A Structure of a Drug Target can initiate and accelerate drug development in many important ways:

I. The Structure of the Target by itself shows immediate novel opportunities for drug design
   e.g. The hexameric arrangement of helices in HIV gp41

II. A Structure of a Target with a Substrate or Co-factor or TS Analog reveals which pockets can be filled by inhibitors and suggests which types of compounds to make
   e.g. HIV protease substrate complex
       Protozoan GAPDH:NAD complex
       Influenza Virus Neuraminidase Inhibitors

III. Structures of the Target with Low MW-low affinity “fragments” show where fragments bind and how to modify and/or link fragments – to achieve higher affinity
     e.g. “Fragment Cocktail crystallography”

IV. The structure of a compound found in a screen in complex with the Target reveals how the compound acts and how it can be modified for better affinity
    e.g. NNRTI’s and HIV Reverse Transcriptase
        Cyclosporin in complex with Calcineurin and Cyclophilin

V. Structures of successive compounds bound to the same Target assist in understanding structure-activity relationships, binding modes and conformational changes: ITERATIVE STRUCTURE-BASED LEAD OPTIMIZATION.
   e.g. Anti-Glaucoma drug targeting carbonic anhydrase

VI. The structure of a Drug Candidate in complex with the Target can be helpful in devising strategies for modifications which MAINTAIN AFFINITY but improve e.g. drug bioavailability or decrease drug toxicity.

VII. The structure of a Drug:Target complex unravels the reasons for DRUG RESISTANCE
     e.g. Gleevec and abl-src kinase
High Affinity for Drug Target

Low Affinity for Homologues of Drug Target
Selective Inhibition
Often, if not always, CRUCIAL

Simplified View of Structure-based Drug Design

PROTEIN STRUCTURE BASED DRUG DESIGN CYCLE

Experimental Functional Genomics ➔ Target Protein Identification ➔ Computational Functional Genomics

Target Protein Overexpression ➔ Three Dimensional Structures of Target Proteins and of Target Protein Drug Complexes

Quantum Chemistry
Molecular Dynamics
Monte Carlo Simulation

Molecular Modeling
Docking
Growing
Linking

3D Databases of Chemicals
Chemical Intuition

Data Base of Target Protein Drug Complexes

Biomolecular Synthesis

Further Development

Functional Studies

Combinatorial Chemistry

X-ray, NMR
Electron Microscopy

Combinatorial Biochemistry
Drug Design
A case study

Structure-Based Inhibitor Design
of
the enzyme
GAPDH
from
the sleeping sickness parasite,
a “Trypanosomatid”

Sleeping Sickness
a.k.a “African Trypanosomiasis”

Blood stream
form of parasite

Lumbar puncture
for diagnosis of parasites in CNS

Tsetse fly

Sleeping sickness is caused by a unicellular eukaryote: Trypanosoma brucei – a “Trypanosomatid”
Other pathogenic trypanosomatids are whole set of Leishmania species.
These cause a spectrum of different tropical diseases, called “leishmaniasis”.
Many enzymes in Trypanosoma brucei and Leishmania species are very similar in amino acid sequence.

With thanks to Wes Van Voorhis
Glycolytic enzymes are critical for the blood stream form of *Trypanosoma brucei*  
(ONLY in this group of parasites most of the glycolytic enzymes are sequestered in a unique organelle: the glycosome)

---

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from the Sleeping Sickness Parasite and the human host

Note the difference in conformation near the ribose of the NAD cofactor in the homologous proteins of host and parasite.
Adenosine is part of the cofactor (co-substrate) NAD of the enzyme GAPDH.

- It is by itself a poor inhibitor of mammalian and *T. brucei* parasite GAPDH.
- Moreover, it inhibits the sleeping sickness parasite enzyme slightly worse than the mammalian enzyme.

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

**Sleeping sickness parasite** GAPDH: Hydrophobic Groove near 2’OH of Adenosine
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

**Human GAPDH**: NO groove near 2’OH of Adenosine

**Sleeping Sickness parasite GAPDH**: Substituent Modeled in Hydrophobic Groove near 2’OH of Adenosine

Previous position of the 2’OH
Selectivity changes of 2'-OH substituted compound versus adenosine

Selectivity of Structure-based Designed GAPDH Inhibitors

Note: Leishmania mexicana GAPDH is ~77% sequence identical to Trypanosoma brucei GAPDH and all residues in the region of interest are identical in these two pathogenic "Trypanosomatids". So these two enzymes are used interchangeably.

