ASMS2024 Anaheim
Monday, June 3rd 2024
Workshop Top-Down Proteomics from the Top-Down Proteomics Interest Group

Top-Down Proteomics: Bridging the Gaps between Academia, Clinical Research, and Biopharmaceutical Industry
Presiders: Fanny Caroline Liu and Yuri van der Burgt

The workshop started with a brief introduction to the Consortium for Top Down Proteomics (CTDP) and the Early Career Researcher (ECR) Committees. The 1st ECR Committee will conclude its term in June 2024, and the 2nd ECR Committee will start in July 2024.

This was followed by updates on the ongoing interlaboratory research initiatives led by members of the 1st ECR (Early Career Researcher) Committee. Kyowon Jeong (University of Tübingen) presented the updates on the Gold-standard Dataset research initiative, and Luca Fornelli (University of Oklahoma) presented the updates on the MS^2 Fragmentation research initiative.

Following the updates on interlab initiatives, a panel discussion was held with 4 experts on biopharmaceuticals and 4 experts on clinical research.
Panelists in Biopharmaceuticals:
Pavel Bondarenko (Amgen)
Lauren Adams (Merck)
Jake Melby (AstraZeneca)
Weijing Liu (Thermo Fisher Scientific)
Panelists in Clinical Research:
Julia Chamot-Rooke (Insitut Pasteur)
Bryon Drown (Purdue University)
Elena Dominguez-Vega (Leiden University Medical Center)
Christoph Gstöttner (Leiden University Medical Center)

The panel discussion began with a question for the audience on how they apply top-down proteomics. In general it was stated that, amongst other applications, TDMS is used for characterizing antibody/drug conjugates and searching for protein biomarkers in human samples. The audience was challenged to ask specific questions to the panelists. This resulted in a lively discussion with good interactions.

What is the role of top down in pharmacokinetics and antibody drug conjugates?

Answer 1: TDMS gives a more complete picture of how the drug conjugate is modified *in vivo*, also when calculating drug-antibody ratio, and how payload increases over time.
Answer 2: TDMS used to identify what modifications are happening to antibodies in the patient? Specifically, how glycoforms are changing over time. Most importantly, it poses new questions about the implications that post-translational modifications may have on the patients.

Given the current community initiatives, what other resources would be valuable for this type of research?
Answer 1: Data processing is a major challenge. Every lab seems to have their own way of doing things, and so a unified way to process the data would be valuable. Also unifying pre-MS separations, such as more standardized methods for reversed-phase LC, SEC, HILIC, CE, etc.
Answer 2: Also would like to see more initiatives focused on front-end standardization, that is, separations that are compatible with nanoflow sources (both native and denatured workflows).
Answer 3: Expanding on the 'golden dataset' initiative, a focus on including data which aims to localize modifications would be important. For example, cation exchange has been useful to separate acidic vs basic variants, but ideally we would like to localize these types of proteoforms directly, without having to collect fractions from cation exchange or other chromatographic methods.

From someone who runs a core facility, switching instrumentation over from bottom-up to top-down is a hassle, particularly the chromatography. How to best optimize this?
Answer: It's best to use commercial columns for top-down separations to make the switch over easier.
Paul Danis recommended MP682 where the authors used bottom-up, top-down, and middle-down approaches, all on the same LC setup.

What advancements do you envision will drive the next five years of drug discovery and biopharmaceutical research?
Answer: new technology such as CD-MS and single cell proteomics is exciting. There have been recent revolutions in oncology research involving precision medicine, but the first step toward being able to bridge the gap between oncologists and mass spec/TDMS practitioners is introducing new tech that can actually identify different proteoforms for precision medicine.

How different are the proteoforms for different individuals in clinical research? Can you comment on that?
Answer: It depends on the protein, but there is a difference between precision medicine (study and targeting of proteoforms) and personalized medicine (targeting the proteoform unique to the patient).

How often do you find proteins that are extensively modified? There is an expectation that top-down will reveal the many concurrent modifications that could exist on a single protein (for example, 8 phosphorylations on one protein), but often times no or few modifications are detected.
Answer: Yes, we call these 'vanilla' proteoforms – many proteins just aren't that modified. But the implication of the modification could be important.

How do you have confidence in the assignments when you have an essentially infinite search space in untargeted TDMS, for example when analyzing serum proteoforms?
Answer: Many software packages will confine the search space, for example the ProSight suite utilizes the Uniprot database and only considers annotated modifications in its assignments. There are still caveats, like hemoglobin for example, which has so many modifications, many of them sequence variants. One person is unlikely to have 500 sequence variants, so there is still a way to go for confidently making these types of assignments.

TDMS is still considered new (compared to bottom-up proteomics), so what do you think it will need to be incorporated into the diagnostic landscape?
Answer 1: Quantitation. The means to quantify individual proteofoms is necessary. We also need more research on determining the structure/function relationship – need to explore what the proteoforms are doing.
Answer 2: on quantitation as well –We need to make progress on this front and could start with the development of good standards for quantitation.
Answer 3: At least for antibody standards, those experiments are now routine and antibody standard kits are commercially available.

Can you share success stories with top-down? Any cool results?
Julia Chamot-Rooke described a discovery where they found a new proteoform of a bacterial protein with a previously unknown modification. The PTM was discovered to help the bacteria cross the epithelial barrier in the human throat and enter the blood stream.
Jake Melby said that identifying proteoforms, especially those that change from cell to cell, and being able to decipher those differences even in an MS1 spectrum is very satisfying.
Bryon Drown discussed differences in phosphorylated kinases, that there are so many proteoforms hiding under a single peak in the MS1 spectrum that you can only see with high-quality fragmentation data – for example a triply phosphorylated form not evident in the MS1.

For proteins that are multiply modified – how to determine what to search for when database searching in order to identify the correct proteoform? Can we integrate bottom-up proteomics in a meaningful way to help quantify or identify some of the proteoforms we see with top down or even

those that are missed because they're only 1% abundant? Similarly, we may detect a proteoform with 8 phosphorylation sites by TDMS, but bottom-up would detect 38 peptides – how to reconcile this?

Answer 1: Often there is no way to observe some lower abundant proteoforms by top-down.

As for how to reconcile data from bottom-up which identifies lots of proteoforms, there should be standardization in how quantitation is approached, so can we learn something from how bottom-up proteomics does it?

Answer 2: Yes, we as a top-down community should define how to quantify.

The problem with quantification is that we all use different software packages, so a quantitative measurement can come from the most abundant charge state, or isotopic envelope, etc. So how do we think about unifying the different approaches which we call quantitation? It doesn't seem that the community is on the same page on this front.

Answer 1: Agreed. There are no guidelines at all, so this needs to be developed. Need to consider what the best approach is.

Answer 2: Yes, we need to put names to how we perform quantitation and create an open discussion surrounding it. For example, bottom-up people have LFQ vs iBAQ and we don't have an equivalent metric.

Answer 3: Agreed, we need to find a name for our metric so that the field is speaking the same language.

How does a modification change the ionization efficiency of different proteoforms? Has anyone encountered this or considered it?

Answer from Neil Kelleher: very rarely is ionization efficiency a problem (citing previous paper of histone proteoforms). The charge states may change slightly across proteoforms but overall the PTMs do not alter the ionization efficiency or distribution of proteoforms.

After the workshop ended, all panelists expressed their enthusiasm about the discussion, and various participants from the audience gave encouraging feedback. In addition, the 2nd CTDP ECR Committee is currently planning new interlaboratory research initiatives and encouraging interested labs/researchers to participate.