Electronic Supplementary Material (ESI) for Chemical Communications

Supplementary Information

Structural landscape of benzoic acid: using experimental crystal structures of fluorobenzoic acids as a probe

Ritesh Dubey, Mysore S. Pavan and Gautam R. Desiraju*
Solid State and Structural Chemistry Unit
Indian Institute of Science, Bangalore 560 012, India

S-1: Data collection and refinement details, crystallization methods

TABLES

T-1: Normalized hydrogen bond distances
T-2: Crystallographic tables of the experimental structures
T-3: Computational details for the ESP charge calculation and polymorph prediction protocols

- For benzoic acid (COMPASS26)
- For 4-fluorobenzoic acid (DREIDING)

FIGURES

F-1: Experimental as well as computed structure overlay diagrams
F-2: Energy versus Density plot of top 100 predicted structures of 4-fluorobenzoic Acid
F-3: ORTEP diagrams for the asymmetric unit of the crystal structures (ellipsoids are drawn at 50% probability)

S-1: Data collection and refinement details, crystallization methods

- Single crystal X-ray diffraction

Single crystal X-ray data for 24DFBA were collected on an Oxford Xcalibur (Mova) diffractometer[1] equipped with an EOS CCD detector using Mo-Kα radiation (λ= 0.71073 Å). The crystal was maintained at the desired temperature during data collection using the Oxford Instruments Cryojet-HT controller[2]. The remaining crystal data were collected on a Rigaku Mercury375R/M CCD (XtaLAB mini) diffractometer using graphite monochromated Mo-Kα radiation, equipped with a Rigaku low temperature gas spray cooler, data were processed with the Rigaku CrystalClear software[3]. Structure solution and refinement were performed using SHELX-97[4] embedded in the WinGX suite[5] and Olex2.[6] Refinement of coordinates and anisotropic thermal parameters of non-hydrogen atoms were carried out by the full-matrix least-squares method. The disorder was analyzed using the PART command in SHELXL-97 and the disordered fluorine were treated and refined with two independent positions, namely A and B. In all cases, except 345TFBA-I, hydrogen atoms were found in difference Fourier maps. In 345TFBA-I, the hydrogen atoms were refined by constraining to geometrical positions. The hydrogen bond tables were created using PLATON.[7] Mercury version 3.0 was used for molecular representations, overlay and packing diagrams.[8]

1. Oxford Diffraction. CrysAlis PRO CCD and CrysAlis PRO RED; Oxford Diffraction Ltd.: Yarnton, England, 2009.
2. Oxford Instruments, Cryojet XL/HT controller; Oxford Diffraction Ltd.: Yarnton, England.
3. CrystalClear 2.0, Rigaku Corporation, Tokyo, Japan.
4. G. M. Sheldrick, *ActaCrystallogr., Sect. A*, 2008, **64**, 112.
5. L. J. Farrugia, *J. Appl. Crystallogr.* 1999, **32**, 837.
6. O. V. Dolomanov, L.J. Bourhis, R.J. Gildea, J.A.K. Howard, H. Puschmann, .*J. Appl. Crystallogr.*, 2009, **42**, 339.
7. A. L. Spek, *J. Appl. Crystallogr.*, 2003, **36**, 7.
8. C. F. Macrae, P. R. Edgington, P. McCabe, E. Pidcock, G. P. Shields, R. Taylor, M. Towler, J. van de Streek, *J. Appl. Crystallogr.*, 2006, **39**, 453.

- Crystallization Methods

All compounds were purchased from Sigma-Aldrich and Alfa-Aesar, and were used without further purification.

All crystallization experiments were carried out at ambient conditions. In 24DFBA and 345TFBA (I), we applied melt crystallization through gradient heating and cooling via oil-bath and obtained good quality single crystals after two to three hours. In 34DFBA and 25DFBA, good quality single crystals were obtained by solvent evaporation using a mixture of hexane-methanol (1:1). Similarly good quality single crystals of 234TFBA were obtained from a mixture of benzene and methanol (1:1). Similar method gave good quality single crystals for (236TFBA and 345TFBA (II)), (245TFBA and 2345TeFBA) and (235TFBA and 24SS) in benzene, chloroform and methanol solvents respectively. In all cases of slow solvent evaporation, crystals were obtained within three to four days.

