Published in final edited form as:
Biochemistry. 2010 June 29; 49(25): 5244–5257. doi:10.1021/bi100391w.

Extensive and varied modifications in histone H2B of wildtype and histone deacetylase 1-mutant *Neurospora crassa*

D.C. Anderson[1,3,*], **George R. Green**[2], **Kristina Smith**[1,4], and **Eric U. Selker**[1]

[1] Institute of Molecular Biology, University of Oregon, Eugene OR 97403

[2] Southern School of Pharmacy, Mercer University, Atlanta GA 30341

Abstract

DNA methylation is deficient in a histone deacetylase 1 (HDA11) mutant (*hda-1*) strain of *Neurospora crassa* with inactivated histone deacetylase 1. Difference 2D gels identified the primary histone deacetylase 1 target as histone H2B. Acetylation was identified by LC/MS/MS at 5 different lysines in wild type H2B, and at 11 lysines in *hda-1* H2B, suggesting *Neurospora* H2B is a complex combination of different acetylated species. Individual 2D gel spots were shifted by single lysine acetylations. FTICR MS-observed methylation ladders identify an ensemble of 20–25 or more modified forms for each 2D gel spot. Twelve different lysines or arginines were methylated in H2B from wild type or *hda-1*; only two were in the N-terminal tail. Arginines were modified by monomethylation, dimethylation or deimination. H2B from wild type and *hda-1* ensembles may thus differ by acetylation at multiple sites, and by additional modifications. Combined with asymmetry-generated diversity in H2B structural states in nucleosome core particles, the extensive modifications identified here can create substantial histone-generated structural diversity in nucleosome core particles.

The lowest level of organization of eukaryotic chromosomes is achieved by wrapping ~146 bp segments of DNA around histone octamers, which consist of a histone H3-H4 tetramer and two histone H2A-H2B dimers (1). The resulting string of nucleosomes can be further condensed into higher order assemblies. It is becoming increasingly clear that the structure of chromatin, both at the level of individual nucleosomes and at higher levels, is functionally important for genetic processes such as gene expression, recombination, DNA methylation and DNA repair. Thus it is of interest to define ways in which chromatin can be modified. In the last decade, a large number of enzymes have been identified that post-translationally

---

[1]**Abbreviations used:** HDA1, histone deacetylase 1; *hda-1*, *Neurospora crassa* strain with mutant HDA1; WT, wild-type *Neurospora crassa*; FTICR, fourier transform ion cyclotron resonance; MS, mass spectrometry; LC/MS/MS, liquid chromatography- tandem mass spectrometry identification of peptides; me, methyl; me2, dimethyl; me3, trimethyl; ac, acetyl; cit, citrulline; TCA, trichloroacetic acid; AUT, acetic acid acid-urea-triton X-100; AU, acetic acid-urea; 2D gel, two-dimensional gel; MALDI-TOF, matrix-assisted laser desorption ionization-time of flight; HPLC, high performance liquid chromatography; ROC, receiver operating characteristic; SVM support vector machine learning, NCP, nucleosome core particle.

[*]To whom correspondence should be addressed: dca0204@gmail.com, (650) 355-1168. Fax (541) 346-5891 (Eric Selker).
[3]Currently at Catalyst Biosciences, South San Francisco CA 94080
[4]Currently at Dept. of Biochemistry and Biophysics, Oregon State University, Corvallis OR 97331

Supporting Information Available: High resolution FTMS spectra comparing methylation of *hda-1* spots a and b (Fig. 2B), *hda-1* spots b and c (Fig. 2C), WT spot a and *hda-1* spot a (Fig. 2D), and annotated MS/MS spectra for peptides identified in Tables 2–7. This material is available free of charge via the Internet at http://pubs.acs.org.

modify histones, introducing chemical modifications, some of which are associated with gene activation and others with repression. The N-terminal tails of the individual histone subunits are prominent targets for modifications, which can alter the higher order structure of chromatin and influence recruitment of effector molecules (2–4). Modifications can also occur within the globular histone core (5), affecting binding of DNA to the nucleosome lateral surface, and interactions with nucleosome remodeling complexes that control mobility of nucleosomes on DNA (6).

Mass spectrometry has revolutionized the study of histones, often allowing site-specific identification of post-translational modifications such as lysine acetylation, mono-, di-, and trimethylation, arginine mono- and dimethylation, deimination, and serine and threonine phosphorylation (7). These and other post-translational modifications constitute the “histone code” that is important in the transmission of genetic information in cells similar to but distinct from the contributions of genes themselves (2,4). Advanced techniques such as electron transfer dissociation (8) and sequential ion-ion reactions (9) have allowed site-specific determination of modifications in highly basic regions of histones such as histone H3.

Epigenetic processes such as DNA methylation, which can result in gene silencing (10), have been linked to histone methylation (11). Patterns of aberrant DNA methylation are commonly observed in human tumors, including a global loss of DNA methylation and CpG island hypermethylation, associated with inactivation of tumor suppressor genes (12–14) and changes in expression of other genes. DNA methylation is conveniently studied in organisms that are amenable to genetic manipulation, such as *Neurospora crassa*, which has only a single gene for each core histone except H4, for which two genes have been identified (15). Having only a single gene for a histone subtype may facilitate association of post-translational modifications with biological phenomena. In contrast, expressed human H2B is composed of at least 11 subtypes in Jurkat cells (16) and 7 subtypes in Hela cells (17), making associations more difficult. DNA methylation is not essential in *Neurospora*, allowing study of a variety of non-lethal mutations affecting methylation. Genes involved in control of DNA methylation in *Neurospora* include DIM-1 (18), the DNA methyltransferase DIM-2 (19), the histone H3 lys9 methyltransferase DIM-5 (20), Heterochromatin Protein 1 (HP1; 21), the histone H3 ser10 phosphatase PP1 (22), and dim-7, which helps target DIM-5 to chromatin (23). Thus covalent modifications to DNA and histones are involved in this process. Trichostatin A is known to block selective DNA methylation in *Neurospora* (24), suggesting that one or more histone deacetylases are also involved in control of DNA methylation. Mutation of the *Neurospora* histone deacetylase gene HDA1 causes a loss of DNA methylation in about half of all methylated regions (25), suggesting that analysis of the histone substrate(s) and changes in modifications associated with inactivation of this enzyme may shed light on its control of DNA methylation.

We used high resolution 2D gels (26) to examine histones affected by HDA1 inactivation, identifying H2B as the most prominent HDA1 histone target in *Neurospora* (25). Since *Neurospora* histone modifications had not been previously characterized by mass spectrometry, we examined overall *Neurospora* H2B post-translational modifications using FTICR mass spectrometry on electroeluted H2B spots isolated from wild type and *hda-1*

strains. We then used LC/MS/MS peptide sequencing to locate modifications on specific peptides, and many on specific residues. To extend the examination of post-translational modifications, results obtained using SEQUEST (27) were compared to modifications identified using the blind search algorithm InsPecT (28). Modifications were compared in H2B samples from wild type and *hda-1* strains, and compared with H2B from other organisms. We report that *N. crassa* histone H2B can be heavily modified, including extensive methylation, diverse lysine acetylation, arginine modifications including methylation, dimethylation and deimination, and several likely *in vitro* modifications.

To examine potential function(s) of the H2B modifications, we mapped them onto the crystal structure of the *Xenopus* nucleosome core particle. The results identify modifications that may affect DNA binding in the nucleosome, as well as modifications that may alter histone subunit interactions, histone assembly, and higher order nucleosome assemblies.

Methods

Neurospora cultures and histone purification

Neurospora crassa was grown in 500 ml culture flasks in Vogel's minimal medium supplemented with 2% sucrose, 1.0 mg/ml alanine and 1.0 mg/ml histidine at 30°C for 24 hr with shaking at 150 rpm. Mycelia were collected on Whatman filter paper and immediately frozen in liquid nitrogen. Chromatin was prepared from the frozen tissues and histones were extracted as previously described for *Saccharomyces cerevisiae* histones, with modifications described below (29). The frozen tissue was ground to a fine powder using a mortar and pestle cooled with liquid nitrogen, and the powder was suspended in APB Buffer (150 mM sodium chloride, 10 mM Tris-HCl, 0.5% Triton X-100, 1.0 mM phenylmethanesulfonyl fluoride, 0.2% β-mercaptoethanol, 1.0 mM sodium pervanadate, 5.0 mM sodium fluoride, pH 8.0). The tissue suspension was pelleted by centrifugation at 14,000 × g for 10 min. The supernatant was discarded and the chromatin pellet was washed twice with APB buffer without 0.5% Triton X-100.

Histones were extracted from the nuclear suspensions as previously described (26). An equal volume of ice-cold 0.4 M H_2SO_4 was added to the nuclear suspension containing 1 mg/ml protamine sulfate in 5% acetic acid, followed by incubation on ice for 24 hr, with occasional mixing. Acid-insoluble material was pelleted by centrifugation at 14,000 × g for 10 min and the supernatant was mixed with 100% TCA to a final concentration of 20% TCA and incubated on ice for 1 hr to precipitate acid-soluble proteins. The precipitate was collected by centrifugation for 10 min, 14,000 × g at 4°C and the protein pellet was resuspended in ice-cold acetone, sonicated in a bath sonicator to disperse the precipitated proteins and recovered by centrifugation at 4°C, 14,000 × g for 10 min. The supernatant was decanted and discarded and the acetone precipitate was air-dried at room temperature. Histones were dissolved in acetic acid-urea loading buffer (8.0 M urea, 5% acetic acid, 5% β-mercaptoethanol, 0.2 mg/ml crystal violet) to a concentration of 5 mg/ml.

2D PAGE and Electroelution of histones from fixed and stained 2D gels

Two dimensional AUT × AU polyacrylamide gels contained acetic acid/urea/Triton X-100 in the horizontal dimension and acetic acid/urea in the vertical dimension and were run as described (26). The gels were stained with Coomassie blue R-250. Histone samples were purified from individual protein spots in fixed and stained polyacrylamide gels as previously described (26). Briefly, individual spots were excised from the gel, equilibrated in elution buffer and eluted from the gel using a custom made electroelution apparatus (30). Eluted proteins were precipitated from solution using an acetone-based ion extraction buffer to remove SDS, Coomassie blue and other contaminants from the sample.

Proteolytic digests of histones

Purified histones electroeluted from 2D gel spots were dissolved in pH 7.8 50 mM sodium phosphate buffer with 50% trifluoroethanol or with 0.1% w/v acid-cleavable RapiGest detergent (Waters Corp., Milford, MA), heat denatured at 95 °C for 20 min. in a heat block and digested for at least 12 hours with 1:10 enzyme: substrate (weight:weight) of different proteases. Progress of the digestions was checked by MALDI-TOF mass spectrometry on an Applied Biosystems Voyager DE instrument. One protease was *Staphylococcus aureus* endoprotease gluC (Calbiochem, San Diego CA, or Worthington Biochemical Corp., Lakewood N.J.), with digestion in pH 7.8 50 mM sodium phosphate buffer for at least 12 hours (31). This pH and buffer, as well as a high enzyme:substrate ratio of 1:10 w:w gluC:histone H2B and relatively long digestion times, were used to enhance cleavage after aspartic acid. To enhance proteolysis at asp beyond the levels seen with endoprotease gluC, other digests included 1:10 w:w endoprotease aspN (Roche Applied Science, Penzberg, Germany) in pH 7.8 50 mM sodium phosphate buffer at 37C for 12–24 hours, or a combination of endoproteinase aspN and endoprotease gluC at pH 7.8. H2B was also digested with clostripain (Worthington Biochemical, Lakewood NJ) in pH 7.6 100 mM sodium phosphate, 10 mM dithiothreitol and 1 mM calcium chloride. *Neurospora* H2B was not reduced and alkylated before proteolysis since it contains no cysteines. The RapiGest detergent was hydrolyzed in 0.1% formic acid before peptides were injected onto a capillary column for LC/MS/MS analysis. Peptides were analyzed with and without propionylation of the N-terminus and lysines. Propionylation, which results in peptide charge reduction and more readily-interpreted MS/MS spectra, was carried out as described for histones (32).

