

Structure-Based Drug Design

A few general principles and case studies

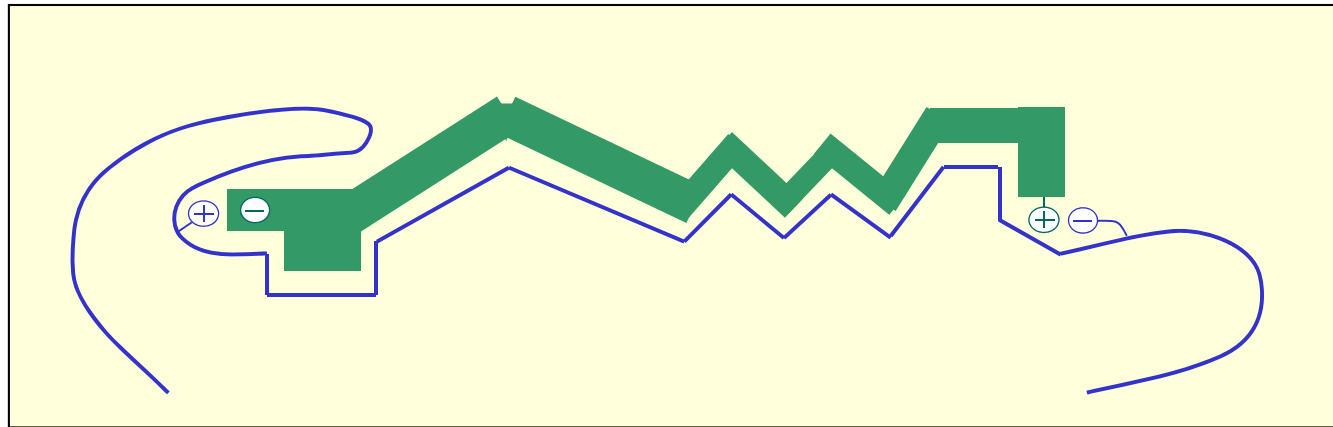
BC530

Fall Quarter 2014

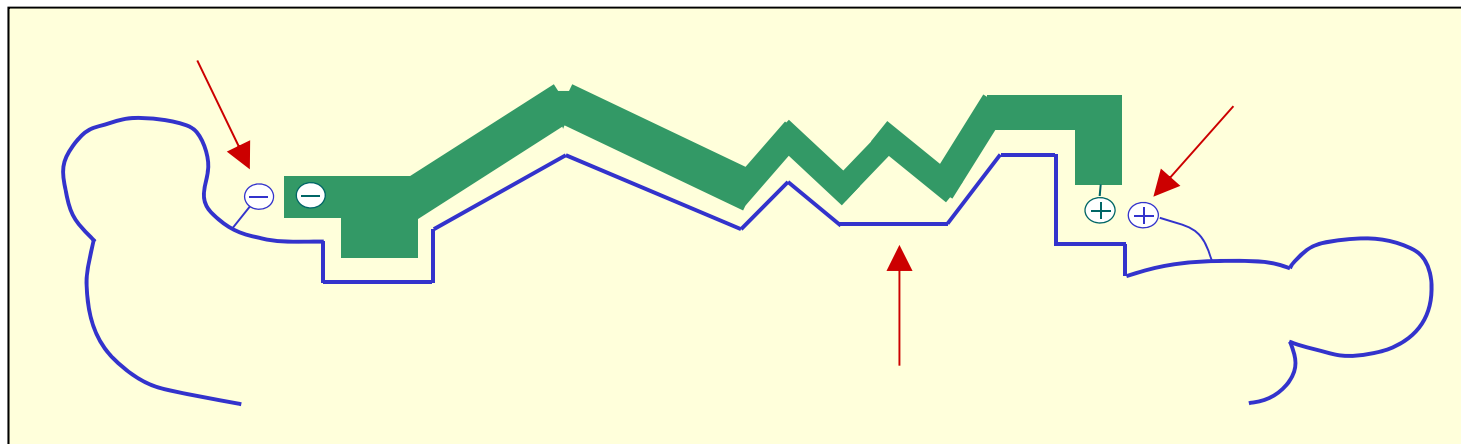
Wim G. J. Hol

<http://www.bmsc.washington.edu/WimHol/>

Simplified View of Structure-based Drug Design



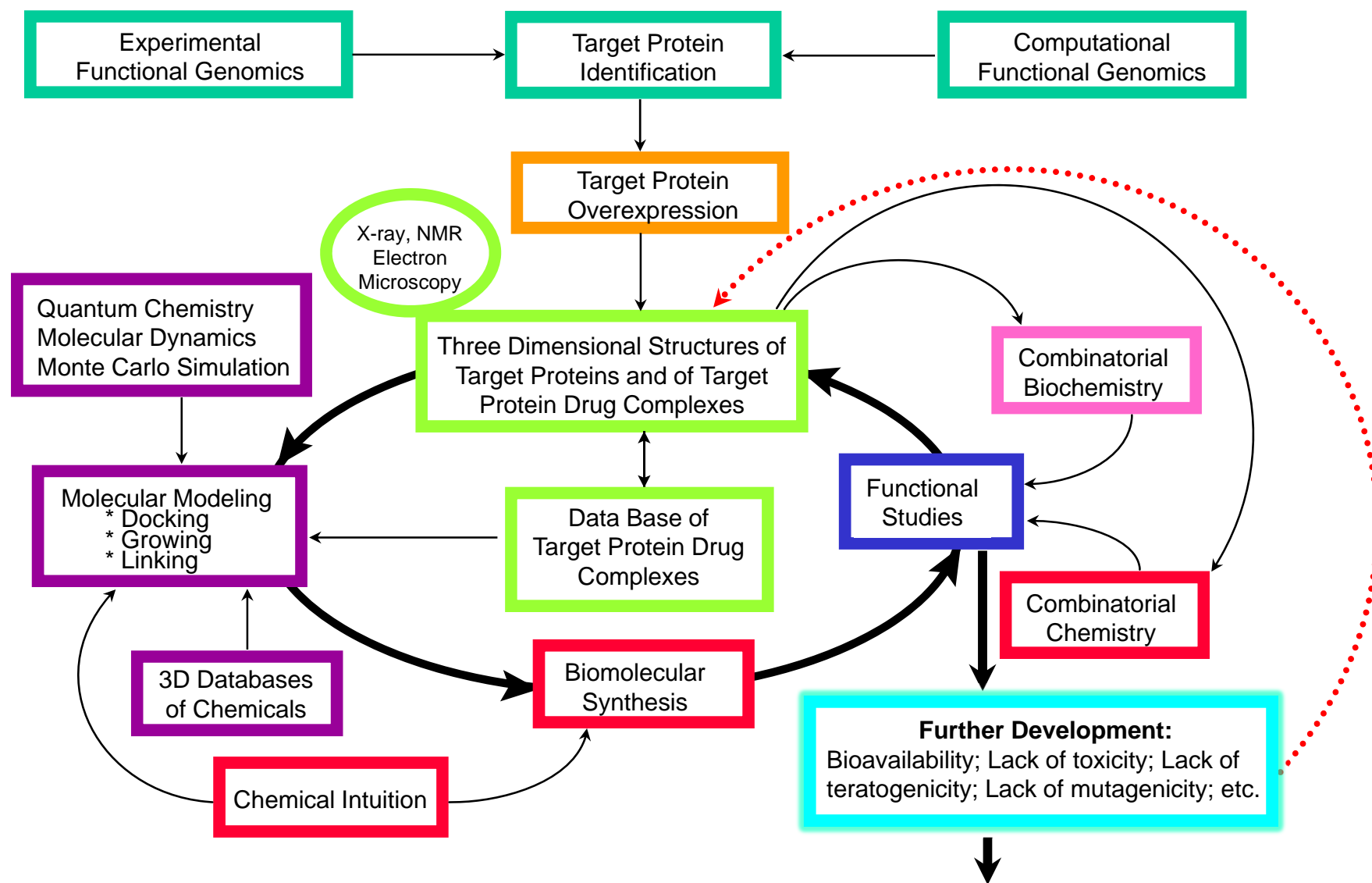
High Affinity for Drug Target



Low Affinity for Homologues of Drug Target

Selective Inhibition is often, but not always (!), CRUCIAL

STRUCTURE-GUIDED DRUG DESIGN



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”



Lumbar puncture
for diagnosis of parasites in CNS



Blood stream
form of parasite



Tsetse fly

Sleeping sickness is caused by a unicellular eukaryote: *Trypanosoma brucei* – a “Trypanosomatid”

Other pathogenic trypanosomatids are whole set of *Leishmania* species.

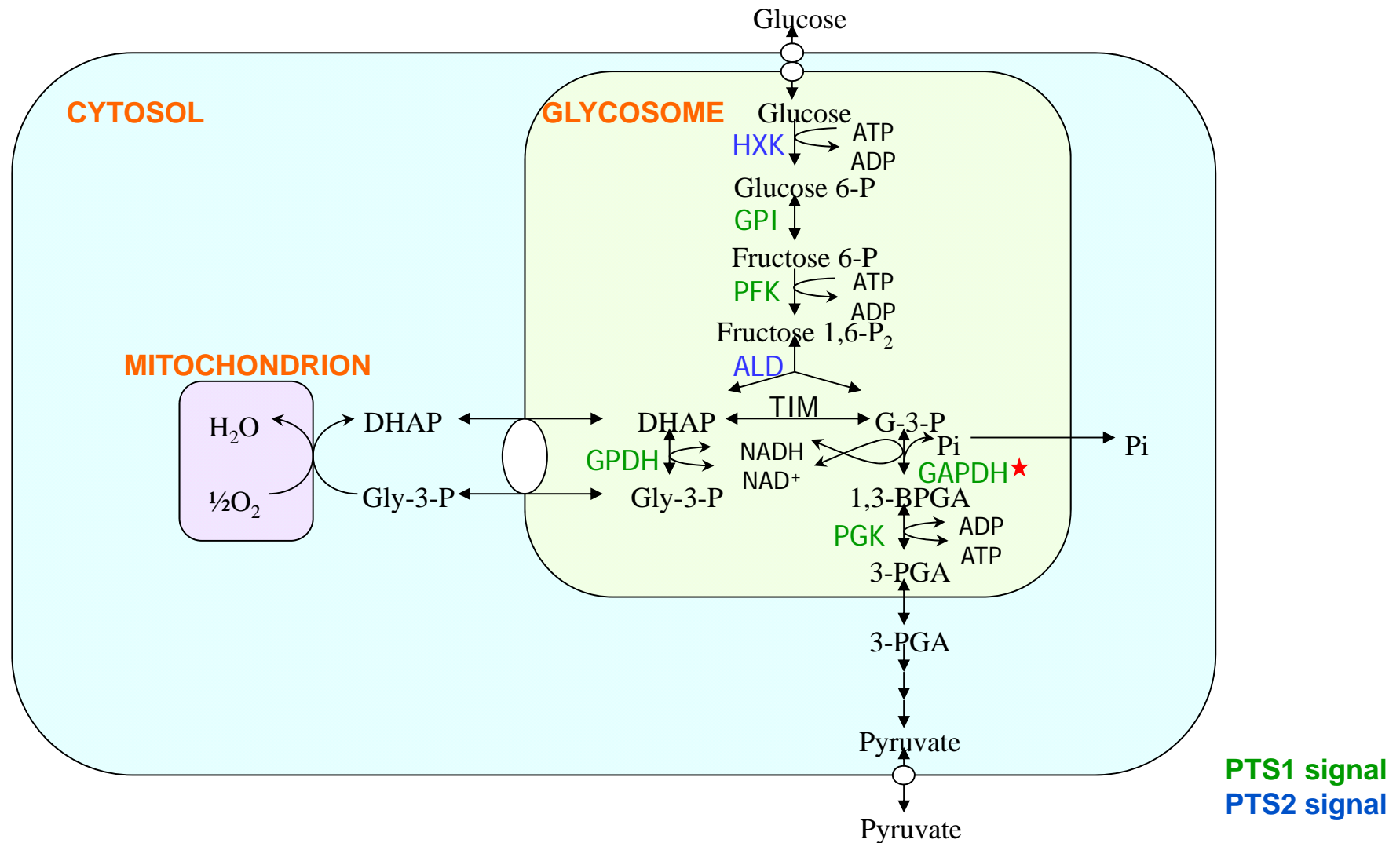
These cause a spectrum of different tropical diseases, called “leishmaniasis”.

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)

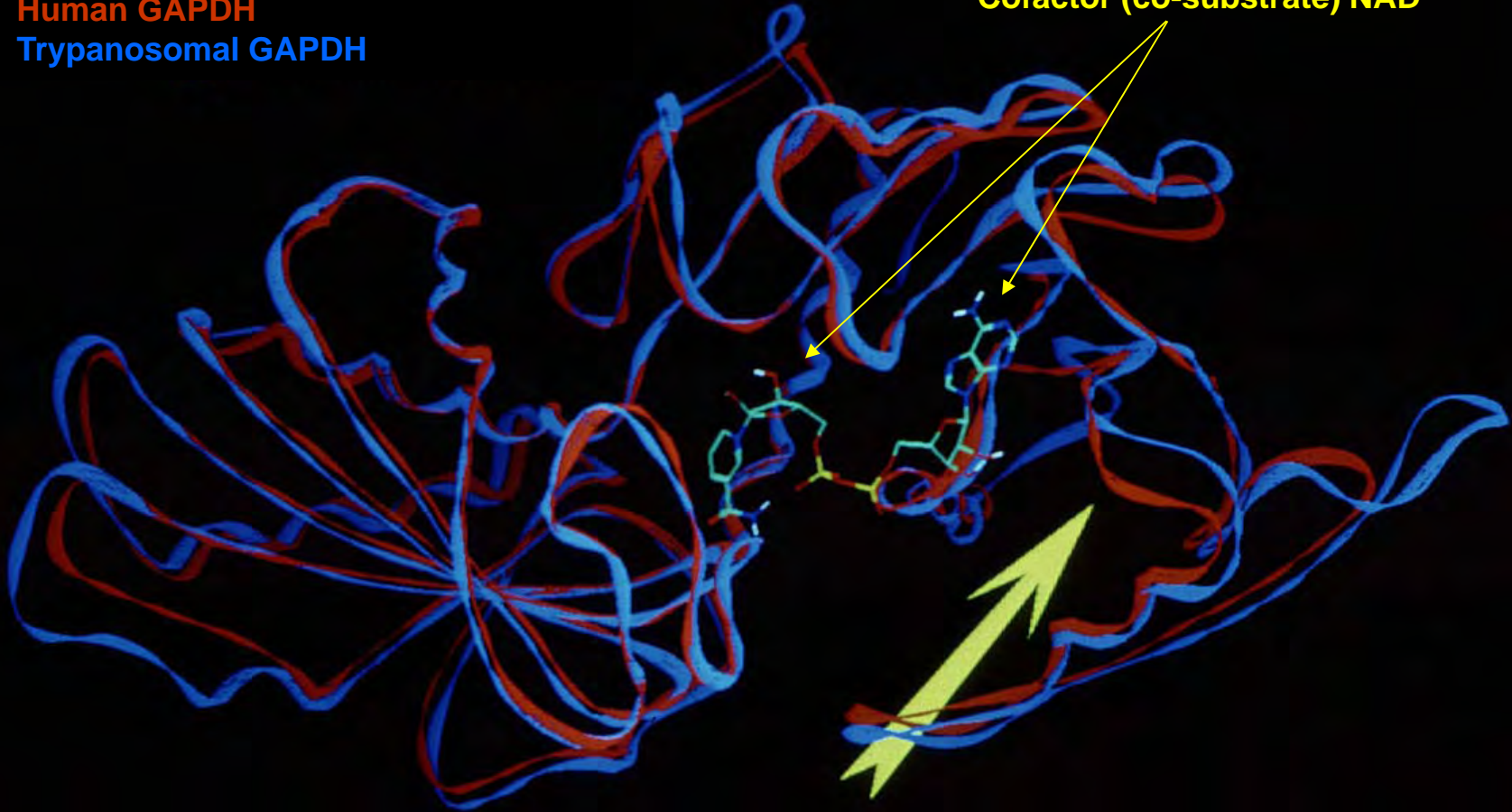


With thanks to Paul Michels and Fred Opperdoes

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

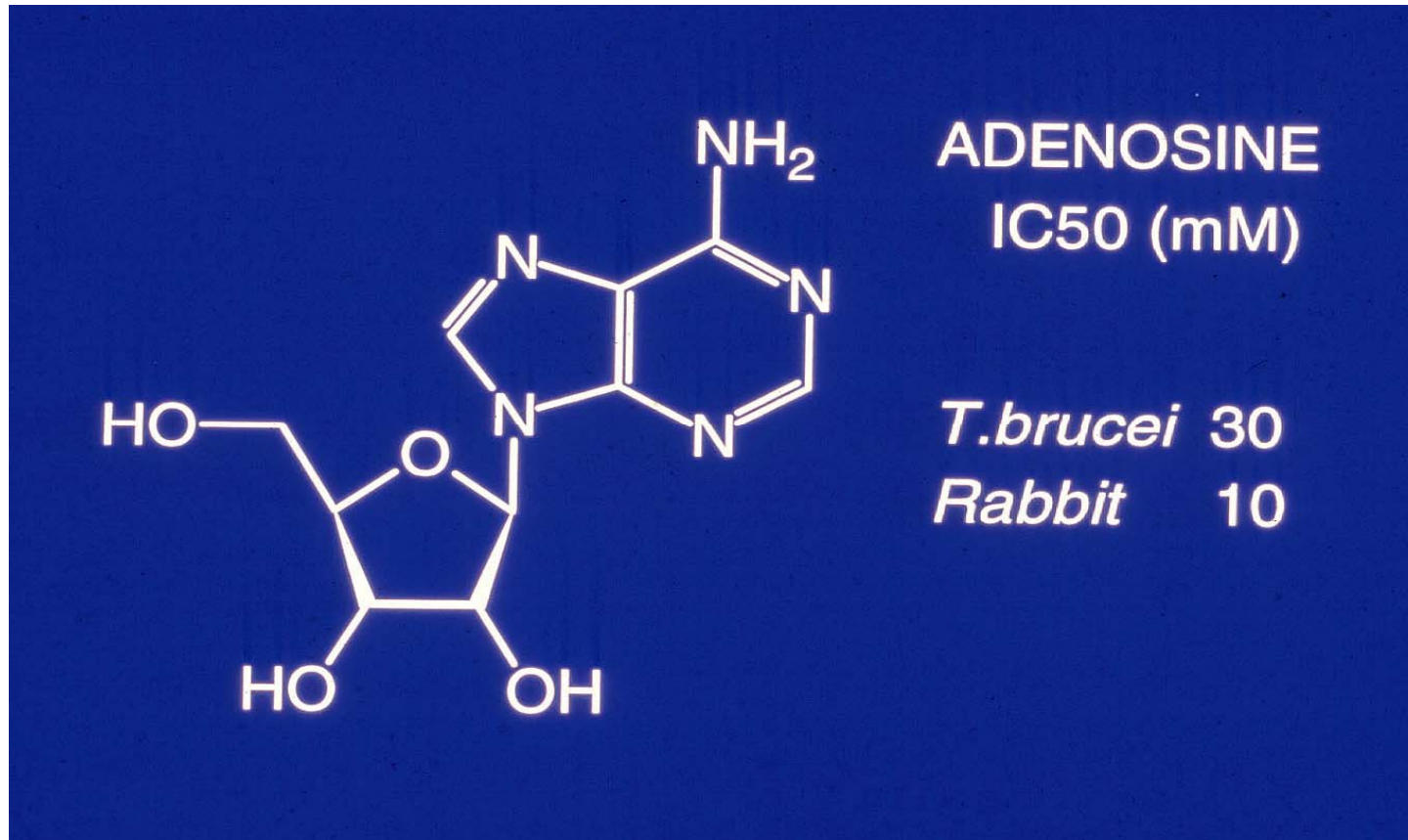
Human GAPDH
Trypanosomal GAPDH

Cofactor (co-substrate) NAD



Note the difference in conformation near the ribose of the NAD cofactor
in the homologous proteins of host and parasite.