T-1: Normalized hydrogen bond distances for the compounds in this study

Compound	D–H···A[#]	r(D–H)/Å	r(D–A)/Å	r(H···A)/Å	∠ D–H···A/°	Symmetry
24DFBA	O1–H1···O2	0.98	1.63	2.616(2)	177	2-x,-y,-z
	C5–H5···O2	1.08	2.45	3.400(3)	146	-1+x,1+y,z
25DFBA	O2–H2···O1	0.9800	1.6500	2.6285(17)	172.00	1-x,1-y,1-z
	C4–H4···O1	1.0800	2.4700	3.377(2)	141.00	-1+x,3/2-y,-1/2+z
34DFBA	O1–H1···O2	0.9800	1.6500	2.619(2)	167.00	1-x,-y,-z
2345TeFBA	O1–H1···O2	0.9800	1.6600	2.636(2)	172.00	-x,1-y,1-z
	O3–H3···O4	0.9800	1.6700	2.6531(18)	175.00	2-x,2-y,-z
	C13–H13···F6	1.0800	2.3300	3.255(2)	142.00	-1+x,1+y,z
245TFBA	O2–H2···O1	0.9800	1.6800	2.6577(16)	174.00	2-x,2-y,1-z
345TFBA(I)	O2–H2···O1	0.9800	1.6700	2.651(3)	176.00	1-x,2-y,-z
234TFBA	O2–H2···O4	0.9800	1.6600	2.623(3)	166.00	1-x,-y,-z
	C10–H10···F1	1.0800	2.4500	3.359(4)	141.00	-1+x,-1+y,z
	C10–H10···O2	1.0800	2.3200	3.319(4)	152.00	-1+x,-1+y,z
236TFBA	O1–H1···O2	0.9800	1.6500	2.628(2)	172.00	-1-x,1-y,-z
	C4–H4···O2	1.0800	2.3800	3.381(3)	153.00	-x,-1/2+y,1/2-z
235TFBA	O1–H1···O2	0.9800	1.6800	2.656(3)	170.00	1-x,-y,2-z
	C4–H4···F1A	1.0800	2.4300	3.287(3)	135.00	-1+x,1/2-y,-1/2+z
345TFBA(II)	O1–H1···O2	0.9800	1.6500	2.633(2)	175.00	-1-x,1-y,2-z
	O3–H3···O4	0.9800	1.6400	2.618(2)	177.00	2-x,-y,1-z
	C9–H9···F1	1.0800	2.3500	3.380(2)	158.00	1-x,1-y,1-z
24SS	O1–H1···O2	0.9800	1.6500	2.623(2)	167.00	3-x,1-y,-z
	C3–H3···F2	1.0800	2.4200	3.219(3)	129.00	1-x,1/2+y,1/2-z

C5–H5···O1	1.0800	2.4000	3.379(3)	150.00	-1+x,y,z
C6–H6···O2	1.0800	2.5000	3.474(3)	150.00	2-x,-y,-z

T-2: Crystallographic tables of the experimental structures

Compound	**24DFBA**	**25DFBA**	**34DFBA**	**234TFBA**	**235TFBA**	**236TFBA**
CCDC No.	881258	881259	881260	881261	881262	881263
Molecular formula	$C_7H_3F_2O_2$	$C_7H_3F_2O_2$	$C_7H_3F_2O_2$	$C_7H_3F_3O_2$	$C_7H_2F_3O_2$	$C_7H_2F_3O_2$
Formula weight	157.09	157.09	157.09	176.09	175.09	175.09
Crystal system	Monoclinic	Monoclinic	Monoclinic	Monoclinic	Monoclinic	Monoclinic
Space group	$P2_1/c$	$P2_1/c$	$P2_1/n$	$P2_1/c$	$P2_1/c$	$P2_1/c$
***a* (Å)**	3.6912 (2)	3.6960(5)	5.5154(12)	5.112(5)	3.6704(9)	6.9301(11)
***b* (Å)**	6.4038 (3)	14.768(2)	5.0528(11)	5.193(5)	15.516(4)	13.517(2)
***c* (Å)**	26.3062 (15)	11.9174(16)	23.144(5)	24.39(2)	11.882(3)	7.0809(12)
***α* (°)**	90	90	90	90	90	90
***β* (°)**	92.908 (5)	93.025(7)	94.792(7)	95.477(11)	93.734(7)	94.132(7)
***γ* (°)**	90	90	90	90	90	90
***V* (Å^3)**	621.02 (6)	649.58 (15)	642.7(2)	644.5(10)	675.2(3)	661.57(18)
***Z*,*Z*`**	4,1	4,1	4,1	4,1	4,1	4,1
***ρcalc* (g/cm^3)**	1.68	1.596	1.623	1.815	1.722	1.758
***F*(000)**	316	312	316	352	348	348
***μ* (mm^{-1})**	0.16	0.15	0.16	0.19	0.18	0.18
***T* (K)**	120(2)	150(2)	150(2)	150(2)	150(2)	150(2)
***λ* (Å)**	0.71073	0.71073	0.71073	0.71073	0.71073	0.71073
Reflns. collected	3872	6701	6090	6117	6781	6431
Unique reflns.	1265	1487	1474	1496	1546	1512
Completeness (%)	99.8	99.9	99.9	99.3	99.8	99.9
R_{int}	0.025	0.027	0.041	0.056	0.059	0.045
R_1 (F^2)	0.046	0.042	0.049	0.041	0.052	0.048
wR_2(F^2)	0.114	0.122	0.127	0.133	0.13	0.13
Goodness-of-fit	1.18	1.06	1.04	1.14	1.04	1.06
Resolution (*2θ*)	52	54	54	54	54	54

