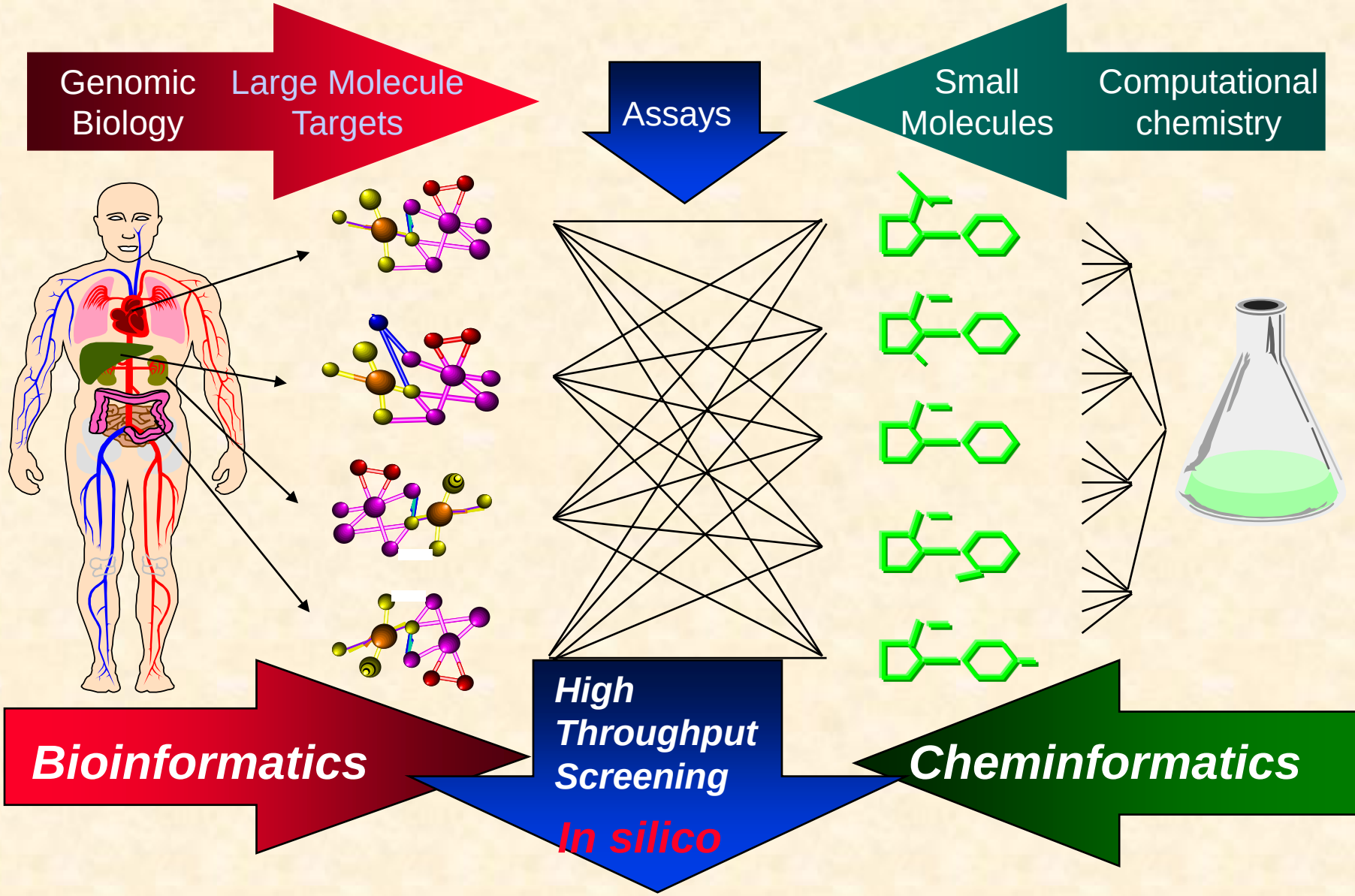
Structure Based Drug Design

G. Narahari Sastry

Molecular Modeling Group
Organic Chemical Sciences
Indian Institute of Chemical Technology
Hyderabad 500 007
INDIA

Seminar on Systems Approach to Bioinformatics, Feb., 18-20, 2004
Bioinformatics centre, Pondicherry University

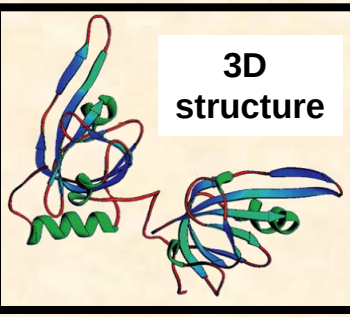
Integration of Chemoinformatics and Bioinformatics

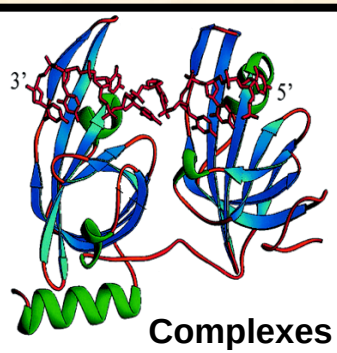


Biological Structure

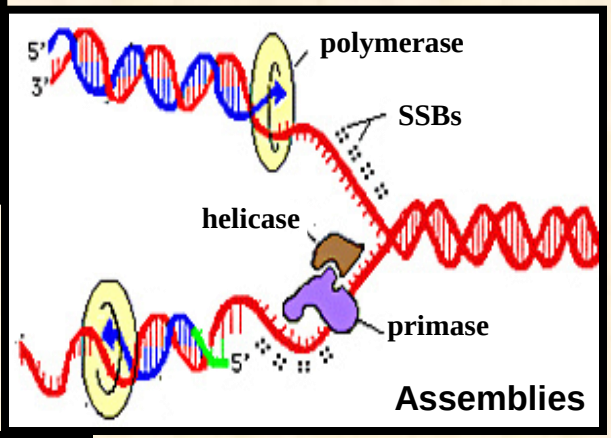
Sequence

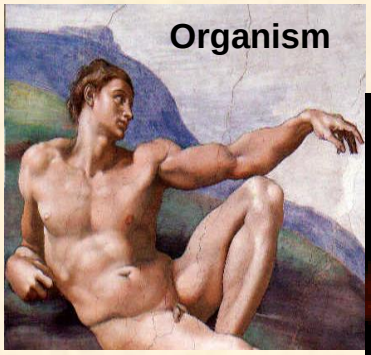
MESDAMESETMESSRSMYN
AMEISWALTERYALLKINCAL
LMEWALLYIPREFERDREVIL
MYSELFIMACENTERDIRATV
ANDYINTENNESSEEILIKENM
RANDDYNAMICSRPADNAPRI
MASERADCALCYCLINNDRKI
NASEMRPCALTRACTINKAR
KICIPCDPKIQDENVSDETAVS
WILLWINITALL

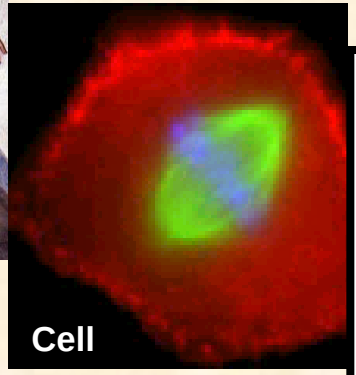


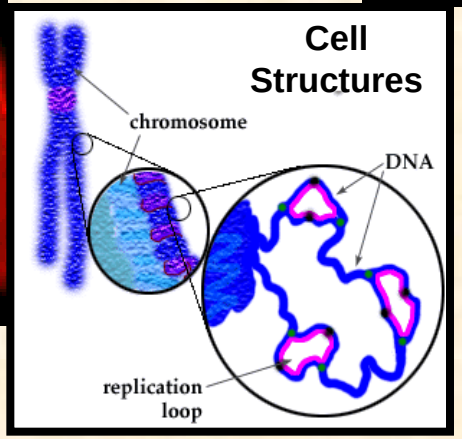


Structural Scales









System Dynamics

Much About Structure

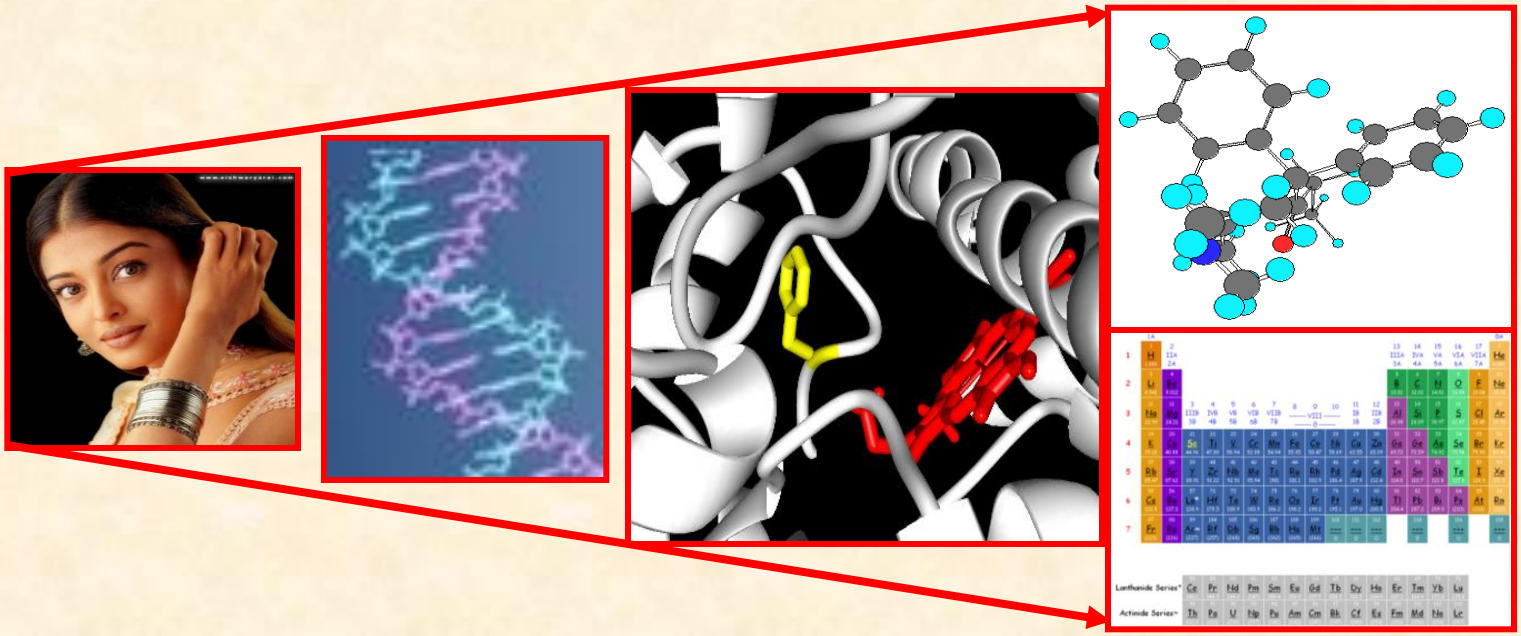
- Structure ⟷ Function
- Structure ⟷ Mechanism
- Structure ⟷ Origins/Evolution
- Structure-based Drug Design

Bottlenecks in developing Structure – Function Relationships

- Structures determined by NMR, computation, or X-ray crystallography are static snapshots of highly dynamic molecular systems
- Biological processes (recognition, interaction, chemistry) require molecular motions and are time dependent.
- To comprehend and facilitate thinking about the dynamic structure of molecules is crucial.

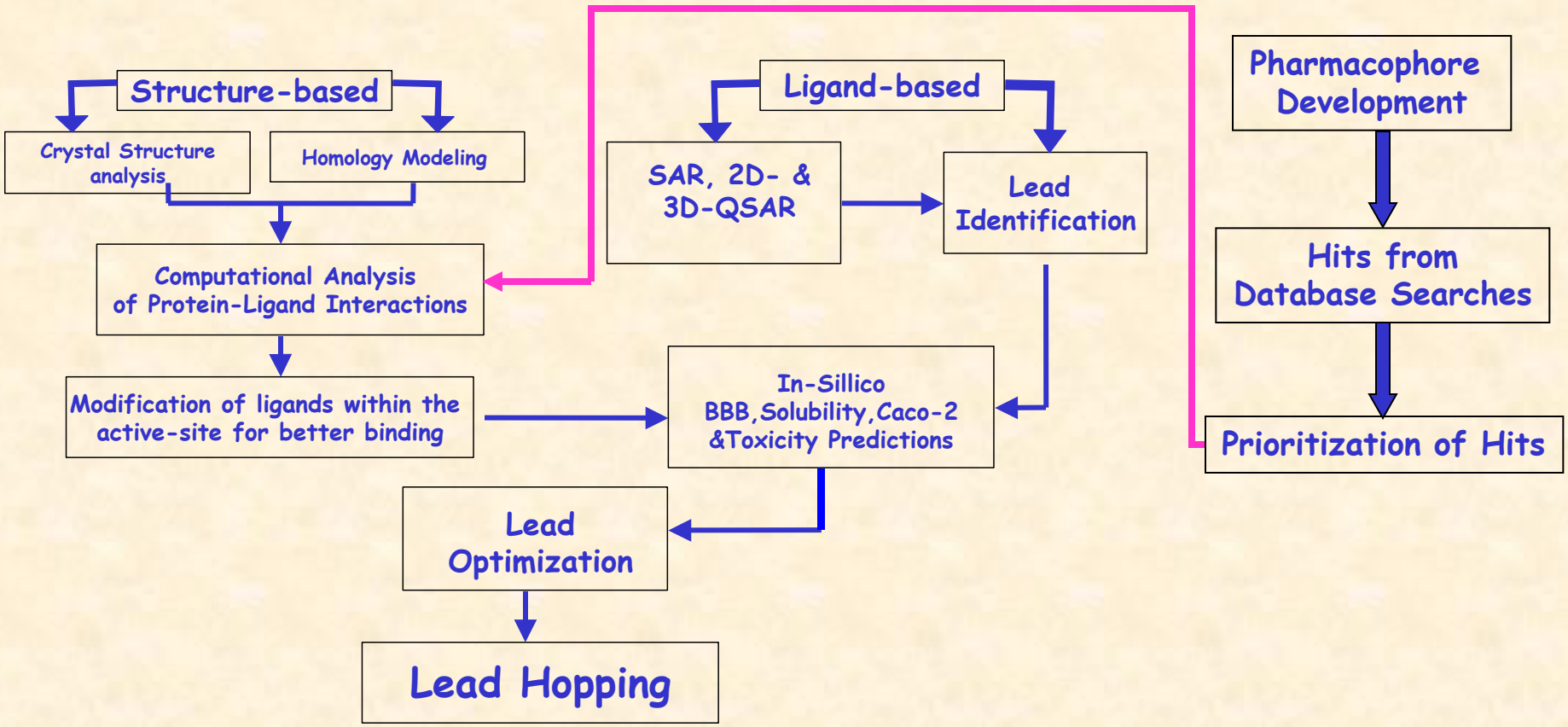
What is Molecular Modeling?

- A science that elucidates and validates experimental evidence through imagination, visualization, and rationalization
- Applied in many areas of research (Academic/Industrial)



Caveat: Is the interpolation and extrapolation reliable?

