CrystEngComm

Accepted Manuscript
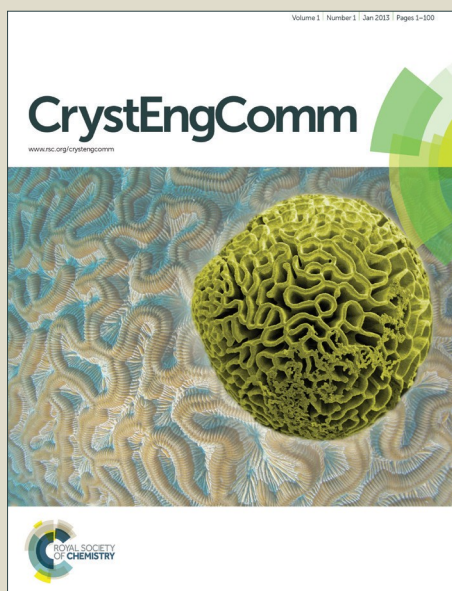
ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/crystengcomm

Crystal structure landscape in conformationally flexible organo-fluorine compounds

Pradip Kumar Mondal,[a] and Deepak Chopra*[a]

The crystal structure landscape for the unsubstituted benzanilide was generated and a number of hypothetical structures accessed with experimentally obtained crystal structures of *mono-*, *di-*, *tetra-* and *penta-*fluorobenzanilides. Thus, chemical modification allows us to access the "high energy" forms of the parent compound thereby delineating the significant role of weak intermolecular interactions.

Crystal structure prediction (CSP) methods for organic molecules have attracted tremendous interest in recent years [1]. These are based on searches for the thermodynamically most feasible crystal structure, and such an evaluation neglects the role of entropy and the kinetics of crystallization [1]. The role of the solvent, temperature, pressure and other related kinetic factors may lead to the formation of alternate crystalline polymorphs [2]. The existence of all possible polymorphic forms in experimentally determined crystal structures is difficult to predict using empirical methods practiced in the prediction of crystal structures of an organic molecule. CSP once conceived to be a challenging exercise [3] has been successfully performed on rigid molecules and the success obtained from the first four blind tests are testimony to this fact (CSP1999 (1st blind test), CSP2001 (2nd blind test), CSP2004 (3rd blind test) and CSP2007 (4th blind test)). For the conformationally flexible molecule (2-3 internal degrees of freedom), the fourth blind test was able to successfully predict the crystal structure of N-(Dimethylthiocarbamoyl)benzothiazole-2-thione [4]. However, the fifth blind test (CSP2010) attempted to predict the crystal structure of a highly flexible molecule (8 internal degrees of freedom) and the tests resulted in a successful prediction [5]. Interestingly, Leusen and co-workers successfully predicted the crystal structure of another molecule in 2011, which could not be predicted in the 2001 blind test [6]. In this study, we have introduced as a new challenge, a moderately flexible organic molecule [7], namely the unsubstituted benzanilide [Scheme 1], and we have evaluated the crystal structure landscape (CSL) [8] for the same. The CSL for a given compound is generated on account of the variations in molecular conformation and arrangement of different packing motifs during the crystallization process. The possible crystal structures that lie in a narrow range of energies. In our case, this molecule is moderately flexible with three internal degrees of freedom [Scheme 1] and contains 26 atoms, [only the elements C, H, N, O and F]. For our target molecule, the landscape was generated with Z' = 1 *only* and the relevant space groups, namely, $P1$, $P\text{-}1$, $P2_1$, Pc, $P2_1/c$, $C2/c$, $P2_12_12_1$, $Pna2_1$, and $Pca2_1$ were *only* investigated. It is to be kept in mind that the experimentally realized crystal structures of fluorinated benzanilides were observed to crystallize in the above-mentioned space groups and hence these only have been considered in the current study. The number of generated crystal structures in the potential energy landscape of the crystal is more than five thousand in number. Hundred lowest energy structures, on thermodynamic considerations, were only analyzed. To map the CSL generated by this molecule we have synthesized a series of different molecules wherein we have introduced fluorine atoms (as an isosteric replacement of the hydrogen atom/atoms) and their experimentally realized crystal structures do enable us to access the "hypothetical" structures predicted for the parent compound. One such study on benzoic acid and cocrystals of benzoic acid, wherein the role of replacement of H with F in accessing "hypothetical" structures for benzoic acid and its cocrystals has been performed [8]. It is noteworthy, to extend the exercise of CSL to relatively flexible molecules containing organic fluorine. This is on account of the fact that such molecules exhibit dynamic disorder [9] in the crystalline lattice. Furthermore, compounds containing fluorine have tremendous applications in all fields of science [10]. It is of interest to note that the simplest possible chemical modification of the molecule can lead to the formation of altered crystal structures wherein the associated difference in energetics is extremely negligible. This has been observed in the case of the deuterated form of pyridine [11]. Thus, the predicted crystal structures need to be verified with the experimental structures, and this requires

