Q&A

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 to realize the promise of precision medicine. As a leading innovator in high-throughput, high-multiplex proteomics technologies, Olink® is also well-placed as a leader in the multiomics revolution. We connected with Dr. Grundberg to ask her to make the case for Olink's current and future impacts in these studies.

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

A. 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 goals are 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) technology[1] for multiplexed protein biomarker research. In brief, the platform is a dual-recognition immunoassay, where two matched antibodies labelled with unique DNA oligonucleotides simultaneously bind to a target protein in solution. This brings the two antibodies into proximity, allowing their DNA oligonucleotides to hybridize, serving as template for a DNA polymerase-dependent extension step. This creates a unique double-stranded DNA "barcode" for the specific antigen and quantitatively proportional to the initial concentration of target protein. The hybridization and extension are immediately followed by PCR amplification. The resulting DNA amplicon can then be quantified using either Next Generation Sequencing (NGS) or microfluidic qPCR, depending on the Olink platform being used. The principle of the PEA method is shown in **Figure 1**.

In a few short years, we have grown from a small company based in Sweden 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 850 peer-reviewed scientific articles.[2]

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?

A. In the central dogma, proteins are most closely associated with biological outcomes, representing both the individual phenotype and how it is influenced by external factors (e.g., environmental conditions and lifestyle choices). This role for proteins is a crucial point for medical research, since most diseases are complex and multi-factorial with contributions from genetics and, as noted, the environment and lifestyle choices). There is a strong, unmet need for large-scale protein profiling, as proteins provide the most actionable targets for therapeutic interventions – in fact, they are the targets of most current drugs.

Q. What are some important applications for proteomics, particularly regarding precision medicine, especially with Olink's platform?

A. In general, protein biomarker 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"

> *"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."*

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. Multiomics is becoming an important development tool but there are many types of omics technologies available now. How are these omics technologies being used and combined with proteomics and for what types of applications?

A. 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 some examples 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. Pioneering groups in this field include for example, 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; we have had the pleasure of working with both of these groups. These and others 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 has been driving ambitious projects like the Human Protein Atlas,[3] as well as publishing some landmark multiomics studies. For example, Uhlén's team have been able to demonstrate that individuals generally maintain a remarkably stable molecular profile over time, but that these profiles vary significantly »
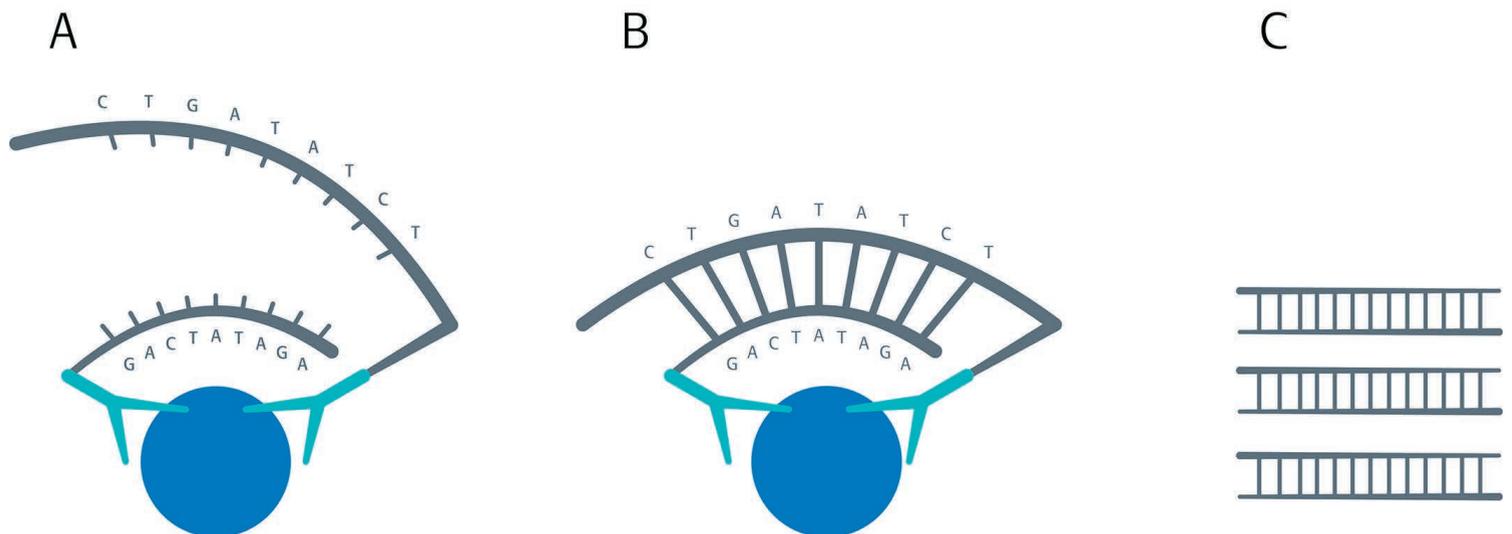
Figure 1: Proximity Extension Assay (PEA) methodology.
(A) Antibody pairs, labelled with DNA oligonucleotides, bind target protein in solution. **(B)** Oligonucleotides that are brought into proximity hybridize and are extended by a DNA polymerase. **(C)** This newly created piece of DNA barcode is amplified by PCR ready for readout by NGS or qPCR. Olink products are for research use only.