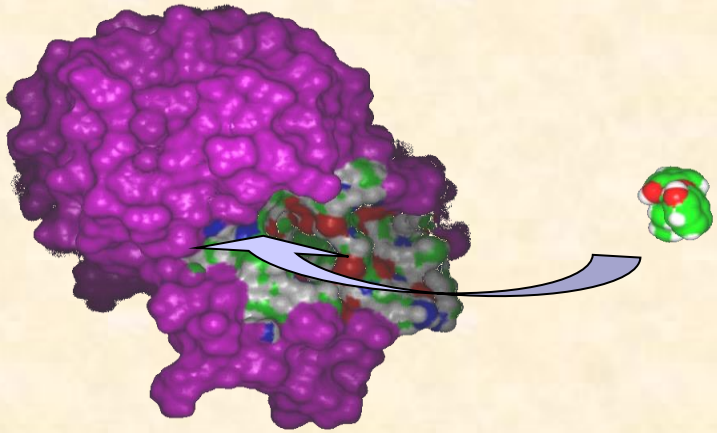
Capabilities of Molecular Modeling

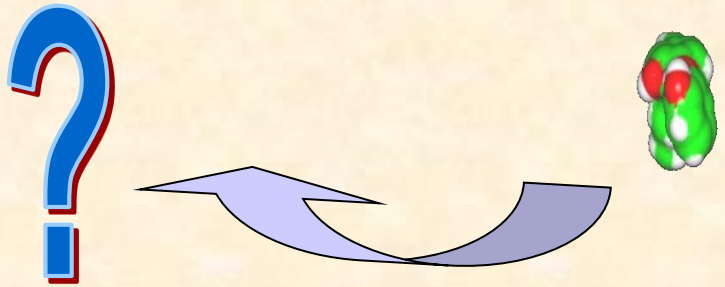


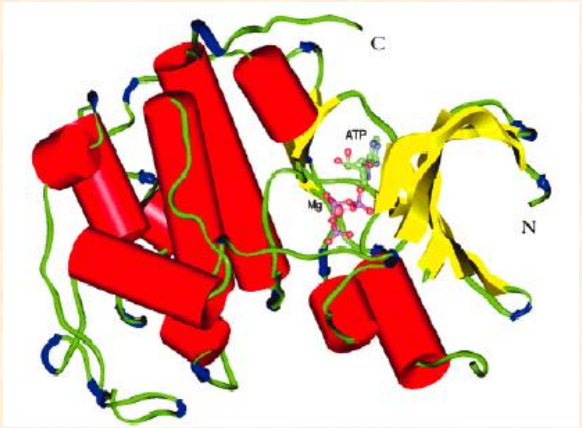
Drug Design

Structure based

Ligand based

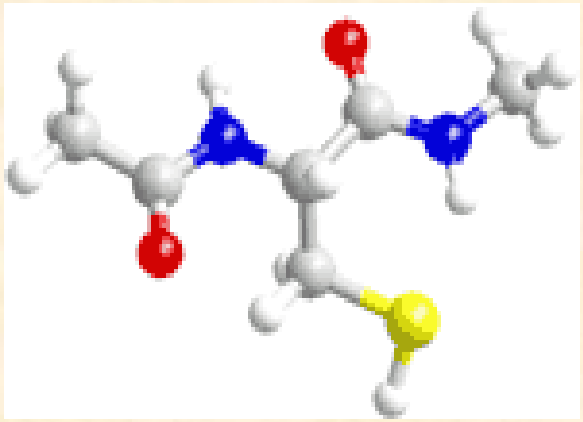






Target (structure) based drug design

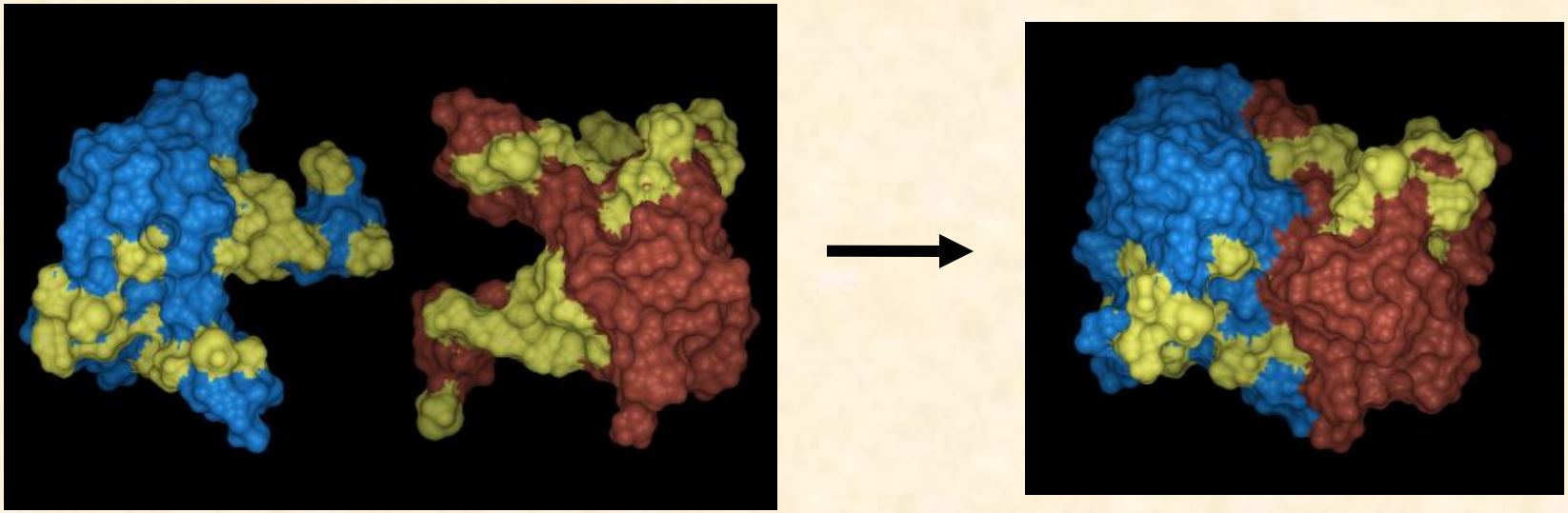
- Receptor structure is known
- Mechanism is known
- Ligands and their biological activities are known/ unknown



Ligand (analog) based drug design

- Receptor structure is not known
- Mechanism is known/ unknown
- Ligands and their biological activities are known

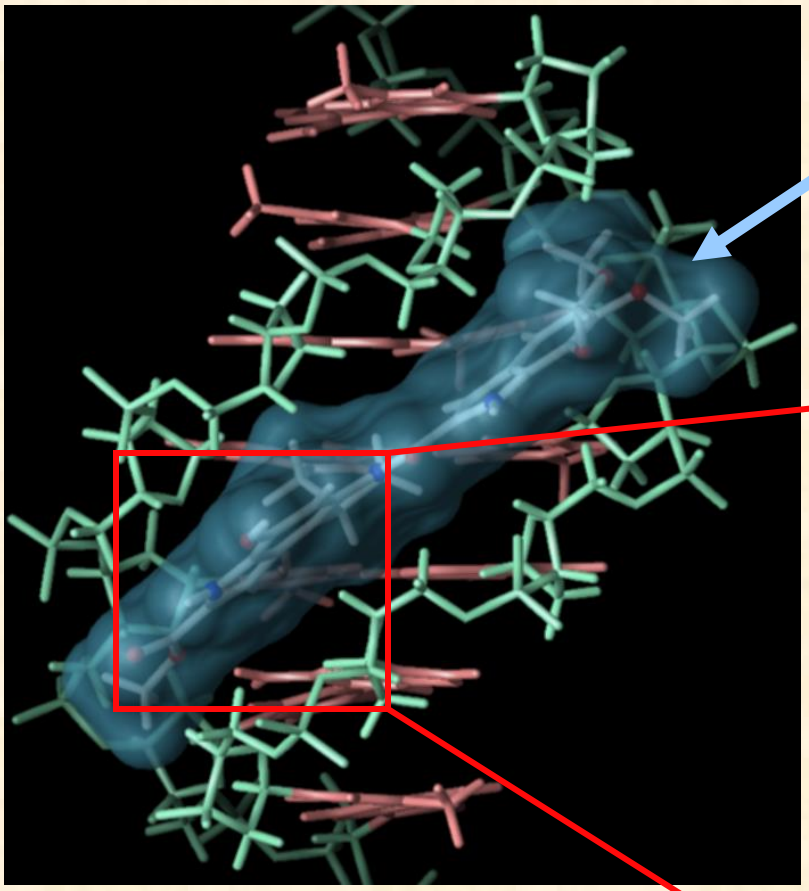
High Resolution Structural Biology



Determine atomic structure

Analyze why molecules interact

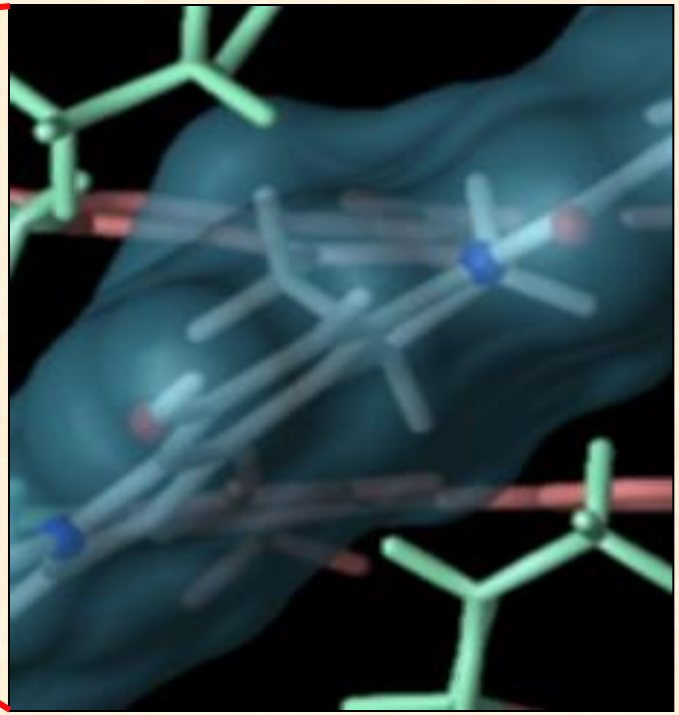
The Reward: Understanding→Control



Anti-tumor activity
Duocarmycin SA

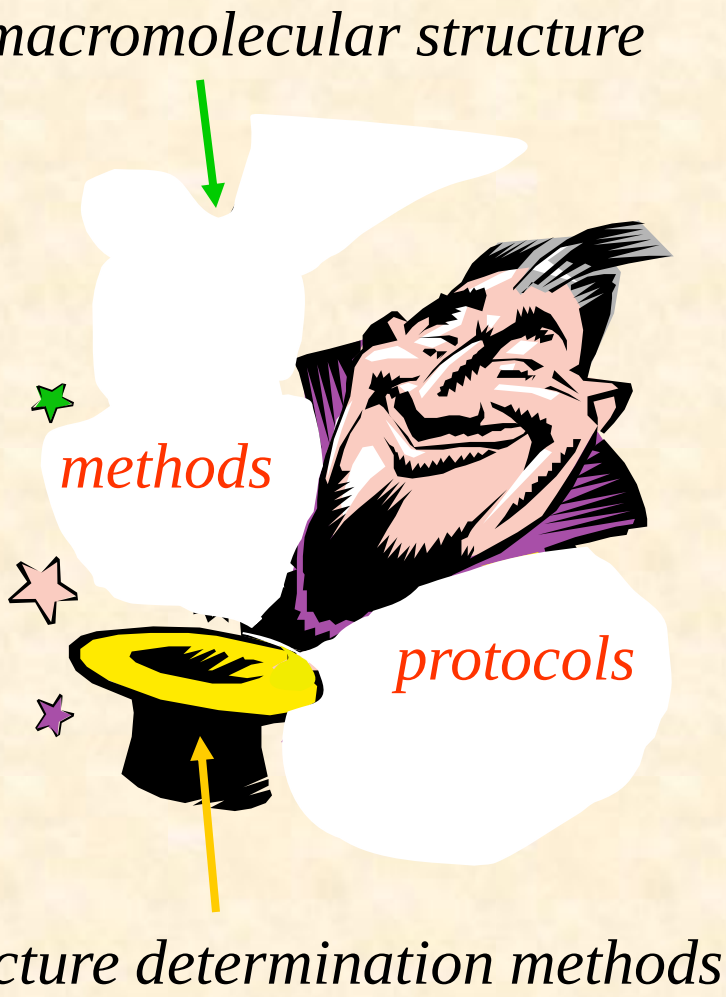
Shape

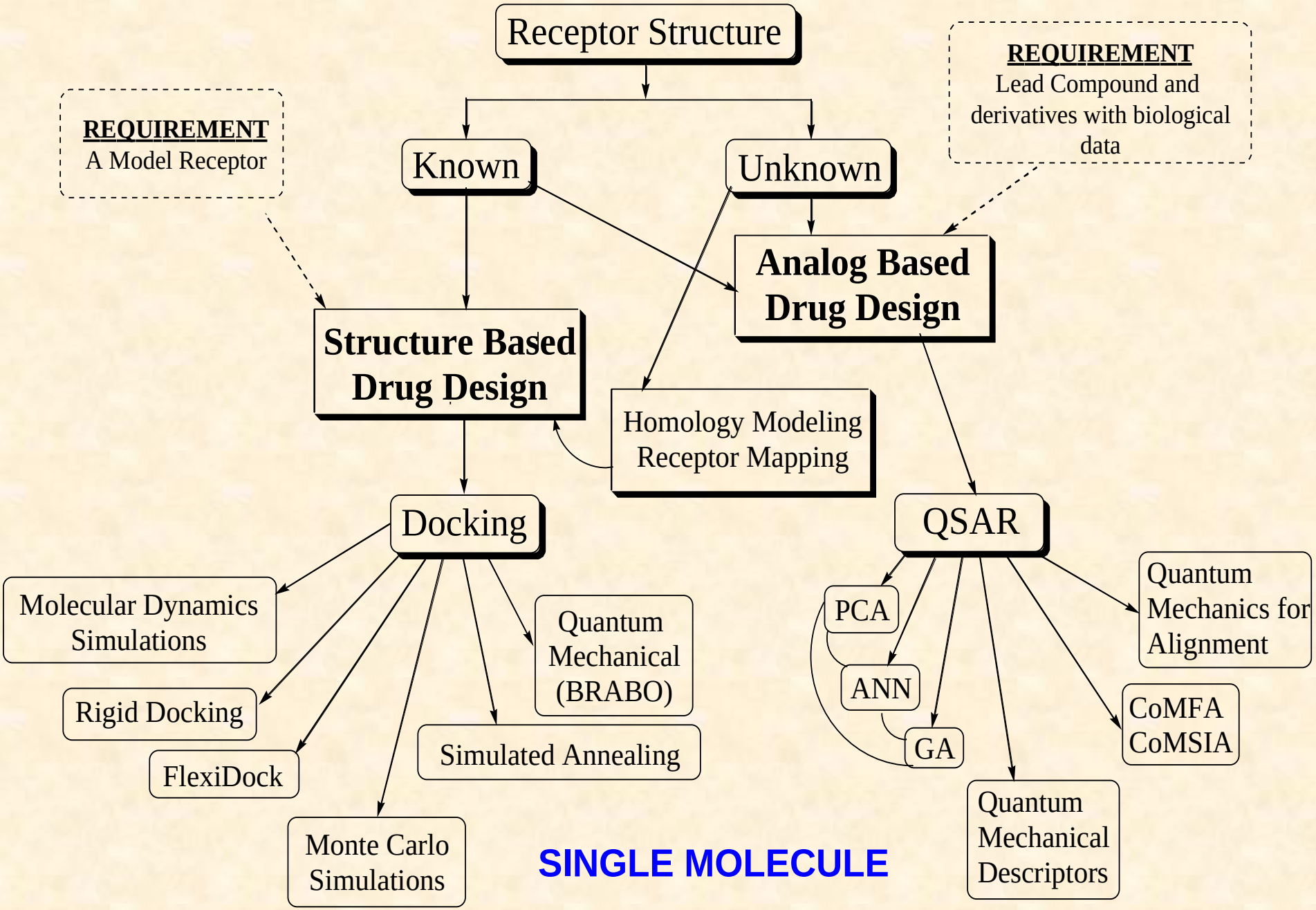
Atomic interactions



CAUTION….

