Lecture 21B • 06/05/12

The terms polypeptide and protein are interchangeable, but a protein tends to refer to a large polypeptide. What does the word large mean? It's large. [what's the cut-off]

Say we had some form of protein that we were going to use this ion-exchange column chromatography approach on. Why is it that, in that approach, we will be able to eventually figure out the identity of the amino acids but not the order that they're in? What is the first step in doing this column chromatography approach? Do we put a protein directly on this column? We first hydrolyze the protein, we cook it up in a strong acid source to force all the amide linkages to break; a polypeptide is a polymer of an amide connected to an amide connected to an amide. The preparation step is to break down all of those connections between the amino acids. Once you've done that, then you place that onto the column. Once you break apart the amino acids, there no order retained whatsoever; it's just a whole bunch of individual amino acids. Chromatography can be used to determine the number and amount of an amino acid in a protein, but cannot be used to determine order, since to prepare for chromatography, a protein is first fully hydrolyzed.

A column is a tube that's filled with some kind of separation material. We have a sample that's loaded on top, and then a solvent is pushed through. The interactions of different amino acids with the material in the column causes some of those amino acids to progress more slowly and, conversely therefore, some are going to progress more quickly. Exactly which ones progress at which rate is going to be dependent on pH, because the pH establishes what proportion of cationic, neutral, or anionic amino acids we have. In theory, if we used more extreme pH values, we could also have a doubly-charged, either a double cation or a double anion, if we had one of our amino acids that had two acids groups on it like aspartic acid, or if we had two amino groups on it like lysine.

How are the results of this chromatography collected and then interpreted? As this stuff is coming off the column, you can't use ninhydrin to identify the amino acids. What is ninhydrin; where have we seen it used or how? It makes [amino acids] turn purple. Ninhydrin is a triketone that, when it reacts with an amino acid – it doesn't matter the stereochemistry and in a sense it doesn't really matter the amino – regardless of the amino acid, it's going to form the same complex. Notice that the R group does not become part of this complex; therefore, each amino acid makes the same derivative, therefore you can't use ninhydrin to identify an amino acid, it can be used to visualize amino acids. [first saw this in electrophoresis] You know you have an amino acid, and you know where you have an amino acid. [In electrophoresis], where we had a paper, across which we applied an electric field. Depending on the pH that was used, the mix of amino acids will have different proportions of their cationic, anionic, and neutral forms. Because there's a different charge on a different amino acids, they move different amounts in different directions, different distances once you do put a charge across the piece of paper. Once they move, you can't see them, so you use the ninhydrin to develop them, similar to TLC. In TLC, we often use light [or a stain]; in electrophoresis, we spray it with in ninhydrin.

For column chromatography, we also use ninhydrin, after the stuff comes off the column. Let's say that I had two different compounds: compound 1 and compound 2. Whether these are beakers or test tubes depends on, in real life, how much protein you've decomposed and how much quantity-wise of amino acids that you have. There's constantly solvent passing through this chromatography column. There's not always stuff coming off of the column; once you put that amino acid soup on there, it does take time even for the less binding of amino acids to make it through the column. If you were to have a line of these test tubes or beakers, you're going to one by one by one by one pass this continuously flowing stream of liquid, which could have a valve on it if you needed to stop it. There might be nothing, nothing, in one beaker you've got a little bit of the first amino acid, the next beaker you have a lot of amino acid, next beaker not very much again. Ideally, the next beaker there'd be nothing, then eventually, in an entirely different beaker, you'd have the next compound that comes through. These different beakers or test tubes could be referred to as different fractions, the different portions of solvent that are coming out. Fractions are different portions of the solvent eluting through the column that may contain one or more solutes. More would be bad, because we're trying to separate these things out. Because amino acids have been studied rather well, we have protocols developed that can be used to successfully separate out these amino acids. To tell you have an amino acid in a fraction, you could use ninhydrin.

