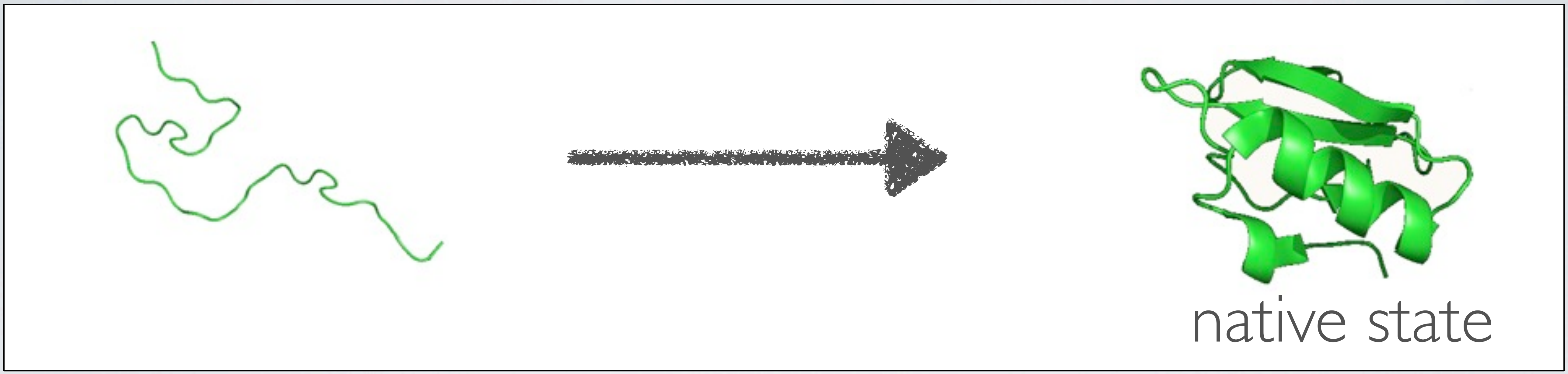
Monte Carlo simulation of proteins: from folding to "nonfolding" to interactions

ATP Colloquium
Nov 2, 2011

Stefan Wallin

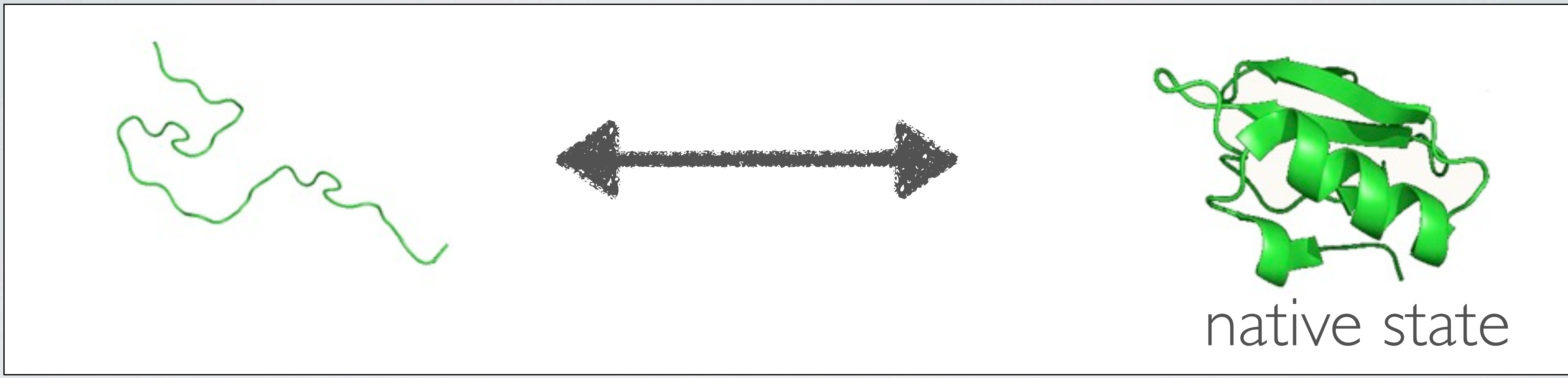
Protein folding



Given the amino acid sequence, what is the 3-dimensional structure of the protein?

Christian B. Anfinsen

Protein folding



Given the amino acid sequence, what is the 3-dimensional structure of the protein?

Christian B. Anfinsen

How does the protein chain find the native state?

153 amino acids

GLSDGEWQLVLNVWGKVEADIPGHGQEVLIRLFKGHPETLE
KFDKFKHLKSEDEMKASEDLKKHGATVLTALGGILKKKGHHE
AEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKHPGDFGADA
QGAMNKALELFRKDMASNYKELGFQG

Number of possible conformations:

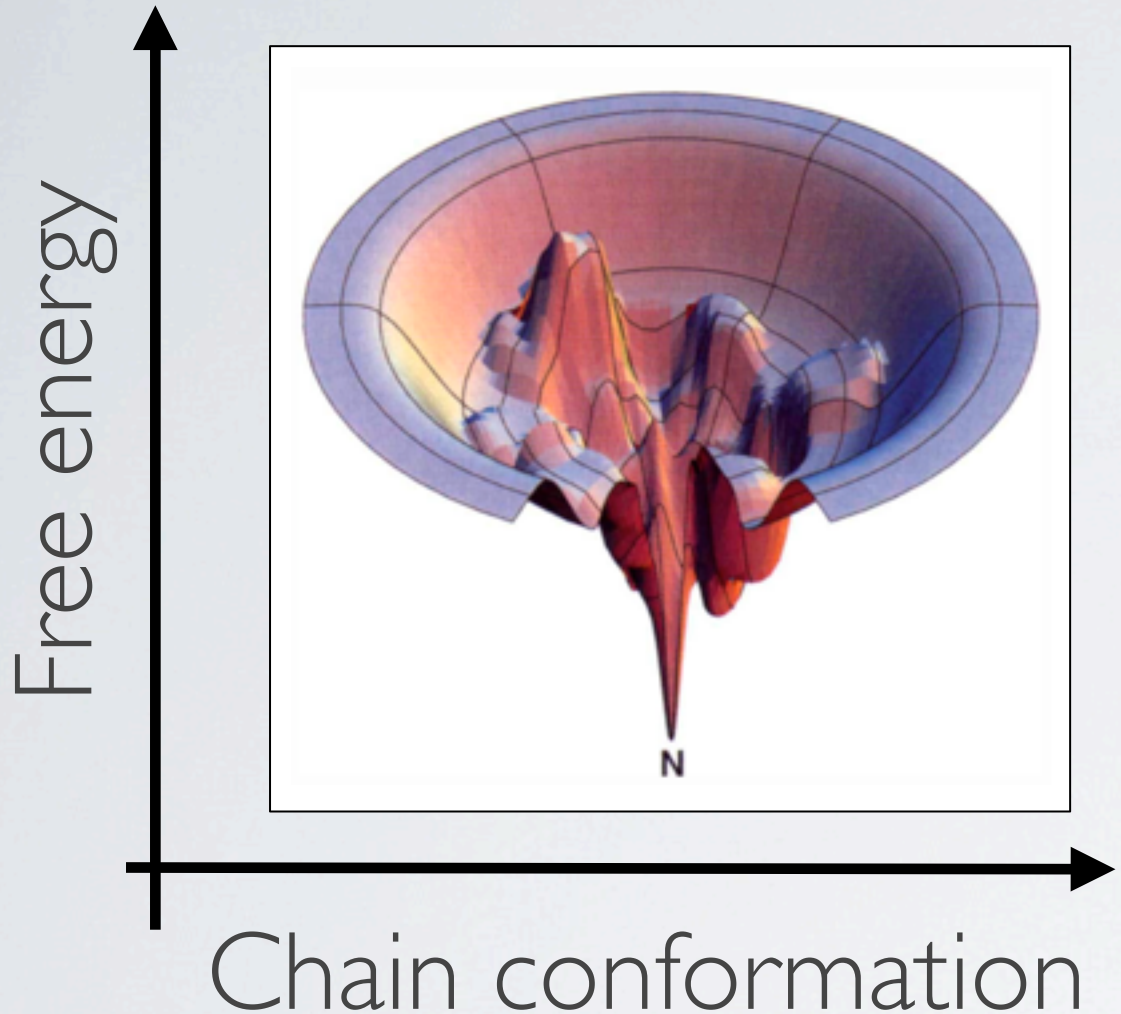
$3^{153} = 3.7 \times 10^{71}$

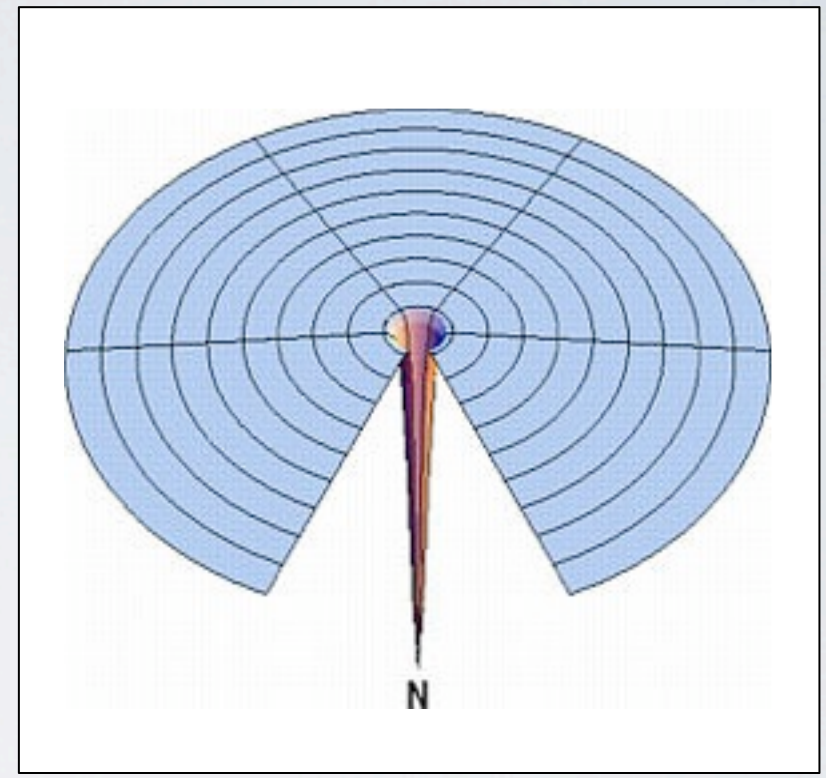
1ps per conformation:

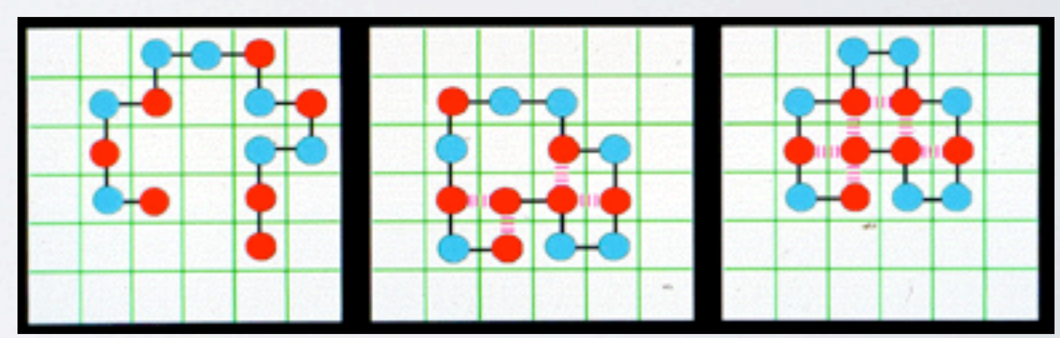
3.7×10^{56} s

"Levinthal's paradox" (1969)

Energy landscape view

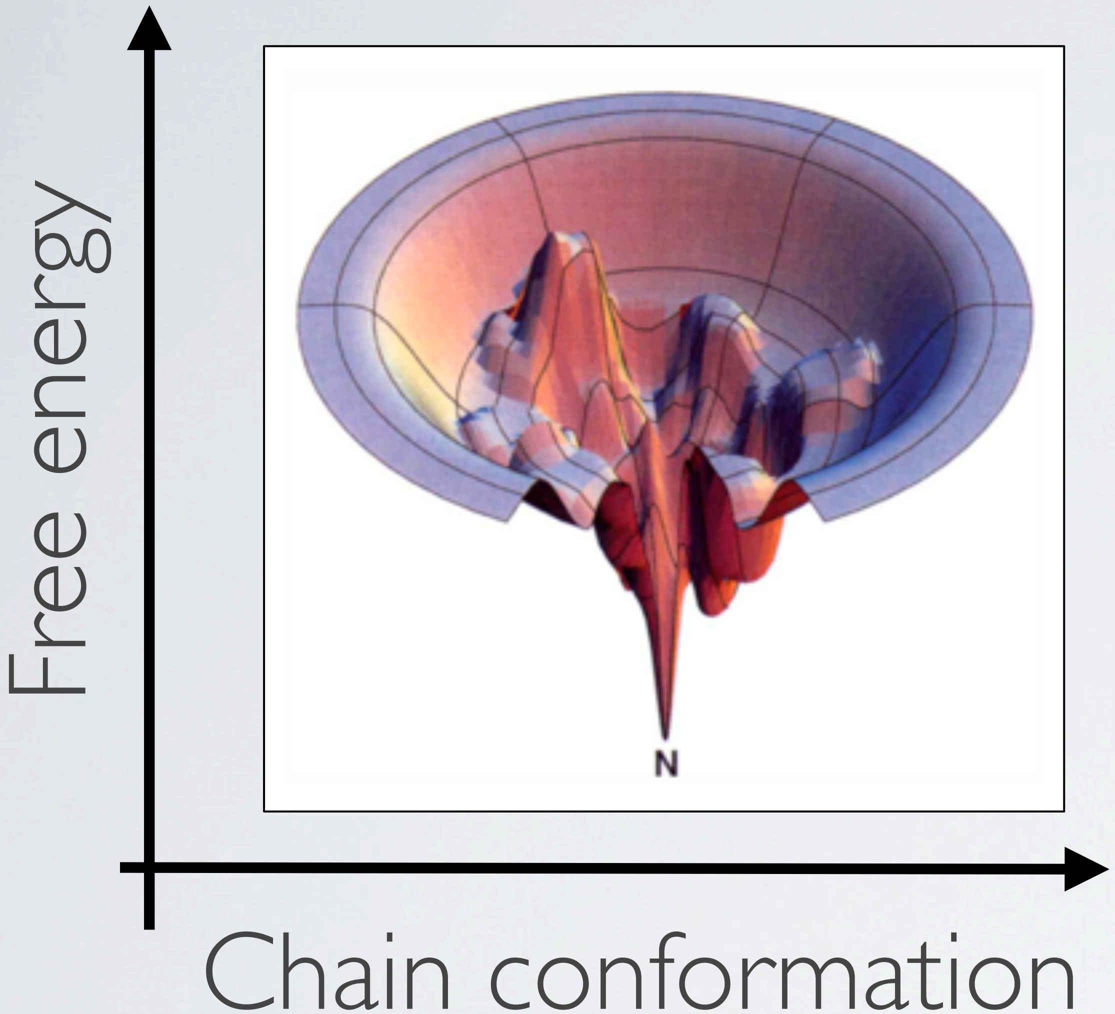


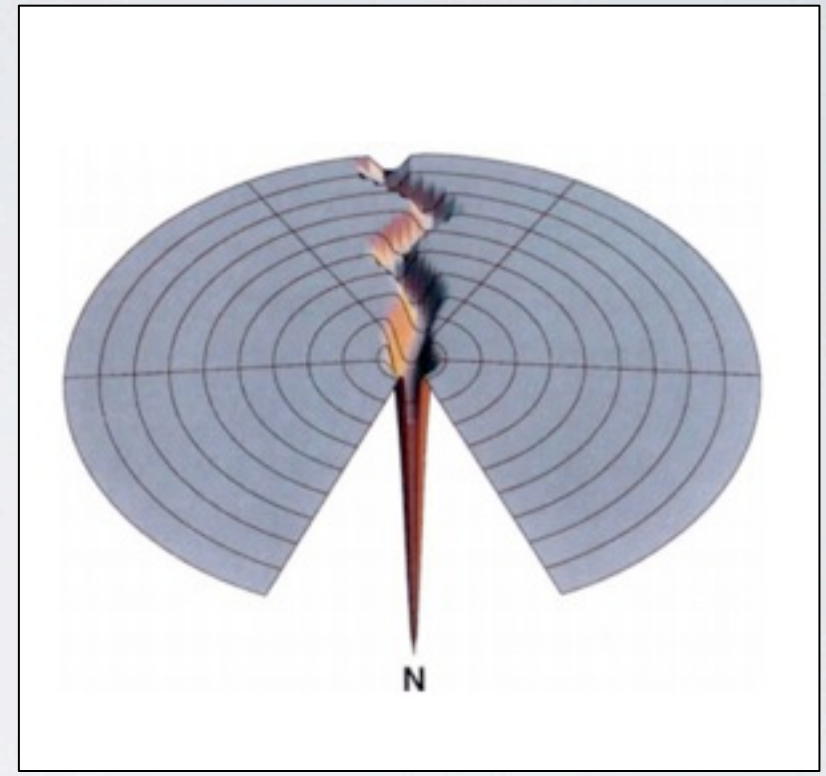


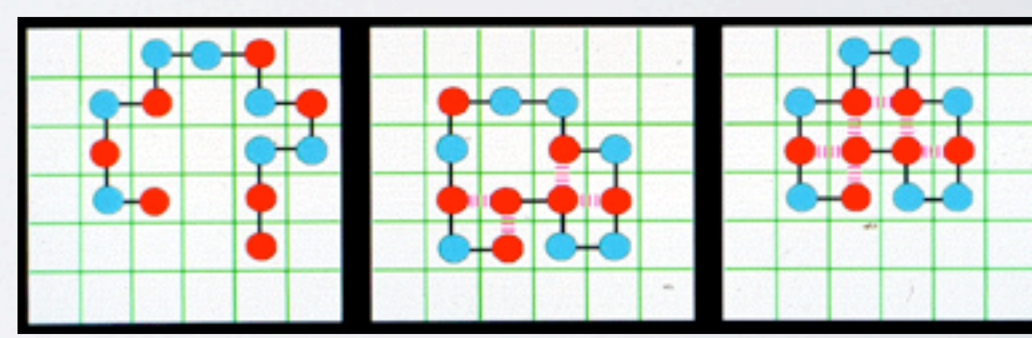


HP model (Lau and Dill, 1989)

