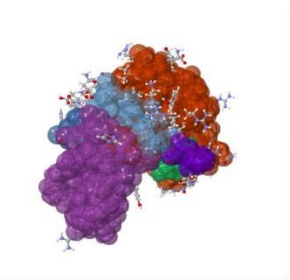
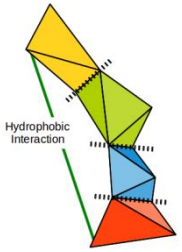
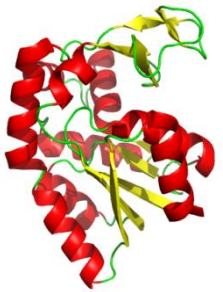
A Combined Molecular Dynamics, Rigidity Analysis Approach for Studying Protein Complexes







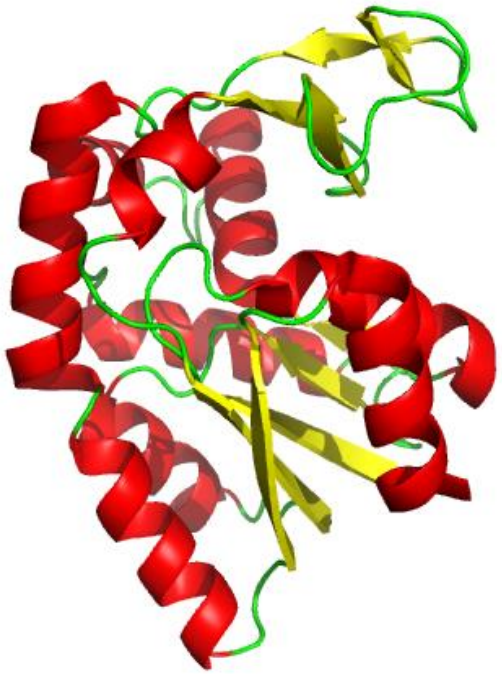
CENTRAL WASHINGTON UNIVERSITY

Brian Orndorff
Dr. Filip Jagodzinski

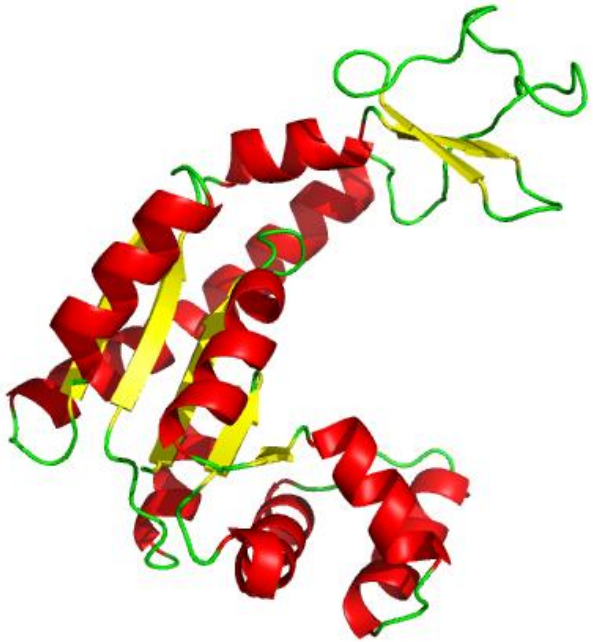
SOURCE
Ellensburg Washington, 15 May 2014

Motivation

Proteins are non-static structures



Barnase
"Closed" Conformation



Barnase
"Open" Conformation

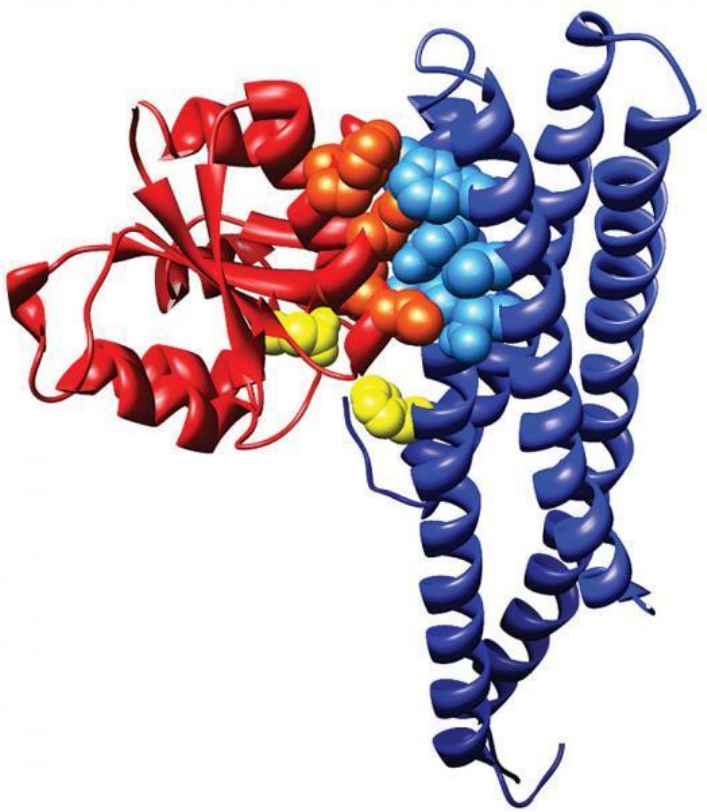
Motivation

Proteins also can exist as complexes, that are composed of two or more associated polypeptide chains.

Motivation

Proteins also can exist as complexes, that are composed of two or more associated polypeptide chains.

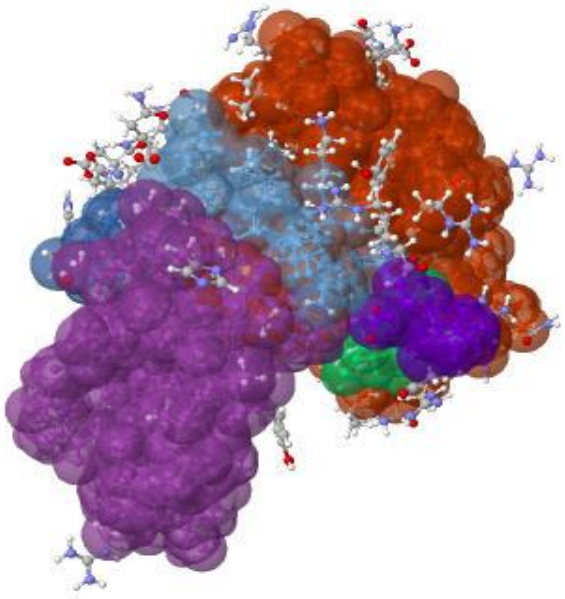
The membrane-bound sensor histidine kinase (shwon in blue) and its response regulator (shown in red) share contact residues (shown as orange and blue spheres).



Our Main Objective: Investigate if (and how) polypeptides in a complex affect each other's flexibilities?

