JCTC Journal of Chemical Theory and Computation

Can the Formation of Pharmaceutical Cocrystals Be Computationally Predicted? 2. Crystal Structure Prediction

Panagiotis G. Karamertzanis,*,†,‡ Andrei V. Kazantsev,† Nizar Issa,‡ Gareth W.A. Welch,‡ Claire S. Adjiman,† Constantinos C. Pantelides,† and Sarah L. Price‡

Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, SW7 2AZ, United Kingdom, and Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, United Kingdom

Received October 13, 2008

Abstract: We report a multistage lattice energy minimization methodology for generating stable packing arrangements of cocrystals containing flexible molecules. In the first approximation, the intermolecular electrostatic interactions are modeled with atomic charges and the molecular deformation energy is interpolated over a set of precomputed quantum mechanical values. At subsequent stages, the accuracy is improved by first using analytically rotated and then conformation-dependent multipole moments, computed from the isolated-molecule charge density, and "on-the-fly" quantum mechanical calculations to compute the intramolecular deformation energy. This multistage approach increases the efficiency of the search and establishes the molecule-dependent error due to the atomic charge representation of the charge density and the neglect of the conformational dependence of atomic multipole moments. The methodology is used to study the lattice energy landscapes of the cocrystals of 4-aminobenzoic acid with 2,2′-bipyridine and 4-nitrophenylacetic acid, as well as the single-component crystals. All single-component, experimentally determined crystal structures within the scope of the search were found at, or very close to, the global minimum. The experimental cocrystal with 2,2′-bipyridine is also predicted to be among the most stable packing arrangements. On the contrary, the lattice energy landscape of the cocrystal with 4-nitrophenylacetic acid contains several low energy structures that are more stable than the experimentally observed form and have different hydrogen bonding motifs. Overall, the methodology can provide worthwhile crystal energy landscapes for multicomponent organic solids and thereby contribute to understanding cocrystal formation.

Introduction

The design of multicomponent crystals with desired properties,[1] and hence with specific molecular arrangements in the solid state, can be seen as the supramolecular analogue of chemical synthesis. However, while chemical synthesis has set the relationship between reactivity and chemical structure on a firm theoretical footing, supramolecular synthesis turns out to be a much more elusive task despite recent progress.[2,3] Although cocrystal formation can often be anticipated on the basis of the complementarity in hydrogen bonding capabilities of the component molecules,[4] our recent analysis of the cocrystals of three pharmaceutically acceptable coformers showed that a significant proportion of the cocrystals did not have the anticipated hydrogen bonding motif.[5] Cocrystals are expected to be formed if there exist packing arrangements containing both molecules that are thermody-

* Corresponding author. E-mail: p.karamertzanis@imperial.ac.uk.
† Imperial College London.
‡ University College London.

10.1021/ct8004326 CCC: $40.75 © 2009 American Chemical Society
Published on Web 04/08/2009

namically more stable than the pure component crystals. Hence, the development of algorithms[6] to predict the structure and thermodynamic stability of single and multicomponent crystals can assist the screening for cocrystals, by showing which, if any, arrangements of the two molecules are thermodynamically plausible. Such algorithms can also promote our understanding of the formation of molecular salts,[7,8] solvates,[9] and hydrates[10,11] and can also be used to rationalize crystallization with more than one chemically identical molecule in the asymmetric unit ($Z' \geq 2$).[12−18]

The most widely used approach for predicting organic crystal structures is to minimize the lattice energy of a large number of systematically or randomly generated candidate structures.[6,19−22] In this approach, the experimentally observed structure is assumed to correspond to the most stable packing arrangement. The identification of polymorphs as low-lying minima in a computational energy landscape is more likely to match physical reality when the entropic contributions to the free energy are also included.[23,24] Such calculations are usually successful in locating the experimentally determined polymorphs, although the rank of the generated structures depends crucially on the model for the crystal energy. Successful predictions[6,25] have been obtained by refining crude search structures with a more accurate energy model, based on either dispersion-corrected periodic density functional methods[26−28] or accurate models to describe the intermolecular contributions to the lattice energy, U_{inter}, and the energy cost of the molecular distortions, ΔE_{intra}, in the crystal.[29−32] It has been established that in the case of hydrogen bonded systems, the predicted relative stability of the structures is very sensitive to the model for the electrostatic contribution to U_{inter}, with the best results obtained using distributed multipole models.[33] The predicted lattice energies often depend significantly on the fine details of the molecular conformation, such as amino group rotation and pyramidalization,[34] which should ideally be optimized simultaneously with the lattice geometry.[29,31] For larger conformational changes, it is also necessary to consider the effect of the molecular conformation on the molecular charge density and consequently the model for the intermolecular interactions.[35−37]

This paper addresses the challenge of using such lattice energy evaluations to predict the range of possible crystal structures for a cocrystal[38] of assumed stoichiometry and for the component molecules and their relative lattice energies, without relying on experimental information. An appealing attribute of the method is that some conformational flexibility, such as rotation around a selected set of single bonds, is modeled from the outset, thus reducing the need to perform multiple searches with rigid conformers.[39,40] The approach is not limited to neutral molecules and is independent of whether the components are solids or liquids at ambient conditions. Therefore, it can be applied to the crystal structure prediction of salts, solvates, and other systems with more than one chemically identical molecule in the asymmetric unit. Finally, unlike an approach previously used for diastereomeric salts[8] and a monohydrate,[10] the method makes no assumptions about the hydrogen bonding between the crystallographically independent entities. This is likely to be particularly important for cocrystals, as it could lead to the identification of unexpected packing motifs that are thermodynamically competitive with the anticipated hydrogen bonding motifs used in cocrystal design.[2,3]

The computational requirements are kept manageable by using three minimization stages of increasing computational cost and accuracy. The first step uses CrystalPredictor[41,42] to generate a large number of crystal structures and to provide a first estimate of their relative stability. It uses an isotropic intermolecular potential for U_{inter} and estimates of ΔE_{intra} interpolated over a set of precomputed ab initio conformational energies. The computational efficiency and parallel implementation of this algorithm make it suitable for searching complex lattice energy surfaces such as those found for cocrystals, salts, and hydrates.

In the second step, a fast reminimization of the most stable unique structures produced by CrystalPredictor is carried out by replacing the point charges with a distributed multipole model to improve the quality of the representation of the electrostatic interactions, particularly the directionality of hydrogen bonds and $\pi-\pi$ stacking. A new algorithm, referred to as DMAflex-Quick, is used for this. It is computationally efficient due to the analytical rotation of the atomic multipole moments carried out to approximate the variation in molecular charge density with conformation. As at stage 1, precomputed interpolation tables are used for the computation of ΔE_{intra}.

At the final stage, the modeling accuracy is improved further by directly computing, with quantum mechanical calculations, the intramolecular energy, and molecular charge density for each conformation that arises in the lattice energy minimization. DMAflex[29] is used for this. The development of the intermediate DMAflex-Quick stage was prompted by the observation that the energy reranking by improving the representation quality of the molecular charge density is often so large that a complete CrystalPredictor search for cocrystals would require an impractically large number of quantum mechanical calculations for the DMAflex minimization of a sufficient number of crystal structures.

To test and illustrate the new algorithm, we apply it to the cocrystals formed by 4-aminobenzoic acid with 2,2′-bipyridine or 4-nitrophenylacetic acid and all three independent single-component crystals. These systems were selected from the set of 26 cocrystals used in part 1 of this paper[5] to provide a stringent test for the algorithm. The total number of degrees of freedom used in the search exceeds 20; this was previously found to be the upper limit, beyond which random UPACK searches frequently missed low energy minima.[11]

There are two known 4-aminobenzoic acid polymorphs. The high temperature polymorph α[43] ($Z' = 2$) comprises $R_2^2(8)$ dimers formed by inversion-related molecules and is kinetically favored at low temperatures under most crystallization conditions (see part 1). The hydrogen bond motif in polymorph β[44,45] ($Z' = 1$) is totally different and based on four-membered rings with alternating carboxylic and amino groups. The 4-aminobenzoic acid:4-nitrophenylacetic acid cocrystal has heterodimers adopting the same $R_2^2(8)$ carboxylic acid hydrogen bonding motif as the α polymorph.

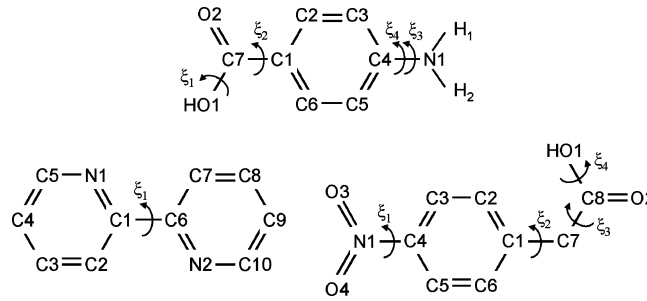
However, the cocrystals of 4-amino- and 4-nitrobenzoic acid with 2,2′- and 4,4′-bipyridine all contain carboxyl···$N_{pyridine}$ hydrogen bonds.[46] In all four of these cocrystals, bipyridine adopts a highly symmetric planar conformation and lies at either an inversion center or a 2-fold axis leading to a 1:2 stoichiometric ratio.

The first question we address is whether the searches are sufficiently complete to consistently find all experimental structures. Then, we analyze how the improvements in the model for the lattice energy change the relative stability of the crystal structures from one stage to the next. This allows us to assess the reranking due to the earlier approximations in the molecular structure, conformational energy, and intermolecular electrostatic interactions. The lattice energy landscapes of the component molecules are examined for the presence of distinct, yet equally thermodynamically favorable, packings. Such crystal energy landscapes, showing that the molecule has many different ways of packing, have been linked[47] to polymorphism and promiscuity in solvate and cocrystal formation.[48] Similarly, by examining the plurality of hydrogen bonding motifs in the cocrystal searches, we discuss whether the experimental structure has any definite energetic advantage over alternatives, and the degree to which cocrystals can be polymorphic. The stability of the low-energy predicted cocrystal structures is finally contrasted with the experimental and theoretically predicted crystal structures of the component molecules to determine whether cocrystal formation could have been predicted a priori.

Computational Methodology. *Extent of Molecular Flexibility.* Forces between molecules are generally weak compared with the forces between covalently bonded atoms. This explains the scarcity of high-energy conformers in experimentally determined crystal structures,[49−51] (There are exceptions for molecules where an intramolecular hydrogen bond in the gas-phase changes to an intermolecular hydrogen bond in the crystal structures[52,53]). Hence, molecular flexibility can often be limited to the distortion of a small number of "flexible" torsion angles denoted as $\boldsymbol{\xi}^{f}$ in the regions around the local conformational-energy minima. If the barrier to interconvert one conformational minimum to another is large, compared with the small energy differences of likely polymorphs, it is both possible and convenient to treat the conformational regions independently.[42]

In Scheme 1, we show the flexible degrees of freedom $\boldsymbol{\xi}^{f}$ that were varied during both the generation of crystal structures and the lattice energy minimizations. In Table 1, we compare the values of $\boldsymbol{\xi}^{f}$ in the experimental crystal structures and in the HF/6−31G(d,p) global conformational minima to assess the effect of the intermolecular forces on the conformations. On the basis of a series of one-dimensional gas-phase quantum mechanical relaxed scans, we modeled conformational changes in the region for which intramolecular energies do not exceed 20−30 kJ mol^{-1} more than the gas-phase global conformational minimum (shown as "range considered in CSP" in Table 1). We confirmed that the resulting region for 4-aminobenzoic acid includes all its conformations found in its 45 crystal structures in the Cambridge Structural Database[54] (CSD hereafter). The CSD analysis showed that the carboxylic group can be displaced by up to 15° ($-15° < \xi_2 < 15°$) from the benzene plane, while the variation in the amine group pyramidalization is greater. We ascertained that the crystal structure prediction searches did not generate any energetically plausible crystal structures containing molecular conformations that lie too close to the border of the conformational region modeled.

Scheme 1. Molecular Structures of 4-Aminobenzoic Acid (top), 2,2′-Bipyridine (bottom left), and 4-Nitrophenylacetic Acid (bottom right)[a]



[a] The flexible torsion angles varied in the generation of crystal structures and lattice energy minimizations are labeled (4-aminobenzoic acid: $\xi_1 \equiv$ H−O1−C7−C1, $\xi_2 \equiv$ O1−C7−C1−C6, $\xi_3 \equiv$ H2−N1−C4−C5, $\xi_4 \equiv$ H1−N1−C4−H2; 2,2′-bipyridine: $\xi_1 \equiv$ N1−C1−C6−N2; and 4-nitrophenylacetic acid: $\xi_1 \equiv$ O3−N1−C4−C3, $\xi_2 \equiv$ C8−C7−C1−C2, $\xi_3 \equiv$ O1−C8−C7−C1, $\xi_4 \equiv$ H−O1−C8−O2). Torsion ξ_4 of 4-aminobenzoic acid is an improper dihedral determining the amine group pyramidalization.

