Large-scale QC modeling using PC GAMESS package.

Alex A. Granovsky

Laboratory of Chemical Cybernetics, M.V. Lomonosov Moscow State University, Moscow, Russia

June 30, 2006

Outline

- Introduction to PC GAMESS
- Modeling of large systems:
  - High-level methods
  - Linear scaling methods
  - QM/MM-based approaches

The PC GAMESS project

- Intel architecture specific high-performance parallel freely-available QC package
- Is being developed in the Laboratory of Chemical Cybernetics at MSU since 1993
- Used by more than 2500 academic/industry research groups all around the world
- Supports Windows and Linux based SMP systems, clusters, and their combinations
- New state-of-the-art parallel algorithms for many QC calculation methods, scalable up to hundreds of nodes.

PC GAMESS vs. GAMESS US:

- **Strongly modified** to achieve the maximum possible performance on Intel-based platforms;
- **Functionally extended** to provide QC methods which are not currently present in the regular GAMESS version;
- **Designed to deal with large molecular systems efficiently**;
- **Runs much faster and requires less resources**

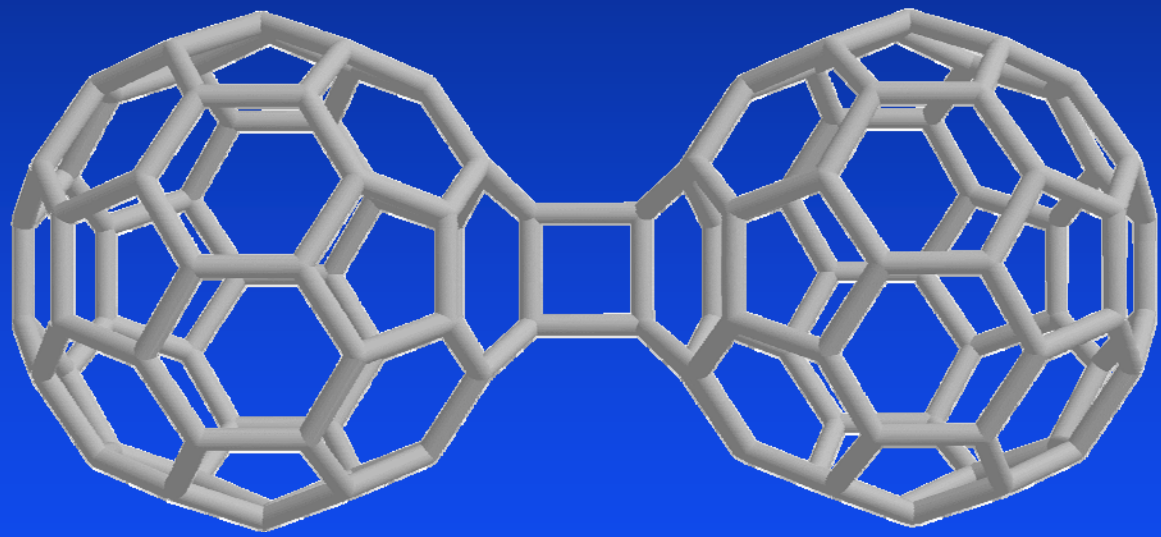
Modeling of large systems, high-level methods

High-level methods, current status

- Routinely applicable to modeling of ground and excited states of medium-size systems:
  - State-of-the-art parallel algorithms for MP2-MP4, MCSCF, and MCQDPT2
  - Allows to model medium size systems (up to 1000-4000 basis functions) with high accuracy and reasonable calculation time

