A polypeptide chain is extended and flexible in the unfolded, enatured state whereas it is globular and compact in the folded, native state.
100 a.a in general
If there are 10 states/a.a $10^{100}$ conformations

- Understand the basic principles of folding.
  How do proteins fold?
- Develop a conformational search method.
  How can we fold proteins using computers?
FACTS:

- Each sequence folds into a unique structure – native structure
- Proteins are functional only in their native state
- Sequence → structure mapping is not yet understood
- Folding is reversible – unfolding and re-folding is possible

PREDICTION

sequence ↔ structure & function

Why is it hard?
- many possible conformations for the protein
- many may have similar energies
- calculated energies are estimates
- hard to tell the correct structure
Degrees of Freedom in Proteins

**Bond length**

1 2

**Bond angle**

**Dihedral angle**

1 2 3 4

DESIGN

Why is it hard?
- many possible sequences
- don’t know what structure each sequence adopts
- calculated energies are estimates
- hard to tell the correct structure
Protein folding problem:
"Predicting 3-dimensional structure from sequence"

- A unique folded structure (native conformation, native fold) is assumed by a given sequence, although infinitely many conformations can be accessed.
- Which? (Protein folding problem)
- How, why? (Folding kinetics)

Basic postulate:
Thermodynamic equilibrium → Global energy minimum

Sequence determines structure.
How?

- Secondary structure preferences (satisfy H bonds)
- Hydrophobic/polar patterning
- Steric complementarity
- Electrostatics

Interactions are both LOCAL and NONLOCAL in sequence
• The molten globule state is an important intermediate in the folding pathway when a polypeptide chain converts from an unfolded to a folded state.

• The molten globule has most of the secondary structure of the native state but it is less compact and the proper packing interactions in the interior of the protein have not been formed.
**CLASSICAL VIEW**

- Transition state (TS) is a molecular structure.
- Rate limiting step corresponds to a TS structure.

The unfolded state is an ensemble of a large number of conformationally different molecules, U1...Un, which undergo rapid interconversions.
The molten globule is an ensemble of structurally related molecules, M1…Mm, which are rapidly interconverting and which slowly change to a single unique conformation, the folded state F.

During the folding process the protein proceeds from a high energy unfolded state to a low energy native state.

The conversion from the molten globule state to the folded state is slow and passes through a high energy transition state, T.

What is TS structure in protein folding?
• (a) Some proteins such as barnase fold through one major pathway whereas others fold through multiple pathways.

• (b) The folding of the enzyme lysozyme proceeds through at least two different pathways.

• Schematic diagram of the structure of the enzyme barnase which is folded into a five stranded antiparallel β sheet (blue) and two α helices (red).
Figure 6.5 (a) Some proteins such as barnase fold through one major pathway whereas others fold through multiple pathways. (b) The folding of the enzyme lysozyme proceeds through at least two different pathways.

Figure 6.6 Schematic diagram of the structure of the enzyme lysozyme which folds into two domains. One domain is essentially α-helical whereas the second domain comprises a three-stranded antiparallel β sheet and two α helices. There are three disulfide bonds (green), two in the α-helical domain and one in the second domain.
Classical view in Folding

- Framework model (Ptittsyn)
- Hydrophobic collapse model (Dill)
- Nucleation – condensation mechanism (Fersht)
• **Models for folding of protein.**

  a) *Framework model*. Protein folding is thought to start with the formation of elements of secondary structure independently of tertiary structure, or at least before tertiary structure is locked in place. These elements then assemble into the tightly packed native tertiary structure either by diffusion and collision or by propagation of structure in a stepwise manner.

  b) *Hydrophobic collapse model for folding*. The initial event of the reaction is thought to be a relatively uniform collapse of the protein molecule, mainly driven by the hydrophobic effect. Stable secondary structure starts to grow only in the collapsed state.

  c) *Nucleation-condensation mechanism*. Early formation of a diffuse protein-folding nucleus catalyses further folding. The nucleus primarily consists of a few adjacent residues which have some correct secondary structure interactions, but is stable only in the presence of further approximately correct tertiary structure interactions. Measurements of the nucleation site by Phi-value analysis agree with predictions using the funnel model, which focuses on the rapid decrease of the conformational dispersity in the course of the reaction. Larger proteins may have several nuclei.

