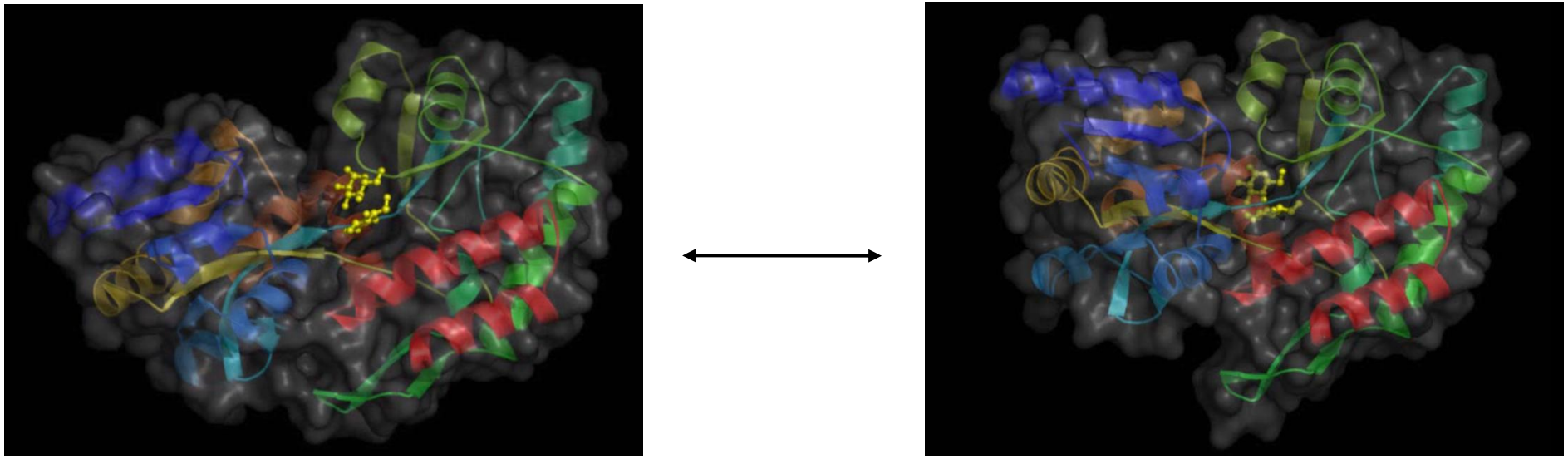
Analyzing Protein Flexibility

An Introduction to Combinatorial Rigidity Methods and Their Applications

Linkage Laboratory
Department of Computer Science
University of Massachusetts Amherst
and Smith College
Prof. Ileana Streinu
Naomi Fox
Filip Jagodzinski

Proteins Are Dynamic Molecular Machines

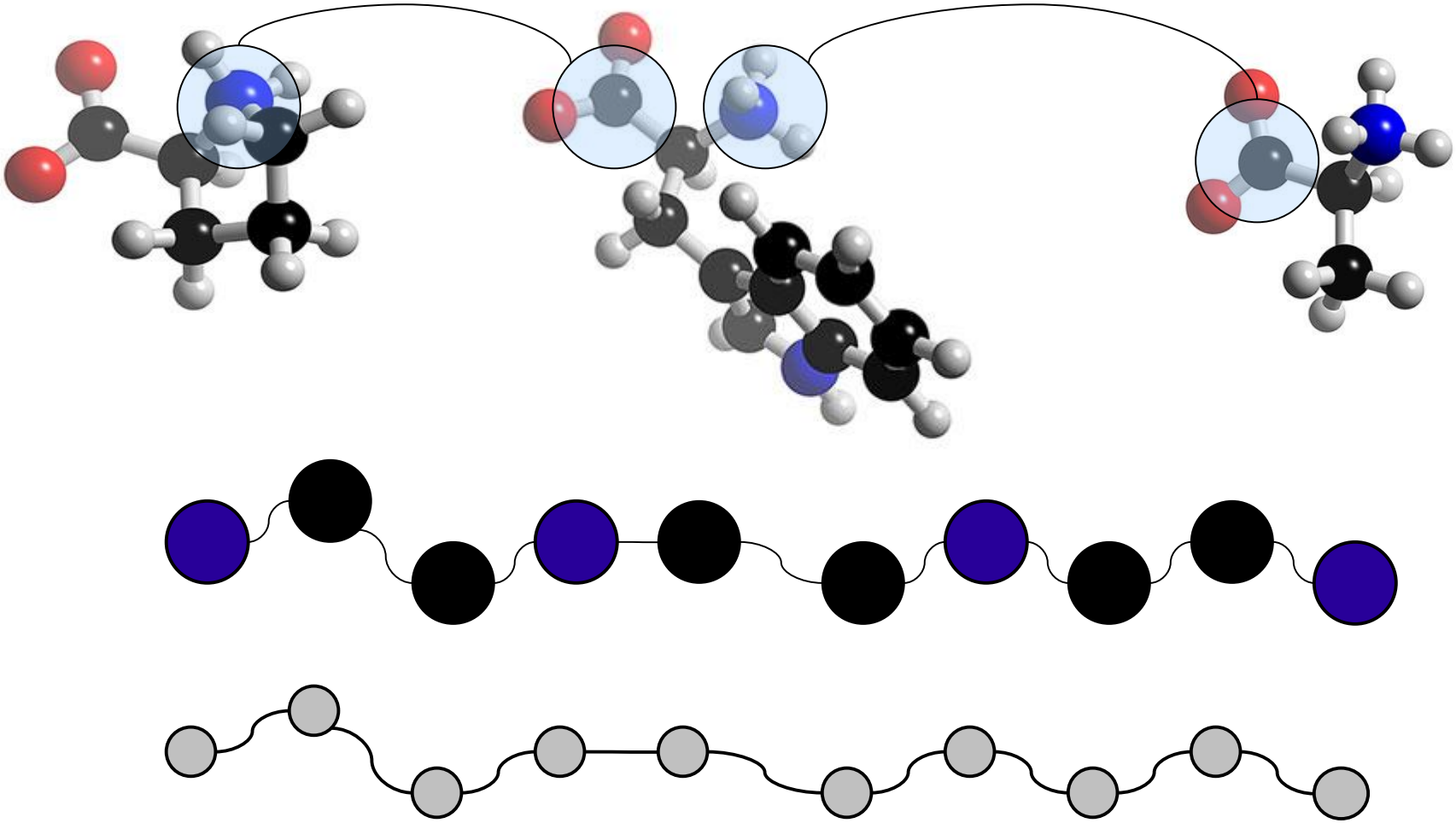


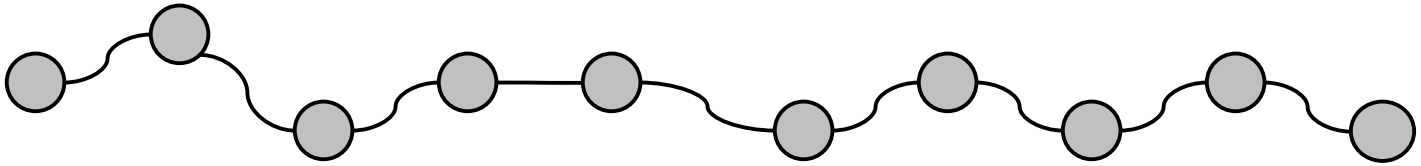
Maltose-binding protein engulfing ligands

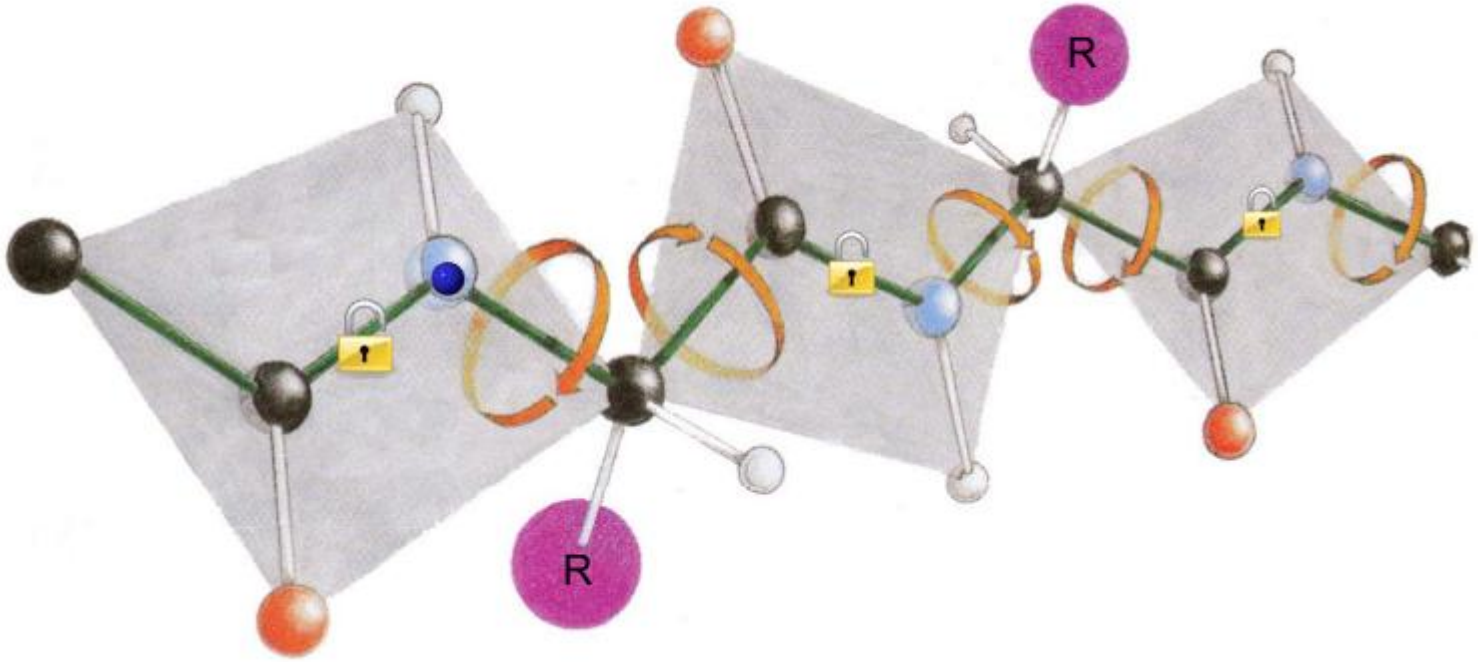
Flores, Echols, Milburn, Hespenheide, Keating, Lu, Wells, Yu, Thorpe, Gerstein (2006). Nucleic Acids Res. 34:D296-301.

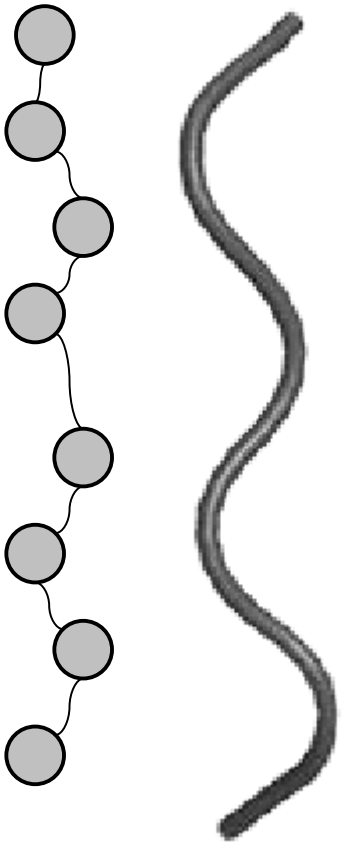
Protein's are ...

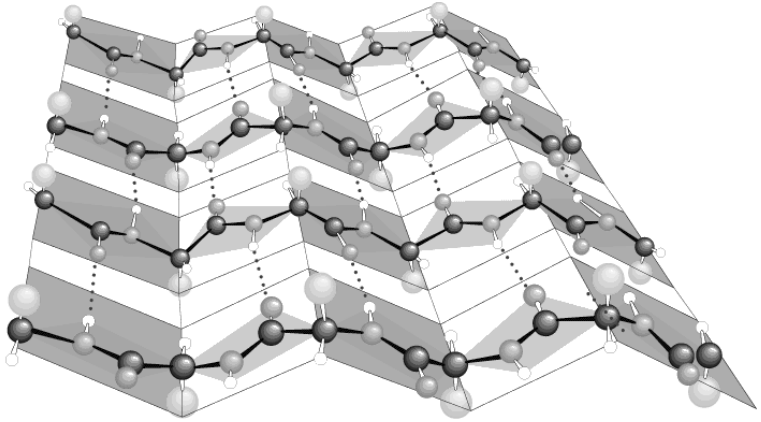
Proline Tryptophan Alanine

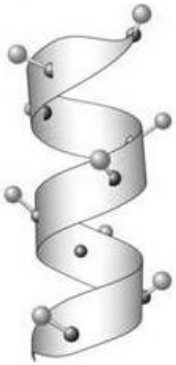




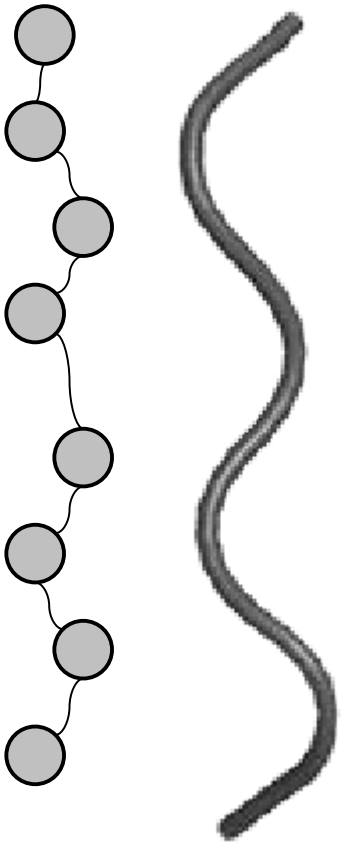

R
R

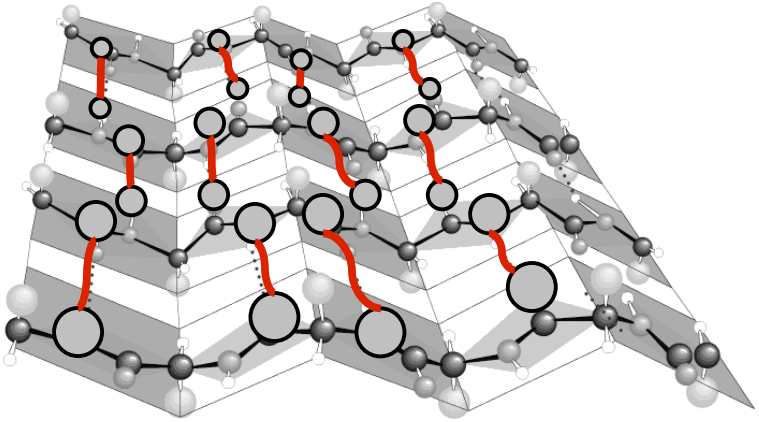


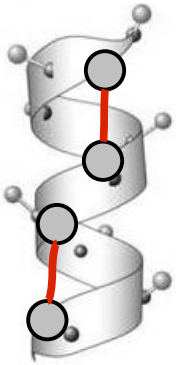




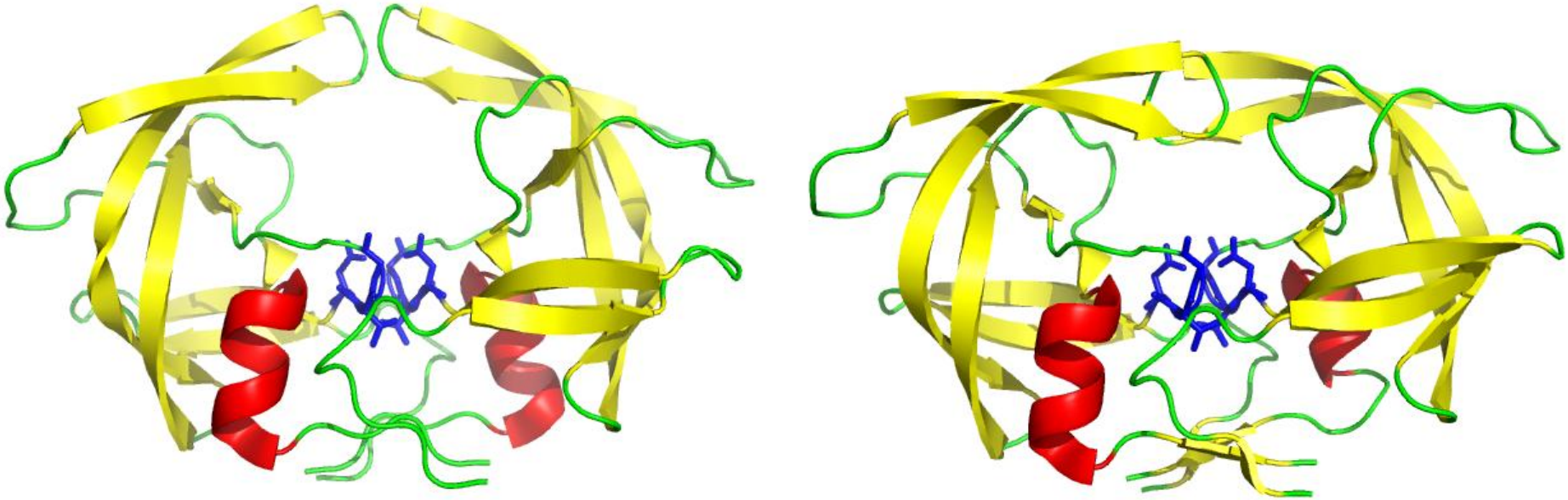
Protein Secondary Structure





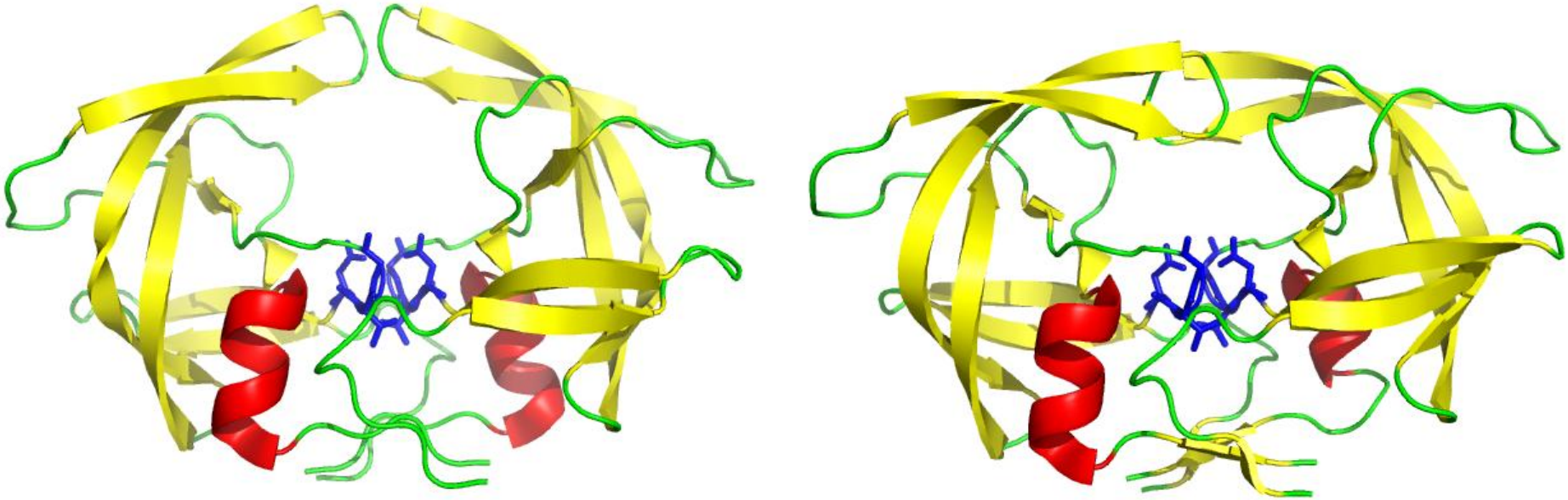


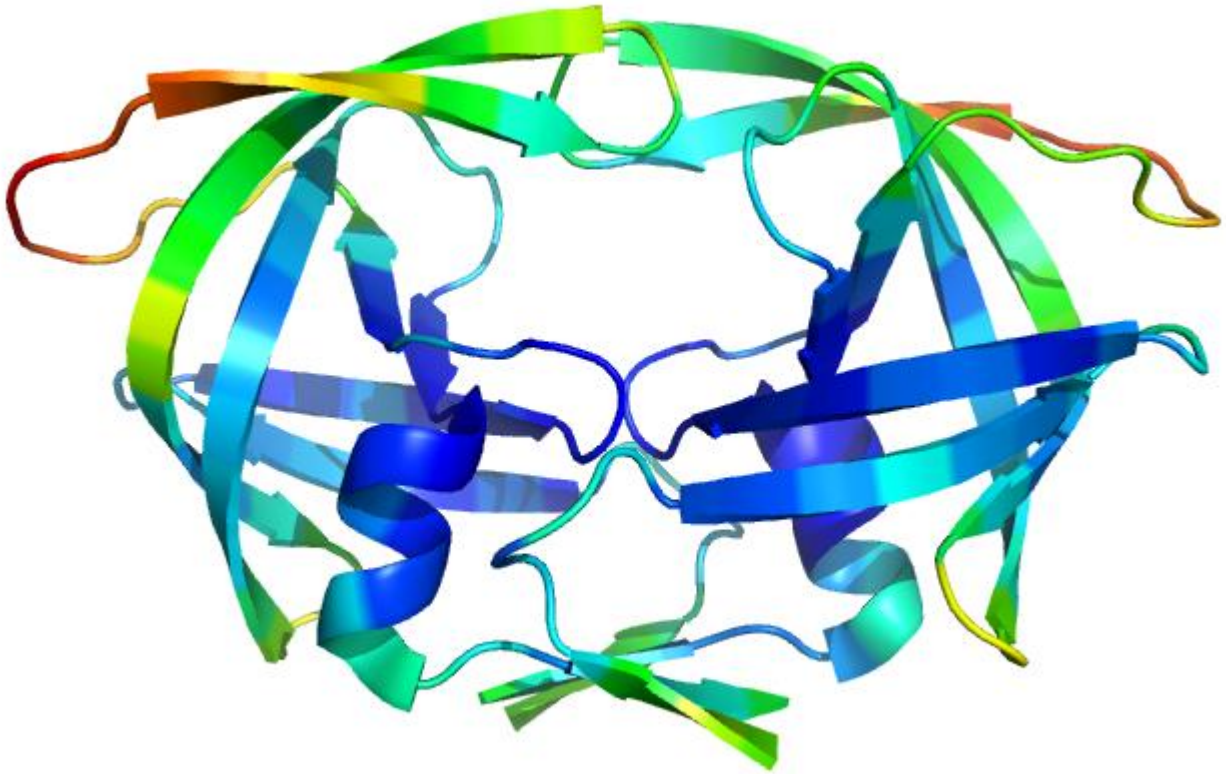
Proteins are dynamic structures



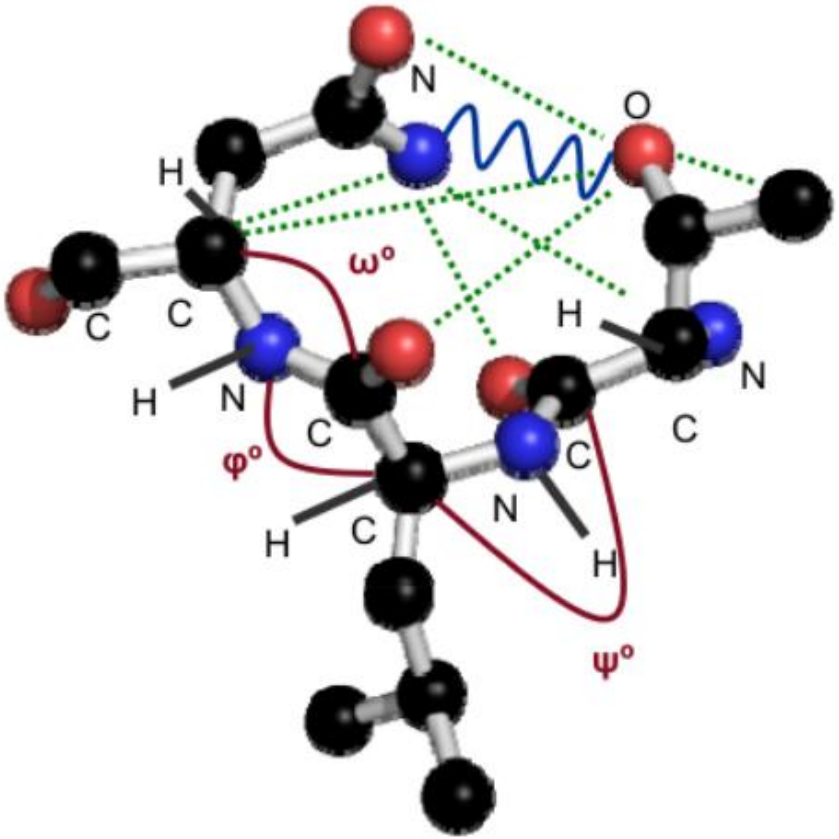
Proteins are dynamic structures

Problem: We cannot observe proteins on the atomic level or their motions directly, but we'd like to understand how they flex or bend.



