**ELIZABETH NOLAN:** What we're going to do today is just discuss a few aspects of cross linking. So we decided it was important to introduce this within recitations this year, because cross-linking comes up time and time again. And there's different ways to do this, and different strengths and limitations to different approaches. So I guess in just thinking about this, what is cross-linking?

So if you say, oh, I'm going to use a cross-linker for my experiment, what does that mean?

**AUDIENCE:** Forming a covalent linkage between two molecules of study.

**ELIZABETH NOLAN:** Yeah. So there's going to be formation of some sort of covalent linkage between two or maybe more-- right? Because some cross-linkers can have more than two reactive groups, OK, of study, right? So we're chemically joining two or more molecules. So why might we want to do this? What are possible applications?

**AUDIENCE:** Study protein-protein interactions.

**ELIZABETH NOLAN:** So that's one. So protein-protein interactions, right. And that could be identifying unknown protein-protein interactions or maybe you know two proteins interact, act but you don't know how, right? And you decide to use cross-linking as a way to probe that. So how might cross-linking help with studying a known protein-protein interaction?

**AUDIENCE:** Start getting an idea of where the proteins are actually interacting or which residues [INAUDIBLE]

**ELIZABETH NOLAN:** Yeah.

**AUDIENCE:** It could allow you to isolate them. [INAUDIBLE]

**ELIZABETH NOLAN:** Right. So maybe there's an unknown one, and you fish that out, because a cross-linker was used, right? And you know what one of them are. Or maybe, say, we know that these two proteins interact somehow, but we don't know how. So is it on an interface on this side versus

maybe the other side versus maybe behind the board, et cetera. And so, there's many ways to study protein-protein interaction. And really, how I'll present cross-linking today is in the context of this particular application, but there are many others.

But if we just think, we've seen a lot of protein-protein interactions in this course, right? So just even today, ClpXP is an example, right? We saw protein nucleotide interaction with the ribosome GroEL GroES is an example of protein-protein interactions, right? And they've been studied by many other methods, like crystallography for instance. But sometimes maybe it's not possible to get a structure, right? And you want to define an interaction surface or know exactly what residues are important.

So here, say, is protein-protein. But that could be generalized to any other type of molecule, like RNA, DNA, right? What about a single protein? So can you use cross-linking to learn more about tertiary structure, quaternary structure? So imagine for instance, rather than two separate proteins, we have one protein where there's some flexible linker. And we have reason to believe these different domains interact. But how do they interact? Again, is it something like this undergoes some conformational change and they're like this versus other possibilities here?

So what about just other applications of cross-linking chemistry before we look at some examples of molecules? So we can capture and identify binding partners, as Lindsey indicated. We can study known interactions. Where else could this come up? While it wasn't defined in this way, we've seen certain technology that takes advantage of cross-linking chemistry often.

**AUDIENCE:** Within the realm of biological things, it's used for-- I mean, if you want to find a functional root. So like bioaccumulation or general bioconjugate chemistry for [INAUDIBLE]

**ELIZABETH NOLAN:** Right. So general. Exactly, general bioconjugate or conjugation chemistry. So maybe you want to attach a tag to a purified protein. Maybe you want to modify an antibody. Similar chemistry can be employed. And likewise, even like from application standpoint, a mobilization. So say you need to make your own resin to do some sort of affinity chromatography and you want to attach a protein or an antibody to that, you can use the types of chemistry shown here.

So we're going to talk about a few different types of cross-linker and the chemistry, and pros and cons. And just as a general overview, I'll describe types. So we just heard the word homobifunctional. So homobifunctional versus heterobifunctional. OK. And this refers to the

reactive groups. So we need to talk about what types of chemistry is going to be used to do cross-linking. So this refers to reactive groups.

And then another classification will be non-specific versus specific. And so, this doesn't refer to, say, the chemical reaction between the cross-linker and whatever it's hitting, but rather whether or not the cross-linking reagent is site-specifically attached to a protein or biomolecule of interest or not.

