2.2-Folding

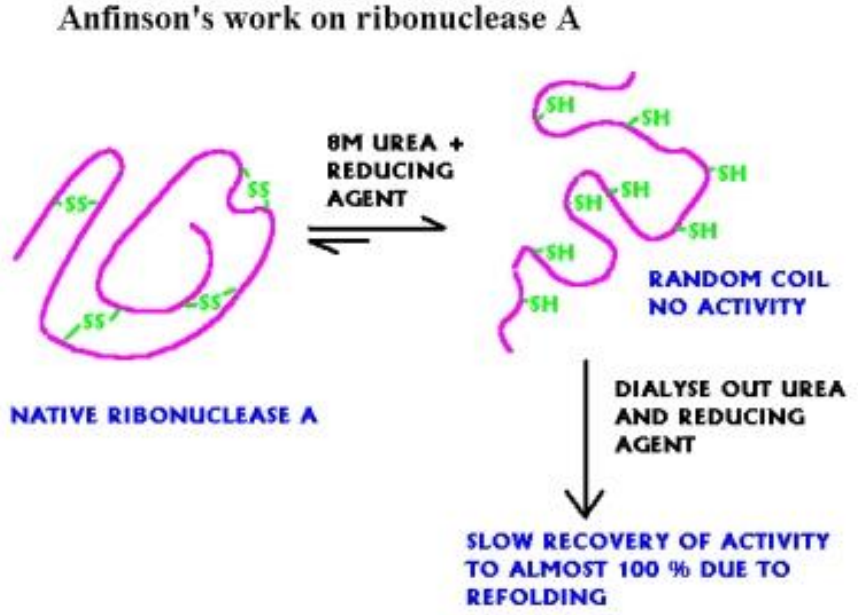
Biophysics – Bioinformatics

Outline

- The folding problem
  - The two –state model and others
- Entropic funnel model
- Folding “in vivo”
- Bioinformatics in folding. Protein structure prediction.

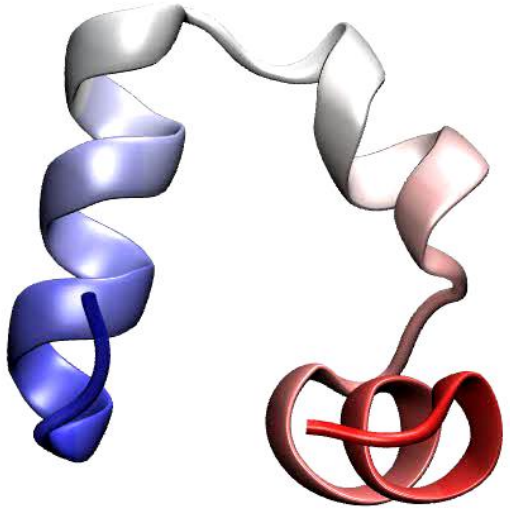
The “folding” problem

- Proteins and Nucleic Acids are synthetized as linear flexible polymers
- Proteins are kept unfolded before reaching their final location
- Folding is spontaneous, quick and cooperative
  - However, protein can get trapped in “misfolded” structures (Amiloids, “folding diseases”)
  - Proteins unfold/fold reversibly under specific conditions
  - Folding “in vivo” is efficient
- Nucleic acids unfold/fold reversibly
- Changes in media conditions (pH, T, solvents,...) denature (unfold) proteins and nucleic acids
- Levinthal paradox
  - 12 conf x torsion → 144 confs. 100 residues → 12^{200} conformations (6.8×10^{215})