Exploring additional hydrophobic grooves near the adenosine binding pocket of Leishmania mexicana GAPDH

Structure leads to “Targeted Combinatorial Chemistry” to fill the grooves optimally

*Note: Leishmania mexicana GAPDH is ~77% sequence identical to Trypanosoma brucei GAPDH and all residues in the region of interest are identical in these two pathogenic “Trypanosomatids”. So these two enzymes are used interchangeably.
Inhibition of *L. mexicana* GAPDH by Adenosine Derivatives

**Principle:**
Make a diverse set of inhibitors by different substituents at three positions R¹, R² and R³ of a so-called “scaffold molecule” (shown above).

<table>
<thead>
<tr>
<th>Compound</th>
<th>R¹</th>
<th>R²</th>
<th>R³</th>
<th>IC₅₀ (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>35</td>
<td>NH₂</td>
<td>OH</td>
<td></td>
<td>250</td>
</tr>
<tr>
<td>36</td>
<td>NH₂</td>
<td>OH</td>
<td></td>
<td>250</td>
</tr>
<tr>
<td>41</td>
<td></td>
<td>OH</td>
<td></td>
<td>inactive</td>
</tr>
<tr>
<td>42</td>
<td></td>
<td>OH</td>
<td></td>
<td>inactive</td>
</tr>
<tr>
<td>48</td>
<td>NH₂</td>
<td></td>
<td></td>
<td>100</td>
</tr>
<tr>
<td>49</td>
<td>NH₂</td>
<td></td>
<td></td>
<td>60</td>
</tr>
</tbody>
</table>

*inactive = inactive at 50 μM.*

Crystal structure of *L. mexicana* GAPDH with “NMDBA”

Clearly visible is the selectivity cleft between Met39 and Val206* (from the neighboring monomer), with the dimethoxybenzamido group of NMDBA inserted into it.

The surface has been color coded according to the electrostatic potential. Red represents negative potential and blue positive potential.

“NMDBA”: A new inhibitor with 10⁵-fold (!) affinity gain compared to the initial inhibitor adenosine
Flexibility in the structure of *L. mexicana* GAPDH

The figure illustrates the displacements of the protein atoms at the inhibitor binding site. In particular, the movement of Met39 effects expansion of the selectivity cleft, and this motion propagates to the other atoms involved in inhibitor binding.

Adaptation of the protein to a ligand is a very common, yet still an often surprising, event.
Influenza Virus Neuraminidase Inhibitors

A classic example of Structure-Based Drug Design (SBDD) on the basis of a Enzyme-Transition State Analog Complex & affinity gain by increasing electrostatic interactions

**Influenza Virus**

Well underfocused electron micrograph of an unstained, frozen, hydrated specimen of A/X31 influenza virus.

Influenza Virus has two surface proteins: haemagglutinin (H) and neuraminidase (N).

Flu Virus is clearly "pleiotropic", i.e. occurs in different shapes.

The bar in the lower left represents 1000 Å.
Influenza Virus Neuraminidase Inhibitors

The enzyme neuraminidase plays a key role in the release of new viruses from the host cell surface. Inhibition of neuraminidase activity appeared to be a good way to decrease the severity of a flu infection.

Enzyme often catalyze reactions by preferential binding of the transition state vs the ground state.

Transition state analogs are potent enzyme inhibitors.

\[ W_n = \text{n-th 4-\(\beta\)-stranded “propeller”} \]

SMITH et al, PROTEIN SCI. 10: 689 (2001) – PDB-code 1F8D.

Modified from Carolyn R. Bertozzi - website: http://grtc.ucsd.edu/lecture42.pdf
The substrate of neuraminidase

N-Acetyleneuraminic acid (pyranose form)

Sialic Acid \equiv N\text{-acetyleneuraminic Acid} \ (VVP \ 2^{nd} \ Ed. \ p \ 213)

(A complex sugar, attached to quite a few human cell surface proteins)

The Reaction catalyzed by neuraminidase

Transition State:
With the Ring-O, the C-2, the Carboxyl carbon and the C3 approximately in one plane.

