Synthesis of libraries of lysine-acetylated nucleosomes and their usage for probing bromodomains

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Yandan Zhao

aus Yunnan, China

Tübingen

2023

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	24.01.2024
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Dirk Schwarzer
2. Berichterstatterin:	Prof. Dr. Gabriele Dodt

Declaration of Authorship

I hereby declare that I alone wrote the doctoral work submitted here under the title “Synthesis of libraries of lysine-acetylated nucleosomes and their usage for probing bromodomains”, that I only used the sources and materials cited in the work. I have explicitly marked all material, which has been quoted either literally or by content from the used sources. Parts of this work were done in collaboration with other scientists whose names are stated in the dissertation.

I declare that I adhered to the guidelines set forth by the University of Tübingen to guarantee proper academic scholarship (Senate Resolution 25.05.2000). I declare that these statements are true and that I am concealing nothing. I understand that any false statements can be punished with a jail term of up to three years or a financial penalty.

Tübingen, den

_____

_____

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr. Dirk Schwarzer for offering me the opportunity to do my PhD in his lab. His expertise, encouragement, unwavering support, and invaluable guidance throughout the entire journey of my doctoral research have been extremely helpful to me.

I'm also grateful to Prof. Dr. Gabriele Dodt for being my second supervisor.

I am very thankful to the group of Prof. Dr. Wolfgang Fischle at the King Abdullah University of Science and Technology in Saudi Arabia for providing the ΔH3 and IntC-ΔH2A-ΔH3 nucleosomes.

I would like to thank Nina Pfahler from the group of Prof. Dr. Thilo Stehle and Dr. Michael Braun from the group of Dr. Marcus Hartmann for helping me to purify the BRD4(1/2)-mTagBFP2 in a large scale.

I also want to acknowledge every person and group that has helped me.

I am deeply grateful to all my current and former colleagues: Julia, Sören, Julian, Stefan, Luisa and Katharina. Thank you for the wonderful working atmosphere and a lot of help in all areas. Special thanks to Sören for providing three bromodomains (BAZ2B-mTagBFP, CREBBP-TurboYFP, BRD3(2)-TurboYFP) and a biotin linker peptide, guiding me in the synthesis of H3 tail depsipeptides, introducing me a lot of techniques in the lab, giving me many valuable suggestions, reading and correcting my thesis. Many thanks to Stefan for his guidance and help in gene cloning, protein expression and purification. I am very thankful to Julia, who provided me with valuable guidance in western blotting and offered me tremendous encouragement. Thanks to Julian for support in Alloc deprotection. I also want to thank Katharina for preparing buffers, gels and bacteria.

Finally, I wish to express my heartfelt gratitude to my parents. Your unwavering love and support have been a constant source of strength and inspiration for me.

Abstract

ε-N-acetylation of lysine residues is one of the most prevalent post-translation modifications of histones, often associated with chromatin remodeling and gene transcription. Bromodomains (BRDs) are conserved protein motifs that can recognize and bind acetylated lysine residues. Commonly, the interactions between BRDs and histone acetylation sites are investigated with peptides, omitting any impact of the nucleosome structure.

In this thesis, libraries of acetylated nucleosomes were generated by protein semi-synthesis and probed with a set of recombinant BRDs. Resorting to sortase-mediated ligation established for semi-synthesis of histone H3, 64 depsipeptides permutating the known acetylation site at K4, K9, K14, K18, K23 and K27 were synthesized and ligated with truncate nucleosomal H3. Protein trans-splicing was used to generate 16 nucleosomes with permuted acetylation sites at K5, K9, K13 and K15 of histone H2A.

The nucleosomal libraries were probed with recombinant BRDs derived from BAZ2B, CREBBP, BRD3(2) and the tandem BRD of BRD4(1/2). In general, mono-acetylated nucleosomes showed similar interaction profiles with these BRDs as reported with peptides. In the case of acetylated H3, additive acetylation marks enhanced nucleosomal recruitment of BRDs, and this effect appeared to result from enhanced direct binding rather than charge neutralization modulating tail accessibility. Nucleosomes with multiple acetylation marks on H3 and H2A showed enhanced recruitment of the tandem bromodomain BRD4(1/2), which likely resulted from independent binding events rather than simultaneous binding of the acetylated H3 and H2A tails.

Zusammenfassung

ε-N-Acetylierung von Lysinresten ist eine der häufigsten posttranslationalen Modifikationen an Histonen und wird oft mit Chromatinremodellierung und aktiver Transkription in Verbindung gebracht. Bromodomänen (BRDs) sind konservierte Proteinmotive, die acetylierte Lysinreste erkennen und binden können. Üblicherweise werden die Wechselwirkungen zwischen BRDs und Acetylierungsstellen an Histonen mit Peptiden untersucht, wobei der Einfluss der Nukleosomenstruktur ignoriert wird.

In dieser Dissertation wurden Bibliotheken von acetylierten Nukleosomen durch Proteinsemisynthese erzeugt und mit einer Reihe von rekombinanten Bromodomänen untersucht. Für die Sortase-vermittelten Ligation, die für die Semisynthese von Histone H3 etabliert wurde, wurden insgesamt 64 Depsipeptide synthetisiert, die die bekannten Acetylierungsstellen an K4, K9, K14, K18, K23 und K27 variieren und im Anschluss mit verkürztem nukleosomalen H3 ligiert. Protein-Trans-Splicing wurde verwendet, um 16 Nukleosome mit variierenden Acetylierungsstellen an K5, K9, K13 und K15 des Histons H2A zu erzeugen.

Die nukleosomalen Bibliotheken wurden mit rekombinanten BRDs der Protein BAZ2B, CREBBP, BRD3(2) und der Tandem-BRD von BRD4(1/2) untersucht. Im Allgemeinen zeigten monoacetylierte Nukleosomen ähnliche Interaktionsprofile mit diesen BRDs wie zuvor mit Peptiden berichtet. Im Fall von acetyliertem H3 führten zusätzliche Acetylierungsmarkierungen zu einer verstärkten nukleosomalen Rekrutierung der BRDs, und dieser Effekt schien eher aus einer direkten Bindung als aus der Neutralisation der Ladung zu resultieren, die die Zugänglichkeit der Histone-Tails moduliert. Nukleosomen mit mehreren Acetylierungsmarkierungen an H3 und H2A zeigten eine verstärkte Rekrutierung von BRD4(1/2), die wahrscheinlich auf unabhängige Bindungsereignisse zurückzuführen ist, anstatt auf gleichzeitige Interaktion mit acetylierten H3- und H2A-Tails.

Contents

Abbreviation

ACN	Acetonitrile
Alloc	Allyloxycarbonyl
Ahx	Aminohexanoic acid
Boc	*tert*-Butyloxycarbonyl
BRD	Bromodomain
BSA	Bovine serum albumin
Cy5	Cyanine 5
DCM	Dichloromethane
DIC	*N,N´*-Diisopropylcarbodiimide
DIPEA	*N,N*-Diisopropylethylamin
DMAP	4-Dimethylaminopyridine
DMF	*N,N*-Dimethylformamide
DNA	deoxyribonucleic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EPL	Expressed protein ligation
Fmoc	Fluorenylmethyloxycarbonyl
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b] pyridinium 3-oxide hexafluorophosphate
HEPES	2-[4-(2-Hydroxyethyl) piperazin-1-yl] ethane-1-sulfonic acid
HPLC	High Performance Liquid Chromatography
IPTG	Isopropyl β-d-1-thiogalactopyranoside
LC-MS	Liquid chromatography–mass spectrometry
NCP	Nucleosome core particle
NMP	*N*-Methyl-2-pyrrolidone
NMM	*N*-Methylmorpholine

PAGE	polyacrylamide gel electrophoresis
PyOxim	[Ethyl-cyano-(hydroxyimino)-acetato-O2]-tri-1 pyrrolidinylphosphonium-hexafluorophosphat
PBS	Phosphate-buffered saline
PCL	Phosphopantetheinyl transferase-catalyzed Ligation
PTM	Post-translational modification
PTS	Protein trans-splicing
SDS	Sodium dodecyl sulfate
SML	Sortase-mediated ligation
SPPS	Solid-Phase Peptide Synthesis
TAMRA	Carboxy-tetramethylrhodamin
TEV	Tobacco etch virus
TFA	Trifluoroacetic acid
TIPS	Triisopropylsilane
Tris	2-Amino-2-(hydroxymethyl) propane-1,3-diol

1. Introduction

1.1 Chromatin structure

In eukaryotes, long strands of DNA need to fit within the cell nucleus, possessing only a tiny volume. The linear DNA of each human cell is about 2 m in length, yet the nucleus is only 5 to 20 µm in diameter [1]. To solve this problem, DNA is tightly folded and wrapped around highly conserved proteins known as histones, eventually yielding an up to 10,000-fold condensed DNA-protein complex called chromatin [2].

Chromatin is a complex consisting of DNA, histones, non-histone proteins, and small RNA in the cell nucleus. As the elementary function of chromatin is packaging DNA, it plays an essential role in preventing DNA from damage and is a factor in regulating gene expression, DNA replication, transcription, repair, and cell division [3] [4]. The basic unit of chromatin is the nucleosome core particle (NCP) (Figure 1B), composed of approximately 147 base pairs of DNA wrapped around an octamer of the four core histones H3, H4, H2A, and H2B in 1.67 left-hand superhelical turns [5]. A strong 4-helix bundle (4-HB) interaction between the H3 molecules connects the two H3-H4 dimers to form the (H3-H4)2-tetramer. This tetramer interacts with H2A-H2B dimers via a weaker 4-HB between H2B and H4 [6]. Noteworthily, the tetramer can be viewed as a stable complex of two H3-H4 heterodimers, which feature an interlocking protein shape known as the “handshake” that is very similar to the structure of H2A-H2B heterodimers. The term “histone fold” currently describes this protein’s structural motif [2]. Each histone protein has a globular domain as well as a flexible, unstructured “histone tail” that stick out from the surface of the nucleosome. In addition to core histones, a linker histone H1 exists, which binds the nucleosome at the entry and exit site of DNA. As a result, the DNA is locked into place, and higher-order structures can be formed [7] [8]. The 11 nm fiber with the “beads on a string” conformation is the most fundamental form of such a formation. A higher-order structure is achieved through the formation of the 30 nm Chromatin fiber with the atypical zigzag pattern. The condensed chromosomes are assembled

during mitosis and meiosis [9] [10] [11] (Figure 1A).

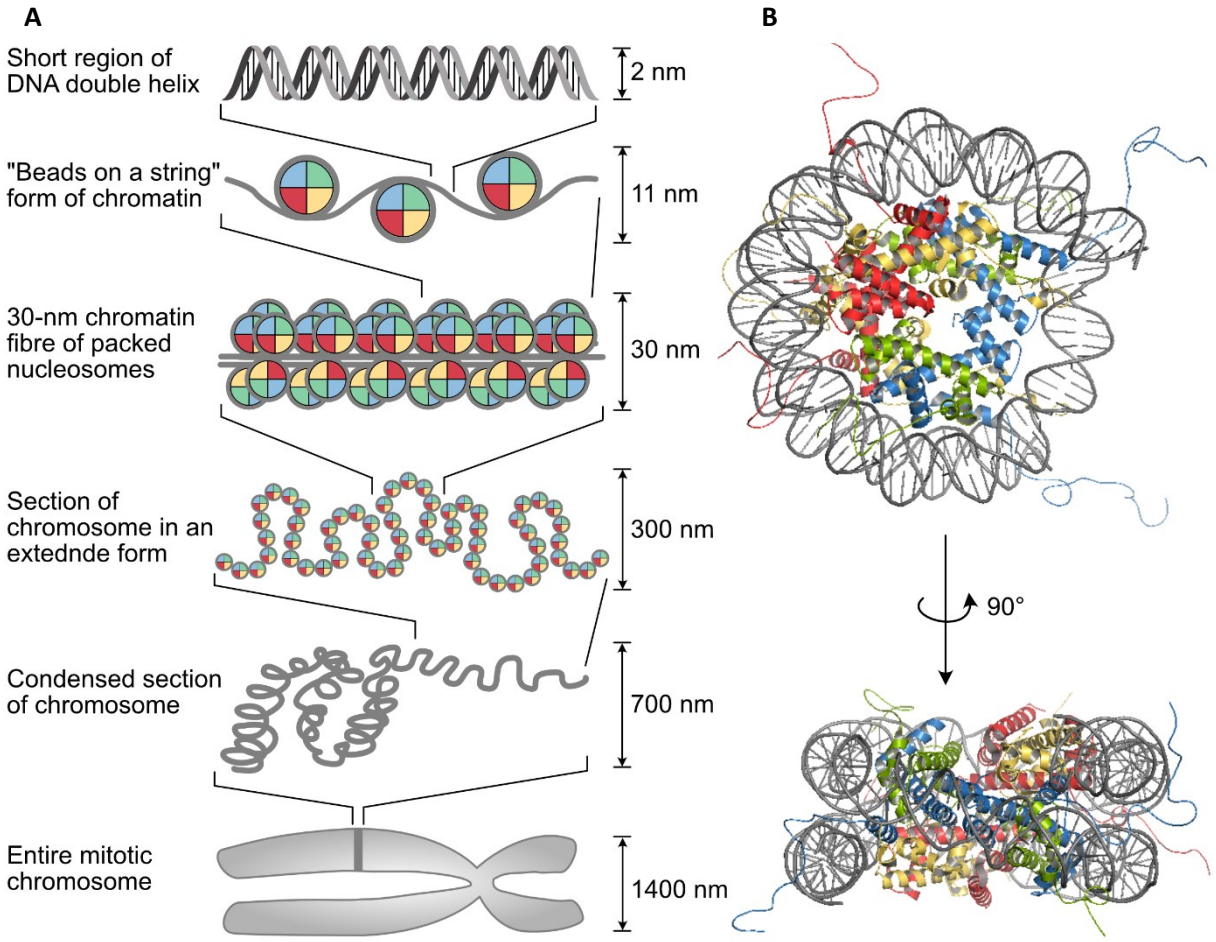


Figure 1. Structure of Chromatin. (A) In eukaryotic cells, during mitosis and meiosis, chromatin condensed as chromosomes, created by a nucleosome-folded, fiber-like structure with a diameter of 30nm. This structure with a so-called “beads on a string” form lets the DNA wrap around histone octamer to conserve DNA storage space (redrawn from [12]). (B) Crystal structure of the nucleosome core particle (DNA is colored in gray, H2A in yellow, H2B in red, H3 in blue, H4 in green) (PDB code: 1KX5).

1.2 Histone post-translational modifications

The histone post-translational modifications (PTMs) that appear in histone globular core domains and their terminal tails can impact gene expression by altering the chromatin structure or recruiting chromatin factors [13]. The histone PTMs, including methylation, acetylation, phosphorylation, ubiquitination, sumoylation, ADP-ribosylation, crotonylation,

and succinylation, have also been considered as epigenetic markers (Figure 2) [14] [15]. Histone modifications are associated with three different types of proteins: 1. “writers”- enzymes that catalyze the formation of PTMs; 2. “readers”- proteins that can recognize and bind specific PTMs through the corresponding reader domains; 3. “erasers”- enzymes that can remove modifications from histones [16]. As the histone PTMs play fundamental roles in many biological processes, particularly DNA damage/repair, mitosis, meiosis, apoptosis, and transcriptional activation/inactivation, dysfunctions of the PTM pathways have been implicated in a number of human diseases, notably cancer, autoimmune disorders, heart disease, Alzheimer’s disease, and Parkinson’s disease [17]. Therefore, the modification pathways of histones could be used as a target for potential treatment of the related diseases.

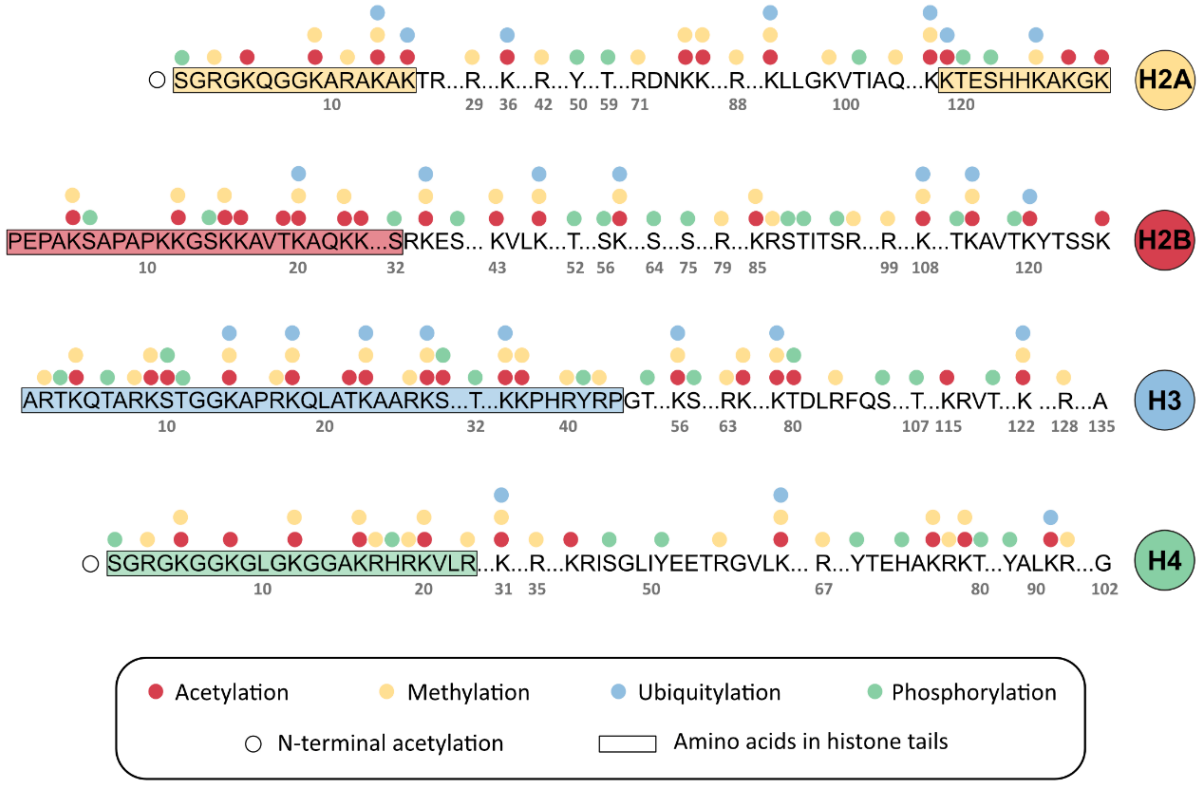


Figure 2. Histones post-translational modifications (PTMs). The four most common PTMs on histones: acetylation, methylation, ubiquitylation, and phosphorylation, are indicated by colored symbols in red, yellow, blue, and green, respectively. Black circles represent the N-terminal acetylation. Amino acids in the histone tails are marked with colored background shading and black borders. Gaps in the sequence are denoted by ellipses. Based on PTMs described in [15] [18] [19].

1.2.1 Histone acetylation and deacetylation

The processes by which an acetyl group is transferred from the cofactor acetyl-coenzyme A to and removed from the ε-amino of lysine side chains is known as histone acetylation and deacetylation. These reactions are catalyzed by two families of enzymes: histone acetyltransferases (HATs) and histone deacetylases (HDACs) [20] [21].

The acetyl group neutralizes the positive charge of the modified lysine residue. As a result, the chromatin structure loosens up as a result of reduced electrostatic connection between histones and the negatively charged DNA backbone [21]. Based on the subcellular localization of HATs, there are two major categories of HATs: type-A and type-B. Type-A HATs are located in the nucleus and are involved in the regulation of gene expression by acetylating nucleosomal histones [22]. Many HATs contain reader domains of acetyllysine, which aid in their ability to detect and attach to acetylated lysine residues on histone substrates. Type-B HATs are cytoplasmic and in charge of acetylating newly produced histones before they are assembled into nucleosomes [23] [24]. These HATs lack a reader domain because their function is to recognize newly generated, unacetylated core histones. For example, newly produced histone H4 is acetylated by type-B HATs at K5 and K12, this type of acetylation is crucial for histone deposition into chromatin before the marks are erased [25]. Type-B HATs are well conserved and show sequence homology with scHat1 of yeast, the original member of this type of HATs [21]. Type-A HATs are further categorized into five families: the GNAT family, the CREBBP/EP300 family, the basal TF family, the MYST family, and the NRCF family [21] [26]. The type-A HATs frequently interact with other proteins in large multiprotein complexes, including other histone-modifying enzymes. These complexes significantly impact the regulation of HAT activity, substrate selectivity, and recruitment [27]. Free scGCN5, for instance, acetylates free histones instead of nucleosomal histones. On the contrary, scGCN5 effectively acetylates nucleosomal histones when it is a part of the so-called SAGA complex [28].

HDACs catalyze the removal of acetyl groups from acetylated histone lysine residues, which reestablish the positive charge of the residue [20]. As a result, histone tails are more tightly

bound to DNA, and local chromatin architecture may be condensed, which may lead to inhibition of gene transcription [22]. Depending on the similarities of sequences and mechanisms, HDACs are divided into four classes: class I, class II, class III, and class IV [27]. Class I HDACs include HDACs 1, 2, 3, and 8, which are most similar to yeast reduced potassium dependency 3 (Rpd 3) proteins [29]. This class is primarily found in the nucleus and is widely expressed. Class II HDACs, corresponding to histone deacetylase 1 (Hda1) in yeast, are further divided into class IIa (containing HDAC 4,5,7 and 9) and class IIb (including HDAC 6 and 10). Except for HDAC6, which is predominantly located in the cytoplasm, class II HDACs shuttle between the nucleus and cytoplasm [30]. Class III deacetylase, known as sirtuin, is analogous to silent information regulator2 (Sir2) and is nicotinamide adenine dinucleotide (NAD^+)-dependent [31]. Class IV HDACs, which only contain HDAC11, are expressed in the brain, testes, and immune cells and have been shown to control the stability of proteins, immunological response, and metabolic illness [30] [32]. Compared to Class III, HDACs of the other three classes have a different catalytic mechanism and require Zn^{2+} as a cofactor [22].

1.2.2 Histone methylation and demethylation

Histone methylation results from the addition of methyl groups from S-adenosyl-l-methionine (SAM) to histone lysine or arginine residues, catalyzed by histone methyltransferases (HMTs). Conversely, histone demethylation is catalyzed by histone demethylases (HDMs) to remove the methyl groups from histones [14]. In histone lysine residues, mono-, di-, or tri-methylation can be found, while histone arginine residues are mono- or di-methylated [33]. Depending on the number and position of methyl groups, histone methylation has different impacts on DNA transcription. For instance, mono-methylation of H3K9, H3K27, H3K79, H4K20, and H2BK5 is correlated with active transcription. By contrast, trimethylation of H3K9, H3K27, and H3K79 is linked to transcriptional silencing [34]. Histone methylation is a crucial factor in controlling DNA transcription, as it either activates or suppresses the recruitment of regulatory proteins to the chromatin [35].

1.2.3 Histone ubiquitination and deubiquitination

The transfer of ubiquitin, a 76 amino acid regulatory protein, to a lysine amino acid in histone core proteins is known as Histone ubiquitination [17]. Deubiquitinating enzymes (DUBs) can reverse the ubiquitination reaction by removing ubiquitin molecules from substrates. Numerous biological processes, including protein degradation, cell-cycle regulation, stress response, protein trafficking, DNA repair, and transcriptional regulation, are associated with ubiquitin. Proteins can be monoubiquitinated or polyubiquitinated, which results from linking additional ubiquitin to one of the seven lysine residues of the previous ubiquitin. However, histone core proteins are mostly monoubiquitinated. Moreover, H2A and H2B are two of the most frequently and highly ubiquitylated proteins in the nucleus [36]. Histone ubiquitination is essential for the regulation of various processes that take place inside the nucleus, such as the preservation of genome stability and transcriptional regulation. For instance, gene repression is closely related to Histone H2A ubiquitination, commonly at lysine 119, which involves Polycomb group repressive complex 1 (PRC1) [37]. Oppositely, the removal of ubiquitin from H2A can be achieved by Ubiquitin-specific peptidase 16 (Usp 16) [38].

1.2.4 Histone phosphorylation and dephosphorylation

Phosphorylation is a reversible PTM, which mainly occurs on serine, threonine, and tyrosine residues [39]. During phosphorylation, the γ-phosphate group of adenosine triphosphate (ATP) is transferred by kinases to the hydroxyl group of the corresponding amino acid side chain by generating adenosine diphosphate (ADP) [14] [21]. The addition of phosphates significantly increases the histone's negative charge, resulting in a more loose association between DNA and histones [17]. Conversely, dephosphorylation catalyzed by phosphatases could strengthen the interaction of DNA with histones. More broadly, histone phosphorylation has been linked to numerous cellular processes, including mitosis, meiosis, DNA repair, apoptosis, and gene expression [39] [40]. For instance, it was demonstrated that the phosphorylation of H3 at S10 and S28 plays an essential role in chromosome condensation and cell-cycle development

during mitosis and meiosis [41] [42]. Additionally, this modification has been associated with transcription initiation in yeast, mammals, and drosophila [17] [43].

1.3 Bromodomains

Bromodomains (BRDs) are one group of epigenetic 'readers', specifically recognizing ε-N-acetylated lysine residues on histone tails. BRDs play a fundamental role in the recruitment of chromatin factors and, consequently, regulate gene transcription [44]. A total of 61 bromodomains were identified in 46 proteins of the human genome [45]. According to their structure and sequence similarity, these bromodomains are divided into eight subfamilies (Figure 3B), which include a variety of transcriptional co-regulators and chromatin modifying enzymes, for instance, HATs (GCN5, PCAF), ATP-dependent chromatin-remodeling complexes (BAZ1B), helicases (SMARCA), SET domain containing methyl-transferases (MLL and ASH1L), transcriptional co-activators (TRIM/TIF1), transcriptional mediators (TAF1), nuclear scaffolding proteins (PB1), and the BET family [46] [47]. BRDs all possess the same conserved fold consisting of a left-handed bundle of four α helices (αZ, αA, αB, and αC), connected by flexible loop regions (ZA and BC loops), which form the acetyl-lysine (Kac) binding site and mediate binding selectivity. Crystal structures showed that the Kac residue is recognized by a central hydrophobic cavity and anchored by a hydrogen bond to a conserved asparagine residue found in the majority of BRDs [45] [48].

1.3.1 Subfamily I of BRDs

The subfamily I consists of acetyl-transferase P300/CBP-associated factor (PCAF) [49], amino-acid synthesis general controller 5-like2 (GCN5L) [50], a transcription factor Fetal Alzheimer antigen (FALZ) [51], and a chromatin remodeling factor cat eye syndrome chromosome region 2 (CECR2) [52], all of them present in the nucleus. The BRD of PCAF has been demonstrated to bind acetylated histone H3 at positions Kac9, Kac14, and Kac36 and histone H4 sites Kac8, Kac16, and Kac36 [45] [53]. The interaction between the BRD of GCN5L2 and mono-/multi-

acetylation sites of histones, as well as H2A (Kac5), H3 (Kac9, Kac14, Kac9-14) and H4 (Kac16, Kac8-16, Kac5-8-12-16) has been also identified [48] [54]. The BRD of FALZ and CECR2 were shown to bind histone H4 (Kac5) and histone H3 (Kac9, Kac14), respectively [45].

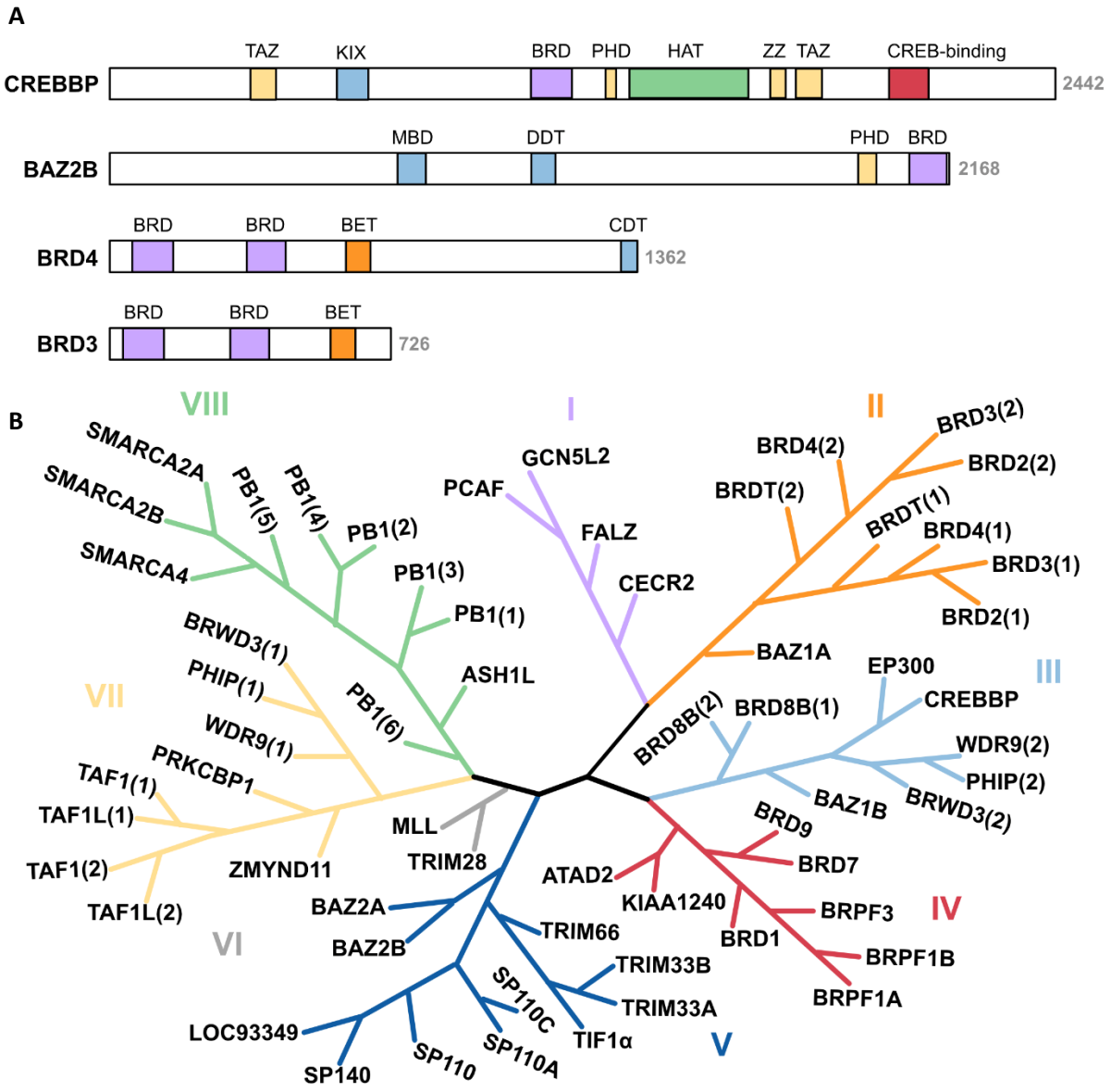


Figure 3. (A) Domain organization of CREBBP, BAZ2B, BRD4, and BRD3. (Zinc finger domains such as transcription adaptor putative zinc finger (TAZ), plant homeodomain zinc finger (PHD), and ZZ-type zinc finger (ZZ) are colored in yellow; bromodomain (BRD) in purple; histone acetyltransferase (HAT) in green; cAMP-responsive element-binding domain (CREB-binding) in red; bromo and extra terminal domain (BET) in orange; other domains such as KID-interacting domain (KIX), methyl-CpG binding domain (MBD), DNA binding homeobox and Different Transcription factors (DDT), and C-terminal domain (CDT) are colored in blue.) (B) Phylogenetic tree of BRDs is based on [45].

1.3.2 Subfamily II of BRDs – the BET family

The subfamily II of BRDs carries the bromo and extra terminal (BET) family, consisting of four proteins: BRD2, BRD3, BRD4 (Figure 3A), and BRDT, sharing a common architecture that includes two N-terminal conserved tandem BRDs (as well as the first bromodomain (BDI) and second bromodomain (BDII)), an extra-terminal (ET) domain and a disparate C-terminal domain (CTD) [55] [56] [57]. The BRD structure forming a hydrophobic cavity acts as the 'reader' of acetylated lysine residues. The ET domain is essential for recruiting specific transcriptional complex proteins [58]. The CTD, which was only found in two of these subfamily members (BRD4 and BRDT), is involved in leading the positive transcription elongation factor (P-TEFb) toward the site of transcription [59]. During mitosis, BET proteins were shown to be recruited to the transcription start sites [60] [61]. In addition, BRD3 showed direct interactions with acetylated lysine residues of the transcription factor GATA1, which plays a vital role in expressing all erythroid and megakaryocyte-specific genes [62].

As the BET proteins control various cellular processes, such as gene transcription, DNA replication, and cell-cycle progression, they are involved in many human diseases, like inflammatory diseases, autoimmune diseases, infections, and cancers [63] [64]. For instance, the expression of several specific cancer-related genes, such as c-MYC, was reported to be deeply associated with BET proteins [65]. BRD3 and BRD4 were shown to interact with WHSC1 and then increase ESR1 transcription, thereby driving resistance to tamoxifen in ER-positive breast cancers [66]. Furthermore, the BRDs of BRD3 and BRD4 are known to part of oncogenic fusion proteins with NUT midline carcinoma (NMC). These BRD3-NUT and BRD4-NUT fusion proteins result from chromosome rearrangements and are the drivers of the global hyperacetylation and suppression of genes required for differentiation [67]. In addition, BRD4 plays an essential role in NF-kB-driven cancers by modulating NF-kb-dependent genes and inhibiting the degradation of Rel A [68]. Moreover, BRD4 is vital for regulating the expression of genes necessary for the transition from the M phase to the early G1 phase [59].

The first and second BRDs of BRD3 and BRD4 were identified to bind to the mono- and multi-

acetylated lysine residues on histone H3 and H4. Notably, BRD3(2) displayed a strong binding affinity to H4-Kac20 with a K_D value of approximately 10.5 µM [46]. Furthermore, multiple acetylated sites may contribute to increased binding affinity for BET proteins compared to monoacetylated histone tails [45]. For instance, the K_D value of BRD4(1) decreased from 810 µM for H4-Kac5 to 2.8 µM for H4-Kac5-8-12-16 [46], meaning these multiple acetylation marks strengthen the affinity by almost 300-fold compared to signal acetylation marks. However, compared with BRD4(1), BRD4(2) showed weaker affinities (K_D =26.6 µM) to the same multi-acetyl sites H4-Kac5-8-12-16 [46], which implies that the first BRD of BRD4 may prefer to be recruited on the acetyl H4 tails. In addition, BRD4(2) showed a slight binding preference for H3-Kac14 with a K_D value of approximately 260 µM [46].

1.3.3 Subfamily III of BRDs

Subfamily III of BRDs is found in the HAT enzymes CREB binding protein (CREBBP) (Figure 3A) and E1A binding protein p300 (EP300), the transcriptional regulator bromodomain containing 8B (BRD8B), the C-terminal domain of the chromatin remodeling factors WD repeat domain 9 (WDR9 domain2), the bromodomain adjacent to zinc finger domain 1B (BAZ1B), the C-terminal domain of the JAK/STAT pathway associated bromodomain-containing protein disrupted in leukemia (BRWD3 domain 2), and the C-terminal domain of the insulin signaling connected pleckstrin homology domain interacting protein (PHIP domain2) [69] [70] [71] [72] [73]. CREBBP/EP300 are regarded as multi-functional transcriptional coactivators involved in a wide range of intracellular processes. For instance, CREBBP/EP300 has been reported to bind to acetylated p53 at position K382ac [53] [74]. Moreover, it has been demonstrated that CREBBP/EP300 binds to various transcription factors and acetylates particular chromatin locations, resulting in chromatin relaxation and consequent transcriptional activation [75] [76]. In addition, CREBBP and EP300 are crucial for the growth of many types of human malignancies [77]. Remarkably, CREBBP was reported to bind to histone H2B (Kac85), histone H3 (Kac14, Kac36, Kac56, Kac9-14) and histone H4(Kac12, Kac20, Kac44) [45] [46].

1.3.4 Subfamily IV of BRDs

Subfamily IV of BRDs includes the transcriptional regulators bromodomain containing 7 (BRD7), bromodomain containing 9 (BRD9), two AAA domain containing protein (ATAD2), KIAA1240 protein (KIAA1240), bromodomain containing protein 1 (BRD1), bromodomain and plant homeodomain (PHD) finger-containing protein 1 (BRPF1), and the bromodomain and PHD finger-containing protein 3 (BRPF1) [78] [79] [80]. The BRD of BRD7 showed weak binding affinity to histone H3 (Kac9, Kac14) and histone H4 (Kac8, Kac12, Kac16) [81].

1.3.5 Subfamily V of BRDs

Subfamily V of BRDs is composed of Bromodomain adjacent to zinc finger domain protein 2A (BAZ2A) and protein 2B (BAZ2B) (Figure 3A), transcriptional repressor tripartite motif-containing 66 (TRIM66), the tripartite motif-containing 33A (TRIM33A) and 33B (TRIM33B), the transcriptional regulator transcriptional intermediary factor 1 (TIF1α), the transcriptional regulators nuclear auto-antigen Sp-100 (SP100), nuclear auto-antigen Sp-110platA (SP110A) and Sp-110B (SP110B), SP140 nuclear body protein (SP140) and the SP140-like protein (LOC93349) [82] [83] [84] [46]. Intriguingly, the BRD of BAZ2B binds explicitly to acetylated histone H3 at position K14 with single-digit micromolar affinity [45].

1.3.6 Subfamily VI of BRDs

Subfamily VI has two members: histone methyl-transferase myeloid/lymphoid or mixed-lineage leukemia (MLL) and the transcriptional co-regulator tripartite motif-containing 28 (TRIM28) [85] [86]. No acetylated lysine of any protein has been identified as the binding site of family VI BRDs.

1.3.7 Subfamily VII of BRDs

Subfamily VII of BRDs contains transcription initiators TAF1 RNA polymerase II TATA box binding protein (TBP)-associated factor (TAF1L), the transcriptional repressing zinc finger MYND domain-containing 11 protein (ZMYND11), N-terminal BRDs of the chromatin remodeling factors WD repeat domain 9 (WDR9 domain 1), the insulin signal related pleckstrin homology domain interacting protein (PHIP domain 1), and the JAK/STAT pathway related bromodomain-containing protein disrupted in leukemia (BRWD3 domain 1) [46] [71] [87] [88]. The BRD of TAF1 was reported to interact with p53 (Kac373/Kac382), which promotes recruiting to the p21 promoter [89]. Furthermore, TAF1 was demonstrated to bind with mono- and multi-acetylated lysine residues on histones H3 (Kac14, Kac9-14) and H4 (Kac8, Kac12, Kac16, Kac5-12, Kac8-16, Kac5-8-12-16) [45] [46].

1.3.8 Subfamily VIII of BRDs

The subfamily VIII of BRDs comprises the methyl-transferase ash1 (absent, small, or homeotic)-like (ASH1L), chromatin remodeling factors SWI/SNF-related matrix-associated actin-dependent regulator of chromatin a2 (SMARCA2), SWI/SNF-related matric-associated actin-dependent regulator of chromatin a4 (SMARCA4), and the Polybromo 1 (PB1) [46] [90] [91] [92]. The BRD of PB1 was shown to bind to several mono-acetylated lysins on histone H3 tail (Kac4, Kac9, Kac14, Kac18, Kac23) with high binding affinity [93] [94].

1.4 Chemical biology strategies for peptide and protein modifications

The study of biomolecules, with a particular focus on protein structure and function, plays a crucial role in the discovery of novel drugs. To access these biomolecules, various chemical biology strategies have been developed. One of the basic strategies is to use native chemical ligation (NCL), expressed protein ligation (EPL), or other chemoselective ligation strategies for proteins and peptides to introduce modified fragments like noncanonical amino acids

containing peptides, fluorescent tags or biophysical probes into the proteins of interest.

1.4.1 Thiol capture ligation

One early example of chemoselective ligation strategies is the thioester ligation (Figure 4A), forming between an unprotected peptide having a C-terminal thioacid and an N-terminal bromoacetic acid-containing peptide. This ligation was applied by Schnolzer and Kent to synthesize the 99-residue HIV-1 protease analog from two about 50-residue unprotected peptides [95]. The HIV-1 protease analogs were also synthesized through thioether ligation (Figure 4B), which involves a C-terminal cysteamine containing unprotected peptide and an N-terminal bromoacetic acid-carrying peptide [96]. Furthermore, a thioether bond can also form between a sulfhydryl group bearing protein and a terminal ε-maleimidocaproyl group carrying synthetic peptide by a Michael addition. This thioether-forming ligation was known as maleimidocaproyl (MIC) ligation (Figure 4C), which was extensively utilized to modify Ras proteins with different functionalities, such as fluorescent labels, contributing to the elucidation of Ras proteins involved molecular signaling networks [97] [98] [99].

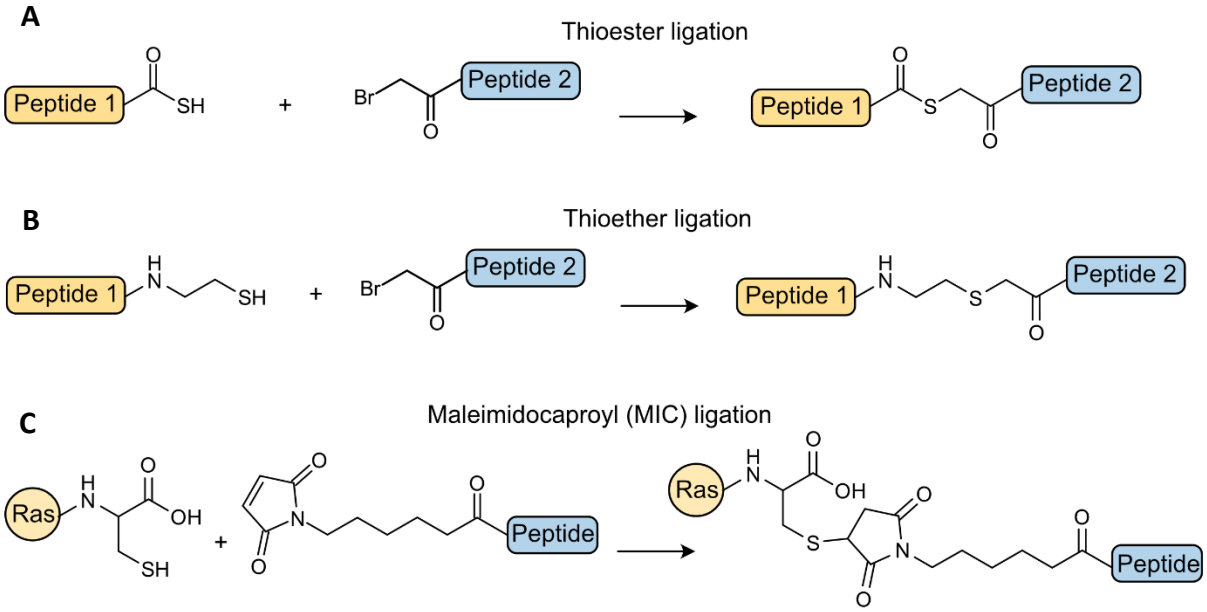


Figure 4. Selected thiol capture ligations, such as (A) thioester ligation, (B) thioether ligation, and (C) maleimidocaproyl (MIC) ligation, have been applied for the semi-synthesis of some functional small proteins.

1.4.2 Native Chemical Ligation

Inspired by the thioester ligation, the native chemical ligation strategy was developed, which is by now the most widespread method of protein semi-synthesis. The NCL forms a native peptide bond between two segments of a C-terminal thioester containing unprotected peptide and a peptide with N-terminal cysteine [100]. This ligation occurs at neutral pH in two steps: firstly, an initial covalent thioester-linked intermediate is generated through a reversible thiol-thioester exchange, which in a second step undergoes a spontaneous, rapid and irreversible intramolecular S→N transfer forming an amide bond between two peptides [101] (Figure 5). As the NCL is a chemoselective ligation strategy, which can link two highly functionalized molecules without the usage of protective groups, NCL is now a powerful technique for the modification and synthesis of proteins, which also surpasses the limitation in lengths of the synthetic peptides achieved by solid-phase peptide synthesis (SPPS) [102] [103].

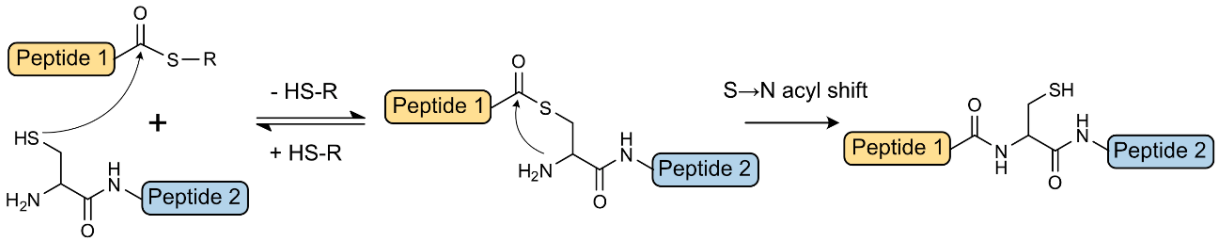


Figure 5. Mechanism of Native chemical ligation.

1.4.3 Expressed Protein Ligation

Like an extension of the NCL method, expressed protein ligation (EPL) is a protein semi-synthesis strategy, which ligates an N-terminal cysteine-containing peptide or protein to a C-terminal thioester carrying protein to form a native peptide bond [104]. EPL relies on a natural phenomenon known as protein self-splicing. An internal domain of the protein, termed intein, is cleaved from the precursor protein, and the C- and N-terminal external proteins, referred to as extein, on either side are coupled [105]. The first step of the intein-mediated protein splicing

is the N→O/S acyl shift at the N-terminal serine or cysteine of the intein to form an ester or thioester bond between intein and N-extein. The ester or thioester is transferred to the downstream N-terminal serine or cysteine of the C-extein, resulting in a branched intermediate. In the final step, cyclization occurs at the N-terminal asparagine residue of the intein, leading to the formation of succinimide at the C-terminus of the intein, which releases the intein, and through an O/S→N acyl shift, the N and C exteins are linked by a native bond [101] [106] [107] (Figure 6A).

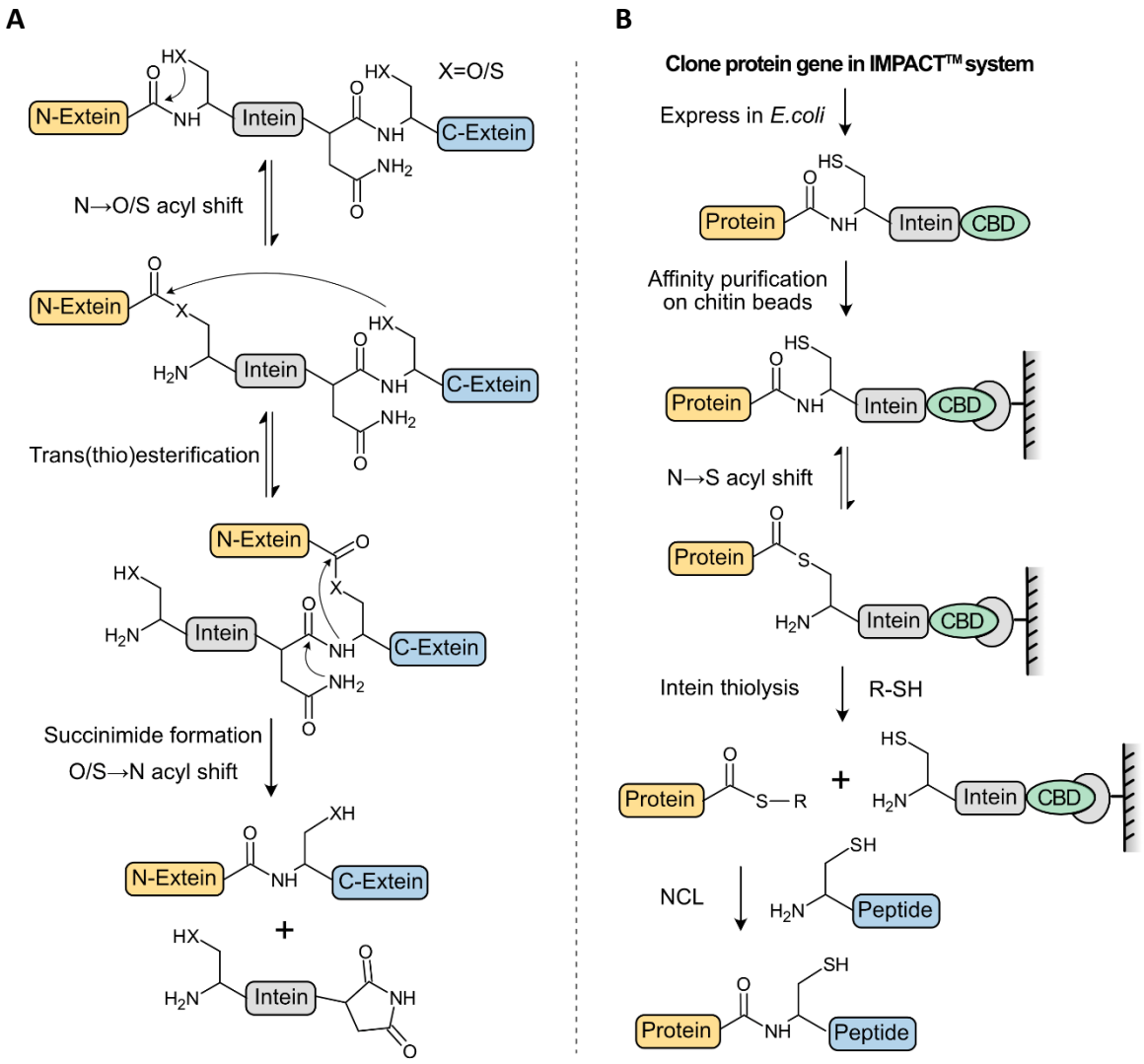


Figure 6. (A) Mechanism of the expressed protein ligation. X represents either the oxygen present in the side chain of serine or the sulfur present in cysteine. (B) Intein-mediated purification with an affinity chitin binding tag (IMPACT™) system.

One of the commercial applications, EPL involved, is the intein-mediated purification with an

affinity chitin binding tag (IMPACT™) system (Figure 6B), which exploits the self-cleavage ability of protein splicing to release the target protein from the affinity tag [108]. Like a mimic of N-extein, the target protein is fused to the intein's N-terminal part, and the intein's C-terminus is fused with a chitin-binding domain (CBD) for affinity purification of the target protein on a chitin resin. To prevent undesired ligation between the target protein and CBD, the C-terminal asparagine of the intein is eliminated, and CBD lacks an N-terminal cysteine or serine residue. After the fusion protein is immobilized on the chitin resin, the target protein is released from the column by the intein undergoing an on-column self-cleavage by adding excess thiols, like 1,4-dithiothreitol (DTT), β-mercaptoethanol (β-ME) or cysteine, at low temperatures. The eluted protein thioester can undergo the NCL and thus form a new ligation product by conjugating with an N-terminal cysteine containing synthetic peptide [109]. The IMPACT™ system, which overcomes the limitation of using protease cleavage, is a powerful tool in protein semi-synthesis.

1.4.4 Protein-trans splicing

As with the EPL, the protein-trans splicing (PTS) method relies on a similar natural phenomenon: protein self-splicing. However, unlike EPL, PTS does not demand the formation of a thioester. Instead, it relies on a specific type of intein: the split inteins, which is split into two domains, namely N-intein (IntN) and C-intein (IntC), which spontaneously reassemble into a functional form, resulting in the splicing of its flanking sequences (N- and C-exteins) [105]. PTS is a convenient and potent tool in protein semi-synthesis by linking the recombinant fusion protein with a functional synthetic peptide [110]. To utilize the PTS approach, the split intein requires a small N- or C-terminal fragment, which can be accessed by SPPS and then ligated with the recombinant protein of interest.

In contrast to the mechanism of EPL, the mechanism of PTS needs an initial assembly of IntN and IntC as an additional first step. The mechanism of PTS (Figure 7) comprises the following five steps: (I) IntN and IntC assemble into an active form. (II) a N→O/S acyl shift occurs at the

N-terminal serine or cysteine residue of the intein, forming an ester or thioester bond between IntN and N-extein. (III) through the trans(thio)esterification between the serine or cysteine residue of C-extein and C-terminal residue of N-extein, these two exteins splice together. (IV) the C-terminal asparagine residue of intein undergoes self-cyclization to form a succinimidyl moiety, resulting in release of the intein. (V) an O/S→N acyl shift occurs at serine or cysteine forming an amide bond to link these two exteins [97] [111] [112].

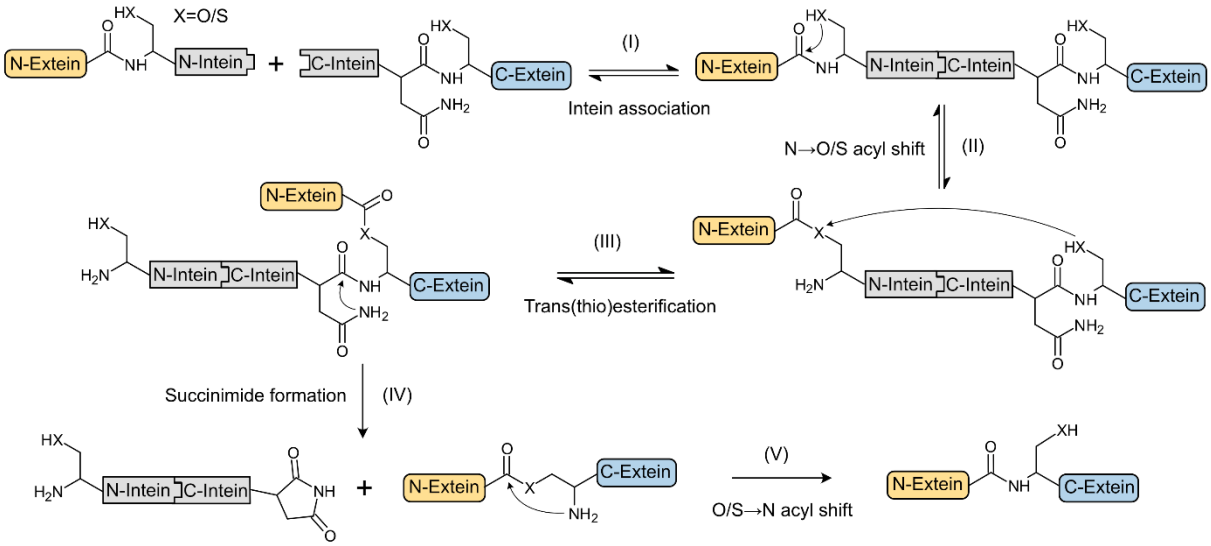


Figure 7. Mechanism of the Protein trans-splicing. X represents either the oxygen present in the side chain of serine or the sulfur present in cysteine.