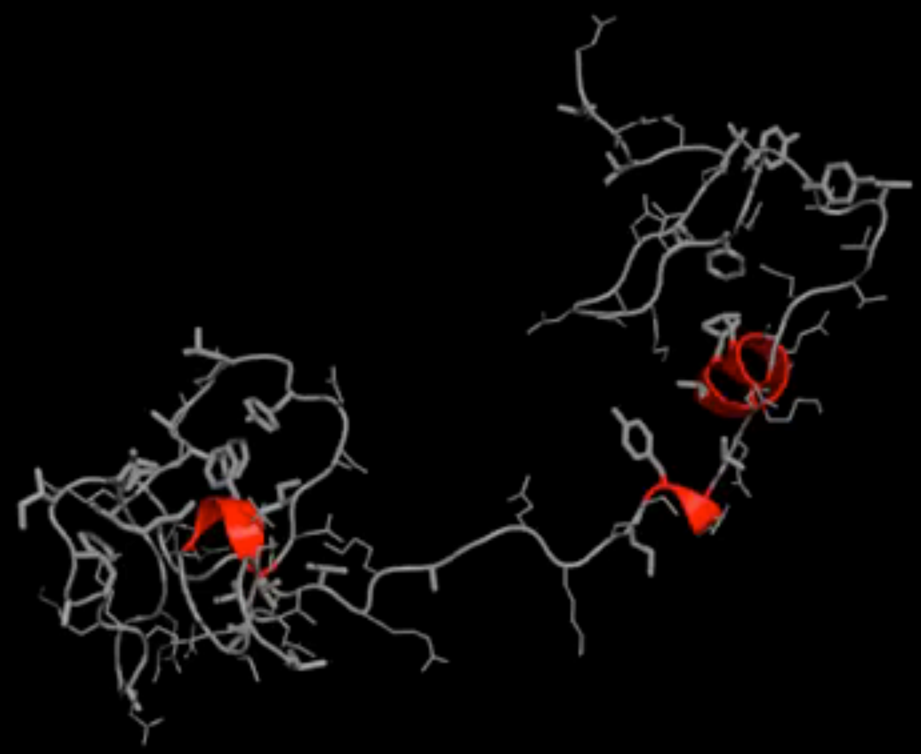
Energy landscape view







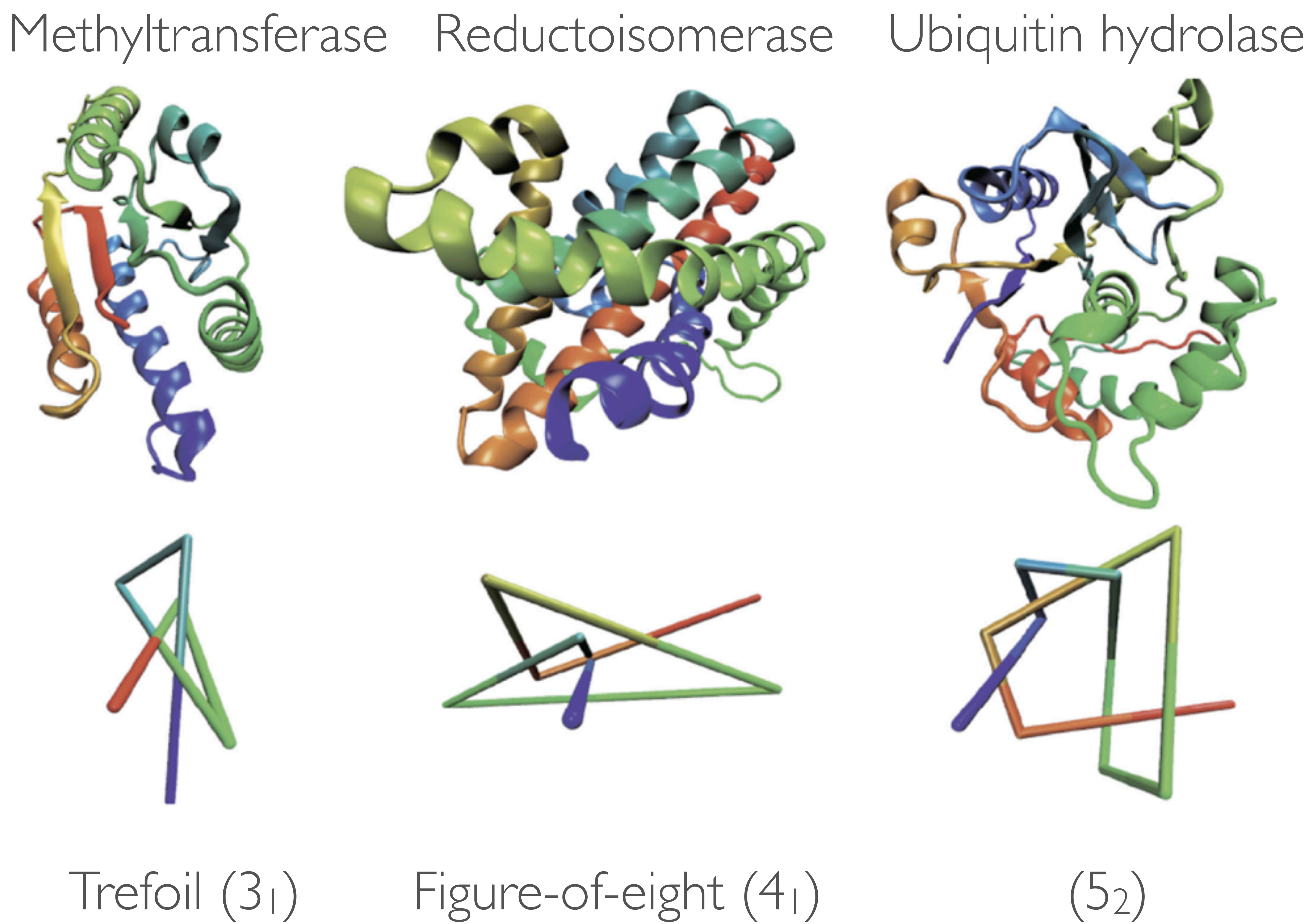
HP model (Lau and Dill, 1989)



Credit: Sandipan Mohanty, Jülich Supercomputing Centre

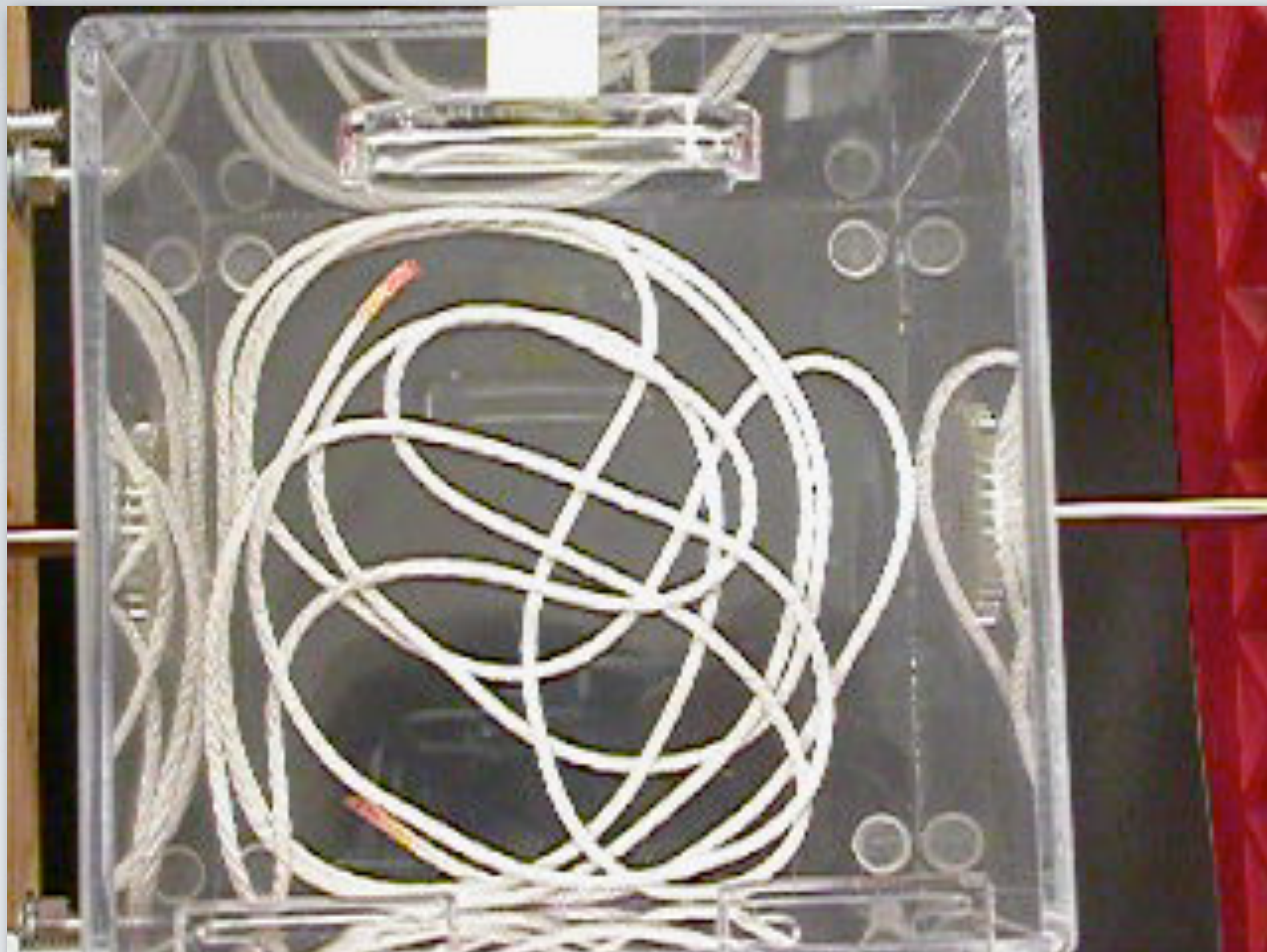
Credit: Sandipan Mohanty, Jülich
Supercomputing Centre

A new challenge in protein folding


Methyltransferase
Reductoisomerase
Ubiquitin hydrolase
Trefoil (3_1)
Figure-of-eight (4_1)
(5_2)

Spontaneous knotting of an agitated string

“Spontaneous knotting of an agitated string”, *Proceedings of the National Academy of Sciences USA*, D Raymer and D Smith (2007).

