Effective and rapid bioactivity profiling using pharmacophore-based parallel screening

Thierry Langer
The Dilemma in Drug Development...

- Success rate is too low
- Time to market is too long

![Diagram showing the stages of drug development from preclinical studies to market launch.](image)

> 100,000 new compounds

Years

850 Mio $
The Dilemma in Drug Development ...

... currently, only 20-30 new drugs are approved by the FDA every year ...

... at this rate it will take 300 years to double the number of drugs that are on the market ...

Chong & Sullivan, Nat. Drug Discov. 2007, 448, 645-646
New uses for old drugs

It takes too long and costs too much to bring new drugs to market. So let’s beef up efforts to screen existing drugs for new uses, argue Curtis R. Chong and David J. Sullivan Jr.

Fast, affordable drug development is a vision that contrasts sharply with the current state of drug discovery—which also neglects too many diseases of the poor. An analysis of 98 approved drugs estimated that it takes an average of 15 years and US$800 million to bring a single drug to market. And despite a doubling in research spending by the US National Institutes of Health (NIH) to $27 billion in 2003, the number of new drugs approved by the US Food and Drug Administration (FDA) each year remains constant at 20–30 compounds.

At this rate it will take more than 300 years for the number of drugs in the world to double.

The current costly and time-consuming paradigm of drug discovery is ill-equipped to combat rapidly emerging diseases, such as avian flu, drug-resistant pathogens and diseases that have a small market size. The solution is to identify new uses for existing drugs. As the pharmacologist and Nobel laureate James Black said, “The most fruitful basis for the discovery of a new drug is to start with an old drug.”

Because existing drugs have known pharmacological and safety profiles and are often approved by regulatory agencies for human use, any newly identified use can be rapidly evaluated in phase I clinical trials, which typically last only 2 years and cost US$6 million. In this way, drug developers can bypass up to 40% of the overall cost of bringing a drug to market.

This back-to-basics approach is growing in popularity. At least 17 existing drugs are in various stages of clinical and animal testing for new uses (see Supplementary information), and another 24 are already being re-marketed by the pharmaceutical industry for new uses.

Although most successful crossovers have been the result of chance observations or educated guesses, exceptions include the antibiotic ceftriaxone, which is a potential treatment for amyotrophic lateral sclerosis, and whose new activity was discovered following the screening of 1,000 compounds from the National Institute of Neurological Disorders and Stroke (NINDS) custom collection in Gaithersburg, Connecticut. In the past, individuals were limited to opening perhaps hundreds of compounds from clinical drug collections like the NINDS Library and the National Cancer Institute in Washington DC, among the 1,000 approved drugs for small-scale screening. In our view, what is needed is a more systematic approach to drug rediscovery that takes advantage of the resources available.

Historically, ‘repurposing’ old drugs has proved successful in introducing new therapies to the developing world. Today, even with the billions of research dollars available to create new drugs through public-private partnerships, and the promise of pay-for-success data, there remains an enormous need for therapies for neglected diseases. A recent example of a repurposed drug is miltefosine, initially developed for breast cancer but now used for treating visceral leishmaniasis. This disease is caused by a sandfly-transmitted parasite and kills an estimated 500,000 people each year. In fact, miltefosine failed phase I trials for tumour reduction and the drug was never approved by the FDA for cancer therapy. However, in vitro and animal studies indicated antiparasitic activity, and phase II proof-of-concept miltefosine as a viable treatment for visceral leishmaniasis.

Cost-cutting

Cost is one reason to revisit existing drugs. Roughly 1,000 of the 10,000 or so drugs ever tested in clinical medicine are covered by patents, so most drugs can affordably be redeplored in the developing world. Safety is another compelling reason. Phase IV clinical studies, which monitor post-marketing safety, cost around $50 million per drug to perform in developed countries and are nearly impossible in countries without an established healthcare infrastructure. Because many existing drug have undergone phase IV surveillance in millions of patients, the same stringent safety standards required by users in developed countries can be offered to patients with neglected diseases in the developing world.

Despite the promise of finding new uses for existing drugs, a comprehensive collection of the approximately 9,000 drugs known to clinical medicine does not exist. This number includes 2,933 unique drugs approved by the FDA since 1906 (ref. 1), 1,037 drugs in the 2006 Physicians Desk Reference, and 7,057 drugs that were either approved abroad or have entered phase I clinical trials, as indicated by a US Adopted Name or International Non-proprietary Name.

Excluding anticancer, pharmaceutical aids, therapeutic plant or animal extracts, and vaccines, we estimate that there are 8,150

Sir James Black

Camille G. Wermuth

SOSA - Selective optimization of side effects approach

Chong & Sullivan, Nat. Drug Discov. 2007, 448, 645-646
Our Aim: Predict Activity Pattern ...

• Modeling of all relevant targets
  – responsible for drug action and side effects
  – build feature-based pharmacophore models

• Compile all models (+ relevant info) into a database
  – Activity profiling of leads / drug candidates
  – Determination of side effects / bio-hazards

• Use this system for development of novel interesting lead molecules and drug candidates
How To Overcome The Efficiency Deficit...

