Chasing Tails: Cathepsin-L Improves Structural Analysis of Histones by HX-MS

Authors
Malvina Papanastasiou, James Mullahoo, Katherine C. DeRuff, Besnik Bajrami, Ioannis Karageorgos, Stephen E. Johnston, Ryan Peckner, Samuel A. Myers, Steven A. Carr, and Jacob D. Jaffe

Correspondence
malpap@broadinstitute.org

In Brief
Cathepsin-L, a protease known to proteolyse histone tails in cells, is introduced as a novel enzyme for HX-MS. Cathepsin-L generates overlapping N-terminal peptides, improving coverage of histone tails that are poorly represented in HX-MS employing pepsin alone. Cathepsin-L/pepsin is employed to investigate in-solution dynamics of H3 and H4 monomers. Rapid deuteration of the tails is indicative of unfolded polypeptides, whereas extensive bimodal distributions detected in the histone-folds reveal cooperative unfolding events.

Highlights
• Cathepsin-L is introduced as a novel protease for HX-MS studies.
• Cathepsin-L improves resolution of traditionally challenging histone tails.
• Cathepsin-L can be readily combined with pepsin for improved protein coverage.
• In-solution dynamics of the H3.1 and H4 monomers reveal extensive EX1 kinetics.
Chasing Tails: Cathepsin-L Improves Structural Analysis of Histones by HX-MS*[S]

© Malvina Papanastasiou†‖, ©James Mullahoo‡, Katherine C. DeRuff‡, Besnik Bajrami‡, Ioannis Karageorgos§¶, Stephen E. Johnston‡, Ryan Peckner‡, ©Samuel A. Myers‡, Steven A. Carr†, and ©Jacob D. Jaffe‡

The N-terminal regions (tails) of histone proteins are dynamic elements that protrude from the nucleosome and are involved in many aspects of chromatin organization. Their epigenetic role is well-established, and post-translational modifications present on these regions contribute to transcriptional regulation. Considering their biological significance, relatively few structural details have been established for histone tails, mainly because of their inherently disordered nature. Although hydrogen/deuterium exchange mass spectrometry (HX-MS) is well-suited for the analysis of dynamic structures, it has seldom been employed in this context, presumably because of the poor N-terminal coverage provided by pepsin. Inspired from histone-clipping events, we profiled the activity of cathepsin-L under HX-MS quench conditions and characterized its specificity employing the four core histones (H2A, H2B, H3 and H4). Cathepsin-L demonstrated cleavage patterns that were substrate- and pH-dependent. Cathepsin-L generated overlapping N-terminal peptides about 20 amino acids long for H2A, H3, and H4 proving its suitability for the analysis of histone tails dynamics. We developed a comprehensive HX-MS method in combination with pepsin and obtained full sequence coverage for all histones. We employed our method to analyze histones H3 and H4. We observe rapid deuterium exchange of the N-terminal tails and cooperative unfolding (EX1 kinetics) in the histone-fold domains of histone monomers in-solution. Overall, this novel strategy opens new avenues for investigating the dynamic properties of histones that are not apparent from the crystal structures, providing insights into the structural basis of the histone code. Molecular & Cellular Proteomics 18: 2089–2098, 2019. DOI: 10.1074/mcp.RA119.001325.

Histones are highly conserved proteins that are integral components of the nucleosome, the repeating unit of chromatin. In the "canonical nucleosome structure," 147 DNA base pairs wrap around a histone octamer that consists of a core H3-H4 tetramer flanked by two H2A-H2B heterodimers (1). Interactions within the octamer core and with the surrounding DNA promote a compact and rigid nucleosome structure bordered by flexible N- and C-terminal tails. An extensive array of covalent post-translational modifications (PTMs)1 decorating the "tails" combinatorially regulate gene transcription, replication, and DNA repair, commonly known as the "histone-code" (2). These PTMs are pivotal in modulating interactions of effector proteins with the nucleosomal DNA and with neighboring nucleosomes to modulate the formation of higher order chromatin structure (3). Decoding the information hidden in this extensive modification network and knowing the structure and dynamics of the tails is crucial in understanding regulation of gene expression.

Histones share a highly similar structure known as the histone-fold composed of the helix-loop-helix motif (4) complemented by highly basic, dynamic, flanking termini that extend outside the nucleosome boundaries. Given their disordered nature, parts of the histone tails are absent in high-resolution X-ray nucleosome structures reported to date (1, 5–7), although extension of the DNA sequence has resulted in improved electron density clarity of the tails (8). NMR spectroscopy has provided complete coverage of either all histone tails in single nucleosome particles (9) or of individual tails in condensed nucleosomal arrays (10–14). However, size limitations (15) prohibit the analysis of samples of increasing complexity (i.e. as would be the case of a histone/reader complex), a space in which cryo-EM has already shown great promise (16). Cryo-EM has been employed successfully for elucidating structures of nucleosomes alone or in complex with chromatin modifiers (17–19) and for deciphering the organization of higher-order chromatin (for a recent review see (20)). Finally, computational studies including coarse-grained models (21, 22) and molecular dynamics simulations (23–25) on the tails have provided additional information to the role these have on internucleosomal interactions and chromatin compaction.

