

Course 1: Introduction

Molecular structure and properties calculations

Computational chemistry
Molecular modeling

*In theory, there is no difference
between theory and practice.
In practice, there is.*

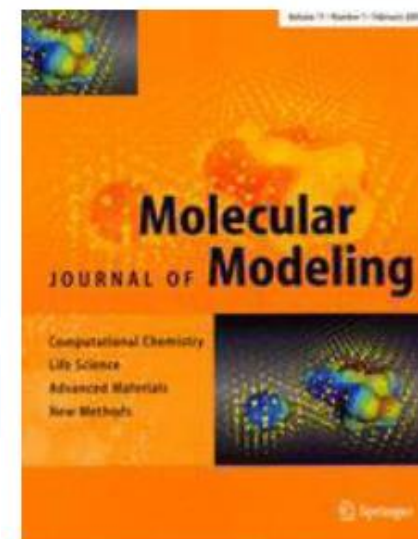
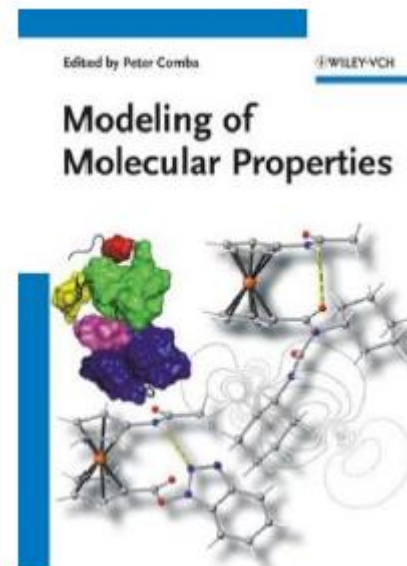
Yogi Berra
(Gaussian "wisdom database")

Computational chemistry (from wikipedia)

- a branch of chemistry that uses computers to assist in solving chemical problems.
- uses the results of theoretical chemistry, incorporated into efficient computer programs, to calculate the structures and properties of molecules and solids.
- it can (in some cases) predict hitherto unobserved chemical phenomena.
- widely used in the design of new drugs and materials.

Molecular modeling

- applied theoretical approaches and computational techniques to model structures and properties of molecular compounds and materials in order to interpret and/or their properties.
- used for studying molecular systems ranging from small molecular units to large biological molecules and material assemblies.



Warren J. Hehre, A Guide to Molecular Mechanics and Quantum Chemical Calculations, Wavefunction, Inc. 18401 Von Karman Ave., Suite 370 Irvine, CA 92612 (<http://downloads.wavefun.com/FAQ/AGuidetoMM.pdf>)

Major goals of computational chemistry

- to create efficient **mathematical approximations** and **computer programs**
 - o to calculate the **properties of molecules**
 - o to apply these programs to concrete physico-chemical systems

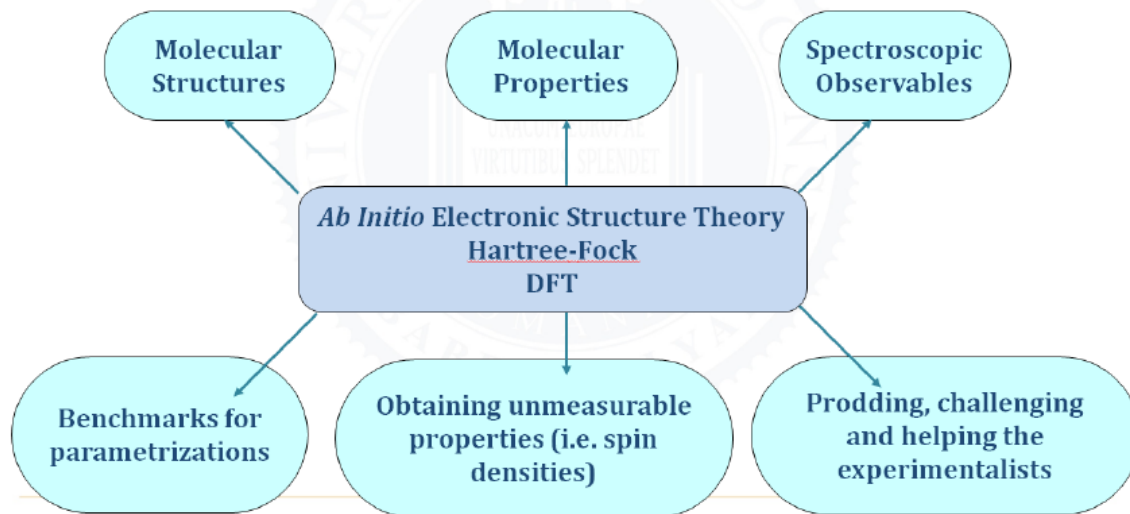
Properties:

- geometrical structure
 - geometrical parameters for ground or excited states or for different phases (gas, liquid, solid)
 - conformers, tautomers,
- absolute and relative energies
- interaction energies
- electronic charge distributions
- dipoles and higher multipole moments
- vibrational spectra (frequencies and intensities)
- UV-Vis, ESR, and NMR spectra
- MO energies and shapes (UPS and XPS) spectra
- etc.



“Computational spectroscopy”

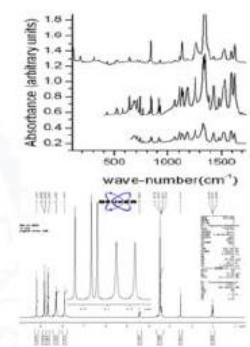
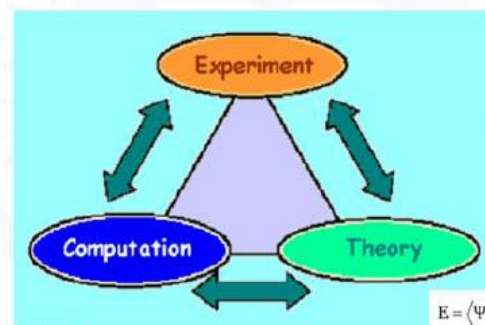
Why computing molecules?



```

1  read(1,10)
2  write(10,10)
3  read(10,10)
4  write(10,10)
5  read(10,10)
6  write(10,10)
7  read(10,10)
8  write(10,10)
9  read(10,10)
10 write(10,10)
11 read(10,10)
12 write(10,10)
13 read(10,10)
14 write(10,10)
15 read(10,10)
16 write(10,10)
17 read(10,10)
18 write(10,10)
19 read(10,10)
20 write(10,10)
21 read(10,10)
22 write(10,10)
23 read(10,10)
24 write(10,10)
25 read(10,10)
26 write(10,10)
27 read(10,10)
28 write(10,10)
29 read(10,10)
30 write(10,10)
31 read(10,10)
32 write(10,10)
33 read(10,10)
34 write(10,10)
35 read(10,10)
36 write(10,10)
37 read(10,10)
38 write(10,10)
39 read(10,10)
40 write(10,10)
41 read(10,10)
42 write(10,10)
43 read(10,10)
44 write(10,10)
45 read(10,10)
46 write(10,10)
47 read(10,10)
48 write(10,10)
49 read(10,10)
50 write(10,10)
51 read(10,10)
52 write(10,10)
53 read(10,10)
54 write(10,10)
55 read(10,10)
56 write(10,10)
57 read(10,10)
58 write(10,10)
59 read(10,10)
60 write(10,10)
61 read(10,10)
62 write(10,10)
63 read(10,10)
64 write(10,10)
65 read(10,10)
66 write(10,10)
67 read(10,10)
68 write(10,10)
69 read(10,10)
70 write(10,10)
71 read(10,10)
72 write(10,10)
73 read(10,10)
74 write(10,10)
75 read(10,10)
76 write(10,10)
77 read(10,10)
78 write(10,10)
79 read(10,10)
80 write(10,10)
81 read(10,10)
82 write(10,10)
83 read(10,10)
84 write(10,10)
85 read(10,10)
86 write(10,10)
87 read(10,10)
88 write(10,10)
89 read(10,10)
90 write(10,10)
91 read(10,10)
92 write(10,10)
93 read(10,10)
94 write(10,10)
95 read(10,10)
96 write(10,10)
97 read(10,10)
98 write(10,10)
99 read(10,10)
100 write(10,10)

```



$$E = \langle \Psi^0 | H | \Psi^0 \rangle = \sum_{i=1}^N H_i + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (J_{ij} - K_{ij})$$

$$H_i = \int \Phi_i^*(x) \left[-\frac{1}{2} \nabla^2 + v(x) \right] \Phi_i(x) dx$$

$$J_{ij} = \iint \Phi_i^*(x_1) \Phi_j^*(x_1) \frac{1}{r_{12}} \Phi_j(x_2) \Phi_i(x_2) dx_1 dx_2$$

$$K_{ij} = \iint \Phi_i^*(x_1) \Phi_j(x_1) \frac{1}{r_{12}} \Phi_i^*(x_2) \Phi_j(x_2) dx_1 dx_2$$

Contributors to the field

Walter Heinrich Heitler (2 January 1904 – 15 November 1981)

- German **physicist** who made contributions to quantum electrodynamics and quantum field theory.
- **brought chemistry under quantum mechanics** through his theory of valence bonding.

