Compact models for compact devices:
Visualisation of SAR using mobile apps

Alex M. Clark, Ph.D.

August 2015
Modern User Experience

- Continuity from 1980s
- Mostly open architecture
- Small touchscreen
- Ultimate mobility
- Deploy on anything
- Very open
- Tied to server
- Unitary functionality
- Company store
MOLECULAR MATERIALS INFORMATICS

Mobile Particularly

Small Data
- Device Local
- presentation
- visualisation
- calculation
- sharing

Medium Data
- custom content
- hard to appify
- searching
- registration
- online sharing
- heavy calculations

Large Data
- Remote API
The Remoteness Problem

- Very tempting to use servers for heavy lifting
- Requests via Internet API

- User identity
- Security issues
- Maintenance burden
- Hostile user experience
  - no signal
  - foreign countries
  - planes
  - wilderness
MOLECULAR MATERIALS INFORMATICS

The Local Solution

- Structure ➔ Property calculation algorithms

- Start with scalars:
  - some are easy
  - others less so

Calculated Properties

- Molecular Formula: $\text{C}_9\text{H}_{10}\text{N}_4\text{O}_2$
- Molecular Weight: 194.1906
- Heavy Atoms: 14
- H-Acceptors: 6
- H-Donors: 0
- Rotatable Bonds: 0
- Log P: -1.0293
- Molar Refractivity: 51.196
- Stereoambiguity: n/a
Substructure Filters

- **PAINS** and **AZ Filters**: both require complex recursive substructure queries
- Overlay visualisation

**PAINS Filters**
Molecule Perception

- Hardcoded algorithms for **stereochemistry** and **tautomer enumeration**
- Displayed by structure annotation
• Prepare data using desktop software (and/or of scripts)
• Build **Bayesian** (ECFP6) model: concise, portable
• Transfer model to mobile app...
Bayesian Models
Predicted numbers are scaled to probability-like predictions: most molecules within the domain are in the range 0..1.

Solubility: 0.630

Probelike: 0.409

hERG: 0.473

Plague: 0.247

Chagas: 0.302

Malaria: 0.123

Tuberculosis: 0.293

• Fast to apply
• Numeric or molecular visualisation
• Overlay uses ECFP6 fingerprints
• With one molecule or many
More is Better

- Extract target:activity collections from ChEMBL
- Compute active/inactive boundaries
- Automatically build Bayesian models...

**ChEMBL**

<table>
<thead>
<tr>
<th>Target</th>
<th><strong>β-Hydroxysteroid dehydrogenase</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-Acylglycerol-3-phosphate</td>
</tr>
<tr>
<td></td>
<td>Acetylcholine esterase</td>
</tr>
<tr>
<td></td>
<td>ADAM10</td>
</tr>
<tr>
<td></td>
<td>Adenosine A1 receptor</td>
</tr>
<tr>
<td></td>
<td>Adrenergic receptor α1</td>
</tr>
<tr>
<td></td>
<td>Aldoketo reductase</td>
</tr>
<tr>
<td></td>
<td>Alkaline phosphatase</td>
</tr>
<tr>
<td></td>
<td>Aminopeptidase N</td>
</tr>
<tr>
<td></td>
<td>Angiotensin converting enzyme</td>
</tr>
<tr>
<td></td>
<td>Anthrax lethal factor Bcl2</td>
</tr>
<tr>
<td></td>
<td>Arachidonate 12</td>
</tr>
</tbody>
</table>

*JCIM* 55, 1246-1260 (2015)
Extreme Appification

- PolyPharma app: work in progress
- Visualise a lot of SAR data... with 2 taps
Demo

- **PolyPharma** app demonstrated live...

- Since the meeting, the app has become available on the iTunes AppStore:

  http://itunes.apple.com/app/polypharma/id1025327772

- For more information, see:

  http://molmatinf.com/polypharma.html
Acknowledgments

• Sean Ekins, Antony Williams, Andy Davis
• Royal Society of Chemistry
• Collaborative Drug Discovery
• Inquiries to info@molmatinf.com

MOLECULAR MATERIALS INFORMATICS
http://molmatinf.com
http://molsync.com
http://cheminf20.org
@aclarkxyz