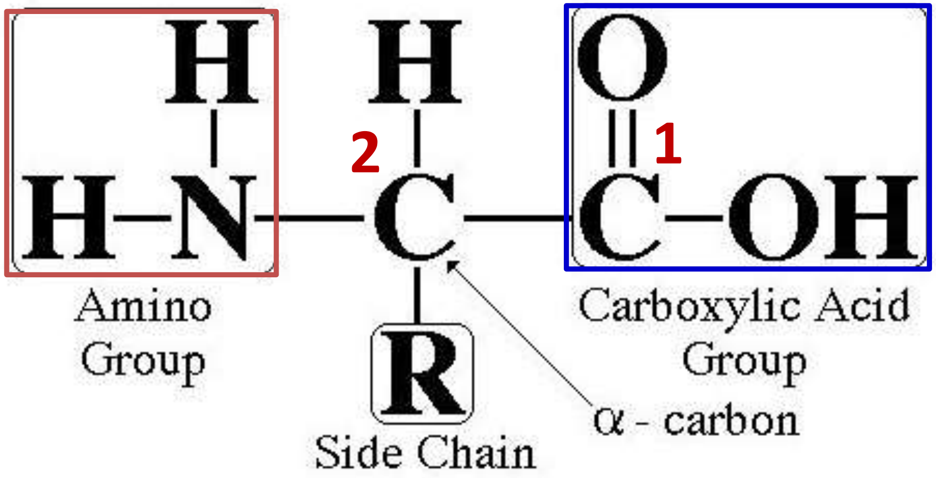
Amino Acids 1

Dr. Nesrin Mwafi
Biochemistry & Molecular Biology Department
Faculty of Medicine, Mutah University

Amino Acid Structure

- Amino acids are biologically important organic molecules that contain both carboxylic acid (-COOH) as well as amine (-NH_2) groups
- The side-chain also called "**R**" group is specific to each amino acid



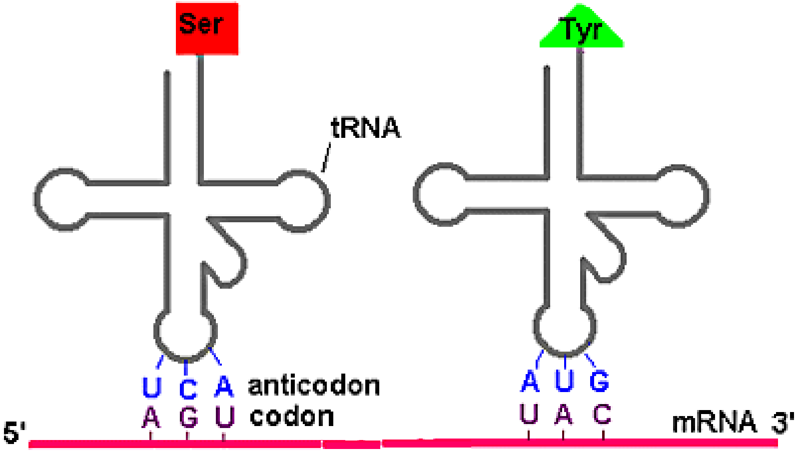
- Amino group is attached to α-carbon (C2)
- C, N, O and H are the key elements of amino acids

Biological significance of Amino Acids

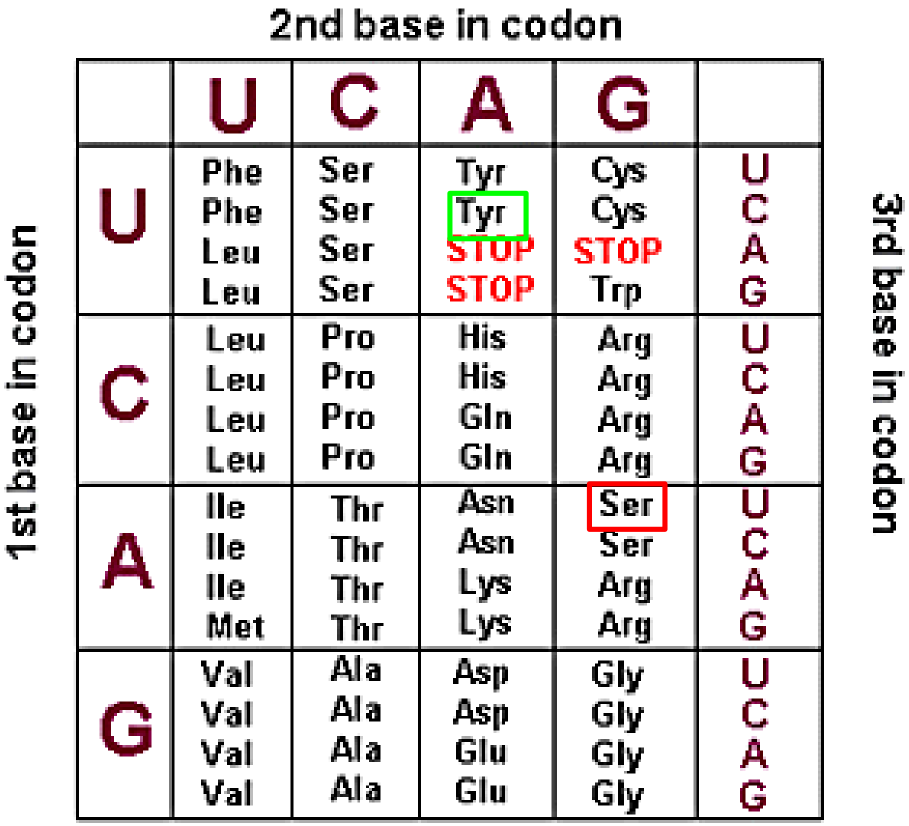
1. Amino acids are N-containing molecules
2. The basic structural building units (monomers) of proteins **(protein role)**
3. Precursors of many biomolecules like neurotransmitters **(non-protein role)**
4. They are also utilized as an energy source

- There are **20 standard** (canonical) amino acids which are encoded directly by triplet codons in the universal genetic code during in vivo protein synthesis process (mRNA translation)

Genetic Code Table



- **The 20 standard** amino acids are known as proteinogenic or natural amino acids


2nd base in codon

1st base in codon

3rd base in codon

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP	STOP	A
	Leu	Ser	STOP	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

Standard Amino Acids List

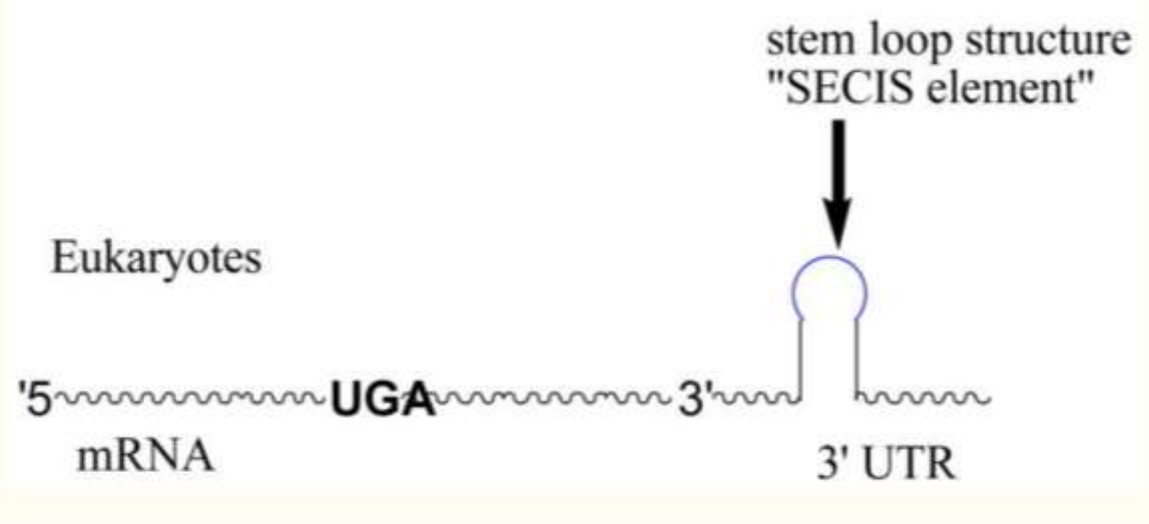
- Histidine
- Isoleucine
- Leucine
- Methionine
- Phenylalanine
- Threonine
- Tryptophan
- Valine
- Lysine

