The Pharmacophore Concept
(Active-site Mapping)

Most therapeutic targets are macromolecules with unknown 3D structures. Often, the structure of a diverse set of compounds that bind to the therapeutic target are known through combinatorial synthesis of compound libraries and high-throughput screening. If the active compounds have sufficiently diverse structures, then one can use the computer to search for common surfaces (the Pharmacophore) that the receptor might recognize (active analog approach).
Things to know, 60 minutes from now

- What is a pharmacophore?
  - Pattern recognition
  - Distance constraints
  - Degrees of freedom and conformational hyperspace, systematic search algorithms

- Quantitative structure-activity relationships (QSAR, COMFA, PCA and PLS)

- How to derive constraints from physical interactions between ligand and receptor?

- How to apply these constraints to predict active 3D conformations of bound ligands?

- How to identify a good (or bad) prediction?

Pharmacophore - Definition

- Paul Erlich, early 1900
  - "a molecular framework that carries (phoros) the essential features responsible for a drug's (pharmacon) biological activity"

- Peter Gund, 1977
  - "a set of structural features in a molecule that is recognized at a receptor site and is responsible for that molecule's activity"

- Wikipedia, today
  - A pharmacophore is a 3-dimensional substructure of a molecule that carries (phoros) the essential features responsible for a drug's (pharmacon) biological activity. Alternatively described as an ensemble of interactive functional groups with a defined geometry. Basically, one tries to talk the protein language by finding the "structural and chemical complementarities" (aka pharmacophore hypothesis) to target receptors.
Quiz: which of these is active?

- How to judge whether or not a molecule might be active, based only on structural info?

Pharmacophore - Constrained Minimization to Overlap Proposed Pharmacophoric Groups

Tests feasibility – if it possible that groups can overlap?

Any solution is not necessarily unique!

Answer depends on starting conformations
Pharmacophore modeling

- Analog-based pharmacophores
  - Unknown receptor
  - Information from active and inactive molecules
  - 3D-QSAR

- Receptor-based pharmacophores
  - Hypothetical receptor site based on function (enzyme)
  - Known (and characterized) receptor (crystal structure)
  - Docking

Quantitative Structure-Activity Relationships (QSAR)

- Activity = function(structural/chemical properties)

  - Example*
    - Relate biological activity to electronics and hydrophobicity
      \[ \log(1/C) = k_1 \log P - k_2 (\log P)^2 + k_3 s + k_4 \]
      - \( C \): concentration of compound that gives a response
      - \( P \): partition coefficient between water and 1-octanol
      - \( k_1, k_2, k_3, k_4 \): constants
      - \( s \): Hammett substituents parameter

**QSAR process**

- Synthesize & test biological activity for diverse set of ligands (including actives and inactives)

- Clever experimental design maximizes information content from a series

- Regression techniques to fit an equation to the data

- Which properties are correlated with activity?

- Cross-validate

**QSAR Objective**

- Correlate 3D structure of a ligand with its biological activity

- Problems
  - Frame of reference
  - Unknown steric interactions
  - Multiple binding modes
  - Underdetermined system
A simple example

Visual Pattern Recognition

- Visual identification of common structural and chemical features among active molecules and those features that are missing in the inactive ones
- Measurement of the 3D aspects of the common features, w.r.t. each other
- Development of a draft pharmacophore
- Validation that the pharmacophore fits the active compounds and fails to fit the inactive ones
- Refinement of the model by applying it to a database of compounds with known activity, until the desired result is reproduced
Visual Pattern Recognition

One six-membered aromatic ring: phenyl or pyridyl

Visual Pattern Recognition

Second 5- or 6-membered ring
Visual Pattern Recognition

Urea group or amide functionality in 2nd ring

Visual Pattern Recognition

Ring systems are side-by-side
Measurements

- From the amide nitrogen to the center of the aromatic ring is approx 5-6 Å
- The aromatic ring and amide group are within 0.5 Å RMSD from planarity

Developing a pharmacophore model

- Phenyl or pyridyl ring
- 5- or 6-member ring with link option
- Additional nitrogen, various bond types
- Distance and planarity constraints
Pharmacophore vs. Active Site Models