If we just think about this non-specific versus specific, if we want to attach a cross-linker at some specific site in a protein, how can we do that? So think back to the ribosome discussion, where unnatural amino acid incorporation was not attached, but was introduced. So that's one possibility. If you have an amino azyl tRNA synthetase and a tRNA that can allow some sort of cross-linker to be introduced site-specifically, and it works for your experimental situation, you can do that. So we saw benzophenone, which is a cross-linker and the evolution of that orthogonal ribosome ribo-x. But let's say you can't do that, right? So for instance as far as I know, there's no tRNA AARS pair for benzophenone in a eukaryotic cell, right? Or maybe in some circumstance. What is something just using standard biochemistry you could do? So what type of residues can be modified in a protein?

**AUDIENCE:** Cysteine.

**ELIZABETH NOLAN:** Yeah. So cysteine, lysine. These are common side chains that are modified. And what would you say is more typically employed if you want to introduce a site-specific modification using chemistry?

**AUDIENCE:** Cysteine.

**ELIZABETH NOLAN:** Cysteine, right? So if you have an individual cysteine that's in the protein or maybe you use site-directed mutagenesis, you know where that cysteine is, and then you can modify it with some reagent there. We'll come back to that in a minute.

So in terms of reactive groups then on the protein, we can think about lysines, right? We have the epsilon amino group, cysteines. We have the thiol. What do we need to think about for our chemistry when thinking about these types of side chains and wanting to do a reaction? So under what conditions do we have a good nucleophile? Pardon?

**AUDIENCE:** [INAUDIBLE]

**ELIZABETH NOLAN:** Yeah. So we need to think about the basicity, right? The PKA of these groups, right? That's very key here for that.

What else do we need to think about? What other factors might govern reactivity, just thinking broadly? So PKA. For your amine, it will be type of amine. For a cysteine, redox will play a role, right? You can't have your cysteine and a disulfide. It needs to be the free thiol form. So these are all things to keep in mind.

So Alex has used a homobifunctional cross-linker. Why did you use a homobifunctional cross-linker?

**AUDIENCE:** It was to stabilize a nanoparticle.

**ELIZABETH NOLAN:** To stabilize a nanoparticle. OK. So very different type of application here.

**AUDIENCE:** Yeah, that's why I didn't mention it.

**ELIZABETH NOLAN:** That's fine. Yeah. We're not doing much with nanoparticles here. But let's say we want to use a non-specific homobifunction. So this was non-specific cross-linker to look at some protein-protein interaction, right?

So if we just suppose, for instance, we have some protein A and we think it interacts somehow with protein B, how can we use cross-linkers to study this? So let's take a look at an example of a homobifunctional cross-linker in terms of design.

So this one will be amine reactive. And its name is DSS here. So effectively, if we want to dissect this structure into different components, what do we have?

**AUDIENCE:** Two leaving groups kind of linking.

**ELIZABETH NOLAN:** So we have two reactive groups, or leaving group, separated by a linker. And in this case, we have two NHS or 6-cinnamyl esters, right? That are amine reactive. So what's the product of reacting an alpha amino group or a lysine epsilon amino group with an NHS ester? What do we get?

**AUDIENCE:** Amide.

**ELIZABETH** An amide, right? We get an amide bond. And then we have this linker or spacer region. OK?

**NOLAN:** Here. So two amine reactive groups and a linker, or spacer. And in this particular case, this linker or spacer is about 11 angstroms and it's flexible. And it's stable and cannot be cleaved. So in the case of Alex's project, this was used to stabilize a nanoparticle. Did you have a pure nanoparticle? Or was this in a very complicated mixture?

**AUDIENCE:** It's very not in this course.

**ELIZABETH NOLAN:** So what's going to happen if this reagent, say, is added to cell lysate? What are you going to get?

**AUDIENCE:** Random cross-linking with a bunch of different lysate proteins [INAUDIBLE].

**ELIZABETH NOLAN:** Yeah. So there's a high, high likelihood of a lot of different cross-links, right? So potentially a big mess, right? High likelihood, right? Because you have no control over where these reactive groups are going to hit. And do most proteins have lysine residues? Yeah. Do all proteins have an alpha amino group? Yeah. Well, some might be modified, but anyhow.

You have very little control with this type of reagent. So then the question is, if you use it, how are you going to fish out your desired protein-protein interaction? Or even if you're working with two purified proteins and they have multiple lysines, you can end up getting multiple cross-links, right? So maybe that's helpful for initially identifying that an interaction exists. But in terms of getting more detailed information in terms of how do these actually interact, that may be tough here. OK? So easy to come by, but potential complications.