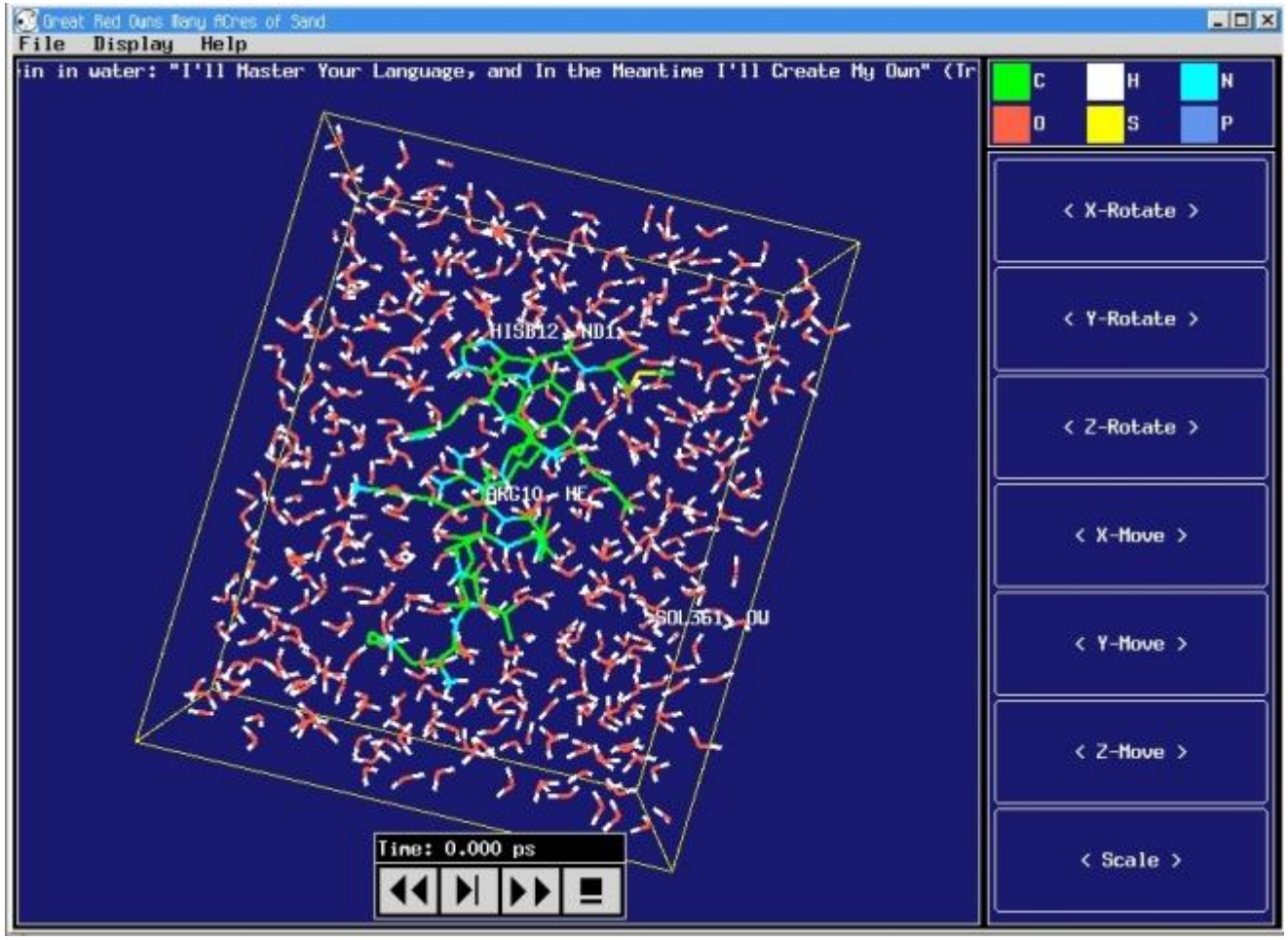


Molecular Dynamics (MD) Simulations



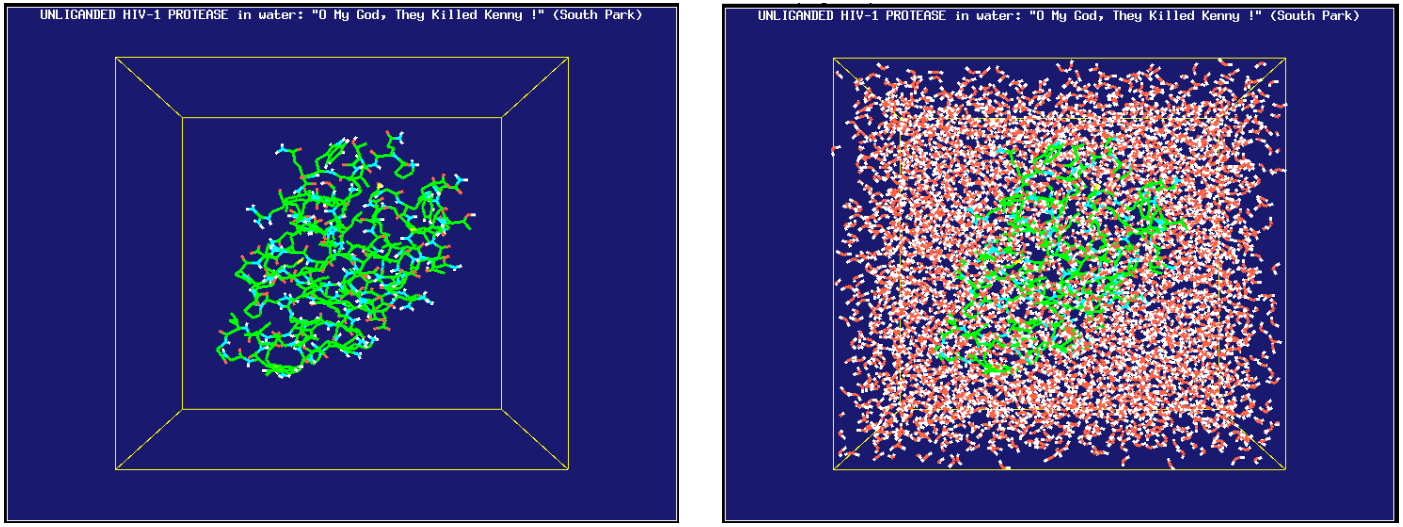
Molecular Dynamics provides the means to solve the equations of particle motion and evaluate these mathematical formulas

Molecular Dynamics Simulation - Demonstration



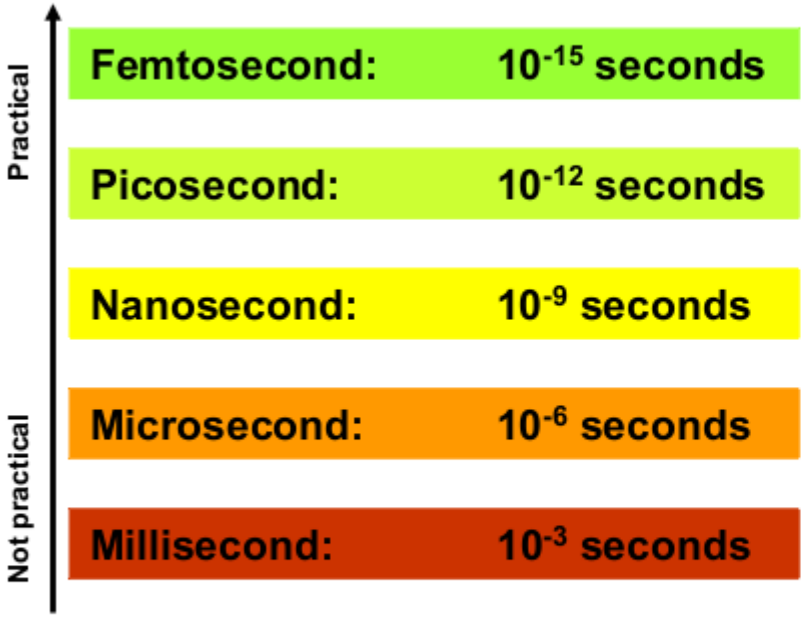
GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation Hess, B., Kutzner, C., van der Spoel, D. and Lindahl, E. J. Chem. Theory Comput., 4, 435-447 (2008)

- Demonstration of Rigidity Analysis
- Introduction to Rigidity Theory in 2-dimensions
- Introduction to Rigidity Theory in 3-dimensions
- Break
- Molecular Modeling and Rigidity Analysis using KINARI
- Potential Applications and ongoing research using Rigidity Analysis
- Concluding Remarks



Check on Molecular Dynamics Simulation

Molecular Dynamics Simulations – Good for?



Local Motions (0.01 to 5 Å, 10^{-15} to 10^{-1} s)

- Atomic Fluctuations
- Sidechain Motions
- Loop Motions

Rigid Body Motions (1 to 10Å, 10^{-9} to 1s)

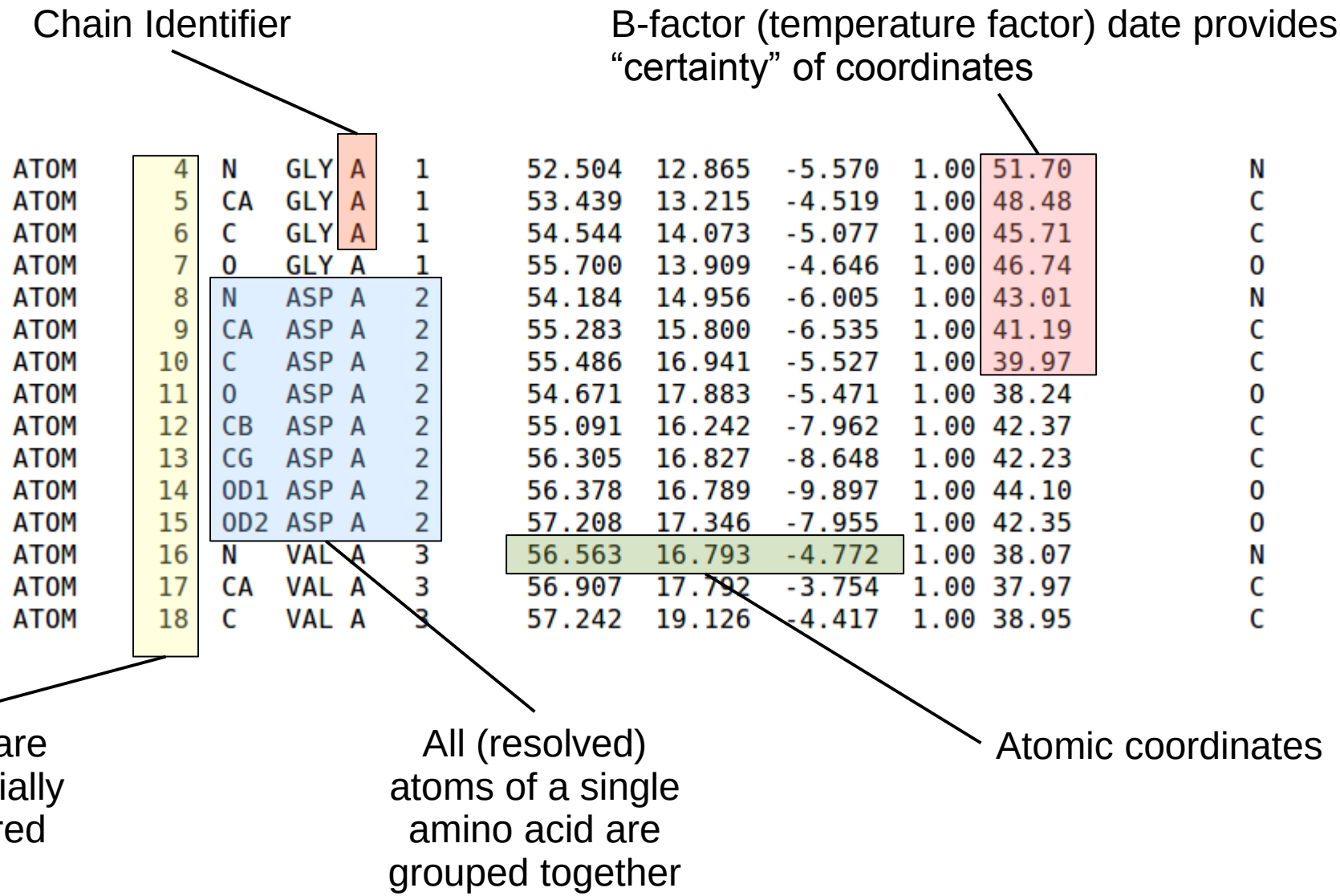
- Helix Motions
- Domain Motions
- Subunit Motions

Large-Scale Motions (> 5Å, 10^{-7} to 10^{4} s)

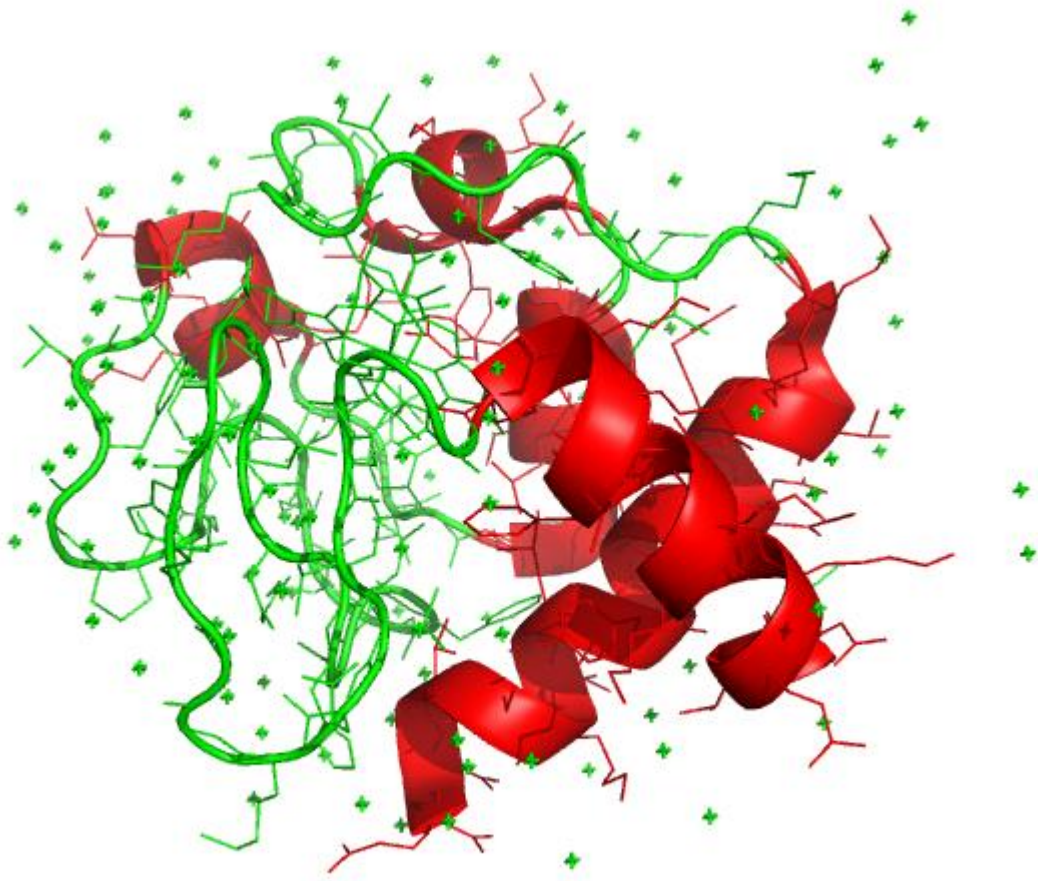
- Helix coil transitions
- Dissociation/Associations
- Folding and unfolding

Part 1:
Demonstrations of
Using Rigidity Analysis

A PDB file ...


Chain Identifier
B-factor (temperature factor) date provides "certainty" of coordinates
ATOM 4 N GLY A 1 52.504 12.865 -5.570 1.00 51.70 N
ATOM 5 CA GLY A 1 53.439 13.215 -4.519 1.00 48.48 C
ATOM 6 C GLY A 1 54.544 14.073 -5.077 1.00 45.71 C
ATOM 7 O GLY A 1 55.700 13.909 -4.646 1.00 46.74 O
ATOM 8 N ASP A 2 54.184 14.956 -6.005 1.00 43.01 N
ATOM 9 CA ASP A 2 55.283 15.800 -6.535 1.00 41.19 C
ATOM 10 C ASP A 2 55.486 16.941 -5.527 1.00 39.97 C
ATOM 11 O ASP A 2 54.671 17.883 -5.471 1.00 38.24 O
ATOM 12 CB ASP A 2 55.091 16.242 -7.962 1.00 42.37 C
ATOM 13 CG ASP A 2 56.305 16.827 -8.648 1.00 42.23 C
ATOM 14 OD1 ASP A 2 56.378 16.789 -9.897 1.00 44.10 O
ATOM 15 OD2 ASP A 2 57.208 17.346 -7.955 1.00 42.35 O
ATOM 16 N VAL A 3 56.563 16.793 -4.772 1.00 38.07 N
ATOM 17 CA VAL A 3 56.907 17.792 -3.754 1.00 37.97 C
ATOM 18 C VAL A 3 57.242 19.126 -4.417 1.00 38.95 C
Atoms are Sequentially numbered
All (resolved) atoms of a single amino acid are grouped together
Atomic coordinates

Cytochrome C

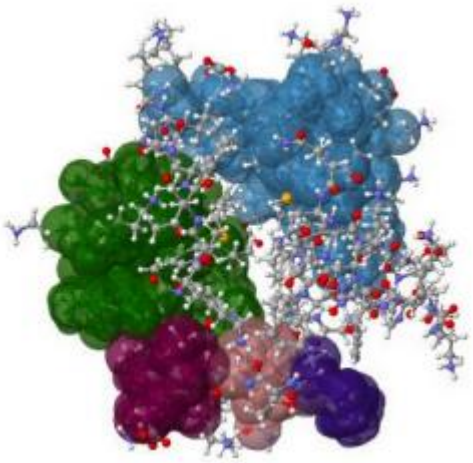


Horse heart cytochrome c (PDB file 1HRC) is a 105 residue heme protein found loosely associated with the inner membrane of mitochondria

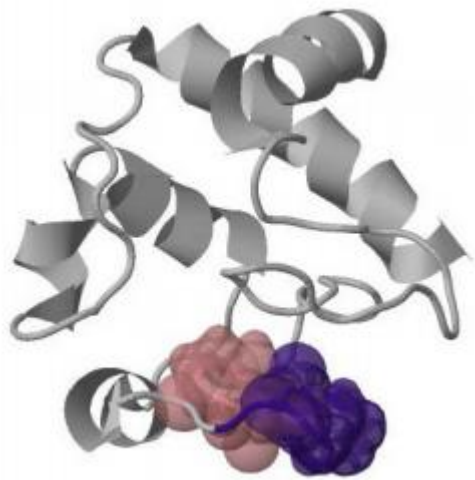
What features of KINARI-Web can we use to determined/investigate its rigidity?

Where are the rigid regions? Hinges? Clusters?

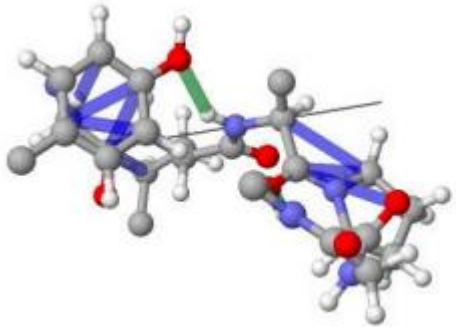
Cytochrome C – Rigidity Results



Default curation and modeling options display the largest rigid clusters as highlighted surfaces in a bar-and-stick model.

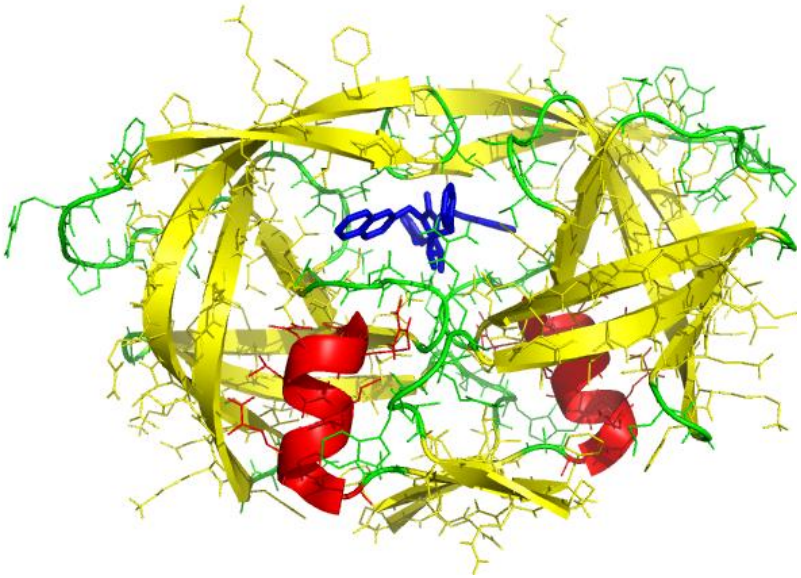


The same protein can be viewed as a cartoon

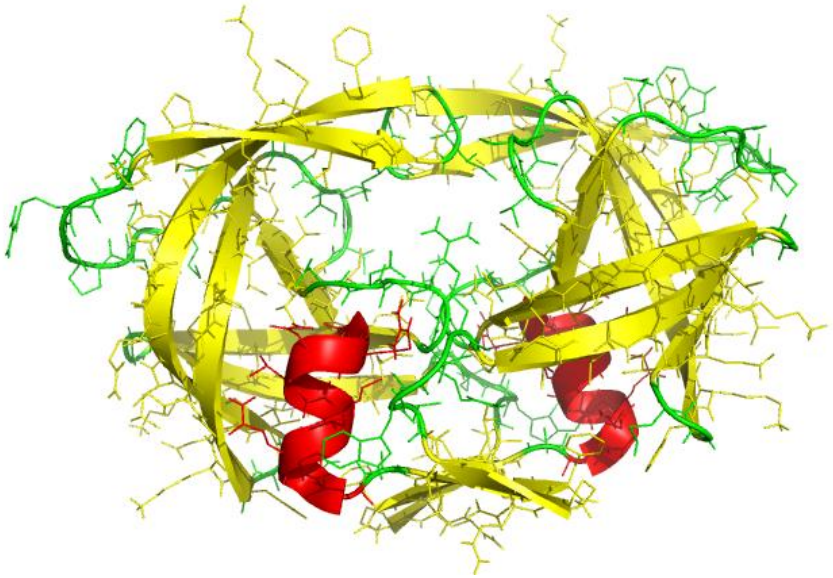


A hinge between two bodies is easily seen when the protein is displayed using a bar-and-stick representation

KINARI-Web: A Server for Protein Rigidity Analysis *Naomi Fox*, *Filip Jagodzinski*, *Yang Li*, *Ileana Streinu*
Nucleic Acids Research, **39** (Web Server Issue), 2011.

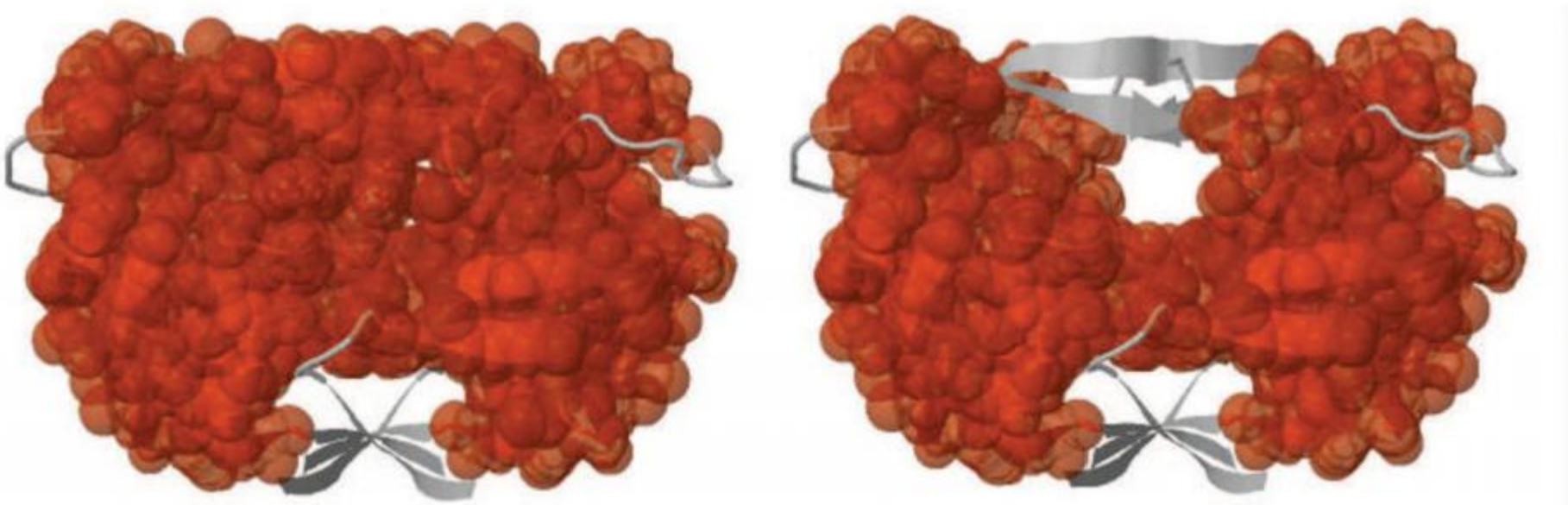


HIV-1 Protease (1HVR) with the ligand (blue)



HIV-1 Protease (1HVR) without the ligand

Can we use rigidity analysis to infer affect of the ligand on the protease's rigidity?