There is a subtle point about molecular symmetry that needs to be considered for cocrystals. The conformational region for 4-aminobenzoic acid contains only positive amine pyramidalization torsion angles (torsion ξ_4) and only describes one of the enantiomers. In the case of single-component crystals, this allows us to reduce the cost of crystal structure prediction without any loss of generality because the other enantiomer will be automatically generated by inversion or mirror plane symmetry operations in racemic space groups, and either enantiomer can be used in chiral space groups. However, the use of only one of the enantiomers of each coformer molecule results in an incomplete search for cocrystals in chiral space groups. Hence, for the cocrystal searches both enantiomers were represented in the conformational regions of both cocrystal formers; thus, the symmetry was exploited only for the purposes of constructing the grid of intramolecular energy values and gradients used by CrystalPredictor and DMAflex-Quick (see section 3.2). Restraining the 4-nitrophenylacetic acid torsion angle ξ_2 to positive values does not limit the scope of the cocrystal search because the molecule is symmetric under half-circle rotation of the nitrobenzene ring with respect to the propionic acid moiety.

Generation of Hypothetical Crystal Structures (CrystalPredictor). Hypothetical crystal structures were generated in 13 densely populated space groups $P1$, $P\bar{1}$, $P2_1$, $P2_1/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$, $Pbca$, $Pbcn$, $C2/c$, Cc, and $C2$ using low discrepancy (Sobol') sequences.[55] This was achieved by varying the lattice variables **X**, which comprise the cell lengths, the cell angles, the position and orientation of one of the fragments (base fragment) for each

Table 1. Flexible Torsion Angle Values in the Experimental Crystal Structures and the Gas-Phase HF/6-31G(d,p) Global Minimum for 4-Aminobenzoic Acid, 2,2′-Bipyridine, and 4-Nitrophenylacetic Acid[a]

	ξ_1	ξ_2	ξ_3	ξ_4
	4-Aminobenzoic Acid			
polymorph α (AMBNAC01[43]), $Z' = 2$	−11.8°	−3.4°	+24.3°	+144.6°
	−8.1°	+1.9°	+27.8°	+142.9°
polymorph β (AMBNAC04[45])	−3.8°	+9.3°	+26.8°	+124.6°
cocrystal with 2,2′-bipyridine (DAQYUQ[46])	+4.8°	−0.7°	+20.8°	+141.6°
cocrystal with 4-nitrophenylacetic acid (RILJOL[97])	+7.7°	+4.5°	+13.9°	+164.5°
HF/6-31G(d,p) gas phase	−0.1°	+0.2°	+22.4°	+137.5°
range considered in CSP	−40° to +40°	−30° to +30°	−50° to +50°	+90° to +180°
number of conformations calculated	5	5	7	7
	2,2′-Bipyridine			
pure-component crystal (BIPYRL04[98])	+180.0°			
cocrystal with 4-aminobenzoic acid (DAQYUQ[46])	+180.0°			
HF/6-31G(d,p) gas phase	+180.0°			
range considered in CSP	+0° to +360°			
number of conformations calculated	36			
	4-Nitrophenylacetic Acid			
pure-component crystal (SEMTAF01[99])	−10.4°	+72.3°	+169.2°	−3.1°
cocrystal with 4-aminobenzoic acid (RILJOL[97])	−4.3°	+43.6°	−154.8°	−2.4°
HF/6-31G(d,p) gas phase	+0.4°	+65.7°	+166.6°	−0.3°
range considered in CSP	−50° to +50°	+20° to +160°	−180° to +180°	−40° to +40°
number of conformations calculated	3	9	17	3

[a] The range of torsion angle values considered in the cocrystal structure prediction and the number of points used to construct the intramolecular energy Hermite interpolants are also given.

molecule in the asymmetric unit, and the flexible torsions $\boldsymbol{\xi}^{\mathrm{f}}$, the latter being restricted to the ranges shown in Table 1. The rest of the degrees of freedom, hereafter referred to as "rigid" and denoted $\boldsymbol{\xi}^{\mathrm{r}}$, were frozen at their values in the HF/6−31G(d,p) global conformational minimum, i.e., the molecules were treated as a set of rigid fragments linked with single bonds around which rotation is permitted.

The generation of hypothetical structures was carried out using CrystalPredictor.[41,42] The asymmetric unit of the 4-aminobenzoic acid:2,2′-bipyridine cocrystal contains one 4-aminobenzoic acid molecule and half-a 2,2′-bipyridine molecule and cannot be located in a search with 1:1 stoichiometry. Hence, we also performed a limited search using one 2,2′-bipyridine molecule and both enantiomers of 4-aminobenzoic acid in the asymmetric unit in space groups $P1$ and $P2_1$. This search should also produce crystal structures in $P\bar{1}$ and the experimental space group $P2_1/c$ with 2,2′-bipyridine lying at an inversion center. PLATON[56] was used to detect cases with higher symmetry than that imposed during lattice energy minimization.

Lattice Energy Minimizations. The minimization problem

$$\min_{\mathbf{X},\boldsymbol{\xi}^{\mathrm{f}}} E_{\mathrm{latt}} = U_{\mathrm{inter}}(\mathbf{X},\boldsymbol{\xi}^{\mathrm{f}};\boldsymbol{\xi}^{\mathrm{r}}) + \Delta E_{\mathrm{intra}}(\boldsymbol{\xi}^{\mathrm{f}}) \qquad (1)$$

is solved at each of the three stages with models that differ in the approximations that they use for the representations of both inter- and intramolecular energy terms and the rigid degrees of freedom $\boldsymbol{\xi}^{\mathrm{r}}$, as summarized in Figure 1. These differences result in a specific formulation for each stage and require separate solution procedures. The methodologies used in the first stage (CrystalPredictor[41,42]) and third stage (DMAflex[29]) have already been presented elsewhere and will therefore be described only briefly here along with the parameters used. The new DMAflex-Quick algorithm used in stage 2 will be discussed in more detail.

Molecular Model and Intramolecular Energy Potential. For both stage 1 and stage 2 minimizations, the values of the rigid degrees of freedom $\boldsymbol{\xi}^{\mathrm{r}}$ are taken from the HF/6−31G(d,p) global conformational minimum and held constant. The intramolecular energy $\Delta E_{\mathrm{intra}}$ is expressed as a function of the flexible torsions using multidimensional restricted Hermite interpolants[42] over a relatively coarse grid of tabulated quantum mechanical values of

$$\Delta E_{\mathrm{intra}}(\boldsymbol{\xi}^{\mathrm{f}}) = \min_{\tilde{\boldsymbol{\xi}},\mathrm{s.t.}\tilde{\boldsymbol{\xi}}^{\mathrm{f}}=\boldsymbol{\xi}^{\mathrm{f}}} E_{\mathrm{intra}}(\tilde{\boldsymbol{\xi}}) - E_{\mathrm{intra}}^{\mathrm{gas}} \qquad (2)$$

and $\partial\Delta E_{\mathrm{intra}}/\partial\boldsymbol{\xi}^{\mathrm{f}}$ covering the conformational range shown in Table 1. The constant $E_{\mathrm{intra}}^{\mathrm{gas}}$ is defined as the global molecular energy minimum $E_{\mathrm{intra}}^{\mathrm{gas}} \equiv \min_{\boldsymbol{\xi}} E_{\mathrm{intra}}(\boldsymbol{\xi})$.

In the third minimization stage, using DMAflex,[29] $\Delta E_{\mathrm{intra}}$ is expressed as a function of $\boldsymbol{\xi}^{\mathrm{f}}$ using "on-the-fly" quantum mechanical calculations, so that the rigid degrees of freedom, $\boldsymbol{\xi}^{\mathrm{r}}$, are allowed to relax in response to changes in the flexible torsion angles using the same isolated-molecule calculation:

$$\boldsymbol{\xi}^{\mathrm{r}}(\boldsymbol{\xi}^{\mathrm{f}}) = \arg\min_{\tilde{\boldsymbol{\xi}},\mathrm{s.t.}\tilde{\boldsymbol{\xi}}^{\mathrm{f}}=\boldsymbol{\xi}^{\mathrm{f}}} E_{\mathrm{intra}}(\tilde{\boldsymbol{\xi}}) \qquad (3)$$

In this work, all quantum mechanical calculations used to optimize molecular structures and compute intramolecular energies were performed at the HF/6−31G(d,p) level of theory using the electronic structure packages GAUSSIAN[57] or GAMESS.[58,59] Hence, the methodology only relies on quantum mechanical estimates of the intramolecular energy $\Delta E_{\mathrm{intra}}(\boldsymbol{\xi}^{\mathrm{f}})$, providing greater accuracy in general than empirical intramolecular force fields.[40,60]

Stage 1: Isotropic Intermolecular Potential (CrystalPredictor). The intermolecular energy is computed with an isotropic atom-atom model. The repulsion−dispersion interactions are modeled with the empirical exp-6 potential developed by Williams.[61] The electrostatic interactions are

	Molecular model	Intramolecular energy model	Intermolecular energy model	
			Electrostatic model	Repulsion-dispersion & other terms
Stage 1 (CrystalPredictor[41,42]) CPU cost: seconds	"Flexible" degrees of freedom (d.o.f.) ξ^f optimized within lattice energy minimization "Rigid" d.o.f. ξ^r frozen to their values in the quantum-mechanically optimized isolated molecule	Approximated with Hermite interpolants on a quantum mechanical (QM) pre-constructed grid	Atomic charges: Computed once for the gas phase global conformational minimum or conformationally dependent using Hermite interpolants on a QM pre-constructed grid	Isotropic, 12-6 or exp-6
Stage 2 (DMAflex-Quick) CPU cost: between 30 min and 3-4 hours			Atomic multipoles up to rank 4: Computed for a reference molecular geometry and rotated analytically with the local environment of the atoms	Isotropic, 12-6 or exp-6 or Any desired improvement that is cost effective in terms of CPU time versus re-ranking, such as use of anisotropic repulsion or C_8 dispersion terms
Stage 3 (DMAflex[29]) CPU cost: between 1 and 3 days	"Flexible" d.o.f. ξ^f optimized within lattice energy minimization "Rigid" d.o.f. ξ^r partially relaxed by isolated-molecule QM optimizations with flexible d.o.f. constrained	Accurate on-the-fly QM calculations	Atomic multipoles up to rank 4: Full conformational dependence with on-the-fly QM calculations of the molecular charge density	CPU intensive improvements such as use of distributed polarizability model

Figure 1. Overview summary of the multistage lattice minimization method. Indicative CPU cost is per crystal structure of the molecules used in this study.

modeled with a set of conformationally invariant atomic charges, fitted to the MP2(fc)/6−31G(d,p) electrostatic potential for the HF/6−31G(d,p) globally optimized molecule in isolation. The electrostatic potential is sampled using GAMESS[58,59] on a dense geodesic mesh of 32 layers, with the first layer at 1.4 times the van der Waals radii and successive layers separated by 0.025 times the van der Waals radii. For consistency with the force field specification,[61] the X−H bond lengths are foreshortened by 0.1 Å for both the fitting of charges and lattice energy minimizations. The repulsion−dispersion interactions are damped to zero at 18 Å using a quintic spline between 15 and 18 Å and summed in direct space. Charge−charge interactions are computed using Ewald summation.[62]

CrystalPredictor minimizes the lattice energy by simultaneously varying both the variables determining the lattice structure and the flexible torsion angles. It uses analytical gradients and a successive quadratic programming algorithm[41,42] that ensures robust and efficient minimization even when the starting point is far away from the minimum structure. Each minimization is very quick, requiring typically no more than a few CPU seconds even for cocrystal structures. In this study, computations were distributed over a 20-node Beowulf cluster. Fifty thousand minimizations were sufficient to provide a reasonable degree of confidence that the search for single-component crystals was complete. However, for the two-component searches, we had to perform significantly more minimizations: 565 000 for the more flexible 4-aminobenzoic acid:4-nitrophenylacetic acid cocrystal and 325 000 for the 4-aminobenzoic acid:2,2′-bipyridine cocrystal in 1:1 stoichiometry (plus 190 000 in 2:1 stoichiometry). Despite the enormous increase in computational effort, we occasionally encountered cocrystal structures within a few kJ mol^{-1} of the global minimum that were located only once, while other minima were much broader[11,40] and were found several tens of times. The stage 1 minima were clustered by comparing intermolecular distances; between 700 and 3000 unique minima were retained for further processing, depending on the system.

