Classical Molecular Dynamics Codes and Coupling of Length Scales

Peter T. Cummings[1,2], Normand Modine[3] and Randy Cygan[4]

[1]Chemical and Biomolecular Engineering, Vanderbilt U.
[2]Center for Nanophase Materials Sciences, ORNL
[3]Center for Integrated Nanotechnologies, SNL
[4]Geochemistry Department, SNL

BES / ASCR / NERSC Workshop
~~*Hilton Washington DC/Rockville Executive Meeting Center*~~
February 9-10, 2010

Classical Molecular Dynamics

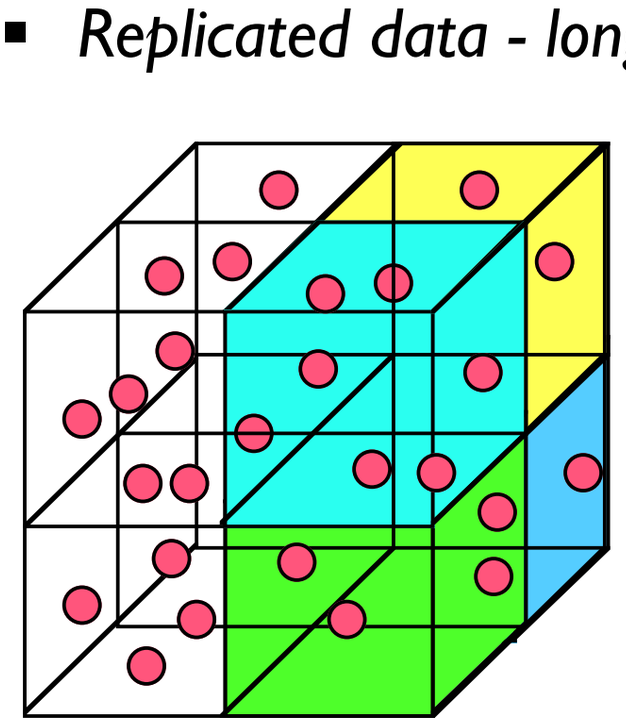
- A reminder...
  - *For N atoms, solve 6N simultaneous 1st-order highly non-linear ordinary differential equations*

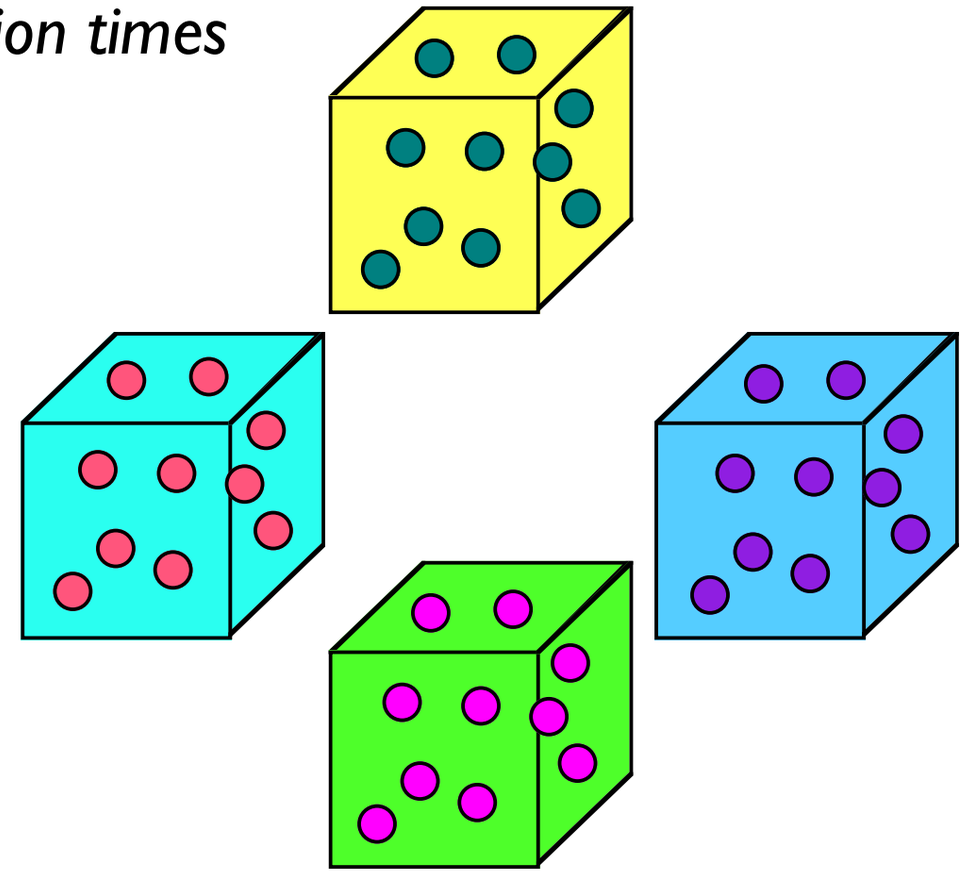
$$\frac{d\mathbf{r}_i}{dt} = \frac{\mathbf{p}_i}{m} \qquad \frac{d\mathbf{p}_i}{dt} = \mathbf{F}_i = -\nabla_{\mathbf{r}_i} U = \underbrace{\left(= \sum_{j \neq i} \mathbf{F}_{ij} \right)}_{\text{simplest case}}$$

  - *Evaluation of right hand side completely dominates calculation*
    - 90-99% of compute time
    - Limits numerical methods applicable
  - *Explicit methods only*
    - Predictor-corrector, Runge-Kutta, specialized algorithms
      - *Predictor-corrector with time step small enough for one corrector cycle*
  - *Infinite but periodic in directions in which periodic boundary conditions imposed*
    - Leads to difficulties in systems with long-ranged forces (Coulombic, dipolar)

Classical Molecular Dynamics

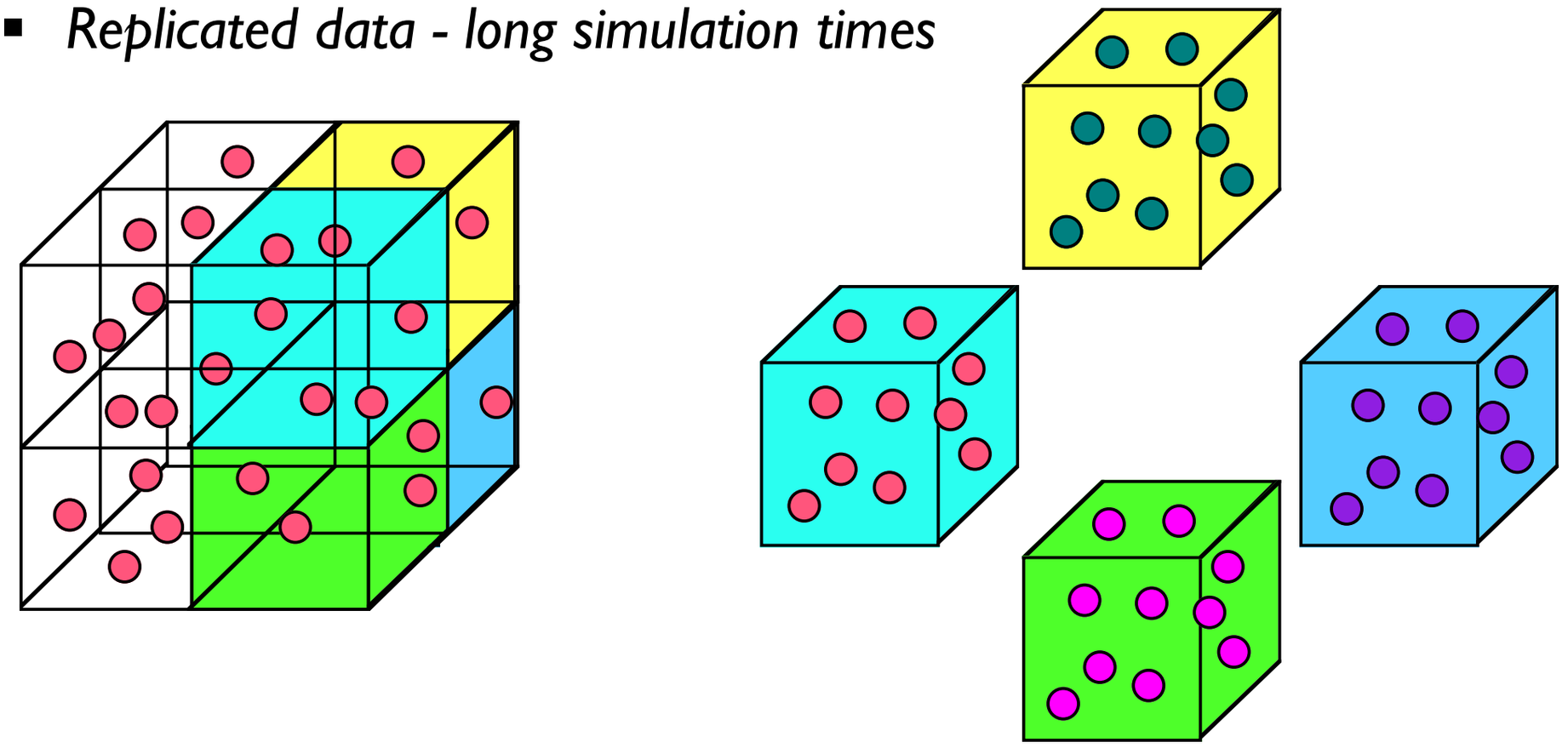
- Parallelization
  - *Domain decomposition - large systems*
  - *Replicated data - long simulation times*





Classical Molecular Dynamics

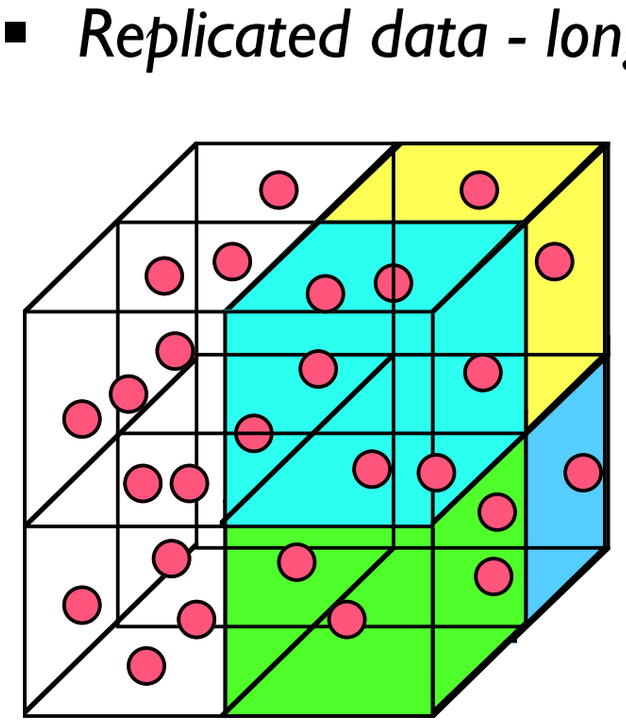
- Parallelization
  - *Domain decomposition - large systems*
  - *Replicated data - long simulation times*

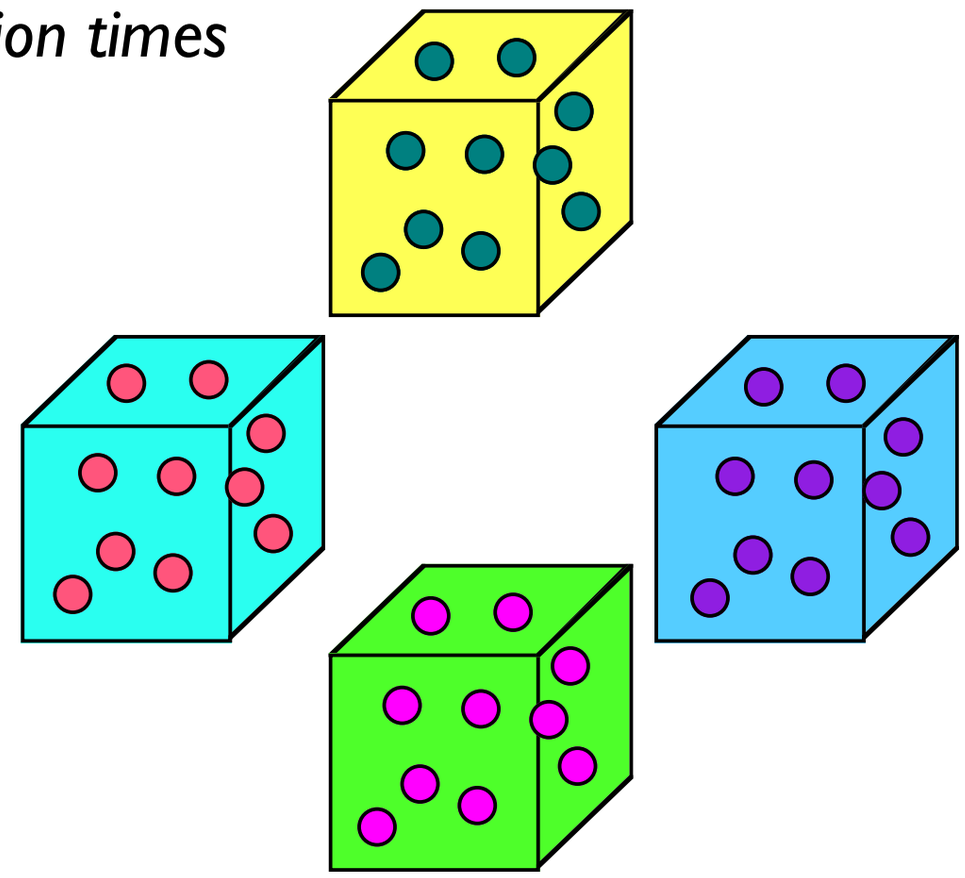


- *Best scaling behavior for large systems*
- *Most complex to code*
- *E.g., LAMMPS*

Classical Molecular Dynamics

- Parallelization
  - *Domain decomposition - large systems*
  - *Replicated data - long simulation times*





