Water clusters, ice and bulk liquid: Improving ab initio structure and energetics

Dario ALFÈ

Department of Earth Sciences, UCL
Department of Physics and Astronomy, UCL
London Centre for Nanotechnology, UCL
Thomas Young Centre, UCL

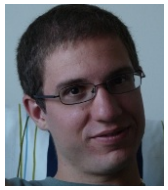
EPSRC Engineering and Physical Sciences Research Council

THE ROYAL SOCIETY

People

Mike Gillan (UCL)

Gábor Csányi (Cambridge)



Albert Bartók-Pártay (Cambridge)

Fred Manby (Bristol)

Mike Towler (UCL)

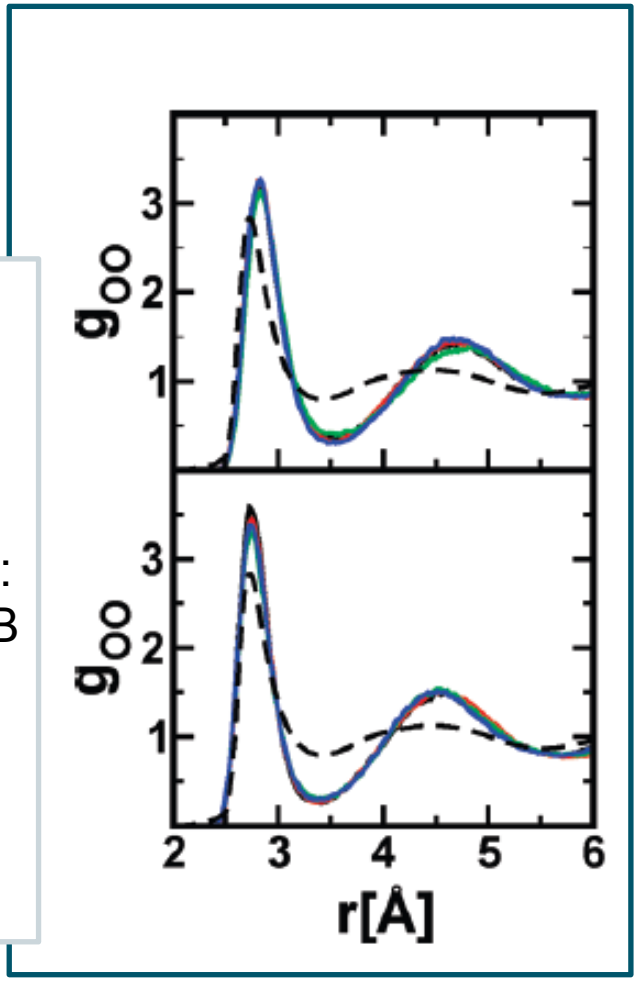
Angelos Michaelides (UCL)

It's a long, long way…

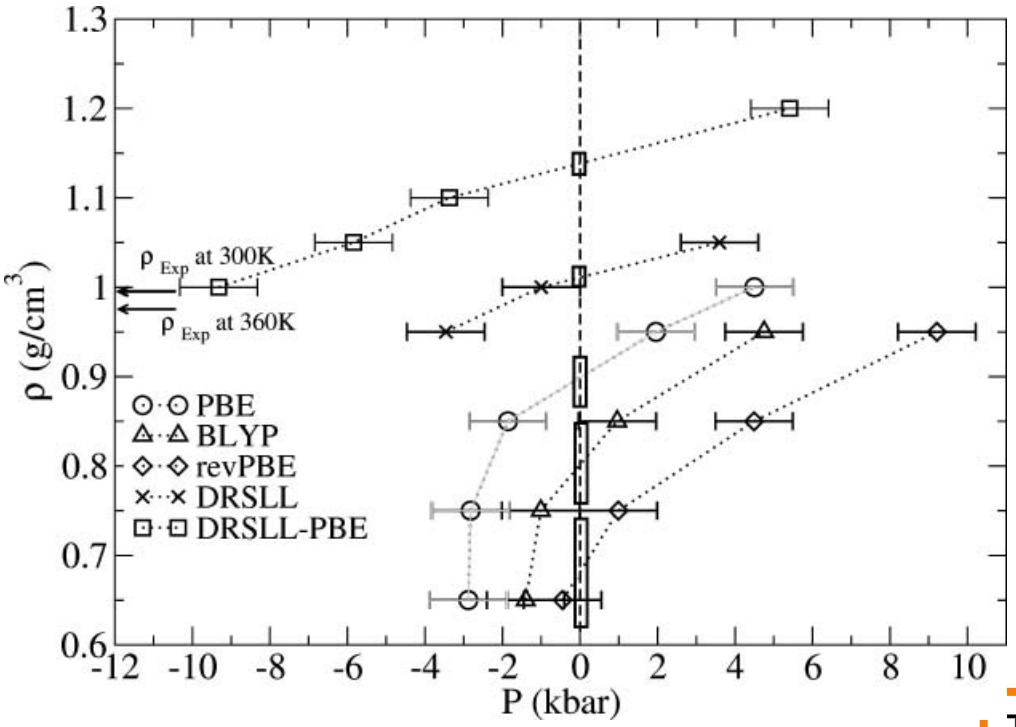
- Classical (rigid, unpolarisable): Bernal-Fowler, SPC, TIPnP, etc: 1933 onwards…
- Classical (flexible, polarisable): Kollman, Dang-Chang, Ponder, Xantheas: AMOEBA, TTMn-F: 1990 onwards…
- *Ab initio*-based: Clementi, Jordan, Szalewicz...: 1976 onwards…
- DFT: Parrinello, Car, Sprik, Tuckerman, Galli and many others:1992 onwards…
- Quantum nuclear corrections with path integrals: Manolopoulos, Car…

