a minimum in the gas phase, in Benzene and Ethyl Acetate an imaginary frequency is observed which we have been unable to remove despite mode following.

CASE 4: Bicalutamide

Bicalutamide, a drug compound used for the treatment of prostate cancer, can be found in the CSD in two anhydrous forms, four cocrystals and one solvate with six different conformers (Figure 11).[27,76] The three folded conformers (N1, C1 and C2) are the stable conformers in the gas-phase all similar in energy. The three unfolded conformers (C3, N2 and S1) are all higher in energy in the gas-phase with when thermal contributions are not taken into account.

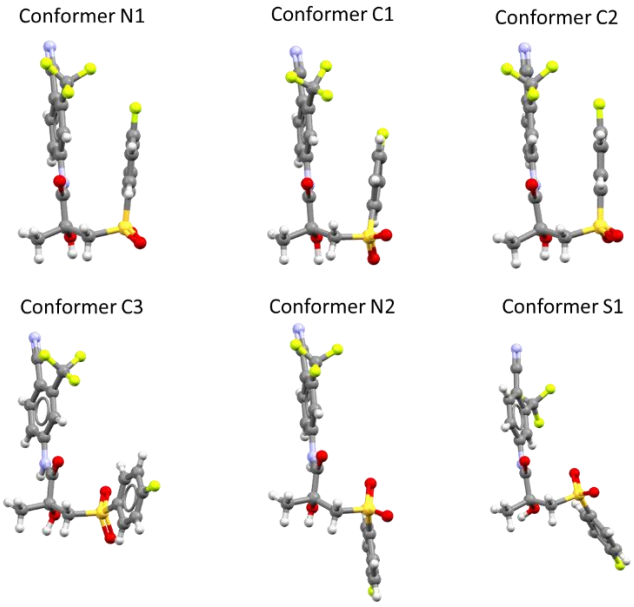


Figure 11: Conformers of bicalutamide. Two anhydrous polymorphs, four cocrystals and a solvate with six different conformers.

The unfolded conformers are always less stable than the folded conformers across all solvents, however their relative energies are much smaller than in the gas phase (Table 7). Particularly in the case of conformer S1, which in water is three times more stable than in the gas-phase. The addition of thermal free energies makes a huge difference to the relative stability of the conformers. Conformer C3 is significantly more stable in the gas phase and all solvents, and with benzene even becomes the most stable conformer. In acetone all the conformers are accessible and within a few kJ/mol of each other when thermal contributions are considered.

Table 7: Conformer* energies of the six conformers of Bicalutamide in the gas phase, benzene, acetone and water.

	Conformer	Gas Phase	Apolar Benzene	Polar Aprotic Acetone	Polar Protic Water
ΔE_{change} (kJ/mol)	N1	0.00	0.00	0.00	0.00
	C1	0.00	0.68	0.66	0.32
	C2	0.00	0.68	0.66	0.31
	C3	16.36	13.65	16.62	22.27
	N2	31.71	24.50	19.14	20.51
	S1	34.58	24.76	14.57	11.63
ΔG_{change} (kJ/mol)	N1	0.00	0.00	0.00	0.00
	C1	-1.31	0.55	1.00	-0.03
	C2	-1.30	0.54	1.00	0.01
	C3	0.06	-1.58	2.27	9.72
	N2	14.97	10.61	4.16	4.09
	S1	17.92	11.58	2.38	2.64

* Conformer N1: anhydrous (JAYCES02), Conformer C1: 4-4'-bipyridine (KIHZOR), Conformer C2: benzamide (EMUKEE) and salicylamide (EMUKII), Conformer C3: trans-1,2-bis(4-pyridyl)ethene (KIHZIL), Conformer N2: anhydrous (JAYCES), Conformer S1: sulfoxide (FAHFIG).

CASE 5: o-Acetamidobenzamide

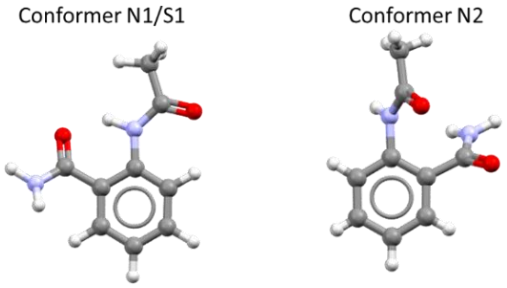


Figure 12: The two conformers of o-acetamidobenzamide. On the left the planar conformer (conformer N1/S1) and on the right the twisted conformer (conformer N2).

o-Acetamidobenzamide has two polymorphs[77] and six solvates.[78] Conformer N1/S1 is the most common experimentally and is found in the alpha polymorph as well as in all of its known solvates (Figure 12). This conformer is planar with an intramolecular hydrogen bond and it is the most stable conformer. Conformer N2 has a twisted geometry and is found in the beta polymorph; this conformer is always the higher in energy (Table 8) but it is significantly stabilised by the inclusion of polar solvents in the calculations (down to just 6 kJ/mol in water from around 30 kJ/mol in the gas-phase). Thermal free energy corrections have a small effect in these conformers.