Stage 2: Improved Electrostatic Interactions (DMAflex-Quick). In stage 2, a key approximation of the model used in stage 1 minimizations is lifted by replacing the atomic charge model with an anisotropic model for the intermolecular electrostatic interactions comprising distributed atomic multipole moments up to rank 4. In the interest of computational efficiency, the multipole moments can be

computed only for a limited set of reference conformations from the MP2(fc)/6−31G(d,p) isolated-molecule charge density[57] using distributed multipole analysis.[63] Each atom is then assigned its own local axis system using two directly connected atoms (or, in the case of terminal atoms, a directly connected atom and a second neighbor). The reference multipole moments are then converted to their Cartesian form and rotated to the local axis system of each atom. In stage 2, the locally expressed reference multipoles are assumed to be constant despite the changes in flexible torsion angles in the course of lattice energy minimization. Thus, the conformational variability of the electrostatic model is accounted for in a limited manner via the analytical rotation of the local atomic multipoles to the molecular axis system of each newly generated conformation:

$$Q^{i}_{k_1 k_2 \cdots k_n}(\boldsymbol{\xi}^{\mathrm{f}};\boldsymbol{\xi}^{\mathrm{r}}) \approx \sum_{k'_1}\sum_{k'_2}\cdots\sum_{k'_n} R^{i}_{k_1 k'_1}(\boldsymbol{\xi}^{\mathrm{f}};\boldsymbol{\xi}^{\mathrm{r}}) R^{i}_{k_2 k'_2}(\boldsymbol{\xi}^{\mathrm{f}};\boldsymbol{\xi}^{\mathrm{r}})\cdots R^{i}_{k_n k'_n}(\boldsymbol{\xi}^{\mathrm{f}};\boldsymbol{\xi}^{\mathrm{r}}) Q^{i,\mathrm{reference}}_{k'_1 k'_2 \cdots k'_n} \quad (4)$$

where the multipole moment of rank n for atom i is computed by using the rotation matrix $\mathbf{R}(\boldsymbol{\xi}^{\mathrm{f}};\boldsymbol{\xi}^{\mathrm{r}})$ that transforms the local axis system to the molecular one. After each such rotation, the multipoles in the molecular axis system are converted[64] to the spherical-tensor formalism[65] and used to model the intermolecular electrostatic interactions with the program DMACRYS.[66,67]

The choice of reference conformations needs careful consideration depending on the conformational transferability[68] of the multipoles. Calculations on *o*-chlorobenzoic acid[69] show that using a single, gas-phase reference conformation can be a poor approximation, and this may also be the case for other systems that contain a flexible functional group in close proximity to other groups. In such cases, it is necessary to compute the multipole moments for the starting conformation or even to update the local multipole moments after every significant change in molecular geometry.[31] However, all the molecules in this study are para-substituted, so it was assumed sufficient to use the multipole moments of the gas-phase HF/6−31G(d,p) global conformational minimum for all DMAflex-Quick minimizations, knowing that the conformational dependence of the multipoles would be accounted for in stage 3.

The lattice energy minimization is formulated as a two-level optimization problem:

$$\min_{\mathbf{X},\boldsymbol{\xi}^{\mathrm{f}}} E_{\mathrm{latt}} = \min_{\boldsymbol{\xi}^{\mathrm{f}}}\left\{\Delta E_{\mathrm{intra}}(\boldsymbol{\xi}^{\mathrm{f}}) + \min_{\mathbf{X}} U_{\mathrm{inter}}(\mathbf{X};\boldsymbol{\xi}^{\mathrm{f}},\boldsymbol{\xi}^{\mathrm{r}})\right\} \quad (5)$$

with the inner minimization being carried out using the rigid-body lattice energy minimization. The outer minimization problem is solved using a Broyden−Fletcher−Goldfarb−Shanno (BFGS), quasi-Newton minimization algorithm,[70] using numerical gradients of U_{inter} with respect to the flexible degrees of freedom $\boldsymbol{\xi}^{\mathrm{f}}$, that are computed using forward finite differences with a 1.5° perturbation. The gradients $\partial\Delta E_{\mathrm{intra}}/\partial\boldsymbol{\xi}^{\mathrm{f}}$ are computed analytically by differentiating the Hermite interpolants as in the first stage. Charge−charge, charge−dipole, and dipole−dipole interactions are calculated with Ewald summation.[62] Repulsion−dispersion and higher multipole contributions are evaluated in direct space with a 30 Å cutoff to ensure that discontinuities in U_{inter}, due to molecules coming in and out of the cutoff distance do not lead to premature termination of the outer optimization.

At this stage, it would, in principle, be possible to also improve the modeling of the other contributions to the lattice energy, for example by introducing anisotropic atom−atom repulsion[71] or other more theoretically justified model intermolecular potentials.[30,72] However, we did not investigate this option; instead stage 2 minimizations were performed using the empirical, isotropic exp-6 model of Williams[61] but with a modified carboxylic proton−pyridine nitrogen repulsion term as discussed in Supporting Information.

The CPU cost per DMAflex-Quick lattice energy minimization is typically between 30 min and 3−4 h on a modern processor, depending on the number of flexible torsion angles. This timing would increase if conformation-specific initial charge density calculations were necessary. The DMAflex-Quick minima were clustered by comparing the 15-molecule (20-molecule for structures with more than one crystallographically independent molecule) coordination sphere[73] with a threshold of 0.4 Å, and the unique minima were passed to stage 3. To speed up the clustering process, structures that differed by 2 kJ mol^{-1} or more in energy were classed as different without further comparisons.

Stage 3: On-the-Fly Quantum Mechanical Calculations To Derive Conformation-Dependent Multipoles and Intramolecular Energies (DMAflex). The accuracy of lattice energies computed at stage 2 is limited because of three main reasons:

- The conformational transferability of the reference multipole moments is limited.[35,36]
- The rigid degrees of freedom $\boldsymbol{\xi}^{\mathrm{r}}$ are not relaxed in response to changes in the flexible torsions and the intermolecular forces.
- The computation of the intramolecular energy $\Delta E_{\mathrm{intra}}(\boldsymbol{\xi}^{\mathrm{f}})$ and its gradients $\partial\Delta E_{\mathrm{intra}}/\partial\boldsymbol{\xi}^{\mathrm{f}}$ is subject to interpolation errors arising from the use of coarse grids.

Stage 3 minimizations use DMAflex[29] to solve the two-level minimization shown in eq 5. Most sources of stage 2 errors identified above are addressed via the use of "on-the-fly" quantum mechanical calculations to compute the HF/6−31G(d,p) $\Delta E_{\mathrm{intra}}$ and the MP2(fc)/6−31 g(d,p) charge density and distributed multipole moments for the specific conformation in each outer minimization step. It is worth noting, however, that the effect of the intermolecular forces on rigid degrees of freedom $\boldsymbol{\xi}^{\mathrm{r}}$ is still neglected at stage 3, as $\boldsymbol{\xi}^{\mathrm{r}}$ are only relaxed within the isolated-molecule optimization.

We found that most of the large change in lattice energy from stage 2 occurs during the first stage 3 lattice energy evaluation using conformation-specific multipoles. Iterating the stage 3 DMAflex optimization to convergence typically further lowers the lattice energy by no more than 3 kJ mol^{-1} per molecule in the asymmetric unit. Hence, in the interest of computational efficiency, we started stage 3 simply by recalculating the lattice energy of all structures passed from stage 2 using multipole moments and rigid degrees of freedom $\boldsymbol{\xi}^{\mathrm{r}}$ recalculated to be consistent with the values of

Table 2. Lattice Energy Minimization of Experimentally Determined Crystal Structures

	lattice energy partitioning[a] (kJ mol^{-1})			lattice and conformational details							
	E_{latt}	ΔE_{intra}		a (Å)	b (Å)	c (Å)	β (deg)	density (g cm^{-3})	rmsd$_{mol}$[b] (Å)		rmsd$_{cs}$[b](Å)
Single-Component Crystals											
4-Aminobenzoic Acid Polymorph α, $Z' = 2$, $P2_1/n$ (CSD ref AMBNAC01,[43] $R = 7.03\%$, RT)											
experimental		2.57	2.08	18.551	3.860	18.642	93.6	1.367			
stage 1	−92.60	1.25	1.32	18.481	3.760	18.919	93.8	1.388	0.049	0.065	0.179
stage 2	−107.21	0.93	1.28	18.467	3.747	18.788	93.2	1.403	0.058	0.072	0.171
stage 3	−107.15	0.73	0.71	18.399	3.772	18.801	92.7	1.398	0.040	0.051	0.164
4-Aminobenzoic Acid Polymorph β, $Z' = 1$, $P2_1/n$ (CSD ref AMBNAC04,[45] $R = 4.99\%$, RT)											
experimental		2.68		6.278	8.583	12.365	100.1	1.389			
stage 1	−92.84	1.85		6.244	8.582	12.576	100.0	1.372	0.023		0.133
stage 2	−109.78	1.31		6.265	8.551	12.182	99.1	1.413	0.018		0.135
stage 3	−112.74	2.82		6.314	8.441	12.300	99.9	1.411	0.013		0.137
2,2′-Bipyridine, $Z' = 1/2$, $P2_1/n$ (CSD ref BIPYRL04,[98] $R = 2.98\%$, 123 K)[c]											
experimental		0.00		5.486	6.166	11.609	95.3	1.326			
stage 1	−79.61	0.00		5.567	6.349	11.442	96.9	1.291	0.013		0.190
stage 2	−81.41	0.00		5.680	6.180	11.490	95.8	1.293	0.013		0.165
stage 3	−81.41	0.00		5.680	6.180	11.490	95.8	1.293	0.013		0.165
4-Nitrophenylacetic Acid, $Z' = 1$, *Pbca* (CSD ref SEMTAF01, $R = 4.1\%$, 110 K)[99]											
experimental		0.76		14.982	7.010	15.885	−	1.442			
stage 1	−99.41	0.72		15.321	7.014	15.829	−	1.414	0.053		0.388
stage 2	−117.79	1.01		15.175	6.810	16.070	−	1.449	0.047		0.241
stage 3	−117.31	1.10		15.176	6.823	16.023	−	1.450	0.052		0.243
Cocrystals											
4-Aminobenzoic Acid:2,2′-Bipyridine, $Z' = 3/2$, $P2_1/c$ (CSD ref DAQYUQ,[46] $R = 5.73\%$, RT)[c]											
experimental		0.48	0.00	15.820	4.309	16.466	107.2	1.332			
stage 1	−259.94	0.69	0.00	16.443	4.140	16.284	103.1	1.323	0.071	0.077	0.472
stage 2	−310.02	1.33	0.00	16.165	4.264	15.787	106.0	1.367	0.081	0.077	0.295
stage 3	−309.68	1.33	0.00	16.167	4.286	15.721	105.9	1.365	0.079	0.077	0.305
4-Aminobenzoic Acid:4-Nitrophenylacetic Acid, $Z' = 1$, $P2_1/a$ (CSD ref RILJOL,[97] $R = 5.75\%$, RT)											
experimental		3.24	4.94	15.272	5.910	17.194	108.8	1.452			
stage 1	−188.89	0.36	4.14	15.154	5.928	16.933	106.0	1.445	0.024	0.065	0.255
stage 2	−217.48	0.79	5.82	15.276	5.876	16.752	104.4	1.452	0.027	0.042	0.322
stage 3	−217.63	2.62	4.93	15.059	5.922	16.850	105.2	1.458	0.027	0.034	0.251

[a] The HF/6−31G(d,p) intramolecular energy ΔE_{intra} is shown separately for each molecule in the asymmetric unit in the order shown in the system title. The experimental ΔE^{intra} refers to the HF/6−31G(d,p) molecular energy of the gas-phase optimized molecule with the flexible torsion angles identified in Scheme 1 constrained to their experimental values. [b] Hydrogen atoms were omitted for both molecular (rmsd$_{mol}$) and coordination sphere (rmsd$_{cs}$) root-mean-square comparison[73] because of the uncertainties in the X-ray experimental determination of their position. For the calculation of rmsd$_{cs}$ we used 15 and 20 molecules in the coordination sphere for structures with one or more crystallographically independent molecule respectively. The rmsd$_{mol}$ for each molecule in the asymmetric unit is given in the order shown in the system title. [c] Minimizations performed by reducing the symmetry to $P2_1$ to have complete molecules in the asymmetric unit. The minimized crystal retained[56] the higher symmetry of the experimental crystal. E_{latt} for the 4-aminobenzoic acid:2,2′-bipyridine cocrystal refers to 2 mol of 4-aminobenzoic acid and 1 mol of 2,2′-bipyridine.

the flexible degrees of freedom ξ^f determined at stage 2 (a "single-point DMAflex minimization"); we then applied clustering to the resulting structures, and then *fully* minimized the 25 most stable, unique ones among these.