- Arginine
- Asparagine
- Glutamine
- Glycine
- Proline
- Serine
- Tyrosine

- Alanine
- Asparatate
- Cysteine
- Glutamate

Standard Amino Acids

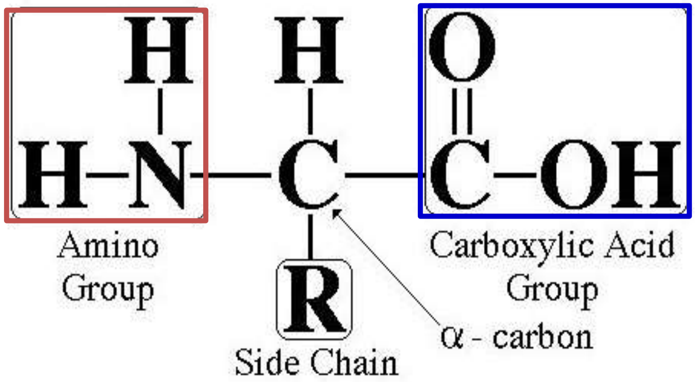
1. They are proteinogenic and natural amino acids (the other proteinogenic amino acids N-formyl methionine, pyrrolysine and selenocysteine are called non-standard or non-canonical amino acids)

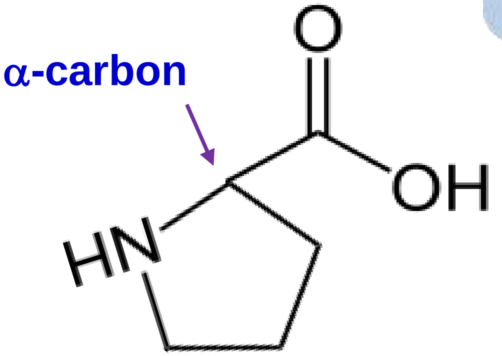


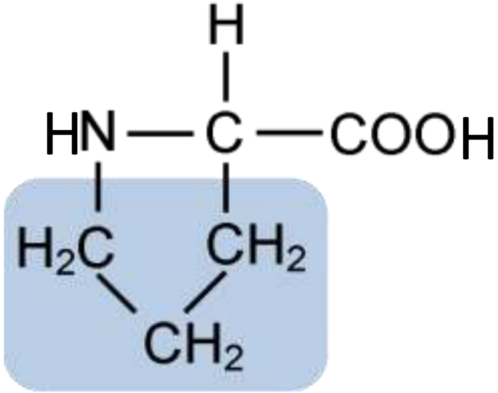
Incorporation of selenocysteine in protein structure by unique mechanism

Standard Amino Acids

2. Known as 2-, alpha- or **α-amino acids** as the primary amino group ($-NH_2$) is attached to α-carbon (the carbon next to –COOH group). Proline is an exception which has a secondary amino group (-NH-)



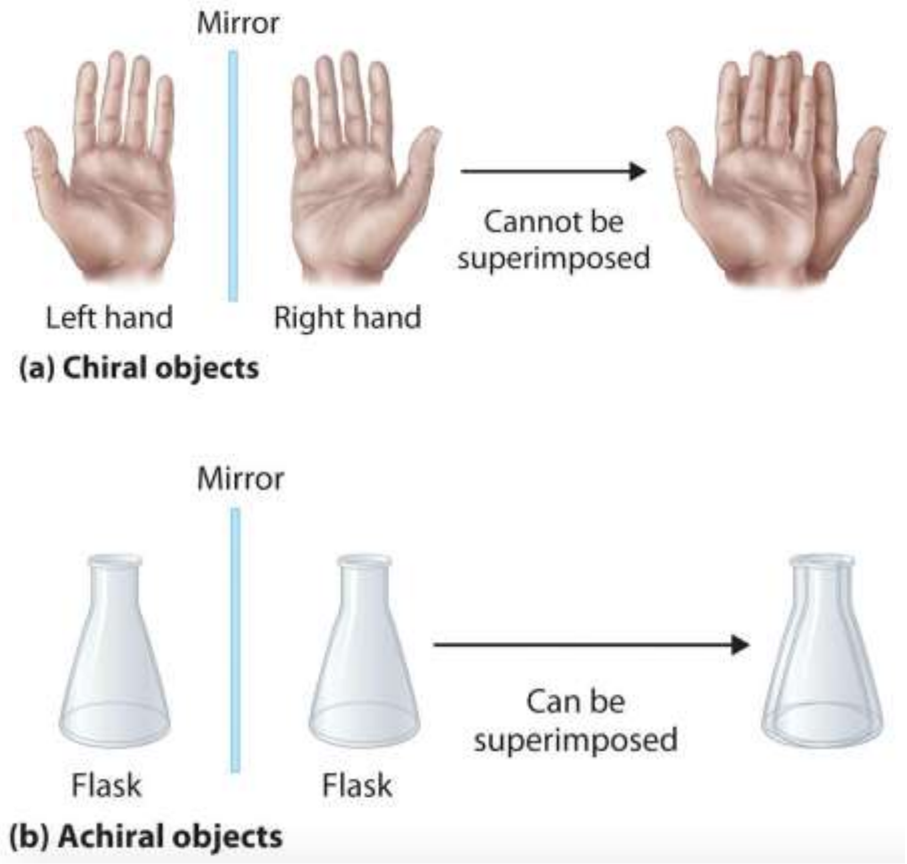




proline

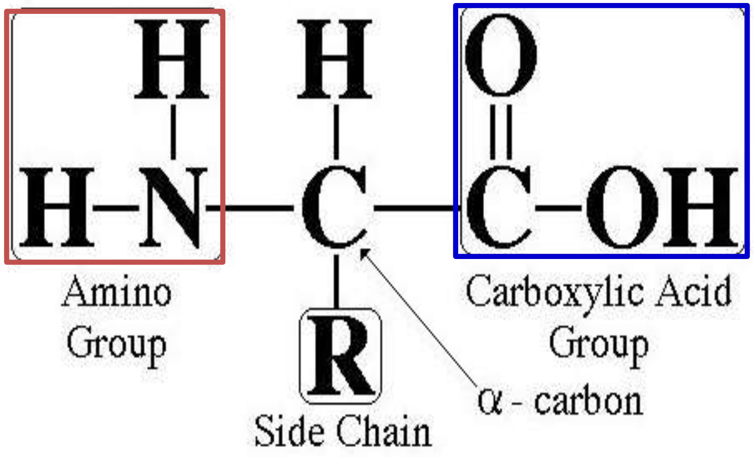
Standard Amino Acids

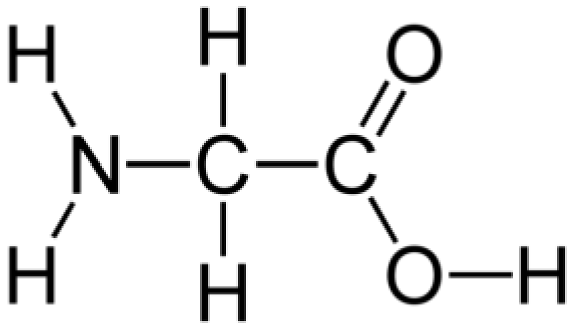
3. They are all chiral molecules (except glycine which has **achiral** C) with **L**- stereochemical configuration (left-handed isomers)



