MELD: a physics based approach to predict protein shapes and interactions

Emiliano Brini
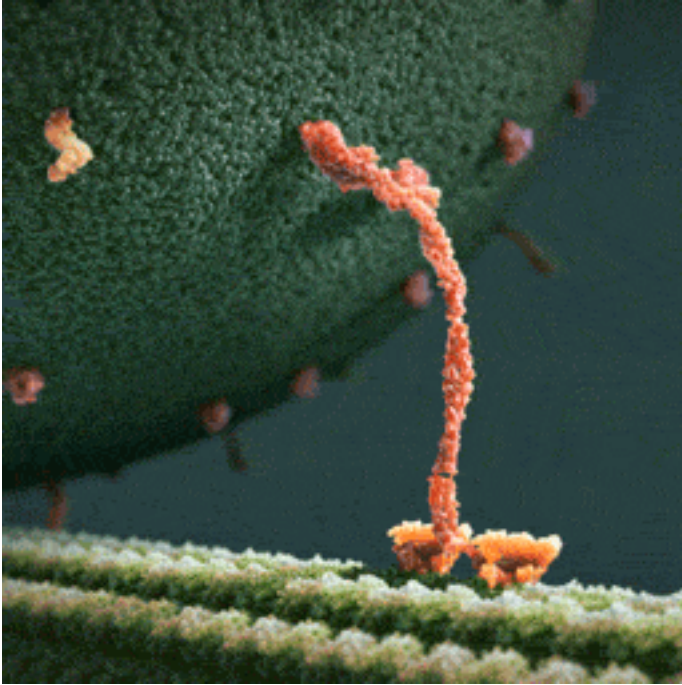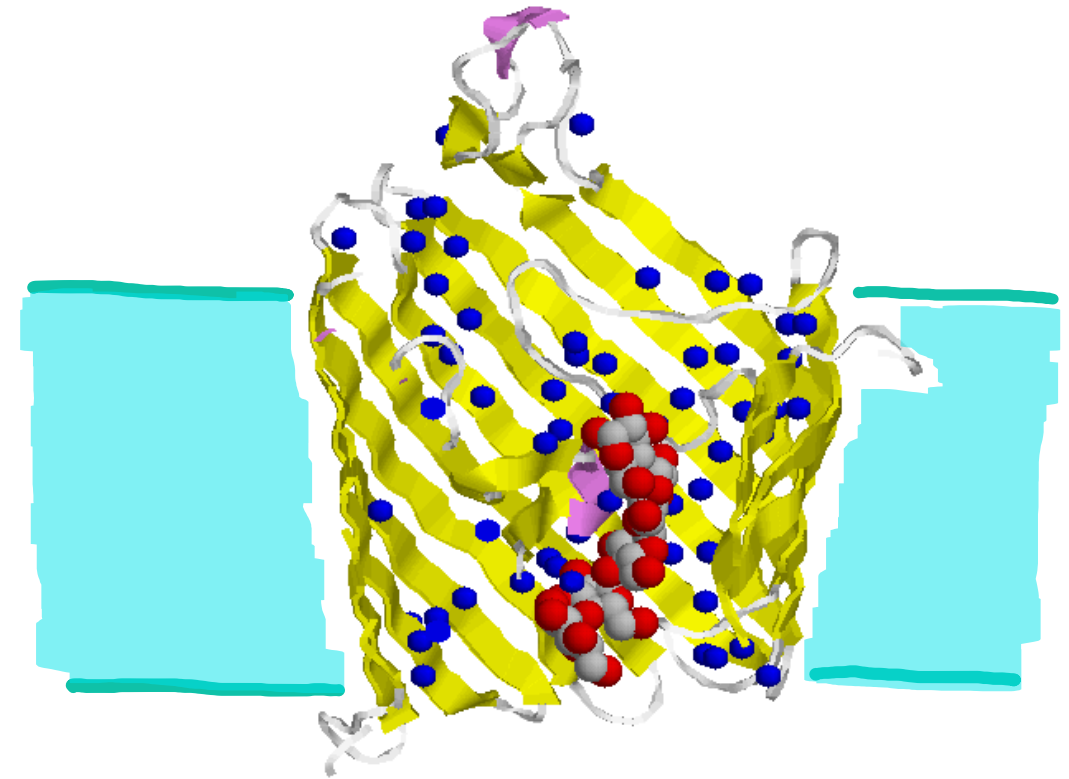
SUNY Stony Brook

emiliano.brini@stonybrook.edu

Proteins are the machines of life

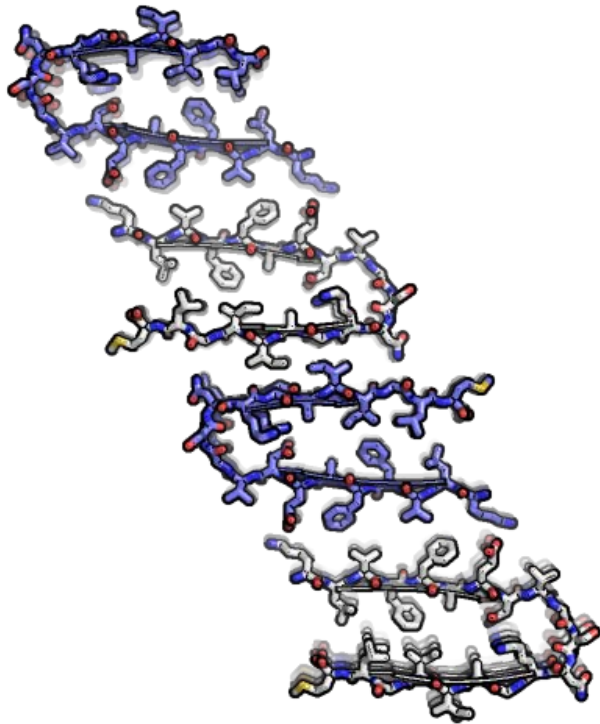Everything that happens in our body happens because a protein supervises it

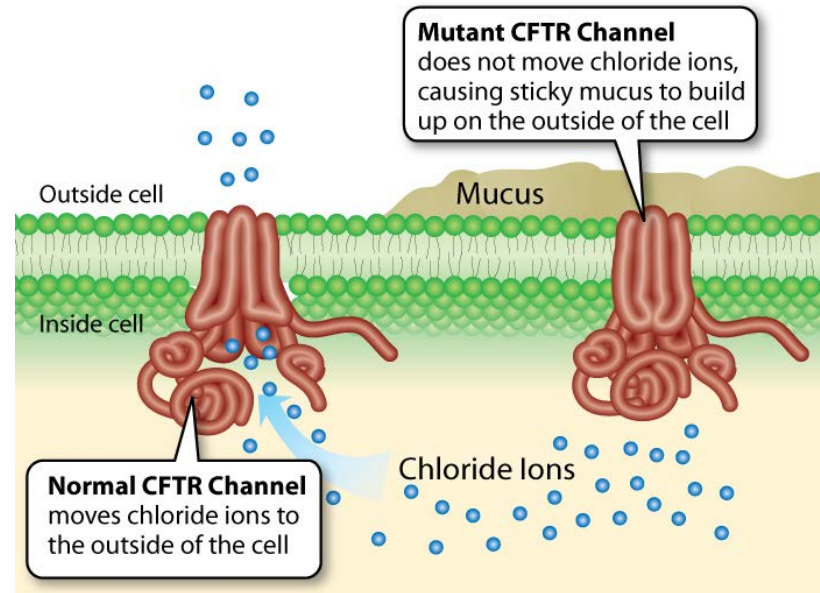Muscles contract because proteins crawl on the top of each other

Proteins allow cell to gather food from their environment

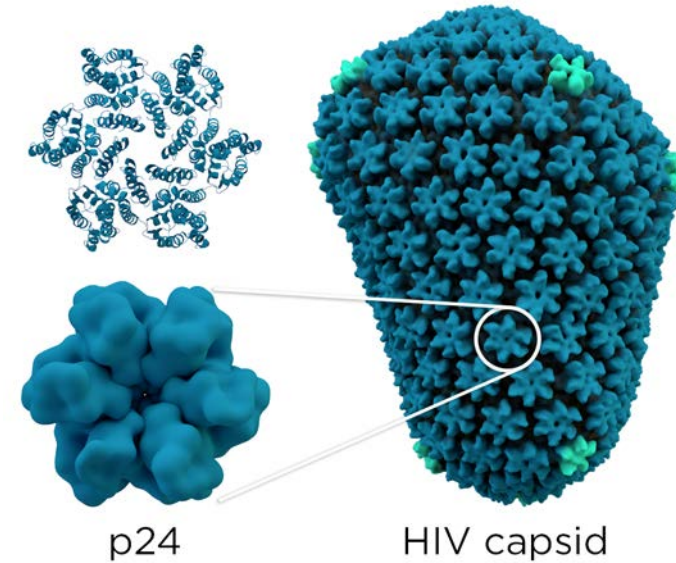# Proteins are the machines of life

Everything that happens in our body happens because a protein supervises it, even when things go wrong

