Electronic Supplementary Material (ESI) for Molecular Omics.

An experimental design to extract more information from MS-based histone studies

Laura De Clerck; Sander Willems; Simon Daled; Bart Van Puyvelde; Sigrid Verhelst, Laura Corveleyn; Dieter Deforce and Maarten Dhaenens.*

Laboratory of Pharmaceutical Biotechnology, Ghent University, Ottergemsesteenweg 460, 9000 Ghent, Belgium

* Corresponding author: maarten.dhaenens@ugent.be

ORCID

Laura De Clerck: 0000-0001-9465-4387
Sander Willems: 0000-0002-7124-610X
Simon Daled: 0000-0001-8129-2217
Bart Van Puyvelde: 0000-0002-7368-1762
Sigrid Verhelst: 0000-0002-6791-3782
Laura Corveleyn: 0000-0002-4167-3581
Dieter Deforce: 0000-0002-0635-661X
Maarten Dhaenens: 0000-0002-9801-3509

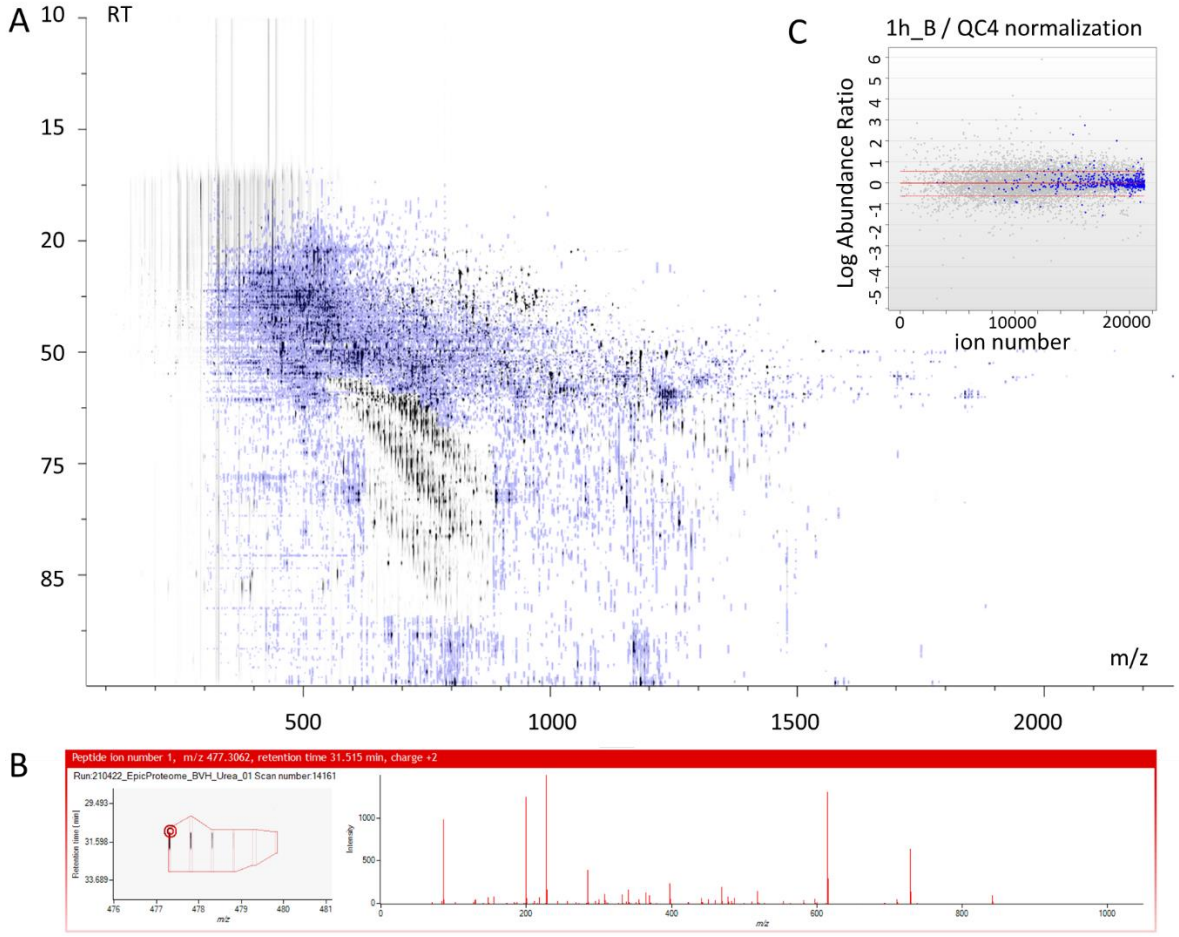
Keywords

histone post-translational modifications; LC-MS/MS; proteomics; identification; HDAC