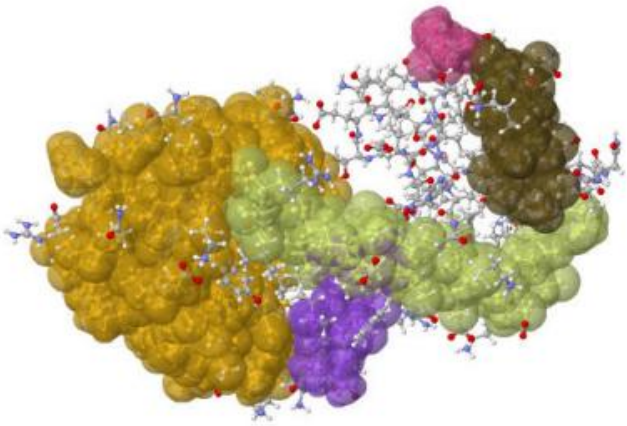
When the ligand is retained in the molecular model of HIV-1 protease (left), the flaps of the dimer are held rigidly in place. When the ligand is remove (right) the flaps are flexible.

KINARI-Web: A Server for Protein Rigidity Analysis *Naomi Fox*, *Filip Jagodzinski*, *Yang Li*, *Ileana Streinu*
Nucleic Acids Research, **39** (Web Server Issue), 2011.

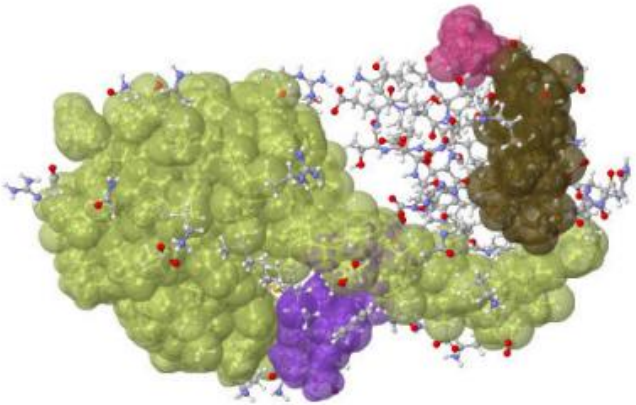
Lysozyme from bactriophage T4

Lysozyme is part of the immune response system; It helps to break apart bacterial cellular walls, thus killing the bacteria.

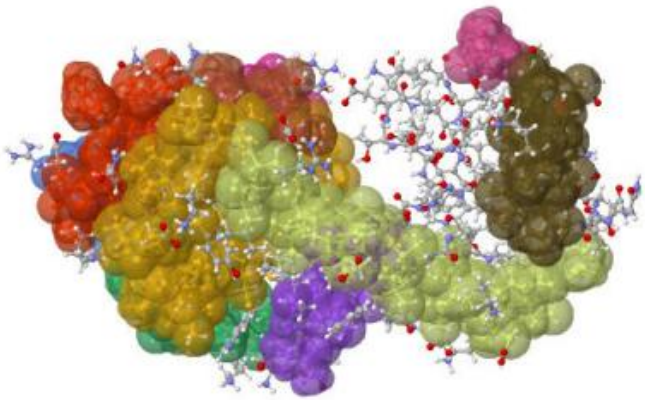
Lysozyme from bactriophage T4 – Rigidity Results



Default Modeling



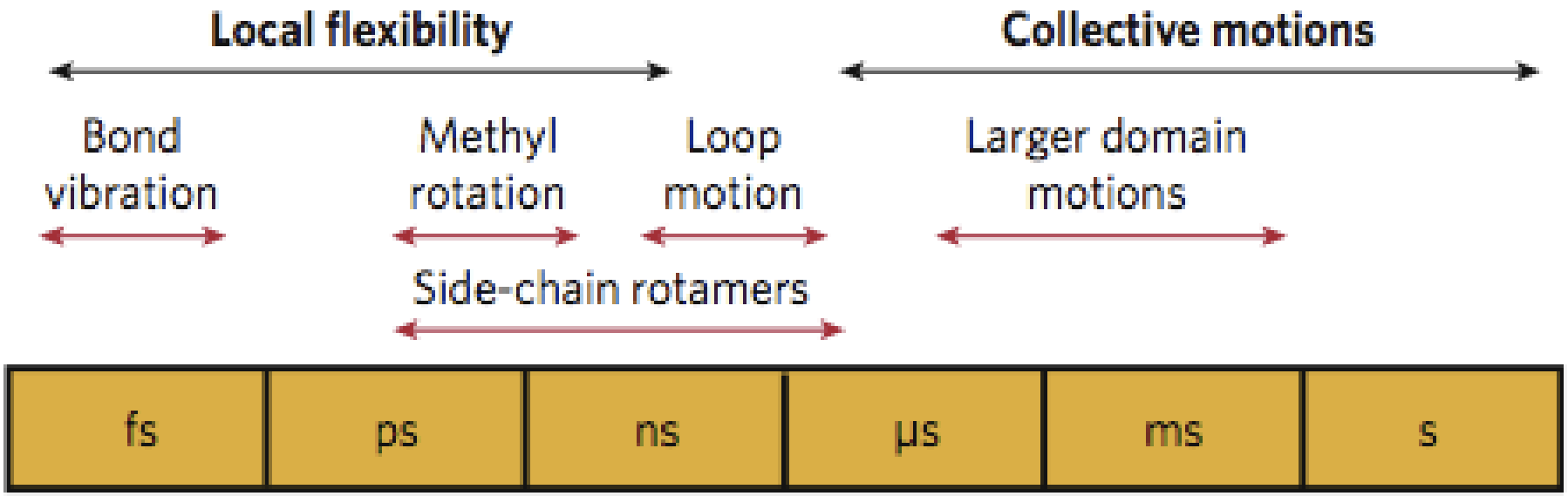
Addition of two hydrogen bonds causes the protein to rigidify



By removing hydrophobic interactions, their effect can be easily seen

KINARI-Web: A Server for Protein Rigidity Analysis *Naomi Fox*, *Filip Jagodzinski*, *Yang Li*, *Ileana Streinu*
Nucleic Acids Research, **39** (Web Server Issue), 2011.

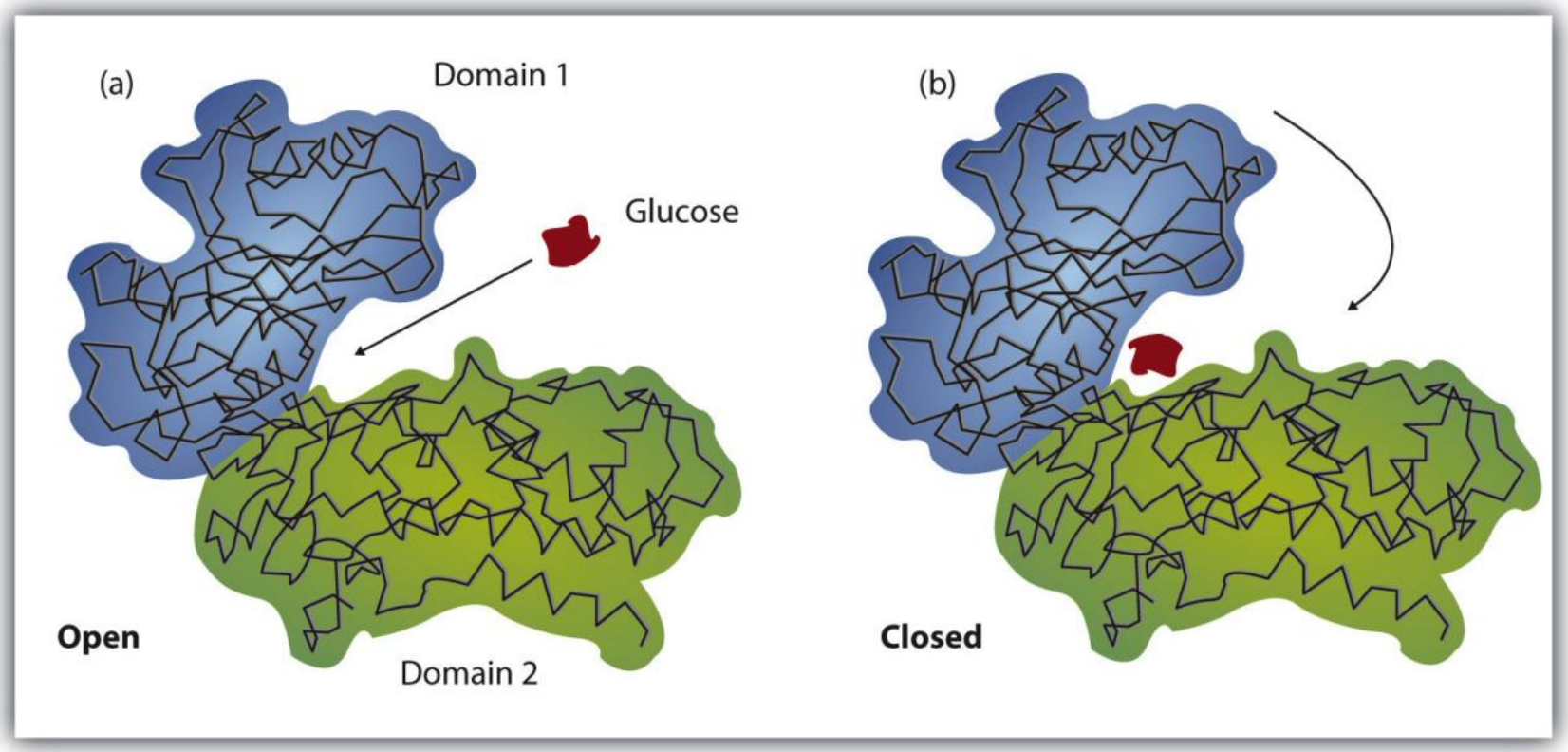
Part 2:
Comparative overview of tools for studying protein flexibility



Henzler-Wildman, Katherine and Kern, Dorothee, *Dynamic personalities of proteins*., Nature, volume 450, number 7172, pages 964-72, 2007.

Large domain, collective motions

Necessary for important protein functions such as Enzyme catalysis, Signal transduction, and Protein-protein interactions



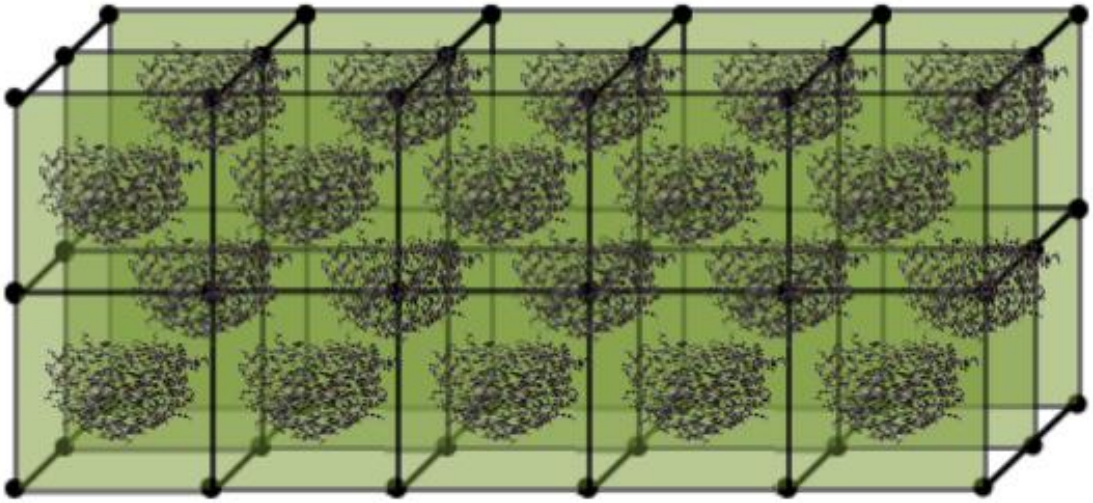
The enzyme hexokinase changes conformation when glucose binds to it.

David W. Ball, John W. Hill, Rhonda J. Scott (2011) The Basics of General, Organic, and Biological Chemistry. Flat World Knowledge.

Methods for studying protein kinematics

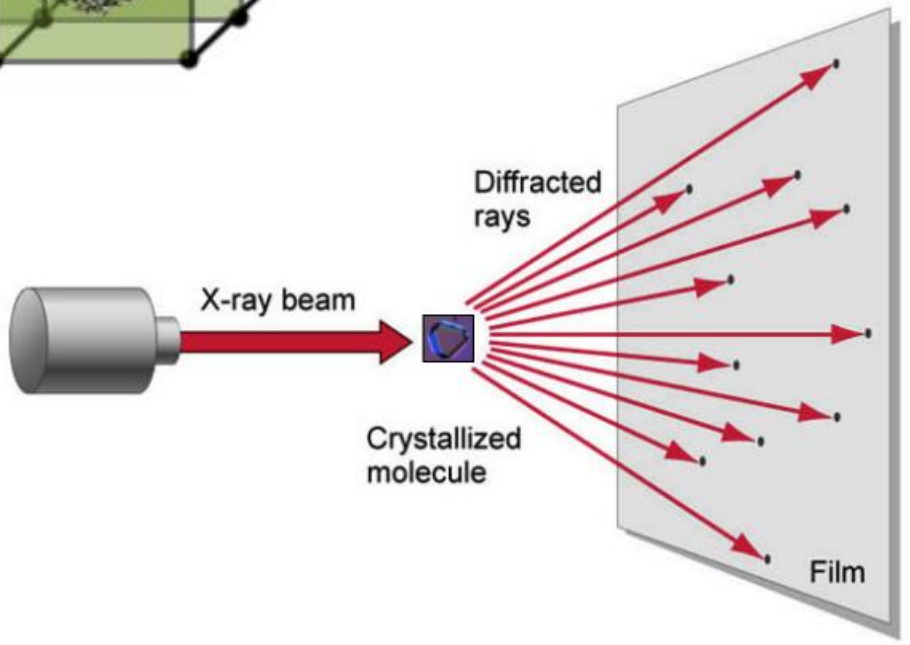
UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

	In Vitro	*In Silico*
Large Domain, Collective Motions	Fluorescence	ENM and NMA Combinatorial Rigidity Analysis
Thermal fluctuations	X-ray Crystallography NMR Cryo-EM	Molecular Dynamics



Step 1:
Crystallize the protein in the native state

Step 2:
Get diffraction patterns



Step 3:
Determine electron density map

Step 4:
Fitting and refinement

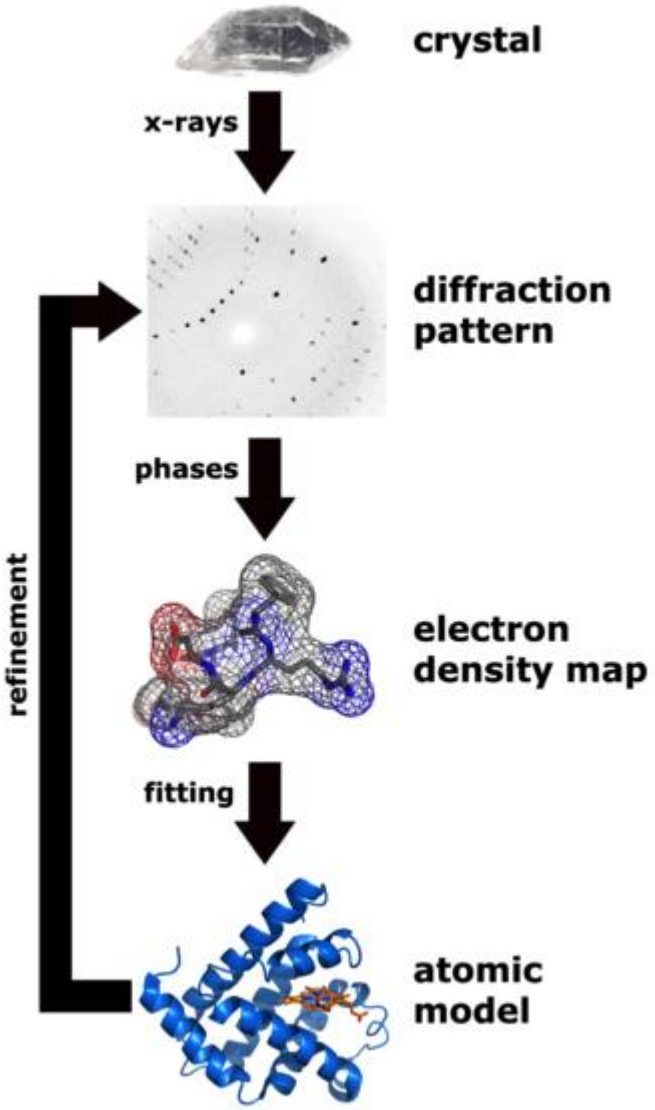
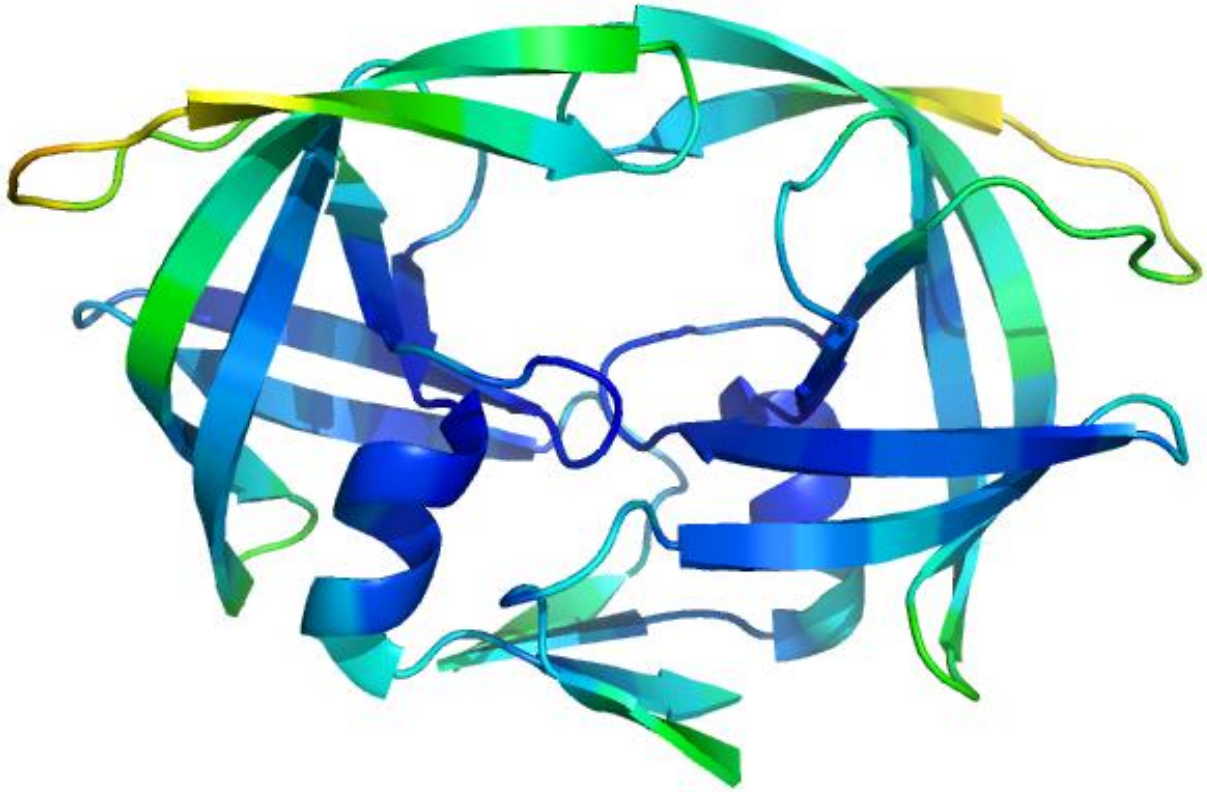


Image from Wikimedia courtesy Thomas Splettstoesser

B-factors – describe the displacement from mean coordinate



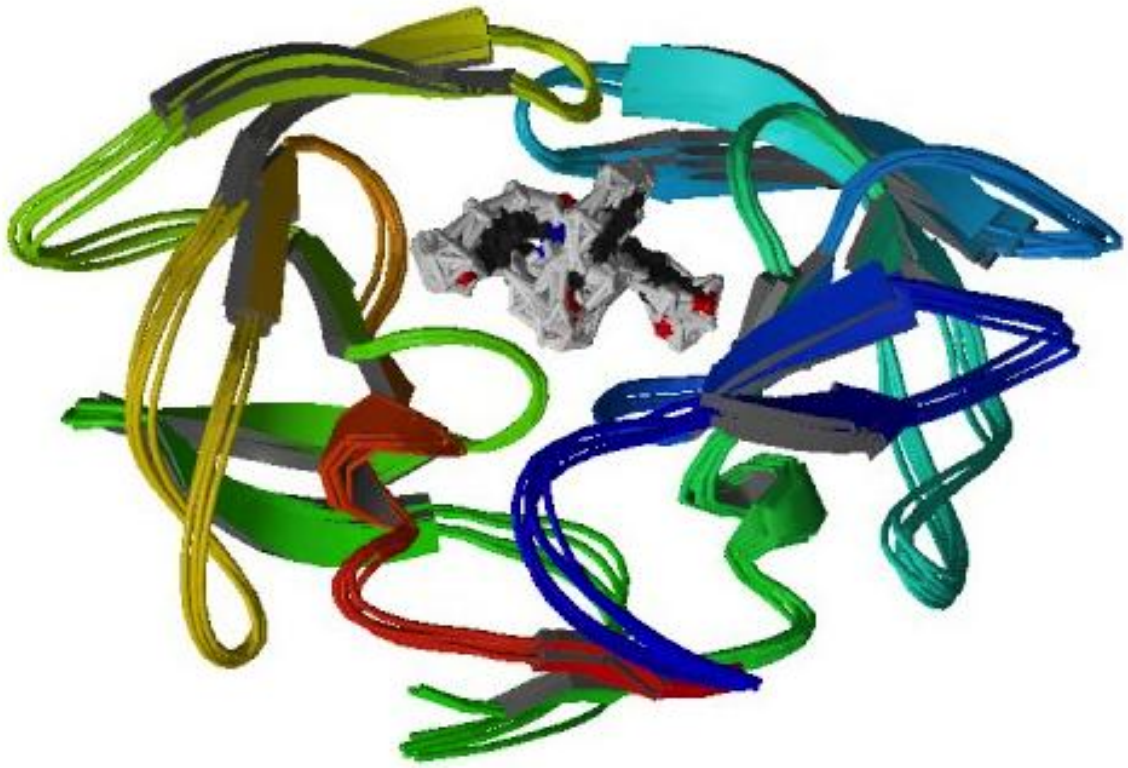
“Missing” atoms – position could not be determined

>70,000 X-ray data files in PDB

Computes coordinates from set of distance constraints

Multiple conformations produced

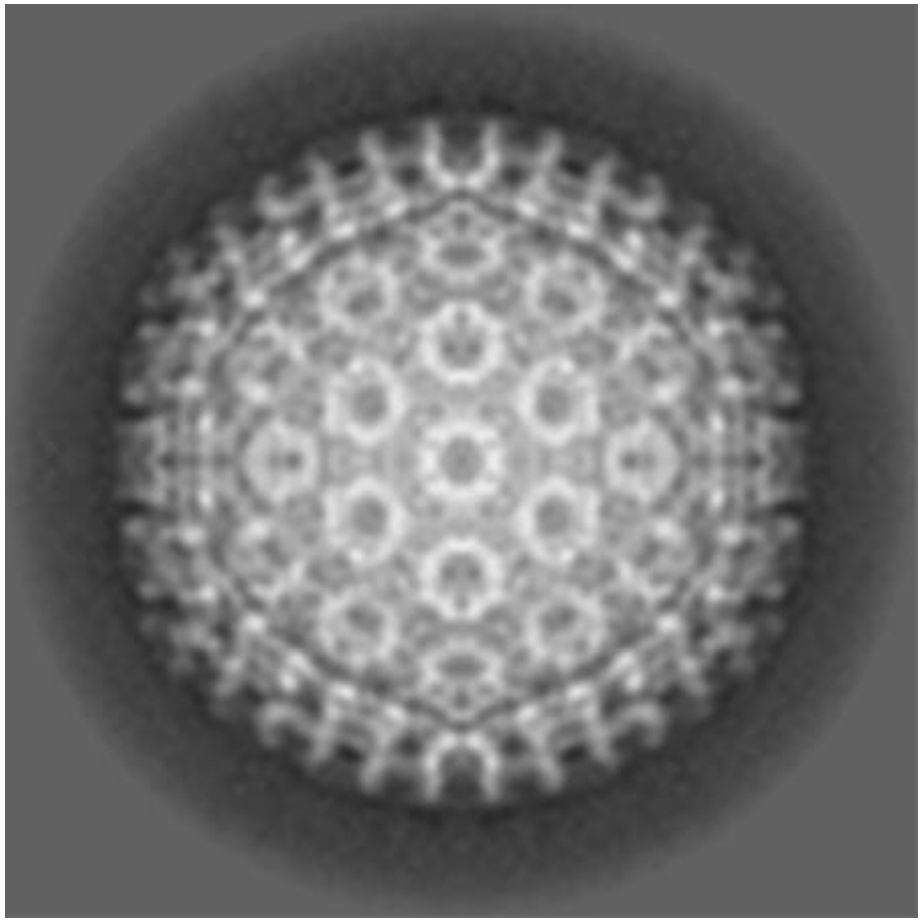
~9,000 NMR-resolved structures in PDB



PDB File 1BVE, HIV-1 protease

Rapidly freeze protein in solution and view under electron microscope

< 400 structures in PDB



EM DATA BANK (EMDB) / 5124

Model all the interatomic interactions in a system and simulate with fs time step.

