

# Computer Aided Drug Design

## —Molecular Simulations I

Qin Xu

<http://cbb.sjtu.edu.cn/~qinxu/CADD.htm>

# Molecular Simulations

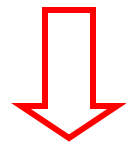
- Molecular Dynamics
  - $dv/dt = F/m$
- Langevin Dynamics
  - $dv/dt = F/m - \gamma v + R/m$
- Monte-Carlo
  - $r * dv/dt = R/m$
  - Markov chain (MCMC)

# Work flow for dynamics simulations

Molecular Dynamics

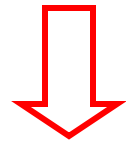
Langevin Dynamics

Set up Initial state



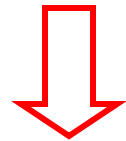
Coordinates,  
velocities

Calculate potentials / forces



QM, MM, CG

Solve equations of motions



$dv/dt = F/m$

Evolve the system



Stochastic movements  
of solvent

$dv/dt = F/m - \gamma v + R/m$

# Statistical Mechanics(统计力学)



- Computer simulation at the microscopic level
  - atomic and molecular positions
  - velocities
  - Effective pair potential energies
- Experimental results on macroscopic properties
  - volume
  - pressure
  - Temperature, internal energy

# Phase space (相空间)

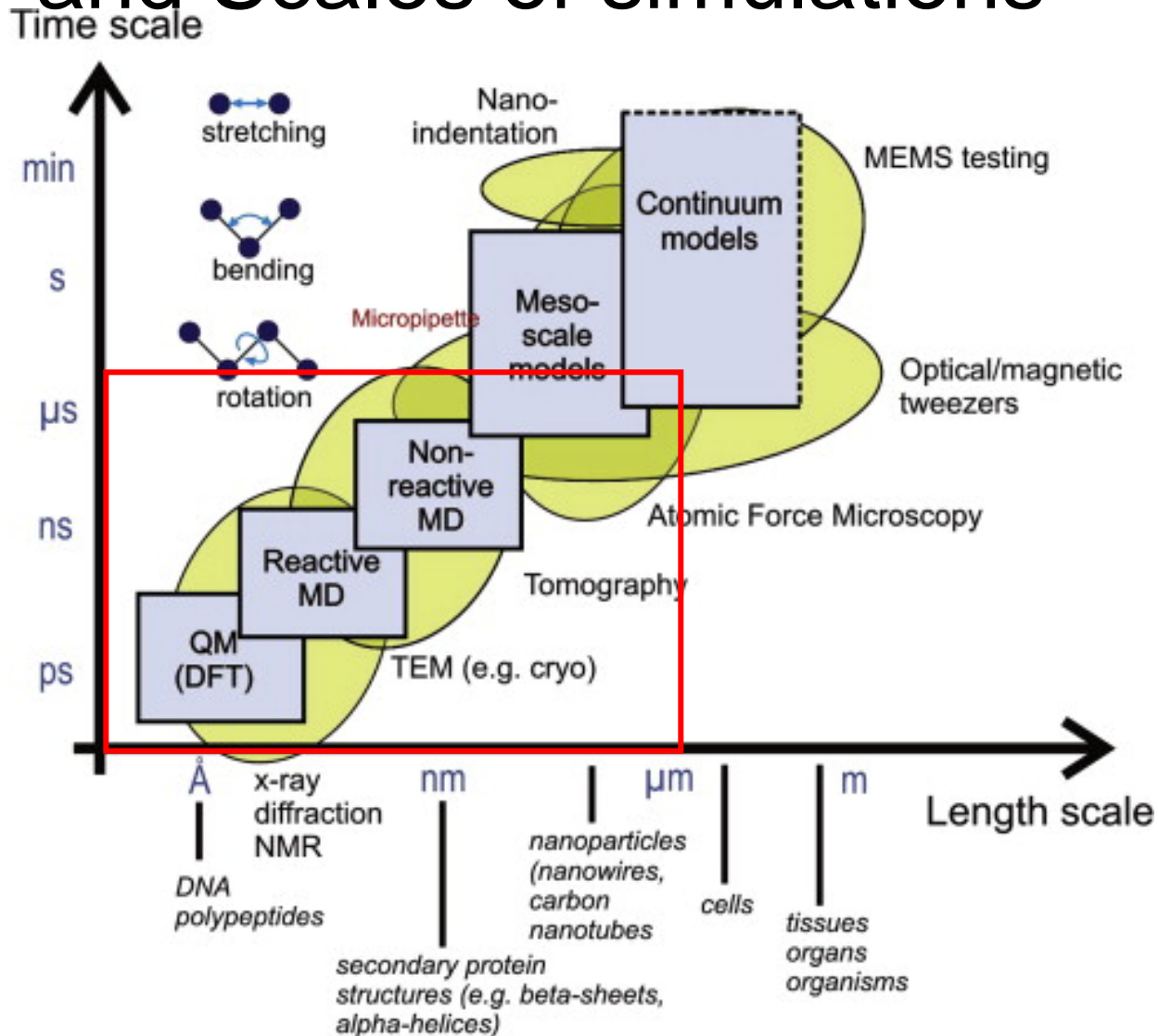
- $N$  Particles in three dimensional space
- $3N$  of degree of freedom
  - $r(x, y, z)$
  - $v(v_x, v_y, v_z)$
- One state has  $3N$  for positions and  $3N$  for momenta, defines a point in the  $6N$ -dimensional **phase space**
  - All possible state= phase space
  - One state = one point in phase space

# Ensembles (系综)

An ensemble can thus be considered to be a collection of the points in phase space

	Micro-canonical (微正则系综)	Canonical (正则系综)	Isothermal-isobaric (等温等压系综)	Grand canonical (巨正则系综)
Fixed thermodynamic variables	$N, V, E$	$N, V, T$	$N, P, T$	$\mu, V, T$
Systems	isolated	closed	closed	opened

# Forces calculation and Scales of simulations



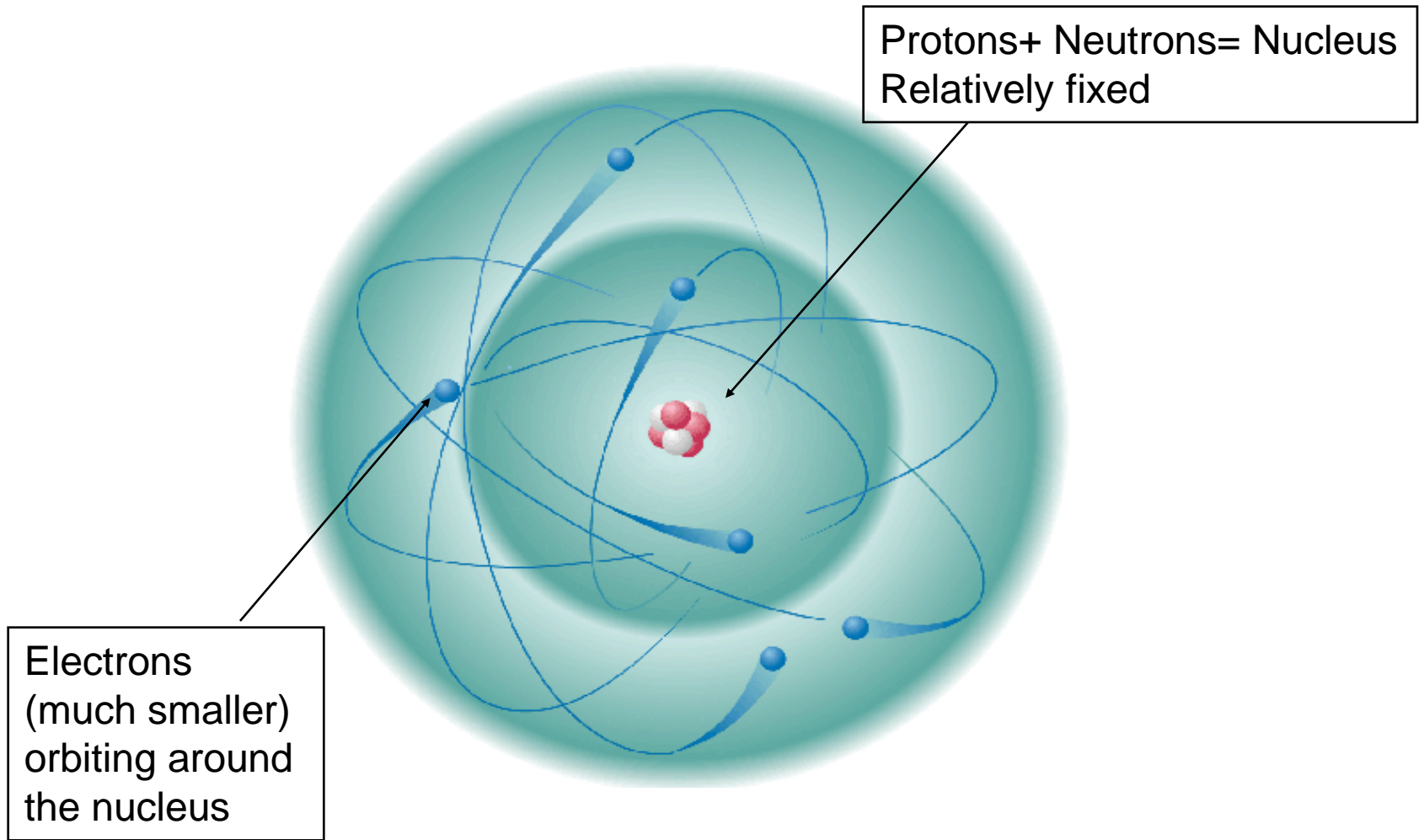
# Quantum Mechanical

## Ab initio calculations

- Born-Oppenheimer approximation
- Single electron approximation
- Hartree-Fock-Roothaan equation
- Separate movements of Nucleus and electrons
- Atomic orbitals
- LCAO-MO (Linear Combination of Atomic Orbital Molecular Orbital method)



# Atoms model



# More QM calculations

- Ab initio calculations
- Semi-empirical method
  - MINDO, MNDO, AM1, PM3
- Density Functional Theory (DFT)
  - PW91, B3LYP

# Relativistic effects (相对论效应)

Sc	[Ar] 3d <sup>1</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1				
Ti	[Ar] 3d <sup>2</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1			
V	[Ar] 3d <sup>3</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1	1		
Cr	[Ar] 3d <sup>5</sup> 4s <sup>1</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1	1	1	1
Mn	[Ar] 3d <sup>5</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1	1	1	1
Fe	[Ar] 3d <sup>6</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1	1	1
Co	[Ar] 3d <sup>7</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1	1
Ni	[Ar] 3d <sup>8</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1
Cu	[Ar] 3d <sup>10</sup> 4s <sup>1</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1	1↓	1↓	1↓	1↓
Zn	[Ar] 3d <sup>10</sup> 4s <sup>2</sup>	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓	1↓



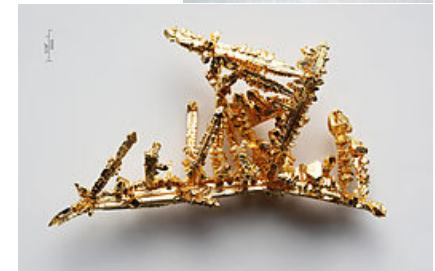
铬: [Cr] 3d<sup>5</sup> 4s<sup>1</sup> and not: [Cr] 3d<sup>4</sup> 4s<sup>2</sup>

铜: [Ar] 3d<sup>10</sup> 4s<sup>1</sup> and not: [Ar] 3d<sup>9</sup> 4s<sup>2</sup>



金: 4f<sup>14</sup> 5d<sup>10</sup> 6s<sup>1</sup>

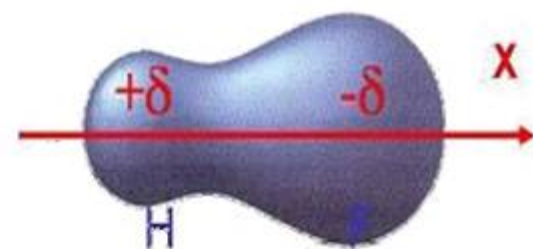
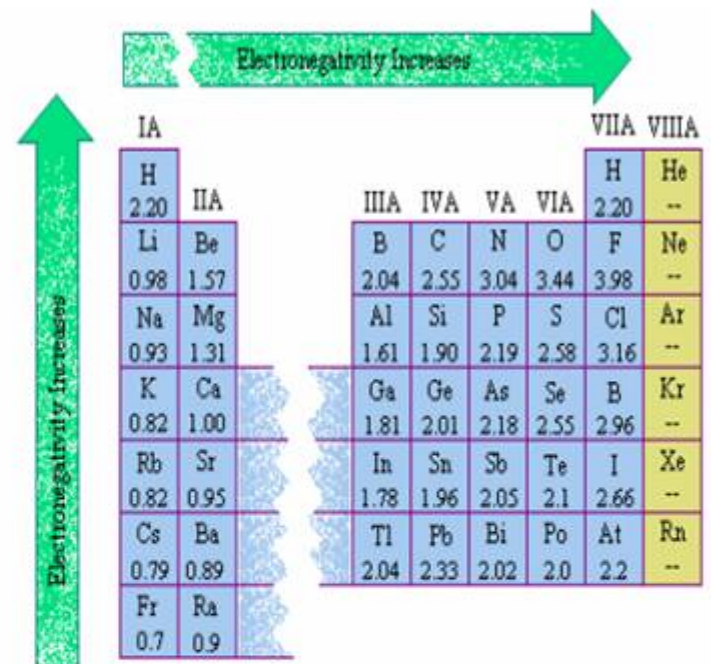
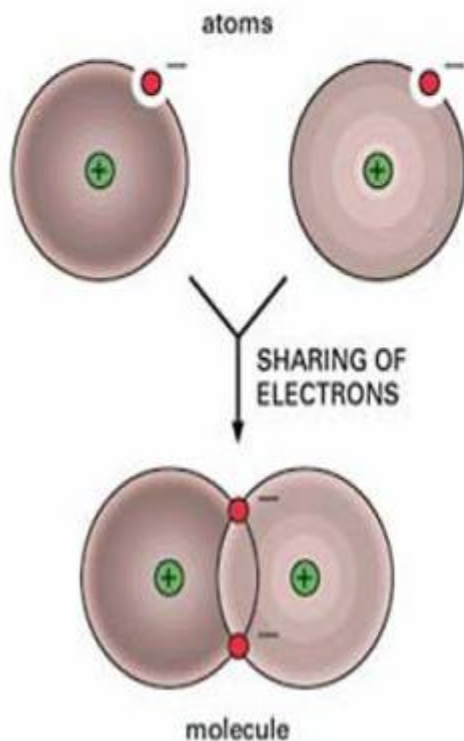
汞: 4f<sup>14</sup> 5d<sup>10</sup> 6s<sup>2</sup>



# “Pair” of Atoms

Types of bonds in protein:

Covalent bond:

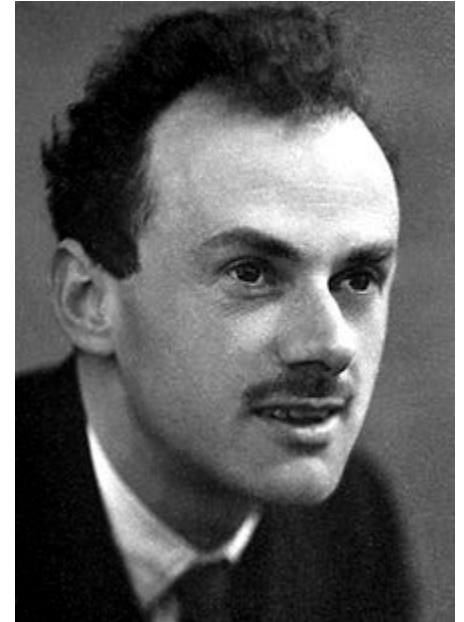


# Origin of Force Fields for Molecular Mechanics

## Quantum Mechanics

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations **much too complicated to be soluble**. It therefore becomes desirable that **approximate practical methods** of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.

