Electronic Supplementary Material (ESI) for CrystEngComm.

Exploring the structural landscape with 'partial' fluoro-substitution as a probe

Shaunak Chakraborty and Gautam R. Desiraju*

Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore-560 012

Contents:

S1: PXRD and DSC data

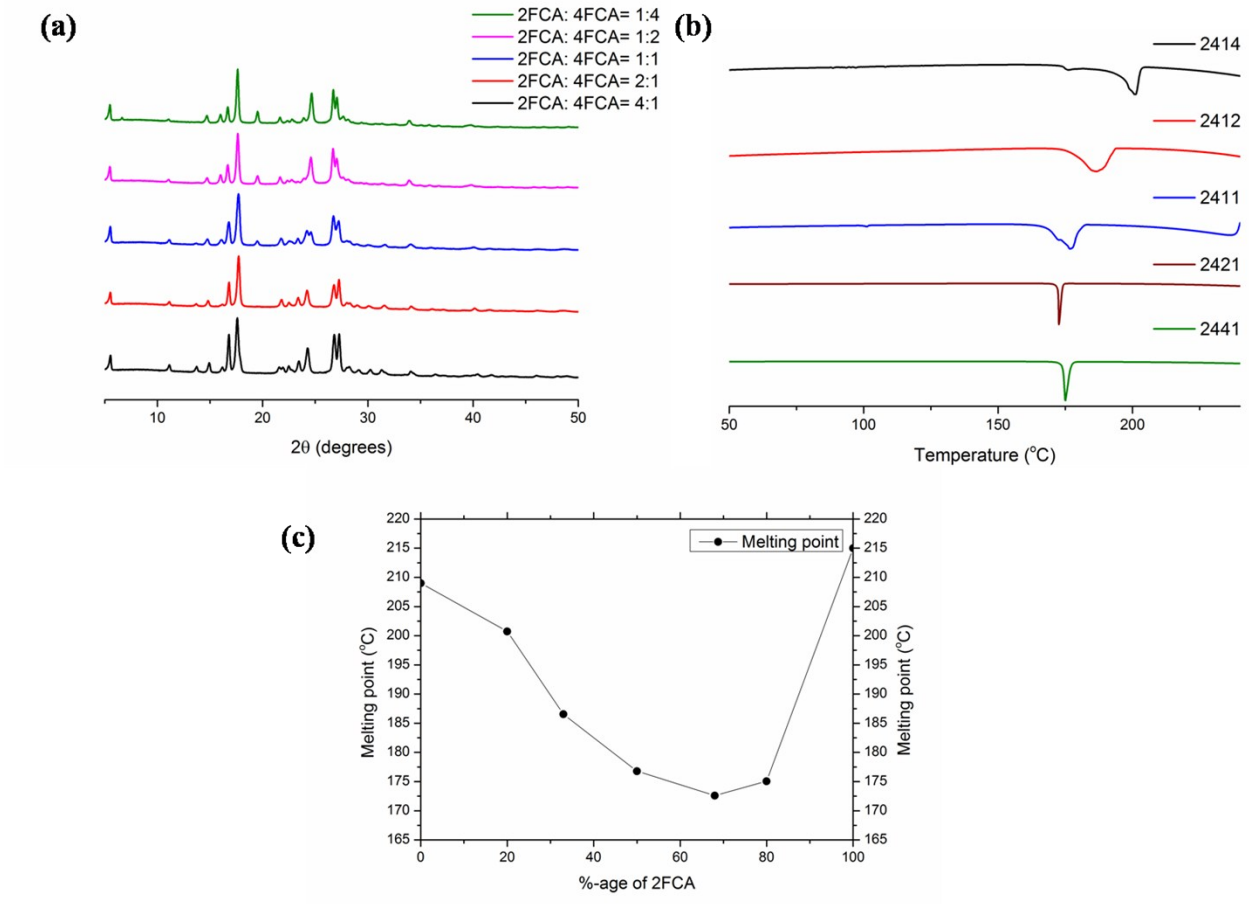


Fig. S1: (a) PXRD patterns of 2FCA and 4FCA ground together in various ratios and (b) DSC profiles of the same. (c) shows the variation of melting points of the ground mixtures as obtained from DSC with increase in percentage of 2FCA. The DSC profiles indicate the formation of a single phase, and the melting points of the mixtures are lower than those of the pure components.

