Table 4-1. Key to Structure. Covalent Structures and Abbreviations of the "Standard" Amino Acids of Proteins, Their Occurrence, and the pK Values of Their Ionizable Groups

Three-le	ame, tter Symbol, etter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	$pK_2 lpha \sim NH_3^{+d}$	pK _R Side Chain ^d
Amino aci Glycine Gly G	COO- H-C-H	r side chains	57.0	7.2	2.35	9.78	
Alanine Ala A	NH ₃ COO ⁻ H-C-CH ₃ NH ₃		71.1	7.8	2.35	9.87	
Valine Val V	COO ⁻ CH ₃ H-C-CH NH ₃ CH ₃		99.1	6.6	2.29	9.74	
Leucine Leu L	COO ⁻ H-C-CH ₂ - NH ₃ ⁺	CH ₃ CH CH ₃	113.2	9.1	2.33	9.74	
Isoleucine Ile I	COO - CH H - C - C* NH3 H	H ₃ 	113.2	5.3	2.32	9.76	
Methionine Met M	COO ⁻ H-C-CH ₂ -0 NH ₃ +		131.2	2.2	2.13	9.28	
Proline Pro P	COO-C3 4 CH2 H N CH2		97.1	5.2	1.95	10.64	
Phenylalan Phe F	ine $COO^ H-C-CH_2$ NH_3^+		147.2	3.9	2.20	9.31	
Tryptophar Trp W	H-C-CH ₂ -NH ₃	$\begin{array}{c c} 3 & & & \\ 2 & 1 & & \\ N & & & \\ H & & & \end{array}$	186.2	1.4	2.46	9.41	

^aThe ionic forms shown are those predominating at pH 7.0 although residue mass is given for the neutral compound The C_{α} atoms, as well as those atoms marked with an asterisk, are chiral centers with configurations as indicated according to Fischer projection formulas. The standard organic numbering system is provided for heterocycles.

 $[^]b$ The residue masses are given for the neutral residues. For the molecular masses of the parent amino acids, add 18.0 D, the molecular mass of H_2O , to the residue masses. For side chain masses, subtract 56.0 D, the formula mass of a peptide group, from the residue masses.

^{&#}x27;Calculated from a database of nonredundant proteins containing 300,688 residues as compiled by Doolittle, R.F. in Fasman, G.D. (Ed.), *Predictions of Protein Structure and the Principles of Protein Conformation*, Plenum Press (1989).

^dData from Dawson, R.M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M., *Data for Biochemical Research* (3rd ed.), pp. 1–31, Oxford Science Publications (1986).

^eThe three- and one-letter symbols for asparagine *or* aspartic acid are Asx and B, whereas for glutamine *or* glutamic acid they are Glx and Z. The one-letter symbol for an undetermined or "nonstandard" amino acid is X.

 $[^]f$ Both neutral and protonated forms of histidine are present at pH 7.0, since its pK_R is close to 7.0.

Table 4-1. (continued)

Name, Three-letter Symbol, and One-letter Symbol	Structural Ma Formula ^a (D	ss Occurre	nce pK ₁	$pK_2 lpha ext{-}NH_3^{+d}$	pK _R Side Chain ^d					
Amino acids with uncharged polar side chains										
Serine COO-	87	6.8	2.19	9.21						
Ser $H-C-CH_2-O$ NH_3^+	Н									
Threonine COO^- H Thr T $H-C-C^*$ NH_3^+ OH	101 CH ₃	.1 5.9	2.09	9.10						
Asparagine ^e COO^- Asn $H - C - CH_2 - C$	O 114	.1 4.3	2.14	8.72						
$\begin{array}{ccc} & \text{NH}_{3}^{+} \\ & \text{Glutamine}^{e} & \text{COO}^{-} \\ & \text{Gln} & \text{I} \\ & \text{Q} & \text{H} - \text{C} - \text{CH}_{2} - \text{C} \\ & \text{I} \end{array}$	$ \begin{array}{ccc} & \text{NH}_2 \\ & \text{O} \\ & \text{H}_2 - C \end{array} $	3.1 4.3	2.17	9.13						
Tyrosine COO^- Tyr Y $H-C-CH_2$	NH ₂ 163 —OH	3.2	2.20	9.21	10.46 (phenol)					
Cysteine COO^- Cys C $H-C-CH_2-SI$ NH_3^+	103 H	1.9	1.92	10.70	8.37 (sulfhydryl)					
Amino acids with charged p	oolar side chains									
Lysine COO^- Lys $H-C-CH_2-C$ NH_3^+	128 H ₂ -CH ₂ -CH ₂ -NH ₃ ⁺		2.16	9.06	10.54 (ε-NH ₃ ⁺)					
Arginine COO^- Arg $H - C - CH_2 - CH_2 - CH_3$ NH ⁺ Histidine COO^-	NH_2 156 CH_2-NH-C NH_3^+	5.2 5.1	1.82	8.99	12.48 (guanidino)					
Histidine ^f COO ⁻ His H $H - C - CH_2$ \downarrow	137 137 137 137	2.3	1.80	9.33	6.04 (imidazole)					
Aspartic acid ^e COO^- Asp D $H-C-CH_2-C$ NH_3^+	O 115	5.1 5.3	1.99	9.90	3.90 (β-СООН)					
Glutamic acid ^e COO ⁻ Glu E H C NH ₃ H NH ₃	O 129	0.1 6.3	2.10	9.47	4.07 (γ-COOH)					