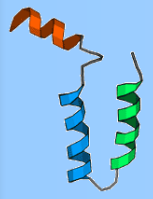
Molecular Modeling And Simulation

Junmei Wang
Department of Pharmacology, University of Texas Southwestern Medical Center at Dallas

Junmei.wang@utsouthwestern.edu



About me

- 1998, Ph.D in Physical Chemistry, Peking University
- 1998-2001, Postdoc in Computational Chemistry, UCSF

  Supervisor: Dr. Peter Kollman
- 2001-2006, Senior Scientist at Encysive Pharmaceuticals Inc
- 2007-, Assistant Professor, Green Center of Systems Biology, UTSW

  Research Interest:

  1. molecular mechanical force field development
  2. Molecular dynamics simulations for biosystems
  3. Computer-aided drug design
  4. Rational protein design

Course Goal

1. To provide an introduction to some current methods in molecular modeling
2. To provide hands-on experience with various molecular modeling software packages
3. To provide some background on the theoretical and computational methods use in molecular modeling
4. To provide some understanding of the capabilities, limitations and reliability of various molecular modeling methods

After the course, I hope each of your could

1. Know the strength and limitation of molecular modeling
2. Better understand papers/seminars related to molecular modeling
3. Conduct basic modeling study using software and resource in the public domain

Assignments and Grade

- Computational assignments - use various molecular modeling packages to investigate selected chemical problems
- Reading assignments – provide a one page summary of selected articles from the current literature involving molecular modeling
- Project and presentation
  - Select a computational chemistry problem from assignment or one related to your current research
  - Carry out the study during the second half of the course
  - 10 min presentation of your problem, computational methods and results
- Grade: 40% assignment, 30% reading assignment and 30% project presentation

What's Molecular Modeling

- **Molecular modeling** is the science and art of studying molecular structure and function through model building and computation.
- **What does wiki say?**

1. **Molecular modelling** encompasses all theoretical methods and computational techniques used to model or mimic the behavior of molecules. The techniques are used in the fields of computational chemistry, drug design, computational biology and materials science for studying molecular systems ranging from small chemical systems to large biological molecules and material assemblies.
2. The simplest calculations can be performed by hand, but inevitably computers are required to perform molecular modelling of any reasonably sized system.
3. The common feature of molecular modelling techniques is the atomistic level description of the molecular systems. This may include treating atoms as the smallest individual unit (the Molecular mechanics approach), or explicitly modeling electrons of each atom (the quantum chemistry approach).

Some Quotes on Molecular Modeling

- “Art is the lie that helps tell the truth” by *Pablo Picasso*
- “A model must be wrong, in some respects, else it would be the thing itself. The trick is to see where it is right” by *Henry A. Bent*

Closely Related Disciplines

1. ***Computational Chemistry*** – The application of mathematical and computational methods to assist in solving chemical problems
2. **Theoretic Chemistry** – usually refer to as quantum chemistry
3. ***Computer-Aided Drug Design*** – The discipline of using computational techniques (including chemical informatics) to assist in the discovery and design of drugs.
4. **Computation Biology** – molecular modeling and molecular simulation applied to study biology systems
5. **Chemical Informatics** –the branch of informatics dealing with all aspects of the representation and use of chemical structures, proteins, and related information, on computer.

Basic Molecular modeling methods

1. Protein homology modeling
2. Energy minimization
3. Molecular dynamics simulation
4. Monte Carlo simulation
5. Free energy calculation
6. Reaction rate calculation
7. Normal mode analysis
8. Molecular docking
9. Pharmacophore modeling
10. 2D and 3D QSAR

Basic Molecular modeling Application

1. Structure determination and refinement
2. Mutant structure prediction
3. Enzyme mechanism
4. Protein folding pathways
5. Structure-based drug design
6. Ligand-based drug design
7. Protein design

Classification of Molecular Systems

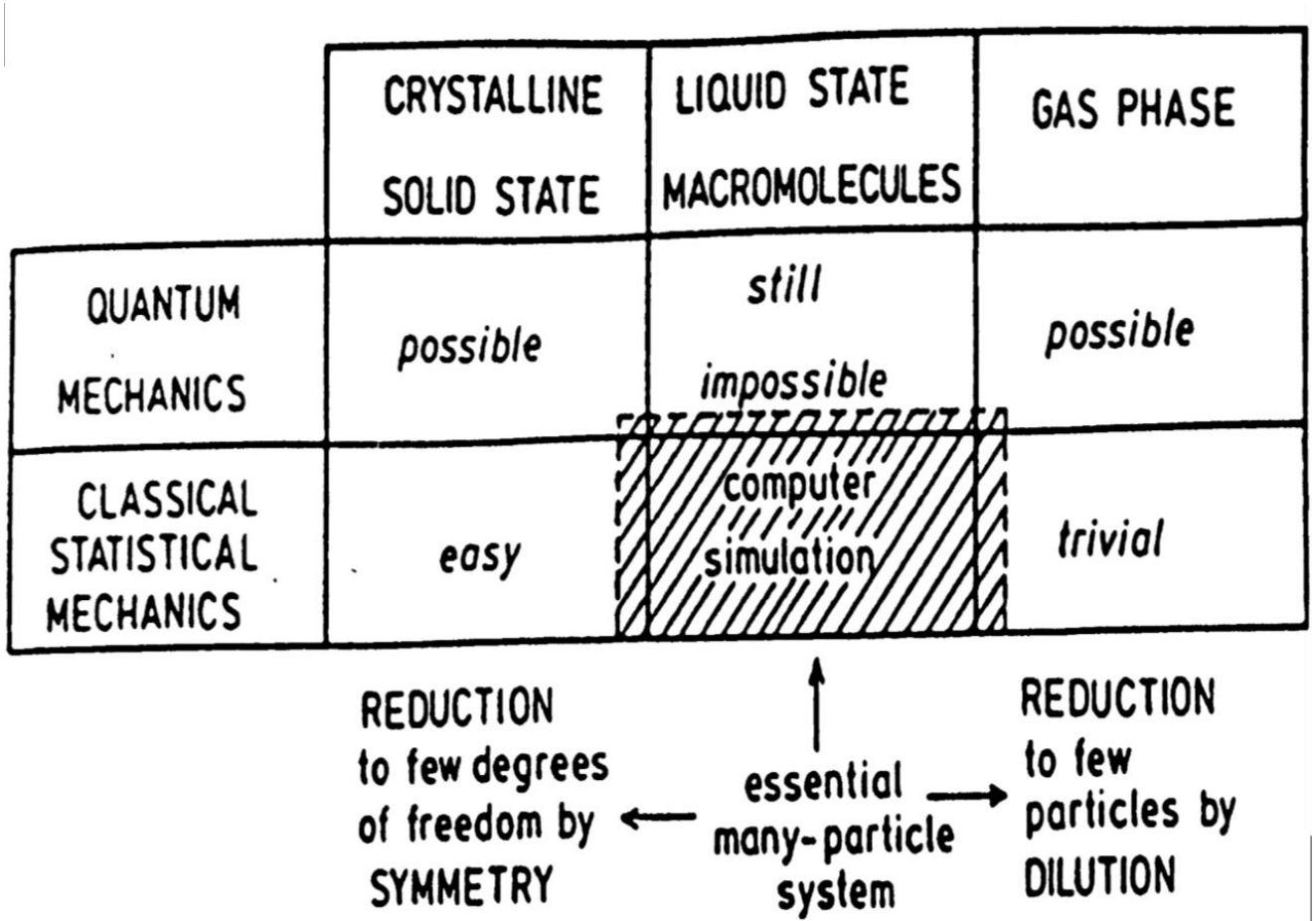


Fig. 1. Classification of molecular systems. Systems in the shaded area are amenable to treatment by computer simulation.

What Is A Model?

The definition of **model** in Oxford English Dictionary : a simplified or idealized description of a system or process, often in mathematical terms, devised to facilitate calculations and predictions.

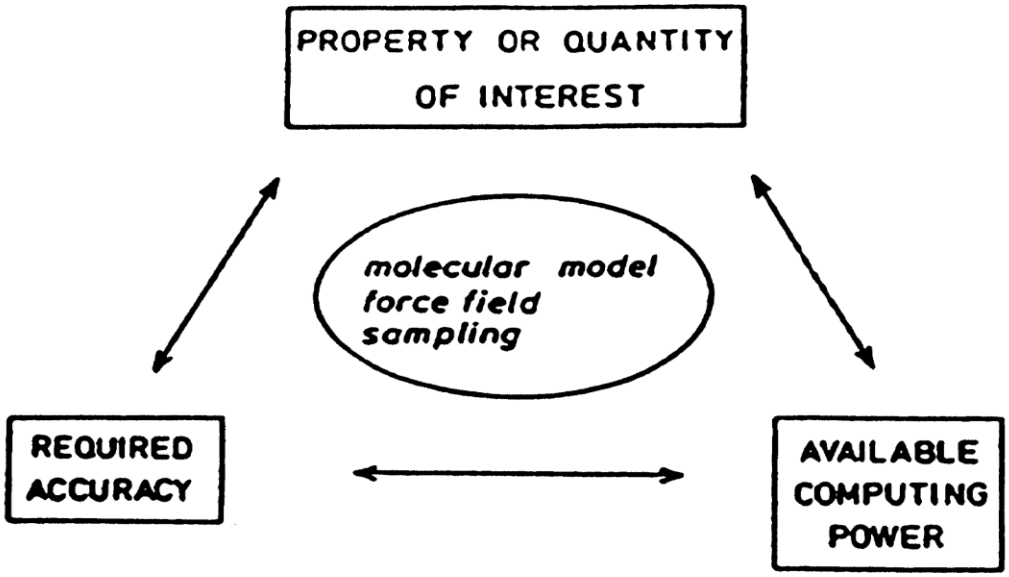


Fig. 3. Choice of molecular model, force field and sample size depends on 1) the property one is interested in (space to be searched), 2) required accuracy of the prediction, 3) the available computing power to generate the ensemble.

Principle of Molecular Modeling and Molecular Simulations

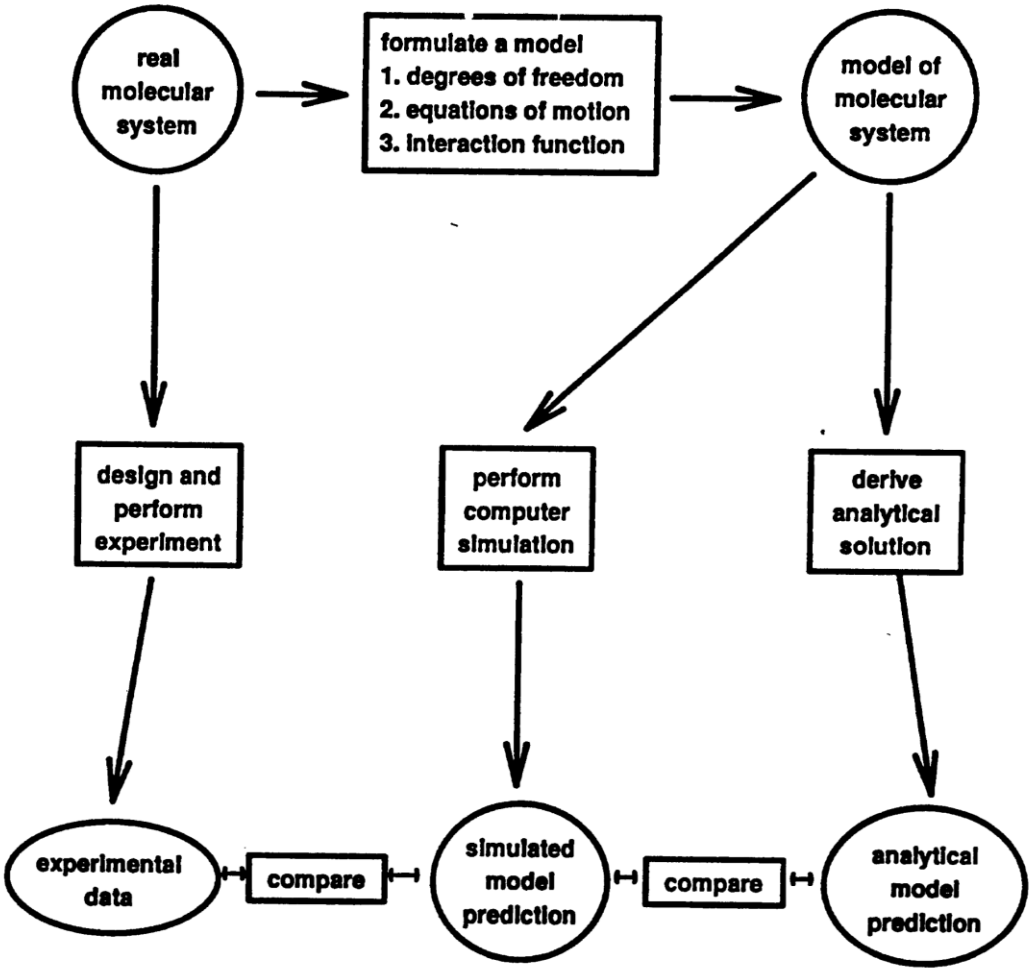
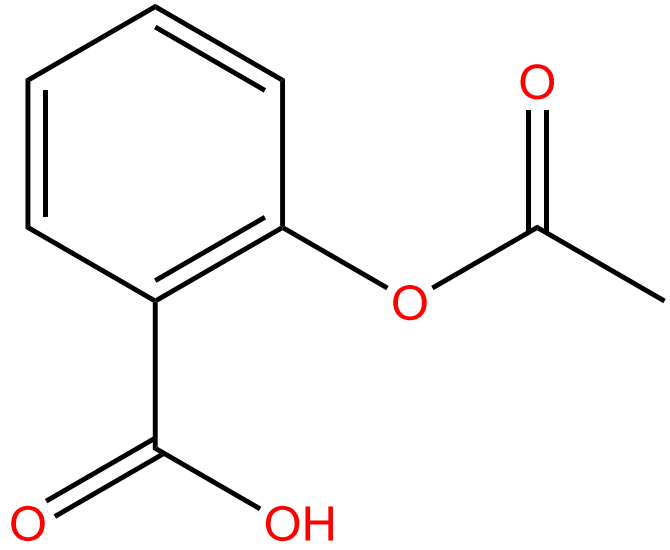


Fig. 1. Molecular models, simulation and experiment.