  All three models may be extended to proteins which have intermediates and multiple transition states on their pathways. The final steps of protein folding usually involve the interlocking of the sidechains.
What is an energy landscape?

According to the principles of thermodynamics, if a system has \( n \) degrees of freedom \( \phi = [\phi_1, \phi_2, \ldots, \phi_n] \), the stable state of the system can be found by determining the set of values \( \phi^* = [\phi^*_1, \phi^*_2, \ldots, \phi^*_n] \) that gives the minimum value of the free energy function

\[
F(\phi) = F(\phi_1, \phi_2, \ldots, \phi_n),
\]

when explored over all possible values of \( \phi \).

Such functions are called energy landscapes. Energy landscapes, per se, are neither new, nor controversial, nor limited to proteins. “Energy landscape” is nothing more than a name for this function. For protein folding, \( \phi \) may be the backbone and sidechain bond angles, for example.

What is an energy landscape?

- Funnel landscape
  - Multiplicity of routes
  - Heterogeneity of routes

Fig. 4. Energy landscapes are free energy \( F(\phi_1, \phi_2, \ldots) \), as a function of the degree of freedom, \( \phi_1, \phi_2, \ldots, \) such as backbone and side-chain bond angles.
NEW VIEW: LANDSCAPES

![Diagram of energy landscape](image)

Keskin et al., Biochemistry 2002
Chaperone Action

Can you fold a protein by unfolding it?

Trapped conformation → Chaperone conformation → Opened conformation → Folded
From a computational point of view:

Protein folding problem

Sequence ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ ✡ 🍒------- créer un modèle de formatage de la traduction en langage naturel de documents. Ce formatage garantit la cohérence et la précision de la traduction en garantissant la conformité avec les normes linguistiques et grammaticales données. Il élimine les fautes de frappe et les erreurs de traduction en garantissant que le texte est traduit de manière appropriée. Il garantit également la compréhension et l'interprétation correctes du document, permettant une lecture naturelle et intuitive. Enfin, il permet la comparaison de documents similaires, en garantissant que les informations sont présentées de manière cohérente et uniforme. En résumé, ce formatage garantit une traduction précise, cohérente et intuitivement compréhensible, ce qui facilite la lecture naturelle et la compréhension des informations contenues dans le document.
• Proteins try to minimize their Gibbs Free Energy

• $G = H - TS$
  $\Delta G = \Delta H - \Delta S$ (at fixed $T$ and $P$)

• For folding
  $\Delta G = \Delta G_N - \Delta G_U \{\text{system} = \text{protein} + \text{water}\}$

Energetics of folding

• From entropic point of view, folding is unfavorable (entropy: max disorder!!)

• Enthalpy (interaction energy) makes folding favorable.

• Water outside the protein is extremely important (solvation) →
  • Entropy (water) increases
Hydrophobicity

- Hydrophobicity is not only the most important physico-chemical characteristic of amino acids, it is also the most poorly defined term.
- The words hydrophobic and hydrophilic literally mean 'afraid of water' and 'fond of water'.
- It is obvious that hydrophobic residues prefer to be in a non-aqueous environment, for example, a lipid bilayer.
- Hydrophilic residues love water, and therefore like to sit at the outside of water soluble proteins.

Hydrophobicity has been measured in many different ways:

- From octanol - water solubility differences
- From sidechain polarities as calculated by quantum chemical techniques
- From the relative distributions at the surface and in the core of proteins
- From the atomic constitution of the sidechains
Hydrophobicty Scale for amino-acids

