Protein Structure

(a) MOLECULAR STRUCTURE
- Primary (sequence)
- Secondary (local folding)
- Tertiary (long-range folding)
- Quaternary (multimeric organization)
- Supramolecular (large-scale assemblies)

(b) Regulation, Signaling, Structure, Movement, Catalysis, Transport
Protein Structure

Primary structure

Amino acid residues:
- Lys
- Lys
- Gly
- Gly
- Leu
- Val
- Ala
- His

Secondary structure

α Helix

Tertiary structure

Polypeptide chain

Quaternary structure

Assembled subunits
Central Dogma of Biology

Gene 1 → Transcription of DNA sequence into RNA sequence → RNA 1
Gene 2 → Transcription of DNA sequence into RNA sequence → RNA 2
Gene 3 → Transcription of DNA sequence into RNA sequence → RNA 3

RNA 1 → Translation (on the ribosome) of RNA sequence into protein sequence and folding of protein into native conformation → Protein 1
RNA 2 → Translation (on the ribosome) of RNA sequence into protein sequence and folding of protein into native conformation → Protein 2
RNA 3 → Translation (on the ribosome) of RNA sequence into protein sequence and folding of protein into native conformation → Protein 3

Protein 1 + Protein 2 + Protein 3 → Formation of supramolecular complex
Protein Folding in the Cell
The Hydrophobic Effect

Dispersion of lipids in H₂O
Each lipid molecule forces surrounding H₂O molecules to become highly ordered.

Clusters of lipid molecules
Only lipid portions at the edge of the cluster force the ordering of water. Fewer H₂O molecules are ordered, and entropy is increased.
The Hydrophobic Effect
it drives binding
The Hydrophobic Effect

it drives folding
Protein Structure
hydrophobic residues are buried

blue: hydrophobic
Protein Folding

Protein unfolding results in loss of function

Native state; catalytically active.

addition of urea and mercapto-ethanol

Unfolded state; inactive. Disulfide cross-links reduced to yield Cys residues.

removal of urea and mercapto-ethanol

Native, catalytically active state. Disulfide cross-links correctly re-formed.
Protein Folding

Protein unfolding results in loss of function

Energy difference

~8-10 kcal mol\(^{-1}\)
Protein Folding

Protein unfolding results in loss of function

![Graph showing the percent unfolded vs. [GdnHCl], M for Ribonuclease A. The graph includes the molecule structures of Urea and Guanidinium chloride. The point Tm indicates the transition midpoint.]
Protein Folding

Protein (un)folding is highly cooperative
Protein Folding

How does a protein arrive at its native structure?

The Levinthal’s paradox: a protein would take $\sim 10^{77}$ years to fold

Two folding models: the hierarchical and through the “molten globule”
Protein Folding

Energy landscape

Diagram showing the energy landscape of protein folding, with various intermediates, states, and aggregates labeled.
Protein Folding

Death by misfolding: the Prion diseases
Protein Folding
Chaperones assist newly synthesized proteins to fold

1. DnaJ binds to the unfolded or partially folded protein and then to DnaK.
2. DnaJ stimulates ATP hydrolysis by DnaK. DnaK–ADP binds tightly to the unfolded protein.
3. In bacteria, the nucleotide-exchange factor GrpE stimulates release of ADP.
4. ATP binds to DnaK and the protein dissociates.

Folded protein (native conformation)
To GroEL system
Partially folded protein

Unfolded protein

Protein Folding

Chaperones assist newly synthesized proteins to fold

1. Unfolded protein binds to the GroEL pocket not blocked by GroES.

2. ATP binds to each subunit of the GroEL heptamer.

3. ATP hydrolysis leads to release of 14 ADP and GroES.

4. 7 ATP and GroES bind to GroEL with a filled pocket.

5. Protein folds inside the enclosure.

6. The released protein is fully folded or in a partially folded state that is committed to adopt the native conformation.

7. Proteins not folded when released are rapidly bound again.
Protein Folding
Chaperones assist newly synthesized proteins to fold

The GroES-EL chaperonin
Protein Folding
Anfinsen’s hypothesis: Sequence determines structure

One sequence, one fold
Protein Folding

One sequence, one fold?

PNAS, 2008, 105, 5057
Intrinsically Unstructured Proteins

Increasing content of stable three-dimensional structure

- **Unstructured** (conformational ensemble), for example, ACTR (no NCBD)
- **Molten globule** (conformational ensemble), for example, NCBD (no ACTR)
- **Linked folded domains** (beads on a string), for example, zinc fingers (no DNA)
- **Mostly folded, local disorder**, for example, eIF4E (N terminus is unfolded)

Folding on target binding

- ACTR–NCBD complex
- Zinc-finger-1–3–DNA complex
- eIF4E–eIF4G complex

Nat. Rev. Mol. Biol., 2005, 6, 199
Intrinsically Unstructured Proteins

Coupled folding and binding