Computing Needs to Model Chemical Transformations: A DOE/BES Perspective

Bruce C. Garrett

Chemical & Materials Sciences Division Director

Thanks to

Chris Mundy

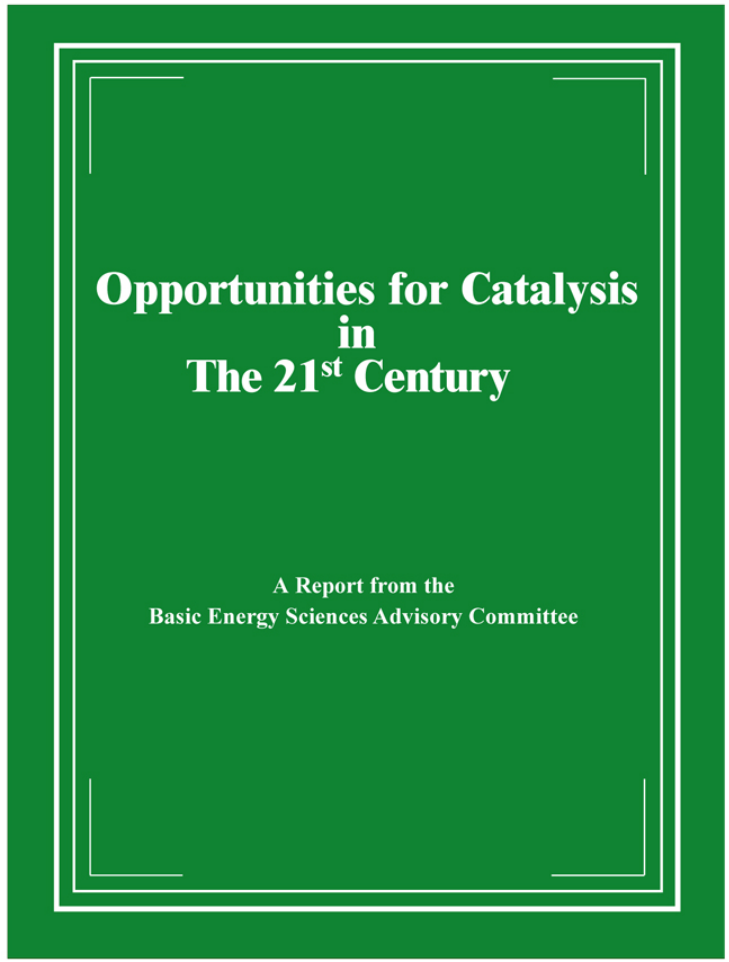
Greg Schenter

Marat Valiev

Sotiris Xantheas

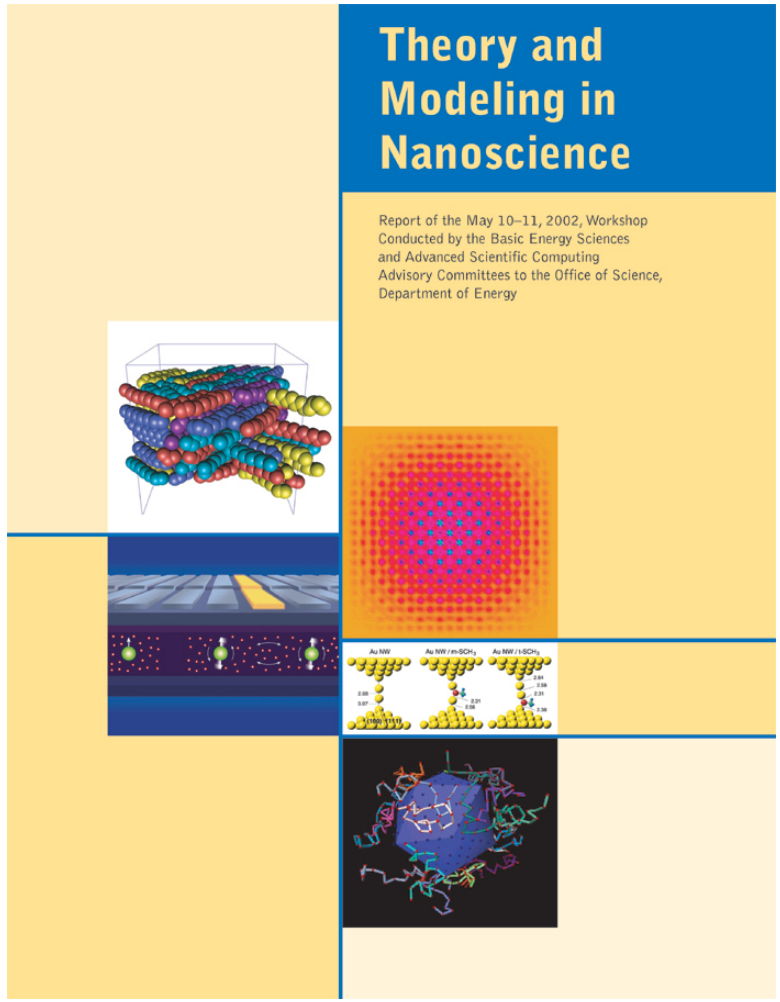
BES workshop on Opportunities for Catalysis - May 2002

- "… the Grand Challenge for catalysis science in the 21st century is to understand how to design catalyst structures to control catalytic activity and selectivity"
- "When scientists and engineers understand how to design catalysts to control catalytic chemistry, the effects on energy production and use and on the creation of exciting new materials will be profound."
- "At the heart of catalysis is the control of chemical transformations, and the ability to predict reaction rates is key to gaining a fundamental understanding of catalytic processes. Thus, computational studies need to put structure and dynamics on *an equal footing*."



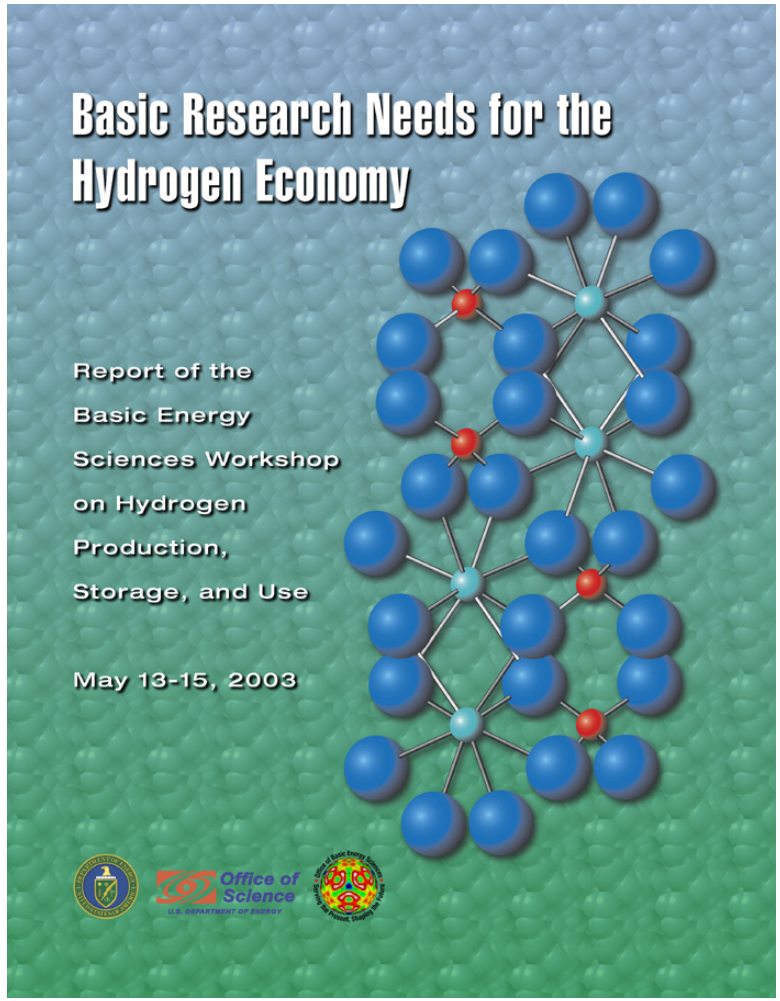
BES and ASCR workshop on T&M in Nanoscience - May 2002

- "Within five to ten years, there must be robust tools for quantitative understanding of structure and dynamics at the nanoscale, without which the scientific community will have missed many scientific opportunities as well as a broad range of nanotechnology applications."
- "Challenges & opportunities include … to simulate self-assembly, the key to large-scale production of novel structures, which typically involves many temporal and spatial scales and many more species than the final product."