performing a large number of crystallization trials to get the required crystal structure as a polymorph [12] for a particular compound of interest. Finally, similar exercises were done on the CSP of paracetamol by Neumann and Perrin in 2009 [13] and its validity with the experimental structures (four polymorphic forms) was also realized paving the pathway to "polymorph instantiation".

Starting from a simple chemical diagram, computations related to CSL were performed with *Materials Studio 6.1* using the COMPASS26 force field. We chose a default set up entitled "fine quality" for packing, geometry optimization, and clustering. This procedure is a crystal structure generation tool to get closely related crystal structures of the parent compound which qualify as polymorphs. Our analysis is thus based on the 100 most stable close-packed structures. The lattice energy of all the experimental crystal structures was calculated by PIXELC module in CLP computer program [14]. The plot of the lattice energy versus density depicts that the space group $P2_1/c$ results in the most efficient packing, as is reflected in the relatively high magnitudes of the density of the compound [see in the ESI F-3].
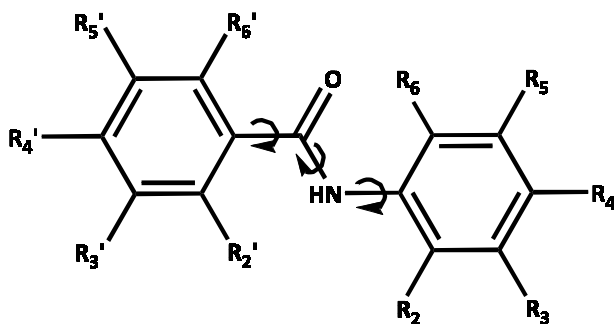
Scheme 1: Chemical scheme of the compound studied. Nomenclature scheme Pmnm'n'; mn = aniline side and m'n' = benzoyl group side; m, n, m' and n' = any value from 2 to 6 or 0 (for absence of fluorine atom in the ring); R and R' = hydrogen or a fluorine atom.

Table 1: Experimental structures of fluorinated benzanilides in the benzanilide landscape

Group	Compound Code	Unit Cell Type	Rank	Space Groups
Group 1	P0000, P0034, P0035, P2334, P3400, P4000, P0040, P4040	5-8-12	4th and 11th	P-1
Group 2	P3500	10-5-20	15th and 65th	$P2_1/c$
Group 3	P0024, P3423, P3435, QUKVUN, P2324, P2423, P3000	5-8-25	48th, 88th, and 93rd	$P2_1/c$
Group 4	P2323	5-9-23	68th	$P2_12_12_1$
Group 5	P2040 (Form 2), P0020	5-5-19	14th, 17th, and 26th	$P2_1$
Group 6	P2335	5-5-12	73rd	$P1$
Group 7	P0026	10-21-5	23rd	$Pna2_1$
Group 8	P2500, P2040 (Form 1), P3020	24-5-8	2nd, 6th, 39th, 50th, and 52nd	$Pca2_1$
Group 9	P2400	5-5-19	24th	Pc

To understand the formation of organic solids, it is of importance to recognize the role of non-covalent interactions in the study of the CSL. In this regard, the presence of strong and well-defined **N-H···O** hydrogen bond is highly effective in the assembly of benzanilide structures [15]. Crystal structures also display the formation of **C-H···O**, **C-H···π** and **π···π** intermolecular interactions in the crystalline lattice. An isosteric replacement of the hydrogen atom with fluorine atoms on

the phenyl ring results in the formation of **C-H⋯F** intermolecular interactions and these are classified as H-bonds when the interaction distance is short [16] and are associated with directional characteristics. Our study involves the formation of alternative packing modes in these isomeric molecules through the process of CSP. We intend to simplify the relation between the theoretically predicted structures of unsubstituted benzanilide and the experimentally realized crystal structures. These are obtained by the change in the position of the fluorine atoms ongoing from *ortho* to *meta* to *para* in different isomeric molecules.