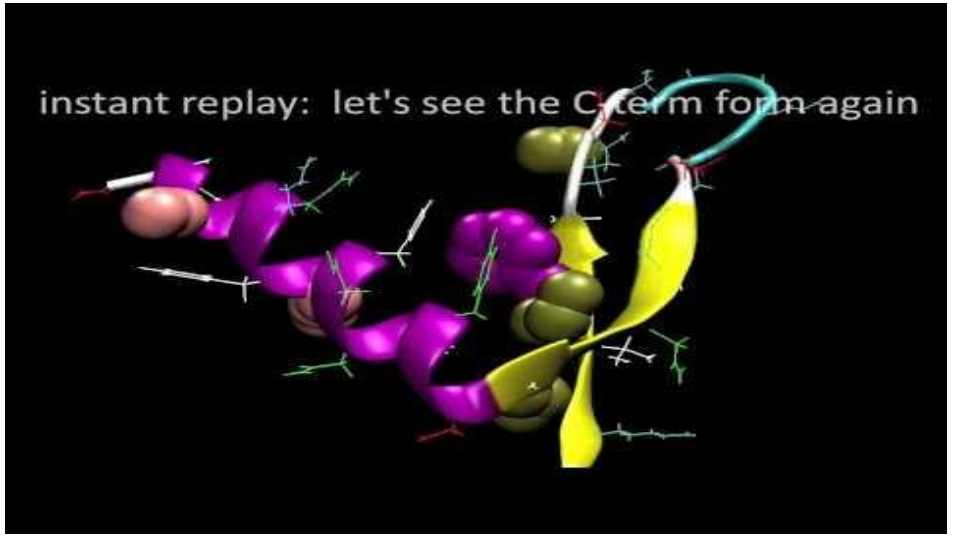
The “folding” problema from Blophysics

- Most macromolecules have a defined folded structure (but not all)
- Folded state is expected to be a thermodynamic mínimum (ΔG)
- Folding should follow a preferred kinetic pathway

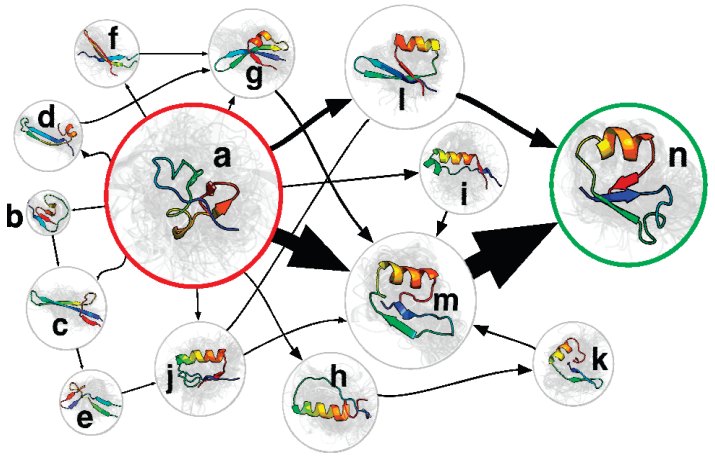


Fragment of Villin deadpiece domain (1QQV)
45 residues
3 µs Free MD (NAMD)

https://youtu.be/gFcp2Xpd29I

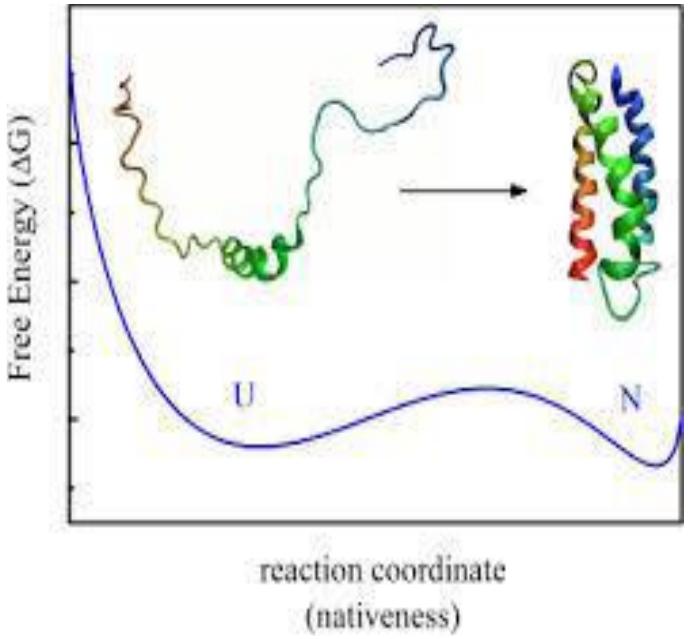


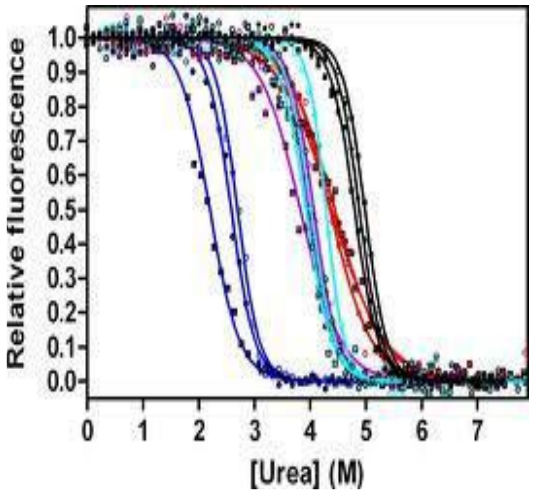
Simulation of millisecond protein folding: NTL9
Folding@home
http://pubs.acs.org/doi/abs/10.1021/ja9090353