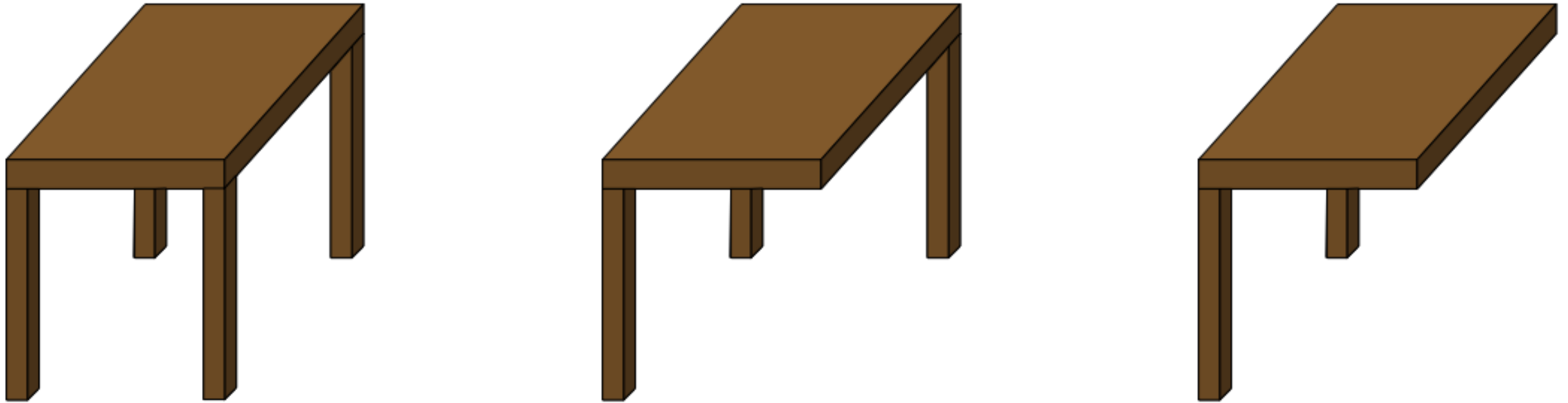Protein aggregates cause Alzheimer's disease

A single "unlucky" mutation causes cystic fibrosis

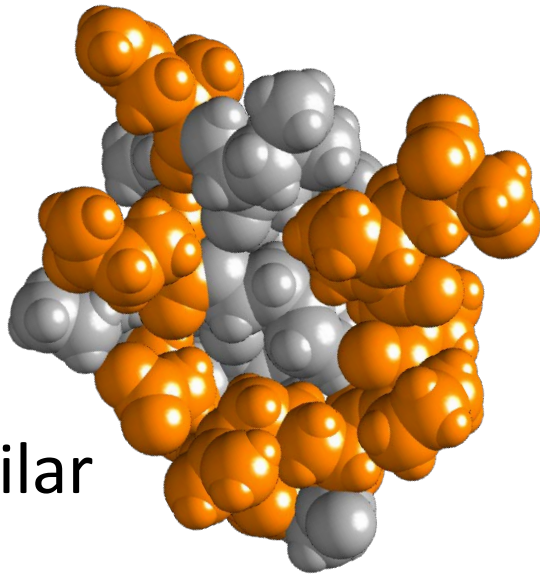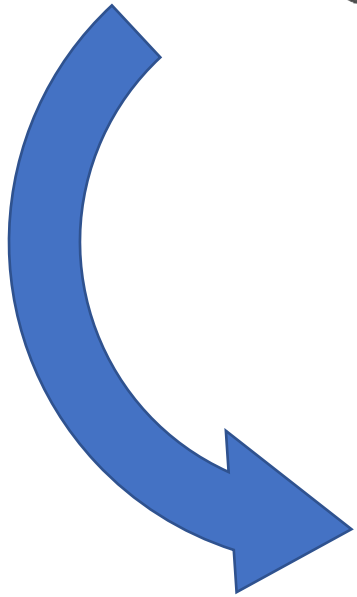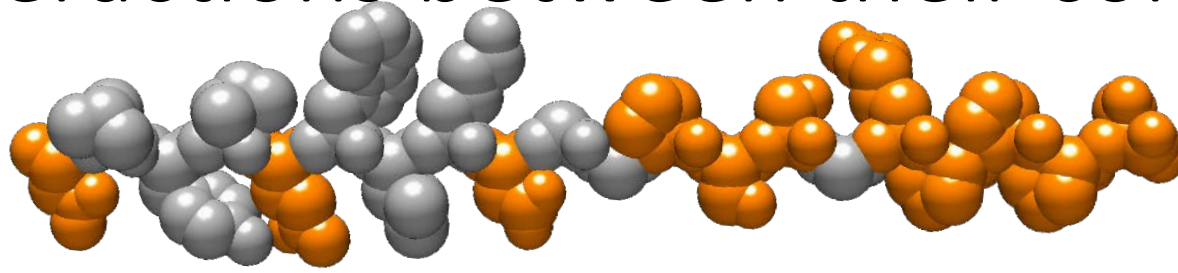Bacteria and virus need proteins to function

Proteins (like any machine) work if they have a proper shape

There is a **shape-function** relation

Sometimes mutations can make things slightly worse, sometime they are deadly

Proteins are long molecules, their shape is determined by the interactions between their components

