Semisynthesis of Ubiquitinated Histone H2B via Native Chemical Ligation

Meihuan Ji

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2022

Committee:

Champak Chatterjee

Jesse Zalatan

Program Authorized to Offer Degree:

Chemistry

University of Washington

Abstract

Semisynthesis of Ubiquitinated Histone H2B via Native Chemical Ligation

Meihuan Ji

Chair of the Supervisory Committee:
Champak Chatterjee
Department of Chemistry

The ubiquitylation of histone H2B at Lys (K) 120 is known to be associated with methylation of H3K4 and H3K79, transcription elongation, and DNA repair. Mechanistic studies of interactions between H2BK120ub and other chromatin-associated proteins, as well investigations of the role of ubiquitylation in chromatin function require milligram scale production of H2BK120ub. However, the direct synthesis of H2BK120ub by solid-phase peptide synthesis (SPPS) has not been possible due to limitations in yields of SPPS for long proteins. In this thesis, truncated H2B(1-116) and Ub(1-75) C-terminal α-thioesters were purified by intein fusion. An H2B(117-125) peptide fragment with an Ala117 to Cys mutation and acetamidomethyl protected Cys, Cys(Acm), attached to Lys120 was synthesized by SPPS. Finally, semisynthesis of H2BK120ub was performed by native chemical ligation (NCL) between various thioester fragments. The high purity of semisynthetic H2BK120ub was confirmed by ESI-MS and 15% SDS-PAGE, the

ubiquitylated histone was used for octamer and mononucleosome reconstitution. The ubiquitylated mononucleosomes will be tested in numerous biological assays in the future to understand how H2BK120ub influences the function of human chromatin.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The thesis could not be completed without the support of my research advisor Dr. Champak Chatterjee, MS facility manager Dr. Martin Sadilek, NMR facility manager Dr. Rajan Paranji, and all of our last/present lab members (Dr. Caroline Weller, Dr. Patrick Shelton, Dr. Sumeet Singh, Dr. Calvin Leonen, Yvonne (Hongyi) Zhu, Andres Reyna, Madeline Currie, Nastassia Parker, Alex Russell, Lena Sobze, Lars McLaughlin, and Dylan Lynch). Through the one-year research, I really learned a lot, both academically and personally: lab techniques, trouble-shooting, revising experiments, explaining data, and well-being.

I would also like to thank my parents and grandma for their endless support. I started my study in the U.S back in 2013. High school, undergraduate, and graduate would be impossible for me without their help.

Chapter 1. INTRODUCTION TO H2BK120UB

1.1 INTRODUCTION TO HISTONE POST-TRANSLATIONAL MODIFICATION

Eukaryotic cells contain genomic DNA packaged in the form of chromatin, which is a large nucleoprotein complex. The nucleosome is the basic structural unit of chromatin and is composed of DNA and an octameric complex of histone proteins [1]. Two copies of H2A, H2B, H3 and H4 (core histone) form the histone octamer, and about 147 base pairs (bp) of DNA is wrapped about 1.65 times around the octamer to result in the nucleosome core particle (Figure 1.1A). This compact packaging of DNA enables the diploid human DNA (~6 million bp) to fit into the small cell nucleus. The N-terminal tails of histone are often covalently modified by post-translational modifications (PTM), for instance, acetylation [2], phosphorylation [3], methylation [4], and ubiquitylation (Figure 1.1B) [5]. Many associated “writer”, “reader”, and “eraser” proteins of these PTMs are crucial for regulating chromatin function [6]. For example, histone H3 Lys 4 trimethylation (H3K4me3) is known to be a mark strongly associated with gene expression [7]. The plant homeodomain (PHD) domain of PHF8 (PHD finger protein 8) acts as a “reader” to recognize H3K4me3 marks in H3 tail, and PHF8 is an “eraser” that demethylates H3K9me2, while the diseases associated with PHF8 include X-linked Intellectual Disability [8].

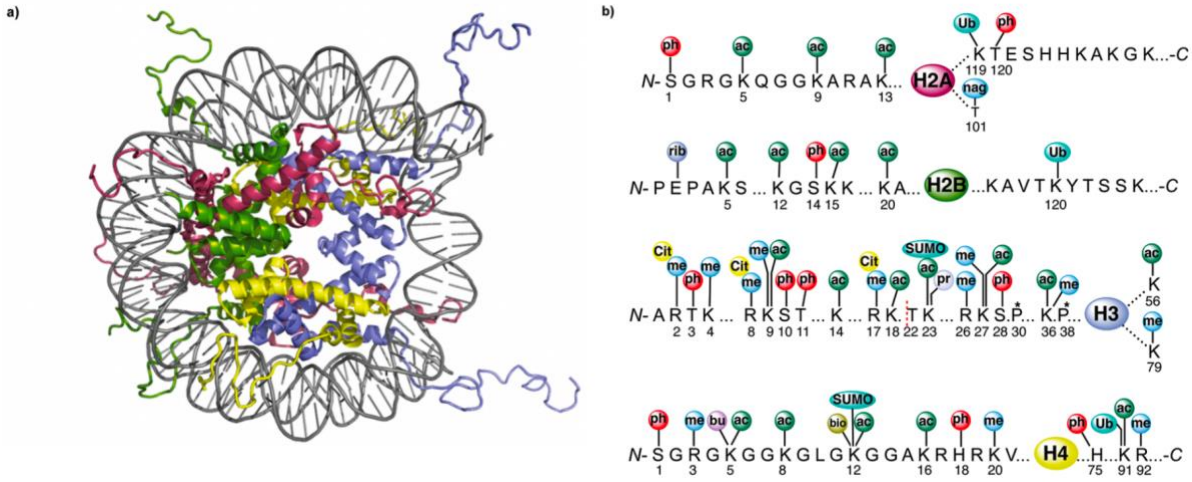

Figure 1.1. The structure of mononucleosome and range of histone PTMs.
(a) The mononucleosome structure at 1.9 Å resolution (PDB code 1KX5), with each histone and 147 bp double-stranded DNA displayed. The histone N-terminal tails stand out of the DNA for PTMs. Red = H2A; green = H2B; blue = H3; yellow = H4. (b) Schematic of the possible PTMs. Chemical groups abbreviation: ac = acetyl; bio = biotinyl; bu = butyryl; cit = ciyrullyl; me = methyl; nag = N-acetylglucosaminyl; ph = phosphoryl; pr = propionyl; rib = ADP-ribosyl; SUMO = SUMOyl; ub = ubiquitinyl [13].

