Chapter 1
Amino Acids

1.1 Introduction

A large number of amino acids are found in nature, most of which are α-amino acids. These occur in free form or as constituents of other biomolecules like peptides, proteins, coenzymes, hormones, etc. Of these, twenty[1] α-amino acids (in fact, nineteen amino and one imino acid) are found to be fundamental to the sustenance of the life forms. These are utilised in the synthesis of peptides and proteins under genetic control, which in turn are vital for life. These α-amino acids are referred to as **coded amino acids** or **primary protein amino acids** or **proteinogenic amino acids.** The rest of the amino acid are referred to as **non-coded amino acids** or **non-protein amino acids**. These are usually formed by post-translational modification, i.e. modified after translation (protein biosynthesis). These modifications are generally essential for the function of the protein. At the present level, we would confine ourselves primarily to the coded or protein amino acids which are also called 'standard' or 'canonical' α-amino acids. Though the coded amino acids primarily serve as the building blocks of proteins, these are precursors to many other important compounds. For example, tryptophan is a precursor of the neurotransmitter serotonin, glycine is one of the reactants in the synthesis of haeme a porphyrin, while serine is a constituent of phosphatidyl serine a phosphoglyceride and so on.

The presence of an acid and amino functional groups in the same molecule makes the α-amino acids show interesting acid-base and other physical properties. These properties in turn are crucial to the structure and the biological roles played by α-amino acids and of the molecules derived from them. In addition to the α-amino acids,

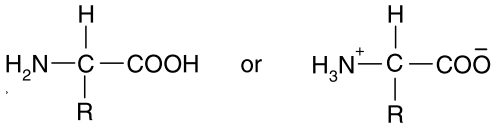
[1] In addition to these, two more amino acids by the names selenocysteine and pyrrolysine have been reported to be found in proteins. These amino acids do have genetic codes (UGA and UAG, respectively) but are not incorporated into the proteins in the usual way. Selenocysteine is found in several enzymes like *glutathione peroxidase, formate dehydrogenase,* etc., while pyrrolysine is found in microbes that produce methane.

V. K. Ahluwalia et al., *Chemistry of Natural Products*,
https://doi.org/10.1007/978-3-030-86698-3_1

many amino acids with amino group at β or γ positions or even more separated from the carboxyl group are also known and have important biological roles.

1.2 Nomenclature of Amino Acids

The amino acids obtained by hydrolysis of proteins with boiling aqueous acid or base, contain an amino as well as a carboxyl group (the amino group being α to the carboxyl group) and are generally known as α-amino acids. In general, an α-amino acid consists of an amino group, a carboxyl group, a hydrogen atom and a distinctive R group also called a side chain, bonded to a carbon atom, which is called the α-carbon. The general formula of an α-amino acid can be represented as shown below. The meaning and the significance of the charged structure are discussed in Sect. 1.5.2.

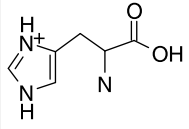


General structure of an α-amino acid

The traditional and well-known names of the common α-amino acids were, generally, given to them by their discoverers and bear no relationship to their chemical structures. These α-amino acids were normally named either on the basis of the source from which they were obtained or on some special property associated with them. For example, glycine derived its name from its sweet taste (Greek *glycos*, meaning sweet), while tyrosine got its name from its source, namely the milk protein, casein (Greek. *tyros*, meaning cheese). The semi-systematic names of substituted α-amino acids are formed according to the general principles of organic nomenclature, by attaching the name of the substituent group to the trivial name of the amino acid for example, N-methylglycine. A list of the coded α-amino acids along with the structures of their side chains, IUPAC names, one and three letter abbreviations (Sect. 1.2.1) is given in Table 1.1.

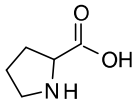
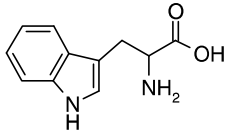
According to IUPAC, an α-amino acid that is otherwise similar to one of the common ones but contains one more methylene group in the carbon chain may be named by prefixing 'homo' to the name of that common amino acid. For example, homoserine and homocysteine are higher homologs of serine and cysteine, respectively.

Table 1.1 Names, symbols and structural formulae of coded α-amino acids

Common name	Three letter symbol	One letter symbol	IUPAC name	Structural formula	Frequency of occurrance[c] (%)
Alanine	Ala	A	2-Aminopropanoic acid	$CH_3–CH(NH_2)–COOH$	7.7
Arginine[b]	Arg	R	2-Amino-5-guanidinopentanoic acid	$H_2N–C(=NH)–NH–(CH_2)_3–CH(NH_2)–COOH$	5.1
Asparagine	Asn	N	2-Amino-3-carbamoylpropanoic acid	$H_2N–CO–CH_2–CH(NH_2)–COOH$	4.3
Aspartic acid	Asp	D	2-Aminobutanedioic acid	$HOOC–CH_2–CH(NH_2)–COOH$	5.2
Cysteine	Cys	C	2-Amino-3-mercaptopropanoic acid	$HS–CH_2–CH(NH_2)–COOH$	2.0
Selenocysteine	Sec[c]	U[c]	2-Amino-3-selenopropanoic acid	$HSe–CH_2–CH(NH_2)–COOH$	–
Glutamine	Gln	Q	2-Amino-4-carbamoylbutanoic acid	$H_2N–CO–(CH_2)_2–CH(NH_2)–COOH$	4.1
Glutamic acid	Glu	E	2-Aminopentanedioic acid	$HOOC–(CH_2)_2–CH(NH_2)–COOH$	6.2
Glycine	Gly	G	Aminoethanoic acid	$CH_2(NH_2)–COOH$	7.4
Histidine[b]	His	H	2-Amino-3-(1H-imidazol-4-yl)-propanoic acid		2.3
Isoleucine[a]	Ile	I	2-Amino-3-methylpentanoic acid	$C_2H_5–CH(CH_3)–CH(NH_2)–COOH$	5.3
Leucine[a]	Leu	L	2-Amino-4-methylpentanoic acid	$(CH_3)_2–CH–CH_2–CH(NH_2)–COOH$	8.5
Lysine[a]	Lys	K	2,6-Diaminohexanoic acid	$H_2N–(CH_2)_4–CH(NH_2)–COOH$	5.9
Methionine[a]	Met	M	2-Amino-4-(methylthio) butanoic acid	$CH_3–S–(CH_2)_2–CH(NH_2)–COOH$	2.4
Phenylalanine[a]	Phe	F	2-Amino-3-phenylpropanoic acid	$C_6H_5–CH_2–CH(NH_2)–COOH$	4.0

(continued)

Table 1.1 (continued)

Common name	Three letter symbol	One letter symbol	IUPAC name	Structural formula	Frequency of occurrance[c] (%)
Proline	Pro	P	Pyrrolidine-2-carboxylic acid		5.1
Serine	Ser	S	2-Amino-3-hydroxypropanoic acid	$HO–CH_2–CH(NH_2)–COOH$	6.9
Threonine[a]	Thr	T	2-Amino-3-hydroxybutanoic acid	$CH_3–CH(OH)–CH(NH_2)–COOH$	5.9
Tryptophan[a]	Trp	W	2-Amino-3-(lH-indol-3-yl)-propanoic acid		1.4
Tyrosine	Tyr	Y	2-Amino-3-(4-hydroxyphenyl)-propanoic acid	$HO–C_6H_4–CH_2–CH(NH_2)–COOH$	3.2
Valine[a]	Val	V	2-Amino-3-methylbutanoic acid	$(CH_3)_2CH–CH(NH_2)–COOH$	6.6

[a]Essential amino acids
[b]Semi-essential amino acids
[c]These are suggested symbols; not officially accepted as yet
[d]From Jones et al. (1991) *CABIOS* 8, 275–282

1.2.1 Representation of Amino Acids

The α-amino acids are commonly represented in terms of ***three letter symbols***. In this convention, the symbol for an amino acid is derived from its trivial name and includes the first three letters of this name (asparagine, glutamine, isoleucine and tryptophan being exceptions). The three letter symbols are written as one capital letter followed by two lowercase letters, e.g. glutamic acid is represented as Glu (not GLU or glu), regardless of its position in a sentence. The three letter symbols of coded amino acids are given in Table 1.1. These are used primarily to save space while representing peptides (obtained by condensation of two or more amino acids) and actually represent the unsubstituted amino acids. However, to represent the derivatives of the amino acids or amino acid residues in a peptide or protein, the symbols are modified by hyphens. The hyphen before the symbol of the amino acid indicates a bond formation with the α-nitrogen, while the one after the symbol of the amino acid is indicative of the bond formation to the carboxyl group.

For example, Ac–Ala represents the N-acetyl derivative of alanine, while Ala-OMe stands for the methyl ester of alanine. –Ala– represents an alanine residue in the peptide (Chap. 2). For example, Gly–Ala–Val represents the tripeptide glycylalanylvaline in which alanine is the central residue. Further, as a convention, these amino acid symbols denote the L-configuration of chiral amino acids (explained later) unless otherwise indicated by the presence of D– or DL–before the symbol and separated from it with a hyphen, for example, D–Ala.

Nowadays, ***one letter system*** is used more often. In this system, each α-amino acid has been assigned a one letter symbol, and the peptide or the protein is represented as a string of letters starting with the N-terminal (Sect. 2.3) amino acid moving towards C-terminal. The one letter symbol for eleven of the twenty coded amino acids is the first letter of their full name and for the rest nine distinct letters have been assigned so as to avoid any ambiguity (Table 1.1). The three letter abbreviations are quite straightforward; however, the relationship of one letter abbreviations to the names of the amino acids is somewhat less obvious. These are commonly used in representing long sequences as these save space and are less likely to be confused. For example, Q, E and G, the one letter symbols of glutamine, glutamic acid and glycine, respectively, are less likely to be confused than the obvious three letter symbols, viz. Gln, Glu and Gly, respectively. The three letter system, on the other hand, has advantages over one letter system as it has provision for representing the protecting groups and other structural details of the amino acids, e.g. Boc-Gly

represents the N-tertiarybutyloxocarbonyl derivative of glycine. Further, one letter symbols are **never** used to represent the individual amino acids in other contexts. A pentapeptide, glycylvalyltyrosylprolylglycine would be represented as Gly–Val–Tyr–Pro–Gly and GVYPG, respectively, in the three letter and one letter conventions. It may be noted here that while writing the name of the peptide, the suffix –**ine** of the amino acid name is replaced by –**yl**. This aspect has been discussed later in detail (Chap. 2).

In representing the amino acid sequences of peptides and proteins as derived from amino acid analysis (Sect. 3.4.1), two additional symbols, viz. Glx and Asx, are used. These ambiguous symbols are used due to uncertainties in the determination of glutamine and asparagine. In the course of their determination (acidic / basic conditions), the Gln and Asn residues may hydrolyse to give Glu and Asp, respectively. It is difficult to ascertain whether the Glu (or Asp) residue determined was originally a Glu (or Asp) or it was obtained from Gln (or Asn). In terms of one letter symbols, Glx and Asx residues are represented as 'X'.

Different atoms of the side chains of the amino acids are usually indicated in terms of Greek alphabets as shown for the amino acid, lysine in Fig. 1.1. It may be noticed that lysine contains an ε-amino group.

The nomenclature of α-amino acids and the conventions of representing them pertain to the ***coded amino acids***. However, the names of some ***non-coded amino acids*** which are closely related to coded amino acids are based on the three letter symbols of the coded amino acids, e.g. L-Hypro represents *trans*-4-hydroxy-L-proline that is obtained by post-translational hydroxylation of proline residue. As could

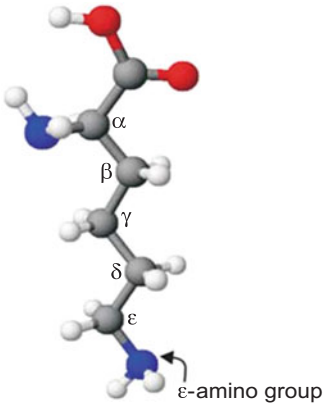


Fig. 1.1 Structure of lysine showing Greek designations of the side chain carbon atoms

be seen that these conventions are used for simplifying the representations of the structures of peptides and proteins. The systematic names and formulae given in Table 1.1 refer to hypothetical forms in which amino groups are unprotonated and carboxyl groups are undissociated. It should not be taken to imply that these structures represent an appreciable fraction of the amino acid molecules under physiological conditions.

1.3 Classification of Amino Acids

In a broad sense, amino acids can be put into two groups—the ***coded amino acids*** and the ***non-coded amino acids***. In the process of protein biosynthesis, the coded amino acids or the ***primary protein amino acids*** are incorporated one by one as per the instructions contained in the 'gene' for the concerned protein. The process is referred to as **translation** as it entails the translation of the genetic code to the amino acid sequence. Once the protein is synthesised, some modifications referred to as **post-translational modifications** can occur in some amino acid residues of the protein. Thus, proteins may contain certain amino acids other than the twenty coded amino acids. These are called as **secondary protein amino acids**. If, however, the modification leads to cross-linking of two amino acids, then these are referred to as **tertiary protein amino acids**. As mentioned earlier, we would confine ourselves primarily to the coded or primary protein amino acids.

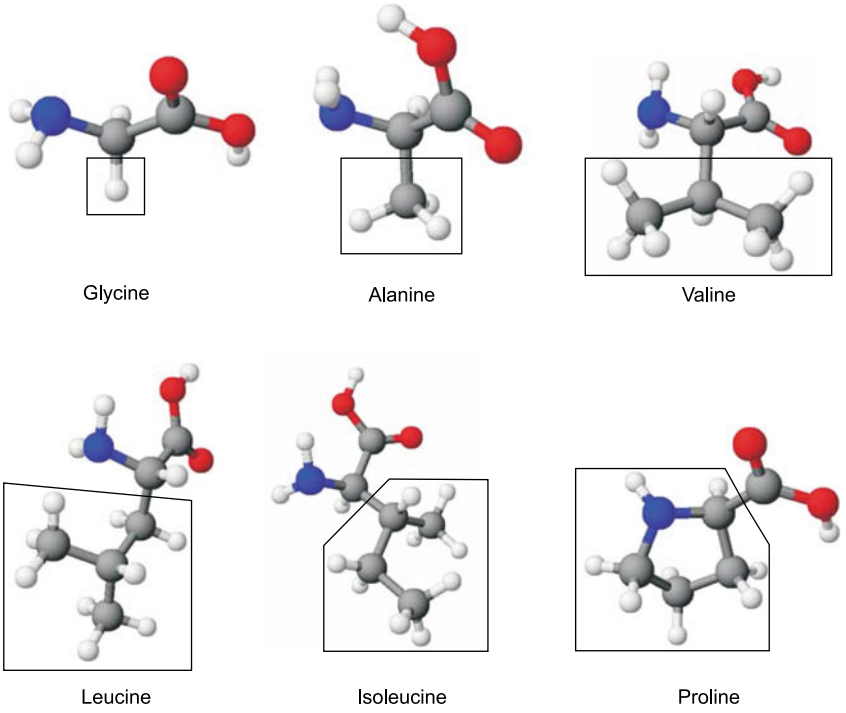
1.3.1 Coded or Primary Protein Amino Acids

The coded or primary protein amino acids with the general structure given earlier (Sect.1.2) contain different side chains varying in size, shape, charge, hydrogen bonding capacity and chemical reactivity. On the basis of structures and nature of the side chains, these amino acids are classified broadly into three groups—non-polar (or apolar), neutral polar and charged polar amino acids.

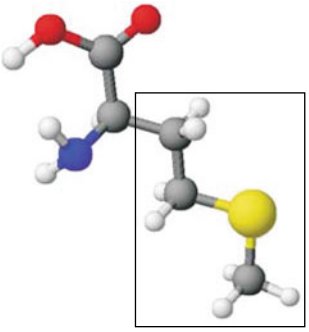
1.3.1.1 Non-polar or Apolar Amino Acids

As many as ten coded amino acids belong to this class. These include glycine, alanine, valine, isoleucine, leucine, proline, cysteine, methionine, tryptophan and phenylalanine. These amino acids are usually located on the interior of the protein as their side chains are hydrophobic in nature. Glycine, alanine, valine, leucine, isoleucine and proline side chains are purely aliphatic. Of these, proline is unique in the sense that it contains an aliphatic side chain that is covalently bonded to the nitrogen atom of the α-amino group, forming an imide bond and leading to a constrained 5-membered ring. The side chains (as shown below) in the case of

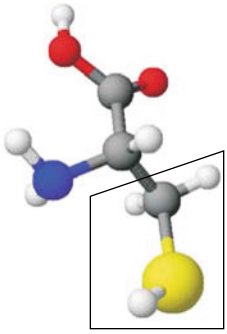
valine, leucine and isoleucine are bifurcated. This bifurcation or branching is close to the main chain and can restrict the conformation of the polypeptide by steric hindrance.



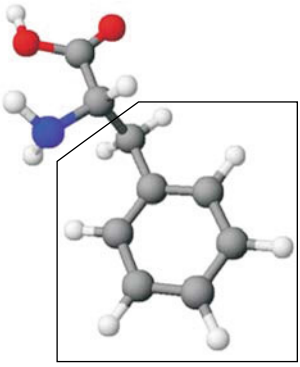
The non-polar side chains generally have low solubility in water because they can have only van der Waals interactions with water molecules. However, tryptophan with a nitrogen atom in its indole ring system is somewhat polar and can have hydrogen bonding interactions with other residues or even solvent molecules. The aromatic ring of phenylalanine is quite hydrophobic and chemically reactive only under extreme conditions, though its ring electrons are readily polarised. The sulphydryl (thiol) group of cysteine can ionise at slightly alkaline pH and can react with a second sulphydryl group to form a disulfide bond. On the other hand, methionine has a long alkyl side chain containing a sulphur atom that is relatively inert as a hydrogen bond acceptor.



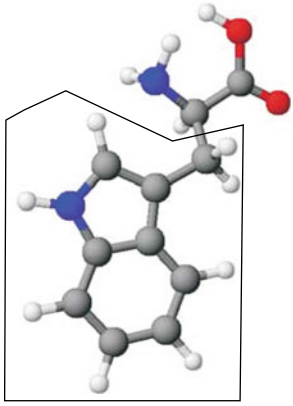
Methionine



Cysteine



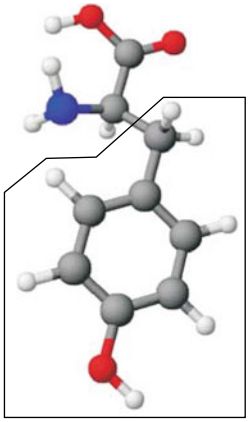
Phenylalanine



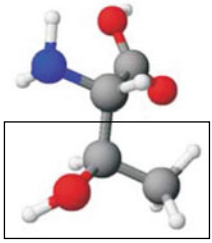
Tryptophan

1.3.1.2 Neutral or Uncharged Polar Amino Acids

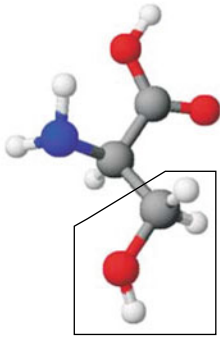
The amino acids belonging to this category have side chains with an affinity for water but are not charged. Serine, threonine, tyrosine, aspargine and glutamine belong to this category. Of these, the first three amino acids have a hydroxyl group, while the latter two have amide groups in their side chain. Due to the presence of hydroxyl groups, serine, threonine and tyrosine can function as both donors and acceptors in hydrogen bond formation. The phenyl ring of tyrosine permits stabilisation of the anionic phenolate form obtained on the ionisation of the hydroxyl proton, which has a pK_a of about 10. Serine and threonine, however, cannot be deprotonated at ordinary pH. Due to the presence of amide group, asparagine and glutamine side chains are relatively polar and can act as donors or acceptors in the formation of hydrogen bonds.



