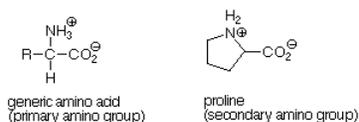


## 1.21: Amino Acids and Peptides

### Amino Acids and Side Chain Structure

Last time we finished our examination of amines. Now we'll look at what happens when a carboxylic acid functional group and an amine functional group are in the same molecule. Our focus will be on the alpha amino acids, those in which the amino group is bonded to the alpha carbon -- the one next to the carbonyl group -- of the carboxylic acid. These are the basic building blocks of proteins and are the most important type of amino acid. While there are many other ways to link an amino group and a carboxylic acid group in a single molecule, we will concern ourselves only with the alpha amino acids.

There are 20 alpha amino acids commonly found in proteins. They are listed in Table 18.1 (p 503) in Brown. When the structures of these molecules are examined, it becomes clear that they share the common structural unit  $\text{RCH}(\text{NH}_3^+)\text{CO}_2^-$  in which R can be either hydrogen (the amino acid is glycine) or one of 19 other possibilities. The one exception to this pattern is proline, in which the R group makes up part of a ring which also includes the amino group and the alpha carbon atom. Since the amino group in proline is involved in two carbon-nitrogen bonds, it is a secondary amino group.



The table is further divided into groups according to the structure of the R group. (The R group is often called the "side chain.") If the R group is made up of only carbon and hydrogen (no heteroatoms), the side chain is regarded as non-polar since there is very little polarity associated with carbon-carbon and carbon-hydrogen bonds. These side chains are hydrophobic (water avoiding) in much the same way that the long hydrocarbon tail of a soap or detergent is hydrophobic. This will be important when we consider how the characteristics of proteins depend upon their folding patterns in an aqueous environment. There are heteroatoms in methionine (sulfur) and tryptophan (nitrogen) but the overall behavior of these amino acids suggests that these heteroatoms contribute very little polarity to the side chain. Side chains which contain more polar functional groups such as amide, alcohol and thiol provide locations for a polar water molecule to hydrogen bond. They are thus somewhat hydrophilic, like the OH groups in a sugar. These side chains are important in making a protein sufficiently water soluble to operate effectively inside a cell.

In two cases (aspartic acid and glutamic acid) the side chain includes a carboxylic acid group in addition to the one next to the amino group. These groups are ionized (present as the carboxylate anion) when the pH is near neutral ( $\text{pH} \sim 7$ ). (We'll take up the acid-base behavior of amino acids shortly.)

Similarly, there are three amino acids whose side chains include an amino group. These amino groups are also ionized (present as the ammonium ion) at neutral pH. The ionized groups are quite polar, and like the ionized ends of soaps or detergents, they make the side chain quite hydrophilic.

### Acid Base Chemistry

At neutral pH (around 7, the typical pH of most body fluids and the pH at which biochemical reactions usually happen) the amino groups in amino acids are protonated to make ammonium ions and the carboxylic acids are ionized to their conjugate bases (carboxylate ions). One way to look at this is to look at a water solution at  $\text{pH} = 7$  as a large reservoir of acid whose  $\text{pK}_a$  is maintained at 7. If an acid with a  $\text{pK}_a$  lower than 7 (like a carboxylic acid,  $\text{pK}_a \sim 5$ ) is dissolved in such a solution, it is the stronger acid and will transfer a proton to the solution and become the carboxylate ion. Thus when the pH is maintained at 7, carboxylic acids are ionized.

In the same way, when an amine (typical ammonium ion  $\text{pK}_a \sim 10$ ) is dissolved in water which is held at  $\text{pH} = 7$ , the water is the stronger acid so the amine is protonated to make the weaker acid. Amines when held at  $\text{pH} = 7$  are protonated to make ammonium ions. Practically, holding the pH at 7 means that the solution is buffered by the inclusion of weak acids and weak bases in sufficient concentration so that the transfer of a few protons does not materially change the  $\text{H}^+$  concentration.