- *Best scaling behavior for large systems*
- *Most complex to code*
- *E.g., LAMMPS*

- *Relatively simple to code*
- *Poor scaling performance due to all-to-all or gather/scatter at each time step*
- *E.g., DL_POLY 2.X*

Classical Molecular Dynamics

- The typical progression in Cummings group
  - *Can I do it in LAMMPS, GROMACS, NAMD,...?*
  - *If yes*
    - Problem solved
  - *If not*
    - Write in-house code
      - *Polarizable fluids*
      - *Force fields models unsupported by standard codes*
      - *Many-body interactions (e.g., TB-SMA)*
      - *External electric field*
    - Parallelize using replicated data
    - Modify LAMMPS to handle specifics
      - *Gnu license, C++, object-oriented, source code available*
      - *May require post-processing of simulation configurations*
- Typical progression for CNMS/NTI user
  - *Move simulation to LAMMPS if at all possible*
  - *LAMMPS is becoming the popular trajectory generator for CMD*
    - Bulk, nano, bio simulations

Classical Molecular Dynamics

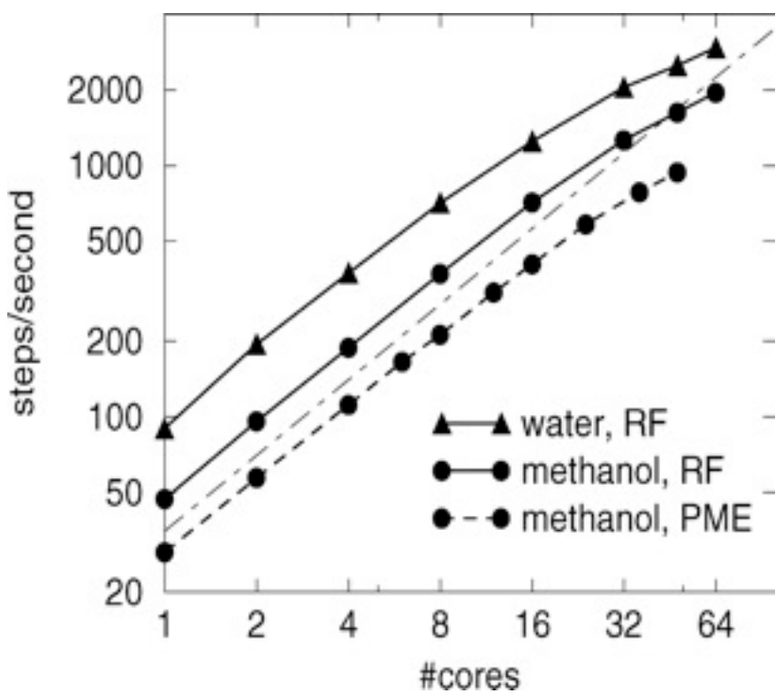
- Short-ranged forces only

$$T_{wall-clock} \approx \left(\frac{N_{atoms}}{N_{proc}} \right)^{\alpha} \frac{1}{PE\left(N_{proc}\right)} N_{timesteps}$$

  - *PE = parallel efficiency, dependent on number of processors*
  - $\alpha \simeq 1+$

- Long-ranged forces
  - *Ewald, particle-mesh Ewald (PME), particle-particle-particle mesh (PPPM)*
  - *Scaling in atom number is in principle* $\sim N \log_2 N$

Scaling for a methanol system of 7200 atoms (circles) and an SPC/E water system of 9000 atoms (triangles), with a cutoff 1 nm, with reaction field (solid lines) and PME (dashed line) with a grid-spacing of 0.121 nm (36 36 36 grid) on a 3 GHz Intel Core2 cluster with Infiniband. The dot-dashed line indicates linear scaling. Hess et al., *J. Chem.Theory Comput.* **2008,** 4, 435-447.



Contrast with First Principles Dynamics

Quantity	Classical MD	*Ab initio* MD	Static
Atom count N	10^3-10^{10}	<1000	>1000
Memory usage	~N*100 bytes	<10GB	>10GB

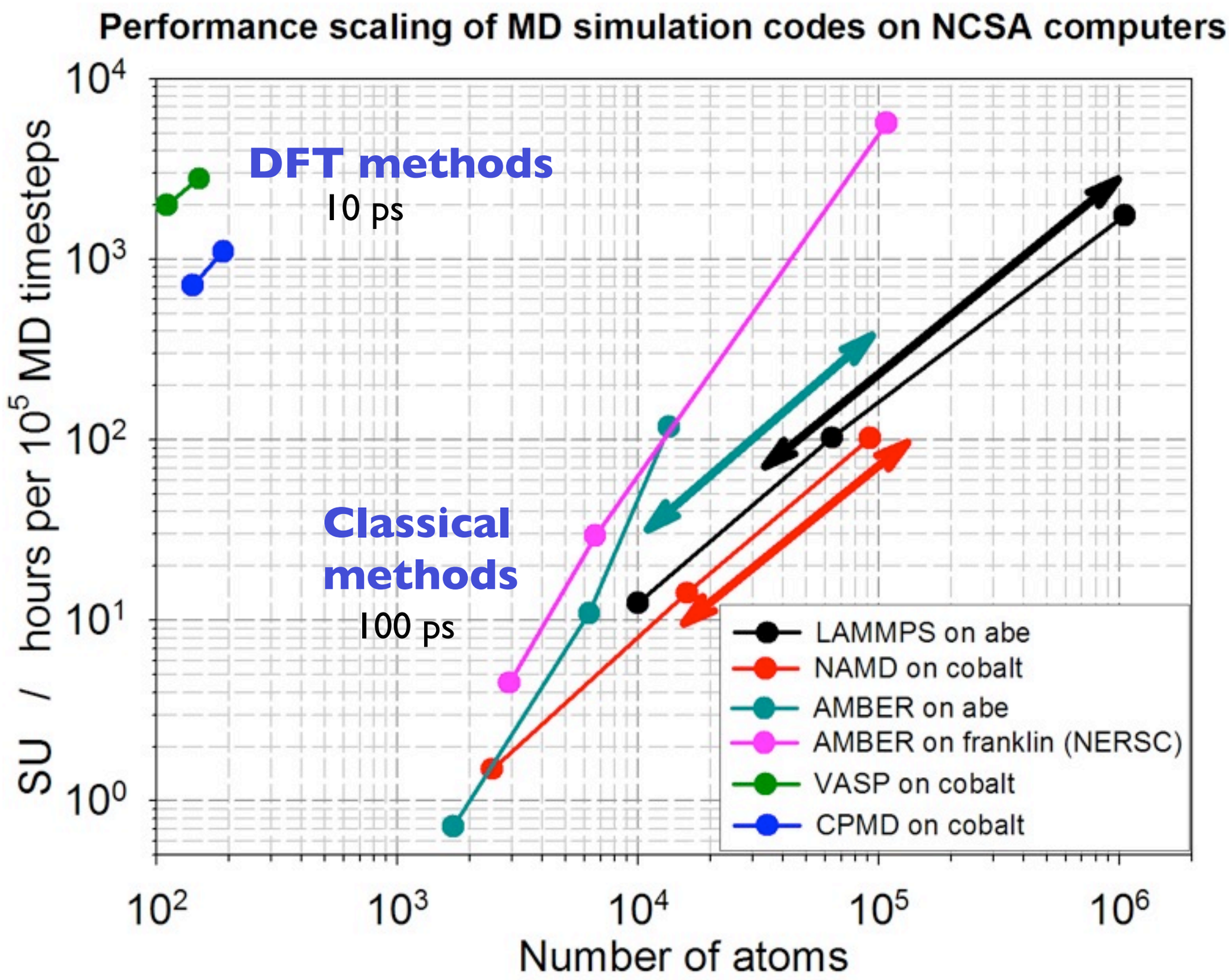
Current	LAMMPS, GROMACS, NAMD, DL_POLY,....	VASP, Qbox, Pwscf	VASP, PEtot
Future	Library-based	???	???

Current	~no linear algebra, N^{1-2}	Dense N^{2-3}	Dense N^{2-3}
Future	~no linear algebra, N^{1+}	Dense N^{2-3}	Sparse N^{1-2}

Practical	Time to reach physically meaningful simulated time (ns-μs-ms)	Time to solution	Awarded hours
Technical	Latency, Bandwidth	Latency, Bandwidth, Dense Linear Algebra	Latency, Bandwidth, Sparse Linear Algebra

- *Thanks to Paul Kent*

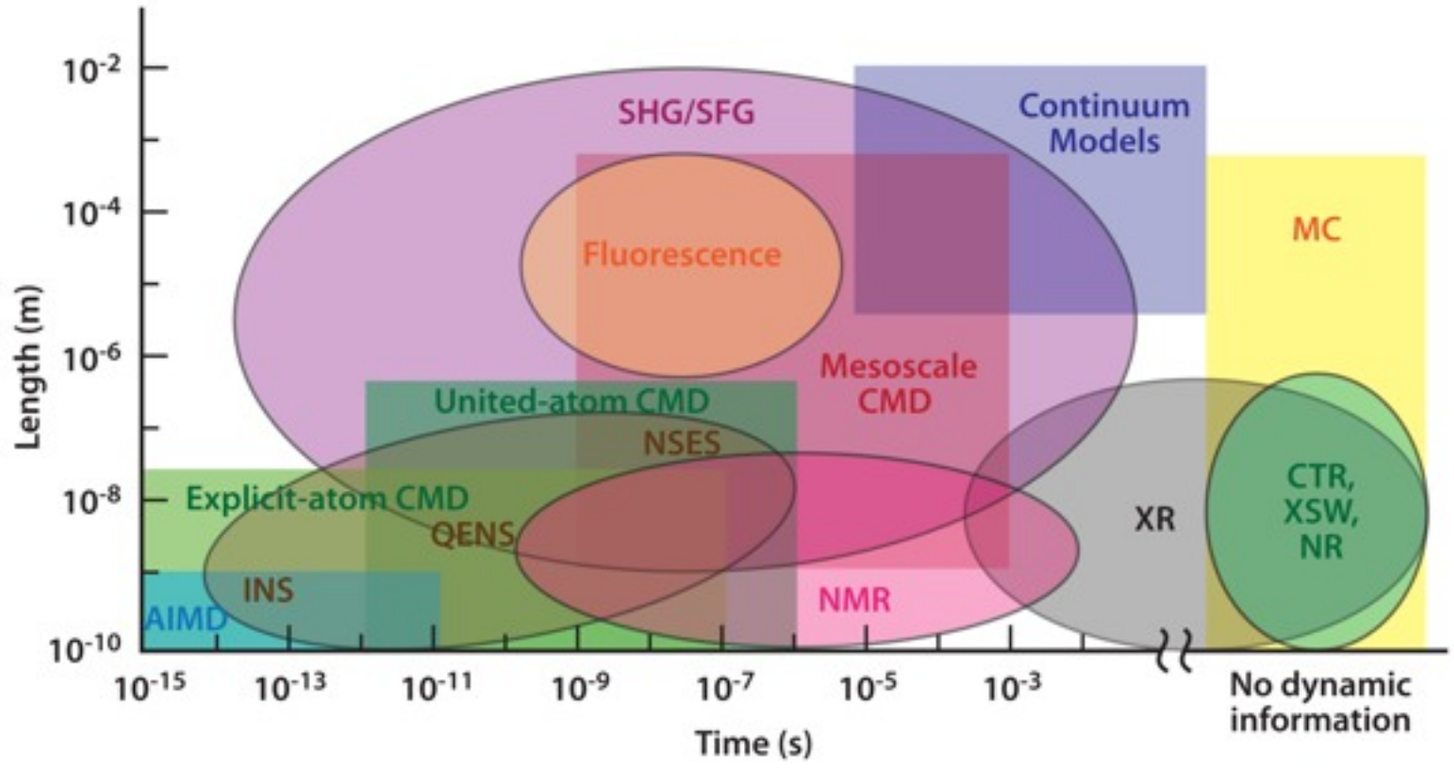
Relative Performance



Courtesy of Andrey Kalinichev via Randy Cygan

Examples of Classical MD

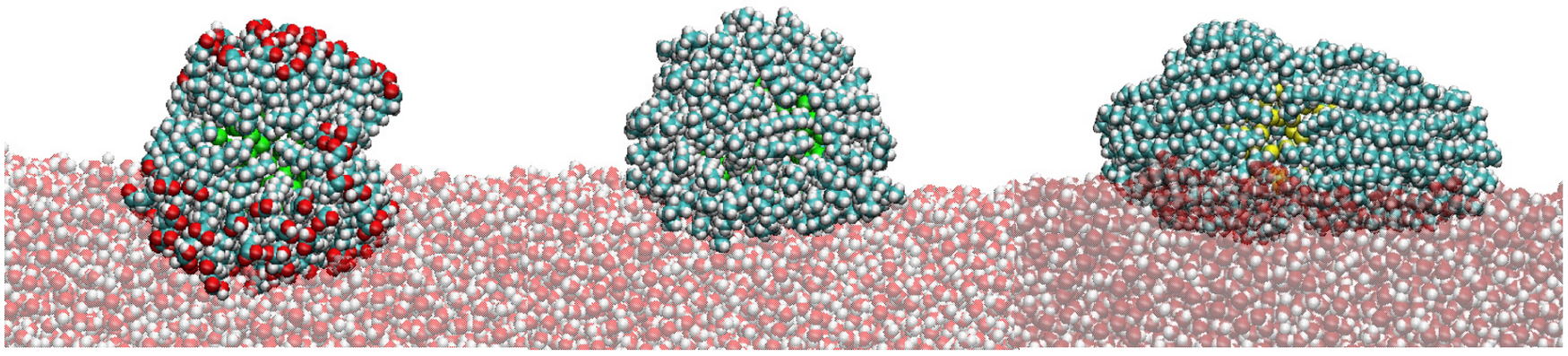
- Computational nanoscience
  - *Interfaces critical*
  - *Systems mixed together at nanoscale*
    - Inorganic, organic, bio
      - *E.g., organic/inorganic hybrid materials, nano-bio*
    - Forcefields are not compatible, and don't work across boundaries
      - *First principles methods needed to calibrate force fields*
  - *New sources of experimental data at nanoscale*
    - Neutrons especially relevant to classical MD
  - *Nanoscience problems are rarely solved just using classical MD*



