Dr. Alexander Wlodawer
Behind the Mask
December 15, 2020

Gabrielle Barr: Good morning, today is December 15, 2020, and I have the pleasure of speaking with Dr. Alexander Wlodawer. Dr. Wlodawer is a senior scientist for the Center of Structural Biology, and he is the head of the Protein Structure Section at the National Cancer Institute [NCI]. Today he's going to talk to us about some of his COVID research. Thank you very much for being with us.

Wlodawer: Thank you.

Barr: So, one of your first projects that you did when the pandemic first hit was you looked at the analysis and the remediation of the crystal structures of SARS-CoV-2 proteins. You've even published on it. Can you talk a little bit about that effort?

Wlodawer: Of course, I'll be happy to talk about this effort. This effort is a follow-up on work that I have been doing together with a number of my colleagues for at least ten years right now. What we've been trying to do was to look at important targets, protein targets in the Protein Data Bank, medically important, proteins like beta-lactamases that destroy antibiotics, and try to make sure that the structures are up to the highest possible standards. In that case, you know, the structures were determined over many years. The methods have changed. The older structures might not be up to the current standards and so on. It's an international group of scientists scattered in several places both in the U.S and out of the U.S. When the pandemic started, the first structures were available literally within two weeks after the genome sequence of SARS-CoV-2 was published by the Chinese scientists.

Barr: That was very fast.

Wlodawer: That was very fast that there was a structure from China. The reason for it is, first of all, they had first access to the sequence, and the second is that since SARS-CoV-2 is quite similar to SARS-1, the old SARS, there was quite a lot of structural work on SARS so they could just follow the procedures. Indeed, within two weeks there was the first structure. That is very different, let's say, than what happened with HIV where it took several years before the first HIV structure came out.

After that there was really an avalanche of structures. A colleague of mine, who is in Poland, Professor Mariusz Jaskolski, immediately suggested that since we are quite experienced in validating structures and looking at structures, we should start looking at the new structures of the COVID proteins that are coming out, and we did. In order to do it properly, what we created was a special database of all the structures of CoV-2 plus a limited number of structures of proteins from other coronaviruses, mostly SARS, in order to validate as many of these structures as we could. There was a limitation there. We did not feel that we can really validate the cryo-EM [electron microscopy] structures. The problem is much more complicated than the validation of cryo-EM structures. There aren't really good procedures for cryo-EM structures.

Barr: Why is it more difficult?

Wlodawer: There are simply no procedures. Crystallography has been around for many years and we know very well how to validate structures. For cryo-EM, it's still a new technology. and there isn't a direct accepted method. Besides our effort involves, I believe seven people. By now, our database has more than 600 structures. It's just physically not possible for us to do everything. What we decided to do is as far as validation is concerned, first of all, the structures…

Barr: Can you define what validation means? I don't think everyone in the audience would know exactly what validation means.

Wlodawer: What validation means is that scientists determine the structures of targets of proteins and nothing is ever perfect. The structures can be of varied quality and there are now various tools to verify that the quality is really very good. So that's the validation now. All the structures that are submitted to the Protein Data Bank (PDB) get validated by the Protein Data Bank, but this is not sufficient validation in our opinion. It's especially important to validate that the drugs or drug-like molecules that are cut and attached to the target enzymes are really described in the best possible way. That is where we come into this whole thing.

This is a very different type of a validation process because all the structures, as I said, were determined in the year 2020, which means that all were determined using modern tools and modern data collection facilities. It's very different than trying to validate structures that were described for the period of 20 years or 25 years. We can also learn right now things which are very general, not just COVID-related, but related to the quality of the structures in the Protein Data Bank. Our hope is to publish the expanded results. We have right now already a couple of publications, but we are preparing another publication with much more expanded results. Those results will not be just important for CoV-2, but that they will be important in general. So that's the question of validations.

Barr: Okay, so you especially looked when you were validating at the ligands. Can you talk a little bit about what those are and why it was so important to understand that?

Wlodawer: That's very important because we really need to know how to design drugs. There is this hope that vaccines will solve all of the problems. What happened, for example, the research on SARS was stopped in a sense prematurely because there was such a good way of stopping the pandemic with just the public health approaches that the interest died down. Unfortunately, because it would have been very useful right now, but there were still quite a lot of compounds that were found to interfere with the life cycle of the virus, and some of the proteins in CoV-2 are very similar to the ones in CoV-1. There were already indications of where to go and there are quite a few compounds right now that have been found to be very good antivirals. They can lead to the design of drugs that might be helpful against this particular virus. But also, I think that by now we really have to be prepared that there may be another epidemic. You know we had already SARS and MERS. Who knows what is next, so we better be prepared.

