SIRT6 AND SIRT7 TARGET IDENTIFICATION USING ACYL-NUCLEOSOMES

A Dissertation

by

WEI WANG

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Wenshe R. Liu
Committee Members,	David P. Barondeau
	Tadhg P. Begley
	Paul D. Straight
Head of Department,	Simon W. North

December 2018

Major Subject: Chemistry

Copyright 2018 Wei Wang

ABSTRACT

For eukaryotes, DNA is maintained in the form of chromatin, which is comprised of millions of repeated nucleosomes. Each nucleosome is assembled by wrapping 147 base pairs (bp) of DNA around a histone octamer containing 2 copies of H2A, H2B, H3 and H4. Histone residues undergo a plethora of post-translational modifications (PTM), which have a great impact on the conformation and organization of chromatin and eventually lead to significant epigenetic changes related to genetic events such as transcription, replication and DNA repair.

SIRT6 and SIRT7 are the newest members of the Sirtuin deacetylase family. With a unique NAD^+ dependent mechanism, Sirtuin enzymes remove histone lysine acetylation and acylations. SIRT6 has been reported as a site-specific histone deacetylase targeting H3K9 acetylation (H3K9ac) and H3K56 acetylation (H3K56ac). SIRT7 was reported to be specific for H3K18 acetylation (H3K18ac). Compared to other Sirtuin members, SIRT6 and SIRT7 exhibit very weak *in vitro* deacetylation activities on acetyl-histones and acetyl-peptides, but become much more active on acetyl-nucleosomes. However, all previous SIRT6 and SIRT7 histone targets were identified using acetyl-peptides as substrates, raising the question whether new active sites could be discovered if acetyl-nucleosomes were used as substrates

The main obstacle to biochemical study on nucleosomal level is the synthesis of homogenous nucleosomes with site-specific PTM. In this work, using unnatural amino acid incorporation technique, we synthesized homogenous histone H3 substrates with

acylation at defined lysine sites, assembled them into acyl- nucleosomes, and screened for targets of SIRT6 and SIRT7 with a novel fluorescence-based deacylation assay. As a result, we discovered H3 K18 and H3 K27 as the new targets for SIRT6, H3 K36 as the new target for SIRT7. The newly identified targets were confirmed *in vitro* using site-specific acetyl-nucleosomes, and *in vivo* using SIRT6 or SIRT7 overexpressed mammalian cell model. Using Chip-Seq assay, H3K36ac was discovered to be downregulated at specific gene promoters, nucleoli and telomeres. SIRT7 deacetylation is activated by extranucleosomal DNA through its electrostatic bridging between histone and enzyme.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Wenshe Liu, and my committee members, Dr. Barondeau, Dr. Begley, Dr. Straight, for their guidance and advice on my research project. I'm really grateful for being introduced to the interdisciplinary research of chemistry and biology. I'm also grateful for the all the support given by administrative staff of the chemistry department and TAMU campus.

I would also like thank all my lab colleagues from the Liu research group for the unforgettable experiences.

I would also like to thank all my families for the patience and support.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Dr. Wenshe Liu [advisor], Dr. David Barondeau and Dr. Tadhg Begley of the Department of chemistry and Dr. Paul Straight of the Department of Biochemistry and Biophysics.

The pyrimidine-tetrazine-FITC conjugated dye for Chapter II was provided by Dr. Alexander Deiters from the Universty of Pittsburgh. Selection of Pyrrolysyl tRNA synthetase library for incorporation of N^{ε}-7-octenoyl-L-Lysine (OcK) was done by Dr. Yu Zeng. Chip-Seq data analysis for Chapter III was provided by Professor Katrin F. Chua from Stanford University. Syntheis of N^{ε}-(7-azidoheptanoyl)-L-lysine hydrochloride (AzHeK) for Chapter III was done by Dr. Yadagiri Kurra. All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by National Institutes of Health, National Science Foundation and Welch Foundation.

TABLE OF CONTENTS

Page

ABSTRACT ii

ACKNOWLEDGEMENTS iv

CONTRIBUTORS AND FUNDING SOURCES v

TABLE OF CONTENTS vi

LIST OF FIGURES viii

CHAPTER I INTRODUCTION AND LITERATURE REVIEW 1

Epigenetics and Histone Post-translational Modifications (PTM) 1
Methods to synthesize homogenous nucleosomes with defined PTM 7
PTM installation through biorthogonal reaction 7
Semi-synthetic methods using Native Chemical Ligation and Expressed Protein Ligation 11
PTM incorporation with amber suppression technique 18
Biochemical studies with modified nucleosome substrates 28
PTM effects on chromatin structure and dynamics 28
Discovery of epigenetic enzyme functions 35
Summary 47

CHAPTER II IDENTIFICATION OF SIRT6 DEACYLATION TARGETS USING ACYL-NUCLEOSOMES 48

Introduction 48
Experimental section 51
Results and Discussion 69
Conclusion 85

CHAPTER III IDENTIFICATION OF SIRT7 DEACYLATION TARGETS USING ACYL-NUCLEOSOMES 87

Introduction 87
Experimental section 89
Results and discussion 121
Conclusion 142

CHAPTER IV CONCLUDING REMARKS AND FUTURE OUTLOOK 149
REFERENCES.....151

LIST OF FIGURES

Page

Figure 1 List of PTM....2

Figure 2 Histone residues (Highlighted) that are found with PTM....3

Figure 3 Methyl-lysine analogs prepared through reaction between cysteine and 2-haloethyl amines. 8

Figure 4 Using biocompatible and biorthogonal free radical reaction to site-specifically introduce PTM to dehydroalanine residue. 10

Figure 5 Semi-synthesis of Histones with PTM at the N-terminal region 13

Figure 6 Semi-synthesis of Histone H3 with T118 phosphorylation 14

Figure 7 Semi-synthesis of Histone H3K56ac 16

Figure 8 Semi-synthesis of H2AK119-Ubiquitin....17

Figure 9 Direct incorporation of Non-canonical amino acid (NcAA) by mutant tRNA synthetase....19

Figure 10 A versatile approach to generate lysine acylation in situ through bio-orthogonal Staudinger Ligation. 21

Figure 11 Incorporation of photocaged monomethyl-lysine and subsequent UV deprotection 22

Figure 12 Incorporation of Allysine precursor and subsequence reductive amination to afford dimethyl-lysine 23

Figure 13 Dipeptide from increase bioavailability of phosphotyrosine and its nonhydrolyzable mimic 26

Figure 14 Incorporation of phosphotyrosine analog and subsequent deprotection....27

Figure 15 Genetic incorporation of δ-thiol-L-lysine precursor and subsequent chemical transformation to afford ubiquitinated protein 27

Figure 16 H4K16ac inhibits chromatin array condensation....28

Figure 17 Histone lysine acetylation and tyrosine phosphorylation at nucleosome entry/exit region increase nucleosome DNA breathing by disrupting electrostatic interaction between histone and DNA.....31

Figure 18 Histone H3K115ac, H3K122ac and H3T118Ph at the nucleosome SIN region disrupt DNA Histone interaction around dyad position.....34

Figure 19 H3K56ac increase binding between DNA, Histone chaperone and H3/H4 tetramer.....36

Figure 20 Histone methyl-transferase ash1 and NSD2 shows distinct site-specificities on peptide, histone octamer and nucleosome. 38

Figure 21 Lysine acetylation on one histone H3 tail promote GCN5 acetyl-transferase reaction on another histone H3 tail.....40

Figure 22 HP1β dimerization mediated by CSD domain facilitates its binding with H3 K9 methyl-nucleosome. 42

Figure 23 53BP1 binds to H4K20me2 and H2AK15Ub nucleosome with multiple interactions.....44

Figure 24 synthetic route of amino acids 1 and 2[276] 51

Figure 25 ^{1}H and ^{13}C NMR of Nε-(7-octenoyl)-lysine (OcK) and Nε-(9-decenoyl)-lysine (DeK)[276].....53

Figure 26 Probing Sirt6-targeted lysine deacylation sites in nucleosome using a chemical biology approach. A terminal olefin-containing fatty acyl-lysine is site-specifically installed in a histone that is assembled into a nucleosome as an active Sirt6 substrate.[276] 70

Figure 27 The genetic incorporation of OcK. (A) Structures of AcK and OcK. (B) A diagram to illustrate fluorescent labeling of Ub-K48oc by a fluorogenic tetrazine dye, FITC-TZ. (C) The selective incorporation of OcK into ubiquitin at its K48 position. BL21(DE3) cells transformed with plasmid pEVOL-OCKRS coding OcKRS and $tRNA^{Pyl}$ and plasmid petDuet-UbK48Am coding ubiquitin with an amber mutation at K48 were grown in LB with or without 1 mM OcK. (D) The Deconvoluted ESI-MS spectrum of Ub-K48oc (calculated molecular weight: 9511.7 Da). (E) The selective labeling of Ub-K48oc by FITC-TZ. Proteins were incubated with 100 μM FITC-TZ for 2 h before they were analyzed by SDS-PAGE. The top panel shows the FITC-based fluorescent image of the gel and the bottom panel shows the same gel after it was stained with Coomassie Blue.[276] 73

Figure 28 *In vitro* reconstitution of Histone H3/H4 tetramers and H2A/H2B dimers.[276] 74

Figure 29 Sirt6 activities on oc-H3-H4 tetramers. (A) SDS-PAGE analysis of 9 purified oc-H3 proteins. (B) The H3-H4 tetramer structure. (C) Sirt6 activities on 9 oc-H3-H4 tetramers. A tetramer (1.6 μM) was incubated with or without 0.4 μM Sirt6 for 3 h before it was labeled with 200 μM FITC-TZ for 3 h and then analyzed by SDS-PAGE.[276] 75

Figure 30 Assembly of wildtype, Octenoylated and Acetylated Nucleosomes.[276] 77

Figure 31 Sirt6 activities on oc-nucleosomes. The image in the top left corner shows the structure of a nucleosome. For all gel images, the top panel shows the FITC-based fluorescent imaging and the bottom panel shows the EtBr-stained DNA from the same gel.[276] 77

Figure 32 SIRT6 site specificity is determined by amino acid sequence. (A) Relative catalytic activities of Sirt6 on five oc-nucleosome substrates. An oc-nucleosome substrate (0.66 μM) was incubated with 0.33 μM Sirt6. The reaction was stopped at 0.5, 1, and 2 h and oc-nucleosome samples at these different times were labeled with FITC-TZ and analyzed in a native agarose gel together with the original oc-nucleosome (0 h). Top panels show FITC-based images, middle panels show same gels based on EtBr staining, and bottom panels show relative integrated FITC-based fluorescent intensities at different reaction times. (B) Sirt6-targeted lysine sites show consensual sequences highlighted in red and blue rectangles.[276] 79

Figure 33 Sirt6 catalyzes deacetylation at H3 K9, H3 K18, and H3 K27. (A) The incorporation of AcK into a nucleosome. (B) Sirt6 activities on H3K9ac-, H3K18ac-, and H3K27ac-nucleosomes. Acetyl-nucleosomes (0.66 μM) were incubated with or without 0.33 μM Sirt6 for 3 h before they was analyzed by a native agarose gel. The top panel shows the EtBr stained agarose gel and the bottom panel shows the pan anti-Kac antibody detected acetylation levels at different samples. (C) Acetylation levels at H3 K9, H3 K18, and H3 K27 in Sirt6-transfected and control cells.[276] 81

Figure 34 Synthetic route of AzHeK 89

Figure 35 NMR spectra of N^{ε}-(7-azidoheptanoyl)-L-lysine hydrochloride (AzHeK) and intermediates 93

Figure 36 N-terminal tag didn't affect refolding of H2A/H2B dimer, H3/H4 tetramer and mononucleosome particles. 98

Figure 37 ESI-MS spectra of recombinant histone Az-H3 and Ac-H3 106

Figure 38 A click chemistry-based approach to profile SIRT7-targeted chromatin lysine deacylation sites. AzHeK that structurally resembles DeK is site-specifically incorporated into a histone that is further assembled *in vitro* with other histones and 601 DNA into an acyl-nucleosome. This acyl-nucleosome can be fluorescently labeled with a strained alkyne dye and subsequently visualized in a 1× TBE native-PAGE gel. When SIRT7 is catalytically active to remove fatty acylation from the incorporated AzHeK to recover lysine, incubating the assembled acyl-nucleosome with SIRT7 will remove the fatty acylation and therefore afford a deacylated nucleosome that cannot be fluorescently labeled and then visualized in a gel. 122

Figure 39 genetic incorporation of AzHeK into Ubiquitin at K48 position. (a) AzHeK incorporated at K48 position of ubiquitin could be site-specifically labelled by DBCO-MB488 through copper-free click reaction (b) test of AzHeK incorporation efficiency with previously reported mmOcKRS/tRNA^{Pyl} pair in E. coli BL21(DE3) CobB^{-} cell line in the following condition: 0.5 mM IPTG, 0.2% (w/v) L-arabinose, 1 mM AzHeK. (c) Labeling test of Ub-K48Az with DBCO-MB488 in the following condition: 5 uM ubiquitin, 0.1 mM DBCO-MB488, PBS buffer, pH 7.4. (d) MALDI-TOF mass of UbK48-AzHeK (measured: 9541.0, calculated: 9540.8).....124

Figure 40 (a) Time-dependent labeling of Ub-K48Az by DBCO-MB488 in the following conditions: 5 uM UbK48Az, 0.1 mM DBCO-MB488, PBS buffer pH 7.4 at r.t. (b) AzHeK stability in the presence of β-ME (Upper lane: Commassie blue staining; Bottom lane: FITC signal detection). (c) Labeling of His-TEV-H3K36Az by DBCO-MB488 caused band position shift on the 15% PAGE gel.....125

Figure 41 SIRT7 actively removes acylation from H3 K18 and H3 K36 in the nucleosome context. (a) SDS-PAGE analysis of eight purified AzHeK-containing H3 proteins. (b) Acyl-nucleosomes that we assembled from eight acyl-H3 proteins, H2A, H2B, H4, and 601 DNA (147 bp). Nu-H3K4az denotes an acyl-nucleosome assembled from H3K4az. All other acyl-nucleosomes are named in the same way. Mononucleosomes are typically shown around the 500 bp DNA position in an EtBr-stained 5% 1× TBE native-PAGE gel. (c) SIRT7 catalyzed deacylation activities on eight acyl-nucleosome substrates. Reaction conditions: we incubated 1 μM acyl-nucleosome with 0.5 μM SIRT7, 0.5 mM BME and 1 mM NAD+ at 37 °C for 2 h before we quenched the reaction by the addition of 20 mM nicotinamide. We then labeled the resulted acyl-nucleosome substrate with 100 μM MB488-DBCO for 1 h and analyzed it fluorescently in an 8% 1× TBE native-PAGE gel. The top panel shows the MB488-DBCO-based fluorescent imaging and the bottom panel shows the EtBr-stained DNA

from the same gel. (d) Quantified deacylation levels at eight H3 lysine sites. We repeated experiments shown in C thrice and calculated the deacylation level at each site by subtracting average MB488-DBCO-based fluorescence intensity of SIRT7-treated samples from that of untreated controls.....127

Figure 42 Effects of DNA and salt on SIRT7-catalyzed deacylation. (a) Free DNA inhibits SIRT7-catalyzed nucleosome deacylation. Reaction conditions: we incubated 1 μM nu-H3K36az with 0.1 μM SIRT7, a varied concentration of 147 bp DNA, 1 mM NAD+ and 0.5 mM BME at 37 °C for 3 h before we labeled the solution with 100 μM MB488-DBCO and analyzed the acyl-nucleosome substrate fluorescently in a 1× TBE native-PAGE gel. The top panel is MB488-DBCO-based imaging that indicates relative acylation levels and the bottom panel is EtBr-based imaging that confirms the nucleosome integrity. (b) Salt inhibits SIRT7-catalyzed nucleosome deacylation. Conditions were as same as in A except we replaced free DNA with NaCl. (c) Free DNA has a binary role on SIRT7-catalyzed deacylation of the H3K36az-H4 tetramer substrate. Reaction conditions: we incubated 2 μM H3K36az-H4 tetramer with 1 μM SIRT7, a varied concentration of 147 bp DNA, 0.5 mM BME, and 1 mM NAD^+ 37 °C for 2 h before we labeled the solution with 100 μM MB488-DBCO and analyzed the tetramer substrate fluorescently in a SDS-PAGE gel. 129

Figure 43 A linker DNA on an acyl-nucleosome substrate improves SIRT7-catalyzed deacylation. (a) A diagram to illustrate how a linker DNA affiliates the binding of SIRT7 to an acyl-nucleosome substrate. (b) The in vitro assembly of nu-H3K18az and nu-H3K36az with different lengths of linker DNAs. The gel showed the EtBr-stained nucleosomes. (c) The catalytic removal of acylation from nu-H3K36az substrates that contained different lengths of linker DNAs. Reaction conditions: we incubated a 1 μM H3K36az nucleosome with 0.1 μM SIRT7, 1 mM NAD and 0.5 mM BME at 37 °C for various times before we quenched the reaction by adding 20 mM nicotinamide, labeled the solution with 100 μM MB488-DBCO for 1 h, and then analyzed the nucleosome substrate fluorescently by 8% 1× TBE native-PAGE gel. (d) The catalytic removal of acylation from nu-H3K18az substrates that contained different lengths of linker DNAs. Reaction conditions were as same as for nu-H3K36az substrates except higher amount of SIRT7 (0.3 μM) was used. (e-f) Quantified deacylation percentage vs time for two acyl-nucleosome substrates. Reactions shown in c and d were repeated thrice. Deacylation was calculated and averaged by subtracting MB488-DBCO-based fluorescence intensity at different time points from that at 0 min.....131

Figure 44 SIRT7 site-specificity test on Az-nucleosomes with 20bp linker DNA. Reaction condition: 1 uM nucleosome, 0.2 uM SIRT7, 1 mM NAD+, 0.5 mM β-ME, 37℃, 2h. 132

Figure 45 Assembly of H3 ac-nucleosomes.....134

Figure 46 Validation of H3K36ac as SIRT7 targets both in vivo and in vitro. (a) SIRT7 deacetylation assay on acetyl-nucleosome in the following condition: 1 uM nucleosome, 0.2 uM SIRT7, 1 mM NAD+, 0.5 mM β-me, 37 ℃, 2 hours. Western blot bands on the upper panel were blotted by site-specific acetyl-histone antibodies, bands on the bottom panel were blotted by wildtype H3 antibodies. (b) global histone H3 acetylation level change after SIRT7 overexpression. EGFP, SIRT7-EGFP and inactive SIRT7-H187Y-EGFP mutant were transiently overexpressed in HEK293t cell, histones were extracted and acetylation levels were estimated by western blot. 135

Figure 47 SIRT7 is a physiologic nucleolar and nucleoplasmic H3K36 deacetylase. (a) Immunoblot showing increased H3K36ac levels in whole cell lysates from SIRT7 knock-down (KD) U2OS cells compared to control cells. (b) ChIP-PCR showing increased H3K36ac levels at promoters of SIRT7 target genes RPS20 and MRPS18b, but not COPS2, in SIRT7 KD and control U2OS cells. Two independent H3K36ac antibodies (Ab 1 and 2) were used. Data show the average of three technical replicates +/- SEM, and are representative of 2-4 biological replicates. Statistical significance was calculated performing Student's t-test. (c) Immunoblot of nucleolar fractionation showing increased levels of H3K36ac in SIRT7 KD cells versus control cells. β-Tubulin, p84, and FBL or UBF are shown as marker controls for cytoplasmic, nucleoplasmic, and nucleolar fractions, respectively. (d) Immunoblot of nucleolar fractionation of SIRT7 knockout versus control U2OS cells. (e) ChIP-PCR showing increased levels of H3K36ac at 18S rDNA sequences in SIRT7 KD cells. Data show the average of three biological replicates +/- SEM. Statistical significance was calculated performing Student's t-test. (f) ChIP-seq enrichment analysis of H3K36ac at repetitive sequences in SIRT7 KD versus control U2OS cells. Bar plot of $\log_{10}$(p-value) shows statistically significant enrichment of H3K36ac in SIRT7-KD cells at rDNA (18S, 28S) and telomeric (TTAGGG) repeats, but not at repetitive sequences from 20 specific centromeres (CT), 6 specific pericentromeres (PCT), or 4 centromeric and 4 pericentric consensus sequences. P-values were calculated performing Chi-squared test. 138

Figure 48 (a) Design of modified 601 DNA for the accessibility assay. (b) Thermal repositioning generated homogenous nucleosome species. Condition: 60 ℃, 30 min. 140

Figure 49 Lysine acetylation at H3 K36 improves DNA unwrapping from the nucleosome core. (a) A diagram to illustrate how H3K36ac increases DNA unwrapping from the nucleosome core and subsequent Pst1 access to an embedded Pst1 restriction site. (b) The *in vitro* assembly of three nucleosomes, nu', nu'-H3K18ac, and nu'-H3K36ac from corresponding histones and a mutant 601 DNA that had a Pst1 restriction site and a linker DNA. The gel shows EtBr-stained nucleosomes. (c) Pst1 digestion progress of three nucleosomes versus time. Reaction conditions: we incubated a 0.49 μM nucleosome with 2 U/uL Pst1 at 37 °C for various times before the reaction was stopped and the solution was analyzed by 5% 1× TBE native-PAGE. The gel was stained by EtBr. The top band in a lane shows the nucleosome with the linker DNA and the bottom band represents the one with the linker DNA removed. (d) The calculated ratio of digested nucleosome to undigested nucleosome versus time. We calculated and averaged the ratios based on integrated fluorescence intensities of original and digested nucleosome bands in gels. 141

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Epigenetics and Histone Post-translational Modifications (PTM)

A central question of biology research is to understand the basic rules governing evolution and development. Each species maintains a distinct program for the development from a single fertilized egg into a complex organism, and this program was stored in DNA, which is the primary carrier of genetic information.[1] Later, scientists start to realize phenotypic differences are not only determined by DNA sequences, but also determined by the epigenetic system, which determines how DNA code is interpreted.[2] For an individual organism, all somatic cells have the same DNA sequence, even though they differ a lot in functions and morphologies.[3] These differences are attributed to changes on epigenetic level, either inherited during reproduction or acquired from the environment.[4]

DNA is compacted in eukaryotic nucleus as chromatin, the histone-DNA complex. Stedman[5] proposed histone acted as a general barrier to gene transcription, and different cells should have different kinds of histones to modulate gene functions. Hence in the early days, histones was regarded as suppressors for DNA transcription and activation of gene required total removal of histones from chromatin, until in 1964, when Allfrey and Mirsky[6] speculated that gene activation was achieved through histone acetylation, not complete removal of histone from DNA. Composition of Chromatin was determined[7] as a string of repeating units of nucleosomes, a 147bp DNA wrapped

histone core containing 2 units of H2A, H2B, H3, and H4 (**Figure 2**).[8] The 2.8 Å nucleosome crystal structure[9] provided detailed structural information about core

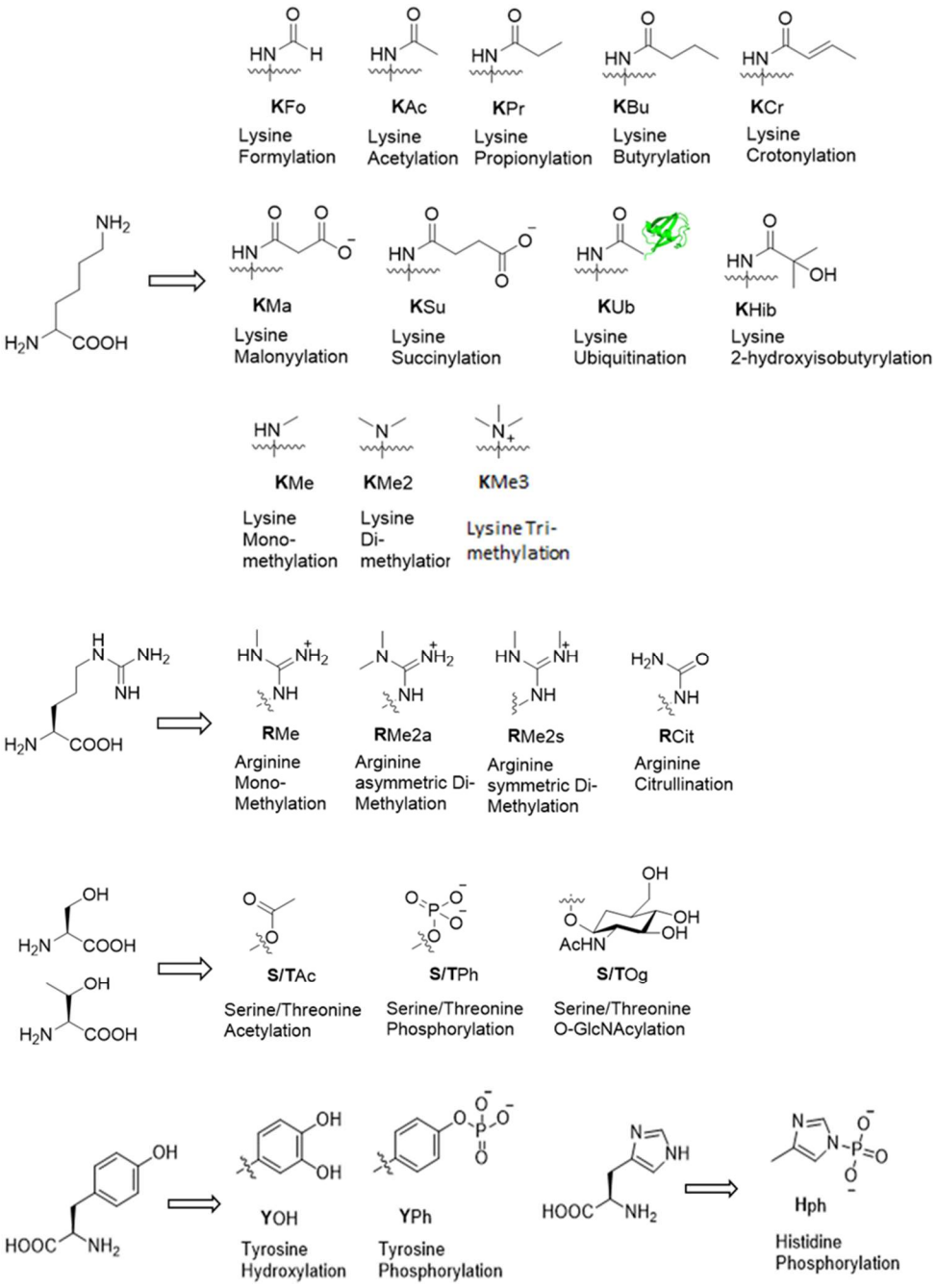


Figure 1 List of PTM

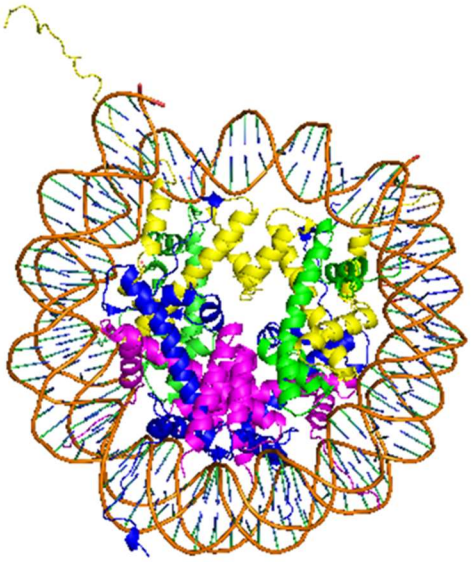


Figure 2 Histone residues (Highlighted) that are found with PTM

histone-DNA interactions and the extended histone N-termini harboring various post-translational modifications (**Figure 2**). It is now widely believed that higher order structure of chromatin is critical for gene regulation. Inactivated genes tend to form highly condensed heterochromatin, while active genes maintain a loose structure called euchromatin that is more accessible for transcription machinery.[10] Function and organization of the chromatin could be modulated through multiple mechanisms, such as

deposition of histone H1 to form higher-order chromatin fibers[11], substitution of histone variants for canonical histones[12], recruitment of histone remodeling complexes to evict or deposit histones[13], and covalent modification on DNA[14] and histones[15]. Covalent post-translational modification (PTM) (**Figure 1**), including methylation on DNA, acetylation, methylation, acylation and ubiquitination on histone lysine, methylation on arginine, phosphorylation on serine and threonine, could either directly alter DNA-histone, inter-nucleosome interactions, or serve as platforms to recruit other gene function regulators. Recombination of various modifications on different residues, together with the genetic information stored in DNA, form an accurate and complicated language for cells to precisely carry out gene functions in all kinds of cellular events. With the help of new technologies in proteomics and high-throughput sequencing, scientists are coming closer to fully understand this ancient but intricate language.

One of the major obstacles of our endeavor to fully understand epigenetic molecular mechanism is the synthesis of histone and nucleosome substrates with post-translational modification at defined sites. Our understanding of PTM function *in vivo* has been pushed forward by chromatin immunoprecipitation assay combined with high-throughput gene sequencing,[16] however, only correlative conclusion was provided by such assay and often the results are confounded by the misread of site-specific PTM antibodies.[17] On the other hand, the need for histone substrates with designated PTM grows, as the development of high-resolution proteomics study has discovered a large number of novel histone PTM. The functional and biochemical difference between these new PTM with the 'classical' PTM such as lysine acetylation and methylation needs to

be elucidated on the histone and nucleosome level.[18, 19] If full length histone with defined PTM is not available, unmodified residues would be mutated to mimic PTM function.[20-22] For example, there was an initial hypothesis assuming that lysine acetylation exerts its function by neutralizing the positive charge of native lysine, thus abolishing the salt bridge interaction between ε-NH_2 group and DNA triphosphate groups. Loss of DNA-histone interaction leads to destabilization of nucleosome structure, which explains the correlation between histone acetylation and active gene transcription.[23] Based on this hypothesis, lysine to glutamine mutagenesis has been used to mimic permanent lysine acetylation, and lysine to arginine mutagenesis has been used to mimic permanent unmodified lysine. Even though researchers have applied this strategy to study functions of acetylation both *in vivo* and *in vitro* for decades,[24-26] doubt will still remain before real homogenous acetyl-histone or nucleosome substrates are synthesized to confirm K to Q mutation has the same effect with real acetylation. With the recent success of synthesizing acetyl-nucleosome, it was found out that glutamine was not able to fully mimic acetyl lysine under many circumstances. For example, semi-synthetic H3K122ac and H3K115ac were refolded into nucleosomes, and these two modification were found to induce nucleosome DNA positioning change under thermal conditions. Nucleosomes with H3K115Q and H3K122Q mutations were also prepared, but these two mimics were not able to demonstrate the same effect with acetyl-nucleosomes, indicating effects other than electrostatic interactions may exist for lysine acetylation.[27] The power of the yeast genetics and limited copy number for yeast histone genes have enabled in vivo observation of phenotypic change with lysine to glutamine

change.[25, 28] Nevertheless, careful interpretation of these results is required because acetylation is transient and dynamic *in vivo*,[29] but glutamine and arginine mimic effect couldn't be reversed. In addition, part of acetylation's function is through the recruitment of specific binding protein called bromodomains,[30, 31] which couldn't recognize glutamine mimic.

Alternatively, histones or nucleosomes with defined PTM used to be prepared by incubating unmodified substrates with histone modifying enzymes.[32, 33] But this approach has only narrow application due to the difficulties of purifying recombinant epigenetic enzyme, which often exists in forms of large complexes.[34-36] Homogenously modified histone is hard to prepare using this method since very few of these enzymes have single defined target site, most have several target sites with distinct reaction efficiency.[37-40] Instead of being assembled from recombinant histone, chromatin substrates have been isolated from nuclear extract, and directly used for *in vitro* assay.[41-43] But the results obtained needs to be reconfirmed by other methods since *in vivo* isolated chromatin may contain other preinstalled modifications, which may affect the biological functions of the sites being studied.[44] Above limitations of traditional approaches have prompted chemists to develop convenient and versatile approach to synthesize histones with defined PTM both *in vitro* and *in vivo*.[45] Currently, full length histones with defined PTM could be synthesized via the following three strategies--direct biorthogonal reactions, semi-synthetic strategy and genetic incorporation, and the modified histone and nucleosome substrates have been utilized to unravel PTM's impact on chromatin dynamics and to probe epigenetic enzyme functions.[46]

Methods to synthesize homogenous nucleosomes with defined PTM

PTM installation through biorthogonal reaction

Most histones are devoid of cysteines, and the only conserved cysteine in H3 has been proved to be inessential for yeast.[47] Therefore, after the unique cysteine has been mutated to alanine, additional cysteine introduced by site-directed mutagenesis could undergo biorthogonal reaction with all kinds of electrophiles to afford PTM mimics. Simon et. al. generated mono-, di-, tri-methyllysine analogs (MLA), (H3K9$_C$me1, H3K9$_C$me2 and H3K9$_C$me3) through selective alkylations between cysteine at H3K9 and 2-haloethyl amines (**Figure 3**). These methyl-lysine analogs have been shown to mimic the function of methyl lysine such as binding with effector protein, recognition by methyl-lysine antibody, and histone methyltransferase crosstalk.[48] Other PTM analogs such as acetyl-lysine and methyl-arginine have also been prepared via cysteine conjugation strategy.[49, 50]

One significant strength of cysteine conjugation is that straight forward generation of large quantity of histone sample enabled the structural study of PTMs' effect on chromatin conformation change.[51] Also this technique is much easier to handle by biologists compared to native chemical ligation and genetic incorporation. However, despite the intrinsic drawbacks of the reaction such as slow rate and possible side reaction, it still remains doubtful whether these analogs would truly replicate the native methyl-lysine. Substitution of methylene with group with sulfur atom would slightly

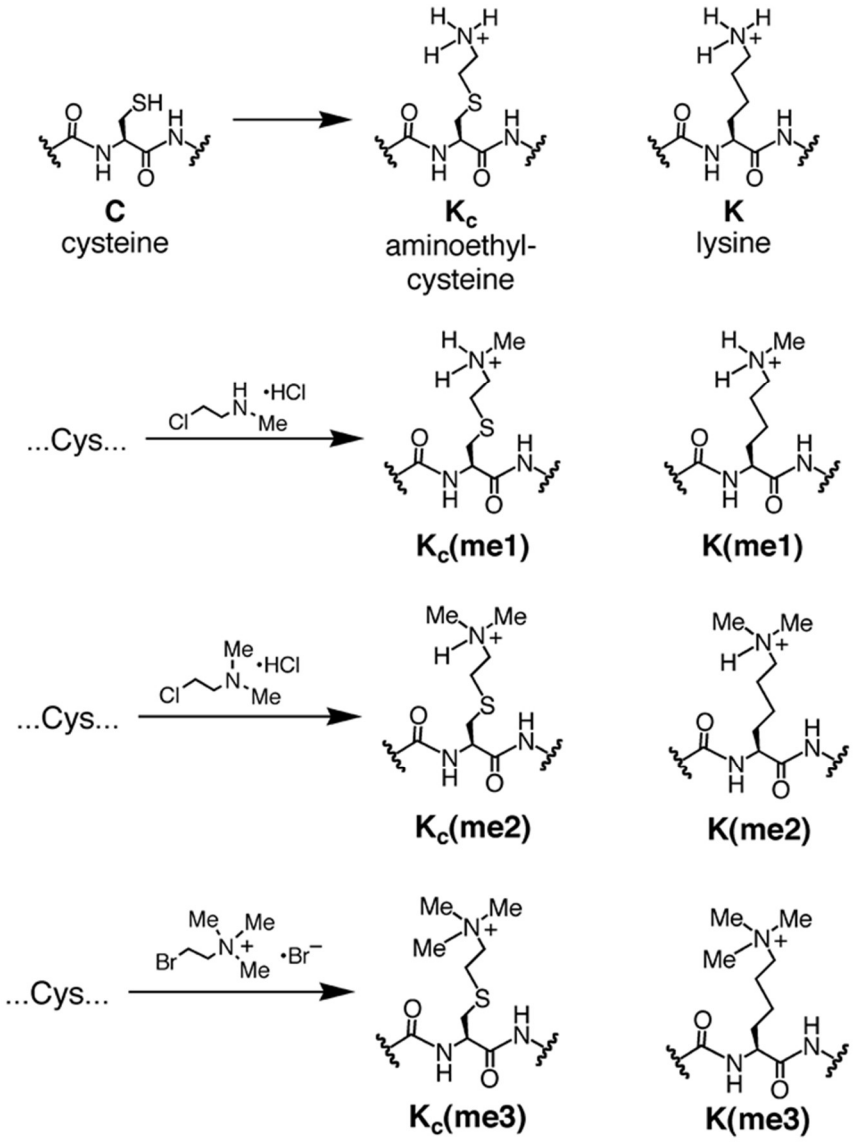


Figure 3 Methyl-lysine analogs prepared through reaction between cysteine and 2-haloethyl amines.

lengthen the lysine side chain, and sulfur's electron withdrawing effect would alter the acidity of ε-NH2 group.[52, 53] The resulting structural change could potentially affect the recognition between methylation analog and effector protein, as is demonstrated in a comparison study of chromodomain binding affinity between authentic methyl-peptide

and its methyl-lysine analog.[54] It is case dependent as to whether K_C(me) could fully mimic methylation lysine's binding feature. Caution needs to be taken before using these analogs to probe methyl-lysine biological function.

Recently, Wright et. al.[55] utilized biocompatible and biorthogonal free radical reaction to site-specifically introduce a wide range of PTM to proteins (**Figure 4**). Alkyl radicals, which were generated through single electron transfer from metals like Zn(0) or In(0), readily reacted with dehydroalanine introduced at different positions with varying levels of secondary structure, to afford modified residues including methyl-lysine, methyl-arginine, phophoryl-serine, and glycosyl-glutamine. Compared to previously reported cysteine reaction with alkyl halides, dehydroalanine reaction with carbon radical displayed much higher efficiency (complete in less than 30 mins) and larger substrate scope. The most significant drawback of this strategy was the lack of stereoselectivity, with almost equal amount of D- and L- amino acid generated in the process. In this paper, dehydroalanine was generated from cysteine introduced by site-directed mutagenesis, which is not applicable for protein containing multiple native cysteine. To resolve this problem, dehydroalanine could be generated from O-phosphoserine introduced by genetic incorporation.[56]

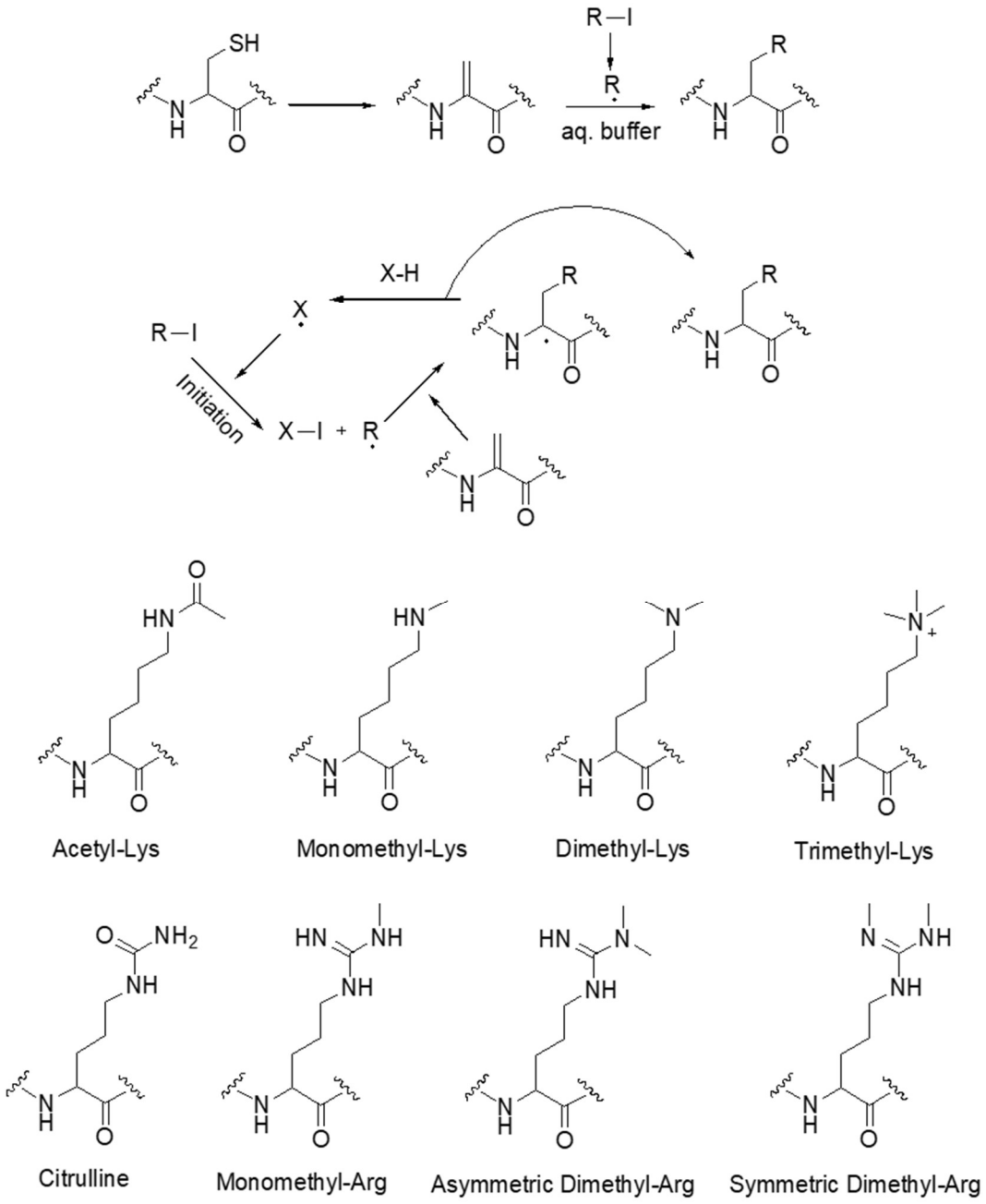


Figure 4 Using biocompatible and biorthogonal free radical reaction to site-specifically introduce PTM to dehydroalanine residue.

Semi-synthetic methods using Native Chemical Ligation and Expressed Protein Ligation

Histone N-terminal tails extend away from nucleosome core in an unstructured conformation, therefore for quite long time, short peptides containing PTM at specific site was treated as mimic for studying PTM function in vitro.[57-60] After the hallmark study of Merrifield to synthesize a fully active 124 residues long ribonuclease A,[61] Solid phase peptide synthesis (SPPS) enabled facile synthesis of numerous long peptides and even small protein.[62] The principle of SPPS is to elongate peptide chain on the solid support through repetitive cycles of deprotecting α-amine group and coupling of protected amino acid to the newly exposed α-amine group. This *de novo* synthesis strategy enabled the site-specific incorporation of various post-translational modifications into peptides, including lysine acetylation,[63, 64] methylation,[65, 66] arginine methylation,[67] Ser/Thr phosphorylation,[68, 69] glycosylation,[70] ADP-ribosylation[71] and ubiquitination.[72] Peptides with PTM were utilized to study PTM reader proteins,[73] PTM modifying enzymes[60, 74] and generation of antibodies highly selective for PTM at specific site.[75, 76] With the discovery of PTM in the histone core, peptide level research no longer meet the requirement for epigenetic study, as secondary structures and allosteric effects must be considered for PTM study at the globular domain.[77] For the N-terminal PTM, the need for full-length histone and nucleosome substrates also grows as additional regulative effects exist beyond the few residues surrounding the PTM being studied.[78] All histone proteins are over 100 residues long, which are difficult to prepare with routine SPPS procedure. This issue was resolved by highly selective peptide ligation approach including Native Chemical Ligation (NCL),[79] Expressed Protein

Ligation (EPL) [80] and Sortase mediated protein ligation.[81] To determine which ligation strategy to use, location and surrounding residues of the PTM need to be considered.

To generate full-length histone with PTM on the N-terminal tail, a short peptide encompassing modified residues is synthesized by SPPS with α-thioester at the carboxylic termini, and the remaining histone fragment is recombinantly expressed with an N-terminal cysteine. Then the short peptide and recombinant fragment are connected by transesterification, a native peptide bond is later formed via intramolecular S to N acyl shift, at the same time leaving a non-native cysteine scar.[82] Normally cysteine mutation would be introduced at less conserved histone residues away from nucleosome core to minimize unwanted functional effect. Otherwise, the cysteine scar could be eliminated by desulfurization to afford native alanine residue,[83] or modified with electrophilic reagent to afford native lysine analog.[84] For example, to incorporate single or multiple acetylations and methylations on H3 K4 H3 K9, H3 K14, H3 K18, H3 K23 H4 K5 and H4 K8, alanine to cysteine mutation was introduced at H3 A25 and H4 A15. Peptide fragments representing H3 1-24 and H4 1-14 with pre-installed PTMs were synthesized by Fmoc/TFA SPPS, and C-terminal fragments representing H3 C25-135 and H4 C15-102 were recombinantly expressed with an extra N-terminal Met residue, which could be autonomously removed before purification. After native chemical ligation step, the artificial cysteine residue could be reduced to native alanine by H2/Raney nickel (**Figure 5**).[85]

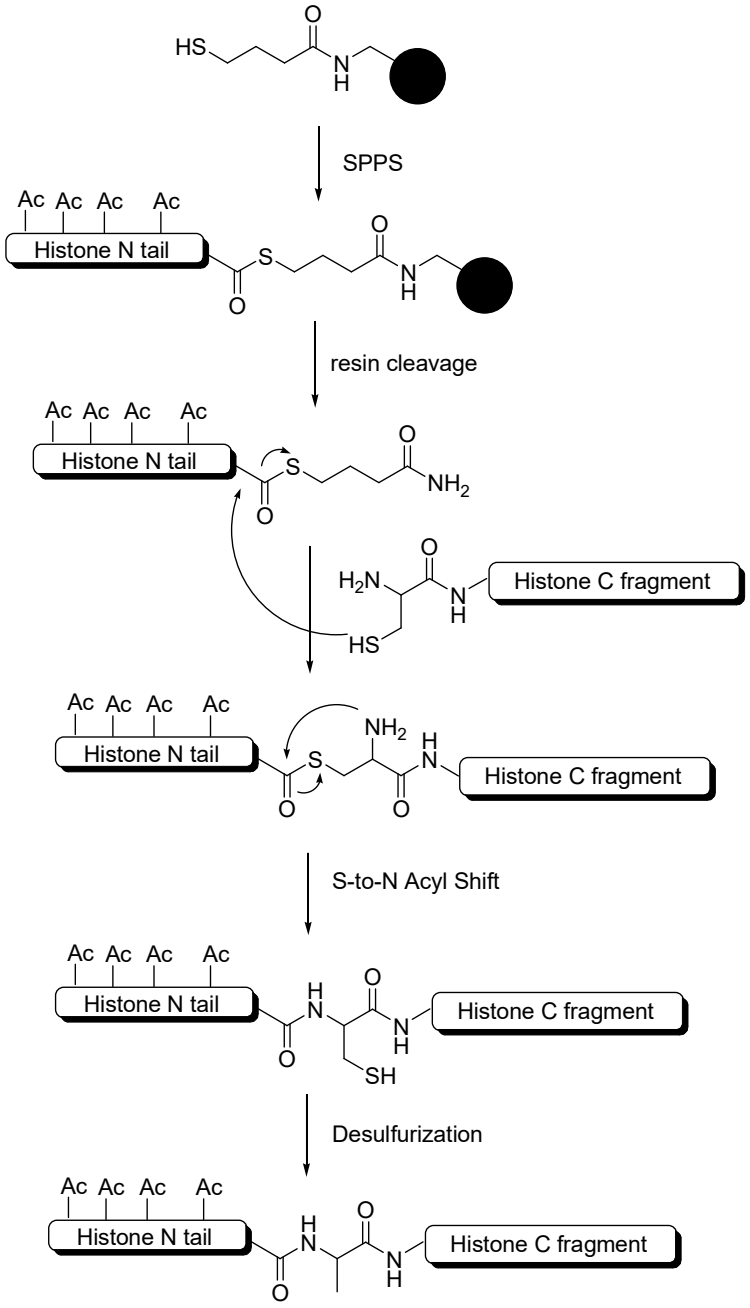


Figure 5 Semi-synthesis of Histones with PTM at the N-terminal region

To install modifications onto histone C-terminal tail, a short C-terminal peptide capped with cysteine is synthesized by SPPS, and the longer N-fragment is recombinantly expressed as an intein-fusion, which is cleaved by thiol reagent to generate α-thioesters. Full length histone is then prepared via Native Chemical ligation between these two parts. For H3, the native cysteine C110 could be taken advantage for the traceless synthesis, PTM could be installed at C-terminal peptide from 109 to 135, which is ligated to 1-109 recombinant α-thioester fragment derived from intein-mediated splicing (**Figure 6**).[86] For H4, alanine 76 could be mutated to cysteine as ligation site, which is later reduced back to alanine to afford the scarless histone.[87]

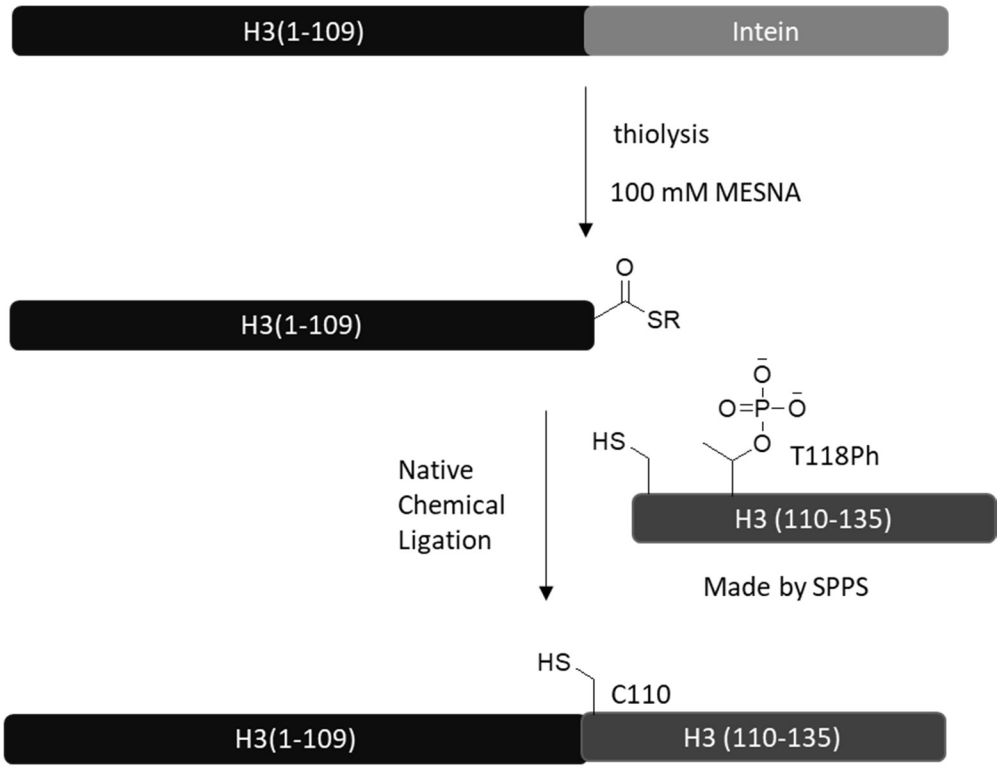


Figure 6 Semi-synthesis of Histone H3 with T118 phosphorylation

For histone modification on the globular domain, sequential ligation strategy is required to combine three fragments, which could either be synthesized by SPPS or recombinantly expressed.[88] To synthesize full length histone H3 with K56 acetylation, Shimko et. al. selected native alanine 47 and 91 as the ligation site and mutated them to cysteine. The C-terminal fragment (91-135) was synthesized with a free cysteine at A91, and regioselectively ligated with the middle fragment (47 -90) with cysteine 47 protected as thiazole form, which was deprotected to free cysteine for the second ligation step with the N terminal fragment (1-46) (**Figure 7**).[89] Since all three fragment could be synthesized by SPPS, similar sequential ligation strategy could be used to incorporate any amount of PTM at random positions.