Representing 2D structures

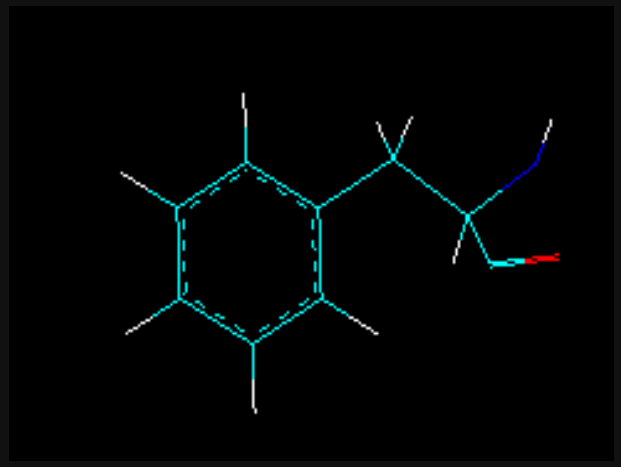
1. Historical representations & structure diagrams
2. Linear notations including SMILES
3. Connection tables
4. Considering structures as mathematical graphs
5. Representation nuances
6. Canonicalization
7. Reactions
8. Generic (Markush) structures
9. Fingerprints



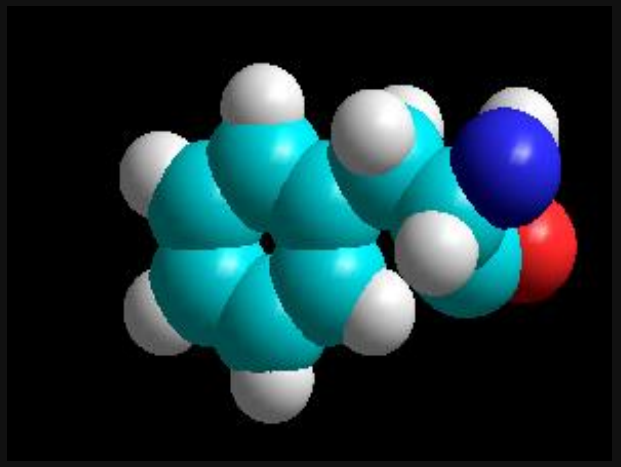
Aspirin

CC(=O)Oc1ccccc1C(=O)O

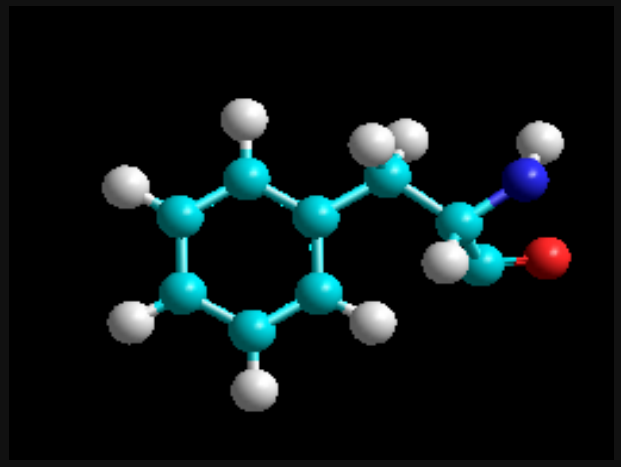
Representing 3D structures



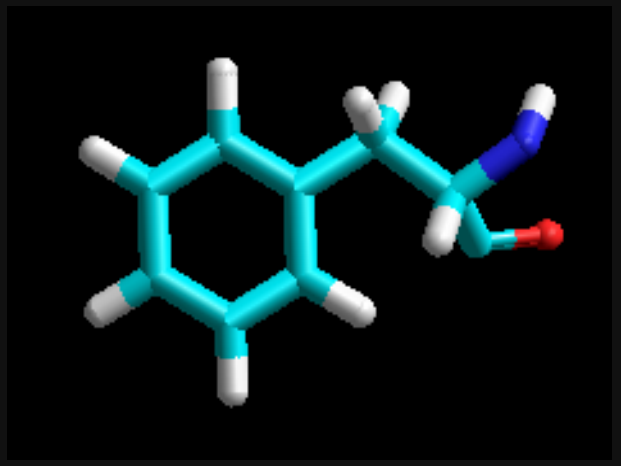
Line



CPK (Space Filling)

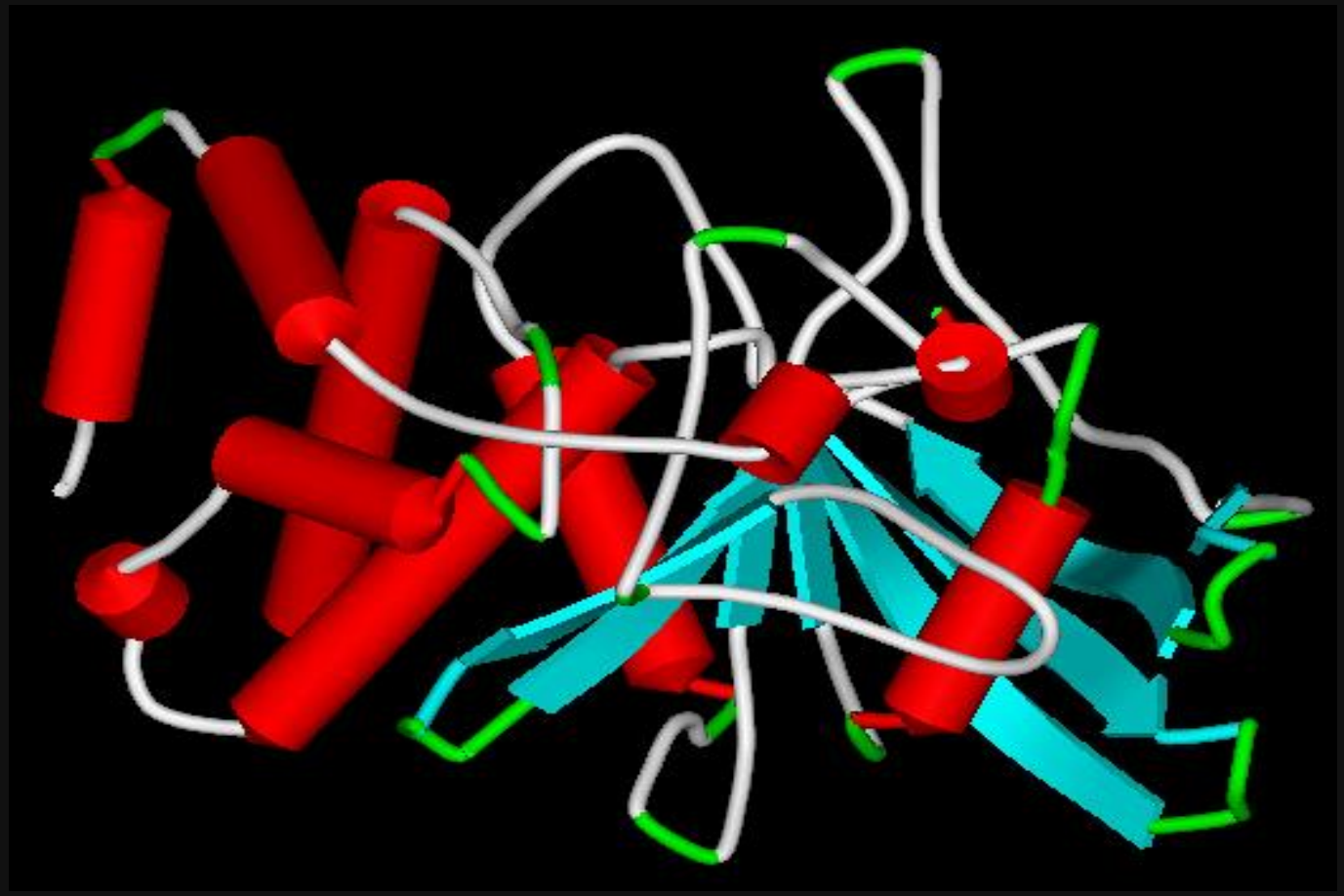


Ball and Stick



Stick

Representing 3D Structures For Proteins



Cartoon

Molecular Surface

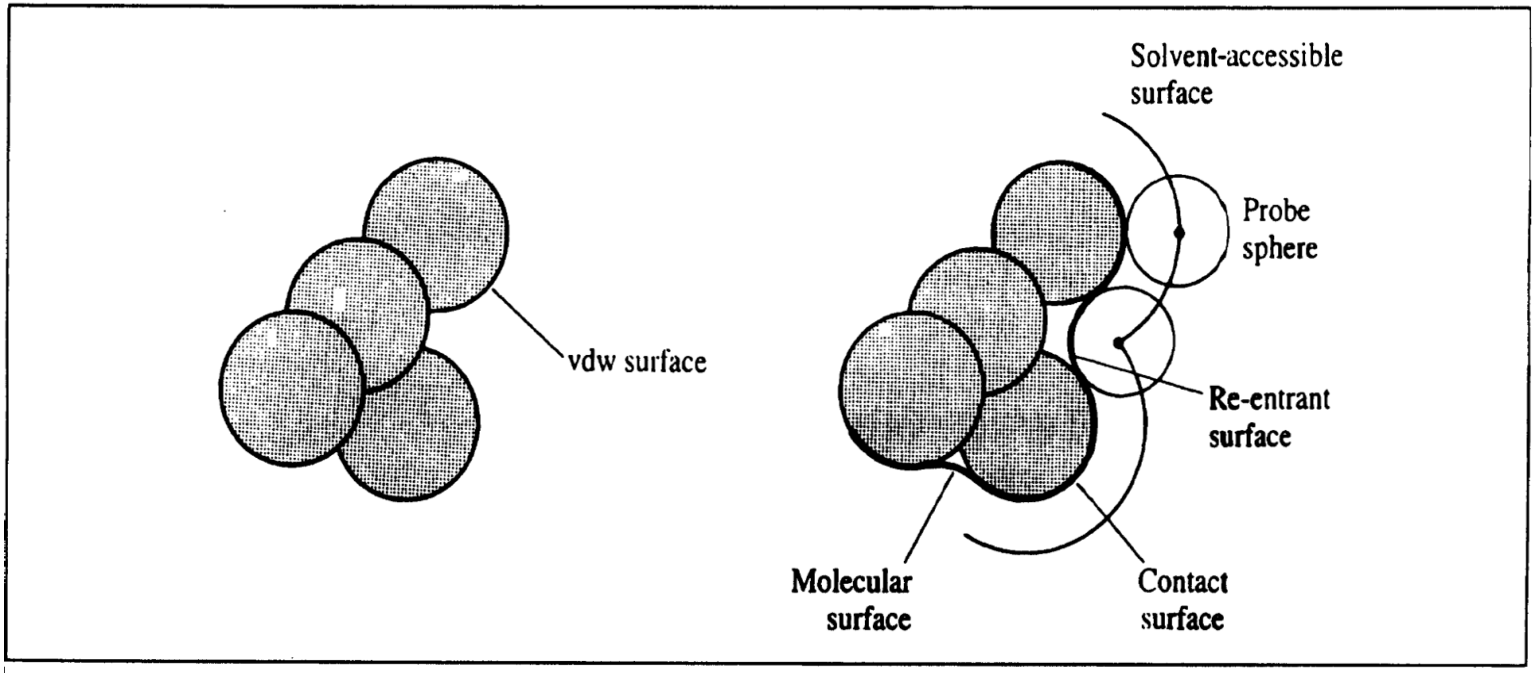
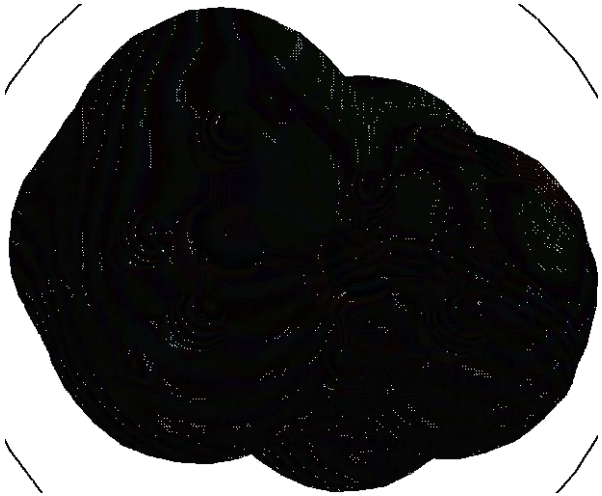


Fig. 1.6: The van der Waals (vdw) surface of a molecule corresponds to the outward-facing surfaces of the van der Waals spheres of the atoms. The molecular surface is generated by rolling a spherical probe (usually of radius 1.4 Å to represent a water molecule) on the van der Waals surface. The molecular surface is constructed from contact and re-entrant surface elements. The centre of the probe traces out the accessible surface.

