Metabolite Spectral Databasing – Lest we forget.

Turning structural data into knowledge

Steve Thomas GSK
European User Meeting
15th June 2011
Ware R&D
1300 staff
Pre-Clinical Development

R&D ➔ PTS ➔ BDD ➔ SID

Structural Identification
Safety implications

The implications of passing an unsafe drug into the clinic are severe.

Getting it wrong is not an option.

What’s on our side?

• Top end equipment
• Experience
• Historical data
• Ability to smell a structural alert from 20 yards!
Historical data storage
A good memory

Grand wizard of metabolism – Frank Hollis

GSK employee:
Late Jurassic - Feb 2009

Contingency plan on Frank’s retirement!
US usage of ACD

- Early adoption from v7 in 2005 to v12 currently
- UK usage in MDR and ChemDev
Drill down from a biotransformation map
Search data from any point of interest

Variety of customers leads to a wide range of ideas as to which data is most important.
Initially designed as a tool for synthetic chemists:

- CFE allows users to database reactions
- Links to data supporting structure assignments
- Structures searchable
  - “Has this been made before?”
- Data searchable
  - “What have I made?”

To be of use in metabolism, the system must cope with 50+ products?
Auto triggered on update.

Fully customisable

- Free text
- Auto fill e.g. Date
- Combiboxes
- Radio (choice)
- Mandatory entry
Stitching databases together.

- **Formula**: C_{26}H_{40}O_6
- **FW**: 435.5815
- **[M+H]+**: 437.289765
- **[M-H]-**: 435.275213
- **Link_ID**: Simvastatin_std
- **Fraction**: 146
- **Species**: Dog
- **Matrix**: Eile
- **Metabolite**: Parent
- **Pt u**: 8.2

SpecDB_RecordDate.Link_ID: Simvastatin_std
SpecDB_RecordDate.Date: 15 April 2010
SpecDB_RecordDate.Species: Human
Populating the map
The database grows..

On data load, CFE retrieves every spectrum for every link for every metabolite.

Over 100 metabolites: most with at least two spectra, some up to 10!
Use of the Entero-Test, a Novel Approach for the Noninvasive Capture of Biliary Metabolites in Dogs
Flexible reporting

- Interpretation sheet possible
- Fragments properly represented
- Structures editable
- Exportable

<table>
<thead>
<tr>
<th>Peak ID</th>
<th>RT (min)</th>
<th>Proposed Structure</th>
<th>Mass (m/z)</th>
<th>Diagnostic Fragment Ions (m/z)</th>
<th>1H NMR Chemical Shifts (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18.082</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.082</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.111</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.136</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22.155</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24.165</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20.177</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.203</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>18.082</td>
<td>15.136</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23.082</td>
<td>22.155</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>28.111</td>
<td>37.115</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>37.115</td>
<td>15.136</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22.155</td>
<td>10.485</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.485</td>
<td>4.203</td>
</tr>
</tbody>
</table>

Diagram of chemical structures with atomic labels and proposed mass values.
<table>
<thead>
<tr>
<th>#</th>
<th>Structure</th>
<th>Compound</th>
<th>Comment</th>
<th>Analyst</th>
<th>Date</th>
<th>Species</th>
<th>Metric</th>
<th>Instrument</th>
<th>Study number</th>
<th>CYP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>Rosoxifene</td>
<td>GSH trapping post</td>
<td>Bianca Steve T</td>
<td>10 February 2011</td>
<td>Human</td>
<td>G-Tot</td>
<td>D0824575_CEDO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>Rosoxifene</td>
<td>GSH trapping post</td>
<td>Bianca Steve T</td>
<td>10 February 2011</td>
<td>Rat</td>
<td>G-Tot</td>
<td>D0824575_CEDO</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>Rosoxifene</td>
<td>GSH trapping post</td>
<td>Bianca Steve T</td>
<td>10 February 2011</td>
<td>Dog</td>
<td>G-Tot</td>
<td>D0824575_CEDO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Future? Biotrans map comparison
Data dependant display
Summary

- Current LinkID approach functional and useful
  - Planned to become single entity

- Valuable way to store, search and share sets of related molecules
  - Body (metabolites)
  - Environment (stability)
  - Chemist (reactions / impurities)

- Further enhancements constantly delivered
  - Searchable assignments from within CFE
Acknowledgements

Ian Peirson
Paul Hubberstey
All of SID-Ware

Frank’s brain