



# HHS Public Access

Author manuscript

*Anal Chem.* Author manuscript; available in PMC 2022 March 30.

Published in final edited form as:

*Anal Chem.* 2021 March 30; 93(12): 5151–5160. doi:10.1021/acs.analchem.0c04975.

## Separation and Characterization of Endogenous Nucleosomes by Native Capillary Zone Electrophoresis–Top-Down Mass Spectrometry

**Kevin Jooß,**

Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, and the Proteomics Center of Excellence, Northwestern University, Evanston, Illinois 60208, United States;

**Luis F. Schachner,**

Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, and the Proteomics Center of Excellence, Northwestern University, Evanston, Illinois 60208, United States

**Rachel Watson,**

EpiCypher Incorporated, Durham, North Carolina 27709, United States

**Zachary B. Gillespie,**

EpiCypher Incorporated, Durham, North Carolina 27709, United States

**Sarah A. Howard,**

EpiCypher Incorporated, Durham, North Carolina 27709, United States

**Marcus A. Cheek,**

EpiCypher Incorporated, Durham, North Carolina 27709, United States

**Matthew J. Meiners,**

EpiCypher Incorporated, Durham, North Carolina 27709, United States

**Amin Sobh,**

Division of Hematology/Oncology, University of Florida Health Cancer Center, Gainesville, Florida 32610, United States

---

**Corresponding Author: Neil L. Kelleher** – Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, and the Proteomics Center of Excellence, Northwestern University, Evanston, Illinois 60208, United States; n-kelleher@northwestern.edu.

### Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.analchem.0c04975>.

Scheme of the developed workflow including Nuc structure, description of optimization of injection volume, results of the separation voltage experiment, comparison of fragmentation between nCZE and nDI, results of the pseudo-MS3 experiment on H3K27me3, regression and LOD of the CZE–TDMS method, analysis of TKO-derived Nucs on the MS1 level, observed and theoretical isotopic distributions of H3 proteoforms, immunoblot analysis of NTKO/TKO cells, relative quantification of H3K27/K36 and H4K16 proteoforms, list of CZE and MS parameters, results of intra- and interday reproducibility experiments, relative quantification of NTKO/TKO-derived Nucs, protocol for the preparation of NTKO/TKO-derived Nucs, and description of optimization of CZE methods (PDF)

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.analchem.0c04975>

The authors declare the following competing financial interest(s): N.L.K. serves as a consultant to Thermo Fisher Scientific and engages in entrepreneurship in the area of top-down proteomics. EpiCypher is a commercial developer and supplier of the semisynthetic modified nucleosomes (dNucs) used in this study.

**Jonathan D. Licht,**

Division of Hematology/Oncology, University of Florida Health Cancer Center, Gainesville, Florida 32610, United States

**Michael-Christopher Keogh,**

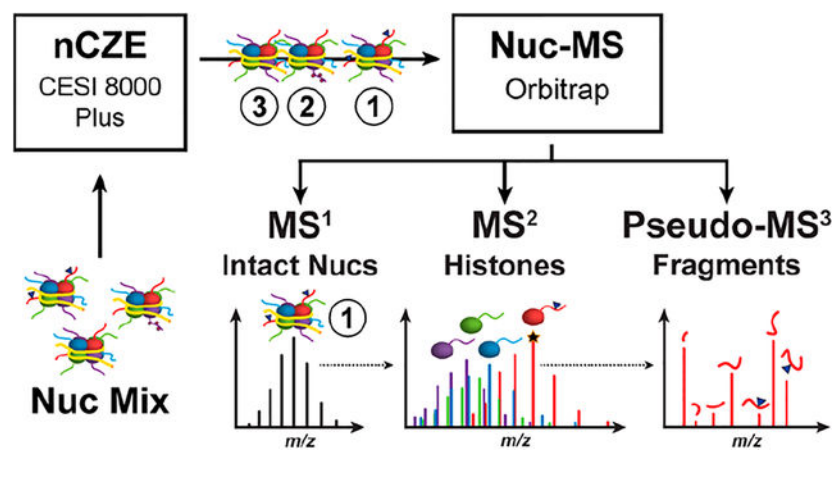
EpiCypher Incorporated, Durham, North Carolina 27709, United States

**Neil L. Kelleher**

Departments of Chemistry and Molecular Biosciences, the Chemistry of Life Processes Institute, and the Proteomics Center of Excellence, Northwestern University, Evanston, Illinois 60208, United States;

**Abstract**

We report a novel platform [native capillary zone electrophoresis–top-down mass spectrometry (nCZE–TDMS)] for the separation and characterization of whole nucleosomes, their histone subunits, and post-translational modifications (PTMs). As the repeating unit of chromatin, mononucleosomes (Nucs) are an ~200 kDa complex of DNA and histone proteins involved in the regulation of key cellular processes central to human health and disease. Unraveling the covalent modification landscape of histones and their defined stoichiometries within Nucs helps to explain epigenetic regulatory mechanisms. In nCZE–TDMS, online Nuc separation is followed by a three-tier tandem MS approach that measures the intact mass of Nucs, ejects and detects the constituent histones, and fragments to sequence the histone. The new platform was optimized with synthetic Nucs to significantly reduce both sample requirements and cost compared to direct infusion. Limits of detection were in the low-attomole range, with linearity of over ~3 orders of magnitude. The nCZE–TDMS platform was applied to endogenous Nucs from two cell lines distinguished by overexpression or knockout of histone methyltransferase NSD2/MMSET, where analysis of constituent histones revealed changes in histone abundances over the course of the CZE separation. We are confident the nCZE–TDMS platform will help advance nucleosome-level research in the fields of chromatin and epigenetics.

**Graphical Abstract**

## 1. INTRODUCTION

Increasing evidence indicates that disruption in the cellular epigenetic machinery may be a key initiator in several types of cancer and other disorders.<sup>1</sup> The mononucleosome (Nuc), the smallest repeating unit of chromatin, resides at the center of epigenetic regulation. It comprises ~147 bp of DNA wrapped around a protein octamer with two copies of each core histone: H2A, H2B, H3, and H4 (Figure S1a).<sup>2</sup> Nucleosome structure (and presumably function) can be altered by covalent modification of the DNA or histone proteins (the latter is termed post-translational modification; PTM) or by replacement of the canonical histones with variants, which have diverse biophysical properties.<sup>3,4</sup> Nuc modification dynamics can regulate key cellular processes central to human health and disease (e.g., transcription, DNA repair, and DNA replication).<sup>5</sup> For this reason, the notion of a “histone code” has gained traction over the last 20 years, given the discovery of effector proteins that simultaneously recognize and bind one or more PTMs on nucleosome assemblies.<sup>6</sup> Unraveling the full landscape of modifications on whole Nucs is thus key to elucidate the mechanisms that exert epigenetic control over a genomic location.

Chromatin and epigenetics researchers generally use histone peptides as Nuc proxies when studying PTM writers, readers, and erasers, obscuring the combinatorial variations that co-occur at the Nuc level and mediate genome transactions.<sup>7,8</sup> The high cost of novel PTM-defined Nuc reagents and lack of a direct compositional readout have slowed Nuc-level biology.<sup>9,10</sup> To address this technology gap, we recently described “Nuc-MS”, a novel method based on top-down mass spectrometry (TDMS) operated in native mode that can decode the protein composition of whole nucleosomes from synthetic and endogenous sources.<sup>11</sup> A major feature of TDMS is the isolation and fragmentation of individual histones, such that PTMs and variant substitutions can be directly assigned as a single nucleosome proteoform.<sup>12</sup> The top-down approach has been extended to whole complexes by our group and others<sup>13–17</sup> and provides a higher-level view of the modification landscape compared to the inference required in bottom-up approaches.<sup>18</sup>

In recent years, native mass spectrometry (nMS) has greatly contributed to structural biology research.<sup>19,20</sup> Here, volatile aqueous buffer systems (e.g., ammonium acetate) are employed to preserve near-native features of proteins and their complexes during electrospray ionization.<sup>21,22</sup> The approach allows one to evaluate the composition of biomolecular structures leading to a better understanding of their biological function and significance.<sup>15,22</sup> In particular, coupling native ionization with TDMS provides insights to subunit stoichiometry, stability, topology, dynamics, and the affinities of protein complexes.<sup>23</sup> The lower extent of protonation during native electrospray ionization (ESI) gives access to a wider range of the MS instrument’s mass-to-charge ratio ( $m/z$ ) region than under denatured ESI, aiding the resolution of different protein signals.<sup>24</sup> However, ion suppression and signal superposition cannot be completely avoided for complex samples. Thus, upfront separation under native conditions is beneficial and highly desired.

