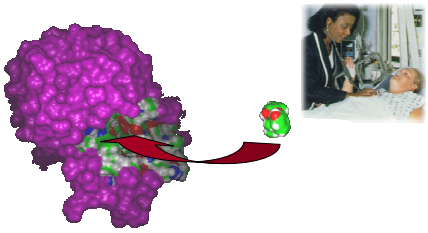
Structure-Based Drug Design

Thomas Funkhouser
Princeton University
CS597A, Fall 2005

Introduction

Drugs
- Molecules that can be introduced to change biological activity



Slide courtesy of Bill Welsh

Introduction

Drug targets
- Enzyme - inhibitors
- Receptor - agonists or antagonists
- Ion channels - blockers
- Transporter - update inhibitors
- DNA - blockers

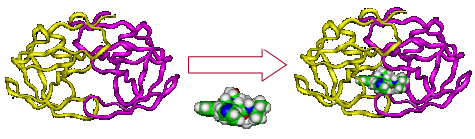
Slide courtesy of Bill Welsh

Structure-Based Drug Design

Goal:
- Given a protein structure,
  and/or its binding site,
  and/or its active ligand (possibly bound to protein),
  find a new molecule that changes the protein's activity

Structure-Based Drug Design

Receptor-based drug design:
- Given a protein structure,
  and/or its binding site,
  and/or its active ligand (possibly bound to protein),
  find a new molecule that changes the protein's activity

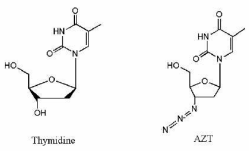


HIV Protease Inhibitor

Example courtesy of Bill Welsh

Structure-Based Drug Design

Ligand-based drug design:
- Given an protein structure,
  and/or its binding site,
  and/or its active ligand (possibly bound to protein),
  find a new molecule that changes the protein's activity



Example courtesy of Joe Corkery

Challenges

Scoring of chemical models
- Activity
- Toxicity
- “Druggability”
- etc.

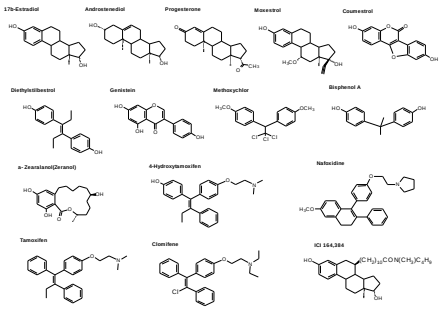
Search of chemical space
- Add/remove/replace chemical groups
- Conformations
- etc.

Outline

Virtual drug screening
- Ligand-based
  - § 2D structure matching
  - § 3D structure matching
  - § QSAR

De novo drug design
- Models
  - § Simulation
  - § Knowledge-based
- Construction algorithms
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

Ligand-Based Drug Screening

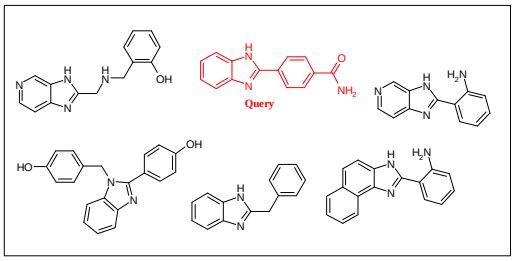


Ligands for Estrogen Receptor Slide courtesy of Bill Welsh

Ligand-Based Drug Screening

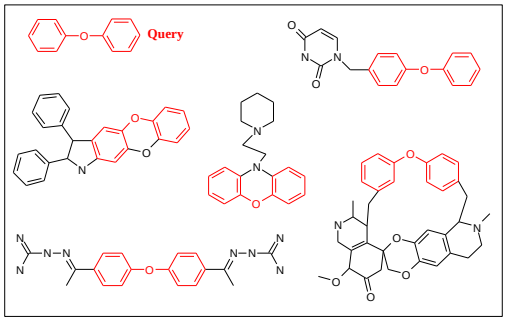
Strategies:
- 2D matching
- 3D matching
- Pharmacore matching
- Quantitative Structure Activity Relationships (QSAR)