Database: AAindex1
Entry: HOPT810101

H NCG791031
D Hydrophilicity value (Hopp-Woods, 1981)
R LIT: 0707585 PDB: 6167991
A Hopp, T.P., and Woods, R.R.
* Prediction of protein antigenic determinants from amino acid sequences
* Proc. natl. acad. sci. USA 78, 3824-3828 (1981)
C LGYH760101 0.985 NXCQ730101 0.886 PVRD50102 0.886
FXQH510194 0.884 NXSQ860101 0.882 KVNH50103 0.881
NADX50103 0.881 GRVR410102 0.874 KITP990103 0.862
VIRH801019 0.879 PVTH500101 0.878 FVQK50102 0.874
VHGR790191 0.874 KOHR800101 0.874 KITP990104 0.843
GUOR801053 0.873 NLIS870101 0.870 KITP990105 0.825
FOQR801091 0.899 KITP990102 0.894 KITP990101 0.893
GUTR850101 0.862 HTRE850101 -0.690 NAGB810103 -0.805
WAGE801181 -0.812 NCDN80105 -0.816 JACR90101 -0.816
NMID89102 0.820 MSES660101 -0.822 NODS50102 -0.825
NNEN801092 -0.826 NADA800101 -0.829 NOOB40103 -0.829
NIDE80103 -0.830 NADA80108 -0.831 CDSG80101 -0.839
EXQH80103 -0.846 BYSV880101 -0.848 WYFW960101 -0.855
KIDR820102 -0.859 BYSV880102 -0.864 BLAP880101 -0.877
EXQH80101 -0.905 PNRK830101 -0.909 HODN880103 -0.958

<table>
<thead>
<tr>
<th>A/L</th>
<th>K/R</th>
<th>R/K</th>
<th>E/D</th>
<th>C/L</th>
<th>G/S</th>
<th>E/T</th>
<th>G/H</th>
<th>B/F</th>
<th>l/V</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.3</td>
<td>3.6</td>
<td>9.2</td>
<td>3.0</td>
<td>-1.0</td>
<td>0.2</td>
<td>3.8</td>
<td>0.0</td>
<td>-0.3</td>
<td>-1.8</td>
</tr>
</tbody>
</table>

-1.8 3.6 -1.3 -2.5 0.0 0.3 -0.4 -3.4 -2.3 -1.5


Hydrophobic effect:

- "Hydrophobic residues attract each other via hydrophobic forces": All atoms attract each other via the Van der Waals interaction.
- Hydrophobicity does not mean "love for hydrophobic atoms", but "fear of water".
- Water doesn't like hydrophobic atoms.
- A water molecule that swims around in water at room or body temperature has on average about 3.5 hydrogen bonds, and it has nearly 6 degrees of freedom.
- A water sitting against the side of the benzene ring in phenylalanine can form only about two hydrogen bonds, and its freedom is also severely restricted. This water molecule compares unfavourably with one in bulk water.
- In a folded protein, many hydrophobic residues pack against each other in the core of the protein. And these hydrophobic residues can no longer make the life of water molecules miserable.
The Hydrophobic Entropy is Due to Ordering of Waters Around the Solute

- The basis of $\Delta s^0 << 0$ that opposes the mixing of oil and water is due to ordering of water by the presence of solute.
- Computer simulations indicate that a solute that is introduced into water at 25 °C orders the first shell of neighboring water molecules. The first shell water molecules orient to avoid wasting hydrogen bonds.
A water molecule has six hydrogen bonding configurations when it is surrounded by four neighboring waters.

When a nonpolar solute replaces the water at the vertex, the central water has only three fully hydrogen-bonded configurations with neighboring water molecules.
Water molecules form tetrahedral hydrogen bonding.
Water molecules form cages especially when there is a nonpolar solute.

Clathrate water cage around a long chain hydrocarbon

Why is the hydrophobic entropy so large?

- Suppose that each first-shell water molecule could form water-water hydrogen bonds in half of its possible orientations around the solute. Experiments show that small solutes are surrounded by about ~15-30 first-shell water molecules. If each 15 water molecules in the shell were restricted independently then the hydrophobic entropy will be:

\[ S = k \ln(W) = k \ln(1/2)^{15} = -86 \text{ J/mol K} \]
So summary for entropy of mixing non-polar with water …

- The entropy of transfer into water is large not because the solute molecule is ordered but because the many surrounding water molecules are ordered.

Why is enthalpy of mixing favorable?

- The relatively favorable enthalpy of inserting a nonpolar solute into water is observed.