One of the extensively utilized split inteins is the M86 mutant of the Ssp DnaB intein, derived from the gene of DNA helicase B of Synechocystis sp. strain PCC6803 [113]. This mini-intein comprises a synthetic 11 amino acids containing N-terminal intein fragment (IntN) and a recombinant C-terminal fragment carrying 143 amino acids (IntC) [114]. It is important to note that incorporating the sequence serine-isoleucine-glutamic acid (SIE) to the asparagine residue of the IntC contributes to a high efficiency of PTS. Compared to the unevolved split intein, this mini-intein shows an approximately 60-fold increased rate in the PTS reaction with higher splicing yields of up to 90% [113]. With these improvements, this intein allows efficient N-terminal labeling of proteins.

1.4.5 Sortase-mediate ligation

Only a few enzymes are able to perform chemoenzymatic ligations of peptides and proteins. The transpeptidase sortase A (SrtA) of *Staphylococcus aureus* is one of them and probably the most commonly used for this purpose. [115]. Surface proteins of gram-positive bacteria play a vital role in absorbing nutrients, escaping host immunity, and sticking to infection sites [116]. SrtA, inserted into the bacterial cell wall, anchors a large number of those surface proteins to the cell wall by recognizing a highly conserved sorting motif Leu-Pro-Xxx-Thr-Gly (LPXTG, where X is any amino acid) in these proteins and cleaving the peptide bond between threonine and glycine residues to form a thioester bond between the cysteine residue of SrtA and threonine residue of the surface protein [117]. The thioester intermediate is then attacked by the α-amino group of the pentaglycine unit of the peptidoglycan to form a stable amide bond. As a result, the target protein is anchored to the cross-bridge on the cell wall peptidoglycan precursor (Figure 8A). Any peptide or protein carrying at least one N-terminal glycine can be substituted for the peptidoglycan. Even though one glycine is frequently sufficient for the SML, additional glycines are reported to benefit the reaction's efficiency [118]. Based on this, Sortase A can ligate peptides or proteins containing a sorting motif and an N-terminal oligo glycine unit (Figure 8B). This sortase A-dependent ligation strategy is known as sortase-mediated ligation (SML).

The wild-type SrtA, composed of 206 amino acids, contains a membrane-spanning region at N-terminus. However, a shortened SrtA, which retains the enzymatic center without the N-terminal membrane-anchoring motif, can be expressed and purified in high yields and soluble active form from E.coil [117] [119]. This truncated SrtA is one of the most commonly exploited enzymes for SML. However, the SrtA catalyzed reaction is reversible, causing low ligation yields and restricting its broader applications. This drawback can be effectively solved by using depsipeptide substrates. In the depsipeptide, the Thr-Gly amide bond in the sorting motif is substituted by an ester bond, leading to a hydroxyacetyl byproduct, which cannot serve as an N-glycine reaction partner driving the reaction equilibrium towards the ligation product [120]. In addition, a few other approaches have been reported to improve SML ligation yields. For

instance, a β-hairpin secondary structure around the sorting motif is formed by incorporating a tryptophan zipper, which prevents the reverse reaction [121]. By extending the sorting motif with additional glycine and histidine, the released GGH fragment can be sequestered by Ni^{2+} complexation, which drives the ligation to completion [122]. Another example involves the proximity-based sortase-mediated ligation (PBSL), in which the target protein is expressed with the sorting motif LPXTG, which links to the short peptide SpyTag. The partner protein of SpyTag is SpyCatcher, which is fused with SrtA and immobilized on an affinity resin. During the purification, the SpyCatcher and SrtA fusion protein can catch the target protein. With the addition of Ca^{2+} and a peptide containing N-terminal triglycine, the ligation is initiated, and the product is eluted. This methodology is demonstrated to enhance the ligation efficiency to over 95% [123].

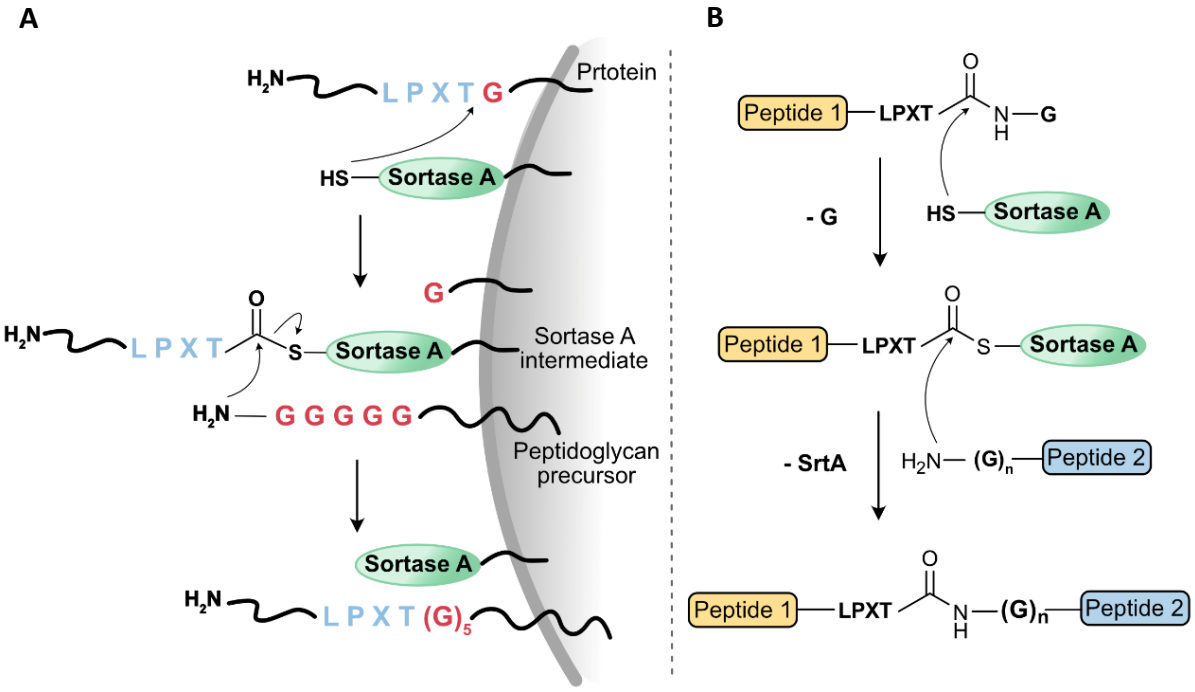


Figure 8. (A) Sortase A in Gram-positive bacteria can anchor surface proteins to the cell wall (redraw from [118]). (B) Sortase-mediated ligation.

A further limitation of SrtA-catalyzed ligation is its strong calcium dependency, which makes SrtA hard to utilize in low Ca^{2+} concentrations and with the presence of Ca^{2+} binding compounds. The Ca^{2+} ion is reported to bind three glutamic acid residues: Glu171 in the β6/β7 loop, as well as Glu108 and Glu105 in the β3/β4 loop, which contribute to stabilizing the β6/β7

loop for a more advantageous position for interaction with the LPXTG sorting motif. Thus, low calcium ion concentration would reduce ligation efficiency. This problem can be addressed by inserting a similar salt bridge into the β3/β4 loop to generate the SrtA mutants, such as SrtA (E105K/E108A) and SrtA (E105K/E108Q), which show the Calcium-independent catalytic activity with slower catalytic kinetics compared to the wild-type SrtA in the presence of Ca^{2+} [124].

Currently, SML is employed for various purposes, such as selective modifications of antibodies, protein circularization, protein labeling, and covalent immobilization of proteins onto solid supports [125]. Notably, SML was demonstrated as a powerful tool for the histone H3 semi-synthesis. Since the sorting motif LPXTG is very similar to the APATG sequence located at positions 29 to 33 at the junction of the globular fold and the tail of histone H3. For instance, Pelaz et al. have successfully applied SML to assemble numerous modified H3 tail peptides on the truncated H3 nucleosomes, generating libraries of H3 nucleosomes in combination with methylation, acetylation, and phosphorylation marks [126] [127].

1.4.6 Phosphopantetheinyl Transferase-Catalyzed Ligation

Phosphopanteteinyl transferases (PPTases) are crucial for modifying fatty acid synthase, polyketide synthases, and nonribosomal peptide synthetases with the 4'-phosphopantetheine (Ppant) cofactor [128]. During the biosynthetic pathways, these transferases can convert the acyl or peptidyl carrier proteins (CPs) from their inactive apo-forms to active holo-forms by transferring the Ppant moiety from coenzyme A (CoA) to an invariant serine residue of CPs [129] (Figure 9A). As a result, the PPant cofactor is covalently ligated with the CP through a phosphoester bond. Due to the irreversibility and stability of this ligation and the relaxed substrate tolerance of PPTases, phosphopantetheinyl transferase-catalyzed Ligation (PCL) was developed for site-specific protein modification with CoA-conjugated probes such as fluorophores or peptides [130] [131]. However, the CP domain, containing 80 to 120 amino acids, is considerably big, which may limit the application of this approach. To address this

problem, several short peptides were discovered as substitutes for CP domains. These peptides, such as tags S6 (GDSLSWLLRLLN) (Figure 9B) and A1 (GDSLDMLEWSLM), each 12 residues in length, as well as an 11-residue peptide with the sequence (DSLEFIASKLA), are identified as effective substrates for the PPTases, such as Sfp and AcpS, catalyzed ligation [132] [133]. It should be noted that the phosphopantetheinylation of S6 catalyzed by Sfp is 442-fold faster than the reaction of Acps. On the contrary, the catalytic efficiency of A1 ligation catalyzed by Sfp is 30-fold slower than the reaction of Acps. Therefore, S6/Sfp and A1/Acps were applied in an orthogonal protein modification strategy, which allows the labeling and imaging of receptor proteins with different tags on the same cell surface [132]. In addition, based on the different catalytic modes and substrate selectivity of Sfp and SrtA, these two enzymes were exploited for orthogonal protein semi-synthesis to assemble three fragments in a one-pot reaction [134].

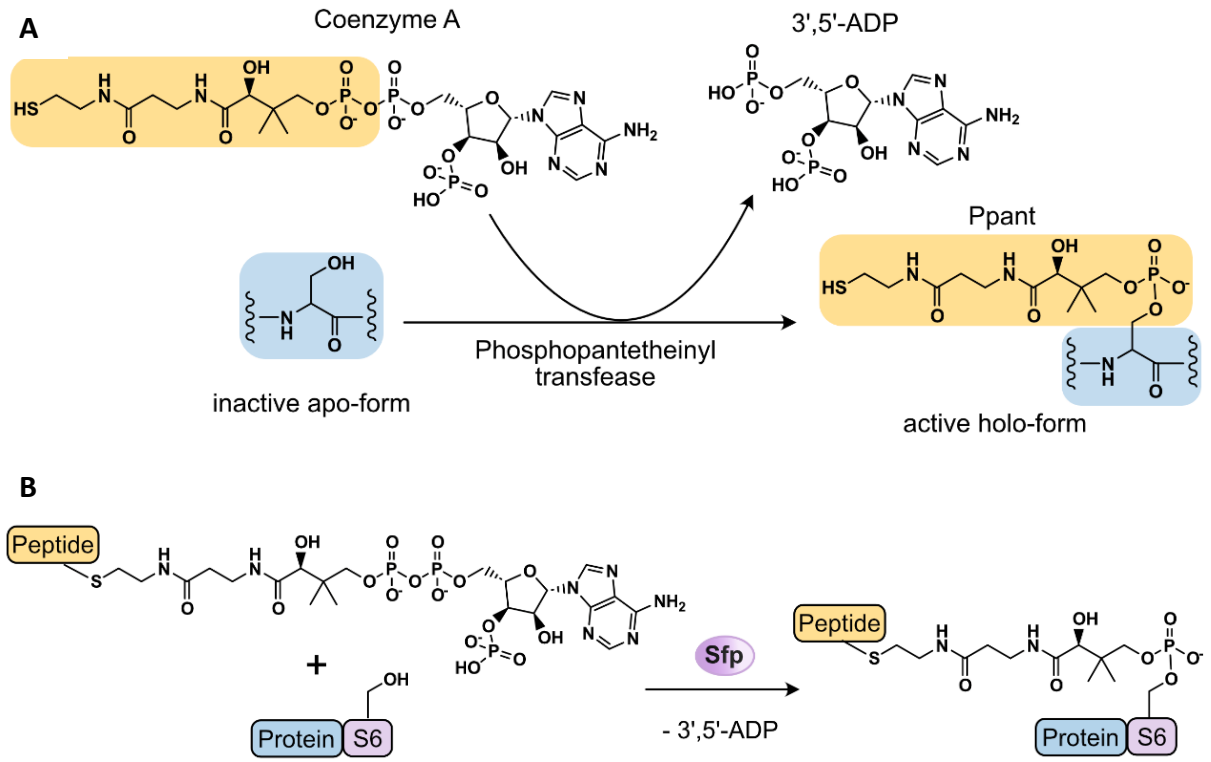


Figure 9. (A) Phosphopanteteinyl transferases (PPTases) transfer a 4'-phosphopantetheine (Ppant) moiety from coenzyme A (CoA) to a serine residue of acyl or peptidyl carrier proteins (CPs), resulting in the conversion of the CPs from their inactive apo-forms to active holo-forms. (B) PPTases Sfp-catalyzed ligation.

1.5 Aims of the study

The primary objective of this thesis was to investigate the impact of different mono- or multi-acetylated lysine residues on the histone H3 tail on the recruitments of selected BRDs including BAZ2B, CREBBP, BRD3(2) and BRD4(1/2). In order to investigate the effect of the nucleosomal context, numerous acetylated H3 nucleosomes need to be generated. Considering that the synthesis of acetylated nucleosome variations in high quantity is laborious, libraries of nucleosomes need to be assembled and established in a straightforward manner. Furthermore, to investigate the recruitments of BRDs in relation to different mono- or multi-acetylation sites on H2A tails, and to further explore the potential crosstalk between acetylated H2A and H3, a large number of acetyl-H2A-H3 nucleosome variants should to be generated and analyzed in the same way.

The aims of the study can be summarized as follows:

(1) Synthesis of 64 acetyl-H3-tail-peptides;
(2) Establishment of acetyl-H3 nucleosomal libraries;
(3) Synthesis of 16 acetyl-H2A-tail-peptides;
(4) Establishment of acetyl-H2A-H3 nucleosomal libraries;
(5) Expression and purification of BRD4(1/2) protein;
(6) Screening these acetylated nucleosomal libraries with fluorescent protein-fused BRDs for investigating their impacts on the BRDs recruitments;
(7) Investigate if observed effect on BRDs recruitment are recapitulated by histone peptides as well.

2. Results

The recruitment of chromatin readers to histones was mainly studied with modified synthetic peptides derived from histone tails. However, the chromatin environment may influence the interaction of histone PTMs with their reader proteins. Hence, nucleosomes represent more physiological probes for investigating these interactions. However, the individual synthesis of numerous nucleosomes with large numbers of PTMs is time- and material-consuming. To overcome this drawback, a high throughput nucleosome semi-synthesis strategy was developed by Pelaz et al., which applied chemoselective ligation strategies to introduce libraries of modified histone tail peptides into preassembled nucleosomes carrying N-terminal truncated histones. Through this nucleosome semi-synthesis strategy, Pelaz et al. have successfully synthesized libraries of nucleosomes carrying different methylation and phosphorylation patterns, aiming to investigate the recruitment of Heterochromatin protein 1 (HP1) [126] [127]. Inspired by this, the strategy of ligation-ready nucleosomes was utilized in this thesis to establish libraries of acetylated nucleosomes and to investigate the effects of the acetylation patterns on the recruitments of BRDs.

2.1 Acetyl-H3 nucleosomal libraries

2.1.1 Assembly strategy of acetyl-H3 nucleosomal libraries

As mentioned in the introduction, histone H3 contains an APATG motif at the interface (residues 29 to 33) between the tail and globular fold, which is very close to the SrtA sorting motif LPXTG. Therefore, sortase-mediated ligation can be applied to establish acetylated H3 nucleosomes by linking the synthesized acetylated H3 tail peptides (residues 1-32) to the ligation-ready truncated H3 (ΔH3) nucleosomes (residues 33-135) (Figure 10). The synthetic acetyl-H3 tail peptides require a C-terminal sorting motif to allow SML to proceed, causing one mutation (A29L) in the final assembled acetylated nucleosome. Notably, to increase the

efficiency of SML, an ester bond substituted the amide bond between threonine and glycine in the sorting motif [126] [127]. The ΔH3 nucleosomes, as start materials, were provided by the group of Prof. Dr. Wolfgang Fischle at King Abdullah University of Science and Technology. The ΔH3 nucleosomes consist of an octamer reconstituted from recombinant H2A, H2B, H4, and ΔH3. The reconstituted histone octamer was wrapped by 147 bp of the Widom 601 DNA sequence with additional biotin and three restriction sites (HindIII, EcoRI, and EcoRV) at the 5' end and a Cy5 fluorophore at the 3' end. This biotin-labeled DNA sequence allows the immobilization of nucleosomes on streptavidin-coated plates with high binding affinity. The three restriction sites can be used to cleave the immobilized nucleosomes from the plates. Cy5 allows the visualization and quantification of immobilized nucleosomes.

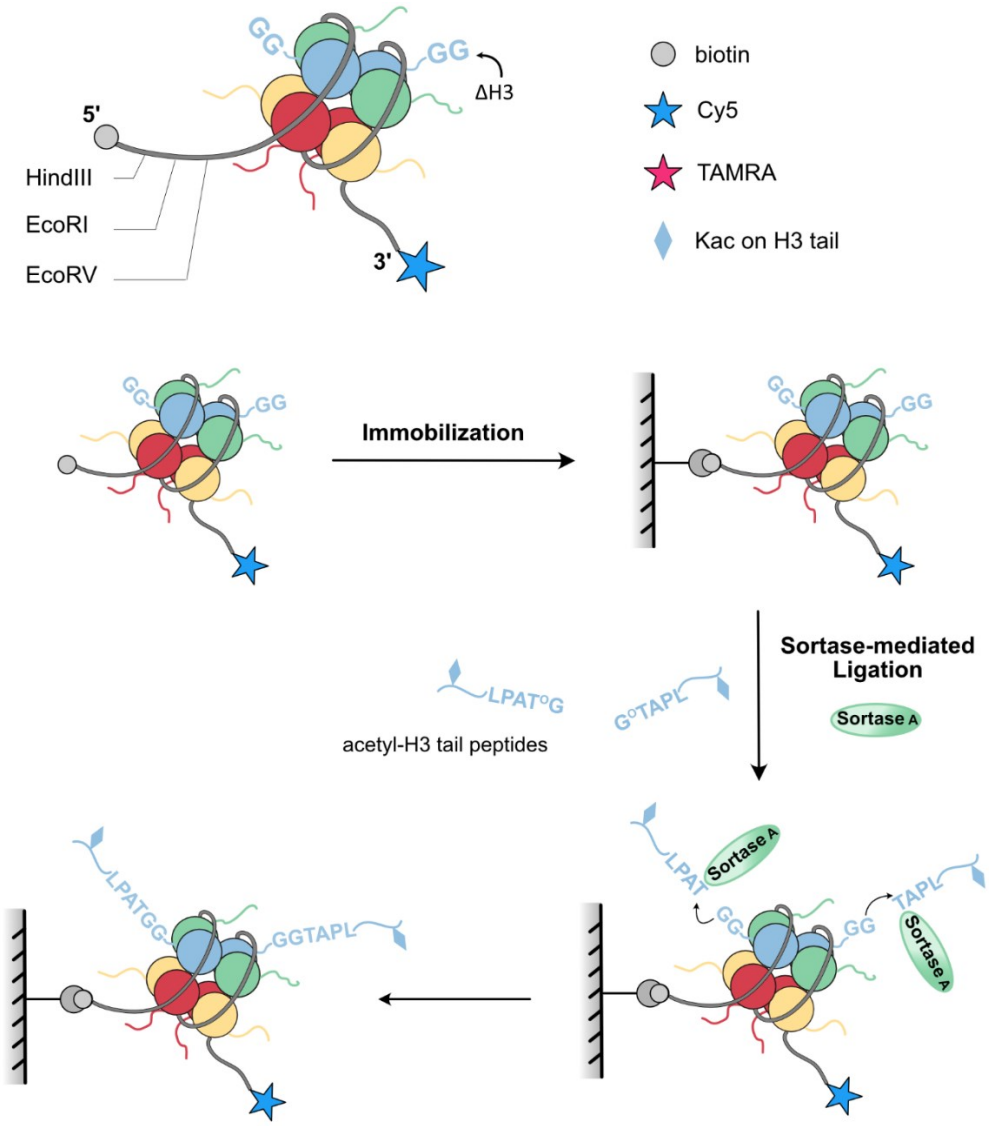


Figure 10. Structure of ligation-ready ΔH3 nucleosome and assembly strategy of acetyl-H3 nucleosomal

libraries. The ligation-ready ΔH3 nucleosome contains full-length H2A (yellow), H2B (red), H4 (green), and truncated H3 (residues 33-135) (blue), which are assembled as an octamer and wrapped by 147 bp 601 Widom DNA containing three restriction sites (HindIII, EcoRI, EcoRV) as well as biotin at 5' end and a Cy5 fluorophore at 3' end. The ΔH3 nucleosome can be immobilized by a biotin tag on the streptavidin-coated plate and then carry out the sortase-mediated ligation (SML) with synthetic acetyl-H3 tail depsipeptides (residues 1-32, A29L). As a result, acetyl-H3 nucleosomes (A29L) were generated on the surface.

2.1.2 Synthesis of acetyl-H3 tail peptides

This thesis focused on one of the most common PTMs: the acetylation of the ε-amino group of lysine. Therefore, six lysine acetylation sites on the histone H3 tail, namely Kac4, Kac9, Kac14, Kac18, Kac23, and Kac27, were chosen to generate 63 variants of acetylated H3 tail peptides. As a control, the unmodified H3 tail peptide was synthesized. As mentioned, all synthesized H3 tail peptides contain the C-terminal sorting motif LPATG to allow sortase-mediated ligation with the truncated histone H3. Noteworthy, these H3 tail peptides are depsipeptides with substitution of the amide bond between threonine and glycine by an ester bond, which makes the SML reaction irreversible and therefore increases the reaction yield. The sequences of the synthesized histone H3 tail peptides are shown in Table 1.

Table 1. Acetyl-H3 tail peptide sequences. Acetylated Lysin (Kac) is marked in red. The sorting motif LPATG is colored in blue. ^{O} represents the ester bond between threonine and glycine.

No.	Name	Sequence
1	H3	**ARTKQTARKSTGGKAPRKQLATKAARKSLPAT^{O}G**
2	H3-Kac4	**ARTKacQTARKSTGGKAPRKQLATKAARKSLPAT^{O}G**
3	H3-Kac9	**ARTKQTARKacSTGGKAPRKQLATKAARKSLPAT^{O}G**
4	H3-Kac14	**ARTKQTARKSTGGKacAPRKQLATKAARKSLPAT^{O}G**
5	H3-Kac18	**ARTKQTARKSTGGKAPRKacQLATKAARKSLPAT^{O}G**
6	H3-Kac23	**ARTKQTARKSTGGKAPRKQLATKacAARKSLPAT^{O}G**
7	H3-Kac27	**ARTKQTARKSTGGKAPRKQLATKAARKacSLPAT^{O}G**
8	H3-Kac4-9	**ARTKacQTARKacSTGGKAPRKQLATKAARKSLPAT^{O}G**
9	H3-Kac4-14	**ARTKacQTARKSTGGKacAPRKQLATKAARKSLPAT^{O}G**
10	H3-Kac4-18	**ARTKacQTARKSTGGKAPRKacQLATKAARKSLPAT^{O}G**
11	H3-Kac4-23	**ARTKacQTARKSTGGKAPRKQLATKacAARKSLPAT^{O}G**
12	H3-Kac4-27	**ARTKacQTARKSTGGKAPRKQLATKAARKacSLPAT^{O}G**
13	H3-Kac9-14	**ARTKQTARKacSTGGKacAPRKQLATKAARKSLPAT^{O}G**
14	H3-Kac9-18	**ARTKQTARKacSTGGKAPRKacQLATKAARKSLPAT^{O}G**

15	H3-Kac9-23	ARTKQTARKacSTGGKAPRKQLATKacAARKSLPAT^{O}G
16	H3-Kac9-27	ARTKQTARKacSTGGKAPRKQLATKAARKacSLPAT^{O}G
17	H3-Kac14-18	ARTKQTARKSTGGKacAPRKacQLATKAARKSLPAT^{O}G
18	H3-Kac14-23	ARTKQTARKSTGGKacAPRKQLATKacAARKSLPAT^{O}G
19	H3-Kac14-27	ARTKQTARKSTGGKacAPRKQLATKAARKacSLPAT^{O}G
20	H3-Kac18-23	ARTKQTARKSTGGKAPRKacQLATKacAARKSLPAT^{O}G
21	H3-Kac18-27	ARTKQTARKSTGGKAPRKacQLATKAARKacSLPAT^{O}G
22	H3-Kac23-27	ARTKQTARKSTGGKAPRKQLATKacAARKacSLPAT^{O}G
23	H3-Kac4-9-14	ARTKacQTARKacSTGGKacAPRKQLATKAARKSLPAT^{O}G
24	H3-Kac4-9-18	ARTKacQTARKacSTGGKAPRKacQLATKAARKSLPAT^{O}G
25	H3-Kac4-9-23	ARTKacQTARKacSTGGKAPRKQLATKacAARKSLPAT^{O}G
26	H3-Kac4-9-27	ARTKacQTARKacSTGGKAPRKQLATKAARKacSLPAT^{O}G
27	H3-Kac4-14-18	ARTKacQTARKSTGGKacAPRKacQLATKAARKSLPAT^{O}G
28	H3-Kac4-14-23	ARTKacQTARKSTGGKacAPRKQLATKacAARKSLPAT^{O}G
29	H3-Kac4-14-27	ARTKacQTARKSTGGKacAPRKQLATKAARKacSLPAT^{O}G
30	H3-Kac4-18-23	ARTKacQTARKSTGGKAPRKacQLATKacAARKSLPAT^{O}G
31	H3-Kac4-18-27	ARTKacQTARKSTGGKAPRKacQLATKAARKacSLPAT^{O}G
32	H3-Kac4-23-27	ARTKacQTARKSTGGKAPRKQLATKacAARKacSLPAT^{O}G
33	H3-Kac9-14-18	ARTKQTARKacSTGGKacAPRKacQLATKAARKSLPAT^{O}G
34	H3-Kac9-14-23	ARTKQTARKacSTGGKacAPRKQLATKacAARKSLPAT^{O}G
35	H3-Kac9-14-27	ARTKQTARKacSTGGKacAPRKQLATKAARKacSLPAT^{O}G
36	H3-Kac9-18-23	ARTKQTARKacSTGGKAPRKacQLATKacAARKSLPAT^{O}G
37	H3-Kac9-18-27	ARTKQTARKacSTGGKAPRKacQLATKAARKacSLPAT^{O}G
38	H3-Kac9-23-27	ARTKQTARKacSTGGKAPRKQLATKacAARKacSLPAT^{O}G
39	H3-Kac14-18-23	ARTKQTARKSTGGKacAPRKacQLATKacAARKSLPAT^{O}G
40	H3-Kac14-18-27	ARTKQTARKSTGGKacAPRKacQLATKAARKacSLPAT^{O}G
41	H3-Kac14-23-27	ARTKQTARKSTGGKacAPRKQLATKacAARKacSLPAT^{O}G
42	H3-Kac18-23-27	ARTKQTARKSTGGKAPRKacQLATKacAARKacSLPAT^{O}G
43	H3-Kac4-9-14-18	ARTKacQTARKacSTGGKacAPRKacQLATKAARKSLPAT^{O}G
44	H3-Kac4-9-14-23	ARTKacQTARKacSTGGKacAPRKQLATKacAARKSLPAT^{O}G
45	H3-Kac4-9-14-27	ARTKacQTARKacSTGGKacAPRKQLATKAARKacSLPAT^{O}G
46	H3-Kac4-9-18-23	ARTKacQTARKacSTGGKAPRKacQLATKacAARKSLPAT^{O}G
47	H3-Kac4-9-18-27	ARTKacQTARKacSTGGKAPRKacQLATKAARKacSLPAT^{O}G
48	H3-Kac4-9-23-27	ARTKacQTARKacSTGGKAPRKQLATKacAARKacSLPAT^{O}G
49	H3-Kac4-14-18-23	ARTKacQTARKSTGGKacAPRKacQLATKacAARKSLPAT^{O}G
50	H3-Kac4-14-18-27	ARTKacQTARKSTGGKacAPRKacQLATKAARKacSLPAT^{O}G
51	H3-Kac4-14-23-27	ARTKacQTARKSTGGKacAPRKQLATKacAARKacSLPAT^{O}G
52	H3-Kac4-18-23-27	ARTKacQTARKSTGGKAPRKacQLATKacAARKacSLPAT^{O}G
53	H3-Kac9-14-18-23	ARTKQTARKacSTGGKacAPRKacQLATKacAARKSLPAT^{O}G
54	H3-Kac9-14-18-27	ARTKQTARKacSTGGKacAPRKacQLATKAARKacSLPAT^{O}G
55	H3-Kac9-14-23-27	ARTKQTARKacSTGGKacAPRKQLATKacAARKacSLPAT^{O}G
56	H3-Kac9-18-23-27	ARTKQTARKacSTGGKAPRKacQLATKacAARKacSLPAT^{O}G
57	H3-Kac14-18-23-27	ARTKQTARKSTGGKacAPRKacQLATKacAARKacSLPAT^{O}G
58	H3-Kac4-9-14-18-23	ARTKacQTARKacSTGGKacAPRKacQLATKacAARKSLPAT^{O}G

59	H3-Kac4-9-14-18-27	**ARTKacQTARKacSTGGKacAPRKacQLATKAARKacSLPAT^{O}G**
60	H3-Kac4-9-14-23-27	**ARTKacQTARKacSTGGKacAPRKQLATKacAARKacSLPAT^{O}G**
61	H3-Kac4-9-18-23-27	**ARTKacQTARKacSTGGKAPRKacQLATKacAARKacSLPAT^{O}G**
62	H3-Kac4-14-18-23-27	**ARTKacQTARKSTGGKacAPRKacQLATKacAARKacSLPAT^{O}G**
63	H3-Kac9-14-18-23-27	**ARTKQTARKacSTGGKacAPRKacQLATKacAARKacSLPAT^{O}G**
64	H3-Kac4-9-14-18-23-27	**ARTKacQTARKacSTGGKacAPRKacQLATKacAARKacSLPAT^{O}G**

The synthesis of the acetylated histone H3 peptides started with the formation of the ester bond between resin-bound glycolic acid serving as glycine substitute and threonine using Steglich esterification [135]. In this reaction, esters are formed from anhydrides in the presence of the 4-dimethylaminopyridine (DMAP) catalyst. The synthesis scheme is shown in Figure 11. The threonine was activated by N, N'-Diisopropylcarbodiimid (DIC) in water- and oxygen-free DCM at 0 °C to form the symmetric anhydride. After coupling glycolic acid to TentaGel® HL RAM resin with activator HATU, the esterification of threonine anhydride and glycolic acid was carried out on the solid phase, which supported the isolation of reaction intermediates. After esterification, a small amount of the product was cleaved off the resin and analyzed by LC-MS. The chromatogram and mass spectrum (Figure 12) demonstrated that the short depsipeptide T^{O}G was successfully synthesized in high yield.

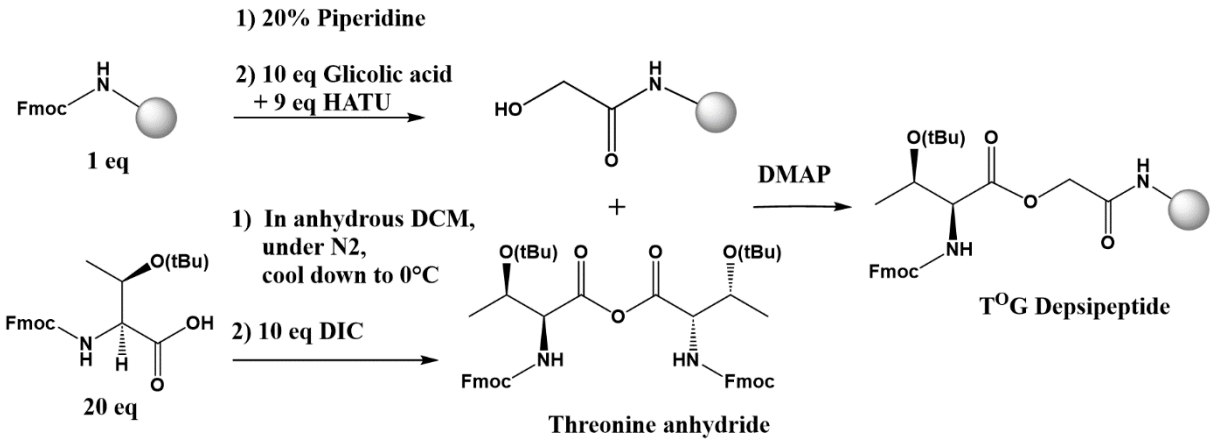


Figure 11. Scheme of T^{O}G depsipeptide formation.

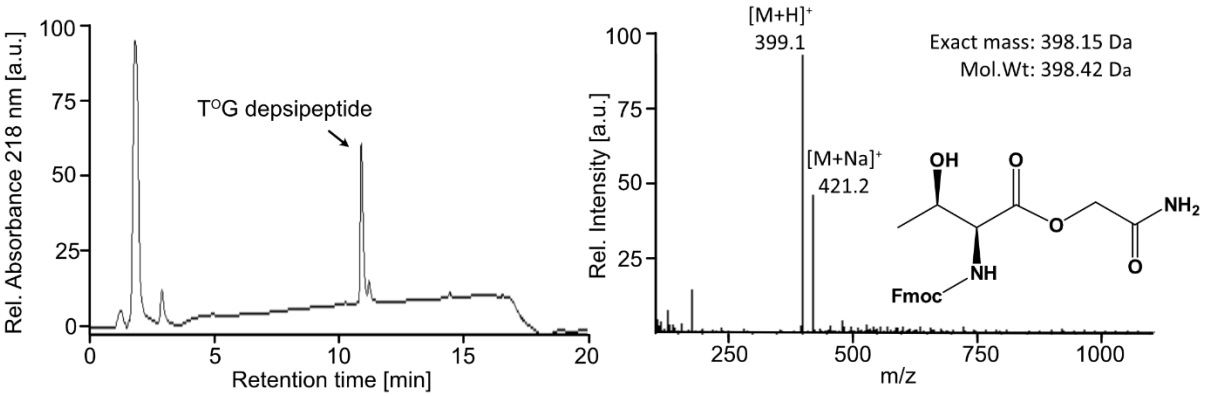


Figure 12. LC-MS results of T^{O}G depsipeptide.

The couplings of the subsequent amino acids were carried out by standard SPPS on an automated peptide synthesizer (Figure 13A). It was reported that the utilization of three dipeptide building blocks (Fmoc-Ala-Thr(psiMe, Mepro)-OH (AT), Fmoc-Gly-(Dmb)Gly-OH (GG), and Fmoc-Gln(Trt)-Thr(psiMe, Mepro)-OH (QT)) is necessary for the synthesis of H3 tail depsipeptides with different PTMs [126] [127]. The pseudoprolines can minimize the aggregation of protected peptides during the synthesis. In this project, instead of the N-alkylated dipeptide GG, the pseudoproline Fmoc-Ser(tBu)-Thr(psiMe, Mepro)-OH (ST) was utilized. These three pseudoprolines AT, ST, and QT (Figure 13B) were applied for the synthesis of acetylated H3 tail peptides, resulting in improved coupling efficiency. The products of peptide synthesis were analyzed by LC-MS and analytical HPLC. These results (Figure 13C and Appendixes 8.1) showed that the synthesis of H3 tail peptides was successful. The purities of these peptides were determined by integrating UV absorption of analytical HPLC chromatograms, resulting in a purity of over 98% for all H3 tail peptides. This level of purity ensures that the synthetic H3 tail peptides are of sufficient quality and suitable for the continuation of the subsequent SML reaction.

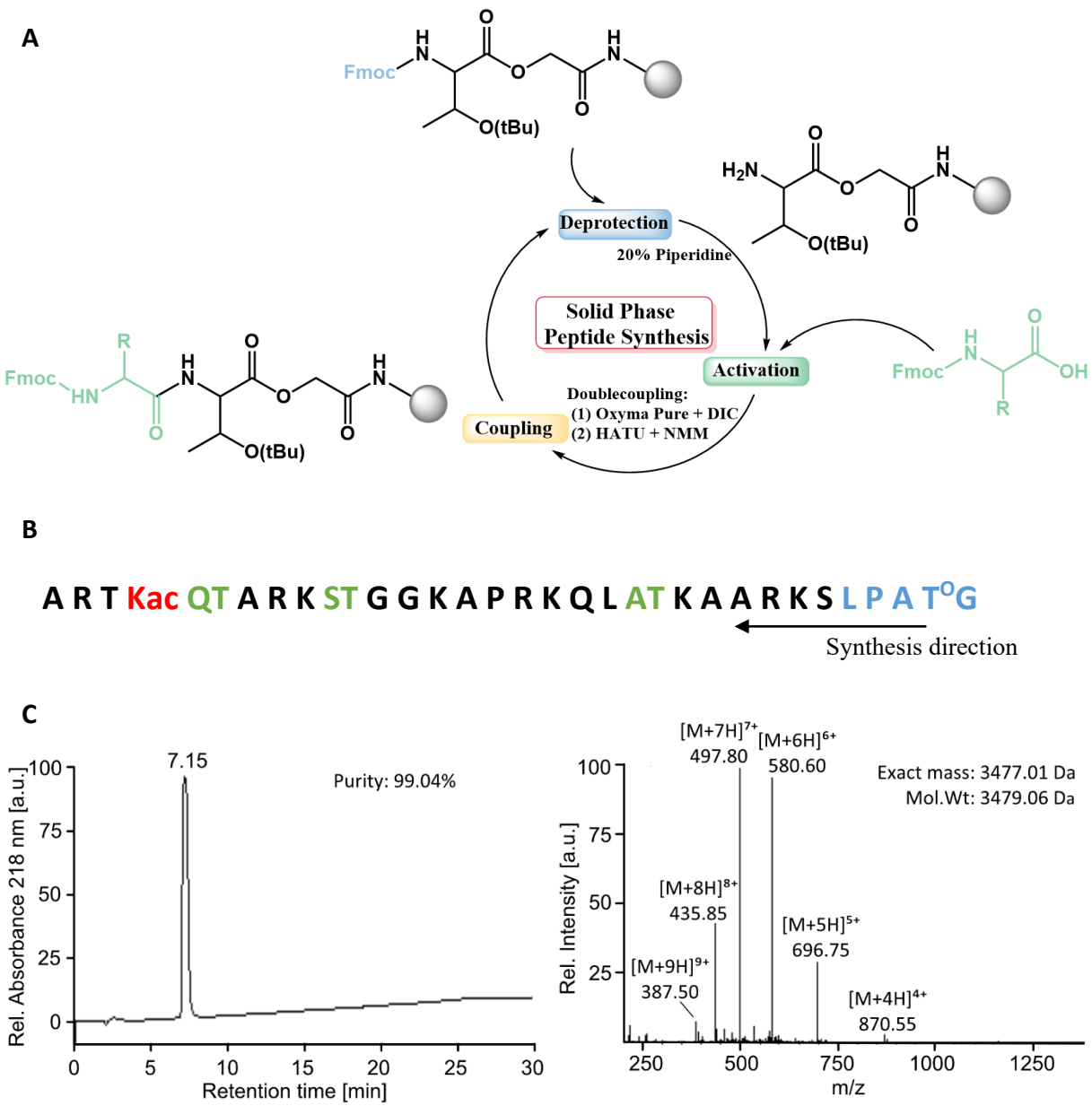


Figure 13. (A) Scheme of solid-phase peptide synthesis in a circle. (B) Sequence of Histone H3 tail peptide No.2 (H3-Kac). Pseudoprolines AT, ST, and QT, colored in green, were coupled manually. (C) Analytical HPLC and LC-MS results of histone H3 tail peptide No.2 (H3-Kac4) as a representative example.

2.1.3 Establishment of acetyl-H3 nucleosomes

Before performing the SML reaction, the biotin-containing ligation-ready ΔH3 nucleosomes were firstly immobilized in streptavidin-coated 96-well plates. This immobilization step facilitates the subsequent SML reactions and pull-down assays, as it allows for convenient removal of excess reactants, catalysts, and further resulting by-products. Additionally, the

solid support is compatible with the different buffer conditions required for the binding assays. The SML reactions were carried out in presence of Ca^{2+} under neutral conditions with 3 µM wild type sortase A and 6 µM depsipeptides at 37 °C while shaking at 400 rpm for 2 hours [126] [127]. After that, the nucleosomes were released and analyzed by Western blot. The results are shown in Figure 14. Compared to the control sample of ΔH3 nucleosomes, a new distinct upper band appeared in the product sample, corresponding to the full-length acetylated histone H3. The intensities and areas of the upper and lower bands were analyzed by the image Lab software package, resulting in 86±6% of conversion for the H3-Kac4 nucleosomes. This result indicated that the SML is an effective strategy to incorporate the acetylated H3 tail depsipeptides into the ΔH3 nucleosomes. Notably, since the recognition sorting motif of wild-type sortase A is LPXTG, these synthetic acetylated histone H3 protein contains Leucin instead of Alanine at the 29th position on its tail.

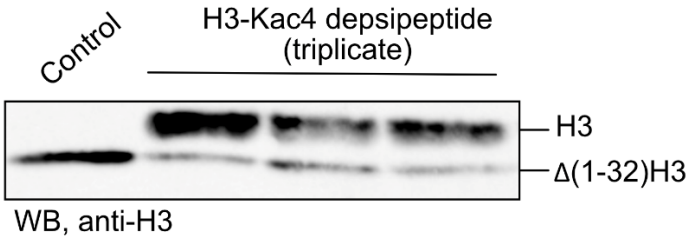


Figure 14. The Sortase-mediated ligation between ΔH3 nucleosomes and H3 tail depsipeptide H3-Kac4 was analyzed by Western blot.

2.2 Pull-down assay of acetyl-H3 nucleosomal libraries with BRDs

After the completion of SML reactions, the library of 64 individual acetylated nucleosomes was available for further BRD binding assays. To investigate the recruitment of BRDs to the nucleosomal library, three BRDs from different subfamilies: BAZ2B from subfamily V, CREBBP from subfamily III, and BRD3(2) from subfamily II, were selected for interaction profiling. The three BRDs constructs were provided by Dr. Sören Kirchgäßner at the Interfaculty Institute of Biochemistry, University of Tübingen [136] [137]. The BAZ2B BRD was fused with mTagBFP, while CREBBP and BRD3(2) were fused with TuroYFP (Figure 15). These fused fluorescent proteins allow the measurement of BRD recruitments to the nucleosomes by determining the

fluorescence intensities. N-terminal His6-Tag and C-terminal Strep-Tag II enable affinity purification of these recombinantly expressed BRDs. The His6-Tag can be removed after the purification by the thrombin cleavage.

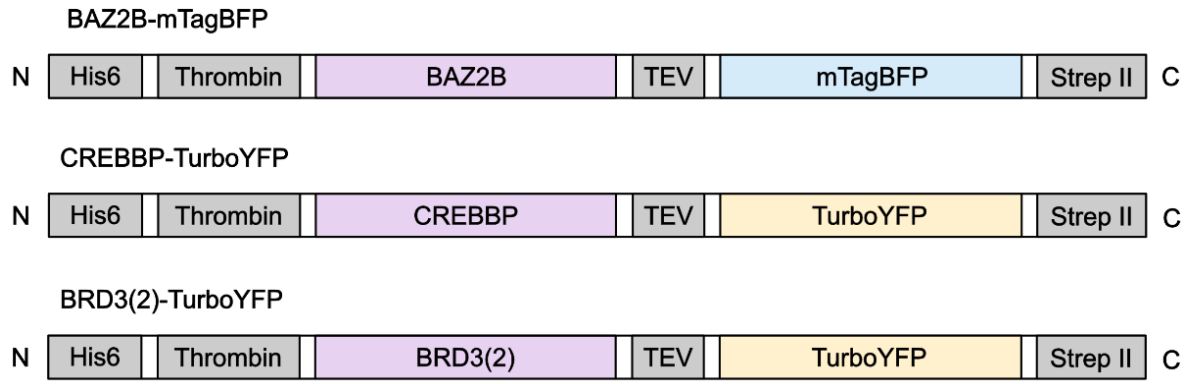


Figure 15. Illustrations of BAZ2B-mTag BFP, CREBBP-Turbo YFP, and BRD3(2)-TurboYFP constructs.

After the establishment of immobilized acetylated H3 nucleosomal libraries, The pull-down assay screening was performed using these fluorescence-labeled BRDs in three replicates. The fluorescence intensity of mTagBFP or TuroYFP was measured using a plate reader to analyze the binding of corresponding BRDs to acetylated nucleosome variants. Simultaneously, the Cy5 signal was detected to normalize the BRD signals to the amounts of immobilized nucleosomes. The mean of the normalized fluorescence signal, obtained by dividing the intensity of mTag BFP or TurboYFP by the intensity of Cy5, served as indicator of the binding of BRDs to different acetyl-H3 nucleosome variations.

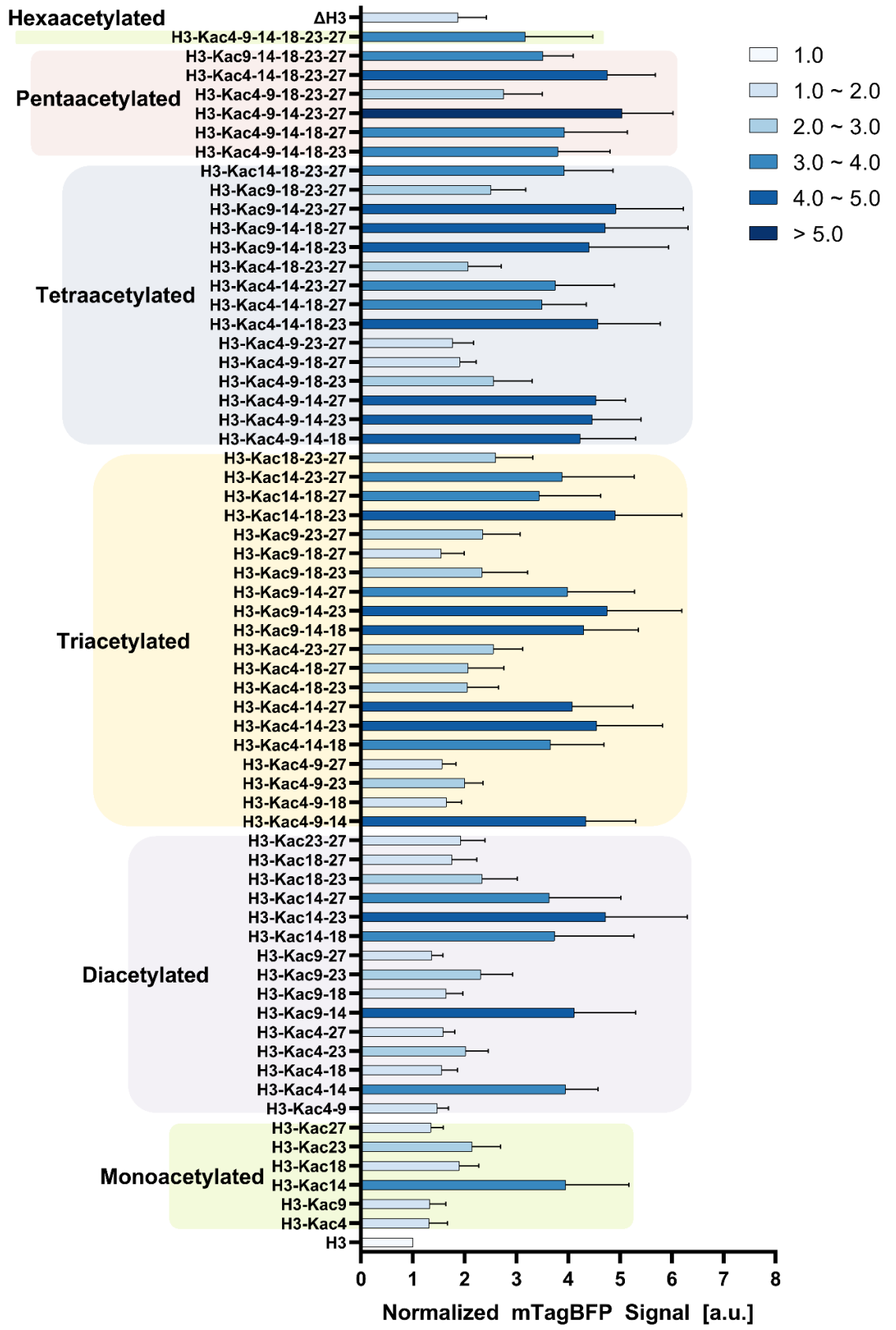


Figure 16. Pull-down assay results of acetyl-H3 nucleosomal libraries with BAZ2B-mTagBFP. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

The results of the pull-down assay with BAZ2B-mTagBFP are shown in Figure 16. According to the number of acetylation marks, the acetyl-H3 nucleosome variations are divided into six groups, from monoacetylated to hexaacetylated nucleosomes. In all groups, nucleosomes containing H3-Kac14 showed pronounced higher normalized mTagBFP signals than other acetylated variants, indicating binding preference of BAZ2B to H3-Kac14 containing nucleosomes. Slightly enhanced binding of BAZ2B to nucleosomes containing H3-Kac18 or H3-Kac23 was further observed with mTagBPF signals, which were 1.9- or 2.1-fold higher than the signal retained on the unmodified nucleosomes. Furthermore, the normalized mTagBFP intensity of BAZ2B increased from 3.9 for H3-Kac14 to 4.7 for H3-Kac14-23, indicating that the diacetylation mark Kac14-23 may facilitate the binding of BAZ2B compared to the single acetylation mark Kac14. In addition, for tri-, tetra-, and penta-acetylated nucleosomes, the highest normalized mTagBPF signals on average were detected for the nucleosomes containing H3-Kac14-18-23, H3-Kac9-14-23-27, and H3-Kac4-9-14-23-27, all of which include the Kac14 and Kac23 modifications on the H3 tail.

Compared to the findings with BAZ2B-mTagBFP, the pull-down assays of acetyl-H3 nucleosomal libraries with CREBBP-TurboYFP (Figure 17) showed a trend of normalized TurboYFP signal intensity that increased with the number of acetylation marks. For the mono-acetylated nucleosomes, the strongest intensity of normalized TurboYFP signals were detected with the H3-Kac14 containing nucleosomes, which were 2.6-fold stronger compared to the unmodified nucleosome. Compared to the normalized TurboYFP intensity of H3-Kac14 monoacetylated nucleosomes, the normalized TurboYFP signals were increased when an additional acetylation site, Kac4 or Kac9, was included. The normalized TurboYFP intensity was enhanced from 2.6 for H3-Kac14 containing nucleosomes to 5.7 or 5.1 for the nucleosomes carrying H3-Kac4-14 or H3-Kac9-14. Furthermore, the highest normalized TurboYFP intensities of tri-, tetra-, and penta-acetylated groups were observed for the nucleosomes containing H3-Kac9-14-23, H3-Kac4-9-14-27, and H3-Kac4-9-14-23-27 with respectively 7.0-, 7.8-, and 7.6-fold enrichments compared to the intensity of unmodified nucleosomes. These observations indicate that the multiple acetylated nucleosomes are more effective in recruiting CREBBP than single acetylated nucleosomes.

