Characterization of Complete Histone Tail Proteoforms Using Differential Ion Mobility Spectrometry

Shliaha, Pavel; Baird, Matthew A; Nielsen, Mogens Møller; Gorshkov, Vladimir; Bowman, Andrew P; Kaszycki, Julia L; Jensen, Ole Nørregaard; Shvartsburg, Alexandre A

Published in:
Analytical Chemistry

DOI:
10.1021/acs.analchem.7b00379

Publication date:
2017

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Characterization of Complete Histone Tail Proteoforms Using Differential Ion Mobility Spectrometry

Pavel V. Shliaha,† Matthew A. Baird,‡ Mogens M. Nielsen,† Vladimir Gorshkov,† Andrew P. Bowman,‡ Julia L. Kaszycki,‡ Ole N. Jensen,*†§ and Alexandre A. Shvartsburg‡§

†Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, and Center for Epigenetics, University of Southern Denmark, DK-5230 Odense M, Denmark
‡Department of Chemistry, Wichita State University, 1845 Fairmount, Wichita, Kansas 67260, United States

ABSTRACT: Histone proteins are subject to dynamic post-translational modifications (PTMs) that cooperatively modulate the chromatin structure and function. Nearly all functional PTMs are found on the N-terminal histone domains (tails) of ∼50 residues protruding from the nucleosome core. Using high-definition differential ion mobility spectrometry (FAIMS) with electron transfer dissociation, we demonstrate rapid baseline gas-phase separation and identification of tails involving monomethylation, trimethylation, acetylation, or phosphorylation in biologically relevant positions. These are by far the largest variant peptides resolved by any method, some with PTM contributing just 0.25% to the mass. This opens the door to similar separations for intact proteins and in top-down proteomics.

Histone proteins (H1, H2A, H2B, H3, and H4) are extensively enzymatically modified by post-translational modifications (PTMs) such as N-methylation (me), N-acetylation (ac), or O-phosphorylation (p) that regulate chromatin structure and function by recruiting proteins involved in replication, transcription, DNA repair, and chromatin compaction. Interpreting the PTM-mediated protein language necessitates analyses of intact histones or their large domains preserving the co-occurring multisite PTM information. In particular, the N-terminal domains (tails) protruding from the nucleosome core are strongly enriched in PTMs and can be cleaved off by endoproteinase Glu-C enzyme.

The staggering challenge here is disentangling numerous isomeric proteoforms that feature the same set of PTMs in different positions (PTM localization variants). While liquid chromatography (LC) techniques such as WCX-HILIC can fractionate histones by me or ac content, many variants still coelute. Tandem mass spectrometry (MS/MS) by any mechanism could characterize the individual variants, but fails for mixtures of three or more variants, since those with internal PTM positions yield no unique fragments.

A growing alternative or complement to LC is ion mobility spectrometry (IMS) based on gas-phase transport properties. Linear IMS approaches (e.g., drift tube or traveling-wave) separate ions by mobility (K) at a moderate electric field (E). While these can resolve many isomers including some localization variants, their power in conjunction with MS is constrained by the intrinsically tight correlation between the mass-to-charge ratio (m/z) and collision cross section within a chemical class. That correlation is much looser for differential or field asymmetric waveform IMS (FAIMS) that sorts ions by the difference of mobilities at high and low field strength.

In FAIMS, a gas flow moves ions through a gap between two electrodes. A periodic asymmetric waveform (CV) superposed on the waveform can equilibrate a given species and let it

Received: January 30, 2017
Accepted: April 13, 2017
Published: April 13, 2017
through the gap into MS analyzer.\textsuperscript{14,15} The resolution is maximized by strong homogeneous field in planar gaps, precise waveforms and He/N\textsubscript{2} or H\textsubscript{2}/N\textsubscript{2} buffers.\textsuperscript{16–19} Another path to higher FAIMS resolution is adding gas modifiers (usually organic volatiles) that cluster with ions in the low-field segment and desorb upon collisional heating in the high-field segment. However, this approach was mostly demonstrated for small singly charged molecules rather than peptides or proteins.\textsuperscript{19,20} One problem with extending it to multiply charged ions such as peptides generated by electrospray ionization (ESI) is the proton transfer to vapor modifiers that have higher proton affinity than N\textsubscript{2}, H\textsubscript{2}, or H\textsubscript{2}.\textsuperscript{21} Some proposed means to mitigate that\textsuperscript{22} would be explored in the future. Here we employ the He/N\textsubscript{2} buffers established in proteomic analyses using FAIMS.