- It seems that inserting a nonpolar solute into water creates cage like ordering in which the first shell water molecules have better hydrogen bonding than tetrahedral (pure) water.
What is the Basis for the Large Hydrophobic Heat Capacity

• For transferring nonpolar solutes from their pure liquid to water, there is a change in heat capacity $\Delta C_p >> 0$.
• This implies that a nonpolar solute surrounded by water has greater ability to absorb enthalpy from the surroundings than the corresponding pure components alone because $\Delta C_p = \partial h/\partial T$.
• First-shell water molecules are ordered (low entropy) and form good hydrogen bonds with other water molecules (low enthalpy). Upon heating, enthalpy and entropy increases. Warming up cool water increases the enthalpy by bending, breaking or loosening water-water hydrogen bonds in the first shell around the solute.

In summary hydrophobicity:

• In water nonpolar solute molecules associate with other. According to simplest argument, pairing up offers the molecules the advantage of reducing the net surface area of contact with water, reducing the number of first-shell water molecules and thus reducing the unfavorable entropy.
Energy Minimization

Step 1: First Parabola
Step 2: Second Parabola

\[ \Phi \text{ Value Analysis} \]

\[ \Delta G = -RT \ln K \]
\[ \Delta G^{\dagger} = -RT \ln k_f \]

\[ \Phi = \frac{\Delta G^{\dagger}}{\Delta G} \]

\( \Phi = 1 \)
At mutation site:
TS has Native-like structure

\( \Phi = 0 \)
TS has Denatured-like structure
Figure 6.21 Schematic diagram of the conformational changes of calmodulin upon peptide binding. (a) In the free form the calmodulin molecule is dumbbell-shaped comprising two domains (red and green), each having two E1 hands with bound calcium ions. (b) In the form with bound peptides (blue) the α-helix linker has been broken, the two ends of the molecule are close together and they form a compact globular complex. The internal structure of each domain is essentially unchanged. The bound peptide binds as an α-helix.

Figure 6.8 Schematic diagram of the enzyme DsbA which catalyzes disulfide bond formation and rearrangement. The enzyme is folded into two domains, one domain comprising five α-helices (green) and a second domain which has a structure similar to the disulfide-containing redox protein thioredoxin (violet). The N-terminal extension (blue) is not present in thioredoxin. (Adapted from J.L. Martin et al., Nature 365: 464–468, 1993.)
Figure 6.9 (a) Peptide units can adopt two different conformations, trans and cis. In the trans-form the C=O and the N-H groups point in opposite directions whereas in the cis-form they point in the same direction. For most peptides the trans-form is about 1000 times more stable than the cis-form. (b) When the second residue in a peptide is proline the trans-form is only about four times more stable than the cis-form. Cis-proline peptides are found in many proteins.

Figure 6.10 Schematic diagram of the structure of the protein cyclophilin, a prolyl peptide isomerase that catalyzes the conversion between cis- and trans-proline peptides. The protein is folded into an eight-stranded antiparallel β-barrel (blue) with two α-helices (red) on the outside. The active site, which has been located by the binding of a proline containing tetrapeptide (green), is on the outside of the β-barrel. (Adapted from J. Jolles et al., Nature 353: 276-279, 1991.)
Figure 6.11 Schematic diagram of the chaperonin GroEL molecule as a cylinder with 14 subunits arranged in two rings of 7 subunits each. The space occupied by one subunit is red and the hole inside the cylinder is blue.

Figure 6.12 (a) Schematic diagram of one subunit of GroEL. The polypeptide chain is folded into three domains. The equatorial domain (green) is the largest domain, comprising 10 α-helices, and is built up from both the N-terminal and the C-terminal regions. The apical domain (blue), which is a β-sandwich flanked by α-helices, is formed by the middle region of the polypeptide chain. The two linker regions between the equatorial and the apical domains form a small intermediate domain (purple) comprising three α-helices. (b) Schematic diagram illustrating the domain arrangement of four subunits in the GroEL molecule: two in each of the seven-membered rings. The equatorial domains form the middle part of the molecule and interact with each other both within each ring and between the rings. The apical domains are at the top and the bottom of the cylinder and form an opening to the interior of the molecule. The small intermediate domains form the thinnest part of the cylinder wall in the middle of each ring. (a) Adapted from F. Blügg et al., Nature 371, 578-586, 1994.)
Figure 6.13 Models of the GroEL molecule in two different functional states based on three-dimensional reconstruction from electron microscopy pictures. A large conformational change of GroEL occurs when GroEL and ATP are bound. The GroES molecule binds to one of the GroEL rings and closes off the central cavity. The GroEL ring becomes larger and the cavity inside that part of the cylinder becomes wider. (Adapted from S. Chen et al., Nature 371: 264–264, 1994.)