Examples of Classical MD

- CINT Classical Molecular Dynamics
  - *Normand Modine, CINT, SNL*
  - *Most CINT CMD is focused on the Nanoparticles in Complex Environments science direction*
    - Polymer Nanocomposites
    - Nanoparticle coatings
    - Interfaces (aqueous, block copolymer, lipid membranes)
    - Environment controls interactions which control organization and properties
  - *CINT Scientists Mark Stevens and Gary Grest*

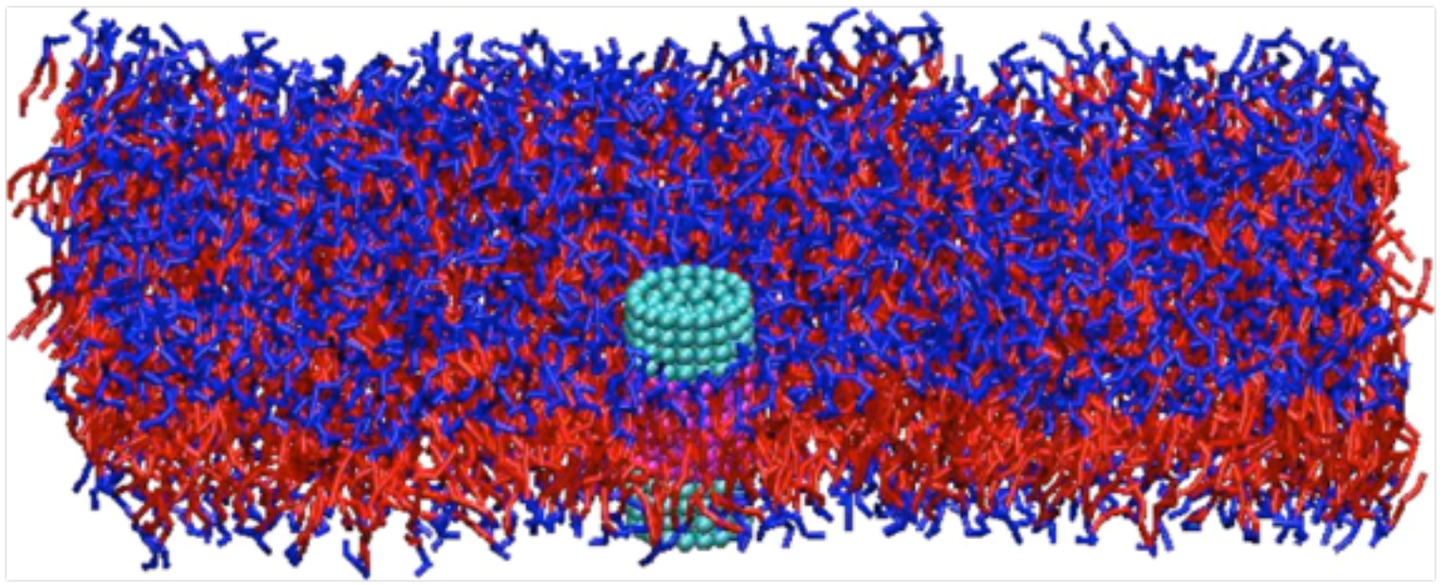
Various Coated Nanoparticles at the Water / Vapor Interface



CINT Classical Molecular Dynamics Case

- Atomistic and Coarse-Grained MD with LAMMPS Code
  - *Typical Systems*
    - 10^4 to 10^6 atoms or elements
    - 1 to 100 ns (atomistic) or 100 to 10,000 ns (coarse-grained)
  - *Typical Requirements*
    - 100 to 1000 Processors for 100 to 1000 Hours
    - 10 to 100 simulations per project
    - 1 GB / processor, 100 to 500 GB storage per project

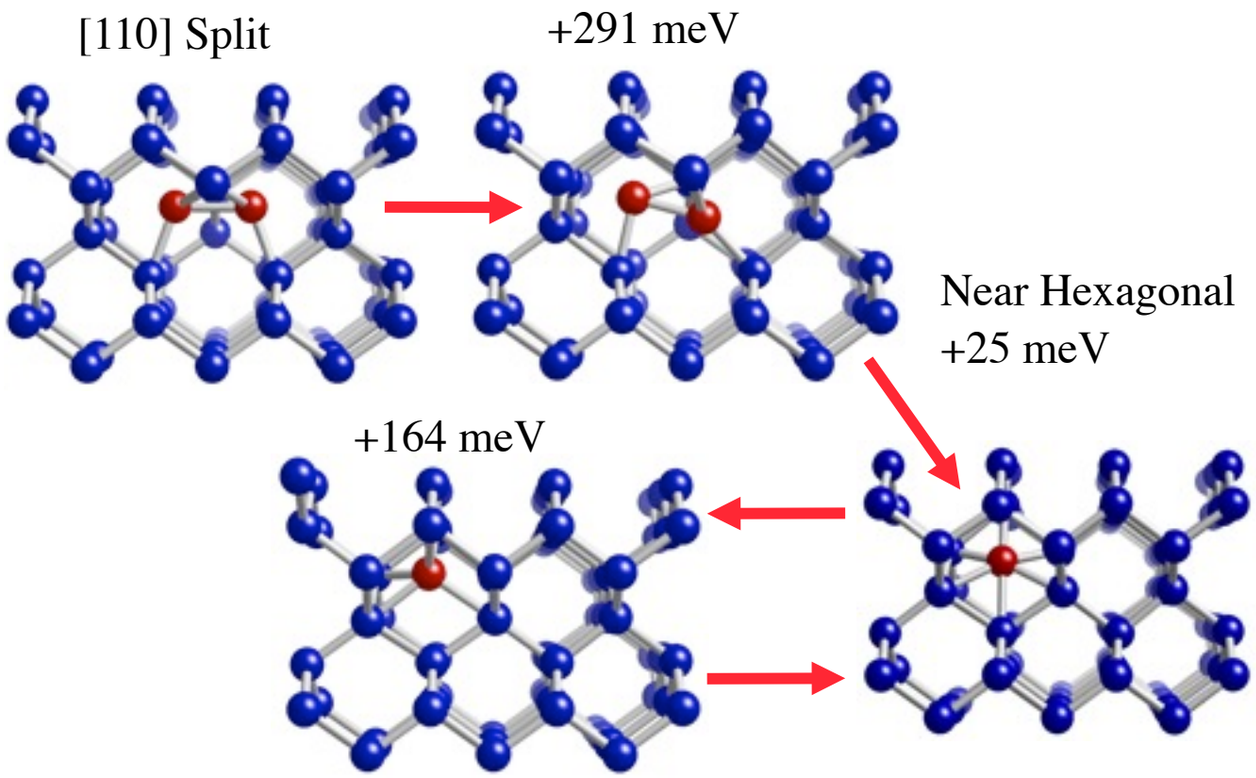
Cylindrical Particle in a Bilayer Lipid Membrane



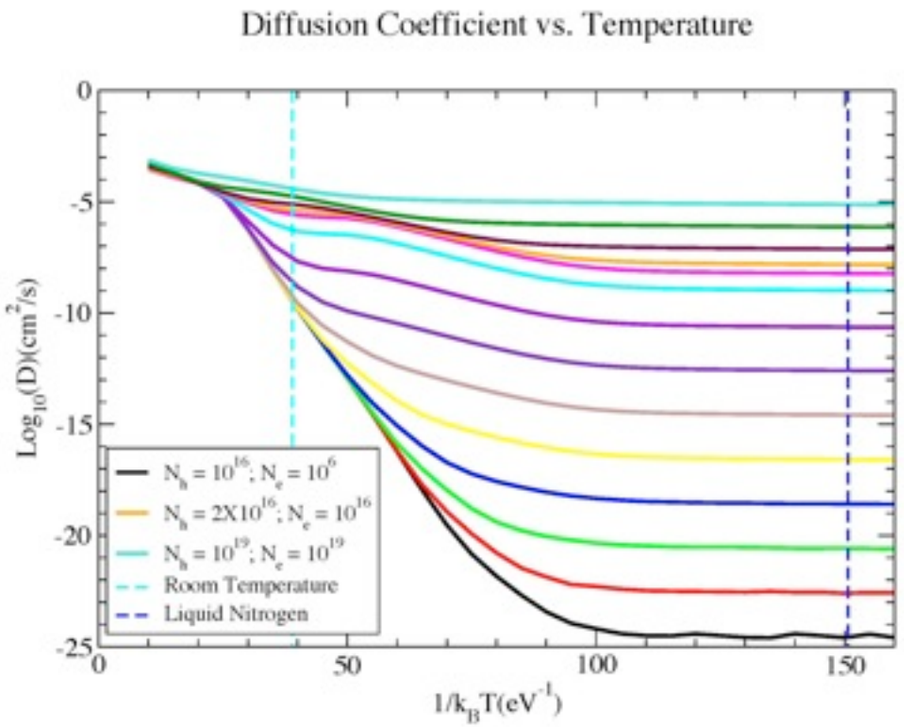
CINT Density Functional Theory (DFT) Case

- Two types of DFT-based Dynamics
  - *DFT Transition State finding for long-time ionic dynamics*
  - *TDDFT for coupled electronic and ionic dynamics*
- Both implemented in the Socorro code

Diffusion of the Neutral Self-Interstitial in Silicon

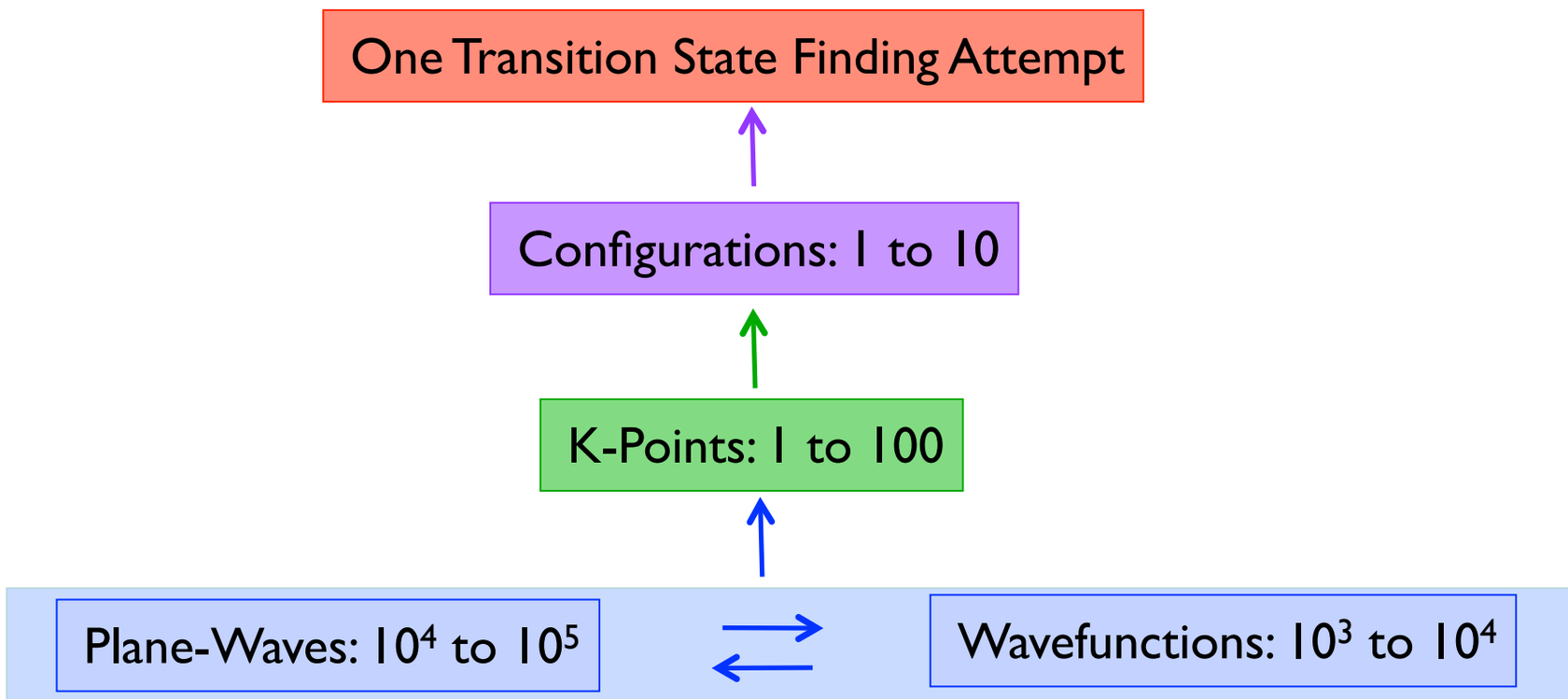


Kinetic Monte-Carlo Based On DFT Results



CINT Density Functional Theory (DFT) Case

- Typical DFT Transition State (TS) system
  - *100 to 1000 atoms with 10 critical configurations*
  - *Finding each critical configuration requires 10 attempts each requiring 1000 to 10000 energy evaluations*
- Typical DFT-TS requirements
  - *100 to 1000 processors for 100 to 1000 hours*
  - *100 simulations per project*