But we can use ninhydrin not just as a crude qualitative test; we could also use it quantitative. There is this popularly named Beer's Law. Absorbance is proportional to three things: something to do with the material itself, it's a fudge factor known as the extinction coefficient; there's the concentration of the solution, the more stuff in it, the darker it should be; and then there's the path length the light itself passes through, the more material you pass through, the more light's absorbed, the darker the material appears to be. Once we have one of these solutions that we collect off of the column, then we could measure the volume of that solution. Ninhydrin's complex, Ruheman's purple, very well studied as well, you could get the absorbance of that purple solution. You would know the path length that you're using; you'd know it's ninhydrin's complex, so you'd know what the extinction coefficient is – measure the absorbance, you could therefore calculate the concentration. You know the concentration and the volume, you know moles. You figure it out for all of the amino acids present, and from that you could get the mole fraction of each amino acid in the entire polypeptide chain.

This method separates out amino acids, but does not retain any order information because you separate all of the amino acids before you do this technique. This technique depends on charge interaction. There is such a thing as size-exclusion chromatography, but size is not the main factor here. There are polarity effects, and you could do a column that separates more on polarity, but because so many of these amino acids are charged, that's why ion exchange material is used because you could more fully resolve or separate amino acids from each other. Once the amino acids come off of this column, the identity of the amino acid is determined by its retention time. Aside from that, we can use these solutions to determine how much of the acid is present; through a spectroscopic observation, we are able to come up with a concentration of the solution which, given the volume of solution, lets you get moles, which lets you get mole fraction.

What about true sequencing? [Let's cover] the Edman degradation, which allows individual amino acids to be cleaved from a polypeptide chain in order, which allows the sequence of a polypeptide to be determined. The reagent is phenylisothiocyanate. [isocyanates appeared in the Curtius and Hofmann rearrangements]. Phenylisothiocyanate means a sulfur analog [or isocyanates]. This will have a very interesting reaction with amino acid chains, with a very, very, very interesting mechanism. We're going to talk about the highlights of that mechanism. If we have an amino acid chain, what's it going to look like? It's going to have the amino end, then I'll just show that we have this carboxylic acid group. Given the phenylisothiocyanate, what's going to be the only logical reaction that happens? Isocyanates [and thus isothiocyanates], what did we previously react them with? What kind of product was created from that? Carbamic acid, which decarboxylates. The point is: the carbon in the middle is going to act like a carbonyl carbon or an imine carbon. These isocyanates are like a carbonyl – a ketone and an imine stuck together. The carbon's going to get attacked, which is the same thing that's going to happen with the phenylisothiocyanate. The first step is to make something that looks like this. The important point here is we first make what is called a thiourea. Urea is a carbonyl with a nitrogen on either end, so it's kinda like a double amide, you could say. Thiourea just means to substitute a sulfur [for the oxygen] instead.

This is where things start to get complicated. Imagine this: we have this nitrogen with a lone pair on it, which we know is heavily going to conjugate with a carbonyl, if we had a carbonyl. A carbon-sulfur double bond has very similar reactivity, so you could imagine that that nitrogen's got some pretty heavy interactions with that sulfur. Of course, we have this carbonyl next door as well, and that is an amide bond. But, we have the right kind of conditions where could have protonate-open-attack-deprotonate, protonate-open-attack-deprotonate, which, it theoretically would be possible, that we try to tautomerize, where the nitrogen would make a double bond with the carbon, pushing the carbon-sulfur bond open. There is this carbonyl right across the way here that the sulfur might be able to attack, which, if it did so, would end up pushing open the other carbonyl. This is not exactly a proper mechanism; more, I'm going to be showing you what happens. Again, nitrogen could try to push open that carbon-sulfur bond, which then can reach out and attack the carbonyl across the way and then push that carbonyl open.

So now, we've formed a ring. We have one carbon in which we have a sulfur, a nitrogen, and an oxygen all attached. That's going to be a really reactive intermediate, and it's kinda like if we were halfway through doing an esterification or deesterification. Imagine that the carbonyl reforms, and instead of kicking the sulfur back open, what if it kicked the chain out? Then we end up with the following: we end up with a five-membered ring that effectively is a thioester. There is a specific name for this – a thiazolinone. The "zol" ending has to do with one of these five-membered nitrogen-containing rings. The "one" part refers to the fact that that ring is also part ketone; "thia" is a prefix meaning the ring itself contains sulfur. So thiazoline means: five-membered ring with a nitrogen in it that also happens to have a carbonyl and a sulfur. How about the benzene ring? It's just a derivative of a thiazolinone.

