

Single, Three Letter Amino Acid Codes

A = Ala = Alanine
 C = Cys = Cysteine
 D = Asp = Aspartate
 E = Glu = Glutamate
 F = Phe = Phenylalanine
 G = Gly = Glycine
 H = His = Histidine
 I = Ile = Isoleucine
 K = Lys = Lysine
 L = Leu = Leucine
 M = Met = Methionine
 N = Asn = Asparagine
 P = Pro = Proline
 Q = Gln = Glutamine
 R = Arg = Arginine
 S = Ser = Serine
 T = Thr = Threonine
 V = Val = Valine
 W = Trp = Tryptophan
 X = Xaa = unknown [non standard – Unk]
 Y = Tyr = Tyrosine

Single Letter DNA Degenerate Code

N = aNy (A, C, G or T)
 Y = pYrimidine (C or T) R = puRine (A or G)
 W = Weak (A or T) S = Strong (C or G)
 K = Keto (G or T) M = aMino (A or C)
 B = "not A" (C, G or T) D = "not C" (A, G or T)
 H = "not G" (A, C or T) V = "not T" (A, C or G)

The Genetic Code			
TTT phe F	TCT ser S	TAT tyr Y	TGT cys C
TTC phe F	TCC ser S	TAC tyr Y	TGC cys C
TTA leu L	TCA ser S	TAA OCH Z	TGA OPA Z
TTG leu L	TCG ser S	TAG AMB Z	TGG trp W
CTT leu L	CCT pro P	CAT his H	CGT arg R
CTC leu L	CCC pro P	CAC his H	CGC arg R
CTA leu L	CCA pro P	CAA gln Q	CGA arg R
CTG leu L	CCG pro P	CAG gln Q	CGG arg R
ATT ile I	ACT thr T	AAT asn N	AGT ser S
ATC ile I	ACC thr T	AAC asn N	AGC ser S
ATA ile I	ACA thr T	AAA lys K	AGA arg R
ATG met M	ACG thr T	AAG lys K	AGG arg R
GTT val V	GCT ala A	GAT asp D	GGT gly G
GTC val V	GCC ala A	GAC asp D	GGC gly G
GTA val V	GCA ala A	GAA glu E	GGA gly G
GTG val V	GCG ala A	GAG glu E	GGG gly G

Amino Acid Properties

(Modified from <http://www.mcb.ucdavis.edu/courses/bis102/AAProp.html>.) This list is provided to remind you of the properties of the side chains of the 20 amino acids present in newly synthesized proteins. The R-groups can be classified in a number of different ways, several of which are described below.

- 1 All - G, A, V, L, I, P, F, Y, W, S, T, N, Q, C, M, D, E, H, K, R
- 2 Polar / hydrophilic - N, Q, S, T, K, R, H, D, E, (C, Y)*
- 3 Non-polar / hydrophobic - (G), A, V, L, I, P, Y, F, W, M, C
- 4 H-bonding - C, W, N, Q, S, T, Y, K, R, H, D, E
- 5 Sulfur containing - C, M
- 6 Charged at Neutral pH Negative / acidic - D, E, (C)
- 7 Charged at Neutral pH Positive / basic - K, R, (H)
- 8 Ionizable - D, E, H, C, Y, K, R
- 9 Aromatic - F, W, Y, (H, but no significant UV absorption)
- 10 Aliphatic - G, A, V, L, I, P
- 11 Cyclic - P
- 12 "Hard to remember" one letter code - W, Y, K, N, Q, D, E

*Note: Amino acids in parentheses have the indicated character to a limited extent.

Some post-translational covalent modifications of amino acid side chains in proteins

- 13 Covalent cross-links (intra- or intermolecular) - C (disulfide bond)
- 14 Phosphorylation (-OH containing) - S, T, Y
- 15 Phosphorylation (via -NH) - H
- 15 Glycosylation - N (N-linked), S, T (O-linked)
- 16 Methylation - K, R
- 17 Acetylation - K
- 18 Ubiquitination - K

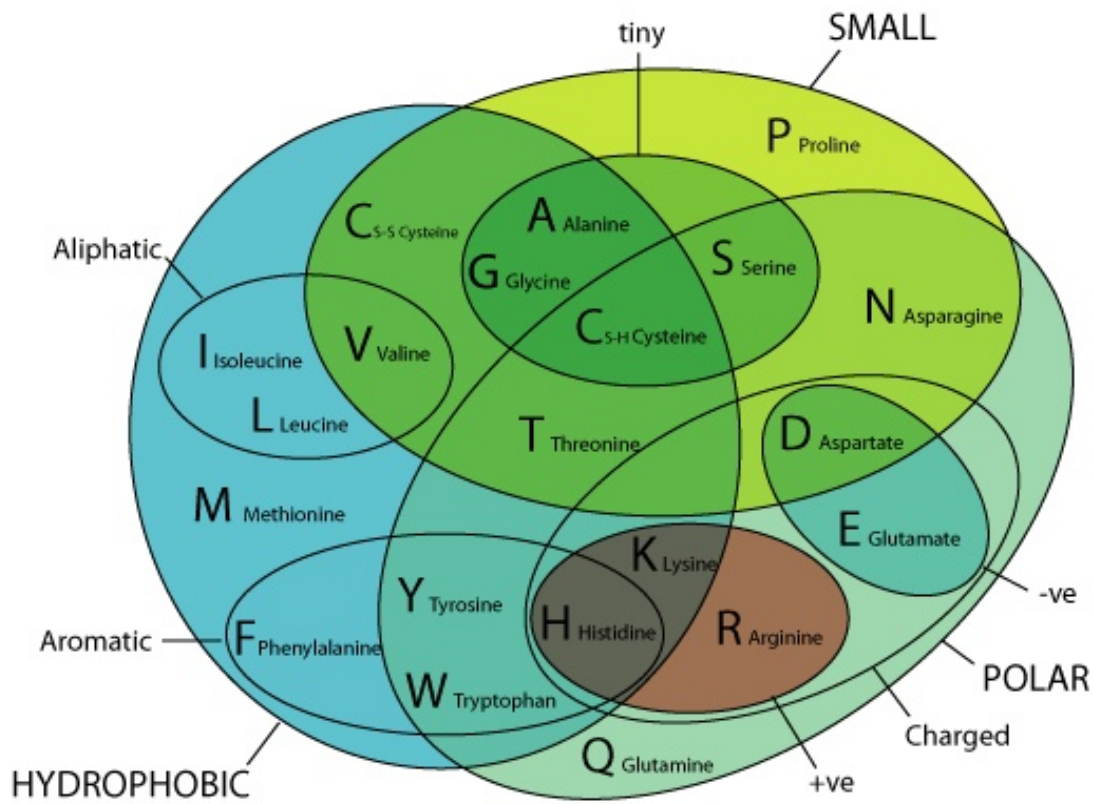


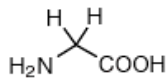
Image from Livingstone, C. D. and Barton, G. J. (1993), "Protein Sequence Alignments: A Strategy for the Hierarchical Analysis of Residue Conservation", *Comp. Appl. Bio. Sci.*, 9, 745-756.

BLOSUM62 Substitution matrix (e.g., used in sequence alignment scoring)

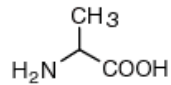
Table shows bonus or penalty score for substituting one amino acid for another

	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W	
C	9																				C
S	-1	4																			S
T	-1	1	5																		T
P	-3	-1	-1	7																	P
A	0	1	0	-1	4																A
G	-3	0	-2	-2	0	6															G
N	-3	1	0	-2	-2	0	6														N
D	-3	0	-1	-1	-2	-1	1	6													D
E	-4	0	-1	-1	-1	-2	0	2	5												E
Q	-3	0	-1	-1	-1	-2	0	0	2	5											Q
H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8										H
R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5									R
K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5								K
M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5							M
I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	-3	1	4						I
L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	-2	2	2	4					L
V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	-3	-2	1	3	1	4				V
F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6			F
Y	-2	-2	-2	-3	-2	-3	-2	-3	-2	-1	2	-2	-2	-1	-1	-1	-1	3	7		Y
W	-2	-3	-2	-4	-3	-2	-4	-4	-3	-2	-2	-3	-3	-1	-3	-2	-3	1	2	11	W

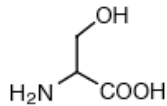
Small



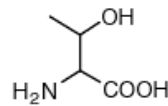
Glycine (Gly, G)
MW: 57.05



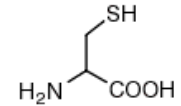
Alanine (Ala, A)
MW: 71.09



Serine (Ser, S)
MW: 87.08, pK_a ~ 16



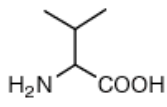
Threonine (Thr, T)
MW: 101.11, pK_a ~ 16



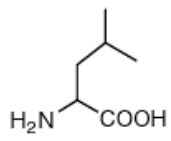
Cysteine (Cys, C)
MW: 103.15, pK_a = 8.35

Nucleophilic

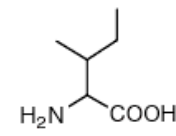
Hydrophobic



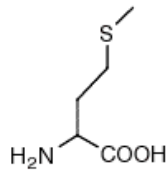
Valine (Val, V)
MW: 99.14



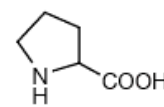
Leucine (Leu, L)
MW: 113.16



Isoleucine (Ile, I)
MW: 113.16

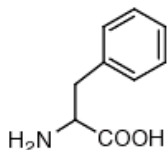


Methionine (Met, M)
MW: 131.19

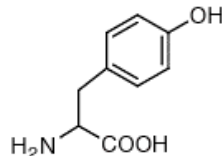


Proline (Pro, P)
MW: 97.12

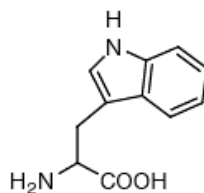
Aromatic



Phenylalanine (Phe, F)
MW: 147.18

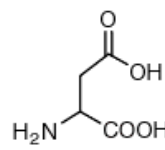


Tyrosine (Tyr, Y)
MW: 163.18

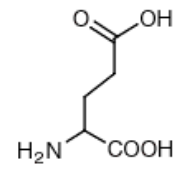


Tryptophan (Trp, W)
MW: 186.21

Acidic

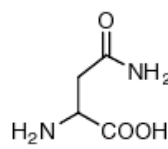


Aspartic Acid (Asp, D)
MW: 115.09, pK_a = 3.9

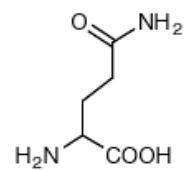


Glutamic Acid (Glu, E)
MW: 129.12, pK_a = 4.07

Amide

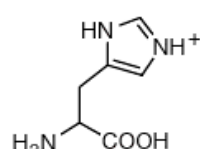


Asparagine (Asn, N)
MW: 114.11

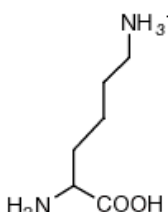


Glutamine (Gln, Q)
MW: 128.14

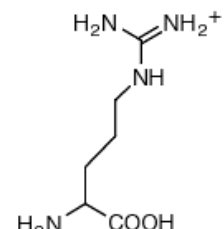
Basic



Histidine (His, H)
MW: 137.14, pK_a = 6.04



Lysine (Lys, K)
MW: 128.17, pK_a = 10.79



Arginine (Arg, R)
MW: 156.19, pK_a = 12.48