Molecular Modeling -- 2020
Lecture 22 -- Protein folding and design and protein-protein docking

Tuesday, April 14

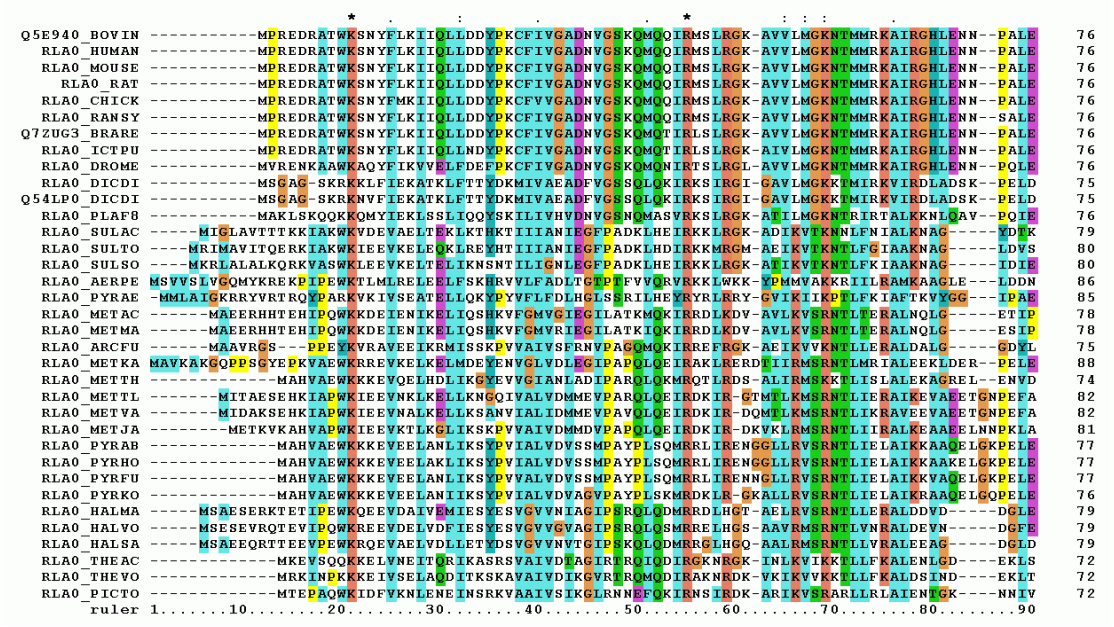
First order of business: term project

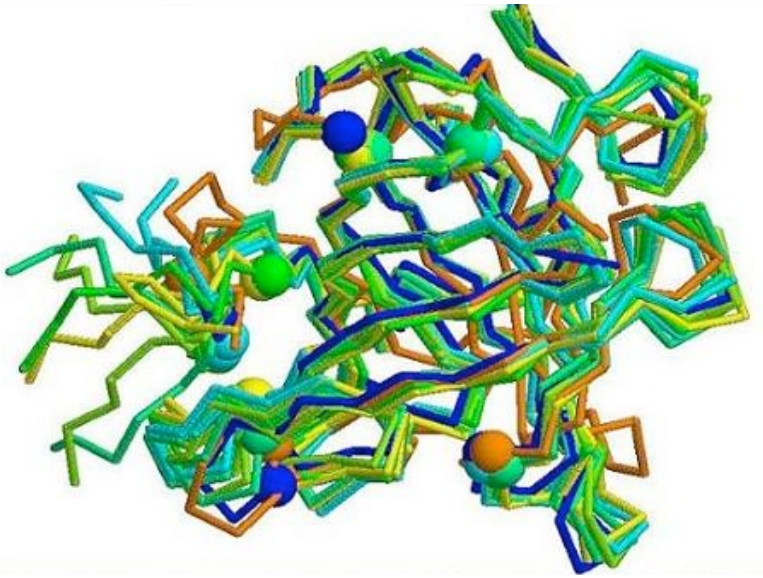
Why design proteins?

- Proteins can do anything!
- Protein-based drugs lack many of the side-effect issues of small molecule drugs.
- Proteins are immunogenic.
- Proteins are environmentally friendly.

Protein folding/ protein design

folding



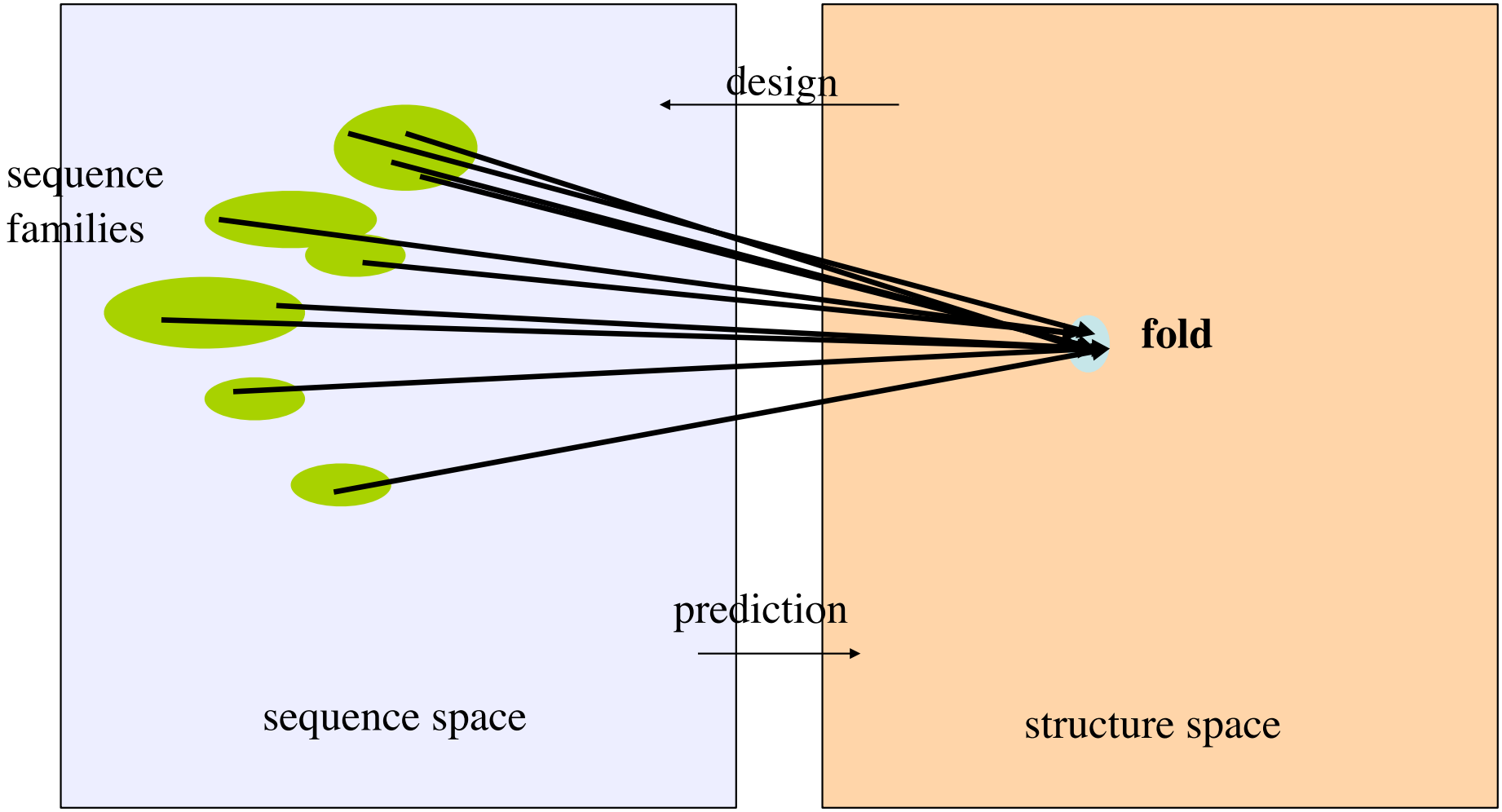


design

sequence

structure

Sequence space maps to structure space like this



.. many-to-one.