Prediction of failure in early stage

Develop new applications

$850\text{ Mio }$

$150\text{ Mio }$

$>> 100,000$

$1$
The Usual Virtual Screening Protocol

$10^x$ molecules against one target

results in a hit list
Why Not Do This?

$10^x$ molecules against $10^x$ targets

... needs a large number of models!
What Is A Pharmacophore?

“A pharmacophore is the 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.”

Feature-based Pharmacophore Models

Totality of universal chemical features that represent a defined binding mode of a ligand to a bio-molecular target

Features: electrostatic interactions, H-bonding, aromatic interactions, hydrophobic areas, ...
Case Study 1: Sigma-1 Receptor

Searching for potent & selective ligands

- No 3D structure of the target available
- No significant homology to other mammalian proteins (homology modeling fails so far ...)
- Many ligands known, however often with low selectivity (sigma-2, hERG channel interaction)
- Clinical potential: antipsychotic, antidepressant, antiepileptic, etc ...
Case Study 1: Sigma-1 Pharmacophore Model

23 compounds
(Ki = 0.1 nM - 35 µM)

4 hydrophobic features
1 positive ionizable

experimental validation?
Case Study 1: Experimental Validation

Hits from the Derwent World Drug Index


<table>
<thead>
<tr>
<th>Compound</th>
<th>Ki [nM]</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR-31717a</td>
<td>2.9</td>
<td>1.2</td>
</tr>
<tr>
<td>Ro-90791</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Ro-91040</td>
<td>0.8</td>
<td>0.5</td>
</tr>
<tr>
<td>Ro-59494</td>
<td>1.4</td>
<td>1</td>
</tr>
<tr>
<td>Carbisocaine</td>
<td>232</td>
<td>11</td>
</tr>
<tr>
<td>VUF-8410</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>
Case Study 1: Experimental Validation

focused library, 4000 compounds

pharmacophore filtering

random selection based on diversity

389 testing 1µM

611 displacement

> 80%

> 50% < 80%

< 50%

Your partner for in-silico drug discovery.
High Throughput Pharmacophore Generation

Experimental basis for model generation: Protein Database (PDB*)

- Large public collection of experimentally determined 3D structure data of biologically relevant macromolecules
- Huge repository, still continuously expanding
  (~ 22,000 crystallized proteins, ~ 4,000 co-crystallized small organic ligands in different conformations)

Problems

- Historically grown (inadequate file format, bad data quality)
- Information on bond type, hybridization status, hydrogen atoms are completely missing

* www.rcsb.org/pdb
Let’s have a look ...
**Implemented Procedure**

1. Detect ligand and clean-up the binding site in the protein (all amino acids within 7Å distance from the ligand)
2. Interpret hybridization status and bond types in the ligand
3. Perform chemical feature recognition for the ligand (H-bond donor, H-bond acceptor, positive ionizable, negative ionizable, hydrophobic, aromatic ring, metal ion coordination)
4. Search for corresponding chemical features of the protein
5. Add interaction features to the model only if a corresponding feature pair is found in the complex
6. Add excluded volume spheres for opposite hydrophobic features

LigandScout Graphical User Interface
LigandScout Graphical User Interface
LigandScout Graphical User Interface
Binding Mode Specificity

One pharmacophore model accounts for one binding mode ...

How to analyze and align these objects?
Alignment By Pharmacophore Points

Methotrexate

Dihydrofolate

Wrong

Correct

Böhm, Klebe, Kubinyi: Wirkstoffdesign (1999) p. 320f
Alignment By Pharmacophore Points

1RX2

1RB3
More Details ...

See Gerhard Wolber’s talk on Monday!

A Novel, Efficient Virtual Screening Algorithm
Using 3-D Chemical Feature Pattern Recognition

COMP 157, Drug Discovery

1:00 PM - 4:15 PM, August 20, 2007, BCEC - 161
Case Study 2: Antiviral Parallel Screening

- 5 viral targets
- 50 pharmacophore models
- 100 antiviral compounds

Will their activity profiles be predicted correctly?
# Case Study 2: Targets

<table>
<thead>
<tr>
<th>Target</th>
<th>Disease</th>
<th>Function</th>
<th>Mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIV protease</td>
<td>HIV infection, AIDS</td>
<td>Cleavage of gag and gag-pol precursor polyproteins into functional viral proteins</td>
<td>Inhibition at active site</td>
</tr>
<tr>
<td>HIV reverse transcriptase (RT)</td>
<td>HIV infection, AIDS</td>
<td>Synthesis of a virion DNA, integration into host DNA and transcription</td>
<td>Inhibition at allosteric site</td>
</tr>
<tr>
<td>Influenza virus neuraminidase (NA)</td>
<td>Influenza</td>
<td>Viral envelope glycoprotein, cleave sialic acid residues for viral release</td>
<td>Inhibition at active site</td>
</tr>
<tr>
<td>Human rhinovirus (HRV) coat protein</td>
<td>Common cold</td>
<td>Attachment to host cell receptor, viral entry, and uncoating</td>
<td>Binding in hydrophobic pocket (capsid stabilization)</td>
</tr>
<tr>
<td>Hepatitis C virus (HCV) RNA polymerase</td>
<td>Hepatitis C</td>
<td>Viral replication, transcription of genomic RNA</td>
<td>Inhibition at various allosteric sites</td>
</tr>
</tbody>
</table>
Ligand-directed Analysis