Content

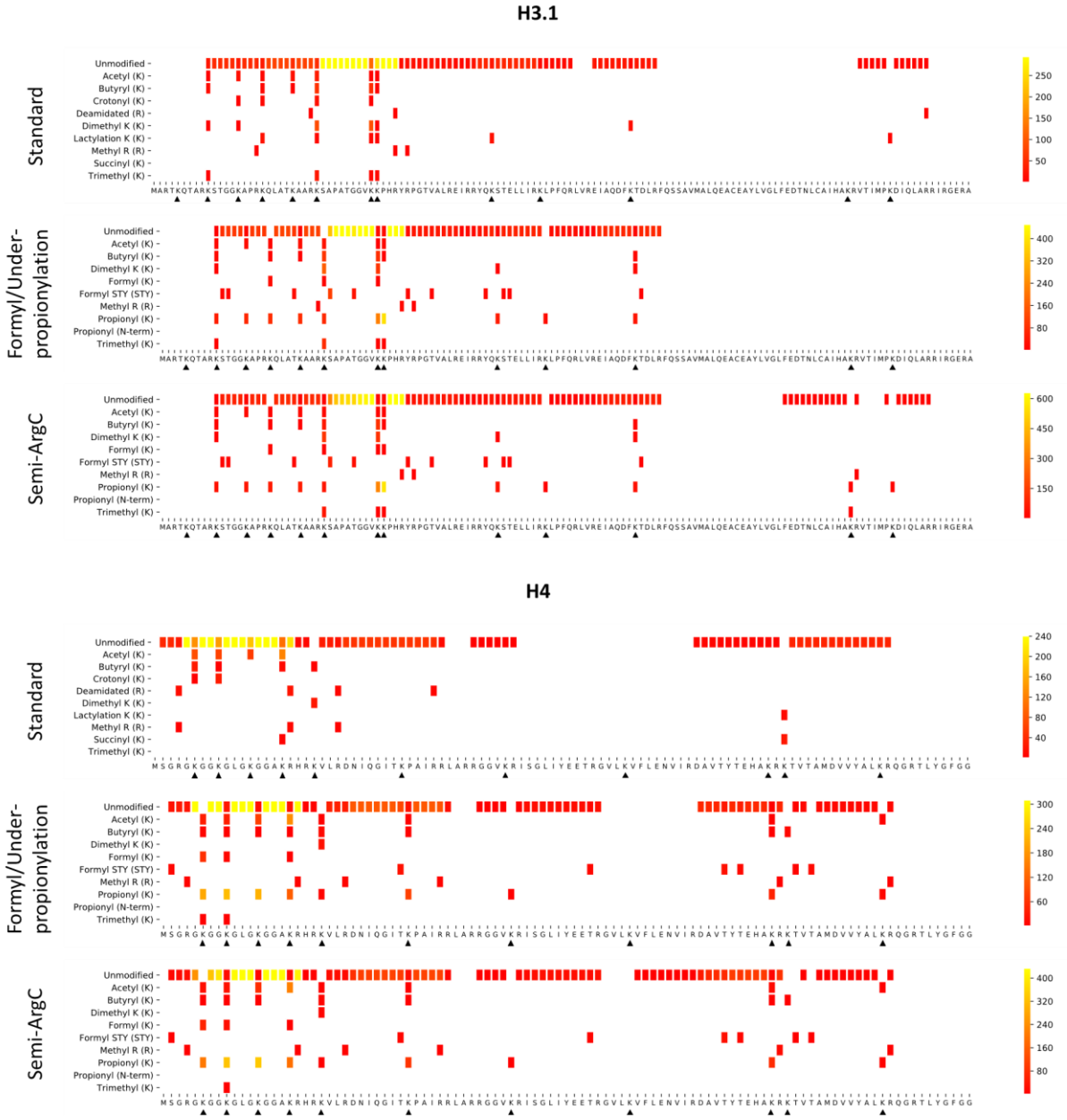
FIGURES

Supplementary Figure 1



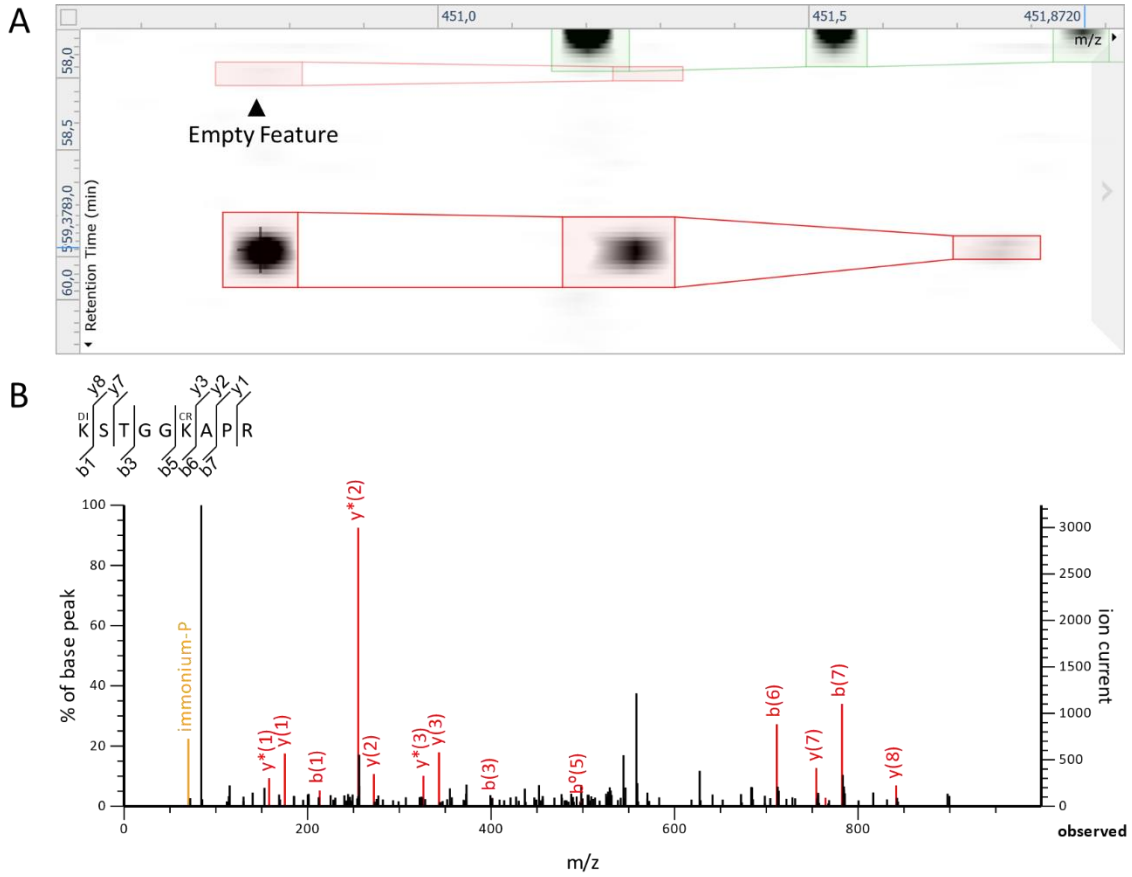
Supplementary Figure 1: Filtering, PSM matching and normalisation done in Progenesis QIP 4.2. A. After alignment and feature detection, ions were filtered based on charge state (2+ to 10+ are retained) and the profile of polyethyleneglycol (PEG) was excised from the run manually (white area between m/z 550-900 and RT 50-90). All retained ions (21293) are highlighted in blue. Time axis equally spaces MS1 scans of the DDA run, resulting in a compressed time line when the cycle time increases due to higher number of targeted precursors. The complete project can be consulted on ProteomeXChange by downloading Progenesis QIP for free and opening the *180329_HDAC_HDDDA_for_HDMSE_180201MD_PEG_Removed.ProgenesisQIPArchive* file (no license required). **B.** A feature (left) is defined as an ion with a restricted retention time window and at least one isotopic peak to determine the charge. The circle (left) depicts the point where the MSMS spectrum (right) was taken. **C.** Normalization of all ion intensities was done using histone peptides (blue dots) against one of the quality control samples (QC4), which is a mixture of all samples in the experiment.