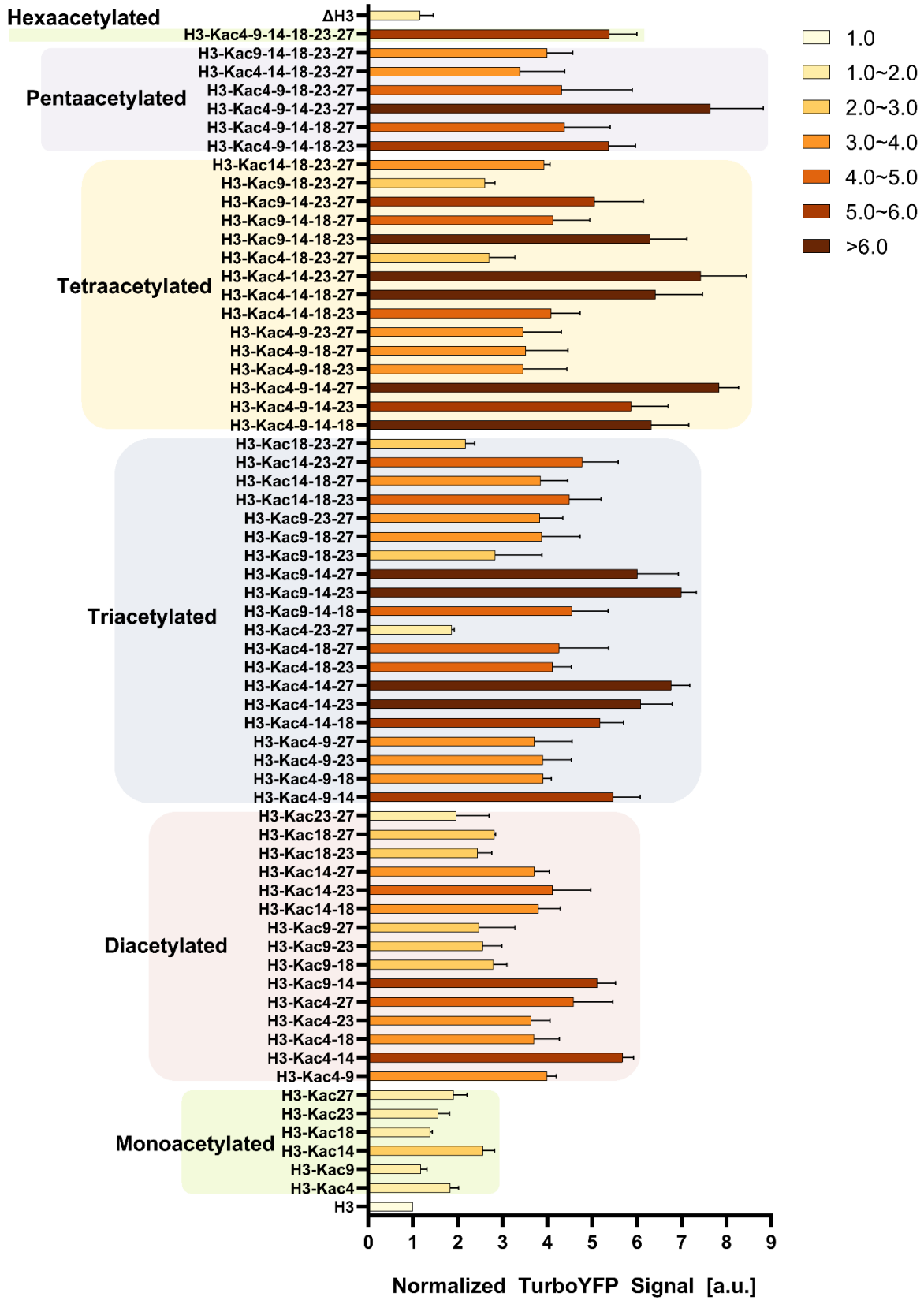


Figure 17. Pull-down assay results of acetyl-H3 nucleosomal libraries with CREBBP-TurboYFP. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

The pull-down assays of acetyl-H3 nucleosomal libraries with BRD3(2)-TurboYFP (Figure 18) showed an interaction profile similar to that of CREBBP-TurboYFP, i.e., the more acetylation sites are present in the nucleosomes, the stronger the corresponding normalized TurboYFP signal of recruited BRD3(2). The BRD3(2) exhibited a subtile binding preference for H3-Kac18 and H3Kac23 containing nucleosomes. This binding preference was enhanced when more acetylation sites were included. For instance, the normalized TurboYFP signal for nucleosomes containing diacetylated H3-Kac18-27, triacetylated H3-Kac18-23-27 and tetraacetylated H3-Kac4-9-18-23 were on average 2.9-, 3.5- and 5.1-fold stronger than the signals recorded with unmodified nucleosome after pull-down. Interestingly, the normalized TurboYFP signal for tetra- to hexa-acetylated H3 nucleosomes did not show a substantial difference, which indicated that the increase of acetylation sites beyond four marks did not enhance BRD recruitment to a major extent.

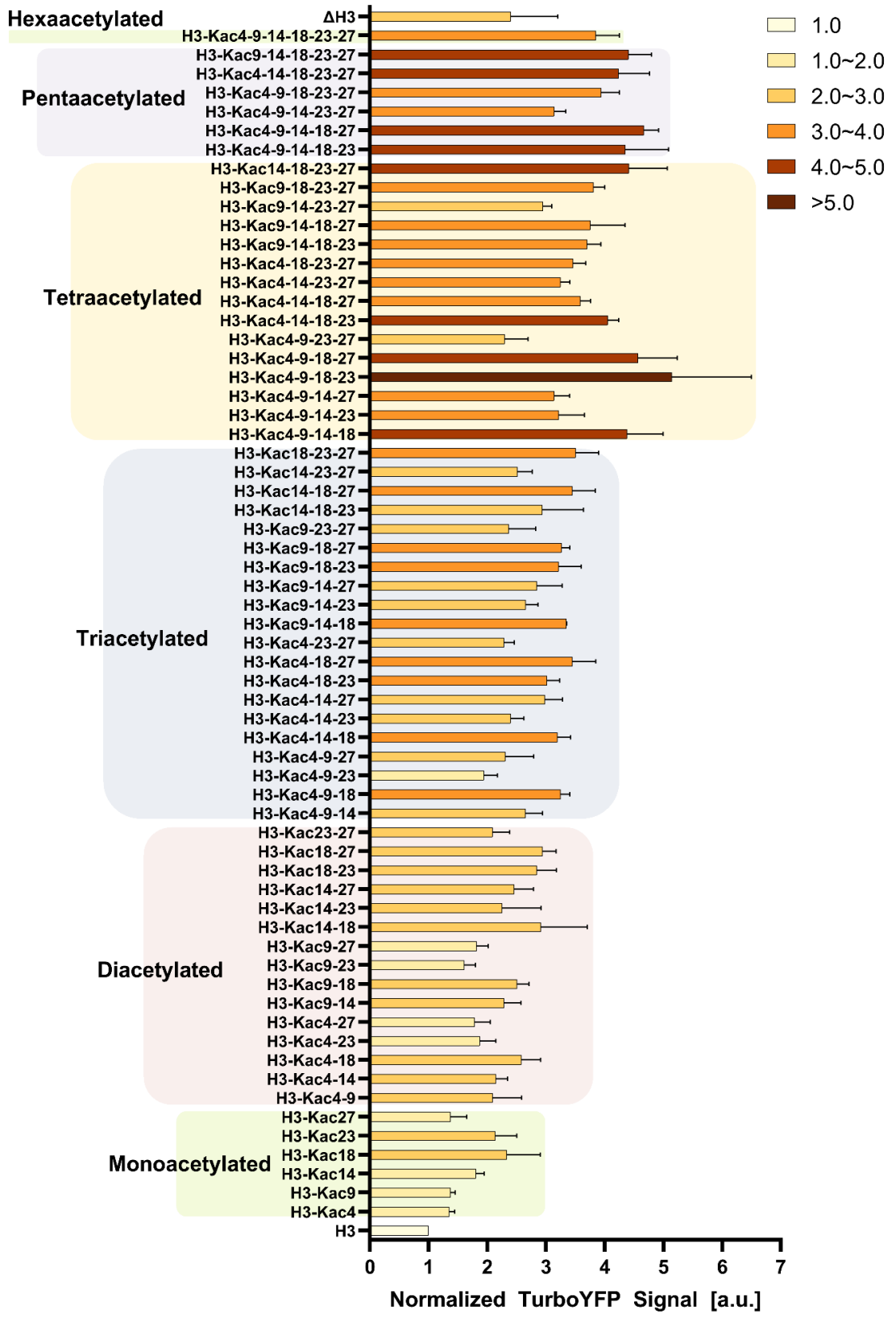


Figure 18. Pull-down assay results of acetyl-H3 nucleosomal libraries with BRD3(2)-TurboYFP. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

2.3 Expression and purification of BRD4(1/2)-mTagBFP2

To further investigate the impact of acetylated nucleosomes on proteins containing more than one bromodomain, BRD4 was chosen as the research target. BRD4 possesses two bromodomains and plays a vital role in diverse cellular processes [63] [64]. Therefore, the following illustrated BRD4(1/2) construct containing the tandem bromodomains of BRD4 (Figure 19A) was designed and expressed. To support the library screening, the BRD4(1/2) sequence was fused with the fluorescent protein mTagBFP2 on gene level, including a TEV protease cleavage site. In addition, a His6-Tag was fused to the N-terminus of the designed protein construct for affinity purification. The BRD4(1/2)-mTagBFP2 was expressed in *E.coli* and purified via metal affinity chromatography. The eluted fractions were analyzed by SDS-PAGE with Coomassie staining (Figure 19 B), and the fractions with high purity were collected, dialyzed, and concentrated. SDS-PAGE and Coomassie staining were used to analyze the purity of the protein (Figure 19C), which demonstrated that the BRD4(1/2)-mTagBFP2 was successfully generated.

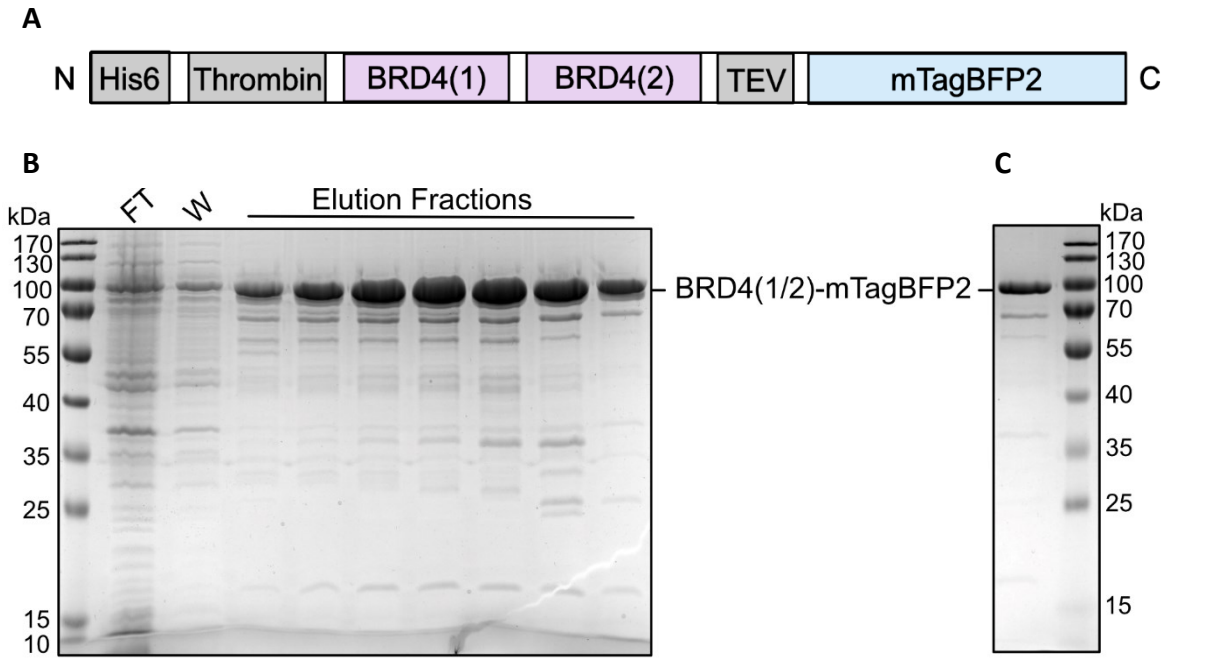


Figure 19. (A) Illustration of the BRD4(1/2)-mTagBFP2 construct. (B) During the purification of BRD4(1/2)-mTagBFP2, samples from flow-through (FT), washing (W), and elution fractions were collected and analyzed by SDS-PAGE and Coomassie staining. (C) The final BRD4(1/2)-mTagBFP2 product was analyzed by SDS-PAGE and Coomassie staining.

2.4 Pull-down assay of acetyl-H3 nucleosomal libraries with BRD4(1/2)

The pull-down screening of acetyl-H3 nucleosomal library with BRD4(1/2)-mTagBFP2 is shown in Figure 20. Similar to BRD3(2), the BRD4(1/2) protein exhibited a subtile preference to nucleosomes containing H3-Kac18, which showed 2.3-fold higher normalized mTagBFP2 intensity compared to the unmodified sample. When combined with the Kac27 mark, BRD4(1/2) retention on nucleosomes increased by 3.6-fold. In addition, in the groups of tri-, tetra-, and penta-acetylated nucleosomes, the strongest normalized mTagBFP2 intensities with values 4.32, 4.39, and 4.41 were detected with the nucleosomes containing H3-Kac18-23-27, H3-Kac14-18-23-27, and H3-Kac9-14-18-23-27. These results support the notion that the multiple acetylation marks are beneficial for the recruitment of BRD4(1/2)-mTagBFP2 protein to nucleosomes. Furthermore, the overall interaction profile followed a trend of enhanced normalized mTagBFP signal intensity that increased with the number of acetylation sites.

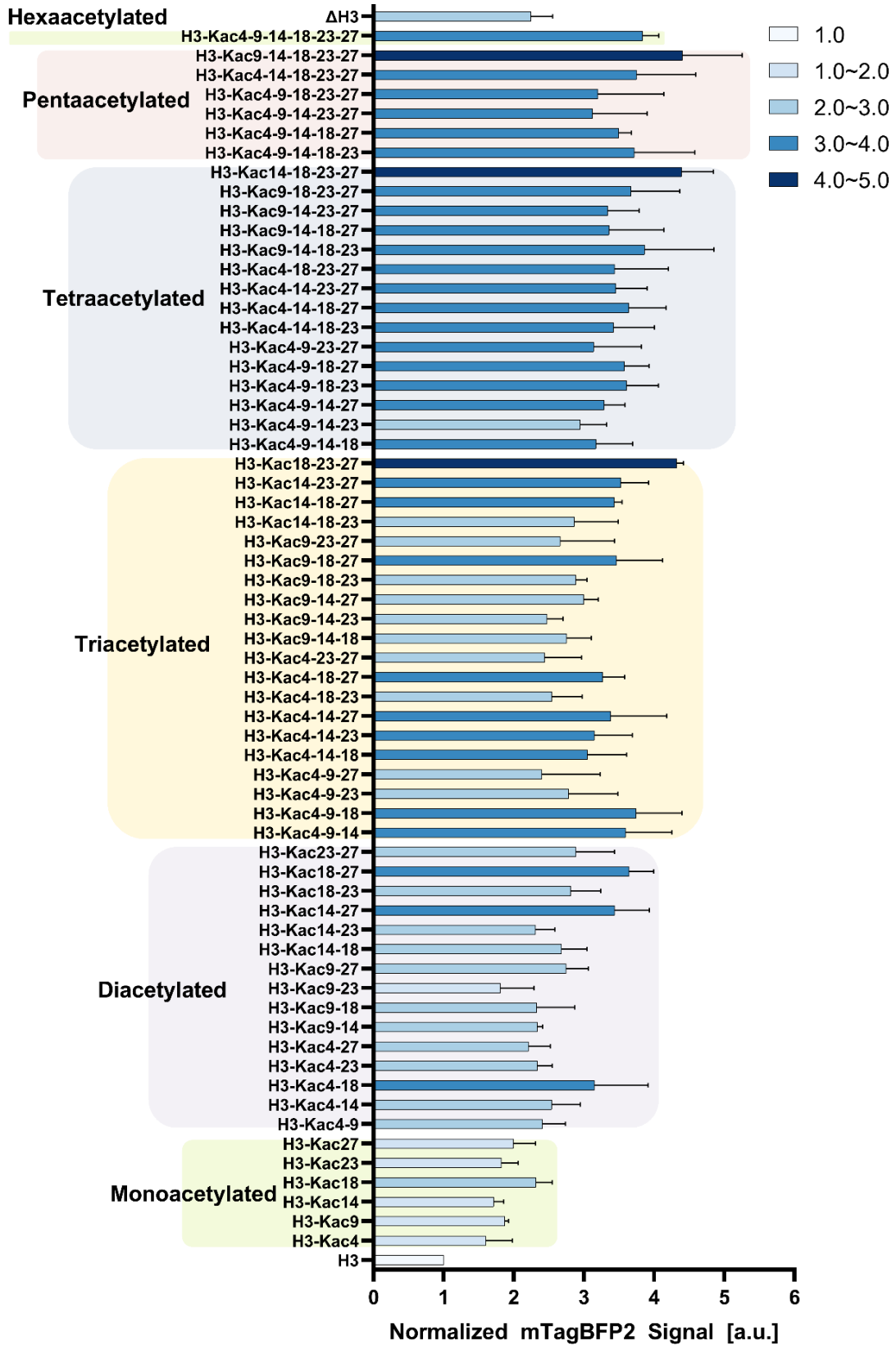


Figure 20. Pull-down assay results of acetyl-H3 nucleosomal libraries with BRD4(1/2)-mTagBFP2. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

2.5 Acetyl-H2A-H3 nucleosomal libraries

2.5.1 Assembly strategy of acetyl-H2A-H3 nucleosomal libraries

To further investigate the potential interaction between acetylation sites and different histones, H2A was included into the design of acetylated nucleosomes. Therefore, an orthogonal chemoselective ligation strategy was required to assemble the acetylated H2A histone to avoid interference with the sortase-mediated ligation of H3 tail. The utilization of protein trans-splicing (PTS) for the semi-synthesis of modified H2A nucleosomes has been reported by Pelaz et al. [126] [127]. Based on this, PTS was employed in this thesis to incorporate acetyl-H2A peptides into truncated H2A. The group of Prof. Dr. Wolfgang Fischle provided the starting material IntC-ΔH2A-ΔH3 nucleosomes. Compared to the ΔH3 nucleosomes, the IntC-ΔH2A-ΔH3 nucleosomes contained the N-terminal truncated H2A lacking residues 1-18 fused with the IntC fragment of the M86 mutant of the Ssp DnaB intein (Figure 21). Through PTS, the corresponding C-terminal IntN containing acetyl-H2A tail peptides (residues 1-18) were spliced with truncated H2A under release of IntC and IntN. To verify the success of PTS, the synthetic H2A peptides were further equipped with the TAMRA fluorophore at the N-terminus. Three additional residues Ser-Leu-Ile and one residue Gly were inserted into the IntN and IntC fragments, resulting in an insert of a GSIE motif into the assembled full-length histone H2A. These alterations were necessary to improve the PTS efficiency. The PTS reaction was firstly carried out in solution. After that, the spliced acetyl-H2A-ΔH3 nucleosomes were immobilized on the streptavidin-coated plates through the biotin tag, and the excess acetyl-H2A tail peptides and cleaved inteins were removed by washing. Subsequently, the sortase-mediated ligation was carried out to introduce the acetyl-H3 tail peptides onto the truncated H3.

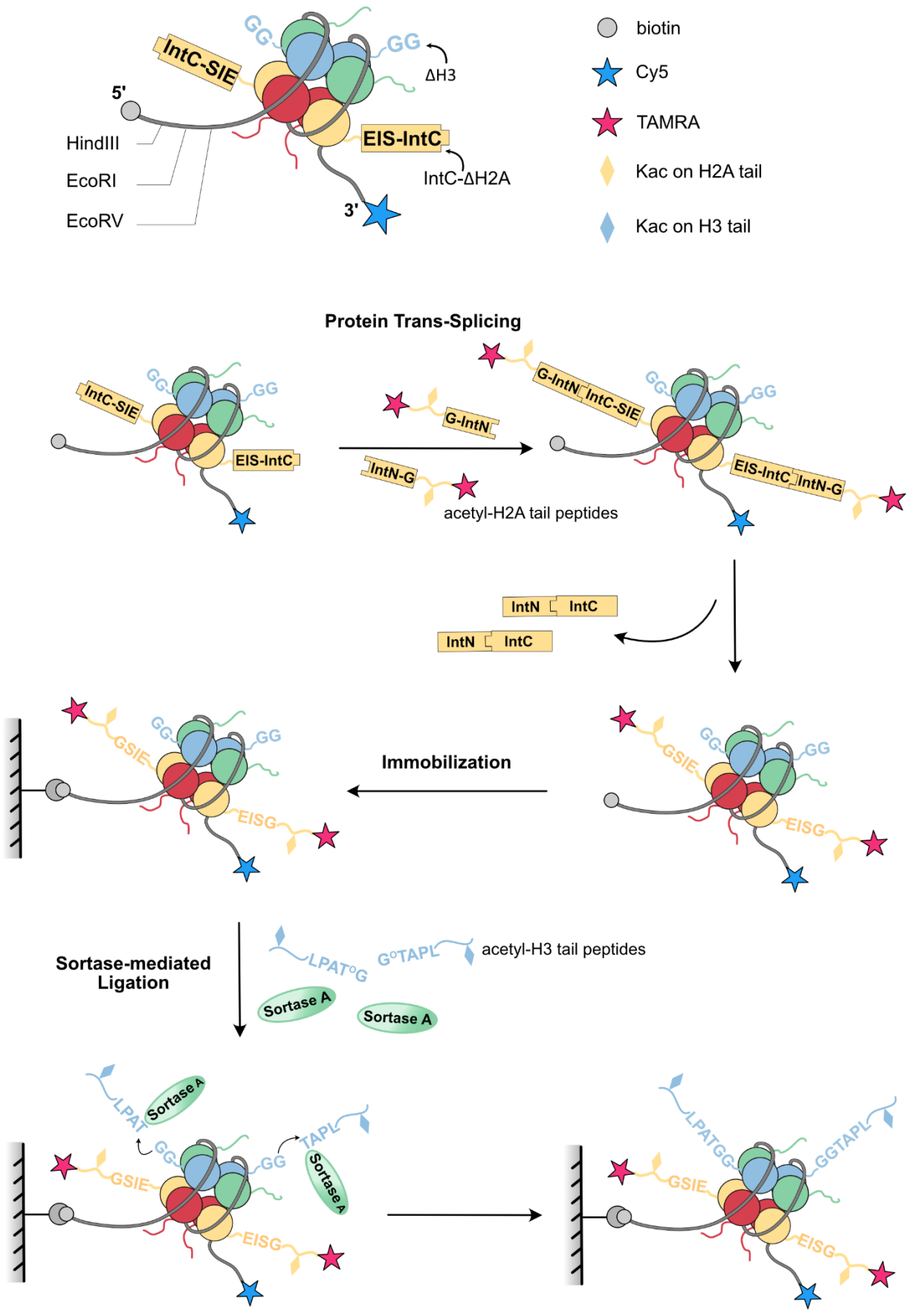


Figure 21. Structure of ligation-ready IntC-ΔH2A-ΔH3 nucleosomes and the assembly strategy of acetyl-H2A-

H3 nucleosomal libraries. The ligation-ready IntC-ΔH2A-ΔH3 nucleosome consists of full-length H2B (red), H4 (green), truncated H3 (residues 33-135) (blue), and truncated H2A (residues 19-129) (yellow) fused with IntC fragment (residues 12-158) of the M86 mutant of the Ssp DnaB intein. These four histones are assembled as an octamer and wrapped by 147 bp 601 Widom DNA containing three restriction sites (HindIII, EcoRI, EcoRV) and biotin at the 5' end and a Cy5 fluorophore at the 3' end. The assembly of acetyl-H2A-H3 nucleosomes starts with the protein trans-splicing (PTS) between the IntC fused truncated H2A and acetyl-H2A tail peptides (residues 1-18) containing the C-terminal IntN fragment (residues 1-11) and an N-terminal TAMRA fluorophore. Three amino acids (SIE) and a single amino acid (G) were separately inserted into the IntC-H2A and the acetyl-H2A tail peptides to facilitate the PTS efficiency, causing an insert of four residues (GSIE) in the spliced nucleosomes. After PTS, the acetyl-H2A-ΔH3 nucleosomes were immobilized on the streptavidin-coated plates and were incorporated with acetyl-H3 tail peptides through sortase-mediated ligation (SML).

2.5.2 Synthesis of acetyl-H2A tail peptides

In total 16 acetylated H2A tail peptides were designed, as shown in Table 2. The H2A tail sequence containing 18 amino acids was C-terminally extended with 11 amino acids of IntN sequence from M86 mutant of Ssp DnaB intein, which can be spliced with the corresponding IntC sequence fused to the N-terminus of truncated H2A, leading to the ligation between the synthetic H2A tail peptides and truncated H2A. Four acetylation sites, K5, K9, K13, and K15, on histone H2A tail were selected, resulting in 16 variants of acetyl-H2A tail peptides, including the unmodified H2A tail peptide. Couplings of all amino acids were performed by an automated peptide synthesizer based on standard SPPS protocols. No additional pseudoprolines were required for the synthesis of acetylated H2A tail peptides with a length of 30 amino acids. After the synthesis, peptides were cleaved off the resin and purified as described in the methods section. The final products were analyzed by analytical HPLC and LC-MS, and the results (Figure 22 and Appendixes 8.2) showed that all acetyl-H2A tail peptides were successfully synthesized and purified with purities between 90% to 98%. It should be noted that the material contains two isomers, 5-TAMRA and 6-TAMRA, resulting in two prominent peaks corresponding to the peptides coupled with different TAMRA isomers as shown in the analytical HPLC traces of some peptides.

Table 2. Acetyl-H2A tail peptides sequences. Acetylated Lysin (Kac) is marked in red. IntN sequence from M86 mutant of Ssp DnaB intein is colored in yellow.

No.	Name	Sequence
1	H2A	TAMRA – S G R G K Q G G K T R A K A K T R S G - C I S G D S L I S L A – OH
2	H2A-Kac5	TAMRA – S G R G Kac Q G G K T R A K A K T R S G - C I S G D S L I S L A – OH
3	H2A-Kac9	TAMRA – S G R G K Q G G Kac T R A K A K T R S G - C I S G D S L I S L A – OH
4	H2A-Kac13	TAMRA – S G R G K Q G G K T R A Kac A K T R S G - C I S G D S L I S L A – OH
5	H2A-Kac15	TAMRA – S G R G K Q G G K T R A K A Kac T R S G - C I S G D S L I S L A – OH
6	H2A-Kac5-9	TAMRA – S G R G Kac Q G G Kac T R A K A K T R S G - C I S G D S L I S L A – OH
7	H2A-Kac5-13	TAMRA – S G R G Kac Q G G K T R A Kac A K T R S G - C I S G D S L I S L A – OH
8	H2A-Kac5-15	TAMRA – S G R G Kac Q G G K T R A K A Kac T R S G - C I S G D S L I S L A – OH
9	H2A-Kac9-13	TAMRA – S G R G K Q G G Kac T R A Kac A K T R S G - C I S G D S L I S L A – OH
10	H2A-Kac9-15	TAMRA – S G R G K Q G G Kac T R A K A Kac T R S G - C I S G D S L I S L A – OH
11	H2A-Kac13-15	TAMRA – S G R G K Q G G K T R A Kac A Kac T R S G - C I S G D S L I S L A – OH
12	H2A-Kac5-9-13	TAMRA – S G R G Kac Q G G Kac T R A Kac A K T R S G - C I S G D S L I S L A – OH
13	H2A-Kac5-9-15	TAMRA – S G R G Kac Q G G Kac T R A K A Kac T R S G - C I S G D S L I S L A – OH
14	H2A-Kac5-13-15	TAMRA – S G R G Kac Q G G K T R A Kac A Kac T R S G - C I S G D S L I S L A – OH
15	H2A-Kac9-13-15	TAMRA – S G R G K Q G G Kac T R A Kac A Kac T R S G - C I S G D S L I S L A – OH
16	H2A-Kac5-9-13-15	TAMRA – S G R G Kac Q G G Kac T R A Kac A Kac T R S G - C I S G D S L I S L A – OH

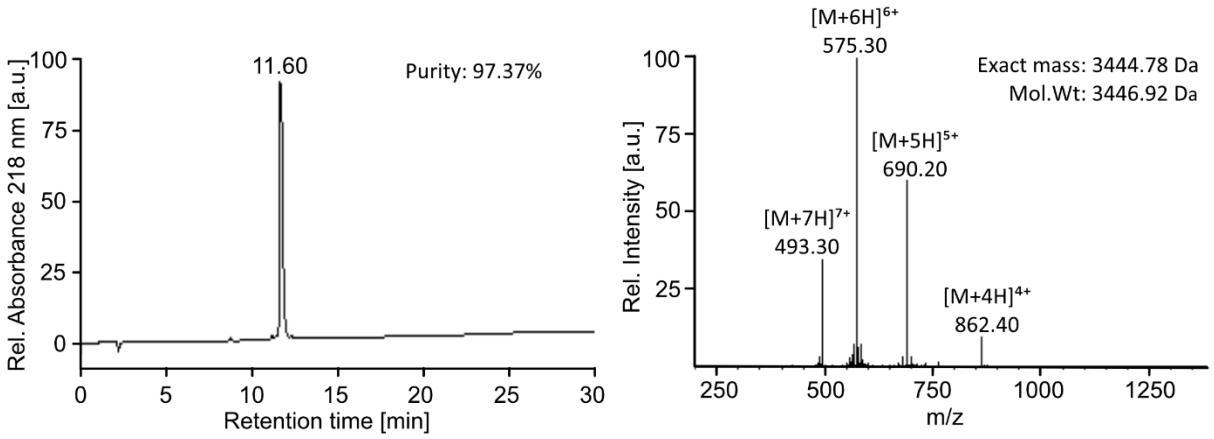


Figure 22. Analytical HPLC and LC-MS results of acetyl-H2A tail peptide No.2 H2A-Kac5 as representative example.

2.5.3 Installation of acetyl-H2A tail peptides on truncate histone H2A nucleosomes

To maintain the activity of the IntC fused truncated H2A histone protein, Protein trans-splicing (PTS) between acetyl-H2A tail peptides and IntC-ΔH2A-ΔH3 nucleosomes was carried out at

5 °C in a neutral buffer. Therefore, the PTS reaction time was prolonged to one week to ensure the reaction was performed thoroughly. Since the PTS reaction is irreversible, extended reaction time did not cause an increase in by-products. After the PTS reactions were completed, the products were analyzed by SDS-PAGE with Coomassie staining, and the results are shown in Figure 23A. In the product lane (left panel), the upper band corresponding to IntC-H2A was very faint, suggesting a high conversion of the IntC-ΔH2A-ΔH3 nucleosomes. A distinct lower band with a strong TAMRA signal shows the acetyl-H2A ligation product. Since the ligated full-length H2A possessed a similar molecular mass as H2B, both proteins cannot be separated by SDS-PAGE. The PTS conversion was further analyzed by Western Blot (WB) (Figure 23B), which exhibited the expected bands with the correct molecular mass of H2A and strong TAMRA signals, and the starting material as control. By comparing the intensities and areas of the upper and lower bands corresponding to starting material and product, respectively, using Image Lab, the conversion was estimated to be approximately 76%.

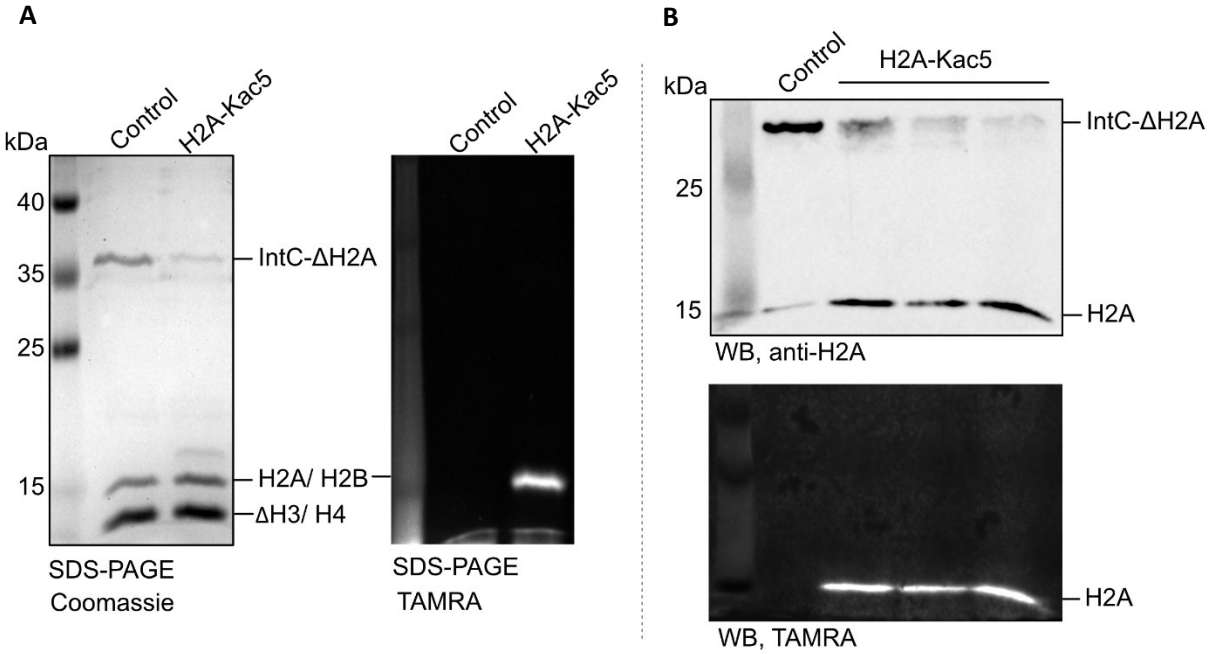


Figure 23. The products of protein trans-splicing between acetyl-H2A tail peptides and IntC-ΔH2A-ΔH3 nucleosomes were analyzed by (A) SDS-PAGE with Coomassie staining and (B) Western Blot (WB).

2.5.4 Installation of acetyl-H3 tail peptides on truncated histone H3 nucleosomes

After the accomplishment of the PTS reaction, acetyl-H2A-ΔH3 nucleosomes were immobilized in streptavidin-coated 96-well plates, and excessive acetyl-H2A tail peptides and the Intein fragments were removed in several wash steps. The sortase-mediated ligation (SML) was then carried out as described before. The product of the SML reaction was released and analyzed by Western blot. The results are shown in Figure 24, in which the distinct upper band corresponding to the full-length H3 occurred in the lane of the product sample, verifying the success of the SML between acetyl-H3 tail peptides and acetyl-H2A-ΔH3 nucleosomes. The conversion of the SML was quantified by Image J, showing 89±1%, which indicated that the previously completed PTS reaction does not influence the efficiency of the SML reaction.

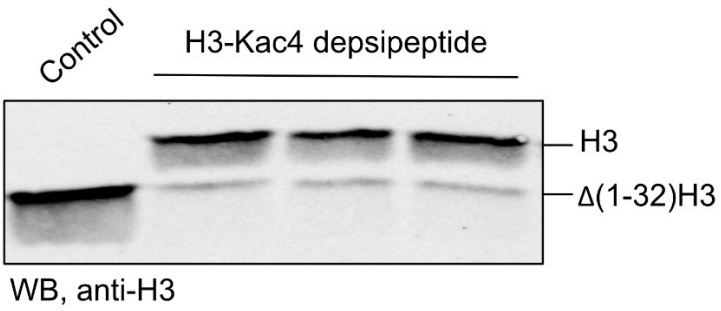


Figure 24. The Sortase-mediated ligation between acetyl-H2A-ΔH3 nucleosomes and acetyl-H3 tail peptides H3-Kac4 was analyzed by Western blot (WB).

2.6 Pull-down assay of acetyl-H2A-H3 nucleosomes with BRD4(1/2)

At first, the impact of H2A acetylation sites on BRD4 recruitment was investigated without potential crosstalk of H3 modifications. Therefore, the 16 acetyl-H2A nucleosomes were equipped with unmodified H3 tail peptides at ΔH3. After the immobilization of acetyl-H2A nucleosomal libraries, the pull-down assay with BRD4(1/2)-mTagBFP2 was carried out as previously described. The library screening results (Figure 25) revealed a subtile binding preference of BRD4(1/2) to the nucleosomes containing H2A-Kac5-9, of which the normalized

mTagBFP2 intensity is almost twice as strong as that of unmodified nucleosomes. Among all 16 acetyl-H2A nucleosome variants, the nucleosome carrying H2A-Kac5-9-15 showed the strongest normalized mTagBFP2 signal on average 2.2-fold higher than that of unmodified nucleosomes. Therefore, these two H2A acetylation variants, H2A-Kac5-9 and H2A-Kac-5-9-15, were selected for combination with different H3 acetylation variations in nucleosomal contexts to further investigate the effect of potential crosstalk between acetylated H2A and H3 on the recruitment of BRD4(1/2).

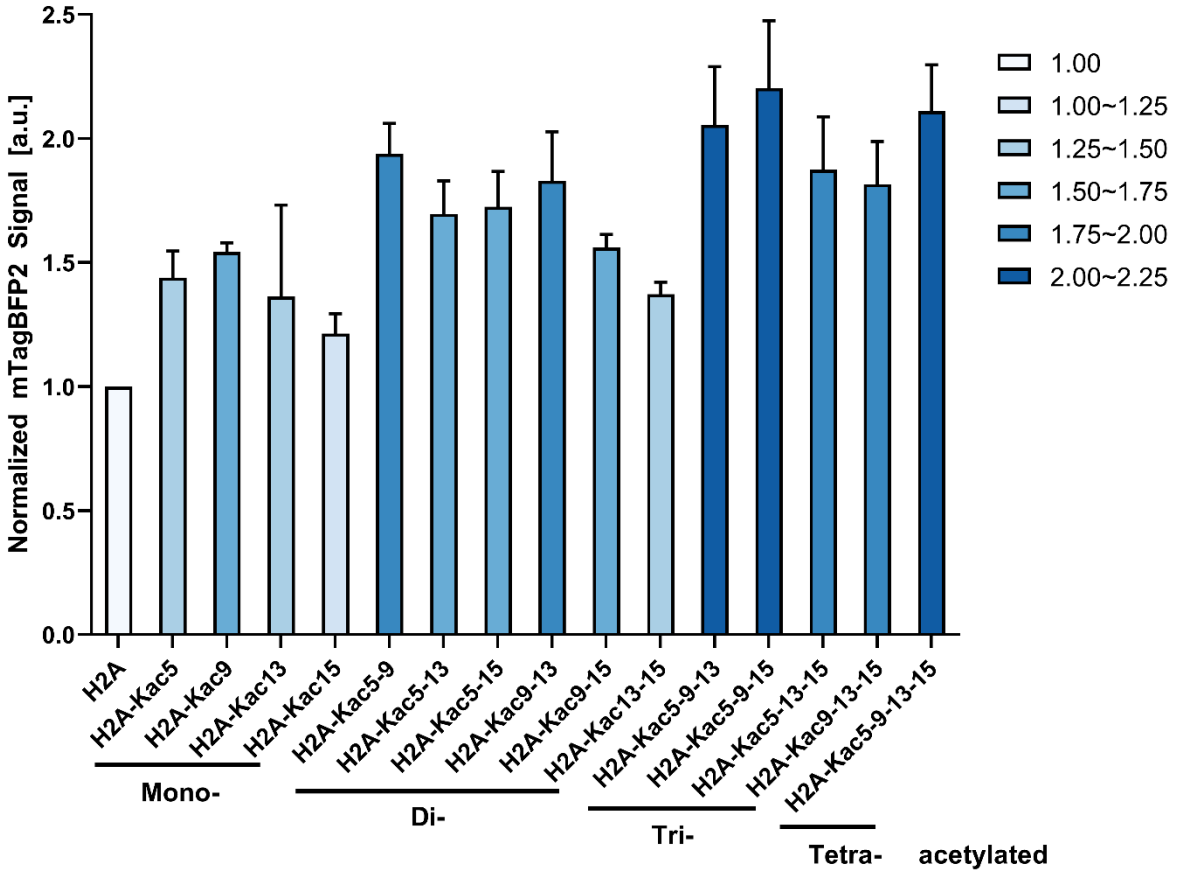


Figure 25. Pull-down assay results of acetyl-H2A nucleosomal libraries with BRD4(1/2)-mTagBFP2. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

After assembly of the two further libraries with H2A-Kac5-9 and H2A-Kac5-9-15 combined with permutated acetylation patterns of H3, pull-down experiments were performed with BRD4(1/2)-mTagBFP2. Similar to the pull-down assay of acetyl-H3 nucleosomal libraries with BRD4(1/2)-mTag BFP2, the pull-downs with H2A-Kac5-9 (Figure 26) and H2A-Kac5-9-15 (Figure

27) containing acetyl-H2A-H3 nucleosomal libraries also showed a binding preference of BRD4(1/2) to nucleosomes containing H3-Kac18. However, the highest normalized mTagBFP2 signal intensities in the monoacetylated H3 group, observed in nucleosomes with H3-Kac18 marks, increased from 2.3 to 3.5 or 3.9 when combined with H2A-Kac5-9 and H2A-Kac5-9-15, respectively. The groups of diacetylated, as well as tri-, tetra-, and penta-acetylated nucleosomes, showed a similar pattern: the presence of the H2A-Kac5-9 or H2A-Kac5-9-15 marks led to higher mTagBFP2 signals. In the H2A-Kac5-9 containing nucleosomal libraries, the highest normalized mTagBFP2 signal for di-, tri-, tetra-, and penta-acetylated H3 group were observed in nucleosomes that contained H3-Kac14-18, H3-Kac18-23-27, H3-Kac9-14-18-23, and H3-Kac9-14-18-23-27, with values of 4.9, 5.2, 5.2 and 5.6, respectively. These values were all higher than those observed in acetyl-H3 nucleosomes lacking H2A acetylation marks. In the case of H2A-K5-9-15 containing nucleosomal libraries, nucleosomes including H3-Kac18-27, H3-Kac9-18-27, H3-Kac9-18-23-27, and H3-Kac9-14-18-23-37 exhibited 6.5-, 6.4-, 6.9- and 7.7-fold increase in normalized mTagBFP2 intensities compared to unmodified nucleosome, respectively. These nucleosomes displayed the strongest binding among each tri-, tetra- and penta-acetylated H3 group for BRD4(1/2). In summary, the additive effect of multiple acetylation marks on H3 on BRD4(1/2) recruitment was further enhanced by multiple acetylation marks of H2A.

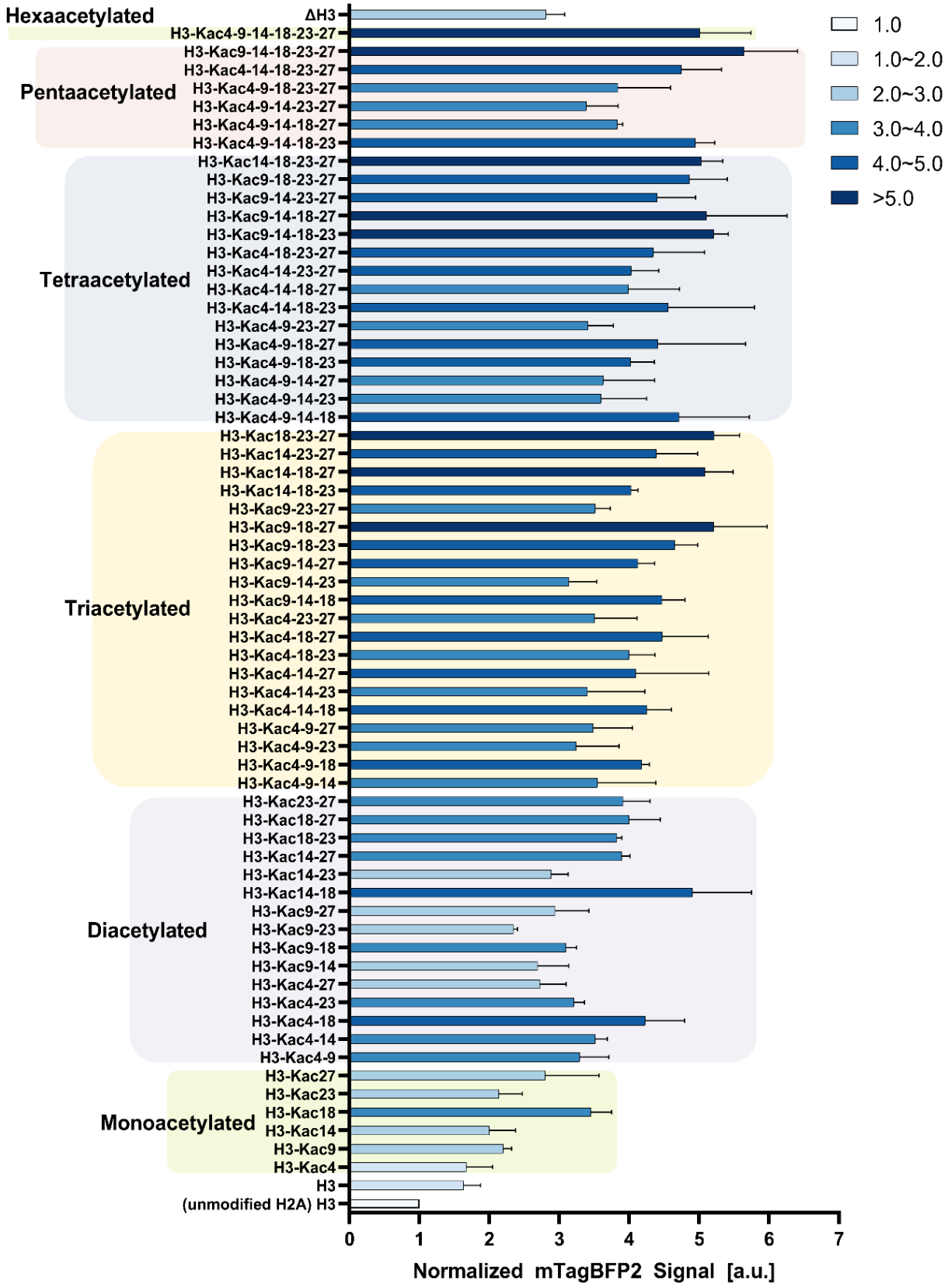


Figure 26. Pull-down assay results of H2A-Kac5-9/acetyl-H3 nucleosomal libraries with BRD4(1/2)-mTagBFP2. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

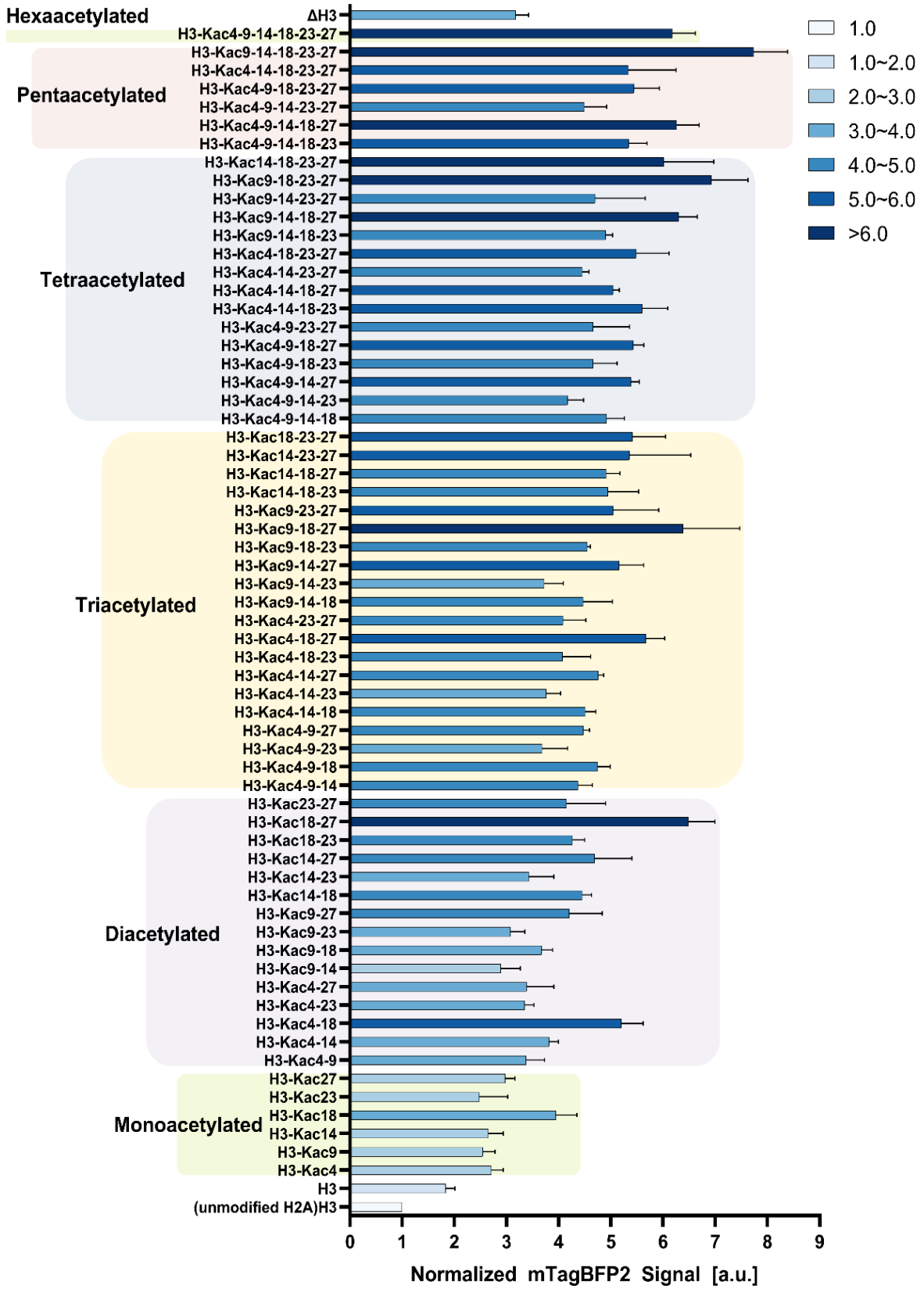


Figure 27. Pull-down assay results of H2A-Kac5-9-15/acetyl-H3 nucleosomal libraries with BRD4(1/2)-mTagBFP2. Values are shown as the average of triplicate measurements, with error bars representing the standard deviations.

2.7 Semi-synthesis of a construct linking acetyl-H2A and H3 tail peptides

2.7.1 Semi-synthesis strategy

A central finding of the nucleosomal library screenings was that the recruitment efficiency of most of the tested BRDs increased with the number of acetylation marks on the histone tail. This additive effect might result from the avidity effects of BRD bindings to multiple acetylation marks or enhanced access to the histone tails as a result of charge neutralization. The latter effect should be restricted to nucleosomes. To explore whether the observed recruitment of BRD4(1/2) under the impact of acetylated H2A and H3 tails was confined to the nucleosomal context, a construct that can display both acetylated H2A and H3 tail peptides without the nucleosome context was designed. The corresponding assembly strategy is illustrated in Figure 28. In order to use the already synthesized H2A and H3 tail peptides, the previous SML and PTS ligation strategies were utilized in the semi-synthesis of this construct. Hence, the designed construct needed to contain both the N-terminal IntC fragment from the M86 mutant of the Ssp DnaB intein and the N-terminal glycines. As a result, the construct required two N-termini in order to correctly display N-terminal IntC and glycines. The linkage of these two parts needed a new ligation strategy to avoid interference with SML and PTS. Consequently, phosphopantetheinyl transferase-catalyzed ligation (PCL) was employed to link a pentaglycine unit as the second N-terminus to the construct. Therefore, an optimized amino acid sequence, referred to as the S6 tag, for coenzyme A (CoA) bioconjugation by PCL using PPTase Spf was introduced into the scaffold protein.

In order to use PCL for installing N-terminal glycines for subsequent SML, a peptide-CoA conjugate was designed for subsequent installation on the IntC scaffold protein. The peptide contained a bridging lysine conjugated to CoA at the side chain and glycine extension at the Nα-amino. In addition, this peptide included two arginines to increase the hydrophilicity. This peptide is referred to as the linker peptide in the following.

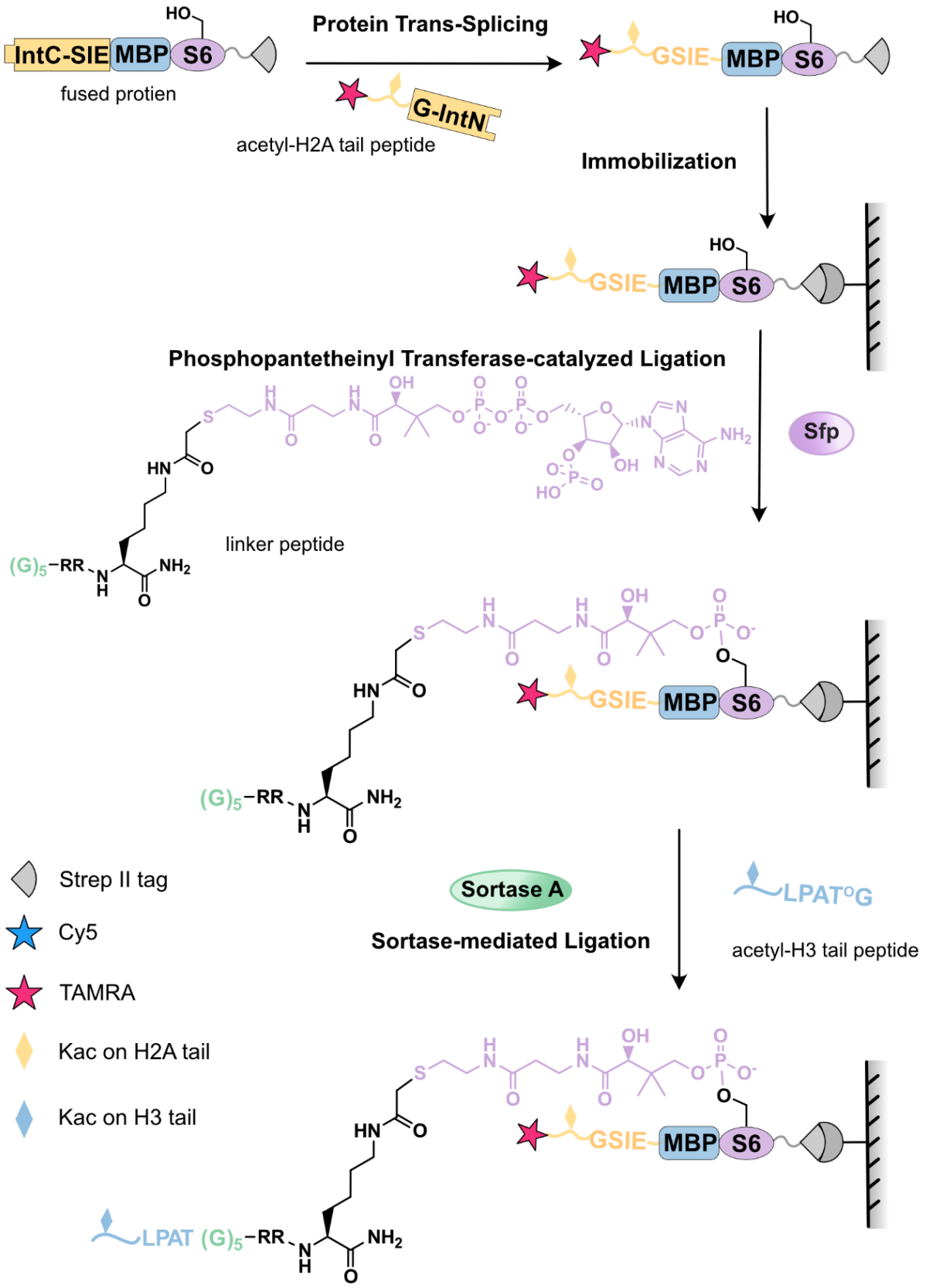


Figure 28. Semi-synthesis strategy of a construct linking acetyl-H2A and H3 tail peptides.

