Quantitative Solution Measurement for the Selection of Complexing Agents to Enable Purification by Impurity Complexation

Citation: Weber, Cameron C., Geoffrey P. F. Wood, Andreas J. Kunov-Kruse, Douglas E. Nmagu, Bernhardt L. Trout, and Allan S. Myerson. "Quantitative Solution Measurement for the Selection of Complexing Agents to Enable Purification by Impurity Complexation." Crystal Growth & Design 14, no. 7 (July 2, 2014): 3649–3657.

As Published: http://dx.doi.org/10.1021/cg500709h

Publisher: American Chemical Society (ACS)

Persistent URL: http://hdl.handle.net/1721.1/97382

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Quantitative Solution Measurement for the Selection of Complexing Agents to Enable Purification by Impurity Complexation

Cameron C. Weber[†], Geoffrey P. F. Wood[†], Andreas J. Kunov-Kruse[†], Douglas E. Nmagu[†], Bernhardt L. Trout[†] and Allan S. Myerson[†]*

[†] Novartis-MIT Center for Continuous Manufacturing and Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

Abstract

The removal of carboxylic acid impurities from amide targets has been accomplished through crystallization featuring a complexing agent in solution. The interaction between the complexing agents and impurities was examined using isothermal titration calorimetry and the free energy of association obtained. From these data, the mechanism of solution association could be elucidated through the use of infrared spectroscopy, density functional theory calculations and the judicious variation of solvent and complexing agent. Furthermore, the calculations employed were able to predict the free energy of association between the complexing agents and the impurity. This association energy was found to correlate with the improvement in purification that could be achieved by the addition of these complexing agents. Optimal complexing agent choice in a rationally selected crystallization solvent produced improvements in purity of 96%, greater than could be achieved by successive crystallizations. The applicability of this complexing agent strategy was demonstrated for the purification of fenofibrate, an active pharmaceutical ingredient, from its major impurity fenofibric acid. Improvements in purity were again greater than 90%, far more than could be achieved by an additional crystallization step. The success of this strategy indicates the importance of solution association in determining the efficiency of complexing agents and suggests an approach towards the rational selection of these compounds *in silico*.

Introduction

Crystallization is the preferred purification technique for the generation of solid phases.[1] The effectiveness of crystallization as a purification process can be hindered by the incorporation of impurities within the target crystal lattice. Impurity inclusion may arise from the adherence of

mother liquor to the crystal, adsorption of impurities onto the crystal or the substitution of impurities into the crystal lattice.[2] Lattice substitution is generally favored by structurally similar compounds and can act as a significant limitation to the further purification of the target compound. Reducing impurity levels is of particular importance for the pharmaceutical and food industries where the presence of even low levels of certain compounds can have significant consequences for the safety, efficacy and regulatory approval of the product.

An approach that has been developed to improve the selectivity of crystallization processes featuring a dominant impurity involves the addition of a compound to the crystallization solution that is able to form an association complex with the impurity. This can prevent the incorporation of the impurity into the crystal lattice of the target by making its interaction with potential lattice sites less favorable, primarily due to steric effects.[3,4] This method has the advantage of improving the target purity without additional crystallization or purification steps unlike other approaches based on selective impurity interaction such as molecularly imprinted solid-phase extraction.[5,6]

The major complication with the complexing agent approach is the selection of an appropriate complexing agent, as this species must be able to selectively interact with the impurity and not the target. This is often confounded by the chemical similarity of many impurities and target compounds. Moreover, the complexing agent should not be incorporated within the crystalline lattice of the target and become an impurity itself. Previous work has focused on cocrystal screening as a method for complexing agent selection. Such investigations have uncovered successful complexing agents for crystallizations of cinnamamide (CAM) and amoxicillin.[3,4] In these examples, complexing agents were chosen if they selectively formed cocrystals with the impurity and not the target. However, this approach may be time consuming when compounds have relatively few known cocrystals listed in structural databases. In addition, it yields only qualitative data that requires further work to rank the effectiveness of the coformers discovered. Moreover, this approach does not account for the solution behavior of the coformers or solvent–solute interactions. Hence it would be extremely valuable to develop a quantitative solution approach whereby the prediction of successful complexing agents could be conducted without the extensive screening of potential compounds.

To study whether this quantitative approach to complexing agent selection could be achieved, model crystallizations with CAM and benzamide (BAM) as the targets and cinnamic acid (CA) and benzoic acid (BA) as the respective impurities were examined. The structures of these compounds are depicted in Figure 1. The chosen acids have been studied as 'tailor-made' additives that influence the crystal growth and habit of the corresponding amides, indicating the strength of interaction between these compounds.[7] The general extent of impurity inclusion in these compounds is relatively small but significant[2,8,9] and these systems have been explored previously for similar purification procedures which led to their selection as models.[4]

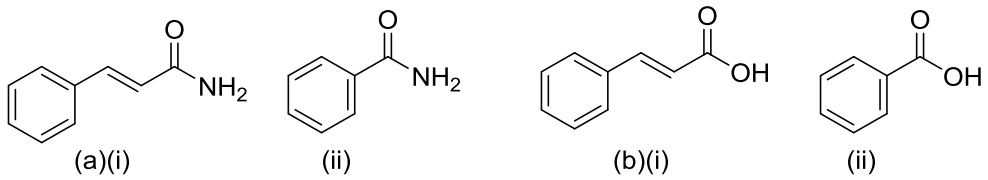


Figure 1. Structures of (a) the target compounds, (i) cinnamamide (CAM) and (ii) benzamide (BAM) and (b) the impurities, (i) cinnamic acid (CAM) and (ii) benzoic acid (BA).

Isothermal titration calorimetry (ITC) was used to obtain quantitative thermodynamic data for the interaction between the impurity and proposed complexing agents in solution. This technique is widely used to examine solution association, due to its sensitivity and universality relative to spectroscopic approaches.[10] ITC also enables the simultaneous determination of the enthalpy and entropy of association from a single experiment, which may give additional insight into the mechanism of interaction. In order for the correct interpretation of some of the experimental observations density functional theory (DFT) calculations have also been performed. These calculations provide insight into the specific species involved in the solution complexation. In addition, the energetic information obtained from the DFT calculations and the mechanistic understanding developed could be utilized to develop a theoretical approach to predict the strength of solution interaction as a means for screening potential compounds.

This study is significant in that it represents, to the best of the authors' knowledge, the first example of a combined experimental and computational approach to the prediction of

complexing agents for the purification of organic crystals. These predictions substantially increase the scope of this methodology as they enable the facile screening of a more substantial range of compounds.