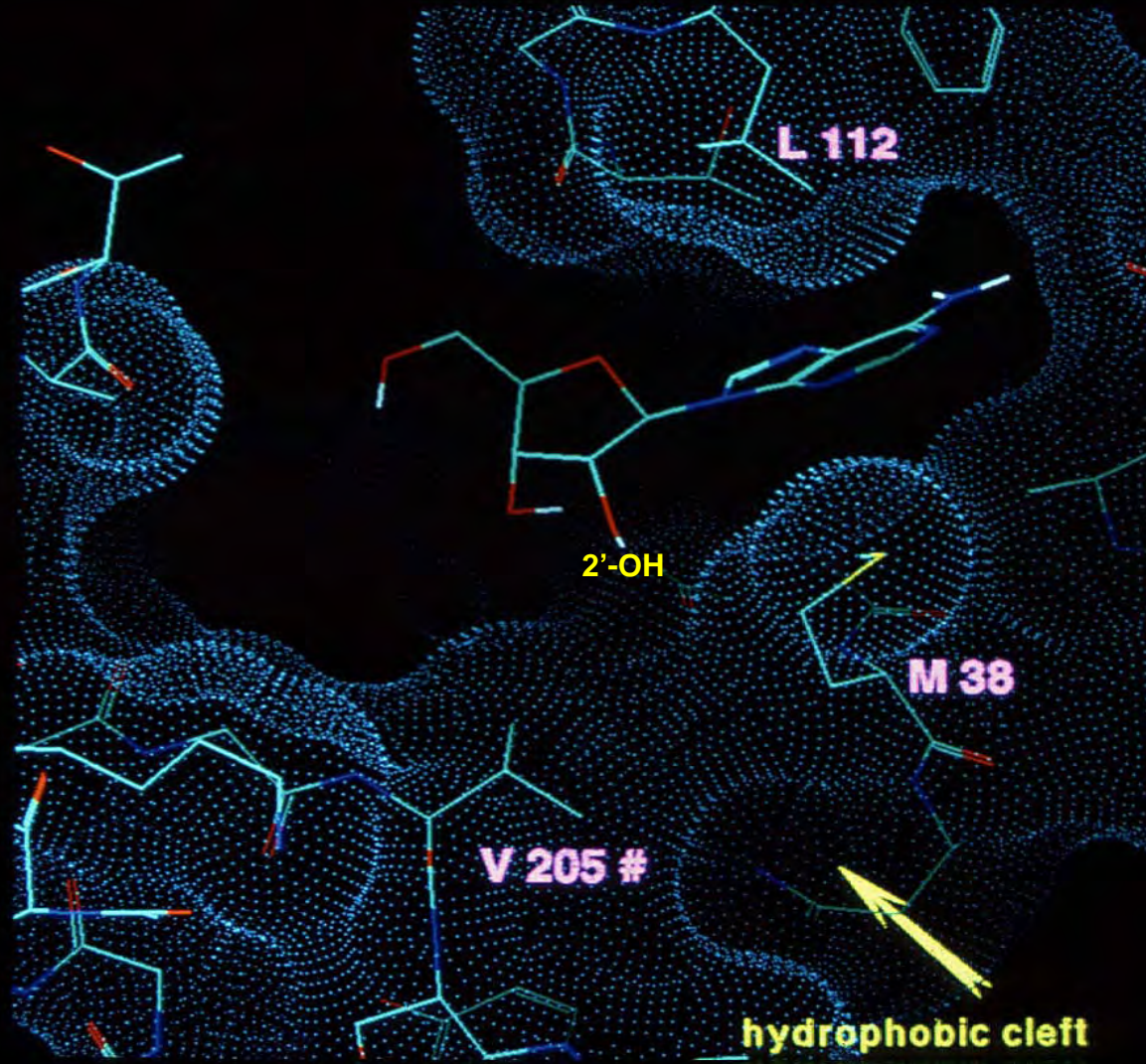
Adenosine – the starting point



- 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)

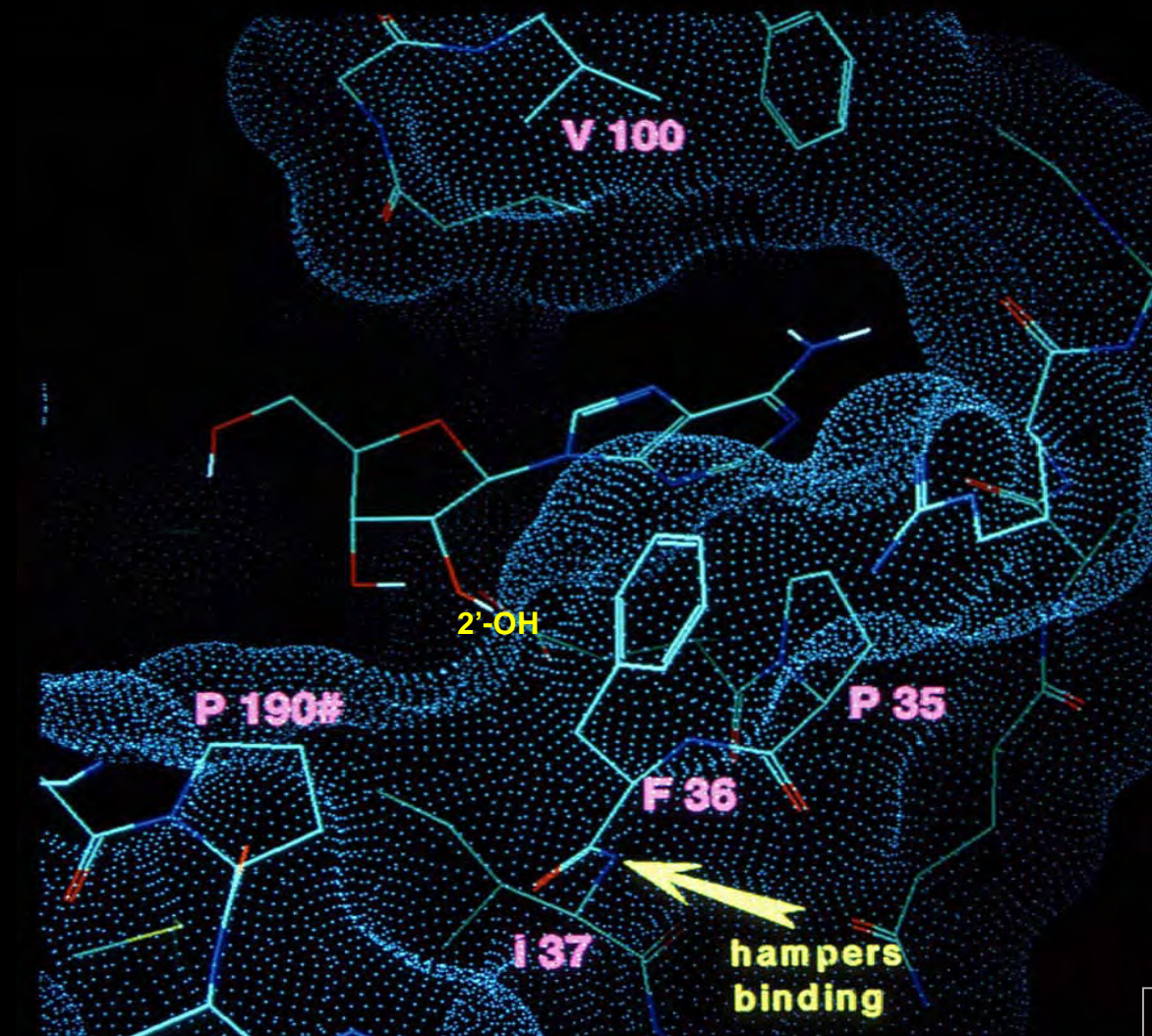
Sleeping sickness parasite GAPDH : Hydrophobic Groove near 2'-OH of Adenosine



Fred Vellieux
Christophe Verlinde

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

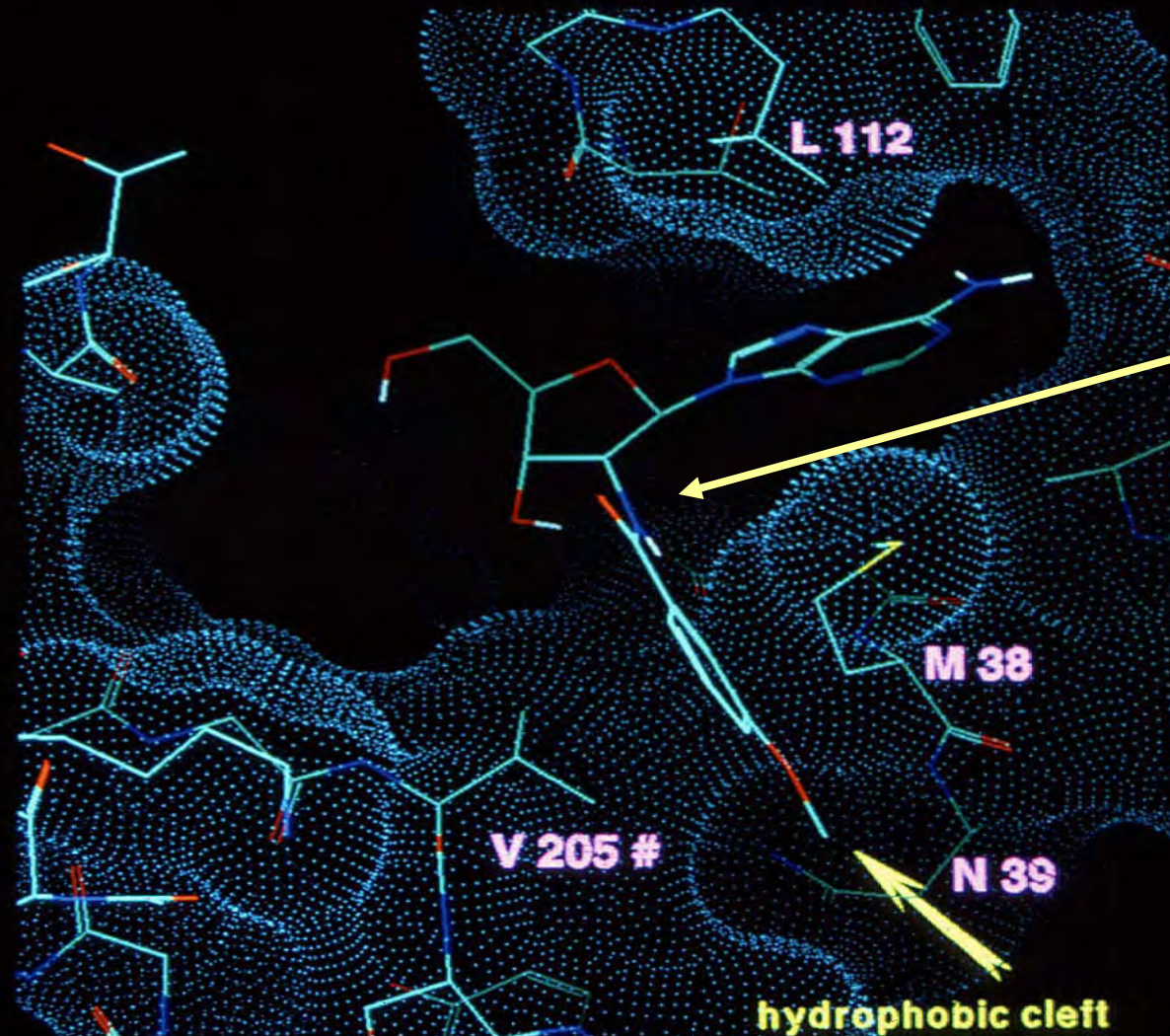
Human GAPDH : NO groove near 2'-OH of Adenosine



Randy Read
Christophe Verlinde

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)

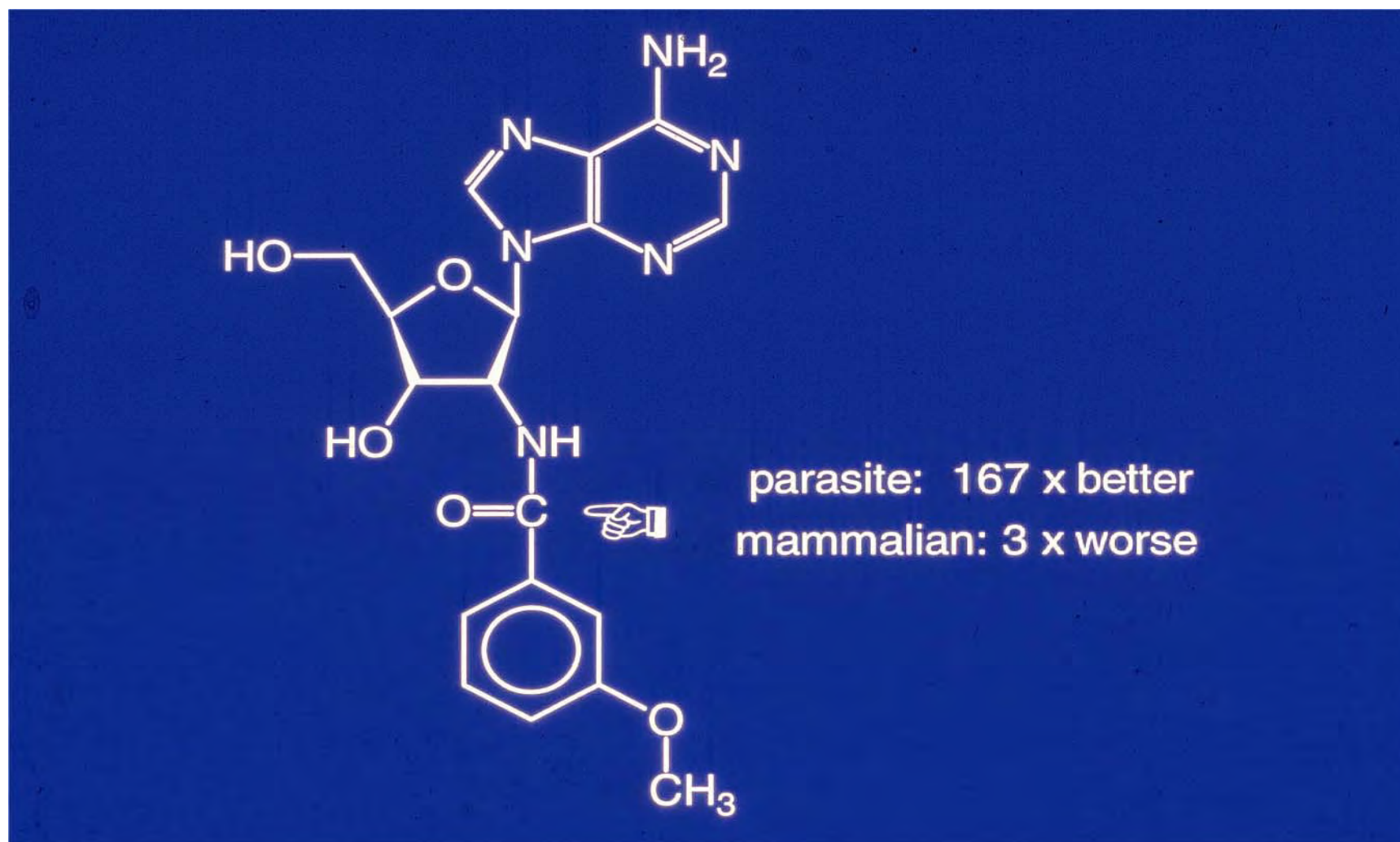
Sleeping Sickness parasite GAPDH : Substituent Modeled in Hydrophobic Groove near 2'OH of Adenosine



Previous
position of the
2'-OH

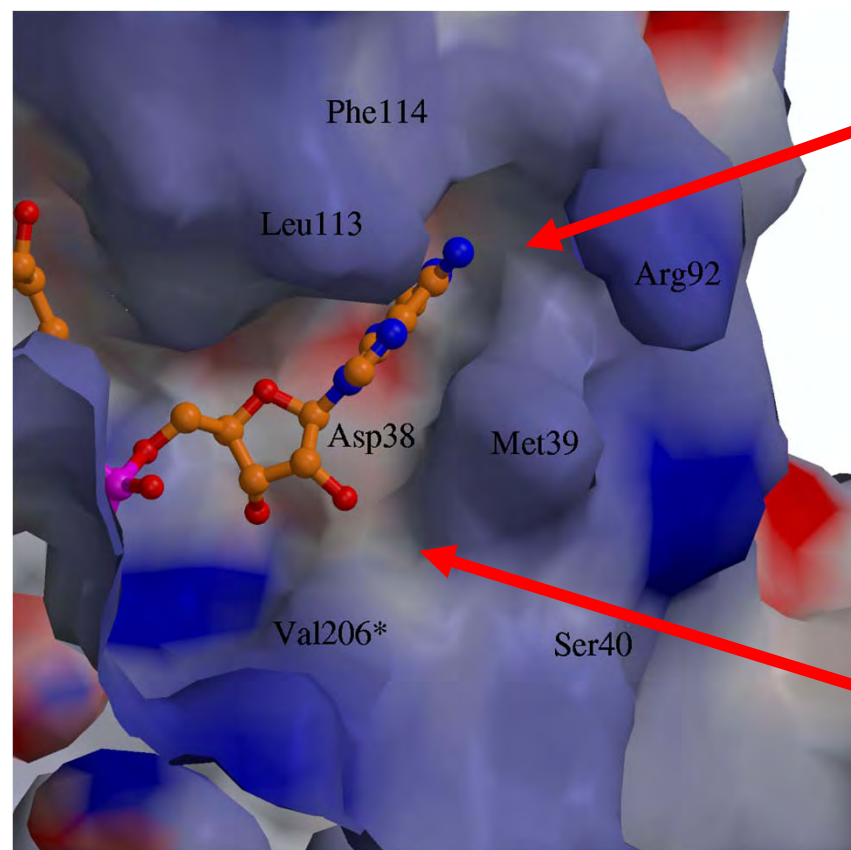
Christophe Verlinde

Selectivity of Structure-based Designed GAPDH Inhibitors



Selectivity changes of 2'-OH substituted compound *versus* adenosine

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



Surface of *L. mexicana** GAPDH
with NAD bound.

Hydrophobic Groove



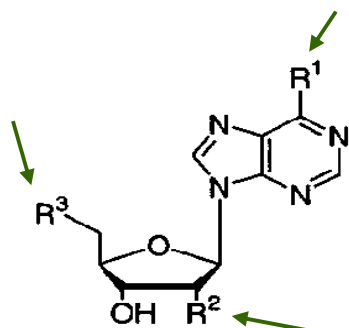
Structure leads to
“Focused Combinatorial Chemistry”
to fill the grooves optimally



Hydrophobic Groove

•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



35, 36, 41, 42, 48, and 49
These numbers are just "short-hand names"
for variant molecules with different
substituents.

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).

Using a "focused
combinatorial library"

Compound	R ¹	R ²	R ³	IC ₅₀ (μM)
35	NH ₂	OH		250
36	NH ₂	OH		250
41		OH		inactive
42		OH		inactive
48	NH ₂			100
49	NH ₂			60

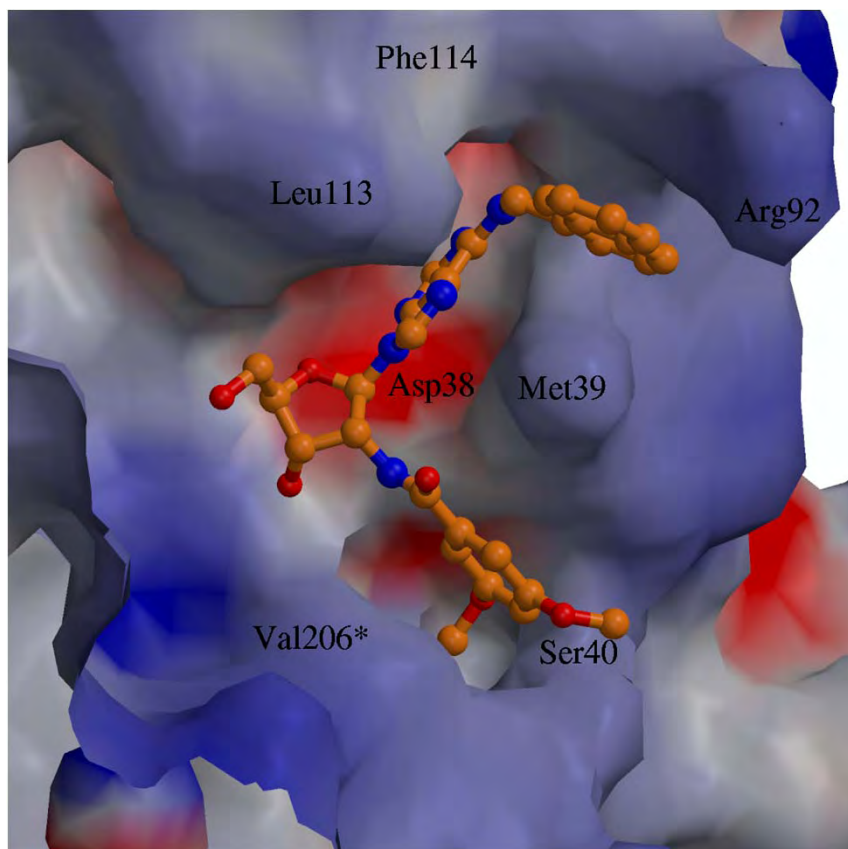
^a Inactive = inactive at 50 μM.

Inhibition of *L. mexicana* GAPDH by Adenosine Derivatives

Crystal structure of L. mexicana GAPDH with “NMDBA”

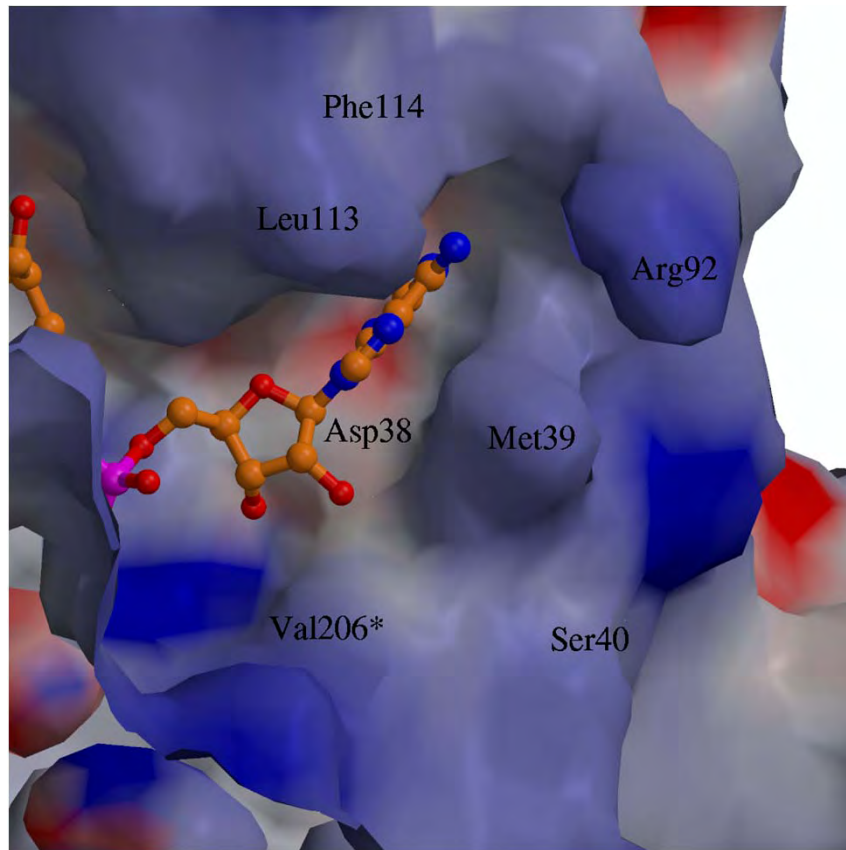
Clearly visible is the selectivity cleft between Met39 and Val206* (from the neighboring monomer), with the dimethoxybenzamido group of NMDBA inserted into it.