To increase the solubility and stability of this IntC-containing protein, the maltose-binding protein (MBP) was included as scaffold, located between the IntC sequence and the S6 tag. Furthermore, a C-terminal Strep-Tag II was included for immobilizing the construct in the streptavidin-coated plates. In addition, the His6-Tag between the S6 tag and Strep II Tag was used for simple purification of the recombinant protein. This protein was named IntC-MBP-S6. The assembly of the whole construct started with PTS ligation between the generated protein IntC-MBP-S6 and the synthetic acetyl-H2A tail. Then, the PCL was carried out to install the linker peptide on the IntC-MBP-S6 construct, followed by the SML on the solid support. The ligation of the acetyl-H3 tail peptide and linker peptide was the final step of the assembly. As a result, a construct displaying acetyl-H2A tail and acetyl-H3 tail peptides should be generated.

2.7.2 Synthesis of the linker peptide

The synthesis scheme of the linker peptide is illustrated in Figure 29. The linker peptide, including N-terminal pentaglycine, two arginines, and one lysine, was synthesized manually based on the SPPS. The α-amino group of all amino acids was temporarily protected by the Fmoc groups except the N-terminal glycine, which was Boc-protected. The ε-amino of the lysine side chain was orthogonally protected by the Alloc group to facilitate the CoA conjugation. The Alloc protection group is base- and acid-stable and can be removed under neutral conditions through the palladium catalyst-amine-borane systems. After removing the Alloc group, the free amino group on the side-chain of lysine was reacted with bromoacetic anhydride under water and oxygen-free conditions. The bromoacetylated peptide was then cleaved off the resin and purified by HPLC. The conjugation between the purified peptide and CoA was carried out in solution under oxygen-free mild base conditions. After the reaction had proceeded to completion, the CoA-conjugated linker peptide was purified again. The final product was analyzed by analytical HPLC and LC-MS, and the corresponding results (Figure 30) demonstrated that the designed linker peptide was successfully synthesized with a purity of 97.91%.

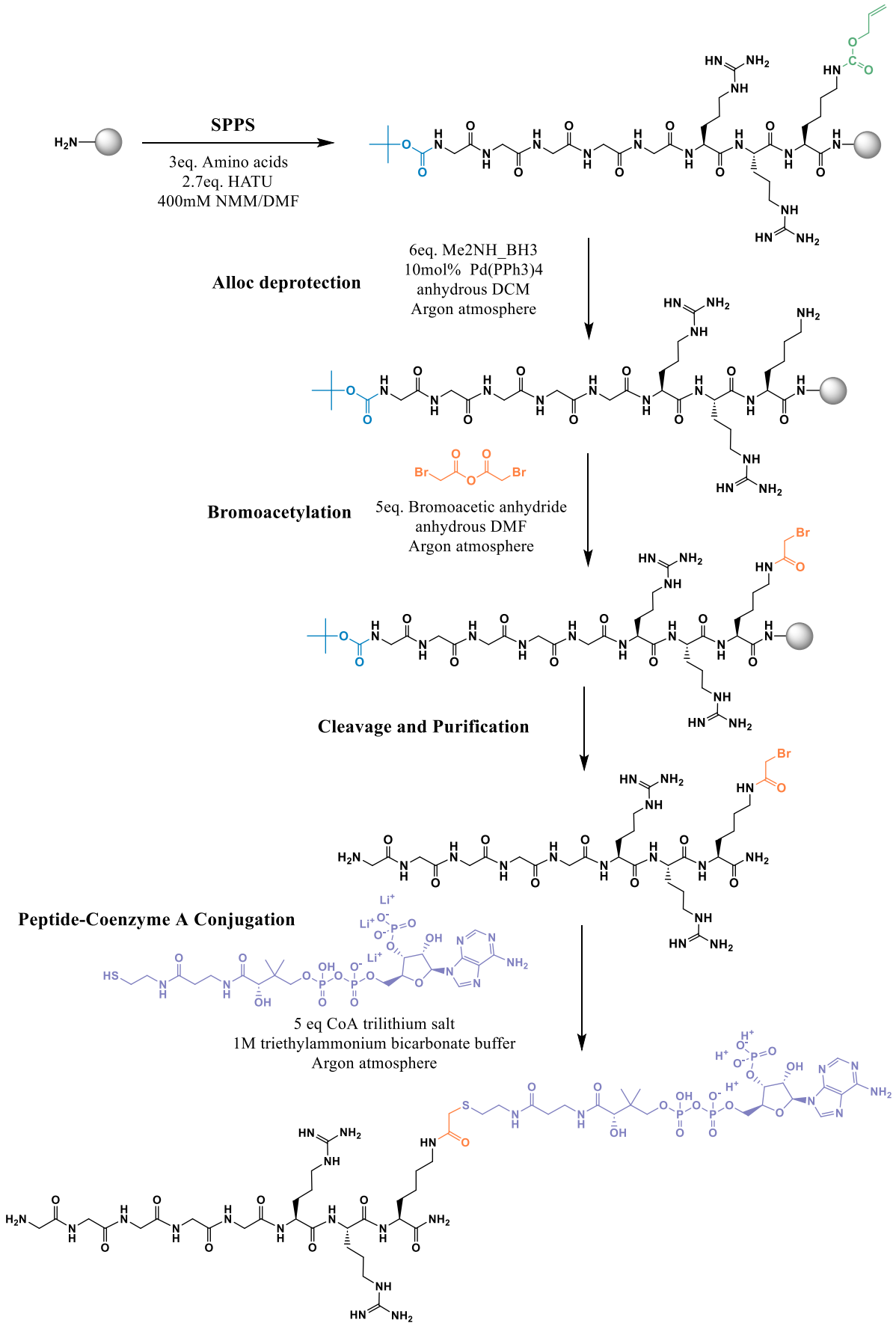


Figure 29. Synthesis scheme of the linker peptide.

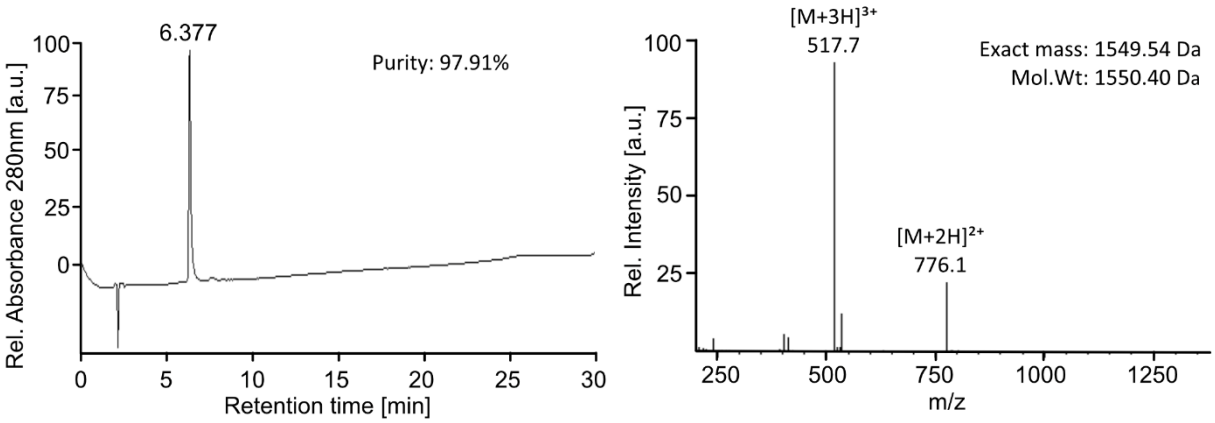


Figure 30. Analytical HPLC and LC-MS analysis of the linker peptide.

2.7.3 Expression and purification of the IntC-MBP-S6 protein

The designed IntC-MBP-S6 protein construct is illustrated in Figure 31A. The whole construct was obtained by gene synthesis and expressed in *E. coli*. The purification of the IntC-MBP-S6 was performed by Ni-NTA affinity chromatography. SDS-PAGE and Coomassie staining were applied to analyze the fractions from each purification step and the pellet from the centrifugation of the lysate (Figure 31B). The purest protein fractions were collected, dialyzed, and concentrated. The final product was then analyzed by SDS-PAGE and Coomassie staining. The analysis showed that the purification of the IntC-MBP-S6 protein was successful, with only a minor impurity of protein without IntC. A significant amount of IntC-MBP-S6 protein remained in the pellet, indicating that this fraction of the protein was not correctly folded.

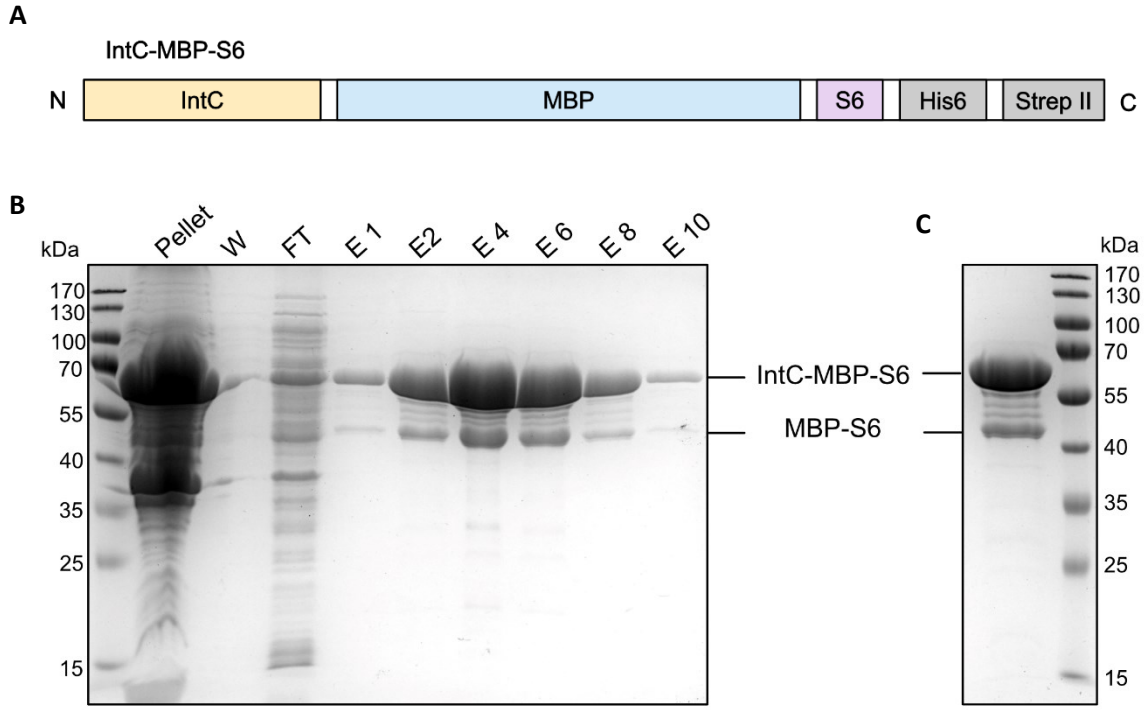


Figure 31. (A) Illustration of IntC-MBP-S6 protein construct. (B) Samples from washing fraction (W), flow through (FT), elution fractions (E1-10) and (C) final protein product were analyzed by SDS-PAGE and Coomassie staining.

2.7.4 Installation of acetyl- H2A and H3 tail peptides on the synthetic construct

With all required parts in hand, installing acetylated H2A and H3 tail peptides on the synthetic construct was performed by subsequent PTS, PCL, and SML as described in the assembly strategy (Figure 28). The corresponding products and intermediates are illustrated in Figure 32A. The PTS reaction was carried out as the first step in a ratio of IntC-MBP-S6: H2A-Kac5 as 1:3 under 25°C. The corresponding samples were collected at selected intervals and analyzed by SDS-PAGE. The results (Figure 32B) showed that after 4 hours, the product (2) band, identified by the TAMRA signal, appeared below the reactant (1) band. After 22 hours, most reactants were converted to products, and the conversion rate no longer increased with time.

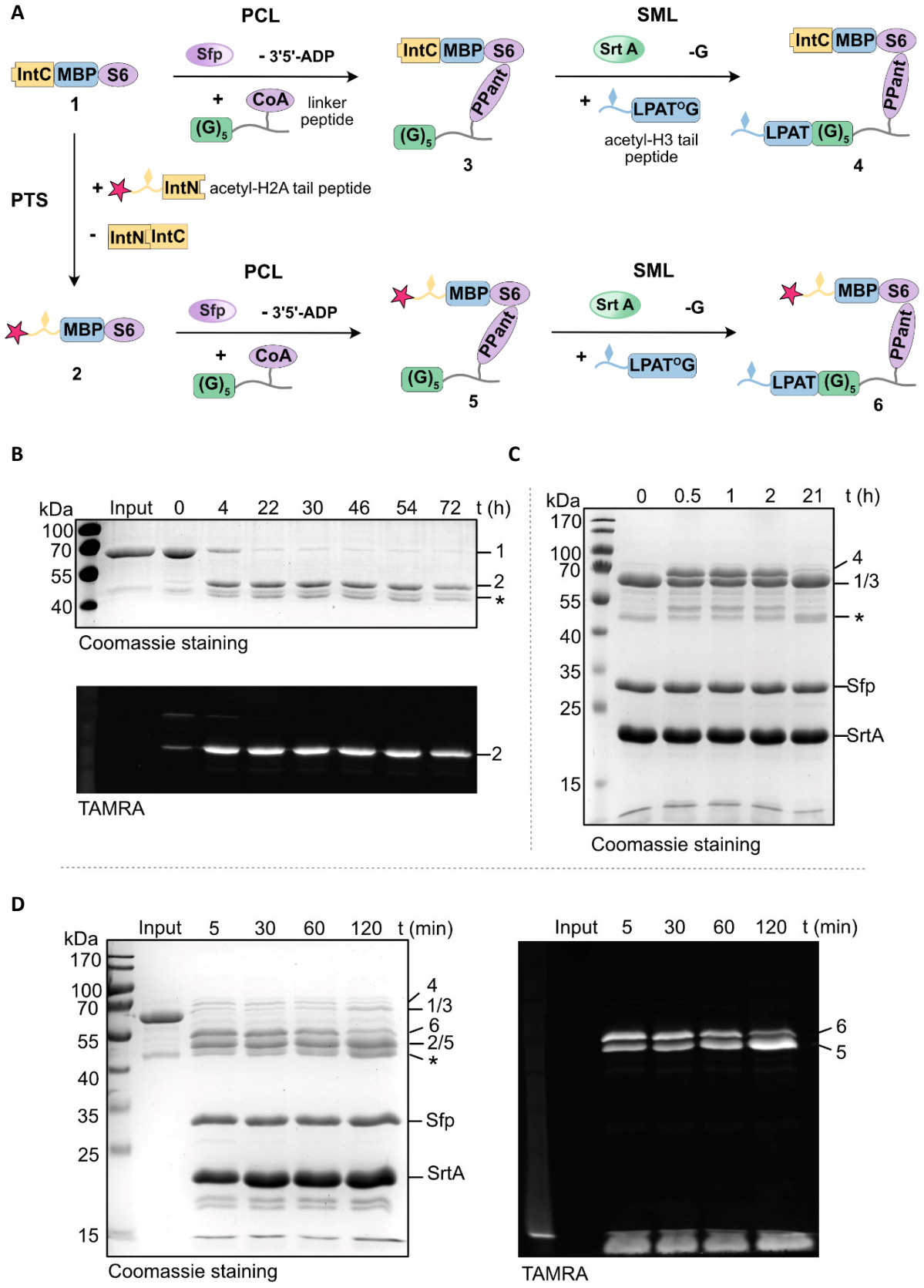


Figure 32. (A) Illustration of the assembly process. The fused protein IntC-MBP-S6 (1) can ligate with acetyl-H2A tail peptides through the protein trans-splicing (PTS). The spliced acetyl-H2A-MBP-S6 (2)

was then catalyzed by the PPTase Sfp to conjugate with the linker peptide, resulting in the generation of the fragment (5), which was followed by the sortase-mediated ligation (SML) to yield the final product (6) containing acetyl-H2A and H3 tail peptides. Meanwhile, the unreacted IntC-MBP-S6 (1) can carry out the phosphopantetheinyl transferase-catalyzed ligation (PCL) to incorporate the linker peptide, leading to the assembly of the construct (3), which was catalyzed by sortase A to ligate with acetyl-H3 tail peptides, resulting in construct (4). (B) The PTS between (1) and acetyl-H2A tail peptides was analyzed by SDS-PAGE with Coomassie staining. The TAMARA signal was detected before Coomassie staining. (C) The PCL between (1) and linker peptides and the SML between linker peptides and acetyl-H3 tail peptides were carried out in one pot. The results were analyzed by SDS-PAGE with Coomassie staining. (D) The PTS and PCL proceeded together, and the products were linked with acetylated H3 tail peptides through the SML. The results were analyzed by SDS-PAGE with Coomassie staining. * represents protein impurity.

Considering that the difference between the molecular masses of the product and reactant of the PCL reaction is too small to be separated by SDS-PAGE, the PCL and SML reactions were performed together in the ratio of 1:2:4 (IntC-MBP-S6: linker peptide: H3-Kac4) at 25°C and then analyzed by SDS-PAGE. The analysis of this reaction (Figure 32C) showed that the total conversion of these two reactions reached about 50% after half hour, with no further increase in conversion over time. Since the PCL is irreversible and the SML is reversible, the end product (4) reversed to intermediated product (3) when the reaction time was prolonged to 21 hours.

In order to shorten the time to assemble the desired construct, PTS and PCL were performed together with 10 µM IntC-MBP-S6, 30 µM H2A-Kac5 peptides, 20 µM linker peptides, and 10 µM Sfp in a final volume of 40 µl under 25 °C for 20 hours. After that, 40 µM H3-Kac4 peptides and 20 µM SrtA were added to perform the SML reaction at 37 °C. After each selected time point, a 10 µl sample was collected and analyzed by SDS-PAGE. The analysis (Figure 32D) showed that after 5 minutes, a new band with the expected molecular mass of the end product (6) displaying a TAMRA signal occurred, indicating that the semi-synthesis of the desired construct linking acetyl-H2A and H3 tail peptides could be achieved, and the designed assembly strategy was feasible. However, due to the reversibility of the SML, the end product was reversed to the intermediate (5) after 2 hours at 37 °C, and the conversion ratio of the whole assembly process was below 50%. In addition, fragments (3) and (4) indicated by faint bands were observed, since the unreacted IntC-MBP-S6 could undergo the PCL and SML

reactions. To increase the conversion rate and get a stable product, the assembly conditions will need to be optimized in the future, and the pulldown assay with this construct can be performed afterward.

The main motivation for establishing this display system of the H2A and H3 tails was to probe if additive effects of multiple acetylation marks on BRD4(1/2)-mTagBFP2 recruitment were nucleosome independent. In this case, avidity effects rather than charge neutralization are likely the cause of the observed effect. In order to obtain this information, a more simple peptide display system for the H3 tail peptides alone was used.

2.8 Pull-down reactions of acetyl-H3 tail peptides with BRD4(1/2)

The recruitment of BRD4(1/2) to isolated H3 tail peptides carrying the same acetylation patterns as selected nucleosomes was examined. Considering that BRD4(1/2) exhibited a binding preference for H3-Kac18 and H3-Kac18-23-27 nucleosomes, corresponding depsipeptides, H3-Kac18 and H3-Kac18-23-27, were chosen. In parallel, the unmodified H3 tail peptide was selected as control. A biotin-containing linker peptide (Figure 33B), supplied by Dr. Sören Kirchgäßner at the Interfaculty Institute of Biochemistry, University of Tübingen [137], was used as scaffold for the immobilization of acetyl-H3 tail peptides on streptavidin-coated plates. Moreover, the biotin linker peptide contains an N-terminal glycine for SML with H3 tail peptides, as well as a Cy5 fluorophore for visualization and quantification of immobilized H3 tail peptides. To facilitate SML, the biotin linker peptides were first ligated with H3-tail peptides in solution and then immobilized on the plates (Figure 33A). The ligated products were analyzed by SDS-PAGE with a 4-20% gradient gel and illuminated in the Cy5 channel (Figure 33C). The intensities and areas of the upper and lower bands, corresponding to ligated H3-linker and linker peptides, were detected and analyzed using the Image Lab software, showing a conversion of approximately 73%. After immobilization, the pull-down reactions were performed with BRD4(1/2)-mTagBFP2 in three replicates. The analysis showed that the normalized mTagBFP2 signal increased from 2.6 for H3-Kac18 to 5.7 for H3-Kac18-23-27 on

average (Figure 33D), indicating that the additive effect of multiple acetylation sites on BRD recruitment is also detectable without the nucleosomal context. This finding supports the notion that the avidity of multiple Kac sites plays a role in BRD4(1/2) recruitment rather than charge neutralization enhancing tail accessibility.

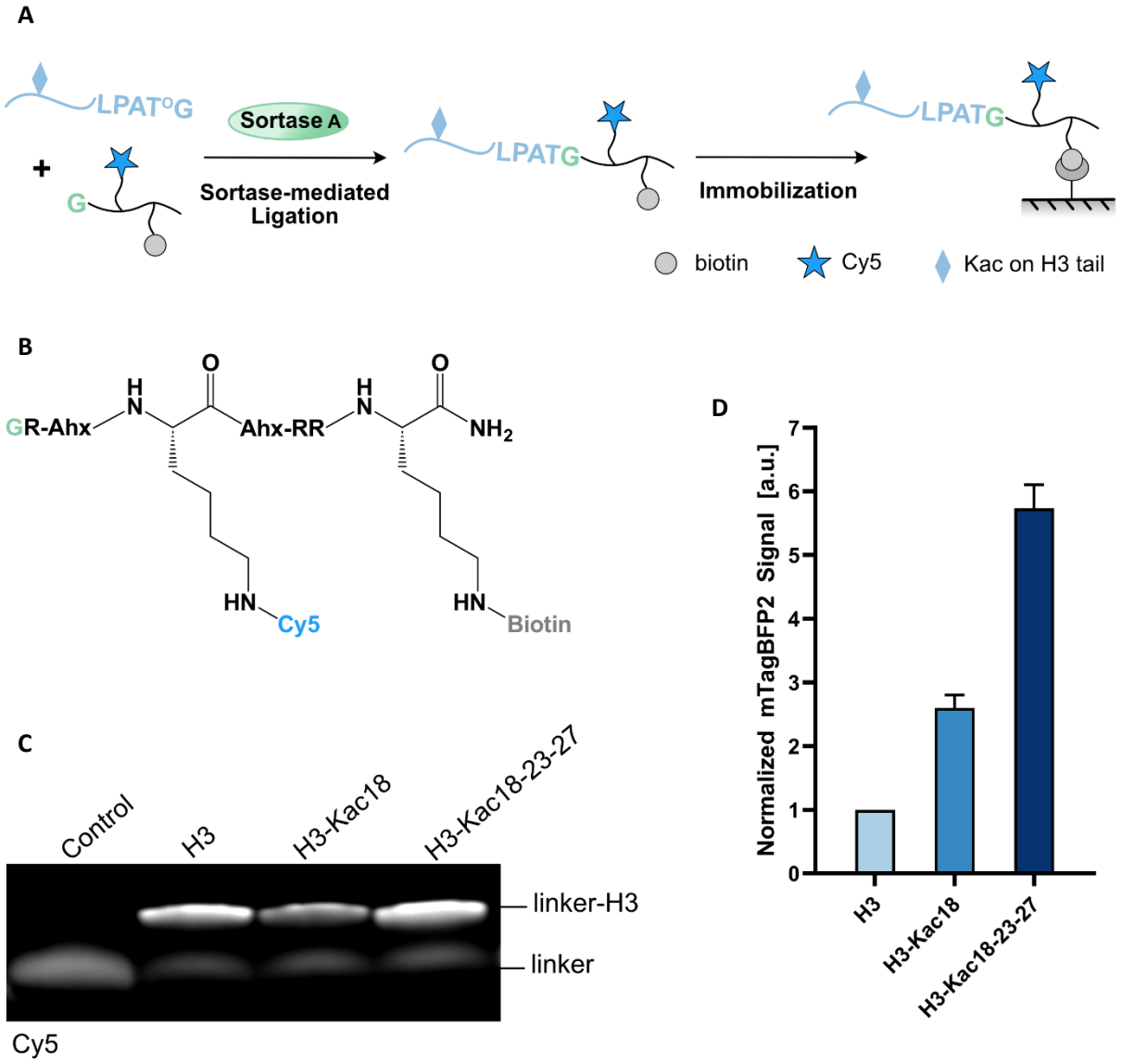


Figure 33. (A) Immobilization scheme of H3 tail peptides. (B) Structure of biotin linker peptide. (C) Biotin linker peptides were separately ligated with unmodified H3, H3-Kac18 and H3-Kac18-23-27, their products were analyzed by SDS-PAGE with 4-20% gradient gel. (D) Pull-down screening results of selected H3 tail peptides with BRD4(1/2)-mTagBFP2.

3.Discussion

3.1 Installation of acetyl-H3 tail peptides on the truncated H3 nucleosomes

By SML, 64 synthetic acetylated H3 tail peptides were installed on the truncated H3 ligation-ready nucleosomes, resulting in 64 variants of the acetyl-H3 nucleosomes. This SML-dependent semi-synthesis strategy circumvents the complex and time-consuming individual reconstitution of numerous modified nucleosome variants. Otherwise, the approximately 86% conversion rate of the ligation reaction between acetylated H3 tail peptides and truncated H3 nucleosomes was achieved within 2 hours by utilizing depsipeptides. The obtained result is consistent with the average SML ratio of 83±3% for 32 modified H3 nucleosomes generated by Pelaz et al. [126] [127]. The only drawback of this installation strategy is the introduction of an A29L mutation since a sorting motif LPXTG (X= any amino acid) is required for the wild-type SrtA-mediated ligation. Compared to the wild-type SrtA, four SrtA mutants, F40, A1-15, A1-22, and F1-21, were reported to accept more amino acids, including Ala, at the first position of the sorting motif and therefore were able to ligate H3 tail peptides containing the native H3 sequence APATG with the truncated H3. However, these engineered SrtA mutants possess low catalytic activities, resulting in a reduced ligation conversion in comparison to the wild-type SrtA [138] [139]. Notably, the F40 sortase was successfully employed to generate a series of modified nucleosomes with an enhanced SML conversion of approximately 90%. This was achieved by utilizing depsipeptides as donor substrates in solution, and using free ΔH3 instead of ligation-ready ΔH3 nucleosomes. The resulting traceless ligated H3 was modified with monoacetylation marks at K9, K14, K18, K23, and K27. These upon assemblies into nucleosomes were applied to investigate the deacetylase selectivity of several HDAC complexes, including the core CoREST complex [140] [141]. However, when the H3 tail depsipeptides strategy was applied with the pre-immobilized truncated H3 nucleosomes, the conversion rate of the F40 sortase did not exceed 50%, even with extended reaction time

[127]. This outcome indicates that SML is more efficient when carried out in solution. Furthermore, purification of the SML product cannot be achieved when the nucleosomes are immobilized in 96-well plates with a binding capacity of 120pmol/well. Assuming that the impact of the A29L mutation on the recruitment of bromodomains may not be significant, and considering the strongly decreased ligation efficiency of the F40 sortase, wild-type SrtA remained the enzyme of choice in this.

3.2 Installation of acetyl-H2A tail peptides on the truncated H2A nucleosomes

To generate acetyl-H2A-H3 nucleosomal libraries, an orthogonal ligation strategy was required. Based on the semi-synthesis strategy of modified H2A nucleosomes developed by Pelaz et al. [126] [127], protein trans-splicing (PTS) was employed to incorporate acetylated IntN-containing H2A peptides into the IntC fused truncated H2A nucleosome. Since the M86 mutant of the Ssp DnaB intein demonstrated a 60-fold higher PTS reaction rate than the unevolved split intein [113], it was utilized to install the acetyl-H2A nucleosomes. The results demonstrated that the M86 mutant maintained its high conversion ratio in the nucleosome context. However, it required additional four residues GSIE inserted into the native H2A sequence. The G(-1) located upstream of IntN was necessary to prevent a thiazoline side product of the PTS that reduces splicing efficiency [142]. Moreover, the SIE incorporated downstream of IntC is required for the high efficiency of the PTS reaction. To achieve traceless H2A semi-synthesis, a special intein with no mutants at C- and N- terminus owning high PTS efficiency is needed, which needs to be generated by protein engineering.

Considering that histones H2A and H2B with similar molecular mass cannot be separated by SDS-PAGE, to verify the success of the PTS reaction, all synthetic acetylated H2A tail peptides were equipped with a TAMRA fluorophore. In addition, a western blot analysis was performed to study the conversion of the PTS reaction, which showed approximately 76% of the reaction yield. This result is in agreement with an average nucleosomal PTS conversion of 78.0 ± 8.6%,

reported by Dr. Diego Aparicio Pelaz. In his study, he generated 8 H2A nucleosomes with acetylation, methylation and phosphorylation marks, and quantified their PTS conversion based on fluorescent readouts [127]. Furthermore, the PTS has no influence on the efficiency of SML, as the conversion ratio of the SML between spliced acetyl-H2A-ΔH3 nucleosomes and acetyl-H3 tail peptides was approximately 89%. These findings indicated that the PTS is a suitable ligation strategy to assemble the acetylated H2A, and the designed orthogonal semi-synthesis strategy for acetyl-H2A-H3 nucleosomes was successful.

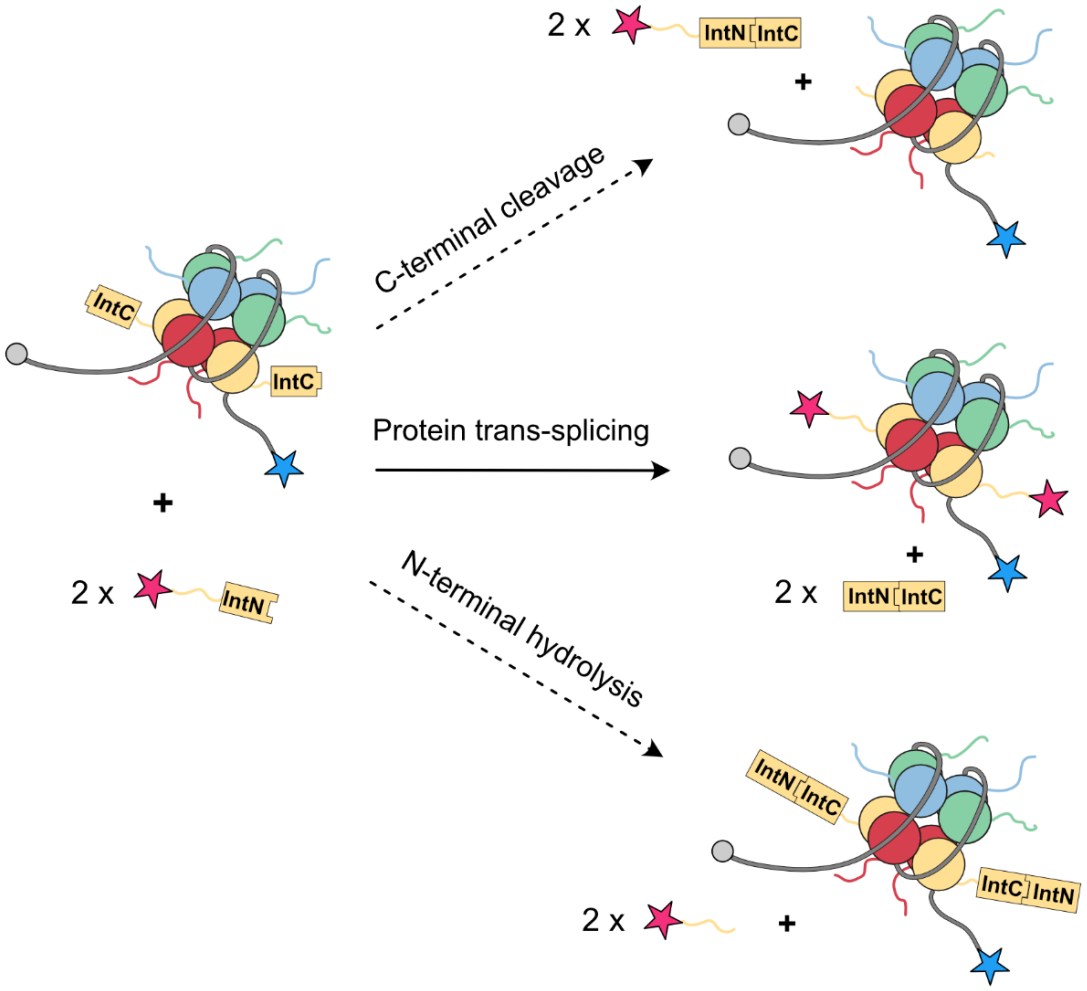


Figure 34. Possible by-products of protein trans-splicing.

In a minor fraction of the IntC-ΔH2A fusion protein, IntC was released hydrolytically without splicing of the IntC peptide. This by-reaction is named C-terminal cleavage. Another possible by-reaction is the N-terminal hydrolysis, which occurs when the branched PTS intermediate does not form rapidly, and the splicing reaction stalls at the N-extein oxy(thio)ester. With the

accumulation of activated oxy(thio)ester, N-extein can be cleaved off the intein by hydrolysis [143]. Dr. Diego Aparicio Pelaz analyzed the PTS products by LC-MS, revealing the presence of only minor C-terminal cleavage byproducts [127]. In this thesis, the PTS between acetyl-H2A tail peptides and IntC-ΔH2A-ΔH3 nucleosomes was analyzed using western blot and SDS-PAGE with Coomassie staining. The results showed neither byproducts of the C-terminal cleavage nor the N-terminal hydrolysis (Figure 34). These findings confirmed that the M86 mutant of the Ssp DnaB intein used in this study is well suited for the semi-synthesis of acetyl-H2A nucleosomes.

3.3 Recruitment of BRDs to acetyl-H3 nucleosomes

Three bromodomain proteins, BAZ2B, CREBBP, and BRD3(2), were selected and fused with the mTagBFP or TurboYFP fluorescent protein to allow investigations of their binding preference to the acetylated H3 nucleosome variants by screening of the nucleosomal libraries. The screen of BAZ2B-mTagBFP showed binding preference of BAZ2B to the nucleosomes containing H3-Kac14, which was in agreement with published data of binding studies with histone peptides. In this case, the H3-Kac14 nucleosomes reflected the findings obtained with peptides. The reported K_D value of the interaction between BAZ2B and an H3-Kac14 containing peptide was 7.6 ± 0.3 µM [45] [46]. In diacetylated nucleosomes, recruitment to H3Kac14 and H3Kac23 appeared to be additive, with the normalized mTagBFP intensity increasing to 4.7 in H3-Kac14-23 nucleosomes compared to 3.9 for the monoacetylated H3-Kac14 nucleosomes. In addition, the highest normalized BFP signals for all multi-acetylated nucleosome groups were observed with nucleosomes containing H3-Kac14-23. One possible explanation for the increased normalized mTagBFP signal of the diacetylated H3-Kac14-23 nucleosomes is that both H3-Kac14 and H3-Kac23 serve as preferred binding sites for BAZ2B. Moreover, the 8-amino acid residue distance between H3-Kac14 and H3-Kac23 offers sufficient space for simultaneous binding of BAZ2B to these two sites. Ultimately, structural investigations will be needed to prove this theory.

For CREBBP, the interaction profiling showed a subtile increased normalized TurboYFP signal for nucleosomes containing H3-Kac14 compared to the unmodified nucleosomes, indicating a potential binding preference of CREBBP for this site. Binding of CREBBP to H3-Kac14 without the nucleosome context was published previously, but the corresponding K_D value detected by isothermal titration calorimetry (ITC) was 734 ± 90.6 µM [45] [46]. This low affinity raises questions regarding the physiological relevance of this interaction. In addition, CREBBP was reported to bind H3-Kac9-14 peptides [144], which was also detected in the nucleosome context. The normalized TurboYFP signal for H3-Kac9-14 containing nucleosomes after CREBBP-TurboYFP pull-down experiments was 5.1-fold higher than that of pull-downs with unmodified nucleosomes. These findings indicated a limited impact of the nucleosomal structure on the CREBBP BRD recruitment. Furthermore, the interaction of the CREBBP BRD with another diacetylated nucleosome variant, the H3-Kac4-14 nucleosomes, resulted in an even higher normalized TurboYFP intensity than that of the H3-Kac9-14 nucleosomes. The binding interaction of CREBBP with H3 nucleosomes was gradually enhanced when more acetylation sites were included, which was indicated by the normalized TurboYFP intensities ranging from 2.6 for H3-Kac14 to 7.8 for H3-Kac4-9-14-27.

Binding of the BRD3(2) BRD to H3-Kac18 peptides was reported with a K_D value of 67.1 ± 4.5 µM [45] [46], consistent with the observed higher normalized TurboYFP signal of H3-Kac18 containing nucleosomes compared to the signal of unmodified nucleosomes after BRD3(2)-TurboYFP pull-downs. Moreover, the H3-Kac23 nucleosomes showed a slightly enhanced normalized TurboYFP signal than the control, indicating the recruitment of BRD3(2).

Similar to the recruitment profile of CREBBP, a trend of enhanced binding of BRD3(2) to nucleosomes with multiple acetylation sites was observed. This trend was evidenced by the rise in normalized TurboYFP intensity from 2.3 for H3-Kac18 nucleosomes to 5.1 for H3-Kac4-9-18-23 nucleosomes. This tendency might be attributed to a relatively weak binding interaction between CREBBP/ BRD3(2) and mono-acetylated nucleosomes. Despite BRD3(2) and CREBBP displaying binding preferences for H3-Kac18 and H3-Kac14 peptides, the corresponding binding affinity (K_D values) remained in the range of double-digit and triple-

digit micromolar values [45] [46]. Consequently, the presence of multiple acetylation sites is necessary for higher binding interaction with BRD3(2) or CREBBP. The hexa-acetylated H3 nucleosomes did not exhibit the highest normalized fluorescence intensities among the screening results of these three BRDs. When the number of acetylation sites increased from four to five, the anticipated rise in normalized fluorescence signal was not observed. Hence, tri or tetra-acetylated nucleosomes might be optimal for recruiting those BRDs to nucleosomal H3.

3.4 Recruitment of BRD4(1/2) under the influence of acetyl-H2A-H3 nucleosomes

The investigation into the recruitment of the tandem BRD of BRD4(1/2) started with the screening of acetyl-H3 nucleosomal libraries. BRD4(1/2) and BRD3(2) showed similar binding preferences. This finding suggests that acetylated H3 nucleosomes may have a weaker effect on BRD4(1) than on BRD4(2) recruitment. Moreover, this screening revealed a binding preference of BRD4(1/2) to H3-Kac18 nucleosomes, as indicated by a normalized mTagBFP2 signal that was slightly stronger than that of the unmodified nucleosomes. Binding of BRD4(1/2) was further enhanced by the inclusion of one or two additional acetylation sites in addition to the monoacetylated H3-Kac18 nucleosomes. Specifically, the normalized mTagBFP2 intensity increased to 3.6-fold for the H3-Kac18-27 nucleosomes and continued to rise to 4.3-fold for the H3-Kac18-23-27 nucleosomes. The observed trend of increasing normalized mTagBFP2 signal correlating with the number of acetylation sites was statistically significant for acetylation marks below three, which may represent a preferred number of acetylation marks for BRD4(1/2) recruitments.

The observed correlation of binding interaction, which is enhanced with the number of acetylation sites, could be explained by a direct preferred interaction of BRD4(1/2) with the nucleosomes containing multiple acetylation sites. Alternatively, these multi-acetylated H3 tails are reduced in positive charge, which might weaken the electrostatic interactions with

the negatively charged DNA. As a result, the nucleosome structure becomes more relaxed, potentially facilitating the accessibility of BRD4(1/2) to the H3 tail. In the first scenario, the observed trend should be observable not just with nucleosomes but also with acetylated H3 tail peptide variants. To further investigate these two scenarios, pull-down reactions were carried out using BRD4(1/2)-mTagBFP2 with immobilized H3-Kac18 and H3-Kac18-23-27 peptides, using unmodified H3 tails peptides as control. The results revealed a significantly increased binding of BRD4(1/2) to H3-Kac18-23-27 compared to H3-Kac18 peptides, indicating that the additional acetylation sites directly contribute to the recruitment of BRD4(1/2).

Building on the aforementioned discovery, the recruitment of BRD4(1/2) to acetylated H2A nucleosomal libraries also needs to be taken into account. The interaction profiling revealed a binding preference of BRD4(1/2) for H2A-Kac5-9 containing nucleosome, which might suggest that the spacer sequence QGG between the H2A-Kac5 and H2A-Kac9 allows simultaneous binding of both sites. Addition of one more acetylation mark at position H2A-K15 resulted in the H2A-Kac5-9-15 containing nucleosome with the highest normalized BFP intensity, indicating the strongest interaction with BRD4(1/2) among the 16 acetylated H2A nucleosomes. Therefore, H2A-Kac5-9 and H2A-Kac5-9-15 were subsequently combined with 64 variants of H3 acetylation sites to generate a library 128 of acetyl-H2A-H3 nucleosomes, which were applied to further explore the impact of potential crosstalk between acetylated H2A and H3 on the recruitments of BRD4(1/2).

Similar to the screening results obtained from the acetyl-H3 nucleosomes with BRD4(1/2), the binding preference of BRD4(1/2) to the H3-Kac18 nucleosomes was also observed under the influence of H2A-Kac5-9 or H2A-Kac5-9-15. Moreover, the normalized mTagBFP2 intensity showed enhanced recruitment that correlated with the addition of H3 acetylation sites in the H2A-Kac5-9 or H2A-Kac5-9-15 containing nucleosomal libraries, implying that the multiple acetylation marks on different histone tails could further facilitate the recruitments of BRD4(1/2). The normalized mTagBFP2 increment when acetyl-H3 nucleosomes combined with H2A-Kac5-9 or H2A-Kac5-9-15 shown in Figure 36 will be discussed later.

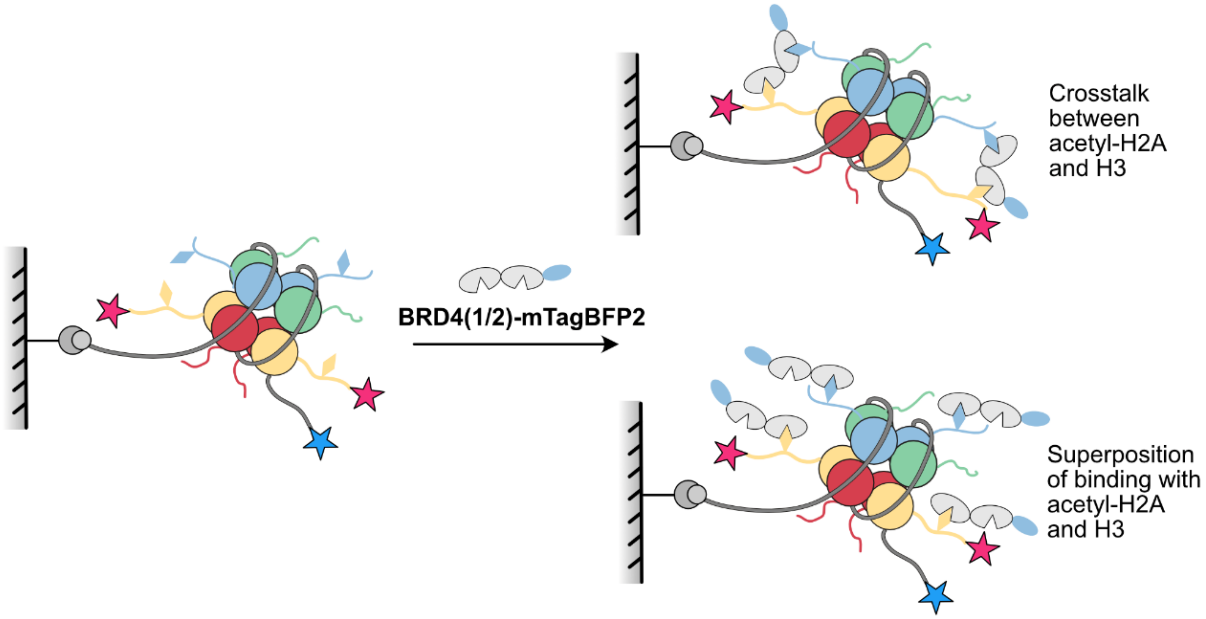


Figure 35. Possible binding models of BRD4(1/2) with acetyl-H2A-H3 nucleosomes.

In the H2A-Kac5-9-15 nucleosomal libraries, the highest normalized mTagBFP2 intensities of di-, tri-, tetra-, and penta-acetylated H3 groups were detected when nucleosomes contained H3-Kac18-27, H3-Kac9-18-27, H3-Kac9-18-23-27, and H3-Kac9-14-18-23-37, respectively. While the nucleosome separately including H3-Kac14-18, H3-Kac18-23-27, H3-Kac9-14-18-23, and H3-Kac9-14-18-23-27 with the combination of H2A-Kac5-9 showed the strongest normalized mTagBFP2 signal in their corresponding di-, tri-, tetra-, and penta-acetylated H3 groups. These nucleosome variants also exhibited higher mTagBFP2 intensity in their corresponding groups within the acetyl-H3 nucleosomal libraries, indicating that acetylated H3 may play a more dominant role in the binding of BRD4(1/2) for the acetyl-H2A-H3 nucleosomes.

Compared to the acetyl-H3 nucleosomes, most acetyl-H2A-H3 nucleosomes exhibited enhanced binding to BRD4(1/2), as indicated by the increase in the normalized BFP signal. This could be attributed to the additional binding of BRD4(1/2) to acetylated H2A, or alternatively, crosstalk between acetylated H2A and H3 exists in some nucleosome variants, facilitating binding to BRD4(1/2) (Figure 35). In the former case, the increase in the normalized mTagBFP2 intensity should be similar for all acetyl-H3 nucleosome variants. Thus, to quantify the enhanced binding of BRD4(1/2) to the corresponding acetyl-H2A-H3 nucleosomes under the

influence of H2A-Kac5-9 or H2A-Kac5-9-15, the normalized mTagBFP2 intensities of H2A-Kac5-9/acetyl-H3 or H2A-Kac5-9-15/acetyl-H3 nucleosomes were divided by the normalized mTagBFP2 intensities of the corresponding acetyl-H3 nucleosome variants without the impact of H2A, yielding the normalized mTagBFP2 increment (Figure 36). The results indicated that the unmodified H3 nucleosome, as control, showed higher or similar normalized mTagBFP2 increment than most of the acetyl-H3 nucleosomes, implying that most nucleosome variants do not exhibit crosstalk between acetylated H2A and H3. Remarkably, the nucleosomes containing H3-Kac14-18 exhibited a 1.8-fold increase in the normalized mTagBFP2 intensity when combined with H2A-Kac-5-9, which was higher than that of the control. However, it should be noted that the acetyl-H3 nucleosomal libraries and acetyl-H2A-H3 nucleosomal libraries are assembled from two types of ligation-ready nucleosomes, which may impact the observed screening results. Therefore, it cannot be postulated with certainty if there is crosstalk between the H2A-Kac5-9 and H3-Kac14-18. To further validate the interaction between these acetylation sites on H2A and H3 tails, it may be necessary to perform the pull-down assay screening without the influence of nucleosome context. Therefore, a construct that can link both acetylated H2A and H3 tail peptides needs to be designed and generated for the pull-down assay with selected BRDs.

Notably, the highest binding affinity of first and second BRDs of BRD3 and BRD4 to histone tail peptides was observed with H4 acetylation marks, specifically H4-Kac5-8 and H4-Kac20 [45] [46]. It is intriguing to investigate the recruitment profile of BRDs to acetylated H4 variants within nucleosome context. However, up to now, the ligation-ready H4 nucleosomes cannot be established. To address this gap, it is necessary to develop a viable semi-synthesis strategy for incorporating modified H4 tail peptides into truncated H4 nucleosomes.

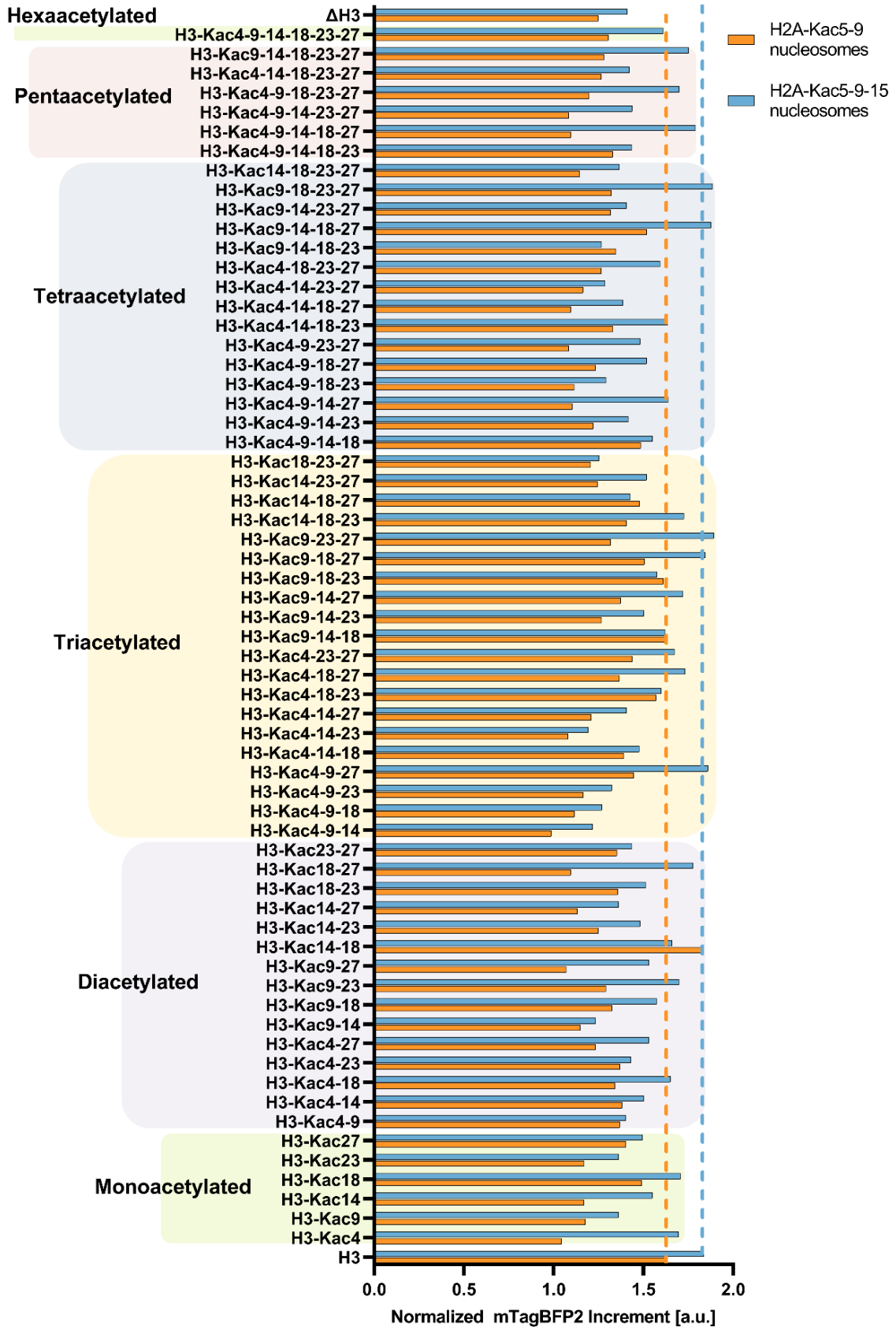


Figure 36. Normalized mTagBFP2 increment when acetyl-H3 nucleosomes combined with H2A-Kac5-9 or H2A-Kac5-9-15.

3.5 Semi-synthesis of a new construct to link both acetylated H2A and H3 tail peptides

To further investigate if the findings of this thesis could also be observed without nucleosome context, a construct was generated to link both synthetic H2A and H3 tail peptides. The construct was composed of an intein C fused IntC-MBP-S6 protein and an N-terminal pentaglycine containing linker peptide, which can be linked together by PCL and orthogonally ligate with acetylated H2A and H3 tail peptides by PTS and SML. The results demonstrated that the three ligation reactions involved 4-fragment assembly was feasible. However, several improvements may be required to obtain a high-yield end product. After the purification of IntC-MBP-S6, a small amount of fused protein lacking the IntC fragment remained, which may cause a by-product with the missing H2A tail peptides. Therefore, it may be worthwhile to remove the impurity, which may be achieved by size exclusion-chromatography. Otherwise, the completion of the PCL reaction could not be confirmed, since the separation of reactant and product was impossible with SDS-PAGE due to their similar molecular masses. To address the issue, one solution would be to label the linker peptide with a fluorophore using an additional coupling step with Fmoc-Lys(5/6-FAM)-OH building block. Moreover, an equipped fluorophore would support the verification of the success of the SML ligation. Besides, the total conversion ratio of SML and PCL reactions did not exceed 50%, and the SML reaction proceeded mostly in the reverse direction after two hours even with depsipeptides. Hence, the reaction conditions need to be further optimized, and the corresponding pull-down assay could be performed afterwards.

