David Davis, Ph.D.

Behind the Mask

September 24, 2021

Barr: Good afternoon. Today is September 24, 2021. My name is Gabrielle Barr, and I'm the Archivist at the Office of NIH History and Stetten Museum. Today, I have the pleasure of speaking with Dr. David Davis. Dr. Davis is a Staff Scientist in the HIV and AIDS Malignancy Branch at the National Cancer Institute. Today, he's going to be speaking about some of his COVID-19 research and experiences, particularly with the main protease [Mpro]. Thank you very much for being with me.

Davis: You're welcome.

Barr: To begin with, can you provide a basic explanation of what a protease is, what they do, and discuss the proteases created by the SARS-CoV-2 virus?

Davis: Sure. There are actually all kinds of proteases in biology. Proteases are a class of proteins which actually cut up other proteins. For example, in the real-world experience, when you eat, there are proteases that are released into the gut that chew up the proteins that we eat. Cells also have proteases that chew up the proteins in cells that either become damaged or non-functional or need to be recycled. However, proteases aren't just to chew up other proteins for digestion. Sometimes it converts one protein, like a form of insulin that's inactive, to an active form. It can also be involved in blood clotting—and in just so many different processes in biology that they are really a very general class. The thing that's important to understand about them is that they cut up other proteins, and there are a variety of different ones and how they do it. That's why you can have a protease, say from your body, cut a protein and do it in a different way with different amino acids than, say, a virus protein. But in general, that's what proteases do.

Barr: Okay, that's interesting. What is important about the Main protease associated with SARS-CoV-2.

Davis: One of the proteases from SARS viruses is also called the Main protease of the SARS virus. The SARS virus actually has two proteases, but the Main protease is responsible for cutting up very large poly protein into smaller pieces. I think the best way to think about this is to think about the building of a wooden house. If you're going to build a wooden house, you collect a whole bunch of wood, then you get a saw, and you cut up the wood into different size pieces in order to build the house. A virus does the same thing. It generates these very long pieces of protein that need to be cut up into smaller pieces that will make up the virus. Main protease is responsible for cutting up the very long protein that's first generated. That long protein comes from the cell that synthesizes the protein off of the viral RNA. When the virus comes in the cell, it has this RNA, and this RNA encodes for the long protein. That long protein is made, and within that long protein is the main protease. The

main protease then starts cutting up all the specific places—very specific cuts—in order for the proteins of the virus to put itself together and then exit the cell and then infect other cells.

Barr: With SARS-CoV-2, there's eleven cleavage areas or something like that. Is that considered a lot, a little, or average?

Davis: Well, I would say it's average. Many viruses generate about that number—anywhere from 10 to 20 cuts. That's pretty normal.

Barr: In your research, you look at the protease, but do you also look at those cleavage areas?

Davis: I don't. I didn't in my research, but that has all been sorted out from the first SARS virus. The main protease was already known from the SARS-1 virus, and this particular protease is only a few amino acids different from SARS-2, so there's already a wealth of information that's known about the protease and where it cuts. In this particular protease, interestingly enough, it cuts after glutamine residues. That is super unique in that this protease cuts after glutamine residues, and there is no known mammalian protein that does that. That really gives us the opportunity to target the main protease, specifically to block the virus. Imagine if you took the saw away from the person building the house. They wouldn't be able to build the house. Stopping the saw, or stopping the main protease, is the goal here for therapeutic intervention. That's why we were interested in studying the main protease, because we thought if we could learn some biology about it, it might also inform us about how we could develop therapeutics.

Barr: Can you define glutathionylation, and what it means for the Main proteus?

Davis: Yeah, sure. Glutathionylation is a mouthful, for sure. Glutathionylation is simply a process where there's a tripeptide produced in almost all living cells. This tripeptide produced is called glutathione. That peptide, while very simple, has a cysteine residue in it. That cysteine residue is very important because it maintains the health of the cell. You could imagine if the cell didn't have glutathione and it was undergoing oxidative processes, the oxidative processes would damage the proteins, the carbohydrates, the cell, and DNA. With glutathione, it can kind of soak up the oxygen radicals and stuff and maintain the health of the cell. Now, in the process of doing that protection—of soaking up the radicals—it leads to the formation of a glutathione attaching to other proteins who have cysteines. It temporarily forms a disulfide bond that attaches to these other proteins. It's believed in the field that this is a way of protecting the protein from being damaged otherwise. In other words, if you don't have a cysteine, then an oxygen species could damage the protein. But if you have a cysteine just form this covalent bond that's temporary, and that can be reversed, then the protein is restored, and its activity is restored as well. Glutathionylationis a process that takes place where some proteins get glutathione attached to them.

