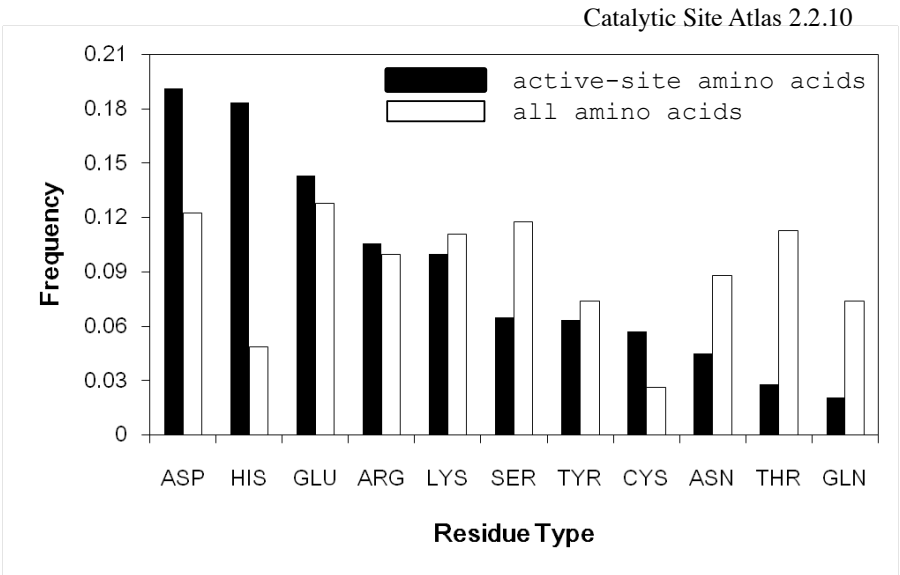
Structural Bioinformatics
or
Computational Structural Biology

黃鎮剛
交通大學
生物資訊所

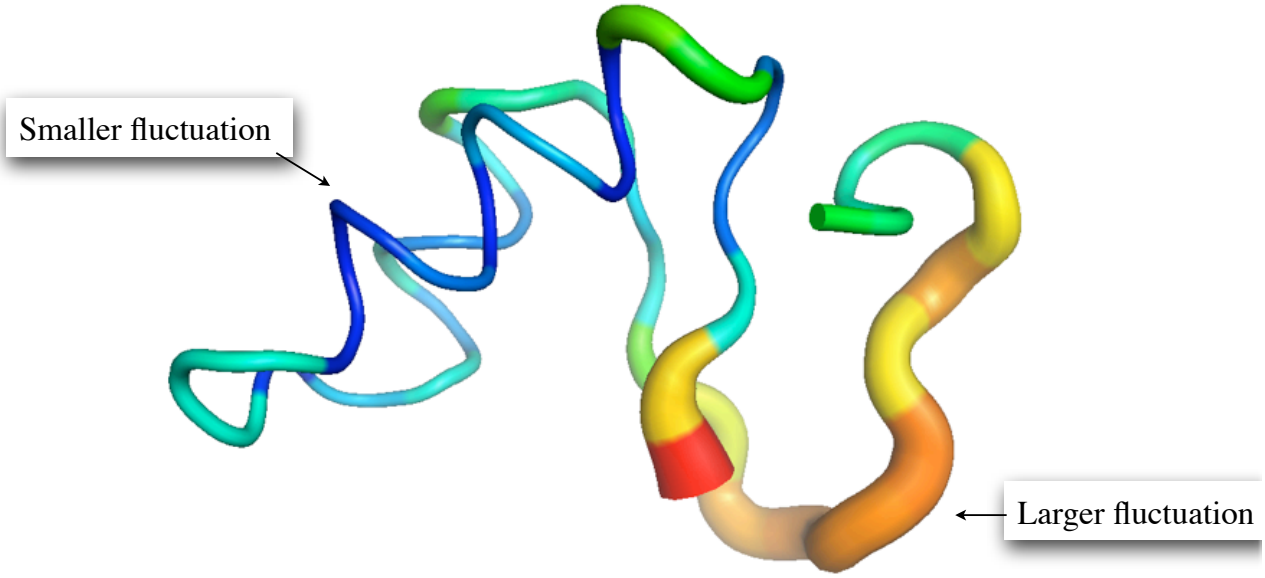
- Molecular Simulation
- Structure prediction
  - 2nd structure prediction; 3D structure prediction; Protein complex prediction
- Protein binding sites prediction
  - Protein - ligand prediction; protein-proetin
- Novel protein structure

From protein structure to active protein

The distribution of amino acid types

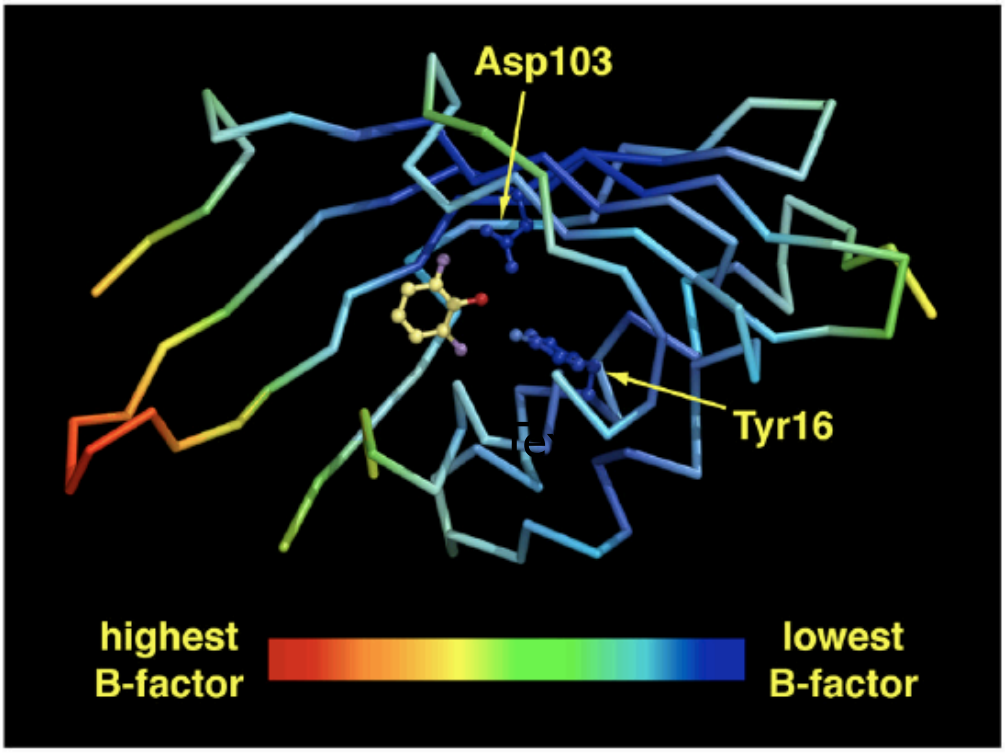


B-factors

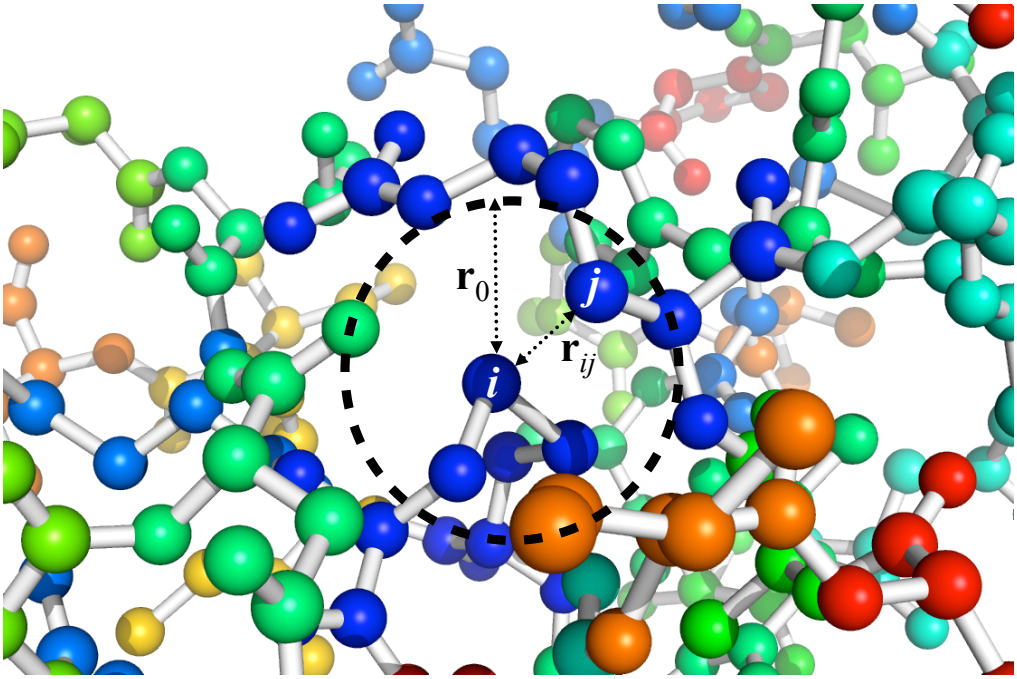


$$B_i = (8\pi^2/3)\langle \delta\mathbf{x}_i \cdot \delta\mathbf{x}_i \rangle$$

C)



JACS (2009) 140: 13696, Testing Geometrical Discrimination within an Enzyme Active Site:



$$v_i = \sum_j h(r_0 - r_{ij}) \qquad (7)$$

Figure 2

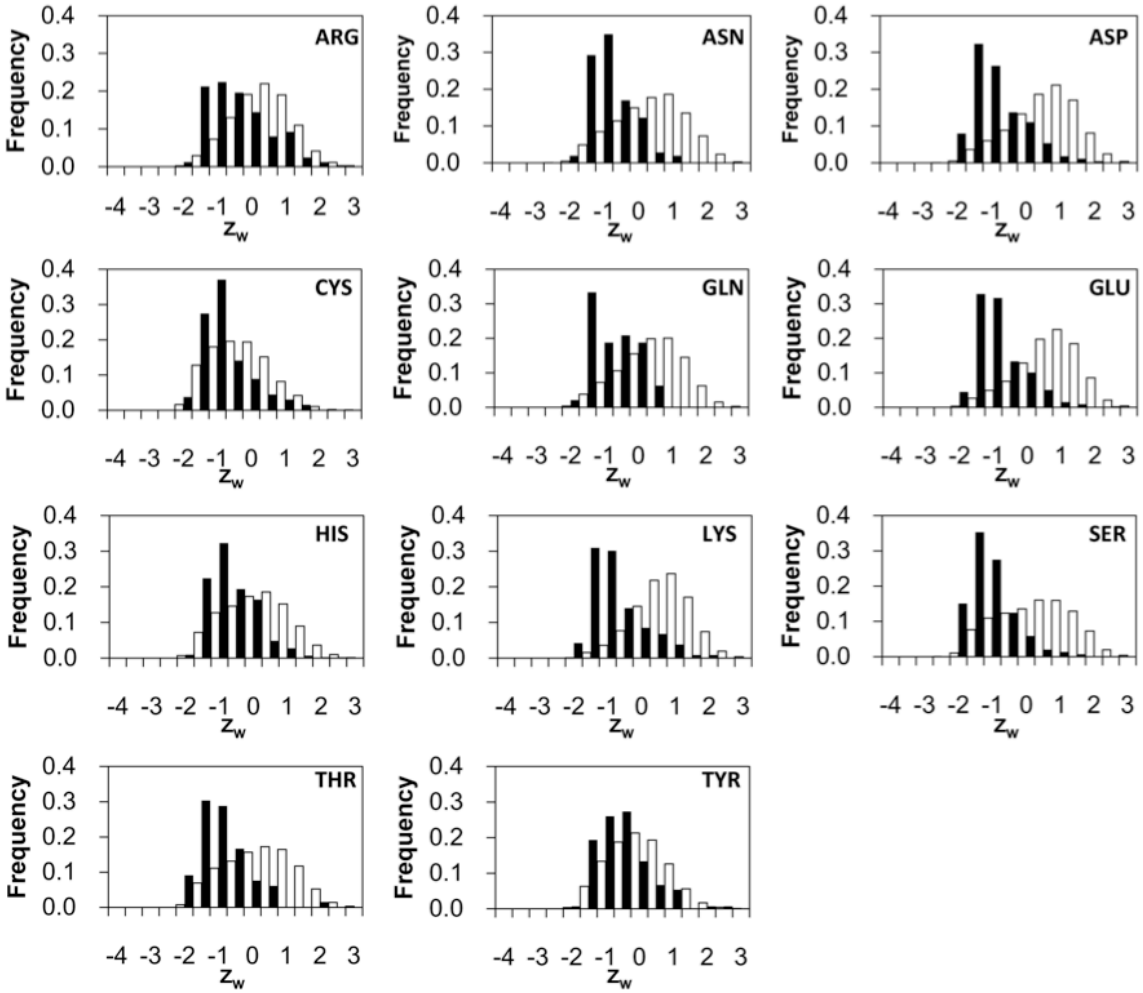


Figure 1

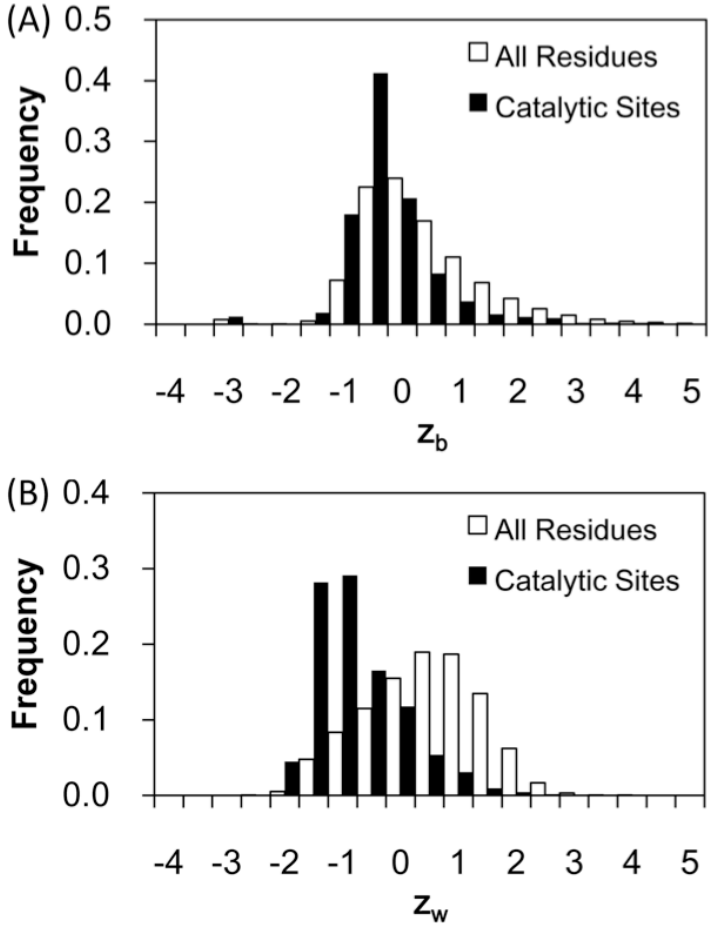


Figure 3

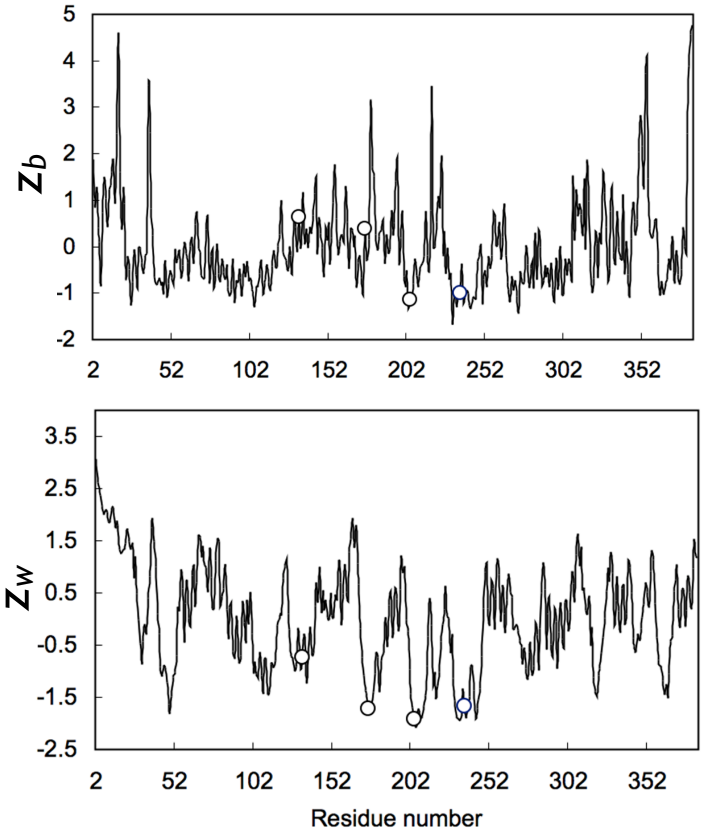
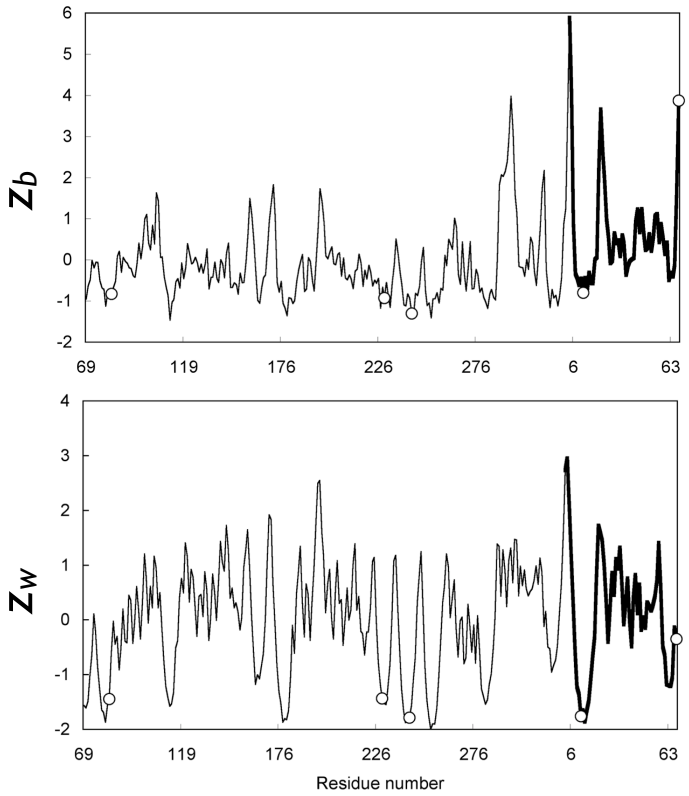
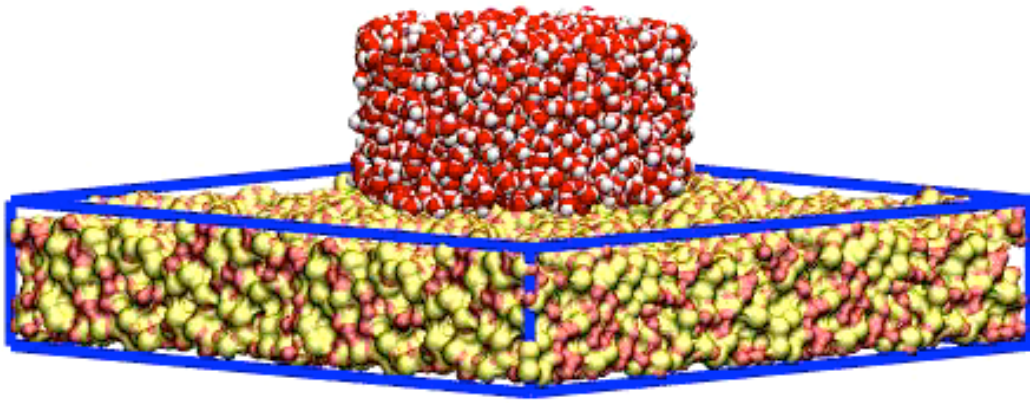
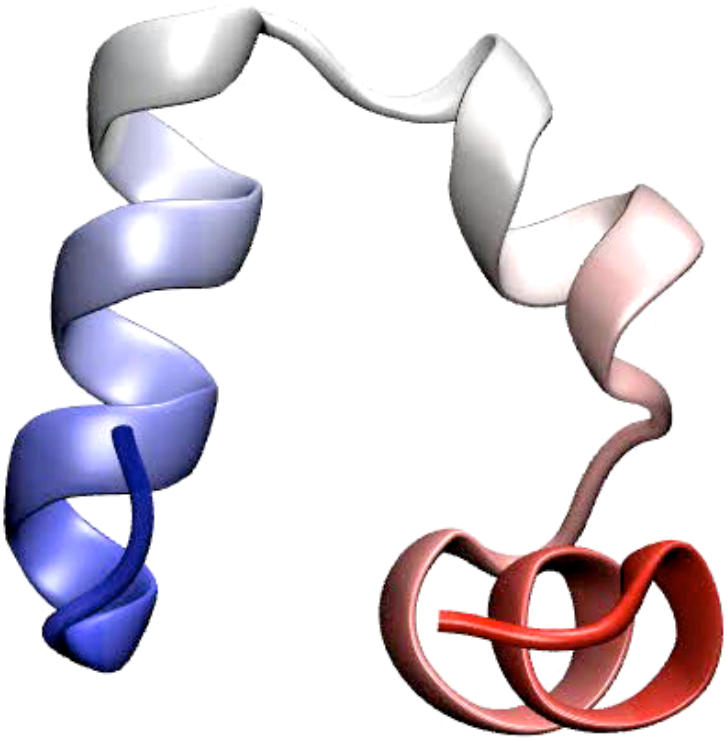


Figure 4





Simulation of water drop in a hydrophobic surface



Villin Folding

Molecular simulation of biological systems

- BIOLOGY
- Computer science
- Physics
- Chemistry
- Statistics
- …

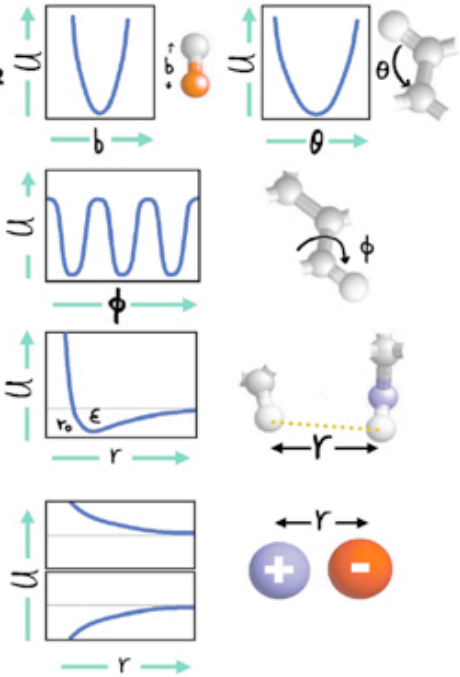