Can only simulate fast timescales

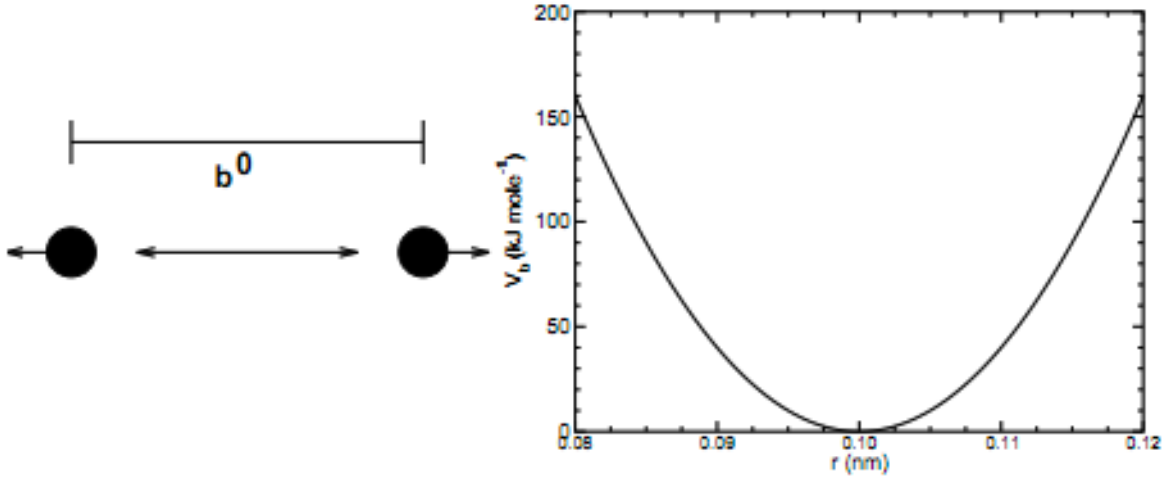


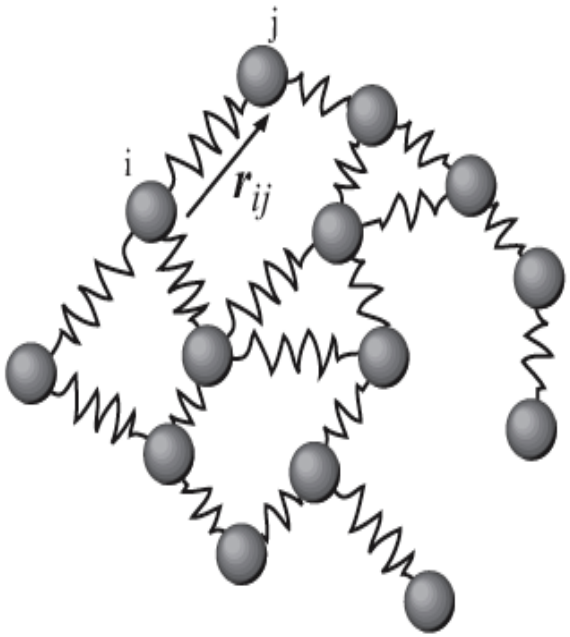
Figure 4.5: Principle of bond stretching (left), and the bond stretching potential (right).

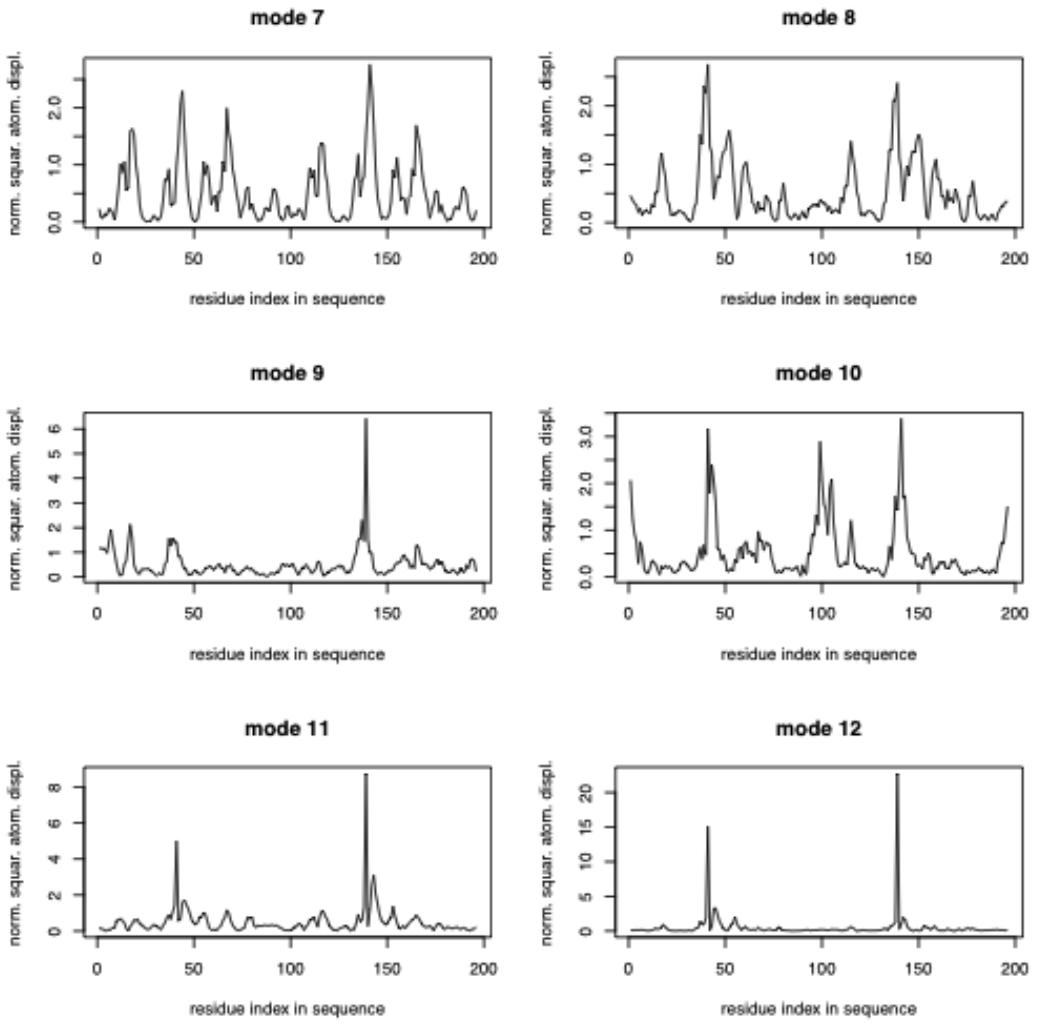
(2010) GROMACS USER MANUAL Version 4.5 http://manual.gromacs.org/

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Build an elastic network model of the protein

Determine the normal modes of the model



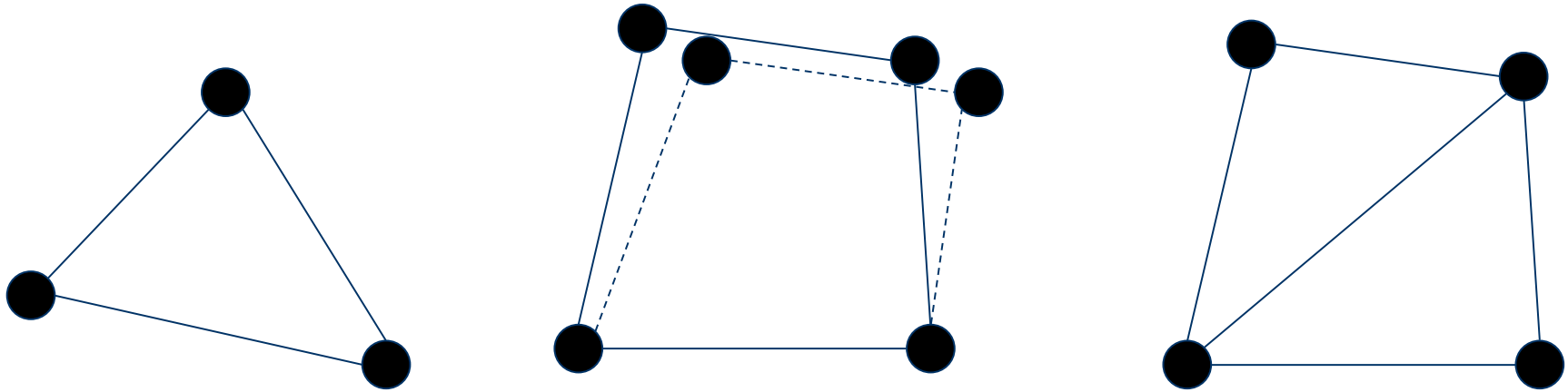


Fluctuations of 1HVR computed using NMA. SM, Sælensminde G, Reuter N. WEBnm@: a web application for normal mode analysis of proteins BMC Bioinformatics. 2005 Mar 11;6(1):52

Chennubhotla, Chakra, Rader, A. J. and Bahar, Ivet, Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies, in: Physical Biology, volume 2, pages S173-S180, 2005.

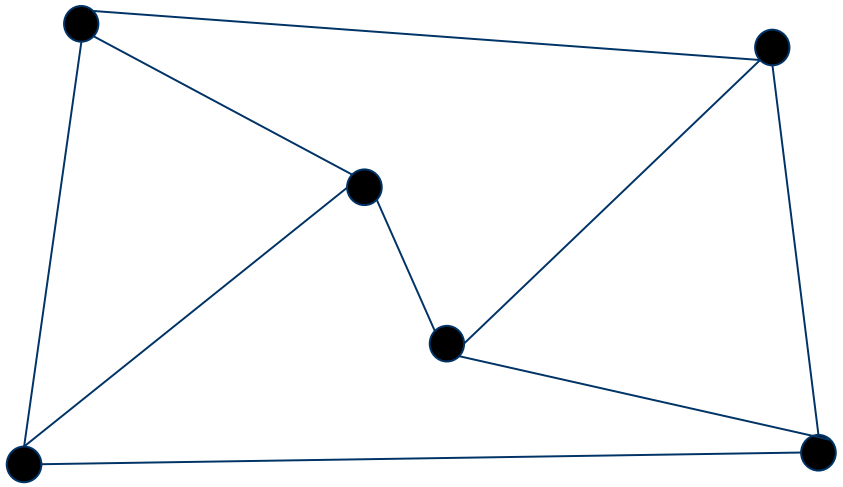
Part 3 :
Introduction to Rigidity Theory in 2 and 3 Dimensions

A framework G(p) is ***rigid*** if it has no continuous deformation.



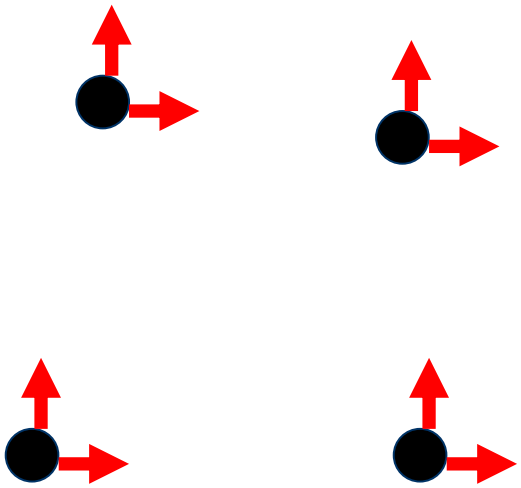
Which frameworks are rigid and which are not?

Is this rigid?



UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

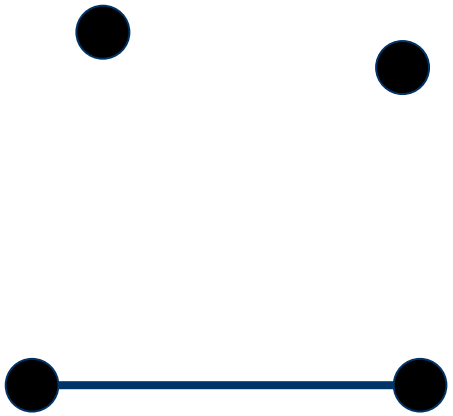
Exercise: Counting Degrees of Freedom



A 2D framework with n points and no edges has 2n degrees of freedom (DOF).

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Exercise: Counting Degrees of Freedom in 2D

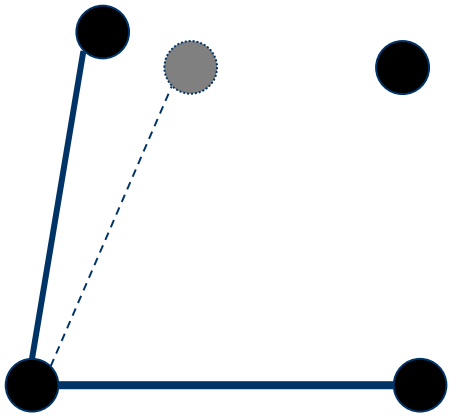


Adding a bar reduces the number of degrees of freedom by 1.

Framework has 7 DOF

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Exercise: Counting Degrees of Freedom in 2D

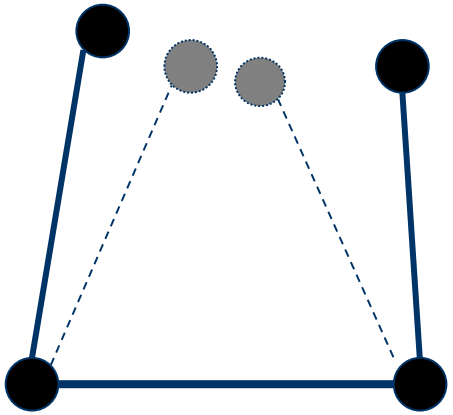


Adding a second bar again reduces the number of degrees of freedom by 1.

Framework has 6 DOF

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Exercise: Counting Degrees of Freedom in 2D

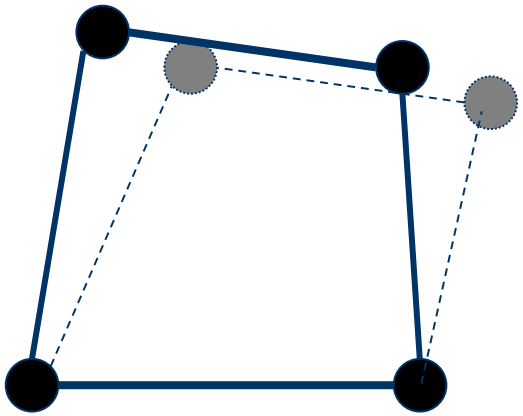


Adding a third bar again reduces the number of degrees of freedom by 1.

Framework has 5 DOF

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

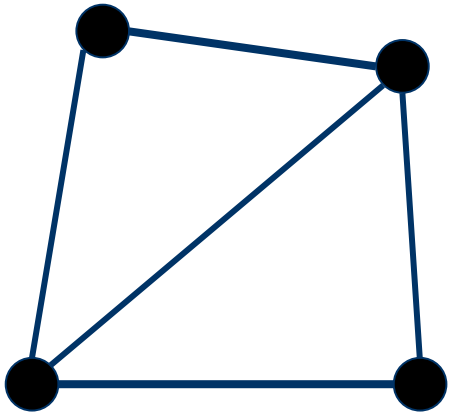
Exercise: Counting Degrees of Freedom in 2D



Adding a fourth bar again reduces the number of degrees of freedom by 1.

Framework has 4 DOF

Exercise: Counting Degrees of Freedom in 2D

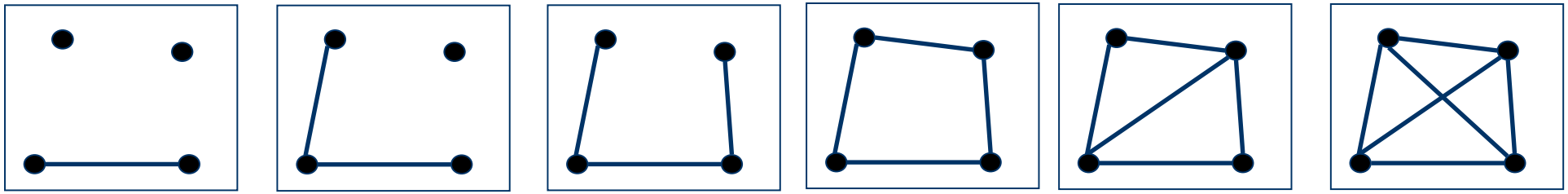


Adding a fifth bar again reduces the number of degrees of freedom by 1.

Framework has 3 DOF

UMASSCS DEPARTMENT OF COMPUTER SCIENCE

Exercise: Counting degrees of freedom in 2D



7 DOF | 6 DOF | 5 DOF | 4 DOF | 3 DOF | 3 DOF

UMASSCS DEPARTMENT OF COMPUTER SCIENCE

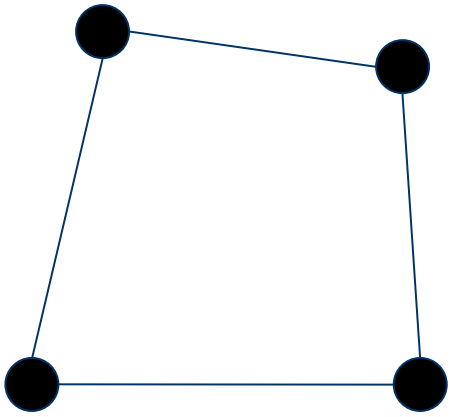
The Maxwell-Laman theorem

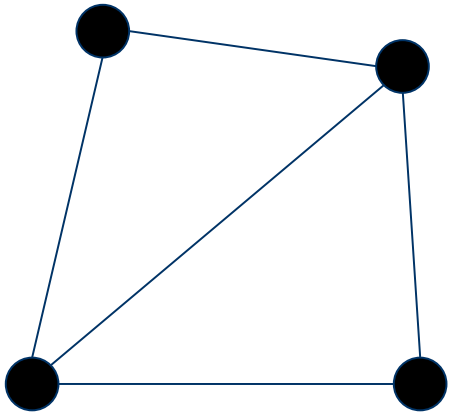
A framework G(p) is generically minimally rigid

iff

- For every subset of vertices, |E'| <= 2|V'| - 3
- |E| = 2|V|-3

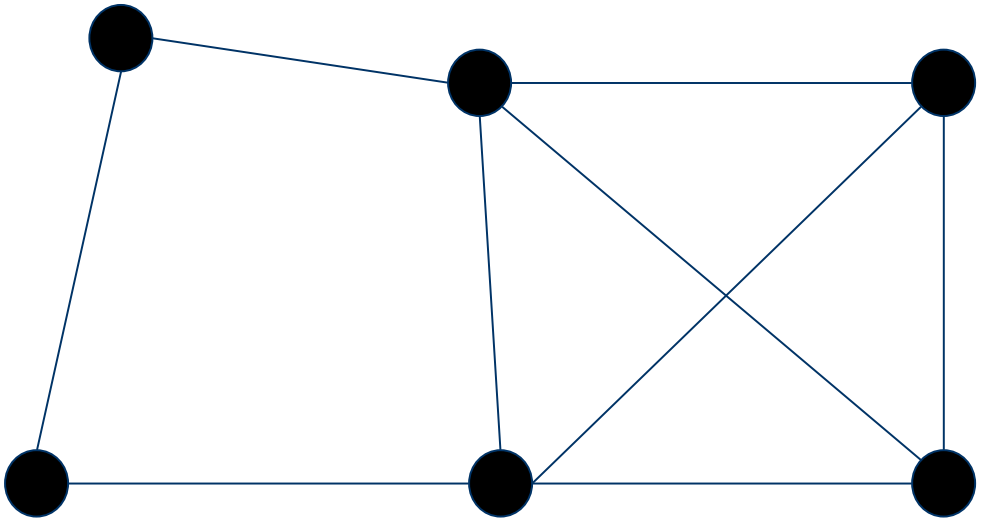
|V| = n = 4
|E| = m = 4
|E| < 2n-3





|V| = n = 4
|E| = m = 5
|E| = 2n-3

Rigid components



n = 6

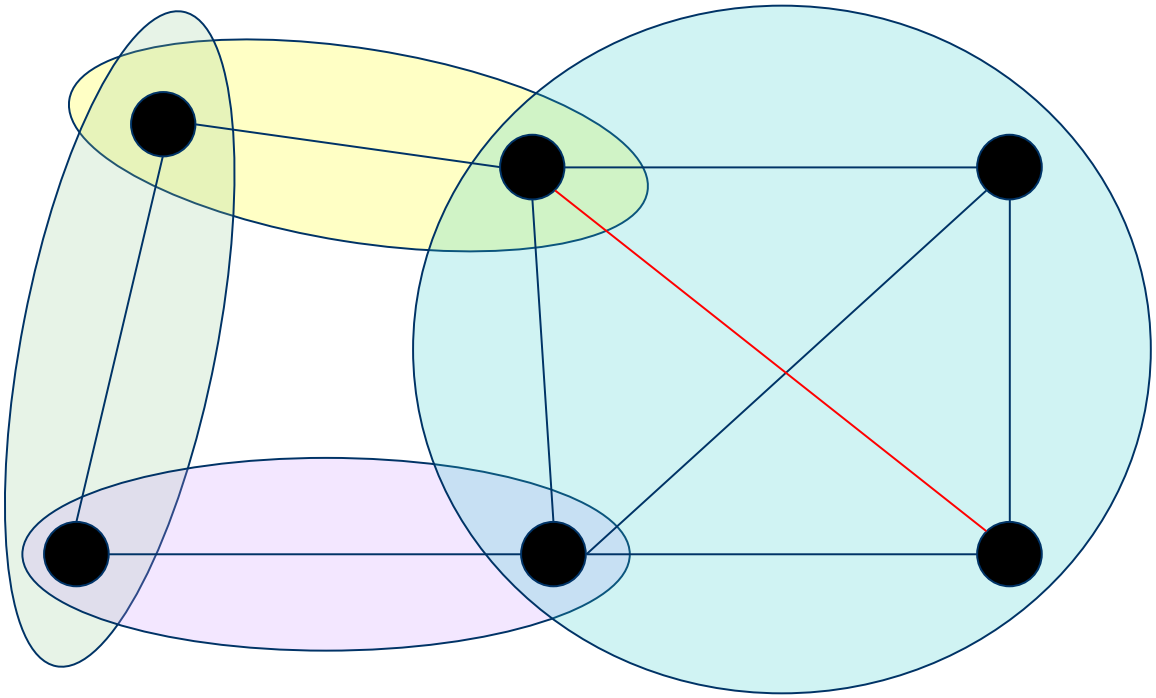
m = 9

2n-3 = 9

Is this graph rigid?

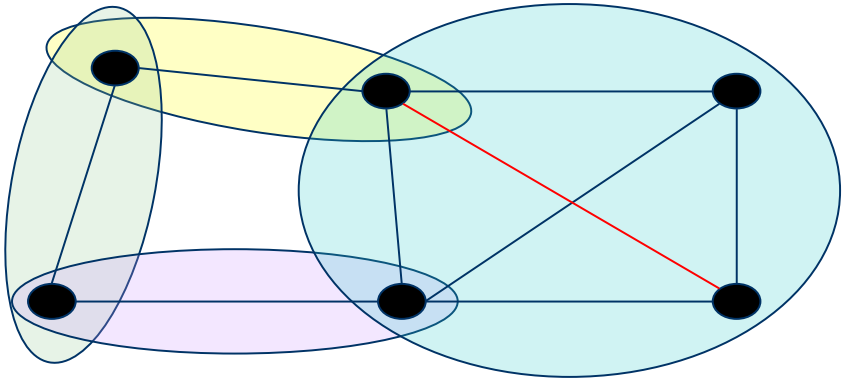
Rigid Components

Rigid components are maximal sets of vertices which are rigid to each other.

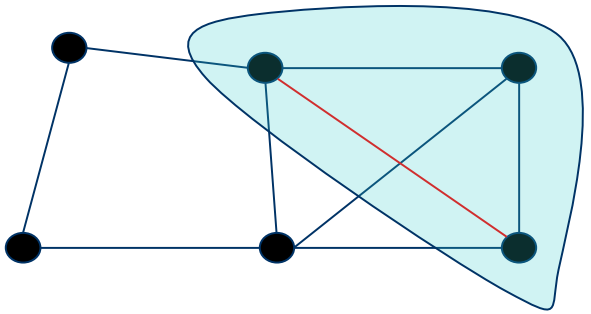


Rigid components

Rigid zcomponent	Maximal set of vertices that are rigid to each other.
(2,3)-component	Maximal set of vertices containing a (2,3)-tight graph.



Not a component.



Component is rigid

↔

Component contains a spanning Laman subgraph.

Can we extend Laman to 3D?

Will 3n-6 edge counts work for 3D?

Why 3n-6?

For each vertex:

Translations along x,y,z axes.

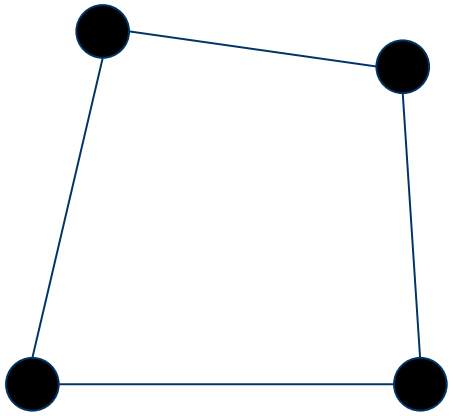
For the entire framework:

Rigid translations along x,y,z axes

Rigid rotations around x,y,z axes

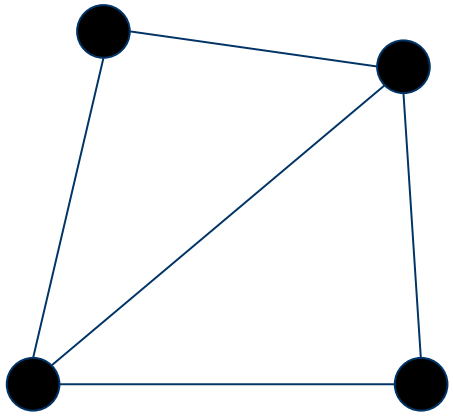
Can we extend Laman to 3d?