Experimental Section

Materials. All compounds were used as received. trans-Cinnamic acid (CA), cinnamamide (CAM), benzoic acid (BA), benzamide (BAM), 2-amino-4,6-dimethylpyrimidine (ADPy), isonicotinamide (INA), dimethylglyoxime (DMG), 2-aminopyridine (2-AP), 2-hydroxypyridine (2-HP), 2-picolinic acid (2-PA), 1,3-diphenylguanidine (DPG), 1,3-di-*o*-tolylguanidine (DOTG), 1,3-diphenylurea (DPU), triethylamine (TEA), fenofibrate (FF), 2-(4-(4-chlorobenzoyl)phenoxy)-2-methylpropanoic acid (fenofibric acid, FFA), trifluoroacetic acid, deuterium oxide, water, methanol and acetone were all purchased from Sigma Aldrich. Dimethylsulfoxide (DMSO) and ethyl acetate were obtained from BDH. 200 proof ethanol was obtained from Koptec. Ethanol-d_1 was purchased from Acros Organics.

Isothermal Titration Calorimetry (ITC) Experiments. ITC experiments were conducted using a TA Instruments Nano ITC calorimeter. In a typical ITC experiment: the reference cell was filled with ethyl acetate, the sample cell filled with complexing agent (10 mM) with 10 μL aliquots of CA solution (100 mM) injected from the syringe stirrer every 300 s for a total of 20 injections at a stir speed of 250 rpm. The resultant heat output was subtracted from a blank injection of the same CA solution into ethyl acetate and the data fitted to an independent binding model, assuming 1:1 stoichiometry, using NanoAnalyze software. As the fitted constants were generally small, at least three experiments were conducted for each complexing agent using at least two different concentrations in the sample cell and, when solubility allowed, employing the complexing agent in the syringe to ensure reproducibility. This procedure was followed for BA, CAM and BAM titrations and for titration in ethanol, with ethanol substituted for ethyl acetate. For systems with large values of K, such as CA or BA with DOTG or DPG, lower concentrations (~ 1 mM of complexing agent, 10–25 mM CA/BA) were used.

Crystallization Experiments. BAM (2.044 g, 16.9 mmol), BA (0.226 g, 1.85 mmol) and complexing agent (1.85 mmol), were dissolved in 30 mL ethyl acetate at 50 °C, with the

temperature controlled by a recirculating water bath. The solution was cooled to 30 °C at a rate of 0.5 °C min^{-1} and further to 20 °C at a rate of 0.16 °C min^{-1}. The slurry was then stirred at 20 °C for an additional hour before the solid was filtered, washed with a small volume of cold ethyl acetate and dried at the pump. The collected solid was analyzed using high performance liquid chromatography (HPLC). The same experimental procedure was employed for crystallizations involving CAM and CA except 0.690 g of CAM and 0.077 g of CA were used in ethyl acetate and 2.522 g CAM and 0.280 g CA were used in ethanol.

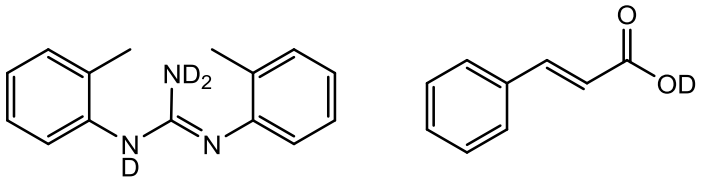
FF (0.504 g, 1.40 mmol), FFA (0.056 g, 0.176 mmol) and DOTG (0.042 g, 0.176 mmol) were dissolved in 10 mL ethanol at 25 °C. To these stirred solutions was added water (10 mL) and the resultant slurry stirred for an additional hour before the solid was filtered, washed with a small volume of a 1:1 ethanol: water (by volume) mixture and then dried at the pump. Second crystallizations were conducted using the same procedure with the solid obtained from the first crystallization with the solvent volume adjusted accordingly.

High Performance Liquid Chromatography (HPLC). The HPLC instrument (Agilent 1100) was equipped with a UV diode array detector. The column used was a YMC-Pack ODS-A 150 × 4.6 mm i.d. column packed with 3 μm particles with 12 nm pore size (YMC America Inc.). The detection wavelength was set at 230 nm for BA and 254 nm for CA and FFA. The samples were analyzed using an isocratic method with a 30/70 water/methanol mobile phase containing 0.1% trifluoroacetic acid for 5 mins for the analysis of CAM and BAM samples and for 25 mins for the analysis of FF samples.

Infrared (IR) Spectroscopy. Solutions of CA, DOTG, di-*o*-tolylguanidinium chloride ($DOTGH^+$) and a 1:1 molar ratio mixture of CA and DOTG were prepared in ethanol (~ 0.11 M). Lithium cinnamate (CA^-) was prepared as a filtered saturated solution due to its poor solubility in ethanol. ATR-FTIR spectra were recorded on a Nicolet 8700 FTIR instrument with a golden gate diamond ATR cell using a KBr beamsplitter and a liquid nitrogen cooled MCT detector. Each spectrum was recorded using 4 scans. All spectra were ATR-corrected using OMNIC 7 software assuming a refractive index equal to ethanol (n=1.361). Reported spectra were obtained by subtraction of the recorded spectrum with the spectrum of the pure solvent without further

manipulation. Due to the low solubility of CA^- in ethanol the intensity of these spectra were enhanced.

To strengthen the spectral interpretation, spectra of the analogous deuterated compounds were measured in ethanol-d_1. Deuterium exchange of the labile protons was achieved by separately dissolving DOTG (0.5 g) and CA (1 g) in 50 mL acetone and adding deuterium oxide (10 g) to each solution. After equilibration of the samples (~ 30 mins) the solvent was removed under reduced pressure. To make $DOTGD^+$-d_4, $DOTGH^+$ (0.5 g) was dissolved in deuterium oxide (10 g), equilibrated and the solvent removed under reduced pressure. The structures of DOTG-d_3 and CA-d_1 are depicted in Figure 2.



(i) (ii)

Figure 2. Structures of the deuterated compounds used for IR experiments, (i) DOTG-d_3 (drawn as the major tautomer)[11] and (ii) CA-d_1.

Computational Methodology. Standard *ab initio* molecular orbital theory[12] and density functional theory[13] calculations were carried out with the GAUSSIAN 09[14] computer program. Geometries of monomers and complexed dimers were initially scanned using the M06/6-31+G(d)[15] level of theory. Geometry optimizations were then carried out using the same level of theory at points on the scans that indicated equilibrium structures. The optimizations were followed by vibrational frequency analyses to ensure that the selected structures corresponded to minima (no imaginary frequencies) on the potential energy surface. All equilibrium geometries found using the M06/6-31+G(d) level of theory are summarized in Table S5 of the Supporting Information.

