Thermodynamics of Protein Folding

- Proteins consist of flexible chains of amino acids. However, for a protein to function correctly, it must have a well-defined conformation.
- Christian Anfinsen won the Nobel Prize in chemistry in 1972 for his work showing that the sequence of amino acids solely determines the way the chain folds itself.

- Though the amino acid sequence of a protein contains the necessary information to create the active conformation of the protein from a newly synthesized chain, the prediction of the conformation from the sequence, the so-called **protein folding problem**, is extraordinarily difficult and is still the focus of much research.
- Solving the problem of how a protein finds its functional conformation will also help us understand why some proteins fold improperly under certain circumstances.
- Misfolded proteins are thought to be involved in a number of diseases, such as cystic fibrosis, Alzheimer's disease, and "mad cow" disease (variant Creutzfeldt-Jakob disease, v-CJD).

NON-POLAR

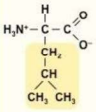
Glycine
(Gly / G)

Alanine
(Ala / A)

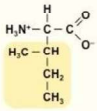
Valine
(Val / V)

Cysteine
(Cys / C)

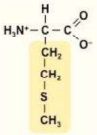
Proline
(Pro / P)



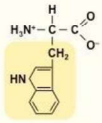
Leucine
(Leu / L)



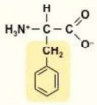
Isoleucine
(Ile / I)



Methionine
(Met / M)

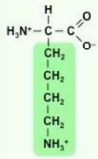


Tryptophan
(Trp / W)

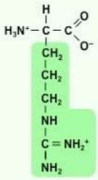


Phenylalanine
(Phe / F)

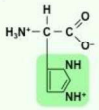
+ CHARGE



Lysine
(Lys / K)



Arginine
(Arg / R)

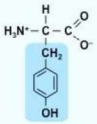


Histidine
(His / H)

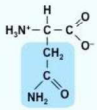
POLAR

Serine
(Ser / S)

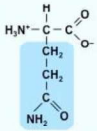
Threonine
(Thr / T)



Tyrosine
(Tyr / Y)

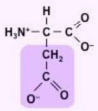


Asparagine
(Asn / N)

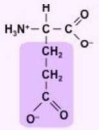


Glutamine
(Gln / Q)

− CHARGE



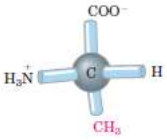
Aspartic Acid
(Asp / D)

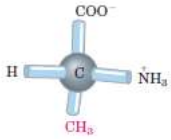


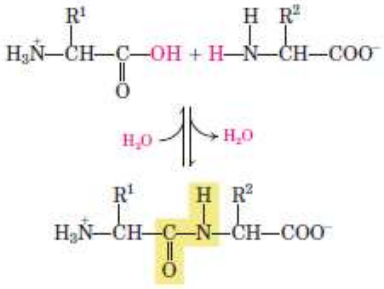
Glutamic Acid
(Glu / E)

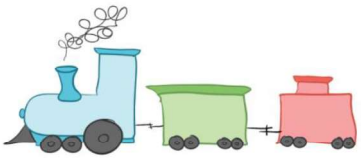
L-Alanine

D-Alanine

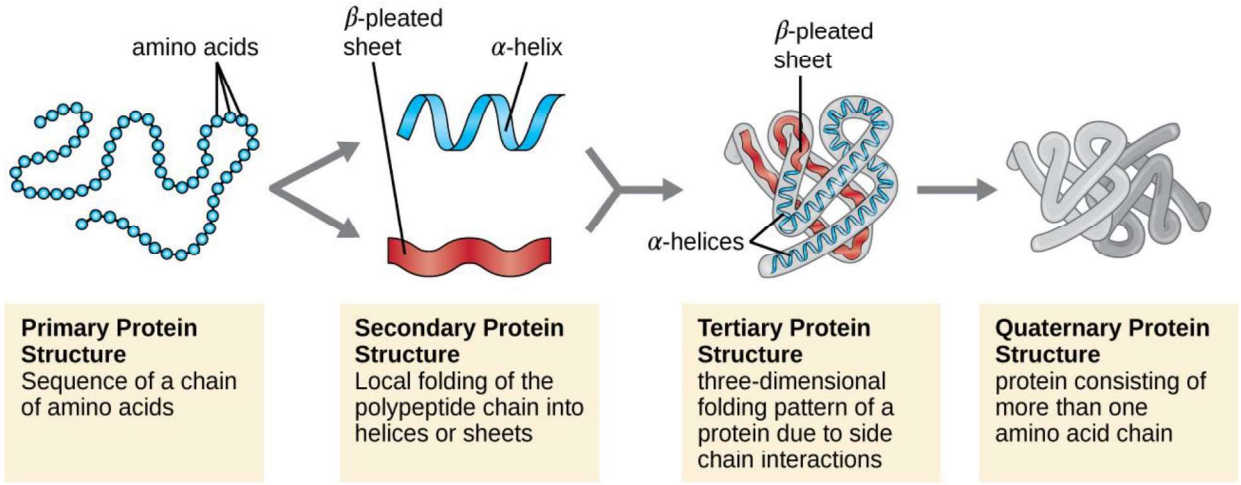


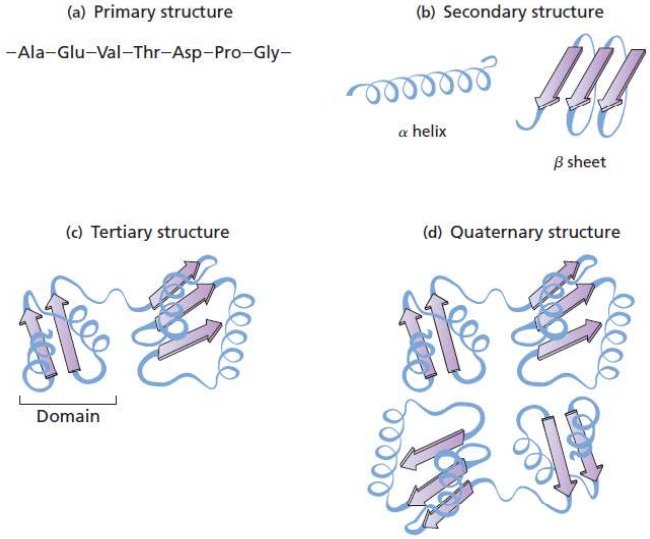






Amino acids link together like the cars of a train





- Weak interactions ensure that the polymer chain preferentially adopts one particular **conformation**, determined by the linear sequence of aminoacids in the chain.
- These noncovalent interactions include **hydrogen bonds** (between polar groups), **ionic interactions** (between charged groups), **hydrophobic interactions** (among nonpolar groups in aqueous solution), and **van der Waals interactions**.

