Omics

How mass spectrometry plays a role in understanding the immune system

An interview with Dr. Susan Klaeger

Susan Klaeger, Ph.D.
The Broad Institute of MIT and Harvard

"Mass spectrometry is one of the few high-throughput technologies, if not the only, that can lead to direct physical detection and identification of thousands of peptides presented on human leukocyte antigens."

– Susan Klaeger, Ph.D.
The Broad Institute of MIT and Harvard

Introduction

Many viruses infect humans, and most are controlled satisfactorily by the immune system with limited damage to host tissues. However, some viruses do cause overt damage to the host, either in isolated cases or as a reaction that commonly occurs after infection. The outcome is influenced by the properties of the infecting virus, the circumstances of infection, and several factors controlled by the host.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of the ongoing coronavirus disease pandemic and one which will be touched upon in this article. To understand the pathogenicity and antigenic potential of SARS-CoV-2 alongside other viruses and to develop therapeutic tools, it is essential to profile the full repertoire of the viral fragments to be presented to the immune system. This is work that Dr. Susan Klaeger along with virologists and computational biologists from The Broad Institute of MIT and Harvard carry out.

Peptide characterization

"We aim to characterize the repertoire of peptides presented on HLA class I molecules and how antigen presentation changes in cancer or viral infection using mass spectrometry."

– Susan Klaeger, Ph.D.

Dr. Klaeger's research focuses on our bodies' immune response in cancer and viral infection, and her work could be used to identify new therapeutic targets. In this interview, we speak to Susan Klaeger and explore how she studies the host immune system as well as gain insights into how she uses mass spectrometry to do so.

Q: Can you describe the immune response to viral infections

A: To begin with, it's important to note that the immune response can be separated into two parts, the innate and adaptive immune response.

The first line of defense against non-self pathogens is the innate or non-specific immune response. This consists of physical, chemical, and cellular defenses. The innate immune response allows the upregulation of different immune cells within the body that engulf the invading pathogen to try to eliminate it.

The second part is the adaptive immune response, which is specific to the pathogen presented and is engaged when using vaccinations. Here, the body needs time to learn what pathogen it is encountering, and the immune response includes B cells that develop antibodies against the pathogen, and T cells.

My collaborators and I are looking at these T cell responses and, more specifically, the human leukocyte antigen class I (HLA-I) – CD8 T cell pathway. Here, HLA-I complexes present protein fragments from pathogens on the cell surface. This process happens in every cell in the body, and it is effectively a signal from a particular cell to say whether it's healthy or infected. If the cell presents as being infected, a CD8 T cell will recognize the viral antigen on the surface as foreign. It will bind to the cell and start releasing cytotoxic proteins that will induce apoptosis in the infected cell, meaning the infected cell dies and will not produce further virus particles.

Antigen presentation also includes the HLA class II pathway, and these molecules are constitutively expressed on professional antigen-presenting cells. They also present antigens to the immune system, often from extracellular sources, and will be recognized by CD4 T cells, which are also known as helper T cells. These T cells perform multiple functions such as recruiting other immune cells like B cells.

Overall, the immune response is quite complex but it's fascinating because there are multiple aspects to it.

Q: What is your research focus and what types of information are you looking for when studying the host immune response?

A: We aim to characterize the repertoire of peptides presented on HLA class I molecules and how antigen presentation changes in cancer or viral infection using mass spectrometry.

Antigen presentation is a key component to many pathways in the body; however, it is particularly challenging to study as nature has created hundreds of possibilities of these molecules and the human population is very diverse. This means that any given person's HLA molecules likely won't be the same as another's. Ultimately, this is a good thing as this allows for different responses to pathogens so that we can have the biggest chance of surviving said pathogens. However, it also means it is very complex to study.

We have been using mass spectrometry to profile the protein fragments (peptides) that are presented by these HLA molecules on the cell surface because the technology enables direct identification of peptides. Furthermore, it provides high throughput, and we can identify the thousands of antigens that are presented on the cell surface.