Supplementary Figure 2



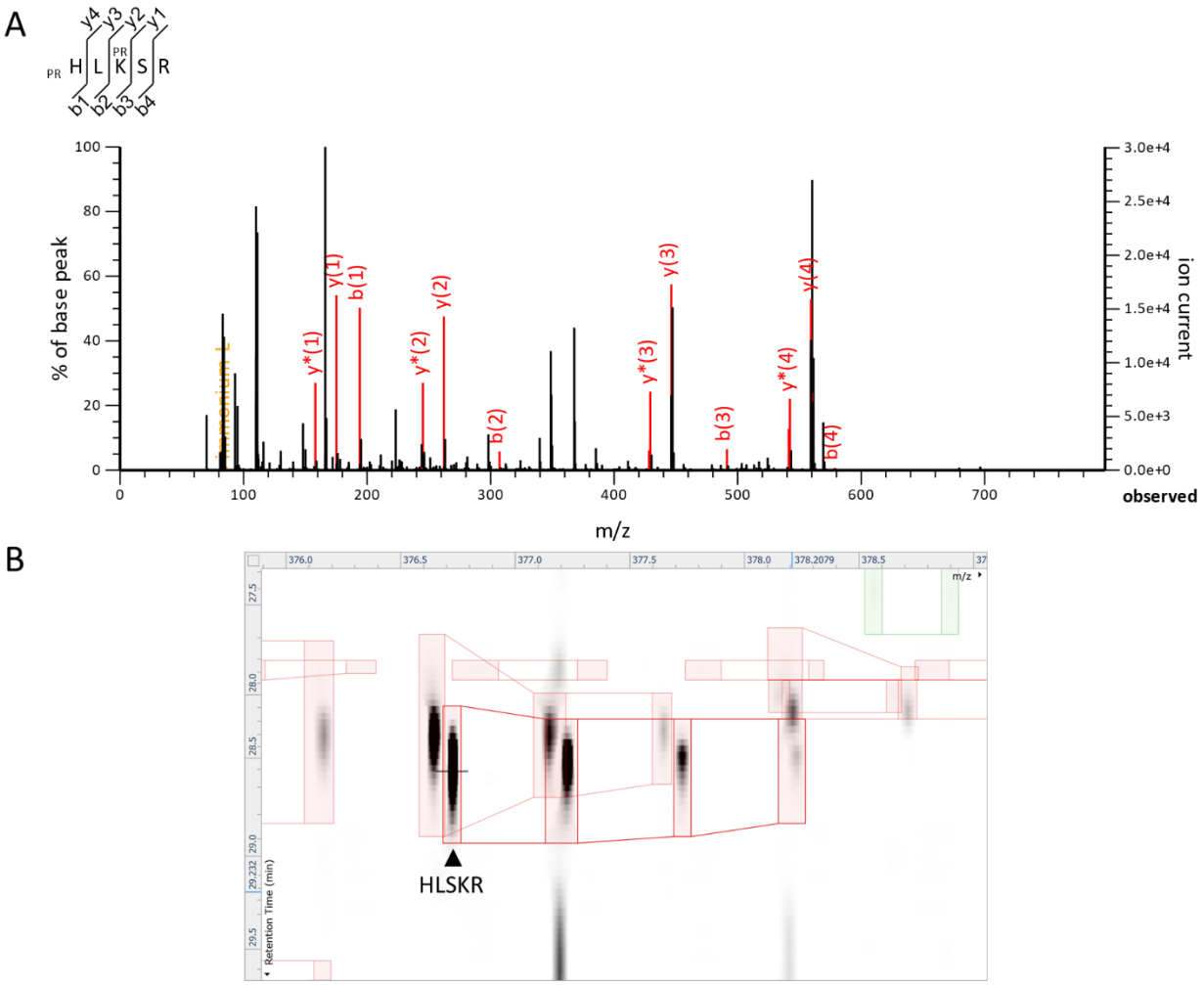
Supplementary Figure 2: Coverage of H3.1 and H4. The coverage plots were created with an in-house script (available at https://github.com/swillems/coverage_plots, accessed on 17/08/2021), which requires the MGF files and the protein database (*.FASTA) used to search the MGF files to map the identified peptides and PTM against the protein backbone. The displayed modifications depend on the search parameters used in Mascot Daemon. The scale to the right of each plot indicates the spectral count (relative to each histone from red to yellow) of each modification/amino acid residue. Black arrow heads indicate the lysine (K) residues.