The outer minimization in DMAflex[29] is solved to convergence with a simplex algorithm[70] because values of the $\partial U_{inter}/\partial \xi^f$ gradients are not readily available. An unfortunate consequence of this is that flexible torsion angles do not generally converge to an accuracy greater than 1−2°. The computational cost for a single-point and a full DMAflex minimization varied between 30 min and 2 h, and 1 and 3 days, respectively, depending on the system.

Results

Reproduction of Experimentally Determined Crystals. The overall reproduction quality of the experimental crystal structures with the models used in the three minimization stages is examined in Table 2. All three models reproduce the known crystal structures well, with errors in the coordination sphere[73] (rmsd$_{cs}$) not exceeding 0.5 Å. It is worth noting that the reproduction accuracy achieved with the models used in stages 2 and 3 is generally comparable.

The molecular conformations are also described well, with errors in the non-hydrogen torsion angles not exceeding 10° with the most accurate (DMAflex) model (Table S3 in Supporting Information). The largest discrepancies are observed for the rotation and pyramidalization of the amine group, although these could well be due to experimental uncertainties as well as problems[74] with the *ab initio* calculation of ΔE_{intra}.

Despite the similar overall reproduction quality with the three models, the predicted absolute and relative lattice energies are significantly different. The atomic charge model consistently underestimates the crystal energies compared with the multipole models, but the underestimation varies with the type of molecule, it being larger for hydrogen bonded systems. The stage 3 model predicts that the β polymorph of 4-aminobenzoic acid is 5.6 kJ mol^{-1} more stable than the α one, whereas stage

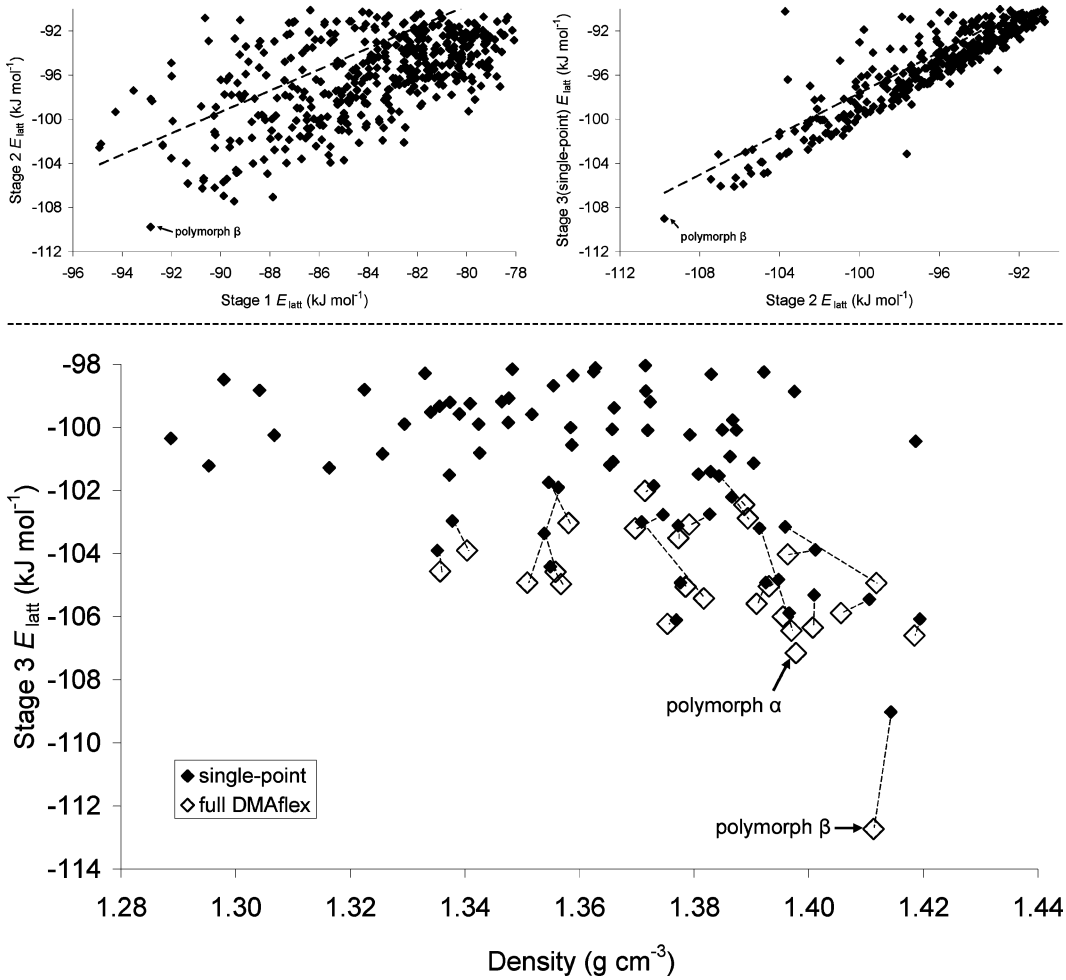


Figure 2. Crystal structure prediction of 4-aminobenzoic acid. The top two figures illustrate the extent of reranking in the three minimization stages; the dashed black lines are linear regressions to guide the eye. The bottom figure shows the lattice energy−density landscape obtained in the third minimization stage with the full and open symbols denoting single-point and full DMAflex minimizations, respectively. The minima obtained when the experimentally determined polymorphs are minimized with the same computational model are also indicated. The experimental $Z' = 2$ polymorph α was fully minimized with DMAflex for completeness, although it was not within the scope of the search. Carrying over the 900 most stable CrystalPredictor structures ($E_{latt,stage1} < -82$ kJ mol^{-1}) to stage 2 was sufficient to locate the majority of structures with single-point DMAflex energies (stage 3 in top right figure) up to 8 kJ mol^{-1} higher than the corresponding global minimum.

1 indicates almost no energy difference between the two polymorphs. The experimental enthalpy difference[75] is 2.5 kJ mol^{-1} at 86 °C and 5 kJ mol^{-1} at 25 °C. Although we cannot confidently extrapolate to 0 K and the modeling of form α only uses a fully ordered structure despite evidence of twinning, disorder, and possibly polytypism,[75] it is certain that the predicted stability difference using multipoles is more realistic compared with the negligible energy difference predicted with atomic charges. We note that the structural reproduction and relative stability of 4-aminobenzoic acid polymorphs using an earlier repulsion−dispersion parametrization[76] in part 1 was significantly worse (see also section S1 in Supporting Information).

With the stage 1 model, both cocrystals are predicted to be marginally (Table 2) less stable than the sum of lattice energies of the most stable polymorphs of the components. With the stage 3 model, the 4-aminobenzoic acid:2,2′-bipyridine cocrystal becomes thermodynamically favorable compared with the most stable single-component polymorphs. However, the 4-aminobenzoic acid:4-nitrophenylacetic acid cocrystal remains unstable, and its stability difference from the single-component crystals increases to 12.4 kJ mol^{-1}.

Crystal Structure Prediction. As shown in Figures 2−6, in all cases the experimentally observed structure was identified as a local lattice energy minimum. This is indicative of the effectiveness of the search algorithm despite the large number of minimization variables.[11] The experimental structures of 4-aminobenzoic acid (Figure 2) and 4-nitrophenylacetic acid (Figure 4) were both found with DMAflex at the corresponding global minima, while the experimental 2,2′-bipyridine crystal structure (Figure 3) was the second most stable structure, only 0.05 kJ mol^{-1} above the global minimum.

Figures 5 and 6 show the crystal structure prediction results for the cocrystals in a form that allows comparison with the pure component lattice energies. The stability of a cocrystal that contains n molecules of A and m molecules of B in the asymmetric unit was expressed as the lattice energy of the

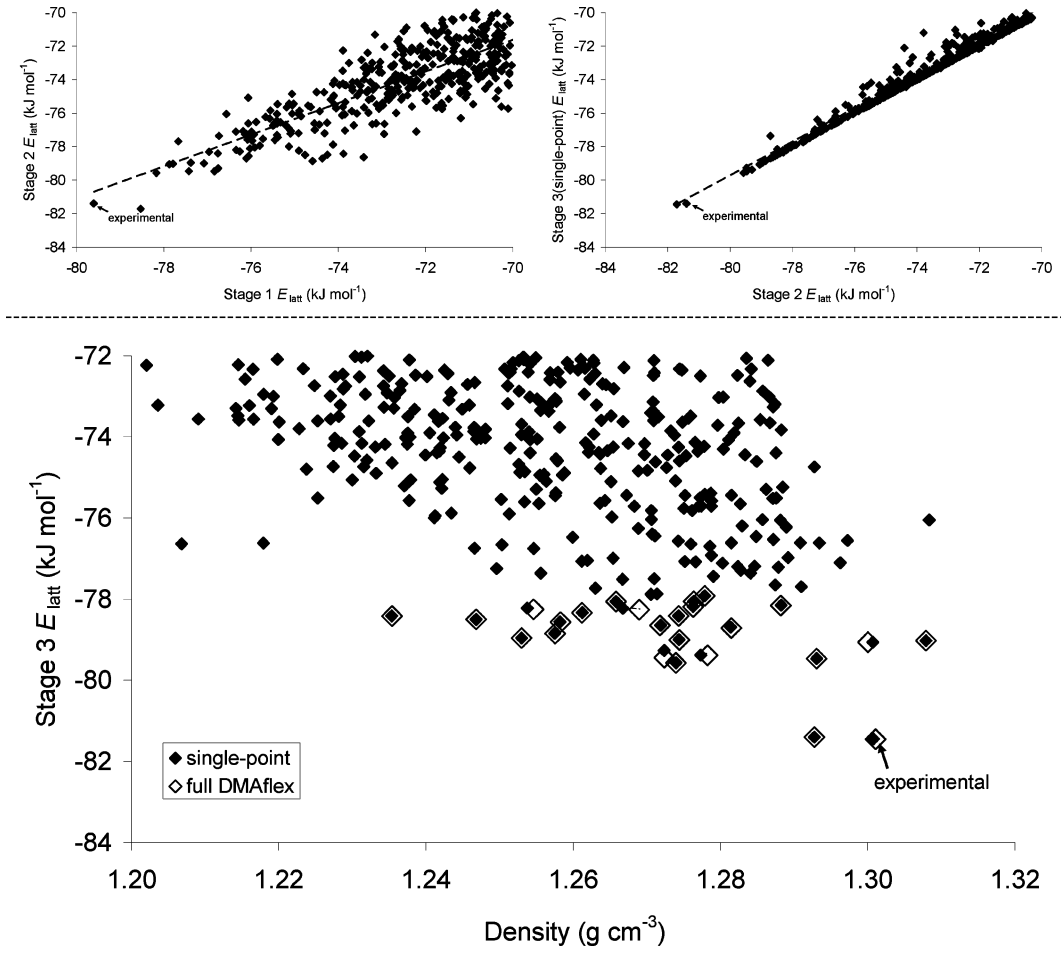


Figure 3. Crystal structure prediction of 2,2′-bipyridine. The top two figures illustrate the extent of reranking in the three minimization stages; the dashed black lines are linear regressions to guide the eye. The bottom figure shows the lattice energy−density landscape obtained in the third minimization stage with the full and open symbols denoting single-point and full DMAflex minimizations, respectively. The minimum obtained when the experimentally determined structure is minimized with the same computational model is also indicated. Carrying over the 700 most stable CrystalPredictor structures ($E_{latt,stage1} < -72$ kJ mol^{-1}) to stage 2 was sufficient to locate the majority of structures with single-point DMAflex energies (stage 3 in top right figure) up to 8 kJ mol^{-1} higher than the corresponding global minimum.

cocrystal (per mol of asymmetric unit) minus n times the predicted global lattice energy minimum of molecule A minus m times the predicted global lattice energy minimum of molecule B. Figure 5 shows that the experimentally determined 4-aminobenzoic acid:2,2′-bipyridine cocrystal was identified as the most stable packing arrangement in the search with 2:1 stoichiometry. However, there is one cocrystal with a 1:1 stoichiometry that is predicted to be more stable than the experimental form.[77] Hence, our predictions do not explain why 4-aminobenzoic acid:2,2′-bipyridine cocrystal adopts a 2:1 stoichiometry, a particularly challenging task given the small energy differences involved.[18] However, as several of the predicted cocrystal structures with both stoichiometries are more stable than the component crystals, the model predicts that cocrystallization is thermodynamically favorable. In contrast, Figure 6 shows that none of the hypothetical 4-aminobenzoic acid:4-nitrophenylacetic acid DMAflex minimized structures are energetically competitive with the DMAflex global lattice energy minima of the components, as all the relative lattice energies are positive. Moreover, the minimized experimental 4-aminobenzoic acid:4-nitrophenylacetic acid cocrystal structure was poorly ranked at 10.4 kJ mol^{-1} above the global minimum, with dozens of structures being more stable, one of which is very similar to the experimental form and 2 kJ mol^{-1} more stable.