Free energies of association at 298K (ΔG^*_{298K}) between a selection of complexing agents and the impurity CA were obtained by combining energetic components from a series of computations. Energies of dimerization (ΔE_0) were computed by taking the difference in electronic energies of the dimers and isolated species from single-point calculations on the optimized structures using

the M06/6-311+G(3df,2p) level of theory. The basis sets employed in this study are sufficiently large to render the basis set superposition error (BSSE) negligibly small.[16] Following the calculation of ΔE_0 the binding free energy at 298 K in the gas-phase (ΔG_{298K}(bind)) was computed by adding entropic and thermal corrections using the rigid rotor harmonic oscillator approximation (RRHO) to this energy. In a small number of cases the free energy of binding using the hindered rotor approximation (HR) has also been calculated. To these energies the free energies of solvation ΔG_{298K}(solv) in ethyl acetate or ethanol have been added using the SM6[17] and cPCM[18] continuum solvation models. In previous studies it has been shown[19] that while continuum models may give poor results for total free energies in solution they can give very good results when combined with either the cluster-continuum solvation model[20-22] or by using cancellation of errors through an isodesmic reaction. The strategy adopted in the present study is the isodesmic approach where relative free energies have been obtained by taking the difference in association between the complexing agents and CA and CA with itself. All components of the final free energies are given in Tables S3 and S4 of the Supporting Information.

Results and Discussion

Selection of Complexing Agents

The selection of a variety of complexing agents able to interact with the impurity was required to screen the relative strength of potential modes of association. These complexing agents are depicted in Figure 3 with their corresponding abbreviations. ADPy, INA and DMG were selected as these compounds have been identified as forming cocrystals with CA and BA based on a search of the Cambridge Structural Database.[4] 2-AP, 2-HP and 2-PA were all investigated as these compounds can potentially form cyclic hydrogen bonds with the carboxylic acid moiety. Of these, 2-AP and 2-HP are able to form an 8-membered ring structure, a frequently observed hydrogen bonding motif for carboxylic acids, while 2-PA is only able to form a less favorable 9-membered ring.[23] It should be noted that 2-HP and 2-PA can exist as isomeric forms. In ethyl acetate solution, however, 2-PA adopts solely the structure depicted in Figure 3 while 2-HP will exist as a mixture of 2-hydroxypyridine and 2-pyridone.[24,25]

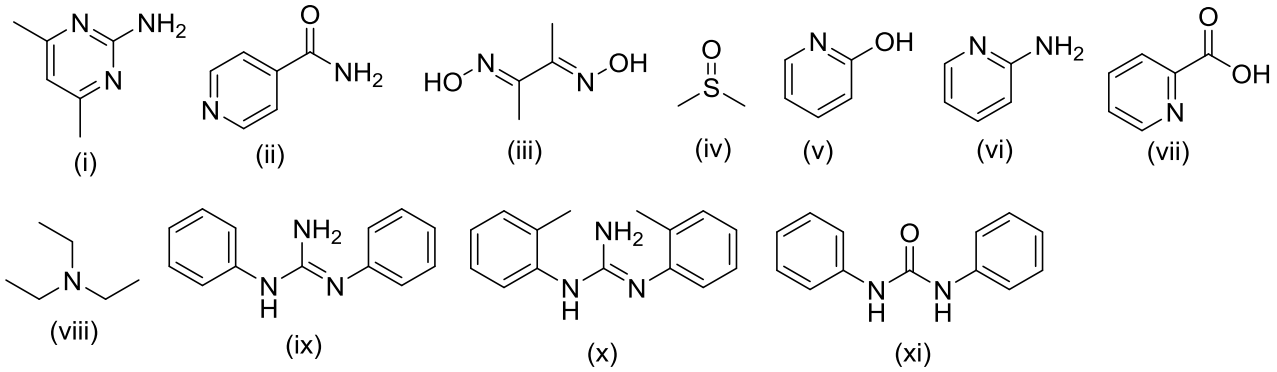


Figure 3. Structures of complexing agents (i) 2-amino-4,6-dimethylpyrimidine (ADPy), (ii) isonicotinamide (INA), (iii) dimethylglyoxime (DMG), (iv) dimethylsulfoxide (DMSO), (v) 2-hydroxypyridine (2-HP), (vi) 2-aminopyridine (2-AP), (vii) 2-picolinic acid (2-PA), (viii) triethylamine (TEA), (ix) 1,3-diphenylguanidine (DPG), (x) 1,3-di-*o*-tolylguanidine (DOTG) and (xi) diphenylurea (DPU).

DMSO is a strong hydrogen bond acceptor but is not Brønsted basic or capable of hydrogen bond donation and therefore would distinguish between pure hydrogen bonding interactions and acid/base reactivity as well as the importance of hydrogen bond multiplicity. TEA is not able to form cyclic hydrogen bonding motifs, however, it can act as a base and should indicate the extent to which basicity is significant. DPG and DOTG feature the guanidine moiety and were selected as guanidinium compounds have been extensively studied as receptors for carboxylates.[26] DPU was chosen as it mimics the hydrogen bond acceptor component of the guanidine moiety, although does not possess the same basicity.

Measurement of Solution Interaction Energy

ITC experiments were conducted to obtain a quantitative measure of the energetics of solution association between the complexing agents described above and the impurities. Tables 1 and 2 summarize the results for interactions between the complexing agents with CA and BA in ethyl acetate. Similar titrations were conducted between these complexing agents and CAM and BAM, however, no substantial interactions could be observed. A detailed list of the experiments conducted involving CAM and BAM is given in Tables S1 and S2 in the Supporting Information. The data in Tables 1 and 2 indicate that the associations between complexing agents and CA or BA display similar trends, with K values for the association with BA being slightly

larger in general. This subtle difference is likely due to the greater acidity of BA (pKa 4.20) relative to CA (pKa 4.44). Due to the structural similarities, this acidity difference indicates BA is likely to be a better hydrogen bond donor as well as a better proton donor.

Table 1. Association constants and thermodynamic data obtained between complexing agents and CA using ITC at 298 K in ethyl acetate. Reported errors are standard deviations from at least 3 replicate experiments.

complexing agent	K	ΔH (kJ mol^{-1})	ΔS (J K^{-1} mol^{-1})	ΔG (kJ mol^{-1})
ADPy	16.7 ± 0.5	−22.6 ± 0.3	−52.4 ± 1.2	−6.98 ± 0.07
INA	2.4 ± 0.5	−19 ± 3	−55 ± 10	−2.1 ± 0.5
DMG	ND[a]	N/A	N/A	N/A
DMSO	9.7 ± 1.6	−11 ± 3	−19 ± 10	−5.6 ± 0.4
2-HP	10.1 ± 0.9	−22.5 ± 0.4	−56.4 ± 1.5	−5.7 ± 0.2
2-AP	24 ± 3	−25 ± 3	−56 ± 8	−7.9 ± 0.3
2-PA	ND[a]	N/A	N/A	N/A
TEA	ND[b]	N/A	N/A	N/A
DPG	480 ± 40	−33.8 ± 1.5	−62 ± 5	−15.3 ± 0.2
DOTG	3000 ± 300	−32.1 ± 0.5	−41.1 ± 1.8	−19.8 ± 0.2
DPU	ND[a]	N/A	N/A	N/A

