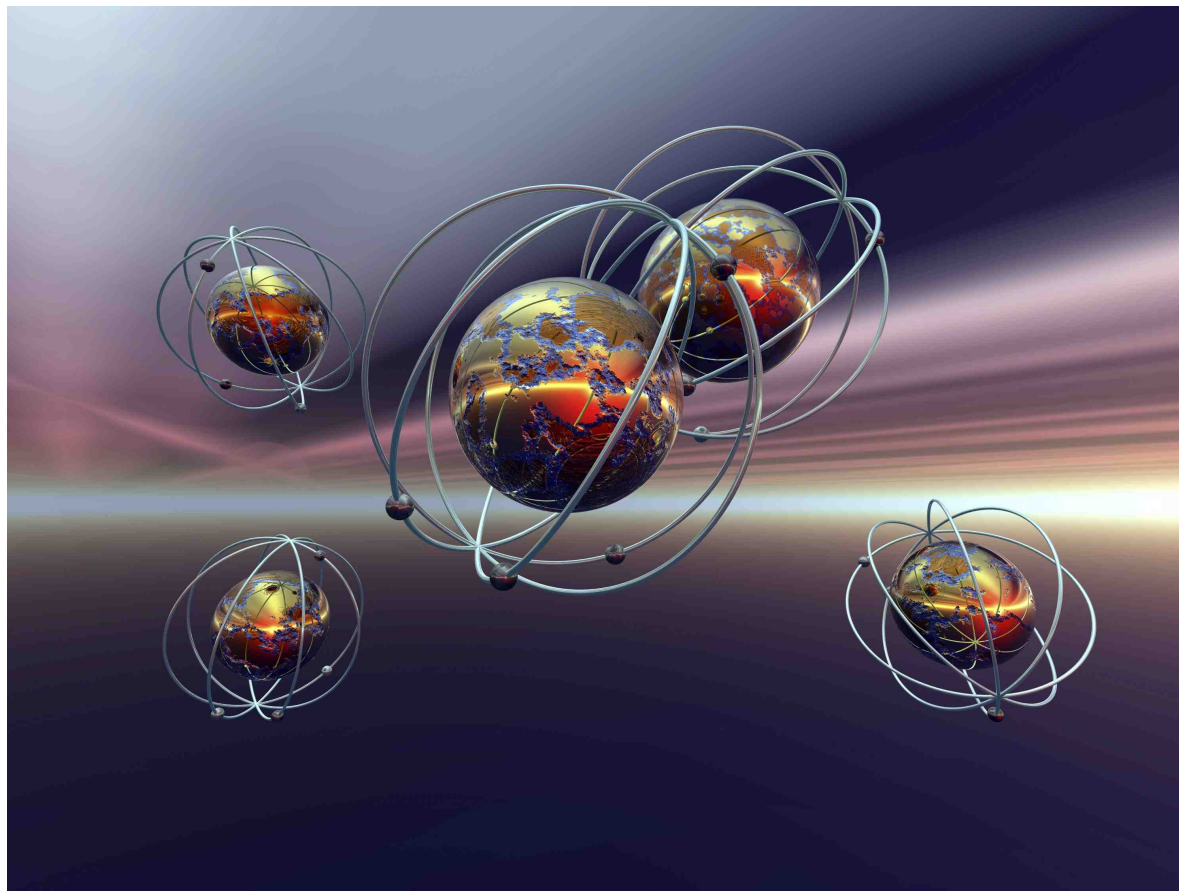


Wintersemester 2015/2016
Biomolecular Engineering/Nanobiophysics Module

LECTURE 6: QM BASICS

FOR COMPUTATIONAL CHEMISTRY



LECTURE 6: QM BASICS FOR COMPUTATIONAL CHEMISTRY

- Milestones of quantum mechanics
- Schrödinger equation
- Methods in quantum chemistry: HF, DFT, SCF
- Geometry optimization, transitional state
- Spectroscopy, NMR
- QM/MM and ONIOM
- Solvent
- Quantum chemistry software
- Case study 1: fluorinated amino acids
- Case study 2: sugar's ring conformations



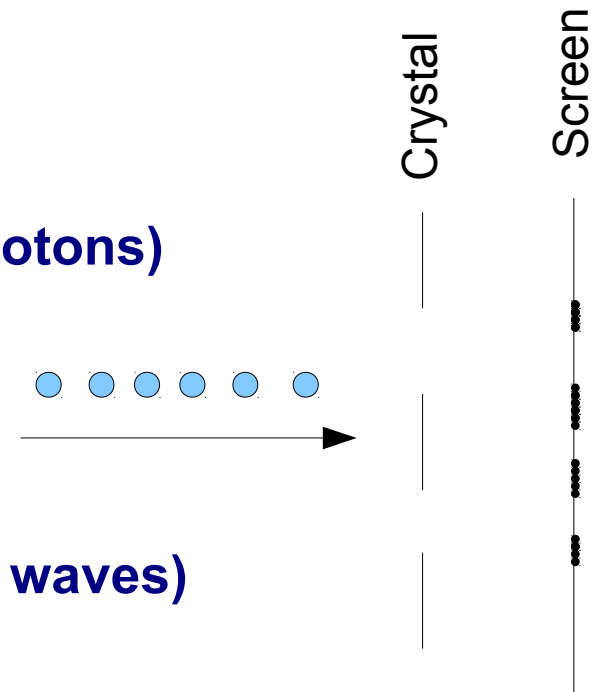
MILESTONES OF QM

- Quantum mechanics describe quantum systems and the laws of their motions (not accessible by classical mechanics)

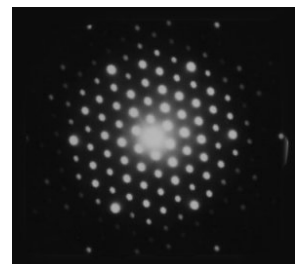
- Planck: Spectrum=f(T) => quantum oscillators

$$E = \frac{h}{2\pi} \omega; \quad h = 6.626\,069\,57 \times 10^{-34} \text{ J s}$$

- Einstein: photo-effect, light consists of quanta (photons)
- Bohr: electrons in atoms are in discrete states
- de Broglie: wave nature of light hypothesis
- Davisson/Jermer: electrons diffraction (de Broglie waves)
- Schrödinger: wave mechanics
- Heisenberg: uncertainty principle
- Compton: effect of the light's scattering



$$\delta x \delta p \geq \frac{h}{4\pi}$$



SCHRÖDINGER EQUATION

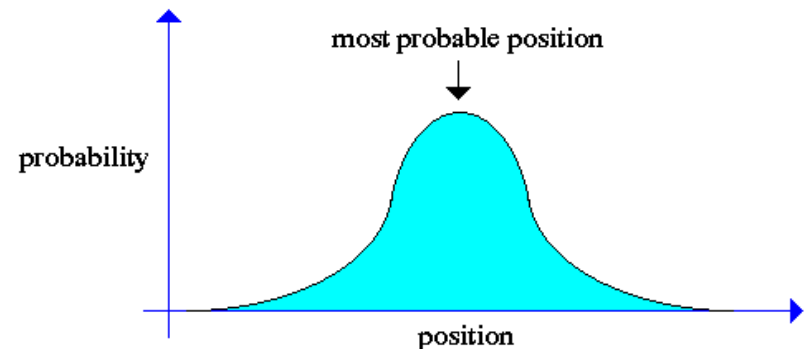
$$H \psi = E \psi ; \quad \langle \psi(x) | \psi(x) \rangle = \text{Probability}(x)$$

$$H = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k} \nabla_k^2 - \sum_i \sum_k e^2 \frac{Z_k}{r_{ik}} + \sum_{i < j} \frac{e^2}{r_{ij}} + \sum_{k < l} \frac{e^2 Z_k Z_l}{r_{kl}}$$

Quantum Wave Function

$$\int \int \int \psi_i \psi_j dx dy dz = \delta_{ij} - \text{Orthonormality}$$

$$\int \int \int \psi_j H \psi_i dx dy dz = \langle \psi_i | H | \psi_j \rangle = E_i \delta_{ij}$$



- **Born-Oppenheimer approximation decouples the Hamiltonian into “slow” nucleus movement and electronic term, because the nuclei move much slower than electron.**
- **Quantum chemistry: application of QM to study chemical behaviour.**
- **Quantum chemistry is a set of approximations!**

ATOMIC ORBITALS

- **Atomic orbitals:** is a mathematical function that describes the wave-like behavior of either one electron or a pair of electrons in an atom - $\Psi(x)$

	s ($l=0$)	p ($l=1$)			d ($l=2$)					f ($l=3$)							
	m=0	m=0	m=±1		m=0	m=±1		m=±2		m=0	m=±1		m=±2		m=±3		
	s	p_z	p_x	p_y	d_{z^2}	d_{xz}	d_{yz}	d_{xy}	$d_{x^2-y^2}$	f_z^3	f_{xz^2}	f_{yz^2}	f_{xyz}	$f_z(x^2-y^2)$	$f_x(x^2-3y^2)$	$f_y(3x^2-y^2)$	
n=1																	
n=2																	
n=3																	
n=4																	
n=5										
n=6					
n=7		

Energetic order: 1s, 2s, 2p, 3s, 3p, 4s, 3d, 4p, 5s, 4d, 5p, 6s, 4f, 5d, 6p, 7s, 5f, 6d, 7p