3.6 Nucleosomal libraries

In this thesis, the acetylated nucleosomal libraries for investigating the BRDs recruitments were successfully established. The establishment of these acetylated nucleosomes was based on the nucleosome semi-synthesis strategy developed by Pelaz et al. [126] [127], which utilized sortase-mediated ligation and protein trans-splicing. Another nucleosomal library was

generated by Nguyen et al., termed DNA-barcoded nucleosome library. They applied expressed protein ligation to generate modified histone variants and then assembled them into a uniquely barcoded DNA-containing nucleosome library. This DNA-barcoded nucleosome library was employed to profile the binding preference of various chromatin factors [145]. Compared to the synthesis strategy of DNA-barcoded nucleosomes, the acetylated nucleosomal libraries utilized the ligation-ready nucleosomes to simplify the intricate process of synthesizing and reconstituting various modified histones individually. However, it is important to note that ligation-ready H2B and H4 nucleosomes are currently not accessible in contrast to the DNA-barcoded library. The readout of the biochemical assay with DNA-barcoded nucleosomal libraries was displayed by chromatin immunoprecipitation, followed by multiplexed DNA-barcode next-generation sequencing (in vitro CHIP-seq), which is highly sensitive. One potential challenge that arises with this approach is establishing correlations across datasets derived from in vitro CHIP-seq using different antibodies [145]. In contrast, the acetylated nucleosome libraries assessed the BRDs recruitment profiles by quantifying corresponding fluorescence intensities. This methodology facilitates a direct comparison of binding interactions involving the same BRD with distinct nucleosome variants. In quantitative terms, the DNA-barcoded nucleosome library, encompassing 54 modified nucleosomes so far [145], is smaller in scale compared to the acetylated nucleosome libraries. In this thesis, a large number of nucleosomes in combination with mono- and multi-acetylation marks were generated and analyzed, which included 64 acetylated H3 nucleosomes, 16 acetylated H2A nucleosomes and 128 nucleosomes with acetylated H2-H3 variants. On the other hand, the DNA-barcoded nucleosomal library focused not only on acetylation marks but also on trimethylation and ubiquitination marks, and included the investigations of H2B and H4 tails [145]. Furthermore, The DNA-barcoded library appears to be limited to binding assays. In contrast, the ligation-ready nucleosome libraries, designed with different variants separately immobilized in wells of microtiter plates, could be used for other assays, such as enzymatic assay [126] [127].

Overall, the nucleosomal libraries take into account the impact of acetylated lysine residues on the nucleosome structure, thereby enhancing the likelihood of obtaining results that

closely reflect the in vivo situation. Furthermore, the generated acetyl-H2A-H3 nucleosomal libraries allow the investigation of potential crosstalk between acetylation sites on H2A and H3 tails. However, the quantity of immobilized nucleosomes was limited by the binding capacity of the streptavidin-coated plates, which consequently restricted the discovery of whole interactions. Additionally, as nucleosomes were reused more frequently, an increase in the amount of free DNA was detected within the nucleosomes. Since the immobilized nucleosomes were quantified by detecting the Cy5 signal, the presence of free DNA may perturb the precise quantification of immobilized nucleosomes. In conclusion, the generated acetylated nucleosomal library is a powerful pre-screening tool for investigating the recruitment of BRDs. The resulting preference binding sites can be further confirmed using techniques such as isothermal titration calorimetry (ITC) or nuclear magnetic resonance (NMR) spectroscopy.

4. Summary and outlook

The semi-synthesis of the acetyl-H3 nucleosomal libraries was successful, which bypasses the complex and time-consuming synthesis procedure of nucleosomes with high quantity variants by using the SML to install the 64 variants of acetyl-H3 tail peptides into the ligation-ready truncated H3 nucleosomes. Based on that, an orthogonal ligation strategy was applied to incorporate the acetylated H2A and H3 tail peptides separately through SML and PTS into the IntC-ΔH2A-H3 nucleosomes, generating acetyl-H2A-H3 nucleosomal libraries. However, the generated acetylated nucleosomes are not entirely traceless, which may potentially affect experimental results. Therefore, a traceless semi-synthesis strategy for the modified nucleosomal library may be necessary. Furthermore, an additional chemoselective ligation strategy for the semi-synthesis of histones H4 and H2B should be considered for further exploration. According to previous investigations with histone tail peptides, H4-Kac marks are the preferred BRD4(1/2) binding sites [45] [46]. Therefore, the generation of acetylated H4 nucleosomal libraries is meaningful for the further investigation of BRDs recruitments. In addition to acetylating the nucleosomes, this semi-synthesis strategy could also be employed to synthesize numerous nucleosomes carrying different PTMs, such as the nucleosomal libraries generated by Pelaz et al. [126] [127]. Overall, the nucleosome semi-synthesis strategy is a powerful tool for investigating the interactions between histone marks with corresponding writers, readers, and erasers.

In this study, the installed acetyl-H3 and acetyl-H2A-H3 nucleosomes were used to investigate the recruitment of several BRDs. The results obtained from pull-down assay screening are consistent mostly with previously published results of binding studies with histone peptides, implying the reliability of the generated nucleosomal libraries. Moreover, the screening results revealed previously unknown binding interactions, such as the binding preference of BAZ2B, CREBBP, BRD3(2), and BRD4(1/2) to nucleosomes respectively containing H3-Kac14-23, H3-Kac4-14, H3-Kac18-23-27, and H3-Kac18. Furthermore, the H3-Kac14-18 nucleosomes showed a significant increase in the normalized fluorescence intensity when combined with

H2A-Kac5-9. To further explore if the observed binding was limited to the nucleosome context, a construct was synthesized to link both acetyl-H2A and H3 tail peptides. However, due to the low yield of the final product, this construct and corresponding reaction conditions are required for optimization. In addition, the K_D value of these observed binding events could be measured by ITC or NMR using peptides and corresponding proteins, which could on the other side verifies whether these results were influenced by the nucleosomal context. In a further study, the focus of research could shift to the impacts of acetylated H2B and H4 nucleosomes on different BRDs. Additional PTMs, such as phosphorylation, methylation, could be included to investigate the recruitments of BRDs under combinatorial different histone marks.

5. Materials

5.1 Equipment

Equipment	Supplier
Autoclave	Systec
Centrifuge (Heraeus Megafuge™ 16R)	Thermo Fisher Scientific
Centrifuge (MiniSpin® plus)	Eppendorf
Centrifuge (Sorvall™ RC6 Plus)	Thermo Fisher Scientific
Electroporator (MicroPulser™)	BioRad
Freeze dryer (Alpha™ 2-4 LDplus)	Martin Christ
Frenchpress (EmulsiFlex-C5)	Avestin
HPLC (Analytical, LC-10A)	Shimadzu
HPLC (Preparative, ProStar 210)	Varian
Imager (ChemiDoc MP)	Bio-Rad
Incubator (Multitron Strandard)	Infors HT
Incubator (BB 150 CO_2)	Thermo Fisher Scientific
LC-MS (LCMS-2020)	Shimadzu
Nano Drop (ND-1000)	PeqLab
Peptide Synthesizer (Syro 1)	MultiSynTech
Plate reader (Infinite® 200 PRO)	TECAN
pH Meter (PB-11)	Sartorius
Rotary evaporator (RV 10 Basic)	IKA
Shaker (Unimax 1010)	Heidolph
Thermomixer (Comfort and Compact)	Eppendorf
Ultrasonic cleaner	VWR
Pipette	Eppendorf

5.2 Chemicals

Amino acid derivatives	Supplier
Fmoc-Ala-OH (A), Fmoc-Asp(OtBu)-OH (D), Fmoc-Arg-(Pbf)-OH (R), Fmoc-Cys(Trt)-OH (C), Fmoc-Gln(Trt)-OH (Q), Fmoc-Gly-OH (G), Fmoc-Ile-OH (I), Fmoc-Lys(Boc)-OH (K), Fmoc-Leu-OH (L), Fmoc-Pro-OH (P), Fmoc-Ser(tBu)-OH (S), Fmoc- Thr (tBu)- OH (T)	GL Biochem
Building blocks	
Fmoc-Ala-Thr (psiMe,Mepro)-OH (AT), Fmoc-Gln(Trt)-Thr (psiMe, Mepro)-OH (QT), Fmoc- Ser (tBu)-Thr (PsiMe, Mepro)-OH (ST) Boc-Gly-OH	Novabiochem
Fmoc- Lys (Ac)- OH (Kac)	Bachem
Fmoc-Lys-OAll*HCl, Fmoc- 6- Ahx- OH	Iris Biotech GmbHF
5(6)-Carboxytetramethylrhodamine (TAMRA)	Carbosynth
Resin	
TentaGel HL RAM, Fmoc Ala TentaGel S Trt resin	Rapp Polymere

All other chemicals are purchased from Rapp Polymere, Novabiochem, Sigma-Aldrich, Roth, Merck, Iris Biotech GmbH, VWR, Carbolution, Biosolve, Promega, Carl Roth, Bachem, Th. Geyer, J. T. Baker, and FisherChem.

5.3 Consumables

Counsumables	Supplier
Amersham™ Hybond® Low Fluorescence 0.2µm PVDF Blotting Membrane	Cytiva WhatmanTM
Amicon Ultra centrifugal filters (MWCO: 30k Da)	Merck
HisTrap™ HP (5ml)	cytiva
Page Ruler Prestained Protein Ladder	Thermo Scientific
Pierce™ Streptavidin Coated High Capacity Plates, Black, 96-Well	Thermo Fischer
Mini-PROTEAN TGX Gel, 4-20%	Bio-Rad
Ni-NTA Agarose	Qiagen

5.4 Antibodies

Antibody (epitope, isotype)	Provider	Catalog
Primary antibody		
Histone H3 (100 Cterm, rabbit-poly-IgG)	abcam	ab1791
Histone 2A ChIP (rabbit-poly-IgG)	abcam	ab88770
Secondary antibody		
IRDye 800 CW donkey anti-rabbit (lgG)	Li-cor	925-32213

5.5 Bacteria

E.Coli Strains	Supplier
BL21(DE3) Competent *E. Coli*	Agilent
XL1-Blue Competent *E. Coli*	Agilent

5.6 Vector

Vector Backbone	Gene	Supplier
Vector for gene synthesis		
pMA-T	BRD4-mTagBFP2	Invitrogen by Thermo Fischer Scientific
pMA-RQ	IntC-MBP-S6	Invitrogen by Thermo Fischer Scientific
Vector for protein expression		
pET21a	IntC-MBP-S6	Merck Millipore
pET28a	BRD4-mTagBFP2	Merck Millipore

5.7 Kit

Kit	Supplier
GeneJET Gel Extraction Kit	Thermo Scientific
PeqGOLD Plasmid Miniprep Kit I	peQlab a VWR brand

5.8 Enzyme and corresponding buffer

Enzyme	Supplier
FastDigest EcoRI	Thermo Scientific
FastDigest HindIII	Thermo Scientific
FastDigest NdeI	Thermo Scientific
T4 DNA Ligase	Thermo Scientific
Buffer	
10x FastDigest Buffer	Thermo Scientific
10x T4 Ligase buffer	Thermo Scientific

5.9 Primers

Sequencing Primers	Sequence	Supplier
T7 promotor fw	TAATACGACTCACTATAGGG	Sigma Aldrich
T7 Terminator rv	GCTAGTTATTGCTCAGCGG	Sigma Aldrich

5.10 Software

Software	Application
Affinity Designer	Figures editing
ApE - A plasmid Editor	DNA sequence analysis
ChemDraw Professional	Chemical structures drawing
Galaxie Chromatography Data System	Preparative HPLC
GraphPad Prism	Data analysis
Image Lab	SDS-PAGE and Western Blot analysis
LabSolution	LCMS and analytical HPLC analysis
Microsoft Office	Data analysis and text edition
PyMOL	Protein structure showing
Zotero	Literature citation

6 Methods

6.1 Peptide synthesis

All peptides were synthesized based on the solid phase peptide synthesis (SPPS) using the Fmoc strategy. The Automated SPPS was performed using the Syro1 peptide synthesizer with the following described cycles. Prior to each coupling, the N-terminal protection Fmoc group was removed by treatment with 20% or 40% piperidine in DMF for several minutes, followed by washing away the excess piperidine with DMF. Next, the double coupling of corresponding amino acids was performed. The first coupling proceeded with HATU as an activator and NMM as a base for 40 minutes, followed by the second coupling with the same amino acids activated by Oxyma Pure and DIC. Excess reagents were removed by washing with DMF several times. These procedures were repeated for the subsequent amino acid coupling until the desired peptide synthesis was completed.

6.1.1 Synthesis of acetyl-H3 tail depsipeptides

The depsipeptide formation was based on the Steglich esterification [135]. TentaGel® HL RAM resin was treated with 20% piperidine (3 x 2 min) to remove the Fmoc group from the N-terminus and then washed with DMF (5 x 1 min). After that, the resin was coupled with glycolic acid (10 eq), activated by HATU (9 eq) in NMM (400 mM in DMF) for 15 min. This coupling procedure was then repeated once. After that, the resin was washed with DMF (3 x 1 min). Fmoc-Thr(OtBu)-OH (20 eq) was added in a round bottom flask equipped with a $CaCl_2$ drying tube and then dissolved with anhydrous DCM. After the mixture was cooled to 0 °C, DIC (10 eq) was added and stirred for 30 min. Solvents were removed with a rotatory evaporator, and the residue was dissolved with DMF. The appropriate amount of mixture and a small amount of DMAP were added to the glycolic acid coupled resin and shaken for 10 min. This step was repeated until the whole mixture was reacted with resin. Finally, the resin was washed with

DMF (3 x 1 min) and DCM (3 x 1 min), dried under vacuum, and stored at 4°C.

The synthetic depsipeptides were applied for the synthesis of acetyl-H3 tail peptides. Three building blocks, Fmoc-Ala-Thr(psiMe, Mepro)-OH, Fmoc-Ser(tBu)-Thr(psiMe, Mepro)-OH, and Fmoc-Gln(Trt)-Thr(psiMe, Mepro)-OH (3 eq), were coupled manually with HATU (2.7 eq) and NMM (400 mM in DMF) for 30 min while shaking. Other amino acids were coupled by automated SPPS.

6.1.2 Synthesis of acetyl-H2A tail peptides

Fmoc Ala TentaGel® S Trt resin was used for the H2A peptides synthesis. All the amino acids were coupled by automated SPPS. The N-terminal fluorophore was manually coupled with 5(6)-TAMRA (2.5 equivalent) and PyOxim (2.5 equivalent) in DMF with overnight shaking.

6.1.3 Synthesis of linker peptide

The linker peptide was synthesized manually based on SPPS. Firstly, TentaGel HL RAM resin was deprotected with 20% piperidine (3 x 1 min) and washed with DMF (5 x 1 min). Next, the resin was coupled with corresponding amino acids (3 eq), activated by HATU (2.7 eq), in NMM (400 mM in DMF) for 30 min. After that, the resin was washed with DMF (3 x 1 min). These steps were repeated until all the amino acids were coupled.

Alloc Deprotection

The removal of the alloc protection group was performed with borane dimethylamine complex (6 eq) and palladium-tetrakis(triphenylphosphine) (10 mol%) in anhydrous DCM under argon [146]. The $Me_2NH{\cdot}BH_3$ was first dissolved in anhydrous DCM, added to the resin under argon, and agitated for a few minutes. The $Pd(PPh_3)_4$ was then dissolved in anhydrous DCM and added to the resin under agitation for 20 minutes. Afterward, the resin was washed with anhydrous DCM (3 x 30 s). The deprotection procedure was repeated once. After the second deprotection,

the resin was washed with DCM (8x 30 s), TFA (0.2%) in DCM (2x 60 s), DCM (5x 30 s), DIPEA (5%) in DCM (3x 60 s) and DCM (5x 30 s).

Bromoacetylation

Bromoacetic anhydride (5 eq) was weighed into one Falcon previously flushed with argon, dissolved in 2 mL anhydrous DMF, and added to the resin. The mixture was then shaken for 30 min at room temperature. This step was repeated once. After that, the resin was washed with DMF (3 x 1 min) and DCM (3 x 1 min). Subsequently, the resin was dried under a vacuum and stored at 4°C.

Peptide-coenzyme A Conjugation

Bromoacetylated peptides and CoA trilithium salt (5 eq) were solved separately in 1 M triethylammonium bicarbonate buffer at pH 8.5, and then these two solutions were mixed slowly. The reaction mixture was stirred under argon at room temperature for 24 h.

6.2 Test cleavage

In order to verify the success of the synthesis, a small amount of resin was added to 50 µL test cleavage cocktail TFA/TIPS/H_2O (18:1:1) and heated at 50 °C for 15 minutes. The mixture was diluted with 1 mL LCMS water and centrifuged to remove some unsolvable resin matrix. The supernatant was analyzed by LCMS with a linear A: B gradient (A: 100% H_2O, 0.1% Formic acid; B: 80% ACN, 20% H_2O, 0.1% Formic acid) from 5% to 95% B with a flow rate of 0.2 mL/min in 20 min.

6.3 Peptide cleavage

Peptide cleavage from the resin and removal of side-chain protecting groups were achieved by treating the resin with 10 mL cleavage cocktail TFA/phenol/TIPS/H_2O (85:5:5:5). The resin was firstly incubated with 7 mL cleavage cocktail for 3 h with shaking, followed by incubation with

another 1.5 mL cleavage cocktail for 30 min, twice. After that, the resin was washed with 3 mL DCM. All the fractions were collected into a 100 mL round bottom flask and concentrated with the rotatory evaporator. The residual solution was then added into 40 mL cold diethyl ether and centrifuged at -4 °C for 10 min at 4000 g. The resulting pellet was resuspended with 20 mL cold diethyl ether and recentrifuged under the same condition. The supernatant was discarded, and the pellet was dried at room temperature. Finally, the pellet was dissolved with 5 mL water and lyophilized.

6.4 Peptide purification and analysis

These crude peptides were purified by HPLC on a preparative column (Reprosil C18, 100 Å, 5 µm, 250 x 20 mm, from Dr. Maisch) with a Varian Pro Star device. The purification was carried out with a linear A: B gradient (A: 100% H_2O, 0.1% TFA; B: 80% ACN, 20% H_2O, 0.1% TFA) from 5% to 95% B with a flow rate of 13 mL/min in 60 min. LCMS was used to confirm the identity of each peptide. The purified peptides were analyzed by LCMS with a linear A: B gradient (A: 100% H_2O, 0.1% Formic acid; B: 80% ACN, 20% H_2O, 0.1% Formic acid) from 5% to 95% B with a flow rate of 0.2 mL/min in 20 min, and their purity was checked with analytical HPLC with a linear A: B gradient (A: 100% H_2O, 0.1% TFA; B: 80% ACN, 20% H_2O, 0.1% TFA) from 5% to 95% B with a flow rate of 1.5 mL/min in 30 min.

6.5 SDS-PAGE

12% polyacrylamide gels (stacking gel: 250 mM Tris-HCl, 40% ROTIPHORESE® Gel 30 (37.5:1), 0.1% SDS, 0.04% APS, 0.1% TEMED, pH 6.8; separating gel: 375 mM Tris-HCl, 12% ROTIPHORESE® Gel 30 (37.5:1), 0.1% SDS, 0.04% APS, 0.1% TEMED, pH 8.8) were prepared in advance and stored at 4 °C. 5x SDS sample buffer (250 mM Tris, 10% SDS, 30% glycerol, 0.5 M DTT, 0.02% bromophenol blue) was added to protein sample and heated for 10 min at 95°C. 6 µL PageRuler Prestained Protein Ladder, used as marker, and protein samples were pipetted into wells of the gel. The gel was previously submerged in SDS-PAGE running buffer (25 mM

Tris-HCl, 192 mM glycine, 0.1% SDS, pH 8.8) in the electrophoresis apparatus. The separation was carried out at 140 V for around 1 h.

6.6 Coomassie Staining

After SDS-PAGE, the gel was washed with warm water (3 x 3 min) on a shaker. Then, the gel was covered with Coomassie Brilliant Blue staining solution (80 mg Coomassie Brilliant Blue G250 and 3 mL 37% HCl in 1 L water), warmed up, and incubated for a half hour at room temperature with gentle agitation. The destaining of the gel was performed by incubation with water overnight on a shaker.

6.7 Western Blot

After SDS-PAGE, the gel was incubated with blot buffer (25 mM Tris-HCl, 192 mM glycine, 0.05 % SDS, 10 % methanol) for 30 minutes. PVDF membrane was previously respectively equilibrated with methanol (15 s), water (2 min), and blotting buffer (10 min). The transfer stack was prepared by sandwiching the PVDF membrane and the gel between filter papers and sponges. The transfer stack was immersed into a transfer tank containing the blotting buffer and ice bag. The western blotting was performed at 30 mA overnight at 4 °C. After that, the membrane was blocked with 5% low-fat powdered milk in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.0) for 30 minutes at room temperature. Then the membrane was incubated with the primary antibody in a dilution 1:1000 in primary incubation buffer (5% BSA in PBST (0.1 % Tween 20 (v/v) in PBS) with 0.05% NaN_3) for 1-2 h at room temperature. Subsequently, the membrane was washed with PBST (3 x 5 s, 3 x 3 min). Afterward, the secondary antibody was diluted with a ratio of 1:10000 in the secondary incubation buffer (5% low-fat powdered milk in PBST with 0.01% SDS). The membrane was then incubated with the secondary antibody solution for 1 h under a shaker at room temperature. Finally, the membrane was washed with PBST (3 x 5 s, 3 x 3 min) and PBS (3 x 5 s). Imaging was carried out with a ChemiDoc MP imaging system (Bio-Rad).

6.8 Agarose gel Electrophoresis

To prepare 1% Agarose gel, 0.5 g agarose powder was mixed with 50 mL 1x TAE (40 mM Tris-HCl, 20 mM acetic acid, 1 mM EDTA, pH 8.5) and heated for 1-3 min until the agarose is completely dissolved. After the agarose solution cooled to about 50 °C, Midori Green (dilution 1:10000) was added for DNA detection, and the agarose mixture was poured into a gel tray with the well comb in place. Once the gel was solidified, the gel was placed into the gel box, and 1xTAE as the running buffer was filled to cover the gel. Gene Ruler ^{TM} DNA Ladder Mix was loaded into the first lane of the gel as the maker, and samples were loaded into the additional wells of the gel. The agarose gel electrophoresis was performed at 100 V for about 1 h.

6.9 Cloning

Transformation of electrocompetent bacteria

1 ng plasmid and 50 µL electrocompetent bacteria (*E. coli* XL1-Blue) were mixed and transferred into a 2 mm electroporation cuvette. The bacteria were pulsed (U = 2.5 kV) with the BioRad MicroPulser^{TM} electroporator and then immediately incubated with 1 mL LB-medium for 1 h at 37°C with shaking of 300 rpm. The diluted transformation mixtures with 3 different concentrations (5% bacteria solution in LB medium, 50% bacteria solution in LB medium, 100% bacteria solution) were plated on LB agar plates with the corresponding antibiotics (ampicillin: 100 µg/mL, kanamycin: 30 µg/mL) and incubated at 37 °C overnight.

Plasmid cloning and isolation

Single colonies were picked from LB agar plates and inoculated into 6 mL LB-medium with corresponding antibiotics. After overnight incubation at 37 °C with the shaking of 180 rpm, the plasmid was isolated through the peqGOLD Plasmid Miniprep Kit I (VMR). The DNA concentration was measured with NanoDrop.

Restriction digestion

The plasmids containing desired genes or vectors pET28a/pET21a were digested with the corresponding restriction enzymes (Thermo Scientific Fast Digest enzymes). The digestion products were purified by agarose gel electrophoresis, and the desired DNAs were extracted with the GeneJET Gel Extraction Kit (Thermo Scientific).

Ligation

The ligation of DNA fragments was performed by incubating the inserts and vectors at a molar ratio of 5:1 with 1 Weiss U of T4 DNA Ligase for 2 hours at 21 °C. The T4 DNA Ligase was then inactivated by heating the mixture at 65 °C for 10 minutes. The ligation mixture was transformed into 50 µL electrocompetent bacteria (*E. coli* XL1-Blue) for plasmid cloning. The sequence of the cloned plasmid was verified by DNA sequencing through Eurofins Genomics.

6.10 Protein expression and purification

Electrocompetent bacteria (*E. coli* BL21(DE3)) were transformed with 1 ng of the plasmid containing desired DNA fragments and expression vectors for recombinant protein expression. A single colony of the transformed bacteria was inoculated into 20 mL of LB medium with the corresponding antibiotics and incubated overnight at 37 °C with shaking at 160 rpm. The precultures were inoculated into 700 mL of LB medium with corresponding antibiotics to obtain a starting OD_{600nm} of 0.1 and incubated at 37 °C with shaking at 160 rpm until the OD_{600nm} reached 0.6. Subsequently, Isopropyl -D-1-thiogalactopyranoside (IPTG) was added into the bacteria solution to induce the protein expression. The protein IntC-MBP-S6 and BRD4(1/2)-mTagBFP2 were expressed at 25 °C for 20 h after induction with 1mM IPTG.

The bacteria were harvested by centrifugation at 5000 rpm for 10 min at 4 °C. The pellets were resuspended in 40 mL lysis buffer (300 mM NaCl, 50 mM Tris-HCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 10% Glycerol, 0.1% Triton X-100, pH 7.5) and homogenized using a homogenizer (EmulsiFlex-C5, Avestin) at 4 °C. The lysate was centrifuged at 15000 rpm for 30 min at 4 °C. The

supernatant with 30 mM imidazole was incubated with 1 mL of 50/50 suspension of Ni-NTA Superflow resin (Qiagen) with rotation for 1 h at 4°C. The resin was washed with 40 mL washing buffer (300 mM NaCl, 50 mM Tris-HCl, 30 mM Imidazole, 10% Glycerol, pH 7.5), and the protein was eluted with 7 mL elution buffer (300 mM NaCl, 50 mM Tris-HCl, 300 mM Imidazole, 10% Glycerol, pH 7.5).

For large-scale protein expression and purification, the bacteria containing the target plasmid were inoculated into 4 x 1.8 L of LB medium with corresponding antibiotics. The protein expression, bacteria culture, and harvest followed the above protocol. The pellets were resuspended in 100 mL of buffer A (300 mM NaCl, 50 mM Tris-HCl, 10% Glycerin, 15 mM Imidazole, pH 7.5) and lysed with a homogenizer. The lysate was then centrifuged at 15000 rpm for 30 min at 4°C. The supernatant was loaded on a 5 mL HisTrap^{TM} HP column and purified by Immobilized metal affinity chromatography with a linear gradient from 0% to 100% buffer B (300 mM NaCl, 50 mM Tris-HCl, 10% Glycerin, 500 mM Imidazole, pH7.5) with a flow rate of 1 mL/min at 4 °C.

Fractions from purification were analyzed by SDS-PAGE and Coomassie-staining. The purified protein was dialyzed against 1 L dialysis buffer (300 mM NaCl, 50 mM Tris-HCl, pH 7.5) overnight at 4 °C and concentrated with centrifugal filters. Protein concentrations were determined using a NanoDrop.

6.11 Installation of acetyl-H3 nucleosomes

Nucleosome immobilization

Black streptavidin-coated High Capacity 96-well plates were washed with NCP buffer (10 mM Tris, 25 mM NaCl, 1 mM EDTA, 2 mM DTT, pH 7.5) (200 µl/well, 3 x 1 min). 50 µl of truncated histone H3 (ΔH3) nucleosomes (provided by Prof. Dr. Wolfgang Fischle at King Abdullah University of Science and Technology) were added into each well and incubated overnight at 4 °C with 300 rpm shaking. After that, nucleosomes were collected and stored at 4 °C for reuse.

The nucleosome-coated wells were washed with NCP buffer (200 µl/well, 5 x 1 min) and stored with 200 µl NCP buffer at 4 °C.

SML between ΔH3 nucleosomes and acetyl-H3 tail peptides

The ΔH3 nucleosome-coated wells were washed with 1x Sortase A (SrtA) buffer (50 mM Tris, 150 mM NaCl, 5 mM $CaCl_2$, pH 7.5) (200 µl/well, 2 x 1 min). Each well was then incubated with 50 µL ligation mixture (3 µM SrtA and 6 µM acetyl-H3 tail peptides in 1x SrtA buffer) for 2h at 37°C while shaking of 400 rpm. Afterward, the ligation mixture was removed. The wells were washed with NCP buffer (200 µl/well, 4 x 1 min) and stored at 4 °C in 200 µL NCP buffer.

The conversion ratio of SML was analyzed by Western Blot (WB). 40 µL 3x SDS sample buffer was previously heated to 95 °C, added into wells, and incubated with ligated nucleosomes for 5 min. After that, the denatured nucleosome sample was collected and analyzed by WB.

6.12 Installation of acetyl-H2A-H3 nucleosomes

PTS between IntC-ΔH2A-ΔH3 nucleosomes and acetyl-H2A tail peptides

IntC-ΔH2A-ΔH3 nucleosomes (provided by Prof. Dr. Wolfgang Fischle at King Abdullah University of Science and Technology) were incubated with acetyl-H2A tail peptides (10 eq) in PTS buffer (50 mM Tris-HCl, 1 mM EDTA, 300 mM NaCl, 2 mM DTT, 10% Glycerol, pH 7.0) at 5 °C for one week. The products were analyzed by SDS-PAGE with Coomassie staining and Western Blot. TAMRA fluorescence was detected before Coomassie staining.

SML between acetyl-H2A-ΔH3 nucleosomes and acetyl-H3 tail peptides

After PTS, the acetyl-H2A tail peptides ligated ΔH3 nucleosomes were immobilized on the black streptavidin-coated High Capacity 96-well plates. The SML was performed following the protocol described above.

6.13 Immobilization of H3 tail peptides

The ligation mixture was prepared by incubating 50 µM biotin linker peptide (provided by Dr. Sören Kirchgäßner at the Interfaculty Institute of Biochemistry, University of Tübingen), 200 µM H3 tail peptides, and 100 µM SrtA in 1x SrtA buffer at 37°C for 15 min. Subsequently, black streptavidin-coated High Capacity 96-well plates were washed with 1x SrtA buffer (200 µl/well, 3 x 1 min, 400 rpm) and then incubated with ligation mixture (50 µl/well) at 25°C for 30 min while shaking of 400 rpm. Afterward, the ligation mixture was removed, and wells were washed with 1x SrtA buffer (200 µl/well, 5 x 1 min, 400 rpm). The immobilized products could be cleaved by incubating with previously heated 3x SDS sample buffer for 5 min. The conversion ratio of SML was analyzed by SDS-PAGE with 4-20% gel.

6.14 Pull-down assay of nucleosomal libraries

The established acetyl-H3 and acetyl-H2A-H3 nucleosomes-coated wells were washed with 200 µl pull-down buffer (2 x 1 min) with 400 rpm shaking at RT, followed by blocking with 100 µL blocking buffer (1 mg/mL BSA in pull-down buffer) for 10 min with 400 rpm shaking. After that, the wells were incubated with BRDs with corresponding conditions (see Table 3). After incubation, wells were washed with pull-down buffer (6 x 1 min) with 400 rpm shaking at RT. Subsequently, 50 µL pull-down buffer was added to each well, and the Intensities of mTagBFP(2), TurboYFP, and Cy5 were measured using a TECAN Infinite 200 PRO plate reader with the following setting (Table 4).

The pull-down assay of acetylated nucleosomal libraries for each BRD was performed in triplicate. The mTagBFP(2) and TurboYFP signal intensities were normalized to the amount of immobilized nucleosomes by dividing them by the Cy5 signal intensities. The resulting mTagBFP(2)/Cy5 and TurboYFP/Cy5 were adjusted to represent fold-changes relative to the corresponding BRD enrichment on unmodified nucleosomes, with the mean of mTagBFP(2)/Cy5 and TurboYFP/Cy5 with unmodified nucleosomes set to 1. The mean and

standard deviation of the normalized mTagBFP(2) and TurboYFP signal for each acetylated nucleosome variant were calculated and visualized using Excel and GraphPad Prism.

Table 3. Pull-down assay conditions.

Protein	Concentration	Incubation	Pull-down buffer
BAZ2B-mTagBFP	1.5 µM	1 h, RT,	100 mM KCl, 20 mM HEPES, 20%
CREBBP-TurboYFP	13.8 µM	400 rpm	Glycerol, pH 7.9
BRD3(2)-TurboYFP	10.5 µM		
BRD4(1/2)-mTagBFP2	10.5 µM	2 h, 4°C,	150 mM NaCl, 1 mM EDTA, 20 mM
		400 rpm	HEPES, 10% glycerol, pH 7.5

Table 4. TECAN Infinite 200 PRO plate reader setting.

Fluorescence	Cy5	mTagBFP(2)	TurboYFP
Excitation Wavelength	649 nm	399 nm	515 nm
Emission Wavelength	680 nm	456 nm	550 nm
Gain		160	
Number of Flashes		25	
Integration Time		20 µs	

6.15 Generation of a construct linking acetyl-H2A and H3 tail peptides

PTS between acetyl-H2A tail peptides and IntC-MBP-S6

5 µM IntC-MBP-S6 and 15 µM acetyl-H2A tail peptides were incubated in PTS buffer at 25 °C. At selected time points, samples were collected and analyzed by SDS-PAGE with Coomassie staining.

PCL and SML

10 µM IntC-MBP-S6, 20 µM linker peptides, 40 µM acetyl-H3 tail peptides, 10 µM Sfp, and 20 µM SrtA were incubated in ligation buffer (50 mM Tris-HCl, 150 mM NaCl, 5 mM $CaCl_2$,

10 mM $MgCl_2$, pH7.5) at 25 °C. Samples were collected at selected time points and then analyzed by SDS-PAGE with Coomassie staining.

PTS, PCL, and SML involved 4-fragments ligation

10 μM IntC-MBP-S6, 30 μM acetyl-H2A tail peptides, 20 μM linker peptides, and 10 μM Sfp were incubated in ligation buffer (50 mM Tris-HCl, 1 mM EDTA, 300 mM NaCl, 5 mM $CaCl_2$, 10 mM $MgCl_2$, 10% Glycerol, 2 mM DTT, pH 7.0) at 25 °C for 20 h. After that, 40 μM acetyl-H3 tail peptides and 20 μM SrtA were added and incubated at 37 °C for 2 h. At selected time points, samples were collected and analyzed by SDS-PAGE with Coomassie staining.

7. References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, “Chromosomal DNA and Its Packaging in the Chromatin Fiber,” *Mol. Biol. Cell 4th Ed.*, 2002, Accessed: Jan. 20, 2023.

[2] J. Widom, “STRUCTURE, DYNAMICS, AND FUNCTION OF CHROMATIN IN VITRO,” *Annu. Rev. Biophys. Biomol. Struct.*, vol. 27, no. 1, pp. 285–327, Jun. 1998, doi: 10.1146/annurev.biophys.27.1.285.

[3] S. Venkatesh and J. L. Workman, “Histone exchange, chromatin structure and the regulation of transcription,” *Nat. Rev. Mol. Cell Biol.*, vol. 16, no. 3, pp. 178–189, Mar. 2015, doi: 10.1038/nrm3941.

[4] E. Li, “Chromatin modification and epigenetic reprogramming in mammalian development,” *Nat. Rev. Genet.*, vol. 3, no. 9, pp. 662–673, Sep. 2002, doi: 10.1038/nrg887.

[5] K. Luger, “Crystal structure of the nucleosome core particle at 2.8 A˚resolution,” *Nature*, pp.251–260. vol. 389, 1997.

[6] C. Bonisch and S. B. Hake, “Histone H2A variants in nucleosomes and chromatin: more or less stable?,” *Nucleic Acids Res.*, vol. 40, no. 21, pp. 10719–10741, Nov. 2012, doi: 10.1093/nar/gks865.

[7] V. Ramakrishnan, “HISTONE STRUCTURE AND THE ORGANIZATION OF THE NUCLEOSOME,” *Annu. Rev. Biophys. Biomol. Struct.*, vol. 26, no. 1, pp. 83–112, Jun. 1997, doi: 10.1146/annurev.biophys.26.1.83.

[8] J. Widom, “Toward a Unified Model of Chromatin Folding,” *Annu Rev Biophys Biophys Chem*. 1989;18:365-95. doi: 10.1146/annurev.bb.18.060189.002053. PMID: 2660830.

[9] S. Rosa and P. Shaw, “Insights into Chromatin Structure and Dynamics in Plants,” *Biology*, vol. 2, no. 4, pp. 1378–1410, Nov. 2013, doi: 10.3390/biology2041378.

[10] T. Schalch, S. Duda, D. F. Sargent, and T. J. Richmond, “X-ray structure of a tetranucleosome and its implications for the chromatin fibre,” *Nature*, vol. 436, no. 7047, pp. 138–141, Jul. 2005, doi: 10.1038/nature03686.

[11] M. P. Scheffer, M. Eltsov, and A. S. Frangakis, “Evidence for short-range helical order in the 30-nm chromatin fibers of erythrocyte nuclei,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 41, pp. 16992–16997, Oct. 2011, doi: 10.1073/pnas.1108268108.

[12] G. Felsenfeld and M. Groudine, “Controlling the double helix,” *Nature*, vol. 421, no. 6921, pp. 448–453, Jan. 2003, doi: 10.1038/nature01411.

[13] M. Lawrence, S. Daujat, and R. Schneider, “Lateral Thinking: How Histone Modifications Regulate Gene Expression,” *Trends Genet.*, vol. 32, no. 1, pp. 42–56, Jan. 2016, doi: 10.1016/j.tig.2015.10.007.

[14] V. Cavalieri, “The Expanding Constellation of Histone Post-Translational Modifications in the Epigenetic Landscape,” *Genes*, vol. 12, no. 10, p. 1596, Oct. 2021, doi: 10.3390/genes12101596.

[15] G. Millán-Zambrano, A. Burton, A. J. Bannister, and R. Schneider, “Histone post-translational modifications — cause and consequence of genome function,” *Nat. Rev. Genet.*, vol. 23, no. 9, pp. 563–580, Sep. 2022, doi: 10.1038/s41576-022-00468-7.

[16] R. Marmorstein and M.-M. Zhou, “Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition,” *Cold Spring Harb. Perspect. Biol.*, vol. 6, no. 7, pp. a018762–a018762, Jul. 2014, doi: 10.1101/cshperspect.a018762.

[17] S. Ramazi, A. Allahverdi, and J. Zahiri, “Evaluation of post-translational modifications in histone proteins: A review on histone modification defects in developmental and neurological disorders,” *J. Biosci.*, vol. 45, no. 1, p. 135, Dec. 2020, doi: 10.1007/s12038-020-00099-2.

[18] B. R. Sabari, D. Zhang, C. D. Allis, and Y. Zhao, “Metabolic regulation of gene expression through

histone acylations," *Nat. Rev. Mol. Cell Biol.*, vol. 18, no. 2, pp. 90–101, Feb. 2017, doi: 10.1038/nrm.2016.140.

[19] C. E. Barnes, D. M. English, and S. M. Cowley, "Acetylation & Co: an expanding repertoire of histone acylations regulates chromatin and transcription," *Essays Biochem.*, vol. 63, no. 1, pp. 97–107, Apr. 2019, doi: 10.1042/EBC20180061.

[20] C. A. Johnson and B. M. Turner, "Histone deacetylases: complex transducers of nuclear signals," *Semin. Cell Dev. Biol.*, vol. 10, no. 2, pp. 179–188, Apr. 1999, doi: 10.1006/scdb.1999.0299.

[21] A. J. Bannister and T. Kouzarides, "Regulation of chromatin by histone modifications," *Cell Res.*, vol. 21, no. 3, pp. 381–395, Mar. 2011, doi: 10.1038/cr.2011.22.

[22] P. Gujral, V. Mahajan, A. C. Lissaman, and A. P. Ponnampalam, "Histone acetylation and the role of histone deacetylases in normal cyclic endometrium," *Reprod. Biol. Endocrinol.*, vol. 18, no. 1, p. 84, Dec. 2020, doi: 10.1186/s12958-020-00637-5.

[23] R. Richman, L. G. Chicoine, M. P. Collini, R. G. Cook, and C. D. Allis, "Micronuclei and the cytoplasm of growing Tetrahymena contain a histone acetylase activity which is highly specific for free histone H4.," *J. Cell Biol.*, vol. 106, no. 4, pp. 1017–1026, Apr. 1988, doi: 10.1083/jcb.106.4.1017.

[24] S. Y. Roth, J. M. Denu, and C. D. Allis, "Histone Acetyltransferases," *Annu Rev Biochem*. 2001;70:81-120. doi: 10.1146/annurev.biochem.70.1.81. PMID: 11395403.

[25] R. E. Sobel, R. G. Cook, C. A. Perry, A. T. Annunziato, and C. D. Allis, "Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4.," *Proc. Natl. Acad. Sci.*, vol. 92, no. 4, pp. 1237–1241, Feb. 1995, doi: 10.1073/pnas.92.4.1237.

[26] S. C. Hodawadekar and R. Marmorstein, "Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design," *Oncogene*, vol. 26, no. 37, pp. 5528–5540, Aug. 2007, doi: 10.1038/sj.onc.1210619.

[27] X.-J. Yang and E. Seto, "HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention," *Oncogene*, vol. 26, no. 37, pp. 5310–5318, Aug. 2007, doi: 10.1038/sj.onc.1210599.

[28] P. A. Grant *et al.*, "Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex.," *Genes Dev.*, vol. 11, no. 13, pp. 1640–1650, Jul. 1997, doi: 10.1101/gad.11.13.1640.

[29] E. Seto and M. Yoshida, "Erasers of Histone Acetylation: The Histone Deacetylase Enzymes," *Cold Spring Harb. Perspect. Biol.*, vol. 6, no. 4, pp. a018713–a018713, Apr. 2014, doi: 10.1101/cshperspect.a018713.

[30] G. Milazzo *et al.*, "Histone Deacetylases (HDACs): Evolution, Specificity, Role in Transcriptional Complexes, and Pharmacological Actionability," *Genes*, vol. 11, no. 5, p. 556, May 2020, doi: 10.3390/genes11050556.

[31] P. Bheda, H. Jing, C. Wolberger, and H. Lin, "The Substrate Specificity of Sirtuins," *Annu. Rev. Biochem.*, vol. 85, no. 1, pp. 405–429, Jun. 2016, doi: 10.1146/annurev-biochem-060815-014537.

[32] H. Chen, C. Xie, Q. Chen, and S. Zhuang, "HDAC11, an emerging therapeutic target for metabolic disorders," *Front. Endocrinol.*, vol. 13, p. 989305, Oct. 2022, doi: 10.3389/fendo.2022.989305.

[33] A. J. Bannister and T. Kouzarides, "Reversing histone methylation," *Nature*, vol. 436, no. 7054, pp. 1103–1106, Aug. 2005, doi: 10.1038/nature04048.

[34] A. Barski *et al.*, "High-Resolution Profiling of Histone Methylations in the Human Genome," *Cell*, vol. 129, no. 4, pp. 823–837, May 2007, doi: 10.1016/j.cell.2007.05.009.

[35] J. L. Miller and P. A. Grant, "The Role of DNA Methylation and Histone Modifications in Transcriptional

Regulation in Humans," in *Epigenetics: Development and Disease*, T. K. Kundu, Ed., in Subcellular Biochemistry, vol. 61. Dordrecht: Springer Netherlands, 2013, pp. 289–317. doi: 10.1007/978-94-007-4525-4_13.

[36] J. Cao and Q. Yan, "Histone Ubiquitination and Deubiquitination in Transcription, DNA Damage Response, and Cancer," *Front. Oncol.*, vol. 2, 2012, doi: 10.3389/fonc.2012.00026.

[37] H. Barbour, S. Daou, M. Hendzel, and E. B. Affar, "Polycomb group-mediated histone H2A monoubiquitination in epigenome regulation and nuclear processes," *Nat. Commun.*, vol. 11, no. 1, p. 5947, Nov. 2020, doi: 10.1038/s41467-020-19722-9.

[38] H.-Y. Joo *et al.*, "Regulation of cell cycle progression and gene expression by H2A deubiquitination," *Nature*, vol. 449, no. 7165, pp. 1068–1072, Oct. 2007, doi: 10.1038/nature06256.

[39] T. Banerjee and D. Chakravarti, "A Peek into the Complex Realm of Histone Phosphorylation," *Mol. Cell. Biol.*, vol. 31, no. 24, pp. 4858–4873, Dec. 2011, doi: 10.1128/MCB.05631-11.

[40] P. Cheung, C. D. Allis, and P. Sassone-Corsi, "Signaling to Chromatin through Histone Modifications," *Cell*, vol. 103, no. 2, pp. 263–271, Oct. 2000, doi: 10.1016/S0092-8674(00)00118-5.

[41] H. Goto, Y. Yasui, EA. Nigg, M. Inagaki, "Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation," *Genes Cells*, Jan. 2002;7(1):11-7. doi: 10.1046/j.1356-9597.2001.00498.x. PMID: 11856369.

[42] C. Prigent and S. Dimitrov, "Phosphorylation of serine 10 in histone H3, what for?," *J. Cell Sci.*, vol. 116, no. 18, pp. 3677–3685, Sep. 2003, doi: 10.1242/jcs.00735.

[43] M. Labrador and V. G. Corces, "Phosphorylation of histone H3 during transcriptional activation depends on promoter structure," *Genes Dev.*, vol. 17, no. 1, pp. 43–48, Jan. 2003, doi: 10.1101/gad.1021403.

[44] S. Muller, P. Filippakopoulos, and S. Knapp, "Bromodomains as therapeutic targets," *Expert Rev. Mol. Med.*, vol. 13, p. e29, Sep. 2011, doi: 10.1017/S1462399411001992.

[45] P. Filippakopoulos *et al.*, "Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family," *Cell*, vol. 149, no. 1, pp. 214–231, Mar. 2012, doi: 10.1016/j.cell.2012.02.013.

[46] P. Filippakopoulos and S. Knapp, "The bromodomain interaction module," *FEBS Lett.*, vol. 586, no. 17, pp. 2692–2704, Aug. 2012, doi: 10.1016/j.febslet.2012.04.045.

[47] Mohd. Muddassir *et al.*, "Bromodomain and BET family proteins as epigenetic targets in cancer therapy: their degradation, present drugs, and possible PROTACs," *RSC Adv.*, vol. 11, no. 2, pp. 612–636, 2021, doi: 10.1039/D0RA07971E.

[48] D. J. Owen, "The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p," *EMBO J.*, vol. 19, no. 22, pp. 6141–6149, Nov. 2000, doi: 10.1093/emboj/19.22.6141.

[49] C. Dhalluin, J. E. Carlson, L. Zeng, C. He, A. K. Aggarwal, and M.-M. Zhou, "Structure and ligand of a histone acetyltransferase bromodomain," vol. 399, 1999.

[50] X.-J. Yang, V. V. Ogryzko, J. Nishikawa, and B. H. Howard, "A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A," vol. 382, 1996.

[51] E. Cavellán, P. Asp, P. Percipalle, and A.-K. Ö. Farrants, "The WSTF-SNF2h Chromatin Remodeling Complex Interacts with Several Nuclear Proteins in Transcription," *J. Biol. Chem.*, vol. 281, no. 24, pp. 16264–16271, Jun. 2006, doi: 10.1074/jbc.M600233200.

[52] N. A. Fairbridge, C. E. Dawe, F. H. Niri, M. K. Kooistra, K. King-Jones, and H. E. McDermid, "Cecr2 mutations causing exencephaly trigger misregulation of mesenchymal/ectodermal transcription factors," *Birt. Defects Res. A. Clin. Mol. Teratol.*, vol. 88, no. 8, pp. 619–625, Aug. 2010, doi:

10.1002/bdra.20695.

[53] L. Zeng, Q. Zhang, G. Gerona-Navarro, N. Moshkina, and M.-M. Zhou, “Structural Basis of Site-Specific Histone Recognition by the Bromodomains of Human Coactivators PCAF and CBP/p300,” *Structure*, vol. 16, no. 4, pp. 643–652, Apr. 2008, doi: 10.1016/j.str.2008.01.010.

[54] B. P. Hudson, M. A. Martinez-Yamout, H. J. Dyson, and P. E. Wright, “Solution structure and acetyl-lysine binding activity of the GCN5 bromodomain,” *J. Mol. Biol.*, vol. 304, no. 3, pp. 355–370, Dec. 2000, doi: 10.1006/jmbi.2000.4207.

[55] G. LeRoy, B. Rickards, and S. J. Flint, “The Double Bromodomain Proteins Brd2 and Brd3 Couple Histone Acetylation to Transcription,” *Mol. Cell*, vol. 30, no. 1, pp. 51–60, Apr. 2008, doi: 10.1016/j.molcel.2008.01.018.

[56] M. Ullah *et al.*, “Molecular Architecture of Quartet MOZ/MORF Histone Acetyltransferase Complexes,” *Mol. Cell. Biol.*, vol. 28, no. 22, pp. 6828–6843, Nov. 2008, doi: 10.1128/MCB.01297-08.

[57] J. Morinière *et al.*, “Cooperative binding of two acetylation marks on a histone tail by a single bromodomain,” *Nature*, vol. 461, no. 7264, pp. 664–668, Oct. 2009, doi: 10.1038/nature08397.

[58] Y. Taniguchi, “The Bromodomain and Extra-Terminal Domain (BET) Family: Functional Anatomy of BET Paralogous Proteins,” *Int. J. Mol. Sci.*, vol. 17, no. 11, p. 1849, Nov. 2016, doi: 10.3390/ijms17111849.

[59] Z. Yang, N. He, and Q. Zhou, “Brd4 Recruits P-TEFb to Chromosomes at Late Mitosis To Promote G_1 Gene Expression and Cell Cycle Progression,” *Mol. Cell. Biol.*, vol. 28, no. 3, pp. 967–976, Feb. 2008, doi: 10.1128/MCB.01020-07.

[60] T. Kanno, Y. Kanno, R. M. Siegel, M. K. Jang, M. J. Lenardo, and K. Ozato, “Selective Recognition of Acetylated Histones by Bromodomain Proteins Visualized in Living Cells,” *Mol. Cell*, vol. 13, no. 1, pp. 33–43, Jan. 2004, doi: 10.1016/S1097-2765(03)00482-9.

[61] A. Dey, F. Chitsaz, A. Abbasi, T. Misteli, and K. Ozato, “The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis,” *Proc. Natl. Acad. Sci.*, vol. 100, no. 15, pp. 8758–8763, Jul. 2003, doi: 10.1073/pnas.1433065100.

[62] J. M. Lamonica *et al.*, “Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes,” *Proc. Natl. Acad. Sci.*, vol. 108, no. 22, May 2011, doi: 10.1073/pnas.1102140108.

[63] J. L. Morgado-Pascual, S. Rayego-Mateos, L. Tejedor, B. Suarez-Alvarez, and M. Ruiz-Ortega, “Bromodomain and Extraterminal Proteins as Novel Epigenetic Targets for Renal Diseases,” *Front. Pharmacol.*, vol. 10, p. 1315, Nov. 2019, doi: 10.3389/fphar.2019.01315.

[64] Y. Xu and C. R. Vakoc, “Brd4 is on the move during inflammation,” *Trends Cell Biol.*, vol. 24, no. 11, pp. 615–616, Nov. 2014, doi: 10.1016/j.tcb.2014.09.005.

[65] J. E. Delmore *et al.*, “BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc,” *Cell*, vol. 146, no. 6, pp. 904–917, Sep. 2011, doi: 10.1016/j.cell.2011.08.017.

[66] Q. Feng *et al.*, “An epigenomic approach to therapy for tamoxifen-resistant breast cancer,” *Cell Res.*, vol. 24, no. 7, pp. 809–819, Jul. 2014, doi: 10.1038/cr.2014.71.

[67] C. French, “NUT midline carcinoma,” *Nat. Rev. Cancer*, vol. 14, no. 3, pp. 149–150, Mar. 2014, doi: 10.1038/nrc3659.

[68] Z. Zou *et al.*, “Brd4 maintains constitutively active NF-κB in cancer cells by binding to acetylated RelA,” *Oncogene*, vol. 33, no. 18, pp. 2395–2404, May 2014, doi: 10.1038/onc.2013.179.

[69] E. Kalkhoven, “CBP and p300: HATs for different occasions,” *Biochem. Pharmacol.*, vol. 68, no. 6, pp. 1145–1155, Sep. 2004, doi: 10.1016/j.bcp.2004.03.045.

[70] Y. Cai *et al.*, “Identification of New Subunits of the Multiprotein Mammalian TRRAP/TIP60-containing

Histone Acetyltransferase Complex," *J. Biol. Chem.*, vol. 278, no. 44, pp. 42733–42736, Oct. 2003, doi: 10.1074/jbc.C300389200.

[71] H. Huang, I. Rambaldi, E. Daniels, and M. Featherstone, "Expression of theWdr9 gene and protein products during mouse development," *Dev. Dyn.*, vol. 227, no. 4, pp. 608–614, Aug. 2003, doi: 10.1002/dvdy.10344.

[72] P. Müller, D. Kuttenkeuler, V. Gesellchen, M. P. Zeidler, and M. Boutros, "Identification of JAK/STAT signalling components by genome-wide RNA interference," *Nature*, vol. 436, no. 7052, pp. 871–875, Aug. 2005, doi: 10.1038/nature03869.

[73] A. Podcheko *et al.*, "Identification of a WD40 Repeat-Containing Isoform of PHIP as a Novel Regulator of β-Cell Growth and Survival," *Mol. Cell. Biol.*, vol. 27, no. 18, pp. 6484–6496, Sep. 2007, doi: 10.1128/MCB.02409-06.

[74] A. Ito *et al.*, "p300/CBP-mediated p53 acetylation is commonly induced by p53-activating agents and inhibited by MDM2," *EMBO J.*, vol. 20, no. 6, pp. 1331–1340, Mar. 2001, doi: 10.1093/emboj/20.6.1331.