Physical Representation



Electrostatic
Potential Map

Molecular Orbital
Representation

Potential Functions

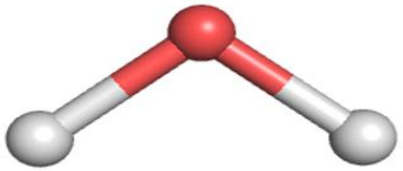
- **Quantum Mechanics**: $\hat{H}\Psi = E\,\Psi$

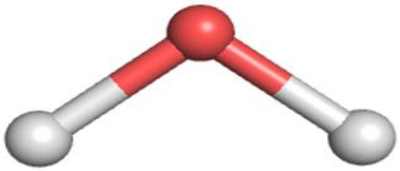
1BE9 (120 aa), ~**1,700** Yrs; **1ERK** (357 aa), ~**900,000** Yrs
(with a DFT (Density Functional Theory) model of B3LYP/6-31G*)

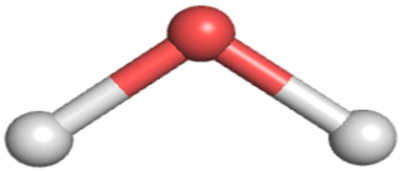
- **Molecular Mechanics**

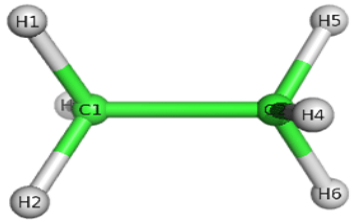
$$V_{potential} = \sum_{\text{bonds}} K_r\left(r - r_{eq}\right)^2 + \sum_{\text{angles}} K_\theta\left(\theta - \theta_{eq}\right)^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}\left[1 + \cos\left(n\varphi - \gamma\right)\right] + \sum_{\substack{i<j \\ \text{van der Waals}}} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}}\right) + \sum_{\substack{i<j \\ \text{electrostatic}}} \frac{q_i q_j}{R_{ij}}$$

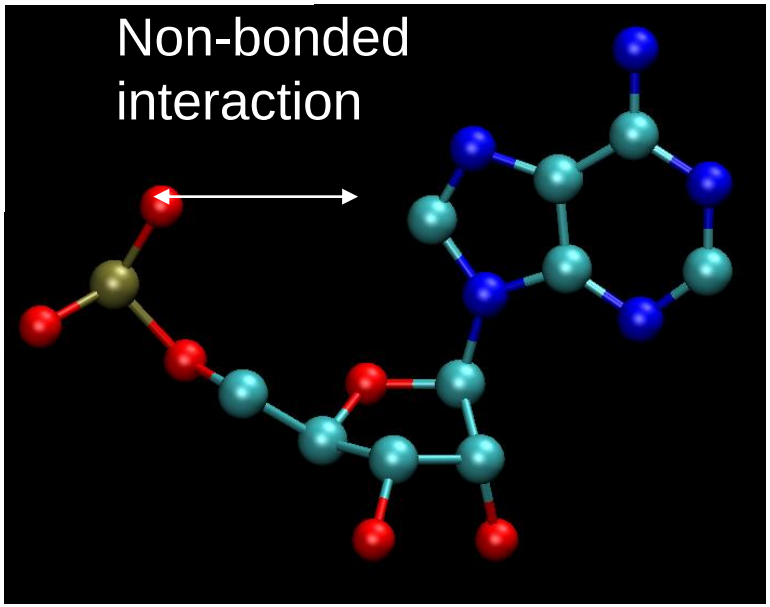
Additive
Harmonic
Effective











Coordinate Systems And File Formats

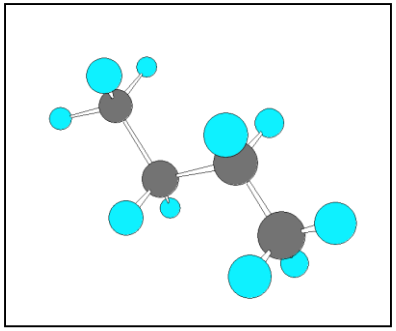
- Coordinate Systems
  1. Cartesian
  2. Internal

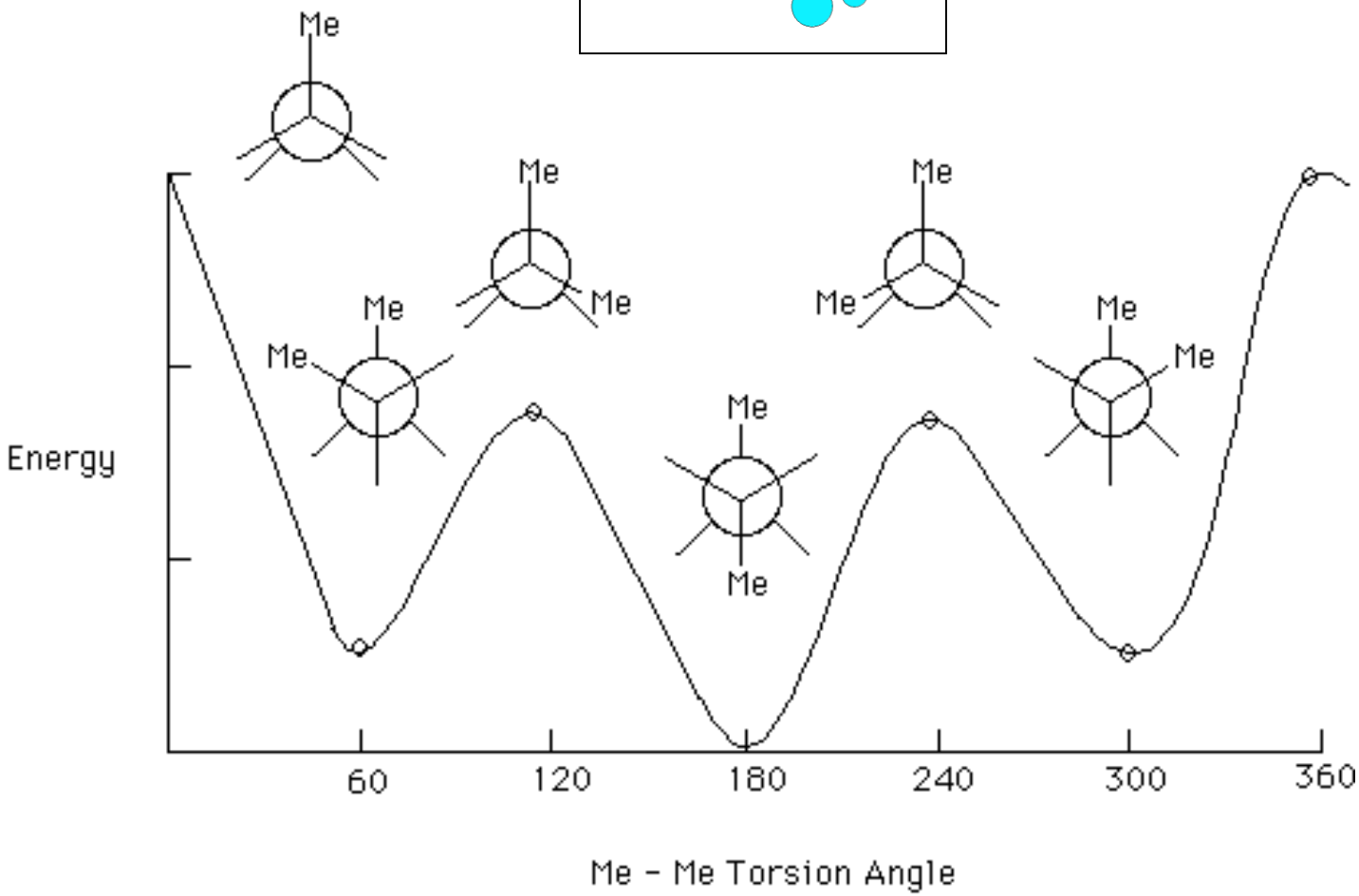
- Molecular Formats
  1. Mol2
  2. Gaussian input
  3. Pdb
  4. Sdf
  5. Smiles
  6. Smarts

Some Computable Quantities of Interest

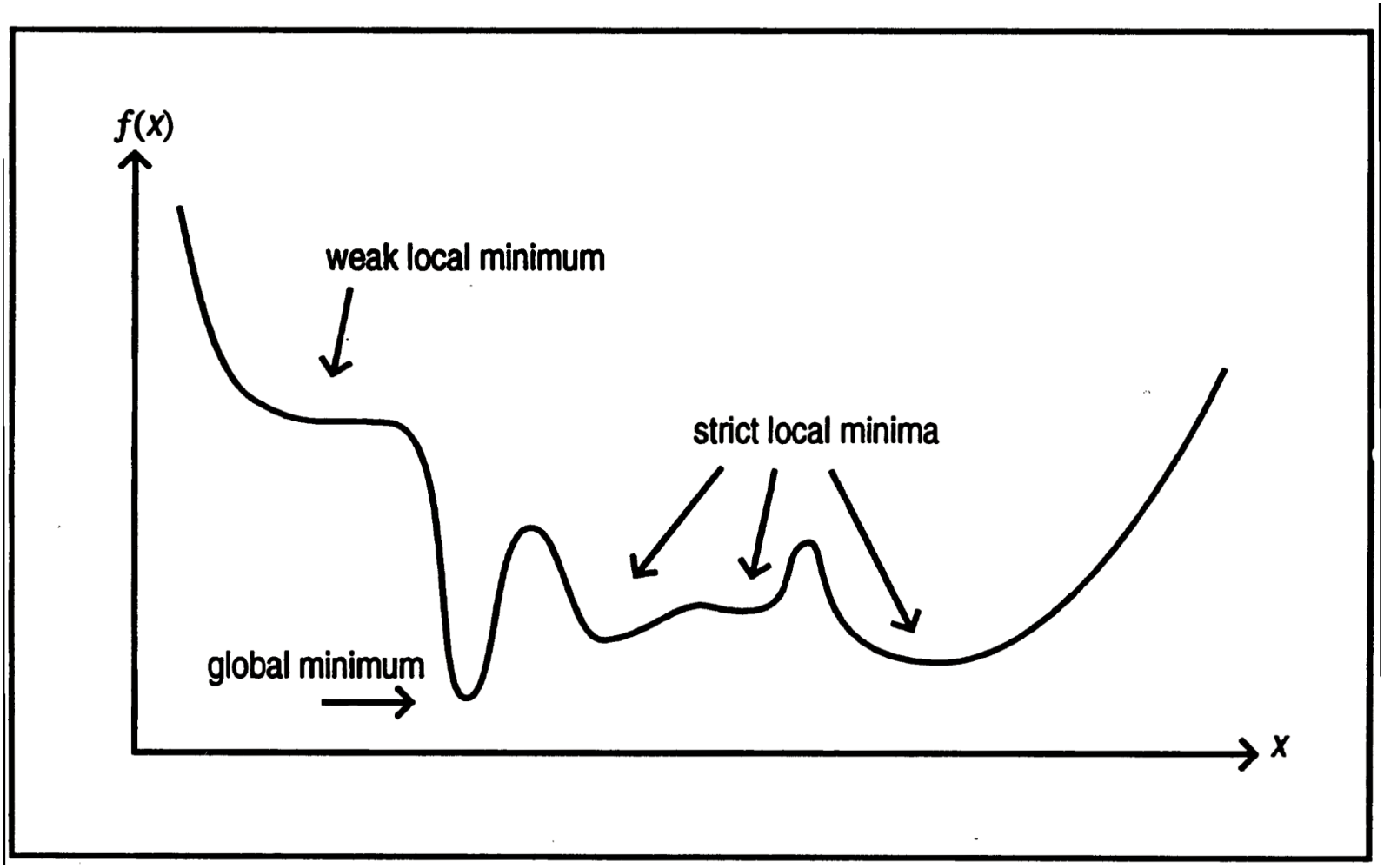
- **Molecular structure** - bond lengths, bond angles, torsion angles, etc.
- **Potential energy surfaces** (PES)
- **Energetics** – heats of formation, dissociation constants, vibrational frequencies, free energy
- **Kinetics** – reaction rate
- **Dynamics** – how does a molecular system relax and fluctuate over time?