1.2 H2BK120UB

The ubiquitylation of histone H2B at Lys 120 is well known to be associated with transcription elongation, DNA repair, and the trimethylation of H3K4 and H3K79 [9]. Any *in vitro* studies of the effect of ubiquitylation on chromatin and interactions between H2BK120ub and other proteins require its production in a milligram scale and high purity. However, the total chemical synthesis of H2BK120ub using solid-phase peptide synthesis (SPPS) is not possible due to technical limitations of synthesizing a maximum ~100 residues on the solid-phase [10]. Many semisynthesis approaches were reported as a solution for generating H2BK120ub. Chatterjee et al. produced H2BK120ub using intein-mediated H2B(1-116)-MES and Ub(1-75)-MES with the aid of an

intein-fusion, and a synthetic H2B(117-125) fragment containing a photocleavable protecting group at K120 for the Ub(1-75) ligation via native chemical ligation (NCL). The H2BK120ub was later used to investigate the stimulatory effect, in nucleosome substrates, on H3K79 methylation by hDot1L [11]. Another semisynthesis approach has also been reported to produce high-quality H2BK120ub without using photochemistry [12]. I pursued the goal of generating H2BK120ub with a Cysteine at the ubiquitylation site to be employed as a suicide substrate that may covalently trap enzymes associated with ubiquitylation and deubiquitylation.

1.3 REFERENCES

[1] Andrews, A. J.; Luger, K. Nucleosome Structure(s) and Stability: Variations on a Theme. *Annu. Rev. Biophys.* **2011**, *40* (1), 99–117.

[2] Phillips, D. M. The Presence of Acetyl Groups of Histones. *Biochem. J.* **1963**, *87*, 258–263.

[3] Stevely, W. S.; Stocken, L. A. Phosphorylation of Rat-Thymus Histone. *Biochem. J.* **1966**, *100* (2), 20C-1C.

[4] Murray, K. The Occurrence of Epsilon-n-Methyl Lysine in Histones. *Biochemistry* **1964**, *3*, 10–15.

[5] West, M. H.; Bonner, W. M. Histone 2B Can Be Modified by the Attachment of Ubiquitin. *Nucleic Acids Res.* **1980**, *8* (20), 4671–4680.

[6] Ruthenburg, A. J.; Allis, C. D.; Wysocka, J. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Mol. Cell* **2007**, *25* (1), 15–30.

[7] Sims, R. J., 3rd; Nishioka, K.; Reinberg, D. Histone Lysine Methylation: A Signature for Chromatin Function. *Trends Genet.* **2003**, *19* (11), 629–639.

[8] Qiu, J.; Shi, G.; Jia, Y.; Li, J.; Wu, M.; Li, J.; Dong, S.; Wong, J. The X-Linked Mental Retardation Gene PHF8 Is a Histone Demethylase Involved in Neuronal Differentiation. *Cell Res.* **2010**, *20* (8), 908–918.

[9] Weake, V. M.; Workman, J. L. Histone Ubiquitination: Triggering Gene Activity. *Mol. Cell* **2008**, *29* (6), 653–663.

[10] Dawson, P. E.; Kent, S. B. Synthesis of Native Proteins by Chemical Ligation. *Annu. Rev. Biochem.* **2000**, *69* (1), 923–960.

[11] Chatterjee, C.; McGinty, R. K.; Pellois, J.-P.; Muir, T. W. Auxiliary-Mediated Site-Specific Peptide Ubiquitylation. *Angew. Chem. Int. Ed Engl.* **2007**, *46* (16), 2814–2818.

[12] Kumar, K. S. A.; Spasser, L.; Ohayon, S.; Erlich, L. A.; Brik, A. Expeditious Chemical Synthesis of Ubiquitinated Peptides Employing Orthogonal Protection and Native Chemical Ligation. *Bioconjug. Chem.* **2011**, *22* (2), 137–143.

[13] Dhall, A.; Chatterjee, C. Chemical Approaches to Understand the Language of Histone Modifications. *ACS Chem. Biol.* **2011**, *6* (10), 987–999.

Chapter 2. PRODUCTION OF H2B(1-116)-MES, H2B(117-125) FRAGMENT, AND UB(1-75)-MES

2.1 INTRODUCTION

For the semisynthesis of H2BK120ub via a NCL approach, two protein C-terminal α-thioesters and a peptide with N-terminal Cys residues were first needed. Hence, in this chapter, H2B(1-116) MES and a fully synthetic H2B(117-125) fragment with an A117C mutation and a Cys residue attached at Lys 120 protected by an acetamidomethyl (Acm) group, and Ub(1-75)-MES were produced. The first ligation of H2B(1-116)-MES and the synthesized H2B(117-125) fragment resulted in full-length H2B with a Cys residue attached at Lys 120, which would be the C-teminal Cys76 residue upon the second ligation to Ub(1-75)-MES. The details of the semisynthesis will be discussed in Chapter 3.

2.1.1 *Overexpression and Intein Fusion Purification*

The transformed *E. coli* BL21(DE3) competent cells contained the LacI gene that expresses the lac repressor, which blocks the Lac operator (LacO) site before the T7 promoter to suppress the expression of our desired target genes. Hence, during the overexpression of H2B(1-116)-GyrA-His_6 and Ub(1-75)-Ava-His_6, Isopropyl β-D-1-thiogalactopyranoside (IPTG), an allolactose mimic, was used for the removal of LacI from the LacO site to enable the binding of T7 RNA Polymerase to the T7 promoter and to induce the expression of the gene of interest (Figure 2.1) [1]. The H2B(1-116) and Ub(1-75) C-terminal α-thioesters could then be purified from intein fusions. The two proteins of interest were hexahistidine-tagged, which allowed binding to the Ni-NTA resin and purification by eluting with high concentrations of the competitive Ni^{2+}-binding agent imidazole. Inteins are single-turnover enzymes that catalyze a protein splicing reaction. This

involves the intein N-terminal Cysteine thiol attacking the C-terminal amide of the N-extein to form an intial α-thioester intermediate, followed by the joining of the two exteins fragments and removal of the intein to yield the final ligated protein. Taking advantage of protein splicing, a C-terminally mutated intein was attached to the H2B(1-116) and Ub(1-75) C-termini, so that they could produce a thioester intermediate but no further splicing was possible due to the absence of a C-terminal extein and catalytic residues in the intein that favor splicing. Upon the addition of an excess of the thiol reagent 2-mercaptoethanesulfonic acid (MESNa), C-terminal α-thioesters were released from the N-terminus of the intein [3].

