

SMALL MOLECULES:

CHEMISTRY, MOLECULAR MECHANICS, CONFORMATIONS
3 in 1

VLAD KHOLODOVYCH
UNIVERSITY OF MEDICINE & DENTISTRY
OF NEW JERSEY (UMDNJ)

KHOLODVL@UMDNJ.EDU
HTTP://WWW2.UMDNJ.EDU/~KHOLODVL

What is life?

VITAL CYCLES

Water

Oxygen

Carbon

Nitrogen

<http://www.kidsgeo.com/geography-for-kids/>

BODY COMPOSITION

	Lean man 70 kg	Obese man 100 kg
Water	60%	47%
Protein	17%	13%
Fat	17%	35%
Remainder	6%	5%

ELEMENTAL COMPOSITION OF THE HUMAN BODY

By Weight*

Oxygen 65%
 Carbon 18%
 Hydrogen 10%
 Nitrogen 3%
 Calcium 1.5%
 Phosphorous 1.0%
 Potassium 0.35%
 Sulfur 0.25%
 Sodium 0.15%
 Chlorine 0.15%
 Magnesium 0.05%
 Iron 0.0004%
 Iodine 0.00004%

By Mole

Hydrogen 63%
 Oxygen 26%
 Carbon 9%
 Nitrogen 1.25%
 Calcium 0.25%
 Phosphorus 0.19%
 Potassium 0.06%
 Sulfur 0.06%
 Sodium 0.04%
 Chlorine 0.025%
 Magnesium 0.013%
 Iron 0.00004%
 Iodine 0.000002%

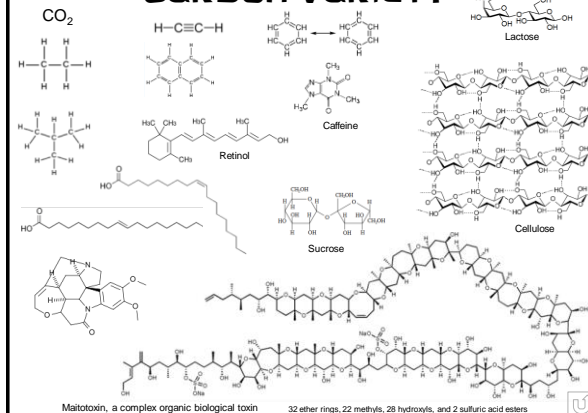
By Mole, Dehydrated

Hydrogen 52%
 Carbon 39%
 Nitrogen 5.6%
 Oxygen 1.6%
 Calcium 0.98%
 Phosphorus 0.77%
 Potassium 0.24%
 Sulfur 0.20%
 Sodium 0.15%
 Chlorine 0.10%
 Magnesium 0.06%
 Iron 0.0002%
 Iodine 0.000008%

*U.S. Bureau of Chemistry and Soils

<http://www.madsci.org/posts/archives/2000-06/962225341.Bc.r.html>

CARBON VARIETY



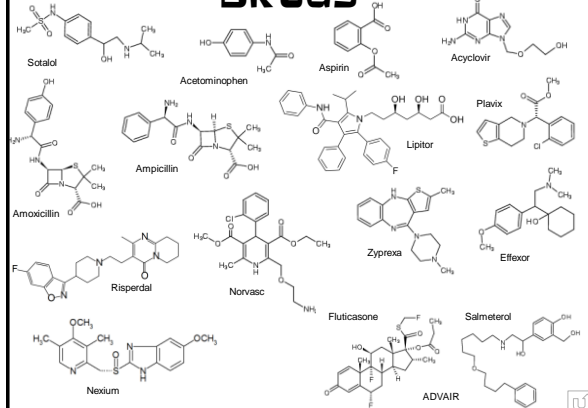
SMALL MOLECULES

- Organic compounds (C,H,O,N,P,S)
- Molecular weight below 1000
- Non proteins
- Non polymers
- Non carbohydrates

WHAT ARE THEY?



DRUGS



SMALL MOLECULES

- Organic compounds (C,H,O,N,P,S)
- Molecular weight below 1000
- Non proteins
- Non polymers
- Non carbohydrates

WHAT ARE THEY?

