上海纽约大学
NYU SHANGHAI

2014 NYUSH Symposium for Undergraduate Research

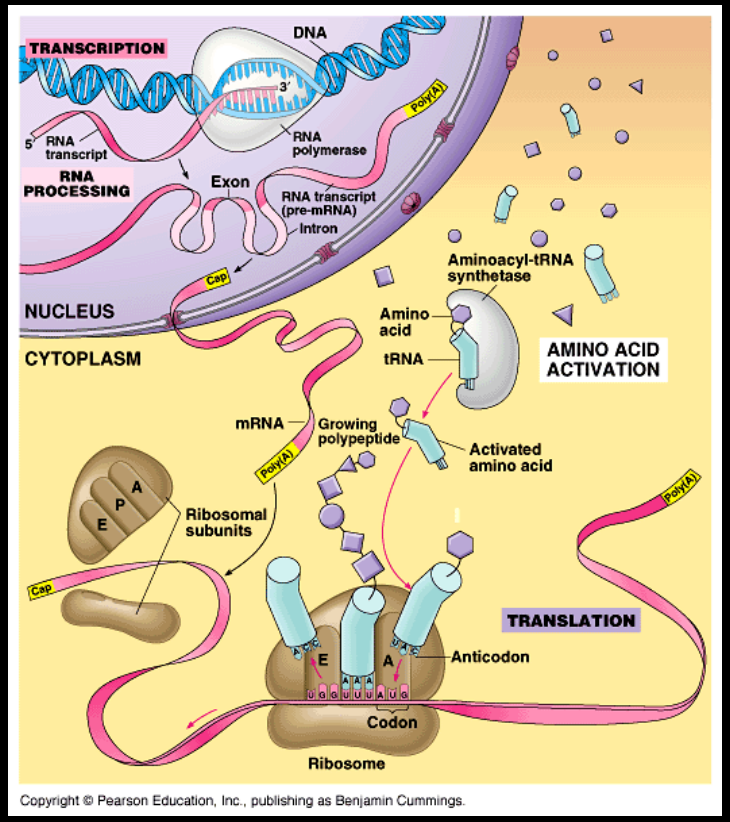
Protein folding

from chemistry to physics

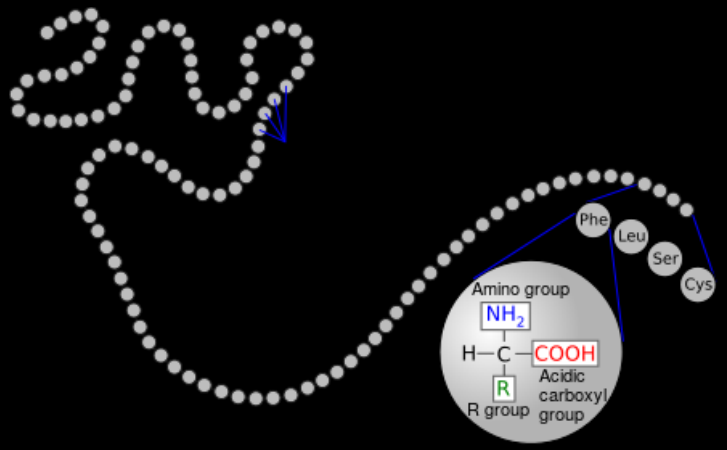
Ye Mei
梅晔
East China Normal University
华东师范大学
NYU-ECNU Center for Computational Chemistry at NYU Shanghai
纽约大学-华东师范大学计算化学联合中心