The extent of reranking in the three minimization stages is molecule-dependent. The least affected is 2,2′-bipyridine (Figure 3), a molecule which lacks hydrogen bond donors and hence the ability to form strongly directional intermolecular interactions. In this case, the point charge model locates the known crystal structure at the global minimum despite its limitations in describing $\pi-\pi$ interactions,[78] and subsequent minimizations give little change. This contrasts with the energy reranking observed for all other systems, which can involve intermolecular hydrogen bonds. The experimental 4-aminobenzoic acid crystal structure was the fifth most stable in the CrystalPredictor search (Figure 2), approximately 2 kJ mol^{-1} above the global minimum that contains carboxylic dimers. The DMAflex model predicts that the experimental crystal, which contains COOH···NH$_2$ hydrogen bonds, clearly stands out as the unique, most stable packing arrangement separated by over 6 kJ mol^{-1} from the second most stable structure (Figure 2, Table 3). The

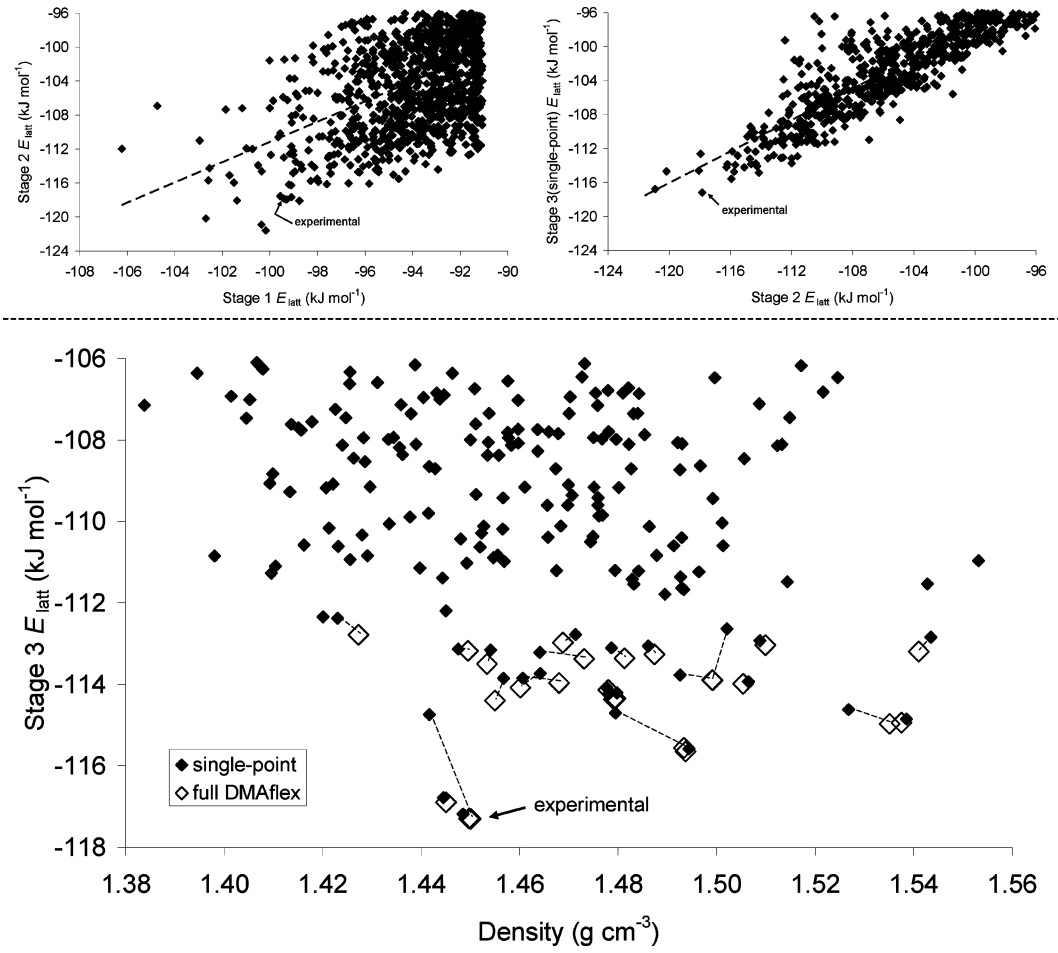


Figure 4. Crystal structure prediction of 4-nitrophenylacetic acid. The top two figures illustrate the extent of reranking in the three minimization stages; the dashed black lines are linear regressions to guide the eye. The bottom figure shows the lattice energy−density landscape obtained in the third minimization stage with the full and open symbols denoting single-point and full DMAflex minimizations, respectively. The structure obtained when the experimentally determined structure is minimized with the same computational model is also indicated. We carried over the 1400 most stable CrystalPredictor structures to stage 2, but the search is unlikely to be complete because of the degree of reranking.

improvements with the computational model are even more evident in the case of 4-nitrophenylacetic acid. Figure 4 shows that the known crystal structure was, at stage 1, poorly ranked and 6 kJ mol^{-1} above the global minimum that contained a catemer hydrogen bond motif. In contrast, the DMAflex model predicts that the experimental structure that contains carboxylic dimers corresponds to the most stable packing arrangement.

Figures 2−6 show that the improvement in the energy ranking of the experimental structures is greater on changing from point charges to distributed multipoles (from stage 1 to stage 2) than in going from analytically rotated to accurate multipoles (from stage 2 to stage 3). Nevertheless, assuming conformational transferability of the intermolecular electrostatic model is also too crude an approximation for the accuracy needed in crystal structure prediction. It is not surprising that the DMAflex-Quick error is greater for 4-nitrophenylacetic acid than 4-aminobenzoic acid, because its structures vary far more in conformation. Several of the hypothetical 2,2′-bipyridine cocrystal structures were also erroneously favored by DMAflex-Quick because they had significantly nonplanar 2,2′-bipyridine conformations, which led to significant errors arising from the analytical rotation of the multipoles of the planar gas-phase geometry. In all cases, the energy lowering caused by completing the DMAflex minimization (indicated by dashed lines in the crystal energy landscapes Figures 2−6) varies in the range of 0−3 kJ mol^{-1} per molecule in the asymmetric unit, which is significantly smaller than the energy changes in the other stages of this multistage improvement in the lattice energy model.

Analysis of Lattice Energy Landscapes. The four most stable predicted crystal structures in each search are given in Table 3 (and the 10 most stable in the Supporting Information), and we have examined the packing arrangement and dominant hydrogen bond motif of the 25 fully DMAflex minimized crystal structures. The majority of 4-aminobenzoic acid hypothetical crystal structures contain carboxylic acid hydrogen bonded dimers as observed in the experimentally determined $Z' = 2$ α form. The chains of OCOH···NH$_2$···OCOH hydrogen bonds found in the β polymorph (at the global minimum) are found only in two other structures. In contrast, all 22 unique, DMAflex-minimized 4-nitrophenylacetic acid crystal structures contain carboxylic acid dimers. In practically all the low energy 2,2′-

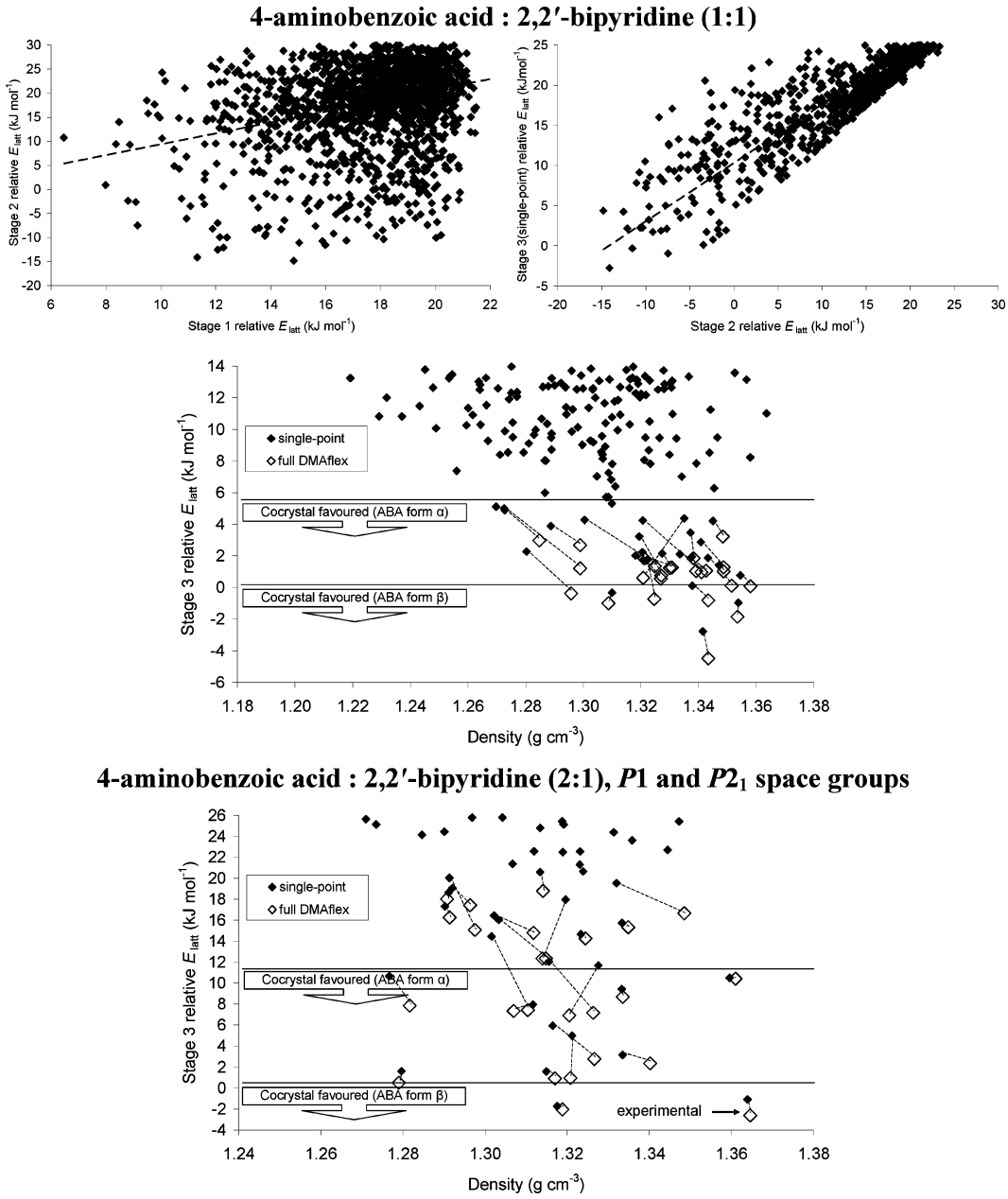


Figure 5. Crystal structure prediction of 4-aminobenzoic acid (ABA):2,2′-bipyridine cocrystal. All lattice energies are reported relative to the sum of the global lattice energy minima of the components with the same computational model. The top two figures illustrate the extent of reranking in the three minimization stages for 1:1 stoichiometry; the dashed black lines are linear regressions to guide the eye. The middle (1:1 stoichiometry) and bottom (2:1 stoichiometry) figures show the lattice energy−density landscape obtained in the third minimization stage with the full and open symbols denoting single-point and full DMAflex minimizations, respectively. The structure obtained when the experimentally determined structure is minimized with the same computational model is also indicated. We carried over the 2000 (1:1) and 750 (2:1) most stable CrystalPredictor structures to stage 2, but the search is unlikely to be complete because of the degree of reranking.

bipyridine structures, the pyridine rings are coplanar, showing that deforming the molecule from this gas-phase, planar conformation does not improve the packing. The hypothetical structure at the global minimum exhibits a tilted edge-to-face (T-shaped) arrangement of the molecular planes and better overlap in the $\pi-\pi$ stacking than the slightly less stable experimental structure.