Conformational Analysis of n-Butane

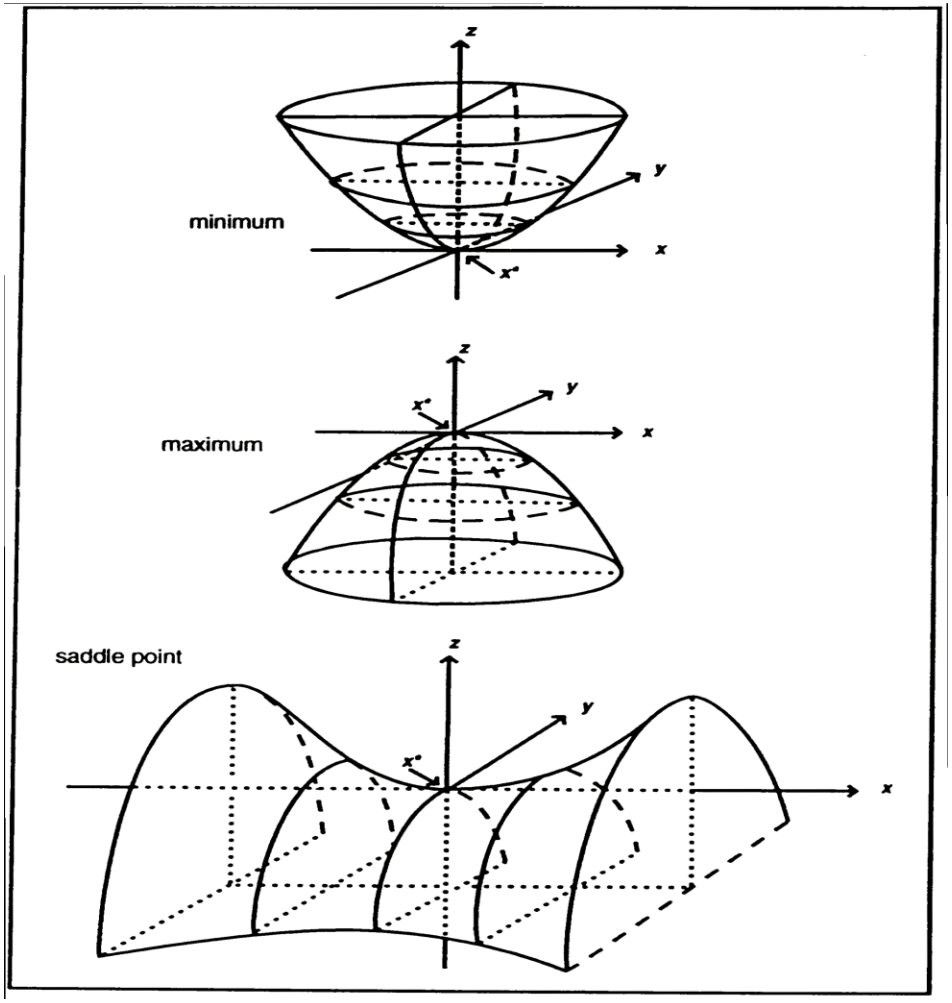


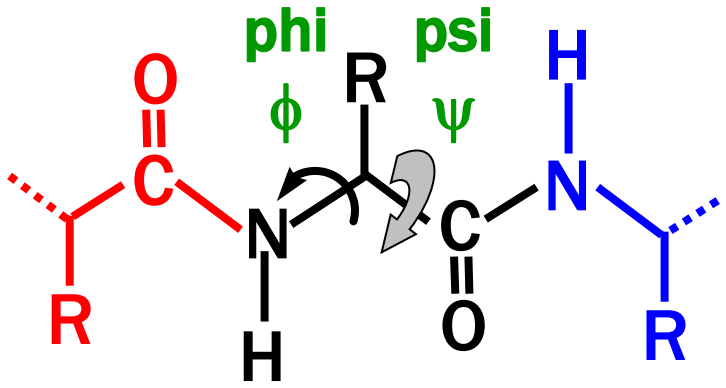


Types of Minima



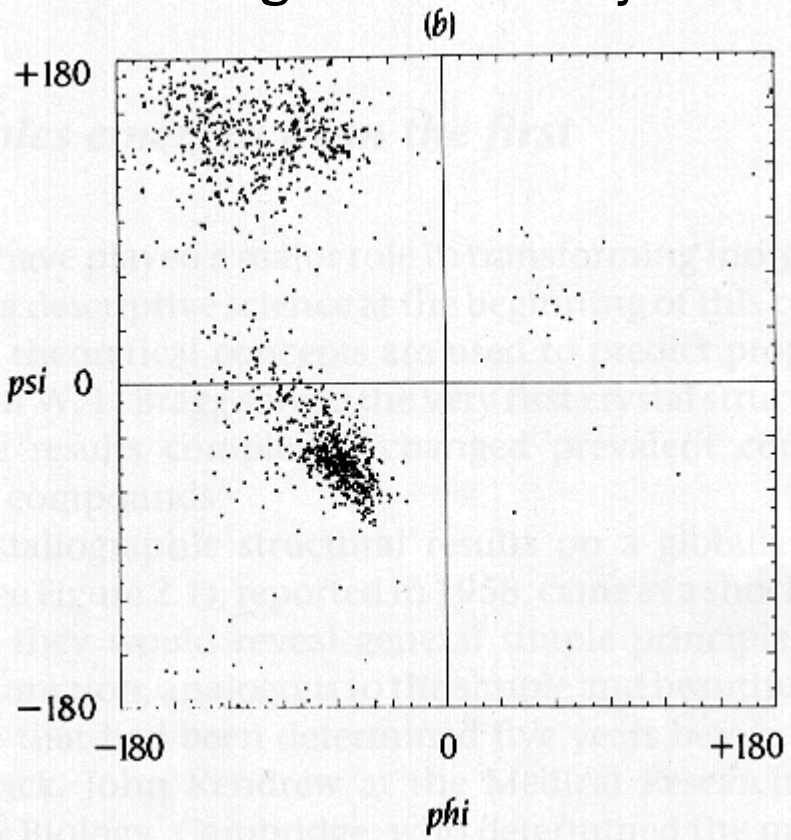
Stationary Points on PES



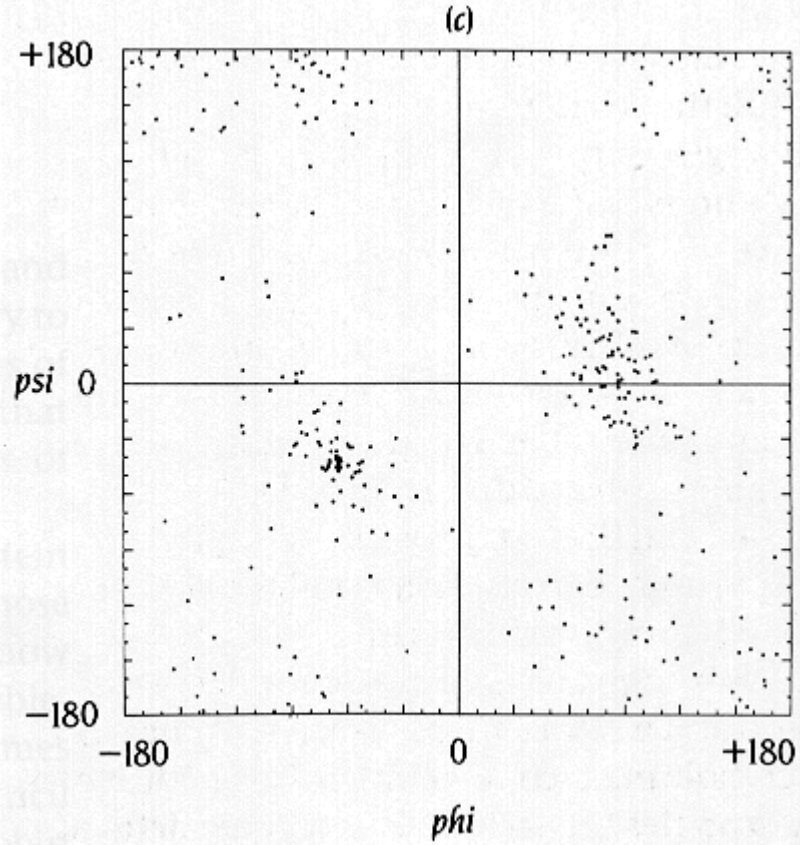


Brandon & Tooze
Introduction to Protein Structure, Figure 1.7b.c
From J. Richardson, *Adv. Prot. Chem.* 34, 174-174 (1981)

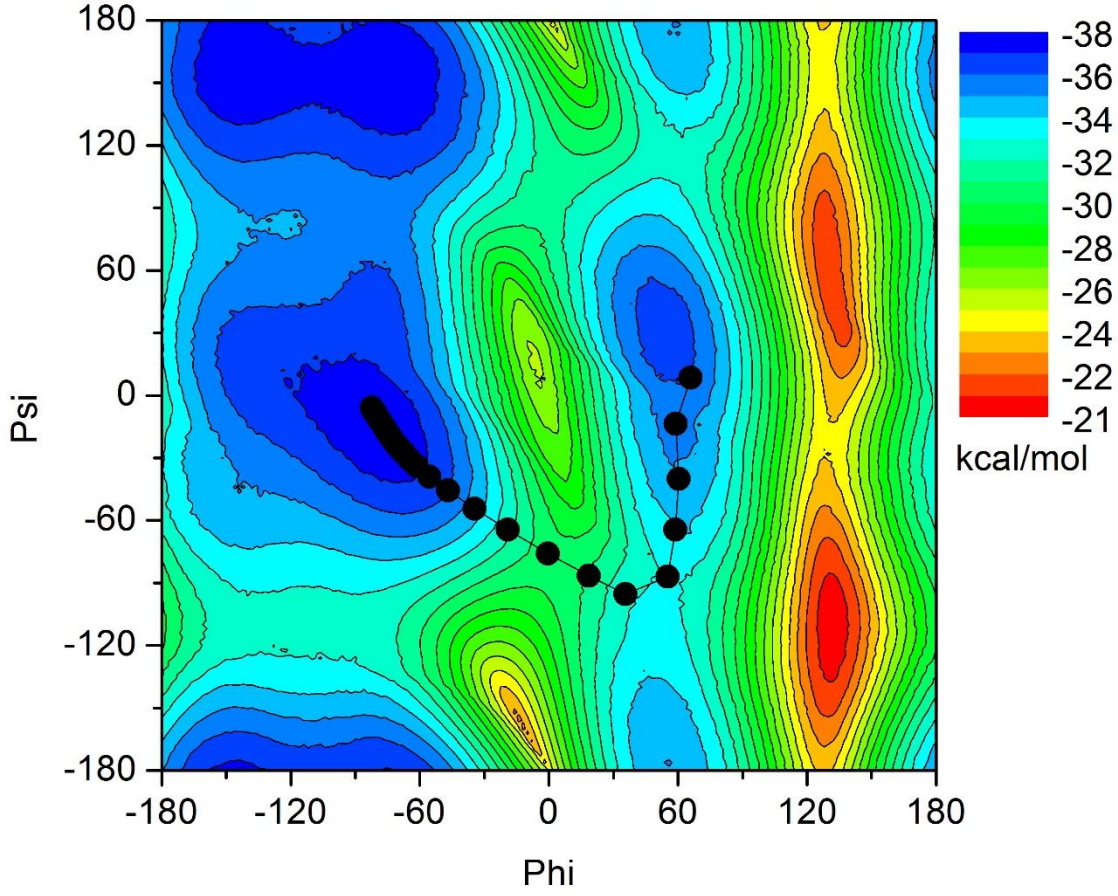
All amino acids (except glycine) from high resolution crystals



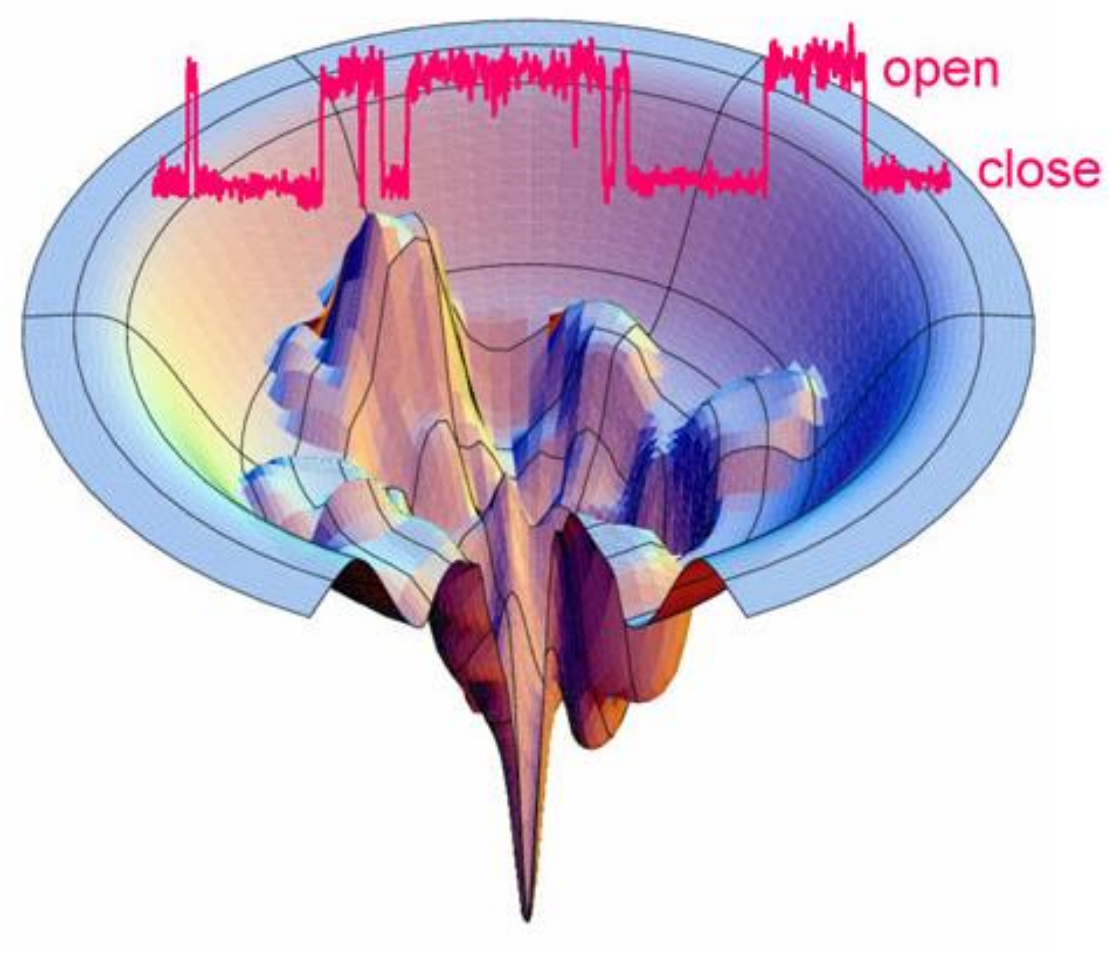
GLY: much more freedom



Potential Energy Surface of A Peptide Bond

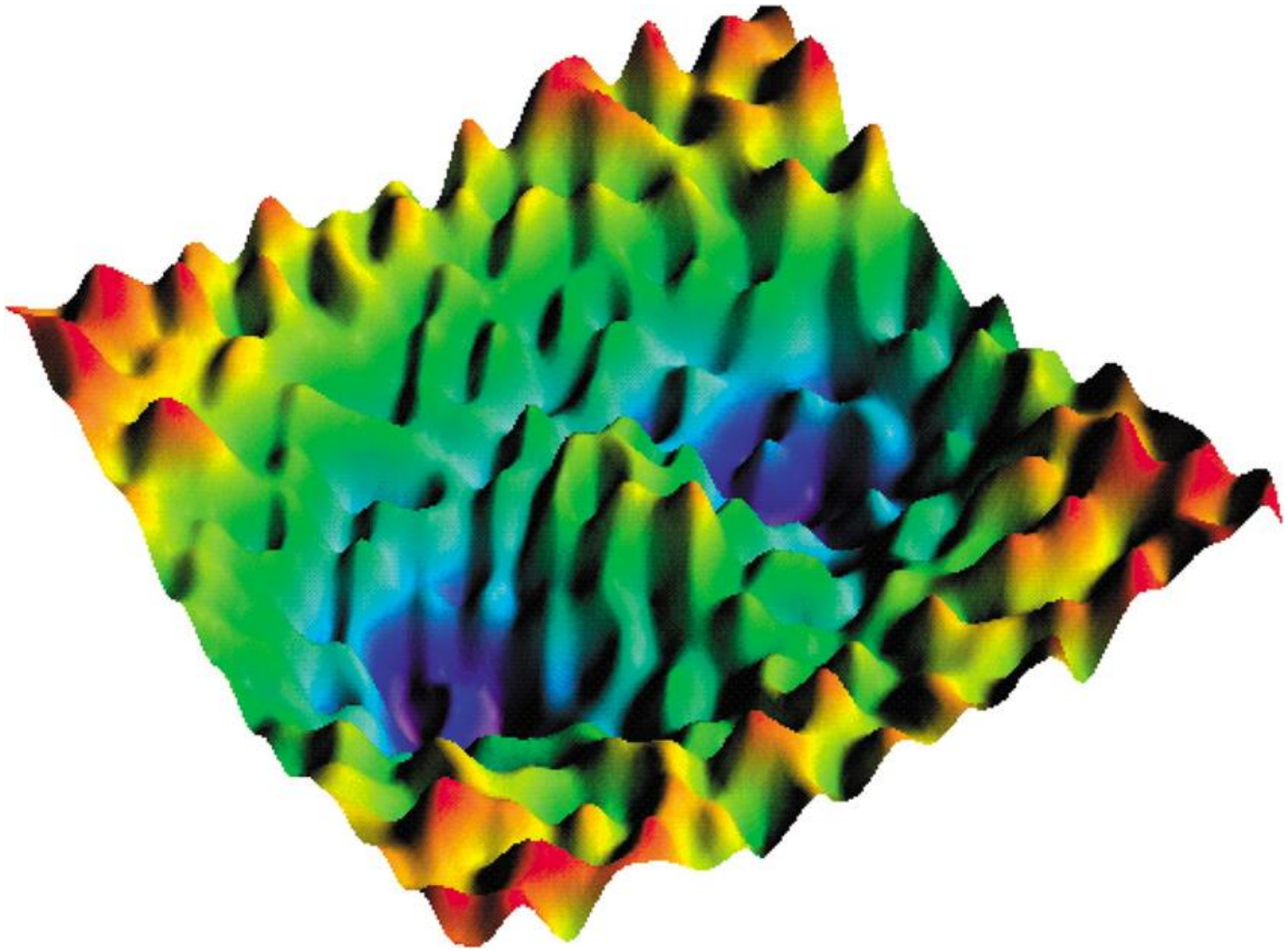


Ion channel activities translated to local movements on the protein energy landscape