To verify such a study, based on prediction, with the experimental structure, we divided the results obtained into a total of nine groups in accordance with their space group and unit cell type. The experimental unsubstituted benzanilide **P0000** [CSD code: **BZANIL02**], has the space group *P*-1 and the following lattice parameters: *a* = 5.352Å, *b* = 7.971Å, *c* = 12.471Å. This structure belongs to Group 1 with unit cell type 5-8-12 and is ranked at 4th and 11th position in the CSL [Table 1]. This structure is held by strong **N-H⋯O** hydrogen bonds and multiple weak **C-H⋯π** interactions. We further consider the monofluorobenzanilides, **P4000** (*P*-1, 5.369Å, 7.862Å, 12.892Å) and **P0040** (*P*-1, 5.349Å, 7.599Å, 12.945Å) respectively. The crystal structure of these molecules match with those obtained from the CSL of benzanilide, ranked 4th and 11th position in the energy-density plot [Figure 1] in the same group. We further consider the difluorobenzanilides (two fluorine atoms present in one or both of the phenyl rings) **P4040** (*P*-1, 5.460, 7.693, 13.038), **P0034** (*P*-1, 5.4385, 7.5939, 12.8178), **P0035** (*P*-1, 5.1300, 8.8893, 11.6782) and **P3400** (*P*-1, 5.4838, 7.7928, 12.6887) are also similar and their crystal structures are also comparable with the CSL of benzanilide. It is of interest to note that the overlay diagram for **P4000**, **P4040**, **P0034** and **P3400** are almost similar with the 4th ranked structure [Figure 2]. In continuation to our hypothesis, we now consider the case of tetrafluorobenzanilides (two fluorine atoms in each of the two phenyl rings), **P2334** (*P*-1, 4.9918, 9.3610, 12.0172), which belongs to the same group (see in the ESI for overlay and RMS deviation, Fig. F-2(a-h)). These structures are also constructed *via* **N-H⋯O** hydrogen bond, but the crystal structures are arranged utilizing other related weak interactions. In **P0034**, **P3400**, **P4000**, **P0040** and **P4040**, the **C-H⋯π** interactions are present in the crystal packing. In the case of **P0035** and **P2334**, the **C-H⋯F** and **π⋯π** are the more significant interactions.
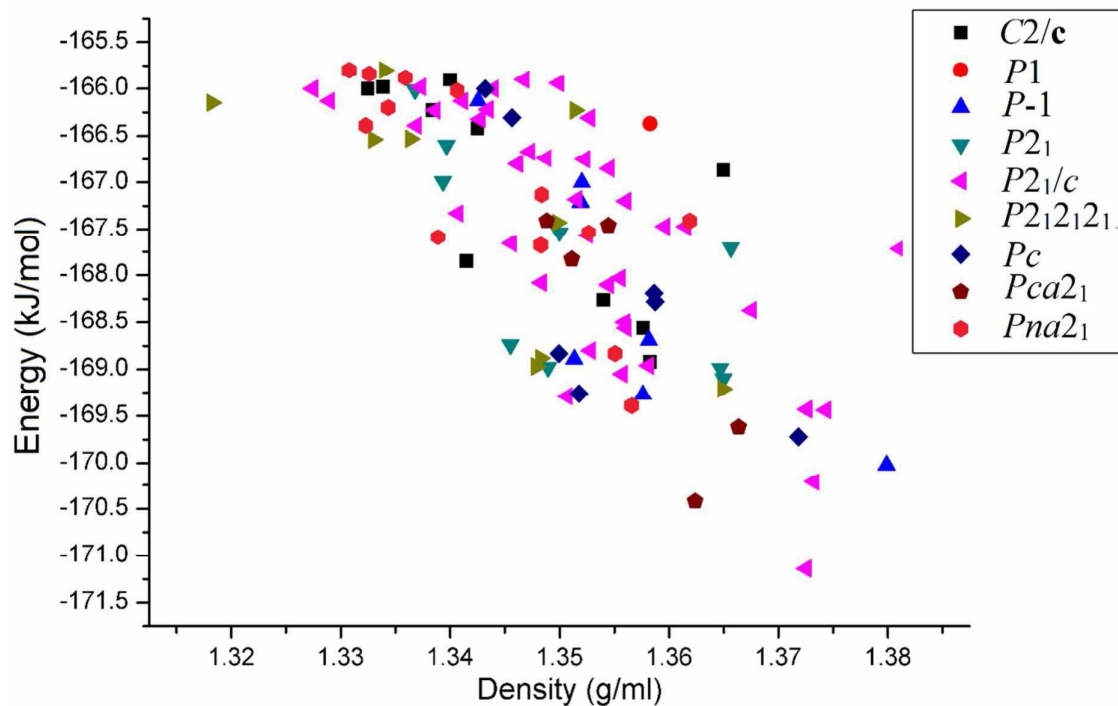
Figure1: Plot of the relative lattice energies versus densities of the top 100 predicted crystal structures for unsubstituted benzanilide in different space groups.