[a] Heat produced was too small for reliable model fitting

[b] Significant heat produced but K value too small to reliably fit

Table 2. Association constants and thermodynamic data obtained between complexing agents and BA using ITC at 298 K in ethyl acetate. Reported errors are standard deviations from at least 3 replicate experiments.

complexing agent	K	ΔH (kJ mol^{-1})	ΔS (J K^{-1} mol^{-1})	ΔG (kJ mol^{-1})
ADPy	16.7 ± 1.3	−23.6 ± 1.9	−56 ± 6	−7.0 ± 0.2
INA	2.3 ± 0.7	−24 ± 8	−70 ± 30	−2.1 ± 0.9
DMG	ND[a]	N/A	N/A	N/A
DMSO	11.8 ± 0.8	−11.3 ± 1.3	−17 ± 4	−6.11 ± 0.17
2-HP	17 ± 4	−19.5 ± 1.6	−42 ± 6	−7.0 ± 0.7

2-AP	31 ± 2	−26 ± 3	−59 ± 10	−8.49 ± 0.18
2-PA	ND[a]	N/A	N/A	N/A
TEA	ND[b]	N/A	N/A	N/A
DPG	1400 ± 200	−38 ± 5	−68 ± 15	−18.0 ± 0.4
DOTG	7500 ± 1200	−34.2 ± 1.5	−41 ± 5	−22.1 ± 0.4
DPU	ND[a]	N/A	N/A	N/A

[a] Heat produced was too small for reliable model fitting

[b] Significant heat produced but K value too small to reliably fit

Tables 1 and 2 illustrate that the compounds possessing guanidine groups gave the strongest interactions, with DPG and DOTG resulting in K values approximately two orders of magnitude greater than the other complexing agents. The complexing agents with the next strongest interactions are those capable of forming 8 membered cyclic hydrogen bond motifs, namely ADPy, 2-HP and 2-AP. The larger K values observed for 2-AP may indicate that basicity is also important as the pKa of the conjugate acid of 2-AP is 7.2 compared to 4.7 for ADPy. That 2-PA does not form a detectable interaction suggests the significance of the cyclic 8-membered hydrogen bond structure for intermolecular association with carboxylic acids. The lack of association for DPU despite its structural similarity with the guanidine moieties suggests that the basicity of DPG and DOTG is essential to their strong solution interactions with CA and BA. Given the result for TEA, a stronger base than DPG and DOTG, it is clear that basicity alone does not account for the association with the carboxylic acids and indicates that a combination of basicity and the ability to form cyclic hydrogen bonds is important for these strong interactions.

The microscopic origins of these interactions can be partially elucidated by further examination of the enthalpy and entropy values. ADPy, INA, 2-HP and 2-AP have changes in enthalpy and entropy upon association that are the same within experimental error. Notably each compound possesses a pyridyl moiety with a pendant hydrogen bond donating group. When compared to DMSO, which can only act as a hydrogen bond acceptor, the more negative enthalpy and less favorable entropy values of the former compounds likely arise from conformational restrictions imposed from multiple interaction sites. The stronger association observed between the acids and DPG arises from an enthalpic effect, likely due to ion-pairing interactions and strengthened

hydrogen bonds as a result of the basicity of DPG. Interestingly, the free energy difference between the acids and DOTG relative to DPG appears to arise from a more favorable change in entropy. The cause of this difference is not immediately apparent but may be due to the reduced conformational flexibility of DOTG resulting from the *o*-methyl substituents leading to pre-organization of DOTG towards the formation of the guanidinium ion.

To further explore the importance of ionization in the mechanism of association for DPG and DOTG, the effect of solvent was examined by ITC. For this study, ethanol was used instead of ethyl acetate and the interactions of DMSO, DPG, DOTG and 2-AP with CA measured. These compounds were selected as DMSO and 2-AP should interact primarily through neutral hydrogen bond complexes while it is hypothesized that DPG and DOTG interact through an ionic mechanism. As ethanol (ε 24.5) possesses a larger relative permittivity than does ethyl acetate (ε 6.0), it favors the ionization of solutes.[27] However, ethanol is a better hydrogen bond donor and acceptor than ethyl acetate and consequently solute–solute hydrogen bonding interactions are generally weaker in ethanol due to competition from the solvent. Consequently it would be anticipated that solutes that interact through a mechanism dominated by their ionization should have this interaction enhanced in ethanol while a neutral process would be disfavored. The ITC data obtained are summarized in Table 3.

Table 3. Association constants and thermodynamic data obtained between complexing agents and CA using ITC at 298 K in ethanol. Reported errors are standard deviations from at least 3 replicate experiments.

complexing agent	K	ΔH (kJ mol^{-1})	ΔS (J K^{-1} mol^{-1})	ΔG (kJ mol^{-1})
DMSO	ND[a]	N/A	N/A	N/A
2-AP	15.8 ± 0.9	−16.3 ± 0.6	−32 ± 2	−6.84 ± 0.13
DPG	21700 ± 1900	−40 ± 3	−51 ± 10	−24.8 ± 0.2
DOTG	57000 ± 3000	−43 ± 3	−52 ± 9	−27.14 ± 0.15

[a] Heat produced was too small for reliable model fitting

It is apparent from these measurements that the strength of association of CA with DPG and DOTG is roughly an order of magnitude stronger in ethanol than in ethyl acetate while the interaction with 2-AP is slightly weaker and no interaction with DMSO can be observed. These data are consistent with the ionization of DPG and DOTG being the predominant mechanism of association. It also indicates that ionization may play a minor role in the association with 2-AP, given the limited reduction in association strength compared to ethyl acetate.

Infrared Spectroscopy

To confirm the existence of an ionic mechanism for the guanidine moieties with the carboxylic acids, IR spectroscopy was used to probe the speciation of CA and DOTG in ethanol. The left hand side of Figure 4 depicts the difference IR spectra of dilute ethanolic solutions of DOTG, CA and a 1:1 molar ratio mixture of the two, as well as their ionic analogs in the form of the chloride and lithium salts respectively. The right hand side shows spectra of the analogous compounds with all mobile protons deuterated prior to dissolution in ethanol-d_1.

