Improving Protein Function Prediction with Molecular Dynamics Simulations

Dariya Glazer

Russ Altman

Motivation

Sometimes the 3D structure doesn't score well for a known function.

The experimental structure may have an altered configuration of atoms.

Possible Solution

Treat molecules as dynamic entities, sample other conformations and test for function

Potential Energy function:

$$U = \sum_{bonds} K_b (b_i - b_o)^2$$

$$+ \sum_{angles} K_\theta (\theta_i - \theta_o)^2$$

$$+ \sum_{dihedrals} K_\phi \left[1 - \cos(n\phi_i + \delta)\right]$$

$$+ \sum_{pairs} \varepsilon \left[\left(\frac{r_o}{r_{ij}} \right)^{12} - 2 \left(\frac{r_o}{r_{ij}} \right)^{6} \right]$$

$$+ \sum_{charges} \frac{q_i q_j}{r_{ij}}$$

Molecular Dynamics

- Given the equations for energy, and the time scale, perform a large numerical simulation to see how the molecule moves.

- Time step = 0.001 ps
- Starting point = experimentally defined structure (already in good local minima)
- Include water molecules around protein
- Goal: study dynamic motion of protein

Numerical Solution for M.D.

$$x(t+\Delta t) = x(t) + v(t)\Delta t + \left[4a(t) - a(t-\Delta t)\right]\Delta t^2/6$$

$$v(t+\Delta t) = v(t) + \left[2a(t+\Delta t) + 5a(t) - a(t-\Delta t)\right]\Delta t/6$$

$$U_{kinetic} = \frac{1}{2}\sum m_i v_i(t)^2$$

$$U_{potential} = \ldots$$

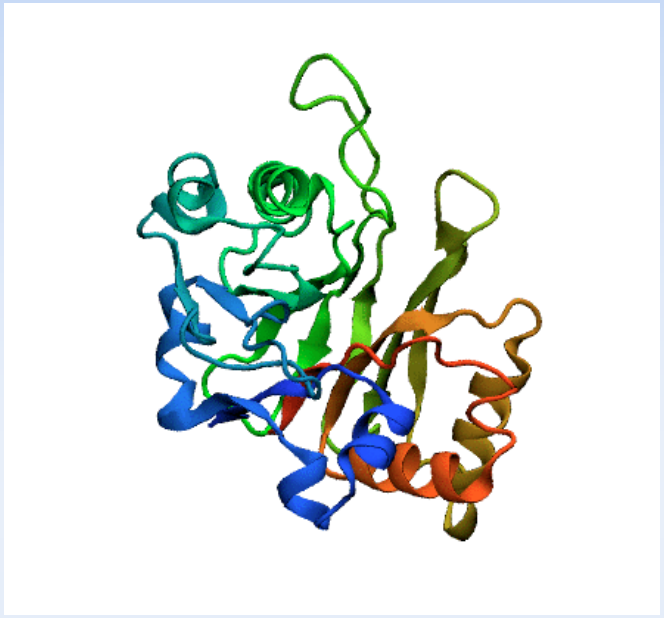
$$U_{total} = U_{kinetic} + U_{potential}$$