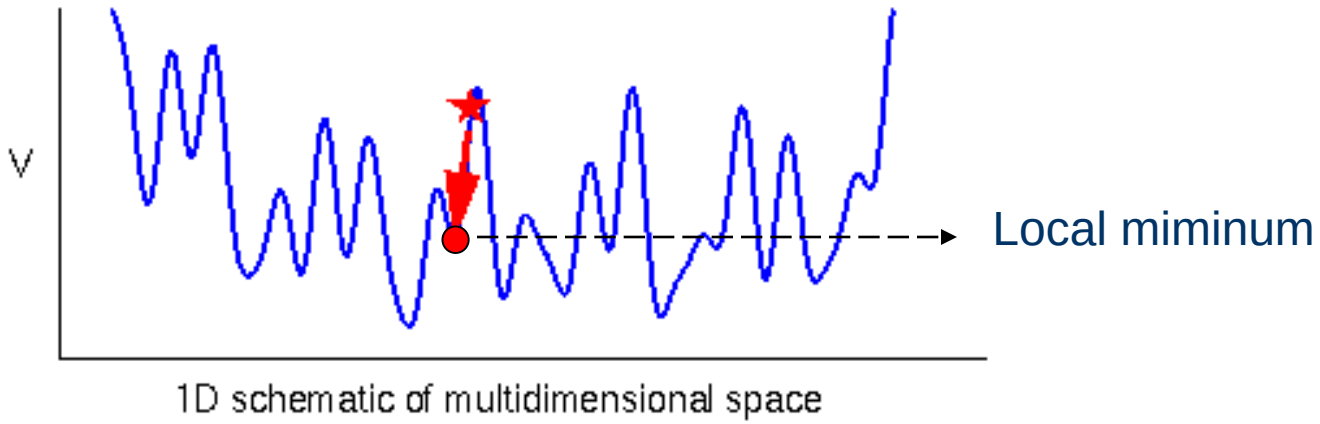


http://www.physiology.vcu.edu/research/mbrg/index.html

Schematic depiction of the potential energy surface of a complex system



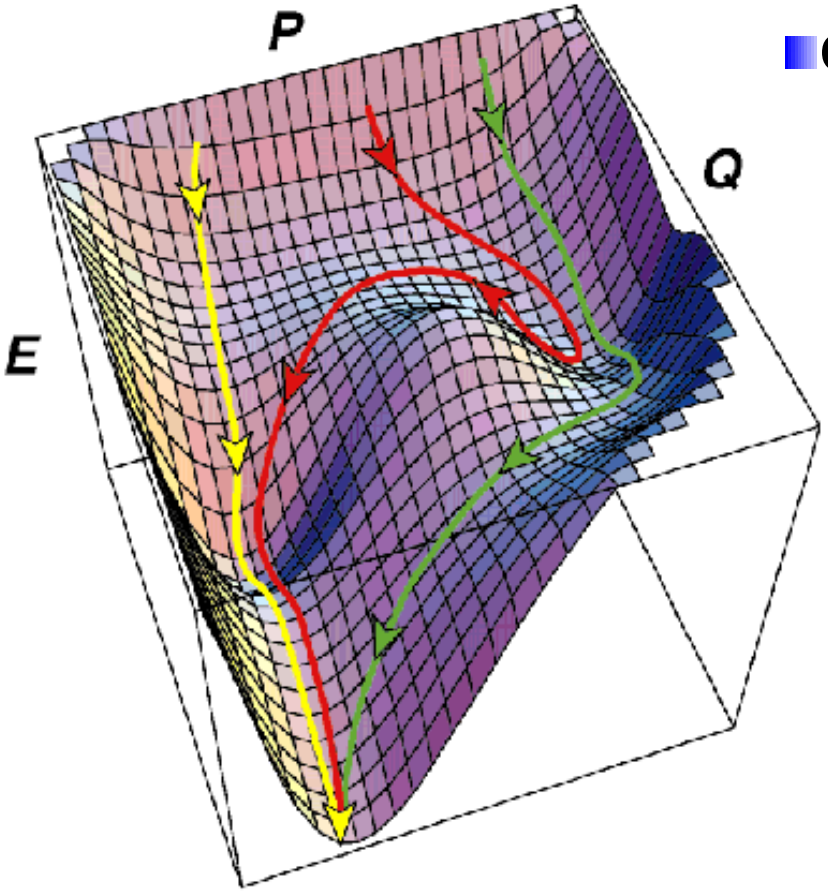
Energy Minimization



- Energy minimization
- Methods
  - First-order minimization: Steepest descent, Conjugate gradient minimization
  - Second derivative methods: Newton-Raphson method
  - Quasi-Newton methods: L-BFGS

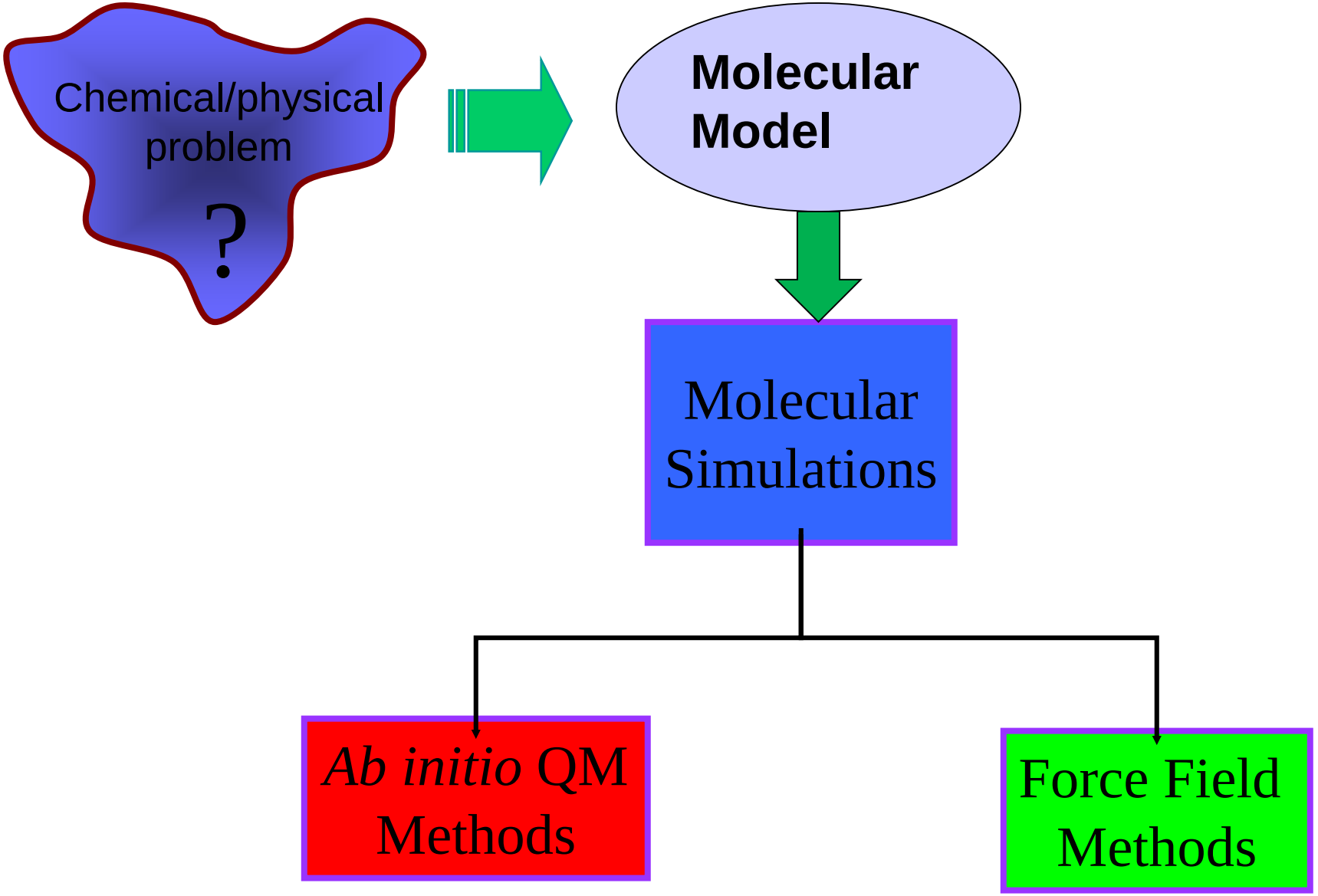
Search Potential Energy Surface

- We are interested in minimum points on Potential Energy Surface (PES)

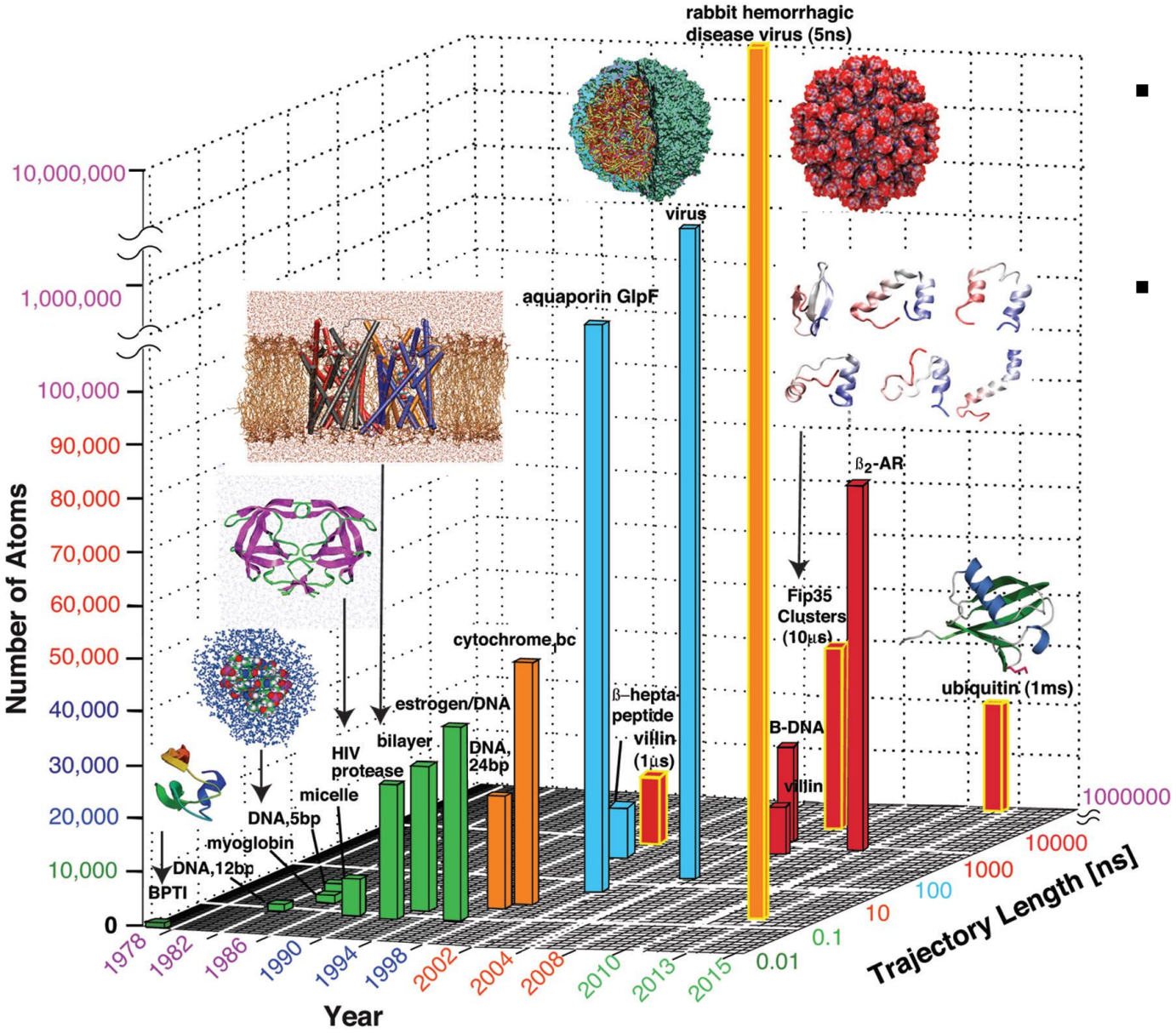


- **Conformational search techniques**
  - Energy Minimization
  - Monte Carlo
  - Molecular Dynamics
  - Others: Genetic Algorithm, Simulated Annealing

Simulations: Modelling Strategies



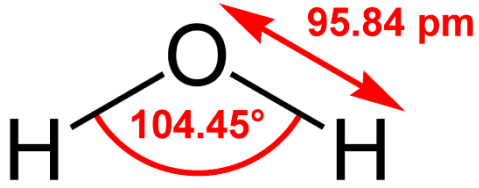
MD Simulations in Studying Biological Systems