Similar likes similar
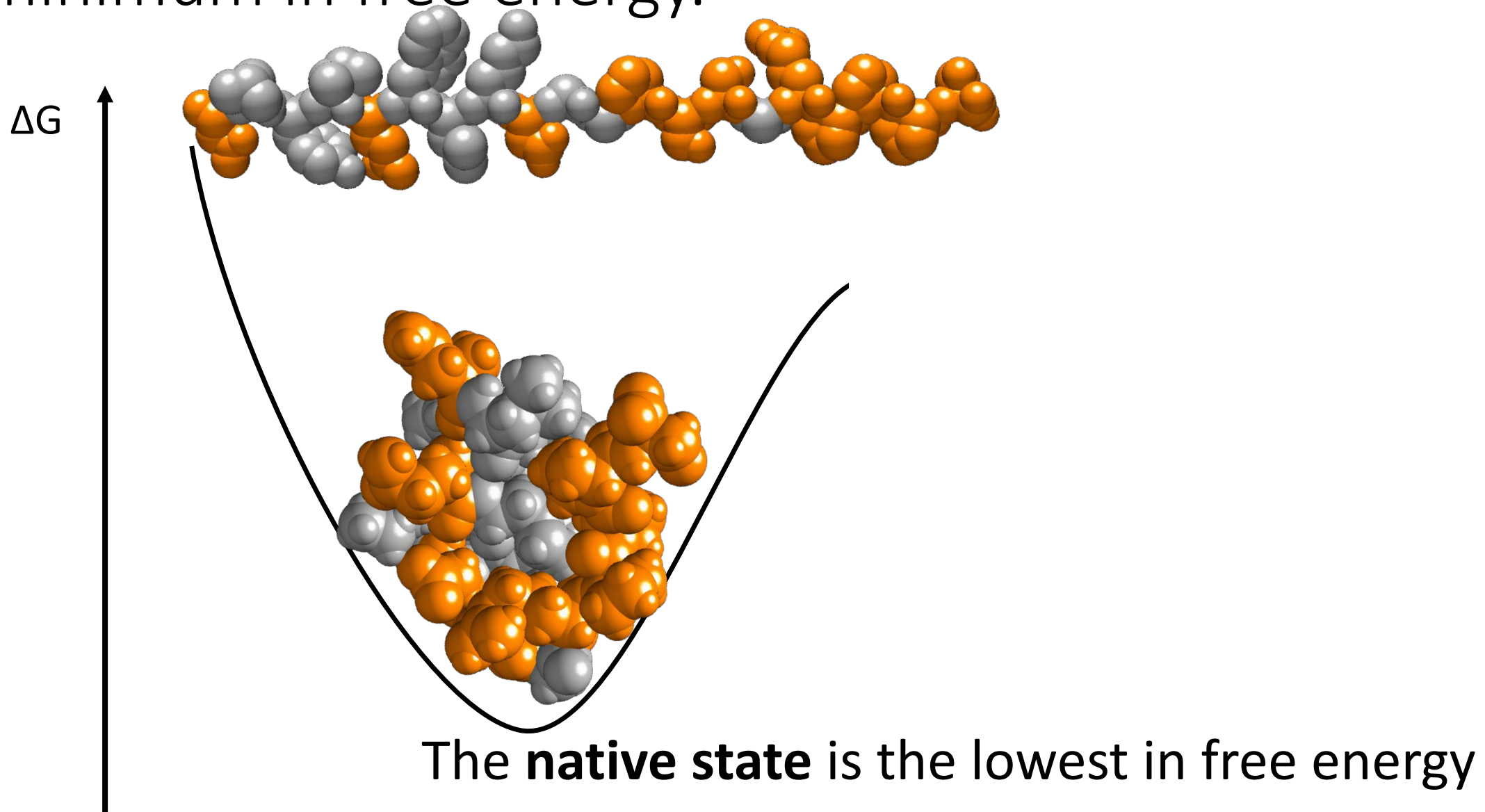
+

Being part of a chain
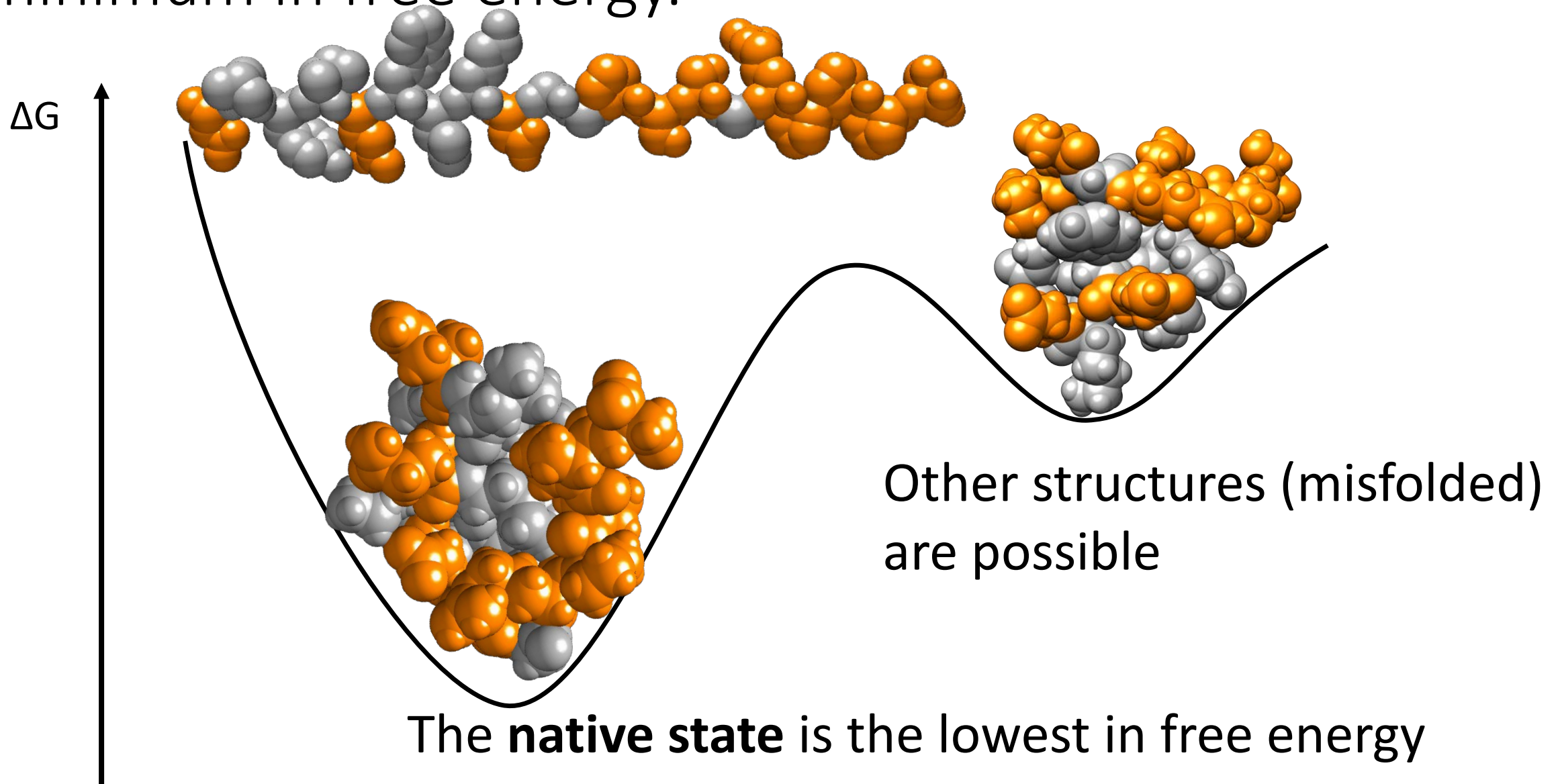
Interactions

↕

Structure

↕

Function

Physics teaches us that the most stable structure is the minimum in free energy.

ΔG

The **native state** is the lowest in free energy

Physics teaches us that the most stable structure is the minimum in free energy.

ΔG

Other structures (misfolded) are possible

The **native state** is the lowest in free energy

Physics works, but the phase space of a protein is huge!

You need to see many of these

before seeing this

We need to reduce the search space!

Knowledge helps reducing the search space

Physics is a really efficient metal detector, but it needs to be guided

Knowledge helps reducing the search space

The treasure is:
-   in a palm forest

Physics is a really efficient metal detector, but it needs to be guided

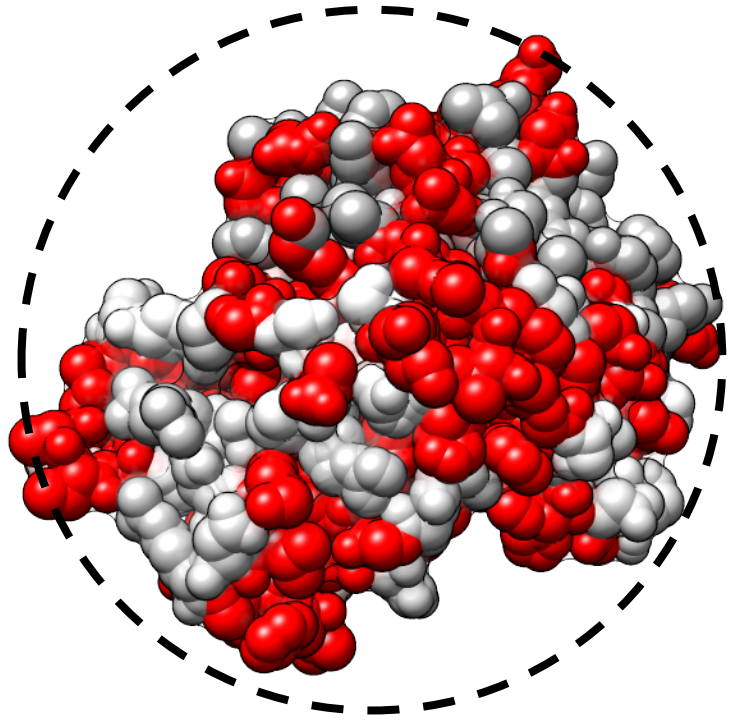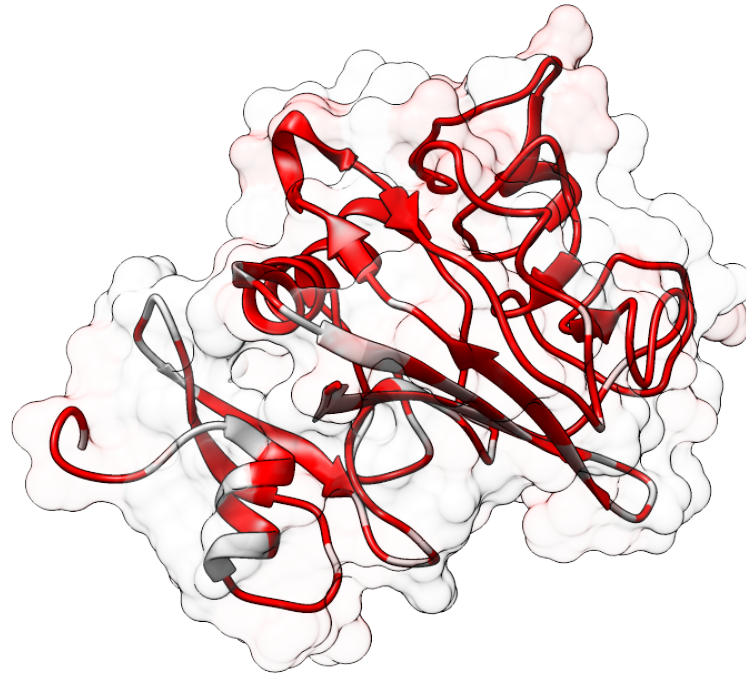# Knowledge helps reducing the search space

The treasure is:
- in a palm forest
- North of the mountains

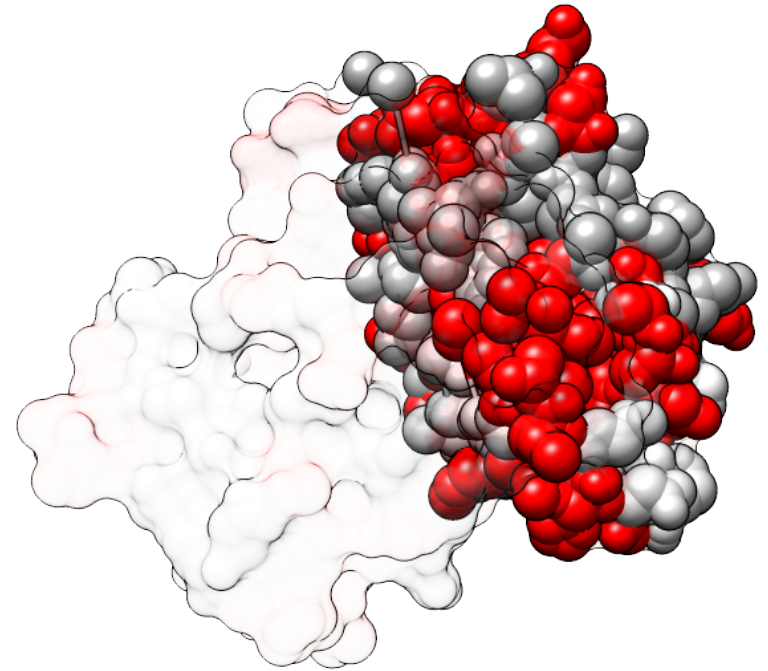Physics is a really efficient metal detector, but it needs to be guided