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

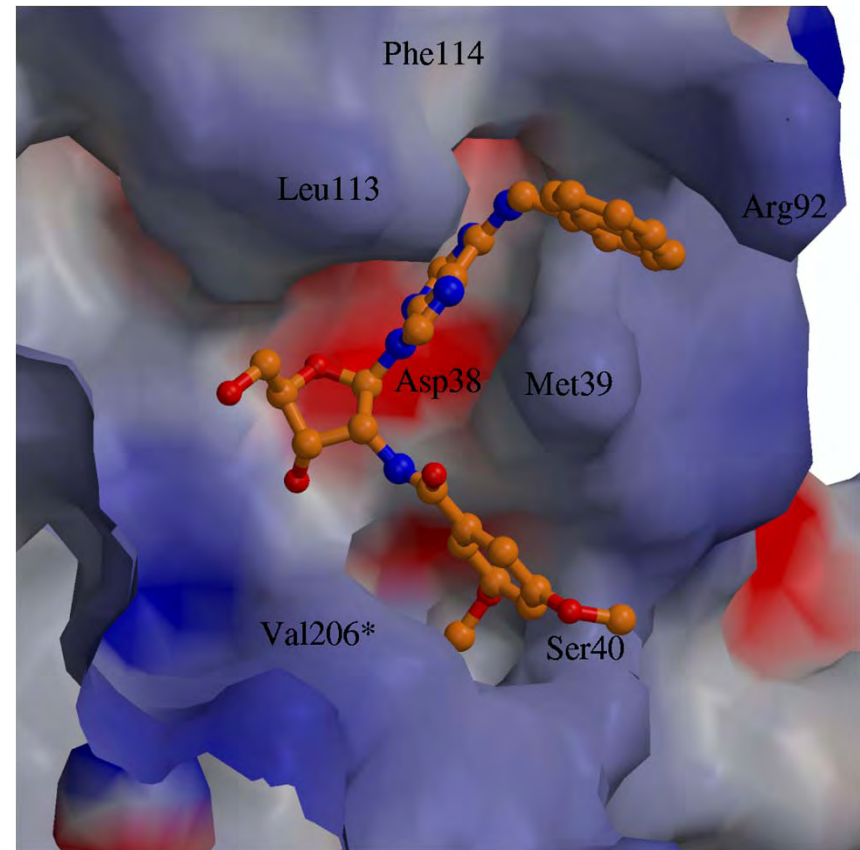


“NMDBA”: A new inhibitor with 10⁵-fold (!) affinity gain compared to the initial inhibitor adenosine

Flexibility in the structure of *L. mexicana* GAPDH

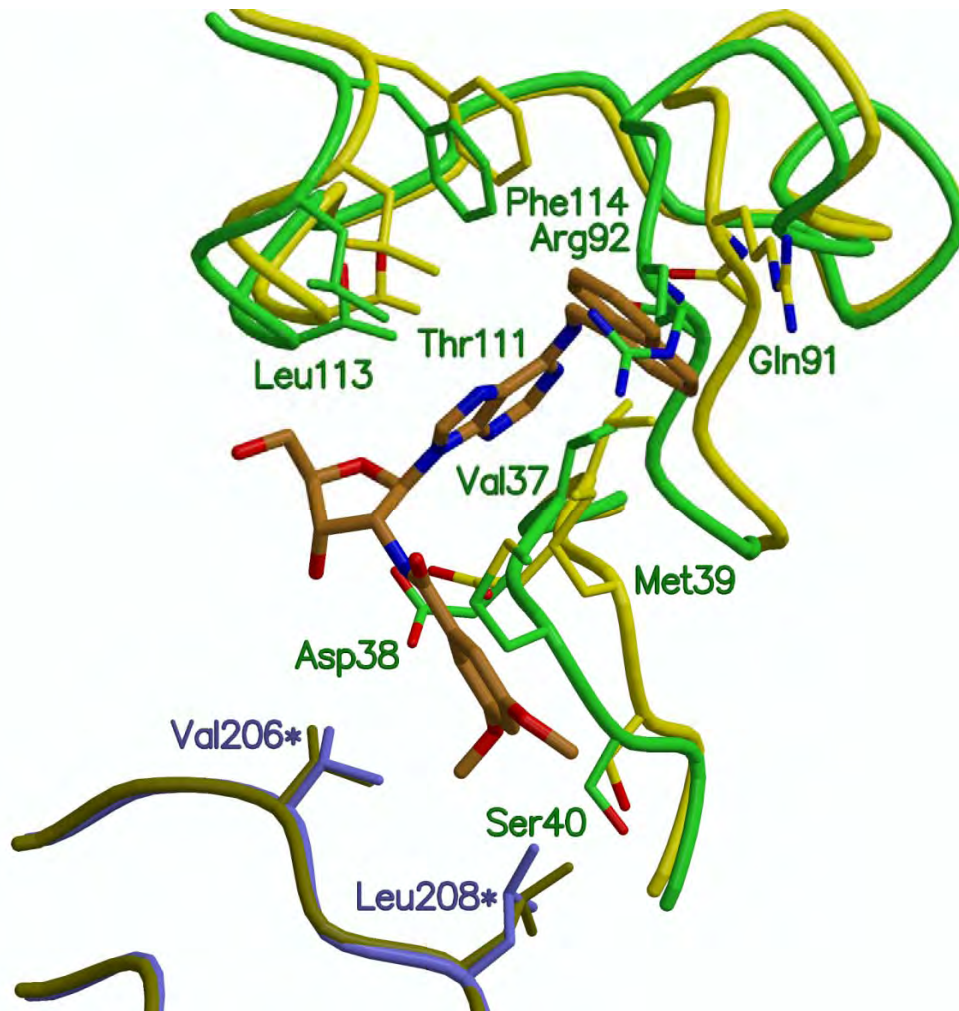


LmGAPDH + NAD



LmGAPDH + NMDBA

Flexibility in the structure of *L. mexicana* GAPDH



GAPDH in complex with NAD: green and violet
GAPDH in complex with TNDBA: yellow and gold
Only TNDBA shown

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.

REFERENCES

Inhibitor Design *Trypanosoma brucei* and *Leishmania* GAPDH

- Aronov, A. M., Verlinde, C. L. M. J., Hol, W. G. J. & Gelb, M. H. (1998). Selective tight binding inhibitors of trypanosomal glyceraldehyde-3-phosphate dehydrogenase via structure-based drug design. *J. Med. Chem.* **41**, 4790-4799.
- Aronov, A. M., Suresh, S., Buckner, F. S., van Voorhis, W. C., Verlinde, C. L. M. J., Hol, W. G. J. & Gelb, M. H. (1999). Structure-based design of sub-micromolar, biologically active inhibitors of trypanosomatid glyceraldehyde-3-phosphate dehydrogenase. *Proc. Natl. Acad. Sci. USA* **96**, 4273-4278.
- Suresh, S., Bressi, J. C., Kennedy, K. J., Verlinde, C. L. M. J., Gelb, M. H. & Hol, W. G. J. (2001). Conformational changes in *Leishmania mexicana* glyceraldehyde-3-phosphate dehydrogenase induced by designed inhibitors. *J. Mol. Biol.* **309**, 423-435.
- Bressi, J. C., Verlinde, C. L. M. J., Aronov, A. M., Shaw, M. L., Shin, S. S., Nguyen, L. N., Suresh, S., Buckner, F. S., Van Voorhis, W. C., Kuntz, I. D., Hol, W. G. J. & Gelb, M. H. (2001). Adenosine analogues as selective inhibitors of glyceraldehyde-3-phosphate dehydrogenase of Trypanosomatidae via structure-based drug design. *J. Med. Chem.* **44**, 2080-2093.

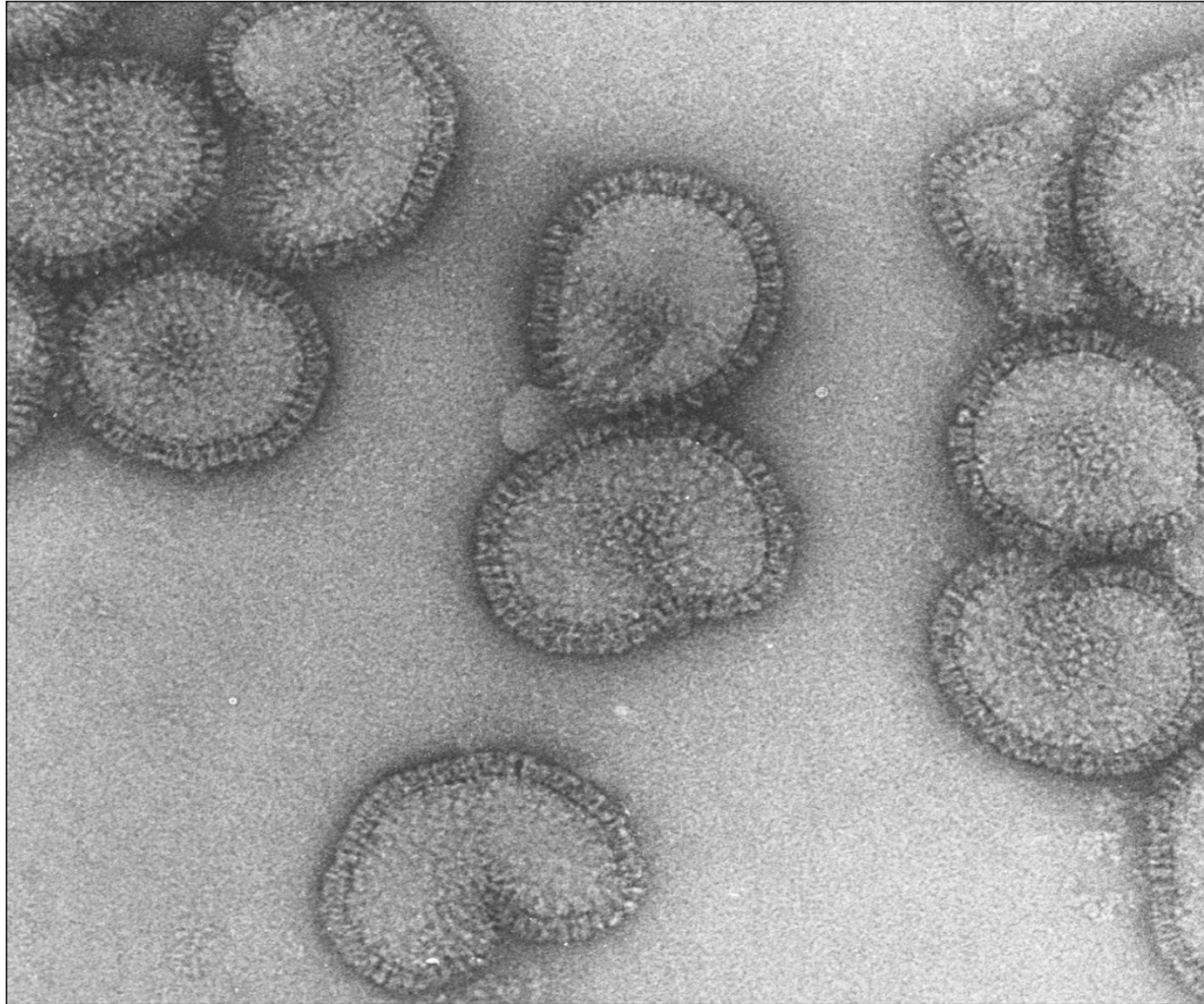
Influenza Virus Neuraminidase Inhibitors

**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



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

Influenza Virus

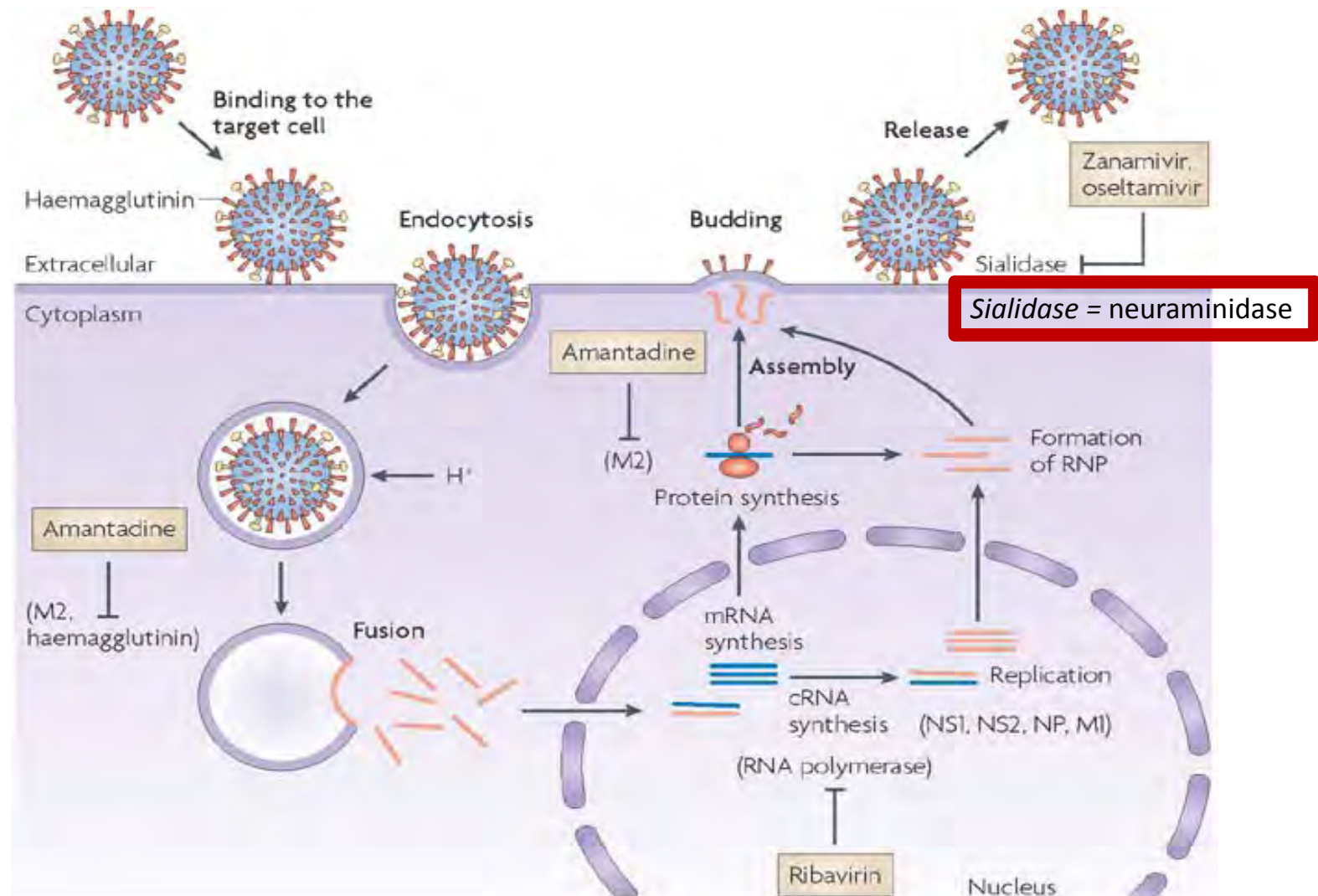
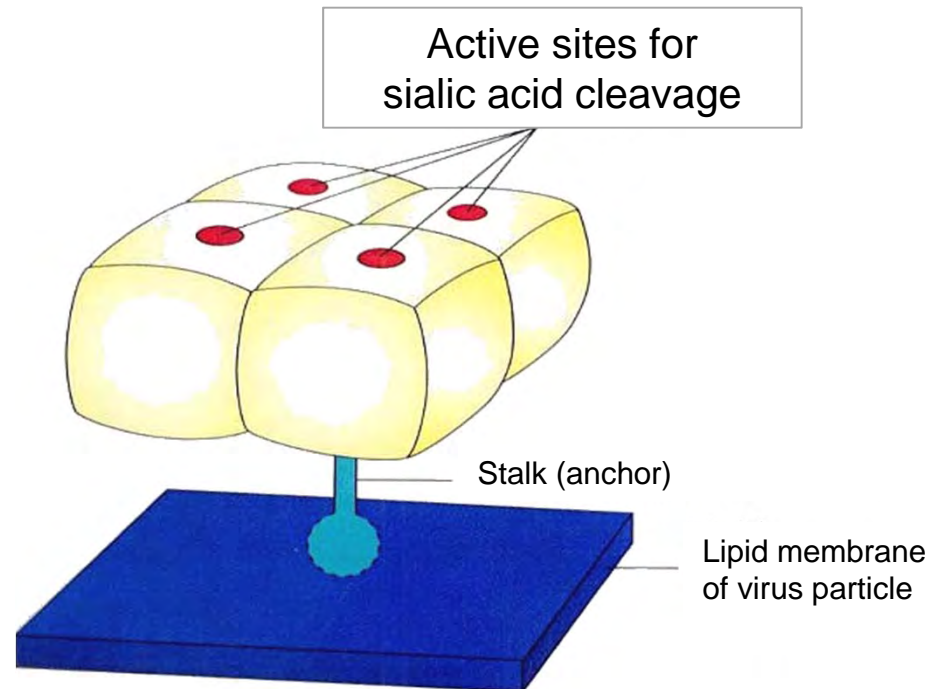


Figure from: von Itzstein, M. (2007). *Nature Reviews Drug Discovery* 6, 967-974

Influenza Virus Neuraminidase.

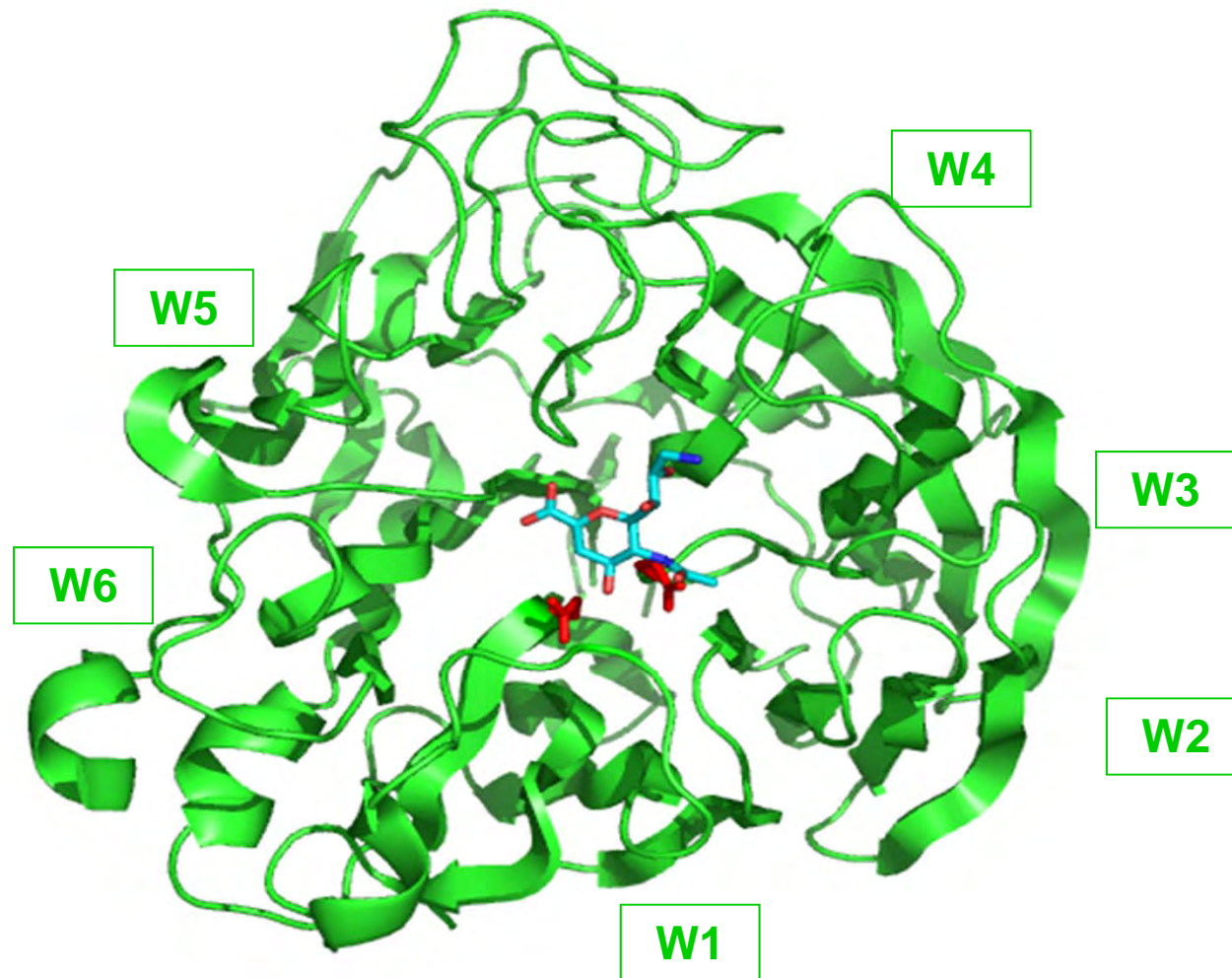


Schematic representation of the neuraminidase tetramer on the surface of the influenza A virus: the “NA spike”

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 way to decrease the severity of a flu infection.