Table 8: Conformer* energies of the four conformers o-acetamidobenzamide in the gas phase, benzene, acetone and water.

	Conformer	Gas Phase	Apolar Benzene	Polar Aprotic Acetone	Polar Protic Water
ΔE_{change} (kJ/mol)	N1/S1	0.00	0.00	0.00	0.00
	N2	37.38	29.26	15.52	9.79
ΔG_{change} (kJ/mol)	N1/S1	0.00	0.00	0.00	0.00
	N2	29.41	29.41	11.48	6.24

* Conformer N1/S1: anhydrous (ACBNZA02), chloroform (CEGBAS), 1-propanol (CEGBEW), 1-butanol (CEGBIA), 2-butanol (CEGBOG), 1,2-dichloroethane (CEGCAT) and 1,4-dioxane (CEGCEX) Conformer N2: anhydrous (ACBNZA01)

4 Discussion

Here we summarise our results and discussion as answers to the three original questions that motivated this study.

a. **How often is conformational change observed in SoNe pairs?** On average only 9% (between 4% and 17%) of solvates have a corresponding neat form thus making a SoNe pair. Of these SoNe pairs, 46% observe a conformational change between the solvate and the neat form conformers. Host compounds of SoNe pairs having a greater degree of flexibility are more likely to undergo a conformational change upon solvation. This observation mirrors well the behaviour of polymorphic compounds crystallising as conformational polymorphs.[23] Thus overall, the degree of molecular flexibility is the major factor impacting whether or not conformational change will occur with the solid form type (i.e. polymorph or solvate) having an impact which is harder to assess with this structural data alone.

b. **Are stable conformers more likely to crystallise as solvates than as neat forms?** Of the conformational SoNe pairs studied, we observe that the majority of conformers observed in either solvates or neat forms are similar in energy. When these conformers are energetically different, we found that there is no preference as to whether the stable conformer will crystallise as a solvate or as a neat form. We would have expected for higher energy conformers to be more prevalent in solvates due the extra added hydrogen bonding possible with the solvent and also because the solvate lattice energy is usually

lower than the lattice energy of the neat form. However, this is not the case. In this regard, the observations are highly system dependent thus each compound and each chemistry may show a different behaviour.

c. **What are typical energy differences in conformational solvates neat form pairs?** Typically, conformer energy differences within conformational polymorphs and SoNe pairs are small (60% differing by less than 5 kJ/mol). Only around 10% of SoNe conformer pairs have conformer energy differences greater than 20 kJ/mol. In our case studies, we have computed the relative energies for some of these systems in more detail by adding vibrational as well as solvent contribution corrections to conformer energies. In many of these systems, we found that the inclusion of vibrational contributions and thus computation of relative conformational free energies can have a significant impact on the stability ranking of these conformers. In many cases, energy differences between conformers shrank significantly upon computation of vibrational contributions. This effect was most significant for Bicalutamide. Accounting for solvation was also found to greatly impact the relative stabilities of conformers; this was most evident for compounds able to form conformers with and without strong non-covalent interactions. For example, higher energy conformers with a broken intramolecular hydrogen bond were found to become significantly more stable when including solvation. The combined effects of the vibrational contributions and solvent environment were found to be striking with high-energy conformers becoming significantly more stable upon inclusion of those. To illustrate this, we have presented the data of our case studies as histograms for relative conformer energies (ΔE_{change}) and conformer free energies (ΔG_{change}) in the gas-phase (blue) and in water (green) in Figure 13. Whilst in the original model (conformer energies in the gas-phase) some conformers differed in energies by almost 60 kJ/mol with a significant amount of them with energies around 30 kJ/mol, most conformers end up differing less than 10 kJ/mol with the final model (conformer free energies in water). This illustrates that high-energy conformer penalties are compensated through interaction with the environment (crystal or solvent) or favourable entropic contributions. These effects are important and should be accounted for in the simulation of conformers and their crystal structures.

