



Proteomics at the heart of multiomic strategies to enable precision medicine

An interview with Ida Grundberg Ph.D., Chief Scientific Officer at Olink® Proteomics

In this Q&A session with Dr. Grundberg, we explore the growing impact of complex multiomic studies to better understand the molecular and cellular mechanisms that underlie health and disease and to develop better and more targeted therapeutics that will help realize the promise of precision medicine. As a leading innovator in high-throughput, high-multiplex proteomics, Olink® is well-placed to greatly impact the multiomics revolution, as Dr. Grundberg makes evident in her discussion on the current and future impacts of these studies.

Q. Could you start by telling us a little about Olink and how it fits in to the multiomics/precision medicine picture?

Olink's mission is to accelerate proteomics together with the scientific community, to understand real-time biology and gain actionable insights into human health and disease. Our goal is to drive the development of better, more targeted therapies and help revolutionize future healthcare through the implementation of precision medicine. We believe that putting high-quality proteomics at the center of multiomic strategies is the key towards achieving these ambitious goals. The company was founded on our innovative Proximity Extension Assay (PEA) technology¹ for multiplexed protein biomarker research and in a few short years, we have developed from a small Swedish company with a few disease-focused panels covering a couple of hundred proteins, to an established global organization offering unique, high-throughput proteomics solutions that enable scientists to measure almost 3000 proteins simultaneously from just a few microliters of blood. Olink customers have now generated close to 900 million protein data points and published well over 800 peer-reviewed scientific articles².

Q. Genomics has dominated biological research over the last couple of decades, but proteomics is becoming more prominent in scientific studies. Could you explain why proteomics is now coming to the fore, and what proteins bring to the table?

In the central dogma, it is proteins that lie closest to biological outcomes, representing both the individual phenotype and how it is impacted by environmental and lifestyle factors. This is crucial for medical research, where most diseases are complex and multifactorial (with contributions from genetics, the environment, and lifestyle choices). There is a strong, unmet need to focus on proteins, as they also provide the most actionable targets for therapeutic interventions - they are the targets of most current drugs.

Q. What are some important applications for proteomics, particularly regarding precision medicine?

If you think about protein biomarkers in isolation, the applications are extremely broad, from basic biological research to drug development and clinical care. Protein biomarker discovery studies using Olink have been used widely in clinical research settings to stratify patient sub-groups, identify predictive, diagnostic, or prognostic protein signatures and to gain deep pathophysiological insights across a broad range of diseases. Where we see the biggest potential for proteomics, however, is in more extensive, "system biology" approaches that combine multiple "omics" methodologies to probe how different cellular and biomolecular components interact and combine to maintain biological processes and their dysregulation in the pathobiology of disease. The multiomics approach gives us the best chance of achieving a much deeper understanding of disease biology that will be necessary to provide actionable insights for developing new, more effective, and better tailored therapies in the future.

Q. You clearly see multiomics as an important development but there are many types of omics technologies available now, how are these being used and combined with proteomics and for what types of applications?

It all started with genomics of course and we now have a broad range of additional approaches being used, epigenomics, transcriptomics, proteomics, metabolomics, lipidomics, mass cytometry etc. Scientists are increasingly using creative combinations of these omics techniques to address a wide range of applications, but there are some examples that come to mind that are particularly exciting and highly promising. We are starting to see some impressive, population health studies that take a systems biology approach to interrogate very large sample cohorts with

multiple techniques to get a truly comprehensive picture. There are some pioneering groups in this field, like those headed by Mathias Uhlén at Stockholm's Royal Institute of Technology (KTH) and Leroy Hood at the Institute for Systems Biology (ISB) in Seattle, both of whom we have had the pleasure of working with. These and other groups around the world have been spearheading the use of large-scale, multiomic studies of healthy populations as a tool to enable knowledge-driven advances in precision medicine. Professor Uhlén's team have been driving ambitious projects like the Human Protein Atlas³, as well as publishing some landmark multiomics studies, for example demonstrating that individuals generally maintain a remarkably stable molecular profile over time, but that these profiles vary significantly among different people⁴. Leroy Hood is a longtime champion of the systems biology approach to drive precision medicine, believing that population studies of seemingly healthy individuals can unlock the secrets of disease biology. Proteomics is a growing part of that work. One study from ISB, for example, looked at a wellness cohort and found that the protein CEACAM5 was able to predict the incidence of several forms of metastatic cancer several years before clinical diagnosis⁵.