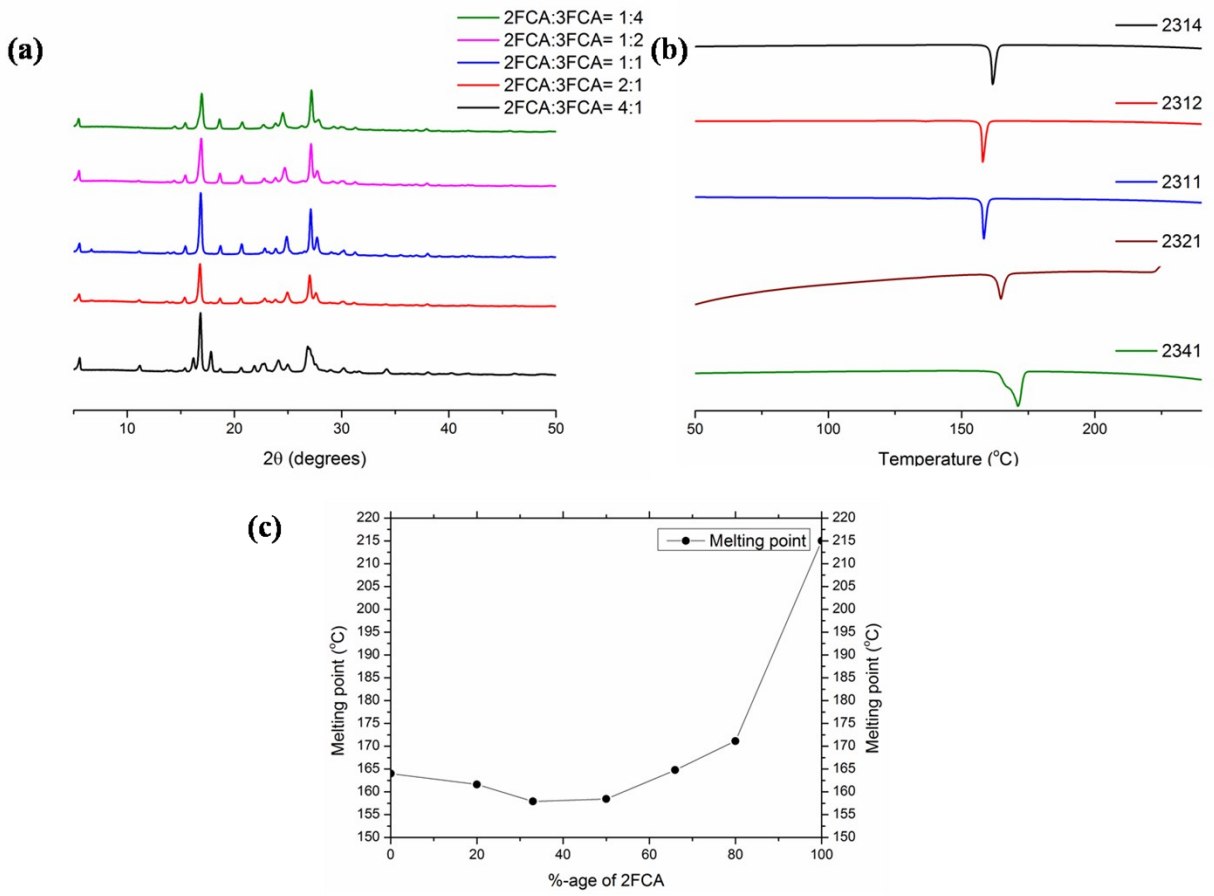


Fig. S2: (a) PXRD patterns of 2FCA and 3FCA ground together in various ratios and (b) DSC profiles of the same. (c) shows the variation of melting points of the ground mixtures as obtained from DSC with increase in percentage of 2FCA. The DSC profiles indicate the formation of a single phase, and the melting points of the mixtures are lower than those of the pure components.

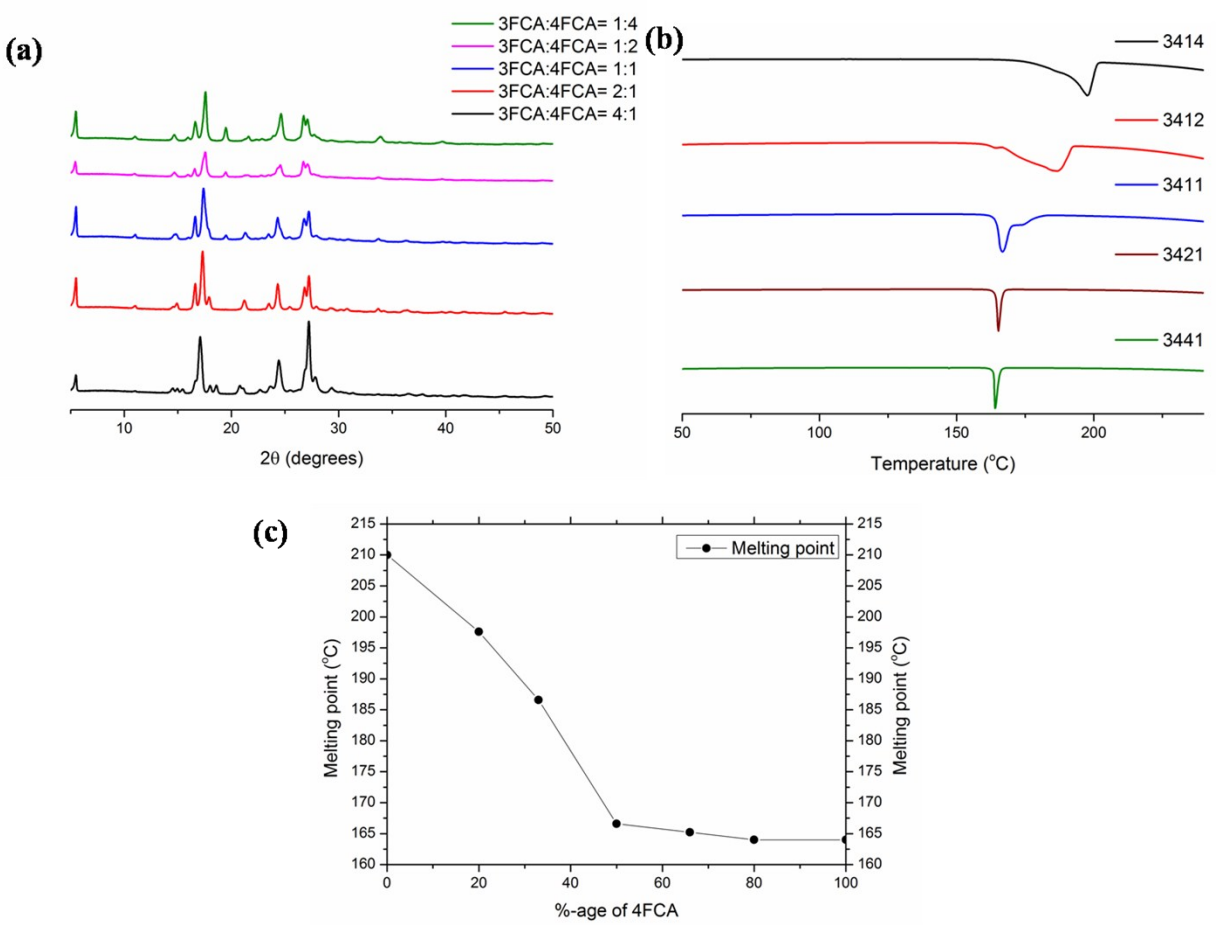


Fig. S3: (a) PXRD patterns of 3FCA and 4FCA ground together in various ratios and (b) DSC profiles of the same. (c) shows the variation of melting points of the ground mixtures as obtained from DSC with increase in percentage of 4FCA. The DSC profiles indicate the formation of a single phase, and the melting points of the mixtures are lower than those of the pure components.

