Computational Structural Biology: Protein Simulation and Structure Prediction

Michael Levitt
Structural Biology
Stanford

http:/csb.stanford.edu/class

INTRODUCTION

- Course Outline.
- Some notes on Class.
- Outline of this lecture.

COURSE OUTLINE

- 10/Jan/01 Basic Physical Principles.
- 17/Jan/01 Principles of Protein Structure (A).
- 24/Jan/01 Principles of Protein Structure (B).
- 31/Jan/01 Physical Simulation of Proteins.
- 7/Feb/01 Protein Structure Prediction.
- 14/Feb/01 Homology Modeling.
- 21/Feb/01 Fold recognition.
- 28/Feb/01 *Ab initio* folding.
- 7/Mar/01 Protein Bio-informatics.
- 9/Mar/01 Genome annotation.

(10 two-hour lectures)

SOME NOTES ON CLASS

- Fill in Registration Page.
- Homework and Final Project.
- Webpage http:/csb.stanford.edu/class
- TA: Michael Sykes sykes@stanford.edu

LECTURE 1. OUTLINE

- Introduction.
- Principles of Biology.
- Protein Folding.
- Protein Architecture.
- Basic Physical Principles.
- Advanced Physical Principles.

PRINCIPLES OF BIOLOGY

Structure Matters.

Evolution.

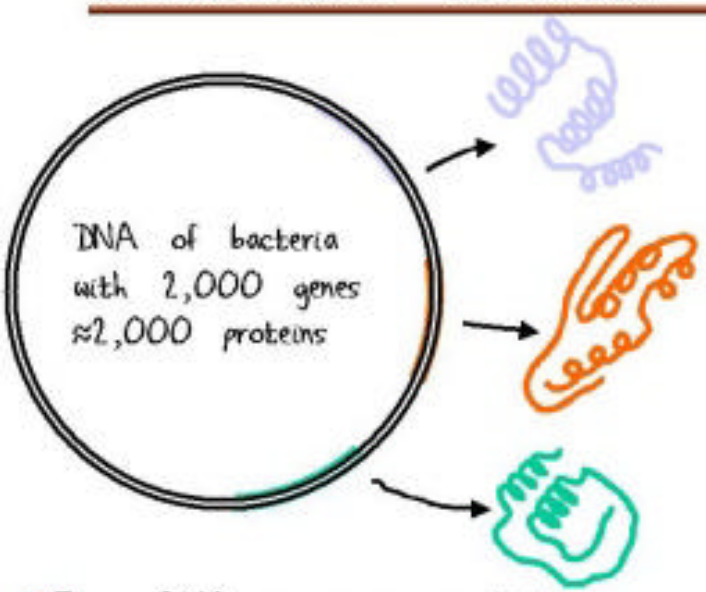
Hierarchy of Stable Subsystems.

STRUCTURAL OVERVIEW OF BIOLOGY

STRUCTURAL OVERVIEW OF BIOLOGY

img009.jpg (400x300x24b jpeg)

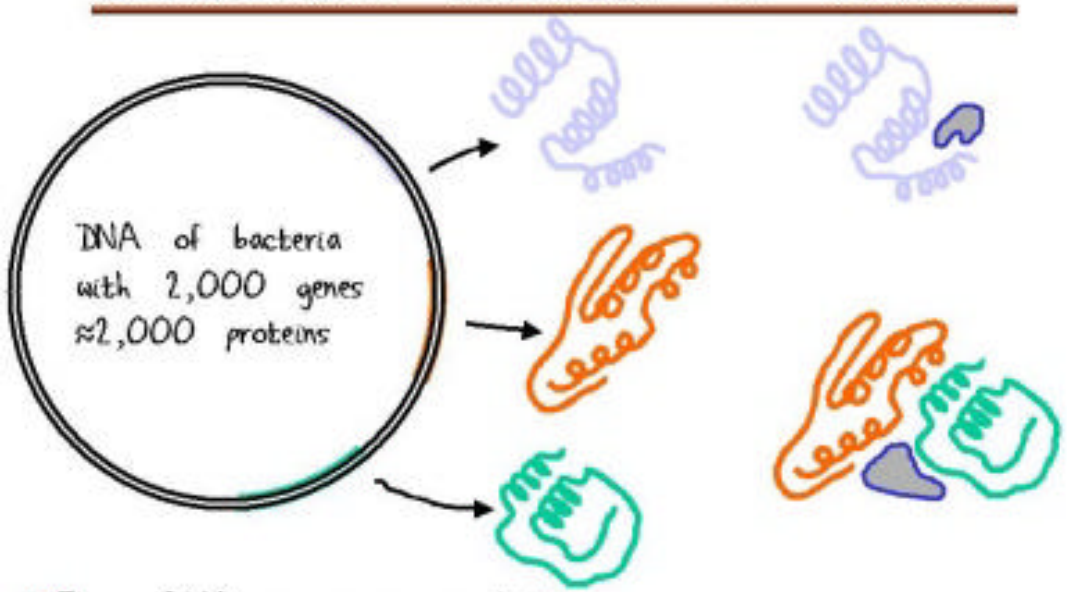
STRUCTURAL OVERVIEW OF BIOLOGY



- From DNA sequence, predict all protein structures

img010.jpg (400x300x24b jpeg)

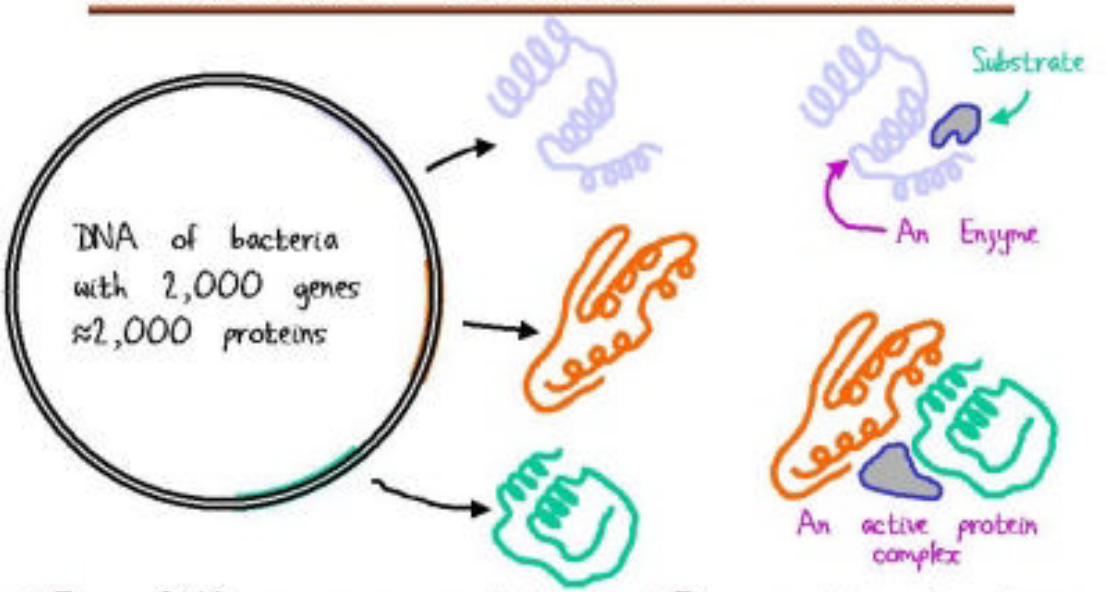
STRUCTURAL OVERVIEW OF BIOLOGY



- From DNA sequence, predict all protein structures

img011.jpg (400x300x24b jpeg)

STRUCTURAL OVERVIEW OF BIOLOGY



- From DNA sequence, predict all protein structures
- From protein structures predict all function.

INFORMATICS OVERVIEW OF BIOLOGY

EVOLUTION: All life-forms are related through common ancestors that were survivors

INFORMATICS OVERVIEW OF BIOLOGY

EVOLUTION: All life-forms are related through common ancestors that were survivors

Life is like a huge software project:

CON

PRO

INFORMATICS OVERVIEW OF BIOLOGY

EVOLUTION: All life-forms are related through common ancestors that were survivors

Life is like a huge software project:

CON
- All is written in Assembly code (DNA)
- Nothing is designed: all is random hacking.

PRO

INFORMATICS OVERVIEW OF BIOLOGY

EVOLUTION: All life-forms are related through common ancestors that were survivors

Life is like a huge software project:

CON
- All is written in Assembly code (DNA)
- Nothing is designed: all is random hacking.

PRO
- Fantastic quality assurance (QA). Darwin
- Uses component that themselves worked best.
- Lots of time: 4,000,000,000 years.

PROTEIN FOLDING

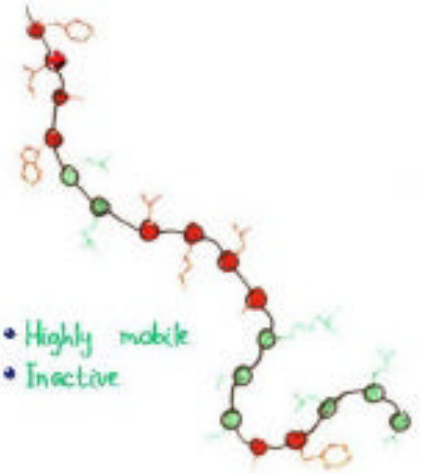
- Spontaneous self-organization.
- Quick, taking a fraction of a second.
- What drives folding?

img018.jpg (400x300x24b jpeg)

PROTEIN FOLDING IS CENTRAL

Sequence -> Structure -> Function

- Unfolded protein is a chain of amino acids

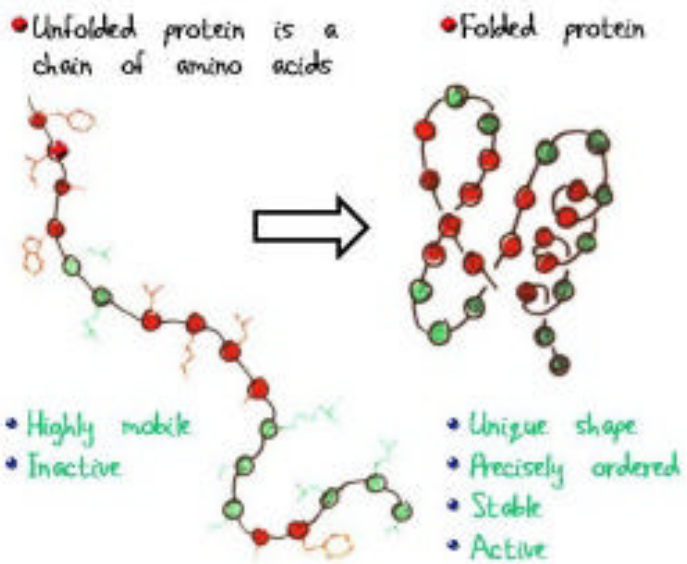


- Highly mobile
- Inactive

img019.jpg (400x300x24b jpeg)