Our "lore" is that a protein should look like a protein

It has a given size
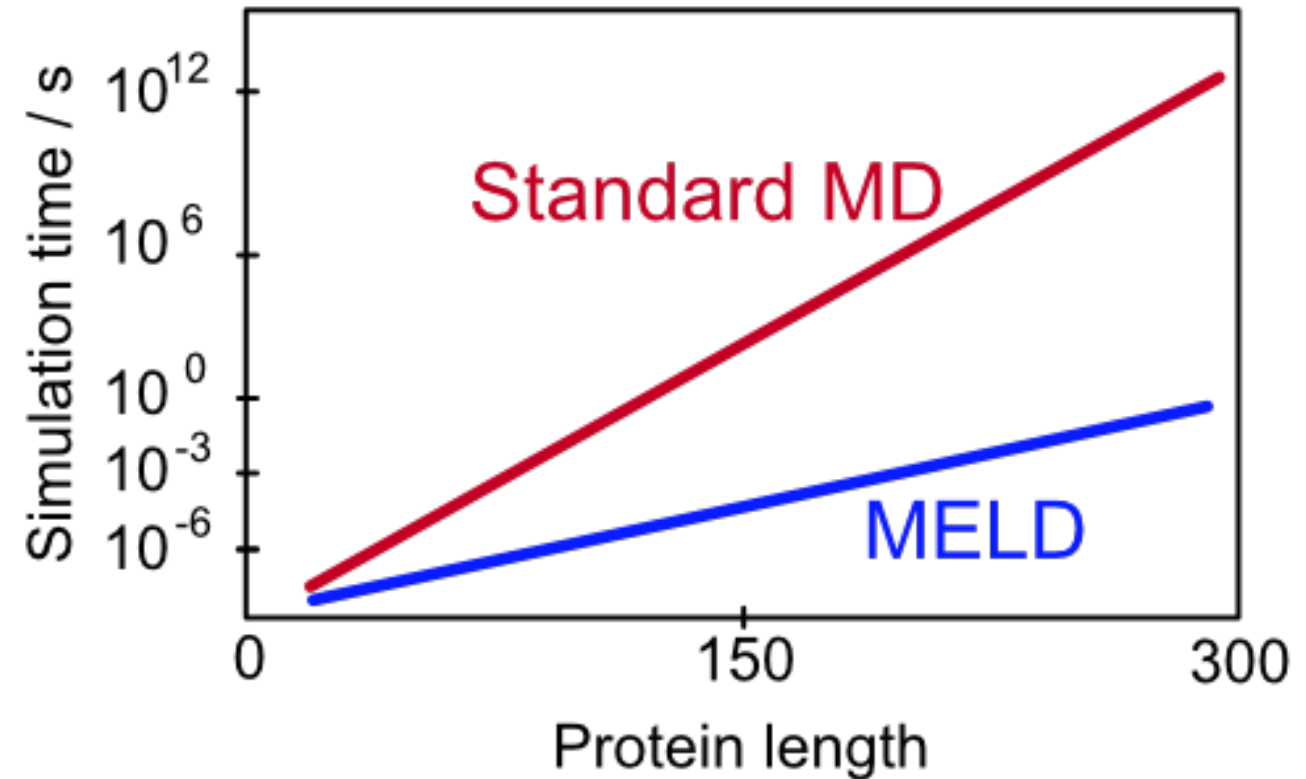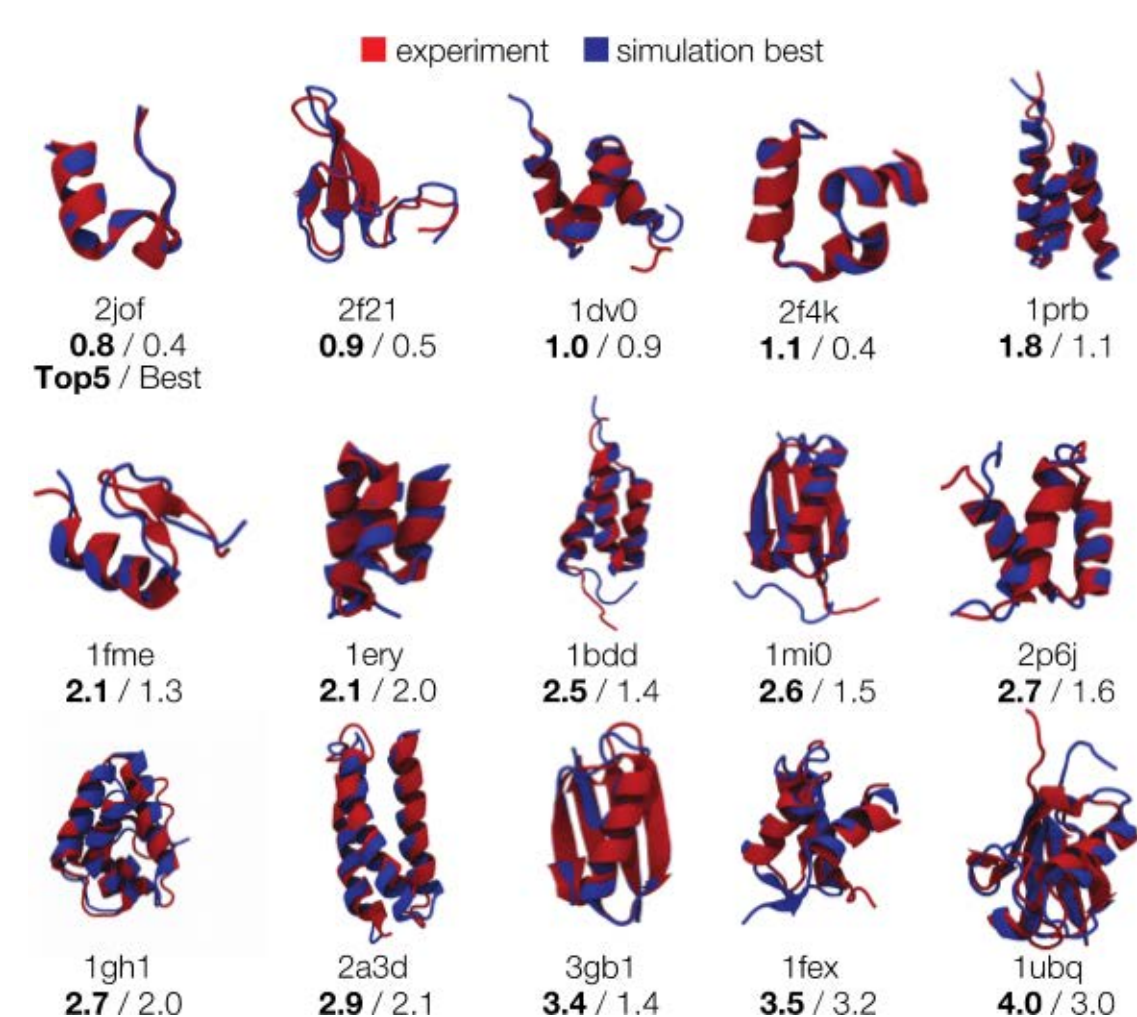
It has a secondary structure

It has a hydrophobic core

We penalize the energy of any structures that don't fit our lore

With MELD we folded small proteins very fast

D. E. Shaw (Anton) and Simmerling (GPU) showed that with physics is possible to fold a set of small proteins, we folded them faster.


experiment ■ simulation best

2jof **0.8** / 0.4 **Top5** / Best	2f21 **0.9** / 0.5	1dv0 **1.0** / 0.9	2f4k **1.1** / 0.4	1prb **1.8** / 1.1
1fme **2.1** / 1.3	1ery **2.1** / 2.0	1bdd **2.5** / 1.4	1mi0 **2.6** / 1.5	2p6j **2.7** / 1.6
1gh1 **2.7** / 2.0	2a3d **2.9** / 2.1	3gb1 **3.4** / 1.4	1fex **3.5** / 3.2	1ubq **4.0** / 3.0

