Supplementary Information

(Experimental Procedures and Figures)

General:

All materials for Fmoc solid phase synthesis were purchased from Novabiochem (Darmstadt, Germany) and used without additional purification. Peptides were synthesized in Protein Technologies (Rainin) PS3 peptide synthesizer and purified using a Varian ProStar HPLC equipped with C18 reverse phase column. Amino acid analysis was performed in the Harvard microchemistry facility or the W. M. Keck facility at Yale University. DNA sequencing analysis was performed in the Sequencing and Synthesis core facility at Johns Hopkins University (Baltimore, MD). All buffering and bacterial media reagents were purchased from Fisher Scientific, Inc. (Hampton, NH). All primers for cloning and site directed mutagenesis were ordered from Integrated DNA Technologies (Coralville, IA) and were used without further purification. Chitin beads were purchased from New England BioLabs. Purification of proteins were carried out using a Pharmacia Biotech (GE Healthcare) FPLC equipped with monoS ion exchange column (GE Healthcare) or Superdex S-75 gelfiltration column (GE Healthcare). ^{14}C acetyl-CoA was purchased from Amersham. Unlabeled acetyl-CoA was purchased from Calbiochem. All fluorescence spectra were measured using a Horiba Jobin Yvon Fluoro MAX-3 spectrofluorometer. All additional chemicals were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO) and used without further purification. Phosphorimage analysis was perfomed on a Typhoon 9410 variable mode imager (GE Healthcare) and quantified using Image Quant TL software.

Synthesis of peptides:

Peptides were synthesized using Rink Amide AM resin (200-400 mesh) on a 0.1-0.2 mmol scale using the Fmoc strategy. Amino acids were double coupled (4 equivalents) with HATU were used (Figure S2). Deblocking and cleavage from resin was performed with a mixture of triisopropyl silane:water:trifluoroacetic acid (TFA) 5:5: 90 for ~3 hrs. Peptides were precipitated using cold ether (~ -20°C) and centrifuged at 3000 rpm for 10 min and the pellets were dissolved in ~ 5-10 mL water, flash-frozen in liquid nitrogen and lyophilized. Crude peptides were purified using reversed-phase HPLC (C18) using a gradient of water: acetonitrile (both 0.05 % TFA) as the mobile phase (10 mL/min) (Figure S5). The fractions were analyzed by MALDI-TOF (Figure S6, S7). Pure fractions were collected together, concentrated and lyophilized.

Circular Permutation (cp) of p300 HAT gene:

The p300 HAT DNA was PCR amplified as two fragments A and B (Figure S1) using Pfu DNA polymerase (Stratagene). Fragment A (encoding residues 1287 to 1522) was amplified using primers with Kpn1 and EcoR1 restriction sites in the forward and reverse

directions, respectively. Fragment B (encoding residues 1561 to 1666) was amplified using primers with Nde1 and Kpn1 restriction sites in the forward and reverse directions, respectively. The primer harboring the Kpn1 site in fragment B contained 21 additional nucleotides that would code for an unstructured loop GGSGGGT (Figure S1, indicated in red). The amplified fragments were digested with Kpn1 and EcoR1 (fragment A) and Nde1 and Kpn1 (fragment B) and ligated to the Nde1/Kpn1 treated pTYB2 vector using T4 DNA ligase and transformed into XL-Blue Supercompetent cells (Stratagene). After isolation and sequencing, excess nucleotides in between the 3' end of the p300 coding sequence and the intein were removed by Quikchange mutagenesis.

For generating catalytically-inactive cp-p300 HAT (cp-p300-F; containing a Y1467F mutation), the same protocol as above was used except: fragment A encoded residues 1287 to 1543 and a Y1467F mutation was introduced using Quikchange mutagenesis.