PROTEIN FOLDING IS CENTRAL

Sequence -> Structure -> Function

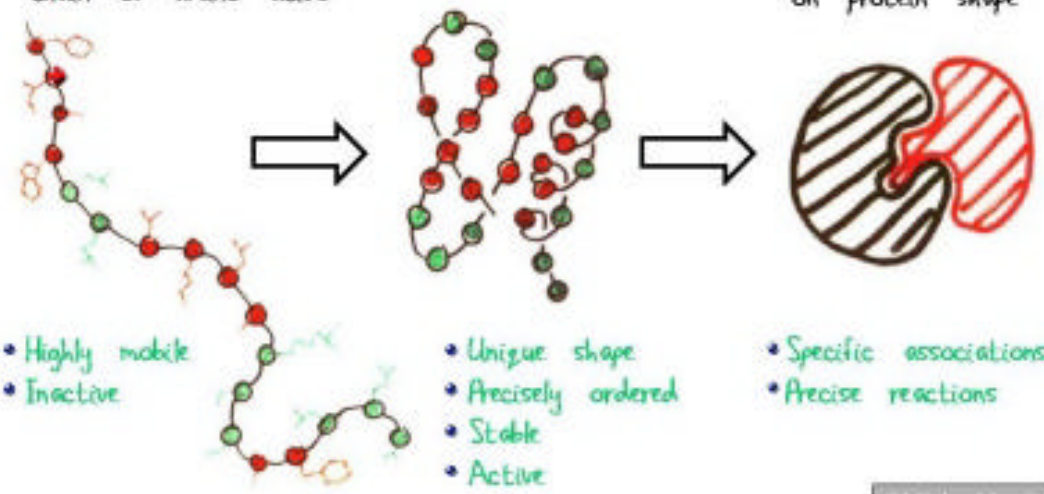


img020.jpg (400x300x24b jpeg)

PROTEIN FOLDING IS CENTRAL

Sequence -> Structure -> Function

- Unfolded protein is a chain of amino acids
- Folded protein
- Function depends on protein shape

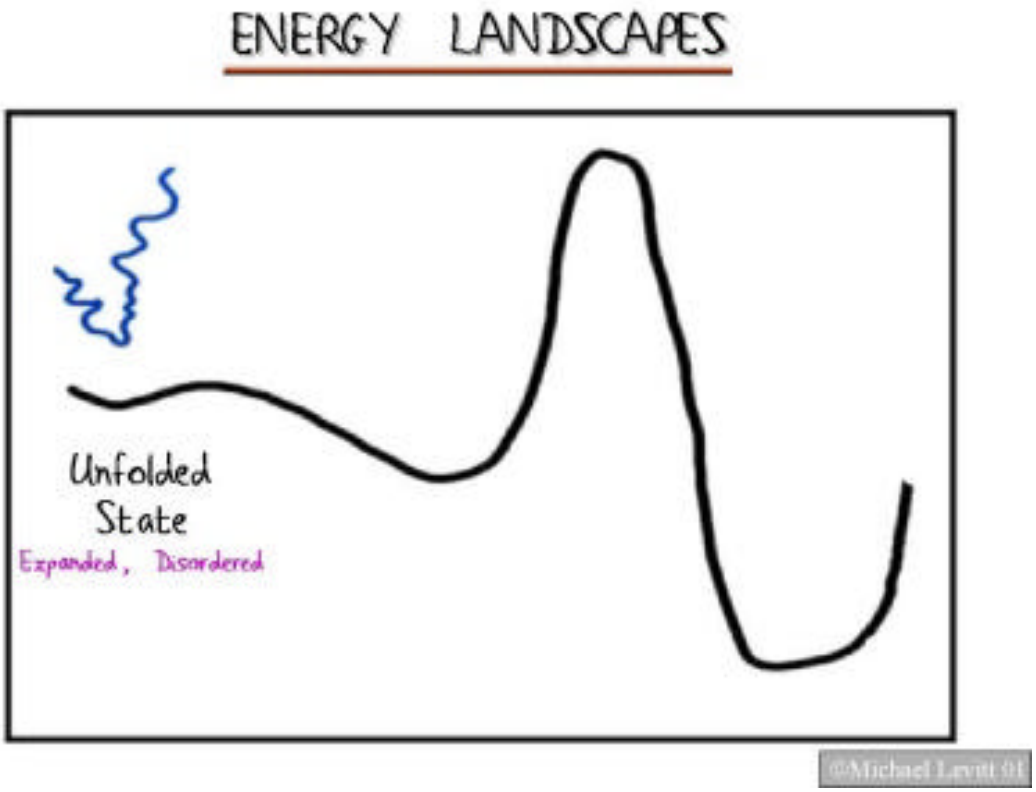


- Highly mobile
- Inactive

- Unique shape
- Precisely ordered
- Stable
- Active

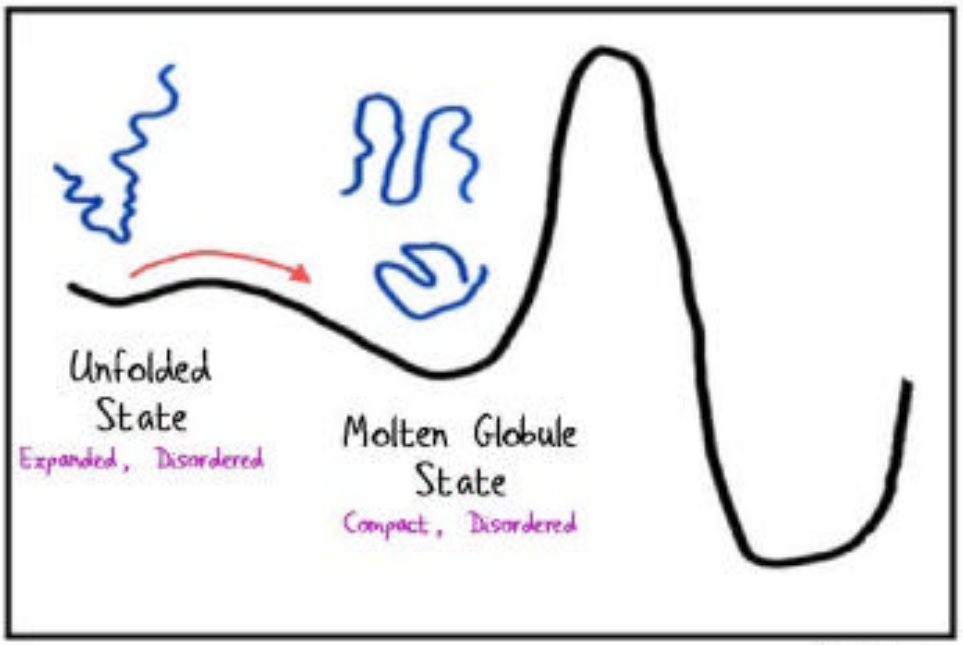
- Specific associations
- Precise reactions

img021.jpg (400x300x24b jpeg)

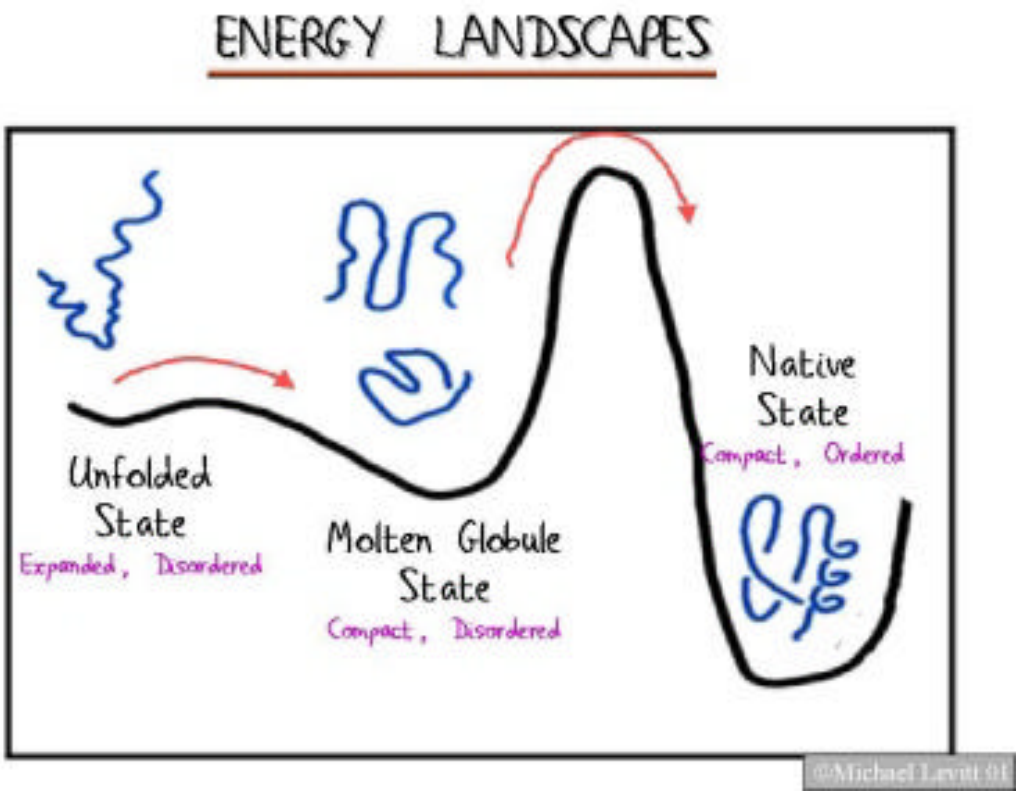


img022.jpg (400x300x24b jpeg)

ENERGY LANDSCAPES

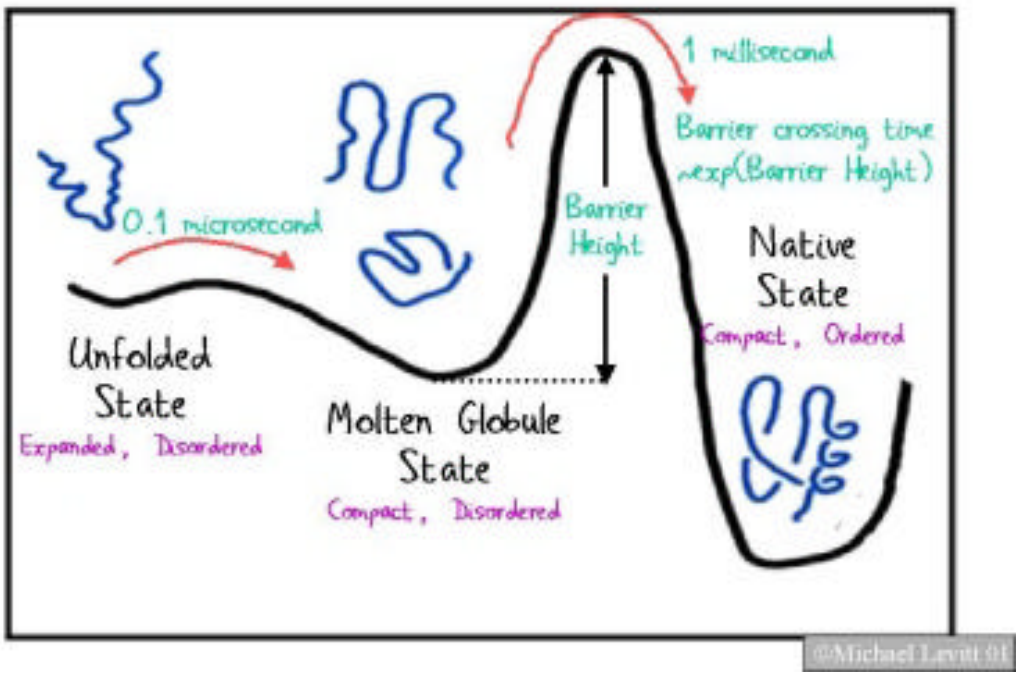


img023.jpg (400x300x24b jpeg)