DRUGS

VITAMINS

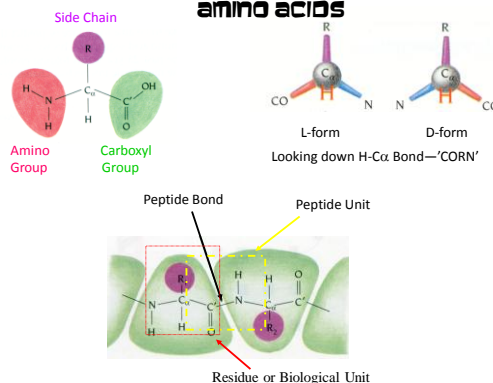
ODORS

LIPIDS

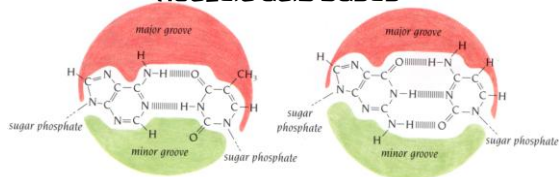
BUILDING
BLOCKS

FOOD ADDITIVES

BUILDING BLOCKS: amino acids



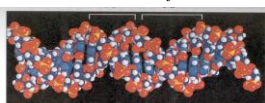
BUILDING BLOCKS: NUCLEIC ACID BASES



ADENINE : THYMINE

GUANINE : CYTOSINE

Minor Groove Major Groove



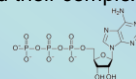
SMALL MOLECULES

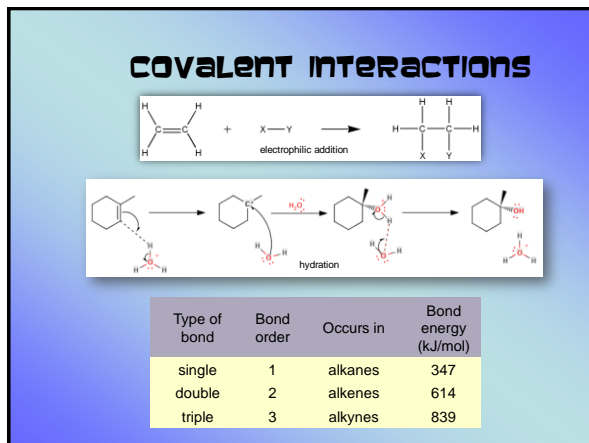
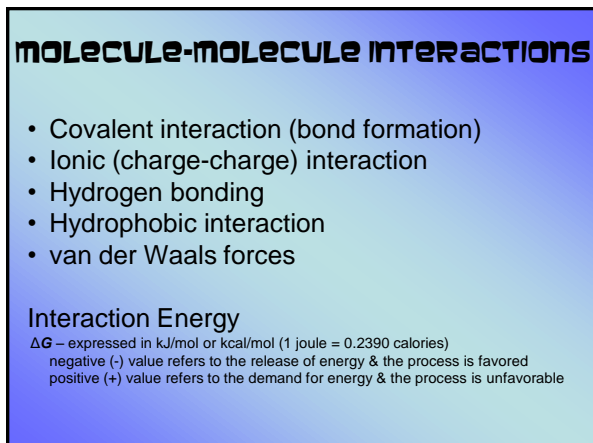
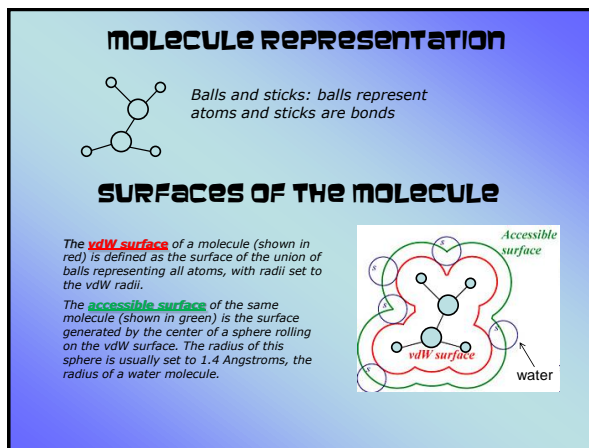
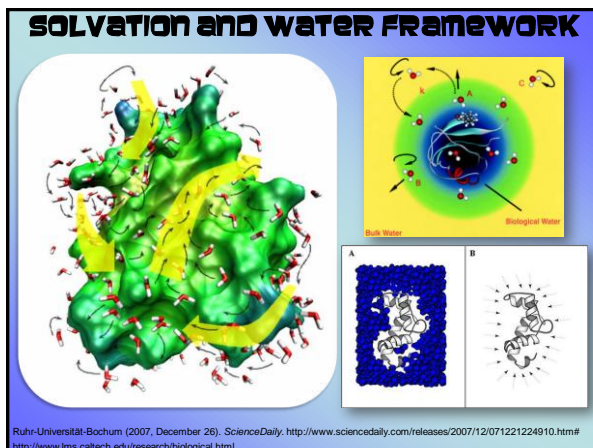
1. Drugs
2. Odors/Fragrances
3. Food additives – taste enhancers, sugar substitutes
4. Vitamins - organic and metalloorganic (complexes with metals)
5. Parts and building blocks of complex organic systems
Amino acids, nucleotide bases
6. Lipids – fatty acids
7. Repeat units of polymers – monomers
8. Miscellaneous – catalysts (metals and their complexes),
live system catalyst - ATP

