The Shape of Things to Come: Structural Biology and Drug Development

Jeremy M. Berg

Fifth Annual Ri.MED Scientific Symposium
October 24, 2011
Knowledge of the 3-D structure of a drug target should allow design of molecules that bind to observed pockets.
The energetics of protein-ligand interactions are complicated and nuanced.

Both proteins and ligands can be quite flexible.

Many target-binding ligands are not good drug candidates.

The structures of many important drug targets are difficult to determine.
Protein-Ligand Binding

Simplistic view
Protein-Ligand Binding

Solvation
Protein-Ligand Binding

Flexibility
Protein-Ligand Binding

Dynamics
Iterative Structure-Based Drug Design
Compound Optimization

* Optimize binding affinity for target (based on structures)
* Additional considerations:
  * Other measures of efficacy
  * Specificity for target compared with other related proteins
  * Physical properties (e.g. solubility)
  * Bioavailability
  * Ease of synthesis
HIV Protease
A Starting Compound
Iterative Design

<table>
<thead>
<tr>
<th>R =</th>
<th>$IC_{50}$ (nmol)</th>
<th>log ($P$)</th>
<th>$c_{max}$ ($\mu$M)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.4</td>
<td>4.67</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
<td>3.70</td>
<td>&lt; 0.1</td>
</tr>
<tr>
<td></td>
<td>0.3</td>
<td>3.69</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>2.92</td>
<td>11</td>
</tr>
</tbody>
</table>
HIV Protease-Drug Complex
Impact of HIV protease inhibitors in combination therapy
Progress in Structural Biology

Protein Data Bank:

Founded in 1971 for storing crystallographic coordinates

13 deposited structures in 1976

80,000 structures expected by end of 2011
Challenges for Structural Biology: Membrane Proteins

* Physiologically important including many drug targets
* Difficult to express and purify to homogeneity
* Difficult to crystallize
Signaling proteins that act by stimulating the exchange of GTP for GDP in associated heterotrimeric G proteins

Large family of membrane proteins characterized by the presence of 7 transmembrane helices

Targets of approximately 40% of known drugs!

2.4 Å structure revealed residues 29-342 (out of 413), bound partial inverse agonist carazolol, bound palmitic acid, and three cholesterol molecules
3-Dimensional Structure
β-2 Adrenergic Receptor

* Protein engineering
* Robotic system crystallization
* Microfocus beamline (Argonne National Laboratory)
* Culmination of two decades of effort by Brian Kobilka beginning with receptor cloning
β-2 Adrenergic Receptor Sequence

MGQPNGSAFLLAPNRSHAPDHDTQQRD
EVWVVMGIVMVLAVLFVNVLVITAIAKFEFLQ
TVTNYFITSALACDLMGLAVVPFGAAMHLKMWTT
FGNEFCEFWSIDVLCTASLETLCVIAVDRYFAITSPFKYQSSL
TKNKARVIILMVWIVSGLTSFLPIQMHWYRATHQEAICNYANETCCDFFT
NQAYAIASSIVSFYVPLVIMAFYVSRFQEAKRQLQKIDKSEGRFHVQNSQVEQDGRTHGRLRRSSKFCL
KEHKALKTLGILNTFTLCWLPFFFIVNIVHVIQDNLIR
KEVYILLNNTGYVNSGFNPILYCR
SPDFRIAFQELLCLRSSLKAYGNGYSSNGNTGEQSGYHVEQKEKNLCCEDLPGETDFVGHQGTVPDSNIDSQGRNCSTNDSL
β-2 Adrenergic Receptor Structure
Carazolol

Palmitate

Cholesterol (X3)

Maltose
Other G Protein Coupled Receptor Structures

- (Bovine rhodopsin)
- Human adenosine A (2A) receptor
- Human histamine H1 receptor
- Turkey β1 adrenergic receptor
- Human dopamine D3 receptor
* Iterative structure-based drug design is now a proven method
* Structures can guide medicinal chemistry to yield novel and efficacious structures
Accurate computational assessment of ligand affinities remains problematic

Improved algorithms and enhanced computational power are available

(Department of Computational and Systems Biology)
* The structures of many potential drug targets are not known
* Progress on structure determination methods continues

(Department of Structural Biology)
Peroxisomes

* Membrane-bound organelles
* House enzymes associated with:
  * Hydrogen peroxide metabolism (catalase)
  * Long chain fatty acid oxidation
  * Plasmalogen, bile acid biosynthesis
  * Purine catabolism
Peroxisomal Protein Targeting

* Targeting sequence: -SKL at carboxyl terminus
* Some conservative substitutions are tolerated: -(S,C,A)-(K,R,H)-(L,M)-COO⁻
Peroxisome Biogenesis Machinery

The Structure of Pex5p

First Pex5p structure previously solved by Greg Gatto, MD, PhD
# The Human Peroxisomal Proteome

