An Introduction to Molecular Visualization with VMD

John Stone

Theoretical and Computational Biophysics Group

University of Illinois at Urbana-Champaign

http://www.ks.uiuc.edu/Research/vmd/

Centre for High Performance Computing,

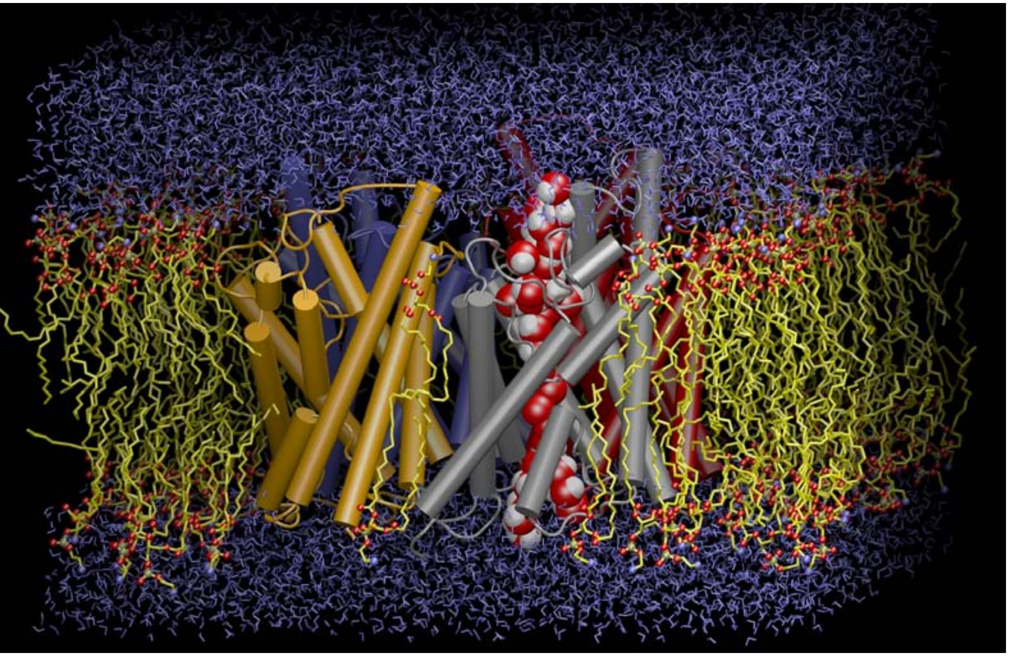
CSIR Rosebank Campus, Cape Town, South Africa, October 27, 2008

Overview

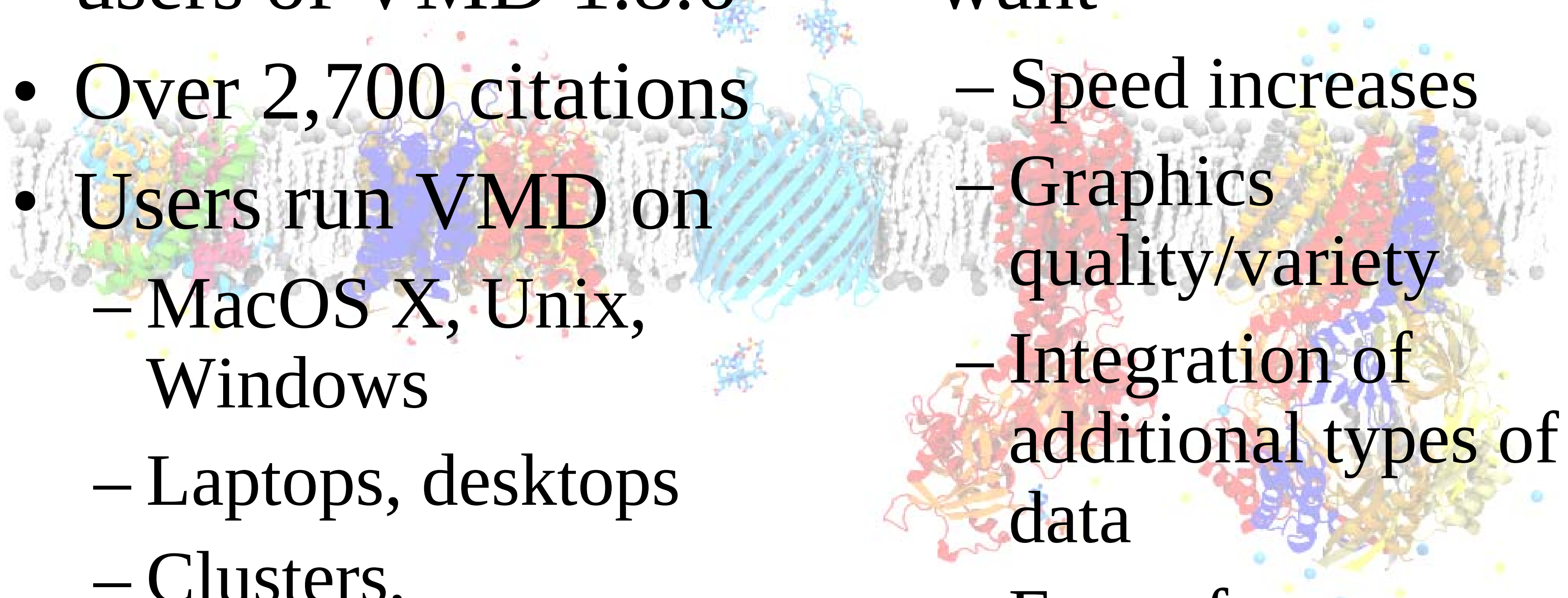
- Will be showing a lot of VMD images, feel free to ask questions
- General visualization concepts and methods
- Specific VMD visualization examples for molecular dynamics trajectories, density maps, etc.
- Overview of VMD scripting, extensibility

VMD

- VMD – “Visual Molecular Dynamics”
- Visualization of molecular dynamics simulations, sequence data, volumetric data, quantum chemistry data, particle systems
- User extensible with scripting and plugins
- http://www.ks.uiuc.edu/Research/vmd/

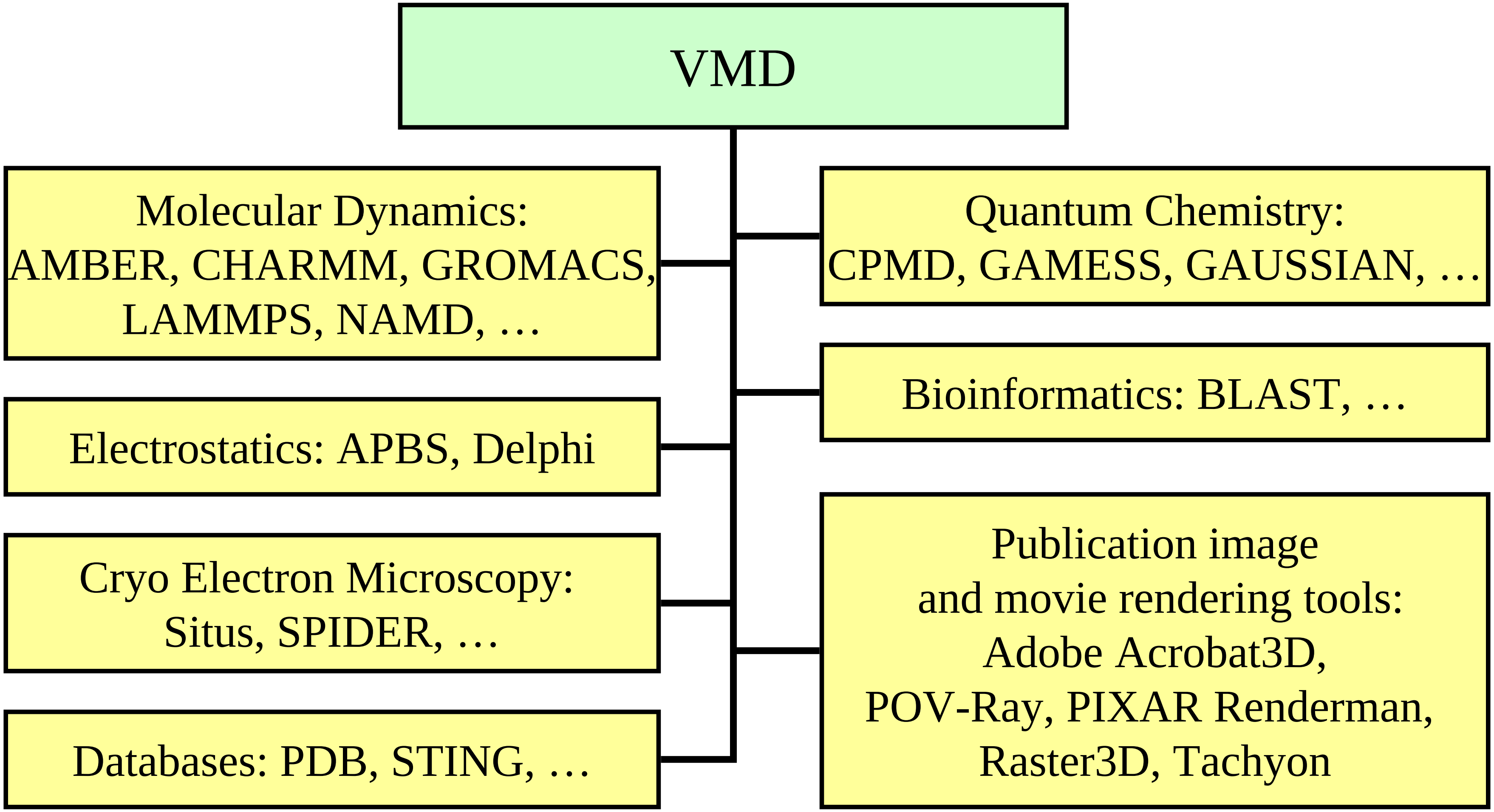


Meeting the Diverse Needs of the VMD User Community



- Over 34,000 registered users of VMD 1.8.6
- Over 2,700 citations
- Users run VMD on
  - MacOS X, Unix, Windows
  - Laptops, desktops
  - Clusters, supercomputers
- Surveyed users want
  - Speed increases
  - Graphics quality/variety
  - Integration of additional types of data
  - Ease of use

VMD's Broad Interoperability With Key Research Tools and Data



VMD

Molecular Dynamics:
AMBER, CHARMM, GROMACS,
LAMMPS, NAMD, …

Electrostatics: APBS, Delphi

Cryo Electron Microscopy:
Situs, SPIDER, …

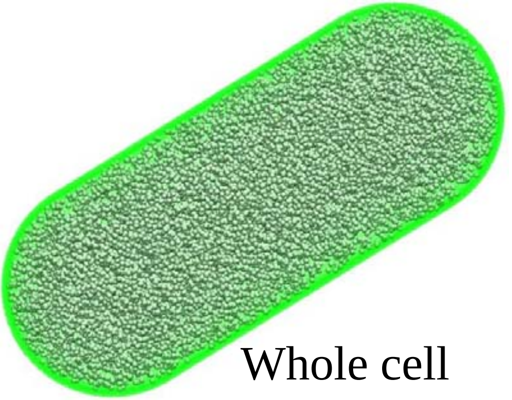
Databases: PDB, STING, …

Quantum Chemistry:
CPMD, GAMESS, GAUSSIAN, …

Bioinformatics: BLAST, …

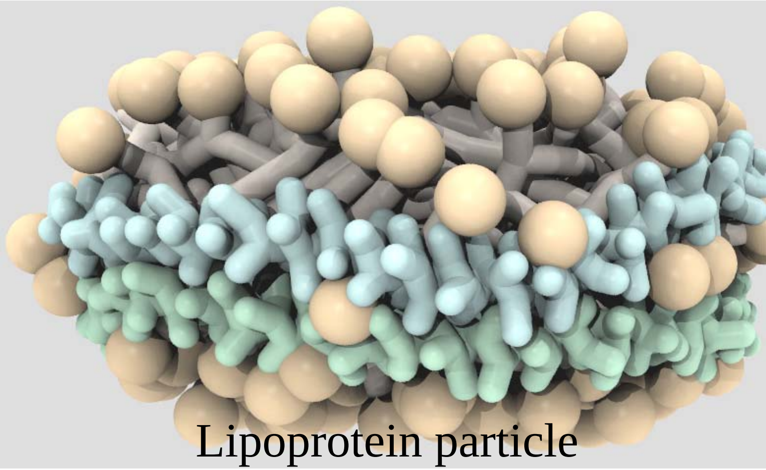
Publication image
and movie rendering tools:
Adobe Acrobat3D,
POV-Ray, PIXAR Renderman,
Raster3D, Tachyon

VMD Advanced Data Handling



Whole cell

Atomic, CG, Particle:
Coordinates, Trajectories,
Energies, Forces,
Secondary Structure, …



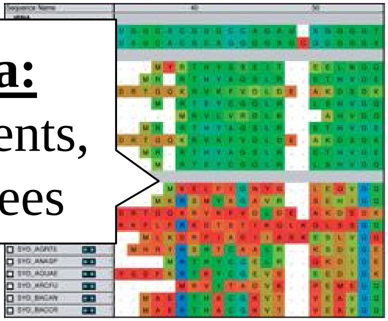
Lipoprotein particle

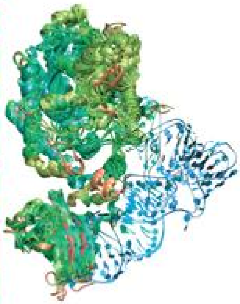
Graphics, Geometry

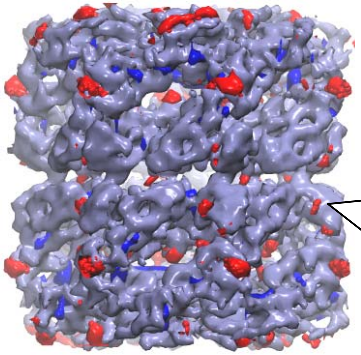
Annotations

VMD

Sequence Data:
Multiple Alignments,
Phylogenetic Trees

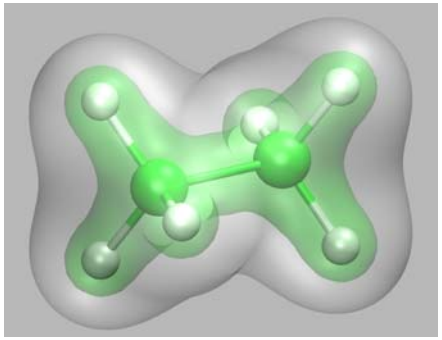






GroEL

Volumetric Data:
Density maps,
Electron orbitals,
Electrostatic potential, …



Ethane

Attributes of the Data We’re Interested in Visualizing

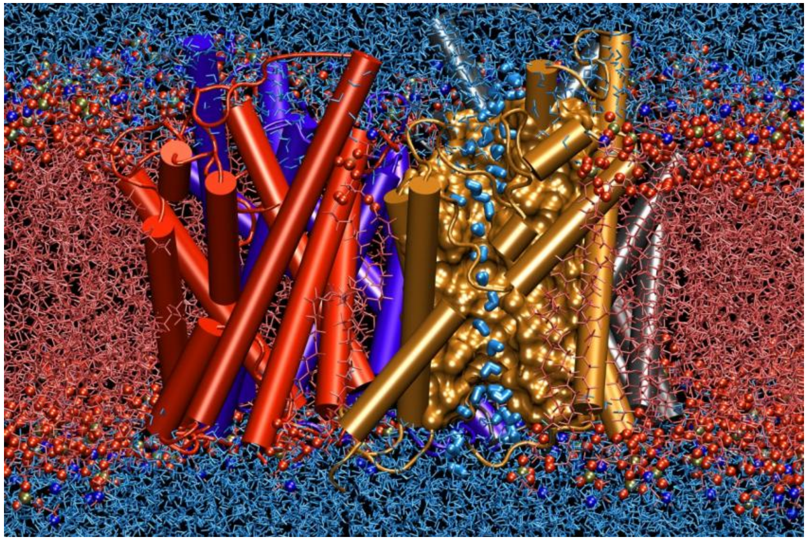
- Multiple types of data
  - Atomic structures, Sequences, Volumetric data, …
- Many attributes per-atom
- Millions of atoms, particles, voxels
- Time varying (simulation trajectories)
- Multiple structures

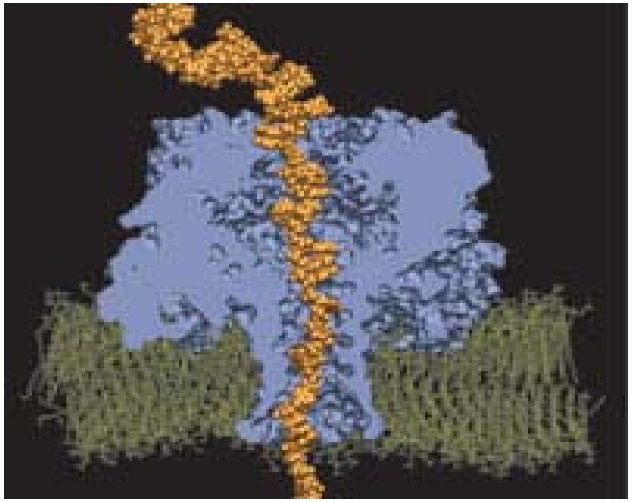
Methods for Visualizing Molecular Data in VMD

- Direct display of atomic structure
- Schematic representations
- Map data to color, texture
- Display data topologically, as graphs, plots
- Combine multiple graphical representations and display modalities together simultaneously