$$F = dU / dx$$

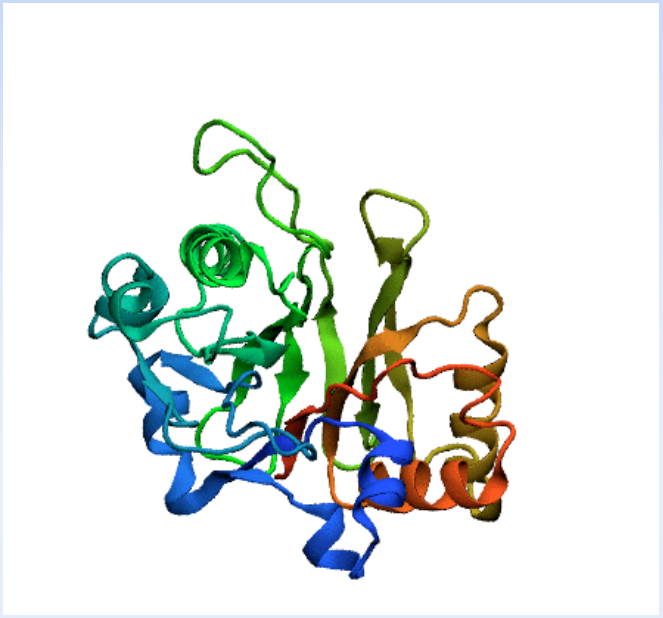
$$a = F / m$$

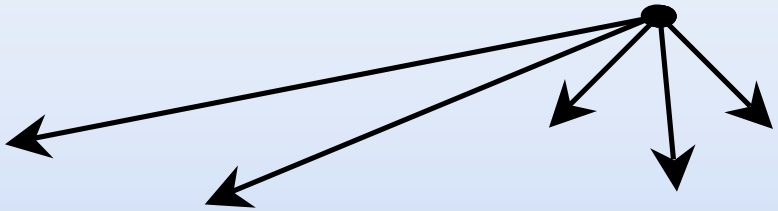
Molecules Come Alive

Static Structure: n = 1



Simulation Trajectory

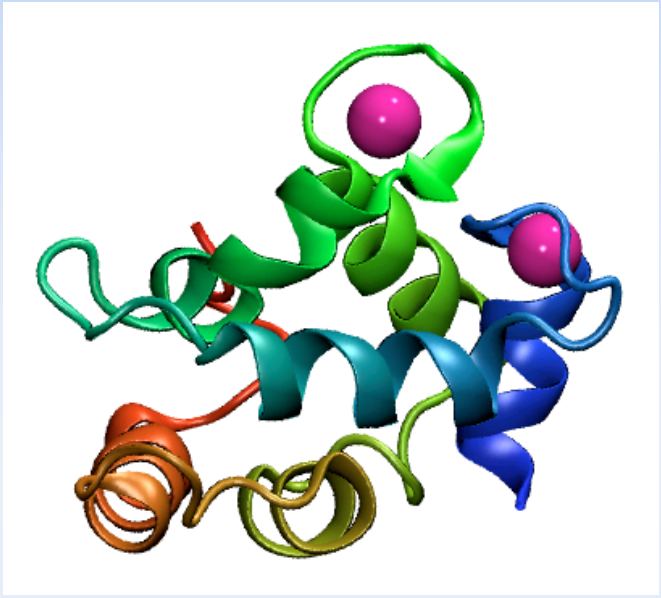




Structural Diversity: n >>>> 1

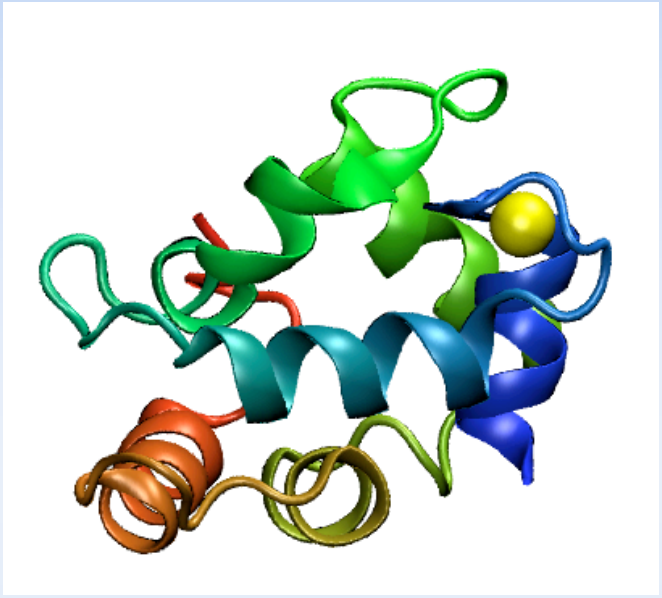
Starting Static Structures: PDB

5 x HOLO



with bound ligand

5 x APO



without the ligand

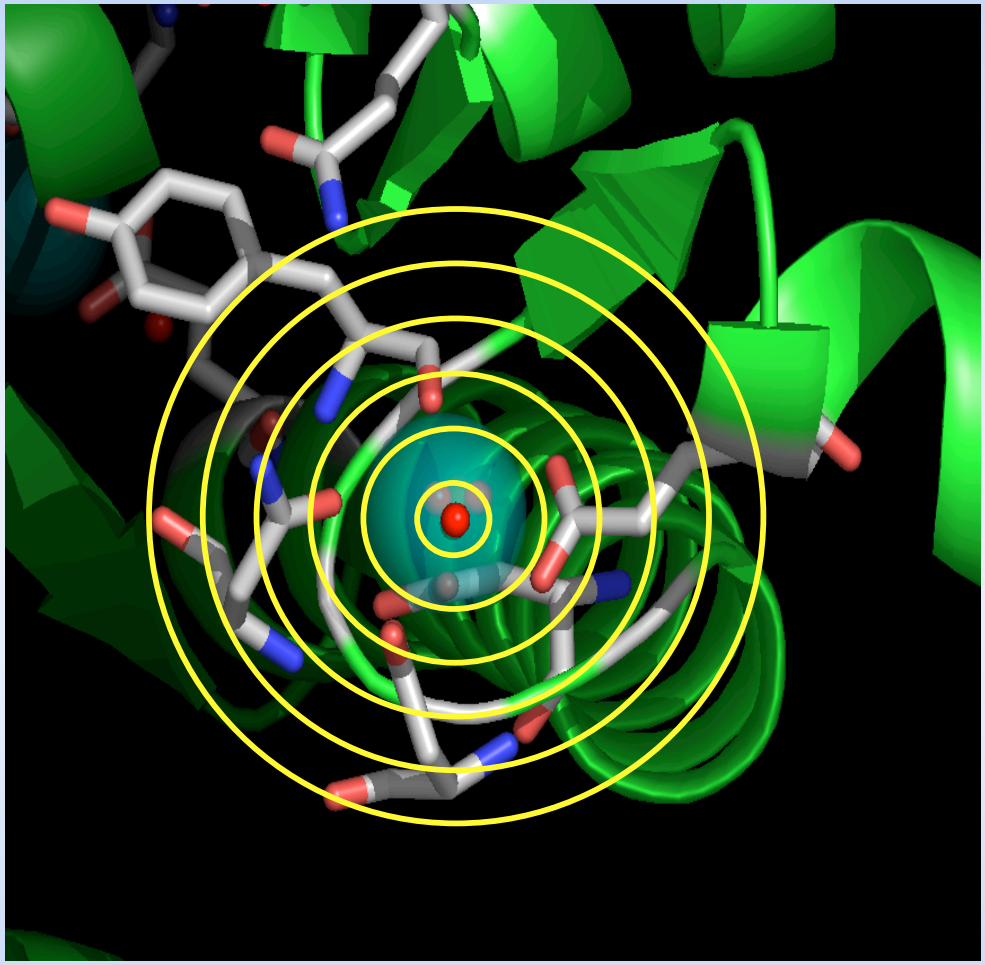
Simulations:

GROMACS, duration 1ns, explicit solvation

Scoring Function:

FEATURE, resolution 1Å

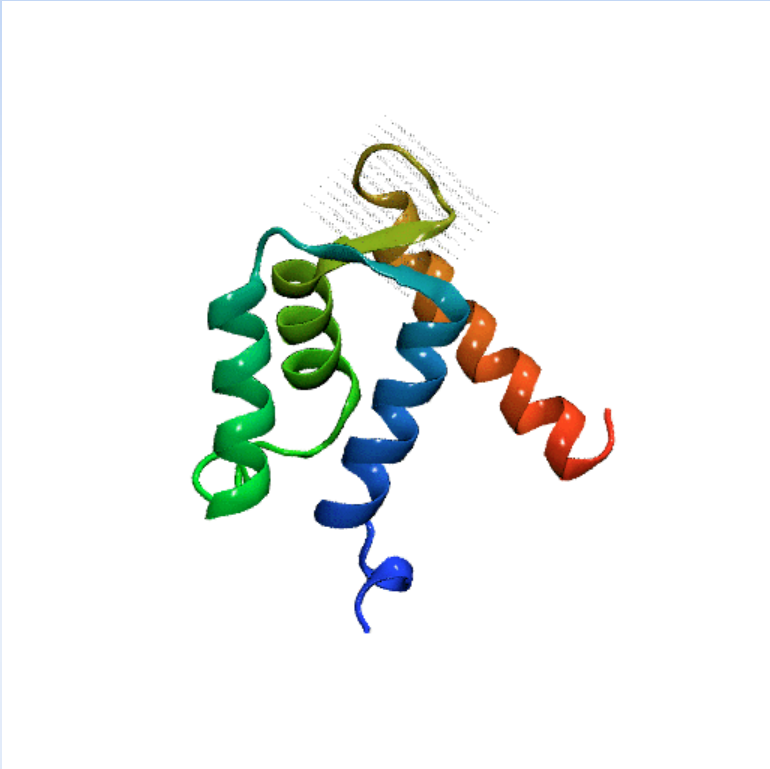
FEATURE & Ca^{2+} Model



MODEL FEATURES		SHELL				
	0	1	2	3	4	5
ATOM-NAME-IS-ANY						
ATOM-NAME-IS-C						
ATOM-NAME-IS-N						
ATOM-NAME-IS-O						
ATOM-NAME-IS-S						
ATOM-NAME-IS-OTHER						
HYDROXYL						
AMIDE						
AMINE						
CARBONYL						
RING-SYSTEM						
PEPTIDE						
VDW-VOLUME						
CHARGE						
NEG-CHARGE						
POS-CHARGE						
CHARGE-WITH-HIS						
HYDROPHOBICITY						
MOBILITY						
SOLVENT-ACCESSIBILITY						
RESIDUE_NAME_IS_GLU						
RESIDUE_NAME_IS_GLY						
RESIDUE_NAME_IS_LEU						
RESIDUE_NAME_IS_LYS						
RESIDUE_NAME_IS_PRO						
RESIDUE_NAME_IS_VAL						
RESIDUE_NAME_IS_HOH						
RESIDUE_CLASS1_IS_HYDROPHOBIC						
RESIDUE_CLASS1_IS_CHARGED						
RESIDUE_CLASS1_IS_POLAR						
RESIDUE_CLASS1_IS_UNKNOWN						
RESIDUE_CLASS2_IS_NONPOLAR						
RESIDUE_CLASS2_IS_POLAR						
RESIDUE_CLASS2_IS_BASIC						
RESIDUE_CLASS2_IS_ACIDIC						
RESIDUE_CLASS2_IS_UNKNOWN						
SECONDARY_STRUCTURE1_IS_3HELIX						
SECONDARY_STRUCTURE1_IS_4HELIX						
SECONDARY_STRUCTURE1_IS_BRIDGE						
SECONDARY_STRUCTURE1_IS_STRAND						
SECONDARY_STRUCTURE1_IS_TURN						
SECONDARY_STRUCTURE1_IS_BEND						
SECONDARY_STRUCTURE1_IS_COIL						
SECONDARY_STRUCTURE1_IS_HET						
SECONDARY_STRUCTURE2_IS_HELIX						
SECONDARY_STRUCTURE2_IS_BETA						
SECONDARY_STRUCTURE2_IS_COIL						

