A roadmap to get host cells proteins analysis by mass spectrometry in a GMP environment

Cyrille C. Chéry
Analytical Development Sciences for Biologicals

Inspired by **patients.**
Driven by **science.**

Acknowledgement

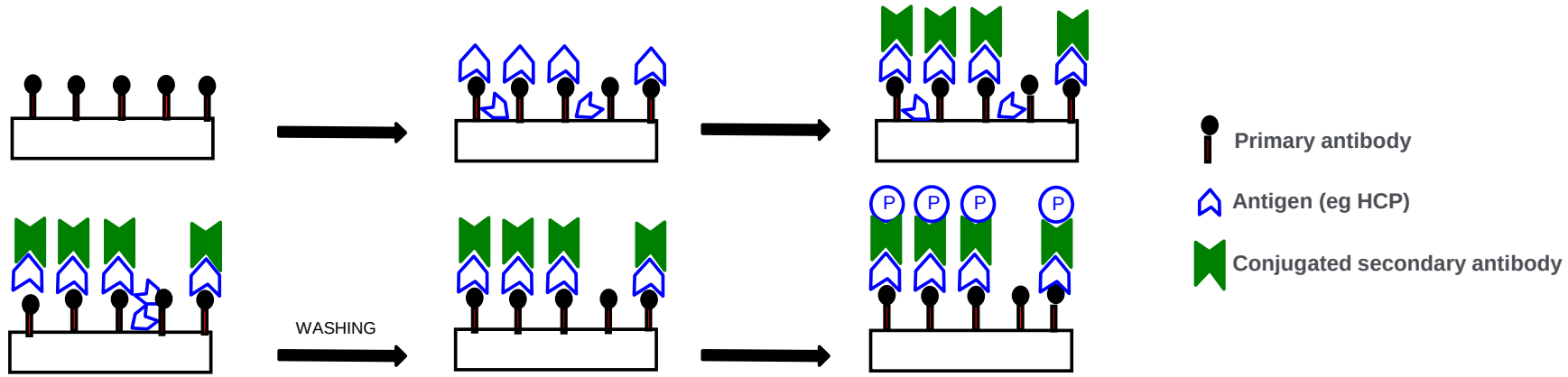
- Fouad BENAMIRA, UCB, compliance expert
- Will BURKITT, UCB, characterization expert
- Christine CARAPITO, University of Strasbourg, biochemical applications of mass spectrometry
- Annick GERVAIS, UCB, initiator of the project
- Somar KHALIL, UCB, mass spec SME
- John O'HARA, UCB, characterization

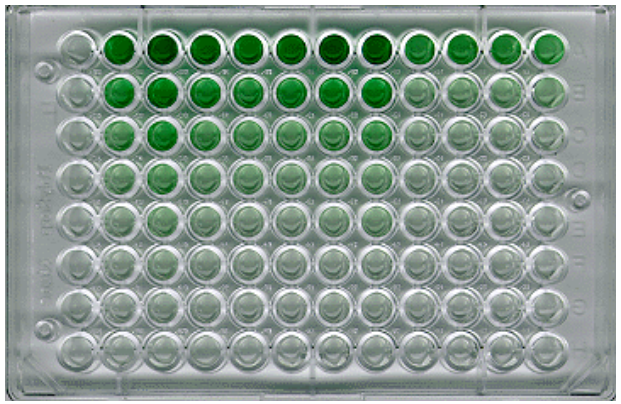
Agenda

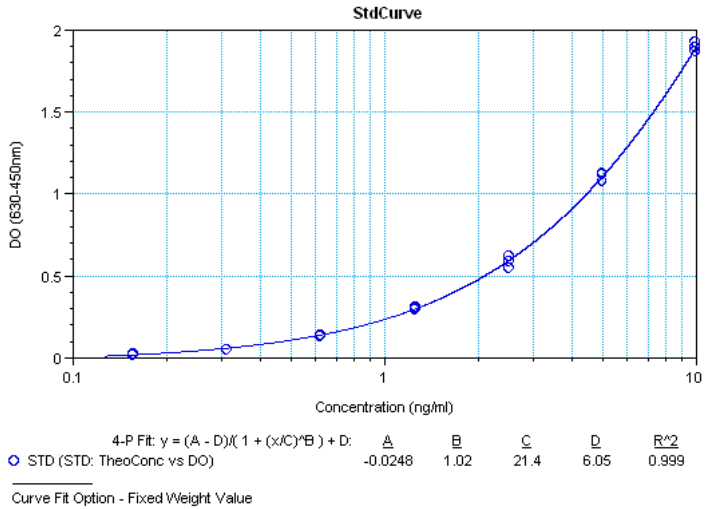
1. ELISA vs. mass spectrometry
2. Qualification challenges
3. Validation challenges
4. Specification challenges