Visualizing Data with Shape

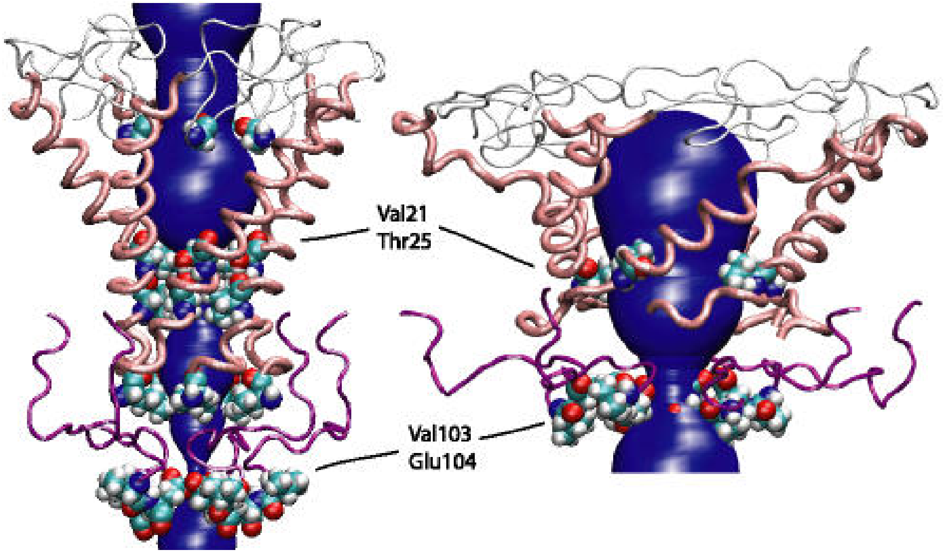
- Direct rendering of geometry from physical data (e.g. atomic structures)
- Indirect rendering of data, feature extraction (e.g. density isosurfaces)
- Reduced detail representations of data (e.g. ribbons, cartoon)
- Use size for emphasis

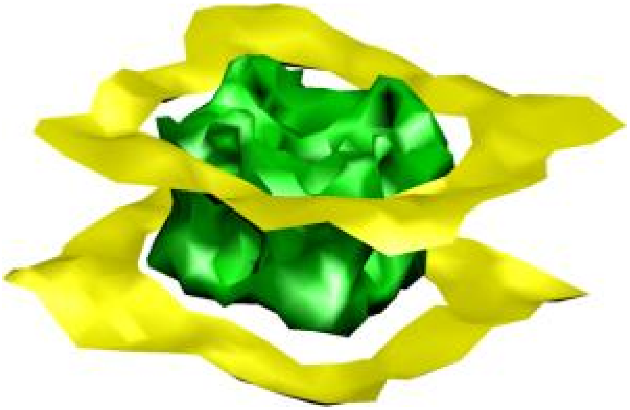


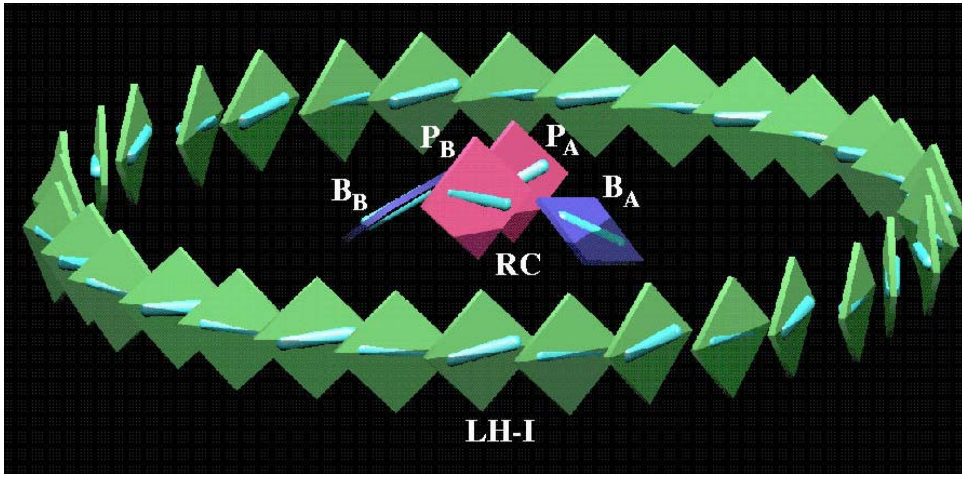


Schematic Representations

- Extract and render pores, cavities, indentations
- Simplified representations of large structural features



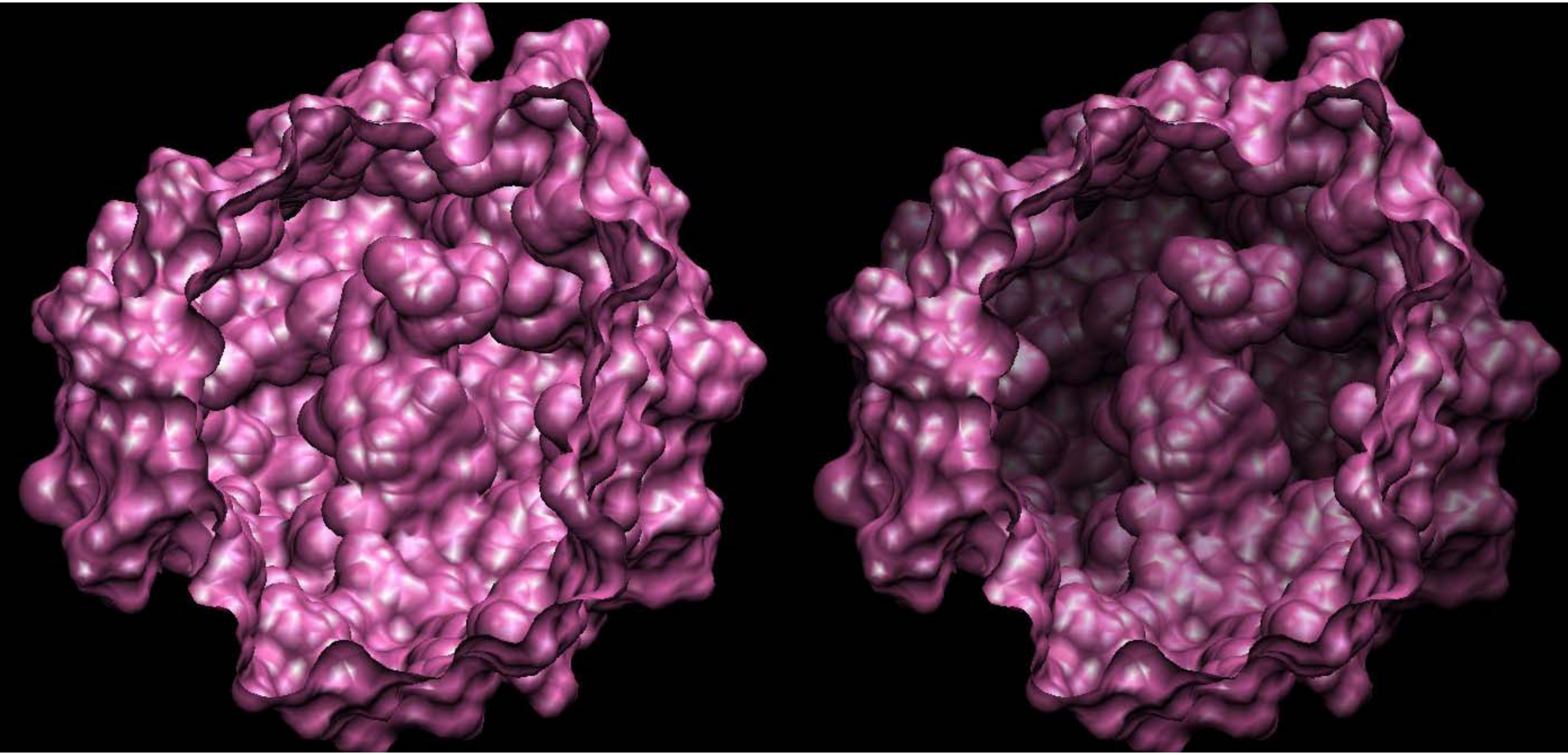


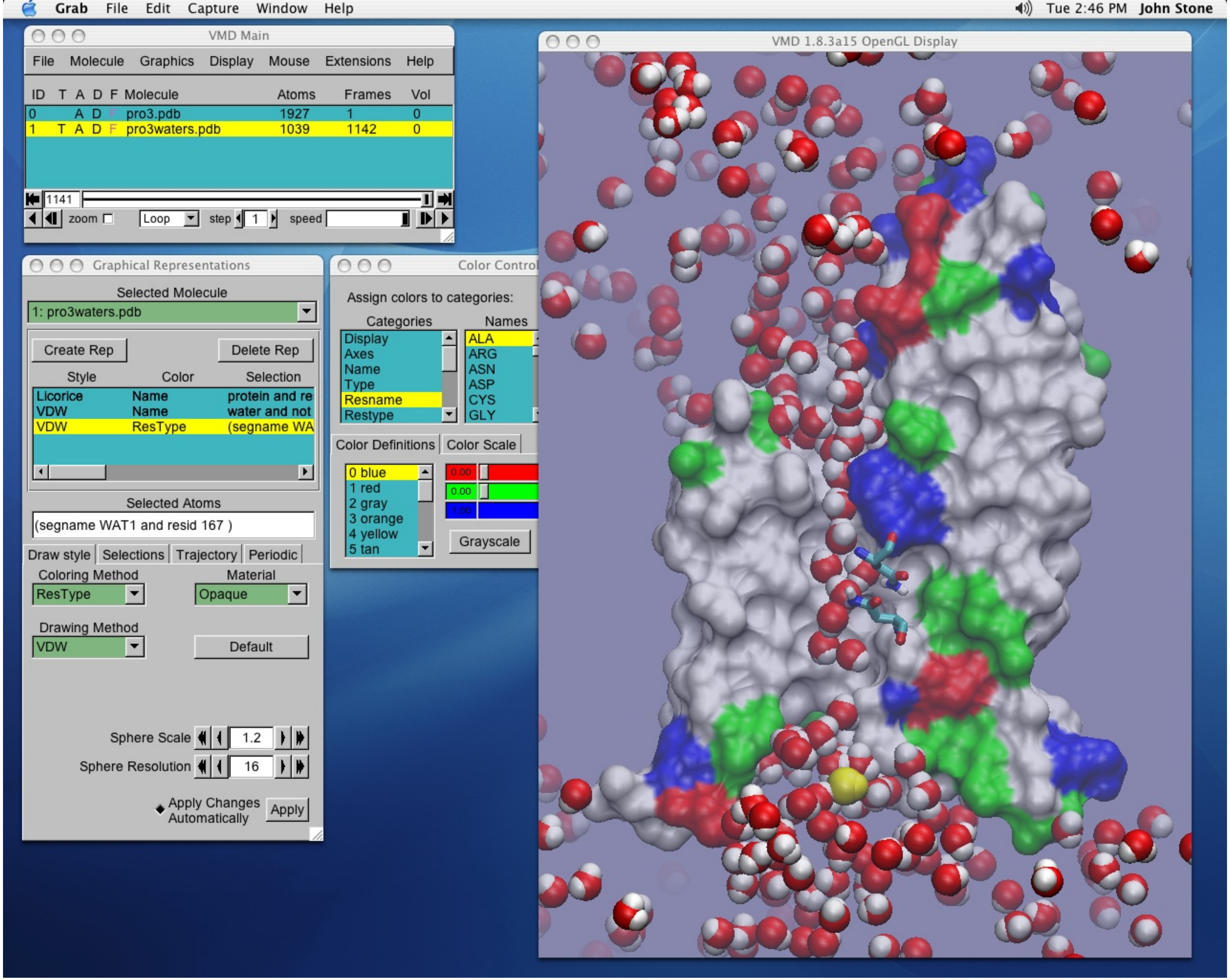


Visualizing Data with Texture and Color

- Direct mapping of properties/values to colors (e.g. color by electrostatic potential)
- Indirect mapping via feature extraction (e.g. color by secondary structure)
- Use saturated colors to draw attention
- Use faded colors and transparency to de-emphasize
- Use depth cueing/fog to de-emphasize background environment

Depth Cueing





ID	T	A	D	F	Molecule	Atoms	Frames	Vol
0		A	D	F	pro3.pdb	1927	1	0
1	T	A	D	F	pro3waters.pdb	1039	1142	0

1141

Loop

step 1

Selected Molecule

1: pro3waters.pdb

Style	Color	Selection
Licorice	Name	protein and re
VDW	Name	water and not
VDW	ResType	(segname WA

Selected Atoms

(segname WAT1 and resid 167)

Coloring Method: ResType

Material: Opaque

Drawing Method: VDW

Sphere Scale 1.2

Sphere Resolution 16

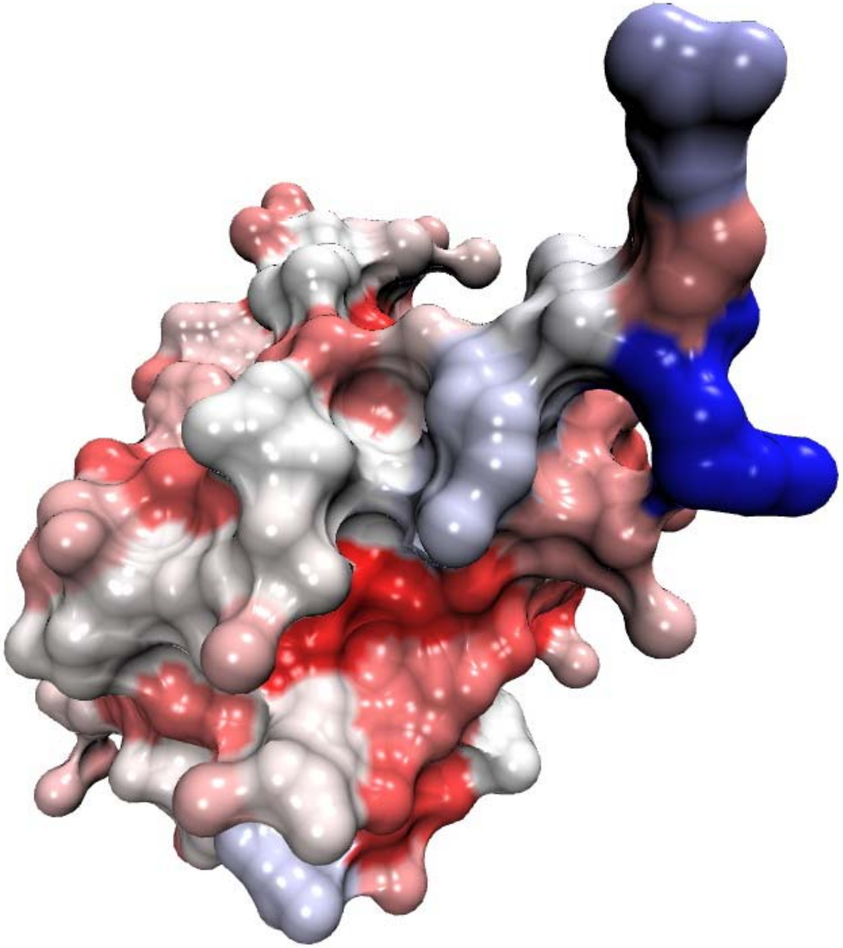
◆ Apply Changes Automatically

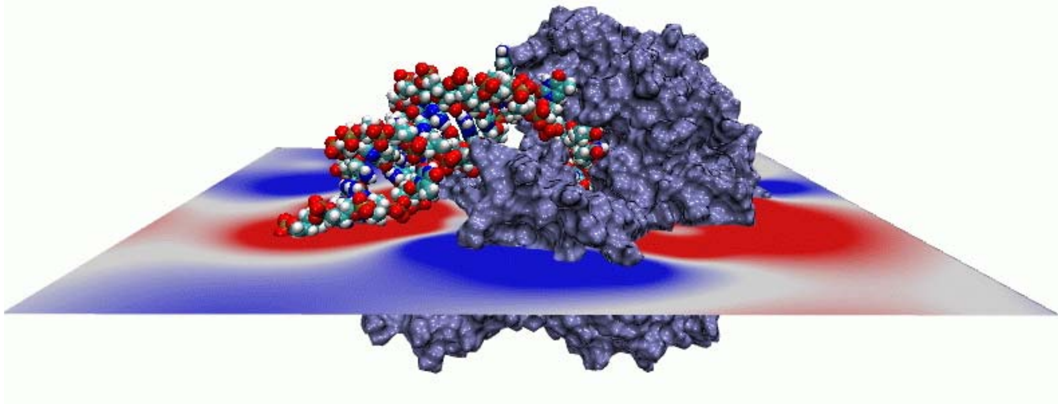
Assign colors to categories:

Categories	Names
Display	ALA
Axes	ARG
Name	ASN
Type	ASP
Resname	CYS
Restype	GLY

0 blue
1 red
2 gray
3 orange
4 yellow
5 tan

0.00
0.00
1.00



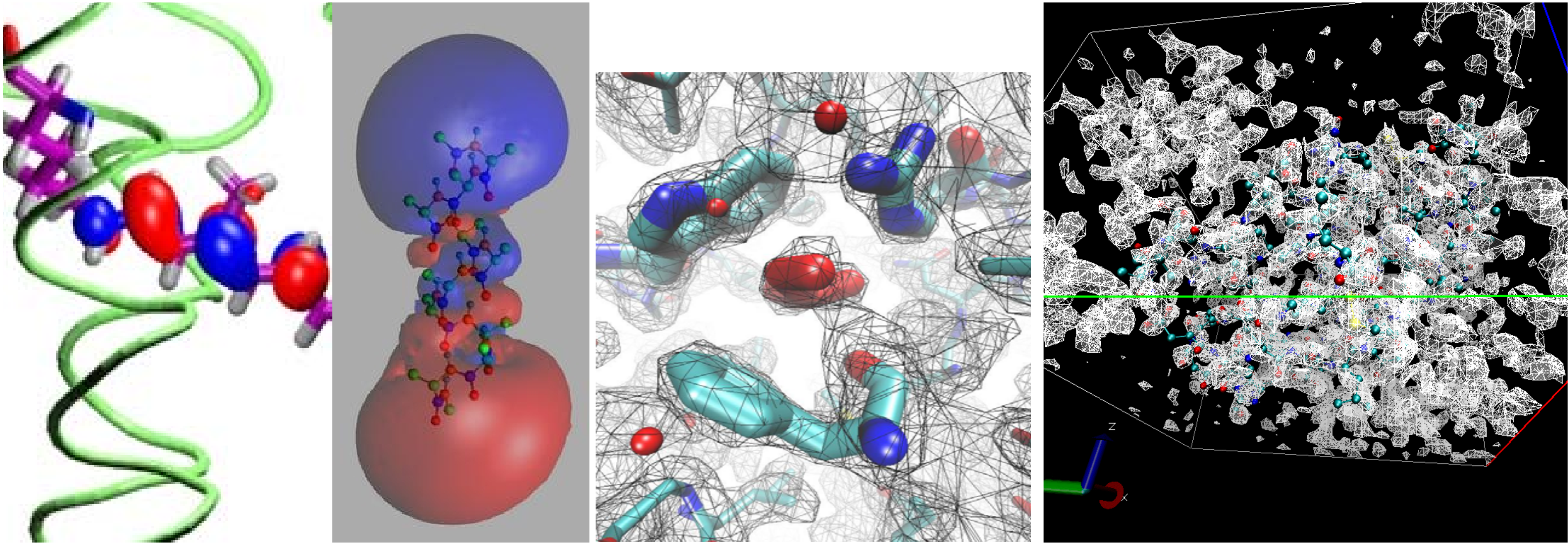


Electrostatic potential for a helicase
obtained with VMD’s PME plugin

Per-residue Solvent-Accessible
Surface Area of Ubiquitin

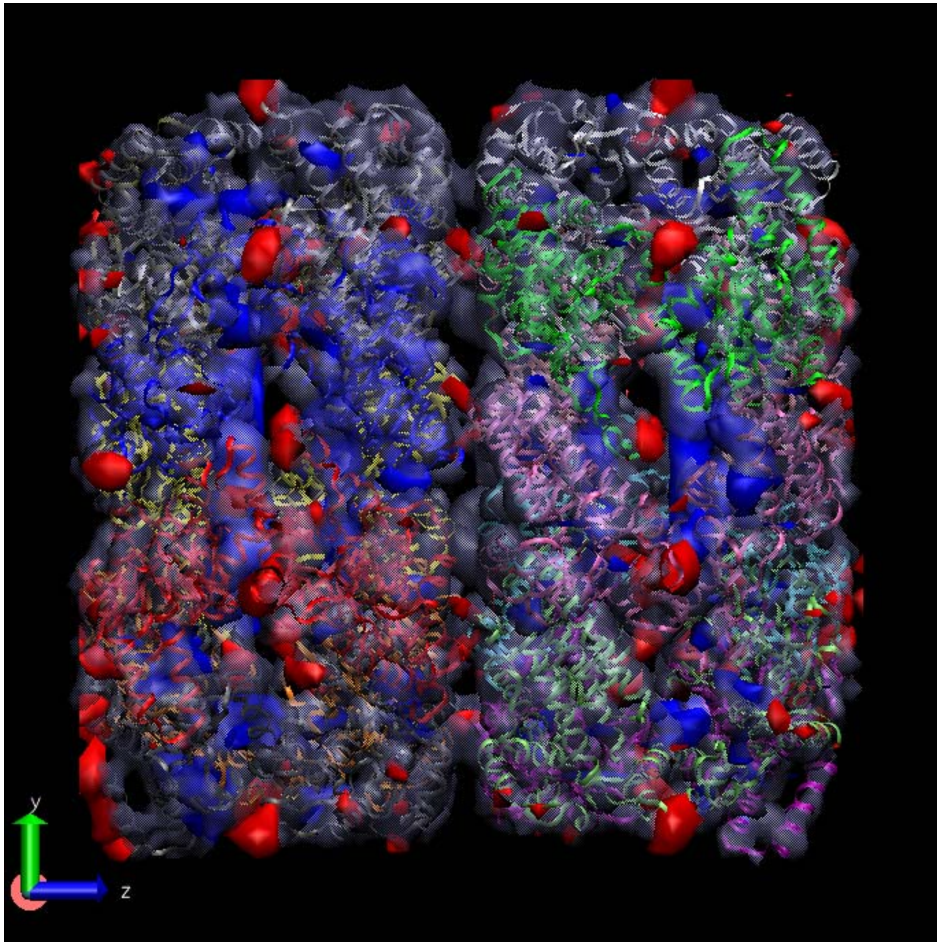
Visualizing Volumetric Data

- Display environment surrounding molecular structure, fields that affect structure and function
- Electron orbitals, electron density, electrostatic potential, temporal occupancy maps



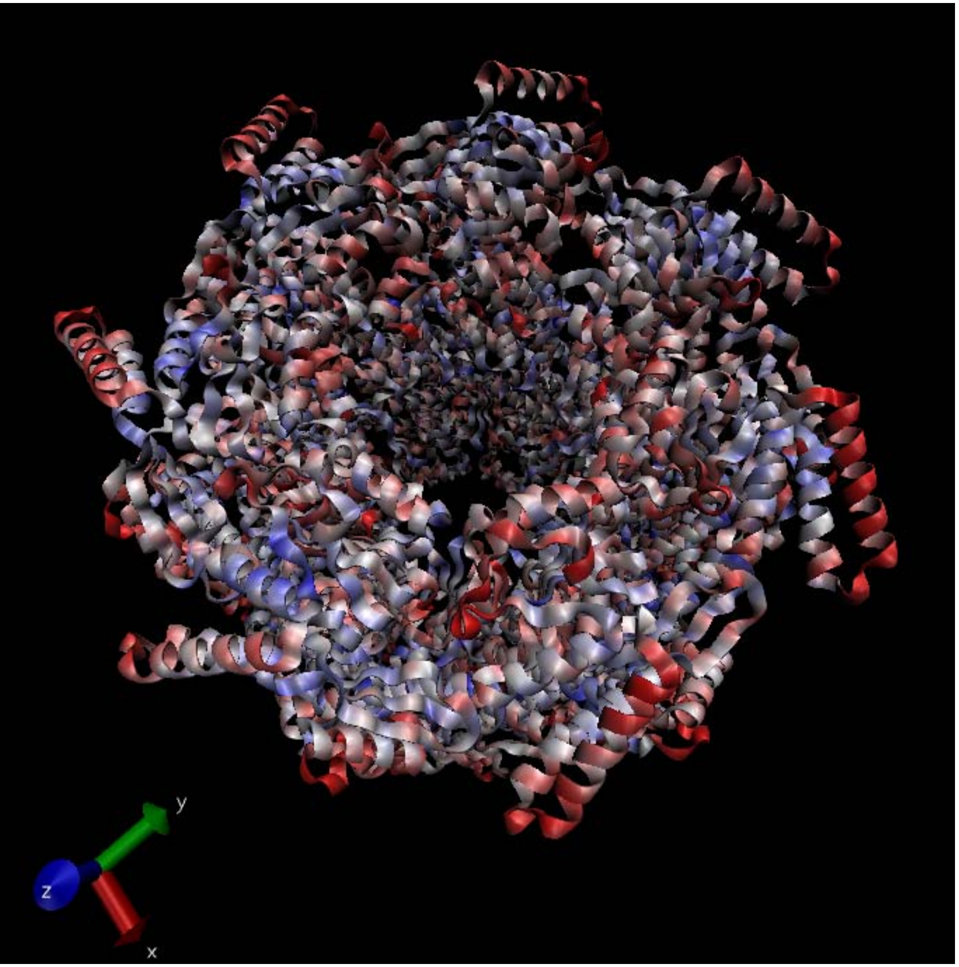
GroEL: Docked Map and Structure

- SITUS:
  - Dock map+structure
  - Synthesize map from PDB
  - Calculate difference between EM map and PDB
- VMD:
  - Load density maps
  - Display isosurfaces
  - Display map/structure alignment error as isosurfaces
  - Texture reps by density or map/structure alignment error

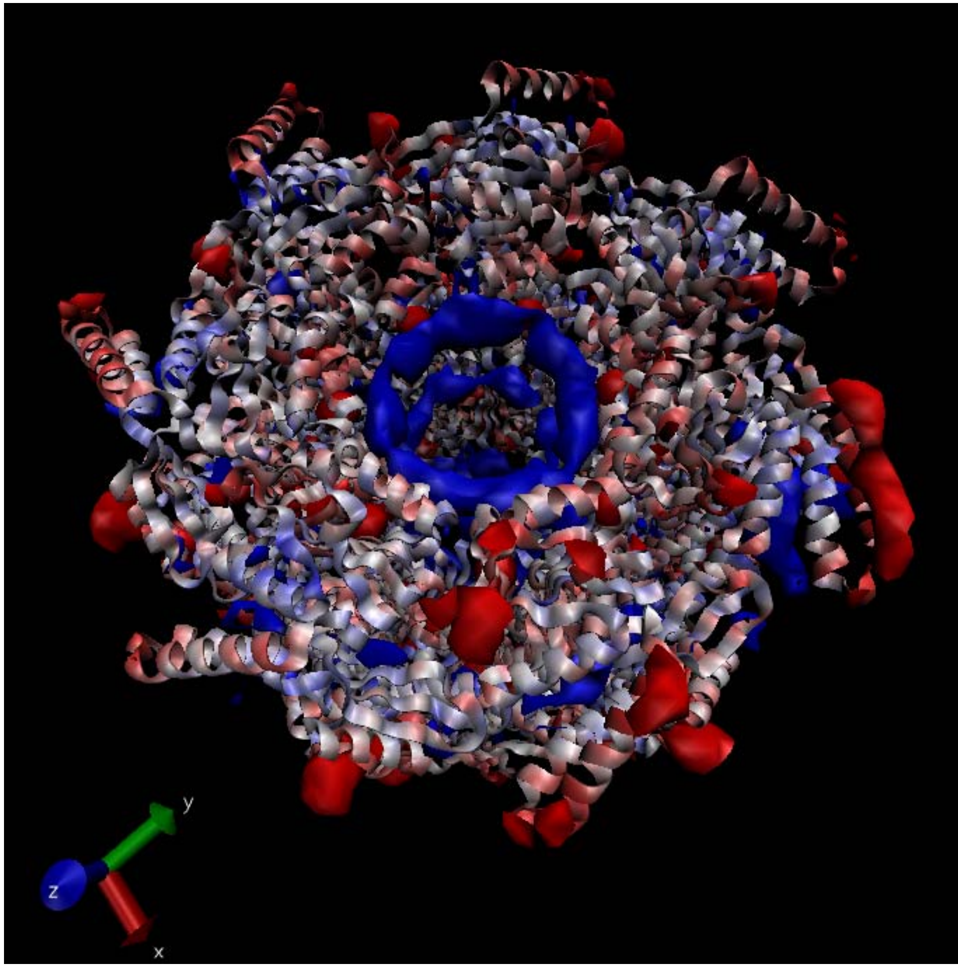


GroEL: Display of Difference, Error

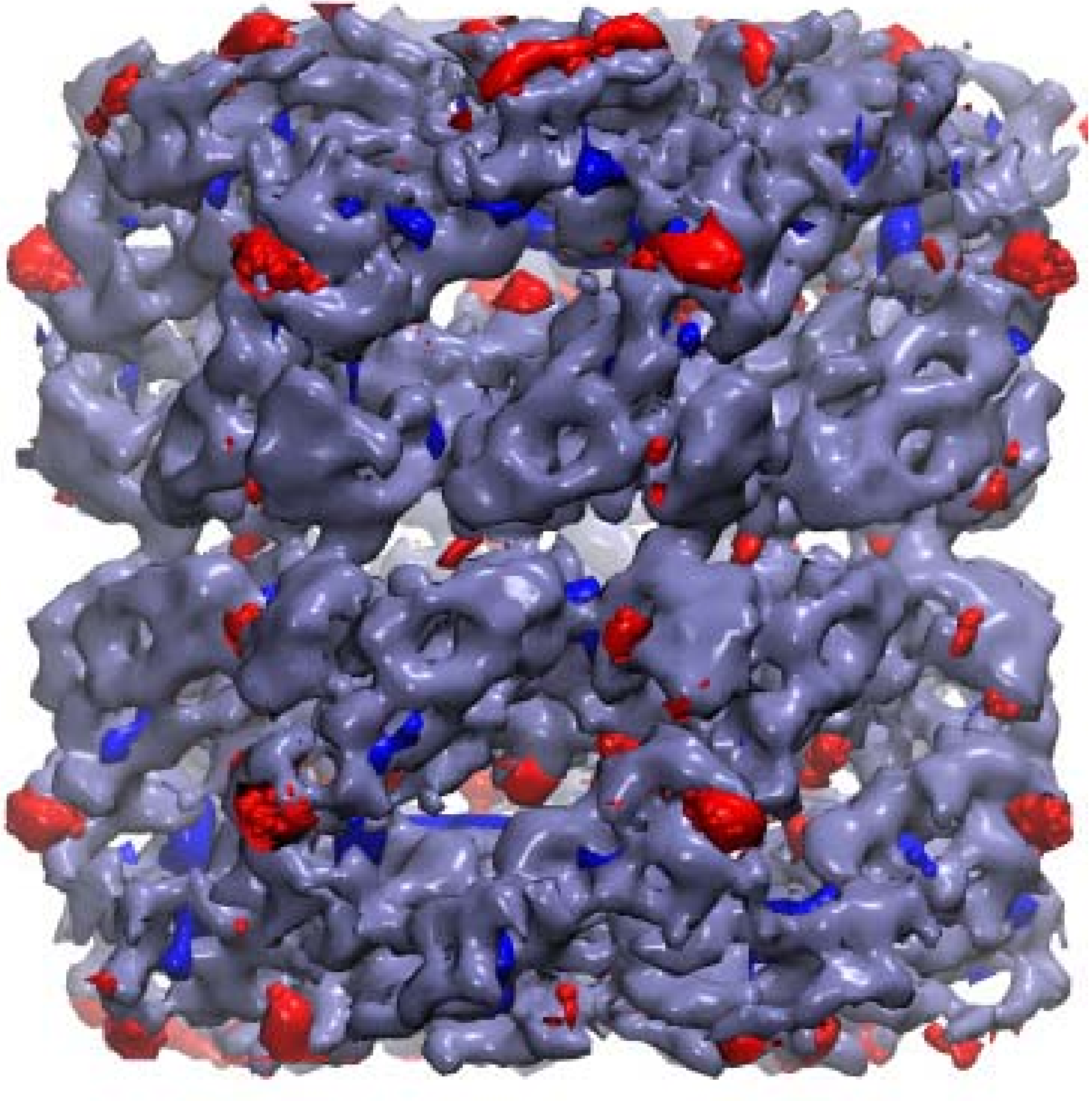
Ribbons textured by difference map



…with difference isosurfaces



GroEL: Difference Isosurfaces

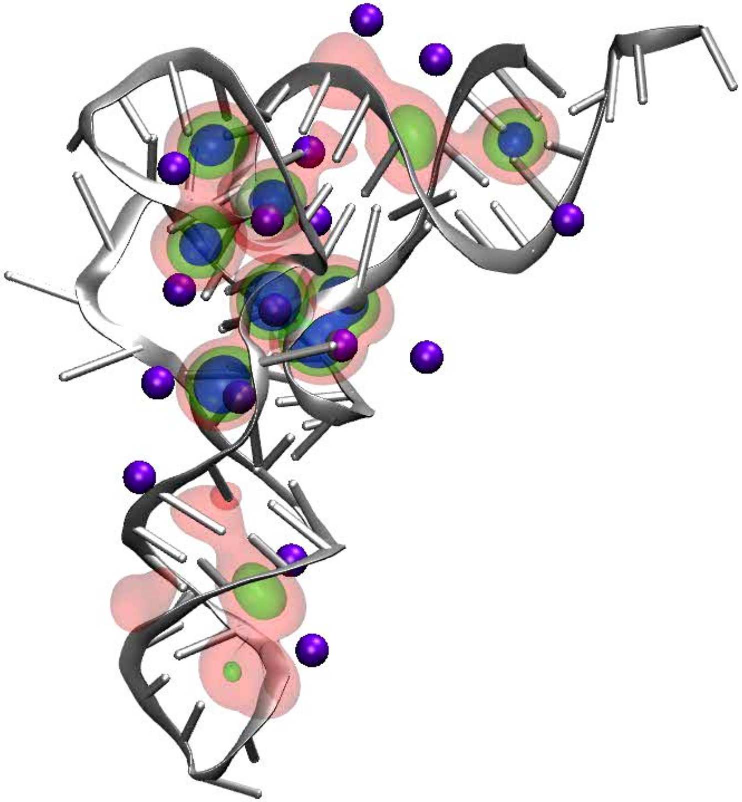


National Center for Research Resources

Beckman Institute, UIUC

Computing Volumetric Properties

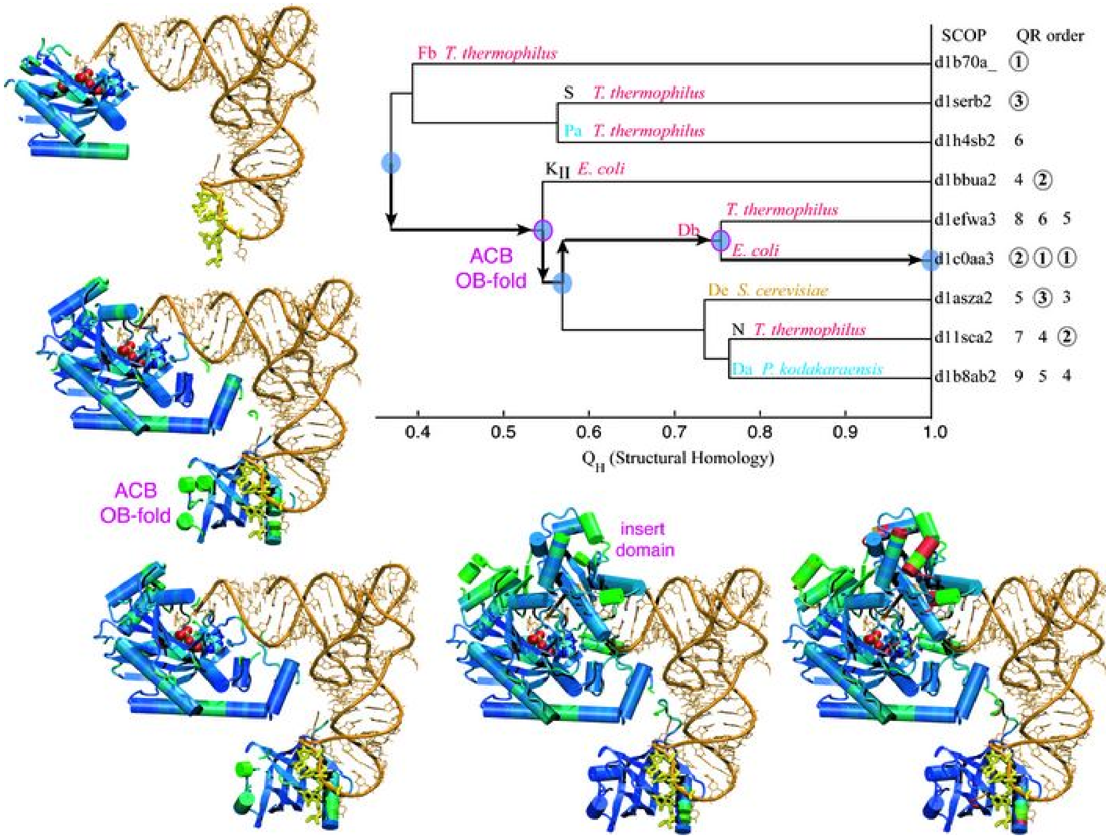
- Compute density, distance, occupancy, potential maps for a frame or averaged over a trajectory
- Example: display binding sites for diffusively bound ions as probability density isosurfaces



tRNA magnesium ion occupancy

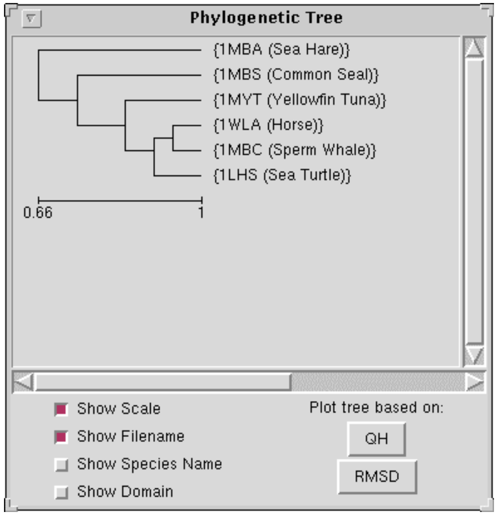
Multiple Structure Alignment

- Study evolutionary changes in sequence and structure of proteins
- Align and superimpose multiple structures
- Color by structural conservation
- Color by sequence conservation
- Display phylogenetic tree, cluster biological form by similarity