Despite the wide applicability of hydrogen-deuterium exchange-mass spectrometry (HX-MS) toward the study of unstructured regions within folded proteins (26–28), there are...
only a few reports employing bottom-up (peptide level) HX-MS to the analysis of histones to-date (29–36) and these mainly focus on features of the histone-fold domains. Histone dynamics of H3 variants and H4 have been reported in the context of tetramers, nucleosomes and nucleosomal arrays (29–31). Further, HX-MS has been employed to study interactions between histone dimers H2A-H2B and H3-H4 in complex with chaperones that facilitate chromatin assembly (32–36). In all these studies, pepsin, the most common enzyme for generating peptide fragments under HX quench conditions (37) was employed and for most histones provided poor coverage of the tails. Neprosin was introduced recently as an alternative enzyme for histone studies producing N-terminal tail peptides 1–38 for H3 and 1–32 for H4 in acidic conditions, (38) demonstrating its potential applicability for future HX-MS studies. Other novel HX-MS proteases that have emerged over the past few years (39–45) have yet to be tested to this class of proteins; their reported cleavage preferences however do not seem to render them appropriate for generating measurable size peptides (7–25 amino acids) given the sequence of the tails. Recently, an alternative approach employing electron transfer dissociation (ETD) in combination with HX-MS top-down and middle-down approaches for this class of proteins was demonstrated (46). Despite the high-resolution histone maps obtained however, the widespread implementation of ETD in HX-MS studies lags typical bottom-up approaches, as the analysis, although seemingly straightforward, presents technical challenges (47–49).

In this study, we sought to investigate the suitability of cathepsin-L, a protease that has been shown to proteolytically process the N-terminal tail of H3 in vivo (50), and to enhance coverage of histone tails in bottom-up HX-MS studies. Cathepsin-L, is one of the 11 cysteine cathepsin proteases that share a common catalytic mechanism and strong sequence similarity with the papain super-family of proteases. Cathepsins are synthesized as propeptides and are processed in the mature form either autocatalytically or by other proteases (51). Their role as degradative enzymes of the lysosome is well established (52), although their observed localization in the nucleus and in tumor microenvironments suggests a variety of extended functions, i.e. in cell development and differentiation (50), as well as in cancer progression (53). Cysteine cathepsins exhibit broad specificity and cleave various proteins in vitro at acidic pH (52). Cathepsin-L has been shown to generate a “primary” cleavage site at H3 Ala21 and additional downstream residues (50, 54), an event of potential epigenetic importance, commonly referred to as “histone clipping” (55). Additionally, cathepsin-L has been shown to generate multiple cleavage sites downstream of H2A Ala21 in vitro (56). Taken together, these data suggest that cathepsin-L may be applicable to the generation of histone N-terminal peptides of appropriate size for HX-MS studies. Inspired by these reports, we investigated the activity of cathepsin-L at HX-MS quench conditions for all core histones. We show enhanced N-terminal coverage for most of the proteins, and developed a method to investigate in-solution dynamics of monomeric H3 and H4 by HX-MS.

EXPERIMENTAL PROCEDURES

Materials—Human recombinant histones (H2A3; Q7L7L0 (M2502S), H2B2E; Q16778 (M2505S), H3.1; P68431 (M2503S) and H4; P62805 (M2504S)) were purchased from New England BioLabs Inc (Ipswich, MA). Mononucleosomes, recombinant human (16–0009) containing H2A1B (P04908), H2B1K (O60814), H3.1 (P68431), and H4 (P62805) were purchased from EpiCypher (Durham, NC). Human recombinant cathepsin-L was purchased either from PromoKine (purity >90%; PK-PRPS7–1135–5, PromoCell GmbH) or from BioVision (purity >90%; 1135–5, Milpitas, CA). Dulbecco’s phosphate buffered saline (PBS; D8537), cathepsin-L inhibitor (RKLW-NH₂, ≥97%; SCG1010), deuterium oxide (99.9 atom %; D; 151882) and guanidine hydrochloride (≥98%; 50950) were purchased from Sigma-Aldrich (Saint Louis, MO). Tris(2-carboxyethyl)-phosphine hydrochloride (TCEP-HCI; 20491) and immobilized pepsin on cross-linked agarose beads (6%; 20343) were purchased from Thermo Scientific (Rockford, IL). Acetone (ACN, 99.9%; A117) and formic acid (FA, >99.5% purity; A117) were from Fisher (Fair Lawn, NJ) and water from J.T.Baker (UT4218–3; Center Valley, PA), Leucine enkephalin (MS Leucine Enkephalin Kit; 700002456) used for lock mass in the MS instrument was from Waters (Milford, MA). SDS-PAGE gels (4–15%, Mini-PROTEAN TGX Precast Protein Gels) were from Bio-Rad (4561086; Hercules, CA) and staining reagent (EZBlueTM Coomassie Brilliant Blue G-250) was from Sigma-Aldrich (G1041).

In Vitro Enzyme Activity Analysis—Individual histones (1 µl of 1 µg/µl, PBS pH 7.5) were diluted with PBS (9 µl, pH ranging from 7.5 to 2.5, TCEP 5 mM), and cathepsin-L (1 µl of 0.5 µg/µl, 25 mM sodium acetate, pH 5.5, TCEP 5 mM) was immediately added to the samples. The enzymatic digestions were allowed to proceed for 2.5 min either on ice (H3) or at room temperature (H4, H2A3 and H2B.1) and were quenched with SCG1010 (1 µl of 240 µM in PBS, pH 2.5). Samples were subsequently subjected to electrophoresis under denaturing conditions and were stained with EZBlueTM following standard protocol.