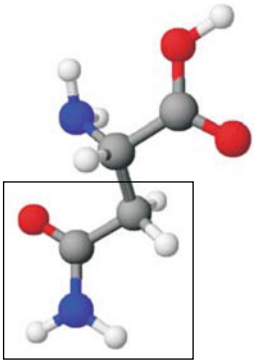
Tyrosine



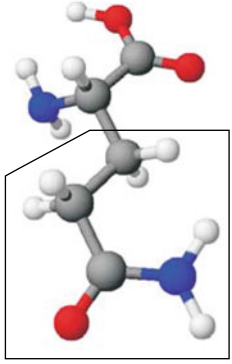
Threonine



Serine



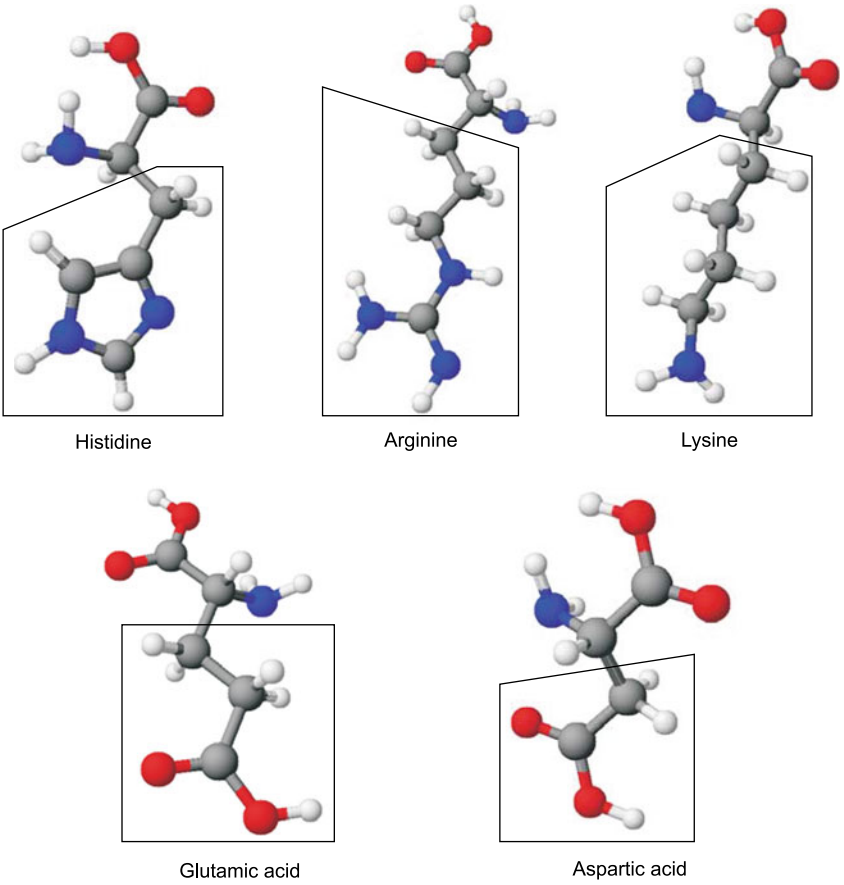
Asparagine



Glutamine

1.3.1.3 Charged Polar Amino Acids

As the name suggests, the members of this group, viz. lysine, arginine, histidine, glutamic acid and aspartic acid, have polar side chains due to the presence of positive or negative charge at the end of their side chain. The lysine ε-amino group has a pK_a value close to 10, while the guanidino group in the arginine side chain has a pK_a value of ~12. Histidine is another basic residue with its side chain organised into a closed ring structure that contains two nitrogen atoms. Glutamic and aspartic acid are two amino acids that contain a carboxyl group in their side chain. These differ only in the number of methylene groups in the side chain, with one and two methylene groups, respectively. Their carboxylate groups are extremely polar and can readily form hydrogen bonds, by acting as a donor or acceptor. These have pK_a values of about 4.5.



1.3.2 Secondary and Tertiary Protein Amino Acids

In addition to the twenty α-amino acids that are the primary components of proteins, some other amino acids exist naturally. These are obtained by modification of the primary protein amino acids. The common modifications include simple derivatisation like hydroxylation, methylation, acetylation, carboxylation and phosphorylation on the side chains of some amino acids. These lead to **secondary protein amino acids or secondary amino acids.** For example, hydroxylysine and hydroxyproline are simply functionalised derivatives of primary protein amino acids, lysine and proline, respectively. These are found in collagen, a common structural protein.

Some secondary amino acids are obtained by replacing the hydrogen atom of OH, SH or NH group in the side chain of the primary protein amino acid by glycosyl,

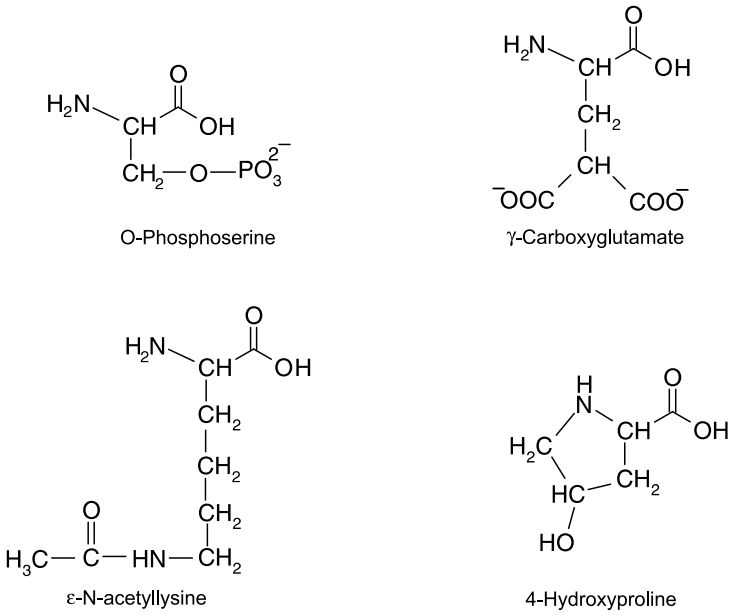


Fig. 1.2 Some secondary protein amino acids

phosphate or a sulphate group. In some cases, larger groups like polymeric carbohydrates or lipids are attached to amino acid residues. Such secondary amino acids are commonly found in glycoproteins. In some other cases, the primary protein amino acids are derivatised at the N-terminal as a methyl derivative to form a secondary amino acid. For example, secondary amino acid like N^{ω}-methylarginine is commonly found in histones, the nuclear proteins. The structures of some secondary protein amino acids are given in Fig. 1.2.

Cross-linking of amino acids in a protein generates **tertiary protein amino acids**. A disulphide bridge formed by the oxidation of the thiol groups of two cysteine residues to give cystine is the most common cross-link occurring in proteins. Sometimes the side chain carboxyl group of a glutamic acid residue can get into a cross-link with the ε-amino group of lysine to give a tertiary protein amino acid, ε-(γ-glutamyl) lysine. The structures of some tertiary protein amino acids are given in Fig. 1.3.

1.3.3 Non-coded or Non-protein Amino Acids

Non-protein amino acids (also called non-standard amino acids) are those amino acids which are neither found in proteins assembled during protein biosynthesis nor are generated by post-translational modifications. This may be due to the lack of a specific codon (genetic code) and t-RNA. Hundreds of such amino acids are known,

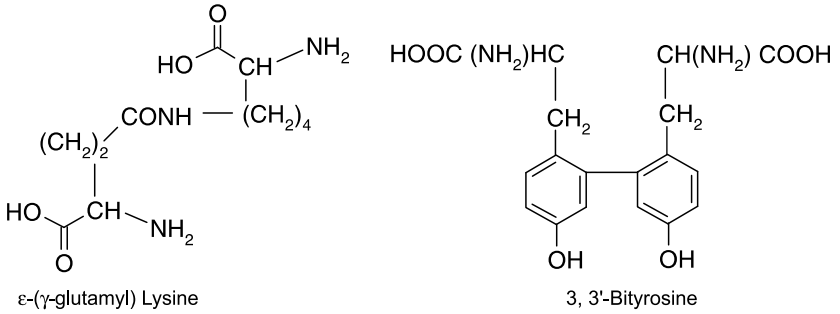


Fig. 1.3 Some tertiary protein amino acids

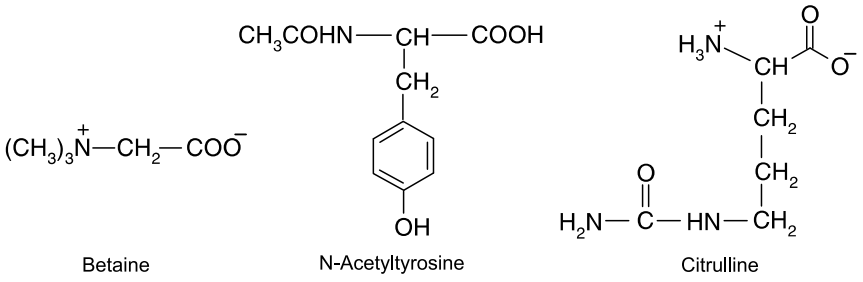


Fig. 1.4 Some non-protein amino acids

and a large number of these are α-amino acids. These non-coded amino acids are found mostly in plants and microorganisms and arise as intermediates or as the end product of the metabolic pathways. For example, N-acetyltyrosine is formed in the metabolism of tyrosine; betaine is involved in glycine biosynthesis and citrulline participates in ornithine cycle. These may also arise in the process of detoxification of the compounds of foreign origin. The structures of some non-protein amino acids are given in Fig. 1.4.

It is difficult to ascribe an obvious direct function to most of these amino acids in an organism. However, most of the functions of these amino acids in plants and microorganisms may be associated with other organisms in the environment. Canavanine-a homologue of arginine in which the δ-methylene group of arginine is replaced by an oxygen atom is found in alfalfa seeds and acts as a natural defence against insect predators.

1.3.4 Essential Amino Acids

In addition to the above-mentioned classification based on the nature of side chains, amino acids are also classified as essential and non-essential amino acids on the basis

Table 1.2 Requirement of essential amino acids

Essential amino acid	Requirement (g per kg dietary protein)
Isoleucine	42
Leucine	48
Lysine	42
Methionine	22
Phenylalanine	28
Threonine	28
Tryptophan	14
Valine	42

of their source in the living system. Of the twenty standard amino acids listed in Table 1.1, in case of humans, more than half of these can be made by the body itself, while the others, called **essential amino acids**, must come from the diet. These are required to maintain the nitrogen balance in the body. In fact, the meaning of the term essential differs from one species to the other. The classification of an amino acid as essential or non-essential does not reflect its importance because all the twenty amino acids are necessary for normal functioning of the body. This classification system simply reflects whether or not the body is capable of manufacturing a particular amino acid. The essential amino acids (for human beings) are isoleucine, leucine, valine, lysine, methionine, phenylalanine, threonine and tryptophan.

Some other amino acids have also been added to this list. These are the ones that are synthesised in the body but not at a rate required for the normal growth of the organism. The amino acids, arginine and histidine, belong to this list of **semi-essential** amino acids because the body does not always require dietary sources for it. Two amino acids, viz. cysteine and tyrosine, occupy ambiguous position in this classification. These are synthesised in the body in adequate amounts but use two essential amino acids, viz. methionine and phenylalanine, respectively, for their synthesis. Therefore, their presence in the diet indirectly decreases the requirement of methionine and phenylalanine.

The requirement of essential amino acids per kilogram of the dietary protein is given in Table 1.2. This is called the reference pattern of the amino acids and acts as a standard to determine the quality of the protein being consumed. Further, a number of essential amino acids are toxic if taken in excess. For example, a diet containing excess of leucine may cause pellagra. Therefore, a caution should be exercised while taking protein or amino acid supplements.

1.4 Stereochemical Aspects of α-Amino Acids

The tetrahedral array of four different groups about the α-carbon atom of α-amino acids (with the exception of glycine) makes it asymmetric-a chiral carbon. This asymmetry produces optical or stereoisomers in the amino acids. It is known that different amino acids contain different side chains (R groups) varying in size, shape,

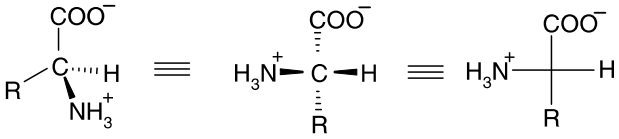
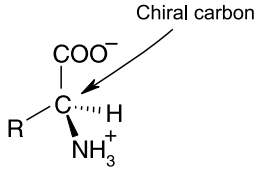


Fig. 1.5 Fischer–Rosanoff representation of an α-amino acid

charge, hydrogen bonding capacity and chemical reactivity. Glycine is the simplest amino acid with just one hydrogen atom as its side chain. Due to the absence of a chiral atom, glycine is not optically active.



It is advisable to draw the structures of amino acids in the **Fischer–Rosanoff** convention. For this, the chiral atom of the amino acid is projected onto the plane of the paper in such a way that the central atom appears as the point of intersection of a vertical line that joins three atoms of the principal chain, i.e. $COO^- - C_\alpha - R$ and a horizontal straight line joining the attached groups, viz. NH_3^+ and H. The central atom is considered to lie in the plane of the paper, the atoms of the principal chain behind the plane from the viewer, and the remaining two groups in front of the plane. The Fischer–Rosanoff representation of an amino acid may be drawn as shown in Fig. 1.5.

1.4.1 Absolute Configuration of α-Amino Acids

The absolute configuration at the chiral carbon atom of α-amino acids is designated by a small size capital letter prefix, D or L. This prefix indicates a formal relationship of the given amino acid to D- or L-serine-a standard used for correlating the configurations of the amino acids. The configurations of D- or L-serine, on the other hand, are related to D- or L-glyceraldehyde, respectively. In other words, correlation of the absolute configuration of the amino acid with D- or L-serine is equivalent to their correlation with D- or L-glyceraldehyde. It is reemphasised here that these D- and

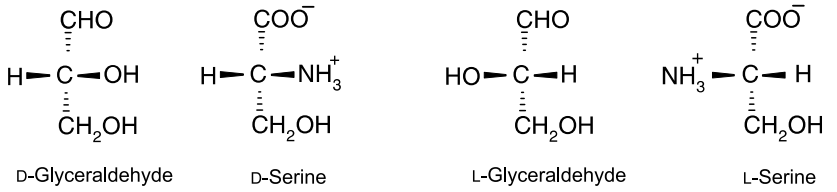


Fig. 1.6 Correlation of absolute configuration of D- and L-serine with D- and L-glyceraldehyde

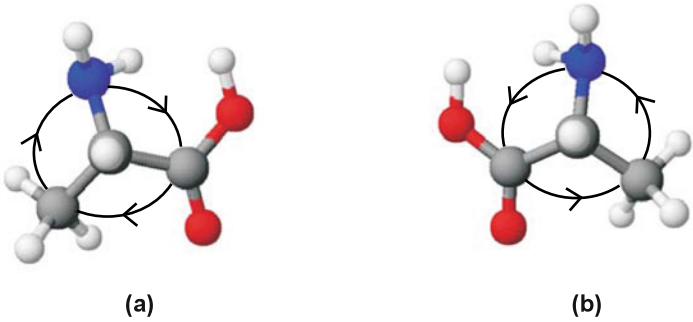


Fig. 1.7 Three-dimensional structures of **a** L-alanine and **b** D-alanine

L-notations do not indicate anything about the optical rotation of the amino acids. The relationship between the two isomeric forms of serine and the corresponding glyceraldehydes is shown in Fig. 1.6. All the natural α-amino acids are found to be L-amino acids.

To ascertain whether an amino acid is a L-amino acid or a D-amino acid, we may use the mnemonic **CORN**, which is an acronym for –**CO**OH, –**R** and –**N**H_2 groups. For this, on looking along the hydrogen-α-carbon bond of an amino acid, if the eyes move **clockwise** as we move through –COOH group, then the –R group and then the –NH_2 group starting at the carboxylic acid group (Fig. 1.7), then the amino acid has L-configuration else it belongs to the D-configuration.

1.4.1.1 The *RS* Notation

In the more general system of stereochemical designation, i.e. the *RS* or *CJP* (Cahn-Ingold-Prelog) system, the ligands of a chiral atom are placed in an order of preference, according to an arbitrary priority scheme. In this scheme, the atom of higher atomic number takes precedence over that of the lower atomic number (e.g. –OH

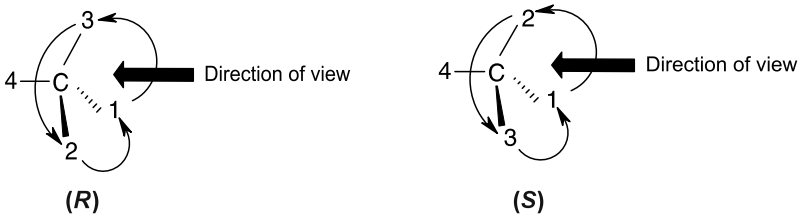


Fig. 1.8 Relative disposition of different substituents around the chiral atom in R and S notations. The numbers indicate the order of preference

precedes $–NH_2$). However, if the first atom of the substituent group happens to be same, then the preference is decided on the basis of the next atom (e.g. $–CH_2OH$ precedes $–CH_3$). These groups are labelled as 1, 2, 3 and 4 (or W, X, Y and Z). On viewing from the side opposite to the least-preferred (4th or Z) group, if the first three groups appear in clockwise order, the chiral centre is designated as *R* (from Latin, *rectus*); if anticlockwise, it is *S* (from Latin, *sinister*), Fig. 1.8.

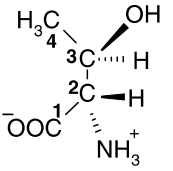
The L-configuration, possessed by the chiral α-amino acids found in proteins, nearly always corresponds to *S* or sinister notation, because in most amino acids the order of preference of the groups around the chiral carbon atom is NH_3^+ (1), COO^- (2), R (3) and H (4), Fig. 1.9a. However, the amino acids L-cysteine and L-cystine are exceptions as these have *R* notation. It is due to the presence of a sulphur atom in the side chain (i.e. $–CH_2SH$). Here, since the atomic number of sulphur is higher than that of oxygen, the group *R* takes precedence over carboxylate ion in these amino acids, Fig. 1.9b. The order is NH_3^+ (1), CH_2SH (2), COO^- (3) and H (4).

A mixture of equimolar amounts of D- and L-amino acids containing one chiral centre is termed racemic and is designated by the prefix DL (no comma between D and L), e.g. DL-leucine or it may be designated by the prefix *rac*- (e.g. *rac*-leucine).