[75] B. M. Dancy and P. A. Cole, "Protein Lysine Acetylation by p300/CBP," *Chem. Rev.*, vol. 115, no. 6, pp. 2419–2452, Mar. 2015, doi: 10.1021/cr500452k.

[76] V. V. Ogryzko, R. L. Schiltz, V. Russanova, B. H. Howard, and Y. Nakatani, "The Transcriptional Coactivators p300 and CBP Are Histone Acetyltransferases," *Cell*, vol. 87, no. 5, pp. 953–959, Nov. 1996, doi: 10.1016/S0092-8674(00)82001-2.

[77] N. G. Iyer, H. Özdag, and C. Caldas, "p300/CBP and cancer," *Oncogene*, vol. 23, no. 24, pp. 4225–4231, May 2004, doi: 10.1038/sj.onc.1207118.

[78] M. D. Kaeser, A. Aslanian, M.-Q. Dong, J. R. Yates, and B. M. Emerson, "BRD7, a Novel PBAF-specific SWI/SNF Subunit, Is Required for Target Gene Activation and Repression in Embryonic Stem Cells," *J. Biol. Chem.*, vol. 283, no. 47, pp. 32254–32263, Nov. 2008, doi: 10.1074/jbc.M806061200.

[79] K. Laue *et al.*, "The multidomain protein Brpf1 binds histones and is required for Hox gene expression and segmental identity," *Development*, vol. 135, no. 11, pp. 1935–1946, Jun. 2008, doi: 10.1242/dev.017160.

[80] M. Ciró *et al.*, "ATAD2 Is a Novel Cofactor for MYC, Overexpressed and Amplified in Aggressive Tumors," *Cancer Res.*, vol. 69, no. 21, pp. 8491–8498, Nov. 2009, doi: 10.1158/0008-5472.CAN-09-2131.

[81] H. Sun *et al.*, "Solution structure of BRD7 bromodomain and its interaction with acetylated peptides from histone H3 and H4," *Biochem. Biophys. Res. Commun.*, vol. 358, no. 2, pp. 435–441, Jun. 2007, doi: 10.1016/j.bbrc.2007.04.139.

[82] M. H. Jones, N. Hamana, J. Nezu, and M. Shimane, "A Novel Family of Bromodomain Genes," *Genomics*, vol. 63, no. 1, pp. 40–45, Jan. 2000, doi: 10.1006/geno.1999.6071.

[83] D. B. Bloch *et al.*, "Sp110 Localizes to the PML-Sp100 Nuclear Body and May Function as a Nuclear Hormone Receptor Transcriptional Coactivator," *Mol. Cell. Biol.*, vol. 20, no. 16, pp. 6138–6146, Aug. 2000, doi: 10.1128/MCB.20.16.6138-6146.2000.

[84] K. Khetchoumian *et al.*, "TIF1δ, a Novel HP1-interacting Member of the Transcriptional Intermediary Factor 1 (TIF1) Family Expressed by Elongating Spermatids," *J. Biol. Chem.*, vol. 279, no. 46, pp. 48329–48341, Nov. 2004, doi: 10.1074/jbc.M404779200.

[85] Y. Dou *et al.*, "Physical Association and Coordinate Function of the H3 K4 Methyltransferase MLL1 and the H4 K16 Acetyltransferase MOF," *Cell*, vol. 121, no. 6, pp. 873–885, Jun. 2005, doi: 10.1016/j.cell.2005.04.031.

[86] H. M. Rowe *et al.*, "KAP1 controls endogenous retroviruses in embryonic stem cells," *Nature*, vol. 463,

no. 7278, pp. 237–240, Jan. 2010, doi: 10.1038/nature08674.

[87] D. A. Wassarman and F. Sauer, “TAFII250: a transcription toolbox,” *J Cell Sci*, Aug. 2001, 114(Pt 16):2895-902. doi: 10.1242/jcs.114.16.2895. PMID: 11686293.

[88] H. Masselink and R. Bernards, “The adenovirus E1A binding protein BS69 is a corepressor of transcription through recruitment of N-CoR,” *Oncogene*, vol. 19, no. 12, pp. 1538–1546, Mar. 2000, doi: 10.1038/sj.onc.1203421.

[89] A. G. Li, L. G. Piluso, X. Cai, B. J. Gadd, A. G. Ladurner, and X. Liu, “An Acetylation Switch in p53 Mediates Holo-TFIID Recruitment,” *Mol. Cell*, vol. 28, no. 3, pp. 408–421, Nov. 2007, doi: 10.1016/j.molcel.2007.09.006.

[90] G. D. Gregory *et al.*, “Mammalian ASH1L Is a Histone Methyltransferase That Occupies the Transcribed Region of Active Genes,” *Mol. Cell. Biol.*, vol. 27, no. 24, pp. 8466–8479, Dec. 2007, doi: 10.1128/MCB.00993-07.

[91] H. K N *et al.*, “Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing,” *Nat. Genet.*, vol. 37, no. 3, pp. 254–264, Mar. 2005, doi: 10.1038/ng1516.

[92] Y. Xue *et al.*, “The human SWI/SNF-B chromatin-remodeling complex is related to yeast Rsc and localizes at kinetochores of mitotic chromosomes,” *Proc. Natl. Acad. Sci.*, vol. 97, no. 24, pp. 13015–13020, Nov. 2000, doi: 10.1073/pnas.240208597.

[93] R. Chandrasekaran and M. Thompson, “Polybromo-1-bromodomains bind histone H3 at specific acetyl-lysine positions,” *Biochem. Biophys. Res. Commun.*, vol. 355, no. 3, pp. 661–666, Apr. 2007, doi: 10.1016/j.bbrc.2007.01.193.

[94] C. Kupitz, R. Chandrasekaran, and M. Thompson, “Kinetic analysis of acetylation-dependent Pb1 bromodomain–histone interactions,” *Biophys. Chem.*, vol. 136, no. 1, pp. 7–12, Jul. 2008, doi: 10.1016/j.bpc.2008.03.011.

[95] M. Schnölzer and S. B. H. Kent, “Constructing Proteins by Dovetailing Unprotected Synthetic Peptides: Backbone-Engineered HIV Protease,” *Science*, vol. 256, no. 5054, pp. 221–225, Apr. 1992, doi: 10.1126/science.1566069.

[96] D. R. Englebretsen, B. C. Garnham, D. A. Bergman, and P. F. Alewood, “A novel thioether linker: Chemical synthesis of a HIV-1 protease analogue by thioether ligation,” *Tetrahedron Lett.*, vol. 36, no. 48, pp. 8871–8874, Nov. 1995, doi: 10.1016/0040-4039(95)01843-7.

[97] T. Durek and C. F. W. Becker, “Protein semi-synthesis: New proteins for functional and structural studies,” *Biomol. Eng.*, vol. 22, no. 5–6, pp. 153–172, Dec. 2005, doi: 10.1016/j.bioeng.2005.07.004.

[98] B. Bader, K. Kuhn, D. J. Owen, H. Waldmann, and A. Wittinghofer, “Bioorganic synthesis of lipid-modi®ed proteins for the study of signal transduction,” vol. 403, 2000.

[99] R. Reents, M. Wagner, J. Kuhlmann, H. Waldmann, “Synthesis and application of fluorescence-labeled Ras-proteins for live-cell imaging,” *Angew Chem Int Ed Engl*, May. 2004, 10;43(20):2711-4. doi: 10.1002/anie.200353265. PMID: 18629997.

[100] P. E. Dawson, T. W. Muir, I. Clark-Lewis, and S. B. H. Kent, “Synthesis of Proteins by Native Chemical Ligation,” *Science*, vol. 266, no. 5186, pp. 776–779, Nov. 1994, doi: 10.1126/science.7973629.

[101] C. P. R. Hackenberger and D. Schwarzer, “Chemoselective Ligation and Modification Strategies for Peptides and Proteins,” *Angew. Chem. Int. Ed.*, vol. 47, no. 52, pp. 10030–10074, Dec. 2008, doi: 10.1002/anie.200801313.

[102] P. E. Dawson and S. B. H. Kent, “Synthesis of Native Proteins by Chemical Ligation,” *Annu. Rev. Biochem.*, vol. 69, no. 1, pp. 923–960, Jun. 2000, doi: 10.1146/annurev.biochem.69.1.923.

[103] E. M. Sletten and C. R. Bertozzi, “Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of

Functionality," *Angew. Chem. Int. Ed.*, vol. 48, no. 38, pp. 6974–6998, Sep. 2009, doi: 10.1002/anie.200900942.

[104] T. W. Muir, D. Sondhi, and P. A. Cole, "Expressed protein ligation: A general method for protein engineering," *Proc. Natl. Acad. Sci.*, vol. 95, no. 12, pp. 6705–6710, Jun. 1998, doi: 10.1073/pnas.95.12.6705.

[105] V. Muralidharan and T. W. Muir, "Protein ligation: an enabling technology for the biophysical analysis of proteins," *Nat. Methods*, vol. 3, no. 6, pp. 429–438, Jun. 2006, doi: 10.1038/nmeth886.

[106] L. Saleh and F. B. Perler, "Protein splicingIn Cis andIn Trans," *Chem. Rec.*, vol. 6, no. 4, pp. 183–193, 2006, doi: 10.1002/tcr.20082.

[107] M. Q. Xu and F. B. Perler, "The mechanism of protein splicing and its modulation by mutation.," *EMBO J.*, vol. 15, no. 19, pp. 5146–5153, Oct. 1996, doi: 10.1002/j.1460-2075.1996.tb00898.x.

[108] R. David, M. P. O. Richter, and A. G. Beck-Sickinger, "Expressed protein ligation: Method and applications," *Eur. J. Biochem.*, vol. 271, no. 4, pp. 663–677, Jan. 2004, doi: 10.1111/j.1432-1033.2004.03978.x.

[109] S. Chong, L. M. Vence, and C. A. Hirvonen, "Single-column purification of free recombinant proteins using a self-cleavable aÃnity tag derived from a protein splicing element," 1997.

[110] J. C. J. Matern *et al.*, "Ligation of Synthetic Peptides to Proteins Using Semisynthetic Protein trans-Splicing," in *Site-Specific Protein Labeling*, A. Gautier and M. J. Hinner, Eds., in Methods in Molecular Biology, vol. 1266. New York, NY: Springer New York, 2015, pp. 129–143. doi: 10.1007/978-1-4939-2272-7_9.

[111] Y. Ding *et al.*, "Crystal Structure of a Mini-intein Reveals a Conserved Catalytic Module Involved in Side Chain Cyclization of Asparagine during Protein Splicing," *J. Biol. Chem.*, vol. 278, no. 40, pp. 39133–39142, Oct. 2003, doi: 10.1074/jbc.M306197200.

[112] G. Volkmann and H. Iwaï, "Protein trans-splicing and its use in structural biology: opportunities and limitations," *Mol. Biosyst.*, vol. 6, no. 11, p. 2110, 2010, doi: 10.1039/c0mb00034e.

[113] J. H. Appleby-Tagoe, I. V. Thiel, Y. Wang, Y. Wang, H. D. Mootz, and X.-Q. Liu, "Highly Efficient and More General cis- and trans-Splicing Inteins through Sequential Directed Evolution," *J. Biol. Chem.*, vol. 286, no. 39, pp. 34440–34447, Sep. 2011, doi: 10.1074/jbc.M111.277350.

[114] W. Sun, J. Yang, and X.-Q. Liu, "Synthetic Two-piece and Three-piece Split Inteins for Protein trans-Splicing," *J. Biol. Chem.*, vol. 279, no. 34, pp. 35281–35286, Aug. 2004, doi: 10.1074/jbc.M405491200.

[115] A. W. Jacobitz, M. D. Kattke, J. Wereszczynski, and R. T. Clubb, "Sortase Transpeptidases: Structural Biology and Catalytic Mechanism," in *Advances in Protein Chemistry and Structural Biology*, Elsevier, 2017, pp. 223–264. doi: 10.1016/bs.apcsb.2017.04.008.

[116] L. A. Marraffini, A. C. DeDent, and O. Schneewind, "Sortases and the Art of Anchoring Proteins to the Envelopes of Gram-Positive Bacteria," *Microbiol. Mol. Biol. Rev.*, vol. 70, no. 1, pp. 192–221, Mar. 2006, doi: 10.1128/MMBR.70.1.192-221.2006.

[117] H. Ton-That, G. Liu, S. K. Mazmanian, K. F. Faull, and O. Schneewind, "Purification and characterization of sortase, the transpeptidase that cleaves surface proteins of *Staphylococcus aureus* at the LPXTG motif," *Proc. Natl. Acad. Sci.*, vol. 96, no. 22, pp. 12424–12429, Oct. 1999, doi: 10.1073/pnas.96.22.12424.

[118] C. P. Guimaraes *et al.*, "Site-specific C-terminal and internal loop labeling of proteins using sortase-mediated reactions," *Nat. Protoc.*, vol. 8, no. 9, pp. 1787–1799, Sep. 2013, doi: 10.1038/nprot.2013.101.

[119] S. Tsukiji, T. Nagamune, "Sortase-mediated ligation: a gift from Gram-positive bacteria to protein

engineering," *Chembiochem*, Mar. 2009, 23;10(5):787-98. doi: 10.1002/cbic.200800724. PMID: 19199328.

[120] D.J. Williamson, M.A. Fascione, M.E. Webb, and W.B. Turnbull, "Efficient N-Terminal Labeling of Proteins by Use of Sortase," *Angew. Chem. Int. Ed.*, 2012, 51: 9377-9380.

[121] Y. Yamamura, H. Hirakawa, S. Yamaguchi, and T. Nagamune, "Enhancement of sortase A-mediated protein ligation by inducing a β-hairpin structure around the ligation site," *Chem. Commun.*, vol. 47, no. 16, p. 4742, 2011, doi: 10.1039/c0cc05334a.

[122] R. David Row, T. J. Roark, M. C. Philip, L. L. Perkins, and J. M. Antos, "Enhancing the efficiency of sortase–mediated ligations through nickel–peptide complex formation," *Chem. Commun.*, vol. 51, no. 63, pp. 12548–12551, 2015, doi: 10.1039/C5CC04657B.

[123]H. H. Wang, B. Altun, K. Nwe, A. Tsourkas, "Proximity-Based Sortase-Mediated Ligation," *Angew. Chem. Int. Ed*, 2017, 56, 5349.

[124] C. Freund and D. Schwarzer, "Engineered Sortases in Peptide and Protein Chemistry," *ChemBioChem*, vol. 22, no. 8, pp. 1347–1356, Apr. 2021, doi: 10.1002/cbic.202000745.

[125]L. Schmohl and D. Schwarzer, "Sortase-mediated ligations for the site-specific modification of proteins," *Curr. Opin. Chem. Biol.*, vol. 22, pp. 122–128, Oct. 2014, doi: 10.1016/j.cbpa.2014.09.020.

[126] D. Aparicio Pelaz *et al.*, "Examining histone modification crosstalk using immobilized libraries established from ligation-ready nucleosomes," *Chem. Sci.*, vol. 11, no. 34, pp. 9218–9225, 2020, doi: 10.1039/D0SC03407J.

[127] D. Aparicio Pelaz, "Establishing a chemical toolbox for designer nucleosomes and its applications for analysis of chromatin binding proteins," PhD Thesis, Eberhard Karls Universität Tübingen, 2017.

[128] R. H. Lambalot *et al.*, "A new enzyme superfamily — the phosphopantetheinyl transferases," *Chem. Biol.*, vol. 3, no. 11, pp. 923–936, 1996, doi: 10.1016/S1074-5521(96)90181-7.

[129] J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, and M. D. Burkart, "The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life," *Nat Prod Rep*, vol. 31, no. 1, pp. 61–108, 2014, doi: 10.1039/C3NP70054B.

[130] T. Nuijens and M. Schmidt, Eds., *Enzyme-Mediated Ligation Methods*, vol. 2012. in Methods in Molecular Biology, vol. 2012. New York, NY: Springer New York, 2019. doi: 10.1007/978-1-4939-9546-2.

[131] D. Rabuka, "Chemoenzymatic methods for site-specific protein modification," *Curr. Opin. Chem. Biol.*, vol. 14, no. 6, pp. 790–796, Dec. 2010, doi: 10.1016/j.cbpa.2010.09.020.

[132] Z. Zhou *et al.*, "Genetically Encoded Short Peptide Tags for Orthogonal Protein Labeling by Sfp and AcpS Phosphopantetheinyl Transferases," *ACS Chem. Biol.*, vol. 2, no. 5, pp. 337–346, May 2007, doi: 10.1021/cb700054k.

[133] J. Yin *et al.*, "Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase," *Proc. Natl. Acad. Sci.*, vol. 102, no. 44, pp. 15815–15820, Nov. 2005, doi: 10.1073/pnas.0507705102.

[134] L. Schmohl and D. Schwarzer, "Chemo-enzymatic three-fragment assembly of semisynthetic proteins," *J. Pept. Sci.*, 2014, 20: 145-151.

[135] B. Neises and W. Steglich, "Simple Method for the Esterification of Carboxylic Acids," *Angew. Chem. Int. Ed. Engl.*, vol. 17, no. 7, pp. 522–524, Jul. 1978, doi: 10.1002/anie.197805221.

[136] S. Kirchgäßner *et al.*, "Synthesis, Biochemical Characterization, and Genetic Encoding of a 1,2,4-Triazole Amino Acid as an Acetyllysine Mimic for Bromodomains of the BET Family," *Angew. Chem. Int. Ed.*, vol. 62, no. 12, p. e202215460, Mar. 2023, doi: 10.1002/anie.202215460.

[137] S. Kirchgäßner, “Synthesis, biochemical characterization and genetic encoding of acetyl-lysine mimicking amino acids for bromodomains and deacetylases,” PhD Thesis, Eberhard Karls Universität Tübingen, 2017.

[138] K. Piotukh *et al.*, “Directed Evolution of Sortase A Mutants with Altered Substrate Selectivity Profiles,” *J. Am. Chem. Soc.*, vol. 133, no. 44, pp. 17536–17539, Nov. 2011, doi: 10.1021/ja205630g.

[139] L. Schmohl, J. Bierlmeier, F. Gerth, C. Freund, and D. Schwarzer, “Engineering sortase A by screening a second-generation library using phage display: Directed evolution of sortase A,” *J. Pept. Sci.*, vol. 23, no. 7–8, pp. 631–635, Jul. 2017, doi: 10.1002/psc.2980.

[140] M. Wu, D. Hayward, J. H. Kalin, Y. Song, J. W. Schwabe, and P. A. Cole, “Lysine-14 acetylation of histone H3 in chromatin confers resistance to the deacetylase and demethylase activities of an epigenetic silencing complex,” *eLife*, vol. 7, p. e37231, Jun. 2018, doi: 10.7554/eLife.37231.

[141] Z. A. Wang *et al.*, “Diverse nucleosome Site-Selectivity among histone deacetylase complexes,” *eLife*, vol. 9, p. e57663, Jun. 2020, doi: 10.7554/eLife.57663.

[142] C. Ludwig, D. Schwarzer, and H. D. Mootz, “Interaction Studies and Alanine Scanning Analysis of a Semi-synthetic Split Intein Reveal Thiazoline Ring Formation from an Intermediate of the Protein Splicing Reaction,” *J. Biol. Chem.*, vol. 283, no. 37, pp. 25264–25272, Sep. 2008, doi: 10.1074/jbc.M802972200.

[143] R. E. Thompson and T. W. Muir, “Chemoenzymatic Semisynthesis of Proteins,” *Chem. Rev.*, vol. 120, no. 6, pp. 3051–3126, Mar. 2020, doi: 10.1021/acs.chemrev.9b00450.

[144] A. Kouskouti and I. Talianidis, “Histone modifications defining active genes persist after transcriptional and mitotic inactivation,” *EMBO J.*, vol. 24, no. 2, pp. 347–357, Jan. 2005, doi: 10.1038/sj.emboj.7600516.

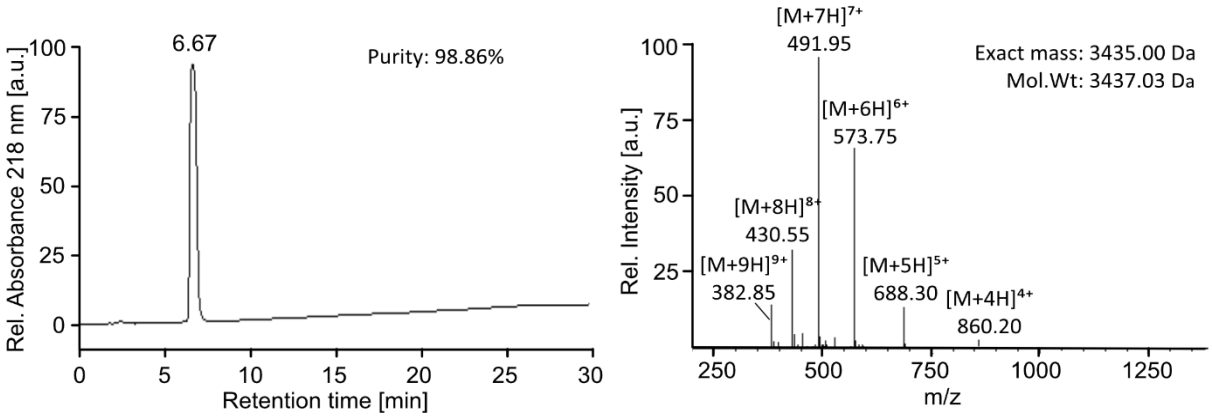
[145]U. T. T. Nguyen *et al.*, “Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries,” *Nat. Methods*, vol. 11, no. 8, pp. 834–840, Aug. 2014, doi: 10.1038/nmeth.3022.

[146] P. Gomez-Martinez, M. Dessolin, F. Guibé, and F. Albericio, “N^{α}-Alloc temporary protection in solid-phase peptide synthesis. The use of amine–borane complexes as allyl group scavengers,” *J. Chem. Soc., Perkin Trans. 1*, pp. 2871-2874, 1999, doi: 10.1039/A906025A.

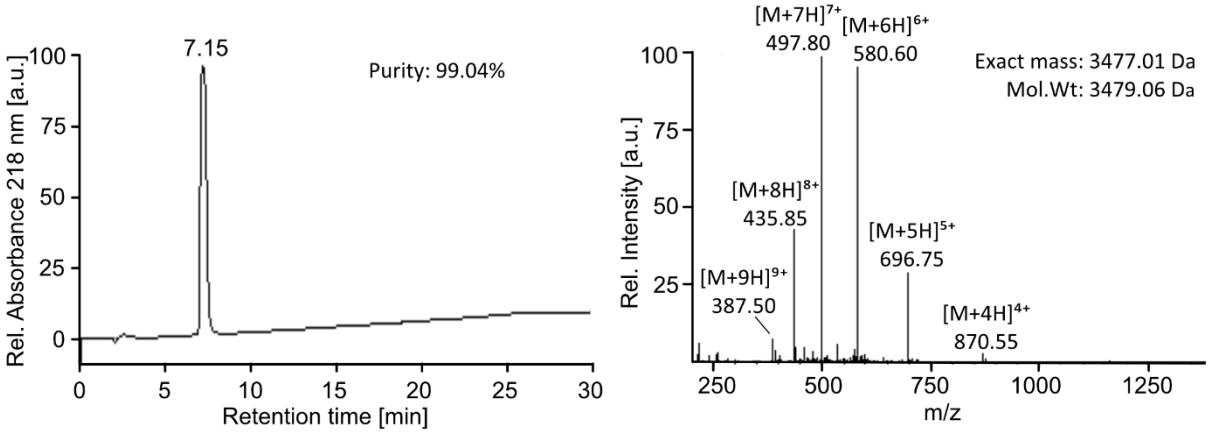
8.Appendixes

8.1 Analytical results of acetyl-H3 tail peptides

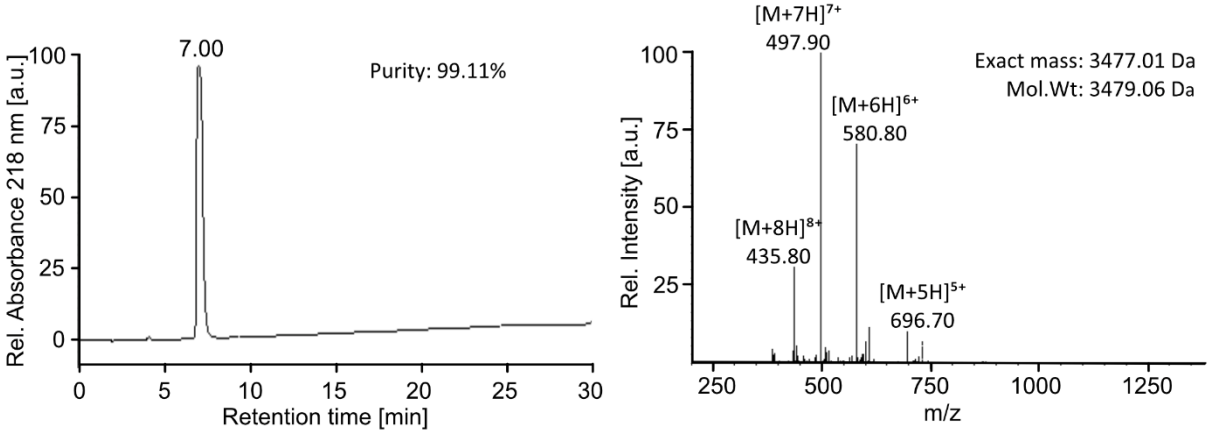
No.1 (H3)



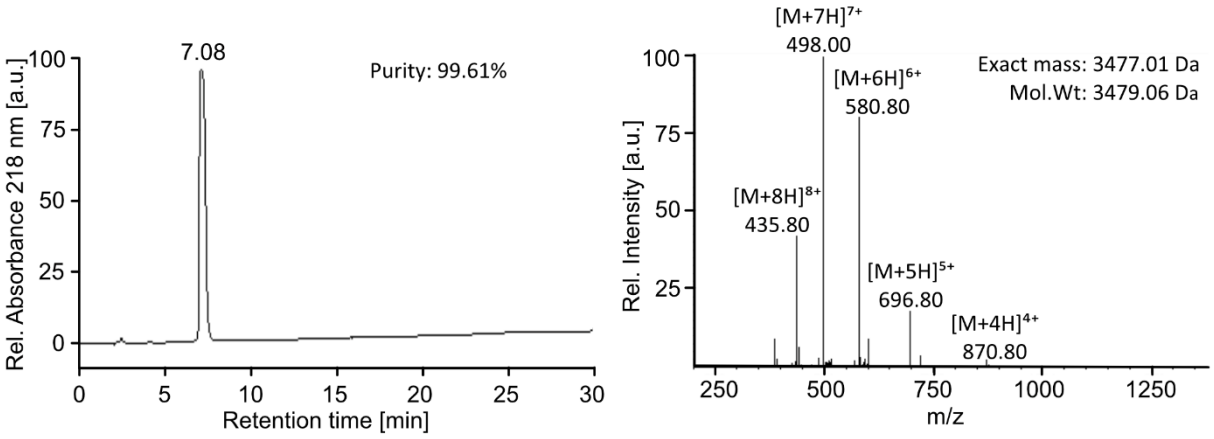
No.2 (H3-Kac4)



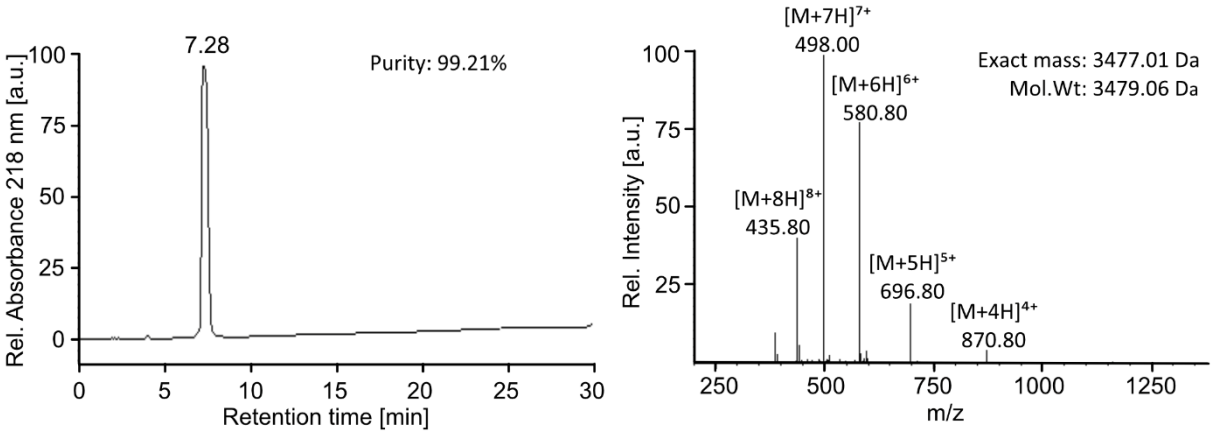
No.3 (H3-Kac9)



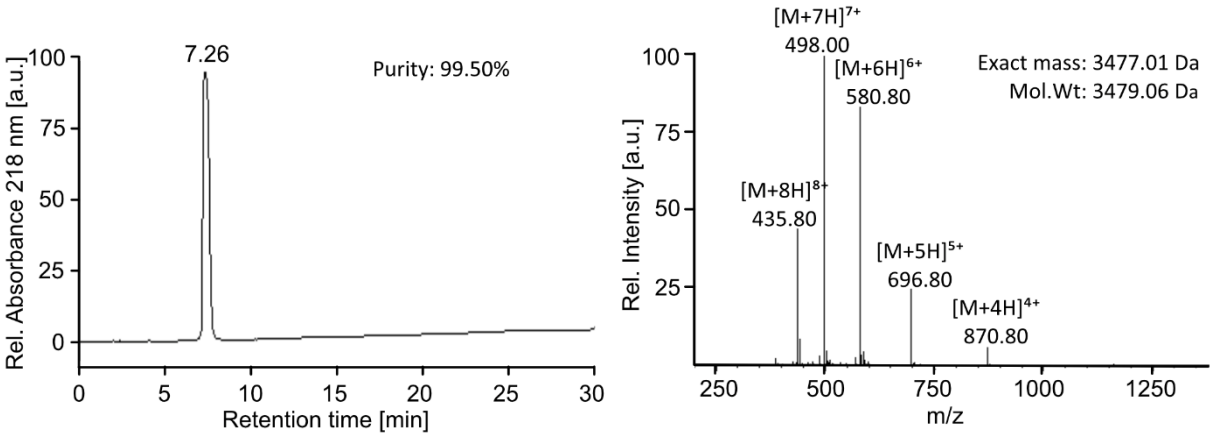
No.4 (H3-Kac9)



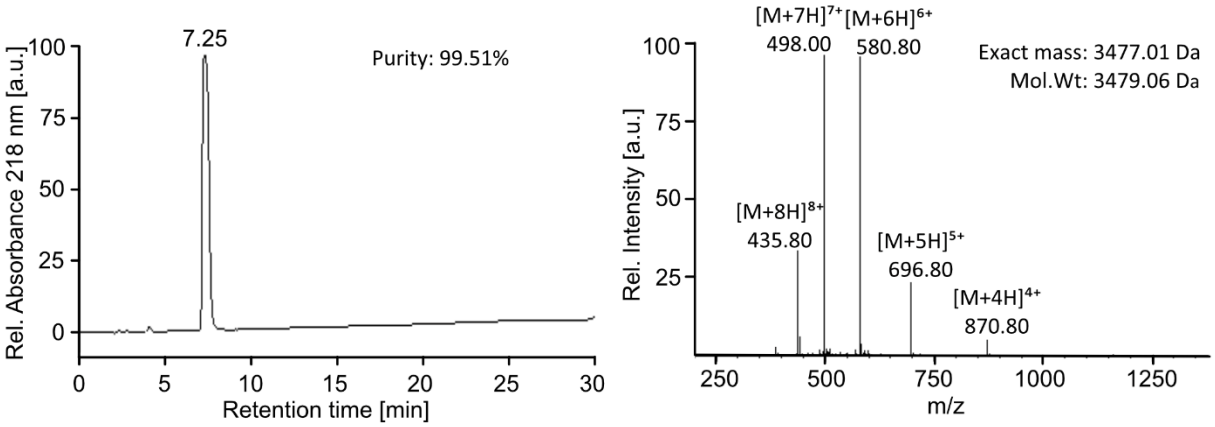
No.5 (H3-Kac18)



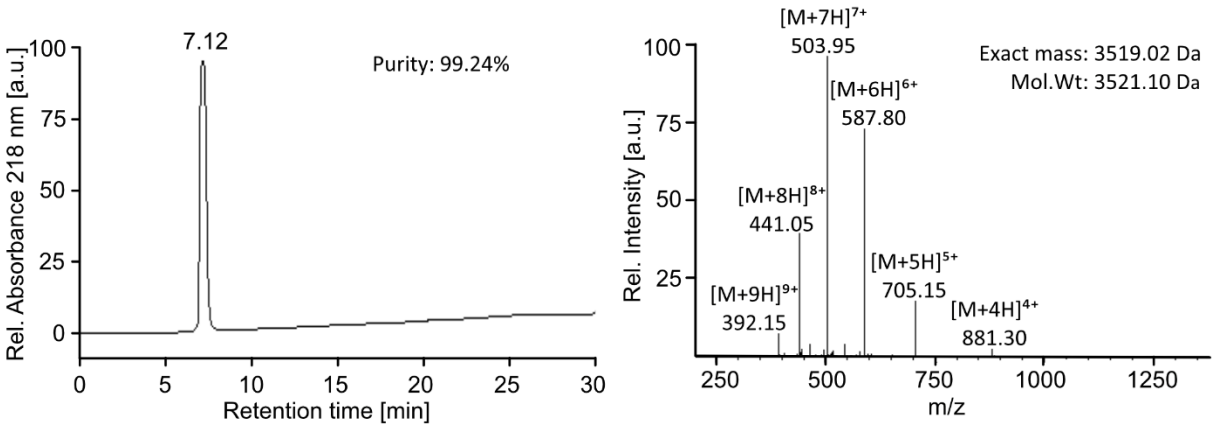
No.6 (H3-Kac23)



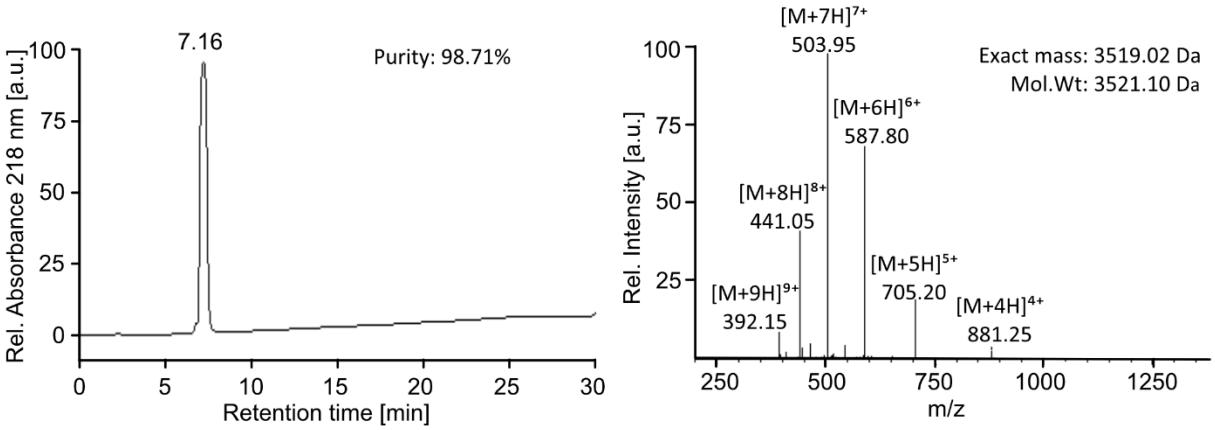
No.7 (H3-Kac27)



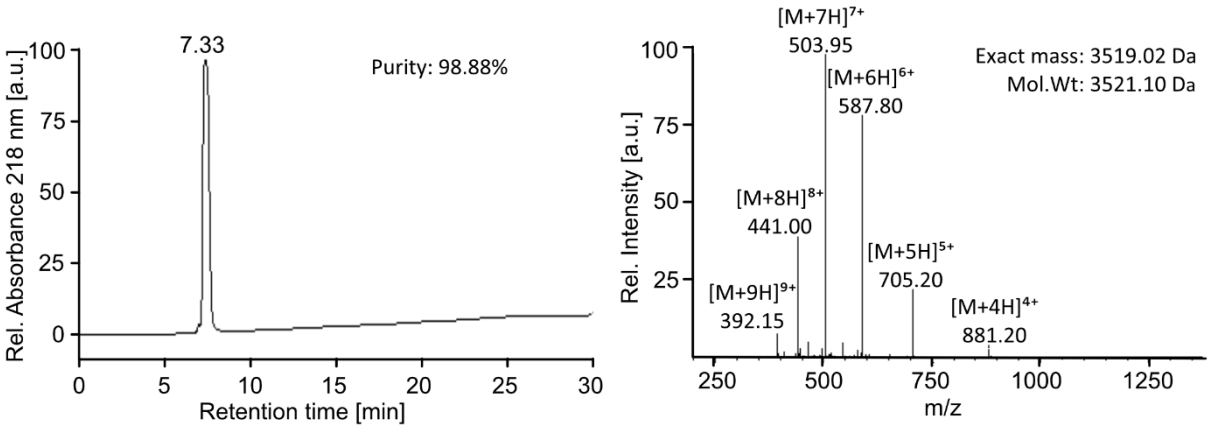
No.8 (H3-Kac4-9)



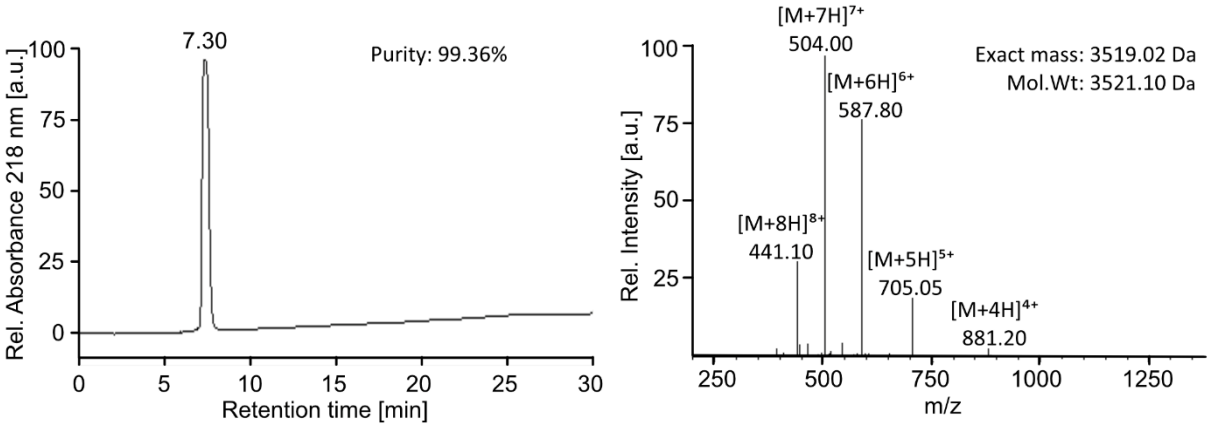
No. 9 (H3-Kac4-14)



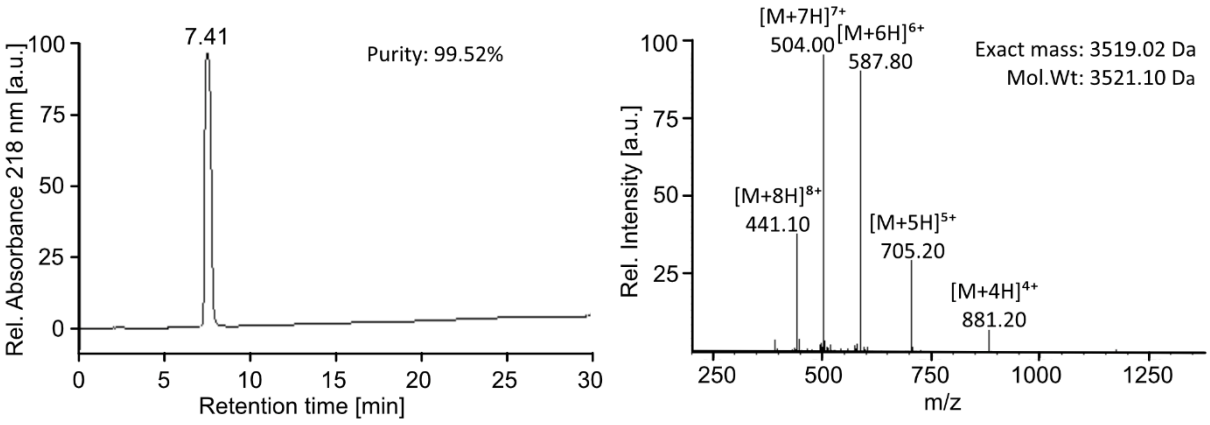
No.10 (H3-Kac4-18)



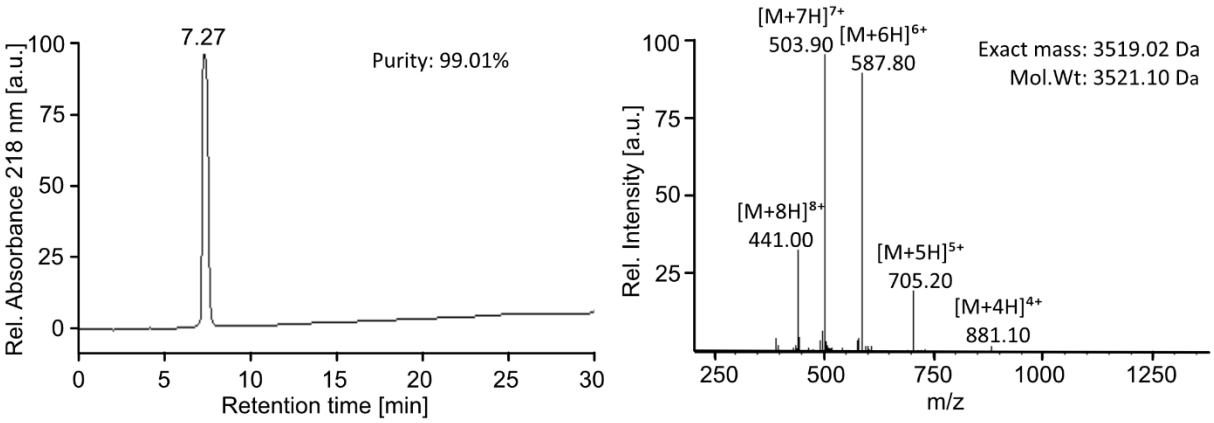
No.11 (H3-Kac4-23)



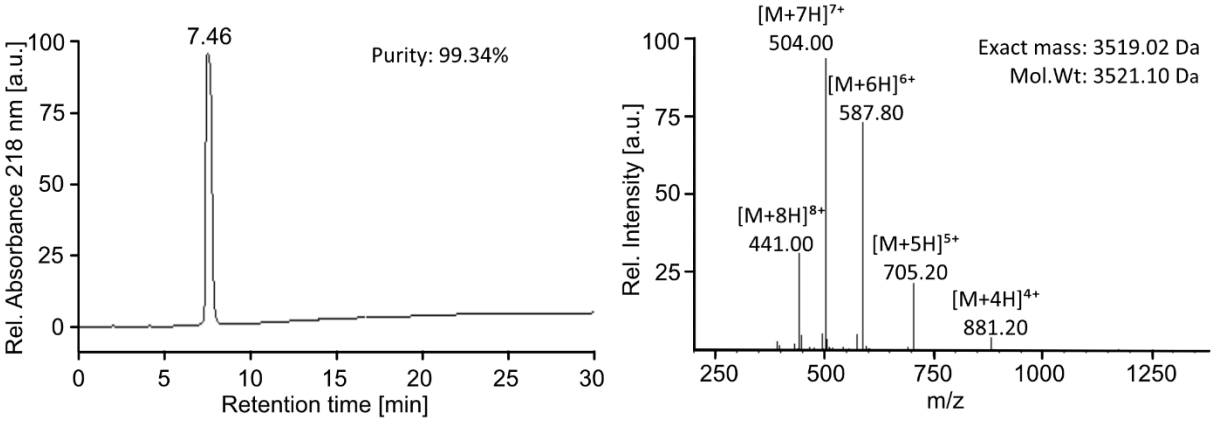
No.12 (H3-Kac4-27)



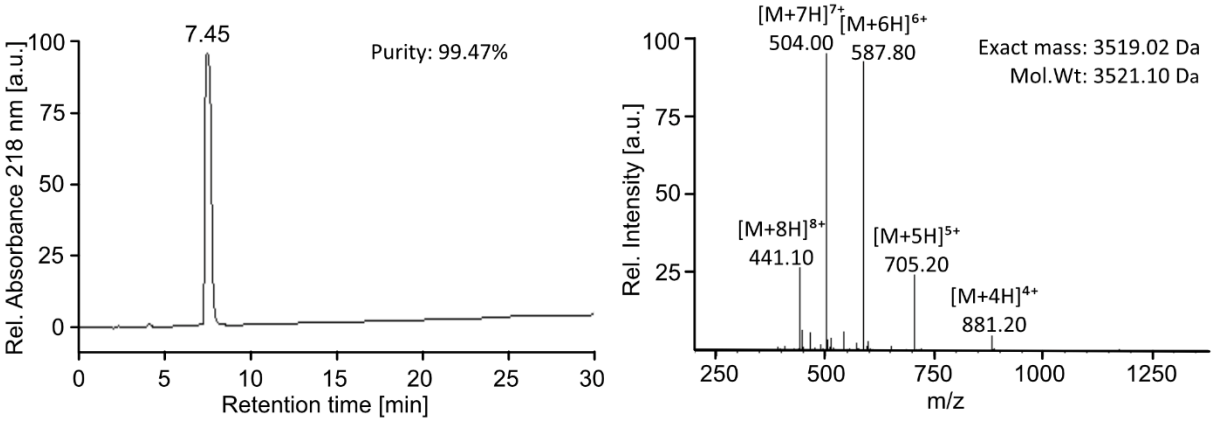
No.13 (H3-Kac9-14)



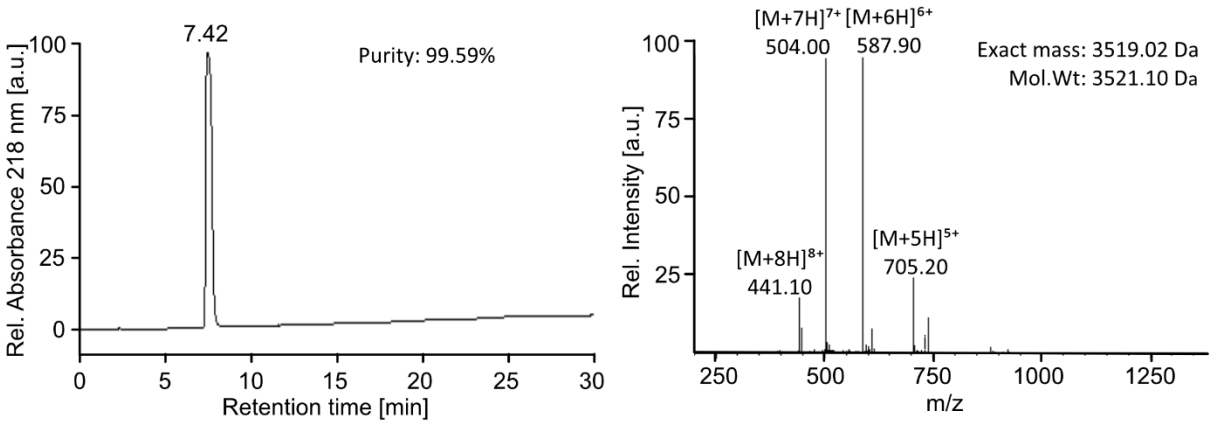
No.14 (H3-Kac9-18)



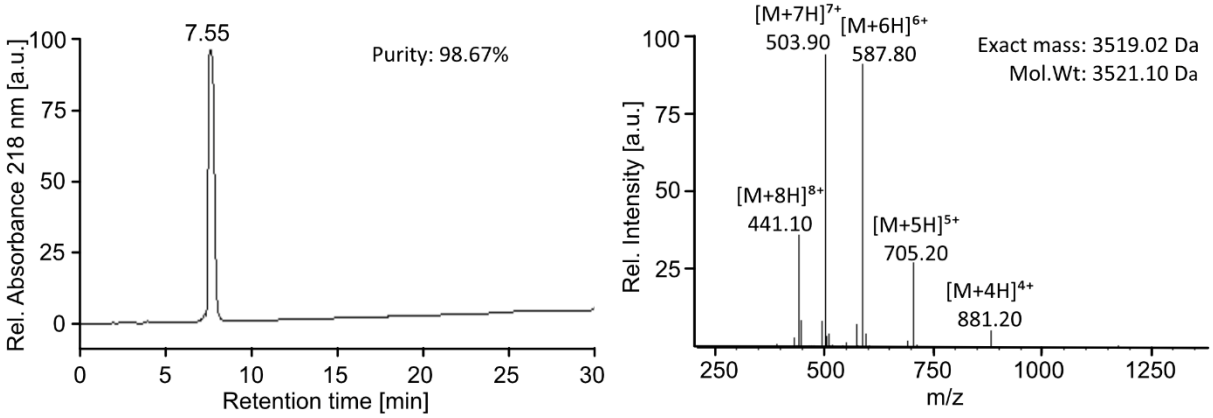
No.15 (H3-Kac9-23)



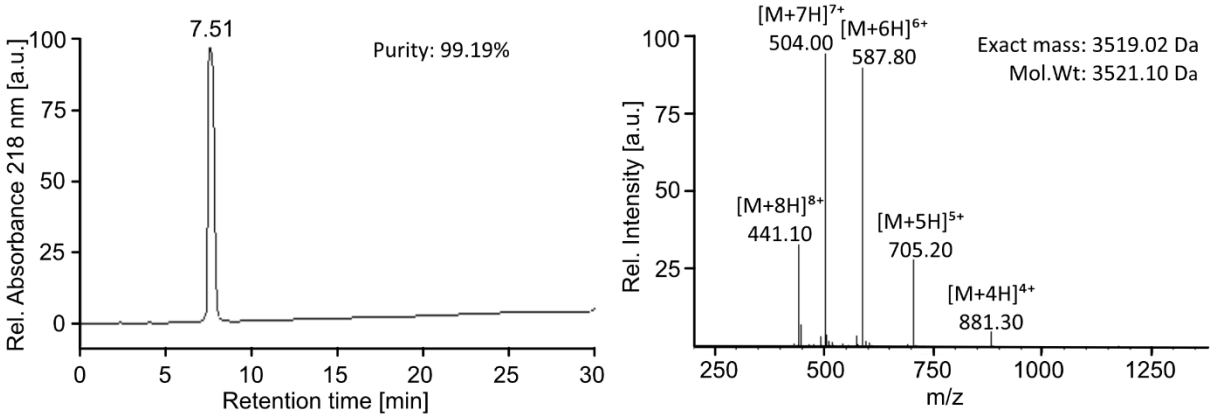
No.16 (H3-Kac9-27)



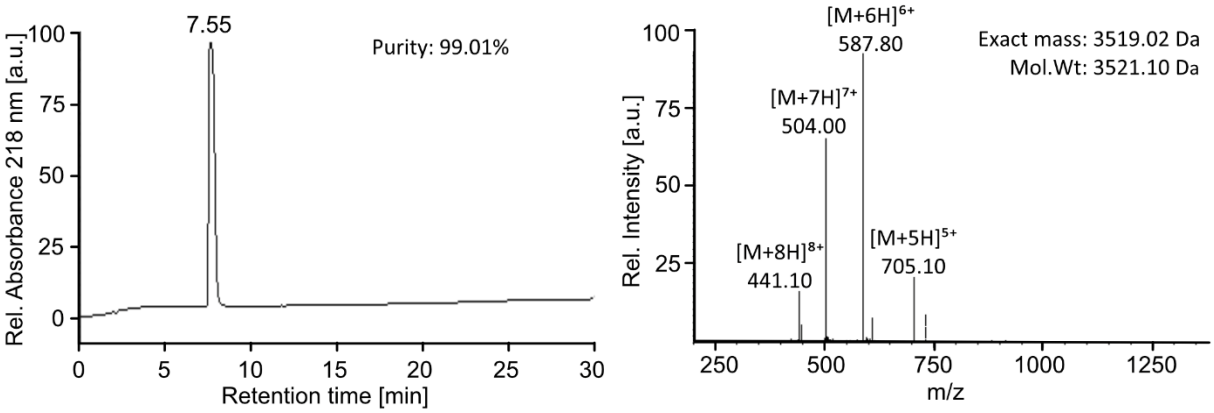
No.17 (H3-Kac14-18)



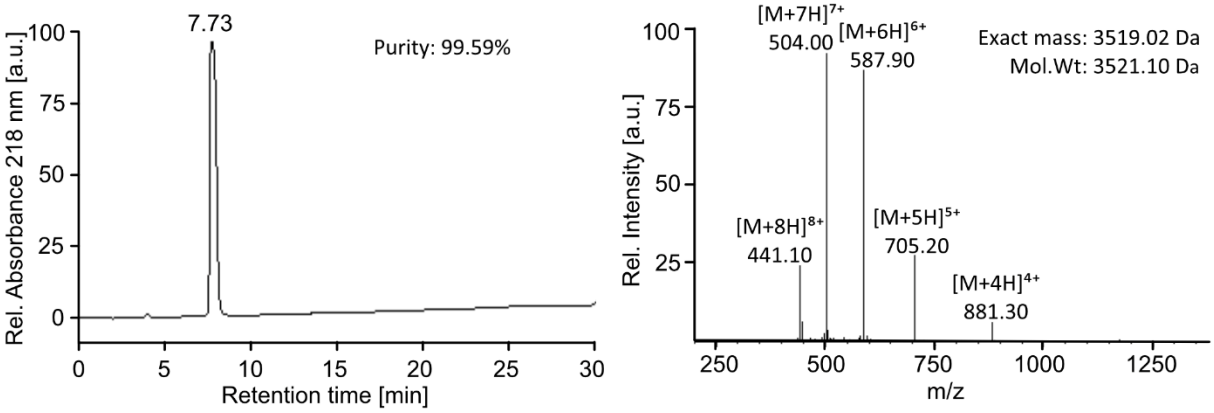
No.18 (H3-Kac14-23)