To synthesize ubiquitinated histone, one popular strategy requires ubiquitin to be converted to α-thioester at C-terminal and undergo native chemical ligation with lysine side chain modified with β-thiol amine functional group. Chatterjee et. al. produced ubiquitin a-thioester lacking the C-terminal glycine, Gly76, and modified H2AK120 with a bromoacetic acid and a protected 1, 2 amino thiol reagent, which together acted as a surrogate form of Gly76. Ligation between these two parts generated native form of H2A ubiquitinated at K120, with the following elimination of the artificial β-thiol group.[72] To accelerate the procedure, ubiquitin Gly76 was surrogated with a cysteine attached to histone lysine side chain. After sequential ligation and desulfurization steps, ubiquitinated histone was synthesized with a G76A mutation without obviously changes to nucleosome structure and function (**Figure 8**).[90]

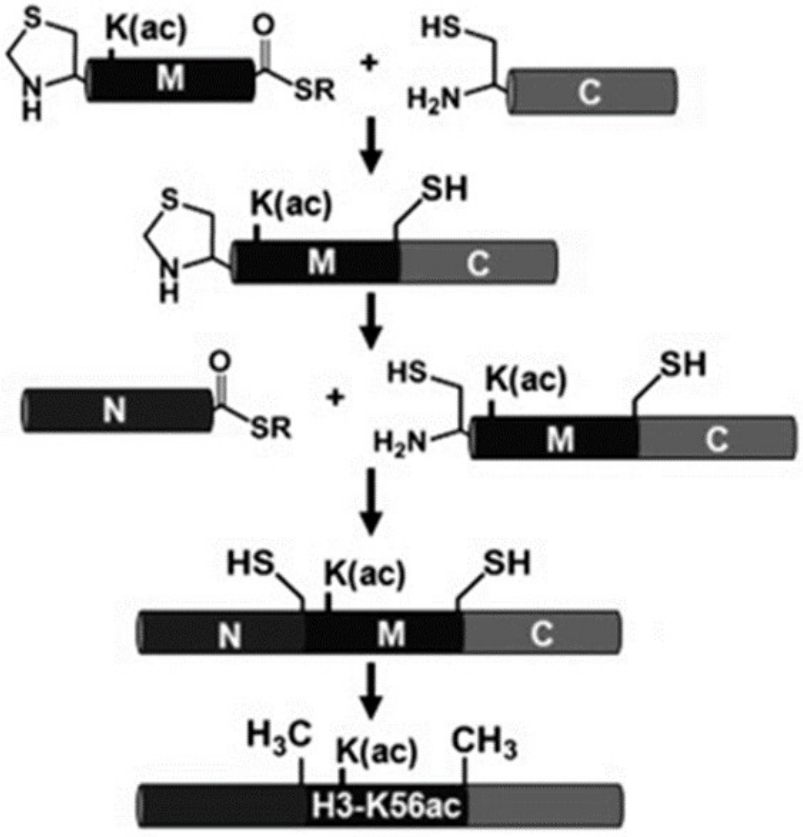


Figure 7 Semi-synthesis of Histone H3K56ac

Despite its power of incorporating multiple PTM simultaneously into histone, the widespread of semi-synthetic strategy was limited by the difficulty and low yield of synthesizing long peptides. Firstly, rounds of protection and deprotection steps are required for the solid peptide synthesis, not only for the α-amino group during peptide elongation, but also for side chains such as lysine.[91] Lengthening of synthetic route and following extra purification processes inevitably bring down the overall yield and consume more time. Moreover, at least one cysteine is required for the ligation, there is only one native cysteine H3C110 which could only be utilized to incorporate

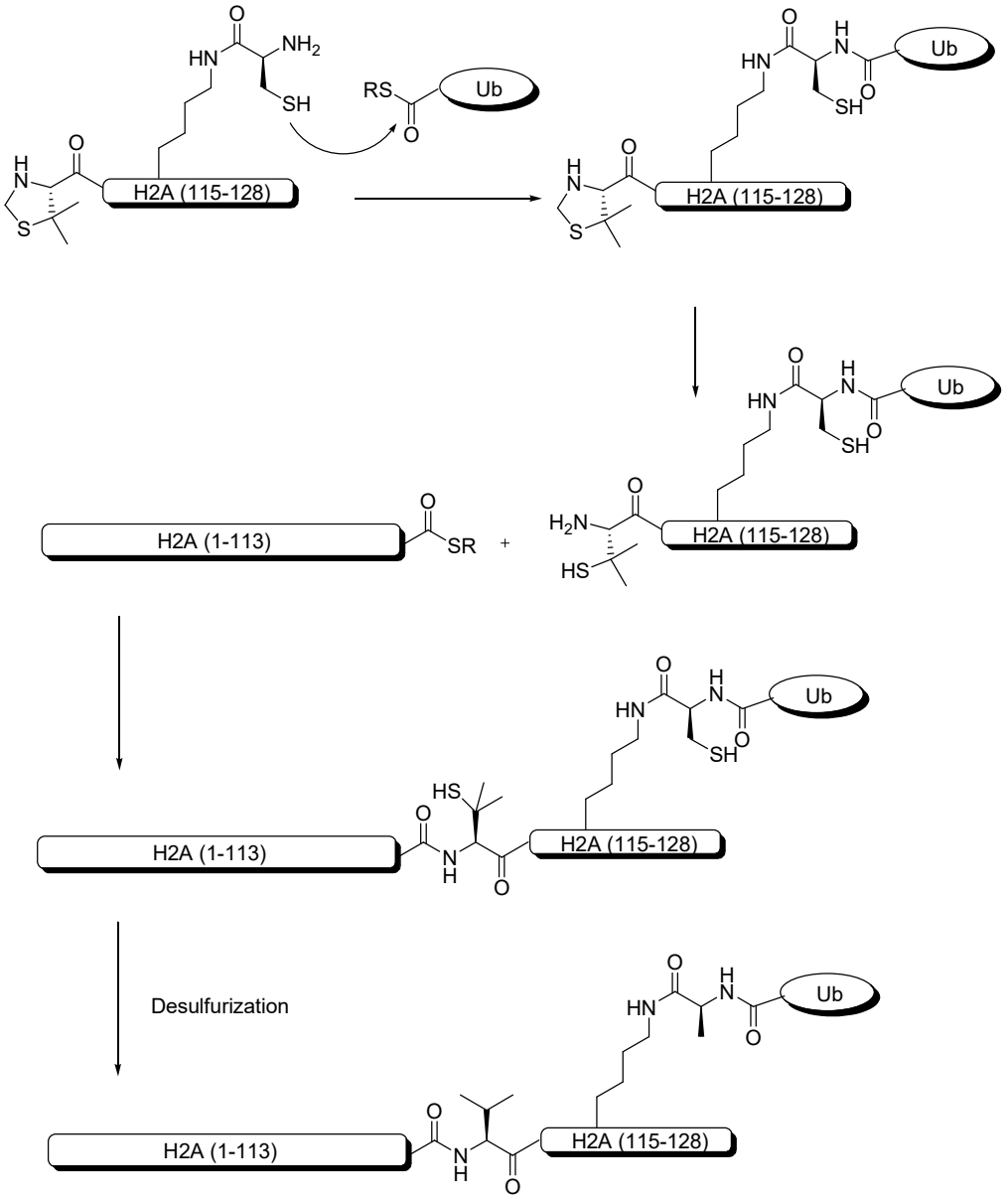


Figure 8 Semi-synthesis of H2AK119-Ubiquitin

modifications encompassing residues from 110 to 135. Native alanines could be mutated to cysteines, but extra desulfurization step is required to convert them back. For many other cases, the artificial cysteine scar is unable to be removed, and extra control experiment is required to exclude the possibility that the cysteine may affect the

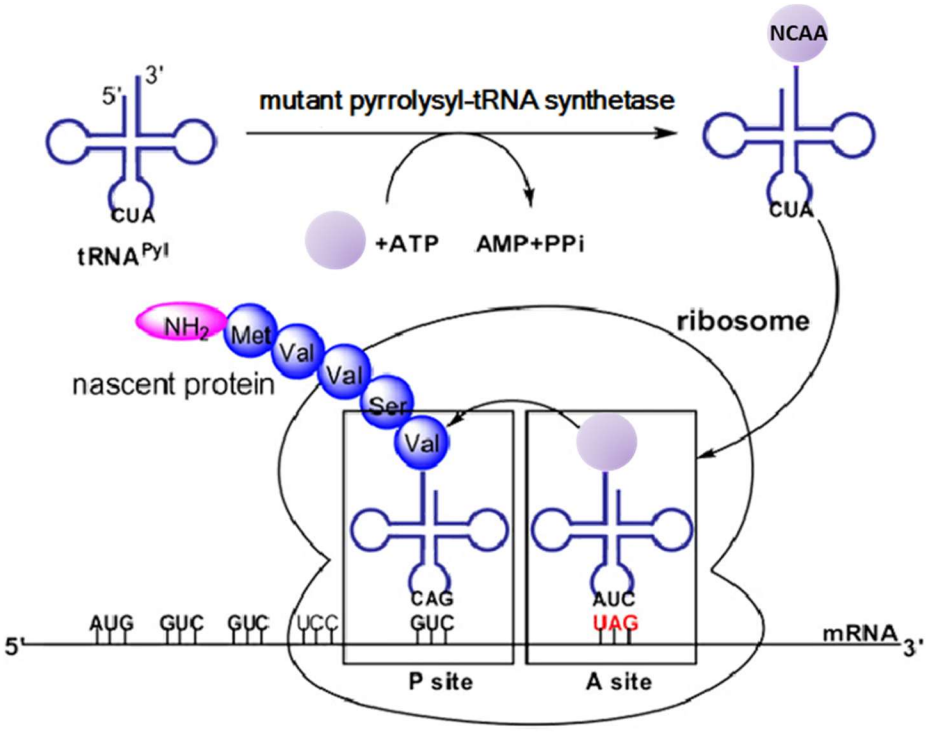
biological function being studied.[22] Low efficiency and yield have hampered the application of semi-synthetic strategy for nucleosome structural studies, which demand larger amount of samples.

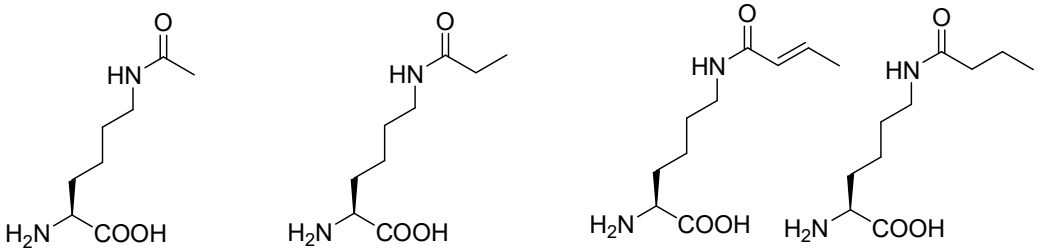
PTM incorporation with amber suppression technique

Over the last decade, genetic code expansion has enabled site-specific incorporation of numerous unnatural amino acids to proteins.[92] One significant contribution of this technique is the generation of proteins with PTM or PTM mimic at desired position, which are indispensable for PTM function elucidation for enzymes and proteins in various biological pathways, especially epigenetics. tRNA and aminoacyl-tRNA synthetase (aaRS) pair, orthogonal to the host endogenous translational machinery, is hijacked for the incorporation of non-canonical amino acid. A wide range of PTM have been either directly incorporated into histones (**Figure 9**), or derived by biorthogonal reaction from pre-installed PTM precursors.[93]

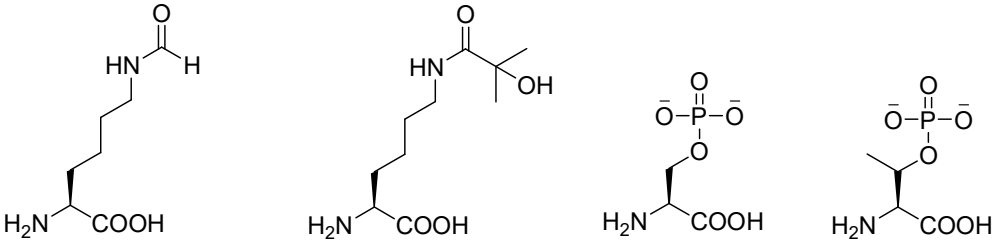
Acetyl-lysine

Genetic incorporation of acetyl-lysine into protein was first enabled by an evolved *Methanosarcina Barkeri* pyrrolysyl-tRNA synthetase (MbPylRS) and its cognate amber suppressor $MbtRNA_{CUA}$ pair. During the library screening, mutations of residues in the binding pocket of MbPylRS tend to favor hydrophobic residues with larger size, possibly due to fact that acetyl-lysine is a substructure of pyrrolysine. At the same time, 20 mM of nicotinamide was added to cell culture to protect nascent acetyl-lysine from removal by endogenous lysine deacetylase CobB.[94] Subsequently, the same system was utilized to incorporate acetyl-lysine to Histone H3 at lysine 56 in the





Acetyl-lysine by AcKRS | Propionyl-lysine by PrKRS | Butyryl- and Crotonyl-lysine by w.t. PylRS



Formyl-lysine by ForKRS | 2-hydroxyisobutyryl-lysine by HibKRS | Phosphoserine by SepRS | Phosphothreonine by ThrRS

Figure 9 Direct incorporation of Non-canonical amino acid (NcAA) by mutant tRNA synthetase

the same time, 20 mM of nicotinamide was added to cell culture to protect nascent acetyl-lysine from removal by endogenous lysine deacetylase CobB.[94] Subsequently, the same system was utilized to incorporate acetyl-lysine to Histone H3 at lysine 56 in the middle region, which was a difficult target for native chemical ligation. This protein was assembled into nucleosome and discovered to facilitate 'DNA breathing' compare to unmodified H3,[95] which means the partial and temporal disengagement of DNA from octamer core, which is called DNA breathing Initial attempt of incorporation of acetyl-lysine to histone H4 was unsuccessful, which was later confirmed to be caused by codon bias and mRNA instability of H4 gene. After codon optimizing and switching to an RNAseE mutant BL21 strain, site-specific K16 monoacetylated H4 was prepared with the yield of 600ug per liter culture.[96] Different strategies have also been applied to increase the yield of acetyl-H4, such as increasing H4 solubility by fusing it to H3 gene, and using C321.ΔA strain.[97] Umehara et. al. evolved the library of Methanosarcina mazei PylRS library with a new selection system based on the killing activity of the toxic ccdB gene. Selected mutant PylRS, together with its cognate $tRNA^{Pyl}$, displayed higher incorporation efficiency for aceyl-lysine than the aforementioned MbPylRS system.[98]

Acyl-lysine

Besides acetylation, histone lysine was discovered to be also modified *in vivo* by short-chain acylation. Crotonyl- (CrK), Propionyl- (PrK) and butyryl-lysine (BuK) were readily incorporated at histone H3K9 by wildtype MmPylRS/$tRNA^{Pyl}$ pair, however, tandem MS/MS analysis discovered unmodified lysine at H3 K9 postion, which was

confirmed to be deacylated by endogenous CobB.[99] Evolution of MmPylRS library generated a refined system, leading to successful incorporation of CrK, PrK and BuK to histone H4.[97] Other newly discovered acylations, including ε-N-formyl-[100] and ε-N-2-Hydroxyisobutyryl-lysine[101] were incorporated by the corresponding evolved PylRS/tRNA` pair. Our group recently developed a versatile strategy to generate various acylations through chemoselective Staudinger ligation with azidohomoleucine, which was site-specifically incorporated by an evolved MmPylRS/tRNA pair (**Figure 10**). Succinyl-lysine was *in situ* generated on histone and nucleosome, avoiding cumbersome steps of PylRS library screening.[102]

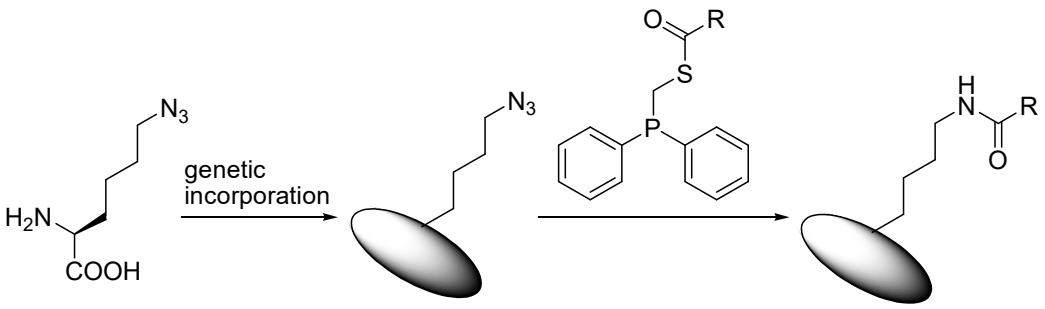


Figure 10 A versatile approach to generate lysine acylation in situ through bio-orthogonal Staudinger Ligation.

Methyllysine

Due to the little structural difference between lysine and methyl-lysine, direct evolution of tRNA synthetase mutant that differentiates between these two was unsuccessful. Methyl-lysine analogs were initially prepared by Michael addition between corresponding thiol reagent and dehydroalanine. MjTyrRS/tRNA^{Tyr} pair was

evolved to site-specifically incorporate phenylselenocysteine, which was transformed into dehydroalanine for thiol conjugation.[103] Our group optimized this strategy by directly incorporating Se-alkylselenocysteine into protein with an engineered MmPylRS/tRNA^{Pyl} pair. The yield was much higher compared to the incorporation of phenyl-selenocysteine.[104] Authentic monomethyl-lysine was introduced into Histone H3 through the incorporation of its Boc protected form N$^{\varepsilon}$-tert-butyloxycarbonyl-N$^{\varepsilon}$-methyl-L-lysine, an efficient substrate for wildtype MbPylRS. Subsequent treatment of dilute TFA afforded monomethylated H3 at K9 position.[105] Alloc protected and photocaged

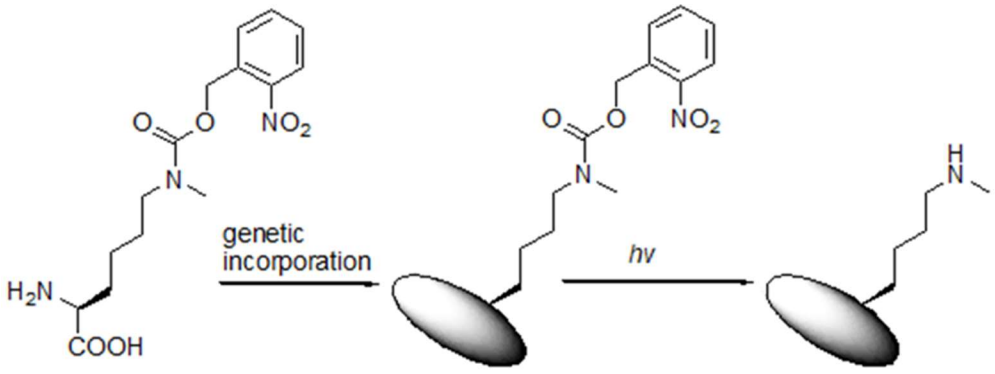


Figure 11 Incorporation of photocaged monomethyl-lysine and subsequent UV deprotection

monomethyl-lysine have also been incorporated with MmPylRS/tRNA^{Pyl} system, and mild deprotection condition was used without affecting other residues (**Figure 11**).[106, 107] As for dimethyl-lysine incorporation, a lengthy synthetic route was designed involving an initial step of protecting all the other lysines with Cbz group, followed by deprotection of Boc-lysine incorporated at the desired position wildtype PylRS, and dimethyl group was introduced via reductive amination on the only free amino group.[108]

The harsh reaction condition and challenging protection and deprotection cycle of multiple lysine residues have prompted us to design much more concise and cleaner synthetic route for dimethyl-lysine incorporation. A precursor of Allysine was incorporated by an evolved MmPylRS, and then dimethyl-lysine was generated through the reductive amination reaction between the released aldehyde group with dimethyl-amine (**Figure 12**).[109]

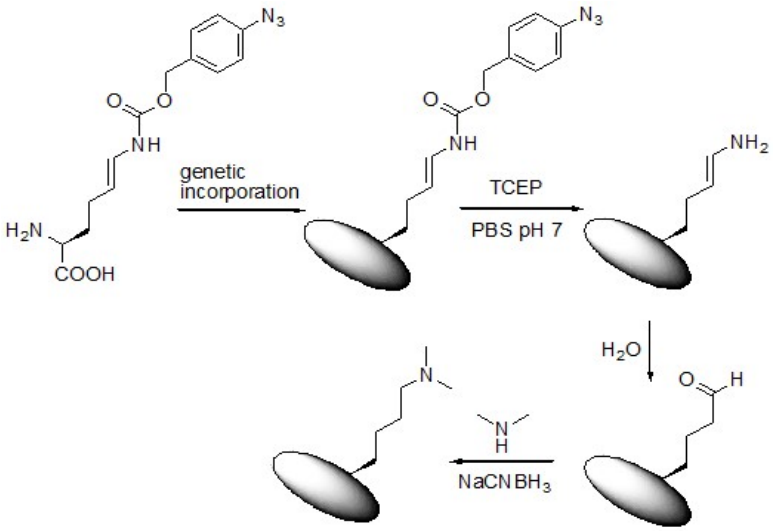


Figure 12 Incorporation of Allysine precursor and subsequence reductive amination to afford dimethyl-lysine

Phosphoserine and phosphothreonine

To genetically incorporate O-phophoserine (Sep),[110] an orthogonal pair $tRNA^{Sep}$-SepRS was engineered from Methanocaldococcus jannaschii (Mj) $tRNA^{Cys}$ and the

mesophilic Methanococcus maripaludis (Mm) Sep-tRNA synthetase (SepRS). Anticodon change to CUA and additional C20U mutation, which generated $tRNA^{Sep}$, retained 40% of its ability to be aminoacylated by Sep compared to original $tRNA^{Cys}$. A negative selection assay discover that simultaneous expression of tRNASep-SepRS pair could not incorporate O-phophoserine into chloramphenicol acetyltransferase in E. coli, due to the failure of native elongation factor EF-Tu to recognize Sep-$tRNA^{Sep}$. After *in vivo* screening assay of EF-Tu library with randomization of 6 residues in the amino acid binding pocket, an active mutant EF-Sep was discovered to enable the incorporation of O-phosphoserine, which was confirmed by successful synthesis of myoglobin and human MEK1 with site-specific serine phosphorylation. Nevertheless, the efficiency is still too low for incorporation of phosphoserine into histone. Lee et. al.[111] further improve the efficiency by optimizing residues at the binding interface between SepRS and $tRNA^{Sep}$, and residues on EF-Sep which were in contact with SepRS. The optimized SepRS9-tRNASep and EF-Sep21 were expressed in E. coli, and facilitated the incorporation of phosphoserine onto histone H3 S10. Milligrams of H3S10ph was purified and assembled into nucleosome for the activity test of GCN5. In another research, Pirman et. al.[112] used an reprogrammed C321.ΔA E. coli strain, in which all 321 TAG stop codon were mutated to TAA stop codon and release factor RF1 was knocked out to enable higher read-through by amino acyl tRNA(CUA). Compared to the original system in BL21(DE3), a new combination with previously optimized SeRS9, EF-Sep21 and a new $tRNA^{Sep\text{-}A37}$ mutant was demonstrated to increase the yield of O-phosphoryl GFP by 9 fold in the C321.ΔA strain. Rogerson et. al. further improve O-

phosphoserine incorporation efficiency through optimizing the affinity between $tRNA^{Sep}$ CUA anticodon region and SepRS substrate binding pocket. Distinct from previous studies, $tRNA^{Sep}$ library containing randomized nucleotides surrounding CUA anticodon was also screened, and best mutant displayed almost 10 times higher activity.[113]

To incorporate phosphothreonine, a Salmonella enterica kinase PduX was first expressed in E. coli to convert threonine to phosphorylated form, and a parallel positive selection coupled with deep sequencing discovered a mutant SepRS that favored the binding of phosphothreonine over phosphoserine. A plethora of proteins were incorporated with phosphothreonine including ubiquitin and CdK2.[114]

Phosphotyrosine

Due to the presence of robust endogenous phosphatases, phosphotyrosine can be dephosphorylated in both the free amino acid and protein context. Hence non-hydrolysable phosphotyrosine analog was initially introduced into protein by genetic incorporation.[115] p-Carboxymethyl-L-phenylalanine (pCMF), through a *M. jannaschii* tyrosyl tRNA synthetase (MjTyrRS) mutant and the *M.jannaschii* amber-suppressor tyrosyl tRNA ($MjtRNA^{CUA}_{Tyr}$), was incorporated into Tyr701 of STAT1 and found to partially mimic the homodimer induction effect of phosphotyrosine-701. Another phosphotyrosine analog phosphorimidate was generated by chemoselective Staudinger reaction between phosphite and a phenyl-azide group introduced by unnatural amino acid incorporation technique, and the resulting modified protein was close enough to be recognized by specific phosphotyrosine antibody.[116] Two recent publications resolve the

difficulty of authentic phosphotyrosine incorporation. Luo et. al. increased the uptake of phosphotyrosine by feeding its dipeptide form to E. coli cell culture and the non-

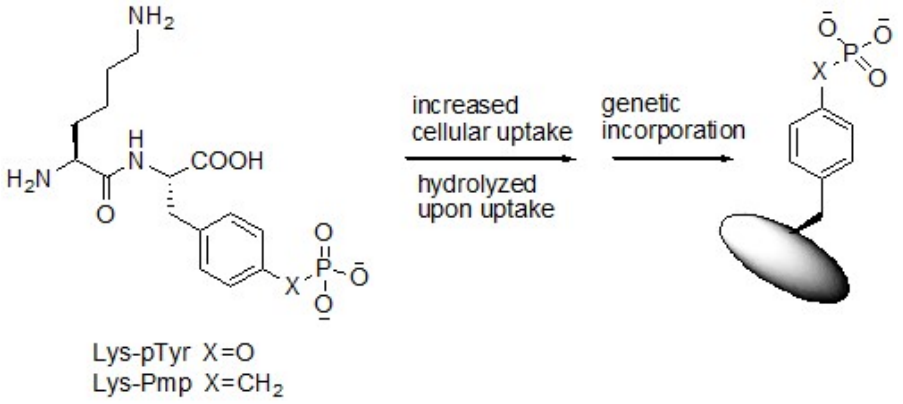


Figure 13 Dipeptide from increase bioavailability of phosphotyrosine and its nonhydrolyzable mimic

hydrolyzable form 4-phosphomethyl-L-phenylalanine was efficiently incorporated into myoglobin with the same MjTyrRS/MjtRNA$^{CUA}_{Tyr}$ pair evolved for pCMF (**Figure 13**). On the other hand, phophotyrosine didn't show up at amber codon position until 1mM sodium orthovanadate was added, which inhibit endogenous phosphatase for the hydrolyzation of newly incorporated phosphotyrsine.[117] Hoppeman et. al. evolved the MmPylRS library and discovered a NpYRS mutant for the specific incorporation of a phosphotyrosine analog containing the phosphoramidate group, which generated authentic phosphotyrosine after hydrolyzation at pH 2 for 36 hours. With this strategy ubiquitin with Tyr59 phosphorylation was produced and discovered to drastically reduce the binding between ubiquintin and E2 ligase (**Figure 14**).[118]

For site-specific ubiquitination, pyrrolysine analog D-Cys-ε-Lys was incorporated into calmodulin with a 1, 2-amino thiol functionality, which could act as the anchor for native chemical ligation with ubiquitin (1-76) α-thioester. Final ligated

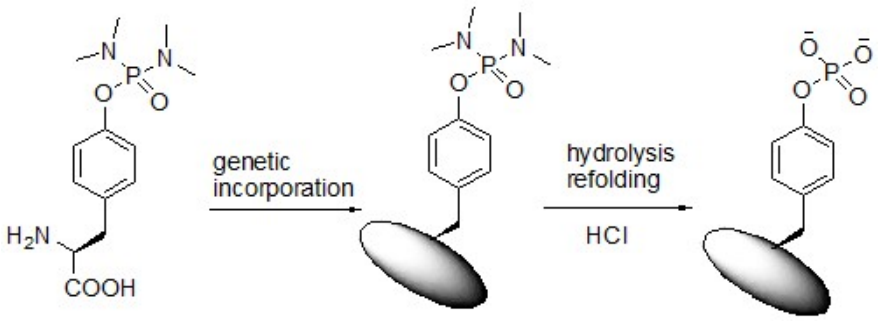


Figure 14 Incorporation of phosphotyrosine analog and subsequent deprotection

protein contained an artificial D-cysteine at the original ubiquitin G77 position.[119] Completely native isopeptide bond between target protein and ubiquitin was constructed with the help of genetic incorporation of δ-thiol-L-lysine precursor δ-thiol-N^{ε}-(p-nitrocarbobenzyloxy)-lysine, which was a substrate for an evolved MbPylRS/$tRNA^{Pyl}$ pair. After deprotection, the released 1,2-aminothiol chemoselectively reacted with ubiqtuin α-thioester to form a ubiquitin dimer. This technique has the potential for traceless synthesis of ubiquitinated histone (**Figure 15**).[120]

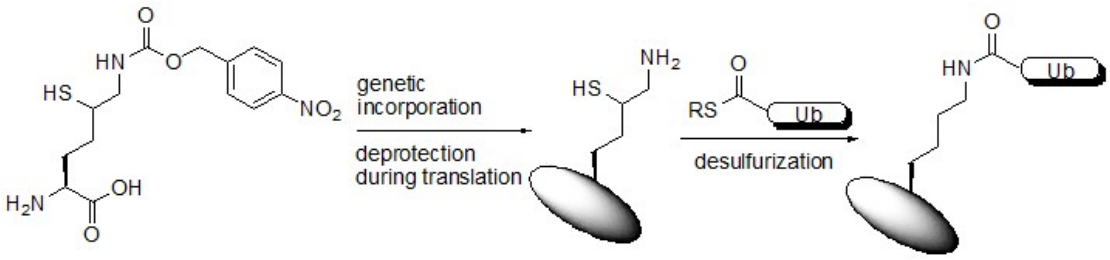


Figure 15 Genetic incorporation of δ-thiol-L-lysine precursor and subsequent chemical transformation to afford ubiquitinated protein

Biochemical studies with modified nucleosome substrates

PTM effects on chromatin structure and dynamics

Besides influencing biological processes through recruiting other non-histone proteins, PTMs on histone N-terminals could directly modulate chromatin conformation and dynamics by interfering with interactions between DNA, histone tails and octamer core,[121] which participate in higher order structure formation. Of the four histone tails, H4 N-terminal tail region from 14 to 19 is uniquely important in chromatin fiber organization.[122] H4K16ac, a prevalent modification marking active gene, happens to be located at this region.[123] Shogren-Knaak et. al.[124] synthesized full length histone H4K16ac with one artificial R23C mutation using native chemical ligation. N-terminal peptide containing K16 acetylation and C-terminal thioester was synthesized by SPPS, and ligated with recombinant C-terminal fragment H4 (C23 to 102). Generated mutant histone H4 was assembled into 12-mer chromatin arrays, and the one containing H4K16ac required much higher concentration of $MgCl_2$ for 30-nm fiber and higher chromosomal domain condensation. (**Figure 16**) At physiological salt concentrations histone N-terminal tails were believed to fold back towards nucleosome core, and

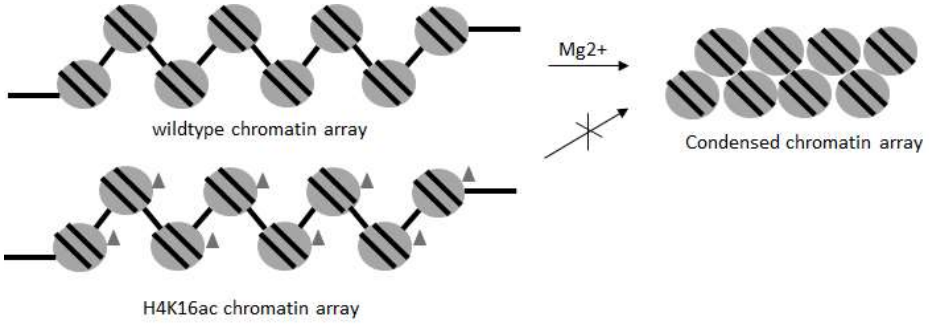


Figure 16 H4K16ac inhibits chromatin array condensation

constrained by interactions with nucleosomal DNA.[125] Acetylation on N-terminal tail may potentially weaken DNA-histone tail interaction and render histone tail more accessible. Tencer et. al.[126] reported that H3K9Ac facilitated binding of ATP-dependent remodeler CHD3 to nucleosome by rendering histone tail more accessible. Histone H3 containing acetylated lysine analog was prepared via cysteine mediated conjugation. H3 N-terminal tail accessibility was evaluated by measuring reaction rate between 5-fluorescein-maleimide and cysteine introduced at two different positions on H3 N-terminal tail, H3 S10 and H3 A21. Acetylation mimetic at H3 K9 resulted increased reaction rate at both positions, indicating a more open conformation was adopted by N-terminal tail.

Histone residues located at the entry-exit site of nucleosome directly interact with nucleotide. PTMs at these sites could potentially interfere DNA breathing.[127] H3K56ac is one of most studied modification at this region, its level has been correlated with transcription regulation,[128] DNA repair[129] and DNA replication[130] Given its numerous important biological functions, the exact function and detailed regulation mechanism were not clear, until the several recent successful attempts were reported to synthesize homogeneous full length protein and nucleosome substrates.[87, 95, 131] Lysine 56 is located in the middle of H3 sequence, and was a difficult object for chemical ligation strategy. Neumann et. al.[95] reported production of full length synthesis of K56 acetylated H3 and nucleosome using genetic incorporation approach. The effects of H3K56ac on nucleosome stability was firstly investigated using a FRET dye labelled substrate, a Cy3

donor was installed on the 5' end of a 147bp nucleosome positioning DNA, the Cy5 acceptor was conjugated to H2AK119C via thiol-maleimide coupling. No FRET signal difference is observed between wildtype nucleosome and H3K56ac nucleosome, indicating K56ac had no effect on nucleosome stability. Therefore, single-pair FRET and alternating excitation selection techniques were used to detect subtle conformational change, and K56ac was found to cause increase in DNA breathing and higher DNA accessibility at the entry-exit site (**Figure 17**). Chromatin fiber formation is also tested but no change was observed between wildtype and H3K56ac nucleosome. On the other hand, H3K56ac was able to increase remodeling rate by 40%.

Using mass spectrometry, Lysine acetylation has been discovered on another histone H3 residue K64, which is also located around the nucleosome entry-exit region.[132] H3K64 is in close contact with nucleosomal DNA and acetylation on this site may transform nucleosome into a more open conformation by interrupting DNA-histone interaction. To investigate its effect on nucleosome stability and remodeling rate, full-length H3K64ac was prepared using amber suppression technique and assembled into nucleosome with DNA labelled at both sites 35bp from 3' and 5' end (**Figure 17**). While increasing salt concentration from 0.5 M to 1 M, H3K64ac nucleosome demonstrated higher tendency of self-disassembly. H3K64ac nucleosome was shown to be more sensitive to chd1, an ATP-dependent chromatin remodeler. In another histone eviction assay, chromatin arrays were assembled using unmodified H3, H3K64ac and H3K64R, and more H3K64ac was found to exist in the evicted histone fraction compared to wildtype and H3K64R mutant. All the above data supports H3K64ac's connection with

active euchromatin, which is further supported by Chip-seq analysis.

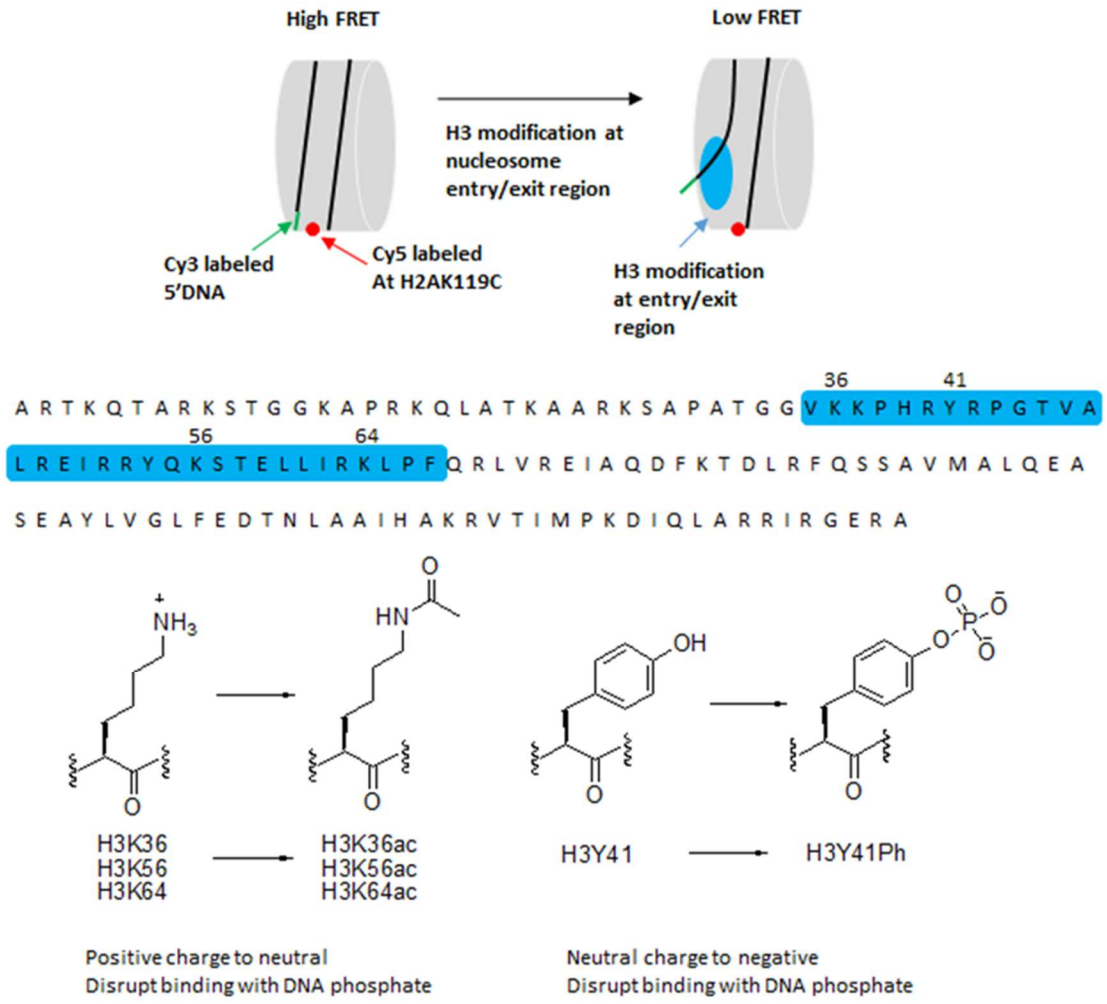


Figure 17 Histone lysine acetylation and tyrosine phosphorylation at nucleosome entry/exit region increase nucleosome DNA breathing by disrupting electrostatic interaction between histone and DNA

Acetyl group on the histone octamer surface possibly weakens the electrostatic DNA-histone interaction by neutralizing the positive charge on lysine residue, serine or threonine phosphorylation, on the other hand, may destabilize DNA-histone interaction

to a larger extent by inducing electrostatic repulsion.[133] Effect of phosphorylation around histone entry-exit site was also evaluated by a full-length H3Y41ph prepared through native chemical ligation (**Figure 17**).[134] H3Y41ph was found to significantly increase DNA accessibility and DNA unwrapping. Moreover, combined H3K56ac and H3Y41ph modifications together increased DNA accessibility over an order of magnitude, indicating that synergistic effect was important for modulating nucleosome structure.

In budding yeast, SWI/SNF histone remodeling complex is involved in the regulation of a vast number of genes.[135] On the other hand, defects caused by loss of SWI/SNF function could be reversed by several histone H3/H4 mutations (SIN mutation) around the nucleosome dyad region,[136] where the Histone-DNA interaction was found to be the strongest.[137] Mutations or PTM around these region could alter nucleosome stability and remodeling motility.[138] Increased nucleosome repositioning,[139] higher susceptibility to DNase[140] and impaired chromatin folding[141] were observed for histone SIN mutations.

H3K115 and H3K122 are located between two SIN motifs, and are positioned to have direct electrostatic interaction with DNA backbone, histone mutants H3K115Q and H3K122Q mimicking lysine acetylation showed reduced transcription silencing at ribosomal DNA and telomeres.[142] The precise role of these mutations on nucleosome stability and dynamics required the synthesis of H3K115ac and H3K122ac nucleosome (**Figure 18**). Manohar et. al.[27] synthesized full length H3K115ac and/or H3K122ac using

the expressed protein ligation strategy. Nucleosome competitive reconstitution assays showed that H3K115ac and/or H3K122ac reduced the free energy of histone octamer binding to DNA. On the other hand, restriction enzyme digestion assay implied that accessibility of DNA around nucleosome dyad region is not affected by H3K115ac or/and H3K122ac. Thermal shifting experiments were carried out with nucleosome assembled with mp2 nucleosome positioning sequence, and nucleosome with H3K122ac demonstrated a slower shifting rate compared to unmodified nucleosome. Nucleosome stability could also be evaluated with magnetic tweezer experiment,[143] in which nucleosome array was extended by an external force and disassembly of nucleosome particle would change the force required to stretch chromatin array. H3K115ac and H3K122ac full length protein were prepared by EPL,[87] and the corresponding chromatin exhibited higher fraction of histone eviction following magnetic tweezer stretch.

H3 T118 was located around the SIN mutation region and T to I mutation caused higher nucleosome motility[144] and DNA accessibility.[136] North et. al.[145] synthesized full length H3T118ph by expressed protein ligation, and found H3T118 dramatically decreased DNA-histone binding, increased SWI/SNF remodeling, and enhanced hMSH2-hMSH6 induced nucleosome disassembly (**Figure 18**). Besides canonical nucleosome particle, H3T118ph protein was also found to induce the formation of two alternative histone-DNA complexes, a nucleosome duplex with two DNA molecules wrapped around two histone octamers, and an altosome complex that contains one DNA molecule wrapped around two histone octamers. *In vivo* evidence is still missing for

these irregular histone-DNA complexes, but possibility is not ruled out that certain histone PTM has the potential to completely alter DNA-histone binding mode.[86]

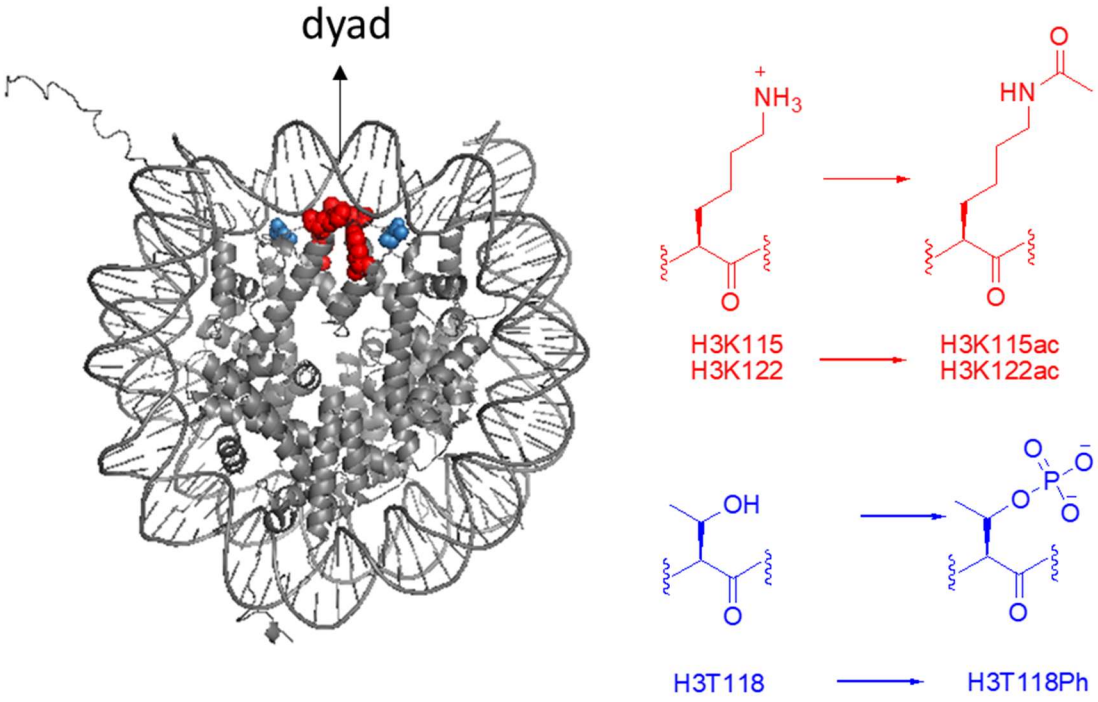


Figure 18 Histone H3K115ac, H3K122ac and H3T118Ph at the nucleosome SIN region disrupt DNA Histone interaction around dyad position

The histone octamer surface surrounding H3 K79, has been reported to be important for the maintenance of repressive heterochromatin.[146] Histone mutations at these regions (72 to 83 for H3 and 78 to 81 for H4), are thus called LRS mutations, as they caused the loss of silencing of rDNA in the genetic screening.[147] Dimethylation or trimethylation of H3K79 (H3K79me2/H3K79me3) has been detected in the transcribed regions of active genes.[148] Lu et. al.[51] prepared $H3K_C79me2$ full length histone analog, and obtained crystal structure of this methylated nucleosome. Global nucleosome

structure was not altered, but the added demethylation on K79 divert the unmodified the lysine from the original position interacting with H4 via weak hydrogen bond, to a new position completely exposed to solvent, and another hydrophobic pocket made up by H4 V70 and H3 L82 was partially uncovered. Modified nucleosome surface could potentially affect binding of proteins involved in transcription regulation. The PTM on other LRS residues, such as H4K77ac and H4K79ac were found to be simultaneously acetylated by high-resolution mass fingerprinting.[149] Full length histone H4 di-acetylated at H4 K77 and H4 K79 was found to increase the DNA accessibility around the nucleosome entry-exit site, but didn't affect overall nucleosome stability.[87]

Discovery of epigenetic enzyme functions

During replication process, newly synthesized histone H3 is pre-acetylated at K56 prior to deposition onto DNA.150 Histone chaperones such as CAF-1, Asf-1 and Rtt106 play important roles in replication-coupled nucleosome assembly pathway,151 but the underneath molecular mechanism had remain unclear. With the full length H3K56ac protein prepared by genetic incorporation, Andrews et. al.152 demonstrated H3K56ac directly affected the binding affinity between histone H3-H4 complex and DNA (Figure 19). Thermodynamic parameters were calculated by monitoring the FRET signal change between histone H3-H4, histone chaperone NAP-1, H2A-H2B dimer and DNA. H3K56ac was discovered to cause a free energy drop of 1.8 kcal/mol for the binding between H3-H4 tetramer and DNA. Once the tetramer was deposited onto DNA, H3K56ac had no effect on the additional deposition of H2A-H2B dimer. In budding yeast, newly synthesized H3-H4 is associated with histone chaperone Asf1 as a

heterodimer, which is presented to histone acetyl-transferase Rtt109 for H3 K56 acetylation.153 The later deposition of H3K56ac-H4 complex to DNA is performed by one of the two histone chaperones, CAF-1 or Rtt106. Winkler et. al.154 synthesize full-length H3K56ac via genetic incorporation approach, and discovered a 2-fold increase of binding between CAF-1 and H3K56ac-H4 complex compared to unmodified H3-H4 complex. Although H3-H4 complex was found to exist in an equilibrium of heterodimer and tetramer, binding ratio was constant between Histone and CAF-1 regardless of the form of H3-H4 complex. Su et. al.155 produce full-length H3KC56ac analog via the cysteine conjugation approach. Binding parameters measured by ITC revealed that H3KC56ac resulted in enhanced binding affinity to Rtt106 histone chaperone. Rtt106 exists as a homodimer and interacts with H3-H4 complex through a two-site binding

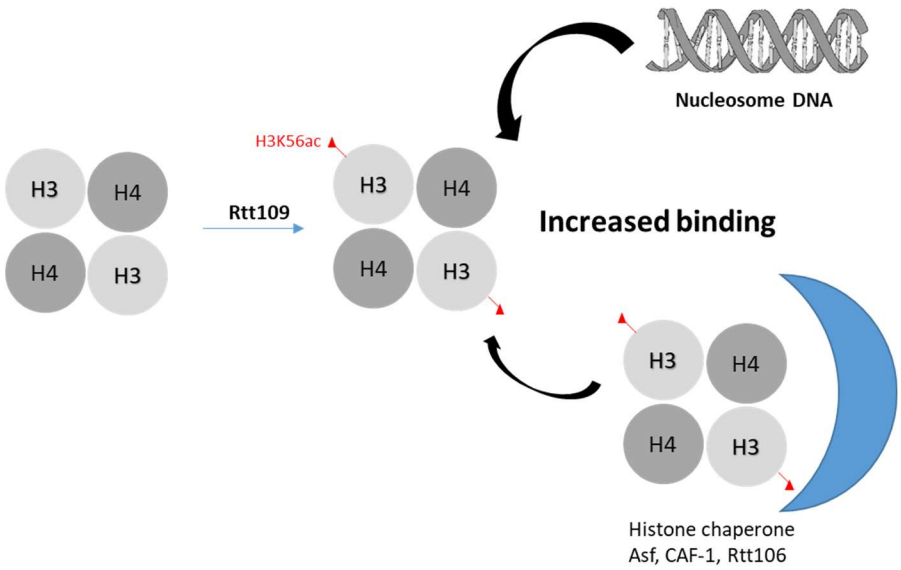


Figure 19 H3K56ac increase binding between DNA, Histone chaperone and H3/H4 tetramer

mode, with a similar 1st dissociation constant Kd1 for both K56ac and unmodified H3-H4 tetramer, but a drastically different second dissociation constant Kd2. NMR study confirmed that Rtt106PH domain (residues 68-301), instead of Rtt106DD domain (residues 1-67), binding to H3K56ac peptide with twofold higher affinity compared to non-acetylated H3K56 peptide. Larger affinity difference (about 20 fold) on H3-H4 tetramer level may be contributed by the conformational entropy increase of αN helix induced by K56ac.

H3K27 methylation, mediated by PRC2 histone methyltransferase complex, plays important roles in repressive transcription regulation and inactive heterochromatin maintenance.[156] The activity of PRC2 has been reported to be antagonized by a trithorax group protein ash1.[157] ash1 was also reported to be a histone methyltrasferase, its targeted sites seemed promiscuous when peptide resembling histone N-terminal or recombinant histone was used as substrates. Multiple sites including H3 K4 and H4 K20 were all active for ash1 catalytic domain.[158, 159] Therefore the exact molecular mechanism of how ash1 antagonize PRC remains elusive, because ash1 targeted sites included both active and repressive transcription mark.[160] Yuan et. al. clarifies ash1's role in regulating PRC2 activity by testing on nucleosomal substrates.[161] On chromatin level, ash1 showed higher activity compared to free histone substrate, and site-specifically dimethylate H3 K36. Full length histone analog H3K$_C$36me3 was then prepared by cysteine conjugation and assembled into mononucleosome. Pre-installed H3 K36 methylation mark greatly impair PRC2's activity on H3 K27. Chip-seq analysis further confirm that H3 K27 methylation and H3 K36 methylation were spatially

exclusive of each other, supporting the conclusion that antagonizing effect exist between methylations on H3 K27 and H3 K36. Another histone methyl-transferase NSD2 also demonstrated distinct target sites on nucleosome substrates and histone octamer along.[162] H3 K36 turned out to be the only active site for NSD2 on nucleosome substrate, while all four histones underwent varied extent of methylation and H4 K44 was the primary target site for histone octamer. Addition of free DNA could also direct NSD2 to dimethylate H3 K36 (**Figure 20**).

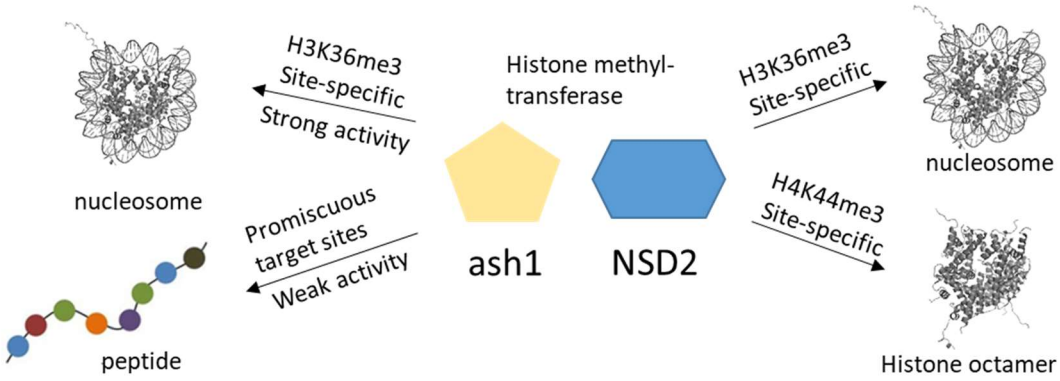


Figure 20 Histone methyl-transferase ash1 and NSD2 shows distinct site-specificities on peptide, histone octamer and nucleosome.

Histone acetyl-transferase GCN5 is one of the first identified histone modifying enzymes,[163] and has been reported to play roles in various biological processes. GCN5 functions in a multi-subunit complex called SAGA and catalyzes the transfer of acetyl group from acetyl coenzyme A to the amino group of histone lysine residues.[36] *In vivo* evidence suggested H3K14ac accumulated on histone H3 following the appearance of H3 S10 phosphorylation on the same N-terminal tail.[164] Peptides spanning residue 1-26 of histone H3 with phosphorylation on serine 10 exhibited 4 to 6 fold higher activity

than unmodified peptide when reacted with yeast GCN5 and [^{3}H] acetyl CoA.[165] This conclusion was further tested with semi-synthetic nucleosome with H3 S10 phosphorylation. A native chemical ligation strategy was designed to combine an N-terminal (1-31) S10 phosphorylated peptide fragment and the remaining C-terminal T32C mutant fragment to afford histone H3 with S10 phosphorylation, which was later used to assemble chromatin array. When reacted with yeast GCN5, phosphorylated nucleosome exhibited 2 fold higher initial velocity compared to unmodified nucleosome, however, when SAGA complex was incubated with phosphoryl nucleosome, no reactivity difference was observed compared to unmodified nucleosome. The failure of H3S10 phosphorylated nucleosome to activate SAGA complex indicated antagonizing effect may exist in interactions between other members of SAGA complex and nucleosome particle.[166]

In another research, Li, et. al.[167] attempted to characterized the kinetics of GCN5 chromatin acetylation, and surprisingly found that the resulting initial velocity data did not fit the standard Michaelis-Menten model, but rather close to a cooperative model. Cooperativity remained when mononucleosome was used as subtrates, but was lost on the H3 peptide substrate, indicating the other copy of H3 N-terminal tail in the same nucleosome particle was indispensable for the cooperativity. This proposition was later confirmed by SAGA kinetic assay on asymmetric nucleosome with one copy of normal H3 and another copy of tailless H3, which showed no cooperativity. SAGA catalytic core GCN5 was reported to contain bromodomain, which has been reported to anchor SAGA complex to hyperacetylated active chromatin region.[168] Bromodomain may play a

role in the observed cooperativity by facilitate binding of SAGA complex to nucleosome via recognizing acetylation mark installed previously. To validate this proposition, histone H3 containing homogenous acetylations on four lysine residues K9, K14, K18 and K23 was synthesized by native chemical ligation, asymmetric nucleosomes with one copy of normal H3 and another copy of tetra-acetylated H3 or another copy of unacetylable H3 were assembled. Consistent with previous result, tetra-acetylated H3 exhibited much higher SAGA activity than unacetylable nucleosome, and mutation Y413A of GCN5 abolish SAGA's cooperativity by impairing bromodomain's binding to acetyl-lysine (**Figure 21**).[169] Other than the bromodomain on GCN5, an H3K4me3

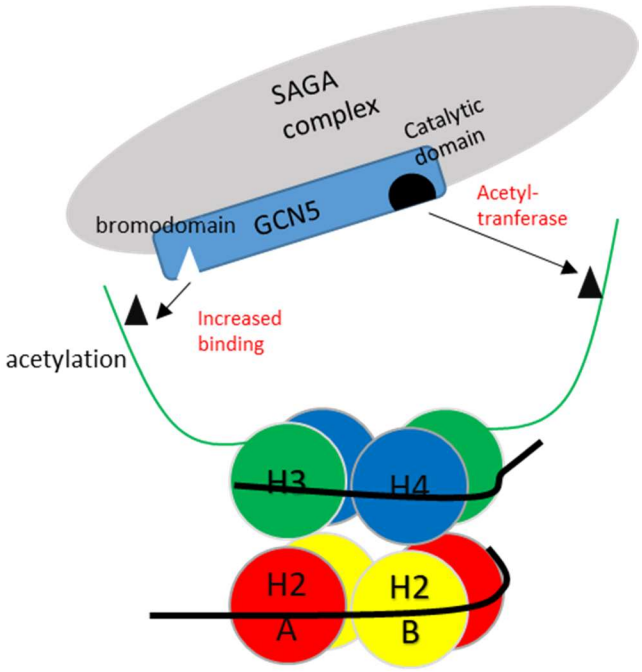


Figure 21 Lysine acetylation on one histone H3 tail promote GCN5 acetyl-transferase reaction on another histone H3 tail

binding domain on one of the SAGA complex member Sgf29, was reported to activate SAGA activity on H3K4me3 nucleosome, which was synthesized through Sortase A mediated protein ligation. In contrast to bromodomain activation in which acetylation on one H3 tail promote acetylation on the other histone H3, H3K4me3 facilitate acetylation on the same H3 N-terminal tail with an in cis mechanism.[170]

H3K9 trimethylation is a crucial mark for repressive heterochromatin conserved through almost all high eukaryotes.[171] Heterochromatin 1 (HP1) has been reported to specifically recognize H3K9me3 on a peptide level *in vitro* and its presence on silenced chromatin was concomitant with H3K9me3 *in vivo* to trigger chromatin condensation.[172] For all HP1 isoforms from *Schizosaccharomyces pombe* to human, a conserved chromodomain (CD) and chromoshadow domain (CSD) were connected by a less conserved hinge domain.[65] Using histone peptides, it was shown that a conserved aromatic cage within the CD binds methylated H3 K9 with low affinity but high sequence specificity and a preference for the trimethylated states.[65] On a nucleosome level, domains other than chromodomain were shown to undertake different roles for different HP1 isoforms. For Drosophila Hp1a, interaction with methylated nucleosome was much lower than binding with corresponding methyl peptide.[173] To obtain H3 K9 methylated nucleosome, unmodified recombinant histone octamer was incubated with Su(VAR)3-9, and 85% percent of the retrieved histone H3 had methylation on K9. No methylation was found in other Histone sites. Drosophila HP1a turned out to have no preference for H3 K9 methylated chromatin over unmodified chromatin, until other HP1a associated proteins such as ACF-1 and SU(VAR)3-9 were added, which interact

with HP1a CSD domain. ACF-1 and SU(VAR)3-9 all had strong binding affinity with chromatin regardless of modification state,[174, 175] indicating a model of HP1a function at specific gene locus was through the recruitment by other chromatin binding proteins. In another study, homogenous H3K9me3 chromatin was generated using EPL.[176] Single-molecule total internal reflection fluorescence microscopy was used to measure the binding dynamics between human HP1α and H3K9me3 chromatin. CSD domain of

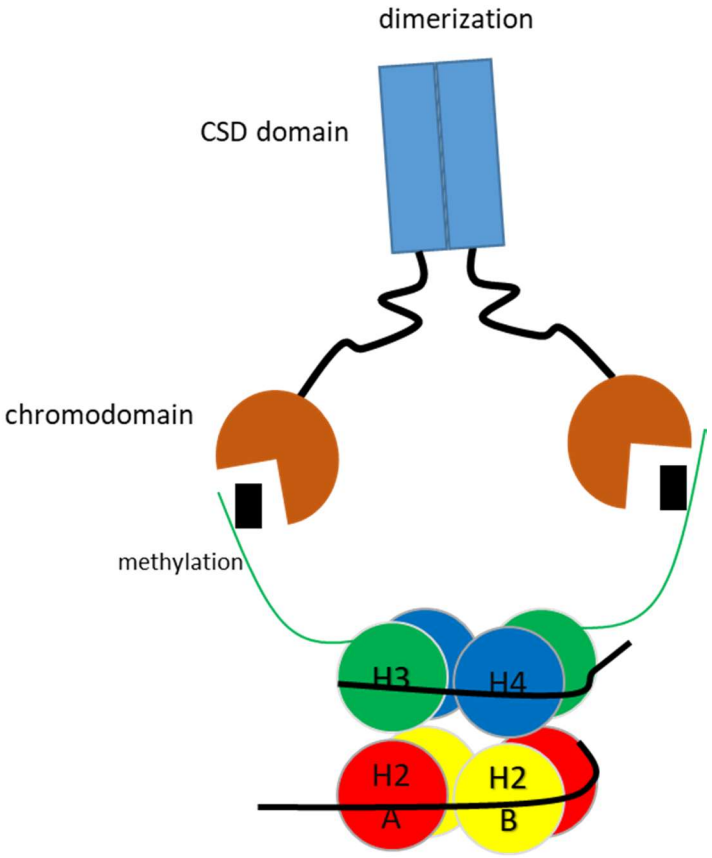


Figure 22 HP1β dimerization mediated by CSD domain facilitates its binding with H3 K9 methyl-nucleosome.