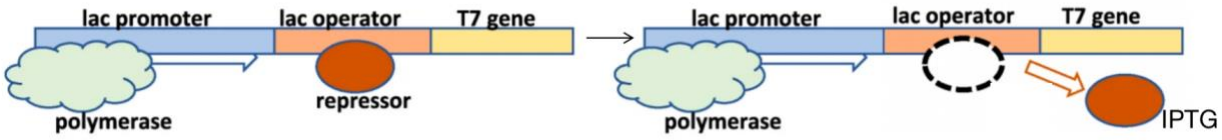

Figure 2.1. **Schematic of IPTG induction.** LacI (repressor) was released upon IPTG addition to allow the binding of T7 RNA Polymerase. Adapted from [2].

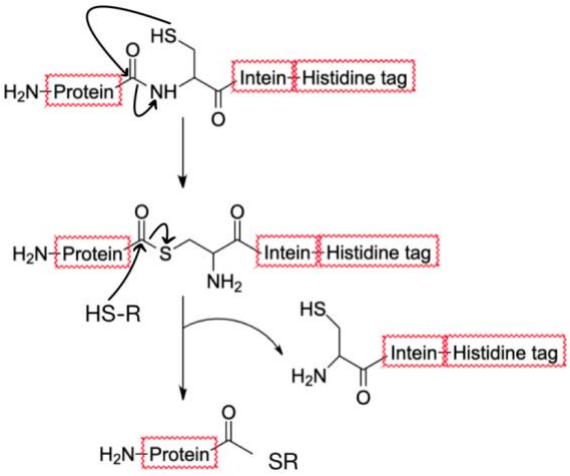

Figure 2.2. **Schematic of C-terminal α-thioester purification from intein fusion.** An *N*,*S*-acyl shift occurs spontaneously by attack on the C-terminal protein amide bond by Cys in the N-terminal intein-Histidine tag fusion. This yields an intein-linked thioster intermediate. The intein-Histidine tag is removed by thiol addition.

2.1.2 *Solid-Phase Peptide Synthesis*

Fmoc-based solid-phase peptide synthesis (SPPS) is a process of coupling N-terminally Fmoc-protected amino acids to the growing peptide chain using condensation reagents such as Oxyma Pure™ (Ethyl cyano(hydroxyimino)acetate) and diisopropylcarbodiimide (DIC). In the coupling reaction, the reaction of DIC and the unprotected amino acid C-terminus forms an *O*-acylurea intermediate, and the addition of Oxyma Pure leads to the formation of an activated Oxyma Pure-ester with the release of *N*, *N*'-diisopropylurea. The regeneration of Oxyma Pure and the formation of a peptide bond occurs upon attack of the Oxyma Pure-amino acid ester by the deprotected N-terminal amine of the growing peptide chain [4].

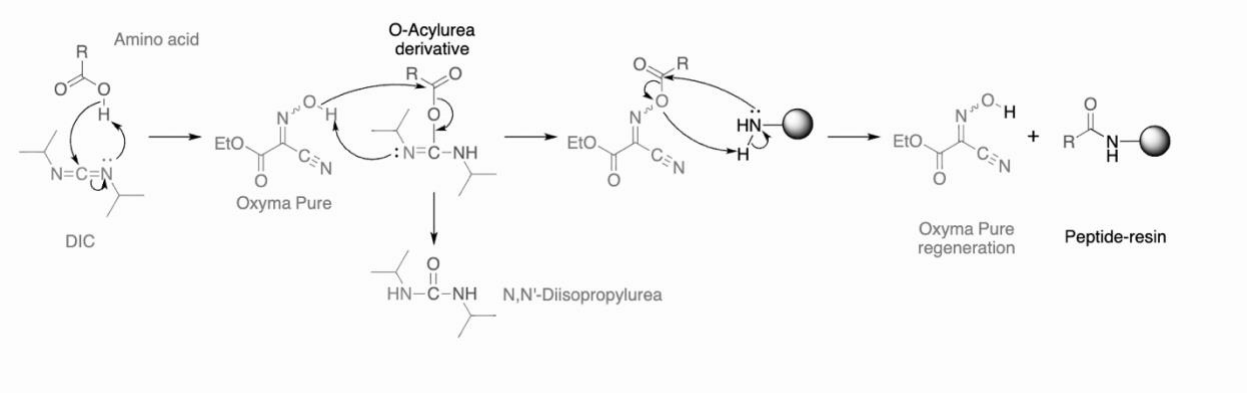

Figure 2.3. **Schematic of peptide coupling.** DIC and the unprotected amino acid react to form an *O*-acylurea intermediate. Further displacement by Oxyma Pure leads to formation of the Oxyma Pure-amino acid ester. Amide bond formation and Oxyma Pure regeneration occur upon aminolysis by the growing peptide chain N-terminus. Adapted from [5].