Different analytical chromatographic<sup>25–28</sup> and electromigrative<sup>29–31</sup> tools theoretically allow native separation. Generally, there are three major requirements to achieve efficient native separation prior to MS detection: (i) the mobile phase or background electrolyte (BGE) must

maintain the nature of the nondenatured protein or complex, (ii) the composition of the mobile phase/BGE must be compatible with electrospray ionization, and (iii) the separation performance and resolution need to be sufficient. Unfortunately, many current techniques suffer in at least one of these aspects. In contrast, capillary zone electrophoresis (CZE) provides high resolution even under native conditions.<sup>23</sup> The separation mechanism of CZE is based on the electrophoretic mobilities of ions in the liquid phase, which are dependent on the charge-to-size ratio of the analytes.<sup>32</sup> CZE has been shown to be well-suited for the separation of intact proteins with only small structural differences such as deamidation events.<sup>33</sup> Another important aspect of CZE is the low sample requirement, typically a few nanoliters per injection.<sup>34</sup> Several CZE–MS interfaces have recently been developed, including sheath liquid, nanoflow sheath liquid, and sheathless porous tip.<sup>35</sup> However, native (n)CZE–TDMS for protein analysis remains a largely unexplored area, with few studies to date.<sup>31,36,37</sup>

The presence of PTMs and histone variants is generally asserted through immunoblots, affinity-based enrichment (e.g., ChIP-seq), and digestion-based proteomics.<sup>38</sup> Here, we report the first native-mode MS platform for the separation and characterization of intact Nucs, their histone subunits, and PTM profiles in a single experiment. The nCZE–TDMS method was initially optimized using synthetic Nucs to enable MS<sup>1</sup>, MS<sup>2</sup>, and pseudo-MS<sup>3</sup> data collection (Figure S1b), and later expanded for the separation of endogenous nucleosomes (endoNucs) using methyltransferase-modulation cell lines as a model system. An important feature of the platform is that Nuc samples in complex buffers were directly injected into the CZE system without any sample preparation, only consuming minor amounts of sample material (<\$0.01 per injection). As a result, sample requirements—and thus cost—are significantly reduced compared to direct infusion, enabling the online separation of Nuc subpopulations based on their charge-to-size ratios. This new data type, coupled with Nuc separation by CZE, provides a wide-lens, semiquantitative view of a cell's epigenetic landscape.

## 2. MATERIALS AND METHODS

### 2.1. Materials.

Water (Optima LC–MS grade), acetic acid (HAc) glacial (Optima LC–MS grade), and hydrochloric acid (HCl) (technical grade) were from Thermo Fisher Scientific (Chicago, IL, U.S.A.). The 7.5 M ammonium acetate (AmAc) stock solution (molecular biology grade) was from Sigma-Aldrich (St. Louis, MO, U.S.A.). Molecular weight cutoff (MWCO) spin filters were from Thermo Fisher Scientific. All solutions used in the CESI 8000 Plus system were passed through 0.2  $\mu\text{m}$  pore size Nalgene rapid-flow sterile disposable filters (Thermo Fisher Scientific).

### 2.2. Synthetic Nucleosome Samples.

All semisynthetic nucleosomes were from EpiCypher (Durham, NC, U.S.A.): unmodified recombinant (r)Nuc (16-0006); H3K27me3 (16-0317), H2BK120ub (16-0370), and H3K4,9,14,18ac (16-0336) designer (d)Nucs; H3.3G34V (16-0347) oncoNuc. For CZE experiments, Nuc samples were kept in their original storage buffer (10 mM Tris–HCL, 1

mM EDTA, 25 mM NaCl, 2 mM DTT, and 20% glycerol) and only mixed and/or diluted in BGE as needed.

For direct infusion experiments via the CESI 8000 Plus device, Nuc samples were buffer-exchanged using Amicon Ultra-0.5 mL centrifugal filters (30 kDa MWCO). In brief, the filter device was equilibrated with 500  $\mu\text{L}$  of 50 mM AmAc and spun for 5 min at 12 000*g*. The Nuc sample was then loaded up to a total volume of 500  $\mu\text{L}$  and spun for 5 min at 12 000*g* or until concentrated to  $\sim$ 50  $\mu\text{L}$ . Eight consecutive washing steps were performed by adding AmAc (40 mM) to 500  $\mu\text{L}$  of mixture and spinning at 12 000*g* for 5 min. The remaining sample volume was transferred to CESI nanovials for analysis.

### 2.3. Endogenous Nucleosome Samples.

TKO and NTKO NSD2 low- and high-expressing cells were prepared as described,<sup>39</sup> with endogenous Nucs extracted and prepared for MS analysis as previously (Method S1).<sup>11</sup> The only deviating step from our previous protocol was a final buffer exchange to 40 mM AmAc for nCZE–TDMS analysis, resulting in a concentration of about 10 mg/mL of endoNuc material.

### 2.4. Capillary Electrophoresis.

A CESI-8000 Plus instrument from SCIEX (Redwood City, CA, U.S.A.) was used, and separation was performed using commercial Neutral OptiMS capillary cartridges (30  $\mu\text{m}$  inner diameter, 90 cm length) containing an integrated sheathless etched porous nanospray tip. The inner wall of the separation capillaries is coated with a binary layer: (i) a hydrophobic coating and (ii) a hydrophobic polyacrylamide surface. In this way, protein adsorption is prevented, and electro-osmotic flow is largely suppressed.

For initial conditioning, the capillary was rinsed (100 psi) with 0.1 M HCl (5 min), BGE (10 min), and water (30 min); the conductive line (CL) was rinsed with water (5 min). Subsequently, both the separation and conductive lines were rinsed (100 psi) with BGE (5 and 3 min, respectively) followed by applying a high voltage of +15 kV and 5 psi supporting flow for 60 min. At the end of this step, the voltage was ramped down over 5 min. Each morning, the capillary was rinsed with fresh BGE and high voltage (+15 kV, 5 psi) was applied for 30 min. For long-term storage, capillaries were rinsed with water and kept at 4  $^{\circ}\text{C}$ .

For Nuc analysis, the capillary was first rinsed with 0.1 M HCl (100 psi, 3 min), followed by BGE (5 min). The CL was filled with 3% HAc (100 psi, 3 min). Hydrodynamic injection was performed at 2.5 psi for 30 s, corresponding to  $\sim$ 11 nL, followed by a water dipping step and the injection of a small plug of BGE (2.5 psi, 10 s). Three different CE methods were developed during this work: (i) standard, (ii) high-resolution, and (iii) high-throughput. A detailed description of each individual method is provided in Table S1. It is advised to replace the BGE every 15 measurements to ensure repeatability of migration times and overall system stability. In addition, clogging of capillaries is an issue that can be potentially encountered with the sheathless system. Procedures to unclog capillaries have been described.<sup>40</sup>

## 2.5. Mass Spectrometer.

The CESI-8000 Plus instrument was hyphenated with a custom Thermo Fisher Q Exactive Orbitrap HF MS with extended mass range<sup>14</sup> (QE-EMR) and a commercial Thermo Fisher Q Exactive Orbitrap MS with ultrahigh mass range (UHMR). Important parameters of the applied tune files for MS<sup>1</sup>, MS<sup>2</sup>, and pseudo-MS<sup>3</sup> experiments for both instruments are provided in Table S2. A Nanospray Flex ion source was changed using an OptiMS Thermo MS adapter from SCIEX for hyphenation of the CESI and orbitrap (OT) instruments. The sprayer tip was positioned 2.5 mm in front of the OT orifice, and an ESI voltage between +1.6 and +1.9 kV was applied during separation. The inlet capillary of the MS instruments was heated to 330 °C (QE-EMR) and 300 °C (UHMR), respectively.