Fritz Wolfgang London (March 7, 1900 – March 30, 1954)

- German-born American **physicist** and professor at Duke University.
- fundamental **contributions to the theories of chemical bonding and of intermolecular forces** (*London dispersion forces*)

Walter Kohn (born March 9, 1923)

- Austrian-born American **theoretical physicist**.
- was awarded, with John Pople, the Nobel Prize in chemistry in 1998.
- played the **leading role in the development of DFT**.

Sir John Anthony Pople (31 October 1925 – 15 March 2004)

- was awarded the Nobel Prize in Chemistry with Walter Kohn in the year 1998.
- doctorate degree in mathematics in 1951.
- Pople considered himself more of a mathematician than a chemist



Sir John Edward Lennard-Jones - a **mathematician** (professor of theoretical physics)

- may be regarded as the **initiator of modern computational chemistry**
- developed the **LCAO** method

John C. Slater – **physicist**

Robert Sanderson Mulliken – **physicist and chemist**

Linus Carl Pauling - **chemist**

- one of the founders of quantum chemistry

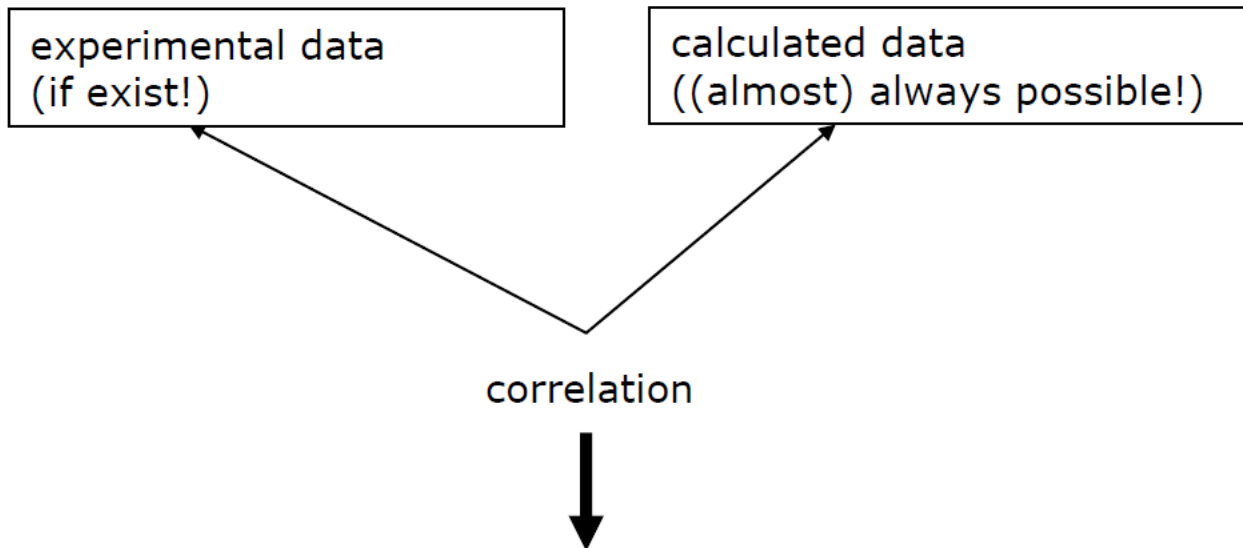
Samuel Francis Boys - **theoretical chemist**

Johann Carl Friedrich Gauss - **mathematician**

STRUCTURE ↔ PROPERTIES



HOW?



Computational chemistry is one of the most rapidly advancing and exciting fields in the natural sciences today.

"Many experimental chemists use various kinds of spectroscopy in their research even though they are not spectroscopists. In a similar manner, more and more scientists are applying computational techniques as another weapon in their arsenal"

Delano P. Chong in *Recent Advances in Density Functional Methods, Part I*, World Scientific, 1995

What shall you learn?

- the theory behind "molecular modelling"
- to use some molecular visualization packages
- to use program packages designed for molecular electronic structure theory
- to do calculations at different levels of theory and to interpret the results
- to make correlations between the experimental and theoretical data

Contents of the course

Theory

- Hartree-Fock Theory
- Basis sets
- Electron Correlation Methods
- Basis set superposition error
- Density Functional Theory

Practical works

- ❖ Potential energy surfaces
- ❖ Geometry optimizations
- ❖ Conformational searches
- ❖ Calculation of vibrational spectra
- ❖ Calculation of NMR and ESR spectra
- ❖ Calculation of UV-VIS spectra
- ❖ Weak intermolecular interactions

Can we do research?

- pure theoretical studies
- coupled experimental and theoretical investigation on the structure and properties of molecular systems

Where can we publish the results?

- Journal of Molecular Structure
- Computational and Theoretical Chemistry
- Journal of Molecular Spectroscopy
- Chemical Physics
- Chemical Physics Letters
- Journal of Molecular Modelling
- International Journal of Quantum Chemistry
- Journal of Computational Chemistry
- Journal of Chemical Physics A
- The Journal of Chemical Physics
- Molecular Physics
- Chemical Reviews
- Theoretical Chemistry Accounts
- Physical Chemistry Chemical Physics
- Journal of Chemical Theory and Computation
- ... and many others