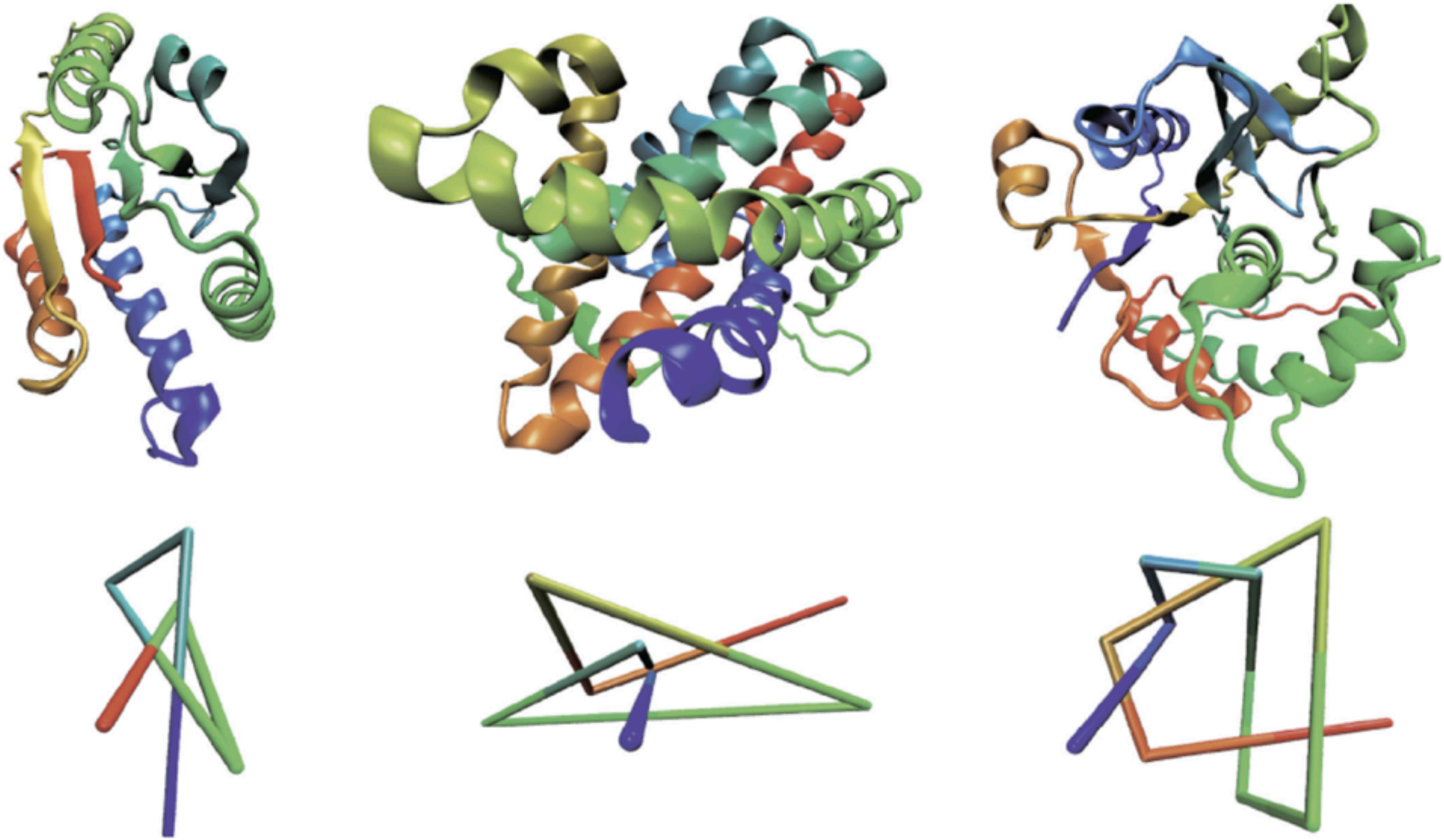


PNAS

Spontaneous knotting of an agitated string

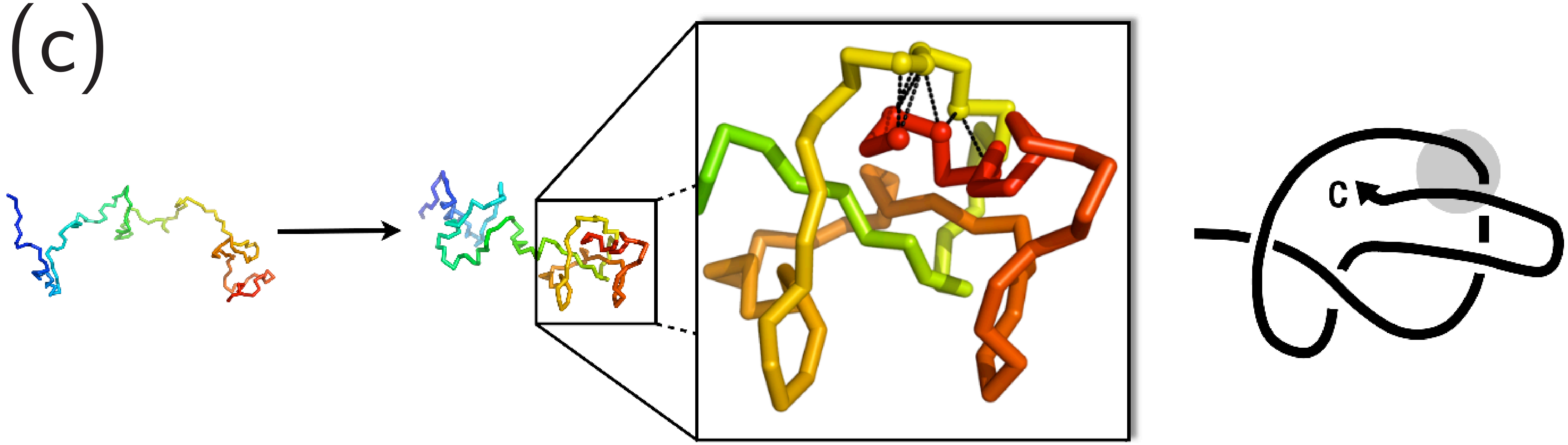
“Spontaneous knotting of an agitated string”, *Proceedings of the National Academy of Sciences USA*, D Raymer and D Smith (2007).

Methyltransferase Reductoisomerase Ubiquitin hydrolase

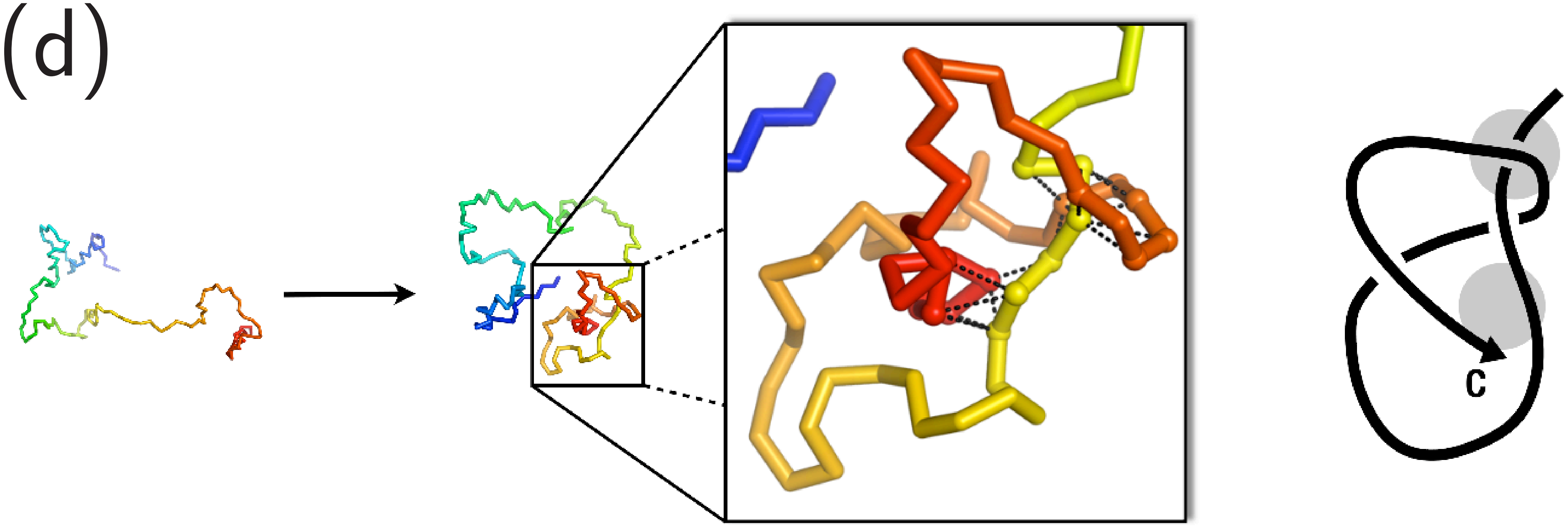


Trefoil (3_1) Figure-of-eight (4_1) (5_2)

(c)



(d)