Standard Amino Acids

- Chiral molecules should contain at least one chiral center (**usually a carbon atom**)
- **Chiral carbon**: asymmetric carbon atom attached to 4 different groups of atoms





Glycine

Isomerization

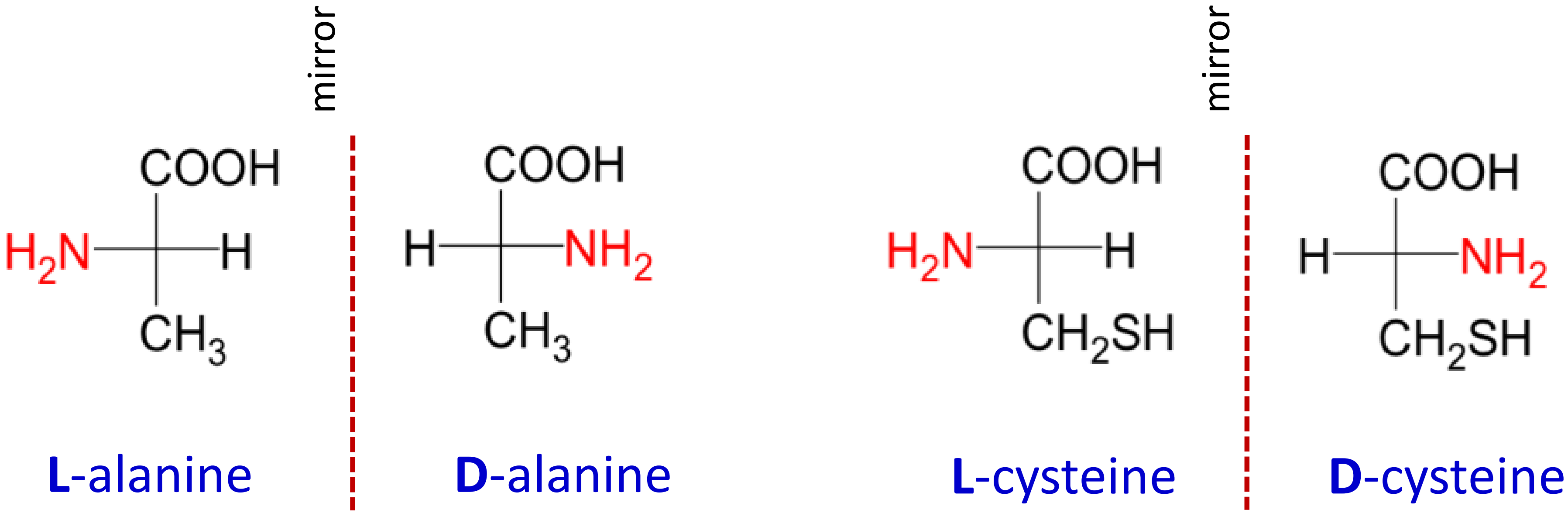
- Isomers: are molecules with same molecular formula but different chemical structures

1. Constitutional (structural) isomers: atoms and functional groups bind together in different ways
2. Stereoisomers (spatial isomers): differ in the configuration of atoms rather than the order of atomic connectivity

D/L Amino Acids

- Enantiomers: are two stereoisomers that are mirror images to each other but not superimposable
- **D-** (dexter)/**L-** (laevus) Nomenclature system: commonly used to assign the configurations in sugars (carbohydrates) and amino acids
- As a rule of thumb: if the amino group is on the right-hand side of α-carbon at Fisher projection, the configuration is D. If it is on the left-hand side, the configuration is assigned as L.

Fischer Projections of Amino Acids



L-alanine **D**-alanine **L**-cysteine **D**-cysteine

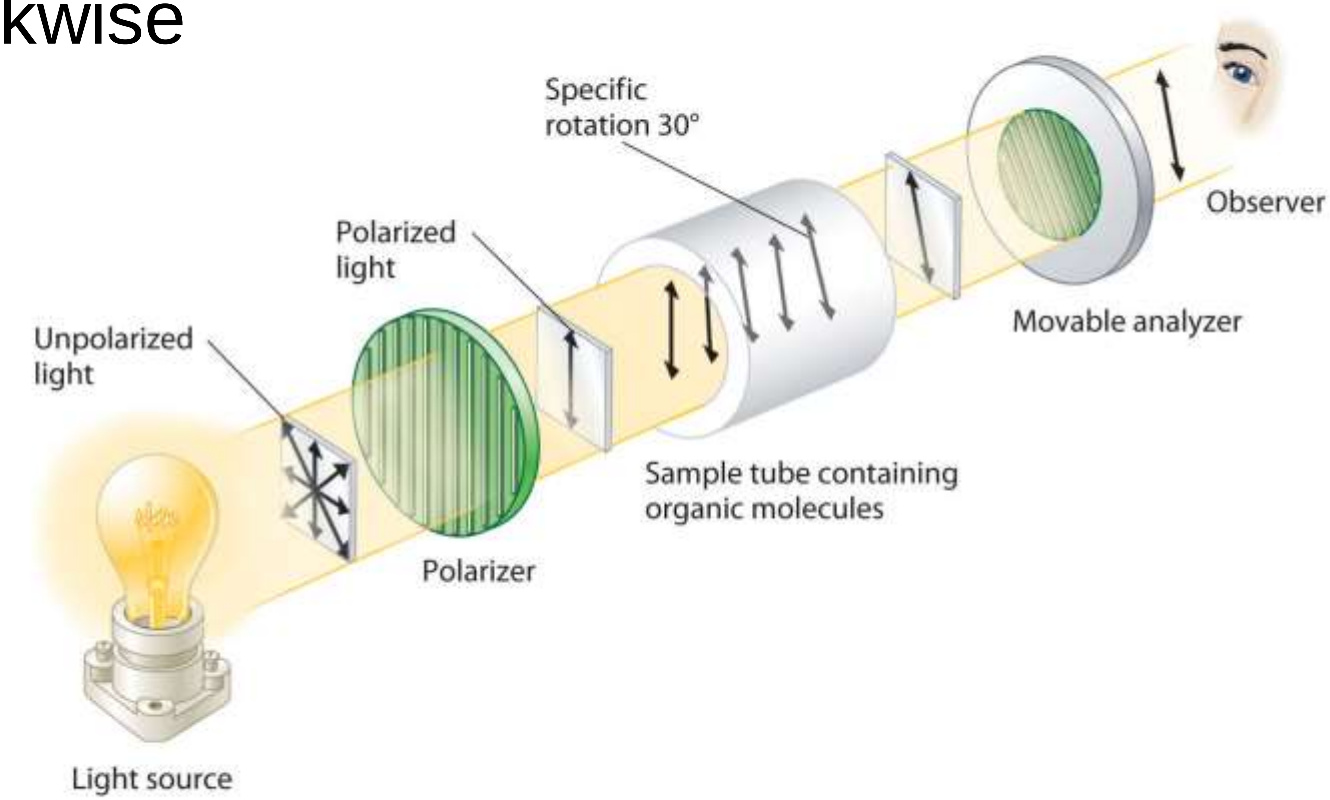
Fisher Projection: is one way commonly used to represent the structure of chiral molecules like carbohydrates and amino acids