## 2.6. Data Analysis.

Spectra were analyzed manually using Thermo Xcalibur 4.0 Qual browser (Thermo Fisher Scientific, Inc.). Figures were created using Adobe Illustrator CC 2015.3. Signal-to-noise ratios (S/N) were calculated as follows:  $S/N = (NL - B)/(N - B)$ , where NL is the signal intensity, B is the baseline intensity, and N is the noise intensity. The sum of all observed charge states for each parameter (NL, N, and B) was used for S/N calculations. In this way, an average S/N for the protein charge state distribution was determined. Least-square regression was performed, and the residual errors were weighted by  $1/x$ , where x is the respective concentration level. The limit of detection (LOD, S/N = 3) was estimated by extrapolation of the lowest measured concentration level. Deconvolution of histone data (MS<sup>2</sup>) was performed using Unidec 3.2.0.<sup>41</sup> Parameters are as follows: data processing, range 500–2500 Th, bin every, 0; charge range, 5–15; mass range, 10–20k Da; sample mass every (Da), 0.05. Peak area values were calculated by integrating the assigned mass range of proteoforms in the deconvoluted mass spectrum. Histone proteoforms were assigned by intact mass and isotopic fit (in-house database). Statistical significance was evaluated using two-sided, two-sample *t* tests, and the resulting *p*-values were compared to Bonferroni-corrected  $\alpha$ -values based on  $\alpha = 0.05\%$ . Mass lists of peptide fragments were created using Xtract (Thermo Fisher Scientific). TDValidator<sup>42</sup> (max ppm tolerance, 16 ppm; sub-ppm tolerance, 7.5 ppm; cluster tolerance, 0.35; charge range, 1–15; minimum score, 0.3; S/N cutoff, 3; Mercury7 limit, 0.0001; minimum size, 2) was used to assign recorded fragment ions to the primary sequence of the histone subunits. Electrophoretic resolution was calculated based on the full width at half-maximum of the respective CZE peaks.

## 3. RESULTS AND DISCUSSION

We first evaluated the performance of the nCZE–TDMS platform with a recombinant Nuc containing the histone H3 trimethylated at lysine 27 (H3K27me3,  $c = 1 \mu\text{M}$ , Figure 1a–c). Importantly, the sample was injected directly in the manufacturer's storage and shipping buffer (10 mM Tris–HCL, 1 mM EDTA, 25 mM NaCl, 2 mM DTT, 20% glycerol: EpiCypher) which was designed to minimize Nuc loss and degradation. The data show that both salts and other buffer components are clearly separated from the Nuc signal in the total ion electropherogram (TIE, Figure 1b), obviating the need for additional sample purification such as solvent exchange to AmAc solutions.<sup>43</sup> A general scheme of the separation system is depicted in Figure 1d. The injection volume was subsequently optimized (Figure S2), and

2.5 psi for 30 s (~11 nL) was selected for further experiments. Composed of both basic histones ( $pI > 10$ ) and acidic DNA strand ( $pI \approx 5$ ),<sup>44</sup> the effective  $pI$  and mobility of Nucs are hard to predict. Therefore, the separation voltage was varied between +12 and +18 kV while keeping the remaining method parameters constant to evaluate the migration behavior of Nucs under native conditions (Figure S3). Higher voltages resulted in an increase in migration time, indicating that the net charge of the Nuc complexes is negative under the given conditions. Consequently, in the applied positive polarity mode they would counter-migrate toward the CESI inlet, but the supplemental pressure is sufficient to drive the Nucs toward the electrospray source.

We next tested histone ejection from the intact Nucs using this platform. An example MS<sup>2</sup> spectrum of H3K27me3 Nucs after ion-source ejection ( $c = 1 \mu M$ , Figure 2a) demonstrates detection of the four histone subunits H4, H2B, H2A, and H3K27me3 (Figure 2b). These ejected histones were subsequently isolated and fragmented by higher-energy collisional dissociation (HCD)<sup>14</sup> and identified by TDValidator (e.g., H2A: Figure 2, parts c and d). Of note, the quality of fragmentation from online CZE–TDMS was comparable to that of native direct infusion analysis of the same sample (Figure S4). Another example showing H3K27me3 fragmentation can be found in Figure S5. The native TD approach assigns ejected histones to their related precursor Nuc. Thus, this procedure constitutes an important gain of information relative to denaturing or proteolysis-dependent histone analyses.<sup>45</sup>

The nCZE–TDMS platform shows high intra- and interday reproducibility in migration time, peak area, and intensity using synthetic Nucs, indicating its potential to support the industrial production/analysis of such reagents (Table S3). Moreover, we tested the LOD and quantitation at the MS<sup>1</sup> level, finding that the linear range of detection is between 10 and 500 nM ( $R^2 = 0.996$ ), with a S/N of  $12.3 \pm 1.0$  (10 nM Nuc sample;  $\hat{=} 110$  amol), resulting in an estimated LOD of 2.4 nM ( $\hat{=} 27$  amol) (Figure S6a–c). In a follow-up experiment we showed adequate MS<sup>2</sup> data for ejected histones from unmodified synthetic Nucs at a concentration of 62.5 nM ( $\hat{=} 0.71$  fmol, Figure S6d).

We next sought to investigate the Nuc separation capabilities of the CZE–TDMS platform. For proof of concept, we mixed tetra-acetylated (H3K4,9,14,18a) and ubiquitinated (H2BK120ub) Nucs prepared (500 nM each) as models of highly divergent Nuc species. Tetra-acetylation/charge neutralization of eight lysine residues (i.e., both histone H3 tails) will decrease the overall net charge of the Nuc complex. The mass shift introduced by two ubiquitins is rather large ( $2 \times 8.5$  kDa), but it is challenging to predict its influence on the overall net charge of the Nuc complex (the ~6.8  $pI$  of ubiquitin is close to the  $pH$  of the developed CZE method). The mixture was analyzed using the “standard” CZE settings (Figure 3a), achieving partial separation ( $R = 0.37$ ,  $n = 2$ ) with H2BK120ub detected first. Under different parameters (evaluated considering separation performance; Discussion S1), migration times and peak widths increase noticeably when decreasing the supplemental pressure from 5.0 to 2.2 psi and electrophoretic resolution is significantly improved from  $R = 0.37$  to 1.09 ( $n = 2$  each, Figure 3b), which becomes even more evident by comparing the mass spectra in Figure 3c–f. The influence of superimposed pressure on CZE flow profiles was described in detail previously.<sup>46</sup> The S/N also improved by ~3-fold, likely due to reduced ion suppression from overlapping peaks and a higher number of scans averaged.

In short, the final “high-resolution” CZE method is based on 40 mM AmAc (pH = 6.8) as BGE and a separation voltage of +18 kV with 2.2 psi of supplemental pressure, resulting in a total run time of about 135 min. With potential for widespread deployment of the nCZE–TDMS platform, we also developed a “high-throughput” CZE method with a total run time of ~20 min tailored for the quality control environment during Nuc manufacturing (see Discussion S2).