2D Structure Matching



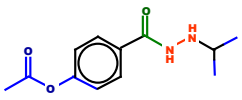
Slide courtesy of Bill Welsh

2D Substructure Matching



Slide courtesy of Bill Welsh

2D Substructure Matching

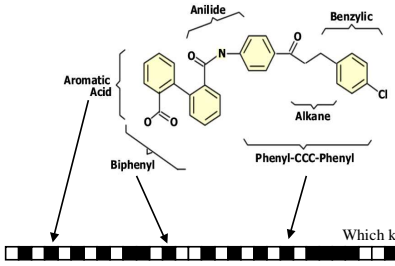


Dictionary of Keys

N-N
O-C(-N)-C
CH_3-Ar-CH_3
C-N-N
N-Ar-Ar-O
N-C-O
N-Ar-O
OH > 1
CH_3 > 1
N > 1
NH
...

Slide courtesy of Bill Welsh

2D Substructure Matching



Slide courtesy of Bill Welsh

2D Substructure Matching

Tanimoto coefficient:

$$T = \frac{(A \cap B)}{(A \cup B)}$$

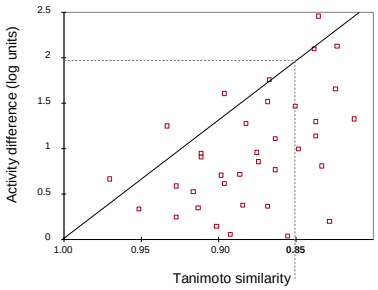
struct A:	00010100010101000101010011110100	13 bits	
struct B:	00000000100101001001000011100000	8 bits	
A & B:	00000000000101000001000011100000	6 bits	A ∩B
A or B:	00010100110101001101010011110100	15 bits	A ∪B

T = 6 / 15 = 0.4

Slide courtesy of Bill Welsh

2D Substructure Matching

*Activity data from Uehling *et al*, J. Med.Chem. 1995



Slide courtesy of Bill Welsh

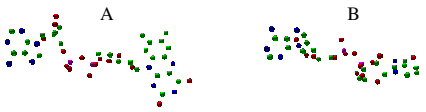
Outline

Virtual drug screening
- Ligand-based
  - § 2D structure matching
  - Ø **3D structure matching**
  - § QSAR

De novo drug design
- Models
  - § Simulation
  - § Knowledge-based
- Construction algorithms
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

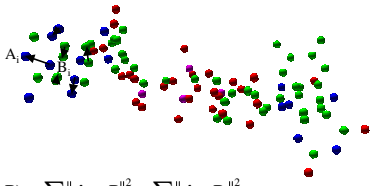
3D Structure Matching

Search database of molecules for ones with similar 3D shape and chemistry



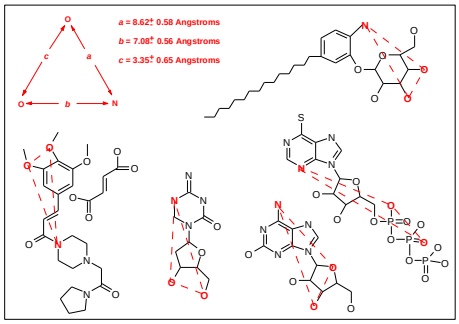
3D Structure Matching

Search database of molecules for ones with similar 3D shape and chemistry



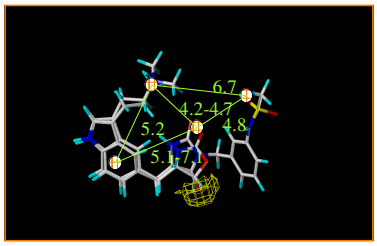
$$d(A,B) = \sum_{A_i \in A} \|A_i - B\|^2 + \sum_{B_i \in B} \|A - B_i\|^2$$