These are just small molecules which have a drug-like character. They are not really necessarily approved drugs, even though some of them are approved for other diseases. And those are, of course,

the first ones that we want to look at. If they look promising, then they can be used very quickly, because lots of work was done on already approving them. That's more or less why we are particularly interested in the binding of small molecules to the targets. We were much less interested in the targets, let's say, for vaccines. Because first of all, most of that work involved cryo-electron microscopy and not x-rays. As I told you, we limited ourselves to what we really could do with the current team. The other thing is that you cannot do everything, so you select what it is that we can do.

Barr: What was the process like in creating some of the resources that you did for this project like the Protein Data Bank? What decisions went into it? How did you go about organizing it?

Wlodawer: You see there is this incredible worldwide resource called Protein Data Bank, and that is accessible to everybody very easily. They are the repository of the structures. They made our job a little bit easier by listing each week all the new structures which are CoV-2 related. We didn't have to search deeply what was new.

But what we decided, since there are a couple other labs that parallel our work, was to create our own database, which is a subset that allows anybody who wants to pull up a structure that is CoV-2 related. They have a choice of either pulling up structures directly from the databank or ones that we have validated or eventually improved. That, of course, involved having an expert on databases. One of our colleagues is involved in this project. He is in charge of making that database and maintaining the database. The rest of us are mostly involved in validating the structures and putting them in the database. It is not a very easy job.

Barr: Can you talk a little bit about the tools, techniques, and the programs that you use to validate these structures?

Wlodawer: We are using fairly standard crystallographic software. By now, the software has been developed extremely well. We have standard graphics programs where we can look at the structures. There is a program called COOT (Crystallographic Object-Oriented Toolkit) that we are using for all the visualization. And then there are several ways of refining structures, going from rough model to highly refined model.

We chose to come to use simply one particular program, which is called Refmac. It's part of an international effort called CCP4 (Collaborative Computational Project No. 4). That's where you find it. There are other ways of doing it. I'm not saying that we have chosen the best one because I think that there are at least three very similar methods of doing it. There is Phoenix, there is Baster, but it was, just again, a question of trying to do everything using the same type of methodology with which we are very familiar. So, familiarity is important here.

One of the things is that for a user, who will just want to look at the structures, there are more or less two questions that the users can ask. One is: what is the structure? Is the protein and here is the ligand, the small molecule, bound to it? Let's see how they interact and how they connect. That's one question.

The second question is: how good is this structure? Can we really believe in every atom in the ligand being perfectly placed? Now in order to do that you have to get the so-called electron density map, which will be the experimental description of the structure. Now the map knows nothing really about the model, but without going into the details because there may be some question of bias. But in general, if you can see both the map and the model, you can see, for example, where the model seems to be going, where the map is really not going, and then you know that those parts may be a little bit of a figment of imagination or they may not be well ordered and so on.

The first thing that is happening is that when you solve a crystal structure, you can place your model in various places in the so-called unit in the cell and it's all completely the same from the point of view of computations. You can put it in the cell, you can put it very far out of the cell and it's still the same. However, if you want to compare, let's say, three different structures, which all were grown in the same crystals, it would be much better if they would be all superimposed. Some years ago, some of my colleagues who are part of this project created a website that can move any structure to "the best place." As I said, such movement changes absolutely nothing to the structure itself but makes it much easier to compare different structures. That's the first thing that we do. We just put everything in the same place. Now since there are many crystal forms of those proteins, you cannot put everything exactly or impose, but that's the first step.

The second step is we go through this structure very quickly. We have various tools which can give you some numerical estimate of how good the structure is. I mean two things, how well is the geometry, of the protein or the ligand, how well it agrees with the standard geometry. I mean if you have a peptide bond, it has certain parameters and you cannot postulate that there will be a peptide bond which will be twice as long. So, there are geometrical restraints.

The second thing is how well does it fit to the electron density, to the experimental data. If we find that everything works perfectly well, we then may check if this structure, if the model is in the standard place in the cell. If it's in the standard place in the cell, we don't do anything more. If it's not, then we may just move it and calculate new parameters for getting the map and stop at that, put it in the database. If we see that there are problems, we will rebuild it, refine it, and then put it in the database. The database does tell the users what was changed, why we did some of these things. If we see that there are some really serious problems, we try first of all to contact the original depositors. Then in some cases, if they think that our concerns are valid, they will themselves improve the structures and put the improved model into the PDB and then we put the improved model straight from what they did into our database or we suggest that we can do it together. There is at least one case where the authors decided that it would be better to just take the structure refined by us, redeposited in the PDB. We do it in various ways.

Barr: Have you found [that] the new models that are coming in differ from the ones that came in April? I was reading your article and you said that a lot of the issues with some of the structures that came in earlier had to do with orientation. There were problems with that, and other things related to that.