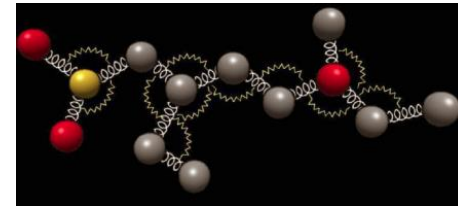
- [20.02.20 PhD scholarships, University of Queensland: Prediction of new electrolytes for improved electrical energy storage](#)
- [20.02.19 Postdoc Position in Molecular Dynamics Simulations ABC Transporter](#)
- [20.02.11 Postdoc Positions in Ab Initio Simulation](#)
- [20.02.11 Open Postdoc/PhD positions at the University of Delaware, USA](#)
- [20.02.07 Multiple PhD and Postdoctoral positions in Computational Biophysics of Protein-Membrane Interactions, Fribourg, Switzerland](#)
- [20.02.05 Three Early Stage Researchers \(ESR\) positions, Maria Skłodowska-Curie Network, Glasgow, Zurich, or Copenhagen](#)
- [20.02.04 Senior Consultant Information Architecture \(m/f/d\)](#)
- [20.02.03 Postdoc and PhD Positions in Theoretical and Computational Membrane Biophysics, Erlangen, Germany](#)
- [20.01.30 Computational Chemistry/Machine Learning/AI Scientist](#)
- [20.01.29 Quantum Chemist, Quantum Computing Company, Boston, MA](#)
- [20.01.28 Postdoctoral Fellow, Computational Antibody Engineering \(Montreal, Canada\)](#)
- [20.01.27 Computational Chemist, New Equilibrium Biosciences, Boston, MA](#)
- [20.01.27 Cheminformatics Scientist \(209420W-01\), Corteva Agriscience, Indianapolis, IN](#)
- [20.01.23 Computational Chemist \(209823W-01\), Indianapolis, IN](#)
- [20.01.23 Postdoctoral fellow at University of Arkansas for Medical Sciences, Little Rock, AR](#)
- [20.01.23 Senior Computational Chemist, Sygnature Discovery, Nottingham or Alderley Park, UK](#)
- [20.01.23 Chemoinformatics/Computational Chemistry Scientist, University of Dundee, UK](#)
- [20.01.23 Senior Cloud Consultant \(m/f/d\)](#)
- [20.01.22 Postdoc in : Machine Learning Aided Homogeneous Catalysis](#)
- [20.01.21 PhD position in Homogeneous Catalysis](#)
- [20.01.17 Ab initio Simulation Specialist, HQS Quantum Simulations, Karlsruhe, Germany](#)
- [20.01.17 Scientific Software Developer - Evotec UK Ltd](#)
- [20.01.16 Computational Chemist, University of Dundee, UK](#)
- [20.01.15 PhD POSITION IN MOLECULAR SIMULATIONS OF PROTEIN COMPLEXES, THE UNIVERSITY OF AUCKLAND, AUCKLAND, NEW ZEALAND](#)
- [20.01.13 IMMEDIATE OPENING FOR A POSTDOC POSITION IN BIOMOLECULAR SIMULATIONS AND MACHINE LEARNING IN NEW YORK CITY, USA](#)
- [20.01.13 Computational Structural Biologist, ProteinQure, Toronto, Canada](#)
- [20.01.13 PhD student in Theoretical Chemistry and Machine Learning](#)
- [20.01.13 Drug Discovery Opportunities at D. E. Shaw Research](#)
- [20.01.12 Postdoctoral Associate, Polarizable Simulations, Virginia Tech Biochemistry, USA](#)
- [20.01.09 Postdoc, Carnegie Mellon University, Modeling of Membrane Protein Allostery and Drugs](#)
- [19.03.25 Computational Chemist Post Doc Position, and Bioinformatics/Computer Programmer in Machine Learning, OPEN at Georgetown University Medical Center, Washington DC, USA](#)
- [20.01.07 Postdoc in Machine Learning in Quantum Chemistry](#)
- [20.01.03 Postdoc, Theoretical/Comp. Chemistry, NYU Shanghai, China](#)
- [19.12.26 2020 Fall Openings in Theoretical and Computational Chemistry Ph.D. Program in Dallas, Texas, USA](#)
- [19.12.23 3-YEAR POSTDOC POSITION in \(COARSE-GRAINED\) MOLECULAR DYNAMICS, Groningen, The Netherlands](#)
- [19.12.21 Computational Catalysis Postdoc](#)
- [19.12.20 X-ray Crystallography Software Development \(Postdoctoral Research Associate\)](#)
- [19.12.20 Scientific Software Developers \(Ref: SD/1219\)](#)
- [19.12.19 Postdoc Fellow - Cryo-EM based molecular dynamics and modelling, AstraZeneca, Boston](#)

METHODS

1. EMPIRICAL APPROACHES

a) Molecular mechanics (MM methods) (empirical force field calculations)

- based on classical-mechanical model of molecular structures
- each atom is simulated as a single particle
- potential energy of molecules is calculated based on a given **force field**
- electrons are not considered explicitly (they find their optimum distribution once the position of nuclei are known) = Born-Oppenheimer approximation
- atoms are considered spheres (with a given **radius** (vdW) and **polarizability**) and a net **charge**
 - vdW radius
 - from equation of state for the atomic gas in question
 - from crystallographic measurements in case of molecular crystals
 - polarizability - from electric susceptibility
 - $\alpha = \chi_e (\epsilon_0 k_B T / p)$ $\chi_e = \epsilon_r - 1$; $V_{vdW} = \alpha / (4\pi\epsilon_0)$
- interactions
 - based on classical potentials
 - determine the spatial distribution of atoms and their energies



van der Waals radii

Element	radius (Å)
Hydrogen	1.2 (1.09) ^[1]
Carbon	1.7
Nitrogen	1.55
Oxygen	1.52

Van der Waals radius

- radius of an imaginary hard sphere which can be used to model the atom for many purposes
- determined from the mechanical properties of gases, from measurements of atomic spacing between pairs of un-bonded atoms in crystals or from measurements of electrical or optical properties (the polarizability and the molar refractivity).

The potential energy of the molecular system:

$$E = E_{\text{covalent}} + E_{\text{non-covalent}}$$

$$E_{\text{covalent}} = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}}$$

$$E_{\text{non-covalent}} = E_{\text{electrostatic}} + E_{\text{van der Waals}} \quad (\text{non-bonded energy})$$

- stretching energy:

$$\sum_{\text{bonds}} k_b (r - r_0)^2 + [\text{higher order terms}]$$

- bending energy:

$$\sum_{\text{angles}} k_\theta (\theta - \theta_0)^2 + [\text{higher order terms}]$$

- torsion energy:

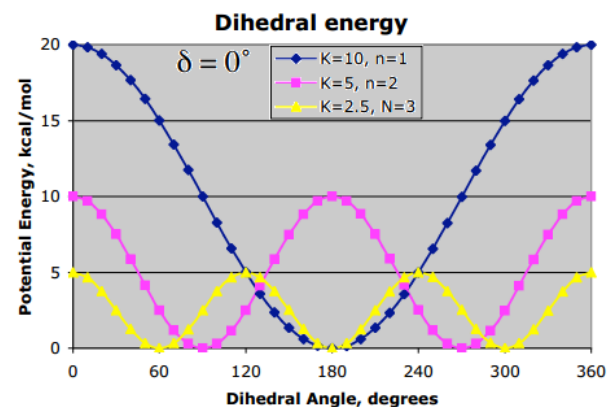
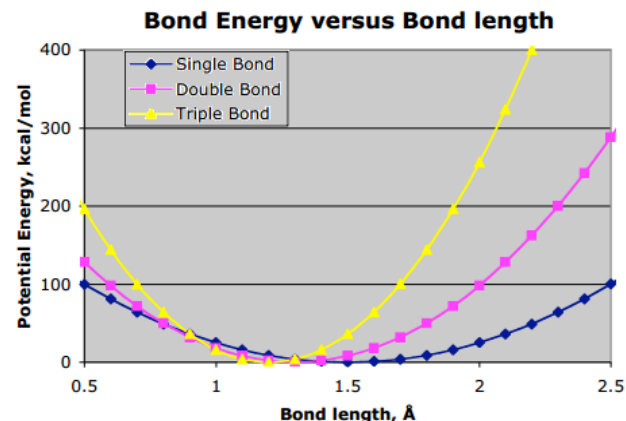
$$\sum_{\text{torsions}} K_{\Phi n} (1 + \cos(n\Phi + \Phi_0)) \quad K_1(1 + \cos(\Phi - \Phi_0)) + K_2(1 + \cos(2\Phi - \Phi_0)) + K_3(1 + \cos(3\Phi - \Phi_0)) + K_4(1 + \cos(4\Phi - \Phi_0))$$

- non-bonded energy:

$$\sum_i \sum_j \frac{q_i q_j}{r_{ij}} + \sum_i \sum_j \epsilon_{ij} \left(-\frac{\sigma_{ij}}{r_{ij}^6} + \frac{\sigma_{ij}}{r_{ij}^{12}} \right) + \text{H bond term}$$

Credit:

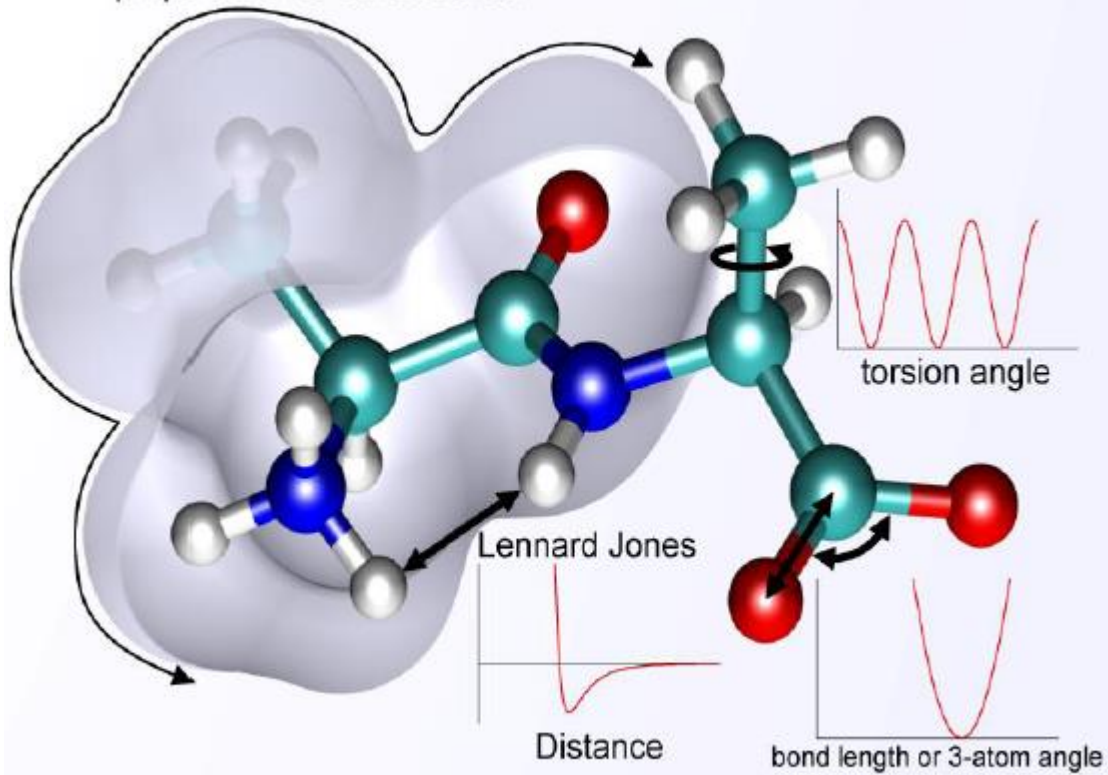
<http://www.ks.uiuc.edu/Training/Workshop/SanFrancisco/lectures/Wednesday-ForceFields.pdf>



water

Hydrophobic effect is roughly proportional to surface area

Continuum solvent model



http://en.wikipedia.org/wiki/Force_field_%28chemistry%29#mediaviewer/File:MM_PEF.png

Each molecular mechanics method is characterized by its particular

force field = interatomic **potential functional** form **and parameters sets**

- used to describe the potential energy of a molecular system
- derived from experiment or high-level quantum mechanical calculations

Force field

- i) *a set of equations* defining how the potential energy of a molecule varies with the positions of its constituent atoms
- ii) *a series of atom types*, defining the characteristics of an element within a specific chemical context. The *atom type* depend on hybridization, charge and types of the other atoms to which it is bonded.
- iii) one or more *parameter sets* that fit the equations and atom types to experimental data. A parameter set contains *force constants* (values used to relate atomic characteristics to energy components) and *structural data* (bond lengths, valence and dihedral angles)

All-atom force fields - contain parameters for each atom in a system

United-atom force field - treat C and H atoms in CH₂ and CH₃ groups as single interaction centers

Disadvantages of MM methods:

- each force field provides **good results for a limited class of molecules** (for which it was parameterized).
 - **No force field is general!**
- since empirical methods do not consider the electrons, they **can not describe bond formation and breaking**
- many molecular properties depending on **subtle electronic effects are not reproducible by MM methods**.

Molecular mechanics methods are generally used for energy minimizations and this is useful for:

- ✓ **energy barriers** between different conformers
- ✓ steepness of the **PES** around a local minimum
- ✓ **protein folding** kinetics
- ✓ protonation equilibria
- ✓ **active site** coordination
- ✓ designing **binding sites**

Results from MM methods

- optimized geometries
- potential energy surfaces
- docking
- vibrational spectra
- thermochemical parameters

See Charmm parameter set!

```
#####
##                               ##
##  Force Field Definition      ##
##                               ##
#####
```

```
forcefield          CHARMM22

vdwtype             LENNARD-JONES
radiusrule          ARITHMETIC
radiustype          R-MIN
radiussize          RADIUS
epsilonrule         GEOMETRIC
vdw-14-scale        1.0
chg-14-scale        1.0
electric            332.0716
dielectric          1.0
```

http://www.studia.pwr.wroc.pl/p/mat/bioinformatics/md_lect_part1_revisedlang2.pdf

Bond stretch parameters for alanine residue

Bond	AMBER94		CHARMM22	
	r_0	k_r	r_0	k_r
CB-HB	1.090	340.0	1.111	322.0
CA-CB	1.526	310.0	1.538	222.5
CA-C	1.522	317.0	1.490	250.0
C-O	1.229	570.0	1.230	620.0
CA-HA	1.090	340.0	1.080	330.0
CA-N	1.449	337.0	1.430	320.0
N-HN	1.010	434.0	0.997	440.0

r_0 is in [Å] and k_r in [kcal·mol⁻¹Å⁻²]

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

Bending parameters for alanine residue

Angle	AMBER94		CHARMM22	
	θ_0	k_θ	θ_0	k_θ
HB-CB-CA	109.50	50.00	110.10	33.43
CB-CA-N	109.70	80.00	113.50	70.00
CB-CA-C	111.10	63.00	108.00	52.00
CB-CA-HA	109.50	50.00	111.00	35.00
CA-C-O	120.40	80.00	121.00	80.00
CA-N-HN	118.04	30.00	117.00	35.00
N-CA-C	110.10	63.00	107.00	50.00
N-CA-HA	109.50	50.00	108.00	48.00
C-CA-HA	109.50	50.00	109.50	50.00

θ_0 is in [deg] and k_θ in [kcal·mol⁻¹deg⁻²]

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

Torsion parameters for alanine residue

Torsion	AMBER94			CHARMM22		
	V	ψ	n	V	ψ	n
O-C-CA-N	0.000	0.0	2	0.000	0.0	2
C-CA-N-HN	0.000	0.0	2	0.000	0.0	2
N-CA-CB-HB	0.156	0.0	3	0.200	0.0	3
C-CA-CB-HB	0.156	0.0	3	0.200	0.0	3
HA-CA-CB-HB	0.156	0.0	3	0.200	0.0	3

V is in [kcal/mol] and ψ in [deg]

$$\sum_{\text{torsions}} K_{\Phi_n} (1 + \cos(n\Phi + \Phi_0))$$

Partial atomic charges in alanine residue

Atom	Amber94	Charmm22
N	-0.4157	-0.47
HN	-0.2719	0.31
CA	0.0337	0.07
HA	0.0603	0.09
CB	-0.1825	-0.27
HB1	0.0823	0.09
HB2	0.0823	0.09
HB3	0.0823	0.09
C	0.5973	0.51
O	-0.5679	-0.51

$$\sum_i \sum_j \frac{q_i q_j}{r_{ij}} + \sum_i \sum_j \epsilon_{ij} \left(-\frac{\sigma_{ij}}{r_{ij}^6} + \frac{\sigma_{ij}}{r_{ij}^{12}} \right)$$

VdW parameters for alanine residue

Atom	AMBER94		CHARMM22	
	σ	ϵ	σ	ϵ
N	1.8240	0.1700	1.8500	-0.2000
HN	0.6000	0.0157	0.2245	-0.0460
CA	1.9080	0.1094	2.2750	-0.0200
HA	1.3870	0.0157	1.3200	-0.0220
CB	1.9080	0.1094	2.0600	-0.0800
HB	1.4870	0.0157	1.3200	-0.0220
C	1.9080	0.0860	2.0000	-0.1100
O	1.6612	0.2100	1.7000	-0.1200

σ is in [Å] and ϵ in [kcal·mol⁻¹]

b) Molecular dynamics (MD methods)

- consists in following the temporal evolution of a system of interacting atoms
 - by integrating their equations of motions: $F_i = m_i a_i$
 - being given a set of positions and velocities => evolution in time is completely determined
- it is a **statistical mechanics method** that provide configurations distributed according to some statistical distribution function.
- forces acting on atoms are derived from the potential:

$$F_i = -\nabla_{r_i} V(r_1, \dots, r_N) \quad E = T + V \text{ -conserved}$$

The simplest choice for V is to write it as a sum of pairwise interactions:

$$V(r_1, \dots, r_N) = \sum_i \sum_{j>i} \Phi(|r_i - r_j|)$$

The simplest choice is the "**pair potential**", in which the total potential energy can be calculated from the sum of energy contributions between pairs of atoms.