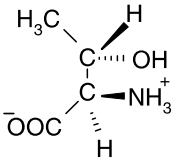
Fig. 1.9 Order of preference of substituents in **a** alanine and **b** cysteine

1.4.2 Amino Acids with Two Chiral Centres

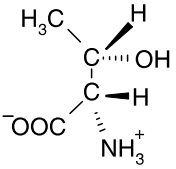
Historically, amino acids with two chiral centres were named by allotting a name (common or trivial) to the first diastereoisomer to be discovered. The second diastereoisomer, when found or synthesised, was then assigned the same name but with the prefix allo-. Since this method can be used only with trivial names, it has now been recommended that the prefix allo- should be used only for the amino acids, isoleucine and threonine.



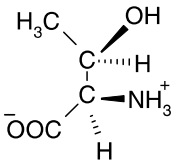
2*S*,3*R*-Threonine
(L-Threonine)



2*R*,3*S*-Threonine
(D-Threonine)

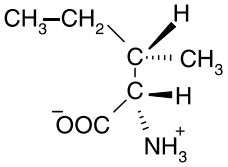


2*S*,3*S*-Threonine
(L-*allo*-Threonine)

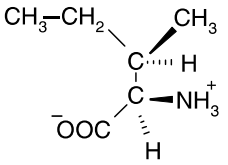


2*R*,3*R*-Threonine
(D-*allo*-Threonine)

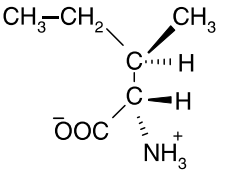
In amino acids with two chiral centres, the D- and L-notations refer to the configuration of the α-carbon atom. Threonine and isoleucine both have a second asymmetric centre at carbon position 3 along their chains. The absolute configuration of L-threonine, the normal protein component, is 2*S*, 3*R*. Its enantiomer (the mirror image), D-threonine, is 2*R*, 3*S*. The diastereomer with 2*S*, 3*S* configuration is called L-*allo*-threonine, while the one with 2*R*, 3*R* configuration is referred to as D-*allo*-threonine. The *allo*-form of a given enantiomer has its configuration inverted at the second chiral atom; for example, L-threonine is 2*S*, 3*R,* while L-*allo*-threonine has 2*S*, 3*S* configuration. On the other hand, the configuration of normal L-isoleucine is 2*S*, 3*S,* while its enantiomer is 2*R*, 3*R*. The diastereomers (2*R*, 3*S* and 2*S*, 3*R*) are called D-*allo*-isoleucine and L-*allo*-isoleucine, respectively.



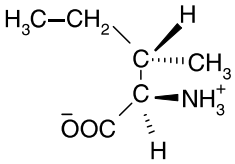
2*S*, 3*S*-Isoleucine
(L-Isoleucine)



2*R*, 3*R*-Isoleucine
(D-Isoleucine)



2*S*, 3*R*-Isoleucine
(L-*allo*-Isoleucine)



2*R*, 3*S*-Isoleucine
(D-*allo*-Isoleucine)

The optically inactive diastereoisomers of amino acids with two chiral centres are represented by a prefix *meso*- or *ms* if the optical inactivity is due to internal compensation, e.g. *meso*-cystine, Fig. 1.10.

The direction of rotation of plane polarised light of specified wavelength in a specified solvent and temperature is denoted with a 'plus' or 'minus' sign in parenthesis. This may be done for emphasis, with or without a configurational symbol D or L, e.g. (+)-glutamic acid or (+)-L-glutamic acid. The sign of optical rotation precedes all other details of the compound like substituents, etc., e.g. (+)-6-hydroxytryptophan. A racemic mixture of amino acids, however, is indicated by (±), e.g. (±)-alanine.

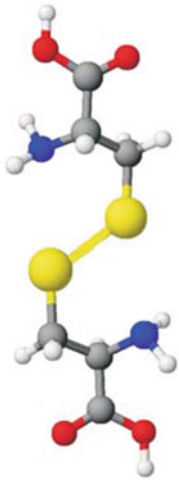


Fig. 1.10
Three-dimensional structure of meso-cystine

1.5 Physical Properties of α-Amino Acids

The physicochemical properties of the amino acids are very important in determining the structure and function of the peptides and proteins derived from them. The presence of an amino group and an acid functional group in the same molecule along with the presence of variable side chains makes amino acids show interesting properties especially the acid-base properties. Further the presence of a chiral carbon makes amino acids show optical properties.

1.5.1 General Physical Properties

Amino acids are colourless crystalline solids that are soluble in water, though to different extents. Cysteine, lysine and proline are highly soluble, while alanine, arginine, glycine and threonine have moderate solubility, and others have a low solubility (from 0.5 to ~ 9 g per 100 g). Tyrosine has an exceptionally low solubility of ~ 0.05 g per 100 g of water at 25 °C. None of the amino acids show a sharp melting point; all with the exception of cysteine and glutamine decomposes over 200 °C. The taste–a subjective property of the amino acids varies from bitter to neutral to sweet. Cysteine, however, has a sulphurous taste. The physical properties of twenty coded α-amino acids have been compiled in Table 1.3.

All the amino acids with the exception of glycine are optically active. The specific rotation ($[\alpha]_D^{25}$) values of the coded amino acids in water and 5 M HCl are also given in Table 1.3. The $[\alpha]_D^{25}$ values in HCl are more positive or less negative than those in water. The optical rotatory dispersion (ORD) and circular dichroic (CD) properties are discussed under spectroscopic properties (Sect. 1.5.3).

1.5.2 Acid-Base Properties of Amino Acids

The amino acids are observed to have low solubilities and high melting points as compared to the compounds containing similar groups and having comparable molar mass as shown below.

Compound	Molar mass (g mol^{-1})	Solubility		Melting point (°C)
		In water	In alcohol	
Lactic acid	90	Very high	Very high	~4
Alanine	89	Moderate	Insoluble	~297
3-Amino-2-butanol	89	Very high	Very high	~9

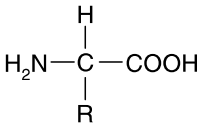
This difference can be explained in terms of the dipolar nature of the amino acid. This is due to the formation of an internal salt by movement of a proton from

Table 1.3 Physical properties[b] of twenty coded α-amino acids

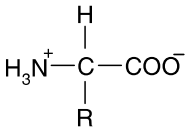
α-Amino acid	Molar mass (g mol^{-1})	Decom. temp. (°C)	Solubility in water (g/100 g 25 °C)	Taste in H_2O (pH = 6.0)	$[\alpha]_D^{25}$		Surface area[c] in $Å^2$	Van der Waal's volume[d] in $Å^3$
					H_2O	5M HCI		
Alanine	89.10	297	16.5	Sweet	+1.6	+13.0	115	67
Arginine	174.21	238	15.0[§]	Bitter	+21.8	+48.1	225	167
Asparagine	132.12	236	3.1	Bitter	−7.4	+37.8	160	148
Aspartic acid	133.11	270	0.5	Bitter	+6.7	+33.8	150	67
Cysteine	121.16	178[a]	v.sol	Sulphurous	−20.0	+7.9	135	86
Glutamic acid	147.14	249	0.84	Tasty	+17.7	+46.8	190	109
Glutamine	146.15	185	3.6	Sweet	+9.2	+46.5	180	114
Glycine	75.07	292	25	Sweet	–	–	75	48
Histidine	155.16	277	7.59	Bitter	−59.8	+18.3	195	118
Isoleucine	131.18	284	4.12	Bitter	+16.3	+51.8	175	124
Leucine	131.18	337	2.3	Bitter	−14.4	+21.0	170	124
Lysine	146.19	224	v.sol	Flat	+19.7	+37.9	200	135
Methionine	149.22	283	3.5	Tasty	−14.9	+34.6	185	124
Phenylalanine	165.20	284	2.97	Bitter	−57.0	−7.4	210	135
Proline	115.14	222	162.3	Sweet	−99.2	−69.5	145	90
Serine	105.10	228	5.0	Sweet	−7.9	+15.9	115	73
Threonine	119.12	253	20.5	Sweet	−33.9	−17.9	140	93
Tryptophan	204.23	282	1.14	Bitter	−68.8	−5.7	255	163
Tyrosine	181.20	344	0.05	Bitter	N.S	−18.1	230	141
Valine	117.15	315	8.85	Bitter	+6.6	+33.1	155	105

[a]as hydrochloride. § at 21 °C
[b]Data adopted from 'The Protein Amino Acids' by P.M. Hardy in chemistry and biochemistry of the amino acids Ed. G.C. Barrett pp. 6–24, Chapman and Hall (1985)
[c,d]From 'Proteins, Structure and Function' by David Whitford pp. 18–22, Wiley, (2005)

carboxylic group to amino group. This dipolar species is called a zwitterion, and the amino acid is said to be in **zwitterionic** form.



Unionised form



Zwitterionic form

In a crystal lattice, the 'ionic' amino acids are held together by strong forces of intermolecular interactions. These contribute to the high melting and boiling points

and solubility of amino acids. The presence of zwitterionic species is also responsible for the electrical conductivity of aqueous solutions of amino acids.

Since the amino acids contain acidic and basic functional groups and both of these groups are capable of ionisation in aqueous solutions, these show interesting acid-base properties. In water, amino acids can act both as acids as well as bases, i.e. these are **amphoteric** in nature. The nature of predominant molecular species present in an aqueous solution of amino acid will depend on the pH of the solution. This aspect is very important in determining the reactivities of different amino acid side chains and their influence on the properties of the proteins.

The acid-base behaviour of amino acids depends on the nature of the side chain. The twenty coded amino acids may be put into two groups for this purpose. The first group includes those amino acids which do not have an ionisable group in the side chain and the second group of amino acids with ionisable group in the side chain.

1.5.2.1 Amino Acids with Non-Ionisable Side Chains

The acid-base behaviour of the amino acids with non-ionisable side chains can be understood by taking alanine as an example. The curve for titration of alanine with NaOH shows two inflection points—one around a pH of 2.3 and another around 9.7 (Fig. 1.11). These correspond to the pK_a values of the carboxyl group and the amino group, respectively.

At a pH of less than 2, both the carboxyl and amino functional groups will be protonated, and the alanine molecule will have a net positive charge (species I). As we raise the pH of the solution, the carboxyl group gets deprotonated, and we get a dipolar species II with no net charge. There would be an equilibrium between species

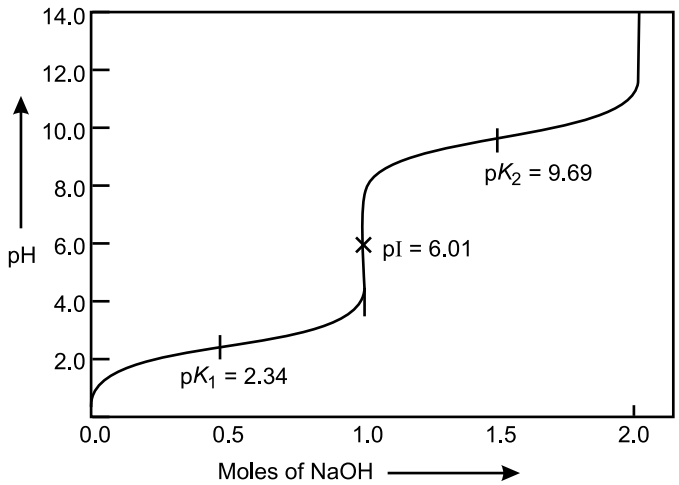
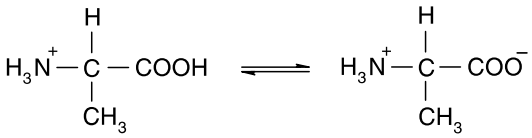


Fig. 1.11 Titration curve for alanine-an amino acid with non-ionisable side chain

I and species II, and the relative amounts of the two being determined by the pH of the solution. The equilibrium can be represented as:

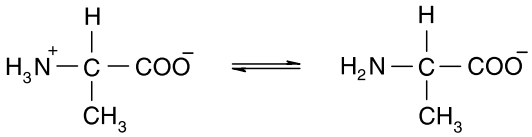


I II

This equilibrium would shift towards species II with further increase in the pH till a characteristic pH is reached, called the **isoelectric point (pI)** where species II would be predominant.

This behaviour is common to all amino acids other than the ones containing an acidic or a basic side chain. Thus, isoelectric point, pI, may be defined as the pH of an aqueous solution of an amino acid at which its molecules on an average have no net charge. The pI for alanine is found to be about 6.0, which means that at a pH of 6.0, alanine would be primarily in the zwitterionic form.

The dipolar ion (species II) is also a potential acid as the $-NH_3^+$ group can donate a proton. Therefore, further rise in the pH would cause this ammonium group to deprotonate and generate an anionic form of the amino acid (species III). There would again be an equilibrium this time between species II and species III, and the relative amounts of the two species again being determined by the pH of the solution. At a pH greater than 10, the anionic species III would be predominant.



II III

The isoelectric point of alanine is related to the pK_a values of the carboxyl and the amino functional group, and in fact, it is the average of the two pK_a values.

$$\text{pI} = \frac{2.34 + 9.69}{2} = \frac{12.03}{2} = 6.015 \simeq 6.0$$

The acid-base properties of the coded α-amino acids are compiled in Table 1.4.

In general, the isoelectric points of the amino acids can be computed by the following formula.

$$pI = \frac{1}{2}\left(pK_i + pK_j\right)$$

Table 1.4 Acid-base properties of coded α-amino acids

α-amino acid	pK_a; CO_2H (main chain) pK_1	pK_a; α-NH_3^+ $pK2$	pK_a; side chain pK_R	lsoelectric pH
Alanine	2.34	9.69	–	6.01
Arginine	2.17	9.04	12.84	10.76
Asparagine	2.02	8.60	–	5.41
Aspartic acid	1.88	9.60	3.65	2.77
Cysteine	1.71	8.18	10.28	5.02
Glutamic acid	2.16	9.67	4.32	3.24
Glutamine	2.17	9.13	–	5.65
Glycine	2.34	9.60	–	5.97
Histidine	1.82	9.17	6.00	7.59
Isoleucine	2.36	9.68	–	6.02
Leucine	2.36	9.60	–	5.98
Lysine	2.18	9.12	10.53	9.82
Methionine	2.28	9.21	–	5.74
Phenylalanine	1.83	9.13	–	5.48
Proline	1.99	10.6	–	6.30
Serine	2.21	9.15	–	5.68
Threonine	2.71	9.62	–	6.16
Tryptophan	2.38	9.39	–	5.89
Tyrosine	2.20	9.11	10.07	5.66
Valine	2.32	9.62	–	5.96

where pK_i and pK_j refer to the pK_1 and pK_2 values, i.e. the pK_a values for the main chain –COOH and the α-NH_3^+ group, respectively, for the neutral amino acids. However, for acidic amino acids, these refer to pK_1 and pK_R, while for the basic amino acids, these represent pK_R and pK_2 values, respectively; pK_R being the pK_a value of the side chain functional group.

The presence of different charged species in an aqueous solution of α-amino acid can be shown experimentally by observing the movement of these species in an electric field, using the technique called **electrophoresis**. In such an experiment, a gel containing a buffer solution of a given pH is set, and a small amount of the amino acid is placed near the centre of the gel. Then an electric potential is applied at its ends. The schematic representation of the set-up of this experiment is given in Fig. 1.12.

On applying the potential, different species present at the pH of the buffer move towards the electrodes depending on their net charge. In the experiment represented by Fig. 1.12, three different amino acids are examined simultaneously at a pH of 6.01. It can be observed that the spot for aspartic acid moves towards anode implying that this amino acid has a net negative charge. Similarly, the spot moving towards the

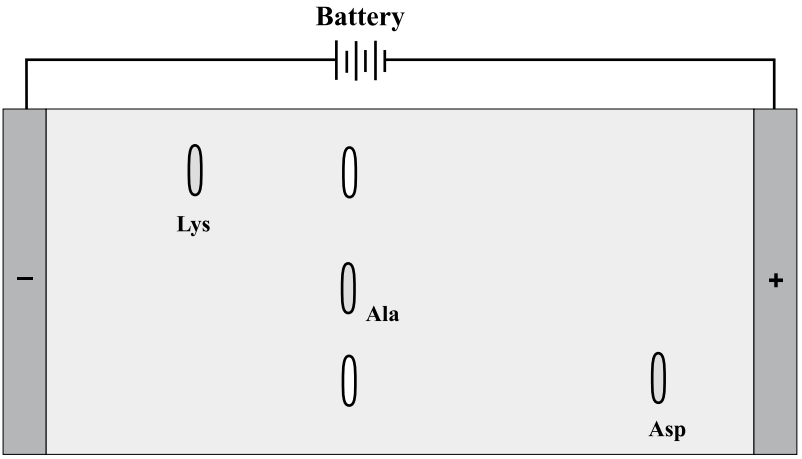


Fig. 1.12 Schematic representation of electrophoresis of a mixture of three amino acids, namely lysine, alanine and aspartic acid in a gel buffered at a pH of 6.01. The blank spots indicate the position of lysine and aspartic acid before applying the electrical field. In case of alanine, the position is same before and after applying the field

cathode refers to an amino acid having a net positive charge, i.e. lysine; while the amino acid spot that does not move indicates that the corresponding amino acid (alanine) is at its isoelectric point, i.e. it is in the dipolar or zwitterionic form.

1.5.2.2 Amino Acids with Ionisable Side Chains

The acid-base behaviour of the amino acids with ionisable side chains is quite complicated. These show complex titration curves with three inflection points one each for the carboxyl group, the amino group and the side chain ionisable group. The titration curve for aspartic acid with ionisable carboxyl group in the side chain is shown in Fig. 1.13.

For amino acids with ionisable side chains, the pK_a values for the carboxyl and the amino group are similar to the ones for the amino acids without ionisable side chains, while the pK_R (the pK_a values of the ionisable group in the side chain) varies with the amino acid. The pK_R for the acidic side chain of aspartic acid (3.65) is close to that of the pK_{COOH}, while the basic side chain of lysine ($pK_R = 10.53$) shows the inflection point close to the amino group ionisation. The equilibria involved in the titration of aspartic acid and lysine can be represented as shown in Fig. 1.14.

The acid-base equilibria of the ionisable side chains of the amino acids are summarised in Table 1.5.

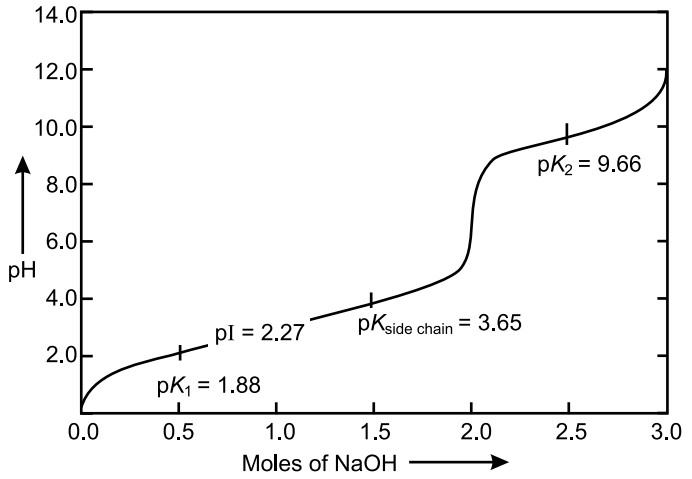


Fig. 1.13 Titration curve for aspartic acid having an ionisable carboxyl group in the side chain

1.5.3 Spectral Properties of α-Amino Acids

Spectroscopy, indisputably, is the most versatile and indispensable tool in the hands of practising chemist. Since the spectroscopic properties of the peptides and proteins depend a great deal on that of the α-amino acids constituting them, it is pertinent to have an understanding of the spectroscopic properties of α-amino acids. The general spectroscopic behaviour of α-amino acids is briefly described below.