Notice something really super-important: the rest of that polypeptide chain is kicked out during this mechanism. I started this whole thing out by saying there's going to be this technique where we can pull one amino acid off at a time; this is it, because as this process occurs – it's kinda like we're swapping out an amide for a thioester – the rest of the peptide chain is kicked out, which means we're dealing with just one R group. Since this reaction – until you put another phenylisothiocyanate in – it doesn't do anything else to the peptide chain, that means after we get this one compound of solution, then we could chew off the next amino acid and identify it, and then the next one and identify it. That's how this Edman degration is useful. We're not quite finished, because it turns out the thiazolinones are not isolatable. They form as an intermediate, but then it decomposes, because yes we exchange an amide for a thioester, be know that amides are less reactive than esters, so what if this compound opens again to try to form an amide.

At this point in the mechanism, the remainder of the polypeptide is ejected from the phenylisothiocyanate derivative. The phenylisothiocyanate derivative itself could be used to identify the amino acid that had been pulled off the chain. Effectively what happens is this: the thiazolinone can do a ring-opening. We form something along the lines of this: the ring opens back up again. Now, effectively, an amide is made again. This thing is called a phenylhydantoin. The main point of all of this is that this thing, that can be isolated, because each derivative is unique depending on which R group ends up on it, you can identify the amino acid.

[I] started this off with the Edman reagent, the Edman degradation. We produce a carboxylic acid derivative by reacting with this isothiocyanate.

There's a series of conversions between different carboxylic acid derivatives, which involve some ring openings and closings. The thiazolinone is an intermediate formed; it just happens to further decompose in solution to make one of these hydantoins that can be isolated, and because each one is different depending on the R group that was removed, that's how you can identify an amino acid. Since one amino acid is removed at a time, that means we can sequence an amino acid chain, starting from the amino end, going towards the carbon terminus. [There are far more modern methods]

Synthesizing amino acids

I'll start with a variation of the malonic ester synthesis called the phthalimidomalonic ester synthesis. Imagine we started out with diethyl malonate – but not diethyl malonate: what if we had alpha-bromodiethyl malonate. Is elimination in this molecule possible? [No] Is Sn2 impossible with this molecule? No, because it's just a secondary substrate, so it may react more slowly than a primary, but it'd still react. Now, we've seen previously the Gabriel synthesis, where we could deprotonate this nitrogen; I'll show it already deprotonated to then alkylate it to make a customized primary amine. What if I put these two different substrates together? Then this Sn2 reaction I'm referring to could occur. This is diethyl N-phthalimidomalonic ester – that still has an alpha proton on it, so you could use a base [ethoxide?] to deprotonate. Let's say we now had a primary alkyl halide. We could substitute and then, if we hydrolyze – once we hydrolize this phthalimide, we release a primary amine from it. Remember that when [a] malonic ester gets hydrolized, we make malonic acid which, with just a little heat and acid, will decarboxylate, so what you'll end up with is an amino acid. The only drawback of this particular approach is, as you can see, I didn't make one particular amino acid or another, I would end up making an enantiomeric mix. But, there are ways to separate out enantiomers.

____

Chromatography can e used to determine the types and quantities of amino acids in a polypeptide, but cannot be used to determine a sequence since to prepare for chromatography a polypeptide is first hydrolyzed.

Ninhydrin is used to visualize amino acids, but not identify them (product of ninhydrin has no R group).

fractions – different portions of the solvent eluting (coming out) from the column that may contain one or (non-ideally) more solutes.

Since the absorbance and volume of a fraction can be measured, and since the path length used to measure a sol'n would be known, the concentration (and therefore moles) of an amino acid can be determined.

– Remember: In chromatography, the identity of an amino acid is determined by retention time.

Sequencing

Edman degradation – Allows individual amino acids to be cleaved from a polypeptide chain in order, allowing sequencing of that chain.