Just in terms of thinking about this in the linker, why is it important to think about the linker and your choice of some reagent here? So what properties does the linker give?

**AUDIENCE:** [INAUDIBLE] to the link, then I guess its flexibility will determine how close the two proteins have to be in space for those to be [INAUDIBLE].

**ELIZABETH NOLAN:** Yeah. So there's some constraints imposed by the linker in terms of how close together or far away are groups that react. What else comes with the linker? How does it affect the properties of the molecule? Alex?

**AUDIENCE:** I was going to say it can dictate how likely you get a cross-linking on the same molecule between two amines. If you make it short enough, so that it can't reach the next lycine or something, then it can prevent [INAUDIBLE]

**ELIZABETH NOLAN:** Yeah. May be able to. So what's an inherent property of a molecule?

**AUDIENCE:** It might affect solubility.

**ELIZABETH NOLAN:** Yeah. Right. It may affect solubility. So linkers can be-- this is a bunch of CH2 groups, relatively hydrophobic, right? There can be more hydrophilic linkers or other strategies. And then the question is, does that matter? Does the solubility properties work with your experiment or not? But imagine if you want to do cross-linking in a live cell, you need that cross-linker to get into the cell. So you need to think about membrane permeability and what happens after that. Here.

So the linker is another critical aspect. And so, if you're ever working with a cross-linker, that's something you want to think about in addition to what types of side chains or what types of biomolecules do you want to modify. So let's look at an example of a heterobifunctional linker. It's not linker. Yeah. Well, it is cross-linker. OK. So this one will have a different type of spacer group. So it will be with a cyclohexyl. So what do we have in this case? Steve?

**AUDIENCE:** So you have an NHS ester and also a maleimide. And then the sulfonate group probably helps the solubility.

**ELIZABETH NOLAN:** Right. So there's a bunch of interesting aspects to this molecule. So we have the NHS ester to react with an amine. Right here we have a maleimide, which will react with thiols. So heterobifunctional, because there's two different reactive groups for different types of side chains. And then, as Steve mentioned, we have this group here. And so, this is to improve water solubility. OK. And then what do we have in this linker region?

**AUDIENCE:** A cyclohexyl instead of the aliphatic--

**ELIZABETH NOLAN:** Yeah. And what does that give?

**AUDIENCE:** Isn't it rigid?

**ELIZABETH NOLAN:** Yeah, exactly. Like cyclohexyl, right? Think about chair conformation, rather than what I have done here. But it will give a more rigid linker, and also shorter than what we see up here. So this is on the order of eight angstroms. So how might this molecule be used? What could you do with it that you can't do with this one?

**AUDIENCE:** Cross-link cysteine and lycine [INAUDIBLE]

**ELIZABETH NOLAN:** Yeah. Well, that's the first point, right? You can have two different groups. One end will react with a cysteine. One with some lysine. So is this specific, or non-specific, or both?

**AUDIENCE:** Probably depends on the context.

**ELIZABETH NOLAN:** Yeah. Right. Could depend on the context. And then from the standpoint of specific cross-linking-- which I would argue is a better use of this compound-- what can you do? Just imagine you have some protein of interest and maybe you want to label it here. And you have some side chain. So site-directed mutagenesis to put in a cysteine. And then you can modify that there, such that you have cross-linking reagent, right? And then you can imagine whatever your experiment is here.

So again, thinking about using this compound in, say, a complicated mixture, like a cell lysate-- you want to see if there's any binding partners or whatever. What's the limitation in terms of reactivity of this amine group that you would use in that second step? Where do you lack control?

**AUDIENCE:** You still can't control for the alpha for the N-terminal reaction, right?

**ELIZABETH NOLAN:** What do you mean by that?

**AUDIENCE:** So if the [INAUDIBLE] is free, then would you have comparable reactivity between the N-terminals, and, for example, your desired lycine?

**ELIZABETH NOLAN:** OK. So that could be an issue. So do lycines and N-terminal alpha amino groups have different reactivity? Do they have different PKAs? And is that something you could control? Maybe, maybe not. But more broadly than that, so you have an issue that it will react, let's say, with any amine, right? Can you control when it reacts?

