= Amino acid =

Amino acids are biologically important organic compounds containing amine (-NH2) and carboxylic acid (-COOH) functional groups , usually along with a side @-@ chain (R group) specific to each amino acid . The key elements of an amino acid are carbon , hydrogen , oxygen , and nitrogen , though other elements are found in the side @-@ chains of certain amino acids . About 500 amino acids are known (though only 20 appear in the genetic code) and can be classified in many ways . They can be classified according to the core structural functional groups ' locations as alpha- (?-) , beta- (?-) , gamma- (?-) or delta- (?-) amino acids ; other categories relate to polarity , pH level , and side @-@ chain group type (aliphatic , acyclic , aromatic , containing hydroxyl or sulfur , etc .) . In the form of proteins , amino acids comprise the second @-@ largest component (water is the largest) of human muscles , cells and other tissues . Outside proteins , amino acids perform critical roles in processes such as neurotransmitter transport and biosynthesis .

In biochemistry , amino acids having both the amine and the carboxylic acid groups attached to the first (alpha-) carbon atom have particular importance . They are known as 2- , alpha- , or ? @-@ amino acids (generic formula H2NCHRCOOH in most cases , where R is an organic substituent known as a " side @-@ chain ") ; often the term " amino acid " is used to refer specifically to these . They include the 23 proteinogenic (" protein @-@ building ") amino acids , which combine into peptide chains (" polypeptides ") to form the building @-@ blocks of a vast array of proteins . These are all L @-@ stereoisomers (" left @-@ handed " isomers) , although a few D @-@ amino acids (" right @-@ handed ") occur in bacterial envelopes , as a neuromodulator (D @-@ serine) , and in some antibiotics . Twenty of the proteinogenic amino acids are encoded directly by triplet codons in the genetic code and are known as " standard " amino acids . The other three (" non @-@ standard " or " non @-@ canonical ") are selenocysteine (present in many noneukaryotes as well as most eukaryotes , but not coded directly by DNA) , pyrrolysine (found only in some archea and one bacterium) and N @-@ formylmethionine (which is often the initial amino acid of proteins in bacteria , mitochondria , and chloroplasts) . Pyrrolysine and selenocysteine are encoded via variant codons ; for example , selenocysteine is encoded by stop codon and SECIS element . Codon ? tRNA combinations not found in nature can also be used to " expand " the genetic code and create novel proteins known as alloproteins incorporating non @-@ proteinogenic amino acids .

Many important proteinogenic and non @-@ proteinogenic amino acids also play critical non @-@ protein roles within the body . For example , in the human brain , glutamate (standard glutamic acid) and gamma @-@ amino @-@ butyric acid (" GABA " , non @-@ standard gamma @-@ amino acid) are , respectively , the main excitatory and inhibitory neurotransmitters ; hydroxyproline (a major component of the connective tissue collagen) is synthesised from proline ; the standard amino acid glycine is used to synthesise porphyrins used in red blood cells ; and the non @-@ standard carnitine is used in lipid transport .

Nine proteinogenic amino acids are called " essential " for humans because they cannot be created from other compounds by the human body and so must be taken in as food . Others may be conditionally essential for certain ages or medical conditions . Essential amino acids may also differ between species .

Because of their biological significance , amino acids are important in nutrition and are commonly used in nutritional supplements , fertilizers , and food technology . Industrial uses include the production of drugs , biodegradable plastics , and chiral catalysts .

= = History = =

The first few amino acids were discovered in the early 19th century . In 1806 , French chemists Louis @-@ Nicolas Vauquelin and Pierre Jean Robiquet isolated a compound in asparagus that was subsequently named asparagine , the first amino acid to be discovered . Cystine was discovered in 1810 , although its monomer , cysteine , remained undiscovered until 1884 . Glycine and leucine were discovered in 1820 . The last of the 20 common amino acids to be discovered was

threonine in 1935 by William Cumming Rose , who also determined the essential amino acids and established the minimum daily requirements of all amino acids for optimal growth .

Usage of the term amino acid in the English language is from 1898 . Proteins were found to yield amino acids after enzymatic digestion or acid hydrolysis . In 1902 , Emil Fischer and Franz Hofmeister proposed that proteins are the result of the formation of bonds between the amino group of one amino acid with the carboxyl group of another , in a linear structure that Fischer termed " peptide " .

= = General structure = =

In the structure shown at the top of the page , R represents a side @-@ chain specific to each amino acid . The carbon atom next to the carboxyl group (which is therefore numbered 2 in the carbon chain starting from that functional group) is called the ? ? carbon . Amino acids containing an amino group bonded directly to the alpha carbon are referred to as alpha amino acids . These include amino acids such as proline which contain secondary amines , which used to be often referred to as " imino acids " .

