Lecture 20A • 06/04/12

Amino acids

Classes of amino acids

This is the structure of the common alpha-amino acids. By convention, we write the carbonyl group up top, just like we do for carbohydrates. By convention, whatever R group or side chain is there, by default they're written in the vertical orientation. So, the amino group ends up on either the lefthand side or the righthand side of this Fischer projection. When oriented to the left, this is known as an L amino acid. There are the 20 common amino acids. There's actually more than that now; the other three are sulfur and selenium analogs of other amino acids. They're often found in what are known as extremophiles, bacteria or other critters which can only exist in conditions which, otherwise, are not conducive to life, but these creatures have adapted over time; part of that adaptation may be in the way that they've incorporated some of these "non-standard" amino acids. There are also what are known as the essential amino acids, which are amino acids that we cannot ourselves synthesize, or, for a couple of them, they're borderline cases where you can do it but it's tough for the body to do [or] it's easy for adults to synthesize but not children. The essential definition of essential amino acids is [they] are those amino acids that the body is not able to produce or produce easily; they must be supplied by other sources, which means food. There are 20 most-common amino acids.

The first class [of amino acids] is of alkyl substituents. I'll give you the simplest one. There are two ways that you can write this: one in which the amino group is in the plane of the paper, and one in which the R group is in the plane of the paper. Since we're about to compare different R groups, the rest of the structures I'm going to draw like this second one, so that that R group can be more clearly seen by putting it in the plane of the paper. We'll see that there are times where it's far more convenient to write it with the amino group in the plane of the paper. This is l-alanine [not aniline]. The next step up would be an alcohol-containing group. Serine, which is a methyl group with an -OH for the R group. Cysteine is a sulfur analog, so we have a sulfur-containing one. There are some amino acids that themselves have a second acid group; R will be containing a carboxylic acid group. Our simplest example of that is going to be aspartic acid, which is to have one carbon and then a carboxylic acid as our sidechain. Notice that these first four examples all have a methyl group as an R group and then they have -H, -OH, -SH, -COOH attached. [three- and one-letter abbreviations]

There's an amide version of the same amino acid we had just before; that's asparagine. We have amino groups. I've shown you one that has a second acid group; I'll show you one that has a second amino group; that is lysine. Lysine has four carbons and then the amino group as the R group. Back to a structurally-related pattern again, there are some benzene-containing side groups, by far the simplest one of which is phenylalanine – alanine was when you have just a methyl group, phenylalanine means, as it says, add a benzene ring to the methyl group of alanine. The last amino acid category are for amino acids that contain heterocycles, which means rings that do not just have carbon as one of the atoms in the ring. The best example of that is going to be proline, which has a five-membered ring; notice that means it is a secondary amine, which means proline has a very different chemistry than the rest of the amino acids. There's your eight classes.

Amino acids can in exist in an anionic form, a cationic form, and then a neutral form which we also call the zwitterionic form. We want to talk about this thing called the isoelectronic point, which is the pH at which an amino acid appears in the greatest percentage in its neutral form. If you had something like alanine, for example, pKa1 – which by convention refers to the acid dissociation – is equal to 2.34, and then pka2 – which is what refers to the amino group, or I should say the conjugate of the amino group – is equal to 9.69. At pH 2.34, what can be said about the forms of this amino acid that exist in solution? [The amino acid and it's acid conjugate will exist in equal concentration]. This that I've written here is the neutral form of the amino acid. As we start to go more acidic, you're going to have that amino group protonated. Although I've written this in the neutral form, the conjugate is therefore going to be this; the amino group has one more proton. It is the acid proton that comes off according to that particular pKa value. The neutral form, which is the zwitterionic form, will be in equal concentration with its cationic or acidic form. What's going to be true, then, at pH = 9.69? There's another proton that comes off, and there is another half-equivalence point effectively going on here. The jump in logic [to establish] this first relationship, when the pH is equal to the pKa, you have roughly 50/50 mix of the acid and its conjugate. That is true at [pH] 2.34, if we're talking about starting with the cationic form, then it's true at 9.69, if we're starting with the neutral form. We'll again have the neutral form show up in this relationship, but now since we're going to more basic conditions, we're talking about that amino group now truly getting deprotonated and ending up with the anionic form. At the first pKa, we have the neutral is equal to cationic [concentration], and at the second pKa the neutral is equal to the anionic [concentration].

If at the first pKa value the pH is there, you're 50/50 one way; if you go to the higher pH to match that second pKa, you're 50/50 the other way. So where do you think you maximize just having the neutral compound? In the middle of those two. That is exactly related to this graph that we generated. We had mole percent or mole fraction as our y-axis; we had pH as our x-axis. We talked about how if we started with low pH, then we're going to have essentially just the acidic form of the molecule which eventually disappears.