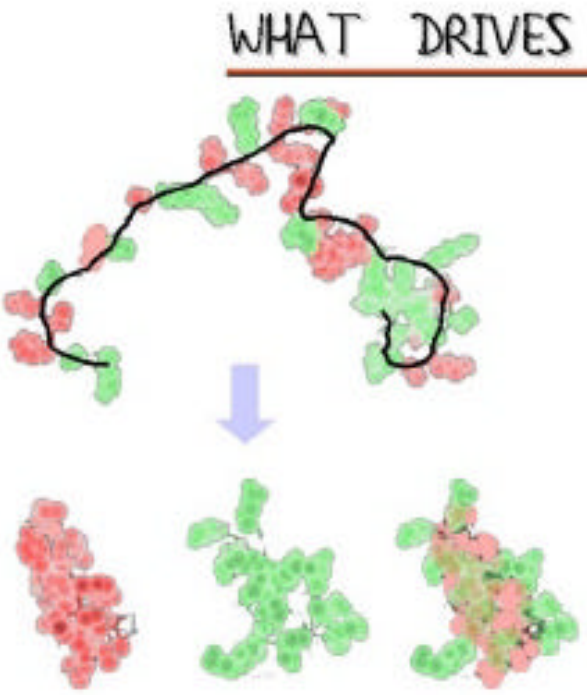


img024.jpg (400x300x24b jpeg)

ENERGY LANDSCAPES

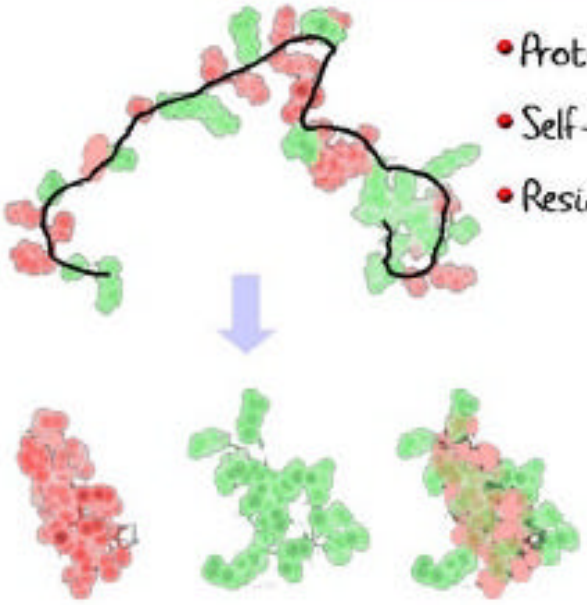


img025.jpg (400x300x24b jpeg)



img026.jpg (400x300x24b jpeg)

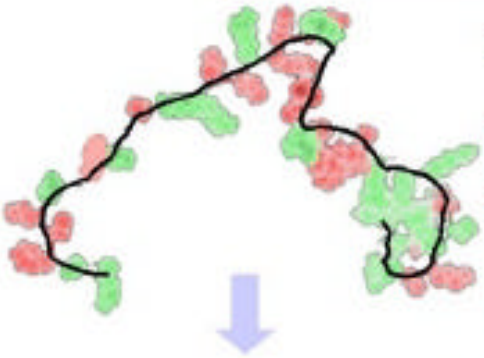
WHAT DRIVES FOLDING?



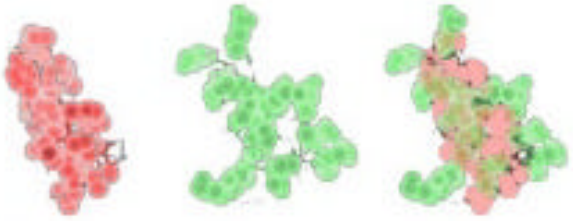
- Protein is a chain.
- Self-avoiding and close packed.
- Residue preferences:
  - Inside/Outside
  - Specific Neighbors

img027.jpg (400x300x24b jpeg)

WHAT DRIVES FOLDING?



- Protein is a chain.
- Self-avoiding and close packed.
- Residue preferences:
  - Inside/Outside
  - Specific Neighbors



Hydrophobic Hydrophilic All Residues

Pink are hydrophobic, like to be away from water

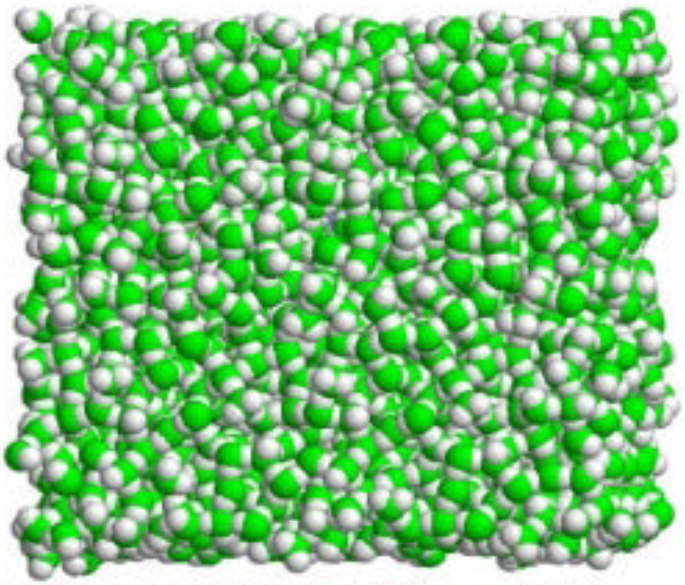
Green are hydrophilic, like contact with water

WHAT DOES A PROTEIN LOOK LIKE?

- Hidden in water?
- A close-packed globular object?
- A set of bond connections?
- A chain of secondary structures?

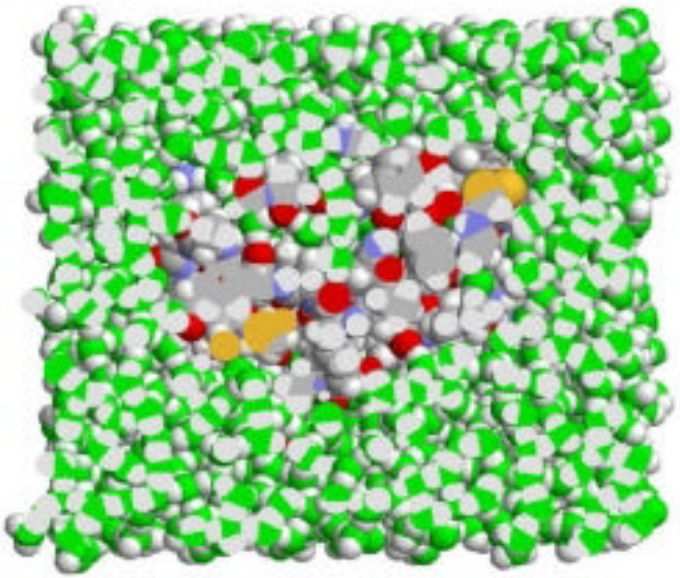
img029.jpg (400x300x24b jpeg)

A BOX OF WATER



Proteins are stable in water. Water molecules and proteins are both just atoms.

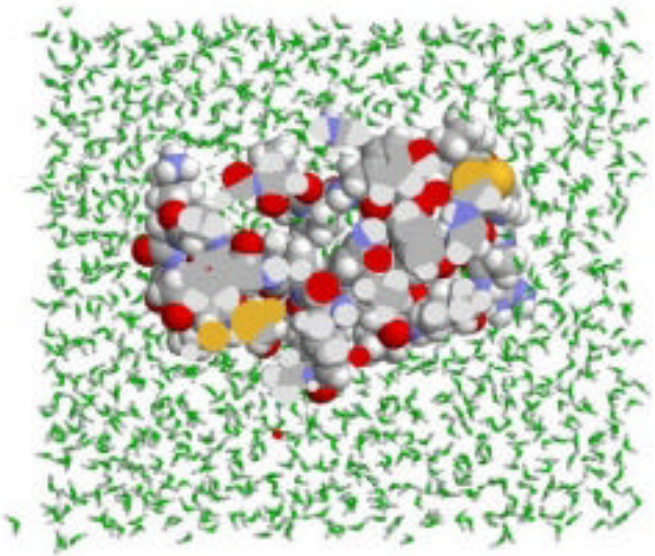
SLICE THE BOX OF WATER



Now you see the protein is closely interacting with the water.

img031.jpg (400x300x24b jpeg)

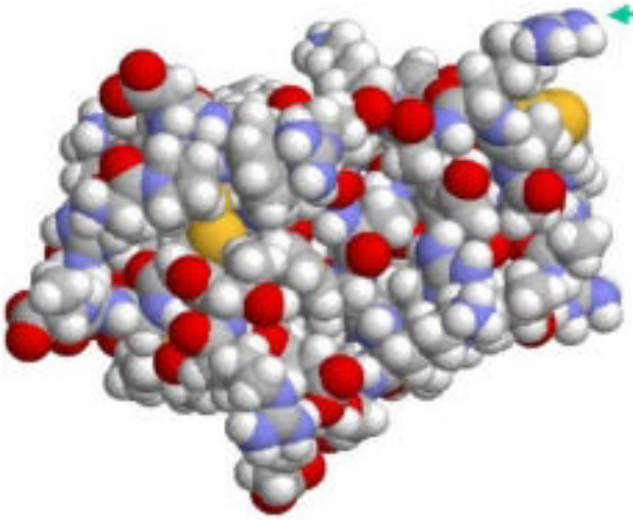
SLICE THE BOX OF WATER



Now you see the protein is closely interacting with the water.

img032.jpg (400x300x24b jpeg)

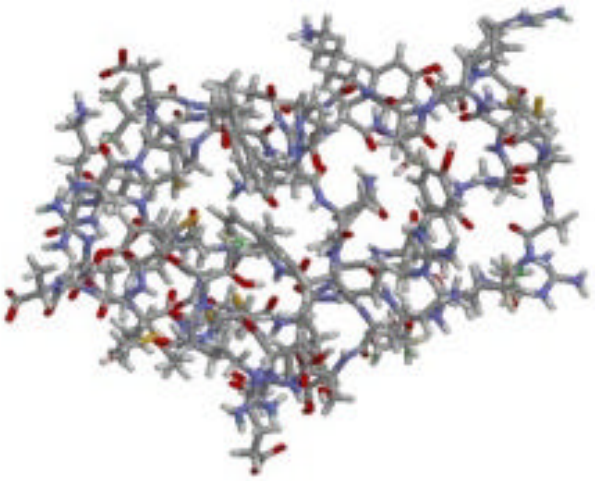
SEE THE PROTEIN ALONE



A protein is a solid object. It acts as a plug to inhibit trypsin using Lys 15.