Expressed protein ligation to generate catalytically-inactive cp-p300 HATs (cp-p300-F):

The pTYB2 plasmid containing circularly permuted, catalytically inactive (Y1467F) p300 HAT in vector was transformed into BL21-CodonPlus-DE3-RIL cells. Four 6 mL overnight cultures were used to innoculate shaker flasks (2L, 2XYT media, carbanecillin (75 mg/L)) which were grown at 37°C. At A_{600}=0.6, the incubator temperature was reduced to 16°C for 20 min, and the cells were induced with 1mM isopropylthiogalactoside (IPTG) and then grown with shaking at 16°C for 20 h. The cells were harvested by spinning at 5000g for 20 min, resuspended in 30 mL lysis buffer (25 mM HEPES pH 7.9, 500 mM NaCl, 10 % glycerol, and 1 mM EDTA). A typical yield was 3.7 g cell paste per L culture.

For protein isolation and ligation, all operations were carried out at 4 °C unless otherwise mentioned. To a suspension of cells (11g) in lysis buffer (30 mL), two anti protease cocktail tablets (Complete Mini, Roche) in 1 mL water and PMSF (20 mg) in 2 mL ethanol were added. Cells were lysed by passing through a French pressure cell twice (18,000 psi) and the contents were then centrifuged at 27,000g for 25 min. The clarified supernatant (~30 mL) was applied to a column containing chitin beads (7.5 mL slurry of beads pre-washed with 10 column volumes of water and buffer A: 25 mM HEPES pH 7.9, 250 mM NaCl and 1 mM EDTA). The contents were gently rocked for 30 min and slowly drained over 30 min. The beads were washed with 15 mL of buffer A followed by buffer B (25 mM HEPES pH 7.9, 500 mM NaCl and 1 mM EDTA). The beads were then washed with 10 column volumes of buffer A and 2 column volumes of ligation buffer (20 mL buffer A + 656.8 mg of mercaptoethanesulfonic acid (MESNa, 200 mM)). N-Cys containing synthetic peptide (17-mer) in ligation buffer was pH adjusted to 7.8 and argon was bubbled through the solution for 5 min prior to applying to chitin beads (estimated final peptide concentration was 2 mM). The column was closed with a septum and argon was purged through the column for ~ 5 min. The ligation was allowed to proceed for 1-2 days under argon atmosphere at room temperature (~ 28°C) when SDSPAGE indicated the ligation proceeded to > 90% completion. The ligation mixture was drained and the beads were washed with 4 column volumes of buffer C (20

mM HEPES pH 7.9, 50 mM NaCl and 1 mM DTT). The washings were combined with the ligation mixture and dialyzed two times against 1 L of buffer C.

To purify cp-p300-F HATs, cation-exchange chromatography (using FPLC with a monoS column) was equilibrated with buffer C. The dialyzed crude protein was injected onto the column and eluted with a gradient of 50 mM NaCl to 1 M NaCl in buffer C at a rate of 0.5 mL/min and 2 mL fractions were collected. Fractions of high purity by SDSPAGE for the desired semisynthetic protein were combined and dialyzed two times against 1L of storage buffer (20 mM HEPES pH 7.9, 50 mM NaCl, 1 mM DTT and 10 % glycerol). The samples were then concentrated by ultrafiltration to ~ 2.5 to 5 mg/mL, flash-frozen and stored at -80°C. Using this protocol, 0-Ac, 3-Ac and 6-Ac, catalytically-inactive, cp-p300-F HATs were generated.

Mass spectrometric analysis of cp-p300-F proteins was performed using MALDI TOF mass spectrometry. The proteins were first loaded onto a C18 Zip-Tip (Millipore), washed with water (0.1% TFA) and then eluted with 2 μL of 75% acetonitrile: 25% water mixture (0.1 % TFA). This process was repeated again using *the same* C18 Zip-Tip. The eluted protein (1 μL) was co-crystallized with sinapinic acid matrix (1 μL) and subjected to laser-desorption ionization. The calculated and observed masses (linear, positive ion mode) are given in Figure S3.

