Overview of the Drug Design and Discovery Workshop

by Foundation for Advanced Education in the Sciences (FAES), NIH

TK Chua
5th Oct 2018
Drug Discovery & Development

Drugs are developed through a series of phases that are critical for ensuring their efficacy and safety before they can be approved for marketing. Here is a summary of the key stages involved in the drug discovery and development process:

**Drug Discovery**
- Target selection
- Lead-finding
- Lead optimisation
- Pharmacological profiling

**Preclinical Development**
- Pharmacokinetics
- Short-term toxicology
- Formulation
- Synthesis scale-up

**Clinical Development**
- **Phase I**
  - Pharmacokinetics
  - Tolerability, side-effects in healthy volunteers
- **Phase II**
  - Small-scale trials in patients to assess efficacy and dosage
  - Long-term toxicology studies
- **Phase III**
  - Large-scale controlled clinical trials

**Regulatory Approval**
- Submission of full data and review by regulatory agencies
- Postmarketing surveillance

The timeline and number of compounds at each stage are as follows:

- **Drug Discovery**: 2–5 years
- **Preclinical Development**: 1.5 years
- **Clinical Development**: 5–7 years
- **Regulatory Approval**: 1–2 years

- **Drug candidates**: ~100 projects
- **Development compounds**: 20 compounds
- **Clinical development phases**:
  - Phase I: 10
  - Phase II: 5
  - Phase III: 2
  - Postmarketing surveillance: 1.2
- **Drug approved for marketing**: 1
In Silico Drug Design Pipeline

Ligand Preparation for screening
  ↓
Databases
  ↓
High Throughput Screening (SBDD and LBDD)
  ↓
ADME/Toxicity estimation
  ↓
Experimental Assay

Abbreviation:
• SBDD-Structure Based Drug Design
• LBDD-Ligand Based Drug Design
• ADME-Absorption, Distribution, Metabolism and Excretion, describes the disposition of a drug in a body.
Ligand Preparation

• Searching for Analogs of a KNOWN compound, e.g. Caffeine
• Databases: ZINC15, PubChem, ChemSpider, etc
• SMILES strings
• ZINC15-Database of commercially available compounds for virtual screening
  • [https://zinc15.docking.org/substances/home/](https://zinc15.docking.org/substances/home/)
• Search unknown structure in online database using Tanimoto70.
Tanimoto Coefficient

• A coefficient measuring similarity between pairs of molecules based on their molecular fingerprints.

• Molecular Fingerprint: Encoding of a molecule into a series of binary digits “based on its structure, such that structurally similar molecules will have similar fingerprints.” (Kristensen TG et al. 2010)

• Tanimoto distance between two molecules (1=exact, 0=nothing in common)

• > 0.7-0.8 = similar
ZINC15: Caffeine-Tanimoto 70
Ligand Preparation

• UNKNOWN Chemical Compounds:
• Draw chemical structures (e.g. ChemSketch, MarvinSketch)
• MarvinSketch understands SMILES strings
• Download the structure from PubChem or ChemSpider as .sdf or .mol files.
Ligand Preparation

Open Babel

- Converting file formats
- .mol (ChemSpider), .sdf (PubChem)
- .mol2, .pdb
- SMILES
- Generate conformers (e.g. Marvin Sketch)
SBDD (AutoDock)

• Designed to predict how small molecules (e.g. ligands or drug molecules) bind to a receptor of known 3D structure.

• Consists of two main programs:

• AutoDock-performs the docking of ligand to a set of grids describing the target protein (receptor).

• AutoGrid-pre-calculates these grids.
SBDD (AutoDock)

- Protein structure (.pdb) required.
- Preparing protein/receptor
  - Delete ligand, water; add hydrogen and charges
  - Saved as .pdb file → .pdbqt file
- Preparing the ligand
  - add hydrogen and charges
  - Saved as .mol2 file → .pdbqt file
- AutoGrid
  - Create .gfp and .glg files
- AutoDock
  - Create .dfp and .dlg files
LBDD (Ligand Based Drug Design)

- Ligand-based (pharmacophore), no protein structure required.
- Methods:
  a) Similarity Search (e.g. Tanimoto Search in ZINC15)
  b) QSAR
  c) Pharmacophore modeling (e.g. PharmaGist)
  d) Shape-based Searching (e.g. Commercial OpenEye Scientific)
- ADME-Tox Prediction (e.g. T.E.S.T)

Chemical Descriptors
- 1D (Mw, polar surface area, etc)
- 2D (Connectivity of atoms in the molecule in terms of the presence and nature of chemical bonds, Cherkasov A. et al. 2014)
- 3D (Geometry: distances, angles, torsions; steric and electrostatic fields, etc)
Similarity Search

• 2D descriptors
• Tanimoto distance between molecules (1=exact, 0=nothing in common)
  • > 0.7-0.8 = similar
• Search unknown structure in ZINC15 database using Tanimoto 70.
QSAR

• Quantitative Structure-Activity Relationships
• IUPAC Definition: mathematical relationships linking chemical structure and pharmacological activity in a quantitative manner for a series of compounds.

• 3D QSAR-CoMFA (Comparative Molecular Field Analysis)
• “Molecular field-based, alignment-dependent, ligand-based method developed by Cramer et al.” (Kunal R. et al. 2015)
• “Correlating molecular field approaches with biological activities” (Cherkasov A. et al. 2014)
Pharmacophore Modeling

• Pharmacophore (IUPAC Definition) – “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response” (Yang SY, 2010)

• Features: Positive/Negative Charges, H-bond donor/acceptor, Hydrophobic group, Aromatic ring

• Alignment

PharmaGist
• Web-based application to detect pharmacophores by alignments of the input ligands (.mol2, http://bioinfo3d.cs.tau.ac.il/pharma/about.html)
• Detects common pharmacophore features
LBDD Programs (Open-Source) to Explore

- DataWarrior
- LiSiCA
- Py-CoMFA (3D-QSAR)

http://www.openmolecules.org/datawarrior/

https://pymolwiki.org/index.php/File:3DResultsOutputtabPyMOLViewer1.png