Figure 6.14 (a) Schematic diagram of the GroES molecule. Seven subunits are linked together in a ring with the same symmetry arrangement as in one of the rings of the GroEL molecule. Two loop regions extend from the core of the subunits (green), one of which (yellow) is flexible and located on the outside of the ring. This loop is hydrophobic and interacts with the GroEL molecule in the GroEL–GroES complex. The other loop (red) covers the central cavity of GroEL when GroES is bound. (b) Schematic diagram of the structure of one subunit of GroES. The core of the subunit structure is a β-barrel (green) comprising two antiparallel β-sheets packed against each other. The mobile loop (yellow) is flexible and the roof β-hairpin loop covers the central part of the seven-membered ring of the GroES molecule. (Adapted from J.E. Hunt et al., Nature 379: 37–42, 1996.)
Figure 6.15 Possible functional cycle of the GnrE-L-GnrES molecule. (a) An unfolded protein molecule (yellow) binds to one end of the GnrE-ADP complex (red) with bound GnrES (green) at the other end. (b and c) GnrES is released from the trans-position and released together with ATP at the cis-position (light red) of GnrEL. (d) ATP hydrolysis occurs as the protein is folding or unfolding inside the central cavity. (e) ATP binding and hydrolysis in the trans-position is required for release of GnrE and the protein molecule. (f) A new unfolded protein molecule can now bind to GnrEL. (Adapted from M. Mayhew et al., Nature 379: 420–426, 1996.)

Figure 6.16 Cell Cycle
Figure 6.17: An schematic diagram of the structure of the cyclin-dependent kinase CDK2. The activation loop is important for catalysis. The N-terminus of the kinase is highly conserved and forms a flexible domain. The X-ray crystal structure of cyclin A and CDK2 shows that the activation loop contains a key serine residue (Ser 10) which is phosphorylated to activate the enzyme. The activation loop is connected to the catalytic domain by a linker region. The active site is located in the hinge region between the catalytic domain and the regulatory domain. The diagram shows the active and inactive states of CDK2 in the presence and absence of cyclin A. (Adapted from J. K. Keyte et al., Nature 376: 515-522, 1995.)

Figure 6.18: The PSTAIR helix undergoes a major conformational change when CDK2 binds to cyclin A. In the inactive form of CDK2 (green), the active site residues Glu 51 is far from the active site. Upon binding of cyclin A to CDK2, the PSTAIR helix (blue) rotates 90° and changes its position so that Glu 51 becomes positioned into the active site. (Adapted from P.D. Jeffery et al., Nature 376: 315-320, 1995.)
Figure 6.29 Schematic diagram of the conformational changes of CDK2 upon cyclin binding. (a) In the active form, the P-N9AII helix (blue) is rotated such that Glu 31 points away from the ATP-binding site (purple) and the T-loop (yellow) blocks the substrate binding site and prevents proteins from binding to CDK2. (b) In the active cyclin-CDK2 complex, the P-N9AII helix is reoriented so that Glu 31 points into the active site and forms a salt bridge to another residue involved in catalysis, Lys 38. The T-loop has changed its conformation and one of its residues, Thr 160, forms hydrogen bonds to the ATP in the active site. The substrate-binding site is now open, proteins can bind, and the cyclin-CDK2 complex can phosphorylate Ser/Thr residues and thereby activate the bound proteins.