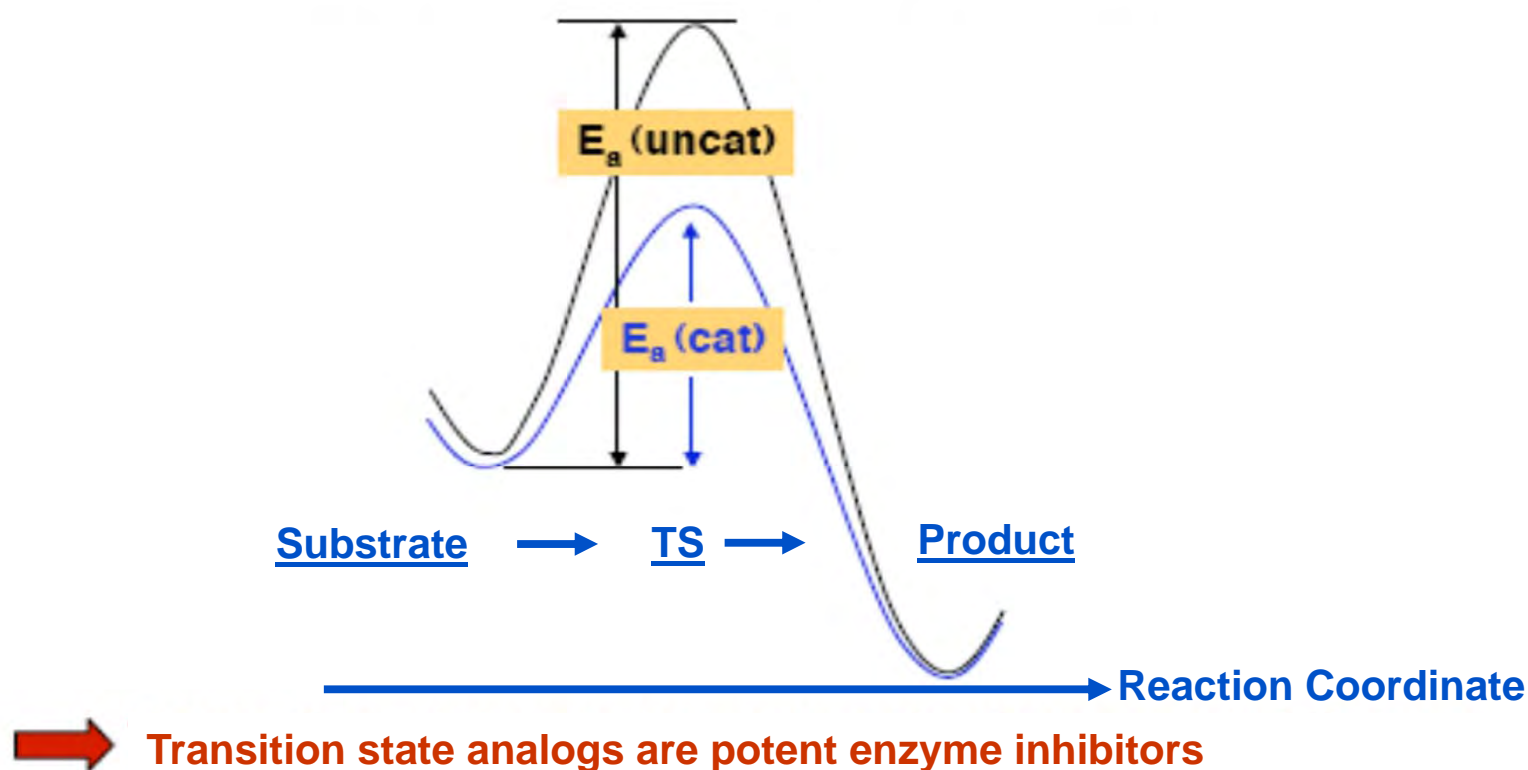
The Three-dimensional Structure of a single Subunit of Influenza Virus Neuraminidase



W_n = n-th 4-β-stranded “propeller”

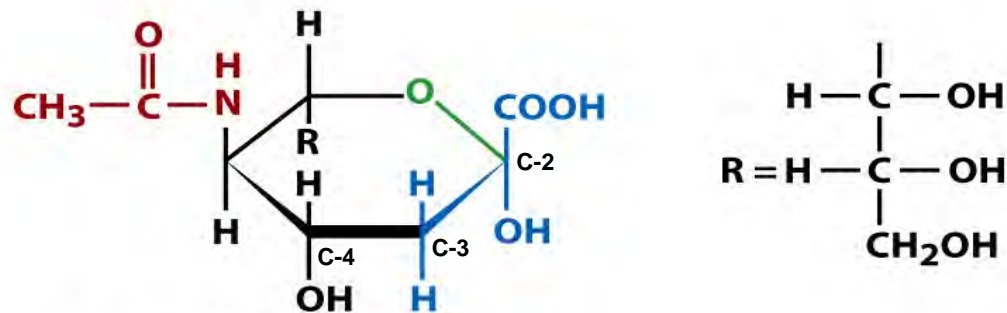
SMITH et al, PROTEIN SCI. 10: 689 (2001) – PDB-code 1F8D.

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



TS = Transition State

The substrate of neuraminidase



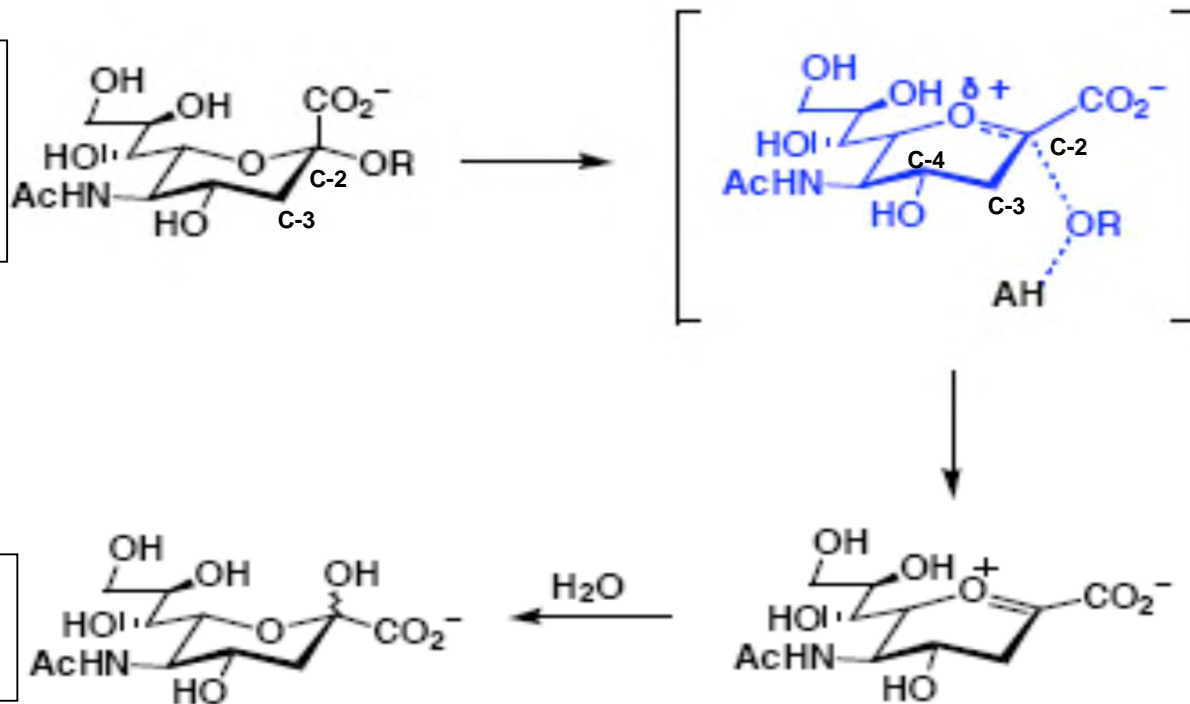
**N-Acetylneuraminic acid
(pyranose form)**

Sialic Acid \equiv N-acetylneuraminic Acid

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

The Reaction catalyzed by neuraminidase

Substrate
(Sialic Acid bound
to a protein
via the 2-OH)

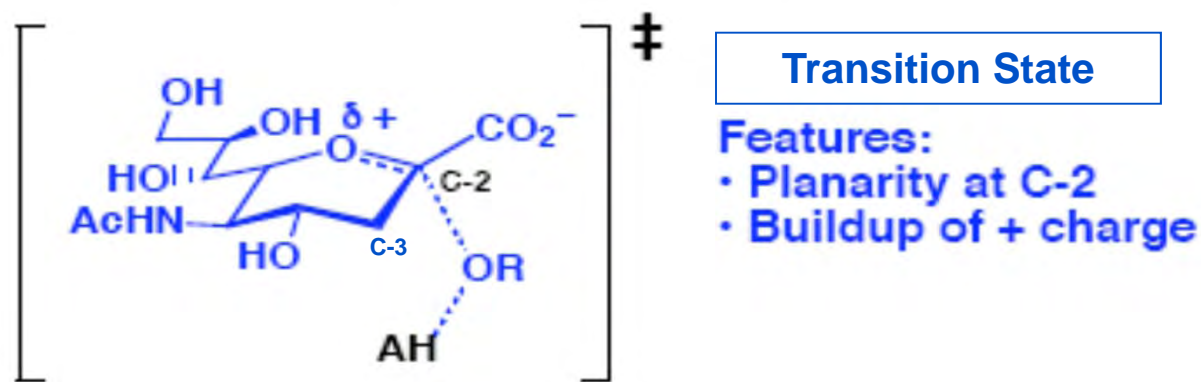


Transition State:
With the Ring-O, the C-2, the Carboxyl carbon and the C3 approximately in one plane.

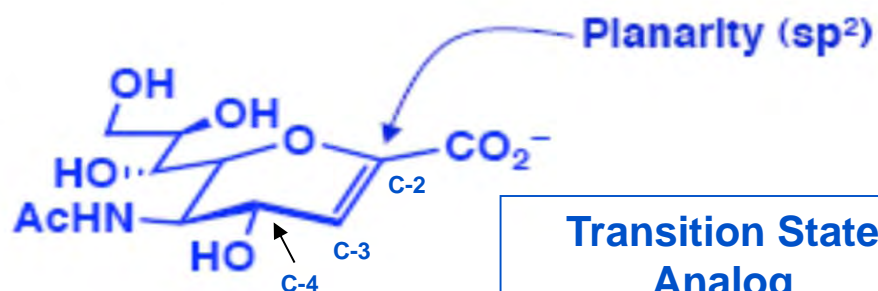
Product
(Free Sialic Acid)

Sialic acid = N-ACETYL-NEURAMINIC ACID

Design of Transition State Analog neuraminidase inhibitors



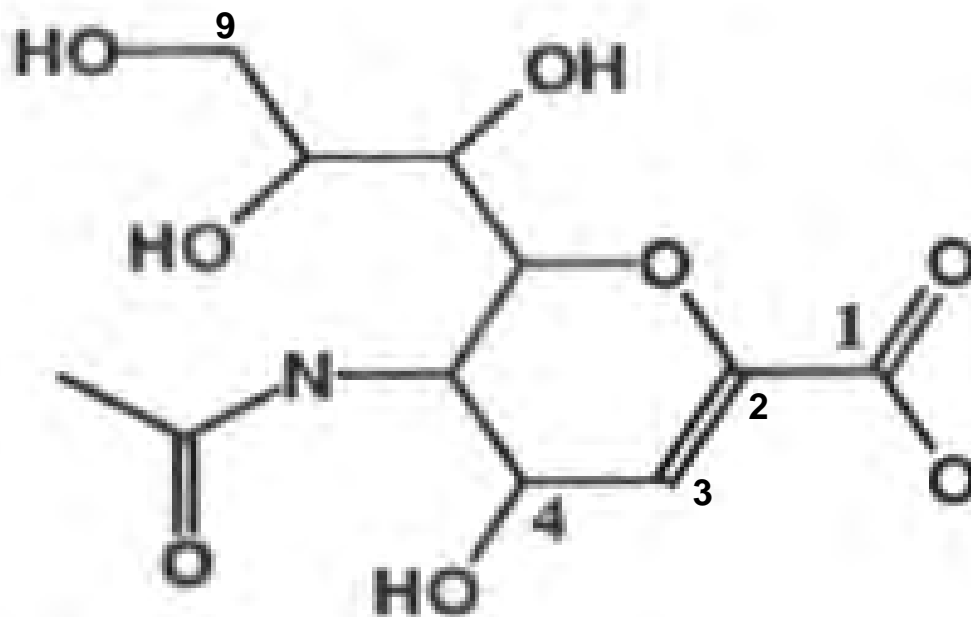
DANA
 $K_i = 10^{-6} \text{ M}$
 Also inhibits human neuraminidase



DANA \equiv
2-DEOXY-2,3-DEHYDRO-N-ACETYL-NEURAMINIC ACID

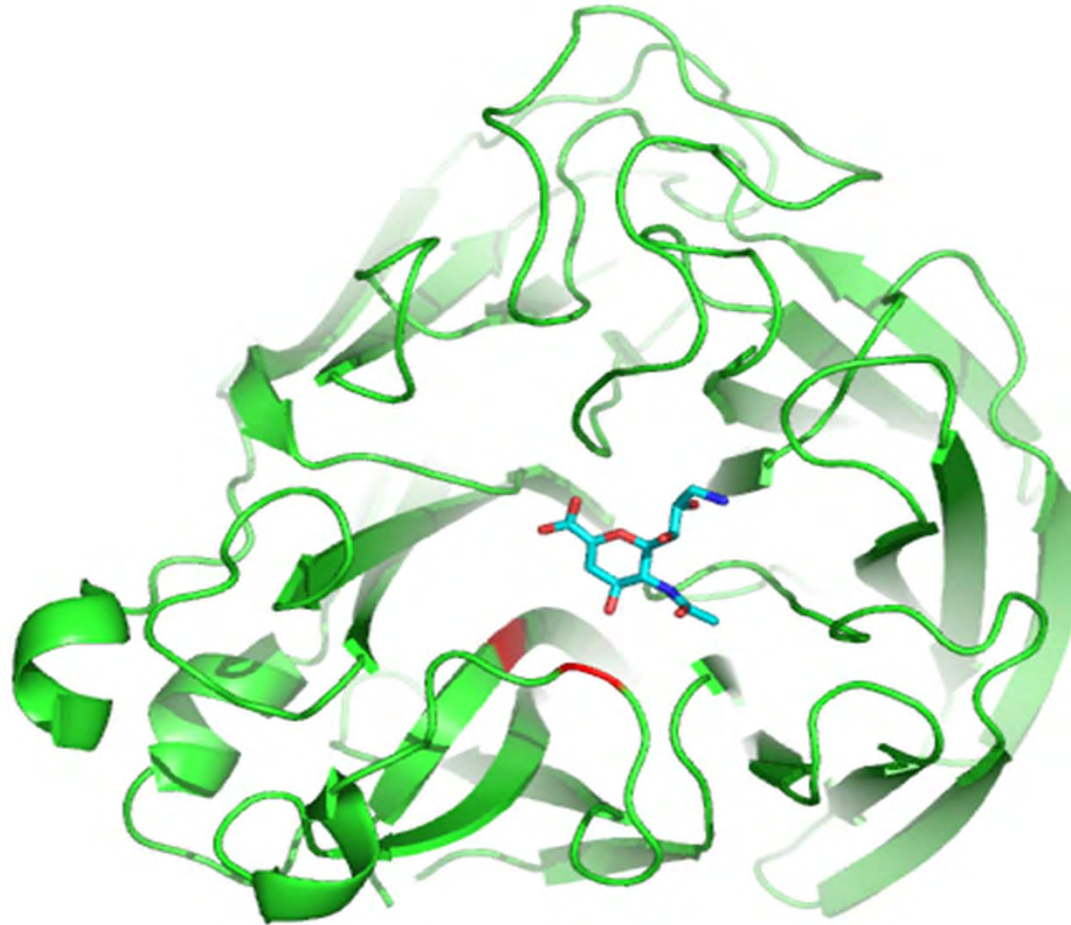
The Starting Point

The Transition State Analog (TSA) DANA



DANA

Influenza Virus Neuraminidase in complex with 9-amino-DANA

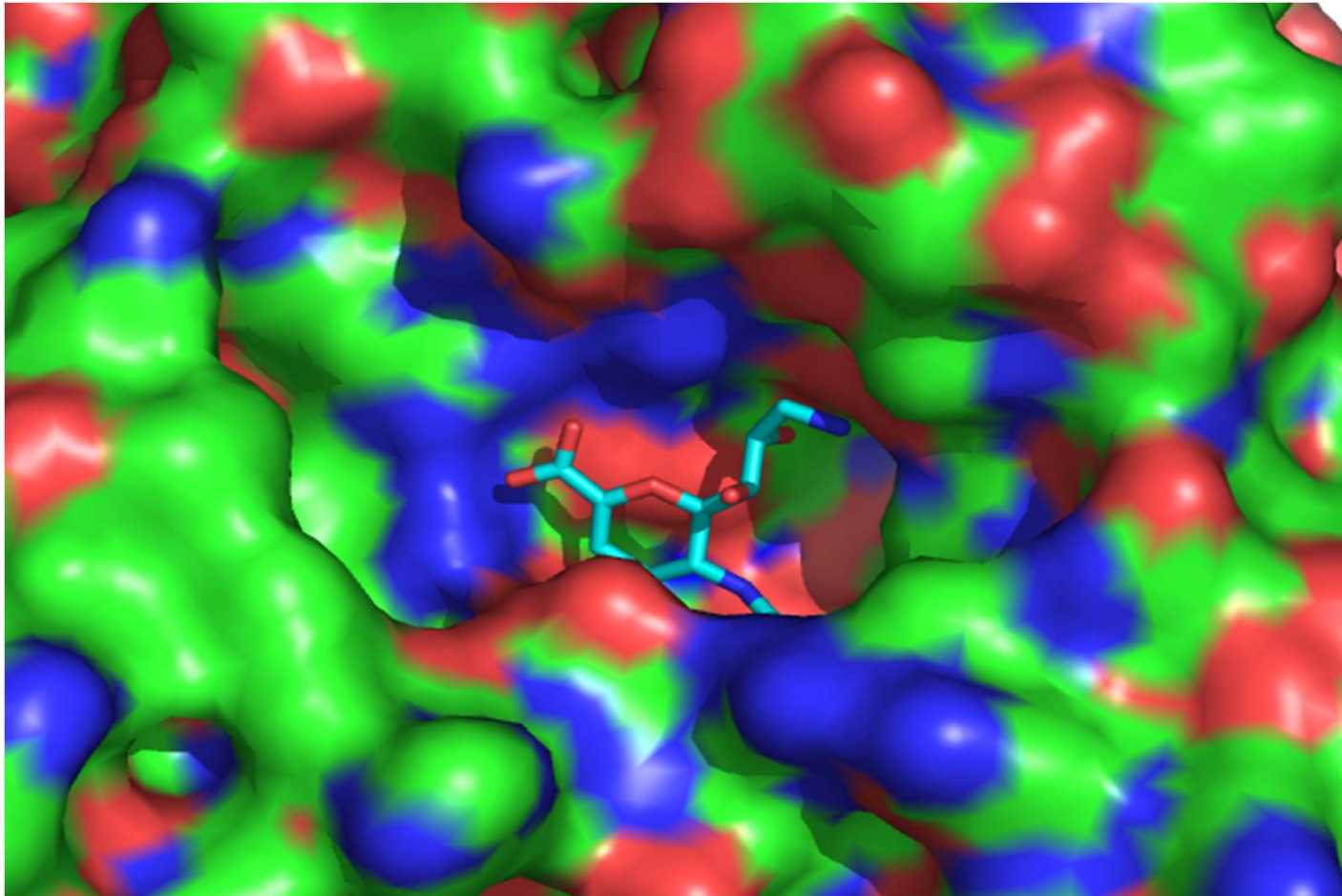


9-amino-DANA =

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

(The 9-amino group is irrelevant for the drug development story)

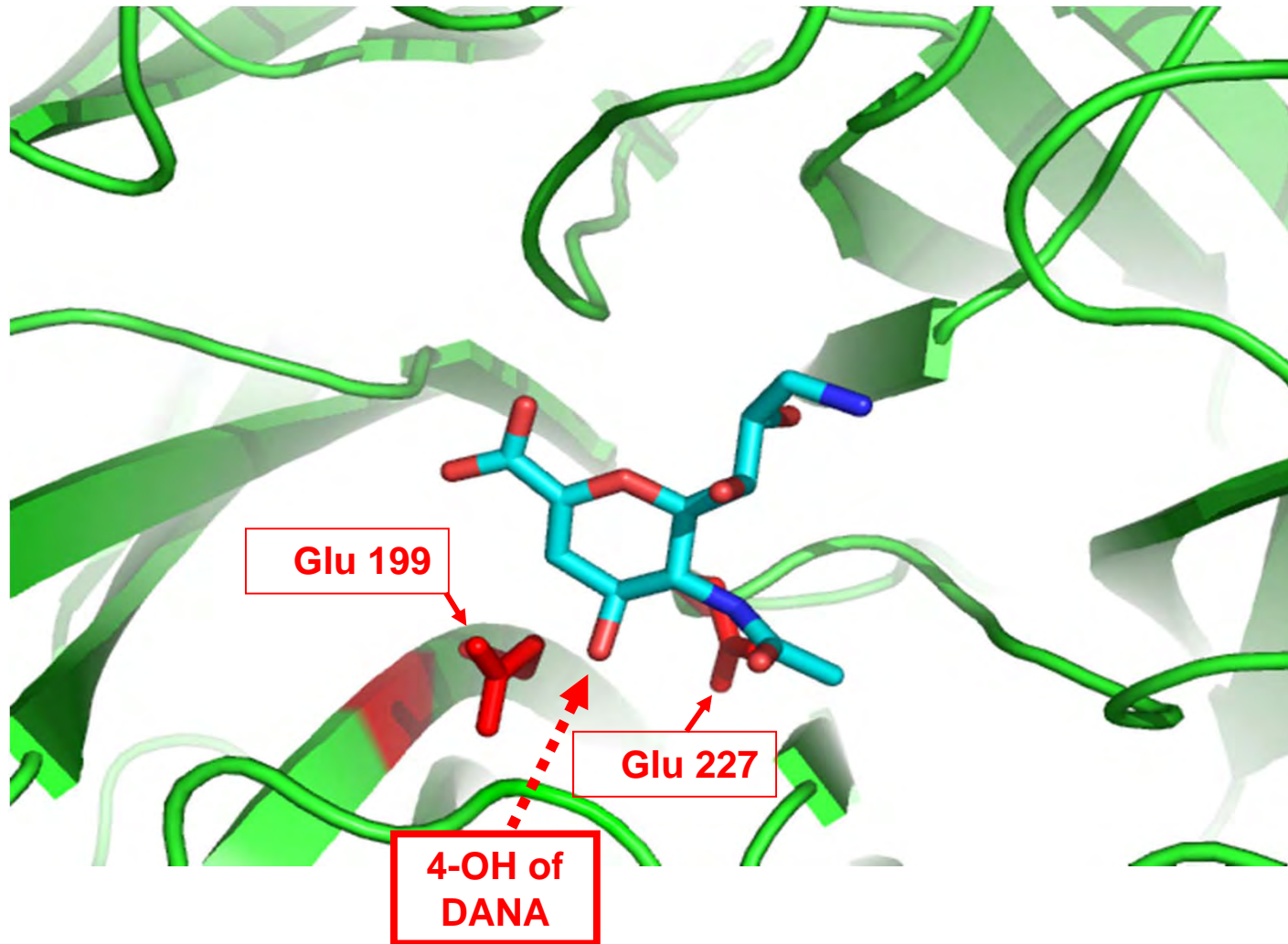
Influenza Virus Neuraminidase in complex with 9-amino-DANA



9-amino-DANA sits clearly in a pocket.