In the majority of the low energy 4-aminobenzoic acid: 2,2′-bipyridine cocrystal structures, all carboxylic acid protons are hydrogen bonded to pyridine nitrogen atoms; in the case of 1:1 stoichiometry, the second pyridine nitrogen competes with the carboxylic carbonyl for the amine hydrogen atoms. Hence, the computational model correctly predicts that the presence of the pyridine groups disrupts the carboxylic dimers and $COOH\cdots NH_2$ hydrogen bonds of 4-aminobenzoic acid that is consistent with the strong tendency of pyridyl nitrogen atoms to hydrogen bond to carboxyl acid donors.[79−81] In the majority of the predicted stable 1:1 cocrystals, 2,2′-bipyridine adopts a nonplanar conformation, whereas it is planar and on an inversion center in several of the 2:1 cocrystals including the experimental form. The rest of the 2:1 cocrystal structures also have both

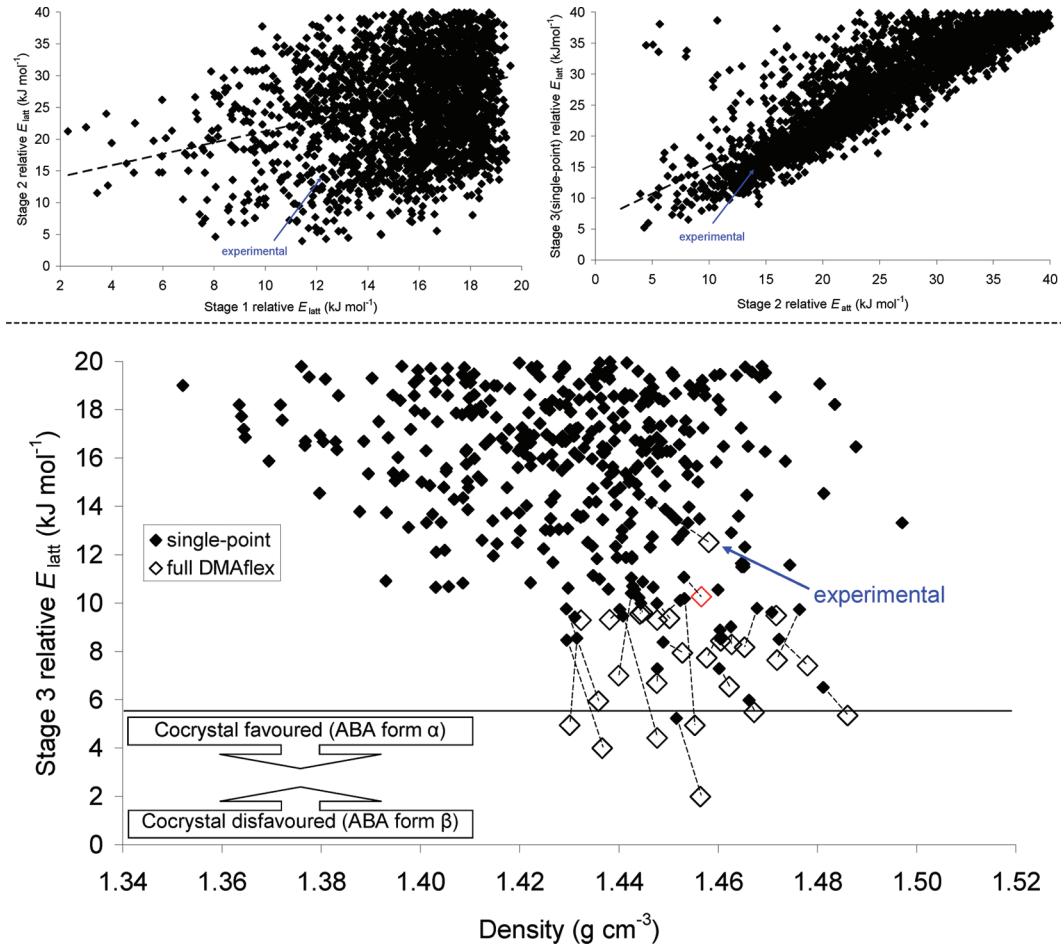


Figure 6. Crystal structure prediction of 4-aminobenzoic acid (ABA):4-nitrophenylacetic acid (1:1) cocrystals. All crystal energies are reported relative to the sum of the global lattice energy minima of the components with the same computational model. The top two figures illustrate the extent of reranking in the three minimization stages; the dashed black lines are linear regressions to guide the eye. The bottom figure shows the lattice energy−density landscape obtained in the third minimization stage with the full and open symbols denoting single-point and full DMAflex minimizations respectively. The structure obtained when the experimentally determined structure is minimized with the same computational model is also indicated; the red diamond shows a hypothetical structure which is similar to the experimental ($rmsd_{cs}$ = 0.34 Å). We carried over the 3000 most stable CrystalPredictor structures to stage 2, but the search is unlikely to be complete because of the degree of reranking.

pyridine nitrogen atoms hydrogen bonded to two carboxylic protons and hence may transform to a slightly more stable symmetric packing if thermal motion were included.

The 4-aminobenzoic acid:4-nitrophenylacetic acid lattice energy landscape exhibits a plurality of hydrogen bonding patterns. A large proportion of the most stable structures contain 4-aminobenzoic acid homodimers. In these structures, the carboxylic acid hydrogen atom of 4-nitrophenylacetic acid is hydrogen bonded to the amine nitrogen; hydrogen bonds are also formed between the amine hydrogen atoms and the carbonyl and nitro groups of 4-nitrophenylacetic acid. The next most dominant hydrogen bond motif is a carboxylic acid heterodimer, with the most stable such structure being found only 3 kJ mol^{-1} above the global minimum. This motif appears in the experimental structure and in 33% of all crystal structures containing two distinct carboxylic acids.[82] There are a limited number of low energy structures in which all hydrogen bond acceptors and donors are associated in chains forming intricate 3-D patterns. There is also a noticeable absence of stable structures containing 4-nitrophenylacetic acid homodimers. The CSD contains only five cocrystals of 4-aminobenzoic acid with carboxylic acids, three of which (including the cocrystal with 4-nitrophenylacetic acid) contain carboxylic acid heterodimers. In the cocrystal with nicotinic acid and the ternary cocrystal with 2,4,6-trinitrobenzoic acid and 1,3,5-trinitrobenzene, 4-aminobenzoic acid does form the homodimers seen in the low energy predicted structures (Table 3). The stronger tendency of 4-aminobenzoic acid to associate in $R_2^2(8)$ carboxylic acid dimers compared with 4-nitrophenylacetic acid is consistent with the differences in electrostatic potential and overall spatial arrangement of the carboxylic groups (see Figure S4 in Supporting Information).

Discussion

Based on our results, the structure generation in stage 1 appears to be sufficiently extensive to locate the experimental crystal structures, including the challenging 4-aminobenzoic acid:2,2′-bipyridine cocrystal that required a search with three crystallographically independent molecules. However, the energy rank of the experimental structures at stage 1 is

Table 3. Most Stable DMAflex Fully-Minimized Crystal Structures[a]

rank, space group[b]	lattice energy partitioning[c] (kJ mol^{-1}) E_{latt}	ΔE_{intra}	conventional lattice cell[d] a (Å)	b (Å)	c (Å)	α, β, γ (deg)	density (g cm^{-3})	packing/dominant hydrogen bond motif
			Single-Component Crystals					
			4-Aminobenzoic Acid					
1, $P2_1/c$	**−112.74**	**2.82**	**6.336**	**8.445**	**12.246**	**100.0**	**1.411**	**OCOH···NH2···OCOH four member rings**
2, $P2_1/c$	−106.60	0.39	3.838	15.271	10.966	92.3	1.419	carboxylic dimers
3, $P2_1$	−106.44	6.12	4.009	14.489	5.661	97.5	1.397	OCOH···NH2···OCOH chains
4, $P2_1/c$	−106.35	0.50	3.707	11.100	15.979	98.5	1.401	carboxylic dimers
			2,2′-Bipyridine					
1, $P2_1/c$	−81.46	0.05	6.017	11.249	12.951	114.6	1.301	T-arrangement, π-π stacking
2, $P2_1/c$, $Z' = 1/2$	**−81.41**	**0.00**	**5.680**	**6.180**	**12.289**	**111.5**	**1.293**	**limited π-π stacking**
3, $Pna2_1$	−79.57	0.00	13.701	5.061	11.744	–	1.274	limited π-π stacking
4, $P2_1/c$, $Z' = 1/2$	−79.47	0.00	3.787	9.735	10.902	93.4	1.293	sheets, limited π-π stacking
			4-Nitrophenylacetic Acid					
1, *Pbca*	−117.30	1.02	15.178	6.822	16.027	–	1.450	carboxylic dimers
2, $P2_1/c$	−116.90	0.82	8.021	15.254	6.805	90.4	1.445	carboxylic dimers
3, $P2_1/c$	−115.65	1.22	6.511	4.793	26.093	98.4	1.494	carboxylic dimers
4, $P2_1/c$	−114.97	2.30	4.755	15.258	10.923	98.5	1.535	carboxylic dimers
			Cocrystals					
			4-Aminobenzoic Acid:2,2′-Bipyridine (1:1 stoichiometry)					
1, $P2_1/c$	−198.68	2.55, 1.34	3.830	13.066	29.100	95.2	1.343	OCOH···N (pyr), NH2···OCOH
2, $P2_1$	−196.04	2.00, 1.35	3.955	12.623	14.512	96.5	1.354	OCOH···N (pyr), NH2···OCOH
3, $P2_1/c$	−195.20	6.66, 1.31	7.041	30.514	7.341	109.6	1.311	OCOH···N (pyr), NH2···OCOH
4, $P2_1$	−195.01	0.21, 1.34	3.856	13.283	14.217	95.2	1.343	OCOH···N (pyr), NH2···OCOH
			4-Aminobenzoic Acid:2,2′-Bipyridine (2:1 stoichiometry)					
1, $P2_1/c$, $Z' = 3/2$	**−309.55**	**1.32, 0.00**	**16.173**	**4.282**	**15.732**	**106.0**	**1.365**	**OCOH···N (pyr), NH2···OCOH·NH2··NH2**
2, $P2_1/c$, $Z' = 3/2$	−308.98	2.12, 0.00	3.904	13.082	21.311	95.2	1.319	OCOH···N (pyr) NH2···OCOH
3, $P2_1/c$, $Z' = 3/2$	−306.44	4.17, 0.00	8.430	4.689	28.583	98.4	1.279	OCOH···N (pyr), NH2···OCOH·NH2··NH2
4, $P2_1/c$, $Z' = 3/2$	−306.02	1.25, 0.00	3.778	21.677	13.509	101.1	1.317	OCOH···N (pyr), NH2···OCOH
			4-Aminobenzoic Acid:4-Nitrophenylacetic Acid (1:1)[e]					
1, $P2_1/c$	−228.05	1.98, 2.72	7.318	15.287	13.954	111.6	1.456	4-aminobenzoic carboxylic homodimers, OCOH(4-nit)···NH2, NH2···O2N
2, $P\bar{1}$	−226.04	3.99, 1.43	5.102	12.082	13.119	105.8, 101.2 101.4	1.437	4-aminobenzoic carboxylic homodimers, OCOH(4-nit)···NH2, NH2···O2N
3, $P2_1/c$	−225.62	4.42, 2.50	13.166	7.471	17.660	122.8	1.448	4-aminobenzoic carboxylic homodimers, OCOH(4-nit)···NH2, HOCO(4-amn,4-nit)···H2N
4, $P2_1/c$	−225.10	4.93, 1.89	4.496	25.887	12.482	90.26	1.455	4-aminobenzoic carboxylic homodimers, OCOH(4-nit)···NH2, HOCO(4-nit)···H2N, NO2···H2N

[a] Rows in Bold Correspond to the Experimental Structures. [b] Z' is given when 2,2′-bipyridine lies at a special position. [c] The HF/6−31G(d,p) intramolecular energy ΔE_{intra} is shown separately for each molecule in the asymmetric unit in the order shown in the system title. [d] Only cell angles not constrained by symmetry given. [e] The DMAflex-minimized experimental crystal lies 10.4 kJ mol^{-1} above the global minimum.

generally poor. The only exception is 2,2′-bipyridine, the only molecule in this study that lacks hydrogen bond donors. The distributed multipole model drastically changes the energy ordering of the hypothetical structures, resulting in a better ranking of the experimental structures, often without any significant structural changes (cf. Table 2). This emphasizes that the reproduction of known crystal structures by lattice energy minimization is a necessary but by no means sufficient condition for the accuracy of the computational model for the lattice energy.