Can we rigidify with 3(4) – 6 = 6 edges?



n = 4, m=4

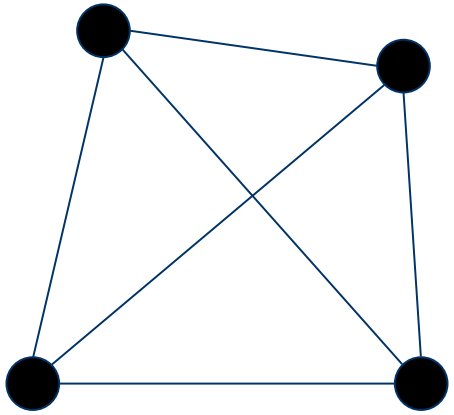
8 DOF



n = 4, m=5

7 DOF

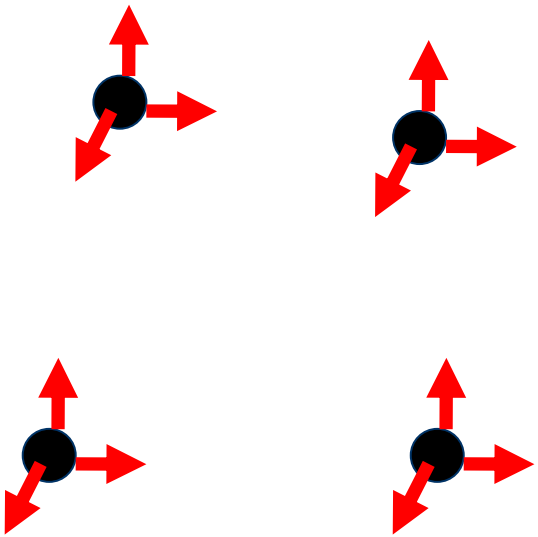
Rigid?



K_4 is rigid

6 DOF

Exercise: Counting Degrees of Freedom in 3D

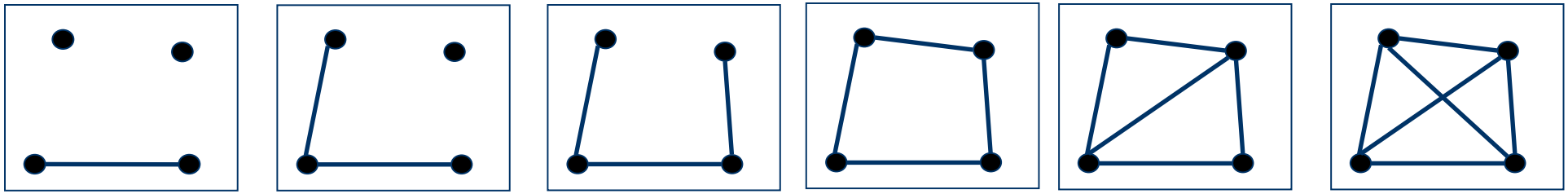


A 3D framework with n points and no edges has 3n DOF.

Framework has 3(4) = 12 DOF.

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

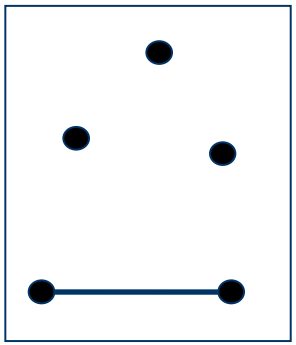
Exercise: Counting Degrees of Freedom in 3D

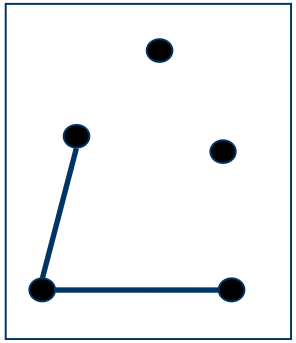


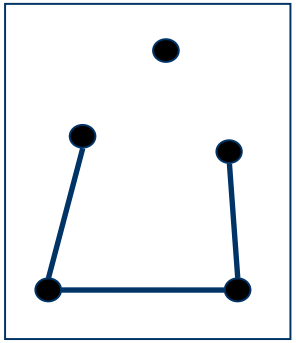
11 DOF | 10 DOF | 9 DOF | 8 DOF | 7 DOF | 6 DOF

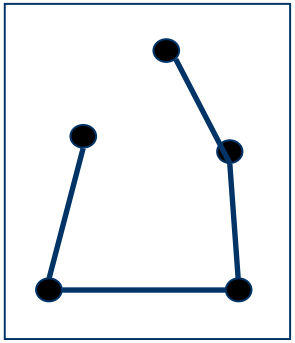
UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

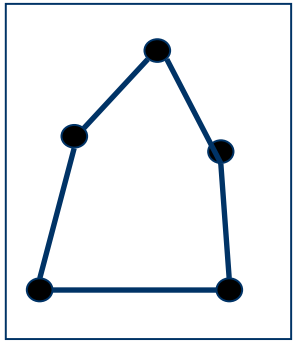
Exercise: Counting Degrees of Freedom in 3D

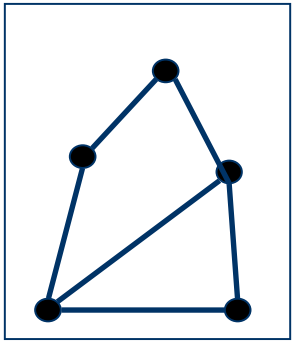

14 dofs

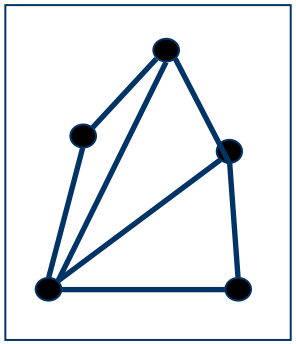

13 dofs

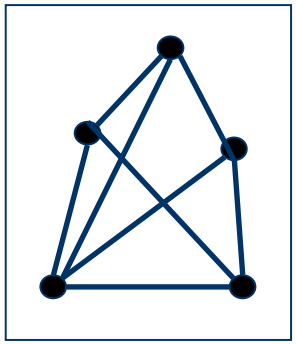

12 dofs

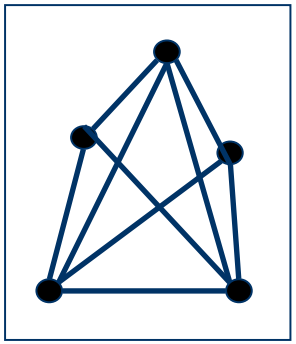

11 dofs

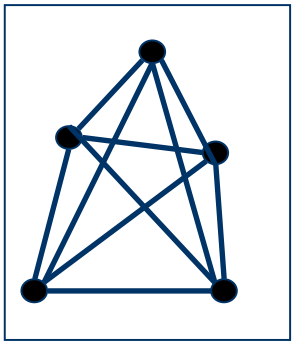

10 dofs


9 dofs


8 dofs


7 dofs

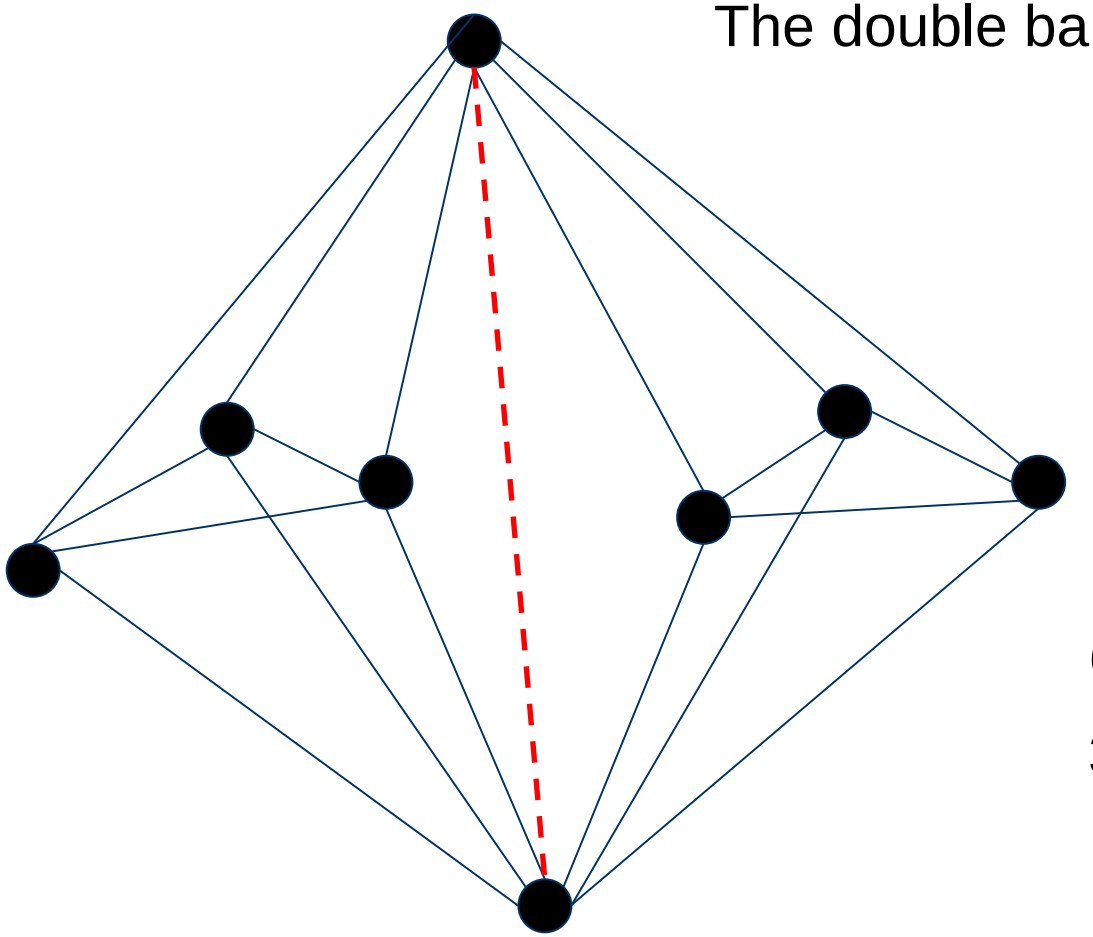

6 dofs


6 dofs

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Counterexample to (3,6)-counts

The double banana



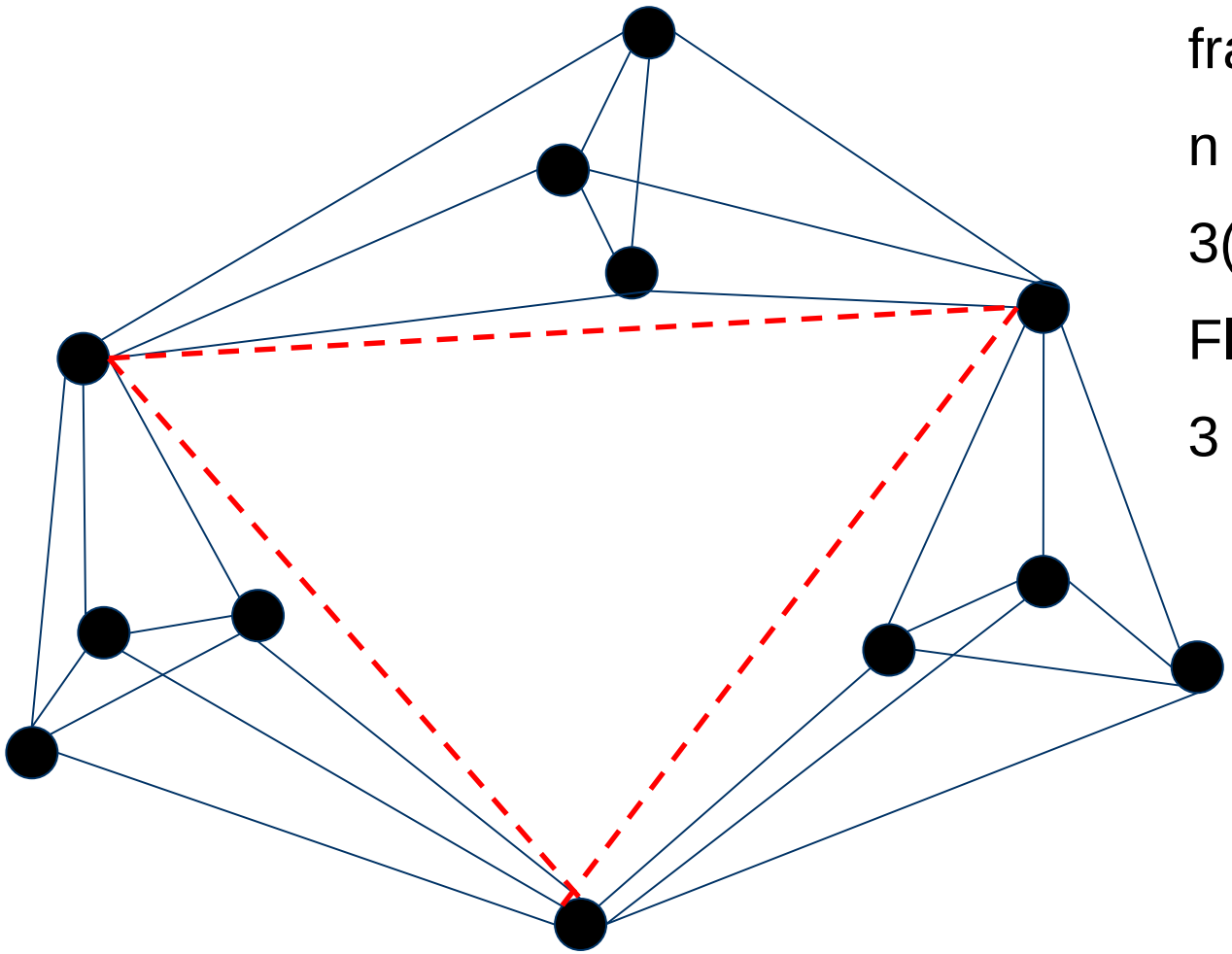
- Edges satisfy (3,6)-counts, but framework is Not rigid

(3,6)-tight

3(8)-6=18=m

Not rigid

The triple banana



For the entire framework:

n = 12, m=27

3(12)-6=36-6=30

Flexible with 9 dofs

3 non-trivial dofs.

Which are the rigid components?

2D vs 3D generic rigidity

2d	3d
G contains a spanning (2,3)-tight subgraph ←→ G is generically rigid.	**Sometimes**, when G contains a spanning (3,6)-tight subgraph, G is generically rigid.
Rigid components ALWAYS induce spanning (2,3)-tight subgraphs.	**Sometimes**, rigid components induce spanning (3,6)-tight subgraphs.

Is there a class of frameworks for which:

G contains a spanning (3,6)-tight subgraph

←→

G is generically rigid in 3D.

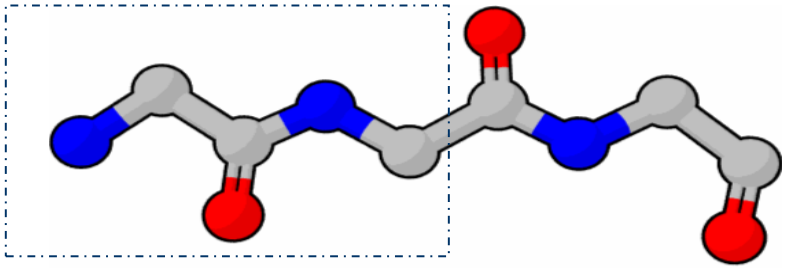
Bar and Joint and Proteins

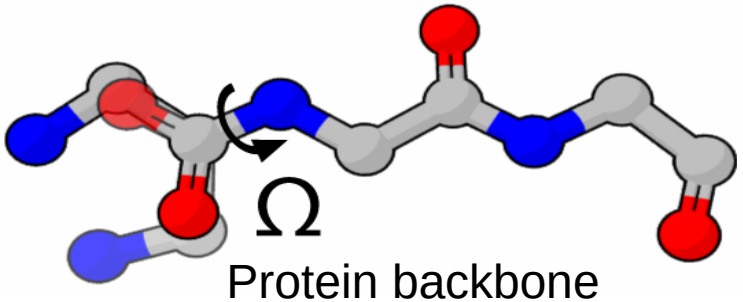
Bars

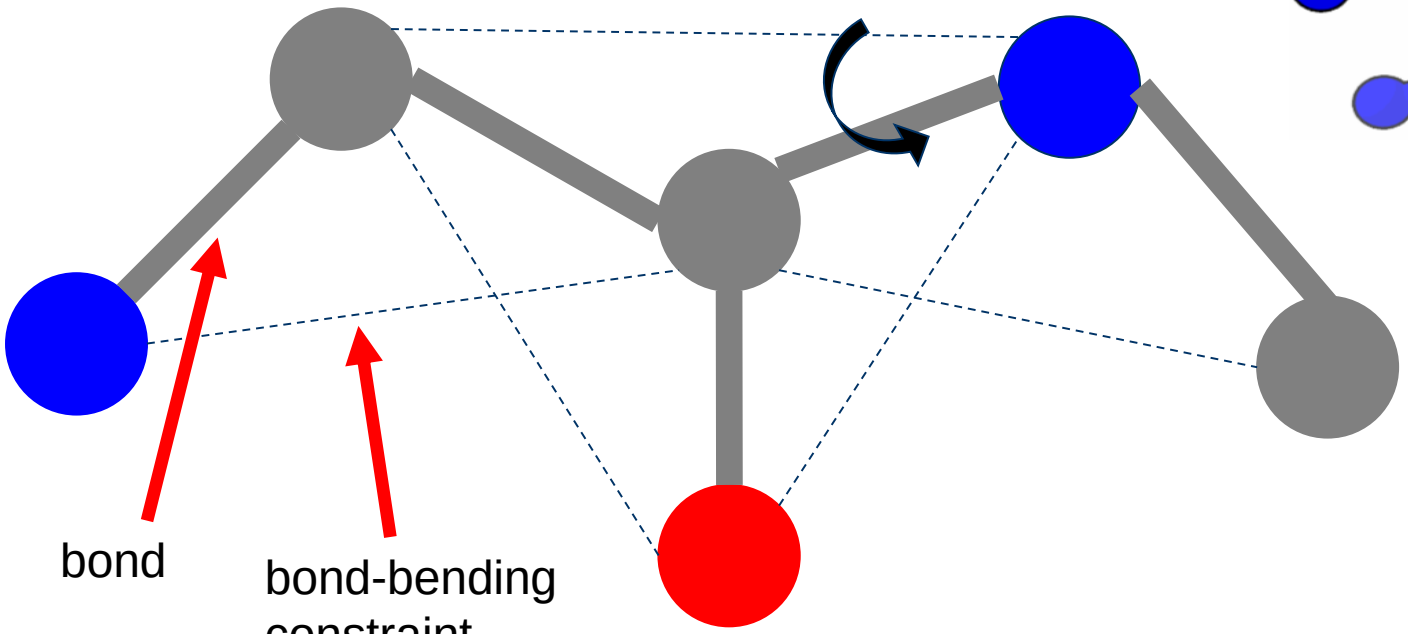
- bonds
- bond-bending constraints

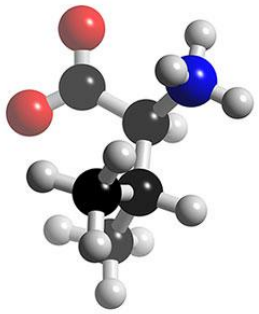
Joints:

- Atoms





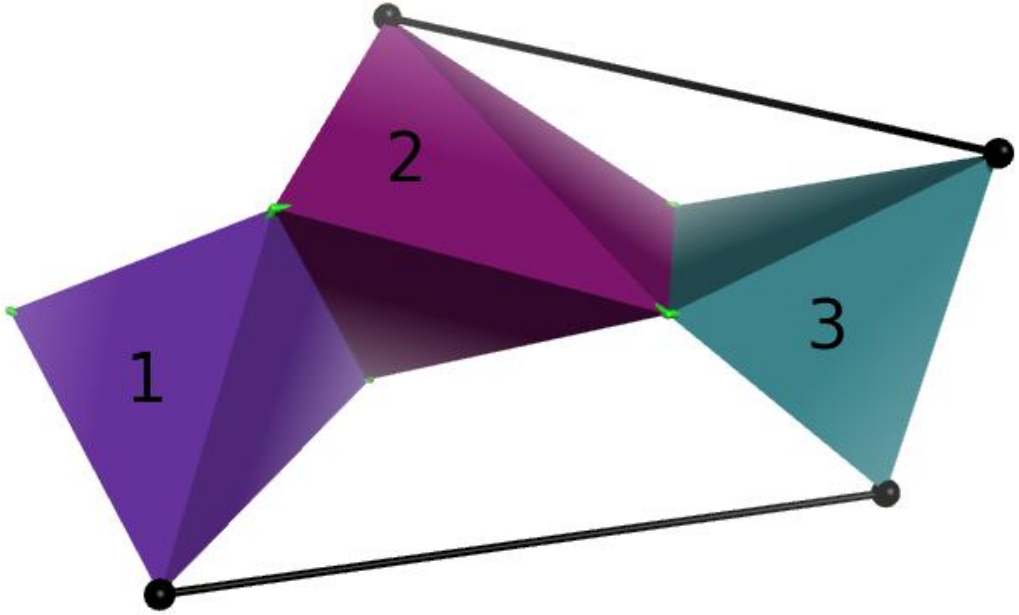


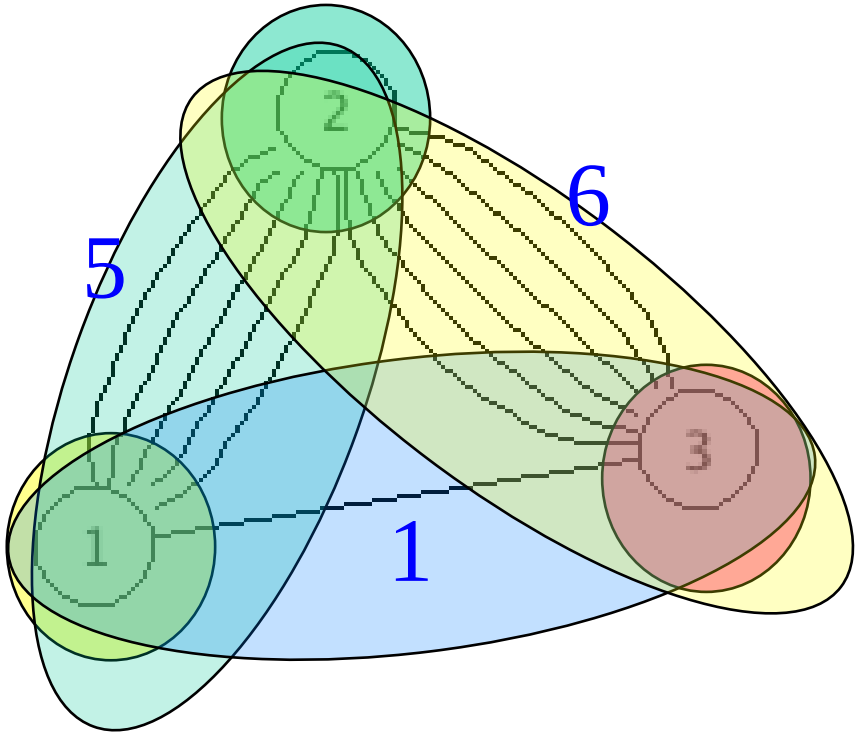


Theorem for 3D Body-Bar-Hinge

A body-bar-hinge framework is generically minimally rigid iff:

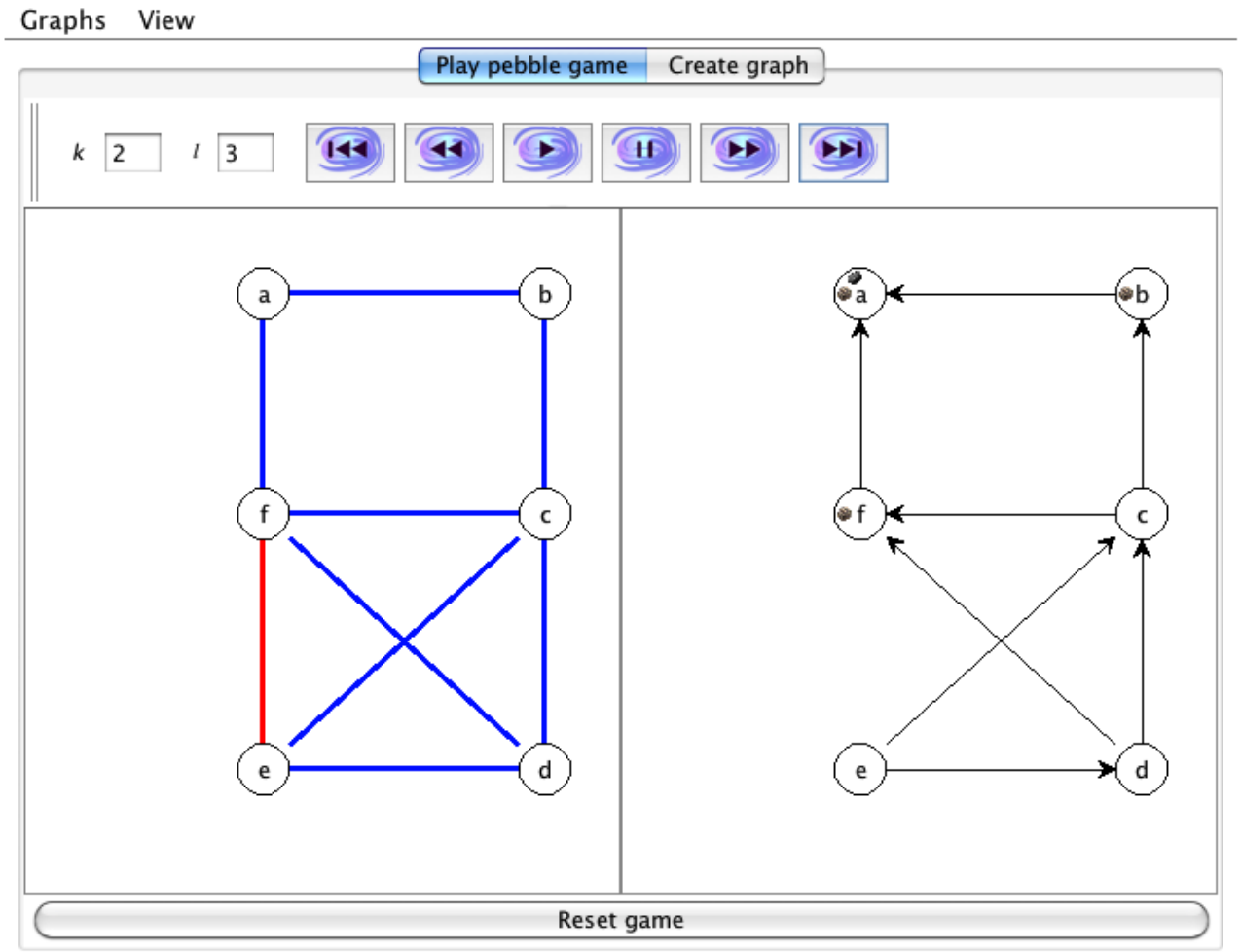
- its associated (multi) graph has exactly 6n-6 edges
- and every subgraph has at most 6n-6 edges





