

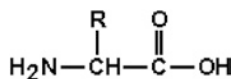
## Chapter 2

# Biosynthesis of Enzymes

### 2.1 Basic Enzyme Chemistry

#### 2.1.1 Amino Acids

An amino acid is a molecule that has the following formula:

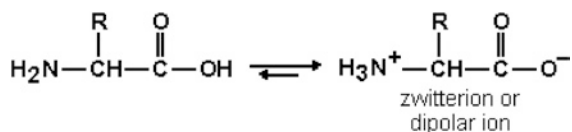


The central carbon atom covalently bonded by amino, carboxyl, and R group in the structure is called the *alpha carbon* ( $\text{C}_\alpha$ ). The side chain R group, vary in chemical composition, size, and interaction with water as reflected in their polarity. There are 20 standard amino acids used as common building blocks for peptides and proteins. The properties and structures of the side chains of these 20 naturally occurring amino acids are shown in Tables 2.1 and 2.2. Eight out of the 20 standard amino acids are called essential amino acids because they cannot be synthesized in our body and must be supplied from outside as food. They are isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine.

Amino acids are amphoteric compounds. They have both a carboxylic acid group and an amino group; so they function as either an acid or a base depending on the pH of the environment. These two functional groups undergo an intramolecular acid base reaction to form a zwitterion (dipolar ion):

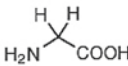
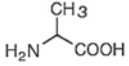
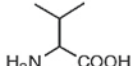
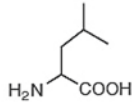
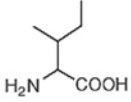
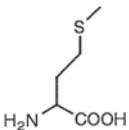
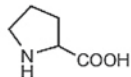
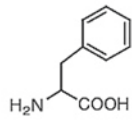
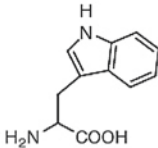
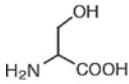
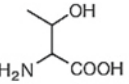
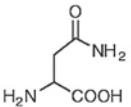
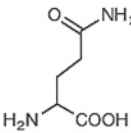
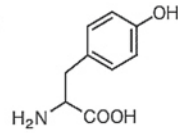
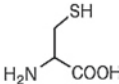
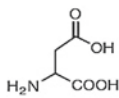
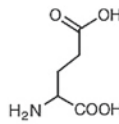
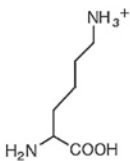
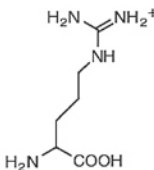
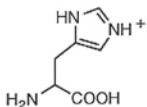
**Table 2.1** Name and properties of the 20 standard amino acids

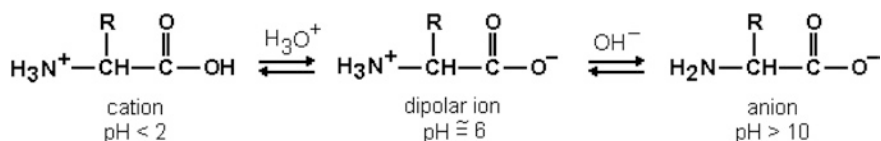
Side chain	Name	Abbreviation	Molecular weight	pK <sub>a</sub> of α-COOH	pK <sub>a</sub> of α-NH <sub>3</sub> <sup>+</sup>	pK <sub>a</sub> of R group	pI
Nonpolar	Glycine	Gly or G	75.07	2.3	9.8		6.1
	Alanine	Ala or A	89.10	2.3	9.9		6.1
	Valine	Val or V	117.15	2.3	9.7		6.0
	Leucine	Leu or L	131.18	2.3	9.7		6.0
	Isoleucine	Ile or I	131.18	2.3	9.7		6.0
	Methionine	Met or M	149.21	2.1	9.3		5.7
	Proline	Pro or P	115.13	2.0	10.6		6.3
	Phenylalanine	Phe or F	165.19	2.2	9.3		5.7
	Tryptophan	Trp or W	204.23	2.5	9.4		5.9
Polar	Serine	Ser or S	105.10	2.2	9.2		5.7
	Threonine	Thr or T	119.12	2.1	9.1		5.6
	Asparagine	Asn or N	132.13	2.1	8.7		5.4
	Glutamine	Gln or Q	146.15	2.2	9.1		5.7
	Tyrosine	Tyr or Y	181.19	2.2	9.2	10.5	5.7
	Cystine	Cys or C	121.16	1.9	10.7	8.4	5.3
Acidic	Aspartic Acid	Asp or D	133.11	2.0	9.9	3.9	3.0
	Glutamic Acid	Glu or E	147.13	2.1	9.5	4.1	3.1
Basic	Lysine	Lys or K	146.19	2.2	9.1	10.5	9.8
	Arginine	Arg or R	174.20	1.8	9.0	12.5	10.8
	Histidine	His or H	155.16	1.8	9.3	6.0	7.6



