ChemModeling

Library Design

Jon Swanson
jon@chemmodeling.com
(636) 329-0300
ChemModeling, LLC
May 1, 2009

The Universe of Potential Drugs is Huge!

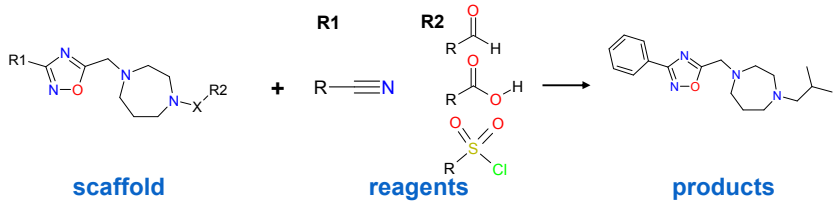
- Estimates vary from 10^{18} to 10^{200}
- Estimate by Regine Bohacek is 10^{60} drug-like molecules*
  - Create a linear compound of up to 30 atoms
  - Use combinations of oxygen, carbon and sulfur for the backbone
  - Add any stable chemical group to the free bonds
  - Add branching, cyclization and stereochemistry
- A lot more than are likely to be made!
- Use Rational Design to focus on highest value compounds
  - Compounds that selectively hit desired targets
  - Eliminate compounds that you wouldn't want even of they did hit
    - Hard to follow-up
    - Poor ADME
    - No patent space
    - Promiscuous
  - Compounds that improve SAR and provide easy optimization

* http://www.nature.com/horizon/chem/chemicalspace/background/explore.html

ChemModeling

Library Designs are Combinatorial

Components of a Combinatorial Library



To Generate a Library-

- Identify scaffold and chemistry to be performed on that scaffold
- Create a reagent pool of acceptable monomers based on the chemistry
- Select a subset of products that spans the activity space
  - Diverse/Representative for general screening libraries
  - Use Docking/Similarity for targeted libraries
- Remove compounds that overlap in activity space with the existing collection
- Generate a final design of the desired size and density from the pool of desirable products

ChemModeling

3

Sources of Scaffolds

From Chemists "chemical intuition"

From "Ideas" Database

From "de novo" Scaffold Generation

Shape-based hierarchical clustering is employed to select diverse scaffolds

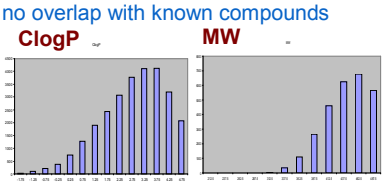
Scaffolds should be vetted for applicability

Scaffold Idea

Representative subset of reagents

"mini" virtual library

Check that library has reasonable property profile and no overlap with known compounds



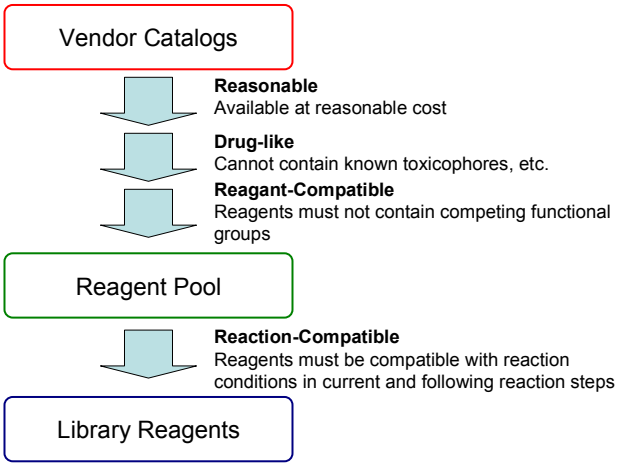
27,417 meet "rule of 5"
14,322 meet "rule of 4.5"
3,918 meet "rule of 4"
using Aldrich reagents

For targeted libraries, use docking or similarity scores to verify desirability of the potential library

ChemModeling

4

Reagents Must Also be Carefully Selected



Depending on the library design goals, custom synthesized reagents may also be used

ChemModeling

5

Products Must Have Reasonable Properties

Lipinski's "Rule of 5" is the best known filtering criteria

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

There are MANY others

=> Rules need to be tailored to specific customers needs

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

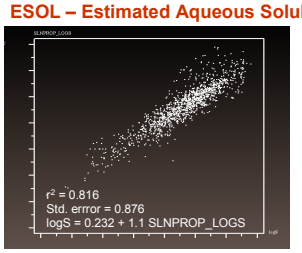
ChemModeling

6

Also Consider Other Factors Relevant to Drug Interactions

- Fragment Based Filters (applied to reagents and products)
  - Unwanted
  - Unstable
  - Toxic groups
- Similarity/Dissimilarity to Known Targets
- Custom Scoring Functions