HP1α was found to induce the self-dimerization, which was important to prolong HP1α's retention time on H3K9me3 chromatin. In another study, NMR spectroscopy revealed that human HP1β bond with H3K9me3 nucleosome solely through interaction between chromodomain and trimethylated lysine.[177] CSD domain and linking hinge region showed no participation in binding, as no chemical shift was observed for NMR spectroscopy. Just like human HP1α, CSD domain of human HP1β could also induce self-dimerization. One single mutation on this domain, I161A could impair human HP1β dimerization and more importantly the condensation process of nucleosome array[178] Two bridged chromodomains could simultaneously bind to two H3K9me3 modifications on the same or two different nucleosomes, which stabilized the heterochromatin conformation via dynamic clustering of chromatin fibers (**Figure 22**).

During double strand break repair process, 53BP1 is recruited to DNA damage site via interaction with nucleosome simultaneously ubiquitinated at H2A K15 and H4K20me2.[179] Exact molecular basis of the binding was determined by a Cryo-EM structure of the complex between a dimer of the tandem Tudor-UDR segment of 53BP1, and doubly modified nucleosome, which was assembled using modified histone mimic prepared by cysteine alkylation of a histone H4K20C mutant to yield H4K$_C$20me2 and chemical ubiquitination of H2A(K15C) mutant to yield H2AK15ub.[180] 3D structure information revealed a cooperative binding mode between nucleosome core, UDR domain and ubiquitin. The 53BP1 UDR domain was in contact with both ubiquitin and nucleosome particle, with the N-terminal part located near the H2B-H4 cleft, the C-terminal part interacted with nucleosome H2A-H2B acidic patch, and the aliphatic

residues binding to the hydrophobic patch of ubiquitin (**Figure 23**). Upon UDR binding, ubiquitin adopted a rigid conformation in contact with both H2B αC helix and 53BP1 UDR domain. On the other hand, 53BP1 tudor domain bound with $H4K_C20me2$ in the nucleosome-independent manner.

Bromodomains, which exist in a variety of epigenetic enzymes and complexes, are a group of protein modules recognizing acetyl-lysine through a highly conserved 4 α-

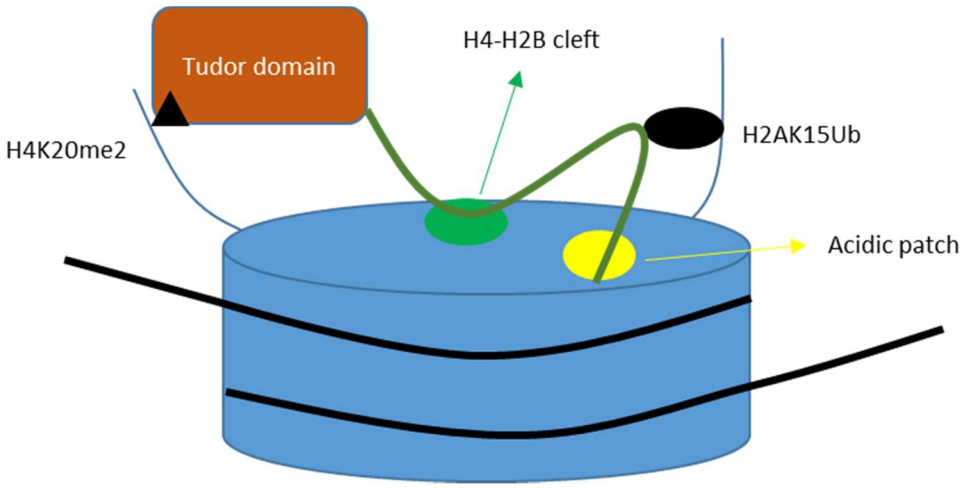


Figure 23 53BP1 binds to H4K20me2 and H2AK15Ub nucleosome with multiple interactions

helices bundle.[181] Researches using acetyl peptides as substrates revealed that bromodomains bind to acetyl-lysine with relatively low specificity and selectivity.[182] On a nucleosomal level, bromodomains appeared to interact with acetyl-nucleosome with more than just acetyl-lysine mark. Miller et. al.[183] synthesized full length histone H3 di-acetylated on K18/K23 and histone H4 di-acetylated on K5/K8 by native chemical ligation. These modified histones were assembled into nucleosomes, and their binding

affinity was measured with BRDT fragments containing either the first bromodomain BD1, the second bromodomain BD2 or both. Results obtained by ITC and methyl-TROSY NMR indicated that only BD1 bound with acetyl-nucleosome via a bivalent interaction mode, which included interactions surrounding the acetyl-lysine binding pocket and additional DNA interactions with basic patches on the bromodomain. Besides BRDT-BD1, other bromodomains also demonstrated binding affinities to DNA molecules, suggesting that in general recruitment of bromodomain to acetyl-lysine is strengthened by extra binding with nucleosomal DNA.

Gene transcription activity has been associated with distinct combinations of PTM.[184] Genome-wide studies demonstrated the spatial connection between active gene and H3K4 methylation and H3 acetylation.[185] But the detailed mechanism of how these PTM coordinate gene transcription has largely unknown. To directly study the functional relevance between PTM and gene regulation, an *in vitro* transcription assay was developed. Firstly recombinant histone was assembled onto template DNA in which multiple copies of 5S rRNA genes flanked transcription binding sequence and TATA box. Then transcription factors and histone modifying coactivators were incubated with the recombinant chromatin, followed by transcription initiation with nuclear extract containing all the other transcription machineries.[186] H3K4me3 was reported to act as one of the anchor to recruit transcription factor TFIID to gene promoter.[187] and it was further illustrated in an *in vitro* transcription assay that chromatin array constructed with full-length H3K$_C$4me3 analog displayed 4-fold higher efficiency compared to array with unmodified H3. *In vitro* chromatin immunoprecipitation analysis showed enhanced

recruitment of PIC (preinitiation complex) component, such as P53, TAF3, RNAPII and TBP around the promoter regions of the chromatin array, which indicated H3K4me3 activate gene transcription by facilitating PIC formation. Installment of H3K4me3 by methyl-transferase SET1 onto active chromatin was later discovered to be synergistically facilitated by transcription factor P53 and co-activator P300,[188] predominantly through direct interaction of SET1-P53 and SET1-P300. A similar assay was also applied to reveal H3R42me2a's (asymmetric arginine dimethylation) activating effect on transcription, possibly through weakening of DNA-histone interaction at entry/exit site.[189]

In vivo presence of H3K122ac was discovered by mass spec in human and bovine histone.[190] Tropberger et. al.[191] produced H3K122ac antibody, and demonstrated by chip-seq analysis that H3K122ac was associated with active gene promoters and enhancers. To investigate molecular mechanism for H3K122ac's transcription activation effect, a full-length H3K122ac was synthesized using genetic incorporation and assembled into chromatin containing VP16-Gal4 binding site for *in vitro* transcription assay. H3K122R chromatin array was used as control with unacetylable arginine residue. When incubated with AcCoA, GAL4-VP16 activator and p300 acetyl-transferase, H3K122ac chromatin showed higher transcription efficiency, possibly through facilitating GAL4-VP16 binding after increased histone eviction from DNA. P300/CBP was identified as the acetyl-transferase responsible for H3K122ac, however, recombinant nucleosome was inactive for P300 *in vitro*, implying other co-activating factors may be required for acetylation at this site.

Summary

Human genome is packed in the form of chromatin, the histone-DNA complex with nucleosome as the basic unit. The accurate interpretation of genome information was carried out in the epigenetics system, which involves multiple levels of regulation such as DNA and histone modification. To study how histone modifications affect chromatin structure and epigenetic enzymes' activity, a fundamental challenge is to synthesize nucleosome substrates site-specifically modified with the PTM. There are three major types of nucleosome synthetic methods, direct biorthogonal reactions, semi-synthetic strategy and genetic incorporation.

Recent studies using *in vitro* assembled intact nucleosome substrates has revealed many new biochemical discoveries that haven't been found with PTM modified peptides. In this thesis, we studied histone deacylase SIRT6 and SIRT7. Both enzymes show very low activity on acetyl or acyl peptide, but are much more active on nucleosome substrates. All previous site-specificity work for SIRT6 and SIRT7 were conducted on acetyl-peptides, which raise the question whether new deacylation targets could be identified using nucleosomes. We answer it in the following two chapter

CHAPTER II

IDENTIFICATION OF SIRT6 DEACYLATION TARGETS USING ACYL-NUCLEOSOMES *

Introduction

As a founding member of NAD^+-dependent histone deacetylases or sirtuins, *Saccharomyces cerevisiae* Sir2 removes lysine acetylation from histones H3 and H4[192] and subsequently impacts genome stability, gene silencing, and yeast lifespan[193]. There are seven Sir2 homologs in mammalian cells, Sirt1-7[194]. Among mammalian sirtuins, Sirt1, Sirt6, and Sirt7 are primarily localized in the nucleus[195]. Their functions have long been considered to remove lysine acetylation from chromatin[196]. However, unlike Sirt1 that displays high deacetylation activities toward acetyl-peptide substrates[197], both Sirt6[198] and Sirt7[199] remove acetylation sluggishly from acetyl-peptide substrates. A recent discovery that Sirt6 has enhanced activities toward fatty acyl-peptide substrates[199] and directly regulates lysine myristoylation of tumor necrosis factor alpha (TNF-α)[200] has led to speculation that Sirt6 serves primarily to remove lysine fatty acylation instead of conventional lysine acetylation. However, both knocking down/out[201, 202] and overexpressing Sirt6[203, 204] in cells have shown clear acetylation pattern changes at H3K9 and H3K56, indicating deacetylation activities of Sirt6 at these two sites. The discrepancy between *in vitro* weak activities of Sirt6 on acetyl-peptide substrates and robust deacetylation of H3K9 and H3K56 by Sirt6 in cells may be due to robust

*Reprinted with permission from "A Chemical Biology Approach to Reveal Sirt6-targeted Histone H3 Sites in Nucleosomes", **Wang, W. W.**, Zeng, Y., Wu, B., Deiters, A., Liu, W. R. *ACS Chem. Biol*. 2016. 11(7), 1973-1981. Copyright (2016), American Chemical Society.

deacetylation of H3K9 and H3K56 by Sirt6 in cells may be due to additional cellular factors that contribute to Sirt6 activation but were missing in *in vitro* analysis.

Indeed, a recent report by Cohen *et al.* showed higher deacetylation activities of Sirt6 on nucleosomes assembled from bulk chicken histones than individual bulk chicken histones themselves[205], implying that the nucleosome structure may activate Sirt6. However, how the nucleosome structure contributes to Sirt6 activation is complicated by the high heterogeneity of bulk chicken histones used in this study, which contain multiple histone isoforms and many modification types and their combinations. To understand how Sirt6 recognizes its acetyl-lysine substrates on the nucleosome level and how the nucleosome structure activates Sirt6 on molecular details, homogenous acetyl-nucleosome substrates are required[45, 90, 166, 206] but have not been applied due to the formidable difficulty of their synthesis using traditional methods. Despite the remaining doubts of Sirt6 biochemistry, mounting evidence has proved critical functions of Sirt6 in metabolism, inflammation, and genome stability[207, 208]. Transgenic mice overexpressing Sirt6 have low LDL cholesterol and triglyceride levels, reduced body fat, improved glucose tolerance, and long lifespan. However, Sirt6-deficient mice show a markedly degenerative phenotype and have short lifespan, hypoglycemia, and defects in DNA repair[209]. Given its role in regulating glycolysis, loss of Sirt6 leads to tumor formation independent of oncogene activation.[209] This tumor-suppressing role is also supported by the fact that Sirt6 is downregulated in several human cancers including pancreatic[210] and colorectal cancers[211]. However, other studies also suggest oncogenic effects of Sirt6. Expression of Sirt6 has been found associated with poor prognosis and

chemosensitivity in patients with non-small cell lung[212], breast[213], prostate[214], and skin cancers[215]. Sirt6 also plays a role in inflammatory pathways, exerting anti-inflammatory effects at the transcriptional level and in bone metabolism, culminating low-turnover osteopenia in Sirt6-deficient mice[216].

It is obvious that Sirt6 plays a number of roles in cells, markedly contrasting to a very few deacetylation targets which have been identified so far. Recently we demonstrated that both Sirt1 and Sirt2 don't have sequence preferentiality when they target lysine for its deacetylation. When assembled in nucleosome, most of native H3 acetylation sites (≥ 10) can be targeted and efficiently deacetylated. Given its high structural similarity with Sirt1 and Sirt2, Sirt6's selective targeting of certain sites is of a surprise. There are many lysine residues in the four histones whose sequence context resembles H3K9 and H3K56. Potentially they are Sirt6-targeted deacetylation sites but have not yet tested. As for H3K56, it is located at the nucleosome core region. Unlike H3K9 that is at the H3 N-terminal tail and flexible when binding Sirt6, the surrounding amino acids of H3K56 have rigid structures. Sirt6 recognition of H3K56 must be dramatically different from that of H3K9. This aspect needs to be addressed using nucleosome substrates that have acetylation specifically installed at H3K9 and H3K56, respectively. In the current study, we showed that Sirt6-targeted sites in nucleosome can be swiftly screened and studied using nucleosome substrates site-specifically installed with a fatty acyl-lysine that can be covalently and fluorescently labeled by a dye.

Experimental section

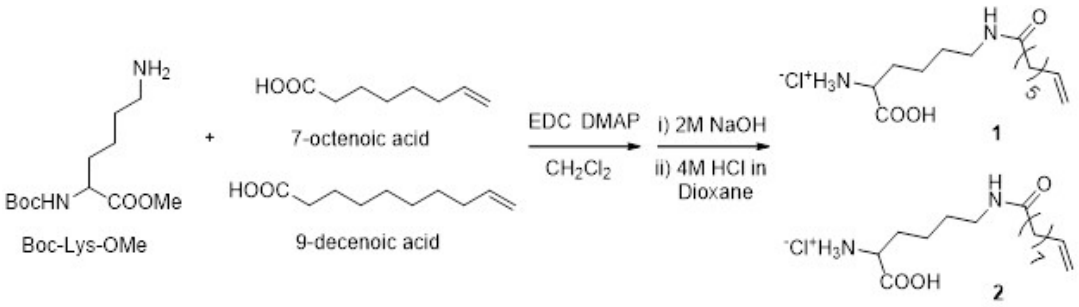


Figure 24 synthetic route of amino acids 1 and 2[276]

Synthesis of N^{ε}-7-octenoyl-L-Lysine 1 and N^{ε}-9-decenoyl-L-Lysine 2. Boc-Lys-OMe was synthesized based on our previously reported method. [217] Boc-Lys-OMe (3.43 g, 13.2 mmol) was dissolved in 40 ml anhydrous CH_2Cl_2, and the solution was injected into a round bottom flask containing 7-octenoic acid (0.94 g, 6.6 mmol) or 9-decenoic acid (1.13 g, 6.6 mmol). EDC (3.26 g, 17 mmol) and DMAP (64.8 mg, 0.53 mmol) were dissolved in 10 ml anhydrous CH_2Cl_2, and injected into reaction flask. The mixture was stirred at room temperature overnight under nitrogen gas flow. And the resulting 7-octenoyl and 9-decenoyl Boc-Lys-OMe were purified by silica gel column chromatography.

7-octenoyl and 9-decenoyl Boc-Lys-OMe were then dissolved in 8 ml THF. 2 M NaOH aqueous solution (20 ml) was added dropwise over 30 minutes into the reaction flasks cooled by ice-water bath. To complete the reaction, the mixtures were stirred in ice-water bath for 1 hour and for another two hours at room temperature. Organic impurities were firstly washed away with ethyl ether (40 ml twice), then pH of the

aqueous phase was adjusted to 3 by adding 3 M HCl. Neutralized products were extracted with ethyl acetate (20 ml twice), and washed with water (40 ml twice) and dried by anhydrous $MgSO_4$. Dried solutions were concentrated to around 5ml by evaporator *in vacuo*, into which 5 ml of 4 M HCl dioxane solution was added. The resulting suspensions were filtered through bucker funnel, and the precipitated products were washed with dry acetone and collected as white solids, N^{ε}-7-octenoyl-L-Lysine **1** (1.87 g 93% yield for two steps) ^{1}H NMR (300 MHz, d6-DMSO) δ 8.43 (s, 3H,), δ 7.84 (t, J=4.5 Hz, 1H,), δ 5.68-5.83 (m, 1H), δ 4.87-5.01 (m, 2H), δ 3.8 (s, 1H), δ 2.98 (q, J=6 Hz, 2H), δ 1.92-2.05 (m, 4H), δ 1.71-1.81 (m, 2H), δ 1.14-1.51 (m, 10H); ^{13}C NMR (75 MHz, d-DMSO):172.36, 171.40, 139.20, 115.13, 52.23, 38.38, 35.79, 33.61, 29.99, 29.01, 28.62, 28.46, 26.68, 22.15). N^{ε}-9-decenoyl-L-Lysine **2** ^{1}H NMR (300 MHz, d6-DMSO) δ 8.46 (s, 3H,), δ 7.86 (t, J=4.5 Hz, 1H,), δ 5.67-5.83 (m, 1H), δ 4.84-5.00 (m, 2H), δ 3.78 (s, 1H), δ 2.97 (q, J=6 Hz, 2H), δ 1.90-2.05 (m, 4H), δ 1.70-1.81 (m, 2H), δ 1.12-1.49 (m, 14H); ^{13}C NMR (75 MHz, d-DMSO): 172,42, 171.35, 139.24, 115.07, 52.23, 38.38, 35.83, 33.60, 29.97, 29.09, 29.04, 28.99, 28.85, 28.67, 25.73, 22.15. (1.97 g, 90% yield for two steps.)

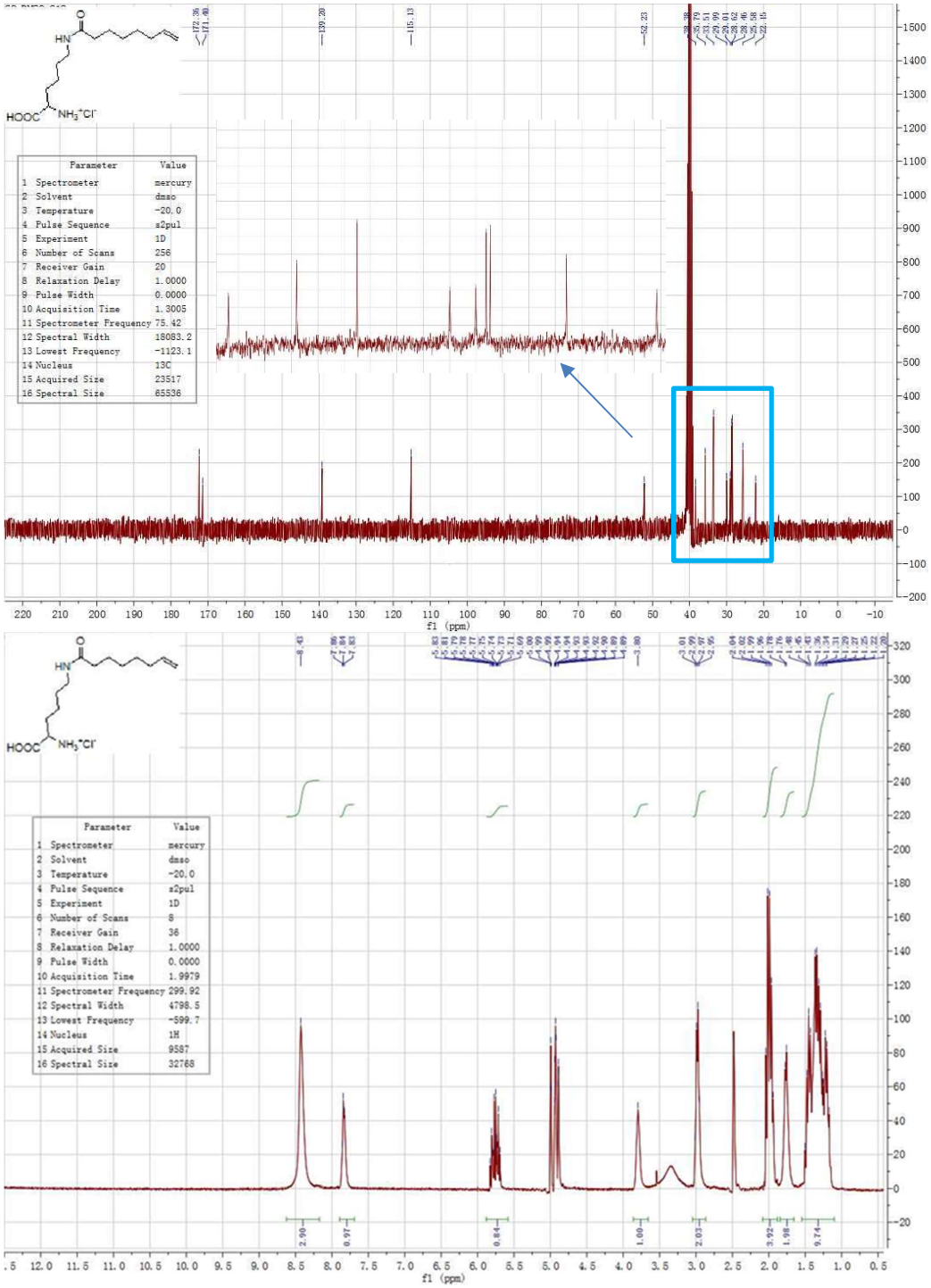


Figure 25 ^{1}H and ^{13}C NMR of Nε-(7-octenoyl)-lysine (OcK) and Nε-(9-decenoyl)-lysine (DeK)[276]

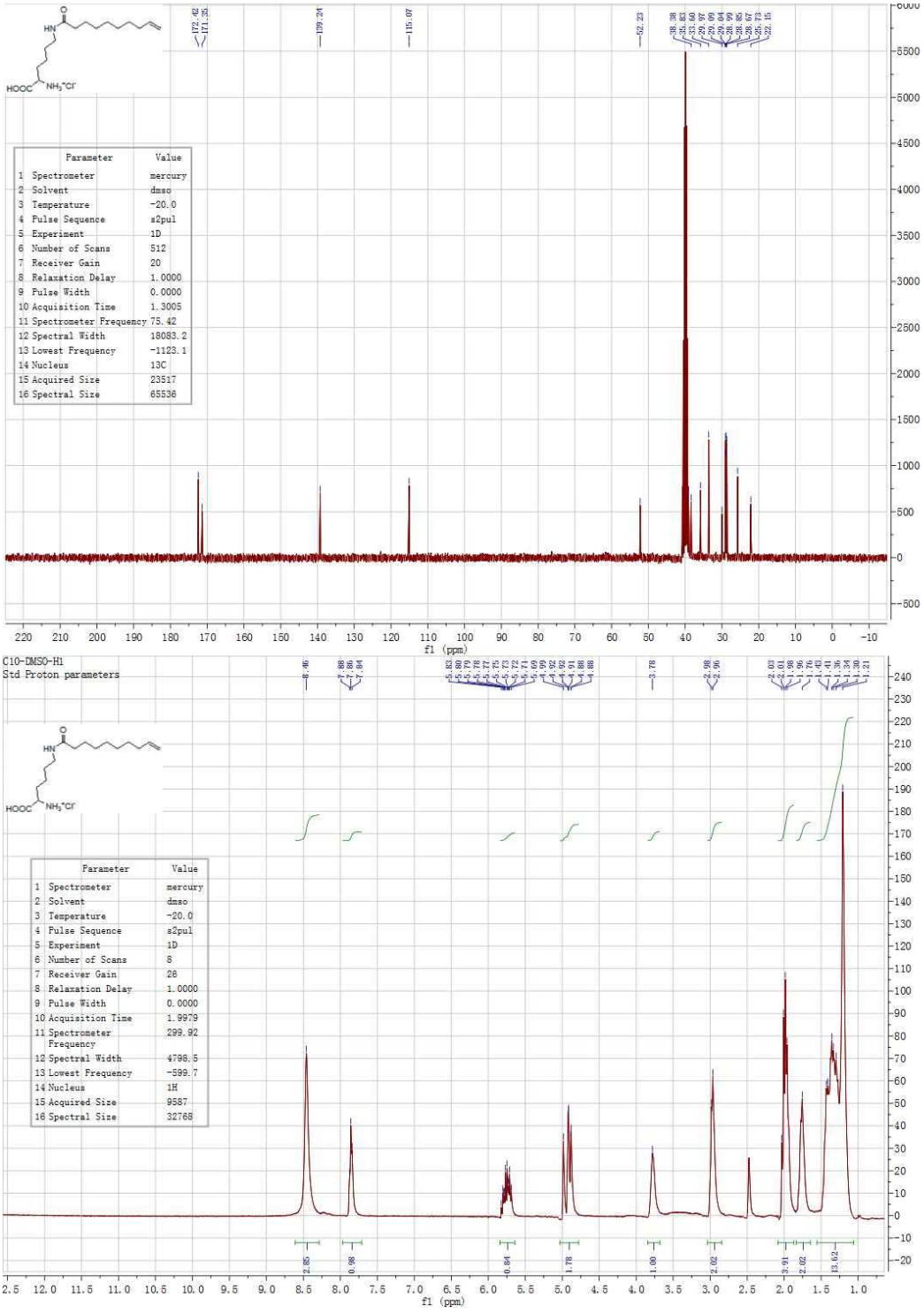


Figure 25 (continued) ^{1}H and ^{13}C NMR of Nε-(7-octenoyl)-lysine (OcK) and Nε-(9-decenoyl)-lysine (DeK)[276]

Screening against mmPylRS mutants for N^{ε}-7-octenoyl-L-Lysine incorporation. Library with active-site mutations of the *Methanosarcina mazei* PylRS gene was constructed by site mutagenesis PCR which has randomization at active site residues (Y306NNK, L309NNK, C348NNK, Y/F/W384). NNK (N=A or C or G or T, K=G or T). The following pairs of primers were used to generate the mmPylRS gene library, pBK-mmPylRS-348NNK-F, ACC ATG CTG AAC TTC NNK CAG ATG GGA TCG GGA TGC ACA CGG, pBK-mmPylRS-348NNK-R, AAA CTC TTC GAG GTG TTC TTT GCC GTC GGA CTC; pBK-mmPylRS-306-309-NNK-F, CTT GCT CCA AAC CTT NNK AAC TAC NNK CGC AAG CTT GAC AGG GCC CTG CCT, pBK-mmPylRS-306-309-NNK-R, CAT GGG TCT CAG GCA GAA GTT CTT GTC AAC CCT. Positive selection: library DNAs were transformed into TOP10 electrocompetent cells containing plasmid pY+ to yield a cell library greater than 1×10^9 cfu, ensuring complete coverage of the pRSL library. Cells were plated on minimal agar plates containing 12 μg/mL tetracycline (Tet), 25 μg/mL kanamycin (Kan), 102 μg/mL chloramphenicol (Cm) and 1 mM NAA. After incubation at 37 °C for 72 h, colonies on the plates were collected and pRSL plasmids were extracted. Negative selection: the extracted plasmids from the positive selection were transformed into TOP10 electrocompetent cells containing plasmid pY- and plated on LB agar plates containing 50 μg/mL Kan, 200 μg/mL ampicillin (Amp), 0.2% arabinose with or w/o 1mM NAA. After incubation at 37 °C for 16 h. Survived cells from plates w/o NAA were pooled to extract plasmids for further selections. The alternative selections were repeated. Final

positive selected colonies grew on LB plates with 102 μg/mL Cm, 25 μg/mL Kan, 12 μg/mL Tet, and with or w/o 1 mM NAA.

Incoporation of N^{ε}-7-octenoyl-L-Lysine into K48 of Ubiquitin. pETDuet vector containing C-terminally His-tagged Ubiqtuin (K48 codon was mutated to amber stop codon TAG) was co-transformed into E. coli BL21(DE3) with a pEVOL vector containing mmPylRS-YASY. Single colony was picked and inoculated in 2YT medium (100 ug/ml Ampicillin 34 ug/ml chloramphenicol). When OD reached to 0.6, 0.5 mM IPTG, 5 mM Nicotinamide, 0.2% (w/v) L-arabinose and 1 mM N^{ε}-7-octenoyl-L-Lysine were added into cell culture. Cells were harvested 5 hours later (expression at 37 ℃) after induction and sonicated in lysis buffer (50 mM sodium phosphate, 250 mM NaCl, 0.1% tritonX-100, Ph 7.4). Ub48OcK was purified by Ni-NTA affinity resin. (15 mg/L, concentration determined by BCA assay) Sample for ESI-MS analysis was prepared by dialyzing into 20 mM NH_4HCO_3 (4℃, pH 7.0) buffer and dried to pellets using the Genevac EZ-Bio Evaporator System. Ub48OcK and w.t. Ubiquitin were both incubated with 0.2 mM pyrimidine-tetrazine-FITC (1XPBS buffer, Ph 7.4) for 4 hours at room temperature. 100% (w/v) TCA were added into reaction buffer (1:4 v:v), and the resulting suspensions were centrifuged at 13000g and 4 ℃ for 20 mins (Eppendorf AG minispin). The resulting protein pellets were washed with 100% acetone for two times to get rid of residual dye. 15 ul 8 M urea buffer (8 M Urea, 20 mM Tris, 500 mM NaCl, pH 8.0) was added to dissolve protein pellets and samples were subjected to SDS-page gel analysis, FITC signals were blotted using ChemiDoc XRS+ system from Bio-Rad, which were compared with coomassie blue staining results.

Incorporation of N^{ε}-7-octenoyl-L-Lysine into Histone H3. pETDuet vectors encoding mutant His-TEV-H3-C110A were co-transformed with Pevol-mmPylRS-YASY into E. coli BL21(DE3) strain. Single colony was picked and inoculated in 2YT medium (100 ug/ml Ampicillin 34 ug/ml chloramphenicol) at 37 ℃. When OD reached to 0.6, 0.5 mM IPTG, 5 mM Nicotinamide, 0.2% (w/v) L-arabinose and 1 mM N^{ε}-7-octenoyl-L-Lysine were added into cell culture. Cells were harvested 10 hours later (expression at 37 ℃) after induction and sonicated in lysis buffer (50 mM Tris, 100 mM NaCl, 0.1% Triton-X100, pH 8.0). Inclusion bodies were collected and washed 3 times with lysis buffer and 3 times with pellet wash buffer (50 mM Tris, 100 mM NaCl, pH 8.0). 6 M guanidinium chloride buffer (6 M guanidinium chloride, 20 mM Tris, 250 mM NaCl, pH 8.0) was then added to dissolve inclusion bodies. Suspensions were then centrifuged. (11000g, 40 min, 4 ℃, Beckman Coulter Allegra X-15R Centrifuge) and supernatants were collected and run through Ni-NTA purification under denatured condition. Purified Histone H3 mutants were dialyzed into MilliQ water at room temperature and dried to pellets by the Genevac EZ-Bio evaporation system, (yield ranging from 5 mg/L to 19 mg/L) concentrations were determined by Pierce BCA protein assay kit).

Expression and purification of Gst-Sirt6. Human Sirt6 fragment was cloned from PQE-80l-Sirt6 (addgene #13739) into PGEX4t-3 vector with an N-terminal GST-tag. And the plasmid was transformed into BL21(DE3) strain. Single colony was picked and inoculated in LB medium (100 ug/ml ampicillin) at 37 ℃. When OD reached to 0.6, cell culture were cooled down in ice-water bath before 0.5 mM IPTG was added. Cell

culture was further incubated at 16 ℃ for 20 hours. Cell pellet was then resuspended in lysis buffer (25 mM Tris, 250 mM NaCl, 1 mM PMSF, 1 mg/ml lysozyme, 1 mM EGTA, 10 mM DTT, pH 7.5). After sonication, cell lysate was centrifuged (11000g, 40 min, 4 ℃, Beckman Coulter Allegra X-15R Centrifuge) and supernatant was collected and run through GST affinity resin at 4℃. Elution was collected and dialyzed into loading buffer (25 mM tris, 20 mM NaCl, 2.5% glycerol, 1 mM DTT, pH 7.5), and purified by Q sepharose FF column FPLC (Buffer A: 20 mM Tris, 2.5% glycerol, 1 mM DTT, pH 8.0, Buffer B: 20 mM tris, 2.5% glycerol, 1 mM DTT, 1 M NaCl, pH 7.5). Purified Sirt6 was concentrated using Amicon Ultra-4 Centrifugal Filter Units and dialyzed into storage buffer (25 mM Tris-HCl, pH 8.0, 100 mM NaCl, 0.05% Tween-20, 1 mM DTT and 20% glycerol pH 8.0) and stored in -80 ℃. (final concentration 15.3 uM, determined by Pierce BCA protein assay kit)

Expression of recombinant human histone His-TEV-H2A, His-TEV-H2B, His-TEV-H3 and His-SUMO-TEV-H4. pETDuet vectors containing Human His-TEV-H2A, His-TEV-H2B, His-TEV-H3-C110A or His-SUMO-TEV-H4 were transcformed into E. coli. BL21(DE3) strain. Single colonies were picked and inoculated in 2YT medium (100 ug/ml Ampicillin 34 ug/ml chloramphenicol) at 37 ℃. When OD reached to 0.6, 0.5 mM IPTG was added to induce histone expression. Remaining protein purification steps were the same with the purification of recombinant mutant Histone H3.

Assembly of Octenoyl-Histone H3/H4 tetramers. Pellets of His-TEV-H3 mutants and Histone His-SUMO-TEV-H4 were redissolved in 6 M guanidinium chloride buffer (6 M guanidinium chloride, 20 mM Tris, 250 mM NaCl, pH 8.0). Concentrations

were determined by UV absorption at 280 nm (Biotek synergy H1 plate reader) and coefficients found on Expasy. 37.5 ug His-SUMO-TEV-H4 and 25 ug His-TEV-H3 (molar ratio=1:1) were added together and diluted using 6 M guanidinium chloride buffer so that total protein concentrations were adjusted to 2 ug/ul. His-SUMO-TEV-H4 and His-TEV-H3 mixtures were dialysed sequentially at 4 ℃ in 2 M NaCl buffer (2 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5), 1 M NaCl (1 M NaCl, 10 mM Tris, 1mM EDTA, Ph 7.5) and 0.25 M NaCl buffer (250 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5). Suspensions were centrifuged (13000g 10min, Eppendorf AG minispin) at 4 ℃, supernatants were collected. Concentrations of His-TEV-H3/His-SUMO-TEV-H4 tetramers were determined by UV absorption at 280nm (Tetramers were denatured by adding Guanidinium Chloride solids). TEV protease was added into His-TEV-H3/His-SUMO-H4 tetramer solution at ratio of 1:30 (TEV protease:Histones w/w). After incubation at 4 ℃ for 24 hours, TEV digestion was confirmed to be finished by SDS-PAGE electrophoresis.

Deacylation assay on Histone H3/H4 tetramers. H3/H4 tetramer concentrations were adjusted to 0.8 uM by adding 0.25 M NaCl buffer (250 mM NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5). NAD+ (1 mM), DTT (1 mM), GST-Sirt6 (0.4 uM) were also added to tetramer reaction solutions. For each control assay, all ingredients were the same except that 0.25 M NaCl buffer was added instead of GST-Sirt6. All tetramer reactions were incubated at 37 ℃ for 3 hours. 10 mM Nicotinamide and 50 mM Tris at pH 8.0 were added to quench the reaction. 20 mM iodoacetate was added to reaction solutions and the solution was incubated at room temperature for additional half an hour to

remove DTT, which may interfere with tetrazine-FITC labeling. Then 0.2 mM Pyrimidine-Tetrazine-FITC (in 5 mM DMSO stock solution) was added to tetramer solutions. The labeling reaction lasted for 8 hours at room temperature. Then TCA precipitation was applied to get rid of unreacted Tetrazine dye. Protein pellets were redissolved in 8M Urea buffer (8 M Urea, 20 mM Tris, 500 mM NaCl, pH 8.0). Deacylation was monitored by directly detect fluorescent H3 bands from SDS-PAGE gel (using Biorad Chemidoc XRS+).

Assembly of site-specific Octenoyl-Nucleosome. His-TEV-H2A and His-TEV-H2B pellets were redissolved by 6 M Guanidinium Buffer, Concentrations were determined by UV absorption at 280 nm (Biotek synergy H1 plate reader) and coefficients found on Expasy. Assembly procedure of His-TEV-H2A and His-TEV-H2B dimer was the same with His-TEV-H3/His-SUMO-TEV-H4 tetramers except that total protein concentration was 4 ug/ul. Mutant H3/H4 tetramers (6.6 uM), H2A/H2B dimer (13.2 uM) and 601-147bp DNA (6.6 uM) (PCR product of a pUC19-601 DNA vector) were added together. NaCl concentrations were adjusted to 2 M using 5 M NaCl solution. Nucleosome was assembled by serial dilution using TE buffer (10 mM Tris, 1mM EDTA, pH 7.5) based on NEB EpiMark Nucleosome assembly protocol (http://www.neb.com/protocols/2012/06/04/epimark-nucleosome-assembly-kit-e5350). The initial 20 ul of Histone-DNA mixture was incubated at room temperature for 30 minutes. Then 7 ul of TE buffer was added to make NaCl concentration to 1.48 M. After 30, 60 and 90 minutes, 13 ul, 27 ul, and 93 ul TE buffer was added to sequentially dilute

NaCl concentration to 0.25 M. After 30 minutes of incubation at room temperature, assembly efficiencies were monitored by 5% 1XTBE Native PAGE gel shift assay.

Deacylation assay on site-specific Octenoyl-Nucleosome. Concentrations of Octenoyl-Nucleosomes were adjusted to 0.66 uM by adding 0.25 uM NaCl buffer, 1 mM NAD+ and 1 mM DTT. GST-Sirt6 (0.33 uM) was also added to deacylation reaction while 0.25 M NaCl buffer was added to control. All nucleosome reactions were incubated at 37 ℃ for 3 hours. 10 mM Nicotinamide and 50 mM Tris at pH 8.0 were added to quench the reaction. 20 mM iodoacetate was added to reaction solutions and the solution was incubated at room temperature for additional half an hour. Then 0.2 mM Pyrimidine-Tetrazine-FITC (in 5 mM DMSO stock solution) was added to nucleosome solutions. Labeling reactions lasted for 8 hours at room temperature. All reaction solutions were directly loaded onto 8% 1X TBE native PAGE gel. Fluorescence strength of separated nucleosome bands were directly blotted using Bio-rad Chemidoc XRS+, and EtBr staining was applied after fluorescence detection.

Time-based deacylation assay on K4Oc, K9Oc, K18Oc, K23Oc and K27Oc nucleosomes. Octenoyl-Nucleosome reaction solutions were prepared as the following recipe (0.66 uM Nucleosome, 0.33 uM GST-Sirt6, 1 mM NAD+, 1 mM DTT, pH 7.5) and equally distributed to 4 tubes labeled as 0 hour, 0.5 hour, 1 hour, 2 hour. And the reactions were quenched at these different time-points by adding 10 mM Nicotinamide, 50 mM Tris (pH 8.0) and 20 mM iodoacetate. After incubation at r.t. for 30 mins, 0.2 mM Pyrimidine-Tetrazine-FITC (in 5 mM DMSO stock solution) was added to nucleosome solutions to label all 4 reactions stopped at different times. Fluorescence

strength of separated nucleosome bands were directly blotted using Biorad Chemidoc XRS+. EtBr staining was done after FITC blot to reveal the total amount of acylated and non-acylated nucleosome. Signal strengths were quantified using Biorad Image Lab software. Ratios of strengths between FITC signal/EtBr signals were plotted against reaction time.

Sirt6 Deacetylation assay on K9ac, K18ac and K27ac nucleosomes. Site specific Acetyl-Nucleosomes were assembled as the same protocol for Octenoyl-nuclesomes. Deacetylation solution were prepared as the following recipe (0.66 uM Nucleosome, 0.33 uM GST-Sirt6, 1 mM NAD+, 1 mM DTT, pH 7.5). Reactions were quenched by adding 10 mM Nicotinamide, and were displayed by 5% 1X TBE Native PAGE electrophoresis. Nucleosome assemblies were confirmed by EtBr staining and nucleosome bands were then directly transferred to PVDF membrane using Bio-rad semi-dry transmembrane protocol (Bio-Rad Trans-Blot Turbo transfer system). The membrane was coated with 5% fat-free milk (10 mL) for 1 h at room temperature and then treated with pan-acetylation antibody from PTM bio-lab (#PTM-101) overnight at 4 ℃ (1:2000, 5 mL). The membrane then was washed by PBST buffer (PBS with 0.1% tween-20, 10 mL) on the shaker six times with 10 min intervals. Then the membrane was treated with secondary antibody (1: 10000, 5 mL) from Jackson ImmunoResearch at room temperature for 1 h. The membrane then was washed by PBST buffer (PBS with 0.1% tween-20, 10 mL) on the shaker three times with 10 min intervals. And then the result was visualized with Pierce ECL ultra-sensitive Western Blotting Substrate (#37070). Images were taken by ChemiDoc XRS+ system from Bio-Rad.

Sirt6 overexpression in HEK293t cell line and overall deacetylation level detection. Human Sirt6 fragment was cloned from PQE-80l-Sirt6 (addgene #13739) into pEGFP vector. 8ug pEGFP-Sirt6 or pEGFP empty vector were added to 60mm HEK293t cell culture dishes and packed into cell by 20 ul lipofectamine 2000 (#11668027 Thermo Fisher Scientific) reagent. Cell medium was changed after 12 hours of incubation. After 48 hours, more than 70% of HEK293t cells showed fluorescence under Microscope, and cells were scraped off the dish and pelleted by centrifuging at 900RPM for 10 minutes at 4 ℃. Histone extractions was performed based on a modified version of a previous protocol[203]. Cell pellets were re-suspended in lysis buffer containing 10 mM HEPES, 150 mM NaCl, 0.5% (w/v) NP-40, protease inhibitor cocktail (#P8340-5ML Sigma), 5 uM TSA, 10 mM Nicotinamide 1 mM DTT and 1mM PMSF. The suspension was incubated on ice for 30 minutes, and centrifuged at 14000 RPM for 15 min at 4 ℃. Supernatant containing the cytoplasmic components was removed, and histone was extracted from the nuclei precipitates by adding 0.2 N HCl and incubating on ice for 30 minutes. 1 M Tris solution (pH 8.0) was added to neutralize the histone solution, and histone concentration was measured by BCA assay and stored in -80 ℃. For western-blot analysis, 5 ug of nuclear proteins were displayed by SDS-PAGE electrophoresis, transferred to PVDF membrane and blotted by site-specific anti-H3K9ac, anti-H3K18ac, anti-H3K27ac, anti-H3K23ac, anti-wtH3 (Acetyl-Histone H3 Antibody Sampler Kit #9927 Cell Signaling Technology, Inc.) and anti-Sirt6 antibodies. (#A302-452A-T, Bethyl Labotaries)

Primers and gene sequences

Sirt6→pGEX4t-3

F: CAATTGGATCCTCGGTGAATTACGCGGC

R: CAATTCCCGGGTCAGCTGGGGACCGCCT

Sirt6→pEGFP

F: TAAGGGAATTCATGTCGGTGAATTACGCGGCGGGGCTGTC

R: AATTGGCGGCCGCTCAGCTGGGGACCGCCTTGGCCTTC

pEGFP-Sirt1→pEGFP

F: ATCGAATTCGAAGCTTGAGCTCG

R: ATAGAATTCATGGTGAGCAAGGGCGAG

For making 601 nucleosome positioning DNA

F: CTGGAGAATCCCGGTGCCG

R: CGTGTCAGATATATACATCCTGT

His-TEV-H3:

atgggcagcagccatcaccatcatcaccacagccaggatccggaaaatctgtacttccaggctcgcaccaaacagactgctcg

taagtccactggcggtaaagcgccgcgtaaacagctggcaaccaaggcagcgcgtaaaagcgctccagctactggcggcgt

gaagaagccgcaccgttatcgcccgggtactgtggctctgcgtgaaatccgccgctaccagaaaagcaccgaactgctgattc

gcaaactgccatttcaacgtctggttcgcgaaattgctcaggatttcaaaaccgacctgcgcttccagtctagcgctgtgatggca

ctgcaagaggcgtctgaggcatatctggttggcctgttcgaagataccaacctggcagcaatccatgcaaagcgtgtaaccatt

atgccgaaagacatccaactggctcgtcgtatccgtggtgagcgtgcgtga

His-TEV-H2A:

atgggcagcagccatcaccatcatcaccacagccaggatccggagaatctctacttccaatctggtcgtggtaaacaaggtggt

aaagcacgtgcaaaggctaagactcgtagcagccgtgccggtctgcagtttccagtgggtcgcgttcaccgtctgctgcgtaaa

ggcaactatgctgaacgtgtgggtgctggtgcaccggtttacctggcagctgtactggaatatctgaccgcagagattctggag

ctggcaggtaacgcagctcgtgataataagaagacccgcatcatcccacgtcacctgcagctggccatccgcaacgatgagg

aactgaacaaactgctgggcaaagttactatcgctcaaggtggcgttctgccgaacatccaggcagttctgctgccgaagaaaa

ccgaatcccaccacaaagcgaaaggtaagtga

His-TEV-H2B:

atgggcagcagccatcaccatcatcaccacagccaggatccggaaaacctgtatttccagtcagaaccagctaagtctgcacc

ggctccgaagaaaggctctaagaaggctgttaccaaggctcagaagaaagatggtaagaaacgcaaacgttctcgtaaagaa

agctattctgtgtacgtgtataaagttctgaaacaagtacatccagacactggcatttccagcaaagcgatgggcattatgaacag

cttcgttaacgatatcttcgaacgtatcgcaggcgaagcgagccgtctggctcactataacaaacgttctaccatcacctctcgtg

aaattcaaactgcagttcgtctgctgctgccaggtgaactggctaaacacgcggttagcgaaggcactaaagcagttaccaaat

acacttcttccaaatga

His-SUMO-TEV-H4:

atgggcagcagccatcatcatcatcacagcagcggcctggtgccgcgcggcagccatatgtcggactcagaagtcaatc

aagaagctaagccagaggtcaagccagaagtcaagcccgagactcacatcaatttaaaggtgtccgatggatcttcagagatc

ttcttcaagatcaaaaagaccactcctttaagaaggctgatggaagcgttcgctaaaagacagggtaaggaaatggactccttaa

gattcttgtacgacggtattagaattcaagctgatcagacccctgaagatttggacatggaggataacgatattattgaggctcac

agagaacagattggtggtggatccgaaaatctgtacttccagtctggtcgtggtaaaggtggtaaaggcctgggtaaaggtggt

gctaagcgtcaccgtaaagtgctgcgcgacaacatccagggtatcaccaaaccagctattcgccgtctggcacgtcgcggtgg

tgtgaaacgcatcagcggtctgatctatgaagaaacccgtggtgttctgaaagtatttctggagaacgttatccgcgatgcggtg

acctacaccgaacacgcgaaacgtaagaccgttactgctatggatgttgtgtacgctctgaaacgccagggtcgtactctgtac

ggtttcggtggctga

GST-SIRT6:

atgtcccctatactaggttattggaaaattaagggccttgtgcaacccactcgacttcttttggaatatcttgaagaaaaatatgaag

agcatttgtatgagcgcgatgaaggtgataaatggcgaaacaaaaagtttgaattgggtttggagtttcccaatcttccttattatat

tgatggtgatgttaaattaacacagtctatggccatcatacgttatatagctgacaagcacaacatgttgggtggttgtccaaaaga

gcgtgcagagatttcaatgcttgaaggagcggttttggatattagatacggtgtttcgagaattgcatatagtaaagactttgaaac

tctcaaagttgattttcttagcaagctacctgaaatgctgaaaatgttcgaagatcgtttatgtcataaaacatatttaaatggtgatca

tgtaacccatcctgacttcatgttgtatgacgctcttgatgttgttttatacatggacccaatgtgcctggatgcgttcccaaaattagt

ttgttttaaaaaacgtattgaagctatcccacaaattgataagtacttgaaatccagcaagtatatagcatggcctttgcagggctg

gcaagccacgtttggtggtggcgaccatcctccaaaatcggatctggttccgcgtggatcctcggtgaattacgcggcggggc

tgtcgccgtacgcggacaagggcaagtgcggcctcccggagatcttcgaccccccggaggagctggagcggaaggtgtgg

gaactggcgaggctggtctggcagtcttccaatgtggtgttccacacgggtgccggcatcagcactgcctctggcatccccga

cttcaggggtccccacggagtctggaccatggaggagcgaggtctggcccccaagttcgacaccacctttgagagcgcgcg

gcccacgcagacccacatggcgctggtgcagctggagcgcgtgggcctcctccgcttcctggtcagccagaacgtggacgg

gctccatgtgcgctcaggcttccccagggacaaactggcagagctccacgggaacatgtttgtggaagaatgtgccaagtgta

agacgcagtacgtccgagacacagtcgtgggcaccatgggcctgaaggccacgggccggctctgcaccgtggctaaggca

agggggctgcgagcctgcaggggagagctgagggacaccatcctagactgggaggactccctgcccgaccgggacctgg

cactcgccgatgaggccagcaggaacgccgacctgtccatcacgctgggtacatcgctgcagatccggcccagcgggaacc

tgccgctggctactaagcgccggggaggccgcctggtcatcgtcaacctgcagcccaccaagcacgaccgccatgctgacct

ccgcatccatggctacgttgacgaggtcatgacccggctcatgaagcacctggggctggagatccccgcctgggacggcccc

cgtgtgctggagagggcgctgccacccctgccccgcccgcccacccccaagctggagcccaaggaggaatctcccacccg

gatcaacggctctatccccgccggccccaagcaggagccctgcgcccagcacaacggctcagagcccgccagccccaaac

gggagcggcccaccagccctgccccccacagacccccccaaaagggtgaaggccaaggcggtccccagctga

601 nucleosome positioning DNA:

Ctggagaatcccggtgccgaggccgctcaattggtcgtagacagctctagcaccgcttaaacgcacgtacgcgctgtccccc

gcgttttaaccgccaaggggattactccctagtctccaggcacgtgtcagatatatacatcctgt

Ubiquitin-His:

atgcagatcttcgtgaagactctgactggtaagaccatcactctcgaagtggagccgagtgacaccattgagaatgtcaaggca

aagatccaagacaaggaaggcatccctcctgaccagcagaggttgatctttgctgggaaacagctggaagatggacgcaccc

tgtctgactacaacatccagaaagagtccaccctgcacctggtactccgtctcagaggtggtcatcaccatcaccatcactga

mmOcKRS:

atggataaaaaaccactaaacactctgatatctgcaaccgggctctggatgtccaggaccggaacaattcataaaataaaacac

cacgaagtctctcgaagcaaaatctatattgaaatggcatgcggagaccaccttgttgtaaacaactccaggagcagcaggact

gcaagagcgctcaggcaccacaaatacaggaagacctgcaaacgctgcagggtttcggatgaggatctcaataagttcctcac

aaaggcaaacgaagaccagacaagcgtaaaagtcaaggtcgtttctgcccctaccagaacgaaaaaggcaatgccaaaatcc

gttgcgagagccccgaaacctcttgagaatacagaagcggcacaggctcaaccttctggatctaaattttcacctgcgataccg

gtttccacccaagagtcagtttctgtcccggcatctgtttcaacatcaatatcaagcatttctacaggagcaactgcatccgcactg

gtaaaagggaatacgaaccccattacatccatgtctgcccctgttcaggcaagtgcccccgcacttacgaagagccagactga

caggcttgaagtcctgttaaacccaaaagatgagatttccctgaattccggcaagcctttcagggagcttgagtccgaattgctct

ctcgcagaaaaaaagacctgcagcagatctacgcggaagaaagggagaattatctggggaaactcgagcgtgaaattaccag

gttctttgtggacaggggttttctggaaataaaatccccgatcctgatccctcttgagtatatcgaaaggatgggcattgataatgat

accgaactttcaaaacagatcttcagggttgacaagaacttctgcctgagacccatgcttgctccaaacctttacaactacgctcg

caagcttgacagggccctgcctgatccaataaaaatttttgaaataggcccatgctacagaaaagagtccgacggcaaagaac

acctcgaagagtttaccatgctgaacttctctcagatgggatcgggatgcacacgggaaaatcttgaaagcataattacggactt

cctgaaccacctgggaattgatttcaagatcgtaggcgattcctgcatggtctatggggataccttgatgtaatgcacggagac

ctggaactttcctctgcagtagtcggacccataccgcttgaccgggaatggggtattgataaaccctggataggggcaggtttc

gggctcgaacgccttctaaaggttaaacacgactttaaaaatatcaagagagctgcaaggtccgagtcttactataacgggattt

ctaccaacctgtaa

mmAcKRS:

atggataagaaaccgctgaatactctgatttctgcaactggtctgtggatgagccgtaccggcaccatccacaagatcaaacac

cacgaggtttcccgtagcaaaatctacatcgaaatggcgtgcggtgaccacctggtggtaaacaactcccgttcttctcgtactg

cacgtgctctgcgccaccacaagtaccgtaagacctgcaagcgctgtcgcgtgtctgatgaagacctgaacaaattcctgacta

aagcgaacgaagatcagacttctgtgaaggtgaaagttgtttctgccccaacccgcaccaagaaagcgatgccgaagtccgtt

gcacgcgctccgaaaccgctggagaacaccgaagccgcacaggcccagccgtctggttctaagttttctccggcaatcccgg

tttctactcaggagtctgtgtctgtgccagcttctgttagcacttctatttcctctatcagcactggtgcgactgcgtccgctctggta

aaaggtaacactaacccgatcaccagcatgtctgctccggttcaggcttctgcaccggcactgactaaaagccagactgaccgt

ctggaggttctgctgaacccgaaagatgaaatcagcctgaactctggcaaaccgttccgtgaactggaatccgaactgctgtct

cgtcgtaagaaagacctgcaacaaatctatgctgaagagcgtgaaaactacctgggtaaactggaacgtgaaatcacccgtttc

tttgtggaccgtggtttcctggaaatcaagtctccgatcctgatcccgctggaatacatcgagcgcatgggtattgataacgacac

cgaactgtccaagcagattttccgtgtggacaagaacttctgcctgcgtccgatgctggcaccgaacctgtacaattacctgcgt

aaactggatcgtgcactgccggacccgatcaaaatctttgaaatcggtccatgctatcgtaaggagagcgacggtaaagaaca

cctggaagagttcactatgctgaacttttgtcagatgggttctggctgcacccgtgaaaatctggaatctatcatcaccgacttcct

gaaccacctgggcattgacttcaaaatcgttggtgattcctgcatggtttacggtgacactctggacgttatgcatggtgatctgga

actgagcagcgctgttgtgggtccgattccgctggatcgtgaatggggtatcgataaaccgtggattggtgctggcttcggtctg

gaacgtctgctgaaagttaagcacgactttaagaacatcaaacgtgctgcgcgttccgagtcctattacaacggcattagcacta
acctgtaa

Results and Discussion

Inspired by the observation made by Lin, Denu, and their coworkers that Sirt6 has enhanced deacylation activities toward peptide substrates containing a long chain fatty acyl-lysine and the observation by Cohen *et al.* that the nucleosome structure may activate Sirt6, we sought to synthesize acyl-H3 proteins with a long chain fatty acyl-lysine and use them to assemble corresponding acyl-nucleosomes as active substrates for Sirt6 targeted site screening. To analyze Sirt6-catalyzed deacylation on these acyl-nucleosome substrates, we could transfer denatured histones to membranes during SDS-PAGE western blot analysis with corresponding fatty acyl-lysine detecting antibodies. But it is critical for us to directly analyze deacylation levels on acyl-nucleosome substrates instead of on acyl-H3 proteins because the dissolution of acyl-nucleosomes during assays to acyl-H3 proteins or their tetramers with H4 will jeopardize our analysis and provide misleading results about what sites are targeted by Sirt6 at the nucleosome level. However, no commercial antibodies for lysine fatty acylation detection are available. An alternative analytical approach is to use mass spectrometry (MS). However, doing MS analysis on intact nucleosomes is very challenging. To overcome this analytical obstacle, we designed a chemical biology approach as illustrated in **Figure 26**. Previously we showed that a tetrazine undergoes inverse electron-demand Diels-Alder reaction readily with an aliphatic terminal olefin under physiological conditions and at ambient temperature. This reaction is bioorthogonal and has been

applied to label proteins with a terminal olefin in live cells[218]. Given the high structural similarity between an aliphatic terminal olefin and an ethyl group, changing the ethyl group at the end of the fatty acyl chain to an olefin is not expected to significantly alter Sirt6 deacylation activity toward this fatty acyl-lysine in the nucleosome context. However, providing this change will allow directly labeling of a corresponding acyl-nucleosome with a fluorogenic tetrazine dye in a native PAGE gel. The acyl-nucleosome in a native PAGE gel can also be stained with ethidium bromide (EtBr) to confirm its nucleosomal state. If Sirt6 is active toward this acyl-nucleosome for its deacylation, the resulting non-modified nucleosome cannot be labeled with a tetrazine dye but is still

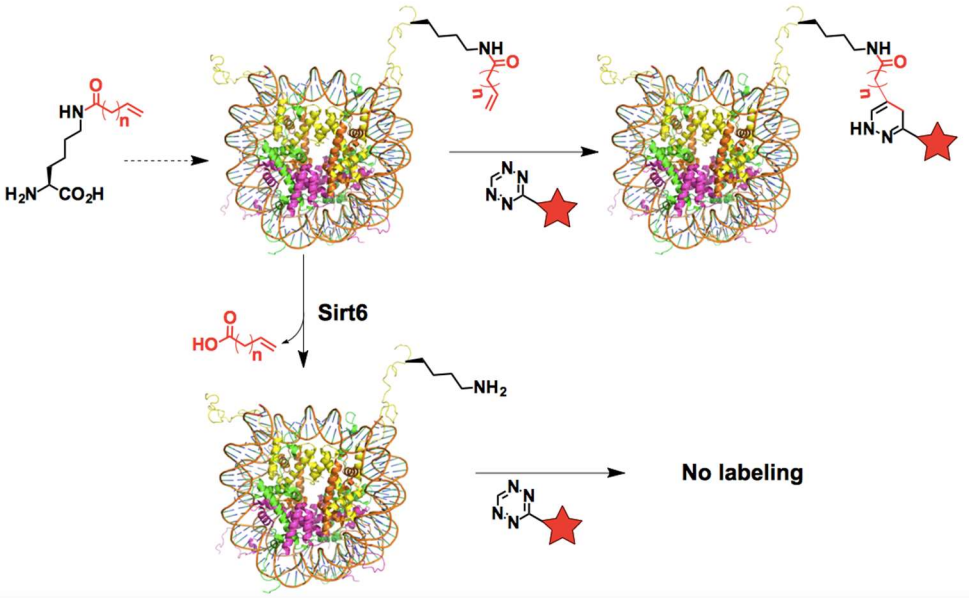


Figure 26 Probing Sirt6-targeted lysine deacylation sites in nucleosome using a chemical biology approach. A terminal olefin-containing fatty acyl-lysine is site-specifically installed in a histone that is assembled into a nucleosome as an active Sirt6 substrate.[276]

visible after staining with EtBr. Therefore by comparing tetrazine-based fluorescent intensities of the original acyl-nucleosome and its Sirt6-treated counterpart, Sirt6 deacylation activity toward this acyl-nucleosome can be obtained. In the following section, we will describe how we applied this approach to fish out Sirt6-targeted H3 lysine sites at the nucleosome level.