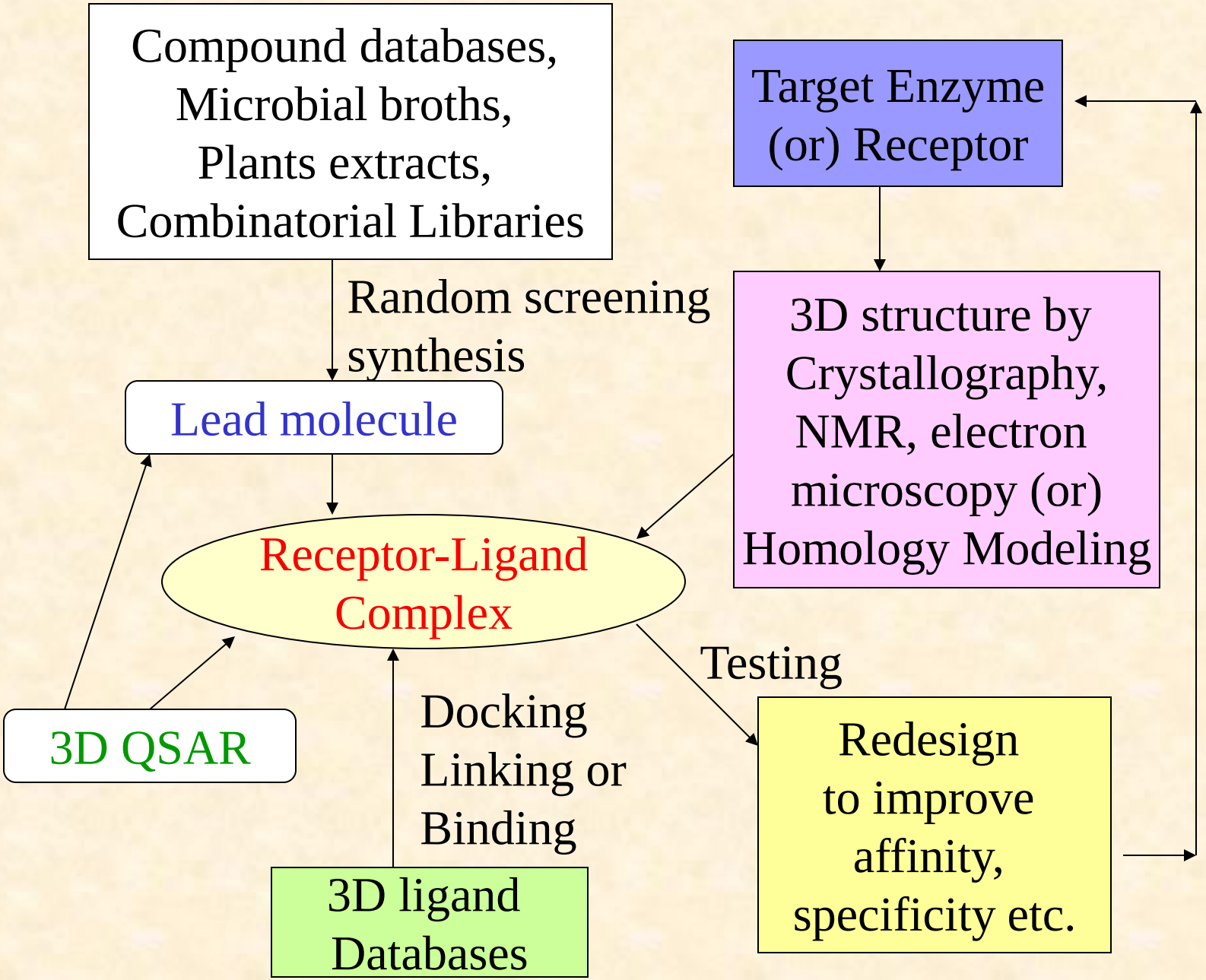
- Don't be a naive user!?!
  - When computers are applied to biology, it is vital to understand the difference between mathematical & biological significance
  - computers don't do biology, they do sums quickly





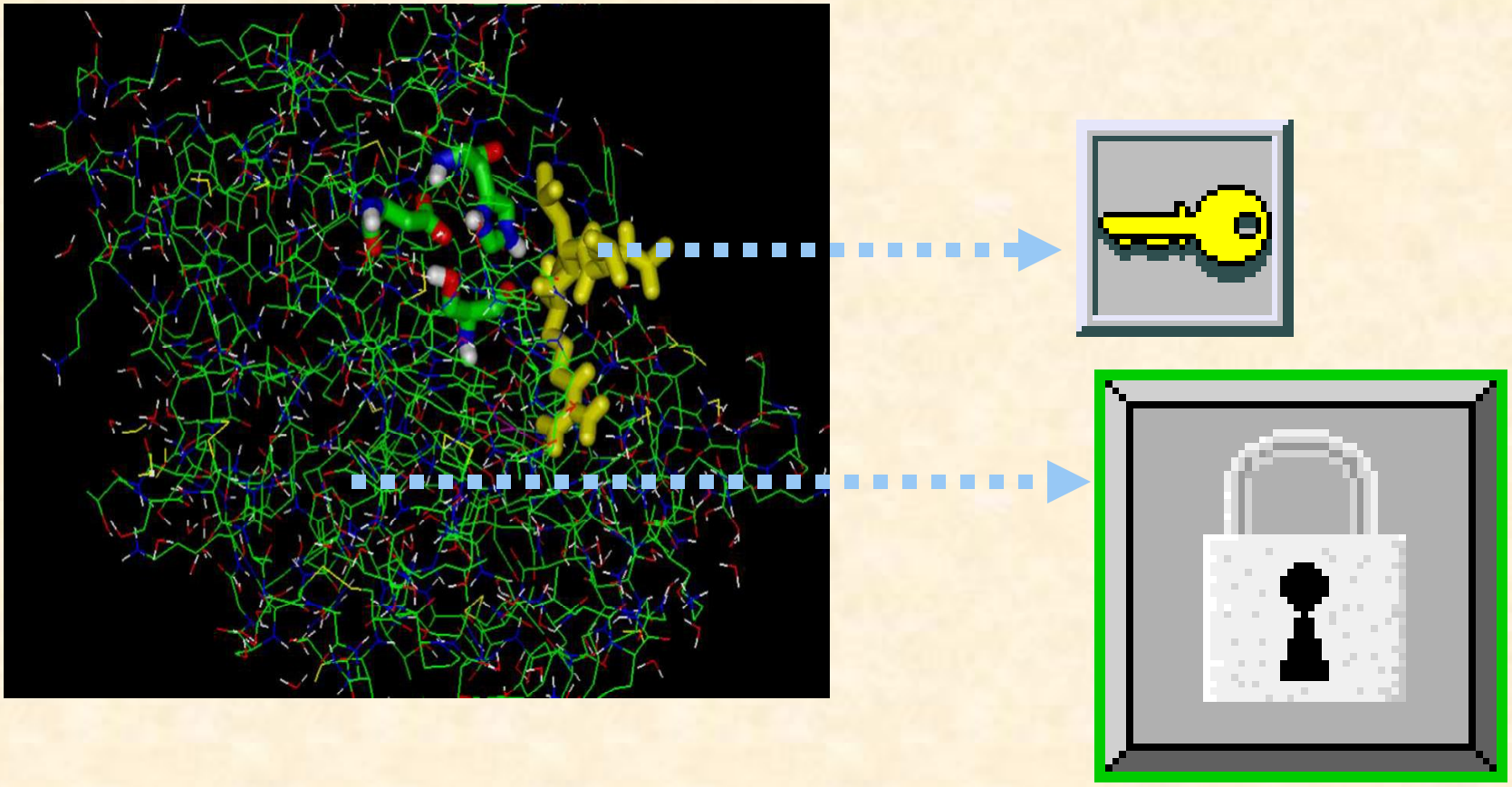
SYBYL, INSIGHT II, CERIUS^2, MOE, AMBER (CDAC), DOCK, AUTODOCK

Structure Based Drug Design



Drug and Target : Lock and Key ?

Most of the drugs “FIT” well to their targets

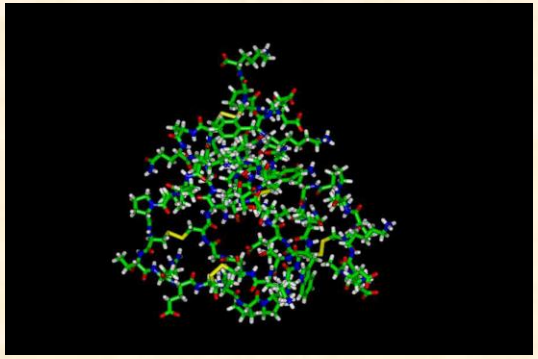


Some “Locks” are known but not all !!

Study of protein crystals give the details of the “lock”.
Knowing the “lock” structure, we can DESIGN some “keys”.

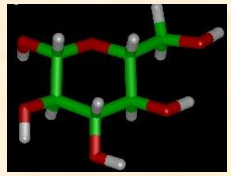
This is achieved by COMPUTER Algorithms

This is called “STRUCTURE BASED DRUG DESIGN”



“Lock” structure (from experiment)

Algorithms



“Key”constructed by computer

Variations on the Lock and Key Model

1- Which structure of the lock should be targeted?

2- Is the binding pocket a good target?

3- Is structure-based design relevant for my receptor?

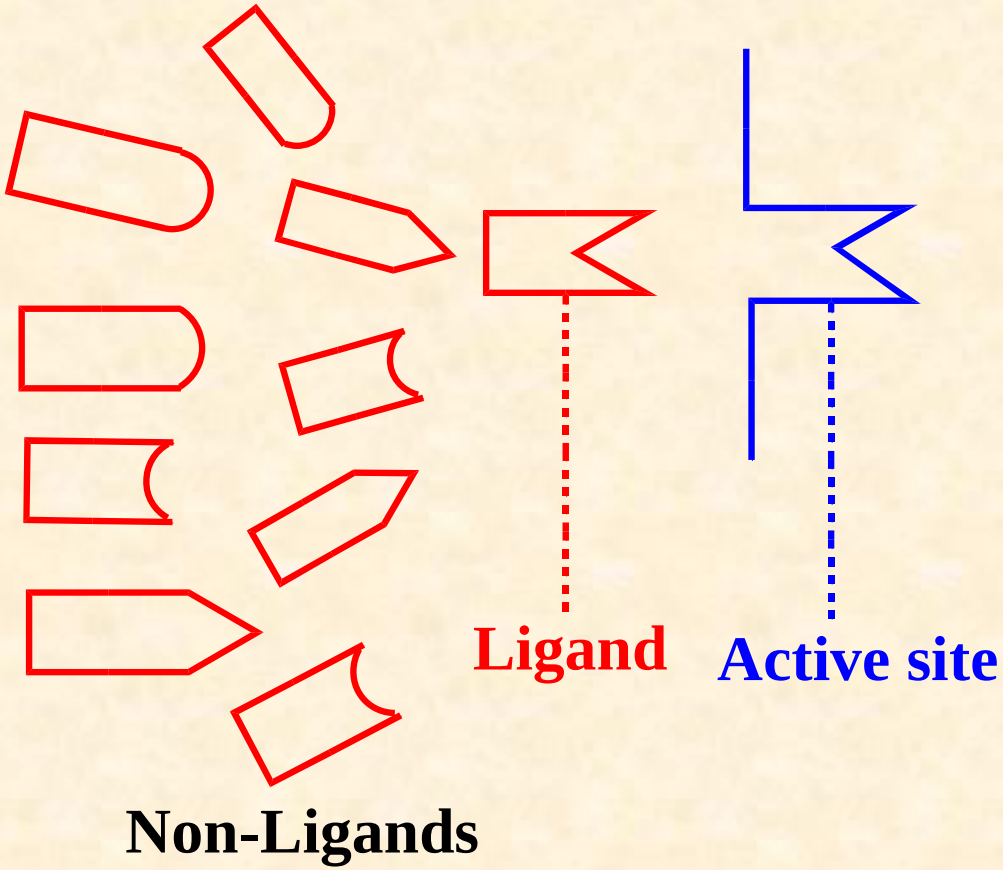
-Is the 3D structure reliable?

-Is the binding pocket static enough?

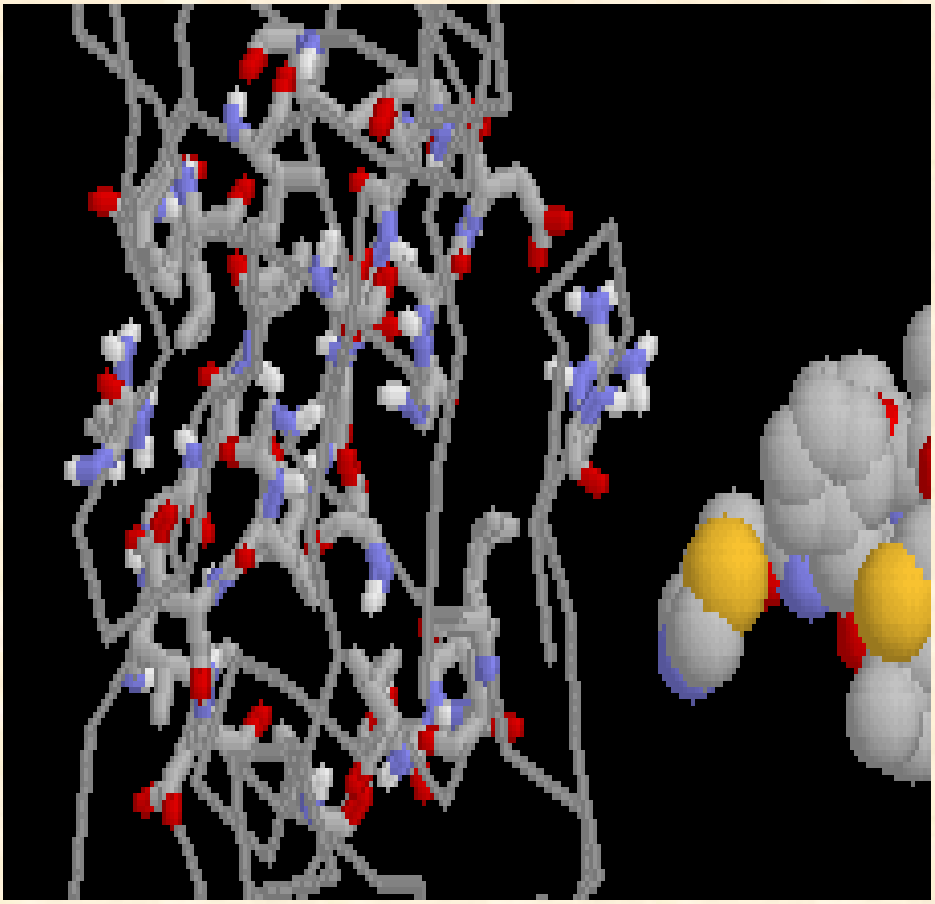
4- Which key fits best?