Figures 2−6 show that the improvement seen by replacing point charges with analytically rotated multipoles is larger than the improvement gained by the use of conformationally dependent multipoles. Hence, the use of DMAflex-Quick followed by a single-point DMAflex minimization is a worthwhile procedure to efficiently locate promising hypothetical crystal structures to perform computationally demanding calculations. In principle, stage 1 structures should be passed in batches of increasing energy to later stages until we generate no additional, distinct minima within a predefined range from the global minimum in the final stage. In practice, if the target was to identify the majority of local minima within 8 and 12 kJ mol^{-1} from the global minimum in the final stage for single- and multicomponent crystals respectively, the objective was achieved only for 4-aminobenzoic acid and 2,2′-bipyridine (Figure 2−3), where we were able to minimize sufficiently large batches of stage 1 structures. The overall energy reranking between stage 1 and the fully DMAflex minimized structures is so severe that all other searches were terminated because of limited computer resources (Figure 4−6). Hence, carrying over several thousand crystal structures to subsequent stages does not guar-

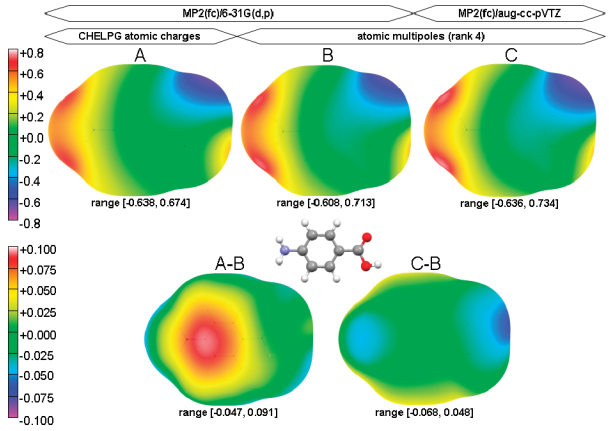


Figure 7. Electrostatic potential (in Volts) (top row) of the HF/6−31G(d,p) globally optimized molecular conformation of 4-aminobenzoic acid on twice the van der Waals radii surface (van der Waals radius for polar hydrogen atoms set to 1 au; radii for other atoms taken from Bondi[94]) computed with the MP2(fc)/6−31G(d,p) atomic charges fitted to the electrostatic potential[95] (A); the DMA[63] atomic multipole moments up to rank 4 derived from the same MP2(fc)/6−31G(d,p) charge density (B) and derived from the MP2(fc)/aug-cc-pVTZ isolated-molecule charge density (C). The bottom row shows electrostatic potential differences A-B and C−B. The maps are displayed using ORIENT[96] for the molecular orientation shown in ball-and-stick.

antee that all low-energy minima are found, even if we assume that the stage 1 search is complete despite its stochastic character.

The reranking from stage 1 to stage 2 is consistent with the large difference in the electrostatic potential around 4-aminobenzoic acid (Figure 7), calculated from an atomic charge and atomic multipole representation of the MP2(fc)/6−31G(d,p) charge density. However, Figure 7 (and also Figure S1 and Table S1 in Supporting Information) shows that increasing the basis set can also have a significant effect, which decreases the predicted stability of 4-aminobenzoic acid structures based on COOH···NH_2 hydrogen bonds relative to those that contain carboxylic dimers. The charge density quality has a smaller impact on the relative stability of 4-nitrophenylacetic acid structures, presumably because all of these contain carboxylic dimers. Thus, the monomer charge-density quality should be carefully considered, particularly when the lower energy structures differ in the hydrogen bonding motif.

This study has relied on empirical, transferable models for the repulsion−dispersion interactions. Part I of this paper showed that the FIT[76] potential severely overestimates the stability difference of 4-aminobenzoic acid polymorphs. This is improved (Table 2) by using the more recent Williams[61] parametrization in this paper, which contains more atom types. However, the Williams potential in conjunction[33,61] with multipole moments underestimates hydrogen bond lengths. This underestimation is particularly severe when the crystal contains COOH···N(pyr) close contacts and hence the modeling of 4-aminobenzoic acid:2,2′-bipyridine cocrystals required a modification of the Williams carboxylic proton - pyridine nitrogen repulsion potential as discussed in Supporting Information. Hence, there is also a noticeable dependence of the computed lattice energy landscape on the repulsion−dispersion parametrization, even when the electrostatic contributions due to strong hydrogen bonding dominate the lattice energy.

Empirically fitted potentials partially absorb the errors in modeling the electrostatic forces and intermolecular polarization effects. Nevertheless, induction is anisotropic and depends strongly on the molecular association in the lattice, which is consistent with the significant difference that the explicit modeling of induction[32] can make in the relative stability of polymorphs with different hydrogen bonding motifs.[53] The hydrogen bonding motif in cocrystals will generally be different from the motif in the component crystal structures; hence, neglecting explicit induction will not cancel out in predicting the thermodynamic grounds for cocrystal formation. The development of molecule-specific models for the repulsion−dispersion[30,72,83] and polarization energies[32] are worthwhile improvements that will complement the accurate representation of the isolated-molecule charge density using distributed multipoles.[33] An advantage of the new search methodology is that it can use such theoretically justified models[30] for calculating the lattice energy at the final stage. An alternative or subsequent improvement in the estimated relative thermodynamic stability would be to investigate the implicit assumption in this study that the differences in vibrational contributions to the free energy are small compared with the lattice energy differences. This assumption is questionable given the density variation of up to 10% in the low energy structures (Figures 2, 4−6).

Despite these limitations, four out of five searches resulted in the experimental crystal structures being very close to the global minimum. The exception is 4-aminobenzoic acid:4-nitrophenylacetic acid cocrystal where the search produced several structures based on 4-aminobenzoic acid homodimers that are more stable than the experimental form. Although this may reflect errors in the relative thermodynamic stability model, it could be a kinetic factor, as self-association of 4-aminobenzoic acid in homodimers has been observed in cocrystals with carboxylic acids.[84] Thus, we cannot exclude the possibility that a homodimer-based polymorph could be found in the future, despite most sets of polymorphic cocrystals[3] in the CSD only differing in the spatial arrangement of identical hydrogen bonded motifs,[85] although exceptions to this generalization do exist.[86] This ability of a computational search to generate alternative hydrogen-bonding motifs makes it a complement to the synthon approach to the design of cocrystals,[87] which relies on the persistence of common hydrogen bonding motifs. On the other hand, cocrystallization of 2,2′-bipyridine with 4-aminobenzoic acid is predicted to almost certainly lead to the disruption of the carboxylic dimers and COOH···NH_2 hydrogen bonds and the formation of carboxylic···N_{arom} hydrogen bonds, in accord with the synthon approach and experimental findings.[80,81]

The proposed methodology allows the efficient identification of stable minima on complex lattice energy surfaces due to the presence of multiple, flexible molecules (or ions) in the asymmetric unit. This knowledge of the lattice energy landscape may prove useful in solving the crystal structures

from powder X-ray data,[88,89] which is often necessary as many cocrystals are synthesized by solid-state grinding. Moreover, the lattice energy landscape can enhance the understanding of the range of possible crystal structures and help to anticipate possible crystallization problems such as disorder and polymorphism.[90] However, the crystallization of cocrystals is also subject to kinetic effects,[91] and hence it is not clear whether all molecular associations generated in the search are really in competition in the nucleation stage of practical cocrystallization processes.

Conclusions

This study presents an extensive, yet efficient, crystal structure prediction methodology to locate thermodynamically stable crystal structures on complex lattice energy surfaces that are typical of multicomponent, flexible systems using conformationally dependent multipole moments and quantum mechanically computed conformational energies. All searches for single- and two-component crystals successfully located the experimentally determined crystal structure and in four out of five cases this lay at, or very close to, the global lattice energy minimum. The ability to predict the structure and thermodynamic stability of cocrystals allows the comparison of the stability of possible cocrystal packing arrangements relative to the single-component crystals. Such applications can greatly enhance our understanding of cocrystal formation, which, at present, relies on a qualitative hierarchy of a small set of chemically intuitive synthons while ignoring the much larger set of secondary, but collectively significant, intermolecular interactions.[92] However, as in part 1 of this paper, it is clear that the energetic advantage of cocrystal formation is small[18] compared with the errors in the model to compute the lattice energies. Nevertheless, this multistage search methodology can be extended to use even more accurate and computationally intensive lattice energy minimization methods[27,93] on a small set of crystal structures that have been found to be energetically competitive.

Acknowledgment. Funding to the Molecular Systems Engineering group from the Engineering and Physical Sciences Research Council (EPSRC) of the UK [EP/E016340] is gratefully acknowledged. Part of this research was carried out within the auspices of the CPOSS project (http://www.cposs.org.uk), funded by the Basic Technology Programme of the Research Councils UK. G.S.K. is acknowledged for funding NI. The assistance of Dr. L. S. Price in preparing this manuscript is gratefully acknowledged.

Supporting Information Available: Sensitivity of the predictions to the model for the intermolecular forces, modification of the Williams repulsion−dispersion parametrization for the 4-aminobenzoic acid:2,2′-bipyridine cocrystal structure prediction, reproduction of flexible torsion angles and hydrogen bond geometries during lattice energy minimization of the experimentally determined crystal structures and tables with the 10 most stable predicted crystal structures for each search. All stage three crystal structures are stored on CCLRC e-Science Centre data portal and are available from the authors on request. This information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Almarsson, O.; Zaworotko, M. J. *Chem. Commun.* **2004**, *17*, 1889–1896.

(2) Aakeroy, C. B.; Salmon, D. J. *CrystEngComm* **2005**, *7*, 439–448.

(3) Vishweshwar, P.; McMahon, J. A.; Bis, J. A.; Zaworotko, M. J. *J. Pharm. Sci.* **2006**, *95*, 499–516.

(4) Etter, M. C. *Acc. Chem. Res.* **1990**, *23*, 120–126.

(5) Issa, N.; Karamertzanis, P. G.; Welch, G. W. A.; Price, S. L. *Cryst. Growth Des.* **2009**, *9*, 442–453.

(6) Day, G. M.; Cooper, T. G.; Cruz Cabeza, A. J.; Hejczyk, K. E.; Ammon, H. L.; Boerrigter, S. X. M.; Tan, J.; Della Valle, R. G.; Venuti, E.; Jose, J.; Gadre, S. R.; Desiraju, G. R.; Thakur, T. S.; van Eijck, B. P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Neumann, M.; Leusen, F. J. J.; Kendrick, J.; Price, S. L.; Misquitta, A. J.; Karamertzanis, P. G.; Welch, G. W. A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; van de Streek, J.; Wolf, A.; Schweizer, B Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem. 2009, in press.

(7) Leusen, F. J. J. *Cryst. Growth Des.* **2003**, *3*, 189–192.

(8) Karamertzanis, P. G.; Price, S. L. *J. Phys. Chem. B* **2005**, *109*, 17134–17150.

(9) Cabeza, A. J. C.; Day, G. M.; Motherwell, W. D. S.; Jones, W. *J. Am. Chem. Soc.* **2006**, *128*, 14466–14467.

(10) Hulme, A. T.; Price, S. L. *J. Chem. Theory Comput.* **2007**, *3*, 1597–1608.

(11) van Eijck, B. P.; Kroon, J. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2000**, *56*, 535–542.

(12) Collins, A.; Cooper, R. I.; Watkin, D. J. *J. Appl. Crystallogr.* **2006**, *39*, 842–849.

(13) Collins, A. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2006**, *62*, 897–911.

(14) Desiraju, G. R. *CrystEngComm* **2007**, *9*, 91–92.

(15) Gavezzotti, A. *CrystEngComm* **2008**, *10*, 389–398.

(16) Mohamed, S.; Barnett, S. A.; Tocher, D. A.; Shankland, K.; Leech, C. K.; Price, S. L. *CrystEngComm* **2008**, *10*, 399–404.

(17) Steed, J. W. *CrystEngComm* **2003**, *5*, 169–179.

(18) Cruz-Cabeza, A. J.; Day, G. M.; Jones, W. *Chem. Eur. J.* **2008**, *14*, 8830–8836.

(19) Lommerse, J. P. M.; Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Mooij, W. T. M.; Price, S. L.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2000**, *56*, 697–714.

(20) Motherwell, W. D. S.; Ammon, H. L.; Dunitz, J. D.; Dzyabchenko, A.; Erk, P.; Gavezzotti, A.; Hofmann, D. W. M.; Leusen, F. J. J.; Lommerse, J. P. M.; Mooij, W. T. M.; Price, S. L.; Scheraga, H.; Schweizer, B.; Schmidt, M. U.; van Eijck, B. P.; Verwer, P.; Williams, D. E. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2002**, *58*, 647–661.