Yes, there might be some challenges

Residual Host Cell Proteins assay by ELISA







ELISA for the quantification of HCP is the current gold standard

But

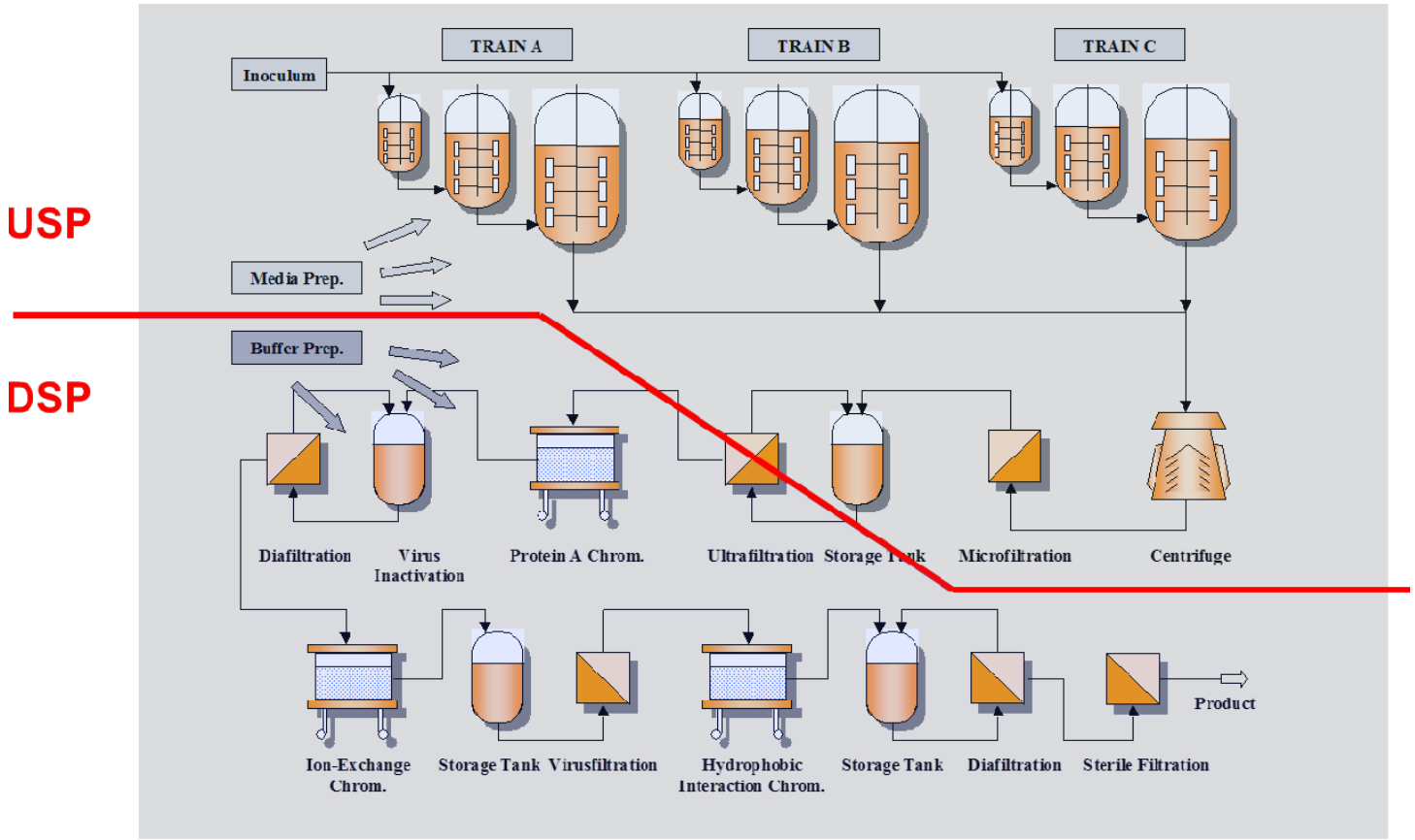
- Production of reagents requires animal sacrifice: against the 3R principle
- ELISA is sensitive to over-immunogenicity of certain HCPs
- Coverage limitation: high % of total HCPs are recognised, but some might be missed
- ELISA gives a global result, no information on the individual HCPs
- The development of a new kit takes ca. 2-3 years

Why mass spec?