Compound	**245TFBA**	**345TFBA(I)**	**345TFBA(II)**	**2345TeFBA**	**24SS**
CCDC No.	881264	881265	881266	881267	881268
Molecular formula	$C_7H_3F_3O_2$	$C_7H_3F_3O_2$	$C_7H_2F_3O_2$	$C_7H_2F_4O_2$	$C_7H_4F_1O_2$
Formula weight	176.09	176.09	176.09	193.58	139.1
Crystal system	Monoclinic	Monoclinic	Triclinic	Triclinic	Monoclinic
Space group	$P2_1/c$	$P2_1/n$	$P\bar{1}$	$P\bar{1}$	$P2_1/c$
a (Å)	7.1042(7)	5.1190(14)	6.5824(9)	4.7350(6)	5.0609(10)
b (Å)	7.1852(7)	6.4295(17)	8.3582(11)	5.9700(7)	5.3645(11)
c (Å)	12.8037(13)	20.343(6)	12.7456(17)	24.651(3)	22.703(5)
α (°)	90	90	105.719(7)	84.773(6)	90
β (°)	102.453(7)	95.574(7)	91.945(7)	87.020(6)	95.007(7)
γ (°)	90	90	98.065(7)	82.885(6)	90
V (Å^3)	638.19(11)	666.4(3)	666.42(16)	688.00(15)	614.0(2)
Z,Z`	4,1	4,1	2,2	2,2	4,1
ρcalc (g/cm^3)	1.833	1.755	1.755	1.869	1.505
F(000)	352	352	352	382	284
μ. (mm^{-1})	0.19	0.18	0.18	0.21	0.13
T (K)	150(2)	150(2)	150(2)	150(2)	150(2)
λ (Å)	0.71073	0.71073	0.71073	0.71073	0.71073
Reflns. collected	6560	5694	7029	7297	5762
Unique reflns.	1466	1532	3058	3149	1408
Completeness (%)	99.9	99.5	99.8	99.8	99.9
R_{int}	0.022	0.082	0.024	0.024	0.041
R_1 (F^2)	0.034	0.06	0.041	0.044	0.055
wR_2(F^2)	0.091	0.161	0.11	0.118	0.144
Goodness-of-fit	1.04	1.06	1.07	1.04	1.08
Resolution (2θ)	54	54	54	54	54

T-3: Computational details for ESP charge calculation and Polymorph prediction protocols

- ❖ Computational details for ESP charge calculation

DMol3 Run Status:

Task: Property
Basis: dnp
Atom Rcut: 3.700 Angstrom
Functional: pbe
Harris: off
Integration Grid: fine
Occupation: fermi
Pseudopotential: none

- Summary of ESP fitting calculations: For BA (COMPASS26)

Number of points : 54644
Total integration weight: 22579.48
Spacing between : 0.25 Ang
Sigma : 0.1471E-02
RMS of V(exact) : 0.1266E-01
RRMS fit : 11.62 %

ESP-fitted charges :

n	Elem	chg	vdW(in)	vdW(ex)
1	H	0.140	1.30	2.80
2	C	-0.099	2.00	3.50
3	H	0.133	1.30	2.80
4	C	-0.209	2.00	3.50
5	H	0.166	1.30	2.80
6	C	0.147	2.00	3.50
7	C	-0.209	2.00	3.50
8	H	0.158	1.30	2.80
9	C	-0.088	2.00	3.50
10	H	0.132	1.30	2.80
11	C	-0.150	2.00	3.50
12	C	0.474	2.00	3.50
13	O	-0.473	1.72	3.22
14	O	-0.545	1.72	3.22
15	H	0.424	1.30	2.80

- Summary of ESP fitting calculations: For 4FBA (DREIDING)

Number of points : 64752
Total integration weight : 23528.34
Spacing between : 0.25 Ang
Sigma : 0.1382E-02
RMS of V(exact) : 0.1151E-01
RRMS fit : 12.01 %

