Computational Chemistry in Drug Design

Xavier Fradera
Overview

- Introduction and background
- Drug Design Cycle
- Computational methods
  - Chemoinformatics
  - Ligand Based Methods
  - Structure Based Drug Design
- Conclusions
Organon

- **Organon** is a mid-size pharmaceutical company
- €3.5 billion in 2005 revenues
- Historical focus on gynecology and fertility
- Products in phase III: **Asenapine** (neuroscience) and **Sugammadex** (anesthesia)
Organon Research Scotland (Newhouse)

- Approx. 300 staff (Chemistry and Pharmacology)
- Focus on neuroscience targets
- Recent investment on ALS facilities
- Computational Medicinal Chemistry (CMC) is part of the Chemistry Department
Computational Medicinal Chemistry (CMC)

• Currently 6 staff
• Role
  - CMC members are part of Hit and Lead Optimization teams
  - **Database** management, compound acquisition, etc.
  - Development of chemoinformatic tools
  - Frequent collaboration with Molecular Design and Informatics (MDI) in Oss
• Some chemists have basic molecular modelling skills which can contribute to projects
• Projects very keen on having CMC support
Drug Research

Genes

Target Discovery
- Target Identification
- Target Validation
- Assay Development

Target

Lead Discovery
- Hit Identification
- Hit Confirmation
- Hit Optimization

Lead

Lead Optimization
- In Vivo Activity
- Bioavailability
- Selectivity

Candidate

Development
Comp. Chem. and Drug Design

• CMC must have an **impact on HO/ LO**
  - Providing **models** that are useful in guiding synthetys → constant interaction with chemists !!
  - Need to **move fast**: projects won’t wait on us !

• Methods to use depend on project and data available
  - **Chemoinformatics** tools, prediction of **physchem properties**, etc.
  - **Ligand-Based Drug Design** applicable specially for targets with no structural data (GPCRs, etc.)
  - **Structure-Based Drug Design** applicable for targets with structural data available (PDB or in-house) → approx. 40% of Newhouse targets
Chemoinformatics

- Combination of external and in-house tools
- Many chemoinformatics tools available to chemists through a web interface
  - Rule of 5
  - Bioisosteres
  - compound clustering
  - ordering of commercial compounds
  - identify scaffolds
  - Etc.
- Project data accessible through an in-house program
Chemoinformatics web interface

- Similar **web interface** for many chemoinformatics tools
- Available internally to all research staff
- Can draw any structure, read sdf files, or retrieve from internal databases
- **Monika** is our in-house program for **Lipinski’s Rule of 5** filters
- Other commonly used properties
  - hERG model
  - Structural alerts
  - pKa, solubility
  - mutalert
Monika (Rule of 5)

- Export menu (e.g. Accord/Excel)
- Selection of cmpds in table
Sygma

- **Systematic Generation of potential Metabolites**
- Molecular structure needed as input (same web interface)
- Uses **reaction rules** derived from MDL Metabolite database
- Useful for **prediction or labile sites**, or to support experimental metabolite ID
Sygma

- Example with Lipitor
- A ranked list of metabolites is generated
- Probability of each rule validated against experimental data
- Many of these metabolites reported in the literature for this compound

E.g., ring closure is a likely reaction for Lipitor
Ligand-Based Methods

• Essential for GPCR and targets without structural information
• Tools available
  - **2D-similarity**: BCI fingerprints, Feature Trees, etc.
  - **3D-Similarity**: MIMIC, FlexS
  - **Pharmacophore alignment**: Gasp, Catalyst, etc
  - **3D-QSAR**: COMFA, etc.
• Tasks
  - Virtual Screening
  - Alignment of ligands → binding mode hypothesis, understand SAR
  - QSAR models
  - Clustering of compounds, diversity, selection of reagents, etc.
Ibis

- **Ibis** is an **in-house topological pharmacophoric fingerprint**
- Each atom type is defined in terms of a basic set of properties (aromatic, hydrophobic, H-donor, H-acceptor, conjugated, non-hydrogen)
- These are used to generate molecular fingerprints and compare them (2D similarity)
- Attachment points can be incorporated in the description → this can be used to search for bioisosteric replacements
- Can retrieve **experimental examples** of bioisoster pairs (in-house, and from BIOSTER database)
- Ibis bioisostere searching available as a web tool

Ibis (bioisosteres)

- Example: bioisosteres of **amide**, ordered by similarity
- When available, **experimental examples** of bioisoster pairs are shown
- Additional information: frequency of bioisosters in WDI, structural alerts.