- Mass spec does not rely on recognition of HCP by antibodies:
  - It relies on the detection of charged ions
  - ***physicochemical*** method rather than ***immunochemical*** method
- It allows the identification of individual HCPs
- It therefore gives greater understanding of HCP clearance, potential risks due to certain HCPs
- The development of an MS method is short compared to ELISA new kit generation
- The detection can be
  - agnostic (e.g., using a so-called data-independent acquisition), but semi-quantitative
  - agnostic (e.g., using a so-called data-dependent acquisition), quantitative
  - or targeted towards certain HCP (e.g., using a so-called multiple reaction monitoring)

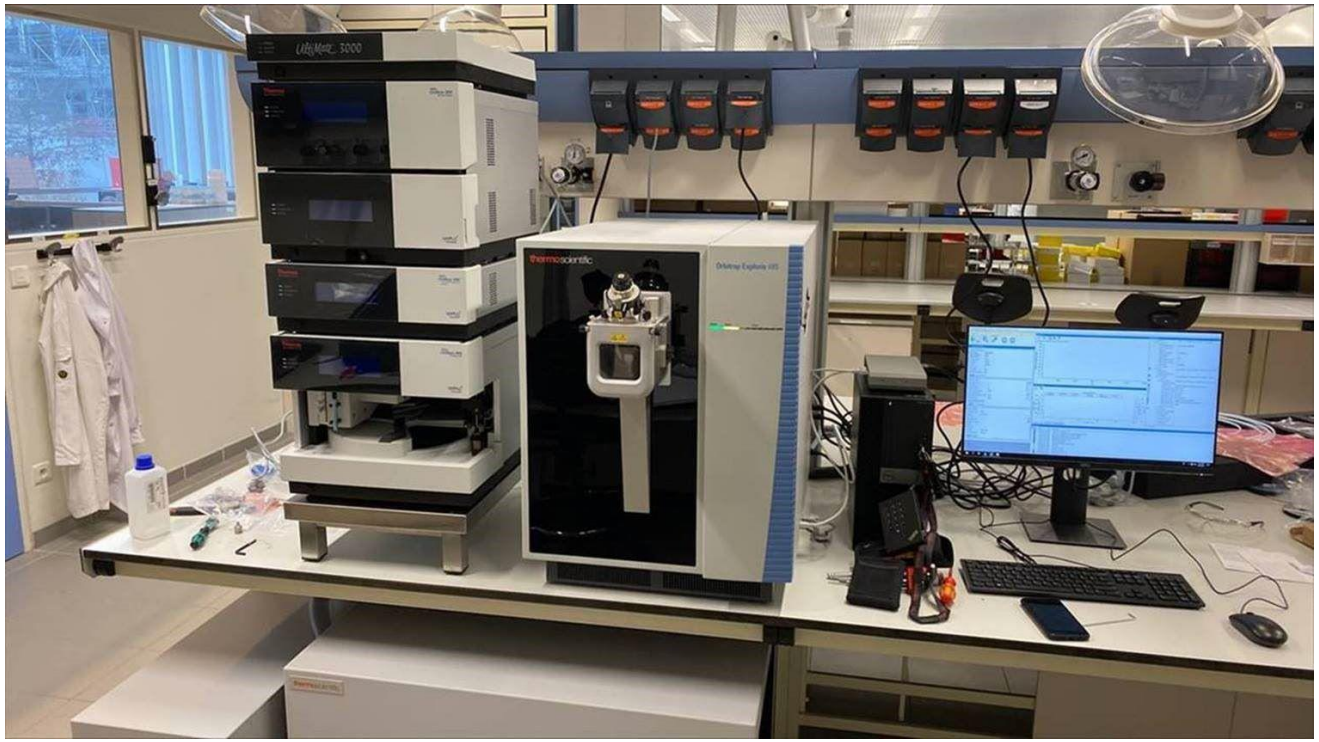
Mass spectrometry is already a characterisation tool

