ChemModeling

Lead Optimization

Mike Lawless
lawless@chemmodeling.com
(636) 329-0333
ChemModeling, LLC
May 21, 2009

Lead Optimization

Goals

- What needs to be optimized?
  - Potency
  - Physiological properties
  - Novelty/IP position
- Reaching one goal may require revisiting others

Scope

- How far from existing leads do you want to go?
  - Retain existing scaffold
  - Move to new lead series

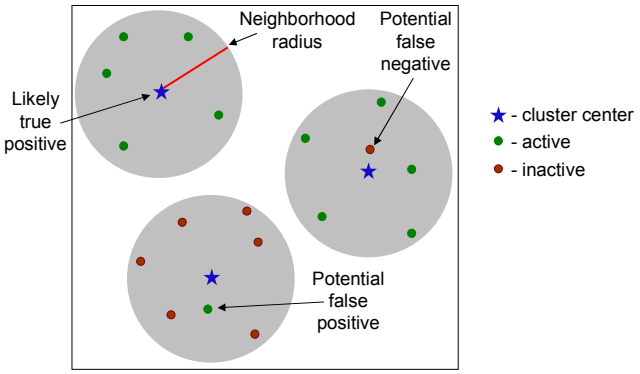
ChemModeling

We have been involved in a large number of lead optimization projects. They ranged from trying to improve the solubility of a potent lead compound to identifying a backup series.

Turning Hits into Leads

Initial screening is performed on a large number of compounds at a fixed concentration

The goal of hit to lead analysis is to identify true positives and potential false negatives



By looking at neighborhood behavior, one can discover possible misidentified compounds

ChemModeling

3

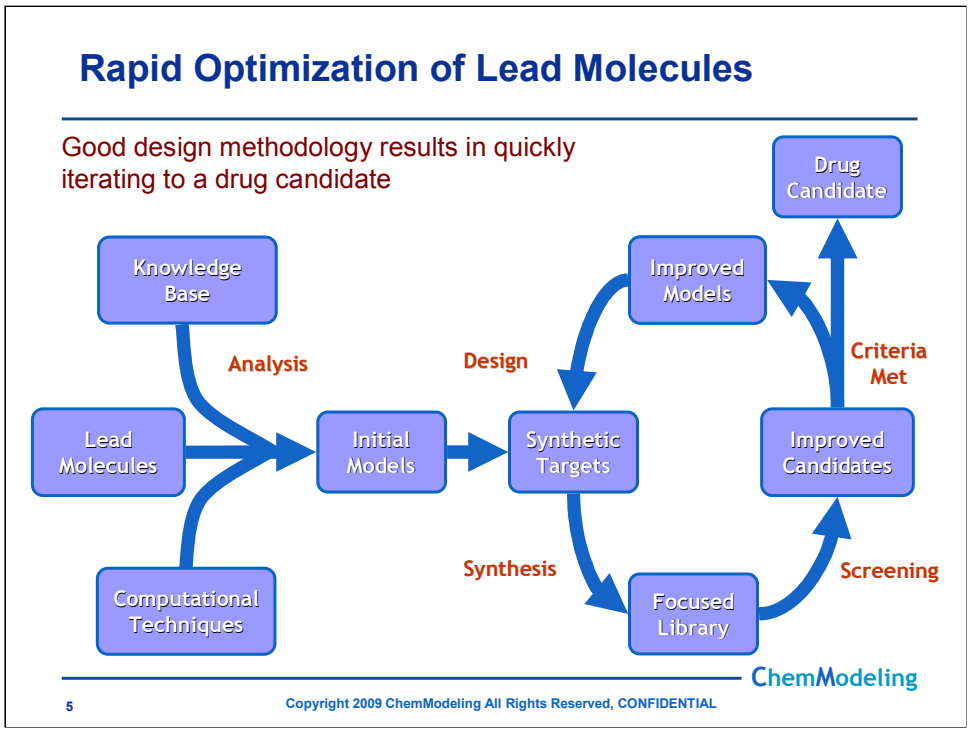
Early in the drug discover process one typically screens a large number of compounds at a fixed concentration to generate percent inhibition data. Analysis of this data is important because there is not always a correlation between percent inhibition and IC50 values. Some of the compounds may not have been pure or the correct structure and so there are good reasons to include additional compounds that didn't show activity in the initial screen or drop compounds that showed activity.