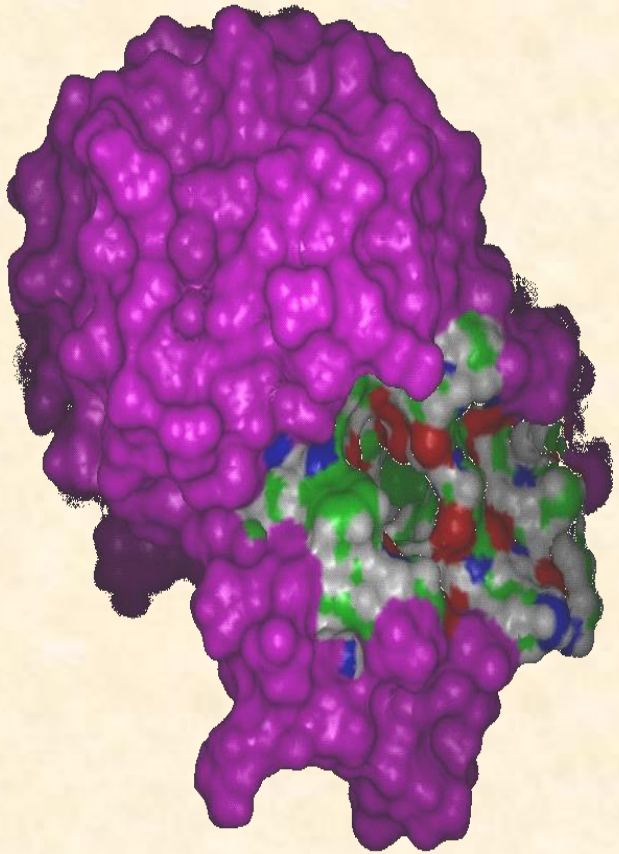
5- What are the prerequisite physicochemical properties for the key for better binding?

The ligand has been identified

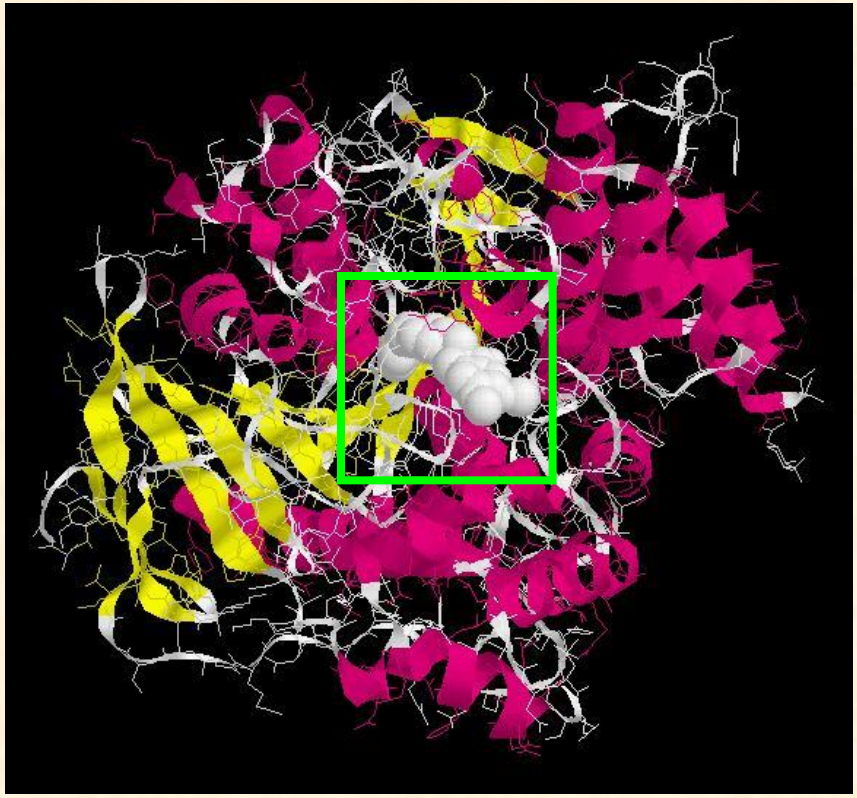


Docking of Ligand to the Active site of Protein





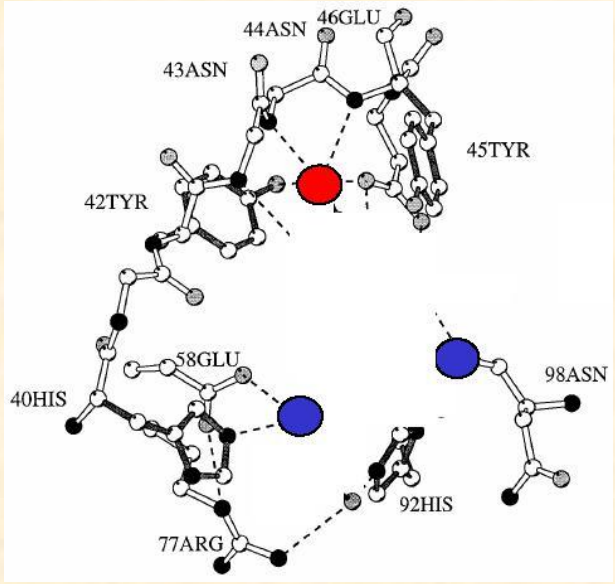
3D Structure of the Complex



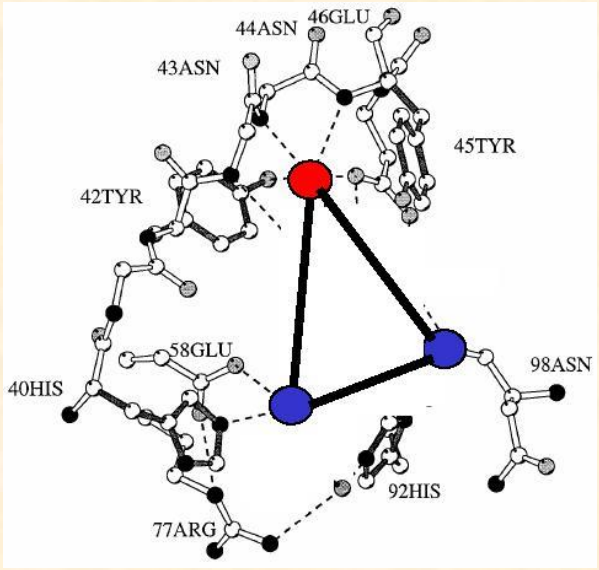
Experimental Information: The active site can be identified based on the position of the ligand in the crystal structures of the protein-ligand complexes

If Active Site is not KNOWN?????

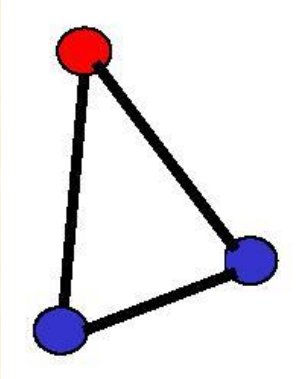
Building Molecules at the Binding Site



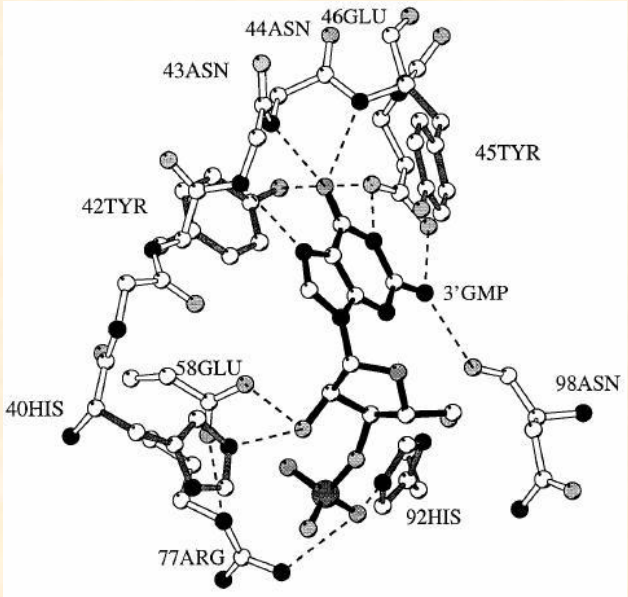
Identify the binding regions



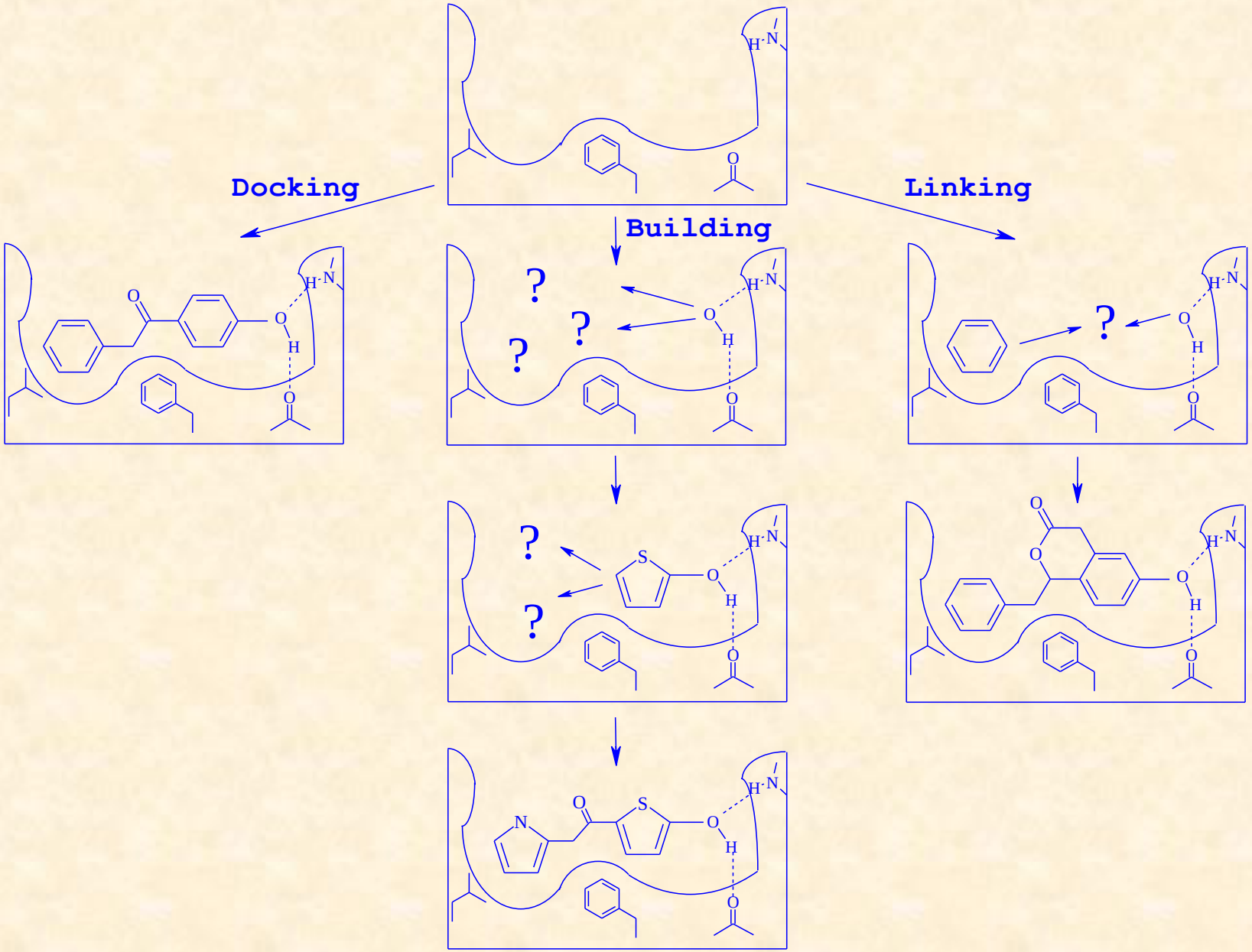
Evaluate their disposition in space



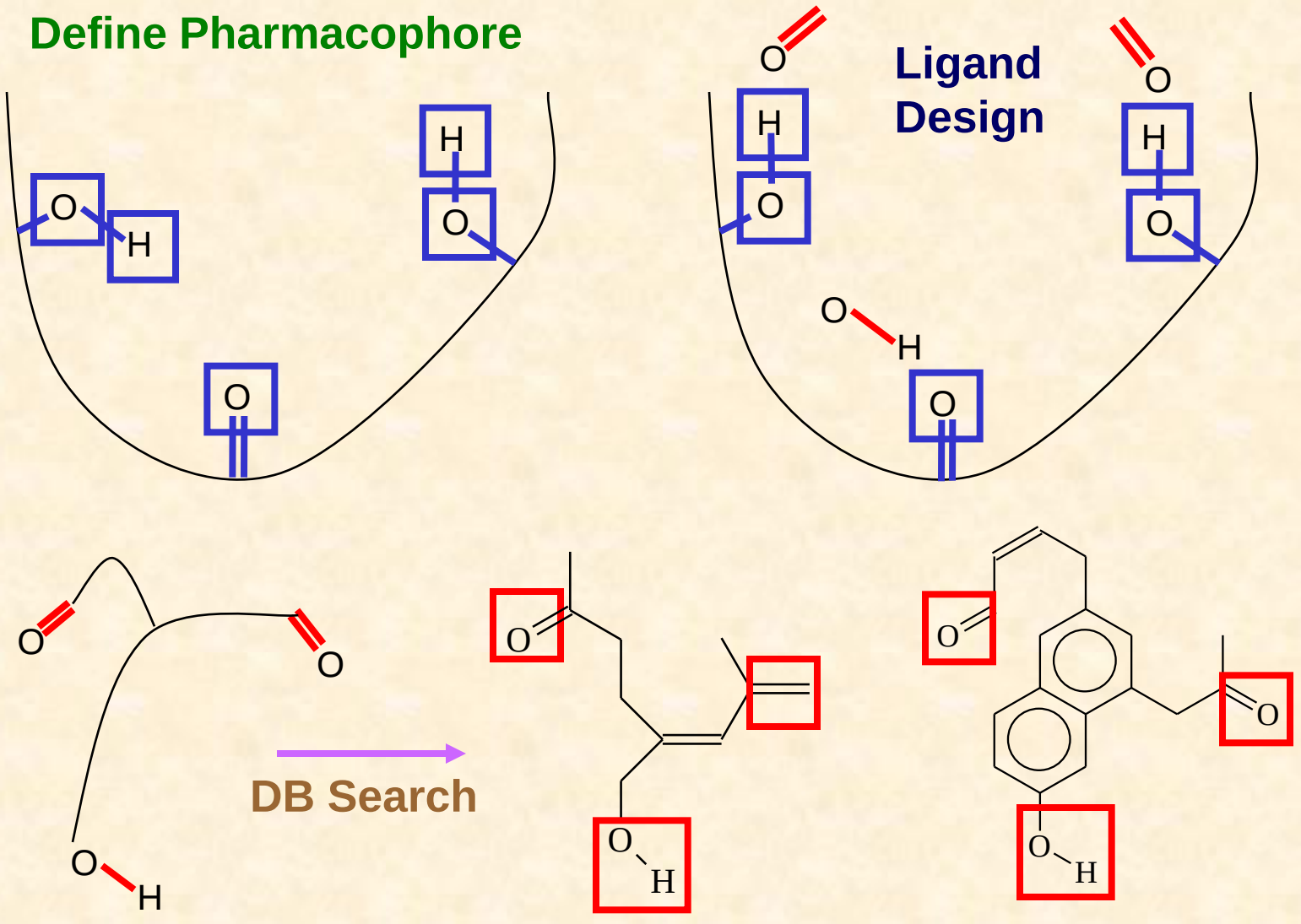
Search for molecules in the library of ligands for similarity