Sialic acid = N-ACETYL-NEURAMINIC ACID \ (VVP \ 2^{nd} \ Ed. \ P. \ 213)

Modified from Carolyn R. Bertozzi - website: http://grtc.ucsd.edu/lecture42.pdf
Design of Transition State Analog neuraminidase inhibitors

![Chemical Structure]

**DANA**

**Features:**
- Planarity at C-2
- Buildup of + charge

**Transition State Analog**

**DANA ≡**

2-DEOXY-2,3-DEHYDRO-N-ACETYL-NEURAMINIC ACID

---

**The Starting Point**

*The Transition State Analog (TSA) DANA*

![Chemical Structure]

DANA
9-amino-DANA =
9-AMINO-2-DEOXY-2,3-DEHYDRO-N-ACETYL-NEURAMINIC ACID
(The 9-amino group is irrelevant for the drug development story)

9-amino-DANA sits clearly in a pocket.
This is the active site of neuraminidase
Influenza Virus Neuraminidase in complex with 9-amino-DANA

TWO NEGATIVELY CHARGED CARBOXYLATES ARE QUITE CLOSE TO THE 4-OH !!!!!
Compound made: 4-guanidino-DANA
A guanidino substituent at the 4-position instead of a hydroxyl

Does it indeed live up to the expectations?
I.e. of being a better inhibitor than DANA?


Inhibitory Properties of modified 4-guanidino-DANA

Based on the structure of the TSA DANA in complex with influenza virus neuraminidase, the compound 4-guanidino-DANA was designed and synthesized.

The $K_i$-values (in M) were as follows:

<table>
<thead>
<tr>
<th></th>
<th>Flu Neura</th>
<th>Human Neura</th>
</tr>
</thead>
<tbody>
<tr>
<td>DANA</td>
<td>$1 \times 10^{-6}$</td>
<td>$1.2 \times 10^{-6}$</td>
</tr>
<tr>
<td>4-guanidino-DANA</td>
<td>$2 \times 10^{-10}$</td>
<td>$1 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

By changing one single functional group:
The affinity for the target flu enzyme was enhanced by a factor of $\sim$10,000;
The affinity for the human homologous enzyme was decreased by a factor of $\sim$100.
The selectivity was improved by a factor $\sim$1,000,000!!!

Properties of 4-guanidino-DANA
Zanamivir (Relenza)

This compound is obviously very hydrophilic:
One guanidinium group & One carboxylate & Three hydroxyls & One –NH-C=O group!

Therefore this medicine is not active when given orally.

However, influenza virus enters host lung cells, so the compound can be administered with an inhalator.

Physical Chemical Requirements of (most) Oral Drugs
The Lipinski "Rules of Five"

"From the 50,427 compounds in the WDI (World Drug Index) File 2245 were selected which are likely to have superior physico-chemical properties.

Poor absorption or permeation are more likely when:
- The MWT is over 500
- There are more than 5 H-bond donors
- There are more than 10 H-bond acceptors
- The Log P is over 5

... orally active therapeutic classes outside the ‘rule of 5’ are: antibiotics, antifungals, vitamins and cardiac glycosides.

...We suggest that these few therapeutic classes contain orally active drugs that violate the ‘rule of 5’ because members of these classes have structural features that allow the drugs to act as substrates for naturally occurring transporters."

Lipinski et al., Advanced Drug Delivery Reviews 46: 3-26 (2001)
For orally available medicines a fine balance is required between:

(i) Sufficient capacity to cross membranes, 
    so it can be taken up from the digestive tract;
(ii) Sufficient water solubility, 
    so it can reach the site of action in sufficient concentrations.

Some other requirements of an ideal medicine are:
(iii) Not being converted to an inactive substance by human enzymes;
(iv) Not being cleared rapidly from the blood;
(v) No teratogenicity;
(vi) No mutagenicity;
(vii) No toxicity;
(viii) And more...

Hence, it is not really a surprise that it is a major challenge to make a new safe, effective, orally available, affordable medicine.

Multivalent Inhibitors of Cholera Toxin (CT)
produced by *Vibrio cholerae*.