In the pH range near neutral the amino acid is in the dipolar ion form. In acidic solution, the carboxylate group becomes protonated and the amino acid is in its cationic form. At basic solution, the ammonium group gives up a proton and the amino acid exists as an anion:

**Table 2.2** Structures of the 20 standard amino acids

Nonpolar Side Chains					
Glycine (G)	Alanine (A)	Valine (V)	Leucine (L)	Isoleucine (I)	
					
Methionine (M)		Proline (P)	Phenylalanine (F)	Tryptophan (W)	
					
Polar Side Chains					
Serine (S)	Threonine (T)	Asparagine (N)	Glutamine (Q)	Tyrosine (Y)	Cysteine (C)
					
Electrically Charged Side Chains					
Acidic Side Chains			Basic Side Chains		
Aspartic Acid (D)	Glutamic Acid (E)	Lysine (K)	Arginine (R)	Histidine (H)	
					



The pH at which an amino acid has overall neutral charge due to an equal proportion of negatively and positively charged groups is called the isoelectric point (pI).

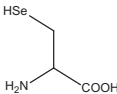
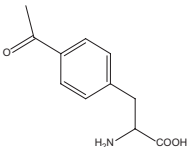
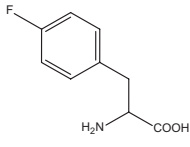
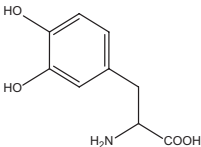
### 2.1.2 Nonstandard Amino Acids

In addition to the 20 standard amino acids, selenocysteine and pyrrolysine are now regarded as the twenty-first and twenty-second amino acids. Selenocysteine (Sec/U) is a cysteine analog but a selenium-containing selenol group instead of a thiol group which provides unique properties, such as lower  $\text{pK}_a$  value (approximately 5.2) compared to that of cysteine (approximately 8.4) (Arner 2010). On the other hand, pyrrolysine (Pyl/O) is an amino acid necessary for methanogenesis pathways (Krzycki 2013). Pyl contains a methylated pyrroline carboxylate in amide linkage to the  $\epsilon$ -amino group of L-lysine (Hao *et al.* 2004).

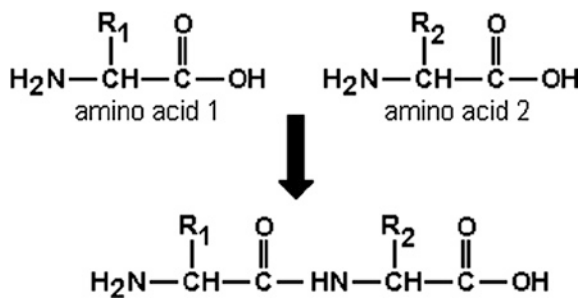
Nonstandard amino acids are either found as minor components of some specialized type of proteins or through modifications of standard amino acids. 4-hydroxyproline and 5-hydroxylysine are among those that are derived from one of the 20 natural amino acids and both are found in the fibrous protein collagen. Homocysteine is formed through the transsulfuration from cysteine (Brosnan and Brosnan 2006). *N*-methyllysine is found in myosin, a muscle protein.

There are also amino acids that occur biologically in either free or combined form. Examples include ornithine and citrulline which are derivatives of arginine and serve as intermediates in the formation of urea, part of amino acid catabolism (Curis *et al.* 2005). Aside from the derivatives of the  $\alpha$ -amino acids found in proteins, some amino acids have their amino group in the  $\beta$  or  $\gamma$  position such as  $\gamma$ -aminobutyric acid (GABA) which serves as neurotransmitter and  $\beta$ -alanine which is an important precursor of the vitamin pantothenic acid. GABA is nowadays used as a starting substrates for bio-based chemicals. Examples of nonstandard amino acids are shown in Table 2.3.

