SAND2008-0818P

A brief survey of the LAMMPS MD code: intro, case studies, and future development

Paul Crozier, Steve Plimpton, Aidan Thompson

February 6, 2008

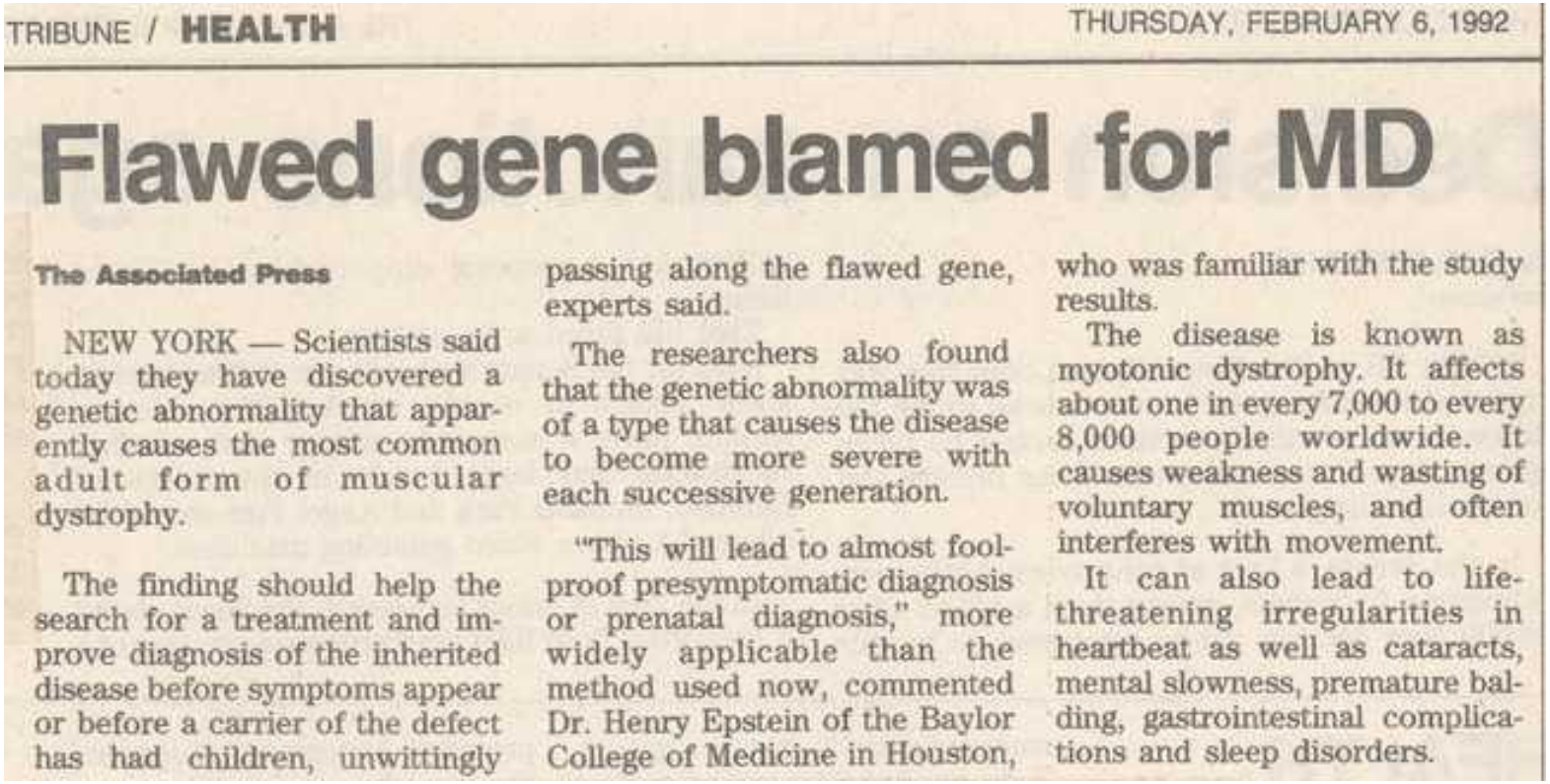
ERDC

Vicksburg, MS

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

Sandia National Laboratories

The MD Disease



TRIBUNE / **HEALTH** THURSDAY, FEBRUARY 6, 1992

Flawed gene blamed for MD

The Associated Press

NEW YORK — Scientists said today they have discovered a genetic abnormality that apparently causes the most common adult form of muscular dystrophy.

The finding should help the search for a treatment and improve diagnosis of the inherited disease before symptoms appear or before a carrier of the defect has had children, unwittingly passing along the flawed gene, experts said.

The researchers also found that the genetic abnormality was of a type that causes the disease to become more severe with each successive generation.

"This will lead to almost foolproof presymptomatic diagnosis or prenatal diagnosis," more widely applicable than the method used now, commented Dr. Henry Epstein of the Baylor College of Medicine in Houston, who was familiar with the study results.

The disease is known as myotonic dystrophy. It affects about one in every 7,000 to every 8,000 people worldwide. It causes weakness and wasting of voluntary muscles, and often interferes with movement.

It can also lead to life-threatening irregularities in heartbeat as well as cataracts, mental slowness, premature balding, gastrointestinal complications and sleep disorders.

Sandia National Laboratories

A brief introduction to MD

- MD: molecular dynamics
- F = ma
- Classical dynamics
- Rapidly grown in popularity and use in research
- Computationally intensive, especially computation of nonbonded interactions
- Uses force fields: mathematical models of interatomic interactions

MD uses empirical force fields

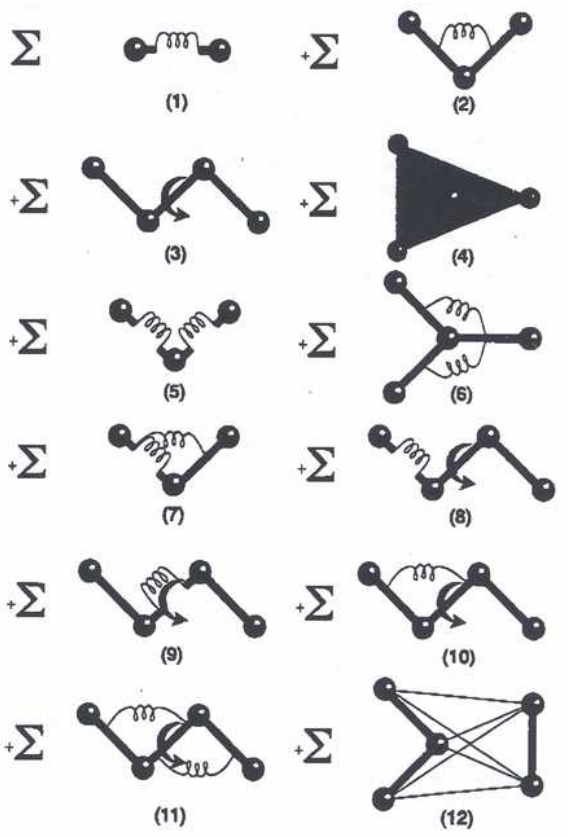
- Particles interact via empirical potentials
  - analytic equations, fast to compute
  - coefficients fit to expt or quantum calcs

$$E = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \qquad r < r_c$$

$$E = K(r - r_0)^2$$

- Potential energy = Φ = f(x)
- Force = -Grad Φ

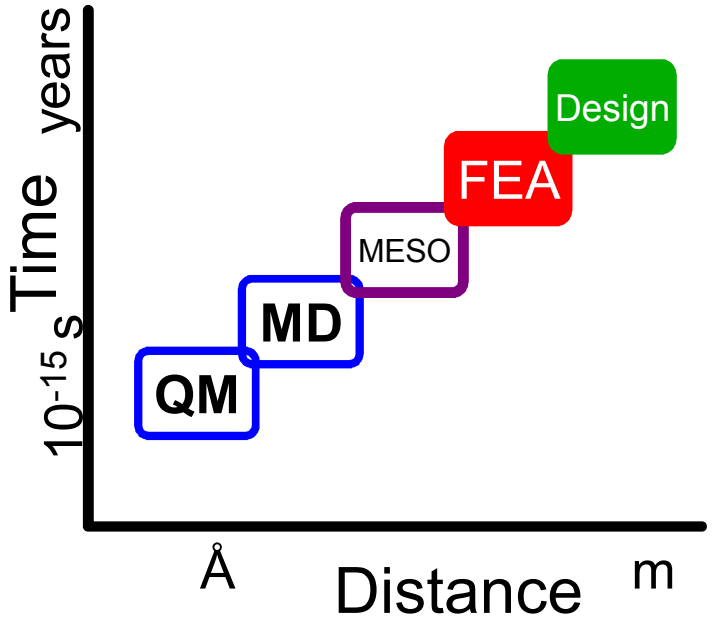
- Pair-wise forces
  - Van der Waals (dipole-dipole)
  - Coulombic (charge-charge)
- Many-body forces
  - EAM, Tersoff, bond-order, ReaxFF
- Molecular forces
  - springs, torsions, dihedrals, ...
- Long-range Coulombic forces
  - Ewald, particle-mesh methods, FFTs



Sandia National Laboratories

MD in the Middle

- Quantum mechanics
  - electronic degrees of freedom, chemical reactions
  - Schrodinger equation, wave functions
  - sub-femtosecond timestep, 1000s of atoms, $O(N^3)$
- Atomistic models
  - molecular dynamics (MD), Monte Carlo (MC)
  - point particles, empirical forces, Newton's equations
  - femtosecond timestep, millions of atoms, O(N)
- Mesoscale to Continuum
  - finite elements or finite difference on grids
  - coarse-grain particles: DPD, PeriDynamics, ...
  - PDEs, Navier-Stokes, stress-strain
  - microseconds → seconds, microns → meters, $O(N^{3/2})$



Sandia National Laboratories

Algorithmic Issues in MD

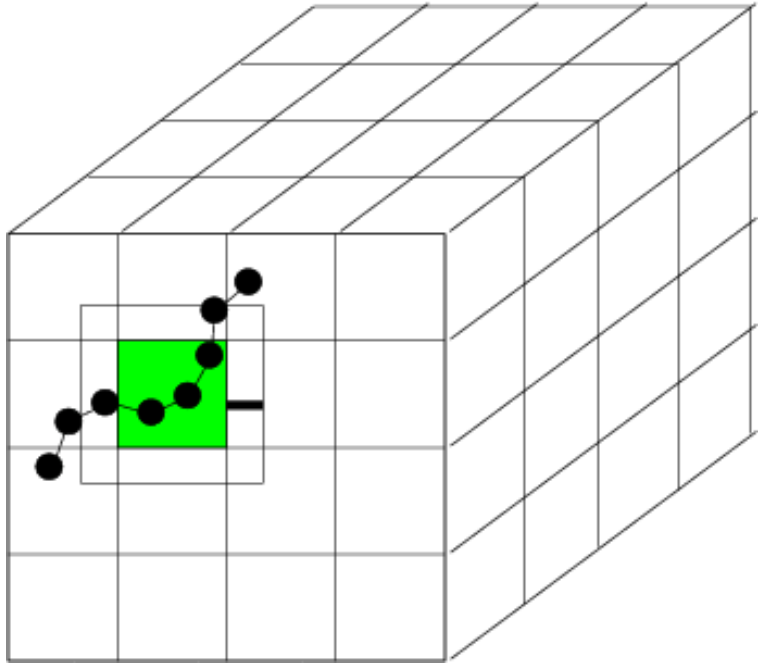
- Speed
  - parallel implementation
- Accuracy
  - long-range Coulombics
- Time scale
  - slow versus fast degrees of freedom
- Length scale
  - coarse-graining

Classical MD in Parallel

- MD is inherently parallel
  - forces on each atom can be computed simultaneously
  - X and V can be updated simultaneously
- Most MD codes are parallel
  - via distributed-memory message-passing paradigm (MPI)
- Computation scales as N = number of atoms
  - ideally would scale as N/P in parallel
- Can distribute:
  - atoms communication = scales as N
  - forces communication = scales as N/sqrt(P)
  - space communication = scales as N/P or $(N/P)^{2/3}$

Parallelism via Spatial-Decomposition

- Physical domain divided into 3d boxes, one per processor
- Each proc computes forces on atoms in its box
  using info from nearby procs
- Atoms "carry along" molecular topology
  as they migrate to new procs
- Communication via
  nearest-neighbor 6-way stencil



- Optimal scaling for MD: N/P
  so long as load-balanced

- Computation scales as N/P
- Communication scales
  sub-linear as $(N/P)^{2/3}$
  (for large problems)
- Memory scales as N/P

Sandia National Laboratories

A brief introduction to LAMMPS

LAMMPS: Large-scale Atomic/Molecular Massively Parallel Simulator

- **Massively parallel, general purpose MD code**
- **Developed at Sandia National Laboratories, with contributions from many labs throughout the world**
- **Freely available for download under GPL**

  lammps.sandia.gov

  Over 20,000 downloads since September 2004

  Open source, easy to understand C++ code

  Easily extensible

Sandia National Laboratories

How to download, install, and use LAMMPS

- **Download page:**
  lammps.sandia.gov/download.html

- **Installation instructions:**
  lammps.sandia.gov/doc/Section_start.html
  go to lammps/src
  type "make *your_system_type*"