- Pharmacophore modeling with assumed ligand groups (A=A’, B=B’, C=C’)
- Active site modeling with receptor groups conceptually linked to ligand (X, Y, Z)

Angiotensin Converting Enzyme (ACE):
A useful 3D example

- Historically important example
  - Evolution of rational drug design
  - Benchmark system for pharmacophore model

- Therapeutically relevant

- Clear illustration of balance between distance map resolution and sampling conformational hyperspace

- Have generated multibillion dollars in sales
ACE “makes” Angiotensin II in vivo

Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Leu
(C-terminus of Angiotensinogen)

Renin

Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu
(Angiotensin I)

ACE

Asp-Arg-Val-Tyr-Ile-His-Pro-Phe (Angiotensin II)
Biological function of ACE

- Overall pressor (pressure-raising) effect
  - Converts decapeptide Angiotensin I to Angiotensin II
    - All = potent vasoconstrictor
    - Stimulates release of steroid aldosterone (increased sodium retention)
  - Hydrolyzes C-terminal peptide from the hypotensive (vasodilator) nonapeptide bradykinin
    - Stimulates antidiuretic hormone (ADH) release

- ACE action increases both vascular resistance and blood volume

- Inhibiting ACE results in antihypertensive effect

How to Inhibit ACE?

- Bradykinin-Potentiating Peptide (BPP)
  - Naturally occurring nonapeptide in Bothrops jararaca venom
  - BPP$_{5a}$ = pGlu-Lys-Phe-Ala-Pro
  - Phe-Ala-Pro also inhibits ACE
  - Many studies to deduce structure-activity relationship with Phe-Ala-Pro mimics
    - Complicated by pKa of titratable groups, etc
    - Simple computer models
  - Several inhibitors emerge as potent
  - Good peptide-analog correlation for captopril
  - Poor peptide-analog correlation for enalapril
Captopril

- designed based upon BPP\textsubscript{5a}
- First ACE inhibitor drug
  - Orally available
  - Potent
- Undesirable clinical side effects
  - Loss of taste
  - Rashes
- At therapeutic doses, does not interact with nervous system and cause side effects
- Similar sulfhydryl moiety as penicillamine, similar side effects

How to accurately deduce an active conformation?

- Apply the active analog hypothesis
- Choose a potent inhibitor (IC\textsubscript{50} < 50nm)
- Scan systematically through all torsions
  - Generate many conformations
  - Reject those with improbably large energy (VdW criterion)
  - # of conformers = (360/angle)^N - rejected
    - N=# of rotatable bonds
  - Small angles or large # of bonds \rightarrow Combinatorial explosion!
- Create distance map (DMAP) relating structural features relevant to binding site
DMAP = Atlas driving distances

Molecule = United States

Atomic geometry = Geography
Grid definition

• How know if two DMAPs are equal?
  – Define a 3D grid around molecule
  • Two conformers evaluate to same grid points?
  • Effective resolution of predicted DMAP
    • ~0.1-0.5 angstrom
    – Fine grid rejects more conformations

• Balance angle resolution with grid spacing!

ACE inhibitors
generated by pharmaceutical industry to provide IP position for drug development.

2D chemical sketches of potent ACE inhibitors (#1–29) and two inactive compounds used as negative controls (#30–31). All chiral centers are in the S configuration unless explicitly noted here as R configuration. Compounds used in clinics have generic and trade names highlighted.
**DMAP definition**

- Du = Zn
- DMAP valid for all active analogs
  - Zn
  - C=O
  - COOH
- DMAP can become many-dimensional for some systems

**Constrained Conformational Search**

- Given a series of potent analogs
  - Most rigid has fewest torsions \( \rightarrow \) fewest conformers
  - Bind same site, common DMAP(s)

- Scan, reject, evaluate DMAP
  - Start with most rigid \( \rightarrow \) keep all valid DMAPs
  - Next most rigid \( \rightarrow \) keep only common DMAPs
  - Iterate until converge on DMAP, or run out of information in the analog series

- Evaluate all analogs w.r.t. the most constrained DMAP (i.e., the final iteration)

- Also superconstrained case, same DMAPs
Overlap of crystal structure of complex of the inhibitor lisinopril with ACE and the predicted enzyme-bound conformation of ACE inhibitors by Mayer et al. JCAMD, 1987. Note overlap between positions of pharmacophoric groups interacting with zinc (orange), C-terminal carboxyl and carbonyl oxygen of amide, the groups targeted by active site modeling. The phenyl group common to enalapril analogs such as lisinopril (white ring) was not constrained (green ring) by analogs at the time.

