Protein structure and folding

Levels of protein structure
Theory of protein folding: Anfinsen’s experiment, Levinthal’s paradox, the folding funnel mode

04.09.2013.

- Amino acids and protein structure
- Protein folding (and thermodynamics)
Protein

**Protein:** A linear polymer of **amino acids** linked together by **peptide bonds** in a specific sequence.

**Linear:** not branched.
**Polymer:** Large molecule made by **covalently** linking several identical or similar units (**monomers**) together.
**Oligomer:** Short polymer (Greek **oligos**, few, little.) (dimer, trimer)
**Polymerisation:** The process by which molecules are linked together to form polymers.

**Function:**
- structural and skeletal proteins (collagene)
- transport function (myosin)
- biochemical functions (enzymes)
- immunological functions (antibodies)
- signal transduction (hormones)

Amino acid

- A group of organic molecules that contains a basic **amino group** (-NH₂), an acidic **carboxyl group** (-COOH), and an organic **R** group (or side chain), which is unique to each amino acid.
Amino acids

- A unit („brick”) within a protein molecule.

- **20 amino acids** building up the proteins within the human body.

- Essential amino acids (9 pcs): the human body can not produce them in a proper amount (methionine).

- To build a 100 amino acid long polymer from the 20 amino acids – the number of the variation is huge ($20^{100} = \sim 1.3\times10^{130}$).

- The order of the amino acids = **amino acid sequence → primary structure**

Peptide bound

Forms the backbone of the protein.
Secondary structure
Forming a spatial-structure: alpha helix & beta sheet

Alpha helix
Single, spiral chain of amino acids stabilized by hydrogen bonds. Rolling up on the surface of an imaginary cylinder (+ or -).
H-bond within the alpha helix

An electrostatic dipole-dipole interaction that involves a hydrogen atom (electronegativity).

Beta sheet

Two or more adjacent parallel polypeptide chains stabilized by hydrogen bonds between the chains.
Tertiary structure

- **Folding**: Forming the final, functional 3D forms of the proteins.

- Under physiological conditions a spontaneous disorder ↔ order transition = „folding”

- Chaperon: A guide or companion to the protein that help to form its tertiary structure.

- Disulfide, hydrogen bound and hydrophobic interactions are stabilizing the folded protein.

Disulphide bounds

A covalent bound (primary) forming between thiol groups (Cysteine).
Hydrophobic interactions

• Highly hydrophobic amino acids:
  Valine, isoleucine, leucin, methionine, phenylalanine, cysteine, tryptophan.
• They are usually non polar molecules.
• minimizing the number of hydrophobic side-chains exposed to water is one of the principal driving forces behind the protein folding
• the hydrophobic amino acids are shielded (buried) from the aqueous solvent.

Tertiary structure
**Quaternary structure**

Two or more peptide chains forming the complex functional form of a protein (e.g. hemoglobin – PDB: 2HHB).

---

**Folding**

Forming the tertiary structure (functional form) of a protein.
Misfolding

The folding is not succesfull (e.g. beta sheets instead of alpha helices) → misfolded proteins

The cell remove the wrong protein → the amount of the functional proteins decrease

The cell will not remove it
→ low functionality (Sickle cell anemia)
→ deposits (plaque) within the cells (Alzheimer disease)

Sickle cell anaemia

α2 (141 aa.)  β2 (146 aa.)