High-level methods, examples

Medium size systems: Fullerene C_{60} and its dimer C_{120}, MP2 calculations

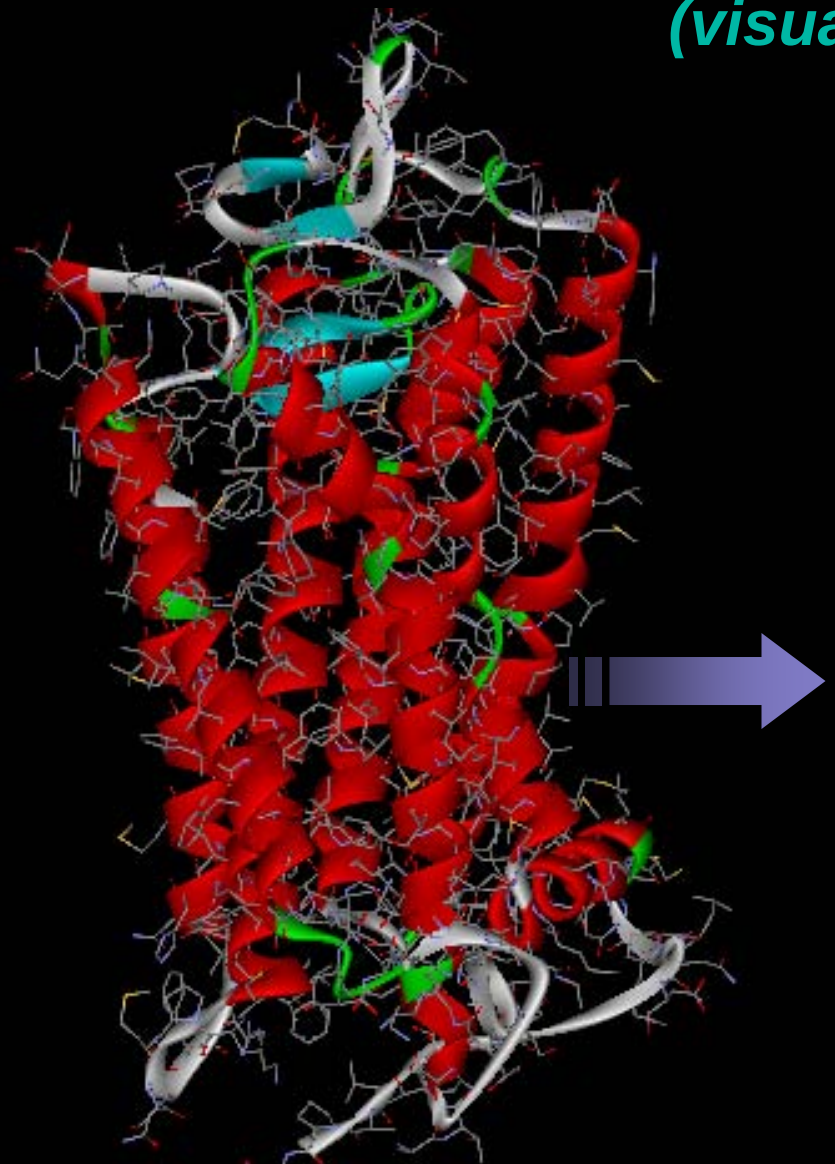


Largest MP2 calculation attempted so far

System	C_{120}		
Basis	cc-pVTZ-f		
Group	D_{2h}		
N	3000		
c	120		
n	240		
N_{nodes}	18		
Dynamic load balancing	off	on	on
Real time data packing	off	on	on
Asynchronous I/O	off	off	on
Total FP operations count	$3.24 \cdot 10^{15}$	$3.32 \cdot 10^{15}$	$3.32 \cdot 10^{15}$
Distributed data size	2.0 TB	2.0 TB	2.0 TB
CPU time on master node, sec	83029	89301	95617
Wall clock time, sec.	150880	110826	95130
CPU usage, %	55	80.5	100.5
Node performance, MFlops/s	1330	1935	2320
Performance, % of peak	27.7	40.3	48.3
Cluster performance, GFlops/s	23.9	34.8	41.7