= = = Isomerism = = =

The alpha amino acids are the most common form found in nature , but only when occurring in the L @-@ isomer . The alpha carbon is a chiral carbon atom , with the exception of glycine which has two indistinguishable hydrogen atoms on the alpha carbon . Therefore , all alpha amino acids but glycine can exist in either of two enantiomers , called L or D amino acids , which are mirror images of each other (see also Chirality) . While L @-@ amino acids represent all of the amino acids found in proteins during translation in the ribosome , D @-@ amino acids are found in some proteins produced by enzyme posttranslational modifications after translation and translocation to the endoplasmic reticulum , as in exotic sea @-@ dwelling organisms such as cone snails . They are also abundant components of the peptidoglycan cell walls of bacteria , and D @-@ serine may act as a neurotransmitter in the brain . D @-@ amino acids are used in racemic crystallography to create centrosymmetric crystals , which (depending on the protein) may allow for easier and more robust protein structure determination . The L and D convention for amino acid configuration refers not to the optical activity of the amino acid itself but rather to the optical activity of the isomer of glyceraldehyde from which that amino acid can , in theory , be synthesized (D @-@ glyceraldehyde is dextrorotatory ; L @-@ glyceraldehyde is levorotatory) . In alternative fashion , the (S) and (R) designators are used to indicate the absolute stereochemistry . Almost all of the amino acids in proteins are (S) at the ? carbon , with cysteine being (R) and glycine non @-@ chiral . Cysteine has its side @-@ chain in the same geometric position as the other amino acids , but the R / S terminology is reversed because of the higher atomic number of sulfur compared to the carboxyl oxygen gives the side @-@ chain a higher priority , whereas the atoms in most other side @-@ chains give them lower priority .

= = = Side chains = = =

In amino acids that have a carbon chain attached to the ? ? carbon (such as lysine , shown to the right) the carbons are labeled in order as ? , ? , ? , ? , and so on . In some amino acids , the amine group is attached to the ? or ? @-@ carbon , and these are therefore referred to as beta or gamma amino acids .

Amino acids are usually classified by the properties of their side @-@ chain into four groups . The side @-@ chain can make an amino acid a weak acid or a weak base , and a hydrophile if the side @-@ chain is polar or a hydrophobe if it is nonpolar . The chemical structures of the 22 standard amino acids , along with their chemical properties , are described more fully in the article on these proteinogenic amino acids .

The phrase " branched @-@ chain amino acids " or BCAA refers to the amino acids having

aliphatic side @-@ chains that are non @-@ linear ; these are leucine , isoleucine , and valine . Proline is the only proteinogenic amino acid whose side @-@ group links to the ? @-@ amino group and , thus , is also the only proteinogenic amino acid containing a secondary amine at this position . In chemical terms , proline is , therefore , an imino acid , since it lacks a primary amino group , although it is still classed as an amino acid in the current biochemical nomenclature , and may also be called an " N @-@ alkylated alpha @-@ amino acid " .

= = = Zwitterions = = =

The ? @-@ carboxylic acid group of amino acids is a weak acid , meaning that it releases a hydron (such as a proton) at moderate pH values . In other words , carboxylic acid groups (? CO2H) can be deprotonated to become negative carboxylates (? CO2 ?) . The negatively charged carboxylate ion predominates at pH values greater than the pKa of the carboxylic acid group (mean for the 20 common amino acids is about 2 @.@ 2 , see the table of amino acid structures above) . In a complementary fashion , the ? @-@ amine of amino acids is a weak base , meaning that it accepts a proton at moderate pH values . In other words , ? @-@ amino groups (NH2 ?) can be protonated to become positive ? @-@ ammonium groups (+ NH3 ?) . The positively charged ? @-@ ammonium group predominates at pH values less than the pKa of the ? @-@ ammonium group (mean for the 20 common ? @-@ amino acids is about 9 @.@ 4) .

Because all amino acids contain amine and carboxylic acid functional groups , they share amphiprotic properties . Below pH 2 @.@ 2 , the predominant form will have a neutral carboxylic acid group and a positive ? @-@ ammonium ion (net charge + 1) , and above pH 9 @.@ 4 , a negative carboxylate and neutral ? @-@ amino group (net charge ? 1) . But at pH between 2 @.@ 2 and 9 @.@ 4 , an amino acid usually contains both a negative carboxylate and a positive ? @-@ ammonium group , as shown in structure (2) on the right , so has net zero charge . This molecular state is known as a zwitterion , from the German Zwitter meaning hermaphrodite or hybrid . The fully neutral form (structure (1) on the right) is a very minor species in aqueous solution throughout the pH range (less than 1 part in 107) . Amino acids exist as zwitterions also in the solid phase , and crystallize with salt @-@ like properties unlike typical organic acids or amines .

= = = Isoelectric point = = =

The variation in titration curves when the amino acids are grouped by category can be seen here . With the exception of tyrosine , using titration to differentiate between hydrophobic amino acids is problematic .

At pH values between the two pKa values , the zwitterion predominates , but coexists in dynamic equilibrium with small amounts of net negative and net positive ions . At the exact midpoint between the two pKa values , the trace amount of net negative and trace of net positive ions exactly balance , so that average net charge of all forms present is zero . This pH is known as the isoelectric point pI , so pI

= ½ (pKa1 + pKa2) . The individual amino acids all have slightly different pKa values , so have different isoelectric points . For amino acids with charged side @-@ chains , the pKa of the side @-@ chain is involved . Thus for Asp , Glu with negative side @-@ chains , pI =