... ANYTHING ELSE ?

SOLVENTS!

WATER, DMSO, ACETONE, ETC.

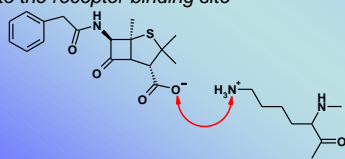




IONIC INTERACTIONS

Strong electrostatic interactions (5-10 kcal/mol)
 Important attractive forces
 Responsible for relative orientation of the molecules to each other especially at the beginning of interaction

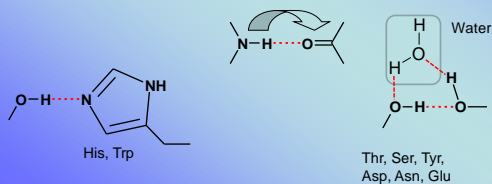
Example:
 Driving force for attraction, orientation and initial binding of the ligand to the receptor binding site



HYDROGEN BONDING INTERACTIONS

(non-ionic/neutral)

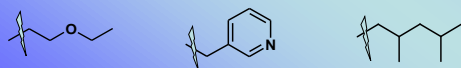
- Non-covalent bonds
- Hydrogen bond strength is distance dependent
- Solvation/desolvation energy balance
- May range from 0.5 – 7 kcal/mol, depending on the binding environment
- Distance between the heteroatoms is about 2.2 – 3.5 Å



HYDROPHOBIC/LIPOPHILIC INTERACTIONS

One of the most important non-covalent interactions in aqueous environment (under physiological conditions)
 Come into play at closer distances than ionic interactions
 Play important role in many properties: poor solubility, effecting drug formulation, distribution, half-life, etc.

Examples:
 Lipophilicity is important for transport across membranes
 Good matching of the hydrophobic regions of a ligand and a receptor results in the lowering of the energy – favorable



VAN DER WAALS INTERACTIONS

- Interactions at a close range
- Dispersion forces attract atoms at close distances due to atomic dipoles
- Repulsive forces play a role at closer distances than van der Waals spheres of both molecules
- Interactions are much weaker (~0.5-1 kcal/mol) than other electrostatic interactions

Example:
 Close contacts (attractive forces) over a large surface area, i.e. at the interface of ligand and binding site, may significantly contribute to the total binding energy

A **FORCE FIELD** is nothing but a set of functions and constants used to calculate the potential energy of the molecule. In the Force Field, the constants k in the formula are related to bonds and angles.

$$E = \sum_{ij} k_{ij} x_i x_j + \sum_{ijk} k_{ijk} x_i x_j x_k + \dots$$

PARAMETERS define the reference points and force constants, allowing for the calculation of the different energies due to deviations from 'natural' values and/or attractive/repulsive interactions between atoms.

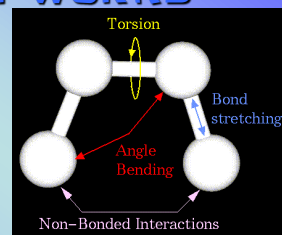
OPTIMIZERS/MINIMIZERS are algorithms to calculate new positions of the atoms after changes have been applied. Different methods and techniques were suggested to overcome the local/global minima problem (Steepest Descents / Conjugate Gradients / Powell / Newton-Raphson / BFGS / Line searches)

HOW IT WORKS

Atoms are spheres

Bonds are springs that can stretch, bend or twist

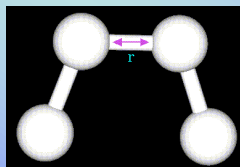
Non-bonded interactions include van der Waals attractions and electrostatic attractions / repulsions



$$\text{Energy} = \text{Stretching Energy} + \text{Bending Energy} + \text{Torsion Energy} + \text{Non-Bonded Interaction Energy}$$

STRETCHING ENERGY

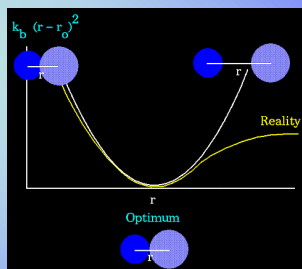
$$E = \sum_{\text{bonds}} k_b (r - r_0)^2$$



estimates the energy associated with vibration around the equilibrium bond length.