img033.jpg (400x300x24b jpeg)

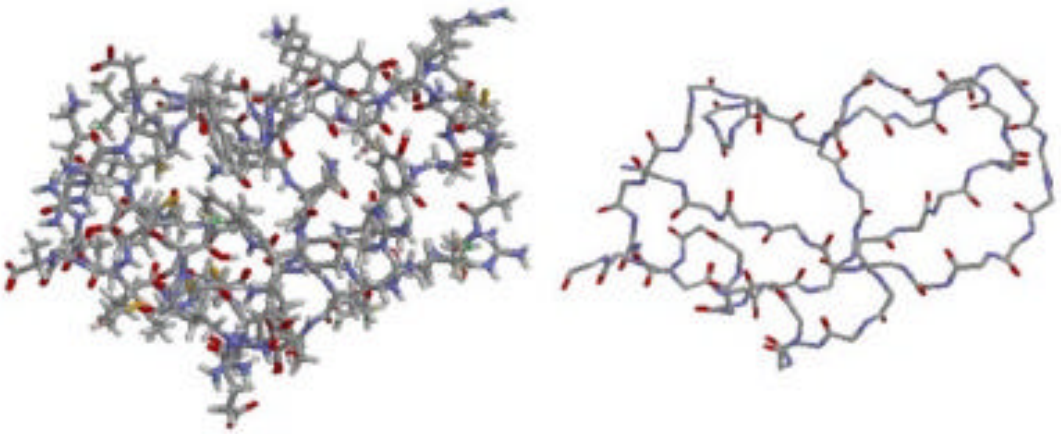
SEE THE BONDED ATOMS



It is hard to see what is important in the structure.

img034.jpg (400x300x24b jpeg)

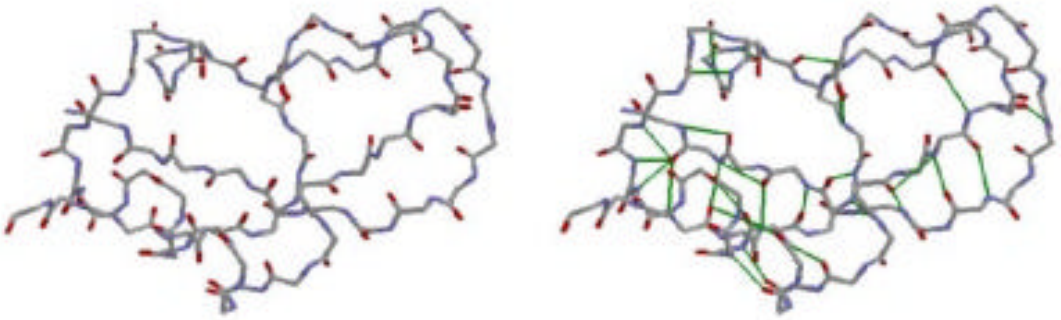
TRY TO SEE THE MAIN CHAIN.



It is still not clear. Where are the ends?

img035.jpg (400x300x24b jpeg)

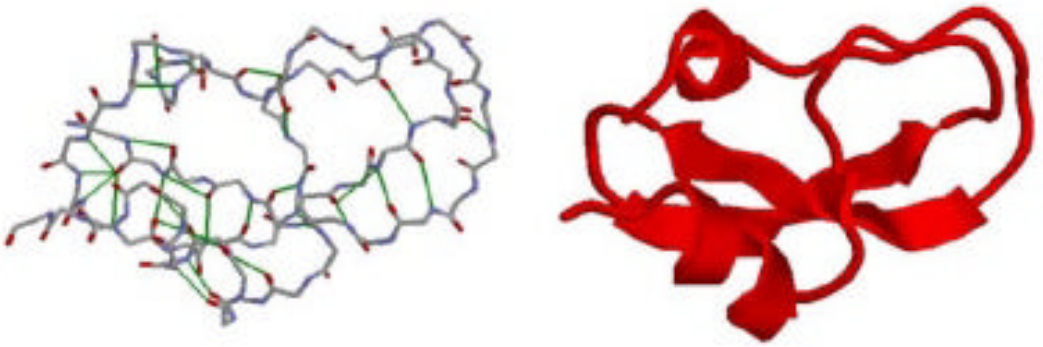
TRY TO SEE THE MAIN CHAIN.



Add the hydrogen bonds as green lines.

img036.jpg (400x300x24b jpeg)

TRY TO SEE THE MAIN CHAIN.



Emphasize the secondary structure (alpha-helices and beta-sheets).

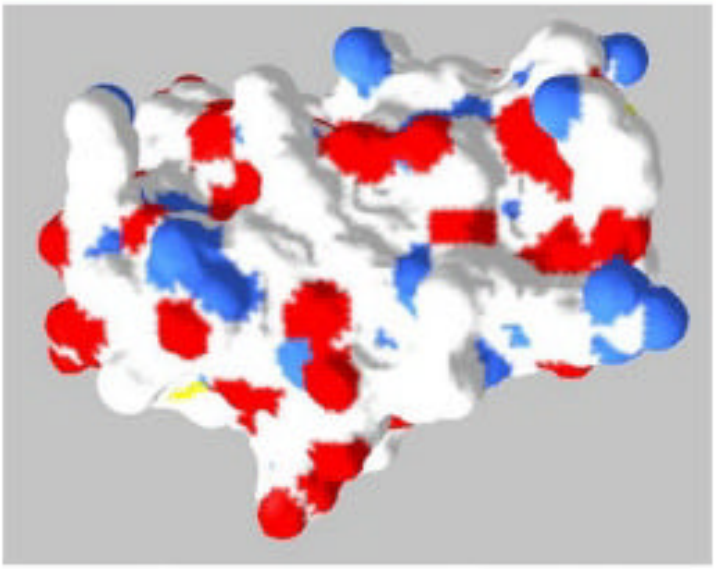
img037.jpg (400x300x24b jpeg)

TRY TO SEE THE CHAIN PATH.

Use color to see the chain direction.
Start (N) is blue. End (C) is red.

img038.jpg (400x300x24b jpeg)

FOR FUNCTION SHOW A CHRISTO MODEL.



Wrap the protein to see the molecular surface. Color by atom type.

BASIC PHYSICAL PRINCIPLES

- Forces between atoms.
- Total energy function.
- Hydrogen bonds.
- Moving over the energy surface.

img040.jpg (400x300x24b jpeg)

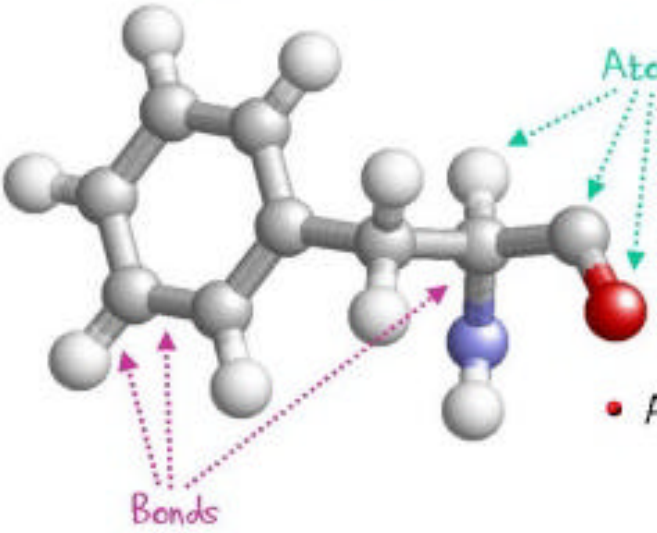
WHAT IS AN ATOM?

Remember we are ignoring quantum mechanics

- A position in space
  (x, y, z) coordinates.
- A solid object. It
  Occupies space
- Carries electric charge.
  Positive or negative

img041.jpg (400x300x24b jpeg)

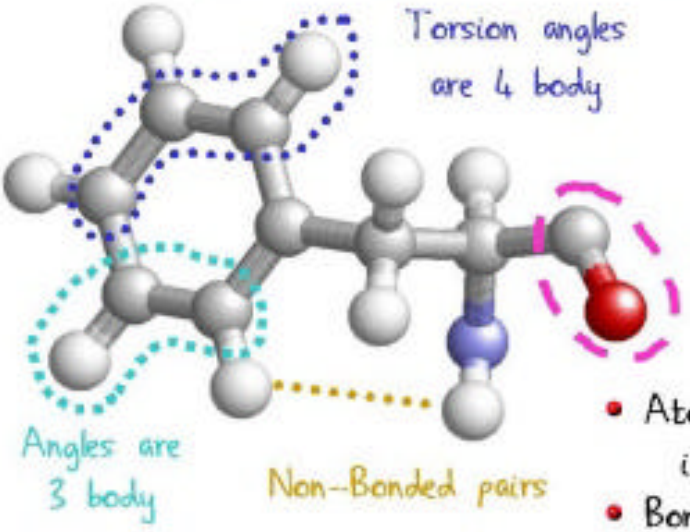
WHAT IS A MOLECULE?



- Atoms are connected by bonds.

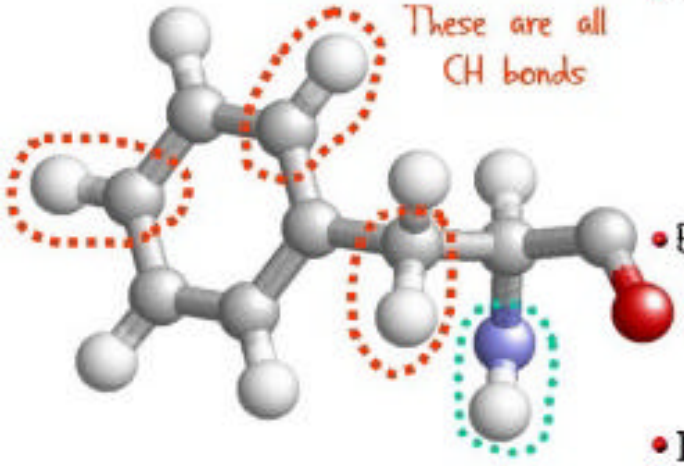
img042.jpg (400x300x24b jpeg)

WHAT IS A MOLECULE?



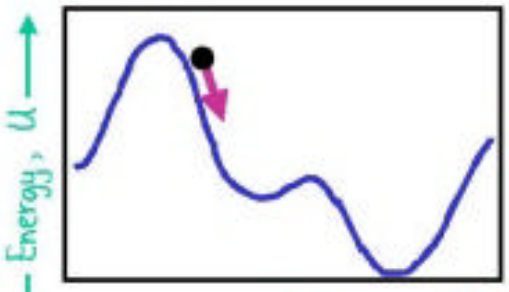
- Atoms are connected in a simple graph.
- Bonded Interactions: 2, 3, 4 body.
- Non-Bonded Interactions.