(21) Day, G. M.; Motherwell, W. D. S.; Ammon, H. L.; Boerrigter, S. X. M.; Della Valle, R. G.; Venuti, E.; Dzyabchenko, A.;

Dunitz, J. D.; Schweizer, B.; van Eijck, B. P.; Erk, P.; Facelli, J. C.; Bazterra, V. E.; Ferraro, M. B.; Hofmann, D. W. M.; Leusen, F. J. J.; Liang, C.; Pantelides, C. C.; Karamertzanis, P. G.; Price, S. L.; Lewis, T. C.; Nowell, H.; Torrisi, A.; Scheraga, H. A.; Arnautova, Y. A.; Schmidt, M. U.; Verwer, P. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2005**, *61*, 511–527.

(22) Rodriguez-Spong, B.; Price, C. P.; Jayasankar, A.; Matzger, A. J.; Rodriguez-Hornedo, N. *Adv. Drug Delivery Rev.* **2004**, *56*, 241–274.

(23) Day, G. M.; Price, S. L.; Leslie, M. *J. Phys. Chem. B* **2003**, *107*, 10919–10933.

(24) van Eijck, B. P. *J. Comput. Chem.* **2001**, *22*, 816–826.

(25) van Eijck, B. P.; Mooij, W. T. M.; Kroon, J. *J. Phys. Chem. B* **2001**, *105*, 10573–10578.

(26) Aubrey-Medendorp, C.; Swadley, M. J.; Li, T. L. *Pharm. Res.* **2008**, *25*, 953–959.

(27) Neumann, M. A.; Perrin, M. A. *J. Phys. Chem. B* **2005**, *109*, 15531–15541.

(28) Neumann, M. A.; Leusen, F. J. J.; Kendrick, J. *Angew. Chem., Int. Ed.* **2008**, *47*, 2427–2430.

(29) Karamertzanis, P. G.; Price, S. L. *J. Chem. Theory Comput.* **2006**, *2*, 1184–1199.

(30) Misquitta, A. J.; Welch, G. W. A.; Stone, A. J.; Price, S. L. *Chem. Phys. Lett.* **2008**, *456*, 105–109.

(31) van Eijck, B. P.; Mooij, W. T. M.; Kroon, J. *J. Comput. Chem.* **2001**, *22*, 805–815.

(32) Welch, G. W. A.; Karamertzanis, P. G.; Misquitta, A. J.; Stone, A. J.; Price, S. L. *J. Chem. Theory Comput.* **2008**, *4*, 522–532.

(33) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.* **2005**, *5*, 1023–1033.

(34) Cabeza, A. J. C.; Day, G. M.; Motherwell, W. D. S.; Jones, W. *Cryst. Growth Des.* **2006**, *6*, 1858–1866.

(35) Koch, U.; Popelier, P. L. A.; Stone, A. J. *Chem. Phys. Lett.* **1995**, *238*, 253–260.

(36) Koch, U.; Stone, A. J *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 1701–1708.

(37) Mannfors, B. E.; Mirkin, N. G.; Palmo, K.; Krimm, S. *J. Phys. Chem. A* **2003**, *107*, 1825–1832.

(38) van Eijck, B. P. *J. Comput. Chem.* **2002**, *23*, 456–462.

(39) Nowell, H.; Price, S. L. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2005**, *61*, 558–568.

(40) Day, G. M.; Motherwell, W. D. S.; Jones, W. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1693–1704.

(41) Karamertzanis, P. G.; Pantelides, C. C. *J. Comput. Chem.* **2005**, *26*, 304–324.

(42) Karamertzanis, P. G.; Pantelides, C. C. *Mol. Phys.* **2007**, *105*, 273–291.

(43) Lai, T. F.; Marsh, R. E. *Acta Crystallogr.* **1967**, *22*, 885–893.

(44) Alleaume, M.; Salascim, G.; Decap, J. *C. R. Hebd. Seances Acad. Sci. Ser. C* **1966**, *262*, 416–417.

(45) Gracin, S.; Fischer, A. *Acta Crystallogr. Sect. E: Struct. Rep. Online* **2005**, *61*, O1242−O1244.

(46) Bowers, J. R.; Hopkins, G. W.; Yap, G. P. A.; Wheeler, K. A. *Cryst. Growth Des.* **2005**, *5*, 727–736.

(47) Price, S. L. *Phys. Chem. Chem. Phys.* **2008**, *10*, 1996–2009.

(48) Aakeroy, C. B.; Beatty, A. M.; Helfrich, B. A.; Nieuwenhuyzen, M. *Cryst. Growth Des.* **2003**, *3*, 159–165.

(49) Allen, F. H.; Harris, S. E.; Taylor, R. *J. Comput. -Aided Mol. Des.* **1996**, *10*, 247–254.

(50) Brameld, K. A.; Kuhn, B.; Reuter, D. C.; Stahl, M. *J. Chem. Inf. Model.* **2008**, *48*, 1–24.

(51) Weng, Z. F.; Motherwell, W. D. S.; Allen, F. H.; Cole, J. M. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2008**, *64*, 348–362.

(52) Cooper, T. G.; Jones, W.; Motherwell, W. D. S.; Day, G. M. *CrystEngComm* **2007**, *9*, 595–602.

(53) Karamertzanis, P. G.; Day, G. M.; Welch, G. W. A.; Kendrick, J.; Leusen, F. J. J.; Neumann, M. A.; Price, S. L. *J. Chem. Phys.* **2008**, *128*, art-244708.

(54) Allen, F. H. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **2002**, *58*, 380–388.

(55) Sobol', I. M. *USSR Comp. Math. Math+* **1967**, *7*, 86–112.

(56) PLATON A Multipurpose Crystallographic Tool, Spek, A. L. Utrecht University: Utrecht, The Netherlands, 2003.

(57) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O. Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Gaussian Inc., Wallingford, CT, 2003.

(58) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

(59) Gordon, M. S.; Schmidt, M. W. Advances in electronic structure theory: GAMESS a decade later. In *Theory and Applications of Computational Chemistry: the first forty years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005; pp 1167−1189.

(60) Brodersen, S.; Wilke, S.; Leusen, F. J. J.; Engel, G. *Phys. Chem. Chem. Phys.* **2003**, *5*, 4923–4931.

(61) Williams, D. E. *J. Comput. Chem.* **2001**, *22*, 1154–1166.

(62) Ewald, P. *Ann. Phys.* **1921**, *64*, 253–287.

(63) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128–1132.

(64) Stone, A. J. *The Theory of Intermolecular Forces*, 1st ed.; Clarendon Press: Oxford, 1996.

(65) This multipole orientation method was checked against the spherical tensor based manipulations in ORIENT.

(66) Willock, D. J.; Price, S. L.; Leslie, M.; Catlow, C. R. A. *J. Comput. Chem.* **1995**, *16*, 628–647.

(67) Welch, G. W. A.; Karamertzanis, P. G.; Price, S. L; Leslie, M. DMACRYS, version 1.05, is a substantial revision of DMAREL.http://www.chem.ucl.ac.uk/cposs/dmacrys/2008.

(68) Price, S. L.; Stone, A. J *J. Chem. Soc., Faraday Trans.* **1992**, *88*, 1755–1763.

(69) Polito, M.; D'Oria, E.; Maini, L.; Karamertzanis, P. G.; Grepioni, F.; Braga, D.; Price, S. L. *CrystEngComm* **2008**, *10*, 1848–1854.

(70) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in FORTRAN 90*; Cambridge University Press: Cambridge, UK, 1996; Vol. 2.

(71) Day, G. M.; Price, S. L. *J. Am. Chem. Soc.* **2003**, *125*, 16434–16443.

(72) Mooij, W. T. M.; van Eijck, B. P.; Kroon, J. *J. Phys. Chem. A* **1999**, *103*, 9883–9890.

(73) Chisholm, J. A.; Motherwell, S. *J. Appl. Crystallogr.* **2005**, *38*, 228–231.

(74) Wang, S. Y.; Schaefer, H. F. *J. Chem. Phys.* **2006**, *124*, art-044303.

(75) Gracin, S.; Rasmuson, A. C. *Cryst. Growth Des.* **2004**, *4*, 1013–1023.

(76) Coombes, D. S.; Price, S. L.; Willock, D. J.; Leslie, M. *J. Phys. Chem.* **1996**, *100*, 7352–7360.

(77) The lowest energy packing configuration that consists of two molecules of 4-aminobenzoic acid per one molecule of 2,2′-bipyridine corresponds to the 1:1 cocrystal global minimum crystallizing concomitantly with the 4-aminobenzoic acid global minimum. This packing configuration is 4.5 and 1.9 kJ mol^{-1} more stable compared with the three molecules crystallizing on their own and the global 2:1 cocrystal global minimum, respectively.

(78) Price, S. L *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 2997–3008.

(79) Steiner, T. *Acta Crystallogr., Sect. B* **2001**, *57*, 103–106.

(80) Aakeroy, C. B.; Hussain, I.; Forbes, S.; Desper, J. *CrystEngComm* **2007**, *9*, 46–54.

(81) Shattock, T. R.; Arora, K. K.; Vishweshwar, P.; Zaworotko, M. J. *Cryst. Growth Des.* **2008**, *8*, 4533–4545.

(82) Allen, F. H.; Motherwell, W. D. S.; Raithby, P. R.; Shields, G. P.; Taylor, R *New J. Chem.* **1999**, *23*, 25–34.

(83) Mooij, W. T. M.; van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G.C. M.; van Eijck, B. P *J. Phys. Chem. A* **1999**, *103*, 9872–9882.

(84) Jebas, S. R.; Balasubramanian, T. *Acta Crystallogr., Sect. E: Struct. Rep. Online* **2006**, *62*, o5621−o5622.

(85) Vishweshwar, P.; McMahon, J. A.; Zaworotko, M. J. Crystal Engineering of Pharmaceutical Co-Crystals. In *Frontiers in Crystal Engineering*, Tiekink, E. R. T., Vittal, J. J., Eds.; John Wiley & Sons Ltd.: Chichester, UK, 2006; pp 25−49.

(86) Sreekanth, B. R.; Vishweshwar, P.; Vyas, K. *Chem. Commun.* **2007**, 2375–2377.

(87) Friscic, T.; Jones, W. *Faraday Discuss.* **2007**, *136*, 167–178.

(88) Tremayne, M.; Grice, L.; Pyatt, J. C.; Seaton, C. C.; Kariuki, B. M.; Tsui, H. H. Y.; Price, S. L.; Cherryman, J. C. *J. Am. Chem. Soc.* **2004**, *126*, 7071–7081.

(89) Schmidt, M. U.; Buchsbaum, C. *Z. Kristallogr.* **2008**, *223*, 418–423.

(90) Copley, R. C. B.; Barnett, S. A.; Karamertzanis, P. G.; Harris, K. D. M.; Kariuki, B. M.; Xu, M. C.; Nickels, E. A.; Lancaster, R. W.; Price, S. L. *Cryst. Growth Des.* **2008**, *8*, 3474–3481.

(91) Desiraju, G. R. *Nat. Mater.* **2002**, *1*, 77–79.

(92) Aakeroy, C. B.; Desper, J.; Elisabeth, E.; Helfrich, B. A.; Levin, B.; Urbina, J. F. *Z. Kristallogr.* **2005**, *220*, 325–332.

(93) Civalleri, B.; Zicovich-Wilson, C. M.; Valenzano, L.; Ugliengo, P. *CrystEngComm* **2008**, *10*, 405–410.

(94) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441–451.

(95) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361–373.

(96) Stone, A. J.; Dullweber, A.; Engkvist, O.; Fraschini, E. Hodges, M. P.; Meredith, A. W.; Nutt, D. R.; Popelier, P. L. A.; Wales, D. J. ORIENT: a program for studying interactions between molecules, version 4.6; University of Cambridge, 2006.

(97) Smith, G.; Lynch, D. E.; Byriel, K. A.; Kennard, C. H. L.; Colin, H. L. *J. Chem. Crystallogr.* **1997**, *27*, 307–317.

(98) Kuhn, F. E.; Groarke, M.; Bencze, E.; Herdtweck, E.; Prazeres, A.; Santos, A. M.; Calhorda, M. J.; Romao, C. C.; Goncalves, I. S.; Lopes, A. D.; Pillinger, M. *Chem. Eur. J.* **2002**, *8*, 2370–2383.

(99) Jackisch, M. A.; Butler, L. G.; Fronczek, F. R. Private communication to the Cambridge Structural Database, 2006.

CT8004326