Lennard-Jones potential:

$$\Phi_{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right]$$

The intramolecular bonding interactions must be also considered:

- stretching energy: $\sum_{bonds} k_b (r - r_0)^2$
- bending energy: $\sum_{angles} k_\theta (\theta - \theta_0)^2$
- torsion energy: $\sum_{torsions} K_{\Phi_n} (1 + \cos(n\Phi + \Phi_0))$

Advantages of MD methods

- can treat very large molecular systems (thousands of atoms)
- used in biophysics (proteins, viruses) and material science

Disadvantages of MD methods

- not able to model bond forming/breaking since electronic structure methods does not enter these models

Those interested in MD:

<http://phys.ubbcluj.ro/~titus.beu/teaching.html>

2. QUANTUM MECHANICAL APPROACHES

a) semiempirical methods (AM1, PM3, PPP, INDO, MINDO, ...)

- **approximate** methods based on **Hartree-Fock formalism** in which
 - some quantities (electron repulsion integrals) are taken from experiment
 - some (small) quantities are neglected
 - some quantities are estimated by fitting to experimental data;

Empirical parameters and functions are used to compensate for errors introduced by neglecting some integrals

- **model only the valence electrons** (or π electrons --> Huckel method)
- core electrons are treated together with the nuclei as one ECP
- $E_{AB} = Z'_A Z'_B e^2 / R_{AB}$ where Z' are reduced nuclear charges
- limited to hundreds of atoms
- can be used to study ground and excited molecular states
- parameterized to reproduce experimental reference data or accurate high-level theoretical predictions (they contain sets of parameters)

Total energy of the molecule is represented as the sum of the electronic energy (net negative) and the core repulsions (net positive).

A "**training set**" of **molecules** is selected, chosen to cover as many types of bonding situations as possible.

- A non-linear least squares optimization procedure is applied with the values of the various adjustable parameters as variables and a set of measured properties of the training set as constants to be reproduced.
- The **measured properties** include heats of formation, geometrical parameters, dipole moments, ionization potentials, etc.

Depending upon the choice of training sets, the exact numbers of types of adjustable parameters, and the mode of fitting to experimental properties, different semi-empirical methods have been developed and are incorporated into semi-empirical software.

- AM1 (Austin Model 1) method of M. J. S. Dewar
- PM3 (Parameter Model 3) method of J. J. P. Stewart
- PM6, PM7
- SAM1
 - MOPAC program (<http://openmopac.net>)

b) non empirical methods

- do not require empirical parameters
- can be used for any molecular system
- limited to tens of atoms
- can be used to study ground and excited molecular states
- reliable for strong (covalent) and weak (non-covalent) interactions

i) *Ab initio* methods  Based on **wavefunction**

ii) DFT methods  Based on **electronic density**

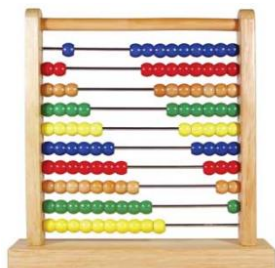
Based on the following principles:

- nuclei and electrons are distinguished from each other
- electron-electron and electron-nuclear interactions are considered explicitly.
- interactions are governed by nuclear and electron charges (potential energy) and electron motions
- interactions determine the spatial distribution of nuclei and electrons and their energies

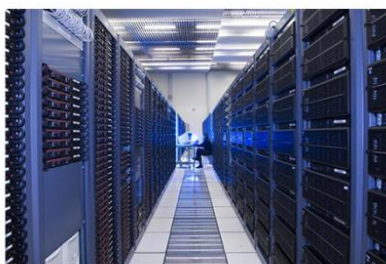
What level of theory?

(level of theory = method + basis set)

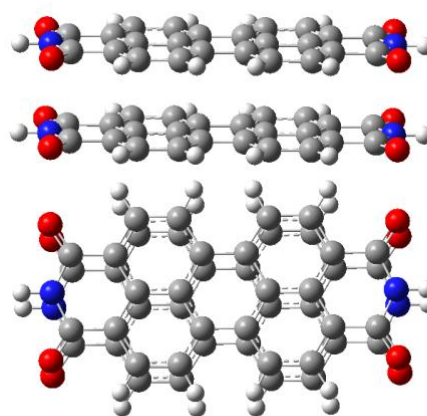
1. determined by the **property** being calculated and the **cost** of the calculation in terms of computer usage.
2. **the higher the level of theory the greater the computational cost.**
For example, Hartree-Fock based ab-initio methods are more costly than semiempirical methods. Correlated methods, depending on the sophistication, add a significant additional cost to the Hartree-Fock method.
3. For the Hartree-Fock and Density Functional methods an additional factor is the **quality (size) of the basis sets** used.
4. In general the larger the number of functions used in the basis set the more flexible (better) the basis set. Each additional function adds extra cost to the calculation.



MM methods



QM methods



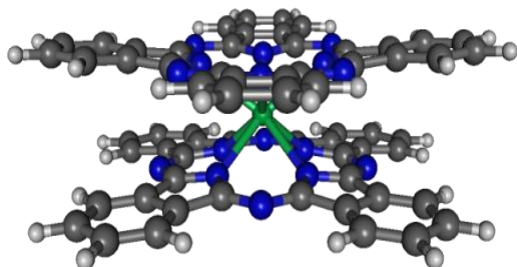
PTCDI dimer

Job CPU time

AM1: 6s

DFT: 55m 22s

MP2: 8h 14m 4s



Job cpu time: 121 days 18 hours 20 minutes 42.0 seconds.

Summary

Methods of computational chemistry

Empirical

MM

MD

Quantum
mechanical

Semiempirical

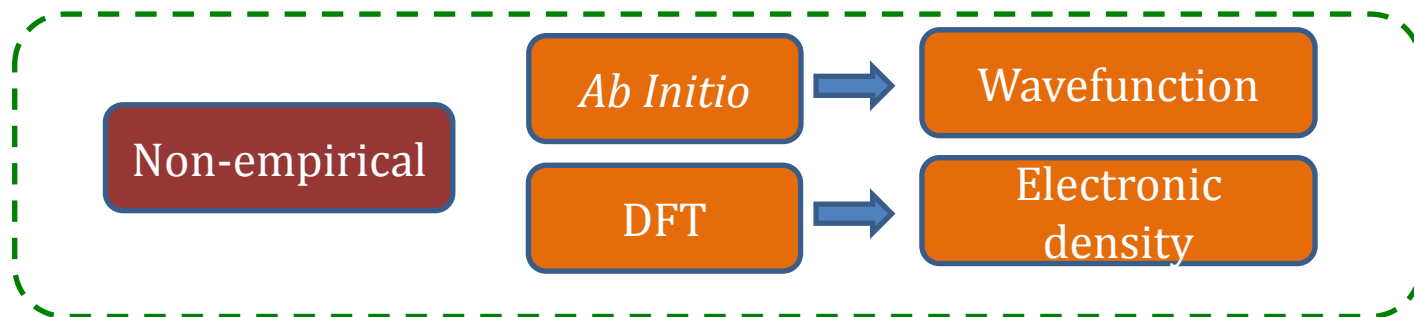
Non-empirical

Ab Initio

DFT

Wavefunction

Electronic
density



Type of calculations

1. Single point calculations

► *energy, wave-function and other requested properties at a single fixed geometry*

- usually done first at the beginning of a study on a new molecule to check out the nature of the wave-function
- it is also frequently carried out after a geometry optimization, with a larger basis set or more superior method than is possible with the basis set and method used to optimize geometry

These calculations are performed to:

- obtain basic information about a given molecule
- check a molecular geometry to be used as a starting geometry in an optimization process
- compute more accurate energies and other properties for an optimized geometry at a lower theoretical level.
- ...