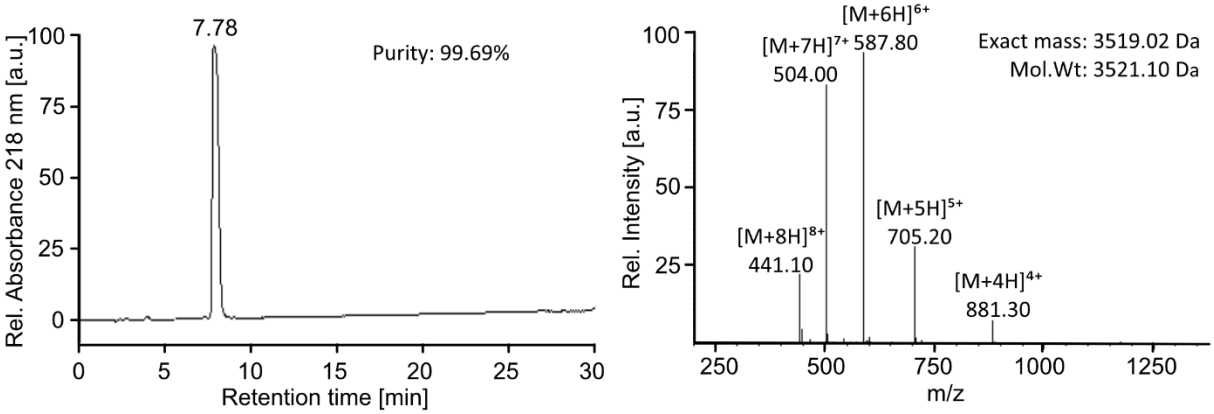
No.19 (H3-Kac14-27)



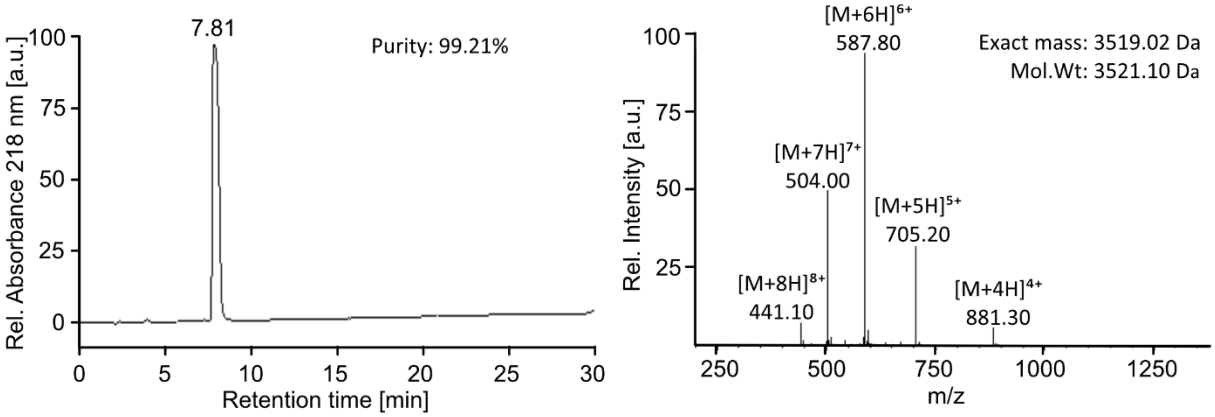
No.20 (H3-Kac18-23)



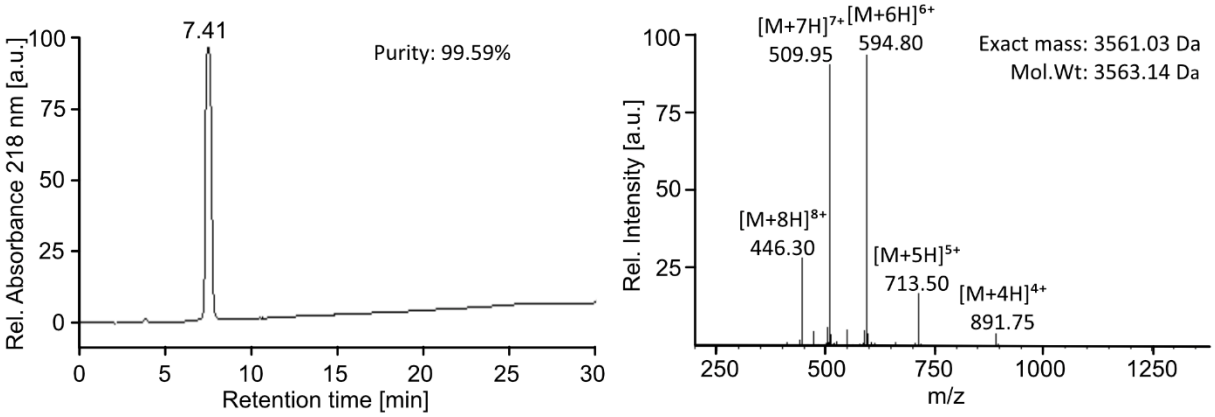
No.21 (H3-Kac18-27)



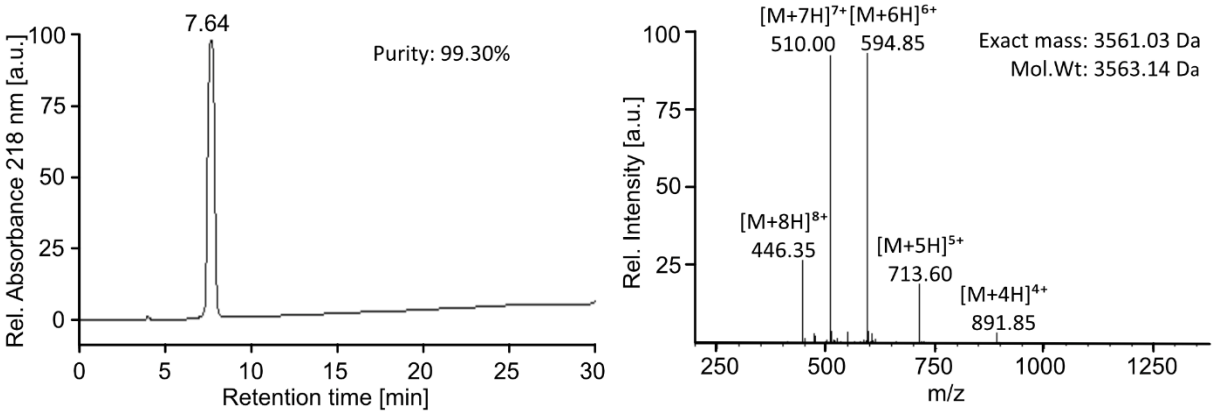
No.22 (H3-Kac23-27)



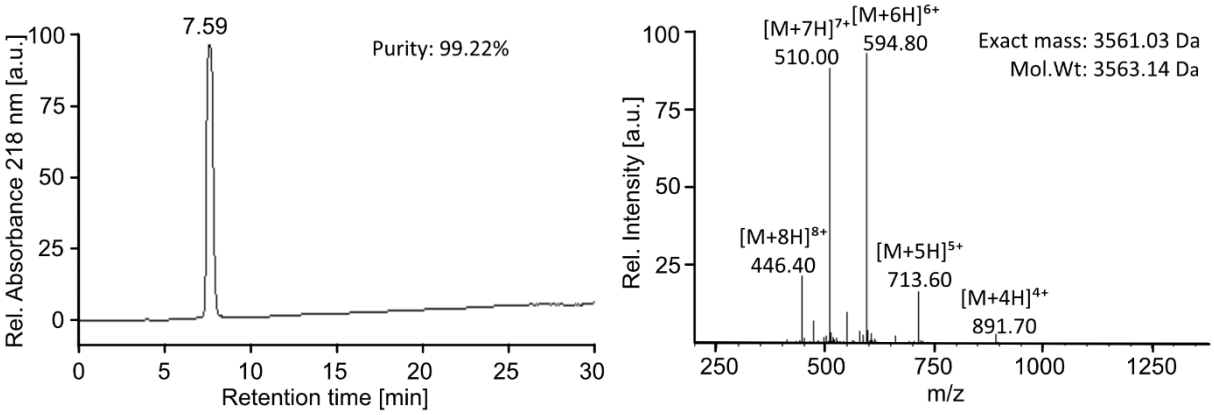
No.23 (H3-Kac4-9-14)



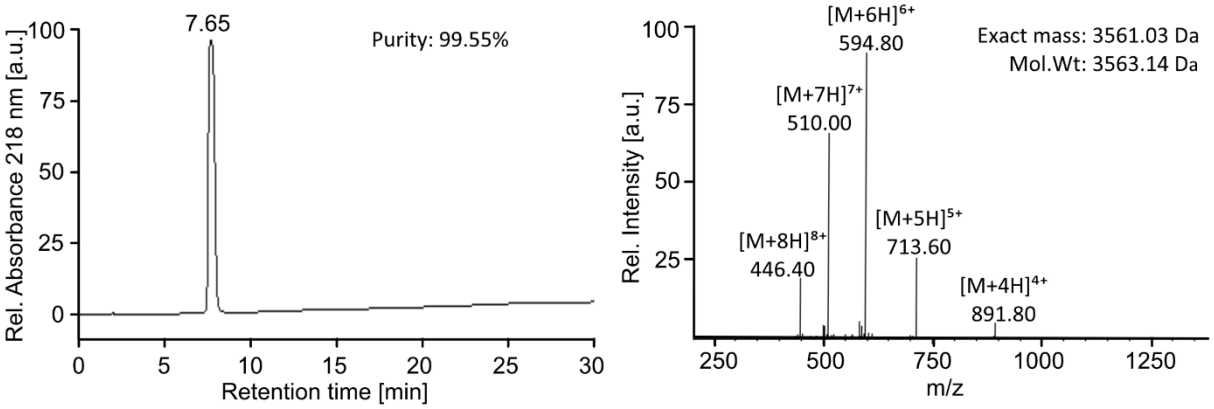
No.24 (H3-Kac4-9-18)



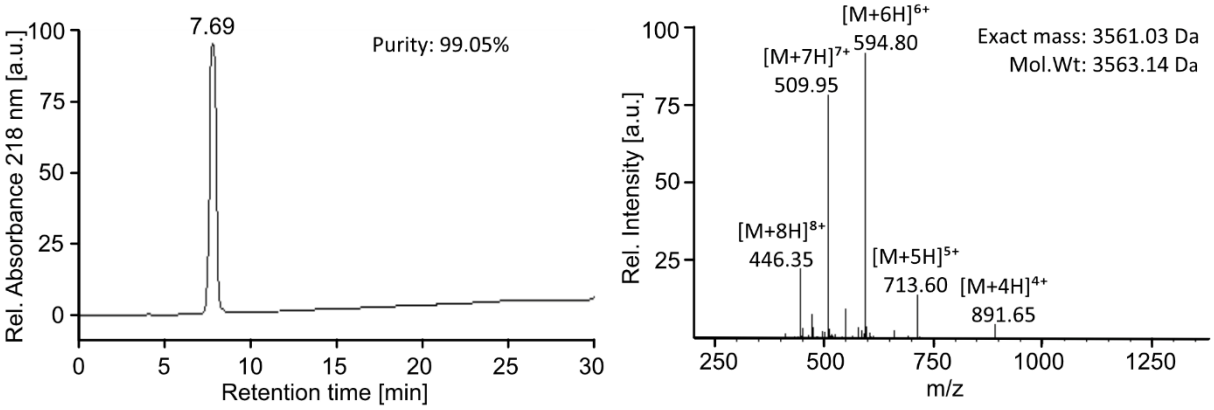
No.25 (H3-Kac4-9-23)



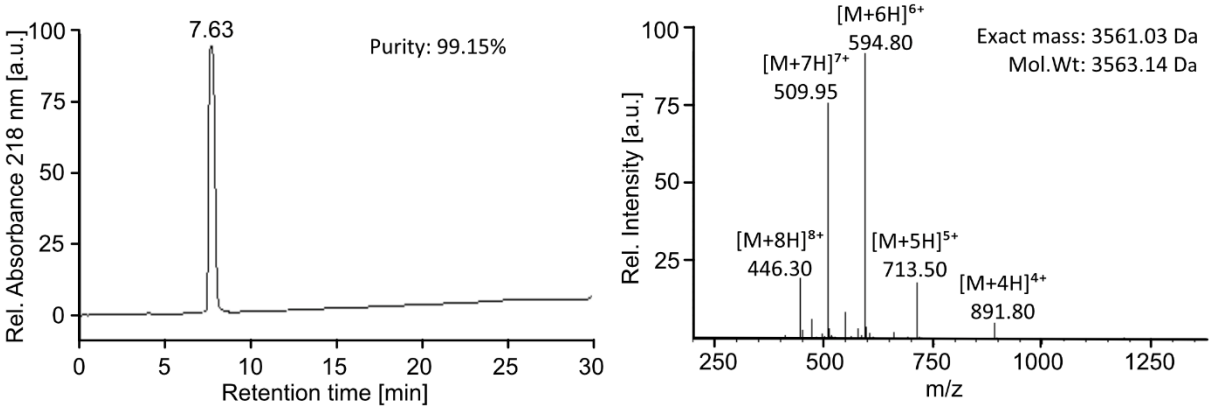
No.26 (H3-Kac4-9-27)



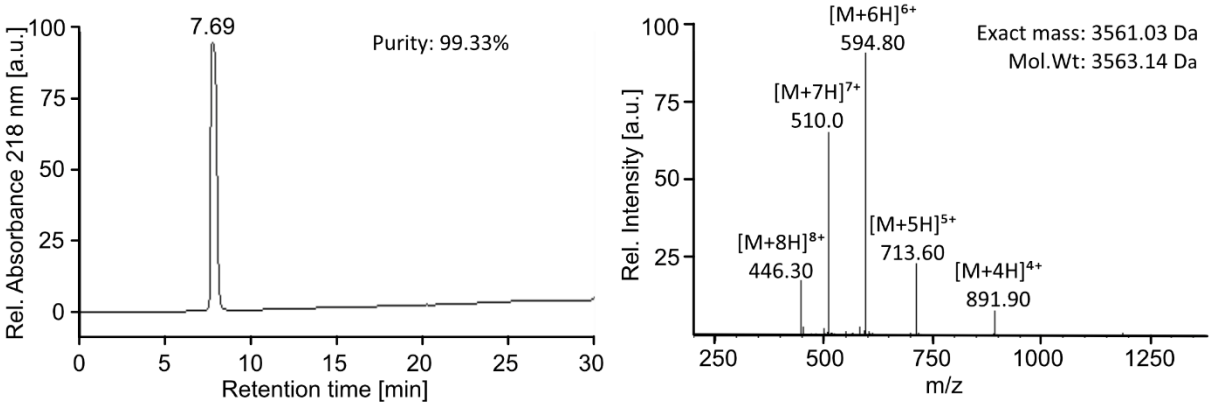
No.27 (H3-Kac4-14-18)



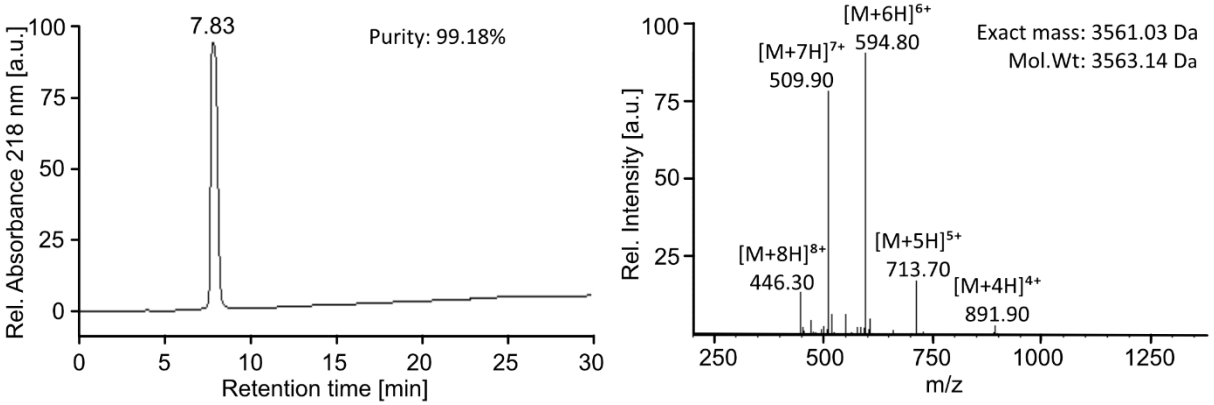
No.28 (H3-Kac4-14-23)



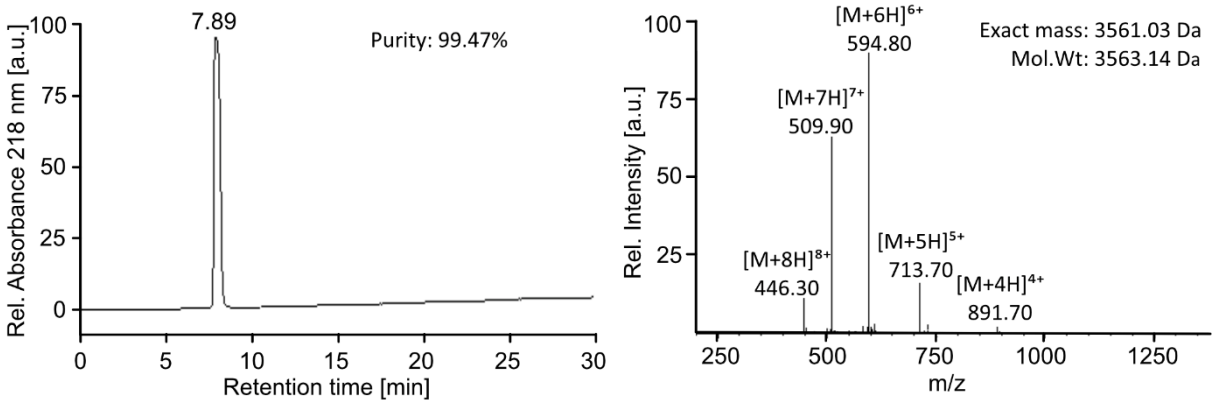
No.29 (H3-Kac4-14-27)



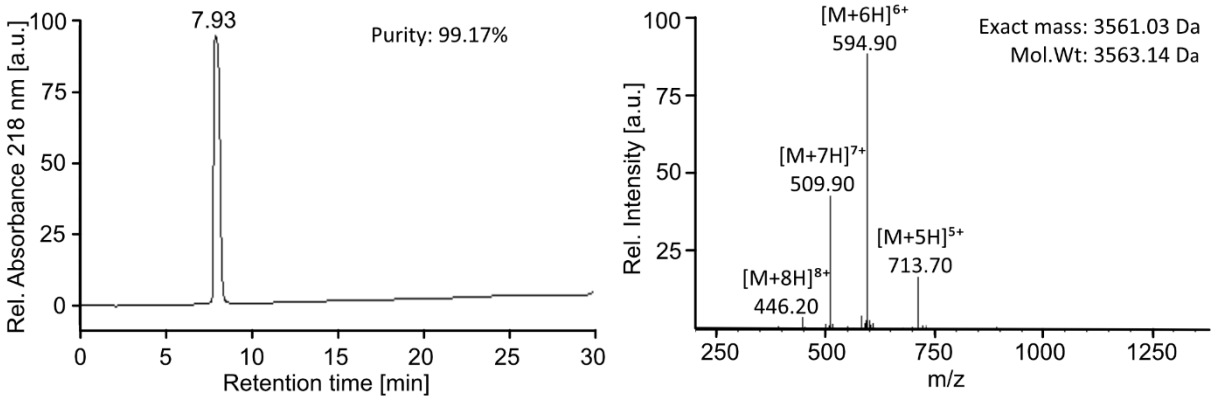
No.30 (H3-Kac4-18-23)



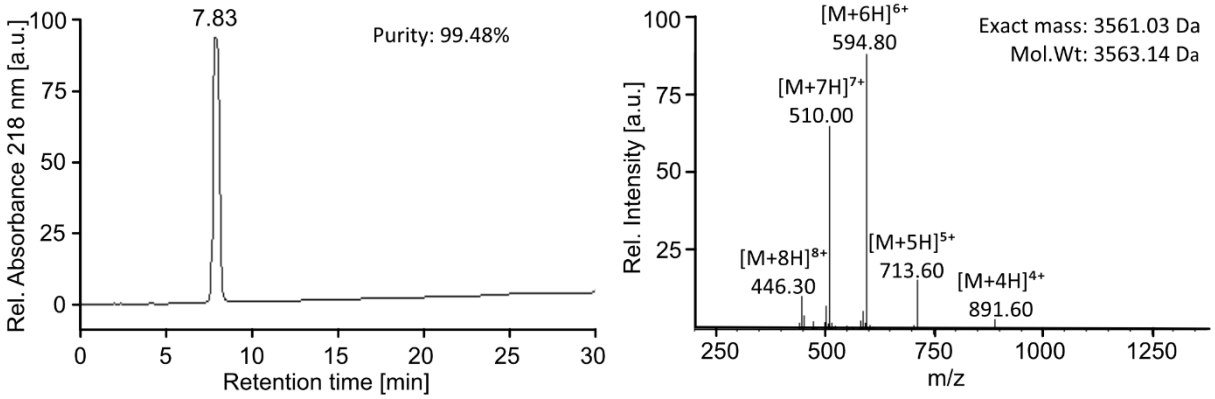
No.31 (H3-Kac4-18-27)



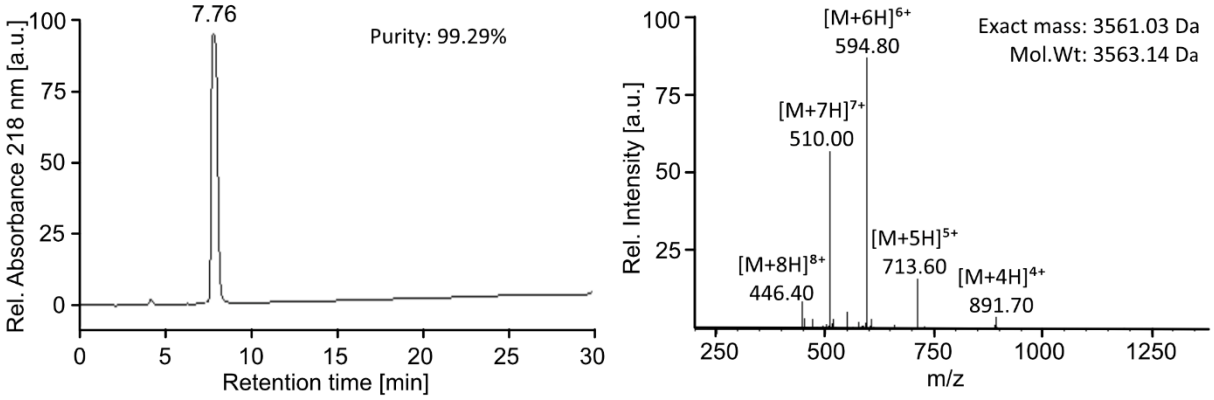
No.32 (H3-Kac4-23-27)



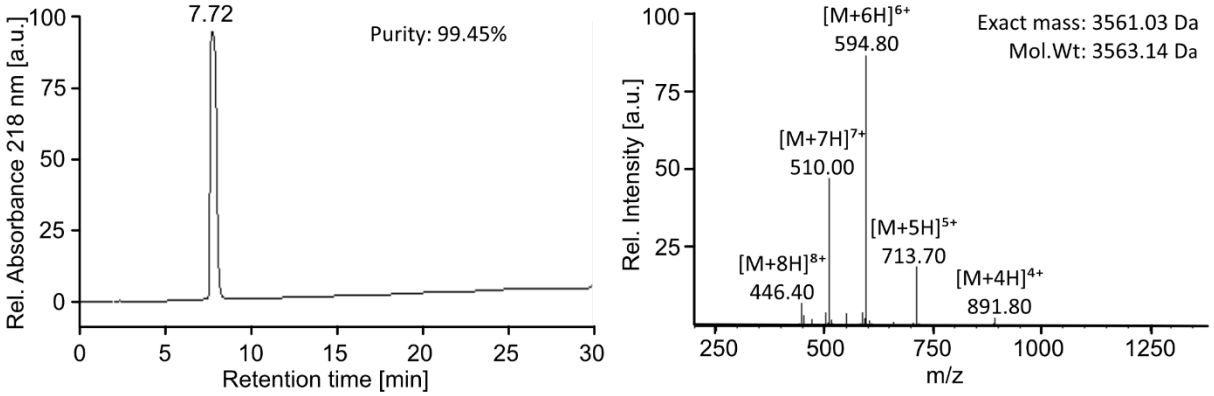
No.33 (H3-Kac9-14-18)



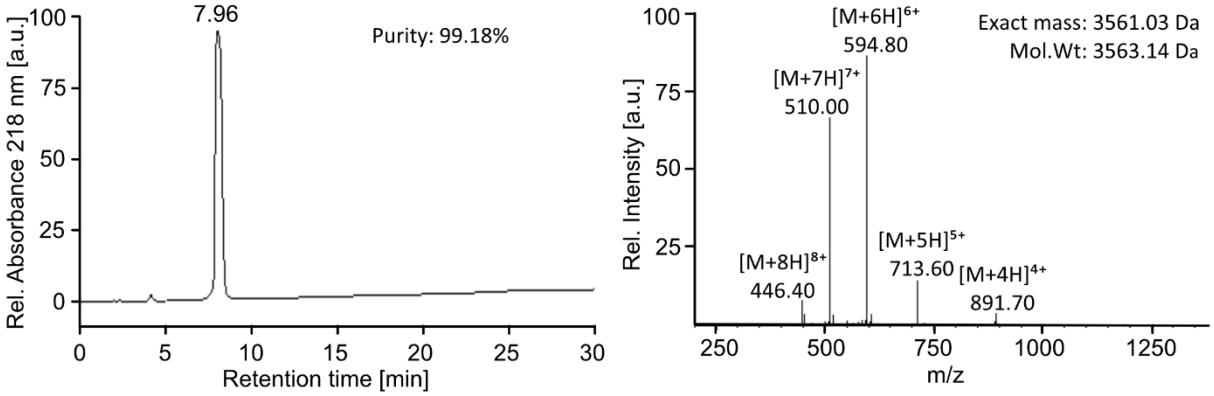
No.34 (H3-Kac9-14-23)



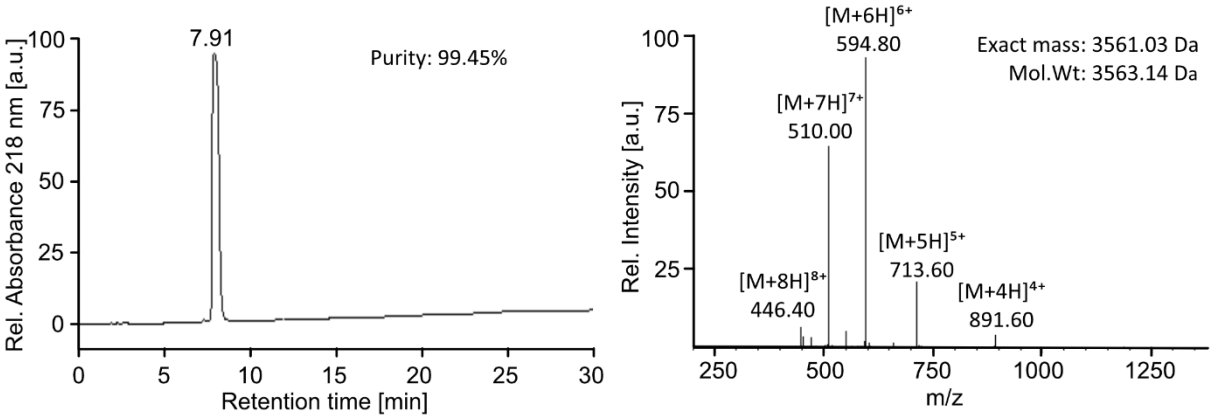
No.35 (H3-Kac9-14-27)



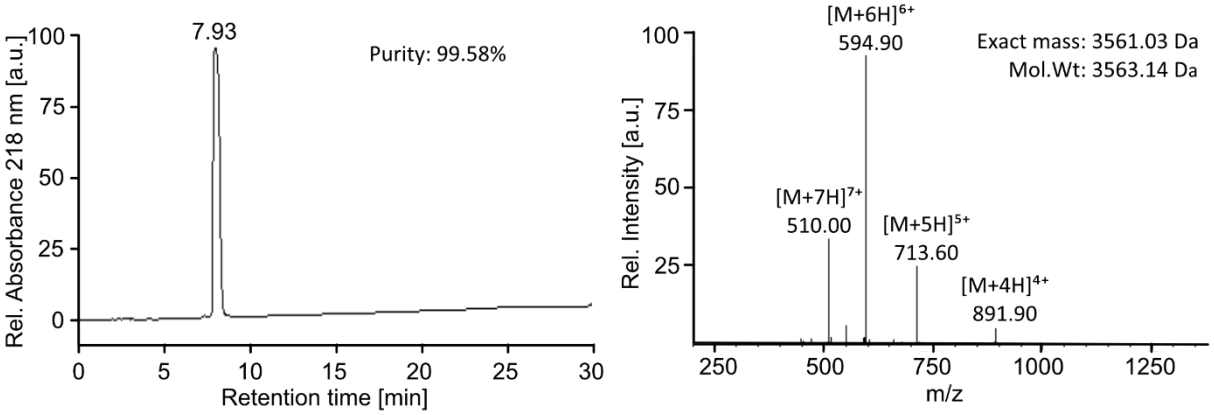
No.36 (H3-Kac9-18-23)



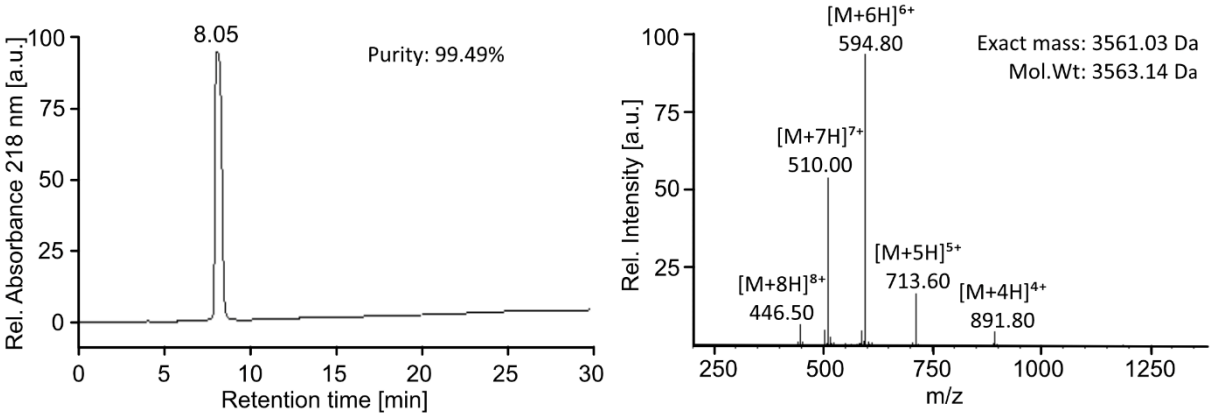
No.37 (H3-Kac9-18-27)



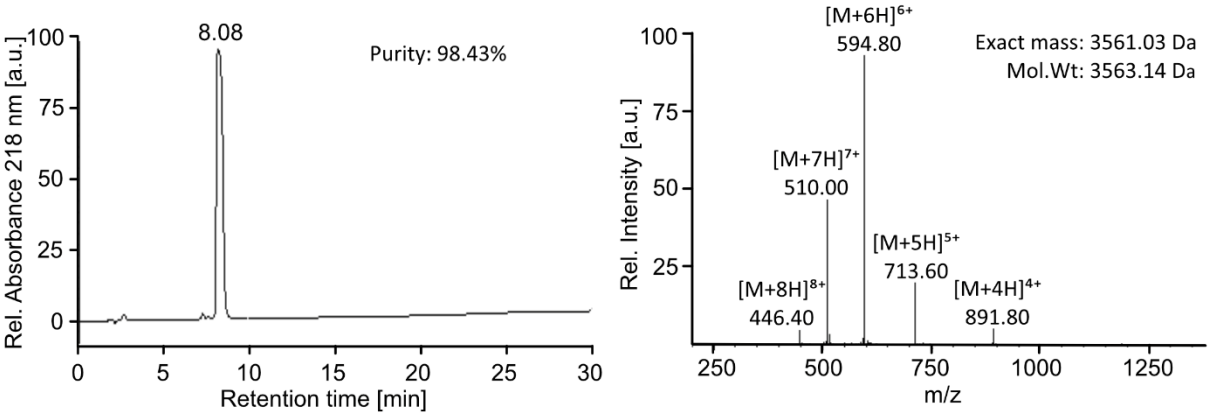
No.38 (H3-Kac9-23-27)



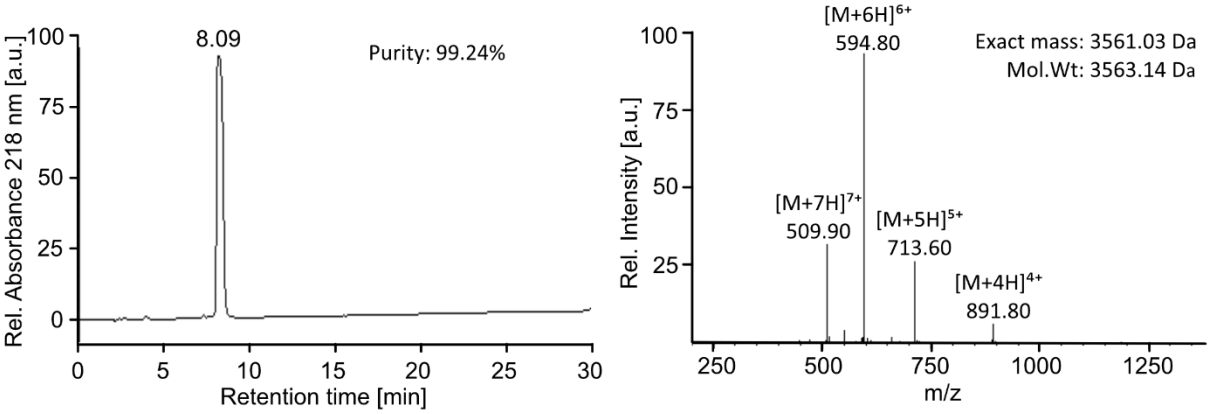
No.39 (H3-Kac14-18-23)



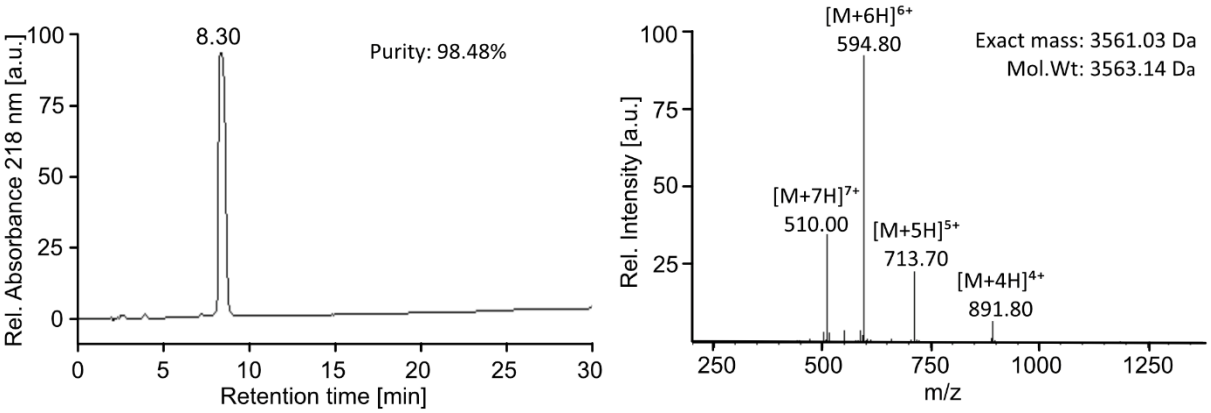
No.40 (H3-Kac14-18-27)



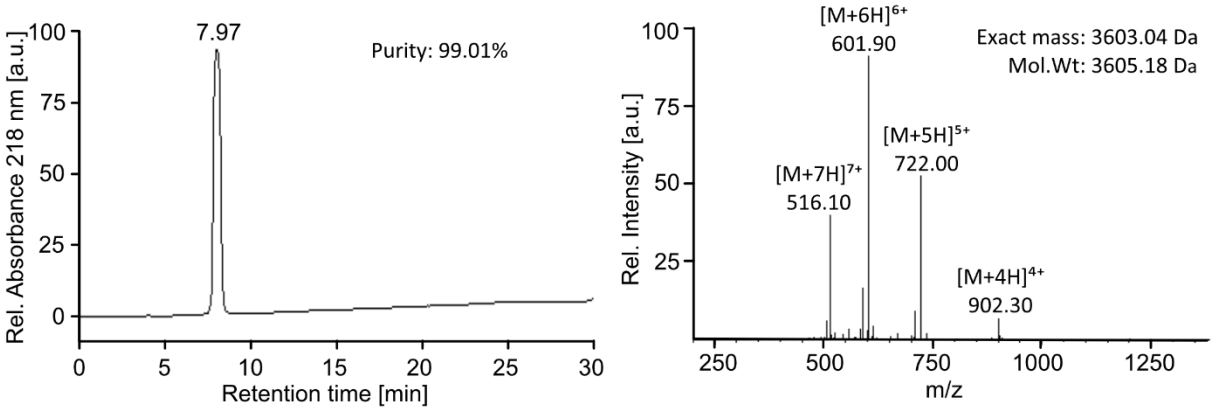
No.41 (H3-Kac14-23-27)



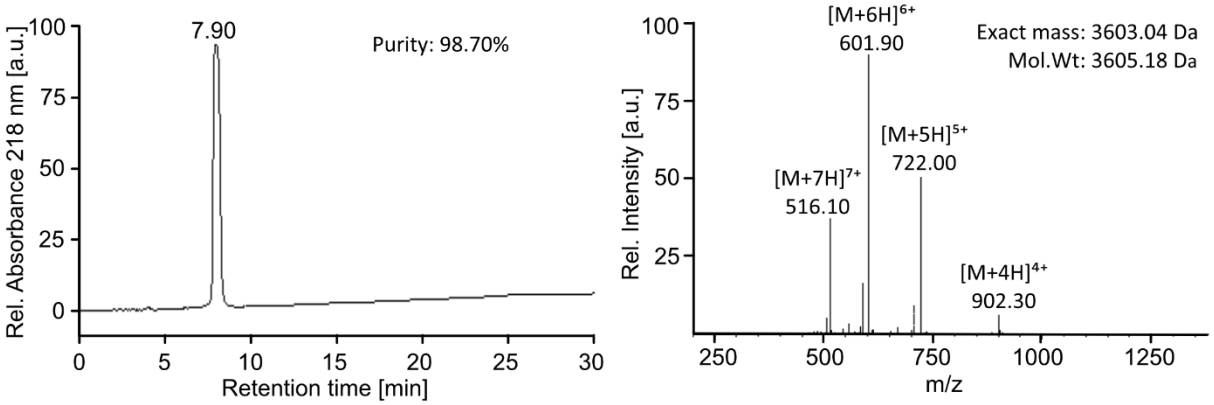
No.42 (H3-Kac18-23-27)



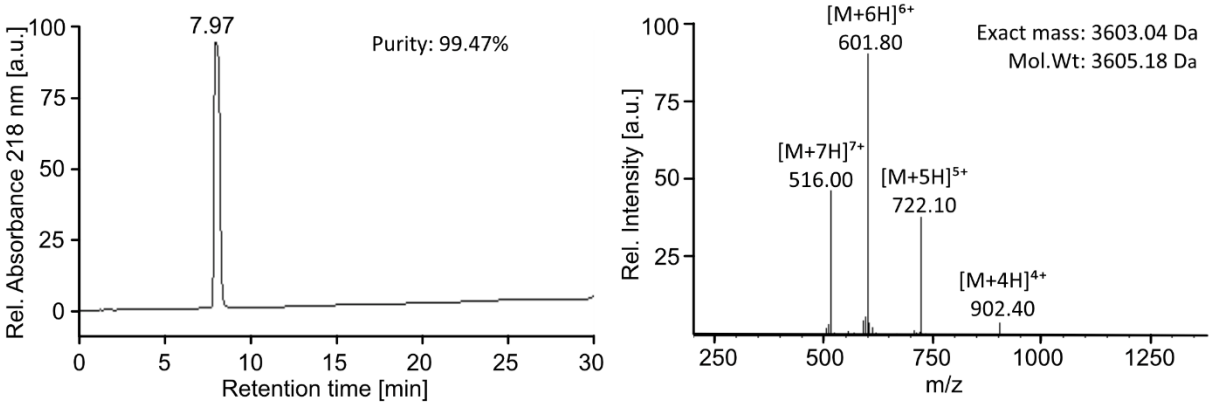
No.43 (H3-Kac4-9-14-18)



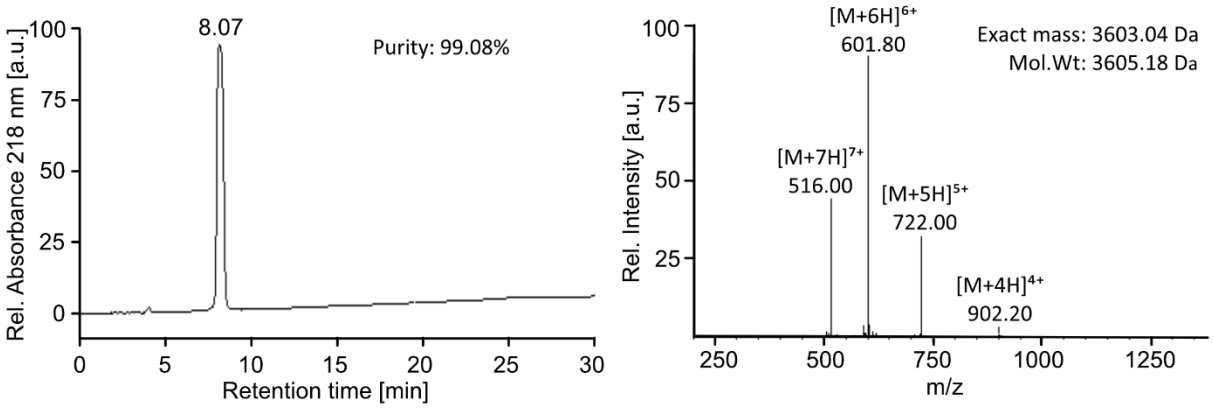
No.44 (H3-Kac4-9-14-23)



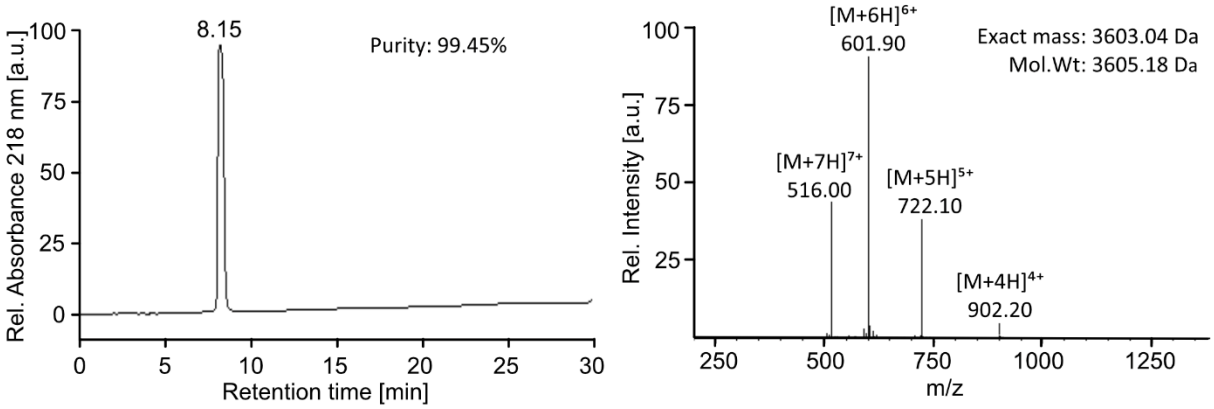
No.45 (H3-Kac4-9-14-27)



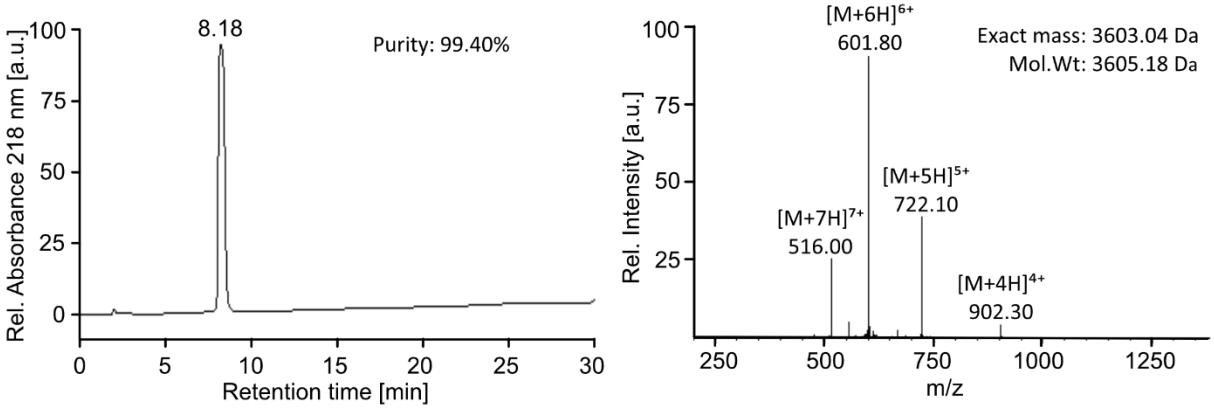
No.46 (H3-Kac4-9-18-23)



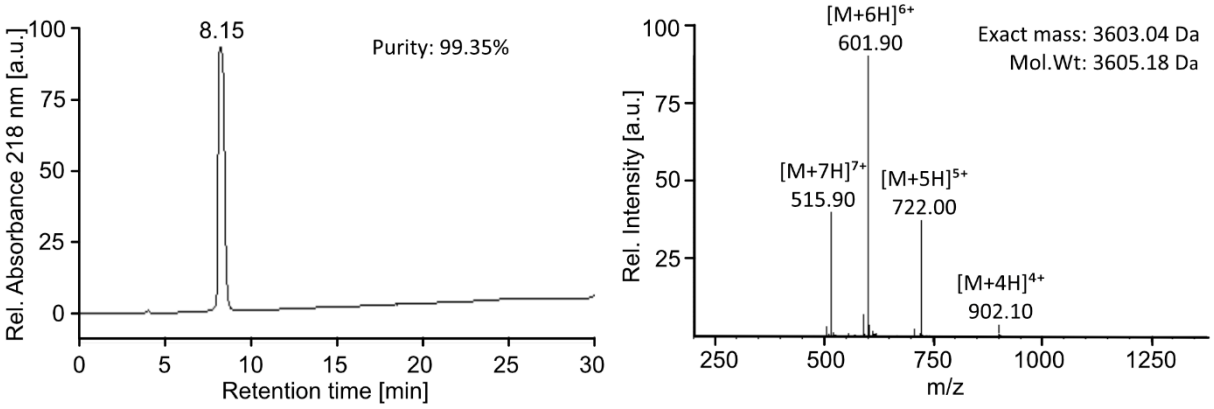
No.47 (H3-Kac4-9-18-27)



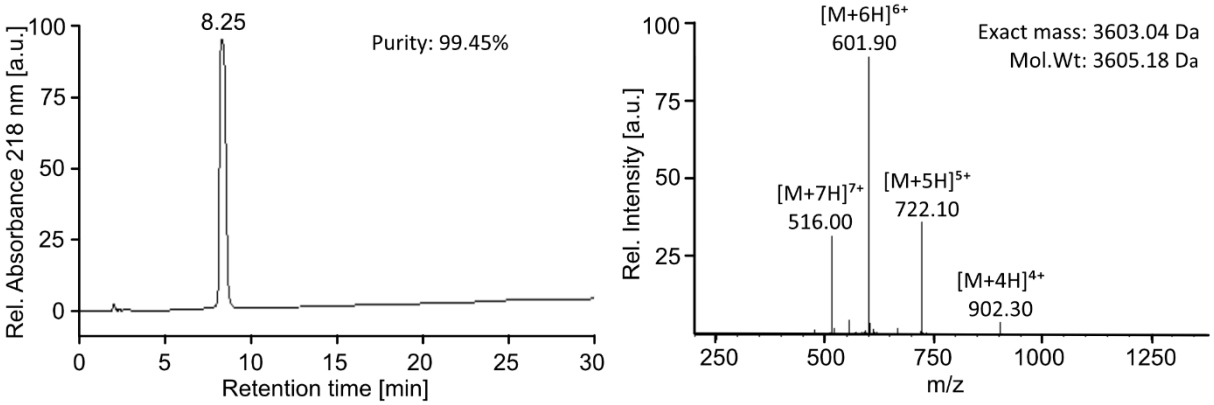
No.48 (H3-Kac4-9-23-27)



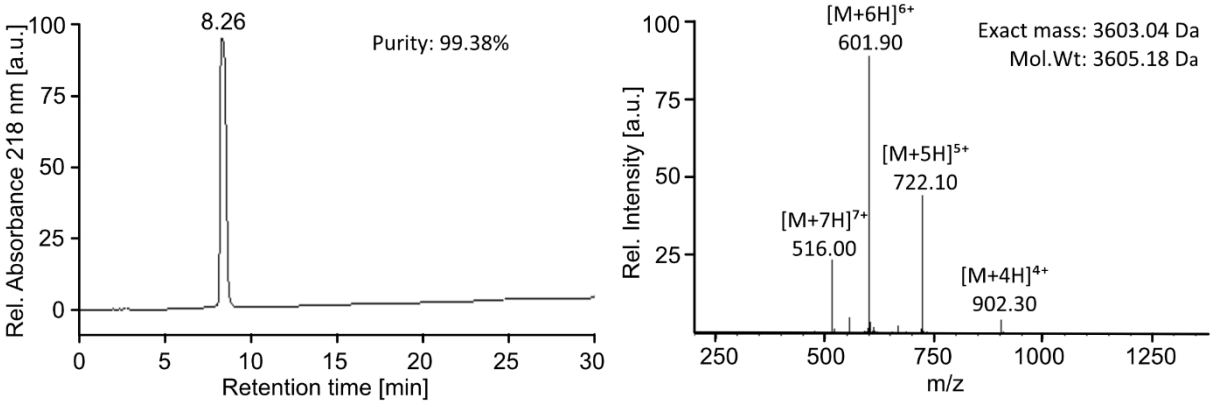
No.49 (H3-Kac4-14-18-23)



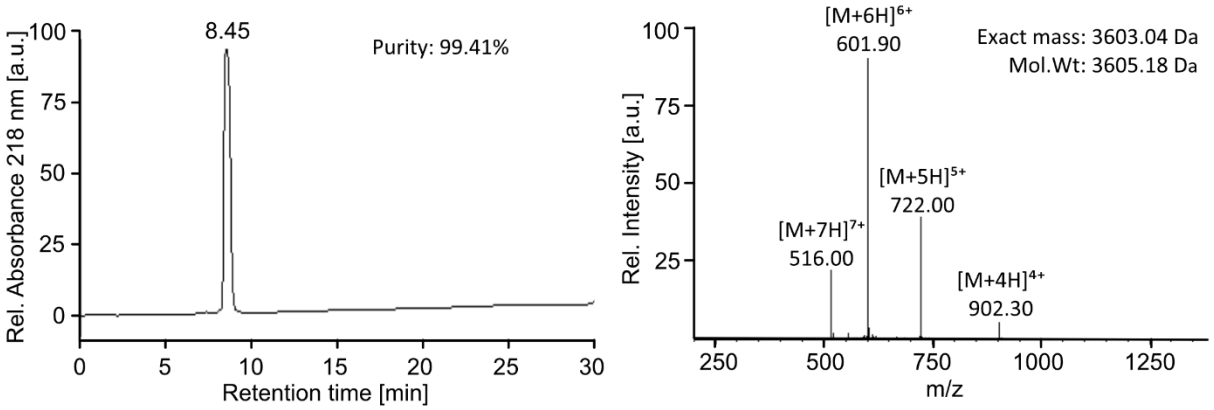
No.50 (H3-Kac4-14-18-27)



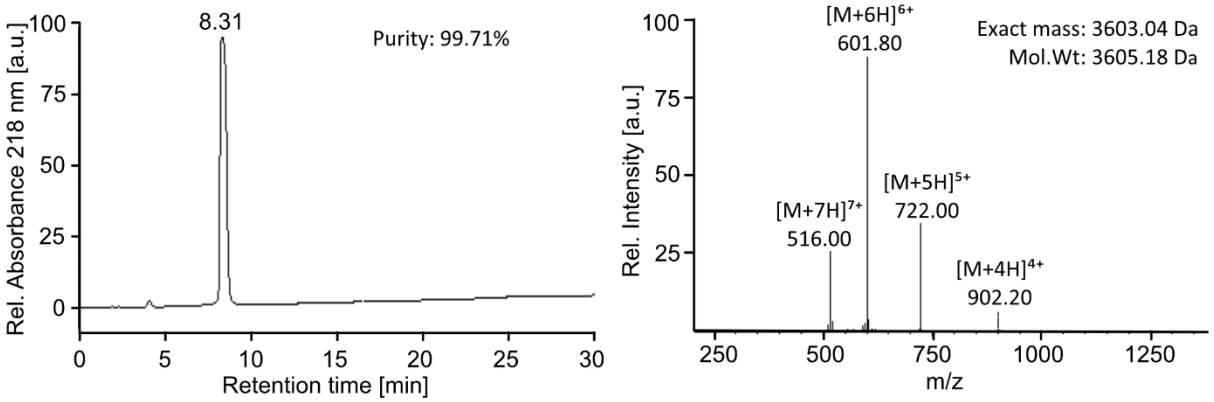
No.51 (H3-Kac4-14-23-27)