LC/MS/MS analysis of histone H2B peptides

Peptides were sequenced and data for identification of post-translational modifications were obtained by tandem mass spectrometry on a ThermoFinnigan (San Jose, CA) LTQ-FT hybrid linear ion trap-fourier transform ion cyclotron resonance mass spectrometer (33) equipped with a New Objective Inc. (Woburn, MA) PV-550 source, after chromatography on a 75 micron internal diameter, 5–10 cm long PicoFrit column (New Objective, Inc.) self-packed with Vydac Denali C18 resin. The column was loaded with the autosampler, and eluted using an Agilent nano-1100 HPLC at a flow rate of 0.3–0.4 ul/min, with the digest eluted for 40–50 min. at 5% acetonitrile-0.1% formic acid, followed by a 3%/min increasing gradient of acetonitrile-0.1% formic acid to 90% acetonitrile, and a 5–10 min. elution at 90% acetonitrile-0.1% formic acid. Peptide precursor ions were analyzed in the ICR cell at a

resolution of 100,000 and MS/MS fragments analyzed in the LTQ linear ion trap. One FT survey scan (maximum fill time 2 sec, target value 5×10^5) triggered parallel LTQ acquisition of MS/MS spectra for the three most intense peptide ions (target value 5×10^3, maximum fill time 0.5 sec). Monoisotopic masses were used for precursor and product ions in Bioworks Browser 3.2 SEQUEST searches, with a precursor ion mass uncertainty of 10 ppm, and the standard ion trap fragment ion mass uncertainty of 1.0 m/z. Dynamic exclusion was activated, with a maximum repetition of 2 over a 30 sec time period, and an exclusion duration of 60 sec. The LTQ was tuned and calibrated within 1–2 weeks of experiments, and the FT was mass-calibrated within 1–2 days of use. MS/MS spectra were plotted using Protein Prospector's MS-Product (prospector.ucsf.edu).

FTICR mass spectrometry

Whole-protein FTICR mass spectra were obtained on histone H2B dissolved in 50% acetonitrile-0.1% trifluoroacetic acid using static nanospray from 4 micron tip internal diameter coated glass capillaries (New Objective Inc., Woburn MA), on the LTQ-FT. The instrument was mass-calibrated before use and operated at a spray voltage just above threshold, which varied for different samples and capillaries from 1.2 – 1.8 kV; data were collected at a resolution of 200,000. To minimize file size, from 2–8 microscans were averaged for each saved scan. The histone H2B charge envelope was collected from 550–1200 m/z in the ICR cell. Whole protein spectra were analyzed with a commercial version of THRASH (34), called Xtract (ThermoFinnigan, San Jose CA), using a 60% fit, a signal/noise setting of 1.3, low sulfur averagine, and a remainder of 1%.

Post translational modification analysis

SEQUEST (27) was used for most analyses of post-translational modifications, searching a database of 7663 protein sequences containing all *Neurospora crassa* histone sequences, human cytokeratin sequences, sequences of the proteases used to digest electroeluted histone H2B, and reversed protein sequences of the *Nostoc punctiforme* proteome (http://genome.ornl.gov/microbial/npun/). Modification monoisotopic masses used for database searching were: acetyl 42.01057 Da, trimethyl 42.04695 Da, dimethyl 28.03130 Da, methyl 14.01570 Da, deimination 0.98402 Da, oxidation 15.99492 Da and phosphorylation 79.96633. Modification types were searched in groups of six different optional modifications. For the above group of six optional modifications (without phosphorylation) used with the above database, the 13-parameter support vector machine learning test ROC score (35) from 10-fold cross-validation was 0.960 + 0.009. Numerous initial searches did not find any phosphorylated H2B sites, so further searches did not consider this modification. Modifications examined by SEQUEST included lysine (unmodified, mono-, di- and trimethylated, acetylated), arginine (unmodified, deiminated, mono- and dimethylated), methionine (oxidized), and serine, threonine and tyrosine (all phosphorylated). The precursor ion spectra of peptides identified with citrulline (Table 4), and of control peptides containing unmodified arg, were individually examined to insure the monoisotopic mass used in the database search was selected correctly by Bioworks Browser; a number of false positives were eliminated by this manual examination. No N-terminal H2B modifications were identified in peptides containing the N-terminus of the protein, in separate searches for mono-, di- or tri-methylation or acetylation. Peptides with termini not

expected for the protease(s) used for digestion were not included in the results. The mass of a precursor ion alone, measured by FTICR-MS, is usually sufficient to derive the composition of modifications on a peptide (36) including distinguishing trimethyllysine from acetyl-lysine. The location of modifications was then derived using SEQUEST and the MS/MS spectrum, using a difference in the SEQUEST parameter delta Cn of at least 0.08 (37) to distinguish the top choices. In many cases there were multiple combinations of modifications at different positions consistent with the MS, MS/MS data, thus our set of modified sites likely underestimates the full population of modifications. Ambiguities in the site of modification are noted in the tables.

Analysis of modifications using InsPecT

Post-translational modifications were independently analyzed using the program InsPecT (28), version 20061212. This software did not allow a precursor ion mass tolerance in ppm, thus a tolerance of 0.008 Da (10 ppm mass accuracy for an 800 Da precursor), and the default product ion tolerance of 0.5 Da. were used. Both one and two modifications were allowed per peptide in separate runs; the inability to analyze three or more modifications per peptide may account in part for the lower number of modifications identified compared to SEQUEST. A maximal modification size of 95 was allowed, and the default p-value cutoff was 0.05. Since cleavage after glu and asp, or before asp could not be specified in the program, searches requiring no enzyme specificity were used, and peptides were selected that had terminal cleavages expected for the proteases used to digest H2B. Searches allowing peptide mass modifications greater than 95 Da were difficult to interpret since extra mass was often added to peptide termini which represented extensions of the peptide sequence in H2B. The InsPecT-calculated p-value reflects the probability that the match between the sequence and MS/MS spectrum is spurious; lower numbers are better. Trimethyl- and acetyl-lysines were manually distinguished by precursor mass accuracy after InsPecT assigned the sequence to the MS/MS spectrum.

Support vector machine learning analysis

SEQUEST results were analyzed using the support vector machine learning program GIST (35). Version 2.0.5 (http://microarray.cpmc.columbia.edu/gist/) was used with a quadratic kernel function, a diagonal factor of 0.03 and a constant of 10. A training set with LTQ-FT data was used for GIST analysis. Each peptide was associated with 13 parameters (38), including peptide protonated mass MH, peptide charge z, peptide ion current, the difference in mass between the predicted mass of the top ranked peptide and the experimentally observed peptide mass (dM), the SEQUEST parameters Xcorr, delta Cn, Sp, RSp, y and b ion match (27), the fraction of the total MS/MS spectrum ion current matched by peaks from the best-fit peptide, the fraction of the total MS/MS spectrum peaks matched by peaks from the best-fit peptide, the total number of MS/MS spectrum peaks, and the sequence homology between the top two best-fit peptides. Probabilities of the correct sequence assignment were calculated using GIST (35). In the context of analysis of purified histones, and since the mass of the FTICR-MS measured precursor ion is usually sufficient to derive the composition of modifications on a peptide (36), the peptide probability is useful distinguishing the top two choices resulting from SEQUEST analysis.

Results

FTICR-MS overview of histone H2B modifications

We wished to characterize posttranslational modifications of Neurospora histone H2B because our preliminary studies using AUT × AU 2D gels implicated modifications of this histone as particularly sensitive to mutation of the histone deacetylase gene *hda-1* (25). We used high resolution, preparative AUT × AU 2D gels to purify acid-soluble proteins extracted from wild type or *hda-1* lysates. The gels separate the core histones H2A, H2B, H3 and H4 due to differential binding of Triton X-100 (26). Histone species differing by charge-altering modifications (potentially acetylation, phosphorylation, deimination, deamidation) appear as spots trailing diagonally to the upper left. Figure 1 shows a comparison of Coomassie-stained 2D gels of histones purified from wild type (WT) and *hda-1* strains. While minor differences can be seen when comparing histones H4 and H3, the most pronounced differences occur with histone H2B. The diagonal ladder of spots a-e for H2B isolated from the *hda-1* strain was extended by one to two additional spots compared to the WT H2B ladder, suggesting the presence of additional modified forms of *hda-1* H2B.

To characterize differences between wild type and *hda-1* H2B, individual H2B 2D gel spots were electroeluted from fixed and stained gels and introduced into the mass spectrometer source using static nanospray. The masses of H2B and various modified forms were obtained at high resolution (R = 200,000). Figure 2 compares the FTICR mass spectrum of WT H2B spot a (the lowest spot in the H2B ladder) with the spectrum of the next higher spot in the WT H2B ladder (WT spot b), to illustrate mass shifts of similar peaks in the two spectra. No unmodified H2B, which would occur at a monoisotopic mass of 14700.97 Da, was observed. Each peak represents a carbon isotope cluster of a particular modified form (or isobaric forms). Both WT spot a and WT spot b spectra consist of ~3 clusters of about 7 peaks, which for WT spot a are centered on peaks 3, 9 and 15 in Figure 2.

Together ~20–25 modified forms are shown for each H2B spot, with the most highly modified forms representing a mass addition of over 400 Da. Individual peaks in the WT spot b clusters are shifted to higher mass by 42 Da compared to corresponding peaks in the WT spot a cluster. The common shift of 42 Da suggests that most peaks in spot b have this additional modification compared to peaks in spot a; the mass of 42 Da is consistent with addition of an acetyl group. An additional modification of 1 Da in spot a (for example from arg conversion to citrulline, or asn or gln deamidation) that is not present in spot b could give rise to a shift of 41 Da, and an additional modification of 1 Da in spot b not present in spot a would give a 43 Da shift. Amounts of wild type H2B in spot c were insufficient to allow acquisition of a good high resolution spectrum.

Monoisotopic masses calculated for each of the major Figure 2 peaks are shown in Table 1. Most peaks are shifted from unmodified H2B by integral multiples of 14 Da or one methyl group. For each peak the mass shifts may thus reflect the composition of methyl groups, but for some peaks the exact composition of methylated species (mono-, di-, or trimethyllysine, and/or mono- and dimethyl arginine) cannot be defined from mass alone. Similar results were obtained comparing *hda-1* H2B spots a and b, with peaks in spot b shifted by +42 Da compared to peaks in spot a, and with appearance of 20–25 peaks spaced by 14 Da (see

supplemental Fig. 2B and right half of Table 1). Comparison of *hda-1* H2B spots b and c gave similar results (supplemental Fig. 2C), while peaks in WT spot a and *hda-1* spot a were not shifted by 42 Da when compared (supplemental Fig. 2D).

LC/MS/MS analysis of N-terminal peptide methylation

To examine the location of methyl groups indicated by experiments illustrated in Figure 2, as well as the location of other modifications, electroeluted protein from all WT H2B spots (and, separately, protein from all *hda-1* spots) was combined and digested with gluC endoprotease, with a combination of aspN and gluC endoproteases, or with argC endoprotease, which can also cleave after lysine residues (39). Digests were then examined by capillary LC/MS/MS on an LTQ-FT instrument. Table 2 shows methylations identified on PPKPADKKPASKAPATASKAPE containing the N-terminus of H2B, and fragments of this peptide, found in wild type and *hda-1* H2B. Some sites of lysine acetylation were also found on methylated peptides. To provide an overall view of the extent of N-terminal modification, results are combined from the different proteolytic digestions. With use of an FTICR instrument to examine peptides, mass accuracy is sufficient to define the composition of most post-translational modifications in peptides derived from proteolytic digests of purified histones. The location in peptides of modified residues was then derived, when possible, from MS/MS spectra.

For wild type H2B, methylated N-terminal peptides could be divided into peptides with 0, 1 or 2 acetylated lysines. No unmodified PPKPADKKPASKAPATASKAPE was observed for either wild type or *hda-1* H2B. Lys 4 was identified separately with 1, 2 and 3 methyl groups, and lys 8 with 1 methyl group. For this N-terminal peptide, for each acetylation state (0, 1 and 2 acetyls), multiple levels of methylation were observed.

For *hda-1* H2B, PPKPADKKPASKAPATASKAPE with 1, 2 or 3 acetylated lys was observed, each containing from 1–4, 2–5 and 1–3 methyl groups respectively. Lys4 was identified separately with 1 and 2 methyl groups. We did not observe this peptide with no acetylated lys, possible due to lower relative levels of the unacetylated form of H2B (Fig. 1B, spot a) than for WT H2B spot a. Often multiple different positional combinations of methyl groups, among the 5 lysines in this peptide, were consistent with the MS/MS spectra. Acetylation was individually identified on lys 4 (see Table 6), lys 8, lys 13 and lys 20; acetylation on lys13 and lys20 were only identified on multiply-acetylated peptides, suggesting they may not be initial sites of acetylation in *hda-1* H2B.

