Quantum Monte Carlo for the Electronic Structure of Atoms and Molecules

Brian Austin
Lester Group, U.C. Berkeley

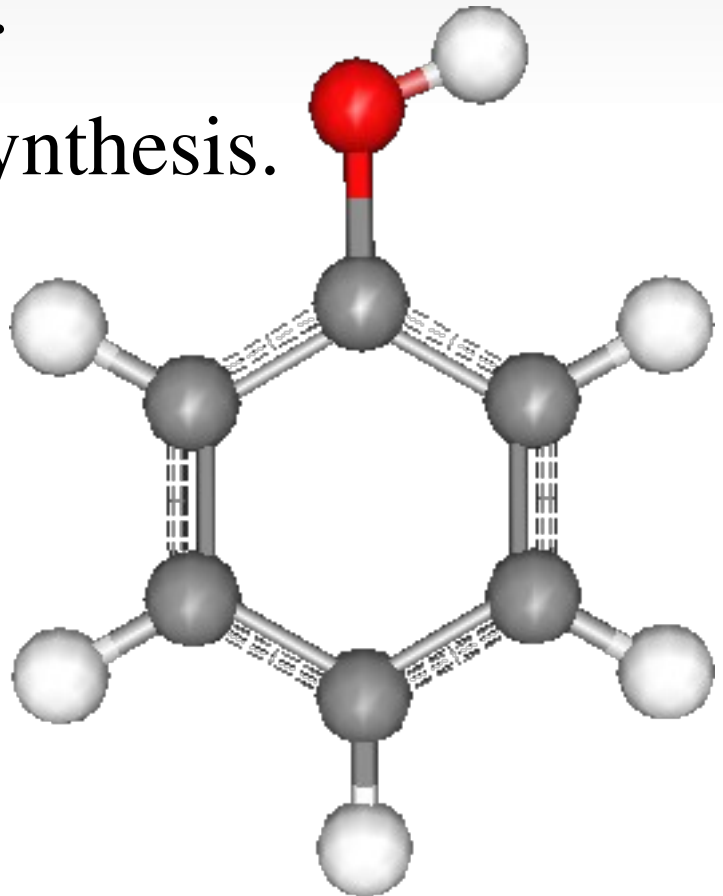
BES Requirements Workshop
Rockville, MD
February 9, 2010

Outline

- Applying QMC to diverse chemical systems
  - Select systems with high interest and impact
  - Phenol: bond dissociation energy
  - Retinal: excitation energy
- Algorithmic details
  - Parallel Strategy
  - Wave function evaluation

O-H Bond Dissociation Energy of Phenol

- Ph-OH $\longrightarrow$ Ph-O$^{\bullet}$ + H$^{\bullet}$ (36 valence electrons)
- Model for oxygenated aromatic species involved in incomplete combustion processes.
- Participates in green plant photosynthesis.
- Broad range of BDEs
  - Experiment: 86.3-90.1 kcal/mol
  - Theory: 86.2-93.4 kcal/mol
  - QMC: 87.6 kcal/mol



Goal: Accurate Thermochemistry by QMC

- Geometry optimization has been elusive in QMC
- Vibrational modes needed for correct themochemistry
- Treat electronic and nuclear motion simultaneously.
  - Avoids Born-Oppenheimer approximation
    - Distinctive feature of QMC
  - Theory is straightforward but needs good wave functions.
  - Sidesteps variance problem associated with forces
  - Correct treatement of vibronic coupling