Empirical force field

$$U = \sum_{\text{All Bonds}} \tfrac{1}{2}K_b(b-b_o)^2 + \sum_{\text{All Angles}} \tfrac{1}{2}K_\theta(\theta-\theta_o)^2$$

$$+ \sum_{\text{All Torsion Angles}} K_\phi[1-\cos(n\phi+\delta)]$$

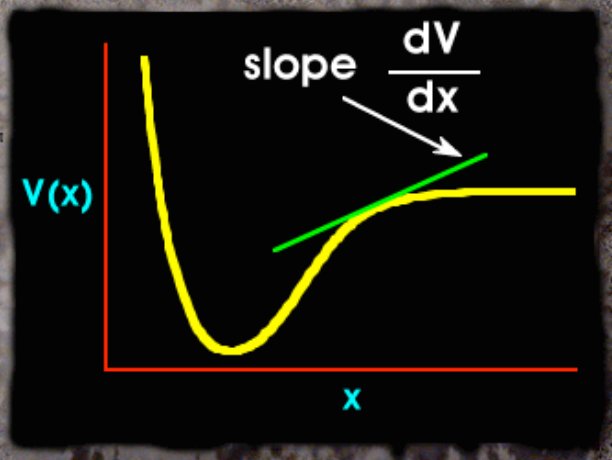
$$+ \sum_{\text{All nonbonded pairs}} \epsilon\left[\left(r_o/r\right)^{12} - 2\left(r_o/r\right)^{6}\right]$$

$$+ \sum_{\text{All partial charges}} 332 q_i q_j / r$$



Levitt. Nature Struct Biol (2001), 8, 392

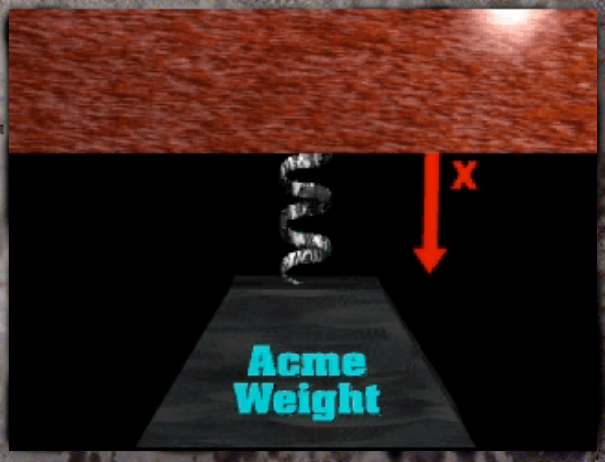
Molecular Dynamics



We need to know the potential energy V(x) to compute the dynamics x(t). How can we describe V(x) for a molecule?