- UCB strategy is to align the characterisation and release MS methodologies, via, among other things, the use of the same instruments
- Mass spectrometry, via data dependent acquisition, is used for process understanding, especially downstream (DSP) clearance of host cell proteins



Qualification of the instrument and the software for QC use

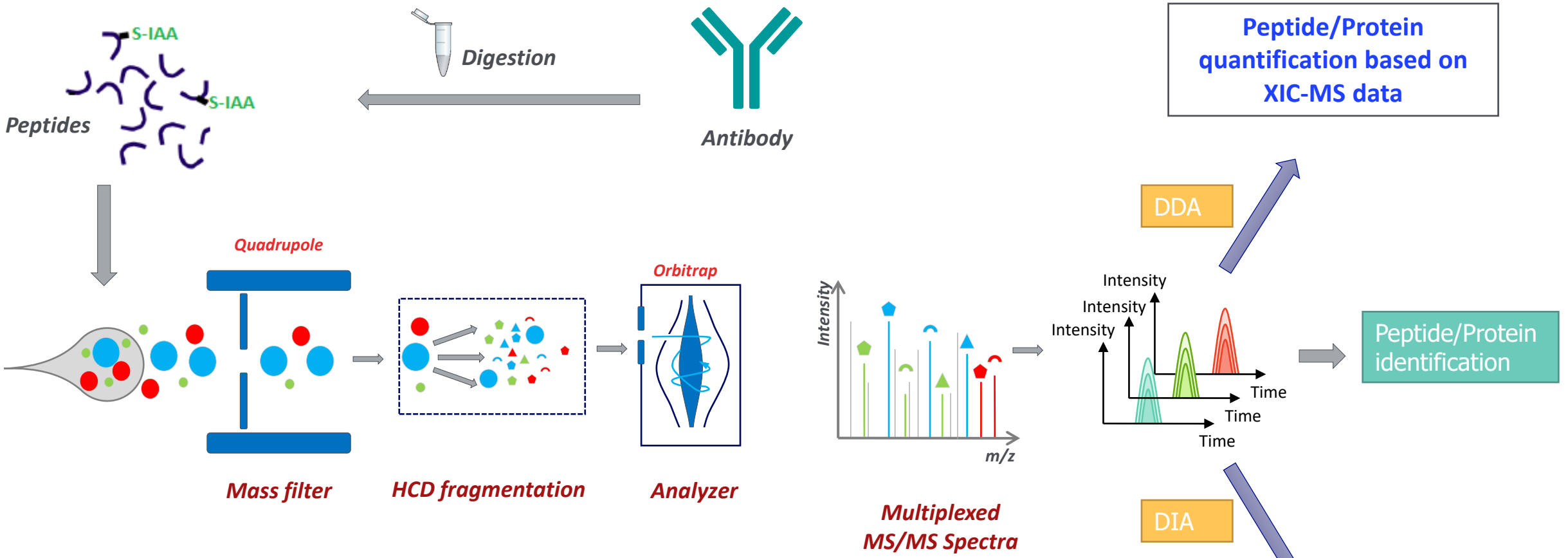
- UCB installed an instrument end 2021, Orbitrap Exploris 480
- GMP qualification of the hardware successful