2.2 Results and Discussion

2.2.1 *H2B(1-116)-MES Production*

With a well-developed and optimized protocol, IPTG-induced H2B(1-116)-GyrA-His_6 was highly overexpressed with 0.35 mM of IPTG after growing for 3.5 h at 25 °C (Figure 2.4A) (~20 g of cell

pellet was collected per 6 L of LB media). By eluting the bound protein from a Ni^{2+} column using an increasing concentration of imidazole solution (20 mM to 500 mM), His-tagged H2B(1-116) was purified (washing with imidazole below 500 mM) and eluted (with 500 mM imidazole). Since the truncated H2B is insoluble, as shown in Figure 2.4C, the majority of H2B(1-116)-GyrA-His_6 was present in the dissolved pellet supernatant (inclusion body). Although the amount of H2B(1-116)-GyrA-His_6 in the cell lysate supernatant was very low (Figure 2.4B), it too was bound to Ni^{2+} and the 500 mM imidazole elution fractions of inclusion body associated and soluble H2B(1-116)-GyrA-His_6 were combined prior to thiolysis as the yield of H2B(1-116)-MES α-thioester was very low (~5 mg per batch). As shown in Figure 2.5, after 16 h of thiolysis reaction at 4 ºC, all intein fusion in the cell lysate supernatant and the majority of protein in the dissolved pellet supernatant were converted into H2B(1-116)-MES. As not all of the intein fusion was released in dissolved pellet supernatant, extra MESNa power could be added, and/or the reaction time could be extended until no H2B(1-116)-GyrA-His_6 remained. For the second Ni^{2+} purification, the column flow-through was collected, compared to imidazole elutions for the first Ni^{2+} purification, because the goal was to remove the His-tagged protein from cleaved H2B(1-116)-MES α-thioester. As a result, only H2B(1-116)-MES was seen in the 15% Coomassie-stained SDS-PAGE gel after the purification. After dialysis to remove some of the Guanidine-HCl and Na_2HPO_4, the purity of crude H2B(1-116)-MES was checked by the gel. Because the yield of H2B(1-116)-MES was so low (the best pure yield was 8 mg per batch), the crude product did not undergo further purification by RP-HPLC to avoid losses. Instead, the pure H2B(1-116)-MES C4 RP-HPLC chromatogram and ESI-MS spectrum were only obtained once for verification purposes (Figure 2.6). C4 analytical column was chosen over C18 because of the hydrophobicity of H2B(1-116)-MES . 6 M

Guanidine-HCl was used to dissolve the sample and resulted in a strong buffer/salt absorbance at ~2 min into the run.

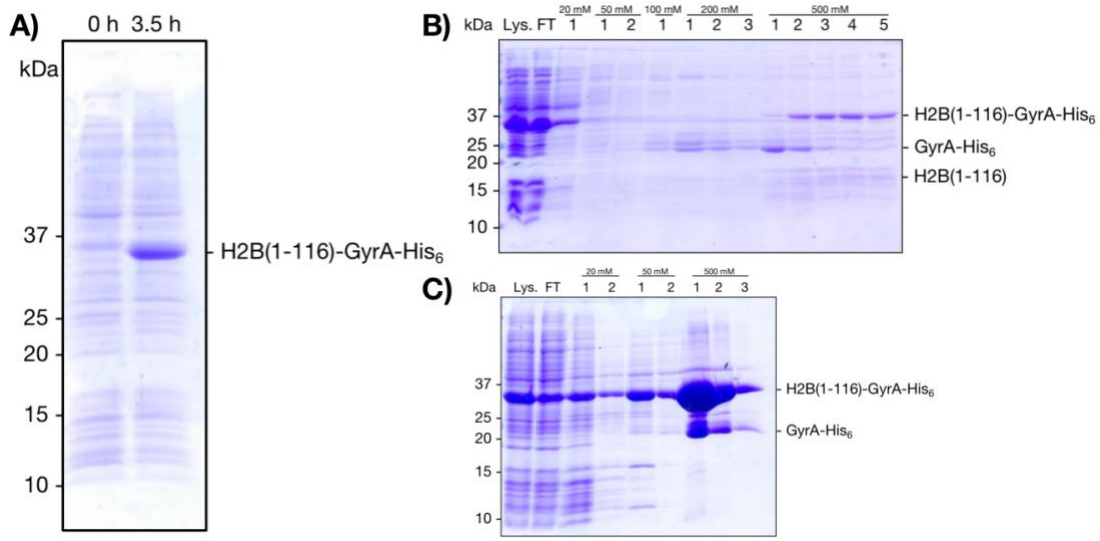

Figure 2.4. **Overexpression and Purification of H2B(1-116)-GyrA-His$_6$.** 15% Coomassie-stained SDS-PAGE gel of (A) H2B(1-116)-GyrA-His$_6$ overexpression by IPTG induction; (B) first Ni^{2+} column purification of cell lysate supernatant; (C) first Ni^{2+} column purification of dissolved pellet supernatant. Lys. = cell lysate supernatant. FT = flow-through after protein binds to the column.

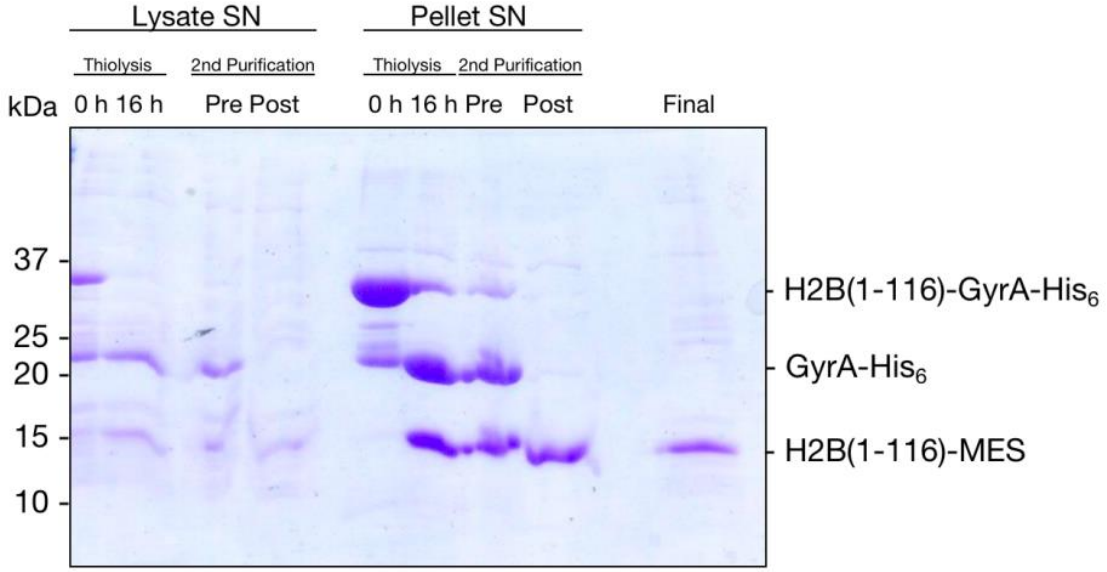

Figure 2.5. **15% Coomassie-stained SDS-PAGE gel of crude H2B(1-116)-MES.** Thiolysis was performed by adding MESNa powder to the dialyzed pooled fractions containing H2B(1-116)-