Dec. 6, 2014

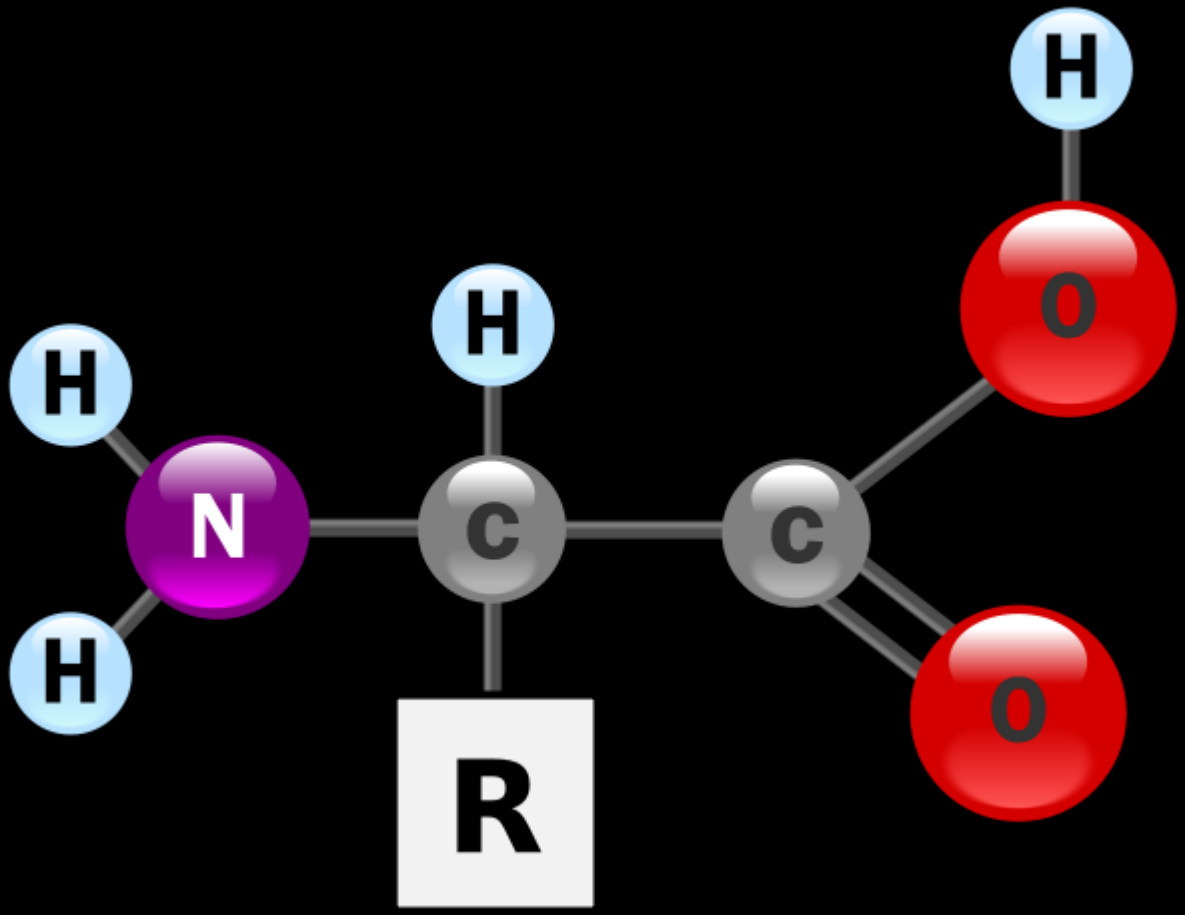
Protein makes up about 16% of our total body weight, which is provided by meat, egg, dairy products, tofu …