- Protein folding and stabilization depend on several noncovalent forces including the hydrophobic effect, hydrogen bonding, van der Waals interactions, and charge–charge interactions.
- Although noncovalent interactions are weak individually, collectively they account for the stability of the native conformations of proteins.
- The weakness of each noncovalent interaction gives proteins the resilience and flexibility to undergo small conformational changes.

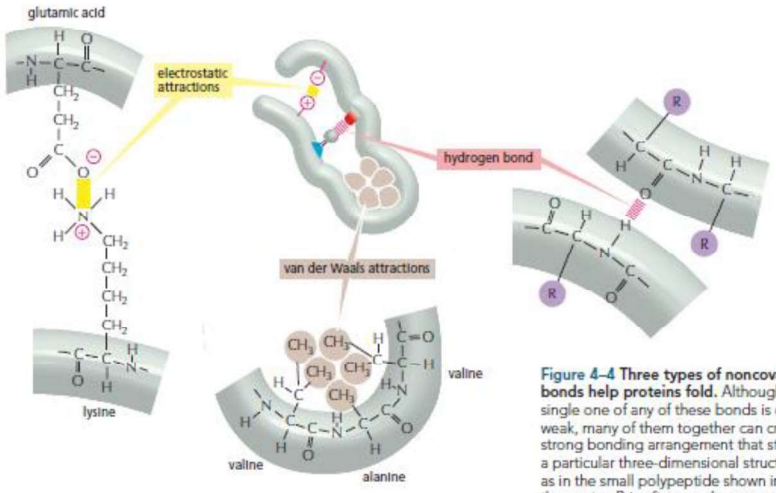
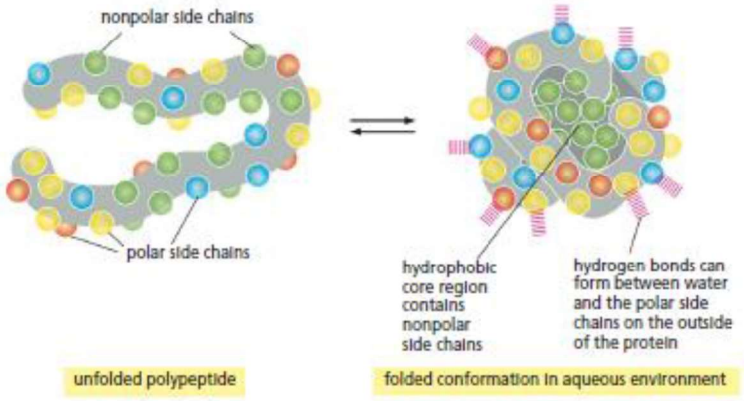


Figure 4–4 Three types of noncovalent bonds help proteins fold. Although a single one of any of these bonds is quite weak, many of them together can create a strong bonding arrangement that stabilizes a particular three-dimensional structure, as in the small polypeptide shown in the center. R is often used as a general designation for an amino acid side chain. Protein folding is also aided by hydrophobic forces, as shown in Figure 4–5.

Long polypeptide chains are very flexible, as many of the peptide bonds that link the carbon atoms in the polypeptide backbone allow free rota-tion of th



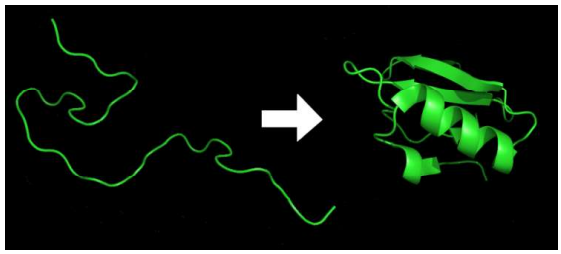
Levinthal's paradox

- To appreciate the complexity of the mechanism of protein folding, consider a small protein consisting of a single chain of 100 amino acids in a well-defined sequence.
- Statistical arguments lead to the conclusion that the polymer can exist in about 10^{49} distinct conformations, with the correct conformation corresponding to a minimum in the energy of interaction between different parts of the chain and the energy of interaction between the chain and surrounding solvent molecules.
- In the absence of a mechanism that streamlines the search for the interactions in a properly folded chain, the correct conformation can be attained only by sampling every one of the possibilities.
- If we allow each conformation to be sampled for 10^{-20} s, a duration far shorter than that observed for the completion of even the fastest of chemical reactions, it could take more than 10^{21} years, which is much longer than the age of the Universe, for the proper fold to be found. However, it is known that proteins can fold into functional conformations in less than 1 s.
- The preceding arguments form the basis for *Levinthal's paradox* and lead to a view of protein folding as a complex problem in thermodynamics and chemical kinetics: how does a protein minimize the energies of all possible molecular interactions with itself and its environment in such a relatively short period of time?

- A polypeptide chain adopts a conformation corresponding to a minimum Gibbs energy, which depends on the **conformational energy**, the energy of interaction between different parts of the chain, and the energy of interaction between the chain and surrounding solvent molecules.
- In the aqueous environment of biological cells, the outer surface of a protein molecule is covered by a mobile sheath of water molecules, and its interior contains pockets of water molecules.
- These water molecules play an important role in determining the conformation that the chain adopts through hydrophobic interactions and hydrogen bonding to amino acids in the chain.
- The conformations existing under a given set of conditions are usually the ones that are thermodynamically the most stable.
- The folding process is thus energetically favorable, as it releases heat and increases the disorder of the universe.

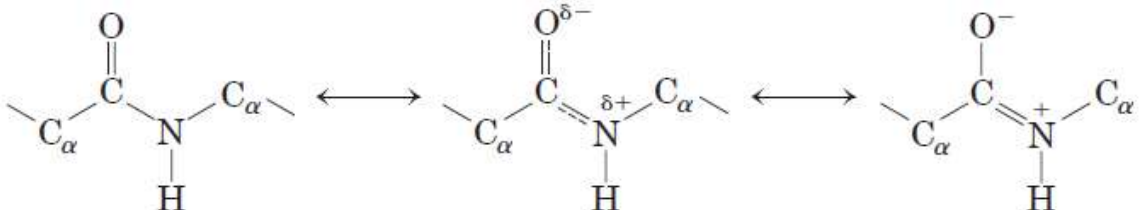
Constraints on possible conformations

- Each protein can adopt huge number of potential conformations since every amino acid residue has a number of possible conformations and since there are many residues in a protein.
- However, under physiological conditions most proteins fold into a single stable shape known as its **native conformation**.
- The biological function of a protein depends on its native three-dimensional conformation.
- A number of factors constrain rotation around the covalent bonds in a polypeptide chain in its native conformation.
  - These include the allowed rotation about the certain covalent bonds, steric hinderence, presence of hydrogen bonds and other weak interactions between amino acid residues.