One cycle:

$F_i \Phi_i = \epsilon_i \Phi_i$ in matrix form: $FC = CE$ (SCF cycle)

2. Geometry optimization calculations

► *calculation of the wave-function and energy at a starting geometry and then searching for the geometry which will give the lowest energy (local or global energy minimum)*

- Used to locate the lowest energy geometry (structure) of a molecule or molecular complex
- Used to locate transition structures on the potential energy surfaces

Two cycles:

Geometry cycle

SCF cycle

G0

SCF0

G1

SCF1

...

Gn

SCFn



Final geometry



Final wavefunction

Potential energy surface (PES)

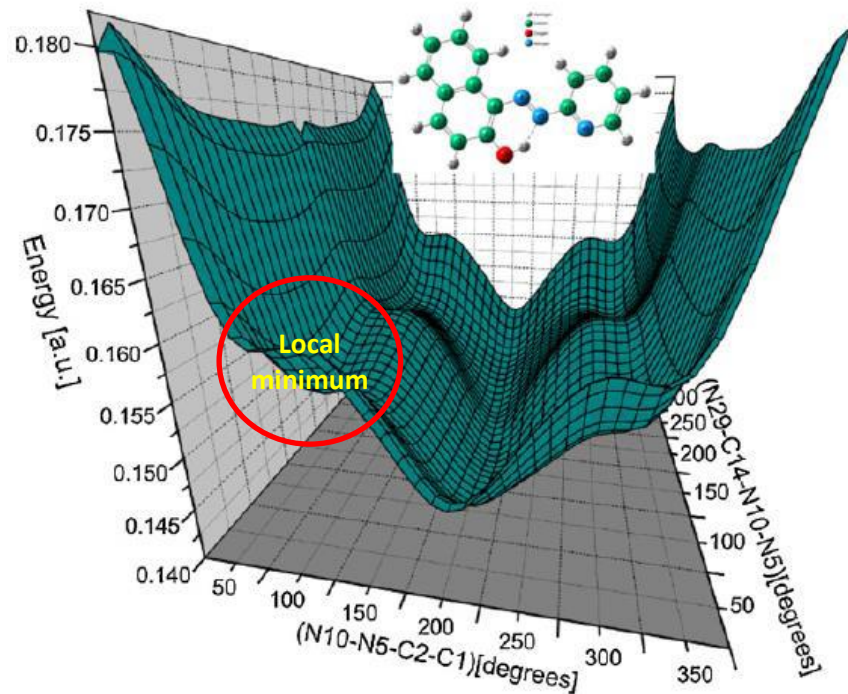
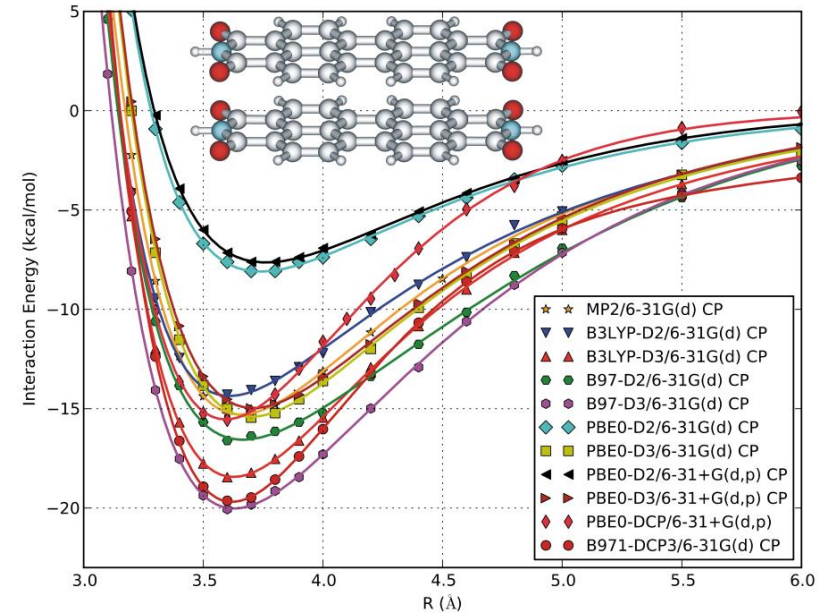


Fig. 2. PES of PAN scanned at PM6 level of theory.



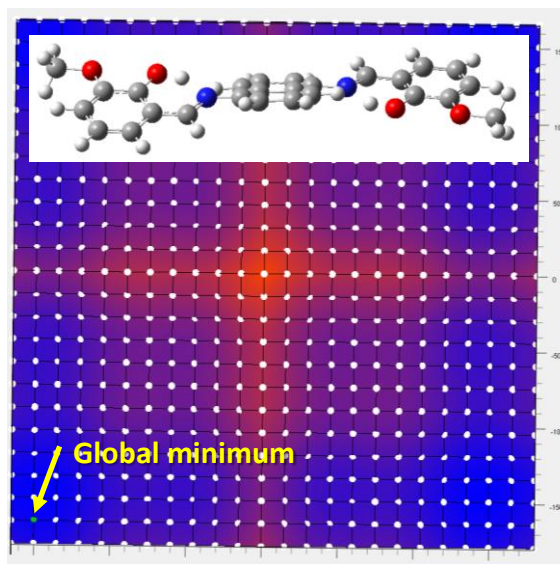
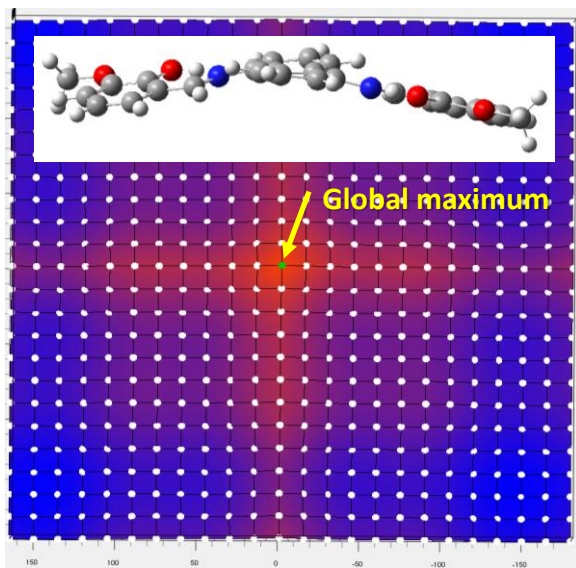
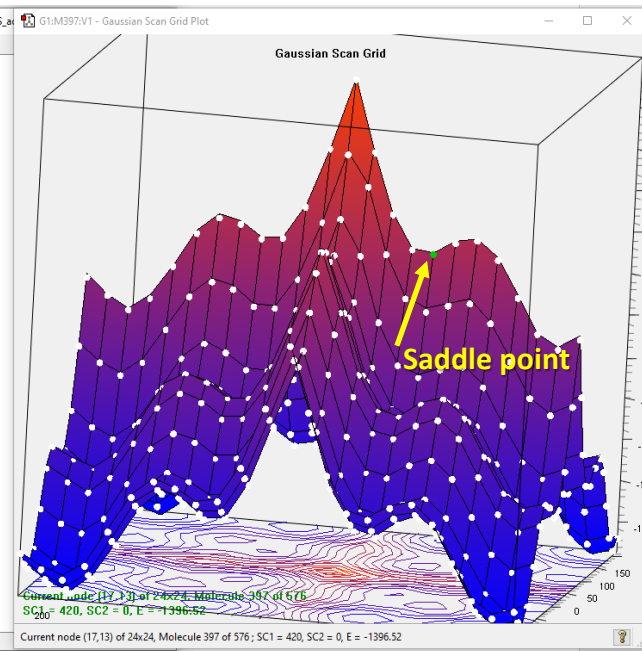
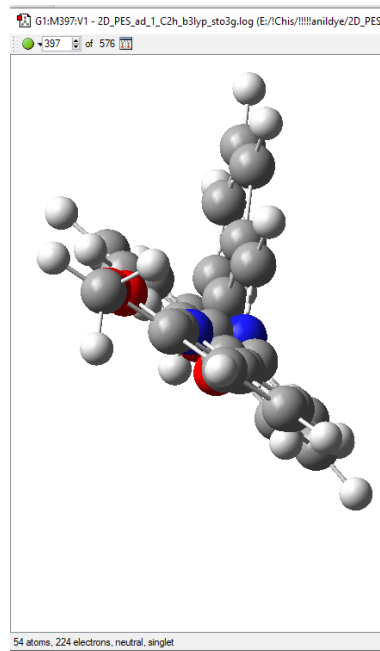
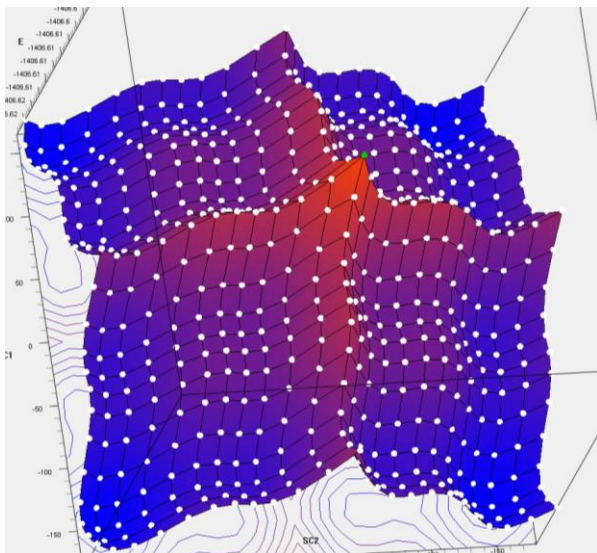
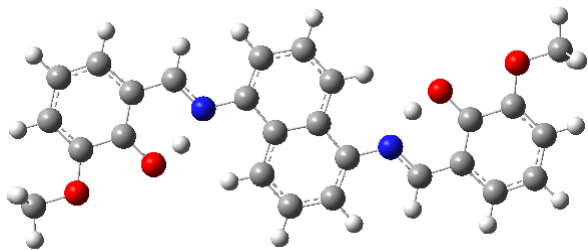
local minimum → the lowest energy point *in a limited region* of PES