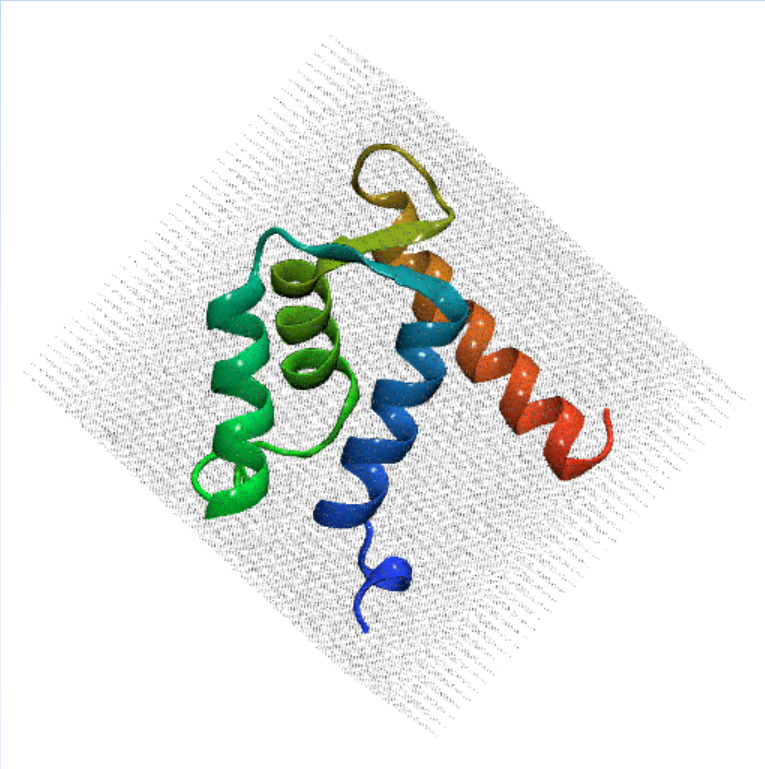
Threshold score of this model is 50

Simulation Analysis



Local Grid: to observe dynamics of a known binding site

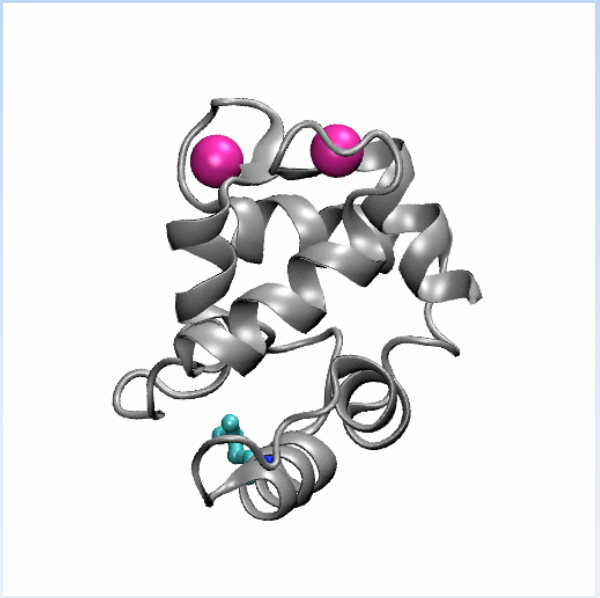
- Keep only the topmost score for each frame

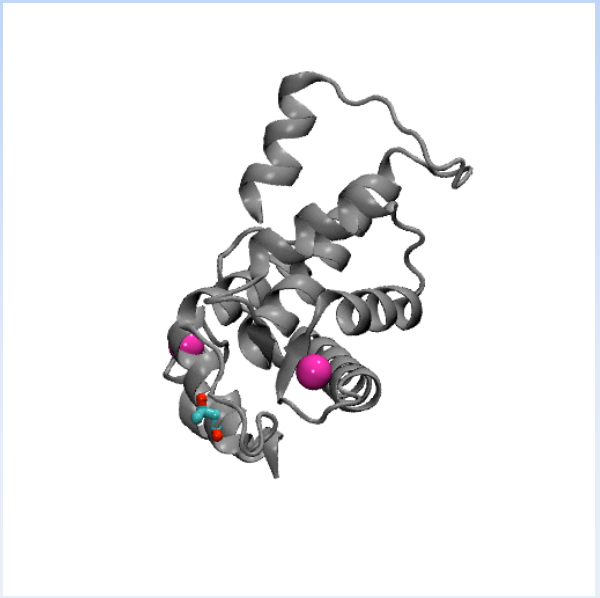


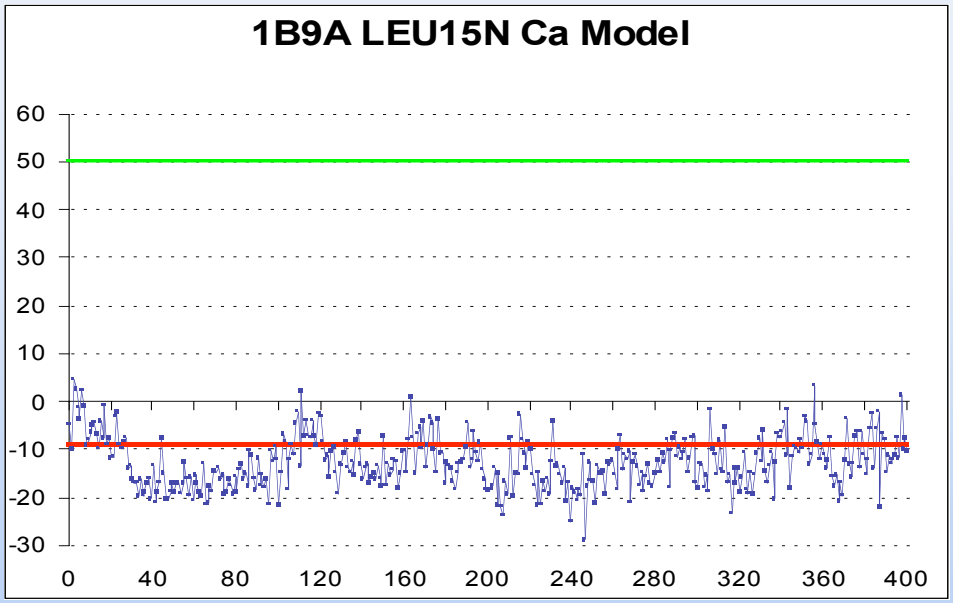
Global Grid: to find all possible binding sites in a given structure