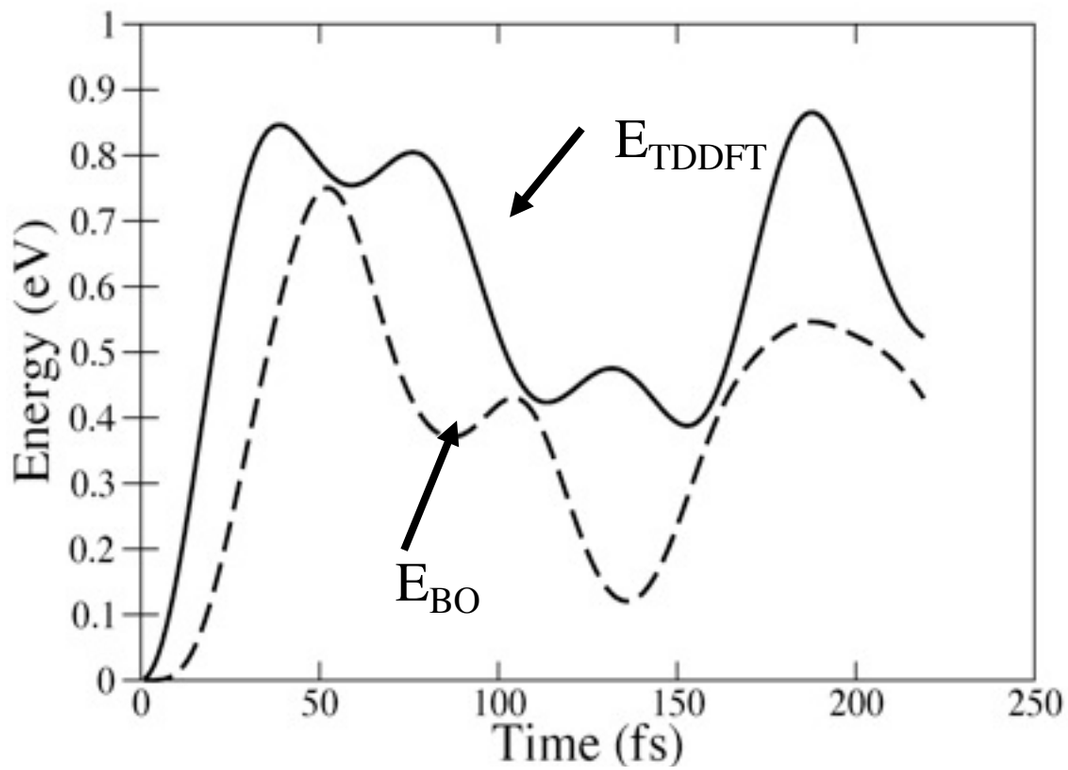


CINT Density Functional Theory (DFT) Case

- Use a real-time TDDFT capability with simultaneous ionic dynamics to study interactions between electrons and ions
  - *Integrate the time-dependent Kohn-Sham equations*
  - *Move ions according to TDDFT forces*
- Allows exchange of energy, momentum, etc. and calculation of thermodynamic and transport quantities

$E_{TDDFT} - E_{BO}$ can be considered to be the instantaneous thermal energy of the electrons

TDDFT Run for 32 Atoms of Al

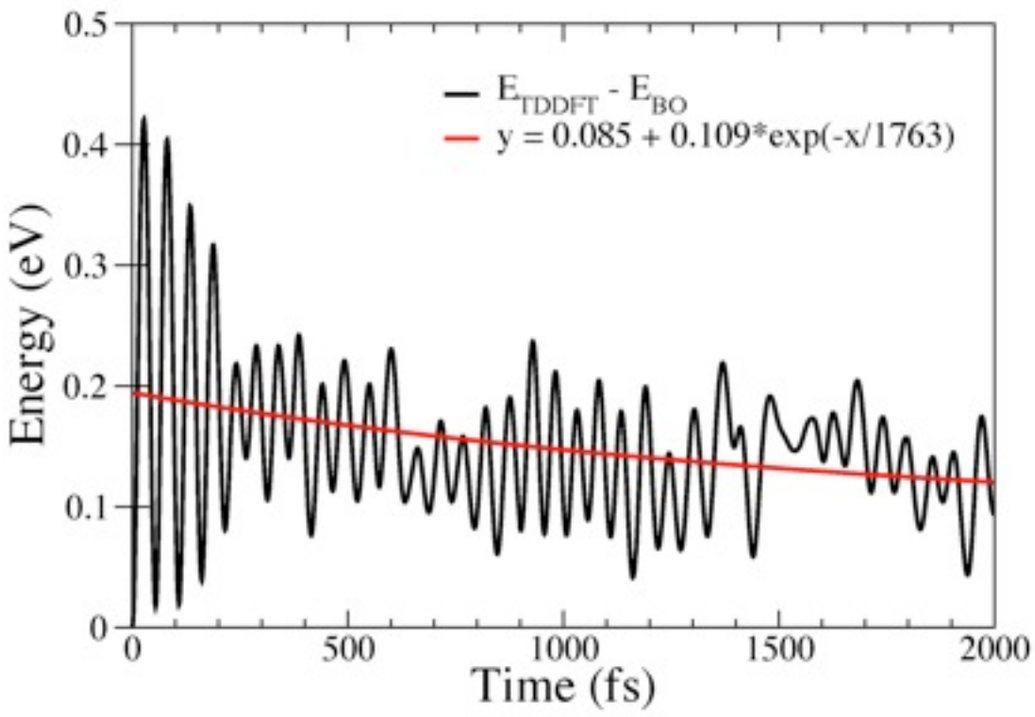


CINT Density Functional Theory (DFT) Case

- Typical TDDFT systems
  - *100 to 1000 atoms for 100 to 1000 fs*
  - *1 attosecond time step → 10^5 to 10^6 time-steps*
- Typical TDDFT requirements
  - *100 to 1000 processors for 100 to 5000 hours*
  - *1 to 10 simulations per project*

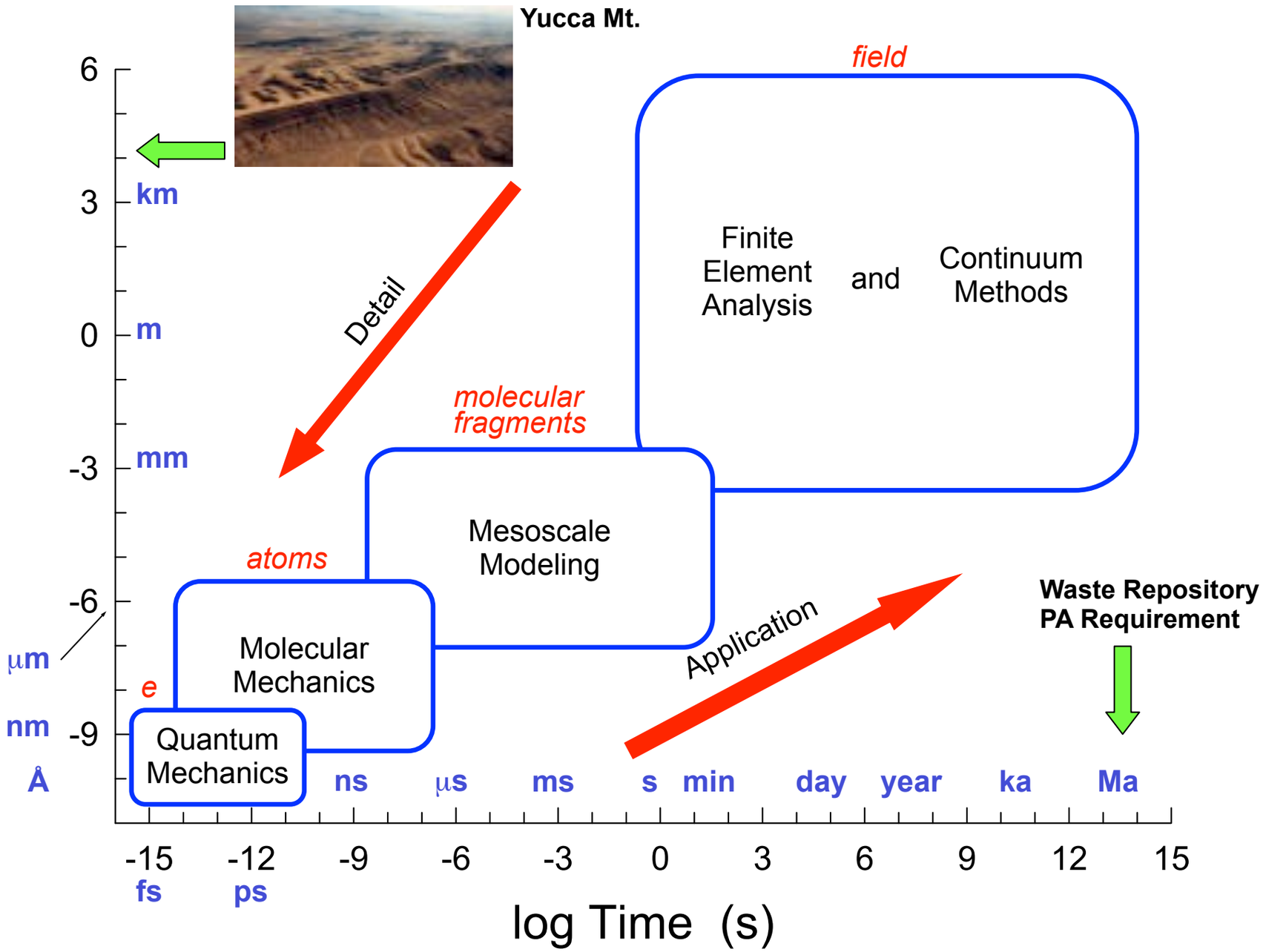
Heat Exchange Between Electrons and Ions in a TDDFT Simulation

- ***Electrons transfer energy to ions with time constant τ_{ep}= 1.8 ps***
- ***Agrees with 1.5-2.0 ps equilibration time from experiment (Kandyla, Shih, and Mazur, 2007)***



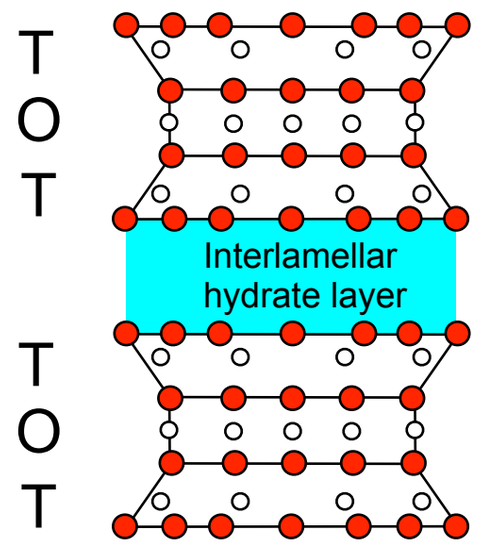
Examples of Classical MD

- Computational geochemistry
  - *Randy Cygan, SNL*



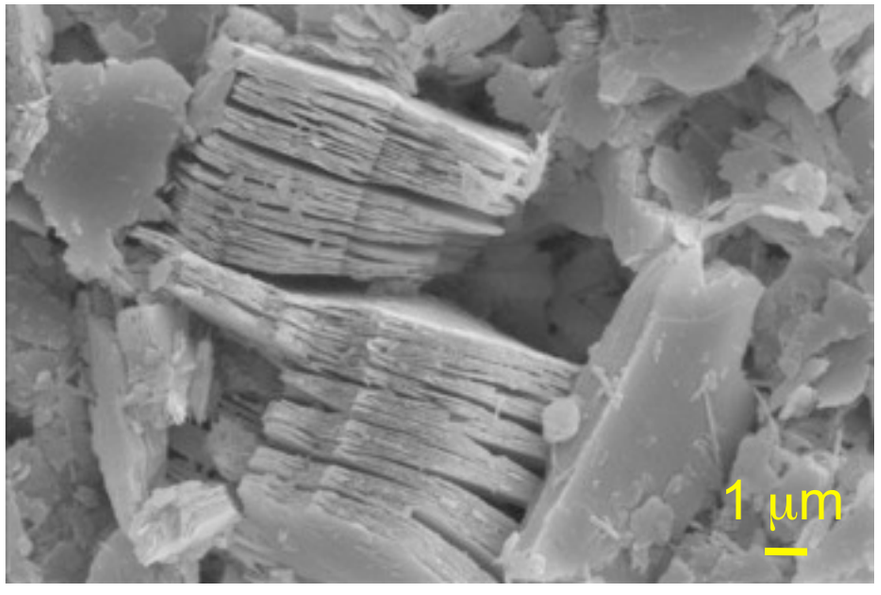
Atomistic Simulation of Clays and Clay Processes

- Crystal structure models of clay minerals are typically unknown
  - *Nanocrystalline (cryptocrystalline) materials (less than 1 μm grain size)*
  - *No large single crystals for X-ray diffraction refinements*
  - *Hydrogens positions are often unknown (require neutron diffraction analysis) and control sorption process*
  - *Complex chemistry with multicomponent systems, cation disorder, and vacancies*
  - *Low symmetry (monoclinic or triclinic)*
  - *Stacking disorder complicates structural analysis*



with M^+

- Atomistic simulations of clay minerals are non-trivial
  - *Require accurate empirical energy forcefield;quantum methods are too costly*
  - *Large unit cells or simulation supercells are required (>100 atoms)*
  - *Significant electrostatic fields associated with layer structure*
  - *Validation of models is difficult*



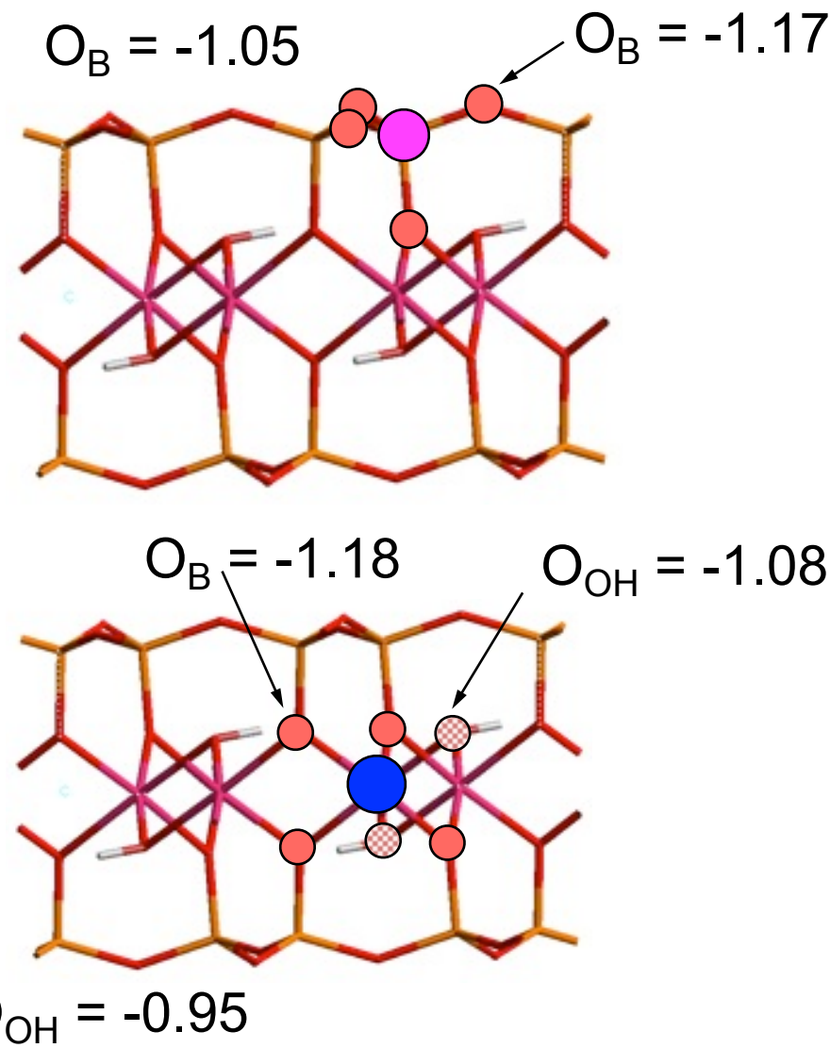
Forcefield for Modeling Clays and Hydrated Phases

