HPC Challenges in Quantum Chemistry

Kenneth Ruud

Department of Chemistry

University of Tromsø

N-9037 Tromsø

Norway

Outline

- What is Quantum Chemistry?
- Why is Quantum Chemistry important?
- Some HPC Challenges in Quantum Chemistry:
  - Molecular Size
  - Molecular Motion
  - Molecular Stability
- A Scandinavian approach: Dalton
- Concluding remarks

The fundamental equation

- The Schrödinger equation

$$H \quad \Psi \quad = E \ \Psi$$

The fundamental equation

- The Schrödinger equation

$$H \; \Psi \; = E \; \Psi$$

- The electronic Hamilton operator

$$H = -\frac{e^2}{2m_e} \sum_i \nabla_i^2 + e^2 \sum_{iK} \frac{Z_K}{R_{iK}} + \frac{e^2}{2} \sum_{ij} \frac{1}{r_{ij}} + V_{\mathrm{NUC}}$$

The fundamental equation

- The Schrödinger equation

$$H \Psi = E \Psi$$

- The electronic Hamilton operator

$$H = -\frac{e^2}{2m_e} \sum_i \nabla_i^2 + e^2 \sum_{iK} \frac{Z_K}{R_{iK}} + \frac{e^2}{2} \sum_{ij} \frac{1}{r_{ij}} + V_{\mathrm{NUC}}$$

- The electronic wave function is a function of 3N variables

The Hartree–Fock approximation

- We assume an independent particle approximation

$$F\left(\phi_j\right)\phi_i = E_i\phi_i$$

- Iterative procedure involving matrix diagonalizations
- F built from one- and two-electron integrals

More sophisticated methods

- The HF approximation ignores electron correlation
- The HF approximation scales as $N^{2.3}$
- More sophisticated approaches scale worse
- Electron correlation important for chemical reactions
- Only *ab initio* methods can describe bond breaking and bond formation

Why is Comp. Chem. important?

- Represents a third independent way of doing chemistry: Theory, experiment and modeling

Why is Comp. Chem. important?

- Represents a third independent way of doing chemistry: Theory, experiment and modeling
- Allows for ’fast’ screening of molecules (non-linear optics, drug design)

Why is Comp. Chem. important?

- Represents a third independent way of doing chemistry: Theory, experiment and modeling
- Allows for ’fast’ screening of molecules (non-linear optics, drug design)
- Allows for an unambiguous interpretation of experimental observations

Why is Comp. Chem. important?

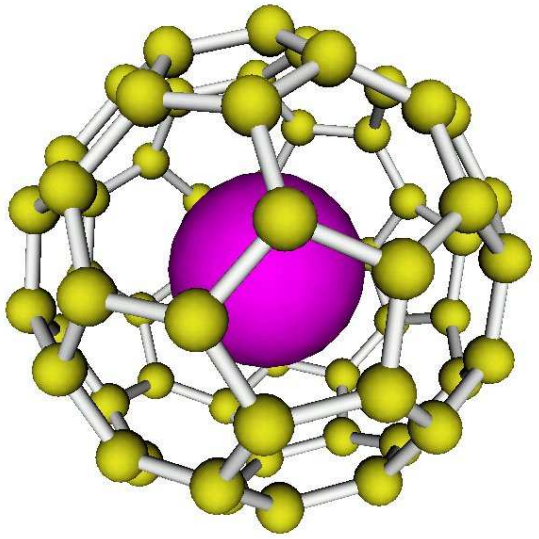
- Represents a third independent way of doing chemistry: Theory, experiment and modeling
- Allows for ’fast’ screening of molecules (non-linear optics, drug design)
- Allows for an unambiguous interpretation of experimental observations
- Enables understanding of chemical reactions at a molecular level

Why is Comp. Chem. important?

- Represents a third independent way of doing chemistry: Theory, experiment and modeling
- Allows for ’fast’ screening of molecules (non-linear optics, drug design)
- Allows for an unambiguous interpretation of experimental observations
- Enables understanding of chemical reactions at a molecular level
- Isolates individual contributions to a molecular property

Challenge 1: Molecular size

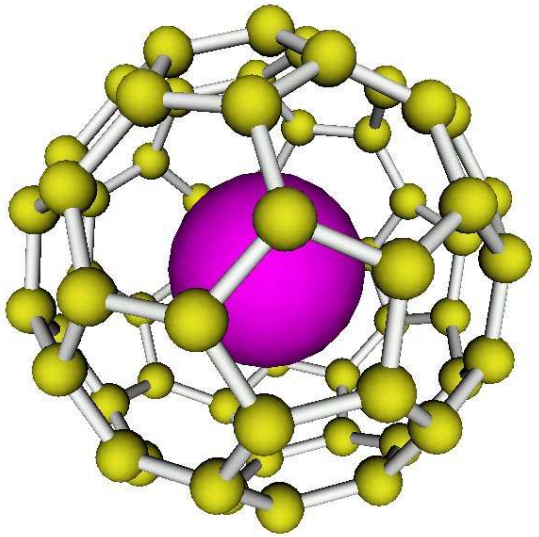
- By chemical standards, *ab initio* methods still work on small molecules



- The reason:

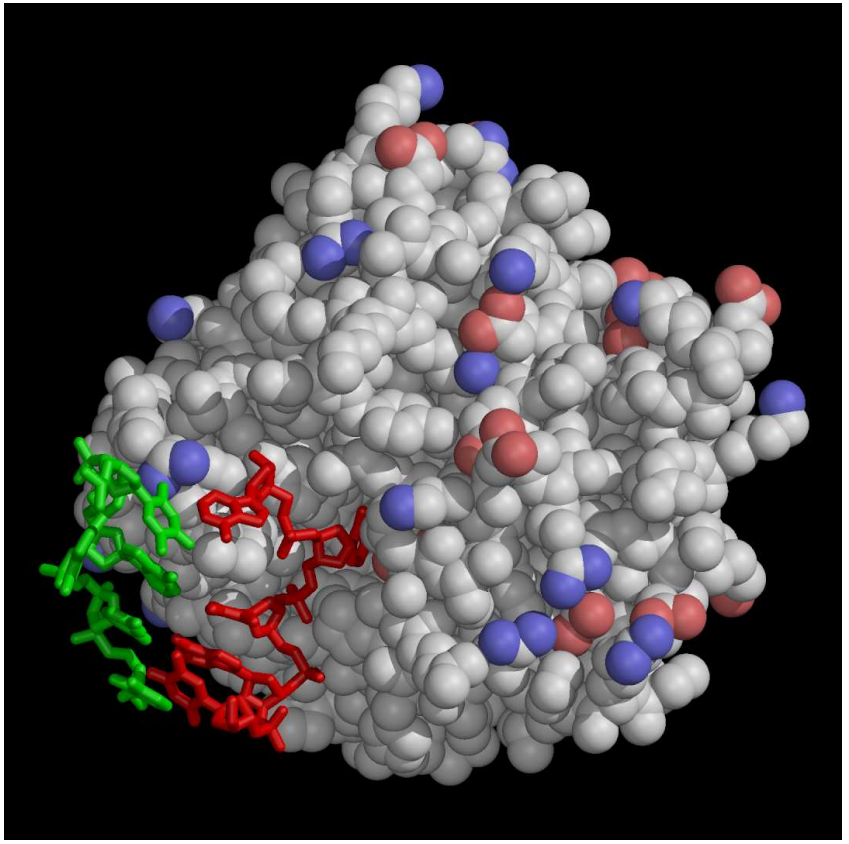
Challenge 1: Molecular size

- By chemical standards, *ab initio* methods still work on small molecules



- The reason: Unfavorable scaling (at least N^2)

What we would like to do:



The hydrogen atoms have been omitted

Picture by courtesy of Magne Olufsen (NORSTRUCT)

Linear scaling technology

- Using methods from astrophysics, the long-range interactions can be made to scale linearly with system size

Linear scaling technology

- Using methods from astrophysics, the long-range interactions can be made to scale linearly with system size
- The approach parallelizes very well

Linear scaling technology

- Using methods from astrophysics, the long-range interactions can be made to scale linearly with system size
- The approach parallelizes very well
- New (parallel) bottleneck: matrix diagonalization

Linear scaling technology

- Using methods from astrophysics, the long-range interactions can be made to scale linearly with system size
- The approach parallelizes very well
- New (parallel) bottleneck: matrix diagonalization
- Convergence problems a challenge

Linear scaling technology

- Using methods from astrophysics, the long-range interactions can be made to scale linearly with system size
- The approach parallelizes very well
- New (parallel) bottleneck: matrix diagonalization
- Convergence problems a challenge
- Better functionals

HPC challenges/needs

- A lot of computer power

HPC challenges/needs

- A lot of computer power
- More work on linear scaling technology

HPC challenges/needs

- A lot of computer power
- More work on linear scaling technology
- Efficient, parallel matrix multiplication and matrix diagonalization routines

HPC challenges/needs

- A lot of computer power
- More work on linear scaling technology
- Efficient, parallel matrix multiplication and matrix diagonalization routines
- Routines for operations on sparse matrices

Challenge 2: Molecular motion

- Chemistry has two parts:

Challenge 2: Molecular motion

- Chemistry has two parts:
  - Properties of molecules

Challenge 2: Molecular motion

- Chemistry has two parts:
  - Properties of molecules
  - Reaction of molecules

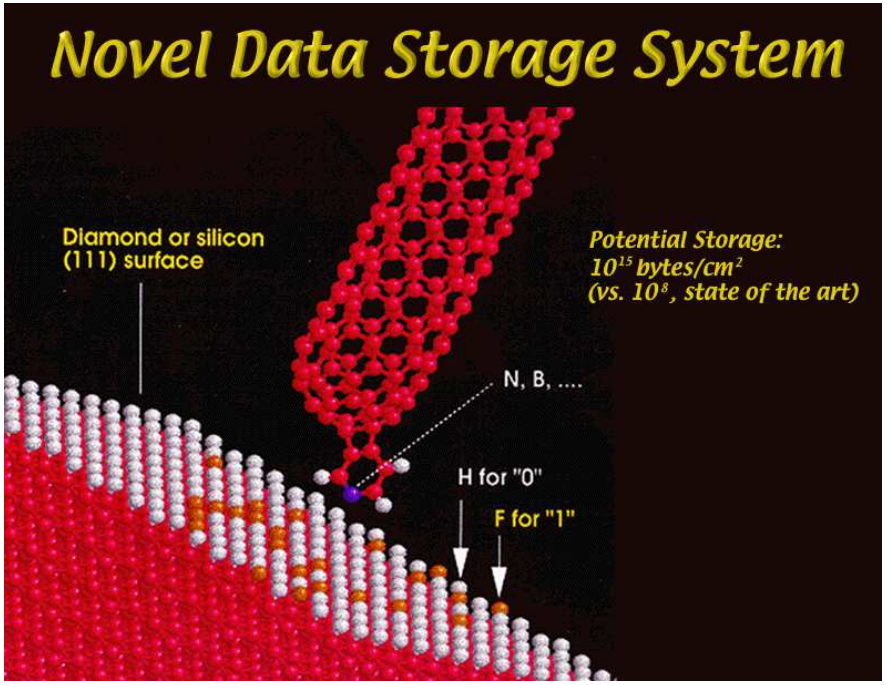
Challenge 2: Molecular motion

- Chemistry has two parts:
  - Properties of molecules
  - Reaction of molecules
- Currently no computationally efficient method that can correctly break bonds