Nguyen, H et al ; JACS 136.40 (2014): 13959-13962.
Lindorff-Larsen, K et al ; *Science* **2011**, 334, 517
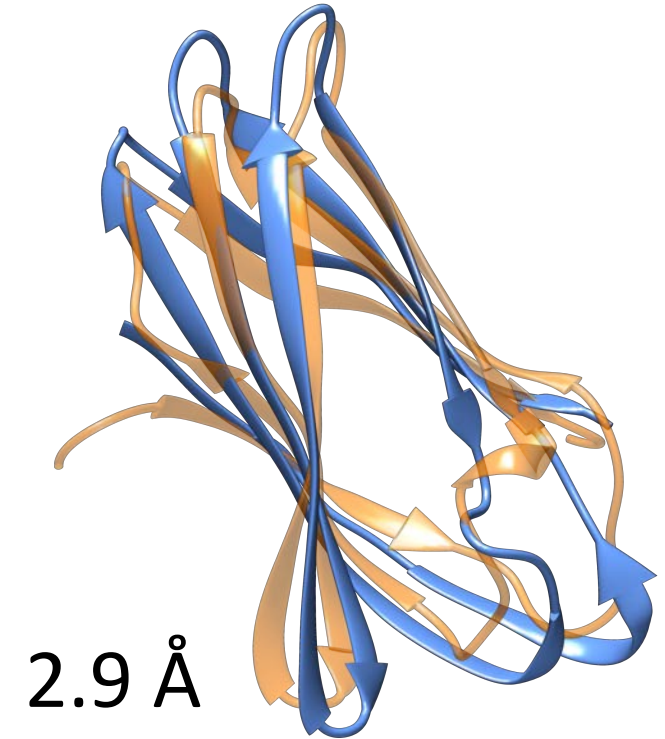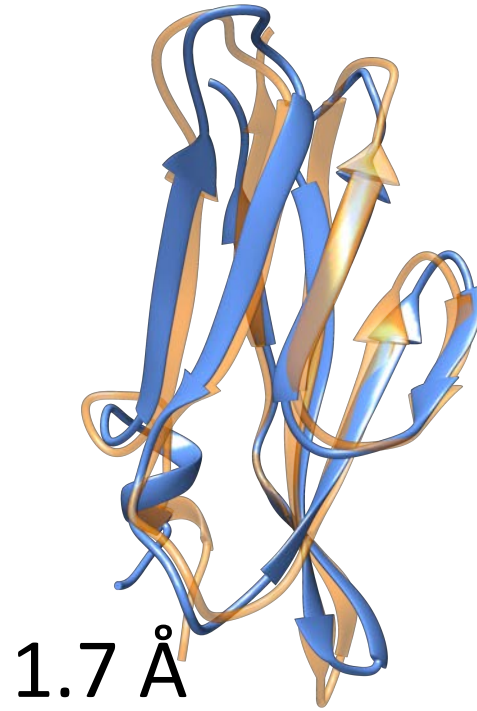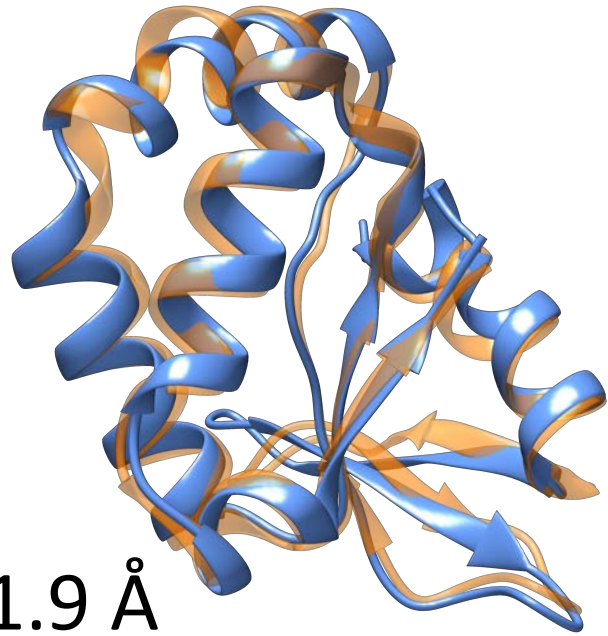Perez, A et al ; PNAS 112.38 (2015): 11846-11851.

We got some good results in CASP11 (2013)

1.4 Å

1.0 Å

0.7 Å

5.8 Å

Experimental data →

1.2 Å

Perez, A et al; *Science adv.* 2.11 (2016): e1601274.
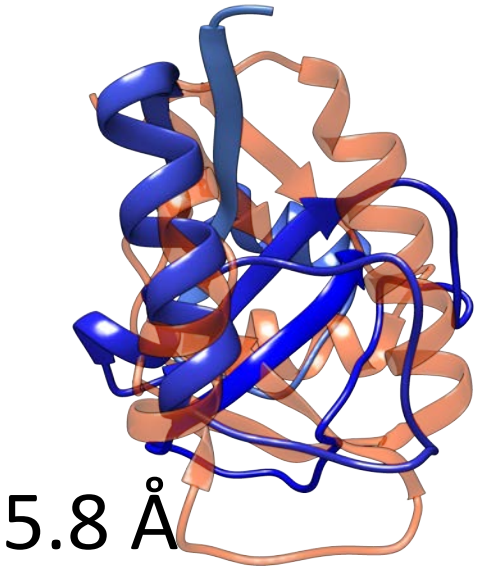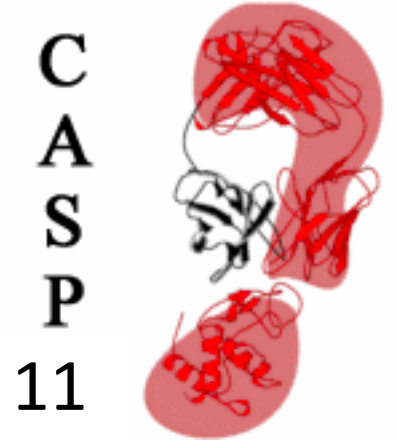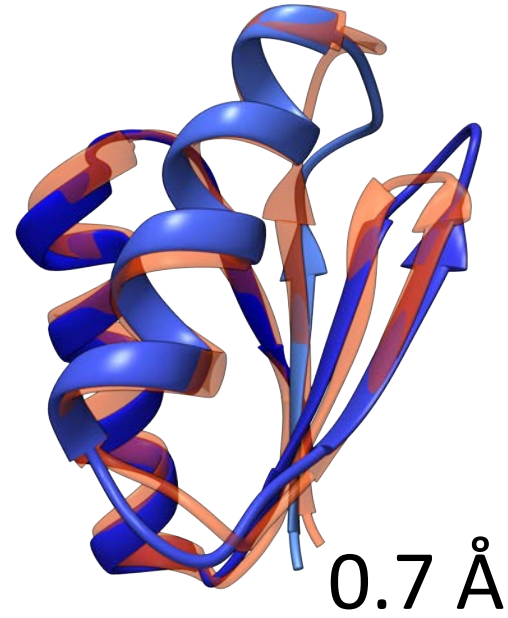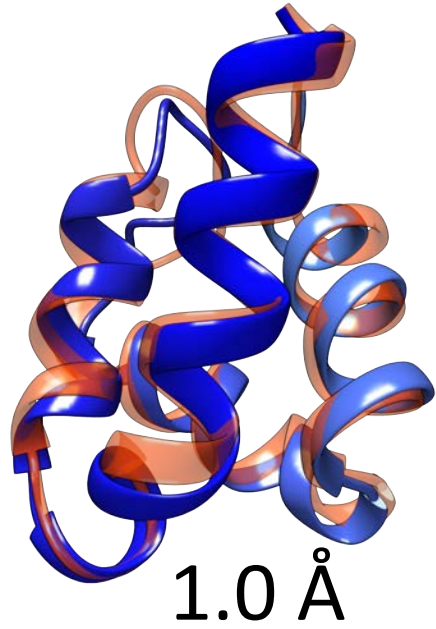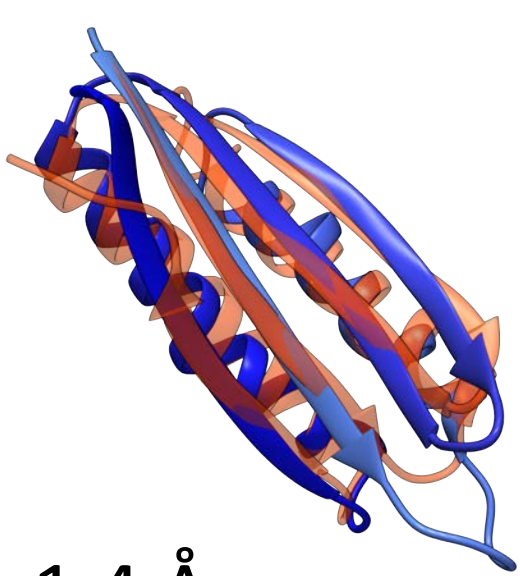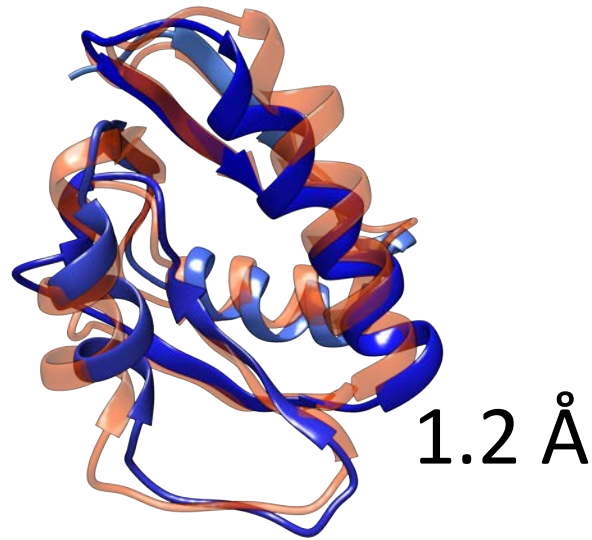
Thanks to BW we hit some great success in CASP 12 (2015)