Thus design is "easier" than prediction.

Short history of protein design

Site-directed mutagenesis, engineering (J. Wells, 1980's-90's)

Coiled coils, helix bundles (DeGrado, 1980's-90's)

Binary patterning (Hecht, 1990's)

Extreme protein stabilization (Mayo, 1990's)

Binding pocket design (Hellinga, 2000)

New fold design (Kuhlman & Baker, 2002-4)

Protein-protein interface design (Gray & Baker, 2004)

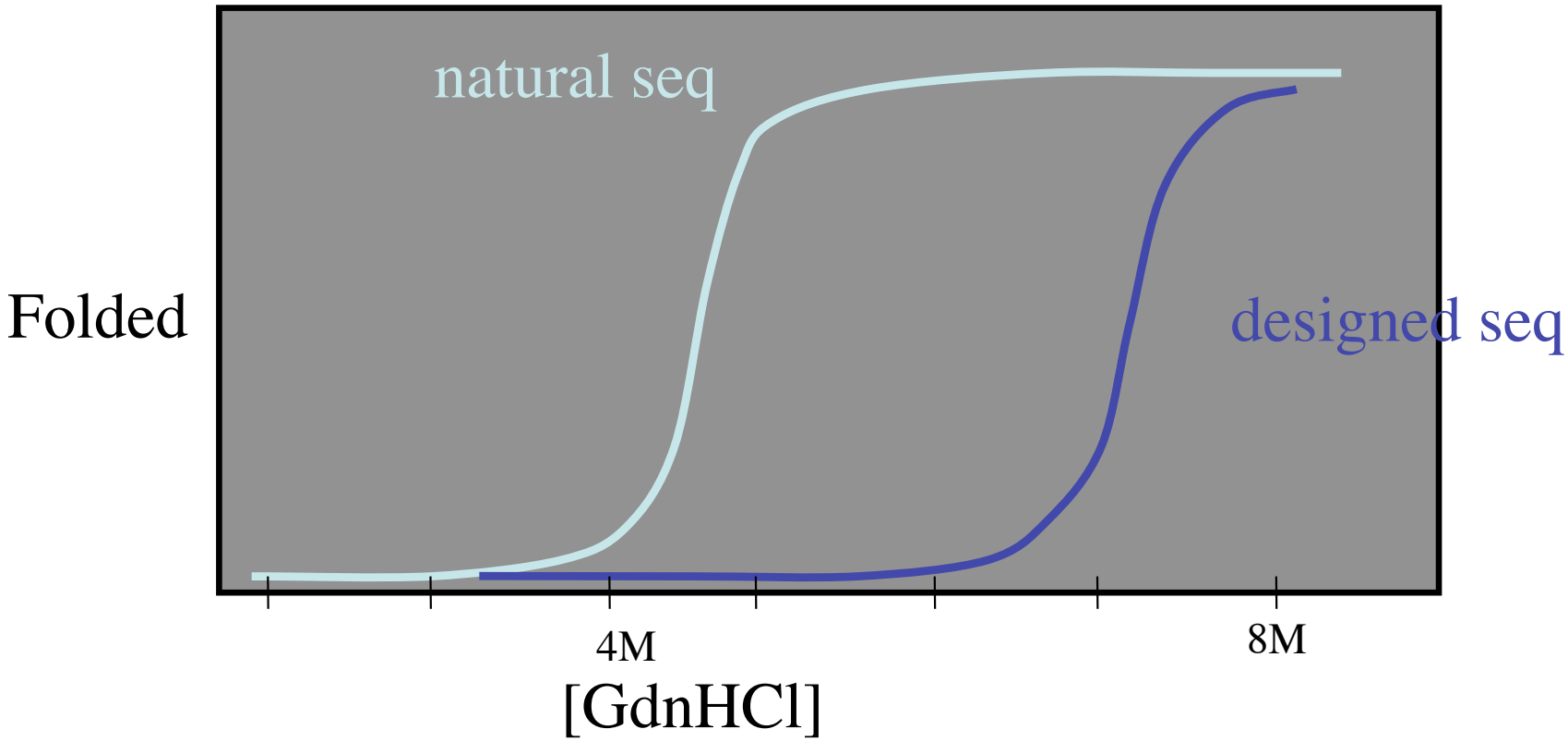
Open source protein design algorithm EGAD (Pokala, 2005)

Enzyme design (Baker, 2008)

Flexible backbone protein design (Kortemme, 2009)

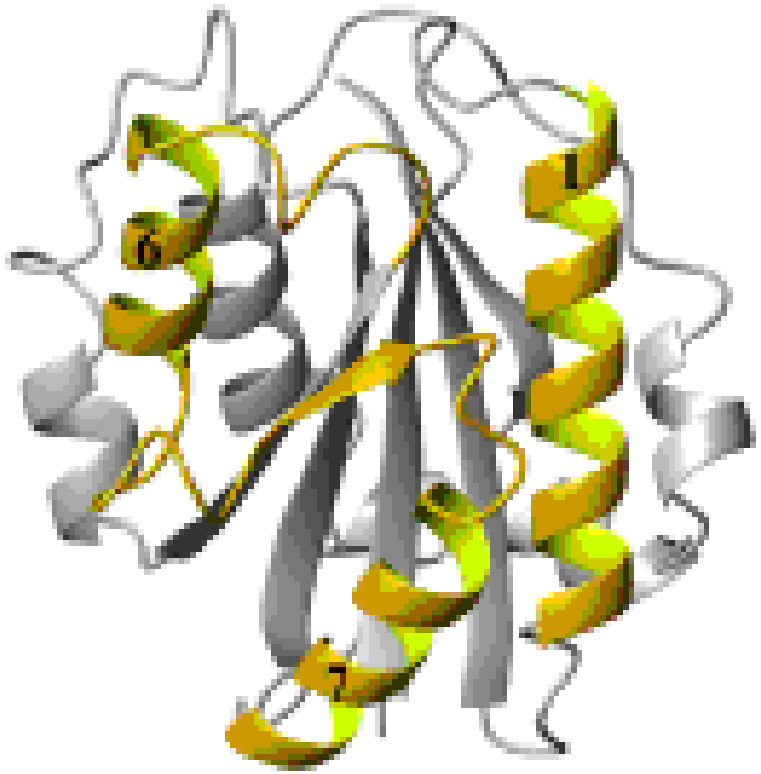
Continuous rotamer design (Donald, 2012)