Markov state model

Two state model: $\mathrm{F} \rightleftharpoons U$





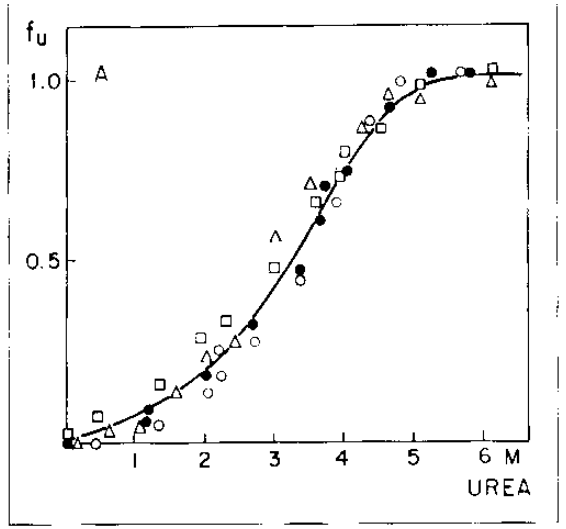


FIGURE 6–7. Urea-induced unfolding of bovine carbonic anhydrase B at pH 3.6 (A) and at pH 7.5 (B). At pH 3.6 in the absence of urea, the protein is in the molten globule state and unfolds by a one-step process. At pH 7.5, the protein is in the native state and unfolds in two steps (native–molten globule and molten globule–unfolded transitions) without additional intermediates. The fraction of unfolding is given by $f_u = (x - x_0)/(x_u - x_0)$, where x is the value of a measured parameter, x_0 is its value in the absence of urea, and x_u is its value at high concentration of urea. (A) ●, the increase of intrinsic viscosity [η], , the decrease of I_{350}/I_{360} (I is the intensity of tryptophan fluorescence at the given wavelength), , the increase of 1/P (P is the polarization of tryptophan fluorescence), , the decrease of the negative ellipticity at 220 nm. Intrinsic viscosity and the spectrum and polarization of tryptophan fluorescence both reflect the compactness of the molecule, while the ellipticity at 220 nm reflects its secondary structure. (Adapted from Rodionova et al., 1989)

$$\Delta G = \Delta G_w - m[D]$$

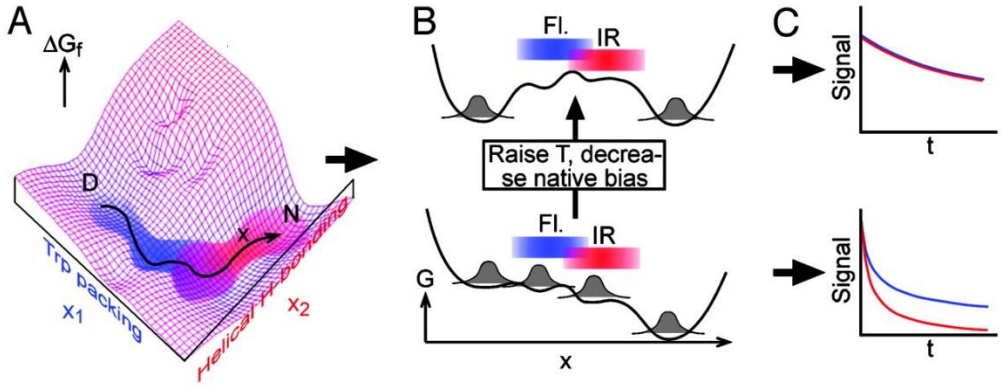
$$\Delta\Delta G = m\,\Delta\left[D_{1/2}\right]$$