**Table 2.3** Structures of nonstandard amino acids

Selenocysteine	p-Acetylphenylalanine	p-Fluorophenylalanine	L-DOPA
			

**Fig. 2.1** Formation of peptide bond



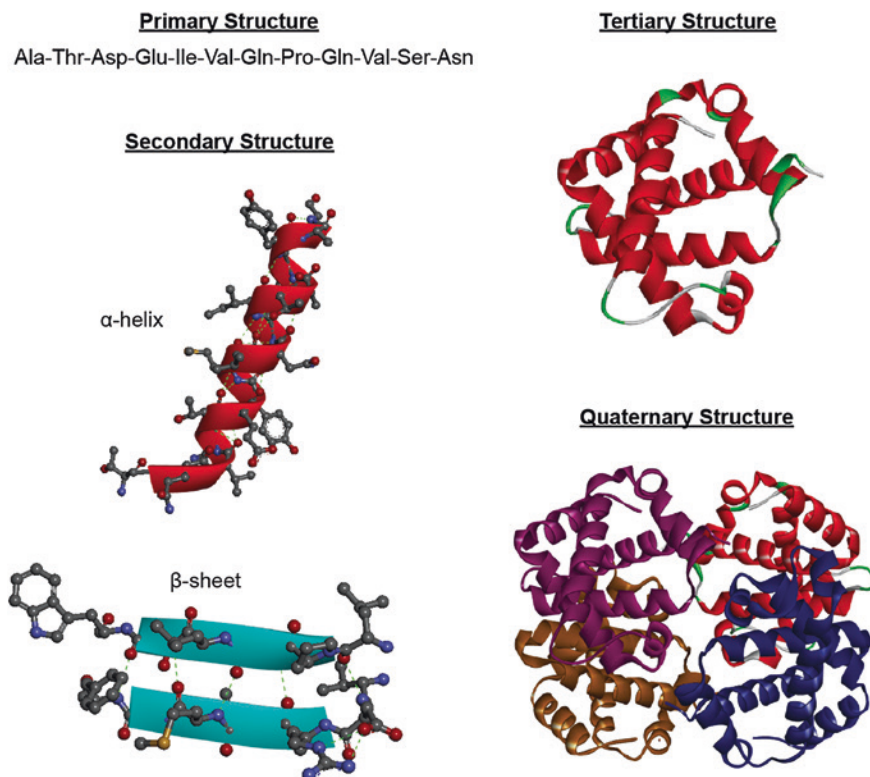
### 2.1.3 Proteins

Amino acids can form amide bonds by condensation between carboxyl group and amino group as shown in Fig. 2.1. The amide bonds are specifically called the peptide bonds.

If two amino acids are condensed, the product is called as dipeptide. When another amino acid condenses to this dipeptide, a tripeptide is formed. In this manner, a chain of amino acids can be linked to make a *polypeptide* or a *protein*. On the basis of their physical characteristics, proteins can be classified into globular, fibrous, and membrane-bound classes. Globular proteins have polypeptide chains that are tightly folded into spherical shapes. They are soluble in aqueous systems and diffused readily. Fibrous proteins, on the other hand, are physically tough and water insoluble. The polypeptide chains in fibrous proteins are arranged in extended, parallel form along a single axis to give fiber structures. They function as structural or protective elements in the organism. The protein  $\alpha$ -keratin for example is the major component of hair, feathers, nails, and skin. Another fibrous protein is collagen which is the major component of tendons.

Membrane proteins are also biologically important proteins. Within a cell and from cell to cell, membrane composition varies. Membrane proteins are found in a highly asymmetric environment which gives them unique properties. Outside the membrane is aqueous while the membrane interior is hydrophobic. Due to the two-dimensional surfaces of membrane proteins, they will concentrate or localized cellular components that regulates the nature and directionality of cell signals. However, their structural information has been known on a relatively small number of membrane proteins since isolating and crystallizing them is difficult.

There are four levels of protein structure (see Fig. 2.2) which are organized hierarchically from so-called primary structure to quaternary structure. The *primary structure* is the amino acid sequence of the polypeptide. The polypeptide chain that results through condensation reactions of amino acids retains a charged amino group at one end of the chain, N-terminus, and a charged carboxylate at the other end, C-terminus. The individual amino acids in a protein are numbered starting from the *N-terminus*.



**Fig. 2.2** Levels of protein structure

*Secondary structure* of protein refers to local sub-structures. *Alpha helix* and the *beta strand* (or *beta sheet*) are the two main types of secondary structure. These secondary structures are formed through specific patterns of hydrogen bonds between the main-chain amide and carboxylate groups.