LC/MS/MS analysis of globular domain and C-terminal lysine methylation

Proteolytic digests of wild type and *hda-1* H2B were examined for methylations not in the N-terminal tail, using both underivatized and propionylated H2B. Table 3 shows the identification of methylated peptides in this region. In wild type H2B methylated lysines included K45 or K54, K57, K97, K127, and K137 at the C-terminus. With the exception of trimethyl-lys 97, modifications were methyl or dimethyl groups, and occurred either alone or in combination with other lys or arg methyls, or citrulline. For HDA1-inactivated H2B, globular domain lysines were methylated or dimethylated at K54, K57, K90, K96, K97, and K119. Some of these occurred in combination with methylated arginine. Figure 3A shows a

graphic overview of lysine methylations (and arg modifications) in wild type and *hda-1* H2B. The identified methylations underestimate the total methylated residues since in some cases the MS/MS spectra cannot distinguish some (particularly adjacent) methylated sites, which are not indicated on the Figure. Overall methylation appears to differ between WT and *hda-1* H2B at 9 sites (Fig. 3A).

LC/MS/MS analysis of arginine modifications

Modifications of arginine in histones are of particular interest since arginine dimethylation may be involved in gene activation and repression, while arginine deimination to citrulline may lead to transcriptional repression (40). Citrulline has been observed in histones H3, H2A and H4 (41) but not to our knowledge in H2B (42). We thus examined different H2B peptide maps for the presence of modified arginine. LC/MS/MS analysis identified a several modified arginines in both wild type and *hda-1* H2B, shown in Table 4. In wild type H2B, monomethyl-arginine was identified at R103, while dimethyl-arginine was identified at R110. In *hda-1* H2B, monomethyl-arginine was identified at R103. Our LC/MS/MS analysis did not distinguish symmetric from asymmetric arginine dimethylation.

Citrulline, which results from deimination of arginine catalyzed by peptidyl arginine deiminases (43–45), is identified by a mass gain of 0.984 Da relative to arginine. SEQUEST-identified citrulline peptide precursors were manually checked to insure the monoisotopic peak was selected. Most of the arginines identified as modified appear to be only partially modified, as cognate unmodified peptides containing the site were also observed.

Analysis of lysine acetylation in histone H2B

We used LC/MS/MS to examine differences of acetylation between H2B from WT and *hda-1* H2B. The monoisotopic masses of acetyl- and trimethyl-lysine differ by 0.036 Da or 18 ppm for a 2 kDa peptide; this difference in precursor ions is readily detectable using an LTQ-FT (33). Identified acetylated lysines not in the N-terminal tail of H2B are shown in Table 5. Lysines 96 and 97 were partially acetylated in both H2B samples. In addition, acetylated lysines were identified at positions 54, 57 and 119 of *hda-1* H2B. An MS/MS spectrum of the *hda-1* H2B peptide PPKPADKKPASKATATASKAPE identifying both dimethylation of K4 and acetylation at lysines 8 and 13 is shown in figure 3B.

To further examine differences between WT and *hda-1* H2B, individual 2D gel spots similar to those in Figure 1 were electroeluted from two different batches of *Neurospora* in separate experiments and digested in solution with a combination of gluC and aspN endoproteases. The lowest three gel spots were analyzed from both wild type and *hda-1* H2B. Results from the two separate experiments are combined in Table 6. In both experiments numerous H2B peptides were identified (data not shown) but acetylated lysines were identified only in the N-terminal tail, in the peptides DKKPASKAPATASKAPE and DAGKKTAASG. Combining both experiments, wild type H2B had acetylated lysines at lys 8, 13, and 20, (and possibly at lys 9) while *hda-1* H2B had acetylated lysines at lys 8, 13, 20 and 29 (and possibly at lys 9). Generally, more acetylated sites were identified in the higher spots b and c than in spot a. Most sites were identified from peptides with one or two acetyl groups. Table

6 (bottom) summarizes additional N-terminal peptides from analyses of pooled spots in separate experiments. No additional sites of acetylated lysine were identified in wild type H2B, while acetylated lysines were identified at lys 4, 8, 9, 13 and 20 for *hda-1* H2B. A summary of all of the sites of acetylation identified in wild type and *hda-1* H2B is shown in Figure 3C. For spots a and b, peptides without lys acetylation were often observed for H2B from both strains, suggesting individual H2B lysines are only partially acetylated in these spots.

For H2B isolated from both strains, no unmodified PPKPADKKPASKAPATASKAPE was identified in experiments summarized in Table 3 and in the bottom of Table 6. This explains the observation in Figure 2 that no unmodified H2B was isolated from either strain. A summary of all modifications detected by SEQUEST is shown in Fig. 4B.

Analysis of histone H2B modifications using the blind search algorithm InsPecT

To extend the analysis of modifications in H2B, we analyzed the LC/MS/MS data using the program InsPecT (28). This algorithm searches for post-translational modifications of peptides without a prior specification of the modification mass, and has been used to examine heavily modified proteins such as lens crystallins (46). It may thus be useful for identification of modifications not specified in the usual SEQUEST searches. To distinguish between acetyl- and trimethyl-lysine, the modification giving the lowest precursor mass error was selected. Fewer modification sites were identified than with SEQUEST, and are summarized for H2B isolated from both strains in Table 7 and Figure 4A. Some of the modifications identified by SEQUEST were also observed with InsPecT, often at the same sites, including lys dimethylation, and lysine acetylation or trimethylation. Modifications not examined in standard SEQUEST searches, but identified by InsPecT, include methionine sulfone, asparagine and glutamine deamidation, and glu methyl esterification.

Discussion

We have performed the first detailed characterization by mass spectrometry of histone H2B from *Neurospora crassa*, a filamentous fungus important in the history of genetics and biochemistry and a model organism useful for genetic dissection of epigenetic phenomena such as DNA methylation (11). High resolution 2D acid-urea-triton × acid-urea gels show differential modification of wild type and *hda-1* H2B, identifying histone H2B as a substrate for this enzyme. Our results show *Neurospora* H2B has some novel features, and a number of modification differences between wild type and *hda-1* H2B are observed. We have identified a variety of post-translational modifications, including some likely due to solution chemistry (met oxidation, asn and gln deamidation), preparation artifacts (methyl esterification), or of unexplained origin (met sulfone formation). Known histone modifications such as phosphorylation (47–48), ADP-ribosylation (49), biotinylation (50) or methylation of the N-terminus of the histone (51) were not observed.

H2B methylation

We detected no unmodified H2B. Instead, the smallest modification was a single methyl group for both wild type and *hda-1* H2B. Each H2B 2D gel spot appears to represent an

ensemble of 20 or more modified forms, not counting isobaric species, with many peaks spaced by 14 Da, part of apparent methylation ladders. These ladders are organized further into 3 or more clusters of about 7 methylated species each. The centers of the methylation clusters are shifted by 98 Da, which could be due to addition of 7 methyl groups. If the highest mass 14 Da-shifted peaks were assigned entirely to methylation (Table 1), ca. 30 or more methyl groups could be simultaneously present in a small subset of H2B molecules. By comparison with the maximal number of methyl groups identified by LC/MS/MS in Fig. 3A (19 for WT H2B, 19 for *hda-1* H2B), more individual sites of methylation may remain to be identified. Since at the protein level it is more difficult to distinguish 3 methyls from an acetyl group, and in view of the LC/MS/MS results, some of the modified forms contain a mixture of methyl and acetyl groups. For individual gel spots, much of the heterogeneity in methylation appears associated with the N-terminal peptide PPKPADKKPASKAPATASKAPE, for which no unmodified specie was observed. The state of methylation of this peptide explains the observation of no unmodified wild type or *hda-1* H2B. Further heterogeneity in methylation is due to different methylated states of the peptide ASKLAAYNKKSTISSRE. Methylated sites in H2B include mono-, di- and trimethylated lysine, as well as mono- and dimethyl arginine, and glutamate methyl esters. Symmetric and asymmetric arg methylation were not distinguished in these experiments. Methylation may be important as transitions between mono-, di- and trimethyl-lysine may control dynamic processes such as transcription and DNA repair (52).

Of the 12 different sites of methylated lysine identified in wild type and *hda-1* H2B, 8 (lys 54, 57, 90, 96, 97, 119, 127, 137) are not in the N-terminal tail, roughly residues 2–32 (53) but instead are in the globular domain of H2B. As discussed below, some of these methylated lysine residues could be involved in DNA binding or histone subunit interactions and assembly. Interestingly, Xiong et al (54) detected N-terminal H2B methylation and dimethylation, and confirmed our observation of mono-, di- and trimethylation of lys 4.

Lysine acetylation

We initially examined modifications of *Neurospora* histones for changes correlated with absence of the HDA1 histone deacetylase. 2D gels show that the most noticeable changes in histones are in H2B, with the appearance of additional spots shifted to higher mass. Clusters of methylated H2B observed by FTICR-MS are shifted up by 42 Da in each higher gel spot, consistent with the shift being due to an additional acetylation for each higher spot. This is consistent with similar early observations from 1D AUT gels (55–56) run on HPLC-isolated protein fractions. However, acetylation of H2B lysines can be complex, as we have found 5 different acetyl-lysine sites for wild type H2B, and 11 for *hda-1* H2B. The high ratio of different acetylated sites to number of acetylated lysines in the 2D gel spots suggests many different individual acetylated species of H2B may exist in *Neurospora*. About half of the acetylated lysine sites are in the N-terminal tail while about half are in the globular region, where modifications may be associated with control of dynamics of histone-DNA interactions (6).

In different histone preparations, different sites of acetylation were observed for wild type and *hda-1* H2B, ranging from acetylation only in the N-terminal tail for the two main

experiments combined in Table 6, to acetylation including the globular domain and closer to the C terminus (Table 5). This range of results suggests uncontrolled variables affecting lysine acetylation, which could include differences in the physiological state of *Neurospora* cultures grown and harvested at different times, small differences in histone isolation procedures, proteolytic digestion procedures or levels of added histone deacetylase inhibitors, or other variables. This may explain the results of Xiong et al (54) who observed acetylation of WT lysines only in the N-terminal tail, at positions 8, 13, 20, 29 and 30. This variability makes identification of substrate site(s) for HDA1, expected to be lysines acetylated in *hda-1* H2B but not in wild type H2B, more difficult. Current candidates include lys 4, 9, 54, 57 and/or 119. Yeast HDA1 deacetylates histone H2B as well as histone H3, with deacetylation of H2B at lysine 16, which (based on sequence alignments in Figure 5) is most homologous to *Neurospora* H2B lys 24. We cannot rule out this lysine as a substrate for HDA1 since our sequence coverage did not include residues 24–25. We have found that methylations at 9 different sites differ between wild type and *hda-1* H2B. Three of these sites (K4, K57, K119) also differ in acetylation. The acetylation state of the substrate lysine(s) for HDA1 may be linked to the presence of modifications at other H2B sites, a possibility previously suggested (57).

Arginine modifications

We have observed three modified states of arginine in *Neurospora* H2B, including monomethyl and dimethyl arg, and arg deiminated to citrulline. Citrulline has been observed previously in histones H2A, H3 and H4 (42) but not to our knowledge in H2B. Ca^{2+}-dependent peptidyl arginine deiminase 4 catalyzes arginine deimination in histones H2A, H3 and H4 in human cells (42) and is thought to have a broad substrate sequence specificity (58). By alignment (59) of *Neurospora* H2B with an H2B of known structure (Table 8), all modified arginines appear to be in the globular domain of the protein rather than in the N-terminal tail. Previous sites of histone arg deimination have been in the N-terminal tails of histones H2A, H3 and H4 (42, 44), thus the location of a site of arg deimination in a histone globular domain may be novel. Arg methylation or dimethylation may prevent deimination by peptidyl arginine deiminases (44, 60), and arg deimination may antagonize arg methylation. Arginine deimination to citrulline may lead to transcriptional repression (40). A functional interaction between arg methylation and deimination in H2B is unclear, since at least so far, the modifications have not been found on the same arginines.