GyrA-His_6, at a final concentration of 100 mM MESNa, pH 7.5, at 4 ºC. Second purification removing H2B(1-116)-GyrA -His_6 and GyrA-His_6 was achieved using Ni^{2+} column. The purity of crude H2B(1-116)-MES after dialysis is labeled as "Final". SN = supernatant.

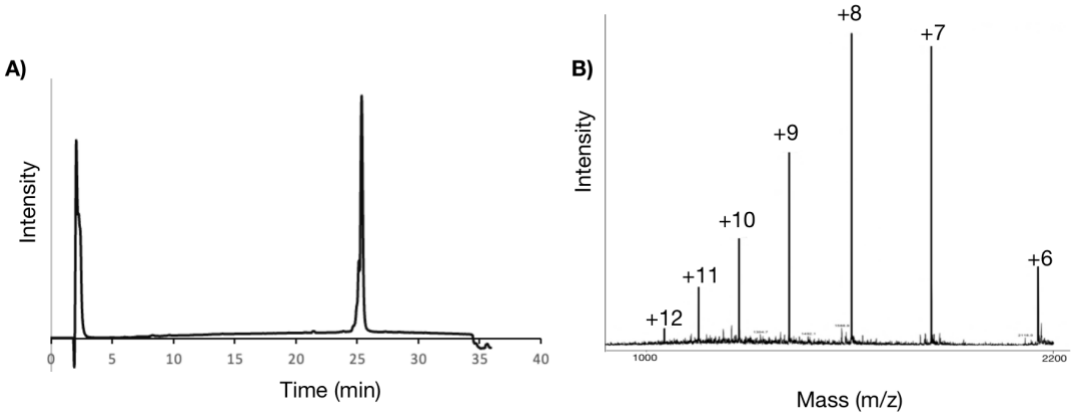

Figure 2.6. **RP-HPLC Purification of H2B(1-116)-MES.** (A) C4 analytical RP-HPLC chromatogram of purified H2B(1-116)-MES α-thioester with the gradient of 0-73% B, 30 min; (B) ESI-MS of purified H2B(1-116)-MES. Solvent A = 0.1% TFA in Millipore water. Solvent B = 0.1% TFA in 90% Acetonitrile prepared in Millipore water.

2.2.2 *Ub(1-75)-MES Production*

By revising an existing protocol, IPTG-induced Ub(1-75)-Ava-His_6 was overexpressed well with 0.75 mM of IPTG after growing for 4 h at 30 ºC (Figure 2.7A) (~30 g of cell pellet was collected per 4 L of LB media). As the truncated Ub(1-75)-Ava-His_6 protein was very soluble, only the cell lysate supernatant was saved for purification. By passing through the Ni^{2+} column using an increasing concentration of imidazole solution (5 mM to 250 mM), His-tagged Ub(1-75) was purified (washing with imidazole below 250 mM) and eluted (with imidazole at 250 mM). As shown in Figure 2.7B, Ub(1-75)-Ava-His_6 was present in the 50 mM (wash) and 250 mM (elution) fractions. After 18 h of thiolysis reaction at 30 ºC, the majority of Ub(1-75)-Ava-His_6 in both pooled fractions were converted into Ub(1-75)-MES (Figure 2.7C). As not all of the intein fusion was released, extra MESNa power could be added, and/or the reaction time could be extended until no Ub(1-75)-Ava-His_6 remained. Since the yield of Ub(1-75)-MES was very high (~50 mg

of pure product per batch), there was no need for the second Ni^{2+} purification as for H2B(1-116)-MES production. Because the band of Ub(1-75)-MES was very thick on the 15% Coomassie-stained SDS-PAGE gel, Ub(1-75)-OH could not be distinguished from it. Later ESI-MS spectrum after RP-HPLC (Figure 2.8) showed that Ub(1-75)-OH (minor impurity as seen from ESI-MS) was present in the purified sample. However, it could not be separated from Ub(1-75)-MES α-thioester by RP-HPLC since the pure C18 analytical chromatogram showed a single peak. As the pH was not regulated after adding MESNa powder for thiolysis reaction versus re-adjusting pH during H2B(1-116)-MES production, the high content of Ub(1-75)-OH inside the pure product compared to almost no H2B(1-116)-OH in the crude H2B(1-116)-MES indicated that maintaining the thiolysis reaction at pH 7.2 may be crucial for reducing the hydrolysis of Ub(1-75)-MES α-thioester.