-- Dirac, 1929



Nobel Prize in Physics  
for 1933

Paul Dirac  
Erwin Schrödinger

# Various Types of Force Fields

- Lattice models
- Off-lattice models
- United-atom force fields (OPLS, MM, AMBER, CHARMM19)
- All-atom force fields (OPLSAA, AMBER94, CHARMM22)
- Next generation force fields: Polarizable Force Fields (Polarizable OPLSAA, AMBER2002, TINKER PFF)

# Source of Parameters

- Experimental data
  - Spectra (bond stretching, angle bending ...)
  - X-ray crystal structures (geometry)
- High-level QM calculations
  - Solvent simulation to match physical properties
  - Ab initio calculations (partial charges, geometry ...)
  - Free energy perturbation in comparison to experimental data

# Force Fields: Typical Energy Functions

$$U = \sum_{\text{bonds}} \frac{1}{2} k_r (r - r_0)^2$$

Bond stretches

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

Angle bending

$$+ \sum_{\text{torsions}} \frac{V_n}{2} [1 + \cos(n\phi - \delta)]$$

Torsion rotation

$$+ \sum_{\text{improper}} V(\text{improper torsion})$$

Improper torsion (sp<sup>2</sup>)

$$+ \sum_{\text{elec}} \frac{q_i q_j}{r_{ij}}$$

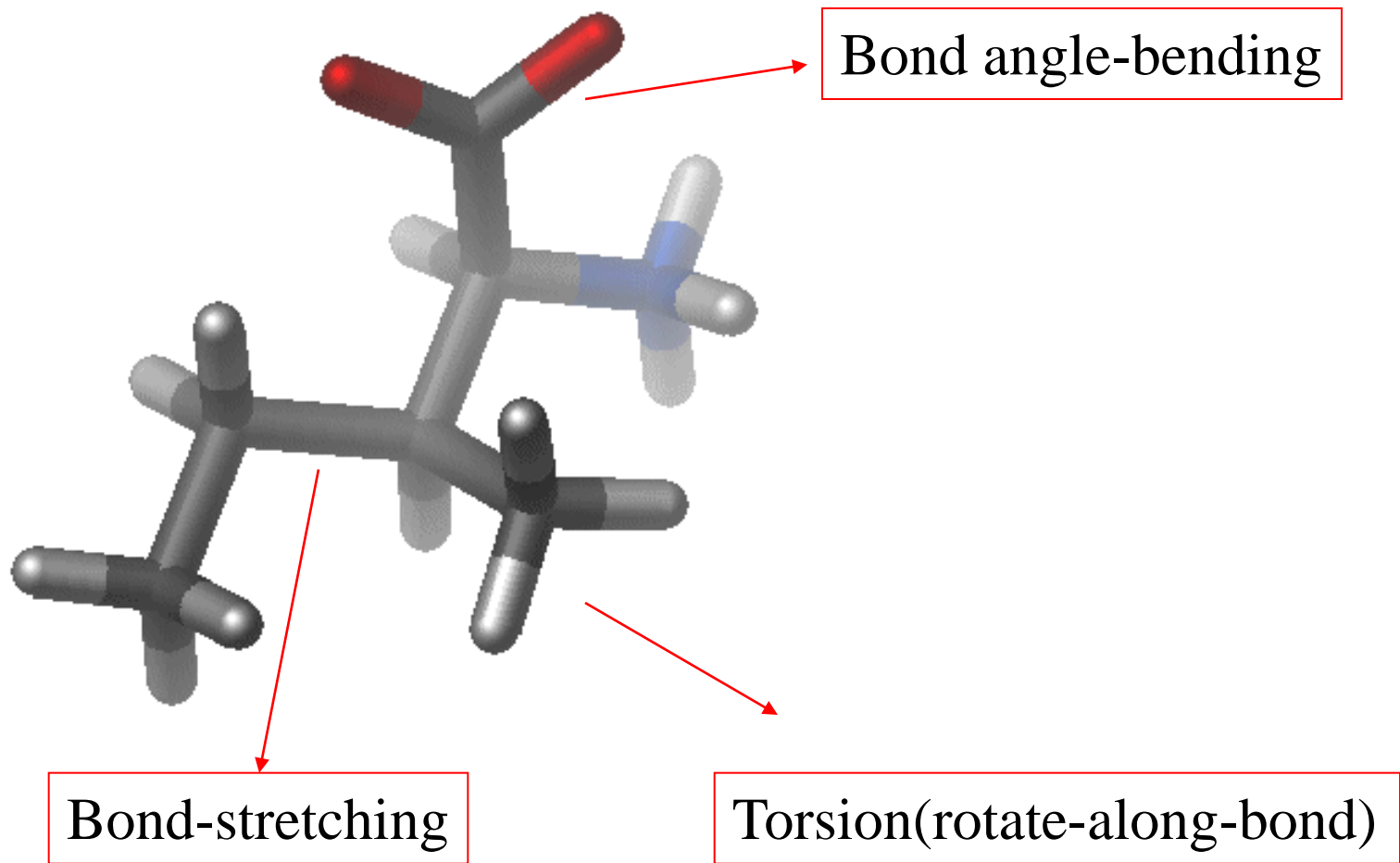
Electrostatic interaction

$$+ \sum_{\text{LJ}} \left[ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right]$$

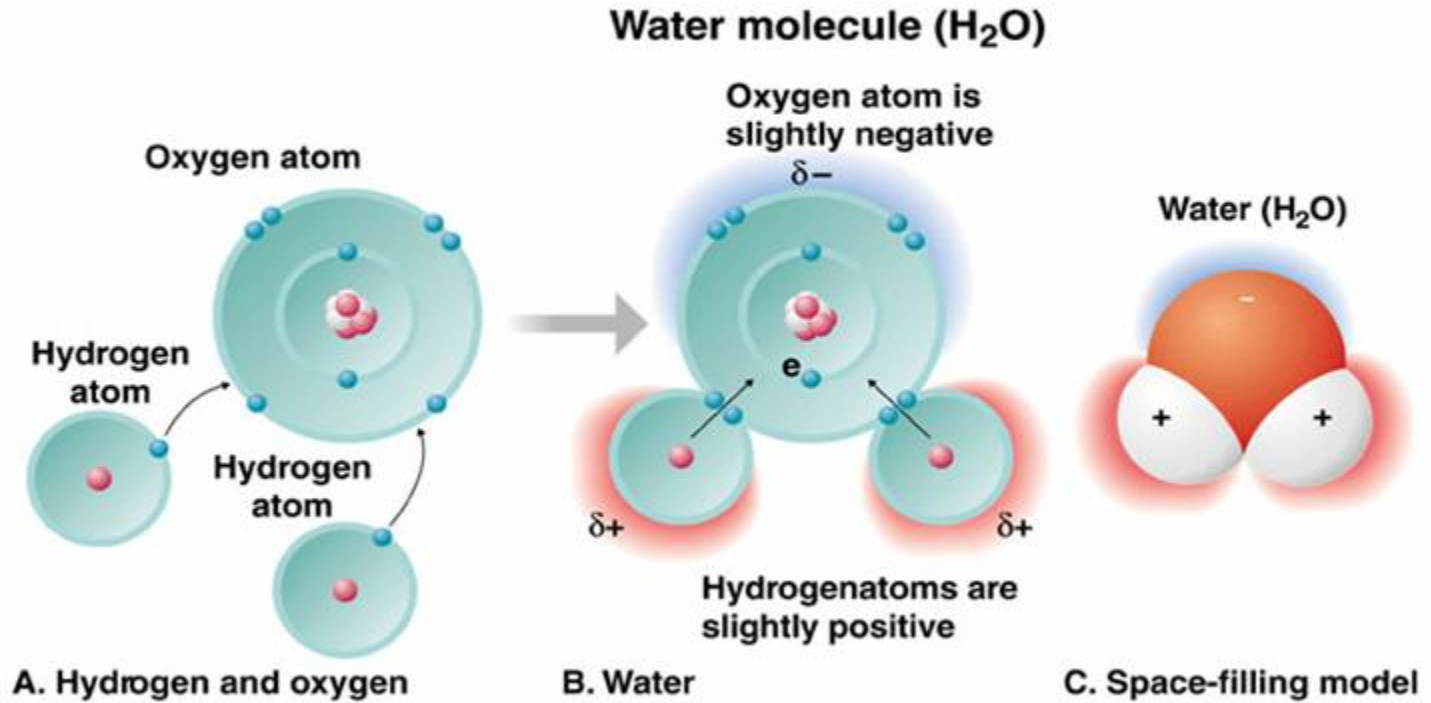
Lennard-Jones interaction



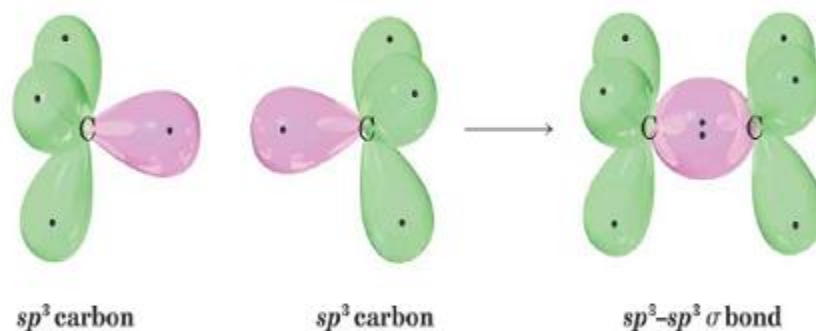
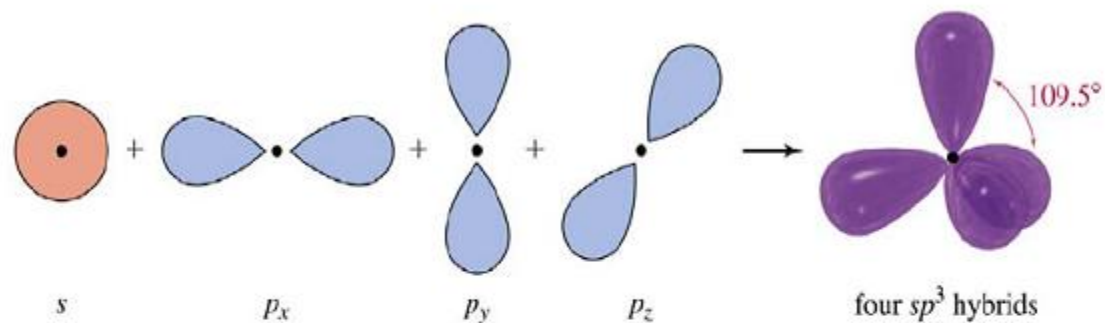
# Bonded potentials



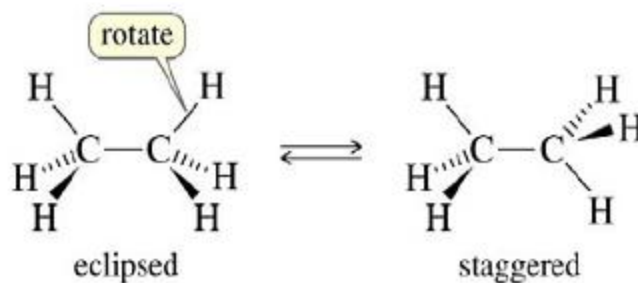
# Water molecule, H<sub>2</sub>O



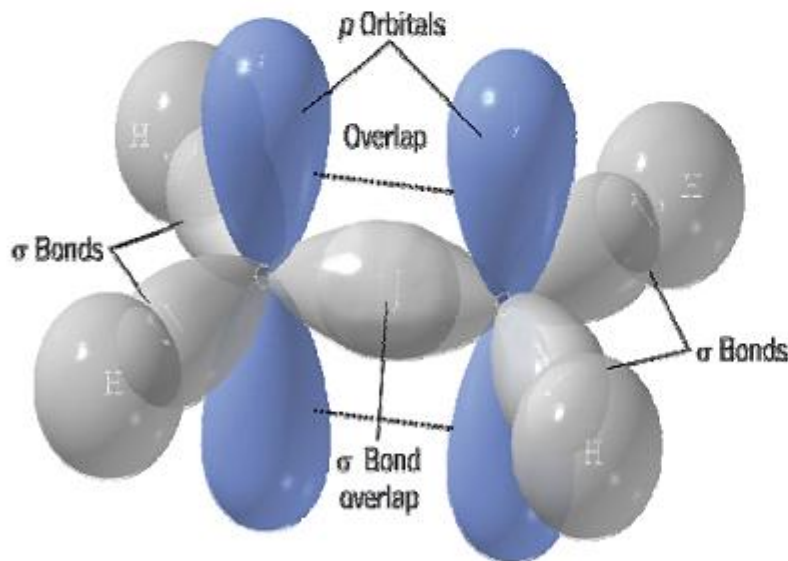
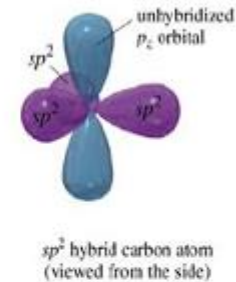
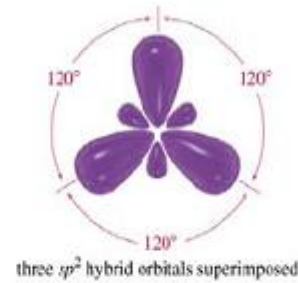
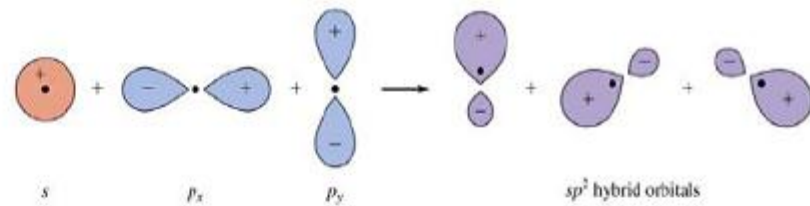
# Sp<sup>3</sup> hybridization: (tetrahedral structures)



“easy” rotation around the bond!.



# Sp<sup>2</sup> hybridization: (planar structures)

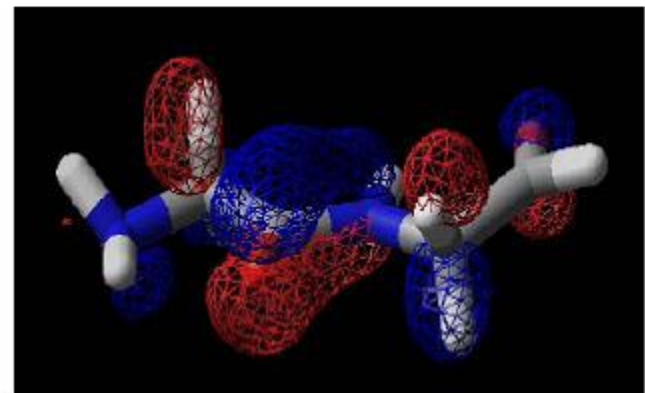
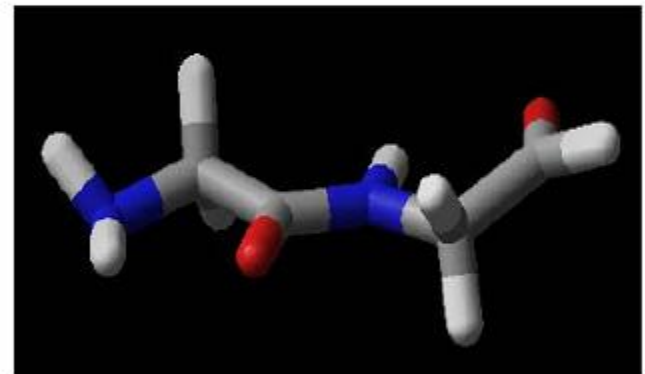
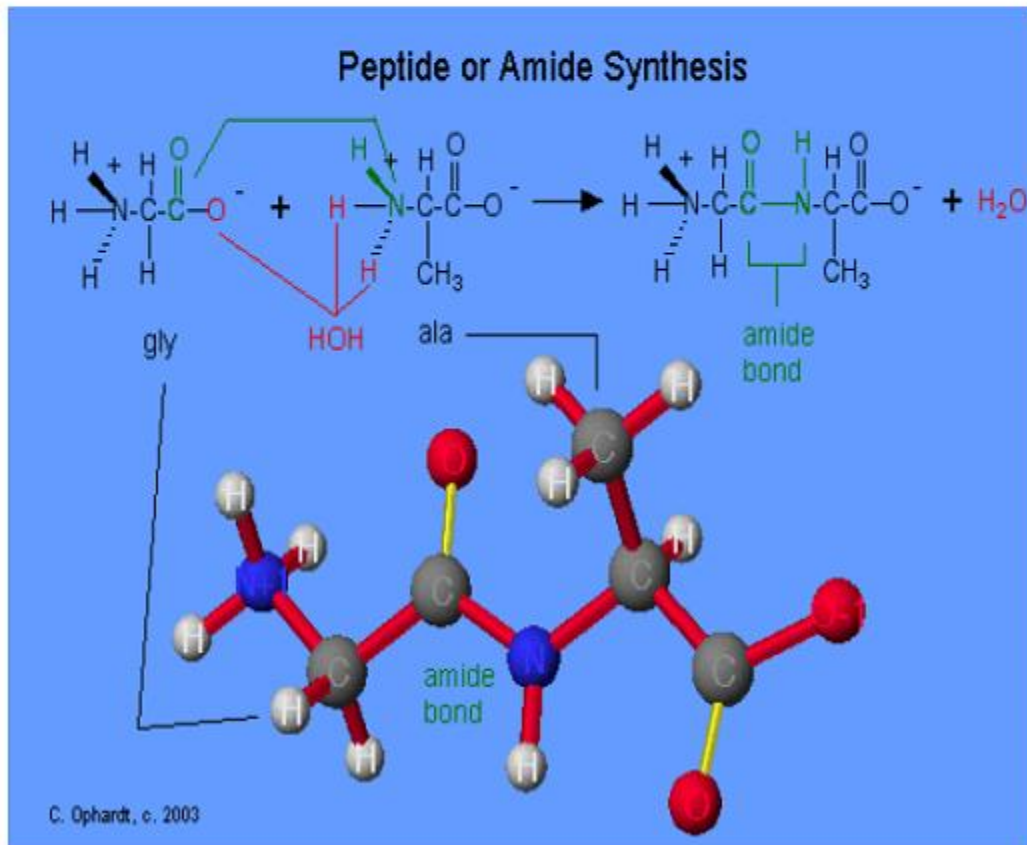
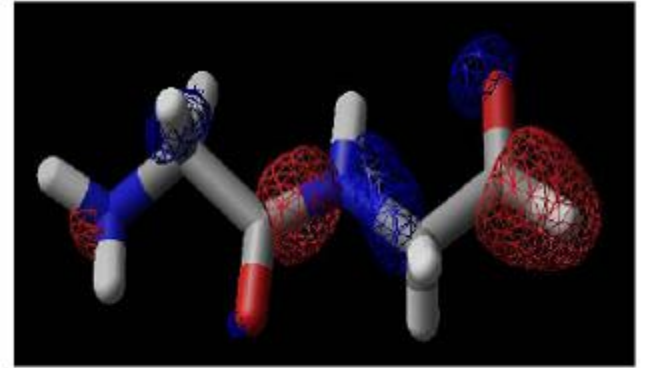


$\pi$  bond

“difficult” rotation around the bond!.