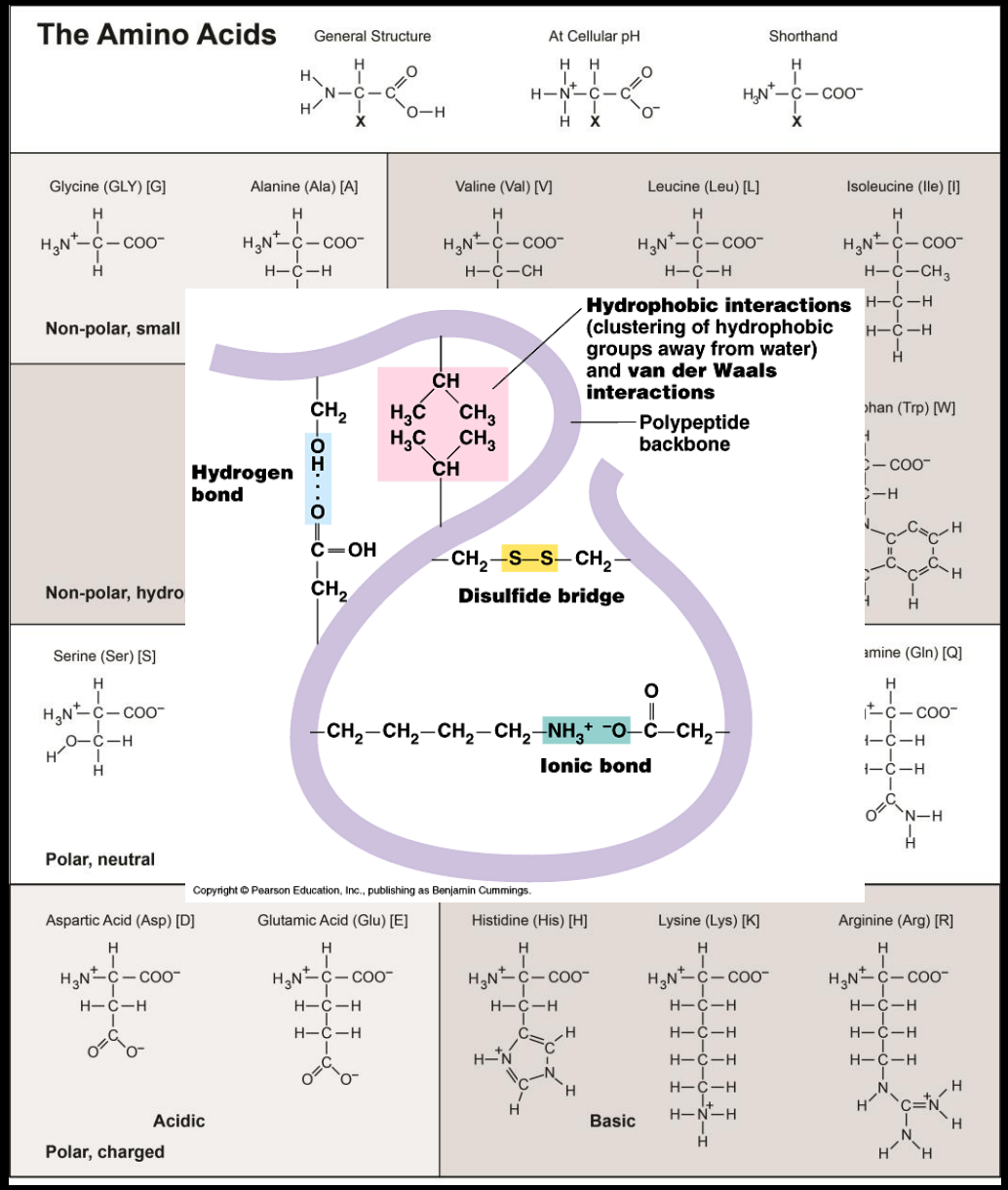
Protein is synthesized in ribosome by decoding genetic information in DNA/RNA



Protein is one type of the natural polymers built from amino acids


H
H
H
N
C
C
O
O
H
R

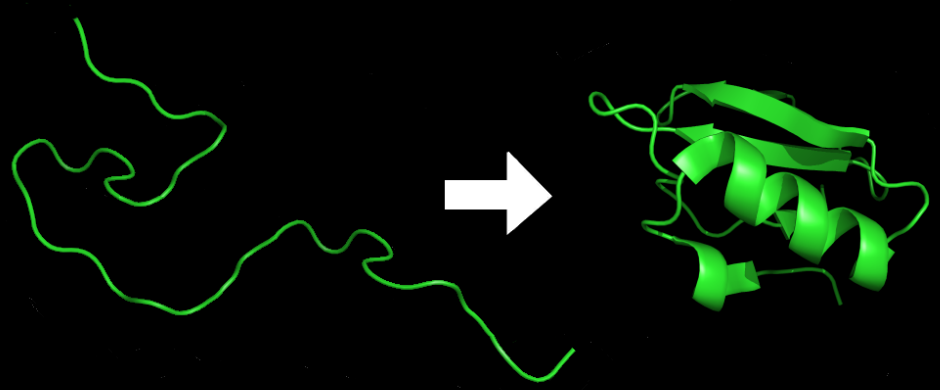
The Amino Acids



Each protein exists as an unfolded polypeptide or random coil when translated from a sequence of mRNA to a linear chain of amino acids. This polypeptide lacks any stable three-dimensional structure.

Protein folding is the physical process by which a polypeptide folds into its characteristic and functional three-dimensional structure from random coil.

Amino acids interact with each other to produce a well-defined three-dimensional structure, the folded protein, known as the native state.