Lead Optimization Process

- Collect knowledge of the target (Knowledge Base)
  - Literature data
  - Customer proprietary/screening data
- Select and implement methodology
  - Physical property profiling/predictions
  - Virtual docking
  - Similarity searching
  - QSAR/CoMFA
- Create and prioritize list of synthetic targets
- Screen compounds and analyze results
  - Iterative process
  - Each set of results provides new knowledge to focus future design

ChemModeling

The best way to optimize a lead depends on the particular problem at hand. Therefore it is important to use information about the target to validate the methods to be used and choose ones for which the actual physical data correlates with the calculated predictions. This is an iterative process and the optimal methods may change as we get closer to a drug candidate.



This is an overview of the complete lead optimization process. It is a knowledge driven approach and medicinal chemistry input is sought in each step of the process.

Physical Property Profiling

Typically libraries are designed to meet Lipinski's "Rule of 5"

Poor absorption or permeation of an orally administered drug is more likely to occur if any two of these criteria are violated:

- Molecular weight is greater than 500
- Lipophilicity is high (ClogP is greater than 5)
- Number of Hydrogen bond donors is greater than 5
- Number of Hydrogen bond acceptors is greater than 10

But in reality properties need to be tailored to target being addressed

Properties of Oral Drugs Categorized by Gene Family

	90% MW	90% ClogP	90% HBD	90% HBA	90% Rbonds
Aminergic GPCRs	460	5.6	2	6	8
Ion Channels	430	4.7	3	6	7
Nuclear Hormone Receptors	495	7.3	2	6	10
Peptide GPCRs	752	6.5	8	10	17
Phospho-diesterases	465	5.2	2	8	9
Protein Kinases	505	5.7	4	7	9
Serine Proteases	572	4.8	4	8	12

Hopkins, et al, Nature Biotechnology **2006,** 7, 805-815

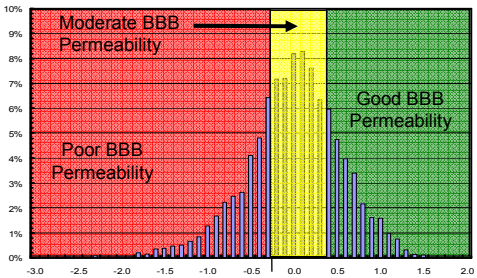
ChemModeling

6

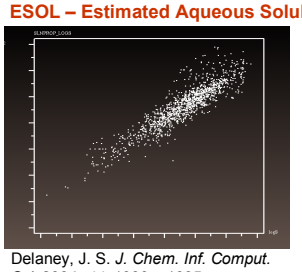
It is important to understand which properties correlate with activity. It may be that a certain range of properties are required for activity or a particular physical property might correlate directly with activity. By using knowledge of compounds that have been screened, it may be possible to identify these trends.

ADMET Predictions

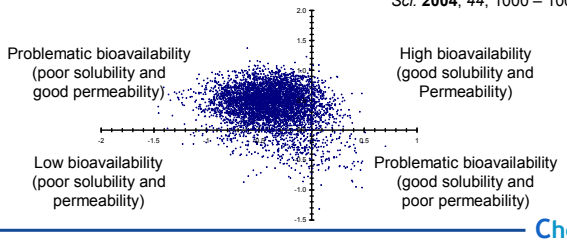
Consider ADMET properties early to minimize problems downstream



ESOL – Estimated Aqueous Solubility



Delaney, J. S. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1000 – 1005.



ChemModeling

7

ADMET properties can be difficult to predict. Ideally, if a particular property is important then it is best to use the data for the lead molecules to generate a local model.