- **To perform a simulation:**
  lmp < *my_script.in*

Sandia National Laboratories

How to get help with LAMMPS

1. **Excellent User's Manual:**
   lammps.sandia.gov/doc/Manual.html

2. **User's e-mail list:**
   lammps.sandia.gov/mail.html

3. **Contact the developers:**
   Steve Plimpton, sjplimp@sandia.gov
   Aidan Thompson, athomps@sandia.gov
   Paul Crozier, pscrozi@sandia.gov

Sandia National Laboratories

Force fields available in LAMMPS

- **Biomolecules:** CHARMM, AMBER, OPLS, COMPASS (class 2), long-range Coulombics via PPPM, point dipoles, ...
- **Polymers:** all-atom, united-atom, coarse-grain (bead-spring FENE), bond-breaking, …
- **Materials:** EAM and MEAM for metals, Buckingham, Morse, Yukawa, Stillinger-Weber, Tersoff, Al-REBO, ...
- **Mesoscale:** granular, DPD, Gay-Berne, colloidal, peri-dynamics, ...
- **Hybrid:** can use combinations of potentials for hybrid systems: water on metal, polymers/semiconductor interface, colloids in solution, …

Sandia National Laboratories

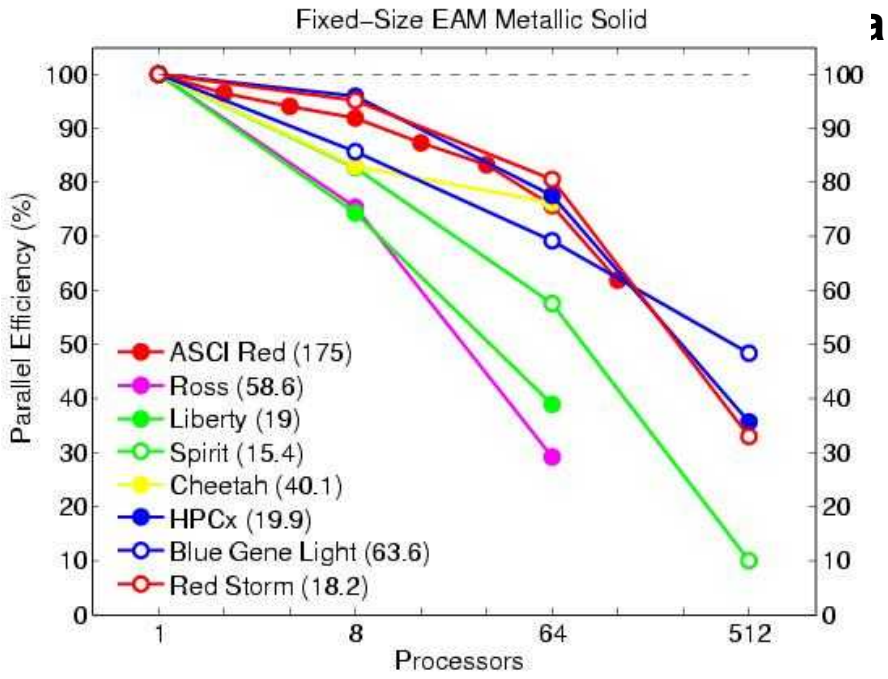
Easily add your own LAMMPS feature

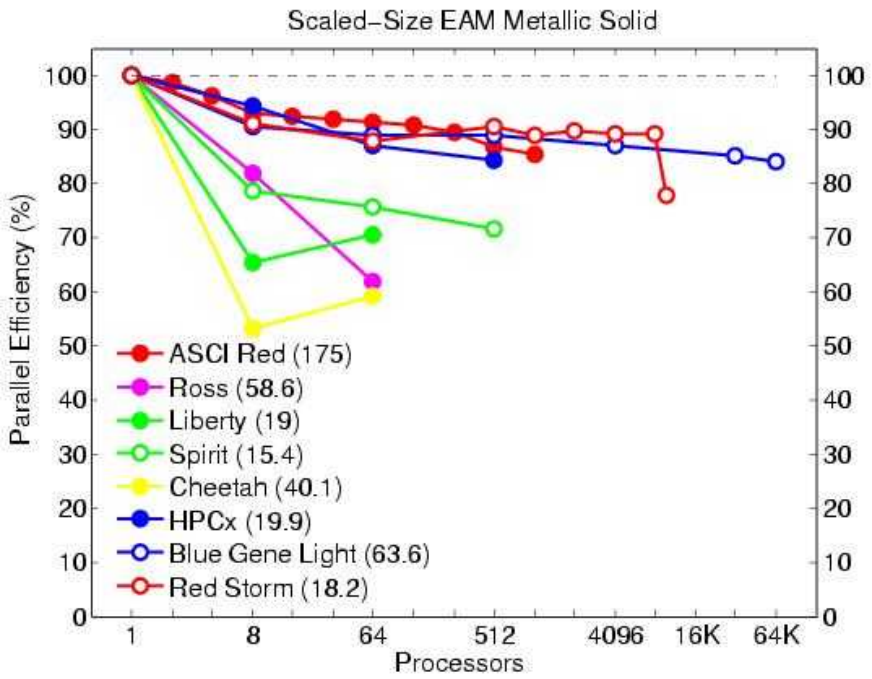
- New user or new simulation → always want new feature not in code
- Goal: make it as easy as possible for us and others to add new features
  - called “styles” in LAMMPS:
  - particle type, pair or bond potential, scalar or per-atom computation
  - "fix": BC, force constraint, time integration, diagnostic, ...
  - input command: create_atoms, set, run, temper, ...
  - 75% of current 100K lines of LAMMPS is add-on styles
- Enabled by C++
  - "virtual" parent class for all pair potentials
  - defines interface: compute(), coeff(), restart(), ...
  - add feature: add 2 lines to header file, add files to src dir, re-compile
  - feature won't exist if not used, won't conflict with rest of code
- Of course, someone has to write the code for the feature!

Sandia National Laboratories

LAMMPS's parallel performance

- **Fixed-size (32K atoms) and scaled-size (32K atoms/proc) parallel efficiencies**





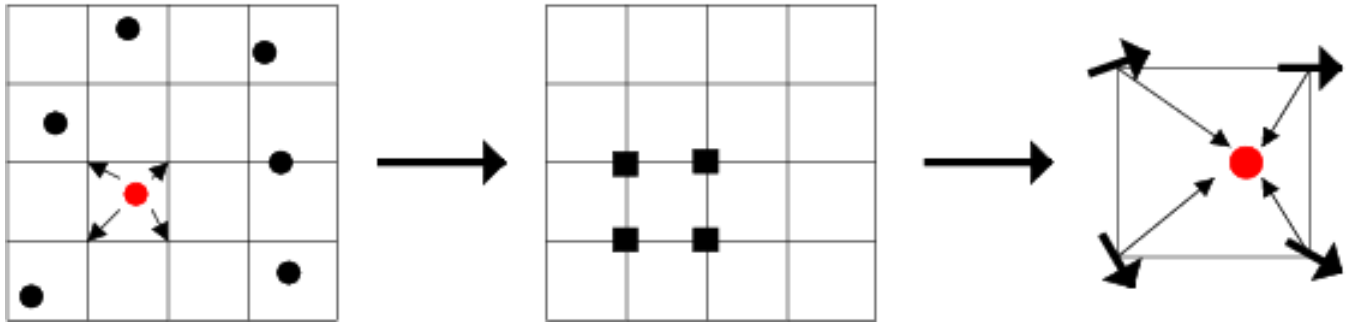
- **Billions of atoms on 64K procs of Blue Gene or Red Storm**
- **Opteron processor speed: 5.7E-6 sec/atom/step (0.5x for LJ, 12x for protein)**

Sandia National Laboratories

Particle-mesh Methods for Coulombics

- Coulomb interactions fall off as 1/r so require long-range for accuracy

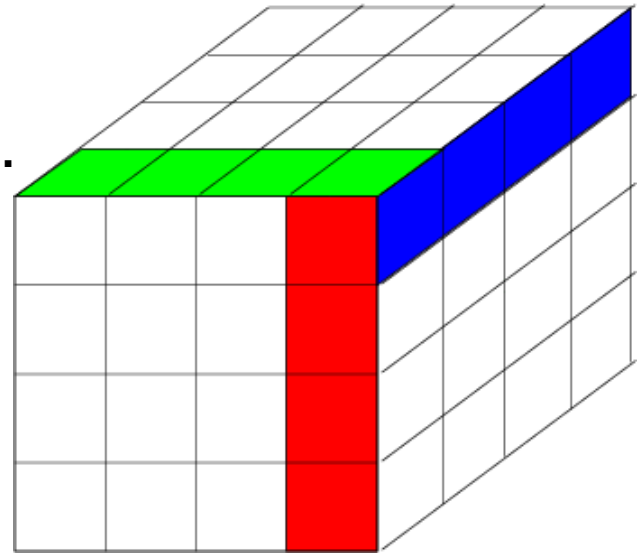
- Particle-mesh methods:
  - partition into short-range and long-range contributions
  - short-range via direct pairwise interactions
  - long-range:
    - interpolate atomic charge to 3d mesh
    - solve Poisson's equation on mesh (4 FFTs)
    - interpolate E-fields back to atoms



- FFTs scale as NlogN if cutoff is held fixed

Parallel FFTs

- 3d FFT is 3 sets of 1d FFTs
  - in parallel, 3d grid is distributed across procs
  - perform 1d FFTs on-processor
    - native library or FFTW (www.fftw.org)
  - 1d FFTs, transpose, 1d FFTs, transpose, ...
    - "transpose" = data transfer
    - transfer of entire grid is costly



- FFTs for PPPM can scale poorly
  - on large # of procs and on clusters

- Good news: Cost of PPPM is only ~2x more than 8-10 Angstrom cutoff

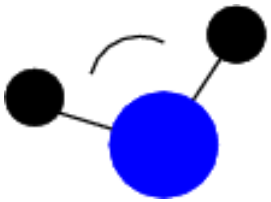
Sandia National Laboratories

Time Scale of Molecular Dynamics

- Limited timescale is most serious drawback of MD
- Timestep size limited by atomic oscillations:
  - C-H bond = 10 fmsec → ½ to 1 fmsec timestep
  - Debye frequency = 10^{13} → 2 fmsec timestep
- A state-of-the-art "long" simulation is nanoseconds to a microsecond of real time
- Reality is usually on a much longer timescale:
  - protein folding (msec to seconds)
  - polymer entanglement (msec and up)
  - glass relaxation (seconds to decades)

Extending Timescale

- SHAKE = bond-angle constraints, freeze fast DOF
  - up to 2-3 fmsec timestep
  - rigid water, all C-H bonds
  - extra work to enforce constraints

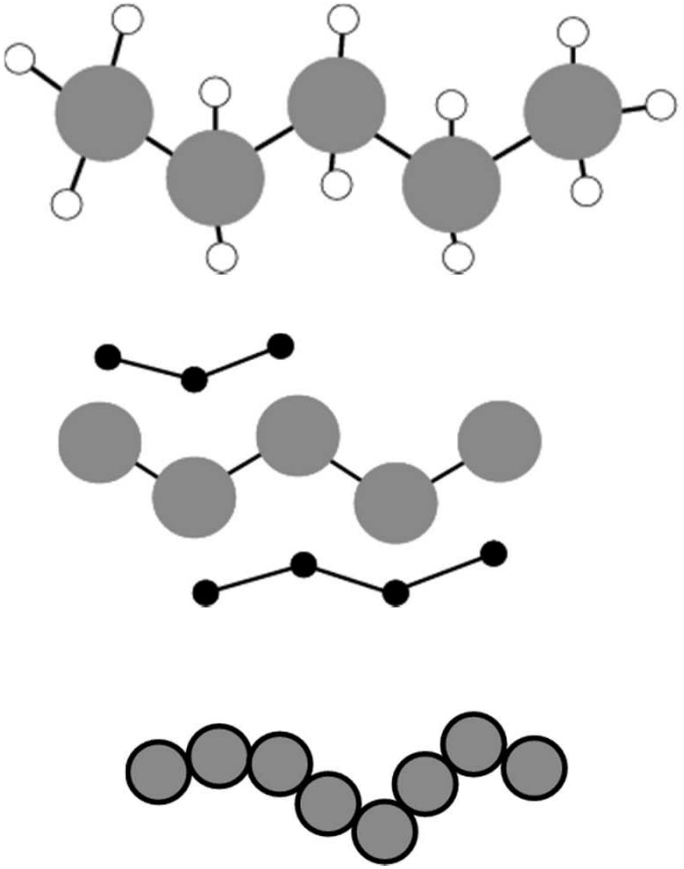


- rRESPA = hierarchical time stepping, sub-cycle on fast DOF
  - inner loop on bonds (0.5 fmsec)
  - next loop on angle, torsions (3-4 body forces)
  - next loop on short-range LJ and Coulombic
  - outer loop on long-range Coulombic (4 fmsec)