Now, my background in my studies of HIV protease, I learned all about this and studied this. And we found that this occurs with the HIV protease, so we decided to explore that with Mpro. We found that, actually, you can do this process experimentally in the lab. We can take the main protease, which we make from bacteria, isolate,

and purify that, and then we can modify it with glutathione. Then we can study its activity. There are enzymatic activity assays that have been developed for the previous SARS protease. What we did is we went ahead and modified the protease with glutathione and then we studied whether or not it was inactive or not. We found that it was inactivated using this process, and we found that two or three of the cysteines were becoming modified. Then we subsequently learned, as you know in the paper, that the cysteine 300 is the cysteine that's sensitive to modification. If the protease is modified at cysteine 300, then the protease can no longer come together and dimerize so it remains inactive.

Barr: Could you define what dimerize means for the audience?

Davis: Yeah. It's very, very common in viruses for their enzymes to require dimerization for activity. Unlike some mammalian proteins whose proteases can be just monomers and active, for whatever reason viruses generate their enzymes as to form dimers. If they don't come together and dimerize, they're not active. It just turns out that both HIV and this SARS virus encode proteases that require dimerization for activity. That was already known for SARS protease, but we didn't know whether or not modification was going to do anything. We found that if you modify cysteine 300, then the two subunits of protease can't come together. They can't form the dimer. Think of it as a pair of scissors but the two halves stay apart, so it can't cut up the proteins, and so they're inactive.

Barr: That's really a great finding. What were some of the challenges that you encountered in your study?

Davis: Probably the greatest challenge of all was that this protease contains 12 cysteine residues out of the 306 amino acids that it is, whereas HIV protease only had two. It was easy to modify this enzyme with glutathione. But identifying which of the cysteines were being modified and leading to that inactivation—that was the challenge, because all of these cysteines are in what's called their reduced form. And all of them can be reactive. When you do these chemical analyses to look at the modification, you've got to make sure that you're not leading to artifacts and things like that. It was a challenge to identify which cysteine out of the 12 it was. We learned that because we saw there was this inactivation in leading to dimerization. We suspected that one of the cysteines is more reactive than the rest. When we lowered the pH, we were able to more specifically cause the modification of 300. That allowed me to then identify 300 as the target. But we also then made a form of the protease that didn't have cysteine 300; it had a different amino acid there—serine. When we put serine there, then we couldn't inhibit the dimerization or the activity of the enzyme anymore, and that's how we really identified that cysteine as being important.

Barr: What are some implications of your research? Obviously drug discovery, but are there people currently working on therapeutics based on what you and your team have found?

Davis: Well, if there are people working on that, based on what I've found and published, I'm not aware of them. But I wouldn't be surprised if some are. However, we are actually pursuing that now. In collaboration with NCATS [National Center for Advancing Translational Sciences] and some structural biologists, we are looking at that region of cysteine 300 and looking for compounds that can modify and occupy that target. We know simply

by modifying with glutathione—remember, that's only three amino acids, not a very big molecule—just by modifying that cysteine, we can inactivate Mpro. It is a target separate from the active site. It is a target that we can go after now. One of the things we can do is look for molecules that can modify cysteine residues, and there are a lot of drugs out there that can do this. Some of them are already approved for other therapies and stuff. One of the things is that it gives us an alternative target. The active site is always the most promising target that people go after. One of the reasons I looked at this wasn't so much as a target; I was really curious about the biology of the virus and whether or not this virus regulates its protease similar to HIV. You might ask yourself why the virus needs to regulate its protease and why it can't just be on all the time. One answer to that is, if you were going to be making more viruses in a cell, would you want to do that in a cell that's undergoing a lot of oxidative attack? I don't think you would, because then the virus coming out could actually not be infectious. The virus only wants to undergo processing in a healthy cell. A healthy cell would have the Mpro being active; in an unhealthy cell it would be inactive. That's what we think the biology of that's for. We haven't proven that; that's just a hypothesis. But at the same time, we can say, even though it's using that to control itself, we want to permanently turn it off. Instead of doing it through glutathione, which can come on and off, we can try to find something that targets and permanently inhibits it. And that's what we're currently looking into right now.