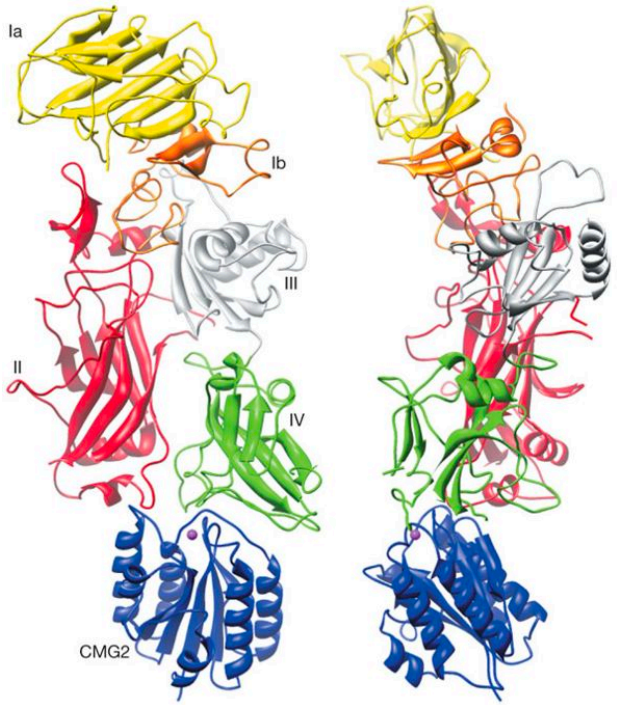
Proteins can be made super-stable



Malakauskas SM and Mayo SL (1998) "Design, Structure, and Stability of a Hyperthermophilic Protein Variant." *Nature Struct. Biol.*, **5**, p.470.

Distinct conformational states can be stabilized.





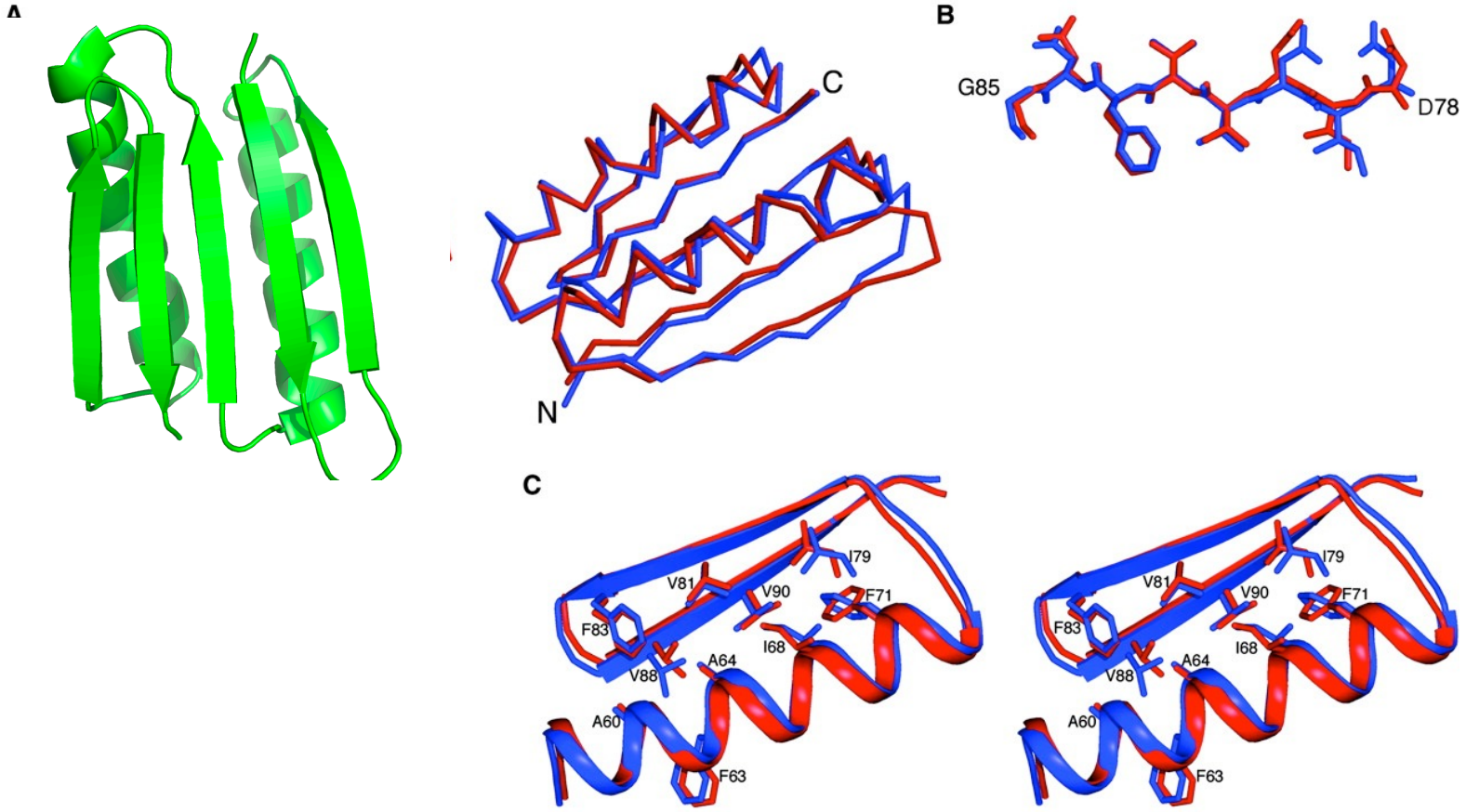
αMβ2 integrin I domain in 2 conformations

2 crystal structures are known. They differ in the highlighted region.

Shimaoka *et al* designed sequences for each form, **open** and **closed**. The two designs were shown to have different physiological properties.

Shimaoka, M., Shifman, J. M., Takagi, J., Mayo, S. L., Springer, T. A. (2000) "Computational design of an integrin I domain stabilized in the high affinity conformation." *Nature Struc. Biol.* **7**(8), 674-678.

New folds have been designed

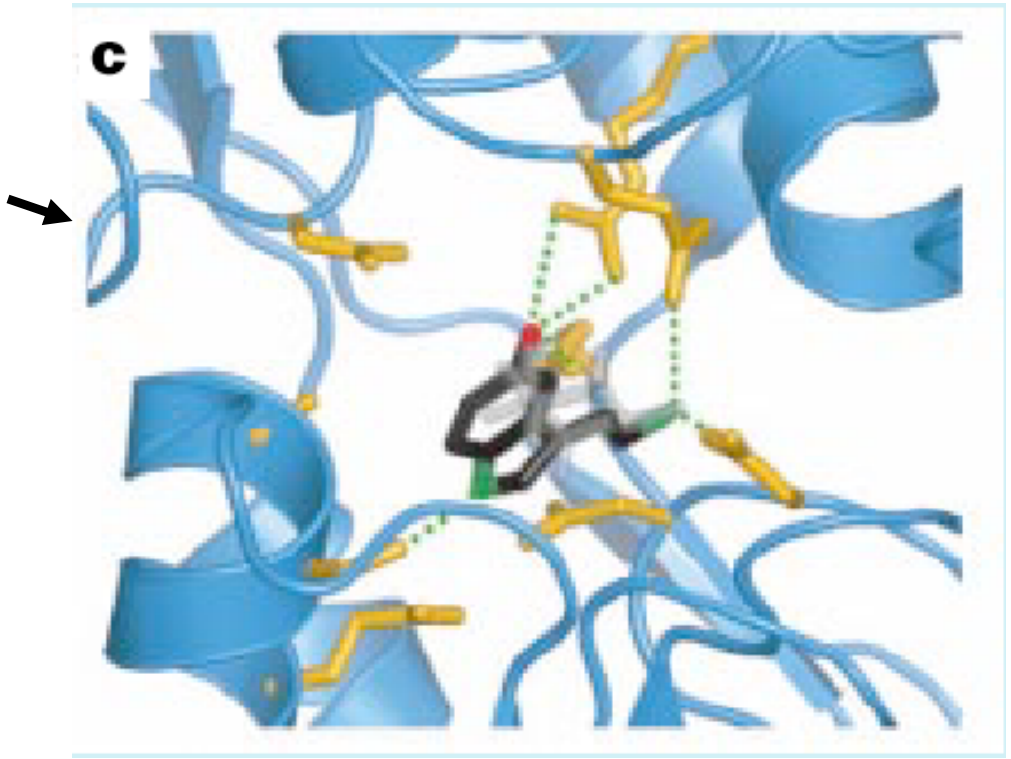


New proteins can be designed that have never been seen before. The designs are accurate (compare red and blue above) and they are highly stable.