Arginine methylation may modulate intermolecular interactions, due to increased side chain bulk and loss of a hydrogen bond donor (61), and may be a mark for pluripotency (62). Histone arginine dimethylation may be involved in gene activation and repression (40). Three potential arginine methyltransferases, homologs of human PRMT1, 3, and 5, have been identified in the *Neurospora* genome (63) but to our knowledge have not been studied in detail. Homologs of these enzymes have however been studied in other organisms. The homolog of PRMT1 in the fungus *Aspergillis nidulans*, RmtA, is specific for histone H4 arg3 (64). PRMT1-catalyzed histone asymmetric arginine dimethylation is involved in gene activation (45), while PRMT5 mediates symmetrical dimethylation of arginine 3 on histone H2A and/or H4 tails (65), which is associated with gene repression (45). Interestingly, arginine methylation may be indirectly linked to DNA methylation, as PRMT5 and its

substrate MBD2 are recruited to CpG islands in a DNA methylation-dependent fashion *in vivo*, and the substrate histone H4 R3 is dimethylated at these foci (66). Fission yeast PRMT3 is a ribosomal protein that catalyzes the formation of asymmetric (type I) dimethylarginine (67), has the ribosomal protein rpS2 as a substrate in mouse cells (68), and is not currently known to modify histones. Thus it is currently unclear which enzyme catalyzes monomethylation of arginine in H2B.

Mapping modified residues onto the structure of histone H2B in the *Xenopus* nucleosome core particle

The existence of new modifications at a number *Neurospora* H2B sites (Figure 5) raises the question of their potential function. Cosgrove et al.(6) proposed that single modifications of histone residues on the surface of the nucleosome modulate nucleosome dynamics by changing histone-DNA interactions, affecting DNA mobility and perhaps the exchange of histone variants into a nucleosome core particle (NCP). To examine this possibility for *Neurospora* H2B, we mapped the post-translationally modified lys and arg residues to residues in the highly homologous *Xenopus* H2B. Residues in the N-terminal tail were not considered due to the significant sequence differences in this region. The sequence of *Neurospora* H2B was aligned (59) in Table 8 with that of the *Xenopus* H2B used for the crystal structure of the NCP at 1.9A (53). There appear to be at least three classes of modified residues in *Neurospora* H2B.

Assuming (due to the high sequence similarity in this region) that the globular core structure of the *Neurospora* NCP is similar to that of the *Xenopus* NCP, the location of the modified *Neurospora* H2B residues close to DNA (Figure 6A) suggests they may be involved in indirect (K54, K57, R68, K96) or direct (K97) DNA binding. HDA1 deacetylation of acetyl-lys 54 or 57 could thus tighten DNA binding to nucleosomes and perhaps decrease DNA mobility in NCPs. In contrast, R68 deimination to citrulline, by removing the side chain charge, could weaken DNA-histone interactions, providing a different type of mechanism for regulating DNA mobility in a NCP.

A second class of locations for modified residues includes H2B arg and lys at or near subunit interfaces in the NCP structure. Lys 54 is near the H2A-H2B interface and appears to bind directly to H2A leu 99. *Neurospora* H2B arg 103 lies at the H2B-H4 interface, where it binds to H4 his 75; methylation of R103 could change H2B-H4 subunit interactions. Modification of lys 90, buried within the core of H2B, could affect the assembly and structure of H2B.

A third class of locations for modified H2B residues includes those exposed at or near the surface of the NCP, including lys 54, lys 57, arg 103, arg 110, lys 119 and lys 127 (Figure 6B). Modification of these residues could affect the binding of chromatin remodeling or other complexes to the NCP, NCP-NCP interactions, and/or formation of higher order DNA structures. The existence of multiple modified states of surface H2B residues suggests regulation of these interactions.

Comparison with histone H2B from other organisms

Post-translational modifications of a number of histone H2Bs have been reported (5, 6, 51, 69–72), allowing their comparison to *Neurospora* histone H2B (Figure 5, wild type and *hda-1* H2B are combined). Sequences were aligned using T-coffee (59). New (to H2B) modifications discovered here in *Neurospora* H2B include arginine deiminated to citrulline, arg methylation and dimethylation, trimethyl lysines, a more extensive set of modifications in the globular core, a C-terminal methyl group, and overall more extensive methylation. Included are additional modifications reported by Xiong et al (54) including N-terminal methylation and dimethylation, and lys 30 acetylation. More detailed study of the non-*Neurospora* H2B's could in the future find additional modifications. A number of H2Bs are ubiquitylated near the C-terminus; we did not examine *Neurospora* H2B for this modification. Tree frog histone H2B has been reported as being unmodified (73). Tetrahymena H2B has been reported to have its N-terminal alanine mono-, di- and trimethylated and K3 and K4 acetylated and trimethylated (74). Garcia et al. (75) have used FTICR-mass spectrometry to obtain complex envelopes of modified forms of the histone H3.2 variant from HeLa cells. However their broadband spectrum of H2B from asynchronous HeLa S3 cells contained mainly unmodified H2B variants and a total of 5 isotopic clusters, suggesting far less complexity of modified forms than for H3.2 (17).

Trimethyllysine was identified at 4 different lysines in *Neurospora* H2B, but to our knowledge has not been reported in yeast, calf , human or *Arabidopsis* histone H2B. Nine different sites of dimethyllysine were observed in *Neurospora* H2B, while only single sites were annotated in calf (5,69), human(71), and *Arabidopsis* H2B (51); none were reported in yeast H2B (70). After sequence alignment, novel methylation sites in *Neurospora* were seen at lys 90 and at the C-terminus at lys137.

All of the histone H2B's can be acetylated at a number of different sites: 11 for *Neurospora* H2B, 6 for yeast and human H2B, 8 for calf H2B and 7 for *Arabidopsis* H2B. From the 2D gel in Figure 1 and the 42 Da mass shifts between spots, we surmise that a subpopulation of *Neurospora hda-1* H2B may have as many as 5–6 simultaneous acetylated lysines, or as many as 3–4 in wild type H2B. Combined with the 11 different observed acetylation sites in *Neurospora* H2B, many combinations of individual acetylated species of *Neurospora* H2B are possible. Other organisms can generate H2B sequence variants using multiple genes. Possessing only a single histone H2B gene, *Neurospora* may achieve high effective sequence diversity with a more extensive use of post-translational modifications. The N-terminal H2B tails are asymmetric in the *Xenopus* NCP crystal structure (53), with one tail extending into solution and the second binding between the two strands of DNA wrapped around the histone octamer, creating (with multiple N-terminal tail modifications) a mechanism for generating even more diversity at the structural level. Given the different structures of the two NCP H2B N-terminal tails, it is possible that modifications in one tail are different from those in the other tail in a single NCP, creating even greater diversity of NCP structural states. Further diversity could be created by differing DNA sequences bound in individual NCPs.

Other modifications

Met 59 and met 62 of human H2B isolated from nickel-treated cells have been observed as sulfoxides, and a single deamidated gln was also observed (76). The *in vivo* significance of met oxidation to met sulfoxide is unclear, since this can occur in solution as well as *in vivo* (77). Identification of met sulfone in a histone peptide is unexpected since this should require prior exposure to a strong oxidizing agent. The *in vivo* significance of methyl esterification and deamidation is also unclear. The 2D gels were destained in methanol, and glutamate methyl esterification has been observed when staining gels with Coomassie blue in the presence of acid and methanol (78). Deamidation of both asn and gln can occur spontaneously in solution (79). An algorithm capable of simultaneously searching a large number of different modifications on a peptide, as well as allowing the presence of multiple simultaneous modifications, will be advantageous for more in-depth analysis of highly modified histone peptides.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Stephen Tanner (UCSD) for advice on the use of InsPecT and for a number of requested bug fixes.

Funding Support: This work was supported by the National Institutes of Health (GM-35690) to EUS, and an Oregon Medical Research Foundation grant to DCA.

References

1. McGhee JD, Felsenfeld G. Nucleosome structure. Ann Rev Biochem. 1980; 49:1115–1156. [PubMed: 6996562]
2. Strahl B, Allis C. The language of covalent histone modifications. Nature. 2000; 403:41–45. [PubMed: 10638745]
3. Green GR. Phosphorylation of histone variant regions in chromatin: Unlocking the linker? Biochemistry and Cell Biology. 2001; 79:275–287. [PubMed: 11467741]
4. Fischle W, Wang Y, Allis CD. Binary switches and modification cassettes in histone biology and beyond. Nature. 2003; 425:475–479. [PubMed: 14523437]
5. Zhang L, Eugeni E, Parthun MR, Freitas M. Identification of novel histone post-translational modifications by peptide mass fingerprinting. Chromosoma. 2003; 112:77–86. [PubMed: 12937907]
6. Cosgrove M, Boeke J, Wolberger C. Regulated nucleosome mobility and the histone code. Nature Struct Mol Biol. 2004; 11:1037–1043. [PubMed: 15523479]
7. Garcia B, Shabanowitz J, Hunt D. Characterization of histones and their post-translational modifications by mass spectrometry. Curr Opin Chem Biol. 2007; 11:66–73. [PubMed: 17157550]
8. Mikesh L, Ueberheide B, Chi A, Coon J, Syka J, Shabanowitz J, Hunt D. The utility of ETD mass spectrometry in proteomic analysis. Biochim Biophys Acta. 2006; 1764:1811–1822. [PubMed: 17118725]
9. Coon J, Ueberheide B, Syka J, Dryhurst D, Ausio J, Shabanowitz J, Hunt D. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. Proc Natl Acad Sci U S A. 2005; 102:9463–9468. [PubMed: 15983376]
10. Richards E, Elgin S. Epigenetic Codes for Heterochromatin Formation and Silencing: Rounding up the Usual Suspects. Cell. 2002; 108:489–500. [PubMed: 11909520]

11. Freitag M, Selker E. Controlling DNA methylation: many roads to one modification. Current Opinion in Genetics & Development. 2005; 15:191–199. [PubMed: 15797202]
12. Jones P. DNA methylation and cancer. Oncogene. 2002; 21:5358–5360. [PubMed: 12154398]
13. Esteller M. DNA methylation and cancer therapy: New developments and expectations. Curr Opin Oncol. 2005; 17:55–60. [PubMed: 15608514]
14. Szyf M, Pakneshan P, Rabbani S. DNA demethylation and cancer. Cancer Lett. 2004; 211:133–143. [PubMed: 15219937]
15. Hays S, Swanson J, Selker E. Identification and Characterization of the Genes Encoding the Core Histones and Histone Variants of *Neurospora crassa*. Genetics. 2001; 160:961–973. [PubMed: 11901114]
16. Bonenfant D, Coulot M, Towbin H, Schindler P, van Oostrum J. Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. Mol Cell Proteomics. 2006; 5:541–552. [PubMed: 16319397]
17. Siuti N, Roth M, Mizzen C, Kelleher N, Pesavento J. Gene-specific characterization of human histone H2B by electron capture dissociation. J Proteome Res. 2006; 5:233–239. [PubMed: 16457587]
18. Foss H, Roberts C, Selker EU. Mutations in the dim-1 gene of Neurospora crassa reduce the level of DNA methylation. Mol Gen Genet. 1998; 259:60–71. [PubMed: 9738881]
19. Kouzminova E, Selker EU. dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in Neurospora. EMBO J. 2001; 20:4309–4323. [PubMed: 11483533]
20. Tamaru H, Selker EU. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. Nature. 2001; 414:277–283. [PubMed: 11713521]
21. Freitag M, Hickey P, Khlafallah T, Read N, Selker E. HP1 is essential for DNA methylation in neurospora. Mol Cell. 2004; 13:427–434. [PubMed: 14967149]
22. Adhvaryu KK, Selker EU. Protein phosphatase PP1 is required for normal DNA methylation in Neurospora. Genes Dev. 2008; 22:3391–3396. [PubMed: 19141471]
23. Lewis Z, Adhvaryu K, Honda S, Shriver A, Selker EU. Identification of DIM-7, a protein required to target the DIM-5 H3 methyltransferase to chromatin. Proc Natl Acad Sci USA. 2010 Apr 19.
24. Selker EU. Trichostatin A causes selective loss of DNA methylation in *Neurospora*. Proc Ntl Acad Sci USA. 1998; 95:9430–9435.
25. Smith K, Dobosy J, Do D, Riefsnyder J, Anderson D, Green G, Selker E. in preparation.
26. Green GR, Do DP. Purification and analysis of variant and modified histones using 2D PAGE. Methods Molecular Biology. 2008; 464:285–302.
27. Eng J, McCormack A, Yates J. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom. 1994; 5:976–989. [PubMed: 24226387]
28. Tanner S, Shu H, Frank A, Wang L, Zandi E, Mumby M, Pevzner P, Bafna V. InsPecT: Identification of Posttranslationally Modified Peptides from Tandem Mass Spectra. Anal Chem. 2005; 77:4626–4639. [PubMed: 16013882]
29. Wyatt H, Liaw H, Green GR, Lustig AJ. Multiple roles for *Saccharomyces cerevisiae* histone H2A in telomere position effect, *spt* S sppression and DSB repair. Genetics. 2003; 164:47–64. [PubMed: 12750320]
30. Green GR, Poccia DL, Herlands L. A multisample device for electroelution, concentration and dialysis of proteins from fixed and stained gel slices. Analytical Biochemistry. 1982; 123:66–73. [PubMed: 7051898]
31. Sorensen S, Sorensen T, Breddam K. Fragmentation of proteins by S. aureus strain V8 protease. FEBS Lett. 1991; 294:195–197. [PubMed: 1684551]
32. Garcia B, Busby S, Shabanowitz J, Hunt D, Mishra N. Resetting the epigenetic histone code in the MRL-lpr/lpr mouse model of lupus by histone deacetylase inhibition. J Proteome Res. 2005; 4:2032–2042. [PubMed: 16335948]
33. Syka J, Marto J, Bai D, Horning S, Senko M, Schwartz J, Ueberheide B, Garcia B, Busby S, Muratore T, Shabanowitz J, Hunt D. Novel linear quadrupole ion trap/FT mass spectrometer:

performance characterization and use in the comparative analysis of histone H3 post-translational modifications. J Proteome Res. 2004; 3:621–626. [PubMed: 15253445]