m: Related with change of hydrophobic surface

Same profile independent on measured property

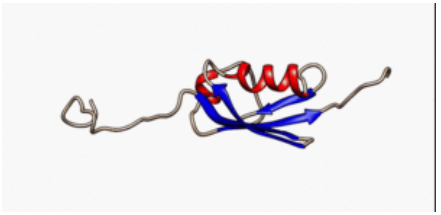
Exceptions to the rule

- Downhill folding
  - No energy barrier
  - Usually small α-rich proteins

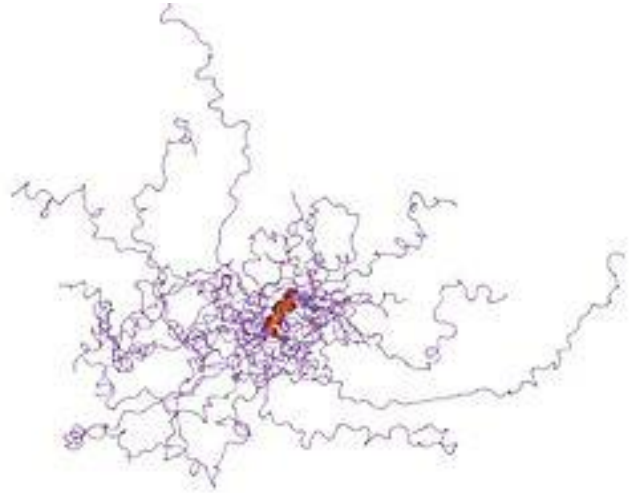


PNAS, 96, 5897, 1999; Nature 442, 317, 2006

- Intrinsically disordered proteins
  - Fold on binding to targets
    - Conformational selection
  - Alternative folded structures
  - Usually studied by NMR
  - Hybrid proteins



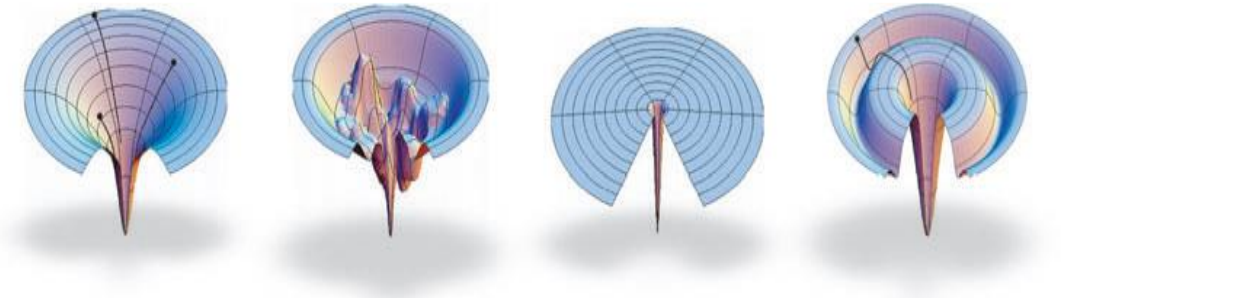
SUMO-1 (1asr)