Q. You seemed to suggest earlier that combining multiple omics technologies can provide insights that aren't possible otherwise, could you expand on that?

I think one of the most powerful combinations, which will be extremely important for more efficient drug development, is the combination of genomics and proteomics, or "proteogenomics" as it is now known. The combination of proteomics and large-scale genetic analyses such as Genome Wide Association Studies (GWAS) can be used to identify specific genetic variants that affect the plasma levels of individual proteins. These are called protein Quantitative Trait Loci (pQTLs) and can occur within or close to the gene encoding the affected protein (cis) or be located distally from the protein-encoding gene (trans). The trans-pQTLs can provide valuable new insights into disease pathophysiology, identifying previously unknown regulatory pathways involved in disease biology or drug action.

The cis-pQTLs are even more powerful though, as they provide unique insights into whether the protein being studied is a cause or consequence of the associated phenotype. Having a statistically significant link between the measured amount of a protein and a disease phenotype for example gives you a potentially valuable biomarker, but there is no way of knowing whether the change in protein level is involved in causing the disease or is simply a response to the disease biology. If you have a genetic variant linked to the levels of protein within the same gene and both of those things are associated with the disease phenotype, that is the best in vivo evidence you can get that the protein is causal in the disease, because the genetic association of course can only work in one direction. Scientists use a statistical method called Mendelian Randomization to test for these associations of genes, proteins, and phenotypes, using the random meiotic assortment of alleles containing genetic variants as a kind of natural substitute for the cases and controls that you would normally have in a randomized clinical trial. Being able to confidently identify proteins that are causally involved in disease is extremely important because that is the essential first property for any new potential drug target.

These types of GWAS/proteomic studies are becoming more frequent and producing a wealth of actionable data for clinical research and drug discovery & development. Large numbers of samples are usually needed to maximize the statistical power of the data, and this is also leading to a collaborative approach where different groups come together and combine their data for the best outcomes. There is an independent consortium of scientists who have both genetic and Olink-generated protein data, for example, called SCALLOP⁶ (Systematic and Combined AnaLysis of Olink Proteins) and they have published some important studies identifying new pQTLs for a number of different diseases^{7,8}. As an additional benefit, the cis-pQTLs also provide strong genetic validation that your proteomics technology is accurately measuring the protein you think you are, and we have been seeing a high frequency of those being identified in the studies referred to.

Other very large ongoing projects using this proteogenomic approach include the UK Biobank Pharma Proteomics Project (UKB-PPP Consortium), which people may well have heard about. They are adding proteomics to complement the huge amount of genetic data they have, using the Olink® Explore platform to measure ~3000 proteins in 53,000 participants from the study. These large-scale collaborative studies, where academic research and pharma work in tandem, may happen more often in the future and obviously have the potential to make a huge impact.

Q. There is clearly some exciting and significant work going on with these proteogenomic studies, but are you seeing any other interesting multiomics approaches where proteomics plays a pivotal role?

Another approach that we are seeing more frequently is studies combining tissue-specific or single cell transcriptomics with plasma proteomics. RNAseq, for example, can provide important insights into the biology that's happening at the site of disease, but that usually requires a biopsy or some other invasive procedure to acquire a sample. Ideally, everyone would like to use a simpler, minimally invasive method to get at relevant biomarkers, but looking at RNA markers in the circulation rarely correlates with what is observed more locally. What we do see, however, is that there is often quite a better correlation between circulating protein levels in blood and what the disease site RNAseq data is saying. This approach has been used across a range of disease areas from dermatology to inflammatory bowel disease but has probably found most use so far in oncology and cancer immunotherapy studies.