As a next step, the native CZE–TDMS platform was applied for the analysis of endogenous nucleosomes (endoNucs) derived (see the Materials and Methods section) from an isogenic knockout and overexpression system (respectively, TKO and NTKO cell lines<sup>47,48</sup>) of NSD2/MMSET, a histone methyltransferase implicated in t(4;14)+ multiple myeloma pathogenesis. In TKO cell lines, the rearranged IGH-NSD2 allele is inactivated by homologous recombination, leaving one wild-type copy of the NSD2, which yields low levels of H3K36me2 (NSD2<sup>±</sup>; H3K36me2<sup>lo</sup>). In NTKO cells, the wild-type NSD2 allele is inactivated leaving the IGH-NSD2 allele and high levels of H3K36me2 (IGH-NSD2; H3K36me2<sup>hi</sup>).<sup>47,48</sup>

The complexity of endoNucs cannot be successfully resolved by traditional direct infusion nMS, which yields a low-abundant and indistinct protein signal (data not shown). In the nCZE–TDMS system, endoNucs showed increased migration time relative to synthetic Nucs. This indicates a higher electrophoretic mobility, so the supplemental pressure was increased from 2.2 to 3.0 psi (Figure 4a and Figure S7). The endoNucs migrated over a large time window (~70–80 min), which was divided into 5 min subsections and individually deconvoluted to search for distinct populations.

The raw and deconvoluted spectra in Figure 4b–d reflect a variety of distinct peaks that increase in mass (192 332–228 854 Da) over the length of the CZE run. The mass differences between the detected peaks reveals up to 60 “nucleoforms” that differ by a base pair [NTKO, 620.2 ± 18.5 Da; TKO, 617.5 ± 22.4 Da; ( $m_{GC} = 618.4$  Da,  $m_{AT} = 617.4$  Da)], likely reflecting processive micrococcal nuclease (MNase) digestion products. In addition, a series of non-baseline-resolved satellite peaks were observed with a mass shift of ~300 Da, indicating single-nucleotide differences (Figure 4b). These results represent the first report of base pair resolution and accurate mass measurement of intact Nucs from endogenous sources.

The same samples were analyzed at the histone level (MS<sup>2</sup>,  $n = 3$  per sample), revealing significant proteoform-level differences between NTKO and TKO cells (Figure 5a–d, Table S4). Notably, we detected H3.2 as the predominant H3 subtype in these cells (proteoforms are labeled accordingly; Figure S8), although the presence of minor amounts of H3.1 cannot be excluded due to their closely related intact masses. The average methylation equivalent (ME) of H3.2 was decreased by ~11% in TKO cells, consistent with their reduced expression of NSD2 and lower levels of H3K36 methylation.<sup>49</sup> Interestingly, the H3.2 4× ME proteoform is highly elevated in TKO, as compared to 6× and 7× ME in NTKO, suggesting NSD2 crosstalk with pre-existing PTMs. To provide additional context for our top-down results, we analyzed the histone mark differences between TKO and NTKO cells by Western blot (Figure S9) and bottom-up proteomics<sup>51</sup> (Figure S10). Close examination of



the K27–K36 peptide reveals a complex picture. In TKO cells, K36me2 indeed decreases by ~81%; however, there is a concomitant increase of ~46% in K27me2, ~95% in K27me3, and ~216% in K27ac, influencing the overall average methylation equivalent. For this reason, the two data types are complementary: the bottom-up information provides mark-specific quantitation; top-down provides a picture of the overall proteoform landscape.

Moreover, H4 acetylation (most likely at K16)<sup>11</sup> is 21% elevated in NTKO, suggesting increased areas of active transcription when NSD2 is overexpressed.<sup>50</sup> Bottom-up proteomics was performed to validate this observation showing an increase of H4K16 acetylation in NTKO by 18% (Figure S11), which is in accordance with the top-down results. Furthermore, the relative abundance of all H3.3 proteoforms is increased in NTKO cells, with 3–5× ME readily detected. Finally, the nCZE–TDMS platform directly quantifies H2A and H2B variants and their proteoforms (Table S4), revealing modification stoichiometry differences, with potential functional implications. Specifically, there is a higher degree of acetylation on H2A.1C and H2A.1B/E in NTKO versus TKO cells (17% and 48% increase, respectively), consistent with the H4 observation. In terms of H2B, there is a lower abundance of variants H2B.1-K and H2B.1-C/E in NTKO relative to TKO cells, suggesting unexplored functional roles for H2B variants in epigenetic regulation.

Interestingly, time-resolved data shows that the relative ratio between histones changes significantly over the course of the CZE separation (Figure 5, parts f and g; Tables S5 and S6). In particular, we observed an increase in overall H2B, with a concomitant decrease in H2A and H3, as the Nuc particles increased in size at higher migration times. This is difficult to explain given our understanding of nucleosome structure and histone stoichiometry/exchange therein. However, inconsistent MNase digestion yields Nucs with varying DNA length (Figure 4), and it is possible that noncovalent interactions between histones and nucleotides in various nucleoforms differentially impact efficiency of histone ejection (e.g., H2B is less tightly bound, explaining its elevated relative abundance with increasing DNA length).

At the proteoform level, some distinct changes over the course of the CZE separation were also observed (Tables S7 and S8). For example, an increase in H4-K20met2-NtAc (+14%) and H2A.1-B/E (+16%) but decrease in H4-R3met1-K20met3-NtAc/H4-K20met1-K16Ac-NtAc (–42%), H2A.1 (–34%), and H2A.1-C-Q104met1 (–30%) was detected at higher migration times for NTKO (and analogous changes in the proteoforms profiles for TKO).

#### 4. CONCLUSION

In summary, we have developed the nCZE–TDMS platform to achieve attomole-level Nuc characterization with emphasis on either high resolution or throughput. In this proof-of-concept study, we demonstrate its potential to separate Nucs containing different proteoforms, which has not been shown by other techniques to date. With \$0.01 of material consumed per 20 min run and minimal sample preparation, the system is a promising candidate for quality control of semisynthetic Nucs. We challenged the platform with endoNucs from an isogenic methyltransferase knockout/overexpression cell line system (NTKO/TKO), detecting changes in the H3 methylation profile consistent with NSD2 status

and proteoforms indicating potential crosstalk with other PTMs, including H4K16ac. Considering the CZE separation performance, at the MS<sup>1</sup> level we achieved base pair resolution and detected ~60 distinct “nucleoforms”, associated with differential enzymatic cleavage of DNA by MNase. At the histone level (MS<sup>2</sup> data) changes in proteoform abundances over the course of the nCZE separation were observed for both NTKO and TKO, demonstrating the capabilities to resolve different Nuc subpopulations. We envision extending into the area of endogenous Nucs from cells and immunoprecipitated specimens where limiting amounts of starting material are available. In future studies, we aspire to simplify the complexity of nucleoforms by optimizing the digestion process to get a better overview of proteoforms changes. In conclusion, nCZE–TDMS has the potential to serve as an important tool in the field of epigenetics, native proteomics,<sup>17</sup> and in the quality control of complex synthetic biomolecules.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

This work was supported by the National Institute of General Medical Sciences P41GM108569 for the National Resource for Translational and Developmental Proteomics at Northwestern University and NIH Grants R01AM115739 (Kelleher lab) and R01CA195732 (Licht Lab), RF1AG063903 from the National Institute on Aging (Kelleher lab), P30DA018310 from the National Institute on Drug Abuse (Kelleher lab), a Leukemia and Lymphoma Society Specialized Center of Excellence Grant (Licht Lab), NIH Grants R44GM116584 and R44CA212733 (EpiCypher), and an AACR Myeloma Fellowship (A.S.). Bottom-up (histone) analysis was performed by the Northwestern Proteomics Core Facility, supported by NCI CCSG P30CA060553 awarded to the Robert H. Lurie Comprehensive Cancer Center, instrumentation award (S10OD025194) from NIH Office of Director, and the National Resource for Translational and Developmental Proteomics supported by P41GM108569. L.F.S. is a Gilliam Fellow of the Howard Hughes Medical Institute. Research in this publication is also supported by Thermo Fisher Scientific and a fellowship associated with the Chemistry of Life Processes Predoctoral Training Grant T32GM105538 at Northwestern University. We thank SCIEX for their support including Dr. Fang Wang for the valuable discussions and insightful suggestions throughout this research project.