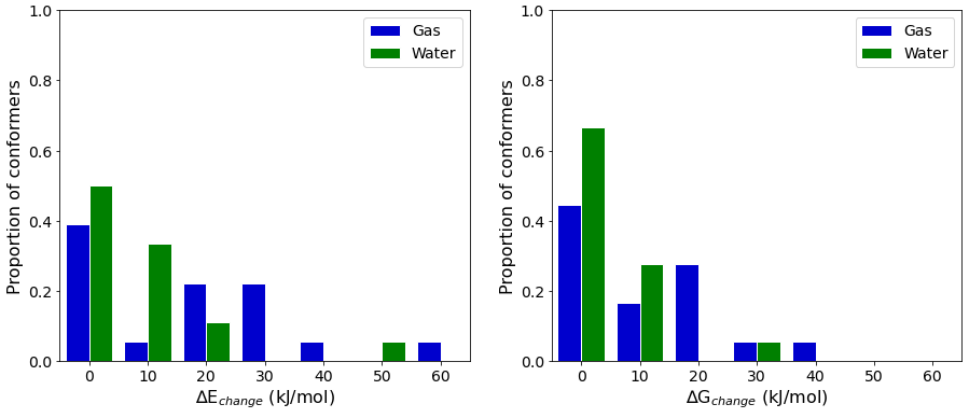


Figure 13: Distributions of relative conformer energies (left) and relative conformational free energies (right) for conformers in the five case studies. A total of 18 conformers were included, some are of high initial conformer energy in the gas phase (Bin size = 10 kJ/mol).

5 Conclusions

Our study finds that conformational change in SoNe pairs is relatively common, occurring in 46% of SoNe pairs. Perhaps as expected, highly flexible molecules have been shown to be more likely to undergo conformational change upon solvation. Our analysis of conformations in SoNe pairs mirrors very well previous analyses of conformations in conformational polymorphs. Thus, polymorphs or solvates have a similar potential to induce a conformational change of the compound under study upon crystallisation.

Regarding the energy differences between different conformers in SoNe pairs, the majority are very small (<2.5 kJ/mol). When energy differences are larger, there are no trends as to whether the most stable conformers will crystallise as a neat form or a solvate. Most importantly, we have shown that high energy conformers can be significantly stabilised by solvent effects as well as thermal contributions to the conformer energies. This is an important observation, especially for Crystal Structure Prediction of polymorphs and solvates; conformer selection for the computational generation of crystal packings may be better guided by conformer free energies in solution than by conformer energies in the gas-phase.

6 Supporting Information

Supporting information is available including information obtained from the CSD searches in Section 3.1 (refcodes, molecular overlay RMSD and DOFlex), energy data from Section 3.2 and energy data from the case study examples in Section 3.3 using additional solvents.

7 Bibliography

(1) Hilfiker, R. Polymorphism: In the Pharmaceutical Industry. *Polymorphism: In the Pharmaceutical Industry*. 2006, pp 1–414. https://doi.org/10.1002/3527607889.

(2) Healy, A. M.; Worku, Z. A.; Kumar, D.; Madi, A. M. Pharmaceutical Solvates, Hydrates and Amorphous Forms: A Special Emphasis on Cocrystals. *Adv. Drug Deliv. Rev.* **2017**, *117*, 25–46. https://doi.org/10.1016/j.addr.2017.03.002.

(3) Khankari, R. K.; Grant, D. J. W. Pharmaceutical Hydrates. *Thermochim. Acta* **1995**, *248* (C), 61–79. https://doi.org/10.1016/0040-6031(94)01952-D.

(4) Rodríguez-Spong, B. General Principles of Pharmaceutical Solid Polymorphism A Supramolecular Perspective. *Adv. Drug Deliv. Rev.* **2004**, *56* (3), 241–274. https://doi.org/10.1016/j.addr.2003.10.005.

(5) Gillon, A. L.; Feeder, N.; Davey, R. J.; Storey, R.; Amy, L.; Gillon, A. L.; Feeder, N.; Davey, R. J.; Storey, R. Hydration in Molecular Crystals - A Cambridge Structural Database Analysis. *Cryst. Growth Des.* **2003**, *3* (5), 663–673. https://doi.org/10.1021/cg034088e.

(6) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **2016**, *72* (2), 171–179. https://doi.org/10.1107/S2052520616003954.