The genetic incorporation of N^{ε}-(7-octenoyl)-lysine (OcK). In order to genetically incorporate a terminal olefin-containing fatty acyl-lysine into H3, we exploited the amber suppression-based noncanonical amino acid mutagenesis approach similarly as in the synthesis of acetyl-H3 proteins. Denu *et al.* showed that Sirt6 has an enhanced deacylation activity toward lysine with a fatty acyl chain length of 8-14 carbons[199]. Based on this observation, we synthesized two terminal olefin-containing fatty acyl-lysines OcK (**Figure 27A**) and N^{ε}-(9-decenoyl)-lysine (DeK). OcK can be dissolved in water up to 2 mM; nonetheless however, DeK is barely soluble. We did not synthesize any acyl-lysine with a chain length long than 10 carbons given its insoluble tendency in water. Since we typically use 1 mM noncanonical amino acid for expressing noncanonical amino acid-containing proteins, OcK is a viable choice. For its incorporation, mutant *M. mazei* pyrrolysyl-tRNA synthetases that selectively charge $tRNA^{Pyl}$ with OcK were first identified from a pyrrolysyl-tRNA synthetase gene library in which codons for four active site residues, Y306, L309, C348, and Y384, of pyrrolysyl-tRNA synthetase were randomized. A widely adopted and double-sieved selection protocol was applied for the selection. The mutant, together with OcK and $tRNA^{Pyl}$, that displays the best amber suppression efficiency in *E. coli* has mutations as

L309A/C348S and is coined as OcKRS. A plasmid pEVOL-OcKRS that contains genes coding OcKRS and $tRNA^{Pyl}$ was then constructed. Together with another plasmid petDuet-UbK48Am that contains a gene coding ubiquitin with an amber mutation at K48 and a C-terminal 6×His tag, was used to transform *E. coli* BL21(DE3) cells. Growing the transformed cells in LB supplemented with 1 mM OcK afforded the expression of full-length ubiquitin Ub-K48oc to a level of 15 mg/L. In the absence of OcK, there was only a minimal amount of full-length ubiquitin expressed, confirming that OcKRS is specific for OcK (**Figure 27C**). The purified Ub-K48oc was analyzed by electrospray ionization-MS that gave a molecular weight of 9511.5 Da, agreeing well with the theoretical molecular weight at 9511.7 Da (**Figure 27D,**). To examine whether OcK in Ub-K48oc can be selectively labeled with a fluorogenic tetrazine, a previously synthesized tetrazine dye FITC-TZ (**Figure 27B**) was employed. After incubation with 100 μM FITC-TZ for 2 h, Ub-K48oc exhibited strong fluorescence in a SDS-PAGE gel, which was in striking contrast to no detected fluorescence for wild type ubiquitin, confirming selective labeling of OcK in a protein by a tetrazine dye (**Figure 27E**).

Sirt6 activities on OcK-containing H3-H4 (oc-H3-H4) tetramers. We next proceeded to synthesize mono-acylated H3 (oc-H3) proteins with OcK incorporated at K4, K9, K14, K18, K23, K27, K36, K56, and K79, respectively. *E. coli* BL21(DE3) cells transformed with pEVOL-OcKRS and a petDuet-H3 vector coding H3 with an amber mutation at a designated site were grown in 2YT supplemented with 1 mM OcK to produce an Oc-H3 protein. For OcK incorporation at K4-56 of H3 (H3K4-56oc), an N-terminal 6×His tag followed by a TEV protease digestion site was introduced for

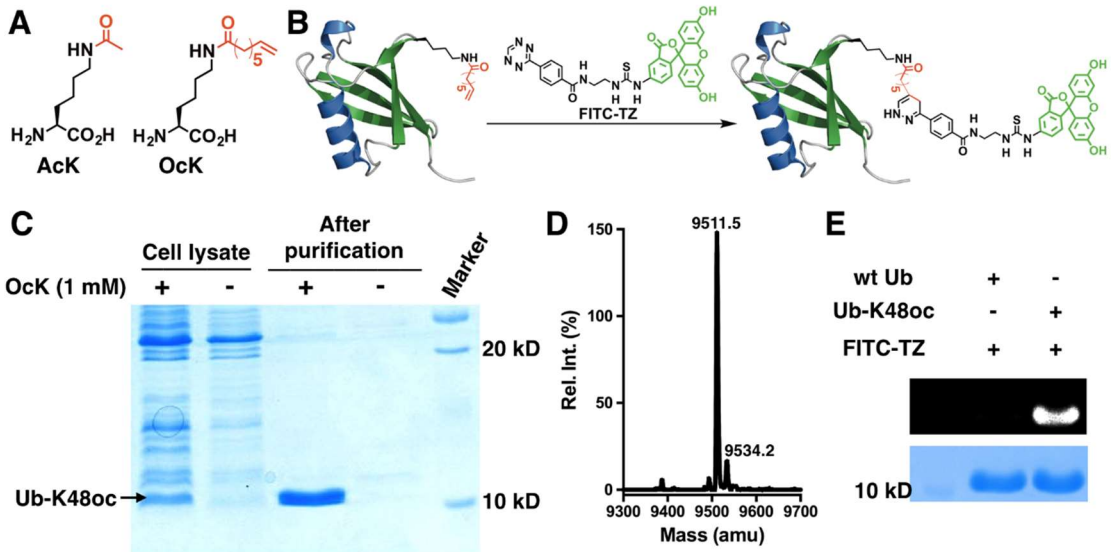


Figure 27 The genetic incorporation of OcK. (**A**) Structures of AcK and OcK. (**B**) A diagram to illustrate fluorescent labeling of Ub-K48oc by a fluorogenic tetrazine dye, FITC-TZ. (**C**) The selective incorporation of OcK into ubiquitin at its K48 position. BL21(DE3) cells transformed with plasmid pEVOL-OCKRS coding OcKRS and $tRNA^{Pyl}$ and plasmid petDuet-UbK48Am coding ubiquitin with an amber mutation at K48 were grown in LB with or without 1 mM OcK. (**D**) The Deconvoluted ESI-MS spectrum of Ub-K48oc (calculated molecular weight: 9511.7 Da). (**E**) The selective labeling of Ub-K48oc by FITC-TZ. Proteins were incubated with 100 μM FITC-TZ for 2 h before they were analyzed by SDS-PAGE. The top panel shows the FITC-based fluorescent image of the gel and the bottom panel shows the same gel after it was stained with Coomassie Blue.[276]

affinity purification with Ni-NTA resins and subsequent digestion by TEV protease to afford intact oc-H3 proteins. For OcK incorporation at K79 of H3 (H3K79oc), a truncation product with translation terminated at K79 predominated and therefore a C-terminal 6×His tag was introduced for easy separation of the full-length H3K79oc from its truncation product. Although we cannot remove the C-terminal 6×His tag from H3K79oc without leaving a scar, this C-terminal 6×His tag is not expected to affect assembling H3K79oc into either a tetramer with H4 or a nucleosome given that the H3

C-terminus is not directly involved in interactions with DNA and other histones. All nine oc-H3 proteins were expressed and purified to homogeneity (**Figure 29A**). They were then assembled together with H4 to form oc-H3-H4 tetramers. (**Figure 28**) Activities of Sirt6 on these tetramers were then analyzed. We tested Sirt6 activities on oc-H3-H4 tetramers instead of oc-H3 proteins themselves due to a concern that the 0.5 M arginine that is required to solubilize H3 may affect Sirt6 function. After incubating 0.4 μM Sirt6 with an oc-H3-H4 tetramer (1.6 μM) for 3 h, the treated tetramer was labeled with FITC-TZ, analyzed by SDS-PAGE, and then fluorescently imaged. The original tetramer without Sirt6 treatment was labeled and analyzed by SDS-PAGE in parallel as a control. As shown in **Figure 29C**, there is no significant fluorescent intensity difference on H3 between Sirt6-treated and untreated samples for all tetramers except the tetramer assembled from H3K9oc. H3K9oc does show weaker fluorescence than its untreated counterpart, indicating that Sirt6 targets this site for deacylation although the activity is still weak.

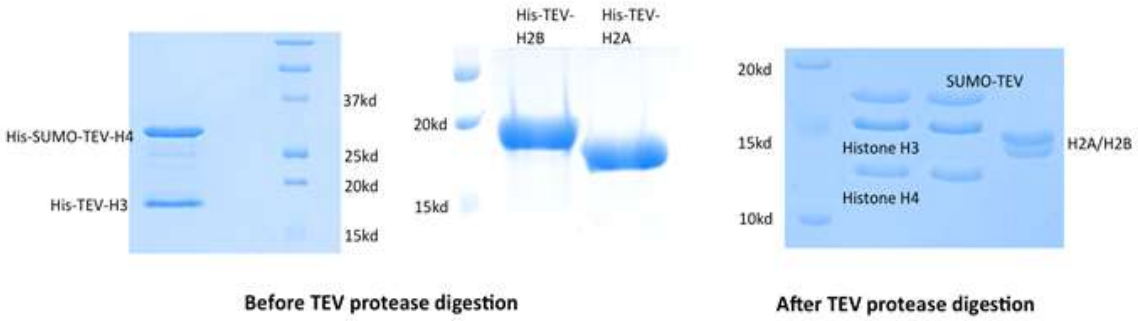


Figure 28 *In vitro* reconstitution of Histone H3/H4 tetramers and H2A/H2B dimers.[276]

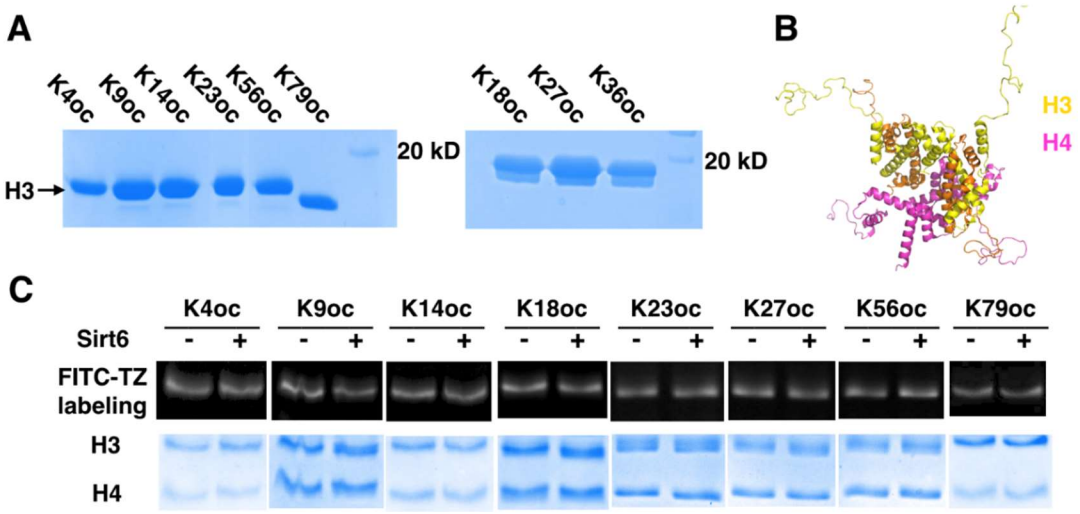


Figure 29 Sirt6 activities on oc-H3-H4 tetramers. (A) SDS-PAGE analysis of 9 purified oc-H3 proteins. (B) The H3-H4 tetramer structure. (C) Sirt6 activities on 9 oc-H3-H4 tetramers. A tetramer (1.6 μM) was incubated with or without 0.4 μM Sirt6 for 3 h before it was labeled with 200 μM FITC-TZ for 3 h and then analyzed by SDS-PAGE.[276]

Sirt6 activities on OcK-containing nucleosome (oc-nucleosome) substrates.

We next moved on to assemble oc-nucleosomes from oc-H3 proteins and used them as substrates for Sirt6 in hope that the nucleosome structure will significantly activate Sirt6 so that a clear picture of Sirt6-targeted H3 sites can be revealed. Following a standardized protocol developed by Luger *et al.*[219], we integrated all 9 oc-H3 proteins into their corresponding oc-nucleosomes. (For full nucleosome gels see **Figure 30**) A wild type nucleosome assembled from an unmodified H3 protein was also made as a control. Deacylation activities of Sirt6 on these oc-nucleosome substrates were then analyzed. The analysis setup was to incubate an oc-nucleosome (0.66 μM) with 0.33 μM Sirt6 for 3 h and then label it with FITC-TZ and to analyze the labeled product in a native agarose gel in parallel with the FITC-TZ-labeled original oc-nucleosome (**Figure**

31). After imaging the gel via FITC fluorescence, the same gel was then stained with EtBr and imaged again via EtBr fluorescence. For all 9 assembled oc-nucleosomes, incubating with Sirt6 didn't cause their dissemblance. All oc- nucleosomes with and without treatment with Sirt6 and the control nucleosome showed strong EtBr-stained DNA bands around 500 bp, agreeing well with what the literature has reported about the migration of a mononucleosome in a native PAGE gel. All nucleosomes have two EtBr-stained bands, indicating one major and one minor nucleosome conformer. This is also an observation reported previously[219]. After incubating with Sirt6, H3K4oc-, H3K9oc-, H3K18oc-, H3K23oc-, and H3K27oc-nucleosomes displayed much lower FITC-labeling fluorescence than their original counterparts. H3K9oc- and H3K18oc-nucleosomes showed almost no FITC labeling after their reactions with Sirt6, indicating close to total removal of the fatty acyl group. However, H3K36oc-, H3K56oc-, and H3K79oc-nucleosomes that reacted with Sirt6 showed no obvious FITC labeling difference from the original untreated nucleosomes. It is apparent that Sirt6 can target H3K4, H3K9, H3K18, H3K23, and H3K27 for their deacylation and that the nucleosome core structure significantly improves Sirt6's catalytic activities at these sites. In our *in vitro* setup, H3K56 is obviously not a preferential deacylation site for Sirt6.

Sequence preferences of Sirt6-targeted lysine sites. To compare catalytic deacylation activities of Sirt6 on the five detected H3 sites, we also carried out a time-based analysis. H3K4oc-, H3K9oc-, H3K18oc-, H3K23oc-, and H3K27oc-nucleosomes were incubated with Sirt6, quenched by 10mM Nicotinamide at different times, labeled with FITC-TZ, and then analyzed in native PAGE gels. Their FITC-labeling intensities

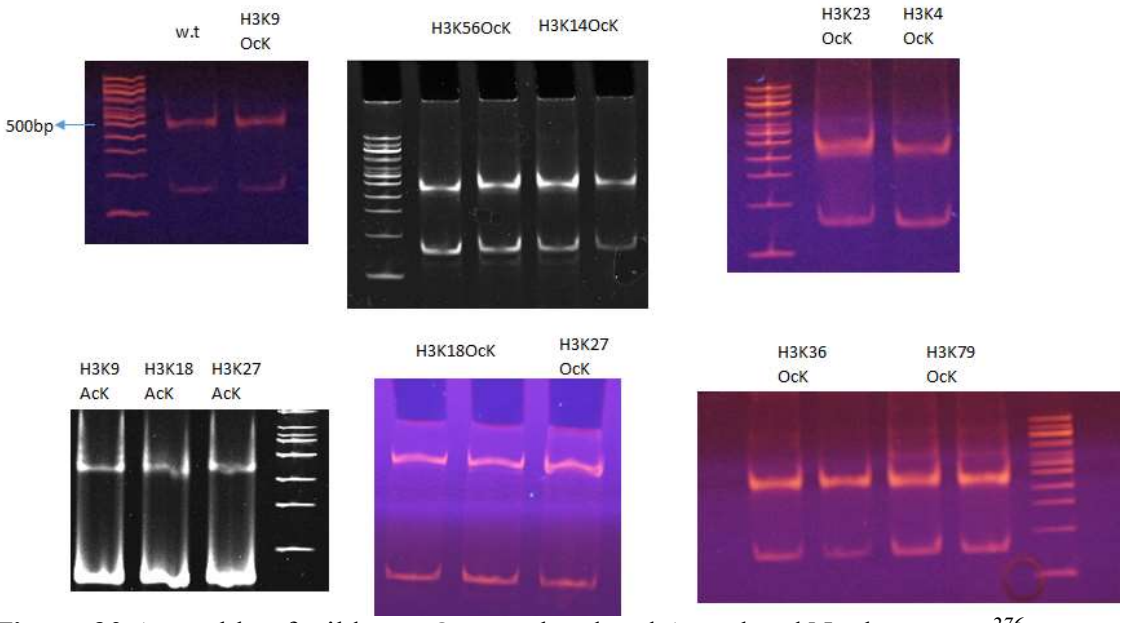


Figure 30 Assembly of wildtype, Octenoylated and Acetylated Nucleosomes.[276]

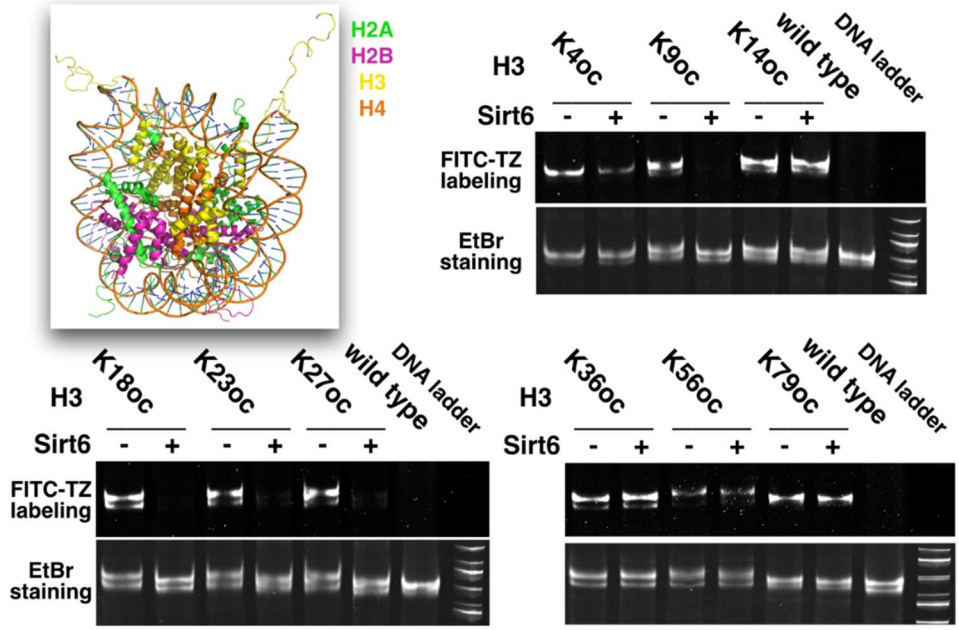


Figure 31 Sirt6 activities on oc-nucleosomes. The image in the top left corner shows the structure of a nucleosome. For all gel images, the top panel shows the FITC-based fluorescent imaging and the bottom panel shows the EtBr-stained DNA from the same gel.[276]

were then integrated and plotted for comparison. As shown in **Figure 32A**, Sirt6 has similar catalytic efficiency in removing fatty acylation from H3 K9 and H3 K18. After H3K9oc- and H3K18oc-nucleosomes (0.66 μM) reacted with 0.33 μM Sirt6 for 2 h, the fatty acyl group was quantitatively removed from the two nucleosomes. The deacylation activity of Sirt6 on H3 K27 is a little lower than on H3 K9 and H3 K18 but still comparable. There was residual fatty acylation left on the tested H3K27oc-nucleosome (0.66 μM) after its incubation with Sirt6 (0.33 μM) for 2 h. For the two sites at H3 K4 and H3 K23, Sirt6 displays similar deacylation activities although Sirt6 activities on these two sites are much weaker than on H3 K9, H3 K18, and H3 K27. After H3K4oc- and H3K23oc-nucleosomes (0.66 μM) reacted with 0.33 μM Sirt6 for 2 h, a significant amount of these two nucleosomes were still fatty acylated. This time-based assay allows us to divide Sirt6-targeted H3 lysine sites into two groups with the first group including H3 K9, H3 K18, and H3 K27 and the second group including H3 K4 and H3 K23. Sirt6 preferentially targets the first group for deacylation but is also catalytically active toward the second group. These two groups of Sirt6-targeted lysine sites also have converged consensual sequences. As shown in **Figure 32B**, H3 K9, H3 K18, and H3 K27 all have a preceding arginine residue right at its N-terminal side and there is an N-terminal serine or threonine residue followed by H3 K4 and H3 K23. These two concensual N-terminal adjacent residues apparently determine catalytic activities of Sirt6 on the five targeted lysine sites.

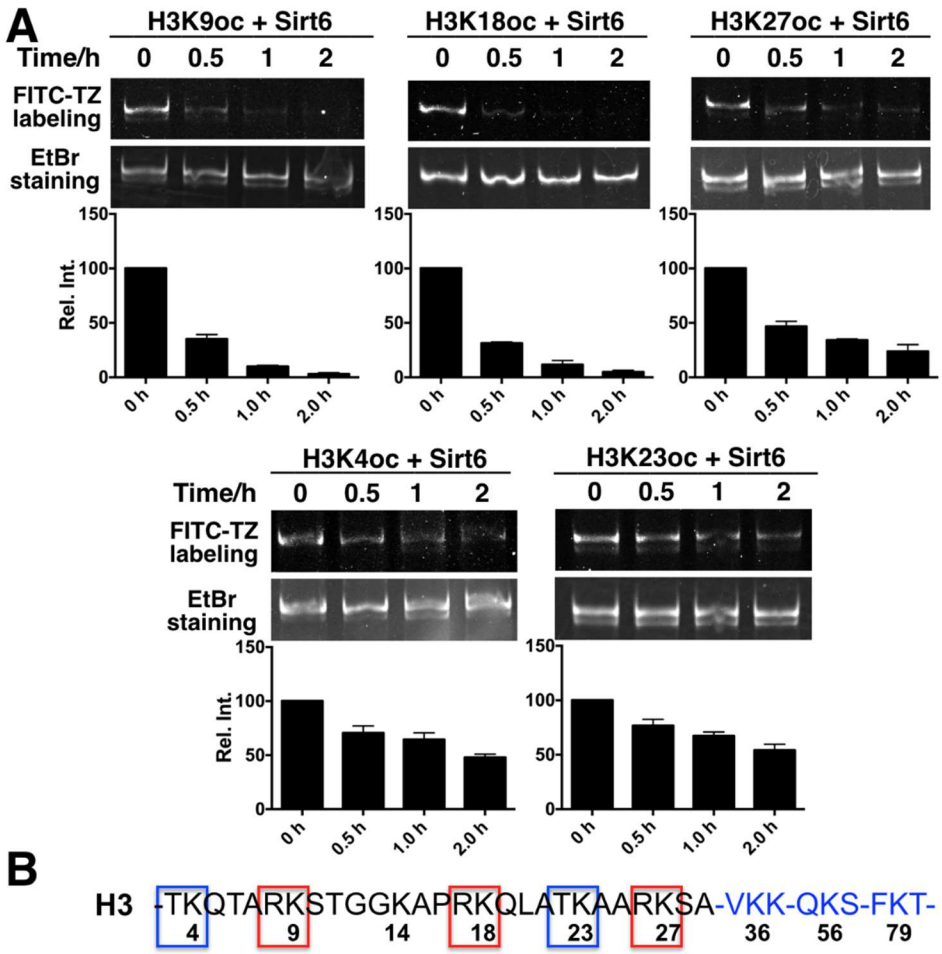


Figure 32 SIRT6 site specificity is determined by amino acid sequence. (A) Relative catalytic activities of Sirt6 on five oc-nucleosome substrates. An oc-nucleosome substrate (0.66 μM) was incubated with 0.33 μM Sirt6. The reaction was stopped at 0.5, 1, and 2 h and oc-nucleosome samples at these different times were labeled with FITC-TZ and analyzed in a native agarose gel together with the original oc-nucleosome (0 h). Top panels show FITC-based images, middle panels show same gels based on EtBr staining, and bottom panels show relative integrated FITC-based fluorescent intensities at different reaction times. (B) Sirt6-targeted lysine sites show consensual sequences highlighted in red and blue rectangles.[276]

Sirt6 catalyzes deacetylation at H3 K9, H3 K18 and H3 K27 in cells. With knowing that H3 K9, H3 K18, and H3 K27 are preferential Sirt6-targeted sites for deacylation, we next move on to demonstrate that Sirt6 indeed catalyzes deacetylation at

these three sites in an *in vitro* biochemistry setup and in cells. Three corresponding acetyl-nucleosomes were assembled from H3K9ac, H3K18ac, and H3K27ac, respectively, together with H2A, H2B, H4, and P601 DNA (**Figure 30**). These acetyl-nucleosomes (0.66 μM) were then reacted with 0.33 μM Sirt6 for 3 h before they were analyzed in a native PAGE gel for EtBr staining and then transferred to a membrane for probing with a pan anti-Kac antibody (**Figure 33B**). In the EtBr stained gel, parallel samples that were incubated with and without Sirt6 showed similar intensities, indicating that Sirt6 did not dissolve any acetyl-nucleosome substrate. In the western blot analysis, the pan anti-Kac antibody showed very different detect detection abilities toward three acetylation sites. Although all three original acetyl-nucleosomes were detected by the pan anti-Kac antibody, signals for H3K9ac- and H3K27ac-nucleosomes were much stronger than that for H3K18ac-nucleosome. Notwithstanding different detection abilities of the pan anti-Kac antibody toward three acetylation sites, all three acetyl-nucleosome samples that reacted with Sirt6 displayed much decreased acetylation levels in comparison to their original counterparts. For H3K9ac- and H3K18ac-nucleosomes, no signal was detected, indicating total removal of acetylation; For H3K27ac-nucleosome, there was residual acetylation after its reaction with Sirt6 but much weaker than acetylation of the original nucleosome. To demonstrate that Sirt6 targets H3K9, H3K18 and H3K27 for their deacetylation in cells, a Sirt6-coding pEGFP vector was constructed and transiently transfected into 293T cells for overexpressing Sirt6. Cell lysates from Sirt6-transfected and control cells were then separated and probed by specific antibodies raised for H3K9ac, H3K18ac, and H3K27ac. As shown in **Figure**

33C, Sirt6-transfected cells have much weaker acetylation at H3K9, H3K18, and H3K27 than control cells. This result confirms that all three sites are indeed targeted by Sirt6 for deacetylation in cells.

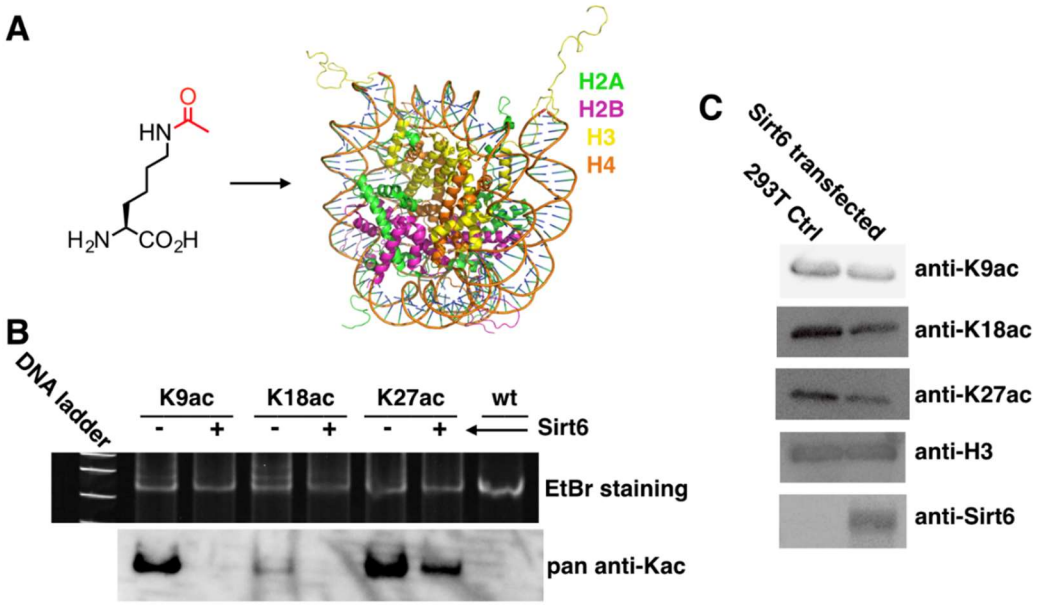


Figure 33 Sirt6 catalyzes deacetylation at H3 K9, H3 K18, and H 3K27. (**A**) The incorporation of AcK into a nucleosome. (**B**) Sirt6 activities on H3K9ac-, H3K18ac-, and H3K27ac-nucleosomes. Acetyl-nucleosomes (0.66 μM) were incubated with or without 0.33 μM Sirt6 for 3 h before they was analyzed by a native agarose gel. The top panel shows the EtBr stained agarose gel and the bottom panel shows the pan anti-Kac antibody detected acetylation levels at different samples. (**C**) Acetylation levels at H3 K9, H3 K18, and H3 K27 in Sirt6-transfected and control cells.[276]

Discussion. As a chromatin associated protein and a NAD^+-dependent sirtuin enzyme, Sirt6 has long been recognized as a histone deacetylase but with puzzlingly low activities toward peptide substrates. We show in this report that Sirt6 has close to undetectable deacetylation activities toward all 9 tested monomeric acetyl-H3 substrates. Although Sirt6 has been recently discovered to have enhanced activities toward a peptide lysine with long chain fatty acylation, our efforts to make oc-H3 proteins and

fold them to corresponding oc-H3-H4 tetramers did not yield active substrates for Sirt6 either. However, deacylation activities of Sirt6 toward several H3 lysine sites were immensely improved when corresponding acetyl- or oc-H3 proteins were assembled into their nucleosomal forms. Although the focus of our current study is to identify Sirt6-targeted H3 lysine sites for deacetylation, our result raise an interesting question as to what are the real substrates of Sirt6 in cells? Are they free histones or their nucleosomal forms? In fact, in promoter and active transcribed gene body regions where Sirt6 are found primarily positioned[220], chromatin is in a dynamic process of assembly and disassembly[221]. Sirt6 has the chance to interact with both free histones and nucleosomes. Our results imply that nucleosomes are real substrates of Sirt6 but do not rule out that Sirt6, as part of a large protein complex, catalyzes efficient deacetylation of free histones. As a matter of fact, a lot of histone epigenetic modifiers are found in large multi-subunit complexes that contain various histone or DNA recognition domains[222-224], although there has been none of these types of reports about Sirt6. Another interesting question raised by the current study is how the nucleosome structure activates Sirt6. Sirt6 has a highly positively charged C-terminus. Chua *et al.* showed that the cleavage of this C-terminus from Sirt6 does not deactivate the enzyme but abolishes its binding to chromatin[225]. One explanation of this observation is that the C-terminus is involved in nonspecific electronic interactions with negatively charged DNA for recruiting Sirt6 to chromatin or nucleosome. However, DNA may also neutralize the positively charged C-terminus and therefore allows Sirt6 to interact with positively charged histone tails for deacetylation. This aspect needs to be explored further. Another important feature of

Sirt6 the current study reveals is the preferential sequence contexts around its targeted lysine sites. Unlike Sirt1, Sirt2, and Sirt3[197, 226], Sirt6 prefers to deacetylate lysine that is preceded with a positively charged arginine residue. This may be attributed to the negatively charged residue E189 at a loop region of Sirt6 that directly interacts with protein or peptide substrates. All other sirtuins have a glycine residue at this site that may define their nonspecific recognition of lysine for deacetylation. Among all four histones, lysine residues that undergo acetylation and also have a preceding arginine residue include H2A K36 and H4 K20. We predict they are also Sirt6 targeted deacetylation sites. Other lysine sites that can be potentially targeted by Sirt6 for deacetylation are H2B K12 and H2B K16. These two lysine sites naturally undergo acetylation and are preceded with a positively charged lysine residue that potentially interacts with E189 of Sirt6 for strong binding.

Besides revealing important biochemical features of Sirt6, the current study also has strong biological implications. Reports have indicated that Sirt6 is frequently recruited to relevant gene promoters and represses gene transcription via removing acetylation at H3 K9[227, 228] and H3 K56[229, 230]. Our results show that H3 K18 and H3 K27 are targeted by Sirt6 as well for deacetylation. H3K18ac and H3K27ac are associated with active gene expression[231]. Targeting these two sites for deacetylation is expected to reinforce the transcriptional repression role of Sirt6. Functional redundancy of histone acetylation on chromatin regulation has been observed. For example, a single mutation of an N-terminal H3 lysine to glycine has little repression effect on gene expression in yeast. However, simultaneous mutations of all N-terminal lysines to

glycine significantly reduces yeast gene expression[232]. Thus, it is reasonable for Sirt6 to remove acetylation from multiple lysines to fulfill a strong gene transcription suppression role. By removing acetylation from H3 K9, H3 K18, and H3 K27, Sirt6 may also regulate functions of proteins that recognize acetylation at these three lysine sites. One bromodomain YEATS recognizes acetylation at H3 K9, H3 K18, and H3 K27 with similar binding affinities[233]. It is highly possible that Sirt6 regulates functions of YEATS-containing proteins such as AF9 and ENL by modulating acetylation at H3 K9, H3 K18, and H3 K27. Di- and trimethylation at H3 K27 (H3K27me2/3) has long been known for shutting down transcription[234]. H3K27me3 is an important mark for inactivated X chromosome. H3K27ac is associated with active transcription[235] and is apparently antagonistic to the repression of gene transcription by H3K27me2/3. Sirt6 may serve as an important mediator to switch H3K27 between these two antagonistic states.

In cancer biology, H3K18ac is an important prognostic mark. Downregulation of H3K18ac has been found correlated with poor prognosis of pancreatic adenocarcinoma[236] and low grade prostate cancer[214]. Patients with high H3K18ac show low resurgence of prostate cancer after treatment[237]. Incidentally overexpression of Sirt6 is observed in prostate cancer cell line PC-3 and inhibition of Sirt6 leads to apoptosis of PC-3[214]. Our study provides the missing link between downregulation of H3K18ac and high expression of Sirt6 in PC-3. It is highly possible that Sirt6 changes prostate cancer development through its deacetylation of H3 K18 and Sirt6 itself can be a potential anti-cancer drug target.

Previous studies showed that Sirt6 catalyzes deacetylation at H3 K56 on purified H3 and Sirt6 knockout mouse tissue had enhanced H3K56ac. Our results detected sluggish removal of acetylation and long chain fatty acylation from H3 K56 by Sirt6. According to current published nucleosomal structures[238], H3 K56 directly interacts with the wrapping DNA. In this nucleosomal format we have also adopted in our study, H3 K56 is blocked from access by Sirt6. However, in cells Sirt6 may involve additional cellular factors for direct deacetylation of H3 K56 on free H3 or positioning of DNA wrapping[239] could be different and expose H3 K56 for deacetylation. It is also possible that H3K56-targeting histone acetyltransferases/deacetylases are regulated by Sirt6 in cells.

Conclusion

As a member of a highly conserved family of NAD^+-dependent histone deacetylases, Sirt6 is a key regulator of mammalian genome stability, metabolism, and life span. Previous studies indicated that Sirt6 is hardwired to remove histone acetylation at H3K9 and H3K56. However, how Sirt6 recognizes its nucleosome substrates has been elusive due to the difficulty of accessing homogenous acetyl-nucleosomes and the low activity of Sirt6 toward peptide substrates. Based on the fact that Sirt6 has an enhanced activity to remove long chain fatty acylation from lysine, we developed an approach to recombinantly synthesize histone H3 with a fatty acylated lysine, N^{ε}-(7-octenoyl)-lysine (OcK), installed at a number of lysine sites and used these acyl-H3 proteins to assemble acyl-nucleosomes as active Sirt6 substrates. A chemical biology approach that visualizes OcK in nucleosomes and therefore allows direct sensing of Sirt6 activities on its acyl-nucleosome substrates was also formulated. By combining these two approaches, we

show that Sirt6 actively removes acylation from H3 K9, H3 K18, and H3 K27, has relatively low activities toward H3 K4 and H3 K23, and sluggishly removes acylation at H3 K14, H3 K36, H3 K56, and H3 K79. Unlike Sirt1 and Sirt2 that show no sequence preferentiality for their targeted lysine sites, Sirt6-targeted lysine sites display obvious sequence similarities with H3 K9, H3 K18, and H3 K27 preceded with an arginine residue and H3 K4 and H3 K23 preceded with a serine residue. Overexpressing Sirt6 in 293T cells led to downregulated acetylation at H3 K18 and H3 K27, confirming these two novel Sirt6-targeted nucleosome lysine sites in cells. The discrepancy between the sluggish *in vitro* activity of Sirt6 on H3 K56 and the high downregulation of H3 K56 acetylation in Sirt6-overexpressed cells may be due to additional cellular factors involved in direct deacetylation of H3 K56 by Sirt6 or regulation of H3 K56-targeting histone acetyltransferases/deacetylases by Sirt6 in cells. Given that downregulation of H3 K18 acetylation is correlated with poor prognosis of several cancer types and H3 K27 acetylation antagonizes repressive gene regulation by di- and trimethylation at H3 K27, our current study implies that Sirt6 may serve as a target for cancer intervention and regulatory pathway investigation in cells.

CHAPTER III

IDENTIFICATION OF SIRT7 DEACYLATION TARGETS USING ACYL-NUCLEOSOMES

Introduction

There exist seven sirtuins in humans, namely SIRT1-7.[240, 241] Four of these, SIRT1, 2, 6, and 7, can translocate to the nucleus, and therefore potentially remove acylation from chromatin for regulating chromatin epigenetic marks.[202, 242-244] Among these four sirtuins, SIRT7 is perhaps the least studied. Accumulating work suggests that SIRT7 has complex effects on cellular homeostasis, oncogenic potential, and cellular aging pathways.[245] Multiple studies have observed high SIRT7 levels in cancers, including head and neck squamous cell carcinoma, colorectal cancer, early state and metastatic breast cancer, thyroid carcinoma, ovarian cancer, and gastric cancer, consistent with oncogenic functions of SIRT7.[246-255] Moreover, SIRT7 represses transcription of several tumor suppressive genes through deacetylation of histone H3 lysine 18 (H3 K18ac), and depletion of SIRT7 is sufficient to reduce malignant properties of cancer cells and inhibit tumor growth in mice.[251, 256] SIRT7 is enriched in nucleoli, and associates with both euchromatic and heterochromatic ribosomal RNA genes (rDNA).[244, 257, 258]. At euchromatic rDNA genes, SIRT7 deacetylates the RNA polymerase I (Pol I) subunit PAF53, which leads to enhanced interactions with rDNA for active transcription.[244, 259] This SIRT7-dependent rRNA synthesis may help support the high ribosome biogenesis, proliferative capacity, and metabolic demands of cancer cells.[260] However, SIRT7 also has tumor suppressive functions, and SIRT7-deficient

mouse embryonic fibroblast cells show increased cell viability and cell-cycle entry into S and G2/M.[261-264] Moreover, SIRT7 is recruited to DNA double-strand breaks (DSBs) and protects against DNA damage by promoting recruitment of 53BP1 for initiating Non-homologous End Joining.[265, 266] Recently, SIRT7 was also shown to guard against genomic instability due to rDNA rearrangements in nucleoli, by maintaining rDNA heterochromatin silencing.[258] This function of SIRT7 is important for preventing senescence of human cells, which can contribute to tissue dysfunction in many aging-related pathologies and favor tumor growth in cancers. Thus, SIRT7 has pleiotropic effects on genomic stability and cellular homeostasis pathways that can impact on cancer and aging biology.

Although much evidence has established SIRT7 as a *bona fide* histone deacetylase, molecular details of SIRT7 interactions with chromatin for deacetylation have not been directly investigated. Two recent studies showed that both DNA and RNA can activate SIRT7.[267, 268] However, the mechanism of this activation process is yet to be illustrated. In the current study, we re-constituted *in vitro* acyl-nucleosomes that have acyl-lysines installed site-specifically at a number of native histone H3 lysine sites and used these recombinant acyl-nucleosomes as substrates to systematically investigate SIRT7 recognition of histone lysines for deacylation in the physiologic context of chromatin. In addition to confirming the SIRT7 deacylation activity on H3 K18, we have uncovered a novel histone substrate of SIRT7, H3 K36. We show that SIRT7 is a highly active, physiologic H3 K36 deacylase, and this activity is nucleosome or chromatin dependent and can be significantly enhanced when free DNA is appended to the acyl-

nucleosome substrate. This novel activity of SIRT7 may serve a pivotal role in guarding against genomic instability and human cellular senescence in cancer and aging-related pathology.

Experimental section

Synthesis of N^{ε}-(7-azidoheptanoyl)-L-lysine hydrochloride (AzHeK)

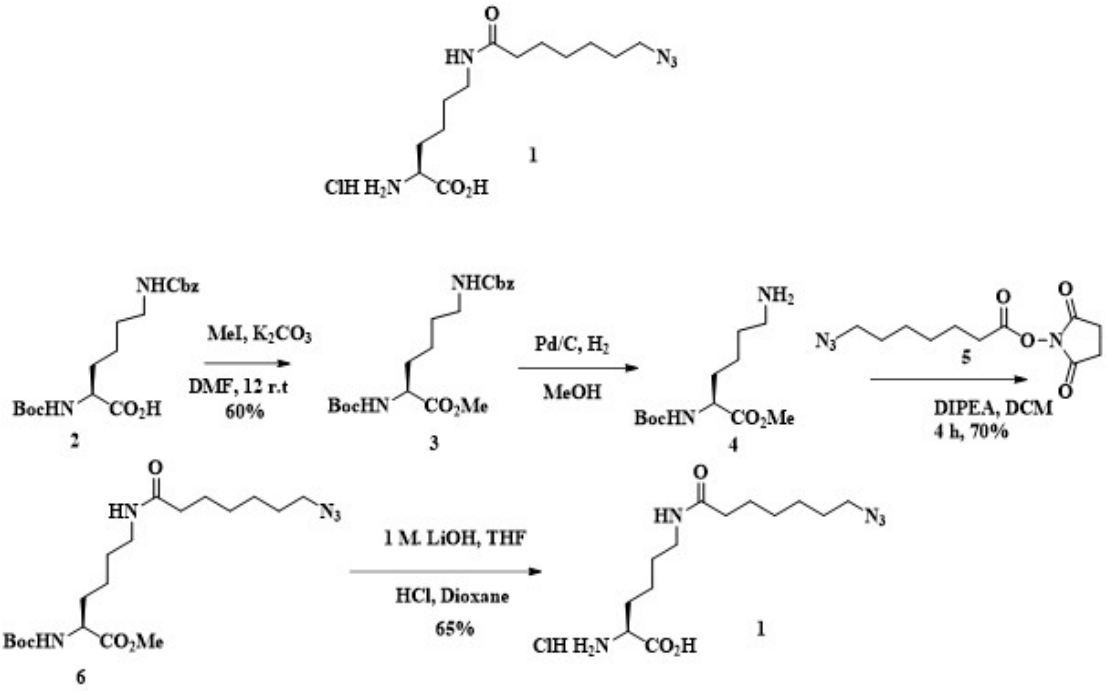


Figure 34 Synthetic route of AzHeK

Synthesis of compound 3: 9.51 g of Boc-Lys(Cbz)-OH (**2**) (25 mmol, Sigma-Aldrich) was dissolved in dry DMF, followed by the addition of K_2CO_3 (7.26 g, 52.5 mmol). 3.44 mL of MeI (55 mmol) was gradually added and was let to stir overnight at 80 ^{o}C. Then the reaction mixture was quenched by the addition of 50 mL H_2O and extracted by EtOAc 30 mL for 3 times. The EtOAc layers were combined, washed with brine, and

dried by anhydrous Na_2SO_4. The crude product was then purified by the silica gel flash column chromatography (eluted at 10% EtOAc/hexane) to give an oil product (**3**) (9.6 g, 97%).

7-azidoheptanoic acid: To 9.14 g of bromoheptanoic acid (44.0 mmol) was added 150 mL anhydrous DMF. To the stirring solution was added 9.14 g NaN_3 (140 mmol, 3.2 eq.) at room temperature. The clear solution was warmed to 60 °C for a period of 6 h. The solution was then cooled to room temperature and diluted with 250 mL DCM. The organic layer was washed with cold 1N HCl (7 x 100mL) and brine (2 x 250mL), then dried over anhydrous Na_2SO_4. Filtration and removal of solvent afforded 7.09 g crude 7-azidoheptanoic acid in 94% yield which was of suitable purity for continuing to next step.

7-azidoheptanoic acid-OSu (5): 7-azidoheptanonic acid (5.0 g, 29.1 mmol) was added to a solution of *N*-hydroxysuccinimide (3.69 g, 32.1 mmol) in dichloromehane (100 mL). The DCC (6.62 g 32.1 mmol) in dichloromethane was then added and the mixture was left stirred overnight at room temperature. Dicyclohexylurea was removed by filtration and the filtrate was concentrated under reduced pressure to yield 7-azidoheptanoic acid-OSu white crystals. This compound was directly used without further purification.

Synthesis of compound 6: The solution of Boc-Lys(Cbz)-OMe (**3**) (3.9 g, 10 mmol) in methanol (100 mL) was added palladium on activated carbon (Pd 10%, 0.6 g, 0.6 mmol). The mixture was stirred at room temperature with hydrogen bubbled through, where the reaction progress was monitored by TLC analysis. The reaction was terminated as the starting protected amino acid on TLC disappeared. Then the reaction mixture was

filtered through a celite cake packed on a Buchner funnel and the flowed-through solution was concentrated under reduced pressure to afford colorless oil that was characterized as the desired product (**4**). The product was directly used in the next step without further purification.

To a solution of the above amine (~9.15 mmol) in anhydrous dichloromethane (90 mL) cooled in an ice bath was added *N, N*-diisopropylethylamine (2.80 mL, 16.07 mmol) dropwise, which was followed by adding a solution of 7-azidoheptanoic acid-OSu **(5)** (2.69 g, 11.07 mmol) in dichloromethane (10 mL) dropwise over 20 min. The mixture was then stirred at room temperature for 12 h, and it was washed with sodium hydroxide solution (0.5 *N*, 20 mL) and brine (20 mL x 2), dried (Na2SO4), evaporated, and flash-chromatographed (EtOAc/hexanes, 1:3) to give **6** (3.2 g, 79% for two steps).

Synthesis of compound 1: The solution of **6** (3.0 g, 7.2 mmol) in a 1:1 mixture of methanol and THF (24 mL) was added into LiOH aqueous solution (1.0 M, 14 mL) and stirred at room temperature. The reaction progress was monitored by TLC analysis. After 5 h, the spot representing the starting ester on TLC plate disappeared. The reaction mixture was then diluted with EtOAc/H_2O. After partition, the organic layer was discarded and the aqueous layer was acidified by the addition of 3M HCl until pH was adjusted to 3-4. The mixture was than extracted with EtOAc twice. The organic layers were combined, washed with brine, dried over anhydrous Na_2SO_4, and concentrated to afford clear sticky oil as the desired product which was subjected directly to next step without chromatography purification.

The afforded acid from the previous step (2.7 g, 7.26 mmol) in dioxane (20 mL) was added into the HCl solution (4 M in dioxane, 3.6 mL). The mixture was stirred at room temperature. The reaction progress was monitored by TLC. After reaction for 3 h, the HCl solution was removed under reduced pressure. The afforded white solid was confirmed as compound **1** (1.5 g, 60% for two steps). ^{1}H NMR (D_2O, 300 MHz).4.10-3.98 (m, 1H), 3.31-3.21 (m, 4H), 2.25-2.28 (m, 2H), 2.12-1.82(m, 2H), 1.67-1.45 (m, 8H), 1.32-1.25 (m, 4H).