After 80 years of intensive effort by 1000's of researchers…

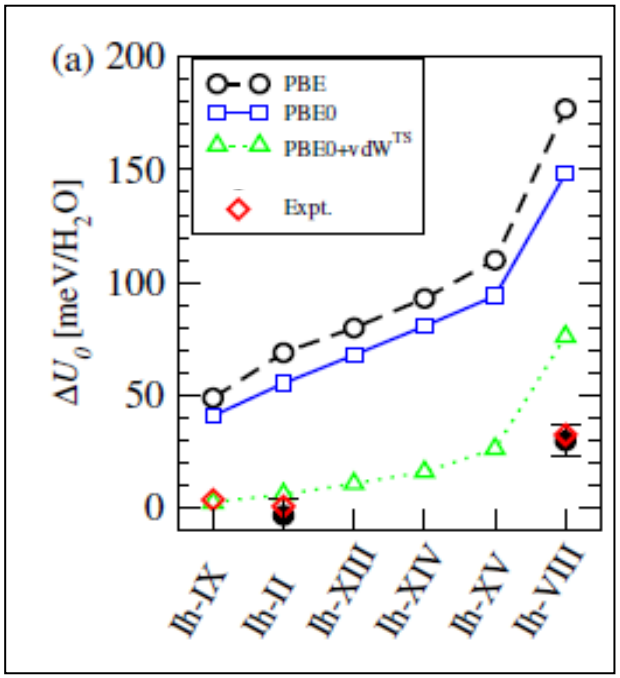
O-O rdfs of bulk liquid water: many DFTs make water over-structured, with a very low diffusion coefficient: Schmidt *et al.* JPCB 2009. Top: BLYP, Bottom: PBE, compared with experiment (dashed)



Predicted density of water can be 30% too low with some functionals (icebergs would sink!)



Wang, Román-Pérez, Soler, Artacho, Fernández-Serra, J. Chem. Phys. (2011)



Relative energies of ice structures: Santra, Alfè, Car, Michaelides, Scheffler *et al.*, PRL 2011

Transition pressures between ice structures can be off by a factor 10

Parts of the energy

1st-order electrostatics	Coulomb interaction between unperturbed charge distributions of monomers	2-body
Polarisation	2nd-order electrostatics (NB: not just dipolar polarisability)	Many-body
Exchange-overlap	Pauli repulsion of closed shells	Mainly 2-body
Dispersion	Non-local electron correlation	Mainly 2-body
Monomer deformation	Bond stretch, bond-angle bend	1-body

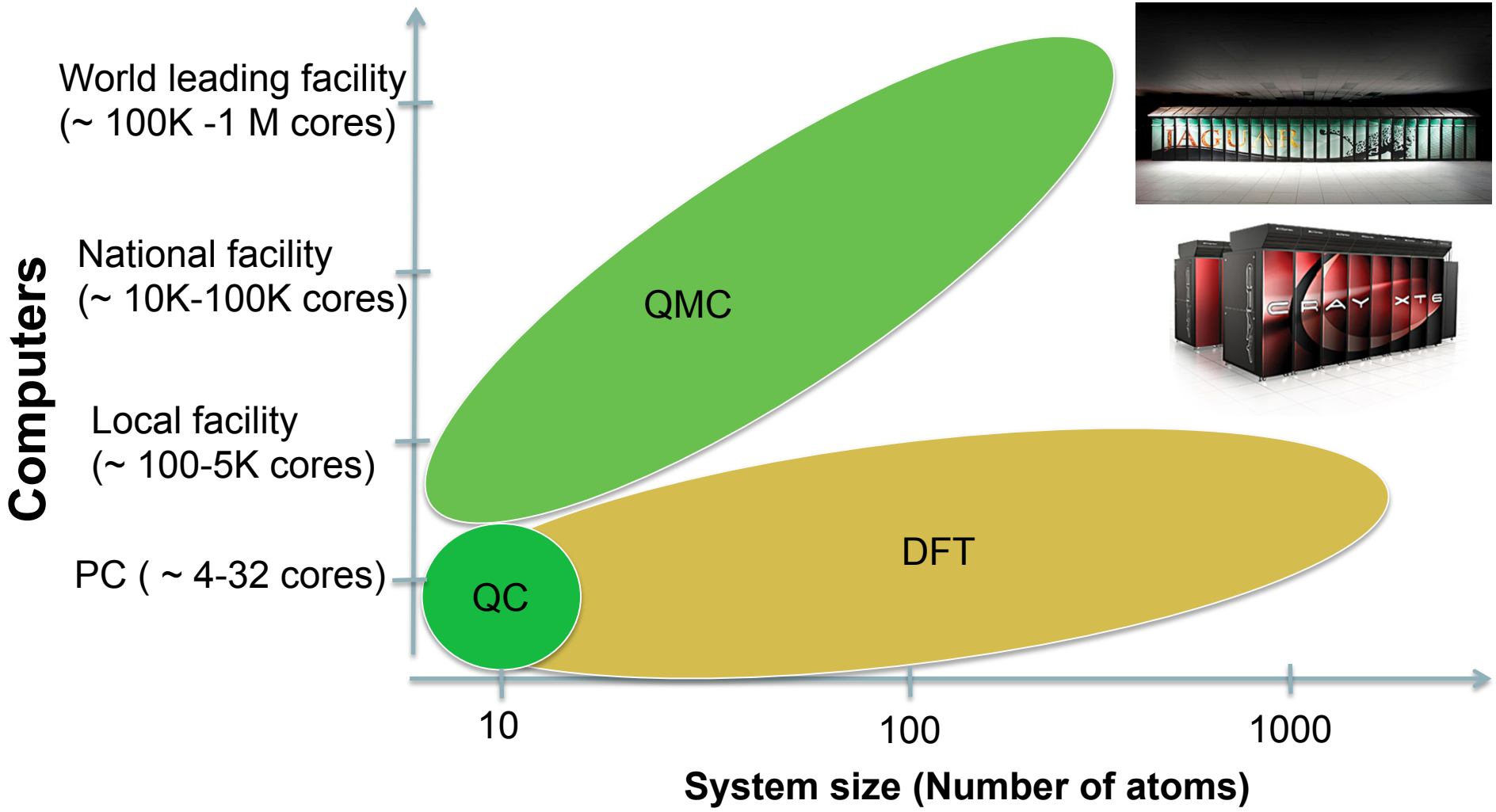
The many-body expansion:

$$E_{\text{tot}}(1,\ldots N) = \sum_{i} E^{(1)}(i) + \sum_{i<j} E^{(2)}(i,j) + \sum_{i<j<k} E^{(3)}(i,j,k) + \ldots$$