Barr: That's really interesting. Is there a possibility that this research on SARS-CoV-2 could be applied to other coronaviruses?

Davis: Yes and no. I mean, all of the coronaviruses that have been studied thus far that I know of have a main protease that dimerize. Their sequences in some cases are dramatically different. They're still cleaving after a glutamine residue, and they have similar areas of sequence, but they can be very different. Whereas SARS-1 and SARS-2 main protease only have about eight or so amino acids different that aren't even that important, they both have cysteine 300. There are these other coronaviruses of other animals—of cows, of bovine, or whatever—that don't have a cysteine 300 to target. You couldn't universally target cysteine 300, but you might be able to universally target the pocket once we eventually find something that could bind there and not have to necessarily need the cysteine residue there.

Barr: You've worked a lot with HIV, but you've also worked with other viruses as well. Can you talk about how your experiences with HIV and other viruses informed your research with SARS-CoV-2?

Davis: Yeah, well, if it wasn't for my experience on HIV protease, and HTLV-1 [human T cell leukemia virus type 1] protease, HIV-2 protease, Rous sarcoma virus [RSV] protease, and so on—if it wasn't for my experience with those, none of this work would have happened. It just so happens that my colleague down the hallway was producing main protease so they could try to target it with drugs. He was struggling with producing that protein and I offered to help sort out that problem. In the process of me sorting out the problem, I decided, "Hey, I'm learning about this protease, and I'm seeing it's very similar to HIV, and that it dimerizes and has these cysteines, and it cuts up a poly protein." I thought to myself, "Could this protease behave much like the HIV protease?" I really was surprised, honestly, when it came out to be the case that there is a single cysteine that if you modify, it inhibits its activity, and you can reverse it. You can reverse it with a cellular enzyme called glutaredoxin. It really looks like a biological process that's taking place inside the cell. It's not a unique thing to HIV but rather

extends to the virology community out there. It really seems to be something that when the viruses evolve, it's an attractive thing for them to control their enzyme activity in order to optimally produce virus. We hope to get to that to prove that at some point, but right now, we're just working with the enzyme. We don't have the virus in our lab right now or anything. Not here.

Barr: When did your colleague contact you? When did you start on this?

Davis: I believe it was in March. In March, I heard that he was working on it and having trouble. I heard that from my boss. I went to him and said, “Hey, I heard you having trouble, I can help you out.” He said, “Okay,” and then I went to my boss and asked if I can stay on, because we knew they might be having a shutdown coming. Within two weeks, I started working with him, and within two weeks, they had to shut down. I was able to stay on and keep coming in to work on the Main protease. I was the only one here in the lab and Haydar was the only one in his lab, and we were working together on this. I really consider myself pretty lucky in being able to do that.

Barr: That is very, very lucky. Can you talk about your contributions to studies of whether small molecules, such as GRL-0920, GRL-1720, and GRL-2420 block SARS-CoV-2 infections? Could you explain their basic characteristics? How did those research studies incorporate some of the principles you're investigating in your own research?

Davis: Like I mentioned earlier, most of the focus on developing an inhibitor to Main protease is to target the active site. That active site, in fact, involves one of the cysteines, so that is another similarity to the site we found. Now the active site cysteine, if you modify it, then the enzyme can't work anymore. It can't cut anymore. Back when SARS-1 was being studied, these drugs were developed to target the active site cysteine. They were shelved when SARS-1 went away, and no money was put into further research to develop drugs because they thought we didn't need them anymore. That coronavirus was gone, and we didn’t need it anymore. I think that's one of the mistakes of the funding for the science. Sometimes you need to fund science without knowing what your end result is going to be because it might be useful someday. Anyways, the funding dried up for that but there were some molecules that had potential. When this came along, Dr. Hiroaki Mitsuya’s group, along with Arun Ghosh of Purdue University, took those compounds off the shelf and started studying them with SARS-CoV-2 protease and saw that it inhibited that enzyme. I collaborated with them. They needed my help to show that when they added the drug to the enzyme, it was in fact attaching to the enzyme. I use mass spectrometry to do that. Then I subsequently use mass spectrometry to identify glutathione being attached to cysteine 300—that's a similarity there in that. And so, in a way, both of these targeting mechanisms involve a cysteine, but two different ways of inhibiting. One inhibits the active site, and the other prevents the two from coming together to form dimer. You can imagine a one-two punch—if you could have the two drugs inhibiting both processes at the same time, it might be very effective. These drugs are small molecules, much like most drugs that eventually get approved, like for HIV. They're small molecules, but they need to be refined, so that they can target very, very specifically the main protease and not have side effects where they react with other cysteines of other proteins that are important. I know right now the studies are going on into the toxicity of these drugs and things like that. That's how that part relates.