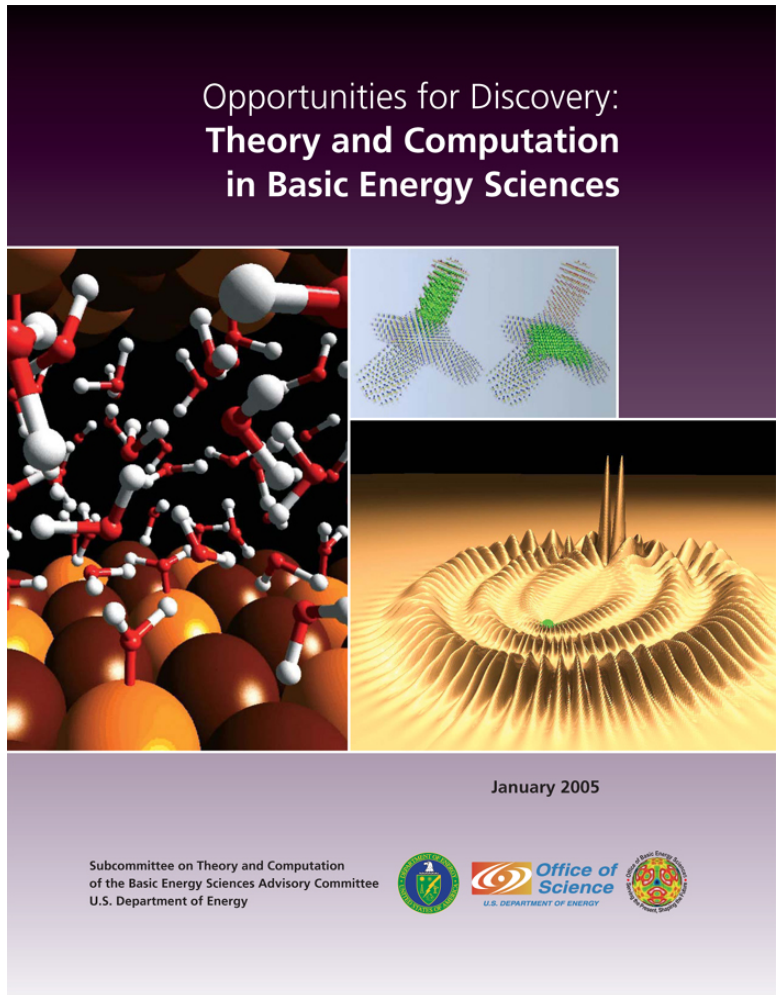
Basic Research Needs for the Hydrogen Economy - May 2003

- "Catalytic performance is a limiting factor for many essential elements of the hydrogen economy (including fuel cell efficiency, storage kinetics, and production capacity) … A fundamental understanding of the chemical and physical processes involved in catalysis at the atomic level in any of these areas would drive collateral progress in all three …"
- "Theory, simulation, and computation closely coupled with well-designed experiments are critical for bridging the vast gap between the present state of the art and the science and technology needed for a hydrogen economy."
- "… the challenge for theory is to develop a fundamental understanding of the controlling factors of catalytic reaction mechanisms from quantum chemical, kinetics, atomistic, and continuum modeling."



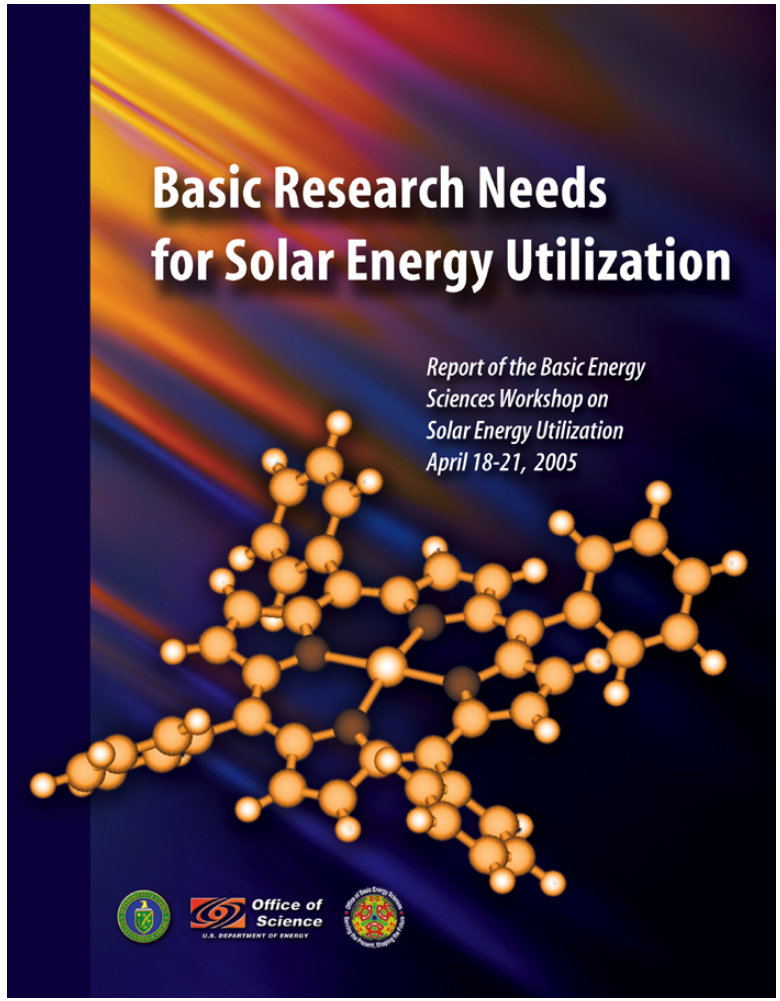
Report from Subcommittee on Theory & Computation in Basic Energy Sciences January 2005

- "… taking advantage of newly emerging opportunities in theory and computation will increase the pace of discovery across the entire scientific scope of the BES portfolio."
- "An even larger number of scientific opportunities are associated with <u>developing computational tools required to understand and control chemical reactions in condensed-phase environments and at interfaces</u>."



Basic Research Needs for Solar Energy Utilization - April 2005

- "… cross cutting issues include … finding catalysts that enable inexpensive, efficient conversion of solar energy into chemical fuels"
- "Solar energy conversion processes … are characterized by activated catalytic processes, which cannot be simulated on the short time scale of molecular dynamics simulations. In this case, approaches like first-principles molecular dynamics … need to be supplemented by approaches for finding chemical reaction pathways both at zero and at finite temperature."



Basic Research Needs for Advanced Nuclear Energy Systems - Aug 2006

- "The fundamental challenge is to understand and control chemical and physical phenomena in multi-component systems from femto-seconds to millennia, at temperatures to 1000ºC, and for radiation doses to hundreds of displacements per atom"
- "… understanding of electron-driven radiation effects, particularly those involving water, would clearly benefit from further fundamental modeling studies."
- "The scientific challenge is to <u>develop a predictive capability (e.g., chemical accuracy for equilibrium constants and rate constants) for modeling solutions and interfacial phenomena</u> for actinide-containing systems under extreme conditions of temperatures, pH, and high radiation fields"

Basic Energy Sciences Research Needs

- Series of BES reports identify need for advances in theory, modeling, and simulation
- Prevalent need for advances to understand chemical transformations in complex systems
- Processes in condensed phases and interfacial systems are critical in a number of areas important to DOE:
  - Hydrogen production, storage, and use
  - Solar energy utilization
  - Advanced concepts for nuclear energy
- Catalysis (a.k.a. chemical reactions) is a cross-cutting theme in all these areas

Modeling Chemical Transformations: What are the Challenges?