The variants separated by any method need to be identified. Assigning the FAIMS spectral peaks using standards can be realistic in targeted applications, but global analyses and discovery of novel proteoforms call for peptide identification by MS/MS. Large multiply charged modified peptides are optimally sequenced by electron transfer dissociation (ETD) that severs the backbone, but not weaker PTM links, producing informative c- and z-fragments,\textsuperscript{22} which enable detailed analyses of complete histone H3 tails.\textsuperscript{24,25}

Previous IMS and FAIMS studies focused on the localization variants of peptides in the bottom-up range, with the largest separated species being methylated or acetylated H3 and H4 mini-tails with ~20–25 residues.\textsuperscript{26,27} The ETD characterization of FAIMS-resolved variants was limited to even smaller peptides (~1.5 kDa).\textsuperscript{28} Such segments cannot reveal the full PTM information encoded in histones. The position of a single PTM would intuitively affect the 3-D geometries of larger peptides less, diminishing their resolution by IMS methods. Here we demonstrate widely applicable FAIMS separation of complete isomeric 5.4 kDa histone tails that contain nearly all histone PTMs—a major capability for proteomics and epigenetic research.

\section*{EXPERIMENTAL SECTION}

\textbf{Materials and Reagents.} Reagents were from Sigma, except 2-Cl-(Trt)-Cl resin (GenScript), 9-fluorenylmethyl carbazate (TCI chemicals), Fmoc-Lys(Me, Boc)-OH (AnaSpec), Fmoc-Lys(Me)\textsubscript{3}-OH (GL Biochem), Fmoc-Lys(Me, Boc)-OH (AnaSpec), Fmoc-Lys(Ac)-OH, Fmoc-Ser(PO(OBzl)OH)-OH, Fmoc-Thr(PO(OBzl)OH)-OH, and Fmoc-Tyr(PO(OBzl)OH)-OH (Merck) and VA-042 mini-trap column (GE-healthcare) and dried under vacuum.

\textbf{Table 1. PTM Localizations in H3 Tails.} These standards were obtained by producing two peptides of 25 residues each via Fmoc SPPS and joining them via native chemical ligation (NCL).\textsuperscript{28} The Fmoc hydrazide beads were prepared in-house with 0.49 mmol/g loading. The part I (ARTKQTARKSTGAPARTK43QT6ARK9S10TGGK14APRK18QLATK23AARK27-K36APATGGVK38KPHRY39RPGTVALRE) was produced on Wang resin with Glu residue attached. The amounts of 10 $\mu$mol were used in Intavis ResPep SPPS machine using the program specified in the SI. The peptides were detached from resin by the cleavage solution (80% TFA, 2.5% H\textsubscript{2}O, 2.5% DTT, 15% triisopropylsiline) and precipitated by adding 100 volumes of ice-cold diethyl ether. The resulting part I contained a hydrazine group at its C-terminus, while the part II started with Cys. Both parts were purified on the Waters XSelect CSH C18 column using Akta Purifier 900 (HPLC conditions in the SI).

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
PTM & positions \\
\hline
Me & K4, K9, K23 \\
me3 & K4, K9, K23, K27, K36 \\
Ac & K9, K14, K18, K27, K36 \\
P & T3, T6, S10, S28, Y41 \\
\hline
\end{tabular}
\end{table}

In the NCL step, the parts I and II were resuspended at 10 mg in 100 $\mu$L of 6 M GndCl (pH 2.0), with 3.4 mg MPA added for part II. The hydrazine in part I was converted to an azide by oxidation for 20 min at $-15$ °C on an acetone–ice bath by adding 10 $\mu$L of 0.5 M NaNO\textsubscript{2}. The two solutions were then combined and pH adjusted to 7.0 using 6 M NaOH. After 4–6 h, the reaction was stopped by adding 20 $\mu$L of 0.5 M TCEP (pH 7.0) for 20 min. The product was purified on a PD-mini trap column (GE-healthcare) and dried under vacuum.