among different people.[4] Leroy Hood is a longtime champion of the systems biology approach to drive precision medicine, while promulgating the hypothesis that population studies of seemingly healthy individuals can unlock the secrets of disease biology. Proteomics is a growing part of ISB's 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 diagnosis.[5]

Q. Can you expand on your earlier comment that combining multiple omics technologies can provide insights that are not otherwise possible?

A. 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 either 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 – after all, genetic associations 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 and 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. For example, SCALLOP[6] (Systematic and Combined AnaLysis of Olink Proteins) is an independent consortium of scientists who have both genetic and Olink-generated protein data. This consortium has published 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 being seeing a high frequency of those being identified in the studies referred to.

> *"One multiomics study that comes straight to mind came out from Imperial College London last year; the study identified potential protein biomarkers for lung cancer."*

Other very large ongoing projects using this proteogenomic approach include the UK Biobank Pharma Proteomics Project (UKB-PPP Consortium). 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. Exciting and significant work is going on with these proteogenomic studies, but are you seeing any other interesting multiomics approaches where proteomics plays a pivotal role?

A. 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. What we do see is a correlation between circulating protein levels in blood and what the disease site RNAseq data reveals. This approach has been used across a range of disease areas from dermatology to inflammatory bowel disease but has probably found the most use so far in oncology and cancer immunotherapy studies.

One multiomics study that comes straight to mind came out from Imperial College London last year; the study identified potential protein biomarkers for lung cancer.[9] Most of these protein biomarkers are also associated with smoking history but lost independent significance after statistical adjustment for smoking. On the other hand, 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/ß-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?

A. 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 outcomes.[10]

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. Such strategies 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 work led to interesting and important discoveries, like mapping local and systemic immune system responses to reveal a spatial dichotomy in the innate immune response,[11] detailed mapping of the immune response to the BNT162b2 mRNA vaccine in humans[12] and a systems biology characterization of the rare MIS-C complication that can occur in SARS-CoV-2 infected children.[13]

> *"Future technical innovations will likely take us on exciting new paths. By their very nature, 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."*

Q. How do you see the future of multiomics approaches in terms of challenges and potential developments?

A. There are some trends developing in multiomics that should be highly relevant. First, 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. We like to think that Olink has made a big contribution to overcoming these hurdles. Work such as ours 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 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 on exciting new paths. By their very nature, 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! JoPM

Ida Grundberg

Ida received her PhD at Uppsala University from the prestigious research group which developed Olink's key technology and founded Olink. She continued as senior scientist at Olink's RnD for commercialization of the patented technology developed during her PhD. Following the market release of Olink´s multiplex product, Dr. Grundberg joined the commercial organization and in 2015 she transferred to head the U.S. market entry for her scientific understanding of Olink's products, coupled with commercial experience. In 2020, Ida took on the role as Chief Scientific Officer, leading the global Scientific Affairs team at Olink.

Background reading
For a case study, see free download of Proteomic plasma biomarkers for predicting response and resistance with immune checkpoint inhibitors in cancer patients @ https://www.selectscience.net/application-articles/proteomic-plasma-biomarkers-for-predicting-response-and-resistance-with-immune-checkpoint-inhibitors-in-cancer-patients/?artid=53146&u=ABAF77CB-5DEA-4152-91C3-87ECAB4542BD&error=1#pop

References
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. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0095192
2. The latest list of publications citing the use of Olink technology is available at https://www.olink.com/resources-support/publications/
3. The Human Protein Atlas website: https://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. https://www.nature.com/articles/s41467-020-18148-7
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. https://www.nature.com/articles/s41598-020-73451-z
6. SCALLOP consortium web page: https://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. https://www.nature.com/articles/s42255-020-00287-2
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. https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.121.056663
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. https://aacrjournals.org/cancerres/article-abstract/81/13/3738/667088/Prospective-Identification-of-Elevated-Circulating?redirectedFrom=fulltext
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. https://www.cell.com/cell-reports-medicine/fulltext/S2666-3791(21)00115-4
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. https://www.science.org/doi/10.1126/science.abc6261
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. https://www.nature.com/articles/s41586-021-03791-x
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. https://www.cell.com/cell/pdf/S0092-8674(20)31157-0.pdf