*Tertiary structure* is a three-dimensionally folded structure due to secondary structure elements and interactions between side chains of the amino acids.

An assembly of several polypeptide chains into one large protein molecule is called the *quaternary structure*. Hemoglobin was the first oligomeric protein for which the complete tertiary and quaternary structures were determined.

## 2.2 Biosynthesis of Enzymes

### 2.2.1 Biosynthesis Mechanisms

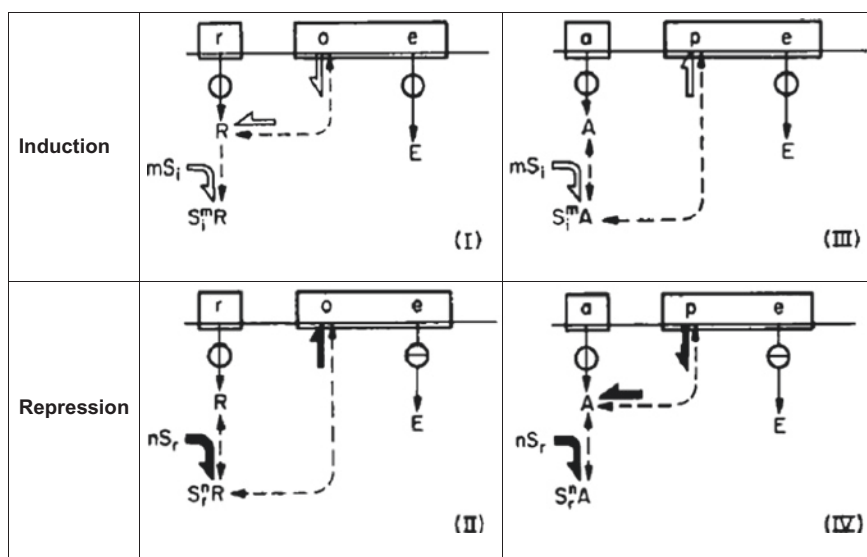
Enzyme biosynthesis can be classified into two types: the *constitutive* and the *inducible*. Enzyme that is produced at a constant rate is called constitutive enzyme. In contrast, inducible enzyme is expressed only under specific conditions

in which it can adapt to some environmental disturbance, thus its synthesis is not constant. One type of enzyme regulation is *induction* where enzyme expression is promoted by the presence of substrate or inducer. Contrary to the *induction* is *repression* where enzyme synthesis is suppressed by the presence of repressor molecules, for example intermediates or products of the enzyme reaction.

The foundation of enzyme production might be attributed to a genetic disturbance. The biosynthesis of an enzyme depends on (a) transcription of the genetic information from DNA or RNA into messenger RNA (mRNA) and (b) translation of the mRNA into polypeptide based on the genetic codon information in the mRNA. This is known as the operon model proposed by Jacob and Monod (1961).

The switching on and off for enzyme formation within a cell depends on the presence or absence of an effector molecule, which in turn is determined by the cell's environment. *Negative control* and *positive control* are the two main control mechanisms for the enzyme regulation. Four models of the switch are shown in Fig. 2.3 (Toda 1981).

For negative control of enzyme synthesis, the repressor (or its precursor) attaches to an operator region ( $o$ ) in the DNA by itself ( $R$ ) or as a complex ( $S_r$ - $R$ ) with an effector ( $S_r$ ). This binding inhibits the transcription process, thus the reduced enzyme ( $E$ ) synthesis.

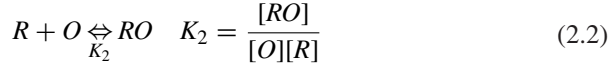
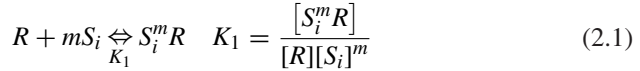


**Fig. 2.3** Four types of enzyme regulation. *Dashed lines* indicate operator gene-repressor binding or promoter gene-activator binding. The *bold arrows* indicate the binding or the detachment of repressor or activator from the genes in the presence of an effector molecule.  $A$  activator;  $a$  gene of activator;  $E$  enzyme;  $e$  gene of enzyme;  $m$  and  $n$  stoichiometric constants;  $p$  promoter gene;  $R$  apo-repressor or repressor;  $r$  repressor gene;  $S_i$  inducer;  $S_r$  co-repressor (Toda 1981)

When an activator stimulates protein synthesis, this enzyme regulation is called positive control. The activator molecule ( $A$ ) by itself or as a complex ( $S_i$ - $A$ ) with an effector ( $S_i$ ) can bind on a promotor ( $p$ ) in the DNA to enable transcription.