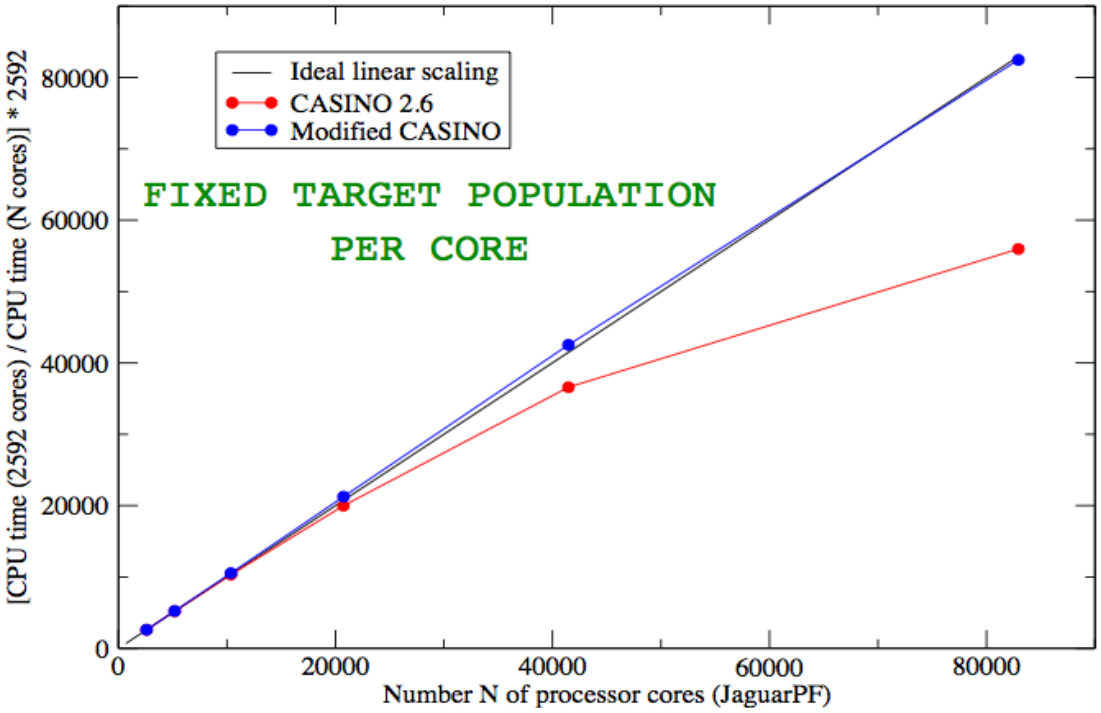
where $E^{(2)}(i,j) = E_{\text{tot}}(i,j) - E^{(1)}(i) - E^{(1)}(j)$, etc...

Cost and limitations of going beyond DFT

- Quantum Chemistry methods (MOLPRO code)
  - MP2 (**N^4**)
  - CCSD(T) (**N^7**), molecules and small clusters. Accuracy: within 1 meV of exact for water dimer.
- Quantum Monte Carlo (CASINO code)
  - Cost is **N^3** (same as DFT, but with a **big prefactor**)
  - Molecules and extended systems
  - Metals and insulators
  - Efficient use of large parallel machines


World leading facility
(~ 100K -1 M cores)
National facility
(~ 10K-100K cores)
Local facility
(~ 100-5K cores)
PC (~ 4-32 cores)
Computers
QMC
DFT
QC
10
100
1000
System size (Number of atoms)
JAGUAR
CRAY XT6

QMC scaling on JaguarPF (Cray XT6, 300,000 cores at ONRL)

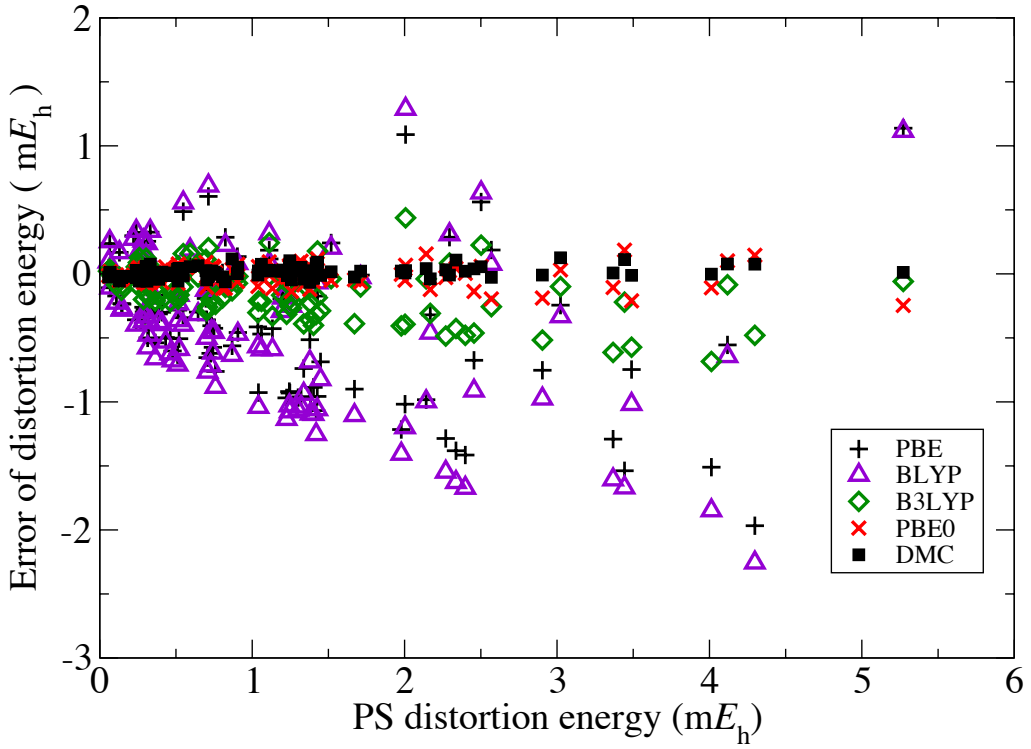


M.J. Gillan, M.D. Towler and D. Alfè, "*Petascale computing opens new vistas for quantum Monte Carlo*", Psi-k Highlight of the Month, February 2011.