34. Horn D, Zubarev R, McLafferty F. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. J Am Soc Mass Spectrom. 2000; 11:320–332. [PubMed: 10757168]
35. Pavlidis P, Wapinski I, Noble W. Support vector machine classification on the web. Bioinformatics. 2004; 20:586–587. [PubMed: 14990457]
36. Freitas M, Sklenar A, Parthun M. Application of mass spectrometry to the identification and quantification of histone post-translational modifications. J Cellular Biochem. 2004; 92:691–700. [PubMed: 15211567]
37. Washburn M, Wolters D, Yates J. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. Nat Biotech. 2001; 19:242–247.
38. Anderson D, Li W, Payan D, Noble W. A new algorithm for the evaluation of shotgun peptide sequencing in proteomics: support vector machine classification of peptide MS/MS spectra and SEQUEST scores. J Proteome Res. 2003; 2:137–146. [PubMed: 12716127]
39. Mitchell W, Harrington W. Purification and properties of clostridiopeptidase B (clostripain). J Biol Chem. 1968; 243:4683–4692. [PubMed: 4971659]
40. Wysocka J, Allis C, Coonrod S. Histone arginine methylation and its dynamic regulation. Front Biosci. 2006; 11:344–355. [PubMed: 16146736]
41. Nakashima K, Hagiwara T, Yamada M. Nuclear localization of peptidylarginine deiminase V and histone deimination in granulocytes. J Biol Chem. 2002; 277:49562–49568. [PubMed: 12393868]
42. Hagiwara T, Hidaka Y, Yamada M. Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes. Biochemistry. 2005; 44:5827–5834. [PubMed: 15823041]
43. Fujisaki M, Sugawara K. Properties of peptidylarginine deiminase from the epidermis of newborn rats. J Biochem (Tokyo). 1981; 89:257–263. [PubMed: 7217033]
44. Cuthbert G, Daujat S, Snowden A, Erdjument-Bromage H, Hagiwara T, Yamada M, Schneider R, Gregory P, Tempst P, Bannister A, Kouzarides T. Histone deimination antagonizes arginine methylation. Cell. 2004; 118:545–553. [PubMed: 15339660]
45. Thompson P, Fast W. Histone citrullination by protein arginine deiminase: is arginine methylation a green light or a roadblock? ACS Chem Biol. 2006; 1:433–441. [PubMed: 17168521]
46. Wilmarth P, Tanner S, Dasari S, Nagalla S, Riviere M, Bafna V, Pevzner P, David L. Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility? J Proteome Res. 2006; 5:2554–2566. [PubMed: 17022627]
47. Ajiro K. Histone H2B phosphorylation in mammalian apoptotic cells. An association with DNA fragmentation. J Biol Chem. 2000; 275:439–443. [PubMed: 10617636]
48. Cheung W, Ajiro K, Samejima K, Kloc M, Cheung P, Mizzen C, Beeser A, Etkin L, Chernoff J, Earnshaw W, Allis CD. Apoptotic phosphorylation of histone H2B is mediated by mammalian sterile twenty kinase. Cell. 2003; 113:507–517. [PubMed: 12757711]
49. Burzio L, Riquelme P, Koide S. ADP ribosylation of rat liver nucleosomal core histones. J Biol Chem. 1979; 254:3029–3037. [PubMed: 218926]
50. Hassan Y, Zempleni J. Epigenetic regulation of chromatin structure and gene function by biotin. J Nutr. 2006; 136:1763–1765. [PubMed: 16772434]
51. Bergmüller E, Gehrig P, Gruissem W. Characterization of post-translational modifications of histone H2B-variants isolated from Arabidopsis thaliana. J Proteome Res. 2007; 6:3655–3668. [PubMed: 17691833]
52. Shahbazian M, Zhang K, Grunstein M. Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. Mol Cell. 2005; 19:271–277. [PubMed: 16039595]
53. Davey C, Sargent D, Luger K, Maeder A, Richmond T. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution. J Mol Biol. 2002; 319:1097–1113. [PubMed: 12079350]
54. Xiong L, Adhvaryu K, Selker EU, Wang Y. Mapping of Lysine Methylation and Acetylation in Core Histones of *Neurospora crassa*. Biochemistry. 2010 in press.

55. Waterborg J, Robertson A, Tatar D, Borza C, Davie J. Histones of Chlamydomonas reinhardtii. Synthesis, acetylation and methylation. Plant Physiol. 1995; 109:393–407. [PubMed: 7480339]
56. Waterborg J. Steady-state Levels of Histone Acetylation in *Saccharomyces cerevisiae*. J Biol Chem. 2000; 275:13007–13011. [PubMed: 10777603]
57. Robyr D, Suka Y, Xenarios I, Kurdistani SK, Wang A, Suka N, Grunstein M. Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. Cell. 2002; 109:437–446. [PubMed: 12086601]
58. Arita K, Shimizu T, Hashimoto H, Hidaka Y, Yamada M, Sato M. Structural basis for histone N-terminal recognition by human peptidylarginine deiminase 4. Proc Natl Acad Sci U S A. 2006; 103:5291–5296. [PubMed: 16567635]
59. Notredame C, Higgins D, Heringa J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol. 2000; 302:205–217. [PubMed: 10964570]
60. Raijmakers R, Zendman AJ, Egberts W, Vossenaar E, Raats J, Soede-Huijbregts C, Rutjes F, van Veelen P, Drijfhout J, Pruijn G. Methylation of arginine residues interferes with citrullination by peptidylarginine deiminases in vitro. J Mol Biol. 2007; 367:1118–1129. [PubMed: 17303166]
61. McBride A, Silver P. State of the arg: protein methylation at arginine comes of age. Cell. 2001; 106:5–8. [PubMed: 11461695]
62. Torres-Padilla M, Parfitt D, Kouzarides T, Zernicka-Goetz M. Histone arginine methylation regulates pluripotency in the early mouse embryo. Nature. 2007; 445:214–218. [PubMed: 17215844]
63. Borkovich K, et al. Lessons from the Genome Sequence of *Neurospora crassa*: Tracing the Path from Genomic Blueprint to Multicellular Organism. Microbiology and Molecular Biology Reviews. 2004; 68:1–108. [PubMed: 15007097]
64. Trojer P, Dangl M, Bauer I, Graessle S, Loidl P, Brosch G. Histone methyltransferases in *Aspergillus nidulans*: evidence for a novel enzyme with a unique substrate specificity. Biochemistry. 2004; 43:10834–10843. [PubMed: 15311944]
65. Ancelin K, Lange U, Hajkova P, Schneider R, Bannister AJ, Kouzarides T, Surani M. Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol. 2006; 8:623–630. [PubMed: 16699504]
66. Le Guezennec X, Vermeulen M, Brinkman A, Hoeijmakers W, Cohen A, Lasonder E, Stunnenberg H. MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Mol Cell Biol. 2006; 26:843–851. [PubMed: 16428440]
67. Bachand F, Silver P. PRMT3 is a ribosomal protein methyltransferase that affects the cellular levels of ribosomal subunits. EMBO J. 2004; 23:2641–2650. [PubMed: 15175657]
68. Swiercz R, Cheng D, Kim D, Bedford M. Ribosomal protein rpS2 is hypomethylated in PRMT3-deficient mice. J Biol Chem. 2007; 282:16917–16923. [PubMed: 17439947]
69. Thorne A, Sautiere P, Briand G, Crane-Robinson C. The structure of ubiquitinated histone H2B. EMBO J. 1987; 6:1005–1010. [PubMed: 3036486]
70. Parra M, Kerr D, Fahy D, Pouchnik D, Wyrick J. Deciphering the roles of the histone H2B N-terminal domain in genome-wide transcription. Mol Cell Biol. 2006; 26:3842–3852. [PubMed: 16648479]
71. Beck HC, Nielsen E, Matthiesen R, Jensen L, Sehested M, Finn P, Grauslund M, Hansen A, Jensen O. Quantitative proteomic analysis of post-translational modifications of human histones. Mol Cell Proteomics. 2006; 5:1314–1325. [PubMed: 16627869]
72. Ahn S, Diaz R, Grunstein M, Allis C. Histone H2B deacetylation at lysine 11 is required for yeast apoptosis induced by phosphorylation of H2B at serine 10. Mol Cell. 2006; 24:211–220. [PubMed: 17052455]
73. Kawasaki H, Isaacson T, Iwamuro S, Conlon J. A protein with antimicrobial activity in the skin of Schlegel's green tree frog Rhacophorus schlegelii (Rhacophoridae) identified as histone H2B. Biochem Biophys Res Commun. 2003; 312:1082–1086. [PubMed: 14651982]
74. Medzihradszky K, Zhang X, Chalkley RJ, Guan S, McFarland MA, Chalmers MJ, Marshall AG, Diaz RL, Allis CD, Burlingame AL. Characterization of *Tetrahymena* Histone H2B Variants and Posttranslational Populations by Electron Capture Dissociation (ECD) Fourier Transform Ion

Cyclotron Mass Spectrometry (FT-ICR MS). Mol Cell Proteomics. 2004; 3:872–876. [PubMed: 15199121]

75. Garcia B, Pesavento J, Mizzen C, Kelleher N. Pervasive combinatorial modification of histone H3 in human cells. Nature Methods. 2007; 4:487–489. [PubMed: 17529979]
76. Karaczyn A, Golebiowski F, Kasprzak K. Truncation, deamidation, and oxidation of histone H2B in cells cultured with nickel(II). Chem Res Toxicol. 2005; 18:1934–1942. [PubMed: 16359184]
77. Hoshi T, Heinemann S. Regulation of cell function by methionine oxidation and reduction. J Phyisiology. 2001; 531:1–11.
78. Haebel S, Albrecht T, Sparbier K, Walden P, Komer R, Steup M. Electrophoresis-related protein modification: alkylation of carboxy residues revealed by mass spectrometry. Electrophoresis. 1998; 19:679–686. [PubMed: 9629898]
79. Peters B, Trout B. Asparagine deamidation: pH-dependent mechanism from density theory. Biochemistry. 2006; 45:5384–5392. [PubMed: 16618128]

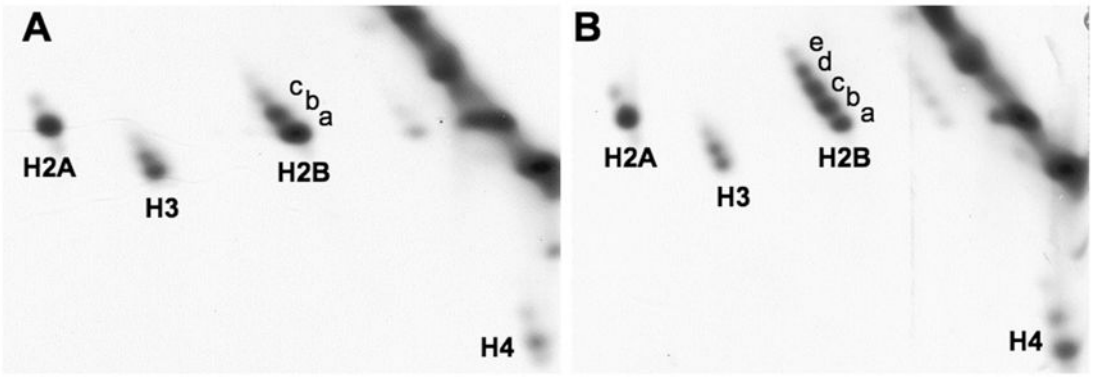

Figure 1. High resolution two-dimensional polyacrylamide gel electrophoresis of histones isolated from wild type and *hda-1* strains of *Neurospora crassa*

Histones from wild type (**A**) and *hda-1* (**B**) strains of *Neurospora crassa* were resolved in acetic acid-urea-Triton X-100 (AUT) in the horizontal dimension and acetic acid-urea (AU) in the vertical dimension, with the direction of protein migration indicated by the arrows. Proteins were stained with Coomassie blue R-250. All four core histones (H2A, H2B, H3, H4) were observed. Wild type H2B appears as a diagonal ladder of 3 labeled spots with a fourth faint spot, while *hda-1* H2B appears as a diagonal ladder of five spots.