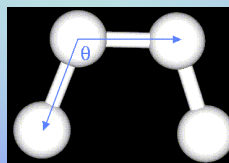
k_b and r_0 are unique for each bonded atom pair based on their types (C-C, C-O,....)

based on Hooke's law, k_b controls the stiffness of the bond spring; r_0 defines its equilibrium length



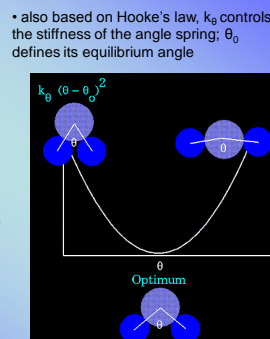
BENDING ENERGY

$$E = \sum_{\text{angles}} k_\theta (\theta - \theta_0)^2$$



estimates the energy associated with vibration about the equilibrium bond angle

The wider the shape of parabola (k is smaller), the more energy is required to deform an angle (or bond) from its equilibrium value

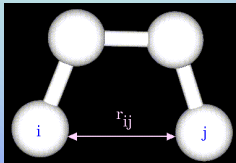


also based on Hooke's law, k_θ controls the stiffness of the angle spring; θ_0 defines its equilibrium angle

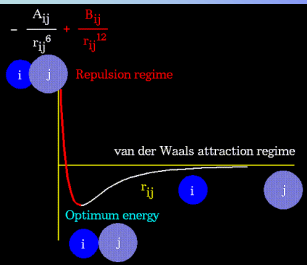
NON-BONDED ENERGY

$$E = \sum_i \sum_j \frac{-A_{ij}}{r_{ij}^6} + \frac{B_{ij}}{r_{ij}^{12}} + \sum_i \sum_j \frac{q_i q_j}{r_{ij}}$$

van der Waals term Electrostatic term



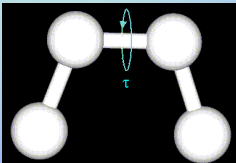
accounts for repulsion, van der Waals attraction, and electrostatic interactions



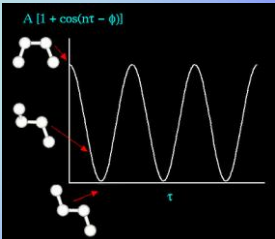
van der Waals attraction: at short range, and rapidly dies off at a few Angstroms. **Repulsion:** the distance between interacting atoms is slightly less than the sum of their contact radii. These effects are often modeled using a 6-12 equation.

The **electrostatic contribution** is modeled using a Coulombic potential.

TORSION ENERGY

$$E = \sum \text{torsions} [1 + \cos(n\tau - \phi)]$$


• is modeled by a simple periodic function and primarily used to correct the remaining energy terms rather than to represent a physical process.



• represents the amount of energy that must be added to or subtracted from the Stretching Energy + Bending Energy + Non-Bonded Interaction Energy terms to make the total energy agree with an experiment or rigorous quantum mechanical calculations.

For example, a model of dihedral angle of ethane may be used as a model for any H-C-C-H bond.

- FF is empirical methods, so no “correct” form of force field can exist.
- FF should be considered as single entity. Usually separate term is not transferable, though it is possible.
- Molecular Mechanics energies have no meanings as absolute quantities. Only difference in energy between two or more conformations have meaning.

MOLECULAR MECHANICS TUTORIAL

CONFORMATIONAL ANALYSIS

Derek Barton 1950: Analysis of substituted cyclohexanes



Biological and chemical properties depend on 3D structures (conformers) that molecule can adopt.

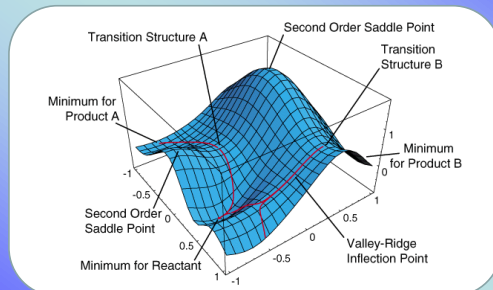
Conformational analysis is the study of the conformations of a molecule and their influence on its properties.