This is the active site of neuraminidase

Influenza Virus Neuraminidase in complex with 9-amino-DANA

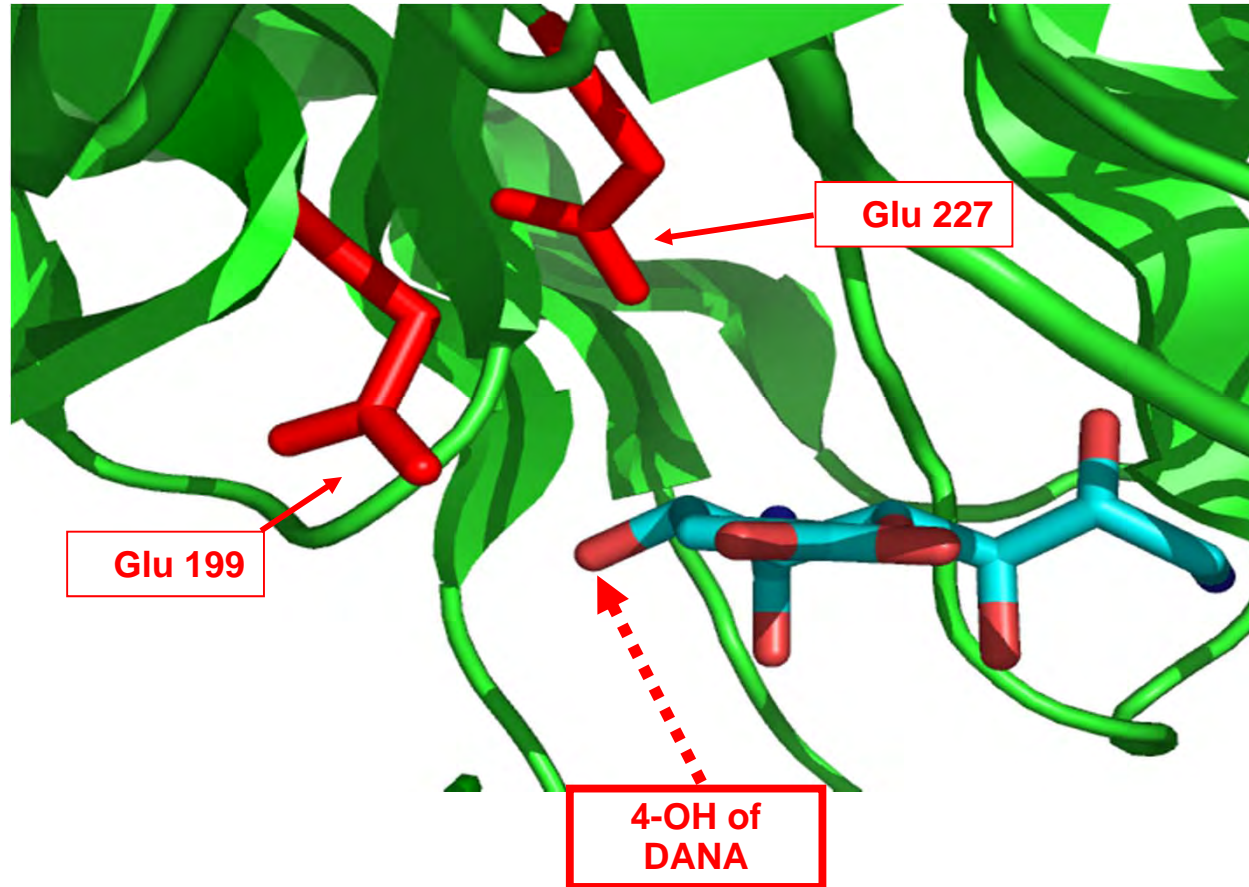


View of two key Neuraminidase residues near the 4-OH of 9-amino-DANA

Influenza Virus Neuraminidase in complex with 9-amino-DANA

Close-up and 90 degrees rotated

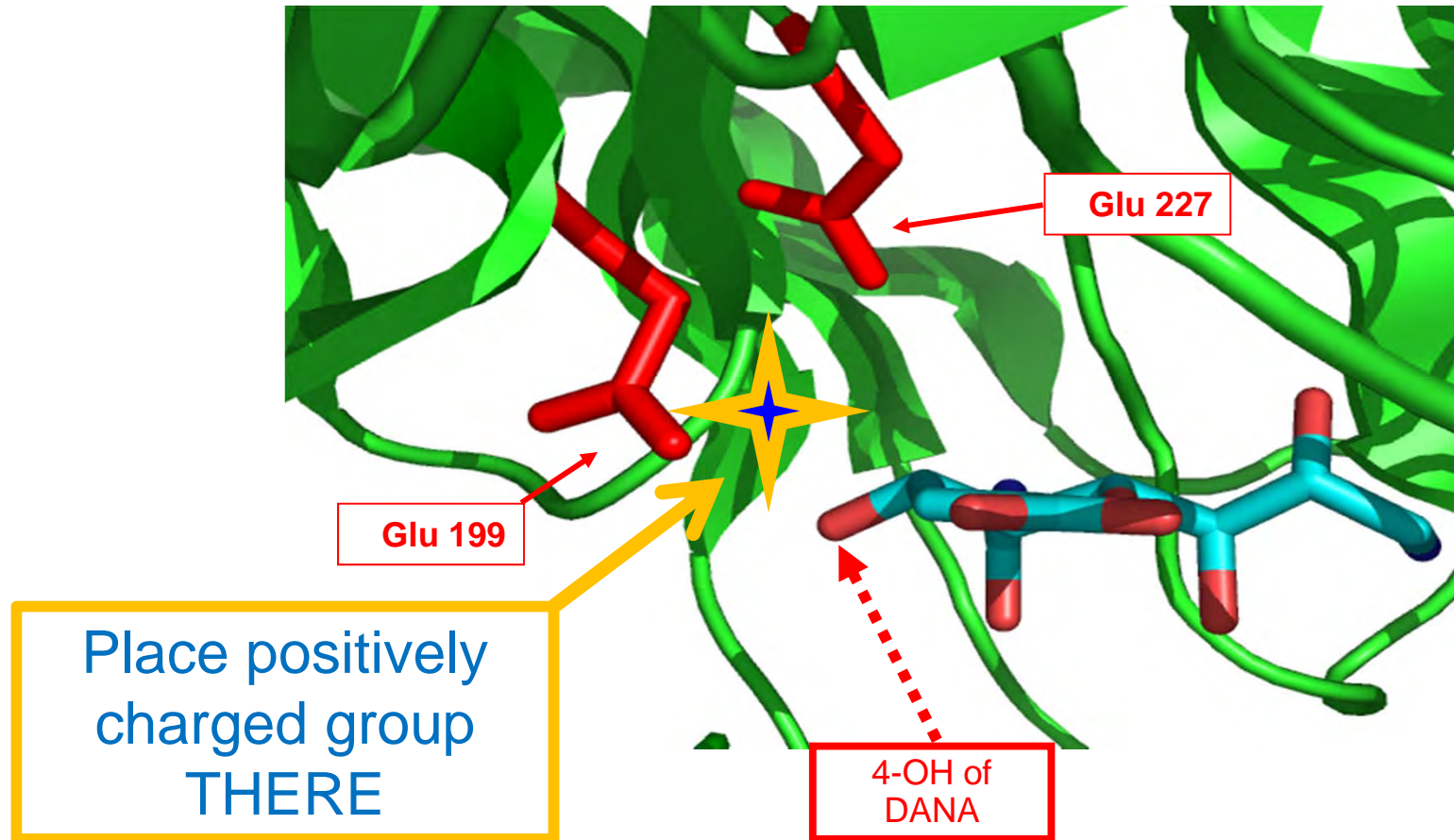
Influenza Virus Neuraminidase in complex with 9-amino-DANA



Two negatively charged carboxylates are quite close to the 4-OH

Bivalent Neuraminidase Inhibitors

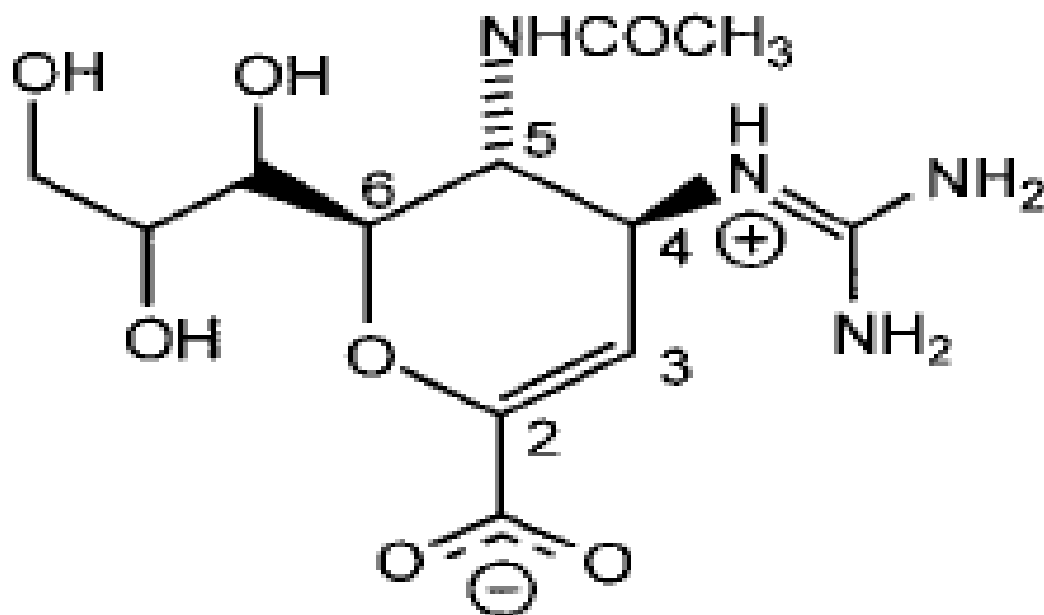
Influenza Virus Neuramininidase in complex with 9-amino-DANA



Quite obvious what to do to gain affinity – TWO compounds made

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:

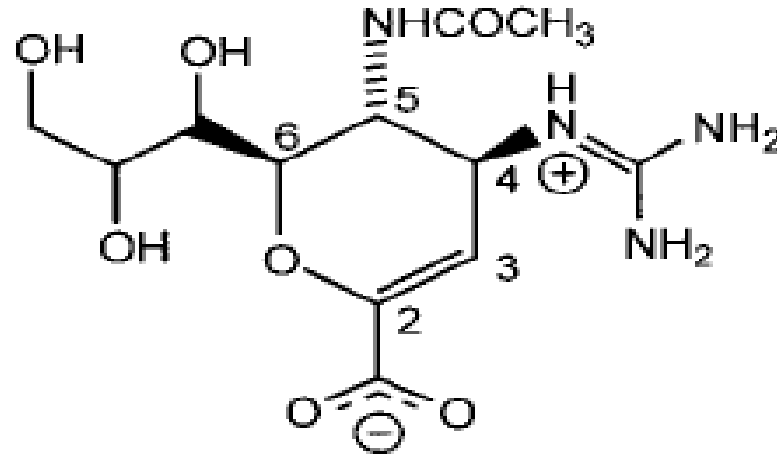
	<u>Flu Neura</u>	<u>Human Neura</u>
DANA	1×10^{-6}	1.2×10^{-5}
4-guanidino-DANA	2×10^{-10}	1×10^{-3}

By changing one single functional group:

- The affinity for the target flu enzyme was enhanced by a factor of ~10,000;
- The affinity for the human homologous enzyme was decreased by a factor of ~100.
- The selectivity was improved by a factor ~1,000,000!!!

Properties of 4-guanidino-DANA

Zanamivir (Relenza)



Zanamivir

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 (four) Lipinski “Rules of Five”

“From the 50,427 compounds in the WDI (World Drug Index) File2245 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."

Medicines have to fulfil many requirements

Drugs are VERY Precious compounds

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)

A toxin 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 traveler's diarrhea

-

The power of multivalency

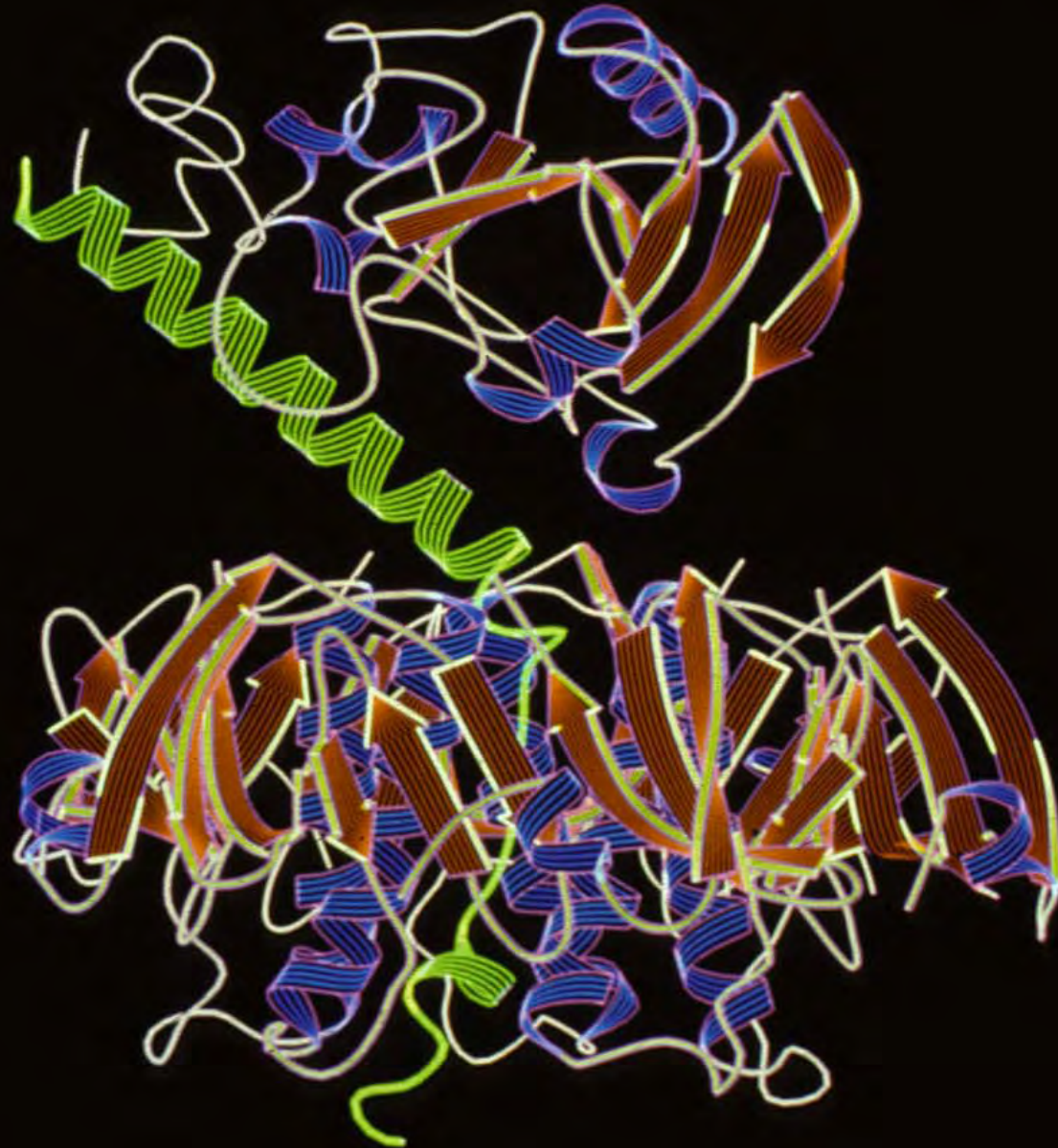
Minke, J. The role of waters in flexible docking strategies for carbohydrate derivatives: heat-labile enterotoxin, a multivalent test case. *J. Med. Chem.* 42, 1778-1788 (1999).

Fan, A 10^5 gain in affinity for pentavalent ligands of *E. coli* heat-labile enterotoxin by modular structure-based design. *J. Am. Chem. Soc.* 122, 2663 (2000)

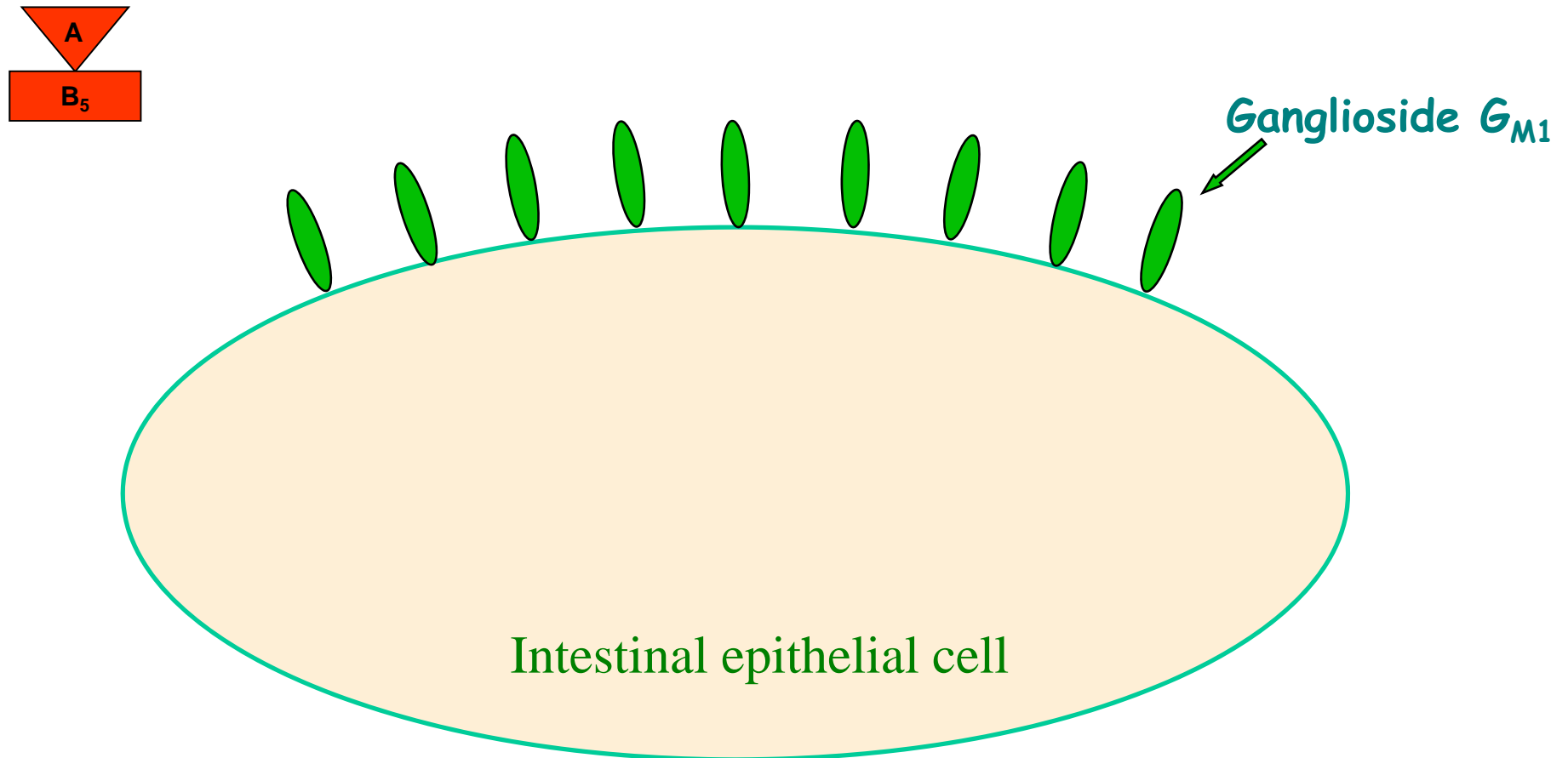
Heat-labile Enterotoxin (LT), a very close relative of Cholera toxin (CT)

A subunit

B pentamer



CT and LT Receptor Binding

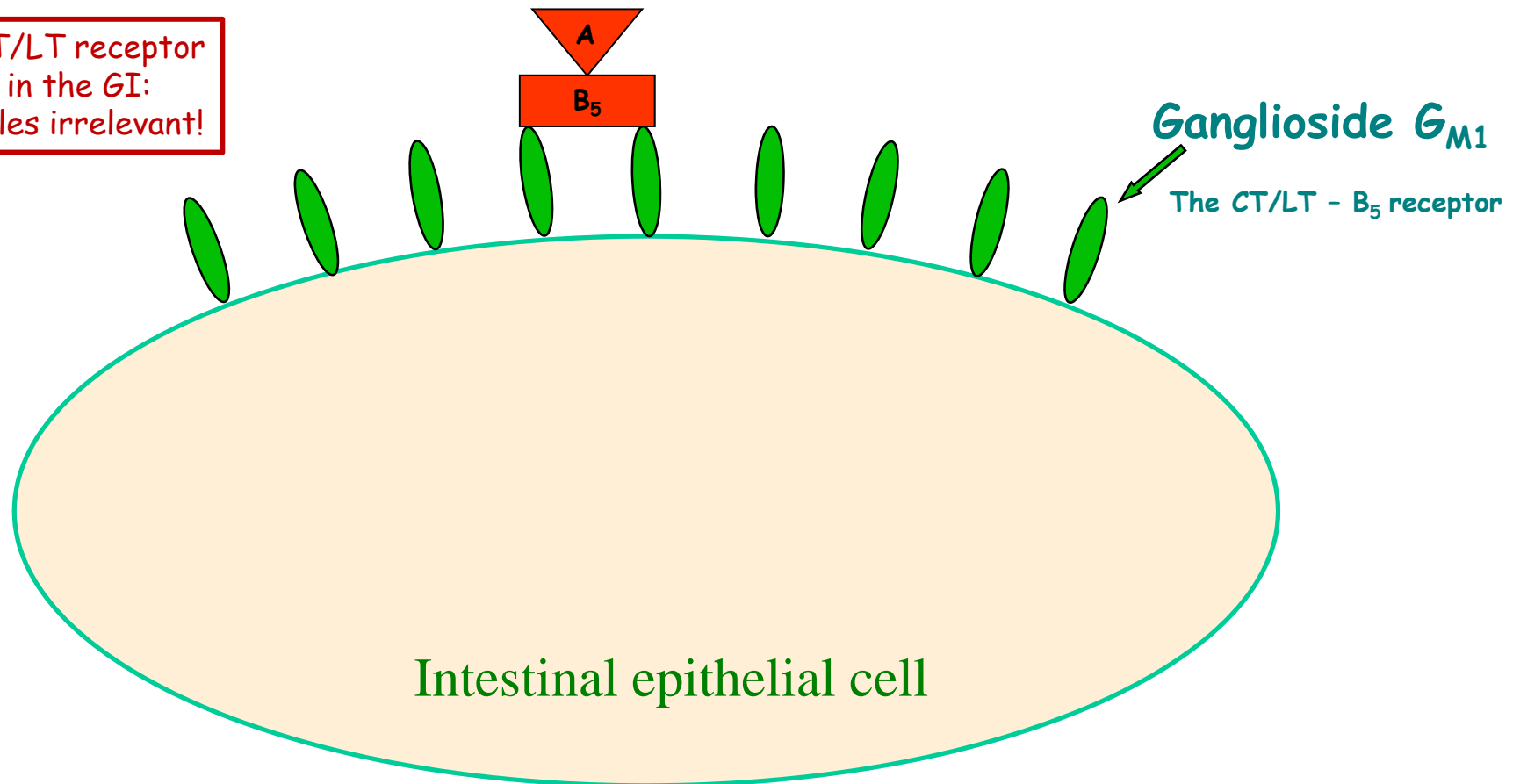


CT : Cholera

LT : Traveller's & Children's diarrhea

CT and LT Receptor Binding

For blocking CT/LT receptor
recognition in the GI:
The Lipinski rules irrelevant!