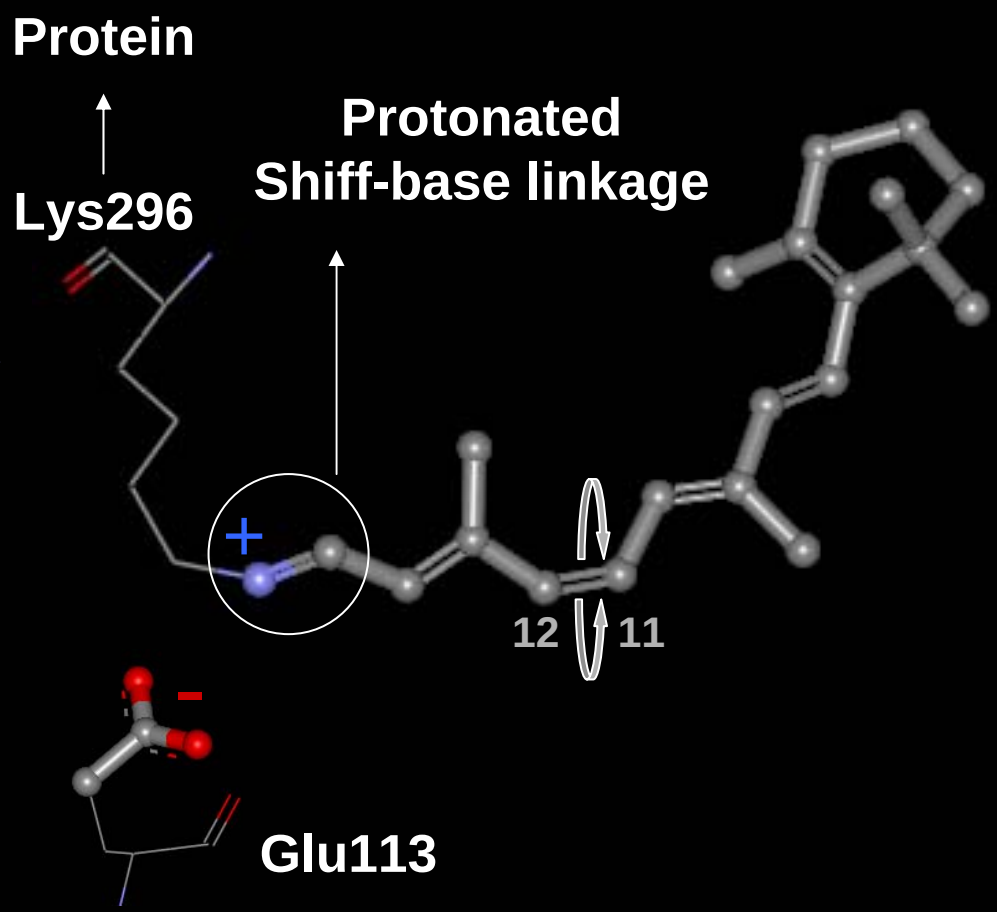
Pentium 4C 2.4 GHz / 1024MB / 120GB / Gigabit Ethernet

Medium size systems: structure and electronic spectra of retinal molecule in rhodopsin photoreceptor protein (visual pigment)



Rhodopsin ~ *5500 atoms*

11-cis retinal chromophore



Electronic spectra of retinal molecule in rhodopsin photoreceptor protein

- Retinal molecule in protein (rhodopsin) environment - photosensitive receptor
  - 77 atoms (H, C, N, O), 258 electrons
  - approx. 5500 atoms in the protein
- four-state QM/MM MCQDPT2 calculation for high-quality description of the excited states and electronic transition moments
  - Number of basis functions (N) **715**
  - Number of FP operations ~ $1.5 \cdot \mathbf{10^{15}}$
- Less than 2 days on the single dual-Nocona workstation (this is one of the largest calculation of that kind ever attempted so far).

High-level methods, future plans

- No linear scaling approaches at present but will be implemented in the future:
  - Local correlation methods with linear scaling
  - Resolution of Identity (RI)-based methods.