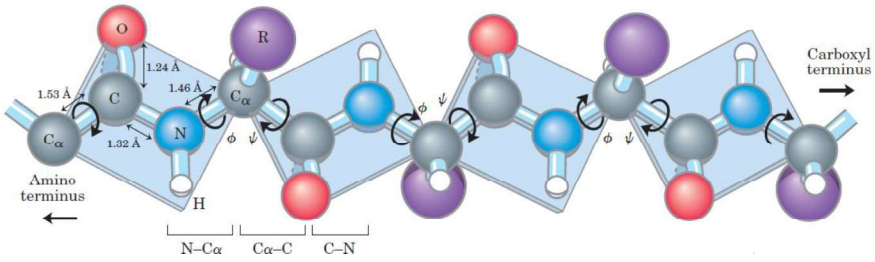


Constraints on possible conformations

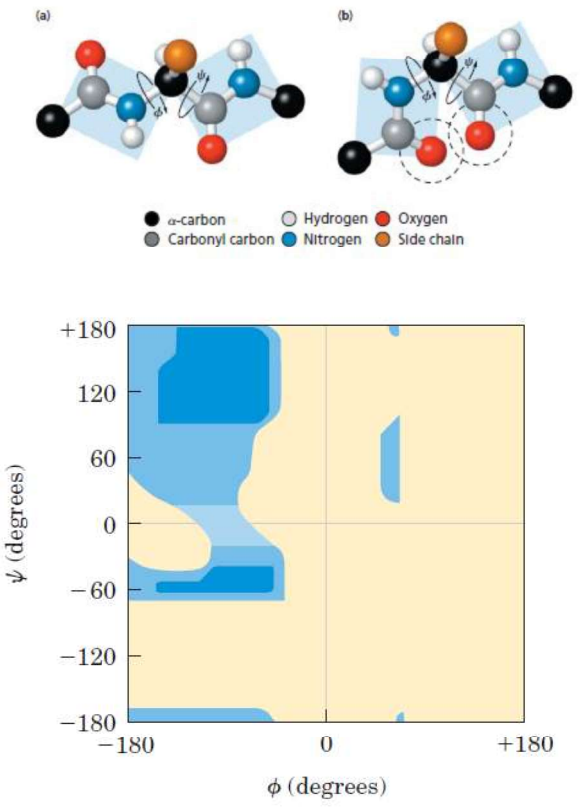
- The carbons of adjacent amino acid residues are separated by three covalent bonds, arranged as C_α —C—N—C_α.
- Peptide C—N bonds are unable to rotate freely because of their partial double-bond character.
- Rotation is permitted about the N—C_α and the C_α—C bonds.
- The backbone of a polypeptide chain can thus be pictured as a series of rigid planes with consecutive planes sharing a common point of rotation at C.
- The rigid peptide bonds limit the range of conformations that can be assumed by a polypeptide chain.



- By convention, the bond angles resulting from rotations at C_α are labeled *as* φ (phi) for the N—C_α bond and ψ (psi) for the C_α—C bond.
- Again by convention, both φ and ψ are defined as 180° when the polypeptide is in its fully extended conformation and all peptide bonds are in the same plane.
- In principle, and can have any value between -180° and 180°, but many values are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains.
- Allowed values for and are graphically revealed when φ is ψ plotted versus in a **Ramachandran plot.**

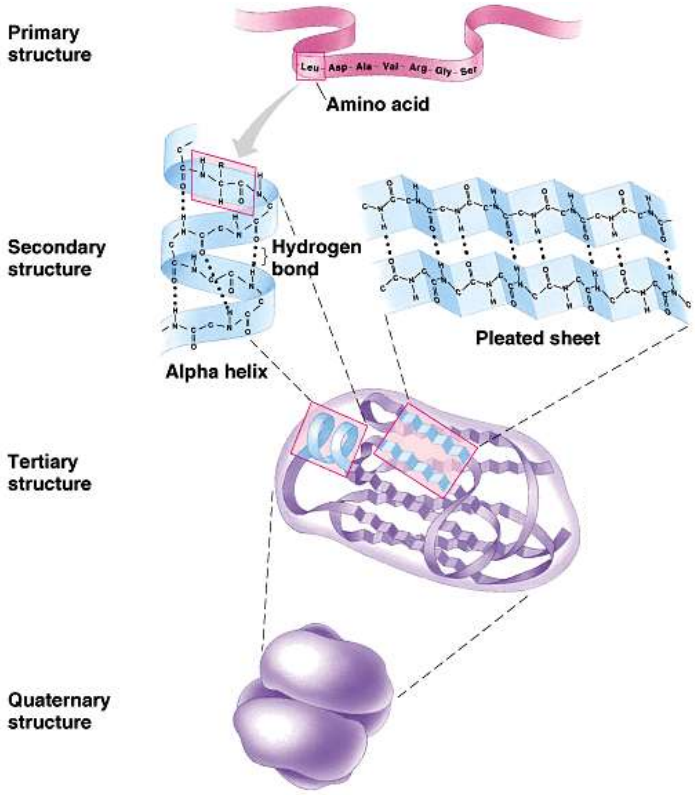


- Sterically permited values of φ and ψ are plotted in Ramachandran plot.
- Conformations deemed possible are those that involve little or no steric interference, based on calculations using known van der Waals radii and bond angles.
- The areas shaded dark blue reflect conformations that involve no steric overlap and thus are fully allowed; medium blue indicates conformations allowed at the extreme limits for unfavorable atomic contacts; the lightest blue area reflects conformations that are permissible if a little flexibility is allowed in the bond angles.
- The asymmetry of the plot results from the L stereochemistry of the amino acid residues. The plots for other L-amino acid residues with unbranched side chains are nearly identical.