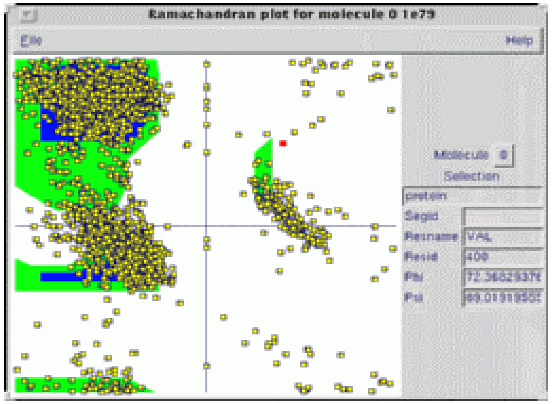
Visualizing Data Topologically

- Data relationships indicated by grouping (e.g. phylogenetic trees)
- Abstract or schematic representation, e.g. Ramachandran plot



☑ Show Scale
☑ Show Filename
☐ Show Species Name
☐ Show Domain

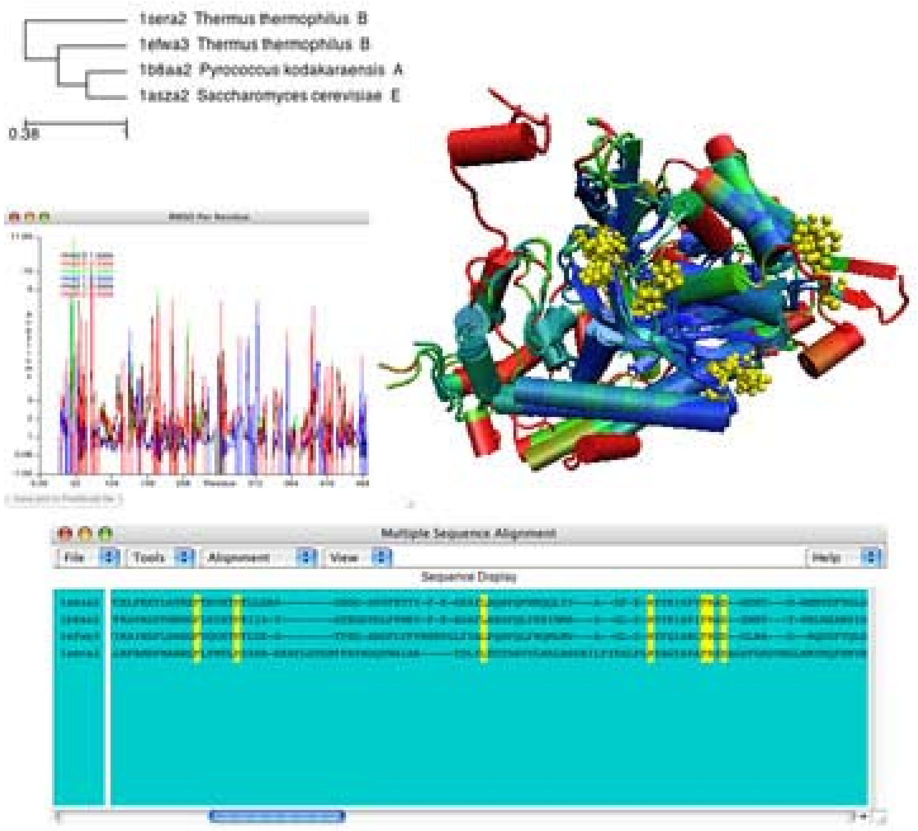
Plot tree based on:

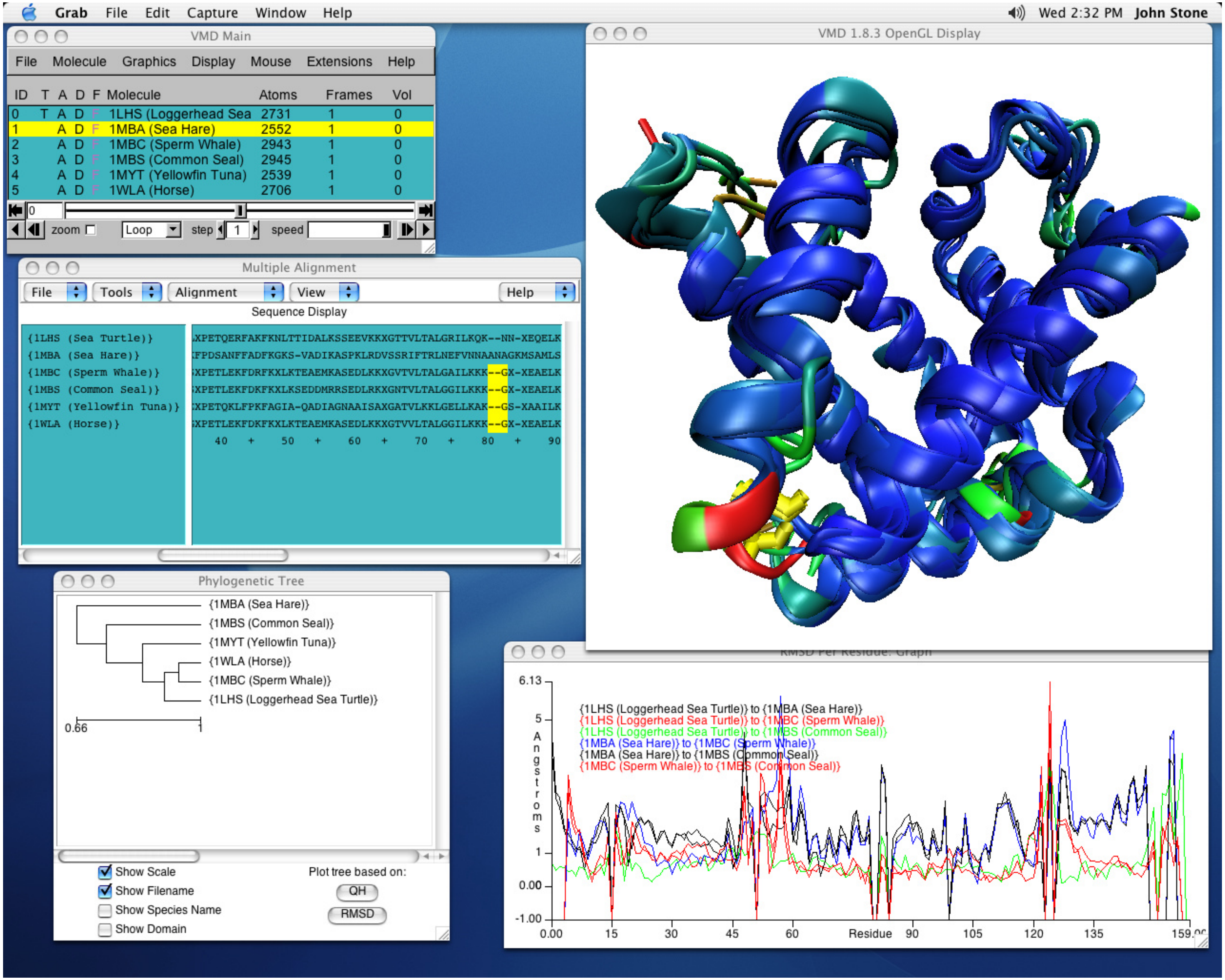


Molecule 0
Selection
protein
Segid
Resname VAL
Resid 408
Phi 72.3602937
Psi 89.0191955

Bringing it all together…

- Aligned sequences and structures, phylogeny
- Simultaneous use of shape, color, topology, and interactivity
- Multiple simultaneous representations
- Multiple data display modalities
- Selections in one modality can be used to highlight or select in others





ID	T	A	D	F	Molecule	Atoms	Frames	Vol
0	T	A	D	F	1LHS (Loggerhead Sea	2731	1	0
1		A	D	F	1MBA (Sea Hare)	2552	1	0
2		A	D	F	1MBC (Sperm Whale)	2943	1	0
3		A	D	F	1MBS (Common Seal)	2945	1	0
4		A	D	F	1MYT (Yellowfin Tuna)	2539	1	0
5		A	D	F	1WLA (Horse)	2706	1	0

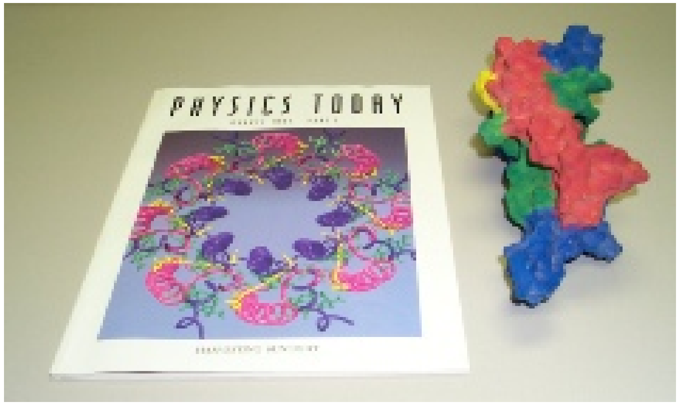
Sequence Display

```
{1LHS (Sea Turtle)}       XPETQERFAKFKNLTTIDALKSSEEVKKXGTTVLTALGRILKQK--NN-XEQELK
{1MBA (Sea Hare)}         FPDSANFFADFKGKS-VADIKASPKLRDVSSRIFTRLNEFVNNAANAGKMSAMLS
{1MBC (Sperm Whale)}      XPETLEKFDRFKXLKTEAEMKASEDLKKXGVTVLTALGAILKKK--GX-XEAELK
{1MBS (Common Seal)}      XPETLEKFDKFKXLKSEDDMRRSEDLRKXGNTVLTALGGILKKK--GX-XEAELK
{1MYT (Yellowfin Tuna)}   XPETQKLFPKFAGIA-QADIAGNAAISAXGATVLKKLGELLKAK--GS-XAAILK
{1WLA (Horse)}            XPETLEKFDKFKXLKTEAEMKASEDLKKXGTVVLTALGGILKKK--GX-XEAELK
                             40    +    50    +    60    +    70    +    80    +    90
```

☑ Show Scale
☑ Show Filename
☐ Show Species Name
☐ Show Domain

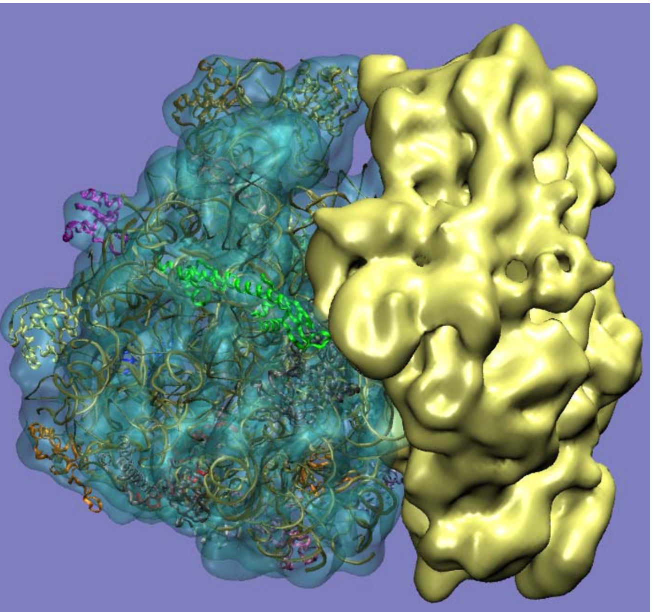
What else can we do?

- Enhance visual perception of shape
  - Motion, interactive rotation
  - Stereoscopic display
  - High quality surface shading and lighting
- Enhance tactile perception of shape
  - Print 3-D solid models
  - Interactive exploration using haptic feedback

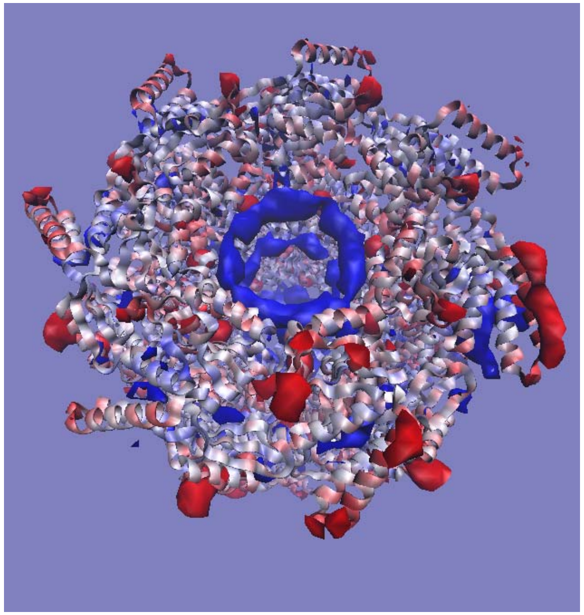


VMD Representation Examples

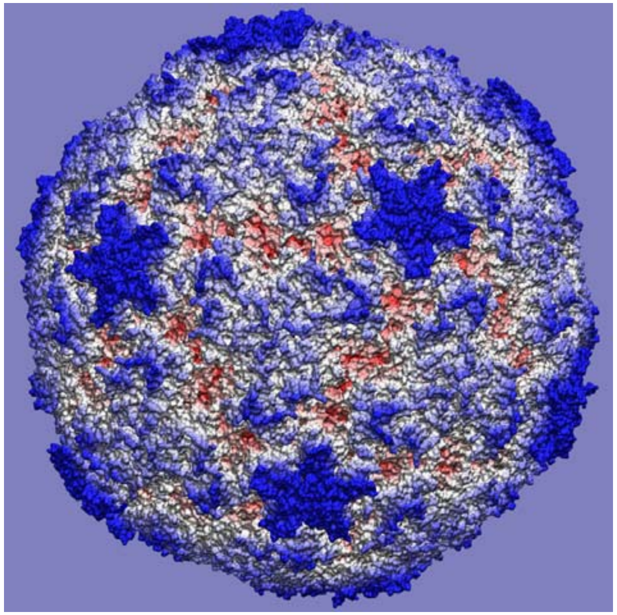
- Draw atomic structure, protein backbone, secondary structure, solvent-accessible surface, window-averaged trajectory positions, isosurfaces of volumetric data, much more…
- Color by per-atom or per-residue info, position, time, electrostatic potential, density, user-defined properties, etc…