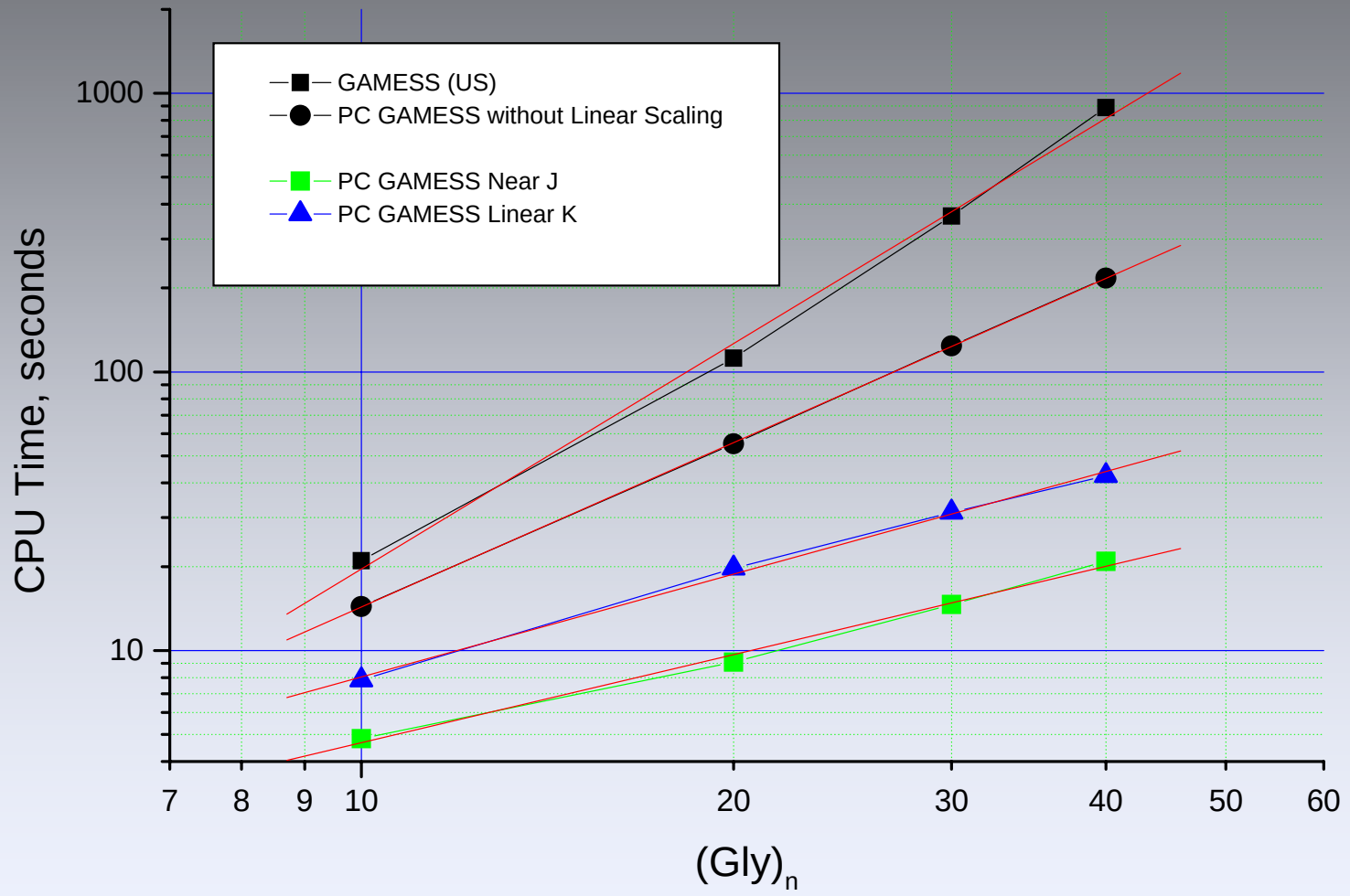
Modeling of large systems - linear scaling methods for HF, DFT, CIS, TDHF, and TDDFT

Linear scaling methods, current status

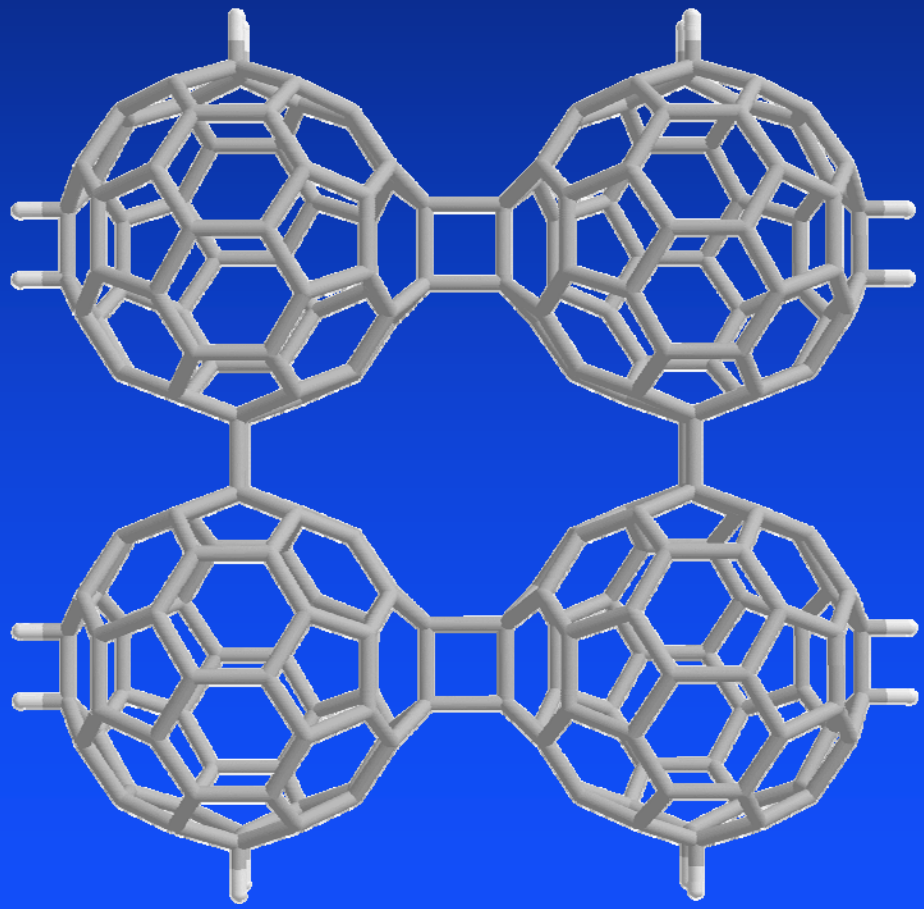
- Fast linear scaling Coulomb (J matrix) assembly via QFMM and near-field J engine
- State-of-the-art grid-based Exchange-Correlation integration for linear scaling DFT and TDDFT
- State-of-the-art linear scaling HF exchange (K matrix) assembly for HF, hybrid DFT, CIS, TDHF, and TDDFT