### 2.2.2 Mathematical Modeling

The enzyme regulation can be described by assuming the binding equilibrium between effectors (denoted as  $S_r$  for the co-repressor and  $S_i$  for the inducer), repressors ( $R$  and  $RS_r^n$ ) and the free operator gene ( $O$ ) for the induction system (Toda 1981):



Total material balance for  $R$  and  $O$  gives

$$[R_t] = [R] + [S_i^m R] \quad (2.3)$$

$$[O_t] = [O] + [OR] \quad (2.4)$$

The enzyme biosynthesis rate ( $Q$ ) or transcription rate is proportional to the ratio of free operator genes. For induction,

$$Q_i = \frac{[O]}{[O_t]} = \frac{1 + K_1[S_i]^m}{1 + K_1[S_i]^m + K_2[R_t]} \quad (2.5)$$

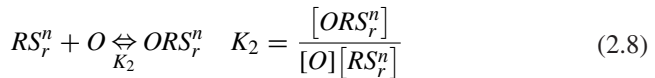
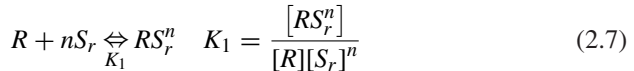
From Eq. 2.5 and for high concentration of inducer;  $Q_i \cong 1.0$  which means full expression for enzyme biosynthesis.

However, for inducer concentration  $S_i \cong 0$ ,

$$Q_i \cong \frac{1}{1 + K_2[R_t]} \neq 0 \quad (2.6)$$

This means minimal expression of enzymes.

For the repression system





$$Q_r = \frac{[O]}{[O_t]} = \frac{1 + K_1[S_r]^n}{1 + K_1[S_r]^n + K_1K_2[S_r]^n[R_t]} \quad (2.9)$$

### Further Discussion

1. Explain how  $\alpha$ -helix and  $\beta$ -sheet are formed. What are the roles of  $\alpha$ -helix and  $\beta$ -sheet in enzyme structure and function? Helix might be used for flexible motion and sheet as rigid motion. Can we get new helix or sheet by changing one or more amino acids in helix or sheet, and how about the properties of the new helix or sheet?
2. How can nonstandard amino acids be incorporated into enzymes? What are the advantages that can be expected using nonstandard amino acids?
3. Induction and repression mechanism can be mathematically expressed in many ways. What are the advantages of expressing the mechanism mathematically and what applications?
4. Explain the induction mechanism and repression mechanism, respectively based on positive control mechanism.
5. Can we engineer inducible enzyme to constitutive enzyme? What could be the advantages and potential problems?

### References

- Arnér ESJ. Selenoproteins - what unique properties can arise with selenocysteine in place of cysteine? *Experimental Cell Research*, 2010, 316:1296–1303.
- Brosnan J and Brosnan M. The sulfur-containing amino acids: an overview. *Journal of Nutrition*, 2006, 136 (6 Suppl): 1636S-1640S.21.
- Curis E, Nicolis I, Moinard C, Osowska S, Zerrouk N, Bénazeth S and Cynober L. Almost all about citrulline in mammals. *Amino Acids*, 2005, 29(3):177–205.
- Hao B, Zhao G, Kang P, Soares J, Ferguson T, Gallucci J, Krzycki J and Chan M. Reactivity and chemical synthesis of L-pyrrolysine — the 22nd genetically encoded amino acid. *Chemistry & Biology*, 2004, 11:1317–1324.
- Jacob F and Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 1961, 3:318–356.
- Krzycki JA. The path of lysine to pyrrolysine. *Current Opinion in Chemical Biology*, 2013, 17:619–615.
- Toda K. Induction and repression of enzymes in microbial culture. *Journal of Chemical Technology and Biotechnology*, 1981, 3:775–790.

Fundamentals of Enzyme Engineering

Yoo, Y.J.; Feng, Y.; Kim, Y.-H.; Yagonia, C.

2017, X, 209 p. 95 illus., 11 illus. in color., Hardcover

ISBN: 978-94-024-1024-2