D/L Amino Acids

- Most naturally occurring sugars are D-isomers while most naturally occurring amino acids are L-isomers (amino acids of protein)
- D-amino acids polypeptides (right-handed isomers) are components of bacterial cell walls to resist digestion by other organisms

Optical Activity

- Enantiomers are optically active and can rotate the polarized light plane either clockwise or counterclockwise



Polarimeter is used to measure optical rotation

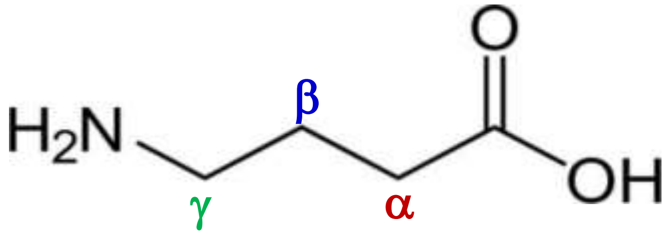
Optical Activity

- **(+)/(-) nomenclature system**: if one enantiomer rotates the light clockwise, it is labeled (+) or (*d*) (dextrorotatory). The second mirror image enantiomer is labeled (-) or (*l*) laevorotatory
- D/L system should not be confused with +/- or *d*/l system. For example, D-isomer might be levorotatory
- 9 of 19 L-amino acids commonly found in proteins are dextrorotatory
- **Racemic mixture** contains equal amounts of each enantiomer (net rotation is zero)

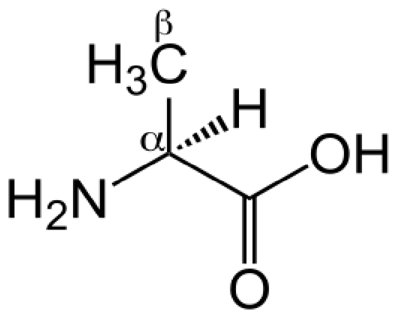
Classification of Amino Acids

- >300 amino acids classified in many ways:

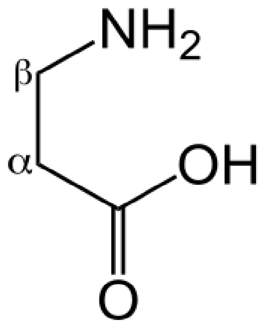
1) Standard and non-standard amino acids
2) α, β, γ and δ amino acids



γ-aminobutyric acid (GABA) is the inhibitory neurotransmitter in the brain



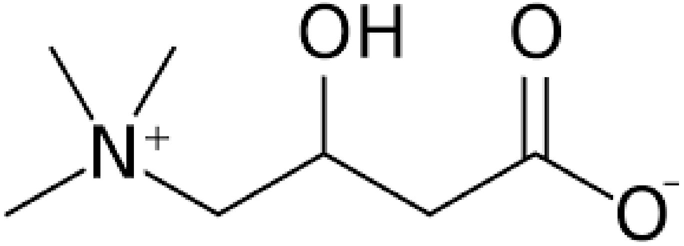
α-alanine



β-alanine

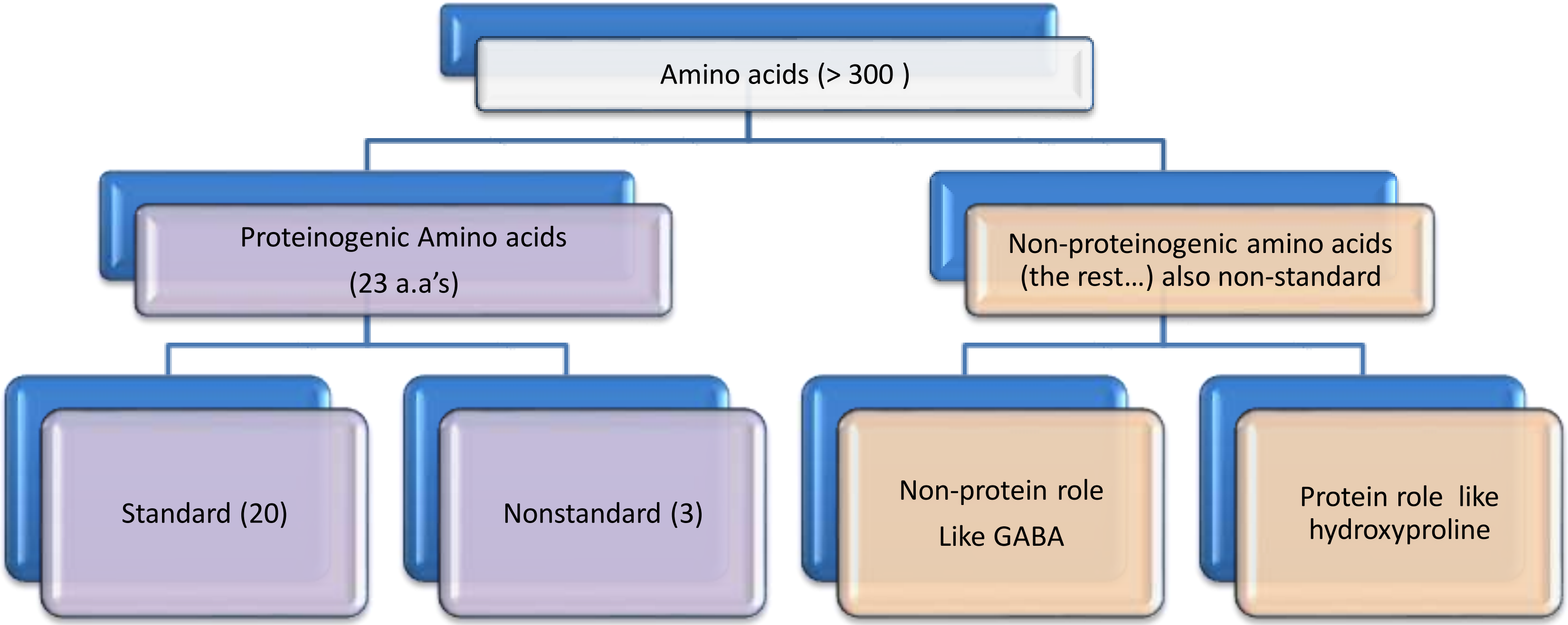
Classification of Amino Acids

3) Proteinogenic and non-proteinogenic amino acids (non-proteinogenic amino acids either have non-protein role like GABA and carnitine or have a protein role but formed by post-translational modification of protein like hydroxyproline)



Carnitine has a role in lipid transportation and fat metabolism

Classification of Amino Acids



Amino acids (> 300)

Proteinogenic Amino acids (23 a.a's)

Non-proteinogenic amino acids (the rest…) also non-standard

Standard (20)

Nonstandard (3)