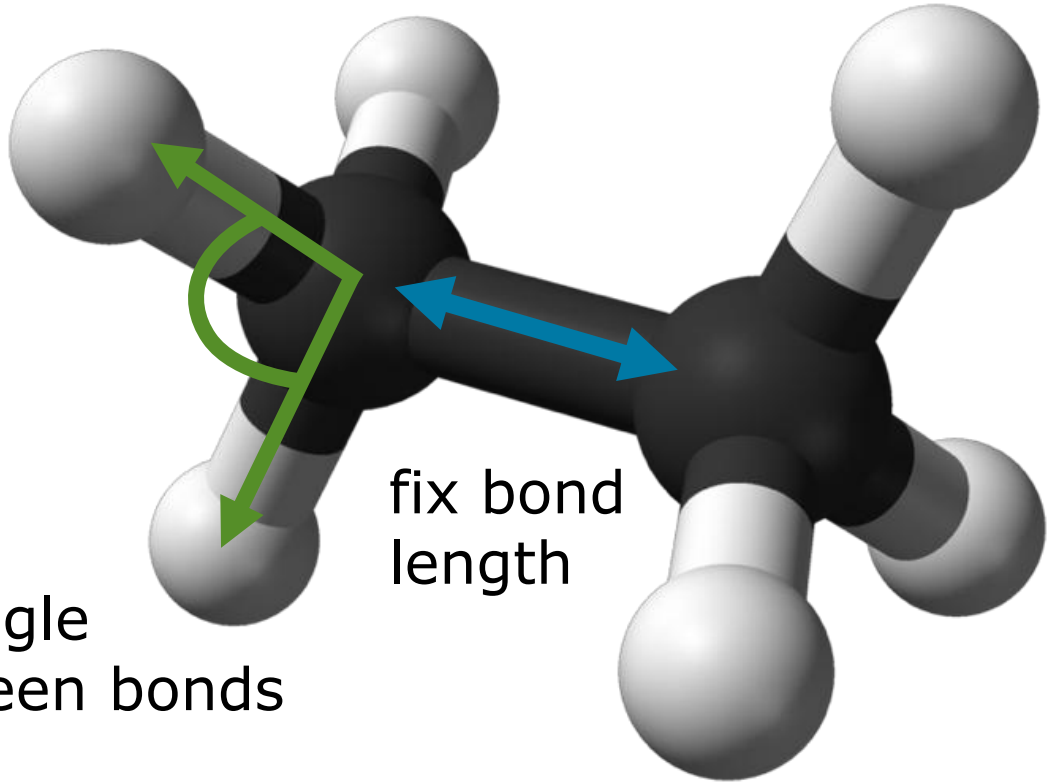
Tay 84, 89; White and Whiteley 87; Katoh and Tanigawa 09

Educational site and demos
http://linkage.cs.umass.edu/pg/pg.html



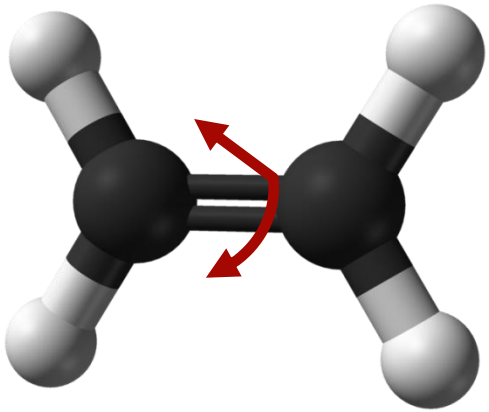
Part 4:
Molecular Modeling and Rigidity Analysis using KINARI

Covalent bonds impose constraints



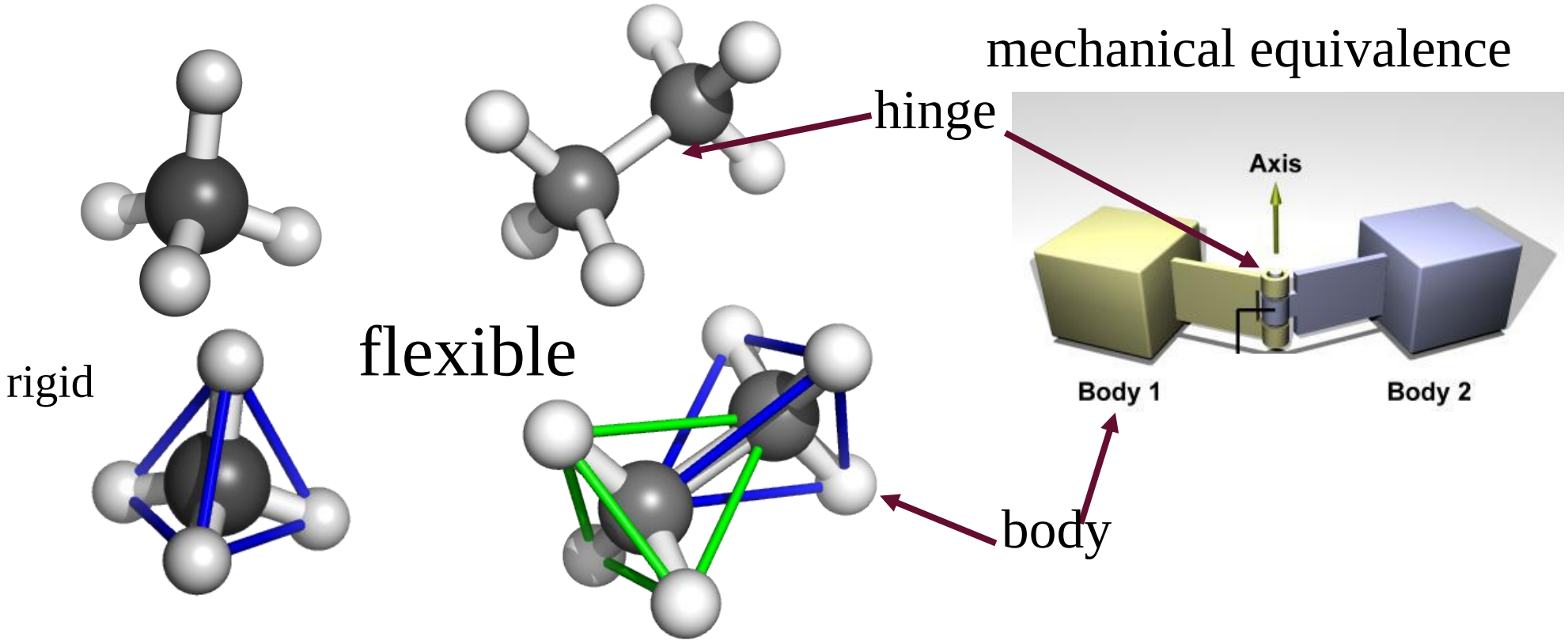
fix angle between bonds

fix bond length

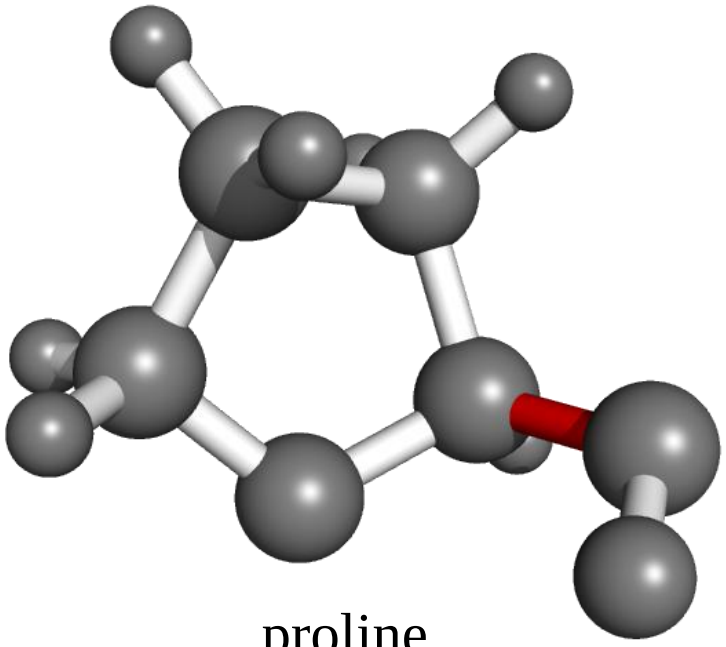


peptide and double bonds fix dihedral angle

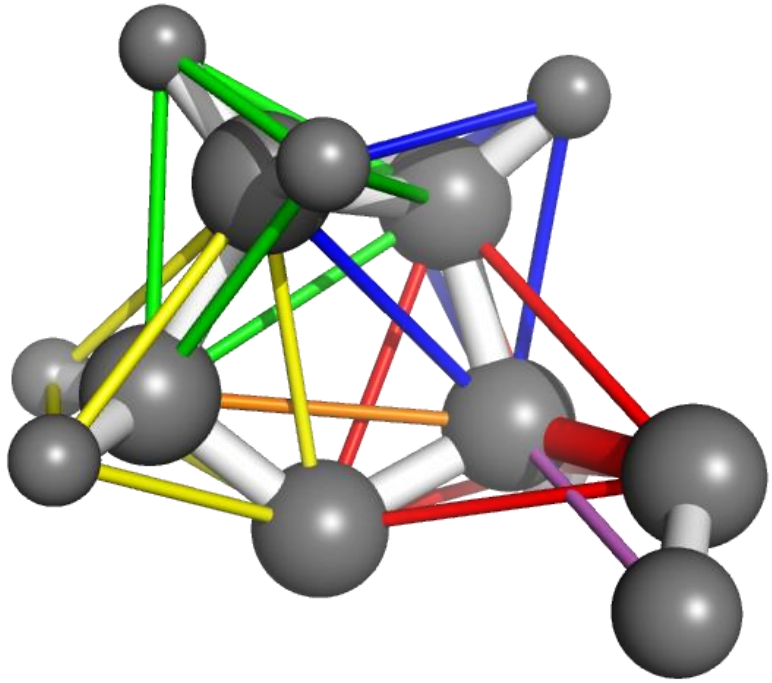
- An atom and its covalent-bonded neighbors form a rigid body.
- Rotatable covalent bonds act as hinges.



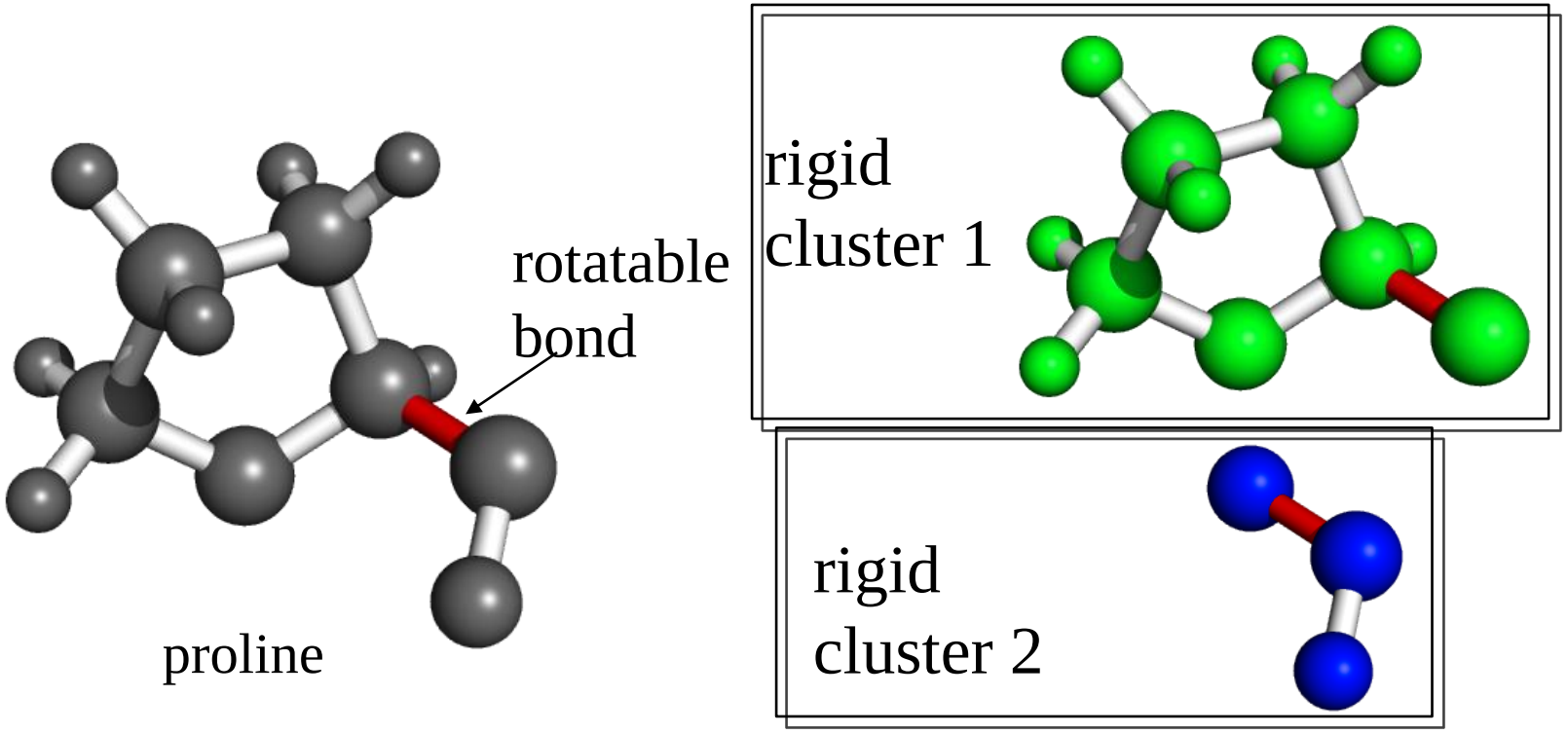
- We know how to determine trivial bodies.
- Would like to find **maximal** bodies.
- A body is maximal if there are no other atoms which are rigidly attached to it.



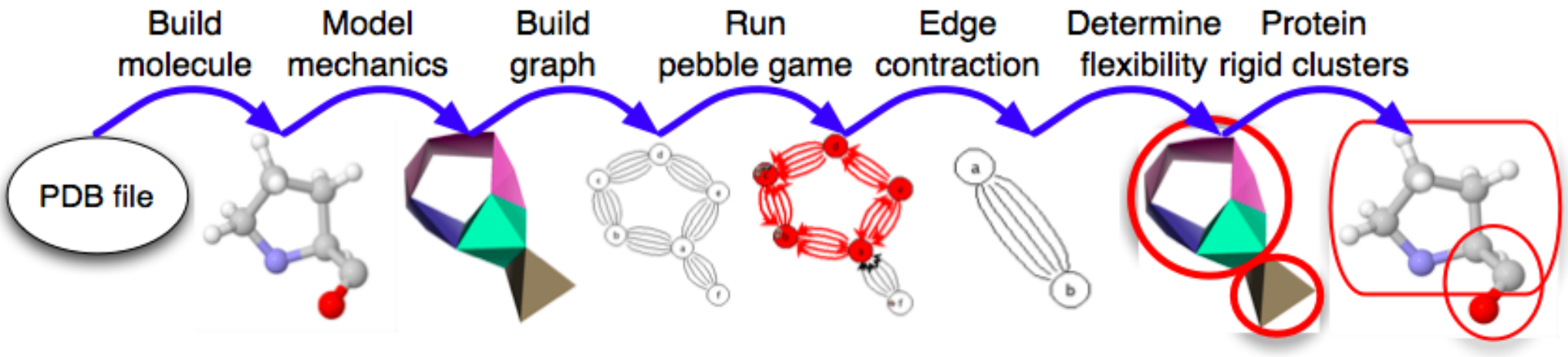
proline



- A *body* or ***rigid cluster*** is a maximal set of atoms and all bonds and interactions that hold the atoms rigidly together.

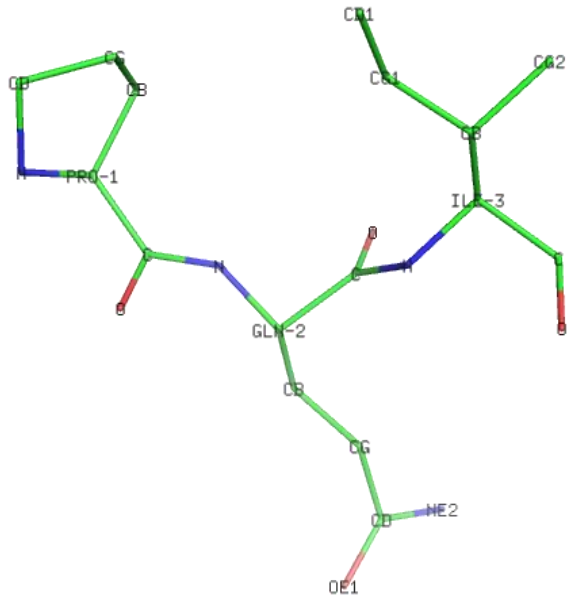


Steps of Protein Rigidity Analysis



PDB File Input

```
ATOM      1  N   PRO A   1      52.574  58.851  -7.646  1.00 34.60           N
ATOM      2  CA  PRO A   1      51.842  59.784  -6.815  1.00 34.88           C
ATOM      3  C   PRO A   1      52.146  59.438  -5.356  1.00 36.01           C
ATOM      4  O   PRO A   1      53.031  58.612  -5.150  1.00 35.05           O
ATOM      5  CB  PRO A   1      50.391  59.581  -7.189  1.00 34.07           C
ATOM      6  CG  PRO A   1      50.353  58.166  -7.724  1.00 32.00           C
ATOM      7  CD  PRO A   1      51.621  58.242  -8.540  1.00 31.72           C
ATOM      8  N   GLN A   2      51.488  60.047  -4.359  1.00 35.47           N
ATOM      9  CA  GLN A   2      51.577  59.564  -2.997  1.00 35.15           C
ATOM     10  C   GLN A   2      50.109  59.273  -2.694  1.00 33.50           C
ATOM     11  O   GLN A   2      49.368  60.231  -2.829  1.00 37.45           O
ATOM     12  CB  GLN A   2      52.076  60.630  -2.075  1.00 37.17           C
ATOM     13  CG  GLN A   2      52.024  59.911  -0.749  1.00 42.02           C
ATOM     14  CD  GLN A   2      52.639  60.592   0.439  1.00 43.50           C
ATOM     15  OE1 GLN A   2      53.841  60.484   0.675  1.00 44.41           O
ATOM     16  NE2 GLN A   2      51.824  61.272   1.246  1.00 43.72           N
ATOM     17  N   ILE A   3      49.555  58.119  -2.306  1.00 30.33           N
ATOM     18  CA  ILE A   3      48.107  57.932  -2.171  1.00 27.03           C
ATOM     19  C   ILE A   3      47.924  57.639  -0.692  1.00 26.95           C
ATOM     20  O   ILE A   3      48.736  56.921  -0.102  1.00 25.73           O
ATOM     21  CB  ILE A   3      47.623  56.740  -3.053  1.00 24.36           C
ATOM     22  CG1 ILE A   3      47.870  57.072  -4.513  1.00 24.07           C
ATOM     23  CG2 ILE A   3      46.116  56.491  -2.898  1.00 22.78           C
ATOM     24  CD1 ILE A   3      47.651  55.899  -5.455  1.00 21.74           C
```



Includes atoms and coordinates from X-ray crystallography or NMR experiments

Usually, does not include hydrogen atoms

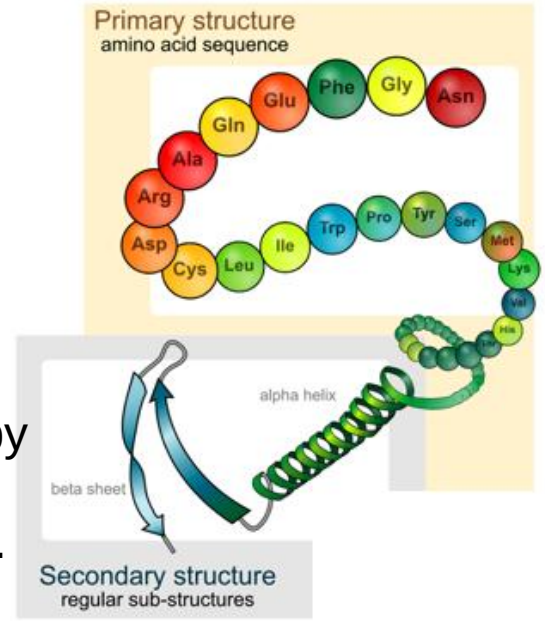
UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

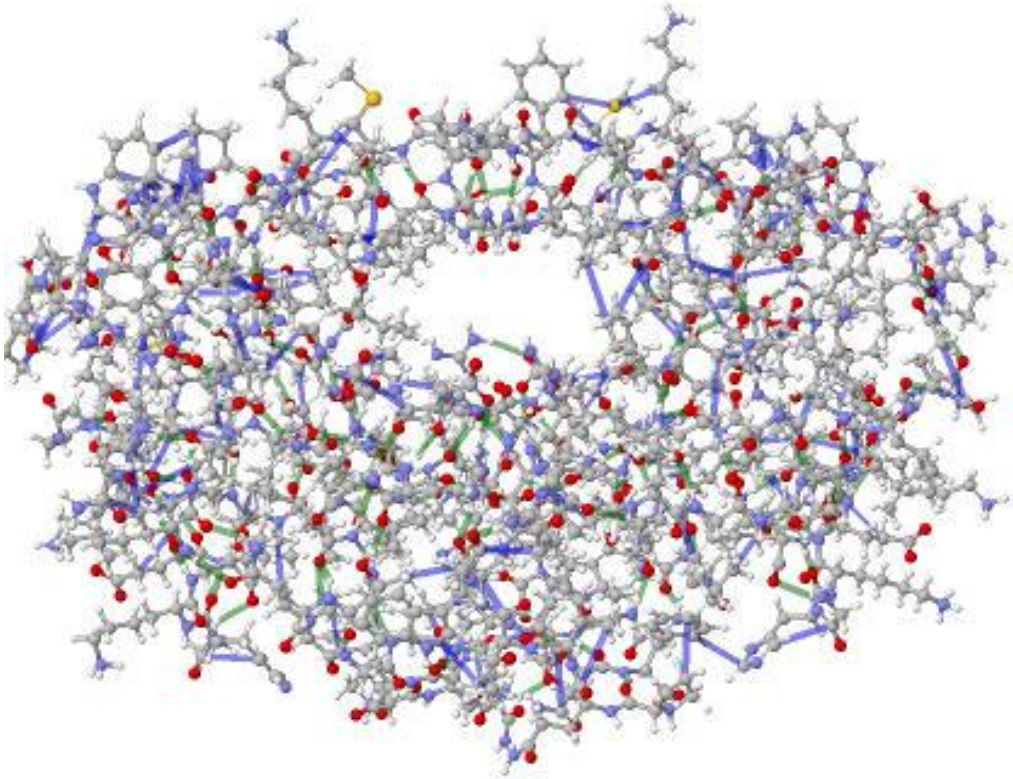
Important interactions for holding a protein's 3D shape

- Covalent bonds
- Hydrogen bonds
- Hydrophobic interactions

Main chain and side chains connected by covalent bonds.

Secondary structures held by backbone hydrogen bonds.