Challenge 2: Molecular motion

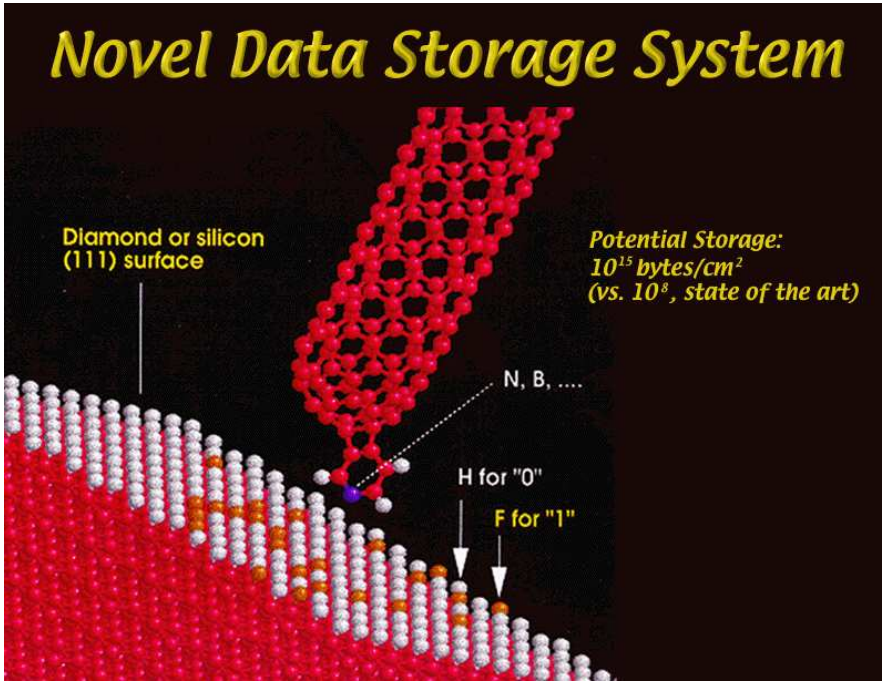
- Chemistry has two parts:
  - Properties of molecules
  - Reaction of molecules
- Currently no computationally efficient method that can correctly break bonds
- This remains a quantum-chemical challenge

How do molecules react and interact?



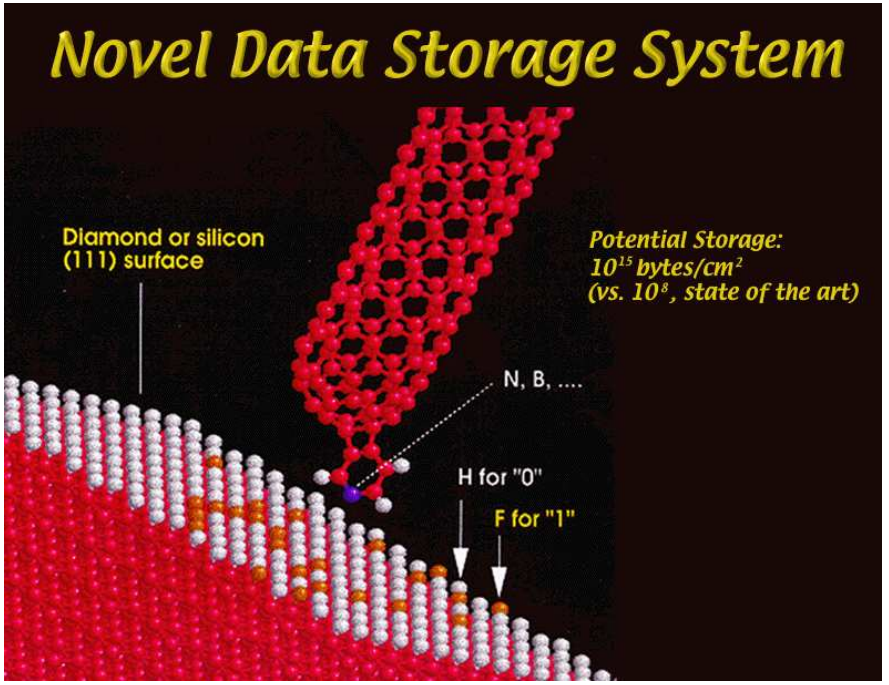
- 3M-6 vibrational coordinates

How do molecules react and interact?