The isomeric difluorobenzanilide, **P3500** ($P2_1/n$, 10.179, 5.150, 20.053) belongs to group 2, in the centrosymmetric $P2_1/c$ space group and the crystal structure is similar to that of the 15th and 65th ranked benzanilide in the energy-density landscape. The third group is the 5-8-25 structure type in the same centrosymmetric $P2_1/c$ space group. Monofluorobenzanilide, **P3000** ($P2_1/c$, 8.069, 5.391, 23.238), difluorobenzanilide, **P0024** ($P2_1/n$, 5.4223, 7.6977, 25.4353), tetrafluorobenzanilides, **P3423** ($P2_1/n$, 5.1818, 8.312, 25.739), **P3435** ($P2_1/n$, 5.1818, 8.312, 25.517) **P2324** ($P2_1/n$, 9.456, 4.7786, 24.253), **P2423** ($P2_1/n$, 8.8818, 4.9233, 24.9499) and pentafluorobenzanilide (five fluorine atoms in one ring) **QUKVUN** ($P2_1/c$, 4.982, 9.724, 25.775) belong to this structure type with rank 48th, 88th and 93rd as obtained from the energy-density ranking [see ESI F-2(i-k)]. The third group with $P2_1/c$ space group is primarily seen in the experimentally determined crystal structures of highly fluorinated benzanilides, **P3423**, **P3435**, **QUKVUN**, **P2324** and **P2423,** wherein the **C-H⋯F** and **π⋯π** are the decisive interactions. The **π⋯π** interaction is mainly predominant in substituted benzanilides containing a higher number of fluorine atoms. This is because the incorporation of such an electronegative atom changes the uniform electron density distribution on the benzene ring [Figure 4]. The **C-H⋯π** interactions are significant for the crystal structure of **P0024**. Detailed investigation on the energetics and topological characteristics of different supramolecular constructs utilizing weak interactions for these class of molecules shall be reported separately [17].

The fourth group belongs to the orthorhombic crystal system and $P2_12_12_1$ space group. Molecules containing tetrafluorinated benzanilides (two fluorine atoms in each ring) **P2323** ($P2_12_12_1$, 5.0295, 8.838, 24.4547) occupies the 68th position in the energy-density ranking list (see ESI F-2(l)). The crystal structure is arranged via **C-H⋯F** and **π⋯π** interactions. Monofluorinated benzanilide **P2040 (Form 2)** ($P2_1$, 4.9617, 5.4859, 19.174) as well as the higher difluorobenzanilide **P0020** ($P2_1$, 5.421, 6.258, 15.534, 97.45) belong to the group 5 which present specific positions in the ranking list, namely 14th, 17th and 26th respectively [see in the ESI F-(m-o) for overlay diagrams]. Overlay Diagram for **P2040 (Form 2)** is almost similar to the 26th ranked structure [Figure 3].