The water monomer

The deformation energy of the H_2O molecule as function of bond lengths and bond angle is known essentially exactly from quantum chemistry

Some DFTs (e.g. PBE, BLYP) give a very poor description, but DMC is very good

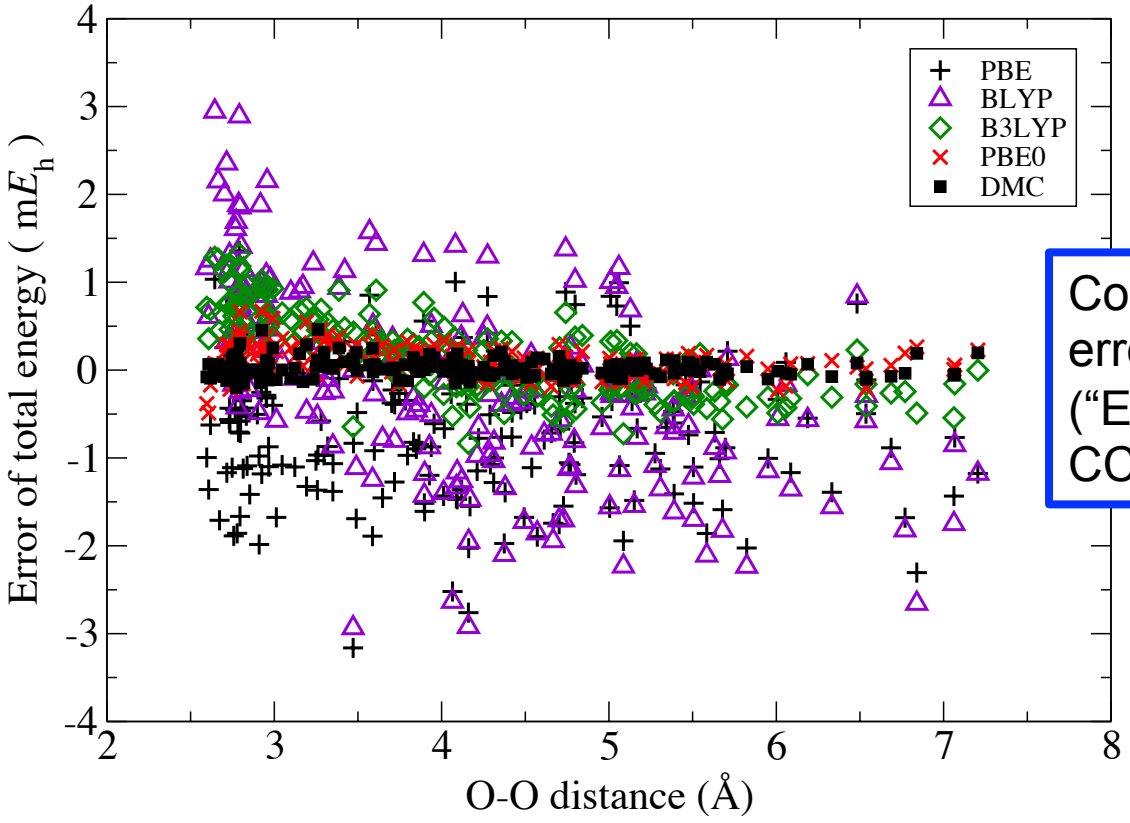


Error in approximate deformation energy vs benchmark

Benchmark PS is H. Patridge and D. W. Schwenke JCP, **106**, 4618 (1997)

The water dimer: DMC

Errors of total energy of H_2O dimer for thermal sample of 198 configurations drawn from AMOEBA m.d. simulation of liquid



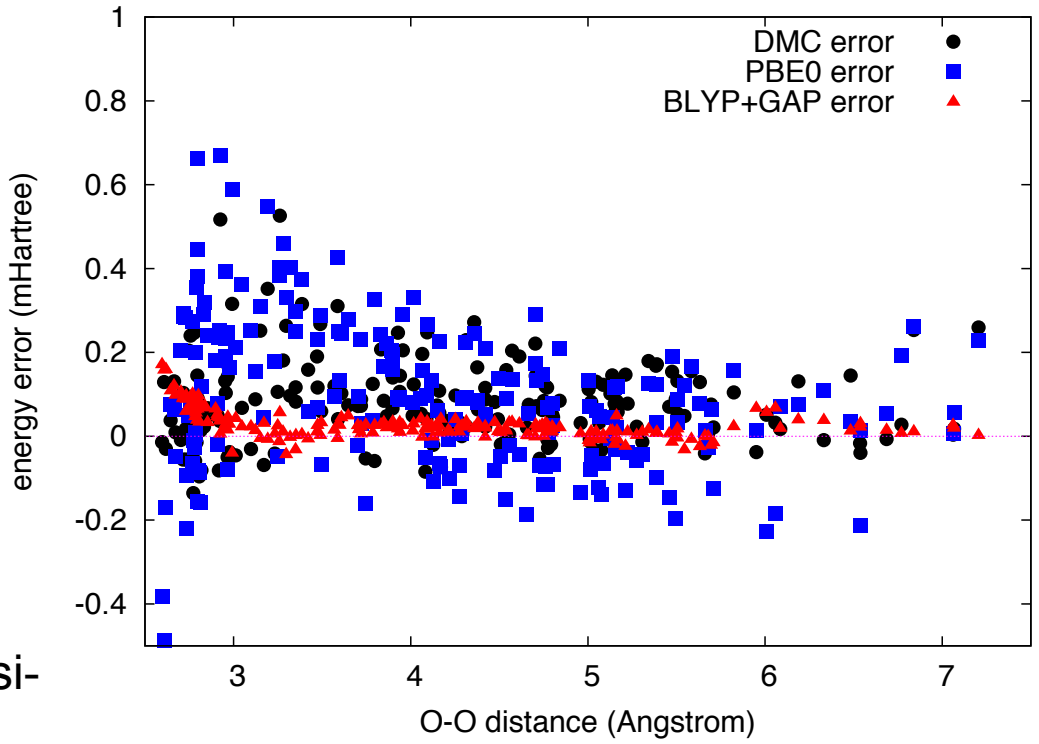
Comparison of DFT errors with DMC errors. ("Errors" = deviation from CCSD(T) at CBS limit.)