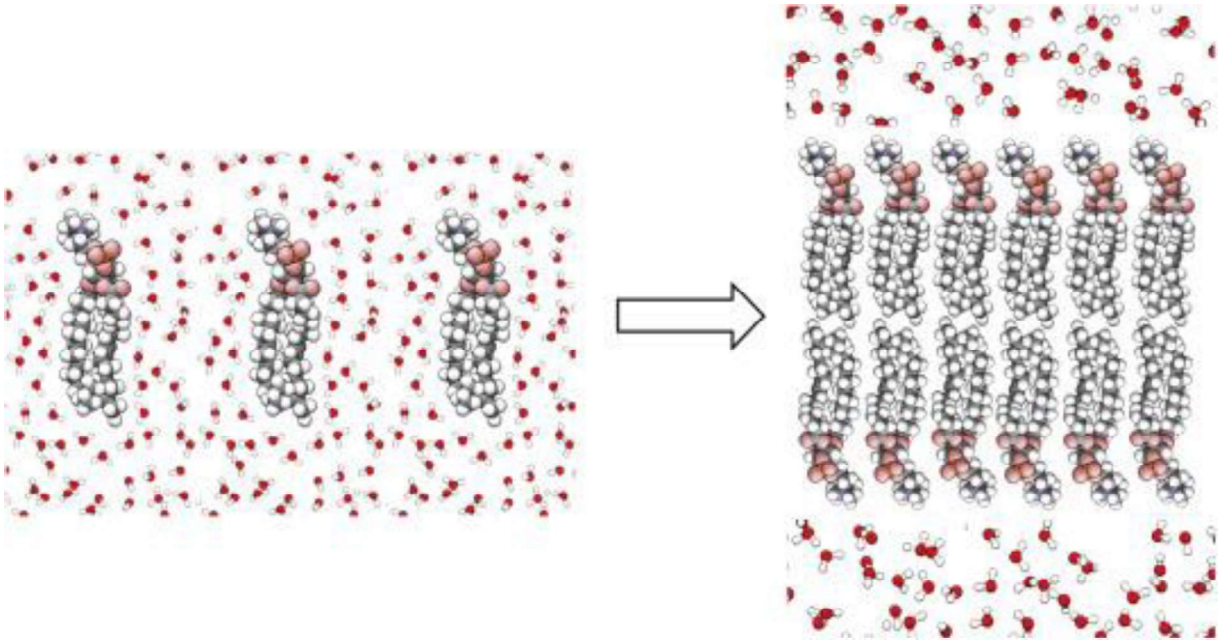


Ramachandran plot for L-Ala residues.

Protein Folding



Hydrophobic Effect

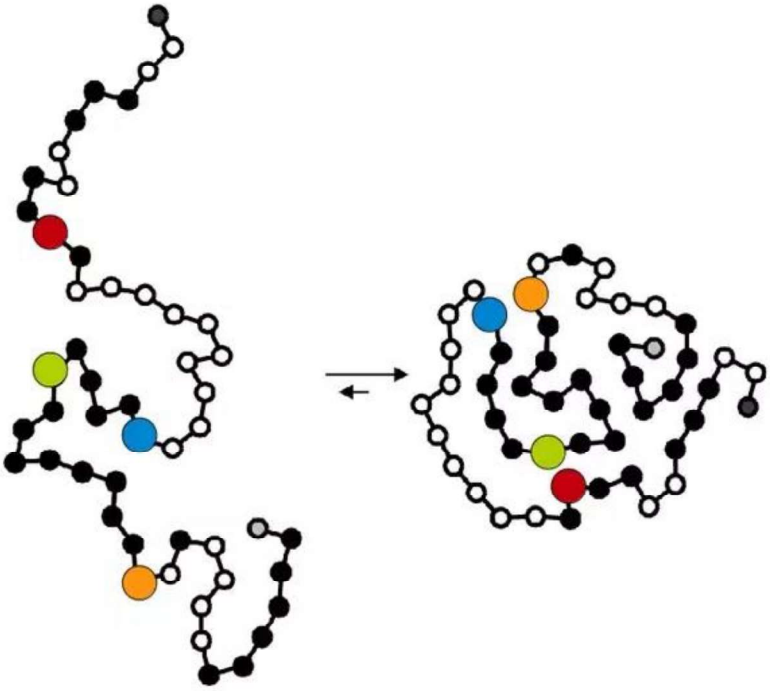


Thermodynamics of folding

- At constant temperature,

$$\Delta G^\circ_{\mathrm{d}} = \Delta H^\circ_{\mathrm{d}} - T\Delta S^\circ_{\mathrm{d}},$$

- ΔG°d is the difference in Gibbs free energy between the unfolded state (denatured state) and the folded (native) state of the protein.



ΔH? ↓

ΔS? ↓

ΔG? ↓

$\Delta G = \Delta H - T\Delta S$

- Main driving forces for folding:
  - ✓Hydrophobic burying
  - ✓Formation of H bonds

- ΔG°d alone tells us nothing about the relative magnitudes of ΔH°d or ΔS°d ; an infinite number of combinations of these thermodynamic functions would be consistent with a given value of ΔG°d .
- Of course, from many of these combinations of ΔH°d and ΔS°d, only one combination will actually describe the system under study.
- Temperature, pH and presence of a denaturant (and its concentration) change either the ΔH°d or ΔS°d or both component affecting ΔG°d.

Protein Stability

- Through the folding process an equilibrium exists between the folded (native) state and the unfolded (denatured) state.
- The transition may affected by the temperature, pH or chemical denaturant concentration.
- However, all the information required for a protein molecule to fold into its native form will be present in the amino acid sequence.

