Choice of
Theoretical Method

1

General Considerations

- Resources
  - Software
  - Computer
- "Expense" (Time and Money)
  - Optimization/Single Point Calculations
  - Basis Set Choice
  - Method Choice
- Accuracy
  - Comparison
  - Recommendations

2

Resources

Software

- Methods "packaged" with software
  - See software comparison sheet
- Basis functions "packaged" with software
  - Can download from
    http:www.emsl.pnl.gov:2080/forms/basisform.html
- Available parameters for atom types
- Size of molecule

3

Resources

Computer

- Speed
  - ‣ Fast chips/IO
  - ‣ Parallel processors
- Memory use for optimization algorithm (D = degrees of freedom like bond length, bond angle, etc.)
  - ‣ Scales as D for Conjugate Gradient, Fletcher-Reeves, Polak-Ribiere
  - ‣ Scales as D^2 for Simplex, Powell, quasi-Newton, Fletcher-Powell

4

Expense

Time and Money

- Optimization
  - ‣ t(optimization) $\approx 5 \times D^2 \times t$(single point)
- Basis Set (N = number of orbitals, M = number of atoms)
  - ‣ MM / MD scales as M^2
  - ‣ AM1 / PM3 scales as N^2 to N^4
  - ‣ HF scales as N^2 to N^4
  - ‣ DFT scales as N^3
  - ‣ MP2 scales as N^5
  - ‣ MP4 scales as N^7
  - ‣ Full CI scales as $N!$

5

- Method
  - ‣ For single point calculation of C_6H_6 using cc-pVTZ basis set on Cray SVI (Young):

PM3	11 s	11 Mb
HF	10 min	42 Mb
MP2	32 min	43 Mb
MP4	36 hr	2.1 Gb
G2	37 hr	2.5 Gb
CBS-APNO	63 hr	6.5 Gb
QCISD	10 hr	1.9 Gb
G3	20 hr	1.6 Gb

CBS is Complete Basis Set
QCISD is Quadratic Configuration Interaction

6

Accuracy

References

- J B Foresman and Æ Frisch, *Exploring Chemistry with Electronic Structure Methods*, 2nd ed, Gaussian (1996), ISBN 0-9636769-4-6, pp 147-148 and 157-158.
- D C Young, *Computational Chemistry*, Wiley (2001), ISBN 0-471-33368-9, pp 137-141.
- R Janoschek, *Pure Appl Chem*, **73** (9), 1521-1553 (2001).
- NIST Computational Chemistry Comparison and Benchmark Data Base at
  - ‣ http://srdata.nist.gov/cccbdb

7

Accuracy

Considerations

- Molecule Set
  - ‣ G2
    - – Originally 125 calculations
    - – 1st and 2nd row heavy atoms
  - ‣ Extended G2 sets
  - ‣ G3
    - – 299 calculations from G2/97 set
- Parameters
  - ‣ Mean Absolute Deviation (MAD)
    - – Average of deviations neglecting sign
  - ‣ Standard Deviation
    - – ~2/3 results lie within StdDev of experimental value
  - ‣ Largest Errors
    - – Worst case values

8

Accuracy

Energy Comparision for Basis Functions
[MAD/(kcal mol^{-1}), Foresman/Frisch]

- HF
  - ‣ 93.3 for STO-3G
  - ‣ 58.4 for 3-21G(d)
  - ‣ 51.0 for 6-31G(d)
  - ‣ 46.7 for 6-31+G(d,p)
- B3LYP
  - ‣ 7.9 for 6-31G(d)
  - ‣ 3.9 for 6-31+G(d,p)
  - ‣ 3.1 for 6-311+G(2d,p)

9

Accuracy

Energy Comparision for Standard Methods
[MAD/(kcal mol^{-1}), Young]

- Mechanics
  - *E* not calculated in MM2 / MM3
- Semi-empirical
  - 18.8 for AM1
  - 17.2 for PM3
- *ab initio*
  - 51.0 for HF/6-31G(d)
  - 7.9 for B3LYP/6-31G(d)

10

Accuracy

Energy Comparision for "Mixed" Methods
[MAD/(kcal mol^{-1}), Foresman/Frisch]

- AM1 optimization followed by single point
  - 54.2 for HF/6-31G(d)
  - *18.8 for AM1*
  - 10.5 for B3LYP/6-31G(d)
- B3LYP/6-31G(d) optimization followed by single point
  - *7.9 for B3LYP/6-31G(d)*
  - 4.0 for B3LYP/6-31+G(d,p)
  - 3.2 for B3LYP/6-311+G(2d,p)
  - 2.7 for B3LYP/6-311+G(3df,2df,2p)