## ABBREVIATIONS

<b>AmAc</b>	ammonium acetate
<b>BGE</b>	background electrolyte
<b>CL</b>	conductive liquid
<b>CZE</b>	capillary zone electrophoresis
<b>dNuc</b>	designer nucleosome
<b>EIE</b>	extracted ion electropherogram
<b>endoNucs</b>	endogenous nucleosomes
<b>ESI</b>	electrospray ionization
<b>fwhm</b>	full width at half-maximum
<b>HAc</b>	acetic acid

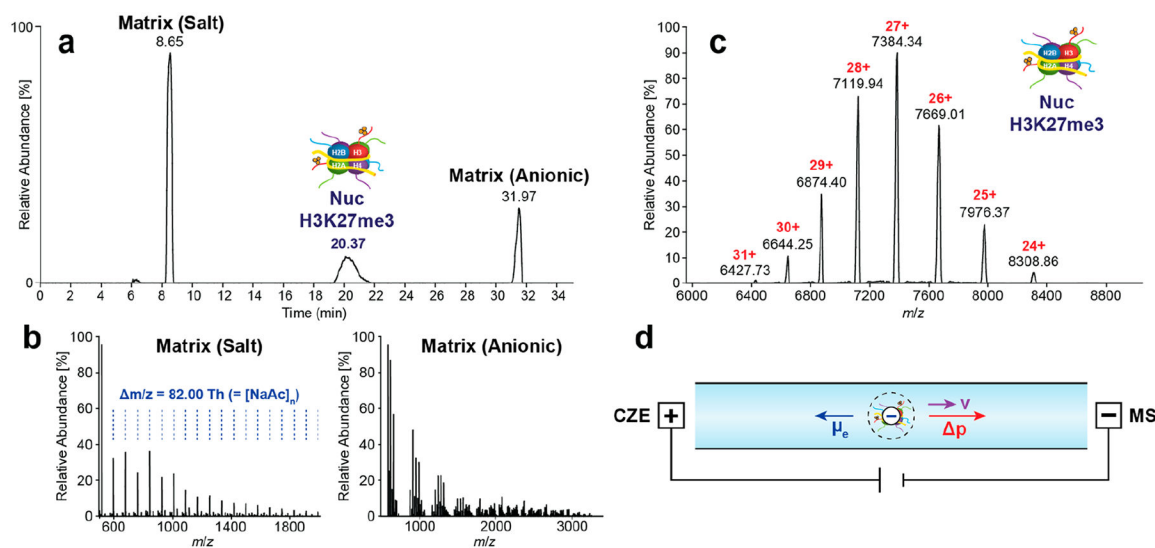
<b>HCD</b>	higher-energy collisional dissociation
<b>HCl</b>	hydrochloric acid
<b>LOD</b>	limit of detection
<b>ME</b>	methylation equivalents
<b>MNase</b>	micrococcal nuclease
<b>MWCO</b>	molecular weight cutoff
<b>nCZE</b>	native capillary zone electrophoresis
<b>nMS</b>	native mass spectrometry
<b>Nuc</b>	mononucleosome
<b>OT</b>	orbitrap
<b>PTM</b>	post-translational modification
<b>QE-EMR</b>	Q Exactive HF MS with extended mass range
<b>rNuc</b>	recombinant nucleosome
<b>SL</b>	separation line
<b>S/N</b>	signal-to-noise ratio
<b>TIE</b>	total ion electropherogram
<b>TDMS</b>	top-down mass spectrometry
<b>UHMR</b>	ultrahigh mass range

## REFERENCES

- (1). Portela A; Esteller M *Nat. Biotechnol* 2010, 28 (10), 1057–1068. [PubMed: 20944598]
- (2). Richmond TJ Hot papers - Crystal structure - Crystal structure of the nucleosome core particle at 2.8 angstrom resolution by K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, T. J. Richmond - *Comments. Scientist* 1999, 13 (23), 15.
- (3). Patel DJ; Wang ZX *Annu. Rev. Biochem* 2013, 82, 81–118. [PubMed: 23642229]
- (4). Zink LM; Hake SB *Curr. Opin. Genet. Dev* 2016, 37, 82–89. [PubMed: 26826795]
- (5). Bannister AJ; Kouzarides T *Cell Res* 2011, 21 (3), 381–395. [PubMed: 21321607]
- (6). Jenuwein T; Allis CD *Science* 2001, 293 (5532), 1074–1080. [PubMed: 11498575]
- (7). Strelow JM; Xiao M; Cavitt RN; Fite NC; Margolis BJ; Park KJ *J. Biomol. Screening* 2016, 21 (8), 786–794.
- (8). Connelly KE; Weaver TM; Alpsy A; Gu BX; Musselman CA; Dykhuizen EC *Nucleic Acids Res* 2019, 47 (5), 2289–2305. [PubMed: 30597065]
- (9). Muller MM; Muir TW *Chem. Rev* 2015, 115 (6), 2296–2349. [PubMed: 25330018]
- (10). Nguyen UTT; Bittova L; Muller MM; Fierz B; David Y; Houck-Loomis B; Feng V; Dann GP; Muir TW *Nat. Methods* 2014, 11 (8), 834–840. [PubMed: 24997861]
- (11). Schachner LF; Jooss K; Morgan MA; Piunti A; Meiners MJ; Kafader JO; Lee AS; Iwanaszko M; Cheek MA; Burg JM; et al. *Nat. Methods* 2021, 18, 303–308.