Wlodawer: Yes, of course. Remember that most of these structures were done very quickly and quick and good doesn't necessarily go together. Some of the earlier structures had more problems than the

later ones. In principle we didn't find too many problems with the proteins themselves because the structures are solved by so-called molecular replacement. People do not go very deep into the original data because you can solve the structure on the basis of the existing model. If the existing model is good, presumably the structure should be quite good and with very rare exceptions, we didn't find many problems with the protein part. We did find some problems with the way the inhibitors were described. We have improved some of them. We are still not finished with our analysis. We're still working on the analysis of what we have learned from all of these structures. That part is by no means finished.

Barr: Has anything surprised you by working on this project?

Wlodawer: No, there weren't too many surprises, because quite a lot was known from previous work on SARS and MERS. We did not expect real big surprises. There were some crystallographic surprises. For example, there is a structure, which is very nice, where two molecules are packed in the crystal in such a strange way that the terminus of one molecule goes right into the active site of another molecule, beautifully describing how a substrate would bind. In this particular case the protease, the main protease, is the substrate of its own, so it cleaves itself off. But why did it happen in the crystal? That's serendipity, but beautiful serendipity. I must say that people who managed to get those crystals were quite happy. The paper describing that particular structure just came out from the group of Natalie Strynadka in Vancouver so that was a kind of a crystallographic surprise. But no, there were no really major surprises that we saw.

Barr: You had mentioned that you thought a lot of the participation was due to the pandemic and that people were maybe more readily willing to share a lot of their data and their thoughts at the speed that they did because of COVID-19. Do you think that this process of submitting your data and your ideas into a database and having it validated will extend to other pathogens in the future?

Wlodawer: I hope so because this was very rapid and quite a lot of structures were deposited long before any papers came out. It's a little bit frustrating because there isn't enough information sometimes in the coordinates that were deposited. There is some information in the core with the coordinates, but sometimes not enough to answer all the questions. This is a very good example of international cooperation that is truly amazing, how quickly these things came out. I would say that this might be quite applicable maybe in the future.

Again, I am going right back to HIV because that was, as you know, another pandemic that worked on a very different scale 30 years ago. In that case the first structure was published by the group from Merck. The only thing that the other people could do was to look at the stereo picture in the Nature paper. Even coordinates were not deposited early. So, lots of things have changed since then. Right now, you can in principle keep your coordinates out of circulation until you publish the paper, but in this case that was very rare. Almost everything was released immediately. Also, there are certain groups involved in so-called structural genomics and those groups have to release all of their coordinates immediately anyway regardless of whether they are COVID or not.

Barr: What kind of groups, government groups?

Wlodawer: I think groups like government groups. There is a group at the synchrotron at Argonne for example, Joachimiak's group that is a structural genomics group that has really contributed a very large number of structures to this particular project. In that case they had no choice. They had to release everything immediately, not that they wanted to keep things to themselves. The problem with that approach was that these groups were slower in publishing. But they did publish some very good papers.

Barr: What has been the most significant challenge today?

Wlodawer: Right now, from our point of view, the challenge was simply that there were so many structures for our very limited group. Remember that was not our major project, and with the exception of one scientist—who retired a few months into the project and then could really concentrate on it—we all had to do other things as well. In a sense we were doing it in so-called spare time of which there wasn't very much. Our main challenge was finding time to do it.

The second challenge was that in some of the structures, we felt that it would be very good to go back to the original data, not just starting with already processed data, but go to the original data. Now there are repositories where original data can be put, but you know people don't normally do it, so we had to convince people that they would give us access to their original data. Again, some people are more willing to share original data than others. Getting that type of access would slow some of the processes. But again, I must say that in most cases the response has been superb.

Barr: That’s really wonderful. What other COVID activities and studies are you working on currently?

Wlodawer: We started a project on trying to make little decoy constructs, hybrid constructs of proteins that could be used for detection of the viruses and antibodies. I managed to get a grant from the NIAID COVID projects to do it. We're just starting that particular work in the lab so that will be another COVID project. I work at the National Cancer Institute. In principle we are supposed to work on cancer, but there is already an established procedure in our Institute allowing us to work on other projects. For example, HIV is not exactly a cancer project, but our Institute has done a lot of HIV work, I think first class HIV work; not talking about my lab, but others in the labs in the Institute. Right now, we were also allowed to jump into the COVID projects, even if we have to slow down some of the cancer-related work. I'm doing my best to do both.

Barr: So, you're continuing doing a lot of the cancer research?

Wlodawer: We're going to be doing our cancer research and that is also one of the reasons why. There is a very limited number of people in the project that I was telling you about. I decided not to get my colleagues here, other than myself, involved directly in the project in order to not take them away from what they are supposed to be doing. Strangely enough in this particular project, almost everybody is a very senior scientist. We feel that maybe as very senior scientists, we are fairly useless in the lab right now, so we can spend more of our time on this particular project.