Estimation of Binding Constants (K_d) of acetonyl-CoA with catalytically inactive cp-p300 HAT-F forms:

A typical assay buffer contained in a final volume of 200 μL: 150 mM HEPES pH 7.9, 150 mM NaCl, 2 mM DTT and 0.4 μM HAT and varying amounts of acetonyl-CoA (0 - 24 μM). A fluorescence spectrum (excitation at 295 nm, emission measured in the range of 310- 400 nm, slit width for incident light = 2 nm and for emission light = 4 nm) was obtained for every concentration of the ligand at 25˚C. The relative drop in fluorescence at 340 nm was plotted against the concentration of acetonyl-CoA (Figure S4). The data was fit to the quadratic equation:

$$F = F_0 - (F_0 - F_{min}) (P_{tot} + L + K_d) - ((P_{tot} + L + K_d)^2 - 4*P_{tot}*L)^{1/2} / 2P_{tot}$$

where F is the observed fluorescence, F_0 and F_{min} are the intial and final fluorescence intensities respectively; P_{tot} is the total p300 HAT concentration, and L is the total concentration of the ligand.

Production of Catalytically Active cp-p300 HATs:

Overexpression: In order to overcome the toxicity of p300 HAT to *E. coli*, catalytically-active cp-p300 HAT was co-expressed with a yeast deacetylase, Hst2. First, the *Hst2* gene was subcloned between Kpn1 and Pst1restriction sites in a PDHK29 vector. This *Hst2*-containing plasmid (conferring kanamycin resistance) was co-transformed (50 ng of each vector) with cp-p300 HAT in a pTYB2 vector (conferring ampicillin resistance) in BL21-CodonPlus-DE3-RIL cells and grown on LB-Agar plates containing ampicillin and kanamycin. For overexpression, 6 x 2L of 2XYT media containing kanamycin (40 mg/L) and carbenicillin (75 mg/ L) were inoculated with 12 x 6 mL overnight cultures. Cells were grown to an A_{600}=0.6 at 37 °C, then cooled to 16 °C, induced with 1 mM IPTG, and grown for 20 h at 16°C. Lysis and ligation were performed as described for cp-p300-F

except 38-mer peptides were used during ligation at ~1 mM final concentration and ligations were carried out for 3 days.

The concentrated ligation mixtures (1 mL) were applied to a superdex S-75 (GE Healthcare) gel filtration column (pre-equilibrated with the elution buffer: 20 mM HEPES pH 7.9, 50 mM NaCl and 1 mM DTT)) and eluted at 0.4 mL/min while collecting 0.5 mL fractions. Purified fractions containing cp-p300 were combined. After concentration to ~ 1-3mg/mL, the solutions were stabilized by adding glycerol (10 % final), flash frozen and stored at -80°C. Final protein concentrations of semisynthetic cp-p300 proteins were determined by SDSPAGE using Coomassie staining compared to a BSA standard curve. Note that these gels (Fig. 3) show trace higher molecular weight contamination likely from DnaK and GroEL, and lower molecular weight contamination (<10%) which may represent minor unligated or proteolyzed cp-p300 HAT. These proteins are presumed to be less pure than the Y/F cp-p300 HAT proteins because of the reduced recombinant production due to host cell toxicity.

Histone acetyltransferase (HAT) assays:

These reactions were performed analogous to previously described procedures.[1] Assay buffer contained in a final volume of 30 μL: 50 mM HEPES pH 7.9, 0.1mM EDTA, 50 μg/mL BSA, 150 mM NaCl, 1 mM DTT, 500 μM H4-15mer (substrate peptide), 6 nM cp-p300 HAT and 20 μM of ^{14}C-acetyl-CoA. The mixture except ^{14}C-acetyl-CoA was incubated at 30°C for 5 min. The reaction was initiated by the addition of 6 μL (2.8 μCi/mL) of ^{14}C-acetyl-CoA at 30°C. After 1-6 min, reactions were quenched by vortexing with 6 μL of gel-loading dye. The quenched reaction mixtures (13 or 15 μL) were loaded onto a 16 % Tris-Tricine gel and the acetylated peptide was electrophoretically separated from the protein. The gel was thoroughly washed with water to remove excess ^{14}C-acetyl-CoA and dried in a gel drier. The dried gel was then exposed to a phosphorimager plate for ~ 36 hrs and quantified relative to a ^{14}C-BSA standard. Product formation was shown to be linear with time and enzyme concentration in the range investigated and less than 10% of the limiting substrate was converted to product. Duplicate measurements were generally within 20%. Miminal autoacetylation is detected under these conditions, but if it were to occur, it would tend to attenuate the rate enhancement seen for the hexaacetylated protein.[1b] Note that a standrd, loop-deleted p300 HAT, known to be 5- to 10-fold more active than wt,[1] loop-containing, unacetylated p300 HAT, is also assayed for comparison to the cp-p300 proteins and the rates shown in Figure 3c.