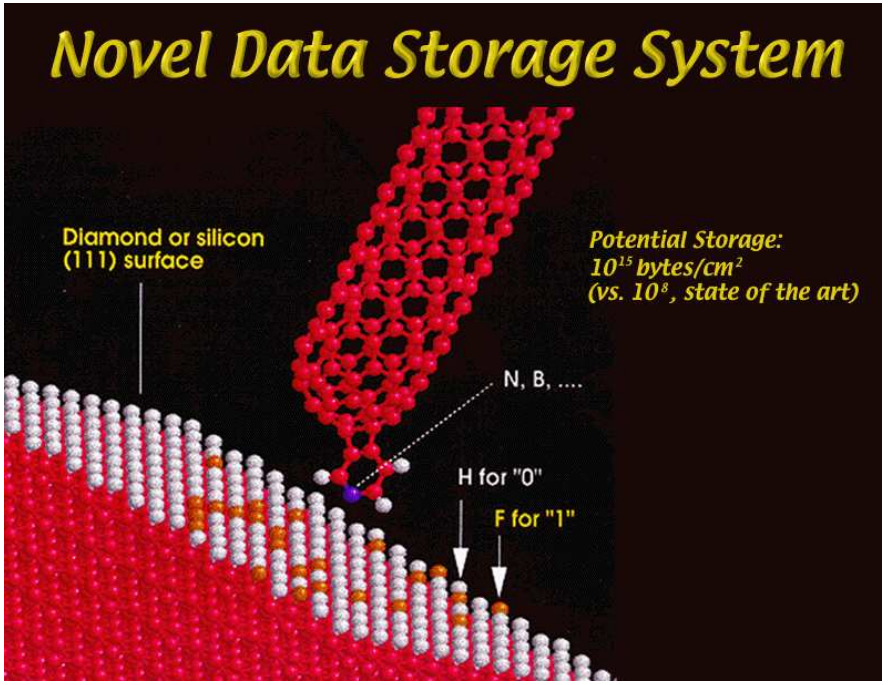
- 3M-6 vibrational coordinates
- Easily parallelizable

How do molecules react and interact?



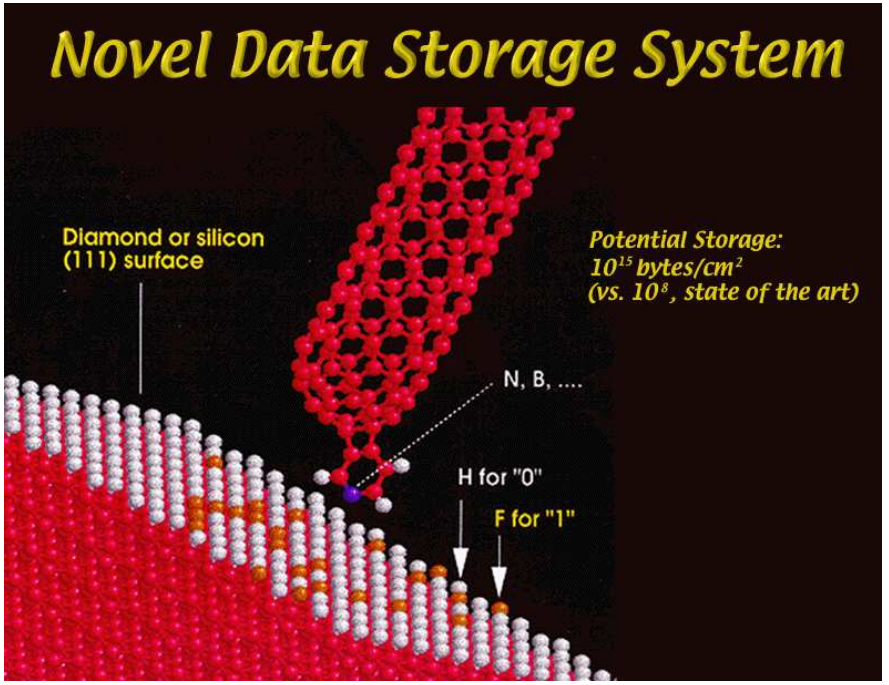
- 3M-6 vibrational coordinates
- Easily parallelizable
- Analytic approaches possible

How do molecules react and interact?



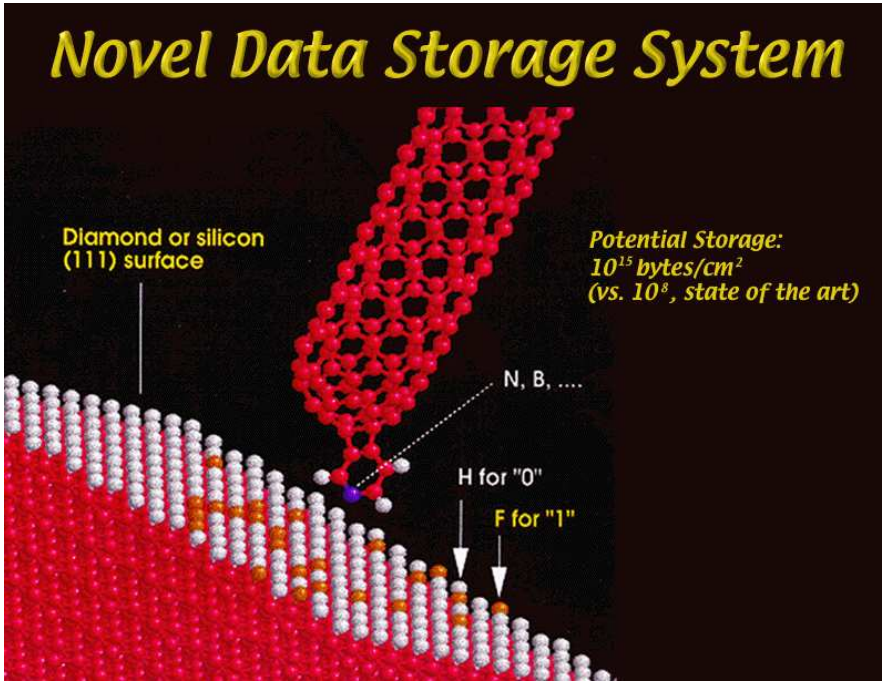
- 3M-6 vibrational coordinates
- How do the molecules interact?