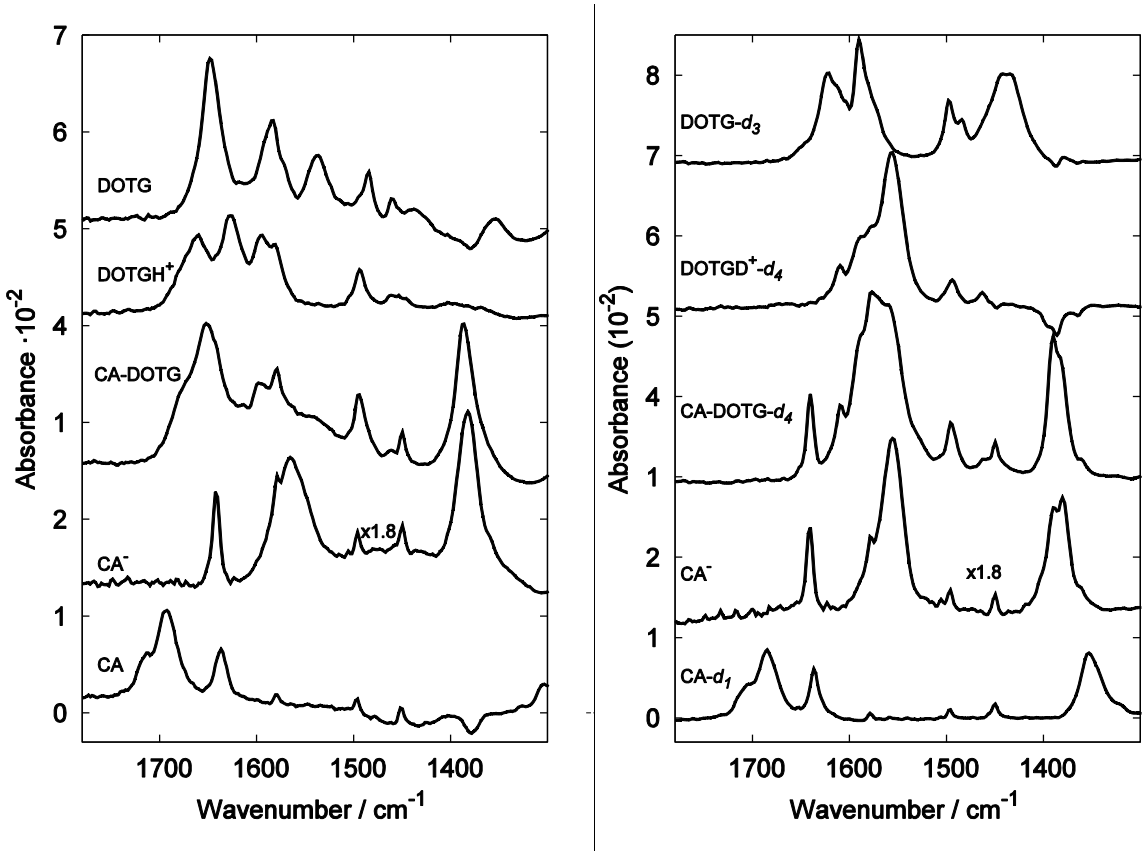


Figure 4. FTIR difference spectra of solutions of DOTG, $DOTGH^+$, 1:1 molar ratio mixture of DOTG and CA, CA^- and CA in ethanol (left) and the deuterated analogs in ethanol-d_1 (right). The original full range spectra are given in Figures S1 and S2 of the Supporting Information.

In the following discussion δ and ν are used to describe bending and stretching modes respectively. In the IR region depicted, the neutral DOTG shows intense bands with peaks at 1648, 1584, 1537 and 1485–1460 cm^{-1} corresponding to $\delta_{N\text{-}H}$, phenyl $\nu_{C=C}$ + $\delta_{C\text{-}H}$, $\delta_{NH2,scissor}$, mixed modes of $\delta_{N\text{-}H}$ and phenyl $\nu_{C=C}$ + $\delta_{C\text{-}H}$ respectively. The major bands for CA are at 1714 and 1694 cm^{-1} due to antisymmetric $\nu_{O=C=O}$ modes of the dimer and monomer respectively, as well as a band at 1637 cm^{-1} due to $\nu_{C=C}$ from the vinyl group adjacent to the carboxylic acid. When an equimolar mixture of both CA and DOTG is prepared the spectral features change significantly indicating the formation of a complex between the two. One significant difference is the disappearance of the two $\nu_{O=C=O}$ modes of the carboxylic acid and the appearance of an intense shoulder at 1675 cm^{-1}. This could suggest a strong neutral complex where the carboxylic $\nu_{O=C=O}$ modes shift downwards, hence the experiments with the deuterated substrates were conducted. The absorption at 1675 cm^{-1} was found to shift downwards upon deuteration, indicating this absorption is not due to a pure $\nu_{O=C=O}$ mode, as can be observed by looking at the $\nu_{O=C=O}$ modes in CA-d_1. Further, comparing the mixture to the spectrum of $DOTGH^+$ indicates the peak at 1675 cm^{-1} is present in the spectrum confirming that this mode is indeed an N-H bending of the protonated $DOTGH^+$ ion.

The very strong absorption of the symmetric $\nu_{O=C=O}$ mode at 1382 cm^{-1} in both the complex and CA^- provides further evidence for the deprotonation of CA in the CA-DOTG complex. One major difference between the spectrum of CA^- in ethanol and the complex is the change in intensity of the antisymmetric $\nu_{O=C=O}$ mode at 1565 cm^{-1}. This mode is much less IR active in the complex where it is only present as a weak shoulder band. The reduced intensity arises from coupling of the $\nu_{O=C=O}$ mode with a $DOTGH^+$ $\delta_{N\text{-}H}$ mode, indicative of a hydrogen bonding interaction between the ionic components. This coupling and change in intensity is consistent with modes calculated using DFT methods and shown in Figure S3 of the Supporting Information. Collectively this IR data indicates that the speciation of CA and DOTG in

equimolar mixtures in ethanol is dominated by the formation of a hydrogen bonded CA^--$DOTGH^+$ ion pair, as was hypothesized based on the ITC data.

Computational Results

Density functional theory (DFT) calculations in combination with continuum solvation models have been used to calculate the free energy of association between a selection of the complexing agents and CA. The final free energies of association have been computed by taking the difference between these values and the free energy of association between CA and itself. The FTIR experiments showed that there is a significant presence of CA dimers in ethanol, validating this assumption. The complexing agents we chose to explore include those that gave detectable K values, as given in Tables 1 and 3. In particular, the free energy of association between CA and the following complexing agents has been calculated in the presence of ethyl acetate: ADPy, INA, DMSO, 2-HP, 2-AP, DPG, and DOTG. In addition, the free energies of association between CA and the complexing agents 2-AP, DPG and DOTG have been calculated in the presence of ethanol. The results of these calculations are given in Figure 5, which displays the correlation between the experimentally derived ΔG values and the computed ones. Each species on the graph is labeled by its abbreviation. To distinguish between the different solvents used each species has been labeled with either a subscripted *ea* (for ethyl acetate) or a subscripted *eth* (for ethanol). To test the hypothesis of an ion-pairing mechanism the free energies of association for both the neutral pairs of DPG and DOTG (no identifying labels in the graph) and the ion pairs (identified with ± superscripts) have been explicitly calculated.