3D Substructure Matching



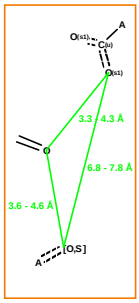
Slide courtesy of Bill Welsh

3D Substructure Matching

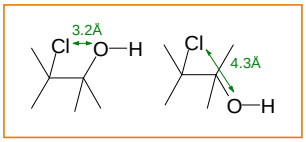


Slide courtesy of Bill Welsh

3D Substructure Matching



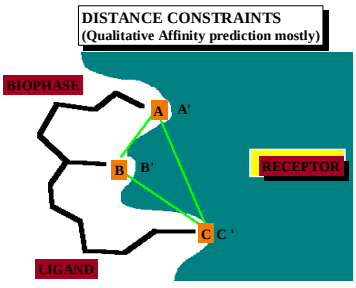
Rigid



Flexible

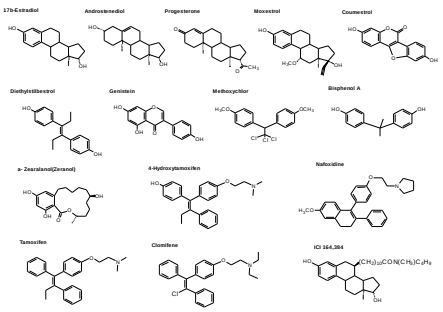
Slide courtesy of Bill Welsh

3D Substructure Matching



Slide courtesy of Bill Welsh

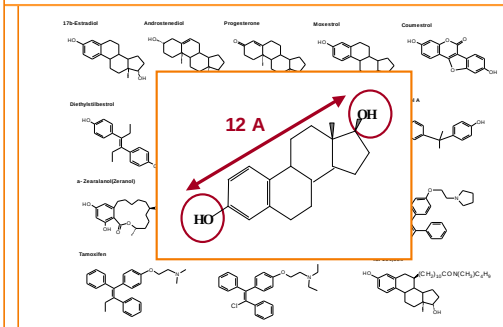
Pharmacore Matching



Ligands for Estrogen Receptor

Slide courtesy of Bill Welsh

Pharmacore Matching



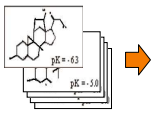
Ligands for Estrogen Receptor Slide courtesy of Bill Welsh

Outline

Virtual drug screening
- Ligand-based
  - § 2D structure matching
  - § 3D structure matching
  - Ø QSAR ←
- De novo drug design
- Models
  - § Simulation
  - § Knowledge-based
- Construction algorithms
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

QSAR

Learn model for activity as function of descriptors (properties) computed from molecules

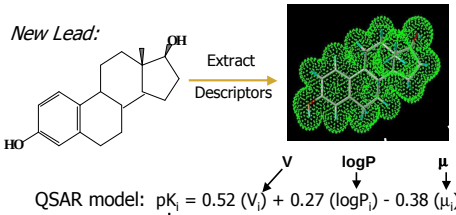


Compound	Activity (pK_i) "Y"	Descriptors (X_i) Mol. Vol. (Å^3)	LogP	Dipole Mom (μ)
1	2.34	420	2.8	0.97
2	1.89	332	4.6	2.23
3	0.23	198	-0.3	3.36
4	3.67	467	3.7	0.45
5	2.55	359	-1.5	1.77
etc.	etc.	etc.	etc.	etc.

Slide courtesy of Bill Welsh

QSAR

Use model to predict activities for new leads from their descriptors



QSAR model: $pK_i = 0.52\ (V_i) + 0.27\ (\log P_i) - 0.38\ (\mu_i)$

Predicted activity of new lead

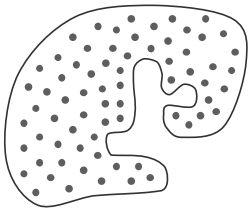
Slide courtesy of Bill Welsh

Outline

Virtual drug screening
- Ligand-based
  - § 2D structure matching
  - § 3D structure matching
  - § QSAR
- De novo drug design ←
- Models
  - § Simulation
  - § Knowledge-based
- Construction algorithms
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