Ribosome, J. Frank



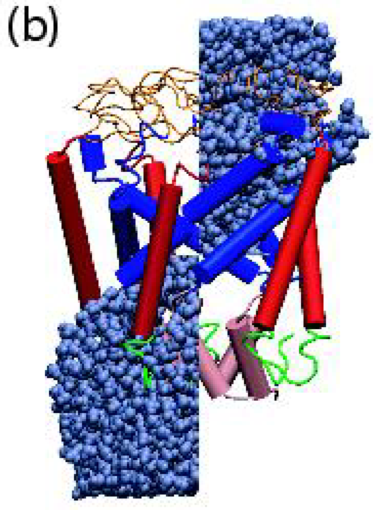
GroEL /w Situs

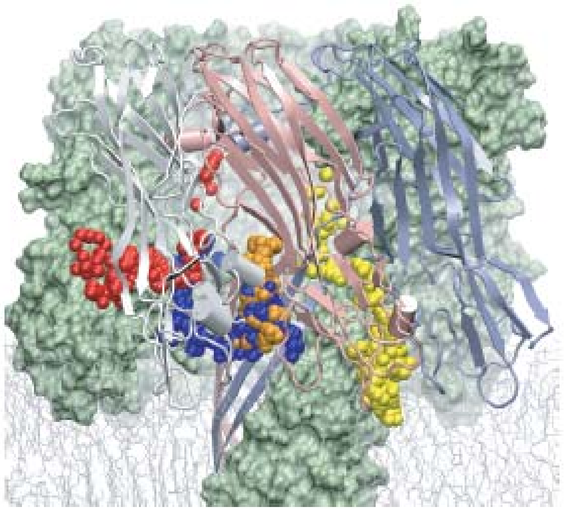


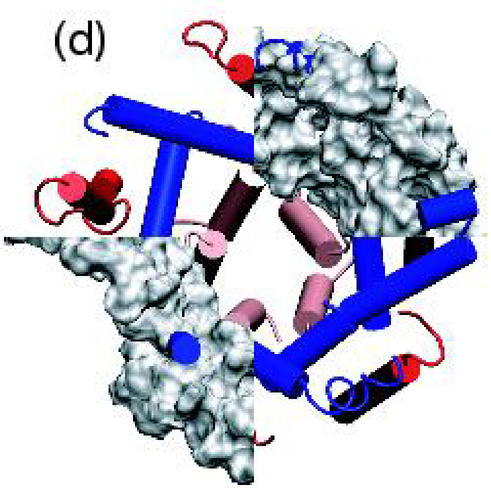
4HRV, 400K atoms

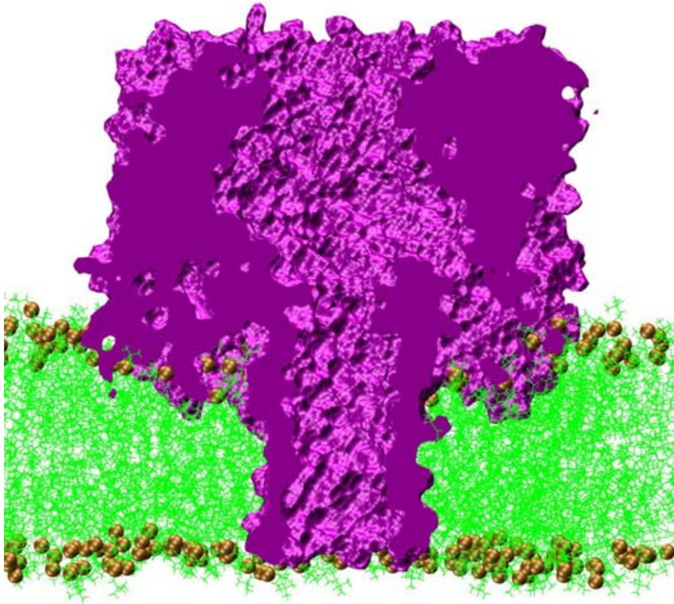
Multiple Representations, Cut-away Views

- Multiple reps are often used concurrently
  - Show selected regions in full atomic detail
  - Simplified cartoon-like or schematic form
- Clipping planes can slice away structure obscuring interesting features



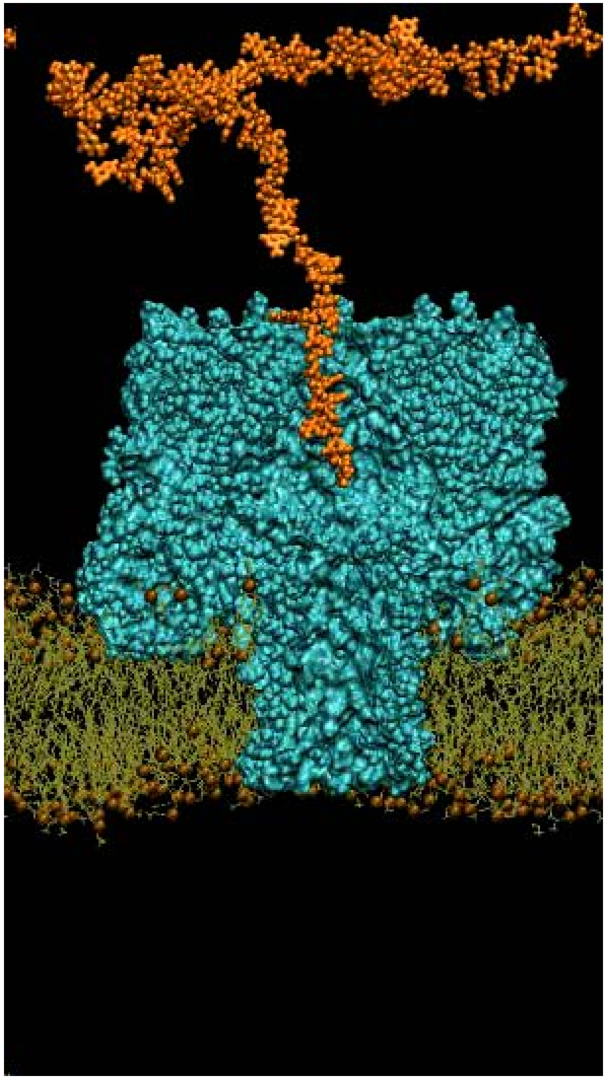






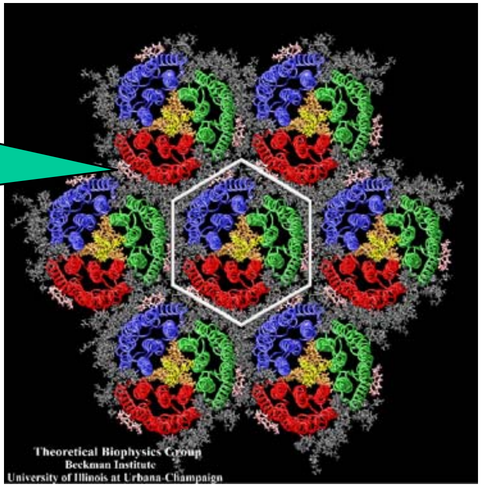
Visualization of Large All Atom Molecular Dynamics Simulations (1)

- All-atom models of proteins, membranes, DNA, in water solution
- 100K to 2M atoms
- 512 CPU jobs run on remote supercomputers for weeks at a time for a 10ns simulation
- Visualization and analysis require workstations with 4-32 GB of RAM, 1-4 CPUs, high-end graphics accelerators

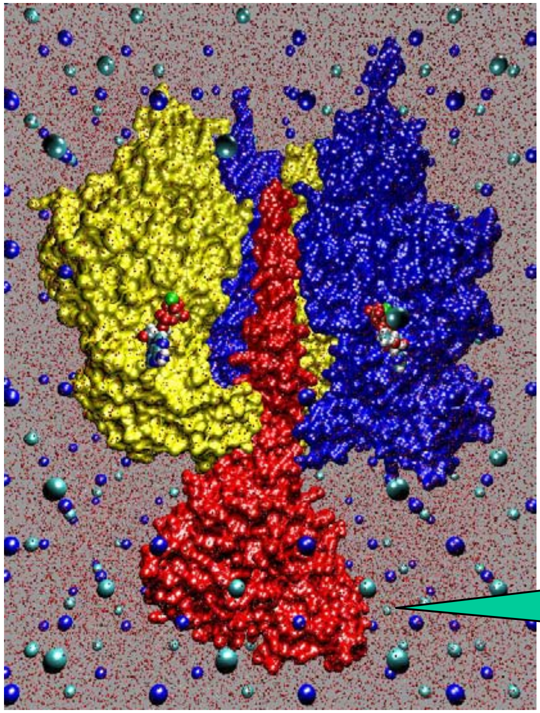


Visualization of Large All Atom Molecular Dynamics Simulations (2)

- Multiple representations show areas in appropriate detail
- Large models: 1,00,000 atoms and up
- Long trajectories: thousands of timesteps
- A 10 ns simulation of 100K atoms produces a 12GB trajectory
- Multi-gigabyte data sets break 32-bit addressing barriers



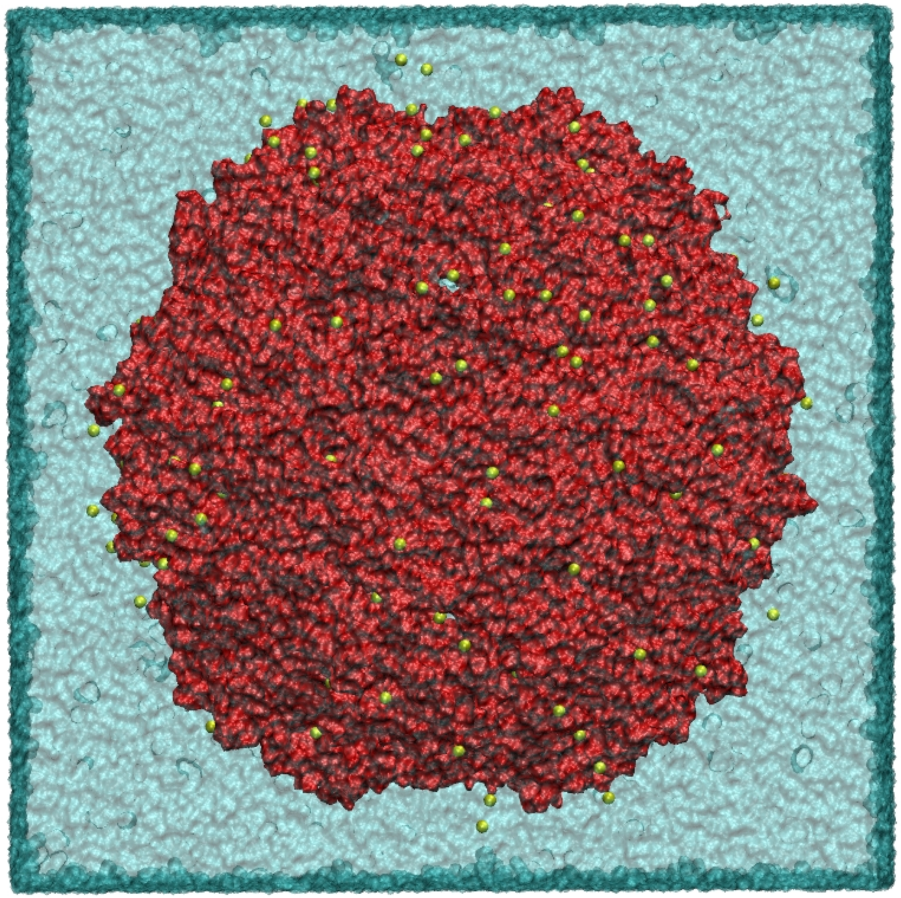
Purple Membrane
150,000 Atoms

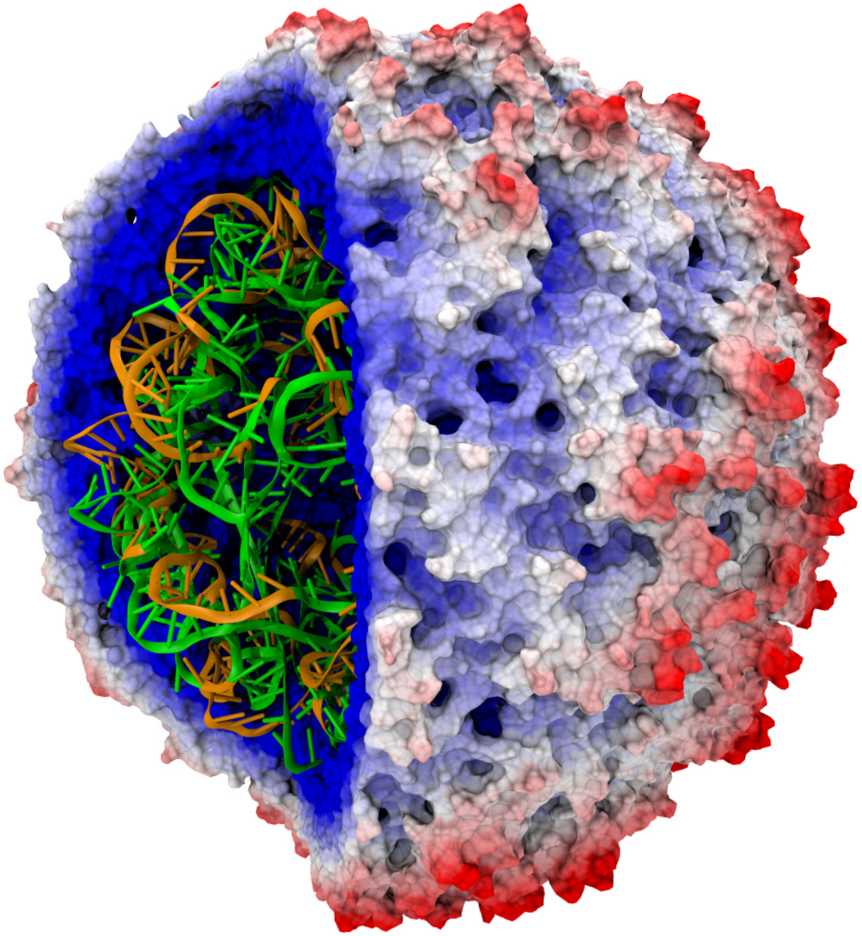


F1 ATPase
327,000 Atoms

Visualization of Large All Atom Molecular Dynamics Simulations (3)

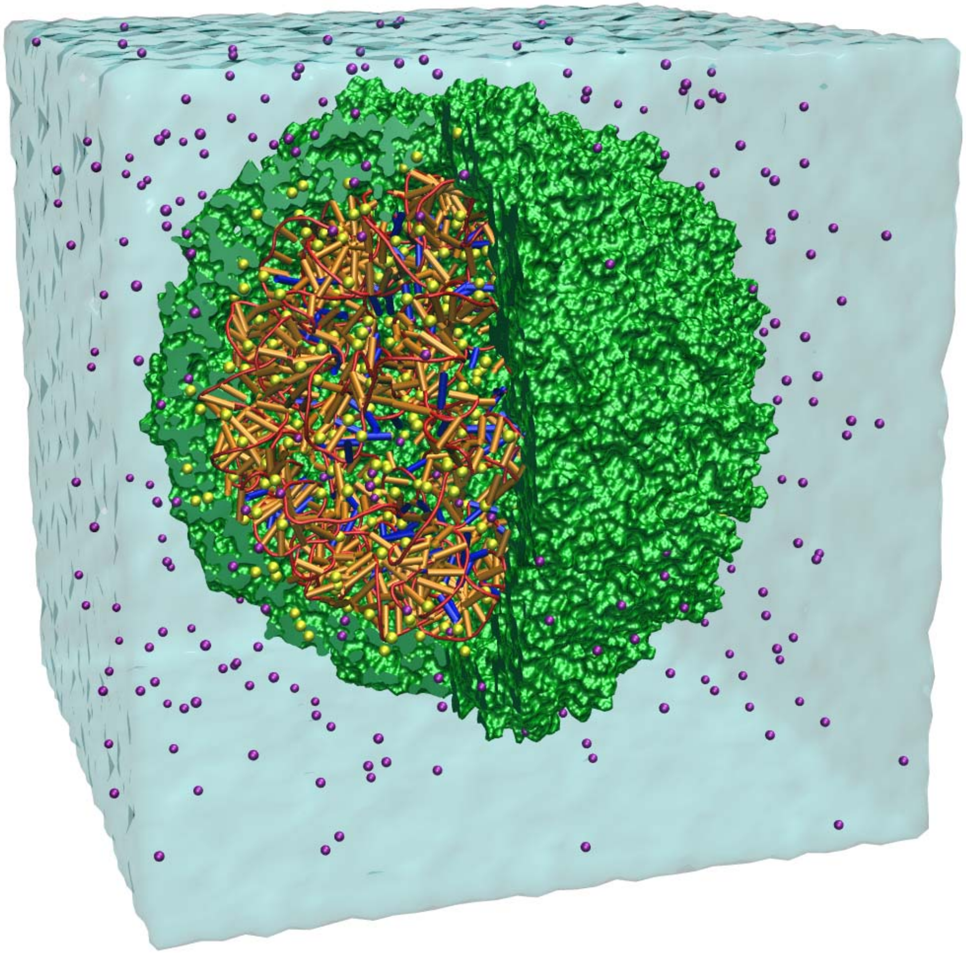
Satellite Tobacco Mosaic Virus 932,508 atoms





Visualizing Coarse-Grain Simulations

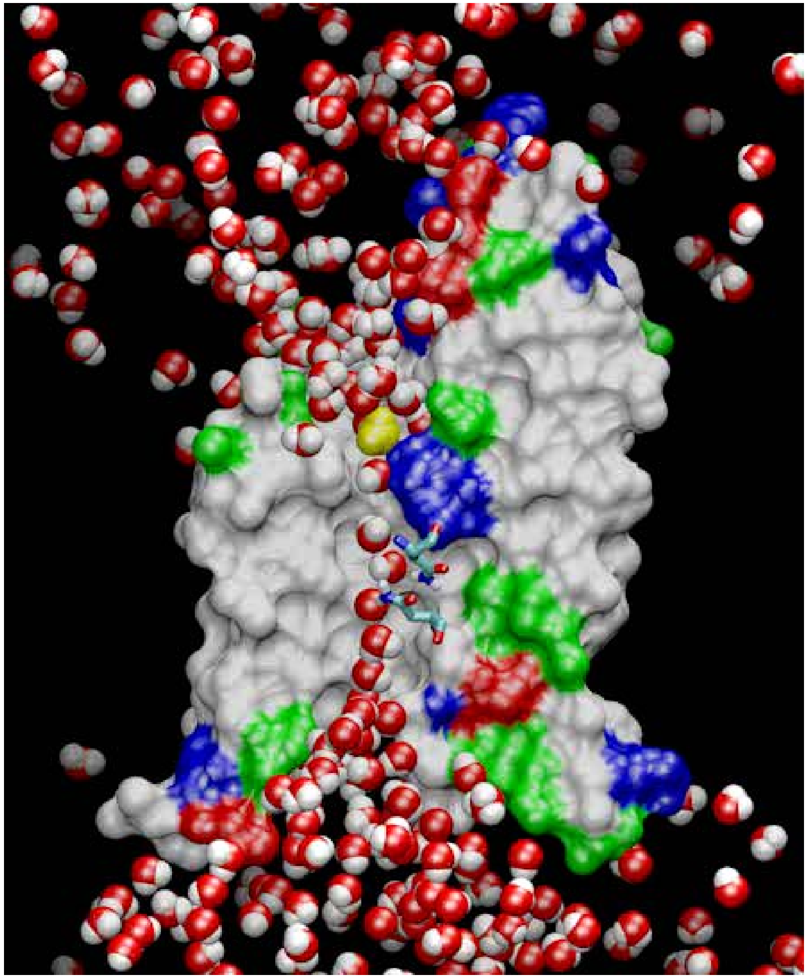
Satellite Tobacco Mosaic Virus, CG Model