Non-protein role Like GABA

Protein role like hydroxyproline

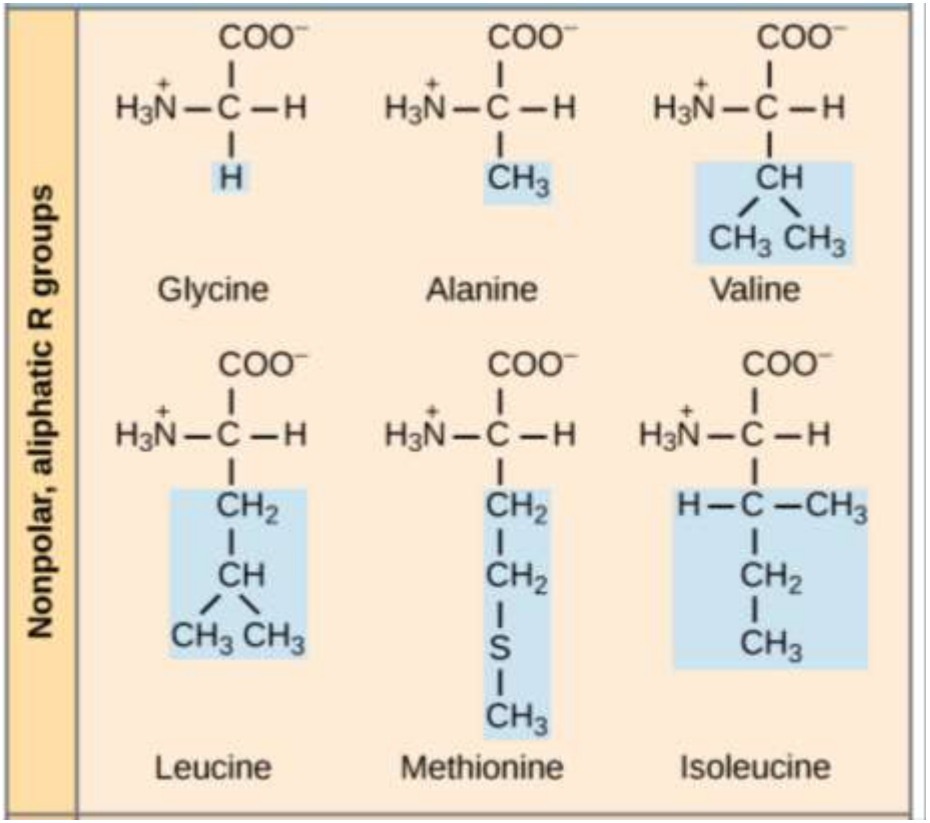
Categories of Standard Amino Acids

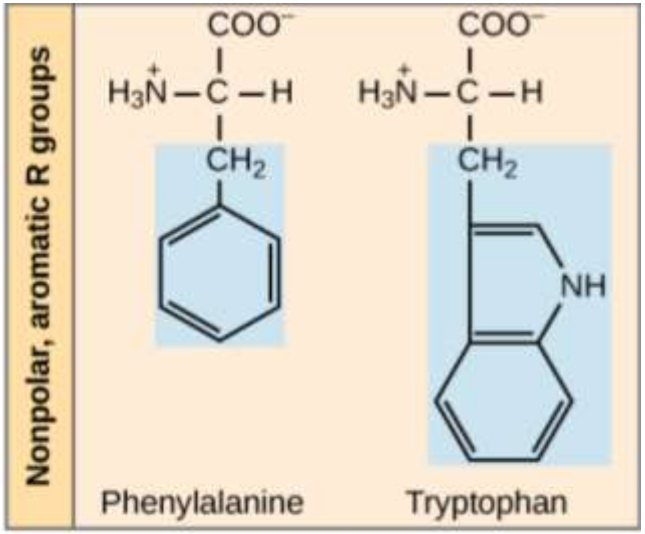
- The 20 standard amino acids are classified into 3 major categories according to the polarities of their “R” groups:

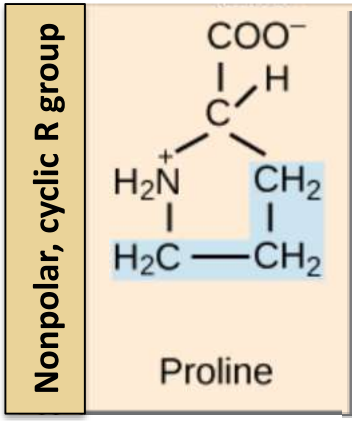
1) **Amino acids with non-polar R groups**
2) **Amino acids with charged polar R groups**
3) **Amino acids with uncharged polar R groups**

Amino acids with non-polar R groups

- 6 amino acids with aliphatic, 2 with aromatic and one with cyclic side chains





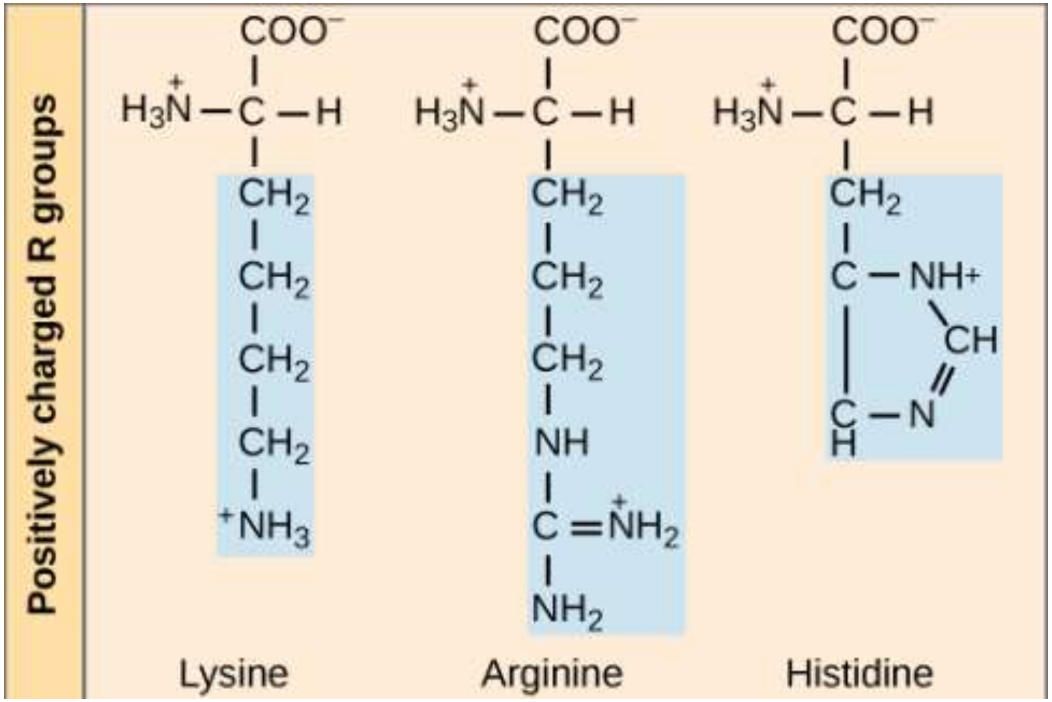


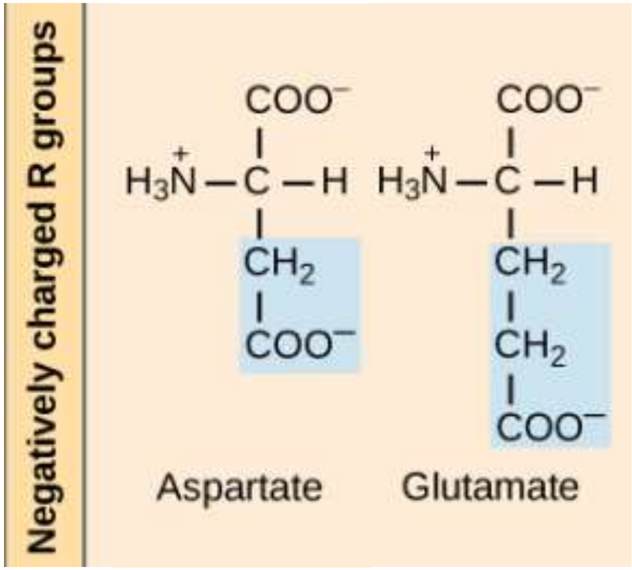
Amino acids with non-polar R groups

- Glycine has the simplest side chain: H atom
- Alanine, valine, leucine and isoleucine have aliphatic hydrocarbon side chains
- Methionine has a thioether side chain (sulfur atom)
- Proline has a cyclic pyrrolidine side chain
- Phenylalanine has a phenyl moiety
- Tryptophan has an indole group