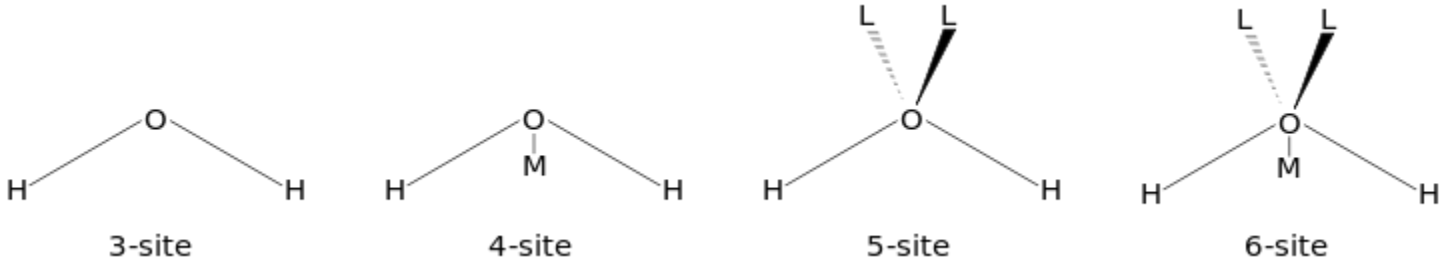


- 1-ms simulations of ubiquitin by D. E. Shaw
- 5-ns simulations of HIV-1 capsid consisting of 64 million atoms using NAMD

Courtesy of Tamar Schlick

Water Models





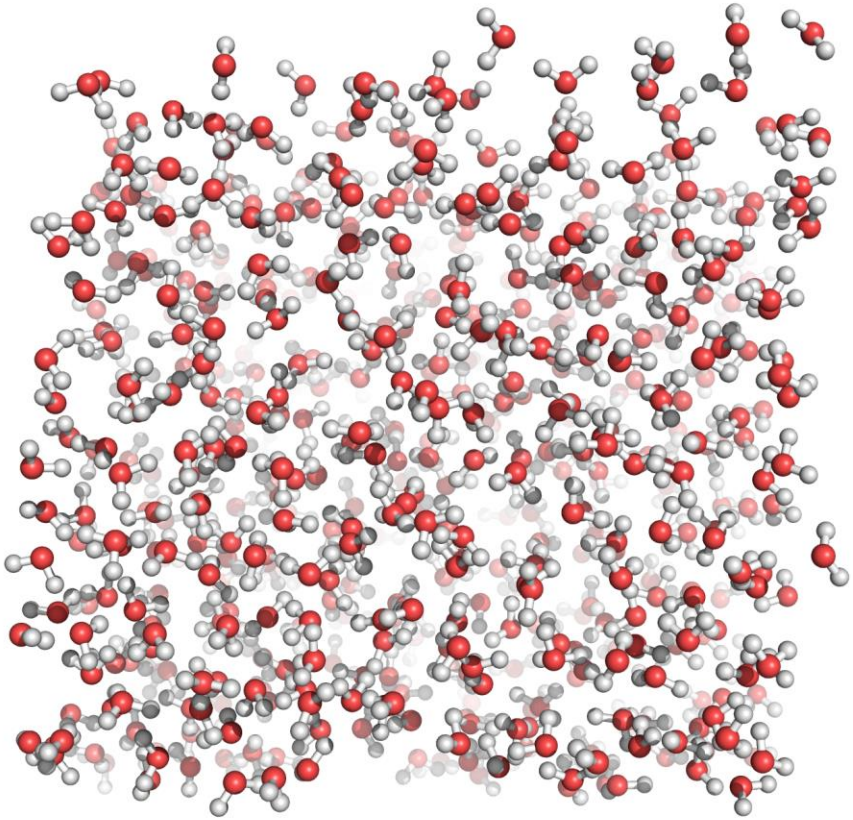
	TIPS	SPC	TIP3P	SPC/E
r(OH), Å	0.9572	1.0	0.9572	1.0
HOH, deg	104.52	109.47	104.52	109.47
A × 10^{-3}, kcal Å^{12}/mol	580.0	629.4	582.0	629.4
B, kcal Å^{6}/mol	525.0	625.5	595.0	625.5
q(O)	−0.80	−0.82	−0.834	−0.8476
q(H)	+0.40	+0.41	+0.417	+0.4238

Molecular Dynamics (MD) Simulations

Why?

Static → Dynamic

562 water molecules



Radial Distribution of TIP3P Water

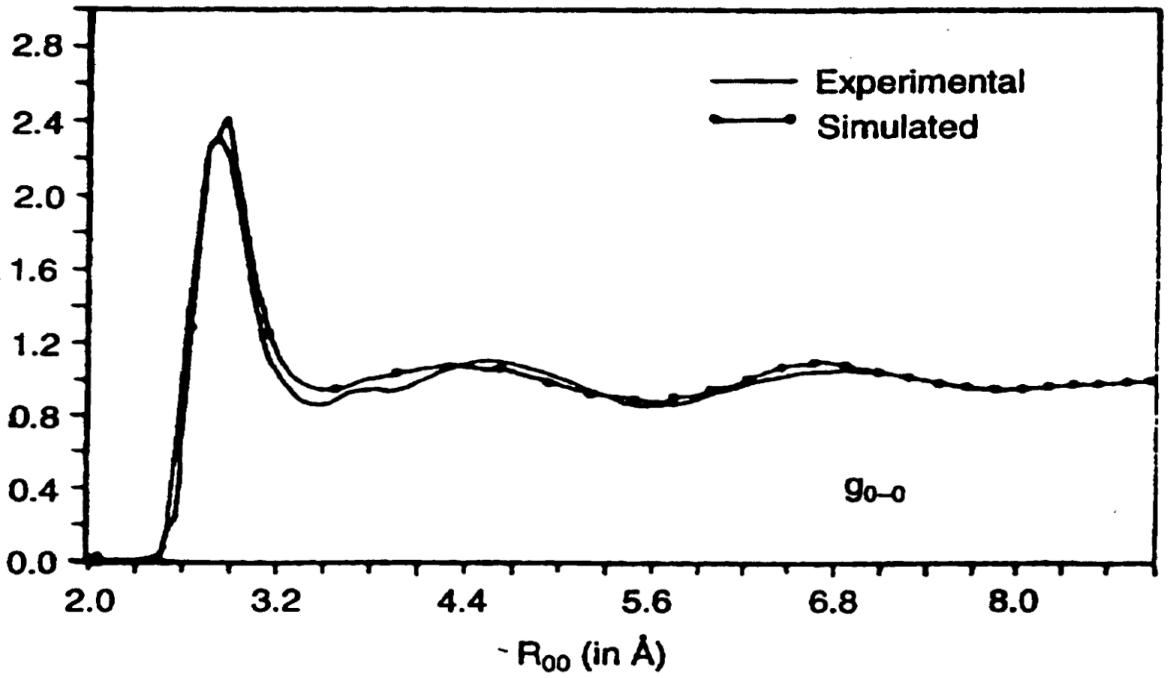
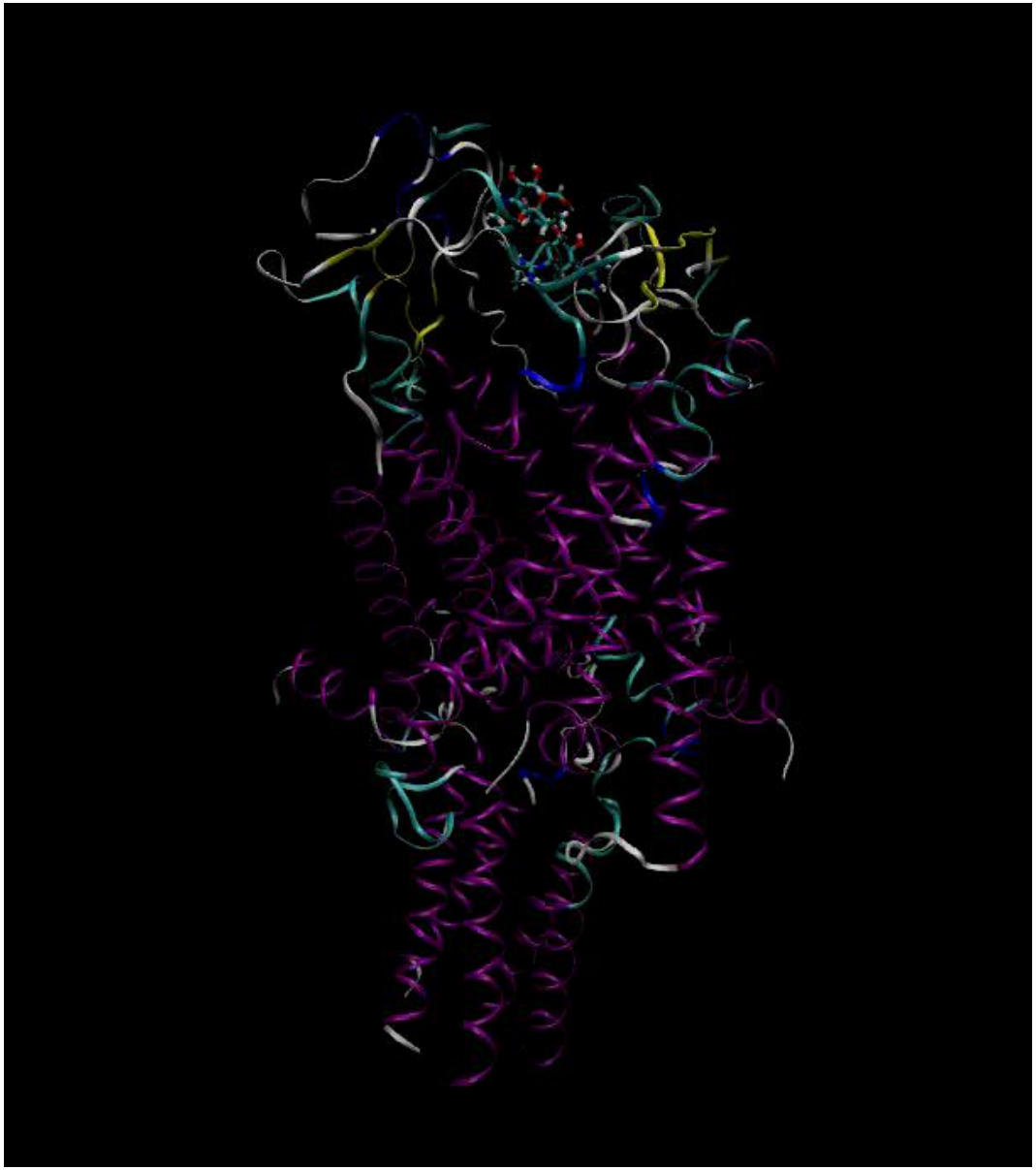


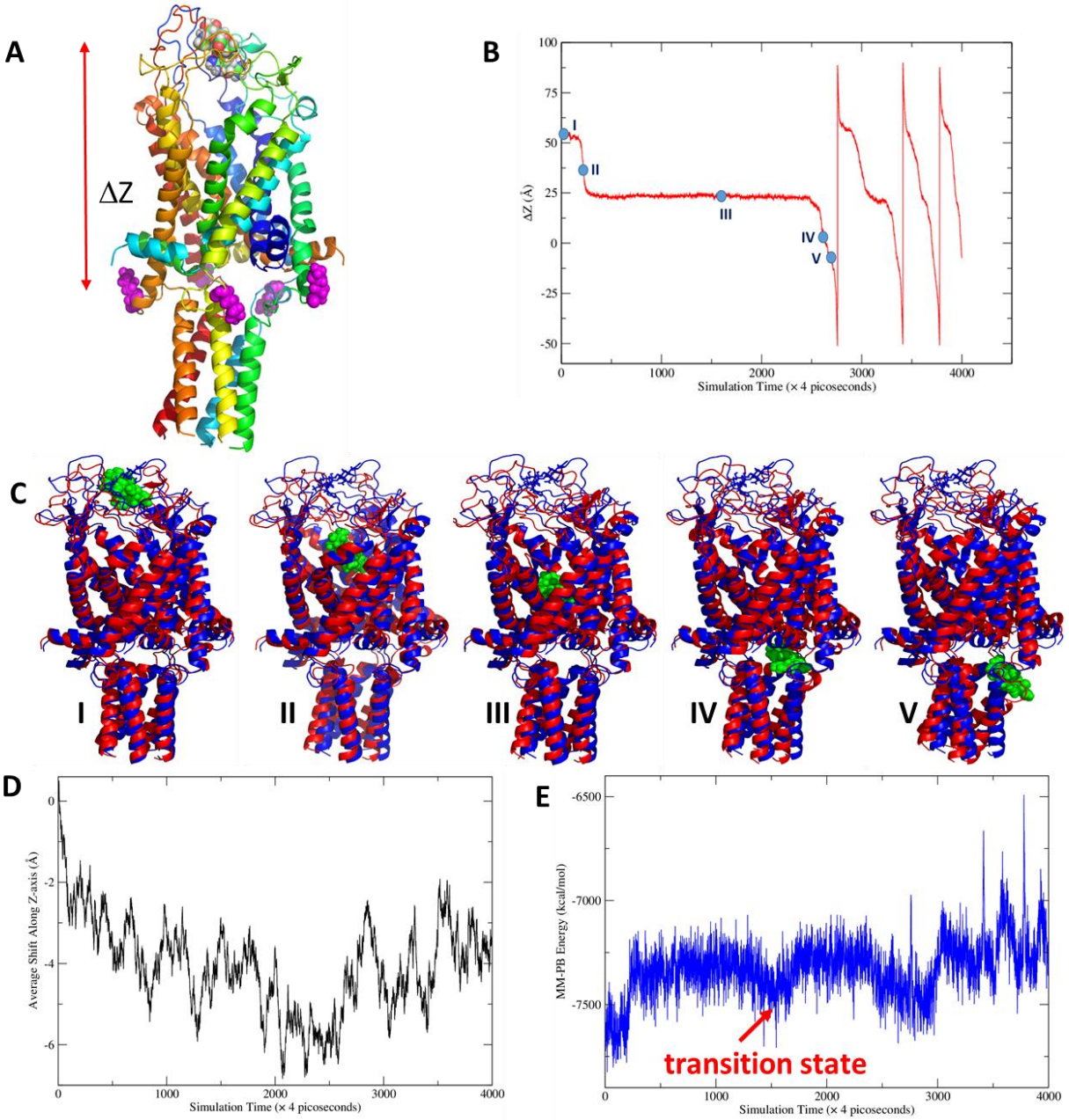
Figure 6.4 Comparison between simulated and experimental O–O radial distribution functions of liquid water (from Lie *et al.* with permission [12]).