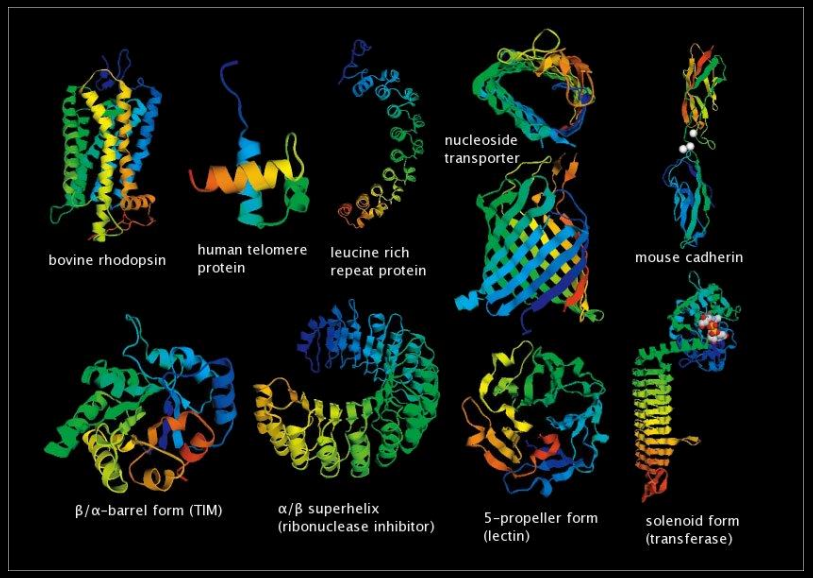


From http://en.wikipedia.org/wiki/Protein_folding

Anfinsen‘s dogma: At least for small globular proteins, the native structure is determined only by the protein's amino acid sequence

-- Christian B. Anfinsen (1961)

The postulate amounts to saying that, at the environmental conditions (temperature, solvent concentration and composition, etc.) at which folding occurs, the native structure is a unique, stable and kinetically accessible minimum of the free energy. It does not depend on its origin.



Methods for determining protein structures:

1. **X-ray Crystallography**
2. **NMR Spectroscopy**
3. **Electron Microscopy**
4. **Computer Simulations**

What if protein folds to a wrong structure?

- Alzheimer's disease, AD

  difficulty in remembering recent events
- Parkinson's disease, PD

  shaking, rigidity, slowness of movement
- Huntington's disease, HD

  cognitive and behavioral symptoms
- Prion disease

  mad cow
- familial amyotrophic lateral sclerosis, ALS

  Ice Bucket Challenge

RCSB PDB | Deposit | Search | Visualize | Analyze | Download | Learn | More | MyPDB Login

RCSB PDB PROTEIN DATA BANK

An Information Portal to 105383 Biological Macromolecular Structures

Search by PDB ID, author, macromolecule, sequence, or ligands

Go

Advanced Search | Browse by Annotations

PDB-101 | wwPDB | EMDataBank

- Welcome
- Deposit
- Search
- Visualize
- Analyze
- Download
- Learn

A Structural View of Biology

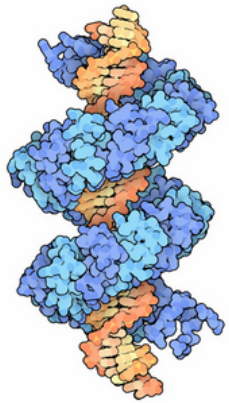
This resource is powered by the Protein Data Bank archive - information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

Use this website to access curated and integrated biological macromolecular information in the context of function, biological processes, evolution, pathways, and disease states.

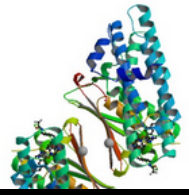
December Molecule of the Month



TAL Effectors

Latest Entries

As of Tuesday Dec 02



New Features

See November 2014 Release Features

Redesigned Home Page

Simple. Clean. Usable. Tools & Functions More Visible.

News

Publications

World AIDS Day Symposium

On December 5, join the RCSB PDB and the François-Xavier Bagnoud Center for a presentation about HIV/AIDS at the molecular level and a panel discussion about clinical practice. » 12/02/14

Art of Science on

A computer simulation is a simulation, run on a single computer, or a network of computers, to reproduce behavior of a system. The simulation uses an abstract model to simulate the system.

The abstract model (or interaction potential) for protein:

Quantum mechanics:

atoms, electrons, …

$$\hat{H}\psi = E\psi$$

The Schrodinger equation was discovered in 1926 by Erwin Schrodinger, an Austrian theoretical physicist. It is an important equation that is fundamental to quantum mechanics.

