On the predictability of supramolecular interactions in molecular cocrystals – the view from the bench†

Mérina K. Corpinot,a Samuel A. Stratford,‡b Mihails Arhangelskis,‡b Jodie Anka-Lufford,‡a Ivan Halasz,c Nenad Judaš,d William Jonesb and Dejan-Krešimir Bučar*ab

A series of theophylline cocrystals involving fluorobenzoic acids was prepared and structurally characterised. The cocrystals display compositions and hydrogen-bond patterns that could not be predicted based on extensive literature/database surveys and the use of other tools. The study demonstrates that – without the use of first-principles crystal structure prediction methods – it is still remarkably difficult to predict and understand the outcomes of cocrystallisation attempts involving small and rigid molecules.

Crystal engineering concerns the understanding, design, assembly and use of crystalline materials. Although great strides have been made in all four spheres in the last few decades, understanding and designing molecular crystals remains a substantial challenge for solid-state chemists. The design of organic crystals relies on the concept of supramolecular synthons, i.e. persistent structural motifs within molecular assemblies that are formed through robust and directional non-covalent interactions, such as hydrogen or halogen bonds. The use of this concept has enabled the straightforward design and assembly of targeted discrete, one-, two- and three-dimensional structures using a single molecular building block. However, the assembly of more complex targeted structures, based on several molecules, is much more demanding and associated with more modest success rates.

Early successful instances of the synthon approach to molecular assembly were informed by crystallographic studies of molecular crystals. It is widely believed that the design and construction of cocrystals involving complex molecules will be similarly enabled by further crystallographic studies focusing on synthon hierarchies in multicomponent crystals. At present, cocrystal design is also aided by the use of statistical tools (e.g. knowledge-based hydrogen-bond propensity calculations or HBPC) and computational methods (e.g. predictions of molecular complementarity, ΔpK_value calculations and calculations of molecular electrostatic potential surfaces or MEPS). Despite the availability of such tools, and despite our evolving understanding of how molecules interact in the solid state, it still remains impossible to empirically and statistically predict (i.e. without first principles calculations) the arrangement of individual or multiple molecules in a crystal lattice. Likewise, it is extremely difficult to predict and understand the outcomes of cocrystallisation attempts involving small and rigid molecules.

† Electronic supplementary information (ESI) available: Crystallisation procedures, computational methods, statistical analyses and crystallographic data. CCDC 1451259–1451274 and 1476648–1476652. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c6ce00293e
‡ The authors contributed equally to this work.

Scheme 1 Molecular structures of thp and FBA cocrystal formers.
difficult to empirically predict the formation of supramolecular synthons in cocrystals composed of molecules containing a broad range of functional groups.4,8,27

In order to facilitate the design of cocrystals containing such complex entities, it will be necessary to considerably deepen our understanding of self-assembly processes in the solid state.

The current contribution suggests that such understanding can only be reached through well-planned crystallographic studies, and by embracing modern computational solid-state methods (such as crystal structure prediction28,29 or CSP) as standard tools in crystal engineering. We present here a case study involving cocrystals of theophylline (thp) and fluorobenzoic acids (FBAs) (Scheme 1), i.e. cocrystals based on small and rigid molecules with multiple functional groups capable of hydrogen bonding. The study demonstrates the difficulty of empirically predicting compositions and synthon hierarchies in such cocrystals, and is complemented by information from crystallographic database surveys, as well as HBPC and MEPS calculations.