Supplementary Figure 3



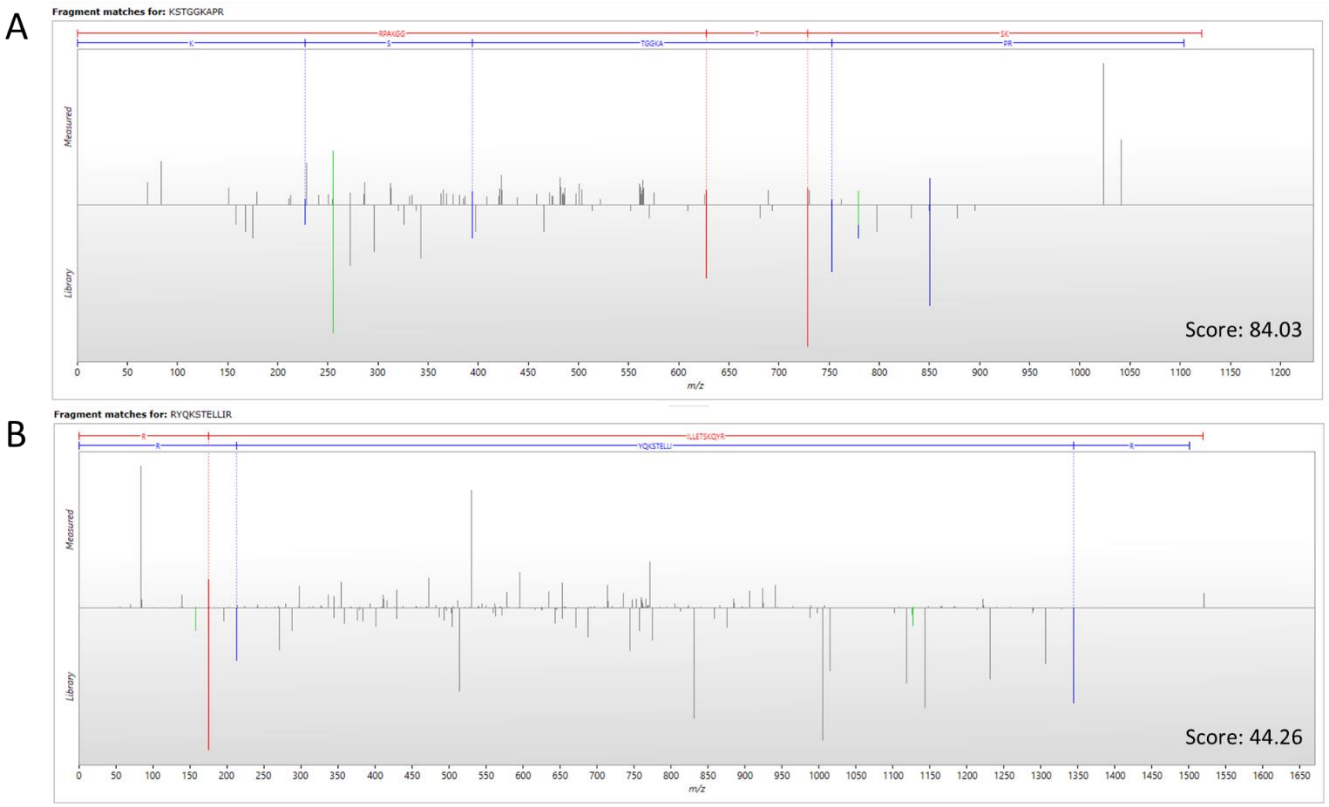
Supplementary Figure 3: Two annotations not matching the expected precursor ion intensity pattern following the standard search. A. For the unmodified peptide H3(64-69; KLPFQR), a total of 11 PSMs were found, 10 of which were eluting at 59,5 minutes. Only one was annotated at 58 minutes, which is the one embedded in the targeted acetylated population. Indeed, this annotation was linked back to a feature which was wrongly created by the ProgenesisQIP software on noise signal. **B.** On the other hand, only one PSM was found for H3(9-18) K9Me2/K14Cr. This peptide could have been decrotonylated by HDAC, since it was recently suggested that HDACs are deacylases, with a wide range of acylations as substrates. However, the spectrum below shows that no diagnostic ion of dimethyl or crotonyl was found in the low mass region and that several larger peaks were not explained (Hseiky *et al.*, 2021). Also, a crotonylation on H3K14 is not reported by UniProt. Therefore, a wrong annotation is more likely the reason for this anomaly.