QMC Thermochemistry of Phenol HPC Requirements

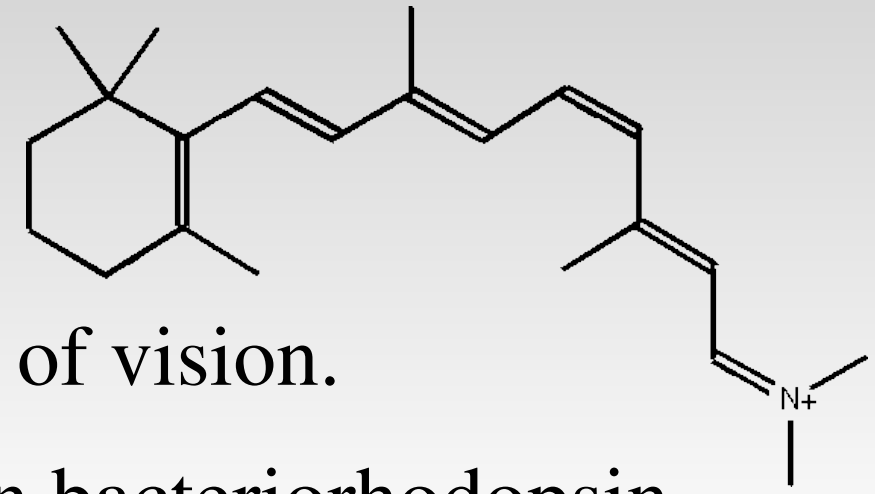
Phenol single point energy

- Calculated on Franklin
- 25,000 core hours
- 4096 cores
- 15 MB memory / core
- 3 GB storage

Phenol thermochemistry

- Experience is limited, cannot estimate CPU time directly
  - Compare to brute-force fit of potential energy surface: ~1000 points
  - 25,000,000 core hours
- Minimal increase in memory requirements.
- 50 GB storage

S_0-S_1 Excitation Energy of Retinal Analogs



- Retinal excitation is intial step of vision.
- Photosynthetic chromophore in bacteriorhodopsin
- Picosecond isomerization; frequent subject of ultrafast spectrocopic experiements
- Large discrepancy between theory and experiment
  - CASPT2 / 6-31G* 2.28 eV
  - Absorbtion spectroscopy 2.03 eV
  - QMC (in progress on Franklin) 1.9(3) eV

Goal: Large Scale Modeling

- Largest calculation to-date: 314 electrons (porphyrin)
- 50x increase in CPU time treats ~1500 electrons
  - 250 carbon atoms
    - Protein reaction centers
    - Add sidechains and counter-ions to retinal model
    - Can include larger protein environment by coupling QMC region to MM force field.
  - 375 silicon atoms (with effective core potentials)
    - Fully correlated model of quantum dots

Large Scale Modeling of Retinal HPC Requirements

Retinal

- Calculated on Franklin
- 500,000 core hours / state
- 4096 cores
- 125 MB memory / core
- 6 GB storage

Rhodopsin binding pocket

- 25,000,000 core hours
- 5 GB memory / core (or modify algorithm)
- 100 GB storage

Algorithmic Details: Parallel Strategy

- QMC offers simple, high level parallelism.
  - $E = \Sigma_i w_i E_L(X_i)$ $\quad\quad$ $E_L(X_i) = H\Psi(X_i) / \Psi(X_i)$
  - Need many X_i to reduce MC error
  - Distribute X_i to many cores
  - Collective MPI routines for averaging
- Load imbalance:
  - Branching creates more walkers on some nodes than others.
  - Addressed by on-the-fly redistribution of walkers
- Nearly all CPU time used to evaluate $\Psi(X_i)$
  - Evaluation of $\Psi(X_i)$ may exploit low-level parallelization

Algorithmic Details: Wave Function Evaluation

- CI Expansion
  - Linear combination of Slater determinants
- Computational Steps
  - Evaluate basis functions
  - Evaluate molecular orbitals
    - Linear transformation of basis functions
  - Evaluate determinant
    - Cubic scaling, but not rate limiting

- Boys-Handy Function
  - Explicit 2-body and 3-body correlation
  - Power series in interparticle distance functions
- Computational Steps
  - Evaluate distance functions
  - Sum over particle pairs, triples
    - Trace over matrix product

- Use sparsity, matrix compression to acheive linear scaling