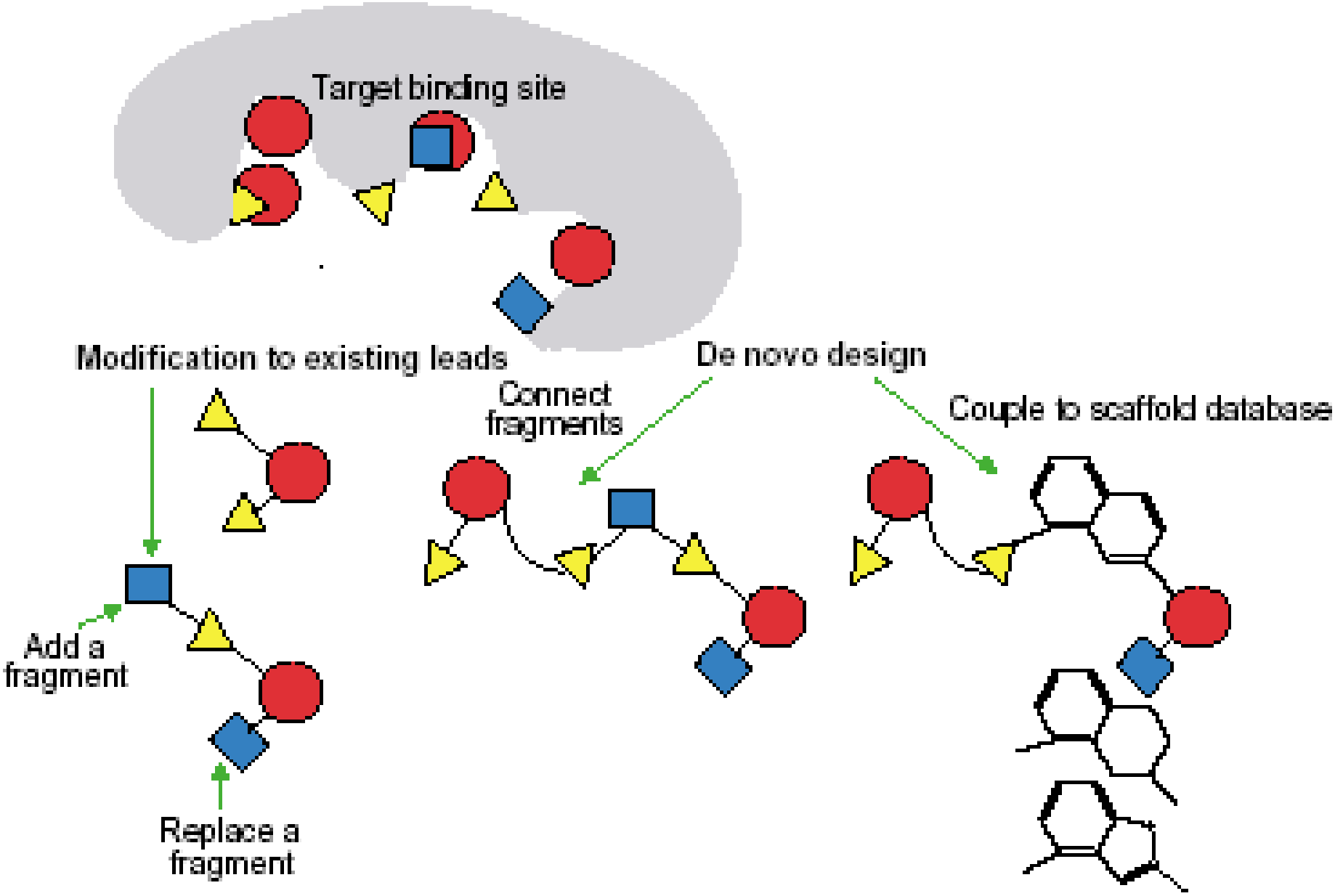


Structure Based Ligand Design



Structure based drug design




Target binding site
Modification to existing leads
De novo design
Connect fragments
Couple to scaffold database
Add a fragment
Replace a fragment

Molecular Docking

- The process of “docking” a ligand to a binding site mimics the natural course of interaction of the ligand and its receptor via a lowest energy pathway.
- Put a compound in the approximate area where binding occurs and evaluate the following:
  - Do the molecules bind to each other?
  - If yes, how strong is the binding?
  - How does the molecule (or) the protein-ligand complex look like. (understand the intermolecular interactions)
  - Quantify the extent of binding.

Molecular Docking (contd…)

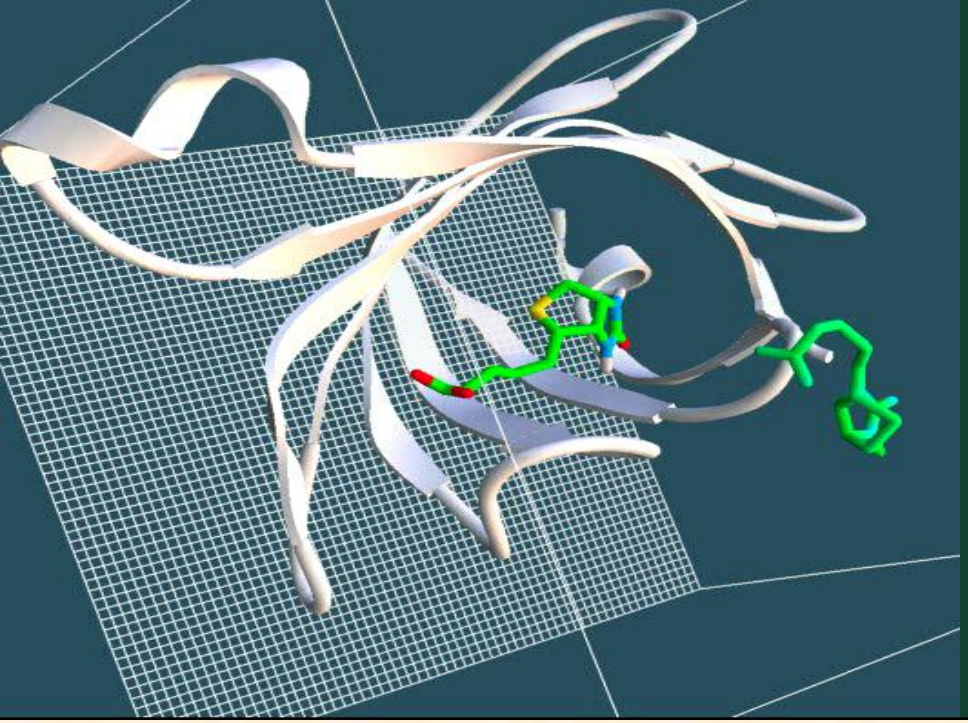
- Computationally predict the structures of protein-ligand complexes from their conformations and orientations.
- The orientation that maximizes the interaction reveals the most accurate structure of the complex.
- The first approximation is to allow the substrate to do a random walk in the space around the protein to find the lowest energy.

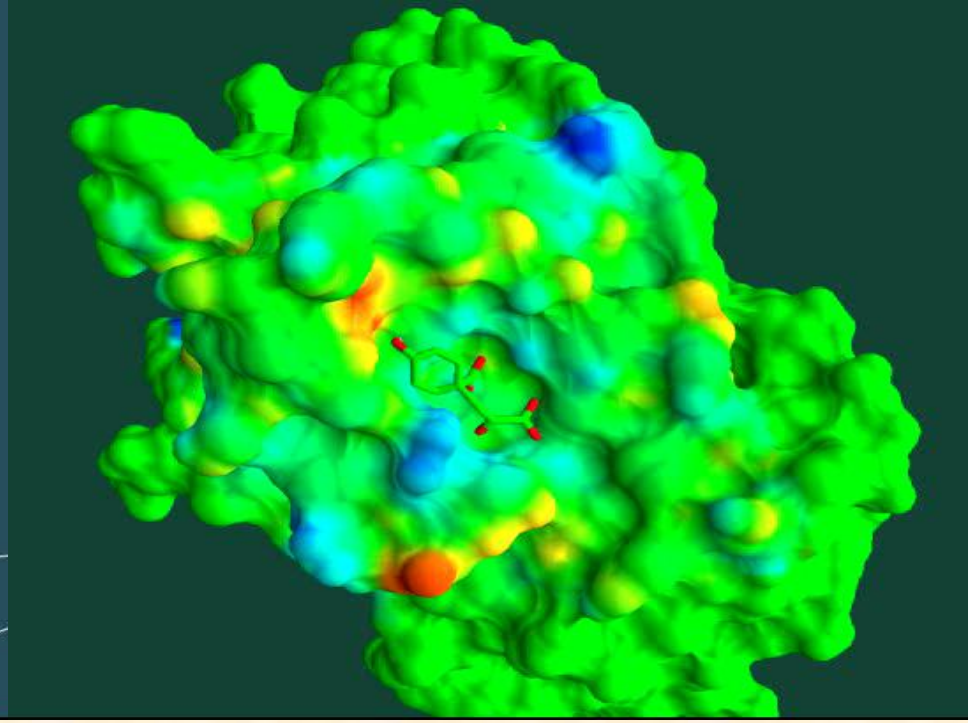
Algorithms used while docking

- Fast shape matching (e.g., DOCK and Eudock),
- Incremental construction (e.g., FlexX, Hammerhead, and SLIDE),
- Tabu search (e.g., PRO_LEADS and SFDock),
- Genetic algorithms (e.g., GOLD, AutoDock, and Gambler),
- Monte Carlo simulations (e.g., MCDock and QXP),

Some Available Programs to Perform Docking

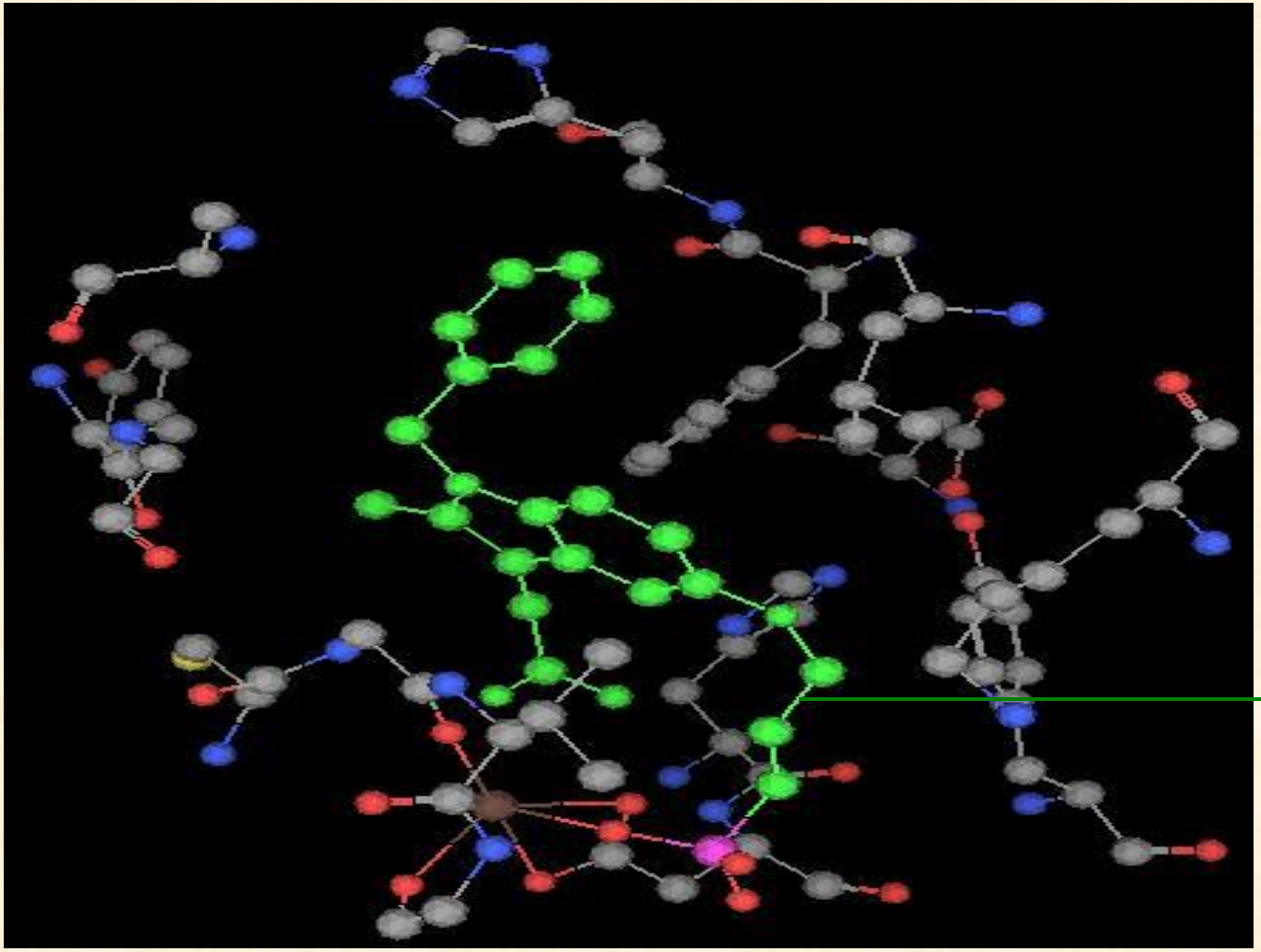
- Affinity
- AutoDock
- BioMedCAChe
- CAChe for Medicinal Chemists
- DOCK
- DockVision
- FlexX
- Glide
- GOLD
- Hammerhead
- PRO_LEADS
- SLIDE
- VRDD