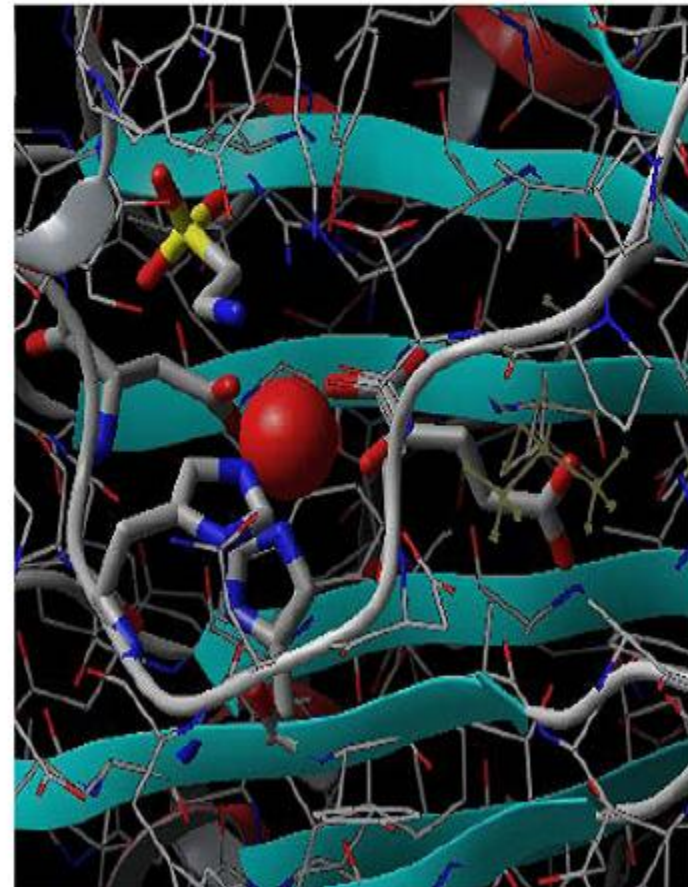
# RESIDUES

Lets look at the protein backbone



# Coordination (Metal) bond: 配位键

Geometry	No. of ligands	hybridization
Octahedral	6	$d^2sp^3$
Trigonal bipyramid	5	$dsp^3$
Tetrahedral	4	$sp^3$
Square planar	4	$dsp^2$
Linear	2	$sp$



# Intermolecular interactions:

weak

London Dispersion (Van der Waals)

Dipole-Dipole

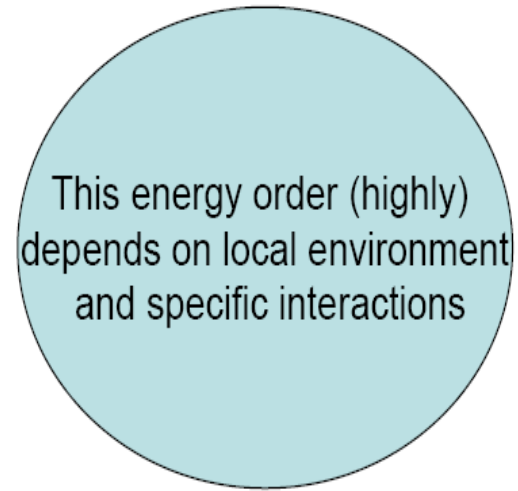
Stacking

Hydrogen Bond

strong

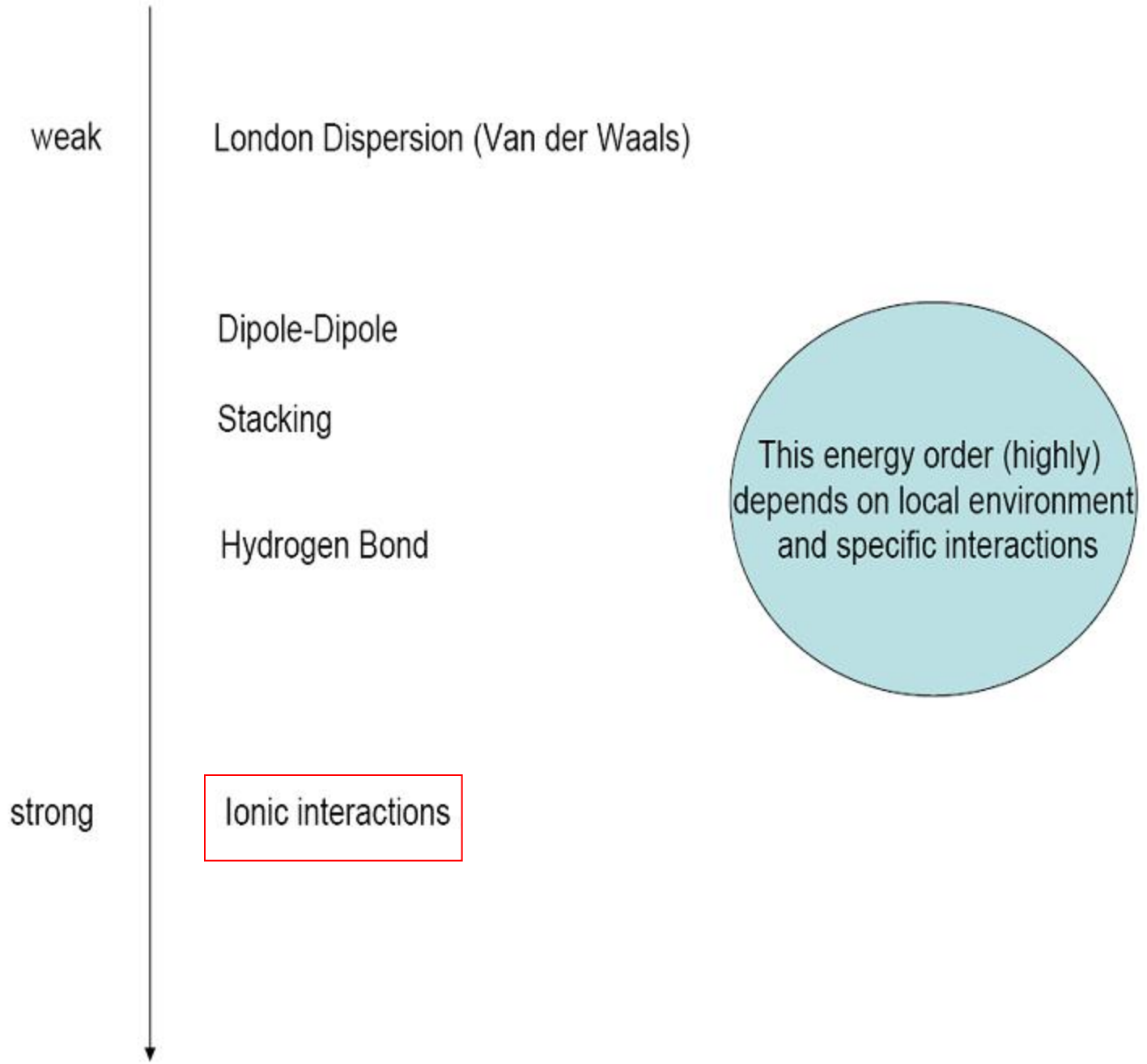
Ionic interactions

# Non-bonded interactions



This energy order (highly) depends on local environment and specific interactions

# Intermolecular interactions:

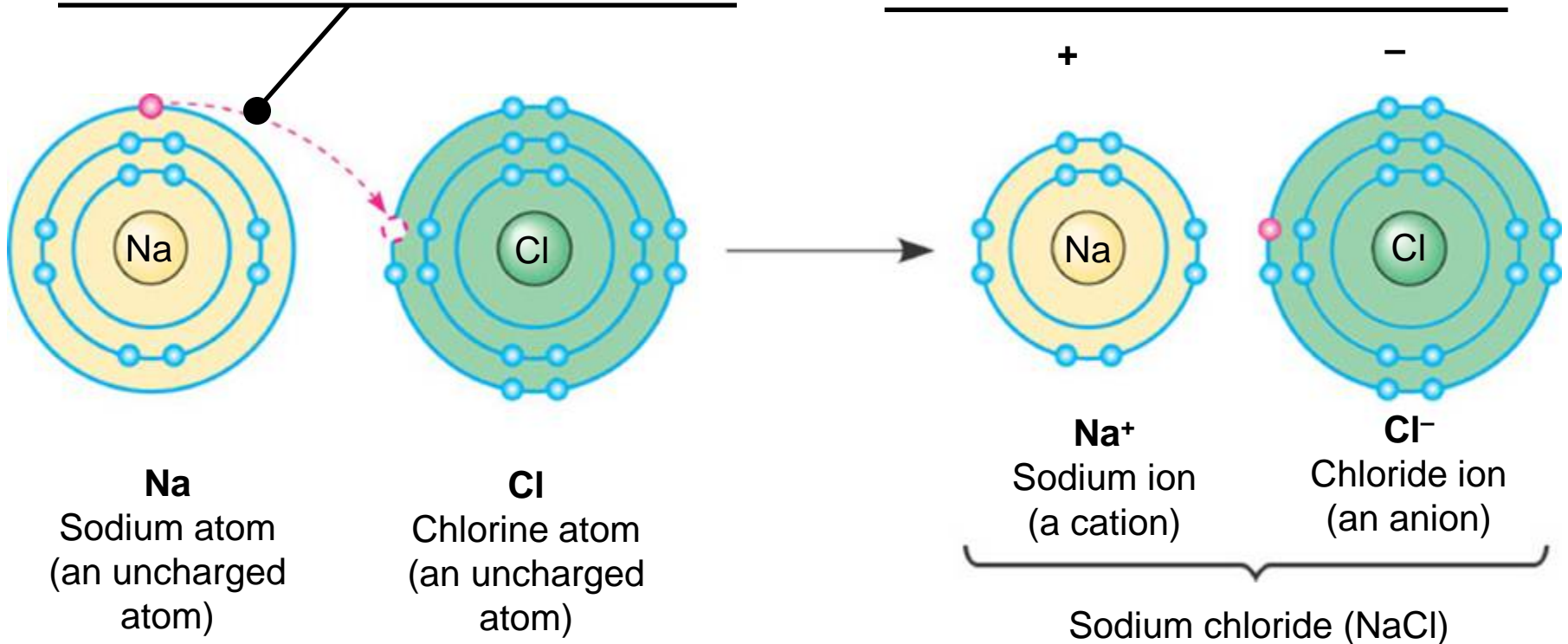




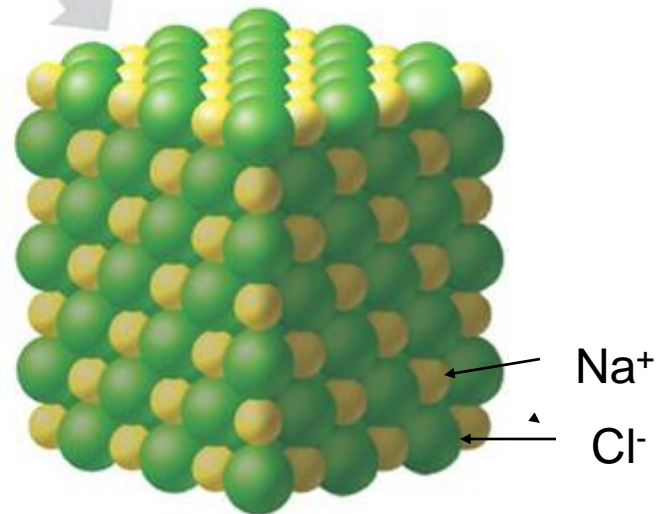
# Electron transfer and ionic bonding

1 The lone valence electron of a sodium atom is transferred to join the 7 valence electrons of a chlorine atom.

2 Each resulting ion has a completed valence shell. An ionic bond can form between the oppositely charged ions.

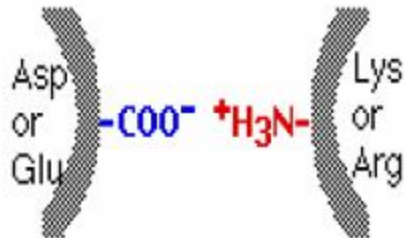


# Sodium chloride crystal

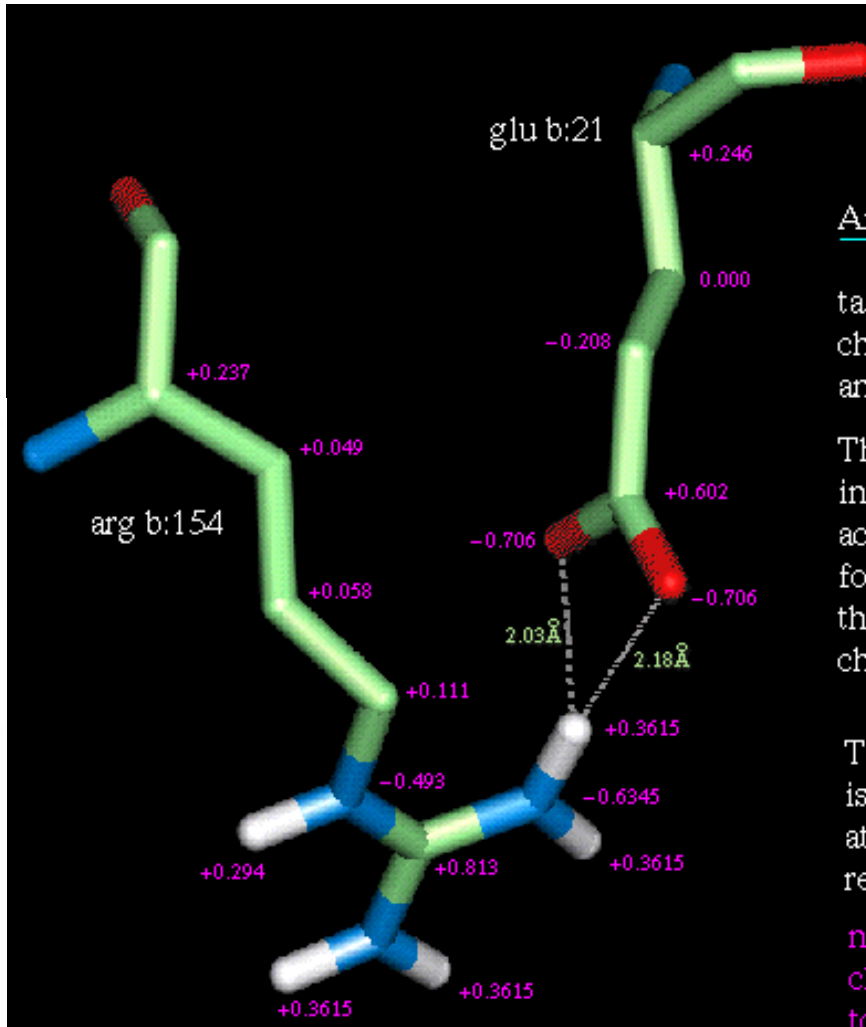


# Salt Bridge

**A relatively weak ionic bond between positively and negatively charged side-chains of proteins.**



**Ionic Interactions**



An example of a salt bridge

taken from the structure of chymotrypsin – due to D.M. Blow and coworkers.

There is an electrostatic interaction between the glutamic acid b:21 which carries a formal charge of -1 unit and the arginine b:154 which has a charge of +1.

The charge of each residue is not confined to a particular atom but spread across the whole residue:

numbers in purple indicate the charge (in electron units) according to the AMBER united atom PEF (Weiner *et al.*, JACS 106:765-784, 1984)

# Intermolecular interactions:

weak

London Dispersion (Van der Waals)

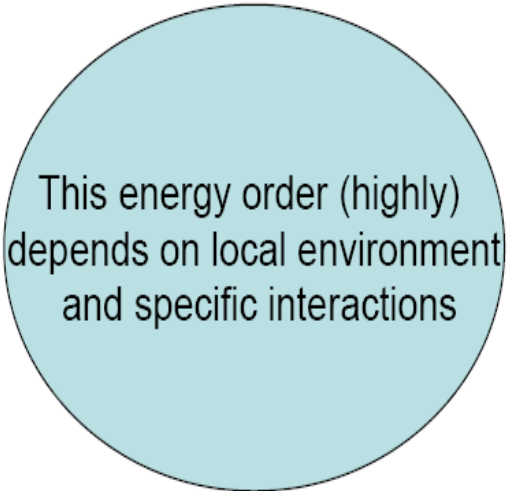
Dipole-Dipole

Stacking

Hydrogen Bond

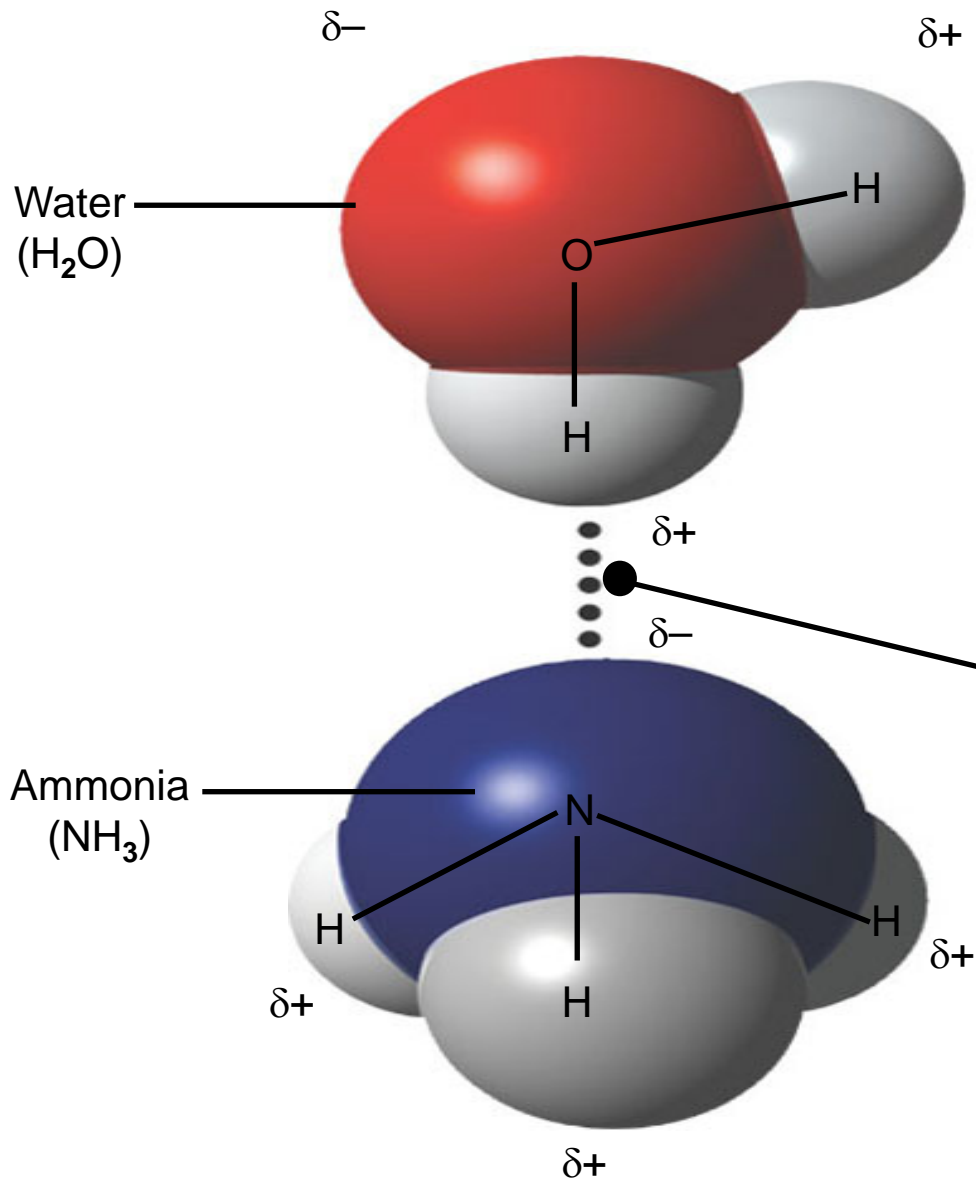
strong

Ionic interactions



This energy order (highly) depends on local environment and specific interactions

# Hydrogen bond



A hydrogen bond results from the attraction between the partial positive charge on the hydrogen atom of water and the partial negative charge on the nitrogen atom of ammonia.

# Hydrogen bond

## The Chemical Foundations of Life

So there is an attraction between the positive and negative sides of a water molecule... this is **hydrogen bonding**.