M. J. Gillan, F. R. Manby, M. D. Towler and D. Alfè, J. Chem. Phys. **136**, 244105 (2012)

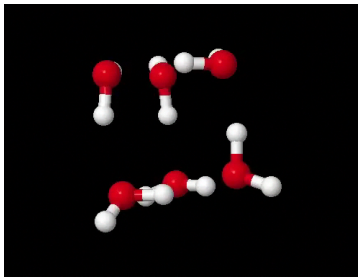
The water dimer: GAP

Correcting DFT with GAP (Csányi *et al.*)

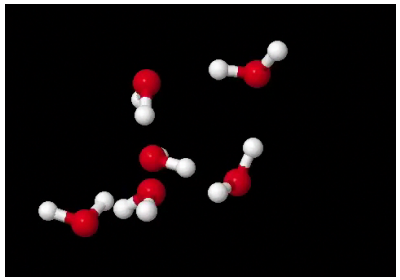
- Based on ideas of Bayesian learning from large databases of energies and forces of system of atoms and molecules: see Bartók, Payne, Kondor, Csányi, PRL (2010)
- For H_2O dimer: 12 degrees of freedom: descriptor space consists of intra- and inter-molecular distances (symmetry is important).
- First GAP for difference between DFT and MP2/ AVTZ, second GAP for basis-set corrections and CCSD(T)-MP2.
- Resulting DFT+GAP gives quasi-exact to within rms of ~ 1 meV.