Intact Protein Analysis and Deconvolution—Intact histones (2 µl, 50 ng/µl in 5% FA) were flow injected and analyzed on a Q Exactive Plus Orbitrap MS (Thermo Fisher Scientific, Bremen, Germany) using a standard m/z range of 400–2000. MS parameters were: spray voltage 2 kV, source temperature 250 °C, AGC target $6e6$, maximum injection time (IT) 200 ms and resolution (at m/z 200) 140,000. Data were acquired in profile mode and ESI mass spectra were deconvoluted using Intact software (Protein Metrics Inc., San Carlos CA).

In-solution Digestions and Analysis Using a Q Exactive Plus MS—Individual histones (1.5 µg, PBS pH 7.5) were diluted with PBS (pH 2.5) and cathepsin-L (0.15 µg, 25 mM sodium acetate, pH 2.5) was added to the samples. For cathepsin-L activity experiments, digestions were performed in the presence of PBS containing TCEP, guanidine HCl, or both (at pH 2.5) at various concentrations (Results and Discussion). The enzymatic digestions were allowed to proceed for 10 min at either 37 °C or 0 °C and were quenched by boiling samples at 95 °C for 5 min. Samples were desalted on Empore C18 silica bead.
Stagetips (57) following standard protocol (58). Elution mixtures were lyophilized and reconstituted in 5% formic acid before nLC-MS/MS analysis. Peptides were injected using a Proxeon Easy nLC 1000 and analyzed on a Q Exactive Plus Orbitrap MS (Thermo Fisher Scientific). Briefly, 2 μl of samples (1 μg/ml) were loaded onto a microcapillary column (360 μm OD × 75 μm ID) with an integrated electrospray emitter tip (10 μm, New Objective) packed with 20 nm of ReproSil-Pur C18-AQ 1.9 μm beads (Dr. Maisch, GmbH). The column was heated to 50 °C using a column heater (Phoenix S&T). Peptides were separated using a flow rate of 200 nl/min and the following gradient: 3% B to 40% B in 30 min, to 90% B in 4 min, keep at 90% B for 6 min and return to initial conditions in 1 min. Samples were analyzed in data-dependent analysis (DDA) mode using a Top-12 method. Ion source parameters were: spray voltage 2 kV, source temperature 250 °C. Full MS scans were acquired in the m/z range 300–2000, with an AGC target 3e6, maximum IT 20 ms and resolution (at m/z 200) 70,000. MS/MS parameters were as follows: AGC target 1e5, maximum IT 50 ms, loop count 10, isolation window 1.6 m/z, isolation offset 0.3 m/z, NCE 27, resolution (at m/z 200) 17,500 and fixed first mass 100 m/z; unassigned and singly charged ions were excluded from MS/MS. Data were acquired in centroid mode in both MS and MS/MS scans. Peptides were identified using PLGS (Waters). For all histones, a nonspecific enzyme search was performed with parameters as described above for SpectrumMill Proteomics Workbench identification. Spectra were recalibrated using Leucine-Enkephalin (556.2771 Da) for lock mass.

Analysis of Deuterium Uptake—Peptide lists were subsequently loaded to DynamX 3.0 (Waters) and deuterium (D) uptake curves of individual peptides were calculated based on the assigned centroid m/z value of the isotopic cluster envelopes. D-uptake values were normalized using experimental values obtained from the analysis of FD controls divided by the deuterium fraction (0.94) in the solvent during the exchange and the theoretical maximum number of exchangeable amides (60). For peptides displaying bimodal distributions, data for each deuteration time point were imported into HX-Express2 (61) and processed using double binomial fittings. Long peptides (>21 residues) that presented convoluted mass envelopes and required triple binomial fitting were excluded from analysis.

RESULTS AND DISCUSSION

Cathepsin-L Activity Profiling at pH 2.5—Cathepsin-L displays optimal activity at slightly acidic pH (~5.5), though activity in the cell nucleus at pH ~7–8 has also been reported (50). To test whether cathepsin-L is functional at pH 2.5 (required during HX-MS quench conditions (37)), we digested the four core histones individually over a pH gradient ranging from basic to acidic to test the enzyme’s activity (Fig. 1A). Cathepsin-L demonstrated cleavage patterns that were substrate and pH-dependent, and interestingly, for all substrates exhibited optimum activity in the lower acidic range (~3.5–2.5). For H3.1, H2A, and H2B, several sub-sequences were detected across the entire pH range tested, although activity against H3.1 was reduced at pH 2.5. Digestions at higher pH values yielded fewer bands indicating lower enzyme activity, in accordance with similar in vitro experiments of cathepsin-L (73.3% sequence similarity) for H3.1 (50). In contrast, cathepsin-L cleaved H4 exclusively at pH values <5.5 with highest intensity fragments detected at pH 2.5. These results confirm previous reports showing that cathepsin-L has maximal activity at acidic pH and that cathepsin-L affects substrate kinetics in a pH-dependent manner (52).