Linear scaling methods, examples

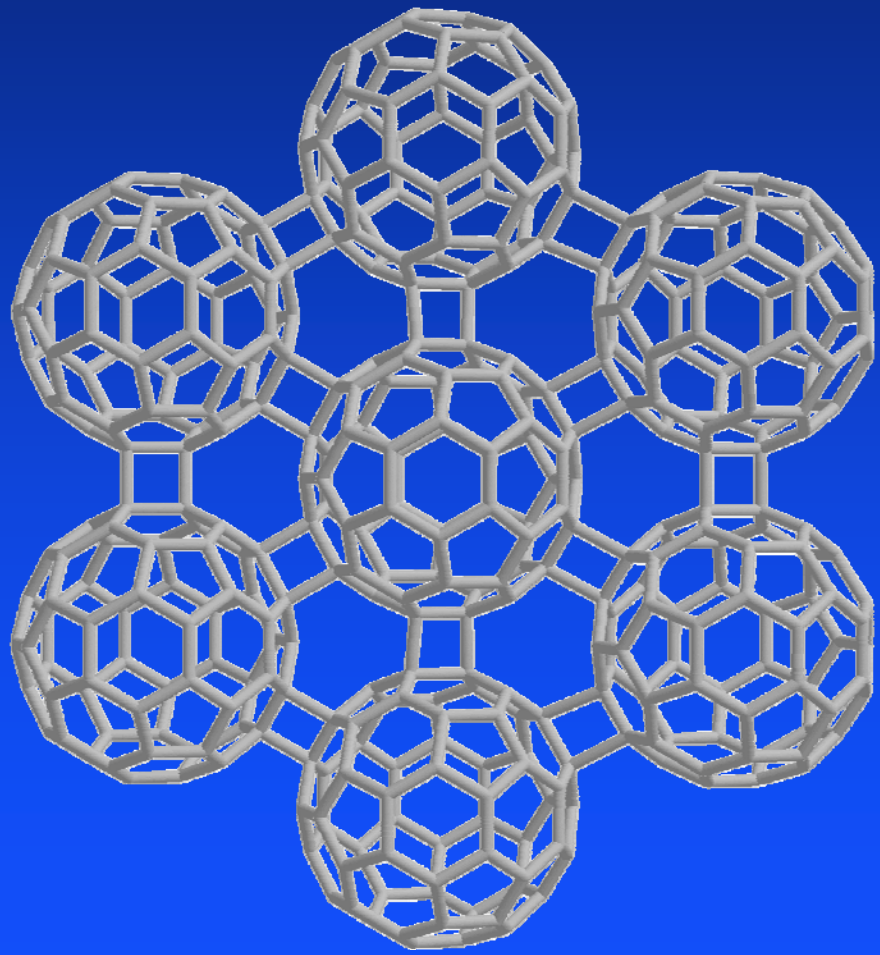
Scaling example, Glycine oligomers (running on 18 nodes)