(7) Threlfall, T. L. Analysis of Organic Polymorphs. *Analyst* **1995**, *120* (10), 2435. https://doi.org/10.1039/AN9952002435.

(8) Braun, D. E.; Griesser, U. J. Why Do Hydrates (Solvates) Form in Small Neutral Organic Molecules? Exploring the Crystal Form Landscapes of the Alkaloids Brucine and Strychnine. *Cryst. Growth Des.* **2016**, *16* (11), 6405–6418. https://doi.org/10.1021/acs.cgd.6b01078.

(9) Braun, D. E.; Karamertzanis, P. G.; Price, S. L. Which, If Any, Hydrates Will Crystallise? Predicting Hydrate Formation of Two Dihydroxybenzoic Acids. *Chem. Commun.* **2011**, *47* (19), 5443–5445. https://doi.org/10.1039/c1cc10762c.

(10) Byrn, S. R.; Zografi, G.; Chen, X. S. Solvates and Hydrates. In *Solid State Properties of Pharmaceutical Materials*; John Wiley & Sons, 2017; pp 38–47. https://doi.org/10.1002/9781119264408.ch3.

(11) Infantes, L.; Chisholma, J.; Motherwell, S. Extended Motifs from Water and Chemical Functional Groups in Organic Molecular Crystals. *CrystEngComm* **2003**, *5*, 480–486. https://doi.org/10.1039/b312846f.

(12) Infantes, L.; Fábián, L.; Motherwell, W. D. S. S. Organic Crystal Hydrates: What Are the Important Factors for Formation. *CrystEngComm* **2007**, *9* (1), 65–71. https://doi.org/10.1039/B612529H.

(13) Brychczynska, M.; Davey, R. J.; Pidcock, E. A Study of Methanol Solvates Using the Cambridge Structural Database. *New J. Chem.* **2008**, *32* (10), 1754–1760. https://doi.org/10.1039/b810063m.

(14) Brychczynska, M.; Davey, R. J.; Pidcock, E. A Study of Dimethylsulfoxide Solvates Using the Cambridge Structural Database (CSD). *CrystEngComm* **2012**, *14* (4), 1479–1484. https://doi.org/10.1039/c1ce05464c.

(15) Cruz-Cabeza, A. J.; Wright, S. E.; Bacchi, A. On the Entropy Cost of Making Solvates. *Chem. Commun.* **2020**, *56* (38), 5127–5130. https://doi.org/10.1039/D0CC01050B.

(16) Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. Structure Prediction, Disorder and Dynamics in a DMSO Solvate of Carbamazepine. *Phys. Chem. Chem. Phys.* **2011**, *13* (28), 12808–12816. https://doi.org/10.1039/c1cp20927b.

(17) Kazantsev, A. V; Karamertzanis, P. G.; Adjiman, C. S.; Pantelides, C. C.; Price, S. L.; Galek, P. T. A.; Day, G. M.; Cruz-Cabeza, A. J. Successful Prediction of a Model

Pharmaceutical in the Fifth Blind Test of Crystal Structure Prediction. *Int. J. Pharm.* **2011**, *418* (2), 168–178. https://doi.org/10.1016/j.ijpharm.2011.03.058.

(18) Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. Towards Prediction of Stoichiometry in Crystalline Multicomponent Complexes. *Chem. - A Eur. J.* **2008**, *14* (29), 8830–8836. https://doi.org/10.1002/chem.200800668.

(19) Cruz-Cabeza, A. J.; Karki, S.; Fábián, L.; Friić, T.; Day, G. M.; Jones, W. Predicting Stoichiometry and Structure of Solvates. *Chem. Commun.* **2010**, *46* (13), 2224–2226. https://doi.org/10.1039/b922955h.

(20) Cruz Cabeza, A. J.; Day, G. M.; Motherwell, W. D. S.; Jones, W. Amide Pyramidalization in Carbamazepine: A Flexibility Problem in Crystal Structure Prediction? *Cryst. Growth Des.* **2006**, *6* (8), 1858–1866. https://doi.org/10.1021/cg0601756.

(21) Iuzzolino, L.; McCabe, P.; Price, S. L.; Brandenburg, J. G. Crystal Structure Prediction of Flexible Pharmaceutical-like Molecules: Density Functional Tight-Binding as an Intermediate Optimisation Method and for Free Energy Estimation. *Faraday Discuss.* **2018**, *211*, 275–296. https://doi.org/10.1039/c8fd00010g.