Sensitivity

"…the challenge is that we need very sensitive instrumentation to dig deeper into the sample amounts we have available and still identify the antigens we are interested in. We are constantly trying to improve sensitivity, which will hopefully translate to identifying actionable targets."

– Susan Klaeger, Ph.D.

In my research, we are using this approach in different ways. To begin with, we have employed a single allele approach to learn the rules of antigen processing and presentation that lead to the presentation of a peptide on a particular allele. As I mentioned previously, each individual's HLA molecules will not be the same and every cell presents up to six different alleles, typically two alleles each of the classical types HLA-A, HLA-B, and HLA-C, which increases the complexity. We have found that by taking one allele each and by analyzing the peptides binding to that HLA molecule, we can systematically study antigen presentation and use that information to train neural networks for improved antigen prediction.

For other research questions, we use cells infected with a virus or tumor cells and we analyze the peptides on the surface to understand which peptide is being presented, why the peptide is being presented, and if it can be responsible for eliciting an immune reaction through the CD8 T cells.

Q: What are the main challenges in this area of research?

A: The immune system is very complex and there are many different versions of these molecules. Each of the alleles binds to different peptides so there are multiple possibilities of how such a peptide can be presented. Moreover, it's important to note that a particular peptide may not necessarily be abundant on the cell surface. Therefore, the amount of material we need to do mass spectrometry and to detect some of the therapeutically interesting but potentially lower abundant antigens is quite high. In the case of the SARS-CoV-2 project we've done in our study, it was particularly challenging because to infect this large amount of cells we also need a lot of virus that has to be produced first. Moreover, working with an active virus is only possible in high safety level settings (BSL-3), so experiments need to be planned according to time and space available in these specialized labs. We also had to adjust the protocol to make sure that the cells with the virus weren't infectious anymore before we could perform HLA-I peptide isolation.

On the mass spectrometry side, the challenge is that we need very sensitive instrumentation to dig deeper into the sample amounts we have available and still identify the antigens we are interested in. We are constantly trying to improve sensitivity, which will hopefully translate to identifying actionable targets. Also, HLA presented peptides differ from our routine MS analytes, and algorithmic searching of the MS/MS spectra is also more complex compared to our common discovery approaches.

Q: In your study, you investigate class I human leukocyte antigen (HLA-I) in SARS-CoV-2. What is HLA-I and why is it important?

A: The human leukocyte antigen (HLA) proteins are encoded by the highly polymorphic major histocompatibility complex (MHC) gene complex in humans. Class I HLA molecules are present as transmembrane glycoproteins on the surface of almost all nucleated cells. Intact class I molecules are made up of an alpha heavy chain bound to a beta-2 microglobulin molecule, and there are over 16,000 different alleles known in the human population.

These proteins regulate the immune system by presenting short fragments of 8–11 amino acids from inside the cell on the surface. Each HLA allele has a distinct peptide binding motif with fixed amino acid residues in certain positions of the peptide. Prior to presentation on the surface, peptides are generated through proteasomal processing of proteins present in the cell and loaded onto HLA-I complexes in the endoplasmic reticulum. If the presented peptide is recognized as a foreign antigen by a circulating cytotoxic T cell, these T cells are activated and destroy the antigen presenting cell.

Enabling greater data insights faster

"The speed and sensitivity of the Orbitrap Exploris 480 mass spectrometer have benefited this research tremendously."

– Susan Klaeger, Ph.D.

SARS-CoV-2 particles enter the host's cell, and the virus uses the host cells' machinery to reproduce and spread. So viral proteins are generated inside the host cell and will likely be degraded by the host proteasome. Fragments of the viral proteins in turn will be presented on HLA molecules akin to host proteins. Ideally, infected cells will be recognized by the immune cells and eliminated. Hence, it is not surprising that viruses often interfere with the host's antigen processing and presentation machinery to avoid recognition. We were interested to see which viral peptides are presented and which SARS-CoV-2 proteins give rise to peptides presented on the surface. Moreover, we wanted to know at what point in time after infection do these peptides appear and whether their identity differs early or late in infection. Lastly, we were interested to learn which of these peptides can elicit an immune response.

Q: How did you use mass spectrometry in your research?