We can use this idea at any pH. For example, if an amino acid is dissolved in water which is held at  $\text{pH} = 2$ , the solution is a stronger acid than the carboxylic acid which would be formed by transferring a proton to the carboxylate ion. The carboxylic acid ( $\text{pK}_a \sim 5$ ) is formed as the weaker acid. Such a molecule would have only a positive charge from the ammonium ion. Similarly, in basic solution ( $\text{pH} > 11$ ) the solution is a weaker acid than the ammonium ion, so the ammonium ion transfers a proton to the solution and becomes the amino group.

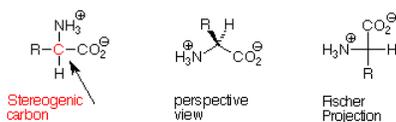
Since there are small variations in the specific  $pK_a$  values of amino and carboxylic acid groups in amino acids, the exact pH at which the predominant species is the zwitterion (the molecule with one positive ammonium ion and one negative carboxylate ion) varies somewhat. This pH is called the isoelectric point (pI) because it is the pH at which the amino acid is as likely to be attracted to a positive electrode as to a negative one. The pI values for the common amino acids are given in Table 18.2 (p 506 in Brown).

Notice that the acidic amino acids have low pI numbers. This makes sense because it will take a fairly strongly acidic solution to ensure that one of the carboxylate ions is protonated. Similarly, for basic amino acids the pI values are higher since it will take a fairly basic solution to ensure that one of the ammonium ions has lost a proton and the positive charge.

## Stereochemistry

For all of the amino acids except glycine, the alpha carbon atom is a stereogenic carbon atom (four different groups attached). In two cases there is also another stereogenic carbon atom in the molecule. Only one of the two possible enantiomers is found in nature in the cases of the amino acids which include stereogenic carbon atoms. In all these cases the absolute configuration of the alpha stereogenic carbon is S.

It became possible to determine absolute configurations well after the stereochemical relationships between amino acids and sugars had been worked out. That work showed that if we orient a Fischer projection of an amino acid with the carboxylate ion group at the top and the R group at the bottom, we find that the ammonium ion is pointed to the left. For this reason the amino acids are considered to have the L configuration (opposite to the D configurations assigned to common sugars). You may wish to verify that that an L-amino acid is also an S amino acid.

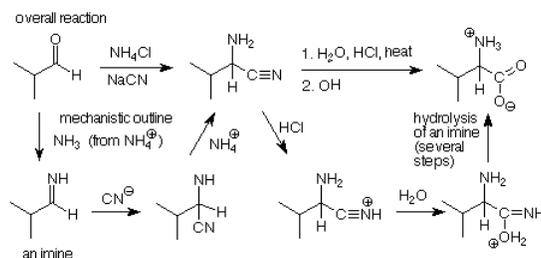


Amino acids are produced in living systems by biochemical pathways which involve multiple enzymes. The enzymes are proteins, themselves made up of L-amino acids so they provide a chiral environment in which only one of the two enantiomers is formed. Laboratory synthesis of amino acids typically does not involve a chiral environment, so equal amounts of the L- and D-amino acids are formed in typical laboratory syntheses. A mixture of equal amounts of enantiomers is called a *racemic* mixture.

## Synthesis

Laboratory syntheses of amino acids are usually related to syntheses of amines and/or carboxylic acids. We'll take a look at one such synthesis, the Strecker synthesis. We won't look at its mechanism in detail, but we will look for similarities with reactions we've seen before.

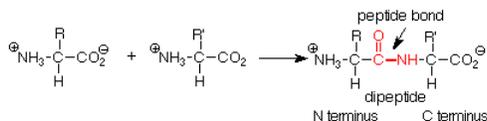
The reaction begins with imine formation from an aldehyde and ammonia. The acid catalysis required for this comes from ammonium chloride, a weak acid. An addition of hydrogen cyanide to the imine follows. This is analogous to the additions of nucleophiles to an aldehyde or ketone which we studied earlier. In this instance, the cyanide ion serves as the nucleophile.