Thp is commonly used as a model compound in studies of pharmaceutical cocrystals,8,10−31 and as such has been primarily cocrystallised with carboxylic acids.34 It is well known that thp engages in cocrystallisation with carboxylic acids through the formation of either an O−H(carboxyl)⋯N(imidazole) hydrogen bond (marked as synthon A in Scheme 1) or a cyclic array of N−H(imidazole)⋯O=C(carboxyl) and O−H(carboxyl)⋯O=C(amide) hydrogen bonds (marked as synthon B in Scheme 1). Our interest in (thp)/(FBA) cocrystals was sparked by our recent investigation of the (thp)/(BA) cocrystal (where BA = benzoic acid, Scheme 1). The latter is a solid reported in numerous prior studies,35,36 but one for which no crystal structure had been reported in the literature at the time we commenced our studies. Initial Cambridge Structural Database37 (CSD) surveys and HBPC suggested that interactions of type A and B are almost equally likely to occur in (thp)/(BA) (see ESI†), while a subsequent single crystal X-ray diffraction study revealed that thp and BA interact through a type B synthon (Fig. 1a).38

Another study of FBA cocrystals performed in our laboratories39 then inspired us to investigate whether minimal structural changes in BA could potentially reorganise the self-assembly process to yield thp cocrystals based on type A synthons. The choice of FBAs as cocrystal formers appeared appropriate, as all mono, di-, tri-, tetra- and penta-substituted FBAs are similar in size and shape to BA due to the relatively small size difference between hydrogen and fluorine atoms (van der Waals radii: 120 pm vs. 147 pm, respectively).40 Moreover, organic fluoride is not likely to act as a hydrogen-bond acceptor,41,42 while short intermolecular C−H⋯F contacts in molecular crystals only make a minor contribution to the cohesive energy of a molecular crystal.43 It was therefore postulated that the presence of fluorine atoms in the (thp)/(FBA) cocrystals would only minimally interfere with the formation of O−H⋯N, N−H⋯O and O−H⋯O hydrogen bonds between the cocrystal components. This hypothesis was also in agreement with results of a recently published cocrystal study involving FBAs.44

The cocrystallisation of thp with all nineteen FBAs was attempted by liquid-assisted grinding45 (LAG) using equimolar amounts of thp and the relevant FBA (see ESI†). The formation of the cocrystals was recognised via powder X-ray diffraction (PXRD) in fifteen out of nineteen cases − only cocrystal screens involving 4FBA, 23diFBA, 24diFBA and 26diFBA did not result in cocrystal formation under the investigated crystallisation conditions.46 The crystal structures of all produced solids were determined either directly from the LAG samples using laboratory PXRD data (by simulated annealing,47 Rietveld refinement48 and CASTEP49 DFT geometry optimisation), or by single crystal X-ray diffraction (SCXRD) using crystals that were obtained through solution-mediated phase transformation50 (SMPT) or by slow solvent evaporation (see ESI†).

Cocrystallisation of thp with 2FBA and 3FBA yielded 1:1 cocrystals, (thp)·(2FBA) and (thp)·(3FBA). Both cocrystals are isostructural with the (thp)·(BA) cocrystal and feature discrete thp:FBA assemblies being held together by type B synthons (Fig. 1b and c). Intriguingly, the mechanochemical synthesis of the (thp)·(2FBA) cocrystal could initially only be achieved in the Cambridge laboratory. Mechanochemical cocrystal screens performed under identical conditions in the London laboratory persistently yielded a cocrystal nitromethane solvate, namely (thp)·(2FBA)·(CH3NO2). The thp and 2FBA components in this solvate were unexpectedly sustained by type A synthon (Fig. 1d). The non-solvated (thp)·(2FBA) cocrystal was finally also obtained in London under SMPT conditions (using acetonitrile as a solvent).

A cocrystal of an unexpected composition was formed in a LAG experiment involving equimolar amounts of thp and 25diFBA where PXRD and SCXRD revealed the formation of a 1:2 cocrystal, (thp)·(25diFBA)2. The structural study also showed
that thp and 25diFBA formed discrete three-component assemblies, which were sustained by synthons of both type A and B (Fig. 2a). Several attempts to produce a 1:1 thp:25diFBA cocrystal either mechanochemically or through solution-based crystallisation methods have, so far, failed.