Convergence of DMAPs for ACE

Comparison with 2004 crystal structure of captopril bound to ACE

<table>
<thead>
<tr>
<th>Predicted model</th>
<th>Crystal structures</th>
<th>Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.188 - 7.812</td>
<td>-1.1</td>
</tr>
<tr>
<td>2</td>
<td>4.812 - 5.312</td>
<td>-0.6</td>
</tr>
<tr>
<td>3</td>
<td>3.582</td>
<td>0.0</td>
</tr>
<tr>
<td>4</td>
<td>4.812 - 5.002</td>
<td>-0.1</td>
</tr>
<tr>
<td>5</td>
<td>3.938</td>
<td>-0.0</td>
</tr>
</tbody>
</table>


Overlap of 3 ACE Inhibitors (Captopril, Enalopril and Lysinopril) bound to ACE. Alignments are based on ACE structure. Note movement of Zinc (orange net) to accommodate SH vs COOH complexes.
**Volume mapping**

体积映射

**Volume Union of active compounds**

体积联合活性化合物

**Receptor Biophase Interface**

受体生物相界面

**Pharmacophore**

药效点

**Extra Volume of inactive compounds**

额外体积无活性化合物


**Analysis of Excluded Volume**

排除体积的分析

图像：

**UNION OF ACTIVE COMPOUNDS - RECEPTOR EXCLUDED VOLUME**

联合活性化合物 - 受体排除体积

=
Analysis of Excluded Volume

Example: Volume Mapping

Fig. 1. Structures of active and inactive amino acid analogues used in volume mapping of the l-methionine binding site of methionine adenosyltransferase.
Use of Predicted Bound ACE inhibitors to Map Enzyme Sterically Allowed Volume

1275 conformations of the inhibitor ramipril (Altace) predicted at the active site of ACE, compared with the experimental structures of lisinopril, enalapril, and captopril. The molecular surfaces of lisinopril (blue), enalapril (red), and captopril (yellow) are shown to overlap significantly with the predicted ramipril structures (green). All three inhibitors bind to the identical active site on ACE as predicted. The predicted ramipril conformations fit within the common volume of the experimentally determined inhibitors. The orange mesh represents an electron density contour encompassing all zinc loci observed in the crystal structures. Only the constrained portion of ramipril is shown.
Concept/language check

- Conformational hyperspace
- Distance Map
- Sampling
- Resolution
- Active analog hypothesis
- QSAR
- Excluded Volume

Recap

- What is a pharmacophore?

- Give an example strategy for sampling conformational hyperspace
  - Pros/Cons to this strategy?

- Quickly outline QSAR
  - How to derive constraints from physical interactions between ligand and receptor?
  - How could you use constraints to predict 3D conformations of a bound ligand?
  - Strategy for identifying a good (or bad) prediction of activity for a particular ligand?
Things to know, 60 minutes from now

- What is a pharmacophore?
  - Pattern recognition
  - Distance constraints
  - Degrees of freedom and conformational hyperspace

- Quantitative structure-activity relationships (QSAR)

- Alignment rule for compounds is essential for 3D QSAR

- How to derive constraints from physical interactions between ligand and receptor?

- How to apply these constraints to predict active 3D conformations of bound ligands?