FORCES BETWEEN ATOMS

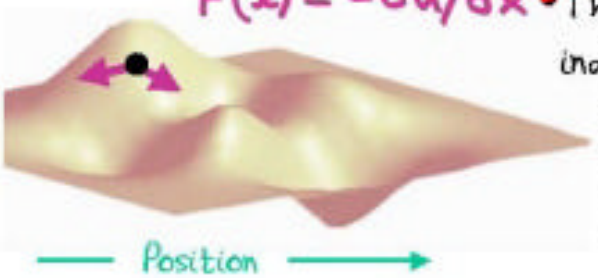


- Total energy is sum of individual contributions: Terms are additive.
- Energy terms transferable: A bond is a bond in any environment.
- Ignore quantum effects: Atoms are like balls, forces like springs.

TOTAL POTENTIAL ENERGY



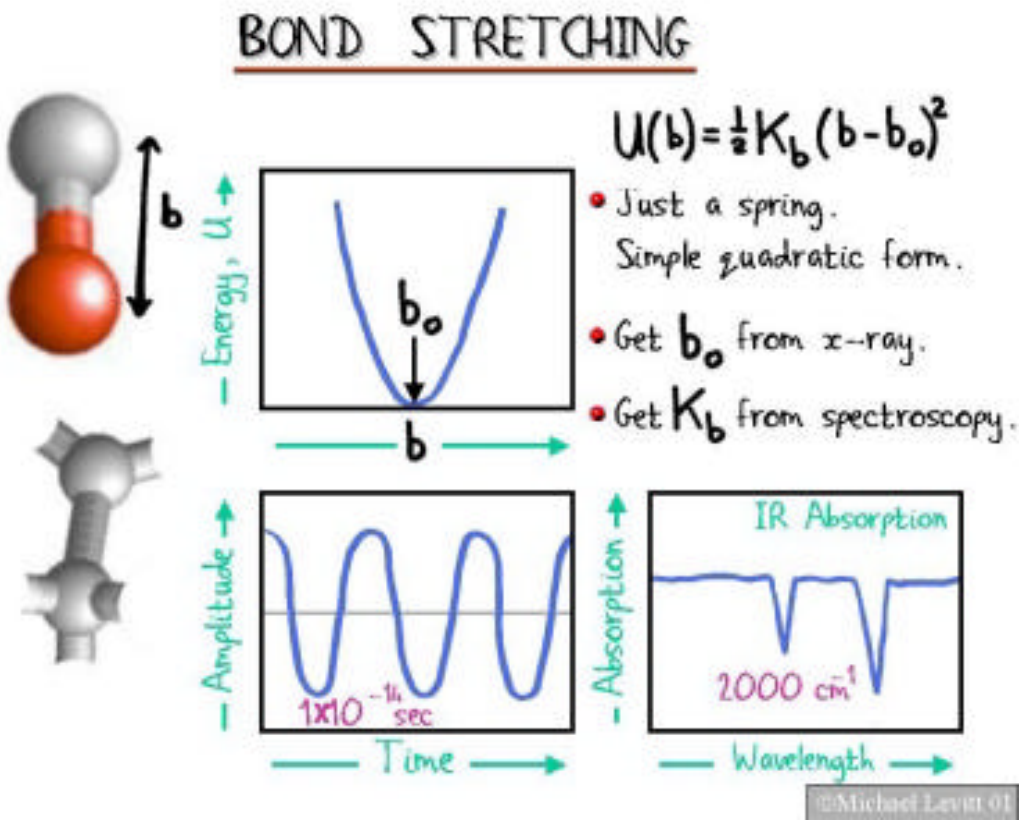
$F(x) = -dU/dx$



Position

- The total potential energy or enthalpy fully defines the system, U.
- The forces are the gradients of the energy.
- The energy is a sum of independent terms for: Bond, Bond angles, Torsion angles and non-bonded atom pairs.

img045.jpg (400x300x24b jpeg)

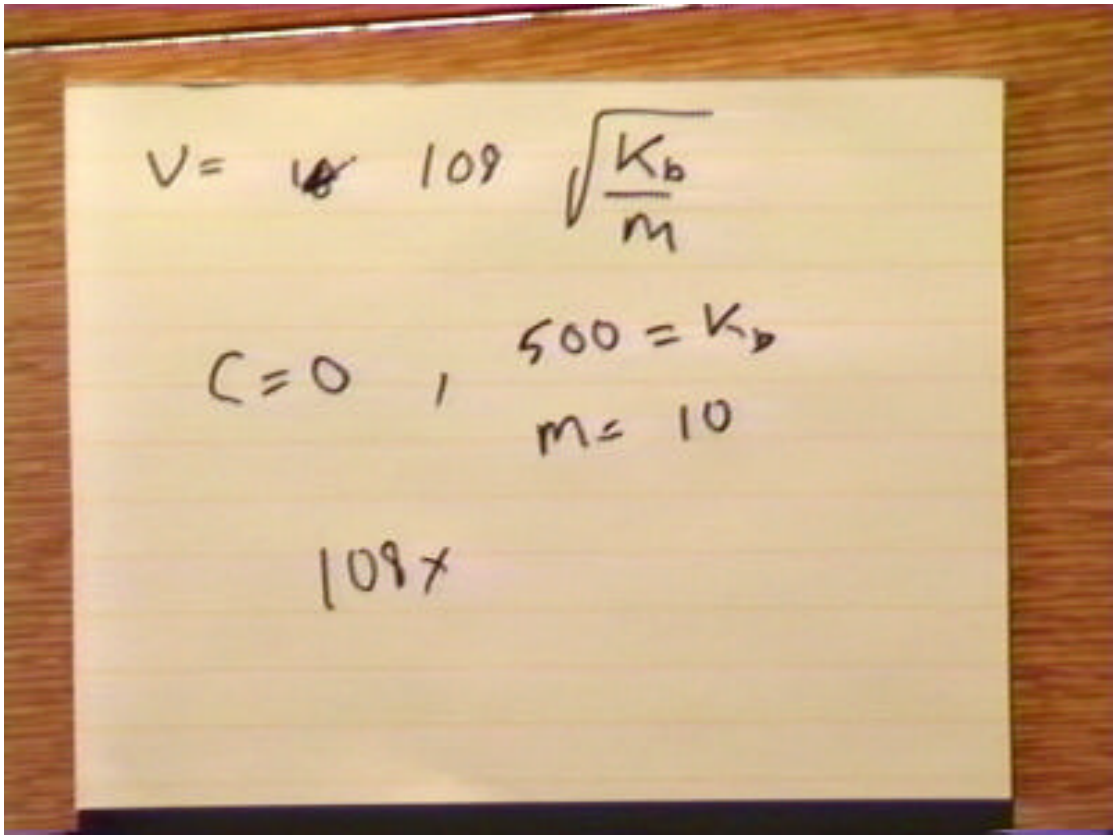


UNITS IN FORCE FIELDS

Quantity	Unit
Energy	kcal/mol
Distance	Å
Mass	A.M.U. (H is 1 AMU)
Time	0.5×10^{-13} secs.
Frequency	cm^{-1}
Charge	e (electron is 1 e)

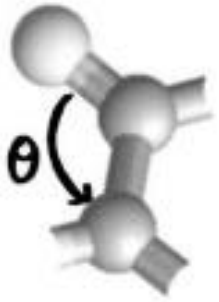
These is definitely a strange combination!

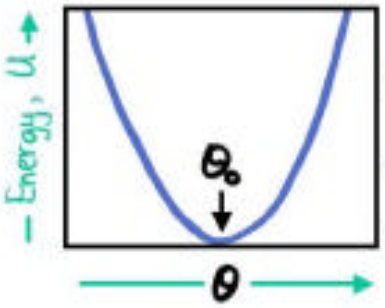
img047.jpg (400x300x24b jpeg)



img048.jpg (400x300x24b jpeg)

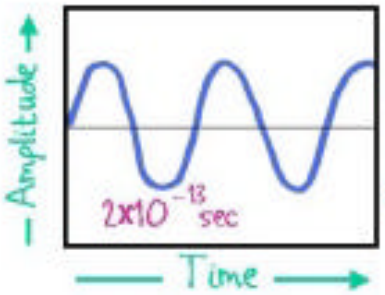
BOND ANGLE BENDING

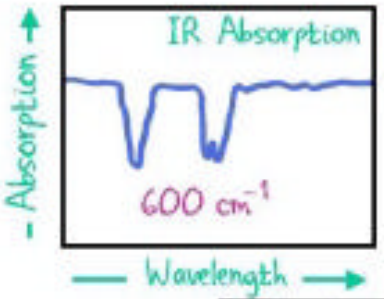




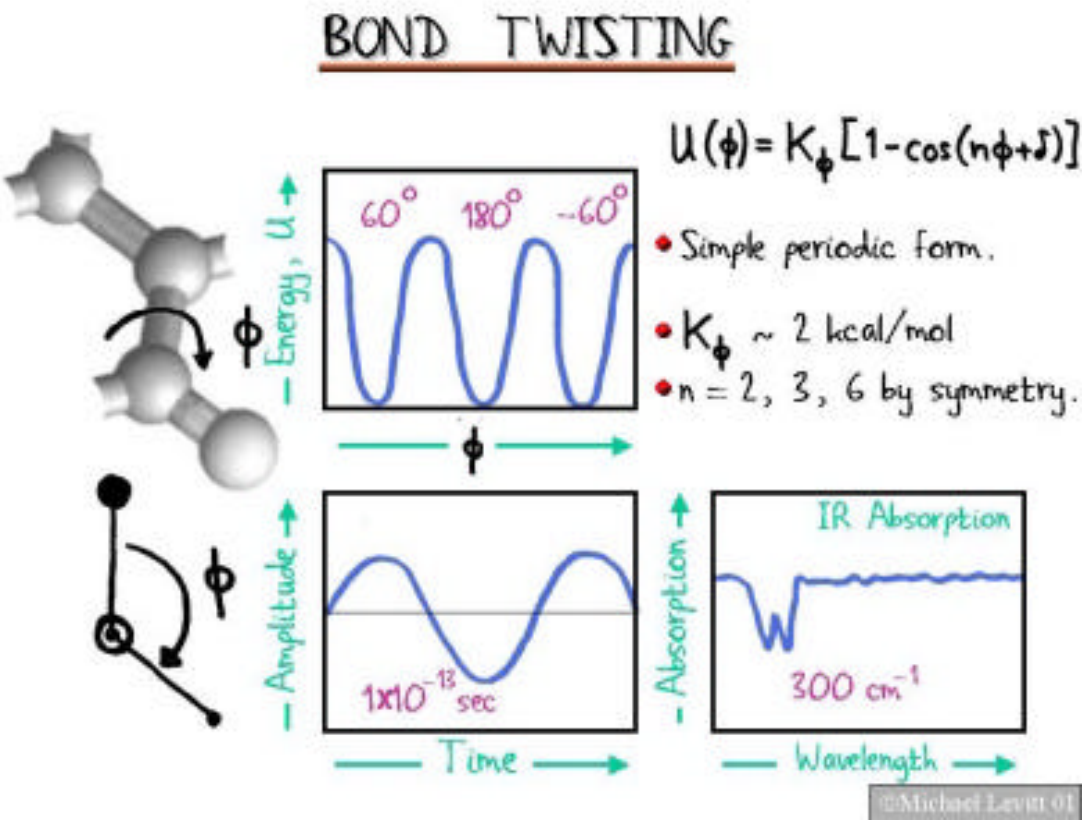
$$U(\theta) = \tfrac{1}{2} K_\theta (\theta - \theta_0)^2$$

- Simple quadratic form.
- Get θ_0 from x-ray.
- Get K_θ from spectroscopy.