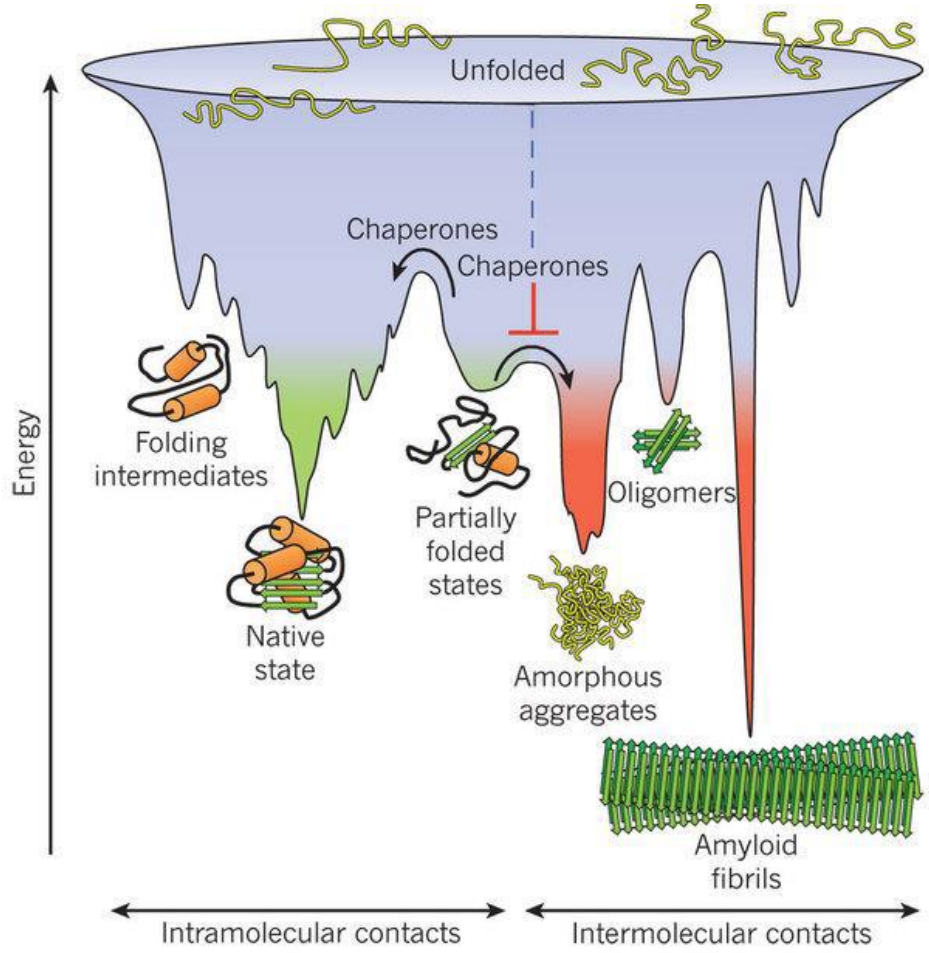
Thylacoid phosphoprotein TSP9

Entropic funnel model

- Concept of “Energy Landscape”
- Entropy and enthalpy compensate building a “funnel”
- Lower energy limits the conformational space
- The model can be adapted to most cases



Fast folder Kinetic traps Conf search Folding intermediate



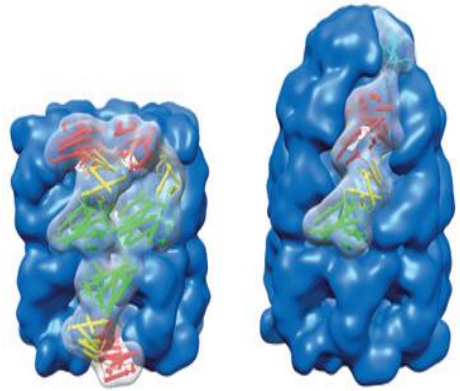
Ann. Rev. Biophys 2008, 37:289

Time scales

- Secondary structure formation << 0.01 s
- “Dry Molten globule”/folding intermediates 0.05-0.1 s
  - “molten globule” is a kind of folding intermediate that has structure but is not compacted. The name comes from traditional folding theories.
- Compaction 0.1 – 1 s
- Isomerization Pro, -SS- formation (1-5 s)

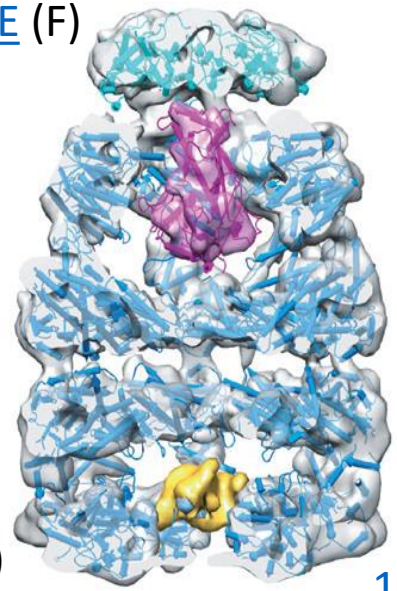
“in vivo” Folding

- “in vivo” folding is more efficient but not quicker
- Relies in keeping unfolded states stable
- Chaperons
  - Stabilize unfolded states, avoiding aggregation
  - Non-specific
- Helper enzymes
  - Accelerate particularly slow changes
  - Peptidyl-prolyl cis-trans isomerases
  - Protein disulfide isomerases



GroEL (1OEL) GroEL+GroES (1SVT)

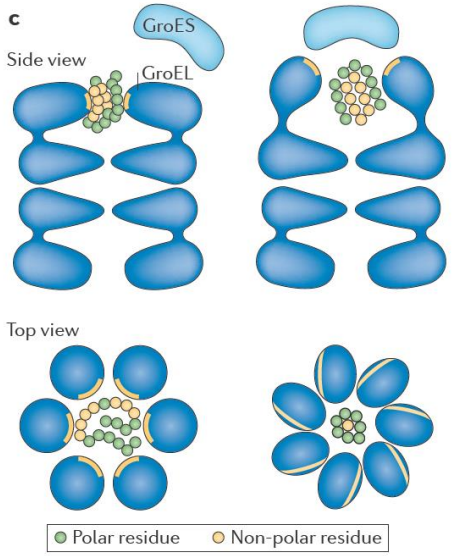
1YUE (F)



gp23 1YUE(U)