Motivation

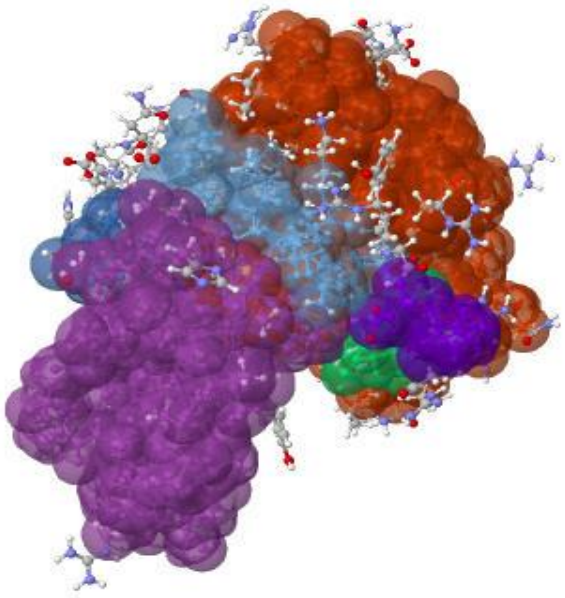
Rigidity analysis of a single Protein crystal structure can be misleading



Rigid cluster decomposition of protein 1LZ1

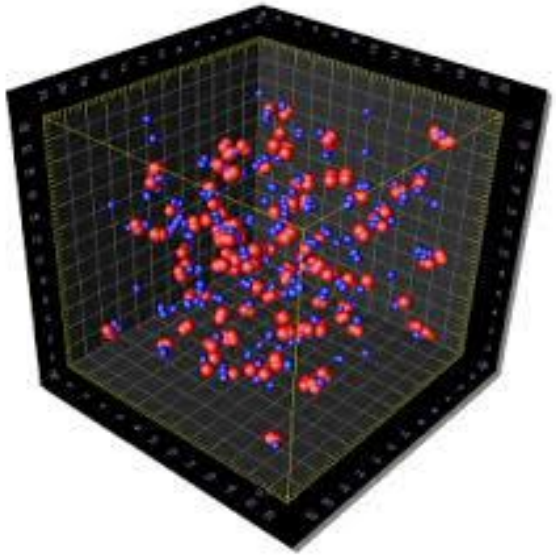
Motivation

Rigidity analysis of a single Protein crystal structure can be misleading



Rigid cluster decomposition of protein 1LZ1

Full-scale Molecular Dynamics simulation runs are computationally intensive



Microsecond (and longer) MD simulations can require thousands of CPUs and/or thousands of wall-clock hours

Our Proposal

Step 1: Use short runs of Molecular Dynamics to generate an ensemble of protein complex conformations

Our Proposal

Step 1: Use short runs of Molecular Dynamics to generate an ensemble of protein complex conformations

Step 2: Analyze the rigidity of a subset of the MD generated conformations

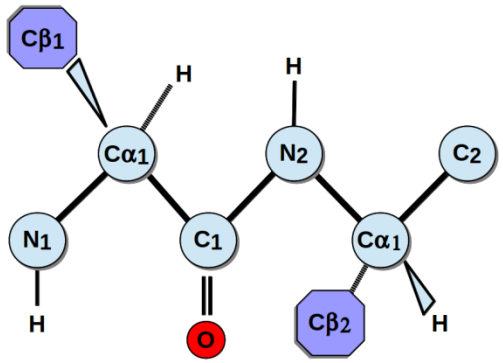
Our Proposal

Step 1: Use short runs of Molecular Dynamics to generate an ensemble of protein complex conformations

Step 2: Analyze the rigidity of a subset of the MD generated conformations

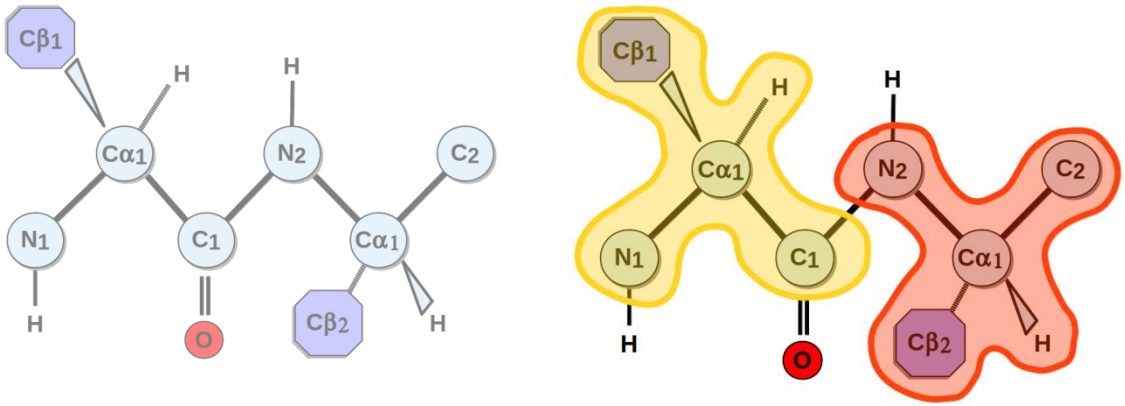
Step 3: Inpsect how the rigidity properties of each conformation in the ensemble of states varies, to infer if (and how) the polypeptides in a complex effect each other's rigidity.