Kuhlman et al.*Science*, v.302(5649), 1364-1368 (2003)

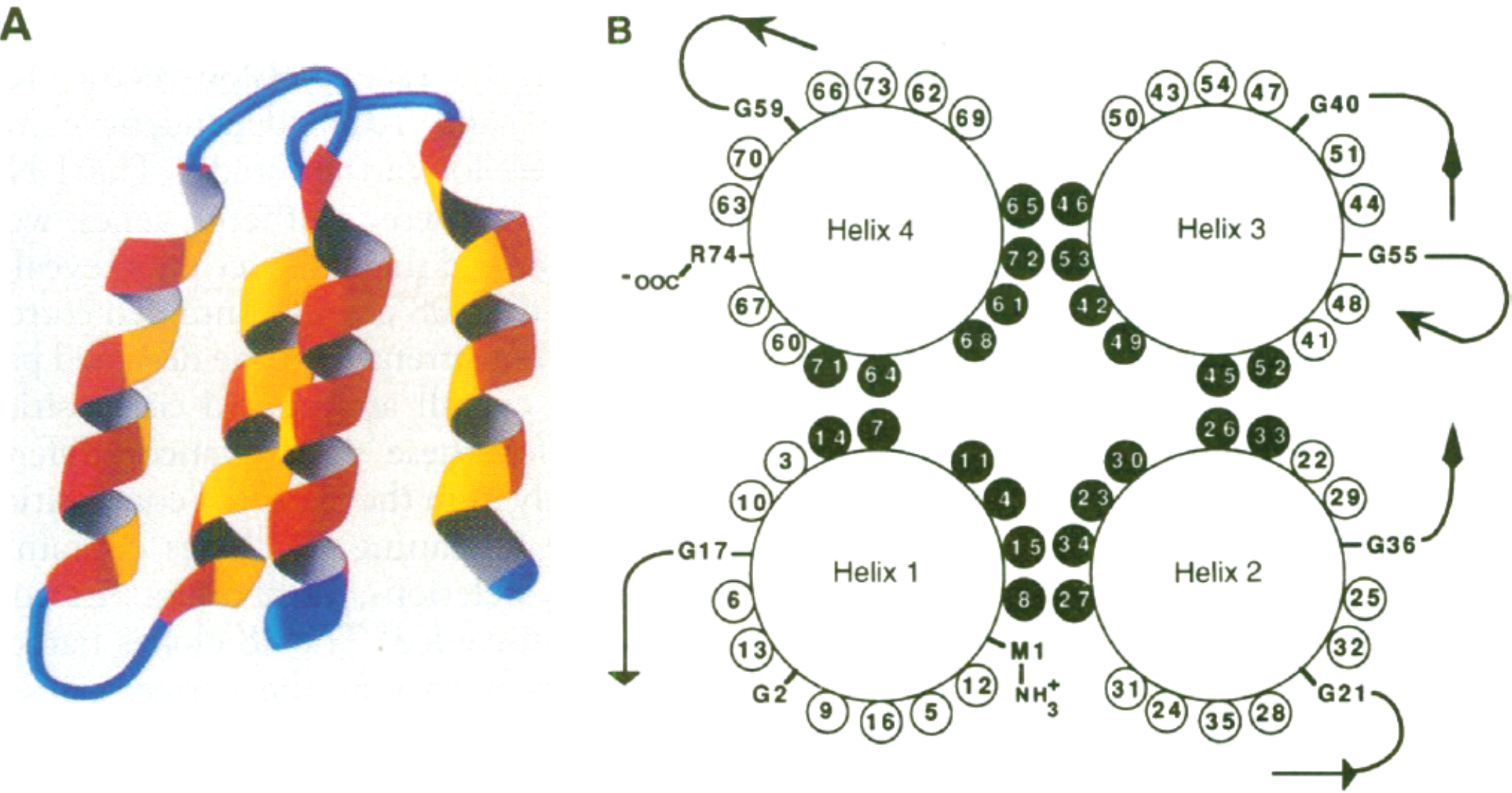
New binding sites can be designed

Used to bind arabinose, now it binds seratonin.



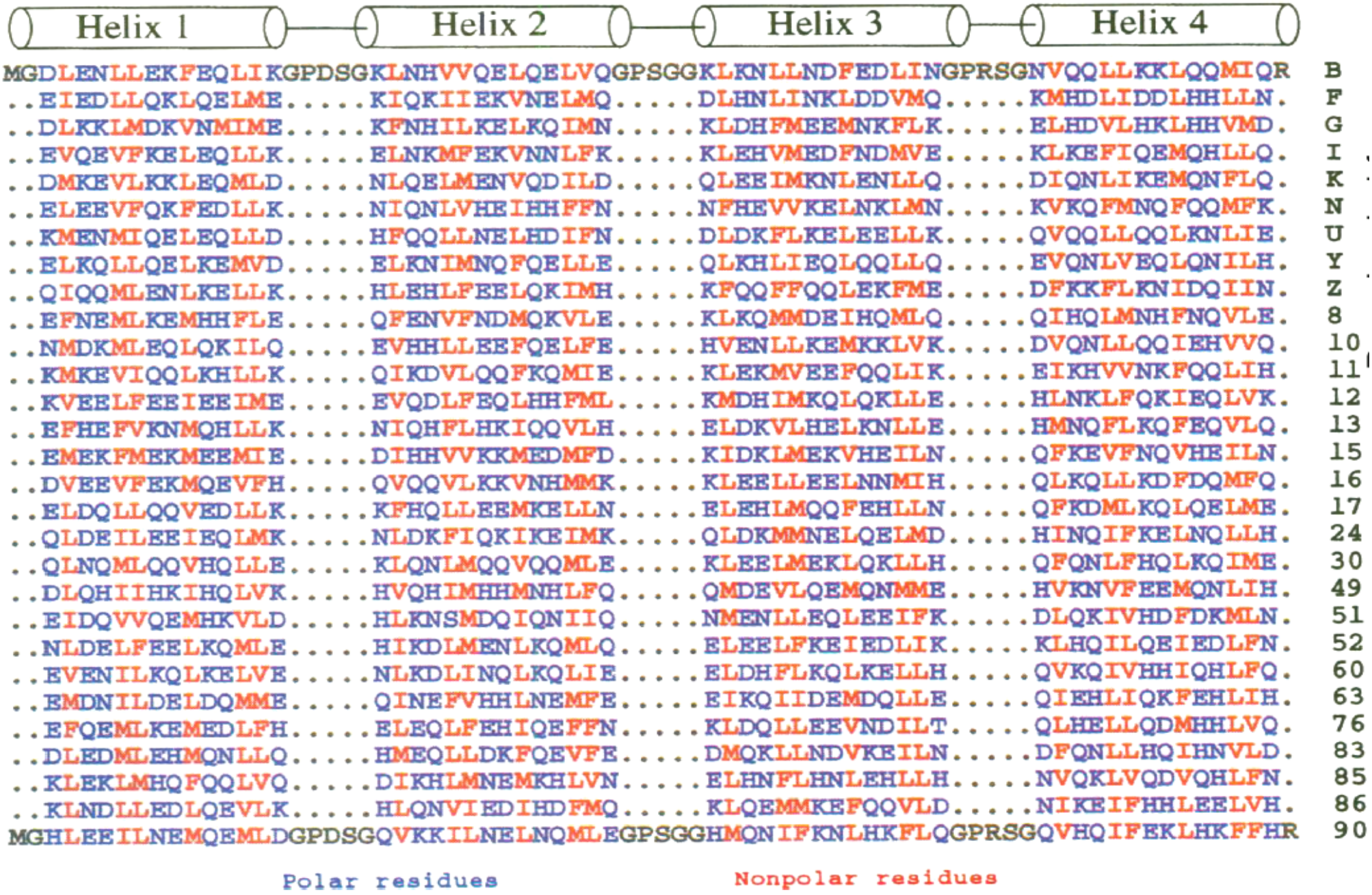
Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. *Nature* 423, 185–190 (2003).