1.9 Å   1.7 Å   2.9 Å
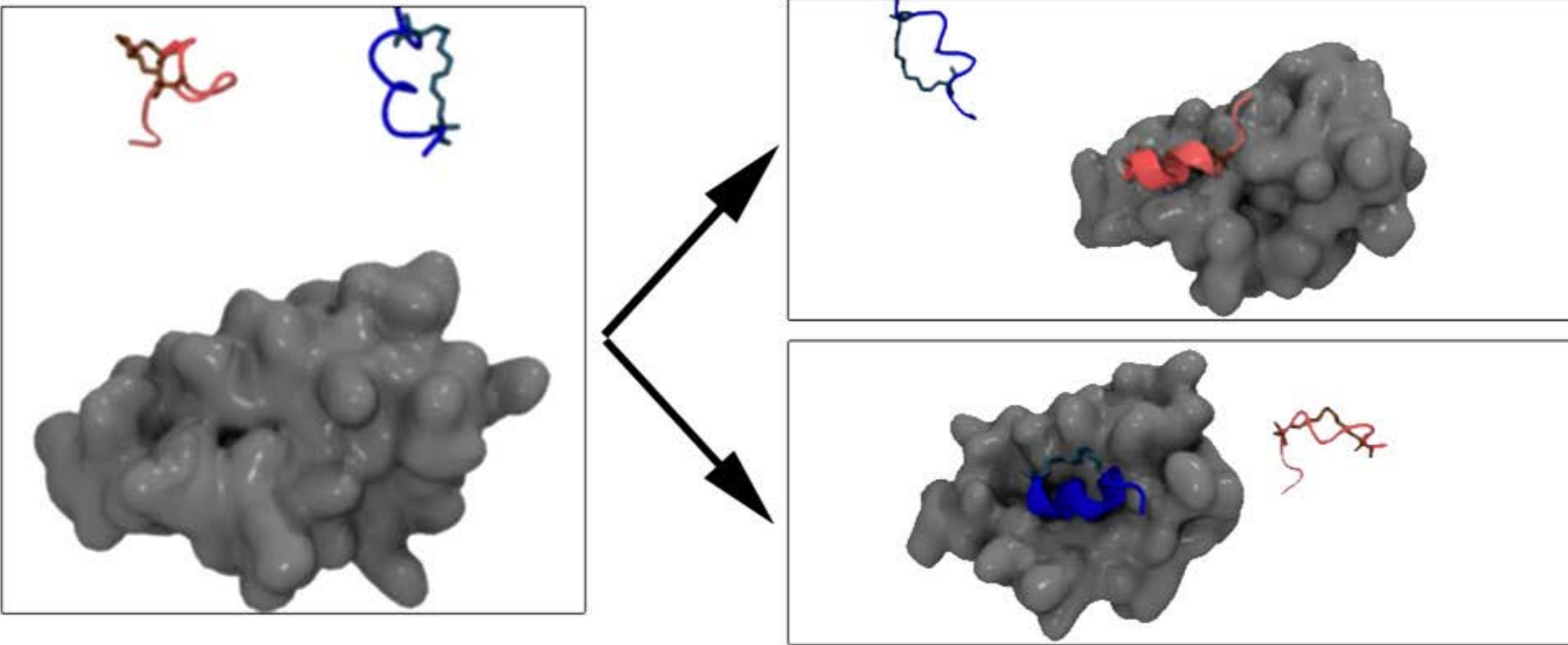
Bigger proteins

Many more proteins than before (120 vs 20)

Many different approaches

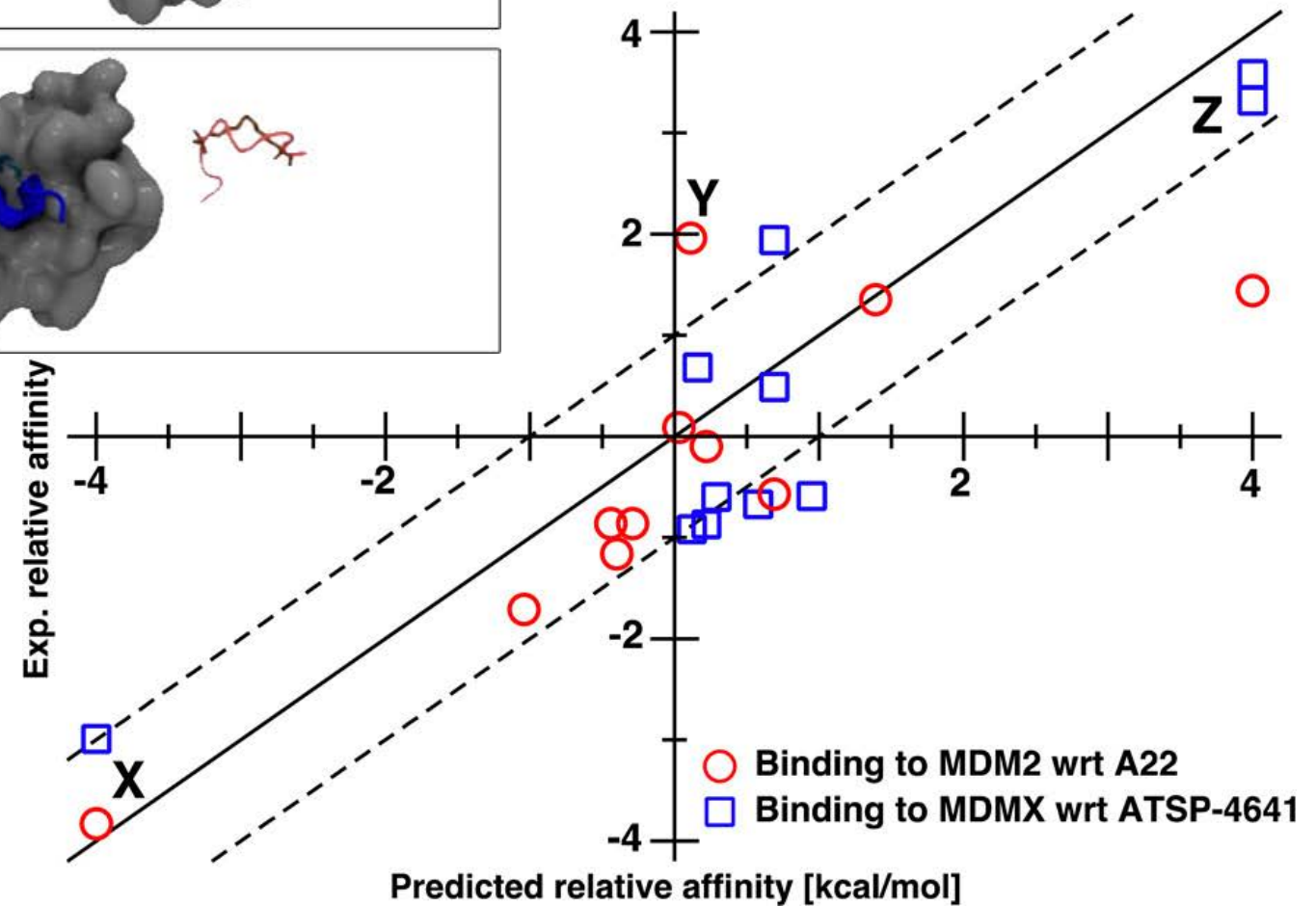**Best structure in all CASP (lowest RMSD)**