…

- The GMP qualification of the software is tougher…

Mass spectrometry principle

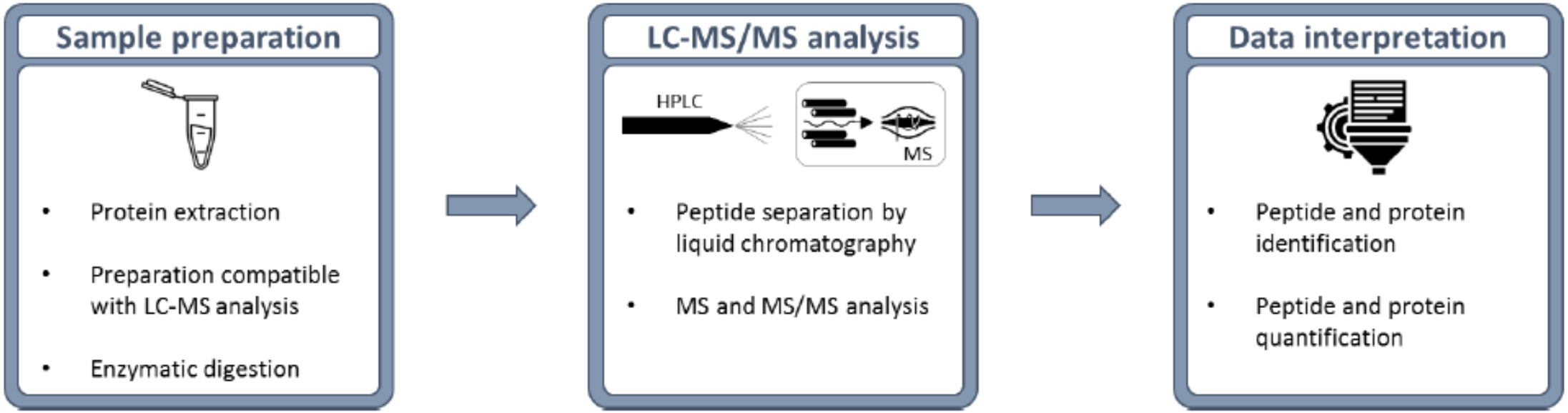


HCD: high energy collisional dissociation
DDA: data-dependent acquisition / DIA: data-independent acquisition
XIC: extracted ion chromatogram

Figure from: S. Hessmann, PhD thesis p.72, Strasbourg University, 2021

Simplified flow

- The sample and data flow for the identification/quantification of HCP are basically the same as for proteomics:

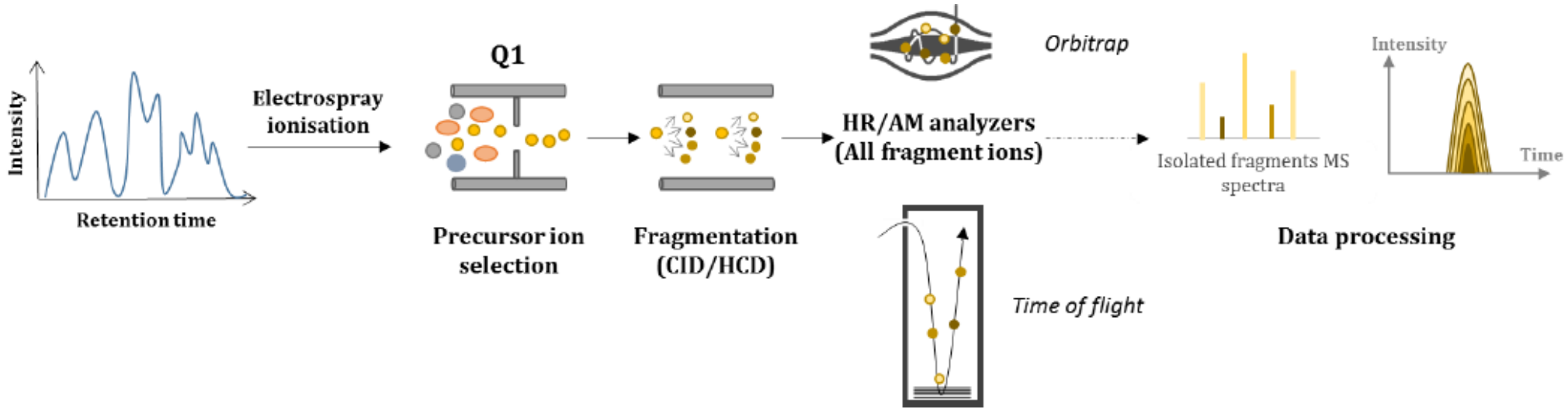


- (Extra challenge for HCP when compared to classical proteomics: dynamic range, ie, presence of the active)

Figure from: S. Hessmann, PhD thesis p.53, Strasbourg University, 2021

Qualification of the software: challenges

- The flow of data is the challenge
- Let us assume ca. 10^2 proteins at Drug Substance, it corresponds to ca. $5x10^2$ peptides
- For certain proteins, one protein can correspond to even 10 peptides…
- This challenge is independent from the type of instrument: Orbitrap or time-of-flight MS



CID: collision induced dissociation / HCD: high energy collisional dissociation
HR: high resolution / AM: accurate mass

Figure from: S. Hessmann, PhD thesis p.72, Strasbourg University, 2021

Note on units

- 1 ppm = 1 (ng HCP) / (mg active)
- 1 ppb = 1 (pg HCP) / (mg active)

Which acquisition mode to choose? DDA or DIA?