1AON

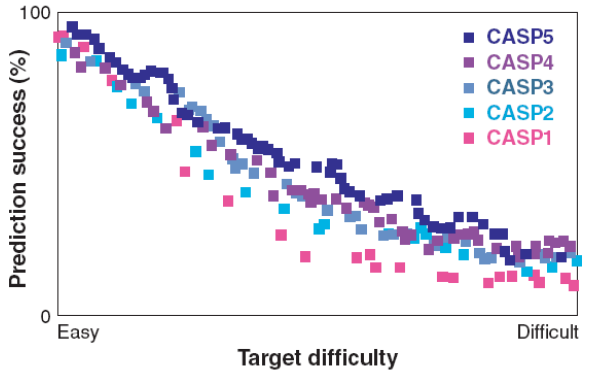
EMD-1548



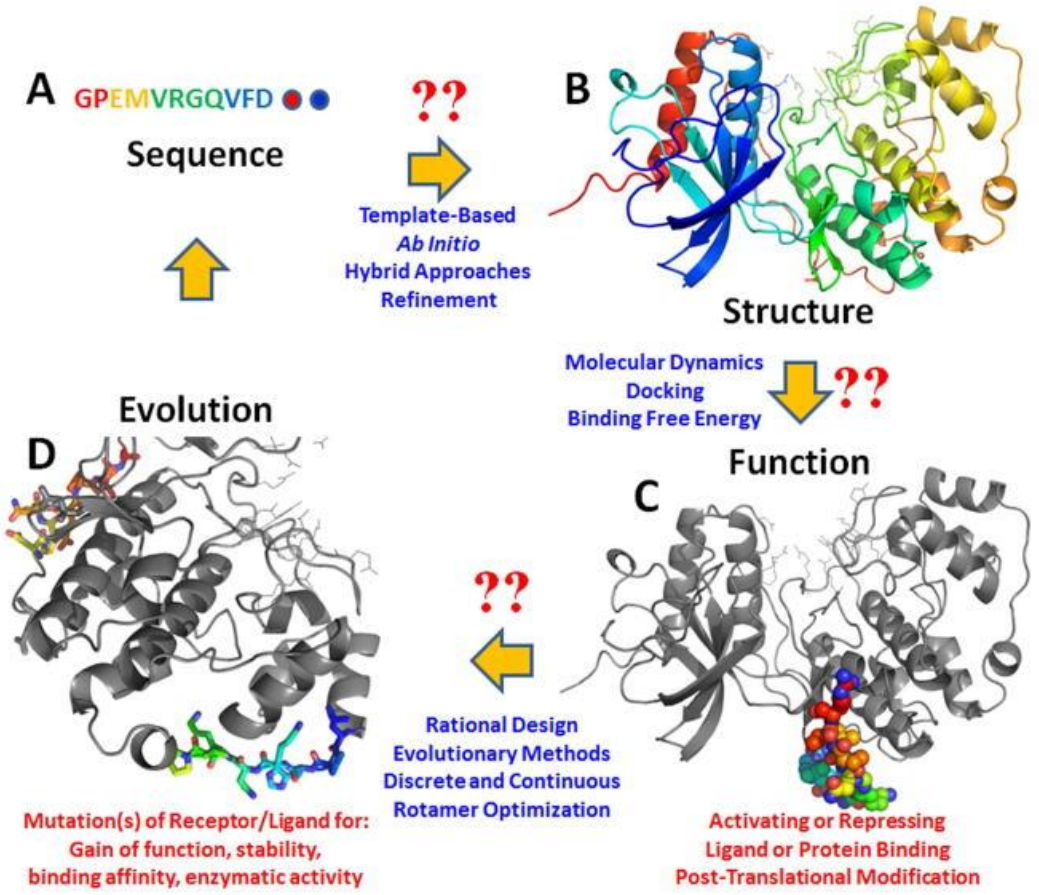
Nat. Rev. Mol Cell Biol. (2013) 14: 630

Bioinformatics in folding

- Protein Structure prediction
  - Ab-initio MD simulations (atomistic and coarse-grained)
  - Contact predictions -> Alphafold
    - Fold recognition
  - Comparative modelling
  - Fragment based prediction
  - Disorder prediction
  - Now almost superseeded by AI!!
- Rational Design
- Directed evolution
- De-novo design



Computer assisted structure prediction (CASP) contests



Trends. Biotechnol 2014 32:99