ESP-fitted charges :

n	Elem	chg	vdW(in)	vdW(ex)
1	F	-0.190	1.72	3.22
2	C	-0.248	2.00	3.50
3	H	0.180	1.30	2.80
4	C	-0.158	2.00	3.50
5	H	0.176	1.30	2.80
6	C	0.050	2.00	3.50
7	C	-0.156	2.00	3.50
8	H	0.168	1.30	2.80
9	C	-0.238	2.00	3.50
10	H	0.179	1.30	2.80
11	C	0.332	2.00	3.50
12	C	0.515	2.00	3.50
13	O	-0.482	1.72	3.22
14	O	-0.551	1.72	3.22
15	H	0.424	1.30	2.80

❖ Computational details of Polymorph Prediction protocols

<u>For BA (COMPASS26)</u>

```
Mode              : New calculation
Protocol          : Packing >> Clustering >> Geometry Optimization >>
Clustering
Version           : 5.0
Build date        : Oct 22 2009
Host              : GRDLAB-PC11
Operating system  : MSWin32
Task started      : Fri Apr 06 10:49:26 2012


---- Packing parameters ----

Search algorithm                : MC Simulated Annealing
```

```
Maximum number of steps         : 7000
Explore torsions                : No
Preoptimize structures          : No
Steps to accept before cooling  : 12
Minimum move factor             : 0.1000E-08
Heating factor                  : 0.02500
Maximum temperature             : 100000.0 K
Minimum temperature             : 300.0 K


---- Cluster analysis parameters ----

Cluster grouping                : Forcefield type
Cutoff                          : 7.000
Number of bins                  : 140
Tolerance                       : 0.1300
Maximum number of clusters      : 2000

---- Geometry optimization parameters ----

Algorithm                       : Smart
Convergence tolerance:
  Energy                        : 0.0001 kcal/mol
  Force                         : 0.005 kcal/mol/A
  Stress                        : 0.005 GPa
  Displacement                  : 5e-005 A
Maximum number of iterations    : 5000
External pressure               : 0 GPa
Motion groups rigid             : NO
Optimize cell                   : YES

---- Energy parameters ----

Forcefield                      : COMPASS26
Electrostatic terms:
  Summation method              : Ewald
  Accuracy                      : 0.0001 kcal/mol
  Buffer width                  : 0.5 A

van der Waals terms:
  Summation method              : Ewald
  Accuracy                      : 0.0001 kcal/mol
  Repulsive cutoff              : 6 A
  Buffer width                  : 0.5 A


Calculation summary:

Number of frames searched           : 2000
Number of clusters found            : 44
Number of frames put into a cluster : 2000 (100 %).
```

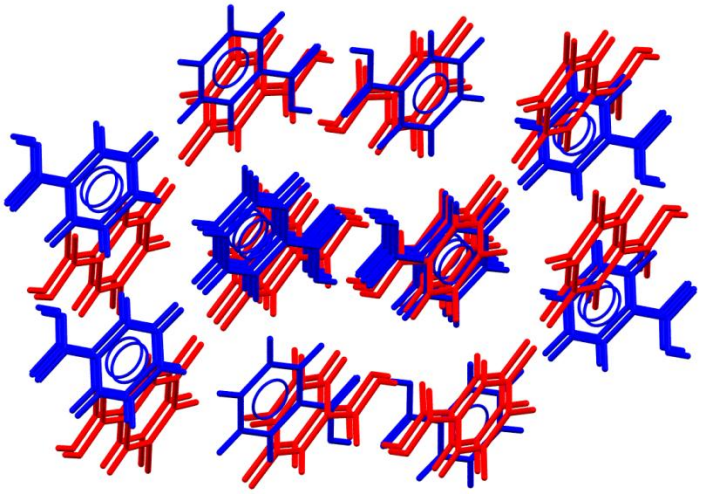
```
Task terminated               : Fri Apr 06 11:42:39 2012
Total CPU time used           : 53:08 minutes

Termination status : Normal
```

For 4FBA (DREIDING)

```
Mode              : New calculation
Protocol          : Packing >> Clustering >> Geometry Optimization >>
Clustering
Version           : 5.0
Build date        : Oct 22 2009
Host              : GRDLAB-PC11
Operating system  : MSWin32
Task started      : Tue Apr 17 21:57:06 2012



---- Packing parameters ----

Search algorithm                : MC Simulated Annealing
Maximum number of steps         : 7000
Explore torsions                : No
Preoptimize structures          : No
Steps to accept before cooling  : 12
Minimum move factor             : 0.1000E-08
Heating factor                  : 0.02500
Maximum temperature             : 100000.0 K
Minimum temperature             : 300.0 K


---- Cluster analysis parameters ----

Cluster grouping                : Forcefield type
Cutoff                          : 7.000
Number of bins                  : 140
Tolerance                       : 0.1300
Maximum number of clusters      : 2000

---- Geometry optimization parameters ----

Algorithm                       : Smart
Convergence tolerance:
  Energy                        : 0.0001 kcal/mol
  Force                         : 0.005 kcal/mol/A
  Stress                        : 0.005 GPa
  Displacement                  : 5e-005 A
Maximum number of iterations    : 5000
External pressure               : 0 GPa
```