CT : Cholera

LT : Traveller's & Children's diarrhea

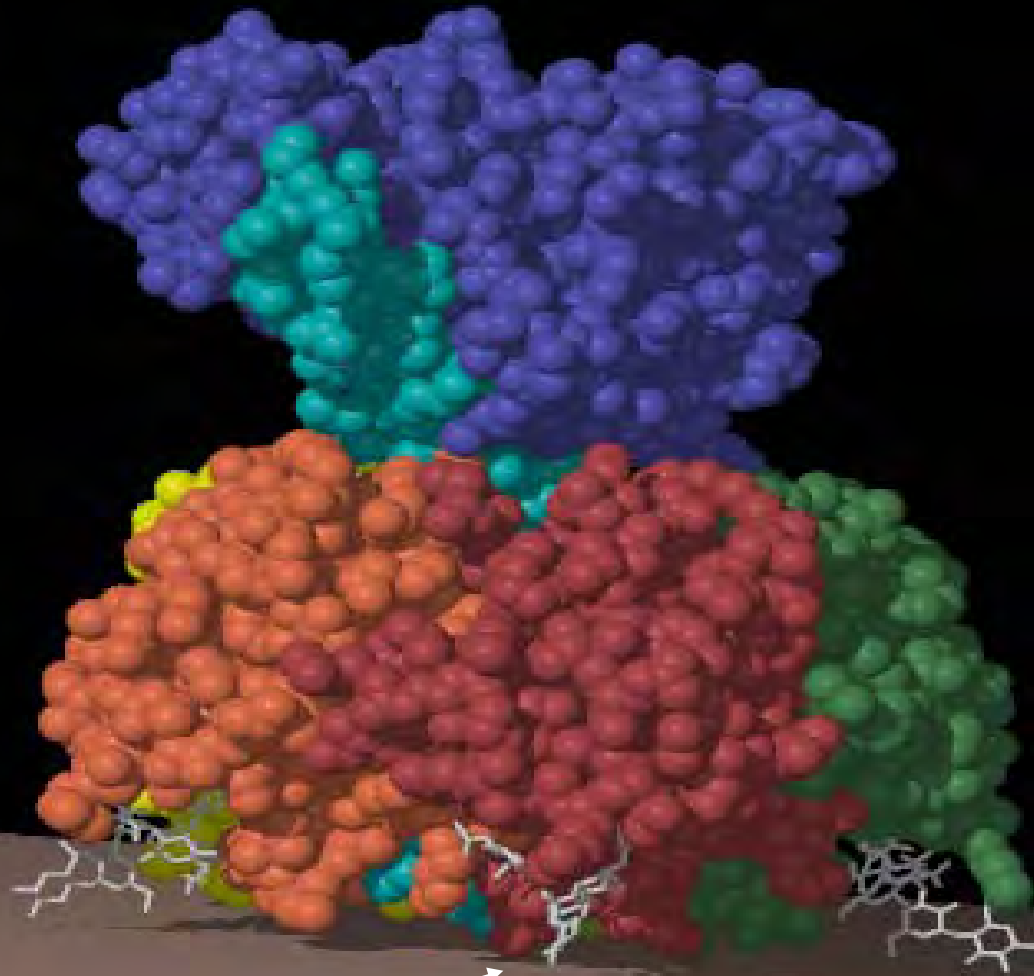
Cholera toxin - G_{M1} Receptor Interaction