1.5.3.1 Mass Spectrometry

In mass spectrometry, the volatile molecules to be analysed are ionised and fragmented. The resulting ions are separated and analysed on the basis of their mass to charge (*m/z*) ratio. Since amino acids are not volatile, these are normally studied as their volatile derivatives like esters. The fragmentation pattern of these derivatives is quite simple to analyse. There are two fragmentation pathways, both involving the cleavage of C–C bond α to the amino group. These are referred to as α-cleavages. The stability of the resulting immonium ions determines the preferred pathway.

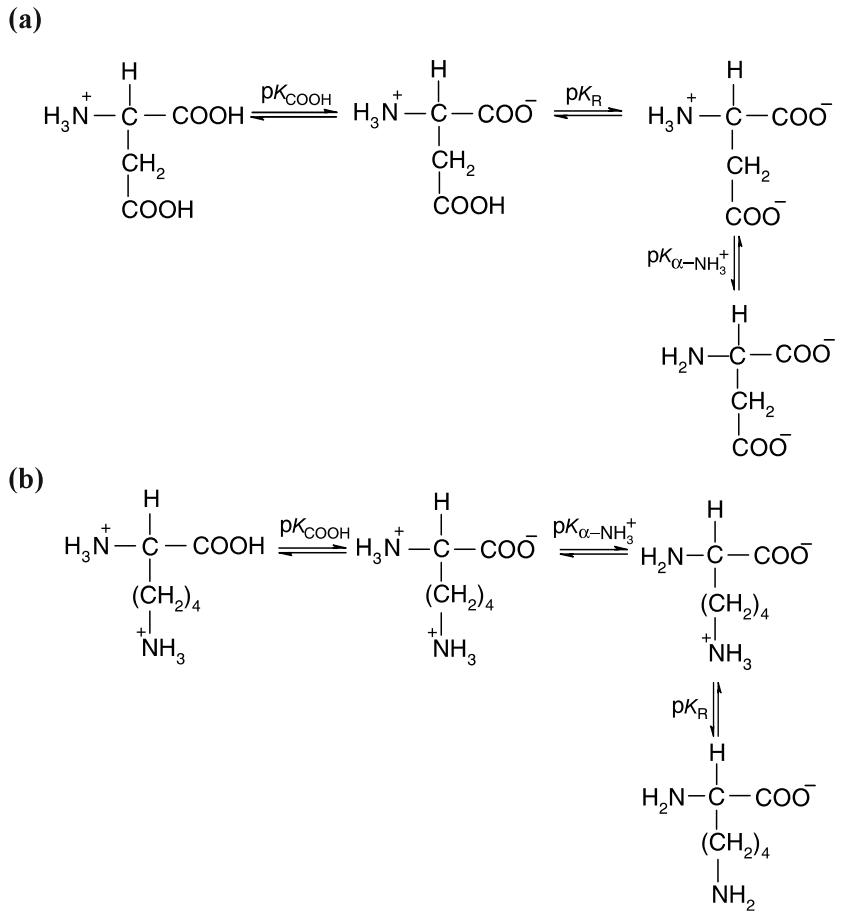
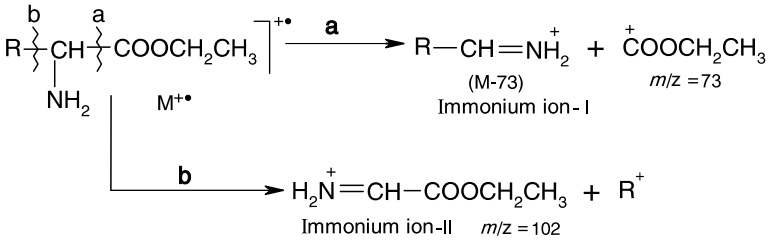


Fig. 1.14 Ionisation equilibria for **a** aspartic acid—an acidic amino acid and **b** lysine—a basic amino acid



In case of ethyl esters, the molecular ion peak is usually of low intensity. The peaks at an *m*/z of 73 (cleavage **a**) and 102 (cleavage **b**) on the other hand are of medium to high intensity. In the cleavage of type **a**, the peak at M-73 is very prominent

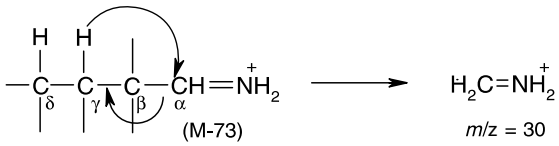
Table 1.5 Acid-base equilibria of the ionisable side chains of α-amino acids[a]

Amino acid residue group	Acid-base equilibrium of the side chain	Representative pK_a value
Aspartic acid/Glutamic acid (side chain carboxyl)	—COOH ⇌ —COO⁻ + H⁺ Carboxylic acid Carboxylate	4.4
Histidine (imidazole)	—CH₂— (HN⁺=NH) ⇌ —CH₂— (N=NH) + H⁺ Imidazolium Imidazole	6.5
Cysteine (thiol)	—SH ⇌ —S⁻ + H⁺ Thiol Thiolate	8.5
Tyrosine (phenolic)	—C₆H₄—OH ⇌ —C₆H₄—O⁻ + H⁺ Phenol Phenolate	10.0
Lysine (side chain amino)	—NH₃⁺ ⇌ —NH₂ + H⁺ Ammonium Amine	10.0
Arginine (guanidinyl)	—NH—C(=NH₂⁺)—NH₂ ⇌ —NH—C(=NH)—NH₂ + H⁺ Guanidinium Guanidino	12.0

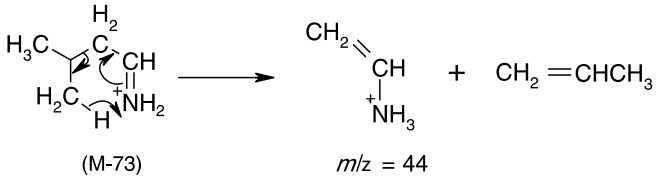
[a]The main chain carboxyl group of amino acid has a pK_a range of ~3.5–4.0, while for the α-amino group, the range is ~8.0–9.0

especially for the amino acids without a side chain functionality. This peak helps in the identification of the amino acid as it provides the mass of R, the side chain.

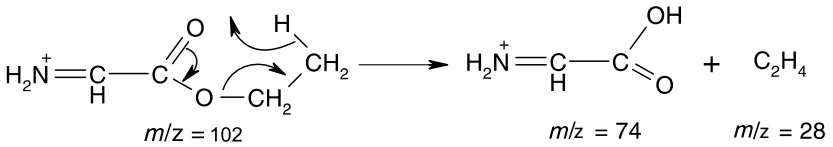
When the side chain contains two or more aliphatic carbon atoms, the M-73 ion ejects an alkene as a consequence of a hydrogen migration from farther carbon atoms and generates a strong peak at *m/*z of 30. Thus, a peak at *m/z* of 30 is indicative of valine, leucine or isoleucine.



In case of leucine, there is a possibility of **Mc Lafferty rearrangement** also to give a peak at *m/z* of 44.



In type **b** cleavage, the M-102 peak (due to R^+) is suggestive of the mass of the side chain. Further, the *m/z* 102 peak can also undergo a Mc Lafferty rearrangement to give a peak at *m/z* of 74.



In case the side chain of the amino acid contains a functional group, then the fragmentation pattern may be similar to the one discussed above with additional fragments. Sometimes the fragmentation pattern may be competitive with the general pattern discussed above.

1.5.3.2 Nuclear Magnetic Resonance (NMR) Spectroscopy

The NMR spectroscopy is extensively being exploited in the study of amino acids, peptides and proteins. The objectives vary from routine analysis to the conformational aspects and structure determination. However, the structural details of the peptides revealed by NMR are second only to the X-ray crystallographic analysis.

The chemical shift positions of the protons (^{1}H) and carbon (^{13}C) nuclei are found to depend significantly on the nature and state of ionisation of different groups of a given amino acid, i.e. whether the amino acid is in the cationic, zwitterionic or the anionic form. In case of peptides and proteins, the chemical shift positions depend on the nature and state of ionisation of functional groups of the neighbouring residues besides depending on solvent and temperature, etc. The chemical shift positions of α, β and other protons of the side chains of different amino acids are compiled in Table 1.6.

1.5.3.3 UV Spectroscopy

The common functional groups (amino, carboxyl and amide) present in the amino acids and peptides are transparent in the UV region, and therefore, amino acids with aliphatic side chains have no absorption above 220 nm in the ultraviolet range. However, the chromophores present in the side chains of a few coded aromatic amino

Table 1.6 PMR chemical shifts of α-amino acids[a,b]

α-amino acid	α-Hydrogen (ppm)	β-Hydrogen (ppm)	γ-Hydrogen (ppm)	δ-Hydrogen (ppm)	0ther protons (ppm)
1. Alanine	3.78	1.49	–	–	–
2. Arginine	3.18	1.87	1.6	3.20	
3. Asparagine	4.00	2.92, 2.87	–	–	–
4. Aspartic acid	4.08	3.01	–	–	–
5. Cysteine	3.82	2.93	–	–	–
6. Glutamic acid	3.75	2.12	2.49	–	–
7. Glutamine	3.77	2.16	2.43	–	–
8. Glycine	3.56	-	–	–	–
9. Histidine	3.97	3.16	–	–	Imidazole protons 7.03, 7.72
10. Isoleucine	3.66	1.97	1.35	1.00	ε = 1.00
11. Leucine	3.70	1.69, 1.73	1.71	0.96	ε = 0.96
12. Lysine	3.37	1.69	1.43	1.69	ε = 2.95
13. Methionine	3.80	2.60	2.12	2.12	–
14. Phenylalanine	3.97	3.20	–	–	Aromatic protons 7.36
15. Proline	4.08	2.06	2.06	–	–
16. Serine	3.94	3.94	–	–	–
17. Threonine	3.58	4.23	1.32	–	–
18. Tryptophan	4.04	3.38	–	–	Indole protons 7.31, 7.22, 7.28, 7.55
19. Tyrosine	3.93	3.05, 3.17	–	–	Aromatic protons 7.19, 6.88
20. Valine	3.62	2.29	1.06	1.06	–

[a]Data adopted from nuclear magnetic spectra of amino acids by G.C. Barrett and J.S. Davies in chemistry and biochemistry of amino acids Ed. G.C. Barrett, pp. 537–539, Chapman and Hall (1985)

[b]The values are for zwitterionic form in D_2O and are relative to DSS (2,2-dimethyl-2-silapentane-5-sulphonate) as internal standard

acids, viz. Tyr, Trp and Phe do absorb in the range of ultraviolet spectrum. The λ_{max} positions and the corresponding molar absorptivity values for these aromatic amino acids are given in Table 1.7.

This allows the application of UV spectroscopy in the quantification of peptides and proteins containing these residues. The pH-dependent spectra of proteins can provide information about number and nature of the tyrosyl residues present in the

Table 1.7 A_{max} and molar absorptivity values for aromatic amino acids at pH = 6.0

Amino acid	λ_{max} (in nm)	Molar absorptivity, ε (in cm^2 mol^{-1})
Tyrosine	193	48,000
	225	8000
	272	1200
Tryptophan	218	35,000
	281	5500
	188	60,000
Phenylalanine	208	8000
	260	150
Histidine	211	5900
Cysteine	250	300

protein. These chromophores provide 'optical handles' to study the conformational aspects of peptides and proteins by the circular dichroic measurements also. Further, the tyrosyl and tryptophanyl side chains show fluorescence spectra in the UV range. For example, tyrosine containing peptides and proteins shows an emission maxima around 310–315 nm on being excited at 275 nm.

1.5.3.4 IR Spectroscopy

The absence of characteristic NH stretching (3300–3500 cm^{-1}) and the carbonyl absorption for –COOH group (1700–1730 cm^{-1}) in the IR spectra of amino acids is indicative of the zwitterionic nature of amino acids. These characteristic frequencies instead are replaced by absorptions around 3070 cm^{-1} and 1560–1600 cm^{-1} that are characteristic of the NH_3^+ and COO^-, respectively. The NH_3^+ group has additional bands in the 1500–1600 cm^{-1} region. In addition to these, most of the amino acids show a medium absorption around 1300 cm^{-1} and weak absorptions in 2080–2140 cm^{-1} and 2530–2760 cm^{-1} regions. When these amino acids combine to give peptides, a number of additional bands characteristic of the peptide bond and of the conformation adopted by the peptide are also observed.

1.5.3.5 Circular Dichroism (CD)

In water, the C_α alkylated amino acids show a positive **cotton effect** with a peak around 200–204 nm at neutral pH. This band called **band 1** shifts to higher wavelengths in acidic and basic pH. In addition, these C_α alkylated amino acids show a negative cotton effect band called **band 2** around 245–250 nm in acidic pH range. The

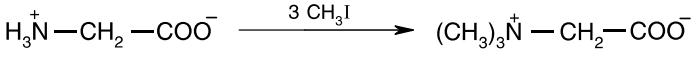
aromatic amino acids (except histidine) and the sulphur containing amino acids, on the other hand, show a maxima above 250 nm in their CD spectra. The CD spectra, like the UV spectra, are sensitive to the state of ionisation of different functional groups in the amino acid.

1.6 Chemical Reactions of Amino Acids

Since amino acids have both an amino as well as a carboxyl group, these show the reactions of the two groups individually as well as in combination. These reactions are outlined below.

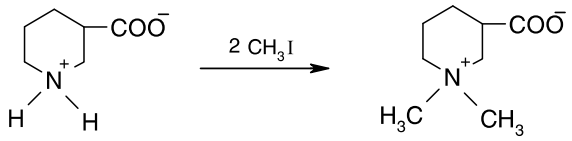
1.6.1 Reactions Due to Amino Group

1. **With alkyl halides**: Amino acids form alkyl derivatives on treatment with alkyl halides. In the presence of excess alkyl halide, these form internal quaternary alkyl ammonium salts. These salts are zwitterionic in nature and are called **betaines.** The term betaine is derived from a naturally occurring compound bearing the same name. Chemically, it is N, N, N-trimethylglycine which can be prepared from glycine in the presence of excess methyl iodide.



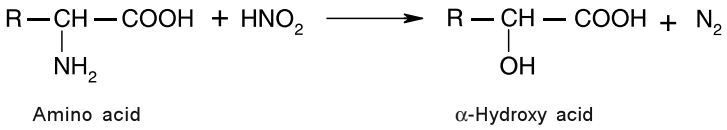
Glycine N,N,N-Trimethylglycine (betaine)

Similarly, the imino acid, proline containing cyclised side chain, reacts in the following manner to give a dimethyl derivative.



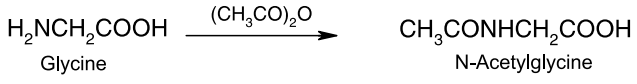
Proline N,N-Dimethylproline

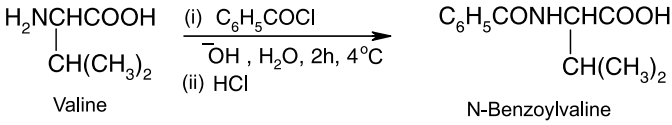
2. **With nitrous acid**: In the reaction of primary aliphatic amines with nitrous acid, the primary amino group is lost as nitrogen through the formation of corresponding diazonium group. The N_2 gas is evolved in quantitative yields. Amino acids with primary amino groups also undergo reaction with nitrous acid in a similar manner and produce corresponding hydroxy acid.



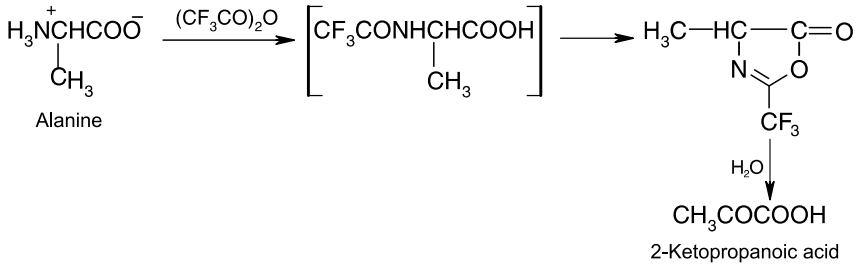
This reaction forms the basis for **Von Slyke determination** of amino nitrogen which in turn gives an estimate of the amount of amino acid.

3. **With acid anhydride**: Amino acids react with acid anhydrides and acid chlorides to produce acyl amino acids. The reaction is carried out under sufficiently alkaline conditions so that a substantial concentration of the free amino group is present. However, it is necessary to acidify the aqueous solution to obtain the acidic product.

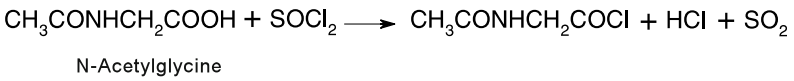




When α-amino acids are treated with trifluoroacetic anhydride, the initial amide rapidly cyclises, and the hydrolysis of the cyclised product gives α-keto acid.

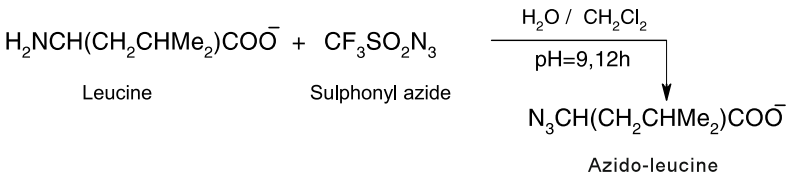


The acyl derivatives have an added advantage that these may be treated with $SOCl_2$ to give acid chlorides, an important step in peptide synthesis.



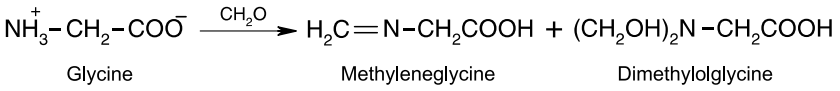
Formation of acid chlorides by direct treatment of amino acid with PCl_5 or $SOCl_2$ is otherwise difficult because of the presence of amino group.

4. **With sulphonyl azide**: The amino group of the amino acids reacts with sulphonyl azide under mild conditions and gets converted into azido group.



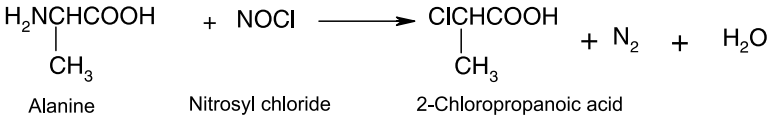
The reaction is significant because it proceeds without loss or change of optical activity.

5. **With formaldehyde**: Amino acids undergo nucleophilic addition with formaldehyde to form a mixture of products.

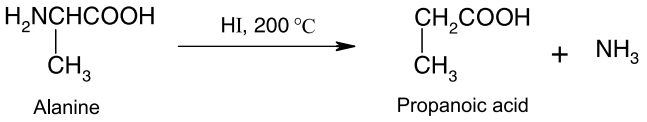


In the derivatives so obtained, the amino group of the amino acid is blocked, and the carboxyl group can be titrated with a standard alkali to determine the amount of the amino acid. Such a titration is called **Sorenson's formol titration**. A direct titration of the amino acid is not possible due to interference from the free amino group.

6. **With nitrosyl halides**: On reaction with nitrosyl chloride and bromide, the amino acids give the respective halo acids.