Fullerene (C_{60}) 2D polymer (tetragonal phase) - modeled by PC GAMESS



Fullerene (C_{60}) 2D polymer (hexagonal phase) - modeled by PC GAMESS



Linear scaling methods, future plans

- Linear scaling QFMM energy gradients for HF, DFT, and TDDFT
- RI-based methods:
  - RI-J based code for pure DFT functionals
  - RI-K based code for hybrid DFT
  - Multipole Accelerated RI (MARI) energies and gradients for super-fast linear scaling DFT and TDDFT
- GW-based methods for excited states and zone theory

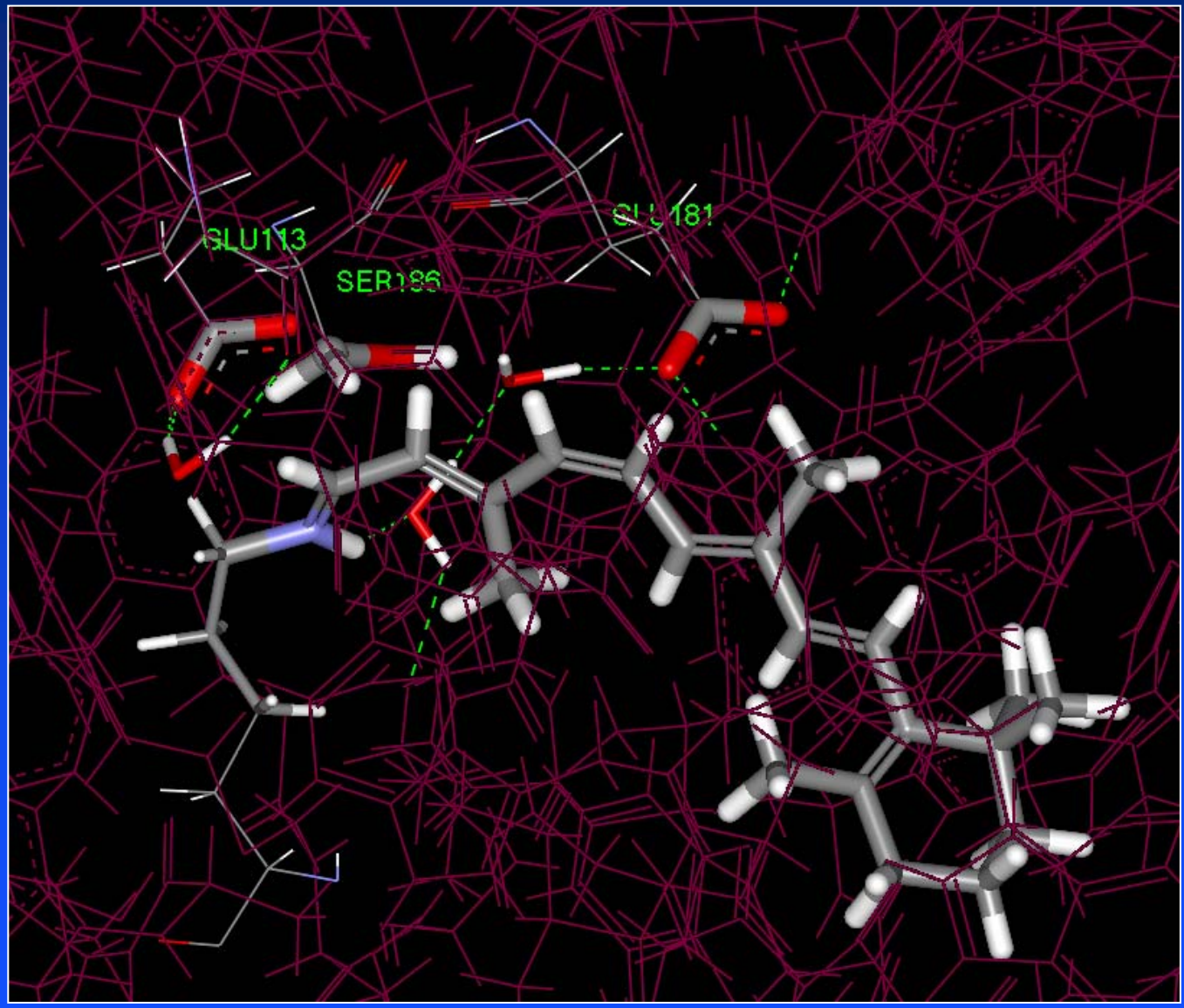
Modeling of large systems - QM/MM-based methods