- Visualization techniques can be used for both all-atom and CG models
- Groups of atoms replaced with “beads”, surface reps, or other geometry
- Display 1/20th the data

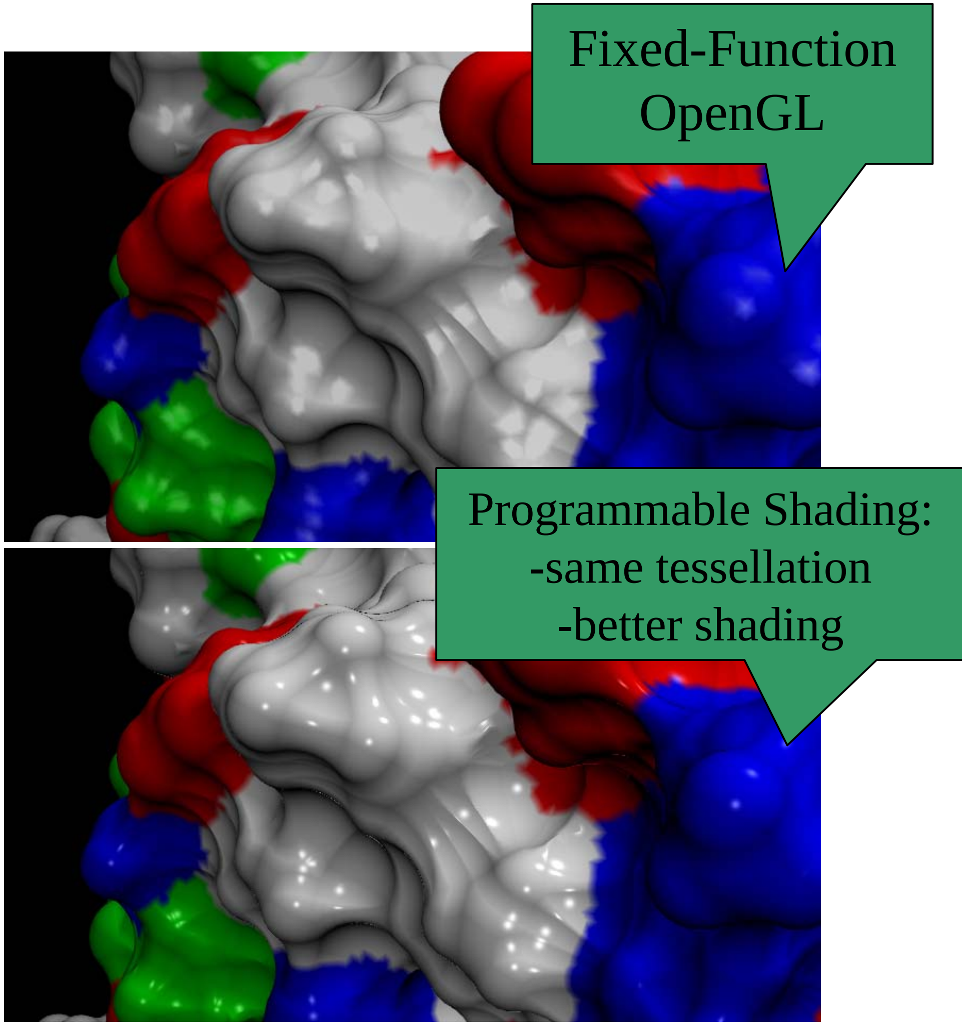
Trajectory Animation

- Motion aids perception of shape, understanding of dynamic processes
- Animate entire model, or just the parts where motion provides insight
- Window-average positions on-the-fly to focus on significant motions
- Selected atoms updated on-the-fly (distance constraints, etc)



Benefits of Programmable Shading (1)

- Potential for superior image quality with better shading algorithms
- Direct rendering of:
  - Quadric surfaces
  - Density map data, solvent surfaces
- Offload work from host CPU to GPU



Fixed-Function OpenGL

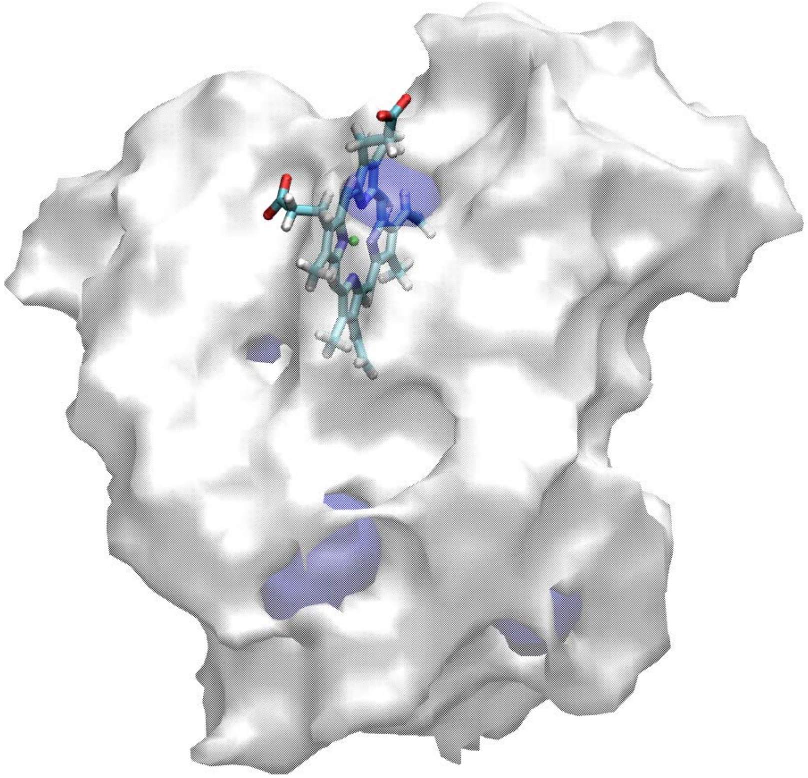
Programmable Shading:
-same tessellation
-better shading

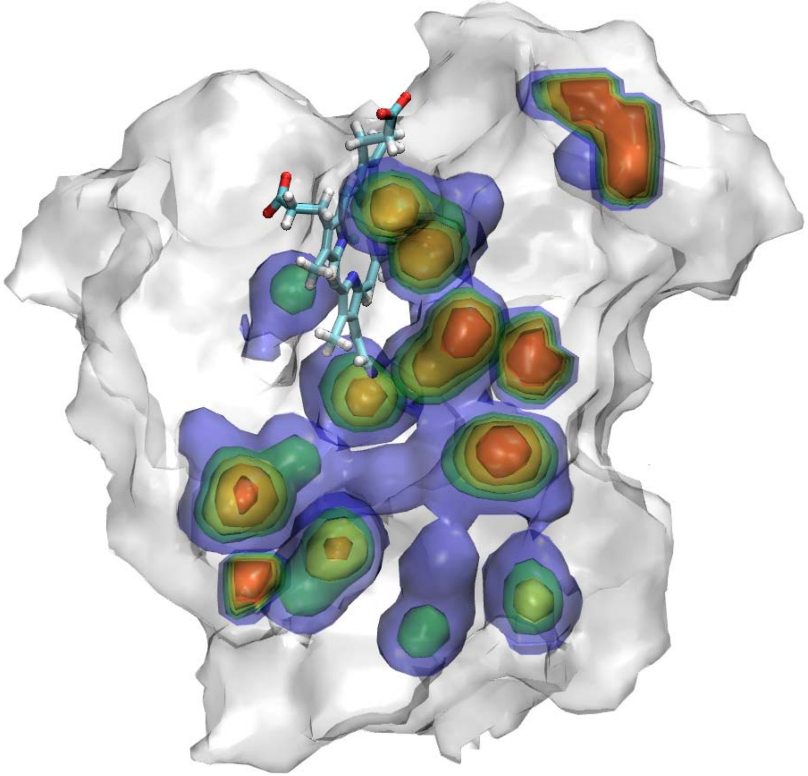
Benefits of Programmable Shading (2)

Myoglobin cavity “openness” (time averaged spatial occupancy)

Single-level OpenGL screen-door transparency obscures internal surfaces

Programmable shading shows transparent nested probability density surfaces with similar performance





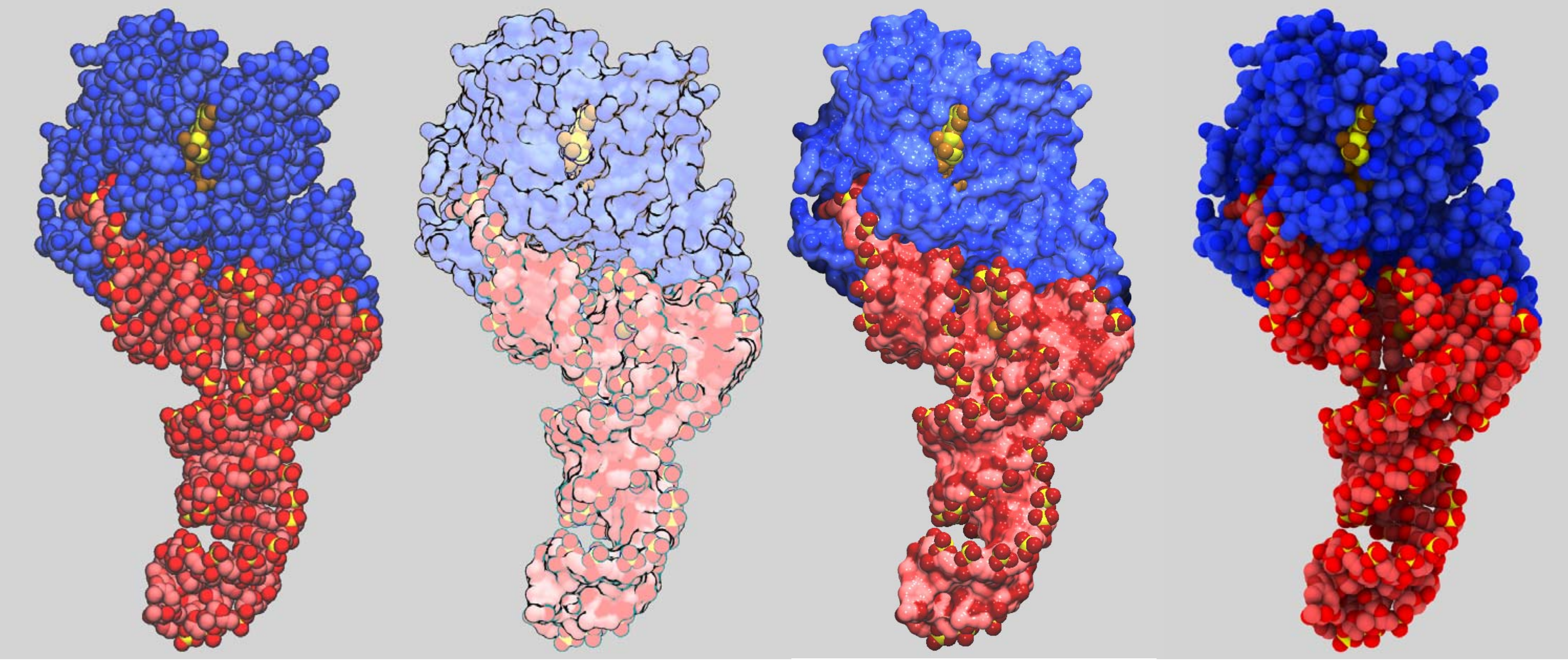
Shading Comparison: EF-Tu

Outline Shader

“Goodsell” Shader

Glossy Shader

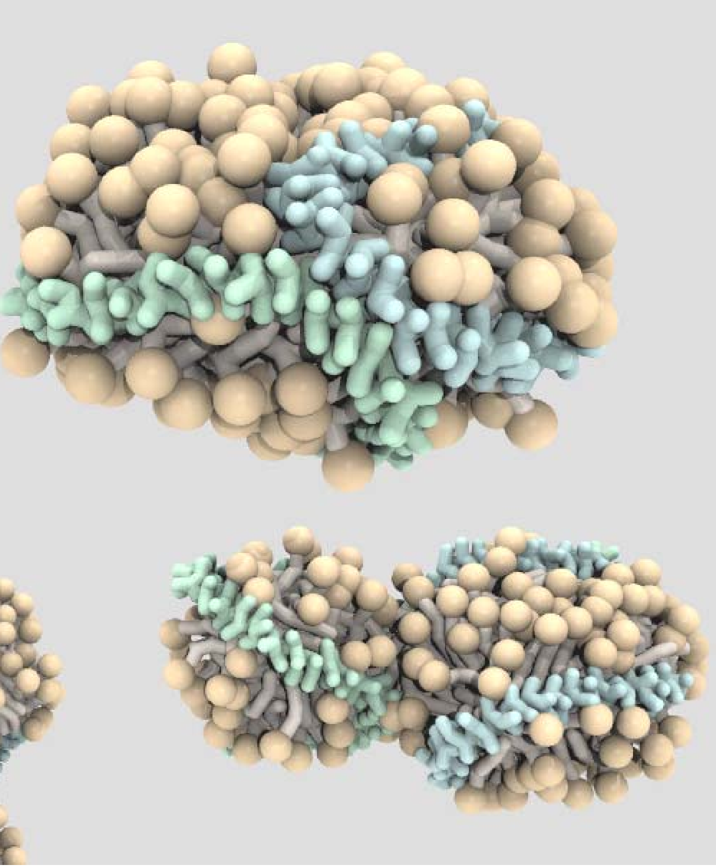
Ambient Occlusion, Shadowing



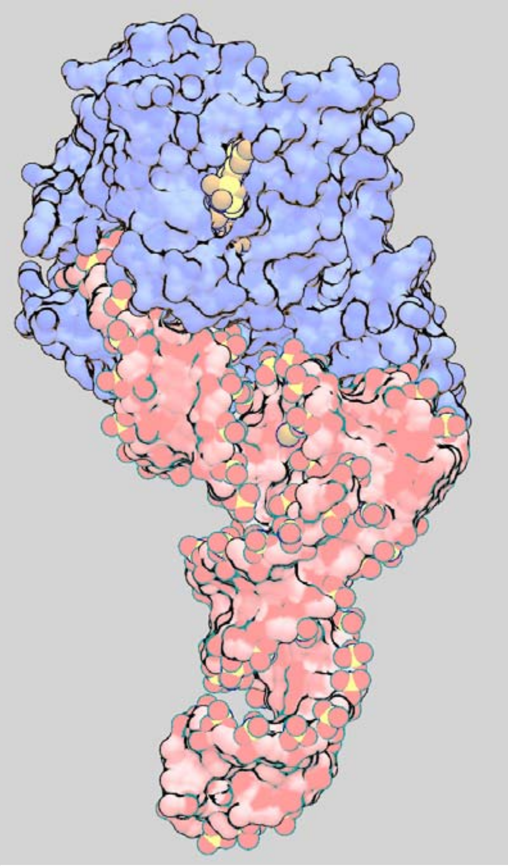
VMD Interactive OpenGL Rendering

VMD/Tachyon

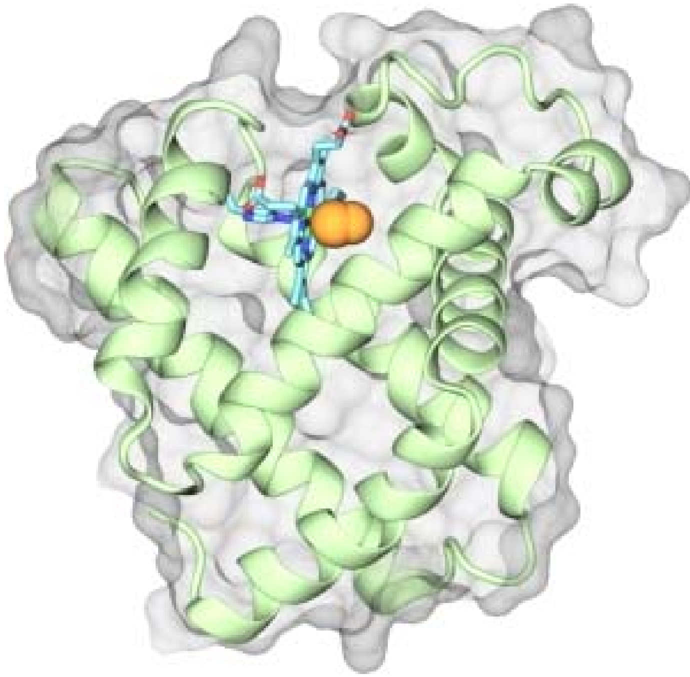
VMD Continues Incorporation of Shading, Lighting Advancements



Lipoprotein particle



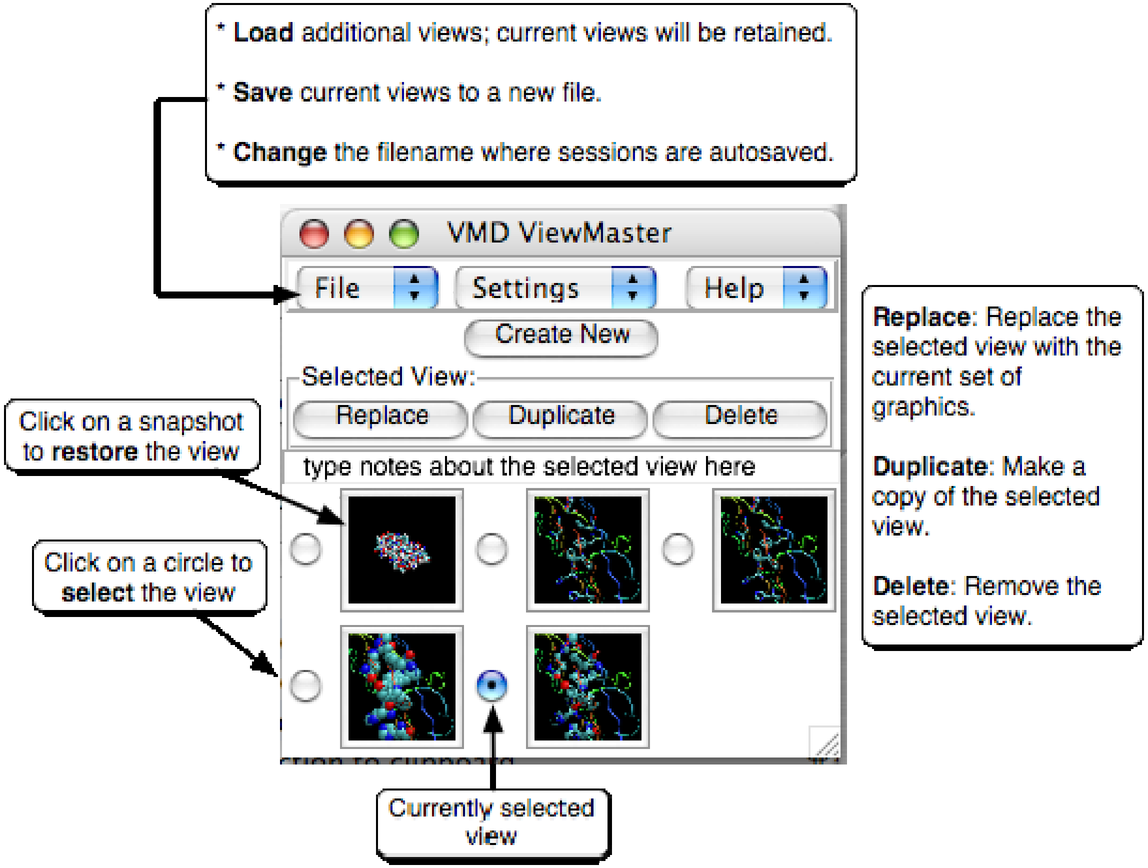
EF-Tu



Myoglobin

Rendering – VMD Supports Technical Communication

- Publication quality image creation
- Quick creation of documentation figures
- Easy creation of movies for lectures



* **Load** additional views; current views will be retained.