Rigidity Analysis



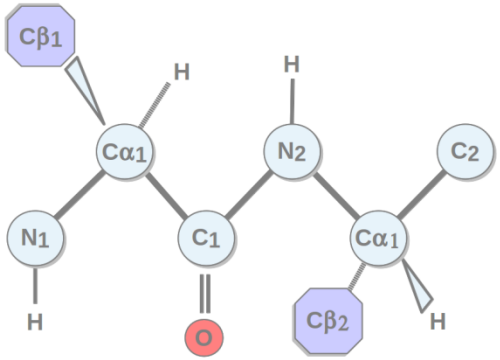
Schematic

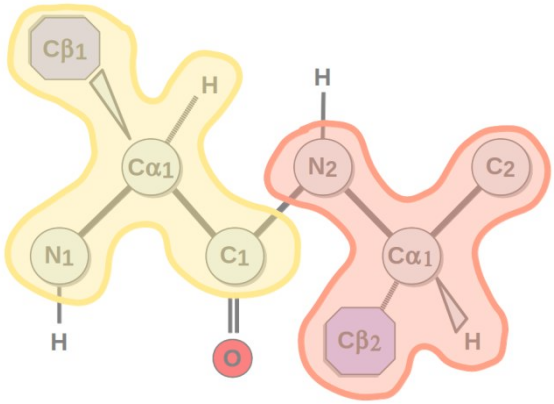
Rigidity Analysis

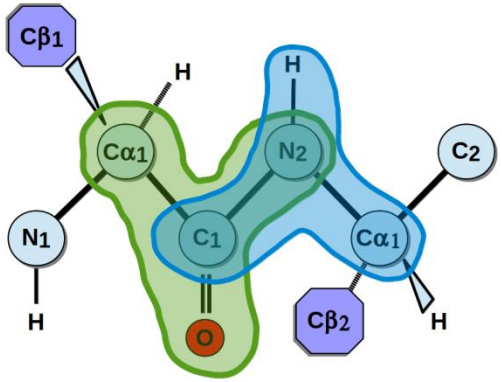


Rigid Bodies

Rigidity Analysis

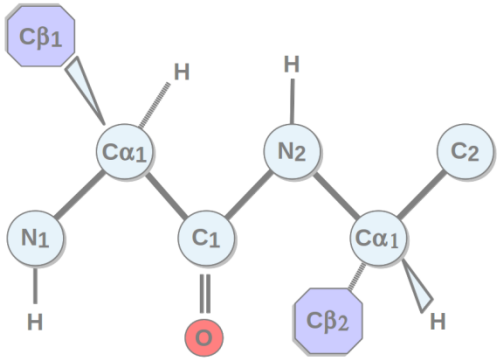


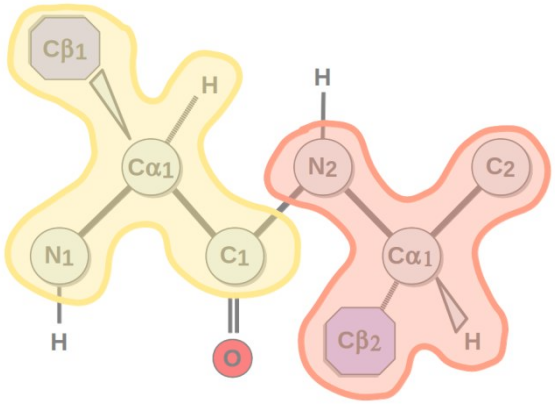


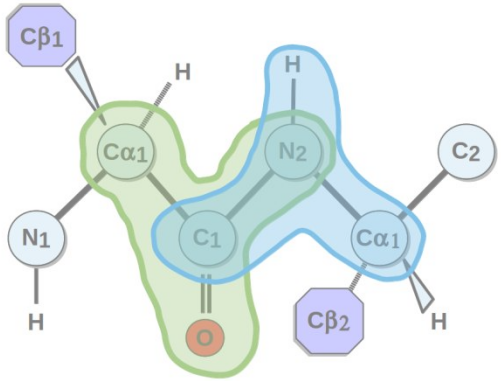


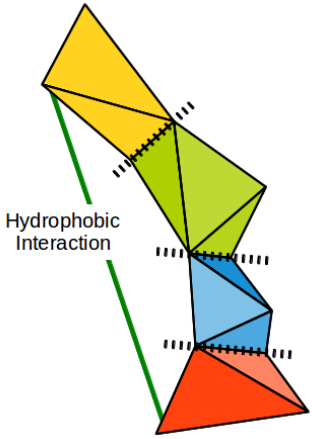
Overlapping Rigid Bodies

Rigidity Analysis



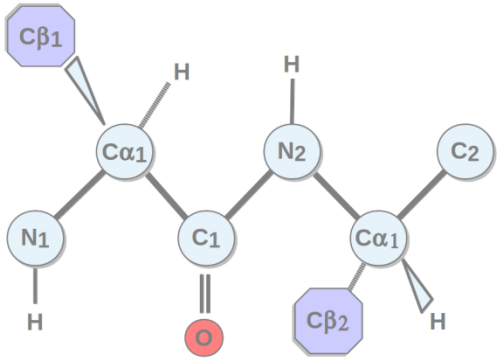


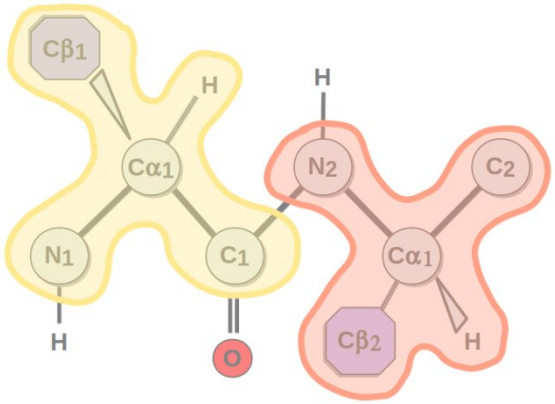


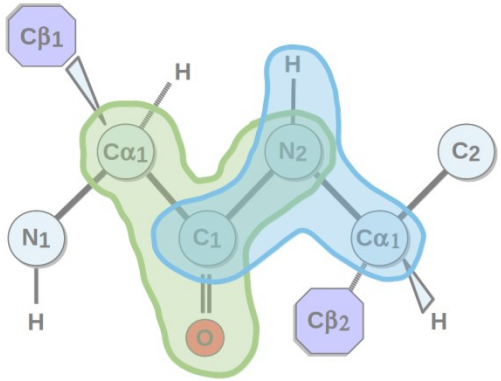


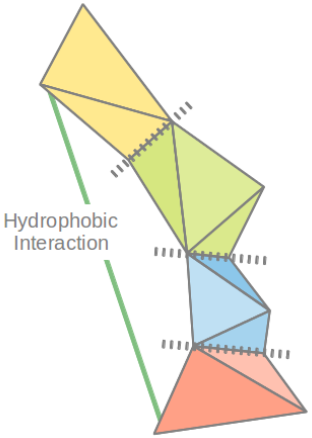
Mechanical Model

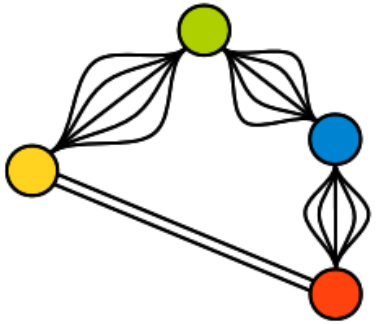
Rigidity Analysis





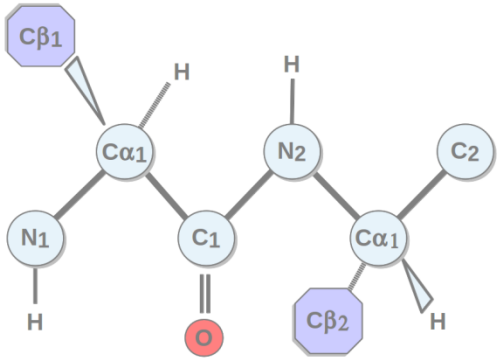


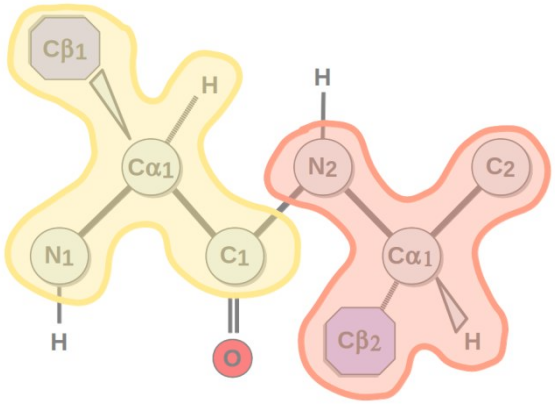


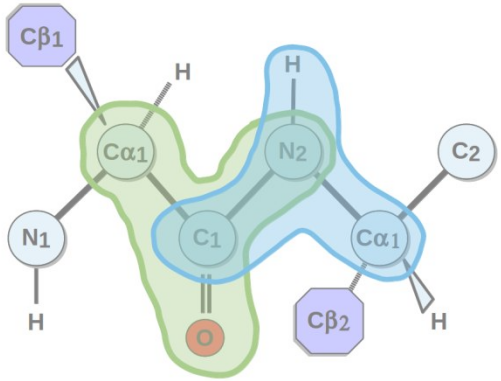


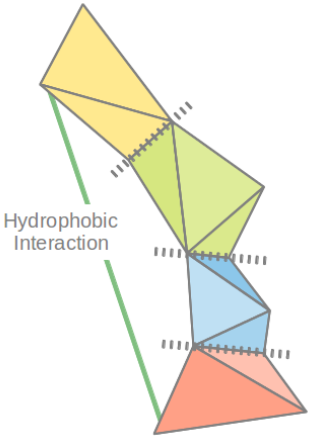
Associated Graph

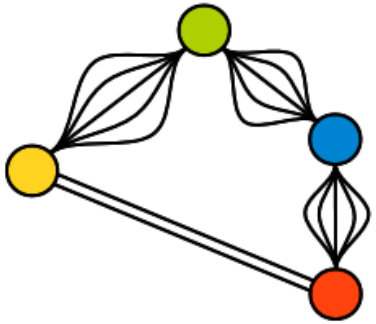