- Rigid body time integration via quaternions
  - treat groups of atom as rigid bodies (portions of polymer or protein)
  - 3N DOF → 6 DOF
  - save computation of internal forces, longer timestep

Length Scale of Molecular Dynamics

- Limited length scale is 2nd most serious drawback of MD → coarse-graining
- All-atom:
  - Δt = 0.5-1.0 fmsec for C-H
  - C-C distance = 1.5 Angs
  - cutoff = 10 Angs
- United-atom:
  - # of interactions is 9x less
  - Δt = 1.0-2.0 fmsec for C-C
  - cutoff = 10 Angs
  - 20-30x savings over all-atom
- Bead-Spring:
  - 2-3 C per bead
  - Δt ←→ fmsec mapping is T-dependent
  - $2^{1/6}\,\sigma$ cutoff → 8x in interactions
  - can be considerable savings over united-atom



Sandia National Laboratories

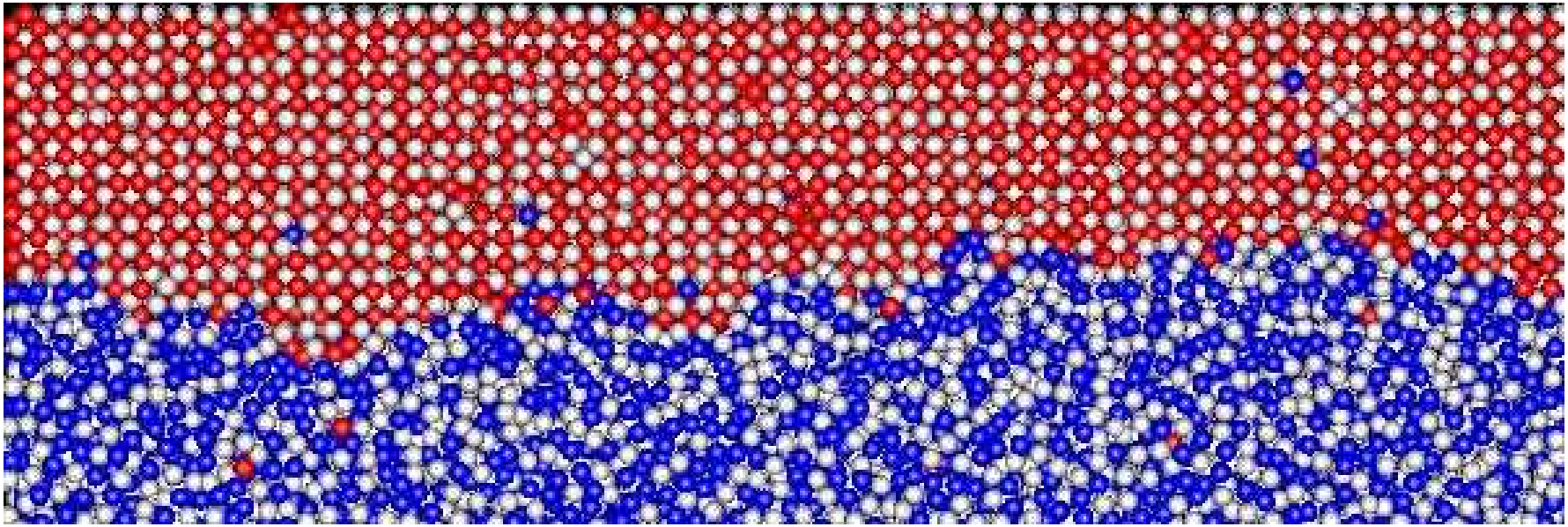
Atomistic Scale Models with LAMMPS

- Interfaces in melting solids
- Adhesion properties of polymers
- Shear response in metals
- Tensile pull on nanowires
- Surface growth on mismatched lattice
- Shock-induced phase transformations
- Silica nanopores for water desalination
- Coated nanoparticles in solution and at interfaces
- Self-assembly (2d micelles and 3d lipid bilayers)
- Rhodopsin protein isomerization

Sandia National Laboratories

Melt Interface in NiAl

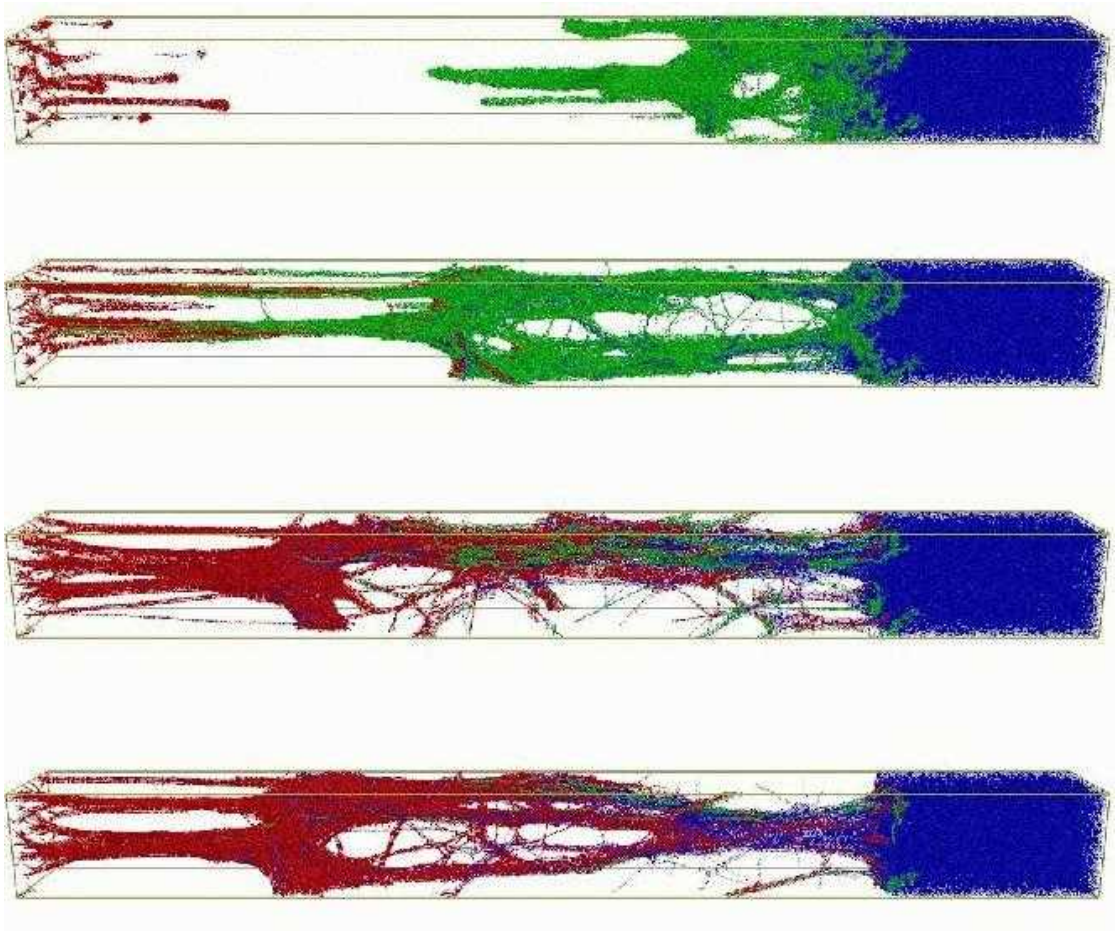
- Mark Asta (UC Davis) and Jeff Hoyt (Sandia)
- Careful thermostatting and equilibration of alloy system
- Track motion and structure of melt interface



Sandia National Laboratories

Polymer Adhesive Properties

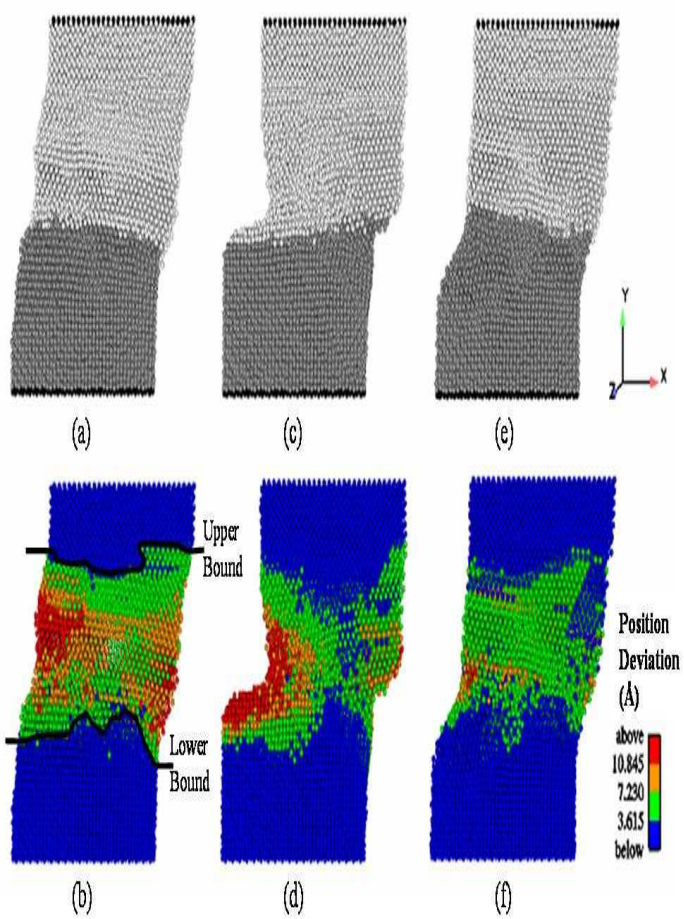
- Mark Stevens and Gary Grest (Sandia)
- Bead/spring polymer model, allow for bond breaking



Sandia National Laboratories

Shear Response of Cu Bicrystal

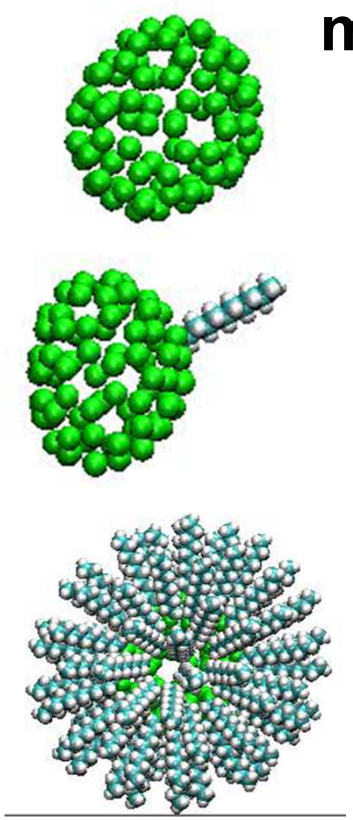
- David McDowell group (GA Tech)
- Defect formation, stress relaxation, energetics of boundary region

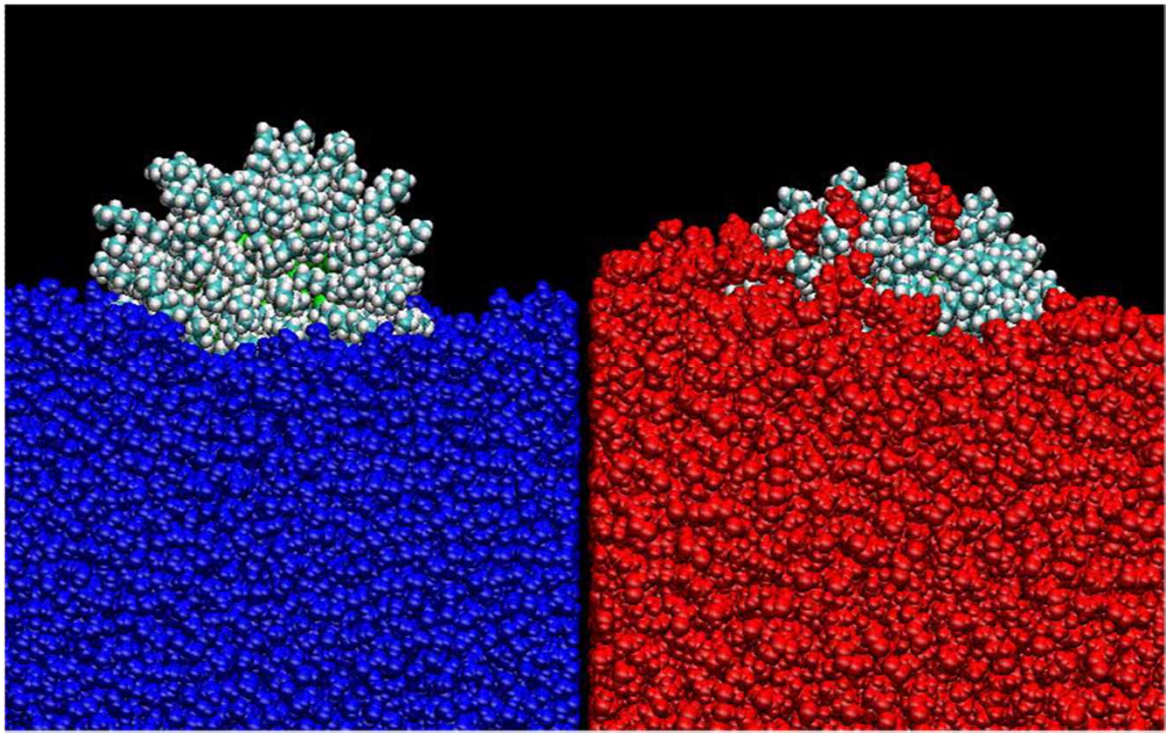


Sandia National Laboratories

Coated Nanoparticles at Interfaces

- Matt Lane, Gary Grest (Sandia)
- S sites on Au nanoparticle, alkane-thiol chains, methyl-terminated, 3 ns sim