References:

1. (a) Thompson, P. R.; Wang, D.; Wang, L.; Fulco, M.; Pediconi, N.; Zhang, D.; An, W.; Ge, Q.; Roeder, R. G.; Wong, J.; Levrero, M.; Sartorelli, V.; Cotter, R. J.; Cole, P. A. *Nat. Struct. Mol. Biol.* **2004**, *11*, 308-315; (b) Karanam, B.; Jiang, L.; Wang, L.; Kelleher, N. L.; Cole, P. A. *J. Biol. Chem.* **2006**, *281*, 40292-40301; (c) Karanam, B.; Wang, L.; Wang, D.; Liu, X.; Marmorstein, R.; Cotter, R.; Cole, P. A. *Biochemistry* **2007,** *46*, 8207-8216; (d) Liu, X.; Wang, L.; Zhao, K.; Thompson, P. R.; Hwang, Y.; Marmorstein, R.; Cole, P. A. *Nature*, **2008**, *451,* 846- 849.

Supplemental Figures

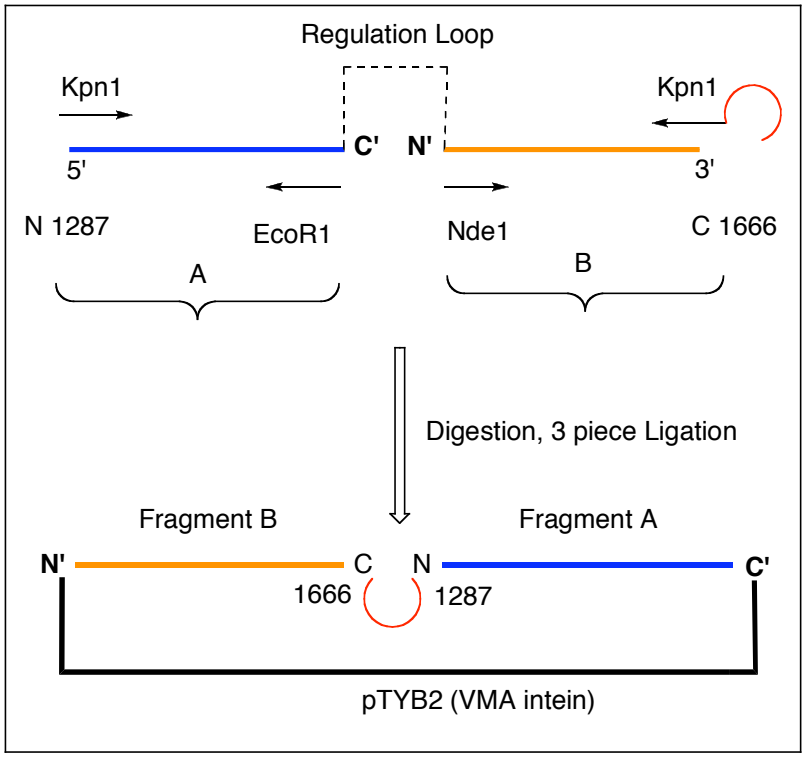

Figure S1. Schematic of circular permutation of p300 HAT.

H_2N—**C** E E E R K R E E N T S N E S T D V T K G D S **K** N A **K** K **K** N N **K** K T S **K** N **K**—$CONH_2$ (For cat-active cp-p300 HAT)

3-Ac : green
6-Ac : green and blue

H_2N—**C** S **K** N A **K** K **K** N N **K** K T S **K** N **K**—$CONH_2$ (For cat-inactive cp-p300 HAT)

H_2N—G R G K G G **K** G L G K G G A A K—$CONH_2$ (H4-15 mer (for assays))

Figure S2. Peptides synthesized for ligations and assay. The selection of triacetylated and hexaacetylated peptides used relate to the fact that the residues are either most sensitive to arginine mutation (triacetylated, P. Thomspon et al NSMB 2004)[1a] or are the earliest to be autoacetylated (hexaacetylated, B. Karanam et al, JBC 2006).[1b] As shown, these p300 loop peptide sequences are identical to the wt sequences except for Cys incorporation at the N-termini which is necessary for ligation.