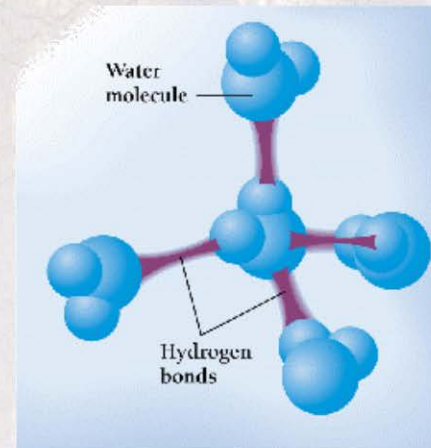
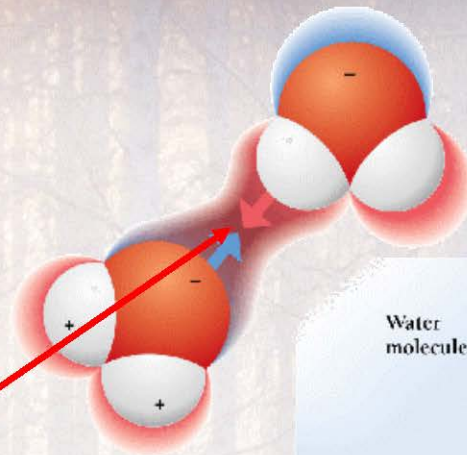
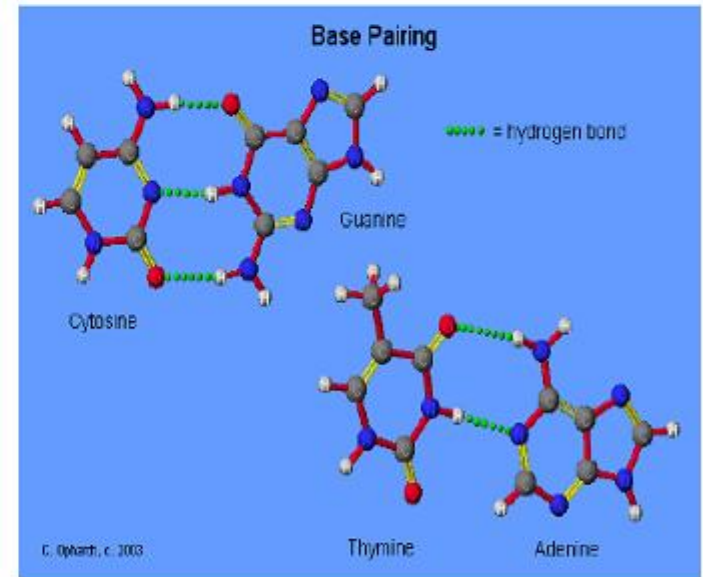
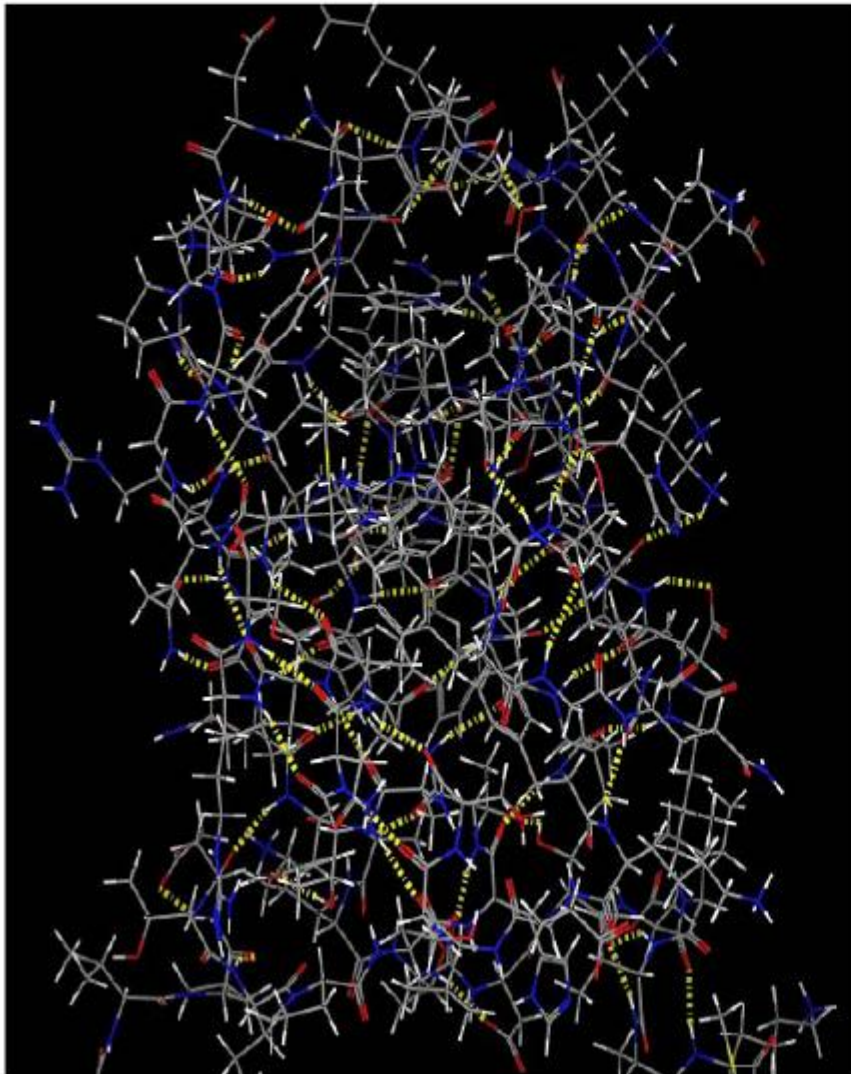
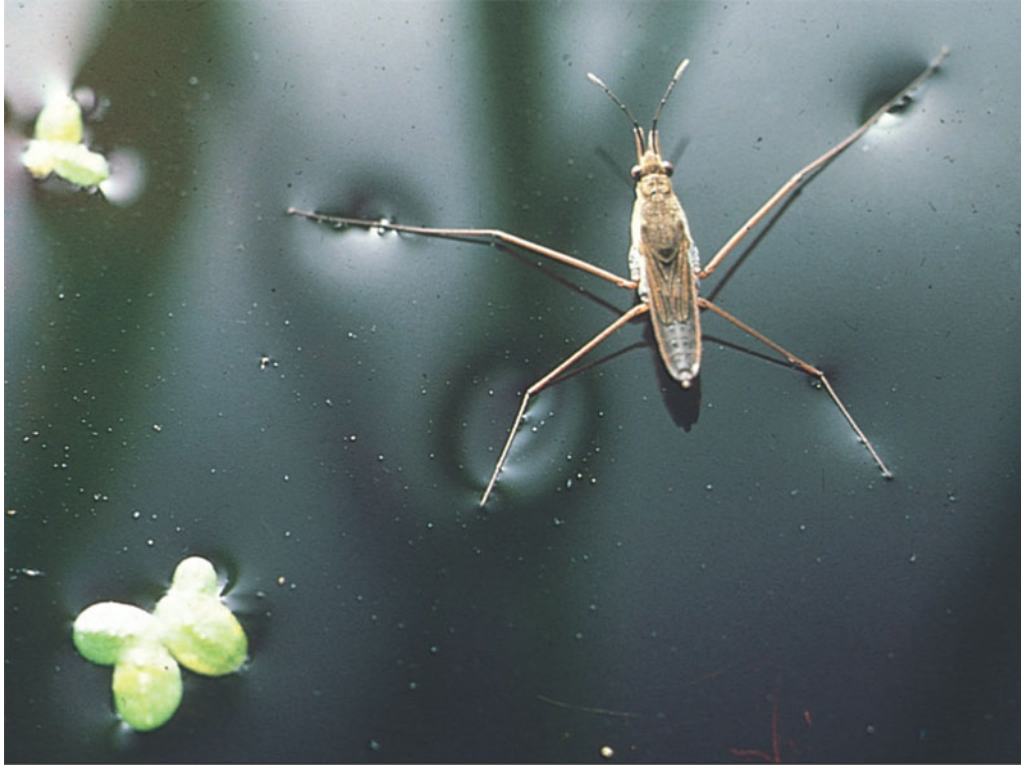


Figure 2-12

# Hydrogen Bond H "shared by two more electronegative groups



# Walking on water



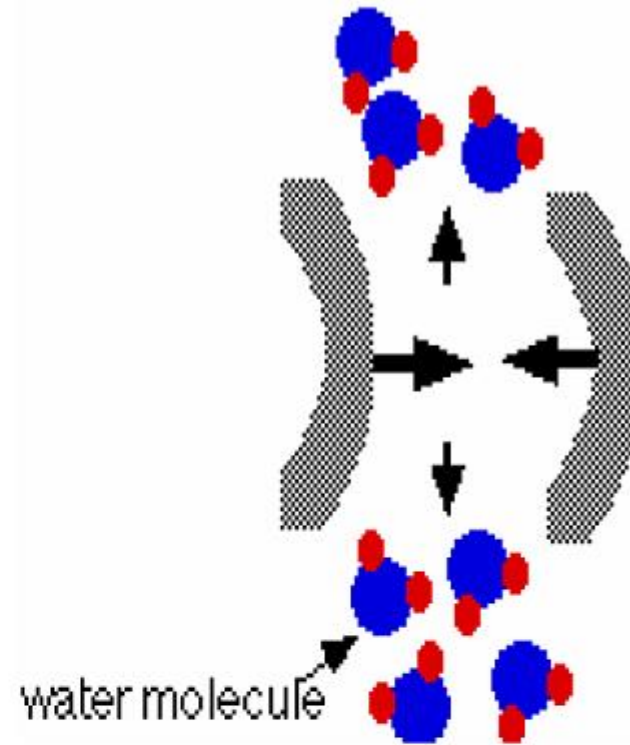
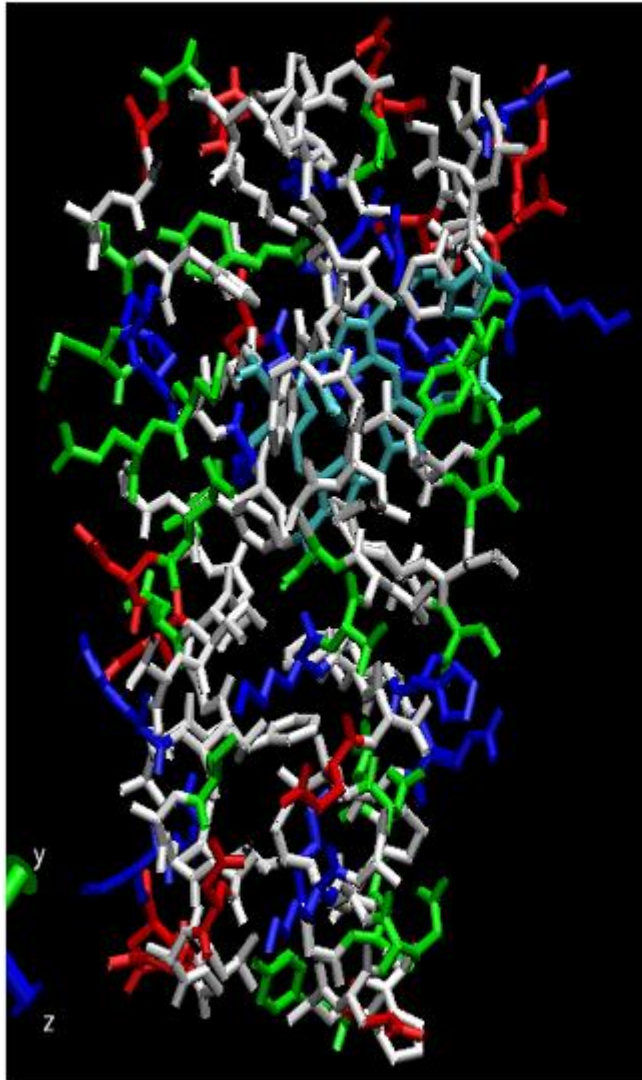


# Hydrophobic Effect



**A droplet of water forms a spherical shape to minimize contact with the hydrophobic leaf.**

Hydrophobic effects (entropic).

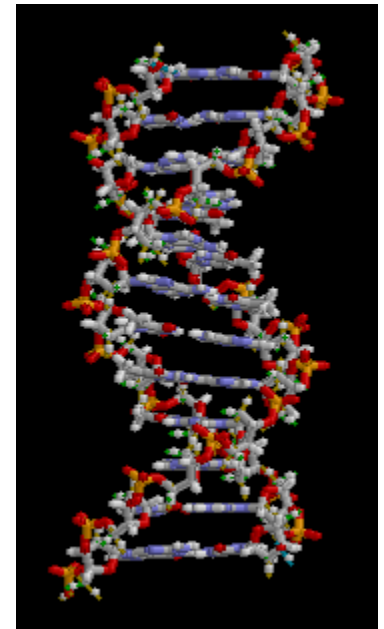
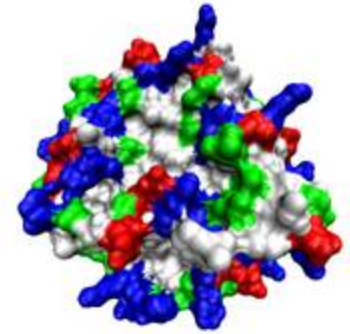


Water molecules and molecules  
Consist of carbon and hydrogen  
Would try to avoid each other!

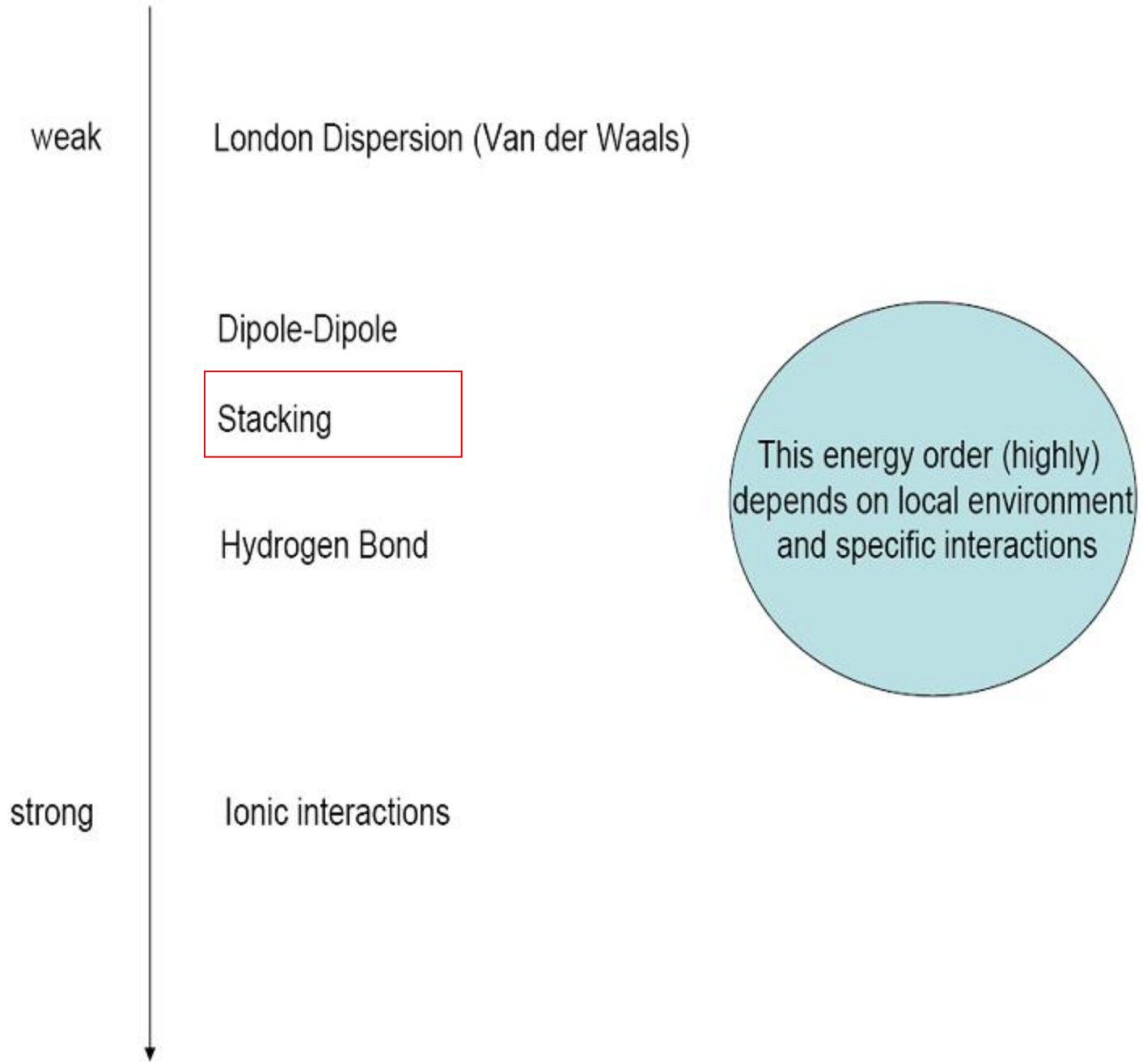
**Basis for globular proteins**

# Hydrophobic Effect in Biology

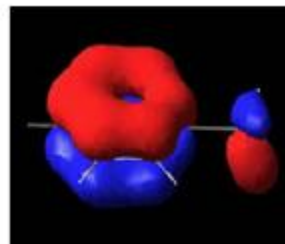
- Protein folding, the hydrophobic
  - Hydrophobic amino acids, such as Ala, Val, Leu, Ile, Phe, Met grouped together within the protein
  - Charged or polar side chains on the solvent-exposed surface where they interact with surrounding water molecules.
- The DNA Double Helices structure
  - backbones of sugars and phosphate groups outside
  - the aromatic base pairs inside



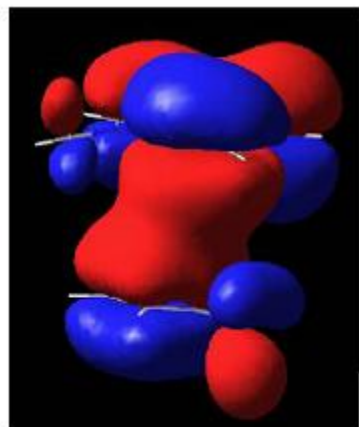
# Intermolecular interactions:



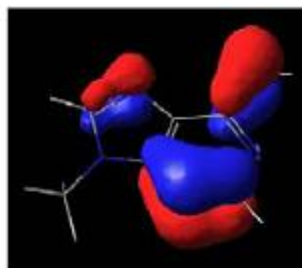
# 1) "Classical" view of stacking.