Rigidity Analysis











A pebble game algorithm is used to identify the rigid regions of the associated graph

The algorithm's results are used to infer the rigid and flexible regions of the protein structure

Data Set

9 protein X-ray structure files were obtained from the Protein Data Bank

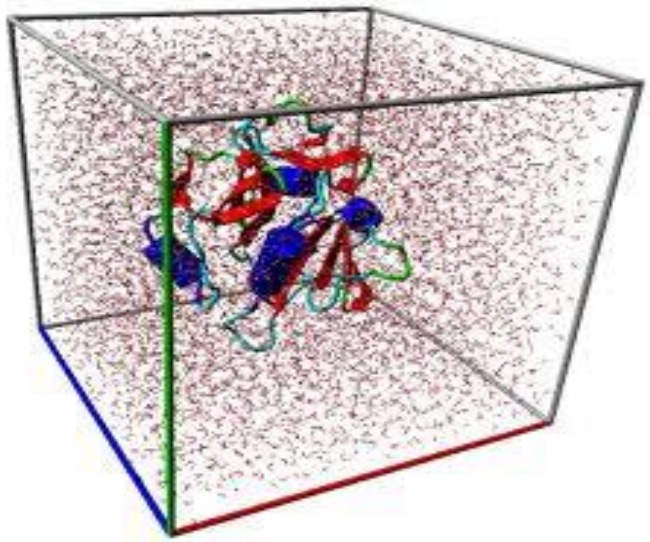
Grouped into 1-chain, 2-chain, and 3-chain sets

PDB	# poly-mers	# chains	chain length(s)	MD run-time	Rigidity Analysis run-time
1lz1	1	1	130	3 min	5 min
4fd9	1	2	92, 92	2 min	5 min
3vb2	1	2	144, 144	4 min	7 min
4bcz	1	2	110, 110	3 min	6 min
3pth	2	2	82, 15	2 min	5 min
3oy6	2	2	253, 23	6 min	4 min
3vz9	2	2	106, 73	1 min	5 min
4g6t	2	2	128, 82	3 min	4 min
4ext	3	3	96, 23, 204	2 min	8 min

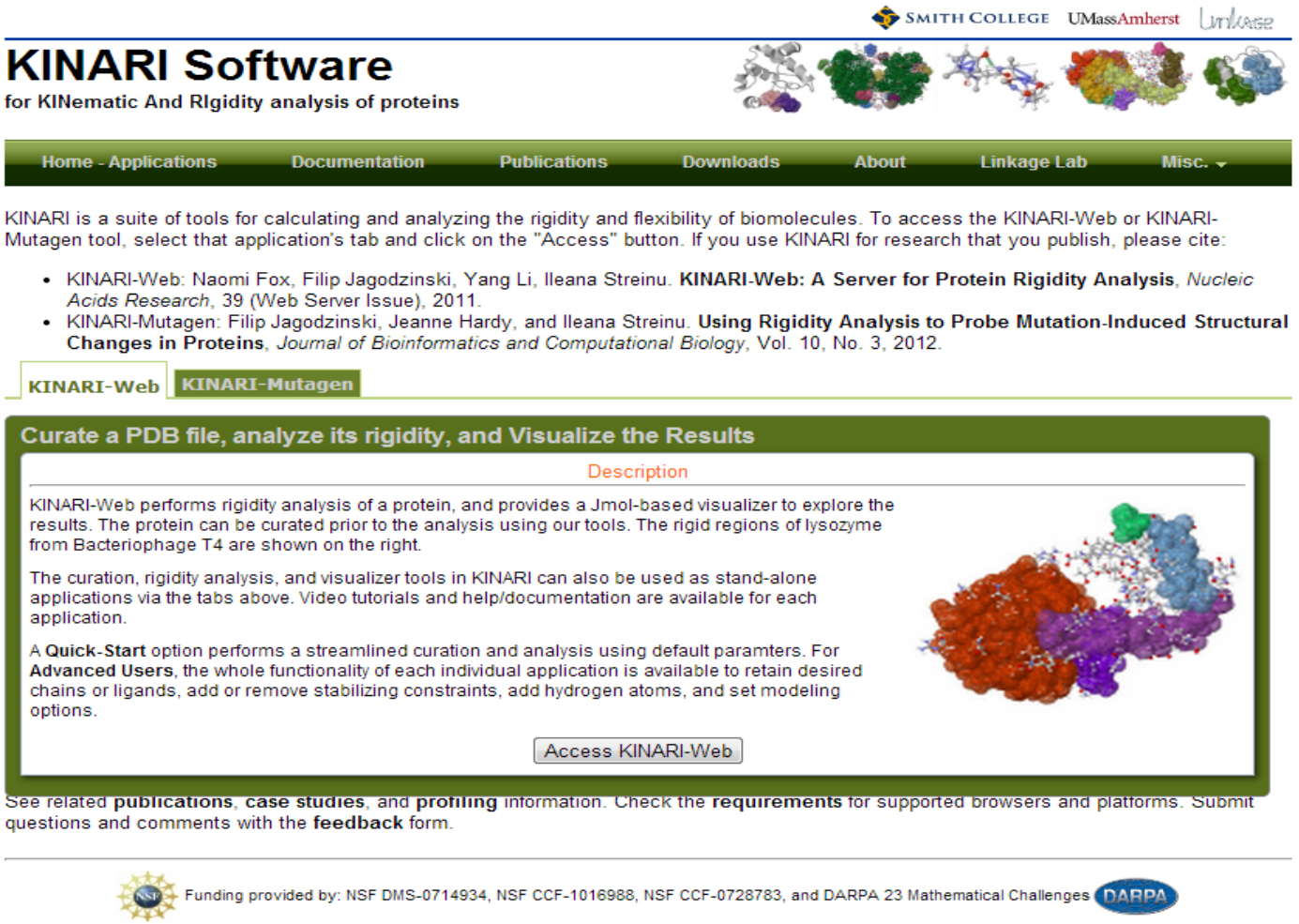
Molecular Dynamics

GROMACS v4.5.5

For each PDB structure, we used 10 conformations randomly selected from a 2 picosecond MD simulation