Figure 5 shows that the computational results give good predictions of the experimental values. This is demonstrated by the high R^2 value of 0.87 for the fitted line. Moreover, the y-intercept is close to zero at 0.16 kJ mol^{-1}. However the magnitude of slope is slightly high at 1.3. The origin of this high value is due to the underestimation of the calculated ΔG of association for $DPG^{\pm}$ and $DOTG^{\pm}$ in ethanol. We have recomputed the gas-phase components of the free energy for $DOTG^{\pm}$ using the hindered rotor approximation but this only shifts the calculated ΔG value by –0.1 kJ mol^{-1}. The most likely reason for the underestimate is a poor description of the solvation energy in ethanol. It is well known that charged species in high dielectric solvents are poorly modeled by continuum solvation methods. Therefore in order to remedy this problem a cluster-

continuum calculation is likely warranted but this is outside the scope of the current work. Despite these shortcomings there is very good agreement between the theoretical and experimental results over a wide-range of ΔG values.

It should be noted that while DMSO was found to have a detectable K value in ethyl acetate the theoretical calculations produced a ΔG value of +5.3 kJ mol^{-1} compared to the experimental value of −5.6 kJ mol^{-1}. The reason for the large difference in the calculated versus experimental numbers is likely due to the inaccuracy of the calculated binding energy between DMSO and CA. In fact, using the computationally more expensive CBS-QB3[28] methodology for the binding energy the calculated ΔG becomes +0.3 kJ mol^{-1}, which is more in line with the experimental findings. Due to this discrepancy the DMSO calculations have not been included on the graph given in Figure 5.

The theoretical calculations performed separately on the ion pairs of DPG and DOTG complexed with CA and their neutral counterparts are also given in Figure 5. From this Figure it is clear that if only the neutral pairings were present in solution then the computed ΔG values would be severely underestimated (see the red squares). In a higher dielectric constant solvent, such as ethanol, the difference in free energies between the charge separated species and the neutral forms should be even larger, which is what is observed. Moreover, the ΔG of 2-AP in ethanol is not greatly affected by the change in solvent further supporting the evidence that an ion-pair mechanism is at work for the guanidinium species. This is in accordance with the IR spectroscopy findings that ions are the dominant solution form in 1:1 mixtures of CA and DOTG in ethanol.

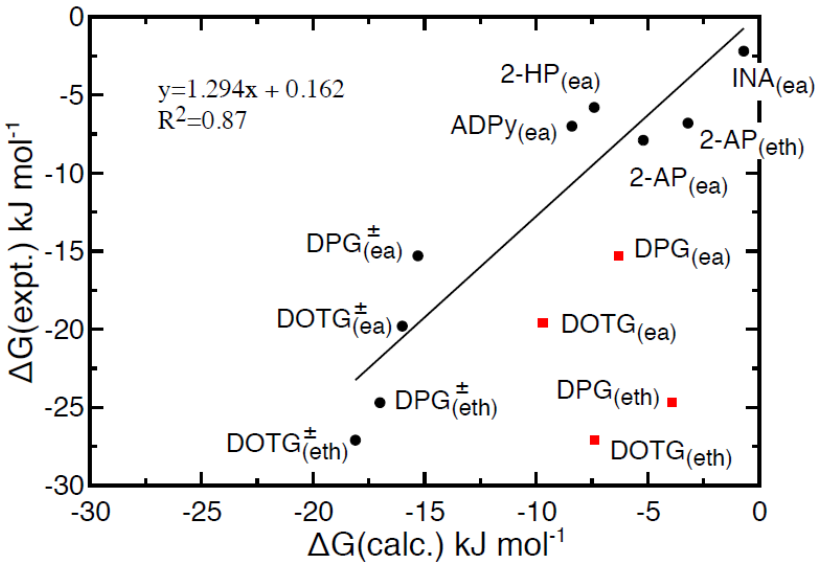


Figure 5: Correlation between the computed ΔG of association between the species labeled and CA. Subscripts *ea* and *eth* refer to the solvents ethyl acetate and ethanol, respectively. The ± superscripts indicate that the calculations were performed on the ion pair dimers. The regression line is fitted through the black circles only.

Purification Results

To determine whether the strength of solution association correlated with the ability of a complexing agent to achieve purification, crystallization experiments utilizing the CAM/CA and BAM/BA systems were conducted. Impure BAM and CAM samples containing 10 wt% of the respective impurities were purified by cooling crystallization. Complexing agents were employed in solution at a molar concentration equivalent to the impurity. The results of these purification experiments in ethyl acetate are summarized in Table 4.

Table 4. Concentration of impurity present after each crystallization in ethyl acetate, as determined by HPLC analysis. Concentration decrease values are relative to the no complexing agent result. Reported errors are standard deviations based on at least 3 replicate experiments. Bold values represent purification at the 95% level of significance.

complexing agent	CA (wt%)	relative CA concentration decrease (%)	BA (wt%)	relative BA concentration decrease (%)

None	0.35 ± 0.03	-	0.33 ± 0.04	-
None – 2nd crystallization	0.022 ± 0.004	94 ± 11	0.047 ± 0.002	86 ± 17
ADPy	0.27 ± 0.04	21 ± 15	0.24 ± 0.02	27 ± 14
INA	**0.23 ± 0.02[a]**	**34 ± 11[a]**	0.25 ± 0.02[a]	24 ± 14[a]
DMG	0.26 ± 0.04	26 ± 15	0.260 ± 0.009[a]	22 ± 13[a]
DMSO	0.21 ± 0.09	39 ± 28	0.24 ± 0.03	28 ± 15
2-HP	0.23 ± 0.08	33 ± 25	0.25 ± 0.04	26 ± 18
2-AP	**0.23 ± 0.03**	**34 ± 12**	**0.238 ± 0.008**	**29 ± 13**
2-PA	0.30 ± 0.11	14 ± 32	0.29 ± 0.03	12 ± 15
TEA	0.36 ± 0.04	−5 ± 14	1.29 ± 0.03	−286 ± 39
DPG	**0.113 ± 0.008**	**68 ± 10**	**0.118 ± 0.004**	**65 ± 15**
DOTG	**0.085 ± 0.003**	**76 ± 10**	**0.113 ± 0.008**	**66 ± 15**
DPU	0.38 ± 0.07[a]	−10 ± 21[a]	N/A[b]	N/A[b]

[a] Complexing agent included in the target compound

[b] DPU was not soluble at the concentration required for the experiment

From Table 4 it is apparent that DPG and DOTG gave the best purification results for both CAM and BAM and only 2-AP, DPG and DOTG yielded significant purification across both systems. No evidence of complexing agent incorporation could be observed in any of these cases. Notably these three compounds also gave the strongest solution association with CA and BA. Furthermore, the order of purity improvement was 2-AP<DPG<DOTG which accords with the relative strengths of solution interaction. None of the complexing agents investigated led to greater purities than could be achieved by repeated crystallization of either CAM or BAM.