Only one theoretically predicted crystal structure out of the hundred obtained structures, containing the triclinic *P*1 space group and ranked 73rd in the CSL (corresponding to group 6) was observed. This nicely matches with the crystal structure landscape of tetrafluorobenzanilides (two fluorine atoms in each ring) **P2335** (*P*1, 4.6457, 5.0544, 11.8597) [Figure 3]. The 23rd position in the rank list is occupied by difluorobenzanilide **P0026** (*Pna*2₁, 9.914, 21.812, 4.923) in group 7 [see ESI F-2(p)] crystallizing in the orthorhombic crystal system.
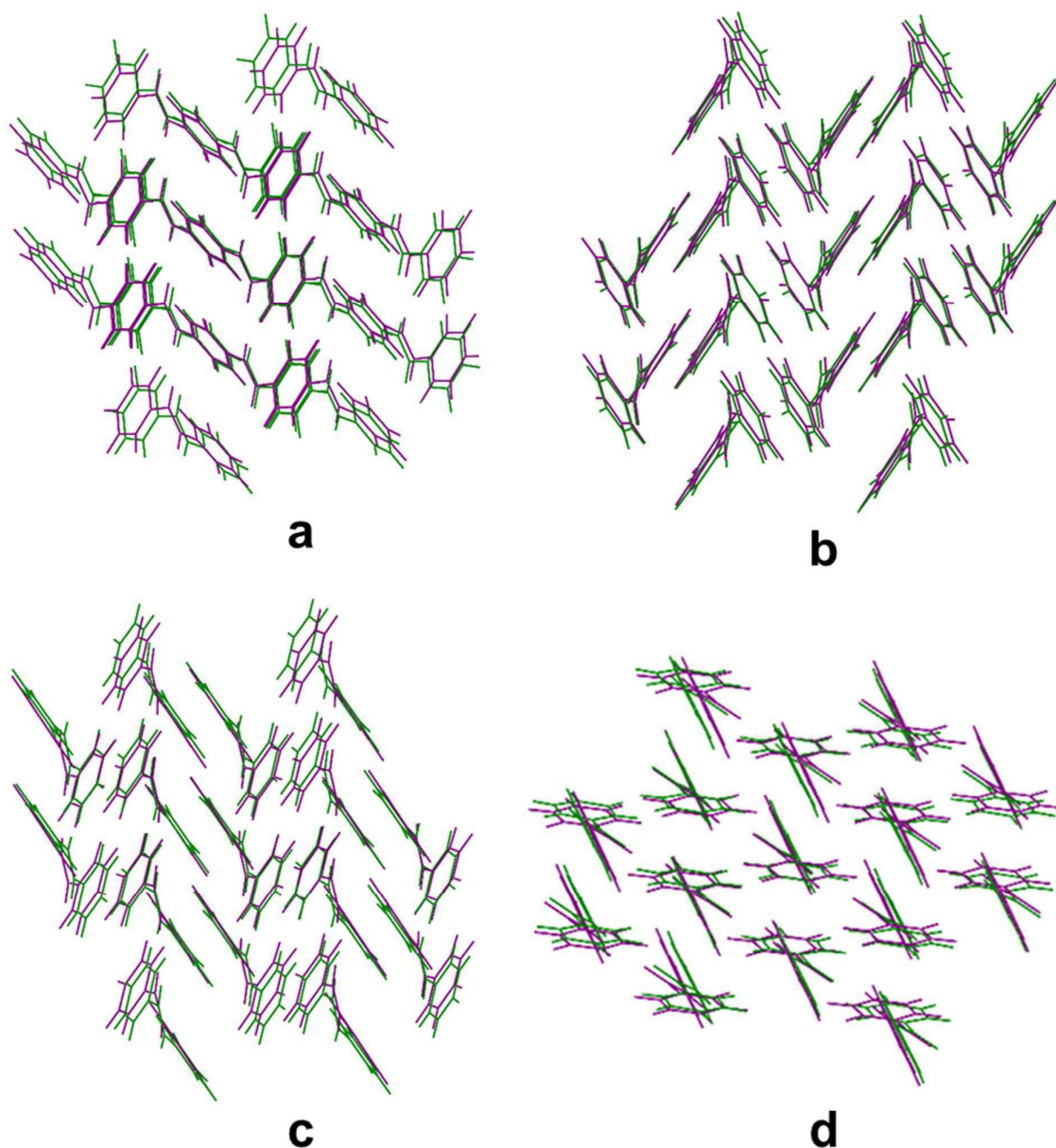
Figure 2: The overlay diagram of the 15 molecules of the experimental (green) and predicted (purple) structure of a) **P0034** and the 4th rank structure with an RMS deviation of 0.328 Å. b) **P4000** and the 4th rank structure with an RMS deviation of 0.264 Å. c) **P3400** and the 4th rank structure with an RMS deviation of 0.370 Å. d) **P4040** and the 4th rank structure with an RMS deviation of 0.344 Å.

Difluorobenzanilides (two fluorine atoms may be present in one ring or both rings) **P2500** (*Pca*2₁, 24.3084, 5.0243, 8.4598), **P2040 (Form 1)** (*Pca*2₁, 25.563, 4.969, 8.250) and **P3020** (*Pca*2₁, 24.660, 5.203, 8.244) belonging to the *Pca*2₁ space group