(22) Görbitz, C. H.; Dalhus, B.; Day, G. M. Pseudoracemic Amino Acid Complexes: Blind Predictions for Flexible Two-Component Crystals. *Phys. Chem. Chem. Phys.* **2010**, *12* (30), 8466–8477. https://doi.org/10.1039/c004055j.

(23) Cruz-Cabeza, A. J.; Bernstein, J. Conformational Polymorphism. *Chem. Rev.* **2014**, *114* (4), 2170–2191. https://doi.org/10.1021/cr400249d.

(24) Wright, S. E.; Bryant, M. J.; Cruz-Cabeza, A. J. Is It Usual to Be Unusual? An Investigation into Molecular Conformations in Organic Crystals. *CrystEngComm* **2020**, *22* (43), 7217–7228. https://doi.org/10.1039/C9CE02001B.

(25) Back, K. R.; Davey, R. J.; Grecu, T.; Hunter, C. A.; Taylor, L. S. Molecular Conformation and Crystallization: The Case of Ethenzamide. *Cryst. Growth Des.* **2012**, *12* (12), 6110–6117. https://doi.org/10.1021/cg301244x.

(26) Campeta, A. M.; Chekal, B. P.; Abramov, Y. A.; Meenan, P. A.; Henson, M. J.; Shi, B.; Singer, R. A.; Horspool, K. R. Development of a Targeted Polymorph Screening

Approach for a Complex Polymorphic and Highly Solvating API. *J. Pharm. Sci.* **2010**, *99* (9), 3874–3886. https://doi.org/10.1002/jps.22230.

(27) Tkachev, V. V; Perlovich, G. L.; Blokhina, S. V; Manin, N. G.; Volkova, T. V; Tkachev, V. V. Polymorphism and Solvatomorphism of Bicalutamide: Thermophysical Study and Solubility. *J. Therm. Anal. Calorim.* **2013**, *111* (1), 655–662. https://doi.org/10.1007/s10973-012-2540-y.

(28) 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. https://doi.org/10.1039/C5SC00522A.

(29) Braun, D. E.; Gelbrich, T.; Kahlenberg, V.; Tessadri, R.; Wieser, J.; Griesser, U. J. Conformational Polymorphism in Aripiprazole: Preparation, Stability and Structure of Five Modifications. *J. Pharm. Sci.* **2009**, *98* (6), 2010–2026. https://doi.org/10.1002/jps.21574.

(30) Alkhamis, K. A.; Obaidat, A. A.; Nuseirat, A. F. Solid-State Characterization of Fluconazole. *Pharm. Dev. Technol.* **2002**, *7* (4), 491–503. https://doi.org/10.1081/PDT-120015052.

(31) Desai, S. R.; Shaikh, M. M.; Dharwadkar, S. R. Thermoanalytical Study of Polymorphic Transformation in Fluconazole Drug. *Thermochim. Acta* **2003**, *399* (1–2), 81–89. https://doi.org/10.1016/S0040-6031(02)00451-3.

(32) Park, H. J.; Kim, M. S.; Lee, S.; Kim, J. S.; Woo, J. S.; Park, J. S.; Hwang, S. J. Recrystallization of Fluconazole Using the Supercritical Antisolvent (SAS) Process. *Int. J. Pharm.* **2007**, *328* (2), 152–160. https://doi.org/10.1016/j.ijpharm.2006.08.005.

(33) Park, H. J.; Kim, M. S.; Kim, J. S.; Cho, W.; Park, J.; Cha, K. H.; Kang, Y. S.; Hwang, S. J. Solid-State Carbon NMR Characterization and Investigation of Intrinsic Dissolution Behavior of Fluconazole Polymorphs, Anhydrate Forms I and II. *Chem. Pharm. Bull.* **2010**, *58* (9), 1243–1247. https://doi.org/10.1248/cpb.58.1243.

(34) Karanam, M.; Dev, S.; Choudhury, A. R. New Polymorphs of Fluconazole: Results from Cocrystallization Experiments. *Cryst. Growth Des.* **2012**, *12* (1), 240–252. https://doi.org/10.1021/cg201005y.

(35) Gorkovenko, E. A.; Kichanov, S. E.; Kozlenko, D. P.; Belushkin, A. V.; Wąsicki, J.; Nawrocik, W.; Mielcarek, J.; Dubrovinsky, L. S.; Lathe, C.; Savenko, B. N. The Pressure-Induced Polymorphic Transformations in Fluconazole. *J. Pharm. Sci.* **2015**, *104* (12), 4164–4169. https://doi.org/10.1002/jps.24644.