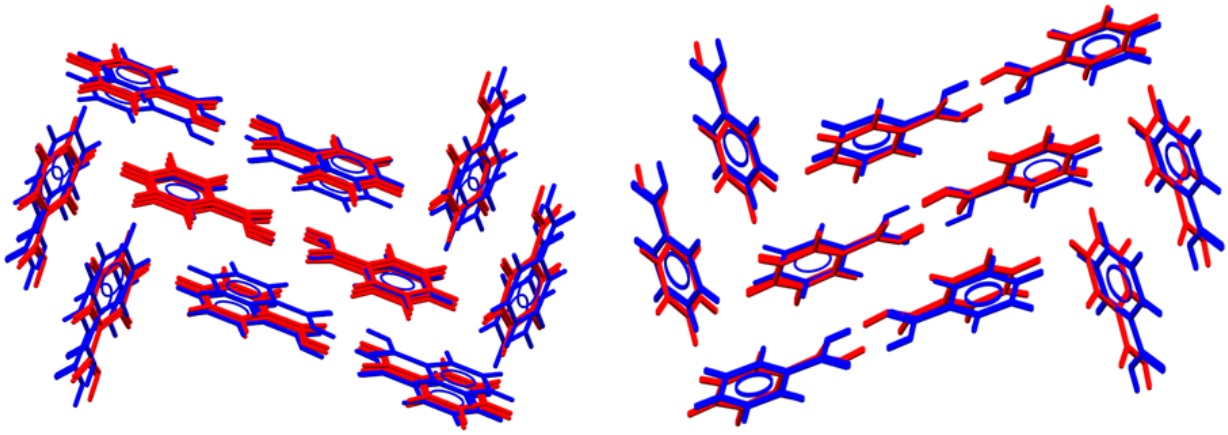
```
Motion groups rigid             : NO
Optimize cell                   : YES

---- Energy parameters ----

Forcefield                      : Dreiding
Electrostatic terms:
  Summation method              : Ewald
  Accuracy                      : 0.0001 kcal/mol
  Buffer width                  : 0.5 A

van der Waals terms:
  Summation method              : Ewald
  Accuracy                      : 0.0001 kcal/mol
  Repulsive cutoff              : 6 A
  Buffer width                  : 0.5 A

Hydrogen bond terms:
  Summation method              : Atom based
  Truncation method             : Cubic spline
  Cutoff distance               : 4.5 A
  Spline width                  : 0.5 A
  Buffer width                  : 0.5 A


Calculation summary:

Number of frames searched           : 2000
Number of clusters found            : 52
Number of frames put into a cluster : 2000 (100 %).



Task terminated               : Tue Apr 17 22:43:07 2012
Total CPU time used           : 45:53 minutes

Termination status : Normal
```

F-1: Experimental as well as computed structure overlay diagrams



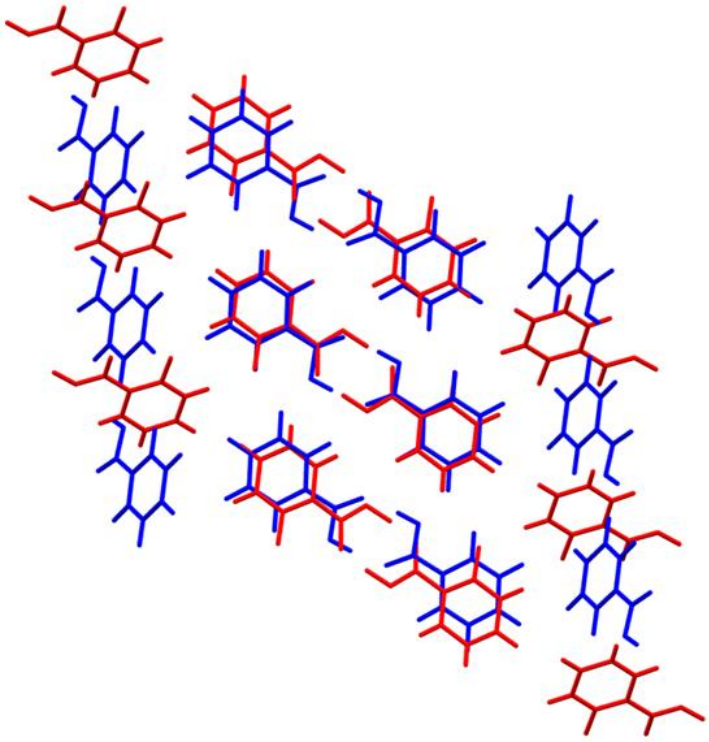
Overlay diagram of BA (exp) and 345TFBA (exp)