CT is a close relative of Heat-Labile Enterotoxin (LT)
produced by enterotoxigenic *E. coli*,
the cause of much of children’s and traveller’s diarrhea
Vibrio cholerae:
- produces CT
- ~4000 victims per year

Enterotoxigenic E. coli:
- produces LT and ST
- ~480,000 victims per year

Heat-labile Enterotoxin (LT),
a very close relative of Cholera toxin (CT)

A subunit

B pentamer
CT and LT Receptor Binding

Intestinal epithelial cell

Ganglioside $G_{M1}$

CT : Cholera

LT : Traveller's & Children's diarrhea

CT and LT Receptor Binding

Intestinal epithelial cell

Ganglioside $G_{M1}$

CT : Cholera

LT : Traveller's & Children's diarrhea
Cholera toxin – $G_{M1}$ Receptor Interaction

A subunit

B pentamer

$G_{M1}$-OS Intestinal cell surface

$IC_{50} = 14 \times 10^{-9} \text{ M}$

Extensive hydrophobic and H-bonding interactions

Joseph Martial

Steve Sarfaty
Ethan Merritt
Five receptor binding sites

Making ligands longer
Ligand-Protein Complex

Pentavalent Ligand
THE PENTAVALENT CONCEPT

Gains in surface-receptor binding inhibition

Erkang Fan and co-workers
**Gains in surface-receptor binding inhibition**

<table>
<thead>
<tr>
<th>One-Unit Linker</th>
<th>Two-Unit Linker</th>
<th>Three-Unit Linker</th>
<th>Four-Unit Linker</th>
</tr>
</thead>
<tbody>
<tr>
<td>240 x Single Finger</td>
<td>3,600 x Single Finger</td>
<td>10,000 x Single Finger</td>
<td>104,000 x Single Finger</td>
</tr>
</tbody>
</table>

Erkang Fan and co-workers

---

**IC$_{50}$ versus EXTENDED(●) & EFFECTIVE (●) DIMENSIONS OF PENTAVALENT LIGANDS**

- ● effective dimensions of ligands
- ● extended dimensions of ligands
- vertical line = distance between non-adjacent binding sites in LT

What if even longer linkers?

Erkang Fan and co-workers
And, indeed, linker too long: less affinity

(A single galactose "finger": IC$_{50} =$ ~ 100 mM = ~100,000 $\mu$M)

<table>
<thead>
<tr>
<th>Linker Units</th>
<th>IC$_{50}$ ($\mu$M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 2</td>
<td>13.26 ± 0.95</td>
</tr>
<tr>
<td>n = 4</td>
<td>1.50 ± 0.10</td>
</tr>
<tr>
<td>n = 6</td>
<td>4.63 ± 0.46</td>
</tr>
<tr>
<td>n = 8</td>
<td>7.25 ± 0.38</td>
</tr>
</tbody>
</table>

Erkang Fan and co-workers
Fragment Cocktail Crystallography

**Principle**

- Protein crystals
- Cocktails of chemical fragments

Protein crystals + Cocktails of chemical fragments → Protein crystal with bound chemical fragment

Probe protein pockets by soaking crystals in well-designed mixtures of 5-10 different chemicals, followed by crystal structure determinations. Followed by "growing" or "linking" the fragments to obtain higher affinity.
ACD Compound Filtering

ACD = Available Chemical Database

9,500 compounds

fragmentation

626 fragments

isolate ring systems

23 frameworks (at connectivity level)

90 compounds

manual selection of compounds

from each framework class

- eliminate mutagens, known poisons
- no highly functionalized compounds
- retain Br containing compounds

Nucleoside 2-deoxyribosyltransferase (Tbru015777AAA) plus Cocktail #4

1,2-DIHYDROBENZO[CD]INDOL-2-ONE

Jürgen Bosch & Christophe Verlinde & Erkang Fan & SGPP

Omni-present glycerol
**REFERENCES**

**Influenza Virus Neuraminidase**

**Cholera Toxin, Heat-labile Enterotoxin**
Computational Approaches

An excellent website with recent tools for Structure based drug design:
http://www.imb-jena.de/~rake/Bioinformatics_WEB/dd_tools.html

Major Journals with plenty SBDD:

J. Medicinal Chemistry
Chemistry and Biology
Nature Reviews Drug Discovery
J. Computer-Aided Molecular Design