(36) Caira, M. R.; Alkhamis, K. A.; Obaidat, R. M. Preparation and Crystal Characterization of a Polymorph, a Monohydrate, and an Ethyl Acetate Solvate of the Antifungal Fluconazole. *J. Pharm. Sci.* **2004**, *93* (3), 601–611. https://doi.org/10.1002/jps.10541.

(37) Alkhamis, K. A.; Salem, M. S.; Obaidat, R. M. Comparison between Dehydration and Desolvation Kinetics of Fluconazole Monohydrate and Fluconazole Ethylacetate Solvate Using Three Different Methods. *J. Pharm. Sci.* **2006**, *95* (4), 859–870. https://doi.org/10.1002/jps.20605.

(38) Basford, P. A.; Back, K. R.; Cram, M.; Docherty, R.; Davey, R. J.; Cruz-Cabeza, A. J. Impact of Crystal Structure and Molecular Conformation on the Hydration Kinetics of Fluconazole. *Cryst. Growth Des.* **2019**, *19* (12), 7193–7205. https://doi.org/10.1021/acs.cgd.9b01066.

(39) Basford, P. A.; Cameron, C. A.; Cruz-Cabeza, A. J. Conformational Change Initiates Dehydration in Fluconazole Monohydrate. *Cryst. Growth Des.* **2020**, *20* (9), 6044–6056. https://doi.org/10.1021/acs.cgd.0c00768.

(40) 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.; et al. *Gaussian 09, Revision D.01*; Gaussian, Inc., Wallingford, CT., 2013.

(41) 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. https://doi.org/10.1021/jp810292n.

(42) Ochterski, J. W. Thermochemistry in Gaussian. *Gaussian Inc Pittsburgh PA* **2000**, *264* (1), 1–19. https://doi.org/10.1016/j.ijms.2007.04.005.

(43) Görbitz, C. H.; Hersleth, H.-P. P. On the Inclusion of Solvent Molecules in the Crystal Structures of Organic Compounds. *Acta Crystallogr. Sect. B Struct. Sci.* **2000**, *56* (3), 526–534. https://doi.org/10.1107/S0108768100000501.

(44) Werner, J. E.; Swift, J. A. Organic Solvates in the Cambridge Structural Database. *CrystEngComm* **2021**, *23* (7), 1555–1565. https://doi.org/10.1039/D0CE01749C.

(45) Werner, J. E.; Swift, J. A. Data Mining the Cambridge Structural Database for Hydrate-Anhydrate Pairs with SMILES Strings. *CrystEngComm* **2020**, *22* (43), 7290–7297. https://doi.org/10.1039/d0ce00273a.

(46) Cruz Cabeza, A. J.; Pidcock, E.; Day, G. M.; Motherwell, W. D. S.; Jones, W. Space Group Selection for Crystal Structure Prediction of Solvates. *CrystEngComm* **2007**, *9* (7), 556–560. https://doi.org/10.1039/b702073b.

(47) Cole, J. C.; Raithby, P. R.; Taylor, R. Prior Likelihoods and Space-Group Preferences of Solvates. *Cryst. Growth Des.* **2021**, *21* (2), 1178–1189. https://doi.org/10.1021/acs.cgd.0c01490.

(48) Bryant, M. J.; Black, S. N.; Blade, H.; Docherty, R.; Maloney, A. G. P.; Taylor, S. C. The CSD Drug Subset: The Changing Chemistry and Crystallography of Small Molecule Pharmaceuticals. *J. Pharm. Sci.* **2019**, *108* (5), 1655–1662. https://doi.org/10.1016/j.xphs.2018.12.011.

(49) Karamertzanis, P. G.; Day, G. M.; Welch, G. W. A. A.; Kendrick, J.; Leusen, F. J. J. J.; Neumann, M. A.; Price, S. L. Modeling the Interplay of Inter- and Intramolecular Hydrogen Bonding in Conformational Polymorphs. *J. Chem. Phys.* **2008**, *128* (24), 244708. https://doi.org/10.1063/1.2937446.

(50) Greenwell, C.; Beran, G. J. O. Inaccurate Conformational Energies Still Hinder Crystal Structure Prediction in Flexible Organic Molecules. *Cryst. Growth Des.* **2020**, *20* (8), 4875–4881. https://doi.org/10.1021/acs.cgd.0c00676.

