Bridging the Dimensions: Seamless Integration of 3D Structure-based Design and 2D Structure-activity Relationships to Guide Medicinal Chemistry

ACS Spring National Meeting. COMP, March 13th 2016
Marcus Gastreich, Matthew Segall, Carsten Detering, Ed Champness, Christian Lemmen
matt.segall@optibrium.com
Overview

- 2-dimensional (2D) structure-activity relationships (SAR)
  - Qualitative: Activity cliffs, matched molecular pair analysis...
  - Quantitative: QSAR models

- 3-dimensional (3D) structure-based design
  - Scoring/affinity prediction
  - Understanding 3D SAR

- Linking 2D and 3D SAR to guide design

- Conclusions
2D Structure-Activity Relationships
Qualitative SAR

Many methods are routinely used for analysis of data to reveal patterns and trends to guide compound optimisation, e.g.

• Clustering
  – Group ‘similar’ compounds to identify series with interesting SAR

• Activity cliff detection
  – Small changes in structure that cause a large change in activity

• Matched molecular pair analysis
  – Pairs of compounds that are identical except for one small change at the same position
Visualising 2D SAR
Card View™

• Freedom from the constraints of ‘chemical spreadsheets’
  – Represent compound relationships

• Work the way you think
  – Cards: Display key compound data
  – Stacks: Summarise and compare data for groups of compounds
  – Links: Highlight compound relationships

• Intuitive visualisation of analyses
  – Clustering, activity cliffs, matched molecular pairs...

• Quickly identify optimisation strategies
Activity Neighbourhood
Activity Cliff Visualisation
Matched Molecular Pair Analysis

Quantitative Structure-Activity Relationships

Principles

\[ y = f(x_1, x_2, x_3, \ldots) \pm \varepsilon \]

• Data
  - Quality data is essential
  - Public data needs very careful curation (and may not be good enough)

• Descriptors, e.g.
  - Whole molecule properties, e.g. logP, MW, PSA...
  - Structural descriptors, SMARTS, fingerprints...

• Statistical fitting or machine learning method, e.g.
  - Partial least squares, artificial neural networks, support vector machines, random forest, Gaussian processes...

• Widely applied to prediction of ADME and physicochemical properties

Please see talk: COMP Wed. 8.40am Room 25C

© 2016 Optibrium Ltd.
Multi-Parameter Optimisation

Probabilistic Scoring

*Integrated assessment of data against project criteria*

Accounts for the uncertainties in all compound-related data (experimental or calculated)

Project specific scoring profile

Compounds ranked by likelihood of success

Histograms for quick visual guide to compound properties

Interactive Redesign

- QSAR models provide estimates of compounds’ properties
- Instant feedback on how properties are likely to change
  - Explore strategies for redesign
- But, important questions
  - “Why is a property value predicted?”
  - “Where can I change this property?”
- Glowing Molecule™:
  - Visual indication of structural influences on predicted properties
3D Structure-Based Design
Visual Understanding of 3D Affinity Data
HYDE: A Different View @ Energetics*

- Physics only
- No calibration to complexes
- Relates to real Free Energies

*CINF 9: Christian Lemmen, “Predicting binding affinity doesn't work, or does it?” Next talk!
HYDE Scoring Function – Concept*

\[ \Delta G^i_{\text{HYDE}} = \sum_{\text{atom } i} \Delta G^i_{\text{dehydration}} + \Delta G^i_{\text{H–bond}} \]

*Reulecke et al., ChemMedChem’08
Hyde - Visual Affinities

HYDE color code:
- +ΔG contribution
- -ΔG contribution
- no ΔG contribution

<table>
<thead>
<tr>
<th>Component</th>
<th>ΔG Contribution (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>receptor carbonyl oxygen</td>
<td>8.2</td>
</tr>
<tr>
<td>ligand aromatic oxygen</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>total desolvation cost</strong></td>
<td><strong>10.6</strong></td>
</tr>
<tr>
<td>receptor aromatic carbons</td>
<td>-5.2</td>
</tr>
<tr>
<td>ligand aromatic carbon</td>
<td>-2.0</td>
</tr>
<tr>
<td><strong>total desolvation gain</strong></td>
<td><strong>-7.2</strong></td>
</tr>
<tr>
<td>receptor amide N dehydrat</td>
<td>6.3</td>
</tr>
<tr>
<td>interaction energy</td>
<td>-7.4</td>
</tr>
<tr>
<td>ligand aromatic N dehydrat</td>
<td>6.4</td>
</tr>
<tr>
<td>interaction energy</td>
<td>-7.5</td>
</tr>
<tr>
<td><strong>total H-bond energy</strong></td>
<td><strong>-2.2</strong></td>
</tr>
</tbody>
</table>
Visual Understanding of 3D Affinity Data

PDB: 1GKC
Linking 2D and 3D SAR to Guide Design
Understanding Activity Cliffs in 3D

PPAR[subscript]γ PDB 4EMA
Understanding Activity Cliffs in 3D
PPAR\textsubscript{\gamma} PDB 4EMA
Understanding Activity Cliffs in 3D

PPARγ PDB 4EMA
Understanding Activity Cliffs in 3D

PPAR\_PDB 4EMA
Understanding Activity Cliffs in 3D

PPARγ PDB 4EMA
Exploration of Virtual Screening Results

HSP90

- Crystal structure PDB ref. 2XJX

- Virtual library generated using STORM workflow in KNIME
  - Amide substitution using Schotten-Baumann reaction on beta resorcylic core
  - Building blocks from vendor catalogues
  - ‘Tail’ of molecule not contributing to affinity

- Resulting library docked with FlexX

- Scored using SeeSAR and HYDE to estimate $pK_i$
Matched Molecular Pair Analysis
Matched Molecular Pair Analysis
Matched Molecular Pair Analysis
Matched Molecular Pair Analysis
HYDE Analysis in SeeSAR

Row731

Hyde pKi e...: 8.142

Row208

Hyde pKi e...: 4.138
Combine with 2D QSAR Predictions
Multi-parameter optimisation
3D View... Optimisation opportunities
HYDE Analysis in SeeSAR
Optimisation Idea
Modify pKa of Nitrogen by amide substitution
Optimisation Idea
Add polar group to phenyl ring
Repose in SeeSAR
Conclusions

• Both 2D and 3D information are important to interpret SAR and guide design

• A seamless combination between these two views of the chemical world maximises the benefits that they bring
  – Understanding SAR from experimental data
  – Analysis of virtual screening/docking results
  – Multi-parameter optimisation of potency, physicochemical and ADMET properties

• For more information:
  – Optibrium: Booth 1227 or outside of room 6E (MEDI)
Free Hands-On Workshop

Seamless Integration of 2D and 3D SAR to Guide MPO

- Where: San Diego Convention Center, Room 15B
- When: Monday 3:30pm to 6pm
- Practical examples with SeeSAR and StarDrop
- Spaces are limited, so please register at Booth #1227 in the Exhibition Hall