The amino nitrile which results from these steps is purified and treated with aqueous HCl, followed by  $\text{OH}^-$ . This converts the nitrile to a carboxylate salt. We can put this reaction in context by thinking of the C-N triple bond as being much like a carbonyl group. That suggests that the electrophilic  $\text{H}^+$  attacks the nitrogen, which is followed by a nucleophilic attack of water on the nitrile carbon. A C=N double bond remains, and its reaction with water is the reverse of imine formation. The outcome is that the C=N double bond is hydrolyzed to a C=O double bond. Finally, neutralization with just enough base gives us the amino acid zwitterion.

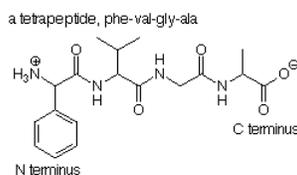
## Peptides and the Peptide Bond

Now let's turn our attention to the way in which amino acids are linked together to form proteins. The key structural element here is the peptide bond. This is an amide linkage which joins the ammonium group of one amino acid to the carboxylate group of another by a new covalent bond. The  $O^-$  of the carboxylate is lost along with two  $H^+$  ions from the ammonium group to form water. This is quite analogous to the formation of an amide by heating a carboxylic acid and an amine. The specific reaction conditions and processes required to do this may be (as we will see) quite sophisticated, but it helps to remember that what is being done is the joining of a carboxylate carbon and an ammonium nitrogen by a new C-N (peptide) bond.



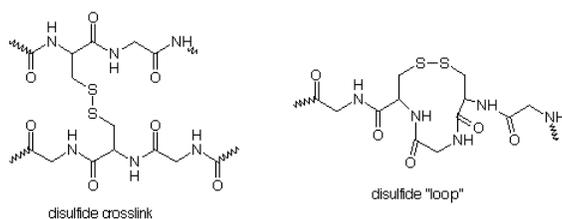
The new compound formed in this way is called a peptide. Our example is a dipeptide, formed from two amino acids. If a third amino acid is connected to the dipeptide by forming a new peptide bond at either the ammonium group or the carboxylate group of the dipeptide, we obtain a tripeptide, and so on. Polypeptides may have many amino acids. Polypeptides with more than 100 amino acids are considered to be proteins.

Since the amino acid whose carboxylic acid group participated in the formation of the peptide bond still has an ammonium group which contains a nitrogen atom, it is called the N terminus of the peptide. The N terminus is conventionally written to the left. Correspondingly, the amino acid which still has a free carboxylate group is called the C terminus and is written to the right. When the order of amino acids in a peptide is written out, it is conventional to write it left to right from the N terminus to the C terminus. The complete order of amino acids in a protein is called its sequence and is conveniently expressed by using the abbreviated names of the amino acids read from N to C terminus.



The sequence is held together by peptide bonds. As a part of an amide functional group, these bonds are difficult to break, so the sequence of a protein is quite stable. While there are many possible ways that a protein chain could be folded, the particular folding pattern adopted by the protein is completely determined by its sequence.

In many cases the folding pattern is "locked in" by disulfide links. As we discussed when we were studying thiols, the presence of SH groups along a protein chain provides an opportunity for crosslinking between chains or the formation of loops within a chain. These disulfide bridges are important in holding the protein chain in a specific folding pattern.



A glance at Table 18.1 tells us that the SH groups necessary to make the disulfide links are found in the amino acid cysteine. Proteins which are stiff and used primarily for structural purposes (keratin in hair, skin and feathers, for example) usually have many disulfide links and thus have high contents of cysteine.

This page titled [1.21: Amino Acids and Peptides](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Kirk McMichael](#).