Amino acids with charged polar R groups

- 3 amino acids are positively charged (basic) and 2 amino acids are negatively charged (acidic)



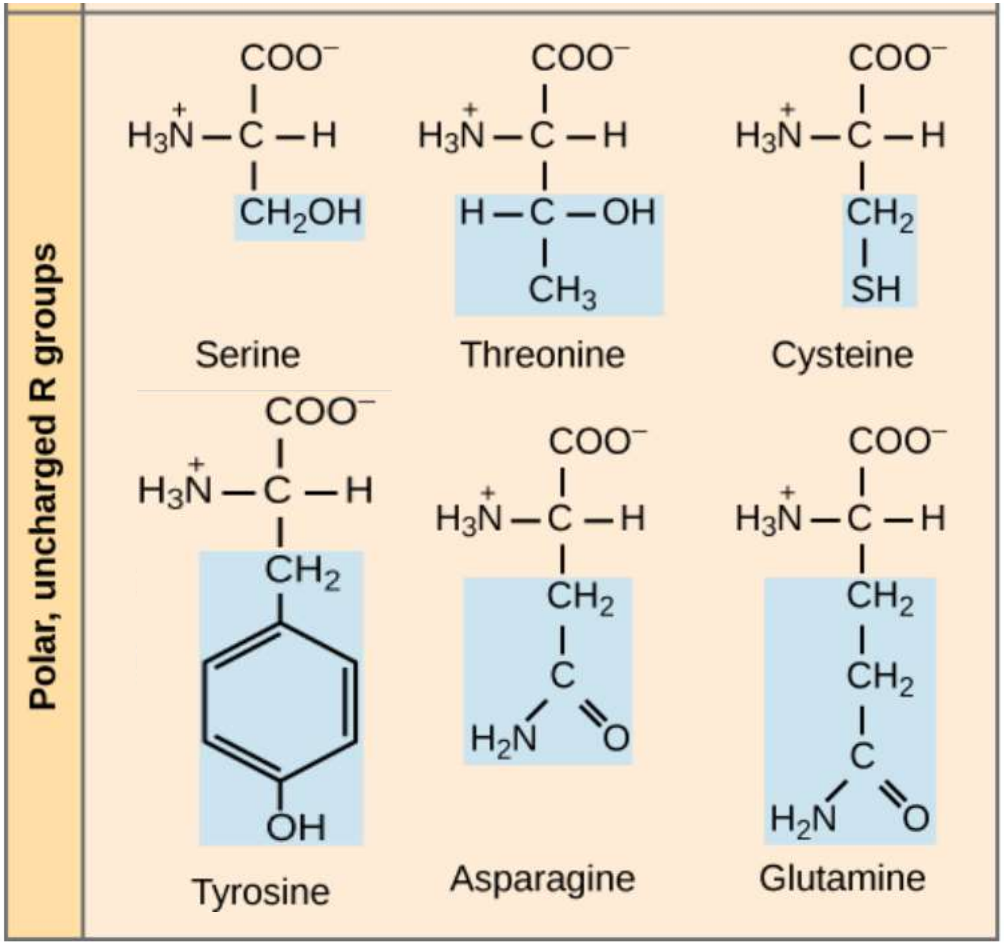


Amino acids with charged polar R groups

- Arginine has a guanidine group
- Lysine has a butyl ammonium side chain
- Histidine has imidazole group
- Aspartic and glutamic acids in their ionized state are called aspartate and glutamate, respectively

Amino acids with uncharged polar R groups

- 6 amino acids with hydroxyl, amide or thiol groups



Amino acids with uncharged polar R groups

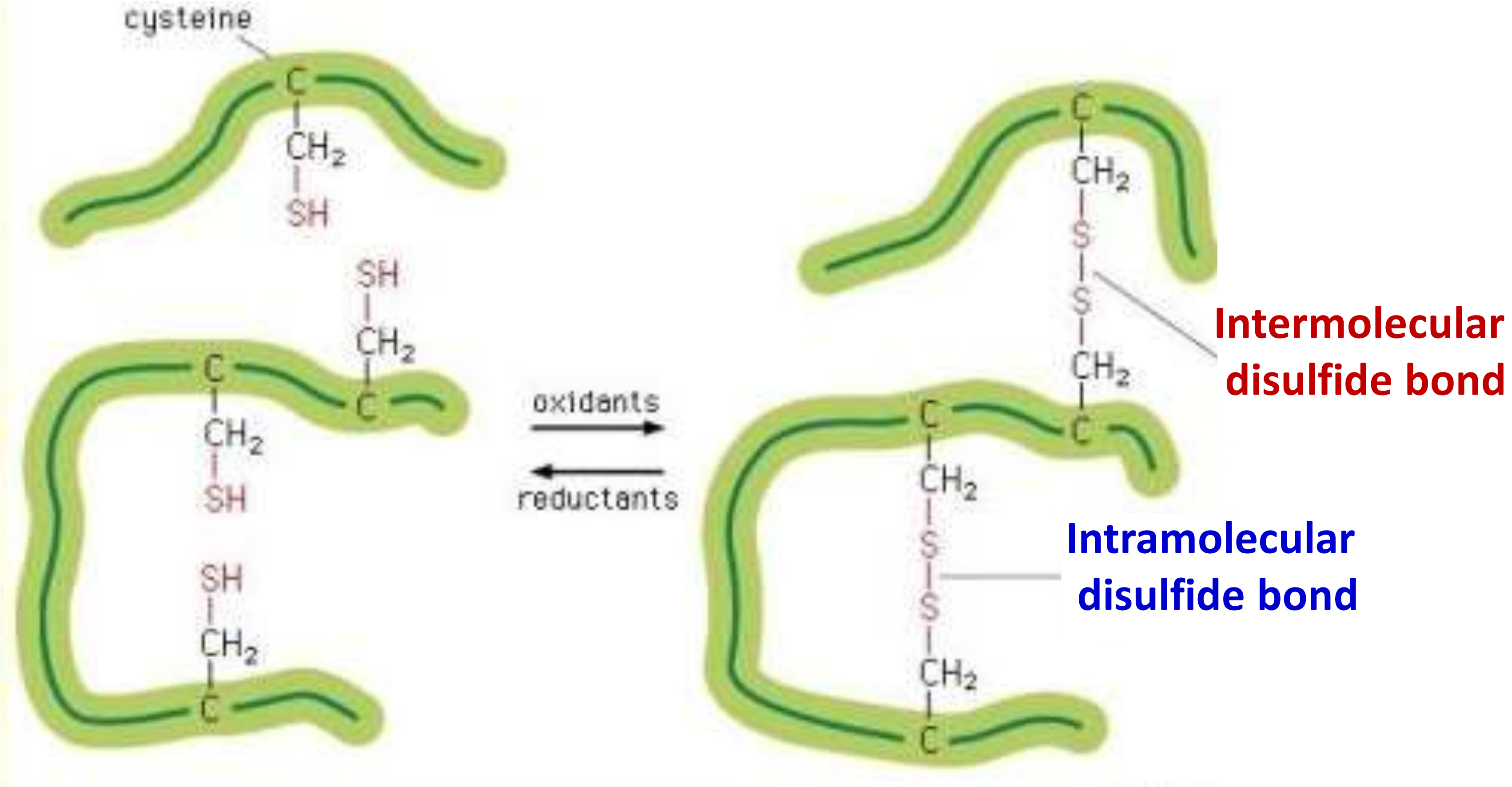
- Serine and threonine bear hydroxyl (-OH) R group

  and Tyrosine is aromatic and has a phenolic group

- Asparagine and glutamine have amide bearing side chains. They are the amide derivatives of aspartic and glutamic acids (OH is replaced with NH_2)
- Cysteine is unique because it has free sulfhydryl **(-SH)** group that can form a disulfide bond (-S-S-) with another cysteine through the oxidation of 2 thiol groups (cystine is the oxidized **dimeric** form). The disulfide bridge in proteins contributes to the stability and overall shape of a protein