De Novo Drug Design

General Strategy:
- ØGiven a protein structure
- Build model of binding site
- Construct molecule that fits model



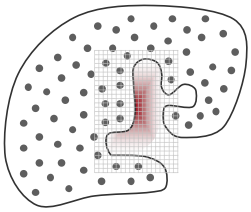
De Novo Drug Design

General Strategy:

- Given a protein structure

ØBuild model of binding site

- Construct molecule that fits model

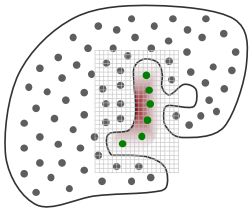


De Novo Drug Design

General Strategy:

- Given a protein structure
- Build model of binding site

ØConstruct molecule that fits model



Outline

Virtual drug screening

- Ligand-based
  - § 2D structure matching
  - § 3D structure matching
  - § QSAR

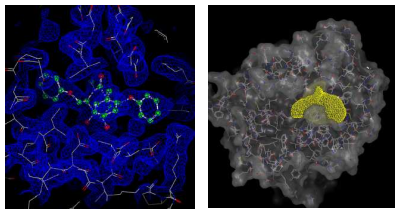
De novo drug design

ØModels ←

  - § Simulation
  - § Knowledge-based

- Construction algorithms
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

Binding Site Models



Example courtesy of Joe Corkery

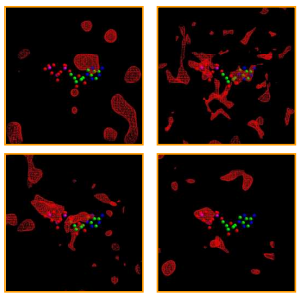
Binding Site Models

Simulation

- e.g., GRID

Knowledge-based

- e.g., X-SITE



Predicted Binding Site Model for 1kp8-1-H-ATP-1-_ using GRID

[Goodford85]

Outline

Virtual drug screening

- Ligand-based
  - § 2D structure matching
  - § 3D structure matching
  - § QSAR

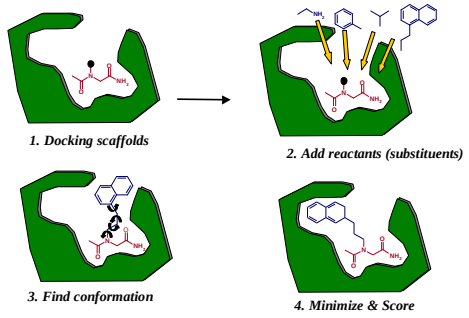
De novo drug design

- Models
  - § Simulation
  - § Knowledge-based

ØConstruction algorithms ←

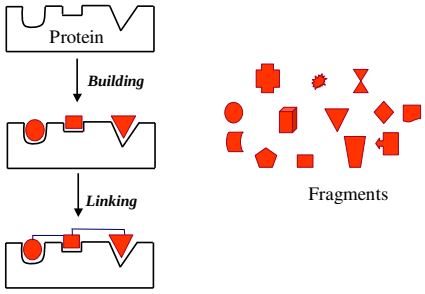
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

Incremental Construction



Slide courtesy of Bill Welsh

Fragment-Based Methods

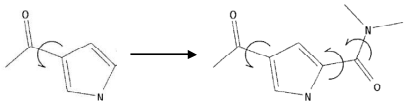


Slide courtesy of Bill Welsh

Stochastic Optimization

Monte Carlo search of chemical space:

- Start from initial drug
- Make random state changes (add/delete/move chemical group)
- Accept up-hill moves with probability dictated by "temperature"
- Reduce temperature after each move
- Stop after temperature gets very small



Summary

Virtual drug screening

- Ligand-based
  - § 2D structure matching
  - § 3D structure matching
  - § QSAR

De novo drug design

- Models
  - § Simulation
  - § Knowledge-based
- Construction algorithms
  - § Incremental construction
  - § Fragment-based
  - § Stochastic optimization

Discussion

?