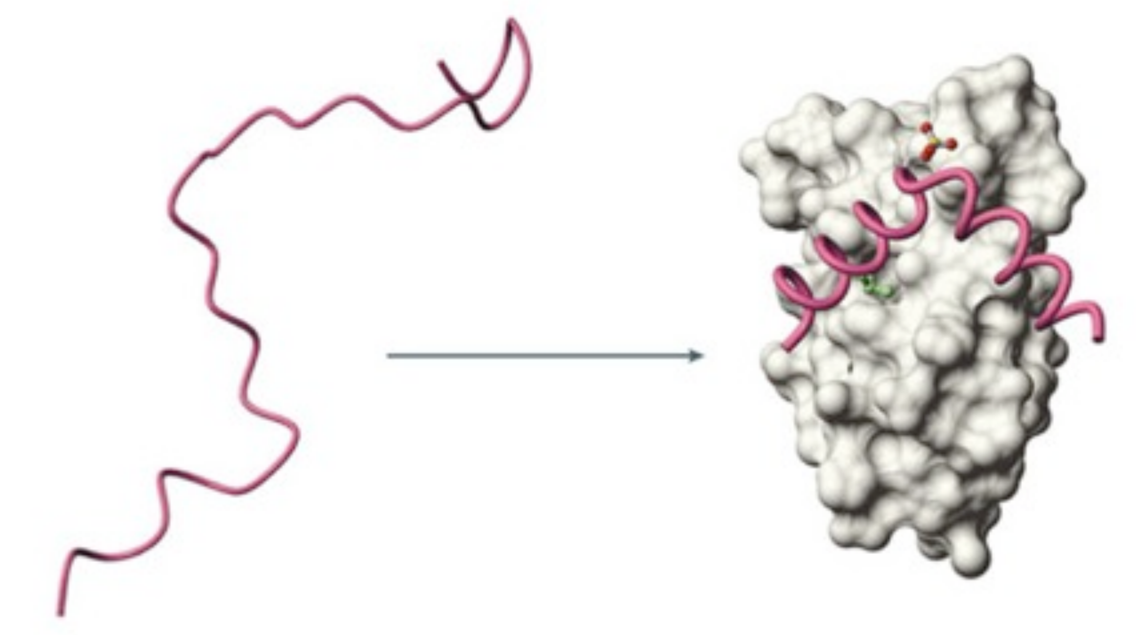
Chan *et al.* Annu Rev Chem Phys (2011)

Sequence → Structure → Function

Intrinsically disordered proteins

- No stable native state, yet fully functional

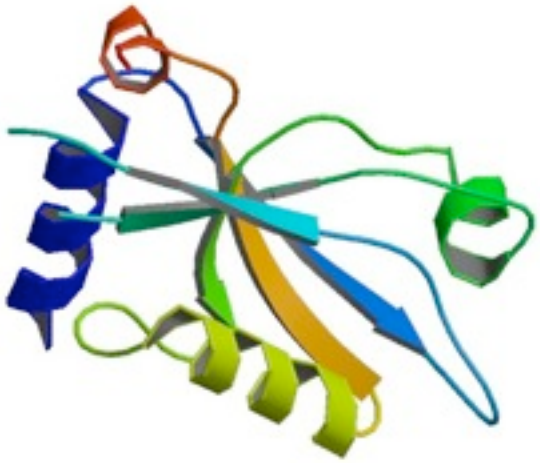
- Coupled folding-binding



pKID binding to the KIX domain of CREB, folding in the process.

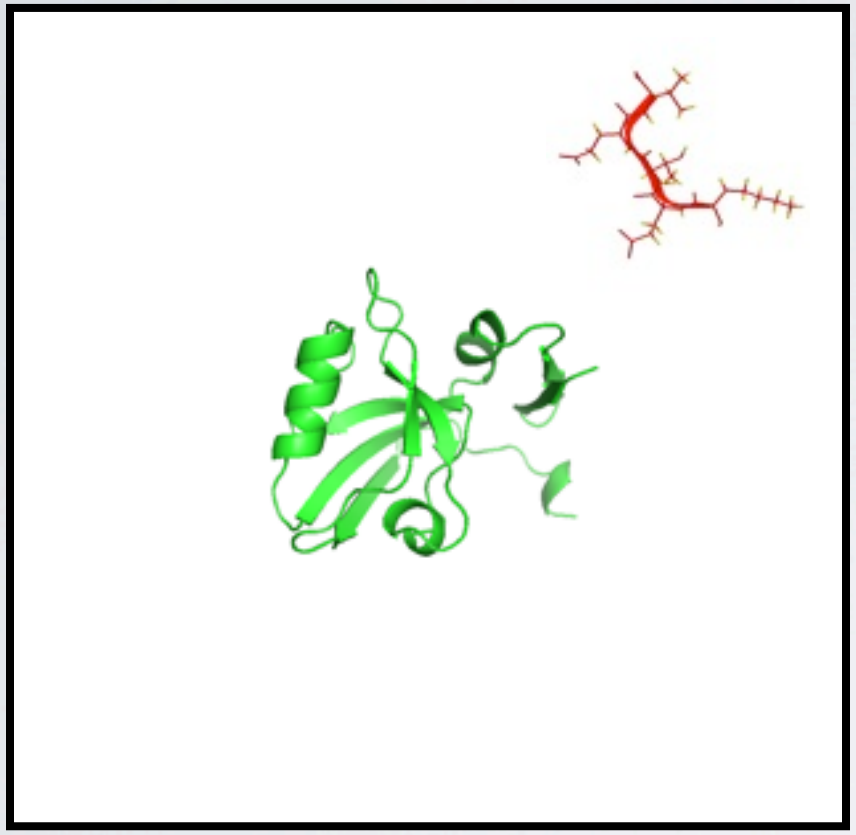
Protein-protein interaction networks:
up to 15-40% of all interactions due to protein-peptide binding

An unsolved problem

AQEASWF +  = ?

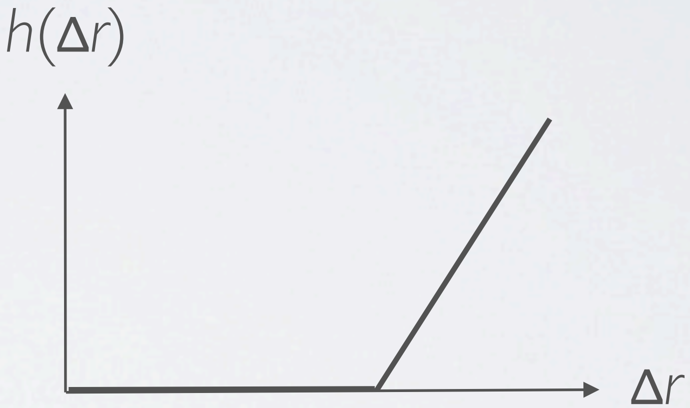
Given a protein structure and a peptide sequence, predict *whether* and *how* the protein can bind the peptide.

Our approach



Protein in native state:

$$E_{fix} = k_{fix} \sum_{C_\alpha} h(|\mathbf{r}_i^{PDB} - \mathbf{r}_i|)$$



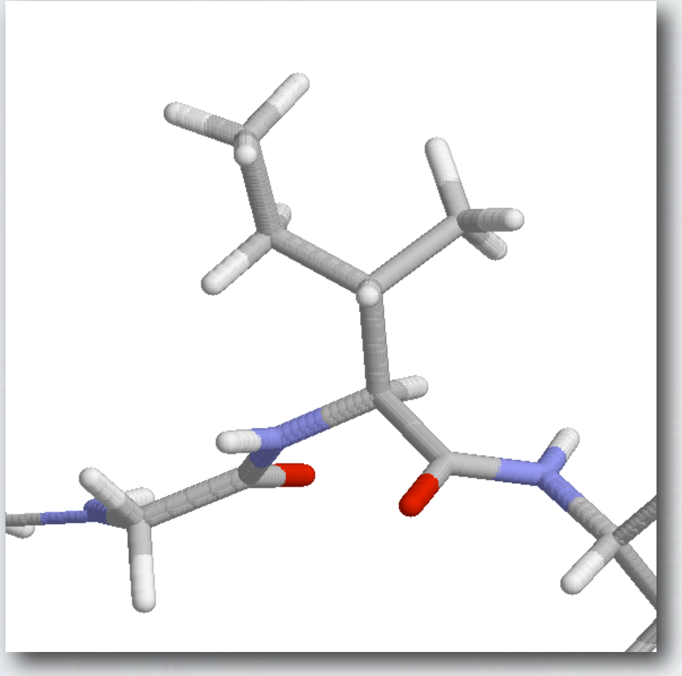
$r_c = 0.5$ Å

Peptide: no added constraints (full chain flexibility)

Box with periodic boundary conditions

Computational model

All protein atoms represented
but no explicit water



Effective energy function:
- excluded volume
- hydrogen bonding
- hydrophobic attraction
- electrostatic attractions