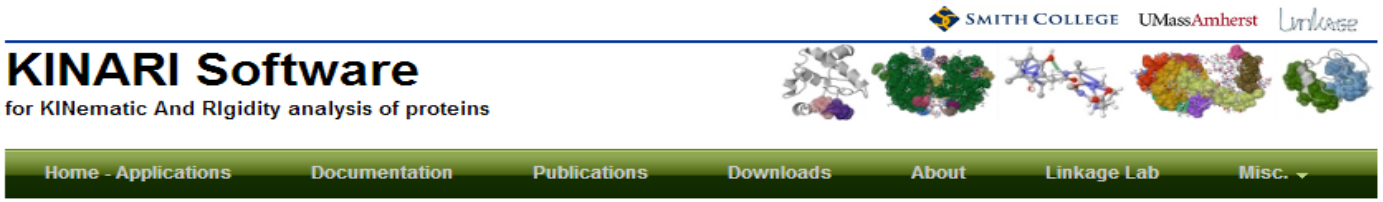


Rigidity Analysis



http://kinari.cs.umass.edu

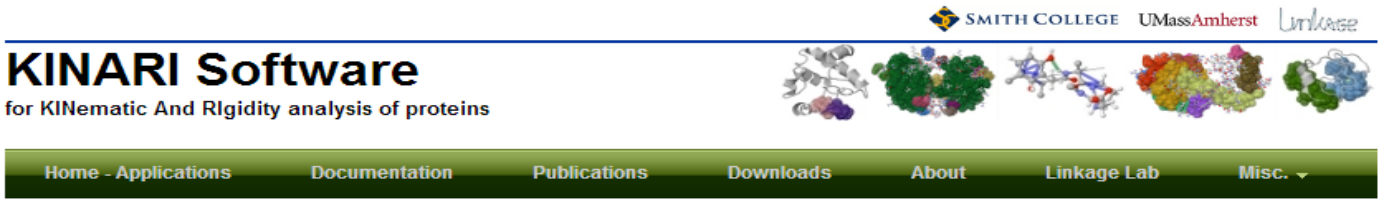
Rigidity Analysis



Default Modeling

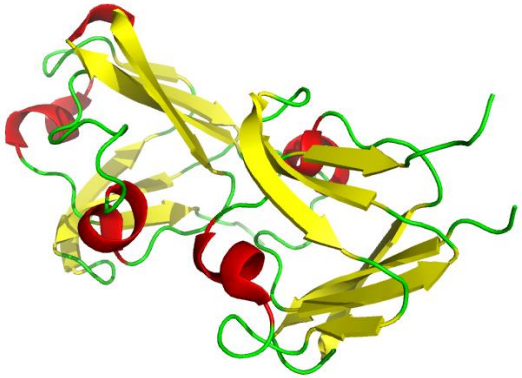
Largest Rigid Cluster (LRC) was recorded for each protein complex conformation

Rigidity Analysis

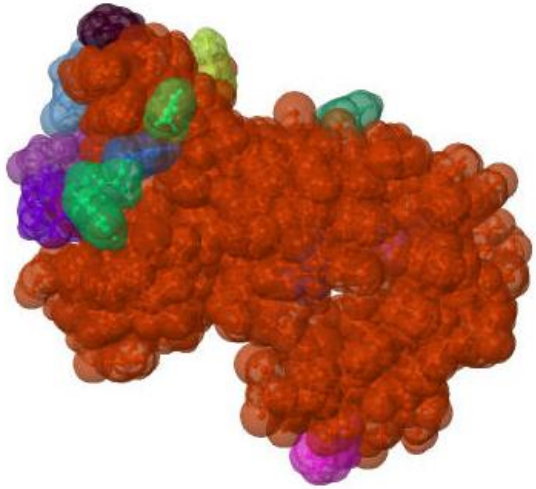


Default Modeling

Largest Rigid Cluster (LRC) was recorded for each protein complex conformation



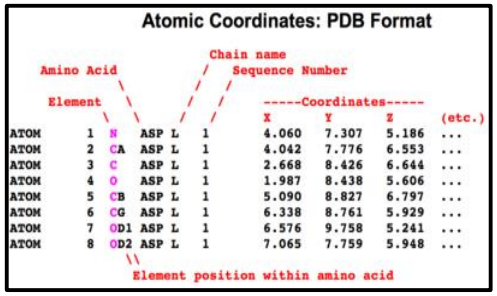
Input protein structure



Calculated Rigid Clusters (same color atoms belong to the same rigid cluster)

Overall Approach

Input



PDB File

Overall Approach

Input

Atomic Coordinates: PDB Format

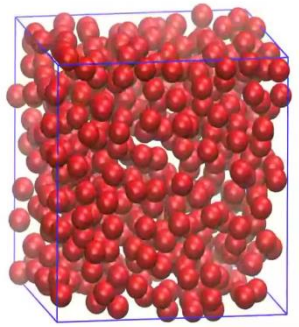
Chain name
Amino Acid / Sequence Number
Element / -----Coordinates-----

						X	Y	Z	(etc.)
ATOM	1	N	ASP	L	1	4.060	7.307	5.186	...
ATOM	2	CA	ASP	L	1	4.042	7.776	6.553	...
ATOM	3	C	ASP	L	1	2.668	8.426	6.644	...
ATOM	4	O	ASP	L	1	1.987	8.438	5.606	...
ATOM	5	CB	ASP	L	1	5.090	8.827	6.797	...
ATOM	6	CG	ASP	L	1	6.338	8.761	5.929	...
ATOM	7	OD1	ASP	L	1	6.576	9.758	5.241	...
ATOM	8	OD2	ASP	L	1	7.065	7.759	5.948	...