**AUDIENCE:** To some extent [INAUDIBLE] pH.

**ELIZABETH NOLAN:** So what are you thinking?

**AUDIENCE:** If you--

**ELIZABETH NOLAN:** So if you think about just experimental design, right? And say you were to try to use pH to control reactivity-- and I'm defining this broadly-- reacting with any amino group. So we're not going to try to do something to selectively label one, right? This is reactive. It will react, right? So would pH change your whole buffer? Or pH change the cell lysate, and then switch to turn on reactivity? Probably not. Probably not, right? That, I'd say is not very likely.

So the issue I'm getting at here is that you have little temporal control or spatial control of an NHS ester. It will react with an amine provided your conditions are appropriate. So just getting back to this pH issue and a little digression, if you want to use something like an NHS ester in a test tube experiment, what you need to think about beyond pH? So what do you need to think about with the buffer?

**AUDIENCE:** You don't have something [INAUDIBLE] buffer [INAUDIBLE] so you might want to use the phosphate buffer, something that doesn't [AUDIO OUT]

**ELIZABETH NOLAN:** So this is a key point. You need to think about cross-reactivity with the buffer. So if you have tris buffer, you have amine. If you have a buffer that's like glycine, there's amine, right? And your buffer concentration in most instances is much higher than whatever the concentration is of the molecule you want to actually modify, right? If you think about 10 million molar tris or 75 million molar tris compared to micromolar or nanomolar of some protein, so you need to have a buffer that's not reactive. You need to have an appropriate PKA. Those are important considerations. You need to know that your reagent is good. Sorry, appropriate pH.

What about the thiol here? What do you need to think about if you're doing a test tube experiment and want to modify a thiol with a maleimide or something else, like iodoacetamide that we saw last time?

**AUDIENCE:** Buffers need to avoid DDT.

**ELIZABETH NOLAN:** So what's DDT?

**AUDIENCE:** [INAUDIBLE]

**ELIZABETH NOLAN:** Right. Or BME, beta mercapto ethanol. Right. Even before that step, what you need to make sure? So what if there's multiple cysteines?

**AUDIENCE:** That they're not [INAUDIBLE].

**ELIZABETH NOLAN:** Right. So either inter- or intramolecular, right? So if a reducing agent's added and the reducing agent is thiol-based, again, you're going to have much more reducing agent than your protein of interest, right? So you don't want your thiol-reactive probe to react with the reducing agent in the buffer here. So that needs to be removed. And then if you remove it, you need to ask, does the thiol stay reduced or is it susceptible to air oxidation? So these are just all practical considerations to keep in mind. If a reaction doesn't work, why doesn't it? And was it something that wasn't right with the buffers there? OK.

So back to this issue of not having much control about timing control for reactivity of these types of groups, what could be done to overcome that? So what other types of cross-linkers are out there? Yeah. Photo-active. Photo-reactive cross-linkers. So what's the idea here?

**AUDIENCE:** [INAUDIBLE] the appropriate [INAUDIBLE]

**ELIZABETH NOLAN:** Yeah. So what do we have? And what can we do? So just the first point to make is that we want to think about specific labeling here. So we can attach site-specifically to a protein or some other biomolecule, maybe it's bi-cysteine modification with something like a maleimide. Maybe it's unnatural amino acid incorporation. And it's chemically inert locally until irradiated. OK? And so, basically irradiating this photo-reactive cross-linker will activate the photo-reactive group, and then you get s-linking. OK. So this type of approach is often used to capture binding partners. It can be used in the test tube or in cells. What are the types of photo-reactive cross-linkers?

**AUDIENCE:** Aryl azides.

**ELIZABETH NOLAN:** Yeah. So aryl azides are one type. What's one we saw in class? Although we, didn't talk about photochemistry. Yeah, benzophenone. And there's some other examples. So what's another example?

**AUDIENCE:** Fluorinated [INAUDIBLE]

**ELIZABETH NOLAN:** Yeah. So they fall in here, right? So we can think about, either just phenyl azides or fluorinated phenyl azides. So another way to do this is to generate carbenes via diazirines here. We'll pretty much focus on these types, which are major types. So where did this idea come from? How new is this type of work to stick a photo-reactive group on a protein, and then use it in a cross-linking application? And where did the idea come from in the first place?

What types of chemists often study photochemistry?

