Analysing selectivity through multi-dimensional activity cliff analysis

Tim Cheeseright
Cresset summary

> Growing and profitable company
  > 20% year on year growth since 2009
  > 20 People, 12 with PhDs

> Primary market pharmaceutical and biotech R&D
  > Software:
    > 14 of the top 20 pharmaceutical companies use Cresset’s technology in their research programmes
  > Consultancy Services:
    > ~200 collaborative projects delivered to global clients

> Secondary markets: agrochemicals, flavours and fragrances, consumer health and fine chemicals
Drug discovery’s similarity hypothesis

> Similar molecules have similar activities
> Small changes lead to small changes
→ QSAR, virtual screening, lead optimization
(Un)Interesting SAR

What about the bits where the similarity hypothesis breaks down?

Nothing happens

Something dramatic happens
Activity cliffs – interesting regions of SAR

> Many names:
  > Disparity (Merck 1990s)
  > SALI (Guha/Drie 2008)
  > Activity Landscapes
  > Activity Cliffs

> Definition:
  > For each pair of molecules \( \kappa = \frac{Act_1 - Act_2}{Distance_{12}} \)
  > Usually distance = 1 – similarity
    > Similarity from 2D fingerprints, tanimoto etc
  > Large \( K \) indicates an activity cliff
Gaining understanding of Activity Cliffs

> Activity cliffs from 2D similarity highly valuable
> But no explanation for why the cliff is present
> Without an explanation we cannot use the cliff to design new compounds with confidence
> True understanding can come from 3D metrics
  > Shape
  > Electrostatics
> What about using 3D similarity from the outset?
Using 3D similarity

> 2D metrics are easy: 1:1 map to topology
> 3D is defined for **conformers**, not for **molecules**
Context is everything

> Don’t need/want **generic** 3D similarity
  > Have activity context – bound to the protein

> Align all molecules to known bioactive reference conformer

> Provides a conformation context to each molecule
3D disparity

1. Generate conformers
2. Align to reference(s)
3. Calculate 3D similarity matrix on aligned conformations

What 3D properties do we want to capture?
Properties of a 3D similarity

> Shape / Sterics

> Electrostatics – substituent effects

Changes to potential interactions from new atoms

Changes induced in retained portions
Detailed electrostatics from XED

> eXtended Electron Distribution gives detailed electrostatic interaction patterns

Separation of π- and σ- charges enables modelling of substituent effects

XED adds p-orbitals to get detailed representation of atoms
Field points

MIP contains too much information to use computationally in a reasonable time

3D Molecular Electrostatic Interaction Potential (MIP)

= Positive

= Negative
Field points

MIP contains too much information to use computationally in a reasonable time.

Field Points provide computationally tractable framework for electrostatic similarity.

3D Molecular Electrostatic Interaction Potential (MIP)

Field Points:
- Red = Positive
- Blue = Negative
- Yellow = Shape
- Orange = Hydrophobic
Alignment, scoring and comparisons

Clique based alignment

Fields
0.66

Cheeseright et al,
*J. Chem Inf. Mod.*, 2006, 665

Shape
0.98

Grant, Gallardo, Pickup,
*J. Comp. Chem.*, 1996, 1653
Alignment, scoring and comparisons

Clique based alignment

Combined 0.82

Fields 0.66

Shape 0.98


Grant, Gallardo, Pickup, *J. Comp. Chem.*, 1996, 1653
3D disparity workflow

1. Generate conformers
2. Align to reference(s)
3. Calculate 3D shape & electrostatic similarity matrix
   > Allow small movements
4. Calculate disparity matrix from similarity numbers
   > Similarity cutoff of 0.95 (Distance cutoff of 0.05)
5. Visualize
   > Difficult – 100 molecules gives 4950 pairs!
Visualization