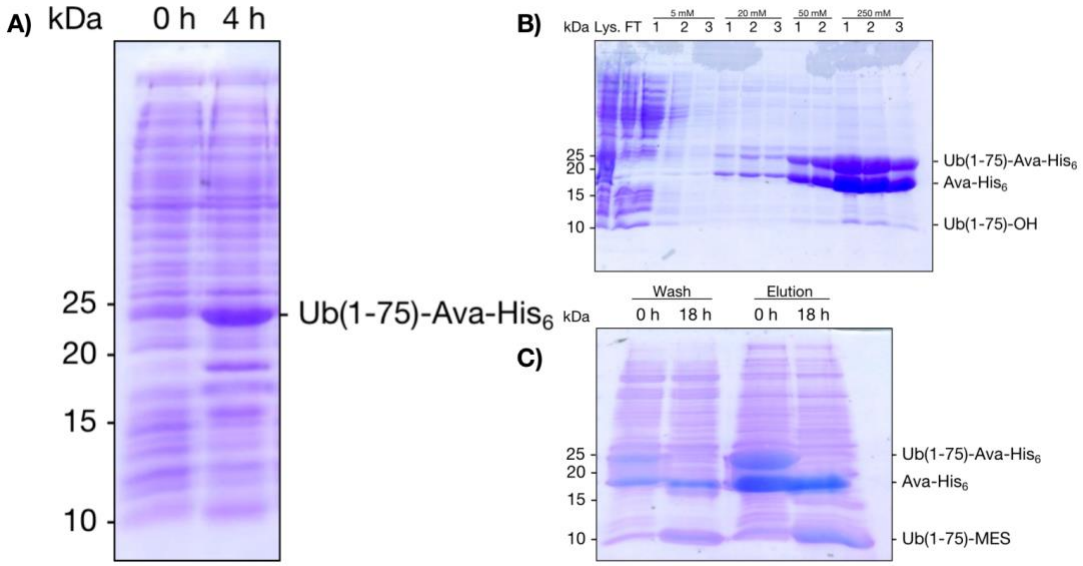

Figure 2.7. **Overexpression and Purification of Ub(1-75)-Ava-His$_6$.** 15% Coomassie-stained SDS-PAGE gel of (A) Ub(1-75)-Ava-His$_6$ overexpression upon induction with IPTG; (B) Ni^{2+} column purification of cell lysate supernatant; (C) thiolysis of dialyzed wash and elution pooled fractions containing Ub(1-75)-Ava-His$_6$ by adding MESNa powder to them at a final concentration

of 100 mM at 30 ºC. Lys. = cell lysate supernatant. FT = flow-through after protein binding to the column.

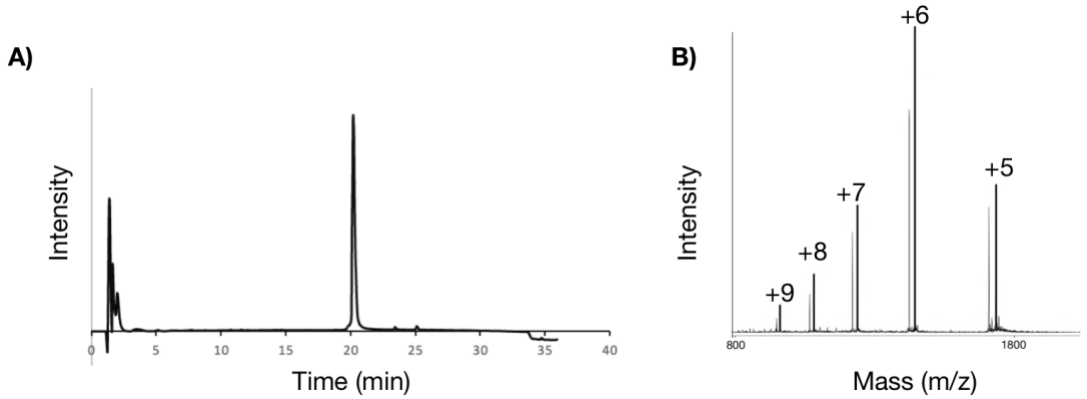

Figure 2.8. **RP-HPLC Purification of Ub(1-75)-MES.** (A) C18 analytical RP-HPLC chromatogram of purified Ub(1-75)-MES α-thioester with the gradient of 0-73% B, 30 min; (B) ESI-MS of purified Ub(1-75)-MES α-thioester. The minor unlabeled peak corresponds to hydrolyzed Ub(1-75)-OH.

2.2.3 *H2B(117-125) Peptide Synthesis*

The C-terminal Lys in H2B(117-125) was purchased pre-loaded on the Wang resin as Fmoc-Lys(Boc)-Wang resin (Novabiochem). After Fmoc deprotection of Lys via deprotonation by a secondary amine (piperazine) [6] and β-elimination of dibenzofulvene, each amino acid was sequentially coupled to the growing N-terminus using DIC/Oxyma Pure activators. K120 was protected by an allyloxycarbonyl (Alloc) group for subsequent Cys(Acm) coupling. The N-terminal Cys was protected by a *t*-butyloxycarbonyl (Boc-) group, which would not be deprotected during Alloc deprotection (Figure 2.9). The completion of the synthesized peptide H2B(117-125) was checked by a test cleavage and ESI-MS, and additional Boc-Cys coupling could be performed without prior Fmoc deprotection to ensure the N-terminal Boc-Cys was completely coupled. The pure intermediate H(Boc)N-CVTKYTSAK-C(O)OH was characterized by RP-HPLC and ESI-MS as shown in Figure 2.10A and B. The Alloc group was then removed from the K120 side-chain of the peptidyl-resin intermediate by palladium-catalyzed allyl transfer to phenylsilane as the

scavenger [7], followed by the Cys(Acm) coupling to result in the desired H2B(117-125) fragment 2 for subsequent NCL (Figure 2.9). 80% (v/v) TFA, 5% (v/v) phenol, 5% (v/v) thioanisole, 5% (v/v) H_2O, 2.5% (v/v) EDT, and 2.5% (v/v) TIS was tested and found to be the optimal cleavage cocktail for peptide release from the resin, instead of the commonly used 95% (v/v) TFA, 2.5% (v/v) H_2O, and 2.5% (v/v) TIS because of its more potent scavenging effect. The RP-HPLC chromatogram and ESI-MS spectrum of pure intermediate **2** are shown in Figure 2.11A and B.