6Glutamic acid → 6Valine  6Glutamic acid → 6Valine
### Protein folding diseases

<table>
<thead>
<tr>
<th>DISEASE</th>
<th>PROTEIN</th>
<th>SITE OF FOLDING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypercholesterolaemia</td>
<td>Low-density lipoprotein receptor</td>
<td>ER</td>
</tr>
<tr>
<td>Cystic fibrosis</td>
<td>Cystic fibrosis trans-membran regulator</td>
<td>ER</td>
</tr>
<tr>
<td>Phenylketonuria</td>
<td>Phenylalanine hydroxilase</td>
<td>Cytosol</td>
</tr>
<tr>
<td>Huntington’s disease</td>
<td>Huntingtin</td>
<td>Cytosol</td>
</tr>
<tr>
<td>Marfan syndrome</td>
<td>Fibrillin</td>
<td>ER</td>
</tr>
<tr>
<td>Osteogenesis imperfecta</td>
<td>Procollagen</td>
<td>ER</td>
</tr>
<tr>
<td>Sickle cell anemia</td>
<td>Haemoglobin</td>
<td>Cytosol</td>
</tr>
<tr>
<td>α1-Antitrypsin deficiency</td>
<td>α1-Antitrypsin</td>
<td>ER</td>
</tr>
<tr>
<td>Tay-Sachs disease</td>
<td>β-hexosaminidase</td>
<td>ER</td>
</tr>
<tr>
<td>Scurvy</td>
<td>Collagen</td>
<td>ER</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>Amyloid β-peptidergase</td>
<td>ER</td>
</tr>
<tr>
<td>Parkinson’s disease</td>
<td>α-synuclein</td>
<td>Cytosol</td>
</tr>
<tr>
<td>Creutzfeldt-Jakob disease</td>
<td>Prion protein</td>
<td>ER</td>
</tr>
<tr>
<td>Familial amyloidoses</td>
<td>Transthyretin lysozyme</td>
<td>ER</td>
</tr>
<tr>
<td>Retinitis pigmentosa</td>
<td>rhodopsin</td>
<td>ER</td>
</tr>
<tr>
<td>Cataract</td>
<td>crystallin</td>
<td>Cytosol</td>
</tr>
<tr>
<td>Cancer</td>
<td>P53</td>
<td>Cytosol</td>
</tr>
</tbody>
</table>

### Protein folding

- Forming the three-dimensional (3D) native **structure** → biological **function**
- The pioneering work of Max Ferdinand Perutz and John Cowdery Kendrew
  - The Nobel Prize in Chemistry 1962 "for their studies of the structures of globular proteins"
Myoglobin (pdb: 1mbn)


Protein folding

As of Tuesday Aug 27, 2013 at 5 PM PDT there are 93436 structures at www.pdb.org
Main questions!

1. How the 3D native **structure** of a protein determined (physical folding code) – what will be the result?
2. Folding **mechanism** - how will the result be produced (kinetics)?
3. How is it possible to **predict** the structure of a protein?

Factors affecting the 3D structure formation

- hydrogen bonds
- van der Walls interactions (close-ranged interactions)
- backbone angle preferences
- hydrophobic interactions
- disulphide bonds
- chain entropy
  - According to the laws of **thermodynamics**, systems tend toward their states of lowest free energy
Gibbs free energy [Joule]

G = H - TS

A spontaneous process is accompanied by a decrease in the Gibbs energy at constant temperature and pressure. At constant temperature and pressure the change in the Gibbs energy is equal to the maximum non-expansion work accompanying a process.

Anfinsen’s experiment

Christian Boehmer Anfinsen (biochemist-USA) - 1957

• sufficient information contained in the protein sequence to determine the protein structure!
• THERMODYNAMIC HYPOTHESIS: The native structure is the GLOBAL minimum of free energy.
Folding mechanism

• Cyrus Levinthal (1968) – the Levinthal’s paradox
  How, despite the huge number of conformations accessible to it, a protein molecule can fold to its one precisely defined native structure so quickly (sometimes microseconds).

• Polymer statistical thermodynamics → more compact, **low-energy** conformational ensembles have **fewer conformations** → funnel-shaped protein-folding energy landscapes

The folding funnel

• **Large number** of folding path with **equal** probability (↔ one folding path in Levinthal’s idea).
• All paths lead directly to the native state (energetic minimum).
• The **depth** of the well symbolize the energetic stabilization of the native state versus the denatured state.
• The **width** of the well symbolize the entropy of the system.
• The **surface** outside the well symbolize the heterogeneity of the random coil state.
• **Directed search**
• The end!