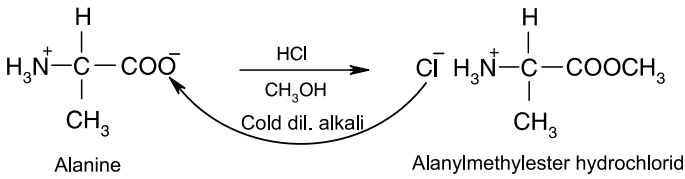


7. **With hydroiodic acid**: On reacting with hydroiodic acid at 200 °C, amino acids lose their amino group as a molecule of ammonia and give an acid as shown below

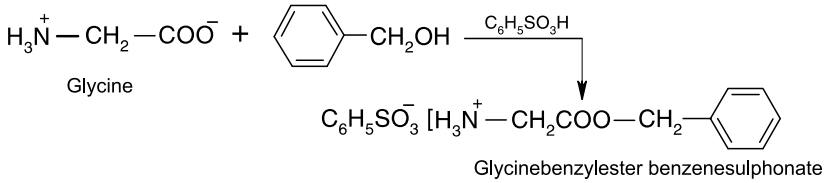


1.6.2 Reactions Due to Carboxyl Group

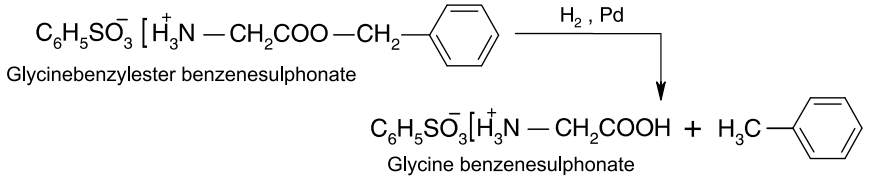
1. **Esterification**: The carboxyl group of amino acids can be easily esterified by treating a suspension of amino acid in an appropriate alcohol with anhydrous hydrogen chloride. The ester product formed can give back the amino acid on treatment with cold dilute alkali. The reaction of methyl alcohol with alanine is shown below.



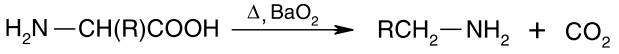
When benzene sulphonic acid is used as a catalyst, benzyl esters are obtained. Water produced in the reaction is removed by azeotropic distillation, avoiding the use of large excess of benzyl alcohol. For example, glycine forms the following benzyl ester on reacting with benzyl alcohol in the presence of benzene sulphonic acid.



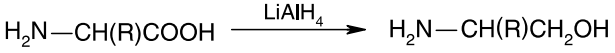
The product formed above may be reacted with H_2 in the presence of palladium to give the amino acid salt from which the free amino acid can be obtained.



2. **Decarboxylation**: Amino acids lose a molecule of carbon dioxide from the carboxyl group on heating with barium peroxide and yield amines.



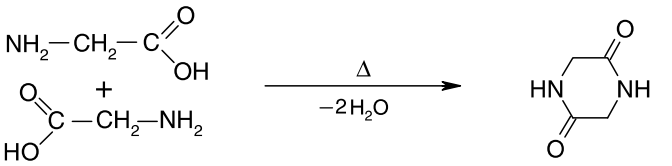
3. **Reduction**: On reduction with $LiAlH_4$ amino acids form amino alcohols with the retention of optical activity.



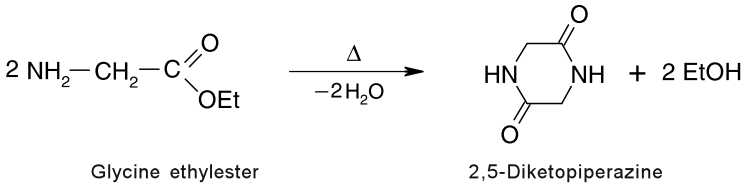
1.6.3 Reactions Due to Both Amino and Carboxyl Groups

1. **Pyrolysis**: On heating amino acids, different products are obtained depending on the position of the amino group.

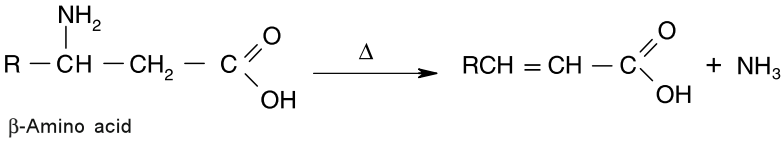
   (i) **α-amino acids**: α-amino acids give 2,5-diketopiperazines on heating. In this reaction, two molecules of α-amino acids react in such a way that the amino group of one forms an amide with the carboxylic group of the other and vice-versa as shown below. It is an example of double amide formation.



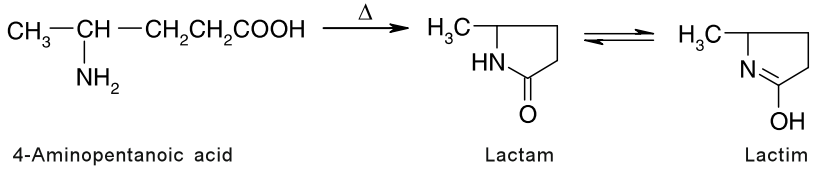
On heating, even the esters of amino acids give 2,5-diketopiperazine.



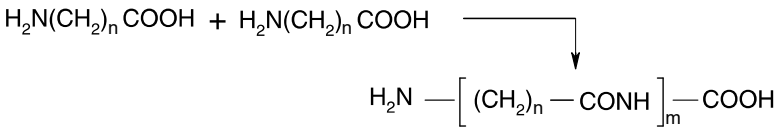
(ii) **β-amino acids**: On heating, β-amino acids eliminate a molecule of ammonia to give α, β-unsaturated acids.



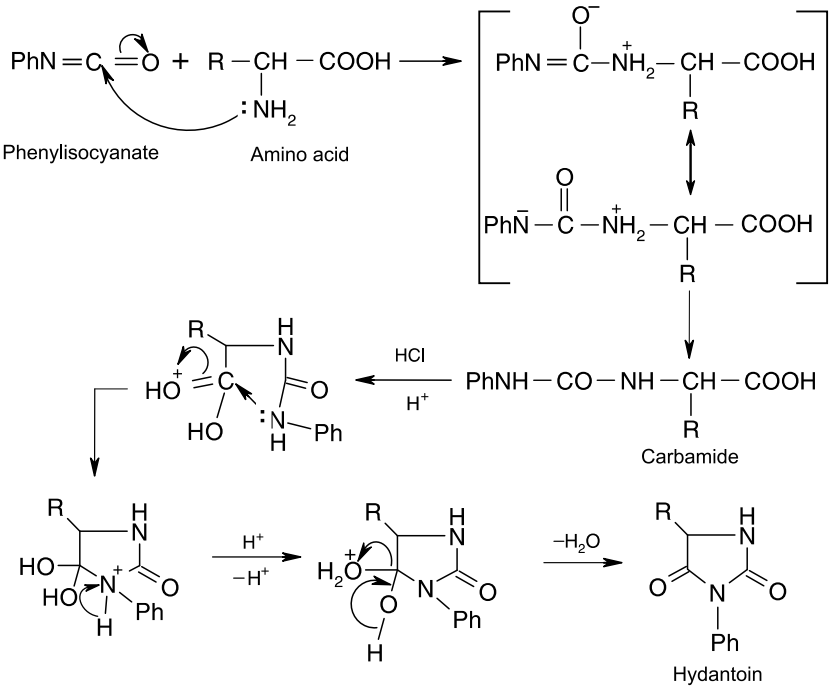
(iii) **γ-amino acids and δ -amino acids**: γ- and δ-amino acids on heating form cyclic amides called lactams. The lactams formed show a special type of tautomerism called **lactam–lactim tautomerism**.



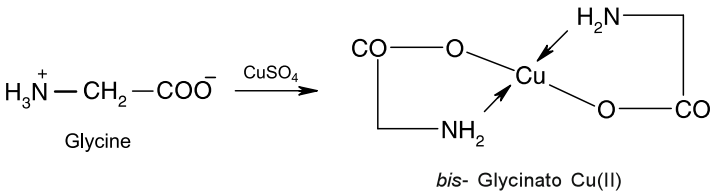
Linear polymeric amides are formed when amino and carboxylic groups are far from each other along the chain.



2. **Formation of hydantoin**: With isocyanates, α-amino acids form carbamides which form hydantoin on warming with HCl.

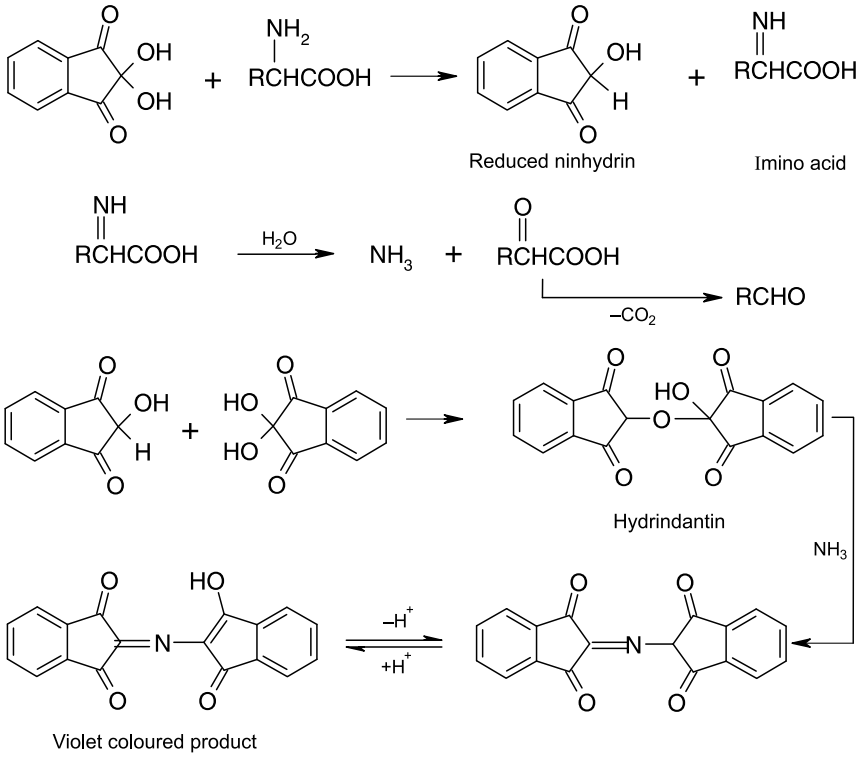


3. **Formation of metal chelates**: Amino acids form chelate structures when mixed with aqueous solution of $CuSO_4$ glycine.



4. **Oxidative deamination**: On reacting with ninhydrin (triketohy-drindene hydrate), the α-amino acids get oxidised into an imino acid which undergoes deamination, releasing ammonia. The reaction is called **oxidative deamination** and is also the path of biodegradation of amino acids. The likely reaction sequence is as given below.

The liberated ammonia reacts with hydrindantin to give the coloured product. This reaction forms the basis of detection of amino acids in their chromatographic determination wherein ninhydrin acts as a spray or detection agent.



1.7 Industrial Preparation of α-Amino Acids

The manufacturing methods of amino acids can be categorised as follows:

- Extraction from hydrolysates of animal or plant proteins
- Chemical synthesis
- Fermentation method.

The two methods, viz. extraction from hydrolysates of animal and plant proteins and fermentation, have been discussed below. The chemical synthesis being elaborate has been discussed separately under Sect. 1.8.

Extraction from Hydrolysates of Animal or Plant Proteins

Traditionally, amino acids were obtained by separating the mixture of amino acids obtained on hydrolysis of plant and animal proteins. In this method, proteins are extracted from plant or animal sources and are purified by dialysis. The free protein (devoid of ionic impurities) is then dried and subjected to acidic or alkaline hydrolysis or to enzymatic degradation with proteolytic enzymes. The separation of the amino acids from protein hydrolysate is then achieved with the help of separation procedures like electrophoresis, ion exchange, paper chromatography, etc.

Though this method provides moderate quantities of coded and post-transitionally modified α-amino acids, it is quite tedious and somewhat expensive. Besides, we cannot get all the amino acids by this method as a number of amino acids are either destroyed or modified during the process of hydrolysis (Sect. 3.4.1.1). More so with the availability of economically viable synthetic methods, this traditional method is losing importance.

Fermentation Method

The fermentation method is used for the commercial production of a number of amino acids. The manufacturing process by fermentation generally comprises three stages, viz.

- Fermentation
- Crude isolation
- Purification.

In the first step of fermentation process, a microorganism is grown in a suitable fermentation medium. In the next step, most impurities contained in the fermentation broth are removed, and in the third step, final purification is performed to obtain the desired product. In a typical production of L-amino acid by direct fermentation, the bacterium is grown in a suitable fermentation medium consisting of a carbon source (e.g. glucose, molasses, alkanes, glycerol, ethanol, etc.), ammonia as a nitrogen source and a small amount of minerals and vitamins as growth factors in a clean and sterile fermentation tank. The pH, temperature and dissolved oxygen content are the control factors in the fermentation process and are to be maintained. Once the growth of the bacteria reaches a certain level, the amino acid starts accumulating in

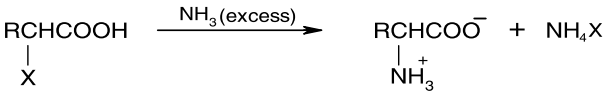
the medium. After the fermentation process is over, the broth is centrifuged or filtered through a membrane filter to separate bacterial cells and proteins. Crude crystals of the desired amino acid are then obtained by crystallisation of the supernatant or filtrate so obtained. Sometimes it is difficult to obtain the crystals in this way. In such cases, first the impurities are removed by means of an ion exchange resin or activated carbon, and then, the filtrate is subjected to crystallisation.

1.8 Chemical Synthesis of α-Amino Acids

A constant supply of optically pure α-amino acids and their analogues is required as these (or the peptides obtained from them) perform important biological functions and also have pharmaceutical importance. For example, α-methyl-3′, 4′-dihydroxyphenylalanine is used in the treatment of Parkinson's disease.

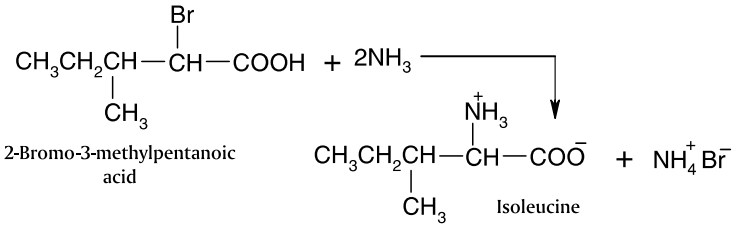
The coded amino acids can be obtained from protein hydrolysates by using suitable separation methods as discussed above. However, this method of obtaining amino acids is quite tedious and expensive. Further, non-coded amino acids cannot be obtained from the protein hydrolysates. Therefore, industrially viable methods for synthesising α-amino acids are required. A wide spectrum of methods is available for synthesising amino acids. Some of the commonly used methods of synthesising α-amino acids are given below.

1. **Amination of α-halo acids**: This is the most general method involving displacement reactions of α-halo acids. It involves amination of a α-halo acid with excess of ammonia (**direct ammonolysis**). However, not all amino acids can be synthesised by this method.

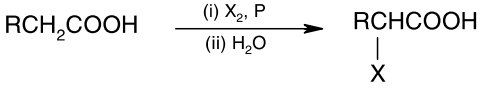


α-Halo acid α-Amino acid

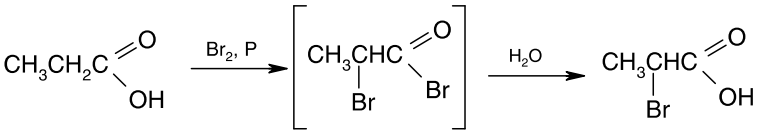
The reaction can also be visualised as alkylation of ammonia with alkyl α-halo carboxylic acids. In this S_N2 reaction, ammonia acts as a nucleophile. Since ammonia can be alkylated to di- and trialkyl stage, the reaction gives a complex mixture of products. Therefore, this method of preparing α-amino acids is usually not preferred. However, by using excess of ammonia, one can get more of monoalkylation product. For example,



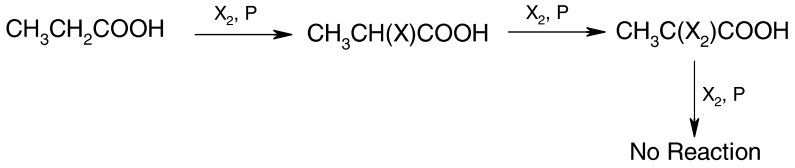
α-halo acids can be conveniently prepared by the **Hell–Volhard–Zelinsky (HVZ) reaction.** Here, carboxylic acids react with chlorine or bromine in the presence of phosphorus or phosphorus halide to yield the α-halo acid. Halogenations occur specifically at the α-carbon because the reaction probably proceeds via enolisation of the acyl halide. The function of phosphorus halide is to convert a little of the acid into acid halide.



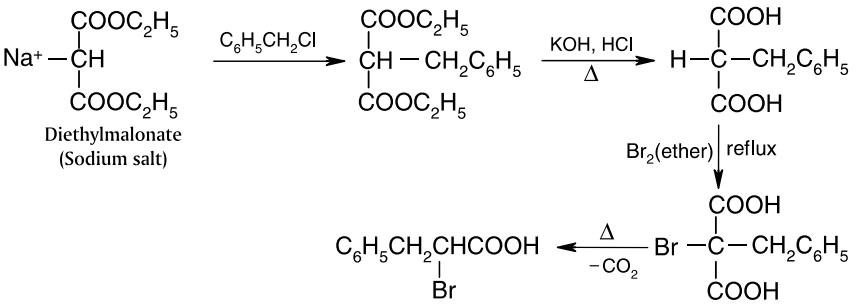
For example,



If more than one molar equivalent of the halogen is taken, then depending upon the availability of α-hydrogen atoms, di or trihalo acids are obtained.

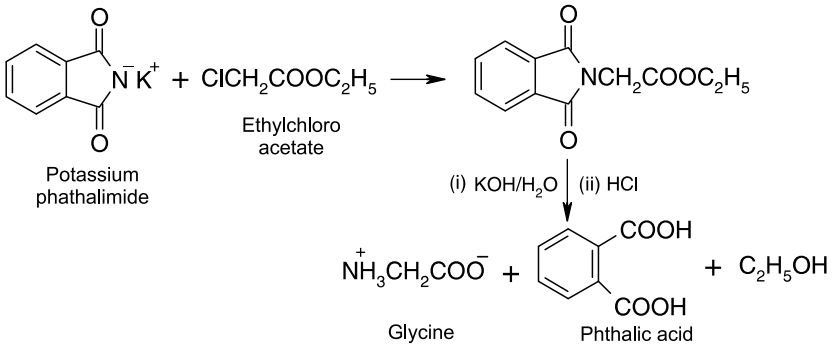


Alternatively, α-haloacids can also be prepared by using modified malonic ester synthesis. In this method, an ester is taken as the reactant instead of an acid. The reaction is given as follows:

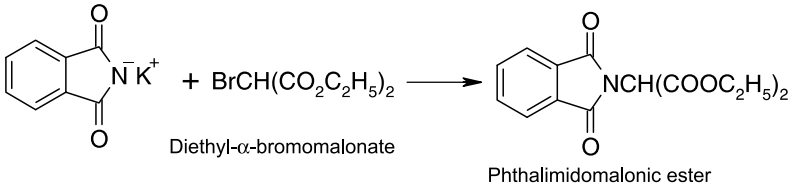


Amino acids like glycine, alanine, valine, leucine, isoleucine and aspartic acid, serine and threonine can be synthesised using direct ammonolysis method. The drawback of this method is that the yields are poor. The method given below is used to obtain better yields.