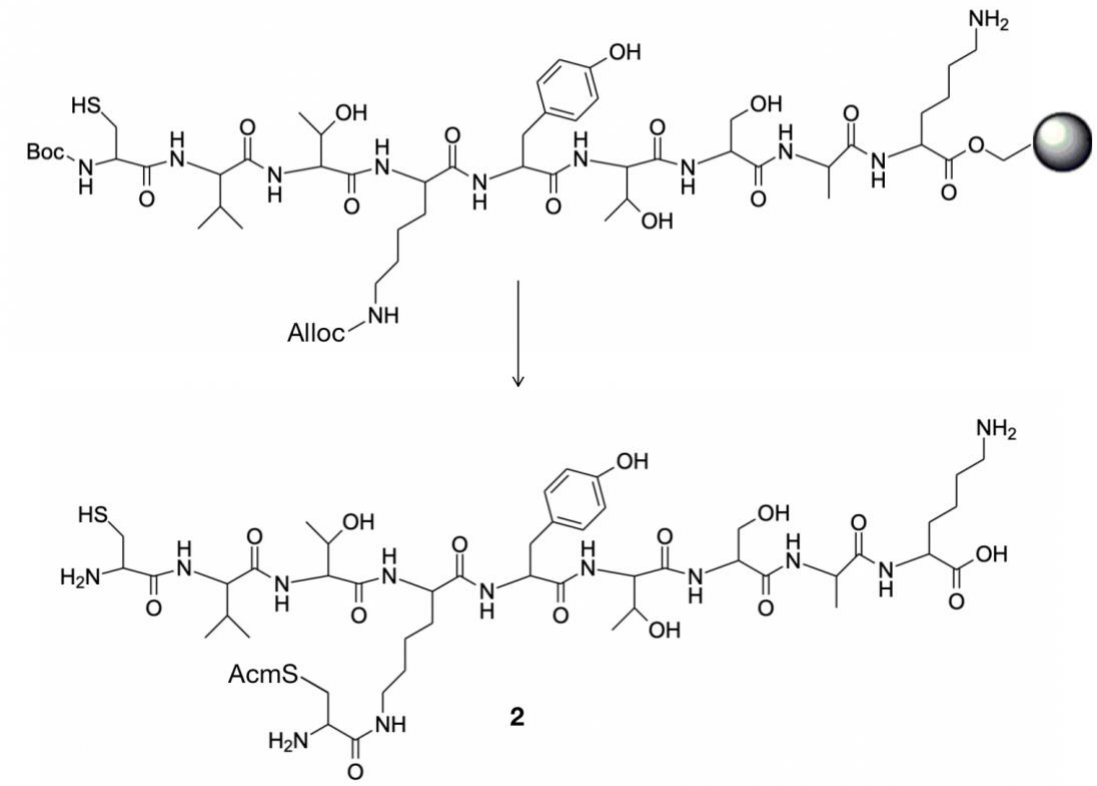

Figure 2.9. **Peptide synthesis of H2B(117-125, A117C).** After H(Boc)N-CVTKYTSAK-Wang resin was synthesized, the Alloc protecting group at K120 was removed using $Pd(PPh_3)_4$ and phenylsilane in DCM. Then Cys(Acm) was coupled at K120. The peptide was finally fully deprotected and cleaved from the resin by a cleavage cocktail containing 80% TFA, 5% phenol, 5% thioanisole, 5% H_2O, 2.5% EDT, and 2.5% TIS.

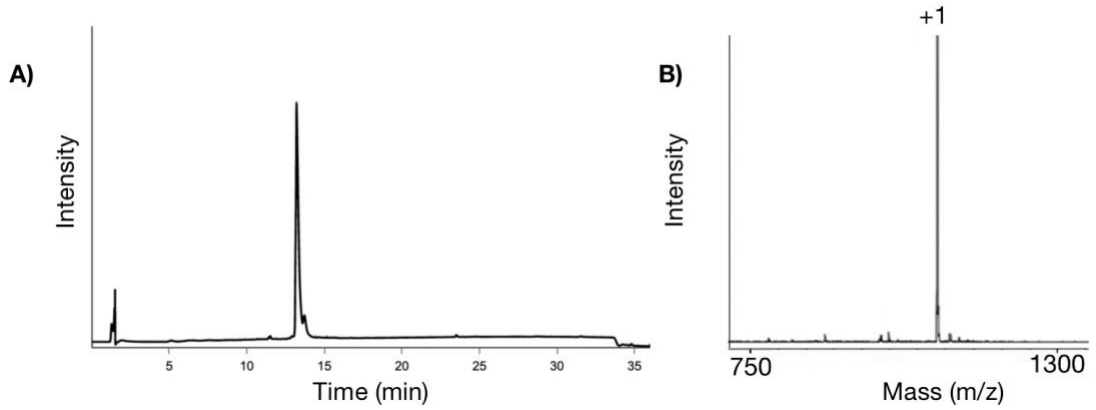

Figure 2.10. **Purification of intermediate H(Boc)N-CVTKYTSAK-C(O)OH.** (A) C18 analytical RP-HPLC chromatogram of purified **H(Boc)N-CVTKYTSAK-C(O)OH** with the gradient of 0-73% B, 30 min; (B) ESI-MS of purified **H(Boc)N-CVTKYTSAK-C(O)OH**. K = K(ε-Alloc).

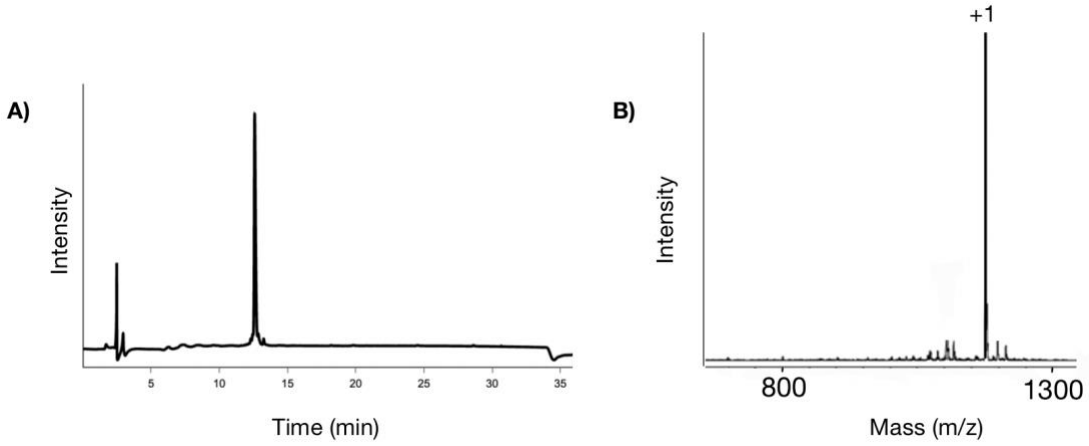

Figure 2.11. **Purification of H2B(117-125) fragment (A117C, K120-C(Acm)), 2**. (A) C18 analytical RP-HPLC chromatogram of purified **2** with the gradient of 0-73% B, 30 min; (B) ESI-MS of purified **2**.