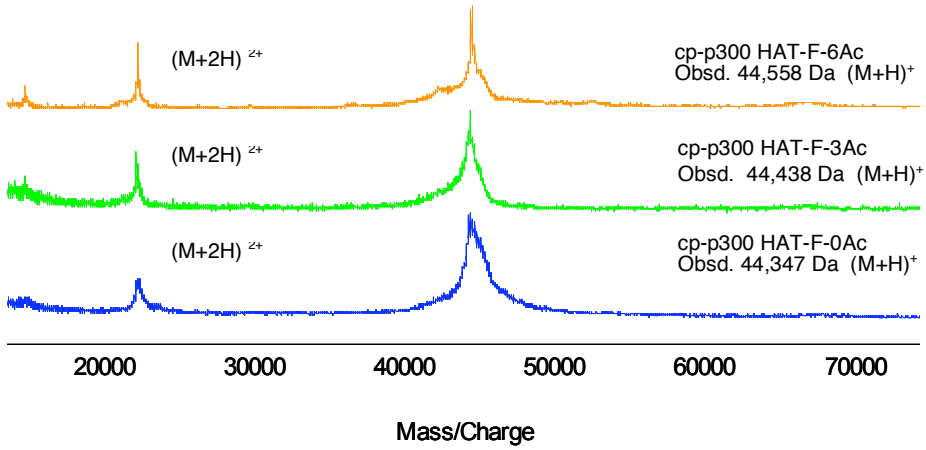

Figure S3. Mass Spectra (MALDI-TOF) of cp-p300 HAT–F $(M+H)^+$: **0Ac** calc: 44,395 Da, obsd. 44,347 Da; **3Ac** calc. 44,521 Da, obsd. 44,438 Da; **6Ac** calc. 44,647 Da, obsd. 44,558 Da. Sinapinic acid was used as the matrix.

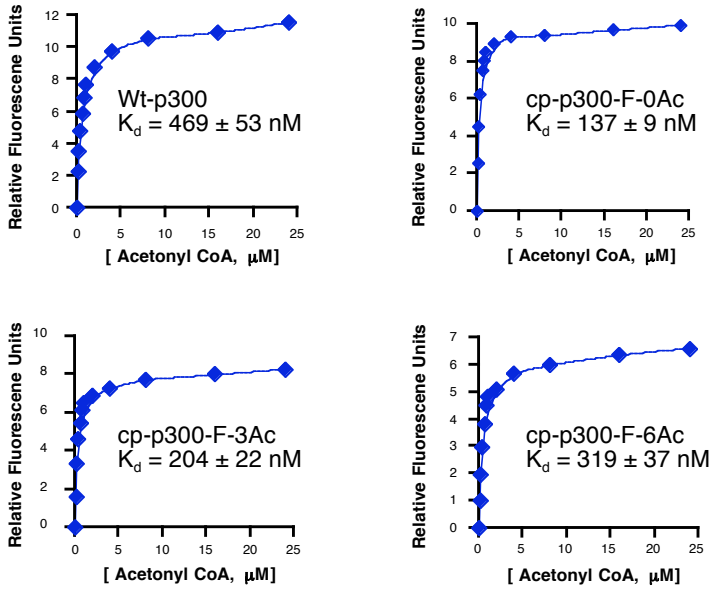

Figure S4. Fluorescence (tryptophan) binding curves for acetonyl-CoA with cp-p300 HATs