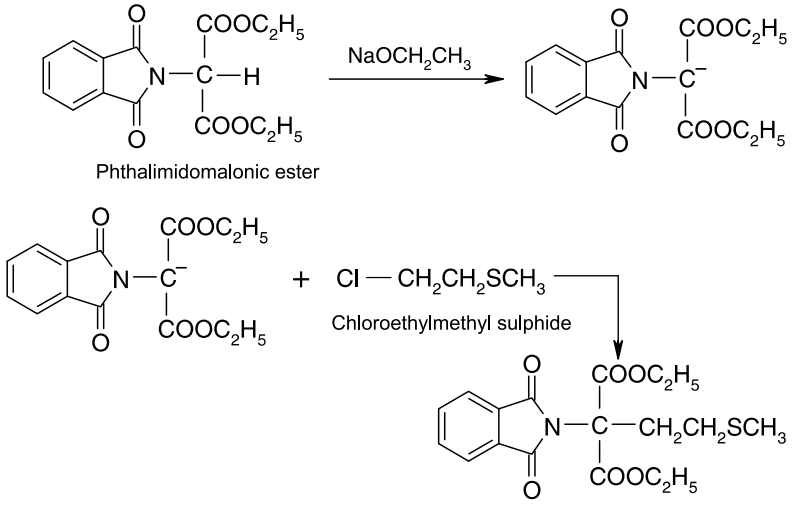
2. **Using potassium pthalimide**: This method is a modification of Gabriel pthalimide synthesis of amines and gives better yields. It is generally used for the preparation of glycine and leucine. In this method, α-haloesters are used instead of α-halo acids. The reaction between α-haloester and potassium phthalimide is carried out in the following steps to obtain α-amino acids.



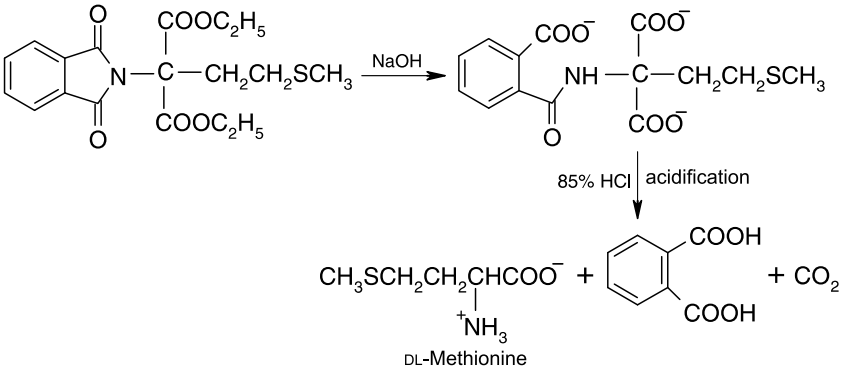
This method can be modified to carry out the synthesis of other amino acids like valine, isoleucine and methionine. It is called as **phthalimido malonic ester method** and makes use of diethyl-α-bromomalonate instead of ethylchloroacetate to prepare an imidomalonic ester.



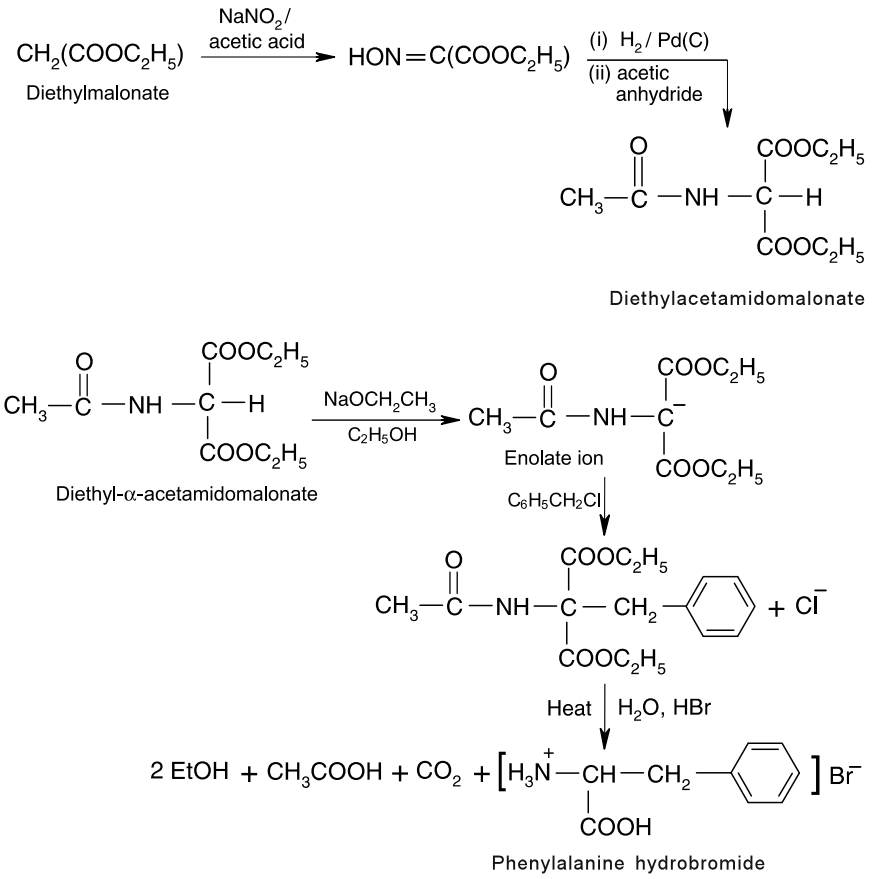
The amino malonate derivative (phthalimidomalonic ester) so obtained is then alkylated. For this, the ester is treated with sodium ethoxide in ethanol to generate an anion which then acts as a nucleophile in the S_N2 reaction with alkyl halide to give the alkyl derivative of malonic ester.



Saponification of the alkyl derivative followed by acidification gives substituted amino malonic acid which decarboxylates to generate an α-amino acid.

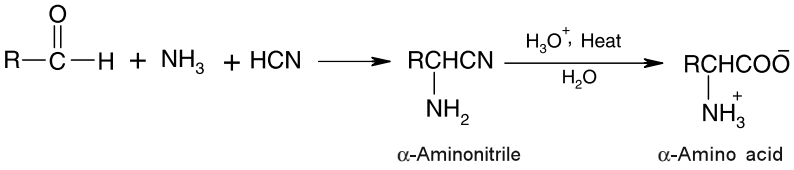


In place of phthalimido derivatives, other amino derivatives (acyl amino) of malonic ester may also be used for preparing α-amino acids. For example, diethyl-α-acetamidomalonate is treated with sodium ethoxide followed by an alkyl /aryl halide to give α-amino acid. Diethylacetami-domalonate can be easily obtained from diethyl malonate as follows.

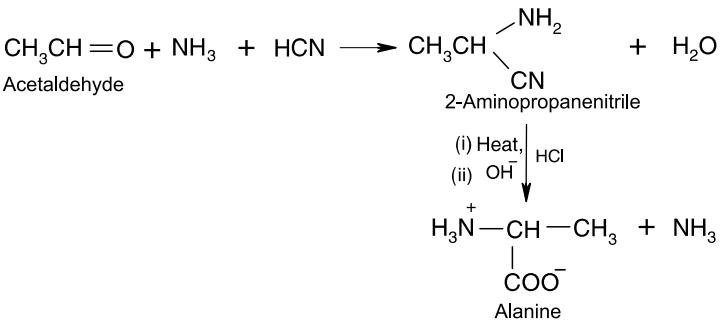


This method is sometimes referred to as **amidomalonate synthesis**. In addition to acylaminomalonate, formamidomalonates, or benzyloxycarbonylaminomalonates can also be used for this reaction. However, the conditions used for the removal of N-derivatives are somewhat different in these cases.

3. **Strecker synthesis**: In this method, α-aminonitriles (prepared by the reaction of an aldehyde with a mixture of HCN and ammonia) are hydrolysed to give α-amino acids.

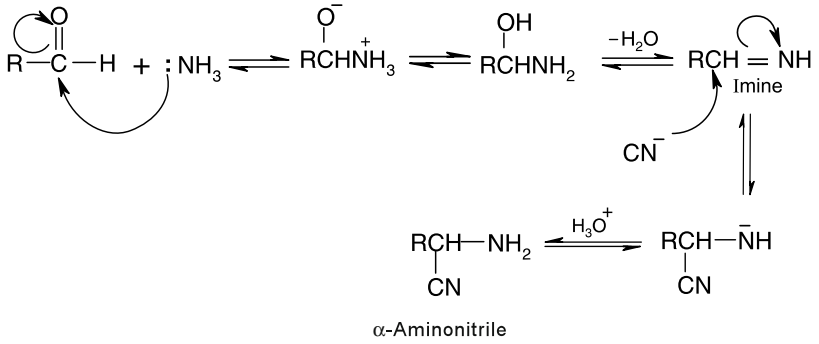


For example, acetaldehyde reacts with a mixture of HCN and ammonia to give 2-aminopropanenitrile, which on hydrolysis gives alanine.

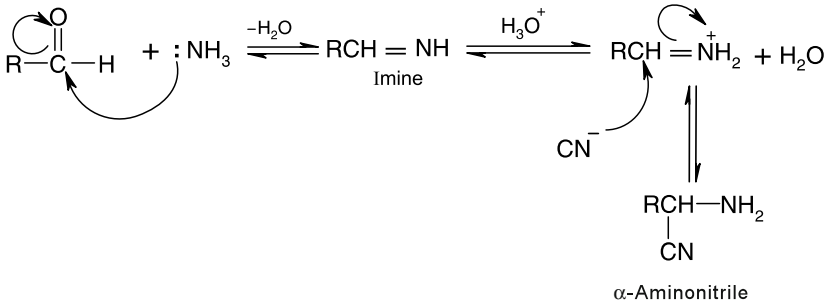


Glycine, leucine, isoleucine, valine, serine, etc., can also be prepared using Strecker synthesis.

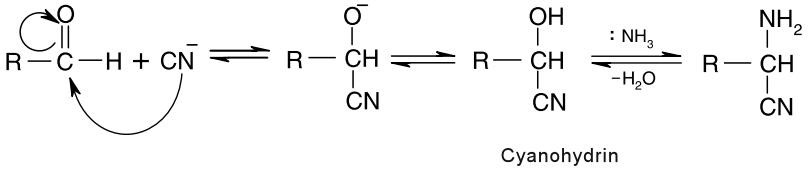
As far as the mechanism of the formation of aminonitrile is concerned, there are two possibilities. In one of the mechanisms, the first step involves formation of imines from the aldehyde and ammonia which is followed by addition of HCN to yield the aminonitrile.



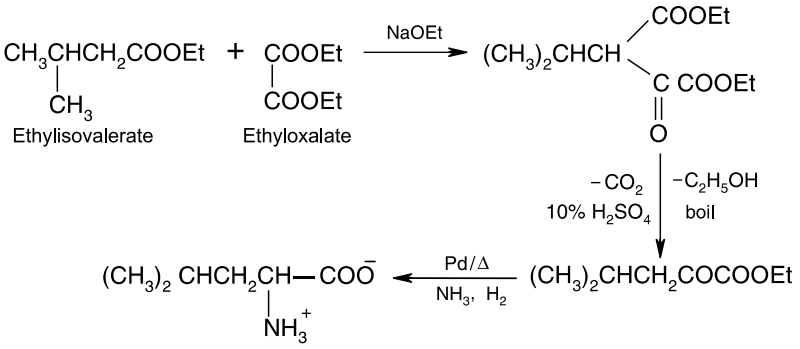
Alternatively, the cyanide ion may react with the conjugate acid of the imine to give the nitrile.



In the second mechanism, the cyanohydrin may be produced first followed by a nucleophilic substitution.

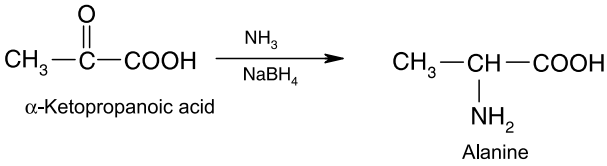


4. **Reductive amination**: α-amino acids can also be prepared by the reduction of aldehydes and ketones in the presence of ammonia. The method is called reductive amination. The reduction is carried out using a catalyst or sodium cyanohydridoborate, $NaBH_3CN$. For example, leucine can be prepared in the following manner starting from ethylisovalerate and ethyloxalate.

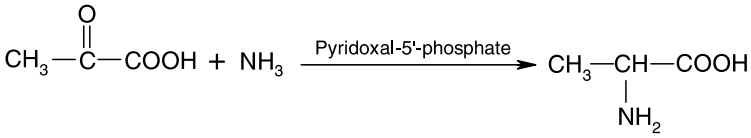


Other α-amino acids prepared by this method are alanine, glutamic acid, etc.

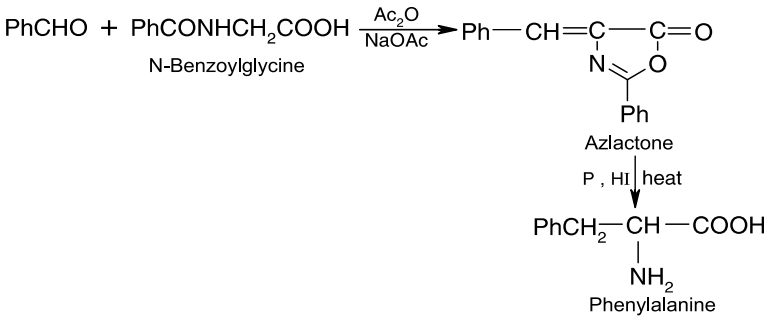
5. **Biosynthetic amination of α-keto acid**: α-keto acids may also be reductively aminated using ammonia and a reducing agent. This pathway is quite similar to the biosynthetic pathway.



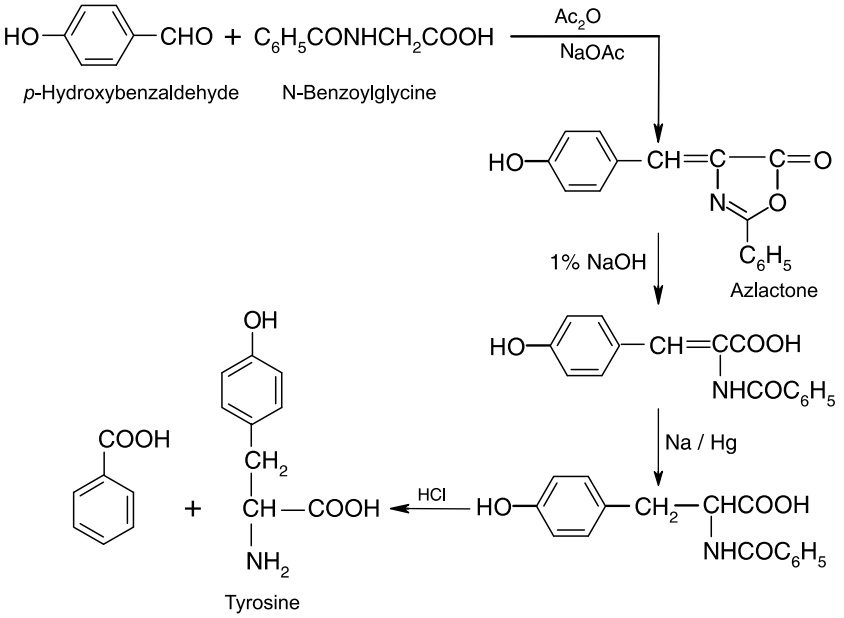
The biosynthetic pathway is as follows:



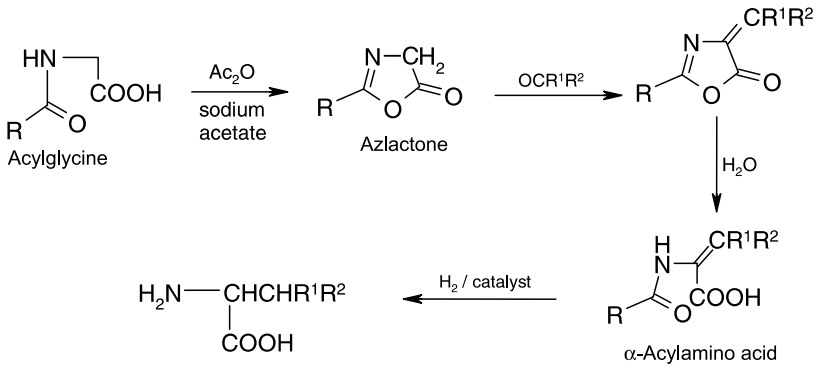
6. **Reduction of azlactones**: Azlactones, prepared by condensation of aromatic aldehydes with N-acyl derivatives of glycine in the presence of acetic anhydride and sodium acetate, can be reduced by phosphorus and HI or other reducing agents to produce amino acids also increasing the chain by two, in the process.



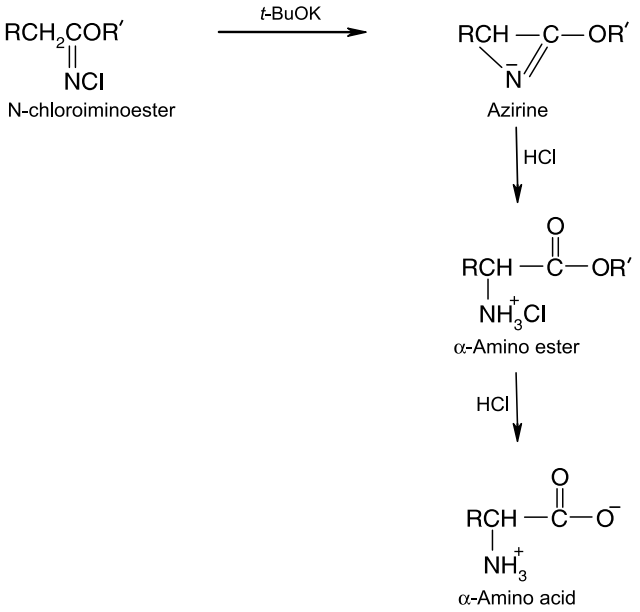
Alternatively, the azlactone is warmed with 1% NaOH to open the ring. The resulting product is reduced with sodium amalgam followed by hydrolysis to give the α-amino acid. This method is good for the preparation of aromatic amino acids like phenylalanine, tyrosine, etc. For example,



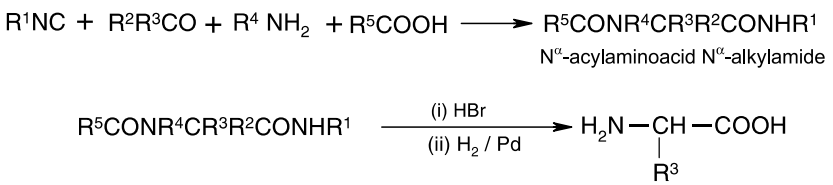
The azlactones may also be formed by intramolecular condensation of acylglycines in the presence of acetic anhydride. The reaction of azlactones with carbonyl compounds followed by hydrolysis gives the unsaturated α-acylamino acids, which on reduction yield amino acid.



