Formation of Biological Structures

Szabolcs Osváth

Semmelweis University

How Big Are They?

2 nm

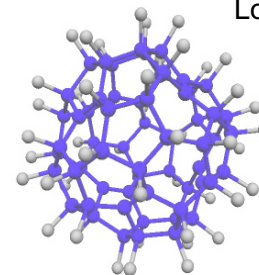Figure 1.1 Physical Biology of the Cell, 2ed. (© Garland Science 2013)

"Plenty of Room at the Bottom"

" The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It is not an attempt to violate any laws; it is something, in principle, that can be done; but in practice, it has not been done because we are too big."
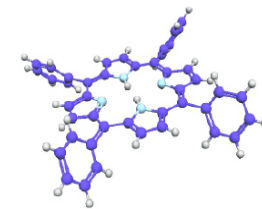
Richard Feynman, 1959

Wave – Particle Duality

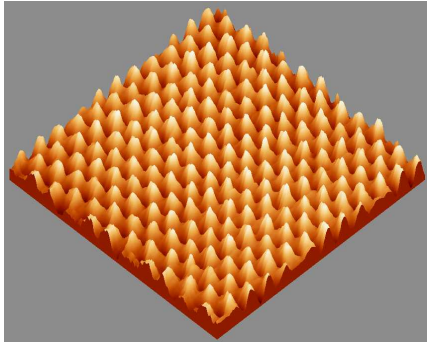Louis De Broglie: $\lambda = h/p$

fluorofulleren
$C_{60}F_{48}$
1632 Da

tetraphenilporfirin $C_{44}H_{30}N_4$

L Hackermuller, S Uttenthaler, K Hornberger, E Reiger, B Brezger, A Zeilinger, M Arndt;   Phys. Rev. Lett. 91 (2003) 90408

1

Wave – Particle Duality
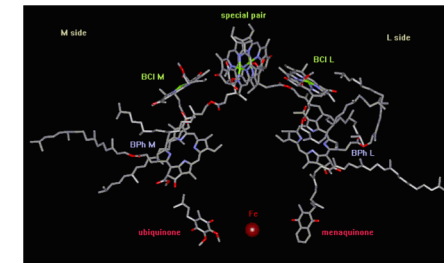

Scanning Tunneling Microscope (STM) image

Structure – Function Relationship

From the molecular level to ecosystems, there is a strong relationship between structure and function of biological systems.
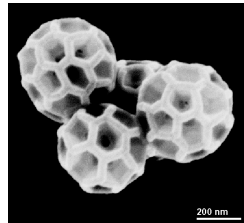
Hartmut Michel, Johann Deisenhofer, Robert Huber
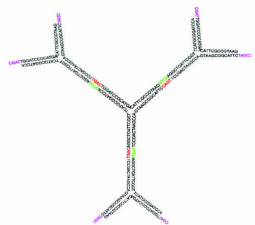1982 – 3D structure of the bacterial photosynthetic reaction center
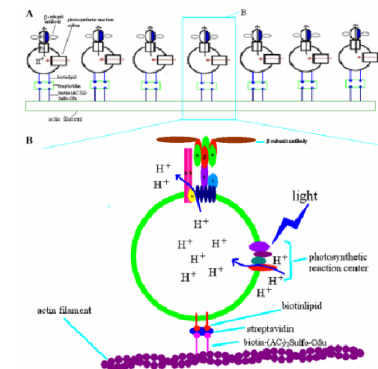1988 – Nobel prize

Self Assembly

Molecular recognition
(e.g. self assembly of DNA molecules into „balls").