Molecular Dynamics

Newton's 2nd Law

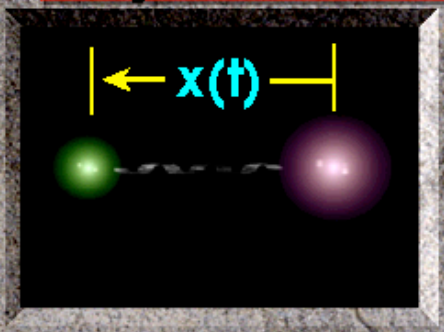


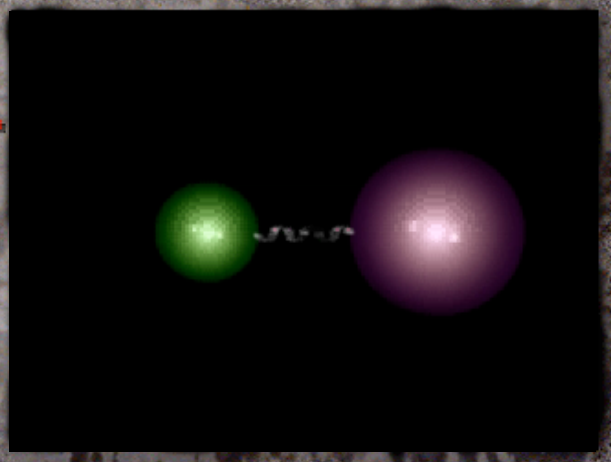
$$-\frac{dV}{dx} = F = ma = m\frac{d^2x}{dt^2}$$

Potential Energy

Position

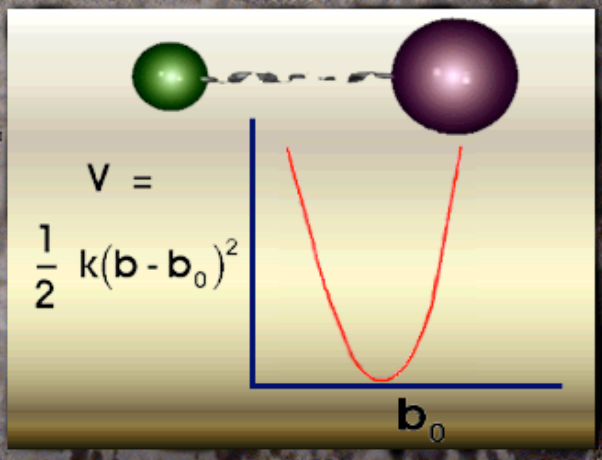
Biomolecular Dynamics





How do we find out the motion of the atoms in a molecule, the x(t)?

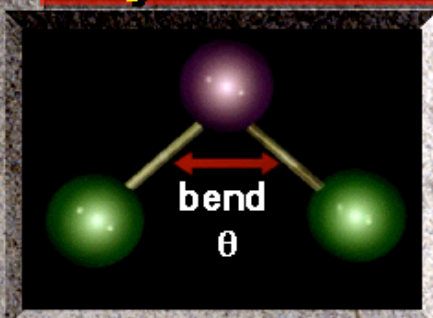
Molecular Dynamics

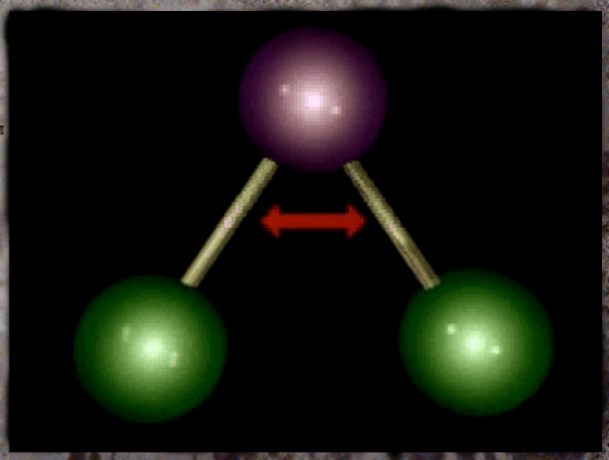


$$V_{bonded} = \sum_{\substack{bonded \\ pair\ of \\ atoms}} \frac{1}{2} k(b - b_0)^2$$

Harmonic Oscillator

Molecular Dynamics

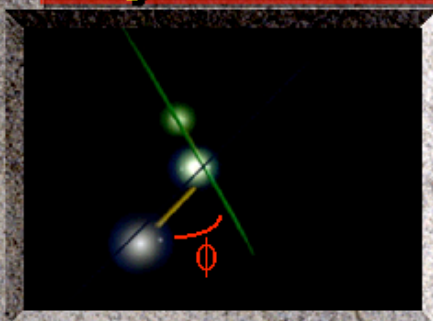


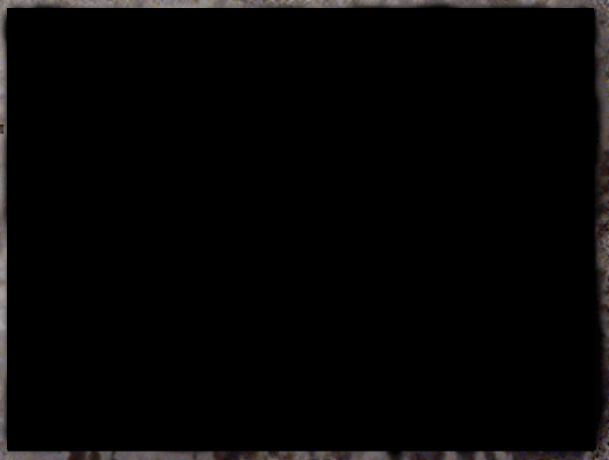


$$V_{bend} = \sum_{\substack{bond \\ angles}} \frac{1}{2} k_\theta (\theta - \theta_0)^2$$

Harmonic Oscillator

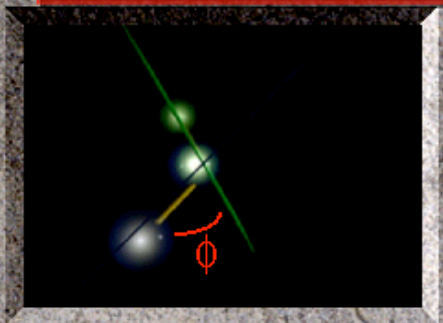
Molecular Dynamics

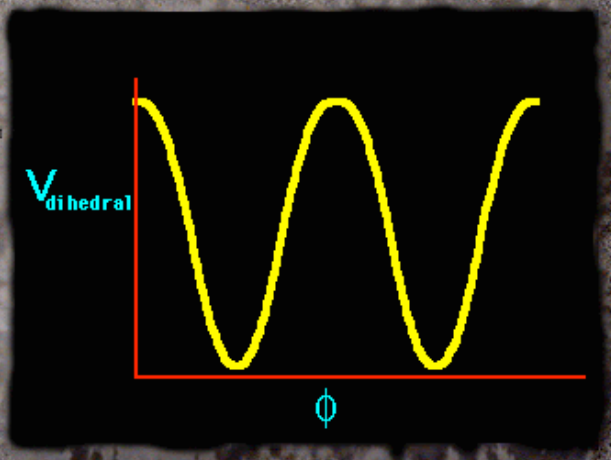




$$V_{dihedral} = \sum_{\substack{dihedral \\ angles}} k_\phi [1 + \cos(n\phi - \delta)]$$

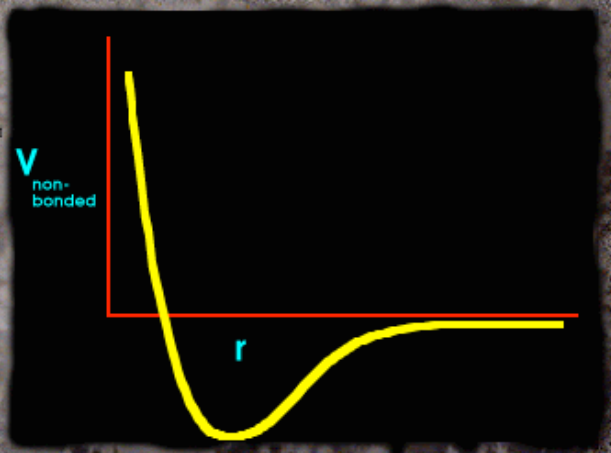
Molecular Dynamics





$$V_{dihedral} = \sum_{\substack{\text{dihedral} \\ \text{angles}}} k_\phi[1+\cos(n\phi-\delta)]$$