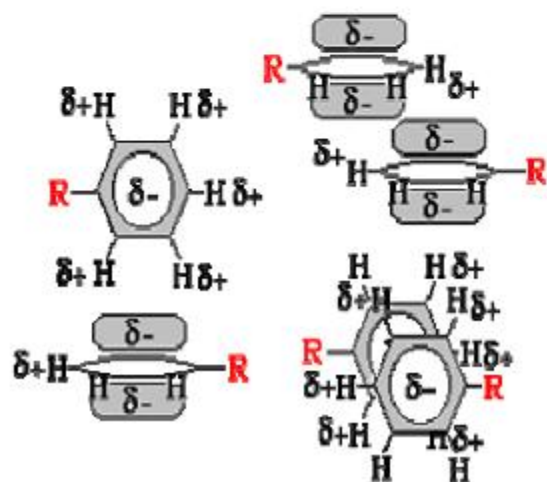
PHE56  
 $E = -0.485$



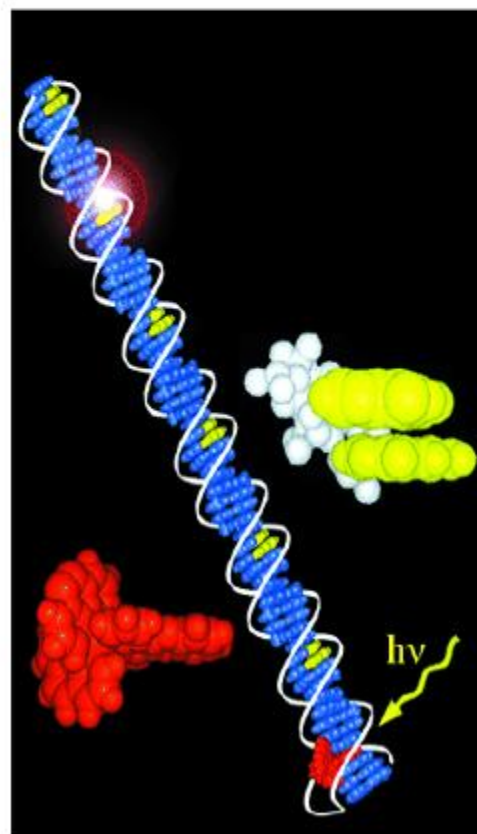
A11  
 $E = -0.489$



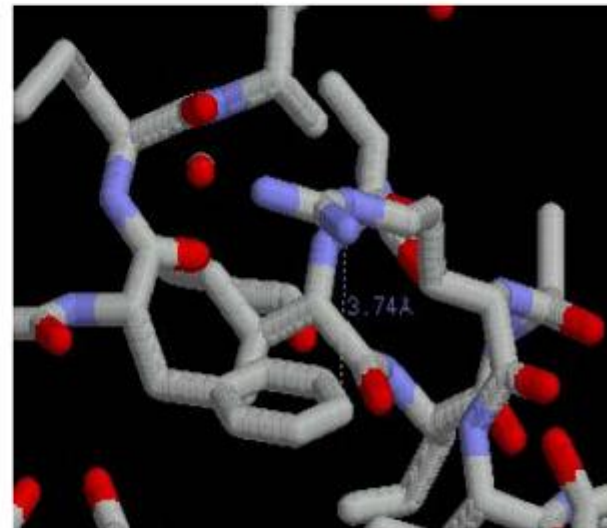
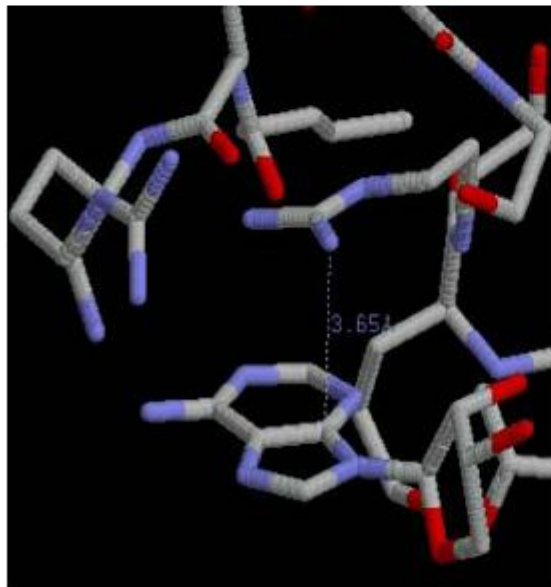
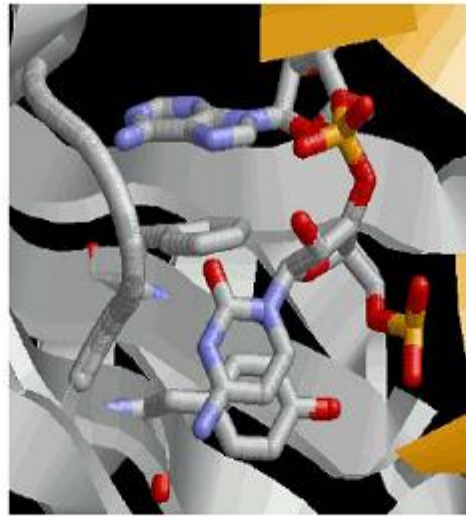
# 1) "Modern" view of stacking.



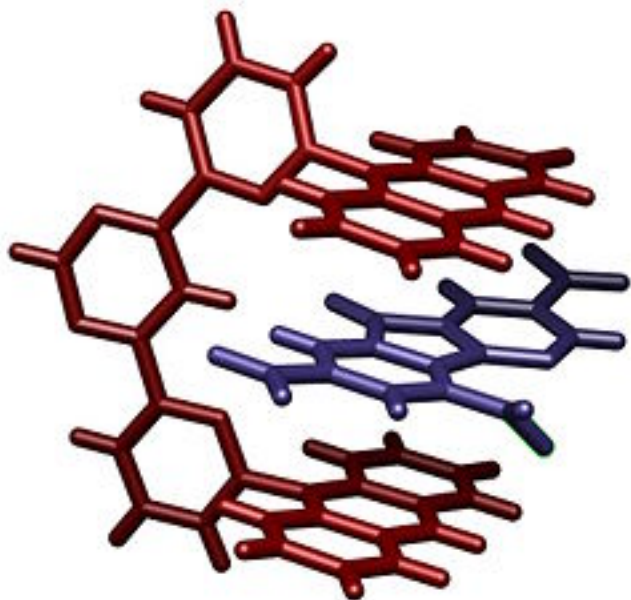
# Dna-mediated electron transfer



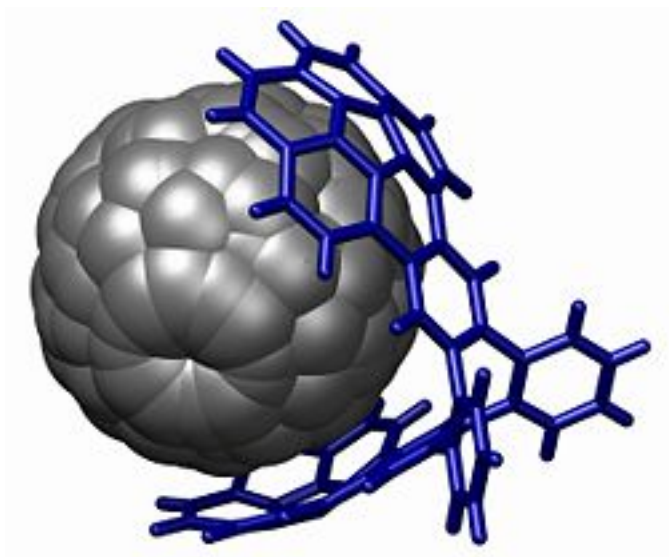
# Stacking examples



Arginine stacking

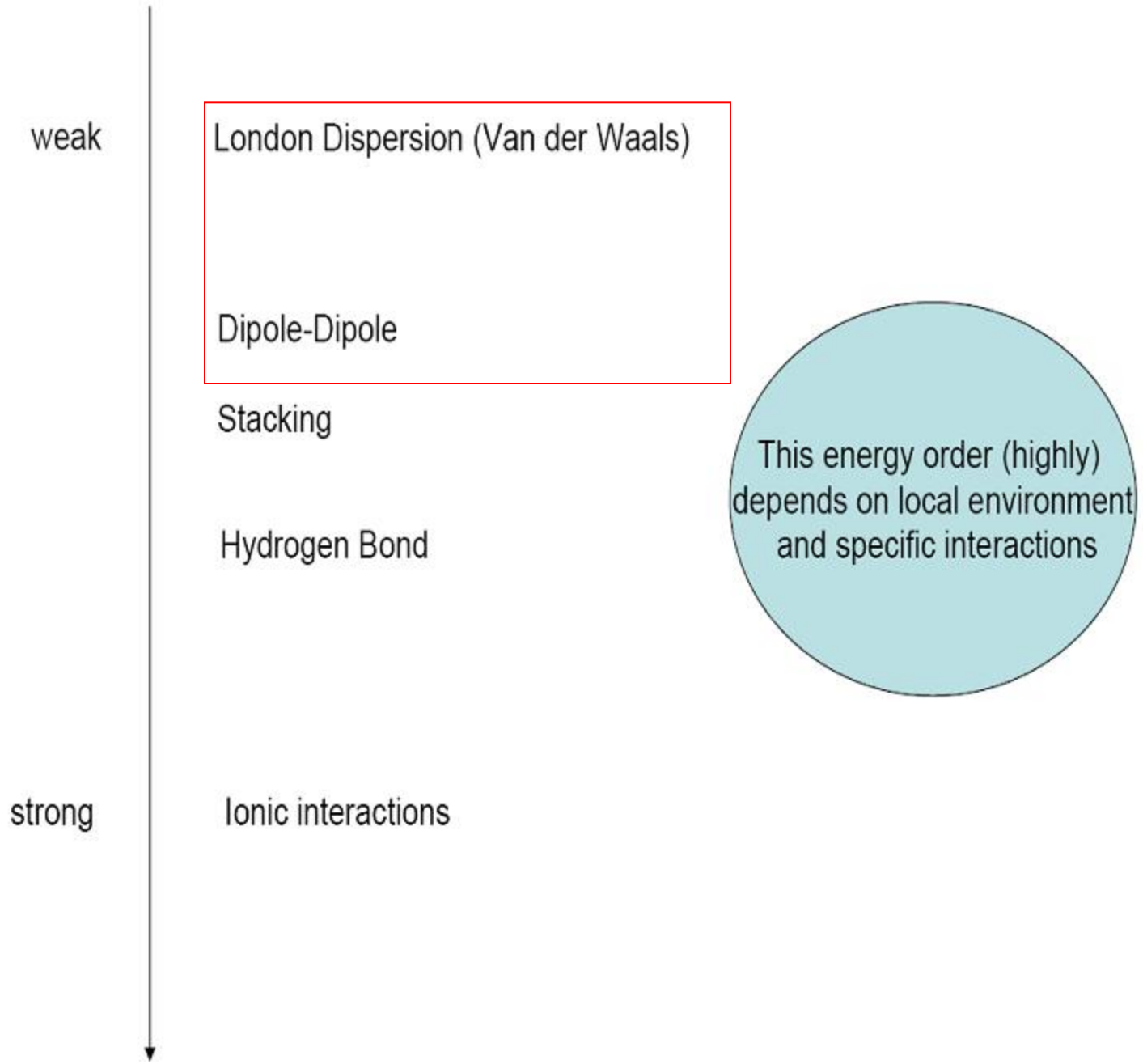


Crystal structure of a trinitrofluorene bound in molecular tweezers through aromatic stacking interactions. Reported by Lehn and coworkers.[\[1\]](#)



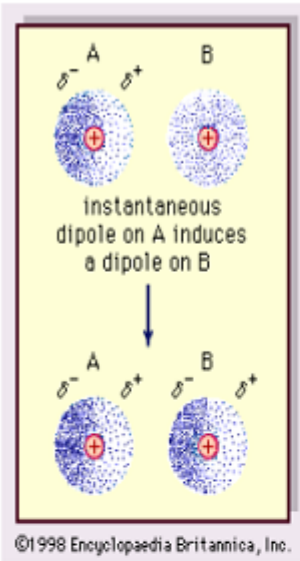
Crystal structure of a fullerene bound in a buckycatcher through aromatic stacking interactions. Reported by Sygula and coworkers.[\[2\]](#)

# Intermolecular interactions:



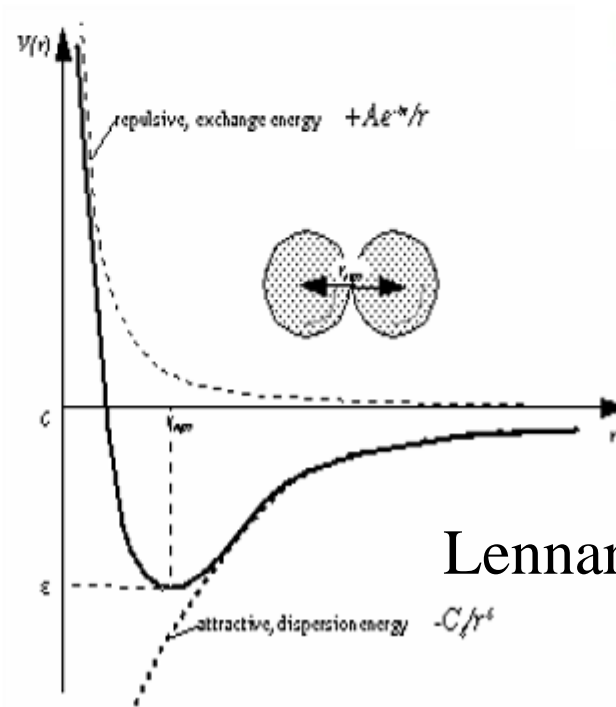
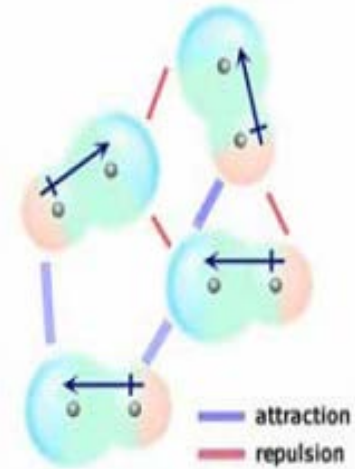


# Dispersion forces (Van der waals): induced/instantaneous dipoles



## Permanent Dipole-Dipole

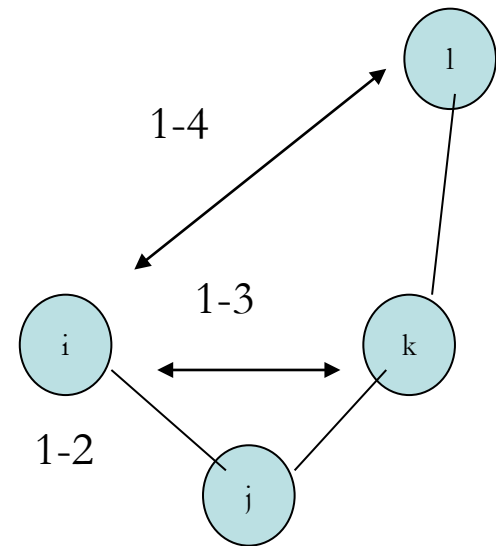
取向力  
诱导力  
色散力



## Lennard-Jones potential

# 1-4 Non-bonded Interactions

- Non-bonded exclusions
  - 1-2 and 1-3 interactions excluded
  - 1-4 interactions partially excluded
- Partly?
  - included in bonded terms already
  - If included, they would be huge to cause stability problem
- 1-4 interaction scalings
  - OPLSAA scales by 0.5 for both electrostatic and LJ
  - AMBER94 scales 0.5 for LJ and 1/1.2 for electrostatic interaction
  - CHARMM22 has special 1,4-terms
- 1-4 terms vs torsion terms



Even though they are non-bonded interactions, 1-4 terms are often calculated along with bonded terms.

# Model construction

- Read in parameters
  - Force field parameters
  - Topology definitions
- Read in sequences
  - *Protein structure file (PSF)*
  - segments
- Read in coordinates
  - .pdb, crd, etc
  - From RCSB PDB, unpublished data
  - X-ray, NMR, homology modeling

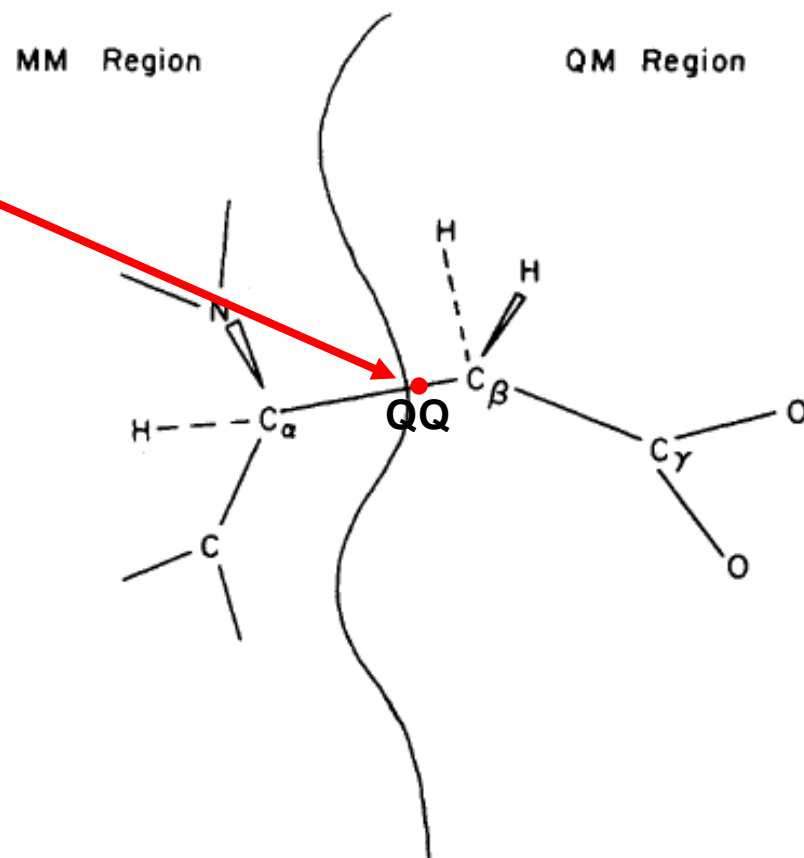
# Revision on models

- Revision on solvated model
  - Mutation
  - Construction of S-S bond
  - Protonation/deprotonation
  - Deletion of unnecessary/unreasonable solvents
  - Modification on coordinates
  - Change origin of coordinates
- Add link-atoms in QM/MM simulations