blue: BA, red: 345TFBA



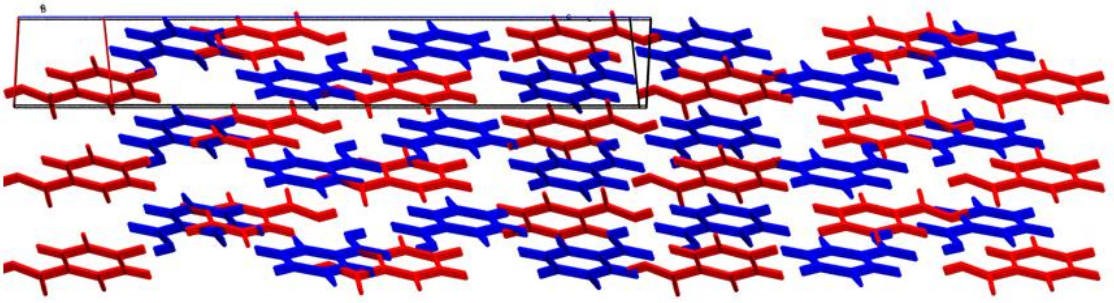
(a) **(b)**

Overlay diagrams of (a) BA (exp) and 5th rank structure (b) BA (exp) and 23rd rank, blue: BA(exp), red: 5th and 23rd respectively



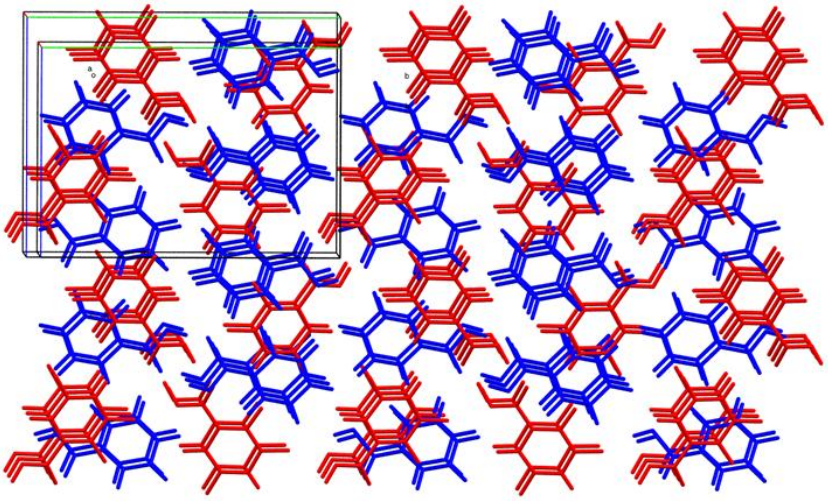
Overlay diagram of 2FBA (6-4-25 structure) and 99th ranked BA.

blue: BA red: 2FBA



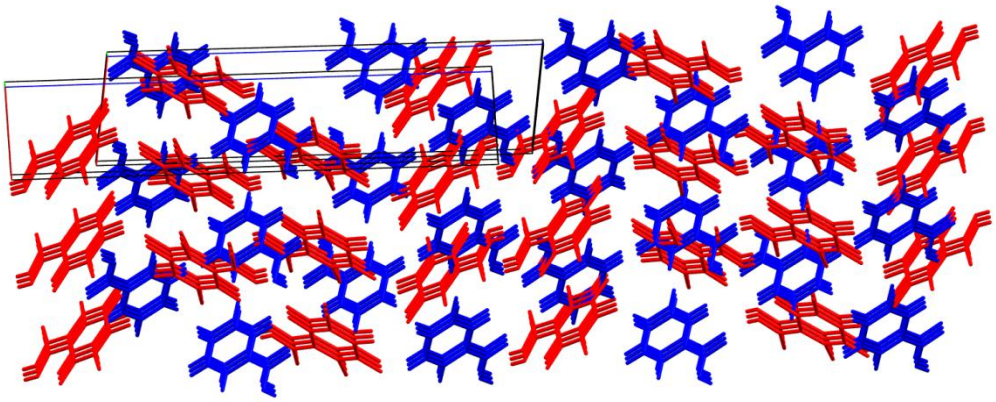
Unit cell overlay diagram of 4FBA (4-6-26 structure) and 55th ranked BA.