\textbf{FAIMS Operation.} Histone tails standards were analyzed employing our FAIMS device with 1.9 mm gap width, coupled to the Thermo LTQ XL ion trap with electrospray ionization (ESI) source and electron transfer dissociation (ETD) stage via a slit aperture/ion funnel interface.\textsuperscript{5,16} The waveform amplitude (dispersion voltage) was 4 kV, the CV scan rate was 0.33 V/min, and the flow of carrier gas (65% He/35% N\textsubscript{2}) was 2 L/min. Samples were delivered for analysis using 20 $\mu$L i.d. glass emitters prepared by chemical etching.\textsuperscript{30} Raw data was processed using MSconvert (Proteowizard package)\textsuperscript{31} and a custom R script.

\textbf{FAIMS/ETD Analyses.} The ETD data were recorded by successively setting the CVs to apexes of major peaks in FAIMS and fragmenting transmitted ions. The initial FAIMS/ETD work has involved 2+ and 3+ precursors, which produced the 1+ fragments only, and used single reaction time ($t_{ETD}$) of 120 ms.\textsuperscript{7,8,32} Present larger peptides with much higher z yield multiply charged fragments with z up to 5. With the limited resolving power of ion trap, rich ETD MS/MS spectra of large peptides include many mass interferences. Longer $t_{ETD}$ permit more consecutive ETD steps, yielding different fragments with overall lower z. To disentangle those interferences and raise the informative content of spectra, all species were fragmented at $t_{ETD}$ of 15, 30, and 150 ms. For utmost specificity, we performed MS/MS in the zoom and ultrazoom modes. As these are less sensitive than the normal mode, each ETD spectrum was acquired for 10 min. The final ETD MS/MS spectra are combinations of regions extracted from the spectra with different $t_{ETD}$.

\section*{RESULTS AND DISCUSSION}

\textbf{Variant Separations.} To evaluate FAIMS separation we produced 18 proteoforms of human H3.1 tail (residues 2–51, monoisotopic mass of 5350 Da) with PTMs in biologically relevant positions (Table 1) by combination of Fmoc solid phase peptide synthesis and native chemical ligation using novel hydrazide chemistry.\textsuperscript{28}

The modified variants (5 $\mu$L solutions in water with 0.1% formic acid) were initially infused individually, and then their FAIMS spectra (plots of signal vs CV) were overlaid. To correct for uncontrolled fluctuations of operating parameters (primarily the ambient pressure and temperature), the CV axes...
of individual FAIMS spectra were linearly scaled using an internal calibrant - the unmodified tail (1 μM). For transferability across devices with unequal gap widths, the CV scale is expressed as the compensation field ($E_C$).

Figure 1 illustrates the experimental concept. Trimethylated tails form charge states $z = 8$–$12$ (Figure 1A), each exhibiting a unique FAIMS spectrum (Figure 1B,C). Overall, the 10+ peaks were most intense and permitted maximum discrimination for monomethylated, acetylated, and phosphorylated variants (Figure 2A–C, S1). Only five variants (K4me, K9me, K23me3, K9ac, K27ac) of the 18 total exhibited secondary features at >20% of the base peak heights at this charge. We baseline-resolved all variants for me, two groups for me3 (K4/K36, K9/K23/K27), three groups for ac (K9, K14/K27, K18/K36), and four for p (S10, S28/Y41, T3, T6). Since binary variant mixtures can be fully characterized using ETD,7,8 the only case not completely solved by FAIMS/ETD here at $z = 10$ is the ternary K9me3/K23me3/K27me3 mixture. We confirmed the robustness and reproducibility of the experimental procedure by ensuring that the spectra for binary and ternary variant mixtures matched the superimposed spectra for individual species (Figures 2D–F, S2, and S3).

**Orthogonality of Separation across Charge States.** The separations in different charge states are mutually independent (orthogonal),26,27,33 hence, the variants overlapping at one $z$ are often resolved at another. For example, the overlaps at $z = 10$ are fully resolved at other charge states: K4me3/K36me3 at $z = 8, 11$, and 12; K9me3/K23me3 and K9me3/K27me3 at $z = 11$; K14ac/K27ac at $z = 11, 12$; K18ac/K36ac and S28p/Y41p at $z = 9, 12$ (Figure S1). The only partly unresolved binary isomeric pair was K23me3/K27me3, with the best (half-maximum) resolution at $z = 11$ (Figure 1C). As the result, the me3 variants were best separated at $z = 11$ (Figure 1C).