**AUDIENCE:** DNA [INAUDIBLE]

**ELIZABETH NOLAN:** More broadly. So physical organic chemistry, right? There's a whole component of photochemistry there. Let's take a vote. 2000? First photo cross-linker. 1990? '80? '70? '60? Just no clue? So around 1962 was the first paper using a photoreactive group on a protein here, Westheimer. And then Jeremy Nulls in 1969 was the first example of an aryl azide. OK? So this work came out of physical organic chemistry and at a time where physical organic chemists were transitioning into enzymology.

So we don't have time to go into a lot of the photochemistry of these different moieties, but it was quite rich there. So how does this work? What types of reactions and groups get modified here in the cross-linking? So let's think about them. So let's consider an aryl azide. So what happens when aryl azides are irradiated with UV light?

**AUDIENCE:** Took all of the nitrogen gas. Get a nitrene.

**ELIZABETH NOLAN:** Get a nitrene. Yeah. So if we just think about nitrenes for a minute, what types of chemistry do nitrenes do? Are they reactive? So can they insert into C-H bonds? N-H bonds? Add to double bonds? Can they do other things as well?

OK. So here we have our protein. What's going to happen? As Steve said, we're going to generate a nitrene. So how does that happen? We irradiate with light to get our nitrene.

So what happens with these aryl azides is some interesting photochemistry when you're at, say, room temperature. So rather than this nitrene reacting, say, with a C-H bond or an N-H bond, it actually undergoes a ring expansion. So what we get-- and this is very fast. So on the order of 10 to 100 picoseconds. So this happens before it has a chance to react with something else. OK. To give us this intermediate.

OK. And so, this species has very different chemistry than a nitrene. And what happens is it will react with nucleophiles. So imagine our amino group to give the cross-link.

So this pathway is the dominant pathway if just an aryl azide is used here. To think about from the standpoint of wanting to do cross-linking. So let's say you attach this aryl azide to a protein of interest, and then you irradiate with light and look for it to cross-link with something, is this an issue? It will form a cross-link. Would you rather have a nitrene reactor or this seven

member reaction with a seven membered ring?

**AUDIENCE:** [INAUDIBLE] nitrene, but it would depend on what you're actually looking at, like what you were investigating.

**ELIZABETH NOLAN:** OK. So why would you argue for the nitrene?

**AUDIENCE:** Because we were talking about the nitrene does have the capacity to do a [INAUDIBLE] So if you wanted to do something like that kind of chemistry, then having this be the dominant pathway would be inefficient.

**ELIZABETH NOLAN:** Yeah. There's a lot more C-H bonds than there are lysines or N-termini. So that's one aspect. We've lost that chemistry. And then to another point, how well did these reactions work? So nitrene reactions are very fast. Relatively speaking, this is kind of sluggish there.

And so the question is, what can be done in order to improve upon this? OK. And Steve mentioned these fluorinated phenyl azides there. And so, photochemical work, unrelated to any sort of biological cross-linking chemistry, showed that if you fluorinate aryl azides you can get nitrene reactivity, rather than this other pathway here. OK? And so, if we just take a look at that, what happens?

For instance, imagine we have this tetrafluoro analog here. We can imagine irradiating this and getting to our nitrene. I'm going to skip the steps. OK. Now what can happen? Imagine we have some C-H bond nearby. We get this cross-link. And this reaction is very, very fast here. Very, very fast.

Can ring expansion occur in this situation? OK. So I'm pointing this out, because the language in the packet was a little strong. If there is something for this nitrene to react with nearby, it will react. But this can undergo ring expansion. It's just much slower than the case above. So the studies I've read say about 170-fold slower there. So it's not that the pathway is completely blocked. It also too depends on the experimental conditions. But anyhow, this is quoted to be near diffusion controlled here for that.

I mean, this is pretty interesting when you think about it, right? Because aryl azides, they can be fed to cells to do unnatural amino acid incorporation, right? They're used in click chemistry for instance, types of conjugation chemistry. But here, the photochemistry can be taken advantage of to give a cross-linker that can be controlled in a temporal manner there.

So what about the benzophenone? What does benzophenone react with after being irradiated with light? So imagine you have your protein. And maybe in this case you did unnatural amino acid incorporation to site-specifically attach a benzophenone. What happens? What happens is that there's formation of a triplet diradical. And what will this do? Here, it's going to react with some C-H bond to get the cross-link.