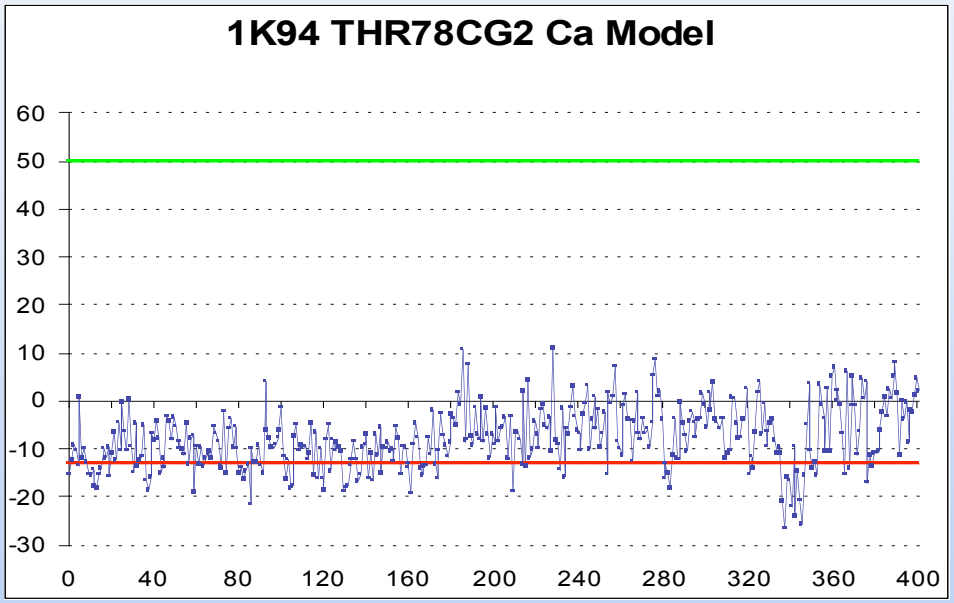
- Keep all scores over 50 for each frame

Local Grid: Non-Sites Score Low



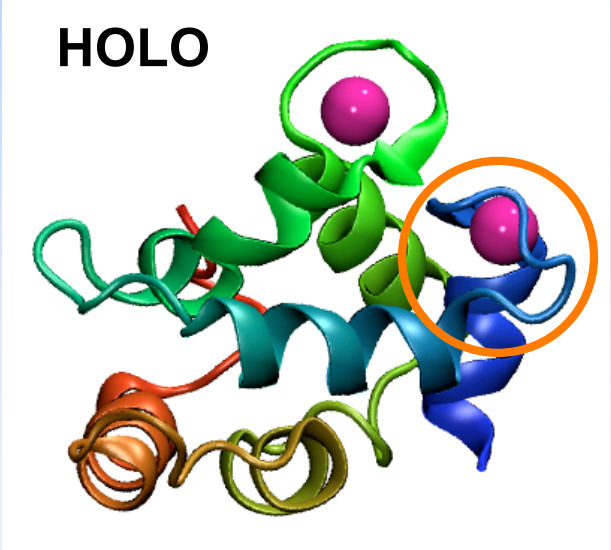


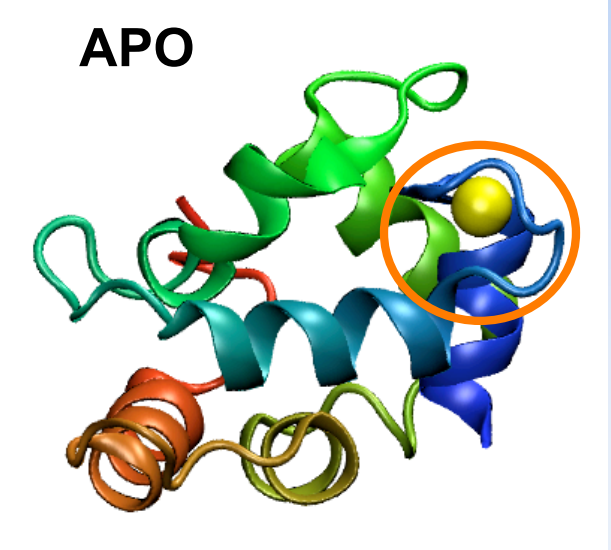


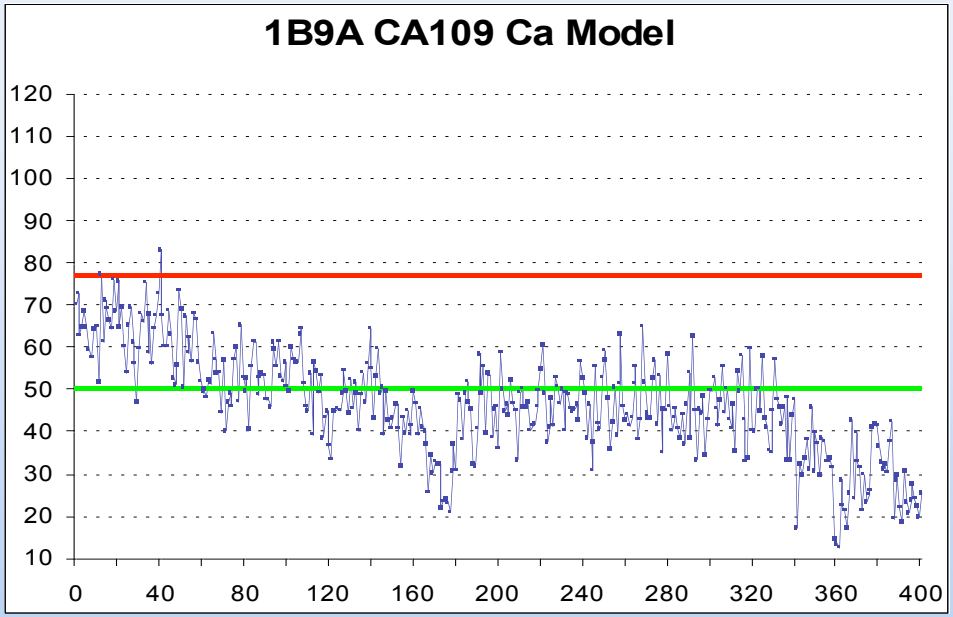


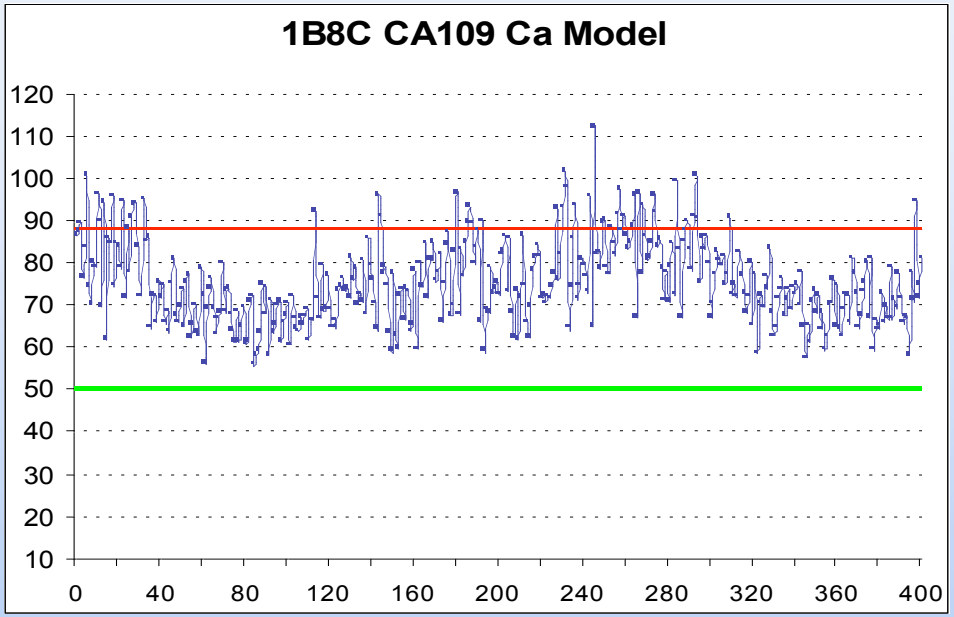
Simulation **Original Score** **Model Threshold**