# Bonds on QM/MM interface

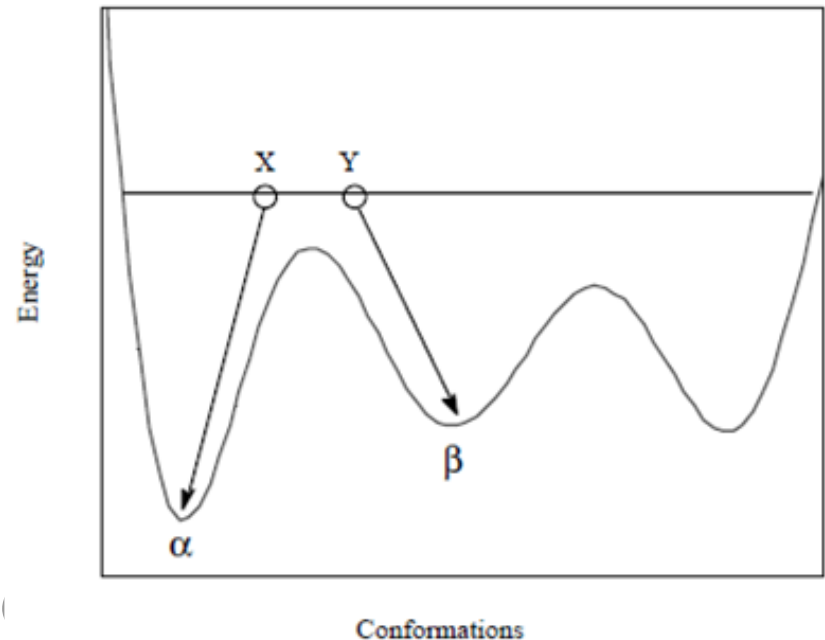
## Link-atom method

- A virtual “link-atom” is added along the covalent bond
- Initially 1Å from the QM atom
- Generally a dummy hydrogen atom
- Satisfy bonding in MM and the valency of the QM atom



# Model optimization

- Original structure under different conditions
  - Temperature
  - Solvent conditions
  - **No hydrogens**
- Revision on structure
  - hydrogens added
  - substrate or ligand ad
  - mutations on residues



# Minimization algorithms

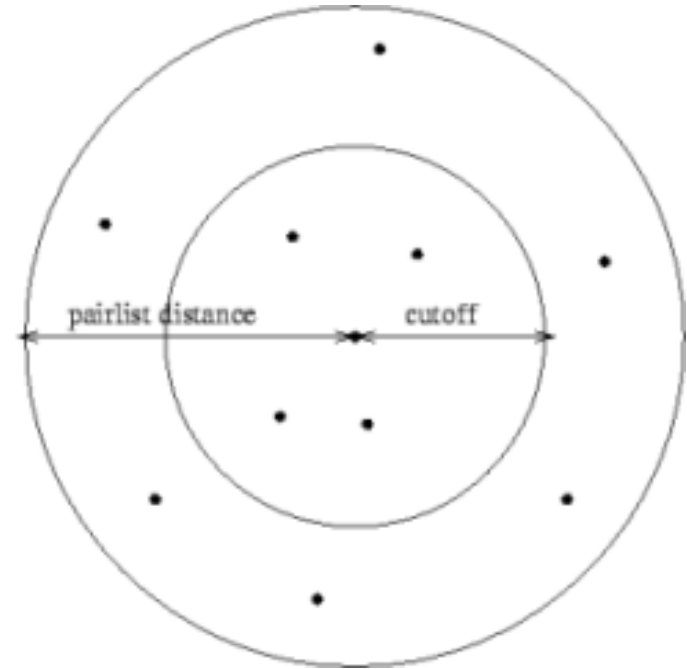
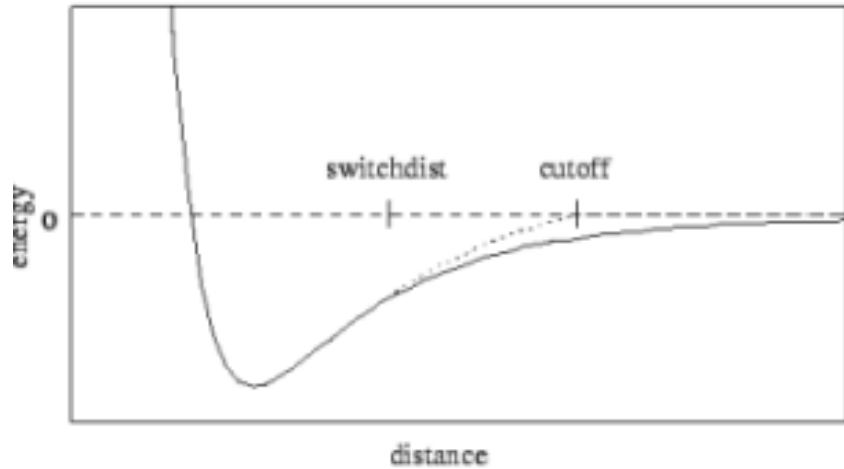
- Steepest-descent (SD)
  - Simplest, first derivative method
  - Poor convergence and accuracy
  - initial steps to relieve bad contacts
- Newton-Raphson (NR)
  - Expensive, second derivative method
  - Good convergence and accuracy
- Adopted-Basis Newton-Raphson (ABNR) and Conjugated Gradients (CONJ)
  - Intermediate first derivative method
  - Very useful for large systems

# Minimization on Energies

- $E_{QM}$ 
  - QM methods
  - Possible for bond breaking-making
  - Expensive, time-consuming
- $E_{MM}$ 
  - Force field
  - Impossible for bond breaking-making
  - Empirical parameters, quick and cheap
- Complicated energy
  - $E_{total} = E_{QM} + E_{MM} + E_{QM/MM}$
  - ONION, SMBP, GSBP



# Potential truncation

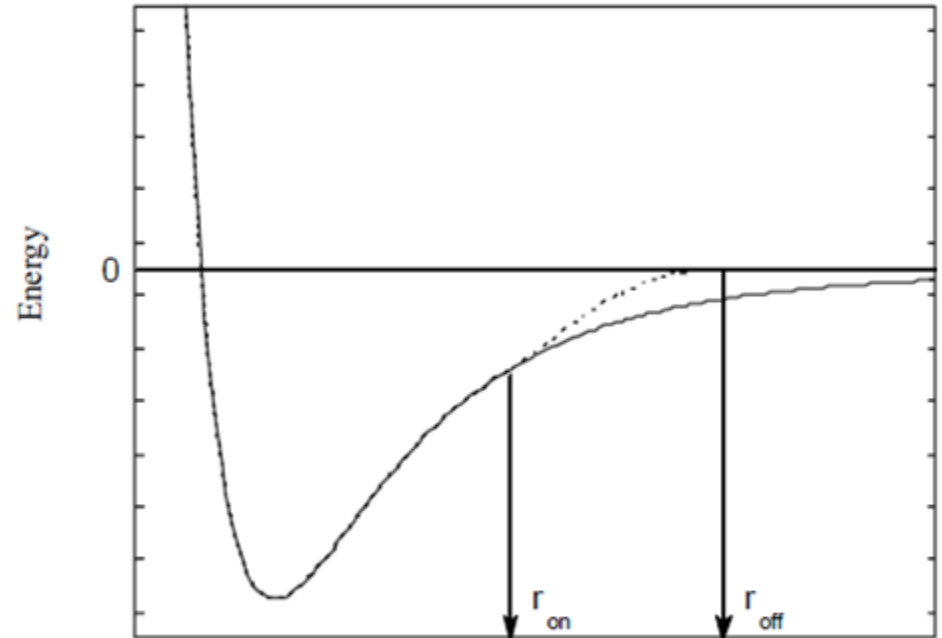


In the MD or MC, in order to save time, we introduce a spherical cutoff. The cutoff distance should be sufficient large to ensure that it is a small perturbation, but it must be no greater than  $BOXL/2$  for the consistency with minimum image convention.

# Truncation for non-bonded interactions

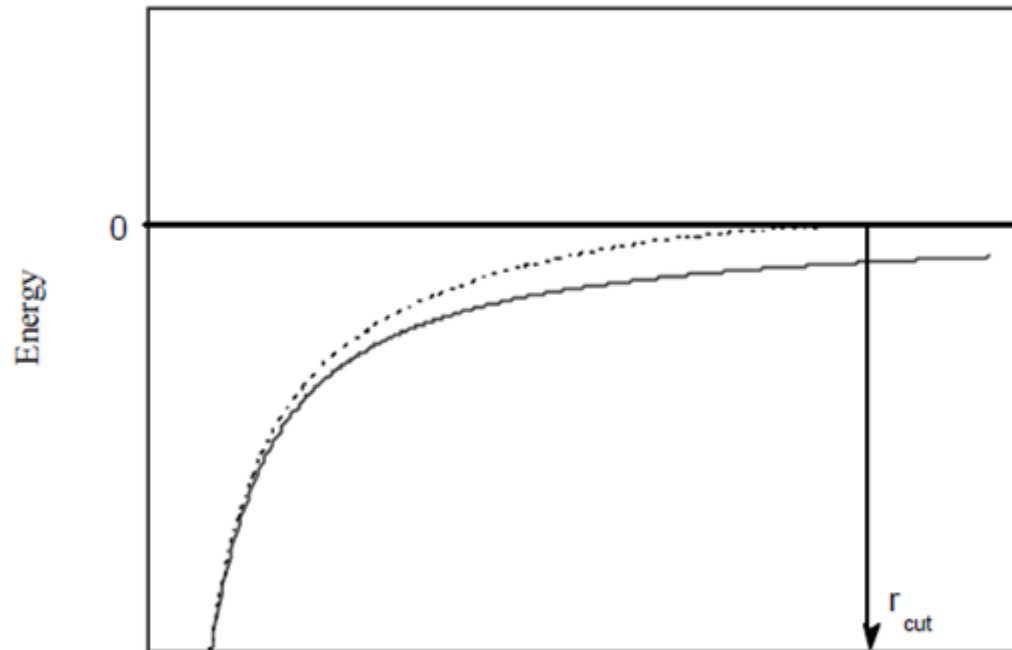
- Assuming the overall contribution from atoms farther than certain CUTOFF distance is quite small
- A gradual truncation in energy leads to  $E=0$  when  $r > r_{\text{cutoff}}$ 
  - SWITCH
    - Limit truncation in a small range [ $r_{\text{on}}$ ,  $r_{\text{off}}$ ]
    - Generally applied to van der waals potential
  - SHIFT
    - Applied to whole range of distances
    - Generally applied to electrostatic potential

# Switch truncation function



$$SW(r, r_{on}, r_{off}) = \left\{ \begin{array}{ll} 1 & r < r_{on} \\ \frac{(r_{off} - r)^2 (r_{off} + 2r - 3r_{on})}{(r_{off} - r_{on})^2} & r_{on} < r < r_{off} \\ 0 & r > r_{off} \end{array} \right\}$$

# Shift truncation function



$$sh(r, r_{\text{cut}}) = \left\{ \begin{array}{ll} \left[ 1 - \frac{2r^2}{r_{\text{cut}}^2} + \frac{r^4}{r_{\text{cut}}^4} \right] & r < r_{\text{cut}} \\ 0 & r > r_{\text{cut}} \end{array} \right\}$$

# Constraints

- An extra energy term
  - Removed when calculating real energy
- Force the molecules into a desired conformation
  - Prohibit dramatic movements
  - Simulate specific interactions
  - Limit the system into desired conformation
  - Umbrella sampling

# Typical constraints

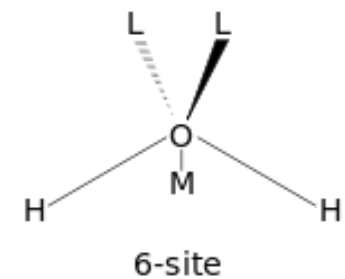
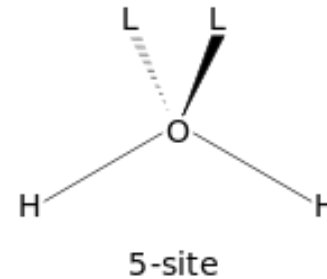
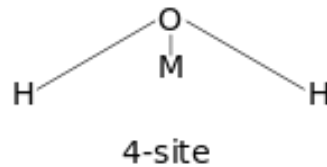
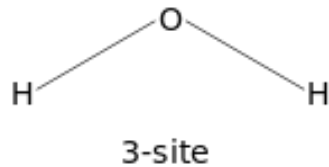
- Harmonic function  $E_{Cr} = \sum K_i (r_i - r_{io})^2$   
 $E_{C\phi} = \sum K_i (\phi_i - \phi_{io})^2$ 
  - Distance or dihedral constraints
  - On boundary condition
  - On specific positions (restraints)
- SHAKE algorithm  $s_{ij} = (r_{ij}^2 - d_{ij}^2) / d_{ij}^2 < \epsilon$ 
  - More rigid than harmonic function
  - Useful to increase time step of molecular dynamics from 0.5 to 1 or 2 fs

# Solvation

- Implicit condition: **continuum**
  - Empirical equations and functionals
  - Poisson-Boltzmann, Generalized Born, GBSA
- Explicit condition: **particles**
  - Add ions or other solvent molecules
  - Programs
    - Preconstructed water droplets or water boxes
    - VMD, packmol, etc
    - Online generator: <http://charmm-gui.org/>
  - Boundary condition
    - Periodic boundary condition
    - Harmonic boundary condition

# Explicit water models

- Simple water models



– TIP3P      TIP4P      TIP5P

- Other models

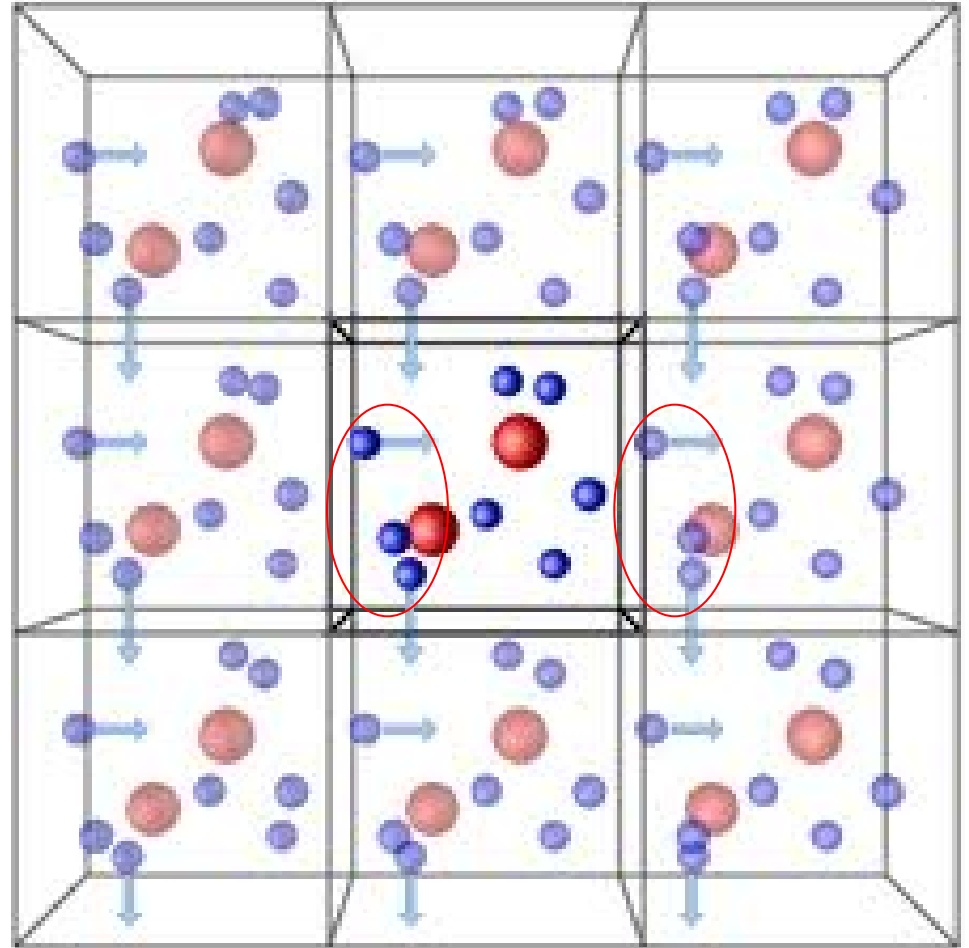
– SPC, SPC/E, Flexible SPC  
– BF, BNS, ST2, GG, MG



# Period boundary condition

- Images

- Around the unit cell (simulation box)
- Only particles in unit cell are recorded and propagated

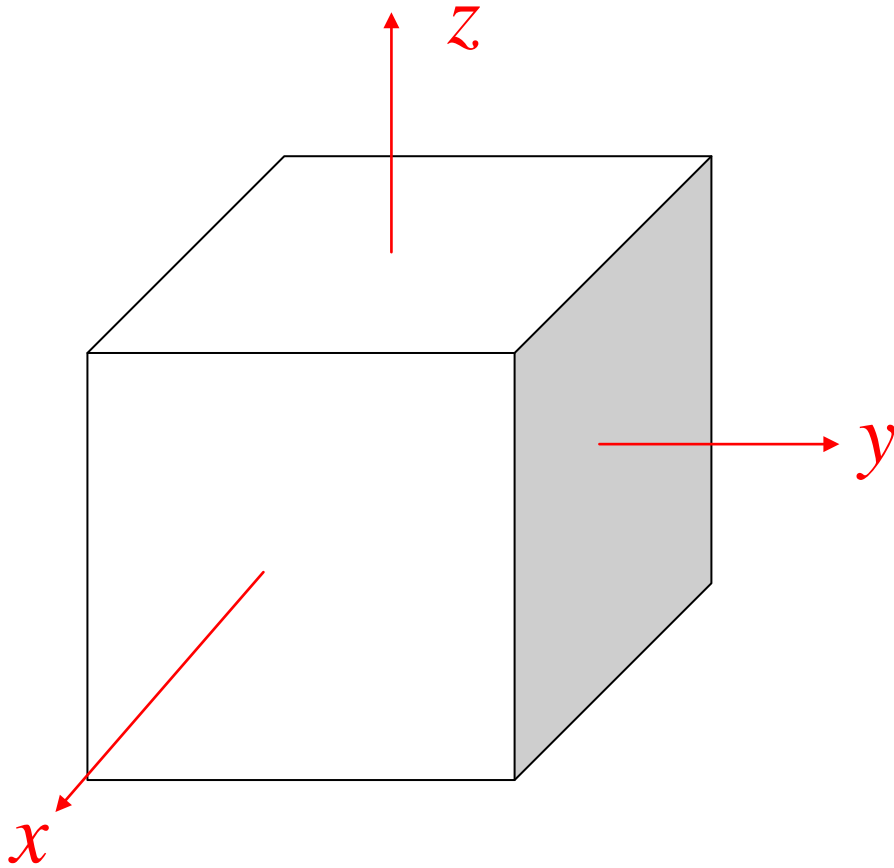


- Conserved

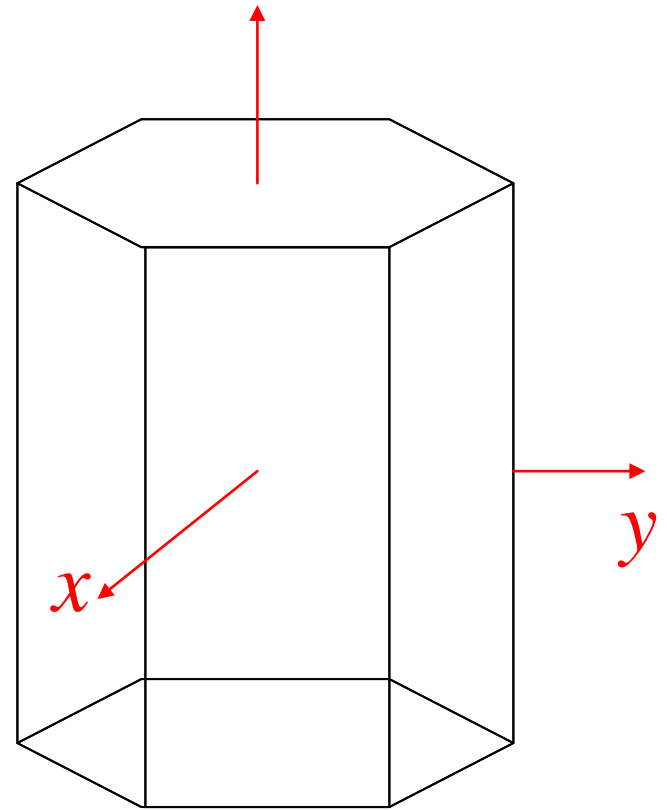
- The number of particles
- Linear momentum of the system