* **Save** current views to a new file.

* **Change** the filename where sessions are autosaved.

Click on a snapshot to **restore** the view

Click on a circle to **select** the view

Currently selected view

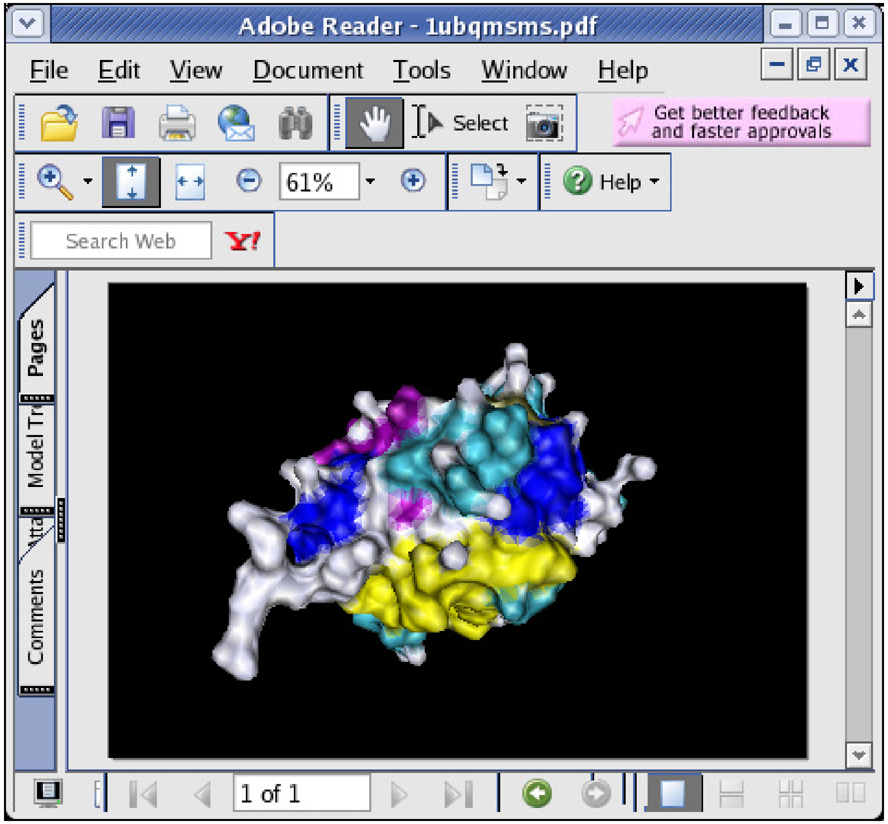
Replace: Replace the selected view with the current set of graphics.

Duplicate: Make a copy of the selected view.

Delete: Remove the selected view.

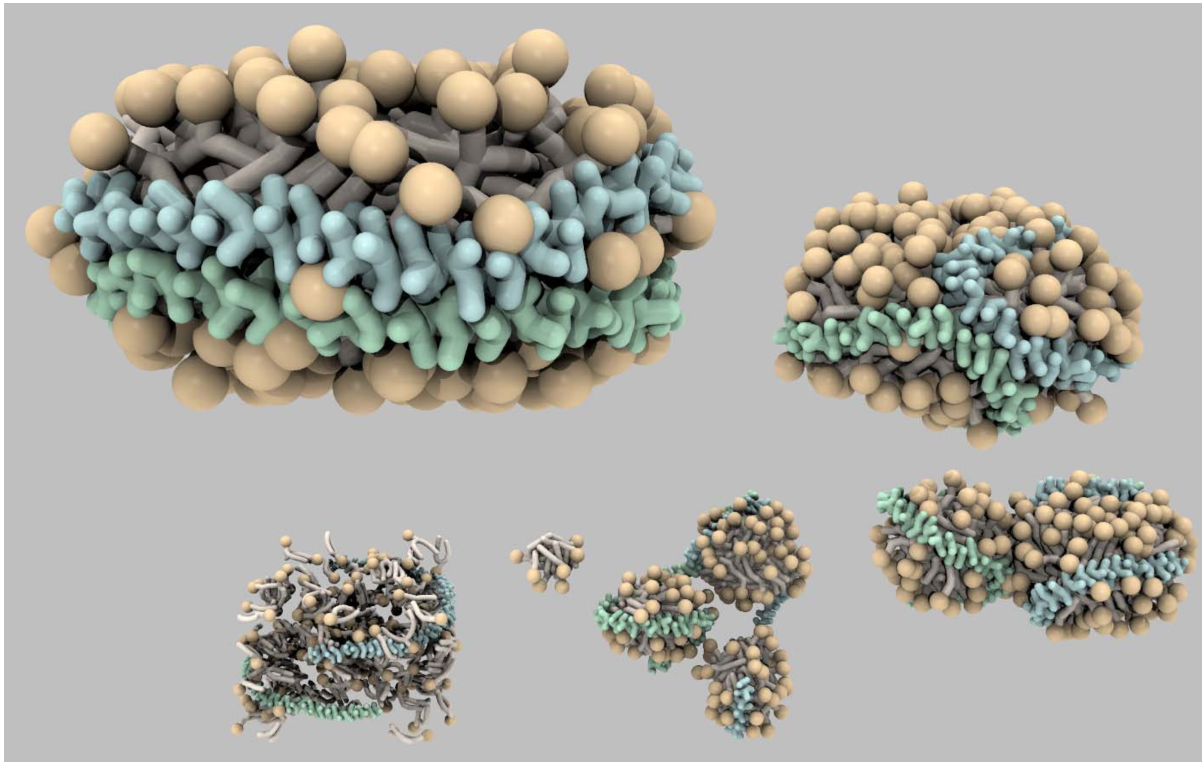
Rendering and Export of VMD Scenes

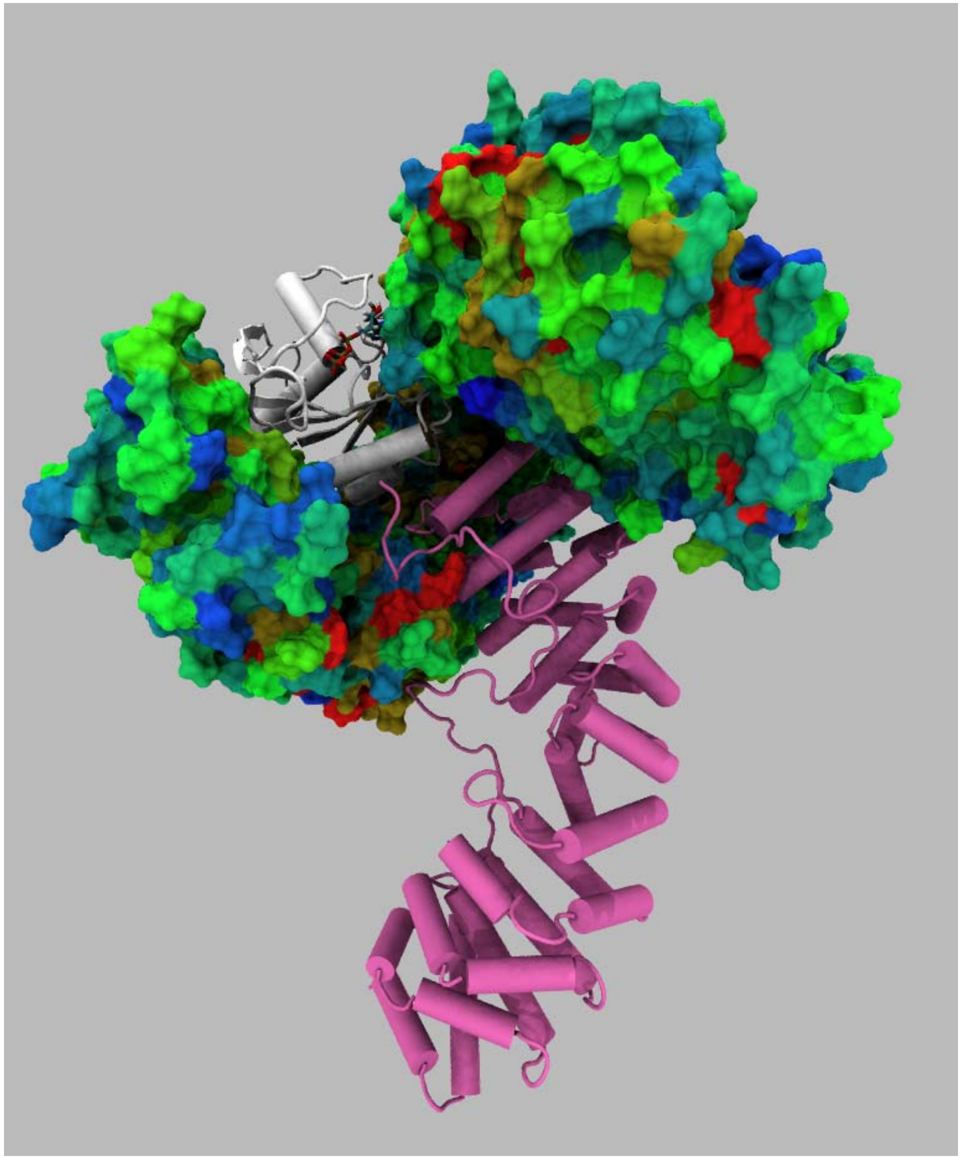
- Acrobat3D 3-D PDF
  - Interactive manipulation in Acroread
- Publication quality renderers:
  - Tachyon
  - POV-Ray
  - Raster3D



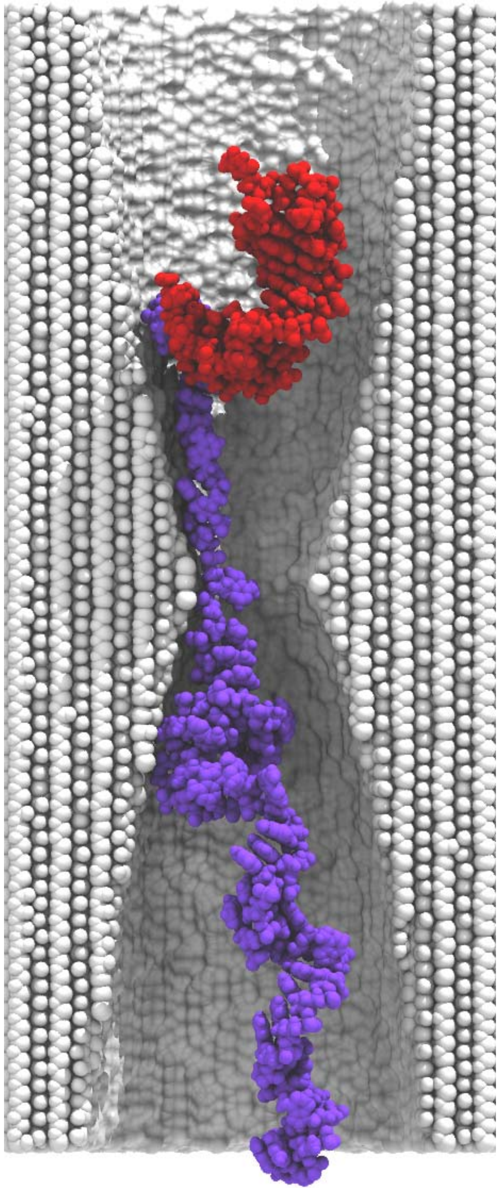
VMD/Tachyon Ambient Occlusion Lighting

- Omnidirectional diffuse lighting
- Improved shape perception
- Tachyon tuned for VMD
- Tachyon AO lighting works with all VMD representations



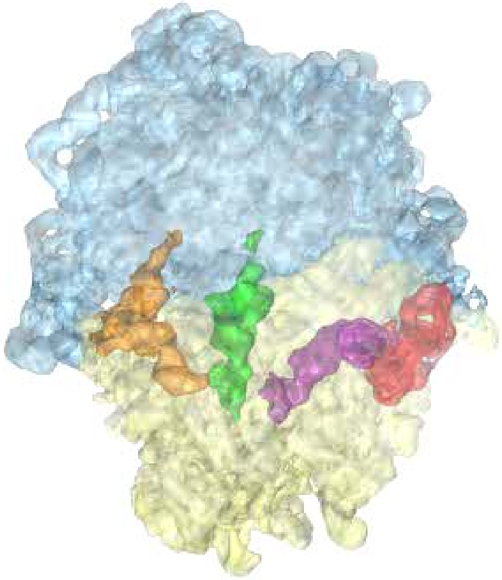


Exportin Cse1p



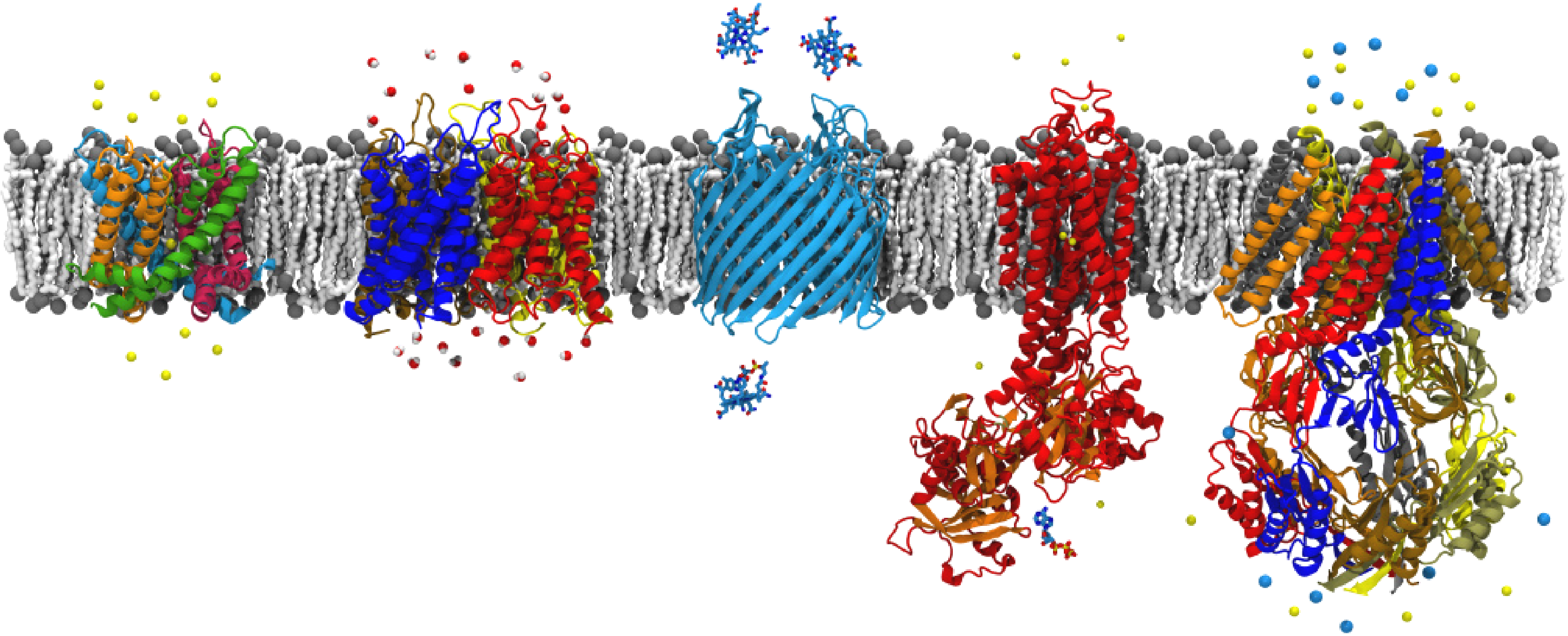
Nanopore

Ribosome Structure Docked to Cryo-EM Map



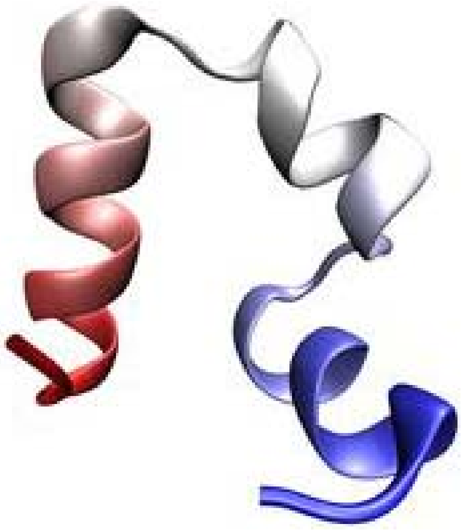
VMD Supports Large Datasets

- 72 million atom structure: 16GB RAM
  ~200 bytes per atom in worst case (water box)
- Load and operate on **entire Protein Databank,** over 40,000 molecules in 25GB RAM