One study that comes straight to mind came out from Imperial College London last year, where they identified potential protein biomarkers for lung cancer⁹. Most of these also associated with smoking history and lost independent significance after statistical adjustment for smoking, but CDCP1 (a protein involved in CD6 T-cell migration and immune function) remained as a smoking-independent predictive risk marker for development of lung cancer. They were able to verify this in a second cohort and show that levels of CDCP1 are elevated several years prior to clinical diagnosis. When they combined the proteomics data with RNA-seq analysis, they identified significant associations between CDCP1 levels and the expression of LRRN3 and SEM1 mRNA, indicating that CDCP1 may be involved in molecular pathways related to WNT/β-

catenin signaling, a canonical signaling pathway associated with carcinogenesis. So, the integration of proteomic and transcriptomic data has the potential to improve early diagnosis of lung cancer by enabling an in-depth pathway analyses of underlying malignant transformation mechanisms.

Q. The COVID-19 pandemic has obviously been a big focus for the past few years, have multiomics studies involving Olink played a part there too?

Yes absolutely! Olink technology was involved in many COVID studies from an early stage in the pandemic and over 60 papers have already been published from those efforts. Olink collaborated with a team from Massachusetts General Hospital (MGH) on a large longitudinal study where proteomics and RNAseq revealed specific proteins and underlying pathobiology associated with severe disease outcomes¹⁰. That study also illustrated the collaborative spirit during the COVID-19 period, as the molecular and associated clinical data were made freely available to the wider scientific community. The whole COVID issue is a prime example of why multiomics approaches with a strong proteomics component are so important to tackle complex biological questions. It became clear early on that patient immune responses were critical in determining the broad spectrum of disease severity and clinical outcomes observed, and thus required comprehensive strategies to map this complexity. That involved looking at inflammatory proteins like cytokines and chemokines of course, but also required a broader discovery mindset to capture diverse molecular components, such as characterizing immune cell populations through approaches like mass cytometry. This led to some interesting and important discoveries, like mapping local and systemic immune system responses to reveal a spatial dichotomy in the innate immune response¹¹, detailed mapping of the immune response to the BNT162b2 mRNA vaccine in humans¹² and a systems biology characterization of the rare MIS-C complication that can occur in SARS-CoV-2 infected children¹³.

Q. There is clearly a lot of work going on already with this multiomics approach, but how do you see the future in terms of challenges and potential developments?

That is obviously a difficult question to answer with certainty, but there are some trends developing that should be highly relevant. Firstly, I'd say that many of the technical barriers to running proteomics at the scale and throughput required for large cohort studies are now being addressed and we like to think here at

Olink that we've made a big contribution to that. This will enable researchers to carry out proteomics at population scale, combining protein data with other omics in ever larger sample cohorts, which will inevitably lead to breakthroughs in disease research and has clear repercussions on future drug development.

It is also fascinating to hear how some of the thought leaders in this field are now envisioning large-scale, broad proteomics initiatives impacting their research. For example, there is great potential for discovery proteomics approaches in large populations to give us more mechanistic insights into unclassified and complex diseases. Biomarker research is beginning to show that in many cases, different diseases exhibit shared disease drivers and biological pathways, which could lead to a different approach to medicine where new types of clinical trials may be conducted based on biomarkers and protein profiles, rather than by disease type.

One inevitable consequence of this growing trend in increasingly large-scale, multiomic studies is a logarithmic growth in the volume and complexity of the data being generated. Advances in biostatistical methods will surely help to meet this challenge, and we are already seeing how things like machine learning algorithms are being used to interrogate large datasets to identify new biomarker signatures and get real biological insights. More sophisticated protein data visualization tools would also be extremely valuable to help integrate protein data into multiomics datasets that will enable better understanding of biological mechanisms and identify more precise disease phenotypes.

As the biomedical research community continues to strive for precision medicine and personalized therapy, proteomics will become ever more integral to reaching that goal. For proteomics data to become even more informative and useful in multiomics strategies, we must reach a higher resolution in our knowledge of the human proteome. This will involve new ways of analyzing protein data and the use of open-source protein databases to further our understanding of protein pathways.