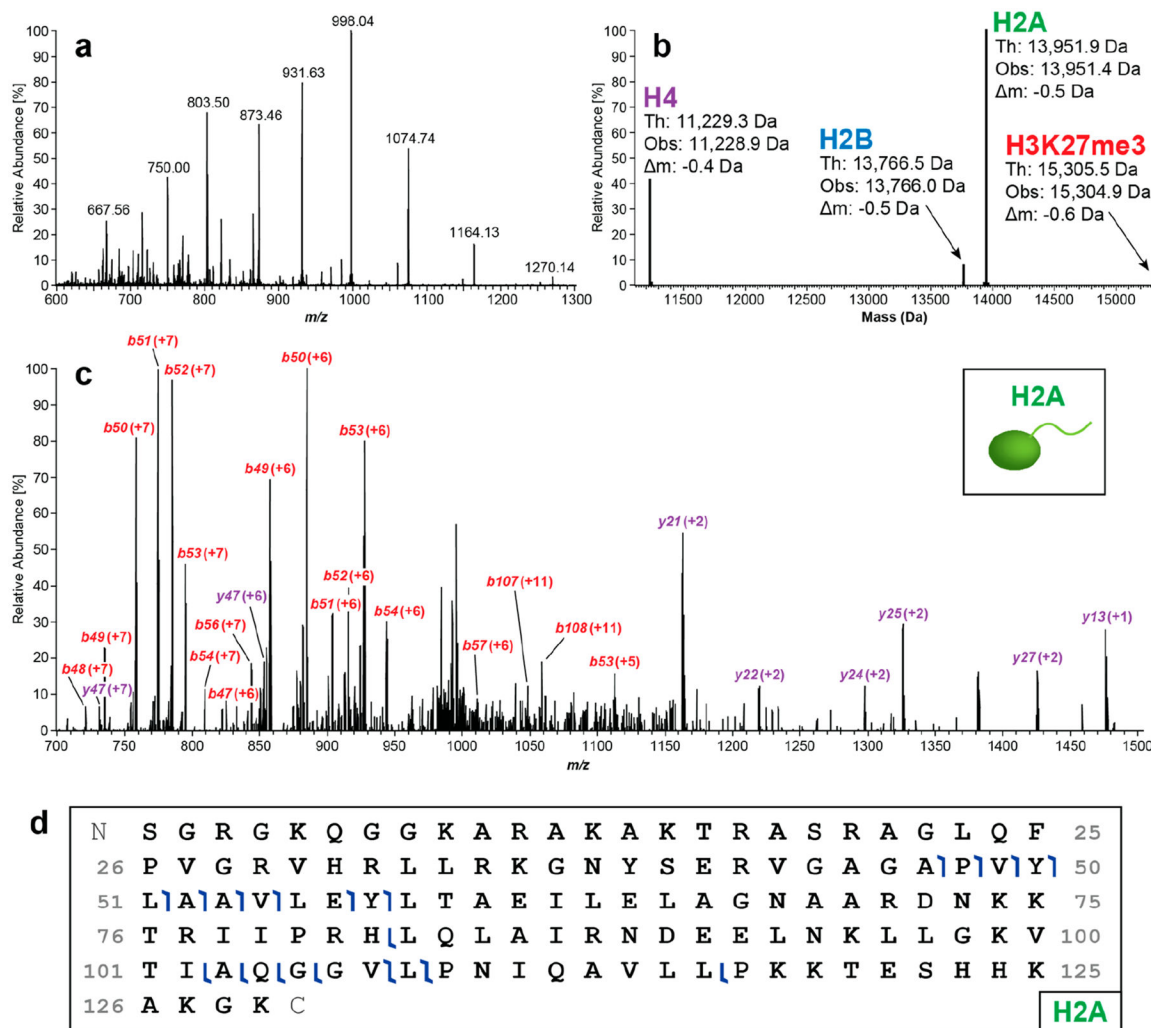
- (12). Smith LM; Kelleher NL *Nat. Methods* 2013, 10 (3), 186–7. [PubMed: 23443629]
- (13). Schachner LF; Ives AN; McGee JP; Melani RD; Kafader JO; Compton PD; Patrie SM; Kelleher NL *J. Am. Soc. Mass Spectrom* 2019, 30 (7), 1190–1198. [PubMed: 30963455]
- (14). Belov ME; Damoc E; Denisov E; Compton PD; Horning S; Makarov AA; Kelleher NL *Anal. Chem* 2013, 85 (23), 11163–11173. [PubMed: 24237199]
- (15). Li HL; Nguyen HH; Ogorzalek Loo RR; Campuzano IDG; Loo JA *Nat. Chem* 2018, 10 (2), 139–148. [PubMed: 29359744]
- (16). Zhang H; Cui WD; Wen JZ; Blankenship RE; Gross ML *J. Am. Soc. Mass Spectrom* 2010, 21 (12), 1966–1968. [PubMed: 20843701]
- (17). Skinner OS; Haverland NA; Fornelli L; Melani RD; Do Vale LHF; Seckler HS; Doubleday PF; Schachner LF; Srzentic K; Kelleher NL; Compton PD *Nat. Chem. Biol* 2018, 14 (1), 36. [PubMed: 29131144]
- (18). Kelleher NL; Lin HY; Valaskovic GA; Aaserud DJ; Fridriksson EK; McLafferty FW *J. Am. Chem. Soc* 1999, 121 (4), 806–812.
- (19). Heck AJ; Van Den Heuvel RH *Mass Spectrom. Rev* 2004, 23 (5), 368–89. [PubMed: 15264235]
- (20). Hilton GR; Benesch JLJR *Soc., Interface* 2012, 9 (70), 801–16. [PubMed: 22319100]
- (21). Fernandez de la Mora *Anal. Chim. Acta* 2000, 406 (1), 93–104.
- (22). Leney AC; Heck AJR *J. Am. Soc. Mass Spectrom* 2017, 28 (1), 5–13.
- (23). Belov AM; Viner R; Santos MR; Horn DM; Bern M; Karger BL; Ivanov AR *J. Am. Soc. Mass Spectrom* 2017, 28 (12), 2614–2634. [PubMed: 28875426]
- (24). Kafader JO; Melani RD; Schachner LF; Ives AN; Patrie SM; Kelleher NL; Compton PD *J. Am. Soc. Mass Spectrom* 2020, 31 (3), 574–581. [PubMed: 31971796]
- (25). Debaene F; Boeuf A; Wagner-Rousset E; Colas O; Ayoub D; Corvaia N; Van Dorsselaer A; Beck A; Cianferani S *Anal. Chem* 2014, 86 (21), 10674–83. [PubMed: 25270580]
- (26). Muneeruddin K; Nazzaro M; Kaltashov IA *Anal. Chem* 2015, 87 (19), 10138–10145. [PubMed: 26360183]
- (27). Gahoual R; Heidenreich AK; Somsen GW; Bulau P; Reusch D; Wuhrer M; Habegger M *Anal. Chem* 2017, 89 (10), 5404–5412. [PubMed: 28398745]
- (28). Kukrer B; Filipe V; van Duijn E; Kasper PT; Vreeken RJ; Heck AJ; Jiskoot W *Pharm. Res* 2010, 27 (10), 2197–204. [PubMed: 20680668]
- (29). Fonslow BR; Kang SA; Gestaut DR; Graczyk B; Davis TN; Sabatini DM; Yates JR 3rd *Anal. Chem* 2010, 82 (15), 6643–51. [PubMed: 20614870]
- (30). Reschiglian P; Zattoni A; Roda B; Roda A; Parisi D; Moon MH; Min BR *Ann. Chim* 2006, 96 (5–6), 253–257. [PubMed: 16856754]
- (31). Gomes FP; Yates JR *Mass Spectrom. Rev* 2019, 38, 445–460. [PubMed: 31407381]
- (32). Jorgenson JW; Lukacs KD *Anal. Chem* 1981, 53 (8), 1298–1302.
- (33). Jooss K; Huhner J; Kiessig S; Moritz B; Neuss C *Anal. Bioanal. Chem* 2017, 409 (26), 6057–6067. [PubMed: 28801824]
- (34). Issaq HJ; Janini GM; Atamna IZ; Muschik GM *J. Liq. Chromatogr* 1991, 14 (5), 817–845.
- (35). Hocker O; Montealegre C; Neuss C *Anal. Bioanal. Chem* 2018, 410 (21), 5265–5275. [PubMed: 29943266]
- (36). Shen XJ; Yang ZC; McCool EN; Lubeckyj RA; Chen DY; Sun LL *TrAC, Trends Anal. Chem* 2019, 120, 115644.
- (37). Mehaffey MR; Xia Q; Brodbelt JS *Anal. Chem* 2020, 92 (22), 15202–15211. [PubMed: 33156608]
- (38). Moller J; de Pablo JJ *Biophys. J* 2020, 118 (9), 2057–2065. [PubMed: 32320675]
- (39). Lauring J; Abukhdeir AM; Konishi H; Garay JP; Gustin JP; Wang Q; Arceci RJ; Matsui W; Park BH *Blood* 2008, 111 (2), 856–864. [PubMed: 17942756]
- (40). Jooss K; McGee JP; Melani RD; Kelleher NL Standard procedures for native CZE-MS of proteins and protein complexes up to 800 kDa. *Electrophoresis* 2021.
- (41). Marty MT; Baldwin AJ; Marklund EG; Hochberg GKA; Benesch JLP; Robinson CV *Anal. Chem* 2015, 87 (8), 4370–4376. [PubMed: 25799115]

