Discovery Studio Catalyst Webinar

Going where no pharmacophore has gone before: fragment-based design and activity profiling

- Reservationless-Plus Dial-In Number (US): (866) 519-8942
- Reservationless-Plus International Dial-In Number: (770) 790-2186
- Conference Code: 8587995760
Discovery Studio®

Fragment-based Design and Ligand Profiler

David M. Lorber, MBA, Ph.D.

November 1, 2007
Why Fragment-based Design

- Compounds designed using this method are more likely to be novel compared to screening of larger compounds, because the smaller fragments explore chemical space more effectively.
- It detects fragments that bind specifically but with low affinity (~100 µM to 10 mM).
- Addressing sub-site specificity and ligand efficiency.
- Low molecular weight compounds represent suitable basis for optimization of ADME properties.
- It is generally easier to build up a small molecule than it is to reduce the size of a large one.

Fragment Based Screening Techniques

- **Computational Methods**
  - Ludi, DS De Novo Evolution, Leap-Frog, RACHEL, AlleGrow

- **Experimental Methods**
  - High Throughput X-Ray Screening
    - SGX Pharmaceuticals
    - Astex Therapeutics
  - SAR by NMR

Can we use a Pharmacophore Model for Fragment Based Screening?
Pharmacophore Screening

- Single hit list
Pharmacophore Screening

- Fragment Libraries
Pharmacophore-based De Novo Design

• For optimization
Two Case Examples

- Methotrexate (3DFR)
- Gleevec® (1IEP)
Pharmacophore Guided FBDD Workflow

Reference Query

Fragment Queries

Fragment Database (*.bdb)

Screen Database (catSearch)

Library Enumeration

Hit Refinement Protocol

Protein Target File

Scoring & Prioritization
FBD workflow
Methotrexate

- Original 6 Featured Pharmacophore Model
- Three fragment-based queries
Methotrexate

- Define linker atoms
Methotrexate Pharmacophore Models

- Increase accuracy by using shape constraints
- Constraints can be applied to individual shape models

Three fragment-based queries
Methotrexate: Fragment Screening

• Catalyst/CAP 2004 Reagents
  - 220,902 compounds
• Minimum shape similarity tolerance adjusted individually for each query
  - comparably small hit lists
Fragment results in 3D
Methotrexate: Results

<table>
<thead>
<tr>
<th></th>
<th>Hits CAP</th>
<th>Hits WDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frag1</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Frag2</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Frag3</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>F1xF2xF3</td>
<td>8,424</td>
<td></td>
</tr>
<tr>
<td>6 features</td>
<td>19</td>
<td>135</td>
</tr>
<tr>
<td>6 features Shape</td>
<td>9</td>
<td>61</td>
</tr>
</tbody>
</table>

- Total of 8,424 combinations of unique ligands using high shape similarity cut-off (0.6 - 0.7)
- Original 6 Featured query retrieves only 19 hits

**Are these 8,424 hits useful or are they noise?**
Methotrexate: De Novo Library

- Library of 8,424 compounds enumerated and further investigated
- Re-screening with six-feature query returns 75% of the library

<table>
<thead>
<tr>
<th>Query</th>
<th>CAP (220,902)</th>
<th>De Novo (8,424)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frag1</td>
<td>24</td>
<td>-</td>
</tr>
<tr>
<td>Frag2</td>
<td>13</td>
<td>-</td>
</tr>
<tr>
<td>Frag3</td>
<td>27</td>
<td>-</td>
</tr>
<tr>
<td>6 features</td>
<td>19</td>
<td>6,207</td>
</tr>
</tbody>
</table>
Refinement and Analysis

- Original Pharmacophore Model built from X-Ray structure of methotrexate
- Hits are actually aligned to query inside the active site of the protein
- In situ pose minimization (CFF)
- Scoring
  - LigScore1&2, Jain, PLP1&2, PMF, Ludi 1,2&3
Examples of Hits Aligned with Methotrexate

- One of the high scoring hits using LigScore1 scoring function
Gleevec

• Original 4 Featured Pharmacophore Model
• Two fragment-based queries
Gleevec

• Increase accuracy by using shape
**Gleevec: Results**

<table>
<thead>
<tr>
<th>Query</th>
<th>Hits CAP (220,902)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frag1</td>
<td>296</td>
</tr>
<tr>
<td>Frag2</td>
<td>247</td>
</tr>
<tr>
<td>5 features</td>
<td>20</td>
</tr>
</tbody>
</table>