global minimum → the lowest energy point *anywhere* on the PES

saddle points → maximum in one direction and minimum in all other directions
→ correspond to transition state structures

all minima + saddle points = **stationary points** $\nabla E = 0$ ($F = 0$)

Geometry optimizations usually locate the stationary points closest to the starting geometry!!!



Conformational search

- necessary to distinguish between a local minimum and the global minimum.

how?

- altering the initial geometry slightly (usually by dihedral angles) and then performing another optimization

3. Frequency calculations

- to identify the nature of stationary points on the PES (check if an optimized geometry corresponds or not to an energy minimum)
- to predict the IR and Raman spectra of molecules (frequencies, intensities and normal modes)
- to compute force constants for a geometry optimization
- to compute zero-point vibrational energies, thermal energy corrections, enthalpy and entropy of the system
- to compute polarizability and hyperpolarizability tensor
- should only be carried out at the geometry obtained from an optimization run and with the same basis set and method.

For a local or a global minimum all the calculated frequencies will be positive (real).

For any stationary point other than a minimum some of the frequencies will be negative (imaginary frequencies).

scaling factors see: CCCBDB (Computational Chemistry Comparison and Benchmark DataBase)

<http://cccbdb.nist.gov/>

4. Magnetic properties calculations

NMR spectra

- chemical shifts, spin-spin couplings

ESR spectra

- hyperfine coupling constants and hyperfine coupling tensors

Molecular properties calculated by Gaussian 09

http://www.gaussian.com/g_ur/m_jobtypes.htm

Atomic charges: **Pop**

Dipole moment: **Pop**

Electron affinities via propagator methods: **OVGF**

Electron density: **cubegen**

Electronic circular dichroism: **TD**

Electrostatic potential: **cubegen**, **Prop**

Electrostatic-potential derived charges: **Pop=Chelp**, **ChelpG** or **MK**

Frequency-dependent polarizabilities/hyperpolarizabilities: **Polar** **CPHF=RdFreq**

High accuracy energies: **CBS-QB3**, **G2**, **G3**, **W1U**

Hyperfine coupling constants (anisotropic): **Prop**

Hyperfine spectra tensors (incl. *g* tensors): **NMR** and **Freq=(VibRot, Anharmonic)**