present at specific positions in the ranking list [Figure 3], namely 2nd, 6th, 39th, 50th and 52nd in group 8 [see ESI F-2(q-s)]. It is indeed interesting to obtain a short contact to the fluorine atom in case of **P2500**, [**C13-H13···F1** hydrogen bond (Neutron value: 2.09 Å, 154˚) [Figure4]. Finally in group 9, the difluorobenzanilide, **P2400** (*Pn*, 5.535, 5.035, 19.29) is similar with the 24th rank predicted crystal structure on the CSL [see ESI F-2(t)]. A short **C12A-H12A···F1A** hydrogen bond (Neutron value: 2.29 Å, 174˚) are responsible for the crystal packing along with the presence of a **C3-H3···O1** hydrogen bond (Neutron value: 2.66 Å, 146˚) [Figure 4].
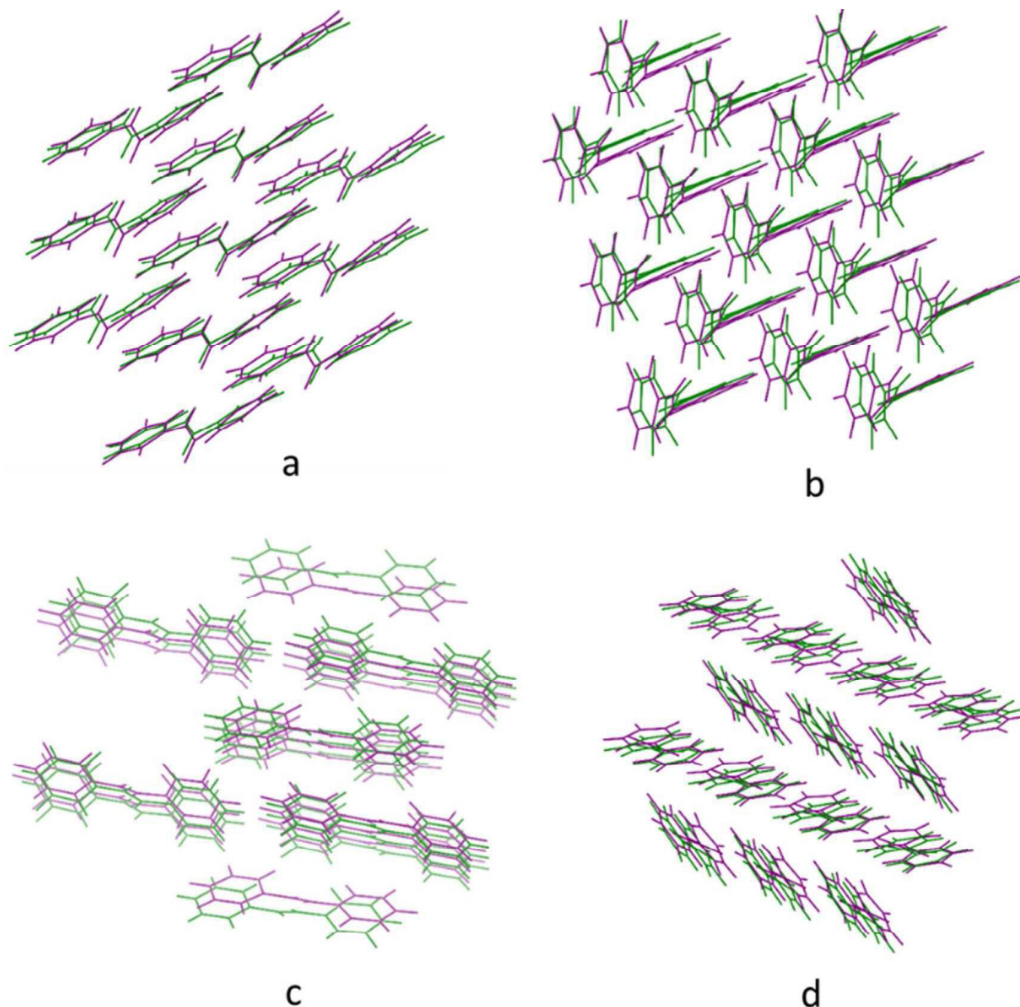

Figure 3: Final overlay diagram of the experimental (green) and predicted (purple) structure of a) **P2040 (Form 2)** and the 17th rank structure depicting 11 out of the calculated 15 molecules with an RMS deviation of 0.328 Å. b) **P2335** and the 73rd rank structure depicting all the 15 molecules with an RMS deviation of 0.458 Å. c) **P2500** and the 6th rank structure depicting all the 15 molecules with an RMS deviation of 0.61 Å. d) **P3020** and the 52nd rank structure depicting all the 15 molecules with an RMS deviation of 0.461 Å.