Some complexing agent incorporation was observed for INA, DMG and DPU experiments although of these only INA with CAM led to significant purification, indicating that these complexing agents are not promising candidates notwithstanding their contamination of the target. TEA gave a surprising result with BAM, yielding a nearly threefold increase in impurity concentration. This result can be rationalized in terms of the basicity of TEA and the poor interaction between the resultant ions leading to coprecipitation of the salt, which would be

poorly soluble, from ethyl acetate solution. Unfortunately TEA is not UV active and therefore its presence cannot be confirmed by HPLC analysis. NMR analysis of the BAM crystallized in the presence of TEA was also not conclusive due to the poor sensitivity of the technique to the small amounts of included TEA expected. However, the increased concentration of BA relative to CA for the crystallization combined with the inherent basicity of TEA means that this explanation likely accounts for this unusual result. It is also important to note that the other complexing agents that did not yield a significant solution interaction, 2-PA, DMG and DPU, also performed poorly with regards to the purification of CAM and BAM.

To examine whether the solvent effects observed for solution interaction strength have any influence on the extent of purification, the crystallization of impure CAM was also attempted using ethanol as a solvent. The complexing agents that were investigated for the crystallization process were those previously studied by ITC and the results are summarized in Table 5. It is apparent that DOTG, 2-AP and DPG again gave significant purification, with no purification afforded by the addition of DMSO. The relative decrease in CA concentration is improved for all three complexing agents compared to their results in ethyl acetate. Notably, this leads to an improvement in purity of 96% for DOTG. The level of CA incorporation when DOTG is used as a complexing agent is now reduced below that which can be obtained by an additional crystallization. This level is significant in that it indicates that for any given purity requirement, the complexing agent approach discussed here will be able to obtain it in fewer crystallizations than in the absence of the complexing agent, resulting in a higher yield of the final product.

Table 5. Concentration of CA present after each crystallization in ethanol, as determined by HPLC analysis. Concentration decrease values are relative to the no complexing agent result. Reported errors are standard deviations based on at least 3 replicate experiments. Bold values represent purification by the complexing agent at the 95% level of significance.

complexing agent	CA (wt%)	relative CA concentration decrease (%)
None	0.67 ± 0.02	-
None – 2nd crystallization	0.077 ± 0.012	89 ± 5
DMSO	0.56 ± 0.09	17 ± 14

2-AP	**0.33 ± 0.03**	**50 ± 6**
DPG	**0.09 ± 0.02**	**86 ± 6**
DOTG	**0.026 ± 0.006**	**96 ± 5**

Finally, it is worth considering the extent to which the solution interaction strength acts as a quantitative predictor of purification efficiency. To visualize this, the improvement in purification for compounds that resulted in significantly enhanced purity was plotted against the free energy of association, shown in Figure 6. The INA/CAM system was excluded due to the incorporation of INA in the CAM crystal. This incorporation probably results in the operation of a different mechanism, for example INA may substitute for CA within the CAM lattice, rather than solution interactions dominating.

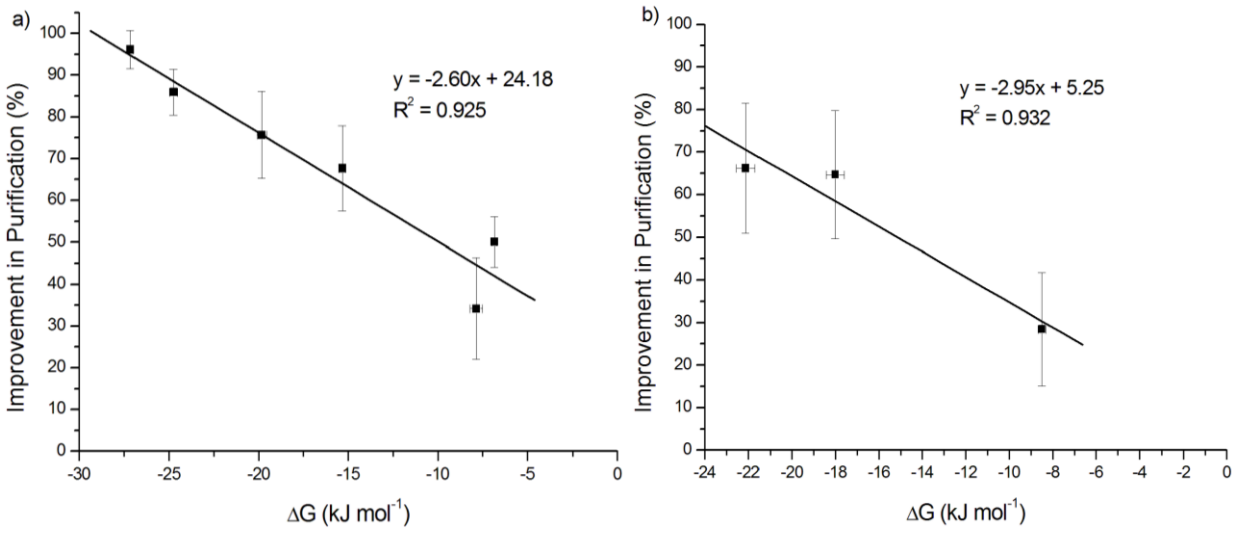


Figure 6. Correlation between the improvement in purification and free energy of association for (a) CAM and (b) BAM. Error bars indicate standard errors.

As can be observed in Figure 6, good correlations between the extent of purification and the free energy of association can be obtained, illustrating the effectiveness of the solution interaction approach at predicting good complexing agents. It is clear from the two fitted lines that the absolute trends differ even between these two closely related systems and it would not be expected that a general relationship between the free energy of association and improvement in purification would hold. Nonetheless, it is apparent that the free energy of association in solution

is a good parameter to employ to predict the performance of a given complexing agent for purification by crystallization and the ability to determine this value theoretically, guided by experimental results, makes this a powerful methodology for the future selection of these compounds.

To investigate whether the selected complex agent could be efficacious in effecting the purification of other related compounds, the purification by crystallization of fenofibrate (FF) was attempted. FF is an active pharmaceutical ingredient (API) used for the treatment of dyslipidemia and hypertriglyceridaemia.[29] A major impurity of FF is fenofibric acid (FFA) which can occur as both a synthetic and degradation impurity.[30] The structures of these compounds are given in Figure 7.

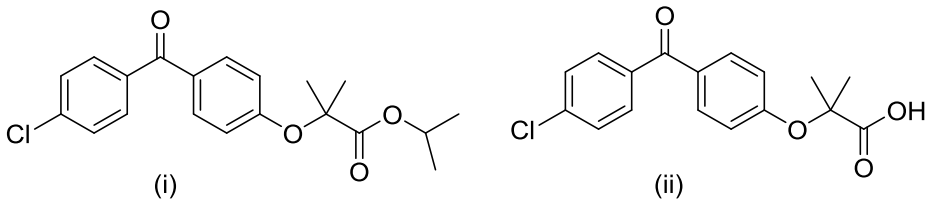


Figure 7. Structures of (i) fenofibrate (FF) and (ii) fenofibric acid (FFA).