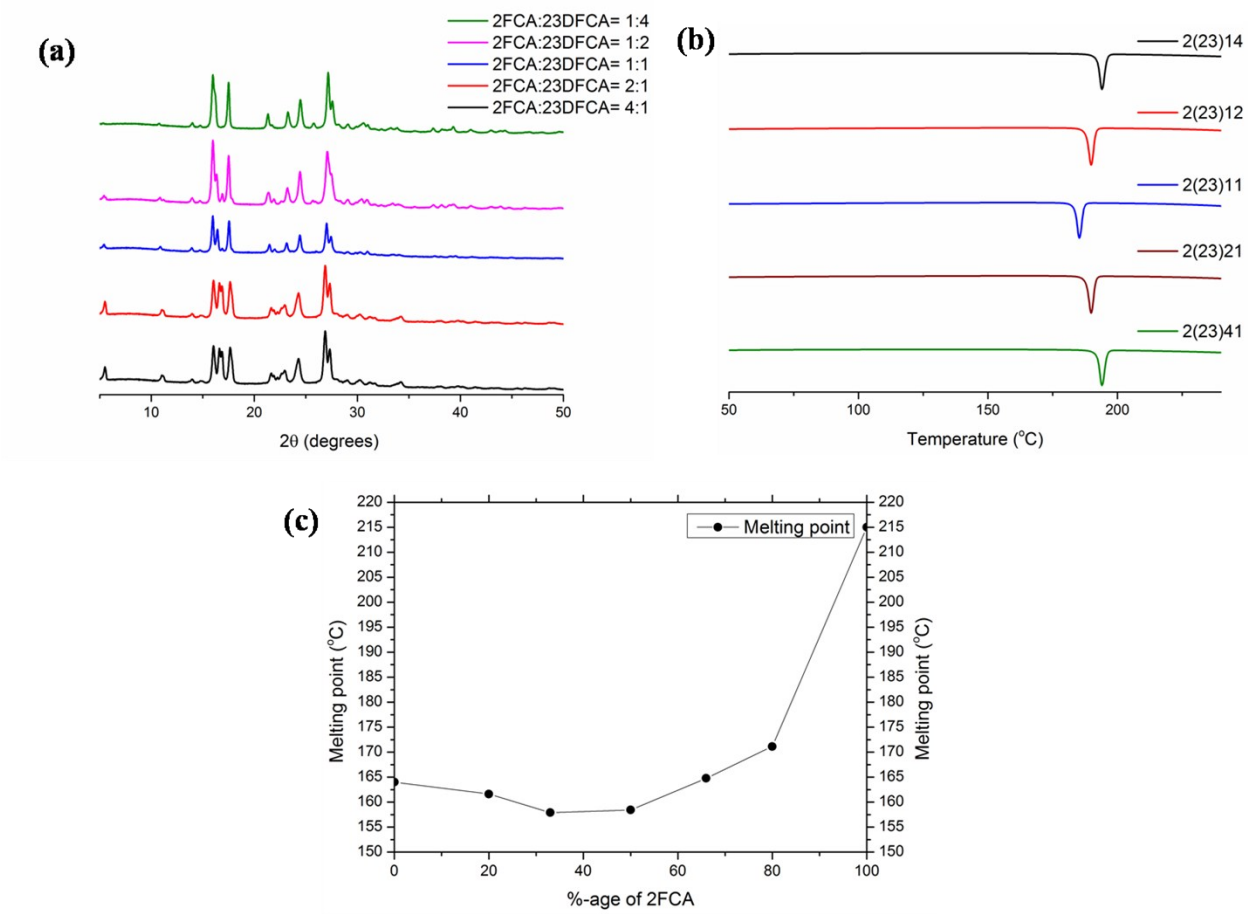


Fig. S4: (a) PXRD patterns of 2FCA and 23DFCA ground together in various ratios and (b) DSC profiles of the same. (c) shows the variation of melting points of the ground mixtures as obtained from DSC with increase in percentage of 2FCA. The DSC profiles indicate the formation of a single phase, and the melting points of the mixtures are lower than those of the pure components.

S2: Computational protocol

We used the Polymorph module incorporated in Materials Studio 6.0, with the unsubstituted CA molecule as input. We used the DREIDING force field and limited our search to the space groups $P2_1/c$ and $P\square$, since these two are the two commonest space groups organic molecules crystallize in.

S3: ESP-fitted charges (DMol^3)

ESP-fitted charges :

n	Elem	chg	vdW(in)	vdW(ex)
1	C	0.687	2.00	3.50
2	C	-0.236	2.00	3.50
3	C	-0.264	2.00	3.50
4	C	0.316	2.00	3.50
5	C	-0.275	2.00	3.50
6	C	-0.070	2.00	3.50
7	C	-0.178	2.00	3.50
8	C	-0.083	2.00	3.50
9	C	-0.257	2.00	3.50
10	O	-0.602	1.72	3.22
11	O	-0.546	1.72	3.22
12	H	0.427	1.30	2.80
13	H	0.164	1.30	2.80
14	H	0.184	1.30	2.80
15	H	0.165	1.30	2.80
16	H	0.131	1.30	2.80
17	H	0.147	1.30	2.80
18	H	0.135	1.30	2.80
19	H	0.156	1.30	2.80

S4: CSP Protocol

```
              Polymorph

Mode              : New calculation
Protocol          : Packing >> Clustering >> Geometry Optimization >>
Clustering
Version           : 6.0
Build date        : Dec  8 2011
Host              : SSCU-PC
Operating system  : Windows
Task started      : Tue Oct 27 23:16:10 2015



---- 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      : All clusters

---- 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    : 500
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


#####################################
Processing space group : P21/C
#####################################



---- Packing ----

Space group                     : P 1 21/C 1
Random number seed              : 464702587
Cooling factor                  : 0.00100
Number of trials made           : 5534
Number of trials saved          : 3767 (68.1%)
Maximum temperature achieved    : 11741.0 K
Final temperature               : 299.8 K

---- Cluster analysis ----

Lowest energy clusters (20 maximum):

Cluster   Frame   Cluster size  Energy (kcal/mol/asym. cell)
    1     3663          4       -2.812
    2     3188          2       -2.720
    3     3741          1       -2.714
    4     3678          4       -2.694
    5     3695          1       -2.693
    6     3186          3       -2.609
    7     3640          1       -2.582
    8     3694          1       -2.573
    9     3176          1       -2.562
   10     3636          4       -2.552
```

11	3674	1	-2.532
12	3665	1	-2.507
13	3693	1	-2.505
14	3175	1	-2.494
15	3740	1	-2.475
16	3744	3	-2.464
17	3639	2	-2.438
18	3174	1	-2.314
19	3739	1	-2.307
20	2668	1	-2.282

Calculation summary:

Number of frames searched : 3767
Number of clusters found : 3533
Number of frames put into a cluster : 3767 (100 %).