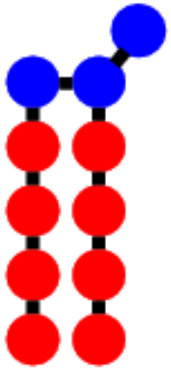
water

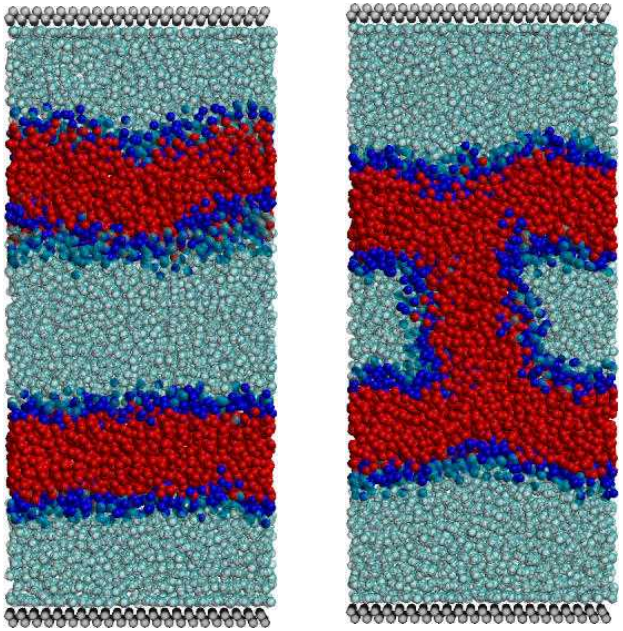
decane

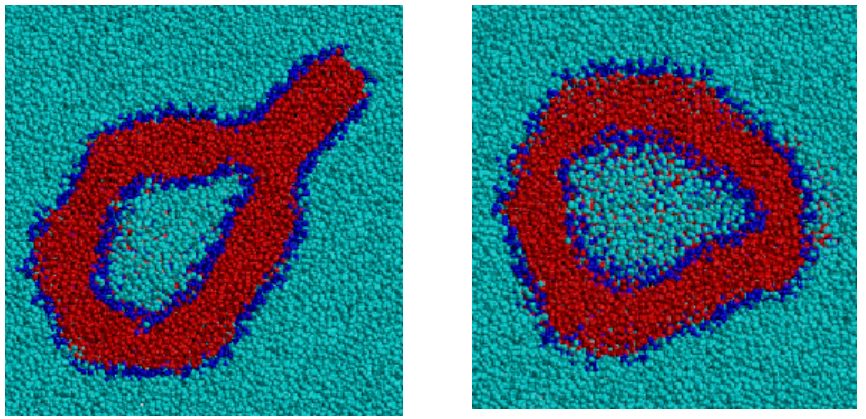
Sandia National Laboratories

3d Membrane Self-Assembly

- Mark Stevens (Sandia)
- Coarse-grain lipid model in monomeric solvent
- Angle terms for rigidity
- Hydrophilic head-group & solvent, hydrophobic tail
- 100Ks of particles for millions of timesteps
- Bilayer & vesicle formation



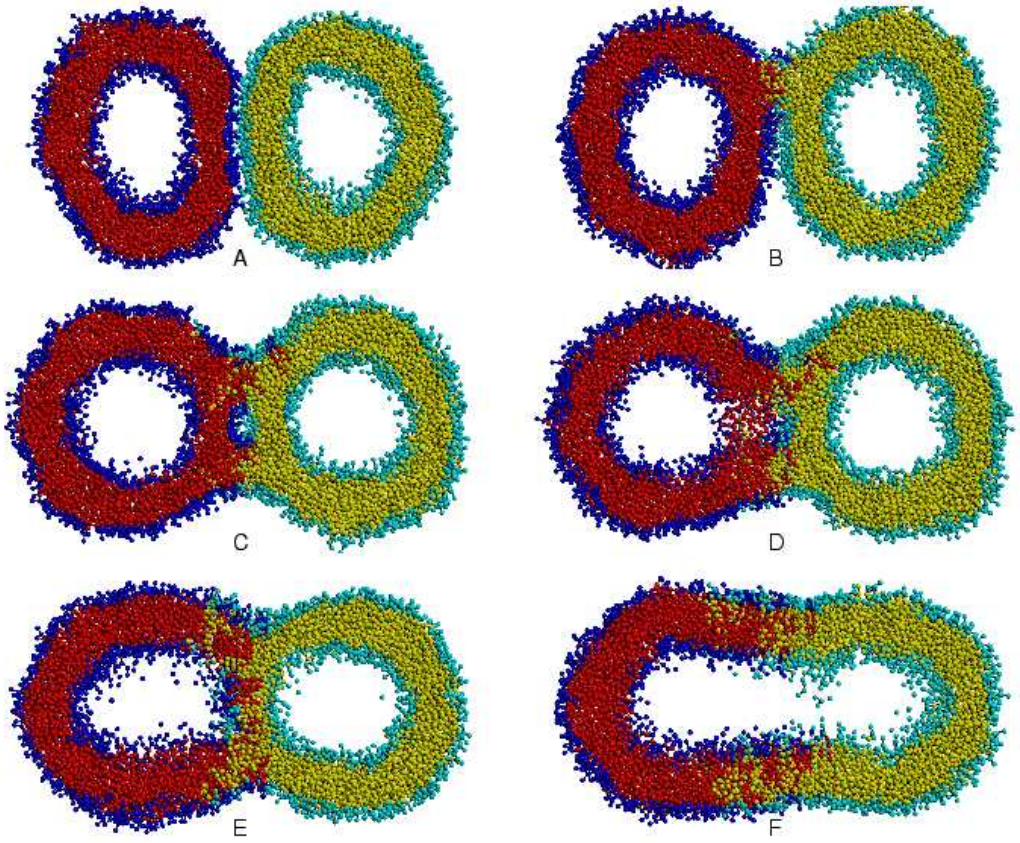


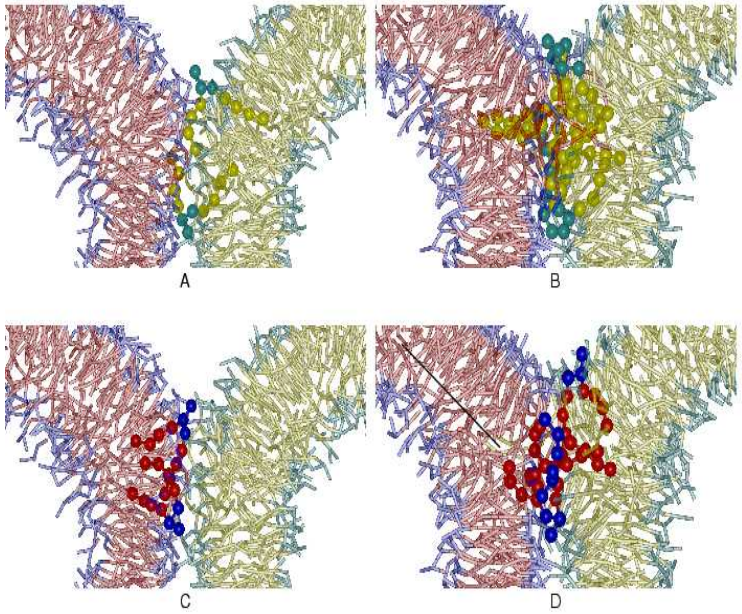


15K monomers for 1M steps

Sandia National Laboratories

Membrane Fusion

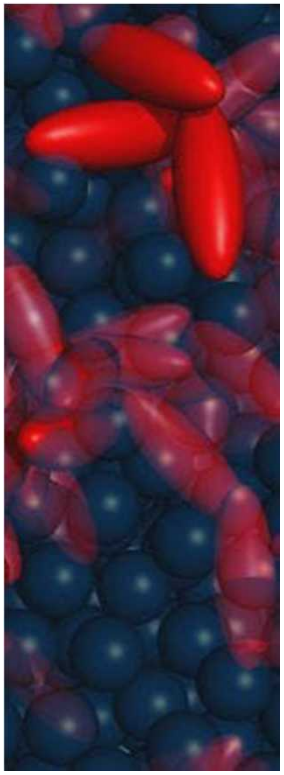


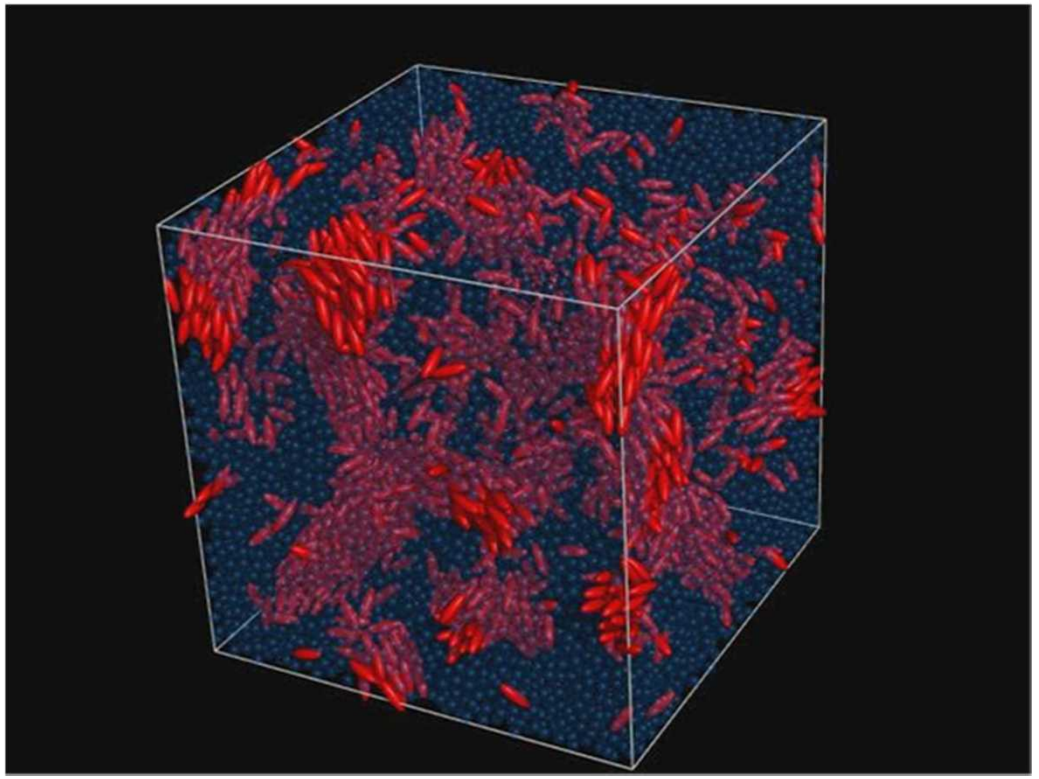


Sandia National Laboratories

Aspherical Nanoparticles

- Mike Brown (Sandia)
- Ellipsoidal particles interacting via Gay-Berne potentials (LC), LJ solvent
- Nanodroplet formation in certain regimes of phase space