- Reaction Energetics
  - Classical potentials (molecular mechanics) are inadequate to treat chemical reactions
  - Accurate energetics of bond-breaking/making require high-order electronic structure methods (i.e., with N^6 scaling or higher) and are difficult for extended systems (e.g., condensed phases & interfaces)
- Reaction Dynamics
  - "Rare events" are difficult to identify without prior knowledge of the potential energy landscape
  - Accurate kinetics and thermodynamics requires sufficient sampling - need to get the statistical mechanics right
  - Quantum nuclear effects (e.g., ZPE and tunneling) increase complexity; accurate quantum dynamics limited to systems with several atoms

Reaction Energetics in Extended Systems: Advances in Computational Tools

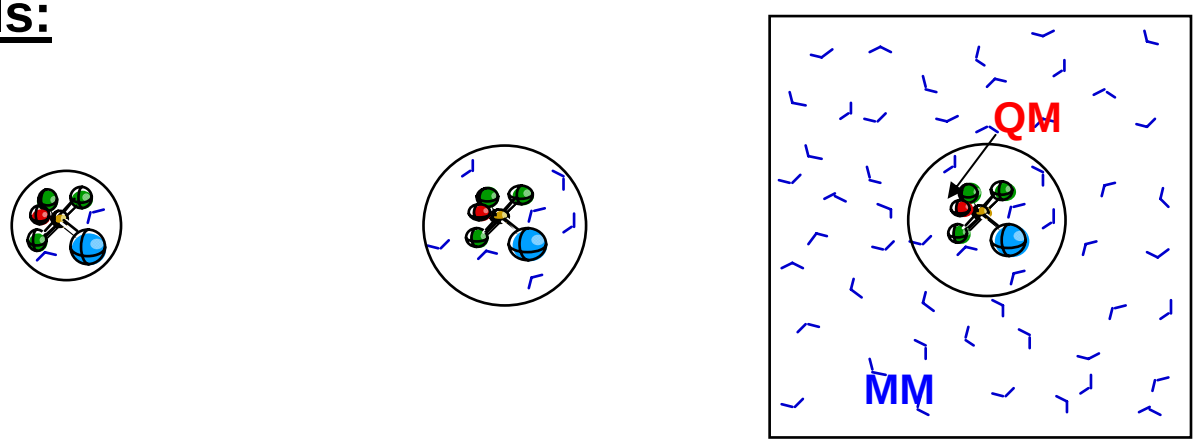
- QM/MM
  - Quantum mechanical treatment of reactive region
  - Molecular mechanics treatment of ‘environment’
  - Self consistent optimization of QM and MM regions
- Layered methods such as ONIOM - Svensson et al. JCP 100, 19357 (1996); ONIOM-PCM - Vreven et al. JCP II5, 62 (2001)
  - Hierarchy of models and levels of theory used to approximate accurate energetics for largest system
- Multiconfiguration molecular mechanics - Kim et al. JCP 112, 2718 (2000); Tishchenko & Truhlar JPCA 110, 13530 (2006)
  - Electronic structure calculations for limited configurations used to parameterize a multi-state EVB-type model
  - Uses interpolation schemes such as Shepard interpolation [Ischtwan & Collins, JCP 100, 8080 (1994)] of ab initio energies and Hessians

Heirarchical Approach to Reaction Energetics

- Combine advantages of layered and QM/MM approaches, e.g., Vreven et al. J. Chem. Theory Comp. 2, 815 (2006)
- Can be systematically improved by applying progressively higher levels theory to larger model systems

Example of 3 levels:

Model:



Theory:	CCSD(T)/CBS MP2/aDZ	MP2/aDZ SCF/6-31G*	QM=SCF/6-31G*

Energetics in Extended Systems: Future Needs for Petascale Computing

- Efficient parallel algorithms for accurate electronic structure calculations - pushing to more accurate levels of theory on larger systems
- Improved MM methods that mimic quantum mechanical methods more reliably (needed for statistical sampling discussed later)
- Implementation of adaptive procedures to refine analytical functional forms ‘on the fly’
  - Parallel adaptation, e.g., parallel trajectories creating a common potential energy surface

The Need for Rare Event Discovery Tools

- Relatively simple systems can exhibit complex, collective dynamics that is difficult to predict *a priori*

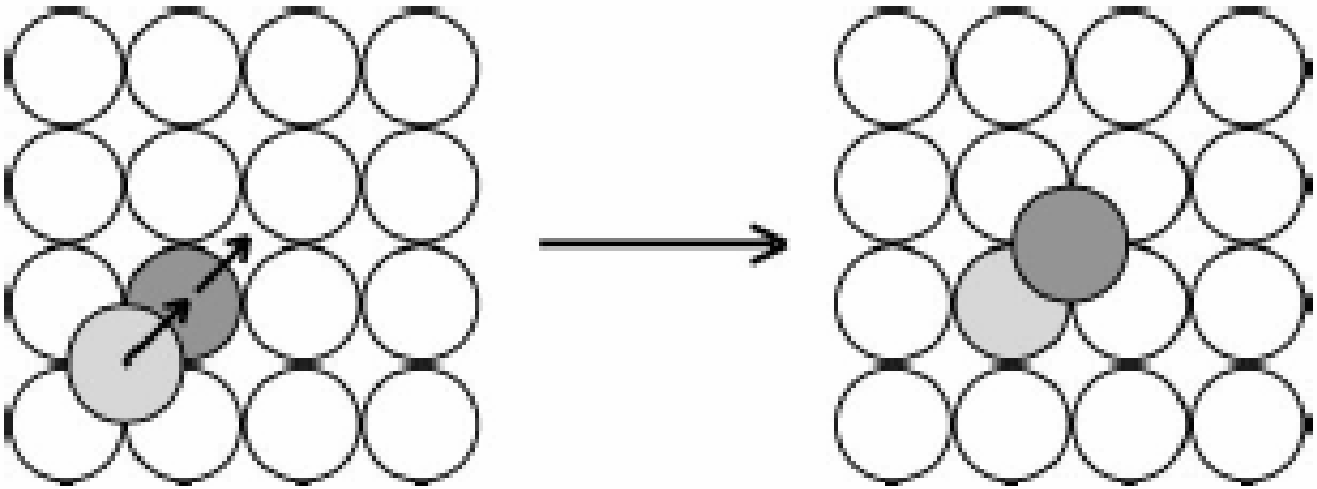


Figure 3 Adatom exchange mechanism on an fcc(100) surface. This, as opposed to a jump, is the preferred adatom diffusion mechanism on a number of fcc metals, demonstrating the complexity of even simple atomistic systems.