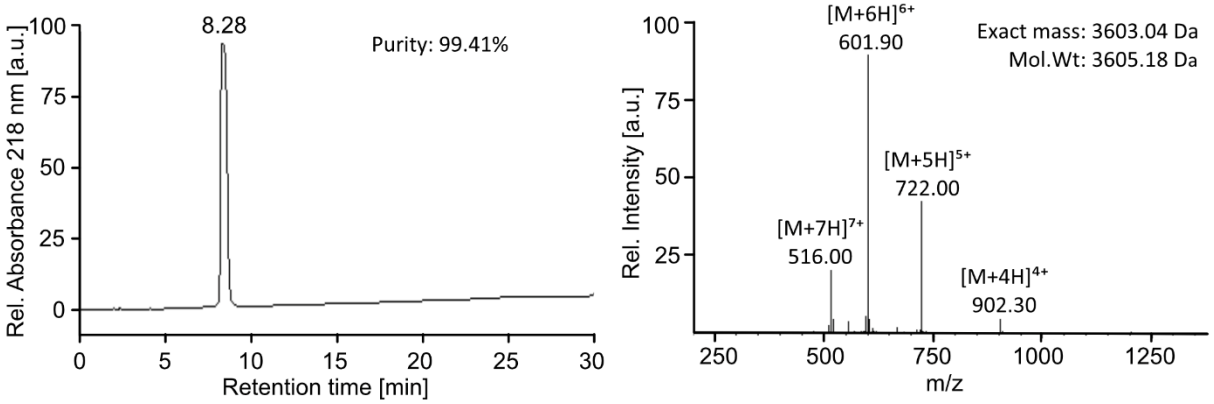
No.52 (H3-Kac4-18-23-27)



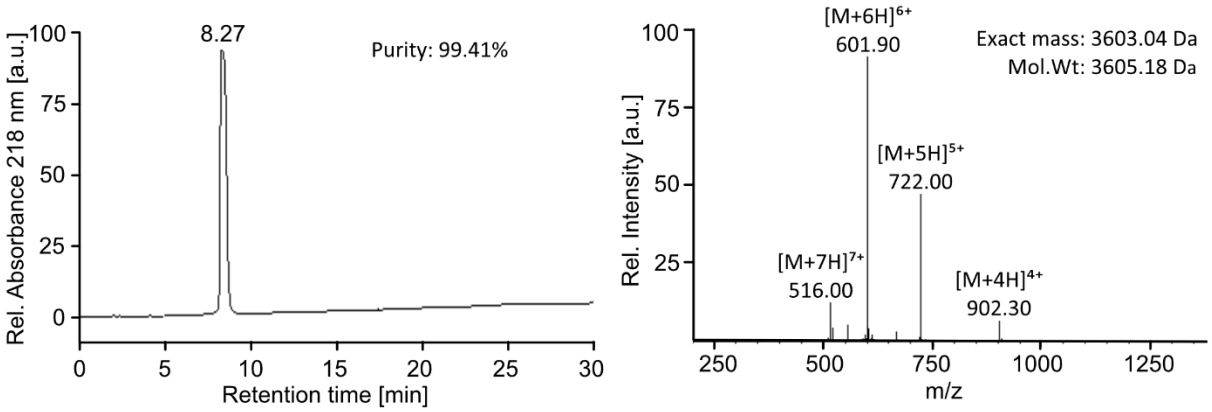
No.53 (H3-Kac9-14-18-23)



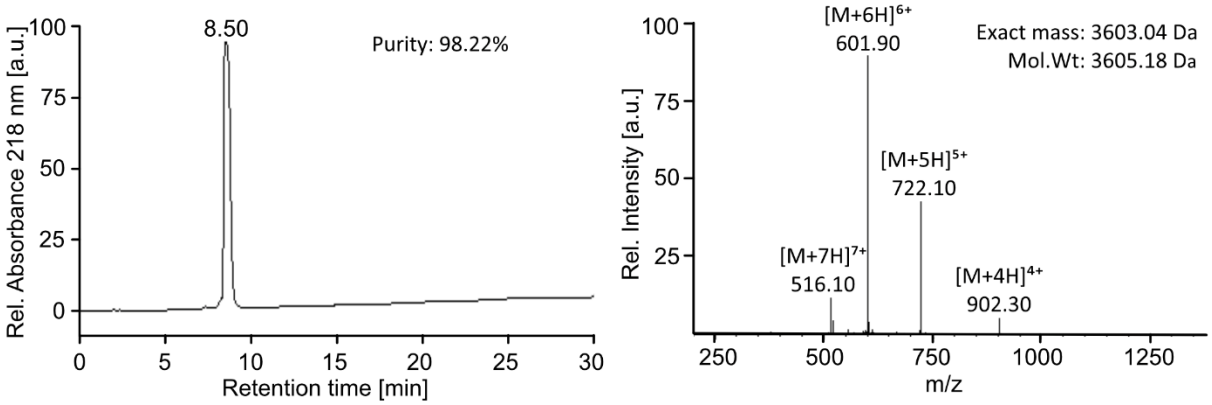
No. 54 (H3-Kac9-14-18-27)



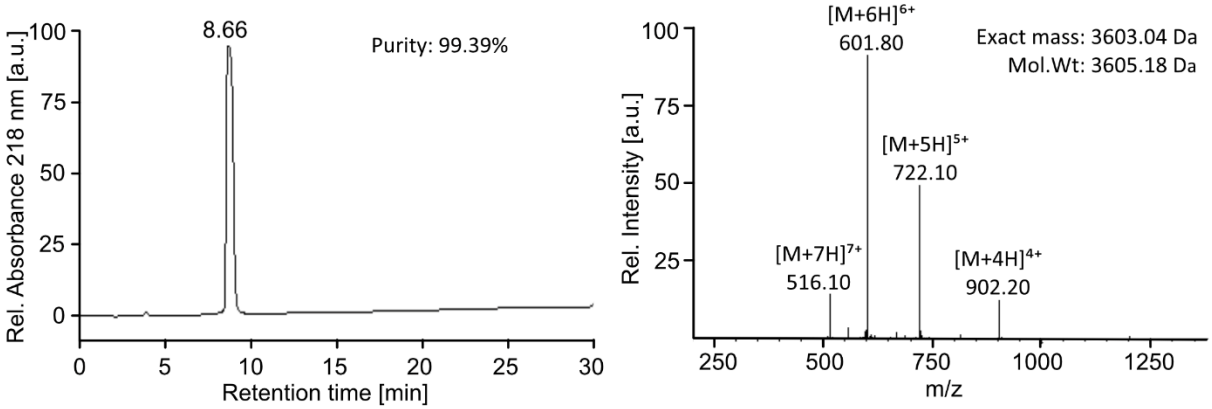
No. 55 (H3-Kac9-14-23-27)



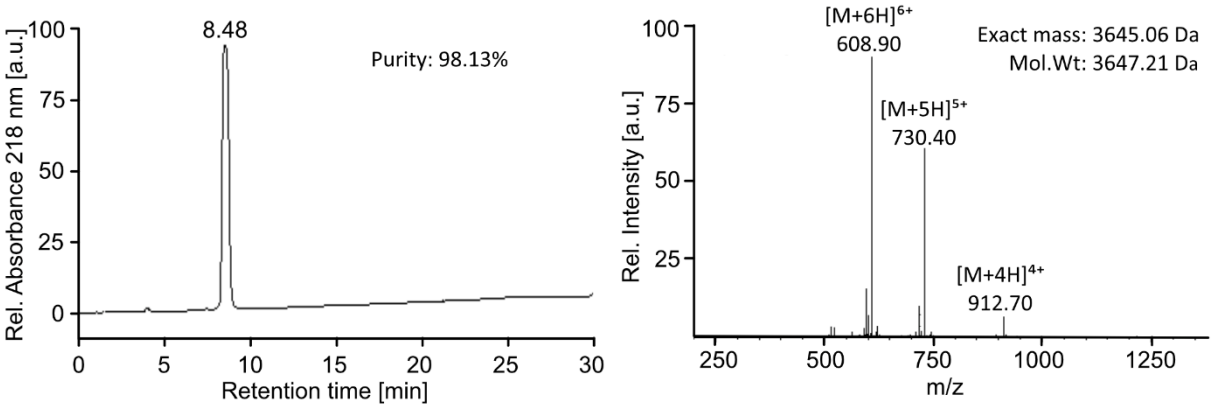
No. 56 (H3-Kac9-18-23-27)



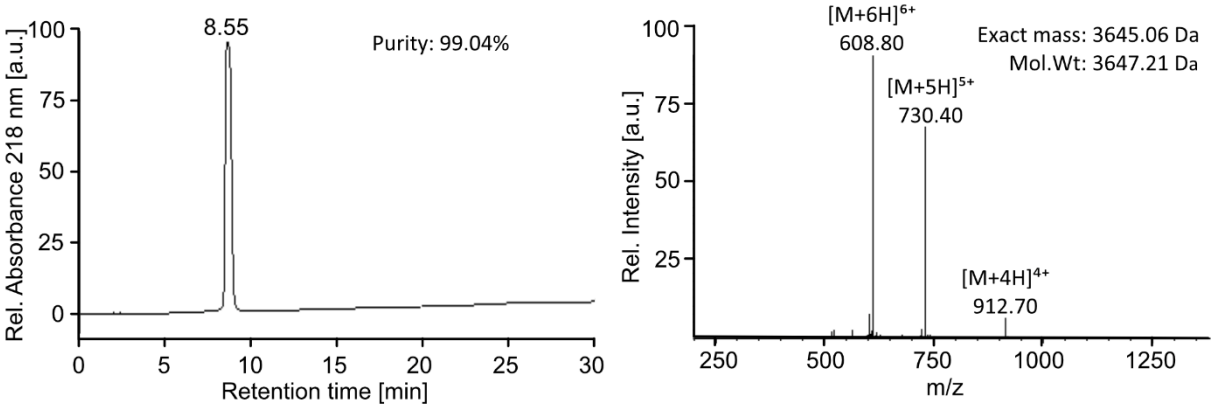
No.57 (H3-Kac14-18-23-27)



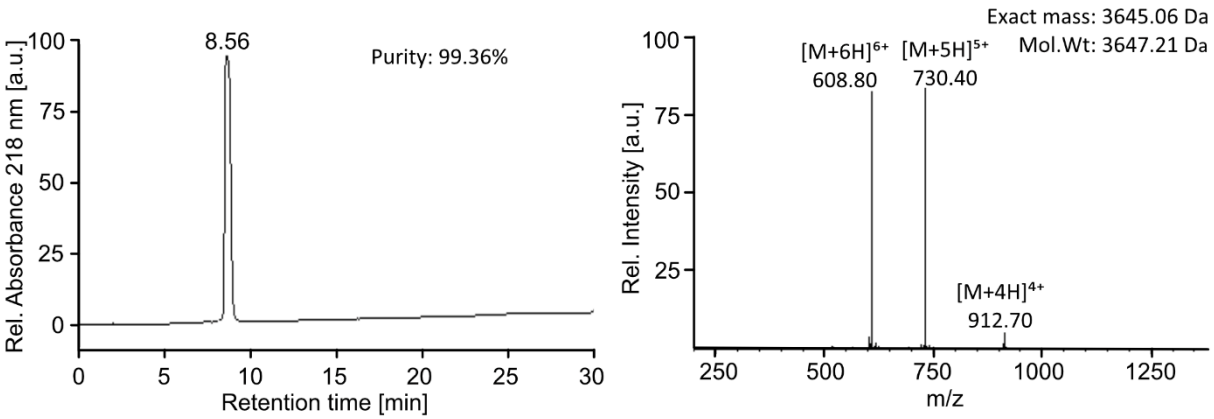
No.58 (H3-Kac4-9-14-18-23)



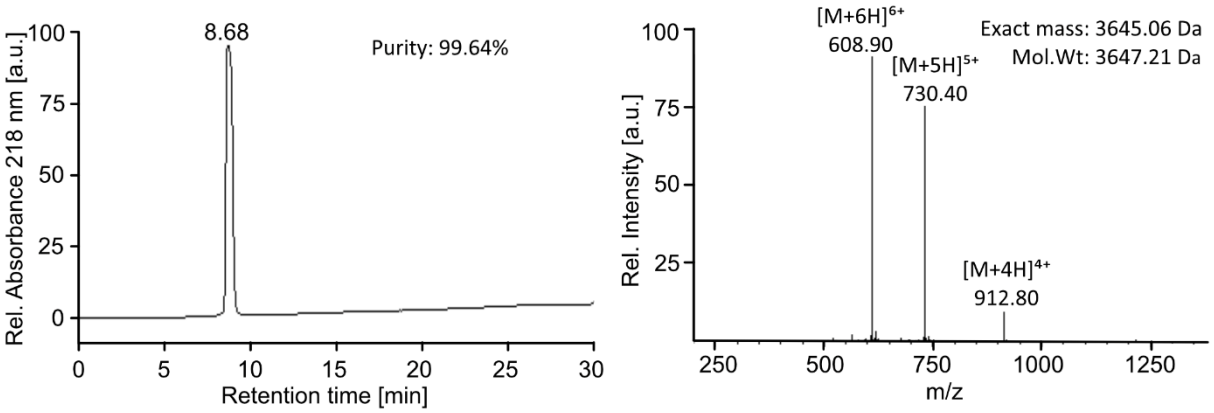
No.59 (H3-Kac4-9-14-18-27)



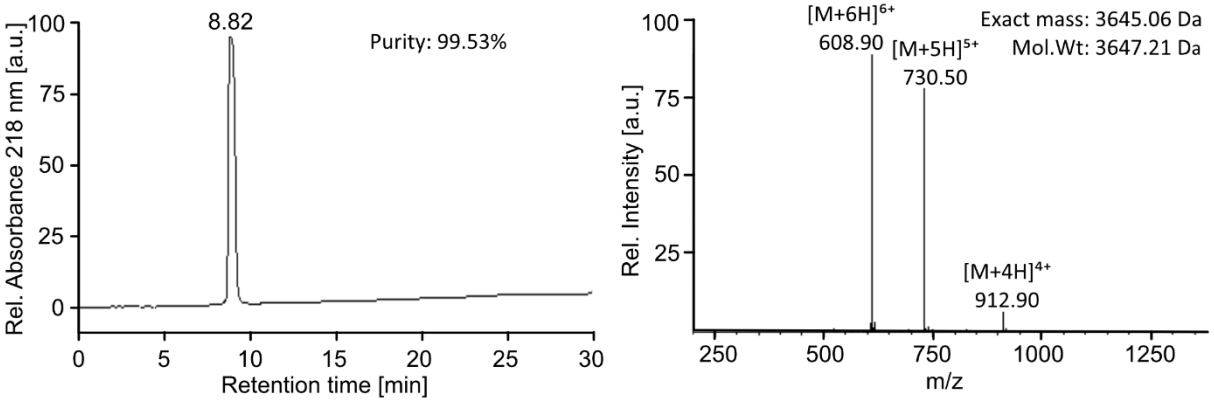
No.60 (H3-Kac4-9-14-23-27)



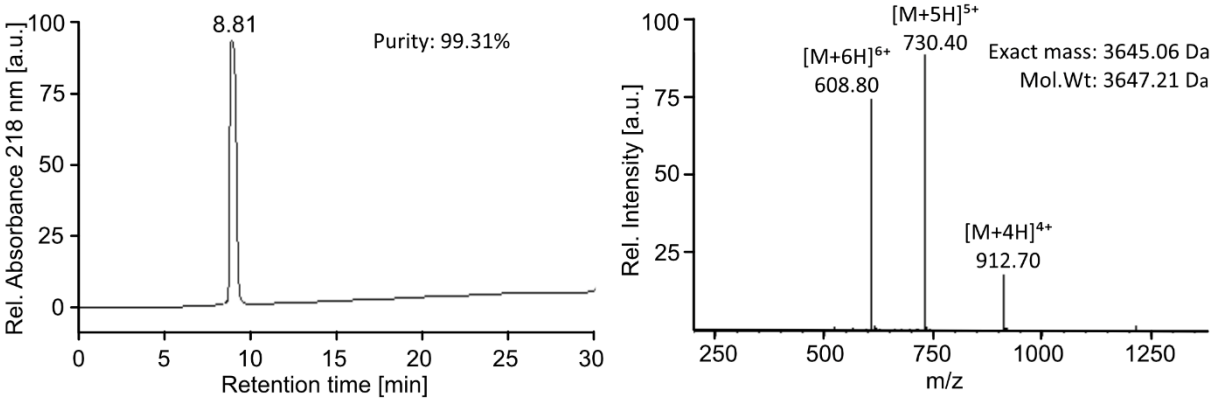
No.61 (H3-Kac4-9-18-23-27)



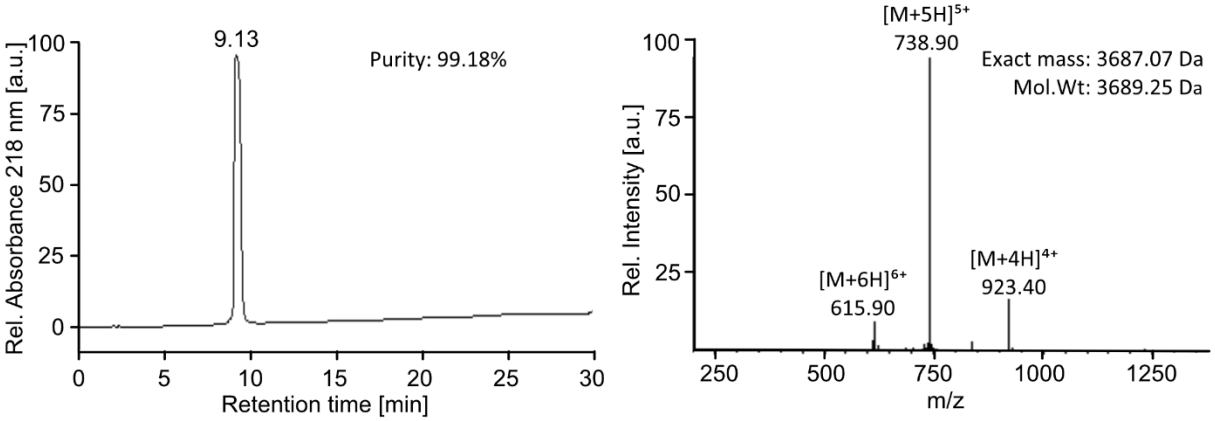
No.62 (H3-Kac4-14-18-23-27)



No.63 (H3-Kac9-14-18-23-27)

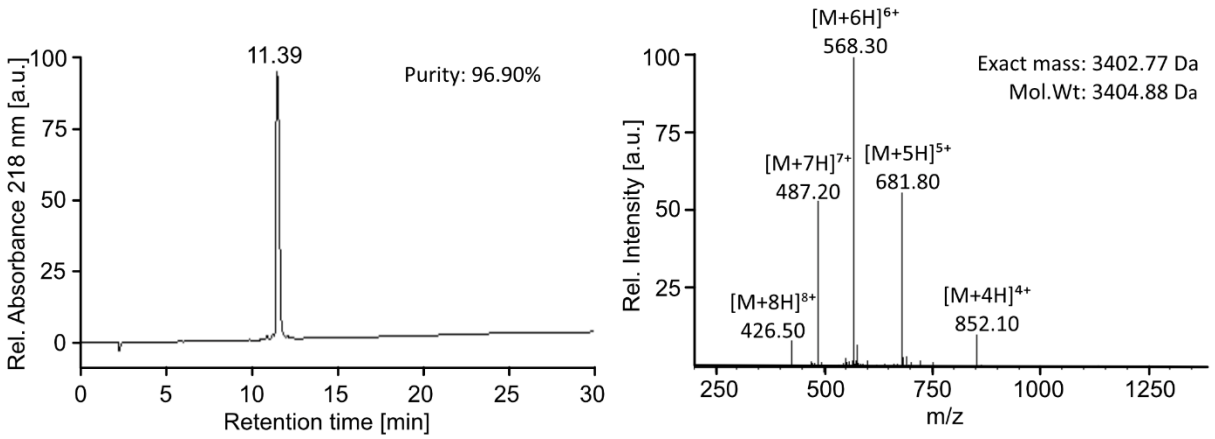


No.64 (H3-Kac4-9-14-18-23-27)

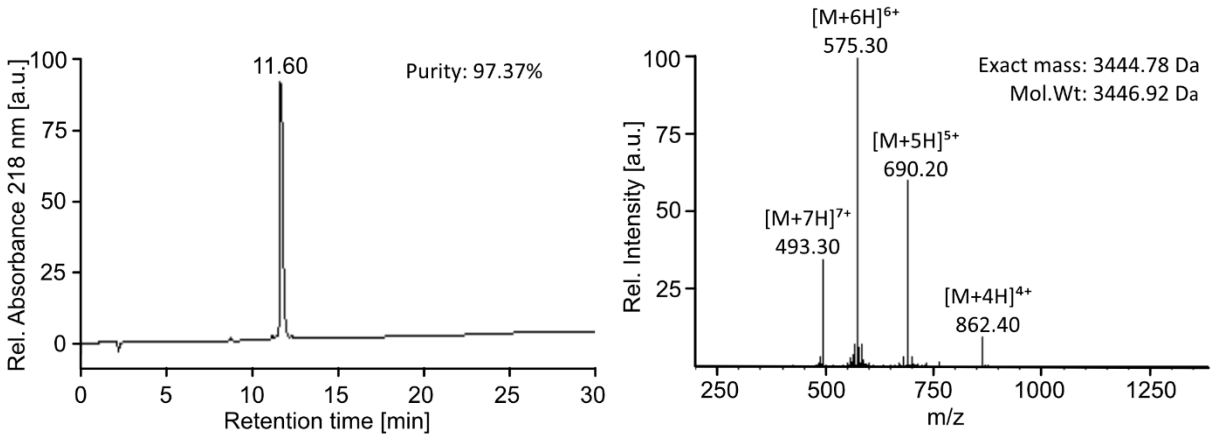


8.2 Analytical results of acetyl-H2A tail peptides

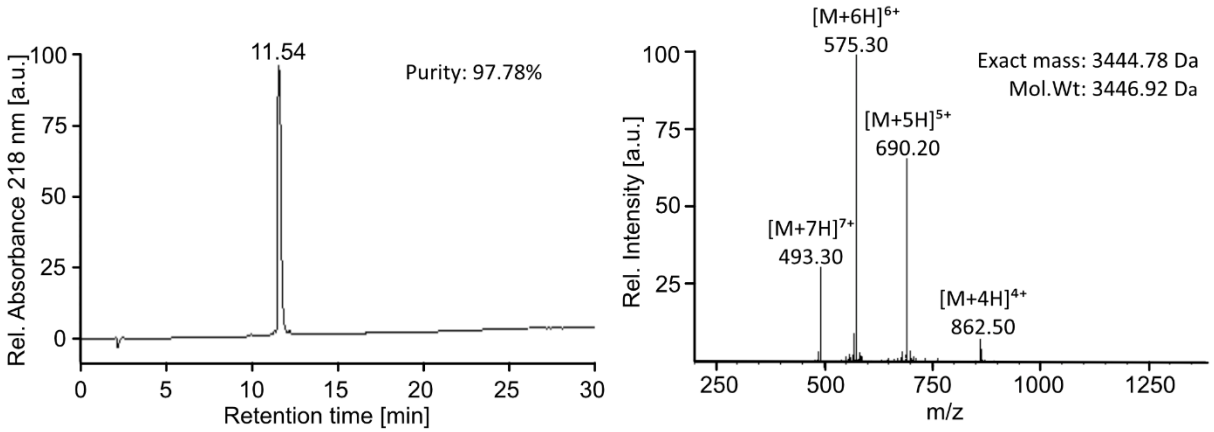
No.1 (H2A)



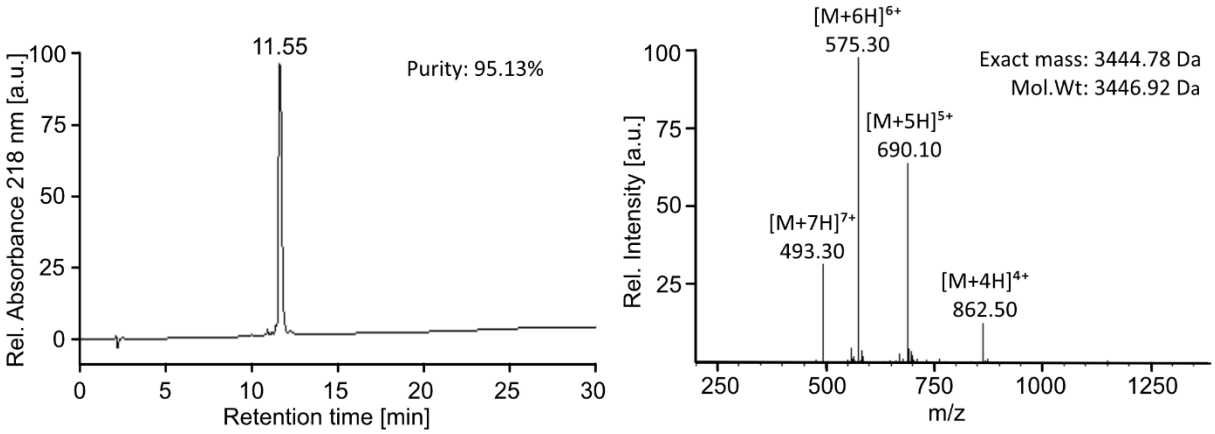
No.2 (H2A-Kac5)



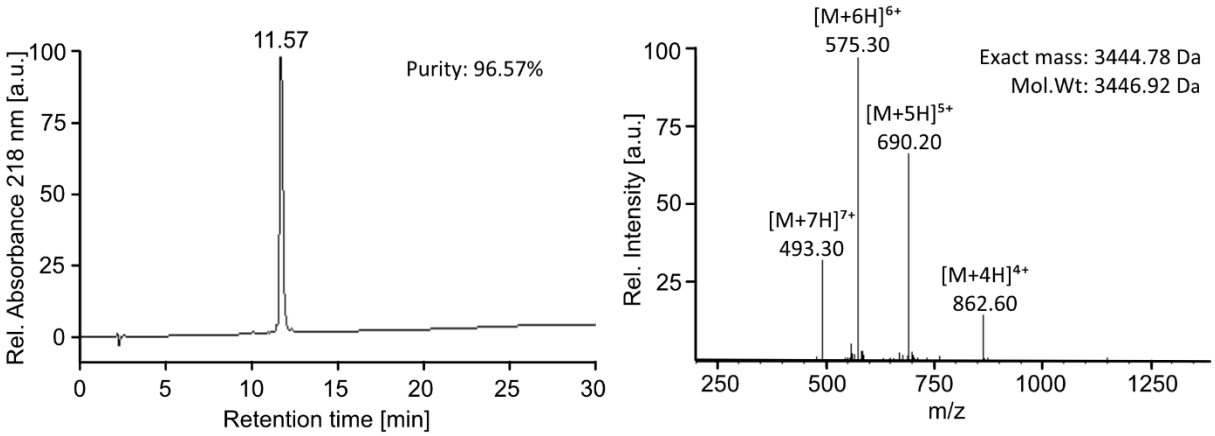
No. 3 (H2A-Kac9)



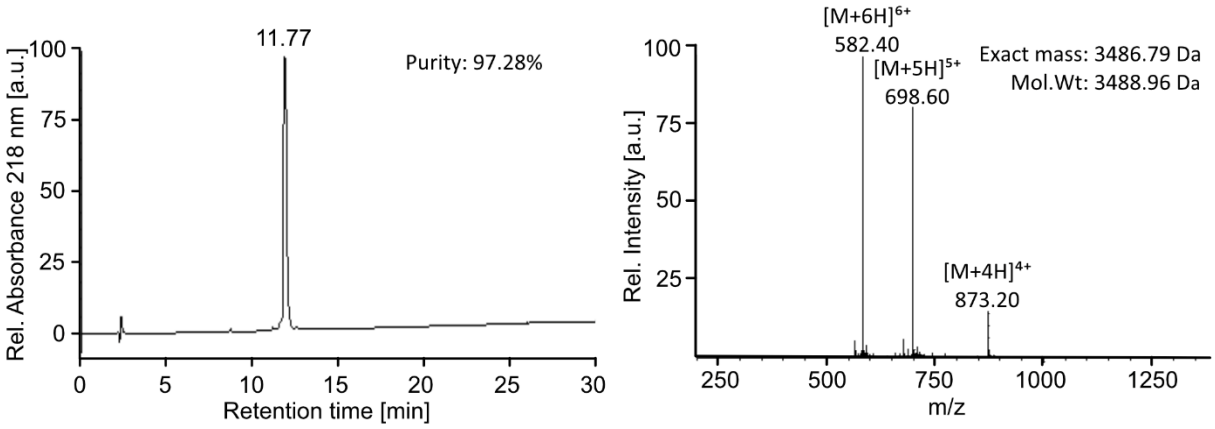
No.4 (H2A-Kac13)



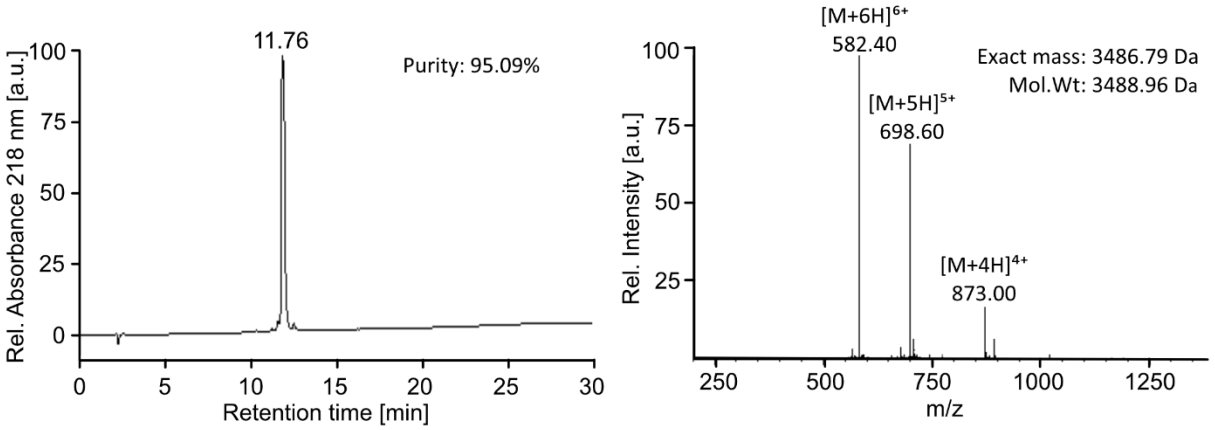
No.5 (H2A-Kac15)



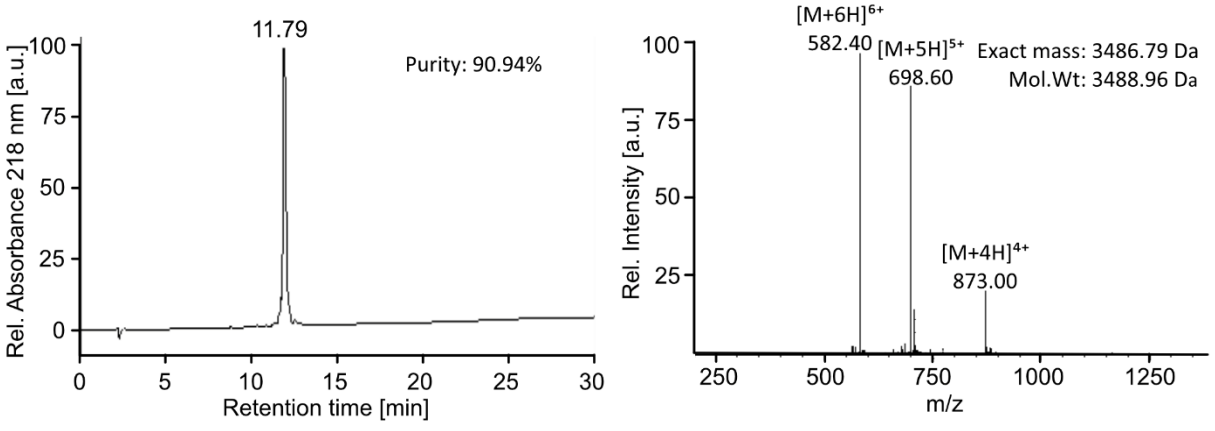
No.6 (H2A-Kac5-9)



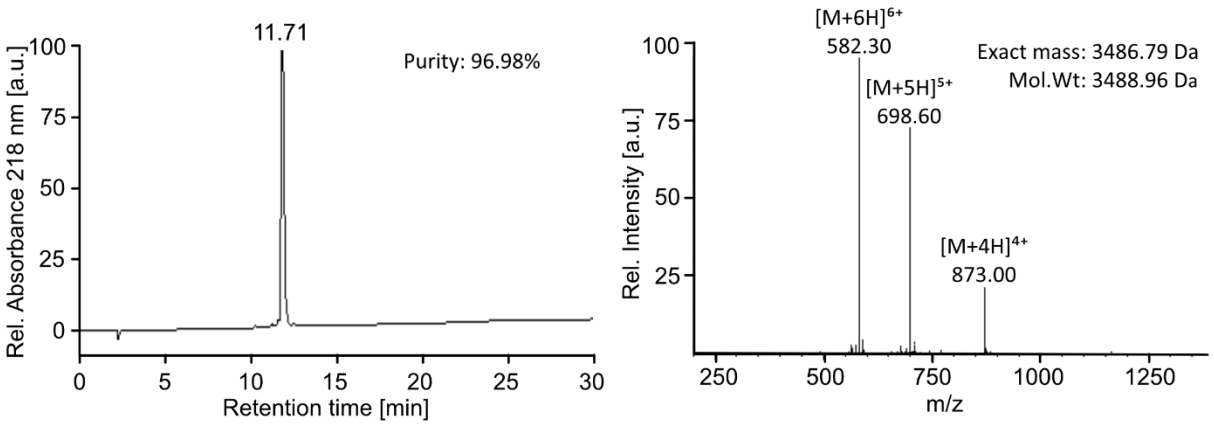
No.7 (H2A-Kac5-13)



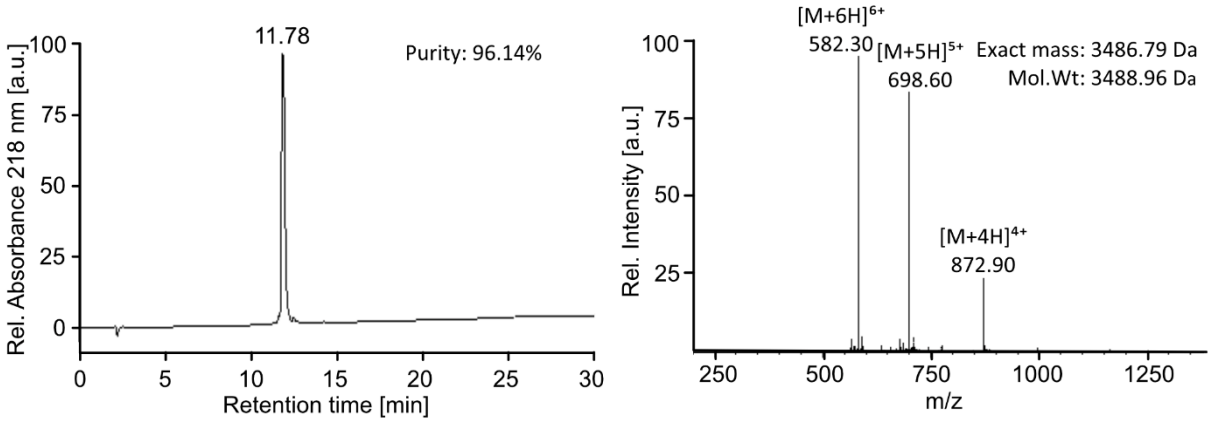
No.8 (H2A-Kac5-15)



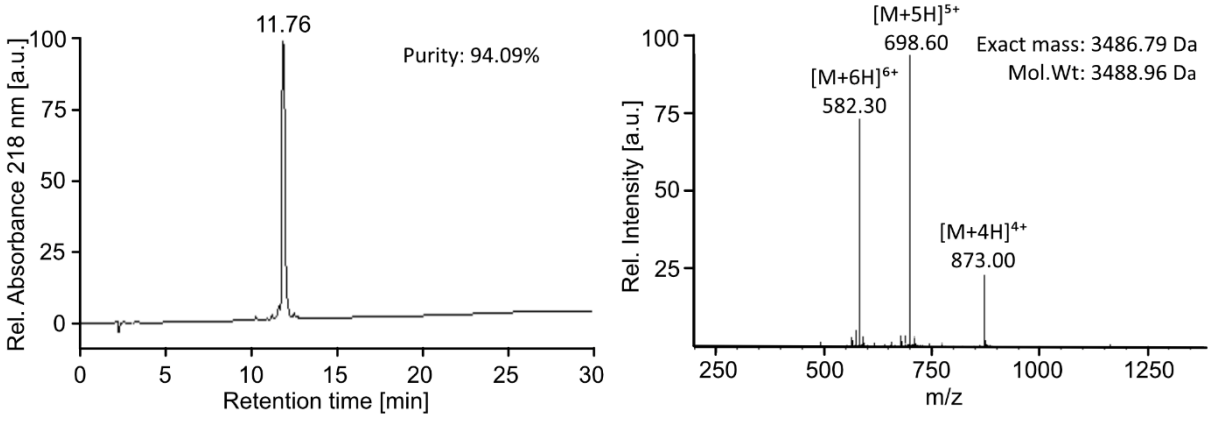
No.9 (H2A-Kac9-13)



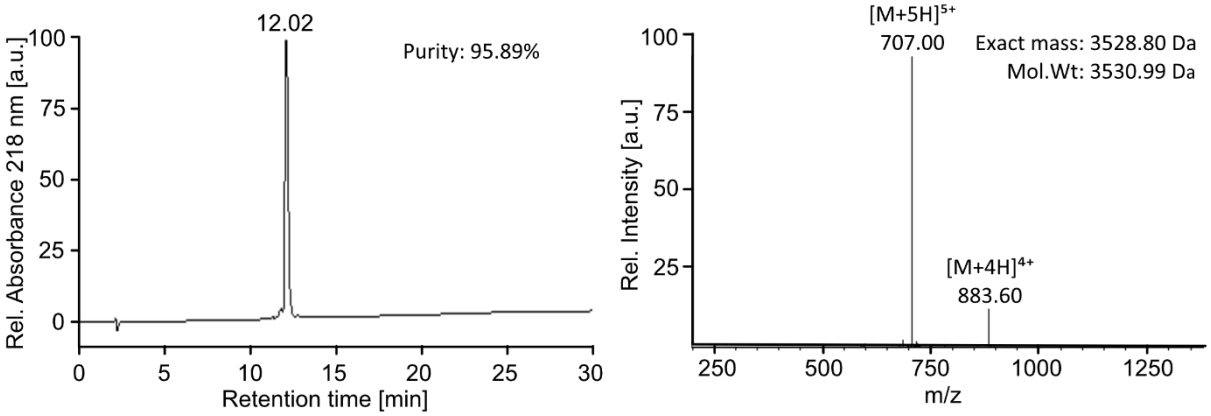
No.10 (H2A-Kac9-15)



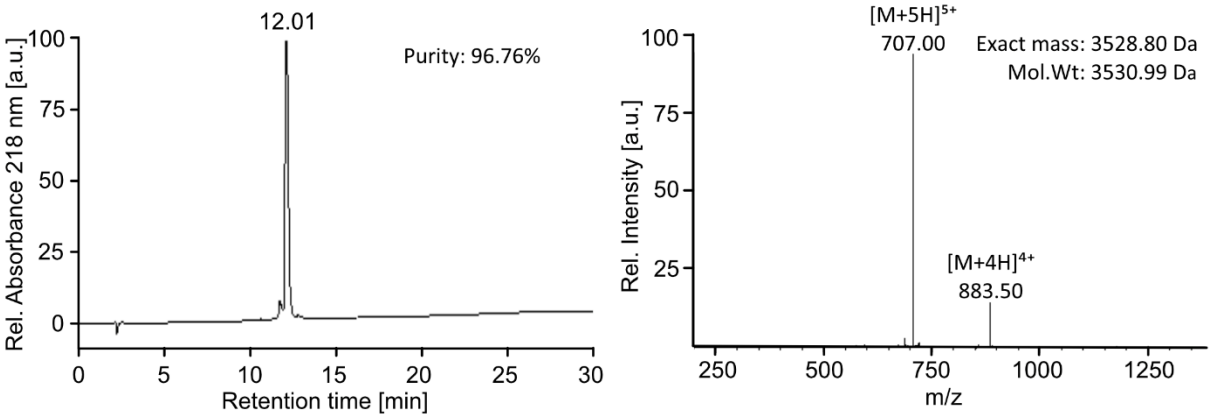
No.11 (H2A-Kac13-15)



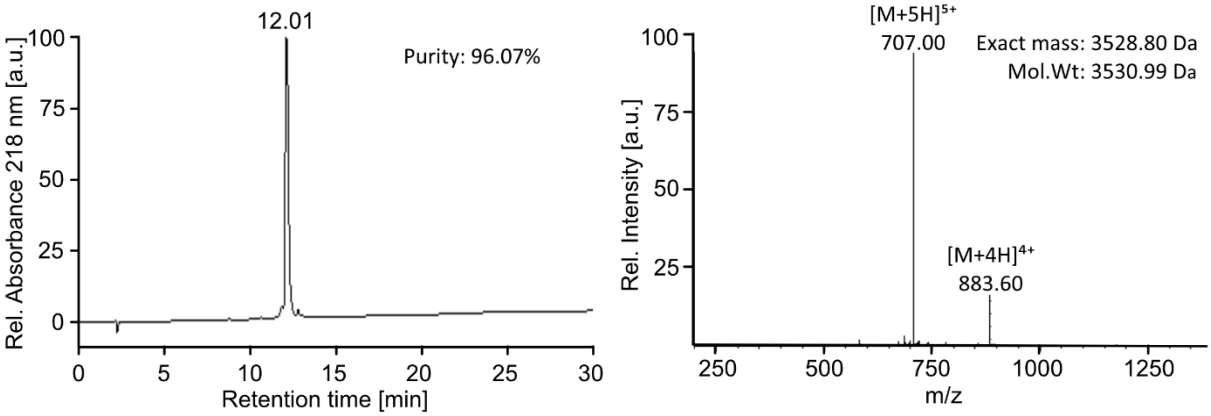
No.12 (H2A-Kac5-9-13)



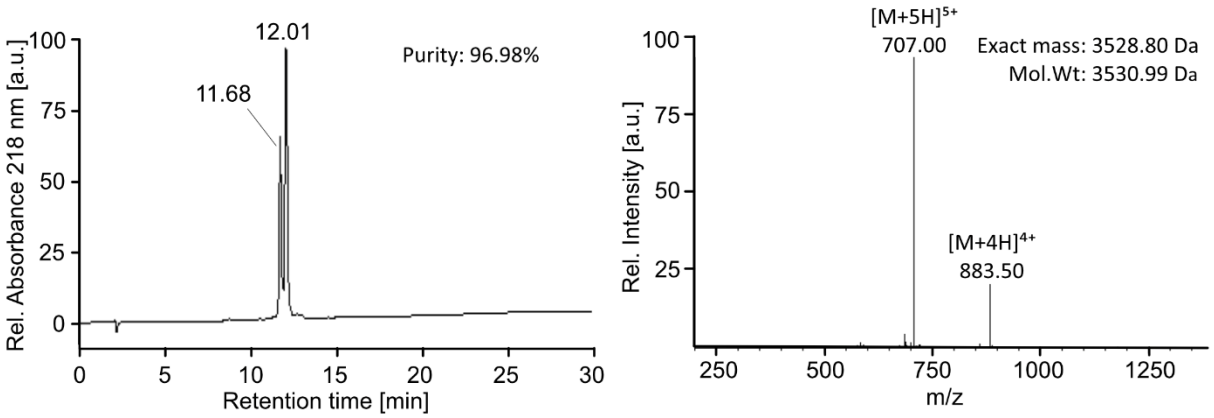
No.13 (H2A-Kac5-9-15)



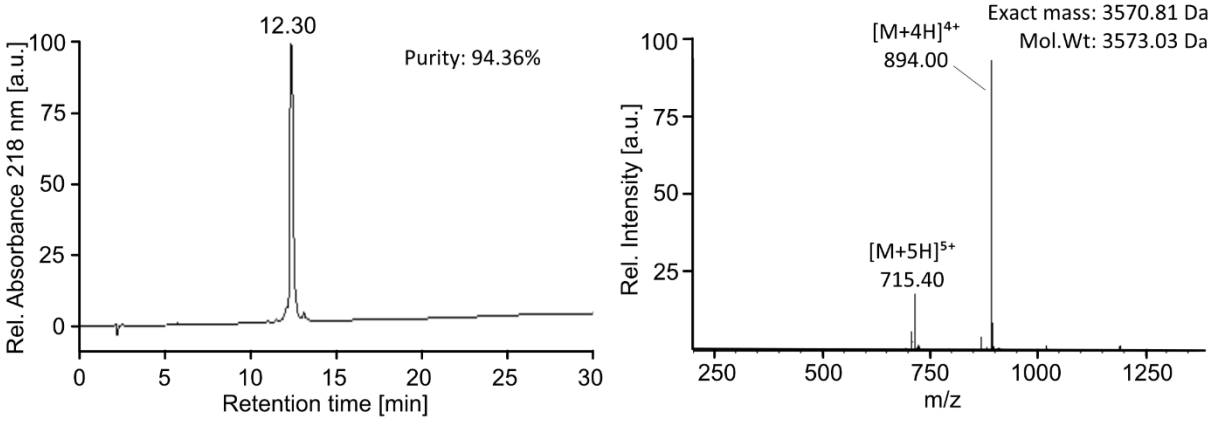
No.14 (H2A-Kac5-13-15)



No.15 (H2A-Kac9-13-15)



No.16 (H2A-Kac5-9-13-15)



8.3 Protein sequences

BAZ2B-mTagBFP

```
  1 MGSSHHHHHH SSGLVPRGSH MSVKKPKRDD SKDLALCSMI LTEMETHEDA
 51 WPFLLPVNLK LVPGYKKVIK KPMDFSTIRE KLSSGQYPNL ETFALDVRLV
101 FDNCETFNED DSDIGRAGHN MRKYFEKKWT DTFKVSGGEN LYFQSGGELI
151 KENMHMKLYM EGTVDNHHFK CTSEGEGKPY EGTQTMRIKV VEGGPLPFAF
201 DILATSFLYG SKTFINHTQG IPDFFKQSFP EGFTWERVTT YEDGGVLTAT
251 QDTSLQDGCL IYNVKIRGVN FTSNGPVMQK KTLGWEAFTE TLYPADGGLE
301 GRNDMALKLV GGSHLIANIK TTYRSKKPAK NLKMPGVYYV DYRLERIKEA
351 NNETYVEQHE VAVARYCDLP SKLGHKLNGG GWSHPQFEK
```

CREBBP-TurboYFP

```
  1 MGSSHHHHHH SSGLVPRGSH MRKKIFKPEE LRQALMPTLE ALYRQDPESL
 51 PFRQPVDPQL LGIPDYFDIV KNPMDLSTIK RKLDTGQYQE PWQYVDDVWL
101 MFNNAWLYNR KTSRVYKFCS KLAEVFEQEI DPVMQSLGGG ENLYFQSGGS
151 SGALLFHGKI PYVVEMEGNV DGHTFSIRGK GYGDASVGKV DAQFICTTGD
201 VPVPWSTLVT TLTYGAQCFA KYGPELKDFY KSCMPDGYVQ ERTITFEGDG
251 NFKTRAEVTF ENGSVYNRVK LNGQGFKKDG HVLGKNLEFN FTPHCLYIWG
301 DQANHGLKSA FKICHEITGS KGDFIVADHT QMNTPIGGGP VHVPEYHHMS
351 YHVKLSKDVT DHRDNMSLKE TVRAVDCRKT YDFDAGSGDT SGGGWSHPQF
401 EK
```

BRD3(2)-TurboYFP

```
  1 MGSSHHHHHH SSGLVPRGSH MGKLSEHLRY CDSILREMLS KKHAAYAWPF
 51 YKPVDAEALE LHDYHDIIKH PMDLSTVKRK MDGREYPDAQ GFAADVRLMF
101 SNCYKYNPPD HEVVAMARKL QDVFEMRFAK MPGGENLYFQ SGGSSGALLF
151 HGKIPYVVEM EGNVDGHTFS IRGKGYGDAS VGKVDAQFIC TTGDVPVPWS
201 TLVTTLTYGA QCFAKYGPEL KDFYKSCMPD GYVQERTITF EGDGNFKTRA
251 EVTFENGSVY NRVKLNGQGF KKDGHVLGKN LEFNFTPHCL YIWGDQANHG
301 LKSAFKICHE ITGSKGDFIV ADHTQMNTPI GGGPVHVPEY HHMSYHVKLS
351 KDVTDHRDNM SLKETVRAVD CRKTYDFDAG SGDTSGGGWS HPQFEK
```

BRD4(1/2)-mTagBFP2

```
  1 MGSSHHHHHH SSGLVPRGSH MNPPPPETSN PNKPKRQTNQ LQYLLRVVLK
 51 TLWKHQFAWP FQQPVDAVKL NLPDYYKIIK TPMDMGTIKK RLENNYYWNA
101 QECIQDFNTM FTNCYIYNKP GDDIVLMAEA LEKLFLQKIN ELPTEETEIM
151 IVQAKGRGRG RKETGTAKPG VSTVPNTTQA STPPQTQTPQ PNPPPVQATP
201 HPFPAVTPDL IVQTPVMTVV PPQPLQTPPP VPPQPQPPPA PAPQPVQSHP
251 PIIAATPQPV KTKKGVKRKA DTTTPTTIDP IHEPPSLPPE PKTTKLGQRR
301 ESSRPVKPPK KDVPDSQQHP APEKSSKVSE QLKCCSGILK EMFAKKHAAY
351 AWPFYKPVDV EALGLHDYCD IIKHPMDMST IKSKLEAREY RDAQEFGADV
401 RLMFSNCYKY NPPDHEVVAM ARKLQDVFEM RFAKMPDEGG ENLYFQSGGG
451 SVSKGEELIK ENMHMKLYME GTVDNHHFKC TSEGEGKPYE GTQTMRIKVV
501 EGGPLPFAFD ILATSFLYGS KTFINHTQGI PDFFKQSFPE GFTWERVTTY
551 EDGGVLTATQ DTSLQDGCLI YNVKIRGVNF TSNGPVMQKK TLGWEAFTET
601 LYPADGGLEG RNDMALKLVG GSHLIANAKT TYRSKKPAKN LKMPGVYYVD
651 YRLERIKEAN NETYVEQHEV AVARYCDLPS KLGHKLN
```

IntC-MBP-S6

```
  1 MSTGKRVPIK DLLGEKDFEI WAINEQTMKL ESAKVSRVFC TGKKLVYTLK
 51 TRLGRTIKAT ANHRFLTIDG WKRLDELSLK EHIALPRKLE SSSLQLAPEI
101 EKLPQSDIYW DPIVSITETG VEEVFDLTVP GLRNFVANDI IVHNSIEGSG
151 GKIEEGKLVI WINGDKGYNG LAEVGKKFEK DTGIKVTVEH PDKLEEKFPQ
201 VAATGDGPDI IFWAHDRFGG YAQSGLLAEI TPDKAFQDKL YPFTWDAVRY
251 NGKLIAYPIA VEALSLIYNK DLLPNPPKTW EEIPALDKEL KAKGKSALMF
301 NLQEPYFTWP LIAADGGYAF KYENGKYDIK DVGVDNAGAK AGLTFLVDLI
351 KNKHMNADTD YSIAEAAFNK GETAMTINGP WAWSNIDTSK VNYGVTVLPT
401 FKGQPSKPFV GVLSAGINAA SPNKELAKEF LENYLLTDEG LEAVNKDKPL
451 GAVALKSYEE ELAKDPRIAA TMENAQKGEI MPNIPQMSAF WYAVRTAVIN
501 AASGRQTVDE ALKDAQTRIT KGSGGSGGDS LSWLLRLLNG SGGSGHHHHH
551 HGSGGSGWSH PQFEK
```

SrtA

```
  1 MQAKPQIPKD KSKVAGYIEI PDADIKEPVY PGPATPEQLN RGVSFAEENE
 51 SLDDQNISIA GHTFIDRPNY QFTNLKAAKK GSMVYFKVGN ETRKYKMTSI
101 RDVKPTDVGV LDEQKGKDKQ LTLITCDDYN EKTGVWEKRK IFVATEVKLE
151 HHHHHH
```

Sfp

```
  1 MEIYGIYMDR PLSQEENERF MTFISPEKRE KCRRFYHKED AHRTLLGDVL
 51 VRSVISRQYQ LDKSDIRFST QEYGKPCIPD LPDAHFNISH SGRWVIGAFD
101 SQPIGIDIEK TKPISLEIAK RFFSKTEYSD LLAKDKDEQT DYFYHLWSMK
151 ESFIKQEGKG LSLPLDSFSV RLHQDGQVSI ELPDSHSPCY IKTYEVDPGY
201 KMAVCAAHPD FPEDITMVSY EELLRSGGLE HHHHHH
```