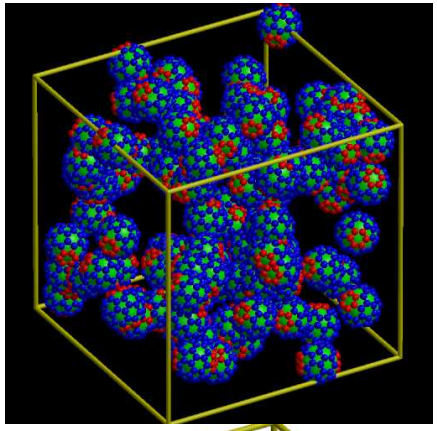


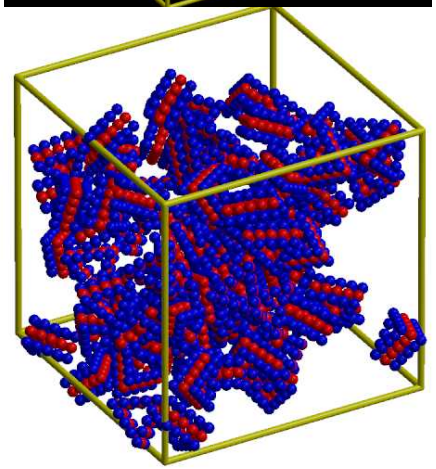


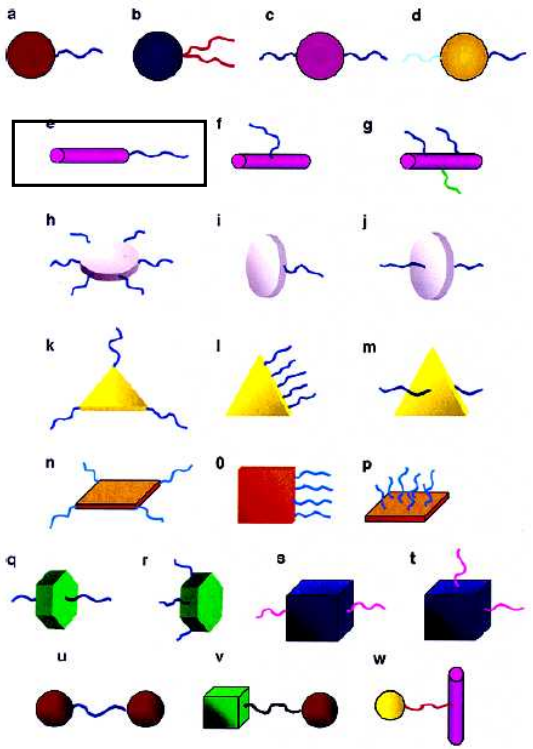
Sandia National Laboratories

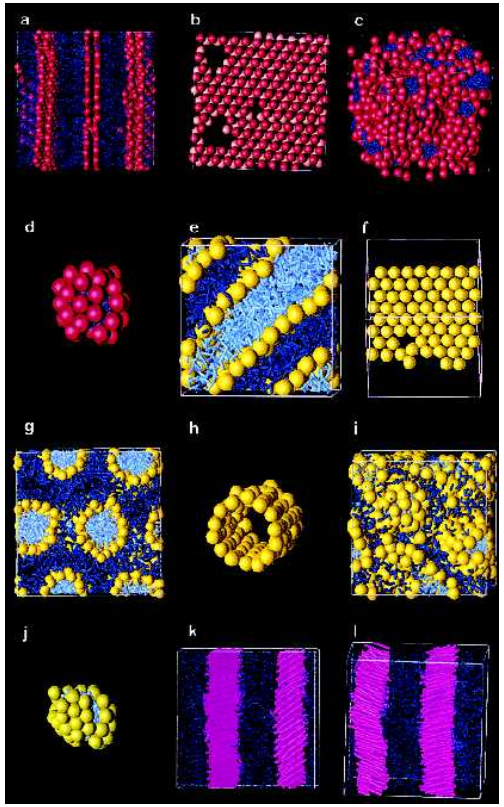
Rigid Nanoparticle Self-Assembly

- Multiple rigid bodies
- Quaternion integration
- Brownian dynamics
- Self-assembly → phases







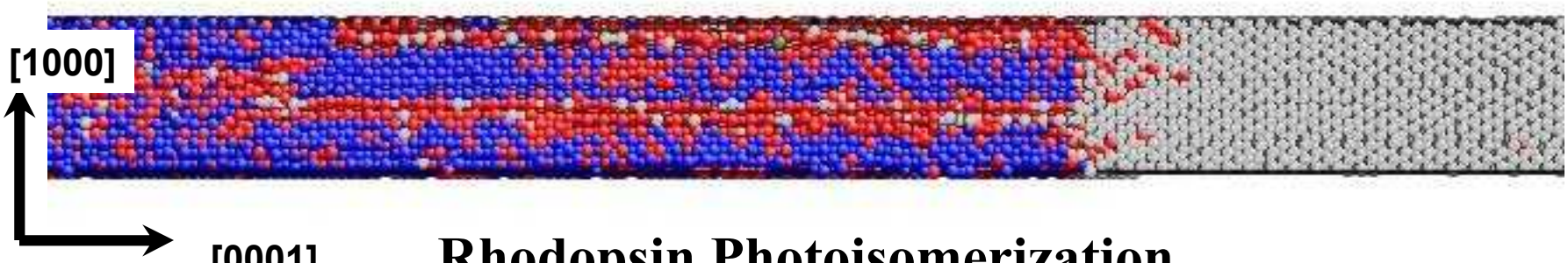


(Sharon Glotzer *et al.*, *Nano Letters*, **3**, 1341 (2003).

Sandia National Laboratories

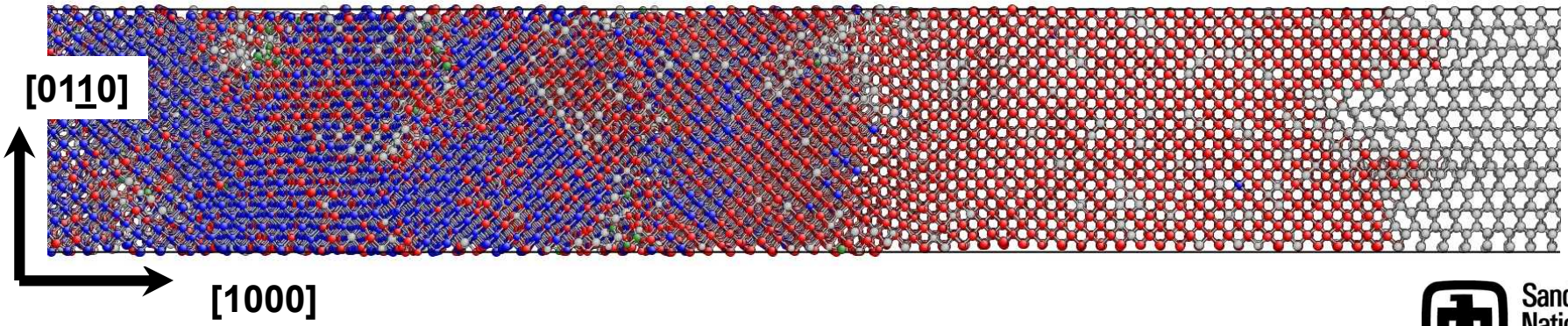
MD Simulation of Shock-induced Structural Phase Transformation in Cadmium Selenide

c-direction: 2-Wave Structure: rocksalt emerges directly from elastically compressed material



Rhodopsin Photoisomerization

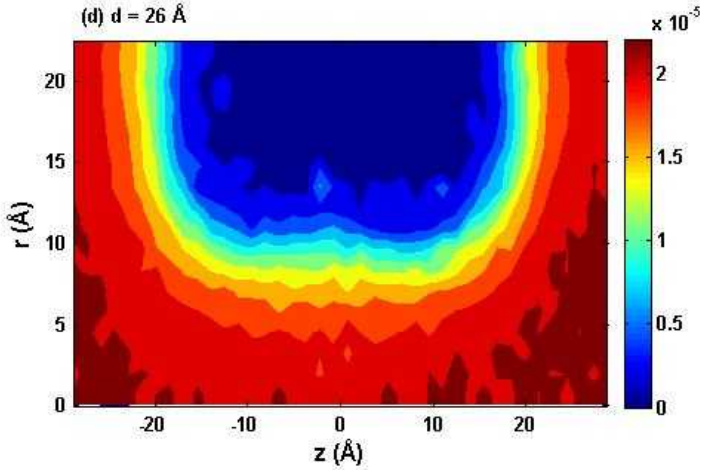
a-direction: 3-Wave Structure: tetragonal region forms between elastic wave and rocksalt phase



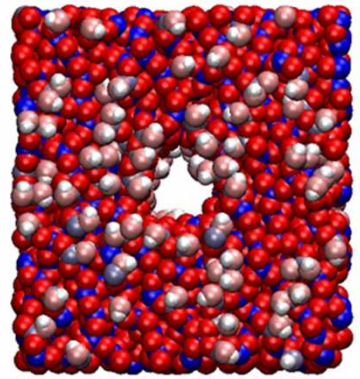
Sandia National Laboratories

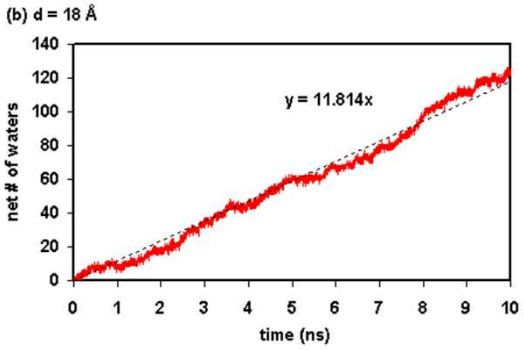
Non-equilibrium MD simulations of brackish water flow through silica and titania nanopores

- Small flow field successfully induces steady state solvent flow through amorphous SiO_2 and TiO_2 nanopores in NEMD simulations.
- Complex model systems built through a detailed processs involving melting, quenching, annealing, pore drilling, defect capping, and equilibration.
- 10-ns simulations carried out for a variety of pore diameters for for both SiO_2 and TiO_2 nanopores.
- Densities, diffusivities, and flows of the various species computed spatially, temporally, and as a function of pore diameter.

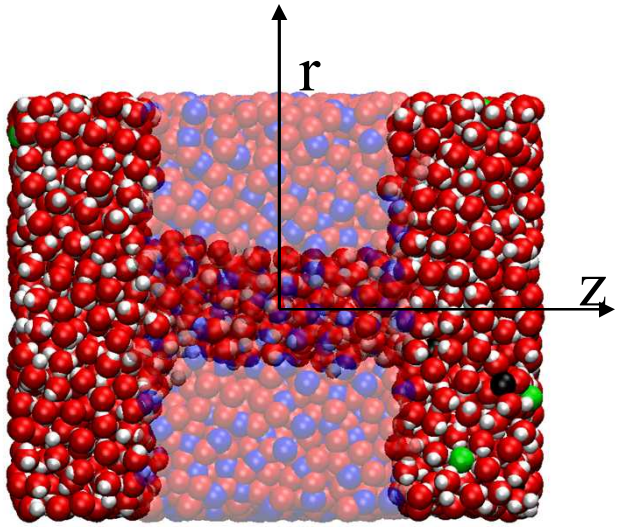


Spatial map of water diffusivities in a 26 Å TiO_2 nanopore.





Water flux through an 18 Å TiO_2 nanopore.

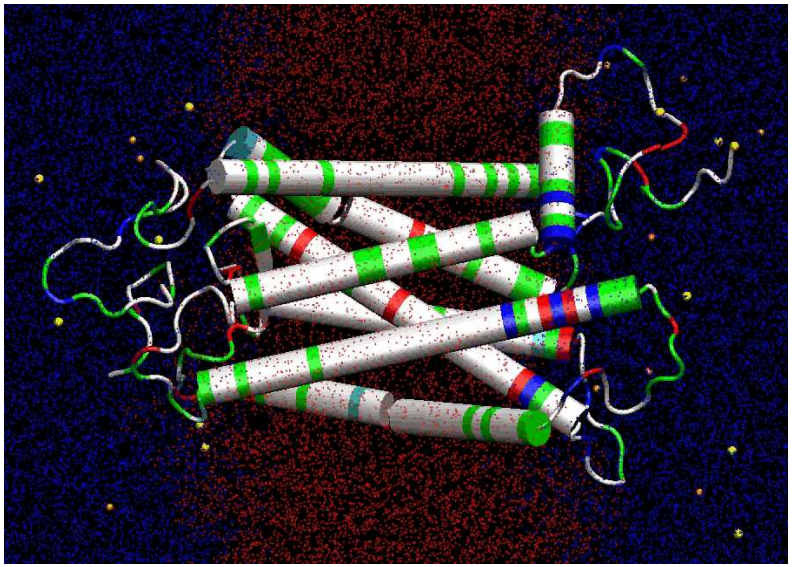


- Water is tightly bound to hydrophilic TiO_2 surface, greatly hampering mobility within 5 Å of the surface.
- Simulations show that amorphous nanopores of diameter at least 14 Å can conduct water as well as Na+ and Cl- ions.
- No evidence of selectivity that allows water passage and precludes ion passage --- functional groups on pore interior may be able to achieve this.