To assess said orthogonality systematically, we tabulated the $E_C$ values for tallest peaks of all variants with trimethylation, acetylation, and phosphorylation PTMs in every charge state (Table S1) and computed their pairwise correlations (Table S2). For example, the plot of $E_C$ for five me3 variants at $z = 10$ versus those at $z = 11$ reveals a linear correlation of $R^2 = 0.01$ (Figure S4a), and the average $R^2$ for all 30 charge state pairs (10 pairs for every PTM) is 0.25, suggesting no significant correlation. This conclusion is confirmed by rigorous statistical tests (see SI) and clearly seen in Figure S4b.

**Separation of Acetylated and Trimethylated Peptides.** Attachment of me3 (42.047 Da) and ac (42.011 Da) to lysine residues is mutually exclusive and has different functional outcomes. For example, (i) H3K9me3 or H3K27me3 are repressing marks, whereas H3K9ac or H3K27ac activate transcription,34,35 (ii) H3K36me3 occurs predominantly in the coding regions of genes and participates in transcriptional elongation, whereas H3K36ac is localized in the promoters and facilitates transcription initiation.36 Baseline MS resolution of

![Figure 1. ESI MS spectrum for a trimethylated histone tail (A) and FAIMS spectra for $z = 10$ (B) and 11 (C).](image)

![Figure 2. FAIMS spectra for monomethylated, acetylated, and phosphorylated histone tails at $z = 10$ (A–C), same for the spectra for selected tertiary mixtures (D–F). The spectra for individual variants are normalized in (A)–(C), vertically scaled to match the features for the mixture in (D)–(F). The sums of spectra for all variants are also shown in (D)–(F).](image)
the tails with me3 and ac requires resolving power of $\sim 3 \times 10^5$ (and $\sim 8 \times 10^5$ for whole H3) in theory, and is commonly precluded by spectral congestion in practice. However, me3 and ac may well favor distinct 3-D geometries, enabling FAIMS separation. Of our three tail pairs with me3 or ac on same residue, two (PTM on K27 or K36) were resolved baseline and one (PTM on K9) nearly so (Figures 3 and S5).

**Confirmation by ETD.** The ETD data unequivocally prove the variant resolution seen at the precursor level. For instance, all fragments observed at high S/N ratio corroborate the separation of T3p/S10p/Y41p mixture (Figure 4).

Parallel results for other representative binary and tertiary mixtures (K4me1/K9me1/K23me1; K4me3/K27me3; K9ac/K14ac/K36ac; K9ac/K18ac/K27ac; T6p/S10p/S28p) show the general utility of procedure (Figure S6). For utmost specificity, we performed MS/MS in the zoom mode. While the ultrazoom mode delivers yet greater MS resolution, a limited scan range of 100 m/z units and low signal demand long acquisitions; we obtained one spectrum to prove feasibility (Figure S7).

**CONCLUSIONS**

We have demonstrated near-universal complete separation of PTM localization variants (involving methylation, acetylation, and phosphorylation at distinct residues) for peptides above 5 kDa, where PTMs amount to 0.25–1.5% of the peptide mass. The maximum $E_c$ difference between principal peaks is $\sim 10–15\%$ with all PTMs (e.g., for K4me/K9me at 10+, K4me/36me3 at 9+, K9/K14ac at 11+, or T6p/S10p at 10+), same as or only moderately below the values for methylated or acetylated mini-tails (2–3 kDa). Hence, the separation space narrows quite gradually (if at all) for longer peptides, encouraging us to extend this FAIMS approach to yet larger peptides and intact proteins. In fact, the 5.4 kDa histone tails explored here approach the size of small proteins (e.g., insulin at 5.8 kDa). In principle, the orthogonality of separations between charge states expands the overall FAIMS peak capacity for polypeptides in proportion to the number of charge states $(n)$ output by the ion source. With regular ESI, $n$ scales roughly as the (peptide mass)$^{1/2}$ or faster, increasing from 2–3 for typical tryptic peptides ($\sim 1–1.5$ kDa) to at least 5 for present tails to $\sim 10–15$ for smaller proteins ($\sim 10–20$ kDa). That number can be augmented by generating higher $z$ via supercharging and lower $z$ via proton stripping. This trend largely underlies the success of this work and will facilitate the separation of localization variants for proteins. The main problem here is the ubiquitous conformational multiplicity (at charge states besides 10+ here), which would be more important for real samples with unequal variant abundances. We will seek to reduce that multiplicity by optimizing the ESI and FAIMS conditions.