(51) Greenwell, C.; McKinley, J. L.; Zhang, P.; Zeng, Q.; Sun, G.; Li, B.; Wen, S.; Beran, G. J. O. Overcoming the Difficulties of Predicting Conformational Polymorph Energetics in Molecular Crystals: Via Correlated Wavefunction Methods. *Chem. Sci.* **2020**, *11* (8), 2200–2214. https://doi.org/10.1039/c9sc05689k.

(52) Nyman, J.; Yu, L.; Reutzel-Edens, S. M. Accuracy and Reproducibility in Crystal Structure Prediction: The Curious Case of ROY. *CrystEngComm* **2019**, *21* (13), 2080–2088. https://doi.org/10.1039/C8CE01902A.

(53) Becke, A. D. Density-functional Thermochemistry. II. The Effect of the Perdew–Wang Generalized-gradient Correlation Correction. *J. Chem. Phys.* **1992**, *97* (12), 9173–9177. https://doi.org/10.1063/1.463343.

(54) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785789.

(55) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate Spin-Dependent Electron Liquid Correlation Energies for Local Spin Density Calculations: A Critical Analysis. *Can. J. Phys.* **1980**, *58* (8), 1200–1211. https://doi.org/10.1139/p80-159.

(56) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627. https://doi.org/10.1021/j100096a001.

(57) Sousa, S. F.; Fernandes, P. A.; Ramos, M. J. General Performance of Density Functionals. *J. Phys. Chem. A* **2007**, *111* (42), 10439–10452. https://doi.org/10.1021/jp0734474.

(58) Austin, A.; Petersson, G. A.; Frisch, M. J.; Dobek, F. J.; Scalmani, G.; Throssell, K. A Density Functional with Spherical Atom Dispersion Terms. *J. Chem. Theory Comput.* **2012**, *8* (12), 4989–5007. https://doi.org/10.1021/ct300778e.

(59) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1–3), 215–241. https://doi.org/10.1007/s00214-007-0310-x.

(60) Karton, A.; Tarnopolsky, A.; Lamère, J. F.; Schatz, G. C.; Martin, J. M. L. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical . *J. Phys. Chem. A* **2008**, *112* (50), 12868–12886. https://doi.org/10.1021/jp801805p.

(61) Walker, M.; Harvey, A. J. A.; Sen, A.; Dessent, C. E. H. Performance of M06, M06-2X,

and M06-HF Density Functionals for Conformationally Flexible Anionic Clusters: M06 Functionals Perform Better than B3LYP for a Model System with Dispersion and Ionic Hydrogen-Bonding Interactions. *J. Phys. Chem. A* **2013**, *117* (47), 12590–12600. https://doi.org/10.1021/jp408166m.

(62) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27* (15), 1787–1799. https://doi.org/10.1002/jcc.20495.

(63) Becke, A. D. Density-Functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *J. Chem. Phys.* **1997**, *107* (20), 8554–8560. https://doi.org/10.1063/1.475007.

(64) Chai, J. Da; Head-Gordon, M. Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2008**, *10* (44), 6615–6620. https://doi.org/10.1039/b810189b.

(65) Chai, J. Da; Head-Gordon, M. Systematic Optimization of Long-Range Corrected Hybrid Density Functionals. *J. Chem. Phys.* **2008**, *128* (8). https://doi.org/10.1063/1.2834918.

(66) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46* (7), 618–622. https://doi.org/10.1103/PhysRev.46.618.

(67) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A Direct MP2 Gradient Method. *Chem. Phys. Lett.* **1990**, *166* (3), 275–280. https://doi.org/10.1016/0009-2614(90)80029-D.

(68) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 Energy Evaluation by Direct Methods. *Chem. Phys. Lett.* **1988**, *153* (6), 503–506. https://doi.org/10.1016/0009-2614(88)85250-3.

(69) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-Direct Algorithms for the MP2 Energy and Gradient. *Chem. Phys. Lett.* **1990**, *166* (3), 281–289. https://doi.org/10.1016/0009-2614(90)80030-H.

(70) Grimme, S.; Steinmetz, M. Effects of London Dispersion Correction in Density Functional Theory on the Structures of Organic Molecules in the Gas Phase. *Phys. Chem. Chem. Phys.* **2013**, *15* (38), 16031. https://doi.org/10.1039/c3cp52293h.

(71) Goerigk, L.; Grimme, S. A Thorough Benchmark of Density Functional Methods for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *Phys. Chem. Chem. Phys.* **2011**, *13* (14), 6670–6688. https://doi.org/10.1039/c0cp02984j.