QM/MM methods, current status

- Several QM/MM approaches are already implemented in the PC GAMESS:
  - Mechanical embedding
  - Electronic embedding
- Different types of QM/MM:
  - Classical QM/MM (Jim Kress)
    - Several empirical force fields, including UFF
  - QM/DIM (Diatomics In Molecules) (A. Bochenkova)
    - Nonempirical many-body force fields
  - QM/EFP (GAMESS US)
    - Semiempirical and nonempirical corrections to H_1

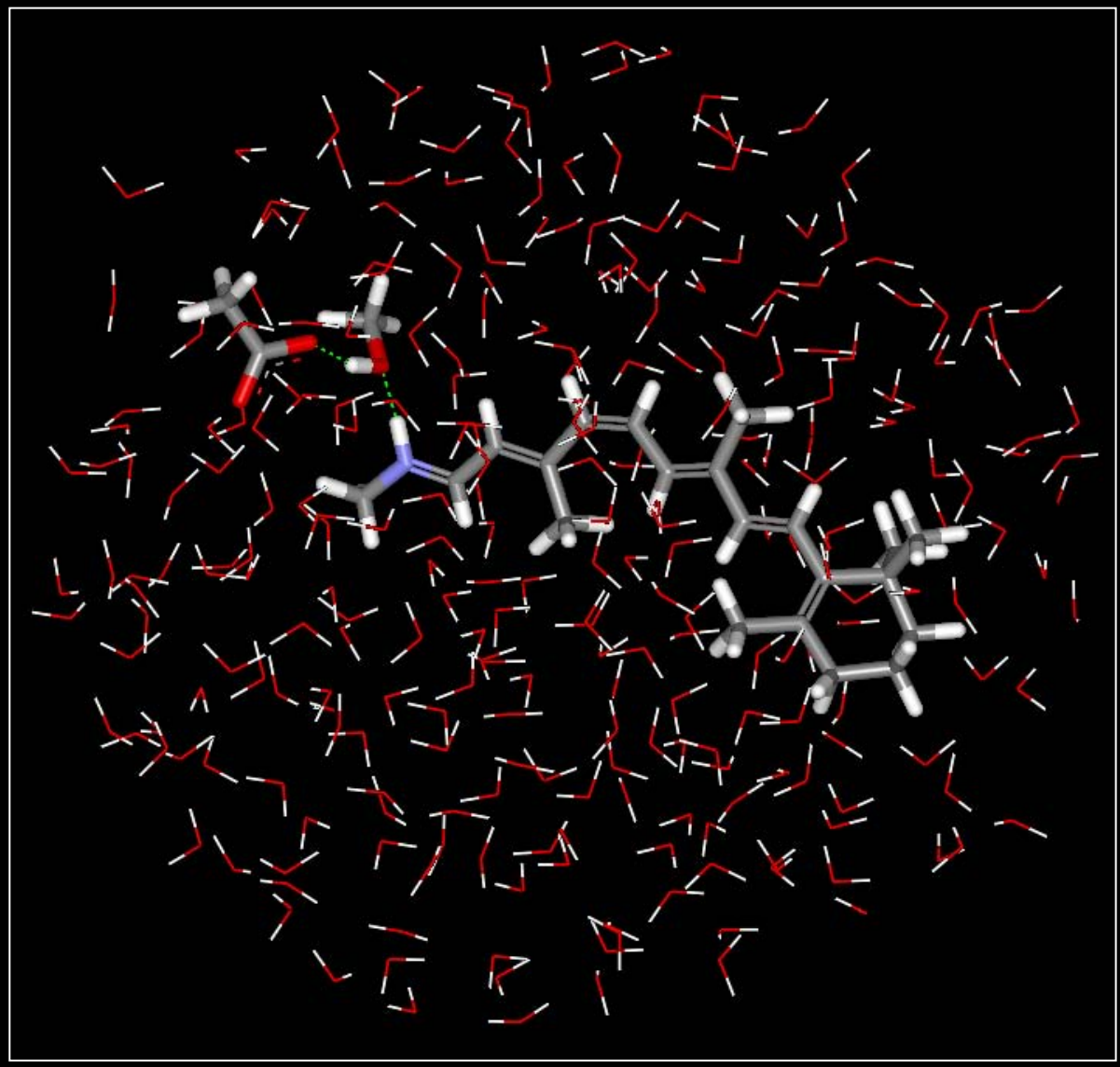
QM/MM methods, examples

Equilibrium structure of Retinal molecule in Rhodopsin protein (5500 atoms)