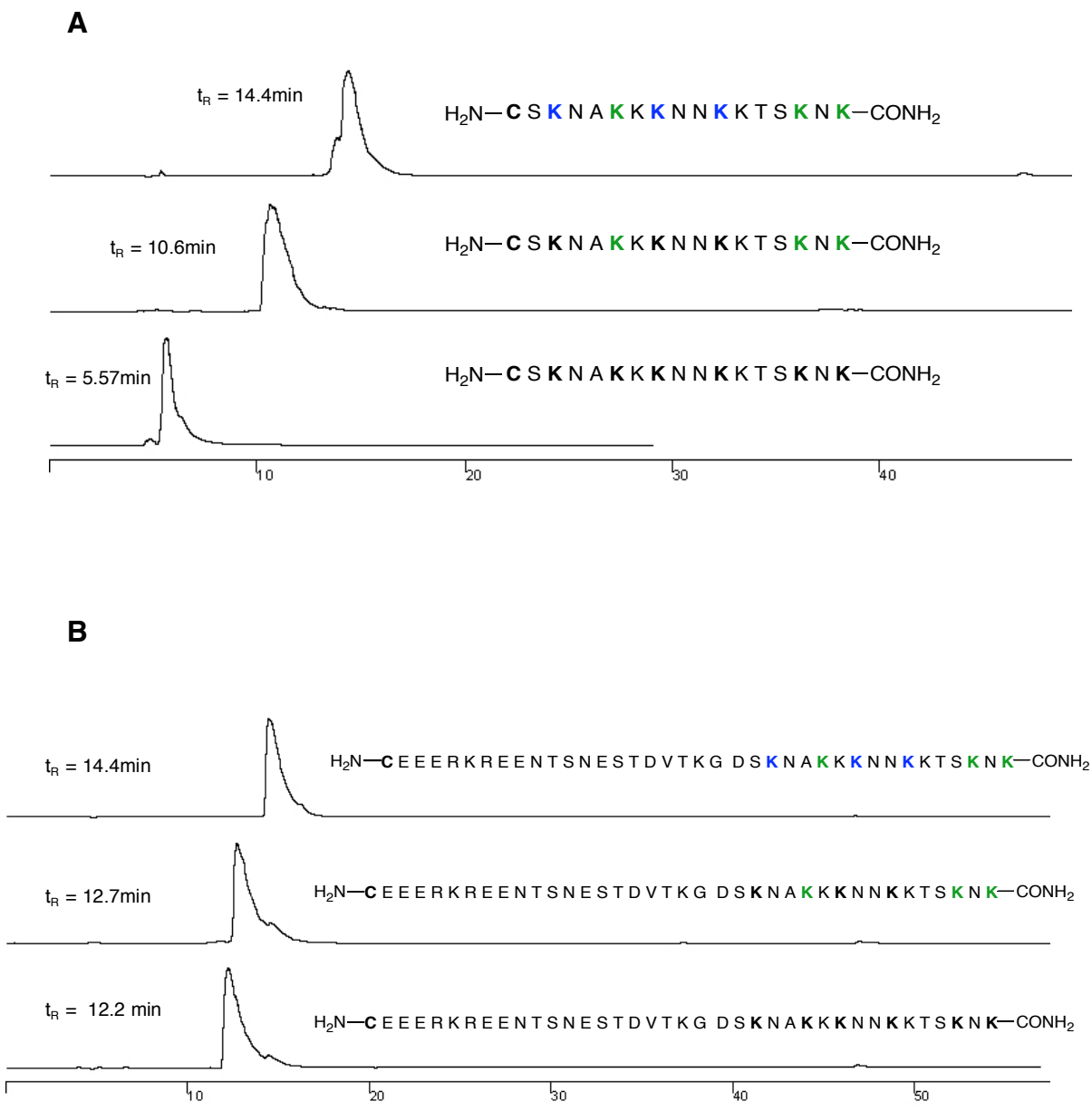

Figure S5. HPLC traces of peptides. Panel A: Peptides for the semisynthesis of cp-p300-F; Panel B: Peptides for the semisynthesis of catalytically active cp-p300. All purifications were done on a C18 reversed-phase column using a gradient of water: acetonitrile (both 0.05 % TFA) as the mobile phase (10 mL/min).