- Both modes are non-targeted, ie, do not focus on a certain HCP list
- Thus, both modes are able to detect "unexpected" HCP (as long as above LOD)

- DIA, data-independent acquisition: all precursor ions within a certain mass/z window are fragmented
- Pro: No bias towards certain ions / Con: Link between MS1 and MS2 lost
- There is no software solution available, to our best knowledge, that is GMP-compliant

- DDA, data-dependent acquisition: the most intense ions are fragmented
- Pro: Link between MS1 and MS2, this is a plus for complex mixtures / Con: bias towards most intense ions
- There might be GMP-compliant software
- UCB currently applies this approach

- (There are targeted approaches, like MRM, multiple-reaction monitoring
- Disadvantage: one might be "blind", ie, miss an HCP that is not on the target list)

The amount of data says it all

Above 10 ppm

Protein Accession	Quantity ppm
A0A3L7HRL5	25.411
A0A8C2LCT4	24.769
A0A3L7HKZ9	16.452
A0A8C2QCP7	13.637
G3HEE4	13.375
A0A061IGC3	10.182

6% by number, i.e., 6 individual HCP out of ca. 100

103.8 ppm

1-10 ppm

Protein Accession	Quantity ppm
A0A8C2LL08	7.956
A0A3L7I8L8	7.829
A0A8C2N173	6.342
A0A3L7I619	6.299
Q9JKY1	5.790
A0A3L7IGF2	5.175
A0A061IR24	5.022
A0A3L7H6B5	4.592
A0A8C2QFD1	4.471
A0A3L7GQU0	3.506
A0A8C2M6Q7	2.822
P17244	2.416
A0A061HZT7	2.294
A0A8C2MFT3	1.505
A0A061IHK7	1.344
A0A3L7HUM4	1.271
A0A3L7I3L8	1.226
A0A8C2N5M3	1.127
A0A8C2LA82	1.123
G3I6T1	1.008

19% by number

73 ppm

At least 3 peptides behind each protein

0.1-1 ppm

Protein Accession	Quantity ppm
A0A8C2QG32	0.929
G3HLK3	0.891
A0A8C2MNP2	0.886
A0A061I466	0.772
A0A8C2M060	0.762
G3HCW9	0.727
G3HN14	0.627
G3H7K5	0.582
A0A8C2QHB3	0.568
A0A8C2L9G7	0.545
A0A8C2M8H1	0.518
A0A061I5B9	0.426
A0A061I1H8	0.383
A0A8C2MDL6	0.379
G3HDI8	0.368
A0A061I094	0.358
G3II69	0.350
A0A3L7ICZ2	0.285
A0A8C2LMG1	0.285
G3GXZ0	0.264
A0A3L7IHU5	0.261
A0A061HUH1	0.258
G3H160	0.250
A0A8C2MY60	0.241
A0A3L7IC53	0.231
A0A8C2LPD2	0.198
A0A2Z6LBQ0	0.173
G3HNJ3	0.171
P62629	0.169
A0A3L7GSA7	0.163
A0A061IAL7	0.158
A0A3L7IDS1	0.149
A0A3L7H0F5	0.145
A0A3L7ICE6	0.136
G3H1W4	0.135
A0A3L7HCT3	0.131
G3HP75	0.131
A0A8C2QFV1	0.126
A0A3L7H3G0	0.126
A0A8C2QII0	0.123
A0A061I	0.120

39%

14.5 ppm

10-100 ppb

LOD < 100 ppb < LOQ

Protein Accession	Quantity ppm
G3I2M1	0.099
A0A3L7HBY6	0.097
A0A8C2ME67	0.096
A0A3L7IAN1	0.096
A0A8C2MCT4	0.093
G3HR88	0.089
A0A8C2M148	0.077
A0A3L7II69	0.073
A0A8C2M8H3	0.069
A0A061I523	0.061
A0A061I443	0.056
A0A8C2M162	0.056
A0A8C2LND3	0.054
G3INC5	0.054
G3HQP8	0.053
A0A3L7GLZ2	0.050
A0A061IQJ9	0.047
A0A061HW36	0.045
A0A061IMX1	0.044
G3I1H5	0.040
A0A8C2MK02	0.039
A0A3L7IC78	0.038
G3GXB0	0.038
A0A8C2LG22	0.036
A0A061IRB6	0.033
A0A8C2QF63	0.028
G3IA94	0.027
G3HT19	0.026
A0A8C2LRM9	0.023
G3GTT2	0.022
A0A3L7HUQ4	0.022
A0A3L7HJ58	0.022
G3H2W6	0.018
A0A8C2M350	0.016
G3GZA7	0.016
A0A8C2QPC5	0.015
A0A3L7IBK4	0.012
A0A8C2MC57	0.012