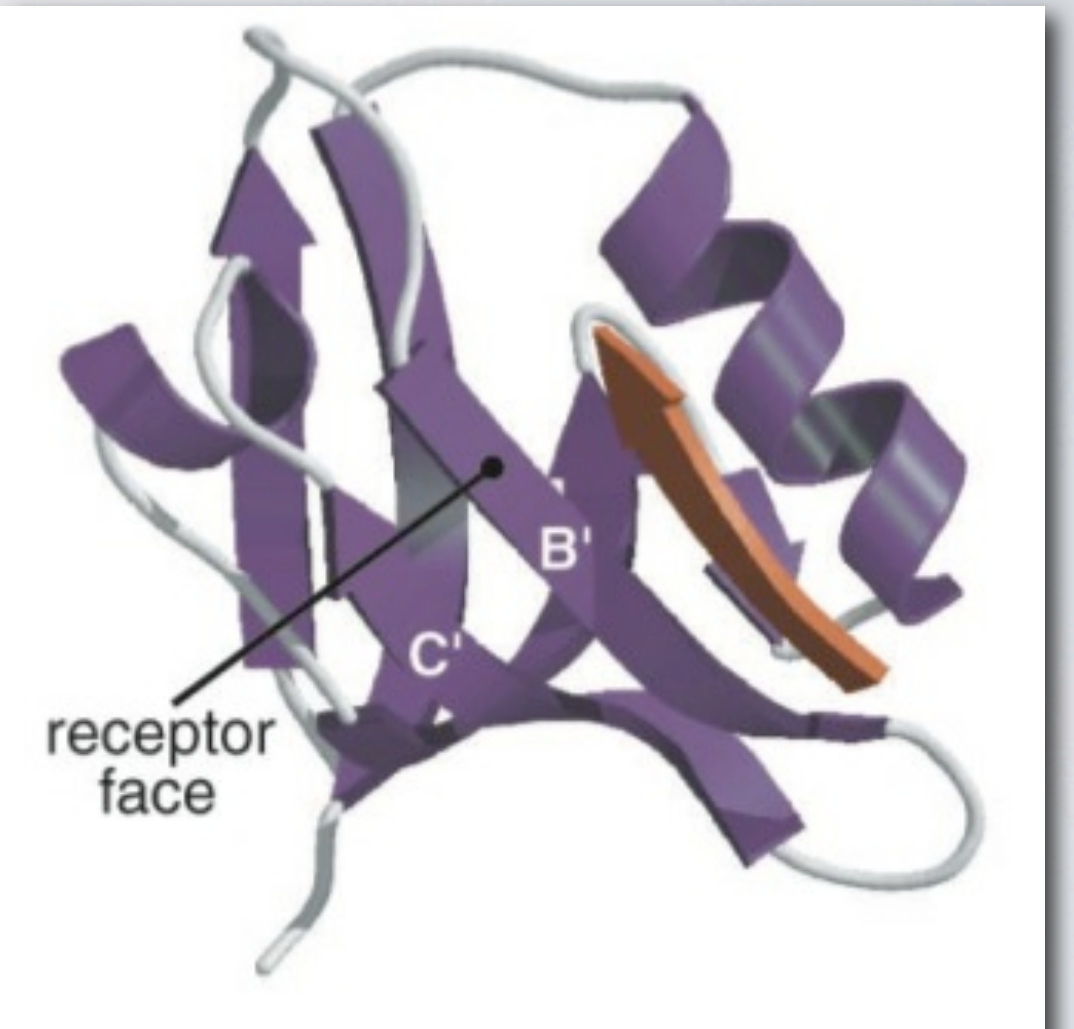
Sampling method:
- <u>Monte Carlo</u>
- Molecular Dynamics

"makes things as simple as possible, but not simpler"

PDZ domains

~70-90 amino acid domains

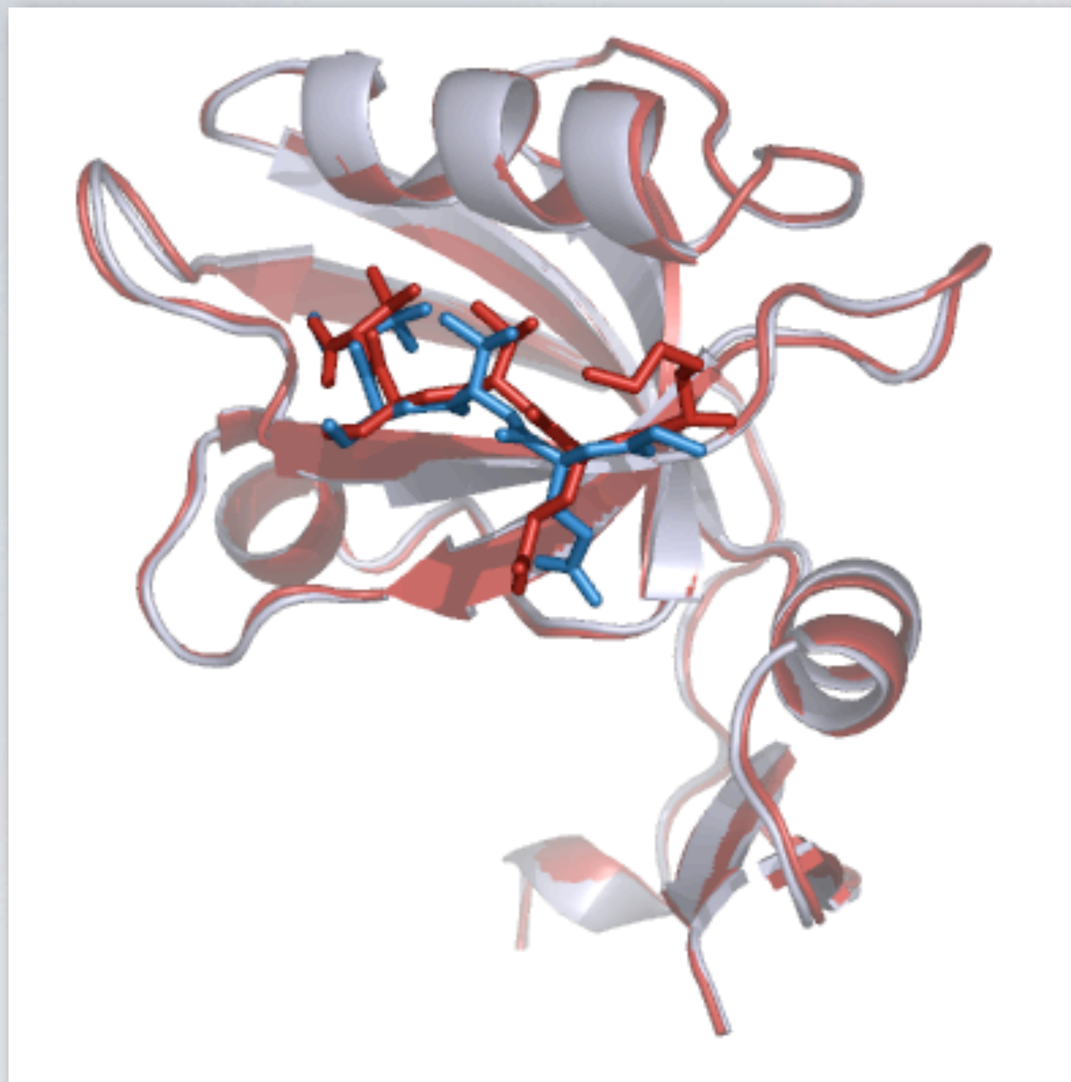
Domain	*Homo sapiens*	*Mus musculus*
PDZ	918	771
PTB	141	115
SH2	352	323
SH3	894	738

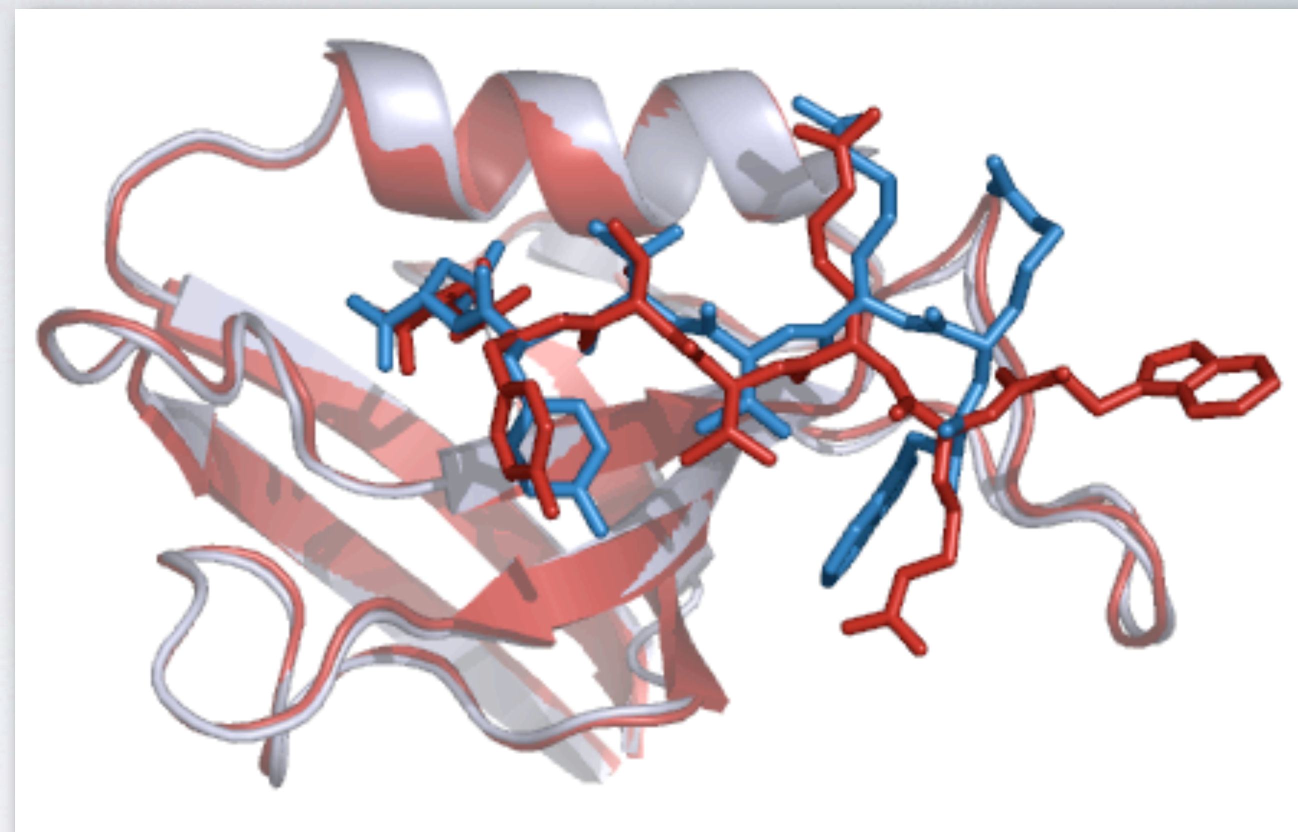


syntrophin - peptide

Peptide-binding groove
- binds C-terminal peptides

Common in signaling and regulatory processes



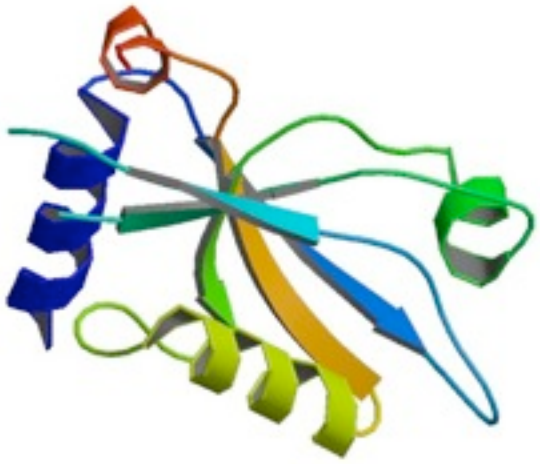


Minimum-energy conformation
X-ray structure

→ *binding mechanism*
→ *binding specificity and promiscuity*
→ *character of the bound state*

conformational changes in domain (allostery)