> Existing ways to visualize
  > Table & Matrix views
## Top pairs table

<table>
<thead>
<tr>
<th>Pair</th>
<th>Good Activity</th>
<th>Bed Activity</th>
<th>Dispersion</th>
<th>Similarity</th>
<th>PBF</th>
<th>Delta Activity</th>
<th>Delta LE</th>
<th>Delta LL</th>
<th>Delta TPSA</th>
<th>Delta Lipophilicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-26</td>
<td>0.84</td>
<td>4.39</td>
<td>-49</td>
<td>0.951</td>
<td></td>
<td>-2.45</td>
<td>-0.094</td>
<td>-0.994</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3-56b</td>
<td>6.36</td>
<td>6.62</td>
<td>-1.2</td>
<td>0.947</td>
<td></td>
<td>-2.18</td>
<td>-0.067</td>
<td>-0.07</td>
<td>3.2</td>
<td>-0.1</td>
</tr>
<tr>
<td>2-55</td>
<td>6.36</td>
<td>6.42</td>
<td>-30.4</td>
<td>0.089</td>
<td></td>
<td>-1.07</td>
<td>-0.055</td>
<td>-0.026</td>
<td>0</td>
<td>-1.2</td>
</tr>
<tr>
<td>3-65b</td>
<td>6.38</td>
<td>6.32</td>
<td>-30.7</td>
<td>0.939</td>
<td></td>
<td>-2.86</td>
<td>-0.074</td>
<td>-0.057</td>
<td>0</td>
<td>-0.7</td>
</tr>
<tr>
<td>3-55a</td>
<td>0.24</td>
<td>0.32</td>
<td>-38.4</td>
<td>0.992</td>
<td></td>
<td>-1.82</td>
<td>-0.059</td>
<td>-0.056</td>
<td>0</td>
<td>-0.3</td>
</tr>
<tr>
<td>3-2</td>
<td>6.77</td>
<td>4.30</td>
<td>-34.8</td>
<td>0.951</td>
<td></td>
<td>-2.38</td>
<td>-0.102</td>
<td>-0.108</td>
<td>0</td>
<td>3.3</td>
</tr>
<tr>
<td>5-51a</td>
<td>8.24</td>
<td>6.62</td>
<td>-34.4</td>
<td>0.963</td>
<td></td>
<td>-1.72</td>
<td>-0.061</td>
<td>-0.07</td>
<td>3.2</td>
<td>9.2</td>
</tr>
</tbody>
</table>
Disparity matrix

Coloured by Disparity
Strong colours $\rightarrow$ SAR

Green $\rightarrow$ Activity Increase
Red $\rightarrow$ Activity Decrease
Visualization

> Existing ways to visualize
  > Table & Matrix views
  > Graph view (Guha/van Drie 2008)
  > Activity landscapes (Bajorath)
Activity View

Current Focus Compound

Comparator compound

Ten Nearest Compounds, height = distance

Shade = Disparity
Strong colours = Strong SAR
Electrostatic comparison
Electrostatic comparison

Difference plot – Regions where each molecule has stronger electrostatics
Selectivity Cliffs

> Selectivity often as important as potency
> Look at what structural changes caused large changes in selectivity
> Use Selectivity Endpoint as Activity?

\[
\kappa \approx \frac{\Delta \text{Selectivity}}{\Delta \text{Structure}} = \frac{\left( \frac{\text{Activity}_\beta}{\text{Activity}_a} \right)_A - \left( \frac{\text{Activity}_\beta}{\text{Activity}_a} \right)_B}{(1 - \text{Similarity})}
\]

> What about 3 activities?
> How would we visualize that?
Activity View – 2 activities

Red/Green = ↓ pKᵢβ, ↑ pKᵢα → PI3Kα selective
Selectivity matrices – 2 activities

Red = ↓ pK$_i\beta$
Green = ↑ pK$_i\alpha$
→ more α selective

GDC-0941
pK$_i\alpha$ = 9.06
pK$_i\beta$ = 8.02
Sim 0.90

6
pK$_i\alpha$ = 9.06
pK$_i\beta$ = 7.02
Application to Adenosine Receptor Antagonists

> 3 Activities – A1, A2a, A3 receptors
> Ligands aligned to x-ray structures 3PWH, 3EML
> 89 cmpd sub-set with high 3D similarity (>0.7)
Disparity Matrix – 11,748 data points
Disparity Matrix – focus on highly similar pairs
Why?
Limitations

> 2 Activities work well
> 3 is OK
> 7 is too many!
Limitations
Conclusions

> Activity Cliff/Disparity analysis provides quick insights into SAR
  > Focus on understanding the reason for a cliff
  > Drive design decisions

> Multiple ways to navigate the data
  > Compound focus
  > Most significant changes
  > Global overview
  > Cluster analysis

> 2D and 3D both useful
  > 2D provides insights into conformational changes
  > 3D provides insights into electrostatic effects

> Visualizing multiple activities simultaneously allows selectivity analysis
  > Large amounts of data difficult to visualize
Acknowledgements

> Mark Mackey
> Nigel Palmer
> Rae Lawrence
> Susana Tomasio
> Giovanna Tedesco
Thank you!

Questions Welcomed

Follow Cresset

tim@cresset-group.com