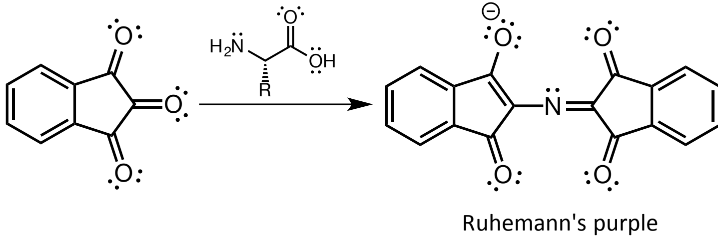
Reagent:

At this point of the mechanism, the remainder of the polypeptide is ejected from the PITC derivative. The PITC derivative can therefore be used to identify the amino acid that had been pulled off the chain.

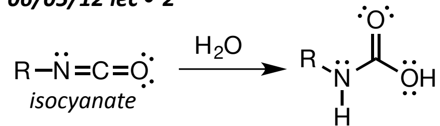
____

Structures

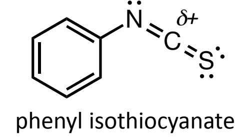
06/05/12 lec • 1



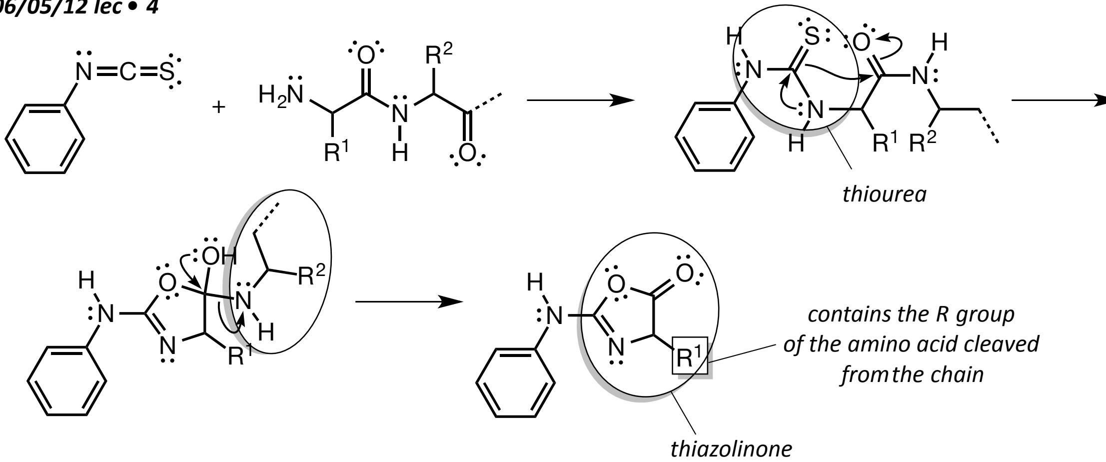
06/05/12 lec • 2



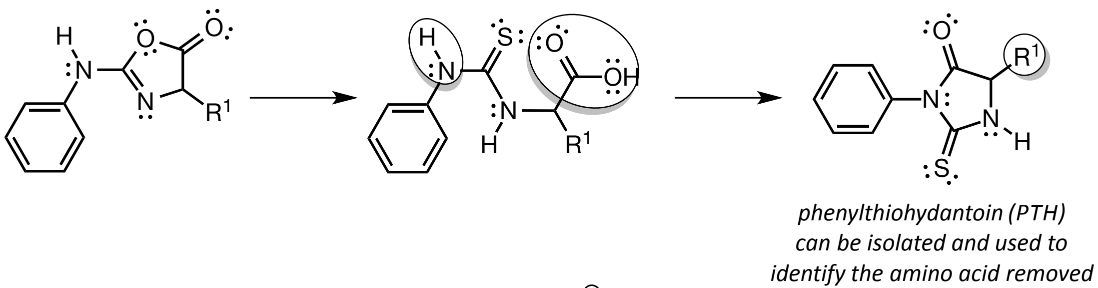
06/05/12 lec • 3



06/05/12 lec • 4



06/05/12 lec • 5



06/05/12 lec • 6