Different views of Docking

Ligand in Active Site Region

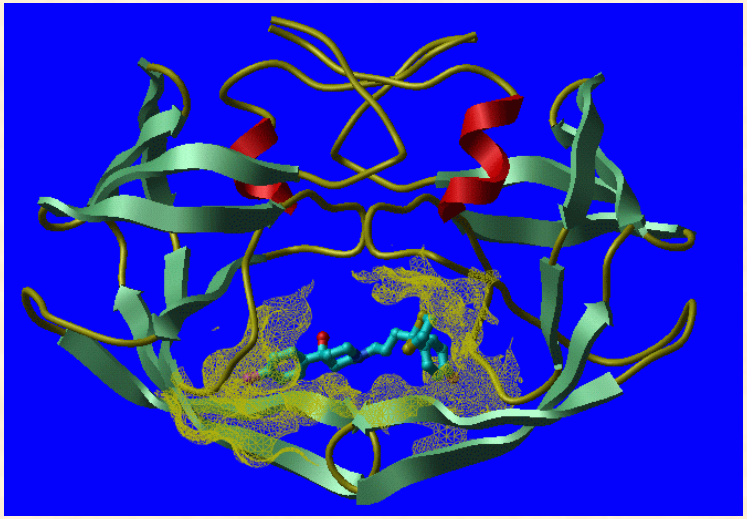


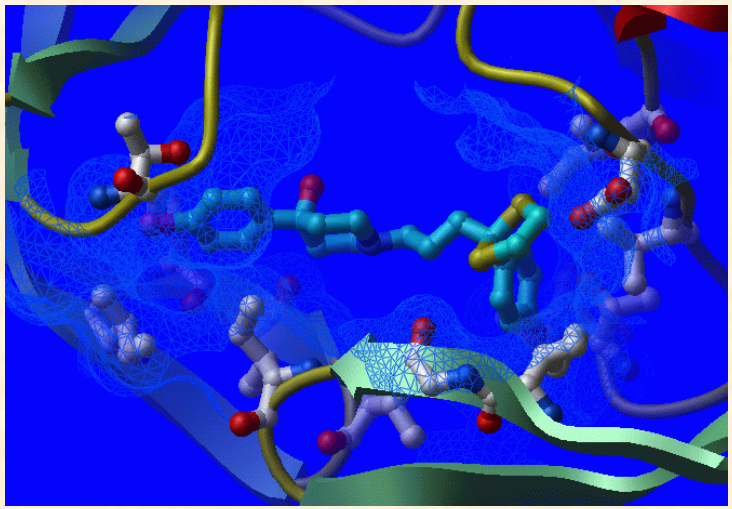
Active site residues

Histidine 6; Phenylalanine 5; Tyrosine 21; Aspartic acid 91; Aspartic acid 48; Tyrosine 51; Histidine 47; Glycine 29; Leucine 2; Glycine 31; Glycine 22; Alanine 18; Cysteine 28; Valine 20; Lysine 62

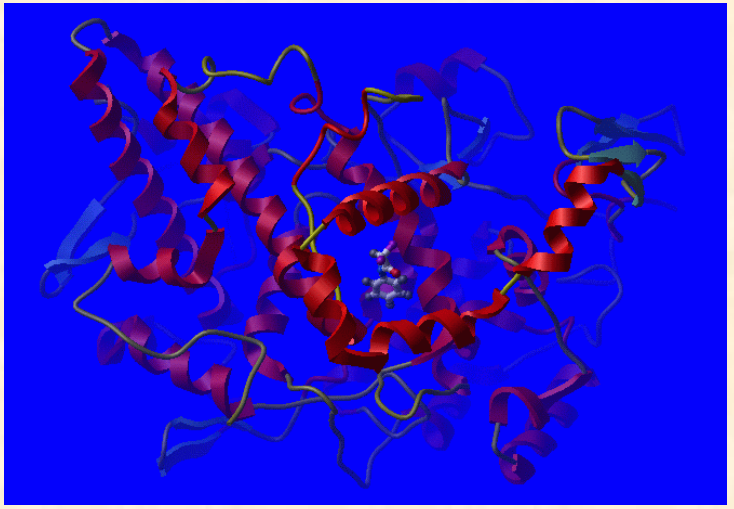
Examples of Docked structures

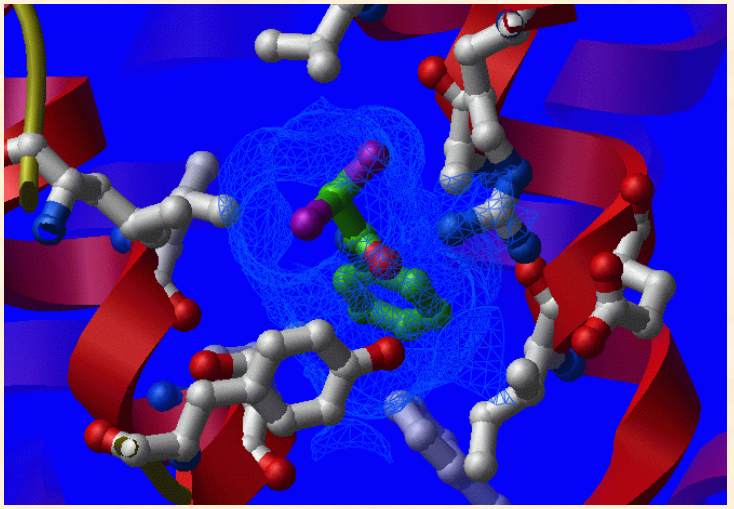
HIV protease inhibitors





COX2 inhibitors



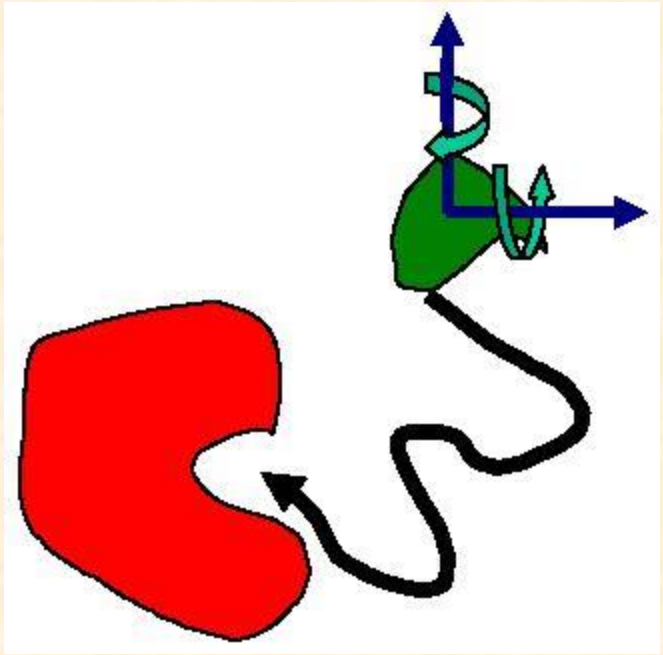


Approaches to Docking

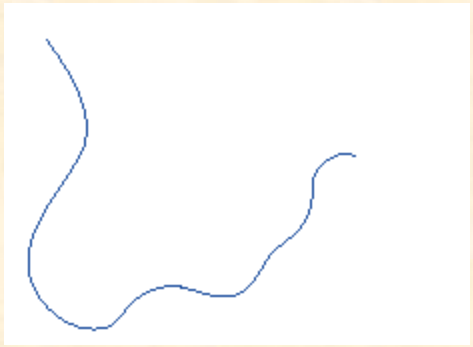
- Qualitative
  - Geometric
  - shape complementarity and fitting
- Quantitative
  - Energy Calculations
  - determine minimum energy structures
  - free energy measure
- Hybrid
  - Geometric and energy complementarity
  - 2 phase process: rigid and flexible docking

Rigid Docking

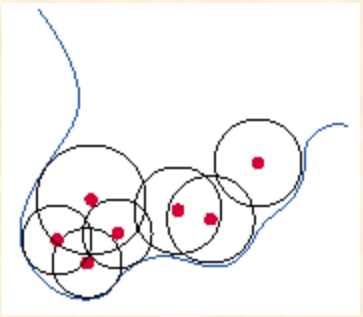
- Shape-complementarity method: find binding mode(s) without any steric clashes
- Only 6-degrees of freedom (translations and rotations)
- Move ligand to binding site and monitor the decrease in the energy
- Only non-bonded terms remain in the energy term
- Try to find a good steric match between ligand and receptor



- Describe binding site as set of overlapping spheres



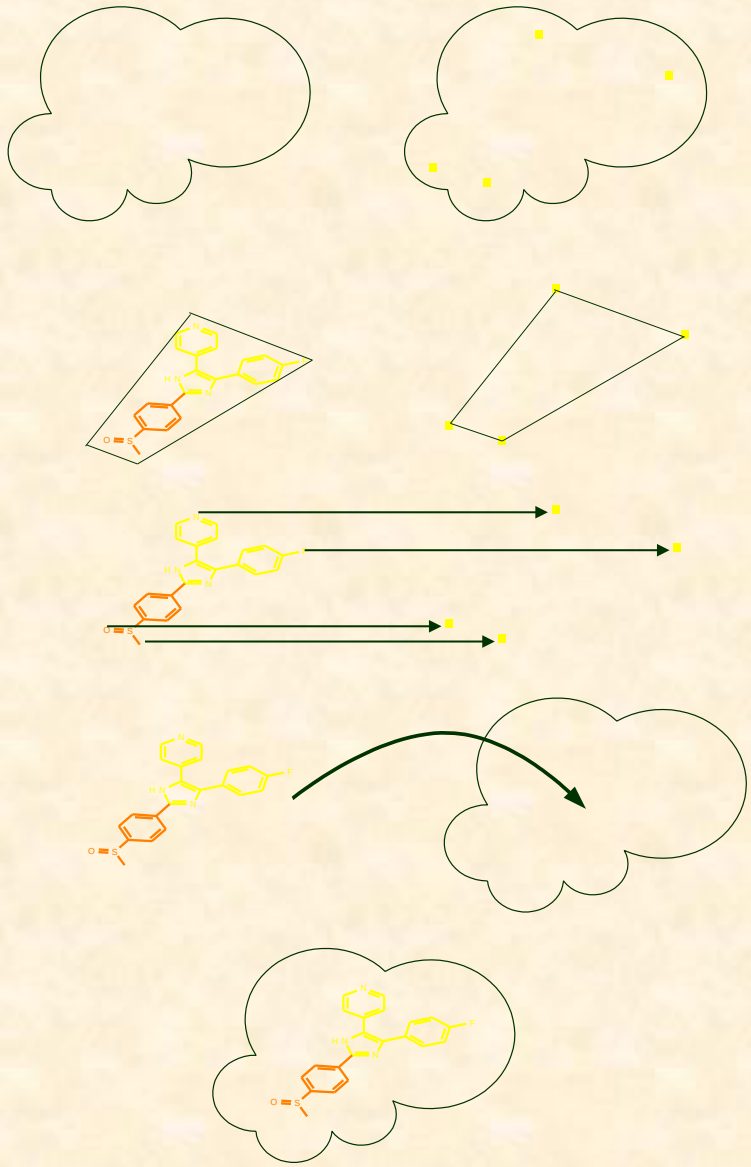
binding site



overlapping spheres

- Both the macromolecule and the ligand are kept rigid; the ligand is allowed to rotate and translate in space
- In reality, the conformation of the ligand when bound to the complex will not be a minima.

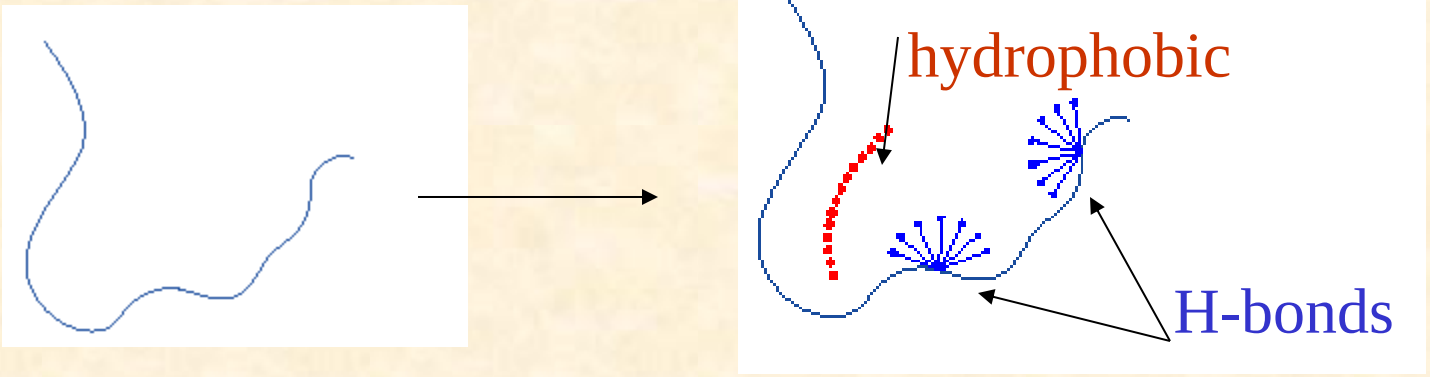
The DOCK algorithm in rigid-ligand mode



1. Define the target binding site points.
2. Match the distances.
3. Calculate the transformation matrix for the orientation.
4. Dock the molecule.
5. Score the fit.