- Five-feature query returns 20 hits
- Total of 73,112 combinations from fragment search results using default (low) shape similarity cut-off
Can We Recover Gleevec?
Examples of De Novo Structures
FBDD Workflow All-in-one Example

Reference Query → Fragment Queries

Fragment Database (*.bdb) → Screen Database (catSearch) → Library Enumeration → Hit Refinement Protocol → Protein Target File → Scoring & Prioritization
Fragment Databases

- Using pre-existing catalyst databases
  - Accelry’s (CAP reagents)
  - In-house
- Creating a Fragment database on the FLY
FBDD Workflow All-in-one Example

Reference Query → Fragment Queries → Screen Database (catSearch) → Library Enumeration

Fragment Database (*.bdb) → Hit Refinement Protocol

Protein Target File → Scoring & Prioritization
Enumerate From Fragments

- New protocol that joins fragment hits based on proximity of linker atoms
  - Atoms that are within 0.5Å are merged if they are from the same type or replaced if they are of different types
  - The current version of protocol does not fuse rings; only single atoms can be connected
  - The protocol will check for correct valence and chemistry
  - Incorrect chemistries will be corrected using the molecular toolkit from PP
FBDD Workflow All-in-one Example

Reference Query

Fragment Database (*.bdb)

Fragment Queries

Screen Database (catSearch)

Library Enumeration

Hit Refinement Protocol

Protein Target File

Scoring & Prioritization
Hit Refinement

- Adjusting similarity tolerance and Minimum Fit value
- Apply HTS Filter
  - Non-organic atom types,
  - Reactive substructures
  - Incorrect Chemistries
  - Lipinski violations
  - ADME/Tox Filter
- Mapping structures back to all-feature query
- Docking and Scoring
- Diversity Analysis
Typical Pharmacophore Modeling Experiment

$10^x$ molecules against one target

results in a hit list
Well Why Not Do This!

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

... needs a large number of models!
Why Ligand Profiling

- Address **selectivity issue** in drug design
- Identify potential **side-effects, toxicity, or metabolic pathways** early on in your research process
- Screen against databases of known multiple therapeutic targets to discover **new applications for a known compound or marketed drug**
- Search for a potential modes of action of acquired compounds (**parallel screening**)
- Design privileged structure against desired targets, while minimizing probability of binding to any others: **library profiling**
What Do We Need For Ligand Profiling

3D Ligands

Pharmacophore Collection

Target 1
Target 2
Target 3
Target 4
Target 5
...
Target n
Ligand Profiling in DS 2.0

• Three new protocols
  - Ligand Profiler
  - Add to Ligand Profiler Database
  - Remove from Ligand Profiler Database

• Two databases
  - Hypo DB
    - Inte:Ligand pharmacophore database*
      - 1846 individual pharmacophore models
      - 187 targets
  - Ligand Profiler DB
    - User pharmacophore database

*Inte:Ligand, correct as of October 4th 2007
Ligand Profiling in DS 2.0: HypoDB Content

1846 Models / 187 Unique targets

HypoDB Content - By Protein Type

- EC1 - Oxydo reductases: 15%
- EC2 - Transferases: 24%
- EC3 - Hydrolases: 38%
- EC4 - Oxydo Lyases: 2%
- EC5 - Isomerases: 1%
- EC6 - Ligases: 1%
- Bacterial proteins: 2%
- Transduction factor receptors: 2%
- Intracellular transduction: 0%
- Surface signals: 0%
- Transmitters: 1%
- Cellular level: 0%
- Antibody: 2%
- Toxins: 0%
- Extracellular proteins: 2%

1846 Models / 187 Unique targets
HypoDB Database Integration within DS Environment

- HypoDB database closely integrated with DS environment
- Database is searchable
- Database can be selectively screened
Ligand Profiling in DS 2.0

• Ligand Profiler Workflow
  - Input ligand set in SD format from file or GUI
    • Conformations generated on-the-fly
    • Pre-computed conformations can be used
  - Input pharmacophores from CHM file on disk, list of files on disk, or database
    • Fitting done on-the-fly for Ligand/Pharmacophore combinations

• Pharmacophore databases
  - Sets of XML describing and referring to CHM models
  - User DB is located in
    <Sci Tecic_root>/apps/accelrys/ds/public/data/LigandProfilerDB/Standard
  - If installed, Inte:Ligand DB is located in
    <Sci Tecic_root>/apps/accelrys/ds/public/data/HypoDB
Ligand Profiling in DS 2.0

