Database-assisted design of cocrystals

László Fábián, Cambridge Crystallographic Data Centre & Pfizer Institute for Pharmaceutical Materials Science

24th March 2009 – ACS Spring Meeting, Salt Lake City, USA
Difficulties of cocrystal design

- Multiple functional groups that could form strong synthons (competition)
- No functional groups with reliable heterosynthons
- Range of pharmaceutically acceptable cocrystal formers is restricted
- Homologous compounds behave differently despite having the same synthons

Screening experiments
Aims

• To reduce the number of screening experiments (cost and time of cocrystal product development)
• To find design strategies that are complementary to the synthon approach
• Cocrystals in the Cambridge Structural Database contain relevant information
• This information can be extracted by statistical analysis
Data set preparation

- 973 structures / 1949 residues
  - at least two different neutral, non-solvent residues
  - inclusion compounds of molecules with >30 and <6 heavy atoms removed
  - cocrystals of most frequent residues removed to reduce bias

- 130 molecular descriptors
  - simple atom and bond counts
  - size and shape
  - electrostatic properties
  - surface area descriptors

Statistical methods

• Data on pairs of molecules from cocrystals
  – each pair of molecules is an observation of 2 x 130 ‘compatible’ properties

• Correlations
  – Pearson’s and Spearman’s correlation coefficients
    \((130 \times 130 / 2)\) ⇒ filtering interesting descriptor pairs
  – density plots ⇒ visualisation of property relationships

\[\text{Cryst. Growth Des. 9, 1436 (2009)}\]
Statistical methods

Scatter plot

Density plot
Results – polarity

**FPV** = Volume of polar atoms in the molecule / Molecular volume
Polar atoms: N, O, S and H atoms bonded to them

**FNO** = (No. of N atoms + No. of O atoms) / No. of non-H atoms
## Results – polarity

<table>
<thead>
<tr>
<th>Descriptor ((p))</th>
<th>Dipole</th>
<th>PV</th>
<th>PSA</th>
<th>FPV</th>
<th>FPSA</th>
<th>FNO</th>
<th>LogP*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r(p_1, p_2))</td>
<td>0.28</td>
<td>0.22</td>
<td>-0.14</td>
<td>0.37</td>
<td>-0.01</td>
<td>0.30</td>
<td>0.08</td>
</tr>
<tr>
<td>(\rho(p_1, p_2))</td>
<td>0.39</td>
<td>0.30</td>
<td>-0.08</td>
<td>0.41</td>
<td>0.01</td>
<td>0.31</td>
<td>0.10</td>
</tr>
</tbody>
</table>


Results – polarity

- **Solubility**: most cocrystals in the CSD were single crystals grown from solution
- **Dipole-dipole interactions**: large + large dipole preferred over small + large
- **Segregation of hydrophobic and hydrophilic regions**: preferred topology depends on volume ratios (see Ward & Horner, *CrystEngComm* 2004, 6, 401)
Results – polarity

JOHPUR

IXUDIO

HEJHOT

FIFLAI
### Results – shape

<table>
<thead>
<tr>
<th>Descriptor ((p))</th>
<th>(L)</th>
<th>(M)</th>
<th>(S)</th>
<th>(S/L)</th>
<th>(M/L)</th>
<th>(LMS)</th>
<th>(V/LMS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(r(p_1,p_2))</td>
<td>0.17</td>
<td>0.04</td>
<td>0.19</td>
<td>0.38</td>
<td>0.41</td>
<td>-0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>(\rho(p_1,p_2))</td>
<td>0.16</td>
<td>0.03</td>
<td>0.22</td>
<td>0.40</td>
<td>0.38</td>
<td>-0.02</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Box model:
Results – shape
Results – shape

JOHPUR

HEJHOT

IXUDIO

FIFLAI
Application

• The aim is to increase the efficiency of cocrystal screening experiments

• The CSD records the results of successful experiments, but no negatives
  – important descriptors have been identified
  – direct fitting of a scoring function is not possible
  – large differences between the values of these descriptors are rare for pairs of molecules that form a cocrystal
Application – difference cut-offs

$\Delta \mu < 5.94 \text{ D}$

$\Delta(FNO) < 0.294$

$\Delta(S/L) < 0.275$

$\Delta(M/L) < 0.294$

9th deciles of differences from CSD cocystal data
Test data

- 233 pairs of compounds
  - pairs selected by considering possible synthons
  - all pairs screened for cocrystal formation using various grinding and solution growth experiments
  - unpublished results of Shyam Karki and Tomislav Friščić (Cambridge University, Bill Jones’ group) → not in the CSD data set
  - 52 pairs formed cocrystals (22%, label ‘y’)
  - 181 pairs did not form cocrystal (label ‘n’)

www.ccdc.cam.ac.uk
Cocrystal formation is considered likely if all four descriptors (Dipole, FNO, S/L, M/L) differ by less than the 90% cut-off values among the molecules.

Success rate increased from 22% to 39%; the number of experiments reduced to half, lost only 15% of cocrystals.

<table>
<thead>
<tr>
<th>Cocrystal / Likely or not</th>
<th>Y</th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>44</td>
<td>70</td>
<td>114 (49%)</td>
</tr>
<tr>
<td>Unlikely</td>
<td>8</td>
<td>111</td>
<td>119 (51%)</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>181</td>
<td>233</td>
</tr>
<tr>
<td></td>
<td>OH</td>
<td></td>
<td>NOH</td>
</tr>
<tr>
<td>------------</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>OH</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td>NOH</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td>COOH</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
</tr>
</tbody>
</table>
Conclusion

• Analysis of cocrystals in the CSD identified molecular properties that influence cocrystal formation

• Simple filtering rules derived from known cocrystals were shown to increase the efficiency of cocrystal screening experiments

• Formation of a cocrystal is likely if supramolecular synthons, molecular polarity and molecular shapes all favour it
Acknowledgments

• Tomislav Friščić
  – screening data
  – shape / synthon experiments

• Shyam Karki
  – screening data

• Bill Jones, Neil Feeder, Frank Allen
  – discussions & support

• Pfizer
  – funding
Synthons vs shapes

• Specific experiments to test the importance of supramolecular synthons and shape similarity (Tomislav Friščić)
  – Same strong synthon, matching/mismatching shapes
  – Similar shapes, strong/weak heterosynthons
Same synthon, different shapes
Same synthon, different shapes
Same synthon, different shapes

- O
- H
- N
- O
- H
- N
- O
- H
- N
- O
- H
- N
- O
- H
- N
Design of cocrystals

- **Supramolecular synthons**: homomolecular vs heteromolecular

---

*Org. Lett.* 9, 3133 (2007)