Our investigations also led to the discovery of a 1:1 cocrystal comprised of thp and 35diFBA, (thp)·(35diFBA). This material surprisingly exhibits discrete four-component assemblies featuring a O–H(carboxyl)⋯O==C(urea) synthon (marked as type C in Scheme 1 and shown in Fig. 2b) that are statistically less likely to occur than synthons A and B, as determined by HBPC (see ESI†). More predictable synthons were, however, found in the 1:1 (thp)·34diFBA cocrystal wherein thp and the FBA interact via type A synthon (Fig. 2c).

Attempts to produce cocrystals using LAG and equivalent amounts of thp and tri-, tetra- and penta-FBAs yielded 1:1 cocrystals, namely (thp)·(234triFBA), (thp)·(235triFBA), (thp)·(236triFBA), (thp)·(245triFBA), (thp)·(246triFBA), (thp)·(245triFBA), (thp)·(2345tetFBA), (thp)·(2346tetFBA), (thp)·(2356tetFBA) and (thp)·(23456pFBA). Their crystal structures (except the one involving thp and 245triFBA) featured discrete four-component assemblies composed of two thp and two FBA molecules. In these assemblies, the thp molecules interact with each other via cyclic arrays of N–H(imidazole)⋯O==C(amide) hydrogen bonds, while the FBAs bind to the thp dimer through synthons of type A (some examples are featured in Fig. 3, others are illustrated in the ESI† file). Surprisingly, the crystal structure of (thp)·(245triFBA) exhibited assemblies that were held together by synthons of type B (Fig. 4a).

The (thp)·(FBA) cocrystals described above clearly demonstrate the unpredictability of synthon hierarchies in two-component cocrystals with several hydrogen-bonding functional groups. The observed supramolecular interactions could neither be predicted empirically nor with the use of HBPC and MEPS calculations. In particular, HBPC has been performed on all structurally characterised solids using Mercury 3.7 and the CSD 5.37 (see ESI†). Although the calculations encouragedly predicted the formation of the observed hydrogen bonds between thp and the FBA in each case, the synthon hierarchies could not be foreseen (see ESI†). In addition, DFT calculations of MEPS were also performed to predict optimal hydrogen-bond donor–acceptor pairs (see ESI†). The values of the electrostatic potentials on the hydrogen-bond donors and acceptors could, unfortunately, not be correlated to the formation of the observed synthons in the cocrystal series (see ESI†). We observed, however, that under the initially studied LAG conditions, almost all tri-, tetra- and penta-FBAs form cocrystals based on type A synthons, while the mono- and di-FBAs participated in hydrogen bonding through type A, B and C interactions (see ESI†).

The difficulty of predicting supramolecular synthons in cocrystals composed of thp and FBA is further highlighted by the discovery of polymorphs of four (thp)·(FBA) cocrystals, namely (thp)·(34diFBA), (thp)·(245triFBA), (thp)·(246triFBA) and (thp)·(23456pFBA). These polymorphs were observed during attempts to grow single crystals of the mechanochemically obtained crystal forms by crystallisation through slow solvent evaporation (see ESI†). Structural analyses revealed that the four solids exhibited both packing and synthon polymorphism.51 Packing polymorphism was observed in the case of (thp)·(34diFBA), whereas (thp)·(245triFBA), (thp)·(246triFBA) and (thp)·(23456pFBA) displayed synthon polymorphism. In particular, the synthon polymorphism of (thp)·(245triFBA) was realised through the formation of the type A synthon (Fig. 4a). On the other hand, synthon polymorphism in the cases of (thp)·(246triFBA) and

![Fig. 2 Perspective views of thp·FBA assemblies in the crystal structures of: a) (thp)·(25diFBA), b) (thp)·(35diFBA) and c) (thp)·(34diFBA) featuring distinct supramolecular synthons.](image1)