Barr: Many of those drugs were described as “indole.” What does that mean?

Davis: Yeah, indole is just an organic structure that's two rings with a few double bonds—with a nitrogen it’s called an indole. That particular structure is favorable for binding to the main protease active site. In addition to that, the indole can have an ester formed that can be reactive toward the cysteine. They make these indole esters, and the ester reacts at the cysteine and then the chemical modifies. By modifying it, it inhibits it. They're trying to make it so that when it modifies it stays on permanently and doesn't come off. In the initial drugs that was a little difficult—the GRL-0920 drugs would modify the enzyme temporarily and then it would eventually pop off. They saw that long-term, that's probably not the best way to go. They're still studying derivatives of the indoles, with different attachments put on to try to get the binding in the pocket tighter so that it fits in there snug like a key. Then it would work really well as an inhibitor. A recent paper from Arun Ghosh, the first author, came out with some new derivatives that target the main proteases with much higher affinity. That paper just came out recently, so they're clearly moving forward on working on those.

Barr: That's really exciting.

Davis: Yeah. We'll be getting about 15 drugs in the next month or so. Right now, we can't order because the fiscal year has ended. But in the next couple of months, we're going to get about 15 drugs that we're going to test to targeting cysteine 300. We can specifically study that because we have the serine 300 protease. If a drug works on both, we know we're probably not targeting the 300. But if it only works on the wild type—the native enzyme—then we believe we're targeting, then we know we're targeting cysteine 300.

Barr: Can you speak more about other COVID-19 studies or initiatives that you've been a part of? You've already talked a little bit about some of the future projects that you hope to continue with, but did you want to maybe speak more about that?

Davis: Actually, this is really the only area that I've been working on, as far as where COVID is the main protease. In fact, a lot of our research has gotten back to what we normally do now that people are able to come back in and stuff. I'm not involved in other initiatives, just the one to pursue moving forward what we learned from the first paper that's been published.

Barr: In addition to being a scientist, you're also a person living through the pandemic like everyone else. Can you speak a little bit about how COVID has affected you personally, both the challenges and the opportunities? What was it like to be on campus by yourself in the hallway or things like that?

Davis: I consider myself one of the lucky ones, because while all of my colleagues were told to stay home, because I actually had a project on COVID, I was allowed to come in. You might think I wouldn’t want to come in because I don’t want to get COVID. Well, when we were coming in, there was no one around. We were very safe—probably safer than being at home. It was extremely comfortable and very easy to come in—there was no traffic or anything. It was easy to focus on the research because there were no distractions from other people. Normally I'm training and helping other people do various things and there were none of those distractions, so I

was able to move very fast on the project. I was glad to be participating in something that might ultimately help towards this pandemic or a next pandemic. That all felt very good. Some people love to stay home and work from home, and that's great. I'm not one of those people. It would have been torture for me. I really think I'm really lucky. Whereas my son, Taylor, he loves working from home. He wants to stay home and work from home forever, but that's not going to happen. As far as how it affected me personally, my mom passed away in February but not of COVID. And yet, we were restricted from seeing her prior to her becoming ill because she was in the hospital. You could only see her for 15 minutes, and it had to be across a table with a divider. It was 15 minutes, and that's it—that's all you could see. She lives in Vermont, and that's a 10-hour drive for me to see her 15 minutes. Those kinds of restrictions definitely affected our seeing my mom. Also, it caused some rifts in the family because some people didn't want to wear masks and others did. There was definitely a challenge with the hospice people wanting everyone to wear masks. I have a big family, eight brothers and sisters. They're all across the board—some are super afraid of getting infected and others don't even care. That clashed when we had to take care of seeing my mother before she passed. That definitely interfered with that. My two sons and my wife have been spared so far from COVID, but just recently I found out my oldest sister is infected with COVID. She has the Pfizer vaccine, but she's got some symptoms. She's able to stay at home right now, and it looks like she's going to be okay. But she has fever, cough, and stuff like that. It seems it's starting to infiltrate even to my family. There was a long time when I didn't know anybody who even got exposed to COVID, never mind gotten sick, and now that's dramatically changed. I know many people that have been infected. I know friends whose family members have died and things like that. It's definitely becoming much closer to home.