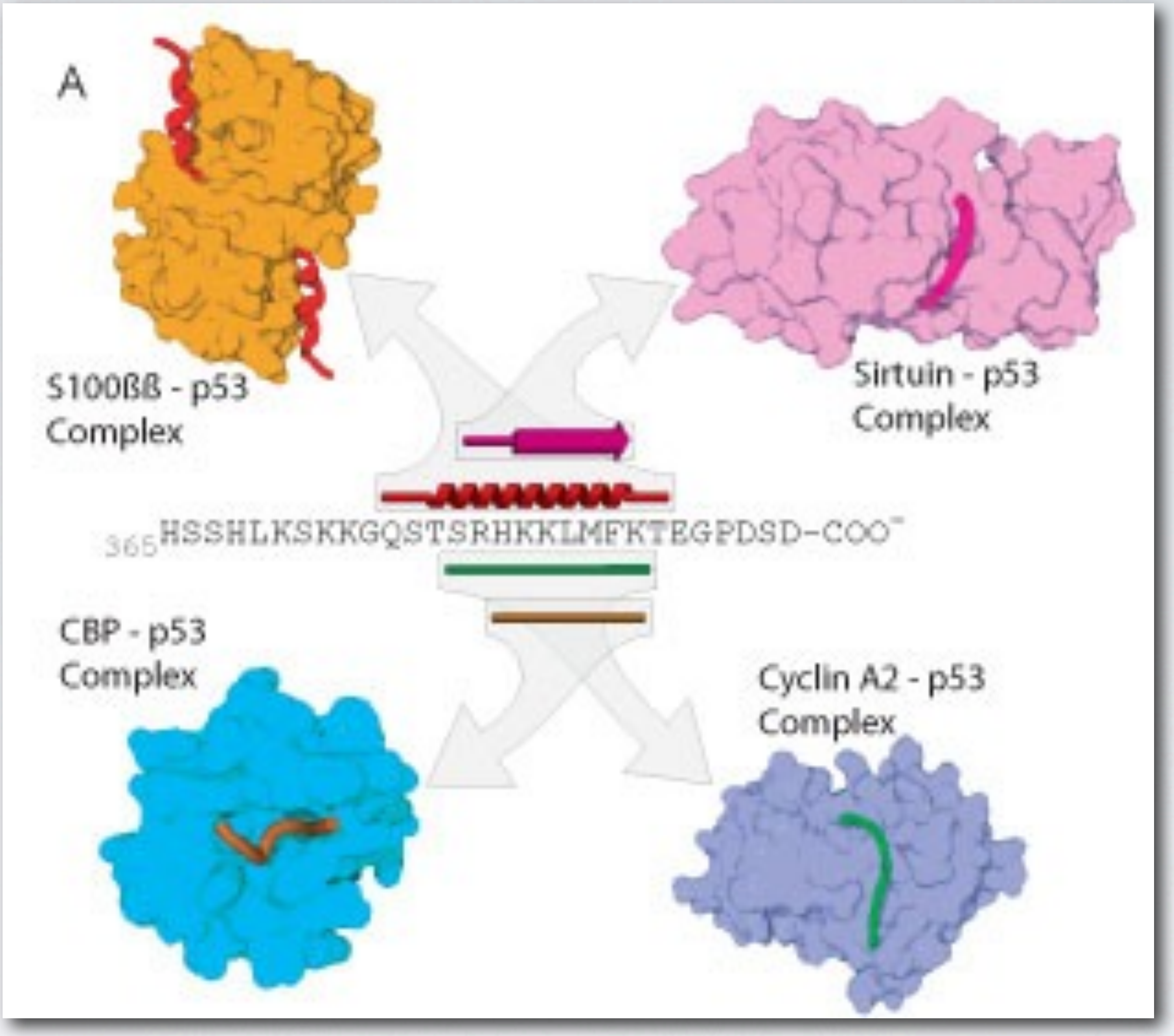
An unsolved problem

QEASWF +  = ?

Given a protein structure, predict *whether* and *how* a peptide sequence can bind to the protein.

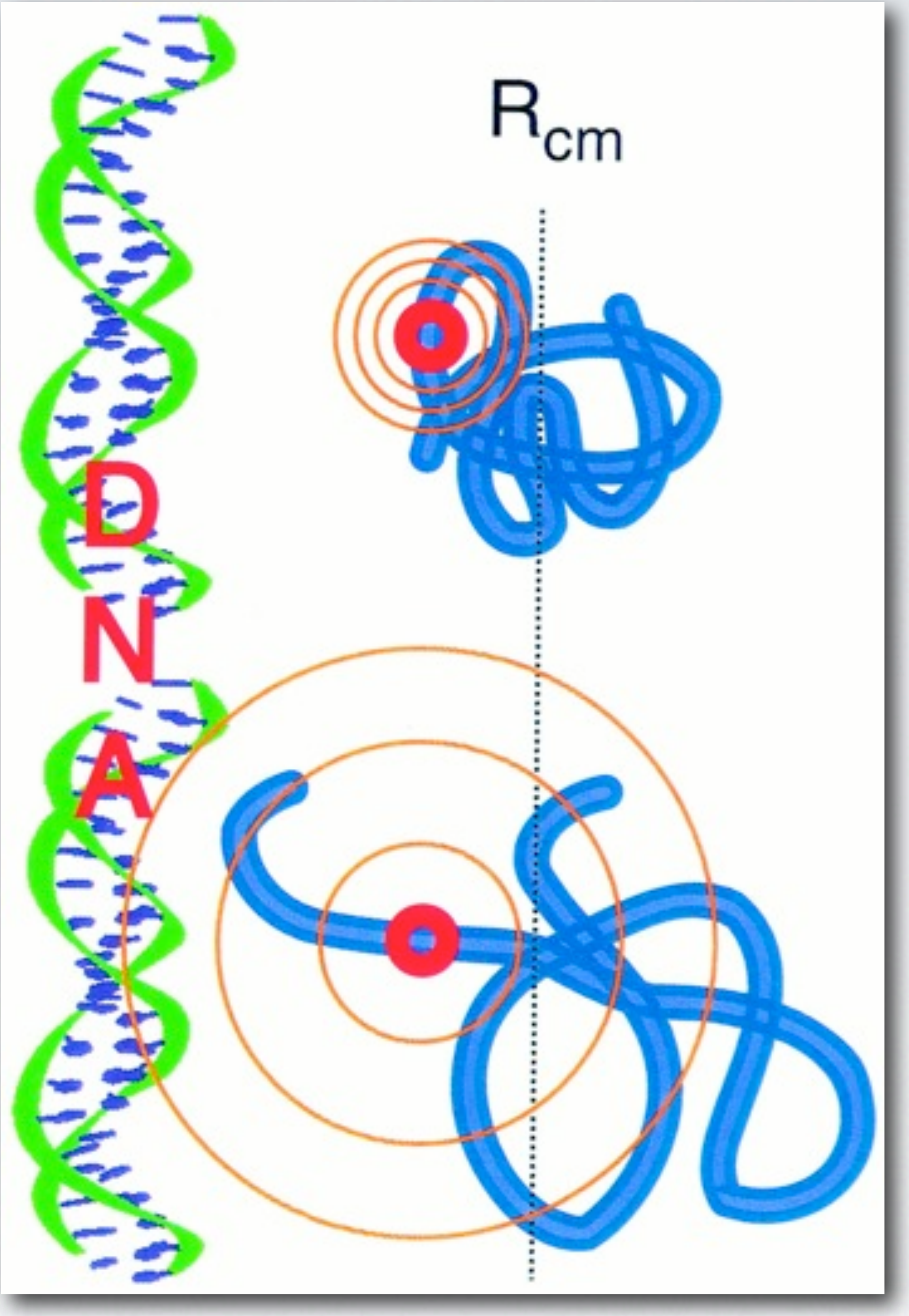
Biological advantages of intrinsic disorder