Sequence patterning is often enough to get the protein to fold.



(Kamtekar et al, Science, 1993)

All of these sequences fold to 4-helix bundle.



Kamtekar, Satwik, et al. "Protein design by binary patterning of polar and nonpolar amino acids." Science 262.5140 (1993): 1680-1685.

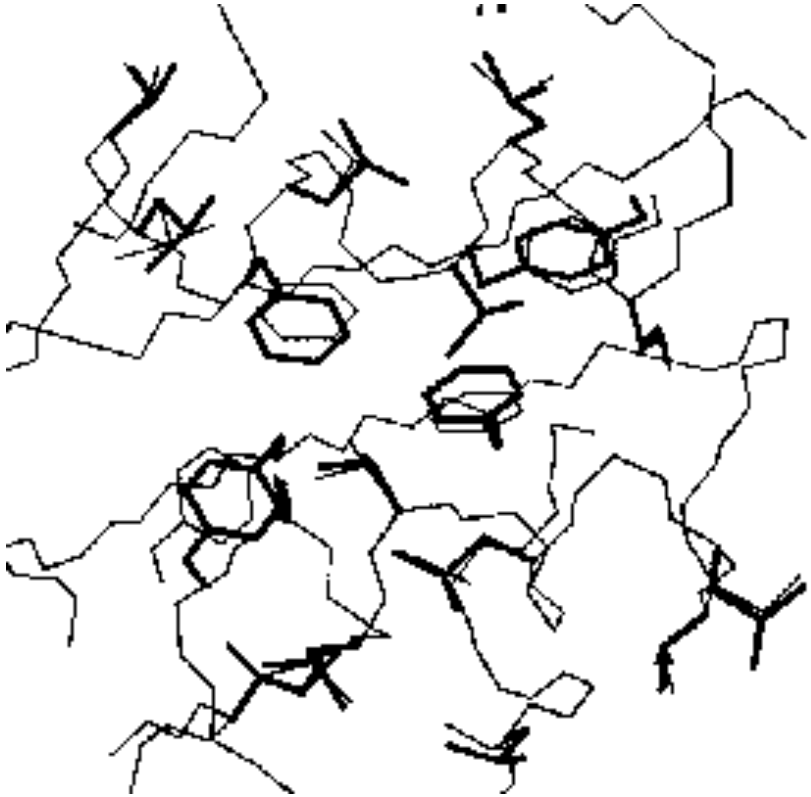
Computational design of sidechains using Dead-End Elimination

1. Select positions for mutating.
2. Choose a *palette* of amino acids at those positions.
3. For all selected positions, all palette amino acids, try all **rotamers**.
4. Chose the sequence whose rotamers give the lowest energy.

rotamer libraries facilitate protein design

Given the sequence and only the backbone atom coordinates, accurately model the *positions of the sidechains*.

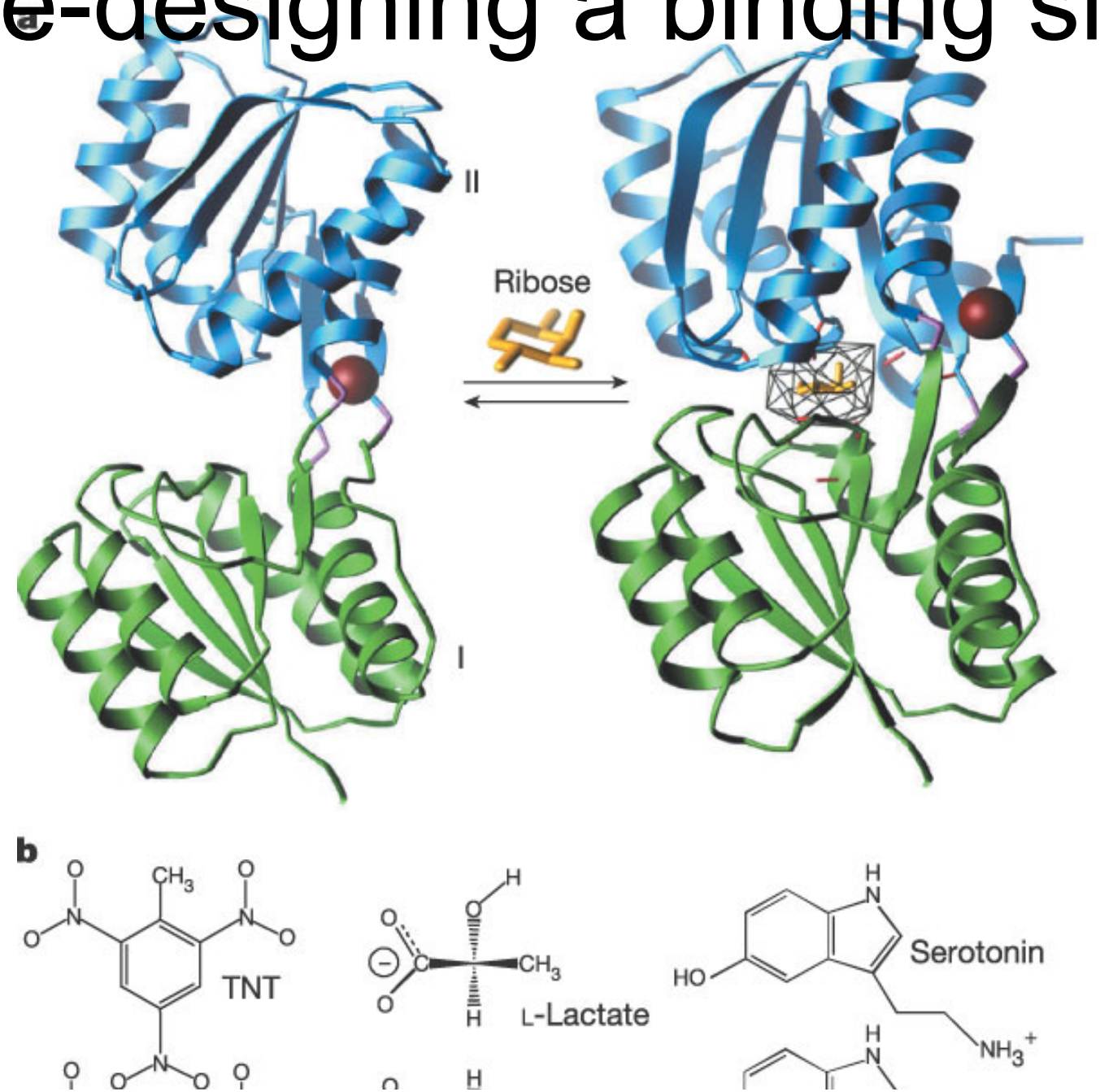
fine lines = true structure
thick lines = side chain predictions using the method of Desmet et al.