Results Agree With Experiments



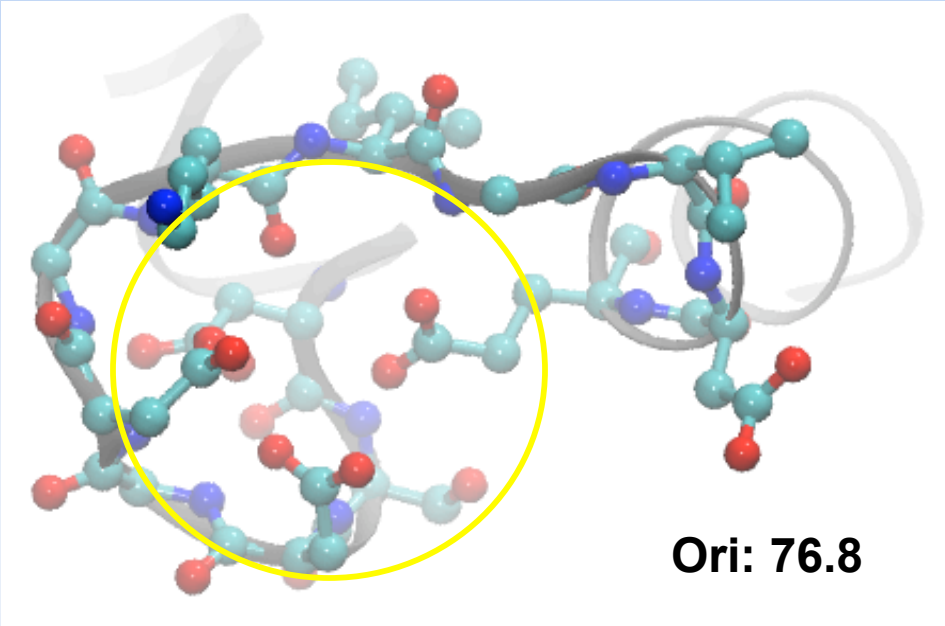


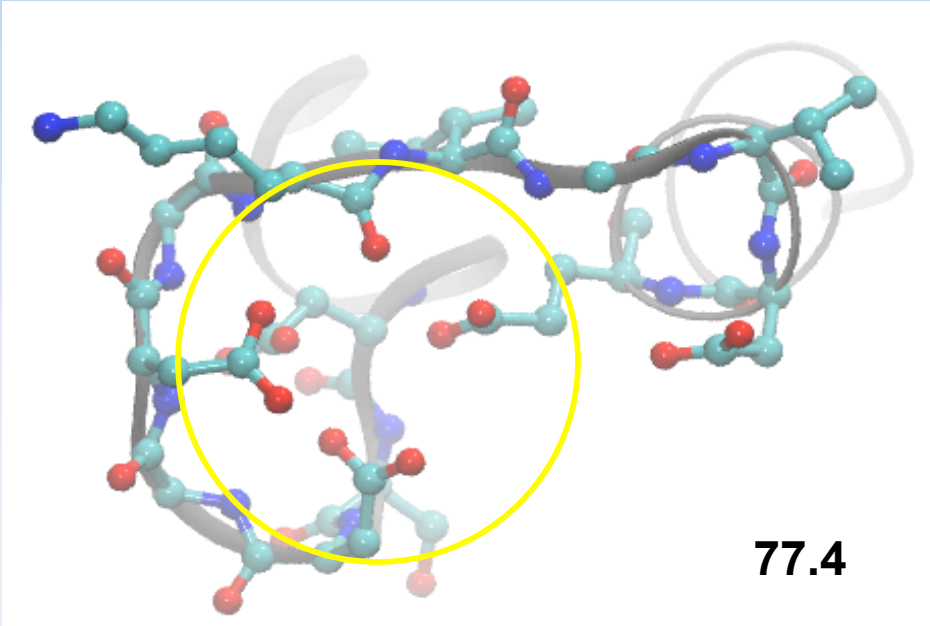


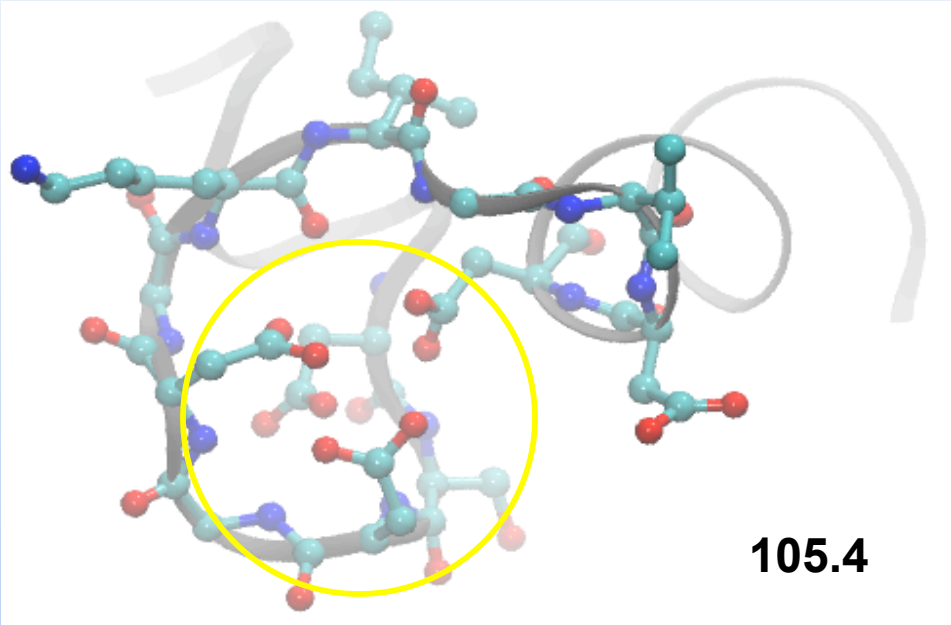


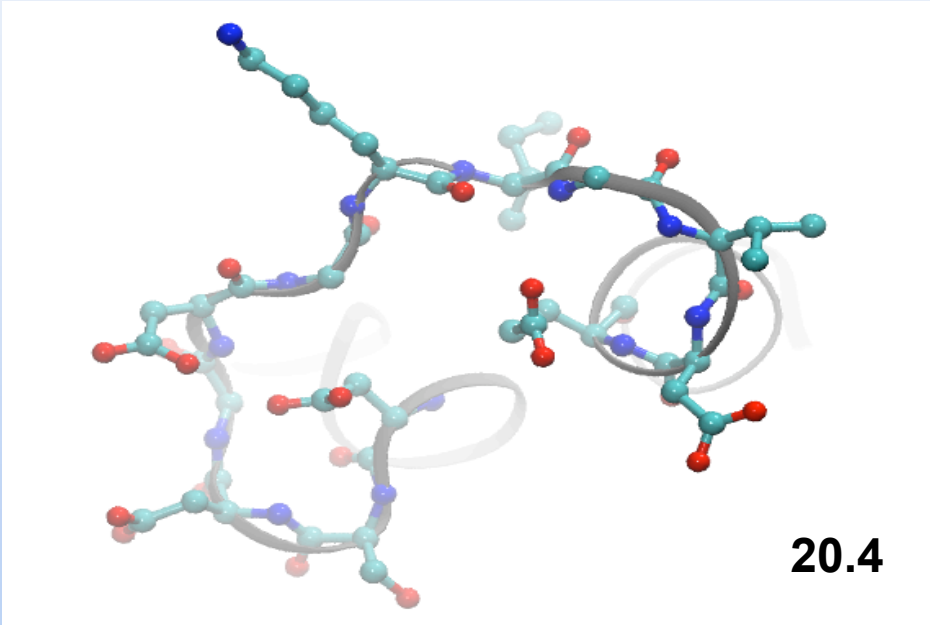
Simulation **Original Score** **Model Threshold**