A: Mass spectrometry is one of the few high-throughput technologies, if not the only, that can lead to direct physical detection and identification of thousands of peptides presented on HLA.

As these HLA-I eluted peptide mixtures (the immunopeptidome) are very complex because the analytes are similar in size but differ in their amino acid composition, we use ultra-high sensitivity instrumentation from Thermo Fisher Scientific—the Thermo Scientific™ Orbitrap Exploris™ 480 mass spectrometer with FAIMS (high field asymmetric ion mobility spectrometry). The speed and sensitivity of the Orbitrap Exploris 480 mass spectrometer have benefited this research tremendously. The addition of FAIMS has further increased coverage of the presented immunopeptidome as it separates coeluting peptides and hence reduces interference of identifications. In our lab, we perform a small offline fractionation of the HLA-eluted peptides and then separate each fraction in an 86 min online gradient ranging from 6% acetonitrile to 90% acetonitrile during which peptides eluting from the analytical column and are analyzed in the instrument. Additional method parameters, such as a fit filter available for the Orbitrap Exploris 480 mass spectrometer, benefit HLA peptide identification further.

The next steps for us would be to understand which of the MS-identified peptides can lead to an immune response. We typically select a few candidates that could warrant further follow-up on immunogenicity assays. In one such assay, the enzyme-linked immunosorbent spot (ELISpot) assay, one can measure cytokine secretion after incubation of cells with a synthetic peptide, which is a signal that immune cells react to the antigen.

Q: What information did mass spectrometry reveal about the role of HLA-I in SARS-CoV-2 immunity, and how can this information be used for the development of potential therapeutics?

A: We now know that T cell-mediated immunity plays an important role in controlling SARS-CoV-2 infection; however, the full repertoire of naturally processed and presented viral epitopes is not yet fully characterized. While we looked at which peptides from SARS-CoV-2 are presented by HLA-I on the surface of infected cells, other groups are evaluating potential candidate antigens based on prediction and T cell assays. Of course, the ultimate goal is to identify which peptide is presented and identify which peptide is helping to eradicate the infected cell.

We did this in collaboration with biologists here at Boston University, where we infected two cell line models. Unfortunately, we were unable to do it on patient material as this is particularly problematic due to the high infectivity of the virus and lack of a BSL-3 lab standard. However, we found collaborators that could infect the cells with live virus, so we have chosen two model system A549 cells, which is a lung cancer cell line and has also been used to study SARS-CoV-2 infection as well. The other cell line we used was HEK293T cells, which we chose because the alleles it presents are the most frequent in the human Caucasian population.

Moving forward with both cell lines, we isolated the peptides presented on HLA and subjected them to mass spectrometry. We then searched the spectra against a database comprised of human protein sequences as well as SARS-CoV-2 proteins, including canonical and noncanonical sequences. We obtained information on the expression of noncanonical sequences from collaborators at the Department of Molecular Genetics, Weizmann Institute of Science, Rehovot in Israel.

After supplementing our database with this information, we searched our mass spectrometry data and identified peptides from SARS-CoV-2—which for me was already a success.

We found that the peptides presented on the surface of infected cells were derived from spike and nucleocapsid proteins as well as non-structural proteins of SARS-CoV-2. Additionally, we found peptides derived from non-canonical open reading frames, particularly from internal out-of-frame ORFs in spike and nucleocapsid proteins. These sequences are not currently captured in therapeutics and vaccine design but can be of interest moving forward. We also want to follow up further to understand why the immune system is presenting these fragments and not others.

Besides identification of epitopes, we also characterized the dynamics of antigen presentation. We examined various time points up to 24 hours post-infection, and we found that all SARS-CoV-2 derived peptides could be observed over that time frame but did change in their abundance on the surface over time.

We also performed standard bottom-up proteomics mass spectrometry experiments on the whole proteome over these time points and found that viral infection might interfere with antigen processing pathways to hide from detection by the immune system.

Taken together, these biological insights can facilitate the selection of peptides for immune monitoring and vaccine development.

Learn more at **thermofisher.com/immunopeptidomics**