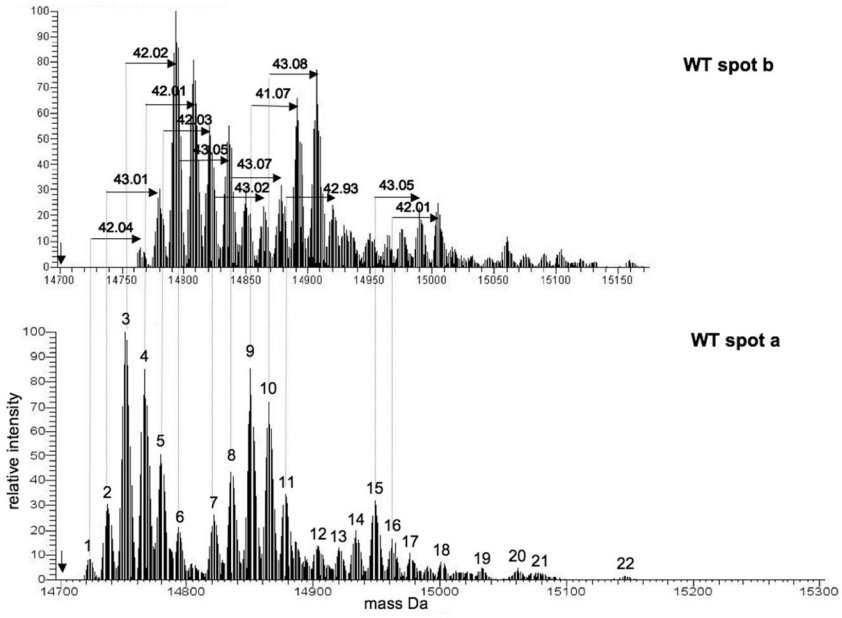

Figure 2. FTICR mass spectrometry of electroeluted histone H2B 2D gel spots

The monoisotopic mass of unmodified H2B is 14700.97 Da (vertical arrow). Each single spectral peak represents an isotope cluster for one or more modified forms of histone H2B. Listed mass shifts are based on monoisotopic masses calculated by THRASH (34), which are aligned with the most abundant isotopic peak of each isotope cluster for clarity. The spectra are aligned to allow comparison of clusters of peaks. The high resolution mass spectrum is shown for the lowest (WT spot a) and next upper (WT spot b) wild type histone H2B spots. Mass shifts for each peak are in Table 1. The y-axis represents relative abundance.

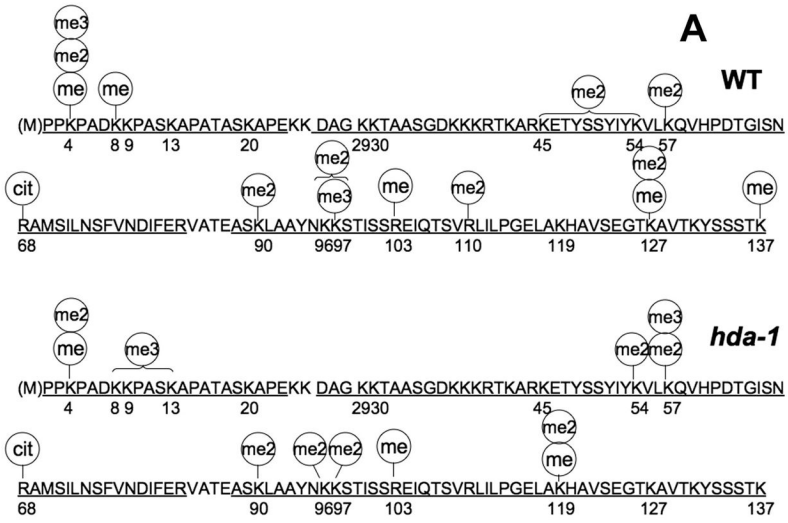

Figure 3A

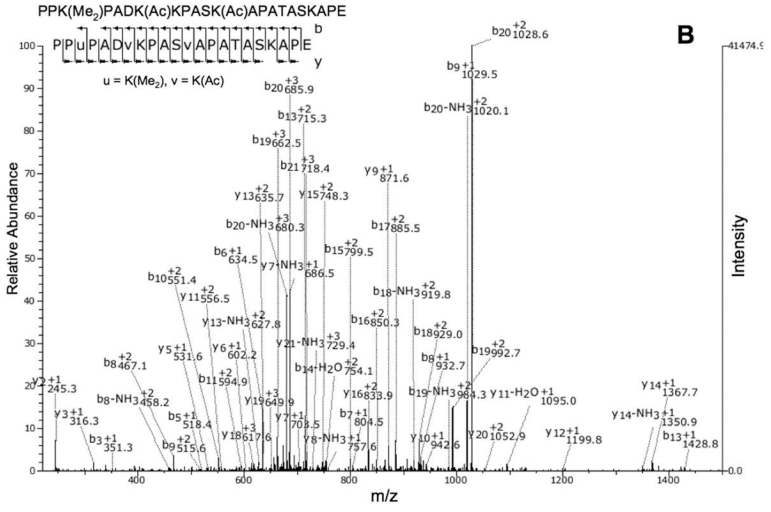

Figure 3B

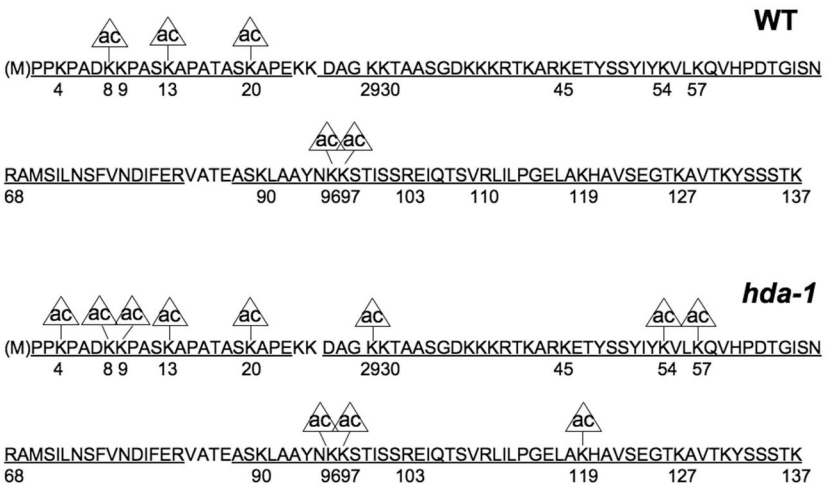

Figure 3C

Figure 3. Post-translational modifications identified on *Neurospora* histone H2B

A. Lysine and arginine modifications identified in *Neurospora* histone H2B. Modifications are indicated by circles above the modified H2B residue (me, methyl; me2, dimethyl; me3, trimethyl; ac, acetyl, cit, arg modified to citrulline). The top sequence indicates modification sites in wild type (WT) H2B; the lower sequence shows modification sites in *hda-1* H2B. In some cases MS/MS spectra could not distinguish which of several alternative lysines are methylated; several of these sites are indicated with brackets spanning the potential modified lysines. The sequence coverage in combined LC/MS/MS experiments is indicated by underlined residues. **B**. MS/MS spectrum identifying the doubly acetylated, doubly methylated *hda-1* H2B N-terminal peptide PPK(Me_2)PADK(Ac)KPASK(Ac)APATASKAPE using SEQUEST **C**. Summary of lysine acetylation in wild type and *hda-1* H2B. Acetylation at each of the sites is partial since other modifications (or no modification) were identified at each of the indicated sites. The sequence coverage from individual peptides is underlined. Acetylated sites in *hda-1* H2B but not in WT H2B are potential substrate sites for the HDA1 histone deacetylase.

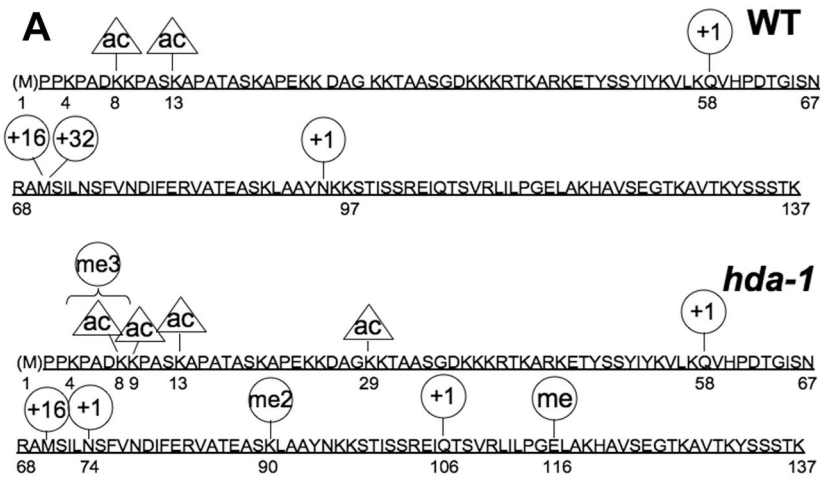

Figure 4A

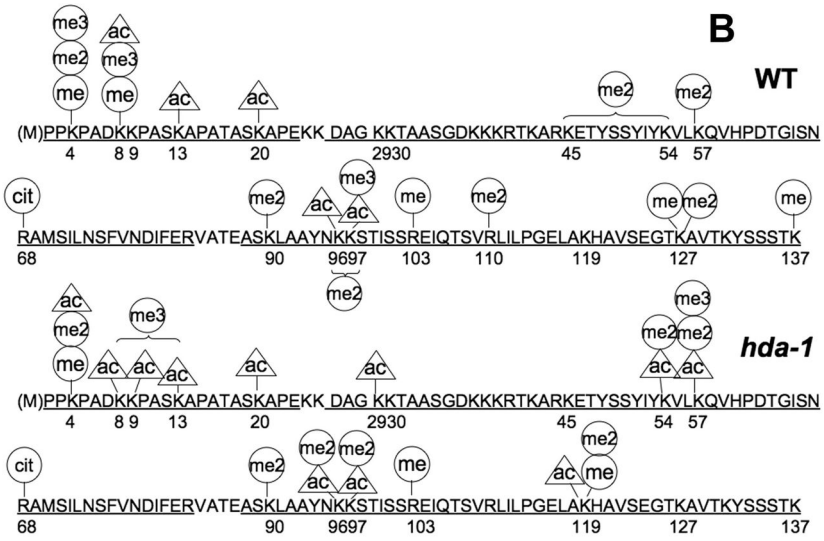

Figure 4B

Figure 4. Comparison of *Neurospora* histone H2B post-translational modifications identified by two algorithms

A. Modifications identified by the blind search algorithm InsPecT (28). InsPecT identified several modifications not observed in standard SEQUEST searches, including deamidated asn or gln (circle with +1), met sulfone (circle with +32), and a glutamate methyl ester (me). Other modifications were oxidized met (circle with +16), dimethyl lys (me2), trimethyl lys (me3), and acetyl-lys (ac). Sequence coverage of identified peptides is indicated by underlined residues. Acetyl- and trimethyllysine were distinguished by precursor ion mass.