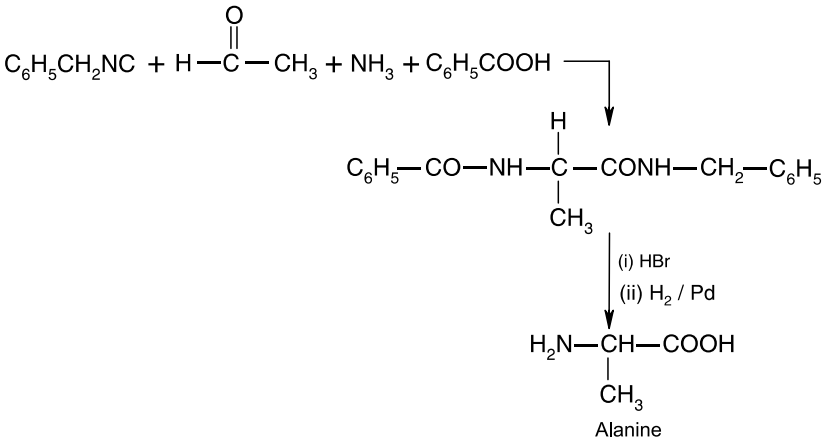
7. **By rearrangement of N-haloiminoesters**: N-haloiminoesters undergo Neber type of rearrangement in the presence of a base to first form an azirine intermediate and then α-amino orthoesters or α-amino esters. Both of these hydrolyse to give α-amino acids. The reactions can be given as under.



8. **Ugi 'Four-component condensation' method or 4CC synthesis**: In this relatively recent method of synthesising α-amino acids, four components—an aldehyde/ketone, a primary amine, an isocyanide and a carboxylic acid—'condense' to form N^{α}-acylamino acid N-alkyl amide. The groups attached to the amino and the carboxyl groups are then removed to get the α-amino acid.

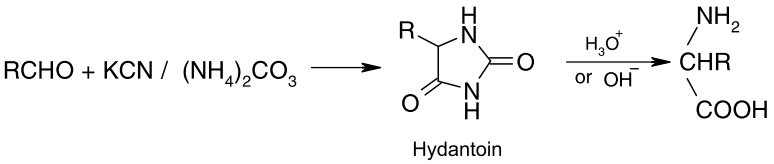


For example, alanine can be synthesised as follows:

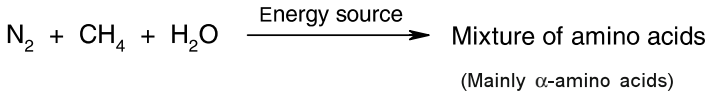


This method though proves to be quite advantageous in preparing certain amino acids, it has not been able to replace the conventional methods of synthesising α-amino acids.

9. **Bucherer–Berg's synthesis**: In this method, α-amino acids can be prepared by the hydrolysis of a hydantoin obtained from the reaction of an aldehyde, with KCN and $(NH_4)_2CO_3$ as given below.



10. **Miller–Urey method**: In this method, atmospheric components like nitrogen, methane and water combine in the presence of a high energy source to give a mixture of amino acids, most of them being α-amino acids.

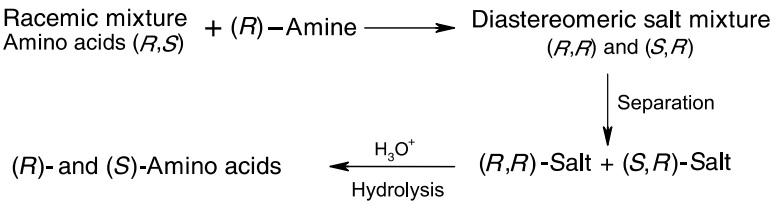


However, the yields are quite low, and it is not possible to prepare a desired amino acid preferentially.

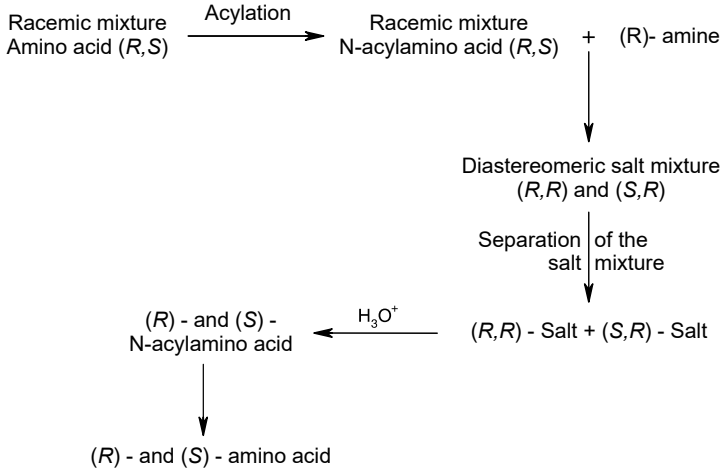
1.8.1 Enantiomeric Resolution of α-Amino Acids

All the amino acids, except glycine, synthesised by the above methods are obtained as racemic mixtures. It is necessary to resolve these mixtures in order to get a pure D- or L-enantiomer. The methods used for resolution are discussed below.

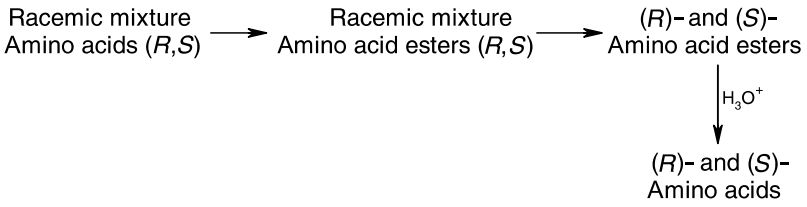
(a) **Amine salt formation**: In this method, the racemic mixture is mixed with an enantiomer (either *R* or *S*) of a naturally occurring optically active amine like strychnine or brucine. This leads to the formation of a diastereomeric pair of salts, which can be separated on the basis of differences in their physical properties like solubility (crystallisation), melting point, etc. The separated diastereomers are identified spectroscopically using circular dichroism, etc. These separated salts on acidification precipitate respective amino acids. The separation with *R*-amine can be represented as follows.



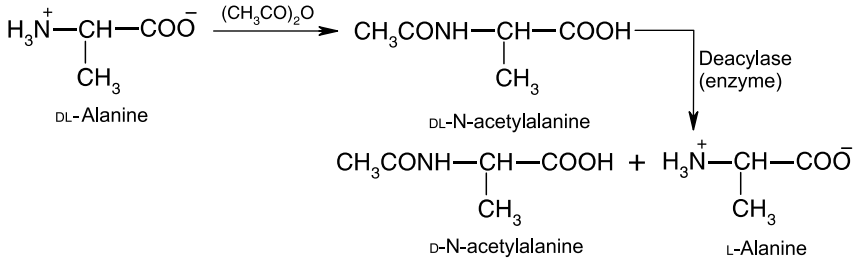
In place of direct salt formation, the racemic amino acids may first be converted into their N-acyl derivatives and then separated by salt formation with an optically active base.



(b) **Ester formation**: In a similar method, the racemic mixture of amino acids may be converted into diastereomeric esters which are then separated by crystallisation and the amino acids recovered by hydrolysis.



(c) **Enzymic resolution of amino acids**: The biological catalysts, i.e. the enzymes can be used to resolve a racemic mixture of amino acids. Certain enzymes (called *deacylases*) obtained from living organisms can selectively catalyse the hydrolysis of one of the enantiomeric N-acylamino acids. For example, the enzyme obtained from hog kidney cleaves the acyl group from the L-enantiomer, while the D-enantiomer remains unaffected. Thus, to resolve a racemic mixture of amino acids, it is converted into an N-acyl derivative, and then, the mixture is hydrolysed with the help of the deacylase enzyme.



The liberated amino acid is then precipitated from ethanol, while the N-acyl derivative remains in solution.

(d) **Chiral ligand exchange (LE) liquid chromatography (LC)**: It is a common chromatographic method employing chiral stationary phases for the separation of enantiomers. The chiral discrimination is attributed to the exchange of one stationary selector ligand in the bidentate complex on the stationary phase by an anylate ligand forming a ternary mixed complexes. The complex formation may be represented as shown in Fig. 1.15.

The differential stability of the mixed copper complexes with the L- and D-enantiomers of the amino acids leads to the differences in their retention times and hence the separation. A trace of the chromatograph showing the separation of a pair of enantiomers is given in Fig. 1.16.

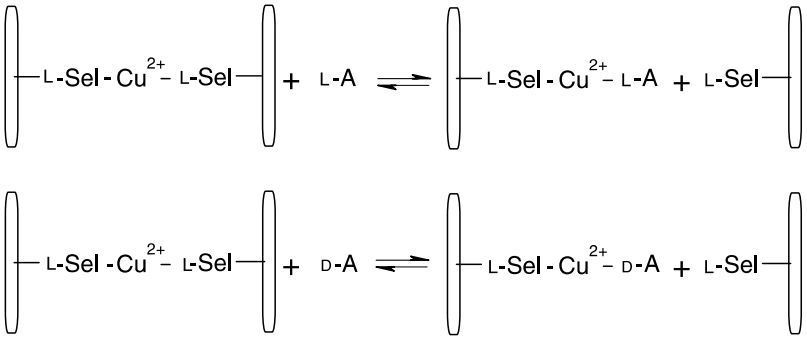


Fig. 1.15 Schematic representation of chiral ligand exchange liquid chromatography. (Sel represents the chiral selector ligand)

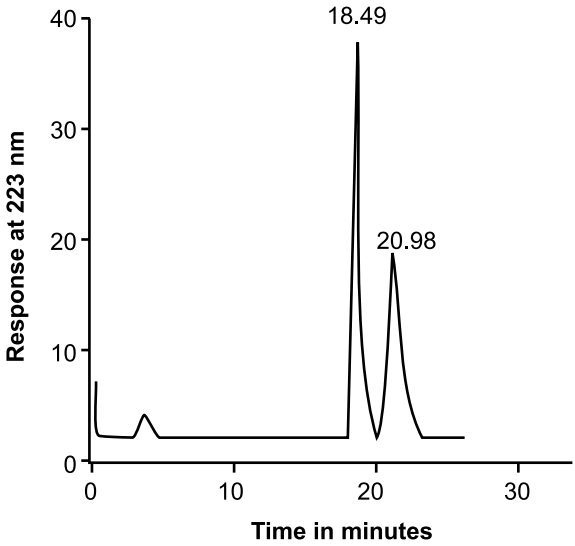


Fig. 1.16 Representative chiral ligand exchange liquid chromatograph

1.8.2 Asymmetric Synthesis of α-Amino Acids

As mentioned earlier, the conventional methods of synthesising α-amino acids give racemic mixtures. Though these can be separated by the procedures outlined above, an ideal situation would be the one where we can prepare optically pure α-amino acids. Nature has evolved selective methods for synthesising optically pure amino acids and also of selectively distinguishing between the enantiomers. In nature, enzymes—the biocatalysts aid in the synthesis of optically pure amino acids from prochiral precursors following a process called **asymmetric synthesis**. It has always been a challenge for a synthetic organic chemist to emulate nature in this respect. Asymmetric synthesis refers to a situation wherein in a reaction sequence at least one

reaction is **stereoselective**, intended to preferentially or exclusively yield a particular configuration of the amino acid. In principle, any of the reactions outlined above for the synthesis of α-amino acids can be used for the purpose.

A number of strategies have been successfully worked out to obtain optically pure α-amino acids. These can be broadly put into two categories.

(a) Using chiral catalyst or specific metal ion complexation
(b) Asymmetric induction using chiral auxiliaries.

Using Chiral Catalyst or Specific Metal Ion Complexation

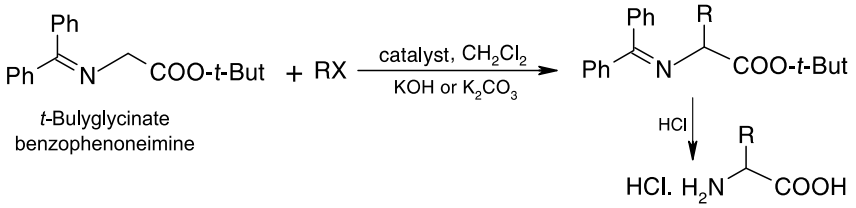
A number of synthetic procedures belonging to this class have been evolved. Some of these are detailed below.

(i) **O'Donnell's phase-transfer catalytic alkylation method**: In this method, pseudoenantiomeric quaternary amines derived from the *Cinchona* alkaloid are utilised for the phase-transfer catalytic alkylation of *t*-butylglycinatebenzophenoneimines with alkyl halides. The products obtained can be deprotected with acid to generate α-amino acids in appreciable enantiomeric excess. **Enantiomeric excess** (ee) is defined as given below.

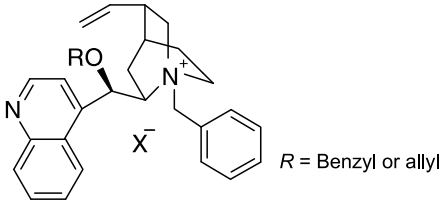
$$\%ee = \frac{R\text{-}enantiomer - S\text{-}enantiomer}{R\text{-}enantiomer + S\text{-}enantiomer} \times 100$$

A reaction yielding a mixture of 95% of *R*-isomer and 5% of *S*-isomer is said to have a 90% ee.

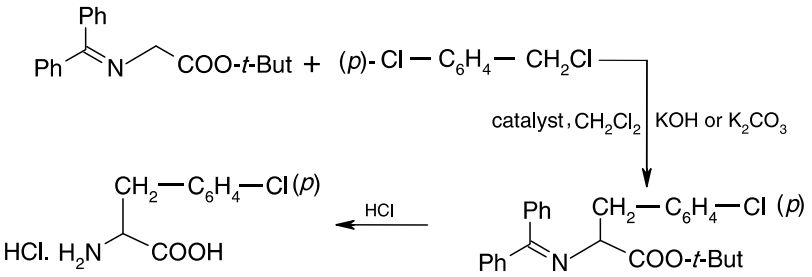
$$\frac{95-5}{95+5} \times 100 \Rightarrow \frac{90}{100} \times 100 = 90\%\text{ee}$$



In a specific example, the catalyst used in the reaction mentioned has the following structure.

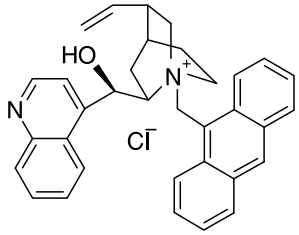


The phase-transfer alkylation of *t*-butylglycinatebenzophenoneimine using 4-chlorobenzylchloride as the alkyl halide yielded 4-chloro-L- phenylalanine in 99% enantiomeric excess.

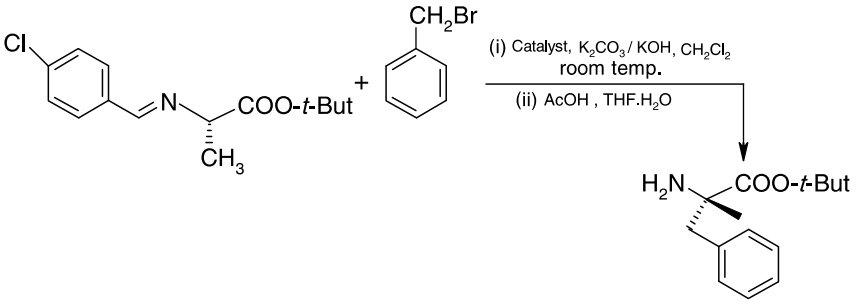


4-Chloro-L-phenylalanine hydrochloride

Minor modifications of the *Cinchona* alkaloid-based quaternary amine phase-transfer catalyst can be used to achieve major enhancements in the asymmetric induction. In particular, use of N-9-anthraceny-lmethyl salts, instead of the N-benzylated adducts used above, generally afforded enantioselectivities greater than 90%.

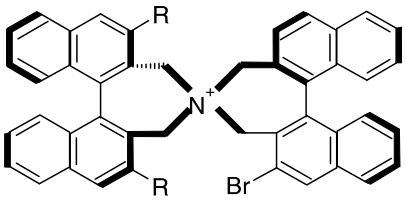


This methodology has been used by Lygo to synthesise α-methyl-α-alkylamino acids from the alanine-derived aldimines.



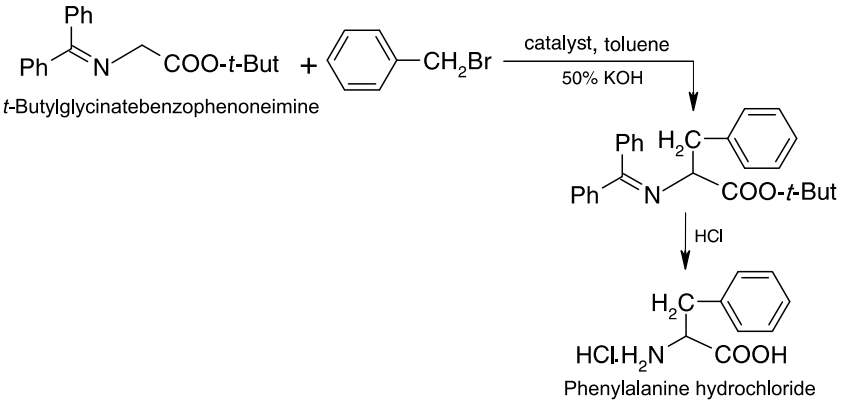
α-Methyl-α-alkylamino acid

(ii) **Non-*Cinchona* alkaloid-derived catalysts**: Following the success of the alkaloid-based catalysts, a number of novel phase-transfer catalysts have been rationally designed for the synthesis of α-amino acids. It has been found that readily assembled biaryl ammonium salts are highly effective catalysts for the phase-transfer alkylation of *t*-butylglycinate benzophenone imine. These afford high yields and enantiomeric excess greater than 90%. The steric bulk of the binaphthyl substituents is responsible for the reactivity and asymmetric induction of the chiral spiro ammonium catalysts.

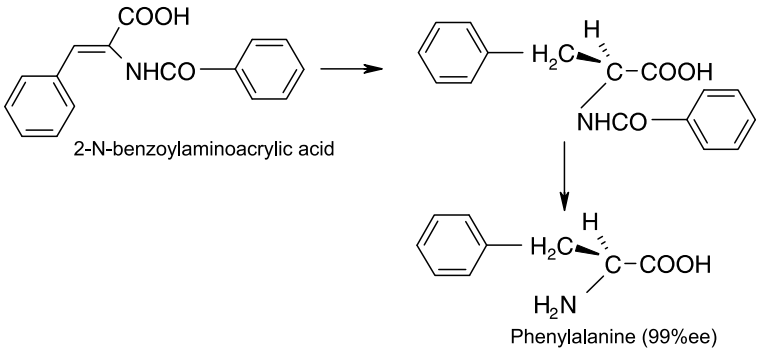


R= β-Naphthyl

The alkylation of *t*-butylglycinate benzophenone imine with benzyl bromide using the catalyst given above gives 96% enantiomeric excess of phenylalanine with a yield of 95%.