Molecular mechanics:

atoms, bonds, angles, torsions, …

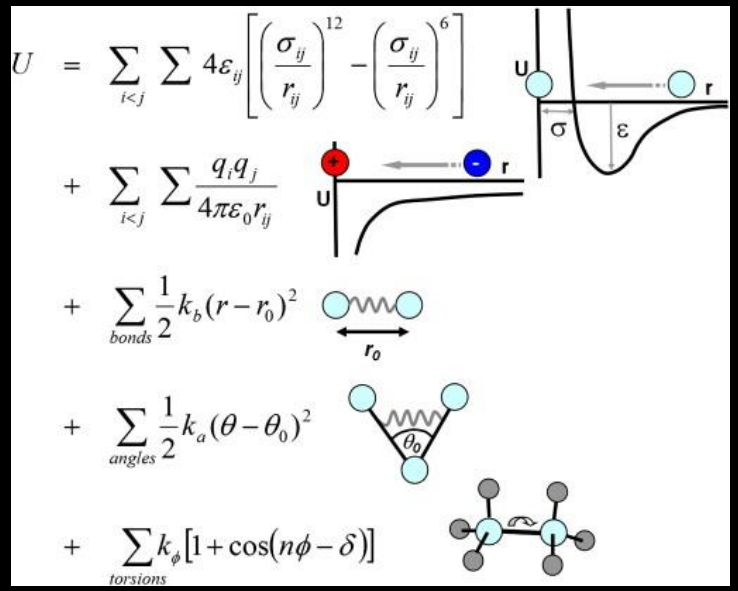
$$U = \sum_{i<j} \sum 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

$$+ \sum_{i<j} \sum \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$

$$+ \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2$$

$$+ \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2$$

$$+ \sum_{torsions} k_\phi [1 + \cos(n\phi - \delta)]$$



The first protein simulations appeared in 1977 with the simulation of the bovine pancreatic trypsin inhibitor (BPTI) .
McCammon, J. A., Gelin, B. R., and Karplus, M. *Nature* 267, 585 (1977)
Simulation time: 8.8 psec (1 psec = 10^{-12} sec)
Now a 10^{-3} sec simulation is possible

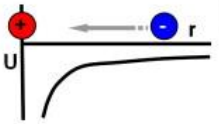
The Nobel Prize in Chemistry 2013

The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel *"for the development of multiscale models for complex chemical systems"*.

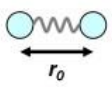
A physical and mathematical view of protein folding is a process of finding the conformation(s) with the lowest (free) energy in a HIGH-dimensional phase space at a certain temperature and pressure.

$$U = \sum_{i<j} \sum 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right]$$

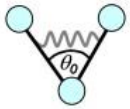
$$+ \sum_{i<j} \sum \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}$$



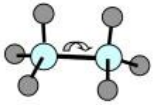
$$+ \sum_{bonds} \frac{1}{2} k_b (r - r_0)^2$$



$$+ \sum_{angles} \frac{1}{2} k_a (\theta - \theta_0)^2$$



$$+ \sum_{torsions} k_\phi [1 + \cos(n\phi - \delta)]$$



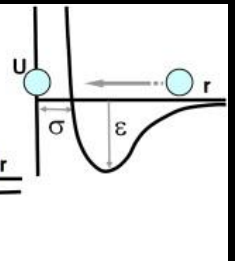
Methods:

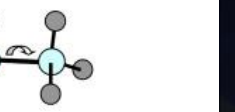
Molecular Dynamics, Monte Carlo

The system has rare chance to travel across the (free) energy barrier, especially when the barrier is higher than kT.

A physical and mathematical view of protein folding is a process of finding the conformation(s) with the lowest (free) energy in a HIGH-dimensional phase space at a certain temperature and pressure.

$$U = \sum_{i<j}\sum 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \sum_{i<j}\sum \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} + \sum_{bonds}\frac{1}{2}k_b(r-r_0)^2 + \sum_{angles}\frac{1}{2}k_a(\theta-\theta_0)^2 + \sum_{torsions}k_\phi[1+\cos(n\phi-\delta)]$$