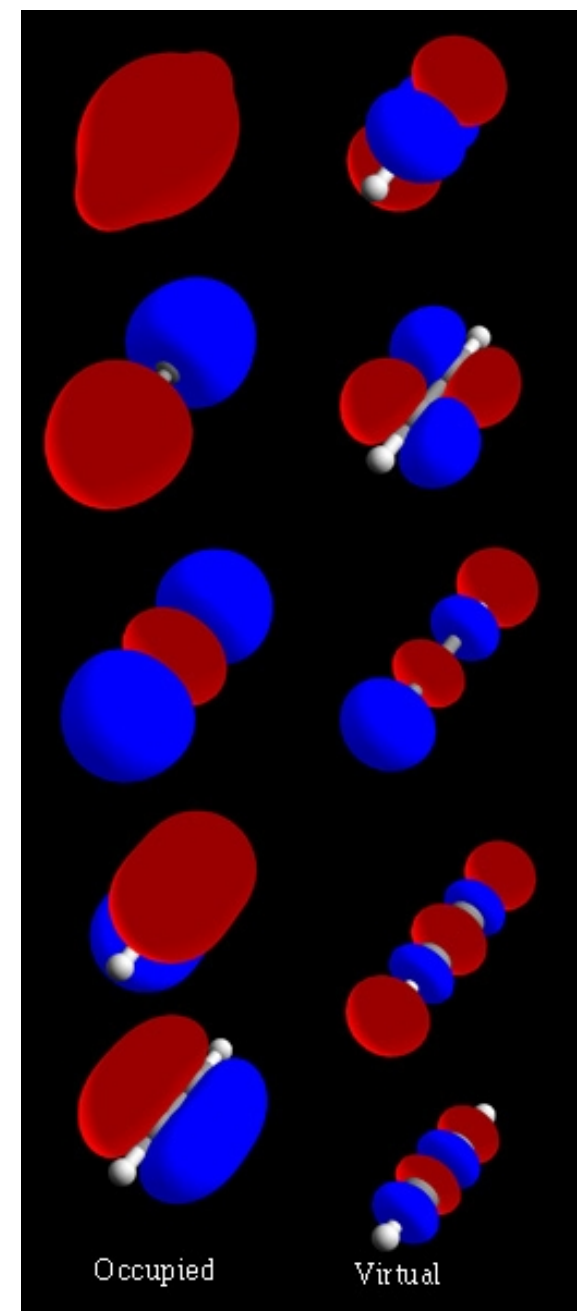
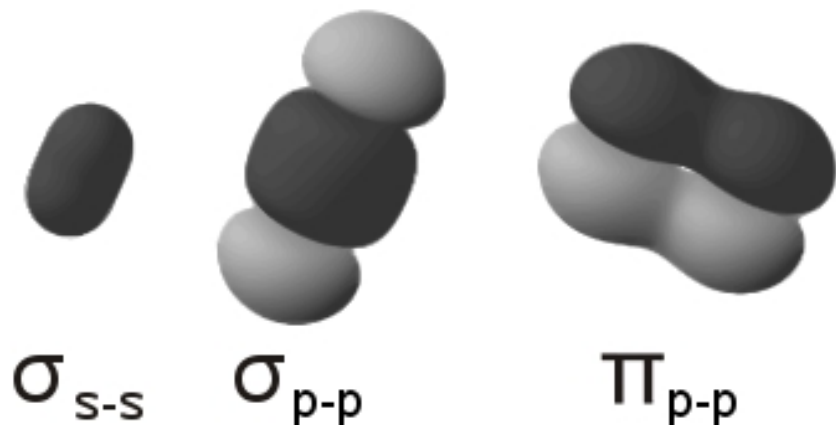
- **Spin:** is a quantum number that parameterizes the intrinsic angular momentum of a given particle (each AO has 2 electrons)

MOLECULAR ORBITALS

● Molecular orbitals: a linear combination of atomic orbitals (LCAO)

- Occupied/Virtual
- LUMO - lowest unoccupied MO
- HOMO – highest occupied MO
- Bonding/Antibonding/Non-bonding
- σ and π -symmetry

$$\phi = \sum_{i=1}^N a_i \varphi_i$$



C_2H_2 orbitals

BASIS SET

● A basis set in quantum chemistry is a set of functions to describe the molecular orbitals, which are expanded as a linear combination of such functions with the weights or coefficients to be determined.

$$\phi = \sum_{i=1}^N a_i \varphi_i$$

Depending on the basis set the methods are classified into:

- Hartree-Fock methods
- Density functional theory
- Semiempirical methods
- Different approaches for electronic correlation

HARTREE-FOCK METHOD

- Hamiltonian can be decomposed into one-electron Hamiltonians:

$$H = \sum_{i=1}^N h_i$$

$$h_i = -\frac{1}{2} \nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}} \quad - \quad \text{one-electronic Hamiltonian}$$

- Solution is a Slater determinant:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \quad \chi_i \quad - \quad \text{Spin-orbital}$$

HARTREE-FOCK METHODS CLASSIFICATION

- Number of electrons taken into account
- Number of Slater determinants
- Number of electronic states
- If $\{a_i\}$ are known

$$\phi = \sum_{i=1}^N a_i \varphi_i$$

$$\chi(\vec{r}) = (x - x_0)^l (y - y_0)^m (z - z_0)^n e^{-\alpha|\vec{r} - \vec{r}_0|} \quad - \quad \text{Slater type orbital (STO)}$$

- Very difficult to calculate integrals (1, 2, 4-electron) with STO

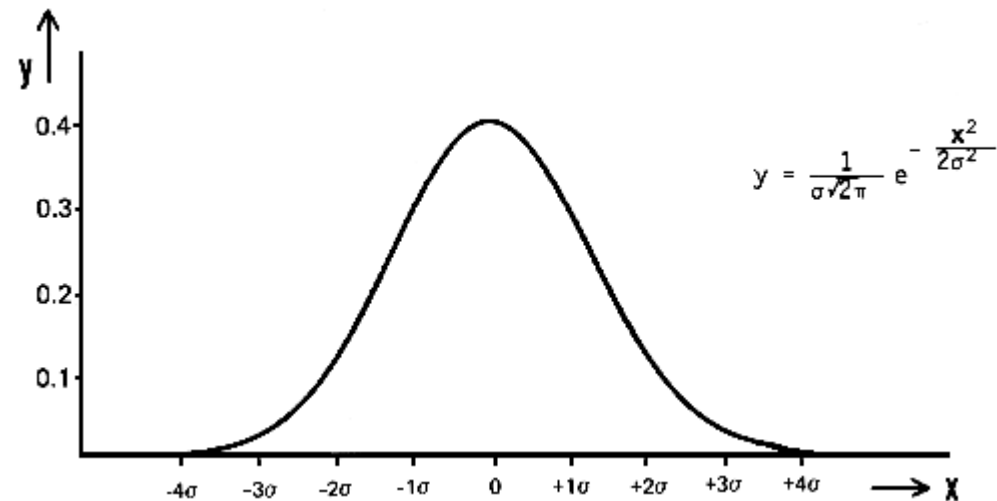
HF METHODS: GAUSSIAN APPROXIMATION

➤ Gaussian function:

$$\chi(\vec{r}) = (x - x_0)^l (y - y_0)^m (z - z_0)^n e^{-\alpha|\vec{r} - \vec{r}_0|^2}$$

- Integrals are much simpler
- More functions than STO
- More functions to represent single points
- Use of contracted functions
- Fall off more rapidly at long range

$$\chi(\vec{r})_{STO} = \sum_1^N c_i \chi(\vec{r})_{Gauss, i}$$



STANDARD BASIS SETS (POPLE BASIS SETS)

- **Minimal: each atom, just enough Gaussian functions:**
 - H: 1s
 - C: 1s, 2s, 2px, 2py, 2pz
 - Cl: 1s, 2s, 2px, 2py, 2pz, 3s, 3px, 3py, 3pz
- **STO-3G: 3 Gaussian primitives per function**
- **Split valence (double-zeta)**
 - Two sizes of contracted function for each atomic orbital
 - H: 1s, 1s'
 - C: 1s, 1s', 2s, 2s', 2px, 2px', 2py, 2py', 2pz, 2pz'
- **6-31G - example**
 - 6 primitive Gaussian in one contracted *core* function
 - 2 contracted functions in the *valence region*: of 3 and 1 primitive Gaussians

STANDARD BASIS SETS: POLARIZATION

- Shift of the symmetry of an AO in relation to nucleus (f.i. bond)

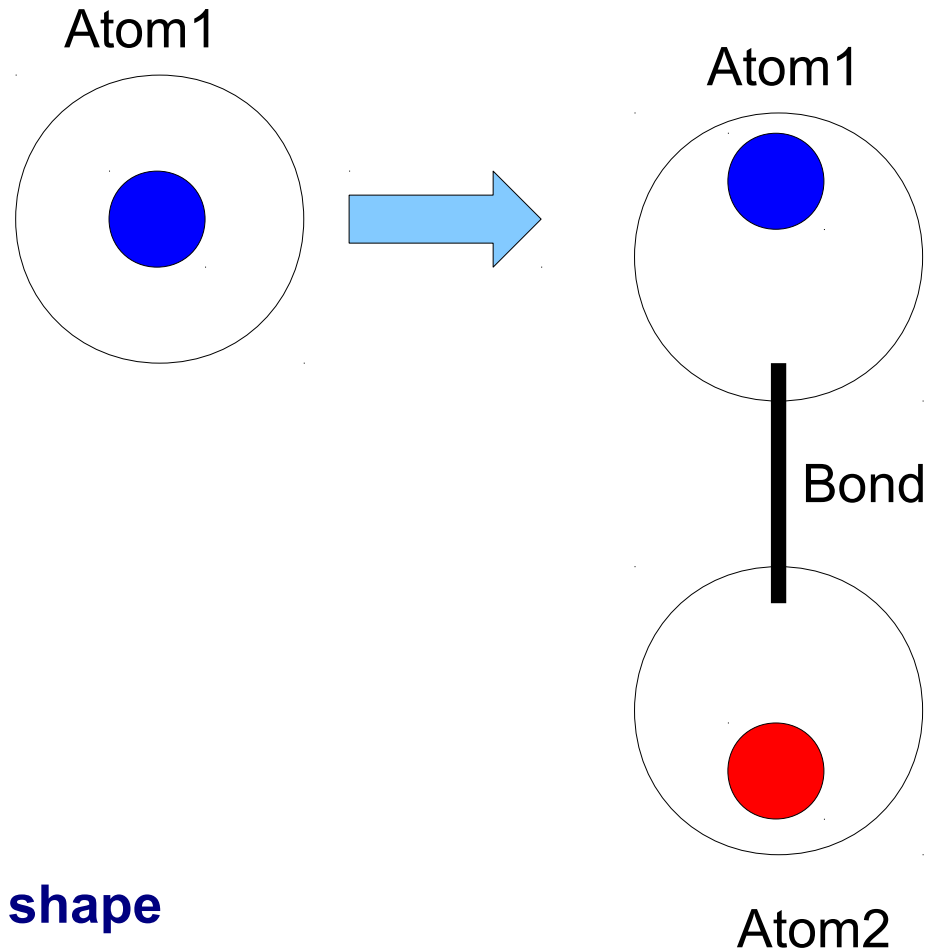
- Higher angular momentum:

- p on H
- d on C
- f on Fe

- Allow orbitals to change shape

- Examples:

- 6-31G (d) = 6-31G*
- 6-31G (d,p) = 6-31G**
- 6-31G (3d, 2p) more flexible angular shape



STANDARD BASIS SETS: DIFFUSE FUNCTIONS

- Represent shallow AO for better approximation of AO 'tails', are important for bigger molecular systems

- Improve tail of orbitals:

- 6-31+G: diffuse functions are added on heavy atoms
- 6-31++G: diffuse functions are added on heavy atoms and hydrogens

- Triple-split valence:

- 6-311G
- More flexible shape

- Other examples:

- 6-311++G (3df, 3pd)

...

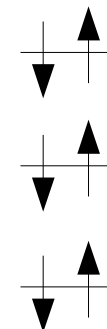
- Correlation consistent methods: cc-pVDZ, cc-pVTZ, aug-cc-pVDZ etc.