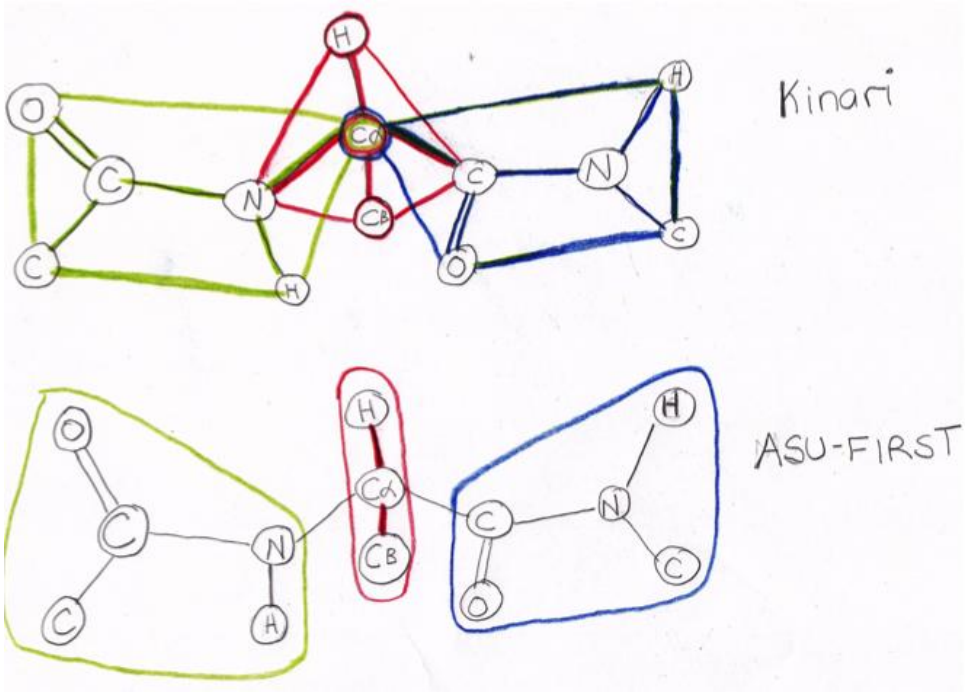




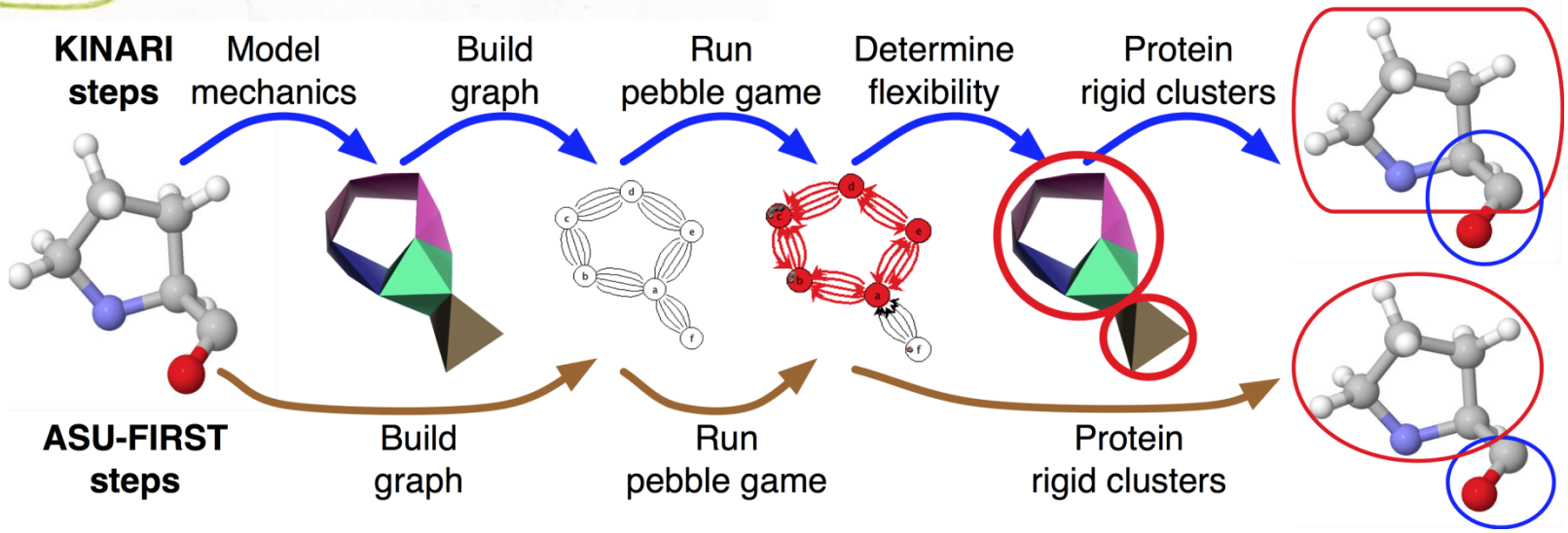
3D folded shape held together by hydrogen bonds and hydrophobic interactions.

- Covalent bonds – modeled as hinges, remove 5 degrees of freedom.
- Hydrogen bonds – include only strong hydrogen bonds. Model as hinges.
- Hydrophobic interactions – weaker interactions. Model with 2 'bars' removing 2 degrees of freedom.

Differences in Systems



ASU-FIRST bodies do not overlap at hinges



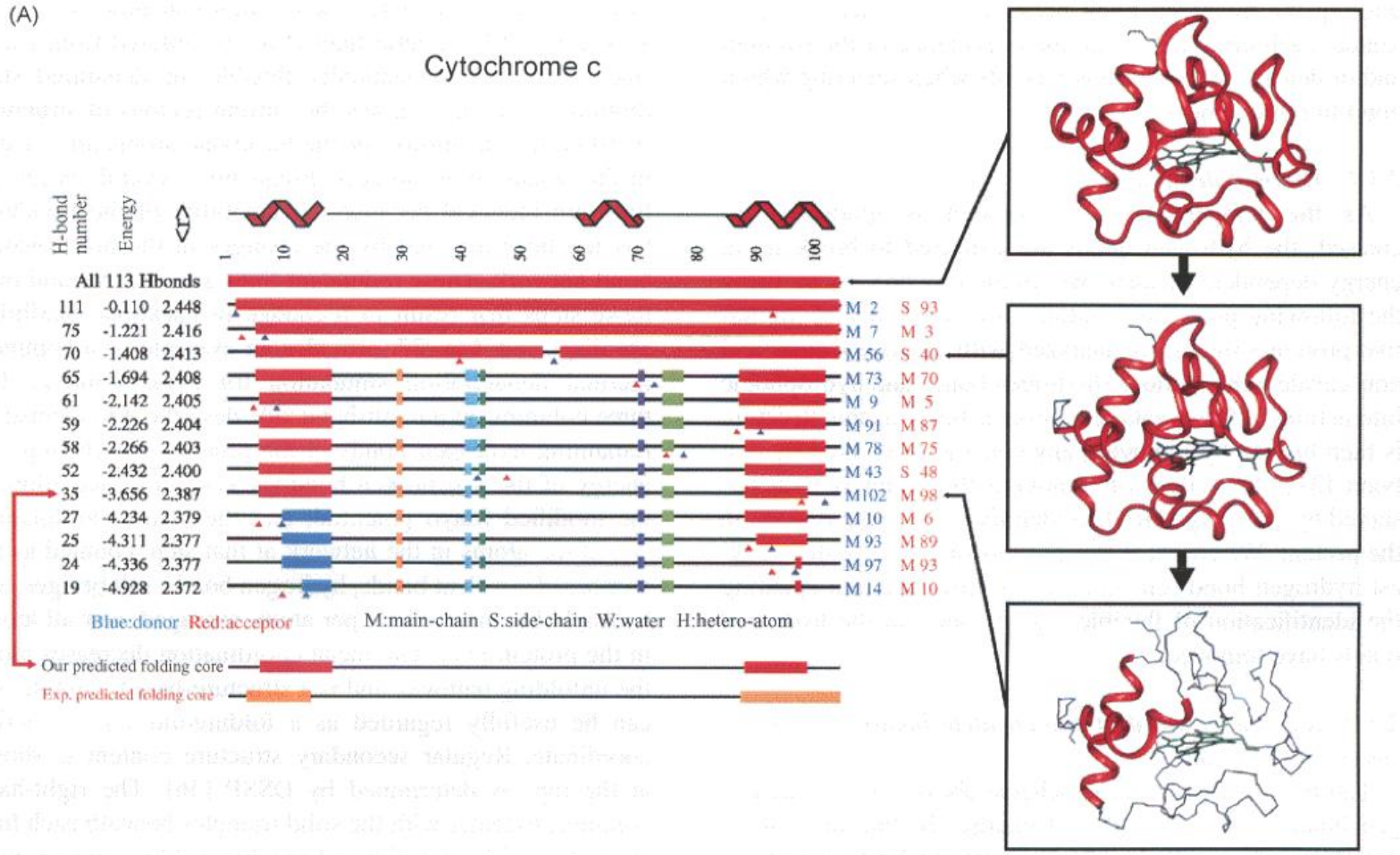
- Body-bar-hinge file (xml file)
  - Set of bodies – lists of atoms
  - Set of bars – pairs of bodies, pairs of atoms
  - Set of hinges – pairs of bodies which share a hinge along the axis between pair of atoms

Part 5: Applications of Protein Rigidity Analysis

- Dilution Analysis (Hespenheide, *et al*)
- Redundancy Analysis (Fox, Streinu)
- FRODA Geometric Simulation (Wells, *et al*)
- Probabilistic Roadmaps (Thomas, *et al*)
- Thermophile stability (Gohlke, Radestock)
- Flexibility of RNA (Fulle, Gohlke)
- KINARI Mutagen (Jagodzinski, Streinu)

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Observing unfolding pathway



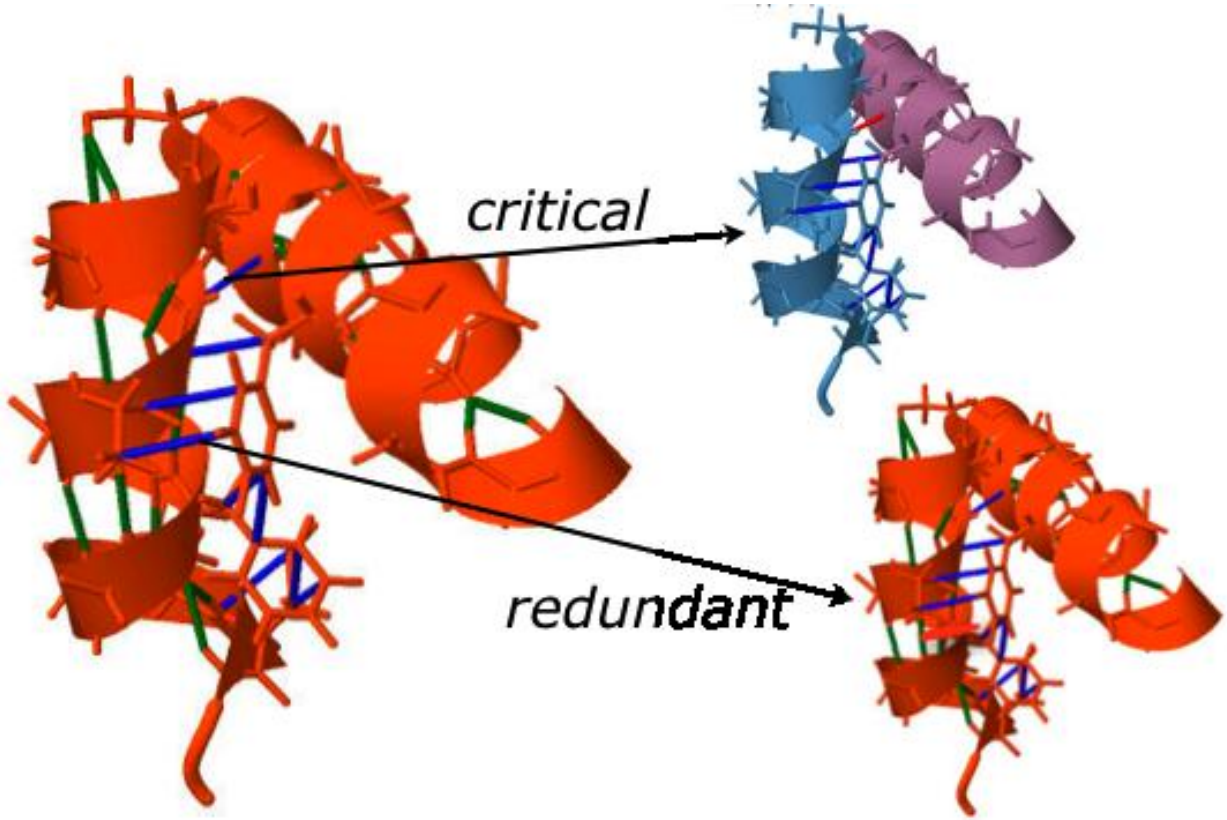
Hespenheide, B. M, Rader, A. J., Thorpe, Michael F. and Kuhn, Leslie A., *Identifying Protein Folding Cores: Observing the Evolution of Rigid and Flexible Regions During unfolding*, in: J Mol Graph Model, volume 21, number 3, pages 195-20, 2002.

Redundancy Analysis

Motivation: Contributions of noncovalent interactions:

- Responsible for holding a protein's 3D shape.
- break and form during normal fluctuations

Our contribution: Classify each interaction in terms of its contribution to the rigidity.



Fox, Naomi and Streinu, Ileana, Redundant Interactions in Protein Rigid Cluster Analysis, in: 1st IEEE International Conference on Computational Advances in Bio and medical Sciences (ICCABS). Feb. 3-5, 2011, Orlando, FL, 2011.

Cluster Redundancy Score

redundant interactions in cluster i

weight of interaction j

Redundancy score of cluster i

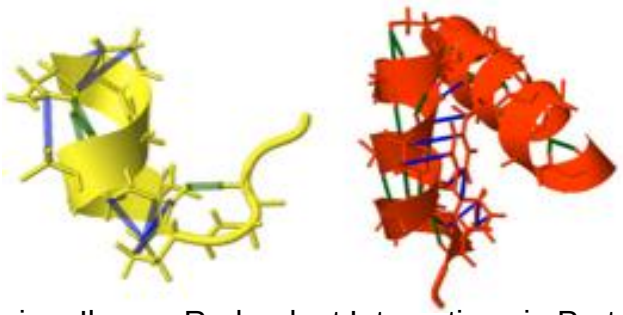
$$\Phi(i) = \frac{\sum_{j \in R(i)} w_j}{\sum_{k \in N(i)} w_k}$$

Sum of weights of all redundant interactions

Sum of weights of all interactions

noncovalent interactions in cluster i

Example: 1HRC



Cluster	Atoms	H-Bonds redun / all	HPh-inter. redun / all	Redun. Score
orange	251	0 / 23	6 / 10	0.089
left yellow	88	1 / 5	6 / 9	0.40

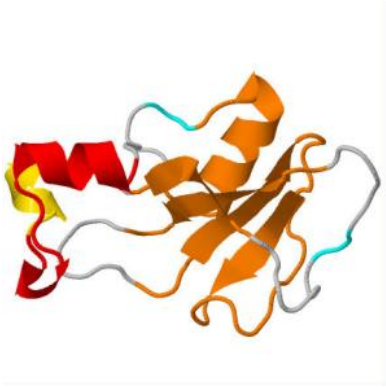
$$\frac{5*0+2*6}{5*23+2*10} = \frac{12}{135} \approx .089$$

Fox, Naomi and Streinu, Ileana, Redundant Interactions in Protein Rigid Cluster Analysis, in: 1st IEEE International Conference on Computational Advances in Bio and medical Sciences (ICCABS). Feb. 3-5, 2011, Orlando, FL, 2011.

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

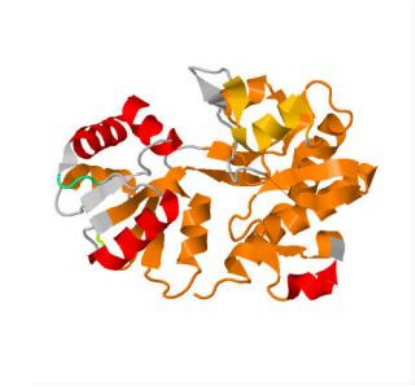
Almost all hydrogen bonds were critical.

The majority of hydrophobic interactions were not critical.



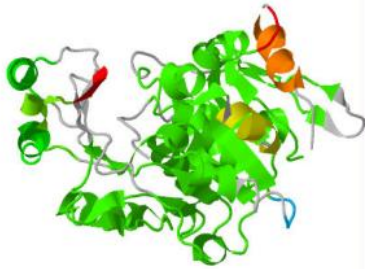
1A2P

1EX6



1FTO

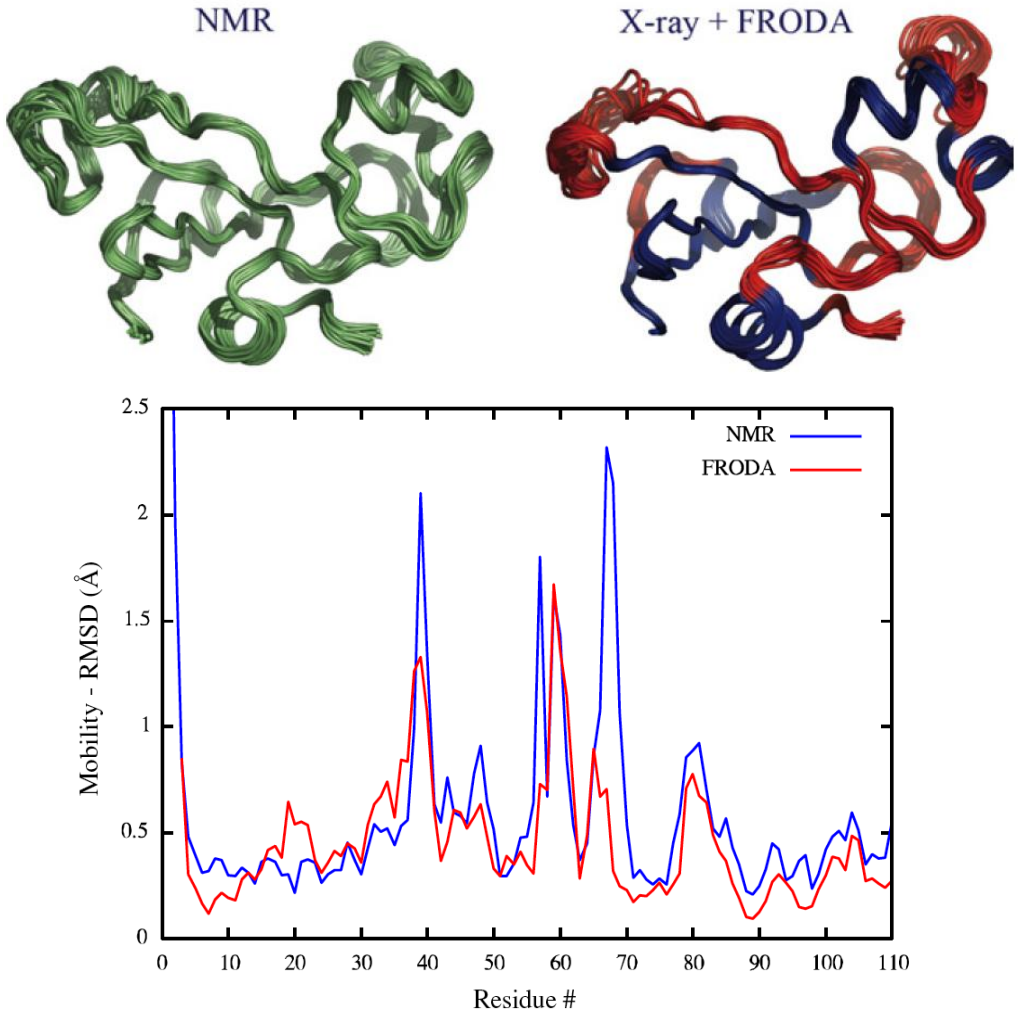
1KKL



1LIO

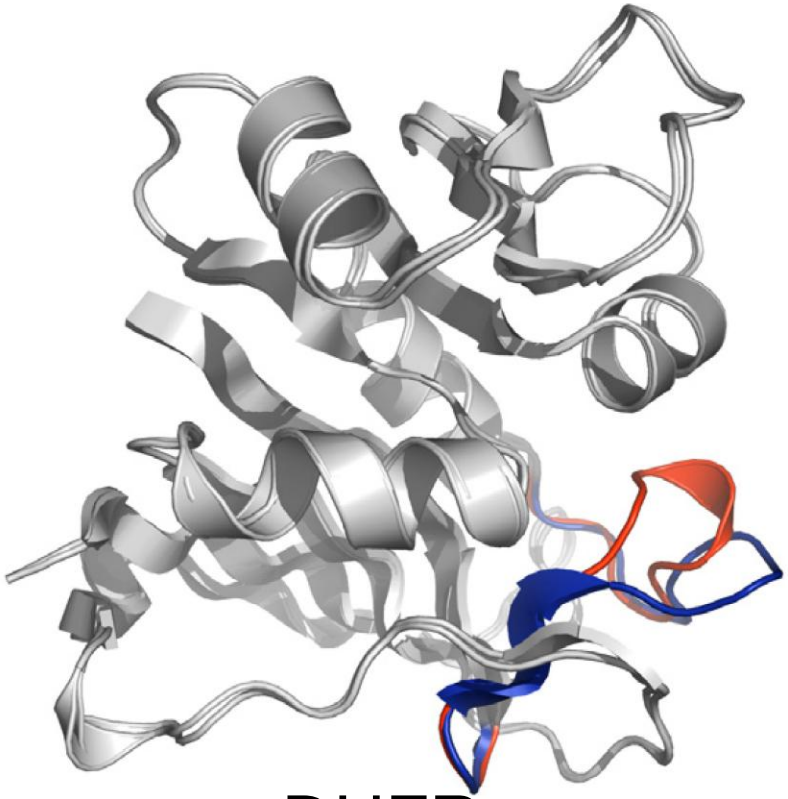
PDB	Cluster	Atoms	H-Bonds redun / all	HPh-inter. redun / all	Redun. Score
1FTO	1	1648	0 / 121	59 / 91	0.15
	2	203	0 / 16	0 / 6	0.00
	3	95	1 / 10	0 / 0	0.10
1A2P	1	903	0 / 74	24 / 45	0.10
	2	62	0 / 4	3 / 4	0.21
	3	32	0 / 2	3 / 4	0.33
1J4P	1	1302	0 / 44	159 / 197	0.52
	2	79	0 / 8	1 / 1	0.05
	3	47	0 / 2	5 / 5	0.50
1HKA	1	1193	0 / 91	39 / 65	0.13
	2	128	0 / 9	3 / 6	0.11
	3	30	0 / 1	2 / 3	0.36
1KKL	1	1157	0 / 74	44 / 80	0.17
	2	148	0 / 17	0 / 1	0.00
	3	129	0 / 15	1 / 2	0.03
1CBU	1	1152	0 / 91	47 / 67	0.16
	2	55	0 / 5	0 / 0	0.00
	3	34	0 / 3	0 / 0	0.00
1EX6	1	879	0 / 51	47 / 70	0.24
	2	618	5 / 41	33 / 45	0.31
	3	125	0 / 15	0 / 0	0.00
1DVR	1	470	0 / 32	32 / 37	0.27
	2	203	0 / 20	4 / 8	0.07
	3	191	0 / 16	10 / 13	0.19
1FIN	1	1881	0 / 124	103 / 143	0.23
	2	491	0 / 28	16 / 35	0.15
	3	94	0 / 11	0 / 0	0.00
1LIO	1	2451	0 / 88	268 / 336	0.48
	2	223	0 / 10	30 / 33	0.52
	3	72	0 / 6	2 / 2	0.12
1CMK	1	2996	1 / 207	133 / 199	0.19
	2	120	0 / 9	3 / 5	0.11
	3	77	0 / 5	0 / 3	0.00
16PK	1	3680	3 / 300	118 / 206	0.13
	2	148	0 / 16	0 / 2	0.00
	3	76	0 / 7	0 / 0	0.00

Use rigid clusters to generate motion for Barnase (110 residue)

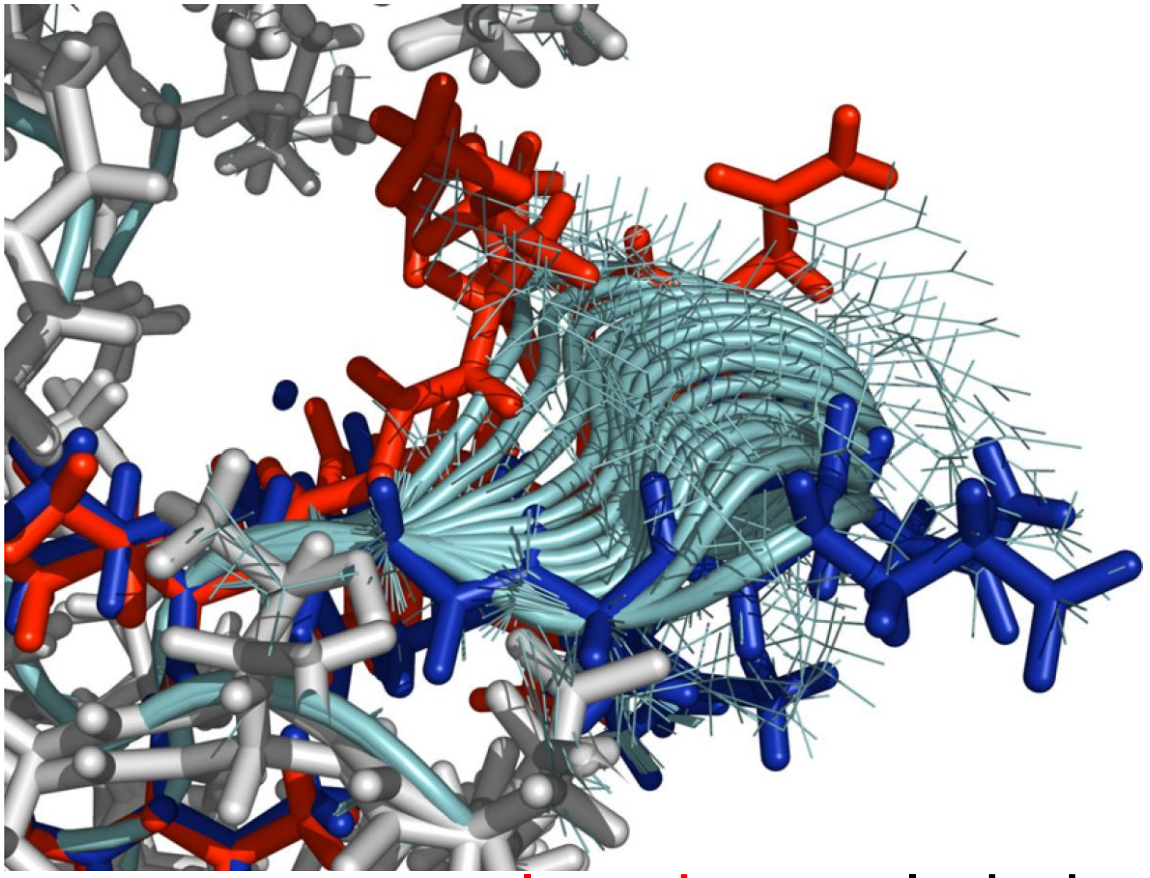


Stephen Wells, Scott Menor, Brandon Hespenheide, and MF Thorpe. Constrained geometric simulation of diffusive motion in proteins. Physical Biology 2:S127-S136, 2005.