## Range of Pex5p-PTS1 Dissociation constants:

1.6 nM -> 25 μM

## Key proteins:

- **Acyl-CoA oxidase 1**: 5.6 nM
- **Catalase**: 1.2 μM

<table>
<thead>
<tr>
<th>Protein (Gene Name)</th>
<th>Sequence</th>
<th>$K_a$, nM (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acyl-CoA oxidase 1 (ACOX1)</td>
<td>YHKSLSKL</td>
<td>5.6(1.0)</td>
</tr>
<tr>
<td>Acyl-CoA oxidase 2 (ACOX2)</td>
<td>YLQSNRSKL</td>
<td>257(5)</td>
</tr>
<tr>
<td>Acyl-CoA oxidase 3 (ACOX3)</td>
<td>YVGSSKL</td>
<td>1.6(1.0)</td>
</tr>
<tr>
<td>L-Bifunctional enzyme (EHADH)</td>
<td>YAGSPSSLSL</td>
<td>1096(156)</td>
</tr>
<tr>
<td>D-Bifunctional enzyme (HSD17B4)</td>
<td>YILKDYYKL</td>
<td>195(62)</td>
</tr>
<tr>
<td>Sterol carrier protein 2 (SCP2)</td>
<td>YLQPGNAKL</td>
<td>380(26)</td>
</tr>
<tr>
<td>3,2-trans-enoyl-CoA isomerase (PECl)</td>
<td>YFLSDKSL</td>
<td>14(3)</td>
</tr>
<tr>
<td>2,4-dienoyl-CoA reductase 2 (DECR2)</td>
<td>YFASFSKSL</td>
<td>60(2)</td>
</tr>
<tr>
<td>Δ-3,5-Δ-2,4-Dienoyl-CoA isomerase (ECH1)</td>
<td>YKTVTFSL</td>
<td>120(32)</td>
</tr>
<tr>
<td>trans-2-enoyl-CoA reductase (PECR)</td>
<td>YTFLKEAKL</td>
<td>21(3)</td>
</tr>
<tr>
<td>2-Hydroxyphytanoyl-CoA lyase (HPCL2)</td>
<td>YHMLTRNML</td>
<td>4467(1289)</td>
</tr>
<tr>
<td>Methylacyl-CoA racemase (AMACR)</td>
<td>YSNVKQASL</td>
<td>1023(295)</td>
</tr>
<tr>
<td>Carnitine acetyltransferase 1 (CRAT)</td>
<td>YQSHPRAKL</td>
<td>59(7)</td>
</tr>
<tr>
<td>Carnitine octanoyl transferase (CROT)</td>
<td>YQLMNTHSL</td>
<td>590(46)</td>
</tr>
<tr>
<td>Acyl-CoA thioesterase 2 (PTE1)</td>
<td>YPQVSESKL</td>
<td>316(129)</td>
</tr>
<tr>
<td>Acyl-CoA thioesterase 1B (ZAP128)</td>
<td>YEIGTPSKVL</td>
<td>646(51)</td>
</tr>
<tr>
<td>Catalase (CAT)</td>
<td>YAAREKANL</td>
<td>1230(314)</td>
</tr>
<tr>
<td>Peroxisidoxin (PRDX5)</td>
<td>YAPNIISSL</td>
<td>128(60)</td>
</tr>
<tr>
<td>D-amino acid oxidase (DAO)</td>
<td>YSRMPPSHL</td>
<td>372(182)</td>
</tr>
<tr>
<td>D-aspartate oxidase (DDO)</td>
<td>YTPIPSNSL</td>
<td>148(46)</td>
</tr>
<tr>
<td>Hydroxyacid oxidase 1 (HAO1)</td>
<td>YNLVQFSRL</td>
<td>93(23)</td>
</tr>
<tr>
<td>Hydroxyacid oxidase 2 (HAO2)</td>
<td>YISAKTLKL</td>
<td>463(58)</td>
</tr>
<tr>
<td>Very-long-chain acyl-CoA synthetase (SLC27A2)</td>
<td>YGKPAKAKL</td>
<td>288(85)</td>
</tr>
<tr>
<td>Dihydroxyacetone phosphate acyltransferase (GPNAT)</td>
<td>YIKLAAQKL</td>
<td>186(56)</td>
</tr>
<tr>
<td>Isocitrate dehydrogenase (IDH1)</td>
<td>YIHCPPKKL</td>
<td>13490(2340)</td>
</tr>
<tr>
<td>Alanine-glycolate aminotransferase (AGXT)</td>
<td>YPSLGAHL</td>
<td>240(52)</td>
</tr>
<tr>
<td>Pipeolic acid oxidase (PIPOX)</td>
<td>YNPPVVSIL</td>
<td>33(7)</td>
</tr>
<tr>
<td>Epoxide hydrolase 2 (EPHX2)</td>
<td>YPPAVNARL</td>
<td>5012(515)</td>
</tr>
<tr>
<td>Glutathione S-transferase class κ (GSTK1)</td>
<td>YVQQPRPRL</td>
<td>2.6(0.4)</td>
</tr>
<tr>
<td>Polyamine oxidase (PAOX)</td>
<td>YQFOKNSKL</td>
<td>562(59)</td>
</tr>
<tr>
<td>Malonyl-CoA decarboxylase (MLYCD)</td>
<td>YVAAQTCKL</td>
<td>1995(292)</td>
</tr>
<tr>
<td>3-Hydroxy-3-methylglutaryl-CoA lyase (HMGLC)</td>
<td>YHKATSRSL</td>
<td>79(11)</td>
</tr>
<tr>
<td>Nudix hydrolase specific for CoA (NUDT7)</td>
<td>YINFMASAL</td>
<td>234(62)</td>
</tr>
<tr>
<td>Insulysin (IDE)</td>
<td>YPPGLNSKL</td>
<td>22(3)</td>
</tr>
<tr>
<td>Lon protease (LONP)</td>
<td>YSVVSSKH</td>
<td>162(23)</td>
</tr>
</tbody>
</table>

Debdeep Ghosh, PhD
Trypanosomes contain a novel organelle termed the glycosome. A peroxisome variant housing the enzymes of the glycolytic pathway. Trypanosomal Pex5p is a potential drug target to kill trypanosomes in the bloodstream of infected individuals.
Peptide Binding by Human vs Trypansomal Pex5p
"You can observe a lot just by watching"

Lawrence Peter "Yogi" Berra