Voter, Montalenti and Germann, Annu. Rev. Mater. Res. ***32****, 321, 2002*

Rare Event Discovery: Advances in Computational Tools

- Accelerated molecular dynamics - Voter et al. Ann. Rev. Mat. Res. 32, 321 (2002)
  - Hyperdynamics - bias potential raises energies in wells, not at transition states
  - Parallel replica dynamics - transition detection scheme that is based on molecular topology, not basins/barriers; many trajectories run in parallel
  - Temperature accelerated dynamics - raise temperature, but only allow processes that would occur at original temperature; uses harmonic TST to correct dynamics to original temperature.
- Metadynamics [Iannuzzi et al. PRL 90, 238302 (2003)]
  - Time-dependent self-adapting bias potential fills in wells - repulsive Gaussians that are functions of collective variables
  - Recovers free energy surface of the complex system

Rare Event Discovery: Future Needs for Petascale Computing

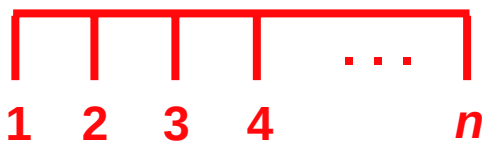
- Implementation of methods to accelerate dynamics in molecular dynamics algorithms
  - Parallel MD approaches with efficient scaling needed for MM potentials
  - Efficient scaling of MM potentials

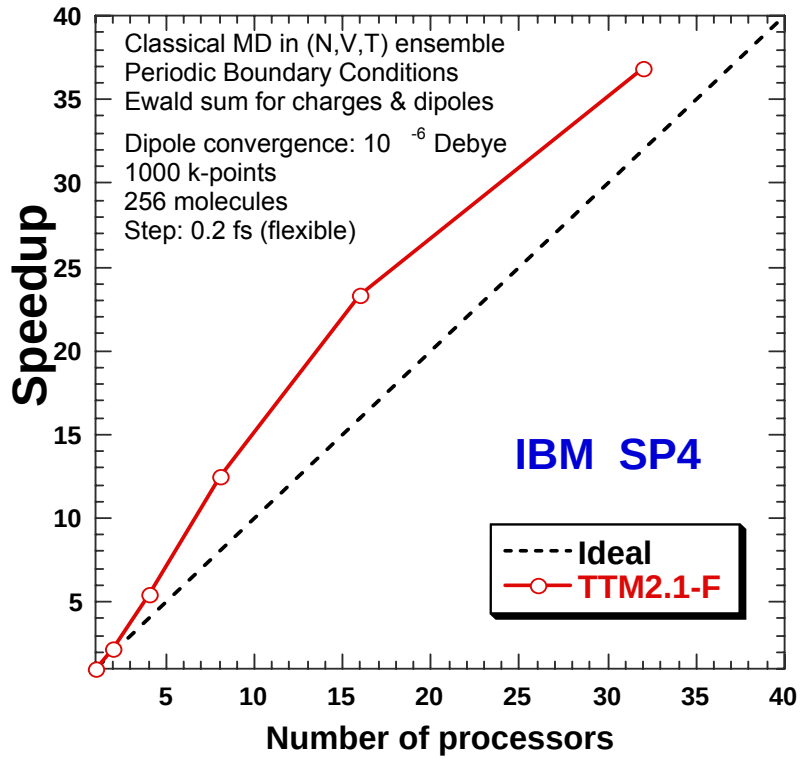
Parallel Implementation of the TTM2-F (v. 2.1) Potential for Water

Fanourgakis, Tipparaju, Nieplocha and Xantheas, *Theor. Chem. Acc.* **117**, 73 (2007)

Global arrays

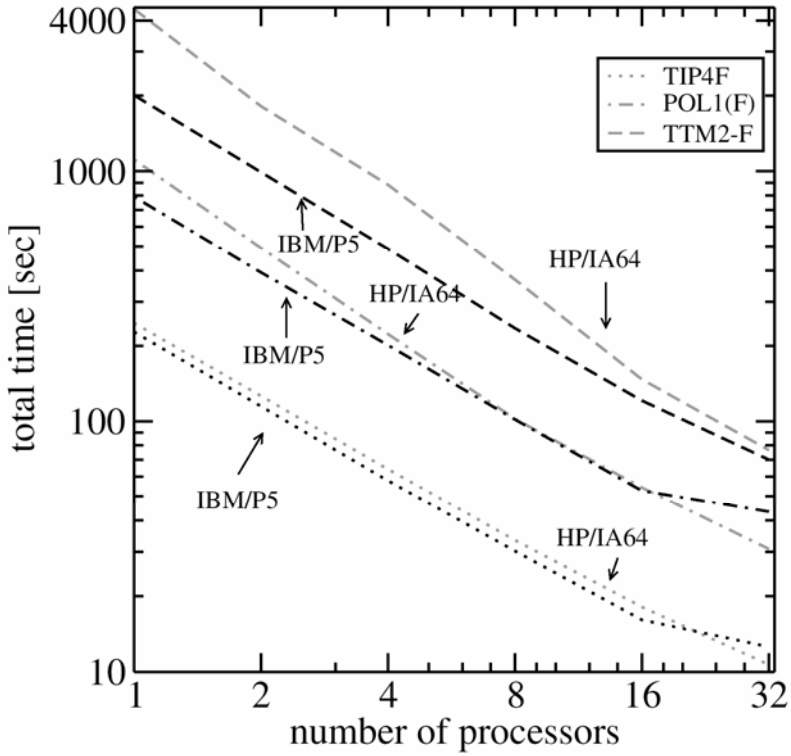
Nieplocha *et al.* (PNNL)





Time to solution

comparison with other classical potentials



Only **implementation to date of a POLARIZABLE force field which exhibits scaling for > 8 processors**

Reaction Pathways and Energetics in Condensed Phases: Advances in Computational Tools

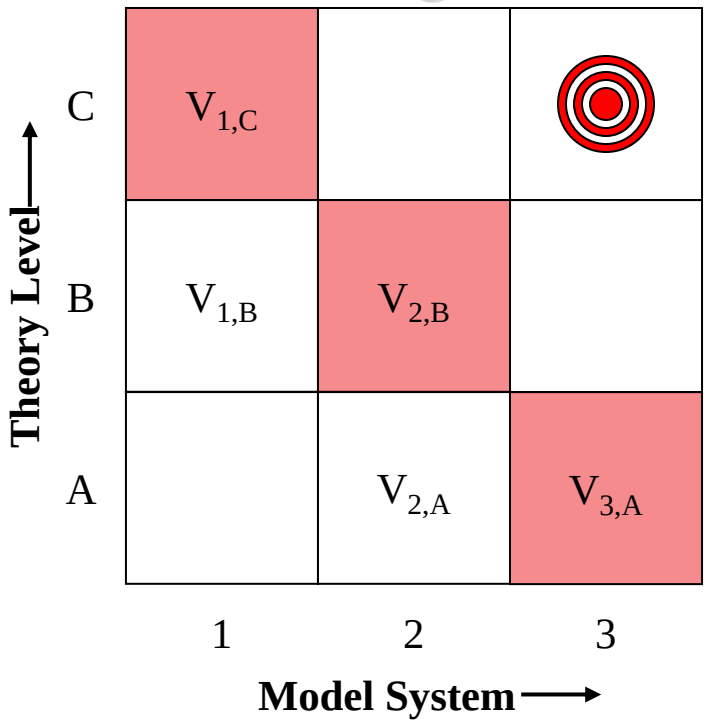
- Potential of mean force (PMF) for predefined coordinates includes the average affect of 'solvent' on reaction energies
  - Predefined coordinates typically taken as the gas-phase reaction path or minimum energy path (MEP)
  - Condensed phase can alter reaction path and even the mechanism
  - Efficient algoritms for PMF with QM/MM methods
- MEPs on PMFs
  - Large scale QM/MM optimizations
  - Nudged elastic band method for MEP
- Metadynamics can be used to determine free energy surfaces and collective coordinates important in the reaction

PMF Calculations with Hierarchical Approach to Energetics

$$
\begin{aligned}
V_{3,C}(\mathbf{x},\mathbf{S}) &\approx V_{1,C}(\mathbf{x}) - V_{1,B}(\mathbf{x}) \\
&\quad + V_{2,B}(\mathbf{x}) - V_{2,A}(\mathbf{x}) \\
&\quad + V_{3,A}(\mathbf{x},\mathbf{S}) \\
\partial W(\mathbf{x})/\partial \mathbf{x} &= \langle \partial V_{3,C}/\partial \mathbf{x} \rangle_{\mathbf{S}} \\
&\approx \partial V_{1,C}/\partial \mathbf{x} - \partial V_{1,B}/\partial \mathbf{x} \\
&\quad + \partial V_{2,B}/\partial \mathbf{x} - \partial V_{2,A}/\partial \mathbf{x} \\
&\quad + \langle \partial V_{3,A}/\partial \mathbf{x} \rangle_{\mathbf{S}}
\end{aligned}
$$

x = coordinates of reactive system

S = solvent coordinates



- $V_{1,C}$ and $V_{2,B}$ are computationally intensive, but only need to be calculated once for each x since average over solvent configurations (S) is for fixed x.
- $V_{3,A}$ involves QM calculations in solvent averages so we require an efficient numerical algorithm to reduce the number of QM evaluations