Figure 6.20 Space-filling diagram illustrating the structural changes of CDK2 upon cyclin binding. (a) The active site is in a cleft between the N-terminal domain (blue) and the C-terminal domain (purple). In the inactive form this site is blocked by the T-loop. (b) In the active cyclin-bound form of CDK2 the T-loop has changed its structure, the active site is open and available, and Thr 160 is available for phosphorylation.
Figure 6.21 Schematic diagram of the conformational changes of calmodulin upon peptide binding. (a) In the free form the calmodulin molecule is dumbbell-shaped comprising two domains (red and green), each having two EF hands with bound calcium (yellow). (b) In the form with bound peptides (blue) the α-helix linker has been broken, the two ends of the molecule are close together and they form a compact globular complex. The internal structure of each domain is essentially unchanged. The bound peptide binds as an α-helix.

Figure 6.22 Schematic diagram of the structure of ovalbumin which illustrates the serpin fold. The structure is built up of a compact body of three antiparallel β sheets, A, B, and C, surrounded by α helices. The polypeptide chain is colored in sections from the N-terminal to facilitate following the chain tracing in the order green, blue, yellow, red and pink. The red region corresponds to the active site loop in the serpins which in ovalbumin is protruding like a handle out of the main body of the structure. (Adapted from R.W. Carrell et al., Structure 2: 257–270, 1994.)
Figure 6.23 Active site loop regions in three forms of the serpins

Figure 6.24 The function of the enzyme phosphofructokinase. (a) Fructose-1,6-bisphosphate is a key enzyme in the glycolytic pathway: the breakdown of glucose to pyruvate. One of the end products in this pathway, phosphofructokinase, is an allosteric feedback inhibitor to this enzyme and ATP is an activator. (b) Phosphofructokinase catalyzes the phosphorylation by ATP of fructose-1,6-bisphosphate to give fructose-2,6-bisphosphate. (c) Fructose-2,6-bisphosphate, which has a structure similar to phosphofructokinase, is also an inhibitor of the enzyme.
Figure 6.23 Schematic diagram of the structure of one dimer of phosphofructokinase. Each polypeptide chain is folded into two domains (blue and red, and green and brown), each of which has an α/β structure. Helices are labeled A to M and β strands 1 to 11 from the amino terminus of one polypeptide chain, and respectively from 'A' to 'M' and '1' to '11' for the second polypeptide chain. The binding sites of substrate and effector molecules are schematically marked in gray. The active site of one subunit is linked to the active site of the other subunit of the dimer through the 6-8 loop between helix F and strand 6. (Adapted from T. Schiemer and P.R. Evans, Nature 343: 140-145, 1990.)

Figure 6.26 The quaternary structure of phosphofructokinase. (a) The four subunits are packed in an arrangement of four dimers (A-B and C-D) with a central interaction between the dimers. The packing of the subunits is such that the dimers are extended and tightly packed against each other and the packing contacts are different in the R and T states. The orientation of the dimers with respect to each other in the T (and mixture of the C-D dimer) and B (green contour) diffuse by a solution of 7M. (b) The dimers are close together in the T state and there are direct hydrogen bonds between two β-strands, and between the A-I dimer (blue) and one from the C-D dimer (green). Hydrogen bonds are shown in yellow. (c) The dimers are further apart in the R state and there is an gap between the two dimers in the mixture of the A-I dimer from the two dimers, which is filled by water molecules (red). These water molecules form bridges between the dimers by making hydrogen bonds to the C-β1 and N-β1 groups of the two β-strands.
Figure 6.27 Conformational changes in the active site of phosphofructokinase. (a) In the active R state the phosphate group of the substrate fructose-6-phosphate, F6P, (red) forms a salt bridge to an arginine residue, Arg 162, of a small β-helix (orange). This salt bridge contributes substantially to a poised binding of the substrate to the enzyme. (b) In the inactive T state the helix has been partially unwound and changed its orientation so that Arg 162 points away from the substrate binding site. Instead a negatively charged glutamate residue, Glu 161, points towards the phosphate binding site of the substrate molecule. Repulsive forces between the negative charges of Glu 161 and the phosphate of F6P prevent binding and result in a thousandfold lower affinity for F6P when the enzyme is in the T state compared with the R state.