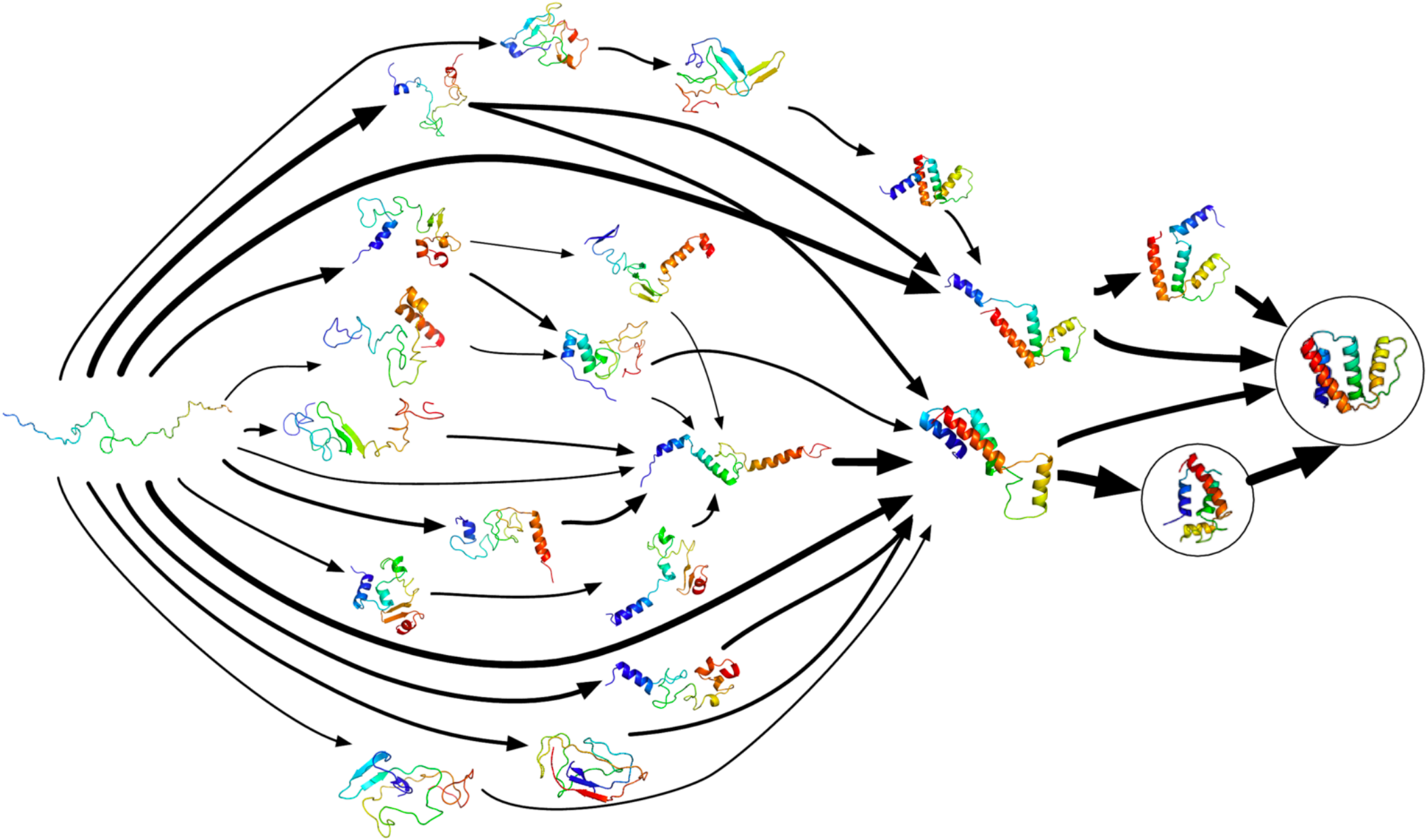
Side chain prediction using the **Dead-End Elimination** algorithm

Desmet et al, Nature v.356, pp339-342 (1992)