Amino acids with uncharged polar R groups

- **Disulfide bond** is a covalent linkage formed between the sulfhydryl groups (SH) of **two cysteine residues (**after oxidation) to produce a **cystine** residue



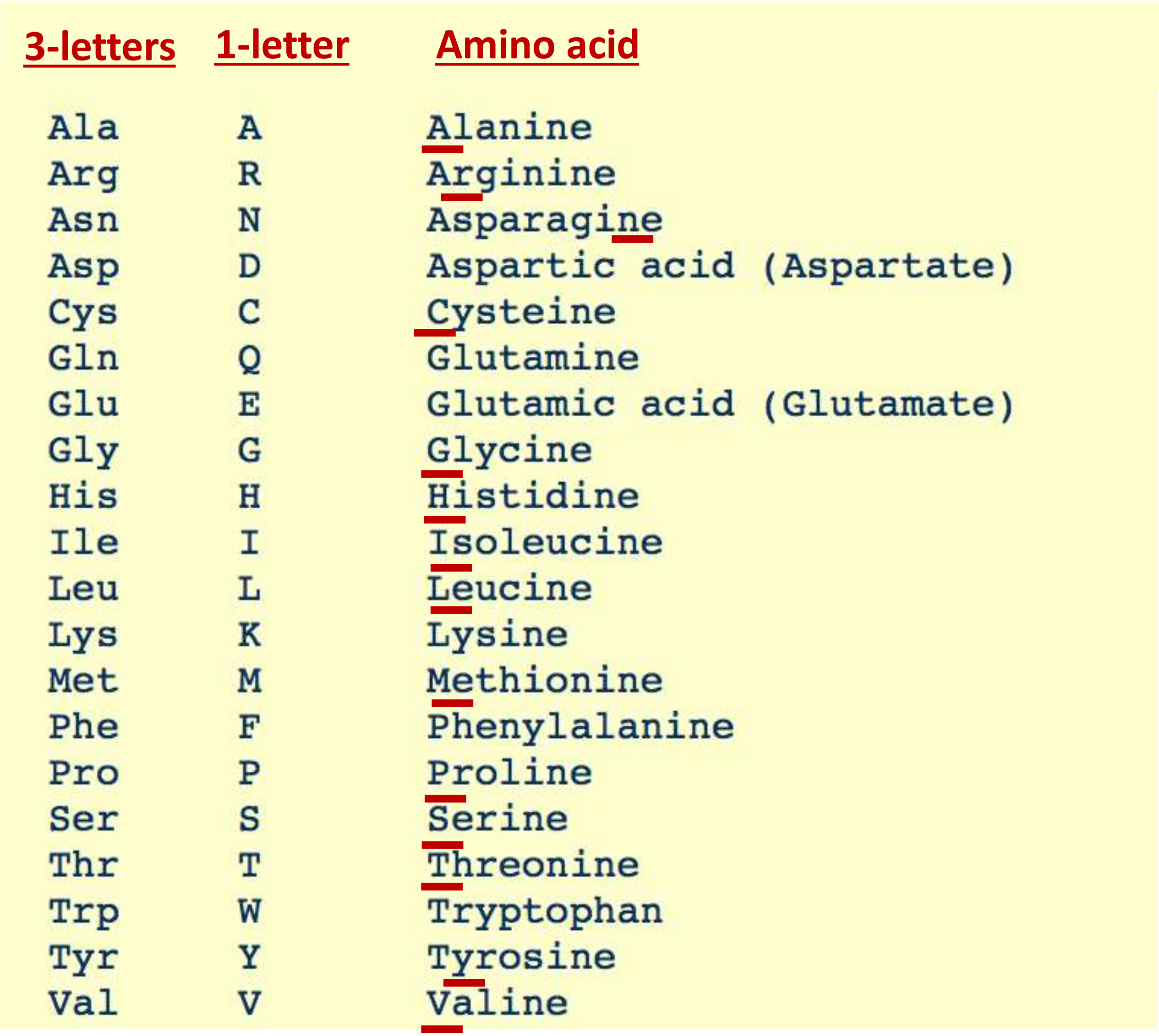
Intermolecular disulfide bond

Intramolecular disulfide bond

Amino acids with uncharged polar R groups

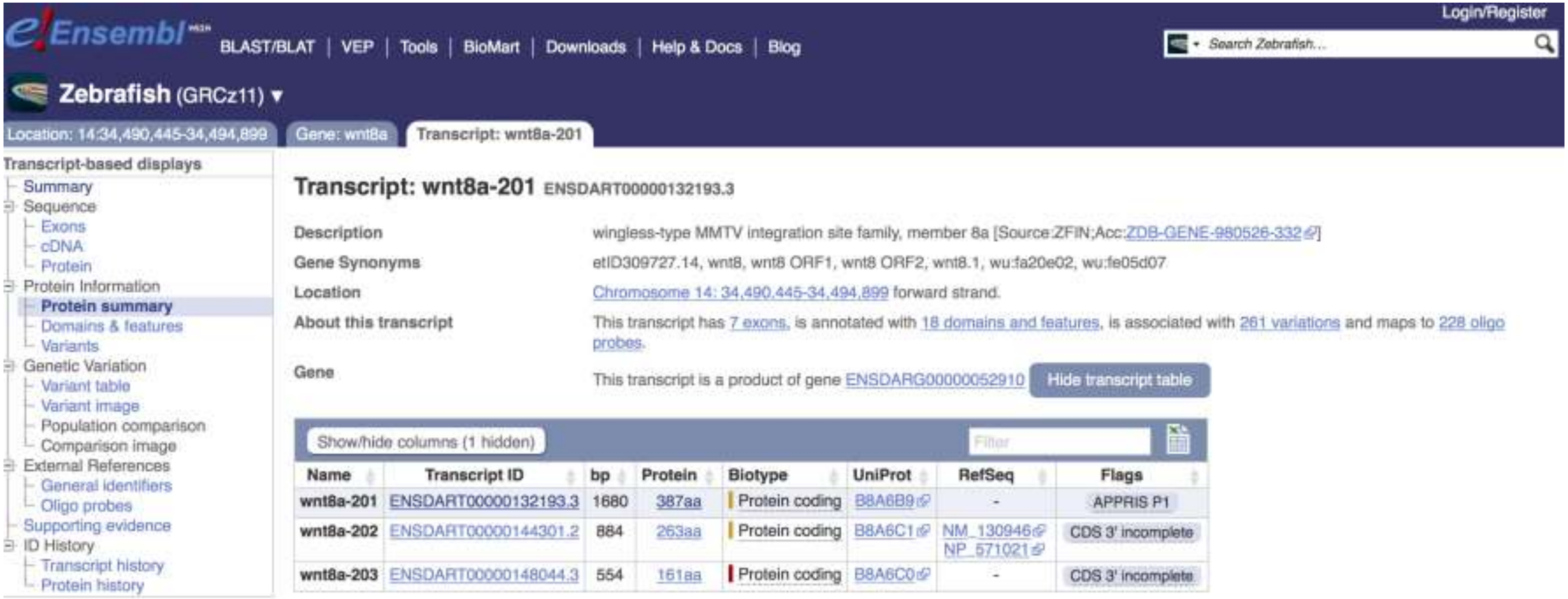
- Cysteine residues may be separated from each other by many amino acids in the primary sequence of a polypeptide or may even be located on two different polypeptides. The folding of the polypeptide chain(s) brings the cysteine residues into proximity and permits covalent bonding of their side chains.
- Disulfide bond could be **intramolecular** (2 cysteine residues on the same polypeptide chain) or **intermolecular** (2 cysteine residues on two separate/ different polypeptide chains)