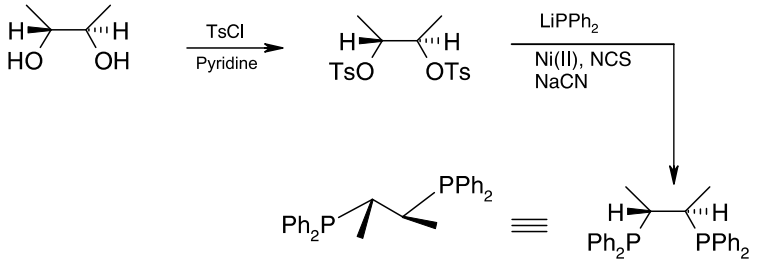
(iii) **Using Rhodium (I) complexes as catalyst**: In this method, a chiral catalyst is used for homogenous catalytic hydrogenation of suitable α-N-acylaminoacrylic acids or esters to obtain N-acylamino acids which on hydrolysis yield the α-amino acid in good enantiomeric excess.



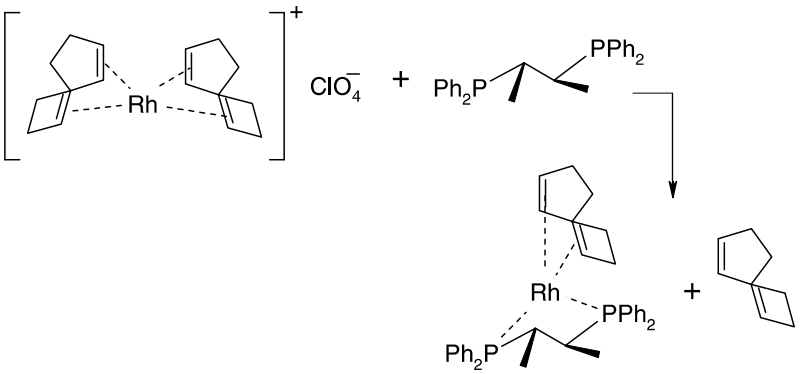
In this example, the starting α-N-acylaminoacrylic acid gives the amino acid, phenylalanine, in 99% enantiomeric excess when the reaction is performed in tetrahydrofuran as solvent.

The catalyst used in this reaction is a version of the Wilkinson's catalyst [$Rh(Ph_3P)_3Cl$] that is suitably modified to act as an asymmetric catalyst. A chiral ligand called (*S*,*S*)-chiraphos [(2*S*,3*S*)-bis(diphenylphosphino) butane] is used to chelate the rhodium metal ion to make the desired chiral complex. This chiral complex in the presence of hydrogen and a suitable solvent yields the active chiral hydrogenation catalyst. The following scheme represents the formation of the active catalyst.

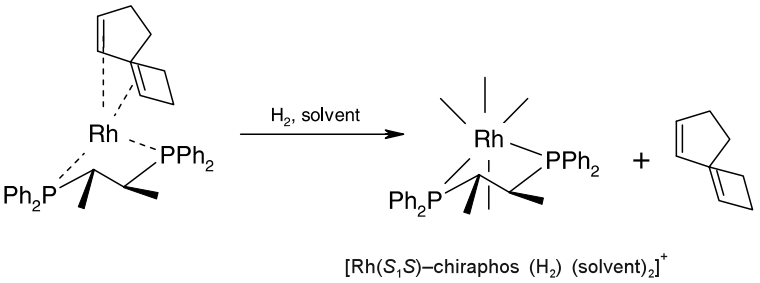
Synthesis of (*S*,*S*)-chiraphos, the Ligand



Formation of Chiral Complex



Formation of the Active Catalyst

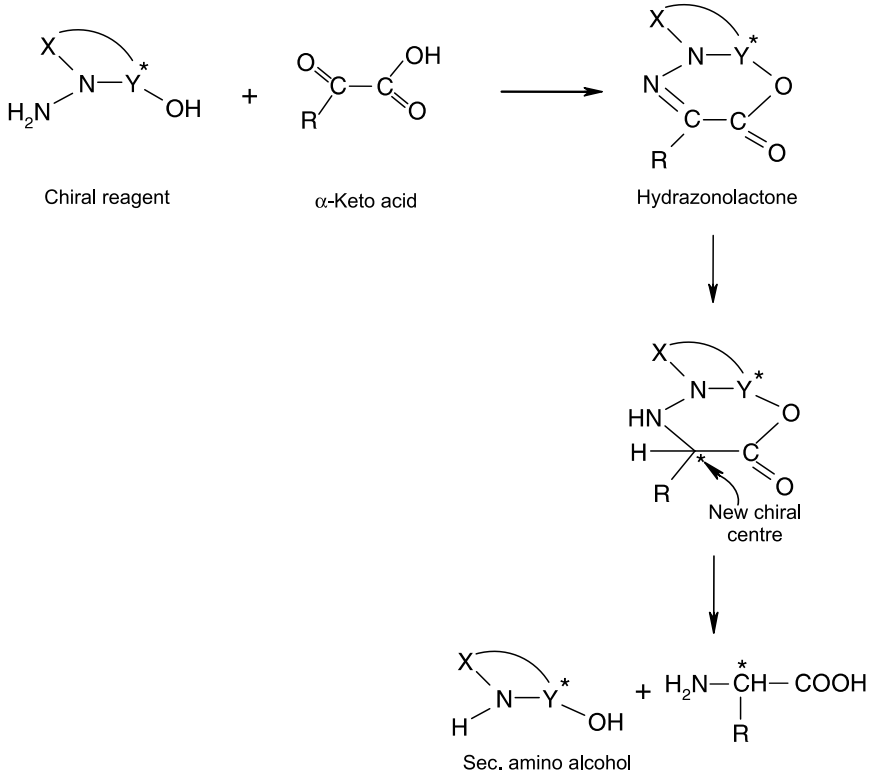


In this methodology, the whole molecular framework of the active catalyst is twisted into a single chiral conformation. The prochiral olefine coordinates with this complex in a specific orientation and undergoes asymmetric hydrogenation. This approach can also be used to synthesise selectively deuterated amino acids by using D_2 in place of H_2.

Asymmetric Induction Using Chiral Auxiliaries

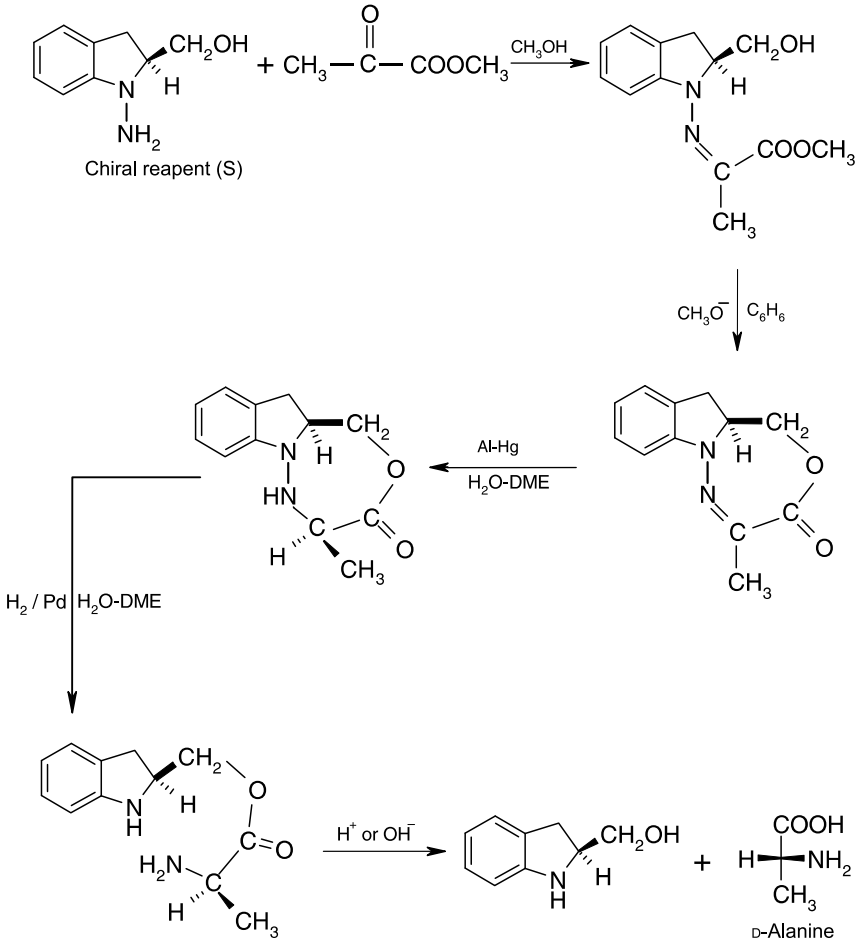
In this approach, optically pure substances are prepared from achiral molecules using chiral reagents containing an asymmetric centre. The chiral reagents have diastereotropic interactions with the reactant molecule and generate the desired asymmetric product. A number of strategies have been used for asymmetric induction. Some of these are given below.

(i) **Corey's methodology (asymmetric induction)**: According to this methodology, optically pure α-amino acids are prepared from the corresponding α-keto acids which act as their precursors. It uses a chiral reagent which when combined with the α-keto acid forms a hydrazonolactone ring. The strategy can be outlined as follows:

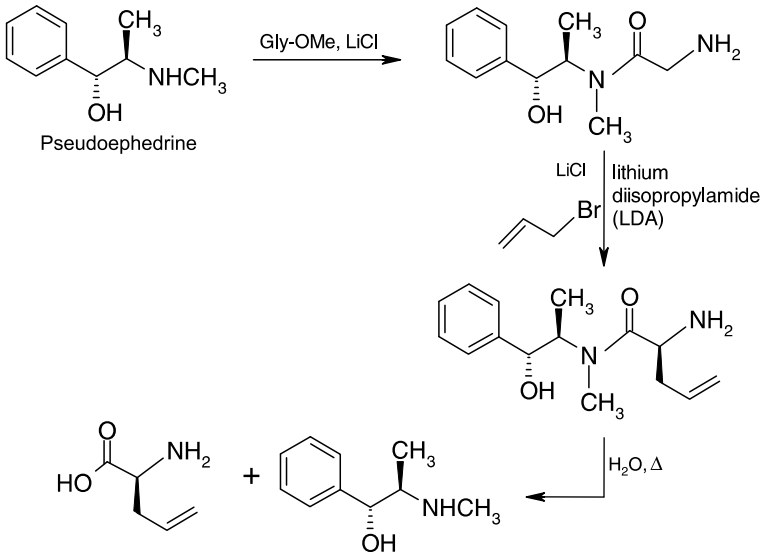


The stereospecific reduction of the N = C double bond produces the chiral carbon of the targeted amino acid, which is obtained by hydrogenolysis of the intermediate. The original chiral reagent is regenerated from the chiral secondary amino alcohol obtained.

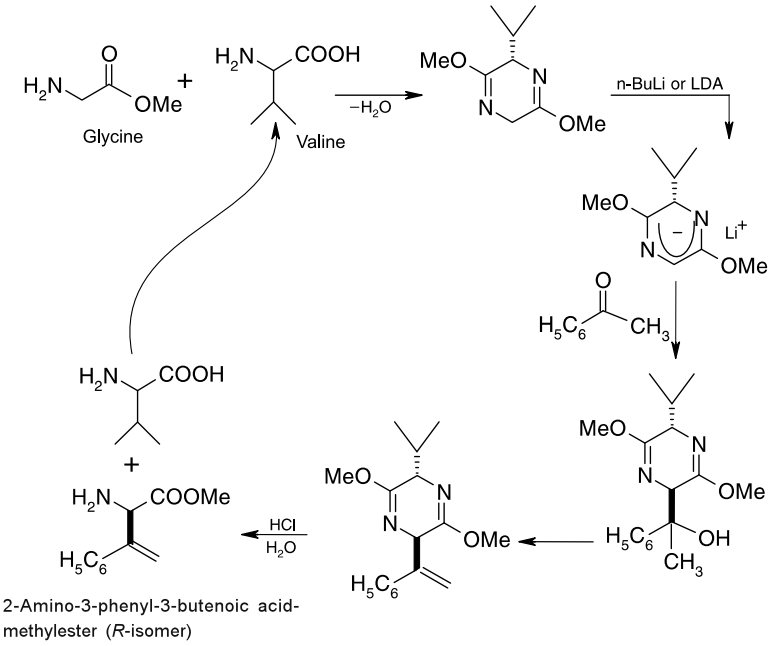
D-alanine can be synthesised with an optical purity of 80% from methylacetoformate ($CH_3COCOOCH_3$) using a chiral reagent with a bicyclic indoline structure.



(ii) **Using pseudoephedrine as chiral auxiliary**: In this method readily available and inexpensive chiral auxiliary, pseudoephedrine reacts with glycine methyl ester to give pseudoephedrine glycinamide methyl ester. The secondary amino group of pseudoephedrine forms an amide bond with the carboxyl group of glycine methyl ester. The alkylation of this substrate with a wide variety of electrophiles proceeds with excellent diastereoselectivity with good yields.



(iii) **Schollkopf's Bis-Lactim approach**: In this method, an amino acid, valine, is used as a chiral auxiliary. The auxiliary condenses with glycine to form a bis-lactim ether. The enolisation and electrophillic reaction with an aldehyde or ketone gives an intermediate with high level of distereoface selectivity.



The dehydration and hydrolysis of this intermediate give the amino acid with an enantiomeric purity of >95%, and the chiral auxiliary is regenerated.

1.9 Industrial Applications of α-Amino Acids

The proteinogenic amino acids find varied industrial applications, and to meet the demands estimated to be over two million tons, these α-amino acids are extensively produced. These are used in medicine to treat dietary deficiencies and in the food industry as antioxidants, artificial sweeteners, flavour enhancers and in cosmetics. In addition, these are required as cheap starting materials in chemical synthesis, and there is a reasonable demand of radiolabelled α-amino acids (to be used as tracers in animal and human studies) also. The principal application of α-amino acids is as flavour enhancer as it consumes about two thirds of all the global production. The annual demand for amino acids used in pharmaceutical products mainly for intravenous and enteral nutrition is to the tune of ~15,000 tons. Table 1.8 lists commercial and industrial applications of some α-amino acids.

Glutamic acid is probably the best-known commercially produced amino acid which is sold as monosodium glutamate (MSG)—a flavour enhancer. MSG was discovered in 1908 by Dr. Kikunae Ikeda as a basic taste substance of kelp—a traditional seasoning in Japan. Glutamic acid was the first α-amino acid to be produced commercially. It was prepared by extraction from acid hydrolysates of wheat and soya proteins. This was followed by production of a number of amino acids by the same extraction procedure. In late 1950s, fermentation technology was established and used for the commercial production of MSG and marked the beginning of modern amino acid production. This technology was triggered by the discovery of glutamic acid in the exhausted medium of *Cornebacterium glutamicum*. Since

Table 1.8 Industrial applications of α-amino acids

Amino acids	Industrial applications
Cysteine, Tryptophan and Histidine	Antioxidants
Methionine, Lysine and Threonine	Nutritive additive in soya products
Phenylalanine and Aspartic acid	Constituents of artificial sweetener aspartame
Lysine	Nutritive additive used in breads
Glutamic acid	Meaty flavoured food additive, meat tenderiser
Glycine and Alanine	Flavour enhancer
Serine and Arginine	Cosmetics
Arginine, Leucine, Isoleucine, Proline Valine, Tryptophan and Tyrosine	Infusions
Arginine, Aspartic acid, Asparagine, Glutamine, Histidine, Methionine and Phenylalanine	Therapy

then fermentation technologies for a number of amino acids have been established. However, several amino acids like L-leucine, hydroxy-L-proline, L-tyrosine and L-cystine are still being manufactured by extraction in addition to fermentation and chemical synthesis.

Exercises

1. Suppose a new α-amino acid having a $–CH_2F$ group as its side chain is isolated and is named as beucine.

   (a) Write down the structures of the L- and D-forms of the amino acid beucine.
   (b) Assign a suitable class to the amino acid.
   (c) Write the three letter code and suggest a one letter code for beucine.
   (d) Comment on its optical activity.
   (e) Ascertain the absolute configuration (*R* or *S*) for the α- carbon atom.

2. The next higher homologue (with one more $–CH_2$ group in the side chain) of beucine (Q. 1) is synthesised. Give its structure and assign suitable configuration (in *R*/*S* system) to the asymmetric carbon atom.
3. Draw four stereoisomers of the amino acid threonine and assign *R*/*S* configurations to the asymmetric carbon atoms.
4. Compute pI values for the amino acids Asp, Val and His (Refer to Table 1.4 for the required pK_a values).
5. Amino acid, serine has a pI value of 5.68. If the carboxyl group has a pK_a value of 2.2, compute the acid dissociation constant for the α-amino group.
6. Using Tables 1.1 and 1.4, write down the structures of the amino acid tyrosine at a pH of

   (a) 2.0
   (b) 6.6
   (c) 11.0

7. The isoelectric point of glutamine is considerably higher than that of glutamic acid. Explain.
8. A mixture of amino acids contains Asp, Asn, Glu, Gly, His, Lys and Val.

   (a) Can this mixture be separated by electrophoresis?
   (b) Predict the order in which these would be obtained on the gel that is buffered at a pH of 6.0.

9. A mixture of amino acids contains glutamic acid, histidine and leucine. Suggest a suitable pH at which electrophoresis of the mixture would give a good separation.
10. Suggest suitable spectroscopic methods that can be used to differentiate between the amino acids alanine and phenylalanine.

11. What differences would you observe in the UV, NMR and CD spectra of the two amino acids mentioned in question 10.
12. Using Tables 1.1 and 1.4, identify

    (a) The most acidic amino acid
    (b) The most basic amino acid.

    Justify the choices made.
13. A given sample of alanine produced 1.4 g of N_2 in the Von Slyke determination.

    Compute the amount of alanine (M_m(Ala) = 89.1 g mol^{-1}) in the given sample.
14. How would you prepare the following α-amino acids using the methods indicated against them

(a) Valine	Strecker synthesis
(b) Leucine	Acetamido malonate synthesis
(c) Isoleucine	Reductive amination
(d) Tyrosine	Reduction of azlactones
(e) D-Alanine	Chiral auxiliary
(f) L-Alanine	4CC synthesis
(g) L-Phenylalanine	Phthalamido malonic ester synthesis.

15. How would you prepare the following α-amino acids using the starting materials indicated against them

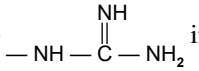
(a) Lysine	1,4-Dibromobutane
(b) Leucine	Isobutyl alcohol
(c) Glutamic acid	α-Ketoglutaric acid
(d) Proline	Adipic acid.

16. Give the products formed in the reactions of valine with the following:

    (a) Ethanol; H^+
    (b) CH_3COCl; pyridine followed by hydrolysis
    (c) Heating at high temperature
    (d) Phenylisocyanate followed by HCl.

17. A synthetic peptide is represented as 'PEPTIDE' using one letter code for amino acids.

    (a) Will it move towards anode or cathode during electrophoresis?
    (b) Will it absorb UV radiation above 240 nm?

18. Arginine with a guanidine group  in the side chain is the most basic amino acid. Explain.
19. The p*K*a values for β-alanine are 3.55 and 10.24. Compare these values with that of alanine (cf. Table 1.4), and explain the differences.

20. In strongly acidic solution, alanine exists as a mono cation with two potential acidic groups (α-NH_3^+ and –COOH).

    (a) Which of these two groups is more acidic?
    (b) How would the structure of alanine change on adding a base?