Efficient Algorithm for PMFs with QM/MM

- PMFs for the S_N2 Reaction CH_3Cl + OH^- using the QM/MM : MM/MM protocol:

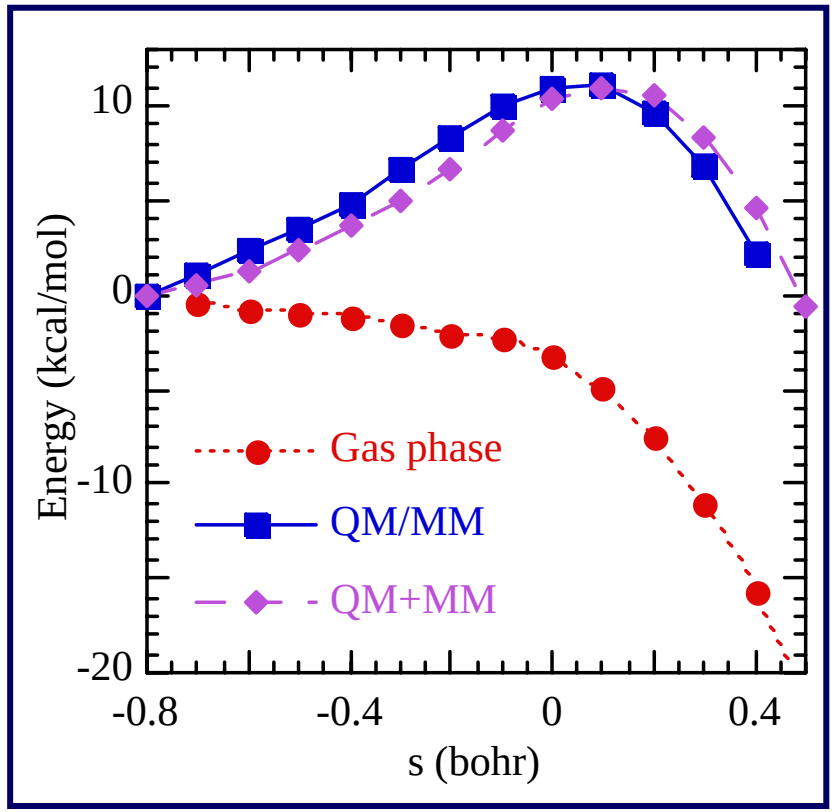
  $\langle \exp\{-(E_J^{QM/MM} - E_I^{QM/MM})/kT\}\rangle_{(I)} \approx$

  $\langle \exp\{-(E_J^{MM/MM} - E_I^{MM/MM})/kT\}\rangle_{(I)} \times$

  $\langle \exp\{-(E_J^{QM/MM} - E_J^{MM/MM})/kT\}\rangle_{(I)} /$

  $\langle \exp\{-(E_I^{QM/MM} - E_I^{MM/MM})/kT\}\rangle_{(I)}$

- MM/MM-only averages use millions of configurations, averages involving QM/MM require significantly less configurations (100s of QM calculations)

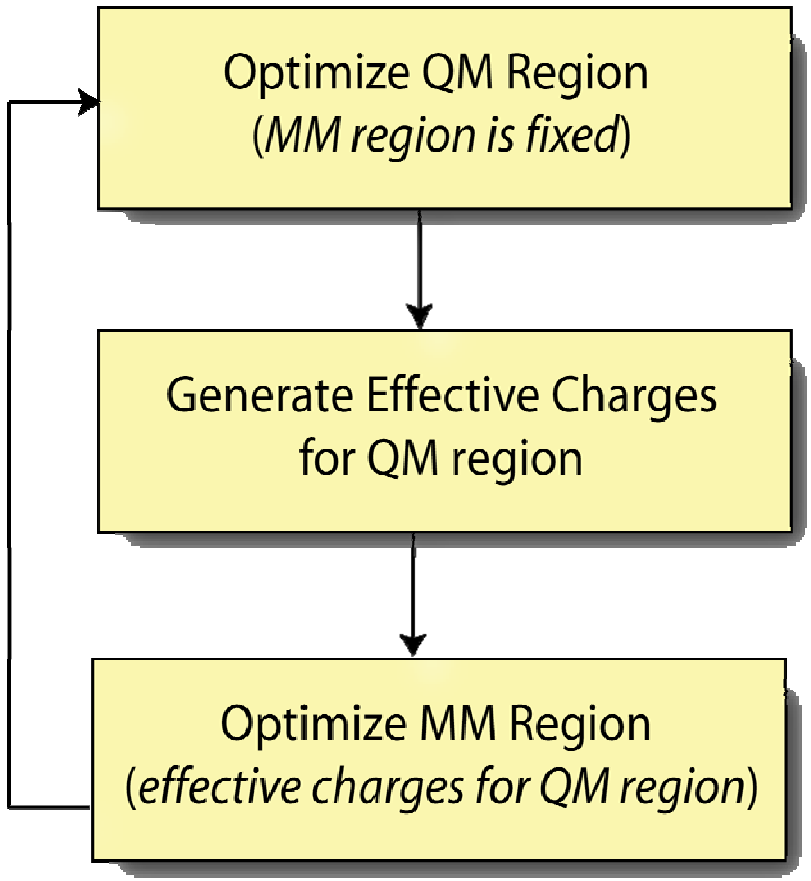


Dupuis, Schenter, Garrett, Arcia, J Mol. Struc. - Theochem 632, 173 (2003)

Large Scale QM/MM Optimizations

Marat Valiev

- Computational Challenges
  - many optimization steps due to many degrees of freedom (10^3-10^6)
  - each step requires expensive evaluation of QM region
- Key Observations
  - most of degrees of freedom are in MM region
  - movement of MM atom affects little the electronic structure of the QM region
- Algorithm
  - alternating optimizations of QM and MM regions
  - efficient relaxation of MM region based on effective charge model of the QM region
  - self-consistent loop ensures correct optimized geometry



NEB-QM/MM approach for Reaction Pathway Calculations

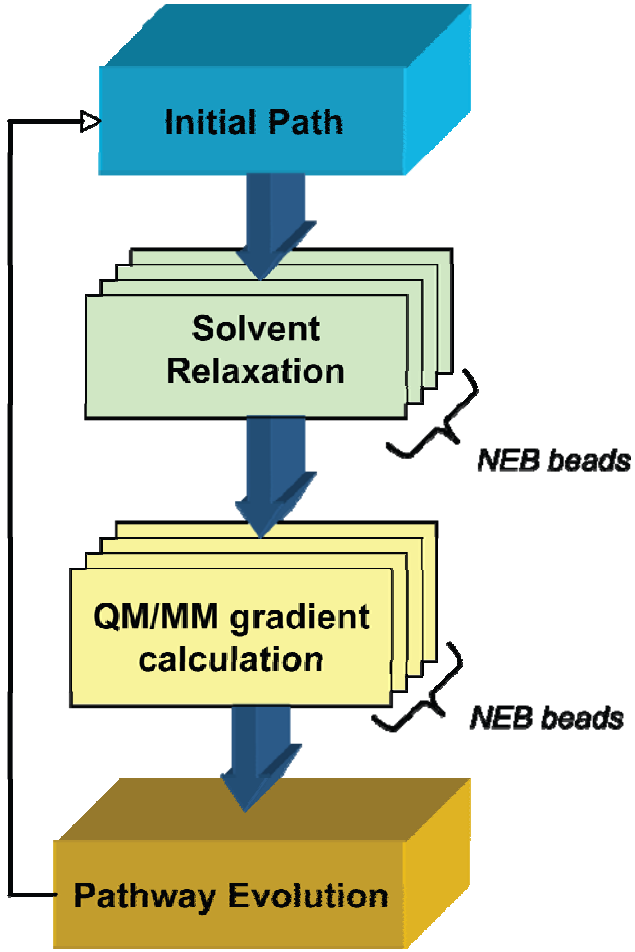
- Separation into “solute” (reaction region) and “solvent” degrees of freedom
- Equlibrium “solvent” representation

$$W(\mathbf{R}_a,\beta) = -\frac{1}{\beta}\ln\int e^{-\beta V(R)}\delta(\mathbf{R}-\mathbf{R}_a)d\mathbf{q}$$