Light - Driven Swimming Structure

2

Biopolymers

reaction	$t_{1/2}$ @ 25 ºC	$t_{1/2}$ @ 100 ºC	typical number of monomeric units in a polymer molecule	number of different monomers
DNA hydrolysis	140 000 years	22 years	$3 \cdot 10^9$ (human DNA)	4
RNA hydrolysis	4 years	9 days	few dozen (tRNA)	4
protein hydrolysis	400 years	5.5 weeks	few hundred	20

Biopolymers

NUCLEIC ACIDS PROTEINS

ALPHABET — nucleotides — amino acids

WORDS — codon — α-helix — β-strand
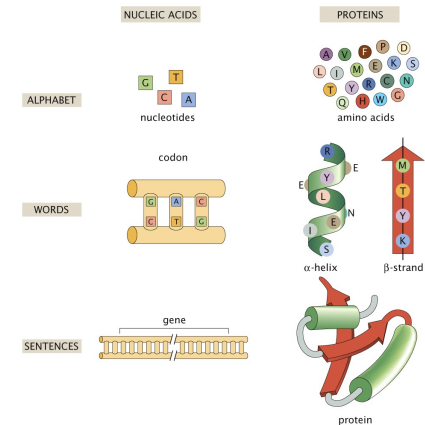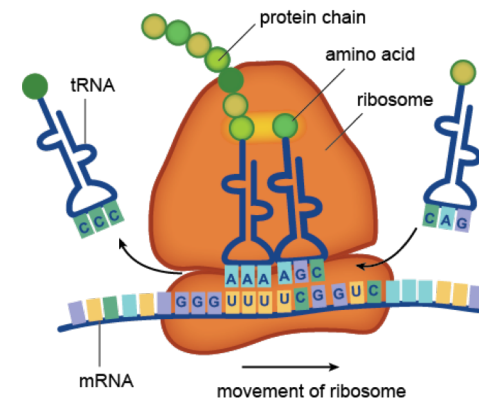
SENTENCES — gene — protein

Figure 1.2 Physical Biology of the Cell, 2ed. (© Garland Science 2013)
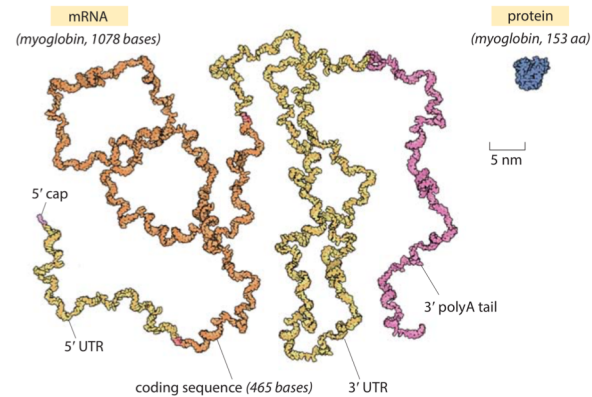
Role of RNA in Living Systems

- messenger (mRNS)
- ribosomal (rRNS)
- transfer (tRNS)
- regulator
- enzime (ribozim)
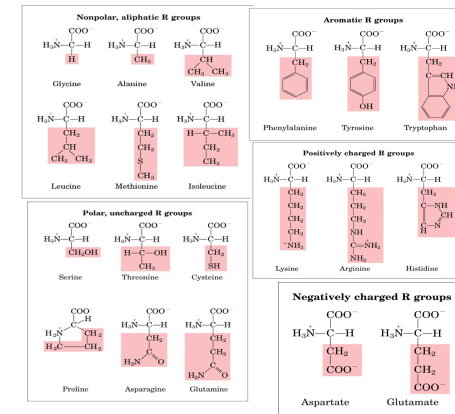- switch (riboswitch)
- virus gene RNS

Ribosome Function

protein chain
amino acid
tRNA
ribosome
mRNA
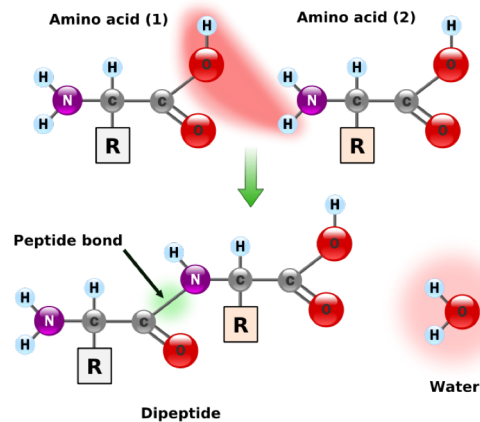movement of ribosome