How do molecules react and interact?



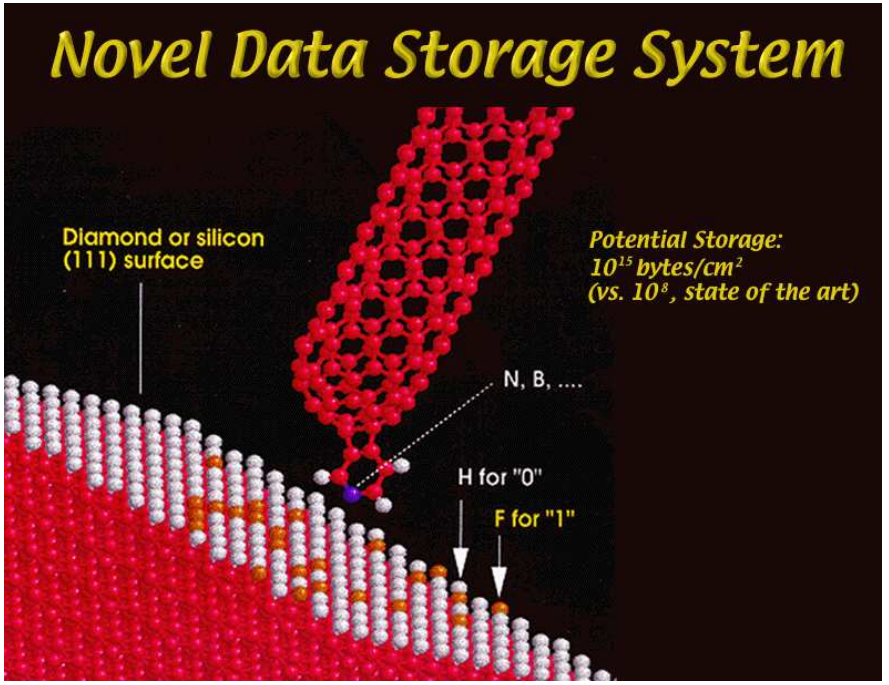
- 3M-6 vibrational coordinates
- How do the molecules interact?
- Time has direction

How do molecules react and interact?



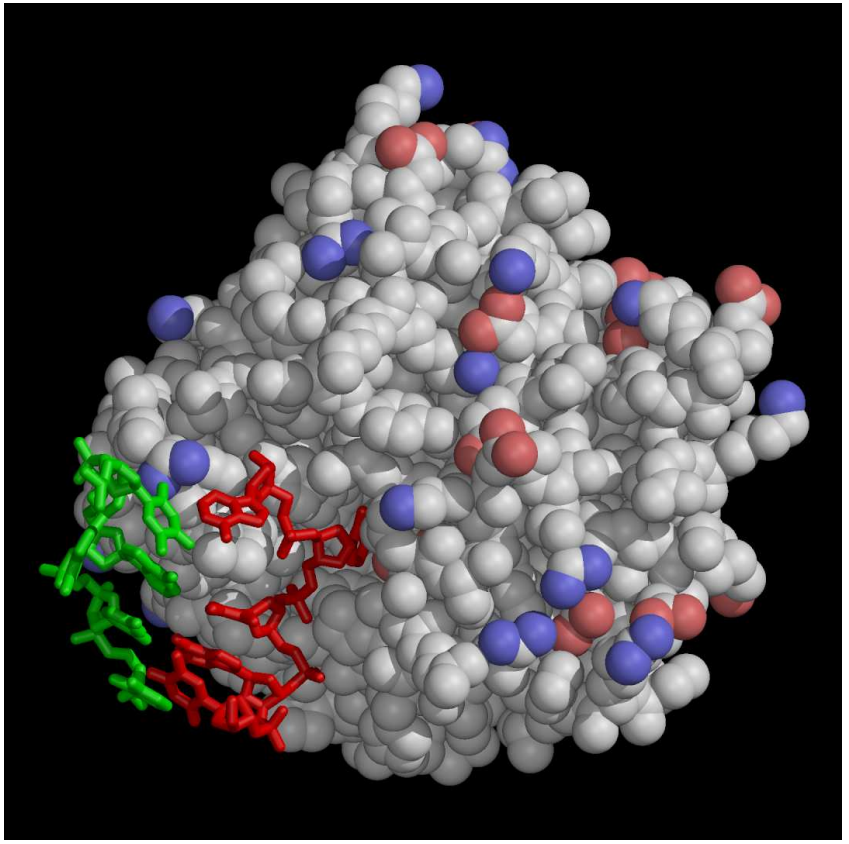
- 3M-6 vibrational coordinates
- How do the molecules interact?
- Time has direction
- Parallelization difficult

How do molecules react and interact?