The main goal of CA is to locate all preferred conformations that determine biological or chemical behavior of molecule.

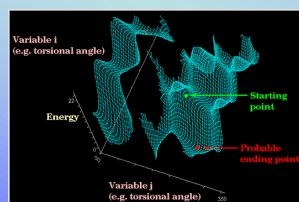
On practice it means that we have to find points with the minimal values on the molecular energy surface.

MOLECULAR ENERGY PROFILE

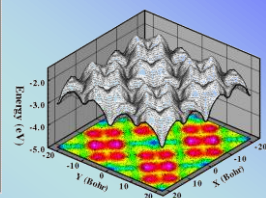
Energy is a function of the degrees of freedom in a molecule (i.e. bonds, angles, and dihedrals).



DIRECTION TO GLOBAL MINIMUM



Single global minimum



Multiple almost equal minima

TECHNIQUES FOR CONFORMATIONAL SEARCH

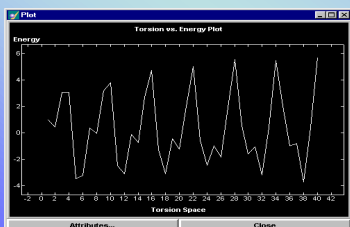
- Systematic Search
- Stochastic Search
(random perturbation methods)
- Distance geometry
- Simulated Annealing

SYSTEMATIC SEARCH

Systematic Conformational Search generates molecular conformations by systematically rotating bonds in a molecule by discrete increments.

Purpose: generating a collection of reasonable molecular conformations, which may or may not be at local minima.

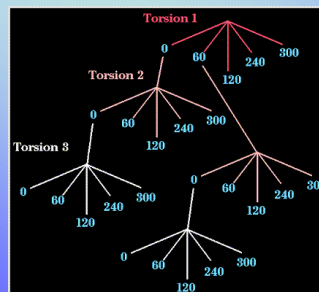
Penalty: van der Waals energy exceeds a threshold (usually, 10-15 kcal/mol).



"Torsion driving"

Energy profile for single dihedral angle.

COMBINATORIAL SEARCH TREE



Rotation around: 3 bonds

Range: 0 to 360°

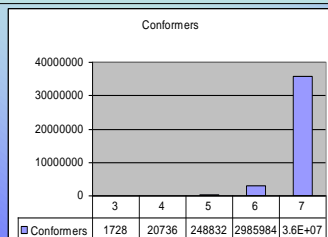
Step: 60°

Number of conformers
360/60*360/60*360/60=216

COMBINATORIAL EXPLOSION

$$\text{Total number of conformations} = \prod_i^N \frac{360}{\theta_i}$$

θ_i , dihedral increment for bond i



3 rot. bonds
Step: 30°
Confs: 1728
Min rate: 1
conf/sec
Time: 29min

5 rot. bonds
Step: 30°
Confs: 248832
Min rate:
1conf/sec
Time: 69 hours

Time for 7 rot.
bonds: 415 days

SYSTEMATIC SEARCH

Advantage:

Explore all conformational space systematically. All possible minimal conformers can be found.

Lack:

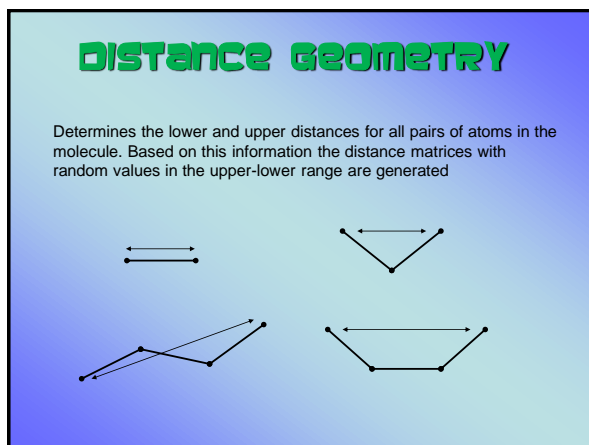
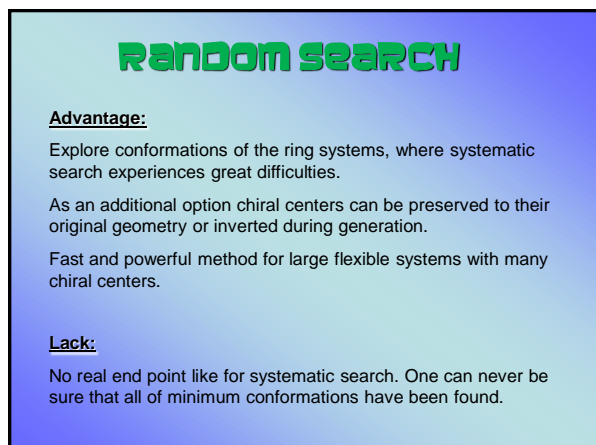
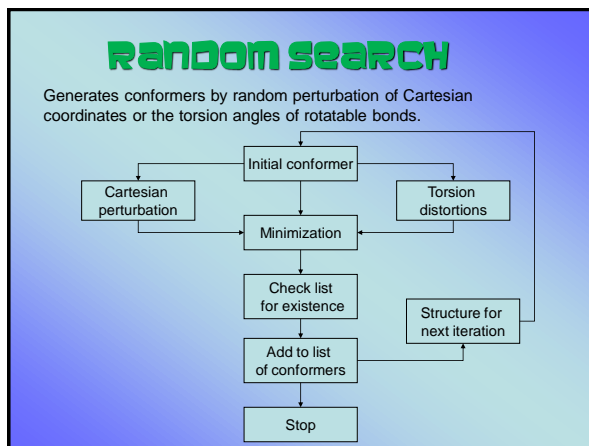
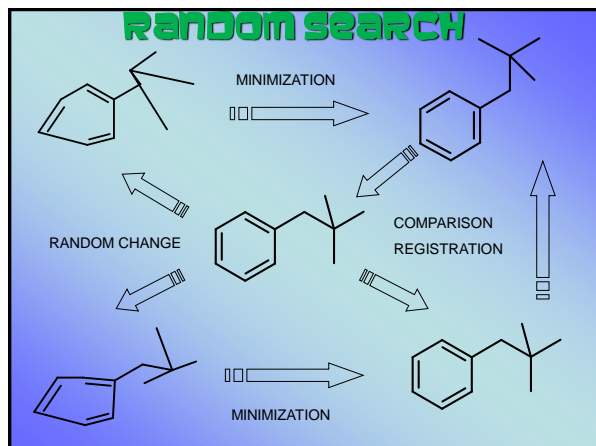
Time consuming: combinatorial explosion.

Can not be used for a large system and has a great limitation for the ring systems.

Partial solution:

Restricted systematic search (energy cutoff)

Fragment based approach rather than atom building blocks



DISTANCE GEOMETRY

Advantage:

Very accurate in the systems for which experimental data are available (NMR, IR, X-Ray crystallography)

Useful for refining the structure of proteins and nucleic acids where the amount of data is huge for manual processing.

Can generate several conformations that are consistent with experimental data and so provides an additional information of possible conformations

Lack:

Requires experimental data

SIMULATED ANNEALING

Simulated annealing is a global optimization technique based on the Monte Carlo method.

Generate small random changes in the current state and then accepting or rejecting each new state according to the Metropolis criterion:

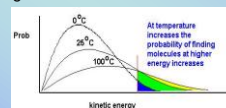
- Moves that decrease the energy of the system are always accepted
- Moves that increase the energy of the system are accepted according to probability p :

$$p = e^{\left(\frac{-\Delta E}{kT}\right)} \quad \text{where } \Delta E = E_1 - E_0$$

and E_1 is an energy of the new state, E_0 is an energy of the current state T is the "temperature" of the simulation, and k is Boltzmann's constant

- Moves accepted if probability value is larger than random chosen number between 0 and 1.

At high temperatures, many states will be accepted, while at low temperatures, the majority of these probabilistic moves will be rejected



CONFORMATIONAL SEARCH TUTORIAL

WHERE TO GET THEM FROM?

Crystallographic data - CDC Cambridge, UK

Vendors – Sigma-Aldrich, etc

Assay specific libraries – kinases, TimTec

Chemical building blocks – Maybridge

Structural conformers – MOE, about 600k molecules with over 45million precalculated conformers

Internet

e-Molecules <http://www.emolecules.com>

ChemSpider <http://www.chemspider.com>

PubChem <http://pubchem.ncbi.nlm.nih.gov>