---- Geometry optimization ----

WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.

Total number of frames : 3533
Frames with lowest energy:

Order	Frame	Energy (kcal/mol/asym. cell)
1	2700	-40.869
2	3406	-40.869
3	3249	-40.869
4	2030	-40.869
5	1537	-40.869
6	1653	-40.869
7	2381	-40.869
8	3213	-40.869
9	443	-40.861
10	318	-40.640
11	1906	-40.589
12	3336	-40.589
13	3483	-40.589
14	2152	-40.589
15	197	-40.589
16	2717	-40.589
17	815	-40.589
18	2909	-40.589
19	2639	-40.513
20	916	-40.406

---- Cluster analysis ----

Lowest energy clusters (20 maximum):

Cluster	Frame	Cluster size	Energy (kcal/mol/asym. cell)
1	2700	8	-40.869
2	443	1	-40.861
3	318	1	-40.640
4	1906	8	-40.589
5	2639	1	-40.513
6	916	13	-40.406
7	1726	1	-40.397
8	3323	1	-40.389

```
    9     2101          7       -40.375
   10      465         64       -40.226
   11      405         16       -40.211
   12     1729         13       -40.156
   13     2548          1       -40.152
   14     2171         14       -40.116
   15     3186          4       -40.115
   16      542          4       -40.092
   17     2879          7       -40.088
   18     1485          8       -40.030
   19     1441         18       -40.008
   20       33         30       -40.004

Calculation summary:

Number of frames searched           : 3533
Number of clusters found            : 745
Number of frames put into a cluster : 3533 (100 %).

#####################################
Processing space group : P-1
#####################################



---- Packing ----

Space group                     : P -1
Random number seed              : 485603013
Cooling factor                  : 0.00100
Number of trials made           : 4482
Number of trials saved          : 2605 (58.1%)
Maximum temperature achieved    : 3864.8 K
Final temperature               : 299.9 K

---- Cluster analysis ----

Lowest energy clusters (20 maximum):

Cluster   Frame   Cluster size  Energy (kcal/mol/asym. cell)
    1     2169         13       -3.313
    2     2176          5       -3.289
    3     2204          1       -3.286
    4     2200          1       -3.214
    5     2156          5       -3.162
    6     2203          1       -3.135
    7     2163          2       -3.111
    8     2596          7       -3.105
    9     2206          1       -3.072
   10     2471         12       -3.071
   11     2201          1       -3.061
   12     2550          3       -3.057
   13     2202          1       -3.047
   14     2585          4       -3.039
   15     2150          1       -3.034
   16     2478          3       -3.019
   17     2567          4       -3.013
   18     2151          1       -2.995
   19     2147          1       -2.993
   20     2593          1       -2.989

Calculation summary:

Number of frames searched           : 2605
```

Number of clusters found   : 2349
Number of frames put into a cluster : 2605 (100 %).

---- Geometry optimization ----

WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.
WARNING No progress. Optimization with current method stopped.

Total number of frames   : 2349
Frames with lowest energy:

Order	Frame	Energy (kcal/mol/asym. cell)
1	1176	-40.646
2	2072	-40.646
3	1965	-40.646
4	1651	-40.646
5	272	-40.646
6	999	-40.646
7	621	-40.646
8	1841	-40.646
9	1893	-40.646
10	1231	-40.646
11	1780	-40.646
12	1641	-40.646
13	924	-40.646
14	554	-40.646
15	1074	-40.646
16	588	-40.646
17	1324	-40.646
18	1672	-40.646
19	2336	-40.646
20	2176	-40.646

---- Cluster analysis ----

Lowest energy clusters (20 maximum):

Cluster	Frame	Cluster size	Energy (kcal/mol/asym. cell)
1	1176	287	-40.646
2	708	2	-40.646
3	710	1	-40.645
4	1987	1	-40.642
5	1861	1	-40.638
6	967	1	-40.635
7	2041	1	-40.598
8	2249	1	-40.566
9	2114	133	-39.912
10	1633	2	-39.912
11	1496	1	-39.910
12	643	1	-39.907
13	1479	1	-39.906
14	1963	1	-39.902
15	533	1	-39.877
16	1974	1	-39.869
17	1375	1	-39.854
18	1318	1	-39.850
19	1516	1	-39.836
20	2111	1	-39.814

Calculation summary:

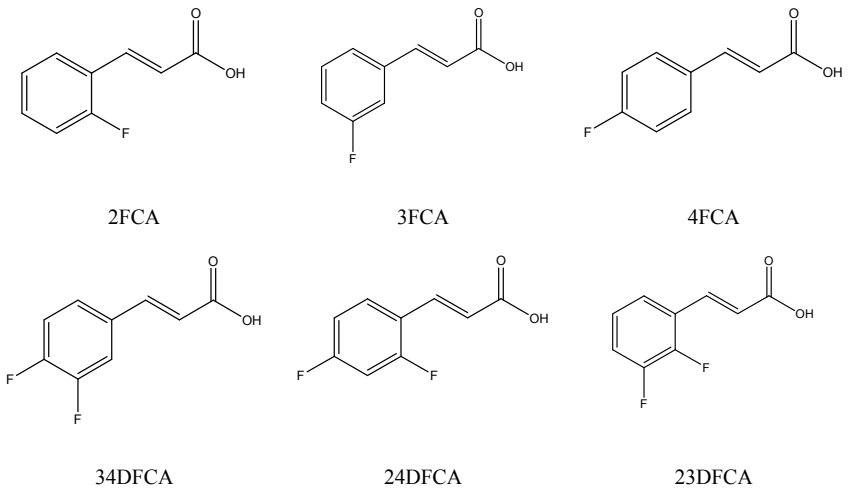
```
Number of frames searched           : 2349
Number of clusters found            : 169
Number of frames put into a cluster : 2349 (100 %).


Task terminated              : Wed Oct 28 00:27:13 2015
Total CPU time used          : 1:04:00 hours

Termination status : Normal
```