ESOL – Estimated Aqueous Solubility



Plot of ESOL predicted solubility implemented in slnProperty versus the experimental logS values for compounds used as a training set for ALOGPS program from the ALOGPS website.

log(S) = 0.16 – 0.63 * ClogP
– 0.0062 * MW + 0.066 * RotBonds
– 0.74 AromaticFraction

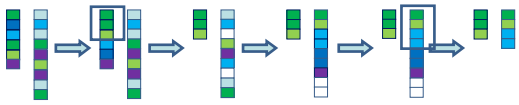
AromaticFraction is fraction of heavy atoms in aromatic 6-membered rings

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

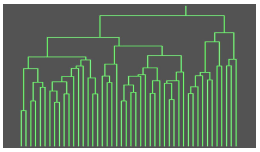
ChemModeling

From the Final Pool a Design is Extracted

A number of methods are used, from simple extraction of a full matrix…



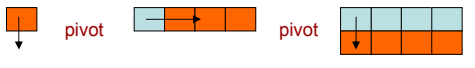
Sort first reagent | Select the top n | Drop products not containing top n | Sort the next reagent | Select the top m | Select products using these nxm



If design is larger than needed, use hierarchical clustering of topomer distances to select diverse reagent subset

to the use of sophisticated mult-objective design programs

OptDesign Uses a Pivoting Method

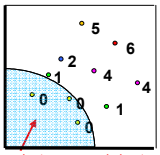


Select an R1 that meets design criteria | Select R2 that combine with R1 to produce locally optimal solution | Select R1 that combine with R2 to produce locally optimal solution

Repeat until a full design is produced

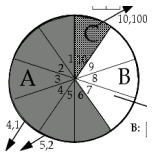
Two techniques used to bias selections

Pareto Ranking



select compounds best meeting all criteria

Roulette Selection



ChemModeling

Library Design is Done Efficiently

A series of proprietary programs are used to insure fast, efficient and reproducible results

Libraries are built from a library definition file that is composed of a scaffold definition and library qualifiers

Library Definition File

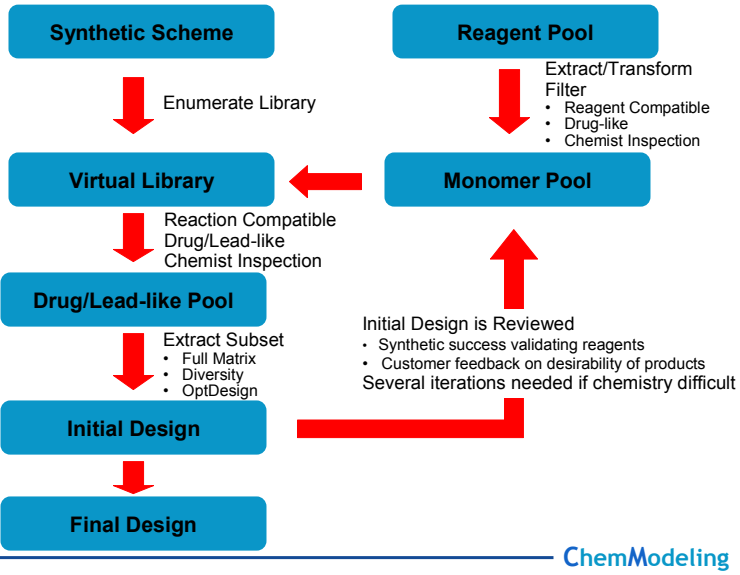
Name	*unique name of library*
Source	*CSLN or library (if making sublibrary)*
Type	*used to group libraries*
Synthons	*lists of reagents at each attachment*
Qualifiers	*can be applied to reagents and products*
Properties	*required property ranges*
2D Inclusion	*required fragments*
2D Exclusion	*disallowed fragments*
List Inclusion	*use specific reagents*
List Exclusion	*disallow specific reagents*
Usage	*ensures efficient use of reagents*

Automating the details means the focus is on the chemistry aspects of the process

ChemModeling

9

The Overall Library Design Process



ChemModeling

10

Summary of the Complete Process

- Design is an Iterative Process
  - What pleases the chemist often doesn't please the computer and vice versa
  - All designs are a balance of competing requirements
- Design is a Step-wise Process
  - A design should start with a good choice of scaffold
  - The preliminary design provides feedback as to the final library size and properties
  - It may not be advisable to proceed to the final design in every case
  - Modification of scaffold and/or reagents may in some cases greatly improve the design
- Decisions that need to be made
  - Source of the synthetic route
  - Source of the reagent lists
  - Exact values for the design constraints
  - Acceptable number of iterations at each stage and criteria to continue forward

ChemModeling