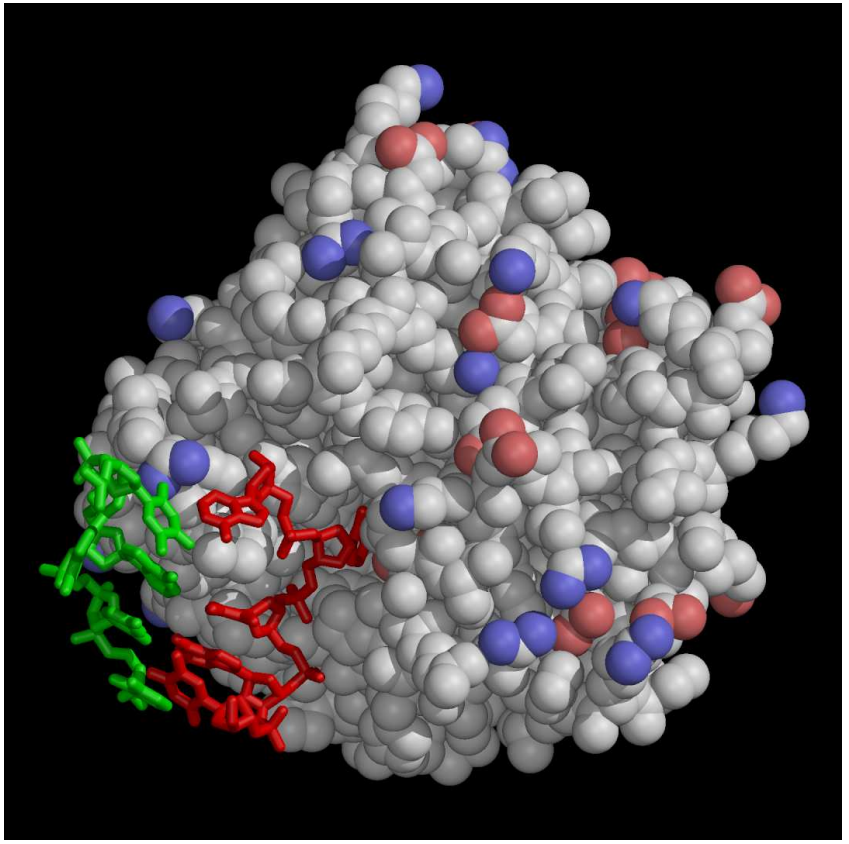
- 3M-6 vibrational coordinates
- How do the molecules interact?
- Time has direction
- Parallelization difficult
- Long time scales

Challenge 3: Molecular stability



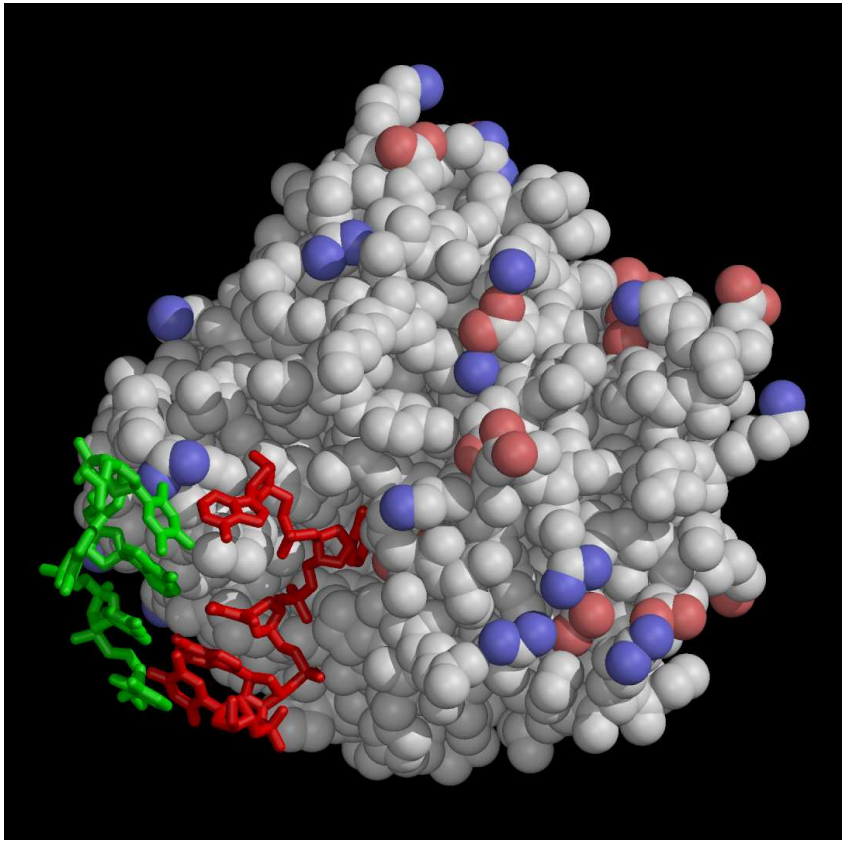
- Is this the most stable structure?

