

Paradigm shift in protein science

Central paradigm

- old idea: sequence determines structure, which determines function
- new idea: sequence determines structure *which determines dynamics* which determines function

Native state(s) of a protein

- old idea: a single, static structure (native state)
- new idea: a structural ensemble
- ubiquitin example: induced fit (different states caused by binding) vs conformational selection (protein exists in multiple forms, substrate binds the correct state)

Flow of genetic information

- DNA is a chain of 4 nucleotides
- proteins are chains of 20 amino acids
- dogma of molecular biology: DNA \rightarrow RNA \rightarrow protein
- genetic code: possible codons (triplets) of 4 nucleotides: $4^3 = 64$; single base changes usually result in same amino acid or an amino acid with similar physical properties

Protein structure

- protein is a chain of amino acids
- amino acids share a backbone, distinguished by the side chains
- classified as hydrophobic, hydrophilic, amphipathic
- hydrophobic
 - engage in van der Waals interactions only; cannot form hydrogen bonds with water
 - normally on interior of protein, packing against each other
 - basis of **hydrophobic effect**
- hydrophilic: form hydrogen bonds with water
- amphipathic: can form hydrogen bonds, but have a large hydrophobic component

Protein flexibility

- bond length and bond angles are (nearly) fixed
- peptide group is planar
- flexibility comes from the torsional angles phi (ϕ) and psi (ψ) (two degrees of freedom on the backbone per residue)

Side chain flexibility

- backbone flexibility comes from ϕ and ψ
- do torsional rotations exist in side chains? how many degrees of freedom are there for each side chain?
- what is the effect of chain torsional rotation?
- effects of side chain rotation are local, so it is less constrained

Protein conformations

- a conformation is a 3D form that a protein can adapt without covalent bond breakage
- the set of all possible conformations are is called the conformational space; this space is high-dimensional
- Ramachandan plot and steric constraints: brighter areas show preferred combinations of ϕ and ψ

Folding and the native structure

- native state is 3D, compact, and lowest energy
- folding is the process by which a protein finds a single native state within a large conformational space
- Levinthal paradox; conformational space is huge, but protein folds almost instantaneously