“Multi-specificity” of a p53 peptide



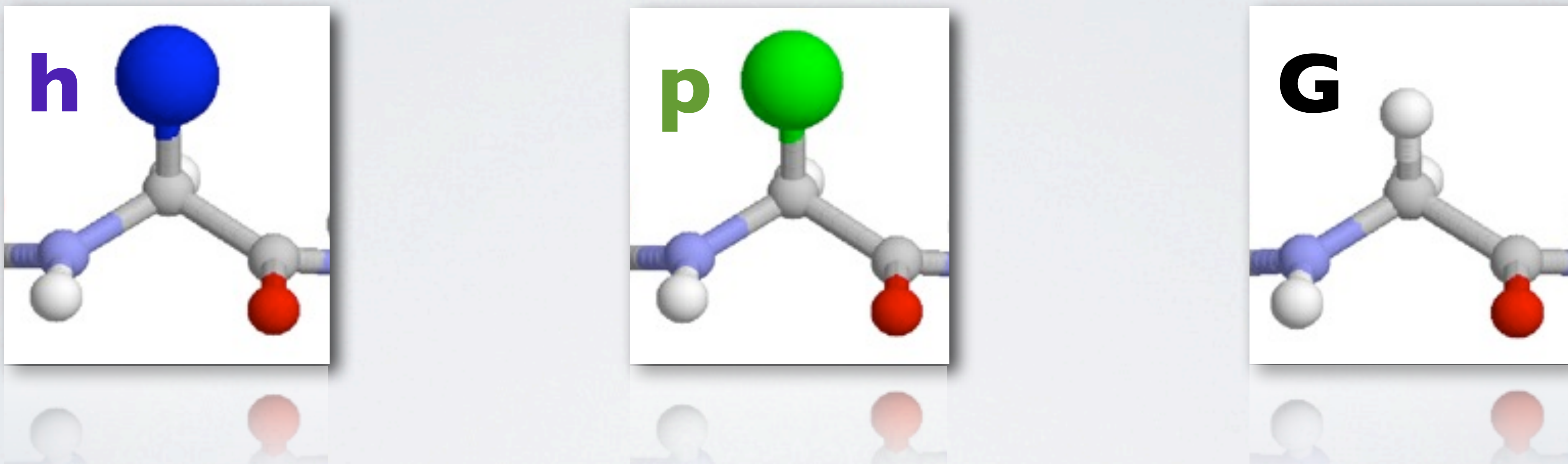
[*Iskra Staneva*, collaboration with Peking University]

A fly-casting mechanism?



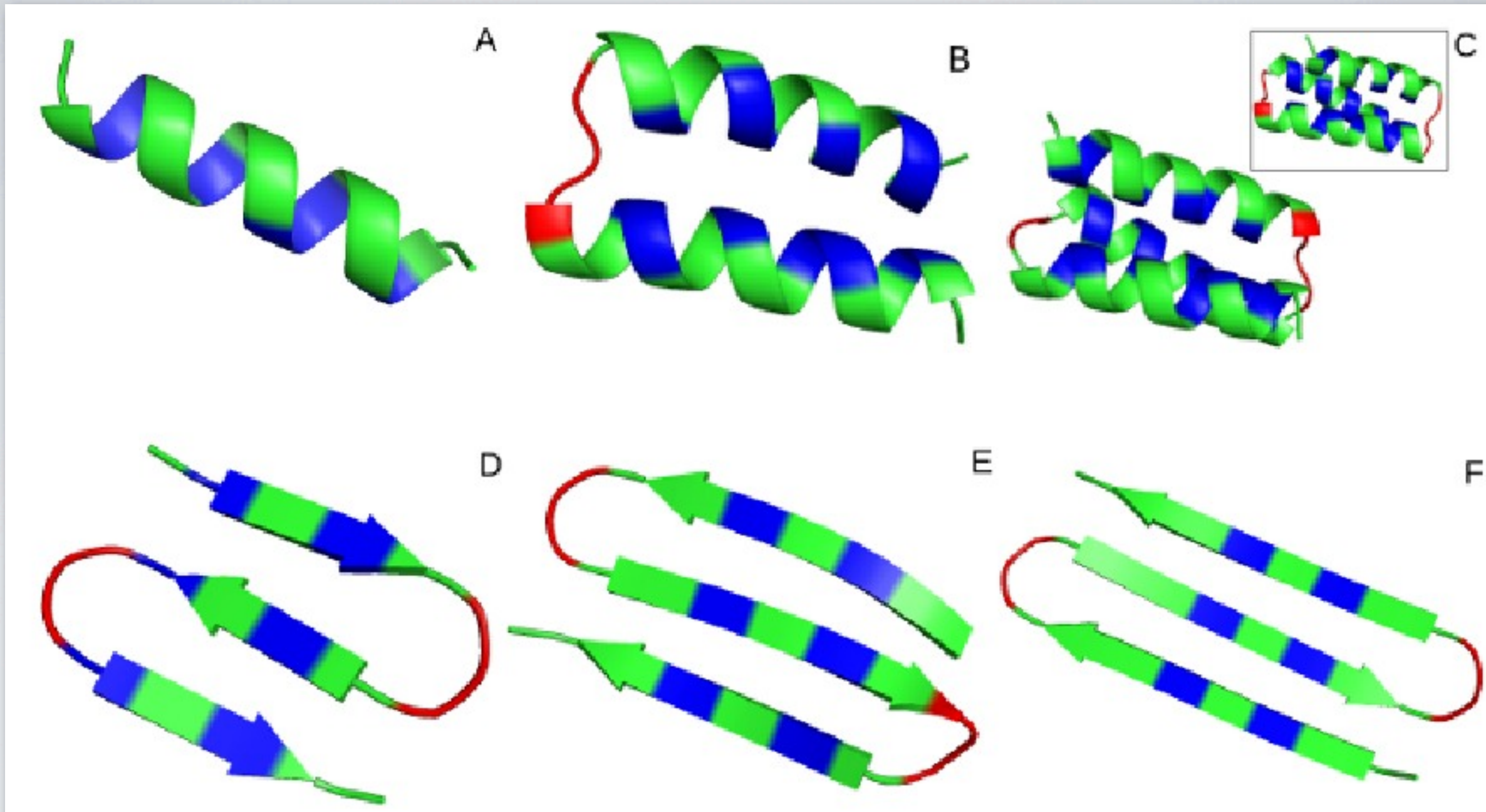
Shoemaker, Portman, & Wolynes *PNAS* **97**, 2000.

Continuous hydrophobic/ polar model

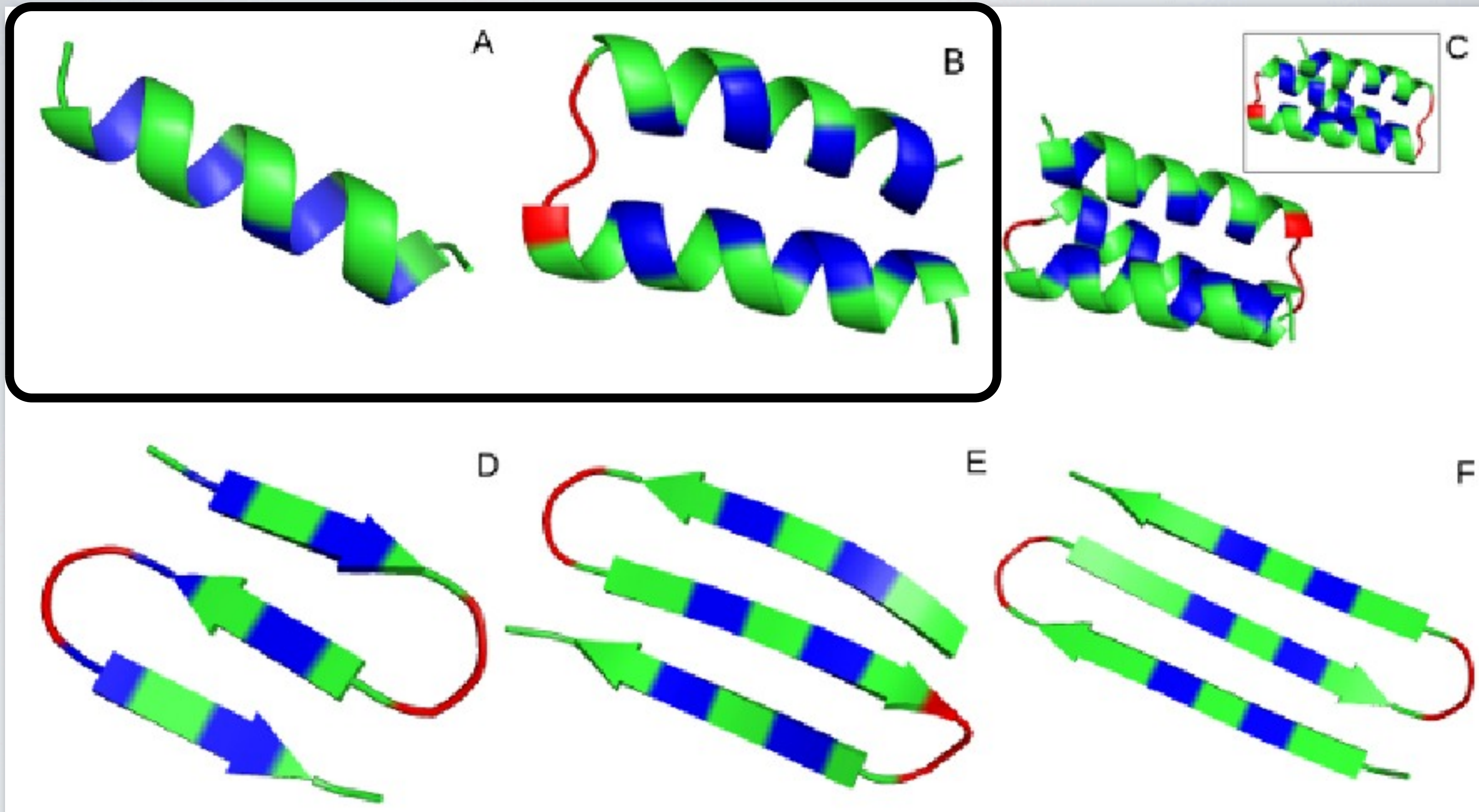


– Excluded-volume
– Hydrogen bonding (secondary structure)
– Hydrophobic attraction: **h ••• h**

[Arnab Bhattacherjee]



A
B
C
D
E
F


A
B
C
D
E
F

Thank you!