Let's say you have this guy here and you want to do a cross-linking experiment. So we can imagine some different possibilities. What do you think you'll get out? Are you only going to get out your desired cross-link? What might happen?

2:00, 1:50 on a Friday. Let's get some jumping jacks. Come on. Should I dismiss all of you, because there is a major energy low today, I have to say there. Yeah. Do you think you'll get one product, 10 products, 100?

**AUDIENCE:** There will be lots of side reactions.

**ELIZABETH NOLAN:** Right. There's still the possibility for many side reactions, right? And you always need to be aware of that. So if you cross-link something, the next question is, is this something that's actually relevant or not? Or is it an artifact there? So the analysis can be very complicated. And so, that's just something to think about. Say you have cross-linked species from cell lysate, what are you going to do to analyze that?

Just think about some of the things that have come up in other contexts here. We talked about protease digest and mass spec for looking at substrates of GroEL GroES, that's something that can be applied. And there's many sophisticated new tools to get a lot of information out of the mass spec, which we won't talk about. But having tags within the cross-linker, right? So then you need to ask, how well is the coverage going to be? So even after this step, there's a lot more work, which we won't go into details in this recitation today.

What about inherent efficiency of cross-linking in terms of these benzophenone versus the aryl azides? We want to think about relative cross-linking efficiency. Any sense of that?

**AUDIENCE:** I think the benzophenone compared to the diazirine is a lot less efficient. I don't really know [INAUDIBLE]

**AUDIENCE:** I have a question. When we're talking about efficiency, is it purely based on the speed of this reactivity? Or is it also taking into account the different cross-reactions that could occur?

Because it seems like there are more possibilities for more cross-reactions. Even though it might be more reactive, it's not--

**ELIZABETH NOLAN:** Yeah. The former, right? Just thinking about the reaction. There's the possibility of cross-reactions for all of these. They're highly reactive. A nitrene is highly reactive. The benzophenone triplet diradical is highly reactive. A carbene, if you're going to get that from some diazirine is very reactive. And yes, it's something important to think about in terms of your experiment. What is the relative efficiency of the reaction? So I said that aryl azide is a little sluggish compared to the others. Something to consider, right? You know what is the timescale of whatever it is you're trying to trap.

So the wavelengths. What is it about these wavelengths that might be undesirable?

**AUDIENCE:** For in vivo studies, one shifting towards UV means that you can have issues undesirable, like DNA cross-linking stuff, but also it means that it's not going to have deep penetrants [INAUDIBLE] shift towards [INAUDIBLE].

**ELIZABETH NOLAN:** What wavelength would you like?

**JOANNE STUBBE:** I would like it around 650. These are all UV visible interface. And you have hundreds of things that absorb length are very incredible inefficient. [INAUDIBLE] Most people never identify what they get out of the other side. They just see two things stuck together, and that's the extent of it. They never describe the molecular details.

**ELIZABETH NOLAN:** So let's actually just close--

**JOANNE STUBBE:** [INAUDIBLE]

**ELIZABETH NOLAN:** Right. So one of the questions I asked in the discussion section, is it worth the effort if you're going to site-specifically put in a cross-linker? And imagine you find this protein-protein interaction, if one chooses, you can do quite a bit more experiments in terms of where you place this cross-linker and mapping out that interaction region there. And so, that's I think also just a take-home is often you need to put your reactive group in more than one place to really get at the answer to the question you're asking. And so, there's folks around doing that there. But is it 20 positions? Is it 10? Is it 50? Because if you don't know at the beginning, you may

need to do a lot of just systematic trial and error for that. Yeah.

So I think you should all read the packet. And there are some suggestions for reading if you're curious to learn more, one of which is a manual from Thermo. So often, the companies give a lot of good general background information, and there's many different types of chemistry included in that as well. I'll also point out, Ed is here for those of you who don't know Ed. So he'll be presenting next week on cryo-EM. And you should definitely read the fatty acid synthase paper beforehand. The structures are incredible. And fatty acid synthase serves as a base for our discussions of polyketide and polyketide synthases, which is where we'll begin module four in thinking about the biosynthesis of natural products there. OK. Have a good weekend.