Re-designing a binding site

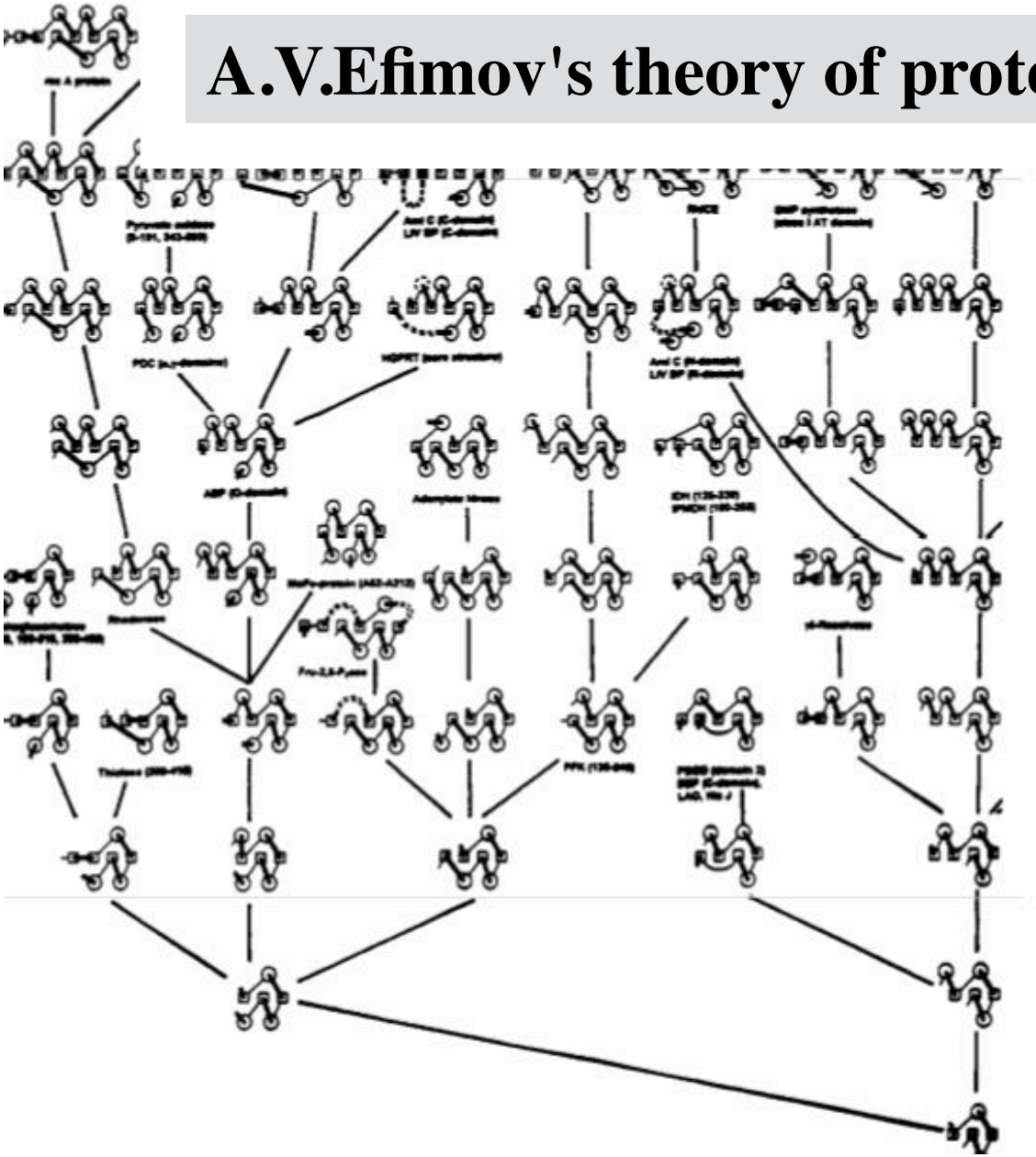


Folding pathway matters



https://commons.wikimedia.org/wiki/File:ACBP_MSM_from_Folding@home.tiff

A.V.Efimov's theory of protein evolution and folding

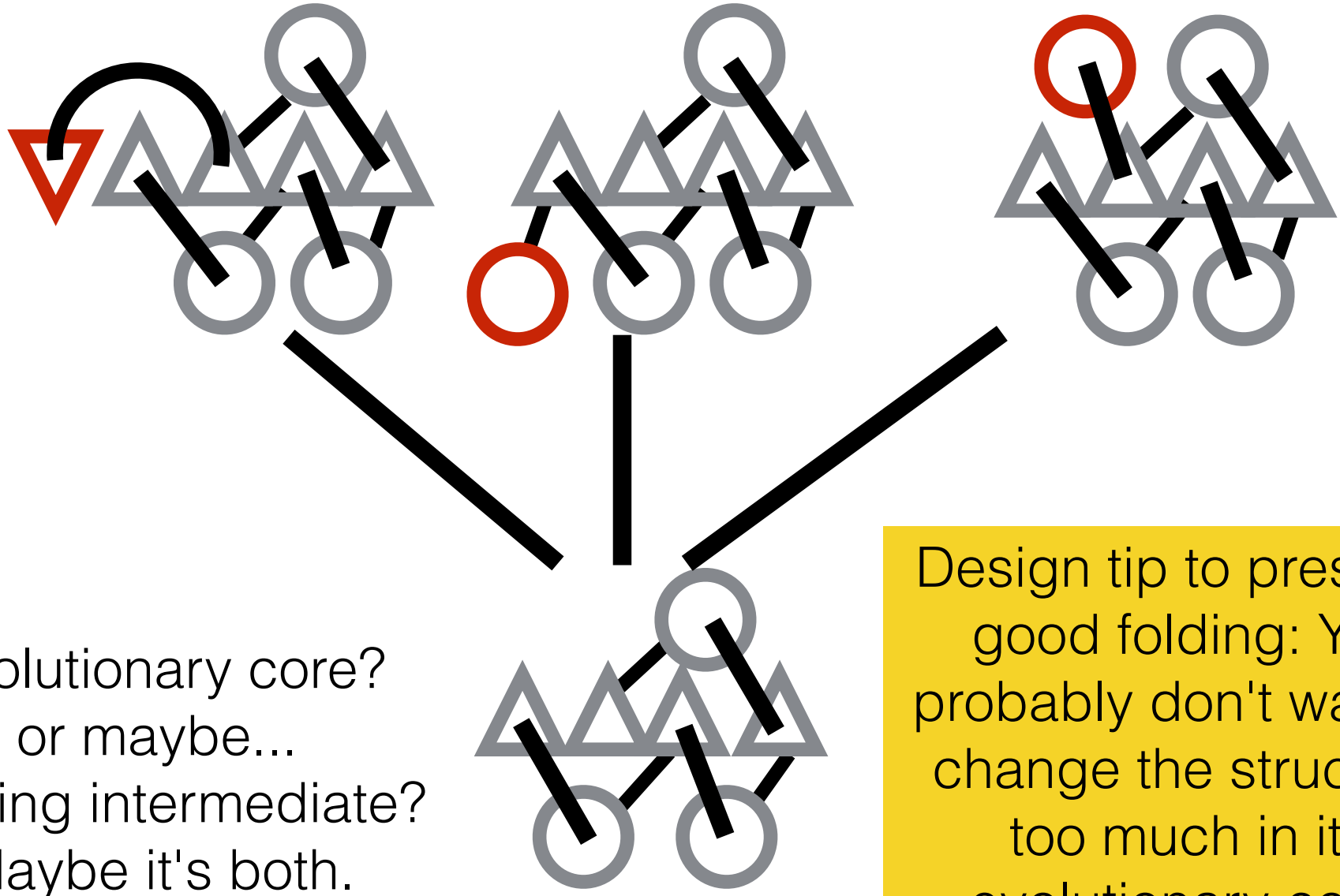


A. Efimov showed that almost all protein structures can be classified as being one of 7 trees, each starting with a motif and "growing" by one secondary structure unit at time.

Does this recapitulate evolution? the folding pathway?

Efimov, A. V. (1997). Structural trees for protein superfamilies. *Proteins: Structure, Function, and Bioinformatics*, *28*(2), 241-260.

Many proteins share common core structures (evolutionary cores)