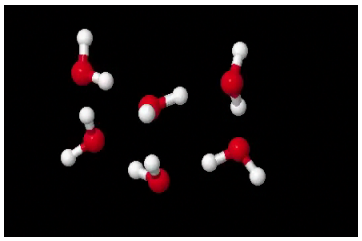
The water hexamer: four isomers



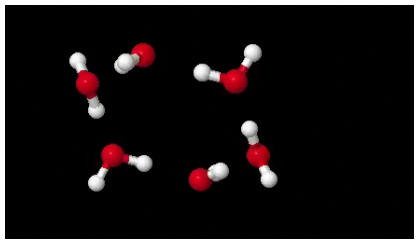
prism



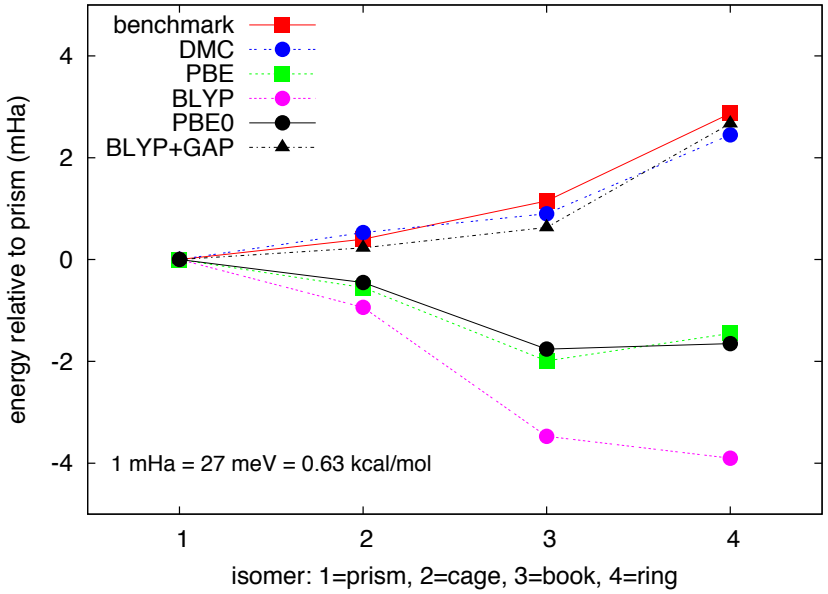
cage

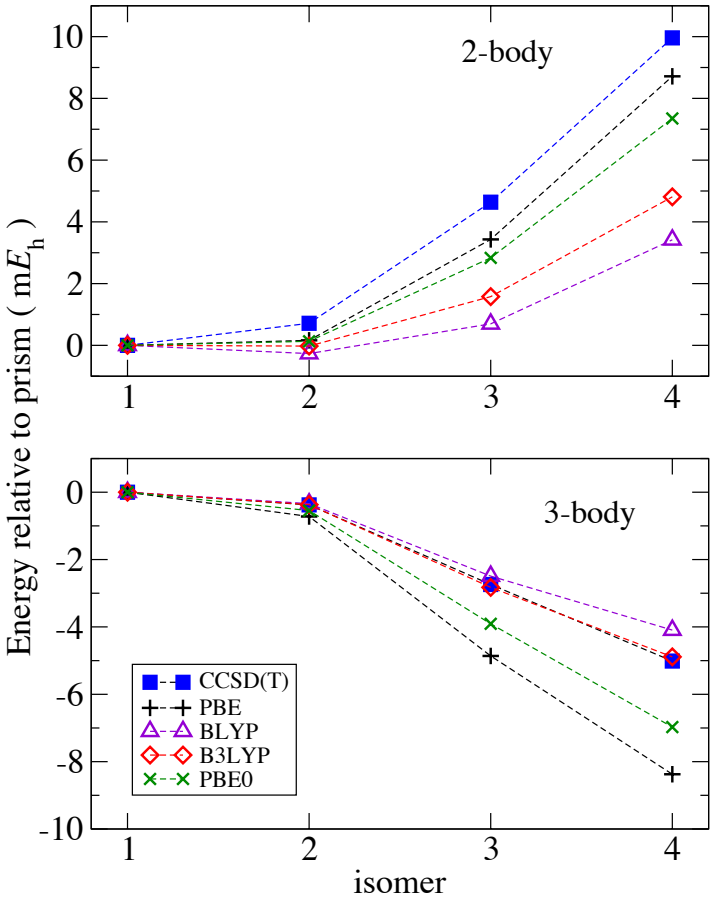


book

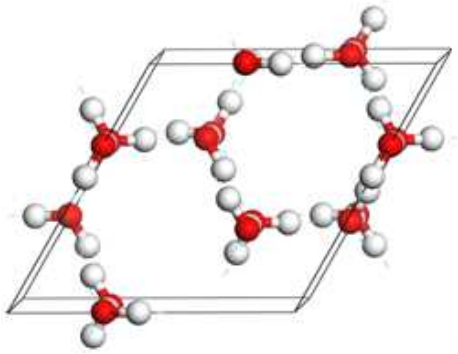


ring

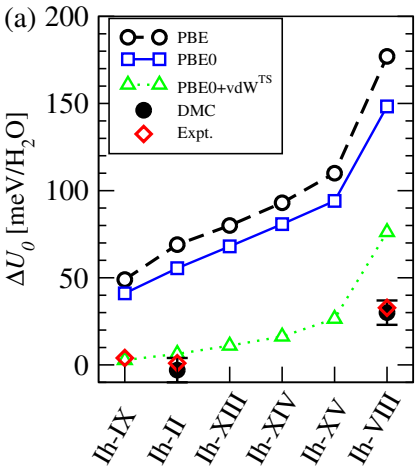




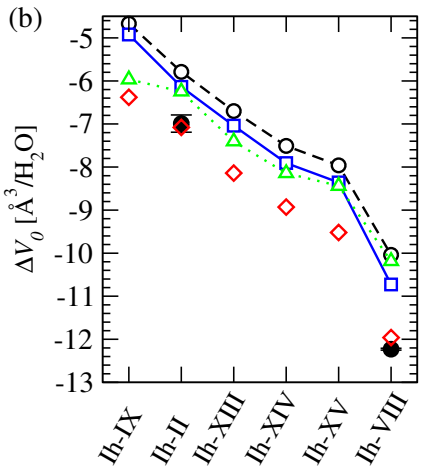
DFT and QMC for ice structures



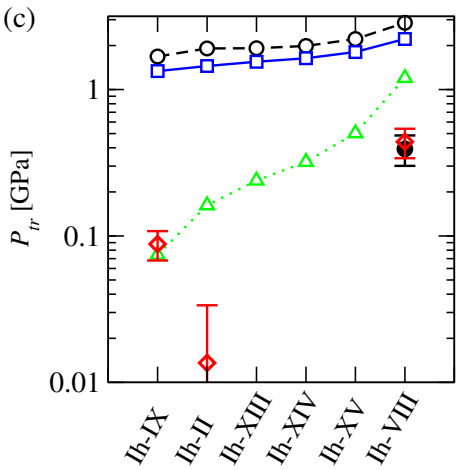
(a) Ih (12 water)



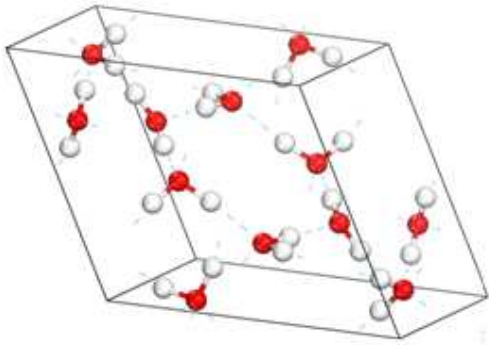
Relative energies



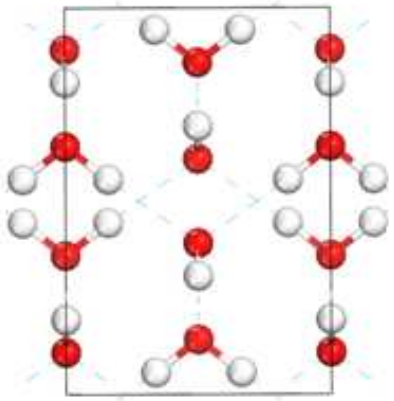
Volumes



Transition pressures



(d) II



h) VIII

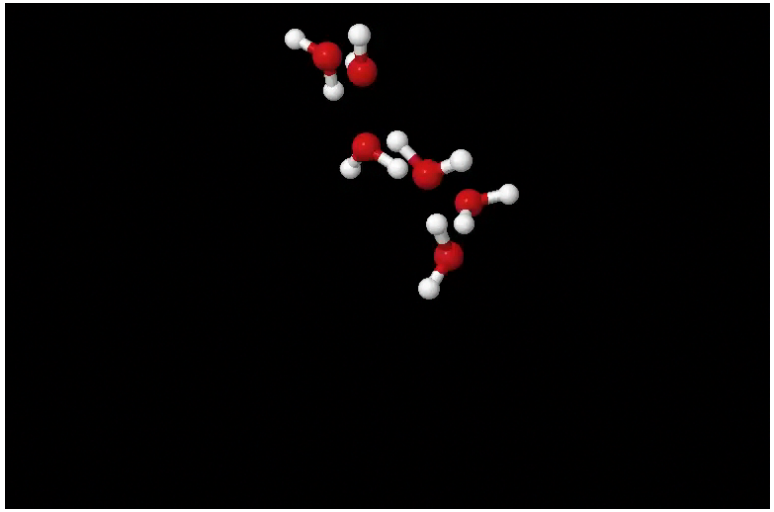
	Ih	II	VIII
Expt	-610	-609	-577
DMC	-605	-609	-575
PBE	-636	-567	-459
PBE0	-598	-543	-450
PBE0+vdW	-672	-666	-596