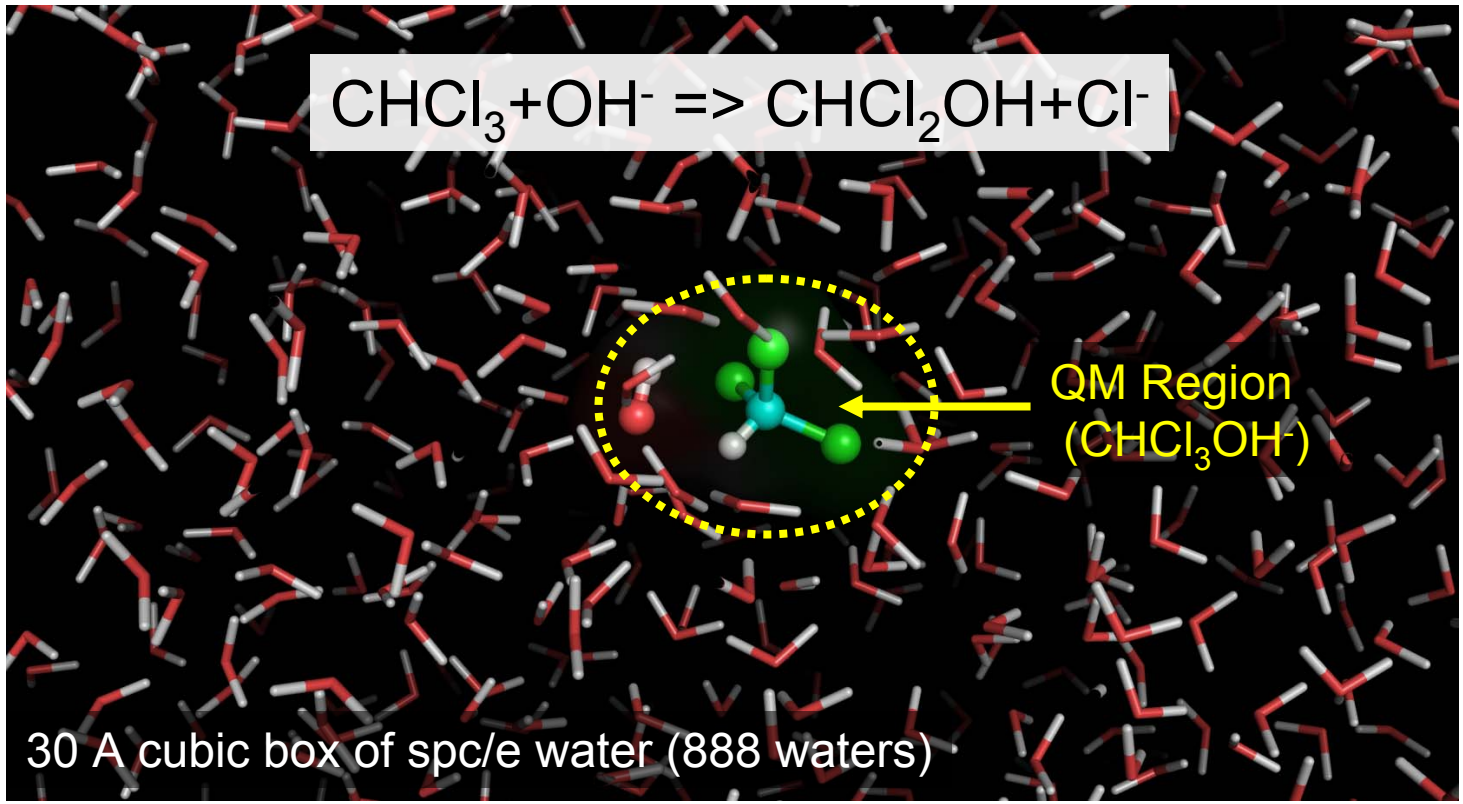
- NEB representation of reaction pathway

$$W^{NEB} = \sum_{n=1}^{Nb} W\left(\mathbf{R}_a^n\right) + \sum_{i=1}^{Nb} k\left|\mathbf{R}_a^n - \mathbf{R}_a^{n-1}\right|^2$$

- Current implementation is based on zero temperature limit, finite temperature implementation is in progress