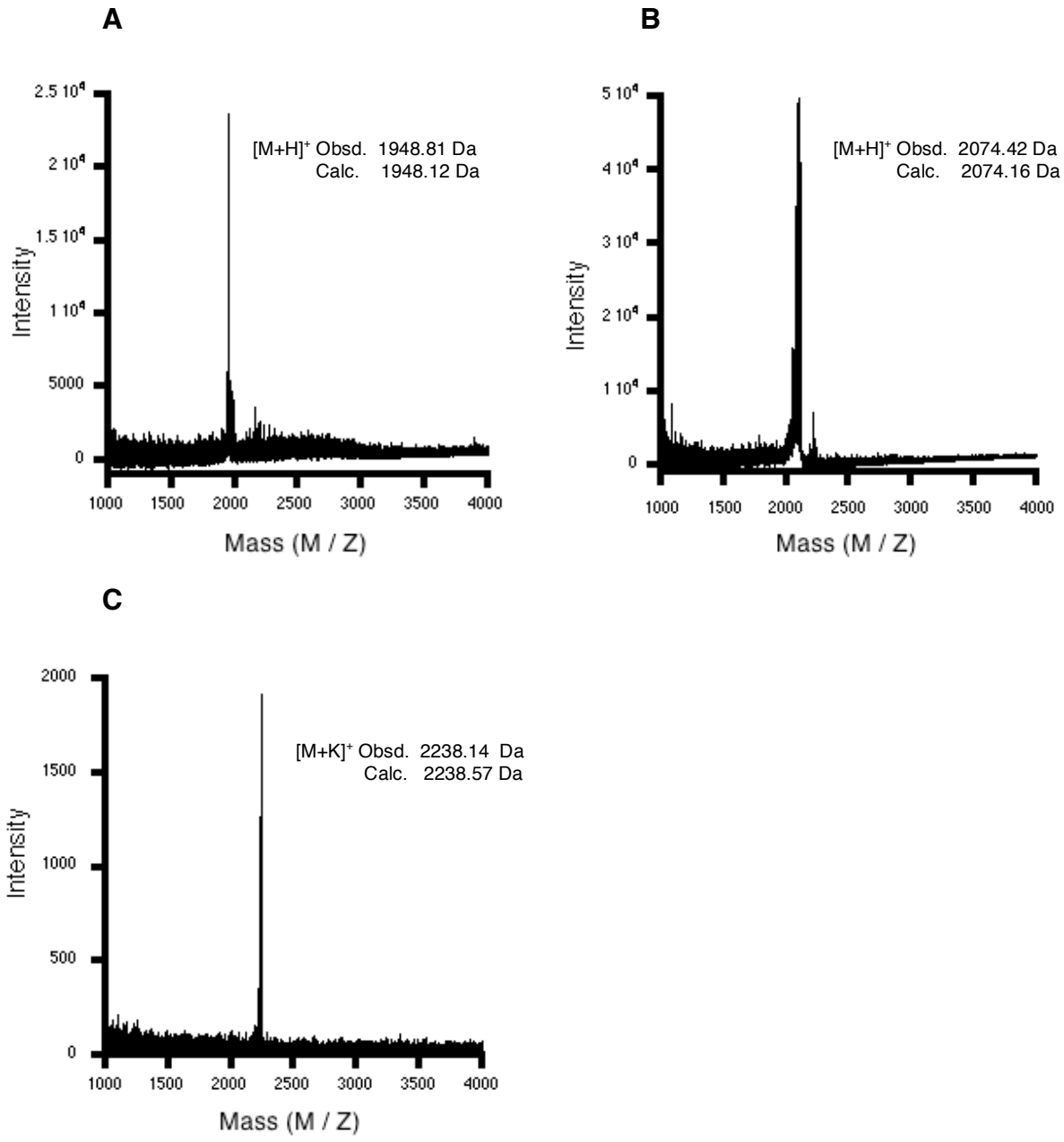

Figure S6. Panel A, B and C represents the mass spectrum (MALDI-TOF, reflector positive ion mode) of 17-mer peptides containing 0, 3 and 6 acetyl lysines respectively. α-Cyano-4-hydroxycinnamic acid (CHCA) was used as the matrix. For sequence information, refer Figure S2.