Molecular Dynamics



charges on atoms

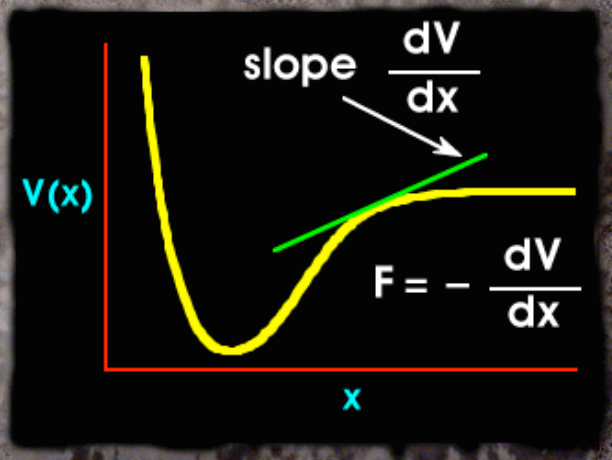
$$V_{nonbonded} = \sum_{\substack{\text{nonbonded} \\ \text{pairs of i, j} \\ \text{atoms}}} V_{nb}(r) + \frac{q_i q_j}{\varepsilon r}$$

dielectric constant

Molecular Dynamics

Force (x) →

$$F(x) = m\frac{d^2x}{dt^2}$$



How can we find out how the atoms in a molecule move? Let's look at a simple example, in which a single atom feels a force F(x) only in the x direction.

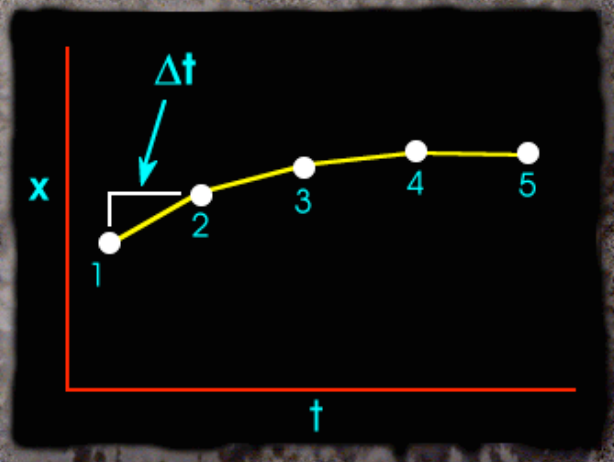
Molecular Dynamics

Force (x) →

$$F(x) = m\frac{d^2x}{dt^2}$$

Now we want to find the trajectory, x(t), of that atom. Assume that at time t, the atom is at position x_1, with velocity $v_1 = \frac{dx_1}{dt}$, and acceleration $a_1 = \frac{d^2x_1}{dt^2} = m^{-1}F(x_1)$.

Molecular Dynamics

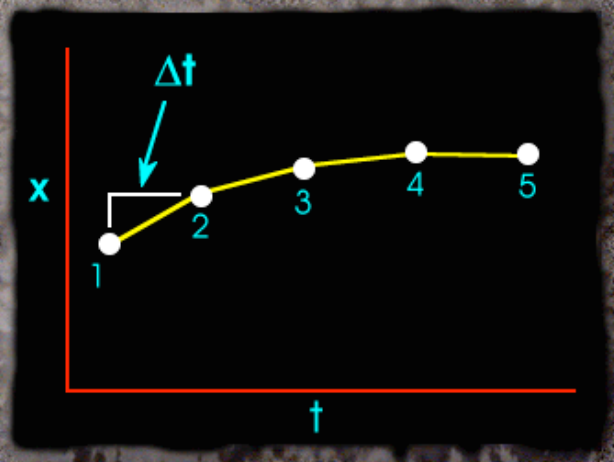


Approximately,

$$x_2 = x_1 + v_1 \Delta t$$

$$v_2 = v_1 + a_1 \Delta t = v_1 + m^{-1} F(x_1) = v_1 - m^{-1} \left. \frac{dV}{dx} \right|_{x_1}$$

Molecular Dynamics



Approximately,

$$x_3 = x_2 + v_2 \Delta t$$

$$v_3 = v_2 + a_2 \Delta t = v_2 + m^{-1} F(x_2) = v_2 - m^{-1} \left. \frac{dV}{dx} \right|_{x_2}$$

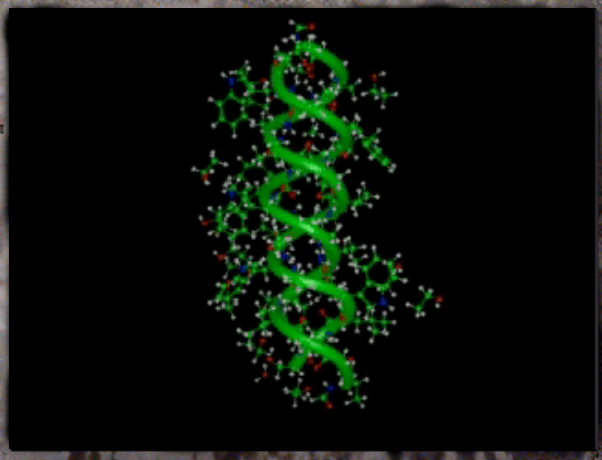
Molecular Dynamics

In general, if we are given the initial position x_1, and velocity v_1, of the atom, and the potential energy $V(x)$ which it feels, we can thus, step by step, compute its positions $x_1(t_1), x_2(t_2), x_3(t_3), \cdots$ by

$$x_i = x_{i-1} + v_{i-1}\Delta t \qquad v_i = v_{i-1} - m^{-1}\left.\frac{dV(x)}{dx}\right|_{x_{i-1}}$$

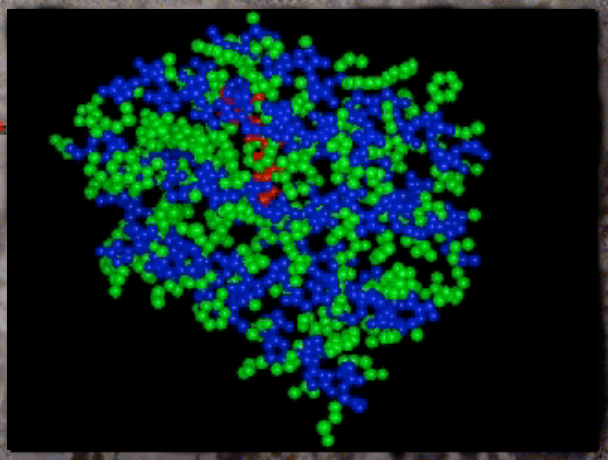
and thus approximately compute its trajectory $x(t)$. We call this process integrating the equations of motion to find the trajectory.

Molecular Dynamics



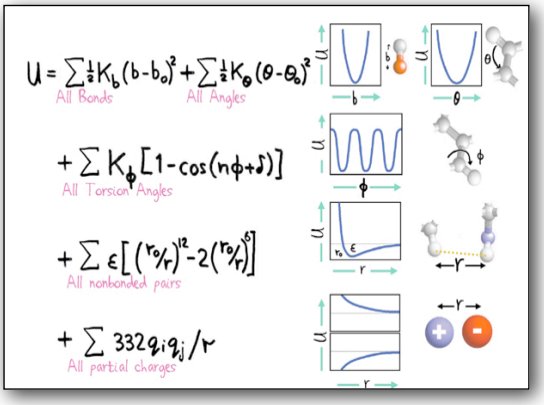
These same ideas may be generalized to compute the trajectories of all the atoms in a biomolecule, if we are given the initial positions and velocities of all the atoms, and the potential energy V of the molecule, such as we just defined, $V = V_{bonded} + V_{bend} + V_{dihedral} + V_{nonbonded}$

Biomolecular Dynamics



In this way, we can at least learn about the short time and distance scale motions involved in chemical reactions, enzyme-substrate interactions, and local interactions of biomolecules with their environment. As computers get faster, and scientists get more clever in using them, the accessible time and distance scales will keep increasing.