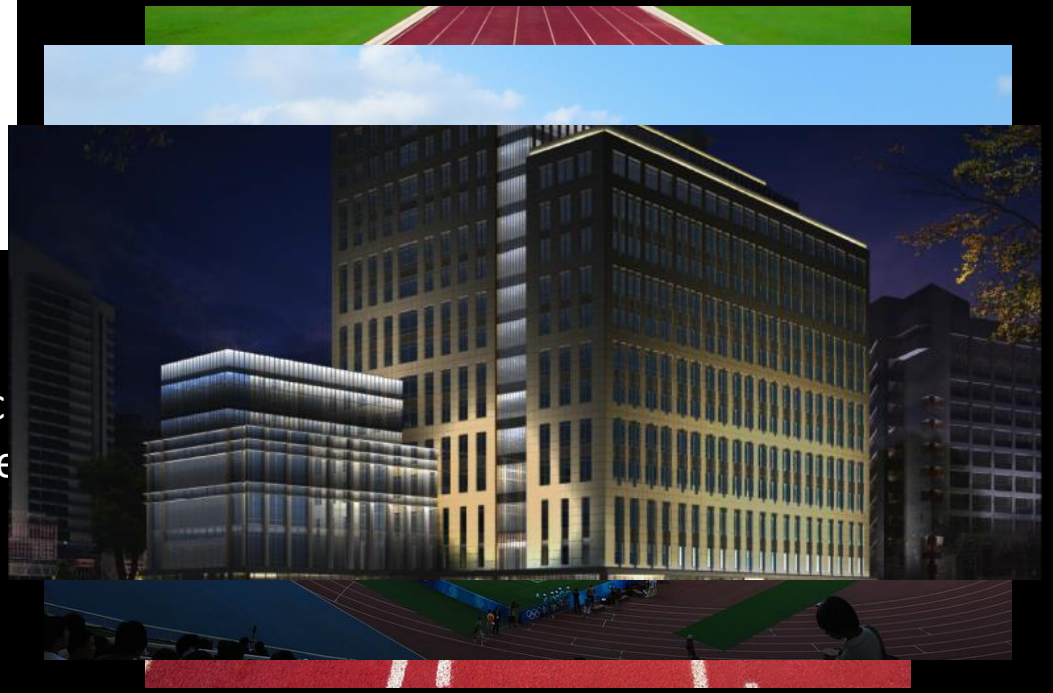
Methods:

Molecular Dynamics, Monte Carlo

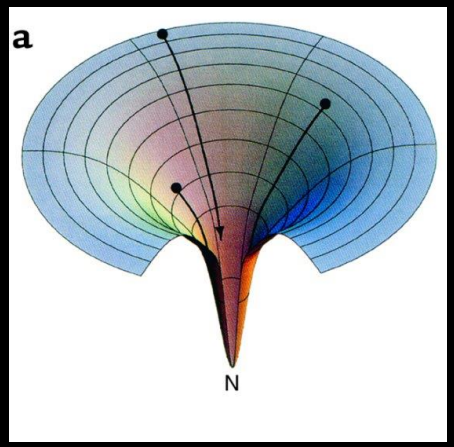
Dimension:

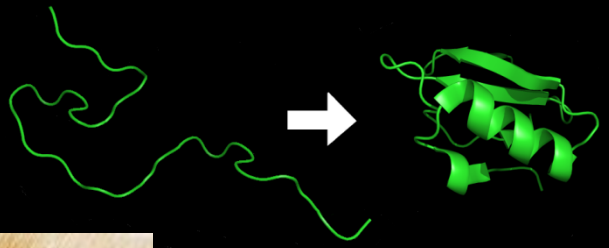
3N-6, N is the number of atoms

6 external DoF (translation and rotation)



The system has rare chance to travel ac
the (free) energy barrier, especially whe
the barrier is higher than kT.

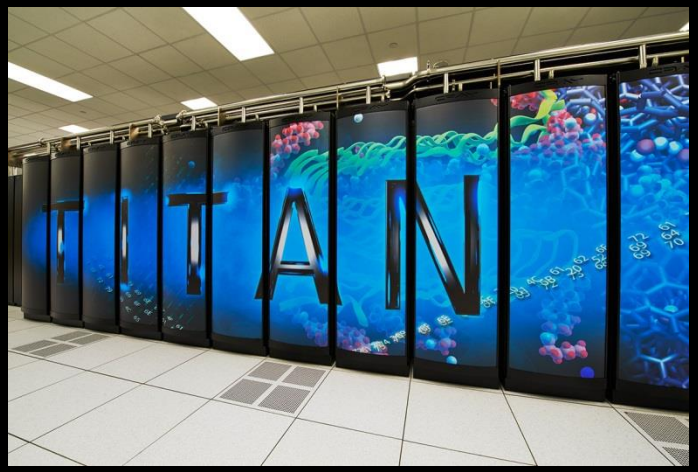

a
N




ORIGINAL MOTION PICTURE SOUNDTRACK
THE
SHAWSHANK
REDEMPTION
MUSIC BY THOMAS NEWMAN


Trainspotting

The workhorses



Thank you