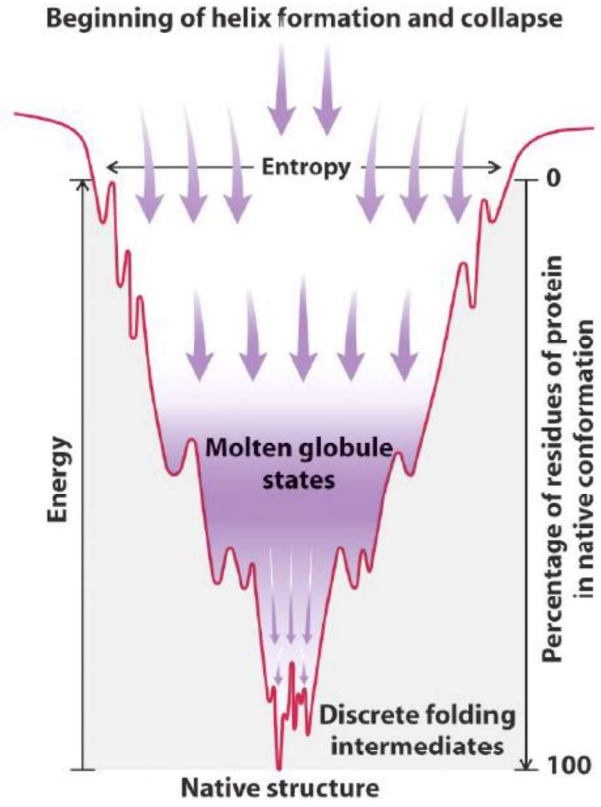
$$F \Leftrightarrow U.$$

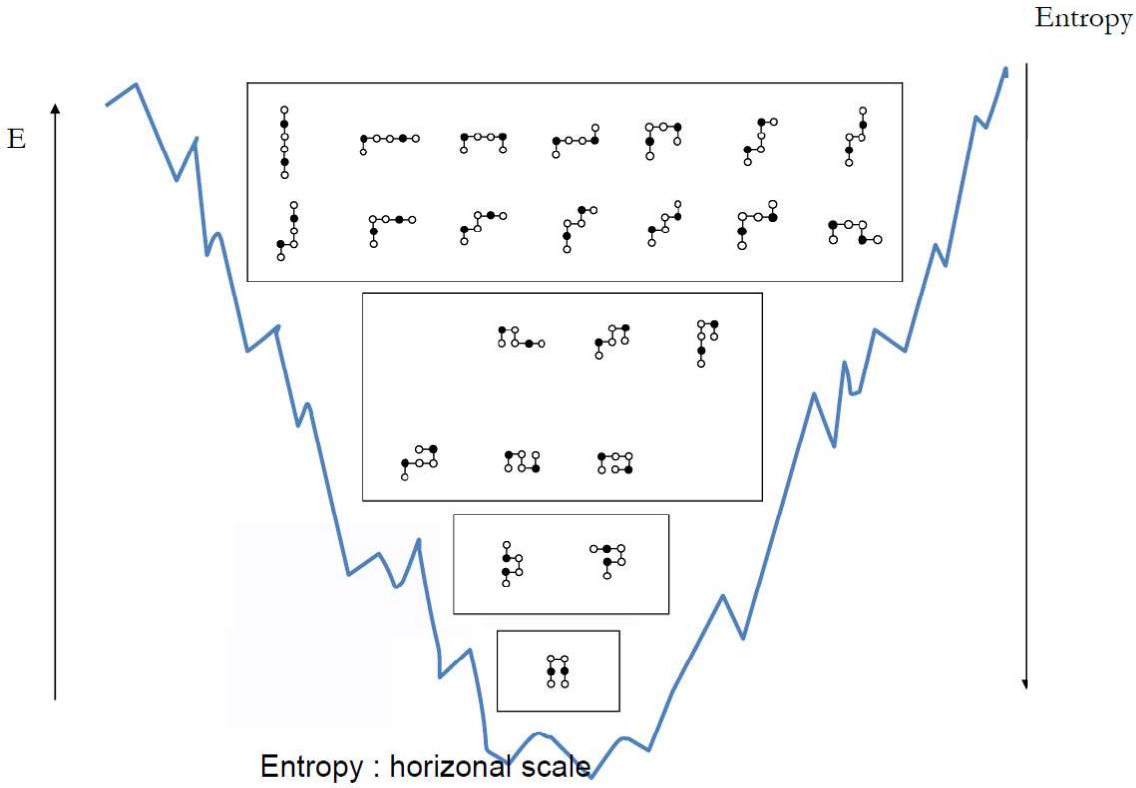
$$K_{eq} = [U]/[F].$$

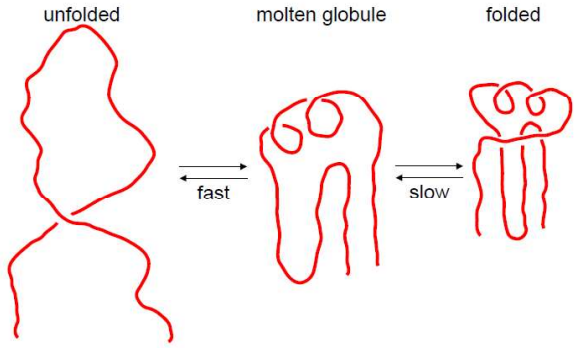
- The free energy difference between the folded state of a protein and its unfolded state is independent of the path!
- Regardless of the process by which a protein folds – in the cell or in a test tube – the free energy difference between folded and unfolded forms is the same (given the same temperature, ion concentrations, pH, etc.).
- But is a catalyst needed to get the reaction to proceed on a biologically relevant time scale?

Folding Event

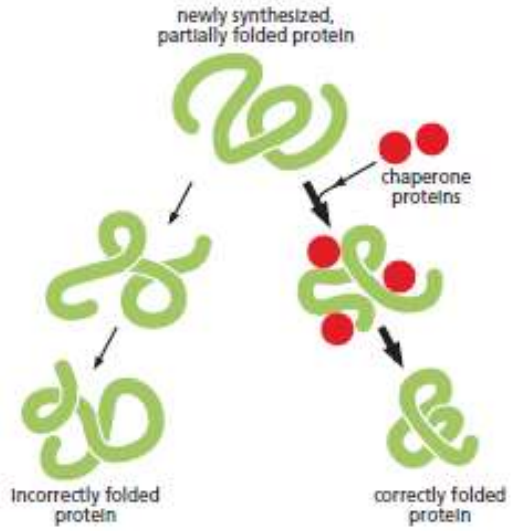
- As the newly synthesized polypeptide emerges from the ribosome, it folds into its characteristic three-dimensional shape.
- Folded proteins occupy a low-energy well that makes the native structure much more stable than alternative conformations.
- Folding is extremely rapid—in most cases the native conformation is reached in less than a second.
- As a protein folds the first few interactions trigger subsequent interactions. This is an example of cooperative effects in protein where the formation of one part of a structure leads to the formation of the remaining parts of the structure.
- As the protein begins to fold, it adopts lower and lower energies.
- In its final, stable, conformation, the native protein is much less sensitive to degradation than an extended, unfolded polypeptide chain. Thus, native proteins can have very long half-lives of many generations.
- The protein folding funnel is the model that best describes the folding of the protein.