❑ CLAYFF

- *specialized semi-empirical fully flexible force field model allowing for realistic exchange of momentum and energy among all atoms – solid substrate and aqueous solution*
  - ➢ Cygan, Liang, and Kalinichev (2004) *J. Phys. Chem. B*, **108** 1255-1266

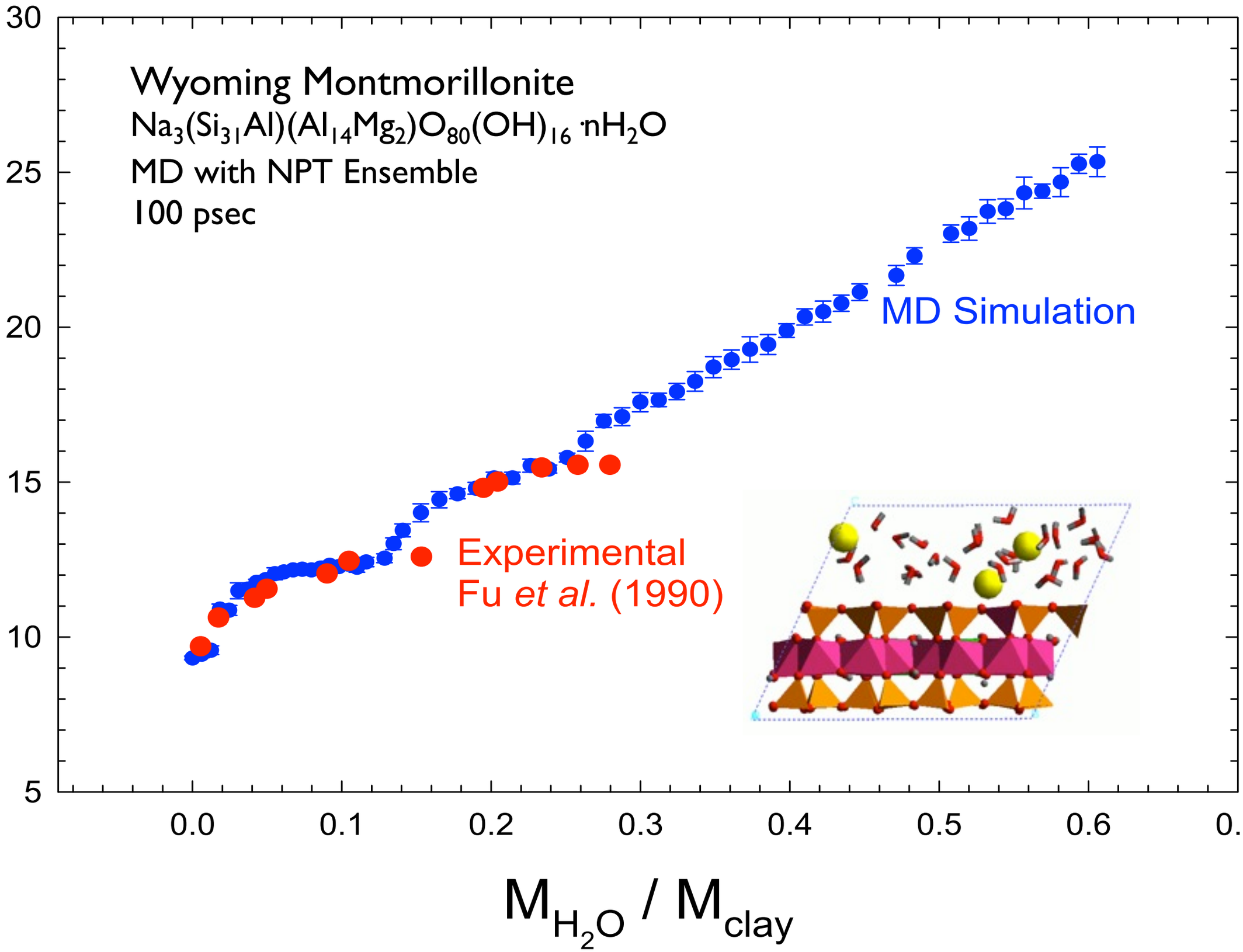
$$U_{ij} = \Sigma\Sigma(A_{ij}/r_{ij}^{12} - B_{ij}/r_{ij}^{6} + q_i q_j /e_0 r_{ij}) + \Sigma\, ½k_b\,(r_{ij} - r_0)^2 + \Sigma ½k_q\,(q_{ij} - q_0)^2$$

Short-range repulsion v-d-Waals Coulombic bond stretching bond bending

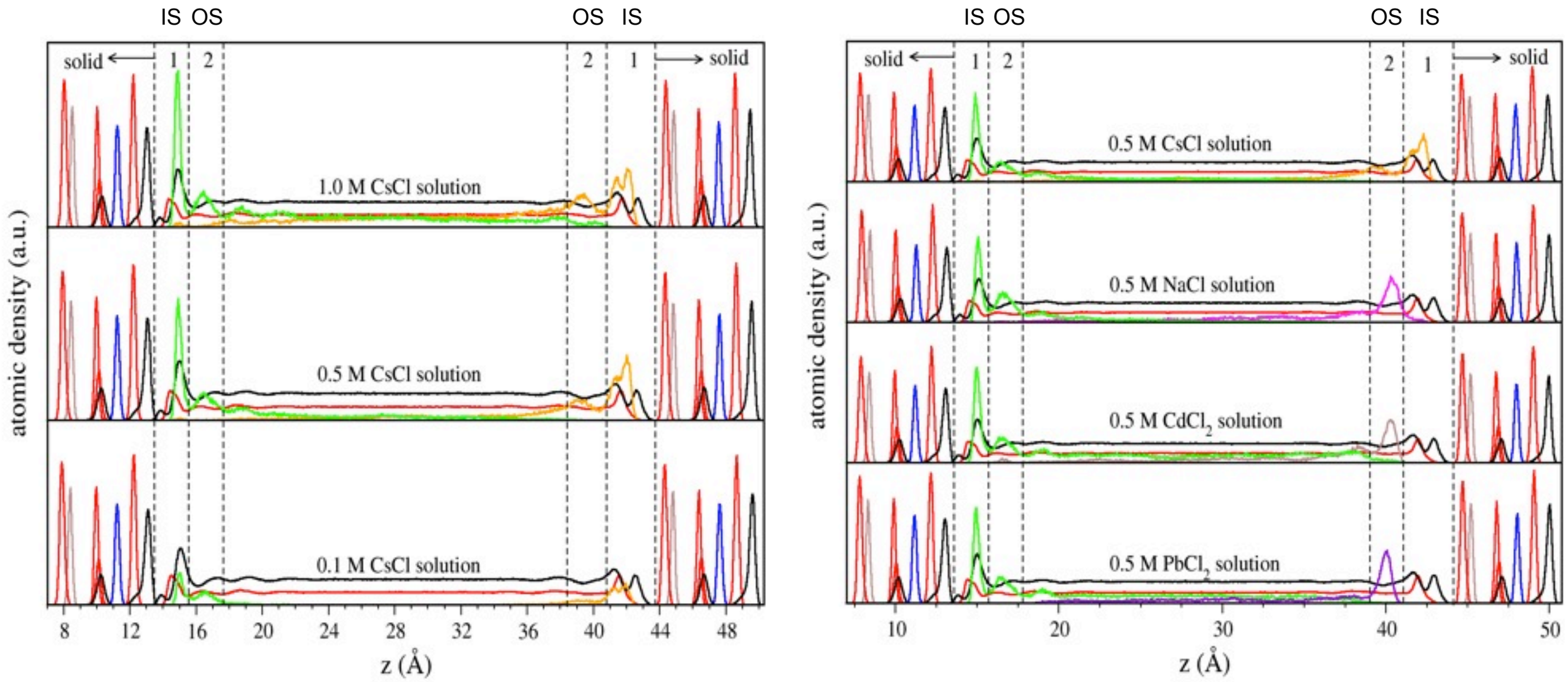


- Simple Point Charge (*SPC*) flexible model for H_2O
- Structural ions: Si, Ca, Al, Fe, Mg, O, OH with partial charges derived from quantum DFT calculations for a number of simple oxides and hydroxides
- Aqueous species: Na^+, K^+, Cs^+, Mg^{2+}, Ca^{2+}, Cl^-, OH^-, SO_4^{2-}, CO_3^{2-}, NO_3^-
- Theoretical models of oxides, hydroxides, clays, and other hydrous materials
- Combination with *CVFF*, *AMBER* or *CHARMM* to model hybrid organic-inorganic systems

Swelling Behavior of Montmorillonite



Density Profiles for Kaolinite Simulations

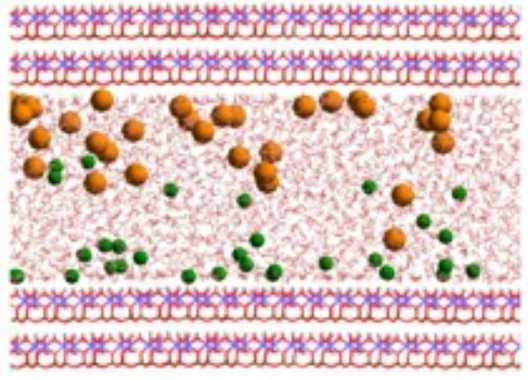


- Profiles calculated from 500 ps of accumulated dynamics after an equilibration period of 600 ps
- Regions named 1 and 2 define inner and outer adsorption shell distances
- Adsorption statistics are obtained by integrating the profiles under regions 1 and 2

Atoms: **Al**, **Si**, **O**, **H**, **Cl^-**, **Cs^+**, **Na^+**, **Cd^{2+}** and **Pb^{2+}**

Derived adsorption statistics: X_{ads}, K_D, site density, etc.