RESTRICTED/UNRESTRICTED; OPEN/CLOSED

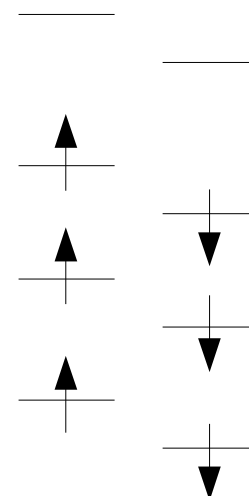
SHELL HF METHODS

➤ Restricted/unrestricted

- α - and β -spins have the same energy or different



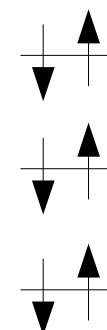
RHF



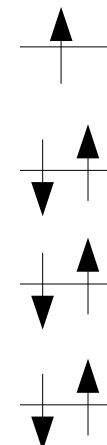
UHF

➤ Open/Closed Shell:

- all orbitals are doubly occupied or not



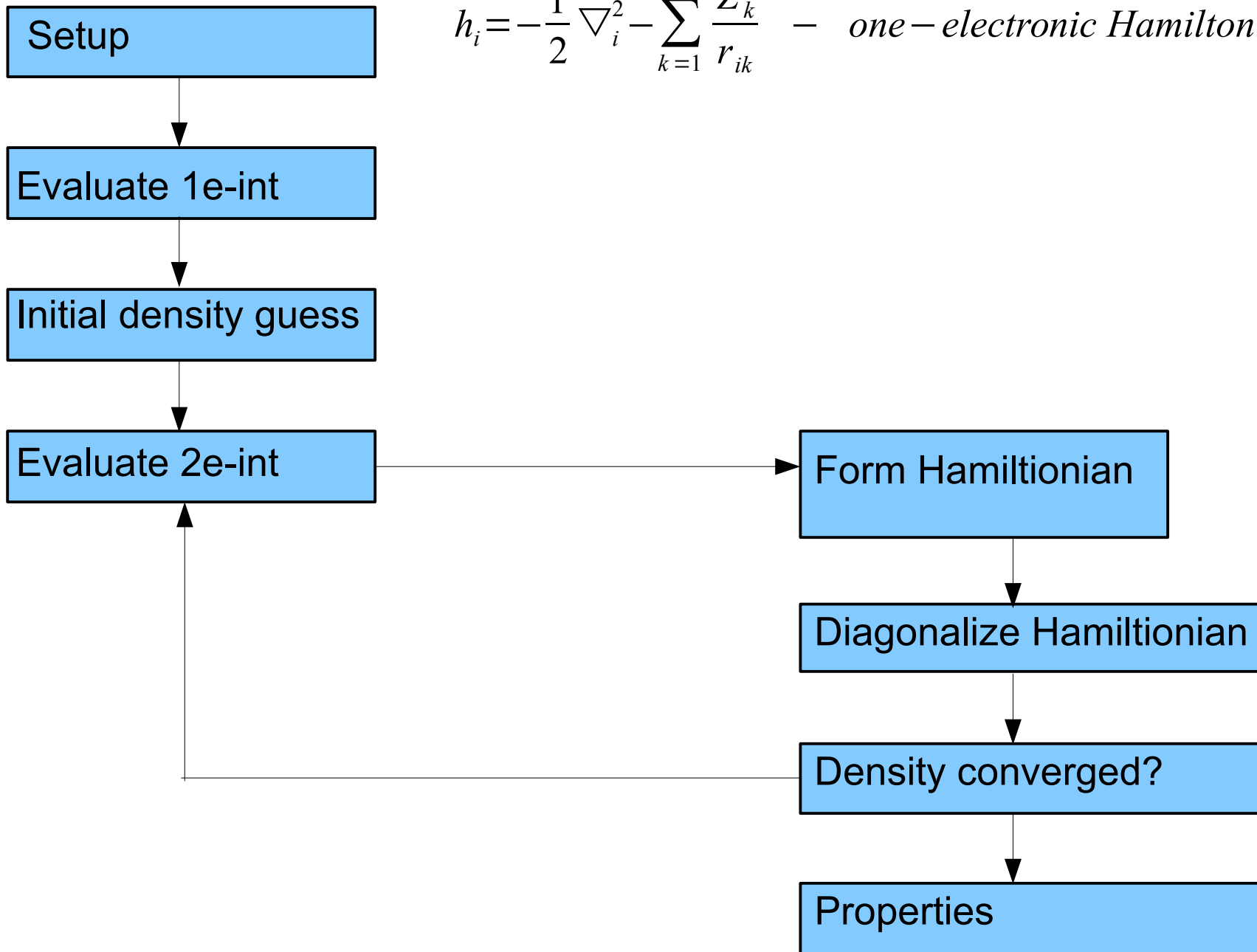
Closed Shell



Open Shell

HF WORKFLOW

$$h_i = -\frac{1}{2} \nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}} \quad - \quad \text{one-electronic Hamiltonian}$$



ELECTRONIC CORRELATION METHODS

- **Electronic correlation: instant interaction between electrons (no in HF!)**

$$h_i = -\frac{1}{2} \nabla_i^2 - \sum_{k=1}^M \frac{Z_k}{r_{ik}} + V_i\{j\} \quad - \quad \text{one-electronic Hamiltonian with correlation}$$

- **Configuration interaction: variational problem for the solution without correlation**
- **Møller-Plesset perturbation theory (MP2, MP3, MP4 ...):**
 $H = F + (H-F)$; (H-F)- perturbation
- **Multiconfigurational self-consistent field**

SEMIEMPIRICAL METHODS

- **HF formalism**
- **Use of empirical parameters (already calculated integrals)**
- **Applied for big systems or as a starting approximation**
- **Faster than HF**
- **Example: Complete Neglect of Differential Overlap (CNDO):**
 - **Only the electrons at the atomic open shells are taken into account.**
 - **The electrons at the closed shells are localized on the point nuclei.**
 - **Basis set corresponds to the occupied states of the atomic orbitals.**
 - **All integrals containing ϕ_i^* and ϕ_j are considered to be zero if $i \neq j$.**
- **Widely used methods: Extended Hückel method, MINDO, MNDO, AM1, PM3, CNDO/2, INDO, NDDO, RM1, SAM1, SINDO, ZINDO**

DENSITY FUNCTIONAL THEORY

- The properties of a many-electron system can be determined by using functionals, which in this case is the spatially dependent on electron density and not wave functions. DFT uses variational principle.

$$H \psi = [T + V + U] \psi ; \quad T\text{- kinetic} ; \quad V\text{- potential} ; \quad U\text{- electron-electron interaction}$$

$$\psi = \psi(\rho) ; \quad \rho \text{ - charge density}$$

$$V[\rho] = \int V(\vec{r}) \rho(\vec{r}) d^3 r ; \quad E[\rho] = T[\rho] + U[\rho] + \int V(\vec{r}) \rho(\vec{r}) d^3 r$$

$$V[\rho]_s = V(\vec{r}) + \int \frac{e^2 \rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3 r' + V_{xc}(\rho(\vec{r})) ; \quad V[\rho]_s \text{ - effective single particle potential}$$

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_s(\vec{r}) \right] \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r}) \quad V[\rho]_{xc} \text{ - exchange-correlation potential}$$

- V_{xc} determines the method (2 components)
- Widely used methods: PBE, BLYP, B3LYP, PW91, M06HF ...

QM METHODS HIERARCHY

Schrödinger equation

MP2, MP4
perturbation theory

Scaling $\geq N^5$

Configuration
interaction

Scaling $\geq N^5$

DFT

Scaling $\geq N^4$

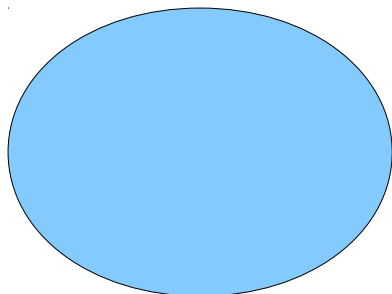
HF

Scaling = N^4

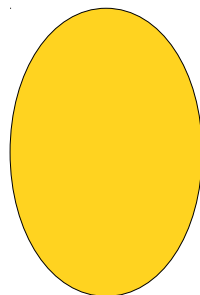
Semiempirical
methods

Scaling = N^3

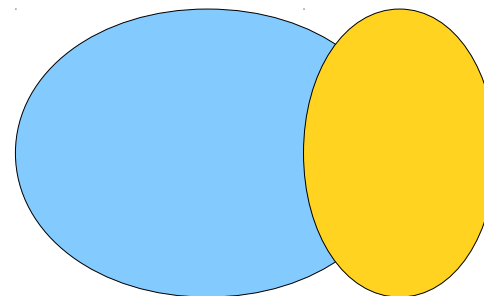
BASIS SET SUPERPOSITION ERROR (BSSE)



Molecule A, basis A



Molecule B, basis B



Complex AB: basis AB

➤ **Basis AB > Basis A + Basis B**

➤ $\Delta\Delta G_{AB} = \Delta G_{AB \cap AB} - \Delta G_{A \cap A} - \Delta G_{B \cap B} > \Delta\Delta G_{AB \text{ exp}}$

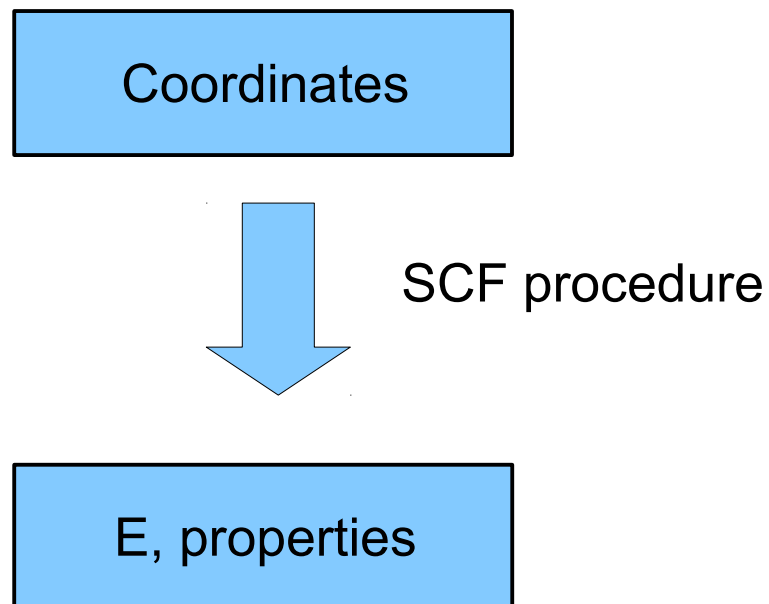
➤ **Counterpoise correction:**

$$\Delta\Delta G_{AB} = \Delta G_{AB \cap AB} - \Delta G_{A \cap AB} - \Delta G_{B \cap AB}$$

$$\Delta\Delta G_{AB} = \Delta G_{AB \cap AB} - \Delta G_{A \cap A} - \Delta G_{B \cap B} + (\Delta G_{A \cap A} - \Delta G_{A \cap AB}) + (\Delta G_{B \cap B} - \Delta G_{B \cap AB})$$