Relationship Between RNA Code and Protein Sequence

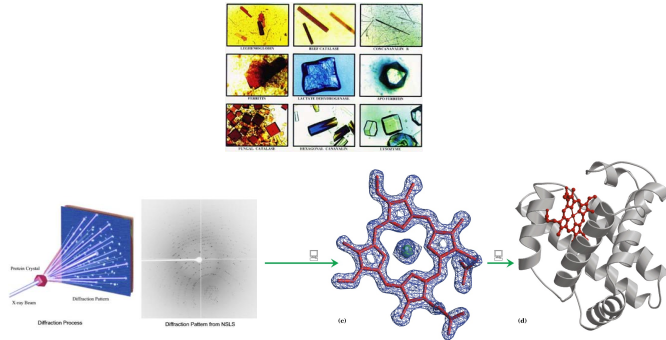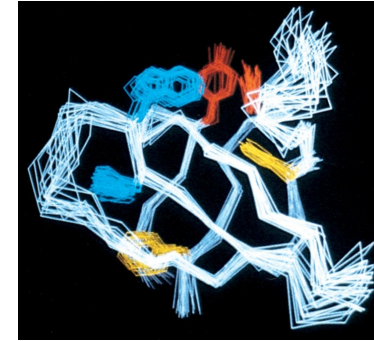## The Twenty Standard Amino Acids

The Peptide Bond

Role of Proteins in Living Systems

- chemical catalysis
- transport
- energy conversion and storage
- coordinated movement
- mechanical skeleton
- immune response
- molecular recognition
- passing information
- gene regulation
- growth and differentiation

X-ray Crystallography

NMR Structure Determination

NMR structure of the 64 amino acid
SH3 domain of the Src protein

Interactions Stabilizing the Native State

short range repulsion

Van der Waals interaction

electrostatic interaction

hydrogen bonding

hydrophobic interaction

disulfide bridge

Short Range Repulsion

Due to the exchange (Pauli) interaction, at short distances there is a strong repulsion between electrons.

The potencial energy of the repulsion increases quickly with decreasing distance ($\sim 1/r^{12}$).

Atoms can be considered hard spheres with a given radius (Van der Waals radius).
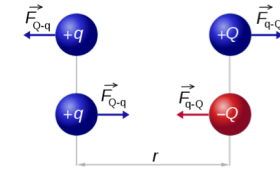
Van der Waals Interaction

Occurs between any two atoms due to the interaction of induced dipole moments.

Dependence on the distance of the interaction energy: ~ $1/r^6$

Electrostatic Interaction

Distance dependence of the interaction energy of the Coulomb force:

$$E = \frac{q \cdot Q}{4 \cdot \pi \cdot \varepsilon_0 \cdot \varepsilon_r \cdot r}$$

The relative dielectric constant inside the protein is approx. 4, and 80 in water.
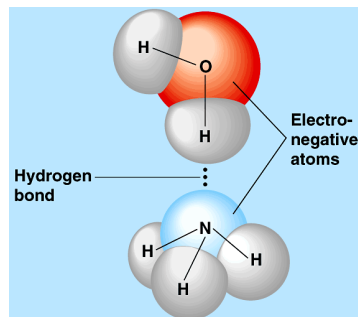
Salt bridges between ion pairs (Lys, Arg and Glu, Asp).

There is a large hydrate shell around charges in water.

Mobile ions can strongly shield charges.