Systematic Search Methodology (Combinatorials Consume CPUs)

1. Prune combinatorial tree ASAP

2. Divide and conquer (start at middle of graph)

3. Use analytical approaches to determine solution space (where to look) rather than stochastic processes to find it

4. Transform problem with constraints on solutions - start with most constrained example to provide tightest bounds on search
Computational Complexity of Systematic Search

The conceptual simplicity of systematic search is in sharp contrast to the combinatorial complexity of its calculation. If $A$ is the torsion angle increment and $T$ is the number of rotatable bonds in the molecule, then the total number of possible conformers to be examined for steric conflict is $360/A^T$. If $N$ is the number of atoms in a molecule, $N(N - 1)/2$ pairwise van der Waals evaluations must be done for each conformation. Consequently, the number of pairwise van der Waals evaluations, $V$, required for a molecule during the course of a systematic search is given by:

$$V = \left( \frac{360}{A} \right)^T \times \frac{N(N - 1)}{2}$$

Table 1
Relative computational complexity of systematic search as a function of the number of torsions and the angle increment. These values were determined using the expression for $V$ (Eq. (1)) and simple hydrocarbons — 5 torsions corresponds to hexane, 10 torsions corresponds to undecane, etc.

<table>
<thead>
<tr>
<th>Number of torsions</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle increment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>$9.1 \times 10^1$</td>
<td>$2.1 \times 10^{13}$</td>
<td>$3.2 \times 10^{14}$</td>
</tr>
<tr>
<td>15</td>
<td>$3.2 \times 10^1$</td>
<td>$9.1 \times 10^4$</td>
<td>$2.2 \times 10^{15}$</td>
<td>$3.5 \times 10^{16}$</td>
</tr>
<tr>
<td>8</td>
<td>$7.4 \times 10^2$</td>
<td>$5.0 \times 10^{11}$</td>
<td>$6.5 \times 10^{26}$</td>
<td>$2.9 \times 10^{27}$</td>
</tr>
<tr>
<td>4</td>
<td>$2.4 \times 10^3$</td>
<td>$5.0 \times 10^{14}$</td>
<td>$6.8 \times 10^{35}$</td>
<td>$3.2 \times 10^{36}$</td>
</tr>
<tr>
<td>2</td>
<td>$7.6 \times 10^5$</td>
<td>$5.0 \times 10^{17}$</td>
<td>$7.1 \times 10^{45}$</td>
<td>$3.5 \times 10^{46}$</td>
</tr>
</tbody>
</table>

Combinatorial Tree Truncation

Fig. 1. The tree structure of systematic search. At each branch point, the possible addition of a new aggregate to the existing partial conformation is evaluated for steric contacts. Gray lines represent "pruning" of the search tree by eliminating from further consideration those branches in which the addition of an aggregate is not sterically allowed at any torsion setting. The process continues until the addition of all aggregates along every branch has been considered.
Distance constraint equations describe the variable distance between any two atoms as a function of a single torsion angle. The square of the interatomic distance between \( a_i \) and \( a_j \) is given by:

\[
d_i^2(\omega) = d_{ij\text{const}} + d_{ij\text{tor}}(\omega)
\]

where the coefficients are defined as follows:

\[
d_{ij\text{const}} = \left| \mathbf{r}_i - \mathbf{r}_j \right|^2 = 2(x_{i1}x_{j1} + x_{i2}x_{j2})
\]

\[
d_{ij\text{tor}}(\omega) = -2(x_{i1}x_{j1} + x_{i2}x_{j2})
\]

\( x_{i1}, x_{i2}, \text{ and } x_{j1}, x_{j2} \) are the three orthogonal components of the vector \( \mathbf{r} \) in Fig. 2. \( x_{i1} = x_{j1} = x_{i2} = x_{j2} = x \).

By a substitution of variables \( x = \tan \frac{\omega}{2} \), Eq. (2) simplifies to

\[
d_i^2(\omega) = (a^2 + b + c)\left(1 + \frac{x^2}{a^2}ight)
\]

and \( r = d_i + d_j \).

The values of \( x \) which minimize or maximize \( d_i^2(\omega) \) are given by

\[
x = \pm \sqrt{\frac{d_j}{d_i}}
\]

These values of \( x \) may be substituted into Eq. (3) to find the minimum and maximum distances between atoms \( i \) and \( j \) as a function of the torsion variable.

\[\text{Fig. 2. Geometric definitions in the evaluation of steric contacts using the distance constraint equations during the addition of an aggregate to a sterically allowed partial conformation: } a_i, a_j, a_k, a_l, a_m, a_n, a_o, a_p, a_q, a_r, a_s, a_t, a_u, a_v, a_w, a_x, a_y, a_z.\]

\[\text{REFERENCES}\]