Flexible Docking

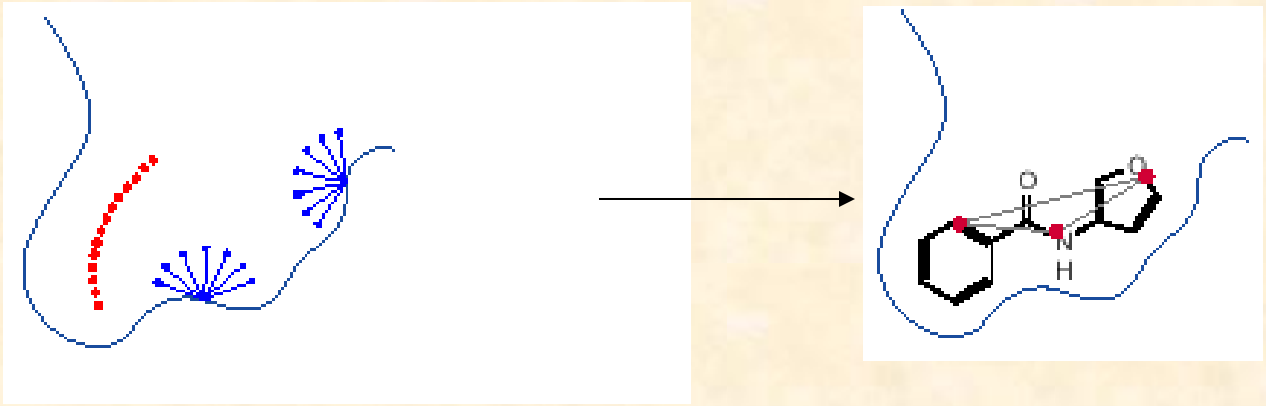
- Dock flexible ligands into binding pocket of rigid protein
- Binding site broken down into regions of possible interactions



binding site
from X-ray

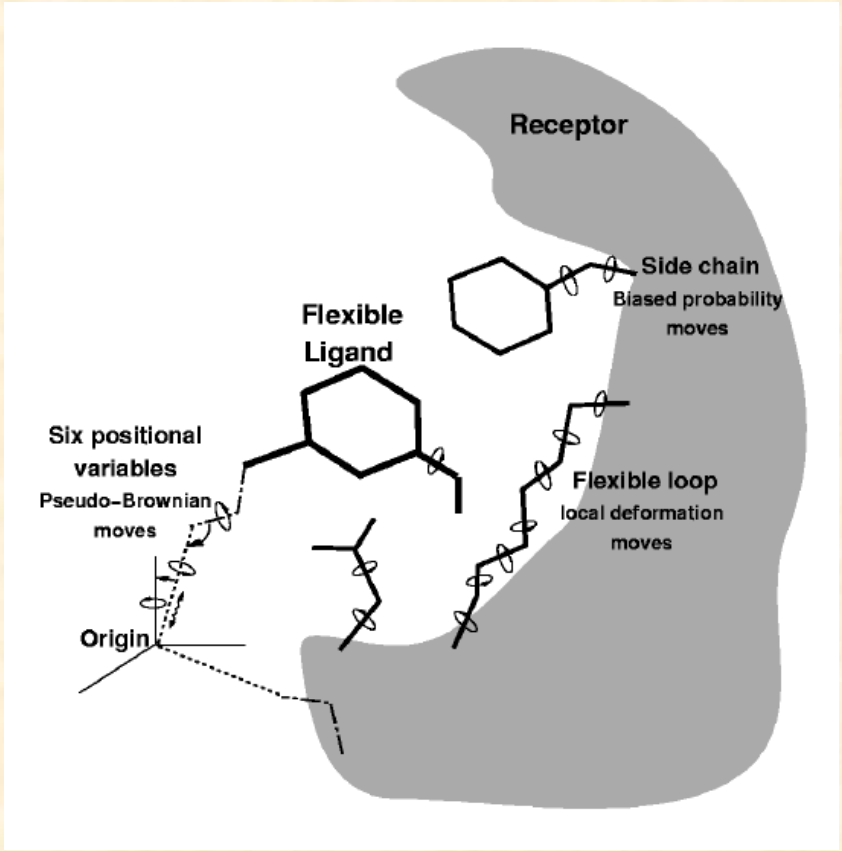
parameterised
binding site

- Then dock the molecule into pocket by matching up interactions with ligand



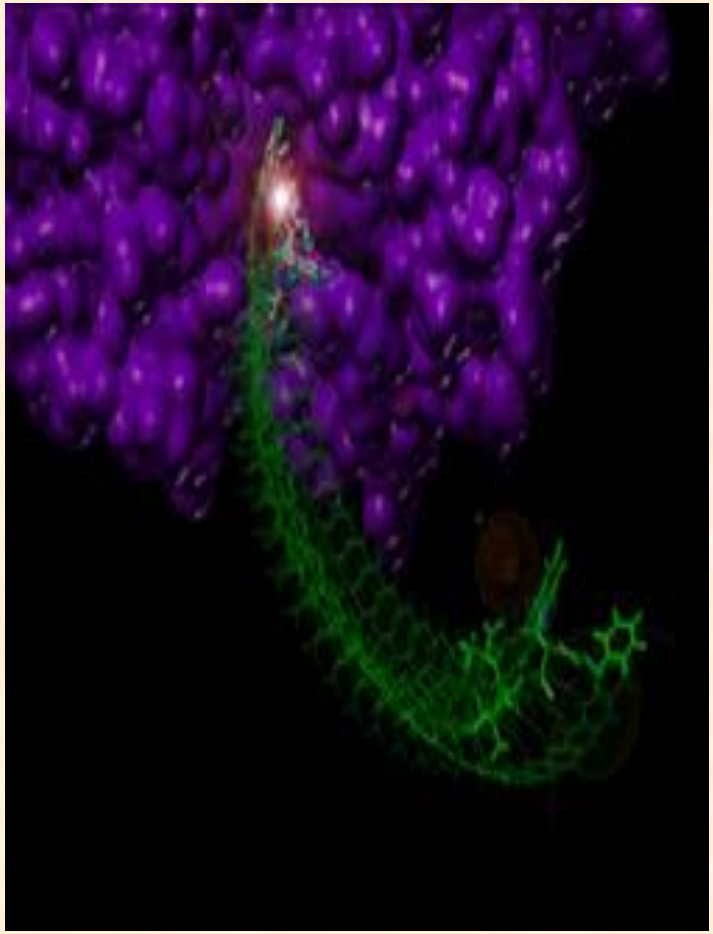
parameterised binding site docked ligand

- Uses “random” translation, rotation, and torsion, and look for a better binding mode.



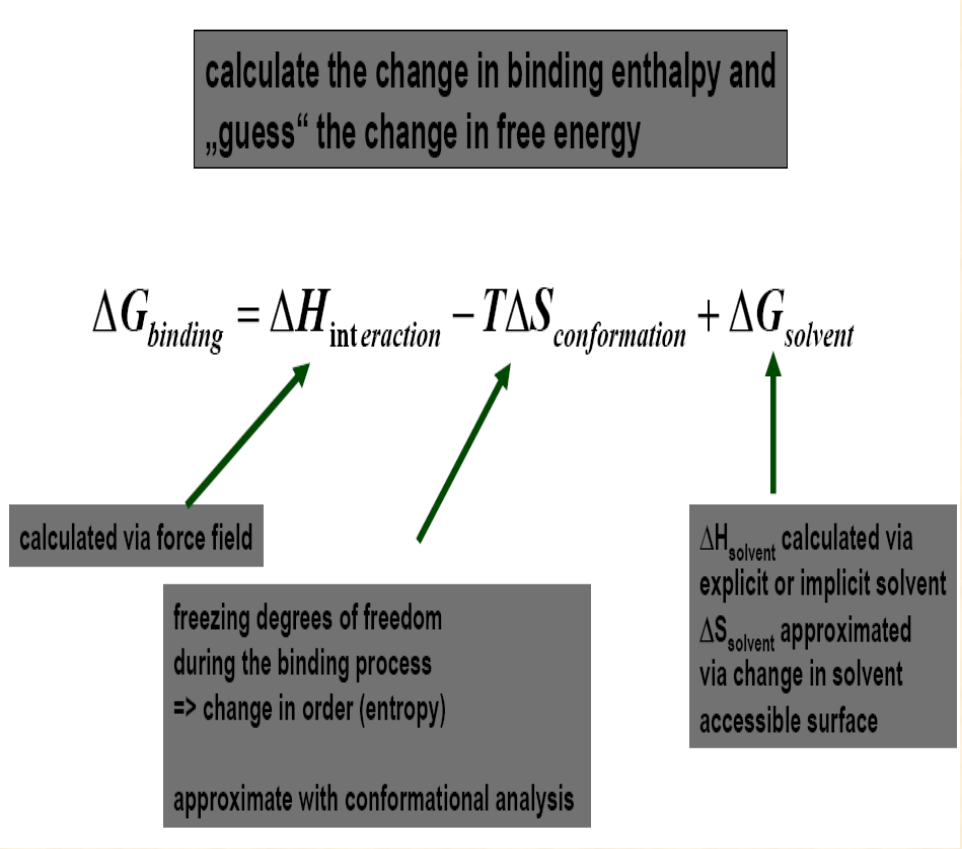
- Even though we have considered the ligand to be flexible, the active site was kept as a rigid structure.
- The side chains of the protein in the vicinity of the active site should be flexible, but computationally more expensive.

Incremental Construction (FlexX)



- A piecewise assembly of ligand within the active site.
- Generate rigid fragments by scissoring the rotatable bonds of known ligands.
- Dock the fragments one by one starting from the larger fragment
- Assemble the whole ligand by reconnecting them and repeat the docking process

Free Energy of Binding



- Dock ligand into pseudo-intercalation site
  - Manual, automatic, and flexible ligand docking
- Energy minimize to determine $\Delta G_{complex}$
- Determine ΔG_{ligand} =interaction energy of ligand with surroundings when explicitly solvated

$$\Delta G_{binding} = \Delta H_{interaction} - T\,\Delta S_{conformation} + \Delta G_{solvent}$$

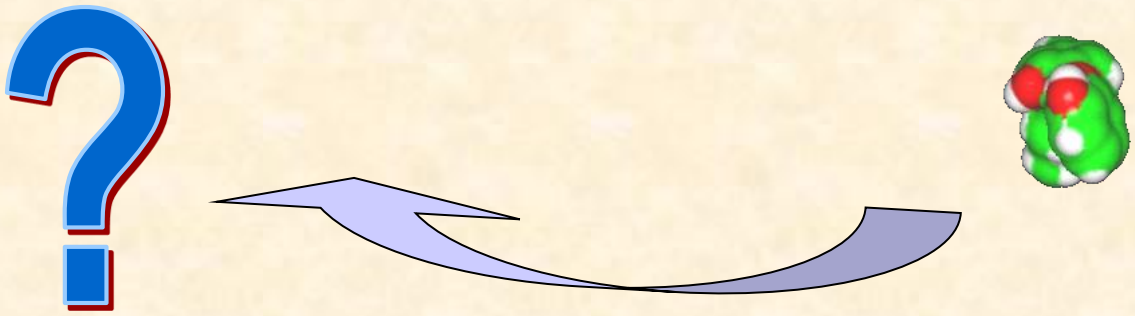
Need for Scoring

Detailed calculations on all possibilities would be very expensive

The major challenge in structure based drug design to identify the best position and orientation of the ligand in the binding site of the target.

This is done by scoring or ranking of the various possibilities, which are based on empirical parameters, knowledge based on using rigorous calculations

Exact Receptor Structure is not always known

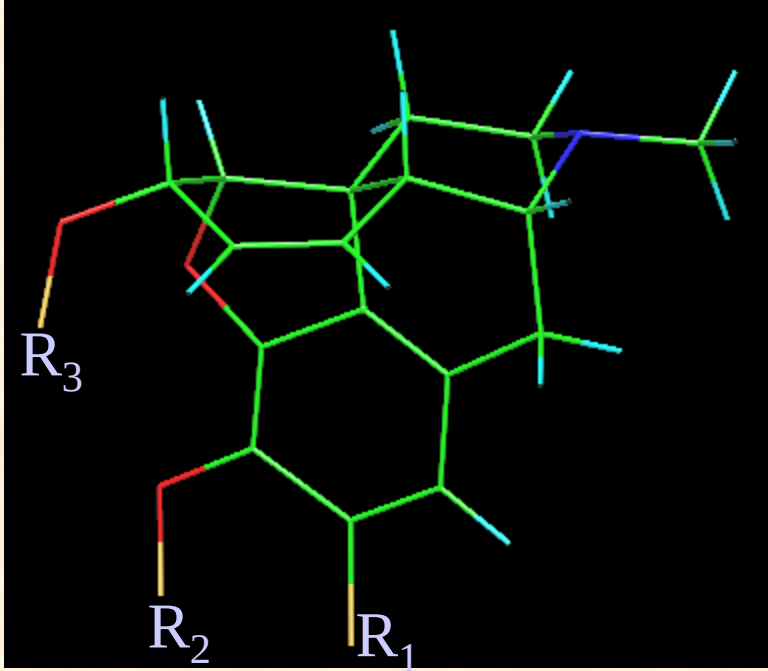


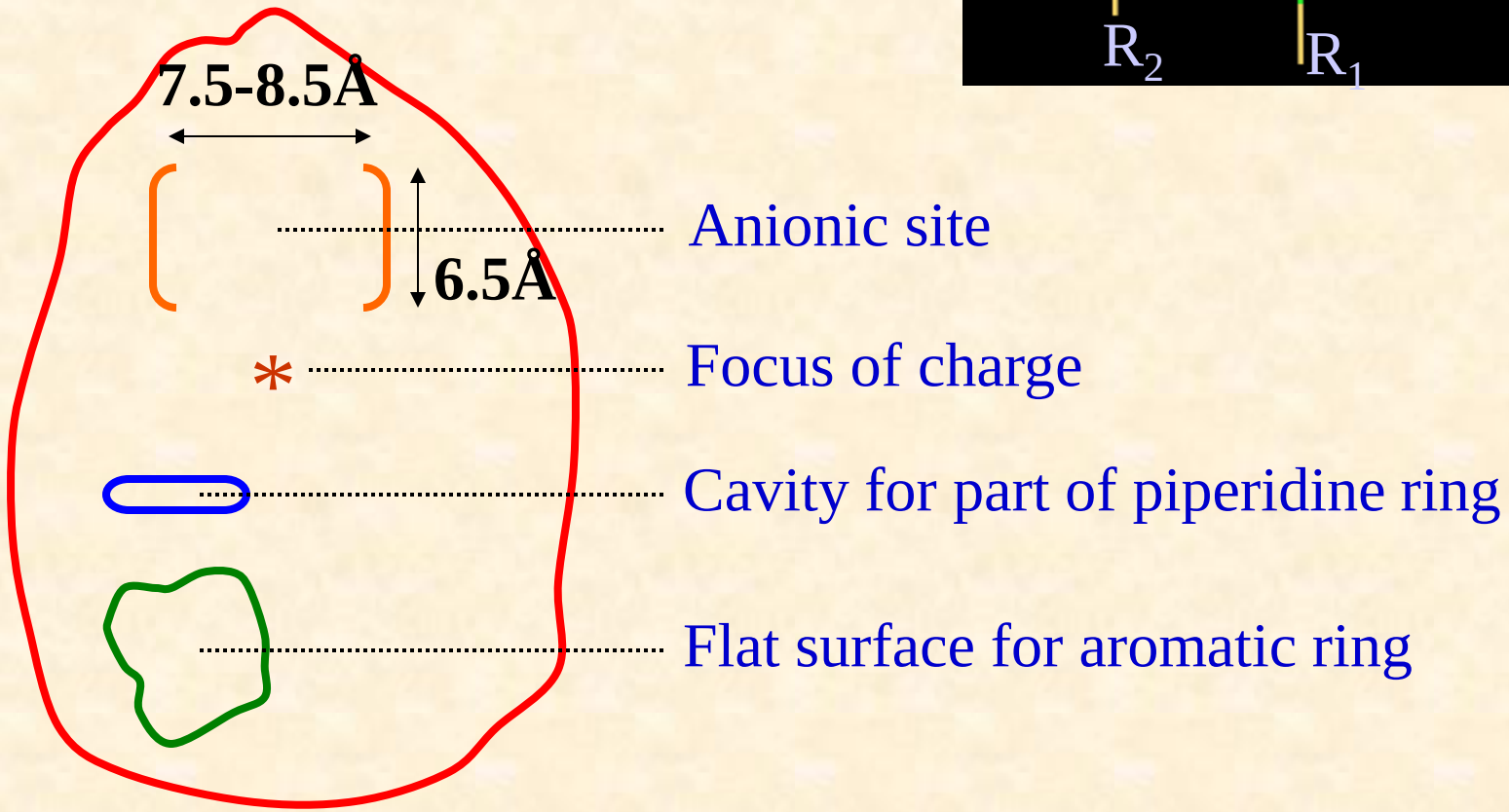
- **Receptor Mapping**

The volume of the binding cavity is felt from the ligands which are active or inactive. This receptor map is derived by looking at the localized charges on the active ligands and hence assigning the active site.