As can be seen from Figure 7, the major functional difference between FF and FFA is the presence of a carboxylic acid moiety on the FFA compared with the ester in FF. This difference in functionality closely mimics that of the CAM/CA and BAM/BA systems studied, hence it would be anticipated that the solution behavior of FF and FFA with DOTG would be similar. To examine the impact of DOTG on the crystallization, impure FF containing 10 wt% FFA was crystallized from ethanol at 25 °C using water as an antisolvent in both the presence and absence of DOTG. The results of these experiments are summarized in Table 6.

Table 6. Concentration of FFA present after each crystallization, as determined by HPLC analysis. Improvements in purification are relative to the no complexing agent result. Reported errors are standard deviations based on 3 replicate experiments.

complexing agent	FFA (wt%)	relative FFA concentration decrease (%)

None	2.46 ± 0.05	-
None – 2nd crystallization	0.74 ± 0.04	70 ± 3
DOTG	0.159 ± 0.006	94 ± 3

As is evident from Table 6, DOTG leads to a significant improvement in the purity of FF compared to the crystallization in the absence of the complexing agent. The improvement in purity observed after a single crystallization with DOTG is substantially greater than could be achieved with two crystallization steps in its absence. Notably, the single crystallization with DOTG represents a 78% improvement in purity even compared to two crystallizations without the complexing agent. The FF results illustrate that once an optimal complexing agent is selected then it is possible to apply this complexing agent to purification processes involving compounds with similar differences in functionality in related solvent systems. It also demonstrates the ability of this rational complexing agent selection approach to be of significance in the design of purification processes for APIs.

Conclusions

The solution association energies, determined for a range of potential complexing agents with CA and BA using ITC, were found to be excellent predictors of the performance of these compounds at improving the resultant purification of a crystallization process. These free energies and the solution mechanism underpinning the association were able to be predicted computationally with a reasonable degree of accuracy using DFT calculations in combination with IR spectroscopy. As a result of understanding the mechanism of association and the solution association strength of the complexing agents with the impurity, a rational choice of complexing agent and solvent enabled an improvement in purification of CAM of 96%, greater than could be achieved by performing an additional crystallization. Furthermore, a 94% improvement in purification of the API FF could be accomplished using DOTG as a complexing agent. This was also greater than could be attained by an extra crystallization step, indicating the potential applicability of this approach to the purification of an extensive range of organic crystals, including pharmaceuticals. These findings demonstrate the feasibility of *in silico* complexing agent screenings which would significantly enhance the scope and applicability of this solution complexation approach to purification.

Associated Content

Supporting Information

ITC data for CAM and BAM, original IR spectra and calculation details. This material is available free of charge via the Internet at http://pubs.acs.org.

Author Information

Corresponding Author

myerson@mit.edu

Acknowledgements

The authors acknowledge Novartis for their support of this project.

References

(1) Giulietti, M.; Bernardo, A. In *Crystallization - Science and Technology*; Andreeta, M. R. B., Ed.; InTech: 2012, p 379.

(2) Meenan, P. A.; Anderson, S. R.; Klug, D. L. In *Handbook of Industrial Crystallization*; 2nd ed.; Myerson, A. S., Ed.; Butterworth-Heinemann: Boston, USA, 2002, p 67.

(3) Hsi, K. H. Y.; Concepcion, A. J.; Kenny, M.; Magzoub, A. A.; Myerson, A. S. *CrystEngComm* **2013**, *15*, 6776.

(4) Hsi, K. H.; Kenny, M.; Simi, A.; Myerson, A. S. *Cryst. Growth Des.* **2013**, *13*, 1577.

(5) Haginaka, J. *TrAC, Trends Anal. Chem.* **2005**, *24*, 407.

(6) Takeuchi, T.; Haginaka, J. *J. Chromatogr. B* **1999**, *728*, 1.

(7) Addadi, L.; Berkovitch-Yellin, Z.; Weissbuch, I.; Lahav, M.; Leiserowitz, L. *Mol. Cryst. Liq. Cryst.* **1983**, *96*, 1.

(8) Vaida, M.; Shimon, L. J. W.; van Mil, J.; Ernst-Cabrera, K.; Addadi, L.; Leiserowitz, L.; Lahav, M. *J. Am. Chem. Soc.* **1989**, *111*, 1029.

(9) Wang, J. L.; Berkovitch-Yellin, Z.; Leiserowitz, L. *Acta Cryst.* **1985**, *B41*, 341.

(10) Schmidtchen, F. P. In *Analytical Methods in Supramolecular Chemistry*; 2nd ed.; Schalley, C. A., Ed.; Wiley-VCH: Weinheim, Germany, 2012; Vol. 1, p 67.

(11) Koll, A.; Rospenk, M.; Bureiko, S. F.; Bocharov, V. N. *J. Phys. Org. Chem.* **1996**, *9*, 487.

(12) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.

(13) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2000.

(14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghayachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowki, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2009.

(15) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215.

(16) Grimme, S. *J. Comp. Chem.* **2004**, *25*, 1463.

(17) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2009**, *113*, 6378.

(18) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comp. Chem.* **2003**, *24*, 669.

(19) Ho, J.; Coote, M. L. *Theor. Chem. Acc.* **2010**, *125*, 3.

(20) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2001**, *105*, 7241.

(21) Pliego, J. R., Jr.; Riveros, J. M. *J. Phys. Chem. A* **2002**, *106*, 7434.

(22) Bryantsev, V. S.; Diallo, M. S.; Goddard III, W. A. C. *J. Phys. Chem. B* **2008**, *112*, 9709.

(23) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120.

(24) Mollin, J.; Kucerova, T. *Chem. Zvesti* **1979**, *33*, 52.

(25) Frank, J.; Katritzky, A. R. *J. Chem. Soc., Perkin Trans. 2* **1976**, 1428.

(26) Schmuck, C. *Coord. Chem. Rev.* **2006**, *250*, 3053.

(27) Reichardt, C.; Welton, T. *Solvents and Solvent Effects in Organic Chemistry*; 4th ed.; Wiley-VCH Verlag & Co. KGaA: Weinheim, 2011.

(28) Wood, G. P. F.; Radom, L.; Petersson, G. A.; Barnes, E. C.; Frisch, M. J.; Montgomery, J. A., Jr. *J. Chem. Phys.* **2006**, *125*, 094106.

(29) McKeage, K.; Keating, G. M. *Drugs* **2011**, *71*, 1917.

(30) Lacroix, P. M.; Dawson, B. A.; Sears, R. W.; Black, D. B.; Cyr, T. D.; Ethier, J.-C. *J. Pharmaceut. Biomed.* **1998**, *18*, 383.

TOC Graphic