We next employed nanoLC-MS/MS to identify peptides generated by cathepsin-L for the four histones (supplemental Fig. S1) and investigated cleavage preferences. Prior to digestion, intact histones were analyzed individually by ESI-MS, confirming that the major proteins species present had com-
been reported in the literature. Our studies indicate that R23 is based homologous to H2A3) at pH 6 reported previously (56). For H3, cleavage at G44 is in line with recognition by cathepsin-L, whereas substitution of P1 and P2 (positive values in Fig. 2), in agreement with previous studies on cathepsin-L cleavage of hemoglobin (64). In H3, the P2 residue (L20) has been shown to be essential for recognition by cathepsin-L, whereas substitution of P1 and P3 has no significant impact on the proteolytic activity of the enzyme (65).

Enhanced Histone N Terminus Coverage by Cathepsin-L/Pepsin under HX-MS Conditions—Next, we assessed the ability of cathepsin-L to generate peptides from the N-terminal regions (<50 amino acids) of histones that have been difficult to access using pepsin in earlier HX studies (29, 30, 31, 50). For H4, proteolysis using cathepsin-L has not yet been reported in the literature. Our studies indicate that R23 is

Fig. 2. Specificity of cathepsin-L derived from histone digestion at pH 2.5. To build the IceLogo, we employed peptides identified from all four histones (TCEP 5 mM, 37 °C, 10 min) and discarded redundant cleavage sites. Peptide sequences were aligned and loaded to the IceLogo web application (63). Peptides originating from histone sequences were used as a reference set. Subsite nomenclature is according to Schechter & Berger (62). Over-represented amino acids around the cleavage site (P1) in the experimental set are indicated with positive values whereas underrepresented amino acids with negative values. Residues are colored based on their hydrophobicity properties.

The "primary" cleavage sites (derived from peptides account-

Molecular & Cellular Proteomics 18.10 2092

Complete and unmodified sequences (~11–15 kDa, supplemental Fig. S2). Histones were digested in triplicate in a slightly reducing environment that enhances the enzyme’s activity (Fig. 1B). The enzyme’s reproducibility was assessed by the number of common peptides identified within individual histones across replicates. Therefore, for H2A 57% of peptides was identified in three replicates, for H2B 53%, for H3 55% and for H4 65% (supplemental Table S1). Of note, another ~20% of peptides were identified in two replicates for individual histones. The moderate reproducibility observed is attributed to notable secondary exopeptidase activity detected (supplemental Fig. S3A), concordant with cysteine cathepsins that are known to contain critical structural elements for both endopeptidase and exopeptidase activities (52). The median pairwise Pearson’s correlation of peptide intensities detected in at least two replicates was 0.93 (supplemental Fig. S3B).

The "primary" cleavage sites (derived from peptides accounting for ~80% of total ion current) are depicted in supplemental Fig. S3C. For H2A3, cleavage at G44 is in line with cathepsin-L in vitro proteolysis experiments of H2A2A (98% homologous to H2A3) at pH 6 reported previously (56). For H3.1, the "primary" cleavage sites lie toward the C terminus of the protein. However, cleavage sites at the proteins’ N terminus agree with H3.1 fragments generated in vivo by cathepsin-L (50). For H4, proteolysis using cathepsin-L has not yet been reported in the literature. Our studies indicate that R23 is...
Flexible H3 and H4 Tails in Monomers—During DNA replication, newly synthesized histones are recognized by chaperones that coordinate their nuclear import and deposit them onto DNA to form nucleosomes. Most histone chaperones have a preference for binding and importing histone heterodimers (69), and HX structural information has been focused mainly on the analysis of heterodimers (34), tetramers (29) and nucleosomes (30, 31) or on heterodimer-chaperone interactions (32–36), disregarding monomeric states. Recent evidence, however, has shown that H3 and H4 are predominantly monomeric in the cytosol and can be rapidly imported into the nucleus bound tightly to importin-β proteins (70). Obtaining structural information at the monomer level, therefore, is crucial for delineating structural elements governing histone-chaperone interactions. We employed our method using dual cathepsin-L/pepsin digestion to probe conformational dynamics of the H3 and H4 monomers. The monomeric states of the proteins were confirmed using size-exclusion chromatography (supplemental Fig. S4). Proteins were incubated in D_2O for different periods of time (10 s, 100 s, 1000 s) and exchange was performed on ice to slow HX rates (71). On quenching, proteins were digested for a total of 5 min using cathepsin-L followed by online pepsin digestion and analyzed by LC-MS (Fig. 4A). The δ-uptake of individual profiles was subsequently calculated; for peptides with bimodal isotope distributions, the fraction of each population was modeled further (Fig. 4B). In total, 36 peptides were generated for H3 (87% coverage) and 56 for H4 (100% coverage).
Cathepsin-L for Histone Tail Analysis by HX-MS