Virtual Docking Strategy

Use crystal structure to identify improvements to lead molecules

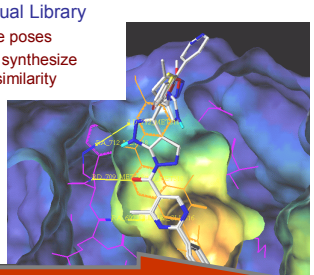
Dock a Diverse Subset of the Virtual Library

- Start with filtered virtual library
- Select by diversity or similarity

	Best	Worst	Avg
P38	-25.6	2.8	-11.6
CDK2	-26.7	2.7	-12.1
Abl	-32.7	-0.8	-17.2
FGFr	-21.9	-5.1	-14.4

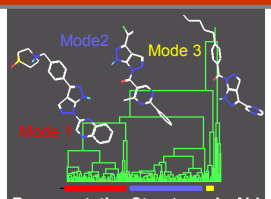
Dock Complete Virtual Library

- Use identified base poses
- Select products to synthesize using scores and similarity



Identify Reasonable Binding Modes

- Cluster on RMS distance between docked structures
- Visually inspect examples from each mode



Representative Structures in Abl

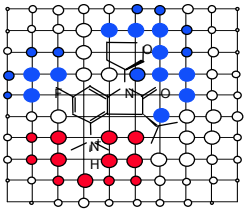
ChemModeling

We have been involved in two types of virtual docking projects. In the first case, lead molecules were docked into the crystal structure in order to identify modifications that would improve potency. In the other case, a large number of compounds were docked in order to identify a new lead series. Multiple crystal structures can be used in order to improve specificity or to mitigate adverse affects.

Similarity Searching

Topomer based searching is effective in searching large virtual libraries

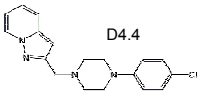
Topomer Fields

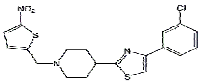


Split molecule into 2-3 fragments
Rule-based alignment of fragments onto a grid
Use probe atom to calculate steric potential at grid points

Topomer Distances

Query Structure





Lead Hop 7.6 uM

Topomer distances are the sum of the pair-wise differences between the fields summed over the fragments plus alignment and steric penalties

An example result from a Tripos validation study is shown above

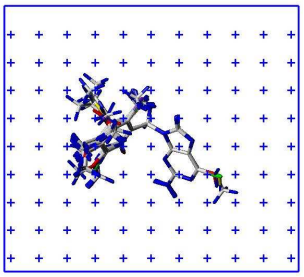
Topomer fields can also be used for CoMFA predictions in virtual libraries

ChemModeling

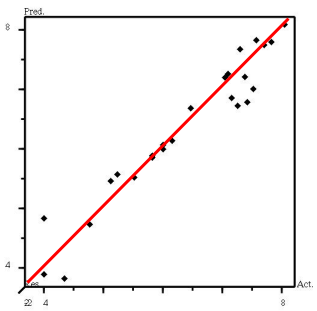
We have several different methods for similarity searching. From project to project different methods are more effect and we will validate these methods with information from the existing leads before using them for lead optimization.

QSAR/CoMFA Model

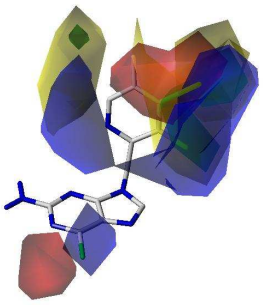
CoMFA models provide more detailed analysis of individual leads and include both positive and negative data in the predictions



Overlay molecules
Compute CoMFA fields



Partial Least Squares Analysis (PLS) to correlate field values with biological activity



Fields indicate areas to increase or decrease steric and electronic properties

ChemModeling

The advantage of a good CoMFA model is that both positive and negative data is incorporated. As the lead optimization process progresses you need to use techniques that provide a more accurate prediction of activity. Techniques that are used to screen compounds are often very approximate because they need to work on very large sets of molecules.

Summary of Lead Optimization

- Methods used will be tailored to customer requirements
  - What is known about the target
  - Quantity and quality of current leads
  - What needs to be optimized
- Large number of computational techniques
  - Multiple methods are likely to be used
  - Alternative methods available for techniques presented earlier
  - Methods validated with customer data to ensure best results
- Coordinate with chemists to ensure best synthetic outcome
- Customer is involved in the entire process
  - Regular meetings to present intermediate results
  - Flexibility to change strategy as information is gained

ChemModeling

We have had considerable experience in optimizing lead. We plan the lead optimization process around the goals of the project and seek input at each step. In this way, the customer can be assured that the result corresponds with their requirements.