For illustration only.
Example 1: Viral Target Screenings

Parallel Screening: A Novel Concept in Pharmacophore Modeling and Virtual Screening†

Theodora M. Steindl,‡§ Daniela Schuster,‡ Christian Laggner,‡ and Thierry Langer*,‡§

Institute of Pharmacy, Computer Aided Molecular Design Group, University of Innsbruck, Imrain 52c, Austria, Center for Molecular Biosciences Innsbruck (CMBI), Peter-Mair-Str.1, A-6020 Innsbruck, Austria, and Inte:Ligand GmbH, Clemens Maria Hofbauer-Gasse 6, A-2344 Maria Enzersdorf, Austria

Received May 18, 2006

Parallel screening comprises a novel in silico method to predict the potential biological activities of a compound by screening it with a multitude of pharmacophore models. Our aim is to provide a fast, large-scale system that allows for virtual activity profiling. In this proof of principle study, carried out with the software tools LigandScout and Catalyst, we present a model work for the application of parallel pharmacophore-based virtual screening on a set of 50 structure-based pharmacophore models built for various viral targets and 100 antiviral compounds. The latter were screened against all pharmacophore models in order to determine if their biological targets could be correctly predicted via an enrichment of corresponding pharmacophores matching these ligands. The results demonstrate that the desired enrichment, that is, successful virtual activity profiling, was achieved for approximately 90% of all input molecules. We discuss descriptors for output validation, as well as various aspects influencing the analysis of the obtained activity profiles, and the effect of the utilized search modus for screening.
Dataset

5 viral targets
- Highly relevant viral diseases
- HIV infection, Influenza, Common cold, Hepatitis C
- Sufficient number of PDB entries
- Diverse inhibitory mechanisms

50 pharmacophore models
- Structure-based pharmacophore models generated
- Manual model check and processing
- Models include shapes, excluded volume spheres, and hydrogen bond acceptors with fluorine atoms
- Pharmacophore model validation

100 antiviral compounds
- Inhibitors both from PDB complexes and from literature
- Conformational model generation within Catalyst-BEST conformer generation algorithm
  - max. 250 conformers
  - 20kcal above the calculated lowest energy conformation

Question: Will they be attributed to the correct activity profiles?
## Parallel Screening Results

<table>
<thead>
<tr>
<th>50 Models</th>
<th>HIV protease polymerase 1 2 3</th>
<th>HIV RT</th>
<th>Influenza NA</th>
<th>HRV coat protein</th>
<th>HCV</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 Inhibitors</td>
<td>HIV protease</td>
<td>HIV RT</td>
<td>Influenza NA</td>
<td>HRV coat protein</td>
<td>HCV</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example 2: Breaking the barriers between Cheminformatics and Computational chemistry

• Example Workflow
  - Generate list of newly registered compounds
  - Tag Chemist name registering the compounds
  - Screen these compounds against pharmacophore databases
  - Generate an activity profile report for each compound registered
  - Email the report to Chemist registering that compound

RUN THIS REPORT DAILY
Conclusion

- Parallel pharmacophore-based virtual screening is a straightforward and rapid method for the assessment of bio-activity profiles
- Fragment based de-novo method offers new and exciting way to generate novel lead compounds using pharmacophore models
New Technical Features

• Exportable shape descriptors
  - Additional QSAR criterion
• Shape-base filtering with pharmacophore searching
  - Rapidly pre-filter ligands based on geometry
• Pharmacophore hypotheses can be clustered by sub-site
  - Identify addition ligand binding modes
• UI for creating and editing advance pharmacophore features
  - Streamlined functionality formerly in Exclude/OR editor
• Customized pharmacophore features
  - Capability builds on customization options available in Catalyst 4.11
• 3D fingerprints
  - Incorporate spatial information as additional descriptors for QSAR
• Distributed computing
  - Increased ability for course-grain and fine-grain parallelization will accelerate database searching and building
• Partial query searching
  - Specify required and “optional,” pharmacophore points for searching
• SMARTS-based pharmacophore queries to Catalyst format converter
  - Easily convert pharmacophores to Catalyst format to add additional data.
Catalyst in Discovery Studio 2.0
Acknowledgements

- Remy D. Hoffmann
- Konstantin Poptodorov
- Nicolas Triballeau-Hugounencq
- Thierry Langer

Thank You