S6: Refinement protocols

We carried out our SS experiments with the following six CA derivatives– 2FCA, 3FCA, 4FCA, 24DFCA, 34DFCA and 23DFCA (Scheme S1). The F-atom occupancies were assigned by the refinement protocols described below.



Scheme S1

1. **Refinement of 24SS:** Structure solution by direct methods showed a CA skeleton with significant residuals at the 2 and 4 positions. Both peaks were labeled as F atoms, but satisfactory refinement could only be achieved when their occupancies were assumed to be fractional. Underlying our refinement protocol is the fact that the asymmetric unit in each SS exhibits only one (averaged) molecule, which means that the 2FCA and 4FCA molecules share the same crystallographic site. Therefore, the asymmetric unit must

contain a total of one F atom, i.e., the occupancies of the F atoms at the 2 and the 4 positions must add up to 1 (Scheme S2). If we assume the F-atom occupancy at the 2 position to be *x* and that at the 4 position to be *y*, they must be related by the equation:

$x + y = 1$

However, each F-atom site is occupied by a H atom as well, and the occupancies of the F atom and the H atom must add up to 1.

Site occupancy factor of the H atom at the 2-position = 1–*x*

Site occupancy factor of the H atom at the 4-position = 1–*y*

2. **Refinement of 23SS:** Structure solution by direct methods showed a CA skeleton with significant residuals at the 2 and 3 positions. Both peaks were labeled as F atoms, but satisfactory refinement could only be achieved when their occupancies were assumed to be fractional. Underlying our refinement protocol is the fact that the asymmetric unit in each SS exhibits only one (averaged) molecule, which means that the 2FCA and 3FCA molecules are distributed in the same crystallographic site. Therefore, the asymmetric unit must contain a total of one F atom, i.e., the occupancies of the F atoms at the 2- and the 3 positions must add up to 1 (Scheme S2). If we assume the F-atom occupancy at the 2 position to be *x* and that at the 3 position to be *y*, they must be related by the equation:

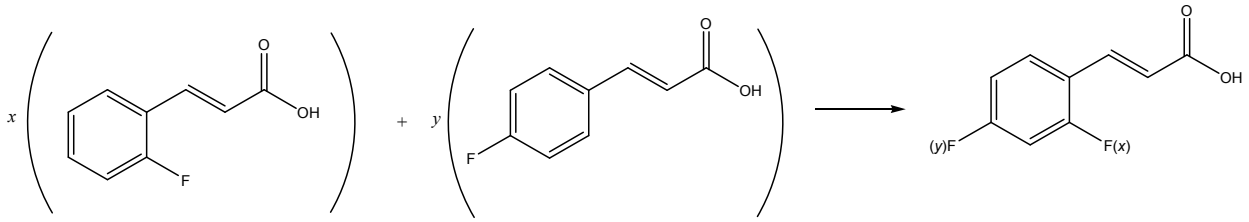
   $x + y = 1$

   However, each F-atom site is occupied by a H atom as well, and the occupancies of the F atom and the H atom must add up to 1.

   Site occupancy factor of the H atom at the 2-position = 1–*x*

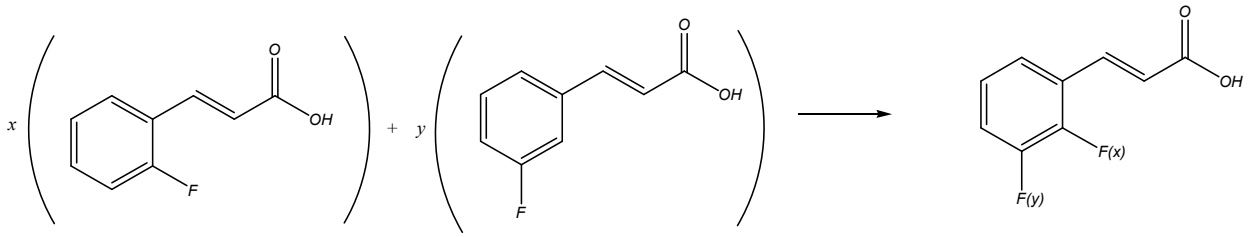
   Site occupancy factor of the H atom at the 3-position = 1–*y*

24SS



$x + y = 1$

23SS



$x + y = 1$

Scheme S2

3. **Refinement of 2(24)SS:** Structure solution by direct methods showed a CA skeleton with significant residuals at the 2 and 4 positions. Both peaks were labeled as F atoms. Underlying our refinement protocol is the fact that the asymmetric unit in each case contains only one (averaged) molecule, which means that there is one full F atom at the 2 position, and only the occupancy of the F atom at the 4 position is fractional. The latter F atom, however, shares the site with a H atom whose occupancy summed with that of the F atom must equal 1 (Scheme S3). The occupancy of this H atom also corresponds to the fraction of 24DFCA present in the crystal structure. Therefore, if the occupancy of this F atom is represented by *x* and that of the corresponding H atom by *y*, they must be related by the equation:

$x + y = 1$

4. **Refinement of 4(24)SS:** Structure solution by direct methods showed a CA skeleton with significant residuals at the 2 and 4 positions. Both peaks were assigned as F atoms. Underlying our refinement protocol is the fact that the asymmetric unit in each case contains only one (averaged) molecule, which means that there is one full F atom at the 4 position, and only the occupancy of the F atom at the 2 position is fractional. The latter F atom, however, shares the site with a H atom whose occupancy summed with that of the F atom must equal 1 (Scheme S3). The occupancy of this H atom also corresponds to the fraction of 24DFCA present in the crystal structure. Therefore, if the occupancy of this F atom is represented by *x* and that of the corresponding H atom by *y*, they must be related by the equation:

   $x + y = 1$

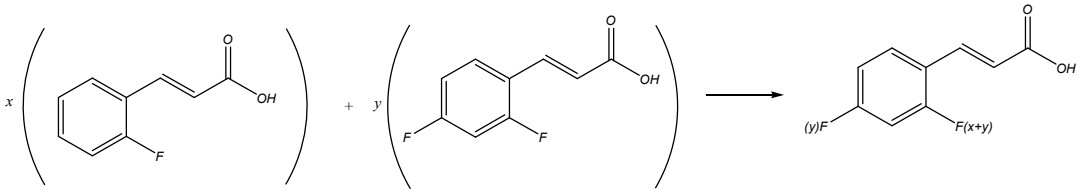
5. **Refinement of 2(23)SS:** Structure solution by direct methods showed a CA skeleton with significant residuals at the 2 and 3 positions. Both peaks were assigned as F atoms. Underlying our refinement protocol is the fact that the asymmetric unit in each case contains only one (averaged) molecule, which means that there is one full F atom at the 2 position, and only the occupancy of the F atom at the 3 position is fractional. The latter F atom, however, shares the site with a H atom whose occupancy summed with that of the F atom must equal 1 (Scheme S3). The occupancy of this H atom also corresponds to the fraction of 23DFCA present in the crystal structure. Therefore, if the occupancy of this F atom is represented by *x* and that of the corresponding H atom by *y*, they must be related by the equation:

$x + y = 1$

6. **Refinement of 3(23)SS:** Structure solution by direct methods showed a CA skeleton with significant residuals at the 2 and 3 positions. Both peaks were assigned as F atoms. Underlying our refinement protocol is the fact that the asymmetric unit in each case contains only one (averaged) molecule, which means that there is one full F atom at the 3 position, and only the occupancy of the F atom at the 2 position is fractional. The latter F atom, however, shares the site with a H atom whose occupancy summed with that of the F atom must equal 1 (Scheme S3). The occupancy of this H atom also corresponds to the fraction of 23DFCA present in the crystal structure. Therefore, if the occupancy of this F atom is represented by *x* and that of the corresponding H atom by *y*, they must be related by the equation:

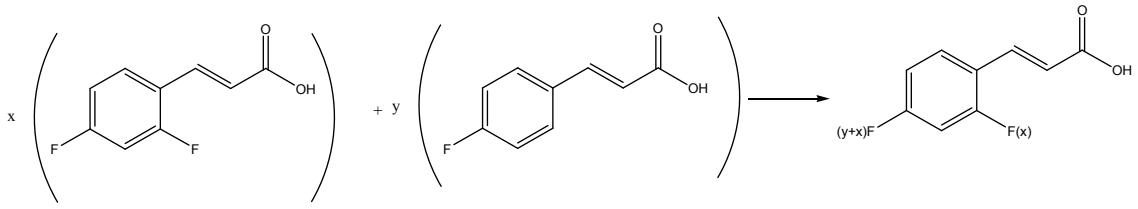
$x + y = 1$

2(24)SS



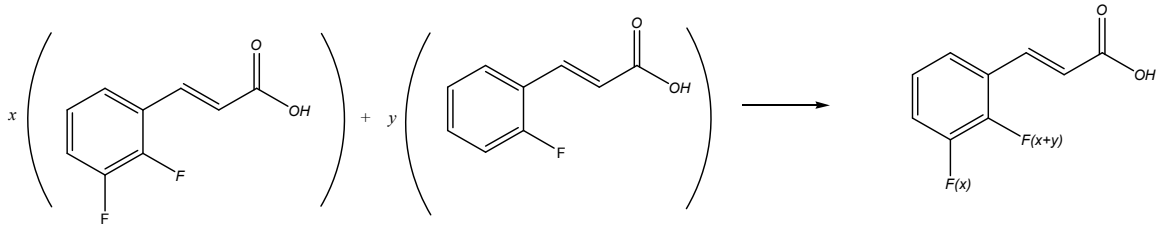
$x + y = 1$

4(24)SS



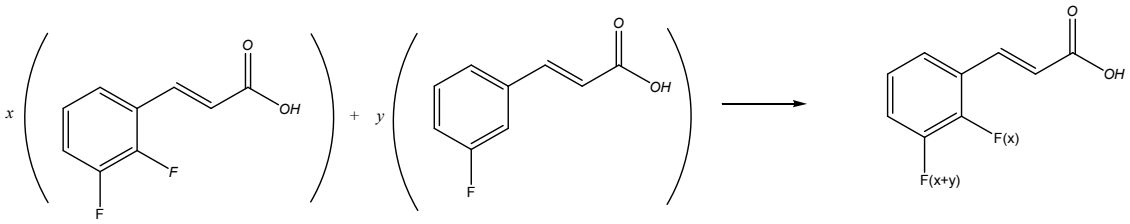
$x + y = 1$

2(23)SS



$x + y = 1$

3(23)SS



$x + y = 1$

Scheme S3

7. **Refinement of 3(34)SS:** Structure solution by direct methods showed a CA skeleton with three significant residuals at the 3, 4 and 5 positions. Underlying our refinement protocol is the fact that there is only one (averaged) molecule in the asymmetric unit, which indicates that the 3 position should contain one full F atom, and the 4 position a fractional one. However, the F atom in the two polymorphic forms of 3FCA is seen to be distributed between the 3 and the 5 positions, and it was expected, by the same token, that

the residuals at the 3 and 5 positions in this case would also correspond to a disordered F atom (Scheme S4). As mentioned before, the F atom at the 4 position should have a fractional occupancy, and a free refinement of it was performed accordingly. The occupancy of this F atom corresponds to the proportion of 34DFCA in the crystal structure. Each of these F atoms share their respective sites with one H atom whose occupancy is equal to that of the F atom subtracted from 1.