A subunit

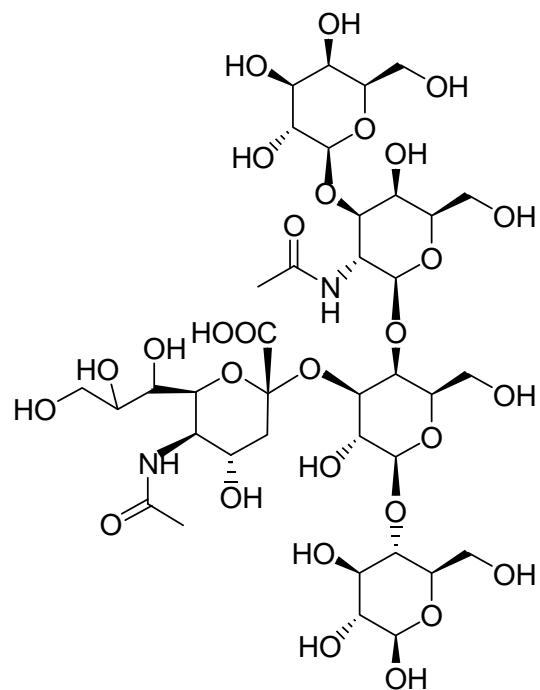
B pentamer

GM1-OS

Intestinal cell surface

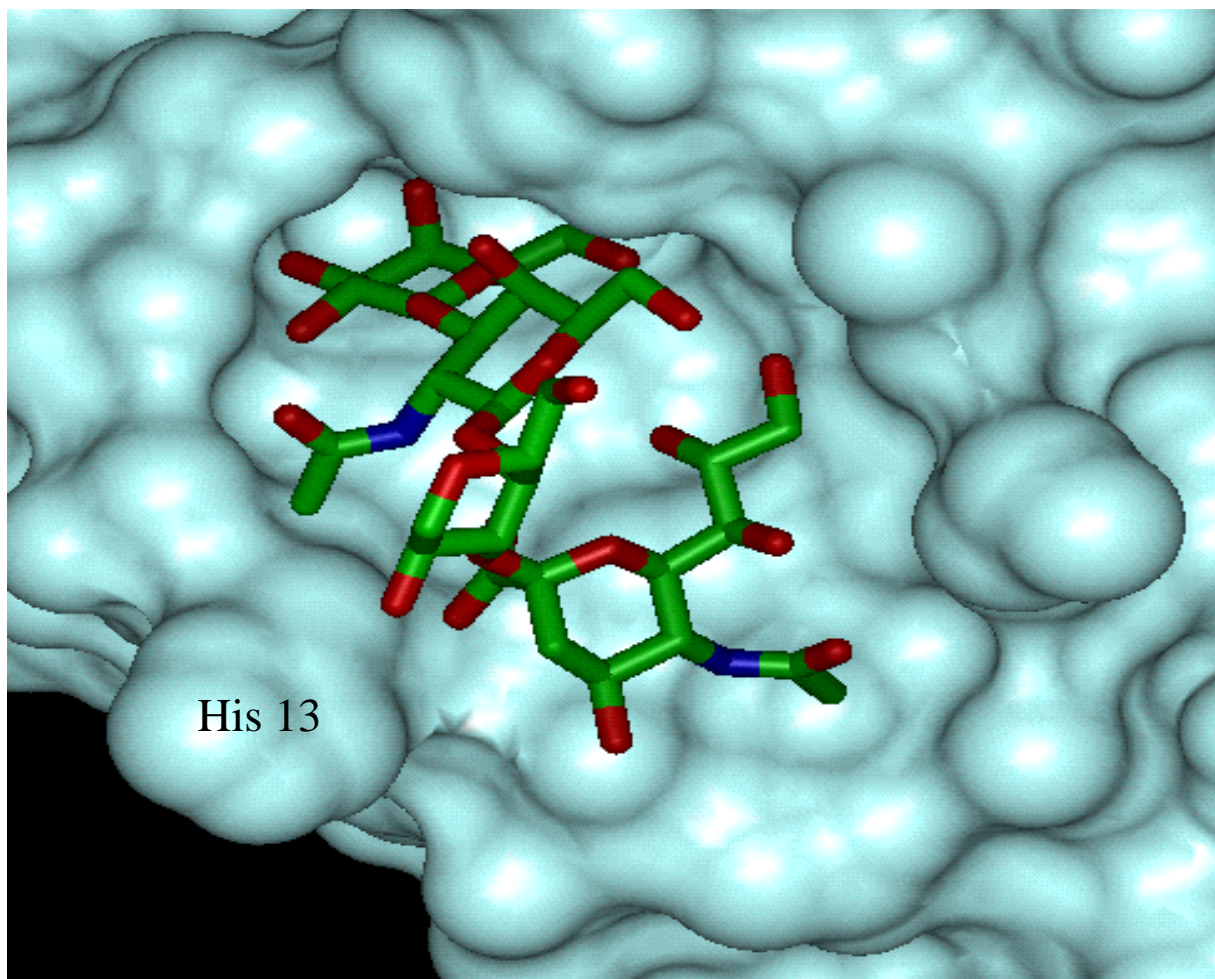


G_{M1} Pentasaccharide bound by CT



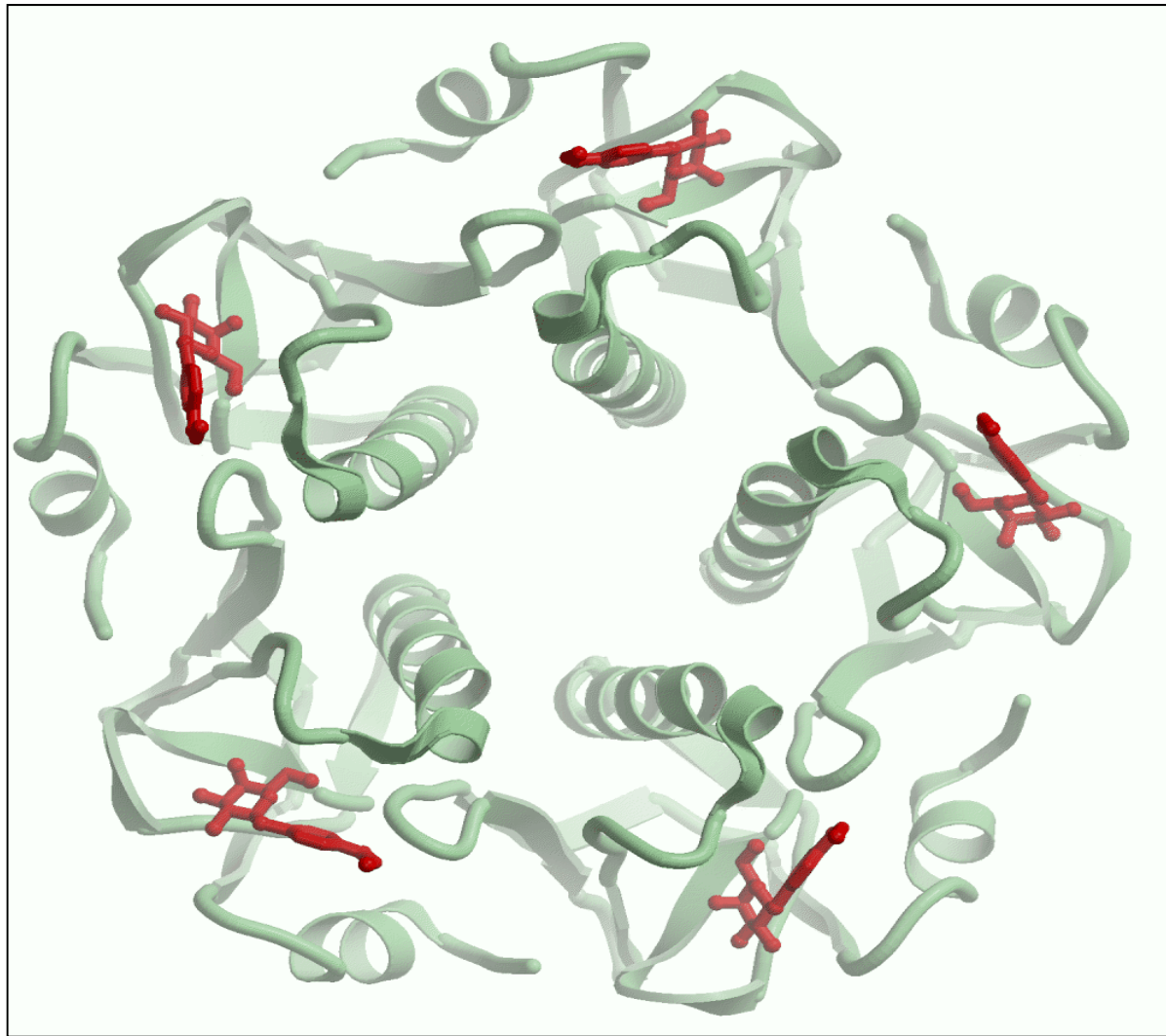
$$IC_{50} = 14 \times 10^{-9} M$$

Extensive hydrophobic and
H-bonding interactions

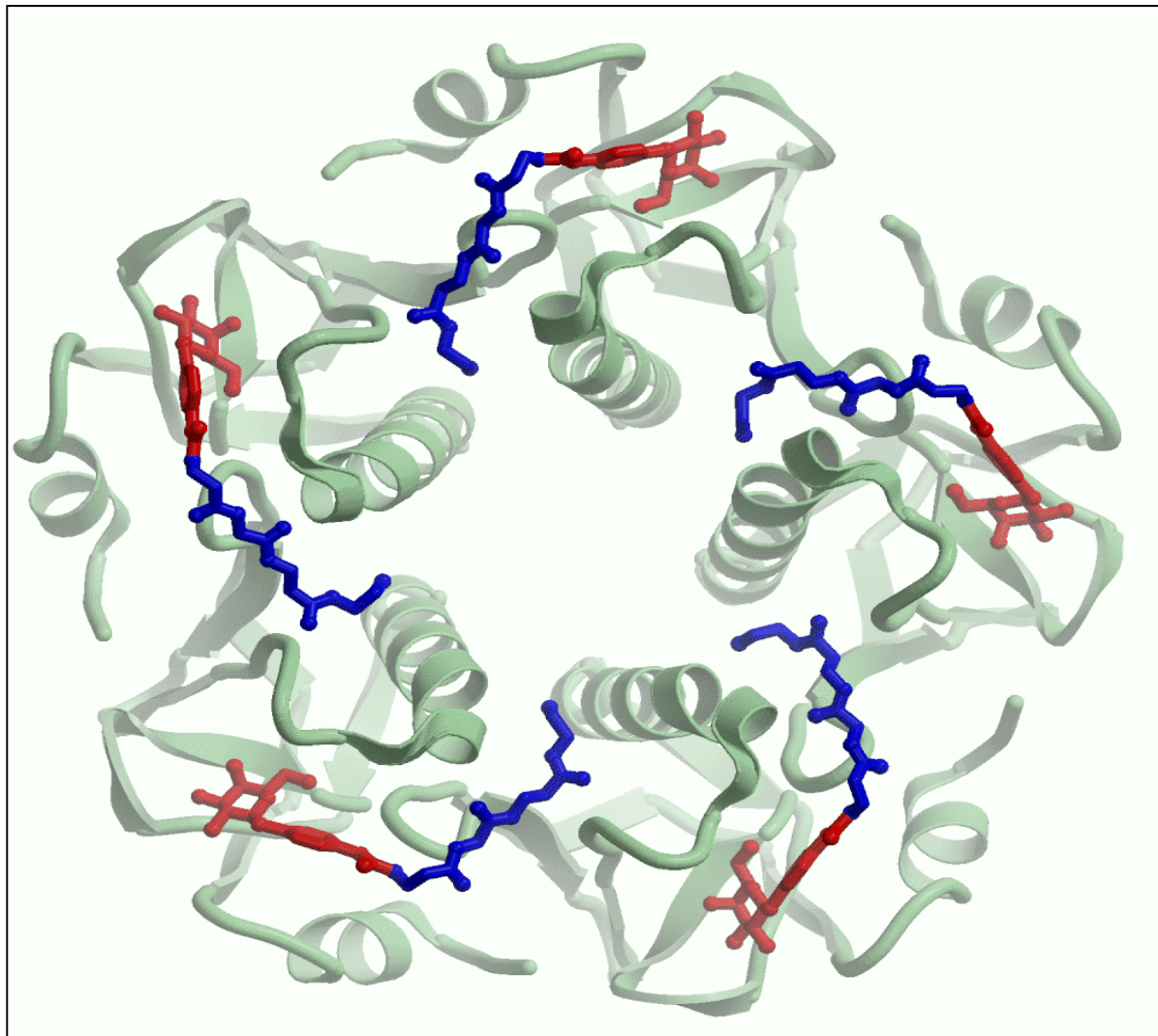


The enemy

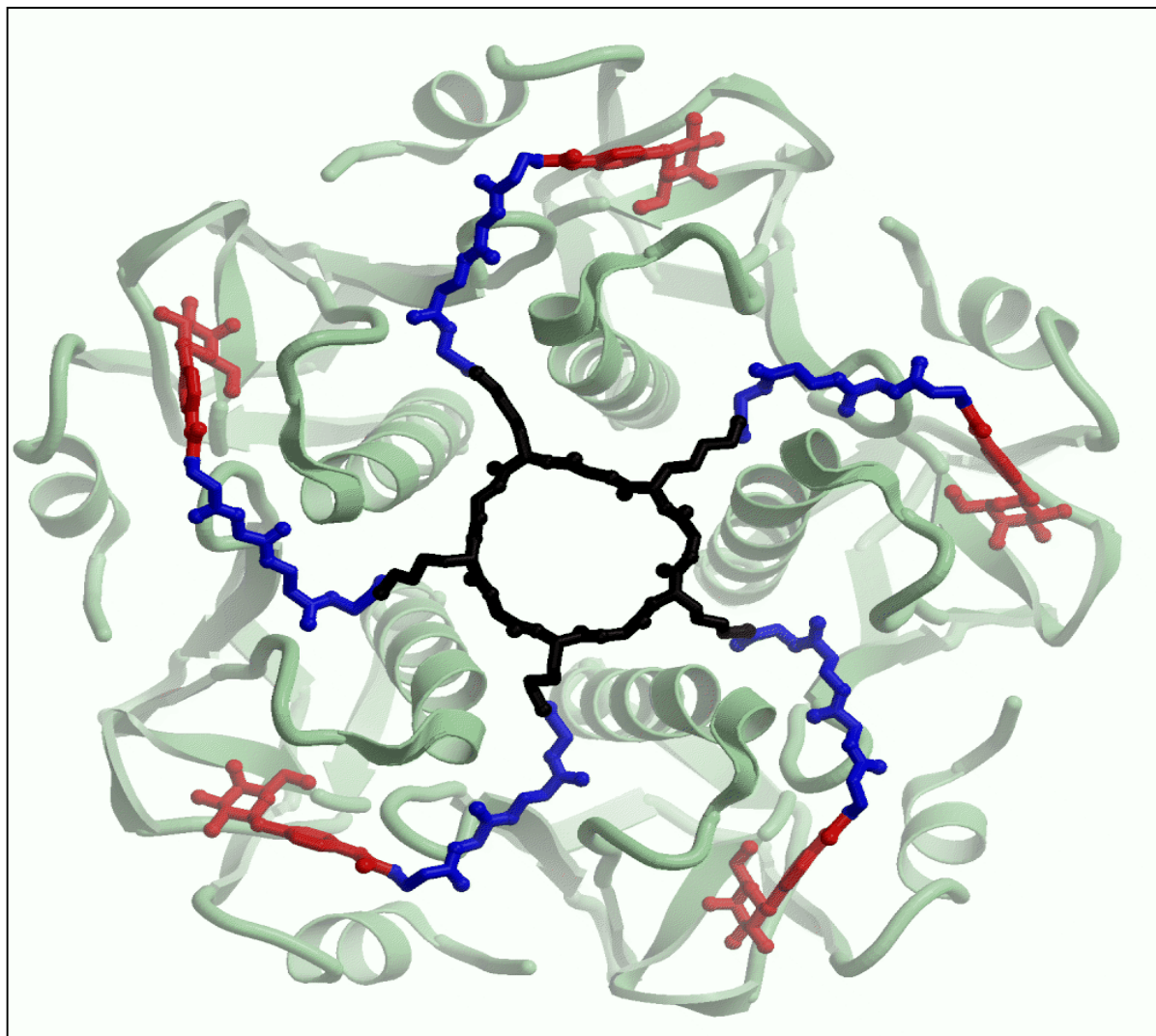
Five receptor binding sites



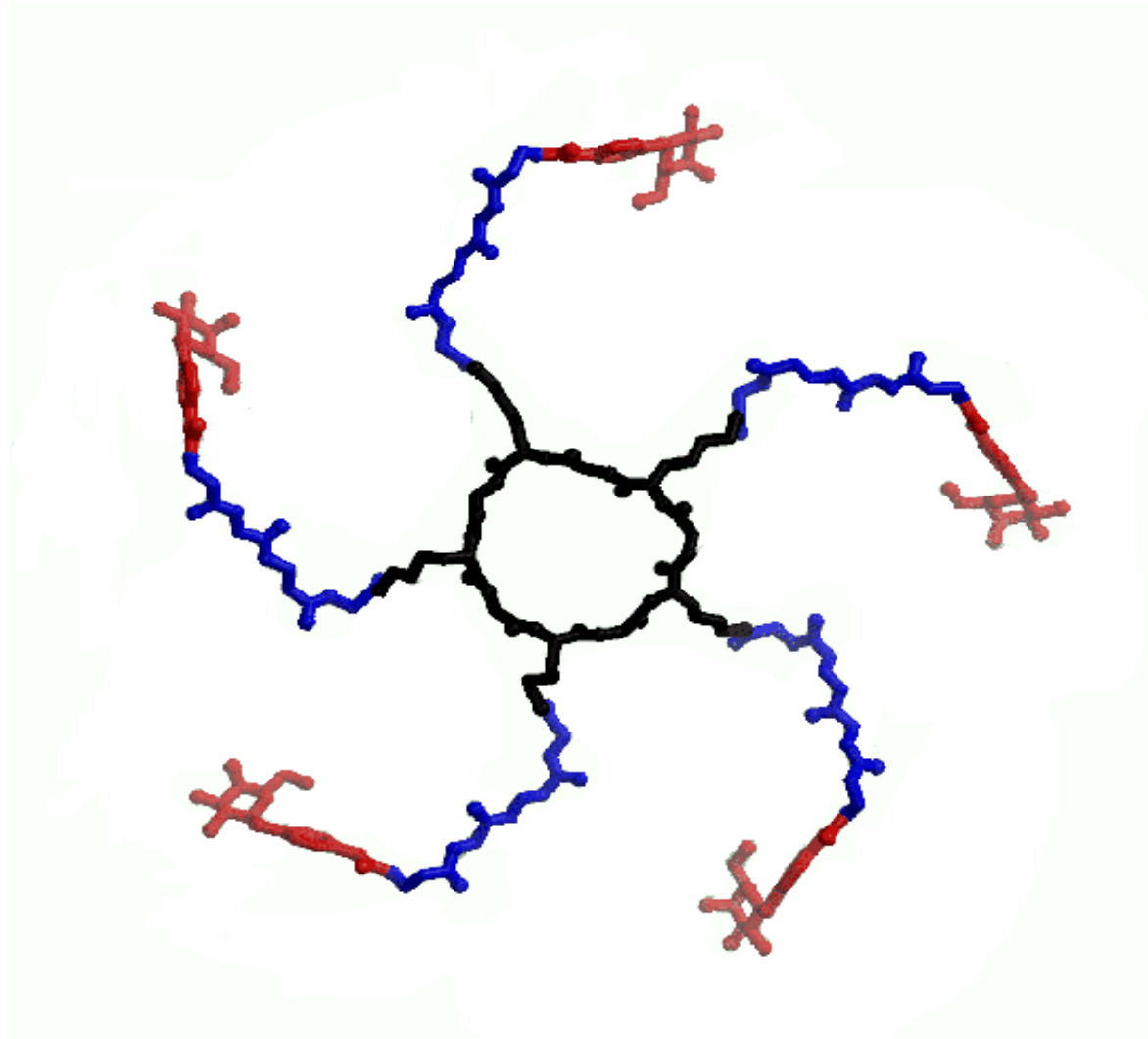
Making ligands longer



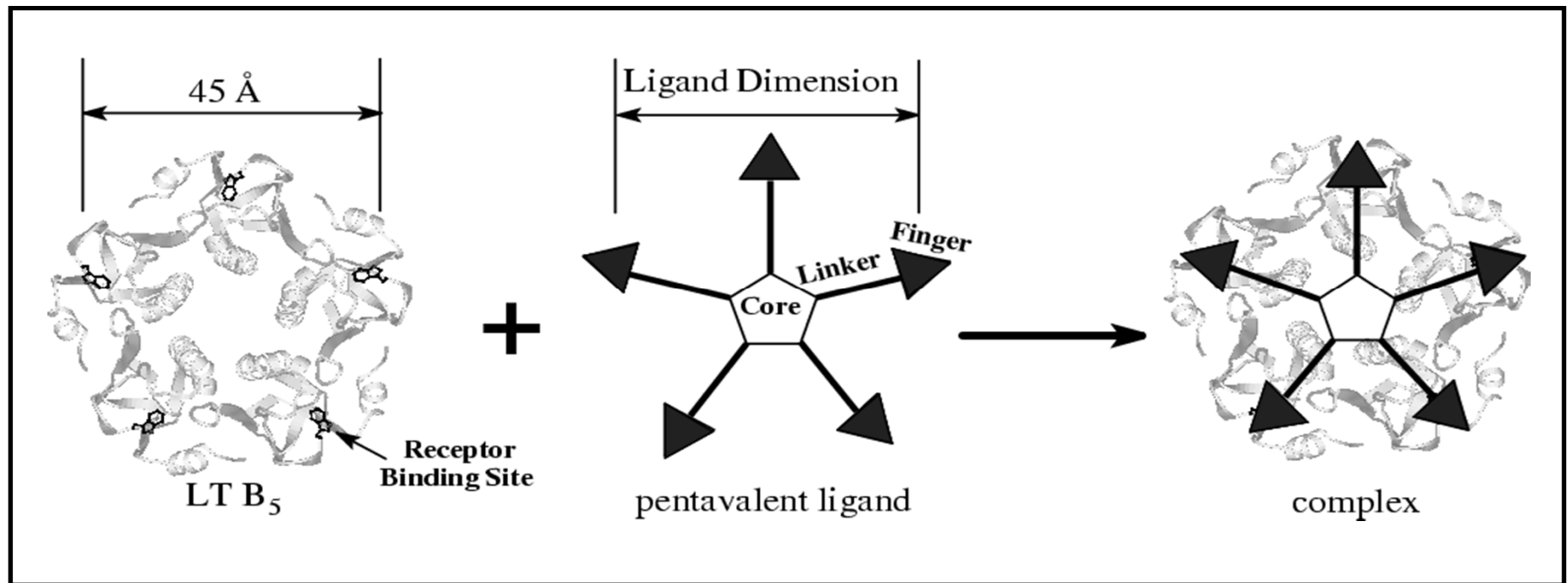
Ligand-Protein Complex



Pentavalent Ligand

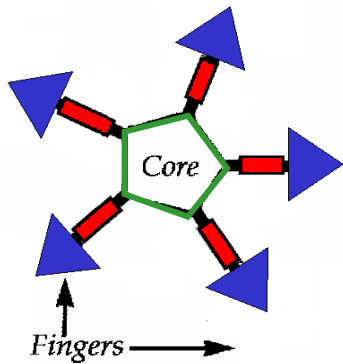


THE PENTAVALENT CONCEPT



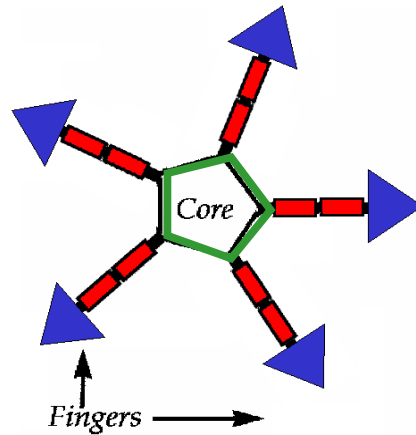
Gains in surface-receptor binding inhibition

One-Unit Linker



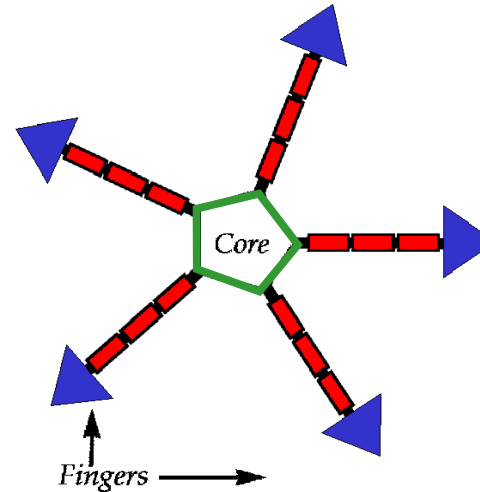
**240 x
Single
Finger**

Two-Unit Linker



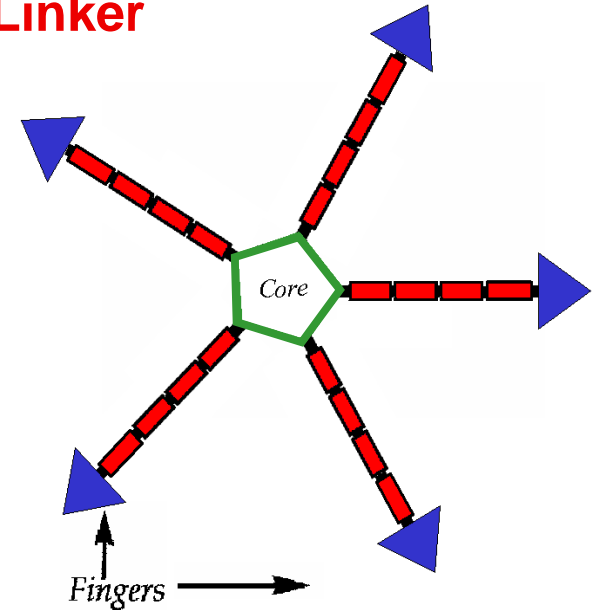
**3,600 x
Single
Finger**

Three -Unit Linker



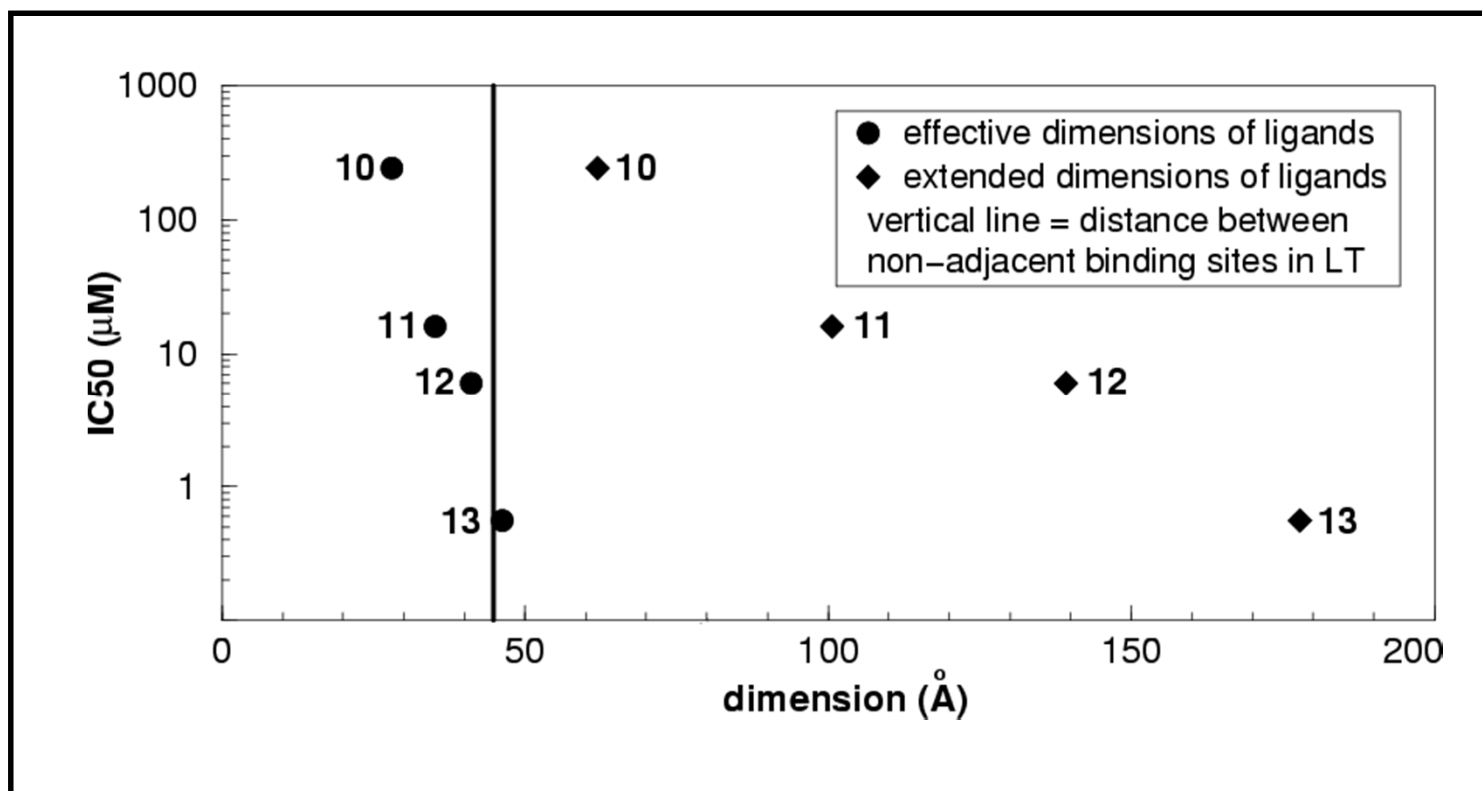
**10,000 x
Single
Finger**

Four-Unit Linker



**104,000 x
Single
Finger**

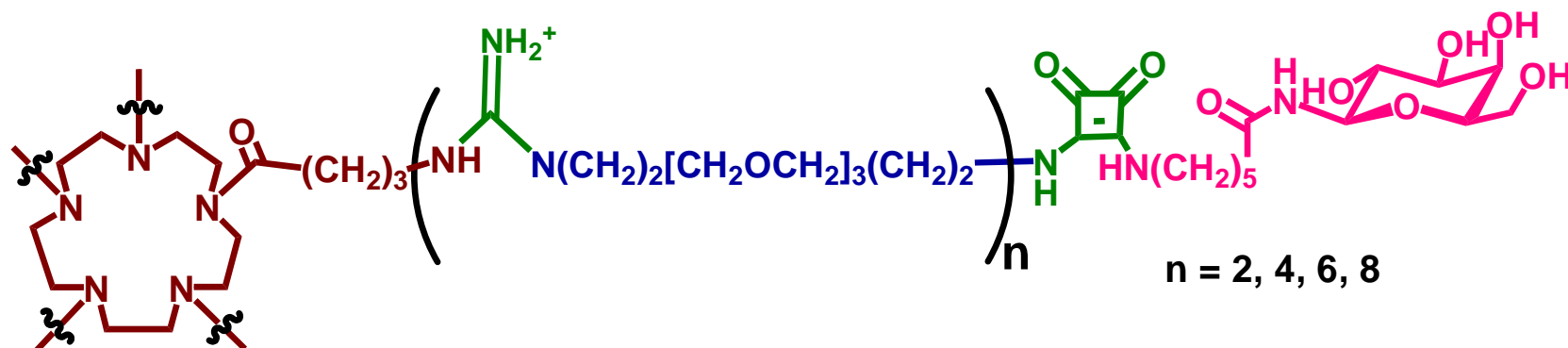
IC_{50} versus EXTENDED(♦) & EFFECTIVE (●) DIMENSIONS OF PENTAVALENT LIGANDS



What if even longer linkers?

Erkang Fan and co-workers

And, indeed, linker too long : *less* affinity



Linker Units

$n = 2$

$n = 4$

$n = 6$

$n = 8$

IC₅₀ (μM)

13.26 ± 0.95

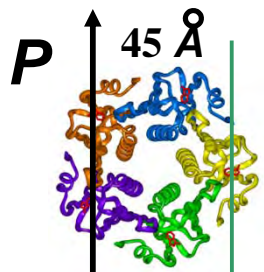
1.50 ± 0.10

4.63 ± 0.46

7.25 ± 0.38

(A single galactose "finger": IC₅₀ = ~ 100 mM = ~100,000 μM)

Erkang Fan and co-workers



And, indeed, linker too long : *less affinity*

IC_{50}

One-unit linker

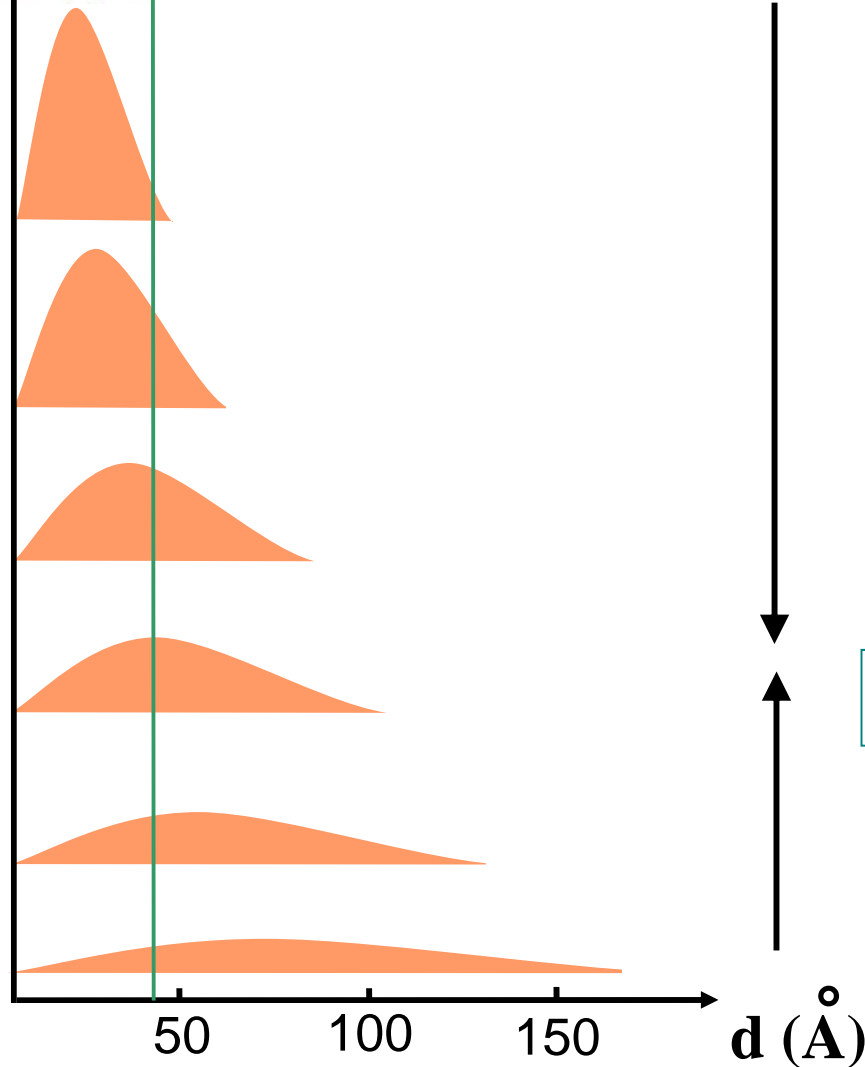
Two-unit linkers

Three-unit linkers

Four-unit linkers

Six-unit linkers

Eight-unit linkers



Fragment Cocktail Crystallography

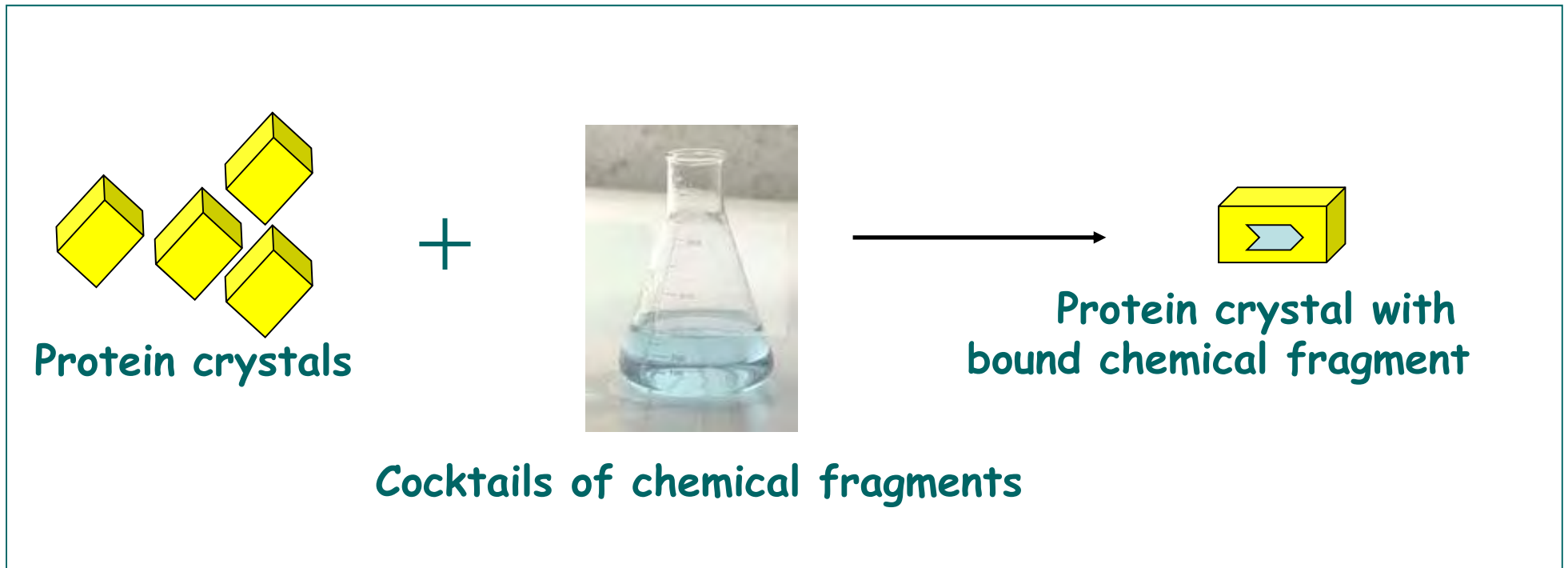


A new tool in drug design

Courtesy of Jürgen Bosch

Fragment Cocktail Crystallography

Principle



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.