Element position within amino acid

PDB File

Generate 10 (only 3 are shown here) conformations using a short MD run



Overall Approach

Input

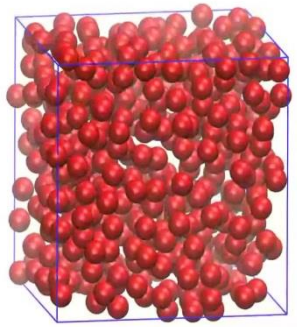
Atomic Coordinates: PDB Format

Chain name / Amino Acid / Sequence Number / Element / -----Coordinates----- / Element position within amino acid

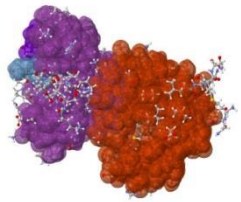
						X	Y	Z	(etc.)
ATOM	1	N	ASP	L	1	4.060	7.307	5.186	...
ATOM	2	CA	ASP	L	1	4.042	7.776	6.553	...
ATOM	3	C	ASP	L	1	2.668	8.426	6.644	...
ATOM	4	O	ASP	L	1	1.987	8.438	5.606	...
ATOM	5	CB	ASP	L	1	5.090	8.827	6.797	...
ATOM	6	CG	ASP	L	1	6.338	8.761	5.929	...
ATOM	7	OD1	ASP	L	1	6.576	9.758	5.241	...
ATOM	8	OD2	ASP	L	1	7.065	7.759	5.948	...

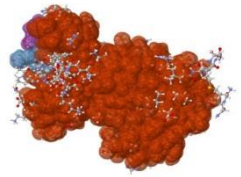
PDB File

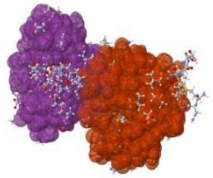
Generate 10 (only 3 are shown here) conformations using a short MD run



Analyze Rigidity of each structure







PDB	# poly-mers	# chains	chain length(s)	MD run-time	Rigidity Analysis run-time
1lz1	1	1	130	3 min	5 min
4fd9	1	2	92, 92	2 min	5 min
3vb2	1	2	144, 144	4 min	7 min
4bcz	1	2	110, 110	3 min	6 min
3pth	2	2	82, 15	2 min	5 min
3oy6	2	2	253, 23	6 min	4 min
3vz9	2	2	106, 73	1 min	5 min
4g6t	2	2	128, 82	3 min	4 min
4ext	3	3	96, 23, 204	2 min	8 min

PDB	# poly-mers	# chains	chain length(s)	MD run-time	Rigidity Analysis run-time
1lz1	1	1	130	3 min	5 min
4fd9	1	2	92, 92	2 min	5 min
3vb2	1	2	144, 144	4 min	7 min
4bcz	1	2	110, 110	3 min	6 min
3pth	2	2	82, 15	2 min	5 min
3oy6	2	2	253, 23	6 min	4 min
3vz9	2	2	106, 73	1 min	5 min
4g6t	2	2	128, 82	3 min	4 min
4ext	3	3	96, 23, 204	2 min	8 min

In each case, the run-time of our combined MD, rigidity analysis approach of our 9 proteins took at most 11 minutes.

Proof of Concept

- Group 1: Range of LRC as % of Max(LRC) is < 30
- Group 2: Range of LRC as % of Max(LRC) is > 30 but < 60
- Group 3: Range of LRC as % of Max(LRC) is greater than 60

PDB ID	Range of size of LRC	Range of LRC as % of Max(LRC)	Group
1lz1	1005	0.67	3
3pth	163	0.41	2
3vb2	827	0.82	3
3vz9	379	0.22	1
4bcz	1408	0.79	3
4fd9	1055	0.55	2
3oy6	228	0.08	1
4g6t	362	0.18	1
4ext	1132	0.41	2

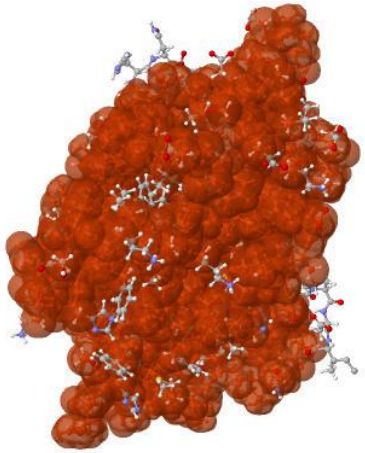
Proof of Concept

In some structures, the rigidity varied **little** across the 10 studied conformations

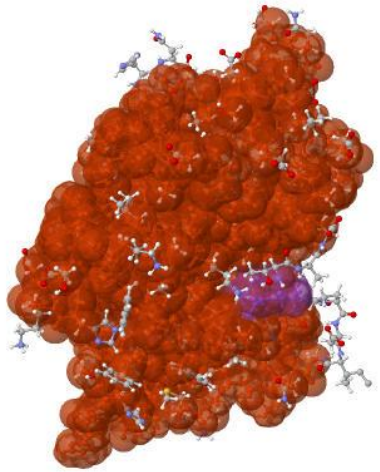
Proof of Concept

In some structures, the rigidity varied **little** across the 10 studied conformations

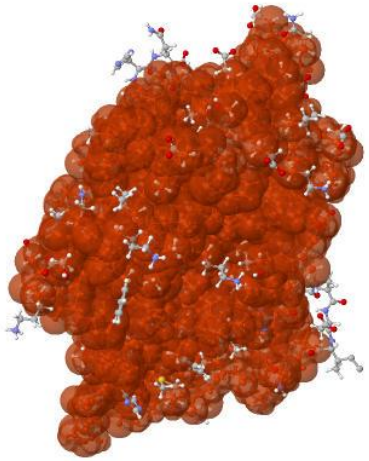
3VZ9 LRC variance : 0.22



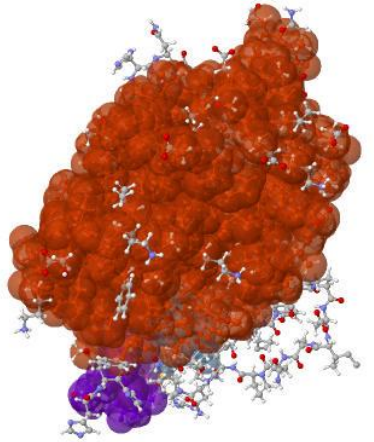
3VZ9 Struct. 3



3VZ9 Struct. 2



3VZ9 Struct. 4



3VZ9 Struct. 5

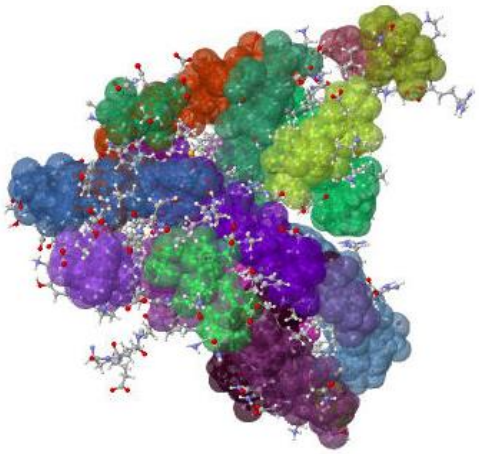
Proof of Concept

In other structures, the rigidity varied **greatly** across the 10 studied conformations