As that acidic form deprotonates, then we end up with the neutral form, which is going to reach some maximum value then, if you get too basic, it itself decrease again. After you pull of the first proton, the next dissociation is going to make the basic form of the amino acid. At these points where there's a crossover, that should be right at the 50% point. At that pH, that corresponds to pKa1, and at the appropriate pH that corresponds to pKa2; that's how we can find these crossover points knowing the pKa values. At the first pKa, 2.34, at that pH, we have 50/50 acidic and neutral; at the next pH value of 9.69 we have 50/50 neutral and basic. As you can see, right in between, that's the maximum amount of the neutral form, so for amino acids like this, the pI is equal to the average of the pKa values. Imagine that you were trying to deal with two simultaneous equilibria, the dissociation of one proton and then another. If you wanted to come up with some overall expression like that, then you would be multiplying equilibrium constants – if you're adding logs, that's the same thing as, in log form, multiplying things together. Short version of this: there's a pI, which is the pH at which you have this maximum amount of the neutral form, and for simple amino acids that just have one amino group, one acid group, you can figure out the pI by averaging the acids strengths of those two groups. For simple amino acids that do not have ionizable side chains, the pI is simply the average of the pKas of the acid and ammonium groups. We can't say amino cause it's already a base; what we mean is the protonated amino, which is called ammonium.

What happens if you have something like lysine or aspartic acid? It's got three pKa values to it: pKa1, which still refers to the acid group, is 2.09; pka2, which still refers to the amino group conjugate, is 9.82. pKa3 – the third value always refers to the side chain – is equal to 3.86. What this means is that there are four forms of aspartic acid; let's write those out. The most acidic form will be [the] fully-protonated form with a positively-charged nitrogen; we should have an overall cationic molecule when it's fully protonated. Take off one proton from this – which is going to be the first proton that comes off? Is it going to be the main acid group on the right, is it going to be the side chain on the left, or is it going to be the protonated amino group itself? The acid group on the right? Why is that going to be the first one to lose a proton? It's the strongest acid – pKa of 2 versus 4 verus 10, 2's a stronger acid, more easily dissociates, so the first proton that dissociates is the one on the right. Here is your neutral form; that, so far, is kind of like a typical amino acid. But now, if you lost one more proton, where's it going to come off of? The side chain, because it's got the second-lowest pKa value – low value, strong acid, wants to dissociate. If we take off one more proton, it's going to be from that carboxylic acid side chain; this is the anionic form. If we got it in basic enough conditions, eventually that amino group gets deprotonated; it's anionic, but it's got a double charge on it.

pI still refers to the point [pH] at which we have the maximum amount of the neutral amino acid. Why do we care so much about that pI point? Because that pH is going to be closer to biological pH than either of the functional groups by its own self – which, since each of these amino acids has slightly pKa values, there's slightly different proportions of each of these amino acids at any pH, including physiological pH, pH = 7.27. We're going to learn in a minute how to determine what exists more of, in terms of which of these forms, for a given amino acid at a given pH. That's why we're learning about these pI values, because the form of an amino acid exists in determines what charges it has on it, which means it determines how it interacts with other amino acids that might be near it, which means it affects protein folding, once we start letting these things do their thing in solution. pI is still supposed to be the point at which you have the maximum of the neutral form. Which pKa values would I use to arrive at that? pKa1 governs this first interchange; pKa3, the second; and pKa2. Adding emphasis to the pKa values that I wrote up here, pKa = 2.09 means it comes off first, cause it's the strongest acid. This comes off last, cause it is the weakest acid. Before, when we were trying to calculate pI, we looked at the acid group and the amino group. We have two acids groups; as we can see, the neutral forms between those two, so the pI is pKa1 + pKa3, their average, not pKa2 this time. A similar thing would happen if we had a molecule like lysine. Lysine is two amino groups, so the pI would be pKa2 + pKa3, the average of the strengths of the conjugates of the amino groups, because that's where the neutral form would lie. pI for this particular compound, aspartic acid, is 2.98, and the pI of alanine is 6.07.

What does the graph look like for this one? It's got three cross points instead of two. It would look something like this. Very similar to the first graph, it's just got three crossover points, because you've got four forms of the amino acid that you're dealing with.

Let's talk about one use of this. Let's say that we have a mixture of l-aspartic acid and l-alanine. If I were to say d-alanine and d-aspartic acid, how would that affect the pKa values? Because there's only one stereocenter, amino acids are enantiomeric. There's one that's called the l form, and there's one that's called the d form, just like sugars. But there's only one stereocenter, so they're easy to make mirror images of. The l forms the biologically active one; what I asked is what would happen to the pKa values if I used the d form instead? Same, because they're enantiomers, and every physical property of enantiomers is the same except optical rotation [and some other new property]. Let's say we had a mixture of those two. How could I separate them? Make one neutral, then do what? Electrophoresis.

Ninhydrin mechanism

Ninhydrin is a triketone, a trione, that technically exists in equilibrium with a hydrate form. [why is it the middle carbonyl that gets converted?] The ketone form, if it encounters any amino acid – I'll simplify and not show stereochemistry because, as you'll find out, it won't matter what the stereochemistry is. You go through a series of protonate-open-attack-deprotonate, or a variation thereof, and we make an imine. Turns out that this will undergo a carboxylate/pseudotautomerization.