B. Summary of all modifications identified in wild type and HDA1-inactivated H2B using SEQUEST, indicated by circles over modified residues: methyl (me), dimethyl (me2), trimethyl (me3), acetyl (ac), and arg modified to citrulline (cit).

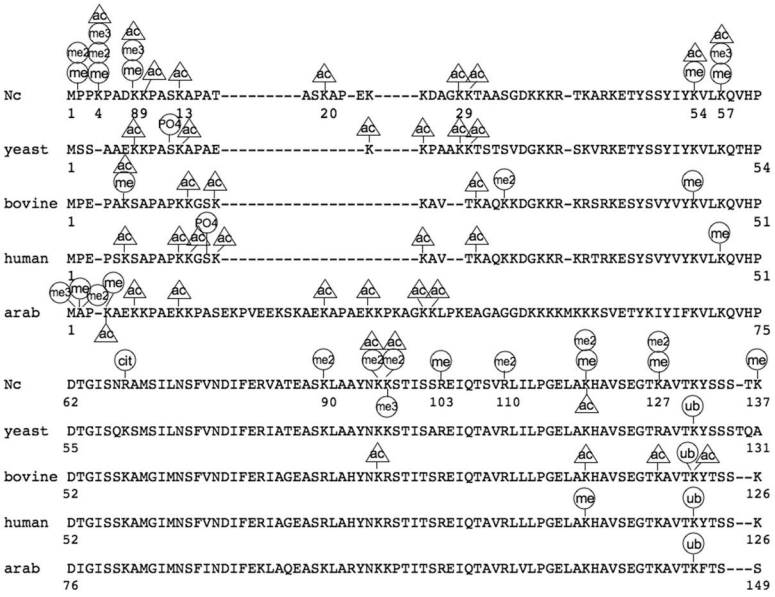

Figure 5. Comparison of *Neurospora* histone H2B post-translational modifications with those of histone H2B from other organisms

H2B sequences from *Neurospora*, yeast (*Saccharomyces cerevisiae*), calf (*Bos taurus*), *Arabidopsis* variant HTB11, and human histone H2B (locus CAB02545) were aligned using T-coffee (59). Numbering starts with the N-terminal met although this is not present in the mature histone. Modifications from wild type and HDA1-inactivated H2B were combined for the *Neurospora* map and include 3 additional modifications reported by Xiong et al (54). The arrow under the *Neurospora* H2B sequence indicates the start of the globular domain.

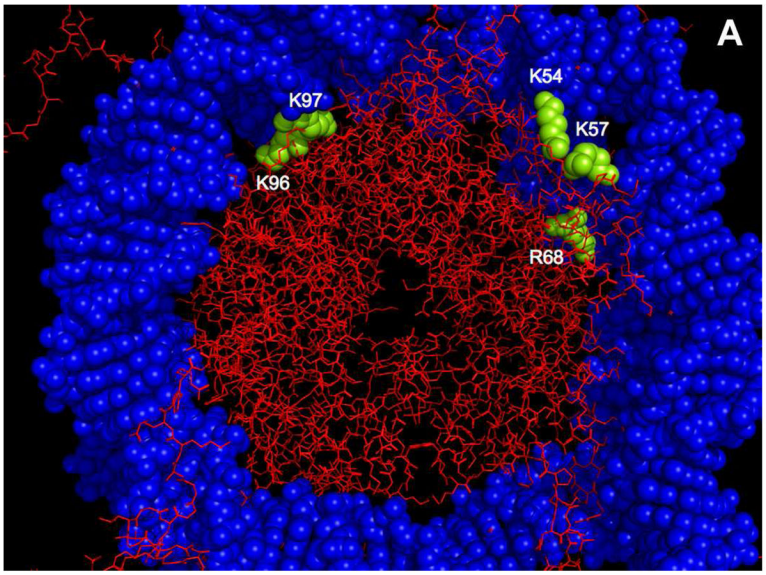

Figure 6A

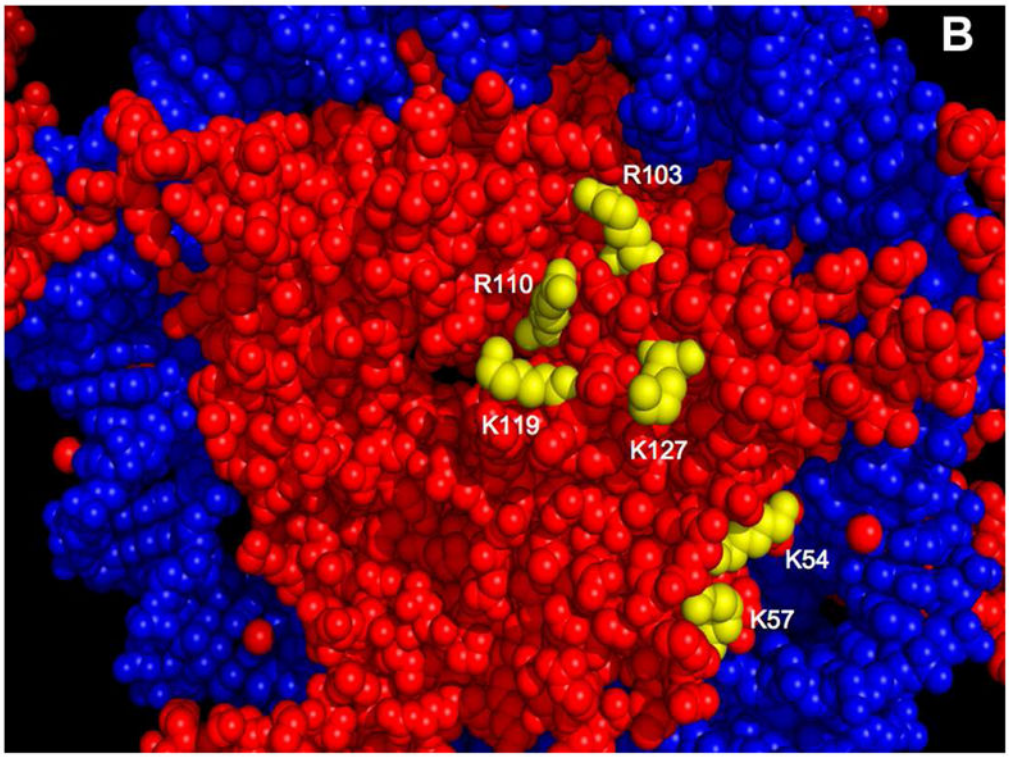

Figure 6B

Figure 6. Predicted location of post-translationally modified *Neurospora* histone H2B residues in the crystal structure of the *Xenopus* nucleosome core particle

Positions of the post-translationally modified *Neurospora* H2B globular domain residues in the structure of the *Xenopus* nucleosome core particle (PDB1KX5) were derived by sequence alignment (59) with *Xenopus* H2B (Table 8). Residues on only one face of the nucleosome core particle structure are shown; similar residues are also on the opposite face.

Histone polypeptides are red; modified residues are yellow, and the two strands of DNA wrapped around the histone octamer are blue. **A.** H2B modified residues predicted to be close to DNA and which may directly or indirectly bind the DNA wrapped around the core histone octamer. K96, K97 and R68 are buried in the histone octamer structure while K54 and K57 are both at the surface of the NCP. **B.** H2B residues predicted to be on or near the surface of the nucleosome core particle (NCP). Numbers are shown for *Neurospora* H2B residues.

Table 1

Mass Shifts of monoisotopic peaks in histone H2B spectra[1]

Shifts in the WT spot a spectrum of Fig. 2A				Shifts in the *hda-1* spot a spectrum of Fig. 2B			
Peak	**Mass, MH^+**	**Shift Da**	**Methyl group equivalents**	**Peak**	**Mass, MH^+**	**Shift Da**	**Methyl group equivalents**
	14701.0	0	peak not observed		14701.0	0	peak not observed
1	14715.0	14.0	1 methyl	1	14715.0	14.0	1 methyl
2	14729.1	28.1	2 methyls	2	14729.0	28.0	2 methyls
3	14743.0	42.0	3 methyls	3	14743.1	42.1	3 methyls
4	14758.1	57.1	4 methyls +1	4	14758.1	57.1	4 methyls +1
5	14771.1	70.1	5 methyls	5	14771.1	70.1	5 methyls
6	14785.0	84.0	6 methyls	6	14785.1	84.1	6 methyls
7	14813.1	112.1	8 methyls	7	14800.1	99.1	7 methyls +1
8	14827.1	126.1	9 methyls	8	14812.1	111.1	8 methyls −1
9	14842.1	141.1	10 methyls +1	9	14827.0	126.0	9 methyls
10	14856.1	155.1	11 methyls	10	14841.1	140.1	10 methyls
11	14869.1	168.1	12 methyls	11	14856.1	155.1	11 methyls
12	14895.1	194.1	14 methyls −3	12	14868.1	167.1	12 methyls −2
13	14912.0	211.0	15 methyls	13	14882.1	181.1	13 methyls −2
14	14925.1	224.1	16 methyls −1	14	14898.0	197.0	14 methyls
15	14940.1	239.1	17 methyls	15	14911.1	210.1	15 methyls −1
16	14954.1	253.1	18 methyls	16	14925.1	224.1	16 methyls −1
17	14968.1	267.1	19 methyls	17	14940.1	239.1	17 methyls
18	14993.1	292.1	21 methyls −3	18	14955.1	254.1	18 methyls +1
19	15024.1	323.1	23 methyls	19	14994.1	293.1	21 methyls −2
20	15052.2	351.2	25 methyls	20	15009.2	308.1	22 methyls −1
21	15068.2	367.2	26 methyls +2	21	15023.2	322.2	23 methyls −1
22	15137.2	436.2	31 methyls	23	15051.1	350.1	25 methyls −1
				24	15067.2	366.2	26 methyls +1
				25	15107.2	406.2	29 methyls −1
				26	15122.2	421.2	30 methyls

Biochemistry. Author manuscript; available in PMC 2015 January 30.

[1] monoisotopic masses were calculated using THRASH (34). Addition of a methyl group adds 14.057 Da, addition of an acetyl adds 42.011 Da, arg deimination or deamidation of asn or gln adds 0.985 Da. The added mass from 3 methyls/one trimethyl, or 1 acetyl group probably cannot be distinguished, as the mass shift of an acetyl from a trimethyl group would be 2.47 ppm for intact H2B, which may be below the mass accuracy of these spectra.

Table 2

Methylation of N-terminal peptides of histone H2B[1].

Peptide MH^+, precursor z	PTM composition	modified sites	Probability
Wild type H2B			
$PPK_4PADK_8K_9PASK_{13}$			
1277.763 3	1 Me	K4-Me	0.86
1291.773 4	1 Me_2	K4-Me_2	0.92
1305.800 3	1 Me_3	K4-Me_3	0.87
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1696.922 2	none	none	0.83
1738.981 3	3 Me	several combinations	0.70
$PPK_4PADK_8K_9PASK_{13}APATASK_{20}APE$			
2201.235 3	1 Me	K4-Me	0.71
2215.242 4	2 Me	several combinations	**
2215.247 5	2 Me	K4-Me, K8-Me	0.54
2229.267 3	1 Me_3	K4-Me_3	0.86
2257.255 4	1Ac+2Me	several combinations	**
2271.278 4	1Ac+3Me	multiple combinations	**
2285.288 4	1Ac+4Me	multiple combinations	**
2313.287 4	2Ac+3Me	multiple combinations	**
***hda-1* H2B**			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1696.939 3	none	none	0.99
1738.962 3	1Me_3	K8, 9 or 13	0.67
$PPK_4PADK_8K_9PASK_{13}APATASK_{20}APE$			
2243.251 3	1Ac+1Me	K4-Me, K8-Ac	0.60
2257.264 4	1Ac+2Me	K4-Me_2, K8-Ac	0.91
2271.280 4	1Ac+3Me	many combinations	**
2285.297 3	1Ac+4Me	many combinations	**
2299.270 3	2Ac+2Me	most have K4-Me_2, K8-Ac	0.75
2313.296 3	2Ac+3Me	many combinations	**
2327.304 3	2Ac+4Me	multiple combinations	**
2341.323 3	2Ac+5Me	multiple combinations	**
2327.266 3	3Ac+1Me	3 combinations	**
2341.286 3	3Ac+2Me	K4-Me_2, K20-,13-Ac &K8- or K9-Ac**	
2355.300 3	3Ac+1Me_3	multiple combinations	**

[1] Residues are numbered starting with the N-terminal met, which is removed *in vivo* for H2B. Me, methyl; Me_2, dimethyl; Me_3, trimethyl; Ac, acetyl. Site-specific modifications are indicated when uniquely identified by MS/MS spectra. The probability of correct sequencing for the top ranked peptide is listed. In most cases the H2B peptide and its composition of modifications are identified by mass alone.