Fragment Cocktail Crystallography

Making the compound library

9,500 compounds



fragmentation

626 fragments

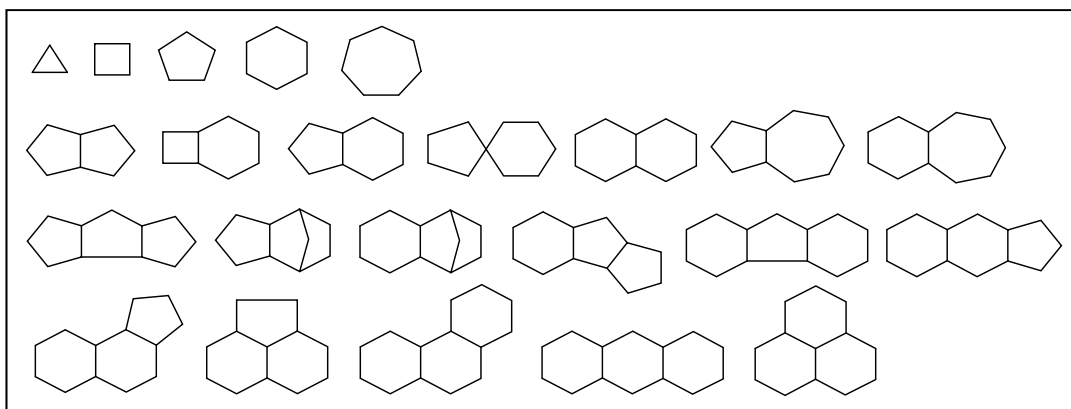


isolate ring systems

23 frameworks (at connectivity level)

ACD Compound Filtering

ACD= Available Chemical Database



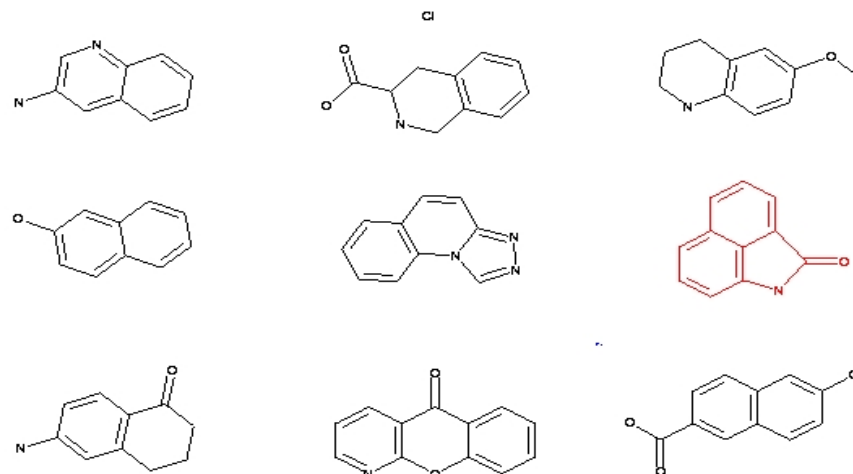
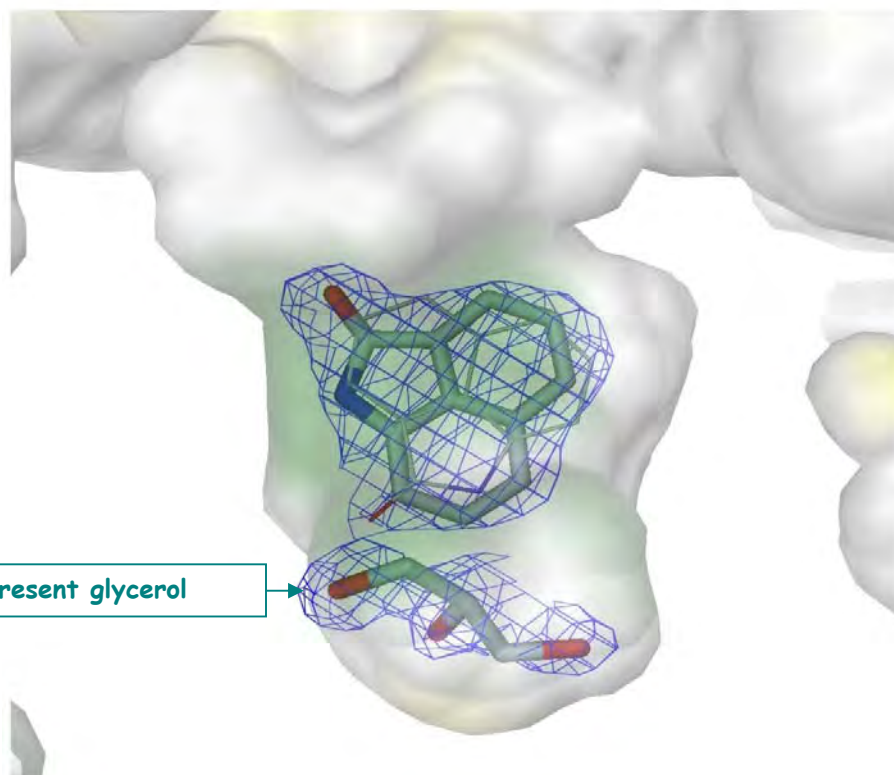
*manual selection
of compounds*

→ 680 compounds
*from each
framework class*

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

Christophe Verlinde, Erkang Fan
<http://faculty.washington.edu/verlinde/>

T. brucei Nucleoside 2-deoxyribosyltransferase Plus Cocktail #4



Left: the electron density
 Right: the compounds in the cocktail
 Conclusion: it is the red one, i.e. 1,2-dihydrobenzo[cd]indol-2-one

Influenza Virus

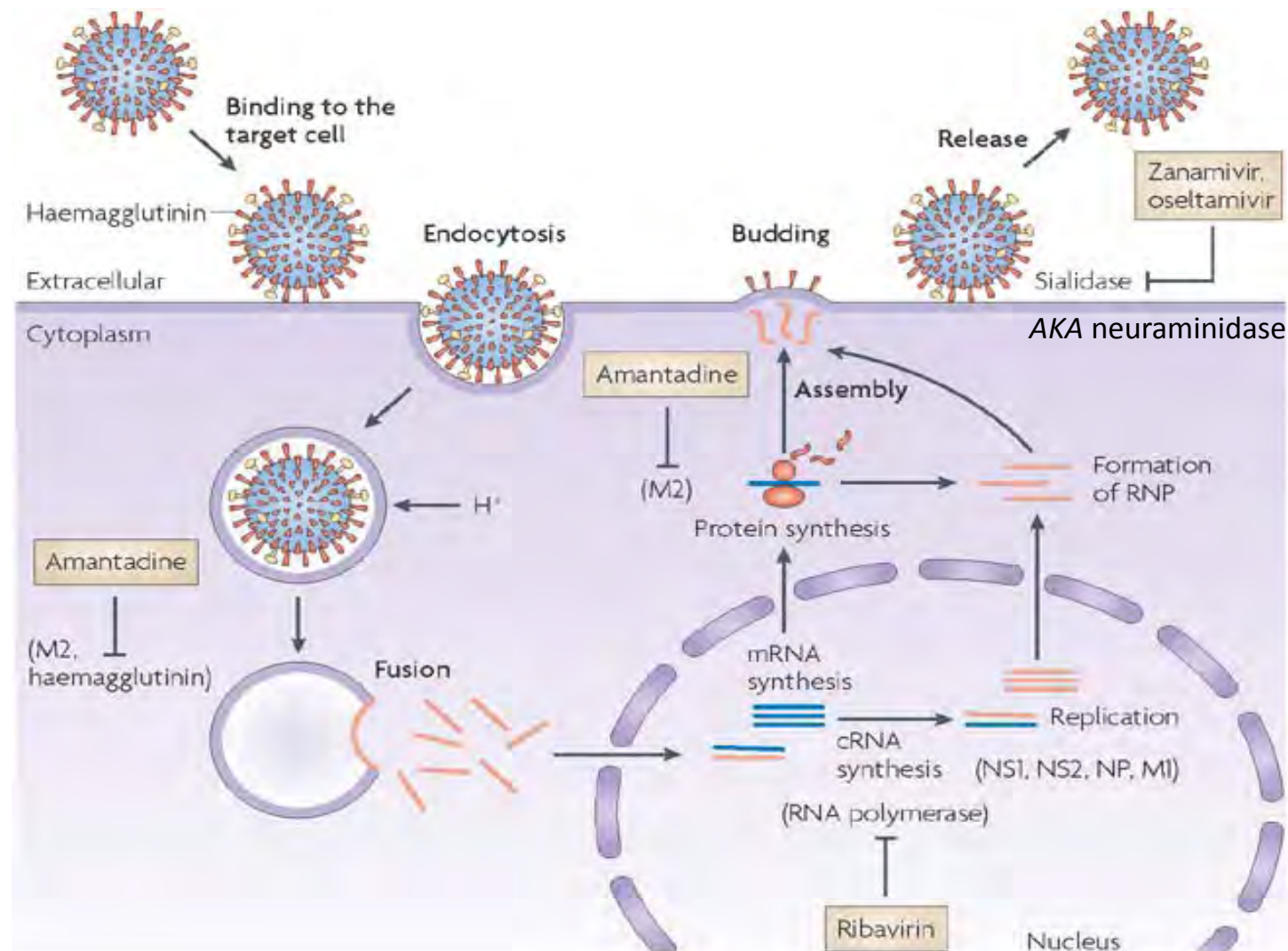


Figure from: von Itzstein, M. (2007). *Nature Reviews Drug Discovery* 6, 967-974

Targeting Influenza Virus RNA-dependent RNA-polymerase

775 fragments in the library

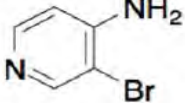
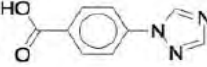
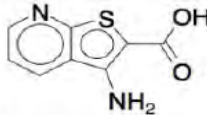
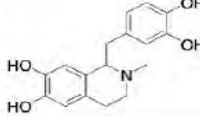
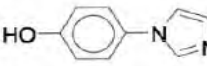
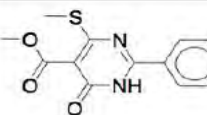
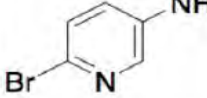
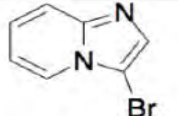
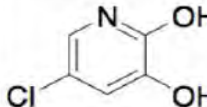
159 fragments are potentially metal chelators

8 fragments from these 159 chelators are found to be a hit in cocktail soaks, verified by individual soaks

1 % hit rate

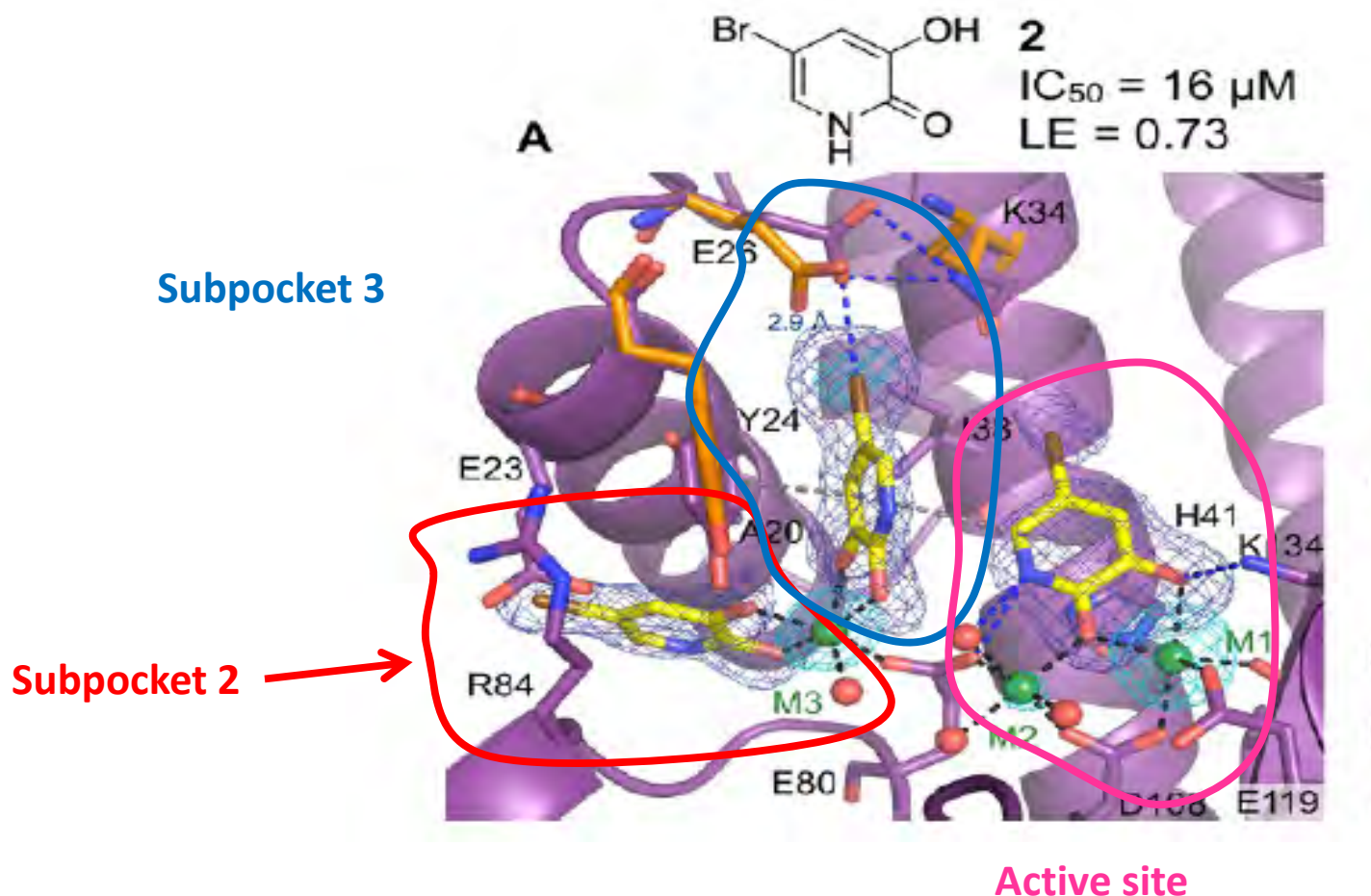
IC₅₀ for each hit were assayed to calculate **ligand efficiency (LE)**

$$LE = \Delta G / N_{\text{heavy}}$$

Compound name	2D-Structure	Binding site	IC ₅₀ (μM)
4-amino-3-bromopyridine		Subpocket 2	>1000
4-(1H-1,2,4-Triazol-1-yl)benzoic acid		Subpocket 2	>1000
3-Aminothiemo[2,3-b]pyridine-2-carboxylic acid		Active site	>1000
D,L-laudanosoline		Active site and subpocket 3	2
4-(1H-imidazol-1-yl)phenol		Subpockets 2 and 3	1000
methyl 4-(methylsulfonyl)-6-oxo-2-phenyl-1,6-dihydropyrimidine-5-carboxylate		Active site and subpocket 3/6	250
6-bromopyridin-3-amine		Subpocket 2	>1000
3-bromoimidazo[1,2-a]pyridine		Outside of active site cleft near Lys73	>1000
5-chloropyridine-2,3-diol		Subpockets 2, 3 and the active site	25

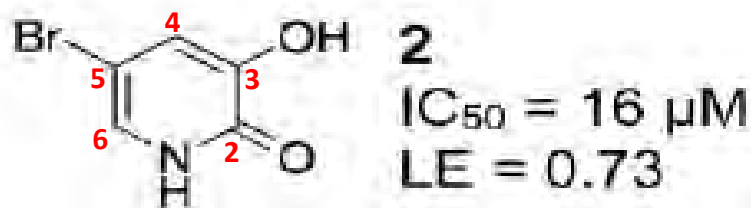
HIT 2 & ANALOGS

Special Length variants of N-terminal endonuclease domain required
Hit 2 is bound three times, and occupies subpockets 2, 3 and active site

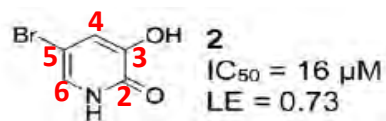


Hit-to-lead development

- Based on Hit 2, design modifications at 4, 5, 6 positions – synthesis – assay – crystal structure



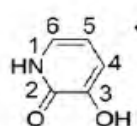
- Substitution at position 4 is compatible with binding predicted by docking, but all substitutions are deleterious experimentally
 - Maybe perturbing the electronic arrangements of chelating oxygens



Binds to subpocket 4, causing side chain rearrangements in subpocket 3

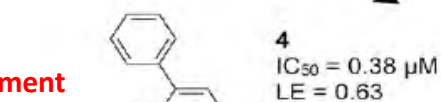
1 molecules bound (unlike 2)

Position 5 points into subpocket 4



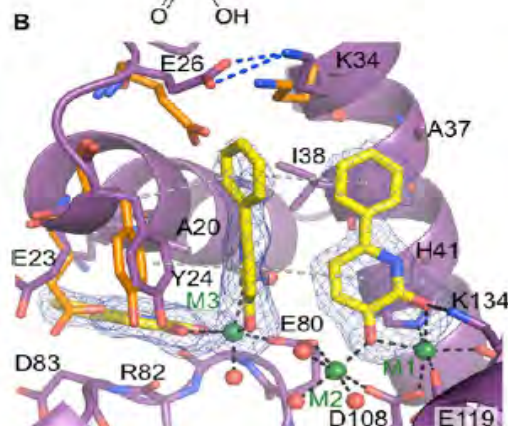
6-substitution

40x improvement

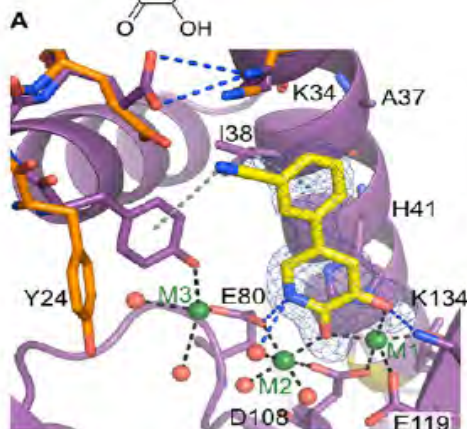
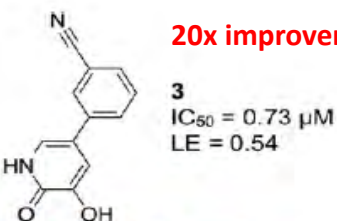


3 molecules bound (like 2 but unlike 3)

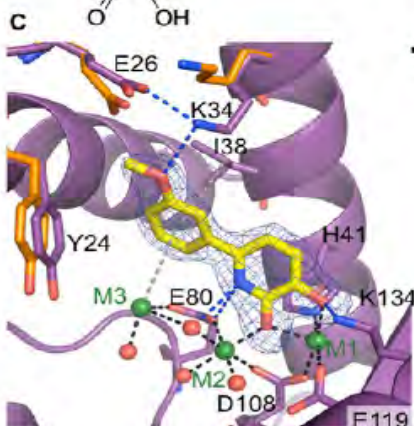
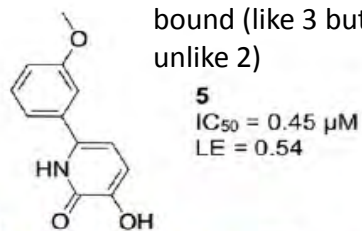
Pyridinone ring flips 180 deg



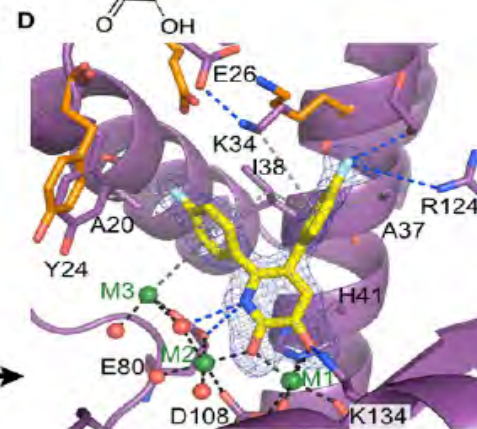
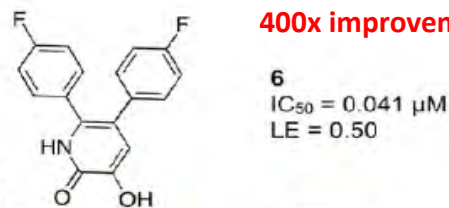
20x improvement



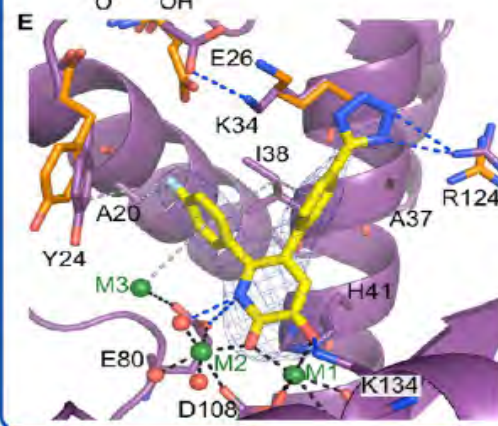
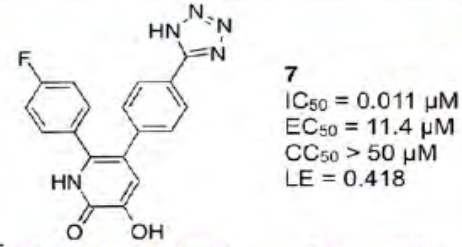
1 molecules bound (like 3 but unlike 2)



400x improvement



1450x improvement



Structural Biology and Drug Development

A marvelous partnership

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

Computational Approaches

An 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

Nature, Science, etc