Molecular Dynamics



Empirical Force Field

$$H = T + U$$

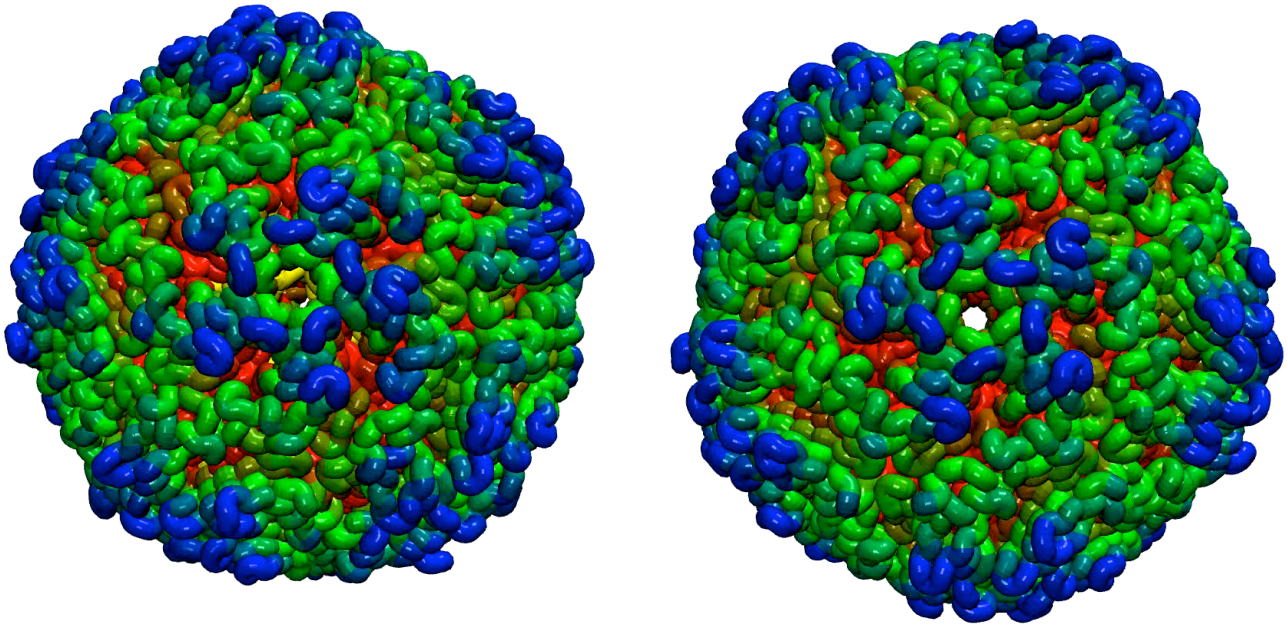
$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}$$