Hydrogen Bonding

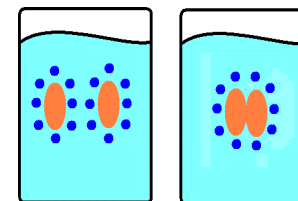Attraction force between a H atom of a more electronegative atom or group (hydrogen bond donor) and another atom bearing a lone pair of electrons (hydrogen bond acceptor).

Hydrophobic Interaction

observed tendency of nonpolar surfaces to adhere in an aqueous solution and exclude water molecule

entropic effect originating from the disruption of hydrogen bonds of liquid water by the nonpolar solute
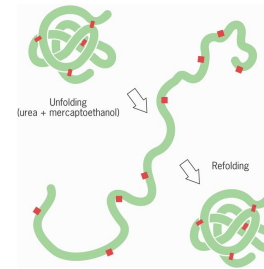
Disulfide Bridge

stabilizes the native structure
decreases the conformational entropy of the unfolded protein:

$$\Delta S = - 2.1 \text{ J/K} - 1.5 \cdot R \cdot \ln n$$

n is the number of AAs between the two bonded AAs.

Anfinsen's Dogma

Refolding of Ribonuclease A   Christian B. Anfinsen

The information of the 3D protein structure is encoded in the 1 D AA sequence.

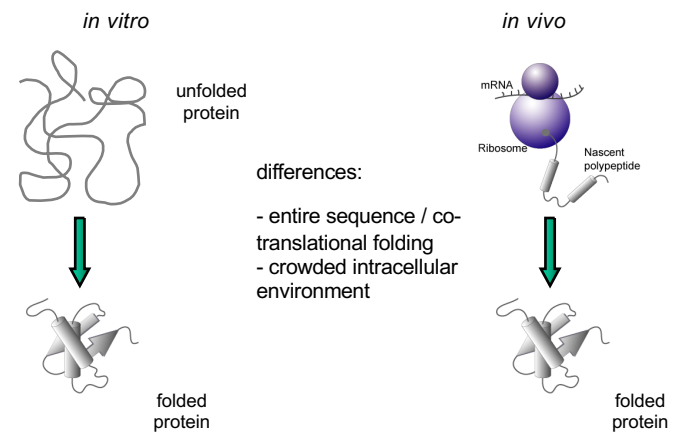## Importance of the Protein Folding Problem

One of the most important questions of molecular biophysics.

We sequence genomes, we build databases, but we can't predict protein structure and function based on the genetic information.

There are roughly two dozen conformational diseases:

Misfolded proteins and deposition of amyloid plaques was observed in various diseases (pl. Creutzfeld-Jakob disease, Alzheimer disease, Parkinson disease).

In vitro and *in vivo* Folding
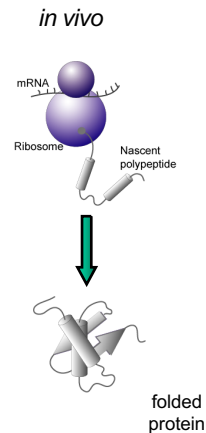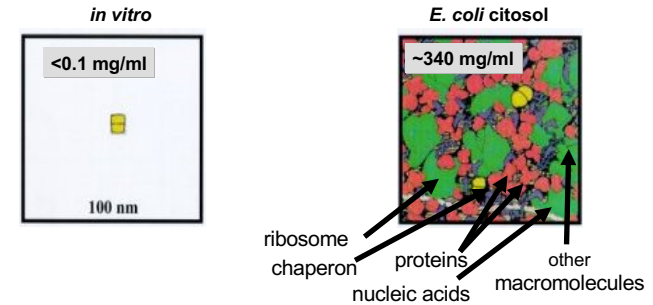

in vitro   *in vivo*

unfolded protein

differences:

- entire sequence / co-translational folding
- crowded intracellular environment

mRNA
Ribosome
Nascent polypeptide

folded protein   folded protein

Co-Translational Folding

in vivo

The N terminal of the nascent polypeptide chain starts to fold before completion of the translation.