Barr: Are you mostly working from home or on campus or a combination?

Wlodawer: No, I am one of those people, who from the first day of the pandemic, found that hiding in my office is by far the best way to do it. For three months I was more or less the only one in the whole big building. Now the two things are: I live very close and I have tools in my office that I would not have at home. The computing access is much better here. I selected from the very beginning the approach that I am on the site and not working from home. Then people started coming here. I'm still basically sheltering in a place, in my office.

Barr: Yes, so personally what have been some opportunities that have come up with the pandemic as well as challenges?

Wlodawer: The main challenge was that for several months it was very hard for my colleagues to really continue their work. That was a very big challenge and that affected the non-COVID projects very much. Unfortunately, my younger colleagues paid the price for the epidemic. From my point of view of course, the challenge was that I got involved in some other projects, which were COVID related. We are going to have an interesting paper coming out any day now as soon as the journal decides not to keep rewriting our paper in their own way. But it's accepted.

Barr: Congratulations.

Wlodawer: And that's on a drug that seems to be working very well with and potentiating the action of Remdesivir. That's coming out, I think, any day now. Yes, we managed because I was here and because we were able to access the facility, the x-ray facility during the early days of the pandemic. We could collect some data that were not possible to collect in other places where people did not have access. In a sense we were lucky that we could do the COVID work as I said. Unfortunately, we paid the price in other work.

Barr: Yes, but this is a fun question. What are you most looking forward to during the holidays this year?

Wlodawer: Oh, basically having fun with my grandkids and after finding a Christmas tree, because it turned out that after toilet paper and paper towels, the next thing that is missing are Christmas trees. There are no Christmas trees to be had right now. One great accomplishment was finding a Christmas tree. Having my grandkids and celebrating the holidays and maybe just forgetting about COVID for a few days will be a really great thing.

Barr: Definitely. Is there anything else you would want to share as an NIH scientist, but also as a person living through the pandemic?

Wlodawer: I mean it's been very difficult for us, because I have grandchildren both locally and in Seattle. The fact that the school is remote for most of the time was a very big problem. You know these are young kids, still in elementary school, very difficult. It's difficult for parents. I can see how our kids are struggling trying to keep both their jobs and taking care of the grandkids. We did have a case of COVID in the family. Fortunately, Remdesivir helped. It was not very easy.

But on the other hand, the Institute is marvelous in the way things could be done. They encouraged us and nobody ever objected to the fact that the budget is being spent in a very different way than what we promised to do. The management has been very supportive. The fact that NIH came up with the grants that were basically awarded very quickly was very important. I think that things were done in a very correct way. I usually complain. I am well known for complaining a lot, but in this case, there is very little that I can complain about. Some of the things we simply had never any control over, like what to do with the kids in school.

Barr: Thank you very much for talking about your projects and your experiments. I wish you the best of luck with them and continued success and good health to your family.

Wlodawer: Thank you very, very much, and it was a real pleasure talking to you. Thank you.

A list of relevant publications:

1. "Ligand-centered assessment of SARS-CoV-2 drug target models in the Protein Data Bank," A. Wlodawer, Z. Dauter, I. Shabalin, M. Gilski, D. Brzezinski, M. Kowiel, W. Minor, B. Rupp and M. Jaskolski, FEBS J., 287, 3703-3718 (2020).

2. "Covid-19.bioreproducibility.org: A web resource for SARS-CoV-2-related structural models," D. Brzezinski, M. Kowiel, D.R. Cooper, M. Cymborowski, M. Grabowski, A. Wlodawer, Z. Dauter, I.G. Shabalin, M. Gilski, B. Rupp, M. Jaskolski and W. Minor. Protein Sci. 30, 115-124 (2021).

3. "A small compound with an indole moiety inhibits the main protease of SARS-CoV-2 and blocks virus replication," S. Hattori, N. Higashi-Kuwata, H. Hayashi, S. Rao Allu, J/ Raghavaiah, H. Bulut, D. Das, Y. Takamatsu, N. Takamune, N. Kishimoto, K. Murayama, K. Hasegawa, M. Li, D.A. Davis, E. N. Kodama, B.J. Anson, E.K. Lendy, A.D. Mesecar, A. Wlodawer, R. Yarchoan, S. Misumi, A.K. Ghosh and H. Mitsuya, Nature Commun., 12:668 (2021).

4. "Crystallographic models of SARS-CoV-2 3CLpro protease: in-depth assessment of structure quality and validation," M. Jaskolski, Z. Dauter, I.G. Shabalin, M. Gilski, D. Brzezinski, M. Kowiel, B. Rupp and A. Wlodawer, IUCr Journal, in press (2021).