MD Simulation of Streptomycin Passing Through the MscL Channel

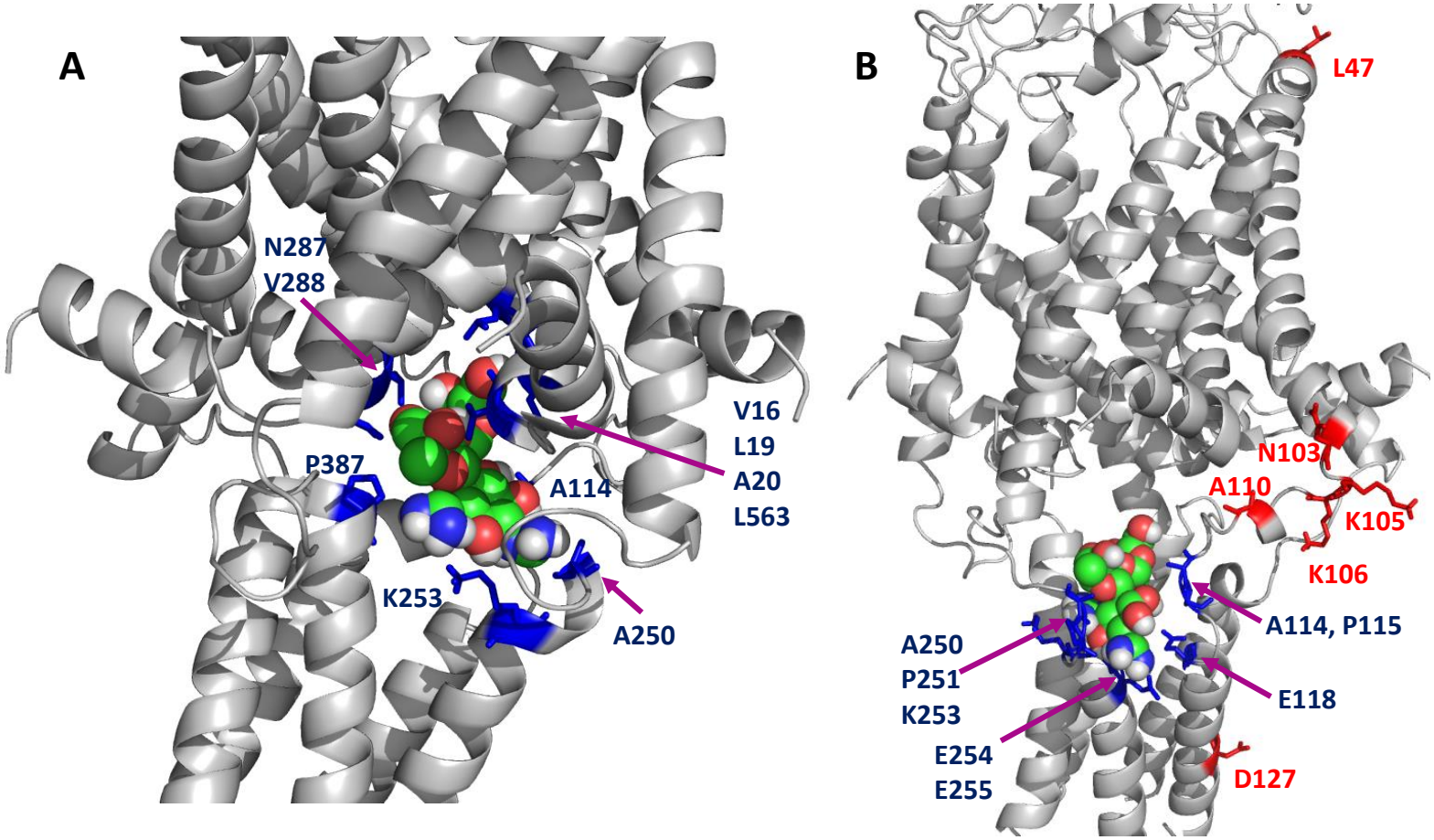


$EEF_X = 0$
$EEF_Y = 0$
$EEF_Z = 0.2$ Volt/Å

“Reaction Coordinate” of DHS Passing Through the MscL Channel



Hot-Spot Residues



Competition Assay (Blind Test)

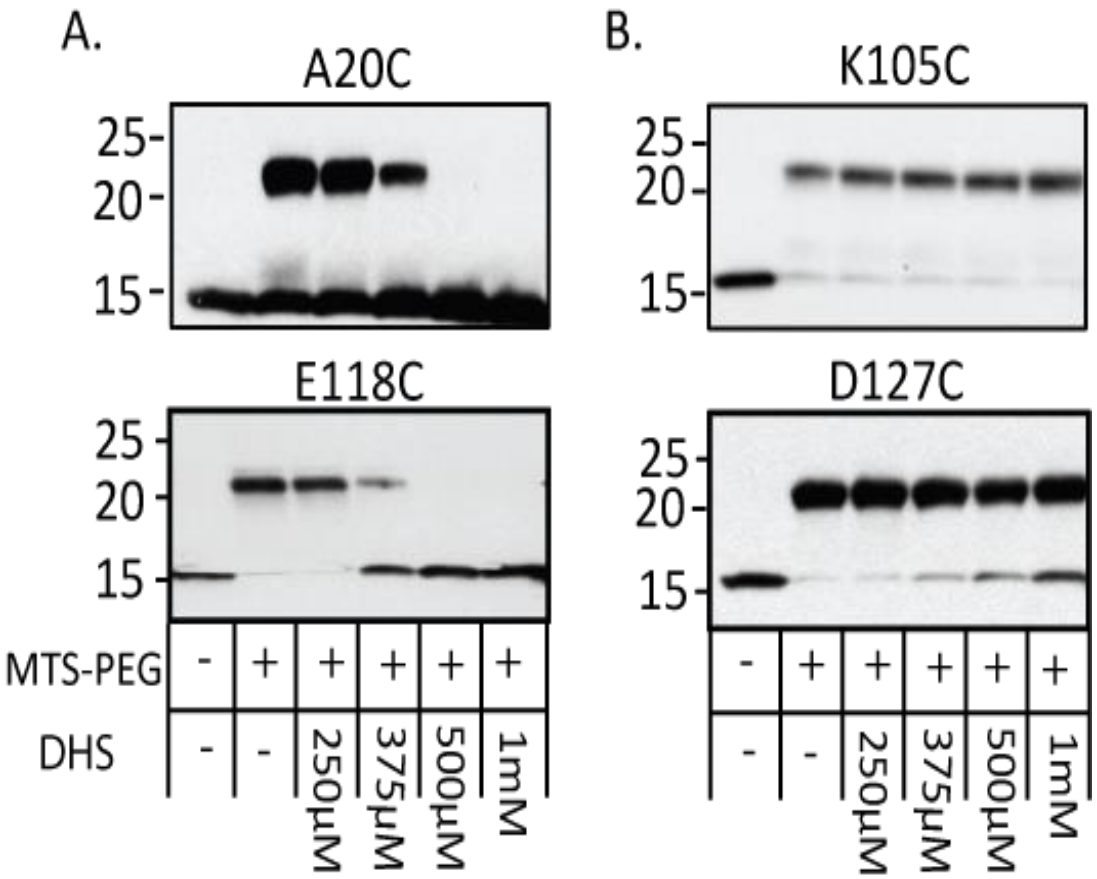
Assay conducted by Blount Lab

Amino Acid	Residue ID	Positive/Negative in Competition Assay
VAL	16	Somewhat positive
VAL	17	Somewhat positive
LEU	19	Positive
ALA	20	Positive
ALA	114	Slightly positive
PRO	115	Positive
THR	116	Slightly positive
LYS	117	Slightly positive
GLU	118	Positive
ASP	18	Positive
VAL	23	Positive
GLU	119	Not Done

Amino Acid	Residue ID	Positive/Negative in Competition Assay
MET	1	Cannot be done
SER	2	Not done
ASN	103	Positive*
LYS	105	Negative
LYS	106	Negative
ASN	127	Negative
GLU	131	Will do assay
ASN	134	Will do assay
ARG	135	Will do assay

*This residue gets buried upon MscL gating

Competition Assay



Positive **Negative**

Robin Wray, Irene Iscla, Junmei Wang* and Paul Blount*, Dihydrostreptomycin directly binds to, modulates, and passes through the MscL channel pore, ***PLOS Biology*, In Press**

DHS Directly Binds to, Modulates, and Passes through the MscL Channel Pore

PLOS | BIOLOGY

RESEARCH ARTICLE

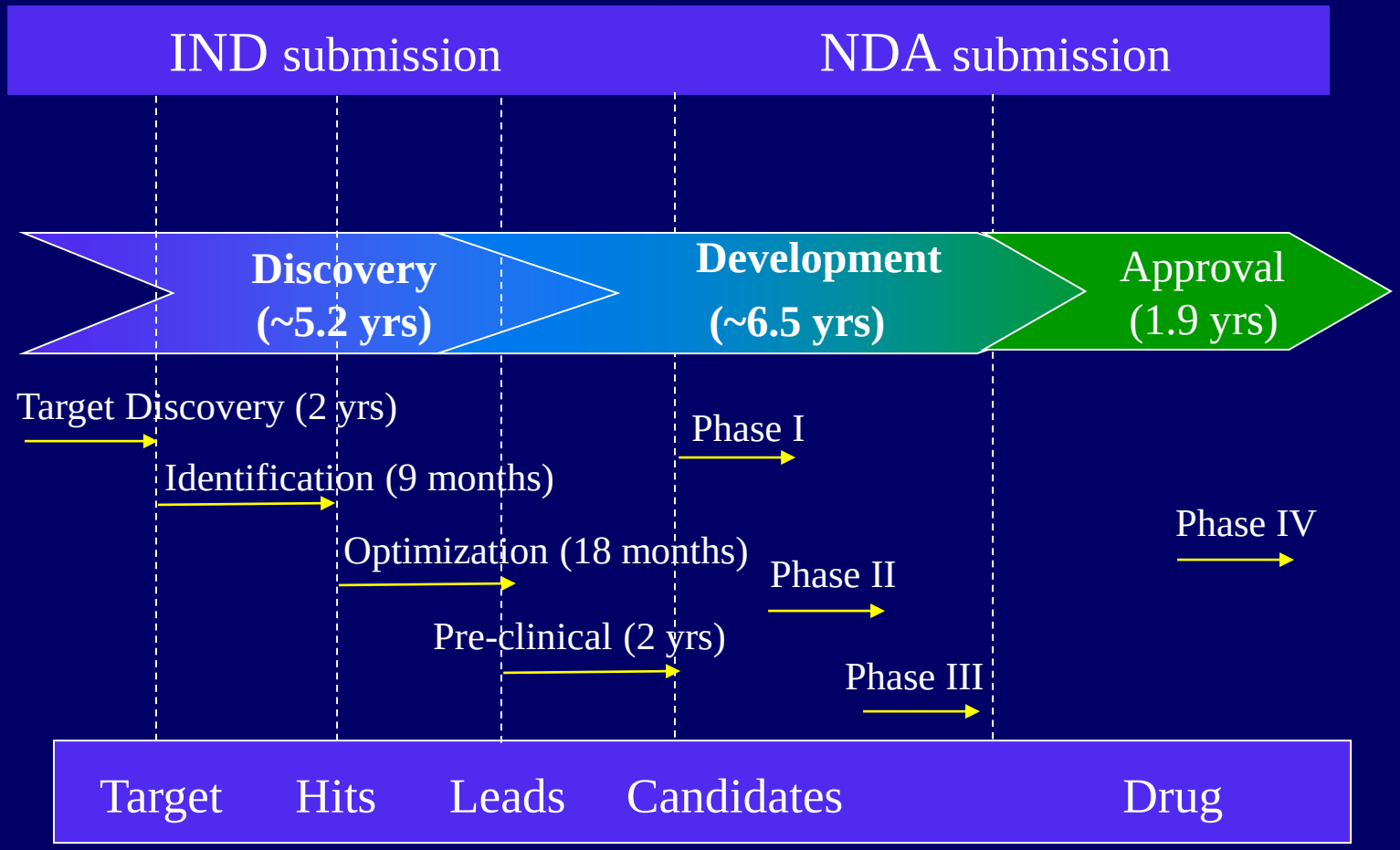
Dihydrostreptomycin Directly Binds to, Modulates, and Passes through the MscL Channel Pore

Robin Wray[1☯], Irene Iscla[1☯], Ya Gao[2], Hua Li[2,3], Junmei Wang[4]*, Paul Blount[1]*

1 Department of Physiology, University of Texas Southwestern Medical Center, Dallas, Texas, United States of America, **2** School of Pharmacy, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, **3** School of Traditional Chinese Materia Medica, Shenyang Pharmaceutical University, Shenyang, China, **4** Green Center for Systems Biology and Biophysics, University of Texas Southwestern Medical Center, Dallas, Texas, United States of America