[isaacs.sourceforge.net/phys/psc.html](http://isaacs.sourceforge.net/phys/psc.html)

# Simulation box

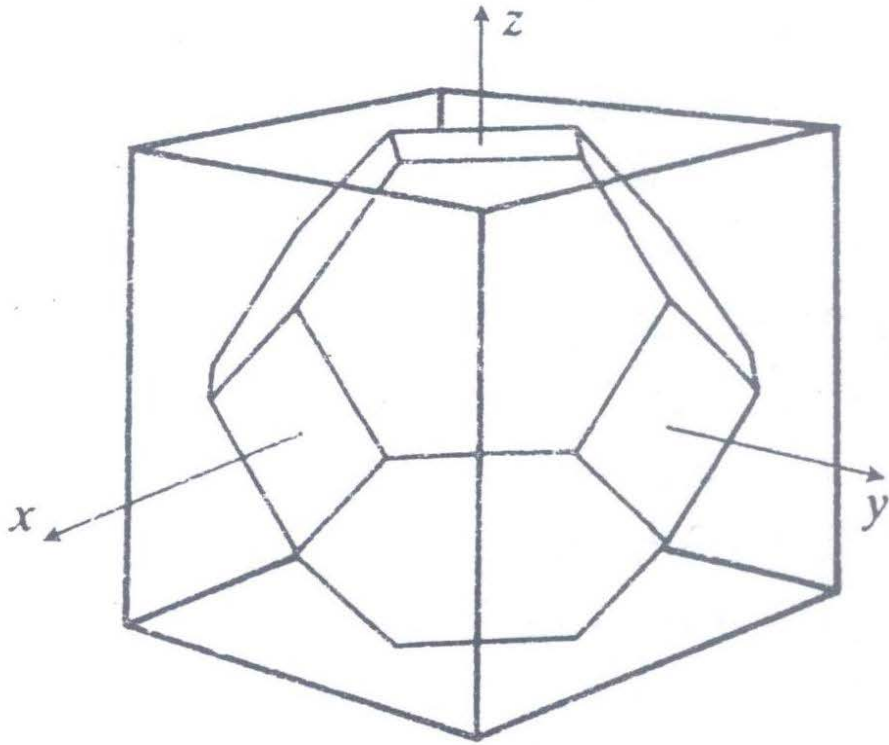


Cube or cuboid



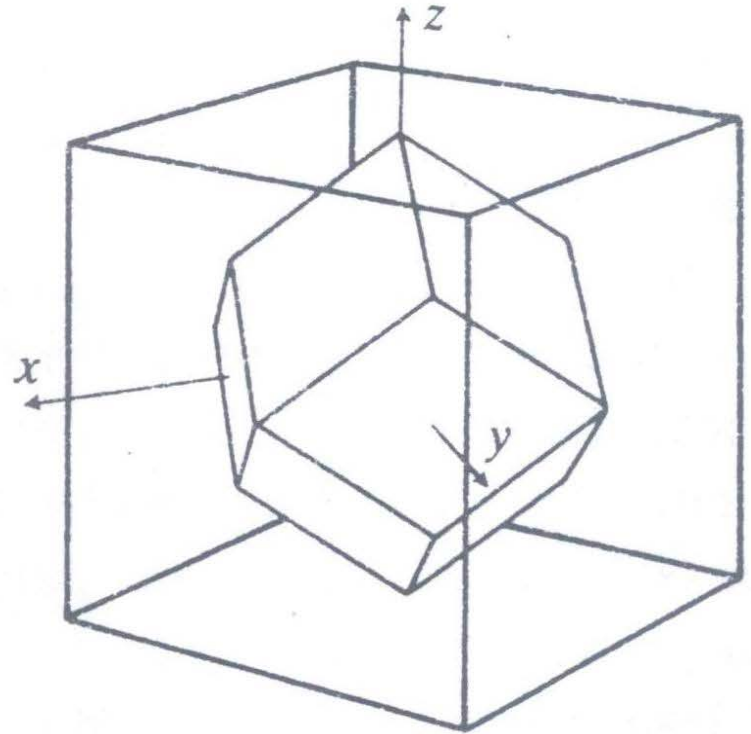
Hexagonal prism

# Simulation box



(a)

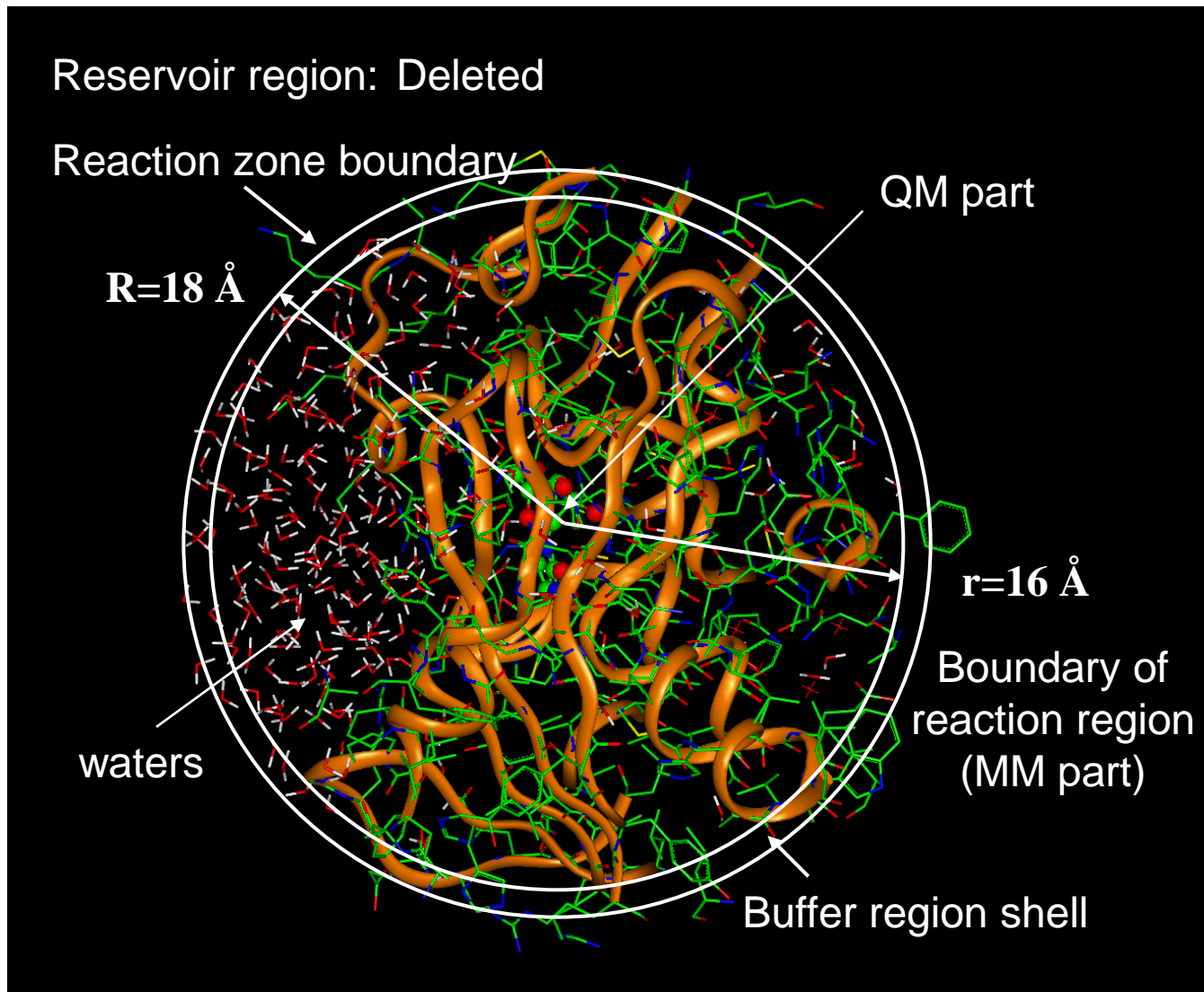
Truncated octahedron



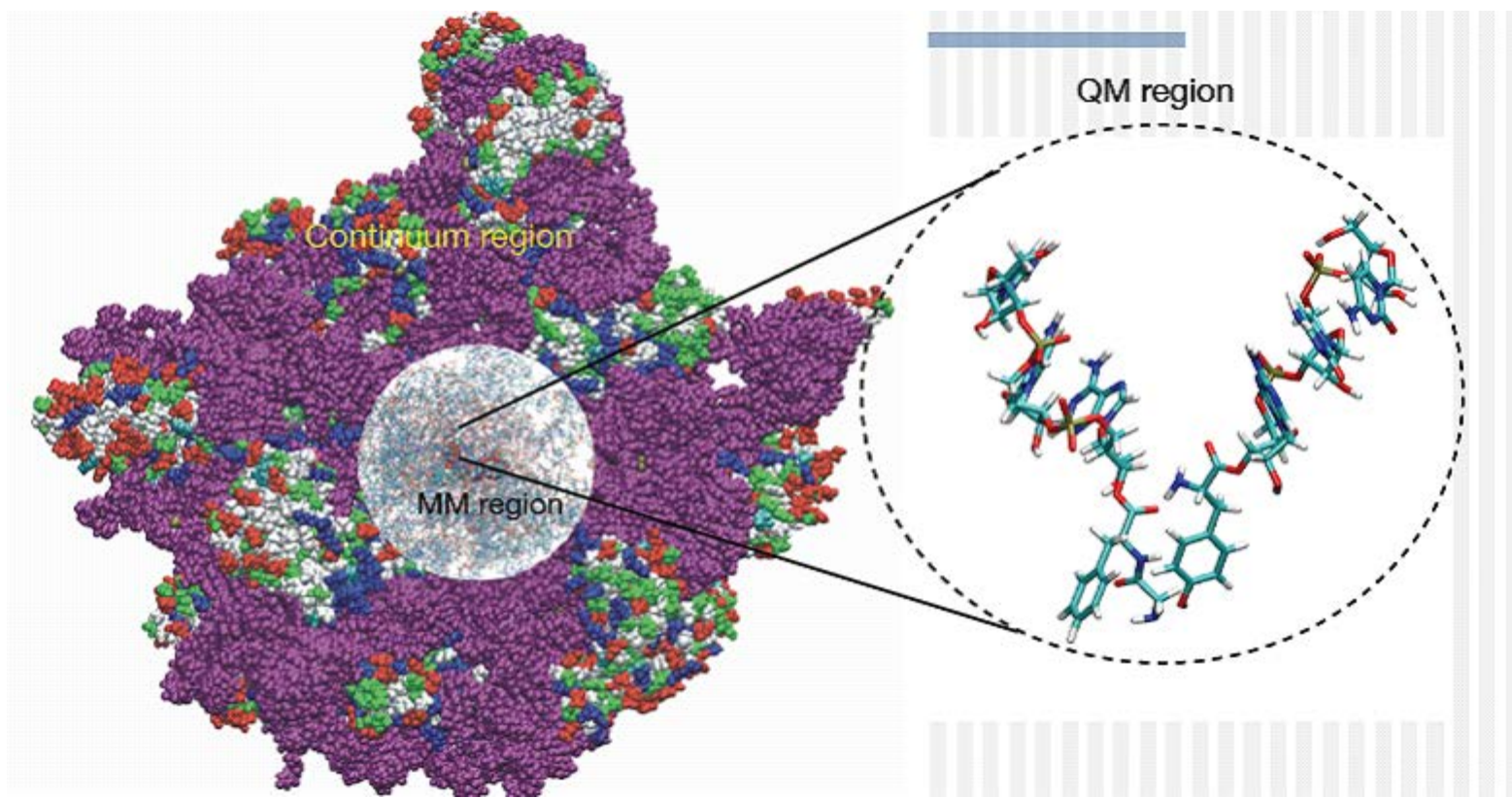
(b)

Rhombic dodecahedron

# Stochastic Boundary System



# GSBP (Generalized Solvent Boundary Potential)



Combine explicit and implicit solvent condition

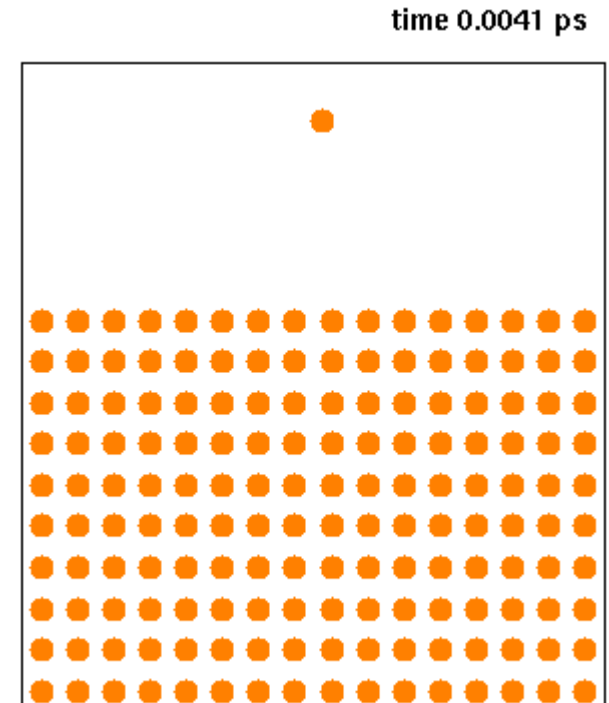
# Molecular dynamics (MD) simulation

- A computational simulation
  - Atoms and Molecules (model)
  - Within a given time (time step)
  - Physical movements
  - Based on interactions
  - By numerically solving the Newton's equations of motion

$$F = m \cdot a$$

(Force = mass • acceleration)

or  $-\nabla_i E = m_i d^2 r_i / dt^2$ ,  $i = 1, \dots, N$



# Flow of Molecular Dynamics Simulations

- Initial structure
  - Model construction and revision
  - Minimization
- Initialization on velocities
- Heating
- Equilibration
- Production run
- Analyze data from trajectories

# Initialization

- Initial velocities
  - a random number distributed uniformly in the  $[-0.5, 0.5]$
  - from a Maxwell-Boltzmann distribution
- Purpose
  - Generate the previous positions before  $t=0$



# Integrating the Equations of Motion in molecular dynamics

Algorithms for solving Newton's equations of motion in given time step ( $\Delta t$ )

## Verlet algorithm

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + (\Delta t)^2\mathbf{F}_i(t)/m_i$$

$$\mathbf{v}_i(t) = 1/2[\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)]/\Delta t$$

## Leap-frog algorithm

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}(t + 1/2\Delta t)\Delta t$$

$$\mathbf{v}(t + 1/2\Delta t) = \mathbf{v}(t - 1/2\Delta t) + \Delta t\mathbf{F}_i(t)/m_i$$

## Velocity Verlet algorithm

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \mathbf{v}(t)\Delta t + 1/2(\Delta t)^2\mathbf{F}_i(t)/m_i$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + 1/2\Delta t[\mathbf{F}_i(t)/m_i + \mathbf{F}_i(t + \Delta t)/m_i]$$

# Verlet algorithm

- To derive it, we start with a Taylor expansion of the coordinate of a particle, around time  $t$ ,

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(t)}{2m} \Delta t^2 + \cancel{\frac{\dot{\mathbf{f}}(t)}{6m}} \frac{\Delta t^3}{3!} + O(\Delta t^4)$$

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - \mathbf{v}(t)\Delta t + \frac{\mathbf{f}(t)}{2m} \Delta t^2 - \cancel{\frac{\dot{\mathbf{f}}(t)}{6m}} \frac{\Delta t^3}{3!} + O(\Delta t^4)$$

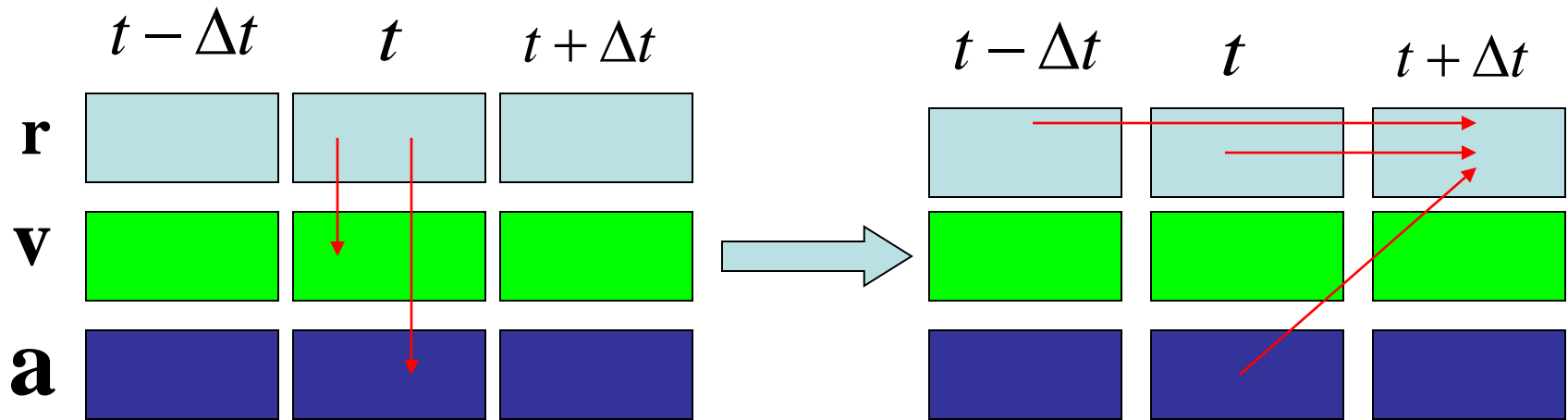
- Summing the two equations in the above slide, we obtain:

$$\mathbf{r}(t + \Delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{\mathbf{f}(t)}{m} \Delta t^2$$