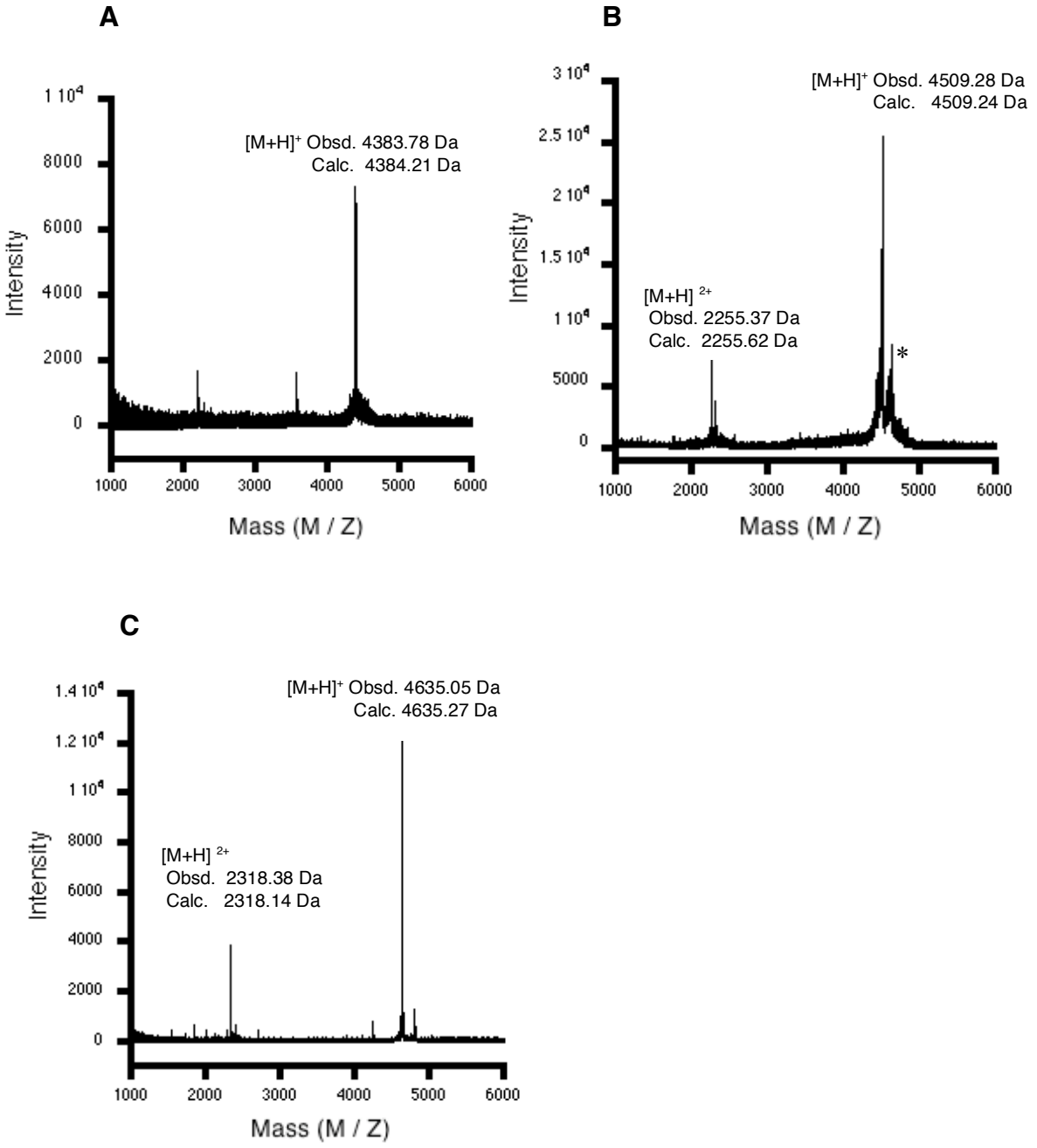

Figure S7. Panel A, B and C represents the mass spectrum (MALDI-TOF, reflector positive ion mode) of 38-mer peptides with 0, 3 and 6 acetyl lysines respectively. 2', 4', 6'-Trihydroxyacetophenone monohydrate (THAP) was used as the matrix. For sequence information, refer figure S2. * TFA adduct