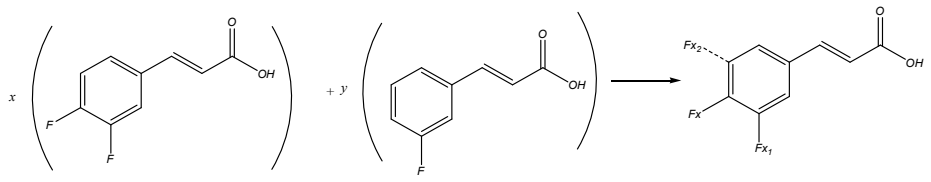
If the F-atom occupancy at the 3 position and that at the 5 position are denoted by x_1 and x_2 respectively, x_1 and x_2 must satisfy the equation:

$$x1 + x2 = 1$$

Therefore, occupancy of the H atom at 3 position = 1– x_1, and that of the H atom at 5 position = 1– x_2

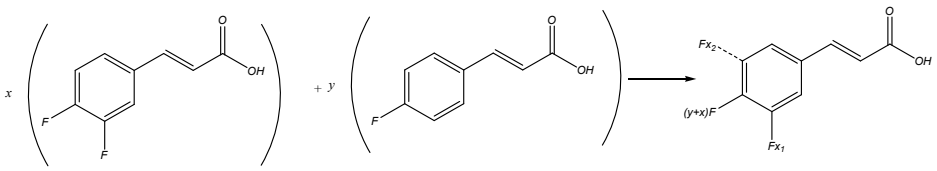
If the F-atom occupancy at the 4-position is denoted by *y,* then the accompanying H atom has an occupancy of 1–*y*

3(34)SS



$x1 + x2 = 1$

4(34)SS



$x + y = 1$

$x1 + x2 = x$

Scheme S4

8. **Refinement of 4(34)SS:** Structure solution by direct methods showed a CA skeleton with three significant residuals at the 3, 4 and 5 positions. Underlying the refinement protocol is the fact that there is only one (averaged) molecule in the asymmetric unit, which indicates that the 4 position should contain one full F atom, and the 3 position a fractional one. The F atom of 3FCA is seen to be distributed between the 3 and the 5 positions, and it was expected, by the same token, that the residuals at the 3 and 5 positions in this case would also correspond to a disordered F atom (Scheme S4). The occupancies of these two F atoms in the 3 and 5 positions summed would give the proportion of 34DFCA in the crystal structure. If the F-atom occupancies at the 3 and 5 positions are x_1 and x_2 respectively, and x moles of 34DFCA are cocrystallized with y moles of 4FCA, then x_1 and x_2 must satisfy the equation:

   $x1 + x2 = x$

   And x and y must satisfy the equation:

   $x + y = 1$

   Each of the F atoms at the 3 and 5 positions share their respective sites with a H atom whose occupancy, summed with those of the F-atom occupancies in their corresponding sites, would amount to 1.

The site occupancy factor of the H atom at the 3 position = $1–x_1$

The site occupancy factor of the H atom at the 5 position = $1–x_2$

9. **Refinement of 2(34)SS:** Structure solution by direct methods showed a CA skeleton with three significant residuals at the 2, 3, 4 and 5 positions. Underlying the refinement protocol is the fact that there is only one (averaged) CA moiety in the asymmetric unit. This would mean that in this case, the occupancies of the F atoms at the 3 and 4 positions would be the same, and the sum of the occupancies at the 3 or 4 position and the 2 position would be 1. But, since in this case the F atom in the 3-position is distributed between the 3 and 5 positions, the occupancies at the 3 and the 5 positions must add up to the occupancy value at the 4 position. The F-atom occupancy at the 2 position corresponds to the proportion of 2FCA in the crystal structure, and the occupancy of the 3 or 4 position to the proportion of 34DFCA (Scheme S5). If x denotes the F-atom occupancy at the 3 or 4 position and y that at the 2 position, they must satisfy the equation:

   $x + y = 1$

   Each of the three F atoms at the 2, 3 and 4 positions share their respective sites with a H atom with fractional occupancy.

   The site occupancy factor of the H atom at the 2 position = $1–y$

   The site occupancy factor of the H atom at the 3 position = $1–x$

   The site occupancy factor of the H atom at the 4 position = $1–x$

10. **Refinement of 4(23)SS:** Structure solution by direct methods showed a CA skeleton with three significant residuals at the 2 , 3 and 4 positions. Underlying the refinement protocol

is the fact that there is only one (averaged) CA moiety in the asymmetric unit. This would mean that in this case, the occupancies of the F atoms at the 3 and 4 positions would be the same, and the sum of the occupancies at the 2 or 3 position and the 4 position would be 1. The F-atom occupancy at the 4-position corresponds to the proportion of 4FCA in the crystal structure, and the occupancy of the 3 /4 position to the proportion of 34DFCA (Scheme 5). If *x* denotes the F-atom occupancy at the 2 or 3 position and *y* that at the 4-position, they must satisfy the equation:

$$x + y = 1$$

Each of the three F atoms at the 2-,3- and 4-positions share their respective sites with a H atom with fractional occupancy.

The site occupancy factor of the H atom at the 2 position = 1–*y*

The site occupancy factor of the H atom at the 3 position = 1–*y*

The site occupancy factor of the H atom at the 4 position = 1–*x*

11. **Refinement of 3(24)SS:** Structure solution by direct methods showed a CA skeleton with four significant residuals at the 2, 3, 4 and 5 positions. Underlying the refinement protocol is the fact that there is only one (averaged) CA moiety in the asymmetric unit. The F atom in the 3-position is distributed between the 3 and the 5 positions. Therefore the sum of the F-atom occupancies at the 3 and the 5 positions would give the proportion of 3FCA in the crystal structure. In this case, the occupancies of the F atoms at the 2 and 4-positions would be the same, and the sum of the occupancies at the 2 or 4 position and the 3 and 5 positions would be 1. The F-atom occupancy of the 2 or 4 position corresponds to the proportion of 24DFCA (Scheme S5).