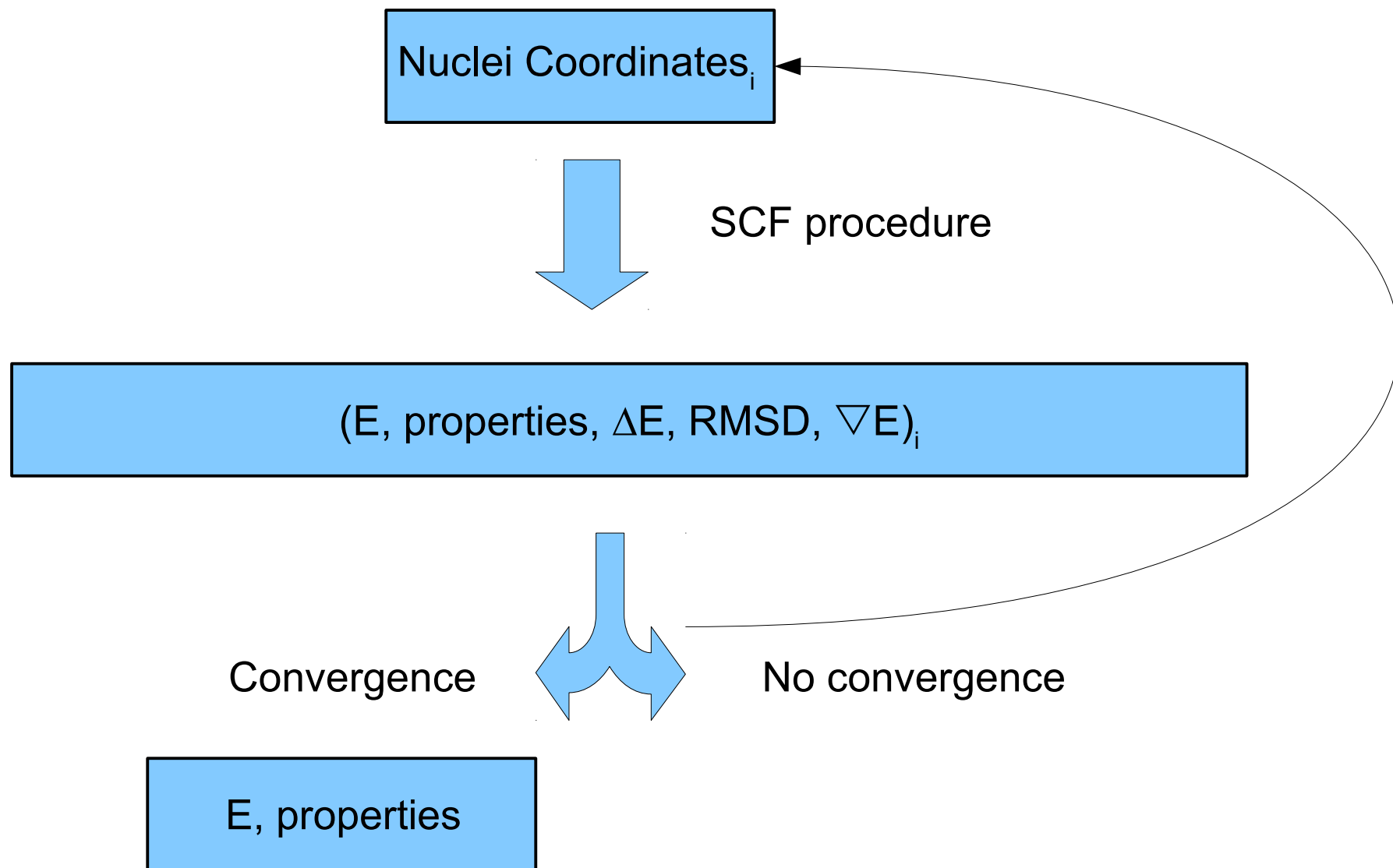
➤ **More calculations than just A, B, AB**

SINGLE POINT CALCULATIONS



- **SP calculations: calculations of the energy and other properties for the given configuration of atomic nuclei**

GEOMETRY OPTIMIZATION



Typical calculation: **Geom. Opt. (Basis1) / SP (Basis 2); Basis 2 > Basis 1**

TRANSITIONAL STATE AND REACTION PATH

➤ Search not for a state with the minimal energy

➤ Input:

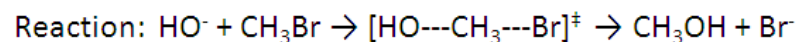
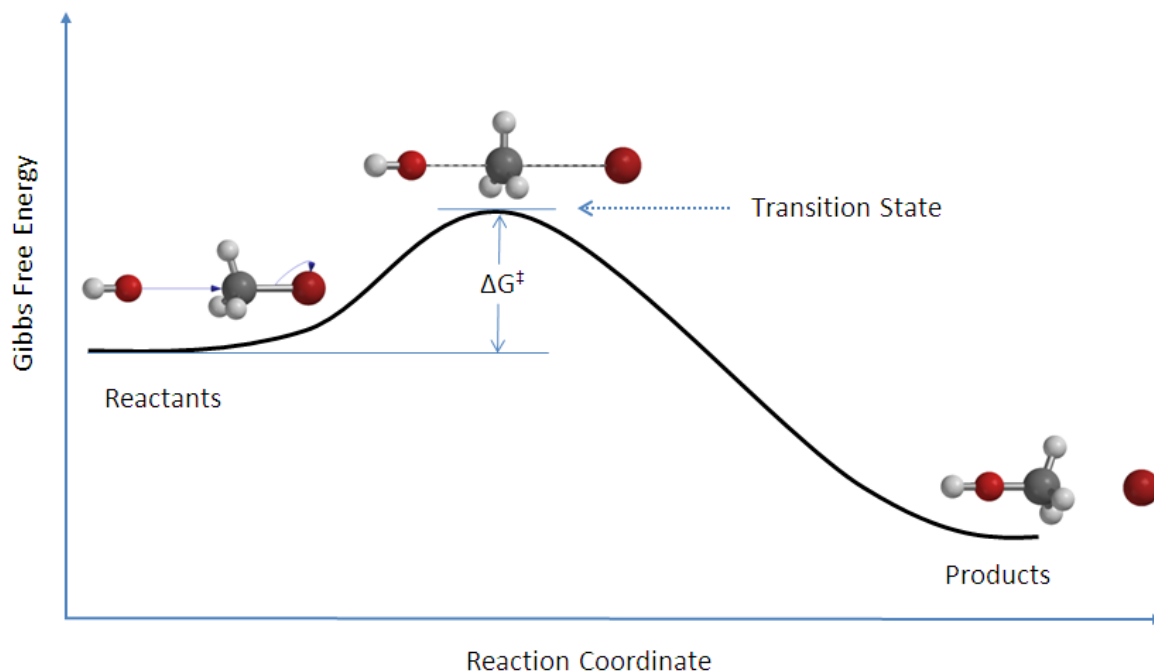
- Reactants

- Products

- * Transitional state guess

- Reaction path step size

➤ Output: **G** (reaction coordinate)



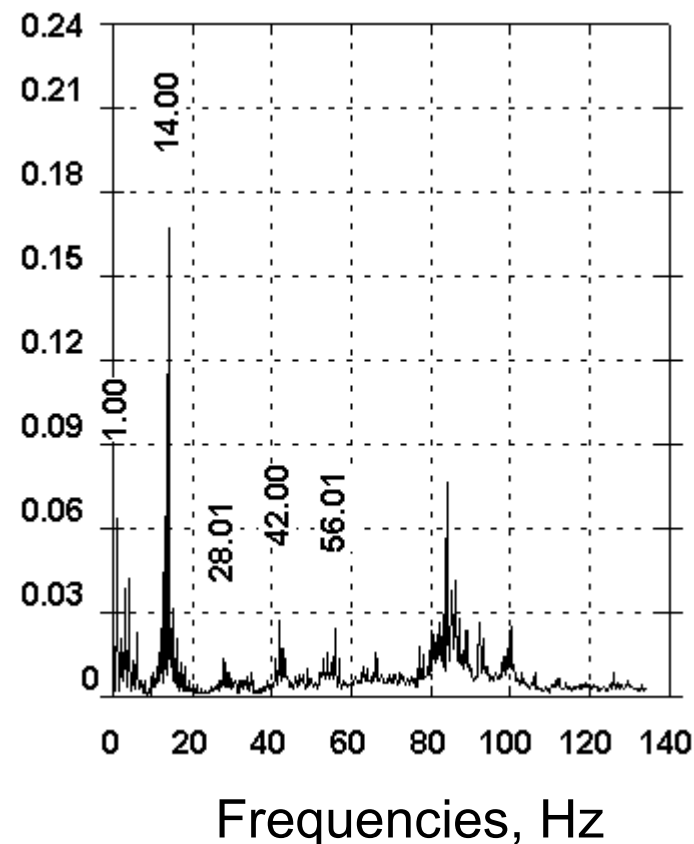
QM AND SPECTROSCOPY

➤ Vibrational frequencies:

- Harmonic vibrational frequencies (from Hessian)
- IR, Raman, NMR spectroscopy
- Anharmonic frequencies

➤ Thermochemistry:

- Partition functions
- Heat capacities
- Absolute entropies
- Energy (by finite temperatures)
- Enthalphy
- Gibbs free energy



VIBRATIONAL ANALYSIS

Mode #	Freq	Infrared	Raman Activity	Depolar-P	Depolar-U
1	255.70	40.6805	0.6059	0.7500	0.8571
2	307.77	174.5255	5.6131	0.7500	0.8571
3	445.38	17.2177	0.5229	0.6012	0.7509
4	912.50	0.0408	0.0066	0.7500	0.8571
5	970.05	14.8655	12.7895	0.2828	0.4409
6	1122.09	93.3675	3.9646	0.4913	0.6589
7	1194.44	4.8695	7.9849	0.6057	0.7544
8	1294.60	8.2303	3.9896	0.7500	0.8571
9	1360.85	90.8538	8.9361	0.5058	0.6718
10	1428.54	0.0794	16.0806	0.7500	0.8571
11	1575.35	0.4018	0.8723	0.4746	0.6437
12	1591.84	17.8479	4.7103	0.4962	0.6633
13	1641.48	4.8304	20.8769	0.7500	0.8571
14	1657.99	5.2437	25.7966	0.7500	0.8571
15	1694.16	2.0041	8.6101	0.7387	0.8497
16	3179.94	61.7585	76.7389	0.2508	0.4011
17	3201.51	21.2194	165.6530	0.0217	0.0424
18	3214.84	66.0077	101.7054	0.7500	0.8571
19	3272.33	41.6335	63.1617	0.7238	0.8398
20	3291.91	48.0026	43.7615	0.7500	0.8571
21	4030.95	23.6374	121.7655	0.3178	0.4823

QM/MM, ONIOM, AB INITIO MD

➤ QM/MM

➤ ONIOM: “an onion-like method”

- Linking atoms between layers
- “REAL” - whole system
- “MODEL” - High level modelled system
- $\Delta E(\text{HIGH}, \text{REAL}) \approx \Delta E(\text{ONIOM})$
- 2 layers example:

$$E_{\text{ONIOM}} = E_{\text{LOW}, \text{MODEL}} + \underbrace{(E_{\text{LOW}, \text{REAL}} - E_{\text{LOW}, \text{MODEL}})}_{\text{Size}} + \underbrace{(E_{\text{HIGH}, \text{MODEL}} - E_{\text{LOW}, \text{MODEL}})}_{\text{Level}}$$