FEATURE Scores vs. Local Structures

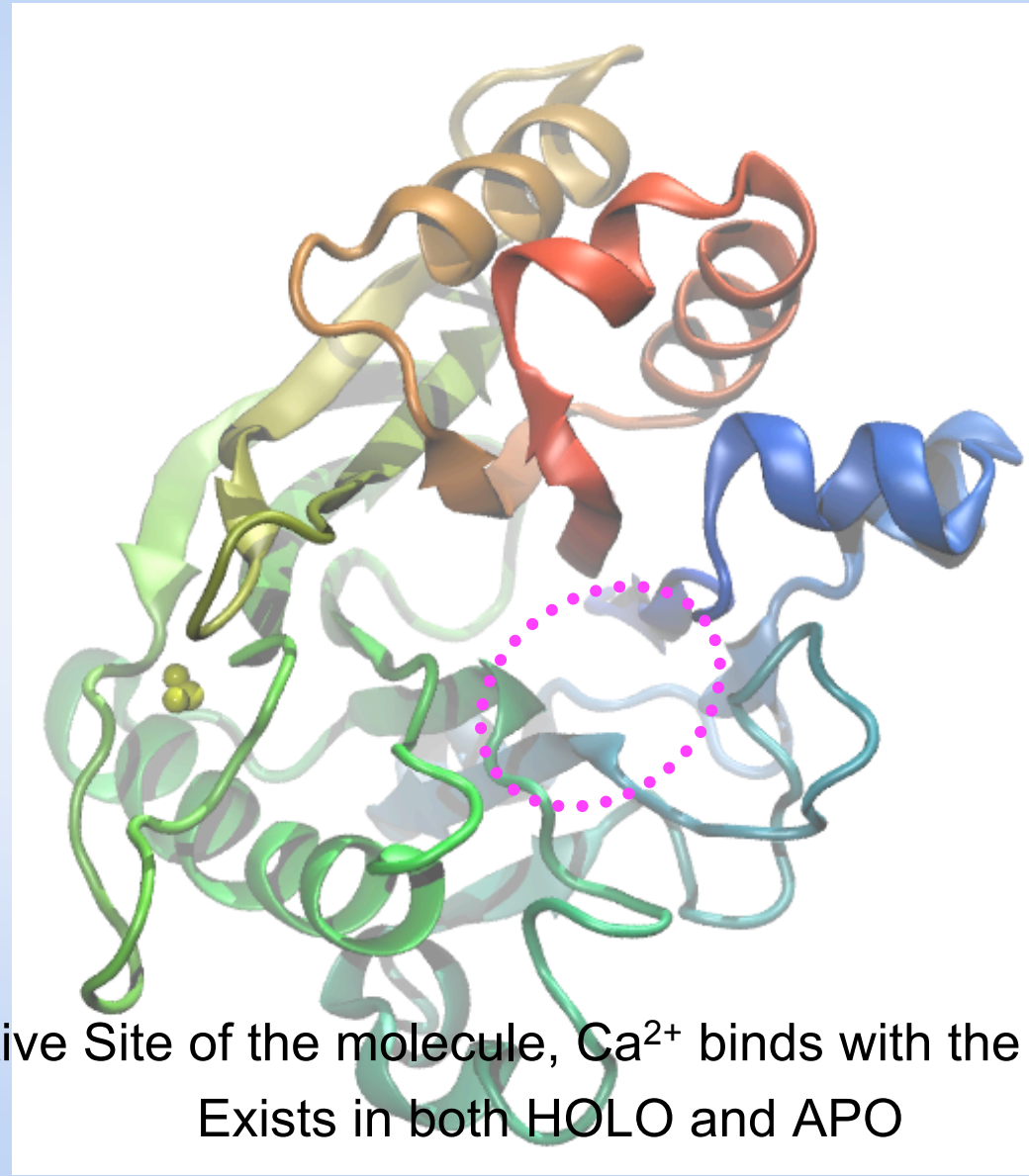


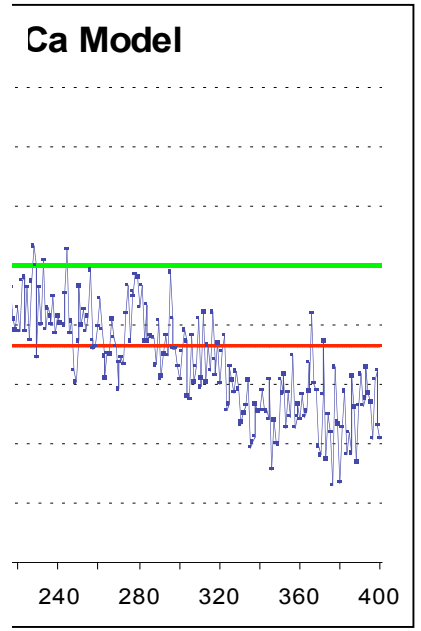






Results: 1DNK





.5ps

ginal Score

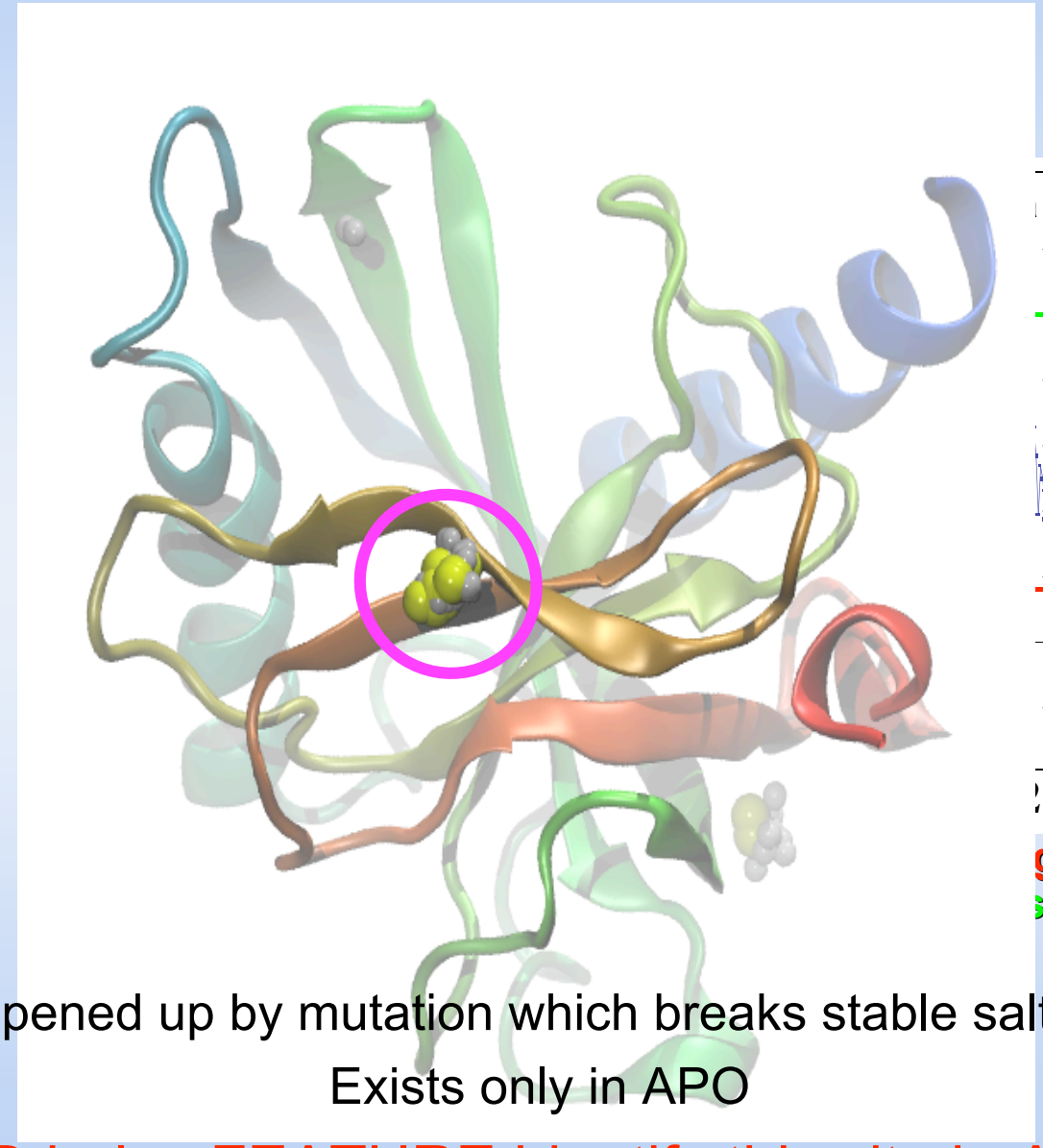
shold

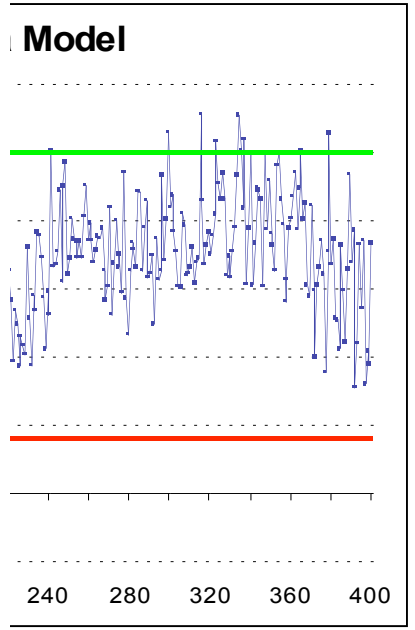
Active Site of the molecule, Ca^{2+} binds with the ligand

Exists in both HOLO and APO

MD helps FEATURE identify this site in both HOLO and APO