If the F-atom occupancy at the 3 position is denoted by x_1, that at the 5 position by x_2 and that at the 2 or 4 position by y, they must satisfy the equation:

$x1 + x2 + y = 1$

Each of the three F atoms at the 2,3, 4 and 5 positions share their respective sites with a H atom with fractional occupancy.

The site occupancy factor of the H atom at the 2 position = 1–y

The site occupancy factor of the H atom at the 3 position = 1–x_1

The site occupancy factor of the H atom at the 4 position = 1–y

The site occupancy factor of the H atom at the 5 position = 1–x_2

12. **Refinement of 234SS:** Structure solution by direct methods shows a CA skeleton with four significant residuals at the 2, 3, 4 and 5 positions. All four residuals were labeled as F atoms, but satisfactory refinement could be achieved only by assuming fractional occupancy. The F atom at the 3 position is distributed between the 3 and the 5 positions. Underlying the refinement protocol is the fact that there is only one (averaged) molecule in the asymmetric unit. In this case, the occupancies of the F atoms at the 2, 3, 4 and 5 positions must add up to one, i.e, if x denotes the occupancy at the 2 position, denotes the occupancy at the 2 position, y_1 and y_2 the occupancies at the 3 and 5 positions respectively, and z the occupancy at the 4-position, x, y_1, y_2 and z must be related by the equation (Scheme S5):

$x + y1 + y2 + z = 1$

Each of the three F atoms at the 2,3, 4 and 5 positions share their respective sites with a H atom with fractional occupancy.

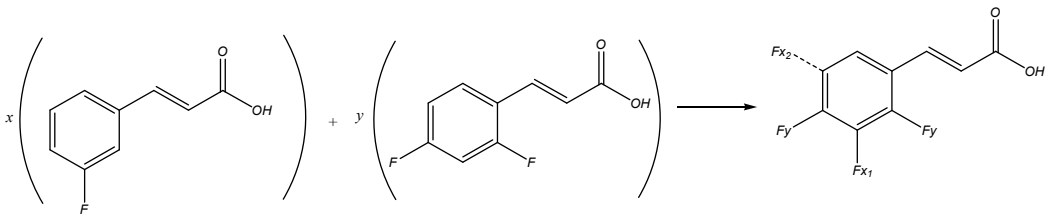
The site occupancy factor of the H atom at the 2 position = 1–x

The site occupancy factor of the H atom at the 3 position = $1–y_1$

The site occupancy factor of the H atom at the 4 position = $1–z$

The site occupancy factor of the H atom at the 5 position = $1–y_2$

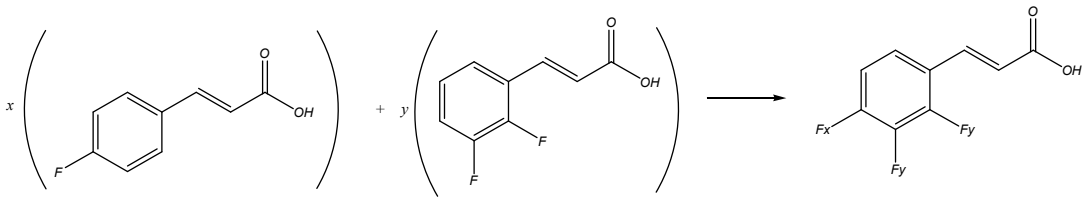
2(34)SS



$x1 + x2 + y = 1$

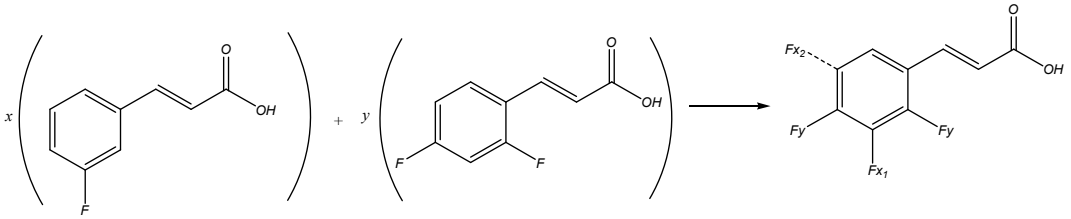
$$x1 + x2 = x$$

4(23)SS



$$x + y = 1$$

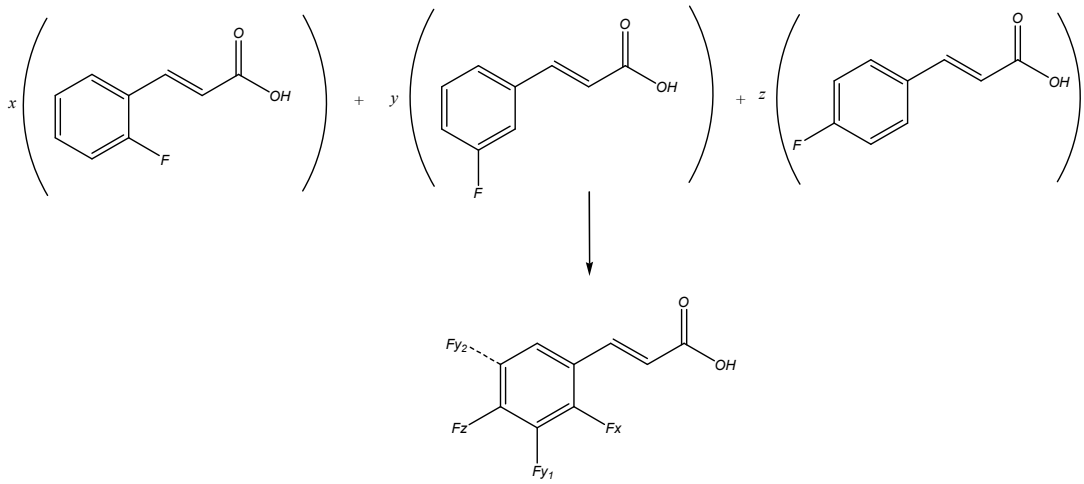
3(24)SS



$$x1 + x2 + y = 1$$

$$x1 + x2 = 1$$

234SS



$$x + y1 + y2 + z = 1$$

Scheme S5