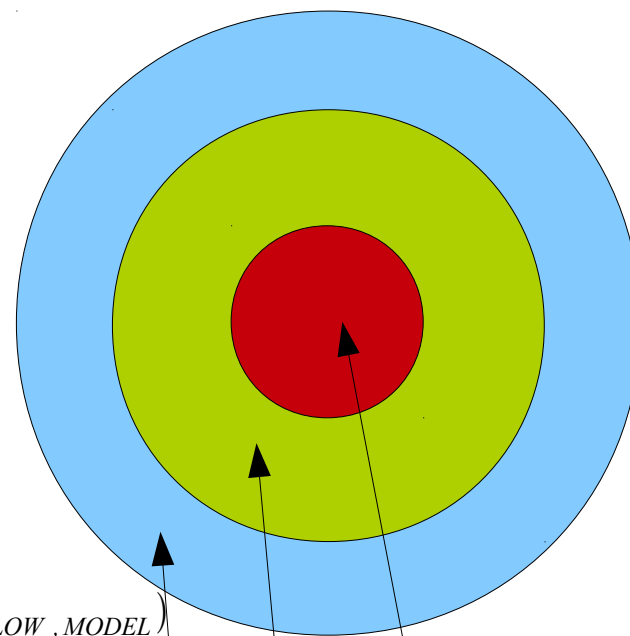
Size

Level

$$E_{\text{ONIOM}} = E_{\text{LOW}, \text{REAL}} + E_{\text{HIGH}, \text{MODEL}} - E_{\text{LOW}, \text{MODEL}}$$

- Only relative energies have meaning
- Different steps for integrations for layers

➤ Car-Parinello MD (CPMD), ~ 10 ps



First layer:

**Bond formation/breaking;
High level method**

Second layer:

**Electronic effect on the 1st
layer; Medium level method**

Third layer:

**Environmental effects on the 1st
layer; Low level method**

SOLVENT EFFECTS

➤ Structures

➤ Energies:

- Isomerization

- Reaction

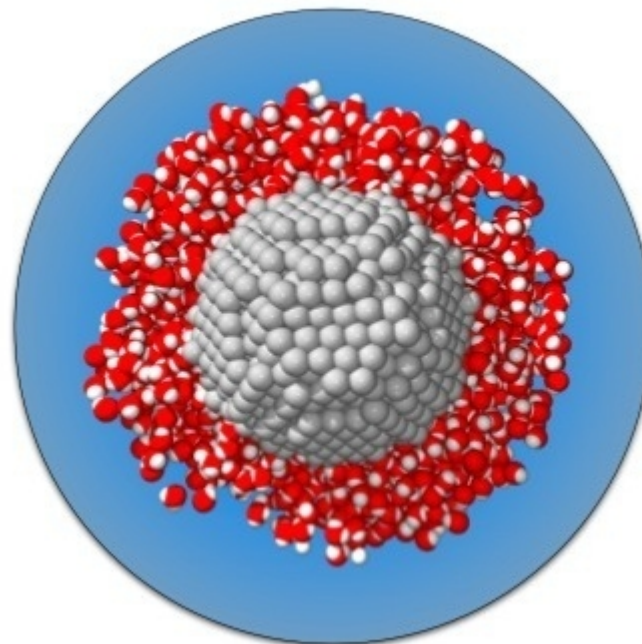
- Bond

- Activation

➤ Spectra:

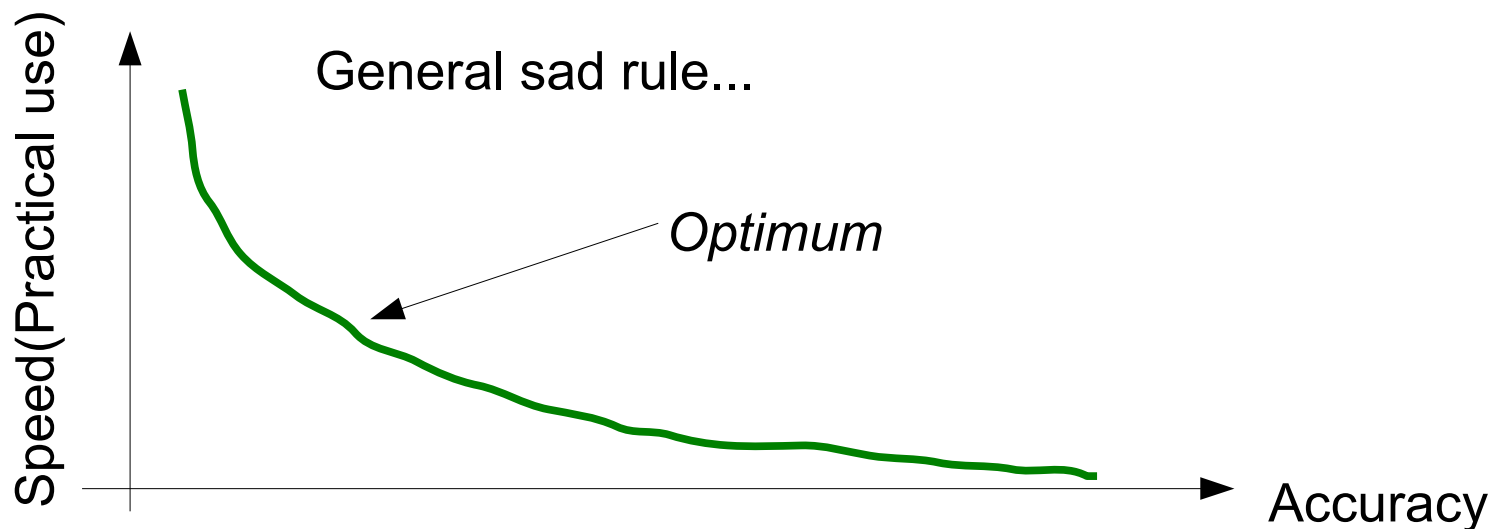
- Vibrational

- Electronic



REQUIREMENTS FOR SOLVENT MODELS

- **Well-defined chemical models:**
 - Results determined by identities of molecules and solvent
 - Systematically calibrated
- **Built on good gas-phase methods**
- **Applicable to a wide range of systems**
 - Few or no empirical parameters
 - Efficient
 - Applicable in principle to any molecule or collection of molecules



SOLVENT MODELS

➤ Continuum models

$$\delta G_{solv} = \delta G_{el} + \delta G_{nonel}$$

- Solvent is uniform dielectric
- Solvent is polarizable
- Resulting electric field
- Size of cavity is considered
- Dispersion and repulsion are added separately or ignored
- Solute interaction with individual solvent molecules is neglected

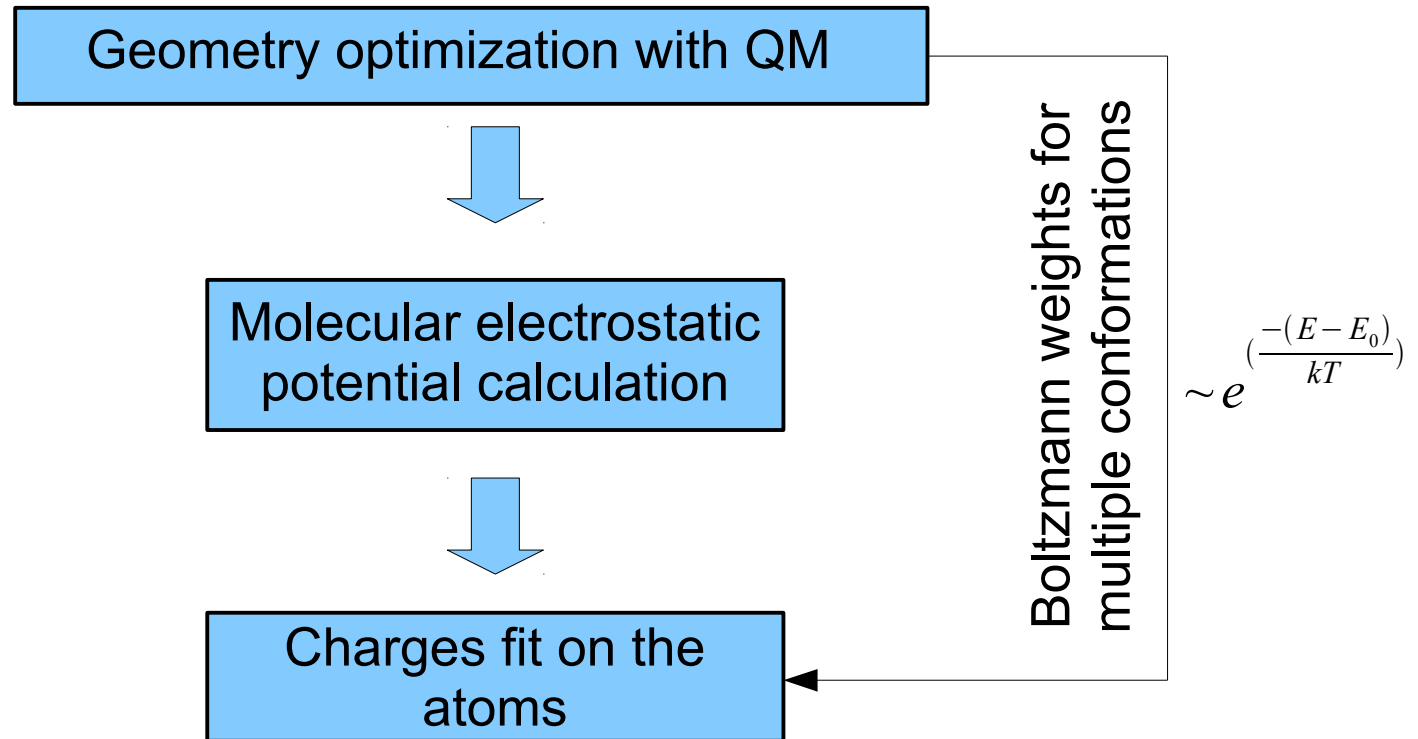
➤ Molecular Mechanics:

- Pure mechanics
- Mechanics with fixed QM solute (gas phase or continuum)
- Mixed QM/MM

➤ Statistical models

CHARGE DERIVATION PROCEDURE FOR MD

- (R)ESP charges (compatible with many ff): (restrained) electrostatic potential charges

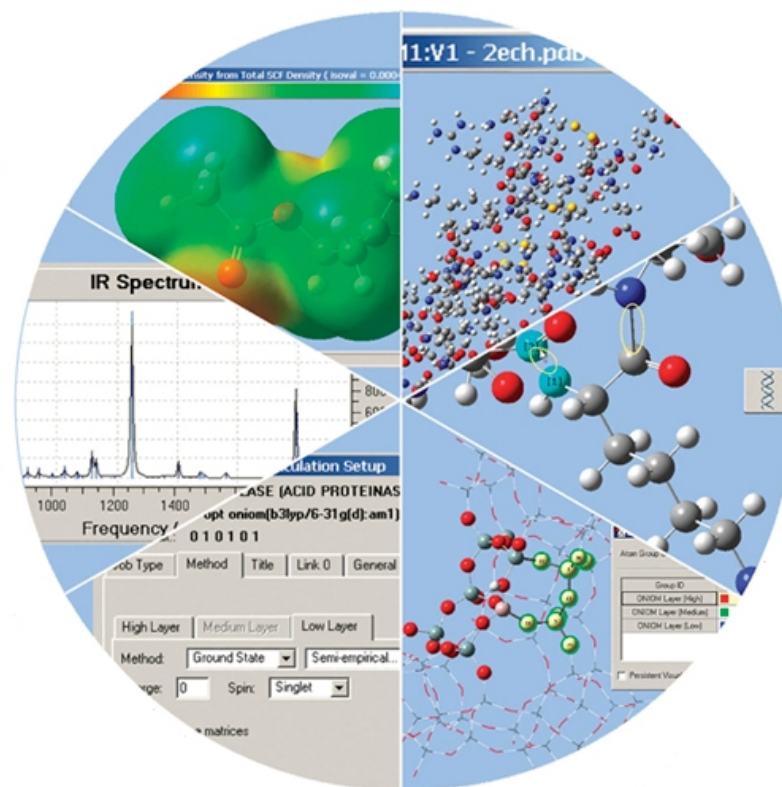


➤ Example: AA side chain parameterization:

- N conformations (angles are sampled with 1° step)
- N Boltzmann weights
- Weighted charges

QUANTUM CHEMISTRY SOFTWARE

- GAUSSIAN
- GAMESS (Open Source)
- JAGUAR
- MOPAC (old versions: Open Source)
- HyperChem
- CPMD (Ab initio dynamics, Open Source)
- R.E.D. III (Open Source, charges derivation)

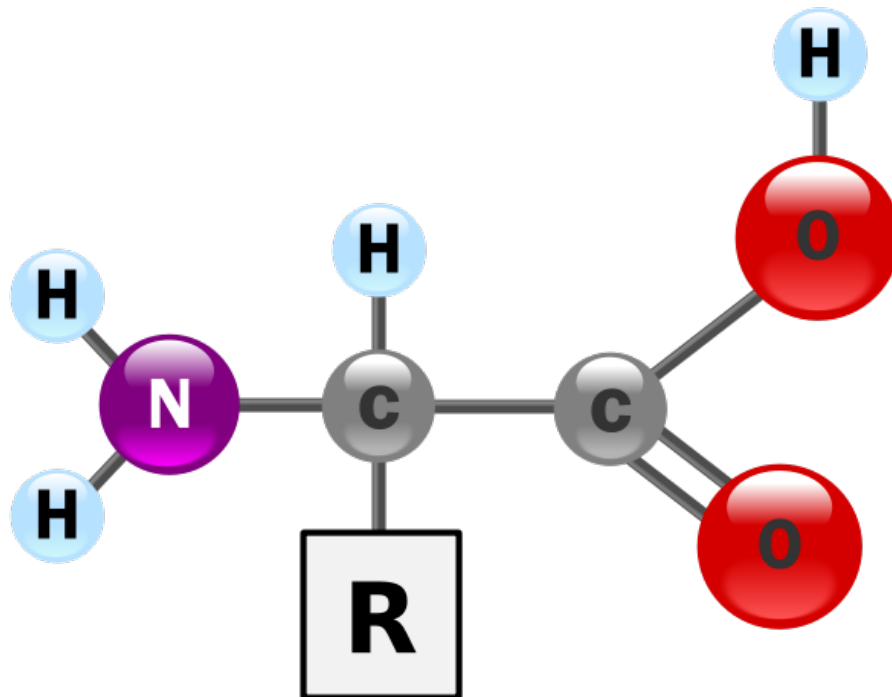


CASE STUDY 1: FLUORINATED AMINO ACIDS

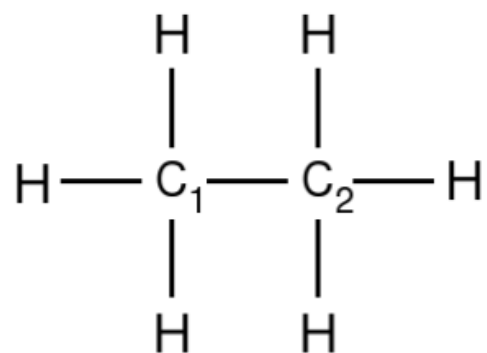
Goals:



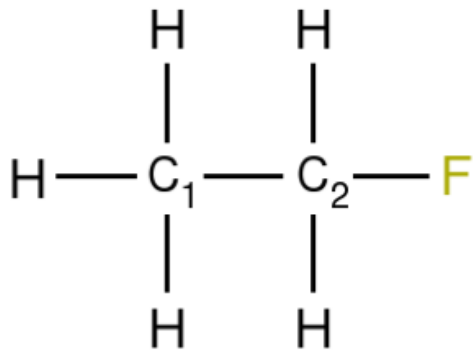
- To characterize changes in aliphatic groups upon fluorination
- Hydrogen-bonds analysis of fluorinated groups



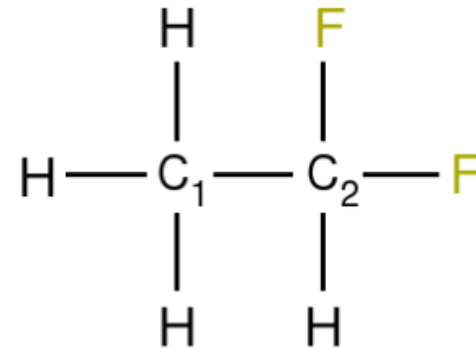
MODEL: FLUORINATED ETHANE DRIVATIVES



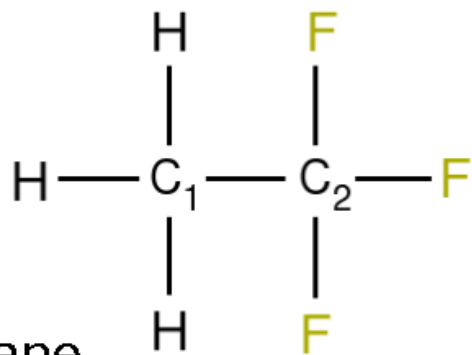
Ethane



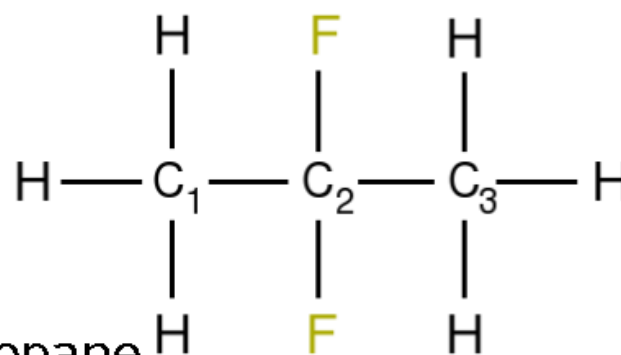
MEthane



DFEthane



TFEthane



2,2DFPropane

Geometry optimization in HF, MP2, BLYP, B3LYP (6-311G++)**

FLUORINATION: GEOMETRY AND CHARGE IMPACT

- Bond lengths decrease $\sim 0.01 \text{ \AA}$ (C-C, C-H), $\sim 0.1 \text{ \AA}$ (C-F)
- Charge redistribution:

$C_F \uparrow \sim 1 e$

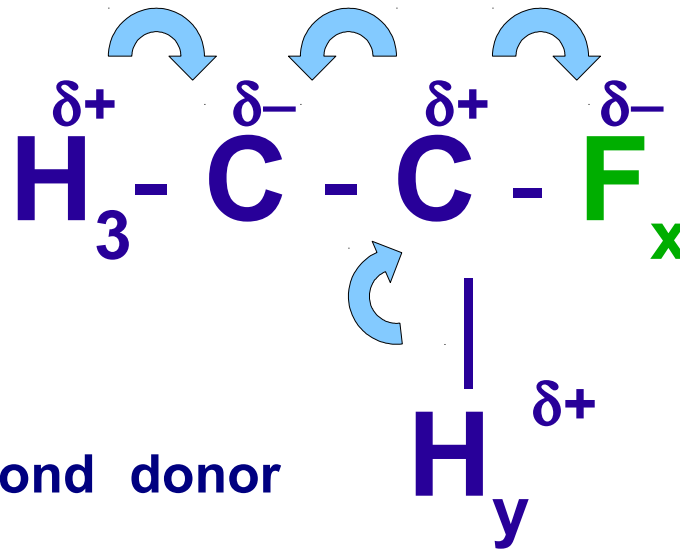
$C_H \downarrow \sim 0.1 e$

$H \uparrow \sim 0.1 e$

$q(H)$ become comparable

with $q(H)$ in amino acid H-bond donor

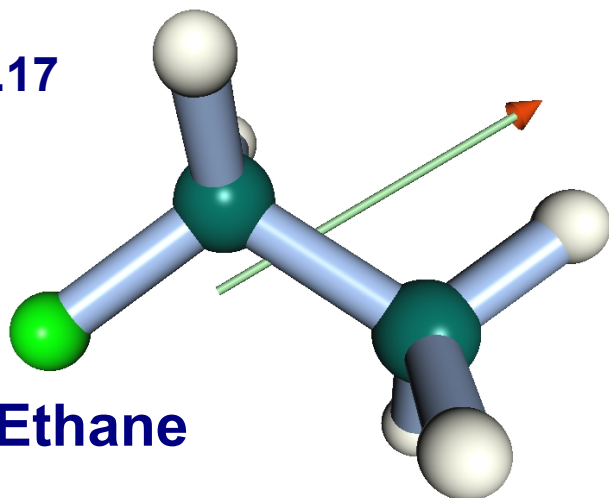
groups



FLUORINATION: DIPOLE MOMENT

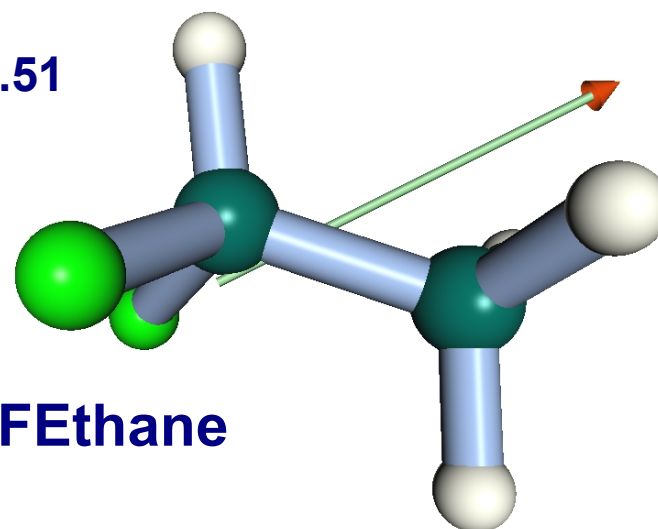
$$\vec{D} = \sum_{i=1}^N q_i \vec{r}_i$$