11

Accuracy

Observations (Foresman/Frisch)

- Semi-empirical energies are more accurate than HF
- DFT energies are rather insensitive to geometry optimization used
- DFT energies are more accurate using highly diffuse/polarized basis sets
- Additional basis sets are not as important for optimization as for predicting properties

12

Accuracy

Other Comparisions for Standard Methods (Young)

- $\Delta_f H$ [StdDev/(kcal mol^{-1})]
  - 0.5 for MM2
  - 0.6 for MM3
  - 8 for AM1 / PM3
  - 4 for HF/6-31G(d)
- Bond Lengths (rms/Å)
  - 0.01(StdDev) for MM2 / MM3
  - 0.048 for AM1
  - 0.037 for PM3
  - 0.032 for HF/6-31G(d)
  - 0.048 for MP2/6-31G*
  - 0.020 for B3LYP/6-31G(d)

13

- Bond Angles (rms/°)
  - 1.0(StdDev) for MM2 / MM3
  - 3.3 for AM1
  - 3.9 for PM3
  - 1.4 for HF/6-31G(d)
  - 1.5 for MP2/6-31G*
  - 1.4 for B3LYP/6-31G(d)
- Dipole Moments (StdDev/D)
  - 0.1 for MM2
  - 0.07 for MM3
  - 0.5 for AM1
  - 0.6 for PM3
  - 0.2 for HF/6-31G(d)

14

Accuracy

Recommendations

- Use DFT with largest basis set for energy calculations following an optimization using
  - DFT with largest basis set
  - DFT with good basis set
  - HF with good basis set
  - HF with small basis set
  - Semi-empirical methods

15

- Use "Multilevel" Methods
  - G2
    - Initial geometry optimization and zero point vibrational energy (E_1) using HF/6-31G(d)
    - Final geometry optimization at MP2/6-31G(d)
    - Diffuse energy term (E_2) using E[MP4/6-311+G(d,p)] - E[MP4/6-311G(d,p)]
    - Polarization energy term on heavy atoms (E_3) using E[MP4/6-311G(2df,p)] - E[MP4/6-311G(d,p)]
    - Residual correlation effect energy term (E_4) using E[QCISD(T)/6-311G(d)] - E[MP4/6-311G(d)]
    - Additional correction for assumption that 2df and diffuse corrections are additive and for 3rd set of f functions on heavy atom and 2nd set of p functions on H atom (E_5) using E[MP2/6-311G+G(3df,2p)] - E[MP2/6-311G(2df,p)] - E[MP2/6-311+G(d,p)] + E[MP2/6-311G(d,p)]

16

- Empirical estimate of remaining correlation energy (E_6) using $(-0.00481)\times$(# valence electron pairs) - $(0.00019)\times$(# unpaired valence electrons)
- $E = (0.8929)E_1 + E$[MP4/6-311G(d,p)] + $E_2 + E_3 + E_4 + E_5 + E_6$
- MAD for E = 1.1 kcal mol^{-1}, $\Delta_f H$ = 1.6 kcal mol^{-1}
- Several flavors of G2

17

- G3
  - E_2 using E[MP4/6-31+G(d)] - E[MP4/6-31G(d)]
  - E_3 using E[MP4/6-31G(2df,p)] - E[MP4/6-31G(d)]
  - E_4 using E[QCISD(T)/6-31G(d)] - E[MP4/6-31G(d)]
  - E_5 using E[MP2/G3large] - E[MP2/6-31G(2df,p)] - E[MP2/6-31+G(d)] + E[MP2/6-31G(d)]
  - E_6 using $(-0.006386)\times$(# valence electron pairs) - $(0.002977)\times$(# unpaired valence electrons)
  - $E = (0.8929)E_1 + E$[MP4/6-31G(d)] + $E_2 + E_3 + E_4 + E_5 + E_6$
  - t(G3) $\approx t$(G2)/2
  - MAD for $\Delta_f H$ = 0.9 kcal mol^{-1}
  - Several flavors of G3

18

- CBS-Q is similar to G2 using
  - MP2/6-31G(d†) for geometry
  - 6-311++G(2df,2p) for HF and MP2
  - MP4/6-31+G(d,p) and QCISD/6-31+G(d†)] for higher order corrections
  - HF/6-31G(d†) for thermo with 0.918 scale factor
  - Empirical correction for electron correlation
  - Empirical correction for spin contamination
  - t(CBS-Q) $\approx$ t(G3)
  - MAD for E = 1.0 kcal mol^{-1}, $\Delta_f H$ = 1.2 kcal mol^{-1}
  - Several flavors of CBS

19