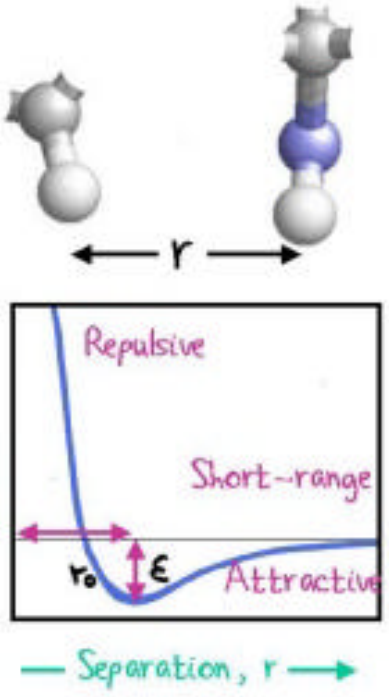


img049.jpg (400x300x24b jpeg)



img050.jpg (400x300x24b jpeg)

VAN DER WAALS INTERACTION



$$U(r) = \epsilon\left[\left(r_0/r\right)^{12} - 2\left(r_0/r\right)^{6}\right]$$

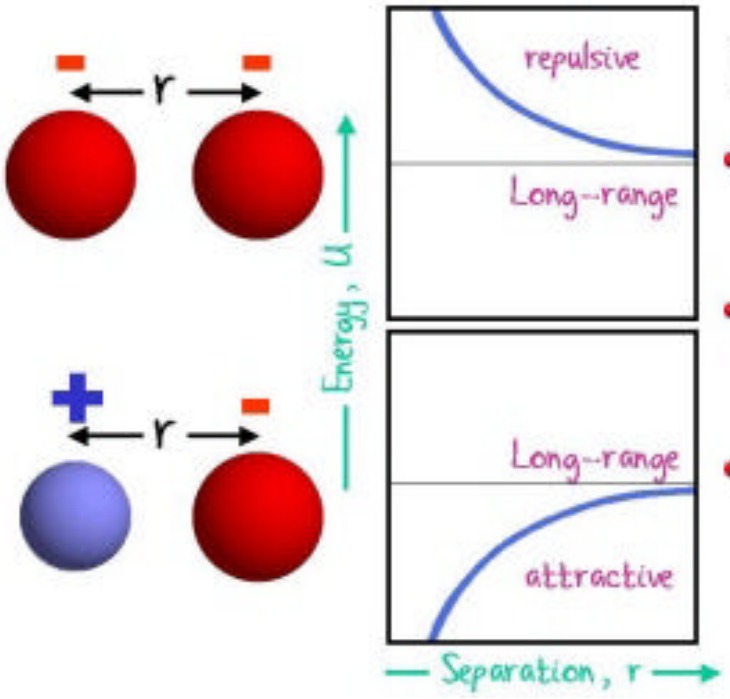
- Weak interaction at 0.2 kcal/mol.

$\epsilon \sim 0.2$ kcal/mol, $r_0 = 3.5$Å

- Same form for all atoms.
- Can be large for a large molecule.

img051.jpg (400x300x24b jpeg)

ELECTROSTATICS INTERACTION

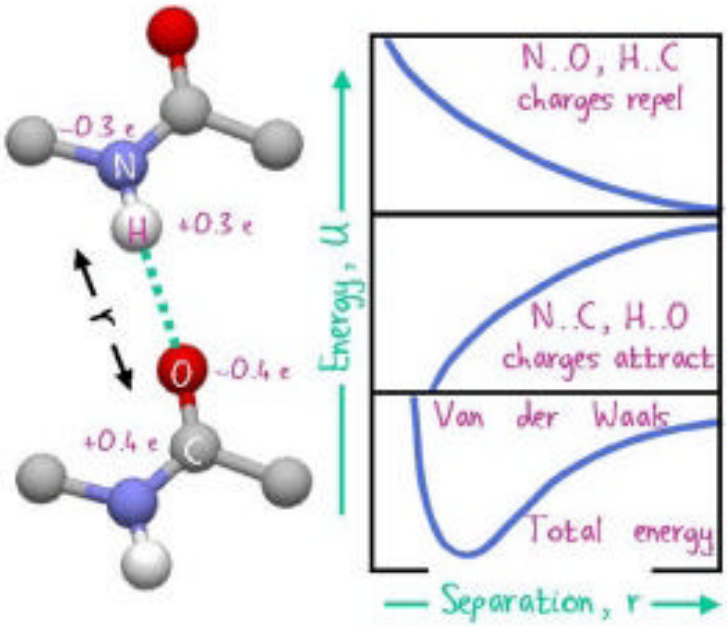


$U(r) = 332 q_i q_j / r$

- Both repulsive and attractive.
- All atoms carry partial charges.
- Charge depends on: nature of atom, state of ionization, environment.

img052.jpg (400x300x24b jpeg)

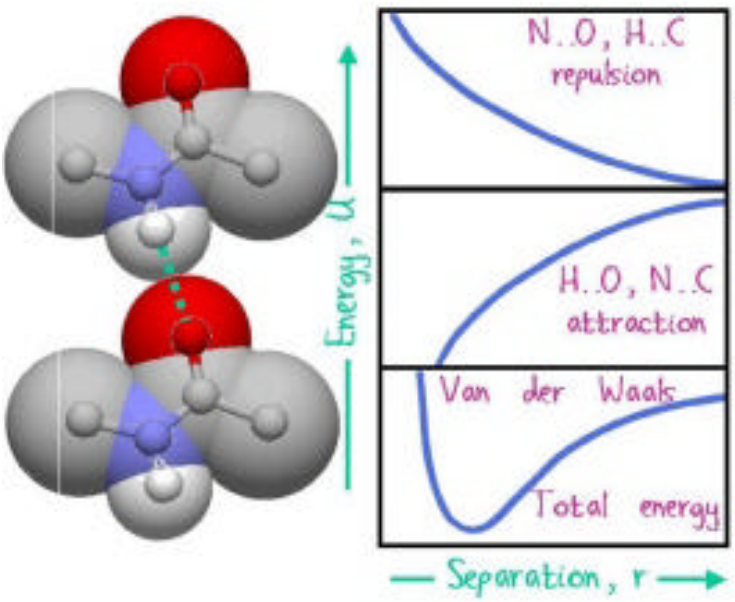
HYDROGEN BONDS



- No special forces.
- Geometry allows good interaction combining van der Waals and electrostatic terms.
- Strong attraction with a minimum 5 kcal/mol deep at r=2.8Å.

img053.jpg (400x300x24b jpeg)

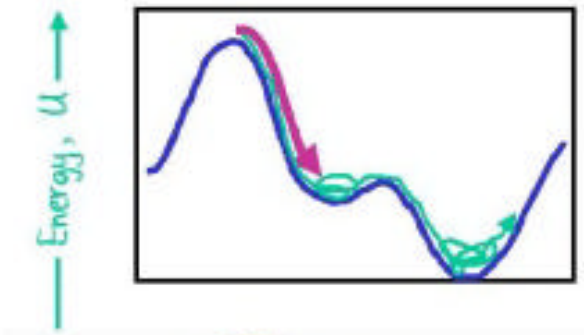
HYDROGEN BONDS

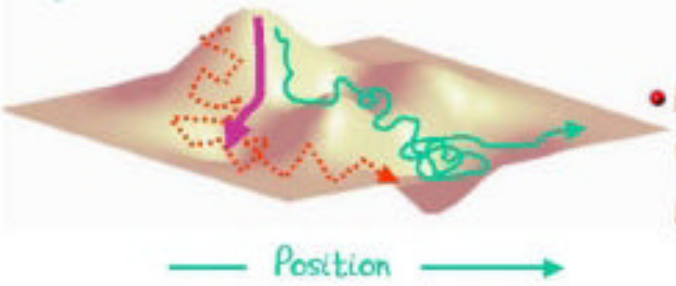


- Very directional.
- This comes from the repulsion between the other atoms in the system.
- Do not need a special directional term even for water molecules.

img054.jpg (400x300x24b jpeg)

MOVING OVER THE ENERGY SURFACE



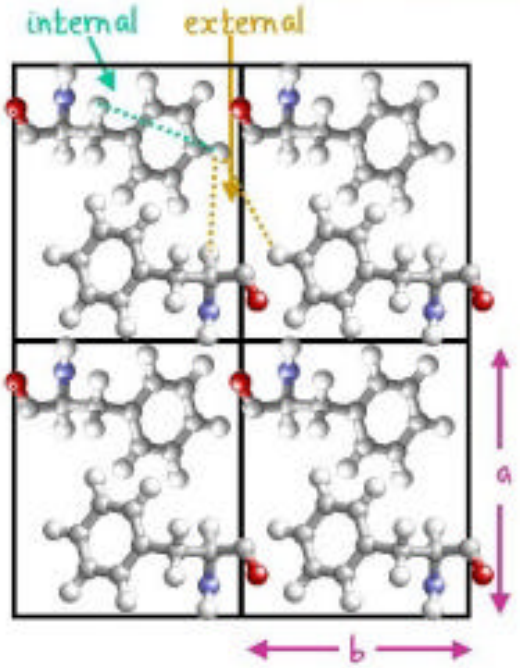


- Energy Minimization drops into local minimum.
- Molecular Dynamics uses thermal energy to move smoothly over surface.
- Monte Carlo Moves are random. Accept with probability $\exp(-\Delta U/kT)$.

ADVANCED PHYSICAL PRINCIPLES

- Simulating crystals.
- Simulating liquids.
- Hydrophobic effect
- Dielectric effect.

SIMULATING CRYSTALS



- $U = U_{internal} + U_{external}$

  $U_{external}$ must include all molecules in lattice.

- Minimize U to calculate:

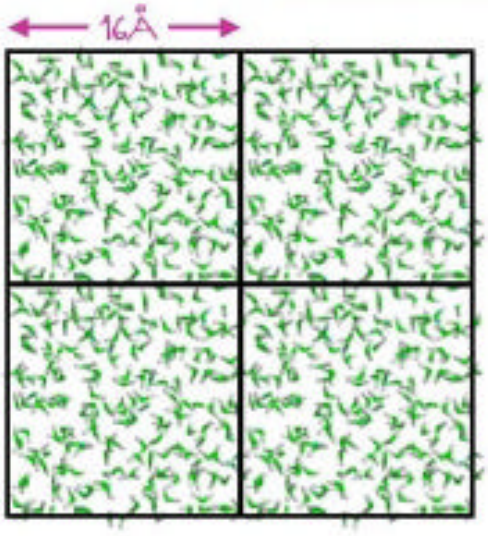
  Unit cell parameters a, b

  Sublimation energy.