D=2.17



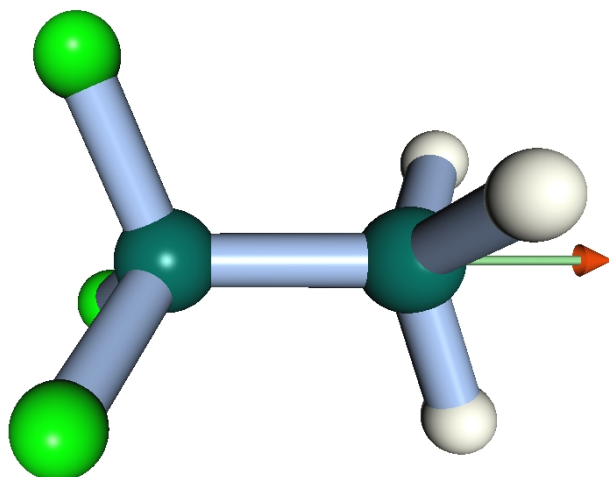
MFethane

D=2.51



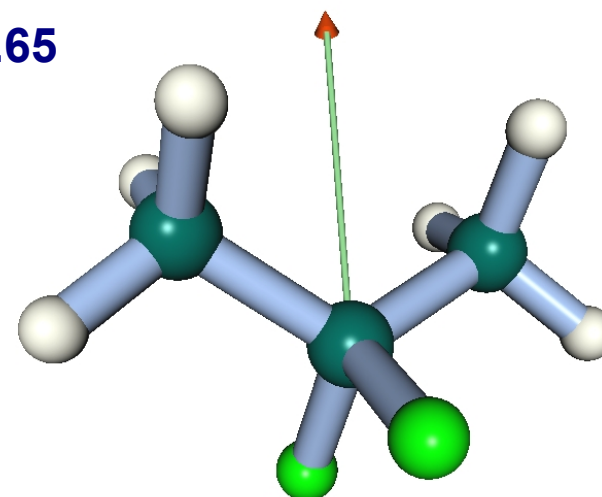
DFethane

D=2.58



TFethane

D=2.65



2,2 DFPropane

B3LYP(6-311G**++)

H-BONDS CHARACTERISTICS

➤ Energy

➤ Length $d(\text{H-A})$

➤ Angle (D-H-A)

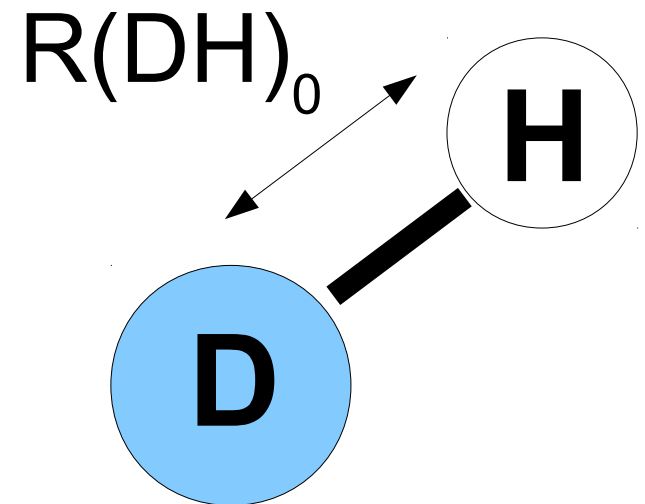
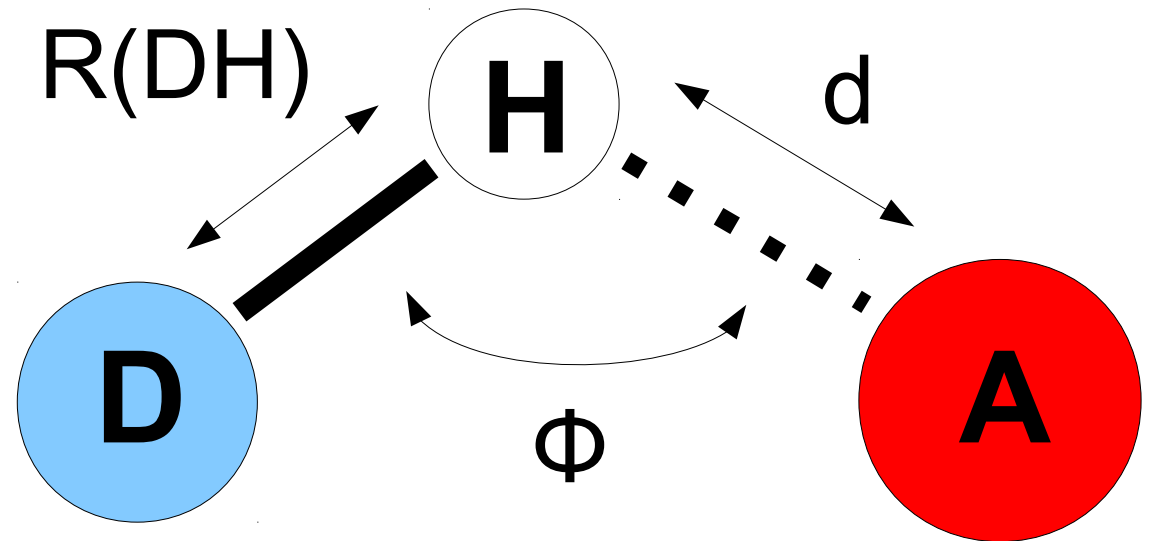
➤ D-H relative shift: $\frac{R(DH) - R(DH)_0}{R(DH)_0}$

➤ Charge transfers:

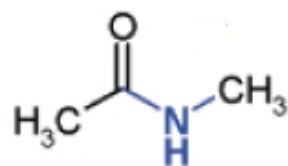
- Acceptor

$$\delta q(A/H) = q(A/H) - q(A/H)_0$$

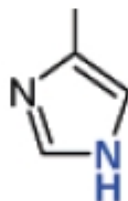
- Hydrogen



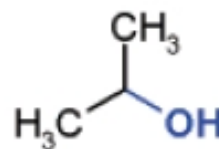
H-BONDS: USED DONORS AND ACCEPTORS



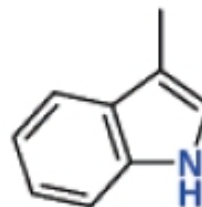
Amide-D



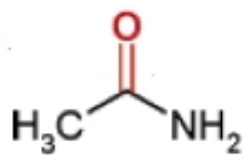
Imidazole-D



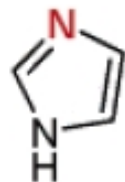
Hydroxyl-D



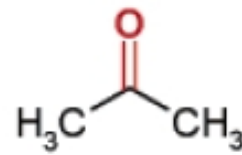
Indole-D



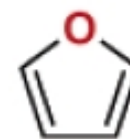
Amide-A



Imidazole-A



Ketone-A

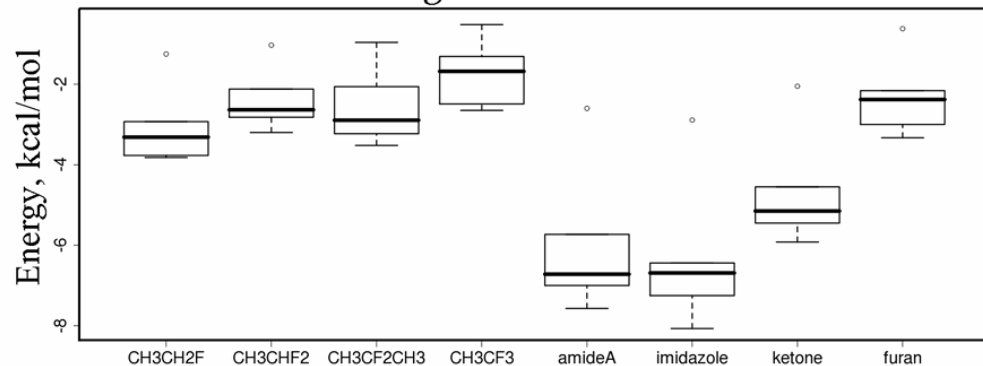


Furan-A

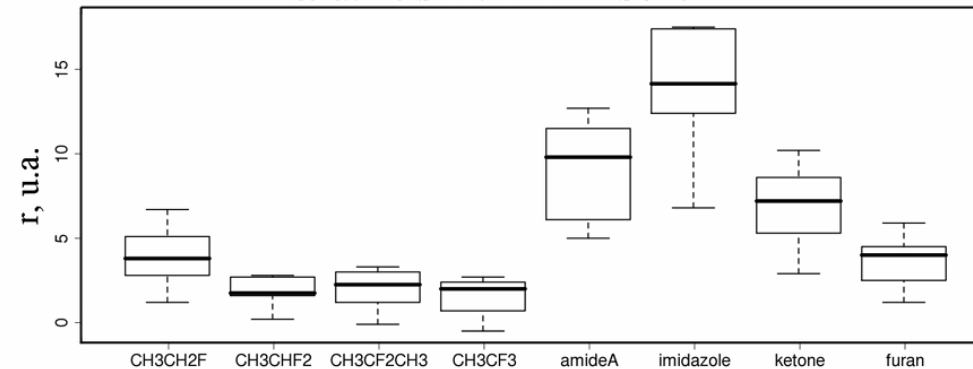
- H HB donor
- HB acceptor

FLUORINATED GROUPS: H-BONDS

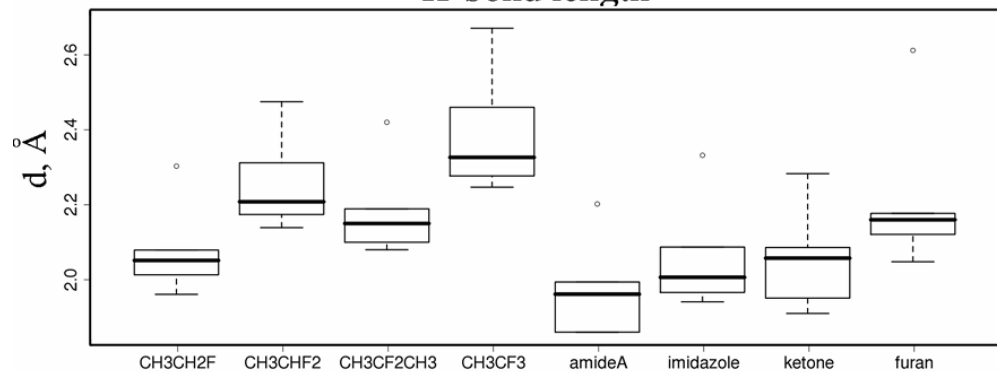
H-bond energies with BSSE correction



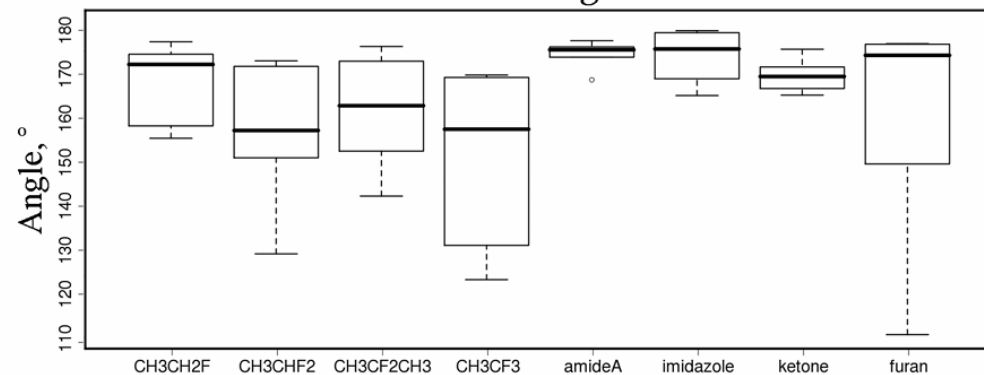
Relative shift in D-H bond



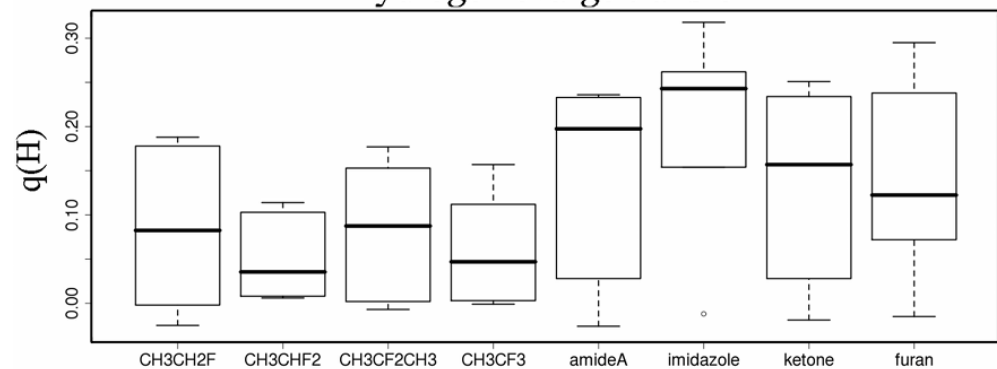
H-bond length



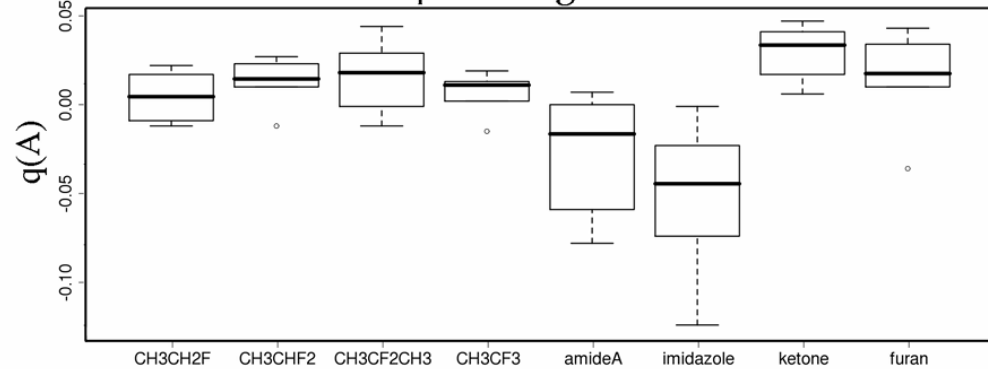
H-bond angle



Hydrogen charge transfer

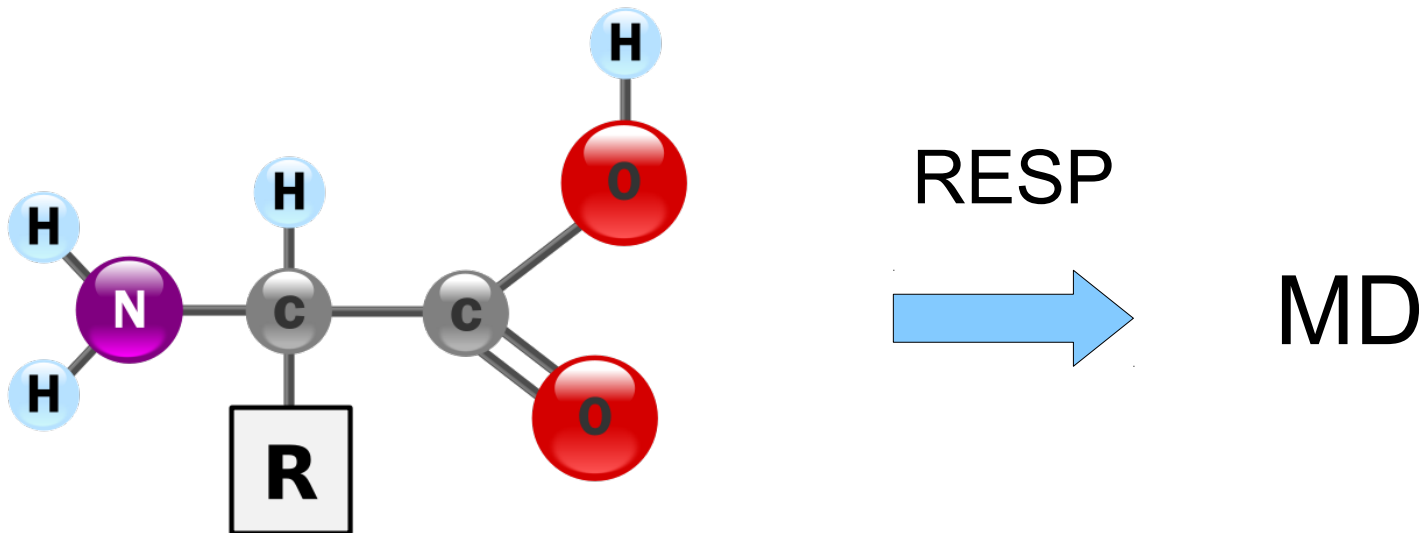


Acceptor charge transfer



CONCLUSIONS

- Geometry changes of ethane fluorinated derivatives have been characterized
- Fluoromethyl groups are weak hydrogen bond donors and acceptors
- F and H atoms in fluoromethyl groups should be considered as potential acceptors and donors of H bonds



CASE STUDY 2: SUGAR RINGS CONFORMATIONS

Goal: conformational analysis of monosaccharide rings
conformations depending on sugar type, sulfation pattern,
environment (adjacent saccharides)

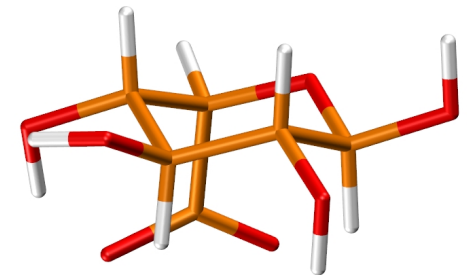


System: Glucuronic/Iduronic acids

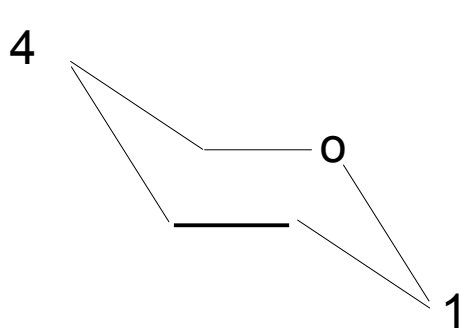
β -D-GlcU



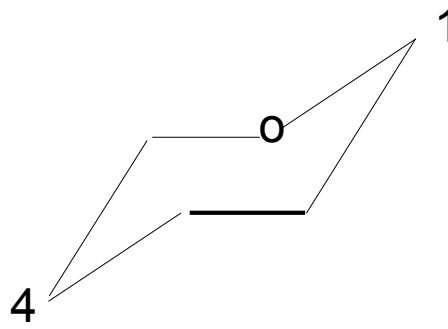
α -L-GlcU



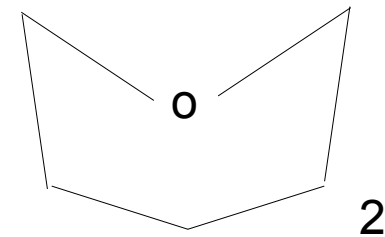
Conformations:



C^4_1 - chair



C^1_4 - chair

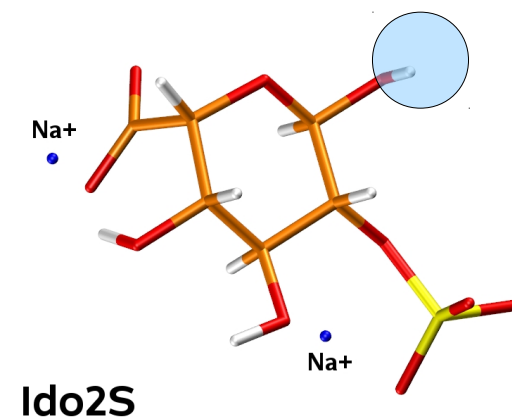
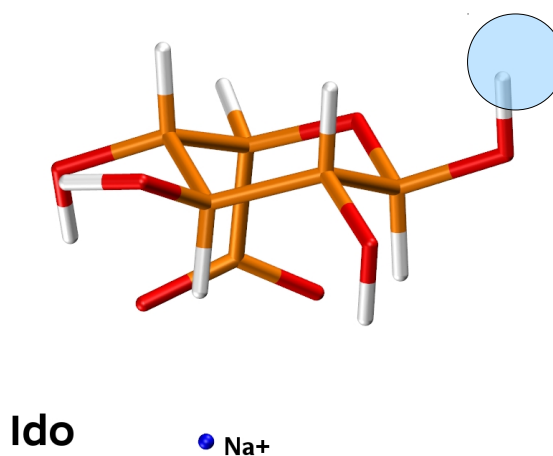
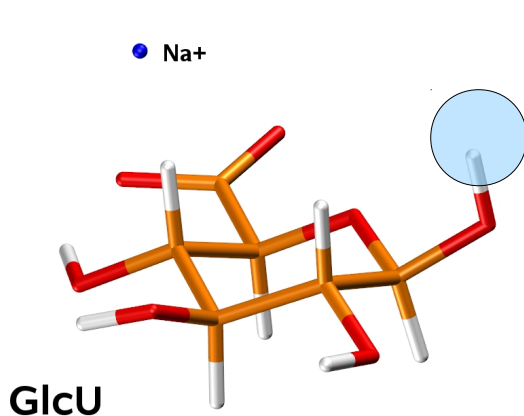


S^2_0 - boat

METHODOLOGY AND RESULTS

➤ Method:

- Counterions addition ($q_{\text{system}}=0$)
- MM Geometry Optimization
- 3-21G Geometry optimization
- B3LYP/6-31+G* Geometry Optimization
- B3LYP/6-311++G** Single Point calculations



$E-E_{\text{min}}$, kcal/mol	GlcU	GlcU-CH ₃	Ido	Ido-CH ₃	Ido2S	Ido2S-CH ₃
C ₁ ⁴ -chair	0.88	0-min	0-min	0-min	0.28	5.78
C ₄ ¹ -chair	0-min	0.70	2.00	4.66	0-min	0-min
S ₀ ² -boat	3.42	5.12	2.33	6.64	2.77	7.18

LECTURE 6: QM BASICS FOR COMPUTATIONAL CHEMISTRY

- Milestones of quantum mechanics
- Schrödinger equation
- Methods in quantum chemistry: HF, DFT, SCF
- Geometry optimization, transitional state
- Spectroscopy, NMR
- QM/MM and ONIOM
- Solvent
- Quantum chemistry software
- Case study 1: fluorinated amino acids
- Case study 2: sugar's ring conformations