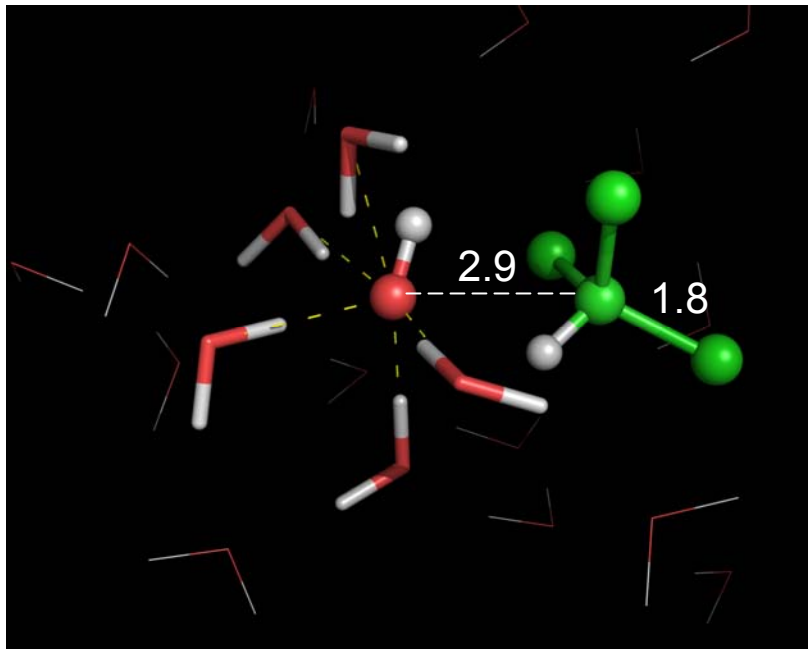


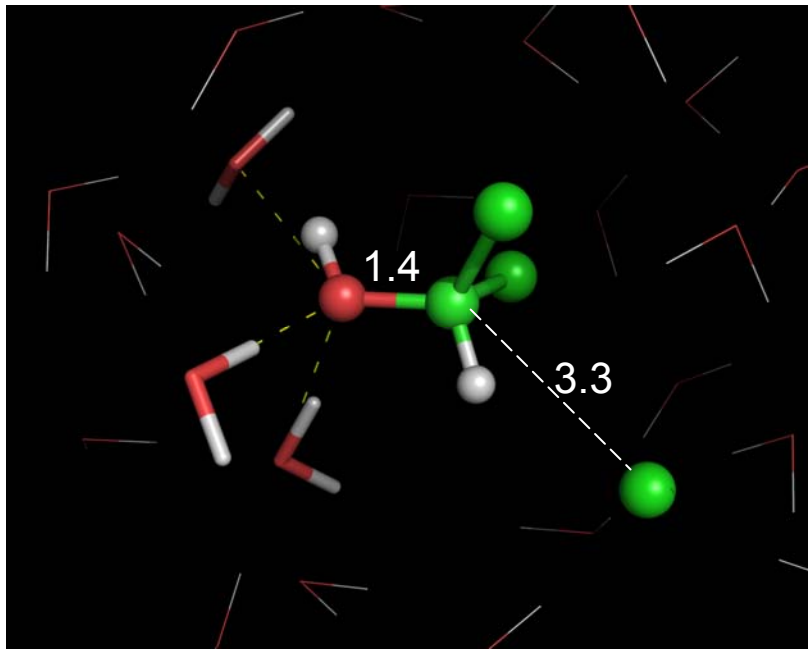
Application to Solution Chemistry



All calculations were performed using NWChem QM/MM module

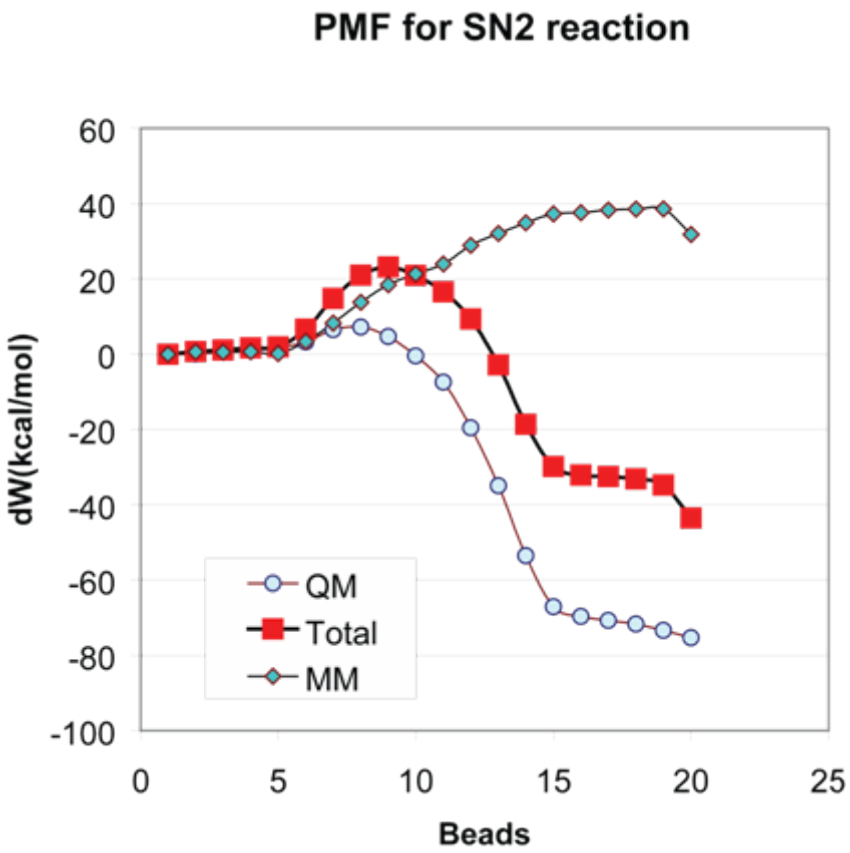
Optimizations





Optimized QM/MM structures of reactant and product states
DFT/B3LYP calculations

Free Energy over NEB reaction pathways



Reaction Pathways and Energetics in Condensed Phases: Future Needs for Petascale Computing

- MC or MD algorithms for statistical sampling to converge ensemble averages
  - 10^6 or more configurations need to be sampled to give reasonable statistics in an ensemble average
  - Need efficient algorithms to reduce number of QM evaluations in averages
  - 2-level parallelization possible: parallel jobs of independent simulations that parallelize the electronic structure calculations

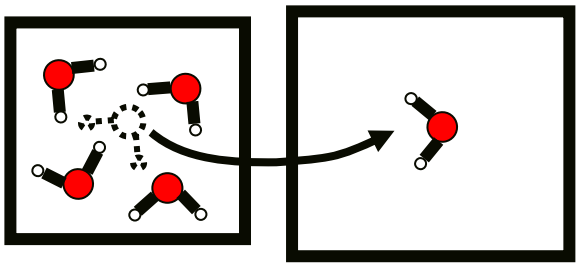
Getting the Statistical Mechanics Right

- It's an imperative - entropy eventually wins and free energies dictate thermodynamic stability and activation barriers
- Need simulation methods that are appropriate for the properties we want to compute (e.g., Gibbs Ensemble for phase equilibria)
- Worry about quantum mechanical effects on nuclear motion (quantum stat mech) when appropriate

Two-Phase Equilibrium Calculations via First-Principles (CP2K) - Mundy et al.

Gibbs Ensemble technique combined with CP2K/QUICKSTEP

Liquid box Vapor box



$T_1 = T_2$ $V_1 + V_2 = V$ $\mu_1 = \mu_2$

- The Gibbs Ensemble is the only correct way to compare to x-ray data of hydrogen bond energetics at different temperatures

- Unique capability allows validation of certain aspects of Saykally and Nilsson experiment and provide additional microscopic insight
- Liquid water predominantly consists of DD moieties but H-bonding is asymmetric

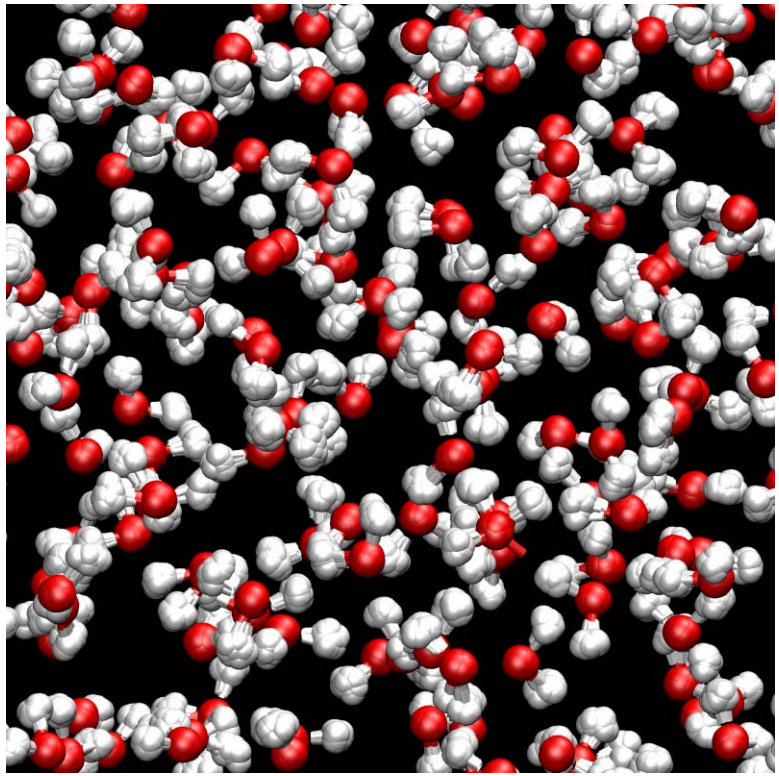
Modeling evaporation at 473K

QuickTime™ and a
YUV420 codec decompressor
are needed to see this picture.