Supplementary Figure 4



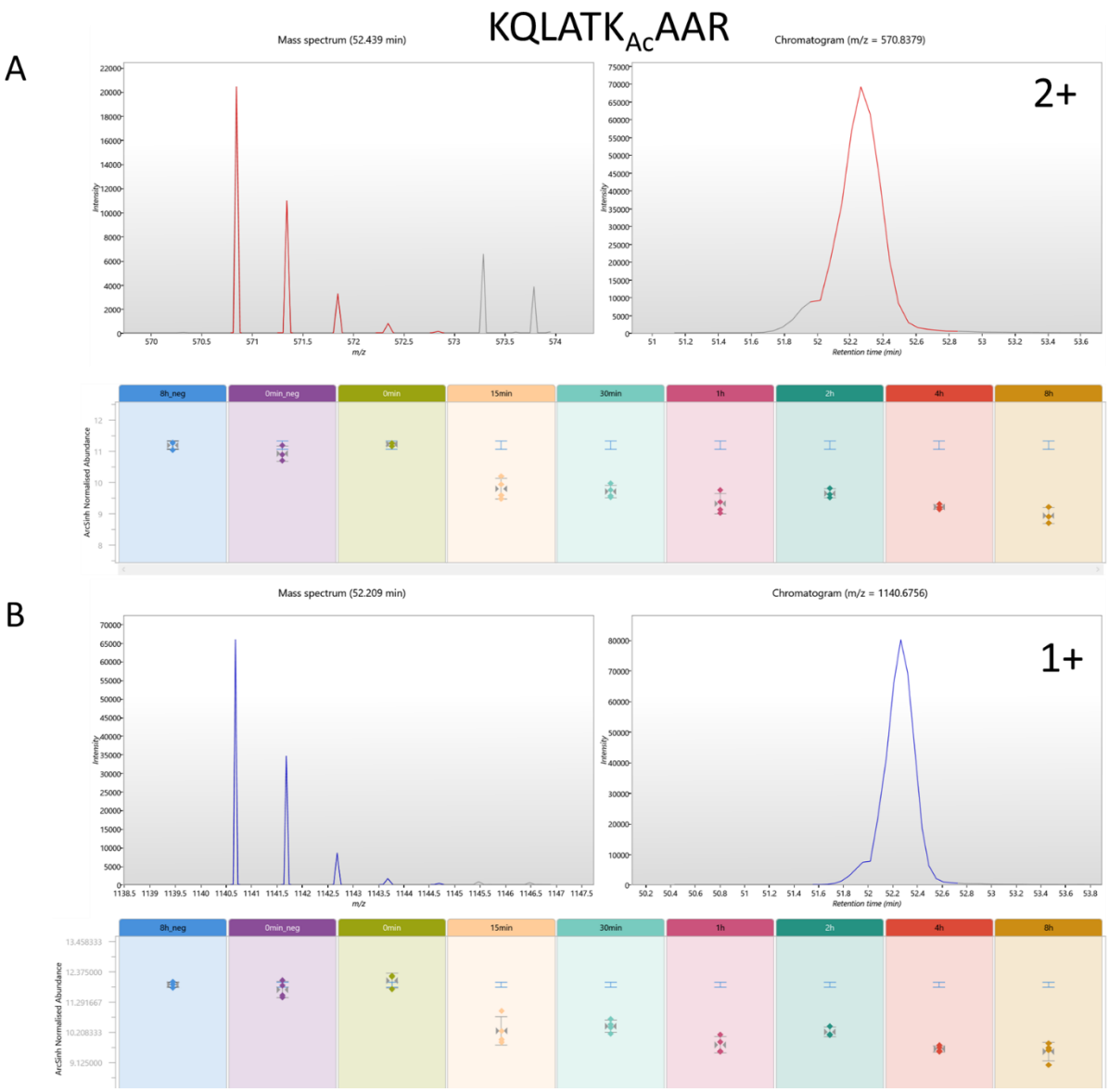
Supplementary Figure 4: One annotation not matching the expected precursor ion intensity pattern following the sample preparation search. A. The Mascot PSM of the HLKSR peptide from H2AZ, displaying full coverage of the short sequence (red fragments), but with several very high unannotated fragments (black) present. This is indicative of a high abundant co-selected precursor that has created a chimeric spectrum. **B.** Closer inspection of the precursor in Progensis QIP showed the co-eluting of a near-isobaric precursor that very plausibly accounts for the unannotated high intensity fragments in the spectrum.

Supplementary Figure 5



Supplementary Figure 5: Mouse Spectral Library matching. A. The fragments matches of the measured (top) spectrum of a precursor ion displaying the profile of an acetylated peptide (**Figure 3C**) and the library (bottum) spectrum of the peptide H3 (K9-18; KSTGGKAPR) K9Me3/S2Ph/K14Ac with a score of 84/100. **B.** The unmodified H3 (53-63; RYQKSTELLIR) was also matched to a precursor from the acetylated ion population (**Figure 3C**) with a score of 44/100. However, the latter has little evidence for being correct. This surfaces the issue of a lack of score thresholding to make spectral library matching a valuable alternative as a search strategy in its current form.

Supplementary Figure 6



Supplementary Figure 6: The proportion of +1 precursor ions. Screenshot from the Progenesis QIP 4.2 project showing the precursor and elution profile of the H3 (K18-24) acetylated peptide in +2 (**A**) and +1 (**B**) charge state. The expression profiles are identical, confirming that the charge state deconvolution applied by Progenesis QIP correctly transferred the identification from the +2 to the +1 precursor ion. Both signals are equally intense, making the +1 precursor ion a perfect backup in case a co-eluting near-isobaric precursor would interfere with the +2 precursor ion (e.g. Supplementary Figure 3). Here, the shoulder at both 1+ and 2+ represents the K_{Ac}QLATKAAR form, which can only be distinguighed in hSWATH acquisition (De Clerck *et al.*, 2019).