Results: 1MJW





.5ps

ginal Score

shold

Site opened up by mutation which breaks stable salt bridge

Exists only in APO

MD helps FEATURE identify this site in APO

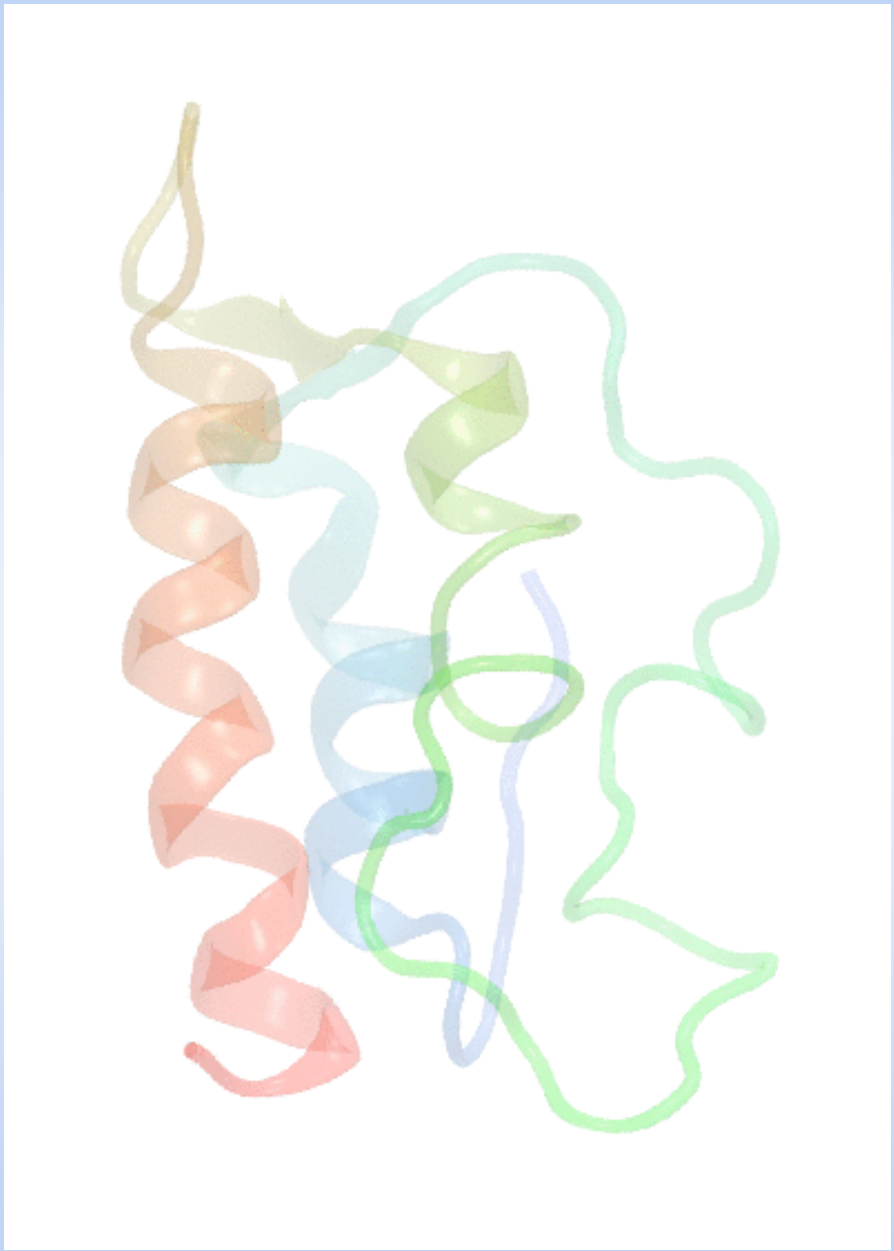
Summary

TOTAL = 12 SITES

Static FEATURE
7 HOLO and 3 APO

MD + FEATURE
10 HOLO and 9 APO

1.4 – 3 Fold Improvement



Protein adopts functional conformations during simulation.