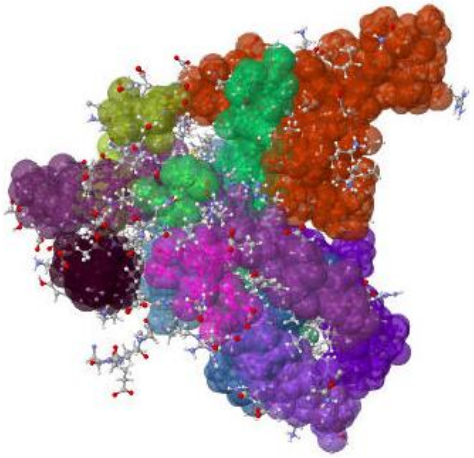
Proof of Concept

In other structures, the rigidity varied **greatly** across the 10 studied conformations

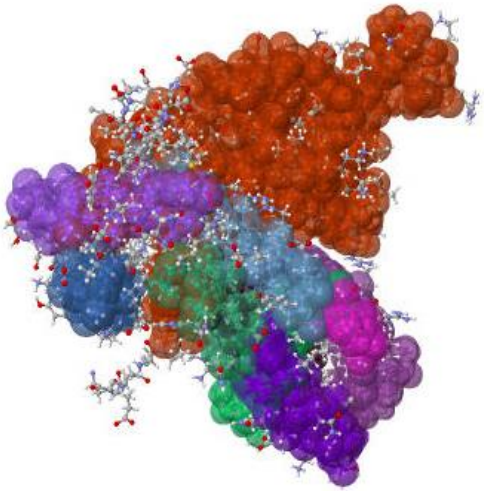
3VB2 LRC variance : 0.82



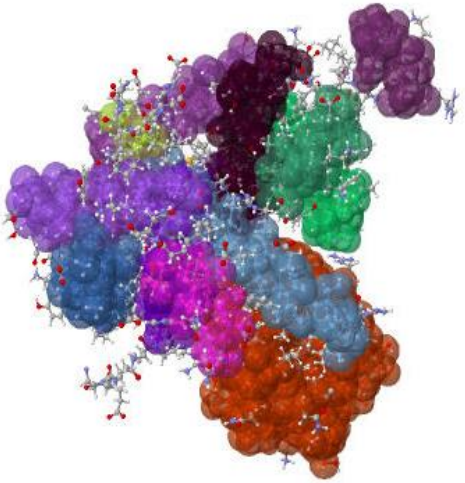
3VB2 Struct. 0



3VB2 Struct. 2



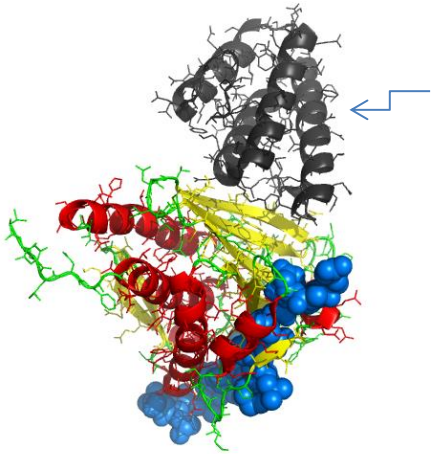
3VB2 Struct. 1



3VB2 Struct. 6

Case Study 1

- Detalied visual analysis of 4EXT

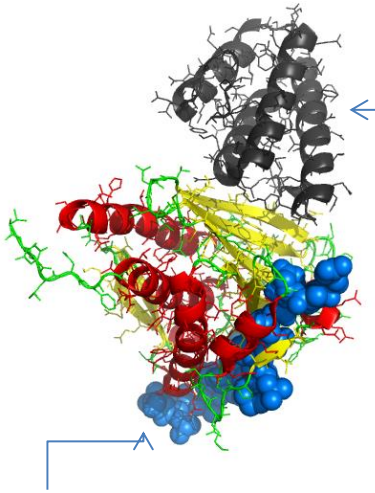


DNA repair protein RVA1 (black) is thought to play a scaffold role in which it recruits DNA polymerase to help repair damaged nucleic acid bases

Lin W, Xin H, Zhang Y, Wu X, Yuan F, Wang Z (Dec 1999). "The human REV1 gene codes for a DNA template-dependent dCMP transferase". Nucleic Acids Res 27 (22): 4468–75. doi:10.1093/nar/27.22.4468. PMC 148731. PMID 10536157.

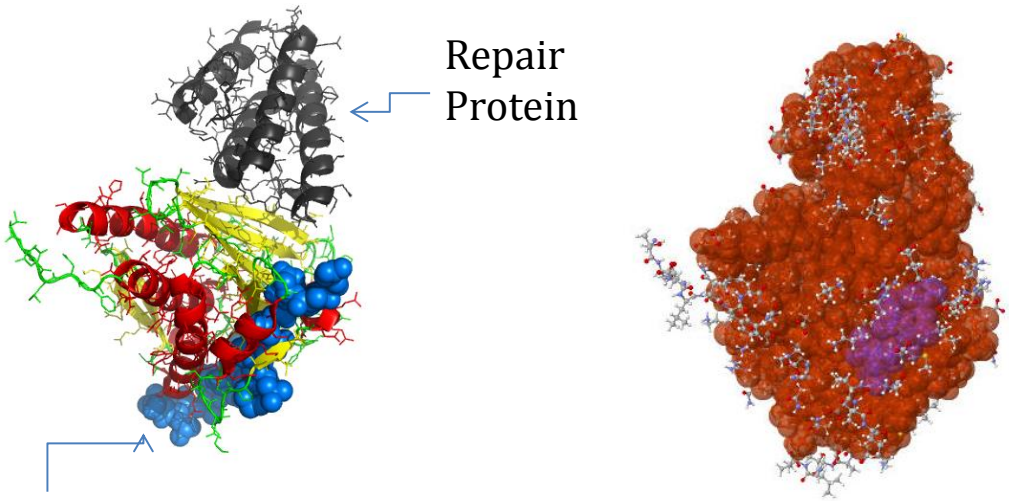
Case Study 1

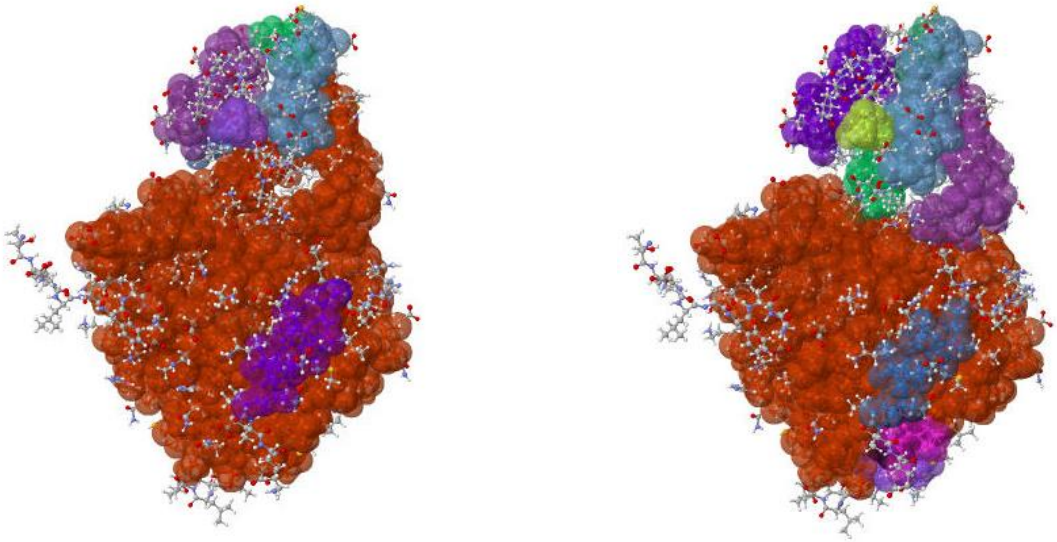
- Detalied visual analysis of 4EXT
- Catalytic subunit (blue) is closely associated with the repair protein (black)