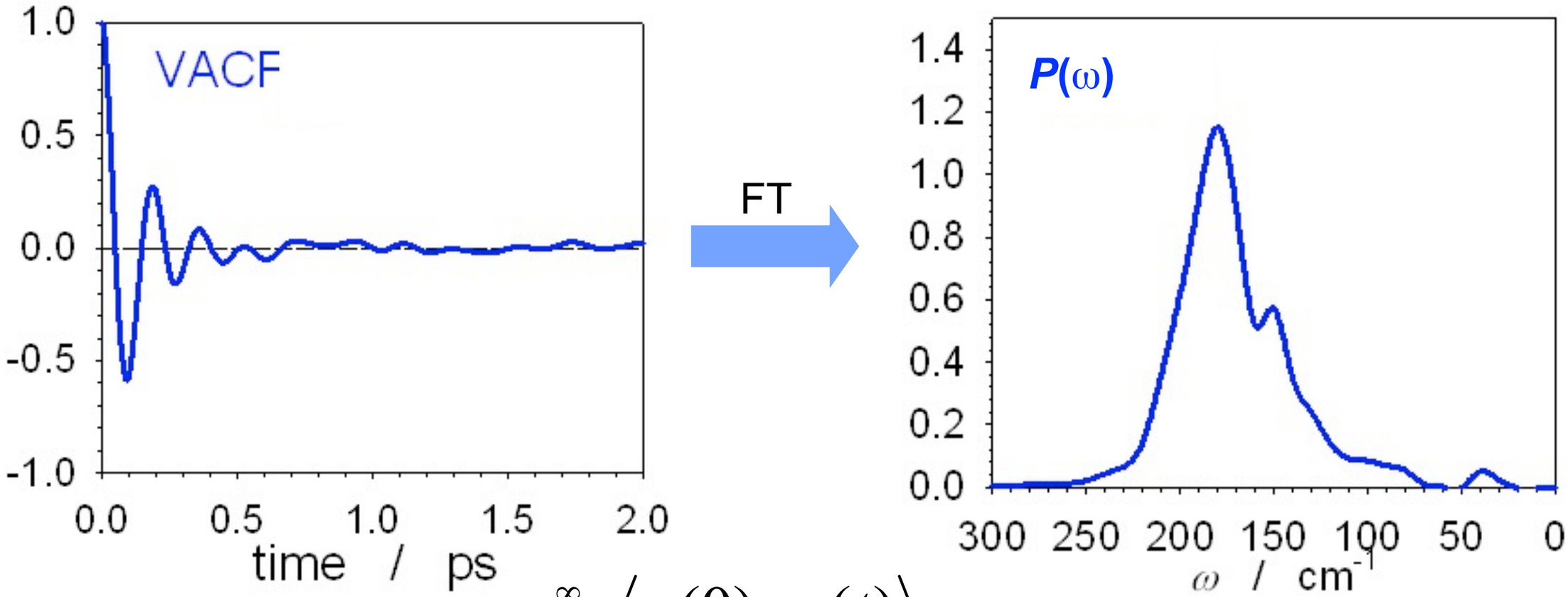
20k atoms



Vasconcelos et al. (2007) Journal of Physical Chemistry C

Dynamics of Individual Atoms
VACFs and Power Spectra

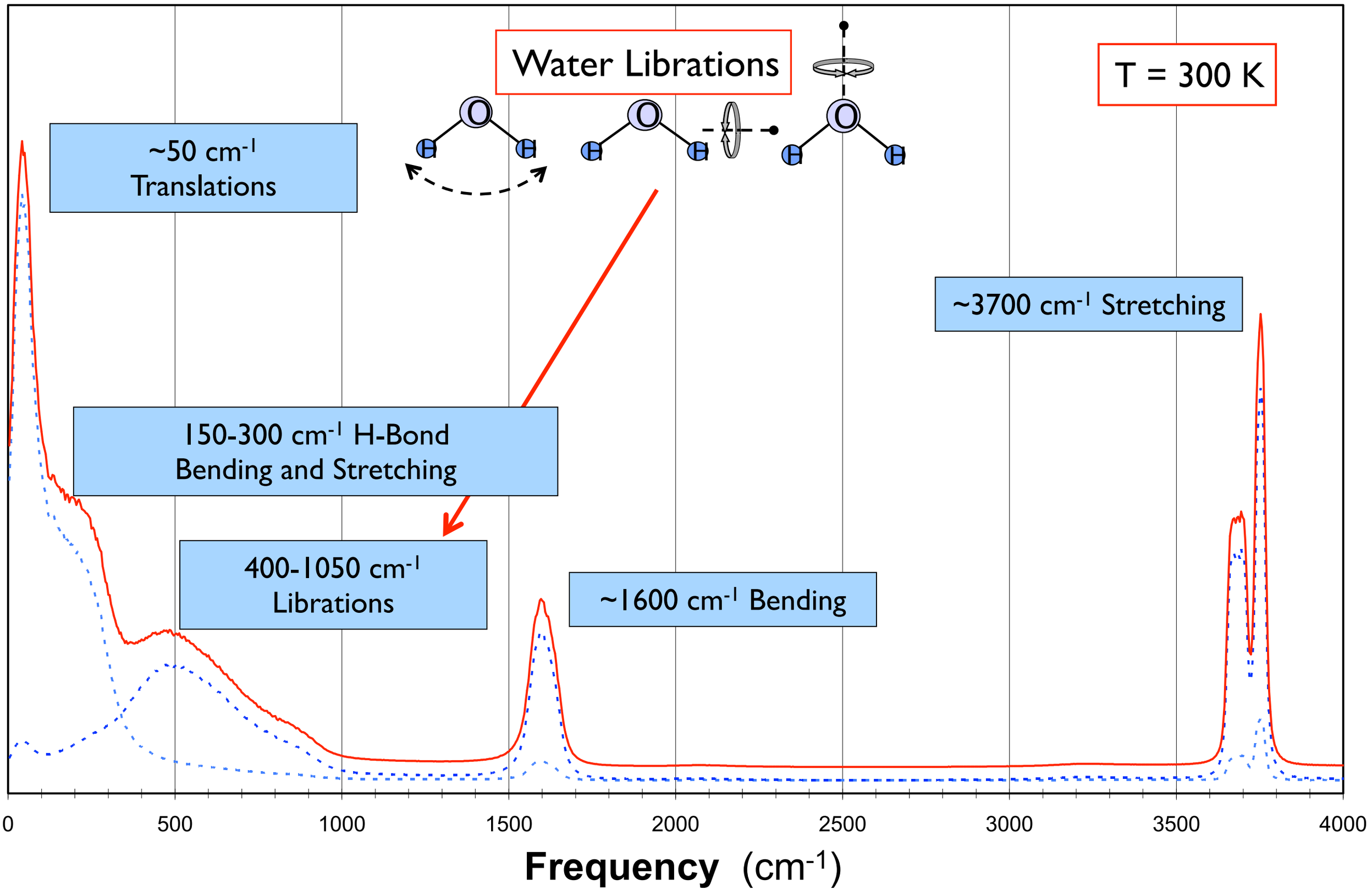
$$\mathrm{VACF} \equiv C_{vv}(t) = \frac{\langle \mathbf{v}(0) \times \mathbf{v}(t) \rangle}{\langle \mathbf{v}(0)^2 \rangle} = \frac{1}{\langle \mathbf{v}(0)^2 \rangle} \int \mathbf{v}(0) \times \mathbf{v}(t)\, d\Gamma$$



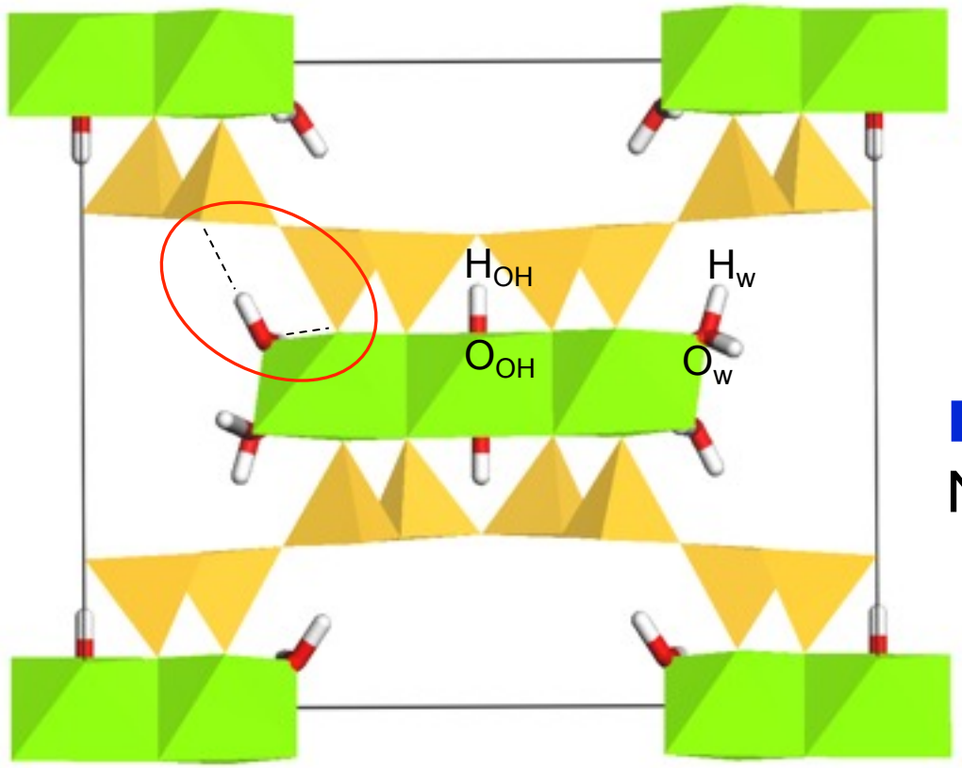
$$P(\omega) = \int_0^\infty \frac{\langle \mathbf{v}(0) \times \mathbf{v}(t) \rangle}{\langle \mathbf{v}(0)^2 \rangle} \cos(\omega t)\, dt$$

VACF = velocity autocorrelation function

Power Spectra of Water



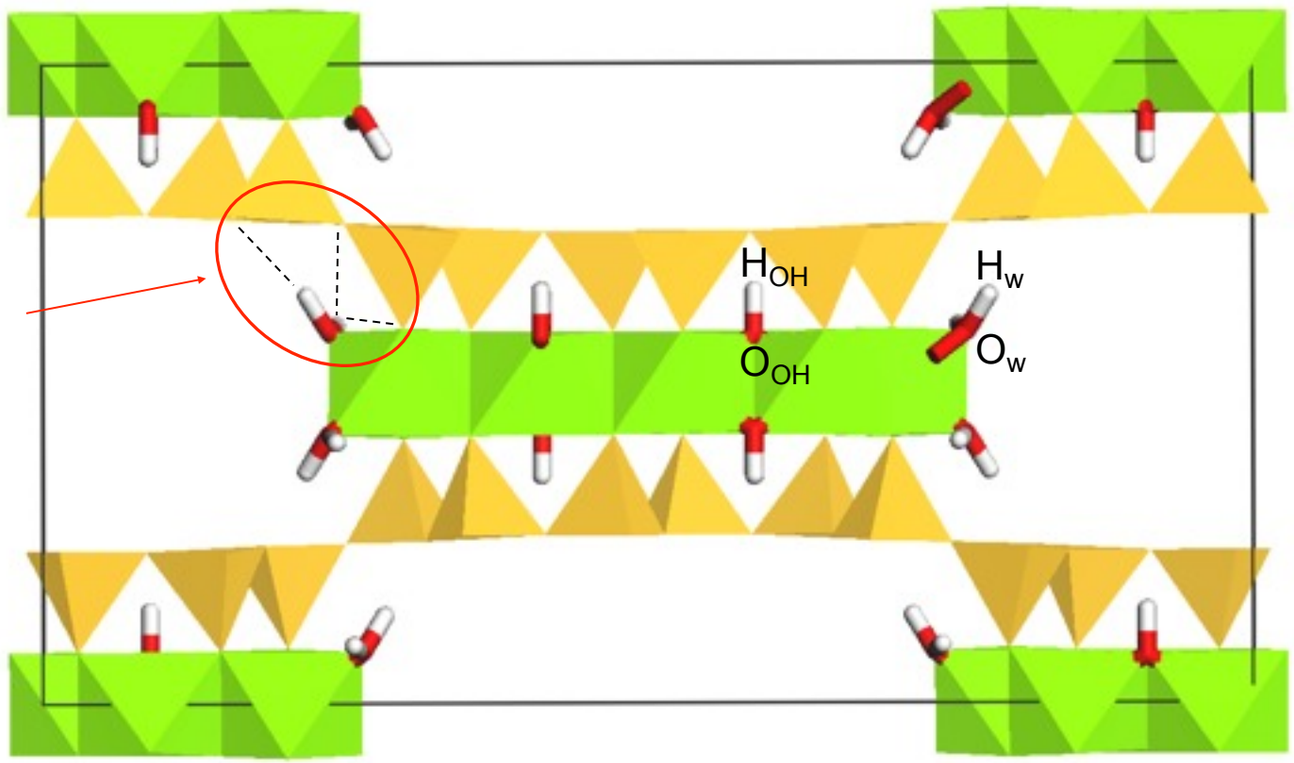
DFT Optimized Structures for Clay Phases