Summary results of peptide identification under HX-MS conditions of histones following digestion with Cathepsin-L, online pepsin and a combination of both

| Unique peptides (#) | Sequence coverage (%) | Average length | Redundancy |
|---------------------|-----------------------|---------------|------------|
| N-terminal (50 AAs) |                       |               |            |
| Cathepsin-L         | 12 5 4 6 100 100      | 31.3          | 7.5        |
| Pepsin              | 9 7 2 5 100 100       | 33.3          | 6.0        |
| Cathepsin-L/Pepsin  | 14 12 5 12 100 100    | 30.1          | 9.6        |
| Total Protein Length| 41 43 15 40 100 85.6  | 27.7          | 8.8        |
| Cathepsin-L         |                       |               |            |
| Pepsin              | 57 91 54 50 100 98.5  | 23.9          | 10.6       |
| Cathepsin-L/Pepsin  | 80 112 60 88 100      | 24.7          | 15.3       |

Fig. 4. HX-MS workflow. A. Experimental scheme for deuterium exchange of histone monomers (H3.1, H4). Samples were prepared and analyzed in duplicates. B. Cartoon representation of d-uptake data analysis. Isotopic distributions (EX2) were processed in DynamX 3.0 (Waters) and the d-uptake (centroid depicted with a black line) of peptides over time was calculated. Bimodal isotopic distributions (EX1) were processed in HX-Express2 (60) and the d-uptake was calculated (red and blue dashed lines correspond to a centroid value calculated for each distribution) along with the fraction of each population in solution (the size of each bar corresponds to the “area under the curve” for each binomial distribution observed). d-uptake values from EX2 and EX1 were normalized to full deuteration controls as described in Mayne (60).

Both monomers demonstrated highly flexible N termini, evident by extensive deuteration at the earliest time point. For H3, all amide protons exchanged with deuterons (peptides 1–21 and 1–22) indicating that this region is unstructured. To our knowledge this is the first time that dynamics of the H3 N terminus has been captured at this peptide resolution by bottom-up approaches. In previous HX-MS studies of H3 and variants, poor coverage of the N terminus was obtained using pepsin (29–31, 34). In the present study, the higher coverage obtained was made possible by use of cathepsin-L. Exchange profiles for the longer overlapping peptides (1–47, 1–48, and 1–49) extending to the first residues of helix αN (46–57) were obtained for the monomers. Deuterium exchange indicated some slight increase in protection (~88% at 10 s versus 100% for 1–1–22 of H3), suggesting participation of amide protons in H-bonds in the 23–49 region.

For H4, the N-terminal tail was characterized by peptides 1–23, 1–24 and few longer (1–38, 1–48, 25–49) that contain segments of helix α1 and loop L1. For peptides 1–23 and 1–24, d-uptake was ~80% throughout the time-course of the experiment. This could be because of some helical content that has been observed in molecular dynamics studies (73) and corroborated further by circular dichroism for 1–23 of unmodified H4 (17% helical content) (74). Further, previous HX studies of the H2A-H2B dimer have shown that ionic strength plays a critical role in secondary structure formation with lower salt concentrations (60 mM NaCl) leading to partial unfolding and higher salt concentrations enhancing H-bond formation (33). These findings suggest that the tail of the H4 monomer may indeed form a stable secondary structure at our experimental conditions (147 mM NaCl). In the absence of salt, the monomeric H4 tail showed rapid exchange at the earliest time point corroborating further the dependence of structure formation as a function of ionic strength; high salt concentration that was required to maintain the H3-H4 tetrameric complex resulted in reduced deuterium uptake (46). Further, in previous studies of the H3-H4 heterodimer and heterotetramer, H4 peptide 1–24 (pepsin generated) showed rapid exchange at the earliest time point (29, 34), suggesting that the chemical environment also plays a major role in the structural conformation of the tails (34).

Extensive EX1 Kinetics in the Histone Folds of H3 and H4 Monomers—Two mechanisms have been proposed to describe HX kinetics of a protein under native conditions, termed EX2 and EX1 (76, 77). In EX2, the HX chemical rate is slower than the protein's structural refolding rate and the protein goes through many unfolding and refolding events before a successful exchange event; in EX1, the HX rate is faster than the refolding rate and exchange will occur on multiple exposed residues prior to refolding. EX2 kinetics are presented with a single isotopic distribution whereas EX1 with two dis-
tinct mass envelopes, with the lower mass envelope representing the “closed” state and the higher mass envelope representing the “open” state of a protein (Fig. 4B). Dynamics of the histone folds (4) were dominated by bimodal isotopic distributions (supplemental Fig. S6), indicating exchange kinetics in the EX1 regime that occur when the rate of refolding is slower than the rate of chemical exchange (77). As a result, for regions in which we observed EX1 kinetics, we included a second heat map for each protein, with the fast-exchanging regions corresponding to the “open” and the slow-exchanging regions corresponding to the “closed” state (Fig. 5A and 5B) and mapped exchange profiles for both conformations on respective structures (Fig. 5C and 5D) (1). To avoid erroneous assignment of EX1 kinetics, blanks were injected to preclude carryover between samples (78) and the StepWave ion guide that has shown to cause artificial bimodal patterns for highly charged peptides (79) was optimized. For all time points we further calculated the relative population of each state in solution (61) (supplemental Table S3, supplemental Fig. S7). For both histones, relative populations changed over time, suggesting EX1 kinetics rather than two distinct but static conformations in solution that would be indicated if the relative population remained constant (61).