Challenge 3: Molecular stability



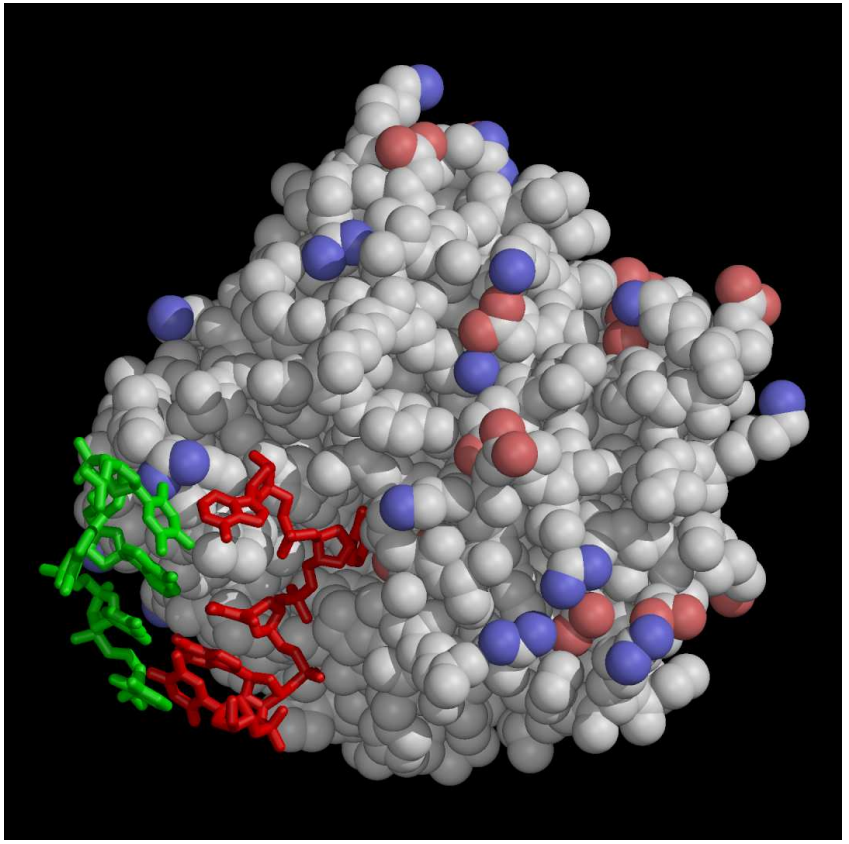
- Is this the most stable structure?
- Which other structures are accessible?

Challenge 3: Molecular stability



- Is this the most stable structure?
- Which other structures are accessible?
- Global minimization

Challenge 3: Molecular stability



- Is this the most stable structure?
- Which other structures are accessible?
- Global minimization
- Decoupled search (parallelizable)

The Dalton program

C. Angeli, K. L. Bak, V. Bakken, O. Christiansen, R. Cimiraglia,
S. Coriani, P. Dahle, E. K. Dalskov, T. Enevoldsen, B. Fernandez,
C. Hättig, K. Hald, A. Halkier, H. Heiberg, T. Helgaker, H. Hettema,
H. J. Aa. Jensen, D. Jonsson, P. Jørgensen, S. Kirpekar, W. Klopper,
R. Kobayashi, H. Koch, O. B. Lutnæs, K. V. Mikkelsen, P. Norman,
J. Olsen, M. J. Packer, T. B. Pedersen, Z. Rinkevicius, E. Rudberg,
T. A. Ruden, K. Ruud, P. Sałek, A. Sanchez de Meras, T. Saue,
S. P. A. Sauer, B. Schimmelpfennig, K. O. Sylvester-Hvid, P. R. Taylor,
O. Vahtras, D. J. Wilson, H. Ågren.

http://www.kjemi.uio.no/software/dalton/dalton.html

Current status of Dalton

- Dalton 1.2.1 distributed to more than 1200 research groups
- Distributed free of charge (but signed license required)
- Dalton 2.0 released March 1, and now distributed to more than 400 research groups
- Can be downloaded from http://www.kjemi.uio.no/software/dalton/dalton.html
- Largest developer sites: Stockholm, Linköping, Oslo, Trondheim, Tromsø, Århus, Odense, København