blue: BA red: 4FBA



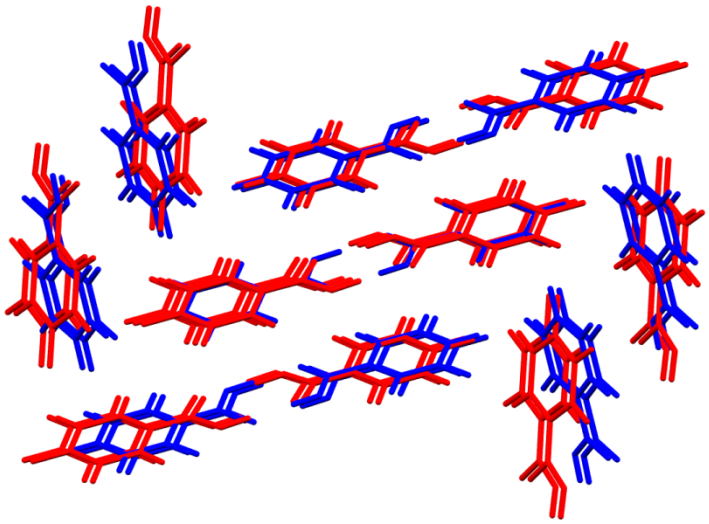
Unit cell overlay diagram of 235TFBA (4-14-12 structure) and 56th ranked BA.

blue: BA red: 235TFBA



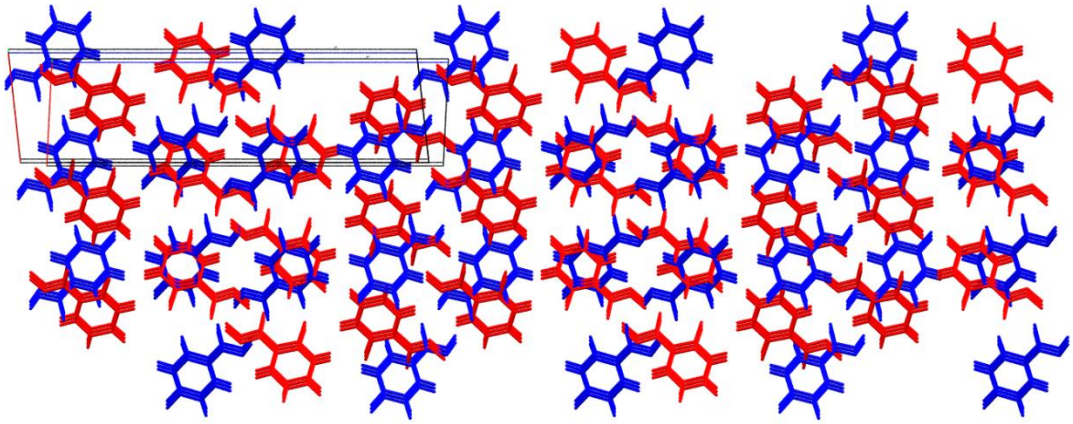
Unit cell overlay diagram of BA (exp) and 2345TeFBA (exp).

blue: BA red: 2345TeFBA



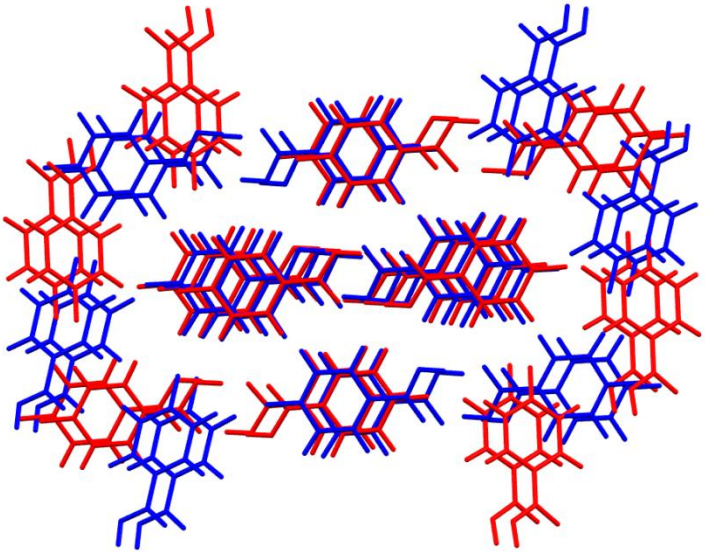
Overlay diagram of BA (exp) and 38th ranked 4FBA.

blue: BA red: 4FBA



Unit cell overlay diagram of 2FBA (exp) and 29th ranked 4FBA.