(72) White, N. G.; Carvalho, S.; Félix, V.; Beer, P. D. Anion Binding in Aqueous Media by a Tetra-Triazolium Macrocycle. *Org. Biomol. Chem.* **2012**, *10* (34), 6951–6959. https://doi.org/10.1039/c2ob25934f.

(73) Gao, J.; Bhadbhade, M. M.; Bishop, R. Polymorphic Crystals Formed by an Achiral Diol under Ambient Conditions. *Cryst. Growth Des.* **2012**, *12* (11), 5746–5756. https://doi.org/10.1021/cg301259f.

(74) Tashkhodjave, B.; Turgunov, K. K.; Tojiboev, A. G.; Makhmudov, U. S.; Ibragimov, T. F.; Saprykina, V. A.; Shakhidoyatov, K. M. Crystal Solvates and Polymorphs of N-(3-Methylthio-1,2,4-Thiadiazol-5-Yl-Aminocarbonylmethyl)Cytisine. *J. Incl. Phenom. Macrocycl. Chem.* **2008**, *61* (1–2), 71–76. https://doi.org/10.1007/s10847-007-9394-0.

(75) Makhmudov, U. S.; Turgunov, K. K.; Tashkhodjaev, B.; Ibragimov, A. A.; Saprykina, V. A.; Shakhidoyatov, K. M. A New Polymorphic Form and Crystal Solvates of N-(3-Methylthio-1,2,4-Thiadiazol-5-Yl-Aminocarbonylmethyl)Cytisine. *J. Incl. Phenom. Macrocycl. Chem.* **2010**, *66* (3–4), 315–318. https://doi.org/10.1007/s10847-009-9638-2.

(76) Vega, D. R.; Polla, G.; Martinez, A.; Mendioroz, E.; Reinoso, M. Conformational Polymorphism in Bicalutamide. *Int. J. Pharm.* **2007**, *328* (2), 112–118. https://doi.org/10.1016/j.ijpharm.2006.08.001.

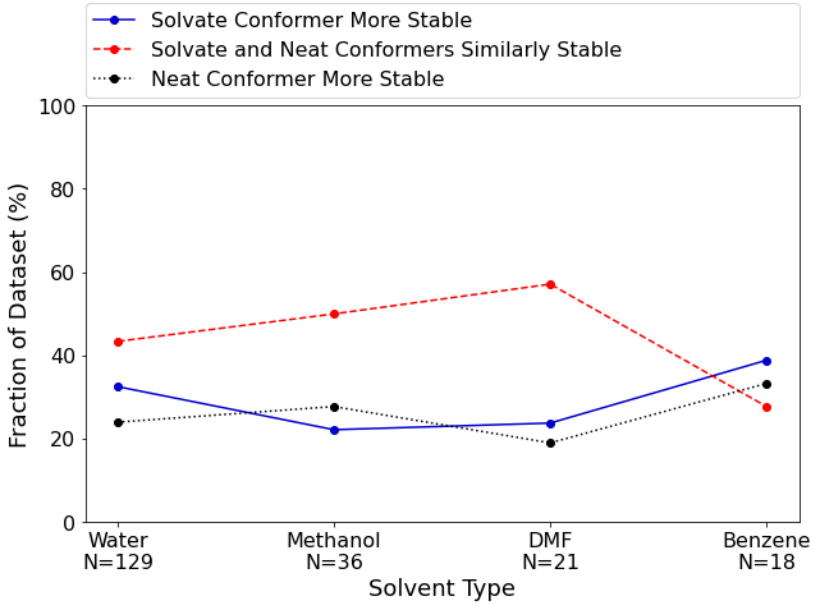
(77) Errede, L. A.; Etter, M. C.; Williams, R. C.; Darnauer, S. M. Solid-State Transformations and Crystal Structure Analysis of α- and β-o-Acetamidobenzamide. *J. Chem. Soc. Perkin Trans. 2* **1981**, No. 2, 233–238. https://doi.org/10.1039/P29810000233.

(78) Barnett, S. A.; Tocher, D. A.; Vickers, M. The Solvates of O-Acetamidobenzamide. *CrystEngComm* **2006**, *8* (4), 313–319. https://doi.org/10.1039/b600288a.

For Table of Contents Use Only

Conformational Change in Molecular Crystals: Impact of solvate formation and importance of conformational free energies.

Sarah E. Wright, Jason C. Cole and Aurora J. Cruz-Cabeza*



Conformational changes in molecular crystals due to solvate formation are investigated through an analysis of the Cambridge Structural Database and DFT calculations.