The histone fold is a conserved structural motif, composed of three α-helices (α1, α2, α3) connected via two loops (L1, L2) that promotes histone dimer formation (4); in H3, there is an extra helix preceding the histone fold termed αN (1). Bimodal distributions for H3 residues 48–125 were detected, entailing helix αN and the histone fold motif (denoted as α1-L1-α2-L2-α3 in Fig. 5A). Exceptions to this were the central and carboxy-terminal part of α3 and the C-terminal of the protein that demonstrated EX2 kinetics. Helix αN (residues 48–61) was mostly rigidly packed (α10.8 in “closed” state at 10 s), evident by the low D-uptake over the time course of the experiment (10%; the remaining fraction (~0.2) was unfolded. Helix αN has previously been shown to adopt its state in a context-dependent manner, appearing unfolded in (H3-H4)2 heterotetramers (29, 34, 46) but stably folded in nucleosomes (1, 30, 46), nucleosomal arrays (31) and in complex with chaperones (34). α1-L1 indicated dynamic regions (~30% overall D-uptake in the time course of the experiment) in contrast to α2 that showed the least extent of deuterium incorporation in the histone fold. Segment α2-L2-α3 that promotes tetramer formation of the H3-H4 histone pairs (1), is mainly unfolded (>0.75 of population, >89% D-uptake), EX1 kinetics for this region have been observed previously in the (H3-H4)2 tetramer at high ionic strength and have been explained by the transient loss of tetramerization, further supported by the abolishment of the biphasic behavior upon complex formation of the H3-H4 dimer with DAXX that stabilized this region (34). Further, this region in monomeric H3 and in an H3-H4 dimer complex with ASF1 has been identified to interact with chaperone NASP (80). Taken together, our data suggest that in physiological concentrations of salt and in the
absence of H4, this region presents a high degree of flexibility that may be required to accommodate the breadth of the histone chaperone network. For H4, peptides within the histone fold exhibited extensive EX1 kinetic profiles with α1 and α3 helices occupying mainly an “open” (>0.7 at 10 s) and helix α2 mainly a “closed” state (−0.8 at 10 s). For α1 and α3, the “closed” population exchanged >40% deuterium at 10 s indicating partially folded elements (supplemental Fig. S5B). This contrasts with α2, where D-uptake was <20% in the time course of the experiment indicating a tightly folded helix (supplemental Fig. S7). EX1 behavior of the α2 helix has been observed in the (H3-H4)2 tetramer, but not when H4 forms a tetramer complex with the H3 variant CENP-A (29). CENP-A appears to rigidify helix α2 and minimize exchange. Further, cooperative unfolding detected previously for both H2A and H2B in the H2A-H2B heterodimer was also localized at helical regions of the histone fold. Similarly to the H3-H4 dimer with DAXX (34), bimodal distributions were no longer observed upon complex formation of the H2A-H2B dimer with Nap1 (33).

Overall, EX1 kinetics seem to be a feature of certain regions of histones, irrespective of whether these are in a monomeric state, engaged in dimers, tetramers or in complexes with chaperones. Differences in experimental conditions (such as temperature and salt concentrations, affecting H-bond formation) render comparison of prior studies to our current work difficult with respect to establishment of the exact boundaries of regions that exhibit EX1 exchange kinetics. However, the EX1 kinetics detected in our study for monomeric histones are extensive and expand the entire protein sequence (omitting the N-terminal tails), suggesting high conformational flexibility, in contrast to histones engaged in complexes where EX1 behavior is confined to short regions. Considering recent evidence showing that H3 and H4 monomers associate with importin-beta proteins and are imported into the nucleus in a monomeric state expands their interaction network further (70). We presume that a high degree of flexibility is a feature of histones required to counterbalance the breadth of chaperones that regulate their nuclear import and assembly into nucleosomes and provide exciting directions for future studies.

In summary, cathepsin-L in tandem with pepsin provided a high-resolution peptide map of histones suitable for HX-MS, including their tails that have thus far eluded crystallographic characterization. The introduction of cathepsin-L as a new tool to facilitate dual-protease techniques could be highly impactful for the study of important biological molecules beyond histones. Such cases may include protein classes known to be challenging for HX-MS, such as antibodies (81), or proteins which have been identified as cathepsinL’s substrates (i.e. elastins, collagens and proteoglycans) (52). For example, the fusion viral glycoprotein F has been reported to be activated through cleavage by cathepsin-L in cells (82) but was presented with moderate coverage in an HX-MS study employing pepsin only (83). We think that the advance we have shown here for histones will encourage the HX-MS community to consider cathepsin-L when probing challenging proteins, and therefore the biggest potential impact is to the HX-MS community rather than the narrow histone structural biology one.

Acknowledgments—We thank Leland Mayne (University of Pennsylvania) and Miklos Guttman (University of Washington) for discussions with respect to the analysis of bimodal distributions.

DATA AVAILABILITY

The original mass spectra may be downloaded from MassIVE (http://massive.ucsd.edu) using the identifier: MSV000083315. The data is directly accessible via ftp://massive.ucsd.edu/MSV000083315.

* This work was supported by a grant from the National Institute of Health (NIH) to J.D.J. at the Broad Institute of MIT & Harvard (U54-HG008097).

This article contains supplemental material. The authors do not declare any conflict of interest.