Most improved refinement in all CASP

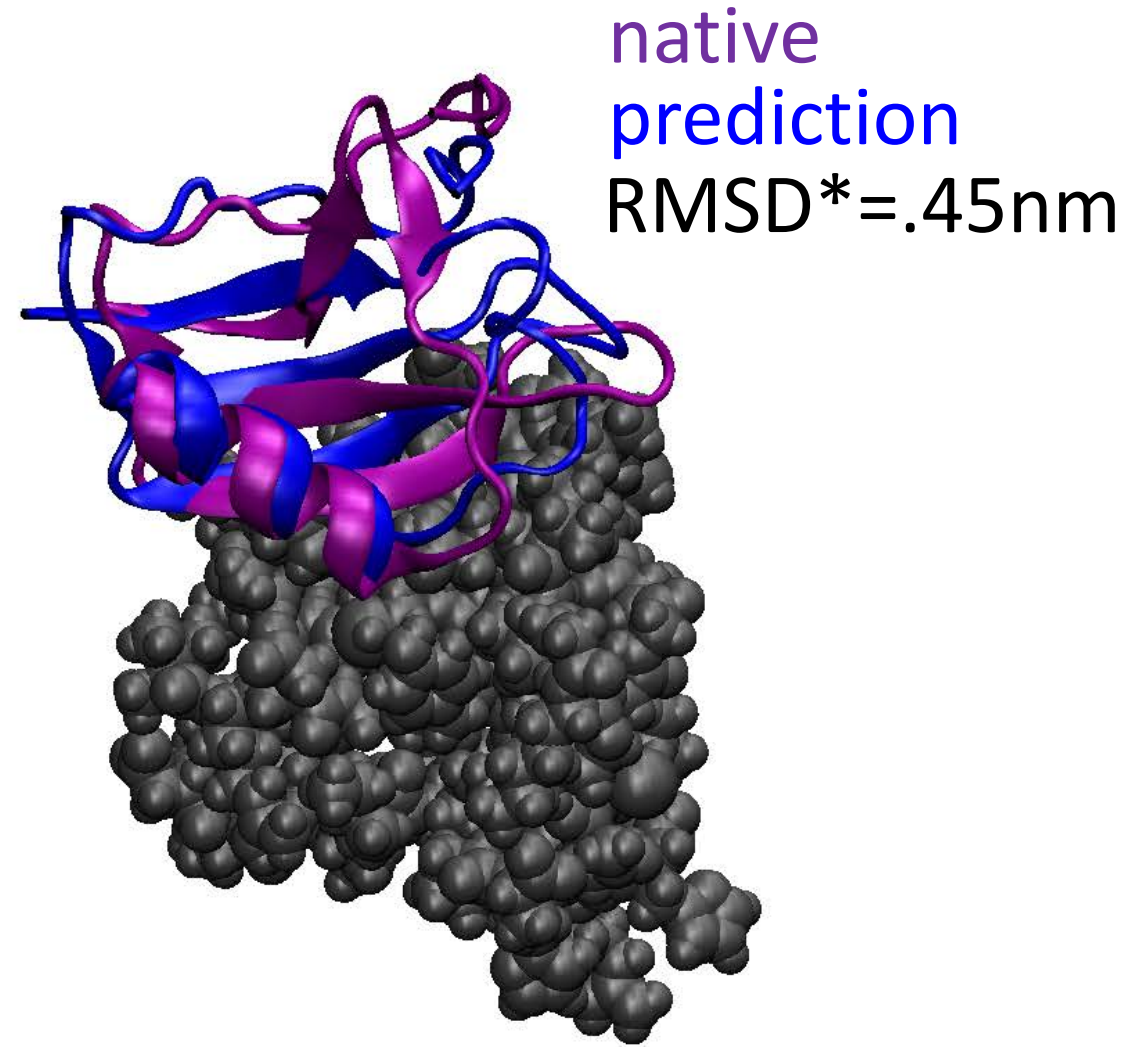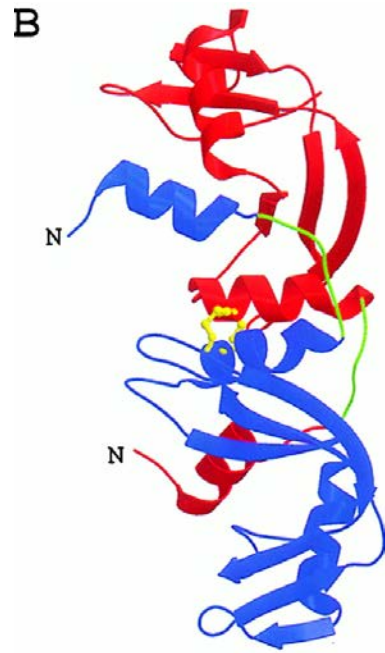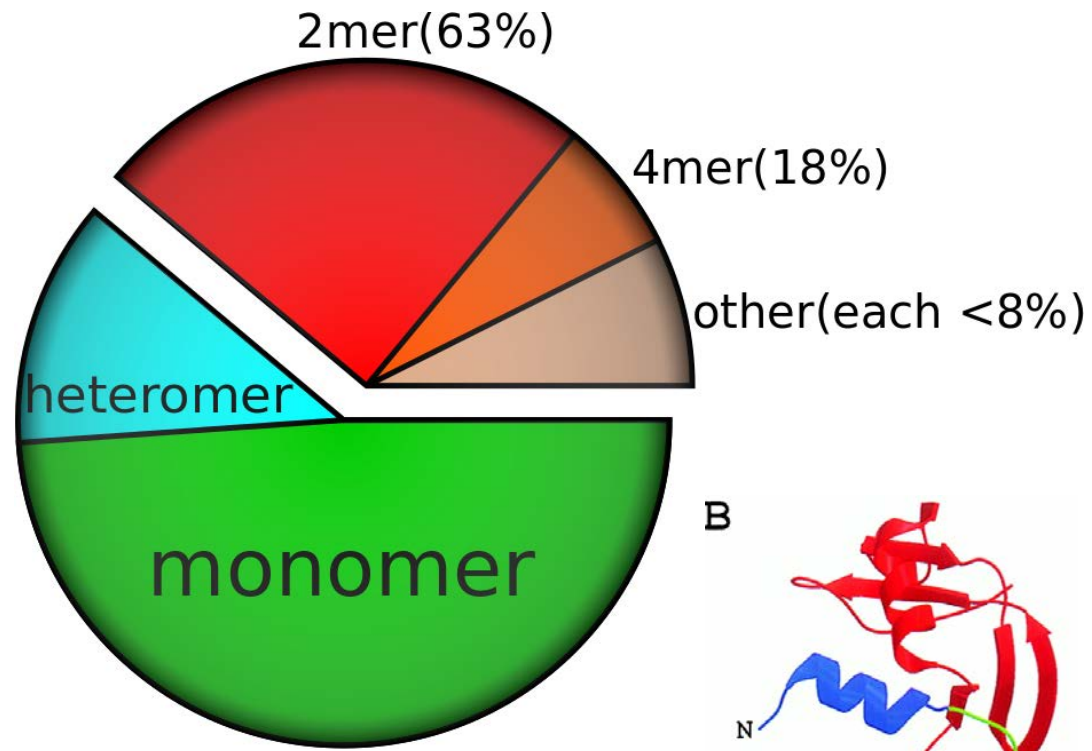# Physics and BW allow us to study drug binding events

Morrone, J et al. *JCTC* (2017).
Morrone, J et al. *JCTC* (2017).

We can look at the preferential binding of two peptides and build a relative binding affinity ranking for a series of compounds

Physics and BW allow us to study protein dimers (CAPRI)

2mer(63%)

4mer(18%)

other(each <8%)

heteromer

monomer

B

N

N

native
prediction
RMSD*=.45nm

Conclusion:

- Knowing the shape of the protein allows to understand it
- Physics is a valuable tool to fold proteins, but needs to be helped. It allows to go beyond folding.
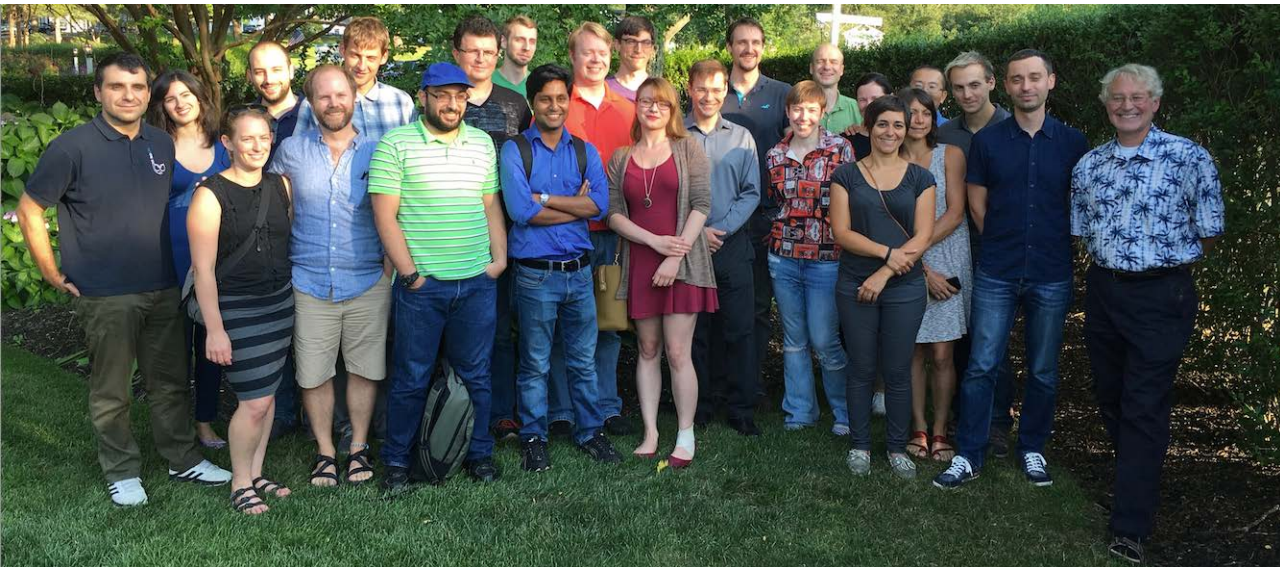- We need BW to be effective

Conclusion:

- Knowing the shape of the protein allows to understand it
- Physics is a valuable tool to fold proteins, but needs to be helped. It allows to go beyond folding.
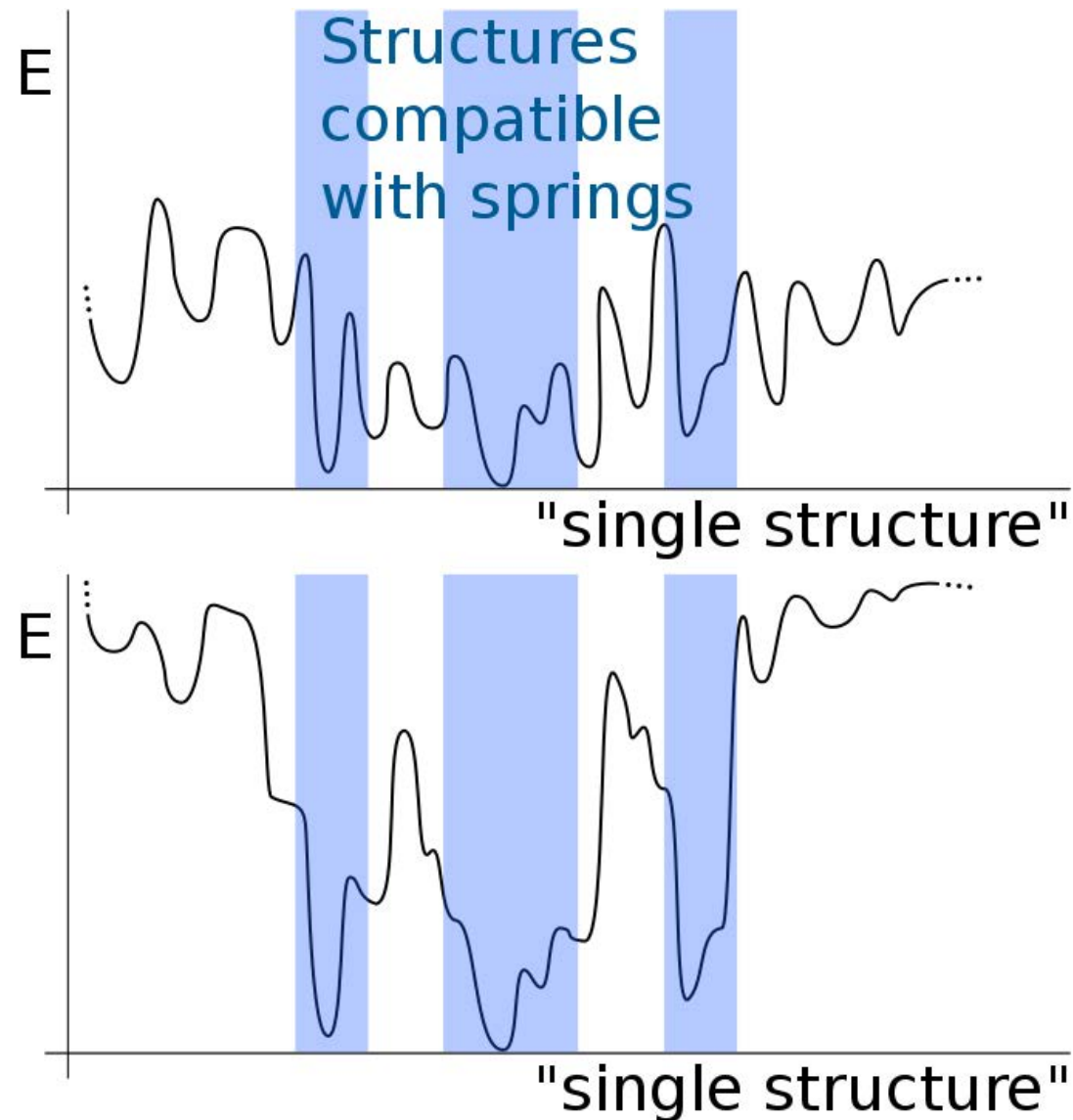- We need BW to be effective

Acknowledgments

Ken Dill
Alberto Perez, Joseph Morrone,
Justin McCallum, Lane Votapka,
James Robertson, Cong Liu
NIH, NSF, Blue Waters

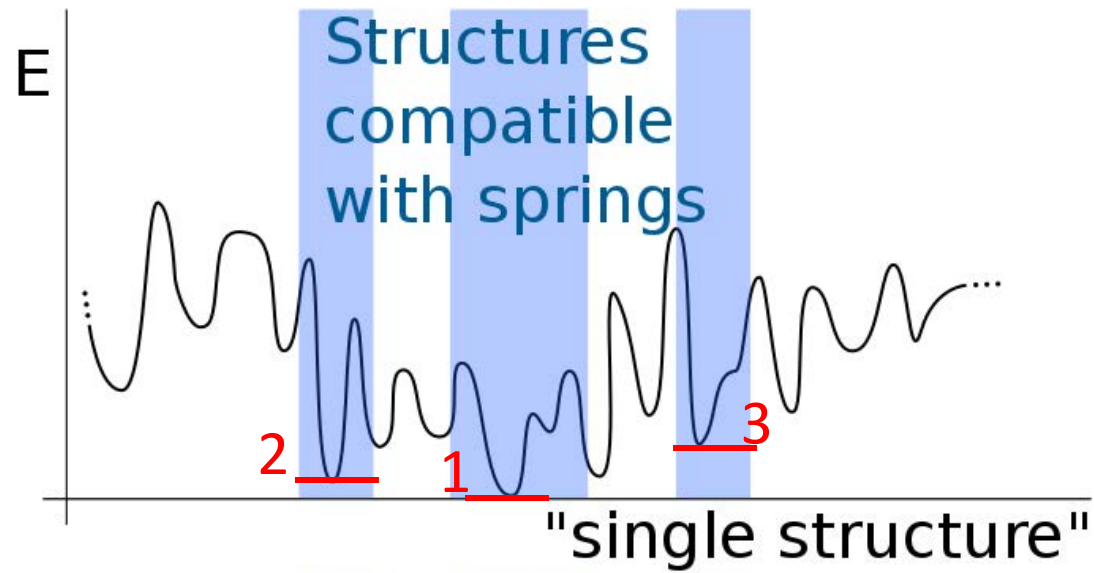# In MELD we use "smart" springs to drive the system without biasing the minima

E

Structures compatible with springs

"single structure"

E

"single structure"

We use flat bottom potential springs to avoid biasing the energy minimum
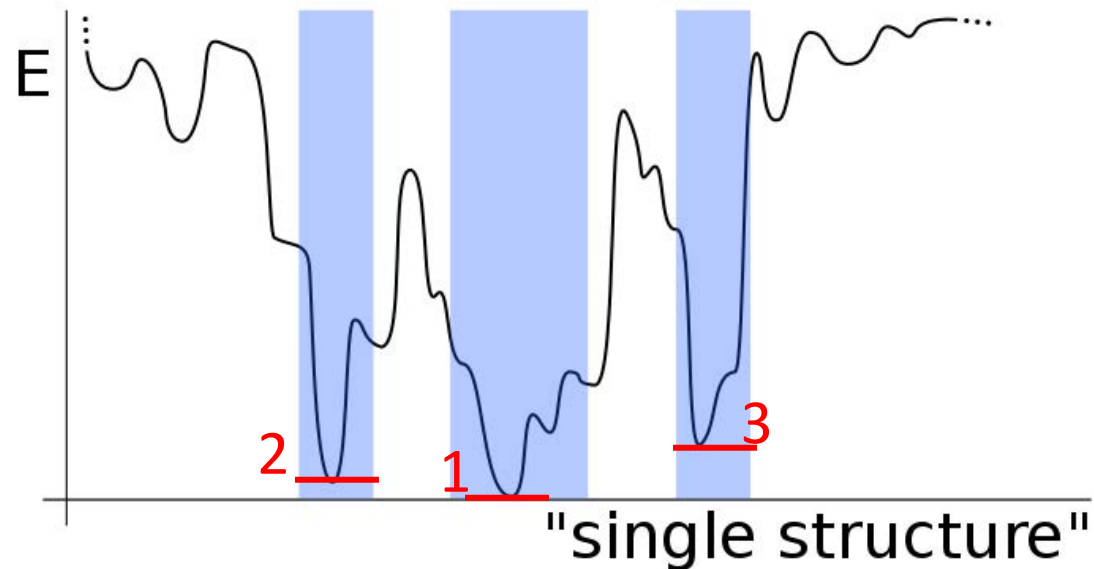
U(r)

r
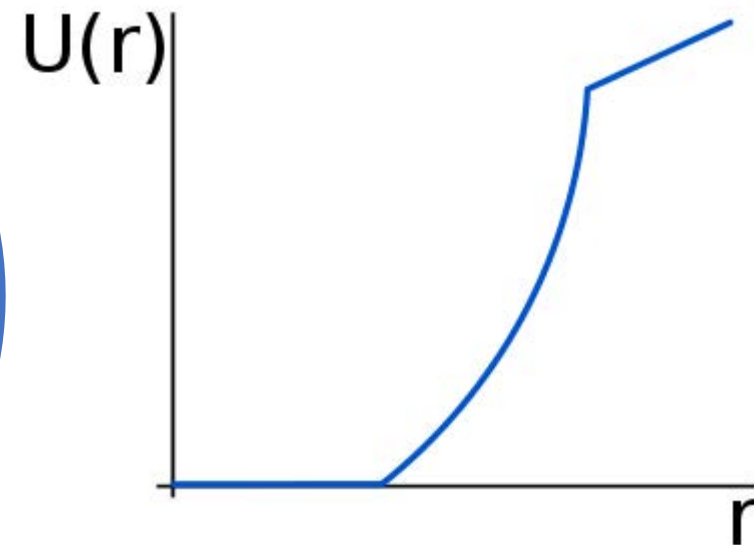
We use subset of springs

$$p(x|R) = p(R|x) * p(x)$$

In MELD we use "smart" springs to drive the system without biasing the minima

$$p_1 = \frac{e^{-E_1/kT}}{\sum_i e^{-E_i/kT}} \qquad \frac{p_1}{p_2} = \frac{e^{-E_1/kT}}{e^{-E_2/kT}}$$

$$p_1^* = \frac{e^{-E_1/kT}}{\sum_i e^{-E_i^*/kT}} \qquad \frac{p_1^*}{p_2^*} = \frac{e^{-E_1/kT}}{e^{-E_2/kT}}$$

We might need to converge only the trend