![Fig. 3 Perspective views of thp·FBA assemblies in the crystal structures of: a) (thp)·(234triFBA) and b) (thp)·(2346triFBA).](image2)
(thp·(23456pFBA) was realised through the occurrence of N–H(imidazole)⋯O=C(urea) hydrogen bonds between two thp molecules, rather than the generally observed cyclic arrays of N–H(imidazole)⋯O=C(amide) hydrogen bonds (Fig. 4b). The serendipitous discovery of the four polymorphs is not particularly surprising, considering the rapidly increasing number of studies reporting cocrystal polymorphs, while a recent study recognised that cocrystals exhibit polymorphism as readily as single-component organic crystals. In light of these findings, it appears that a better understanding of synthon hierarchies in cocrystals could be achieved by surveys of their structural landscapes through polymorph screens, which are not regularly performed at the present time.

So, what is the significance of the intriguing synthon diversity in the presented (thp·FBA) cocrystal series? The unpredictability of supramolecular interactions in the (thp·FBA) cocrystals, and lack of knowledge about the structural and electronic effects (on synthon hierarchies and cocrystal compositions) of fluorine substitutions in the cocrystal former, suggest that cocrystals are not as readily designed and predictably constructed as is so often portrayed in the literature. Supramolecular synthons are an incredibly useful synthetic tool that has greatly contributed to the development of advanced functional materials, and to the preparation of strikingly complex solid-state structures. The current state of the art in cocrystal design, however, does not support high-yielding supramolecular syntheses of cocrystals for cases entailing molecules with more than two functional groups. To improve our ability to better design and more accurately construct cocrystals comprised of complex multifunctional molecules (such as pharmaceuticals), more targeted and systematic crystallographic studies need to be pursued and published. But to fully understand cocrystals – their structural landscapes, synthon hierarchies and thermodynamic properties – computational methods (such as CSP or lattice energy calculations) need to be regularly used in crystal engineering. Comprehensive computational solid-state studies are very intensive (requiring a great deal of CPU time and time to process the generated data) and it can be argued that their use is, nowadays, only justified when dealing with specialty chemicals of particular interest (e.g. drug candidates in development for marketing). Recent reports, however, have shown that even less exhaustive predictions can contribute substantially to experimental cocrystal research. The maturity of contemporary first principles computational solid-state methods and their relevance to crystal engineering is now widely recognised, and it is time that they are equally widely embraced by the experimental community.

MKC and DKB gratefully acknowledge financial support from the UCL Faculty of Mathematical and Physical Sciences. DKB and WJ thank the Royal Society for a Newton International Fellowship and the Isaac Newton Trust (Trinity College, University of Cambridge) for funding. MA thanks the EPSRC for a studentship, while SAS acknowledges funding through the EPSRC CASE scheme with Pfizer. We are grateful for computational support from the UK national high performance computing service, ARCHER, for which access was obtained via the UKCP consortium and funded by EPSRC grant (EP/K013564/1). Lastly, we thank Prof. Sarah L. Price, Dr. Edward O. Pyzer-Knapp, Dr. Rui Guo and Mr Luca Iuzzolino for helpful discussions.

Notes and references

Communication

58 P. G. Karamertzanis, A. V. Kazantsev, N. Issa, G. W. A. Welch, C. S. Adjiman, C. C. Pantelides and S. L. Price, 
5154–5165.
60 D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. 
Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-
Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van 
Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, 
D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. 
Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, 
F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, 
Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, 
H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. 
535–551.
61 K. Hoxha, D. H. Case, G. M. Day and T. J. Prior, 
*CrystEngComm*, 2015, 17, 7130–7141.
62 J. Yang, W. Hu, D. Usvyat, D. Matthews, M. Schütz and 
63 A. V. Kazantsev, P. G. Karamertzanis, C. S. Adjiman, C. C. 
Pantelides, S. L. Price, P. T. A. Galek, G. M. Day and A. J. 
64 The Cambridge Crystallographic Data Centre (CCDC), 