To whom correspondence should be addressed: Broad Institute of MIT and Harvard, 415 Main St, Cambridge, MA 02142. Tel.: (617) 714-8497; E-mail: malpap@broadinstitute.org.

Author contributions: M.P. and S.A.M. designed research; M.P., J.M., K.C.D., B.B., and I.K. performed research; M.P., J.M., K.C.D., I.K., R.P., and S.A.M. analyzed data; M.P., S.A.M., S.A.C., and J.D.J. wrote the paper; B.B. and S.E.J. contributed new reagents/analytic tools; S.A.C. and J.D.J. supervised research.

REFERENCES

1. Lugger, K., Rechsteiner, T. J., Flaus, A. J., Waye, M. M. Y., and Richmond, T. J. (1997) Characterization of nucleosome core particles containing histone proteins made in bacteri a, J. Mol. Biol. 272, 301–311
2. Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. Nature 403, 41–45
3. Pepenella, S., Murphy, K. J., and Hayes, J. J. (2014) Intra- and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. Chromosoma 123, 3–13
4. Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991) The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. U.S.A. 88, 10148–10152
5. White, C. L., Suto, R. K., and Lugger, K. (2001) Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. EMBO J. 20, 5207–5218
6. Tachiwana, H., Kagawa, W., Osakabe, K., Kawaguchi, K., Shiga, T., Hayashi-Takanaka, Y., Kimura, H., and Kurumizaka, H. (2010) Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. Proc. Natl. Acad. Sci. U.S.A. 107, 10454–10459
7. Aguilar-Gurrieri, C., Larabi, A., Vinayachandran, V., Patel, N. A., Yen, K., Reja, R.,EBong, I. O., Schoehn, G., Robinson, C. V., Fugh, B. F., and Parne, D. (2016) Structural evidence for Nap1-dependent H2A-H2B deposition and nucleosome assembly. EMBO J. 35, 1465–1482
8. Davey, C. A., Sargent, D. F., Lugger, K., Maeder, A. W., and Richmond, T. J. (2002) Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. J. Mol. Biol. 319, 1097–1113
9. Zhou, B.-R., Feng, H., Ghirlando, R., Kato, H., Gruschus, J., and Bai, Y. (2012) Histone H4 K16Q mutation, an acetylation mimic, causes structural disorder of its N-terminal basic patch in the nucleosome. J. Mol. Biol. 421, 30–37
10. Kato, H., Gruschus, J., Ghirlando, R., Tjandra, N., and Bai, Y. (2009) Characterization of the N-terminal domain of histone H3 in condensed nucleosome arrays by hydrogen exchange and NMR. J. Am. Chem. Soc. 131, 15104–15105
Cathepsin-L for Histone Tail Analysis by HX-MS