- Similarly

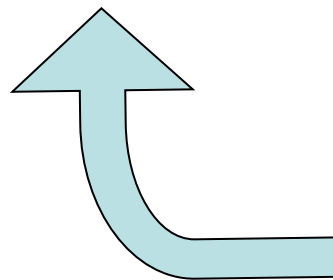
$$\mathbf{v}(t) = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} + O(\Delta t^2)$$

# Verlet algorithm calculation procedure



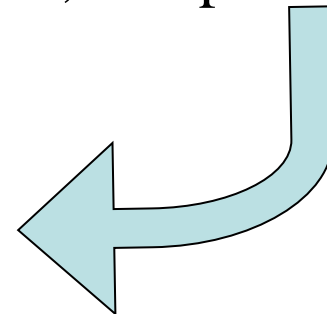
Calculate  $\mathbf{v}$  and  $\mathbf{a}$  at **present** from **present** position

Calculate position at next time step from present and last positions, and present force.



$$t = t + \Delta t$$

Update time



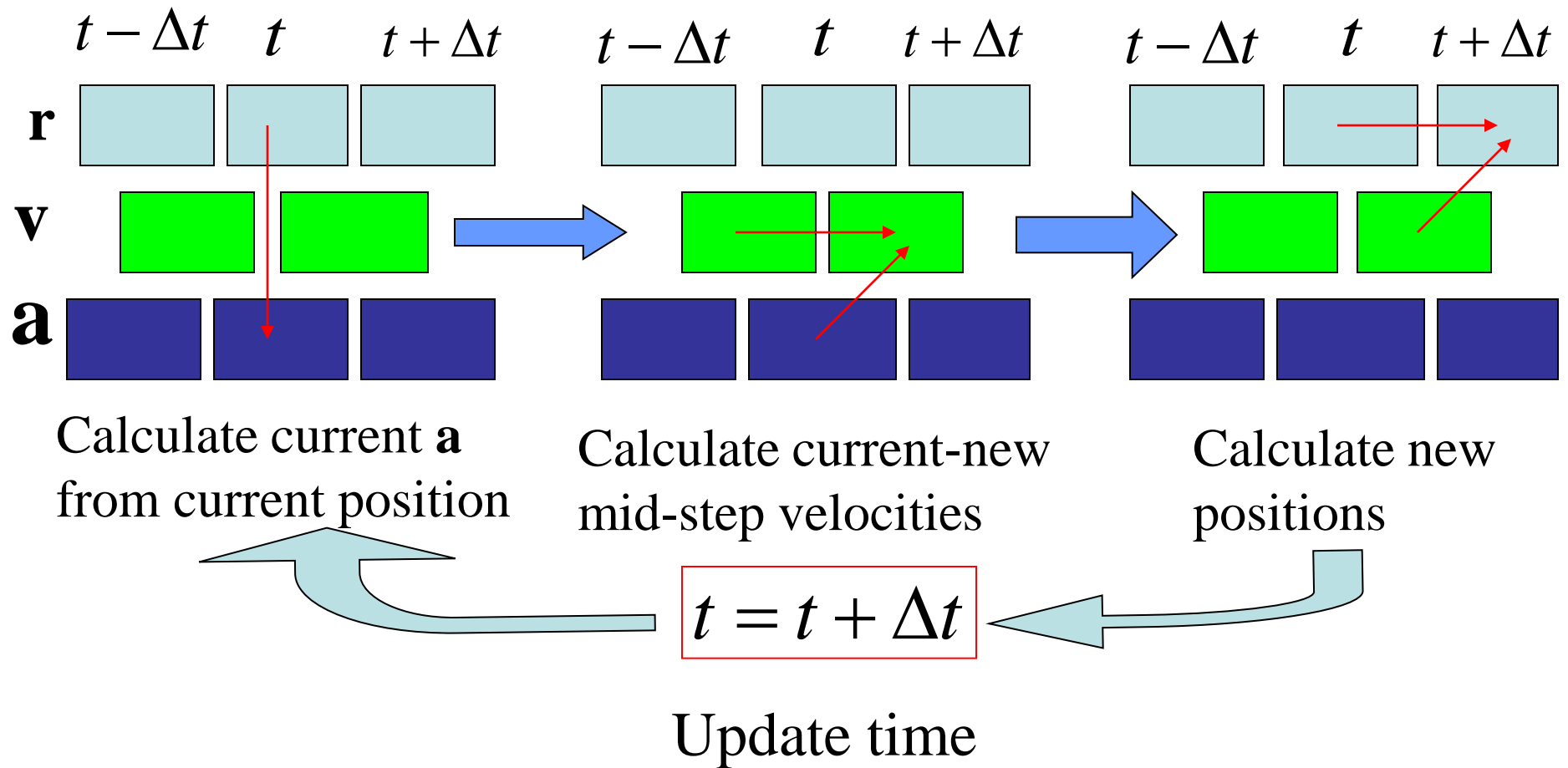
# Advantages of Verlet algorithm

- The Verlet method requires essentially  $9N$  words ( $3[RXOLD(I), RX(I) \& RXNEW(I)]$ ) of storage, make it very **compact**, and it is **simple to program**.
- The algorithm is exactly **reversible in time** and, given conservative forces, is guaranteed to conserve linear momentum. Have excellent **energy-conserving** properties even with long time steps.

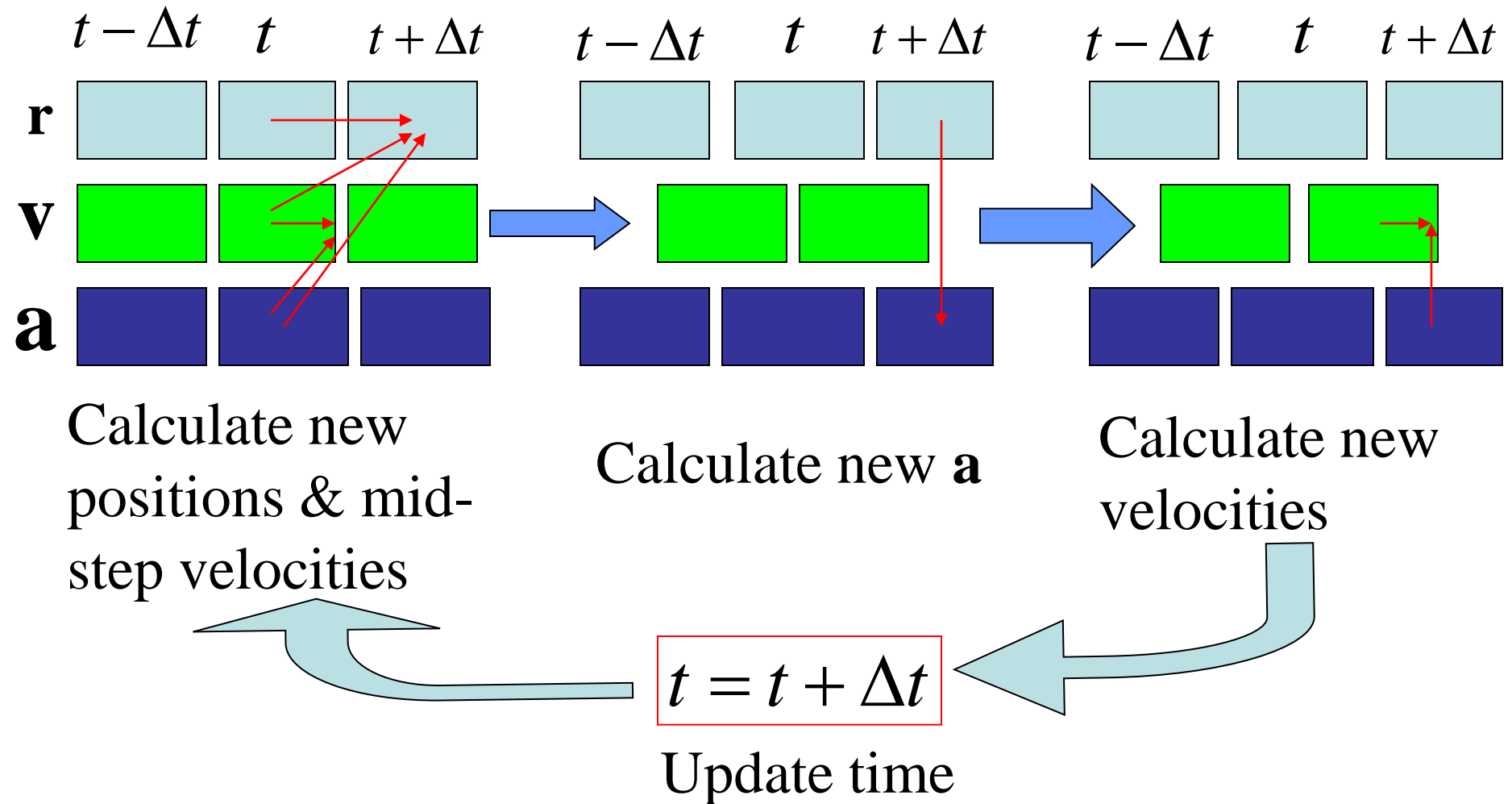
# Disadvantages of Verlet algorithm

- The velocities from Verlet algorithm are subject to errors of order  $\Delta t^2$ . More accurate estimates of  $\mathbf{v}(t)$  can be made, if more variables are stored, but this adds to the inconvenience, namely  $\mathbf{v}(t)$  can only be computed once  $\mathbf{r}(t + \Delta t)$  is known.
- The handling of velocities is rather awkward, and that the form of the algorithm may needlessly introduce some numerical imprecision.
  - **Modifications on basic Verlet scheme**
    - half-step leap-frog algorithm
    - velocity Verlet algorithm.

# Half-step leap-frog scheme---calculation procedure



# Velocity Verlet algorithm---calculation procedure



# A successful MD simulation algorithm

- It should be fast, and require little memory.
- Permit the use of a long time step  $\Delta t$ .
- Duplicate the classical trajectory as closely as possible.
- Satisfy the known conservation laws for energy and momentum, and be time-reversible
- Simple in form and easy to program.



# Other methods

- Beeman's method

- The method stores  $\mathbf{r}(t)$ ,  $\mathbf{v}(t)$ ,  $\mathbf{a}(t)$ , and  $\mathbf{a}(t-\Delta t)$ .

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \left[\frac{2}{3}\mathbf{a}(t) - \frac{1}{6}\mathbf{a}(t - \Delta t)\right]\Delta t^2$$

$$\mathbf{v}(t + \Delta t) = \frac{\mathbf{a}(t + \Delta t)\Delta t}{3} + \frac{5\mathbf{a}(t) - \mathbf{a}(t - \Delta t)}{6}\Delta t^2$$

- Predictor-corrector integration methods

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{a}(t)\Delta t^2}{2} + \frac{\mathbf{b}(t)\Delta t^3}{6} + \frac{\mathbf{c}(t)\Delta t^4}{24} + \mathbf{L}$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}(t) + \mathbf{a}(t)\Delta t + \frac{1}{2}\mathbf{b}(t)\Delta t^2 + \frac{1}{6}\mathbf{c}(t)\Delta t^3 + \mathbf{L}$$

$$\mathbf{a}(t + \Delta t) = \mathbf{a}(t) + \mathbf{b}(t)\Delta t + \frac{1}{2}\mathbf{c}(t)\Delta t^2 + \mathbf{L}$$

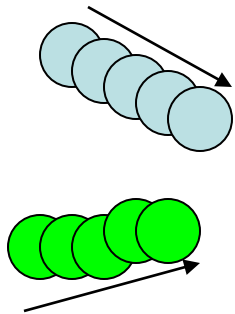
$$\mathbf{b}(t + \Delta t) = \mathbf{b}(t) + \mathbf{c}(t)\Delta t + \mathbf{L}$$

# $\Delta t$ , time step

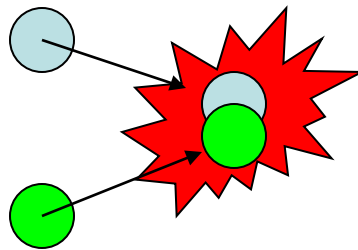
With an appropriate time step, phase space is covered efficiently and collisions occur smoothly.

- Too small and the trajectory will cover only a limit proportion of the phase space.

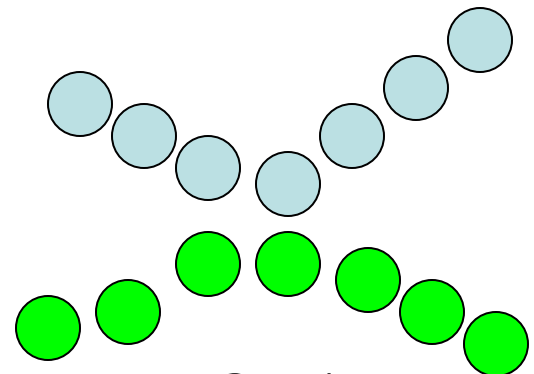
- Too large and instabilities may arise in the integration algorithm due to high overlaps between atoms.



Too small  $\Delta t$

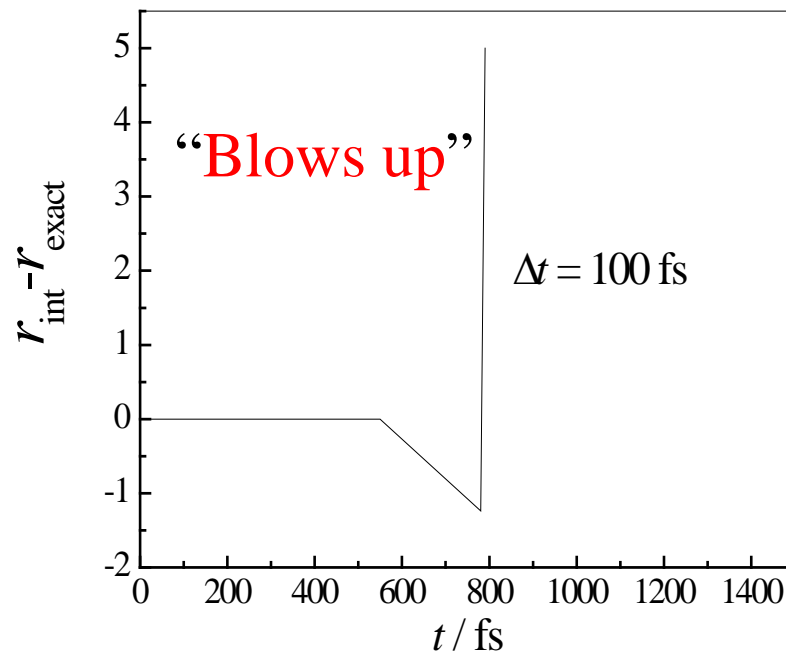
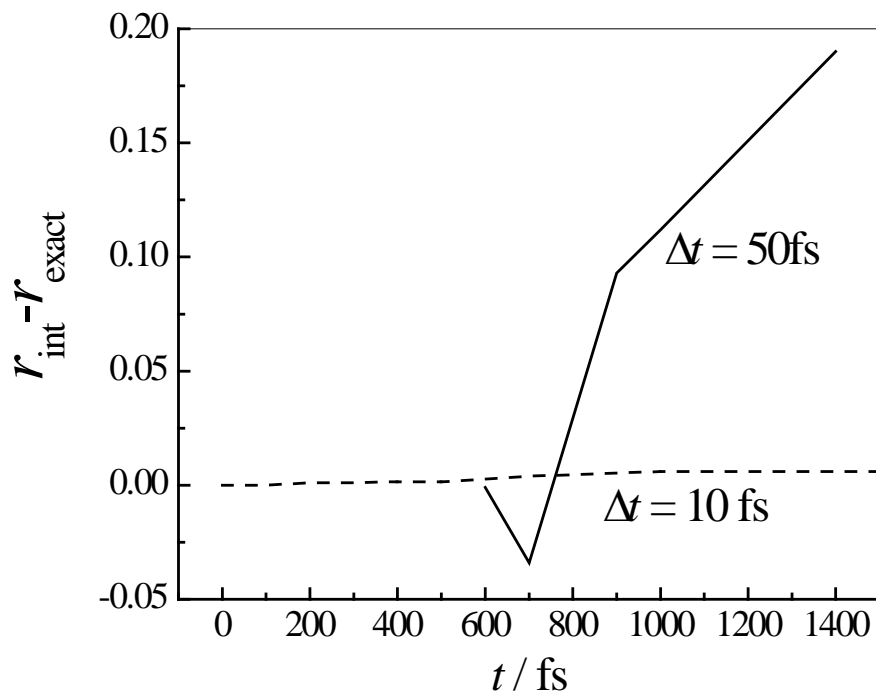


Too large  $\rightarrow$  instabilities



OK!

# Time step



Difference between exact and numerical trajectory for the approach of two argon atoms with different time steps. Ar atoms are moving towards each other along the x axis with initial velocities of 353m/s (corresponds to 300K).

# Choosing time step

- When simulating an atomic fluid the time step should be small compared to the mean time between collisions.
- When simulating flexible molecules a useful guide is that the time step should be approximately one tenth the time of the shortest period of motion.
  - A C-H bond vibrates with a repeat period of approximately 10 fs.
  - MD simulations on biological micromolecule, **0.5~2fs**

# Types of motion present in various systems with suggested time steps

System	Types of motion	Suggested $\Delta t$
Atoms	translation	$10^{-14}\text{s}$
Rigid molecules	Translation & rotation	$5 \times 10^{-15}\text{s}$
Flexible molecules with rigid bond	Translation, rotation & torsion	$2 \times 10^{-15}\text{s}$
Flexible molecules with flexible bond	Translation, rotation, torsion, vibration	$10^{-15}$ or $5 \times 10^{-16}\text{s}$

# Advanced method of MD simulations

- Free energy simulations
- Free energy perturbations (FEP)
- Metadynamics
- Steered Molecular Dynamics (SMD)
- Targeted Molecular Dynamics (TMD)
- Accelerated Molecular Dynamics (aMD)
- Collective Variable-based Calculations
- ○ ○ ○ ○ ○ ○

# Free Energy Simulations

- Potential of Mean Force (PMF)
  - The change of free energy as the function of **specific reaction coordinates**
- Umbrella Sampling method
- WHAM (Weighted Histogram Analysis Method) program