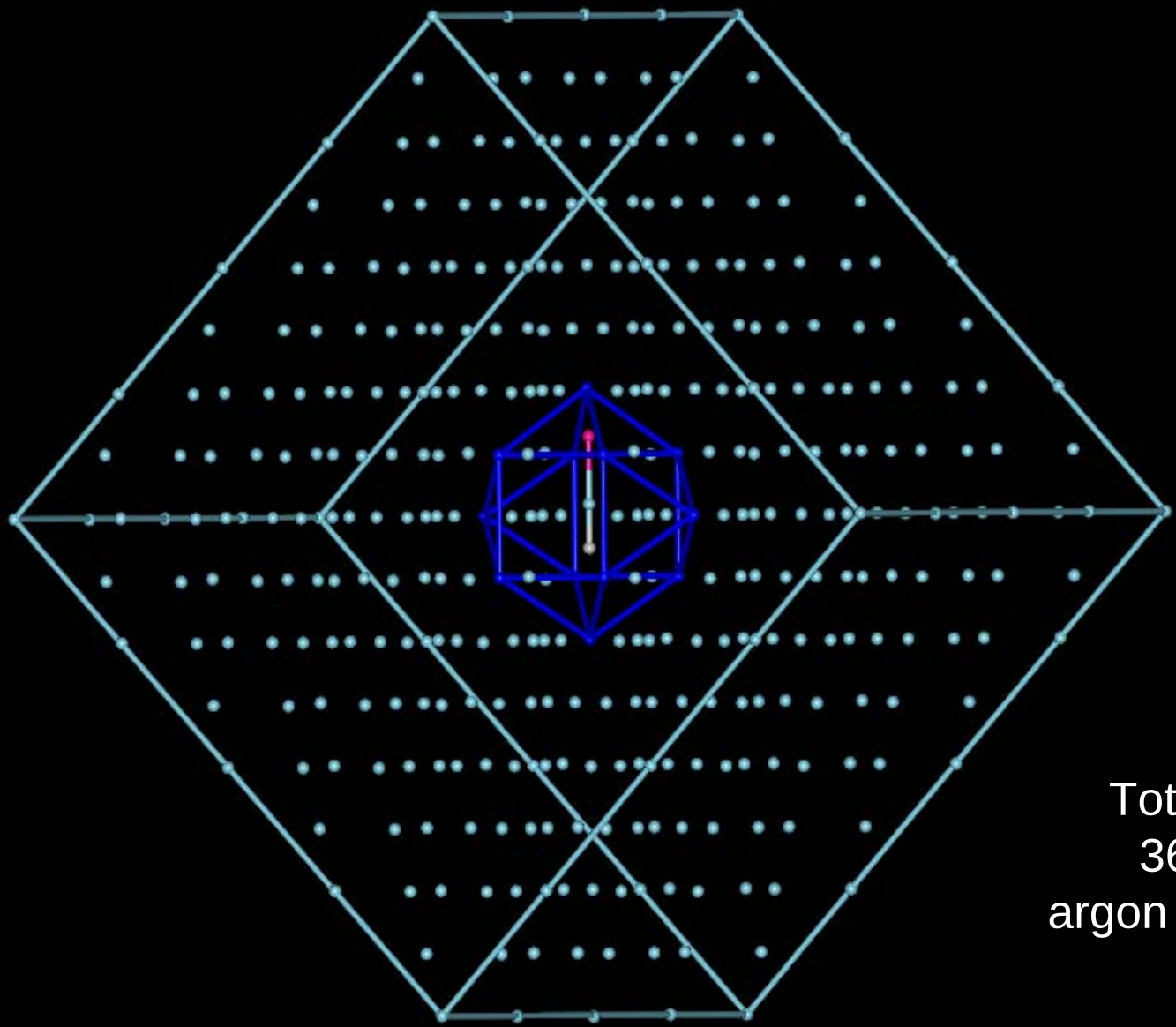
QM/MM methods - modeling effects of the environment

Equilibrium geometry of the retinal molecule in water solution - PC GAMESS QM/EFP results



Totally 275 water molecules

Equilibrium geometry of HArF molecule in argon matrix - PC GAMESS QM/DIM results



Totally
365
argon atoms

QM/MM methods, future plans

- Implement polarizable force fields

The PC GAMESS on the Web:

- **http://classic.chem.msu.su/gran/gamess/index.html**