To further substantiate our understanding of the CSL in fluorine containing molecules, we have undertaken the exercise of obtaining the landscape of crystal structures for decafluoro-substituted benzanilide, where in all the ten hydrogen atoms on the two phenyl rings is replaced by the fluorine atoms. These molecules also provide the required crystal energy landscape for experimental realization of the crystal structures of *mono-*, *di-*, *tetra-* or related poly fluoro-substituted benzanilides [see in the ESI T-4]. These compounds generate a very similar group of crystal structure types and unit cell configurations. It was observed that the CSL generated for deca-fluorinated benzanilide does not produce the required overlay diagram when mapped to the experimentally determined crystal structures. Hence no values of the "similarity index (RMS deviation)" are reported. This may be on account of the variations in the crystal density of the calculated landscape for decafluorinated benzanilides which lies in the range of 2.09 g/ml to 2.21 g/ml, whereas the densities for the experimentally determined crystal structures (corresponding to the presence of different fluorine atoms) lies in the range of (1.30 g/ml-1.71 g/ml) which is closer to the ranges in density (1.31 g/ml to 1.39 g/ml) obtained from CSL of the unsubstituted benzanilide. The energy density graph [ESI F-3], the table [ESI T-5] and the CIF of the first 100 predicted structures corresponding to the deca-fluorinated benzanilide are provided in the ESI.
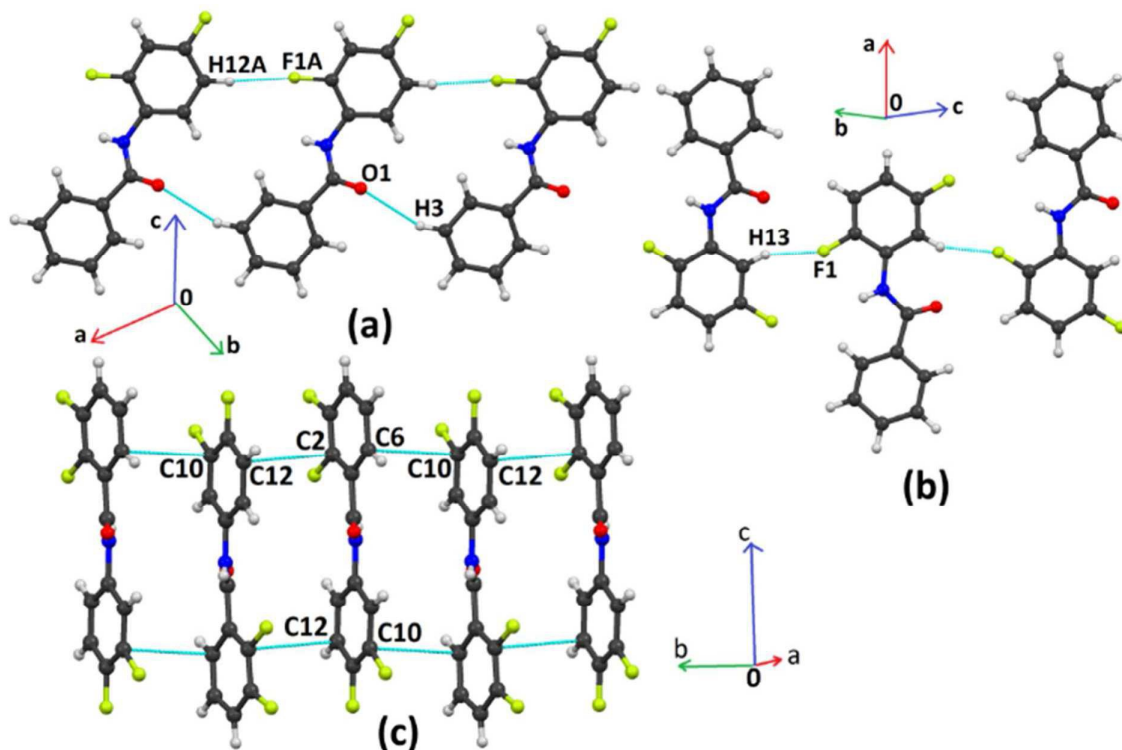

Figure 4: (a) C-H···F hydrogen bond along with **C-H···O** utilizing the (x-1, y+1, z) symmetry in **P2400. (b) C-H···F** hydrogen bond utilizing the (1.5–x, y-1, z+0.5) symmetry in **P2500. (c)** Crystal packing in **P3423** depicting interaction of the electron deficient carbon atom C10 with the electron rich carbon atom C6 [3.343 Å] (a similar feature is observed between C2 and C12 [3.319 Å]).

