Investigating crystal engineering principles using a data set of 50 pharmaceutical cocrystals

Peter A. Wood, Cambridge Crystallographic Data Centre
24th March 2009 – ACS Spring Meeting, Salt Lake City, USA
Co-crystals

- The term “co-crystal” refers to a crystalline material that contains two or more neutral, organic molecules or a neutral molecule plus a salt.
- In the context of pharmaceutical co-crystals, the molecule of most interest is the Active Pharmaceutical Ingredient (API).
- The additional molecule in a multi-component structure is referred to as the “co-former”.
Co-crystals

- Co-crystals are slowly becoming more popular within the pharmaceutical industry
- Potentially greater flexibility and diversity than salts or hydrates
- To use co-crystals effectively, we need to understand & have some control over them
- A large family of co-crystals with a consistent host molecule is invaluable in learning about packing effects, motifs, H-bond competition and so on…
Aims

- First multi-technique analysis of a large family of organic co-crystals
- Wanted to analyse structures using a range of methods to learn as much as possible
- Trying to avoid focusing solely on hydrogen bonds – allow other features to show importance
- What can we learn about co-crystal design in general?
Carbamazepine

- Has anti-epileptic and analgesic properties
- Known to form at least four polymorphs
- Large family of multi-component structures
- UK crystallography’s favourite compound?

Carbamazepine (CBZ)
Developing the dataset

- 37 structures retrieved from the CSD
- 13 new cocrystals determined through advanced co-crystal screening studies
  - only carboxylic acid co-formers were screened

➢ In total, 50 CBZ structures

Analytical tools

• New crystal form analysis tools in Materials module of the CSD visualiser (Mercury) – arising from collaboration with Pfizer Institute for Pharmaceutical Materials Science:
  ➢ Hydrogen-bond motif analysis
  ➢ General packing feature analysis
  ➢ Crystal packing similarity analysis
Hydrogen-bond motif analysis

- Designed to provide search tools for interaction motifs
- Auto-generated sets of motifs
- Frequency of occurrence data (CSD) indicates how common or likely a motif is
General packing feature analysis

- Easy selection of feature
- Simple searching to assess ‘how unusual is this feature?’ & ‘has this been seen before?’
- Can relate structural features to stability
Packing similarity analysis

- Materials module similarity calculation
- Identifies clusters in common between crystal structures
- Can group structures based on level of geometric similarity
Aspects of analysis

- Start with description and analysis of hydrogen bonding interactions within the dataset

- The next talk (by Scott Childs, Renovo Research) will focus more on molecular shape-based packing features
Hydrogen bonding motifs

- Most obvious and well documented motif in CBZ structures is the carboxamide homo-dimer (a) – occurs in all four polymorphs

- Carboxylic acid can provide competition – the hetero-dimer (b)
Hydrogen bonding motifs

- Motif percentage frequencies observed in the CSD:
  - Structures containing both groups
  - CBZ homo-dimer only perturbed by COOH group

- Similar results seen in CBZ dataset for structures containing both groups (27% homo, 65% hetero)
Hydrogen bonding motifs

- Can also use the motif analysis tools to investigate the more complicated motifs present
- In the analysis of co-crystals, the arrangement of the components in a motif is of interest
- Identified ring and chain patterns occurring as well as order of the groups interacting, i.e. CBZ or CF (coformer contact group)

**e.g.** an infinite 2 molecule chain could be CBZ-CBZ-, CBZ-CF-, or CF-CF-
Hydrogen bonding motifs

- Motifs:
  - Intermolecular CBZ…CBZ contacts are not common
Hydrogen bonding motifs

• Role of co-former as a link:
Hydrogen bonding motifs

- Dimensionality:

<table>
<thead>
<tr>
<th>H-bond dimension</th>
<th># of structures</th>
<th>% of structures</th>
<th>Ave. # of Donors</th>
<th>Ave. # of Acceptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>0D</td>
<td>17</td>
<td>42.5</td>
<td>0.8</td>
<td>1.5</td>
</tr>
<tr>
<td>1D</td>
<td>11</td>
<td>27.5</td>
<td>1</td>
<td>1.8</td>
</tr>
<tr>
<td>2D</td>
<td>7</td>
<td>17.5</td>
<td>2.6</td>
<td>2.1</td>
</tr>
<tr>
<td>3D</td>
<td>5</td>
<td>12.5</td>
<td>2.8</td>
<td>2.6</td>
</tr>
</tbody>
</table>

- Preference for discrete aggregates
- High correlation between dimension of H-bonding network & average number of acceptors
Hydrogen bonding

- Lack of acceptor for second carboxamide N-H

Structure 36 – CBZ/benzoic acid (1:1)
Etter’s rules:

- How many CBZ structures violate the first rule?
- Is this unusual?
- Why does it happen?
Etter’s first rule

• Start with evaluation of the CSD to get a broad perspective first

• How do you define a hydrogen bond?
  – Selected very broad limits
  – $Q_A$–H…$Q_B$ [where $Q_A = N, O$ & $Q_B = N, O, S, F, Cl, Br, I$]
  – Contact defined as within sum of vdw radii + 0.3 Å
Etter’s first rule

- Transformed variables to:
  - normalise distances
  - use spherical coordinates, solving angle bias problems

Lommerse et al. (1996) JACS, 118, 3108-3116.
Etter’s first rule

- Clear peak at $\theta = 180^\circ$
- Smaller peaks at $130^\circ$ and $105^\circ$ due to intramolecular contacts
- Peak finishes before $x = 1$
Etter’s first rule

• 97.5% of strong hydrogen-bond donor hydrogen atoms have an acceptor within sum of vdw plus 0.3 Å with D-H…A angle ≥ 90°

• In the CBZ dataset, 12 of the structures (24%) do not have such an acceptor

• Appears likely to be due to steric hindrance
Conclusions

• Developed a dataset of 50 structures containing the same pharmaceutically-relevant compound

• Applying newly developed analytical tools to learn more about co-crystal design

• Motif competition with the carboxamide dimer only seen to be successful using carboxylic acids

• Etter’s first rule broken for 24% of the dataset – indication that molecular shape is crucial
Acknowledgements

• Scott Childs (formerly with SSCI/Aptuit, now CEO of Renovo Research)
• Naír Rodríguez-Hornedo & L. Sreenivas Reddy (University of Michigan)
• Kenneth Hardcastle (Emory University)

• CCDC
  – Elna Pidcock
  – László Fábián
  – James Chisholm
  – Clare Macrae

• Funding: