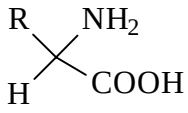
AMINO ACIDS

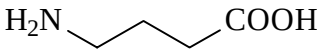
Jack DeRuiter

I. Introduction

Amino acids are organic compounds that consist of a basic amino group and acidic carboxylic acid moiety. The properties of these individual functional groups are described in more detail in the *Carboxylic Acid* and *Amine* Tutorials. Amino acids of concern for biochemical and drug science can be placed into two relatively broad categories: the α-amino acids present in proteins and other substances of biologic interest, and the non-α-amino acids. In the α-amino acids the amino and carboxyl groups are attached to the same carbon atom as shown generally below. In the non-α-amino acids the amino and carboxyl groups are attached to different carbon atoms:



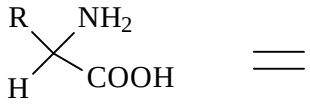
α-Amino Acids



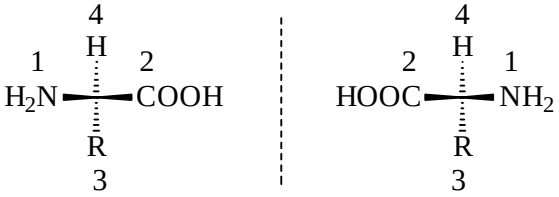
Non-α-Amino Acids

II. Stereochemistry

All of the α-amino acids found in proteins (except glycine) are chiral at the α-carbon, and thus two enantiomeric forms may exist. The α-amino acids found in natural proteins of human origin typical have the S-configuration (L-amino acids):



α-Amino Acids



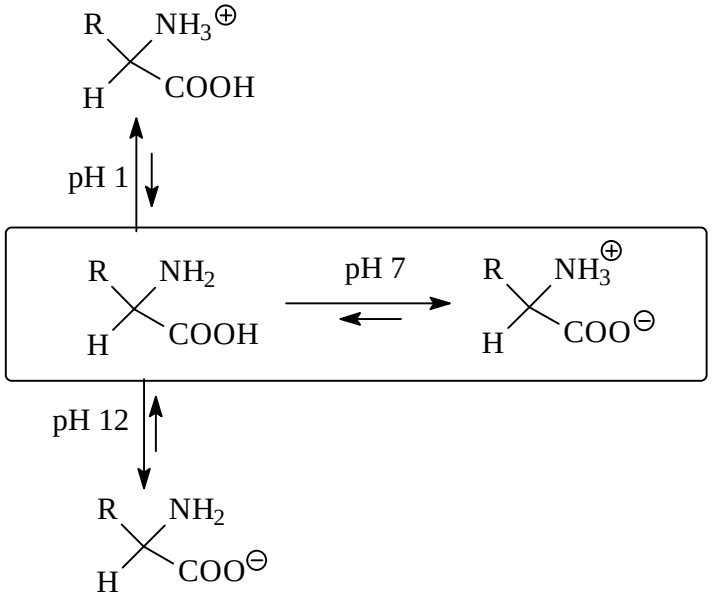
R-Amino Acid (D-form)

S-Amino Acid (L-Form)

As discussed in the *Stereochemistry Tutorials*, these enantiomers have similar physiochemical properties (pKas, water solubilities, etc.) in an achiral environment, but would have different properties in chiral environments (receptors, transporters, etc.)

III. Solubility and Other Physiochemical Properties

Amino acids contain an acidic carboxyl group with a pKa typically in the 2.0-2.4 range, and a basic (usually primary) amino group with a pKa typically in the 9.1-9.8 range (see Table 1). Thus at conditions of very low pH, amino acids exist primarily in the amine-protonated cationic form, and at conditions of very high pH they exist primarily in the carboxyl-unprotonated anion form (see Figure below). Under most physiologic conditions (pH near 7) these compounds exist in a dipolar or "zwitterionic" form:



The pH at which the concentration of cation is equal to the concentration of anion is called the isoelectric point (pI). At the pI value an amino acid may exist as an "internal salt" and thus the pI value typically corresponds to that pH at which the particular amino acid is <u>least</u> water soluble. For a simple amino acid (an amino acid containing only a single amine and carboxyl group), the pI value is calculated as a simple average of the pKa values for the basic and acid groups as shown by the general formula below and example calculation for glycine:

$$pI = [pKa\ (acid) + pKa\ (base)]/2$$

$$pI\ (glycine) = [2.34 + 9.60]/2 = 5.97$$

For amino acids with multiple acidic or basic functional groups (see text below and table at the end of this section), the pI value is calculated by determining the average of the pKa values for the two similar (acidic or basic) functional groups. For example, aspartic acid has two acidic groups with pKas of 2.09 and 3.86. At either of these pH values, the amino group (pKa 9.82) will be essentially completely ionized (protonated), thus this pKa value can be ignored. Calculating then using the two similar or acidic groups, the pI is:

$$pI\ (aspartic\ acid) = [2.09 + 3.86]/2 = 2.98$$

Similarly, the pI value for an amino acid with two basic residues such as lysine is derived from the average of the pKas of the two basic groups:

$$pI\ (lysine) = [8.95 + 10.53]/2 = 9.74$$

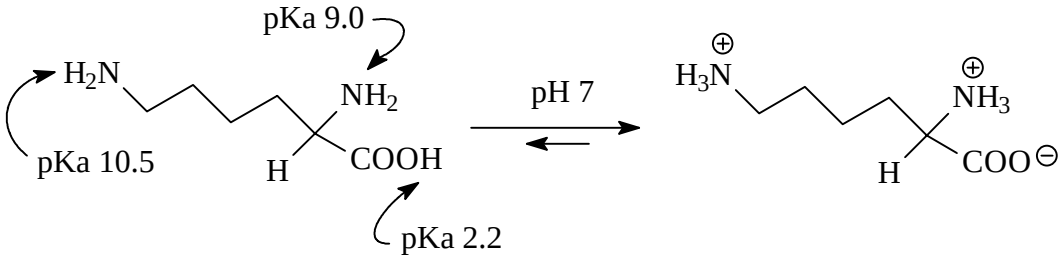
The <u>relative</u> solubilities (and reactivities) of different α-amino acids is more complex than represented by the discussion above due to the presence of R groups of different structure and physicochemical properties (see Table 1). The R substituents of the α-

amino acids vary considerably and may be 1). non-polar aliphatic hydrocarbon groups (Gly, Ala, Val, Leu, Ile), 2). aromatic rings of non-polar nature (Phe) or containing a polar group (Tyr), 3). heterocyclic rings of varying polarity (His, Trp or Pro), or 4). polar groups (Ser, Thr, Cys, Met, Asp, Glu, Asn, Gln, Lys, Arg). Furthermore, the polar functionality in R groups may be acidic, basic or non-ionizable (neutral) over the pH range of 0-12. Each of these types of R groups are summarized below:

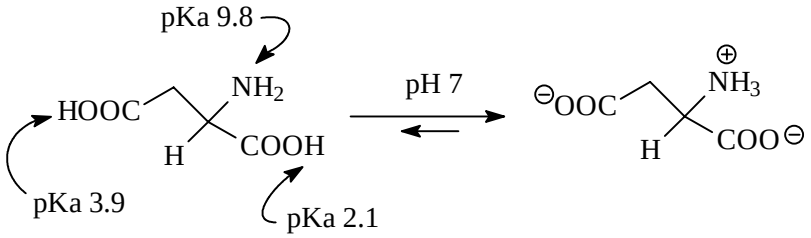
- Non-polar, Non-ionizable R groups: Gly, Ala, Val, Leu, Ile, Phe, Trp, Pro
- Polar, Non-ionizable R groups: Ser, Thr, Met, Asn, Gln
- Polar, Highly Ionizable Acidic R groups (pKas ≈ 4): Asp, Glu,
- Polar, Weakly Ionizable Acidic R groups (pKas >8): Cys, Tyr
- Polar, Basic Ionizable R groups (pKa 6-12): Lys, His, Arg

Among amino acids with non-ionizable R groups, relative water solubility is largely a function of the hydrophobic or hydrophilic nature of the R group. For example, as the hydrocarbon content (hydrophobicity) of the R group increases from glycine (R=H) to leucine (R = $(CH_3)_2CHCH_2$-) and phenylalanine (R = $C_6H_5CH_2$-) water solubility decreases (Table 1). Also amino acids with polar, non-ionizable functional R groups are more water soluble than the corresponding amino acid with a non-polar R group. This is illustrated by comparing the water solubility of serine (R = CH_2OH) to alanine (R = CH_3).

The solubility properties of amino acids with ionizable R groups (acidic or basic) are more complicated than those of non-polar and non-ionizable amino acids. For example, lysine contains an additional basic moiety in its R group and therefore can exist in a number of forms of varying ionization. At physiologic pH the α-amino group would exist primarily in the protonated, cationic form and the α-carboxyl group would exist primarily in the unprotonated, anionic form. The amino moiety present in the R group of lysine has a relatively high pKa (10.53) and thus is highly ionized at physiologic pH. Based on the presence of all three of these ionizable groups, lysine has a pI of 9.74 and therefore bears a net positive charge and displays relatively high water soluble at physiologic pH.



Similarly, an amino acid with an acidic moiety in the R group, such as aspartic acid (pI 2.98) and glutamic acid (pI 3,22), will be predominantly ionized and will display relatively high water solubility at physiologic pH:

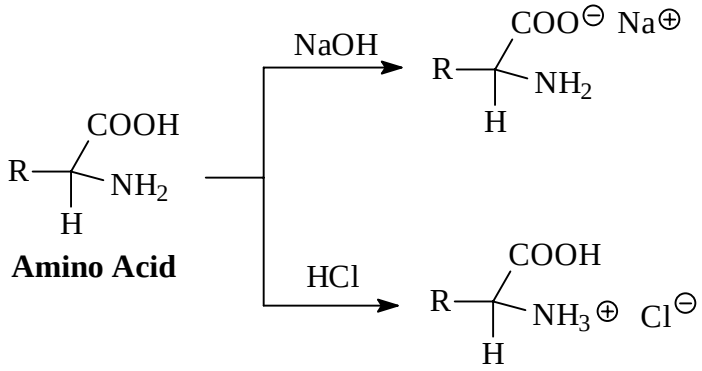


In the crystalline state the α-amino acids exist in their dipolar forms and therefore can participate in strong electronic intermolecular interactions resulting in relatively high melting points.

IV. Reactions of the Amino Acids

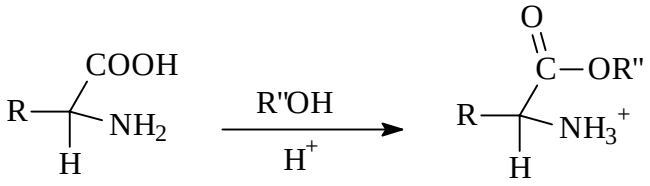
A. Ionization and Salt Formation

As a result of their relatively acidic and basic nature, amino acids will ionize if placed in an environment of adequate basicity or acidity. Thus amino acids in aqueous media of extreme pHs, such as aqueous sodium hydroxide (pH>10) or aqueous HCl (pH<2), will exist primarily in the ionized, conjugate base/acid form. Furthermore, this ionization enhances water solubility by providing an ionic center that can participate in energetically favorable ion-dipole interactions with water. Thus the water solubility of amino acids is "optimized" in aqueous environments where they exist primarily in one ionized form or the other (cationic or anionic). Similarly, salts can be formed from acid or basic moieties present in the amino acids and these salt forms will show enhanced water solubility.



B. Electrophilic Reactions:

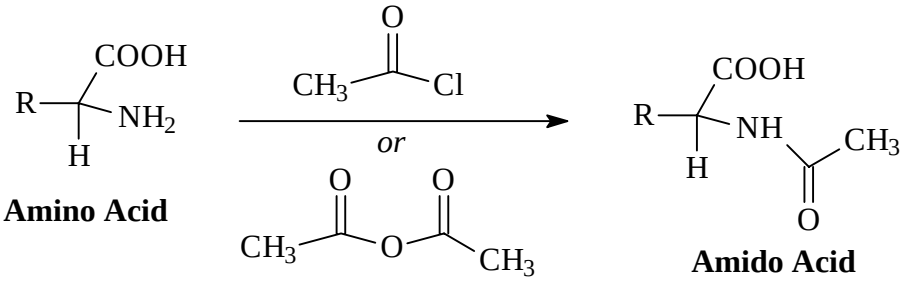
The carboxyl groups of amino acids are capable of functioning as electrophiles under the appropriate conditions, due to the presence of the dipolar carbonyl moiety. Under "dehydrating conditions" nucleophiles can attack the carbonyl group of an amino acid and displace the acid OH group as water (or another good leaving group). Such is the case in esterification reactions performed under acidic conditions. In these reactions an acid is treated with an alcohol (R"OH) which serves as the nucleophile, and an acid (H+) which serves as a catalyst. The acid catalyzes the reaction by 1). Further polarizing the carbonyl moiety through partial protonation, and 2). Providing a proton source for a hydroxyl leaving group (which "leaves" as water). The mechanism of this reaction is described in more detail in the *Carboxylic Acid* Tutorials



Amino Acid **Amino Ester**

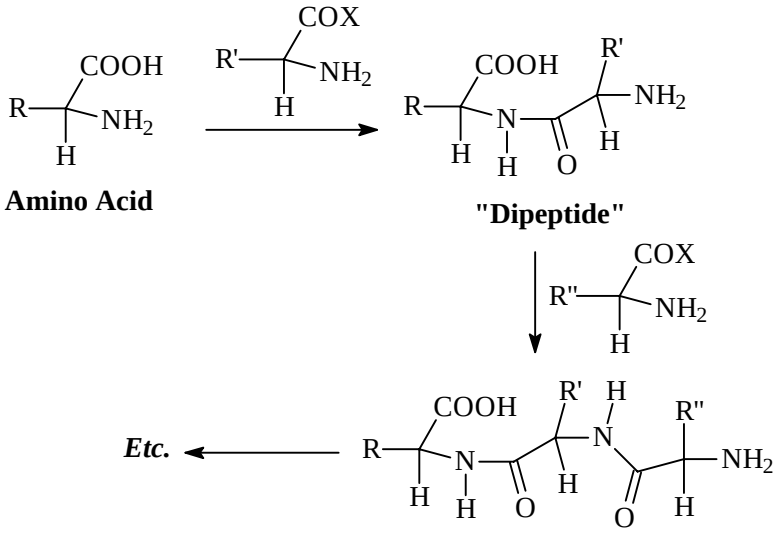
C. Nucleophilic Reactions:

As a result of their valence and bonding order, the amino moieties of amino acids have a NBEs and thus can function as nucleophiles (as well as bases). In the presence of electrophilic species, amines may react to form a substituted amine product; this reaction is referred to as a "nucleophilic substitution" or displacement reaction. For such a reaction to proceed the electrophile must have a "leaving group"; an electronegative functionality capable of displacement from the electrophile with the bonding electrons. Also, to form a neutral product, the amine nucleophile must have at least one hydrogen atom (a primary or secondary amine) to be lost as a proton. These principles are illustrated in the general example below and discussed in more detail in the *Amine* Tutorial:



Amino Acid **Amido Acid**

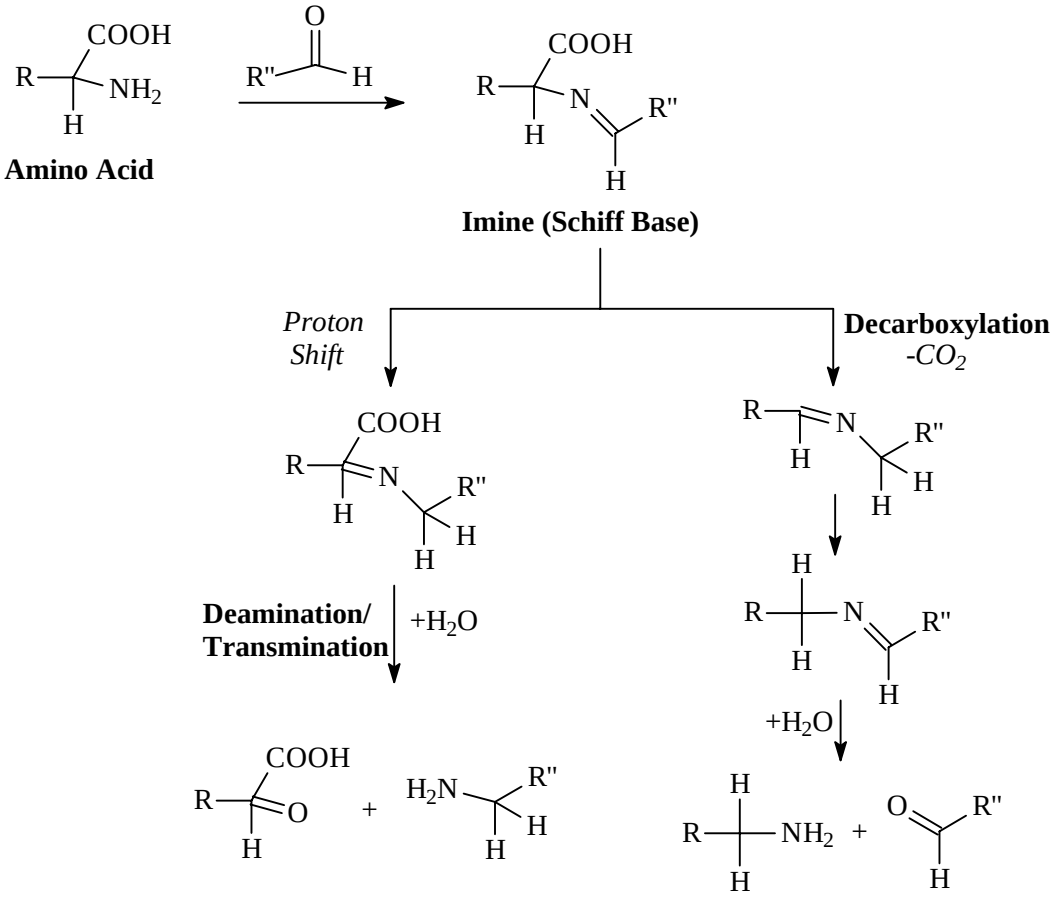
The reaction of nucleophilic amino acids with other electrophilic amino acid derivatives is employed in the synthesis of peptides or proteins as shown below. In this process two amino acids are covalently bound by a amide or "peptide" bond. Sequential reactions can lead to the formation of dipeptides, tripeptides and oligopeptides comprised of a number of amino acids. This process is accomplished biologically on the surface of ribosomes, and can be performed in the laboratory by protein synthetic methodology and instrumentation.



Amino Acid **"Dipeptide"** **"Tripeptide"**

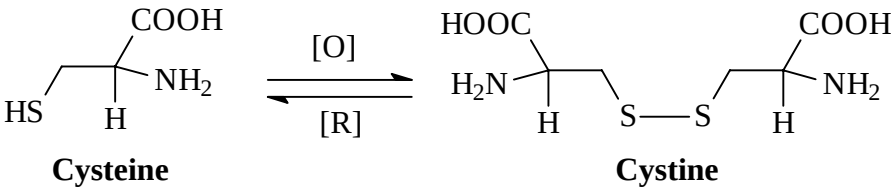
D. Decarboxylation/Deamination

Amino acids can undergo a variety of α-bond cleavage reactions under the appropriate conditions. For example, in the biologic environment the cofactor pyridoxal phosphate in association with decarboxylases or transminases can catalyze such reactions that are very important in amino acid metabolism. These reactions are illustrated generally below and will be discussed in more detail in the appropriate drug tutorials



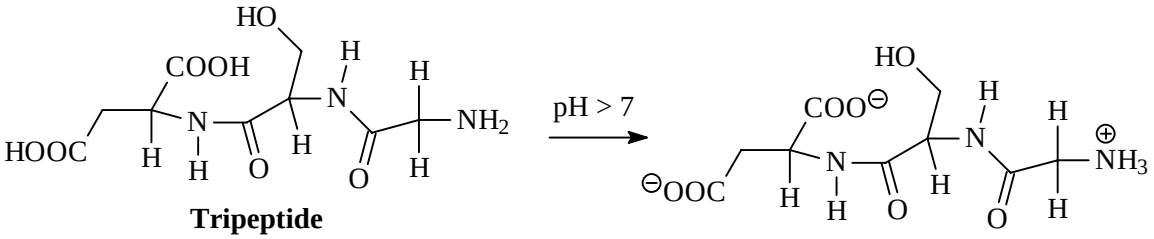
E. Oxidation/Reduction of Sulfur-Containing Amino Acids

Cysteine and cystine can be interconverted readily in the presence of a variety of oxidizing [O] and reducing [R] agents as shown below. These reactions can occur with the individual amino acids or when the amino acids are part of proteins. In proteins these disulfide bonds may be critical determinants of structure and function:



V. Protein Structure and Ionization

In peptides and proteins, the α-amino groups and α-carboxyl groups of the component amino group are involved in "peptide" or amide bonding and thus their individual acid and base properties are lost (except at the C- and N-terminals of the peptide). Thus the overall ionic character of a peptide or protein is determined by the acid and/or base character of moieties in the R side chains of the component amino acids. Consider the tripeptide example below. Other than the C-terminal acid and N-terminal amine, there is only one ionizable functional group: a side chain (R) acid group part of the C-terminal aspartate residue. Thus this protein is acidic and would bear a net negative charge at pH values > 4.



Tripeptide

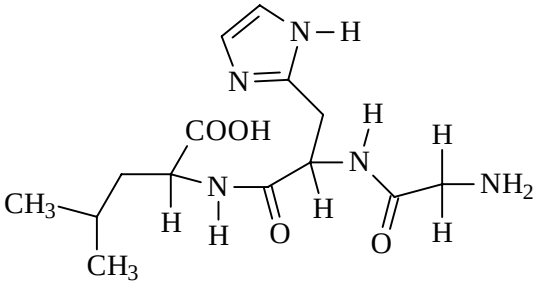
VI. Problems

1. Draw the “naturally-occurring” stereoisomeric form of phenylalanine (phe).

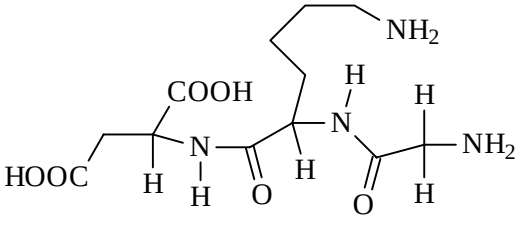
2. Draw the predominant ionic species of arginine at pH 4.

3. Draw the structure of the tripeptide gly-phe-lys (gly N-terminal, Lys C-terminal) at physiologic pH?

4. Draw the predominant species of each tripeptide below at pH values of 3, 7 and 12.



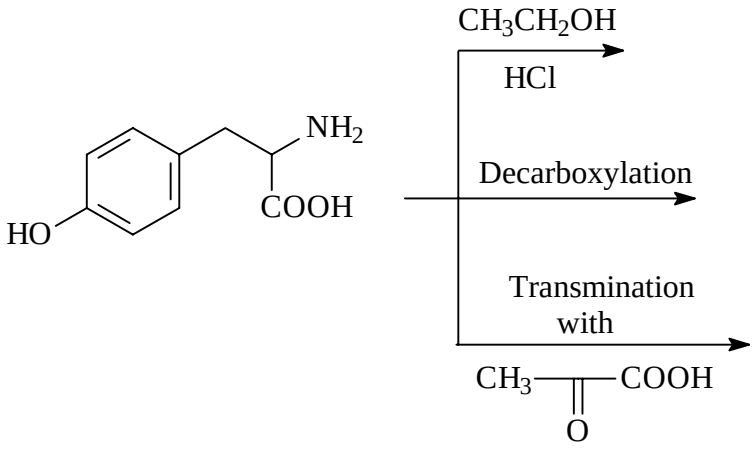
Tripeptide A



Tripeptide B

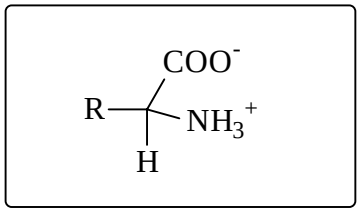
5. Show the disulfide that would form between the tripeptide (C)-gly-ile-cys-(N) and the tripeptide (C)-pro-cys-ala-(N).

6. Show the products for the following reactions with tyrosine

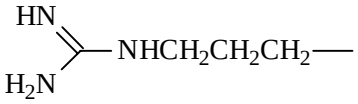
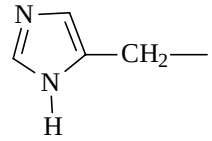
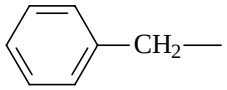
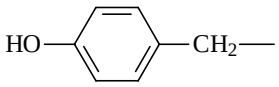


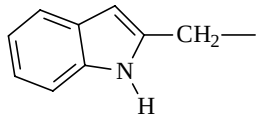
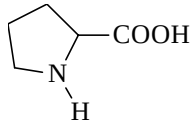
7. Show how the pI values are calculated for arginine and glutamic acid (see Table 1 for values)

Table 1. Amino Acid Structure, pKas and Solubility



Name	Structure (R)	pKa (COOH)	PKa $(NH_3)^+$	PKa (R)	PI	Solubility at pI (g/100 mL H_2O	MP, ^oC
Glycine (Gly, G)	H-	2.34	9.60	-----	5.97	22.5	292
Alanine (Ala, A)	CH_3-	2.35	9.69	-----	6.02	15.8	297
Valine (Val, V)	$(CH_3)_2CH$-	2.32	9.62	-----	5.97	6.8	315
Leucine (Leu, L)	$(CH_3)_2CHCH_2$-	2.36	9.60	-----	5.98	2.4	337
Isoleucine (Ile, I)	$(CH_3CH_2)CH(CH_3)$-	2.36	9.68	-----	6.02	2.1	285
Serine (Ser, S)	$HOCH_2$-	2.21	9.15	-----	5.68	4.3	228
Threonine (Thr, T)	$HOCH(CH_3)$-	2.09	9.10	-----	5.60	1.6	253
Cysteine (Cys, C)	$HSCH_2$-	1.17	10.8	8.3	5.02	Very	-----
Cystine (Cys-Cys)	-CH_2-S-S-CH_2-	1.65/2.26	7.86/9.85	-----	5.06	0.009	258

Methionine (Met, M)	$CH_3SCH_2CH_2$-	2.28	9.21	-----	5.06	3.0	283
Aspartic Acid (Asp, D)	$HOOCCH_2$-	2.09	9.82	3.86	2.98	0.4	269
Glutamic Acid (Glu, E)	$HOOCCH_2CH_2$-	2.19	9.67	4.25	3.22	0.7	247
Asparagine (Asn, N)	H_2NCOCH_2-	2.02	8.8	-----	5.41	2.4	236
Glutamine (Gln, Q)	$H_2NCOCH_2CH_2$-	2.17	9.13	-----	5.70	3.6	184
Lysine (lys, K)	$H_2NCH_2CH_2CH_2CH_2$-	2.18	8.95	10.53	9.74	Very	-----
Arginine (Arg, R)		2.17	9.04	12.48	10.76	Very	230-244
Histidine (His, H)		1.82	9.17	6.0	7.59	4.0	287
Phenylalanine (Phe, F)		1.83	9.13	-----	5.48	2.7	283
Tyrosine (Tyr, Y)		2.20	9.11	10.07	5.67	0.04	342

Tryptophan (Trp, W)		2.38	9.39	-----	5.88	1.1	283
Proline (Pro, P)		1.99	10.60	6.30	154	220	