Sandia National Laboratories

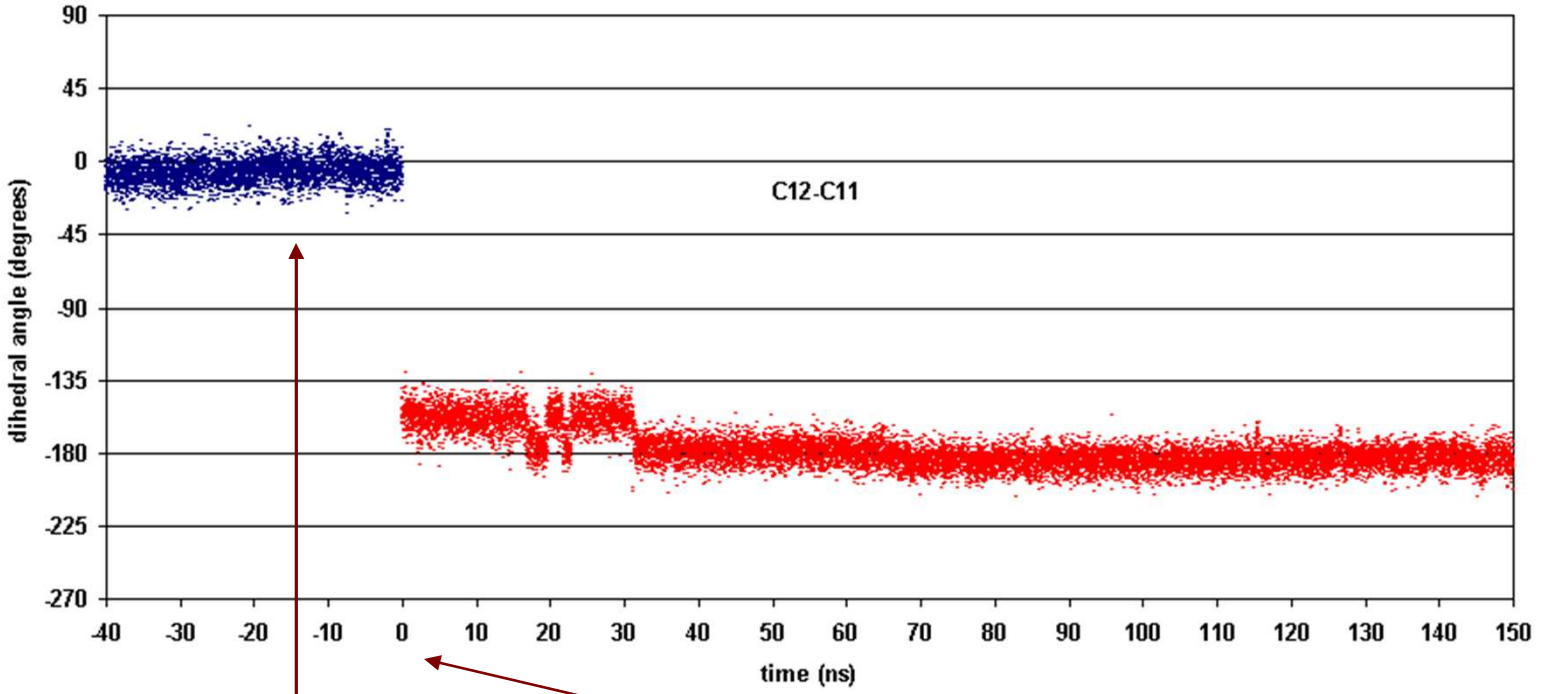
Rhodopsin photoisomerization simulation

- 190 ns simulation
  - 40 ns in dark-adapted state (*J. Mol. Biol.*, **333**, 493, (2003))
  - 150 ns after photoisomerization
- CHARMM force field
- P^3M full electrostatics
- Parallel on ~40 processors; more than 1 ns simulation / day of real time
- Shake, 2 fs time step, velocity Verlet integrator
- Constant membrane surface area
- System description
  - All atom representation
  - 99 DOPC lipids
  - 7441 TIP3P waters
  - 348 rhodopsin residues
  - 41,623 total atoms
  - L_x=55 Å, L_y=77 Å, L_z=94-98 Å



Sandia National Laboratories

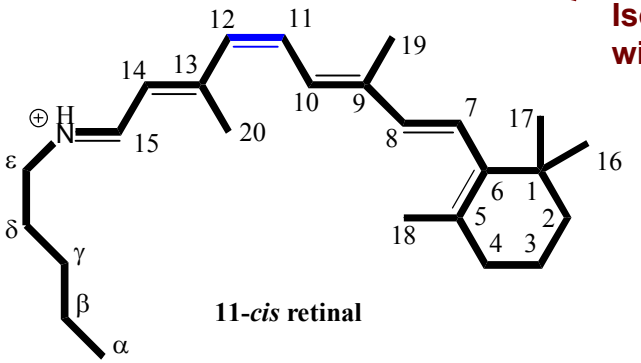
Photoisomerization of retinal

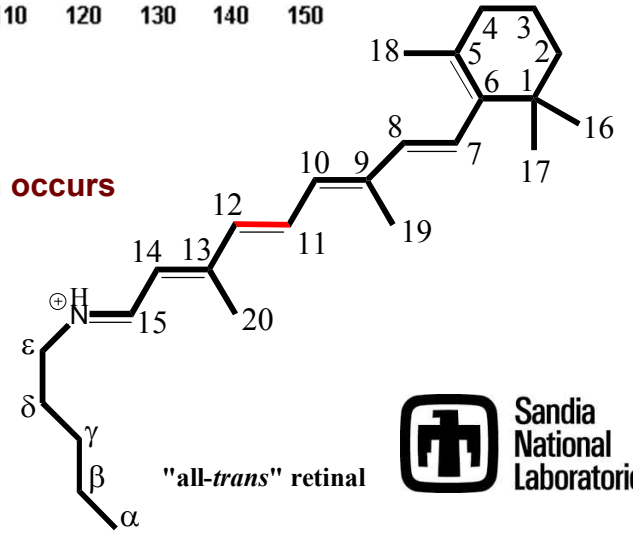


Dihedral remains in *trans* during subsequent 150 ns of relaxation

40-ns simulation in dark-adapted state

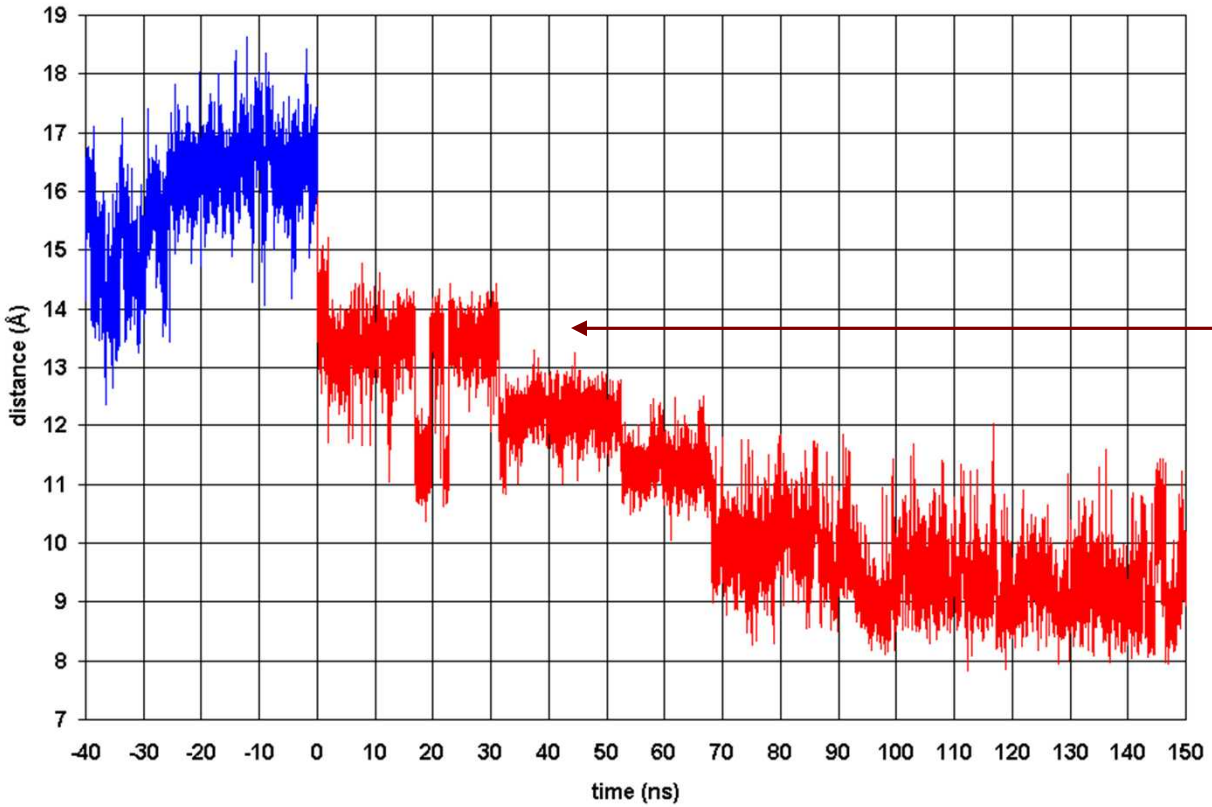
Isomerization occurs within 200 fs.





Sandia National Laboratories

Distance between β-ionone ring and Ala 169



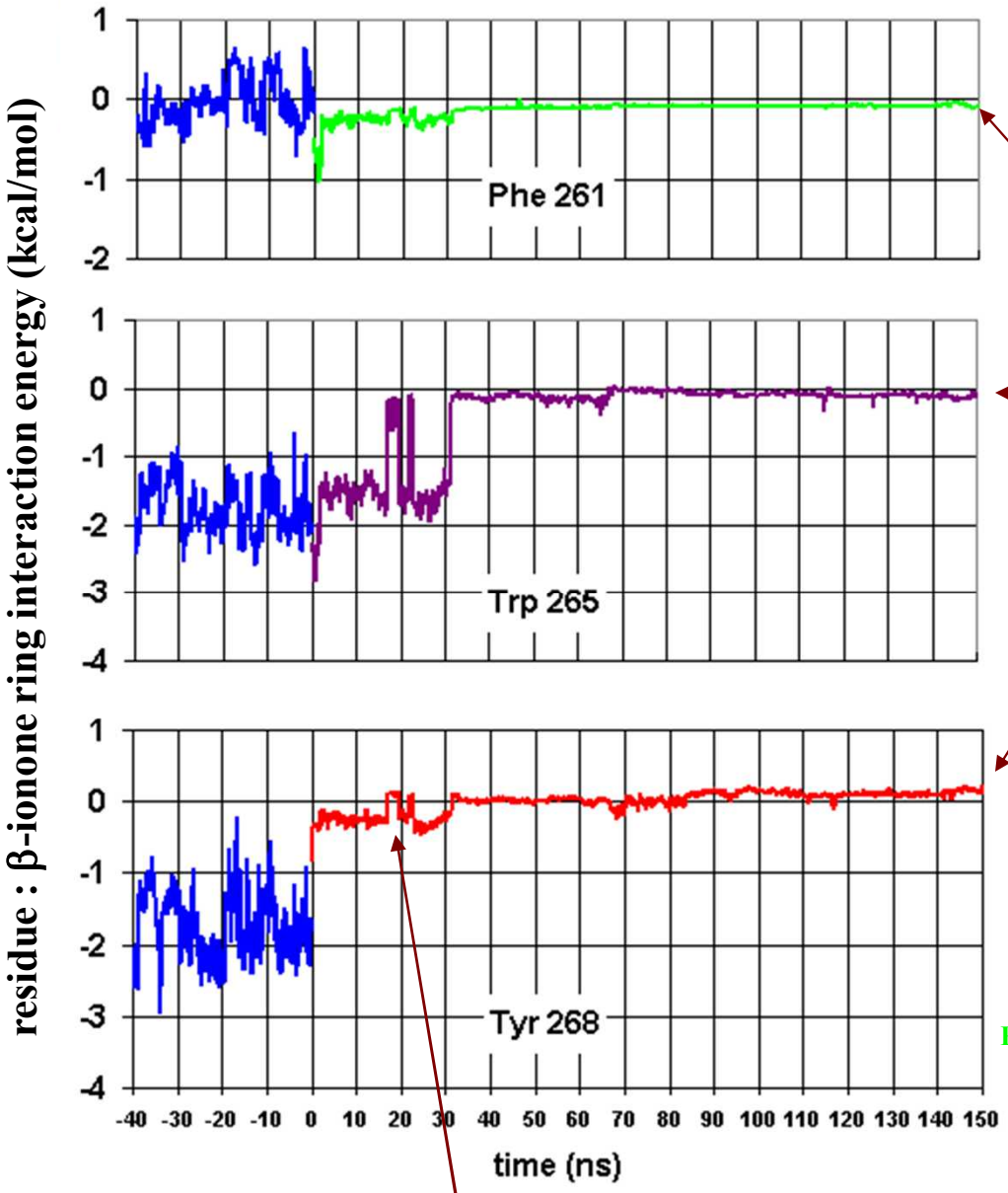
A sharp decrease in the distance is observed, in agreement with cross-linking experiments
(Borhan et al. *Science*, **288**, 2209, (2000))

Correlation with retinal movements indicates that initial decrease in distance is due mostly to movement of the β-ionone ring

Slow rotation of helix 4 will close much of the remaining distance

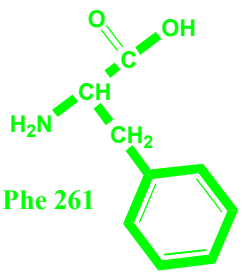
Sandia National Laboratories

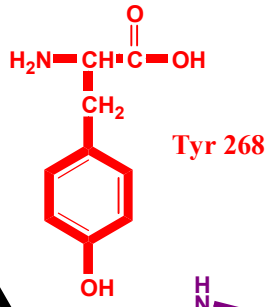
Retinal : helix 6 sidechain interactions

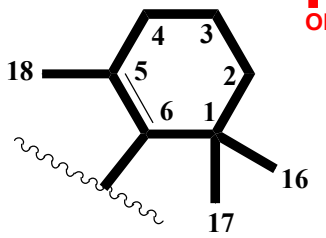


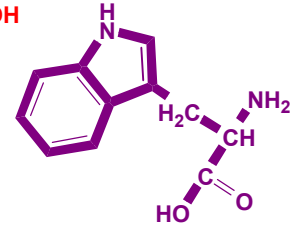
Interactions approach zero --- decoupling of helix 6 from retinal will allow subsequent large-scale helix 6 movement
(Farrens et al., *Science* **274**, 768, (1996))