** composition identified by mass.

Table 3

Histone H2B globular domain lysine methylations.

Peptide MH$^+$, precursor z	PTM composition	modified sites	Probability
Wild type H2B			
K_{45}ETYSSYIYK$_{54}$VLK$_{57}$QVHPDTGISNR			
2826.492 4	none	none	0.96
2854.507 3	2 Me	K45- or K54-Me_2	0.97
2854.489 4	2 Me	K57-Me_2	0.79
ASK$_{90}$LAAYNK$_{96}$K$_{97}$STISSR$_{103}$E			
1853.997 3	none	none	0.91
1882.028 3	2 Me	K96- or 97-Me_2	0.94
1896.041 3	1 Me_3	K97-Me_3	0.87
IQTSVRLILPGELAKHAVSEGTK$_{127}$AVTKYSSSTK$_{137}$			
3499.969 4	none	none	0.996
3513.980 4	1 Me	K127-Me	0.97
3527.984 3	2 Me	K127-Me_2	0.97
GTKAVTKYSSSTK$_{137}$			
1357.733 3	none	none	0.98
1595.859[1] 2	1 Me	K137-Me	0.81
***hda-1* H2B**			
LAK$_{119}$HAVSE			
854.472 2	none	none	0.94
882.469 2	2 Me	K119-Me_2	0.76
TYSSYIYK$_{54}$VLK$_{57}$QVHPD			
1941.009 2	none	none	0.99
1969.012 2	2 Me	K54-Me_2	0.76
1968.988 2	2 Me	K57-Me_2	0.59
1983.100 2	1 Me_3	K57-Me_3	0.72
ASK$_{90}$LAAYNK$_{96}$K$_{97}$STISSR$_{103}$E			
1854.018 2	none	none	0.99
1882.012 2	2 Me	K90-Me_2	0.80
2064.148[1] 2	3 Me	K90-Me_2, R103-Me	0.96
2036.148[1] 2	5 Me	K96,97-Me_2, R103-Me	0.96
IQTSVRLILPGELAK$_{119}$HAVSE			
2161.240 3	none	none	0.87
2175.248 3	1 Me	K119-Me	0.98
2189.267 3	2 Me	K119-Me_2	0.94

[1] identified after propionylation.

Table 4

Modified arginines in histone H2B.

Peptide MH^+, precursor z	PTM composition	modified sites	Probability
Wild type H2B			
$TGISR_{68}AMSILNSFVN$			
1723.881 2	none	none	0.99
1724.864 2	1 cit[1]	R68-cit	0.69
$ASK_{90}LAAYNKKSTISSR_{103}E$[2]			
1854.008 2	none	none	0.88
2092.133 2	1 Me	R103-Me	0.96
2064.134 2	3 Me	R103-Me, K90-Me_2	0.90
2064.116 2	3 Me	R103-Me, K96 or 97-Me_2	0.89
2036.139 2	5 Me	R103-Me, K90-Me_2; K96 or 97-Me_2	0.88
2036.116 2	3 Me, 2 Ac	R103-Me, K96,97Ac, K90-Me_2	0.80
$ASKLAAYNKKSTISSREIQTSVR_{110}LILPGE$			
3188.806 3	2 Me	R110-Me_2	0.52
***hda-1* H2B**			
$DTGISR_{68}AMSILNSFVND$			
1838.907 2	none	none	0.90
1839.893 2	1 cit	R68-cit	0.97
$ASKLAAYNKKSTISSR_{103}E$[2]			
1854.018 2	none	none	0.99
2092.132 2	1 Me	R103-Me[3]	0.95

[1] cit, citrulline;

[2] identified after propionylation, except for unmodified peptide;

[3] all 3 top choices have R103-Me

Table 5

Non N-terminal tail acetylated lysines in histone H2B.

Peptide MH^+, precursor z	PTM composition	modified sites if identified	Probability
Wild type H2B			
$ASK_{90}LAAYNK_{96}K_{97}STISSR_{103}E$			
1896.019 3	1 Ac	K97-Ac	0.92
2036.116[1] 2	3 Me, 2 Ac	R103-Me, K96,97Ac, K90-Me_2	0.80
***hda-1* H2B**			
$ASK_{90}LAAYNK_{96}K_{97}STISSR_{103}E$			
2036.101[1] 2	2Ac, 3Me	K96,97-Ac, R103Me, K90Me_2	0.87
$TYSSYIYK_{54}VLK_{57}QVHPD$			
1941.009 2	none	none	0.99
1983.029 2	1 Ac	K57-Ac	0.72
1983.026 2	1 Ac	K54-Ac	0.68
$LAK_{119}HAVSE$			
854.474 2	none	none	0.93
896.483 2	1 Ac	K119-Ac	0.84

[1] identified after propionylation

Table 6

Wild type and *hda-1* histone H2B 2D gel spot lysine acetylations.

Peptide MH^+, precursor z	PTM composition	modified sites if identified	Prob.
WT spot a			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1696.923 2	none	none	0.99
1738.933 3	1 Ac	K8- or K9-Ac	0.78
WT spot b			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1696.923 2	none	none	0.92
1738.934 3	1 Ac	K8-Ac	0.94
WT spot c			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1738.936 3	1 Ac	K8- or K9-Ac	0.86
1780.941 2	2 Ac	K13; K8 or K9-Ac	0.58
1780.957 2	2 Ac	K20; K9 or K8-Ac	0.70
***hda-1* spot a**			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1696.921 2	none	none	0.98
1738.933 3	1 Ac	K13-Ac	0.69
***hda-1* spot b**			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1696.925 2	none	none	0.99
1738.945 3	1 Ac	K8-Ac	0.87
1738.937 2	1 Ac	K13-Ac	0.61
***hda-1* spot c**			
$DK_8K_9PASK_{13}APATASK_{20}APE$			
1738.932 3	1 Ac	K8-Ac	0.89
1738.930 3	1 Ac	K13-Ac	0.85
1780.942 2	2 Ac	K20; K8- or K9-Ac	0.57
$DAGK_{29}K_{30}TAASG$			
947.480 2	1 Ac	K29-Ac	0.54[1]
Pooled spots, multiple experiments			
***hda-1* H2B**			
$K_8K_9PASK_{13}APATASK_{20}APE$			
1707.935 2	3 Ac	K9, K13, K20-Ac	0.88
$PPK_4PADK_8K_9PASK_{13}APATASK_{20}APE$			
2243.251 3	1Ac+1Me	K8-Ac, K4-Me	0.60
2257.264 4	1Ac+2Me	K8-Ac, K4-Me_2	0.91
2299.270 3	2Ac+2Me	K13, K8-Ac, K4-Me_2	0.75
2313.290 4	2Ac+3Me	K4-Ac; K8, 9 or 13-Ac	0.86[2]

[1] also identified by InsPecT (Table 7);

[2] all top choices have K4-Ac

Table 7

Analysis of histone H2B post-translational modifications using InsPecT[1].

identified peptide	MH+	z	p-value	modification
wild type H2B				
DKKPASKAPATASKAPE	1696.924,	2	0.00143	none
D**K$_8$+42**KPASKAPATASKAPE	1738.934,	3	4.0E-5	K-Ac
DKKPAS**K$_{13}$+42**APATASKAPE	1738.933,	3	0.0056	K-Ac
ASKLAAYNKKSTISSRE	1853.997,	3	7.0E-5	none
ASKLAAY**N$_{95}$+1**KKSTISSRE	1855.017,	4	0.026	N deamidation
TYSSYIYKVLK**Q$_{58}$+1**	1493.796,	2	0	Q deamidation before gluC cleavage
DTGISNRAMSILNSFVN	1838.905,	2	0.00035	none
DTGISNRA**M$_{70}$+16**SILNSFVN	1854.903,	2	2.0E-5	M oxidation
DTGISNRA**M$_{70}$+32**SILNSFVN	1870.900,	2	0.0069	M double oxidation
***hda-1* H2B**				
PP**K$_4$+42**PAD**K$_8$+42**KPASKAPATASKAPE	2271.280,	4	7.0E-5	K-Ac+K–Me$_3$[1]
DKKPASKAPATASKAPE	1696.923,	2	3.0E-5	none
D**K$_8$+42**KPASKAPATASKAPE	1738.935,	2	1.0E-5	K-Ac
DK**K$_9$+42**PASKAPATASKAPE	1738.931,	2	0.0006	K-Ac
DKKPAS**K$_{13}$+42**APATASKAPE	1738.937,	3	0.0030	K-Ac
DAG**K$_{29}$+42**KTAASG	947.479,	2	0.00021	K-Ac
ASKLAAYNKKSTISSRE	1854.019	2	0	none
AS**K$_{90}$+28**LAAYNKKSTISSRE	1882.015,	2	0	K-Me$_2$
I**Q$_{106}$+1**TSVRLILPGELAKHAVSE	2162.220,	3	0.0263	Q deamidation
IQTSVRLILPG**E$_{116}$+14**LAKHAVSE	2175.255,	2	0	E methyl ester
TYSSYIYKVLK**Q$_{58}$+1**	1493.782,	2	0	Q deamidation before gluC cleavage
DTGISNRAMSIL**N$_{74}$+1**SFVN	1839.893,	2	3.0E-5	N deamidation
TGISNRAMSILNSFVND	1838.964,	2	0	none
TGISNRA**M$_{70}$+1**6SILNSFVND	1854.911,	2	0.0013	M oxidation

[1] Acetyl asnd trimethyl groups are distinguished by precursor mass.

[2] mass errors are for 2 acetyl groups (20.2 ppm), 2 trimethyl groups (11.8 ppm), or one of each (4.16 ppm).

Table 8

Mapping of modified *Neurospora* histone H2B residues onto the structure of H2B in the *Xenopus* nucleosome core particle.

T-Coffee alignment of *Neurospora* and *Xenopus* histone H2B sequences

```
Nc              20        29                  54  57  68            83    90    9697    103    110119          127     137
MPPKPADKKKPASKAPATASKAPEKKDAGKKTAASGDKKKRTKARKETYSSYIYKVLKQVHPDTGISNRAMSILNSFVNDIFERVATEASKLAAYNKKSTISSREIQTSVRLILPGELAKHAVSEGTKAVTKYSSSTK
MP-EPAKSAPAPKKGS--------KKAVTKTQ-KKDGKKRRKTRKESYAIYVYKVLKQVHPDTGISSKAMSIMNSFVNDVFERIAGEASRLAHYNKRSTITSREIQTAVRLLLPGELAKHAVSEGTKAVTYTS-AK
xle  1                        26  30      40  43       54           69    76    8283    89      96          105           113  117
```

Location of modified residues

N.c.	***Xenopus***	**comment**
K54	K40	K40 is at the top and bottom H2B surface; its NZ is ~4.7A to guanosine 48 01P, ~4.6A H2a Q24 backbone oxygen at the H2B-H2A interface
K57	K43	K43 is near the top/bottom NCP surface, ~6.7A to thymidine 54 O1P
R68	K54	K54 is on the surface of the histone octamer but inside the DNA; NZ is 6.7A to guanosine 56 O1P, ~6.6A to H3 E94 OE2
K90	R76	R76 is buried in H2B core at center of NCP, abuts small central cavity
K96	K82	K82 is on H2B surface in histone octamer but inside DNA; Nz is 7.65A from adenosine O2P
K97	R83	R83 is on surface of H2B in histone octamer but inside DNA; directly binds adenosine and guanosine 34 phosphodiesters
R103	R89	R89 is at the top/bottom surface of NCP and at the H2B-H4 interface: R89 NH2 is ~3.2A to H4 K77 NZ, NH1 is ~2.6A to H4 H75 carbonyl oxygen
R110	R96	R96 side chain protrudes at the top/bottom surface
K119	K105	K105 side chain protrudes from top/bottom surface
K127	K113	K113 side chain protrudes from the top/bottom surface