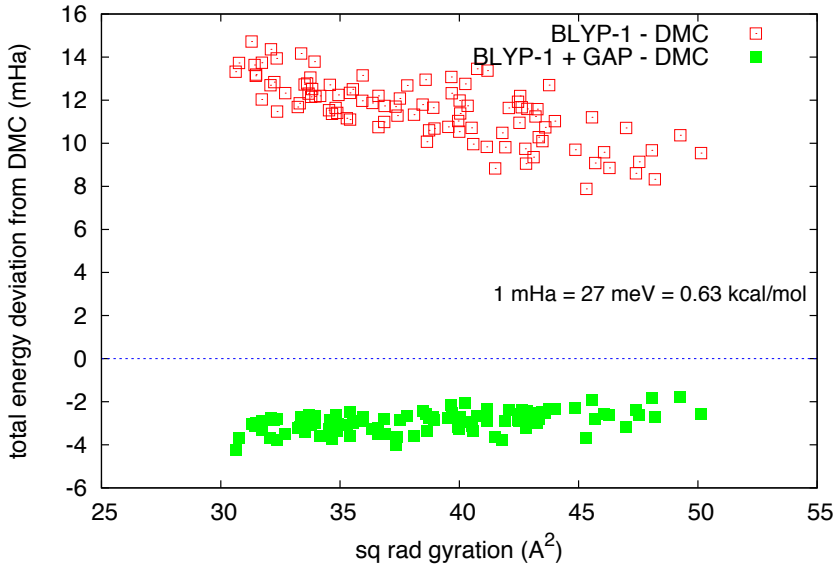
Santra et al. PRL (2011)

Cohesive energies of ice Ih, II and VIII (meV/H_2O)

Hexamer in thermal equilibrium

M.d. simulation with Fanourgakis-Xantheas classical potential (flexible polarisable): T = 200 K, duration = 1 ns, configurations taken every 10 ps.





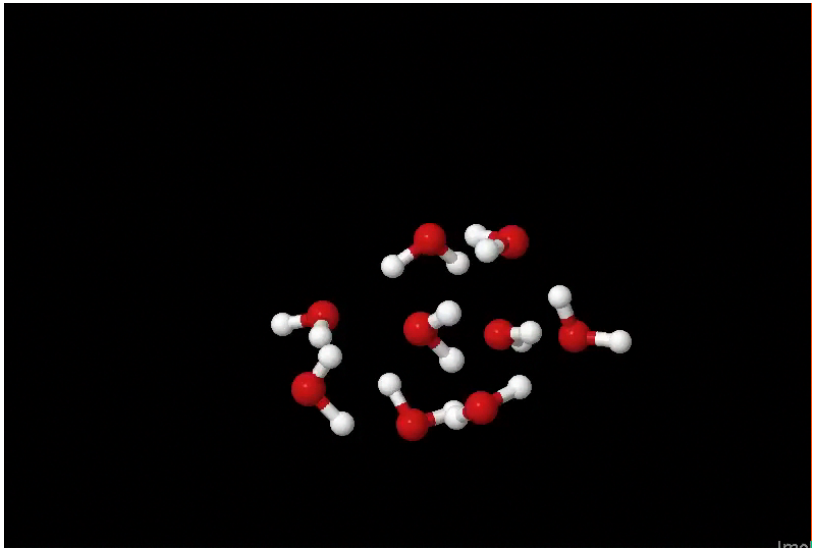
Errors of BLYP (using DMC as benchmarks)
Red points: BLYP corrected for 1-body errors
Green points: BLYP corrected for 1-body and 2-body

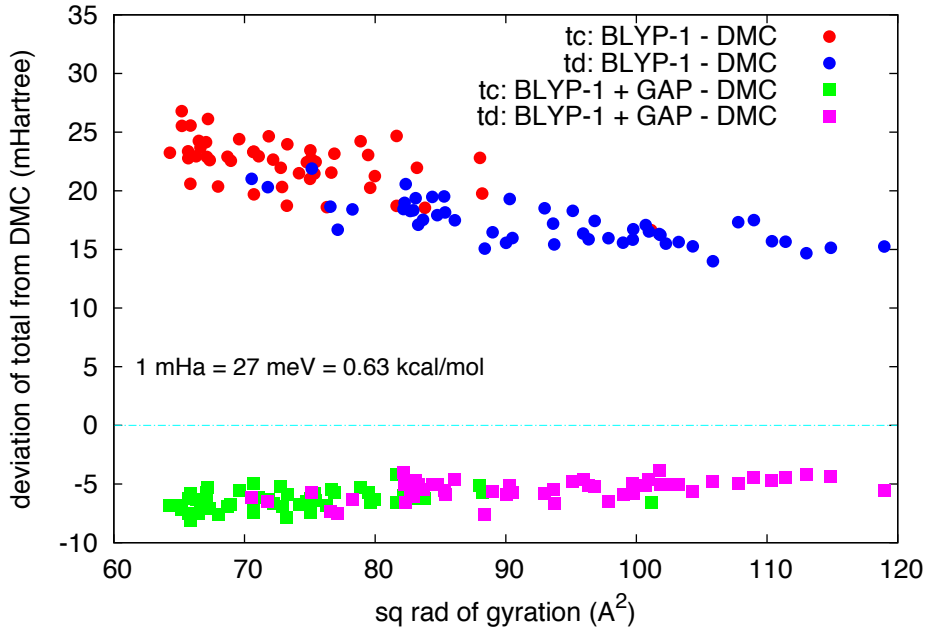
Key points:

- BLYP-1 seriously underbinds
- Underbinding decreases as cluster expands, so BLYP-1 errors make cluster expand in thermal equilibrium
- GAP 2-body correction gives BLYP-2: slight overbinding, must be due to beyond-2-body
- BLYP-2 errors almost config-indep, so do not affect thermal-eqm structure or dynamics.