FRODA Directional Motion



DHFR

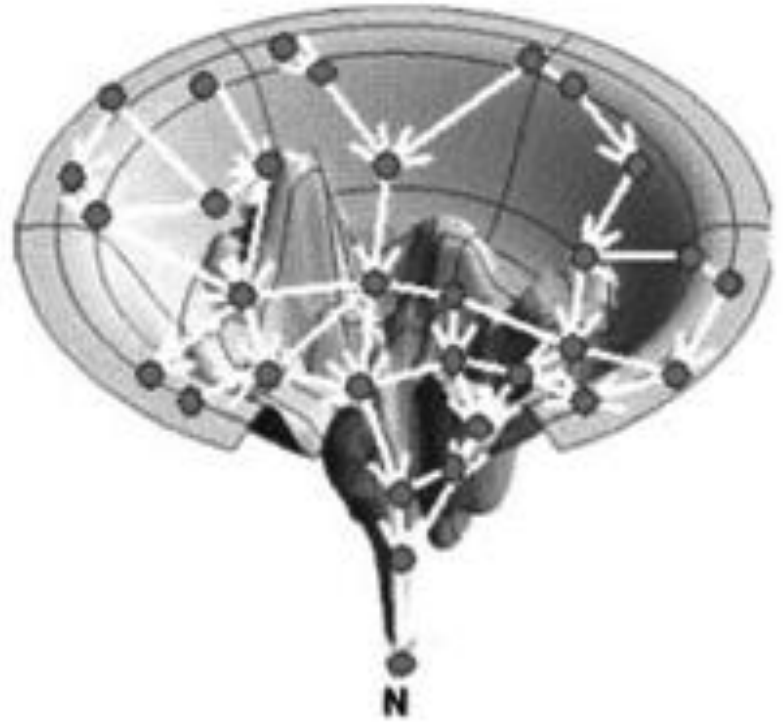


closed to occluded

The enzymatic protein has three very similar conformations (open, closed, and occluded), most significant differences in residues 14-24 (M-20 loop)/Mobile loop in closed (red) and occluded (blue) positions.

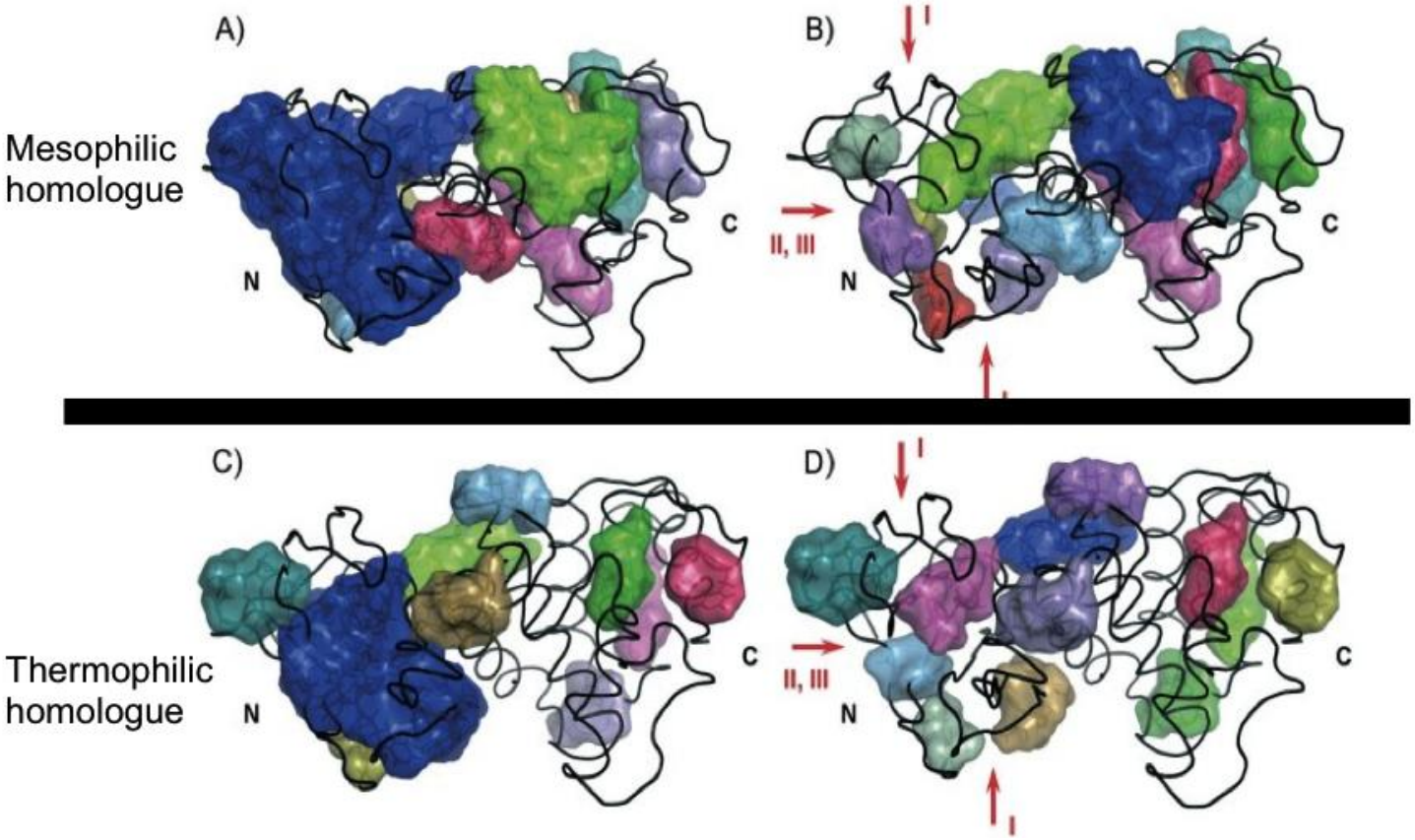
Stephen Wells, Scott Menor, Brandon Hespenheide, and MF Thorpe. Constrained geometric simulation of diffusive motion in proteins. Physical Biology 2:S127-S136, 2005.

- Extension of Probabilistic Roadmap algorithm
- Models protein backbone as a linkage, and samples different conformations
- Uses rigidity analysis to restrict conformers sampled



Thomas, Shawna, Tang, Xinyu, Tapia, Lydia and Amato, Nancy M., Simulating Protein Motions with Rigidity Analysis, in: Journal Of Computational Biology, volume 14, number 6, pages 839-855, 2007.

Do mesophile/thermophile homologs have corresponding states during unfolding?



Thermolysin -like protein (TLP)

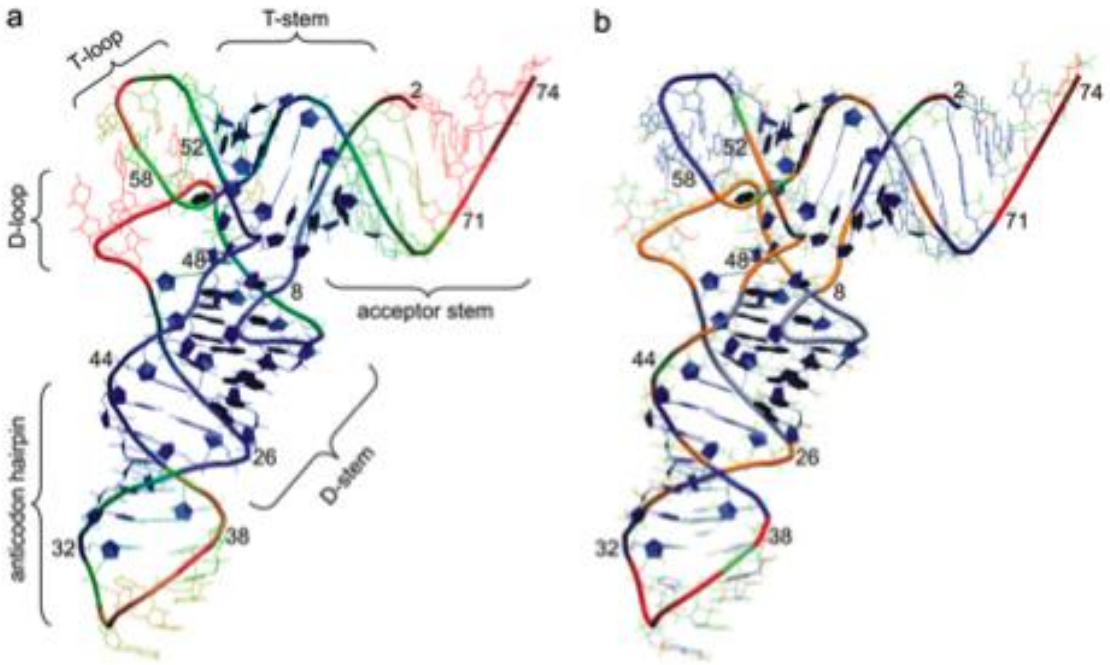
1npc

1kei

Gohlke, Holger and Radestock, S., Exploiting the link between protein rigidity and thermostability for data-driven protein engineering, in: Eng. Life Science, volume 8, pages 507-522, 2008.

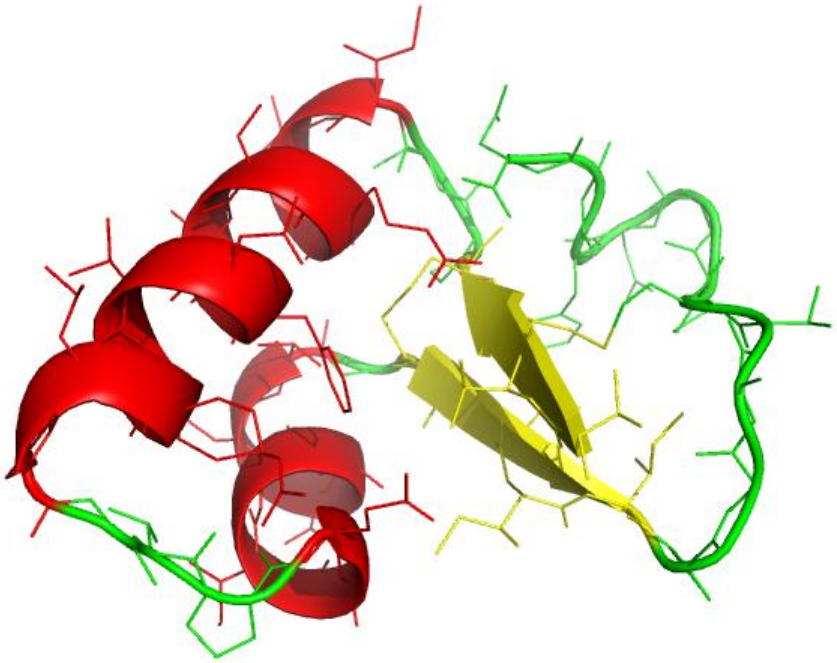
- Different parameterization required for calculating hydrophobics
- Found correspondence between B-values and flexibility index

Colored by B-value



Colored by flexibility index

Fulle, Simone and Gohlke, Holger, Analyzing the Flexibility of RNA Structures by Constraint Counting, in: Biophysics Journal, volume 94, pages 4202-4219, 2008.



Crambin (PDB file 1CRN)

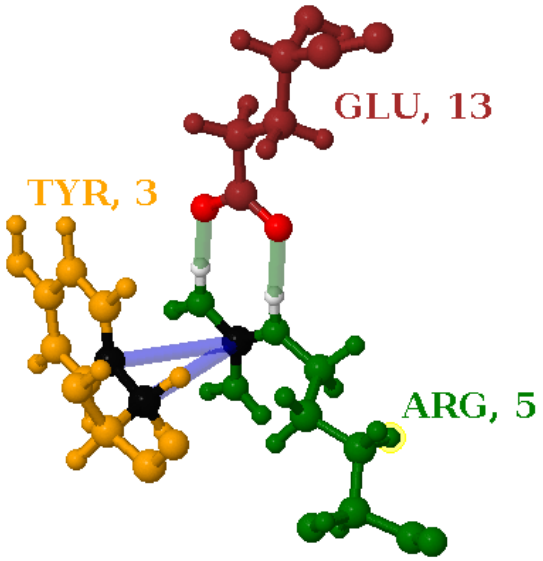
Crambin is a 46 amino acid plant protein, whose crystals diffract to ultra-high resolution.

Can we use Rigidity Analysis to infer critical residues of crambin?

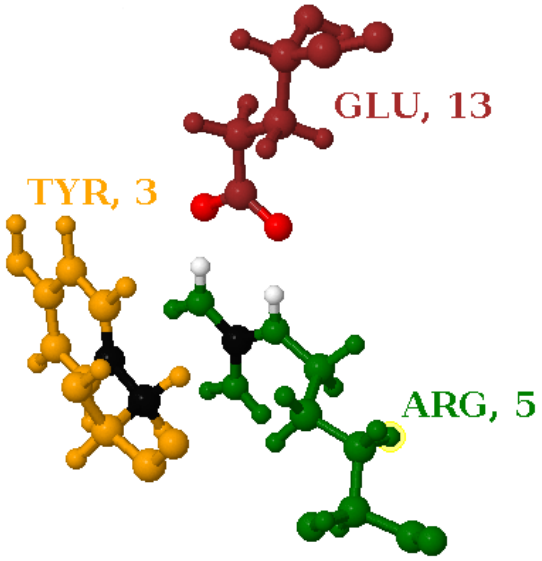
The insight: “mutate” different residues to generate *in-silico* mutants, and perform rigidity analysis on them, to infer which residues, IF mutated, would destabilize the protein, and hence are important

Filip Jagodzinski, Jeanne Hardy, Ileana Streinu, Using Rigidity Analysis To Probe Mutation-Induced Structural Changes in Proteins, CSBW 2011.

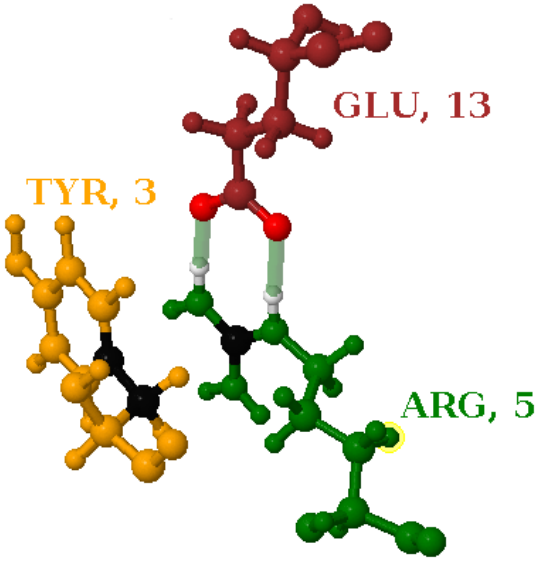
To simulate a substitution to a glycine, all hydrogen bonds and hydrophobic interactions for a residue are removed from the molecular model; this effectively removes the side chain from the molecular model because it cannot contribute stabilizing interactions.



Wild-type α-defensin 1 (PDB File 2PM1) has two hydrogen bonds (light green bars) and two hydrophobic interactions (blue bars) among residues 3, 5, and 13



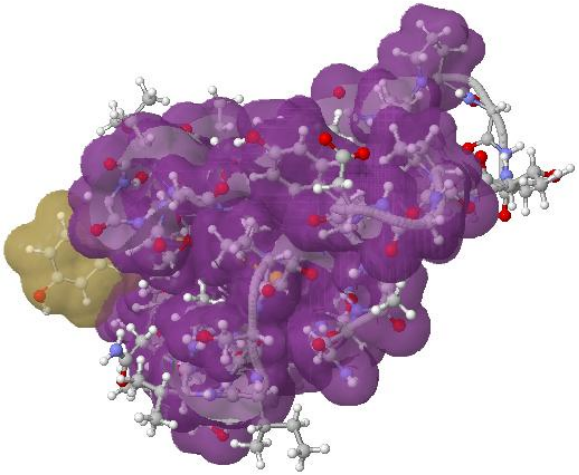
Mutating residue 5 removes hydrogen bonds between residue 5 and 13 and hydrophobic interactions between residue 5 and 3



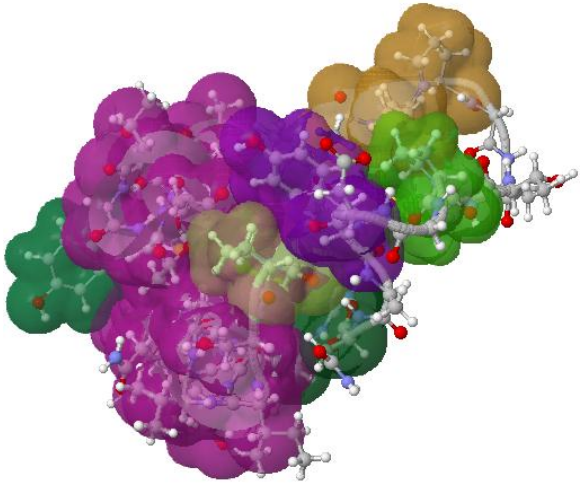
Mutating residue 3 removes hydrophobic interactions that it forms with residue 5

Filip Jagodzinski, Jeanne Hardy, Ileana Streinu, Using Rigidity Analysis To Probe Mutation-Induced Structural Changes in Proteins, CSBW 2011.

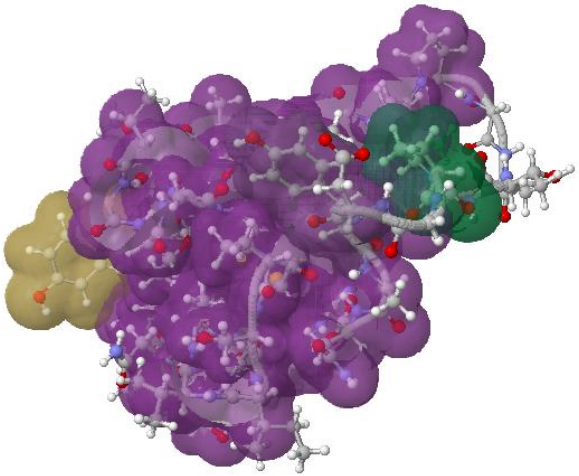
We can compare the rigidity results of mutants to infer which mutations disrupt the protein's rigidity, and hence which residues are critical



The while type of the crambin contains one large rigid cluster (purple)



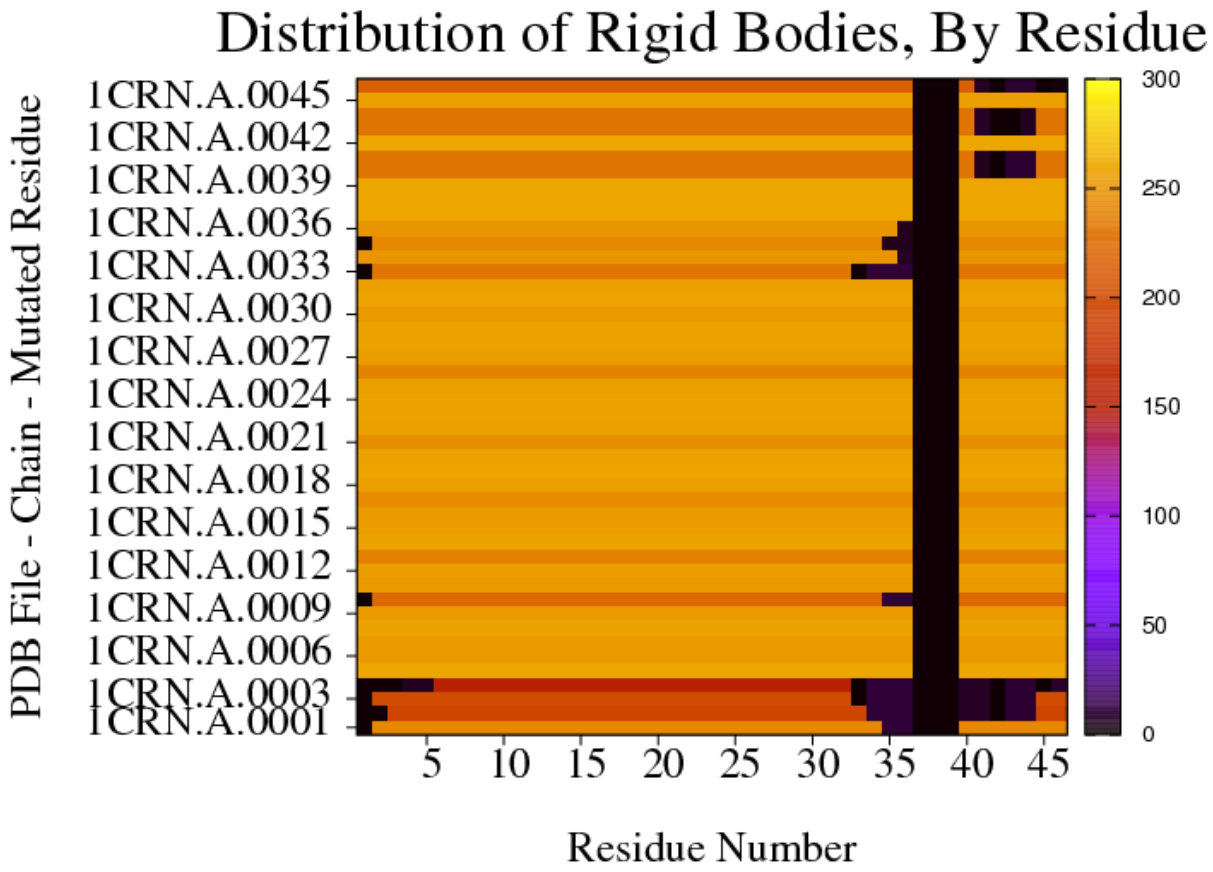
Mutating residue 4 affects the rigid bodies of crambin



Mutating Residue 40 causes the largest rigid body to break down, but not as much as when residue 4 is mutated

Filip Jagodzinski, Jeanne Hardy, Ileana Streinu, Using Rigidity Analysis To Probe Mutation-Induced Structural Changes in Proteins, CSBW 2011.

Rigidity results for different mutants can be compared to infer which residue(s) is (are) critical.



Filip Jagodzinski, Jeanne Hardy, Ileana Streinu, Using Rigidity Analysis To Probe Mutation-Induced Structural Changes in Proteins, CSBW 2011.

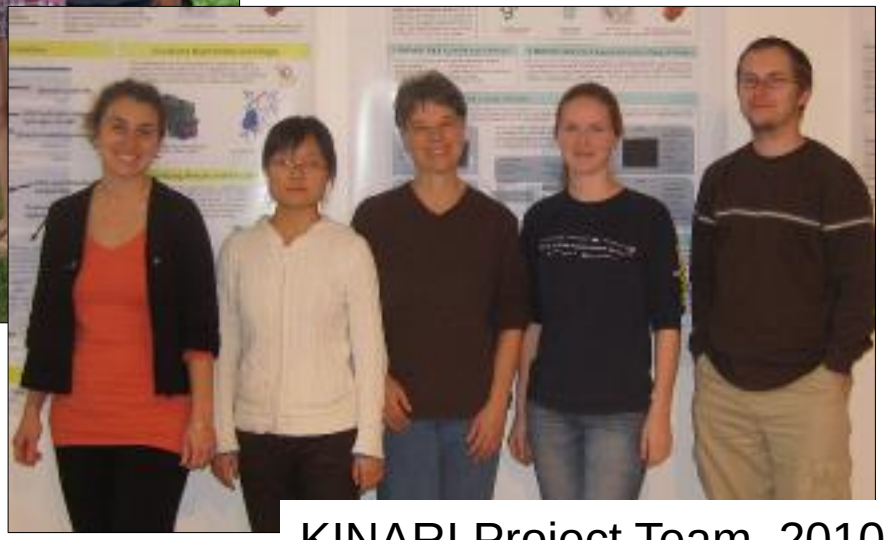
KINARI Mutagen – Correlating results to Web-Lab Experiments and Other Metrics

Filip Jagodzinski, Jeanne Hardy, Ileana Streinu, Using Rigidity Analysis To Probe Mutation-Induced Structural Changes in Proteins, CSBW 2011.

Residues 3, 4, 40, and 41 (among others), have been found to be identical among Crambin and two homologous plant toxins viscotoxin A3 and α1 purothioni

M.M. Teeter, J.A. Mazer, and J.J. L'Italien. Primary structure of the hydrophobic plant protein crambin. Biochemistry, 20(19):5437–5443, 1981.

Linkage Lab, Summer 2011



KINARI Project Team, 2010

UMASSCS
DEPARTMENT OF COMPUTER SCIENCE

Contact
Ileana Streinu, Director
Linkage Laboratory
istreinu@smith.edu
streinu@cs.smith.edu
streinu@cs.umass.edu

http://linkage.cs.umass.edu
http://kinari.cs.umass.edu

Funding
NSF CCF-1016988
CCF-0728783, DMS-0714934
DARPA 23 Mathematical Challenges