Receptor Map Proposed for Opiate Narcotics

(Morphine, Codeine, Heroin, etc.)





Homology modeling

Predicting the tertiary structure of an unknown protein using a known 3D structure of a homologous protein(s) (i.e. same family).

Assumption that structure is more conserved than sequence

Can be used in understanding function, activity, specificity, etc.

Structure modeling (Structure vs. Sequence)

- Homology modeling
- Fold recognition/ Threading

The 3D structures are used to understand protein function and to design new drugs

Mapping by Homology

Sequence 1

-A-K-L-M-

Sequence 2

-I-Q-D-M-

Dynamic Programming

Dayhoff Matrix

Sequence Alignment

-A-K-L-M-

| | |

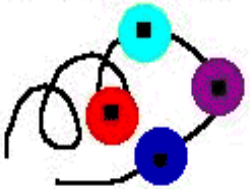
-I-Q-D-M-

Profiling

Sequence

-A-K-L-M-

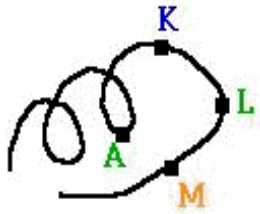
Structure with defined environment for each residue



Dynamic Programming

3D - 1D scoring table

Sequence to structure alignment

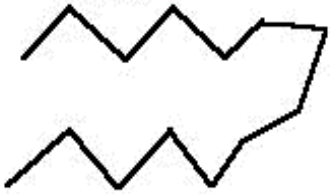


Threading

Sequence

-A-K-L-M-

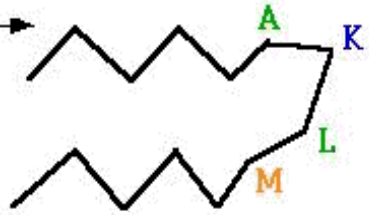
Fold



Dynamic Programming

Pair-wise potential function

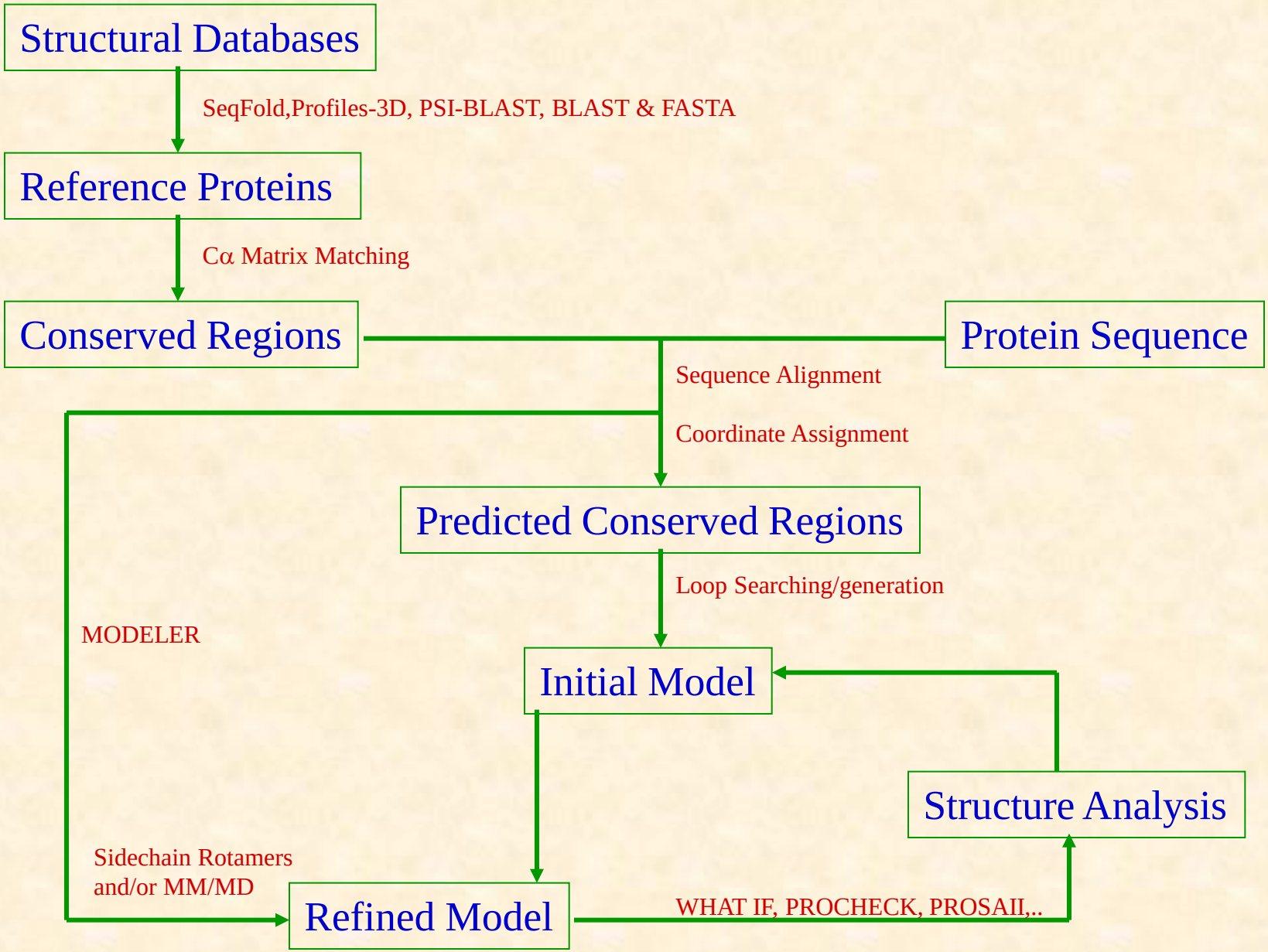
Optimum location of sequence on fold



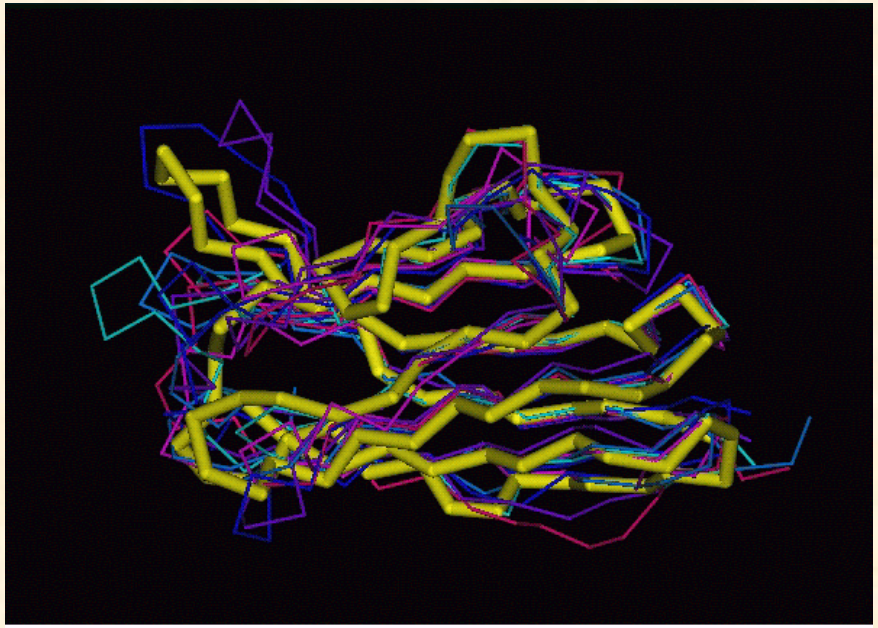
Key step in Homology Modeling

- Alignment
  - Multiple possible alignments
- Build model
- Refine loops
  - Database methods
  - Random conformation
  - Score: best using a real force field
- Refine sidechains
  - Works best in core residues

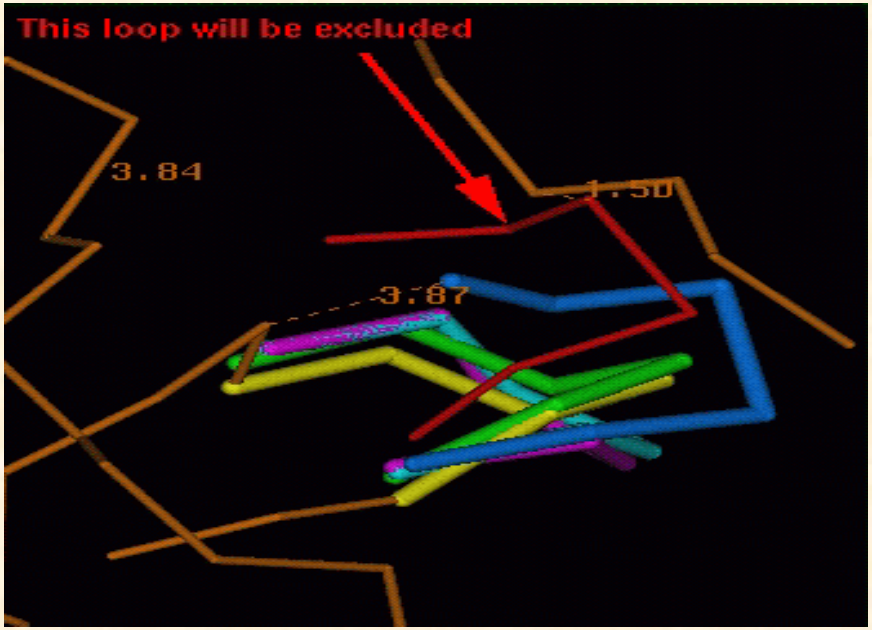
Structure Prediction by Homology Modeling



Generating a framework



Framework for just the target backbone is shown in yellow against the template structures



Fragments which have the right conformation to properly connect the stems without colliding with anything else in the structure

Typical Contributions of CADD to Drug Discovery Projects

- Suggestions of structures which, when retrieved from a compound collection or synthesized, were found upon testing to be active or inactive as predicted
- Development of structure-activity relationships
- Visualization of receptor models, pharmacophoric models, molecular alignments, or data models
- Reanalyzing available data to achieve new insights
- Creative search of available structures to find new leads

- Identification of preferred sites for structure elaboration
- Development of models to improve drug transport, specificity, safety or stability
- Development of mechanistic insights
- Use of leads in one area to derive new leads in a related assay
- Establishment of useful databases of project structures and properties
- Computation of physical or chemical properties to correlate with activities

Kinds of Computational approaches for the discovery of new ligands

- The search in 3D databases of known small molecules
- De novo design

Structure Searching

2D Substructure searches

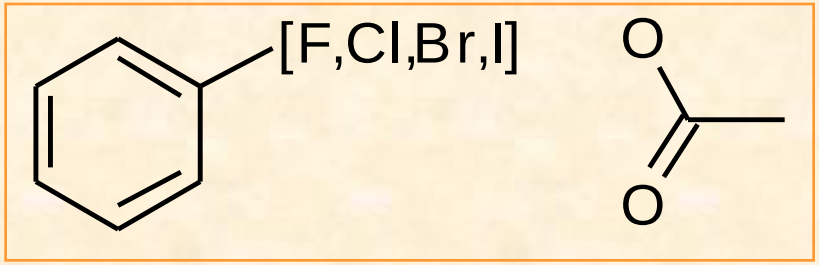
3D Substructure searches

3D Conformationally flexible searches

2D Substructure searches

Functional groups

Connectivity



De Novo Design

1) **Define Interacting Sites**

   HB donor/acceptor regions, Hydrophobic domain, Exclusion volumes

2) **Select Sites**

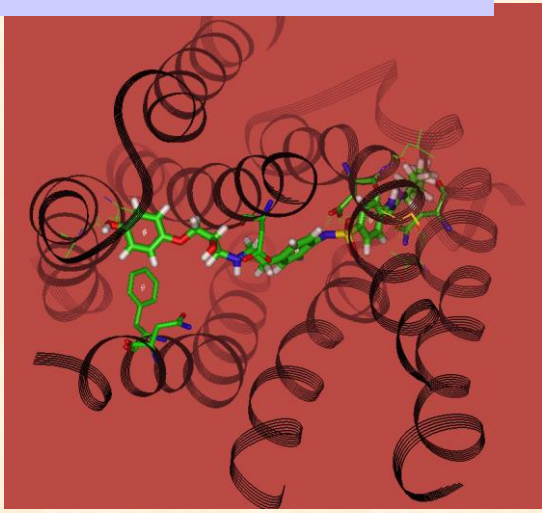
3) **Satisfy Sites**

4) **Join Functional Groups**

5) **Refine Structure**

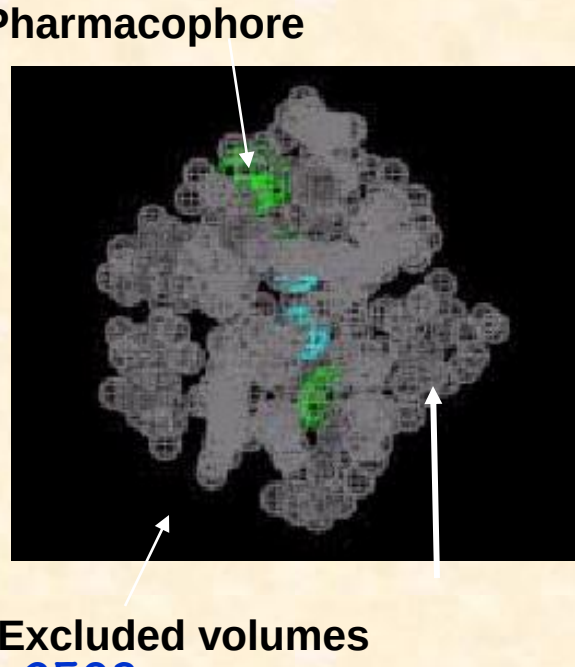
Virtual screening: Target structure based approaches

Protein-ligand docking



- The most promising route available for determining which molecules are capable of fitting within the very strict structural constraints of the receptor binding site and to find structurally novel leads.
- The most valuable source of data for understanding the nature of ligand binding in a given receptor

Active site-directed pharmacophores



- A Pharmacophore based method along with the utilisation of the geometry of the active site for enzyme inhibitors, represented by 'excluded volumes' features,
- Produces an optimised pharmacophore with improved predictivity compared with the corresponding pharmacophore derived without receptor information

Greenidge et. al. *J Med Chem.* 1998, 41, 2503

Pharmacokinetics play an extremely important role in drug development

ADMET

- Absorption
- Distribution
- Metabolism
- Excretion
- Toxicity

Outlook

- Molecular modeling first introduced in the pharmaceutical industries in the early 70's have raised probably unrealistic hopes such as it can "do it all". But it took quite a while before it could deliver
- No doubt, with the ever-expanding new powerful methods available, today's modelers have the requisite potential to bring real benefits to pharmaceutical industry.

- Molecular Modeling and Computational Chemistry are essential to understand the molecular basis for biological activity and has Tremendous Potential to aid Drug Discovery

- A healthy interaction between computational chemists and pharmaceutical industry seem indispensable.

- Structure Based Drug Design is an extremely important tool in the computer aided drug design.

- I Hope that you are convince!


Molecular Modeling Group
IICT
January, 04