The nonamer in thermal equilibrium

M.d. simulation with Fanourgakis-Xantheas classical potential (flexible polarisable): T = 200 K, duration = 1 ns, configurations taken every 10 ps.





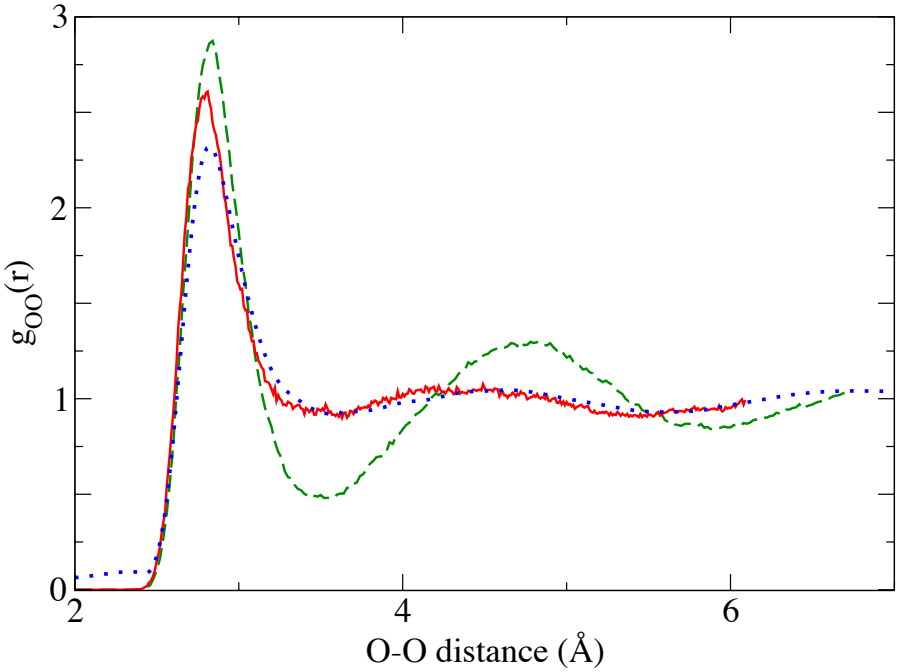
Errors of BLYP (using DMC as benchmarks)
Red and blue: BLYP corrected for 1-body
Green and magenta: BLYP corrected for 1- and 2-body

Similar to hexamer:

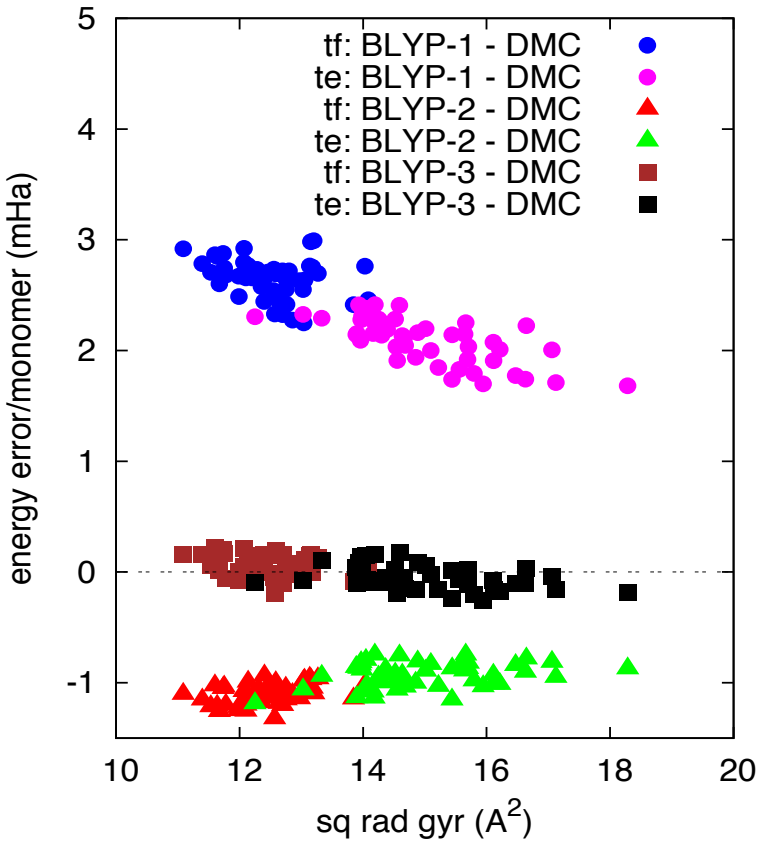
- BLYP-1 greatly underbinds, makes cluster expand
- BLYP-2 = BLYP-1 + GAP somewhat overbinds, because of beyond-2-body errors
- BLYP-2 errors almost indep of config, so ok for thermal eq structure and dynamics

Liquid water

Liquid water with BLYP-2:



The pentadecamer in thermal equilibrium



Where are the errors in DFT functionals?

- There is no single answer – it depends on the DFT
- For BLYP, dominant errors are in 1-body and 2-body; beyond-2-body error is not negligible, but it depends only weakly on configuration. So using GAP to correct 1- and 2-body gives big improvement.
- For PBE, dominant errors appear to be in 1-body and beyond-2-body. So using GAP to correct 2-body is unlikely to achieve much – we will test this expectation.

Where so far? – Where next?

- Must achieve overall description: clusters, crystal, liquid
- QMC succeeds where DFT fails
- Water: the subtle balance between 2-body and beyond-2-body is crucial
- Normal DFT fails, DFT+GAP is accurate for clusters, crystal, liquid – with judicious choice of DFT
- The concept of “statistical benchmarking”: QMC for large thermal samples used to test and calibrate other methods
- Will generalise to many other molecular systems: CH_4, NH_3, HF, $H_2O + CH_4$, $H_2O + HF \rightleftharpoons H_3O^+ + F^-$, …

Project made possible by large allocation of time on JaguarPF and Titan under INCITE program