- VASP DFT code
- GGA with projector-augmented wave

Palygorskite

$Mg_5Si_8O_{20}(OH)_2 \cdot 8H_2O$



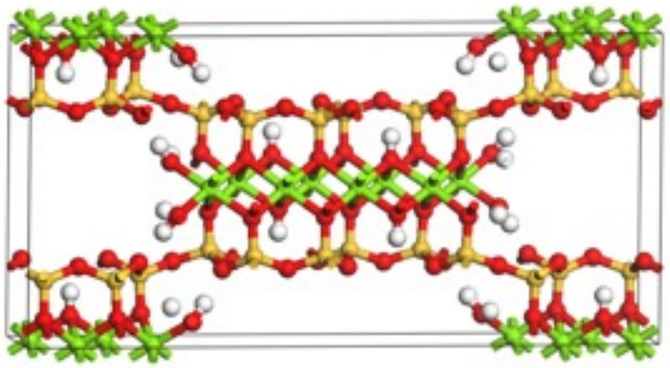
Sepiolite

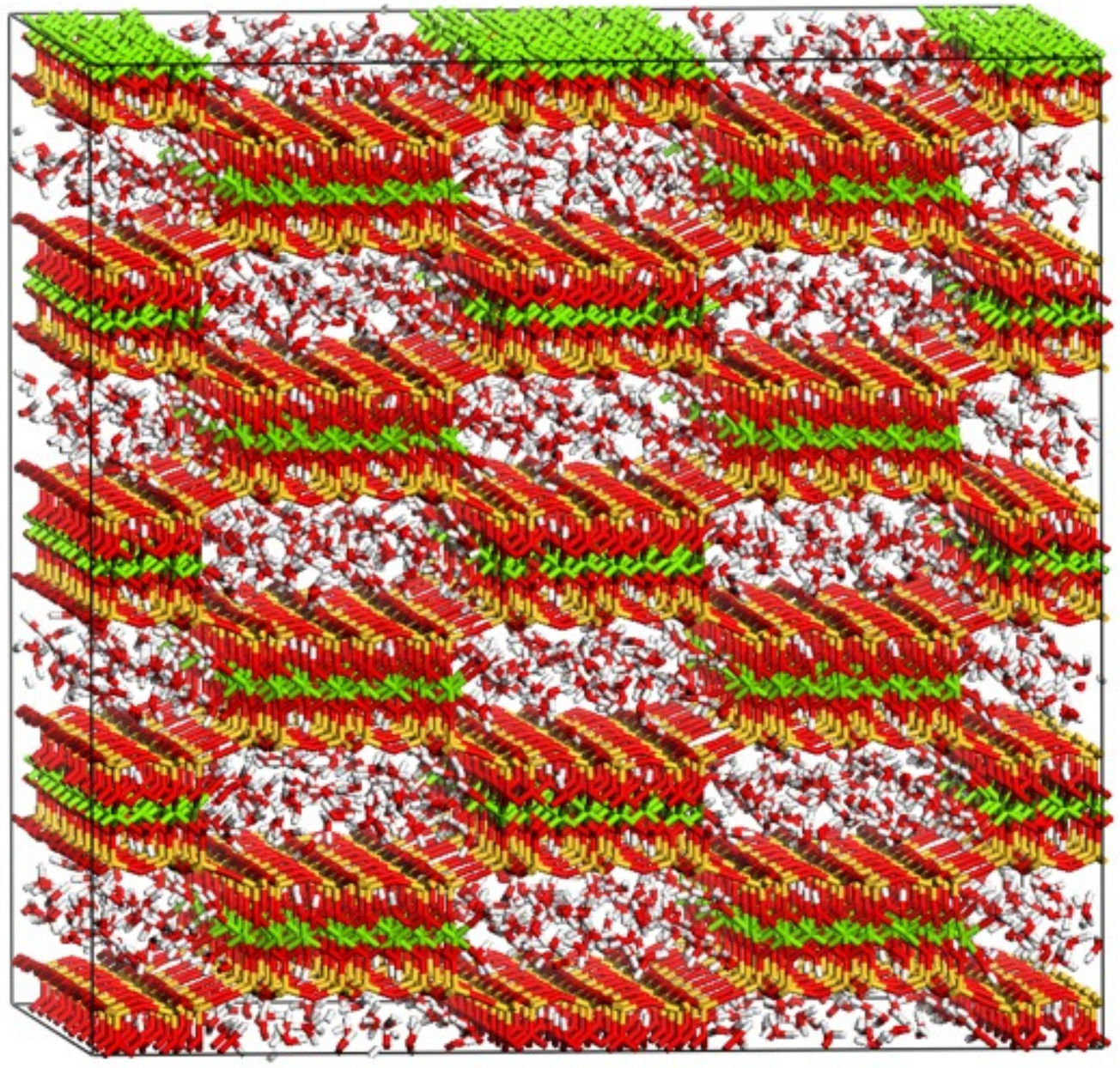
$Mg_8Si_{12}O_{30}(OH)_4 \cdot 12H_2O$

Ockwig et al. (2009)
Journal of the American Chemical Society

Classical and DFT Models for MD

Sepiolite Example





Classical MD — large scale

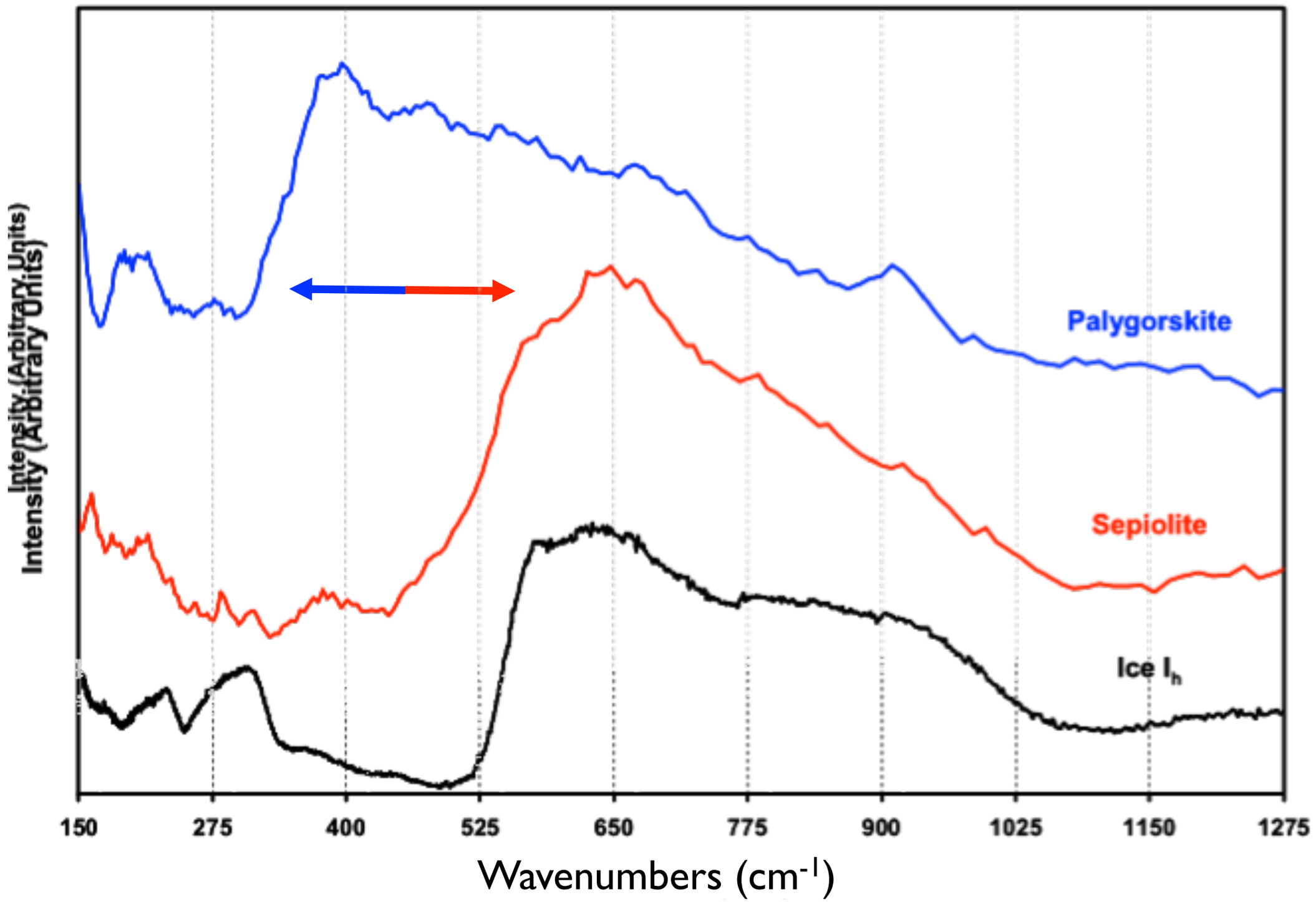
- LAMMPS classical code with CLAYFF
- 250 ps *NVT* and *NPT* MD to equilibrate then 1000 ps for production run
- 40 ps *NVT* MD for VACF calculations
- Structural and vibrational analysis using MD trajectory

Ab Initio MD — unit cell

- VASP DFT code
- GGA with projector-augmented wave
- AIMD for 62 ps *NVT*
- Structural and vibrational analysis using MD trajectory

Sepiolite: 15,040 atoms with 1920 waters
Palygorskite: 20,130 atoms with 2640 waters

INS Spectra for Clay Phases and Ice



Inelastic neutron scattering (INS) data of hydrated palygorskite, hydrated sepiolite, and ice Ih at 90 K

What's Needed for Chemical Accuracy?

GOAL: Develop computational approaches that are highly accurate for the right system. Get the right answer for the right reason.

Predict equilibrium chemistry: ***Selectivity***

Change in K_{eq} @ 298 K

K_{eq} = 1 50:50 ΔG = 0 kcal/mol

K_{eq} = 10 90:10 ΔG = 1.4 kcal/mol

K_{eq} = 100 99:1 ΔG = 2.8 kcal/mol

Predict accurate rates: ***Reactivity***

Absolute rates @ 298 K

Factor of 10 in rate @ 25°C is a change in E_a of 1.4 kcal/mol

Do this in a complex system where the model represents the system accurately under the relevant conditions

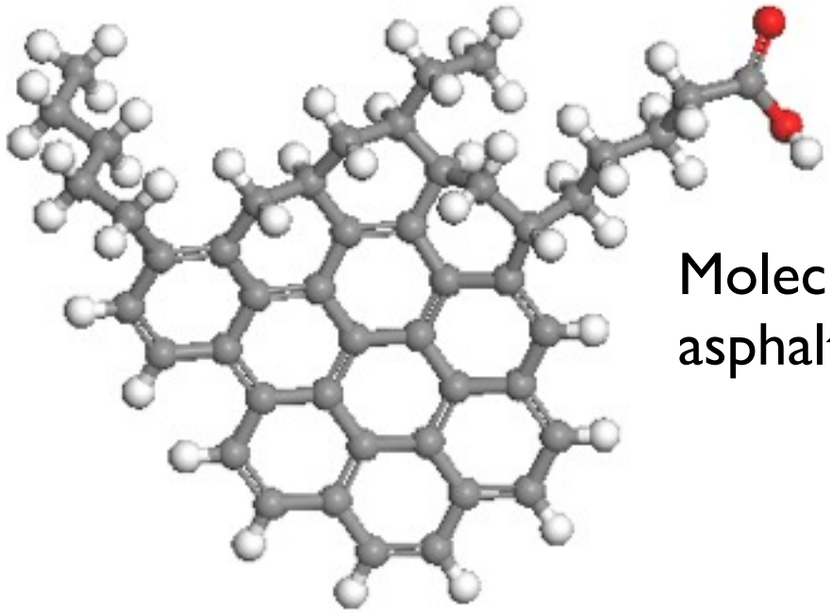
Complexity examples for system size:

- 100x100x100 nm box of water molecules would have 4 x 10^5 H_2O molecules
- Neutral pH requires 10^7 H_2O molecules per H^+/OH^- pair
- Minimum number of atoms in a molecular dynamics trajectory study will be 10^5 to 10^6 atoms for microseconds (10^{-6} s) with femtosecond (10^{-15} s) time steps.

Courtesy of David Dixon

Computational Challenges

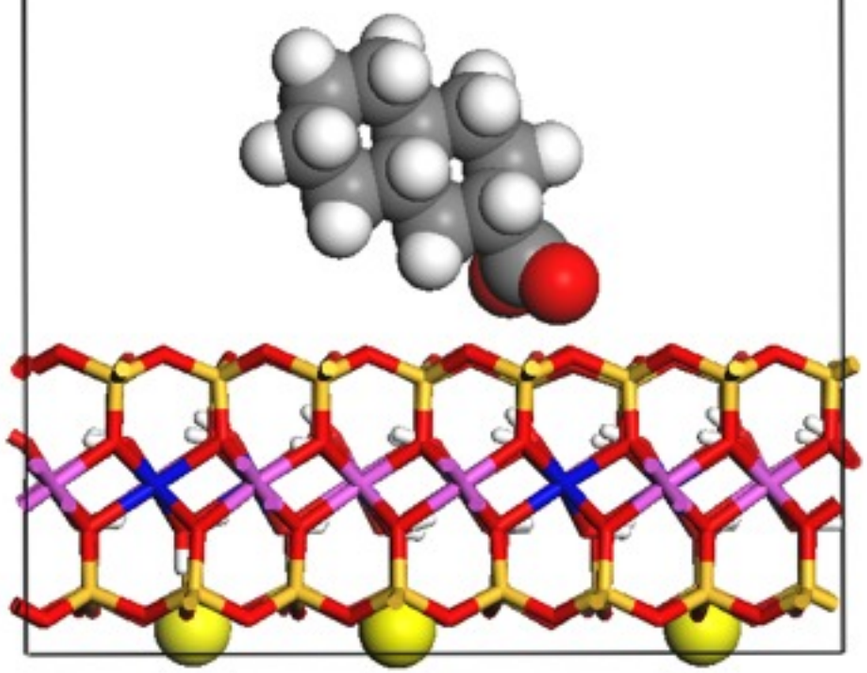
- Enhanced oil recovery
- Carbon sequestration
- Environmental contamination

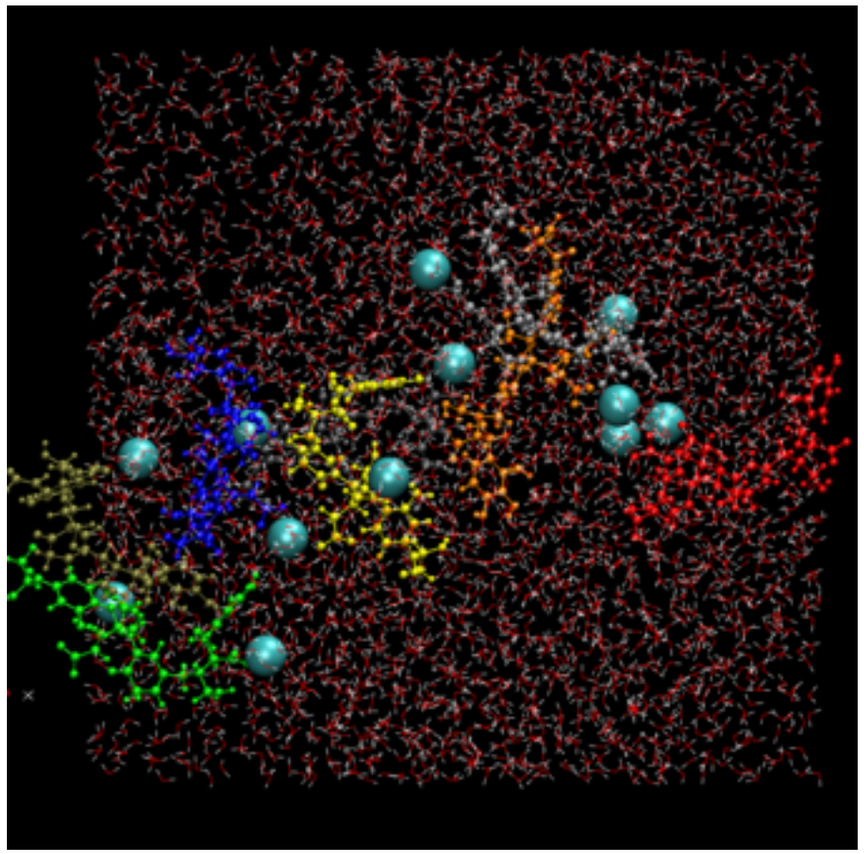


Molecular abstract for an asphaltene (only ~10%)