Amino Acids Abbreviations



3-letters	1-letter	Amino acid
Ala	A	Alanine
Arg	R	Arginine
Asn	N	Asparagine
Asp	D	Aspartic acid (Aspartate)
Cys	C	Cysteine
Gln	Q	Glutamine
Glu	E	Glutamic acid (Glutamate)
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Leu	L	Leucine
Lys	K	Lysine
Met	M	Methionine
Phe	F	Phenylalanine
Pro	P	Proline
Ser	S	Serine
Thr	T	Threonine
Trp	W	Tryptophan
Tyr	Y	Tyrosine
Val	V	Valine

Ensembl Genomic Browser

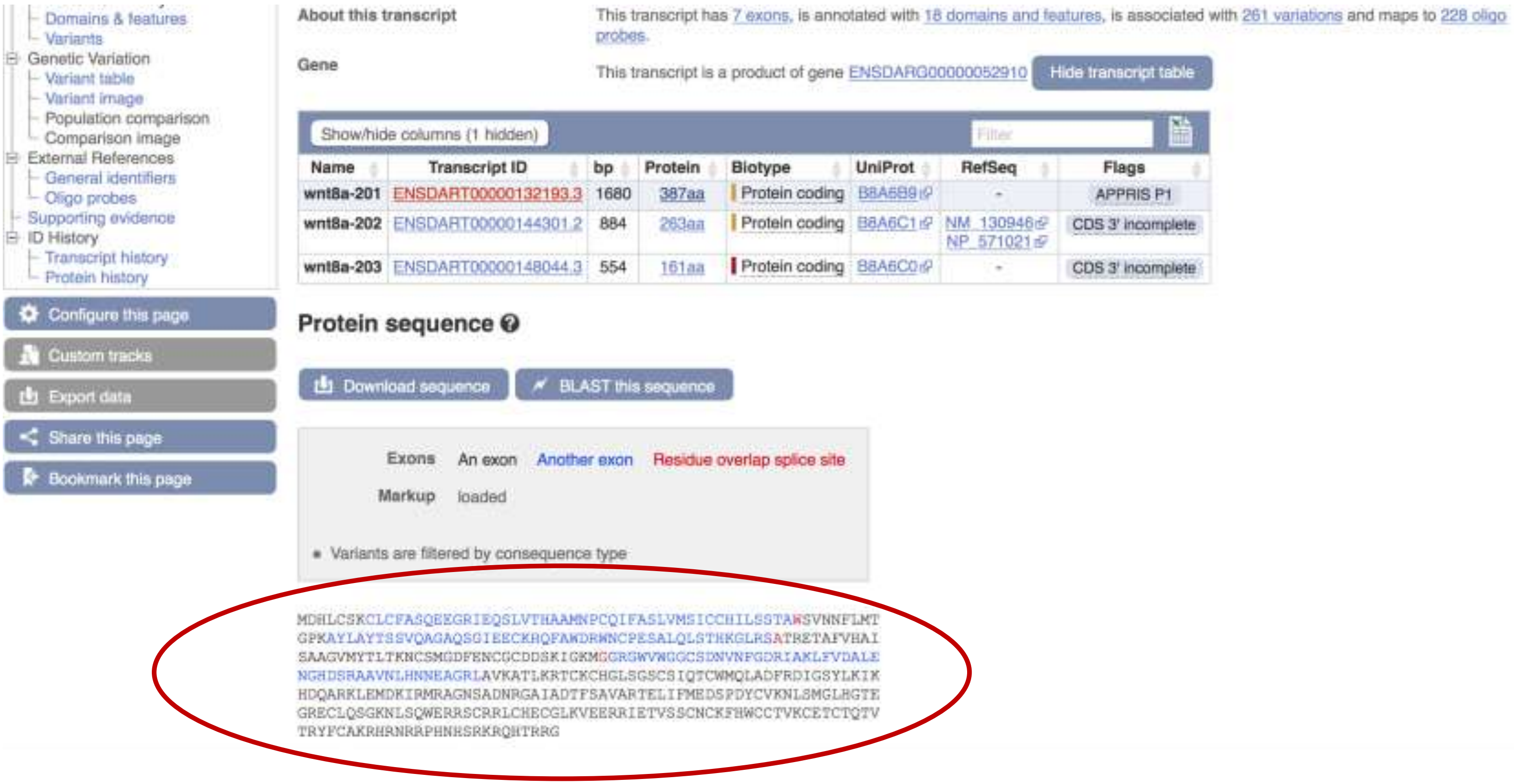


Transcript: wnt8a-201 ENSDART00000132193.3

Description	wingless-type MMTV integration site family, member 8a [Source:ZFIN;Acc:ZDB-GENE-980526-332]
Gene Synonyms	etID309727.14, wnt8, wnt8 ORF1, wnt8 ORF2, wnt8.1, wu:fa20e02, wu:fe05d07
Location	Chromosome 14: 34,490,445-34,494,899 forward strand.
About this transcript	This transcript has 7 exons, is annotated with 18 domains and features, is associated with 261 variations and maps to 228 oligo probes.
Gene	This transcript is a product of gene ENSDARG00000052910

Name	Transcript ID	bp	Protein	Biotype	UniProt	RefSeq	Flags
wnt8a-201	ENSDART00000132193.3	1680	387aa	Protein coding	B8A6B9	-	APPRIS P1
wnt8a-202	ENSDART00000144301.2	884	263aa	Protein coding	B8A6C1	NM_130946 NP_571021	CDS 3' incomplete
wnt8a-203	ENSDART00000148044.3	554	161aa	Protein coding	B8A6C0	-	CDS 3' incomplete

Ensembl Genomic Browser



About this transcript This transcript has 7 exons, is annotated with 18 domains and features, is associated with 261 variations and maps to 228 oligo probes.

Gene This transcript is a product of gene ENSDARG00000052910

Name	Transcript ID	bp	Protein	Biotype	UniProt	RefSeq	Flags
wnt8a-201	ENSDART00000132193.3	1680	387aa	Protein coding	B8A6B9	-	APPRIS P1
wnt8a-202	ENSDART00000144301.2	884	263aa	Protein coding	B8A6C1	NM_130946 NP_571021	CDS 3' incomplete
wnt8a-203	ENSDART00000148044.3	554	161aa	Protein coding	B8A6C0	-	CDS 3' incomplete

Protein sequence

Exons An exon Another exon Residue overlap splice site

Markup loaded

- Variants are filtered by consequence type

```
MDHLCSKCLCFASQEEGRIEQSLVTHAAMNPCQIFASLVMSICCHILSSTAWSVNNFLMT
GPKAYLAYTSSVQAGAQSGIEECKHQFAWDRWNCPESALQLSTHKGLRSATRETAFVHAI
SAAGVMYTLTKNCSMGDFENCGCDDSKIGKMGGRGWVWGGCSDNVNFGDRIAKLFVDALE
NGHDSRAAVNLHNNEAGRLAVKATLKRTCKCHGLSGSCSIQTCWMQLADFRDIGSYLKIK
HDQARKLEMDKIRMRAGNSADNRGAIADTFSAVARTELIFMEDSPDYCVKNLSMGLHGTE
GRECLQSGKNLSQWERRSCRRLCHECGLKVEERRIETVSSCNCKFHWCCTVKCETCTQTV
TRYFCAKRHRNRRPHNHSRKRQHTRRG
```