Hyperpolarizabilities: **Freq**, **Polar**

Ionization potentials via propagator methods: **OVGF**

IR and Raman spectra: **Freq**

Pre-resonance Raman spectra: **Freq** **CPHF=RdFreq**

Molecular orbitals: **Pop=Regular**

Multipole moments: **Pop**

NMR shielding and chemical shifts: **NMR**

NMR spin-spin coupling constants: **NMR=SpinSpin**

Optical rotations: **Polar=OptRot** **CPHF=RdFreq**

Polarizabilities: **Freq**, **Polar**

Thermochemical analysis: **Freq**

UV/Visible spectra: **CIS**, **Zindo**, **TD**

Vibration-rotation coupling: **Freq=VibRot**

Vibrational circular dichroism: **Freq=VCD**

H atom abstraction in propene

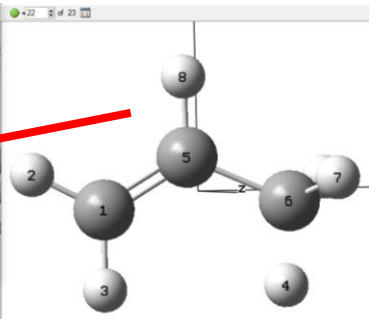
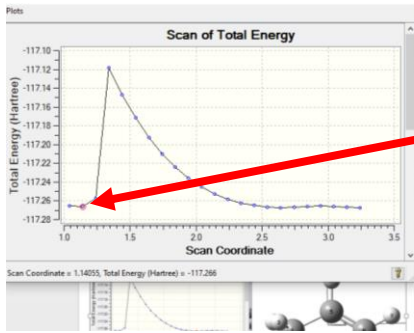
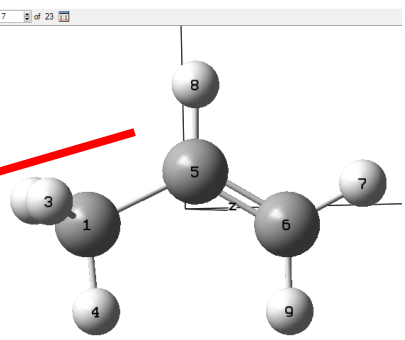
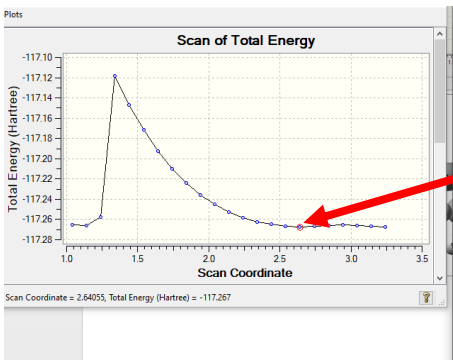
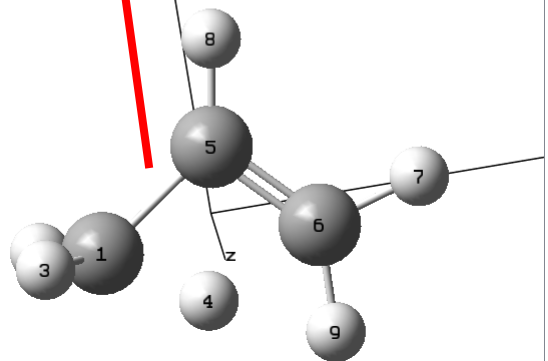
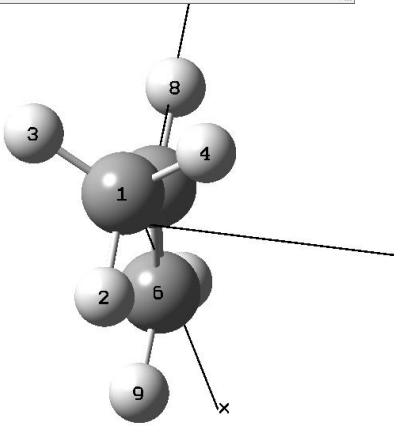
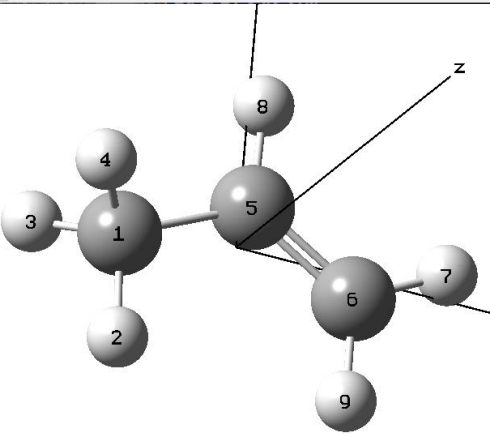
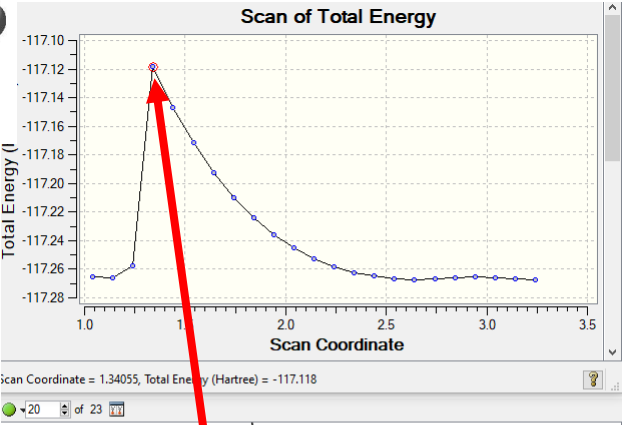
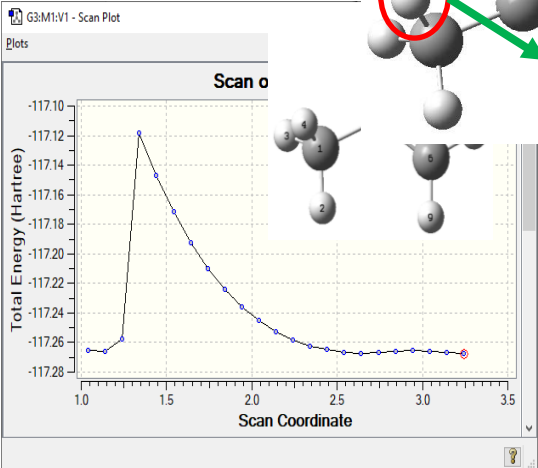
File Edit Options Help

%chk=D:\cspm\c1_Intro\CH3CHCH2_S_b3_321g.chk
opt=modredundant b3lyp/3-21G

Title Card Required

C	-1.23917200	-0.16478700	0.00001000
H	-1.17621200	-1.25844100	-0.00057100
H	-1.81163700	0.14978800	-0.88326900
H	-1.81119800	0.14887900	0.88389700
C	0.13605900	0.46062200	-0.00004200
C	1.28428500	-0.22250900	0.00004600
H	2.24787600	0.27951200	0.00001000
H	0.16249300	1.55107300	-0.00001000
H	1.30164700	-1.31076900	-0.00014500

B 6 4 S 22 -0.1



Molecular properties calculated as energy derivatives

Derivative	Property		
$\frac{dE}{dx_i}$	forces on nuclei	$\frac{d^2 E}{dx_i d\epsilon_\alpha}$	dipole derivatives, harmonic IR intensities
$\frac{d^2 E}{dx_i dx_j}$	force constants	$\frac{d^3 E}{dx_i d\epsilon_\alpha d\epsilon_\beta}$	polarizability derivatives, Raman intensities
$\frac{dE}{d\epsilon_i}$	dipole moment	$\frac{d^2 E}{dB_\alpha dB_\beta}$	magnetic susceptibility and <u>magnetizability</u>
$\frac{d^2 E}{d\epsilon_\alpha d\epsilon_\beta}$	polarizability	$\frac{d^2 E}{dm_{Kj} dB_i}$	magnetic shielding tensors; NMR chemical shifts
$\frac{d^3 E}{d\epsilon_\alpha d\epsilon_\beta d\epsilon_\beta}$	<u>Hiperpolarizability</u>	$\frac{d^2 E}{dI_{Ki} dI_{Lj}}$	spin-spin coupling constants
$\frac{d^3 E}{dx_i dx_j dx_k}$	cubic force constants	$\frac{dE}{dS_i}$	spin densities, hyperfine coupling constants
		$\frac{d^2 E}{dS_i dS_j}$	g tensor

Energetic aspects

$ABCD... \rightarrow \text{nuclei} + \text{electrons}(R = \infty)$ **total energy, E_{tot}**

$ABCD... \rightarrow A + B + C + D...(R = \infty)$ **atomisation energy, E_{at}**

$ABC - D \rightarrow ABC + D(R = \infty)$ **dissociation energy, D_e^{CD}**

$ABCD... \rightarrow ABCD...^{n+} + ne^-(R = \infty)$ **nth-ionisation potential, IP_n**

$ABCD... + e^-(R = \infty) \rightarrow ABCD...^{\ominus}$ **electron affinity, EA**

Program packages in molecular electronic structure calculations

1. Gaussian
<http://www.gaussian.com/>
2. ORCA
<https://orcaforum.cec.mpg.de/>
3. Gamess
<http://www.msg.ameslab.gov/GAMESS/GAMESS.html>
4. NWChem
http://www.nwchem-sw.org/index.php/Main_Page
5. DeFT
<http://lavoisier.dq.ufscar.br/download/chem/dft/>
6. DALTON
<http://www.kjemi.uio.no/software/dalton/dalton.html>
7. Mopac
<http://comp.chem.umn.edu/WWW/MOPAC/MOPAC.html>

See a more complete list of computational chemistry software at:

http://en.wikipedia.org/wiki/List_of_quantum_chemistry_and_solid_state_physics_software

or:

<http://ccl.net/chemistry/links/software/index.shtml>

Molecular structure and properties visualization programs

1. GaussView
<http://www.gaussian.com>
2. Molegro Molecular Viewer
<http://www.clcbio.com/>
3. Mercury
<http://www.ccdc.cam.ac.uk/Solutions/CSDSystem/Pages/Mercury.aspx>
4. Gabedit
<http://gabedit.sourceforge.net/>
5. Molekel
<http://www.cscs.ch/molekel/>
6. Molden
<http://www.ccl.net/cca/software/SOURCES/FORTRAN/molden/index.shtml>

Useful Molecular Databases

1. Drug Bank
<http://www.drugbank.ca/>
2. Spectral Database for Organic Compounds SDBS
http://sdb.sdb.aist.go.jp/sdb/cgi-bin/direct_frame_top.cgi
3. Crystallography Open Database
<http://www.crystallography.net/>
4. Periodic Table for Crystal Structures
http://openmopac.net/PM7_accuracy/Periodic_table_solids.html
5. Fluorophores
<http://www.fluorophores.tugraz.at/substance/>
6. Benchmark Energy and Geometry Database
<http://www.begdb.com/>
7. Dielectric constants of various materials
<http://www.clippercontrols.com/pages/Dielectric-Constant-Values.html>
8. The binding database
http://www.bindingdb.org/validation_sets/index.jsp
9. MyPDB
<http://www.rcsb.org/pdb/mypdb/signup.do>
10. The Protein Model Portal
<http://www.proteinmodelportal.org/?pid=101>