Future technical innovations will likely take us off on exciting new paths as scientists seek to probe the complexities that define our biology, going ever deeper into the different omics landscapes and integrating the wealth of data that results from these efforts to get a complete picture of human health and disease. One thing that I am sure about is that the multiomics strategy with proteomics at its heart is here to stay and will have a profound impact on future therapeutic development and the implementation of precision medicine.

Thank you Ida!

More information and examples of how proteomics is being used at the heart of multiomic studies can be found on the Olink web hub: www.olink.com/application/multiomics

References

1. J. D. Harding, Nonhuman Primates and Translational Research: Progress, Opportunities, and Challenges. *ILAR J.* 58, 141-150 (2017).
1. Assarsson E, Lundberg M, Holmquist G, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. (2014) *PLOS One*, DOI: 10.1371/journal.pone.0095192 [Article link](#)
2. The latest list of publications citing the use of Olink technology is available at www.olink.com/resources-support/publications
3. The Human Protein Atlas website: www.proteinatlas.org
4. Tebani A, Gummesson A, Zhong W, et al. Integration of molecular profiles in a longitudinal wellness profiling cohort. (2020) *Nature Communications*, DOI: 10.1038/s41467-020-18148-7. [Article link](#).
5. Magis A, Rappaport N, Conomos M, et al. Untargeted longitudinal analysis of a wellness cohort identifies markers of metastatic cancer years prior to diagnosis. (2020) *Scientific Reports*, DOI: 10.1038/s41598-020-73451-z. [Article link](#).
6. SCALLOP consortium web page: www.olink.com/scallop
7. Folkersen L, Gustafsson S, Wang Q, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. (2020) *Nature Metabolism*, DOI: 10.1038/s42255-020-00287-2. [Article link](#)
8. Henry A, Gordillo-Maranon M, Finan C, et al. Therapeutic Targets for Heart Failure Identified Using Proteomics and Mendelian Randomization. (2022) *Circulation*, DOI: 10.1161/CIRCULATIONAHA.121.056663 [Article link](#)
9. Dagnino S, Bodinier B, Guida F, et al. Prospective Identification of Elevated Circulating CDCP1 in Patients Years before Onset of Lung Cancer. (2021) *Cancer Research*, DOI: 10.1158/0008-5472.CAN-20-3454. [Article link](#)
10. Filbin M, Mehta A, Schneider A, et al. Longitudinal proteomic analysis of severe COVID-19 reveals survival-associated signatures, tissue-specific cell death, and cell-cell interactions. (2021) *Cell Reports Medicine*, DOI: 10.1016/j.xcrm.2021.100287. [Article link](#)
11. Arunachalam P, Wimmers F, Mok C, et al. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. (2020) *Science*, DOI: 10.1126/science.abc6261 [Article link](#).
12. Arunachalam P, Scott M, Hagan T, et al. Systems vaccinology of the BNT162b2 mRNA vaccine in humans. (2021) *Nature*, DOI: 10.1038/s41586-021-03791-x. [Article link](#).
13. Consiglio CR, Cotugno N, Sardh F, et al. The Immunology of Multisystem Inflammatory Syndrome in Children with COVID-19. (2020) *Cell*, DOI: 10.1016/j.cell.2020.09.016. [Article link](#).



www.olink.com

For research use only. Not for use in diagnostic procedures.

This product includes a license for non-commercial use. Commercial users may require additional licenses. Please contact Olink Proteomics AB for details.

There are no warranties, expressed or implied, which extend beyond this description. Olink Proteomics AB is not liable for property damage, personal injury, or economic loss caused by this product.

Olink® is a registered trademark of Olink Proteomics AB.

© 2019–2021 Olink Proteomics AB. All third party trademarks are the property of their respective owners.

Olink Proteomics, Dag Hammarskjölds väg 52B, SE-752 37 Uppsala, Sweden

V1.0, 2022-06-21