β-ionone ring moved away from helix 6 sidechain rings during photoisomerization







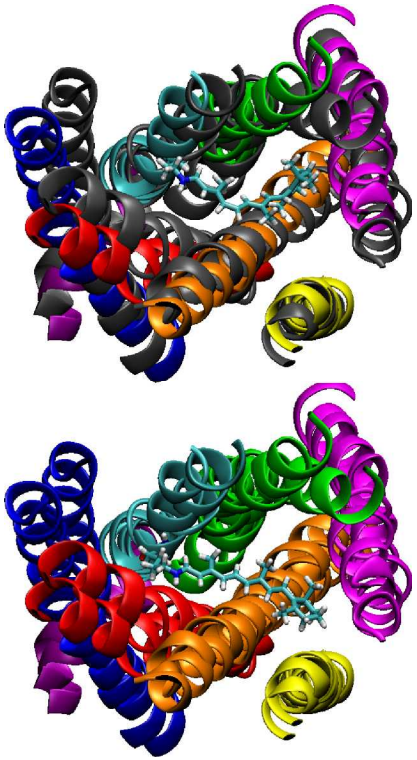


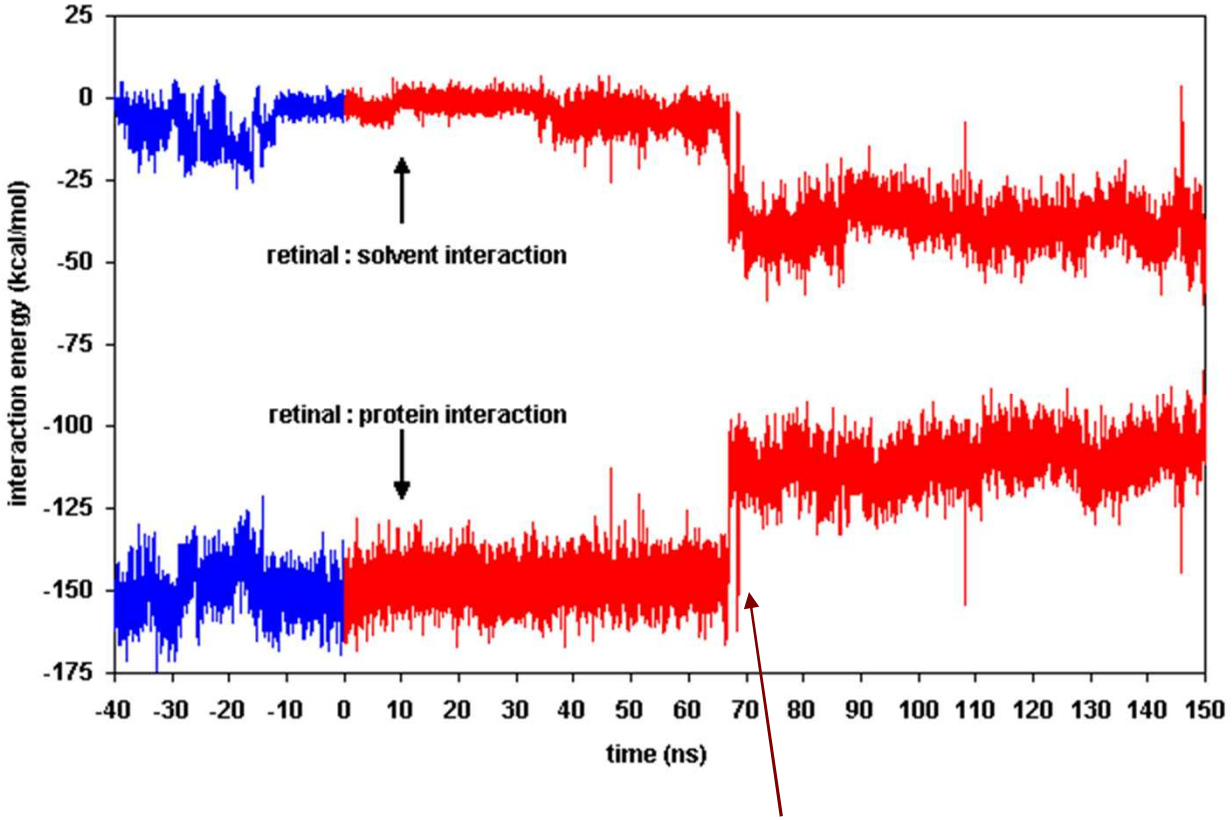
Relaxation of the retinal chain moves β-ionone even further from helix 6 sidechain rings

Sandia National Laboratories

Transition in retinal's interaction environment

Retinal's interaction with the rest of the rhodopsin molecule weakens and is partially compensated by a stronger solvent interaction

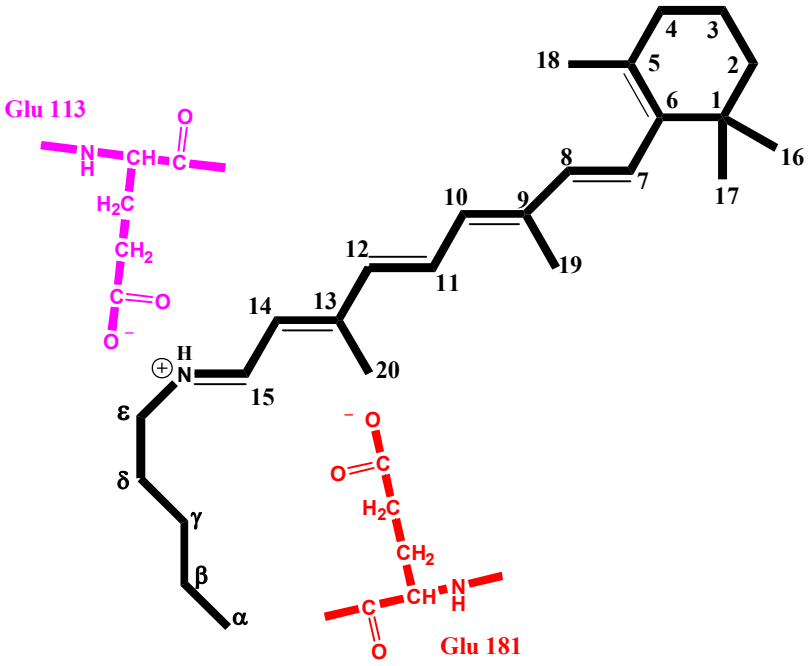


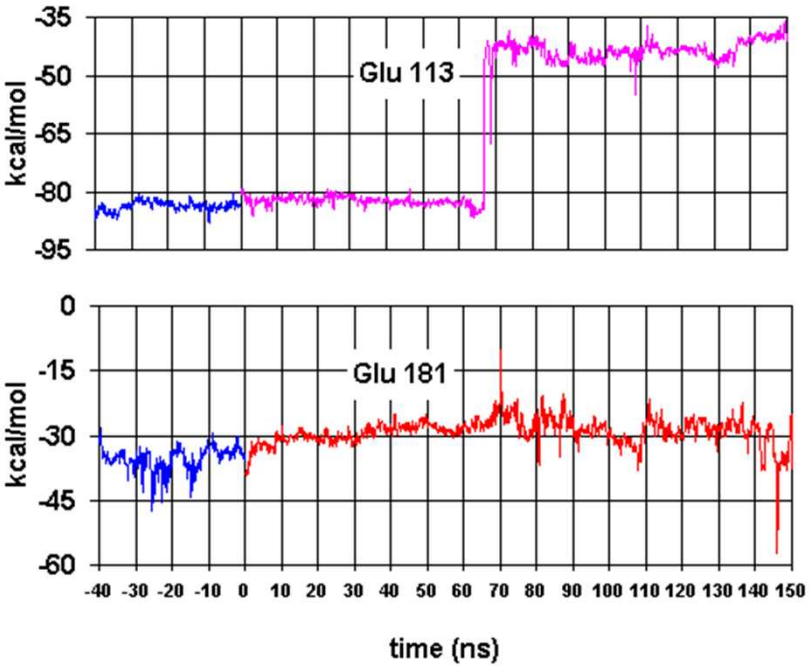


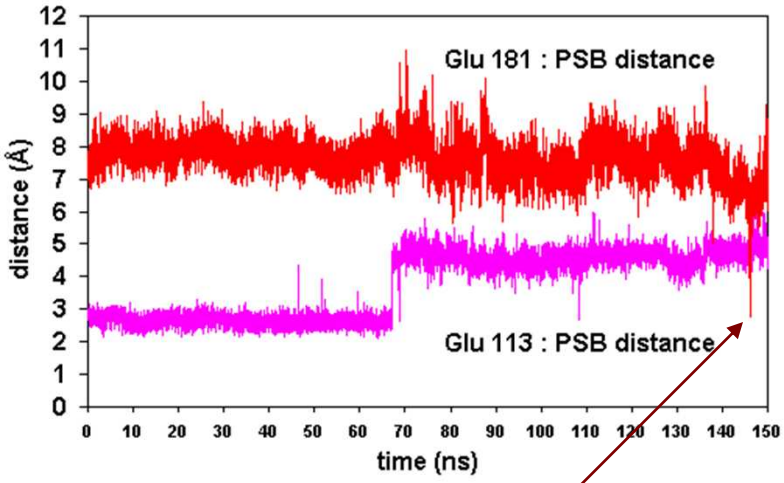
Most of the shift is caused by breaking of the salt bridge between Glu 113 and the PSB

Sandia National Laboratories

Switch of PSB counterion & weakening of interaction with retinal





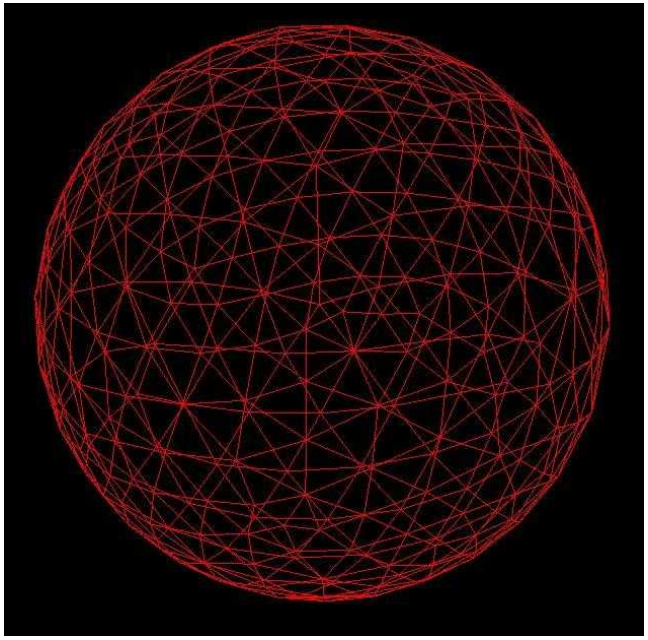


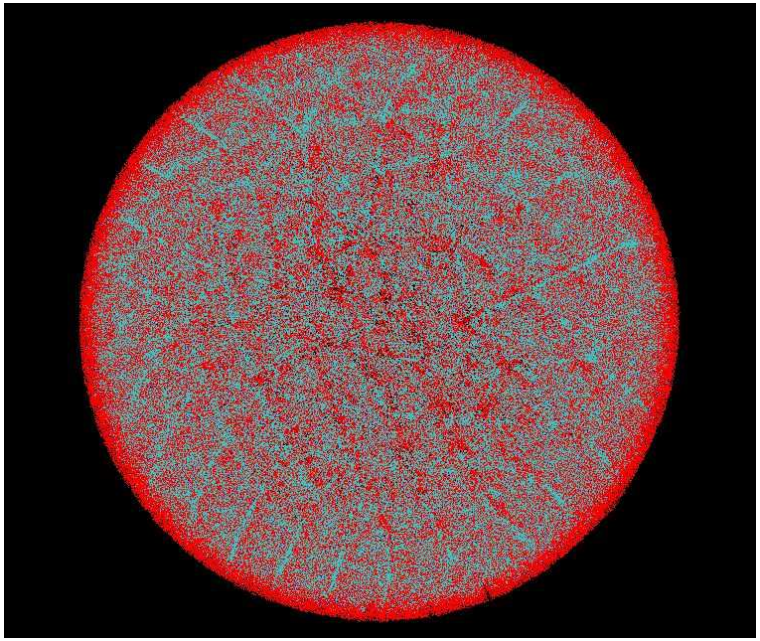
PSB interaction with its counterion weakens as system transitions towards Lumi intermediate
(Pan and Mathies, *Biochemistry* 40, 7929, (2001))

The min value of 2.8 Å in the Glu 181 time series corresponds with the max value of 6.3 Å in the Glu 113 time series (both at *t* = 146 ns)

Sandia National Laboratories

Whole vesicle simulation





- Enormous challenge due to sheer size of the system
  - 5 million atoms prior to filling box with water
  - Estimate > 100 million atoms total
- Sphere of tris built using Cubit software, then triangular patches of DOPC lipid bilayers were cut and placed on sphere surface.