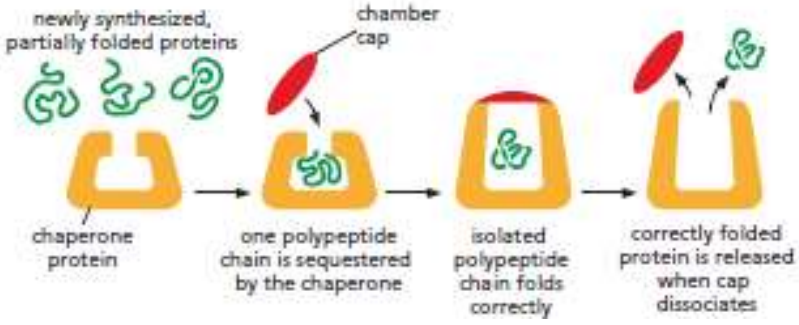



Entropy
E
Entropy : horizonal scale


unfolded
molten globule
folded
fast
slow

- Thus, folding event is rapid and spontaneous and dictated by the aminoacid sequence of the protein.
- However, chaperons generally assist this folding event.
- Some of these chaperones bind to partly folded chains and help them to fold along the most energetically favorable pathway.
- Others form “isolation chambers” in which single polypeptide chains can fold without the risk of forming aggregates in the crowded conditions of the cytoplasm.
- Chaperones make the folding process more efficient and reliable.





- Proteins can denature with the effects of temperature, pH or denaturant.

$$\Delta G = \Delta H - T\Delta S$$

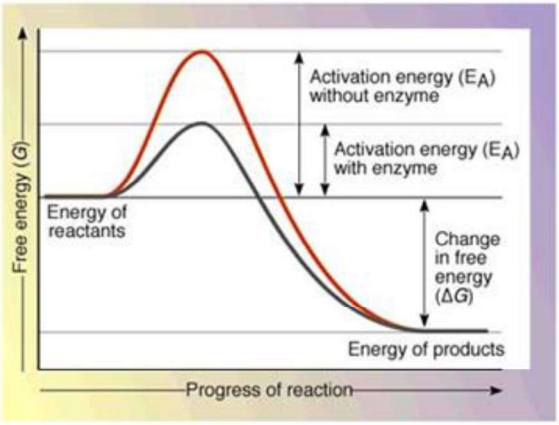
- At high temperatures, for instance; it is harder for ΔH to compensate for the large positive value coming from – TΔS.
- At the melting temperature, also called the heat-denaturation temperature, the fraction of molecules in the folded state equals that in the unfolded state; the free energy difference between them, ΔG, is 0.

- Computational techniques provide important insights into molecular interactions and can lead to reasonable predictions of the functional conformation of a protein.
- For example, in a **molecular mechanics** simulation, mathematical expressions from classical physics are used to determine the structure corresponding to the minimum in the energy of molecular interactions within the chain at the absolute zero of temperature.
- Such calculations are usually followed by **molecular dynamics** simulations, in which the molecule is set in motion by heating it to a specified temperature.
- The possible trajectories of all atoms under the influence of intermolecular interactions are then calculated by consideration of Newton's equations of motion.
- These trajectories correspond to the conformations that the molecule can sample at the temperature of the simulation.
- Those calculations are very difficult and time-consuming, theoretical studies inform experimental studies and vice versa.
- For example, the available data indicate that, in a number of proteins, a significant portion of the folding process occurs in less than 1 ms (10^{-3} s).
- Among the fastest events is the formation of helical and sheet-like structures from a fully unfolded chain. Slower events include the formation of contacts between helical segments in a large protein.

ENZYMES

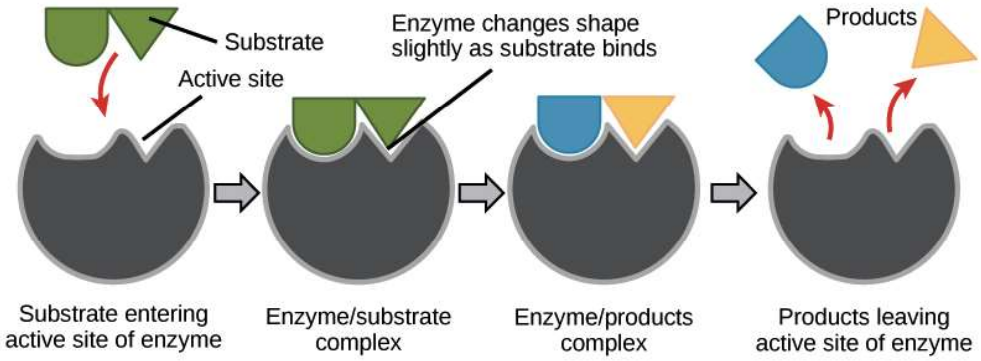
- Enzymes are extraordinarily efficient, selective, biological catalysts (biocatalysts).
- Every living cell has hundreds of different enzymes catalyzing the reactions essential for life—even the simplest living organisms contain hundreds of different enzymes.
- In multicellular organisms, the complement of enzymes differentiates one cell type from another but most of the enzymes are common to all cells.
- These enzymes catalyze the reactions of the central metabolic pathways necessary for the maintenance of life.

- In the absence of the enzymes, metabolic reactions will not proceed at significant rates under physiological conditions.
- The primary role of enzymes is to enhance the rates of these reactions to make life possible.
- Enzyme-catalyzed reactions are 10^3 to 10^{20} times faster than the corresponding uncatalyzed reactions.
- A catalyst is defined as a substance that speeds up the attainment of equilibrium. It may be temporarily changed during the reaction but it is unchanged in the overall process since it recycles to participate in multiple reactions.
- Reactants bind to a catalyst and products dissociate from it.
- It lowers the amount of energy needed in order for the reaction to proceed.