Simbios: National Center for Physics-Based Simulation of Biological Structure

- NIH-funded resource to promote biomedical computing (one of seven)
- Devoted to biological structural simulations at all scales
- Two-fold mission:
  - Perform high quality research
  - Disseminate software and models to others
- Also publish quarterly magazine "Biomedical Computation Review"

NIH Center for Biomedical Computation at Stanford

Home
News
Seminars
People
Collaborating with Simbios
Dissemination & Training
Driving Biological Projects
Biocomputation Research
About Simbios

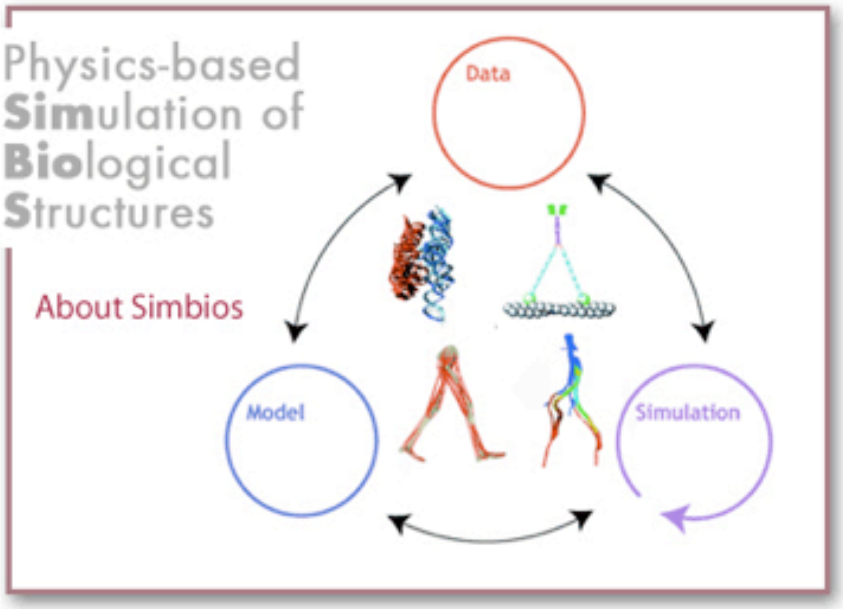
Simbios Sites

Downloads

Simbiome

Biomedical Computation Review

Home



Interested in Collaborating with Simbios?

Highlights of current and upcoming courses, symposiums and workshops

October 25th	Life in Motion Symposium	Clark Center, Stanford University
October 27th	BCATS Symposium	Clark Center, Stanford University

Past Events:

We released OpenSim on August 22nd, view some of the pictures taken at the event.

We held a short course on SimVascular from August 27-31 -- the cardiovascular software tool kit.

Home | About SimTK | How to Contribute

Search Simtk.org: Projects | Go

News | Log In | Create Project | Register

Enabling groundbreaking biomedical research by providing open access to high-quality simulation tools, accurate models and the people behind them.

About SimTK

SimTK, the Simulation Toolkit, is part of the **Simbios** project funded by the National Institutes of Health. **Learn more** ...

Related Sites

NIH Center for Physics-based Simulation

Simbiome

Biomedical Computation Review

Where To Get Downloads

Applications and Models: Free downloadable stand-alone simulation software and models

Core Simulation Technology: Free downloadable source code for the underlying algorithms and computational tools for simulations in a variety of biological application areas

All Projects with Downloads: All available free downloadable software and data

Featured Project

Release of OpenSim 1.0: OpenSim is an open-source software system that lets users develop models of musculoskeletal structures and create dynamic simulations of movement.

Visit the OpenSim Project and download the **new** OpenSim 1.0 release.

Learn more about our latest OpenSim Workshop (8/22/07) on using OpenSim 1.0.

How to Contribute

Biological Application Areas

Biomolecular Simulations:

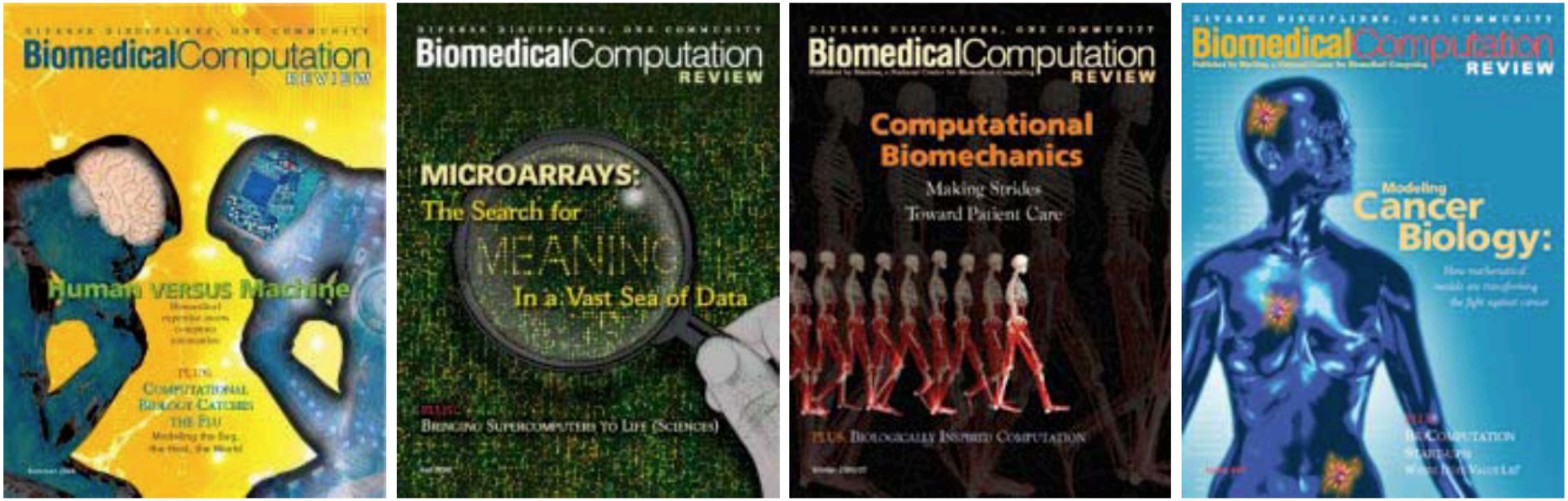
RNA Folding

Myosin Dynamics

Cardiovascular Dynamics

Neuromuscular Biomechanics

Biomedicalcomputationreview.stanford.edu



Covers of the four issues of *Biomedical Computation Review* from this past year.

Conclusions

- We can build statistical models of functional sites in proteins based on 3D structures
- We can use these models to recognize function in new structures
- We can extend the range of our models by adding physics-based simulation
- Challenges:
  - Discovering and labeling new functions
  - Applying in high-throughput to all structures
  - Using for drug design/discovery

Thanks!

russ.altman@stanford.edu

Mike Liang, Shirley Wu, Dariya Glazer,
Inbal Halperin, Jessica Ebert