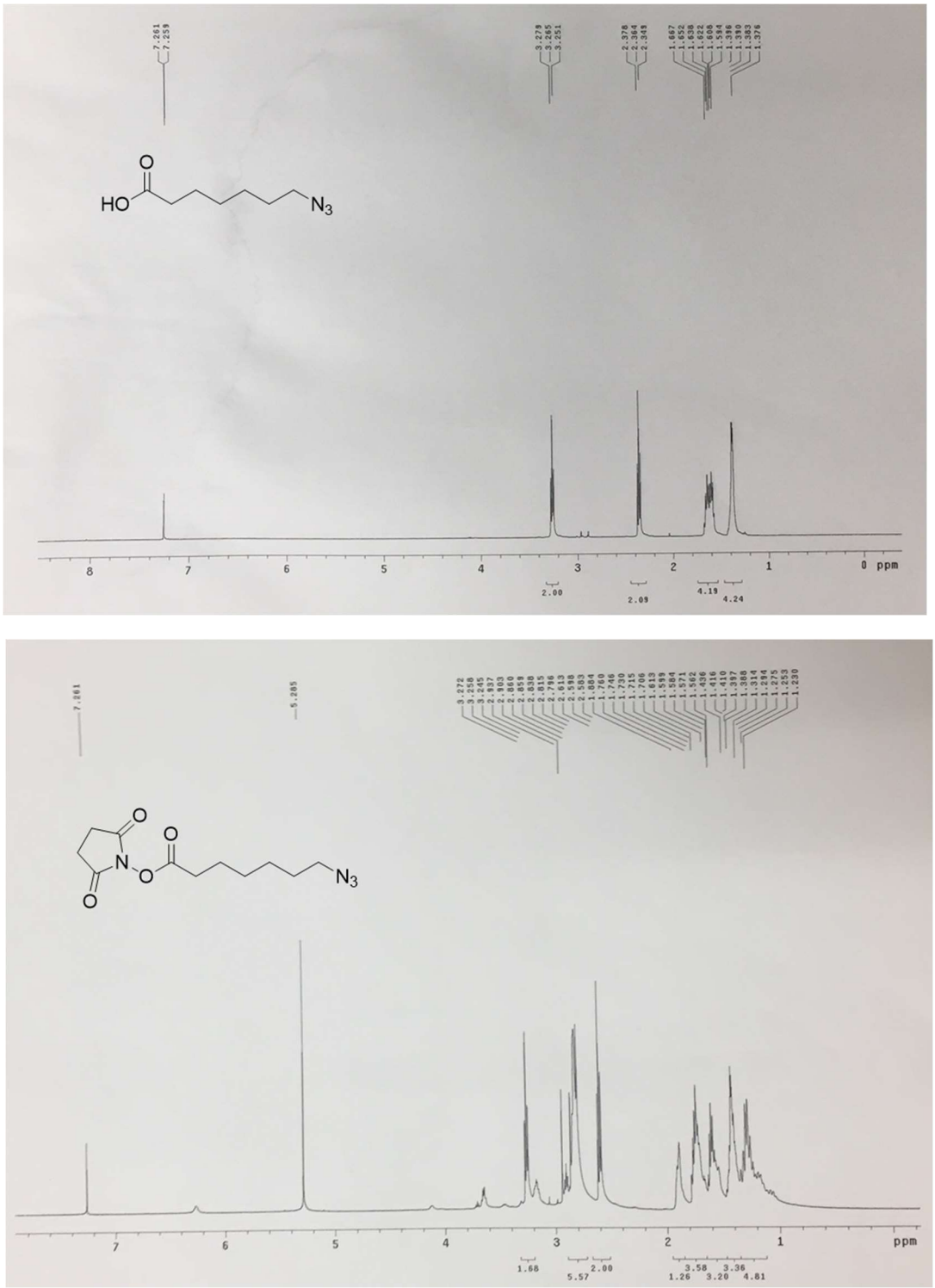


Figure 35 NMR spectra of N^{ε}-(7-azidoheptanoyl)-L-lysine hydrochloride (AzHeK) and intermediates

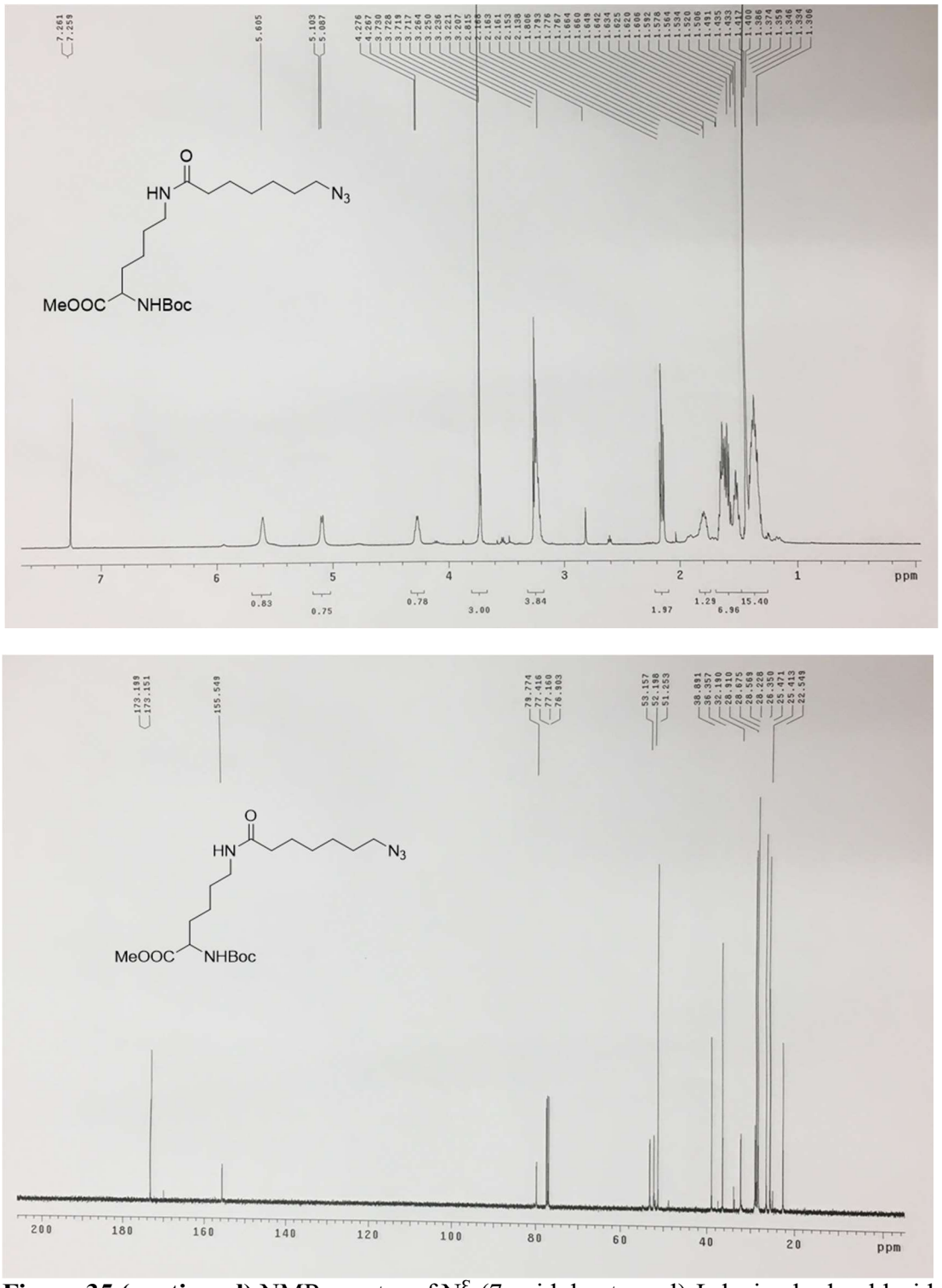


Figure 35 (continued) NMR spectra of N^{ε}-(7-azidoheptanoyl)-L-lysine hydrochloride (AzHeK) and intermediates.

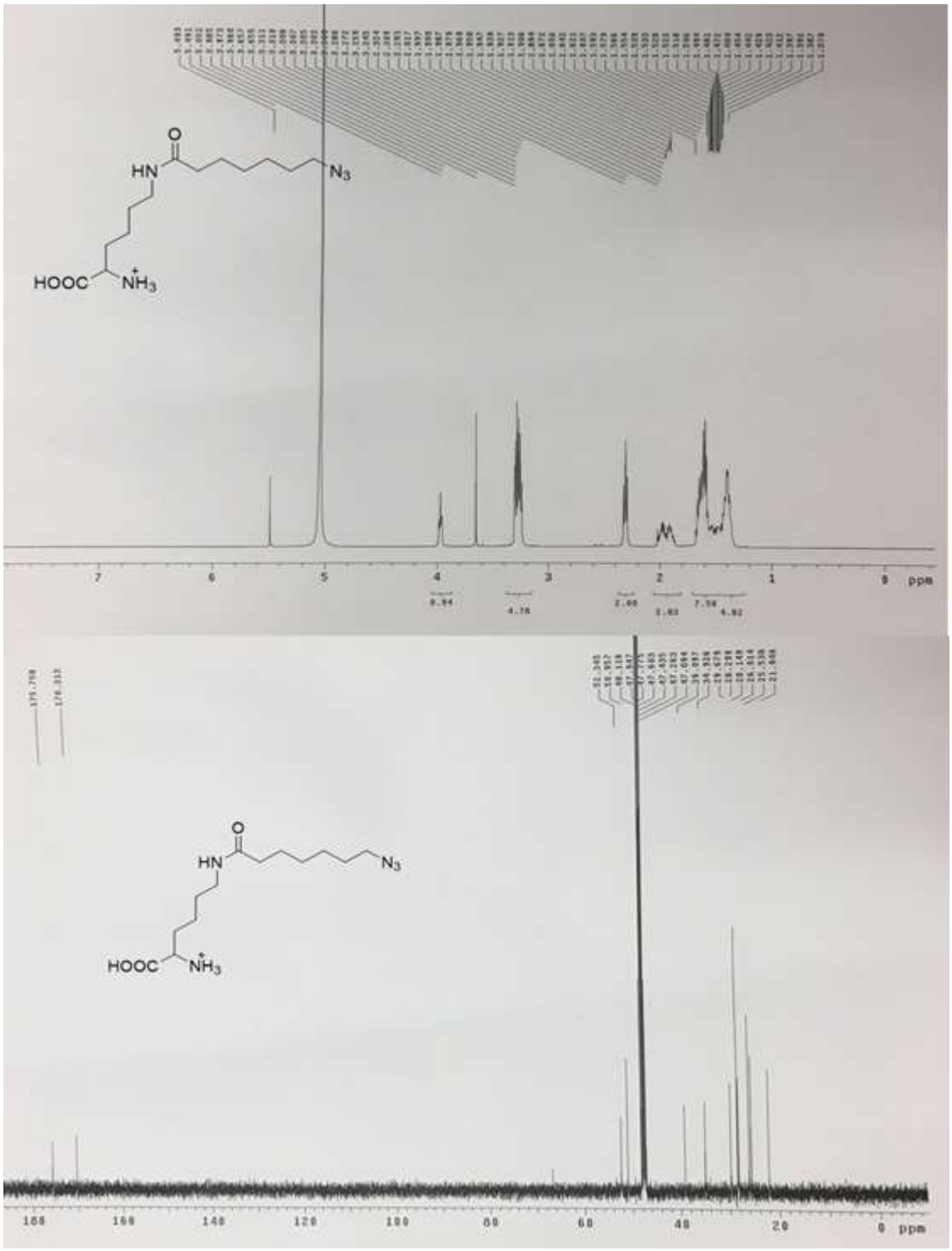


Figure 35 (continued) NMR spectra of N^{ε}-(7-azidoheptanoyl)-L-lysine hydrochloride (AzHeK) and intermediates.

Expression and purification of Ub-K48Az. Ubiquitin gene fragment with a C-terminal 6xHis tag was cloned into pETDuet-1 vector. An amber stop codon was introduced at K48 through site-directed mutagenesis. The resulting pETDuet-1-UbK48am was used to co-transform a CobB$^-$ E. coli BL21 (DE3) strain together with pEVOL-OcKRS. A single colony was picked and inoculated into 2YT medium supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. When OD_{600} reached to 0.6, 0.5 mM IPTG, 0.2% (w/v) L-arabinose and 1 mM AzHeK were added into the culture to induce protein expression. Cells were further incubated for 5 hours at 37 ℃ before cell collection and the following protein purification. Ubiquitin with AzHeK incorporated at K48 was purified in the same procedure as chapter II with a yield of 18 mg/L.[1]

Expression and Purification of histone H3 mutants incorporated with AzHeK. A pETDuet-1 vector with a human H3 gene with amber stop codon introduced at a specific lysine coding position was used to co-transform a CobB$^-$ E. coli BL21 (DE3) strain together with pEVOL-OcKRS. A Single colony was picked and inoculated into 2YT medium supplemented with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. When OD_{600} reached to 0.6, H3 expression was induced at 37 ℃ by adding 0.5 mM IPTG, 0.2% (w/v) L-arabinose and 1 mM AzHeK into the cell culture. Cells were harvested 6 h after induction, and histone H3 was purified according to procedures as chapter II. Yield of Histone H3 with AzHeK incorporated at different sites ranged from 12 mg/L-26 mg/L.

The In Vitro Assembly of Az-nucleosomes for SIRT7 deacylation assays. Recombinant Histone His-TEV-H2A, His-TEV-H2B and His-SUMO-TEV-H4 were expressed and

purified according to procedures reported in a previous publication.[1] All histone pellets were dissolved in 6 M GuHCl buffer (6 M guanidinium hydrochloride, 20 mM Tris, 500 mM NaCl, pH 7.5), and concentration was measured by UV absorption at 280 nm in a BioTek synergy H1 plate reader. To prepare H2A/H2B dimer, His-TEV-H2A and His-TEV-H2B were mixed in the molar ratio of 1:1, and 6 M GuHCl buffer was added to adjust total protein concentration to 4 mg/mL. The denatured His-TEV-H2A/His-TEV-H2B solution was dialyzed sequentially at 4 ℃ against 2 M TE buffer (2 M NaCl, 20 mM Tris, 1 mM EDTA, pH 7.8), 1 M TE buffer (1 M NaCl, 20 mM Tris, 1 mM EDTA, pH 7.8), and 0.5 M TE buffer (0.5 M NaCl, 20 mM Tris, 1 mM EDTA, pH 7.8). The resulting dimer solution was centrifuged at 14000 RPM for 5 min at 4 ℃ to remove precipitates. And its concentration was determined by UV absorption at 280 nm using the BioTek synergy H1 plate reader. Steps of His-TEV-H3/His-SUMO-TEV-H4 tetramer refolding were generally similar to the refolding of the His-TEV-H2A/His-TEV-H2B dimer except that the total protein concentration was adjusted to 2 mg/mL and there was no stirring in the sequential dialysis. To generate a histone octomer, the refolded His-TEV-H2A/Hi-TEV-H2B dimer was mixed with the refolded His-TEV-H3/His-SUMO-TEV-H4 tetramer in a molar ratio of 1:1, and NaCl solid was added to adjust NaCl concentration to 2 M. A 601 nucleosome positioning DNA (147 bp or a longer one with 5bp, 10bp, or 20bp extension at both 5’ and 3’ ends) was prepared by polymerase chain reaction and purified with PCR cleanup kit (#2360250 Epoch Life Science). Purified 601 DNA was redissolved in 2 M TE buffer and added to histone octamer solution in the molar ratio of 0.9:1. Lower molar ratio of DNA may be added if

free DNA is present in the final gel shift analysis. 2 M TE buffer was added to adjust the final DNA concentration to about 2-3 uM. The DNA histone mixture solution was then transferred to a dialysis bag and placed inside about 200 ml 2 M TE buffer. While it was stirred at room temperature, Tris buffer with no salt (20 mM Tris, pH 7.8) was slowly added into the 2M salt buffer through a liquid transfer pump (#23609-170 VWR®). When salt concentration was reduced to about 200 mM (measured by EX170 salinity meter), the dialysis bag would be transferred to nucleosome storage buffer (20 mM Tris, 20 mM NaCl, 1 mM EDTA pH 7.8) for another 4 h, and precipitation was removed through centrifugation at 14000 RPM and room temperature for 10 min. TEV protease was then added to the nucleosome solution (TEV: nucleosome 1:30, w:w) for about 1 h at 37 ℃. to remove all the histone tags. (**Figure 36**) Since the A260 absorption of a histone octamer is much lower than DNA, nucleosome concentration was determined by directly measuring A260 absorption in the BioTek synergy H1 plate reader.

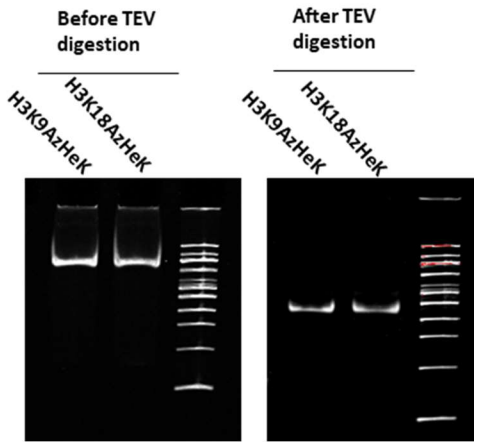


Figure 36 N-terminal tag didn't affect refolding of H2A/H2B dimer, H3/H4 tetramer and mononucleosome particles.

SIRT7 deacylation of Az-nucleosomes. Recombinant His-SIRT7 was expressed and purified from E. coli as reported previously by Lin et al.[2] An az-nucleosome solution was first dialyzed to a SIRT7 assay buffer (20 mM Tris, 150 mM NaCl, pH 7.8) and then supplemented with 1 mM NAD^+ and 0.5 mM β-ME. The total nucleosome concentration was adjusted to 1 uM by adding the SIRT7 assay buffer (20 mM Tris, 150 mM NaCl, pH 7.8), and 0.5 uM SIRT7 was added to initiate the reaction. After 2 h, 20 mM nicotinamide was added to quench the reaction and 0.1 mM DBCO-MB488 (#1190-5 Click Chemistry Tools) was added to label unreacted az-nucleosomes for 1 h at room temperature. Deacylation products were then displayed by 8% Native TBE-PAGE gel electrophoresis, and deacylation percentage was obtained by comparing the fluorescent signal strength of nucleosome bands between the sample incubated with SIRT7 and control sample with no SIRT7 treatment. Fluorescent signals were normalized with EtBr staining signal of nucleosome bands. Both DBCO-MB488 and EtBr fluorescence were imaged and quantified by Bio-Rad Chemidoc XRS+. For time-dependent assay, 20 mM nicotinamide was added at each time point to quench the reaction and then the reaction tube was put on ice until all reactions finished.

SIRT7 deacylation of the H3K36Az-H4 tetramer. For the H3K36Az-H4 tetramer deacylation assay, the reaction solutions were prepared according to the following conditions: 1 uM His-SIRT7, 2 uM H3K36Az-H4 tetramer, 1 mM NAD^+, 0.5 mM β-ME, 20 mM Tris, 150 mM NaCl, pH 7.8 and 0.4 uM, 1 uM, 2 uM, or 4 uM free 147 bp 601 DNA. After incubating this reaction solution at 37 ℃ for 4 h, 20 mM nicotinamide was added to quench the deacylation, and 0.15 mM DBCO-MB488 dye was added to label

unreacted H3K36Az for 1 h at room temperature. 5 mM AzHeK was added to quench the unreacted DBCO dye and 2 U/unit DNase1 (ThermoFisher #M0303S) was added to digest free DNA. All reaction solutions were then flash-frozen by liquid nitrogen and lympholized to powders. The pellets were re-dissolved by 8 M Urea buffer (8 M Urea, 20 mM tris, pH 7.8), and 4X SDS buffer was added to boil the samples to 95℃ for 10 min, which were then separated in 15% PAGE gel. DBCO-MB488 fluorescent signals of histone H3 band was imaged by Bio-Rad Chemidoc XRS+. Deacylation percentage was derived from the DBCO-MB488 signal differences between control sample that was not treated with SIRT7 and reaction samples that were treated with SIRT7 and also had different amounts of DNA in their reaction mixtures. Gels were counterstained by Brilliant Blue G250 and imaged by Bio-Rad Chemidoc XRS+.

SIRT7 deacetylation of Ac-nucleosomes. H3 mutants with acetyl-lysine incorporated at various sites were expressed and purified from a CobB$^-$ E. coli BL21(DE3) strain according to procedures in our previous report.[1] Through PCR, biotinylated primers were used to prepare biotinylated 601 nucleosome positioning DNA, which was used later to assemble biotinylated ac-nucleosomes. For SIRT7 assay with biotinyl-ac-nucleosomes, 5 ul (total volume) of streptavidin resin (#L00353 GenScript) was added to a 20 ul reaction. At room temperature, the mixture was re-suspended 3 times once every 3 min, and centrifuged at 3000 RPM for 4 min. Supernatant was removed carefully by pipets, and 50 ul SIRT7 assay buffer was added to wash streptavidin resin twice. Histone proteins were eluted from streptavidin resin by adding 8 M urea buffer (8 M urea, 20 mM tris, 500 mM NaCl, pH 7.8) and 4X SDS loading buffer. After incubation at room temperature for

5 min, the mixture was centrifuged at 3000 RPM for 4 min and supernatant was boiled up and loaded onto 15% SDS-PAGE electrophoresis. Histone bands were transferred to nitrocellulose membrane by a standard semi-dry method using a Bio-Rad Trans-Blot Turbo transfer system. The membrane was coated with 5 % (w/v) fat-free milk (10 mL) for 1 h at room temperature and then treated with site-specific histone H3 acetylation antibody (#9927, #27683, Cell Signaling Technology) overnight at 4 ℃. The membrane was then washed with PBST buffer (10 ml PBS with 0.1% Tween-20) for 6 times with 2 min intervals, and then the membrane was treated with secondary antibodies for 1 h at room temperature. (#111-035-144, Jackson ImmunoResearch). After being washed by 10 ml PBST buffer for 6 times with 2 min interval, the membrane was treated with Pierce ECL Plus western substrates (#32132 ThermoFisher) and imaged by Bio-Rad Chemidoc XRS+.

Validation of H3K18ac and H3K36ac as in vivo targets of SIRT7. A human SIRT7 gene was cloned from a Pet28a-SIRT7 vector (provided by Dr. Hening Lin's group) to the pEGFP vector (#6085-1 Addgene) to afford pEGFP-SIRT7 and site-directed mutagenesis was used to generate an inactive H187Y mutant form of pEGFP-SIRT7. Then empty pEGFP, pEGFP-SIRT7 and pEGFP-SIRT7-H187Y plasmids were prepared by QIAGEN plasmid midi kits (#12143 QIAGEN), and transfected separately into HEK293T Cell cultures using Lipofectamine 3000 transfection reagents (#L3000008, ThermoFisher). 48 h later when above 80% HEK293T cells exhibited strong fluorescence, 5 uM TSA (#T5582, Sigma-Aldrich) was added to cell culture to induce overall increase of histone acetylation. After additional 4 h of incubation, cells from each group were collected, re-

suspended in a lysis buffer (10 mM Tris, 1 mM EDTA, 10 mM KCl, 1% Triton X-100, 5 uM TsA, 20 mM nicotinamide, protease inhibitor cocktail (#P8340, Sigma) pH 7.5) and vortexed several times while being incubated on ice for 30 min. The lysate was then centrifuged at 14000 RPM for 15 min and supernatant (cytoplasmic portion and nucleic proteins) was removed. Histone proteins were extracted from chromatin pellets by adding 0.2 N HCl and incubating on ice for 20 min. The supernatant that contained acidified histones was retrieved after being centrifuged at 14000 RPM and 4 ℃ for 20 min and neutralized by adding 1 M Tris buffer (pH 8.0). Histone concentration was determined by BCA assay (#23225 ThermoFisher). About 5ug of extracted histone mix was electrophoresed by 15**%** SDS-PAGE and transferred to PVDF membrane (#162-0177 Bio-Rad) according to the standard semi-dry protocol using the Bio-Rad Trans-Blot Turbo transfer system. The membrane was blotted by site-specific acetylation antibodies (#9927, #27683, #9674 Cell Signaling Technology) in the same manner as shown the previous paragraph.

Locus-specific monitoring of H3K36ac change in vivo. For ChIP experiments, control rabbit IgG, anti-H3 (Abcam), anti-H3K36ac (Active Motif or Millipore) were used in combination with Protein A coated magnetic beads (Dynabeads, Invitrogen #10001D). For ChIP sequencing, anti-H3K36ac (Active Motif) and Spike-in anti-H2Av (Active Motif) were used. For immunoblotting, the following antibodies were used: anti-H3 (Covance), anti-H3K36ac (Millipore or Diagenode), anti-FBL (Abcam), anti-UBF (SCBT), anti-p84 (GeneTex), anti-SIRT7 (Covance), anti-β-tubulin (Millipore) in

combination with secondary anti-rabbit or anti-mouse conjugated to HRP (Jackson ImmunoResearch). Nucleolar fractionation was performed as previously described.[3]

For chromatin immunoprecipitation experiments, U2OS cells were crosslinked with 1% formaldehyde for 5 minutes at room temperature and the reaction was quenched by adding glycine to a final concentration of 1.25 M for 5 more min at room temperature. Cells were washed twice with ice-cold PBS and ressuspended and sonicated in ChIP lysis buffer to generate 200-300 bp fragments. Chromatin was diluted with low-SDS RIPA buffer and 25 ug chromatin was incubated overnight at 4 ºC with antibody coated protein A beads. Next day, after 3x washes with RIPA buffer, followed by 1x wash with TE buffer, immunoprecipitated chromatin was reverse crosslinked and DNA was purified with a PCR purification kit. For ChIP-PCR, ChIP DNA was analyzed using SYBR Green Master Mix in a Lightcycler480 II (Roche). The following primers were used: RPS20,[4] COPS2,[4] MRPS18b Forward *ctcacccaggattggtagaaa* - Reverse *acgactcttagatcgcattgtt*, 18S Forward *aattgacggaagggcaccac* – Reverse *tcaatcctgtccgtgtccg.* ChIP-seq library was constructed using the Truseq ChIP kit from Illumina following the manufacturer instructions. Sequencing was performed in a HiSeq Illumina4000 platform with 2x150 reads. 0.05% Drosophila spike-in chromatin (Active Motif) was included for normalization.

For the repetitive sequence analysis, PCT and CT consensus or chromosome-specific sequences, and control sequences were obtained from Eymery et al.[5] along with the 18S and 28S rDNA, and telomeric $(TTAGGG)_8$ sequence. Because these sequences were shorter than the sequencing read length (150 bp), the number of reads containing the

analyzed sequences were counted, and up to two mismatches were allowed. The read counts for each sample were normalized to total read count.

Nucleosome accessibility assay with restriction endonuclease Pst1. Modified 601 nucleosome positioning DNA that contained a Pst1 site at the entry/exit nucleosome region was prepared as what was previously designed.[6] Primers encompassing the mutated Pst1 restriction site were used to prepare modified 601 DNA through polymerase chain reaction and purified by a PCR cleanup kit (#2360250 Epoch Life Science). Mononucleosomes containing w.t. H3, H3K18ac and H3K36ac were assembled similar to az-nucleosomes, except an additional thermal repositioning step at 60 ℃ for 20 min to generate nucleosome samples with homogenous DNA positioning. The final nucleosome solutions were dialyzed to 20 mM Tris buffer (pH 7.9). For Pst1 digestion assay, 10X NEB3.1 buffer (#B7203S, NEB) was added to provide all buffer ingredients and 0.49 uM w.t., H3K18ac and H3K36ac nucleosomes were incubated with 2 U/ul Pst1 (#R0140S NEB) at 37 ℃. At each time point, 14 ul of reaction solution was taken out and added 2.5 ul 100 mM EDTA solution to quench the reaction. Reaction products of each time points were electrophoresed by 5% Native 1X TBE-PAGE gel, and digestion efficiency was estimated by comparing the EtBr staining signal strength between original nucleosome band and the lower digested nucleosome band.

Mass determination of Histone H3 protein using HPLC-MS. Histone pellets were dissolved in 0.1% (w/v) formic acid solution to final concentration of 1 mg/ml. 2 ul of Histone H3 solution was injected into Thermo Scientific Q Exactive Focus Orbitrap LC-MS/MS System equipped with a Sigma BIOshell A400 Protein C4, 3.4 um HPLC

column (10 cm x 2.1 mm). Histone H3 was separated with a water/acetonitrile gradient with flow phase A as water with 0.1% (w/v) formic acid and flow phase B as acetonitrile with 0.1% (w/v) formic acid. From 0 to 10 min, ran HPLC at 0.2 ml/min with linear gradient from 20% B to 60% B. From 10 to 15 min, ran at 0.2 ml/min with linear gradient from 60% B to 100% B. Then ran with 100% B at 0.2ml/min to clean up the column from 15 min to 20 min. ESI-MS data was acquired in the positive ion mode with the following parameters: Resolution: 70000, Range: 600-1400 m/z, AGC:3e6, Max inject time: 150ms, Sheath gas flow rate: 45, Aux gas flow rate: 30, Spray voltage 3.75kv, Capillary temperature: 320, RF value 20, In-source CID 30ev.

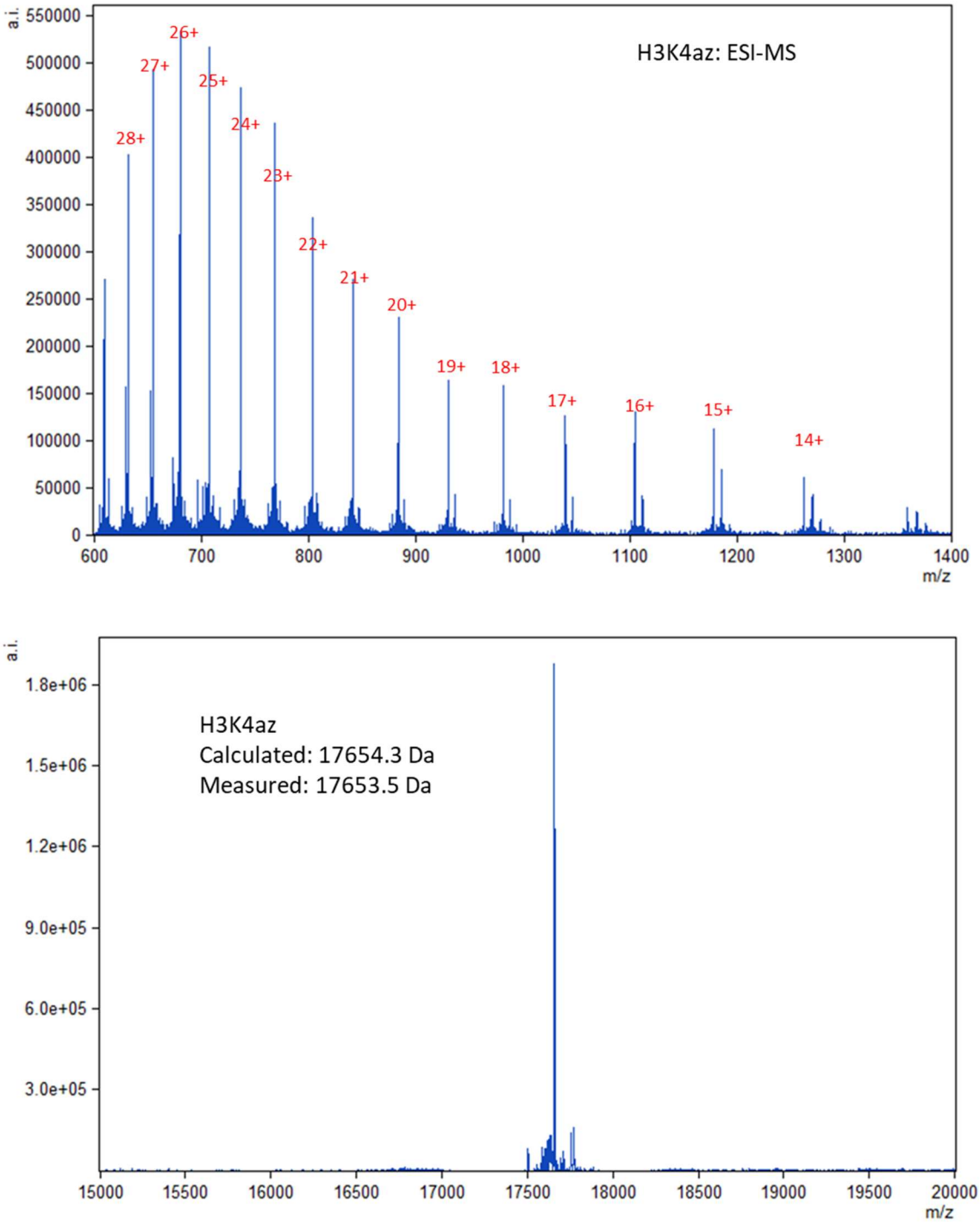


Figure 37 ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

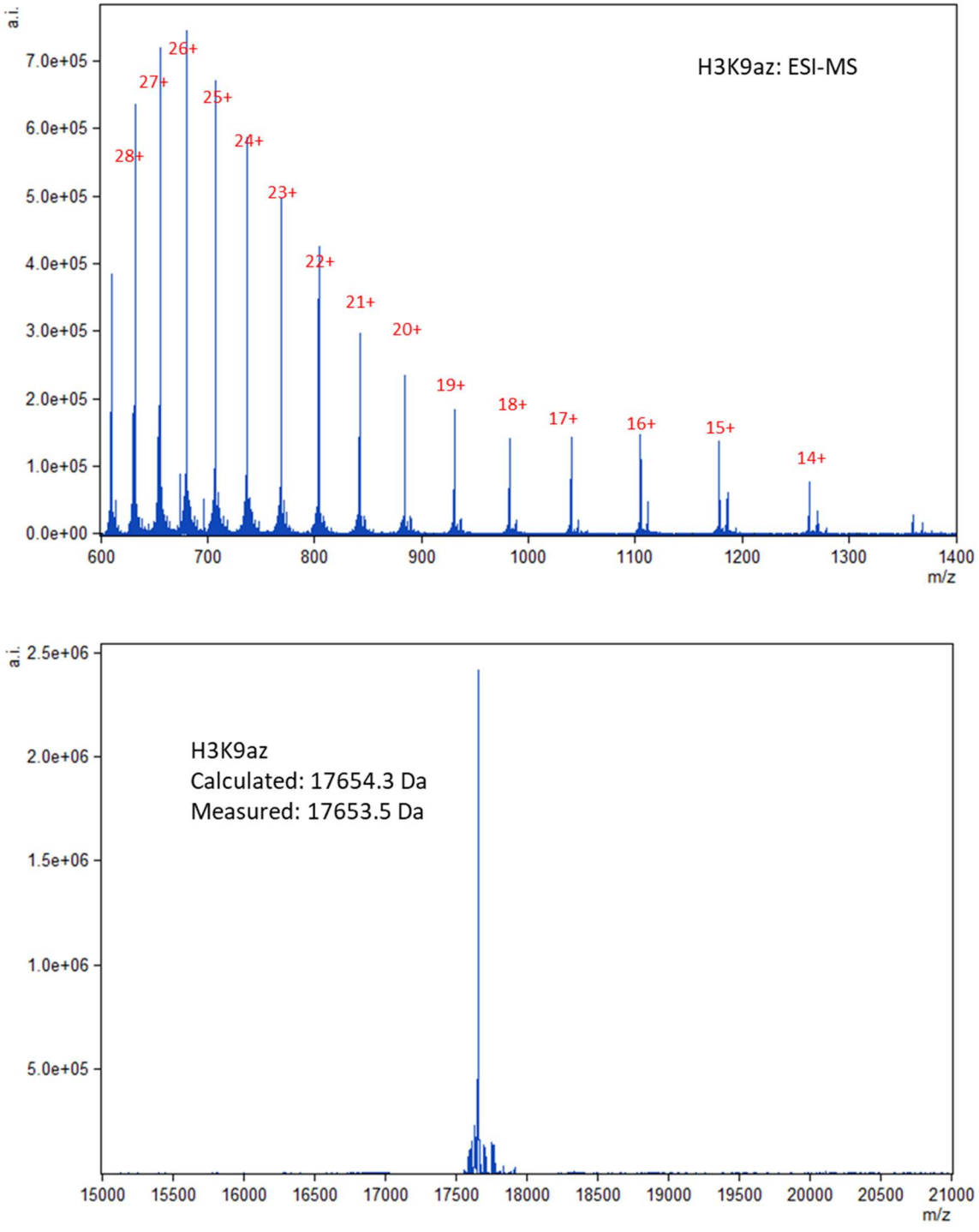


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

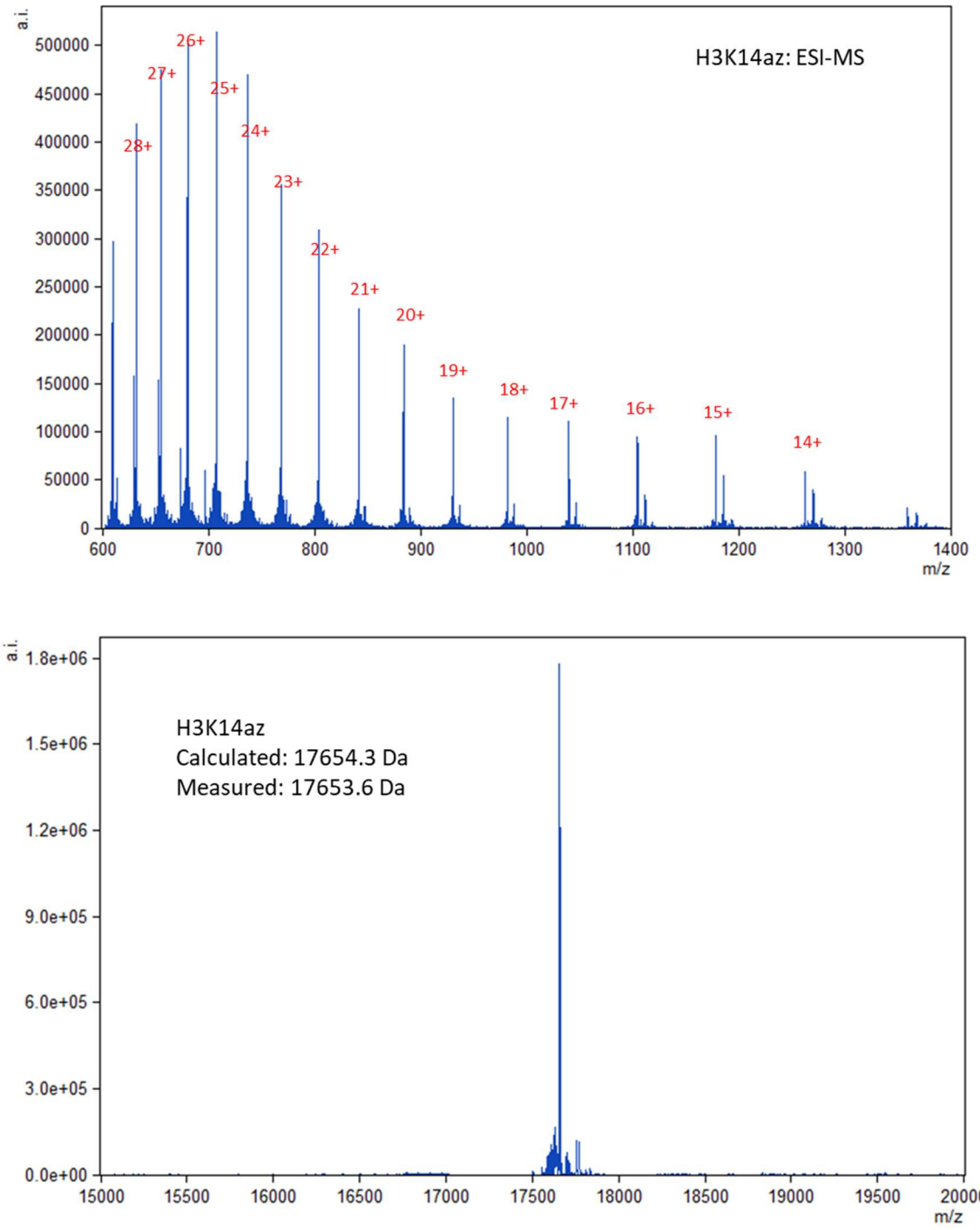
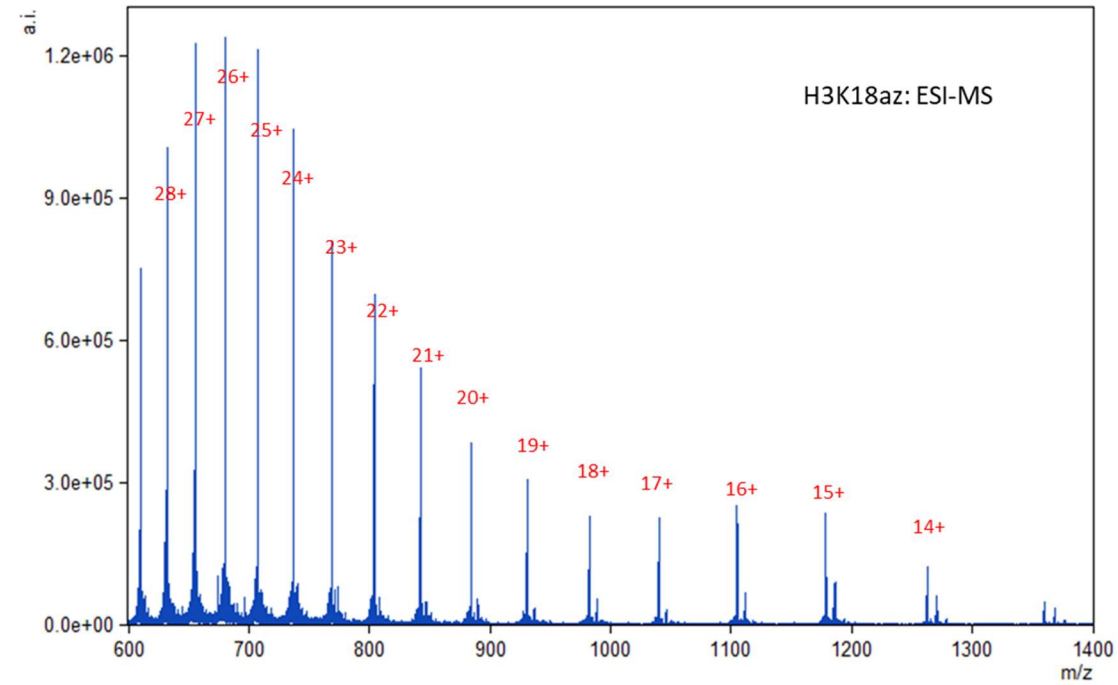