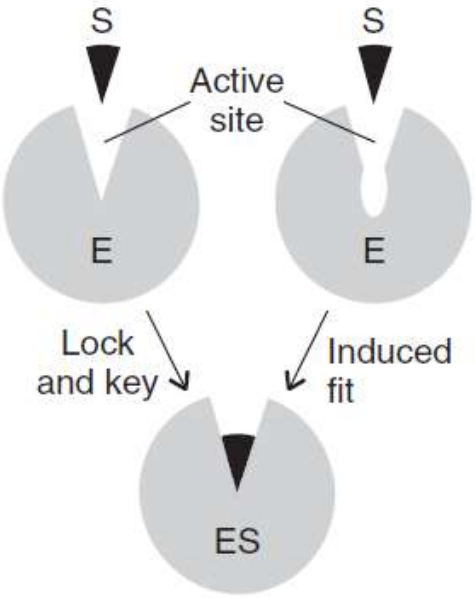


- Enzymes are special biological polymers that contain an **active site**, which is responsible for binding the **substrates**, the reactants, and processing them into products.
- As is true of any catalyst, the active site returns to its original state after the products are released. Many enzymes consist of proteins, some featuring organic or inorganic co-factors in their active sites.

$$E + S \rightarrow ES \rightarrow EP \rightarrow E+P$$

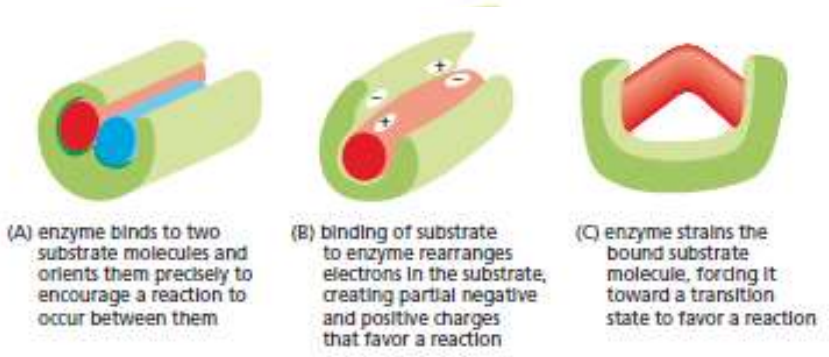


- There are two models that explain the binding of a substrate to the active site of an enzyme.
- In the **lock-and-key model,** the active site and substrate have complementary three-dimensional structures and dock perfectly without the need for major atomic rearrangements.
- Experimental evidence favors the **induced fit model**, in which binding of the substrate induces a conformational change in the active site. Only after the change does the substrate fit snugly in the active site.



Enzyme Mehanism of Action

- In reactions involving two or more substrates, the active site acts like a template or mold that brings the reactants together in the proper orientation for the reaction to occur.
- The active site of an enzyme contains precisely positioned chemical groups that speed up the reaction by altering the distribution of electrons in the substrates.
- Binding to the enzyme also changes the shape of the substrate, bending bonds so as to drive the bound molecule toward a particular transition state.



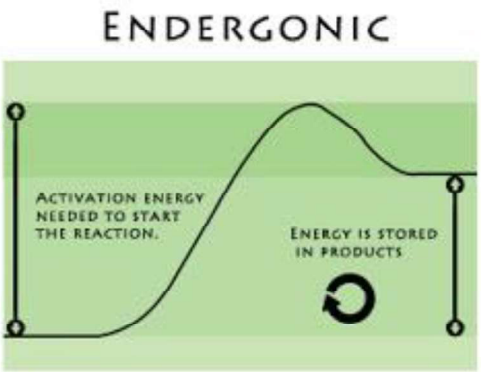
(A) enzyme binds to two substrate molecules and orients them precisely to encourage a reaction to occur between them

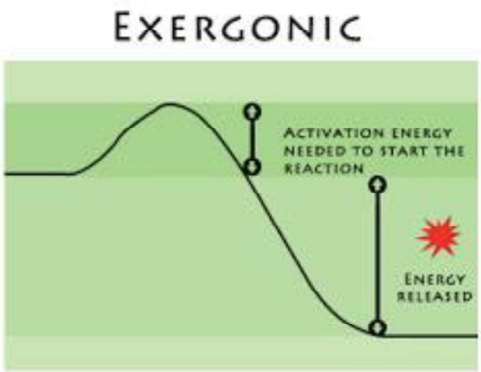
(B) binding of substrate to enzyme rearranges electrons in the substrate, creating partial negative and positive charges that favor a reaction

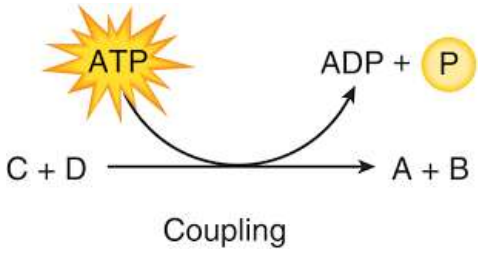
(C) enzyme strains the bound substrate molecule, forcing it toward a transition state to favor a reaction

Many enzymes participate intimately in the reaction by briefly forming a covalent bond between the substrate and an amino acid side chain in the active site. Subsequent steps in the reaction restore the side chain to its original state, so the enzyme remains unchanged after the reaction and can go on to catalyze many more reactions.

- Enzymes can do more than simply increase the rate of a single, highly specific reaction.
- Some can also combine, or couple, two reactions that would normally occur separately.
- This property allows the energy gained from one reaction to be used in a second reaction.
- Coupled reactions are a common feature of many enzymes—the hydrolysis of ATP, for example, is often coupled to less favorable metabolic reactions.



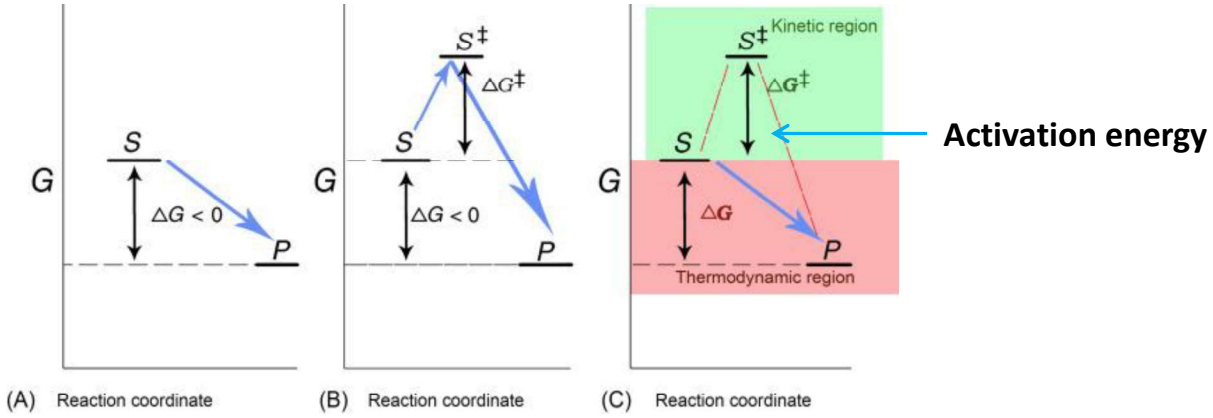




- The activity of enzymes is important for the proper functioning of cells since the organism must be able to catalyze chemical reactions efficiently and selectively.
- In the context of energy flow in living organisms, enzymes catalyze most reactions in metabolic pathways. Acting in organized sequences, they catalyze the hundreds of stepwise reactions that degrade nutrient molecules, conserve and transform chemical energy, and make biological macromolecules from simple precursors.
- Through the action of regulatory enzymes, metabolic pathways are highly coordinated to yield a harmony in the interplay of the many activities necessary to sustain life.
- Thus, enzymes not only make most reactions possible in an intracellular environment, enzymes allow for the control and stabilization of these reactions.