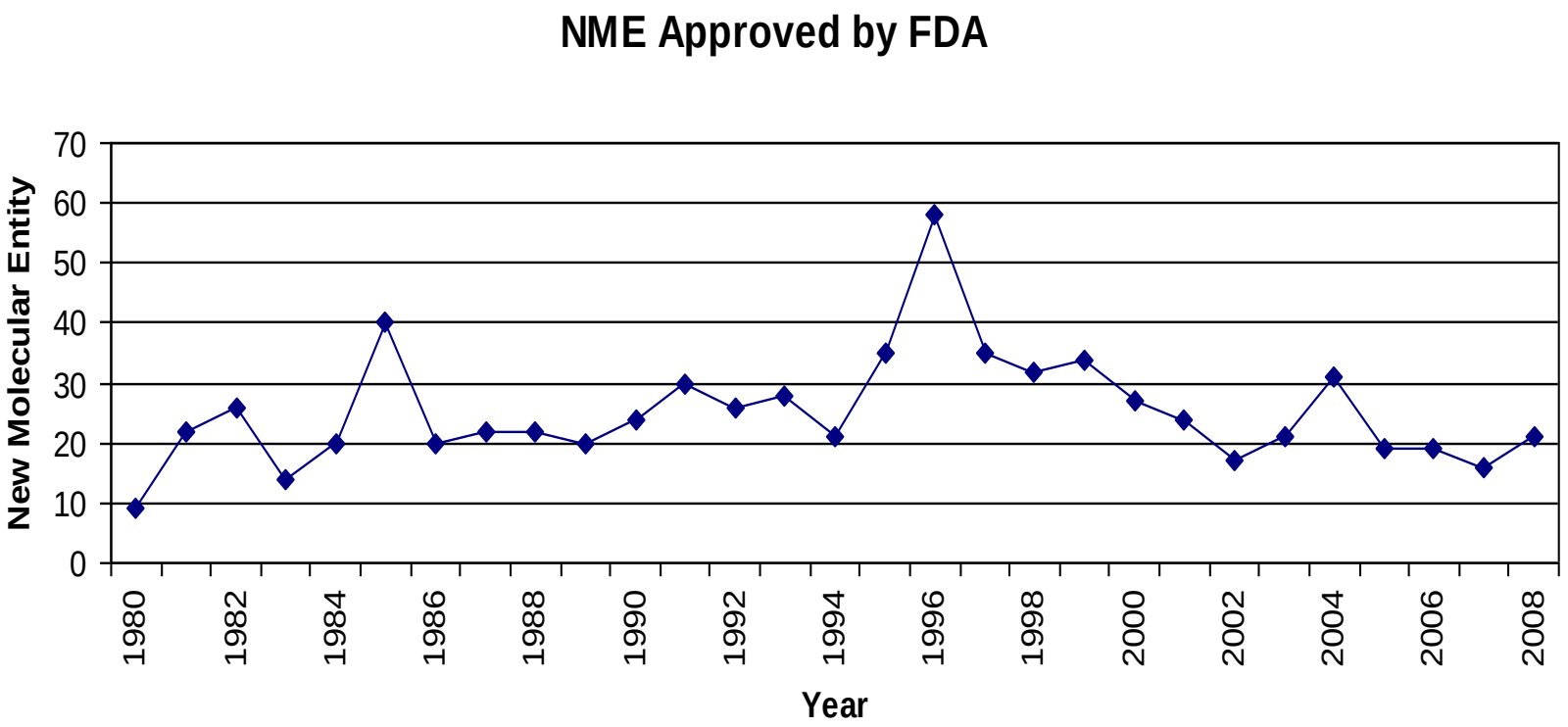
☯ These authors contributed equally to this work.
* Junmei.wang@utsouthwestern.edu (JW); paul.blount@utsouthwestern.edu (PB)

It is getting more difficult to bring a drug into market



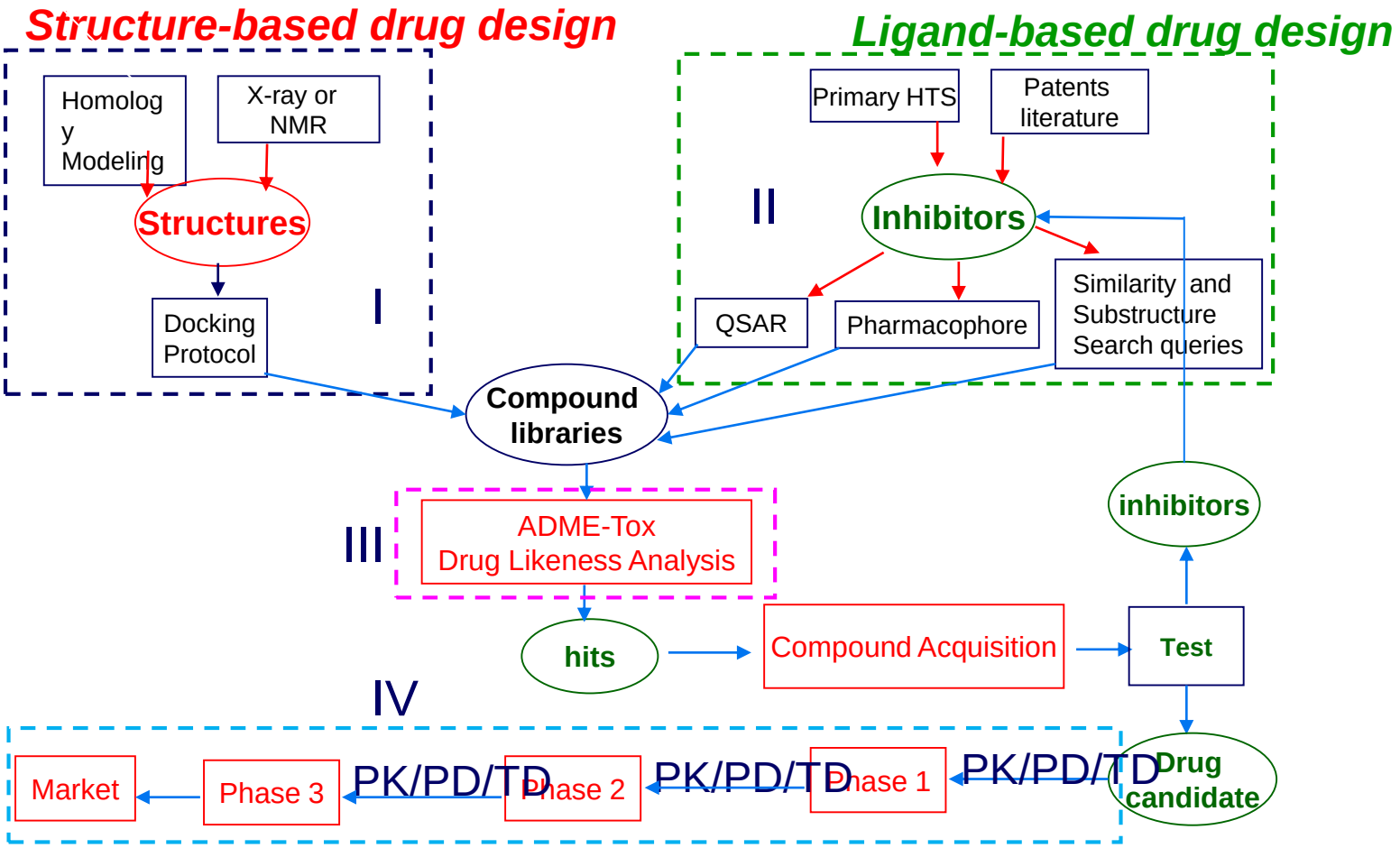
average 13.6 years, $900M spending

Survey of NME Approved Annually by FDA

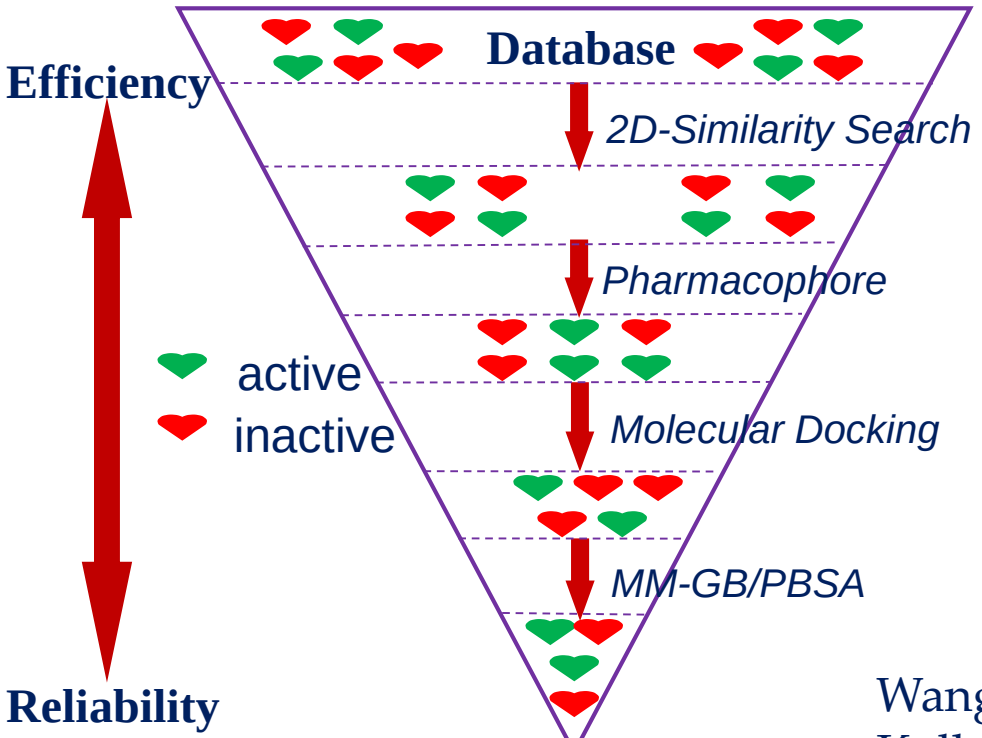


http://www.accessdata.fda.gov/Scripts/cder/DrugsatFDA/

Computer-Aided Drug Discovery And Development

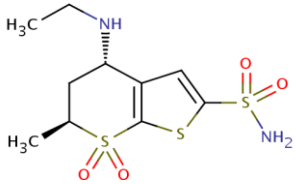


Lead Identification Through Virtual Screening Using A Set of Hierarchical Filters

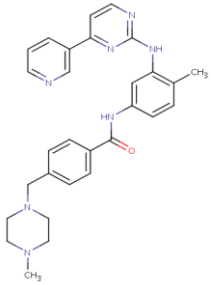


Wang, J.*; Kang, X.; Kollman, P. A.; Kuntz, I. D. *J. Med. Chem.* . 48, 2432-2444

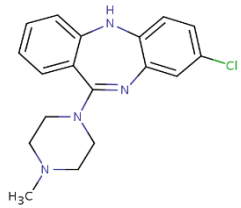
Approved Drugs Designed with CADD



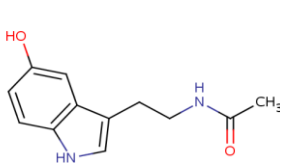
Dorzolamide

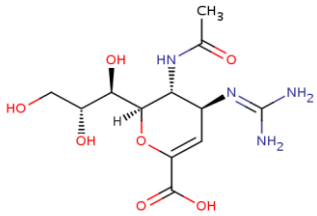


Imatinib

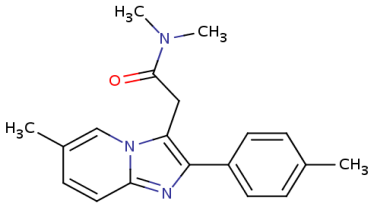


Clozapine

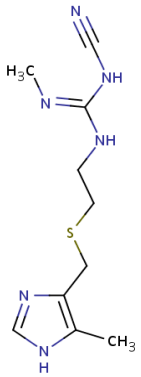


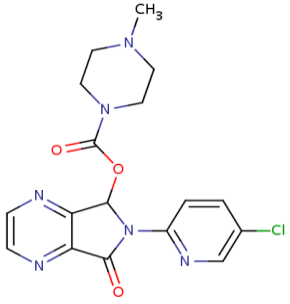


N-Acetyl Serotonin Zanamivir

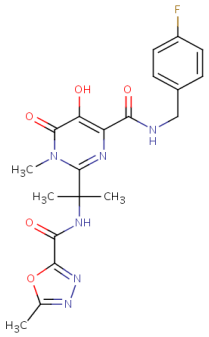


Zolpidem

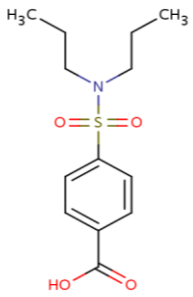




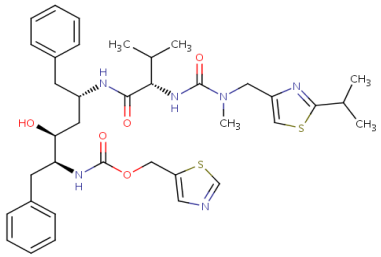
Cimetidine Zopiclone



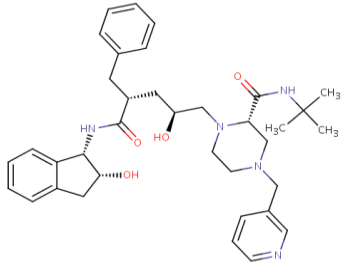
Raltegravir



Probenecid



Ritonavir



Indinavir

Main Stream Molecular Modeling And Molecular Simulation Software

- **Drug Design**
  1. Sybyl (www.certara.com)
  2. Schrodinger (www.schrodinger.com)
  3. Discovery studio (www.accelrys.com)
- **Molecular Docking**
  1. Dock，free http://dock.compbio.ucsf.edu/
  2. Autodock，free www.scripps.edu/mb/olson/doc/autodock/
  3. FlexX， BioSolveI. www.biosolveit.de/flexx
- **Molecular Simulation**
  1. AMBER， www.ambermd.org
  2. CHARMM, www.charmm.org
  3. NAMD, free http://www.ks.uiuc.edu/Research/namd/
- **Graphics**
  1. Weblab Viewlite，Free
  2. VMD，free www.ks.uiuc.edu/Research/vmd/
  3. Pymol，free pymol.sourceforge.net
  4. Chimera, free www.cgl.ucsf.edu/chimera/

Main Journals For Molecular Modeling

1. Journal of Chemical Information and Modeling
2. Journal of Computational Chemistry
3. Journal of Computer-aided Molecular Design
4. Journal of Molecular Graphics and Modeling
5. Journal of Molecular Modeling
6. Bioinformatics
7. PLoS Computational Biology
8. Journal of Medicinal Chemistry
9. Nucleic Acids Research
10. Proteins: Structure, Functions & Bioinformatics
11. Journal of Molecular Biology
12. Bioorganic & Medicinal Chemistry
13. Drug Discovery Today

Books on Molecular Modeling

1. **Molecular Modelling. Principles and Applications**. A. R. Leach, Addison Wesley Longman Limited,1996
2. **Essentials of Computational Chemistry - Theories and Models,** 2nd Edition, Christopher J. Cramer
3. **Molecular Modeling and Simulation: an interdisciplinary guide**, 2nd Edition, Tamar Schlick
4. **Exploring Chemistry with Electronic Structure Methods**, 2nd Edition, James B. Foresman and Aeleen Frisch

Useful Links

- MMFFT
- Protein Data Bank
- Molinspiration

Lab Section

- Basic Unix/Linux Commands
- Chemoffice and ChemBio3D
- Pymol