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3



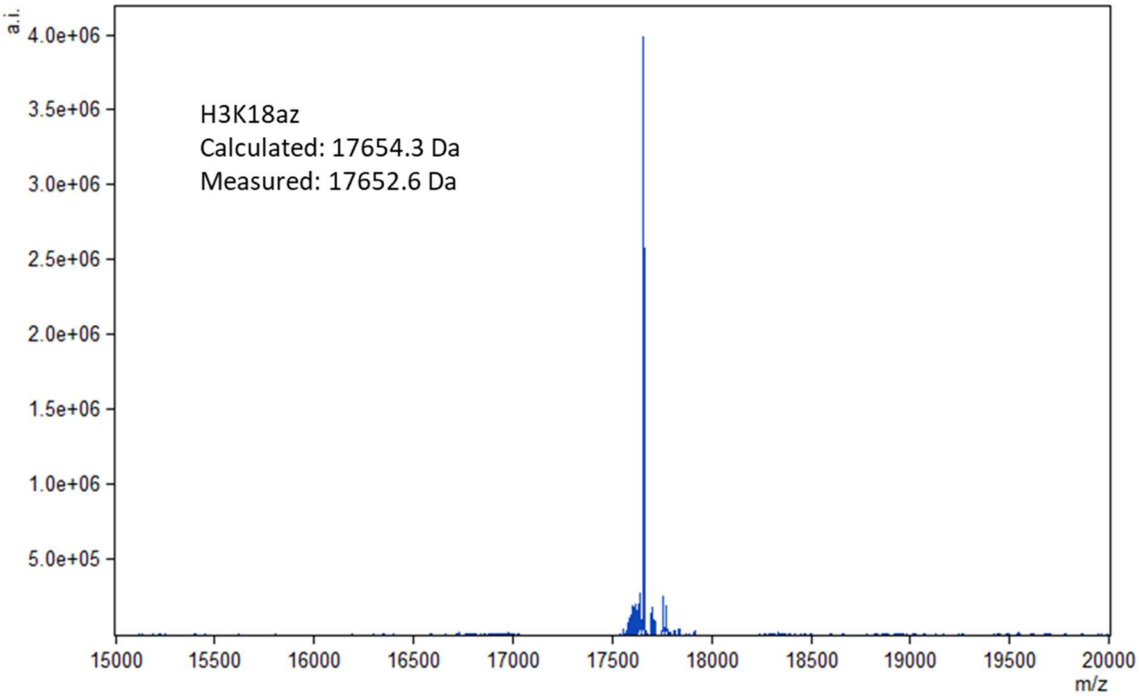


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

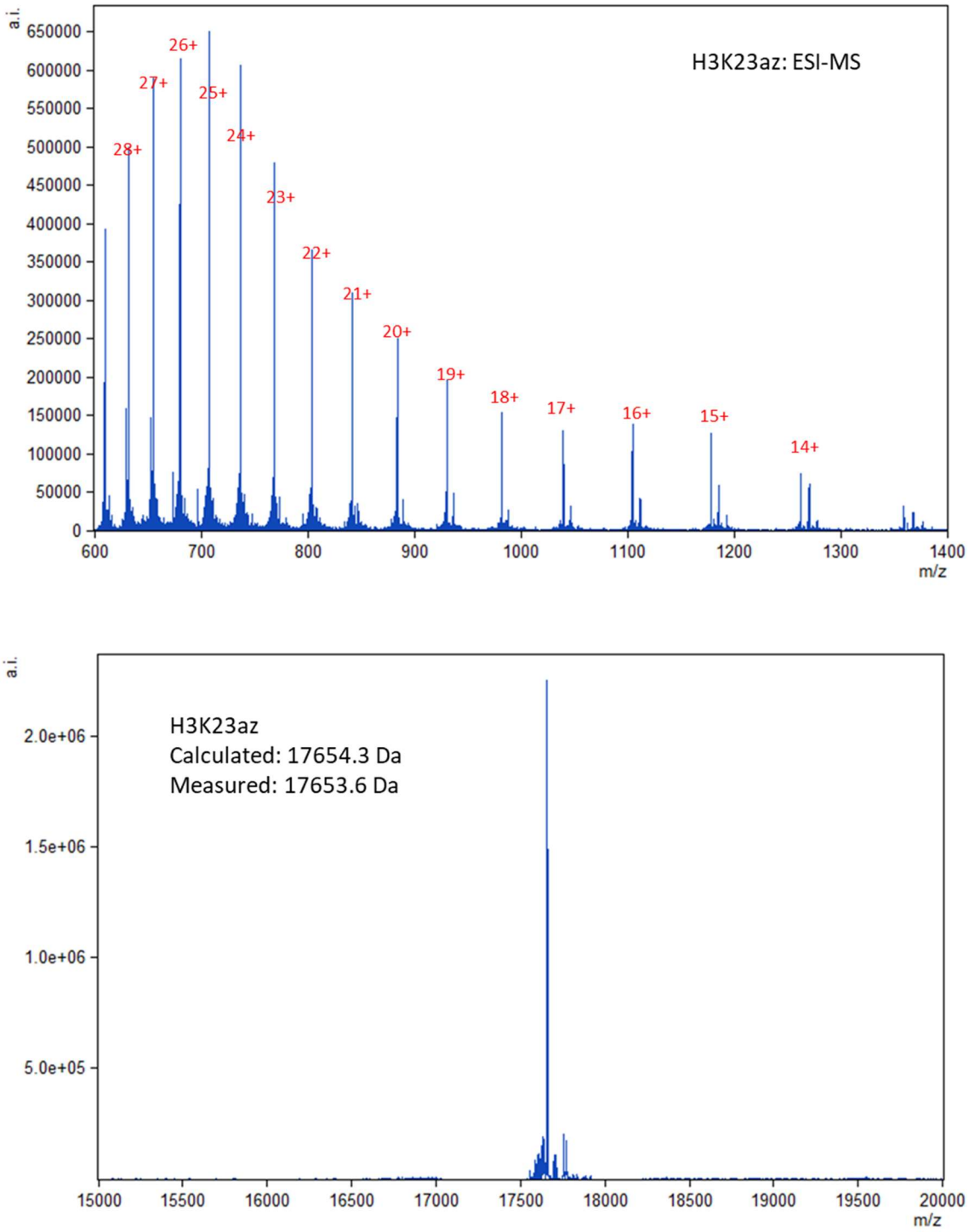


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

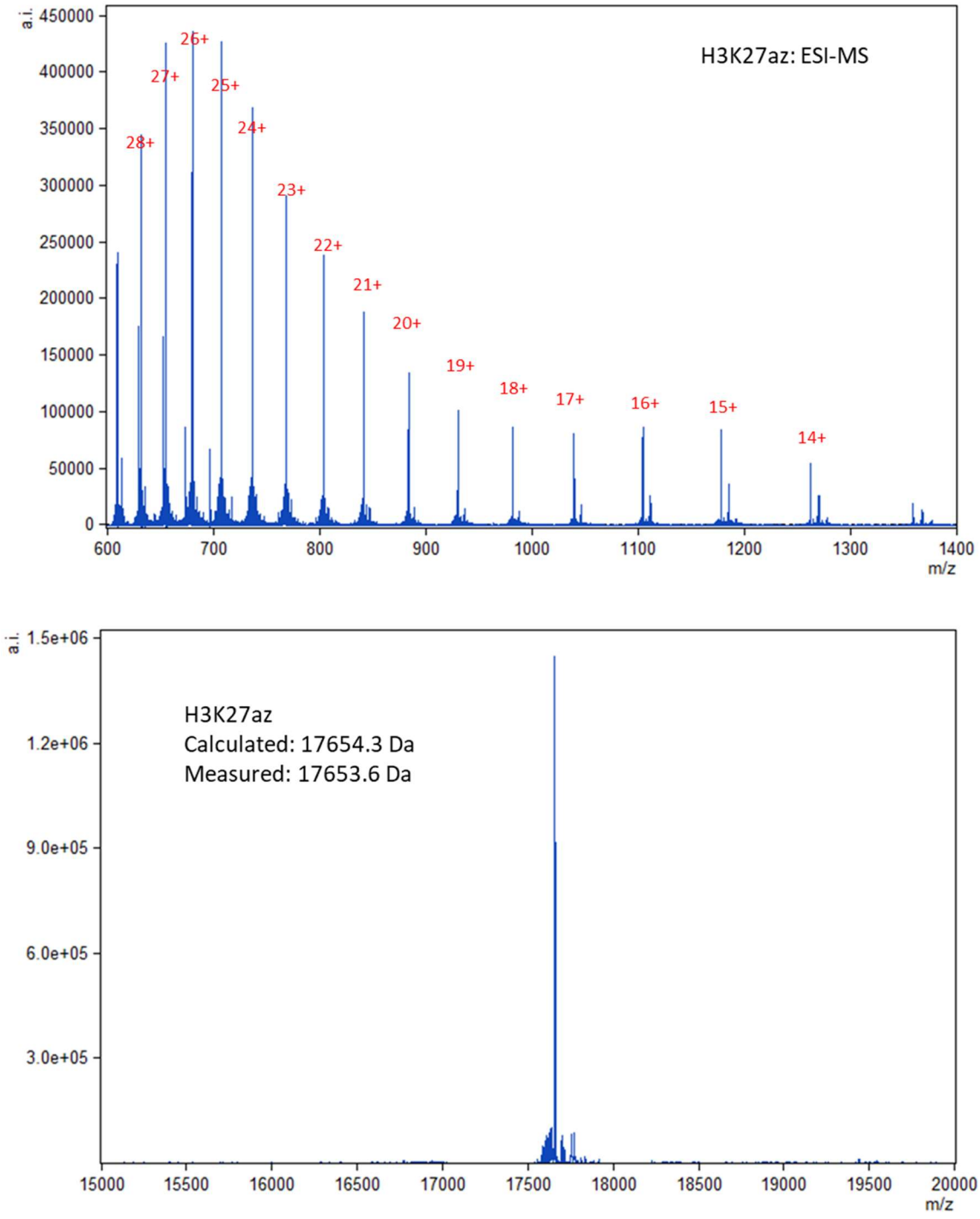


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

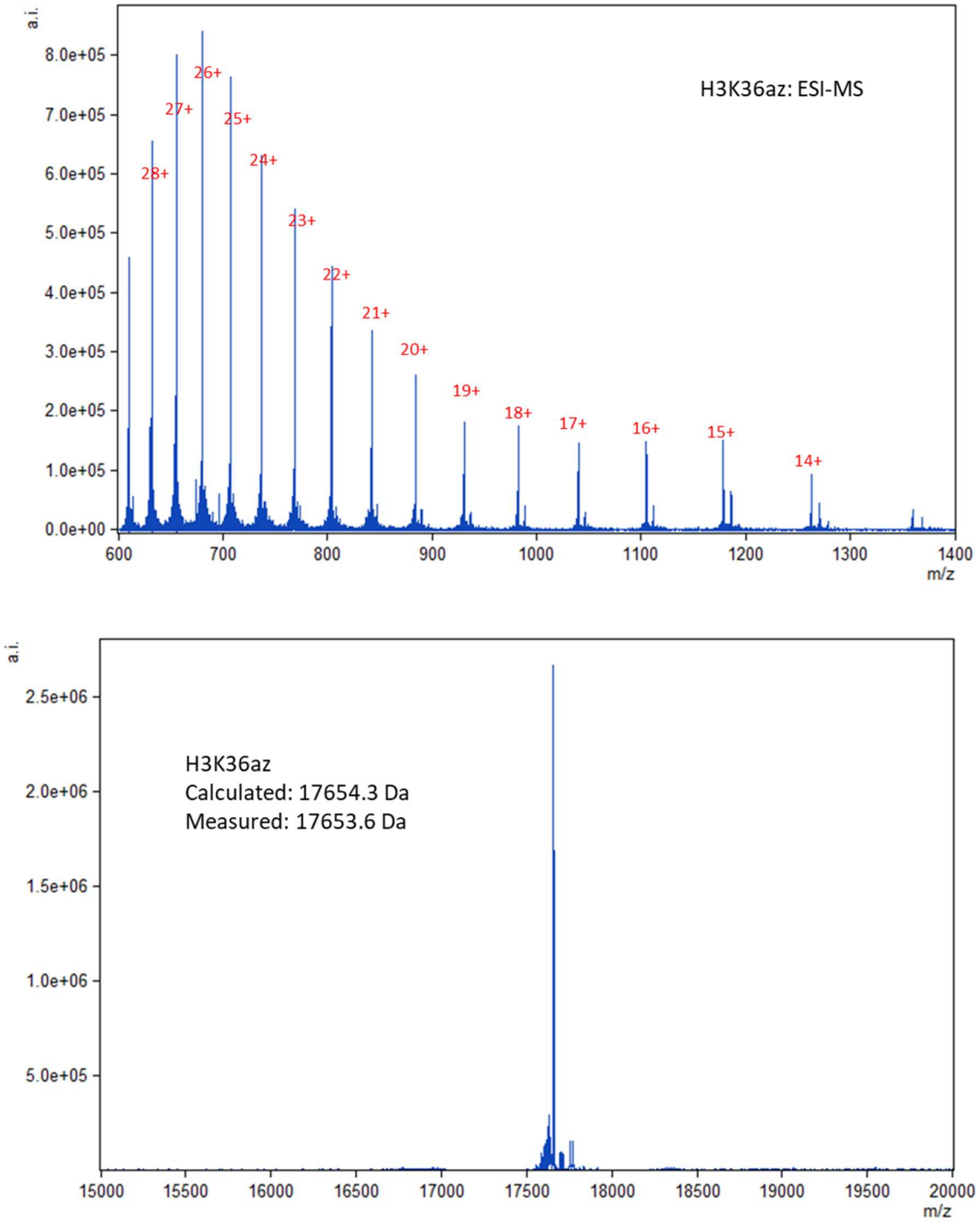


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

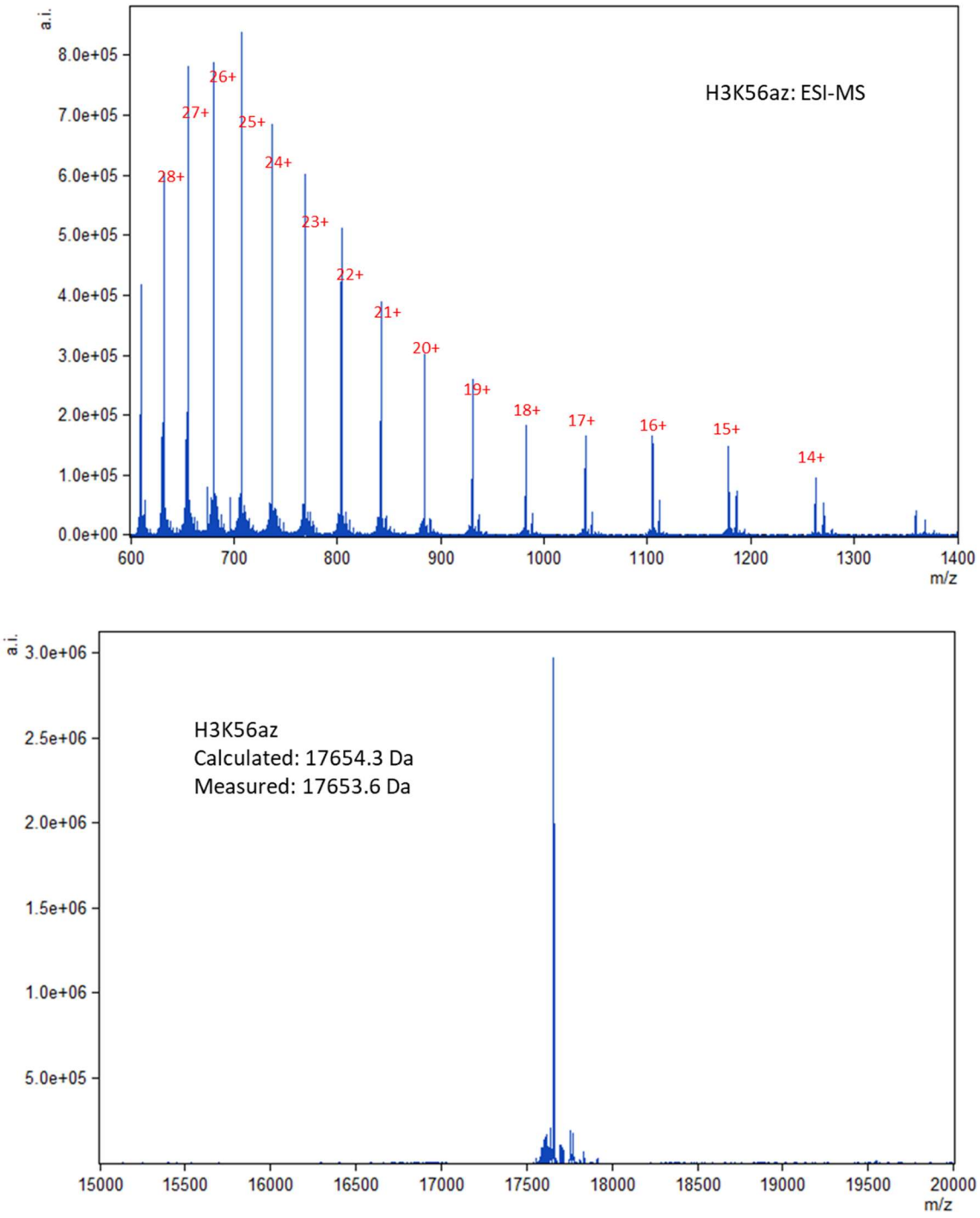


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

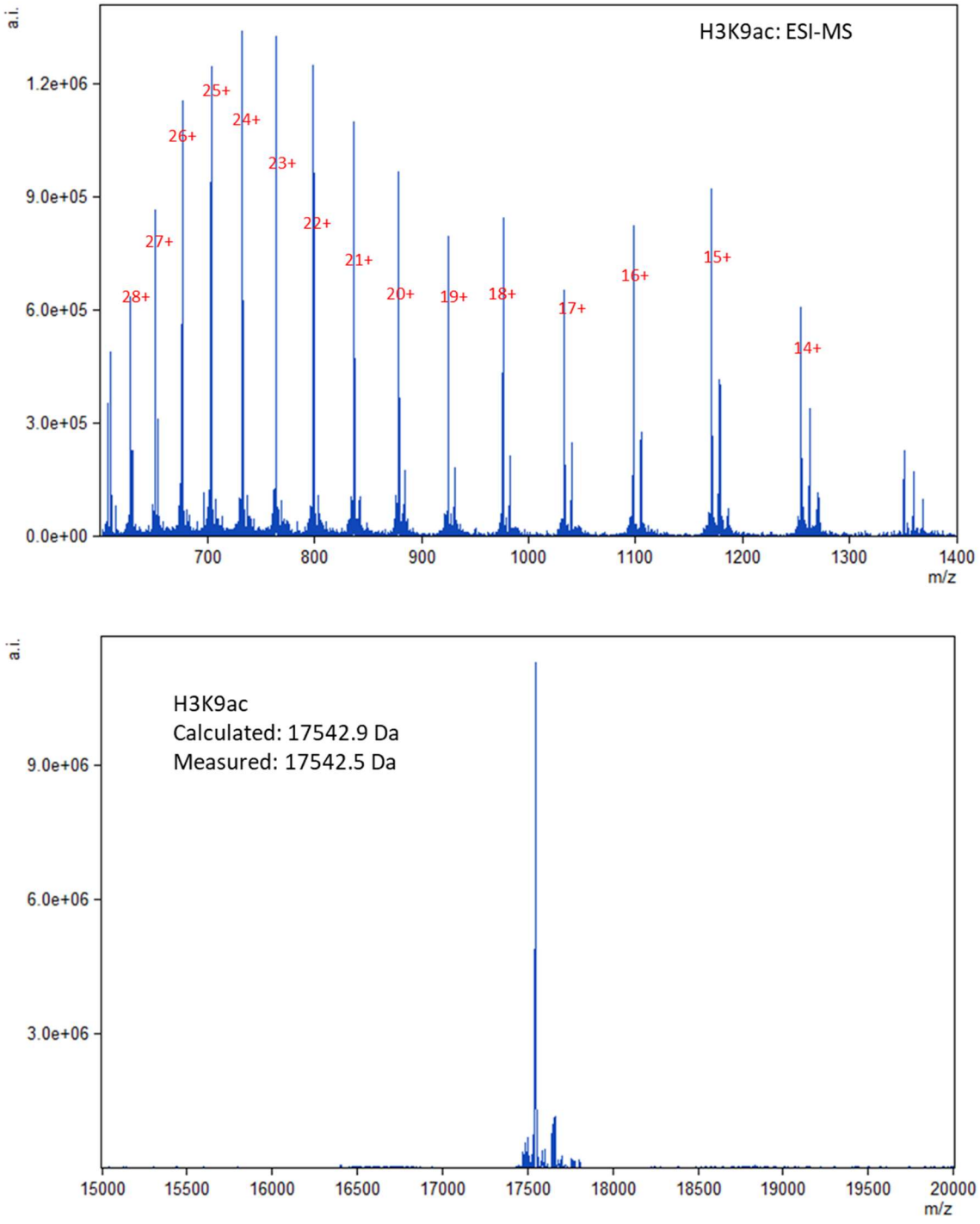


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

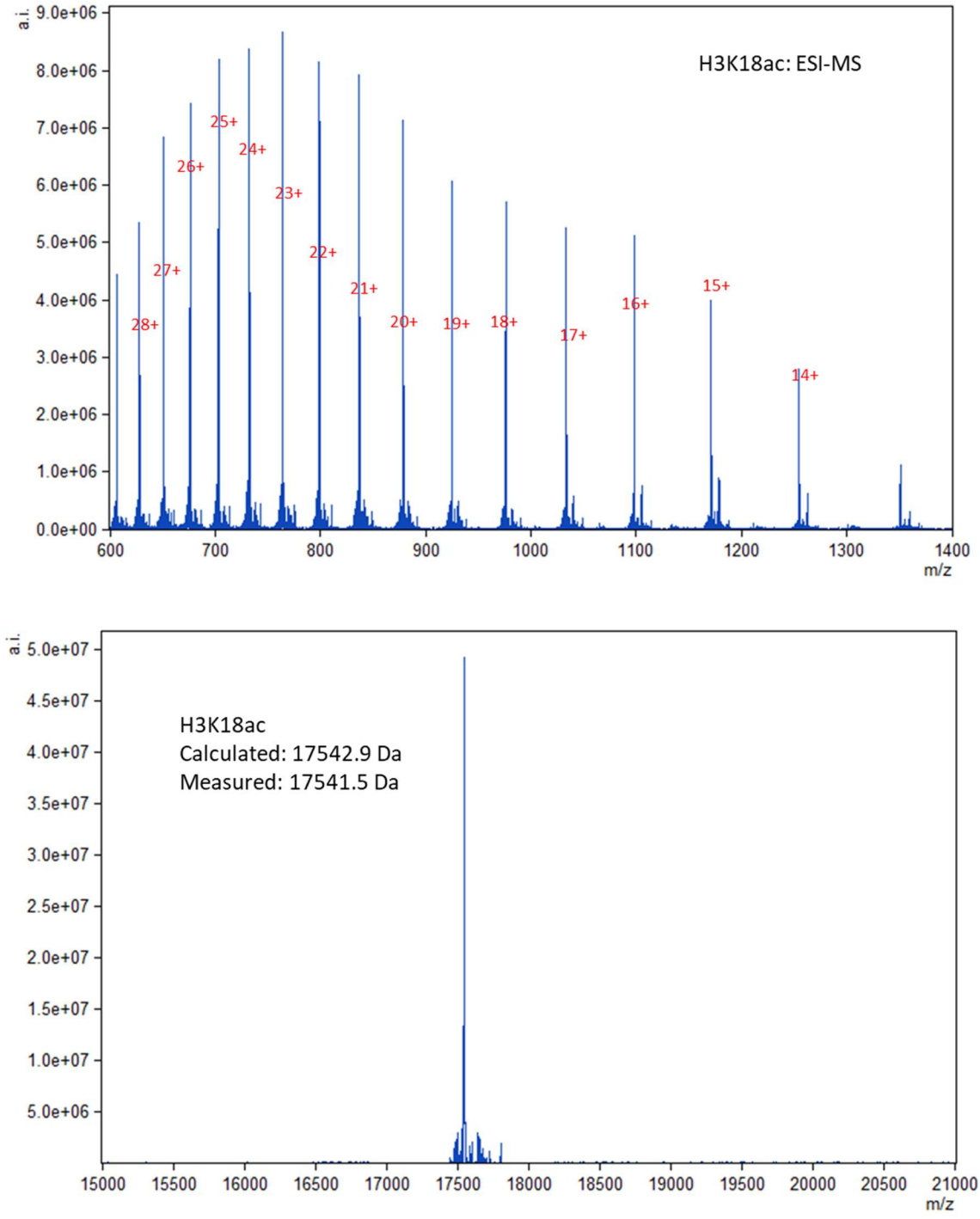


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

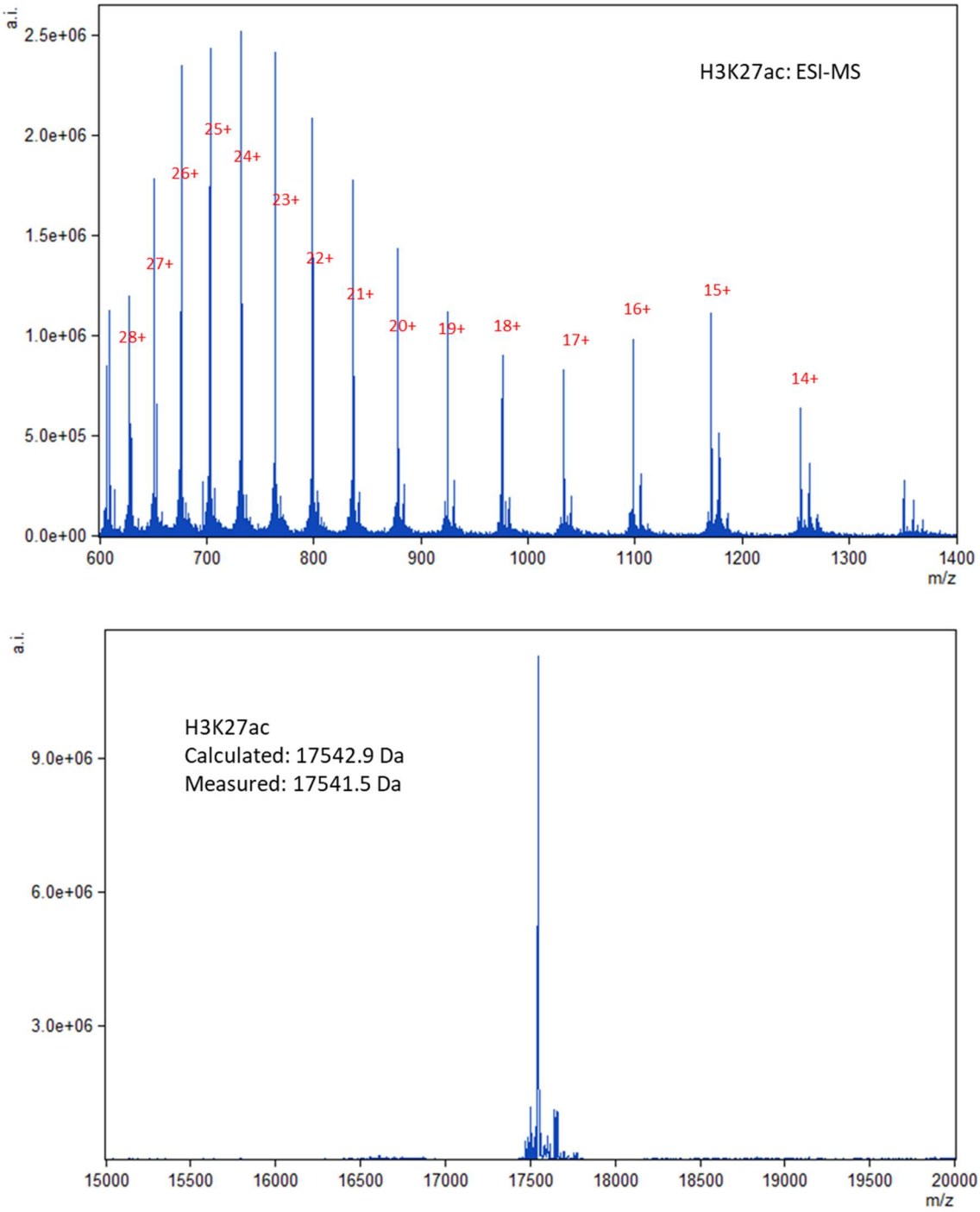


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

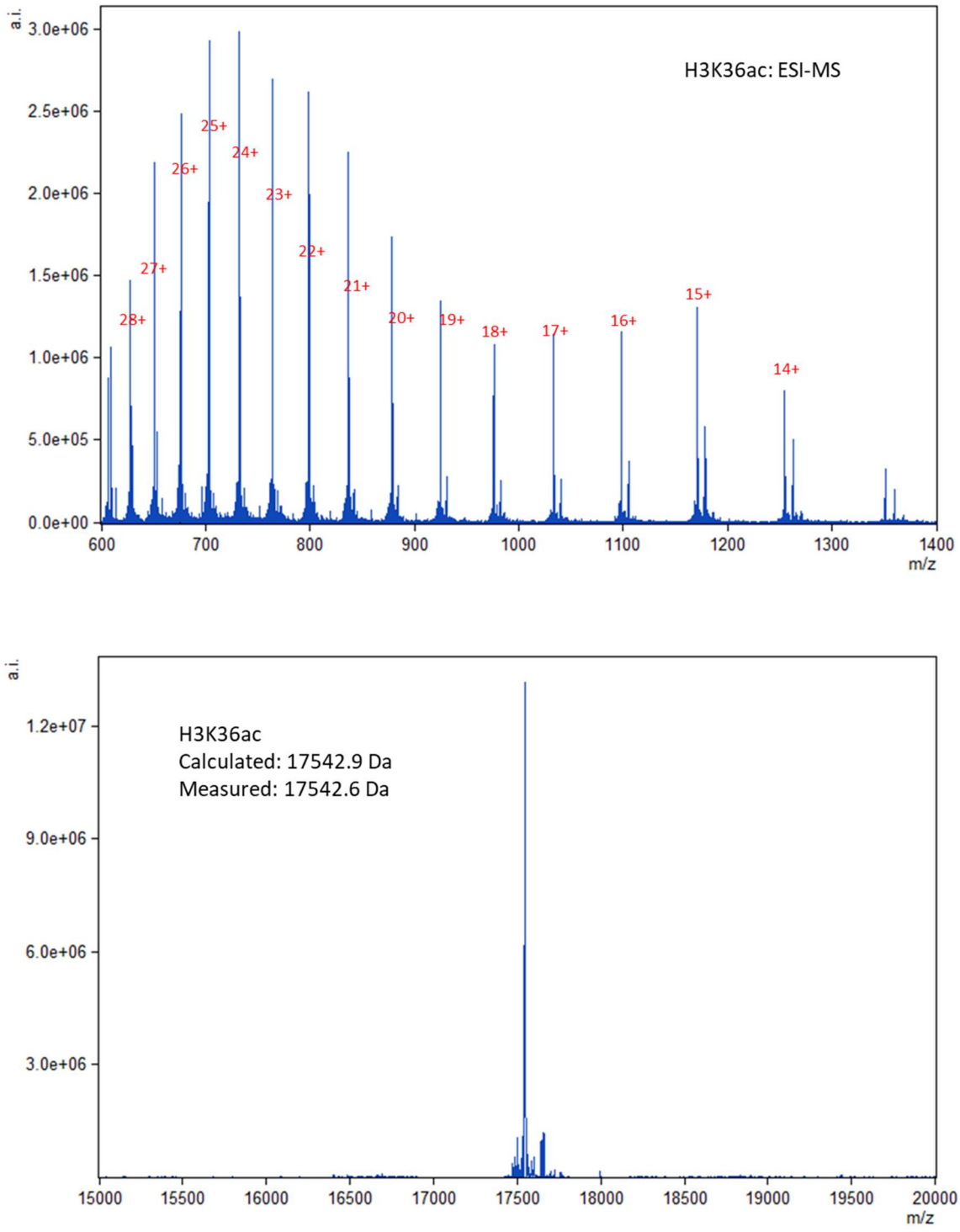
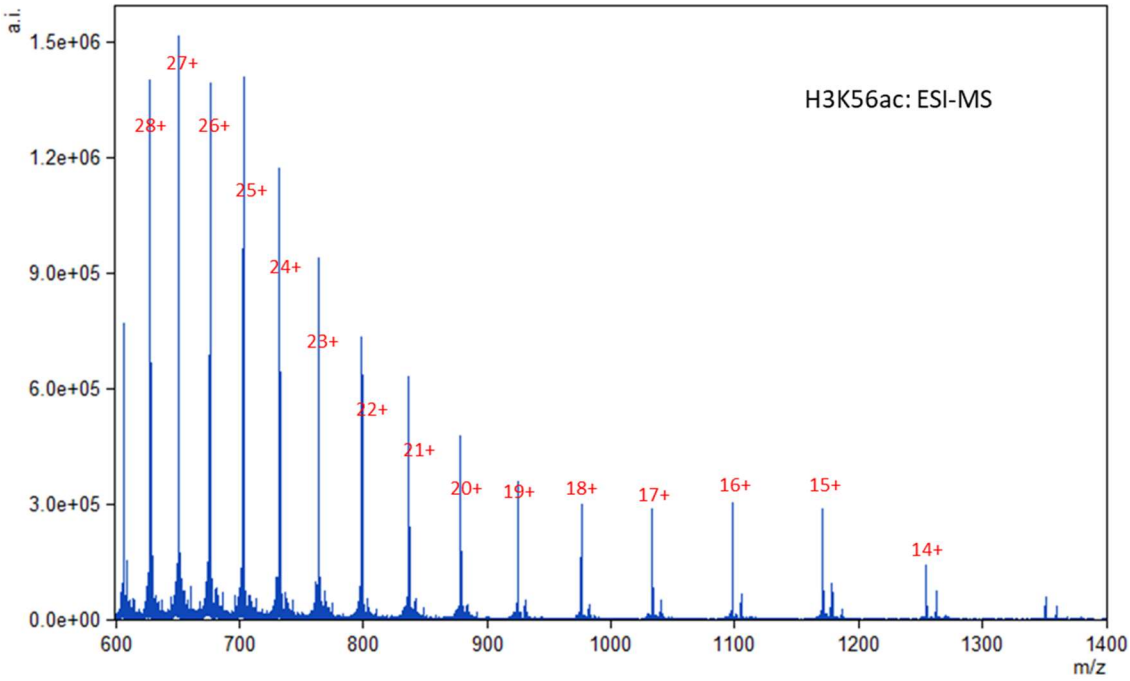


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3



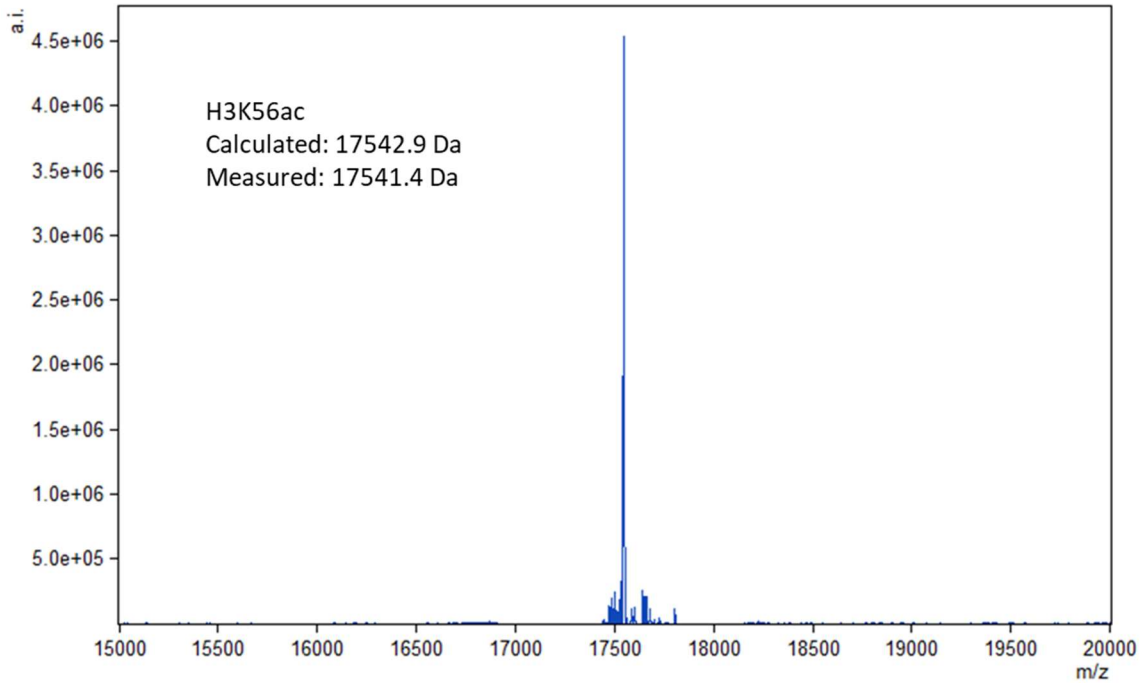


Figure 37 (continued) ESI-MS spectra of recombinant histone Az-H3 and Ac-H3

Primers and gene sequences

pEGFP-SIRT7 → pEGFP-SIRT7-H187Y:

F-primer: TATGGGAACATGTACATTGAAGTCTGTACCTC

R-primer: GAGCTCGGAGATGGCCGTG

601 DNA with no linker DNA attached:

F-primer: CTGGAGAATCCCGGTGCCG

R-primer: ACAGGATGTATATATCTGACACGTGC

601 DNA with 5bp linker DNA:

F-primer: GTAATCTGGAGAATCCCGGTGCCG

R-primer: TAAGGACAGGATGTATATATCTGACACGTGC

601 DNA with 10bp linker DNA:

F-primer: GTAATGTAATCTGGAGAATCCCGGTGCCG

R-primer: TAAGGTAAGGACAGGATGTATATATCTGACACGTGC

601 DNA with 20bp linker DNA:

F-primer: GTAATGTAATGTAATGTAATCTGGAGAATCCCGGTGCCG

R-primer:

TAAGGTAAGGTAAGGTAAGGACAGGATGTATATATCTGACACGTGC

601 DNA with pst1 restriction site:

F-primer: GGACCCTATACGCGCGCGCCCTGCAGAATCCCGGTGCCG

R-primer: TACATGCACAGGATGTATATATCTGACACGTGC

His-SIRT7:

atgcatcaccatcatcaccacgcagccgggggtctgagccgctccgagcgcaaagcggcggagcgggtccggaggttgcg

ggaggagcagcagagggagcgcctccgccaggtgtcgcgcatcctgaggaaggcggcggcggagcgcagcgccgagg

agggccggctgctggccgagagcgcggacctggtaacggagctgcagggccggagccggcggcgcgagggcctgaag

cggcggcaggaggaggtgtgcgacgacccggaggagctgcgggggaaggtccgggagctggccagcgccgtccggaa

cgccaaatacttggtcgtctacacaggcgcgggaatcagcacggcagcgtctatcccagactaccggggccctaatggagtgt

ggacactgcttcagaaagggagaagcgttagtgctgccgacctgagcgaggccgagccaaccctcacccacatgagcatca

cccgtctgcatgagcagaagctggtgcagcatgtggtgtctcagaactgtgacgggctccacctgaggagtgggctgccgcg

cacggccatctccgagctccacgggaacatgtacattgaagtctgtacctcctgcgttcccaacagggagtacgtgcgggtgtt

cgatgtgacggagcgcactgccctccacagacaccagacaggccggacctgccacaagtgtgggacccagctgcgggaca

ccattgtgcactttggggagaggggacgttggggcagcctttgaactgggaagcggcgaccgaggctgccagcagagca

gacaccatcctgtgtctagggtccagcctgaaggttctaaagaagtacccacgcctctggtgcatgaccaagccccctagccg

gcggccgaagctttacatcgtgaacctgcagtggaccccgaaggatgactgggctgccctgaagctacatgggaagtgtgat

gacgtcatgcggctcctcatggccgagctgggcttggagatccccgcctatagcaggtggcaggatcccatttttctcactggcg

actcccctgcgtgctggtgaagaaggcagccacagtcggaagtcgctgtgcagaagcagagaggaggccccgcctgggga

ccggggtgcaccgcttagctcggcccccatcctaggggggctggtttggcaggggctgcacaaaacgcacaaaaaggaagaa

agtgacgtga

Results and discussion

A modified click chemistry-based assay for quick profiling of SIRT7-targeted chromatin lysine deacylation sites. Although originally defined as histone deacetylases, accumulated evidence has established that sirtuins catalyze the removal of a spectrum of acylation types from protein lysines.[199, 267, 269-275] Therefore, they are more generally categorized as protein lysine deacylases. Specifically for SIRT7, studies have demonstrated that it catalytically removes fatty acyl and succinyl groups from histone lysines.[267, 275] Through genetic incorporation of N^{ε}-(7-octenoyl)-lysine (OcK) into histones in conjunction with the alkene-tetrazine cycloaddition reaction to analyze fatty acylation levels in acyl-nucleosome substrates that we assembled *in vitro* from OcK-incorporated histones, we previously revealed that SIRT6 is catalytically active to remove acylation from multiple H3 lysine sites.[276] Given the structural similarity between SIRT6 and SIRT7 and the demonstrated activity of SIRT7 to remove fatty-acylation from protein lysines, a similar approach might be applied to profile SIRT7-targetted H3 lysine deacylation sites. However, the previously described alkene-tetrazine cycloaddition-based method has a significant drawback. The slow reaction (the second order rate constant: ~0.01 $M^{-1}s^{-1}$) between a terminal alkene and a tetrazine requires a long reaction time.[218] Given the delicate nature of acyl-nucleosome substrates, this long reaction time is particularly undesirable. To accelerate the labeling reaction, we envisioned that N^{ε}-(7-azidoheptanoyl)-L-lysine (AzHeK) could be installed in acyl-nucleosome substrates and expected that AzHeK's close mimic of N^{ε}-decanoyl-L-lysine (DeK) would allow efficient recognition by SIRT7 for defatty-actylation. AzHeK has an

azide functionality that reacts favorably with a dibenzocyclooctyne (DBCO) dye (the second order rate constant: ~1 $M^{-1}s^{-1}$)[277-279] for quick labeling of acyl-nucleosome substrates. Therefore, we can potentially use acyl-nucleosomes with AzHeK installed at different H3 lysine sites as SIRT7 substrates for defatty-acylation and then analyze

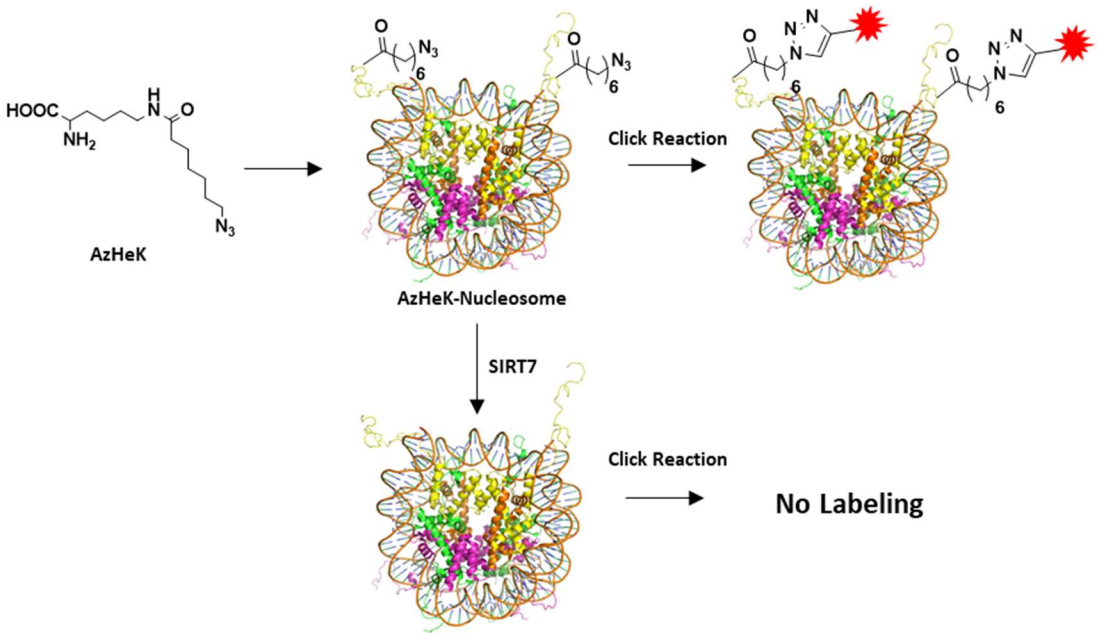


Figure 38 A click chemistry-based approach to profile SIRT7-targeted chromatin lysine deacylation sites. AzHeK that structurally resembles DeK is site-specifically incorporated into a histone that is further assembled *in vitro* with other histones and 601 DNA into an acyl-nucleosome. This acyl-nucleosome can be fluorescently labeled with a strained alkyne dye and subsequently visualized in a 1× TBE native-PAGE gel. When SIRT7 is catalytically active to remove fatty acylation from the incorporated AzHeK to recover lysine, incubating the assembled acyl-nucleosome with SIRT7 will remove the fatty acylation and therefore afford a deacylated nucleosome that cannot be fluorescently labeled and then visualized in a gel.

relative activities of SIRT7 toward different acyl-nucleosomes by probing residual fatty-acylation on these substrates using a DBCO dye when their reactions with SIRT7 finish (**Figure 38**). Accordingly, an effective assay for profiling SIRT7 targeted H3 lysine

deacylation sites can be potentially developed. This novel assay will allow not only straightforward determination of SIRT7 deacylation activities at different nucleosome lysine sites but also their kinetic characterization. In comparison to antibody-based detection, this designed assay is much simpler and more accurate.

The genetic incorporation of AzHeK into a protein allows its click labeling with a strained alkyne dye. Following a synthetic route shown in **Figure 34**, we synthesized AzHeK in a large quantity. For its genetic incorporation, we employed our previously identified OcKRS.[276] We have shown that this MmPylRS mutant actively charges $tRNA^{Pyl}$ with OcK and DeK for their incorporation at amber codon during translation in *E. coli* (data not shown). When growing *E. coli* BL21(DE3) cells that harbored plasmids coding genes for OcKRS, $tRNA_{CUA}^{Pyl}$, and ubiquitin-6×His with an amber mutation at the K48 position in the presence of 1 mM AzHeK, we successfully induced the expression of full-length ubiquitin UbK48az (**Figure 39b**). The expression level was 18 mg/L, equivalent to the expression of Ub with OcK incorporated at K48. Wild type Ub has an expression level around 100 mg/L in the same conditions. On the contrary, we were not able to induce the expression of ubiquitin when AzHeK was absent. After incubating UbK48az with 100 μM MB488-DBCO, a strained alkyne dye for 1 h, we observed a strongly labeled ubiquitin band when we analyzed the protein fluorescently in a SDS-PAGE gel (**Figure 39c**). However, we were not able to label wild-type ubiquitin using the same condition. In comparison to the previous alkene-tetrazine cycloaddition-based labeling that requires a minimum of 8 h for close to 40% labeling of a OcK-containing protein in the presence of a 200 μM tetrazine dye, this new labeling was much quicker

and led to close to quantitative labeling in about 30 min (**Figure 40a**). The azide group in UbK48az was stable at room temperature for about 1 h when 1 mM β -mercaptoethanol (BME) was present (**Figure 40b**). We also characterized the purified UbK48az using MALDI-TOF mass spectrometry (**Figure 40d**). The MALDI-TOF-MS-determined molecular weight (9540.9 Da) agreed well with the calculated one (9540.7 Da).

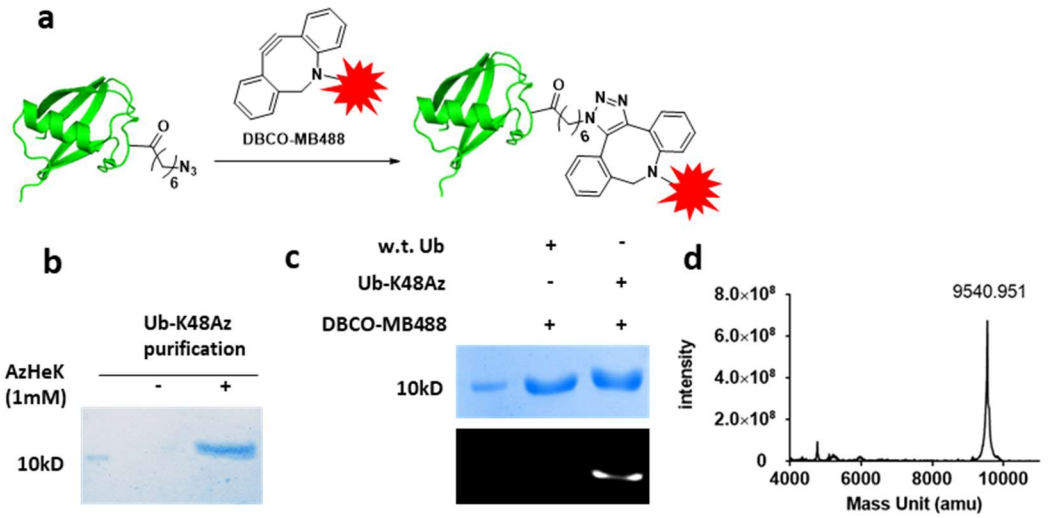


Figure 39 genetic incorporation of AzHeK into Ubiquitin at K48 position. (**a**) AzHeK incorporated at K48 position of ubiquitin could be site-specifically labelled by DBCO-MB488 through copper-free click reaction (**b**) test of AzHeK incorporation efficiency with previously reported mmOcKRS/tRNA^{Pyl} pair in E. coli BL21(DE3) CobB^{-} cell line in the following condition: 0.5 mM IPTG, 0.2% (w/v) L-arabinose, 1 mM AzHeK. (**c**) Labeling test of Ub-K48Az with DBCO-MB488 in the following condition: 5 uM ubiquitin, 0.1 mM DBCO-MB488, PBS buffer, pH 7.4. (**d**) MALDI-TOF mass of UbK48-AzHeK (measured: 9541.0, calculated: 9540.8)

SIRT7 catalyzes removal of acylation from H3 K18 and H3 K36 in the context of nucleosomes. After confirming that OcKRS was efficient in mediating the genetic incorporation of AzHeK at amber codon, we proceeded to recombinantly express H3

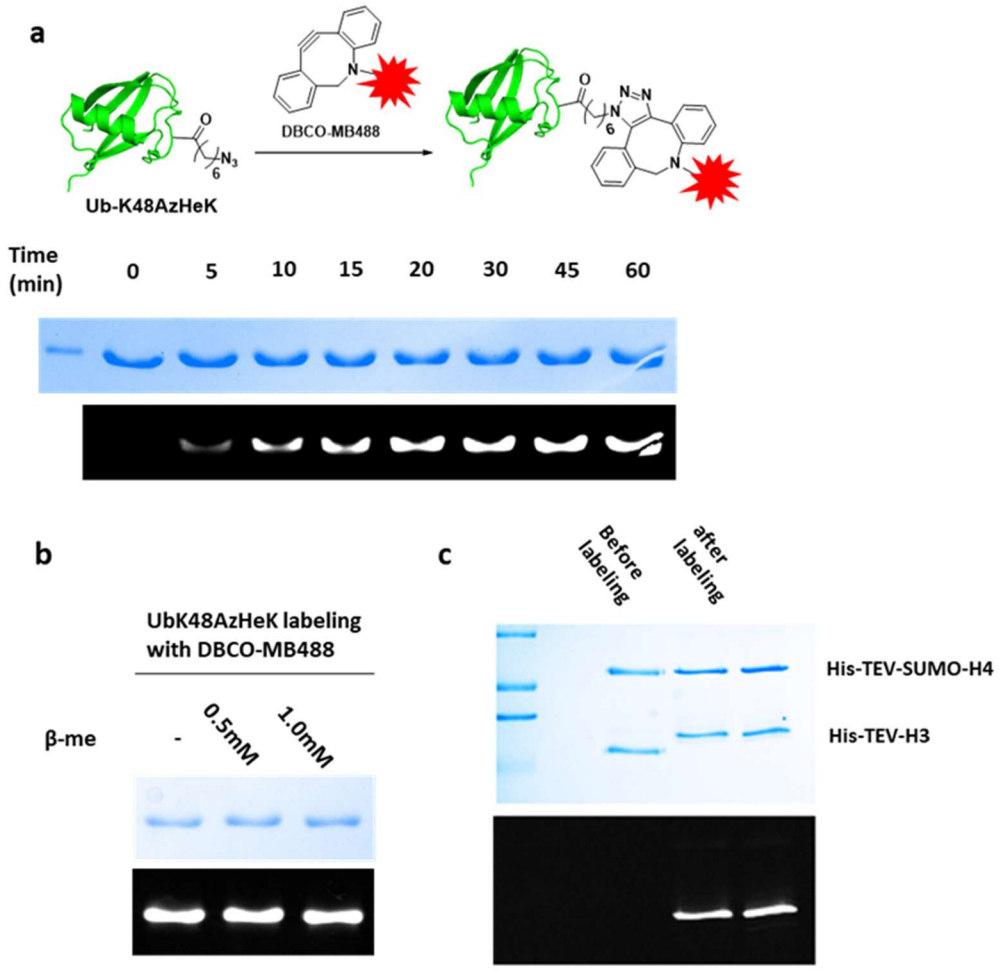


Figure 40 (**a**) Time-dependent labeling of Ub-K48Az by DBCO-MB488 in the following conditions: 5 uM UbK48Az, 0.1 mM DBCO-MB488, PBS buffer pH 7.4 at r.t. (**b**) AzHeK stability in the presence of β-ME (Upper lane: Commassie blue staining; Bottom lane: FITC signal detection). (**c**) Labeling of His-TEV-H3K36Az by DBCO-MB488 caused band position shift on the 15% PAGE gel.

proteins with AzHeK incorporated at K4, K9, K14, K18, K23, K27, K36, and K56, respectively. These AzHeK-containing H3 histones were named as H3K4az, H3K9az, and so on. To recombinantly produce a specific AzHeK-containing H3 protein, we transformed *E. coli* BL21(DE3) cells with pEVOL-OcKRS that contained genes coding

OcKRS and $tRNA_{CUA}^{Pyl}$ and a pETDuet-H3 vector that contained a *N*-terminally 6×His-fusing H3 gene with an amber mutation introduced at a designated lysine coding site and grew the transformed cells in 2YT media supplemented with 1 mM AzHeK. We successfully expressed all eight proteins and affinity-purified them using Ni-NTA resins (**Figure 41a**). Electrospray ionization mass spectrometry (ESI-MS) analysis of purified proteins confirmed the incorporation of AzHeK (**Figures 37**). Using these recombinantly produced acyl-H3 proteins together with H2A, H2B, H4, and 601 DNA,[280] we succeeded in reconstituting corresponding acyl-nucleosomes, namely as nu-H3K4az, etc. that we visualized using ethidium bromide (EtBr) staining in a polyacrylamide gel (**Figure 41b**). We then subjected these *in vitro* assembled acyl-nucleosomes to deacylation by SIRT7. Acyl-nucleosomes treated with and without SIRT7 were then labeled with MB488-DBCO and imaged in 5% 1X TBE Native-PAGE gels. As shown in **Figure 41c**, SIRT7 removed acylation from H3 K18 and H3 K36 but was not active toward other lysine sites. The deacylation activity of SIRT7 on H3 K36 is clearly much higher than that on H3 K18. The fluorescent gel showed total removal of acylation at H3 K36 but residual acylation on H3 K18. H3 K18 was a previously demonstrated lysine deacylation site for SIRT7. However, SIRT7-catalyzed deacylation activity on H3 K36 came as a surprise. We went further to quantify deacylation activities of SIRT7 on all tested sites. In the given deacylation conditions in which a 1 uM acyl-nucleosome was incubated with 0.5 uM SIRT7 and 1 mM NAD^+ at 37 °C for 2 h, the deacylation activity of SIRT7 on H3 K36 was about 4 times of that on H3 K18 based on calculated deacylation percentages (**Figure 41d**).

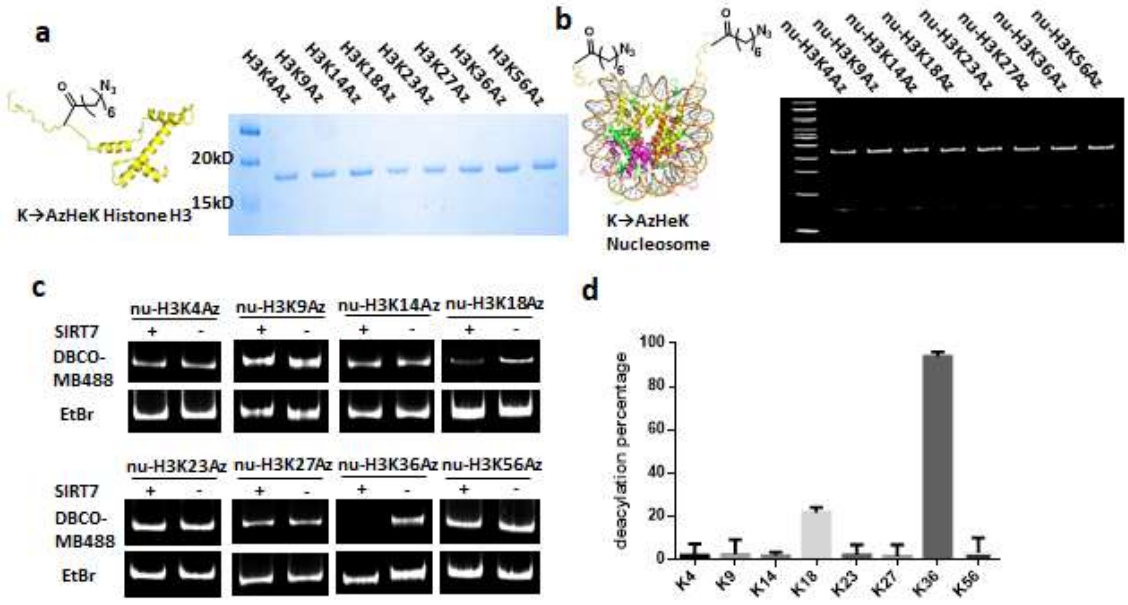


Figure 41 SIRT7 actively removes acylation from H3 K18 and H3 K36 in the nucleosome context. (**a**) SDS-PAGE analysis of eight purified AzHeK-containing H3 proteins. (**b**) Acyl-nucleosomes that we assembled from eight acyl-H3 proteins, H2A, H2B, H4, and 601 DNA (147 bp). Nu-H3K4az denotes an acyl-nucleosome assembled from H3K4az. All other acyl-nucleosomes are named in the same way. Mononucleosomes are typically shown around the 500 bp DNA position in an EtBr-stained 5% 1× TBE native-PAGE gel. (**c**) SIRT7 catalyzed deacylation activities on eight acyl-nucleosome substrates. Reaction conditions: we incubated 1 μM acyl-nucleosome with 0.5 μM SIRT7, 0.5 mM BME and 1 mM NAD+ at 37 °C for 2 h before we quenched the reaction by the addition of 20 mM nicotinamide. We then labeled the resulted acyl-nucleosome substrate with 100 μM MB488-DBCO for 1 h and analyzed it fluorescently in an 8% 1× TBE native-PAGE gel. The top panel shows the MB488-DBCO-based fluorescent imaging and the bottom panel shows the EtBr-stained DNA from the same gel. (**d**) Quantified deacylation levels at eight H3 lysine sites. We repeated experiments shown in C thrice and calculated the deacylation level at each site by subtracting average MB488-DBCO-based fluorescence intensity of SIRT7-treated samples from that of untreated controls.

DNA and salt inhibit SIRT7-catalyzed nucleosome deacylation. Since a previous study reported that DNA activates SIRT7-catalyzed deacetylation on peptide

substrates,[268] we investigated whether it has the same effect on SIRT7-catalyzed nucleosome deacylation. We used nu-H3K36az as a SIRT7 substrate and tested its deacylation by SIRT7 when we provided different concentrations of free 147 bp 601 DNA in the reaction solution. As shown in **Figure 42a**, SIRT7 strongly removed acylation from nu-H3K36az but this activity decreased significantly when we provided free DNA to the reaction solution and was close to be abolished when 2 μM free DNA was present. Thus, DNA apparently inhibits SIRT7-catalyzed nucleosome deacylation. Free DNA is negatively charged. It interacts electrostatically with both a positively charged histone-based peptide substrate and SIRT7 that has positively charged *N*- and *C*-termini. Therefore, it potentially serves as a bridge to bring SIRT7 and a peptide substrate in close proximity for improved deacetylation. We suspected that the wrapping DNA in an acyl-nucleosome substrate might play a similar role to interact electrostatically with SIRT7 for improved deacylation. To test this prospect, we examined SIRT7-catalyzed deacylation of nu-H3K36az in different concentrations of the NaCl salt. If interactions between SIRT7 and nu-H3K36az are electrostatic in nature, increasing ionic strength of the reaction solution by providing NaCl will disrupt the binding of SIRT7 to nu-H3K36az and therefore leads to a lower deacylation activity. As expected, the SIRT7-catalyzed deacylation activity on the nu-H3K36az substrate diminished gradually when more salt was present in the reaction solution (**Figure 42b**). To confirm that free DNA does serve as a bridge to bring SIRT7 and a peptide or protein substrate into close proximity for improved deacylation activity, we tested the SIRT7-catalyzed deacylation activity on a soluble H3K36az-H4 tetramer substrate in the

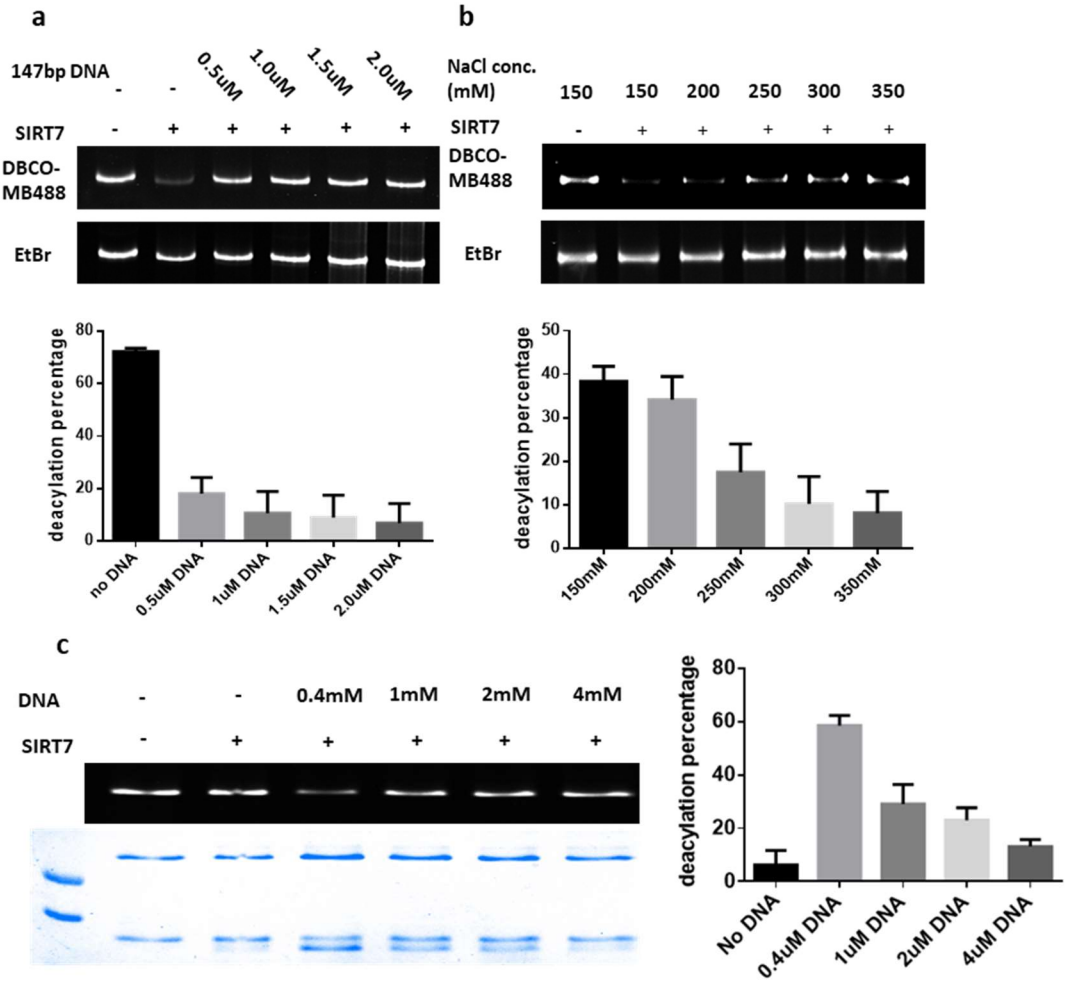


Figure 42 Effects of DNA and salt on SIRT7-catalyzed deacylation. (**a**) Free DNA inhibits SIRT7-catalyzed nucleosome deacylation. Reaction conditions: we incubated 1 μM nu-H3K36az with 0.1 μM SIRT7, a varied concentration of 147 bp DNA, 1 mM NAD+ and 0.5 mM BME at 37 °C for 3 h before we labeled the solution with 100 μM MB488-DBCO and analyzed the acyl-nucleosome substrate fluorescently in a 1× TBE native-PAGE gel. The top panel is MB488-DBCO-based imaging that indicates relative acylation levels and the bottom panel is EtBr-based imaging that confirms the nucleosome integrity. (**b**) Salt inhibits SIRT7-catalyzed nucleosome deacylation. Conditions were as same as in A except we replaced free DNA with NaCl. (**c**) Free DNA has a binary role on SIRT7-catalyzed deacylation of the H3K36az-H4 tetramer substrate. Reaction conditions: we incubated 2 μM H3K36az-H4 tetramer with 1 μM SIRT7, a varied concentration of 147 bp DNA, 0.5 mM BME, and 1 mM NAD^+ 37 °C for 2 h before we labeled the solution with 100 μM MB488-DBCO and analyzed the tetramer substrate fluorescently in a SDS-PAGE gel.

presence of various concentrations of free DNA. The result is shown in **Figure 42c**. Similar to its low activity toward H3K36ac, SIRT7 displayed a marginal deacylation activity on the H3K36az-H4 tetramer when free DNA was absent. However, this activity was significantly enhanced when we added 0.4 μM DNA into the reaction solution. When we added more DNA, this activity diminished gradually until it was almost totally abolished. This phenomenon is similar to what we observed with the nu-H3K36az substrate.

SIRT7 is more active toward acyl-nucleosomes with linker DNAs. In chromatin, a bridging linker DNA that binds loosely to H1 and transcription factors connects two adjoining nucleosomes. To determine whether this linker DNA contributes to SIRT7 binding to an acyl-nucleosome for active deacylation, we synthesized nu-H3K18az and nu-H3K36az nucleosomes with different lengths of linker DNAs and used them as substrates to test SIRT7-catalyzed deacylation (**Figures 43a** & **b**). Appending a linker DNA to nu-H3K36az significantly improved SIRT7-catalyzed deacylation (**Figure 43c**). A 5 bp linker DNA led to more than 3-fold deacylation activity increase (**Figure 43e**). Increasing the linker DNA length from 5 bp to 20 bp slightly enhanced SIRT7 deacylation activity at H3 K36 but not to a significant level. Although appending the nu-H3K18az substrate with a linker DNA also improved SIRT7-catalyzed deacylation at H3 K18, this activity increase was much less significant than what we observed for the nu-H3K36ac substrate (**Figures 43d** & **f**). On the contrary, appending a linker DNA to nu-H3K4az, nu-H3K9az, nu-H3K14az, H3K23az, and H3K27az did not improve their deacylation by SIRT7 (**Figure 44**).

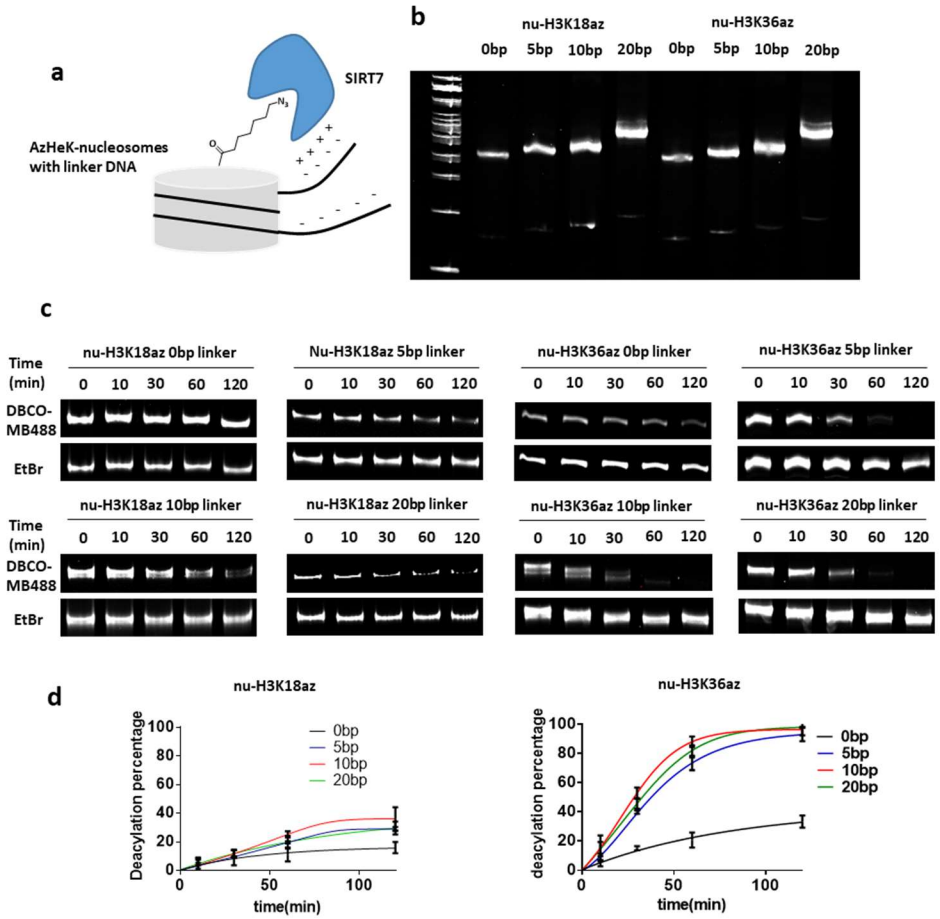


Figure 43 A linker DNA on an acyl-nucleosome substrate improves SIRT7-catalyzed deacylation. (**a**) A diagram to illustrate how a linker DNA affiliates the binding of SIRT7 to an acyl-nucleosome substrate. (**b**) The in vitro assembly of nu-H3K18az and nu-H3K36az with different lengths of linker DNAs. The gel showed the EtBr-stained nucleosomes. (**c**) The catalytic removal of acylation from nu-H3K36az substrates that contained different lengths of linker DNAs. Reaction conditions: we incubated a 1 μM H3K36az nucleosome with 0.1 μM SIRT7, 1 mM NAD and 0.5 mM BME at 37 °C for various times before we quenched the reaction by adding 20 mM nicotinamide, labeled the solution with 100 μM MB488-DBCO for 1 h, and then analyzed the nucleosome substrate fluorescently by 8% 1× TBE native-PAGE gel. (**d**) The catalytic removal of acylation from nu-H3K18az substrates that contained different lengths of linker DNAs. Reaction conditions were as same as for nu-H3K36az substrates except higher amount of SIRT7 (0.3 μM) was used. (**e-f**) Quantified deacylation percentage vs time for two acyl-nucleosome substrates. Reactions shown in c and d were repeated thrice. Deacylation was calculated and averaged by subtracting MB488-DBCO-based fluorescence intensity at different time points from that at 0 min.