Red alerts for reactive or toxic funct. groups

Retrieval of experimental examples of bioisoster pairs
Structure Based Drug Design

• Structural data very useful when available
• Organon has started producing **in-house x-ray structures**
• This is having a big impact in projects
• CMC role essential:
  – Solving x-ray structures
  – understanding structural data and presenting to the chemists
  – modelling: docking, MD, etc.
  – Incorporate external data (PDB, CSD)
  – Design of general/specific fragment libraries for Fragment-Based SBDD
SBDD

- Limitations of receptor-based computational methods
  - **Protein flexibility** difficult to take into account
  - X-ray models give only a **static picture**
  - **Scoring functions** unreliable for predicting affinity
  - Accurate calculation of **binding Free Energies** is still difficult
  - Doesn’t take into account **known SAR** around the ligand
SBDD for LXR

- Liver X receptor (LXR) is a highly flexible receptor
LXR

- At the time, only 1 pdb structure was available
- DOCK generated several orientations, but the top-ranking one couldn’t explain the SAR
- FlexX had difficulties in fitting the ligand into the cavity
- We used the orientations generated with DOCK followed by a “quick and dirty” MD relaxation
- This was successful in predicting orientations for our LO cmpd that were later confirmed by x-ray structures
- Later the protocol was shown to obtain similar results using different protein structures
**LXR**

- **Example: native LXR-ligand (epoxy-cholesterol)**
- DOCK is incapable to recognize the native orientation (1)
- Re-scoring with a quick MD protocol, we were able to rank them correctly (using native and non-native receptor structures)
- 2 protonation states taken into account for His435 (delta/epsilon)

<table>
<thead>
<tr>
<th>Structure</th>
<th>Energy (native)</th>
<th>Energy (non-native)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (native)</td>
<td>-81.6 / -85.4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-69.2 / -72.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>-69.8 / -71.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-77.8 / -79.7</td>
<td></td>
</tr>
</tbody>
</table>
LXR

X-ray orientation compared to min. E snapshot from MD
MacDOCK

Ariadne’s string led Theseus safely out of the Minotaur’s labyrinth. The labyrinth represents reality, the string represents knowledge.
MacDOCK

- First Implementation of a **Similarity-Driven Docking algorithm** (2000)
- In-house modified version of **DOCK 4.0** combined with **MIMIC**
- Maximal use of Available Information
  - **Receptor**: X-ray structure
  - **Ligand**: known binding mode from x-ray structure or SAR
- Similarity between ligand (B) and reference structure (A) is used to correct the docking score D
  
  \[ G_{RB} = D_{RB} \cdot S_{PB} \]

- A 2\textsuperscript{nd} similarity term for covalent docking can also be included
- Similar approaches now implemented in many commercial programs (**FlexX-Pharm**, etc)

Ref: X. Fradera, R.M.A. Knegtel, J. Mestres
Example: PPACK in thrombin

- Receptor steric and electrostatic grids from a non-covalent structure
- “Anchor” structure from a covalent structure
- “Pharmacophoric” structure from a known ligand
- Ligand warhead has to be modified (e.g. ketone to hydroxyl)

Docking of PPACK

Layer 1

\[ D = 0.00 \]
\[ S_A = 0.97 \]
\[ G = -0.97 \]
Docking of PPACK

Layer 4

\[ D = -13.08 \]
\[ S_A \times S_P = 0.48 \]
\[ G = -6.28 \]
Docking of PPACK

Layer 6

\[ D = -36.65 \]
\[ S_A \times S_P = 0.75 \]
\[ G = -27.49 \]
Docking of PPACK

Layer 8

\[ D = -42.47 \]
\[ S_A \times S_P = 0.88 \]
\[ G = -37.37 \]
Docking of PPACK vs Xray Structure

RMSD with respect to X-ray structure is 0.95
Conclusions

• Computational Chemistry is an integral part in Target Validation, Hit-, and Lead-Optimization at Organon
• Important to use all data available
• Many methodological challenges remain
  – Accurate and fast prediction of ligand-protein binding energies
  – Prediction of water solubility
  – Accurate ADME/Tox modelling
  – Etc.