Thermodynamics of Enzymatic Reactions

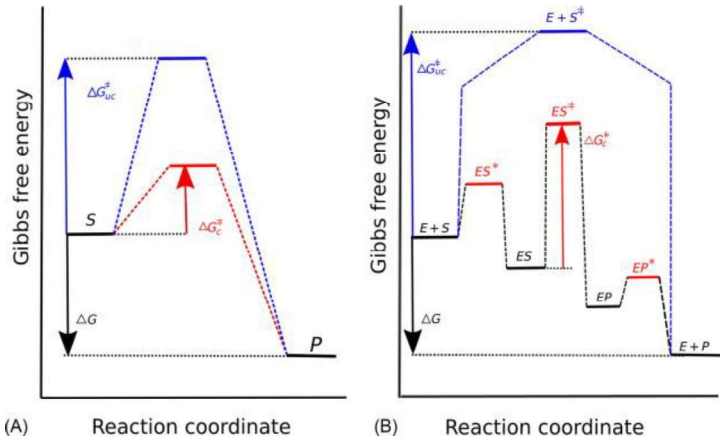
- Given a general chemical reaction of transformation of certain substrates in products (e.g. $S \rightleftharpoons P$), this reaction will proceed spontaneously if, and only if, the overall free energy content of products is lower than that of substrates ($\Delta_r G < 0$).
- This can be easily illustrated in an energy diagram by putting the overall energy level of *S* above than that of *P*.
- However, regardless of the actual value of $\Delta_r G$, kinetic constraints impose that any substrate or reactant must overcome a "kinetic barrier" to upgrade its basal energy state to an "activated" transition state ($S^{\ddagger}$) before the conversion to products can proceed.
- That means that the "path" to diminish the energy from G_S to G_P is not a straight line.
- The energy equivalent of this "kinetic barrier" is the variation of the free energy of substrates to achieve their "activated" transition state ($\Delta G^{\ddagger}$).



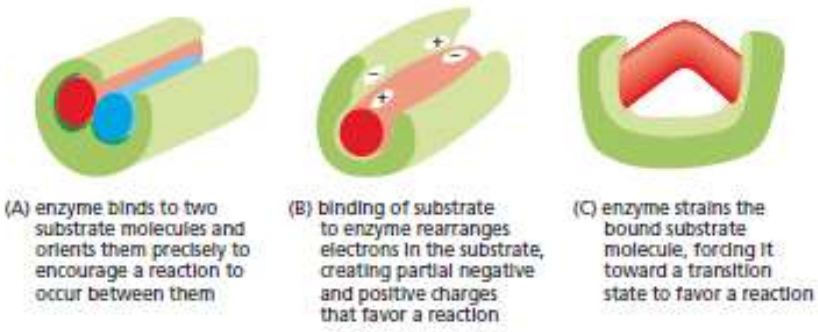
- Notice that $\Delta G^{\ddagger} > 0$ in all the cases.
- If the thermodynamic criterion of spontaneity for a reaction is the $\Delta G < 0$, how could the transition state $S^{\ddagger}$ be reached spontaneously if always $\Delta G^{\ddagger} > 0$?
- $\Delta G^{\ddagger}$, it contains both an enthalpic and an entropic component, according to the equation $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T \Delta S^{\ddagger}$.
- It is the enthalpic component, $\Delta H^{\ddagger}$, which is functionally related to the activation energy (E_a). This highlights the temperature dependence of reaction rates.
- However, this does not apply for enzyme-catalyzed reactions because when departing from their optimal temperature range, enzymes lose their native conformation, and become denatured, which implies that the empirically determined reaction rate will reflect these two different effects of temperature on the reaction rate constant, and on enzyme stability.

- How does the Gibbs energy profile for a chemical reaction change when this reaction is catalyzed by an enzyme *E*?
- As any other catalyzer, an enzyme affects the kinetics of the chemical conversion of reagents into products without affecting the thermodynamic behavior of the reaction.
- In fact, enzymes are extremely potent catalyzers, able to increase reaction rates by several orders of magnitude, typically 10^5–10^6 fold, but in some cases, up to 10^{17} fold.
- An enzyme physicochemically interacts with reagents and products of the reaction in order to accelerate its rate. A mixture of mechanistic descriptions and thermodynamic treatment is perhaps the best option to properly understand how enzymes can achieve such impressive rate enhancements.

$$E + S \rightleftharpoons ES^* \rightleftharpoons ES \rightleftharpoons ES^{\ddagger} \rightleftharpoons EP \rightleftharpoons EP^* \rightleftharpoons E + P$$



- Enzymes manage to diminish $\Delta G^{\ddagger}$ values (from $\Delta G_{uc}^{\ddagger}$ to $\Delta G_{c}^{\ddagger}$) by aligning substrates (proximity and orientation effects), by bending and weakening chemical bonds, by adding or withdrawing protons or electrons, and so forth.



(A) enzyme binds to two substrate molecules and orients them precisely to encourage a reaction to occur between them

(B) binding of substrate to enzyme rearranges electrons in the substrate, creating partial negative and positive charges that favor a reaction

(C) enzyme strains the bound substrate molecule, forcing it toward a transition state to favor a reaction

- For a multistep reaction, spontaneity requires not only a negative free energy change of the global process, but also that the condition of spontaneity occurs in each of the elementary reaction steps.

$$E + S \rightarrow ES \rightarrow EP \rightarrow E+P$$

- Therefore, in a spontaneous enzyme-catalyzed reaction, the Gibbs free energy of substrates, intermediates, and products must satisfy: $G_{E+S} > G_{ES} > G_{EP} > G_{E+P}$.

Happy New Year!