VMD Supports Large Datasets

- Tested with 128 GB data in physical memory
- Loads trajectories faster than 1GB/sec
- New techniques required as simulations generate terabytes per run
  - Transparent operation on datasets much larger than computer's physical memory
  - Transparent access to remote datasets too large to move, e.g. petascale simulations: VMD+remote agents
  - Efficient file formats, compression, parallel I/O, …

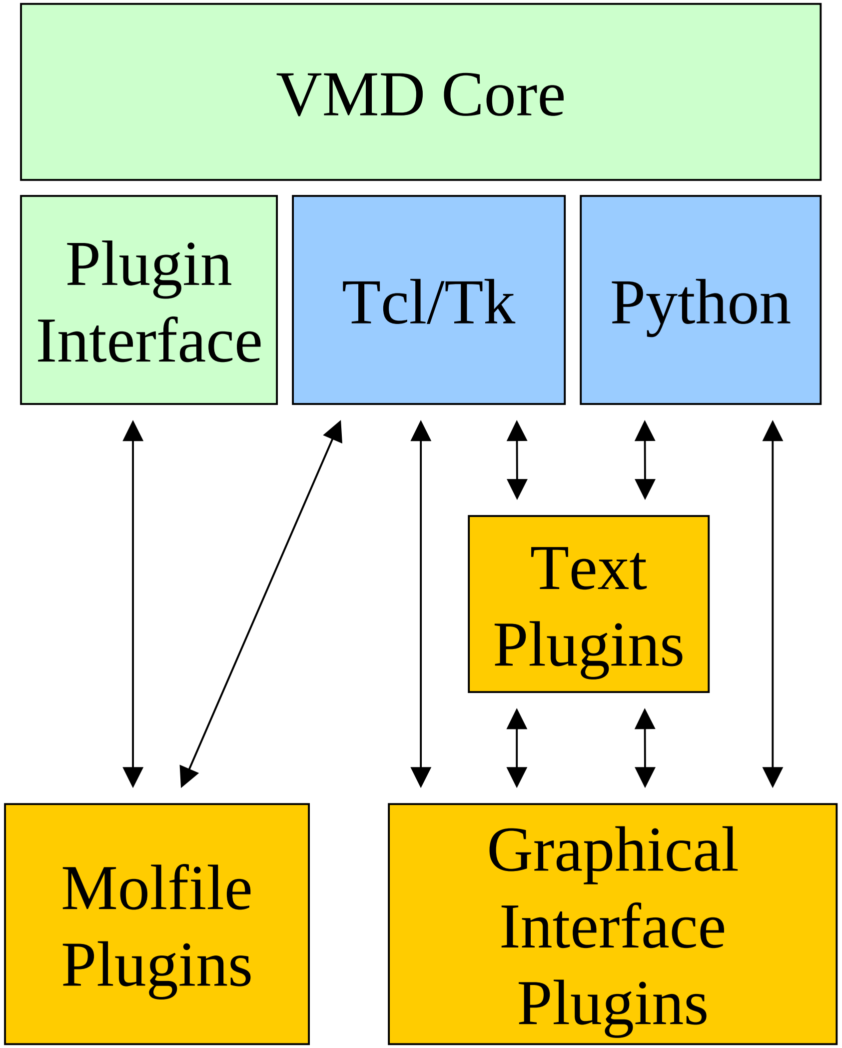


Villin headpiece

380 GB trajectory generated by a 6.9μs folding simulation of 30K atoms

VMD Extensibility

- Embedded Tcl/Tk, Python scripting
- Load VMD as a Python module
- Plugin System
  - Graphical interface extension plugins
  - Molfile plugins
  - Open source license
  - Over 25 developed by the user community
- Continued expansion and revision supports new data types and capabilities



VMD Core

Plugin Interface

Tcl/Tk

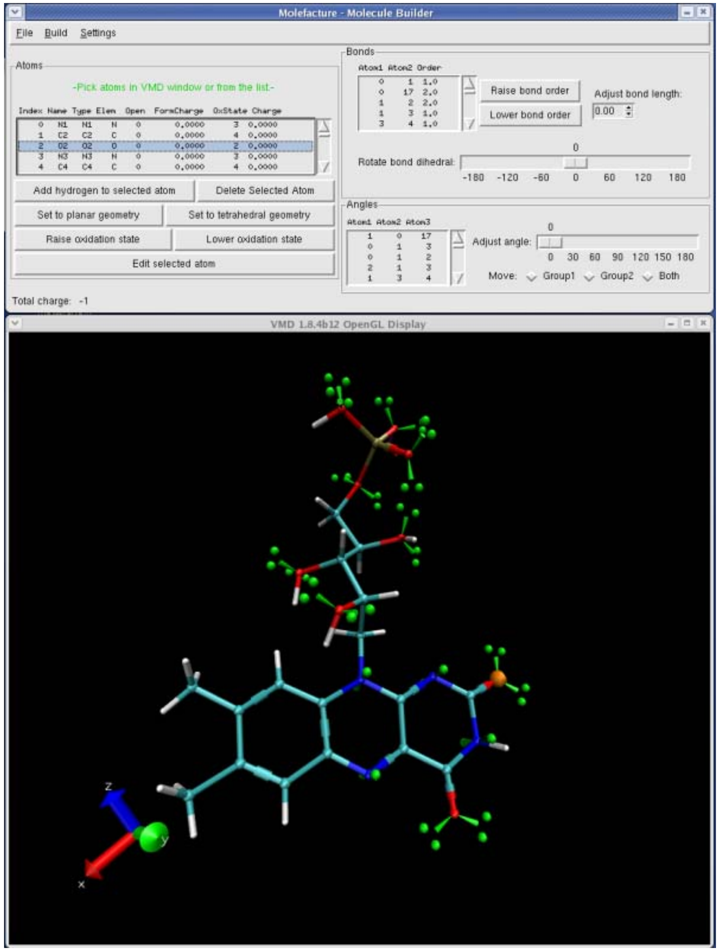
Python

Text Plugins

Molfile Plugins

Graphical Interface Plugins

Extensibility – VMD Adaptable by Researchers, Community Developed Tools



Index	Name	Type	Elem	Open	FormCharge	OxState	Charge
0	N1	N1	N	0	0.0000	3	0.0000
1	C2	C2	C	0	0.0000	4	0.0000
2	O2	O2	O	0	0.0000	2	0.0000
3	N3	N3	N	0	0.0000	3	0.0000
4	C4	C4	C	0	0.0000	4	0.0000

Total charge: -1

Atom1	Atom2	Order
0	1	1.0
0	17	2.0
1	2	2.0
1	3	1.0
3	4	1.0

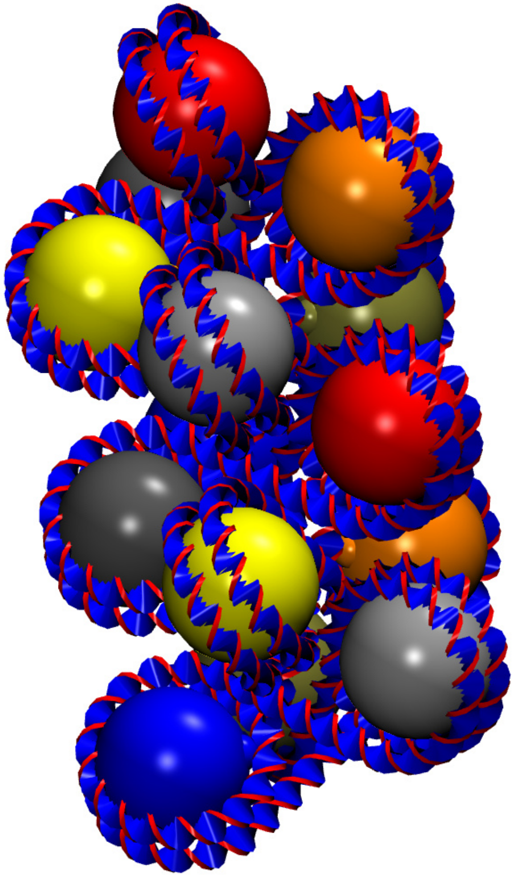
Adjust bond length: 0.00

Rotate bond dihedral: 0

Atom1	Atom2	Atom3
1	0	17
0	1	3
0	1	2
2	1	3
1	3	4

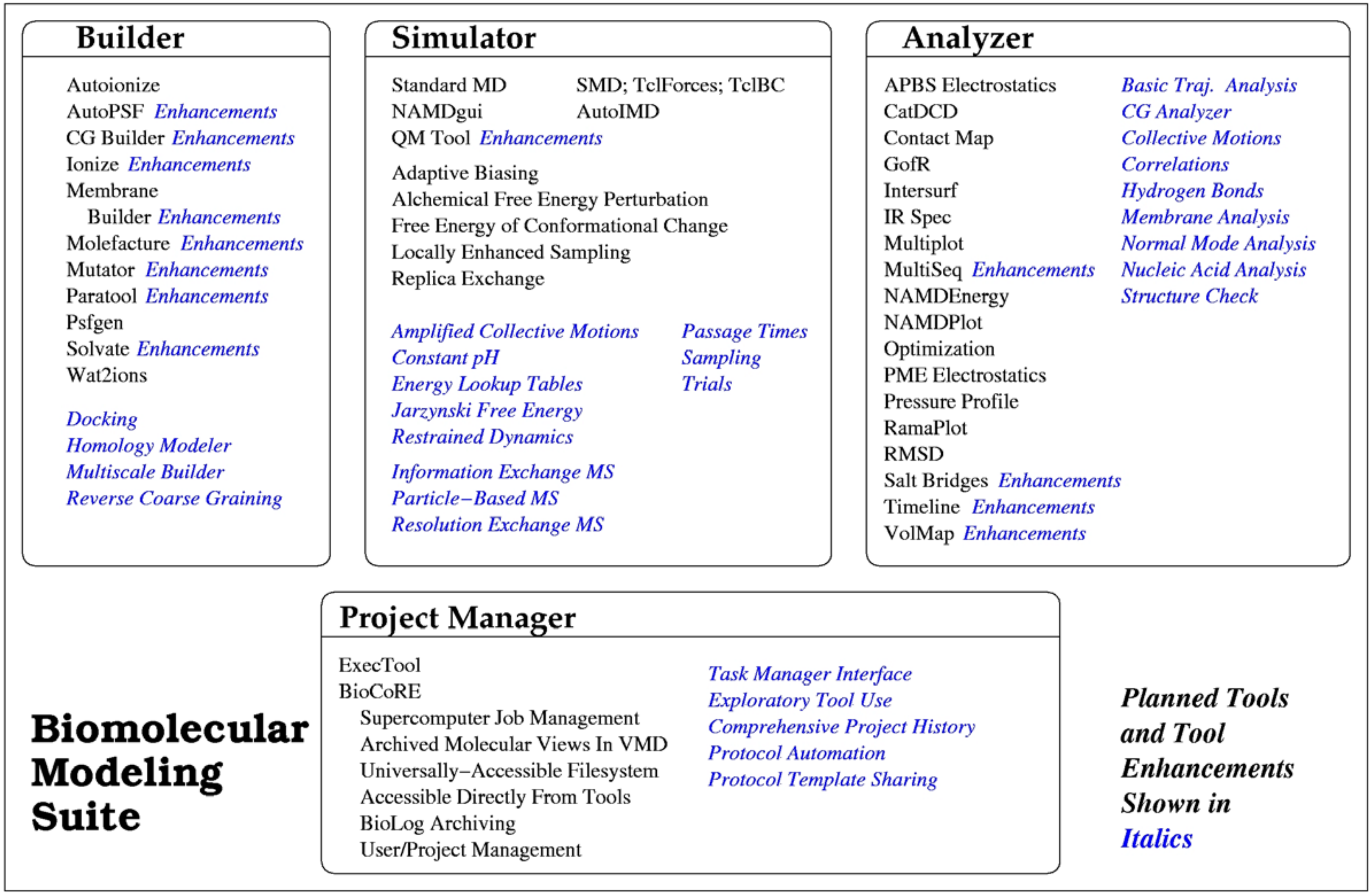
Adjust angle: 0

Molefacture Plugin



“VDNA” Chromatin Rendering Script

VMD Extensibility and Infrastructure Is Fundamental to Development Efforts



Builder

Autoionize
AutoPSF *Enhancements*
CG Builder *Enhancements*
Ionize *Enhancements*
Membrane
 Builder *Enhancements*
Molefacture *Enhancements*
Mutator *Enhancements*
Paratool *Enhancements*
Psfgen
Solvate *Enhancements*
Wat2ions

Docking
Homology Modeler
Multiscale Builder
Reverse Coarse Graining

Simulator

Standard MD
NAMDgui
QM Tool *Enhancements*
SMD; TclForces; TclBC
AutoIMD

Adaptive Biasing
Alchemical Free Energy Perturbation
Free Energy of Conformational Change
Locally Enhanced Sampling
Replica Exchange

Amplified Collective Motions
Constant pH
Energy Lookup Tables
Jarzynski Free Energy
Restrained Dynamics
Passage Times
Sampling
Trials

Information Exchange MS
Particle–Based MS
Resolution Exchange MS

Analyzer

APBS Electrostatics
CatDCD
Contact Map
GofR
Intersurf
IR Spec
Multiplot
MultiSeq *Enhancements*
NAMDEnergy
NAMDPlot
Optimization
PME Electrostatics
Pressure Profile
RamaPlot
RMSD
Salt Bridges *Enhancements*
Timeline *Enhancements*
VolMap *Enhancements*

Basic Traj. Analysis
CG Analyzer
Collective Motions
Correlations
Hydrogen Bonds
Membrane Analysis
Normal Mode Analysis
Nucleic Acid Analysis
Structure Check

Project Manager

ExecTool
BioCoRE
 Supercomputer Job Management
 Archived Molecular Views In VMD
 Universally–Accessible Filesystem
 Accessible Directly From Tools
 BioLog Archiving
 User/Project Management

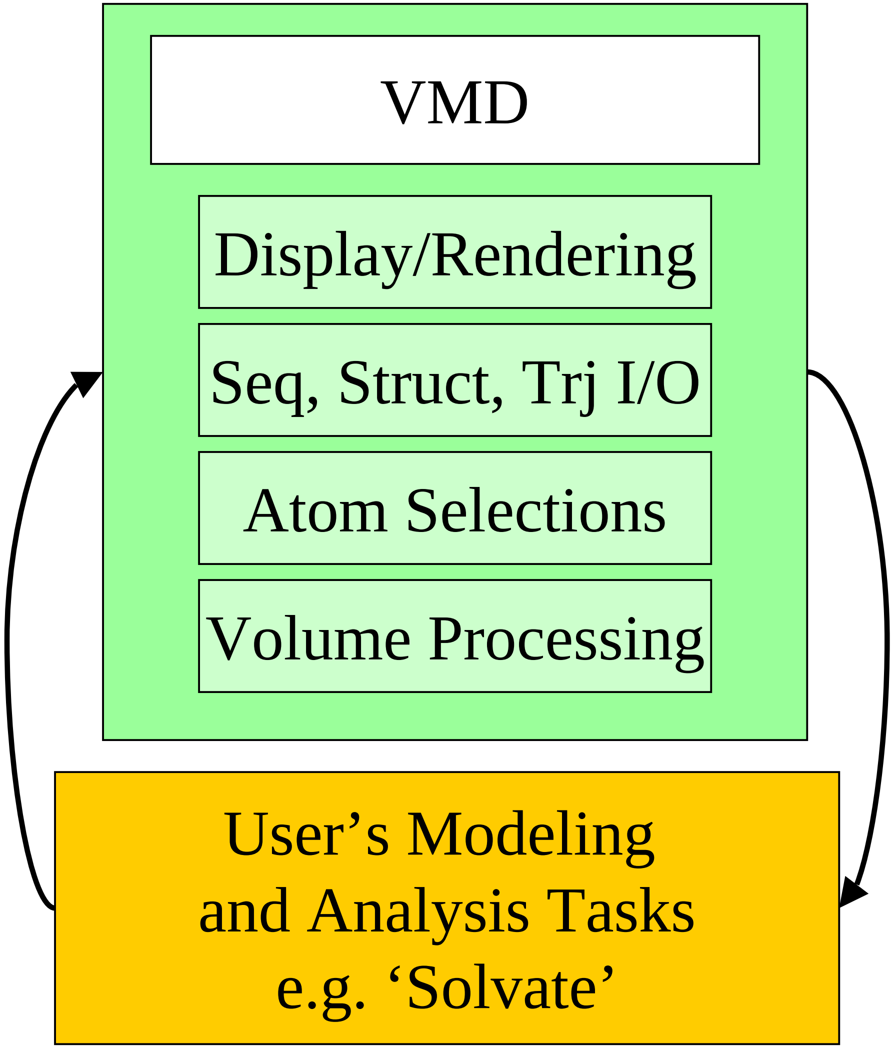
Task Manager Interface
Exploratory Tool Use
Comprehensive Project History
Protocol Automation
Protocol Template Sharing

Biomolecular Modeling Suite

Planned Tools and Tool Enhancements Shown in Italics

VMD Analysis Infrastructure

- VMD does “heavy lifting” for custom analysis tasks
- Link with external tools
- Users do their work via graphical plugins or scripting
- Scripting language revisions will further improve batch mode, and parallel processing capabilities



VMD

Display/Rendering

Seq, Struct, Trj I/O

Atom Selections

Volume Processing

User’s Modeling and Analysis Tasks e.g. ‘Solvate’

VMD Takes Advantage of Emerging Technological Opportunities

- 8- and 12-core CPUs common by 2010…
- Graphics processors (GPUs) have over 240 processing units, and can achieve speedups of 8-30x vs. CPUs
- Parallel processing is now **required** to increase performance
- Several VMD algorithms are now parallelized for multi-core CPUs and GPUs
- Continued developments will more broadly benefit rendering and analysis features of VMD

Many VMD Tutorials Available!

- The easiest way to learn VMD is to work through the tutorials that are linked on the main VMD web site:
  - http://www.ks.uiuc.edu/Research/vmd/