Barr: I'm so sorry to hear about your sister and your mom. That's very hard.

Davis: Yeah.

Barr: Well, we're going to end on a thought-provoking question. I heard someone say that so much of understanding SARS-CoV-2 and how to combat it relies on investigating how the virus operates on a molecular or anatomical level, which many people overlook or don't fully comprehend, particularly the public. How do you show the importance of understanding this step and looking at the virus in this way, and how do you hope that this kind of understanding can be broken down for the public?

Davis: I feel that the general public just needs the condensed version of what has been worked out at the molecular and anatomical level and all of those details. We don't all know how a car is built, or how an electric saw is put together and how it works. If someone were to ask you to build one, you wouldn't know what to do. Yet we know how to use these things and operate them. Just like the building of the house concept, using these analogies, you can say, “Oh, a virus has to put things together to build and get a new one.” You use those analogies because everything is so repetitive in nature and biology. For example, when a virus attacks a cell or infects a cell, I always use the analogy that it's very similar to someone breaking into your house and kidnapping you. The kidnappers have to hold you under control, right? Yet at the same time the kidnappers want to get something from you. The virus goes into a cell, it kidnaps the cell, it wants to make more virus. But it wants to make sure that that cell stays alive, because if the cell dies, it can't make more virus. You might want to jump out of the house windows or whatever to get out, but they're not going to let you jump. They're going to attach

you to the bedpost or whatever, because they're going to want you to go to the ATM and get the money or whatever. If you use these analogies, you can kind of understand better without knowing the anatomical details. What they need to be able to understand is distilled down by saying that there's this protein, it's really important in making the virus into new viruses, and if we blocked that protein, we could stop that process. That is a very direct approach, and that has been used successfully against HIV. In HIV, we require three drugs to block all of the virus replication, and who knows whether or not coronavirus might require multiple drugs? That's why it's good to have the additional targets that we've identified. But definitely distilling it down to analogies and simplicity. They don't need to understand the subatomic level to appreciate that the scientists have gone into the details to make sure that when they develop a drug, it's not going to kill you when you take it. That's what we do. If they really want to learn, they should be trained and go to school, because it takes a lot—a lot—of work to understand down at those levels. There's no need to go into all of the processes of how it was accomplished to know that we did accomplish it. Of course, they do have to trust us.

Barr: Yes, definitely. What about other scientists? There's been talk that people have different levels of specialization, but a lot of people are going into areas with COVID that they never were before in the scientific community. What about for the scientific community that has different types of training?

Davis: Right. I mean, actually, they say you have to know your audience, right? You [do] have to know your audience. Certainly, the nurses and administrators—people that administer drugs and advice to other people—have to be informed properly, so that when they give advice, it's not an opinion but rather a fact. And if they don't have a fact, then they know to tell you to speak to your doctor about specific details. I mean, I'm not an administrative type of person, but definitely disseminate the information when they have some training in school, where they understand concepts like proteases and indoles and things like that, and where they understand some of those things that can be incorporated, so they grasp it better. But in the end, when they're giving that information to people below them, or the lay public or whatever, they're still going to have to give it in a way that makes sense to the other person. If you just use the scientific terms, it's not going to be very effective.

Barr: Definitely. Well, thank you very much for all of your research. I wish you the best of luck and success, and I hope that your sister gets better and that the rest of your family stays safe.

Davis: You too.

Barr: Thank you very much.

Davis: Thank you very much.

https://pubmed.ncbi.nlm.nih.gov/36290777/
https://pubmed.ncbi.nlm.nih.gov/34399606/
https://pubmed.ncbi.nlm.nih.gov/33510133/