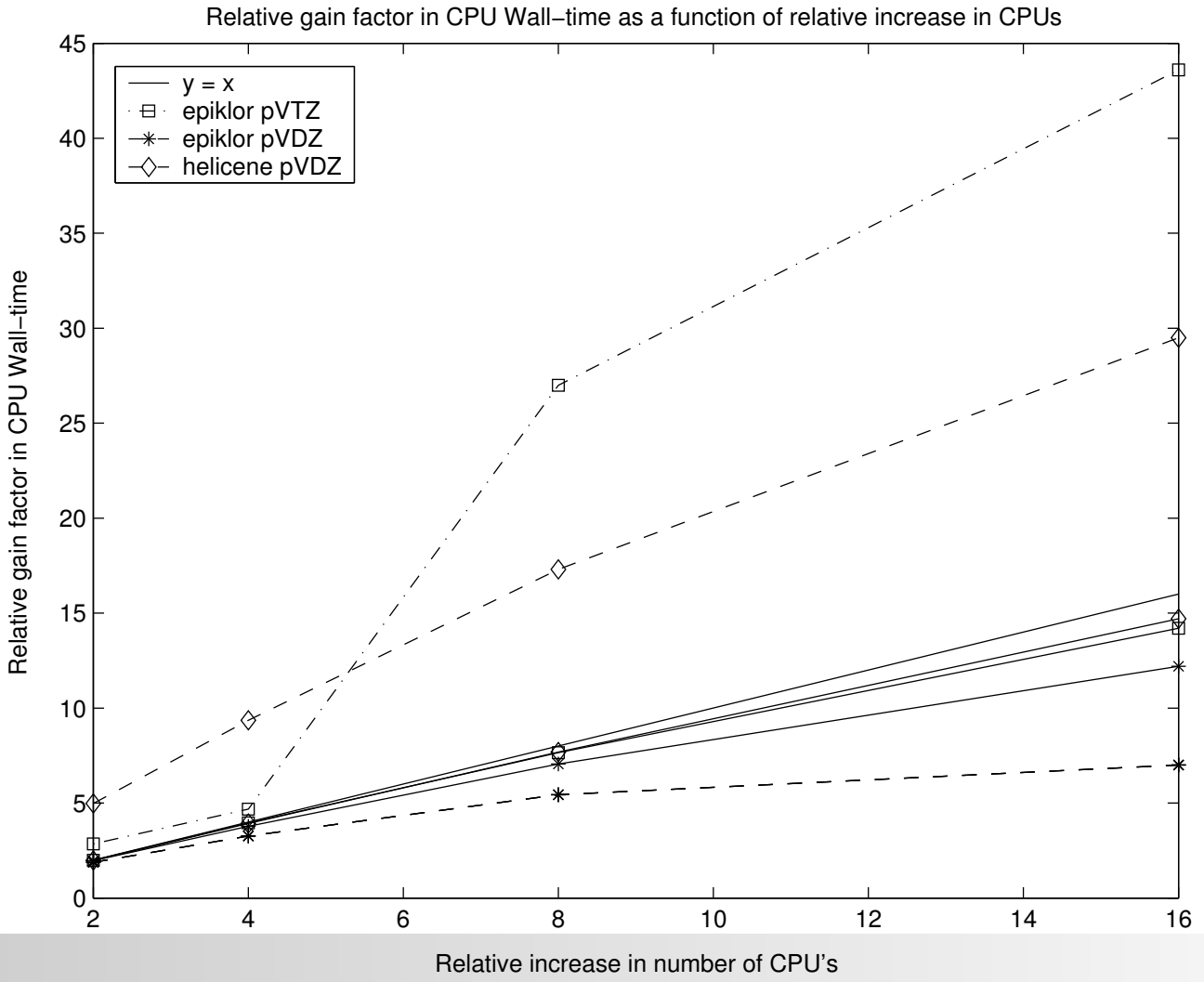
Methodological advances

- As for most other HPC research fields, improvements in algorithms and programs as important as improvements in hardware
- Perhaps even more so for chemistry since hardware not adapted to our problems, and our problems are ’non-standard’
- Not all these methodological advances can lead to publications (developed by others): How do we ensure they are implemented and used?
- Two Dalton examples:
  - Incore storage of two-electron integrals
  - Density fitting

In-core integral calculations

- Much more memory available on parallel machines
- Can be used to store two-electron integrals instead of recalculating these
- Particularly effective with density fitting (three-index integrals)
- Need to know where integrals are stored
- Master keeps track of stored integral batches $(ij \mid **)$
- Send batches stored back to the same slave to be used in the (partial) construction of a Fock matrix

HF energy calculations: scaling



Density-fitting

- One of our four-index quantities is the Coulomb-interaction

$$J_{\mu\nu\rho\sigma} = \int \int \phi_\mu^* (\mathbf{r}_1) \, \phi_\nu (\mathbf{r}_1) \, \frac{1}{|r_1 - r_2|} \phi_\rho^* (\mathbf{r}_2) \, \phi_\sigma (\mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2$$

- Interaction between two densities $\Rightarrow$ we fit one of the densities to a single basis

$$J_{\mu\nu\rho\sigma} \approx \sum_\tau C_\tau \int \int \phi_\mu^* (\mathbf{r}_1) \, \phi_\nu (\mathbf{r}_1) \, \frac{1}{|r_1 - r_2|} \Lambda_\tau (\mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2$$

- Reduces the formal scaling from N^4 to N^3

Density-fitting

- One of our four-index quantities is the Coulomb-interaction

$$J_{\mu\nu\rho\sigma} = \int \int \phi_\mu^* (\mathbf{r}_1) \, \phi_\nu (\mathbf{r}_1) \, \frac{1}{|r_1 - r_2|} \phi_\rho^* (\mathbf{r}_2) \, \phi_\sigma (\mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2$$

- Interaction between two densities $\Rightarrow$ we fit one of the densities to a single basis

$$J_{\mu\nu\rho\sigma} \approx \sum_\tau C_\tau \int \int \phi_\mu^* (\mathbf{r}_1) \, \phi_\nu (\mathbf{r}_1) \, \frac{1}{|r_1 - r_2|} \Lambda_\tau (\mathbf{r}_2) \, d\mathbf{r}_1 \, d\mathbf{r}_2$$

- Reduces the formal scaling from N^4 to N^3
- For a calculation of 1000 basis functions, integral time reduced from 62 minutes to 5 minutes

Concluding remarks

- Ab initio theory is today a third way of chemical research

Concluding remarks

- Ab initio theory is today a third way of chemical research
- However, many molecules of chemical interest is still too large for simulations

Concluding remarks

- Ab initio theory is today a third way of chemical research
- However, many molecules of chemical interest is still too large for simulations
- Chemical reactions remain a problem

Concluding remarks

- Ab initio theory is today a third way of chemical research
- However, many molecules of chemical interest is still too large for simulations
- Chemical reactions remain a problem
- The Scandinavian HPC has the opportunity to enable a concerted hardware/software effort in the field of quantum chemistry through the Dalton program.

Concluding remarks

- Ab initio theory is today a third way of chemical research
- However, many molecules of chemical interest is still too large for simulations
- Chemical reactions remain a problem
- The Scandinavian HPC has the opportunity to enable a concerted hardware/software effort in the field of quantum chemistry through the Dalton program.
- Requires man-power to implement better algorithms developed by others (and develop new algorithms)

Concluding remarks

- Ab initio theory is today a third way of chemical research
- However, many molecules of chemical interest is still too large for simulations
- Chemical reactions remain a problem
- The Scandinavian HPC has the opportunity to enable a concerted hardware/software effort in the field of quantum chemistry through the Dalton program.
- Requires man-power to implement better algorithms developed by others (and develop new algorithms)
- Requires incentives for users to switch to Dalton