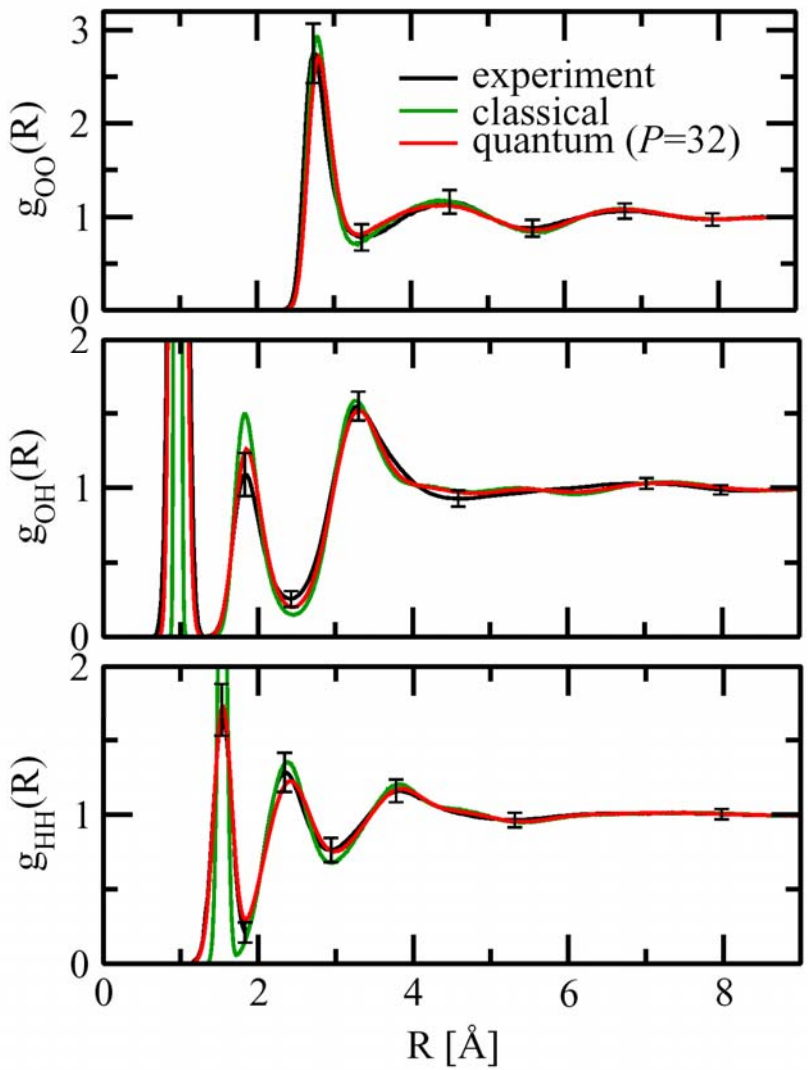
Collaborators:
- University of Minnesota
- University of Zurich
- ETH-Zurich

McGrath *et al.* JPC-A 110, 640 (2006)
McGrath *et al.* Comp. Phys. Comm 169, 289 (2005)
McGrath *et al.* ChemPhysChem 6, 1894 (2005)

Quantum Simulations of Liquid Water (TTM2.1-F)

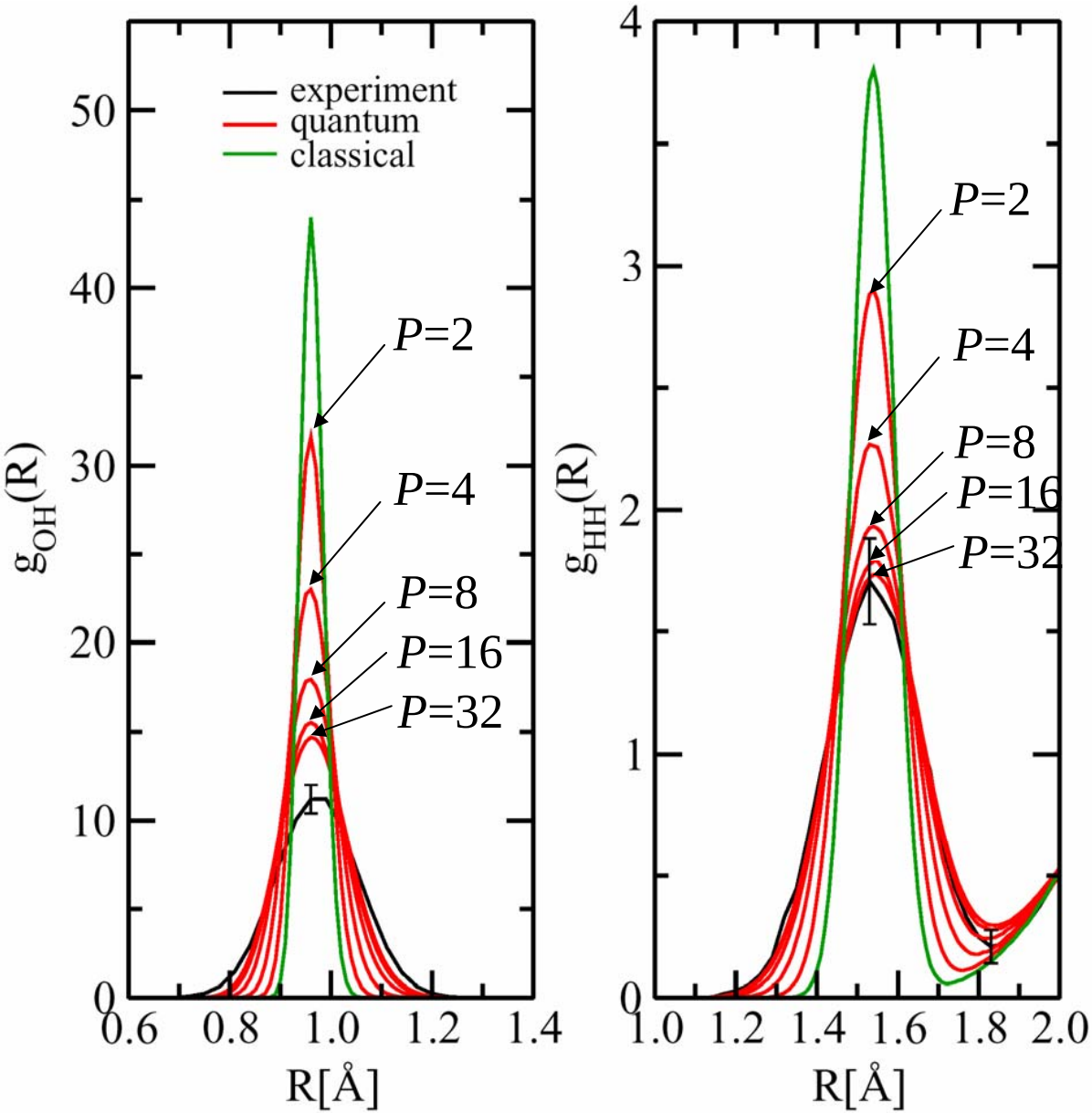


Snapshot of ‘quantum waters’ in centroid MD simulation



Fanourgakis, Schenter and Xantheas, *JCP* **125**, 141102 (2006)

Convergence of Quantum Simulations



Enthalpy of Liquid Water
(T=298.15 K, ρ=0.998 g/cm³)

P	Monomer		Liquid		ΔE (kcal/mol)	
1 (Classical)	3.55		-8.11(2)		-11.67(2)	
	U_0	U_T	U_0	U_T	U_0	U_T
2	5.175(3)	32.501(40)	-5.52(1)	21.20(1)	-10.90(1)	-11.51(4)
4	7.815(5)	20.383(13)	-2.54(1)	9.63(1)	-10.56(1)	-10.97(2)
8	11.048(3)	16.142(7)	0.76(1)	5.63(1)	-10.49(1)	-10.71(1)
16	13.475(6)	15.170(7)	3.12	4.73(1)	-10.56(1)	-10.64(1)
32	14.571(5)	15.048(6)	4.13	4.59	-10.64(1)	-10.67(1)

ΔH (= $-\Delta E + k_B T$)

12.27±0.02 kcal/mol		(P=1, Classical)
11.47	11.93	(P=2)
11.25	11.89	(P=4)
11.06	11.27	(P=8)
11.13	11.21	(P=16)
11.18	11.21	(P=32)

Best estimate: ΔH = 11.2 kcal/mol (exp. 10.51 kcal/mol)

Summary

- Need to move away from the domination of electronic structure in determining computational needs for modeling chemical transformations
  - New strategies that 'reuse' ab initio information or adaptively create functional fits (e.g., MCMM)
- New strategies are needed for statistical mechanics and dynamics algorithms …
  - … that can capitalize on advances in electronic structure methods
  - … that efficiently combine QM and MM methods
- Combination of electronic structure + simulations (statistical mechanics or dynamics) for chemical transformations provides a problem set that *can make immediate use* of petascale computing
- … But do we *need* petascale computing to solve these problems?
  - For example, can we distribute work across a collection of terascale computers instead of requiring a petascale platform?