36%

1.8 ppm

Qualification of the software: challenges

- The sheer flow of data, MS-MS data on typically 10^2 proteins, from ca. $5x10^2$ peptides, is the challenge
- Open-based proteomics software is available, but hardly GMP compatible: not traceability of data, user, etc.
- Some commercial software is available, but
  - Either not up to Part 11 CFR compliance (according to us)
  - Either not up to full proteomics

GMP release method
Tough nuts to crack: validation drives me nuts

- Goal is to validate, according to ICH Q2 (R2) the quantification method
- The analytical target profile (ATP) is likely to be:
  - The reportable result is the total amount of HCP
  - The total error (combination of trueness and precision) accepted on the result could be 50-100 % ("50-100% uncertainty on the reportable result")
- The validation could therefore be performed
  - using a mix (which one? Synthetic mixes exist and are commercially available)
  - or an in-house sample (drug substance), spiked with a synthetic mixture of representative proteins

- We are thinking of a validation focusing on a few HCP (even if the mixture is complex), by selecting analytically challenging analytes, because it is not reasonable to produce data on:
  - Accuracy
  - Linearity of results
  - Range, including quantitation limit
  - Robustness

For ca. 150 analytes…

GMP release method
Specification *on the total HCP?*

- How should a sound specification setting be performed?
  - "<100 ppm **by ELISA**" (from WHO)
  - Keep in mind that the reading principles of ELISA and MS methods are totally different (apples vs. pears)
    - ELISA is based on a immunological principle
    - MS is physico-chemical
- It sounds reasonable, in terms of patient safety, to set the specification on the total HCP present
- The limit could be totally different from 100 ppm

The MS method is able to identify the individual HCP

- Should we risk assess for each HCP identified?
  - Where do we set the threshold: above 1 ppm?
  - Against which reference list of HCP of concern?
    - The industry has build an HCP list, eg, via Biophorum: BPDG HCP risk assessment (*)
- A specification on each single HCP is not reasonable,
  the risk assessment should be enough to have a proper understanding of the risk

Example of the individual results

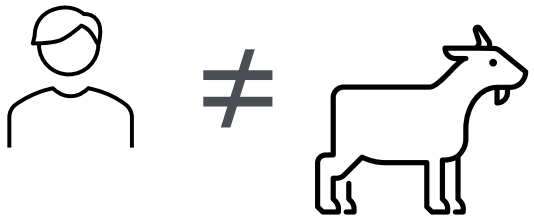
Protein Accession	Quantity ppm
A0A8C2LL08	7.956
A0A3L7I8L8	7.829
A0A8C2N173	6.342
A0A3L7I619	6.299
Q9JKY1	5.790
A0A3L7IGF2	5.175
A0A061IR24	5.022
A0A3L7H6B5	4.592
A0A8C2QFD1	4.471
A0A3L7GQU0	3.506
A0A8C2M6Q7	2.822
P17244	2.416
A0A061HZT7	2.294
A0A8C2MFT3	1.505
A0A061IHK7	1.344
A0A3L7HUM4	1.271
A0A3L7I3L8	1.226
A0A8C2N5M3	1.127
A0A8C2LA82	1.123
G3I6T1	1.008

(*) "High-risk" host cell proteins (HCPs): A multi-company collaborative view Biotechnol Bioeng 2021

The MS method does not give a result linked to immunogenicity
Could we miss something?

Not really!

- The typical limit of detection for an individual HCP by MS is 10 ppb in our laboratory
- An HCP, below 10 ppb, has to be hugely immunogenic to be of any concern
- Furthermore, there is no certainty that it would be picked up by ELISA, because the ELISA antigens stem from animals, not humans…



Literature: LOD ca. 100 ppb

(Temporary) conclusions

- Proteomics MS is promising as a GMP tool
- Main challenges:
  - Software qualification
  - ICH Q2(R2) method validation, ie, finding a sound design
  - Specification setting

Inspired by **patients.**
Driven by **science.**

Thank you