Ratio ≥ 1
90% of the compounds correctly predicted

Ratio < 1
8% more often predicted for one specific false target than for correct one

for 2% of the compounds no activity prediction possible
Pharmacophore-directed Analysis

Model with lowest selectivity:
70% of actives (HIV RT), but 75% from one specific false target (HRV coat protein)
40% active and 60% inactive compounds in hit list

Model with 85% hit rate

Model with highest selectivity:
100% of actives (HCV polymerase 1), no other compounds
100% active and 0% inactive compounds in hit list
Case Study 3: Aspartate Protease Inhibitors

Profiling Aspartate Protease Inhibitors

81 Pharmacophore models

4 Compound datasets:

- Set 1: Known HIV protease inhibitors (89)
- Set 2: Other protease inhibitors (79)
- Set 3: Drug-like non inhibitors (85)
- Set 4: Presumably inactive compounds (85)

Will the activity profiles of the test set antiviral compounds be correctly predicted?

Dataset Examples

Set 1: Known HIV protease inhibitors

Set 2: Other protease inhibitors

Set 3: Protease inhibitor-like true inactives

Set 4: Drug like presumably inactive virtual compounds

Pharmacophore Model Examples

1zlh

6upj
Parallel Screening Results

Underlying Screening Platform

PipelinePilot Script & Catalyst™ DB Search

K. Chuang
J. Benedict
N. Triballeau-Hugounencq
Rémy D. Hoffmann

Your partner for in-silico drug discovery.
Web Based Parallel Screening Platform

2. Choose your data source:
- Molecular File
- Catalyst Database
- Sketched Molecule

Select an input file:

3. Select your options:
- Remove duplicate structures
- Screen all tautomers (not for .pdb)
- Best score by target
- Export To Excel

Submit
Web Based Parallel Screening Platform

Pharmacophore Profiling with HypoScreen

HypoScreen is a Pipeline Pilot WebPort application that allows compounds, either from a molecular file, a set of ligands, or an internal catalyst database or directly sketched, to be screened against HypoPiPi.

1. Pharmacophore selection:
   - glycosidases
   - glycosylases
   - peptidases
   - phosphates
   - phosphodiesterases
   - proteases (aspartic)
   - Sap2 (C. albicans)
   - beta-secretase
   - cathespin D
   - penicillopepsin
   - protease (HIIV-1)

   Screen:
   - only models with shape
   - only models without shape
   - all selected models

2. Data source:
   - Molecule File
   - Catalyst Database
   - Sketched Molecule

   Select an input file:
   users/cac7592/Suppinfo.txt Browse...
Web Based Parallel Screening Platform
Summary

First published examples of applications of extensive parallel screening approach based on pharmacophores

- Multitude of pharmacophore models (up to several thousand ...)
- Large set of molecules (up to several million ...)

Results indicate

- Correct assignment of selectivity in most cases
- Independent of search algorithms used

Fast, scalable in silico activity profiling is now possible!
Our Pharmacophore Database

~ 300 unique targets ready to use*

• Represented in
  ~ 200 ligand-based pharmacophore models
  ~ 2200 structure-based pharmacophore models

• Covering a selection of all major therapeutic classes
• Contains anti-target models for finding adverse effects
• Categorized according to the pharmacological target

* out of ~630 categorized by August 2007
Conclusions ... 

• Parallel pharmacophore-based virtual screening is a straightforward and rapid method for the assessment of bio-activity profiles

• Together with informatics-based molecular building tools, optimized design of novel and promising compounds will become feasible

• Assessment of risks in later development stages becomes possible on a rational & transparent basis
Acknowledgements

• Markus Böhler
• Oliver Funk
• Johannes Kirchmair
• Eva Maria Krovat
• Daniela Ladstätter
• Christian Laggner*
• Patrick Markt
• Claudia Schieferer
• Daniela Schuster
• Theodora Steindl*

• Rémy D. Hoffmann*
• Konstantin Poptodorov
• Nicolas Triballeau-Hugounencq*
• Kareen Chuang

• Fabian Bendix
• Martin Biely
• Alois Dornhofer
• Robert Kosara
• Thua Huynh Buu Thi Hoang
• Thomas Seidel
• Gerhard Wolber*
Thank you for your attention ...

www.inteligand.com

pharmazie.uibk.ac.at/camd
www.inteligand.com

pharmazie.uibk.ac.at/camd