TABLES

Supplementary Table 1

Supplementary Table 1: 27 Additional precursor ions identified in the Semi-ArgC search. Using the N-terminal propionylation and C-terminal amino acid, they can be subclassified into preprocessed or aspecifically cleaved (whether or not because of underpropionylation) or in-source decay fragments.

Sequence	Modifications	Cleavage type
G.KGGKGLGKGGAKR.H	K5 Prop|K8 Prop|K12 Prop|K16 Ac	Lack of N-terminal propionylation indicates in-source decay fragmentation
G.KGGKGLGKGGAKR.H	K5 Prop|K8 Prop|K12 Prop|K16 Ac	
G.KGGKGLGKGGAKR.H	K5 Prop|K8 Prop|K12 Prop|K16 Prop	
G.KGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Prop	
G.GKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Ac	
G.GKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Prop	
G.GKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Prop	
K.GGKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Ac	
K.GGKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Ac	
K.GGKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Prop	
K.GGKGLGKGGAKR.H	K8 Prop|K12 Prop|K16 Prop	
G.KGGKGLGKGGAKR.H	(N-term) Prop|K5 Fo|K8 Prop|K12 Prop|K16 Prop	N-terminal propionylation indicates aspecific cleavage during digestion.
G.KGGKGLGKGGAKR.H	(N-term) Prop|K5 Prop|K8 Prop|K12 Prop|K16 Ac	
G.KGGKGLGKGGAKR.H	(N-term) Prop|K5 Prop|K8 Prop|K12 Prop|K16 Prop	
K.GGKGLGKGGAKR.H	(N-term) Prop|K8 Prop|K12 Prop|K16 Ac	
K.GGKGLGKGGAKR.H	(N-term) Prop|K8 Prop|K12 Prop|K16 Prop	
G.GKGLGKGGAKR.H	(N-term) Prop|K8 Fo|K12 Prop|K16 Ac	
G.GKGLGKGGAKR.H	(N-term) Prop|K8 Fo|K12 Prop|K16 Prop	
G.GKGLGKGGAKR.H	(N-term) Prop|K8 Prop|K12 Ac|K16 Ac	
G.GKGLGKGGAKR.H	(N-term) Prop|K8 Prop|K12 Prop|K16 Ac	
G.KGLGKGGAKR.H	(N-term) Prop|K8 Fo|K12 Prop|K16 Prop	
G.KGLGKGGAKR.H	(N-term) Prop|K8 Prop|K12 Prop|K16 Prop	
K.GLGKGGAKR.H	(N-term) Prop|K12 Prop|K16 Prop	
G.LGKGGAKR.H	(N-term) Prop|K12 Prop|K16 Prop	
R.GKGGKGLGK.G	(N-term) Prop|K5 Prop|K8 Prop|K12 Prop	Cleavage type at the C-terminus cannot be classified.
R.GKGGKGLGKG.G	(N-term) Prop|K5 Prop|K8 Prop|K12 Prop	
R.GKGGKGLGKGGAK.R	(N-term) Prop|K5 Prop|K8 Prop|K12 Prop|K16 Prop	

REFERENCES

De Clerck, L. *et al.* (2019) 'hSWATH: Unlocking SWATH's Full Potential for an Untargeted Histone Perspective.', *Journal of proteome research*, 18(11), pp. 3840–3849. doi: 10.1021/acs.jproteome.9b00214.

Hseiky, A. *et al.* (2021) 'Small Mass but Strong Information: Diagnostic Ions Provide Crucial Clues to Correctly Identify Histone Lysine Modifications', *Proteomes*. MDPI AG, 9(2), p. 18. doi: 10.3390/proteomes9020018.