$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}$$

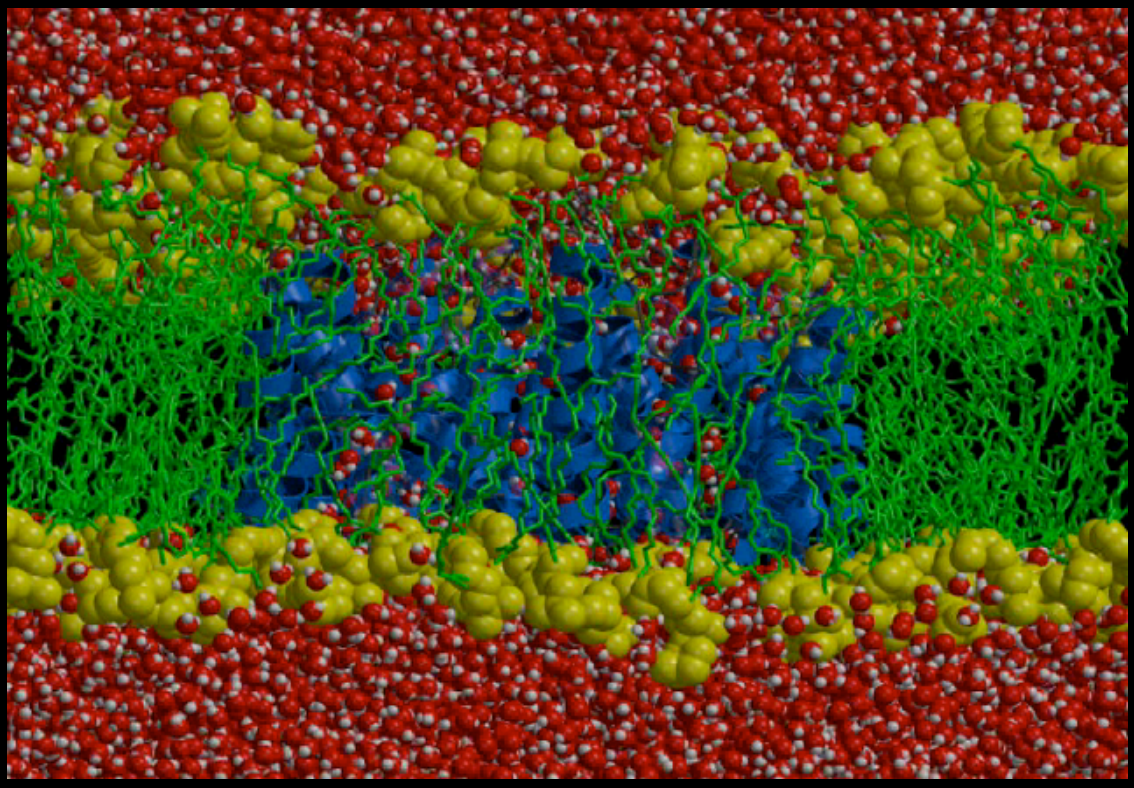
Classical equations of motion

$$\mathbf{q}(t)$$
$$\mathbf{p}(t)$$

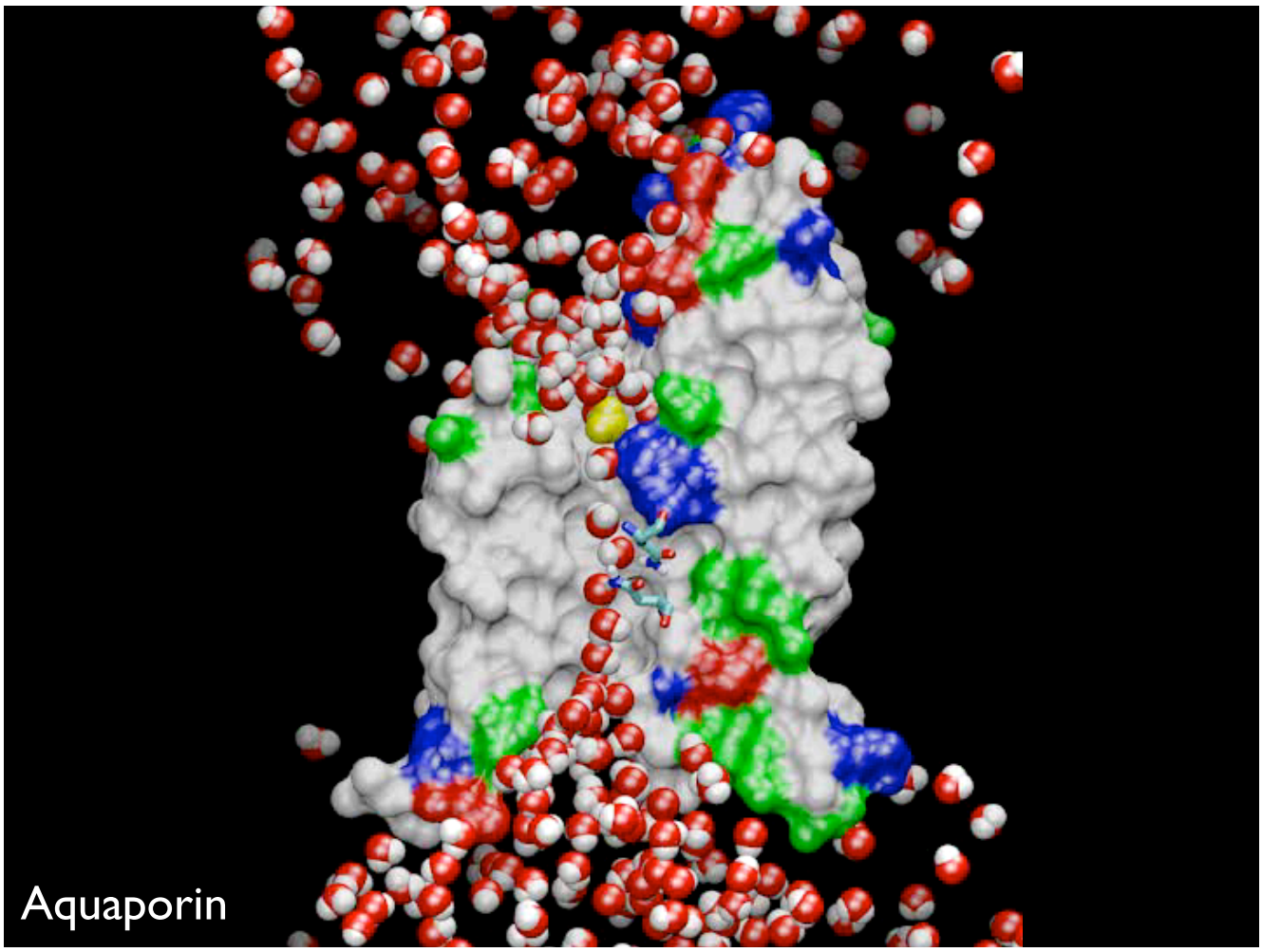
Trajectory integration



Satellite Tobacco Mosaic Virus



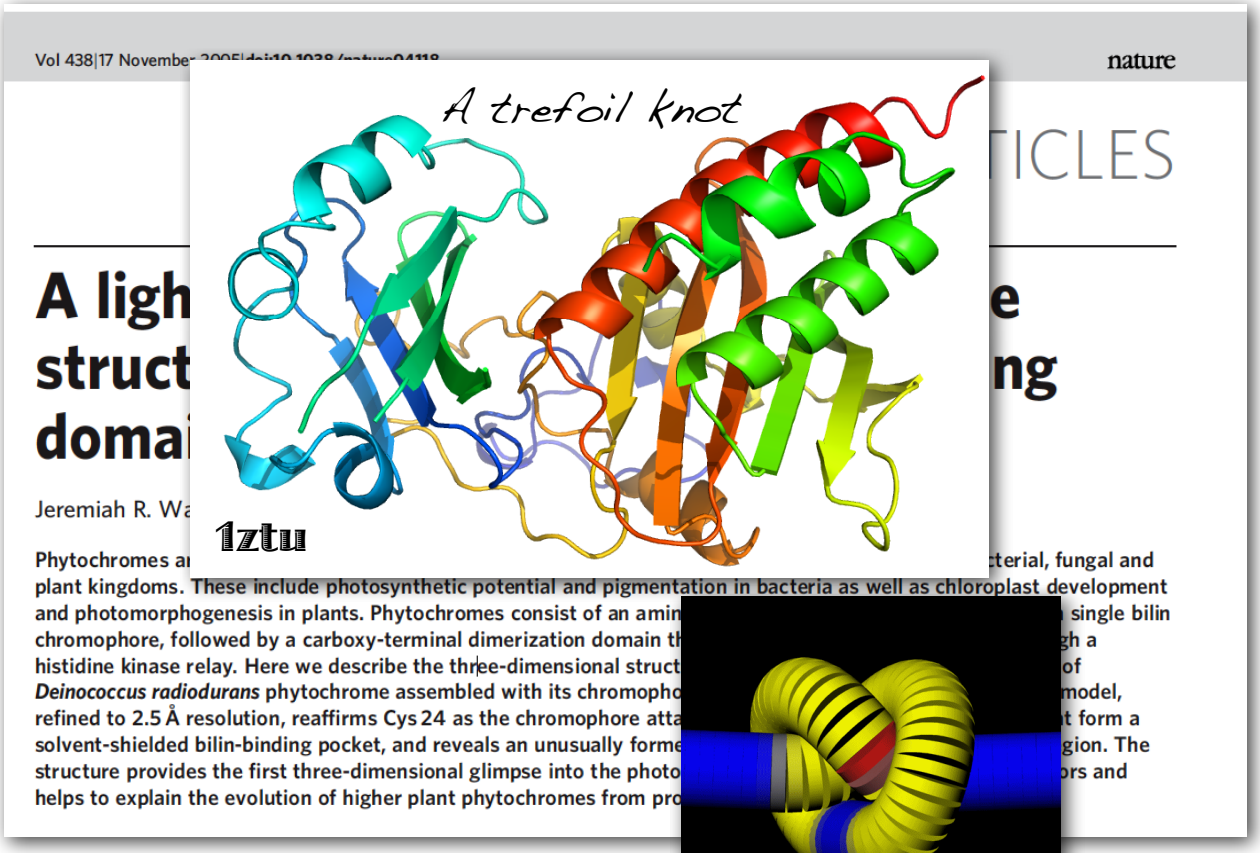
Simulation of protein dynamics



Knotted Proteins

Different types of knots

KnotPlot



A trefoil knot

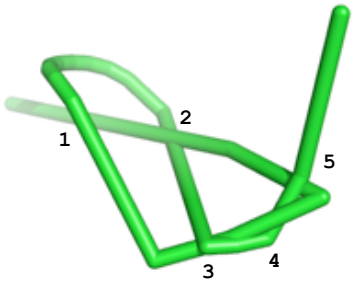
1ztu

Nature (2005) 438, 324

PyMOL

To find knots by a smoothing algorithm

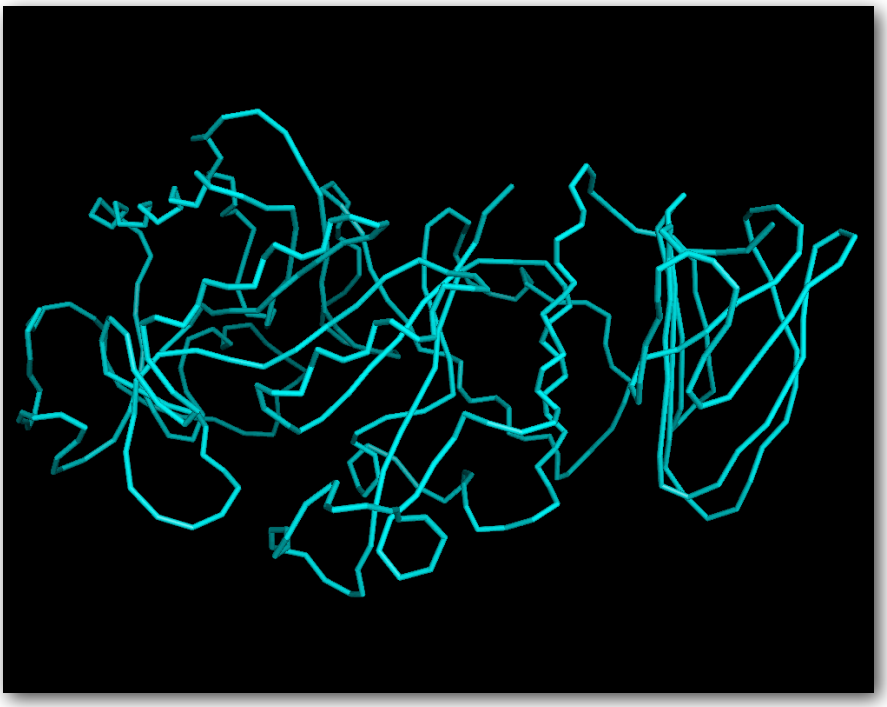
Ubiquitin Hydrolase
1XD3:A



PyMol: show animation

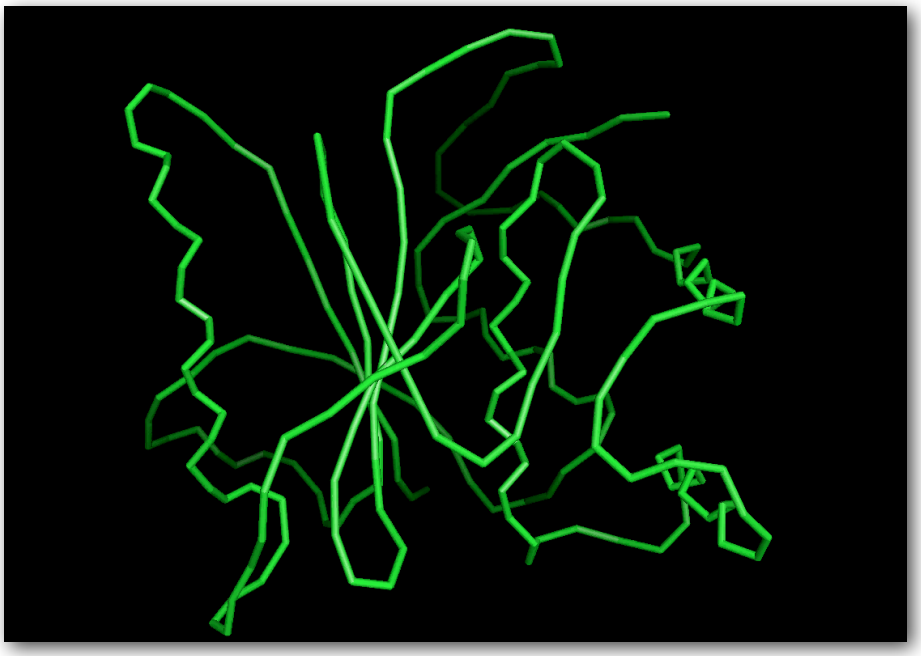
A 5_2 knot

Smooth the protein backbone

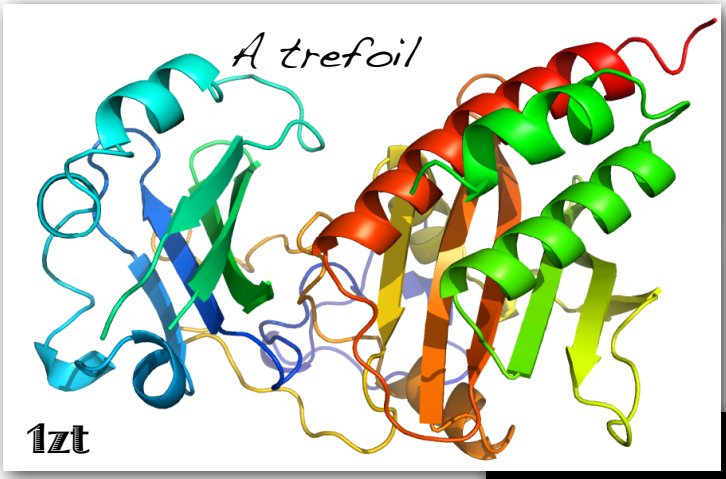


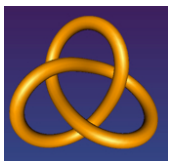
Lai et al. NAR, 2007; Taylor Nature (2000)

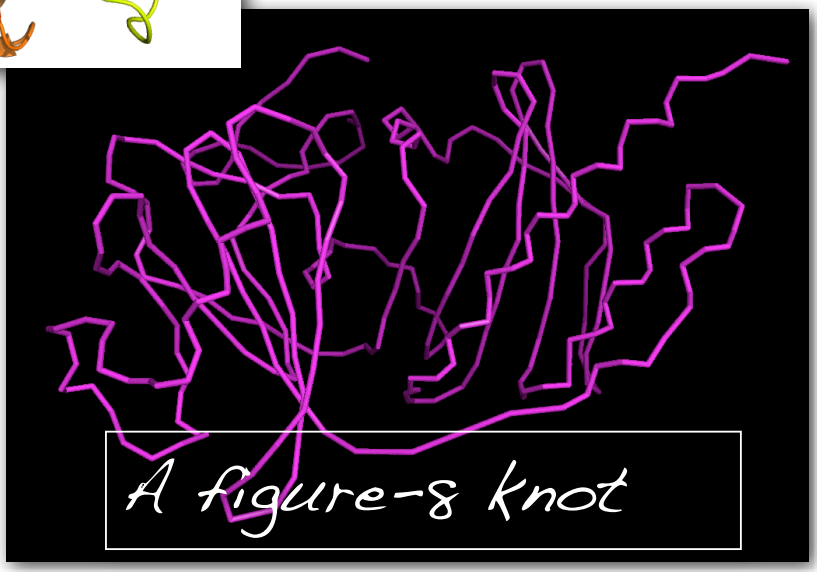
Smooth the protein backbone



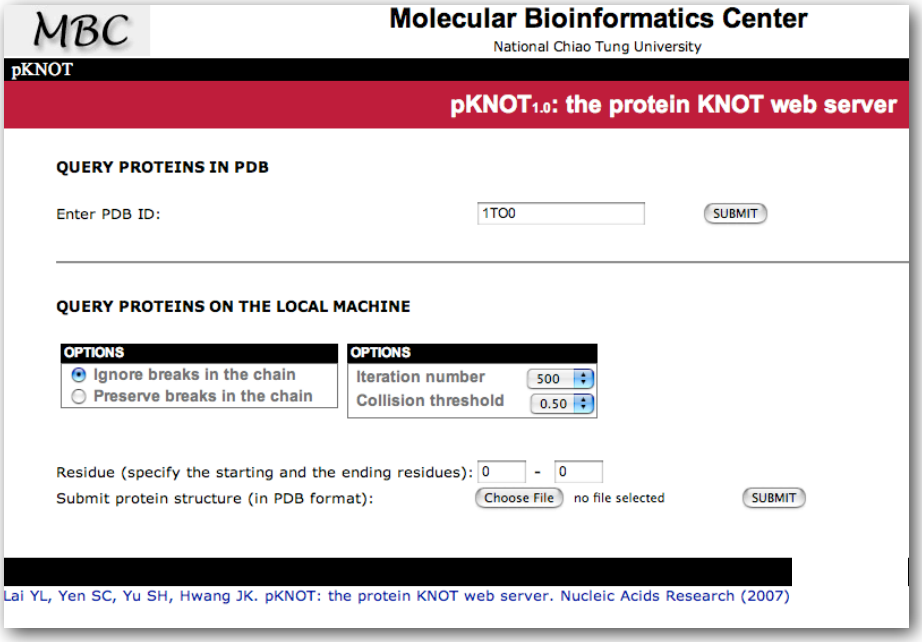
Lai et al. NAR, 2007; Taylor Nature (2000)