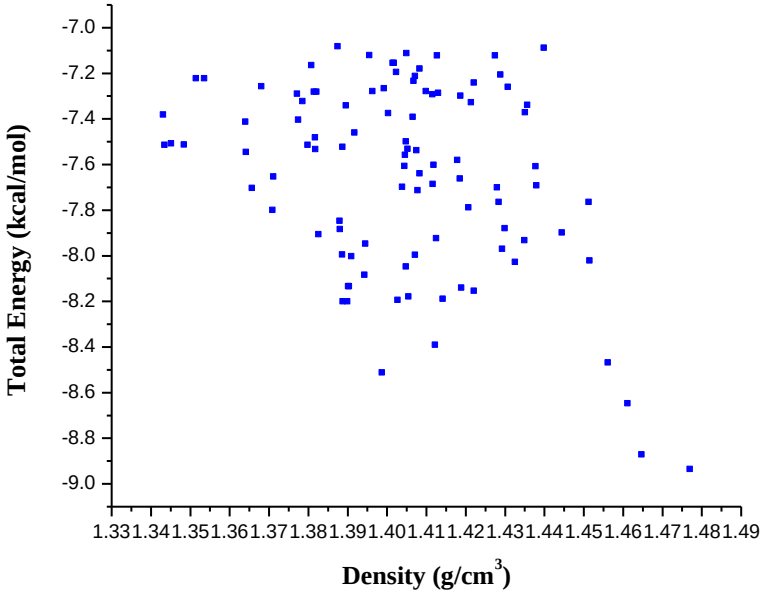
blue: 2FBA red: 29th ranked 4FBA



Overlay diagram of 4FBA (exp) and 9th ranked 4FBA.

blue: 4FBA (exp) red: 9th ranked 4FBA

F-2: Energy density plot of top 100 predicted structures of 4-fluorobenzoic Acid



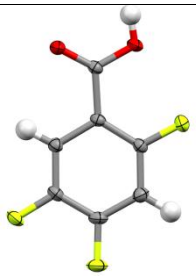
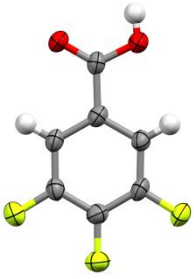
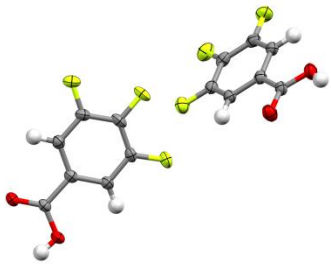
(To simplify the analysis, the CSP was restricted only to the space groups $P2_1/c$ and $P\bar{1}$)

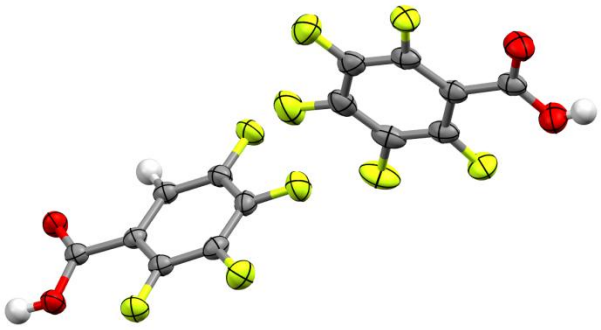
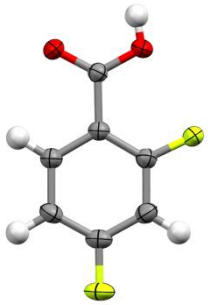
Ranking of different structure types in CSP of 4FBA

DREIDING	
(5-5-22)	3rd, 38th
(6-4-25)	29th, 31st
(4-6-26)	6th, 9th

** (red colored ranks are experimental structures)

F-3: ORTEP diagrams for the asymmetric unit of the crystal structures (ellipsoids are drawn at 50% probability). Disordered structures are indicated with an asterix. Only the asymmetric unit is shown.

 **2,4-Difluorobenzoic Acid***	 **2,5-Difluorobenzoic Acid***	 **3,4-Difluorobenzoic Acid***
 **2,3.4-Trifluorobenzoic Acid**	 **2,3,5-Trifluorobenzoic Acid***	 **2,3,6-Trifluorobenzoic Acid***
 **2,4,5-Trifluorobenzoic Acid**	 **3,4,5-Trifluorobenzoic Acid(I)**	 **3,4,5-Trifluorobenzoic Acid (II)**

	
2,3,4,5-Tetrafluorobenzoic Acid*	**24-Solid Solution (24SS)***