- (42). Fornelli L; Srzentic K; Huguet R; Mullen C; Sharma S; Zabrouskov V; Fellers RT; Durbin KR; Compton PD; Kelleher NL *Anal. Chem* 2018, 90 (14), 8421–8429. [PubMed: 29894161]
- (43). Susa AC; Xia ZJ; Williams ER *Anal. Chem* 2017, 89 (5), 3116–3122. [PubMed: 28192954]
- (44). Marino-Ramirez L; Kann MG; Shoemaker BA; Landsman D *Expert Rev. Proteomics* 2005, 2 (5), 719–729. [PubMed: 16209651]
- (45). Zheng YP; Huang XX; Kelleher NL *Curr. Opin. Chem. Biol* 2016, 33, 142–150. [PubMed: 27371874]
- (46). Jarvas G; Szigeti M; Guttman AJ *Sep Sci* 2018, 41 (11), 2473–2478.
- (47). Popovic R; Martine-Garcia E; Giannopoulou EG; Zhang QW; Zhang QY; Ezponda T; Shah MY; Zheng YP; Will CM; Small EC; et al. *PLoS Genet* 2014, 10 (9), 1004566.
- (48). Zheng YP; Sweet SMM; Popovic R; Martinez-Garcia E; Tipton JD; Thomas PM; Licht JD; Kelleher NL *Proc. Natl. Acad. Sci. U. S. A* 2012, 109 (34), 13549–13554. [PubMed: 22869745]
- (49). Martinez-Garcia E; Popovic R; Min DJ; Sweet SMM; Thomas PM; Zamdborg L; Heffner A; Will C; Lamy L; Staudt LM; Levens DL; Kelleher NL; Licht JD *Blood* 2011, 117 (1), 211–220. [PubMed: 20974671]
- (50). Taylor GCA; Eskeland R; Hekimoglu-Balkan B; Pradeepa MM; Bickmore WA *Genome Res* 2013, 23 (12), 2053–2065. [PubMed: 23990607]
- (51). Zheng YP; Thomas PM; Kelleher NL *Nat. Commun* 2013, 4, 2203. [PubMed: 23892279]

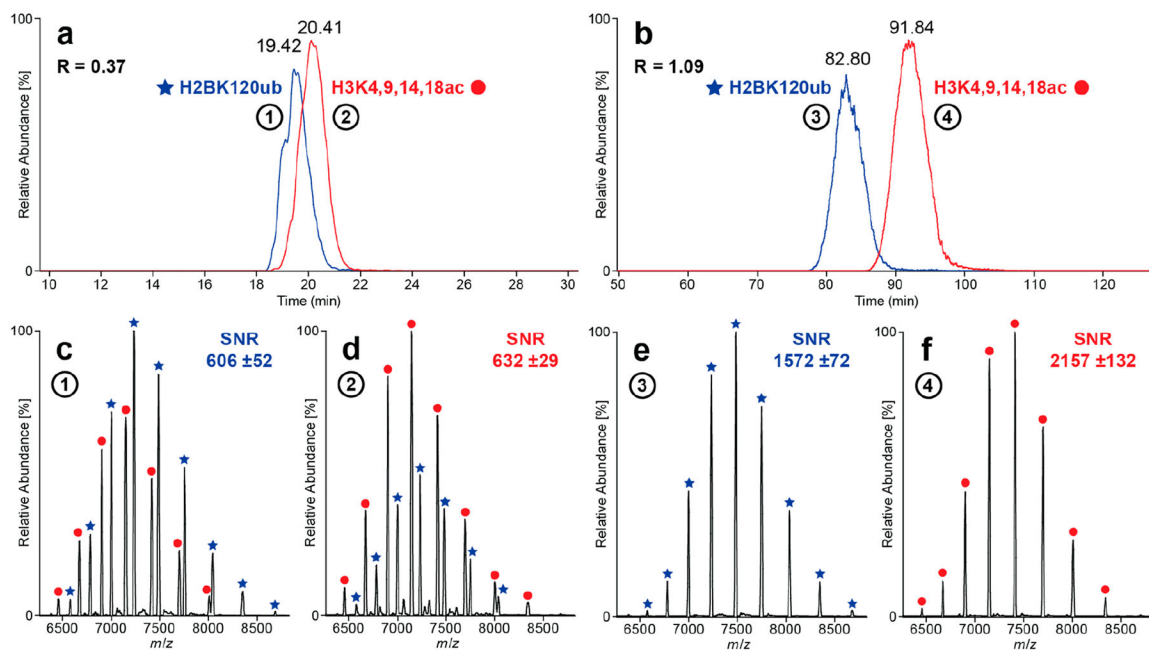


**Figure 1.**

Native CZE–TDMS analysis of Nucs. (a) TIE of Nuc containing H3K27me3 ( $c = 1 \mu\text{M}$ ) showing the separation of the analyte signal and matrix. (b) Mass spectra of high-intensity matrix compounds including sodium acetate (salt) and EDTA (anionic) clusters. (c) Mass spectrum of the intact H3K27me3 nucleosome. (d) Scheme of Nuc migration during analysis. Nucs are negatively charged in AmAc-based BGEs ( $\text{pH} \approx 6.8$ ).