VASP optimization
Napthoic acid adsorption
montmorillonite

~600 atoms

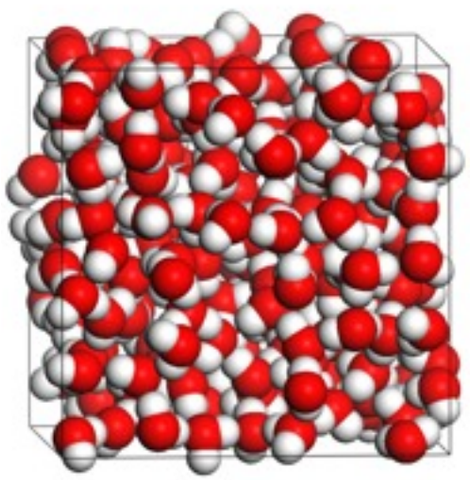


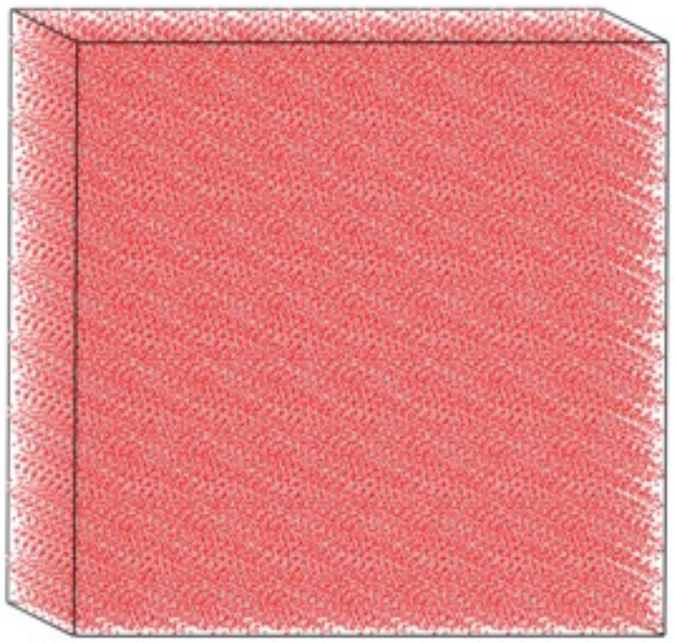


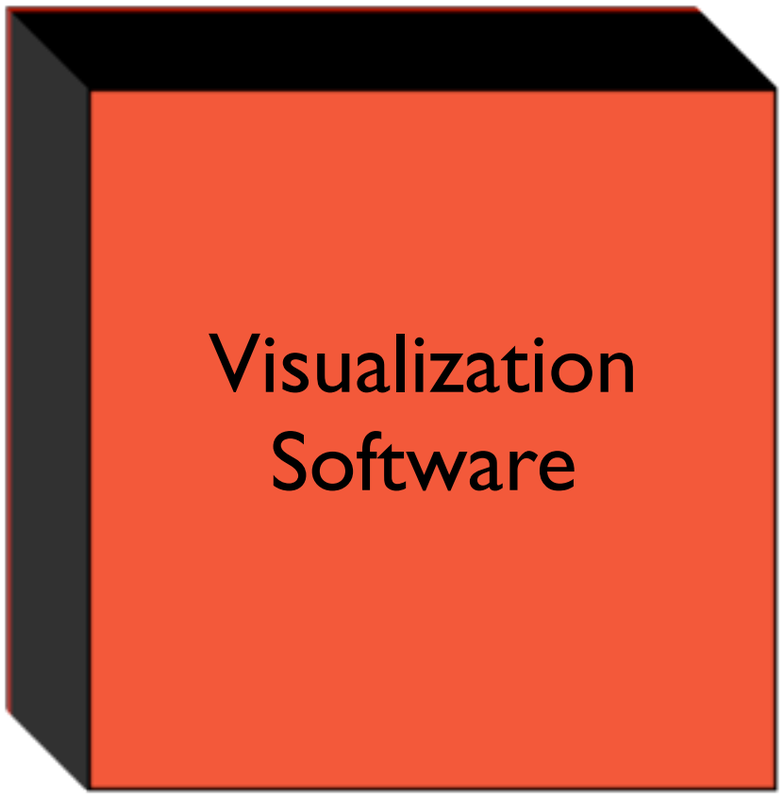
Natural organic matter
M^+ complex formation
Surface adsorption
Aggregation

t = 10 ns

The Future







216 water molecules
19 Å periodic box

110,000 water molecules
150 Å periodic box

10^6 water molecules
450 Å periodic box

Classical MD

100 ns to 1 μs
10^7 to10^8 atoms

Gigascale to
Terascale

Ab Initio MD

100 ps to 1 ns
10^3 to10^4 atoms

Petascale to
Exascale

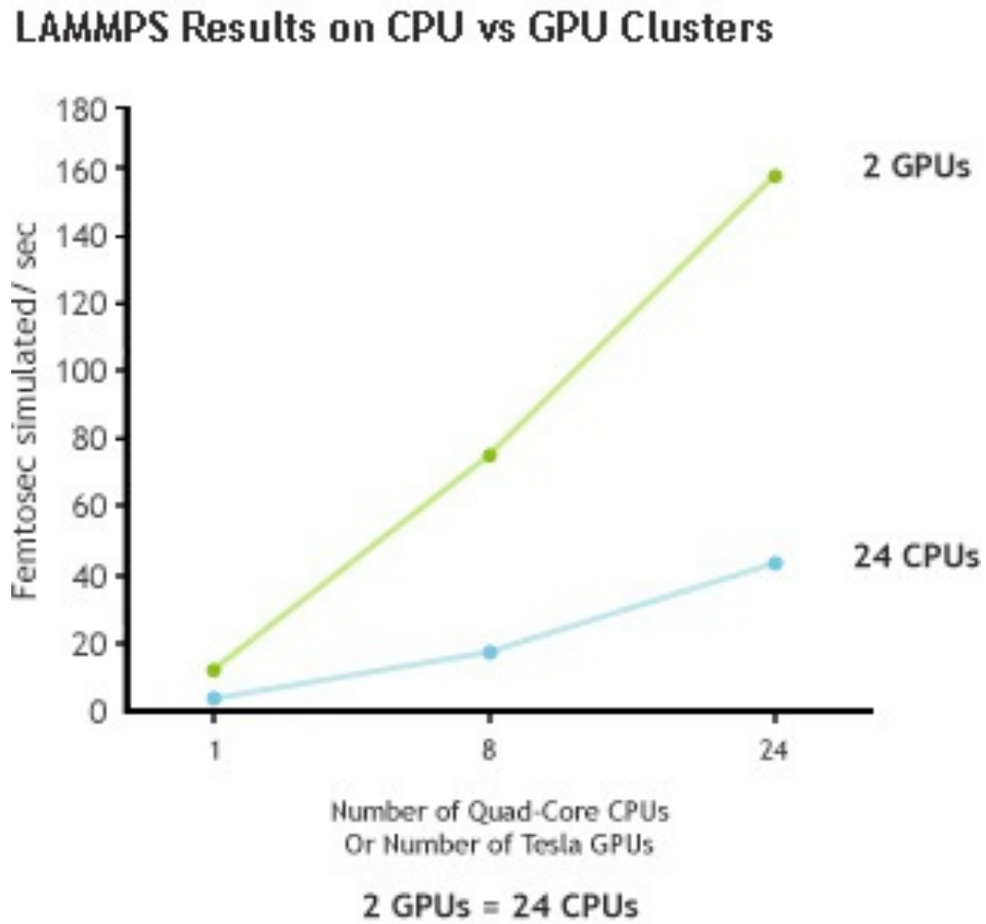
The Future

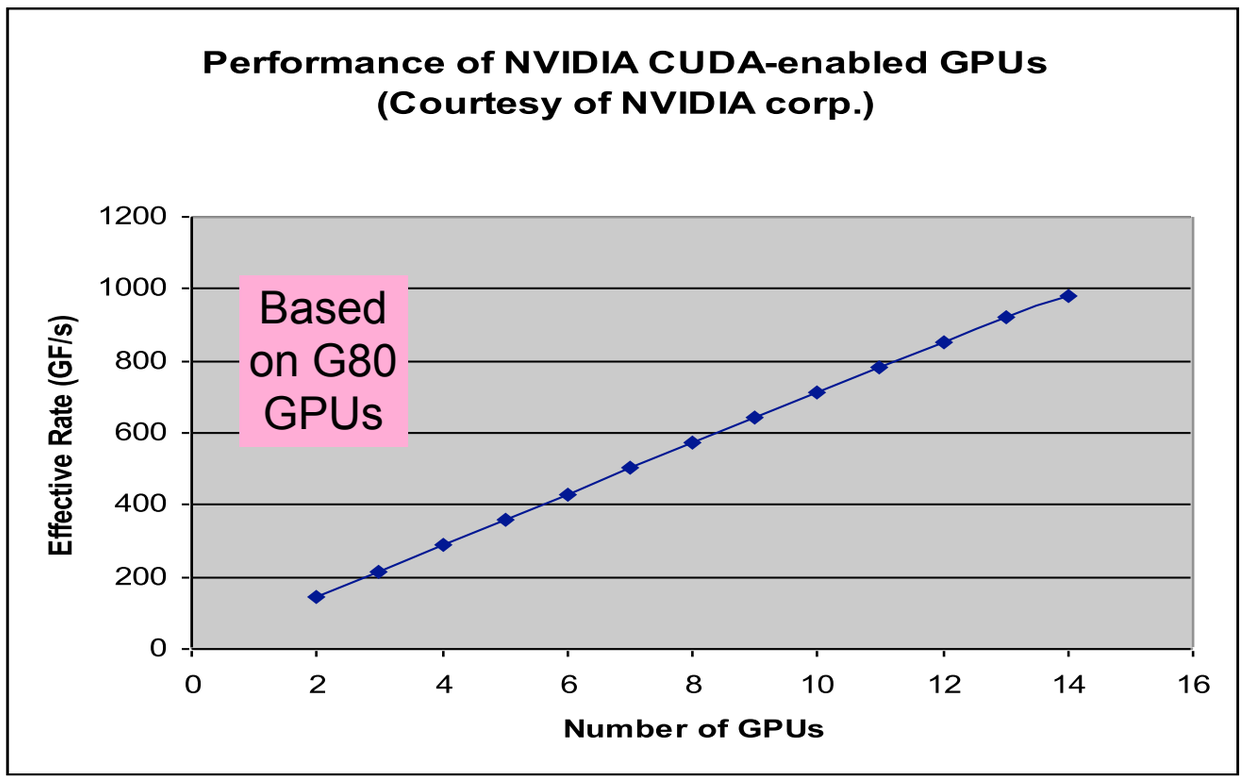
- GPUs and similar processors*
  - *Direct experience (PTC)*
    - Ported one of our in-house MD codes to NVIDIA using CUDA
    - 26-fold speed increase over host CPU
    - Significant reprogramming needed to accommodate GPU limitations
  - *Next generation GPUs will be much more capable for scientific calculations*
    - 64-bit address space and IEEE arithmetic
    - Caches
  - *Many-threaded programming model required*
    - Requires re-thinking algorithms
  - *Approaches*
    - CUDA - vendor specific
    - OpenCL - vendor independent; higher level abstraction → reduced performance
  - *Heterogeneous Multicore Parallel Programming (HMPP)*
    - Applicable to GPUs and multi-cores
    - Generates codelets that can be hand-optimized
  - *Portland Group PGI Accelerator Fortran and C99 compilers*
    - For NVIDIA GPUs

*Courtesy Dave Dixon via Randy Cygan

The Future

- GPUs
  - *MD codes ported to GPUs*
    - LAMMPS
    - HOOMD and HOOMD-Blue
    - FROMACS
    - NAMD
    - AMBER
  - *Many first principles methods*





20x Car-Parrinello (likely greater)

The Future

- Trends
  - *Use of standardized codes (LAMMPS, etc) as trajectory generators*
    - More post-simulation analysis
    - LAMMPS dump file is ~200 bytes/atom
    - 1 Million atoms → 200MB/dump
    - μs → 10^9 timesteps → 10^7-10^8 dumps → 2-20 TB of data
      - *May need to be persistent for months or even years*
      - *Might even wish to keep every timestep (e.g., Green-Kubo)*
  - *Analysis may be almost as computationally expensive as original computation*
  - *Large MD calculations need long jobs*
    - Time steps cannot be parallelized in same way as space
    - There are only 32 million seconds in a year
    - Even if an MD calculation can be parallelized to point where calculation takes no time at all, time for an MD step cannot be lowered beyond time for one communication (E.g., 0.5 μs for on-node, 7 μs for off-node of Cray XT5)
      - *At most $6x10^{13}$ timesteps = $6x10^{13} \times 10^{-15}$ s = 0.06 s for a year of dedicated HPC*
      - *In reality, orders of magnitude less*
      - *24-48 hr runs don't get very far in spanning time scales of interest*

The Wild Card

- Anton
  - *The 800-lb gorilla in the machine room*
  - *Designed and built by D. E. Shaw Research using custom application-specific integrated circuits (ASICs)*
    - Compare with GRAPE-MD
  - *1-2 orders of magnitude faster than standard hardware+ system+app stack*
  - *Game-changing for MD*
    - Pittsburgh Supercomputing Center to host an Anton for allocation by NIH
    - DOE needs an Anton

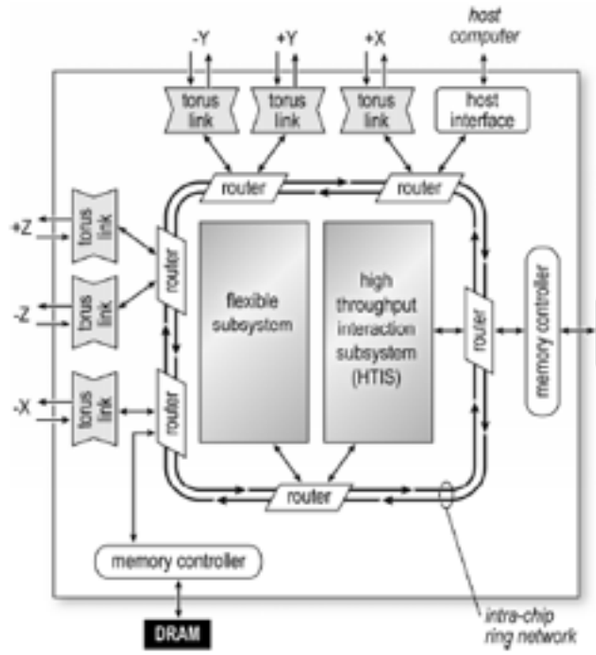


Figure 1. Anton ASIC block diagram.

Table 3. Performance of various MD platforms on the DHFR system. All simulations use the benchmark parameters specified in Table 2 of Shaw et al. [16]. Dates reflect when the measurement was made, which may differ from the date the machine commenced operation.

MD System	Processor cores	Nodes	Time (μs) per Long-Range Time Step
Anton (2008 estimate, RTL simulation)	NA	512	19
Desmond on Cluster (2006)	512	256	1,400
Blue Matter on BG/L (2006)	16,384	8,192	1,700
NAMD on Cluster (2006)	256	128	6,300
MDGRAPE-3 (2003)	NA	12	26,000
GROMACS (2007, single core, 3.2 GHz Xeon)	1	1	181,000