- Calibrate to fit experiment.

img057.jpg (400x300x24b jpeg)

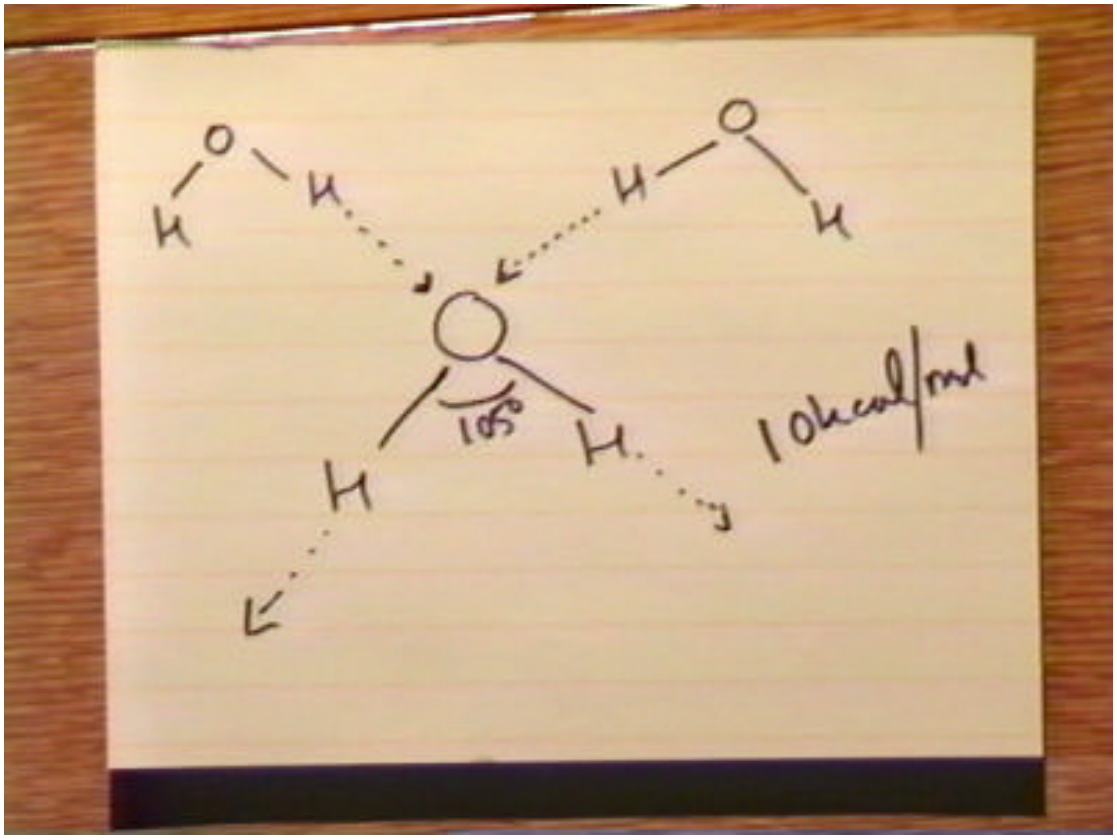
SIMULATING LIQUIDS



Rahman & Stillinger 1970

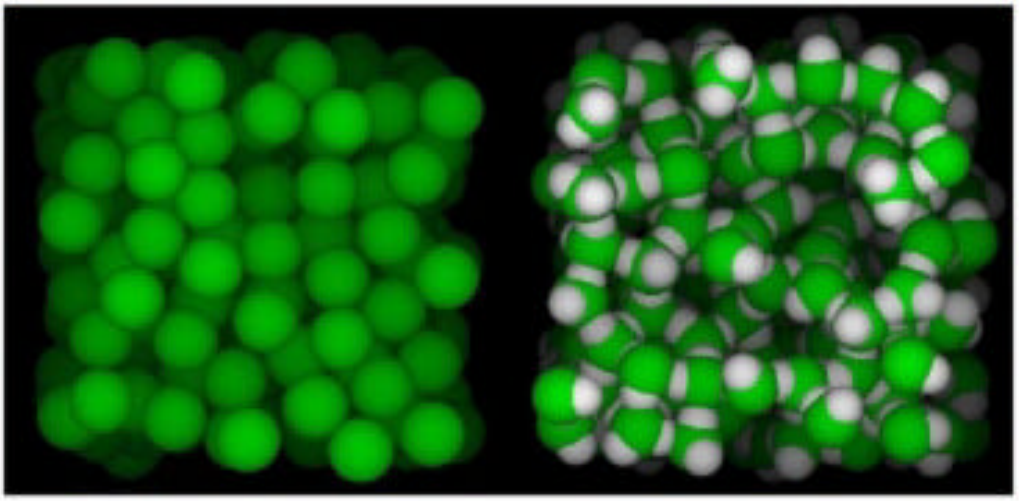
- Periodic box with 216 water molecules. Simulate motion for 100 ps.
- Calculate key experimental properties:
  - Heat of vaporization
  - Structure.
  - Internal pressure
  - Diffusion constant
- Compare with experiment and calibrate.

img058.jpg (400x300x24b jpeg)



img059.jpg (400x300x24b jpeg)

LIQUID EXAMPLES: ARGON & WATER

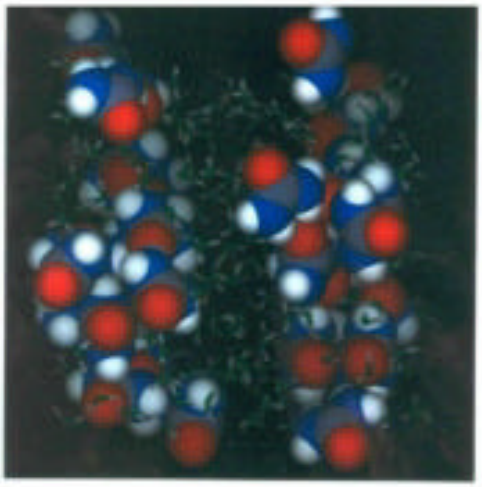


Argon is like a collection of hard spheres.

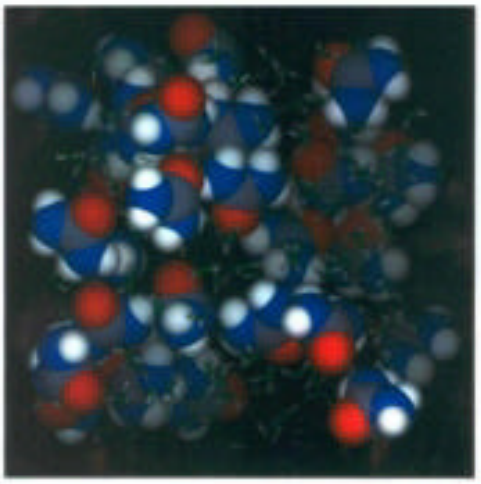
Water has open structure, tetrahedral geometry.

img060.jpg (400x300x24b jpeg)

LIQUID EXAMPLES: UREA



Initially the urea molecules are clustered.

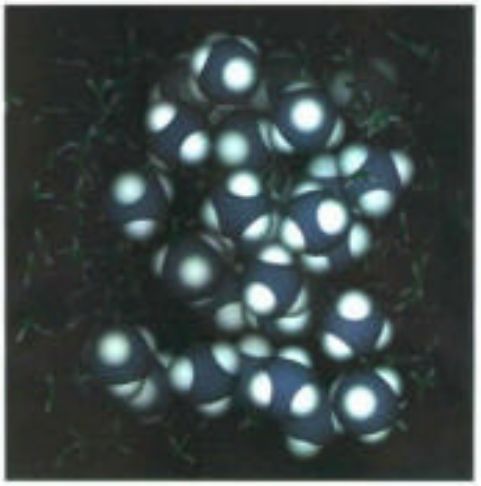


At equilibrium they are well-mixed: urea is soluble.

img061.jpg (400x300x24b jpeg)

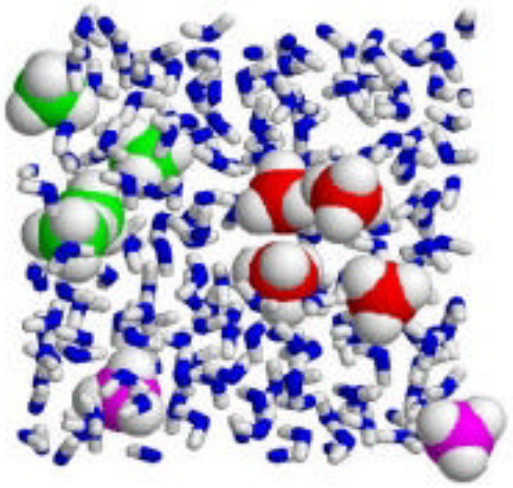
LIQUID EXAMPLES: METHANE

Initially the methane molecules are well-mixed.



At equilibrium they cluster: methane is insoluble.

HYDROPHOBIC EFFECT

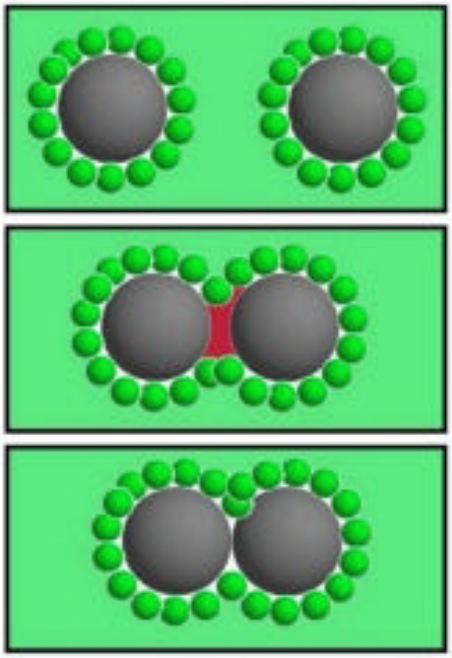


Box with periodic boundaries.