20-30 AAs of the C terminal are protected within the ribosome.

mRNA

Ribosome

Nascent polypeptide

folded protein

Molecular Crowding

in vitro

<0.1 mg/ml

100 nm

E. coli citosol

~340 mg/ml

ribosome
chaperon
nucleic acids
proteins
other macromolecules

In vitro experiments
- lack of binding partner molecules
- lack of posttranslational modifications
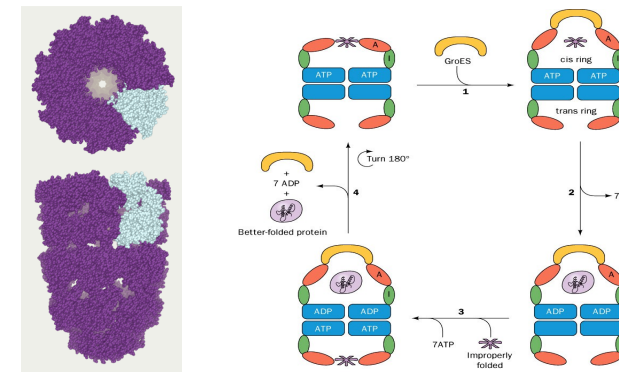- very different physico-chemical environment than in a cell

Effect of Molecular Crowding
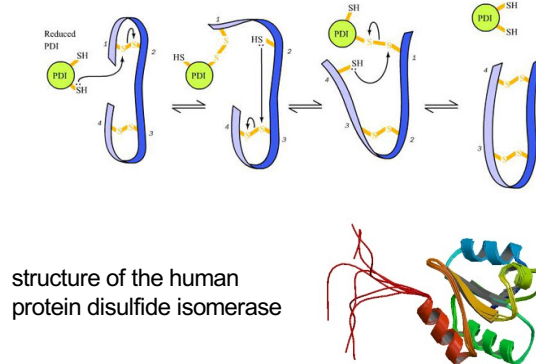
Molecular crowding:
a large fraction of the volume of the cytoplasm is filled by other molecules than water.

- dissociation constants decrease
- speed of protein-protein association increases
- association of fully or partially denatured proteins speeds up

GroEL/ES Chaperon Cycle

GroES

cis ring

trans ring

Turn 180°

7 ADP

Better-folded protein

7ATP

Improperly folded protein

Protein Disulfide Isomerase Function

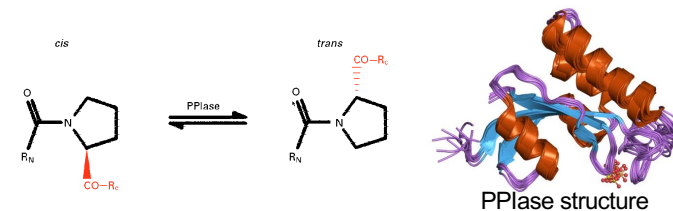structure of the human
protein disulfide isomerase

Proline Cis/Trans Isomerase

Due to the activation barrier between the cis and trans prolines,
the presence of cis prolines in the native structure:
- speeds up early folding steps
- slows down the final formation of the native structure.

PPIase (peptidyl-prolyl isomerase)

PPIase structure

Fate of the Protein in Eukaryotic Cells

cytosol .... protein synthesis and folding,
extracellular volume .... export of folded protein
mitochondrion .... limited protein synthesis
chloroplast .... limited protein synthesis
endoplasmic reticulum .... import of unfolded protein
peroxisome .... import of folded protein
nucleus .... import of folded protein
lysosome .... import of unfolded protein

Levinthal's Paradox - Calculation

Cyrus Levinthal

Consider a protein of 151 AAs. Assume all the 150 bonds
connecting them have only two possible conformations. Assume
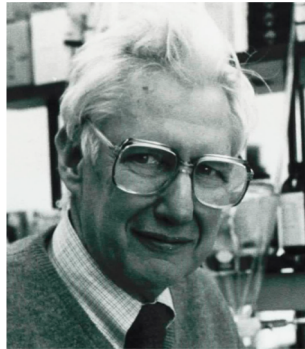that a reorientation of the bonds happens in 10^{-13}s.