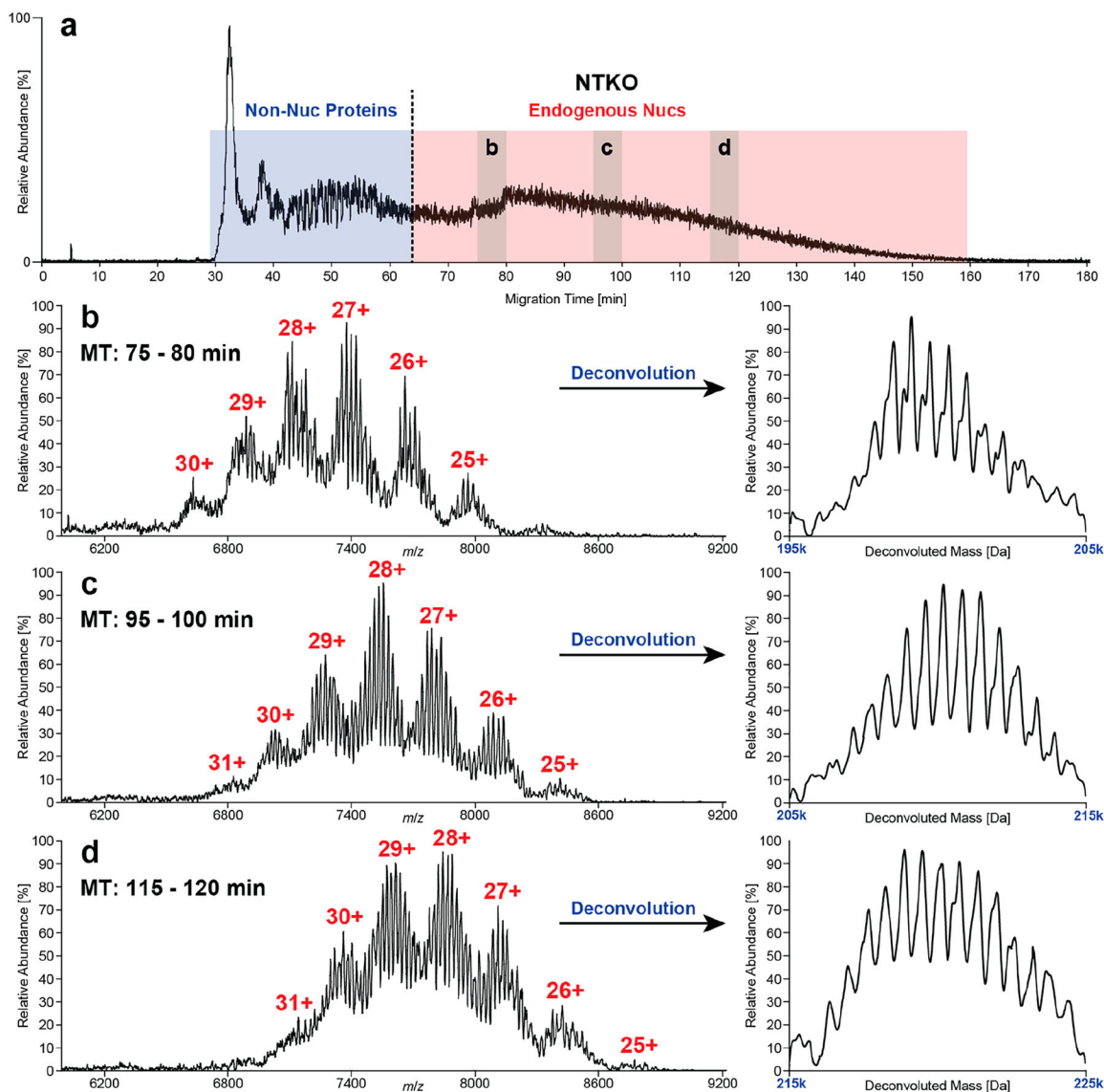


**Figure 2.** Native CZE-TDMS analysis of H3K27me3 ( $c = 1 \mu\text{M}$ ). (a) Mass spectrum of histones ejected from intact Nucs. (b) Deconvoluted mass spectrum of ejected histones including H4, H2B, H2A, and H3K27me3. (c) After histone ejection from the Nuc using in-source dissociation, individual histones were isolated and fragmented by HCD. An example fragment mass spectrum of H2A is depicted. (d) Graphical fragment map of peptides observed from H2A.

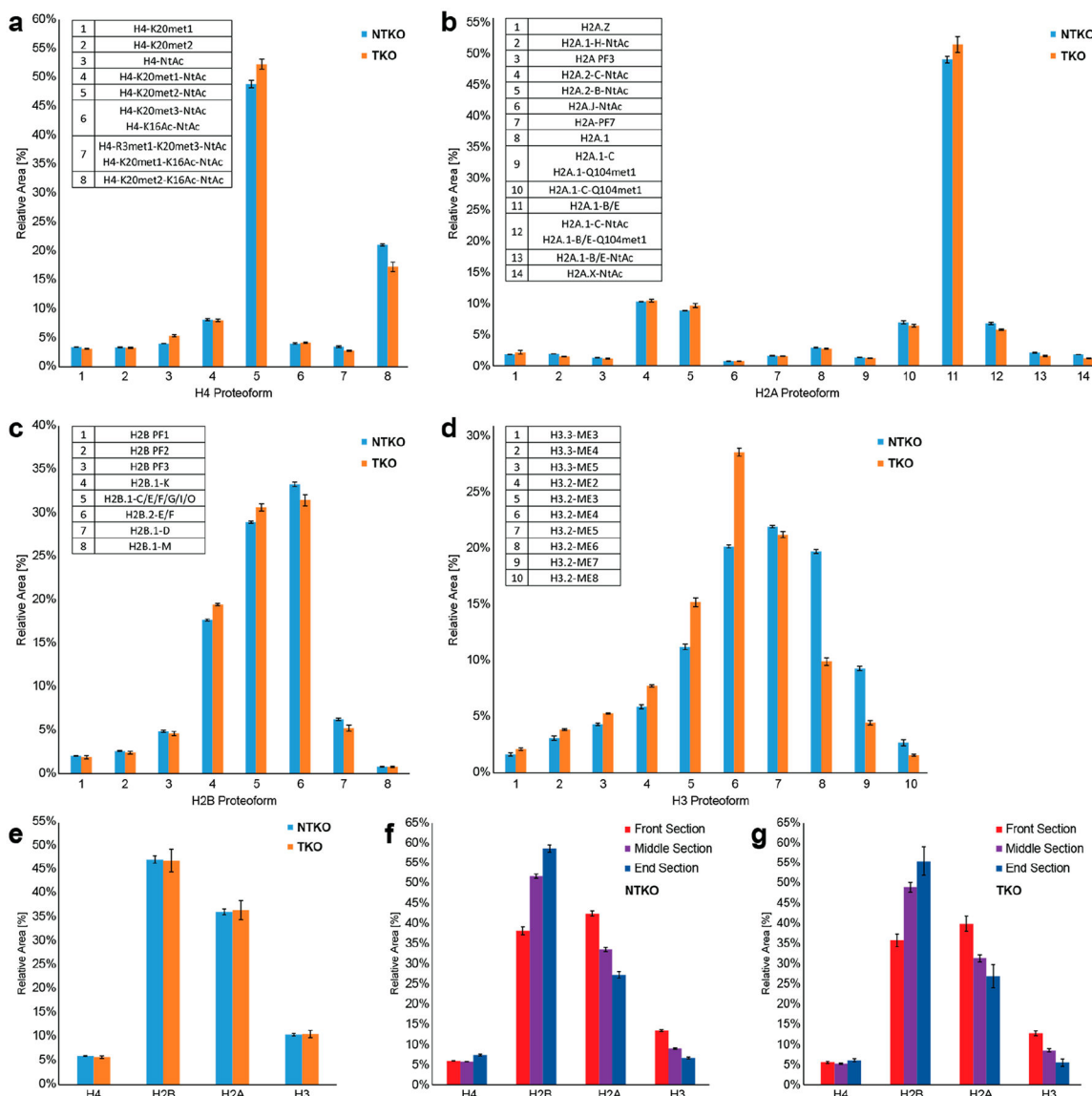


**Figure 3.** Comparison of “standard” and “high-resolution” CZE methods. Extracted ion electropherograms (EIEs, seven highest charge states each) of a sample containing H2BK120ub and H3K4,9,14,18ac analyzed with (a) standard and (b) high-resolution CZE methods. Mass spectra (c–f) generated by averaging the EIE peaks at full width at half-maximum (fwhm). Partial separation ( $R = 0.37$ ,  $n = 2$ ) is obtained using the standard method, and close to baseline separation ( $R = 1.09$ ,  $n = 3$ ) is achieved for the high-resolution method.





**Figure 4.** Separation of endoNucs (~10 mg/mL) derived from NTKO (IGH-NSD2; H3K36me2<sup>hi</sup>) cells. TIE (a) divided into two major regions: (i) non-Nuc proteins and (ii) endogenous Nucs. Raw and deconvoluted mass spectra (b–d) of three exemplary sections (75–80, 95–100, and 115–120 min), where several “nucleoforms” predominated in each. In general, the intact mass of nucleoforms increased with migration time. The average mass shift of  $620.2 \pm 18.5$  Da between neighboring peaks corresponds to a base pair difference in the length of the associated DNA ( $m_{GC} = 618.4$  Da,  $m_{AT} = 617.4$  Da).



**Figure 5.** Relative quantification of NTKO/TKO histones using nCZE–TDMS ( $n = 3$ ). (a–d) Histone proteoform level: averaged mass spectra generated over the entire endoNuc migration range were used, and the total peak areas of H4 (a), H2A (b), H2B (c), and H3.1 (d) proteoforms were set to 100%, respectively. (e–g) Histone type level: (e) comparison between NTKO (IGH-NSD2; H3K36me2<sup>hi</sup>) and TKO (NSD2<sup>±</sup>; H3K36me2<sup>lo</sup>) samples using spectra averaged over the entire endoNuc migration time. Comparison between the front, middle, and end sections of the endoNuc migration time window for NTKO (f) and TKO (g). For this purpose, electropherograms were divided into three equally long sections.