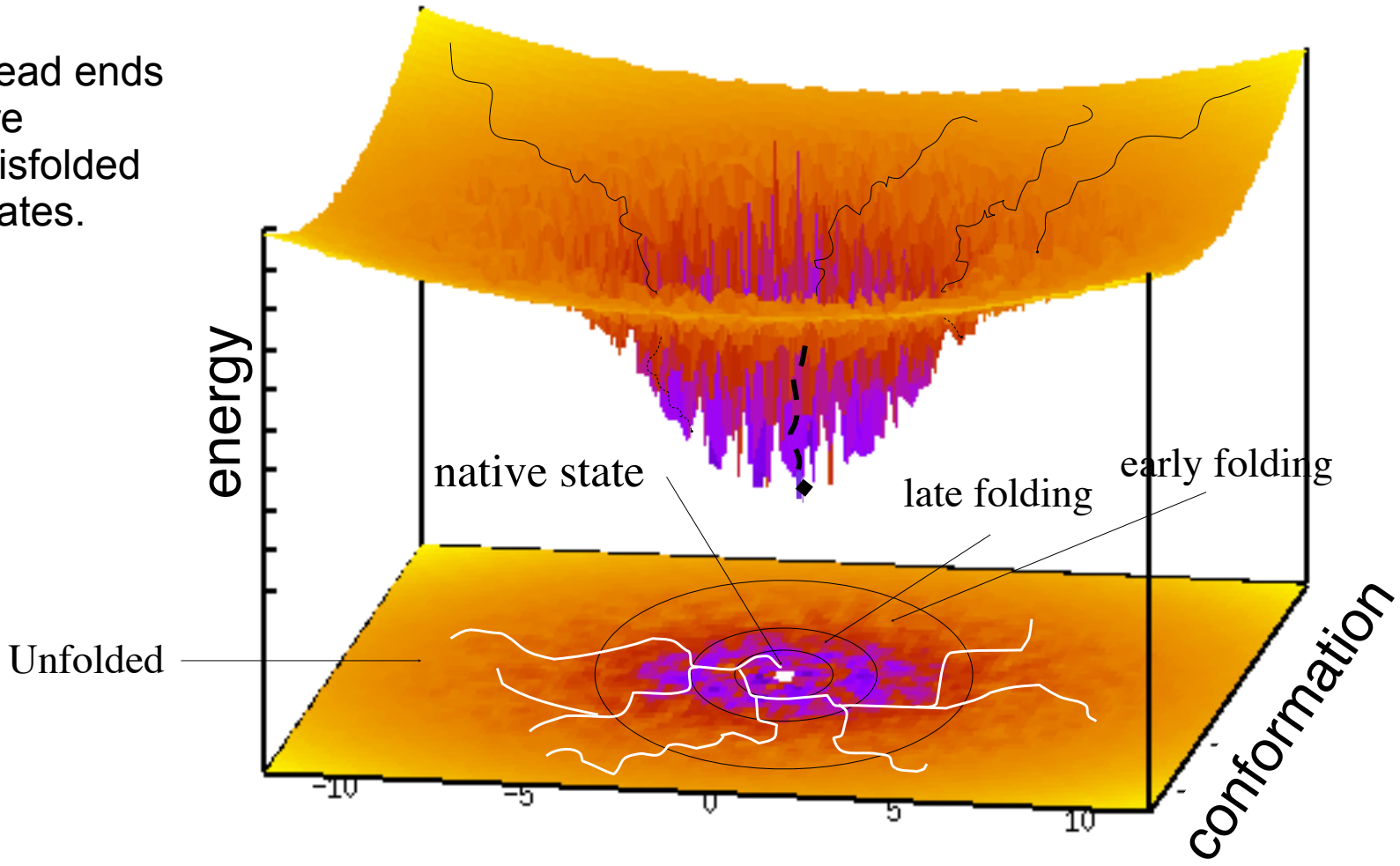


Evolutionary core?
or maybe...
Folding intermediate?
Maybe it's both.

Design tip to preserve good folding: You probably don't want to change the structure too much in its evolutionary core.

The protein folding energy landscape: ***the folding funnel***

Dead ends are misfolded states.



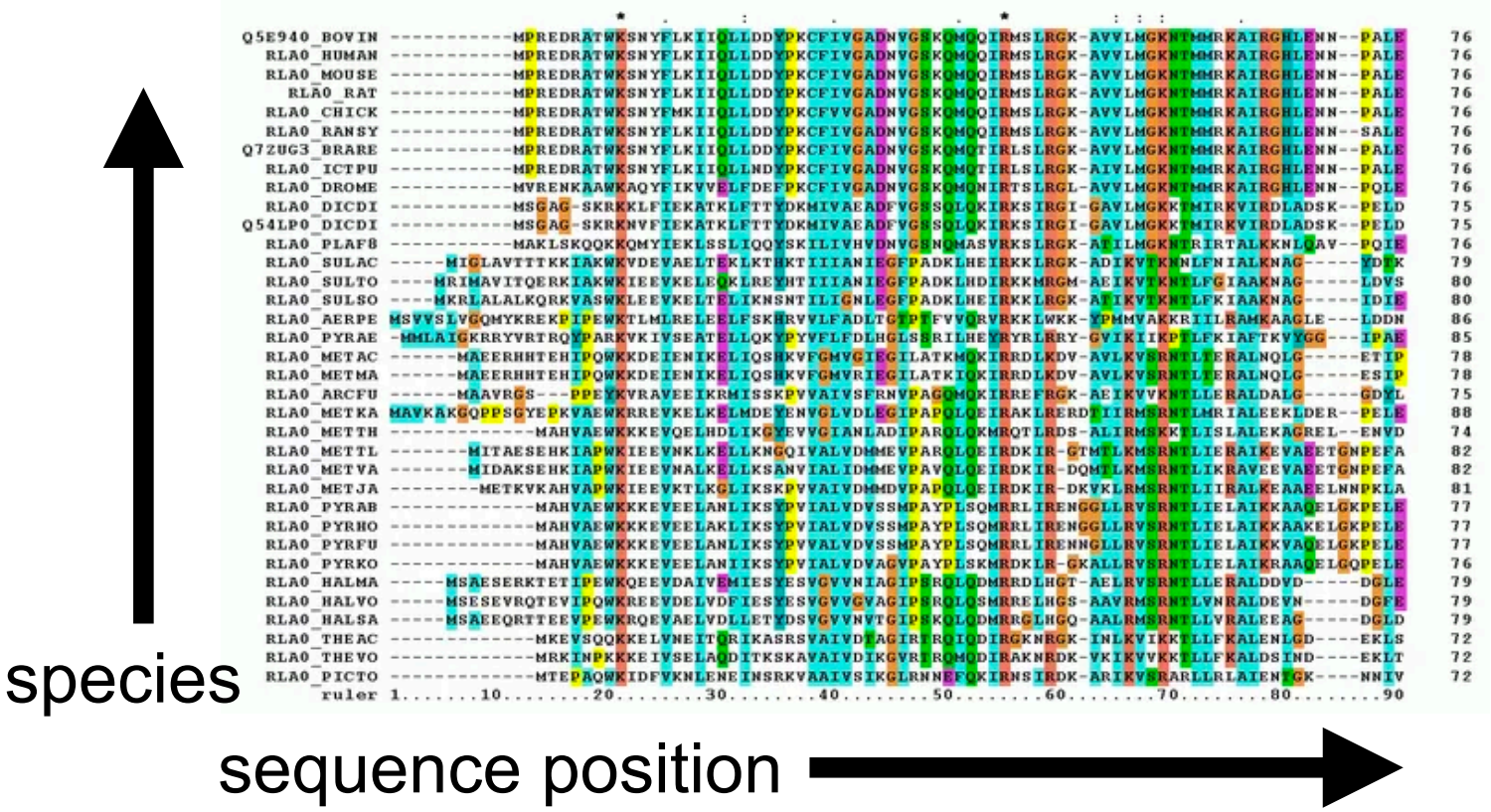
Many paths to the folded state.

Design tip to preserve good folding: You probably don't want create **non-native folding intermediates**. Don't put a helix pattern where it is supposed to be sheet, and vice versa.

Dill, K. A. (1999). Polymer principles and protein folding. *Protein Science*, *8*(6), 1166-1180.

Garcia-Manyes, S., Dougan, L., Badilla, C. L., Brujić, J., & Fernández, J. M. (2009). Direct observation of an ensemble of stable collapsed states in the mechanical folding of ubiquitin. *Proceedings of the National Academy of Sciences*

Requirements for folding are encoded in the sequence history



In protein design, we want to work within the **amino acid palette** defined by millions of years of evolution.

The protein folding pathway must be maintained

Design tip to preserve good folding: You probably don't want to change the structure too much in its **evolutionary core**.

Design tip to preserve good folding: You probably don't want create **non-native folding intermediates**. Don't put a helix pattern where it is supposed to be sheet, and vice versa.

In protein design, we want to work within the **amino acid palette** defined by millions of years of evolution.

Exercise 22.1 -- manual docking

- Download **receptor** 1TIM
- Download ligand **template** 1F2X
- Remove waters and other extraneous molecules.
- Fix the **receptor**. Unfix the **ligand**.
- Display as ribbons. Hide all atoms.
- Select the **ligand**.
- **Dock** using **alt-middlemouse, shift-alt-middlemouse, shift-middlemouse, and middlemouse** dragging.
- Done when you have *maximized number of ligand/ receptor distances between 5Å and 8Å*

The philosophy of expert protein design

- Well-trained intuition is much faster than a random search.
- There are many, many, many right answers. We don't need the very best one.
- No force field is perfect anyway. We can't avoid the need for experimental confirmation. (See Lecture 18)

Work on Homework 4