A random search through the phase space would last:
$$2^{150} * 10^{-13}s = 4.6 \cdot 10^{24} \text{years}.$$

Age of Earth: $4.6 \cdot 10^9$ years
Age of the Universe: $13.7 \cdot 10^9$ years
Proteins typically fold on the ms to s timescale.

Levinthal's Paradox - Conclusion

The phase space of a protein is way too big to find the native structure by random search.

Cyrus Levinthal
1922 - 1990

Kinetic Pathways and Intermediate States

All proteins have a most stable conformation.

The protein can find this conformation by following a kinetic pathway and adopting specific intermediate states.

In vivo, trapping of the protein in intermediate states is prevented by protein disulfide isomerases, peptidyl prolyl isomerases and chaperones.
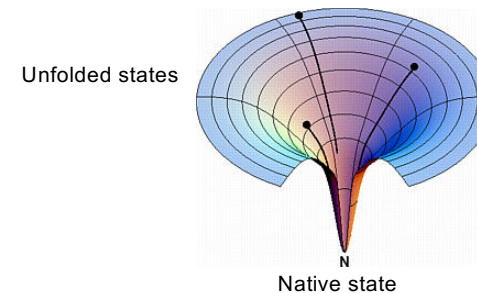
Energy Landscape Models

At constant pressure and temperature every thermodynamic system tends to minimize Free enthalpy (Gibbs free energy).
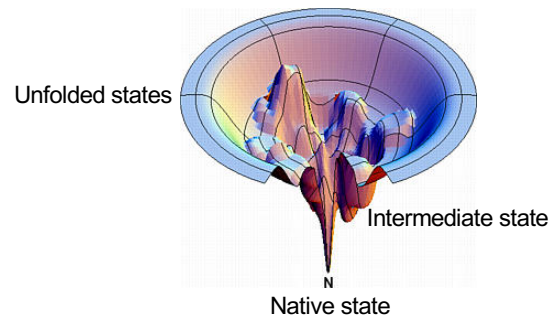
A free enthalpy (Gibbs free energy) value is associated to every conformation of the protein.

The protein does not search through the entire phasespace, but starts to "flow" towards lower free enthalpies.
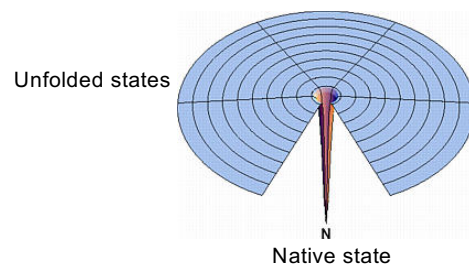
Smooth Funnel