Biological histone samples exhibit much higher complexity and dynamic range than the mixtures of standards in this study. The true variant distribution is convoluted in the analysis with nonuniformities of ionization efficiency, transmission through the FAIMS gap, and ETD yield in total and for specific

![Figure 3. Normalized FAIMS spectra for the tails trimethylated or acetylated at K27 and their mixture.](image)

![Figure 4. FAIMS/ETD analysis of tertiary mixture of phosphorylated tails: FAIMS spectrum (left) and selected ETD spectral windows (from the data at various $t_{ETD}$) comprising the informative fragments with schematic sequences on top (right).](image)
FAIMS scan by stepping through discrete CVs. This approach gases (He or H\textsubscript{2}) with fast Brownian motion than all air instruments in the Orbitrap or FTICR implementations. mass accuracy provided by Fourier-Transform (FT) MS histone samples demand much higher MS peak capacity and other methods are eminently suitable. Analyses of real extended accumulation and multistep fragmentation by ETD peak width is \(\sim 10^{-5}\) Torr, but not to Orbitrap MS that needs ultrahigh vacuum (\(10^{-10}\) Torr). We are currently re-engineering the FAIMS stages and Orbitrap vacuum system to enable addition of high-resolution FAIMS using helium-rich buffers.

Comprehensive analyses also require an LC step in front to partition histone tails into fractions with same number of acetyl and methyl PTMs. The challenge for online LC/FAIMS integration is time scale mismatch: the typical chromatographic peak width is \(\sim 10\) s, while a full high-resolution FAIMS scan takes \(\sim 10\) min. One established solution is replacing a gradual FAIMS scan by stepping through discrete CVs. This approach best fits targeted applications, for example, distinguishing several peptide variants with known CVs such as present histone tails. In external convolution, the same LC gradient is rapidly stepped through several values (with dwell times of a few s) in a mode resembling multiple reaction monitoring while LC peak elutes. An alternative is off-line LC fractionation followed by direct infusion. We foresee the integration of high-definition FAIMS with 2-D LC and FTMS to allow unprecedentedly comprehensive characterization of proteomes and proteoforms.

**ASSOCIATED CONTENT**

1. Supporting Information
   The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.7b00379.

   Description of procedures for histone tail synthesis, additional FAIMS spectra for further variants and charge states, and statistical methods for evaluation of the orthogonality between FAIMS and MS dimensions (PDF).

**AUTHOR INFORMATION**

Corresponding Author
*E-mail: jenseno@bmb.sdu.dk.

ORCID
Vladimir Gorshkov: 0000-0003-0170-5785
Ole N. Jensen: 0000-0003-1862-8528

Author Contributions
The two senior authors of this work.

**Funding**
This research was funded by a grant from the VILLUM Foundation to the VILLUM Center for Bioanalytical Sciences at University of Southern Denmark (O.N.J.), a Lundbeck Foundation Postdoctoral Fellowship (P.S.), and NIH COBRE (P30 GM110761) and NSF CAREER (CHE-1552640) grants (A.S.). A.S. also holds a faculty appointment at the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Russia.

**Notes**
The authors declare the following competing financial interest(s): Alexandre Shvartsburg has interest in Heartland MS that makes FAIMS and ion funnel systems utilized in this work.

**ACKNOWLEDGMENTS**
We thank Drs. Y. C. Huang and Y. M. Li for guiding our synthetic work.

**REFERENCES**

(13) Saba, J.; Bonneil, E.; Pomiès, C.; Eng, K.; Thibault, P. J. Proteome Res. 2009, 8 (7), 3355.