- Hydrophobic molecules cluster in water (not soluble).
- The energy is proportional to the surface area buried in cluster.
- This is not a pairwise additive force.

img063.jpg (400x300x24b jpeg)

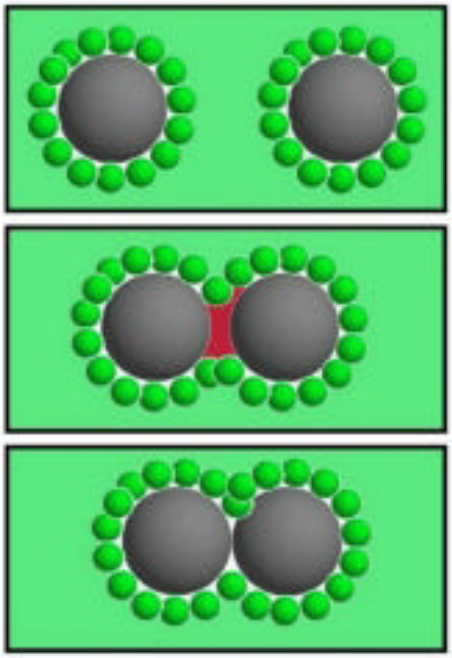
EXPLAINING THE HYDROPHOBIC EFFECT



- Water molecules next to solute molecules are less free & unhappy as a result of entropy.

img064.jpg (400x300x24b jpeg)

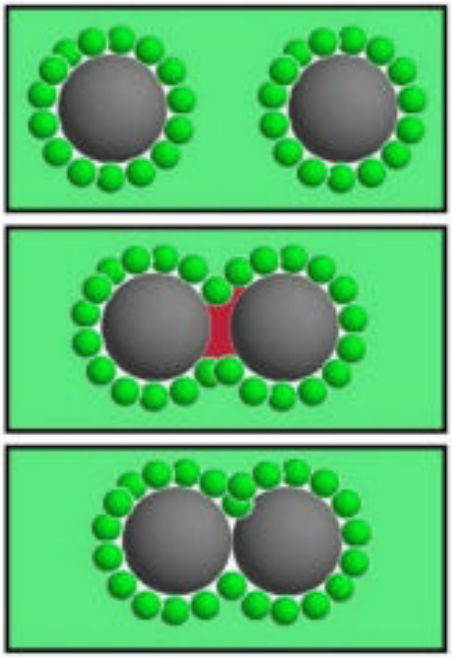
EXPLAINING THE HYDROPHOBIC EFFECT



- Water molecules next to solute molecules are less free & unhappy as a result of entropy.
- When solute separation is slightly less than water radius, there is a void. Voids are bad in liquids.

img065.jpg (400x300x24b jpeg)

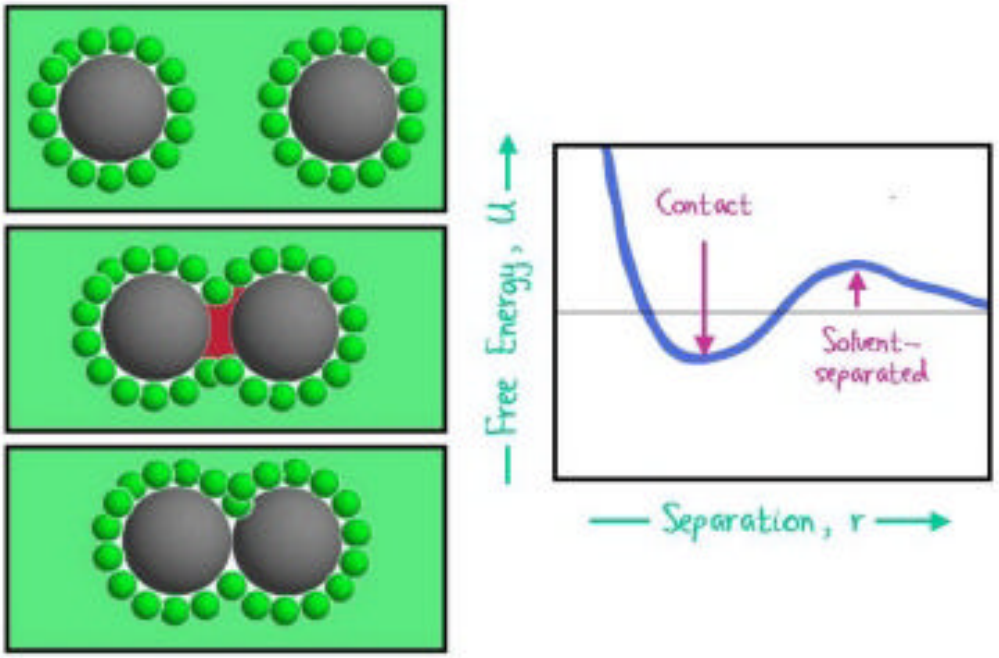
EXPLAINING THE HYDROPHOBIC EFFECT



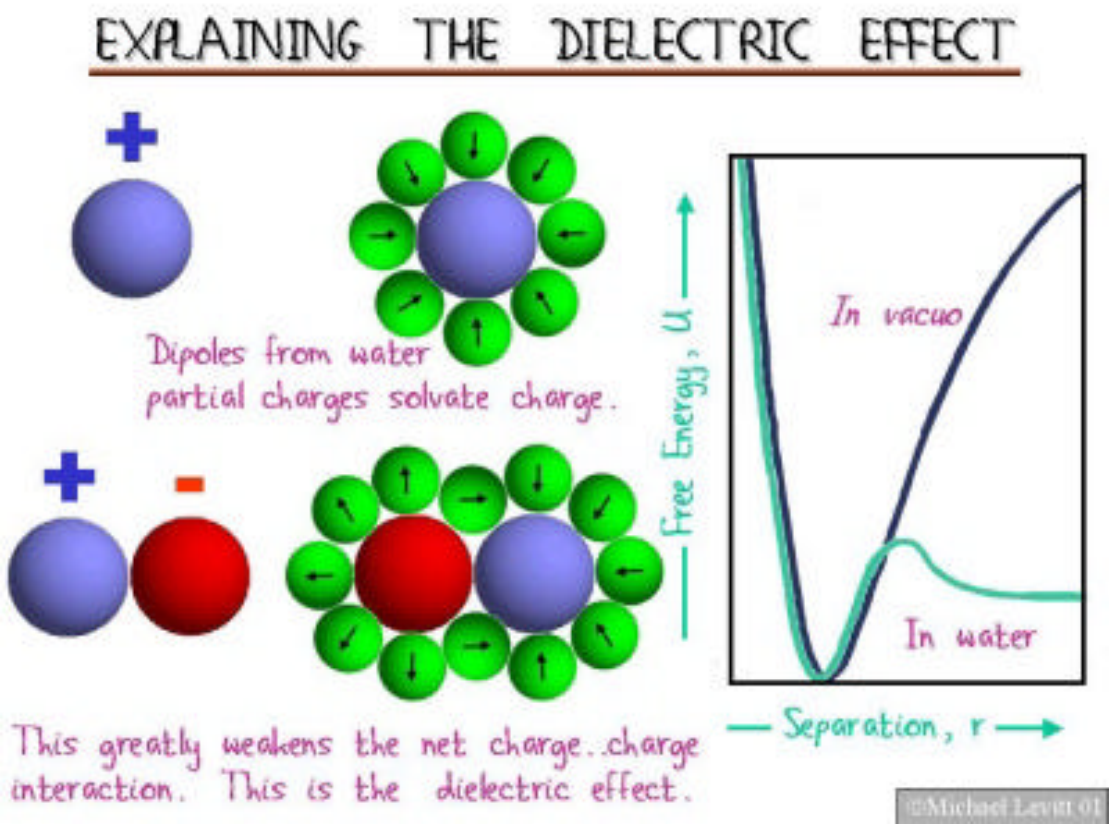
- Water molecules next to solute molecules are less free & unhappy as a result of entropy.
- When solute separation is slightly less than water radius, there is a void. Voids are bad in liquids.
- As the solutes get closer the void disappears. There are fewer unhappy water molecules.

img066.jpg (400x300x24b jpeg)

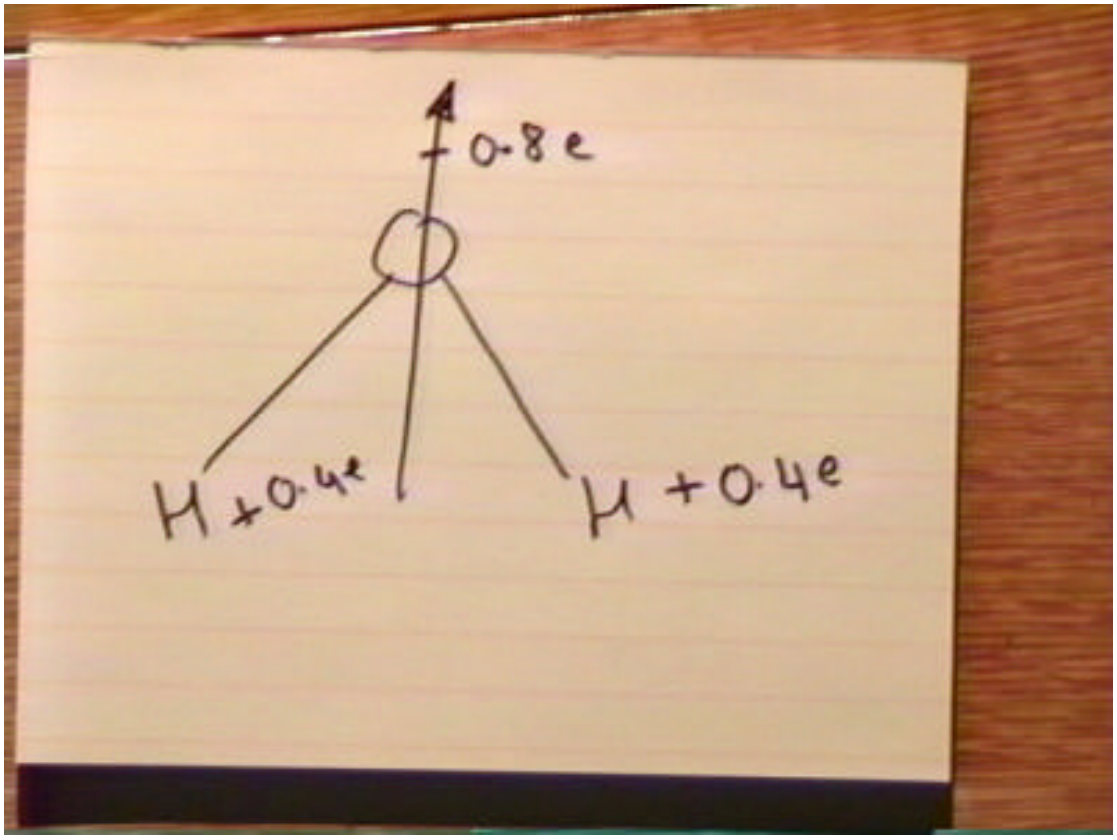
EXPLAINING THE HYDROPHOBIC EFFECT



img067.jpg (400x300x24b jpeg)



img068.jpg (400x300x24b jpeg)



STRENGTH OF INTERACTIONS

Interaction	Energy (kcal/mol)
Van der Waals in water	-0.1
Van der Waals *in vacuo*	-0.3
Hydrogen bond in water	-1.0
Hydrogen bond *in vacuo*	-5.0
Change bond angle by 10 degrees	+2.0
Change bond length by 0.1 Å	+3.0
Torsion barrier about single bond	+3.0
Torsion barrier about double bond	+20.0

HOMEWORK ASSIGNMENT

- Get molecular graphics programs:
  SwissPDB and Rasmol.
- Learn how to download and view a protein.
- Learn how to use the program by
  next Wednesday!
- See class website for help and more details.

THE END

of Lecture 1