51. Turk, B., Turk, D., and Turk, V. (2000) Lysosomal cysteine proteases: more than scavengers. Biochim. Biophys. Acta 1477, 98–111
52. Turk, V., Stoka, V., Vasiljeva, O., Renko, M., Sun, T., Turk, B., and Turk, D. (2012) Cysteine cathepsins: From structure, function and regulation to new frontiers. Biochim. Biophys. Acta 1824, 68–88
53. Olson, O. C., and Joyce, J. A. (2015) Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response. Nat. Rev. Cancer 15, 712–729
54. Duarte, L. F., Young, A. R. J., Wang, Z., Wu, H.-A., Panda, T., Kou, Y., Kapoor, A., Hassan, D., Mills, N. R., Ma‘ayan, A., Narita, M., and Bernstein, E. (2014) Histone H3.3 and its proteolytically processed form drive a cellular senescence programme. Nat. Commun. 5, 5210
55. Dhaenens, M., Gilbert, P., Meert, P., Vossaert, L., and Deforce, D. (2015) Histone tails and the H3 αN helix regulate nucleosome mobility and stability. Mol. Cell. Biol. 27, 4037–4048
56. Anderson, L. C., Karch, K. R., Ugrin, S. A., Coradin, M., English, A. M., Siddoli, S., Shabanowitz, J., Garcia, B. A., and Hunt, D. F. (2016) Analyses of histone proteoforms using front-end electron transfer dissociation-enabled orbitrap instruments. Mol. Cell. Proteomics 15, 975–988
57. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. 2, 1896–1906
58. Mertins, P., Qiao, J. W., Patel, J., Ushidi, N. D., Clauser, K. R., Mani, D. R., Burgess, M. W., Gillette, M. A., Jaffe, J. D., and Carr, S. A. (2009) Integrated proteomic analysis of post-translational modifications by cellular enrichment. Nat. Methods 10, 634–637
59. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods 4, 207–214
60. Mayne, L. (2016) Chapter thirteen - Hydrogen exchange mass spectrometry. Methods Enzymol. 566, 335–356
61. Guttmann, M., Wels, D. D., Engen, J. R., and Lee, K. K. (2013) Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. J. Am. Soc. Mass Spectrom. 24, 1906–1912
62. Schecter, I., and Berger, A. (1967) Pi and Pl are commonly used nomenclature for amino acid residues of a substrate (or an inhibitor). Biochem. Biophys. Res. Commun. 27, 157
63. Colaert, N., Helsen, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009) Improved visualization of protein consensus sequences by iceLogo. Nat. Methods 6, 786–787
64. Lowell, J., Robinson, M. W., Donnelly, S. M., Xu, W., Stack, C. M., Matthews, J. M., and Dalton, J. P. (2009) The importance of pH in regulating the function of the Fasciola hepatica cathepsin L1 cysteine protease. PLoS Negl. Trop. Dis. 3, e639
65. Adams-Gioaba, M. A., Krupa, J. C., Xu, C., Mort, J. S., and Min, J. (2011) Structural basis for the recognition and cleavage of histone H3 by cathepsin L. Nat. Commun. 2, 197
66. Rand, K. D., Zehl, M., and Jørgensen, T. J. D. (2014) Measuring the hydrogen/deuterium exchange of proteins at high spatial resolution by mass spectrometry: overcoming gas-phase hydrogen/deuterium scrambling. Acc. Chem. Res. 47, 3018–3027
67. Kan, Z.-Y., Walters, B. T., Mayne, L., and Englander, S. W. (2013) Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. Proc. Natl. Acad. Sci. U.S.A. 110, 16438–16443
68. Rand, K. D., Zehl, M., Jensen, O. N., and Jørgensen, T. J. D. (2009) Protein hydrogen exchange measured at single-residue resolution by electron transfer dissociation mass spectrometry. Anal. Chem. 81, 5577–5584
69. Keck, K. M., and Pernbort, L. F. (2012) Histone chaperones link histone nuclear import and chromatin assembly. Biochim. Biophys. Acta 1819, 277–289
70. Apta-Smith, M. J., Hernandez-Fernaud, J. R., and Bowman, A. J. (2018) Evidence for the nuclear import of histones H3.1 and H4 as monomers. EMBO J. 37, e98714
71. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange. Proteins 17, 75–86
72. Ferreira, H., Somers, J., Webster, R., Flaus, A., and Owen-Hughes, T. (2007) Histone tails and the H3 αN helix regulate nucleosome mobility and stability. Mol. Cell. Biol. 27, 4037–4048
73. Lins, R. D., and Röthlisberger, U. (2006) Influence of long-range electrostatic treatments on the folding of the N-terminal H4 histone tail peptide. J. Chem. Theory Comput. 2, 246–250
74. Wang, X., Moore, S. C., Lasczczak, M., and Ausió, J. (2000) Acetylation increases the α-helical content of the histone tails of the nucleosome. J. Biol. Chem. 275, 35013–35020
75. Dorigo, B., Schalch, T. Bystricky, K., and Richmond, T. J. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail. J. Mol. Biol. 327, 85–96
76. Englander, S. W., and Kallenbach, N. R. (1983) Hydrogen exchange and structural dynamics of proteins and nucleic acids. Q. Rev. Biophys. 16, 521–565
77. Englander, S. W. (2006) Hydrogen exchange and mass spectrometry: A historical perspective. J. Am. Soc. Mass Spectrom. 17, 1481–1489
78. Fang, J., Rand, K. D., Beuning, P. J., and Engen, J. R. (2011) False EX1 signatures caused by sample carryover during HX MS analyses. Int. J. Mass Spectrom. 302, 19–25
79. Guttmann, M., Wales, T. E., Whittington, D., Engen, J. R., Brown, J. M., and Lee, K. K. (2016) Tuning a high transmission ion guide to prevent gas-phase proton exchange during H/D exchange MS analysis. J. Am. Soc. Mass Spectrom. 27, 662–668
80. Bowman, A., Lercher, L., Singh, H. R., Zinne, D., Timinszky, G., Carlon-agnon, T., and Ladumem, A. G. (2016) The histone chaperone sNAP52 binds a conserved peptide motif within the globular core of histone H3 through its TPR repeats. Nucleic Acids Res. 44, 3105–3117
81. Hudgens, J. W., Gallagher, E. S., Karageorgos, I., Anderson, K. W., Filliben, J. J., Huang, R. Y.-C., Chen, G., Bou-Assaf, G. M., Espada, A., Chalmers, M. J., Harguindey, E., Zhang, H.-M., Walters, B. T., Zhang, J., Venable, J., Steckler, C., Park, I., Brock, A., Lu, X., Pand ey, R., Chandramohan, A., Anand, G. S., Niurdodhi, S. N., Sperry, J. B., Rouse, J. C., Carroll, J. A., Rand, K. D., Leurs, U., Wes, D. D., Al-Najashibandi, M. A., Hageman, T. S., Deredde, D., Wintrode, P. L., Papanastasious, M., Lambris, J. D., Li, S., and Urata, S. (2019) Interlaboratory comparison of hydrogen-deuterium exchange mass spectrometry measurements of the fab fragment of NIST-mAb. Anal. Chem. 91, 7336–7345
82. Pager, C. T., and Dutch, R. E. (2005) Cathepsin L is involved in proteolytic processing of the Hendra virus fusion protein. J. Virol. 79, 12714–12720
83. Wong, J. J. W., Young, T. A., Zhang, J., Liu, S., Leser, G. P., Komives, E. A., Lamb, R. A., Zhou, Z. H., Salafsly, J., and Jardetzky, T. S. (2017) Monomeric ephrinB2 binding induces allosteric changes in Nipah virus G that precede its full activation. Nat. Commun. 8, 781