Case Study 1

- Detalied visual analysis of 4EXT
- Catalytic subunit (blue) is closely associated with the repair protein (black)
- The combined MD, rigidity analysis approach provides previously unavailable information

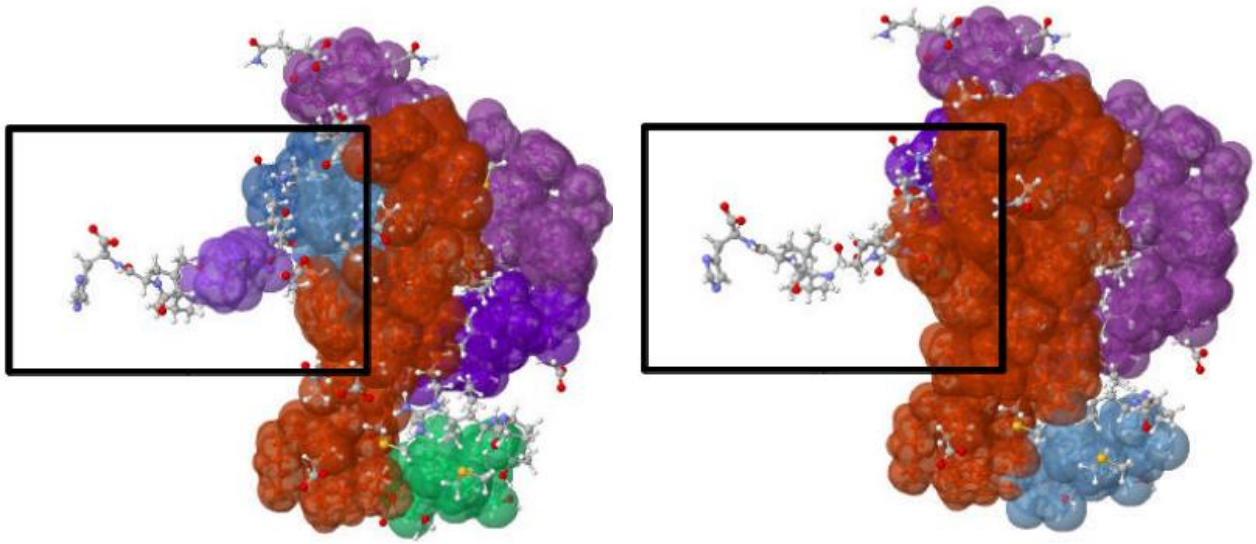




Case Study 2

- Detalied visual analysis of 3PTH

The La family of proteins is characgterized by a nuclear phosphoprotein that recognizes newly synthesized RNA. Structure 3PTH contains a synthetic construct (yet unpublished) coupled with LARP4B.

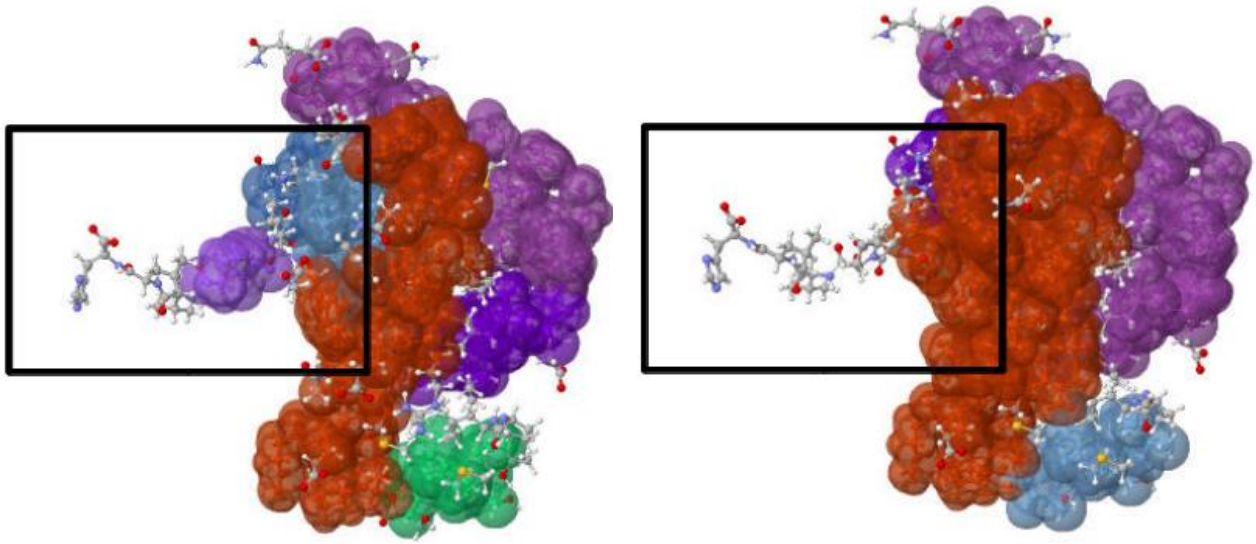


3PTH Struct. 1 3PTH Struct 2

Dong G, Chakshusmathi G, Wolin SL, Reinisch KM. Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. EMBO J. 2004 Mar 10;23(5):1000-7. Epub 2004 Feb 19.

Case Study 2

- Detalied visual analysis of 3PTH
- Ligand has tenuous effect on the protein's rigidity
- Again, the combined MD, rigidity analysis approach provides previously unavailable information



3PTH Struct. 1 3PTH Struct 2

Conclusion

Our combined MD, rigidity analysis approach provides information that neither method (rigidity analysis of single PDB structure, or full-length MD run) alone can provide about the rigidity and flexibility of the studied protein complex

Conclusion

Our combined MD, rigidity analysis approach provides information that neither method (rigidity analysis of single PDB structure, or full-length MD run) alone can provide about the rigidity and flexibility of the studied protein complex

Future Work

Reproduce results for 20, 30, etc. MD generated conformations of each structure, to determine an optimal run-time of MD simulation

Measure complex flexibility using a metric other than Size of Largest Rigid Cluster

Incorporate the entire combined MD, rigidity analysis method into a freely-available online tool

Thank you

Thanks to Dr. Jim Schwing of CWU's Computer Science Department and Dean Kirk Johnson of CWU's College of the Sciences for funding

jagodzinski@cwu.edu
orndorffb@cwu.edu