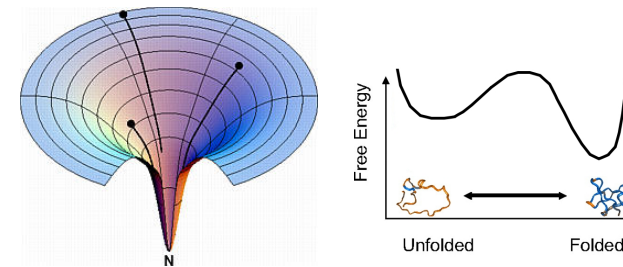Unfolded states

N

Native state

Rugged Funnel

Unfolded states

Intermediate state

N

Native state

Comparison of the Two Folding Models

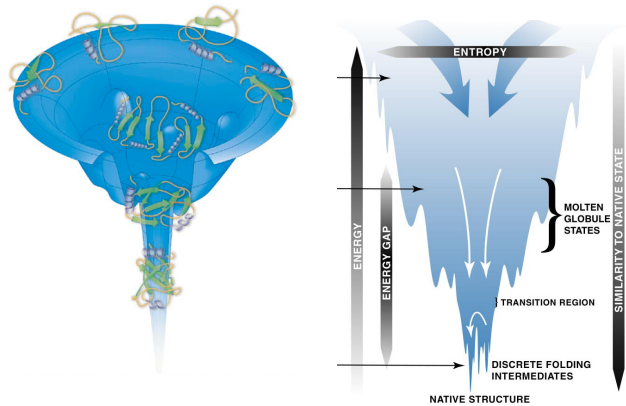| Pathways	Landscape
Given pathways	Energy landscape
Well distinguished intermediates	Multitude of intermediates
Consecutive steps	Parallel folding routes
Classical chemical kinetics applied to protein folding	Statistical physics developed to understand spin glasses

Energy Landscape View of Levinthal's Paradox

Unfolded states

N

Native state

Averaging Less Important Coordinates

Free Energy
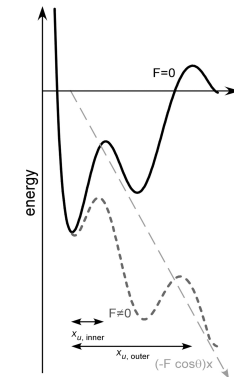
Unfolded   Folded

N

Formation of Ordered Structure by Folding

Effect of Destabilization on the Energy Landscape
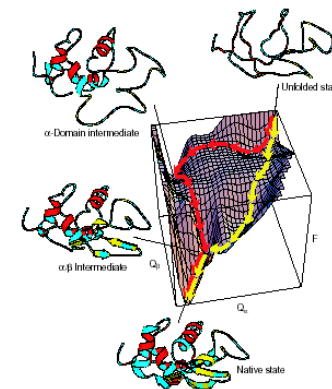
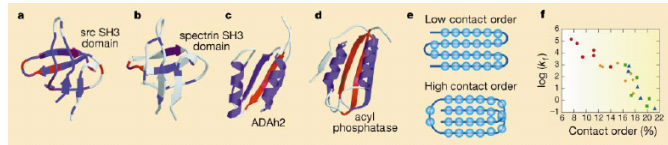Effect of mechanical force

Molten Globule State

Oleg Ptitsyn predicted its existence:

- compact, globular
- native-like secondary structure
- stabilized by non-specific hydrophobic interactions
- similar to the native structure
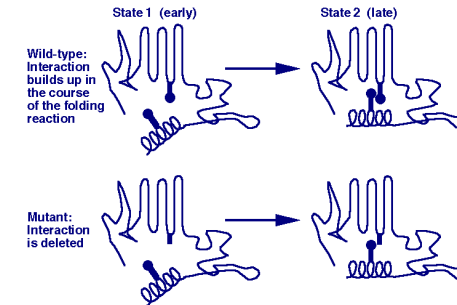- no rigid tertiary structure

Lysozyme Folding
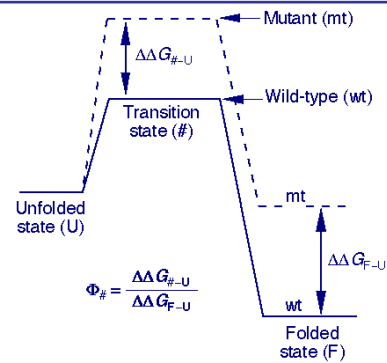
Effect of Native Structure on Folding

Proteins with smaller contact order fold faster.

Correlation between contact order and folding speed is observable through 6 orders of magnitude in time.

Study of the Transient State by Φ Value Analysis

Study of the Transient State by Φ Value Analysis

$$\Phi_{\#} = \frac{\Delta\Delta G_{\#-U}}{\Delta\Delta G_{F-U}}$$

Folding of the BBA5 Mini-protein