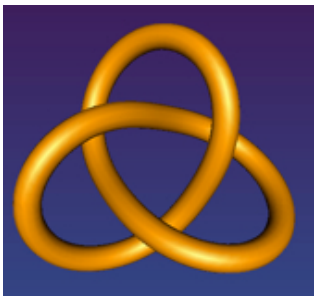


pKNOT: The Protein KNOT Server

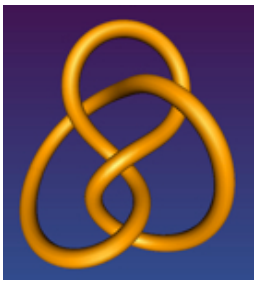


http://pknot.life.nctu.edu.tw/

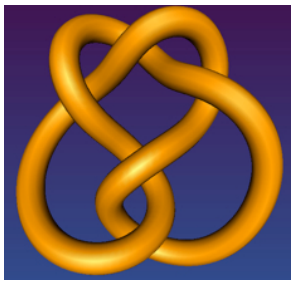
Only 3 types of knot are found in PDB



The trefoil knot



The figure-8 knot



The 5_2 knot

Families of the Knotted proteins

A. The trefoil knot (349)

1. methyltransferase (51)
2. transcarbamylase (20)
3. methionine adenosyltransferase (26)
4. carbonic anhydrase
5. YMPa superantigen (NMR)

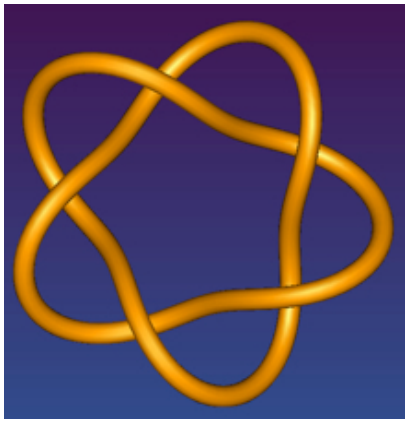
B. The figure-8 knot

1. ketol-acid reductoisomerase
2. the chromophore binding domain of bacterial phytochrome
3. the core protein of bluetongue virus
4. a LIM-Idbl-LID chimeric protein (NMR)

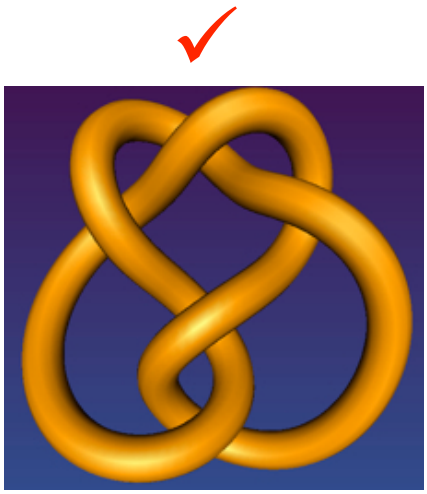
C. The 5_2 knot

1. ubiqutine c-terminal hydrolase

Interestingly, there are 2 types of knots with 5 crossings -- **5_1** and **5_2**, but only **5_2** is found in protein



The 5_1 knot



The 5_2 knot

KnotPlot: to review the knots