In conclusion, the landscape depends on the number of fluorine atoms and the position of the fluorine atoms. When the number of fluorine atoms is less, one or two, then the crystal packing is primarily guided by **C-H···π** interactions and in the case of higher number of fluorine atoms, the prevalence of **C-H···F, F···π,** and **π···π** interactions is important for the formation of the crystal structures. This is a subtle, yet importance structural feature responsible for formation of crystals utilizing weak interactions. The position of the fluorine is also important in the consideration of the factors that are instrumental in the finally obtained crystal structure (for example [**P0034** and **P3400]**, [**P2324** and **P2423]**, [**P4000** and **P0040]** are present in the same group. This entire exercise reflects a pivotal point in the crystal chemistry of organic solids which is equivalent to stating the fact that "chemical modification" leads to the experimental realization of different "forms" of a compound which otherwise are not accessible under routine crystallization conditions for the molecule of interest. The isolation of polymorphs for the poly-fluorinated benzanilides is expected to render deeper insights into the role of weak interactions and facilitate a greater degree of mapping with the crystal structure landscape of the parent compound.

Notes and references

1   (a) S. L. Price, *Chem. Soc. Rev.*, 2014, **43**, 2098-2111. (b) S. L. Price, *Phys. Chem. Chem. Phys.*, 2008, **10**, 1996-2009. (c) G. M. Day, W. D. S. Motherwell and W. Jones, *Cryst. Growth & Des.*, 2005, **5**, 1023-1033. (d) G. M. Day, *Cryst. Rev.* 2011, **17**, 3-52. (e) D. E. Braun, M. Orlova, and U. J. Griesser, *Cryst. Growth & Des.*, 2014, **14**, 4895-4900. (f) D. E. Braun, J. A. McMahon, L. H. Koztecki, S. L. Price and S. M. Reutzel-Edens *Cryst.Growth&Des.*, 2014, **14**, 2056-2072. (g) A. T. Hulme, S. L. Price, and D. A. Tocher, *J.Am.Chem.Soc 2005*, **127**, 1116-1117.

2   S. L. Price, *ActaCryst.*, 2013, **B69**, 313-328.

3   (a) A. Gavezzotti, *CrystEngComm*, 2002, **4(61)**, 343-347. (b) A. Gavezzotti, *Acc.Chem.Res*, 1994, **27**, 309-314.

4   G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. D. Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. Van de Streek, A. K. Wolf and B. Schweizer, *ActaCryst.* 2009, **B65**, 107-125.

5   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. D. Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti, and I. K. Zhitkov, *ActaCryst.* 2011, **B67**, 535-551.

6   (a) M. A. Neumann, F. J. J. Leusen and J. Kendrick, *Angew. Chem. Int. Ed.* 2008, **47**, 2427-2430. (b) H. C. S. Chan, J. Kendrick, and J. J. Leusen, *Angew. Chem. Int. Ed.* 2011, **50**, 2979-2981.

7   C. Ouvrard and S. L. Price *Cryst. Growth & Des.*, 2004, **4**, 1119-1127.

8   (a) R. Dubey, M. S. Pavan and G. R. Desiraju, *Chem. Commun.*, 2012, **48**, 9020-9022. (b) R. Dubey and G. R. Desiraju, *Chem. Commun.*, 2014, **50**, 1181-1184.

9   S. K. Nayak, M. K. Reddy, D. Chopra and T. N. G. Row, *CrystEngComm*, 2012, **14**, 200-210.

10 K. Muller, C. Faeh and F. Diederich, *Science*, 2007, **317**, 1881-1886.

11 S. Crawford, M. T. Kirchner, D. Bläser, R. Boese, W. I. F. David, A. Dawson, A. Gehrke, R. M. Ibberson, W. G. Marshall, S. Parsons, and O. Yamamuro, *Angew.Chem.Int.Ed.*2009, **48**, 755-757.

12 D. Chopra and T. N. G. Row, *Cryst. Growth & Des.* 2005, **5**, 1679-1681.

13 M. A. Neumann and M. A. Perrin *CrystEngComm*, 2009, **11**, 2475-2479.

14 (a) A. Gavezzotti, *New J. Chem.*, 2011, **35**, 1360; (b) A. Gavezzotti, *J. Phys. Chem. B*, 2003, **107**, 2344; (c) A. Gavezzotti, *J. Phys. Chem. B*, 2002, **106**, 4145.

15 D. Chopra and T. N. G. Row, *CrystEngComm*, 2008, **10**, 54-67.

16 P. Panini and D. Chopra, *Cryst. Growth & Des.*, 2014, **14**, 3155-3168.

17 P. K. Mondal and D. Chopra, 2015 (to be communicated).