This decarboxylates; this bond kicks on over to the nitrogen, which kicks on over to the carbonyl, which pushes the carbonyl open. You then have hydrolysis occur. Right now, that's an enol, so it is going to tautomerize [and] reform the ketone up top. At the same time, if we let it keep hydrolyzing, we do have an imine, so even though we formed an imine, we could also unform it, because remember that these proton-open-attack-deprotonate mechanisms are reversible. We hydrolyze, which means we end up with a derivative of ninhydrin. Notice the R group, which identifies the amino acid, is no longer present. Why are we using this reaget? The electrophoresis technique is one in which you can separate amino acids based on their charge, but once you separate them, you have to visualize them, because the amino acids themselves aren't colored [or not enough so to show up on the substrate], or if they are, they're various shades of clear or brown or yellow which – diluted and put on the paper for electrophoresis – you can't see them. What I'm about to show you is a complex that forms due to ninhydrin that very brightly purple in color – forms a complex called Ruheman's purple. That's the purple that forms on these electrophoresis plates once you treat it with ninhydrin, so you can see where the amino acids were, even if you can't identify, by the color, which amino acid it is. But, we'll know the position of the amino acids, which [we] can use that to identify them. Or, we'll then talk about chromatography – there, we'll use ninhydrin again as a visualizing agent to know when it had been eluted, brought off of a chromatography column. There's Beer's law, which lets you relate the absorbance of a solution with its concentration, so we'll be using ninhydrin for that as well.

Once the amino acid reacts and you have this chain reaction hydrolysis that occurs, there's still ninhydrin left in solution, which reacts with the amino-derivatized ninhydrin that just got formed, and that's what ends up causing the formation of this complex. All the amino acids that react with this form this particular complex.

_____

essental amino acids – those amino acids that the body is not able to product (or produce easily) and must be supplied by other sources (food).

20 "most common" amind acids
1) R = alkyl
2) R = alcohol-containing
3) R = sulfur-containing
4) R = carboxylic acid-containing
5) R = amide
6) R = amine-containing
7) benzene-containing
8) heterocycle

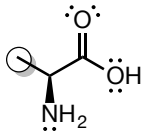
Isolectric point – pH at which an amino acid appears, in greatest percentage, in neutral form.

For "simple" amino acids that do not have ionizable side chains, the pI is simply the average of the pKa's of the acid & ammonium groups.

_____

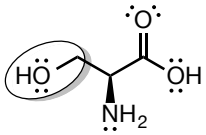
Structures *(remaining structures identical to lecture 19B & lecture 20B)*

06/04/12 lec • 1



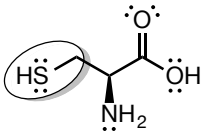
L-alanine

06/04/12 lec • 2



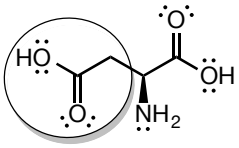
L-serine

06/04/12 lec • 3



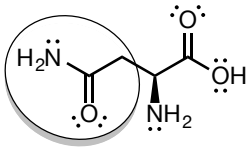
L-cysteine

6/04/12 lec • 4



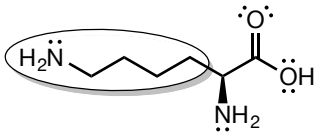
L-aspartic acid

6/04/12 lec • 5



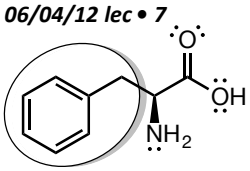
L-asparagine

06/04/12 lec • 6



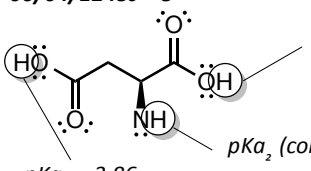
L-lysine

06/04/12 lec • 7



L-phenylalanine

06/04/12 lec • 8

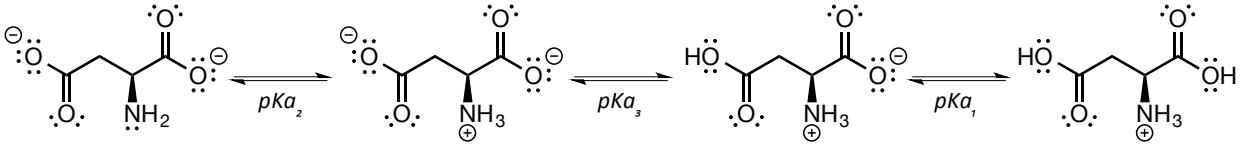


pKa_1 = 2.09 comes off first (strongest acid)

pKa_3 = 3.86

pKa_2 (conjugate) = 9.82 comes off last (weakest acid)

06/04/12 lec • 9



double anionic — *anionic* — *neutral (zwitterionic)* — *cationic*