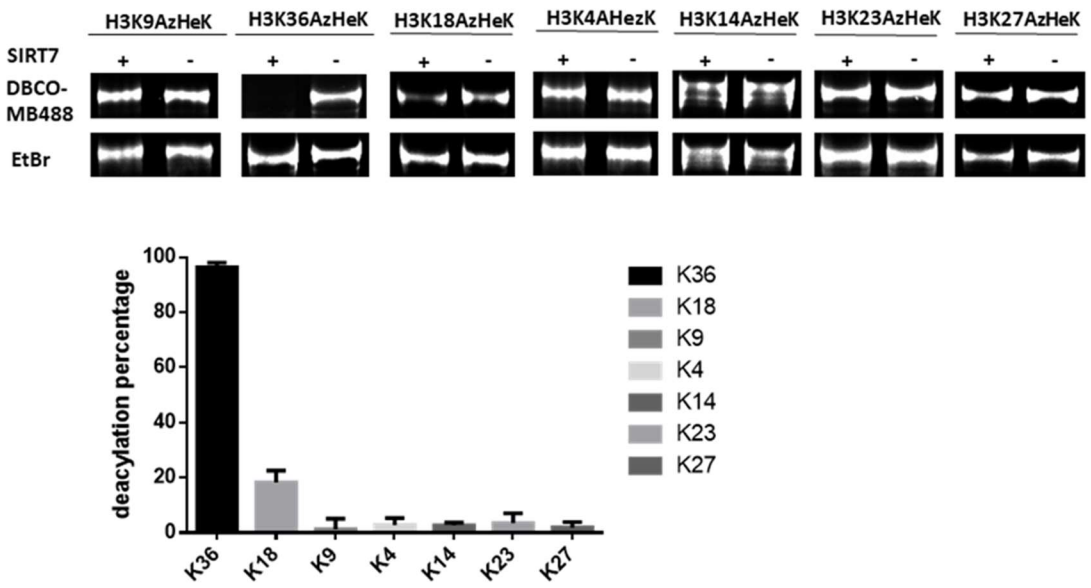


Figure 44 SIRT7 site-specificity test on Az-nucleosomes with 20bp linker DNA. Reaction condition: 1 uM nucleosome, 0.2 uM SIRT7, 1 mM NAD+, 0.5 mM β-ME, 37°C, 2h.

SIRT7 catalyzes removal of nucleosomal H3 K36 acetylation in vitro and in cells.

Our finding that H3 K36 is a preferential SIRT7-targeted lysine site for defatty-acylation led us to ask if SIRT7 is also catalytically active in removing acetylation at H3 K36 both in *in vitro* biochemical assays and in cells. We assembled two *in vitro* nu-H3K36ac nucleosomes using 147 bp 601 DNA with or without a 20 bp linker DNA (**Figure 46**). We then incubated these two nucleosomes with SIRT7 in deacetylation reactions and subsequently detected acetylation levels of the reaction products by Western blotting with an anti-H3K36ac antibody. The result revealed that SIRT7 deacetylated H3 K36 in the context of nucleosomes, and appending a linker DNA significantly improved this deacetylation activity (**Figure 45a**). To test SIRT7 activity on other nucleosomal H3

lysine, sites, we also reconstituted several other acetyl-nucleosomes including nu-H3K9ac, nu-H3K14ac, nu-H3K18ac, nu-H3K27ac, and nu-H3K56ac (**Figure 46**) for use as substrates in SIRT7 deacetylation assays. All acetyl-nucleosomes contained a 20 bp linker DNA for improving SIRT7 activities. Our data confirmed that SIRT7 is catalytically active in deacetylation of H3 K18 but is inert toward all other lysine acetylation sites tested, and the deacetylation activity of SIRT7 on H3 K18 is much weaker than that on H3 K36. To examine the deacetylase activity of SIRT7 on H3 K18 and H3 K36 in cells, we constructed two plasmids, one encoding a SIRT7-EGFP fusion gene and the other a SIRT7-EGFP fusion gene with a H187Y mutation to inactivate the SIRT7 activity, from a mammalian pEGFP vector. Together with the original pEGFP vector, we transiently transfected 293T cells separately with all three plasmids. We separated cell lysates from these three transiently transfected cells using SDS-PAGE and then analyzed different histone acetylation levels using antibodies that recognize acetylation at K9, K14, K23, K27, and K36 of H3, respectively. These assays revealed significantly decreased acetylation levels at H3 K18 and H3 K36 in cells with overexpressed SIRT7-EGFP but not in cells that transfected with the empty control vector or expressing the inactive SIRT7-H187Y-EGFP fusion protein. These data demonstrate that SIRT7 is catalytically active in deacetylating both H3 K18 and H3 K36 in cells (**Figure 45b**).

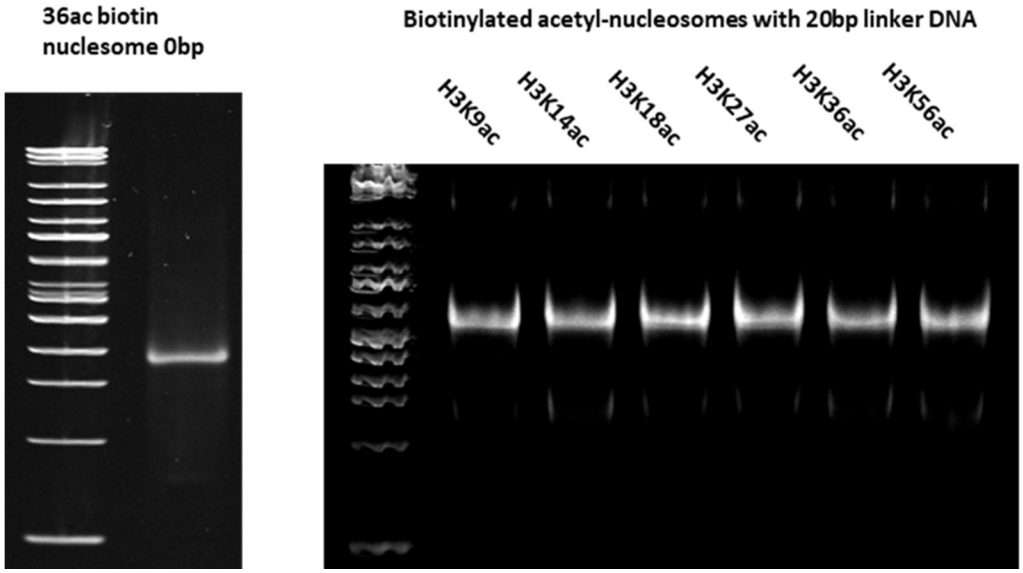


Figure 45 Assembly of H3 ac-nucleosomes

SIRT7 is essential for maintaining low H3K36ac levels at nucleolar rDNA sequences and select gene promoters. We next asked if H3K36Ac is a physiologic substrate of SIRT7, and if endogenous SIRT7 is important for setting baseline levels of H3 K36 acetylation. We generated SIRT7-deficient cells by shRNA-mediated knockdown strategies, and assayed levels of H3K36Ac in whole cell extracts by Western analysis **(Figure 47a)**. We observed significantly increased H3K36 acetylation in the SIRT7-depleted cells. We also examined locus-specific H3K36ac levels at promoters of previously described genomic targets of SIRT7 by chromatin immunoprecipitation (ChIP).[256] In SIRT7-depleted cells, levels of H3K36ac were significantly increased at two ribosomal protein gene targets of SIRT7, but not at another SIRT7 target, COPS2 **(Figure 47b)**. Importantly, we confirmed these findings using an independent H3K36ac antibody, since antibody reagents for H3K36ac ChIP have not been extensively

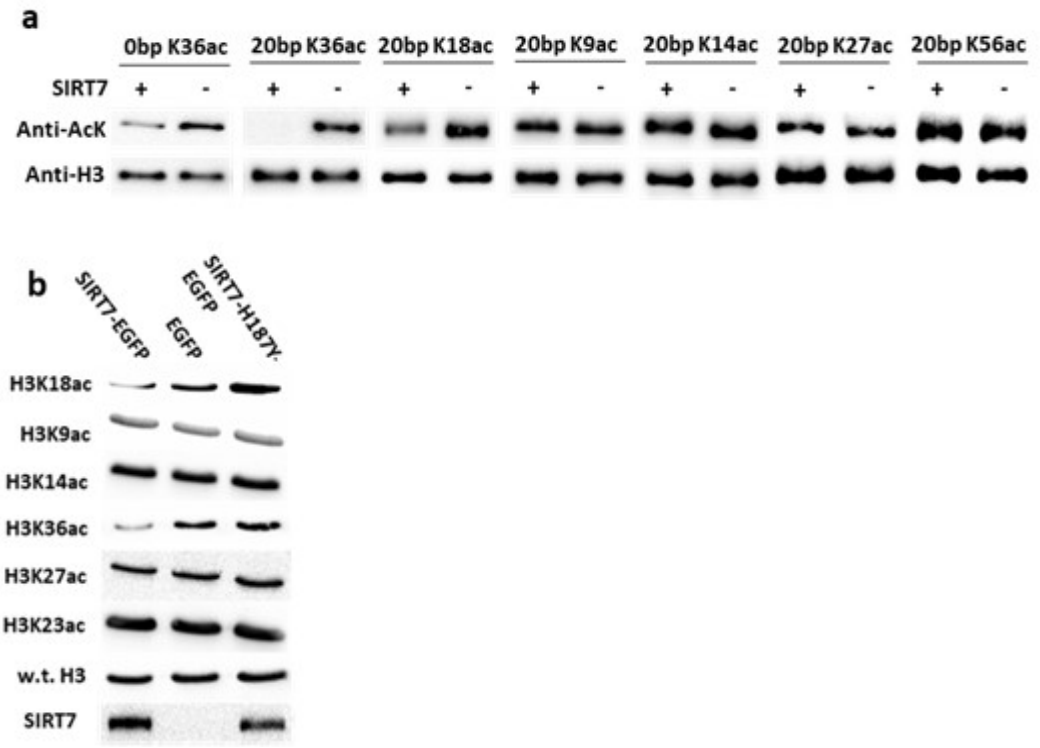


Figure 46 Validation of H3K36ac as SIRT7 targets both in vivo and in vitro. (a) SIRT7 deacetylation assay on acetyl-nucleosome in the following condition: 1 uM nucleosome, 0.2 uM SIRT7, 1 mM NAD+, 0.5 mM β-me, 37 ℃, 2 hours. Western blot bands on the upper panel were blotted by site-specific acetyl-histone antibodies, bands on the bottom panel were blotted by wildtype H3 antibodies. (b) global histone H3 acetylation level change after SIRT7 overexpression. EGFP, SIRT7-EGFP and inactive SIRT7-H187Y-EGFP mutant were transiently overexpressed in HEK293t cell, histones were extracted and acetylation levels were estimated by western blot.

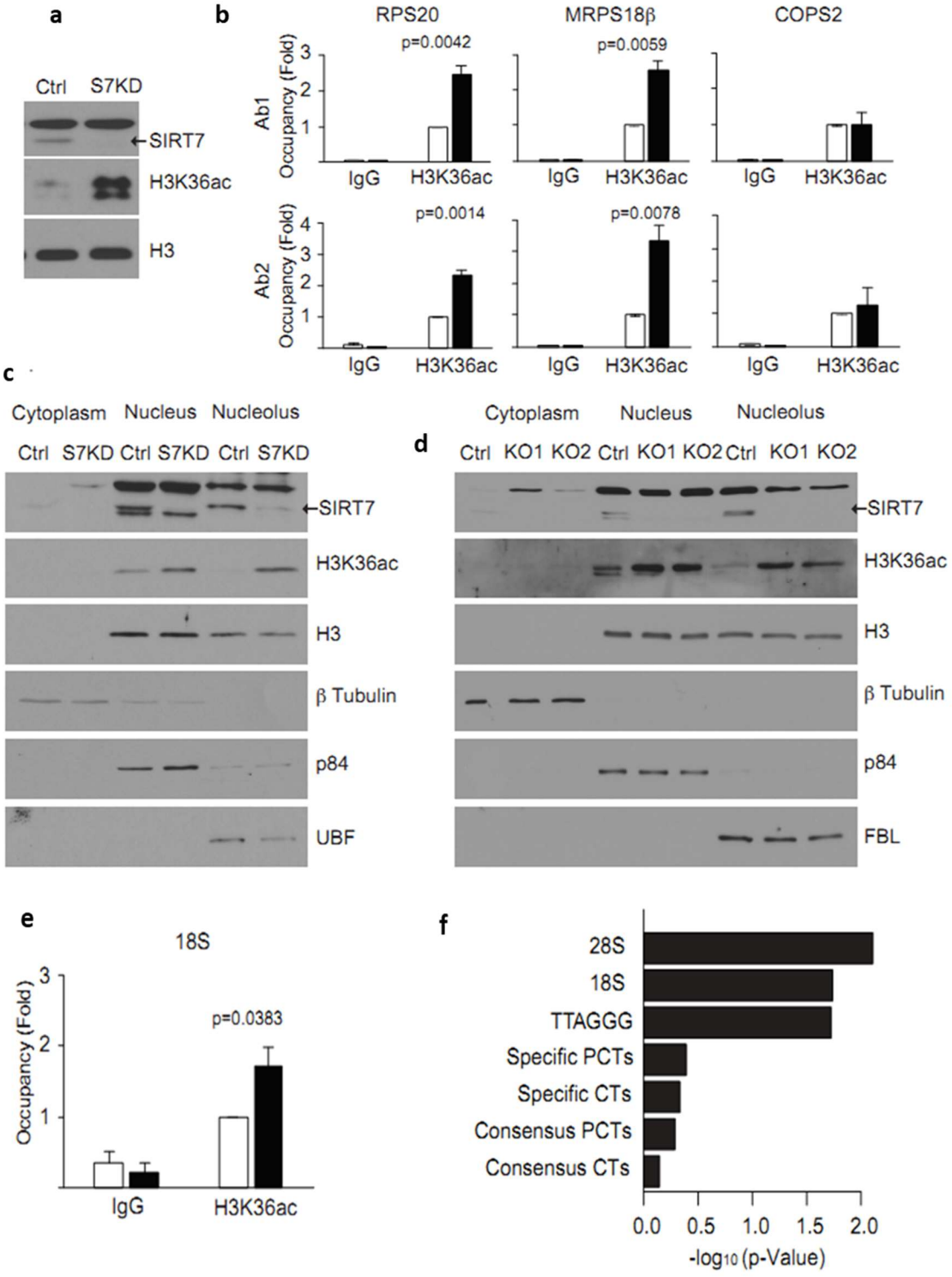
characterized **(Figure 47b)**. SIRT7 is reported to be enriched in nucleoli based on immunofluorescence studies. To examine changes in H3 K36 acetylation specifically in nucleoli, we carried out biochemical fractionation of SIRT7-deficient and control cell extracts into nucleolar and nucleoplasmic (non-nucleolar) fractions. Immunoblot analysis revealed that both nucleolar and nucleoplasmic levels of H3K36ac are dramatically increased in SIRT7-depleted cells compared to control cells **(Figure 47c**).

To further validate these findings in cells completely lacking SIRT7, we used CRISPR/Cas9 targeting to generate two independent SIRT7-knockout cell lines and a non-targeted control-treated cell line. Nucleolar fractionation of the cells confirmed the dramatic increase in H3K36ac levels in both nucleoli and nucleoplasm of SIRT7-knockout cells (**Figure 47d**). Importantly, in the control cells, H3K36ac levels (relative to total H3) were much lower in nucleoli than nucleoplasm, and this inversely correlated with SIRT7 levels in those nuclear compartments. This further underscores the physiological relevance of SIRT7 deacetylase activity on H3K36ac levels in a non-perturbed system. The main DNA components of nucleoli are rDNA genes. To directly assess H3K36 acetylation changes at rDNA, we analyzed H3K36ac ChIPs from SIRT7-deficient and control cells using real-time PCR or unbiased high-throughput sequencing The ChIP-PCR revealed significantly increased H3K36ac occupancy at rDNA sequences (**Figure 47e**). Notably, analysis of the high-throughput sequencing reads from the ChIPs confirmed these findings, and showed that the increase in H3K36ac is quite specific to rDNA compared to other repetitive regions of the genome, as no significant changes in acetylation were observed at repetitive sequence tracts present at over 20 different centromeric and pericentromeric regions, or when queried for eight different centromeric and pericentromeric consensus repeat sequences (**Figure 47f**).[281] Finally, this analysis also revealed an intriguing and unexpected increase in H3K36ac occupancy at TTAGGG repeats, which are present at telomeres. Together, our studies demonstrate that SIRT7 is a novel physiologic H3K36 deacetylase, and is essential for maintaining low H3K36

acetylation levels at rDNA sequences in nucleoli, certain protein-coding genes, and possibly telomeric DNA.

H3 K36 acylation leads to a loose nucleosome complex. H3 K36 is located at the entry and exit sites of DNA in the nucleosome dyad. Its positively charged side chain amine can interact directly with a negatively charged DNA backbone phosphate. This potential electrostatic interaction is expected to contribute to a tightly bound nucleosome complex that prevents the access of H3 K36 by enzymes and transcription factors including SIRT7. The disruption of this electrostatic interaction by H3 K36 acylation will weaken the binding of DNA to the nucleosome core, therefore leading to more frequent unwrapping of DNA from core histones and consequently more accessibility to enzymes and transcription factors. We think this potentially increased unwrapping of DNA from the nucleosome core assists SIRT7-catalyzed deacetylation at H3 K36. This improved unwrapping of DNA from the nucleosome core will also lead to more access of DNA by a nuclease (**Figure 49a**). To test whether H3 K36 acetylation improves DNA unwrapping, we introduced a Pst1 endonuclease restriction site at the entry/exit site of 601 DNA that also contained a 20 bp linker DNA beyond the Pst1 restitution site and used this mutant 601 DNA to reconstitute three nucleosomes including wild type (nu') with no histone modification, nu'-H3K18ac, and nu'-H3K36ac for analyzing effects of acetylation of H3 K18 and H3 K36 on Pst1-catalyzed restriction digestion of the wrapping DNA (**Figure 48** and **Figure 49b**). We introduced the Pst1-cutting site at the entry/exit site in order to maintain the nucleosome integrity after Pst1- catalyzed restriction cleavage of the tail linker DNA fragment. Since the digested nucleosome has

Figure 47 SIRT7 is a physiologic nucleolar and nucleoplasmic H3K36 deacetylase. (**a**) Immunoblot showing increased H3K36ac levels in whole cell lysates from SIRT7 knock-down (KD) U2OS cells compared to control cells. (**b**) ChIP-PCR showing increased H3K36ac levels at promoters of SIRT7 target genes RPS20 and MRPS18b, but not COPS2, in SIRT7 KD and control U2OS cells. Two independent H3K36ac antibodies (Ab 1 and 2) were used. Data show the average of three technical replicates +/- SEM, and are representative of 2-4 biological replicates. Statistical significance was calculated performing Student's t-test. (**c**) Immunoblot of nucleolar fractionation showing increased levels of H3K36ac in SIRT7 KD cells versus control cells. β-Tubulin, p84, and FBL or UBF are shown as marker controls for cytoplasmic, nucleoplasmic, and nucleolar fractions, respectively. (**d**) Immunoblot of nucleolar fractionation of SIRT7 knockout versus control U2OS cells. (**e**) ChIP-PCR showing increased levels of H3K36ac at 18S rDNA sequences in SIRT7 KD cells. Data show the average of three biological replicates +/- SEM. Statistical significance was calculated performing Student's t-test. (**f**) ChIP-seq enrichment analysis of H3K36ac at repetitive sequences in SIRT7 KD versus control U2OS cells. Bar plot of $\log_{10}$(p-value) shows statistically significant enrichment of H3K36ac in SIRT7-KD cells at rDNA (18S, 28S) and telomeric (TTAGGG) repeats, but not at repetitive sequences from 20 specific centromeres (CT), 6 specific pericentromeres (PCT), or 4 centromeric and 4 pericentric consensus sequences. P-values were calculated performing Chi-squared test.


a
Ctrl
S7KD
SIRT7
H3K36ac
H3
b
RPS20
MRPS18β
COPS2
p=0.0042
p=0.0059
p=0.0014
p=0.0078
Ab1
Occupancy (Fold)
Ab2
IgG
H3K36ac
c
Cytoplasm
Nucleus
Nucleolus
Ctrl
S7KD
SIRT7
H3K36ac
H3
β Tubulin
p84
UBF
d
KO1
KO2
FBL
e
18S
p=0.0383
f
28S
18S
TTAGGG
Specific PCTs
Specific CTs
Consensus PCTs
Consensus CTs
-log10 (p-Value)

a shorter DNA, it will run faster than the original nucleosome and can be easily visualized in a 5% 1× TBE native-PAGE gel after EtBr staining. We included nu'-H3K18ac in our analysis to see whether acetylation may generally increase DNA unwrapping from the nucleosome histone core. As shown in **Figure 49c**, Pst1 showed much higher activities to digest DNA in nu'-H3K18ac and nu'-H3K36ac than DNA in nu' that had no histone lysine acetylation. The effect of H3 K36 acetylation is more significant than that of H3 K18 acetylation (**Figure 49d**).

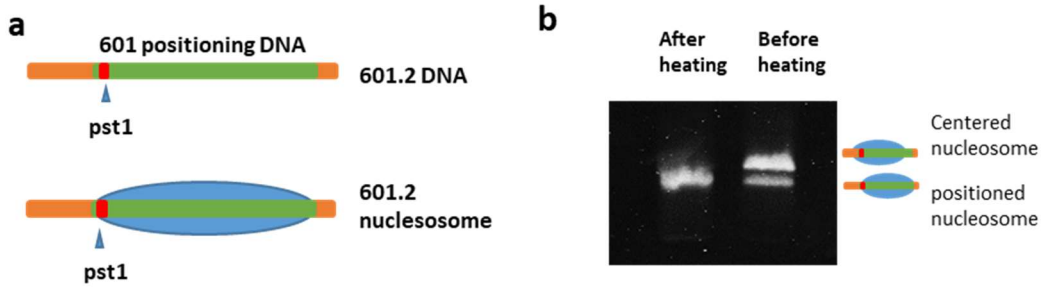


Figure 48 (a) Design of modified 601 DNA for the accessibility assay. (b) Thermal repositioning generated homogenous nucleosome species. Condition: 60 ℃, 30 min.

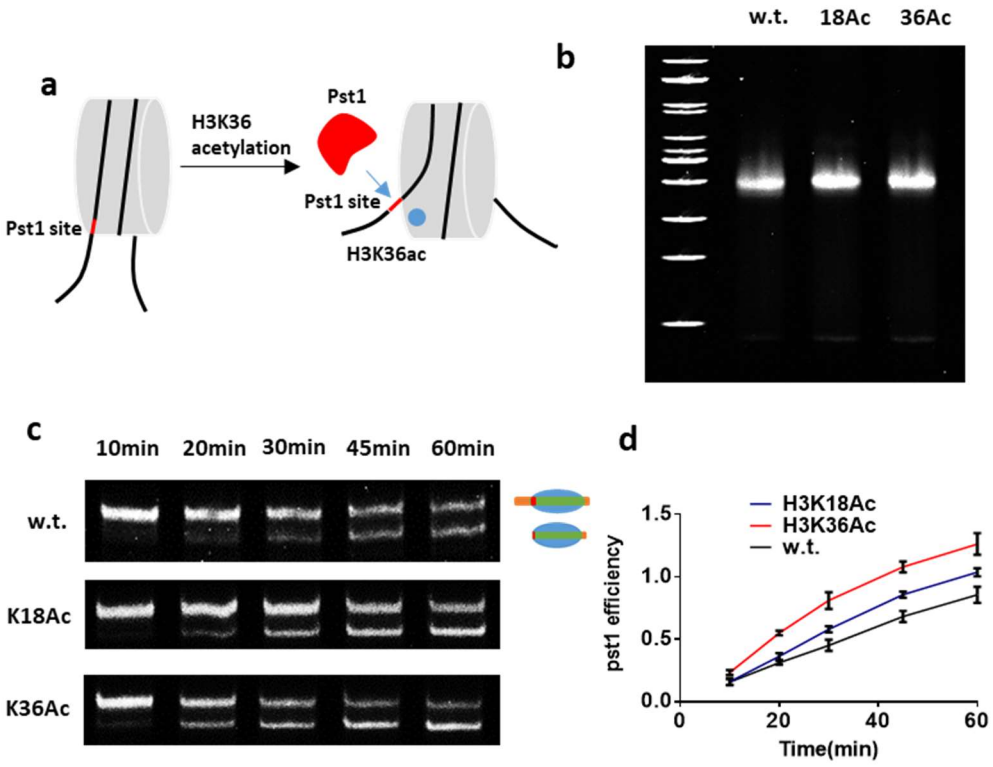


Figure 49 Lysine acetylation at H3 K36 improves DNA unwrapping from the nucleosome core. (**a**) A diagram to illustrate how H3K36ac increases DNA unwrapping from the nucleosome core and subsequent Pst1 access to an embedded Pst1 restriction site. (**b**) The *in vitro* assembly of three nucleosomes, nu', nu'-H3K18ac, and nu'-H3K36ac from corresponding histones and a mutant 601 DNA that had a Pst1 restriction site and a linker DNA. The gel shows EtBr-stained nucleosomes. (**c**) Pst1 digestion progress of three nucleosomes versus time. Reaction conditions: we incubated a 0.49 μM nucleosome with 2 U/uL Pst1 at 37 °C for various times before the reaction was stopped and the solution was analyzed by 5% 1× TBE native-PAGE. The gel was stained by EtBr. The top band in a lane shows the nucleosome with the linker DNA and the bottom band represents the one with the linker DNA removed. (**d**) The calculated ratio of digested nucleosome to undigested nucleosome versus time. We calculated and averaged the ratios based on integrated fluorescence intensities of original and digested nucleosome bands in gels.

Conclusion

For kinetic analysis of histone modifying enzymes, peptide-based substrates are typically used, which accounts for straightforward characterization of both substrates and products using HPLC, mass spectrometry, and other methods. However, these analytical methods are difficult to apply to histone and nucleosome-based substrates. Although Western blot can serve as an alternative analytical approach, accurate quantification based on antibody detection is relatively difficult. Western blot also involves multiple steps that are tedious. Inspired by the finding that some sirtuins are general deacylases that catalytically remove a number of acylation types, we previously demonstrated that a long chain terminal olefin-containing fatty acylation on a nucleosomal lysine in combination with an olefin-tetrazine cycloaddition reaction to fluorescently label a fatty acyl-nucleosome allowed characterization of SIRT6-targeted histone deacylation in the nucleosome context.[276] However, the reaction used in the SIRT6 study was kinetically slow, which required a long labeling time. To resolve this problem, we demonstrated the genetic encoding of AzHeK in the current study for the synthesis of azide-containing fatty acyl-nucleosomes that can be quickly labeled with a DBCO-based dye but still serve as substrates for SIRT7. Using this new click chemistry-based approach, we expediently characterized SIRT7-catalyzed deacylation on multiple acyl-nucleosome substrates. Given that SIRT1, SIRT2, SIRT6, and SIRT7 share similar catalytic features[199] and some Zn^{2+}-dependent HDACs are also active to remove long chain fatty acylation from histone lysines,[282, 283] this newly developed method can potentially serve as an efficient tool to study all these enzymes.

SIRT7 is a demonstrated histone H3 K18 deacetylase in cells. However, it doesn't display appreciable activity when assayed on H3K18ac as a free histone H3 protein. Our studies with 8 other acetyl-H3 histone substrates showed exactly same results. Apparently, SIRT7 is inactive toward a stand-alone acetyl-histone substrate. SIRT7 has an isoelectric point (pI) as 9.8. This high pI makes it highly positively charged at physiological pH. The calculated positive net charges determined by Protein Calculator v3.4 at pH 7.4 for SIRT7 are about 24. Histone H3 naturally has many lysine and arginine residues, which result in also a high pI value as 11.1. At pH 7.4, an acetyl-H3 will have positive net charges about 19. Since both SIRT7 and an acetyl-H3 substrate are highly positively charged at physiological pH, they will naturally have strong electrostatic repulsion that prevents their direct binding for deacetylation. This explains our *in vitro* data and also supports the notion that SIRT7 doesn't natively deacetylate free histones in cells. We have also made a similar observation with SIRT6, a close homolog of SIRT7. Although SIRT6 actively removes acetylation from H3 K9 in cells,[202] it is inactive toward an H3K9ac substrate. Like SIRT7, SIRT6 has a high pI value as 9.3 that makes it highly positively charged at physiological pH and consequently repel the binding of an acetyl-H3 substrate. Although SIRT6 and SIRT7 naturally repel acetyl-histone substrates, their highly positively charged nature makes them easily associate electrostatically with negatively charged DNA in chromatin. As a matter of fact, both enzymes have been characterized as chromatin binding enzymes. In a chromatin setup, not only does chromosomal DNA neutralize positive charges in both SIRT6 or SIRT7 and an acetyl-histone substrate for relieving their electrostatic repulsion, but it also serves as a mediator

to bring SIRT6 or SIRT7 and the acetyl-histone substrate into close proximity for deacetylation. As shown here for SIRT7 and previously for SIRT6, putting H3K18ac and H3K9ac in a nucleosome context brought up their significant deacetylation by SIRT7 and SIRT6, respectively. Therefore, both enzymes are nucleosome-dependent deacylases. DNA in chromatin may also serves to activate SIRT7. A previous report showed that free DNA activates SIRT7-catalyzed deacetylation of an acetyl-peptide substrate.[268] We re-evaluated this potential DNA activating effect using both an acyl-nucleosome substrate and an acyl-H3-H4 tetramer substrate. Providing free DNA clearly inhibited SIRT7-catalyzed deacylation on the acyl-nucleosome substrate. However, adding a small amount of DNA to the acyl-H3/H4 tetramer substrate significantly improved its deacylation by SIRT7 at the beginning but this improvement was offset when additional DNA was added. These results support our conclusion that DNA serves only as a mediator to bring SIRT7 and an acyl-histone substrate into close proximity for deacylation. Free DNA competes against an acyl-nucleosome substrate to bind SIRT7 and therefore inhibits SIRT7-catalyzed deacylation of the acyl-nucleosome substrate. However, when DNA is provided in a small amount to an acyl-H3-H4 tetramer substrate, a single DNA fragment is able to bind both SIRT7 and the acyl-H3-H4 tetramer substrate to form a 1:1:1 complex and therefore serves to bring them to close proximity for interactions and deacylation. When provided in a large amount, additional DNA competes against the formation of a DNA-SIRT7-substrate (1:1:1) complex that is required for deacylation and consequently brings down the SIRT7-catalyzed deacylation of the acyl-H3-H4 tetramer substrate. The interactions between DNA and SIRT7 are largely electrostatic in nature. When we

provided NaCl that disrupts electrostatic interactions, the activity of SIRT7 on an acyl-nucleosome substrate was clearly inhibited. We further confirmed the electrostatic nature of DNA-SIRT7 interactions by appending an acyl-nucleosome substrate with two additional linker DNAs to show an improve deacylation activity by SIRT7. These linker DNAs do not directly participate in the nucleosome complex formation. Their negative charges are less neutralized like the wrapping DNA. With more net negative charges, they were expected to interact more efficiently with SIRT7 than the wrapping DNA for improving SIRT7 binding and recognition of the integrated acyl-histone substrate in the nucleosome context for deacylation. This was exactly what the results showed.

Our current SIRT7 study also points out a critical aspect of epigenetic enzymes that might have been long ignored in their biochemical analysis. Most histone modifying enzymes naturally interact directly or indirectly with chromatin in which nucleosome is the smallest subunit. It is logic to consider that corresponding nucleosome substrates might be minimally required to study these enzymes for resembling their catalytic features in cells. However, for convenient reasons, a lot of biochemical studies that have been carried out for histone modifying enzymes are based on peptide substrates and some are based on histones. How much an activity of a histone modifying enzyme on a peptide or histone substrate deviates from that of a nucleosome or chromatin substrate is a significant concern. Many histone modifying enzymes such as GCN5, MLL1-2, Suv39h1-2, Suv420h1-2, Dot1L, and KDM4D share the highly basic feature similar to SIRT6 and SIRT7, their strong interactions with chromatin for improved activities on histones are expected. There are also a large number of histone modifying enzymes that have strong

acidic features. Most Zn^{2+}-dependent HDACs have relatively low pI values. At physiological pH, they are highly negatively charged and thus expected to bind tightly to histones but resist the association with chromatin. Although peptide substrate-based data have indicated strong deacetylation activities for most Zn^{2+}-dependent HDACs,[284] it is doubtful that these enzymes behave the same on nucleosome or chromatin substrates. Many Zn^{2+}-dependent HDACs are part of transcription factor complexes that associate with chromatin. It is highly possible they are brought into interactions with their histone substrates within chromatin by their associated partners in these complexes to avoid electrostatic repulsion with chromatin DNA. Supported by the SIRT7 study in this work and our previous study with SIRT6, we strongly recommend re-characterizing biochemistry of most histone modifying enzymes using related nucleosome substrates.

As for SIRT7-targetted H3 lysine sites, we were able to successfully confirm the previously discovered H3 K18 site but surprisingly revealed the H3 K36 site. SIRT7 displays a several fold more activity to deacylate H3 K36 than to deacylate H3 K18. We did not observe appreciate activities on other tested H3 lysine sites. H3 K18 and H3 K36 have both low peptide sequence resemblance and low chemical environment resemblance. How SIRT7 recognizes these two dissimilar lysine sites for deacylation is very intriguing. Interestingly, H3 K18 highly resembles H3 K9 and H3 K27 in peptide sequence contexts. All three lysines are in the H3 *N*-terminal tail with much lower structural constrain in comparison to H3 K36 that is at the DNA entry/exit site of the nucleosome dyad. Our previous study showed that SIRT6 was mainly active to deacylate these three H3 lysine sites probably due to their consensual sequence contexts.[276] It is quite puzzling that SIRT7,

a close homolog of SIRT6, deacylates H3 K18 but is totally inert toward H3 K9 and H3 K27. Given that H3 K36 is at the DNA entry/exit site of the nucleosome dyad, unwrapping DNA partially from the nucleosome core will be necessary to leave enough space for SIRT7 to bind and deacylate it. How SIRT7 unwraps DNA and then interacts with both DNA and H3 K36 is perplexing. Before the structure of SIRT7 complexed with an acyl-nucleosome substrate is finally determined, we hesitate to postulate a theory to explain our observation.

H3K36ac is an active transcription mark that accumulates at the transcription start sites.[231] It is enriched in euchromatin, which is possibly due to its potential weakening of the nucleosome complex.[285] Our *in vitro* assembled nu'-H3K36ac did show more DNA unwrapping than the nucleosome without any acetylation. When we overexpressed SIRT7 in cells, we did observe down-regulation of acetylation at both H3 K18 and H3 K36, confirming H3 K36 as a cellular deacetylation target of SIRT7. Importantly, we also showed that loss of SIRT7 in human cells leads to hyperacetylation of H3K36, which can be detected at the level of whole cell extracts, isolated nucleoli, and select SIRT7 target promoters. These findings demonstrate that endogenous SIRT7 has an essential function in determining the physiologic levels of this acetylation mark.

Our biochemical fractionation and ChIP studies also reveal the pivotal role of SIRT7 in maintaining low H3K36ac levels at rDNA in nucleoli. SIRT7 has two distinct functions at rDNA sequences, which exist in either active or silent chromatin states. At active rDNA, where SIRT7 was previously linked to rDNA transcriptional activation, SIRT7-dependent removal of H3K36ac might allow for maintenance of H3K36me3,

which is associated with transcription elongation.[257] By contrast, at silent rDNA regions, H3 K36 deacetylation by SIRT7 might have a novel function in contributing to SIRT7 effects on heterochromatin maintenance and prevention of rDNA instability.[258] Finally, SIRT7 is recruited to DNA DSBs where it contributes to DNA damage responses.[266] An intriguing possibility is that SIRT7-dependent deacetylation of H3 K36 at DSBs might clear the way for generation of H3K36me3, a critical mediator of DNA damage signaling.

CHAPTER IV

CONCLUDING REMARKS AND FUTURE OUTLOOK

In this study, we incorporated *N*$^{\varepsilon}$-(7-octenoyl)-lysine (OcK) and N$^{\varepsilon}$-(7-azidoheptanoyl)-L-lysine (AzHeK) into different lysine sites of Histone H3 (oc-H3), and assembled them into mono-nucleosome substrates. Since SIRT6 and SIRT7 have been reported as nucleosome-dependent Histone deacylases, we screened site-specificity of these two enzymes against nucleosomes with OcK or AzHeK at different lysine sites. Acylation level changes during SIRT6 or SIRT7 incubation were detected through the click reaction either between pyrimidine-tetrazine conjugated FITC and OcK, or between dibenzocyclooctyne conjugated MB488 and AzHeK.

For SIRT6, besides confirming the previously reported H3 K9 acylation, we discovered several other active sites. Overall, the amino acid in front of lysine appears to determine the deacylation activity. For sites with higher activity, H3 K18, H3 K27 and the previously reported H3 K9 all have arginine at the -1 position. H3 K4 and H3 K23 demonstrate relatively weaker activity and they have threonine at the -1 position. However, only H3 K9, H3 K18 and H3 K27 were confirmed as SIRT6 targets in HEK293t cell. For SIRT7, we confirmed the previously reported site H3 K18 and discovered another much more active site H3 K36, which was validated as the *in vivo* target in both SIRT7 overexpressed and SIRT7 knockdown mammalian cell models. SIRT7 was further demonstrated as loci-specifc H3 K36 deacetylase at gene promoters, nucleoli and telomeres. We also revealed that DNA acted as the negatively charged

'bridge' bringing positively charged histone and SIRT7 into proximity through electrostatic interactions, thus extranucleosomal DNA activated SIRT7 deacylation by acting as the additional negatively charged binding surface.

Histone PTM code is a complex system managing the thousands of genes. To understand the independent and combinatorial meaning of all PTM, one fundamental task would be revealing the accurate site-specificity of PTM editing and reading proteins. Our results indicated the necessity of characterizing epigenetic enzymes' site-specificity using the native nucleosome substrates, as many of these enzymes show distinct *in vitro* reactivity between peptide, histone and nucleosomes. Most histone PTM sites are blocked by the wrapping DNA, which requires either the enzymes to have intrinsic DNA binding affinity such as SIRT6 or SIRT7, or other co-activators such as CoREST to facilitate binding between enzymes and nucleosomes. Recently, many epigenetic enzymes were found to be regulated by the existing PTM on the substrate. And characterization of this PTM crosstalk effect is clearly expected to be only correct if nucleosomes are used as *in vitro* substrates, since the space-distribution of histone tails and allosteric effect of other nucleosome components such as DNA must be considered. The biggest obstacle for nucleosome biochemical study is the reconstitution of homogenous nucleosome substrates with site-specific PTM. With the rapid development of histone synthesis techniques such as non-canonical amino acid incorporation, expressed protein ligation and bio-orthogonal reaction, more precise details of epigenetic code will be revealed in the future.

REFERENCES

[1] Avery, O. T., Macleod, C. M., and McCarty, M. (1944) Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii, *J Exp Med 79*, 137-158.

[2] Waddington, C. H. (1957) *The Strategy of the Genes*, Routledge, London.

[3] Laskey, R. A., and Gurdon, J. B. (1970) Genetic content of adult somatic cells tested by nuclear transplantation from cultured cells, *Nature 228*, 1332-1334.

[4] Bird, A. (2007) Perceptions of epigenetics, *Nature 447*, 396-398.

[5] Stedman, E. (1950) Cell specificity of histones, *Nature 166*, 780-781.

[6] Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964) ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS, *Proceedings of the National Academy of Sciences of the United States of America 51*, 786-794.

[7] Kornberg, R. D., and Thomas, J. O. (1974) Chromatin structure; oligomers of the histones, *Science 184*, 865-868.

[8] Richmond, T. J., Finch, J. T., Rushton, B., Rhodes, D., and Klug, A. (1984) Structure of the nucleosome core particle at 7 Å resolution, *Nature 311*, 532.

[9] Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution, *Nature 389*, 251.

[10] Turner, B. M. (2000) Histone acetylation and an epigenetic code, *Bioessays 22*, 836-845.

[11] Thoma, F., Koller, T., and Klug, A. (1979) Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin, *J Cell Biol 83*, 403-427.

[12] Talbert, P. B., Ahmad, K., Almouzni, G., Ausio, J., Berger, F., Bhalla, P. L., Bonner, W. M., Cande, W. Z., Chadwick, B. P., Chan, S. W., Cross, G. A., Cui, L., Dimitrov, S. I., Doenecke, D., Eirin-Lopez, J. M., Gorovsky, M. A., Hake, S. B., Hamkalo, B. A., Holec, S., Jacobsen, S. E., Kamieniarz, K., Khochbin, S., Ladurner, A. G., Landsman, D., Latham, J. A., Loppin, B., Malik, H. S., Marzluff, W. F., Pehrson, J. R., Postberg, J., Schneider, R., Singh, M. B., Smith, M. M., Thompson, E., Torres-Padilla, M. E., Tremethick, D. J., Turner, B. M., Waterborg, J. H., Wollmann, H., Yelagandula, R., Zhu, B., and Henikoff, S. (2012) A unified phylogeny-based nomenclature for histone variants, *Epigenetics Chromatin 5*, 7.

[13] Saha, A., Wittmeyer, J., and Cairns, B. R. (2006) Chromatin remodelling: the industrial revolution of DNA around histones, *Nat Rev Mol Cell Biol 7*, 437-447.

[14] Miller, C. A., and Sweatt, J. D. (2007) Covalent modification of DNA regulates memory formation, *Neuron 53*, 857-869.

[15] Hake, S. B., Xiao, A., and Allis, C. D. (2004) Linking the epigenetic 'language' of covalent histone modifications to cancer, *Brit J Cancer 90*, 761-769.

[16] Rivera, C. M., and Ren, B. (2013) Mapping Human Epigenomes, *Cell 155*, 39-55.

[17] Kouzarides, T. (2007) Chromatin modifications and their function, *Cell 128*, 693-705.

[18] Huang, H., Lin, S., Garcia, B. A., and Zhao, Y. (2015) Quantitative proteomic analysis of histone modifications, *Chem Rev 115*, 2376-2418.

[19] Sabari, B. R., Zhang, D., Allis, C. D., and Zhao, Y. M. (2017) Metabolic regulation of gene expression through histone acylations, *Nat Rev Mol Cell Bio 18*, 90-101.

[20] Wang, X. D., and Hayes, J. J. (2008) Acetylation mimics within individual core histone tail domains indicate distinct roles in regulating the stability of higher-order chromatin structure, *Molecular and Cellular Biology 28*, 227-236.

[21] Hendzel, M. J., Lever, M. A., Crawford, E., and Th'ng, J. P. H. (2004) The C-terminal domain is the primary determinant of histone H1 binding to chromatin in vivo, *J Biol Chem 279*, 20028-20034.

[22] Ferreira, H., Somers, J., Webster, R., Flaus, A., and Owen-Hughes, T. (2007) Histone tails and the H3 alpha N helix regulate nucleosome mobility and stability, *Molecular and Cellular Biology 27*, 4037-4048.

[23] Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993) A Positive Role for Histone Acetylation in Transcription Factor Access to Nucleosomal DNA, *Cell 72*, 73-84.

[24] Megee, P. C., Morgan, B. A., Mittman, B. A., and Smith, M. M. (1990) Genetic-Analysis of Histone-H4 - Essential Role of Lysines Subject to Reversible Acetylation, *Science 247*, 841-845.

[25] Zhang, W. Z., Bone, J. R., Edmondson, D. G., Turner, B. M., and Roth, S. Y. (1998) Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase, *Embo J 17*, 3155-3167.

[26] Watanabe, S., Resch, M., Lilyestrom, W., Clark, N., Hansen, J. C., Peterson, C., and Luger, K. (2010) Structural characterization of H3K56Q nucleosomes and nucleosomal arrays, *Bba-Gene Regul Mech 1799*, 480-486.

[27] Manohar, M., Mooney, A. M., North, J. A., Nakkula, R. J., Picking, J. W., Edon, A., Fishel, R., Poirier, M. G., and Ottesen, J. J. (2009) Acetylation of Histone H3 at the Nucleosome Dyad Alters DNA-Histone Binding, *J Biol Chem 284*, 23312-23321.

[28] Huang, H. L., Maertens, A. M., Hyland, E. M., Dai, J. B. A., Norris, A., Boeke, J. D., and Bader, J. S. (2009) HistoneHits: A database for histone mutations and their phenotypes, *Genome Res 19*, 674-681.

[29] Clayton, A. L., Hazzalin, C. A., and Mahadevan, L. C. (2006) Enhanced histone acetylation and transcription: A dynamic perspective, *Molecular Cell 23*, 289-296.

[30] Delmore, J. E., Issa, G. C., Lemieux, M. E., Rahl, P. B., Shi, J. W., Jacobs, H. M., Kastritis, E., Gilpatrick, T., Paranal, R. M., Qi, J., Chesi, M., Schinzel, A. C., McKeown, M. R., Heffernan, T. P., Vakoc, C. R., Bergsagel, P. L., Ghobrial, I. M., Richardson, P. G., Young, R. A., Hahn, W. C., Anderson, K. C., Kung, A. L.,

Bradner, J. E., and Mitsiades, C. S. (2011) BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc, *Cell 146*, 903-916.

[31] Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999) Structure and ligand of a histone acetyltransferase bromodomain, *Nature 399*, 491-496.

[32] Robinson, P. J. J., An, W., Routh, A., Martino, F., Chapman, L., Roeder, R. G., and Rhodes, D. (2008) 30 nm chromatin fibre decompaction requires both H4-K16 acetylation and linker histone eviction, *Journal of Molecular Biology 381*, 816-825.

[33] Liokatis, S., Klingberg, R., Tan, S., and Schwarzer, D. (2016) Differentially Isotope-Labeled Nucleosomes To Study Asymmetric Histone Modification Crosstalk by Time-Resolved NMR Spectroscopy, *Angew Chem Int Edit 55*, 8262-8265.

[34] Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Aufiero, D. J., Kitabayashi, I., Herr, W., and Cleary, M. L. (2004) Leukemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression, *Molecular and Cellular Biology 24*, 5639-5649.

[35] Ikura, T., Ogryzko, V. V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., Scully, R., Qin, J., and Nakatani, Y. (2000) Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis, *Cell 102*, 463-473.

[36] Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., OwenHughes, T., Allis, C. D., Winston, F., Berger, S. L., and

Workman, J. L. (1997) Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: Characterization of an Ada complex and the SAGA (Spt/Ada) complex, *Gene Dev 11*, 1640-1650.

[37] Kuo, Y. M., and Andrews, A. J. (2013) Quantitating the Specificity and Selectivity of Gcn5-Mediated Acetylation of Histone H3, *Plos One 8*.

[38] McManus, K. J., and Hendzel, M. J. (2003) Quantitative analysis of CBP- and P300-induced histone acetylations in vivo using native chromatin, *Molecular and Cellular Biology 23*, 7611-7627.

[39] Kang, T. H., Park, D. Y., Choi, Y. H., Kim, K. J., Yoon, H. S., and Kim, K. T. (2007) Mitotic histone H3 phosphorylation by vaccinia-related kinase 1 in mammalian cells, *Molecular and Cellular Biology 27*, 8533-8546.

[40] Kalb, R., Mallery, D. L., Larkin, C., Huang, J. T. J., and Hiom, K. (2014) BRCA1 Is a Histone-H2A-Specific Ubiquitin Ligase, *Cell Rep 8*, 999-1005.

[41] Xu, W. T., Edmondson, D. G., and Roth, S. Y. (1998) Mammalian GCN5 and P/CAF acetyltransferases have homologous amino-terminal domains important for recognition of nucleosomal substrates, *Mol Cell Biol 18*, 5659-5669.

[42] Tong, J. K., Hassig, C. A., Schnitzler, G. R., Kingston, R. E., and Schreiber, S. L. (1998) Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex, *Nature 395*, 917-921.

[43] Nishioka, K., Rice, J. C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y. M., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J. T., Allis, C. D., and Reinberg, D. (2002) PR-Set7 is a nucleosome-specific methyltransferase that

modifies lysine 20 of histone H4 and is associated with silent chromatin, *Molecular Cell 9*, 1201-1213.

[44] Zhang, T. Y., Cooper, S., and Brockdorff, N. (2015) The interplay of histone modifications - writers that read, *Embo Rep 16*, 1467-1481.

[45] Holt, M., and Muir, T. (2015) Application of the Protein Semisynthesis Strategy to the Generation of Modified Chromatin, *Annu Rev Biochem 84*, 265-290.

[46] Muller, M. M., and Muir, T. W. (2015) Histones: At the Crossroads of Peptide and Protein Chemistry, *Chemical Reviews 115*, 2296-2349.

[47] Dai, J., Hyland, E. M., Yuan, D. S., Huang, H. L., Bader, J. S., and Boeke, J. D. (2008) Probing nucleosome function: A highly versatile library of synthetic histone H3 and H4 mutants, *Cell 134*, 1066-1078.

[48] Simon, M. D., Chu, F. X., Racki, L. R., de la Cruz, C. C., Burlingame, A. L., Panning, B., Narlikar, G. J., and Shokat, K. M. (2007) The site-specific installation of methyl-lysine analogs into recombinant histones, *Cell 128*, 1003-1012.

[49] Huang, R., Holbert, M. A., Tarrant, M. K., Curtet, S., Colquhoun, D. R., Dancy, B. M., Dancy, B. C., Hwang, Y. S., Tang, Y., Meeth, K., Marmorstein, R., Cole, R. N., Khochbin, S., and Cole, P. A. (2010) Site-Specific Introduction of an Acetyl-Lysine Mimic into Peptides and Proteins by Cysteine Alkylation, *J Am Chem Soc 132*, 9986-9987.

[50] Le, D. D., Cortesi, A. T., Myers, S. A., Burlingame, A. L., and Fujimori, D. G. (2013) Site-Specific and Regiospecific Installation of Methylarginine Analogues

into Recombinant Histones and Insights into Effector Protein Binding, *J Am Chem Soc 135*, 2879-2882.

[51] Lu, X., Simon, M. D., Chodaparambil, J. V., Hansen, J. C., Shokat, K. M., and Luger, K. (2008) The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure, *Nature Structural & Molecular Biology 15*, 1122-1124.

[52] Gellman, S. H. (1991) On the Role of Methionine Residues in the Sequence-Independent Recognition of Nonpolar Protein Surfaces, *Biochemistry-Us 30*, 6633-6636.

[53] Gloss, L. M., and Kirsch, J. F. (1995) Decreasing the Basicity of the Active-Site Base, Lys-258, of Escherichia-Coli Aspartate-Aminotransferase by Replacement with Gamma-Thialysine, *Biochemistry-Us 34*, 3990-3998.

[54] Seeliger, D., Soeroes, S., Klingberg, R., Schwarzer, D., Grubmuller, H., and Fischle, W. (2012) Quantitative Assessment of Protein Interaction with Methyl-Lysine Analogues by Hybrid Computational and Experimental Approaches, *Acs Chem Biol 7*, 150-154.

[55] Wright, T. H., Bower, B. J., Chalker, J. M., Bernardes, G. J. L., Wiewiora, R., Ng, W. L., Raj, R., Faulkner, S., Vallee, M. R. J., Phanumartwiwath, A., Coleman, O. D., Thezenas, M. L., Khan, M., Galan, S. R. G., Lercher, L., Schombs, M. W., Gerstberger, S., Palm-Espling, M. E., Baldwin, A. J., Kessler, B. M., Claridge, T. D. W., Mohammed, S., and Davis, B. G. (2016) Posttranslational mutagenesis: A chemical strategy for exploring protein side-chain diversity, *Science 354*.

[56] Yang, A., Ha, S., Ahn, J., Kim, R., Kim, S., Lee, Y., Kim, J., Soll, D., Lee, H. Y., and Park, H. S. (2016) A chemical biology route to site-specific authentic protein modifications, *Science 354*, 623-626.

[57] Jacobson, R. H., Ladurner, A. G., King, D. S., and Tjian, R. (2000) Structure and function of a human TAF(II)250 double bromodomain module, *Science 288*, 1422-1425.

[58] Moriniere, J., Rousseaux, S., Steuerwald, U., Soler-Lopez, M., Curtet, S., Vitte, A. L., Govin, J., Gaucher, J., Sadoul, K., Hart, D. J., Krijgsveld, J., Khochbin, S., Muller, C. W., and Petosa, C. (2009) Cooperative binding of two acetylation marks on a histone tail by a single bromodomain, *Nature 461*, 664-U112.

[59] Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D. (1999) Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena, *P Natl Acad Sci USA 96*, 14967-14972.

[60] Cook, P. J., Ju, B. G., Telese, F., Wang, X. T., Glass, C. K., and Rosenfeld, M. G. (2009) Tyrosine dephosphorylation of H2AX modulates apoptosis and survival decisions, *Nature 458*, 591-U553.

[61] Gutte, B., and Merrifield, R. B. (1969) The total synthesis of an enzyme with ribonuclease A activity, *J Am Chem Soc 91*, 501-502.

[62] Raibaut, L., El Mahdi, O., and Melnyk, O. (2015) Solid phase protein chemical synthesis, *Top Curr Chem 363*, 103-154.

[63] Hong, L., Schroth, G. P., Matthews, H. R., Yau, P., and Bradbury, E. M. (1993) Studies of the DNA binding properties of histone H4 amino terminus. Thermal

denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA, *J Biol Chem 268*, 305-314.

[64] Owen, D. J., Ornaghi, P., Yang, J. C., Lowe, N., Evans, P. R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A. A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p, *Embo J 19*, 6141-6149.

[65] Jacobs, S. A., and Khorasanizadeh, S. (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail, *Science 295*, 2080-2083.

[66] Li, H., Fischle, W., Wang, W., Duncan, E. M., Liang, L., Murakami-Ishibe, S., Allis, C. D., and Patel, D. J. (2007) Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger, *Mol Cell 28*, 677-691.

[67] Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlautz, H., Luscher, B., and Amati, B. (2007) Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive, *Nature 449*, 933-937.

[68] Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., Hunt, D. F., Funabiki, H., and Allis, C. D. (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation, *Nature 438*, 1116-1122.

[69] MacDonald, N., Welburn, J. P. I., Noble, M. E. M., Nguyen, A., Yaffe, M. B., Clynes, D., Moggs, J. G., Orphanides, G., Thomson, S., Edmunds, J. W., Clayton, A. L., Endicott, J. A., and Mahadevan, L. C. (2005) Molecular basis for

the recognition of phosphorylated and phosphoacetylated histone H3 by 14-3-3, *Molecular Cell 20*, 199-211.

[70] Yuan, Y., Chen, J., Wan, Q. A., Wilson, R. M., and Danishefsky, S. J. (2010) Toward Fully Synthetic, Homogeneous Glycoproteins: Advances in Chemical Ligation, *Biopolymers 94*, 373-384.

[71] van Noort, G. J. V., van der Horst, M. G., Overkleeft, H. S., van der Marel, G. A., and Filippov, D. V. (2010) Synthesis of Mono-ADP-Ribosylated Oligopeptides Using Ribosylated Amino Acid Building Blocks, *J Am Chem Soc 132*, 5236-5240.

[72] Chatterjee, C., McGinty, R. K., Pellois, J. P., and Muir, T. W. (2007) Auxiliary-mediated site-specific peptide ubiquitylation, *Angew Chem Int Edit 46*, 2814-2818.

[73] Voigt, P., and Reinberg, D. (2011) Histone Tails: Ideal Motifs for Probing Epigenetics through Chemical Biology Approaches, *Chembiochem 12*, 236-252.

[74] Smith, B. C., and Denu, J. M. (2007) Sir2 deacetylases exhibit nucleophilic participation of acetyl-lysine in NAD(+) cleavage, *J Am Chem Soc 129*, 5802-+.