Sandia National Laboratories

GRASP (General Reactive Atomistic Simulation Program): A Scalable Molecular Dynamics Code for Reactive Force Fields

Why Reactive Force Fields?

- Material behavior often dominated by chemical processes
- HE, Complex Solids, Polymer Aging
- Quantum methods limited to hundreds of atoms
- Ordinary classical force fields limited accuracy
- We need to have the best of both worlds $\Rightarrow$Reactive force fields

Why GRASP?

- Reactive force fields typically exist as custom serial MD codes
- GRASP is a general parallel MD code for reactive force fields

Sandia National Laboratories

Summary of GRASP Features

- Scalable parallel MD
- Calculates cumulative forces due to arbitrary set of force fields
- TwoBody, ThreeBody, External, EAM, Ewald Sum, PPPM, ReaxFF, Tersoff, Charge Equilibration
- Triclinic Multicell Periodic BC's
- Velocity Verlet, *NVE*, *NVT*
- **Pressure Ensembles (isotropic, anisotropic, fully flexible)**
- **NEMD Shear Deformation**
- Various Minimizers
- Parallel Charge Equilibration
- First Parallel implementation of ReaxFF (anywhere)
- Open Source License
- Not freely downloadable
  - **limited resources**
  - **ReaxFF available to Army researchers**
- CVS Repository: software.sandia.gov

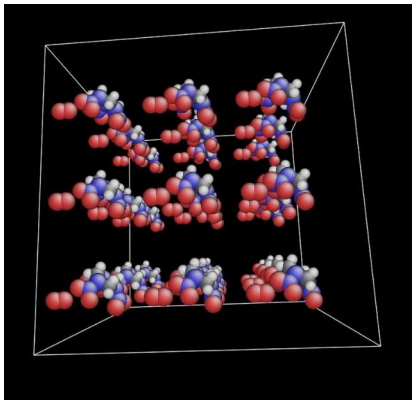
Sandia National Laboratories

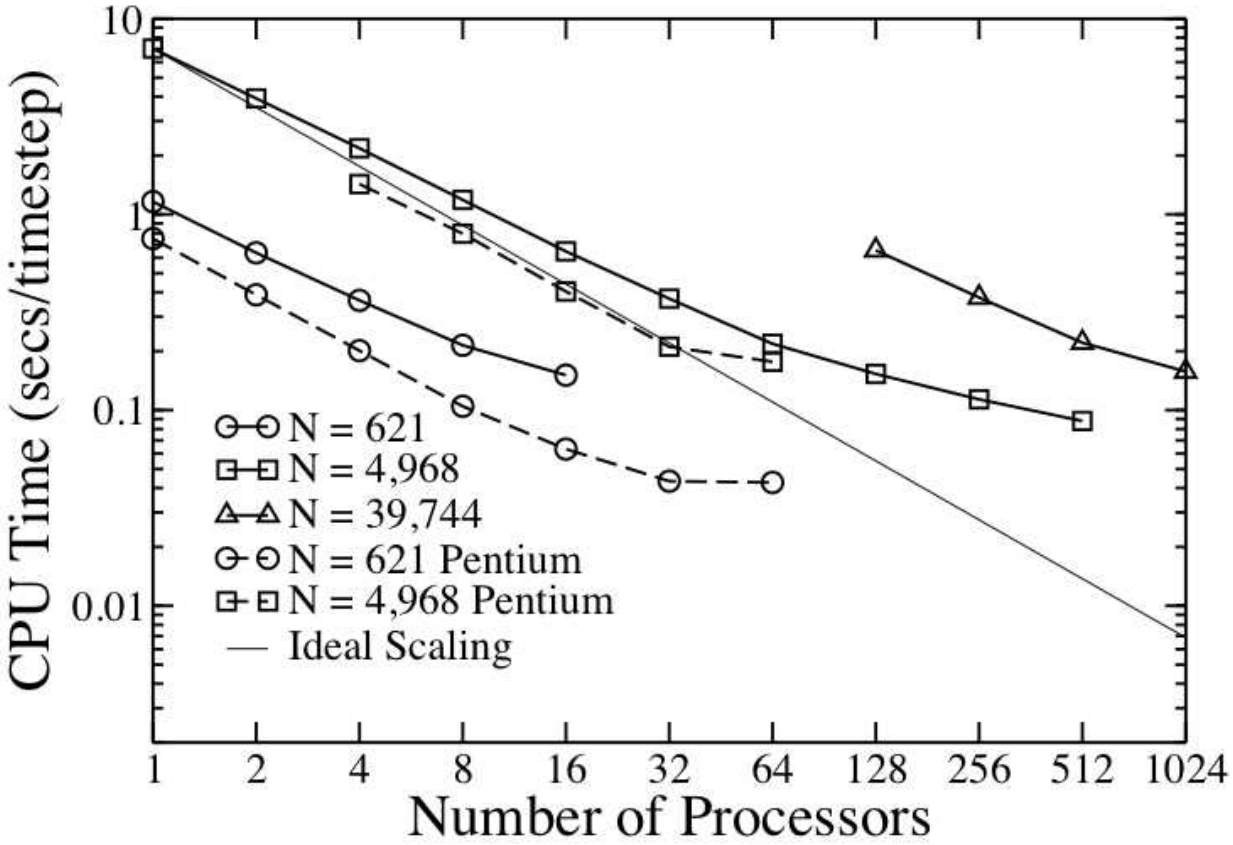
GRASP Performance on BG/L with ReaxFF

Comparison with 3GHz Pentium+Myrinet

RDX Explosive with Oxygen

ReaxFF force field with charge equilibration





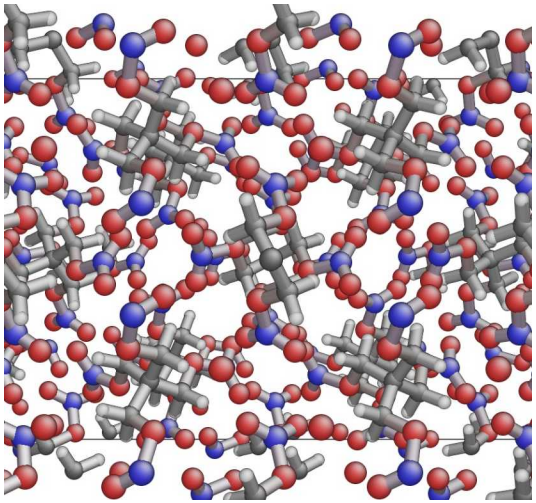
- First parallel implementation
- Mixes C++ and FORTRAN
- Heavy memory use
- Speed on BG/L close to 3 GHz Pentium
- Good weak scaling
- Limited strong scaling

$$O(N/P)\ (1 + O((N/P)^{-1/d}))$$

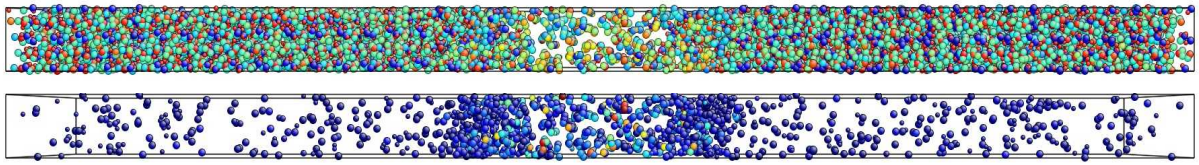
Sandia National Laboratories

GRASP+ReaxFF enables direct simulation of detailed initial energy propagation in HE

- Improved understanding of sensitivity will aid development of more reliable microenergetic components
- Goal: Identify the specific atomistic processes that cause orientation-dependent detonation sensitivity in PETN
- Thermal excitation simulations used as proof-of-concept
- Collaborating with parallel DoD-funded effort at Caltech (Bill Goddard, Sergey Zybin)
- Now running multi-million atom shock-initiated simulations with different orientations
- Contracted Grant Smith to extend his HMX/RDX non-reactive force field to PETN

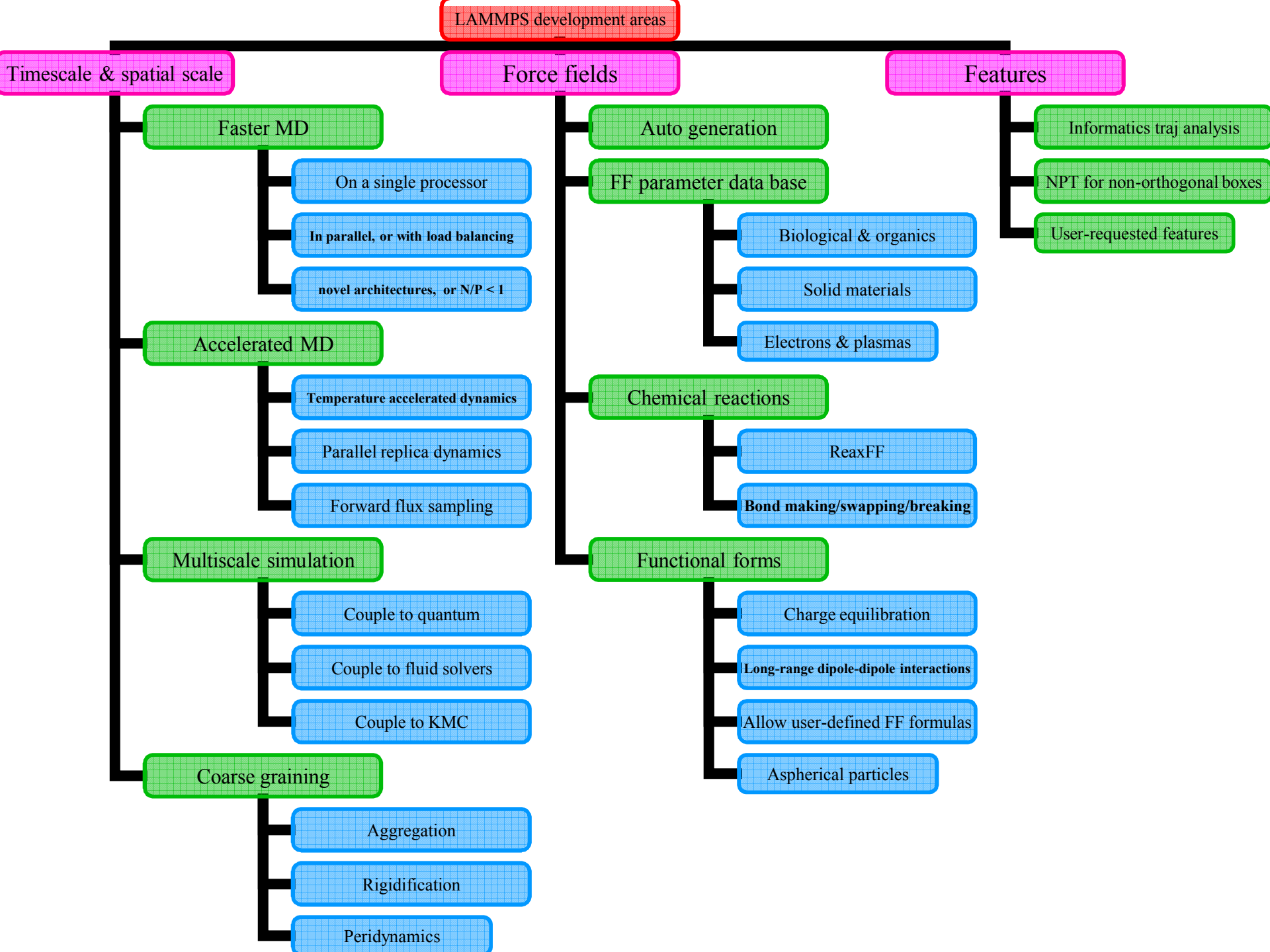


Complex molecular structure of unreacted tetragonal PETN crystal, C (gray), N (blue), O (red), and H (white).



Propagation of reaction front due to thermal excitation of a thin layer at the center of the sample for 10 picoseconds. Top: atoms colored by potential energy. Bottom: atoms colored by temperature (atoms below 1000K are not shown).

Sandia National Laboratories


LAMMPS development areas
Timescale & spatial scale
Faster MD
On a single processor
In parallel, or with load balancing
novel architectures, or N/P < 1
Accelerated MD
Temperature accelerated dynamics
Parallel replica dynamics
Forward flux sampling
Multiscale simulation
Couple to quantum
Couple to fluid solvers
Couple to KMC
Coarse graining
Aggregation
Rigidification
Peridynamics
Force fields
Auto generation
FF parameter data base
Biological & organics
Solid materials
Electrons & plasmas
Chemical reactions
ReaxFF
Bond making/swapping/breaking
Functional forms
Charge equilibration
Long-range dipole-dipole interactions
Allow user-defined FF formulas
Aspherical particles
Features
Informatics traj analysis
NPT for non-orthogonal boxes
User-requested features

A brief survey of the LAMMPS MD code: intro, case studies, and future development

Paul Crozier, Steve Plimpton, Aidan Thompson

February 6, 2008

ERDC

Vicksburg, MS

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

Sandia National Laboratories