[75] Turner, B. M., and Fellows, G. (1989) Specific Antibodies Reveal Ordered and Cell-Cycle-Related Use of Histone-H4 Acetylation Sites in Mammalian-Cells, *Eur J Biochem 179*, 131-139.

[76] Perez-Burgos, L., Peters, A. H. F. M., Opravil, S., Kauer, M., Mechtler, K., and Jenuwein, T. (2004) Generation and characterization of methyl-lysine histone antibodies, *Method Enzymol 376*, 234-254.

[77] Tessarz, P., and Kouzarides, T. (2014) Histone core modifications regulating nucleosome structure and dynamics, *Nat Rev Mol Cell Bio 15*, 703-708.

[78] Ng, M. K., and Cheung, P. (2016) A brief histone in time: understanding the combinatorial functions of histone PTMs in the nucleosome context, *Biochem Cell Biol 94*, 33-42.

[79] Dawson, P. E., Muir, T. W., Clarklewis, I., and Kent, S. B. H. (1994) Synthesis of Proteins by Native Chemical Ligation, *Science 266*, 776-779.

[80] Muir, T. W., Sondhi, D., and Cole, P. A. (1998) Expressed protein ligation: A general method for protein engineering, *P Natl Acad Sci USA 95*, 6705-6710.

[81] Piotukh, K., Geltinger, B., Heinrich, N., Gerth, F., Beyermann, M., Freund, C., and Schwarzer, D. (2011) Directed Evolution of Sortase A Mutants with Altered Substrate Selectivity Profiles, *J Am Chem Soc 133*, 17536-17539.

[82] Maity, S. K., Jbara, M., and Brik, A. (2016) Chemical and semisynthesis of modified histones, *J Pept Sci 22*, 252-259.

[83] Wan, Q., and Danishefsky, S. J. (2007) Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides, *Angew Chem Int Ed Engl 46*, 9248-9252.

[84] Allahverdi, A., Yang, R., Korolev, N., Fan, Y., Davey, C. A., Liu, C. F., and Nordenskiold, L. (2011) The effects of histone H4 tail acetylations on cation-induced chromatin folding and self-association, *Nucleic Acids Res 39*, 1680-1691.

[85] He, S., Bauman, D., Davis, J. S., Loyola, A., Nishioka, K., Gronlund, J. L., Reinberg, D., Meng, F., Kelleher, N., and McCafferty, D. G. (2003) Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis, *Proc Natl Acad Sci U S A 100*, 12033-12038.

[86] North, J. A., Simon, M., Ferdinand, M. B., Shoffner, M. A., Picking, J. W., Howard, C. J., Mooney, A. M., van Noort, J., Poirier, M. G., and Ottesen, J. J. (2014) Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure, *Nucleic Acids Res 42*, 4922-4933.

[87] Simon, M., North, J. A., Shimko, J. C., Forties, R. A., Ferdinand, M. B., Manohar, M., Zhang, M., Fishel, R., Ottesen, J. J., and Poirier, M. G. (2011) Histone fold modifications control nucleosome unwrapping and disassembly, *Proc Natl Acad Sci U S A 108*, 12711-12716.

[88] Maity, S. K., Jbara, M., Mann, G., Kamnesky, G., and Brik, A. (2017) Total chemical synthesis of histones and their analogs, assisted by native chemical ligation and palladium complexes, *Nat Protoc 12*, 2293-2322.

[89] Shimko, J. C., North, J. A., Bruns, A. N., Poirier, M. G., and Ottesen, J. J. (2011) Preparation of Fully Synthetic Histone H3 Reveals That Acetyl-Lysine 56 Facilitates Protein Binding Within Nucleosomes, *Journal of Molecular Biology 408*, 187-204.

[90] Fierz, B., Kilic, S., Hieb, A. R., Luger, K., and Muir, T. W. (2012) Stability of Nucleosomes Containing Homogenously Ubiquitylated H2A and H2B Prepared Using Semisynthesis, *J Am Chem Soc 134*, 19548-19551.

[91] Albericio, F. (2000) Orthogonal protecting groups for N alpha-amino and C-terminal carboxyl functions in solid-phase peptide synthesis, *Biopolymers 55*, 123-139.

[92] Xiao, H., and Schultz, P. G. (2016) At the Interface of Chemical and Biological Synthesis: An Expanded Genetic Code, *Csh Perspect Biol 8*.

[93] Liu, W. S. R., Wang, Y. S., and Wan, W. (2011) Synthesis of proteins with defined posttranslational modifications using the genetic noncanonical amino acid incorporation approach, *Mol Biosyst 7*, 38-47.

[94] Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008) Genetically encoding N-epsilon-acetyllysine in recombinant proteins, *Nat Chem Biol 4*, 232-234.

[95] Neumann, H., Hancock, S. M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J. W. (2009) A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation, *Mol Cell 36*, 153-163.

[96] Kallappagoudar, S., Dammer, E. B., Duong, D. M., Seyfried, N. T., and Lucchesi, J. C. (2013) Expression, purification and proteomic analysis of recombinant histone H4 acetylated at lysine 16, *Proteomics 13*, 1687-1691.

[97] Wilkins, B. J., Hahn, L. E., Heitmuller, S., Frauendorf, H., Valerius, O., Braus, G. H., and Neumann, H. (2015) Genetically Encoding Lysine Modifications on Histone H4, *Acs Chem Biol 10*, 939-944.

[98] Umehara, T., Kim, J., Lee, S., Guo, L. T., Soll, D., and Park, H. S. (2012) N-Acetyl lysyl-tRNA synthetases evolved by a CcdB-based selection possess N-acetyl lysine specificity in vitro and in vivo, *Febs Lett 586*, 729-733.

[99] Gattner, M. J., Vrabel, M., and Carell, T. (2013) Synthesis of epsilon-N-propionyl-, epsilon-N-butyryl-, and epsilon-N-crotonyl-lysine containing histone H3 using the pyrrolysine system, *Chem Commun 49*, 379-381.

[100] Wang, T. Y., Zhou, Q., Li, F. H., Yu, Y., Yin, X. B., and Wang, J. Y. (2015) Genetic Incorporation of N-epsilon-Formyllysine, a New Histone Post-translational Modification, *Chembiochem 16*, 1440-1442.

[101] Xiao, H., Xuan, W. M., Shao, S. D., Liu, T., and Schultz, P. G. (2015) Genetic Incorporation of epsilon-N-2-Hydroxyisobutyryl-lysine into Recombinant Histones, *Acs Chem Biol 10*, 1599-1603.

[102] Wang, Z. P. A., Kurra, Y., Wang, X., Zeng, Y., Lee, Y. J., Sharma, V., Lin, H. N., Dai, S. Y., and Liu, W. S. R. (2017) A Versatile Approach for Site-Specific Lysine Acylation in Proteins, *Angew Chem Int Edit 56*, 1643-1647.

[103] Guo, J. T., Wang, J. Y., Lee, J. S., and Schultz, P. G. (2008) Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins, *Angew Chem Int Edit 47*, 6399-6401.

[104] Wang, Z. Y. U., Wang, Y. S., Pai, P. J., Russell, W. K., Russell, D. H., and Liu, W. S. R. (2012) A Facile Method to Synthesize Histones with Posttranslational Modification Mimics, *Biochemistry-Us 51*, 5232-5234.

[105] Nguyen, D. P., Alai, M. M. G., Kapadnis, P. B., Neumann, H., and Chin, J. W. (2009) Genetically Encoding N-epsilon-Methyl-L-lysine in Recombinant Histones, *J Am Chem Soc 131*, 14194-+.

[106] Wang, Y. S., Wu, B., Wang, Z. Y., Huang, Y., Wan, W., Russell, W. K., Pai, P. J., Moe, Y. N., Russell, D. H., and Liu, W. R. (2010) A genetically encoded photocaged N-epsilon-methyl-L-lysinewz, *Mol Biosyst 6*, 1575-1578.

[107] Ai, H. W., Lee, J. W., and Schultz, P. G. (2010) A method to site-specifically introduce methyllysine into proteins in E. coli, *Chem Commun 46*, 5506-5508.

[108] Nguyen, D. P., Alai, M. M. G., Virdee, S., and Chin, J. W. (2010) Genetically Directing epsilon-N, N-Dimethyl-L-Lysine in Recombinant Histones, *Chem Biol 17*, 1072-1076.

[109] Wang, Z. P. A., Zeng, Y., Kurra, Y., Wang, X., Tharp, J. M., Vatansever, E. C., Hsu, W. W., Dai, S., Fang, X. Q., and Liu, W. S. R. (2017) A Genetically Encoded Allysine for the Synthesis of Proteins with Site-Specific Lysine Dimethylation, *Angew Chem Int Edit 56*, 212-216.

[110] Park, H. S., Hohn, M. J., Umehara, T., Guo, L. T., Osborne, E. M., Benner, J., Noren, C. J., Rinehart, J., and Soll, D. (2011) Expanding the Genetic Code of Escherichia coli with Phosphoserine, *Science 333*, 1151-1154.

[111] Lee, S., Oh, S., Yang, A., Kim, J., Soll, D., Lee, D., and Park, H. S. (2013) A Facile Strategy for Selective Incorporation of Phosphoserine into Histones, *Angew Chem Int Edit 52*, 5771-5775.

[112] Pirman, N. L., Barber, K. W., Aerni, H. R., Ma, N. J., Haimovich, A. D., Rogulina, S., Isaacs, F. J., and Rinehart, J. (2015) A flexible codon in genomically recoded Escherichia coli permits programmable protein phosphorylation, *Nat Commun 6*.

[113] Rogerson, D. T., Sachdeva, A., Wang, K. H., Haq, T., Kazlauskaite, A., Hancock, S. M., Huguenin-Dezot, N., Muqit, M. M. K., Fry, A. M., Bayliss, R., and Chin, J. W. (2015) Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog, *Nat Chem Biol 11*, 496-+.

[114] Zhang, M. S., Brunner, S. F., Huguenin-Dezot, N., Liangl, A. D., Schmied, W. H., Rogerson, D. T., and Chin, J. W. (2017) Biosynthesis and genetic encoding of phosphothreonine through parallel setection and deep sequencing, *Nat Methods 14*, 729-+.

[115] Xie, J. M., Supekova, L., and Schultz, P. G. (2007) A genetically encoded metabolically stable analogue of phosphotyrosine in Escherichia coli, *Acs Chem Biol 2*, 474-478.

[116] Serwa, R., Wilkening, I., Del Signore, G., Muhlberg, M., Claussnitzer, I., Weise, C., Gerrits, M., and Hackenberger, C. P. R. (2009) Chemoselective Staudinger-Phosphite Reaction of Azides for the Phosphorylation of Proteins, *Angew Chem Int Edit 48*, 8234-8239.

[117] Luo, X. Z., Fu, G. S., Wang, R. S. E., Zhu, X. Y., Zambaldo, C., Liu, R. H., Liu, T., Lyu, X. X., Du, J. T., Xuan, W. M., Yao, A. Z., Reed, S. A., Kang, M. C., Zhang, Y. H., Guo, H., Huang, C. H., Yang, P. Y., Wilson, I. A., Schultz, P. G., and Wang, F. (2017) Genetically encoding phosphotyrosine and its nonhydrolyzable analog in bacteria, *Nat Chem Biol 13*, 845-+.

[118] Hoppmann, C., Wong, A., Yang, B., Li, S. W., Hunter, T., Shokat, K. M., and Wang, L. (2017) Site-specific incorporation of phosphotyrosine using an expanded genetic code, *Nat Chem Biol 13*, 842-+.

[119] Li, X., Fekner, T., Ottesen, J. J., and Chan, M. K. (2009) A Pyrrolysine Analogue for Site-Specific Protein Ubiquitination, *Angew Chem Int Edit 48*, 9184-9187.

[120] Virdee, S., Kapadnis, P. B., Elliott, T., Lang, K., Madrzak, J., Nguyen, D. P., Riechmann, L., and Chin, J. W. (2011) Traceless and Site-Specific Ubiquitination of Recombinant Proteins, *J Am Chem Soc 133*, 10708-10711.

[121] Iwasaki, W., Miya, Y., Horikoshi, N., Osakabe, A., Taguchi, H., Tachiwana, H., Shibata, T., Kagawa, W., and Kurumizaka, H. (2013) Contribution of histone N-terminal tails to the structure and stability of nucleosomes, *Febs Open Bio 3*, 363-369.

[122] Dorigo, B., Schalch, T., Bystricky, K., and Richmond, T. J. (2003) Chromatin fiber folding: requirement for the histone H4 N-terminal tail, *J Mol Biol 327*, 85-96.

[123] Wang, Z., Zang, C., Cui, K., Schones, D. E., Barski, A., Peng, W., and Zhao, K. (2009) Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes, *Cell 138*, 1019-1031.

[124] Shogren-Knaak, M., Ishii, H., Sun, J. M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions, *Science 311*, 844-847.

[125] Mutskov, V., Gerber, D., Angelov, D., Ausio, J., Workman, J., and Dimitrov, S. (1998) Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding, *Mol Cell Biol 18*, 6293-6304.

[126] Tencer, A. H., Cox, K. L., Di, L., Bridgers, J. B., Lyu, J., Wang, X., Sims, J. K., Weaver, T. M., Allen, H. F., Zhang, Y., Gatchalian, J., Darcy, M. A., Gibson, M. D., Ikebe, J., Li, W., Wade, P. A., Hayes, J. J., Strahl, B. D., Kono, H., Poirier, M. G., Musselman, C. A., and Kutateladze, T. G. (2017) Covalent Modifications of Histone H3K9 Promote Binding of CHD3, *Cell Rep 21*, 455-466.

[127] Luger, K., Dechassa, M. L., and Tremethick, D. J. (2012) New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?, *Nat Rev Mol Cell Bio 13*, 436-447.

[128] Xu, F., Zhang, K. L., and Grunstein, M. (2005) Acetylation in histone H3 globular domain regulates gene expression in yeast, *Cell 121*, 375-385.

[129] Chen, C. C., Carson, J. J., Feser, J., Tamburini, B., Zabaronick, S., Linger, J., and Tyler, J. K. (2008) Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair, *Cell 134*, 231-243.

[130] Han, J. H., Zhou, H., Horazdovsky, B., Zhang, K. L., Xu, R. M., and Zhang, Z. G. (2007) Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication, *Science 315*, 653-655.

[131] Tan, Y. L., Xue, Y., Song, C. Y., and Grunstein, M. (2013) Acetylated histone H3K56 interacts with Oct4 to promote mouse embryonic stem cell pluripotency, *P Natl Acad Sci USA 110*, 11493-11498.

[132] Di Cerbo, V., Mohn, F., Ryan, D. P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., Richter, F. M., Bannister, A. J., Mittler, G., Michaelis, J., Khochbin, S., Feil, R., Schuebeler, D., Owen-Hughes, T., Daujat, S., and Schneider, R. (2014) Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription, *Elife 3*.

[133] Anderson, K. W., Mast, N., Pikuleva, I. A., and Turko, I. V. (2015) Histone H3 Ser57 and Thr58 phosphorylation in the brain of 5XFAD mice, *Febs Open Bio 5*, 550-556.

[134] Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., Ottesen, J. J., Luger, K., and Poirier, M. G. (2015) Histone Core Phosphorylation Regulates DNA Accessibility, *J Biol Chem 290*, 22612-22621.

[135] Sudarsanam, P., Iyer, V. R., Brown, P. O., and Winston, F. (2000) Whole-genome expression analysis of snf/swi mutants of Saccharomyces cerevisiae, *P Natl Acad Sci USA 97*, 3364-3369.

[136] Kurumizaka, H., and Wolffe, A. P. (1997) Sin mutations of histone H3: influence on nucleosome core structure and function, *Mol Cell Biol 17*, 6953-6969.

[137] Luger, K., and Richmond, T. J. (1998) DNA binding within the nucleosome core, *Curr Opin Struc Biol 8*, 33-40.

[138] Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hughes, T. (2004) Sin mutations alter inherent nucleosome mobility, *Embo J 23*, 343-353.

[139] Fleming, A. B., and Pennings, S. (2001) Antagonistic remodelling by Swi-Snf and Tup1-Ssn6 of an extensive chromatin region forms the background for FLO1 gene regulation, *Embo J 20*, 5219-5231.

[140] Wechser, M. A., Kladde, M. P., Alfieri, J. A., and Peterson, C. L. (1997) Effects of Sin- versions of histone H4 on yeast chromatin structure and function, *Embo J 16*, 2086-2095.

[141] Horn, P. J., Crowley, K. A., Carruthers, L. M., Hansen, J. C., and Peterson, C. L. (2002) The SIN domain of the histone octamer is essential for intramolecular folding of nucleosomal arrays, *Nature Structural Biology 9*, 167-171.

[142] Hyland, E. A., Cosgrove, M. S., Molina, H., Wang, D. X., Pandey, A., Cottee, R. J., and Boeke, J. D. (2005) Insights into the role of histone H3 and histone H4 core modifiable residues in Saccharomyces cerevisiae, *Molecular and Cellular Biology 25*, 10060-10070.

[143] Strick, T. R., Allemand, J. F., Bensimon, D., Bensimon, A., and Croquette, V. (1996) The elasticity of a single supercoiled DNA molecule, *Science 271*, 1835-1837.

[144] Muthurajan, U. M., Bao, Y. H., Forsberg, L. J., Edayathumangalam, R. S., Dyer, P. N., White, C. L., and Luger, K. (2004) Crystal structures of histone Sin mutant nucleosomes reveal altered protein-DNA interactions, *Embo J 23*, 260-271.

[145] North, J. A., Javaid, S., Ferdinand, M. B., Chatterjee, N., Picking, J. W., Shoffner, M., Nakkula, R. J., Bartholomew, B., Ottesen, J. J., Fishel, R., and Poirier, M. G. (2011) Phosphorylation of histone H3(T118) alters nucleosome dynamics and remodeling, *Nucleic Acids Res 39*, 6465-6474.

[146] Park, J. H., Cosgrove, M. S., Youngman, E., Wolberger, C., and Boeke, J. D. (2002) A core nucleosome surface crucial for transcriptional silencing, *Nat Genet 32*, 273-279.

[147] Fry, C. J., Norris, A., Cosgrove, M., Boeke, J. D., and Peterson, C. L. (2006) The LRS and SIN domains: Two structurally equivalent but functionally distinct nucleosomal surfaces required for transcriptional silencing, *Molecular and Cellular Biology 26*, 9045-9059.

[148] Shanower, G. A., Muller, M., Blanton, J. L., Honti, V., Gyurkovics, H., and Schedl, P. (2005) Characterization of the grappa gene, the Drosophila histone H3 lysine 79 methyltransferase, *Genetics 169*, 173-184.

[149] Zhang, L. W., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003) Identification of novel histone post-translational modifications by peptide mass fingerprinting, *Chromosoma 112*, 77-86.

[150] Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005) A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response, *Nature 436*, 294-298.

[151] Ransom, M., Dennehey, B. K., and Tyler, J. K. (2010) Chaperoning Histones during DNA Replication and Repair, *Cell 140*, 183-195.

[152] Andrews, A. J., Chen, X., Zevin, A., Stargell, L. A., and Luger, K. (2010) The Histone Chaperone Nap1 Promotes Nucleosome Assembly by Eliminating Nonnucleosomal Histone DNA Interactions, *Molecular Cell 37*, 834-842.

[153] Donham, D. C., 2nd, Scorgie, J. K., and Churchill, M. E. (2011) The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3/H4-DNA complexes, *Nucleic Acids Res 39*, 5449-5458.

[154] Winkler, D. D., Zhou, H., Dar, M. A., Zhang, Z., and Luger, K. (2012) Yeast CAF-1 assembles histone (H3-H4)2 tetramers prior to DNA deposition, *Nucleic Acids Res 40*, 10139-10149.

[155] Su, D., Hu, Q., Li, Q., Thompson, J. R., Cui, G., Fazly, A., Davies, B. A., Botuyan, M. V., Zhang, Z., and Mer, G. (2012) Structural basis for recognition of H3K56-acetylated histone H3-H4 by the chaperone Rtt106, *Nature 483*, 104-107.

[156] Muller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., Miller, E. L., O'Connor, M. B., Kingston, R. E., and Simon, J. A. (2002) Histone

methyltransferase activity of a Drosophila Polycomb group repressor complex, *Cell 111*, 197-208.

[157] Papp, B., and Muller, J. (2006) Histone trimethylation and the maintenance of transcriptional ON and OFF states by trxG and PcG proteins, *Genes Dev 20*, 2041-2054.

[158] Byrd, K. N., and Shearn, A. (2003) ASH1, a Drosophila trithorax group protein, is required for methylation of lysine 4 residues on histone H3, *P Natl Acad Sci USA 100*, 11535-11540.

[159] Gregory, G. D., Vakoc, C. R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G. A. (2007) Mammalian ASH1L is a histone methyl transferase that occupies the transcribed region of active genes, *Molecular and Cellular Biology 27*, 8466-8479.

[160] Barski, A., Cuddapah, S., Cui, K. R., Roh, T. Y., Schones, D. E., Wang, Z. B., Wei, G., Chepelev, I., and Zhao, K. J. (2007) High-resolution profiling of histone methylations in the human genome, *Cell 129*, 823-837.

[161] Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011) H3K36 Methylation Antagonizes PRC2-mediated H3K27 Methylation, *J Biol Chem 286*, 7983-7989.

[162] Li, Y., Trojer, P., Xu, C. F., Cheung, P., Kuo, A., Drury, W. J., Qiao, Q., Neubert, T. A., Xu, R. M., Gozani, O., and Reinberg, D. (2009) The Target of the NSD Family of Histone Lysine Methyltransferases Depends on the Nature of the Substrate, *J Biol Chem 284*, 34283-34295.

[163] Brownell, J. E., Zhou, J. X., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., and Allis, C. D. (1996) Tetrahymena histone acetyltransferase A: A homolog to yeast Gcn5p linking histone acetylation to gene activation, *Cell 84*, 843-851.

[164] Cheung, P., Tanner, K. G., Cheung, W. L., Sassone-Corsi, P., Denu, J. M., and Allis, C. D. (2000) Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation, *Molecular Cell 5*, 905-915.

[165] Lo, W. S., Trievel, R. C., Rojas, J. R., Duggan, L., Hsu, J. Y., Allis, C. D., Marmorstein, R., and Berger, S. L. (2000) Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14, *Molecular Cell 5*, 917-926.

[166] Shogren-Knaak, M. A., Fry, C. J., and Peterson, C. L. (2003) A native peptide Ligation strategy for deciphering nucleosomal histone modifications, *Journal of Biological Chemistry 278*, 15744-15748.

[167] Li, S., and Shogren-Knaak, M. A. (2008) Cross-talk between histone H3 tails produces cooperative nucleosome acetylation, *P Natl Acad Sci USA 105*, 18243-18248.

[168] Hassan, A. H., Prochasson, P., Neely, K. E., Galasinski, S. C., Chandy, M., Carrozza, M. J., and Workman, J. L. (2002) Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes, *Cell 111*, 369-379.

[169] Li, S. S., and Shogren-Knaak, M. A. (2009) The Gcn5 Bromodomain of the SAGA Complex Facilitates Cooperative and Cross-tail Acetylation of Nucleosomes, *J Biol Chem 284*, 9411-9417.

[170] Ringel, A. E., Cieniewicz, A. M., Taverna, S. D., and Wolberger, C. (2015) Nucleosome competition reveals processive acetylation by the SAGA HAT module, *P Natl Acad Sci USA 112*, E5461-E5470.

[171] Grewal, S. I., and Elgin, S. C. (2007) Transcription and RNA interference in the formation of heterochromatin, *Nature 447*, 399-406.

[172] Verschure, P. J., van der Kraan, I., de Leeuw, W., van der Vlag, J., Carpenter, A. E., Belmont, A. S., and van Driel, R. (2005) In vivo HP1 targeting causes large-scale chromatin condensation and enhanced histone lysine methylation, *Mol Cell Biol 25*, 4552-4564.

[173] Eskeland, R., Eberharter, A., and Imhof, A. (2007) HP1 binding to chromatin methylated at H3K9 is enhanced by auxiliary factors, *Molecular and Cellular Biology 27*, 453-465.

[174] Eberharter, A., Vetter, I., Ferreira, R., and Becker, P. B. (2004) ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts, *Embo J 23*, 4029-4039.

[175] Ner, S. S., Harrington, M. J., and Grigliatti, T. A. (2002) A role for the drosophila SU(VAR)3-9 protein in chromatin organization at the histone gene cluster and in suppression of position-effect variegation, *Genetics 162*, 1763-1774.

[176] Kilic, S., Bachmann, A. L., Bryan, L. C., and Fierz, B. (2015) Multivalency governs HP1 alpha association dynamics with the silent chromatin state, *Nat Commun 6*.

[177] Munari, F., Soeroes, S., Zenn, H. M., Schomburg, A., Kost, N., Schroder, S., Klingberg, R., Rezaei-Ghaleh, N., Stutzer, A., Gelato, K. A., Walla, P. J., Becker, S., Schwarzer, D., Zimmermann, B., Fischle, W., and Zweckstetter, M. (2012) Methylation of Lysine 9 in Histone H3 Directs Alternative Modes of Highly Dynamic Interaction of Heterochromatin Protein hHP1 beta with the Nucleosome, *J Biol Chem 287*, 33756-33765.

[178] Hiragami-Hamada, K., Soeroes, S., Nikolov, M., Wilkins, B., Kreuz, S., Chen, C., De La Rosa-Velazquez, I. A., Zenn, H. M., Kost, N., Pohl, W., Chernev, A., Schwarzer, D., Jenuwein, T., Lorincz, M., Zimmermann, B., Walla, P. J., Neumann, H., Baubec, T., Urlaub, H., and Fischle, W. (2016) Dynamic and flexible H3K9me3 bridging via HP1 beta dimerization establishes a plastic state of condensed chromatin, *Nat Commun 7*.

[179] Fradet-Turcotte, A., Canny, M. D., Escribano-Diaz, C., Orthwein, A., Leung, C. C. Y., Huang, H., Landry, M. C., Kitevski-LeBlanc, J., Noordermeer, S. M., Sicheri, F., and Durocher, D. (2013) 53BP1 is a reader of the DNA-damage-induced H2A Lys 15 ubiquitin mark, *Nature 499*, 50-+.

[180] Wilson, M. D., Benlekbir, S., Fradet-Turcotte, A., Sherker, A., Julien, J. P., McEwan, A., Noordermeer, S. M., Sicheri, F., Rubinstein, J. L., and Durocher,

D. (2016) The structural basis of modified nucleosome recognition by 53BP1, *Nature 536*, 100-+.

[181] Sanchez, R., and Zhou, M. M. (2009) The role of human bromodomains in chromatin biology and gene transcription, *Curr Opin Drug Disc 12*, 659-665.

[182] Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J. P., Barsyte-Lovejoy, D., Felletar, I., Volkmer, R., Muller, S., Pawson, T., Gingras, A. C., Arrowsmith, C. H., and Knapp, S. (2012) Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family, *Cell 149*, 214-231.

[183] Miller, T. C., Simon, B., Rybin, V., Grotsch, H., Curtet, S., Khochbin, S., Carlomagno, T., and Muller, C. W. (2016) A bromodomain-DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT, *Nat Commun 7*, 13855.

[184] Zhou, V. W., Goren, A., and Bernstein, B. E. (2011) Charting histone modifications and the functional organization of mammalian genomes, *Nat Rev Genet 12*, 7-18.

[185] Suganuma, T., and Workman, J. L. (2011) Signals and combinatorial functions of histone modifications, *Annu Rev Biochem 80*, 473-499.

[186] An, W. J., and Roeder, R. G. (2004) Reconstitution and transcriptional analysis of chromatin in vitro, *Chromatin and Chromatin Remodeling Enzymes, Pt C 377*, 460-474.

[187] Vermeulen, M., Mulder, K. W., Denissov, S., Pijnappel, W. W. M. P., van Schaik, F. M. A., Varier, R. A., Baltissen, M. P. A., Stunnenberg, H. G., Mann, M., and

Timmers, H. T. M. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4, *Cell 131*, 58-69.

[188] Tang, Z. Y., Chen, W. Y., Shimada, M., Nguyen, U. T. T., Kim, J., Sun, X. J., Sengoku, T., McGinty, R. K., Fernandez, J. P., Muir, T. W., and Roeder, R. G. (2013) SET1 and p300 Act Synergistically, through Coupled Histone Modifications, in Transcriptional Activation by p53, *Cell 154*, 297-310.

[189] Casadio, F., Lu, X. D., Pollock, S. B., LeRoy, G., Garcia, B. A., Muir, T. W., Roeder, R. G., and Allis, C. D. (2013) H3R42me2a is a histone modification with positive transcriptional effects, *P Natl Acad Sci USA 110*, 14894-14899.

[190] Tan, M. J., Luo, H., Lee, S., Jin, F. L., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z. Y., Rousseaux, S., Rajagopal, N., Lu, Z. K., Ye, Z., Zhu, Q., Wysocka, J., Ye, Y., Khochbin, S., Ren, B., and Zhao, Y. M. (2011) Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification, *Cell 146*, 1015-1027.

[191] Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G. H., Mittler, G., Liu, E. T., Buhler, M., Margueron, R., and Schneider, R. (2013) Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer, *Cell 152*, 859-872.

[192] Imai, S., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000) Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase, *Nature 403*, 795-800.

[193] Brachmann, C. B., Sherman, J. M., Devine, S. E., Cameron, E. E., Pillus, L., and Boeke, J. D. (1995) The SIR2 gene family, conserved from bacteria to humans, functions in silencing, cell cycle progression, and chromosome stability, *Genes & development 9*, 2888-2902.

[194] Herskovits, A. Z., and Guarente, L. (2013) Sirtuin deacetylases in neurodegenerative diseases of aging, *Cell Res 23*, 746-758.

[195] Michishita, E., Park, J. Y., Burneskis, J. M., Barrett, J. C., and Horikawa, I. (2005) Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins, *Mol Biol Cell 16*, 4623-4635.

[196] Gut, P., and Verdin, E. (2013) The nexus of chromatin regulation and intermediary metabolism, *Nature 502*, 489-498.

[197] Blander, G., Olejnik, J., Olejnik, E. K., Mcdonagh, T., Haigis, M., Yaffe, M. B., and Guarente, L. (2005) SIRT1 shows no substrate specificity in vitro, *Journal of Biological Chemistry 280*, 9780-9785.

[198] Pan, P. W., Feldman, J. L., Devries, M. K., Dong, A. P., Edwards, A. M., and Denu, J. M. (2011) Structure and Biochemical Functions of SIRT6, *Journal of Biological Chemistry 286*, 14575-14587.

[199] Feldman, J. L., Baeza, J., and Denu, J. M. (2013) Activation of the protein deacetylase SIRT6 by long-chain fatty acids and widespread deacylation by mammalian sirtuins, *J Biol Chem 288*, 31350-31356.

[200] Jiang, H., Khan, S., Wang, Y., Charron, G., He, B., Sebastian, C., Du, J. T., Kim, R., Ge, E., Mostoslavsky, R., Hang, H. C., Hao, Q., and Lin, H. N. (2013) SIRT6

regulates TNF-alpha secretion through hydrolysis of long-chain fatty acyl lysine, *Nature 496*, 110-+.

[201] Schwer, B., Schumacher, B., Lombard, D. B., Xiao, C. Y., Kurtev, M. V., Gao, J., Schneider, J. I., Chai, H., Bronson, R. T., Tsai, L. H., Deng, C. X., and Alt, F. W. (2010) Neural sirtuin 6 (Sirt6) ablation attenuates somatic growth and causes obesity, *P Natl Acad Sci USA 107*, 21790-21794.

[202] Michishita, E., McCord, R. A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T. L. A., Barrett, J. C., Chang, H. Y., Bohr, V. A., Ried, T., Gozani, O., and Chua, K. F. (2008) SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin, *Nature 452*, 492-U416.

[203] Kugel, S., Feldman, J. L., Klein, M. A., Silberman, D. M., Sebastian, C., Mermel, C., Dobersch, S., Clark, A. R., Getz, G., Denu, J. M., and Mostoslavsky, R. (2015) Identification of and Molecular Basis for SIRT6 Loss-of-Function Point Mutations in Cancer, *Cell Rep 13*, 479-488.

[204] Michishita, E., McCord, R. A., Boxer, L. D., Barber, M. F., Hong, T., Gozani, O., and Chua, K. F. (2009) Cell cycle-dependent deacetylation of telomeric histone H3 lysine K56 by human SIRT6, *Cell Cycle 8*, 2664-2666.

[205] Gil, R., Barth, S., Kanfi, Y., and Cohen, H. Y. (2013) SIRT6 exhibits nucleosome-dependent deacetylase activity, *Nucleic Acids Res 41*, 8537-8545.

[206] Müller, M. M., and Muir, T. W. (2015) Histones: At the Crossroads of Peptide and Protein Chemistry, *Chemical Reviews 115*, 2296-2349.

[207] Kugel, S., and Mostoslavsky, R. (2014) Chromatin and beyond: the multitasking roles for SIRT6, *Trends Biochem Sci 39*, 72-81.

[208] Baur, J. A., Ungvari, Z., Minor, R. K., Le Couteur, D. G., and de Cabo, R. (2012) Are sirtuins viable targets for improving healthspan and lifespan?, *Nat Rev Drug Discov 11*, 443-461.

[209] Mostoslavsky, R., Chua, K. F., Lombard, D. B., Pang, W. W., Fischer, M. R., Gellon, L., Liu, P. F., Mostoslavsky, G., Franco, S., Murphy, M. M., Mills, K. D., Patel, P., Hsu, J. T., Hong, A. L., Ford, E., Cheng, H. L., Kennedy, C., Nunez, N., Bronson, R., Frendewey, D., Auerbach, W., Valenzuela, D., Karow, M., Hottiger, M. O., Hursting, S., Barrett, J. C., Guarente, L., Mulligan, R., Demple, B., Yancopoulos, G. D., and Alt, F. W. (2006) Genomic instability and aging-like phenotype in the absence of mammalian SIRT6, *Cell 124*, 315-329.

[210] Badea, L., Herlea, V., Dima, S. O., Dumitrascu, T., and Popescu, I. (2008) Combined Gene Expression Analysis of Whole-Tissue and Microdissected Pancreatic Ductal Adenocarcinoma identifies Genes Specifically Overexpressed in Tumor Epithelia, *Hepato-Gastroenterol 55*, 2016-2027.

[211] Anders, M., Fehlker, M., Wang, Q., Wissmann, C., Pilarsky, C., Kemmner, W., and Hocker, M. (2013) Microarray meta-analysis defines global angiogenesis-related gene expression signatures in human carcinomas, *Mol Carcinogen 52*, 29-38.

[212] Azuma, Y., Yokobori, T., Mogi, A., Altan, B., Yajima, T., Kosaka, T., Onozato, R., Yamaki, E., Asao, T., Nishiyama, M., and Kuwano, H. (2015) SIRT6

expression is associated with poor prognosis and chemosensitivity in patients with non-small cell lung cancer, *J Surg Oncol 112*, 231-237.

[213] Khongkow, M., Olmos, Y., Gong, C., Gomes, A. R., Monteiro, L. J., Yague, E., Cavaco, T. B., Khongkow, P., Man, E. P., Laohasinnarong, S., Koo, C. Y., Harada-Shoji, N., Tsang, J. W., Coombes, R. C., Schwer, B., Khoo, U. S., and Lam, E. W. (2013) SIRT6 modulates paclitaxel and epirubicin resistance and survival in breast cancer, *Carcinogenesis 34*, 1476-1486.

[214] Liu, Y., Xie, Q. R., Wang, B., Shao, J., Zhang, T., Liu, T., Huang, G., and Xia, W. (2013) Inhibition of SIRT6 in prostate cancer reduces cell viability and increases sensitivity to chemotherapeutics, *Protein Cell*.

[215] Ming, M., Han, W. N., Zhao, B. Z., Sundaresan, N. R., Deng, C. X., Gupta, M. P., and He, Y. Y. (2014) SIRT6 Promotes COX-2 Expression and Acts as an Oncogene in Skin Cancer, *Cancer Research 74*, 5925-5933.

[216] Sugatani, T., Agapova, O., Malluche, H. H., and Hruska, K. A. (2015) SIRT6 deficiency culminates in low-turnover osteopenia, *Bone 81*, 168-177.

[217] Lee, Y. J., Wu, B., Raymond, J. E., Zeng, Y., Fang, X., Wooley, K. L., and Liu, W. R. (2013) A genetically encoded acrylamide functionality, *Acs Chem Biol 8*, 1664-1670.

[218] Lee, Y. J., Kurra, Y., Yang, Y., Torres-Kolbus, J., Deiters, A., and Liu, W. R. (2014) Genetically encoded unstrained olefins for live cell labeling with tetrazine dyes, *Chemical communications 50*, 13085-13088.

[219] Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Preparation of nucleosome core particle from recombinant histones, *Methods Enzymol 304*, 3-19.

[220] Ram, O., Goren, A., Amit, I., Shoresh, N., Yosef, N., Ernst, J., Kellis, M., Gymrek, M., Issner, R., Coyne, M., Durham, T., Zhang, X., Donaghey, J., Epstein, C. B., Regev, A., and Bernstein, B. E. (2011) Combinatorial patterning of chromatin regulators uncovered by genome-wide location analysis in human cells, *Cell 147*, 1628-1639.

[221] Rufiange, A., Jacques, P. E., Bhat, W., Robert, F., and Nourani, A. (2007) Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1, *Molecular cell 27*, 393-405.

[222] Lalonde, M. E., Cheng, X., and Cote, J. (2014) Histone target selection within chromatin: an exemplary case of teamwork, *Genes & development 28*, 1029-1041.

[223] Cai, Y., Jin, J. J., Swanson, S. K., Cole, M. D., Choi, S. H., Florens, L., Washburn, M. P., Conaway, J. W., and Conaway, R. C. (2010) Subunit Composition and Substrate Specificity of a MOF-containing Histone Acetyltransferase Distinct from the Male-specific Lethal (MSL) Complex, *Journal of Biological Chemistry 285*, 4268-4272.

[224] Guenther, M. G., Barak, O., and Lazar, M. A. (2001) The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3, *Mol Cell Biol 21*, 6091-6101.

[225] Tennen, R. I., Berber, E., and Chua, K. F. (2010) Functional dissection of SIRT6: Identification of domains that regulate histone deacetylase activity and chromatin localization, *Mechanisms of Ageing and Development 131*, 185-192.

[226] Smith, B. C., Settles, B., Hallows, W. C., Craven, M. W., and Denu, J. M. (2011) SIRT3 Substrate Specificity Determined by Peptide Arrays and Machine Learning, *Acs Chem Biol 6*, 146-157.

[227] Etchegaray, J. P., Chavez, L., Huang, Y., Ross, K. N., Choi, J., Martinez-Pastor, B., Walsh, R. M., Sommer, C. A., Lienhard, M., Gladden, A., Kugel, S., Silberman, D. M., Ramaswamy, S., Mostoslavsky, G., Hochedlinger, K., Goren, A., Rao, A., and Mostoslavsky, R. (2015) The histone deacetylase SIRT6 controls embryonic stem cell fate via TET-mediated production of 5-hydroxymethylcytosine, *Nat Cell Biol 17*, 545-557.

[228] Zhong, L., D'Urso, A., Toiber, D., Sebastian, C., Henry, R. E., Vadysirisack, D. D., Guimaraes, A., Marinelli, B., Wikstrom, J. D., Nir, T., Clish, C. B., Vaitheesvaran, B., Iliopoulos, O., Kurland, I., Dor, Y., Weissleder, R., Shirihai, O. S., Ellisen, L. W., Espinosa, J. M., and Mostoslavsky, R. (2010) The Histone Deacetylase Sirt6 Regulates Glucose Homeostasis via Hif1 alpha, *Cell 140*, 280-293.

[229] Yang, B., Zwaans, B. M., Eckersdorff, M., and Lombard, D. B. (2009) The sirtuin SIRT6 deacetylates H3 K56Ac in vivo to promote genomic stability, *Cell Cycle 8*, 2662-2663.

[230] Toiber, D., Erdel, F., Bouazoune, K., Silberman, D. M., Zhong, L., Mulligan, P., Sebastian, C., Cosentino, C., Martinez-Pastor, B., Giacosa, S., D'Urso, A., Naar, A. M., Kingston, R., Rippe, K., and Mostoslavsky, R. (2013) SIRT6 recruits SNF2H to DNA break sites, preventing genomic instability through chromatin remodeling, *Molecular cell 51*, 454-468.

[231] Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Peng, W., Zhang, M. Q., and Zhao, K. (2008) Combinatorial patterns of histone acetylations and methylations in the human genome, *Nature genetics 40*, 897-903.

[232] Martin, A. M., Pouchnik, D. J., Walker, J. L., and Wyrick, J. J. (2004) Redundant roles for histone H3 N-terminal lysine residues in subtelomeric gene repression in Saccharomyces cerevisiae, *Genetics 167*, 1123-1132.

[233] Li, Y. Y., Wen, H., Xi, Y. X., Tanaka, K., Wang, H. B., Peng, D. N., Ren, Y. F., Jin, Q. H., Dent, S. Y. R., Li, W., Li, H. T., and Shi, X. B. (2014) AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation, *Cell 159*, 558-571.

[234] Sparmann, A., and van Lohuizen, M. (2006) Polycomb silencers control cell fate, development and cancer, *Nat Rev Cancer 6*, 846-856.

[235] Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., Hanna, J., Lodato, M. A., Frampton, G. M., Sharp, P. A., Boyer, L. A., Young, R. A., and Jaenisch, R. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state, *Proc Natl Acad Sci U S A 107*, 21931-21936.

[236] Manuyakorn, A., Paulus, R., Farrell, J., Dawson, N. A., Tze, S., Cheung-Lau, G., Hines, O. J., Reber, H., Seligson, D. B., Horvath, S., Kurdistani, S. K., Guha, C., and Dawson, D. W. (2010) Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704, *J Clin Oncol 28*, 1358-1365.

[237] Seligson, D. B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M., and Kurdistani, S. K. (2005) Global histone modification patterns predict risk of prostate cancer recurrence, *Nature 435*, 1262-1266.

[238] Richmond, T. J. (1999) Hot papers - Crystal structure - Crystal structure of the nucleosome core particle at 2.8 angstrom resolution by K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond - Comments, *Scientist 13*, 15-15.

[239] Struhl, K., and Segal, E. (2013) Determinants of nucleosome positioning, *Nat Struct Mol Biol 20*, 267-273.

[240] Denu, J. M. (2005) The Sir 2 family of protein deacetylases, *Curr Opin Chem Biol 9*, 431-440.

[241] Houtkooper, R. H., Pirinen, E., and Auwerx, J. (2012) Sirtuins as regulators of metabolism and healthspan, *Nat Rev Mol Cell Bio 13*, 225-238.

[242] Vaquero, A., Scher, M., Erdjument-Bromage, H., Tempst, P., Serrano, L., and Reinberg, D. (2007) SIRT1 regulates the histone methyl-transferase SUV39H1 during heterochromatin formation, *Nature 450*, 440-444.

[243] North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M., and Verdin, E. (2003) The human Sir2 ortholog, SIRT2, is an NAD+-dependent tubulin deacetylase, *Molecular cell 11*, 437-444.

[244] Grob, A., Roussel, P., Wright, J. E., McStay, B., Hernandez-Verdun, D., and Sirri, V. (2009) Involvement of SIRT7 in resumption of rDNA transcription at the exit from mitosis, *Journal of cell science 122*, 489-498.

[245] Kiran, S., Anwar, T., Kiran, M., and Ramakrishna, G. (2015) Sirtuin 7 in cell proliferation, stress and disease: Rise of the Seventh Sirtuin!, *Cell Signal 27*, 673-682.

[246] Martinez-Redondo, P., Santos-Barriopedro, I., and Vaquero, A. (2012) A big step for SIRT7, one giant leap for Sirtuins... in cancer, *Cancer Cell 21*, 719-721.

[247] Lu, C. T., Hsu, C. M., Lin, P. M., Lai, C. C., Lin, H. C., Yang, C. H., Hsiao, H. H., Liu, Y. C., Lin, H. Y., Lin, S. F., and Yang, M. Y. (2014) The potential of SIRT6 and SIRT7 as circulating markers for head and neck squamous cell carcinoma, *Anticancer Res 34*, 7137-7143.

[248] Yu, H., Ye, W., Wu, J., Meng, X., Liu, R. Y., Ying, X., Zhou, Y., Wang, H., Pan, C., and Huang, W. (2014) Overexpression of sirt7 exhibits oncogenic property and serves as a prognostic factor in colorectal cancer, *Clin Cancer Res 20*, 3434-3445.

[249] Aljada, A., Saleh, A. M., Alkathiri, M., Shamsa, H. B., Al-Bawab, A., and Nasr, A. (2015) Altered Sirtuin 7 Expression is Associated with Early Stage Breast Cancer, *Breast Cancer (Auckl) 9*, 3-8.

[250] Geng, Q., Peng, H., Chen, F., Luo, R., and Li, R. (2015) High expression of Sirt7 served as a predictor of adverse outcome in breast cancer, *Int J Clin Exp Pathol 8*, 1938-1945.

[251] Malik, S., Villanova, L., Tanaka, S., Aonuma, M., Roy, N., Berber, E., Pollack, J. R., Michishita-Kioi, E., and Chua, K. F. (2015) SIRT7 inactivation reverses metastatic phenotypes in epithelial and mesenchymal tumors, *Scientific reports 5*, 9841.

[252] McGlynn, L. M., McCluney, S., Jamieson, N. B., Thomson, J., MacDonald, A. I., Oien, K., Dickson, E. J., Carter, C. R., McKay, C. J., and Shiels, P. G. (2015) SIRT3 & SIRT7: Potential Novel Biomarkers for Determining Outcome in Pancreatic Cancer Patients, *PLoS One 10*, e0131344.

[253] Singh, S., Kumar, P. U., Thakur, S., Kiran, S., Sen, B., Sharma, S., Rao, V. V., Poongothai, A. R., and Ramakrishna, G. (2015) Expression/localization patterns of sirtuins (SIRT1, SIRT2, and SIRT7) during progression of cervical cancer and effects of sirtuin inhibitors on growth of cervical cancer cells, *Tumour Biol 36*, 6159-6171.

[254] Wang, H. L., Lu, R. Q., Xie, S. H., Zheng, H., Wen, X. M., Gao, X., and Guo, L. (2015) SIRT7 Exhibits Oncogenic Potential in Human Ovarian Cancer Cells, *Asian Pac J Cancer Prev 16*, 3573-3577.

[255] Zhang, S., Chen, P., Huang, Z., Hu, X., Chen, M., Hu, S., Hu, Y., and Cai, T. (2015) Sirt7 promotes gastric cancer growth and inhibits apoptosis by epigenetically inhibiting miR-34a, *Sci Rep 5*, 9787.

[256] Barber, M. F., Michishita-Kioi, E., Xi, Y., Tasselli, L., Kioi, M., Moqtaderi, Z., Tennen, R. I., Paredes, S., Young, N. L., Chen, K., Struhl, K., Garcia, B. A., Gozani, O., Li, W., and Chua, K. F. (2012) SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation, *Nature 487*, 114-118.

[257] Ford, E., Voit, R., Liszt, G., Magin, C., Grummt, I., and Guarente, L. (2006) Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription, *Genes & development 20*, 1075-1080.

[258] Paredes, S., Angulo-Ibanez, M., Tasselli, L., Carlson, S. M., Zheng, W., Li, T. M., and Chua, K. F. (2018) The epigenetic regulator SIRT7 guards against mammalian cellular senescence induced by ribosomal DNA instability, *J Biol Chem*.

[259] Chen, S. F., Seiler, J., Santiago-Reichelt, M., Felbe, K., Grummt, I., and Voit, R. (2013) Repression of RNA Polymerase I upon Stress Is Caused by Inhibition of RNA-Dependent Deacetylation of PAF53 by SIRT7, *Molecular cell 52*, 303-313.

[260] Blank, M. F., and Grummt, I. (2017) The seven faces of SIRT7, *Transcription 8*, 67-74.

[261] Ashraf, N., Zino, S., Macintyre, A., Kingsmore, D., Payne, A. P., George, W. D., and Shiels, P. G. (2006) Altered sirtuin expression is associated with node-positive breast cancer, *Brit J Cancer 95*, 1056-1061.

[262] De Nigris, F., Cerutti, J., Morelli, C., Califano, D., Chiariotti, L., Viglietto, G., Santelli, G., and Fusco, A. (2002) Isolation of a SIR-like gene, SIR-T8, that is overexpressed in thyroid carcinoma cell lines and tissues, *Brit J Cancer 87*, 1479.

[263] Vakhrusheva, O., Braeuer, D., Liu, Z., Braun, T., and Bober, E. (2008) Sirt7-dependent inhibition of cell growth and proliferation might be instrumental to mediate tissue integrity during aging, *Journal of physiology and pharmacology : an official journal of the Polish Physiological Society 59 Suppl 9*, 201-212.

[264] Vakhrusheva, O., Smolka, C., Gajawada, P., Kostin, S., Boettger, T., Kubin, T., Braun, T., and Bober, E. (2008) Sirt7 increases stress resistance of cardiomyocytes and prevents apoptosis and inflammatory cardiomyopathy in mice, *Circulation research 102*, 703-710.

[265] Vazquez, B. N., Thackray, J. K., Simonet, N. G., Kane-Goldsmith, N., Martinez-Redondo, P., Nguyen, T., Bunting, S., Vaquero, A., Tischfield, J. A., and Serrano, L. (2016) SIRT7 promotes genome integrity and modulates non-homologous end joining DNA repair, *EMBO J 35*, 1488-1503.

[266] Paredes, S., and Chua, K. F. (2016) SIRT7 clears the way for DNA repair, *EMBO J 35*, 1483-1485.

[267] Tong, Z., Wang, M., Wang, Y., Kim, D. D., Grenier, J. K., Cao, J., Sadhukhan, S., Hao, Q., and Lin, H. (2017) SIRT7 Is an RNA-Activated Protein Lysine Deacylase, *ACS Chem Biol 12*, 300-310.

[268] Tong, Z., Wang, Y., Zhang, X., Kim, D. D., Sadhukhan, S., Hao, Q., and Lin, H. (2016) SIRT7 Is Activated by DNA and Deacetylates Histone H3 in the Chromatin Context, *ACS Chem Biol 11*, 742-747.

[269] Jiang, H., Khan, S., Wang, Y., Charron, G., He, B., Sebastian, C., Du, J., Kim, R., Ge, E., Mostoslavsky, R., Hang, H. C., Hao, Q., and Lin, H. (2013) SIRT6 regulates TNF-alpha secretion through hydrolysis of long-chain fatty acyl lysine, *Nature 496*, 110-113.

[270] Du, J., Zhou, Y., Su, X., Yu, J. J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J. H., Choi, B. H., He, B., Chen, W., Zhang, S., Cerione, R. A., Auwerx, J., Hao, Q., and Lin, H. (2011) Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase, *Science 334*, 806-809.

[271] Mathias, R. A., Greco, T. M., Oberstein, A., Budayeva, H. G., Chakrabarti, R., Rowland, E. A., Kang, Y. B., Shenk, T., and Cristea, I. M. (2014) Sirtuin 4 Is a Lipoamidase Regulating Pyruvate Dehydrogenase Complex Activity, *Cell 159*, 1615-1625.

[272] Peng, C., Lu, Z., Xie, Z., Cheng, Z., Chen, Y., Tan, M., Luo, H., Zhang, Y., He, W., Yang, K., Zwaans, B. M., Tishkoff, D., Ho, L., Lombard, D., He, T. C., Dai, J., Verdin, E., Ye, Y., and Zhao, Y. (2011) The first identification of lysine malonylation substrates and its regulatory enzyme, *Molecular & cellular proteomics : MCP 10*, M111 012658.

[273] Park, J., Chen, Y., Tishkoff, D. X., Peng, C., Tan, M., Dai, L., Xie, Z., Zhang, Y., Zwaans, B. M., Skinner, M. E., Lombard, D. B., and Zhao, Y. (2013) SIRT5-

mediated lysine desuccinylation impacts diverse metabolic pathways, *Molecular cell 50*, 919-930.

[274] Tan, M., Peng, C., Anderson, K. A., Chhoy, P., Xie, Z., Dai, L., Park, J., Chen, Y., Huang, H., Zhang, Y., Ro, J., Wagner, G. R., Green, M. F., Madsen, A. S., Schmiesing, J., Peterson, B. S., Xu, G., Ilkayeva, O. R., Muehlbauer, M. J., Braulke, T., Muhlhausen, C., Backos, D. S., Olsen, C. A., McGuire, P. J., Pletcher, S. D., Lombard, D. B., Hirschey, M. D., and Zhao, Y. (2014) Lysine glutarylation is a protein posttranslational modification regulated by SIRT5, *Cell metabolism 19*, 605-617.

[275] Li, L., Shi, L., Yang, S. D., Yan, R. R., Zhang, D., Yang, J. G., He, L., Li, W. J., Yi, X., Sun, L. Y., Liang, J., Cheng, Z. Y., Shi, L., Shang, Y. F., and Yu, W. H. (2016) SIRT7 is a histone desuccinylase that functionally links to chromatin compaction and genome stability, *Nat Commun 7*.

[276] Wang, W. W., Zeng, Y., Wu, B., Deiters, A., and Liu, W. R. (2016) A Chemical Biology Approach to Reveal Sirt6-targeted Histone H3 Sites in Nucleosomes, *ACS Chem Biol 11*, 1973-1981.

[277] Debets, M. F., van Berkel, S. S., Schoffelen, S., Rutjes, F. P., van Hest, J. C., and van Delft, F. L. (2010) Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition, *Chem Commun 46*, 97-99.

[278] Baskin, J. M., Prescher, J. A., Laughlin, S. T., Agard, N. J., Chang, P. V., Miller, I. A., Lo, A., Codelli, J. A., and Bertozzi, C. R. (2007) Copper-free click chemistry for dynamic in vivo imaging, *Proc Natl Acad Sci U S A 104*, 16793-16797.

[279] Gordon, C. G., Mackey, J. L., Jewett, J. C., Sletten, E. M., Houk, K. N., and Bertozzi, C. R. (2012) Reactivity of biarylazacyclooctynones in copper-free click chemistry, *Journal of the American Chemical Society 134*, 9199-9208.

[280] Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Expression and purification of recombinant histones and nucleosome reconstitution, *Methods Mol Biol 119*, 1-16.

[281] Eymery, A., Horard, B., El Atifi-Borel, M., Fourel, G., Berger, F., Vitte, A. L., Van den Broeck, A., Brambilla, E., Fournier, A., Callanan, M., Gazzeri, S., Khochbin, S., Rousseaux, S., Gilson, E., and Vourc'h, C. (2009) A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells, *Nucleic acids research 37*, 6340-6354.

[282] Aramsangtienchai, P., Spiegelman, N. A., He, B., Miller, S. P., Dai, L., Zhao, Y., and Lin, H. (2016) HDAC8 Catalyzes the Hydrolysis of Long Chain Fatty Acyl Lysine, *ACS Chem Biol 11*, 2685-2692.

[283] Kutil, Z., Novakova, Z., Meleshin, M., Mikesova, J., Schutkowski, M., and Barinka, C. (2018) Histone Deacetylase 11 Is a Fatty-Acid Deacylase, *ACS Chem Biol 13*, 685-693.

[284] Schultz, B. E., Misialek, S., Wu, J., Tang, J., Conn, M. T., Tahilramani, R., and Wong, L. (2004) Kinetics and comparative reactivity of human class I and class IIb histone deacetylases, *Biochemistry 43*, 11083-11091.

[285] Mahrez, W., Arellano, M. S., Moreno-Romero, J., Nakamura, M., Shu, H., Nanni, P., Kohler, C., Gruissem, W., and Hennig, L. (2016) H3K36ac Is an Evolutionary

Conserved Plant Histone Modification That Marks Active Genes, *Plant Physiol 170*, 1566-1577.