2.3 Conclusion and Outlook

The generation of H2B(1-116) and Ub(1-75) α-thioesters was mediated by their intein fusions. Some of the shortcomings of the procedures could be relieved by controlling reaction temperature and time, adjusting to thiolysis-favorable pHs, and/or the addition of excess intein-cleaving thiol reagent (MESNa). Reduction of losses in H2B(1-116)-MES yields was achieved by avoiding RP-

HPLC purification of the fusion protein after Ni^{2+} column. The H2B(117-125) C-terminal fragment was synthesized using Fmoc-SPPS. The high purity and homogeneity of the H2B(117-125) fragment was ensured by test cleavage to check coupling efficiency, additional C-terminal Boc-Cys-OH coupling, and the use of a more potent peptide cleavage cocktail that has additional scavenging reagents such as thioanisole and phenol. Further refining the approaches presented in this chapter for producing the components required for the H2BK120ub semisynthesis would be very useful for our and other research groups in generating homogeneously modified H2B in good yields.

2.4 Experimental Procedures

2.4.1 *General methods*

Fmoc-Lys(Boc)-Wang resin, amino acid derivatives, and peptide coupling reagents were purchased from Novabiochem (San Diego, CA). *E. coli* BL21(DE3) competent cells were purchased from Novagen (Madison, WI). DNA sequencing was performed by Eurofins (Louisville, KY). Plasmid mini-prep kit was purchased from Qiagen (Valencia, CA). SpectraPor dialysis membrane was purchased from Spectrum Labs (Rancho Dominguez, CA). All other chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA). Solid phase peptide synthesis (SPPS) was performed on a Liberty Blue Automated Microwave Peptide Synthesizer (CEM Corporation, Matthews, NC). Analytical reversed-phase HPLC (RP-HPLC) was performed on a Varian (Palo Alto, CA) ProStar HPLC with a Grace-Vydac C4 or C18 column (150 x 4.6 mm) at 1 mL/min using 0.1% TFA in water (A) and 90% acetonitrile, 0.1% TFA in water (B) as the mobile phases. Preparative RP-HPLC was performed using a Grace-Vydac (Deerfield, IL) C18 column (250 x 22 mm) at 9 mL/min. Typical analytical gradients were 0-73%

B over 30 min. Mass spectra were obtained on a Bruker Esquire ESI-MS instrument (Billerica, MA).

2.4.2 *Construction of the H2B(1-116) Overexpression Plasmid*

The plasmid pTXB1-H2B(1-116)-GyrA-His_6 [8], containing truncated the *Xenopus* H2B gene H2B(1-116), was generated by PCR amplification of H2B(1-116) using a *Xenopus* H2B expression plasmid template [9], forward primer 5′-GGAATTCCATATGCCTGAGCCAGCCAAGTCCGCTCCAGCCCCG-3′, and reverse primer 5′-GGTGGTTGCTCTTCCGCACTTGGTGCCCTCGGACAC-3′. Followed by *NdeI* and *SapI* digestion and similarly digested pTXB1 vector ligation, the plasmid consists of a C-terminal GyrA intein and a chitin-binding domain (CBD) fusion to H2B(1-116). DNA sequencing was undertaken to confirm the plasmid sequence (Eurofins).

2.4.3 *Construction of Ubiquitin Overexpression Plasmid*

The plasmid pTXB1-Ub(1-75)-AvaDNAE-AAFN-His_6, containing truncated human ubiquitin gene Ub(1-75), was modified from the gift plasmid pTXB1-Ub(1-76)-AvaDNAE-AAFN-His_6 from the Muir lab at Princeton University, by varying a Glycine at the C-terminal [10]. A site-directed mutagenesis (QuikChange kit, Agilent Technologies, Santa Clara, CA) template was used for preparing the plasmid with forward primer 5’-CTGCACCTGGTACTCCGTCTCAGAGGTTGCCTGAGCTATGATACCGAAGTGCTG-3’ and reverse primer 5’- CAGCACTTCGGTATCATAGCTCAGGCAACCTCTGAGACGGAG TACCAGGTGCAG-3’ [11]. The T7 forward primer was used to verify the desired plasmid sequence (Eurofins).

2.4.4 *Overexpression and Purification of H2B(1-116)-MES*

The protocol was revised from McGinty et. al [8]. pTXB1-H2B(1-116)-GyrA-His_6 was transformed into *E. Coli* BL21(DE3) cells. Cells were grown in 6 L of LB media supplemented with 100 µg/mL of Ampicillin with shaking at 240 rpm at 37 ºC until the OD_{600} reached 0.4-0.6. At this point, overexpression was induced by 0.35 mM of IPTG addition and cells grown at 25 ºC for an additional 3.5 h. The cells were harvested by centrifugation at 7,000x*g* for 15 min. The cell pellet was resuspended in twice the dry pellet weight of 1X PBS, pH 7.5, lysed by pulse-sonication on ice for 12 min, and centrifuged at 20,000x*g* for 20 min at 4 ºC. The cell lysate supernatant was filtered and applied to 5 mL of Ni-NTA column pre-equilibrated with 1X PBS, pH 7.5. The cell lysate supernatant on the column was nutated at 4 ºC for 1.5 h to allow time for Ni^{2+} binding. The insoluble pellet was washed with 40 mL of 50 mM Tris, 200 mM NaCl, 1% triton X-100, pH 7.5 at room temperature, and centrifuged at 20,000x*g* for 15 min at 4 ºC. The pellet was dissolved in 40 mL of 6 M urea, 10 mM Tris, 200 mM NaCl, pH 7.5 at room temperature (Dissolving Buffer), and centrifuged at 20,000x*g* for 15 min at 4 ºC. The dissolved supernatant was then applied to 5 mL of Ni-NTA column pre-equilibrated with Dissolving Buffer. The dissolved pellet supernatant column was nutated at room temperature for 1 h. The columns were then washed in buffer containing increasing amounts of imidazole at 4 ºC: cell lysate supernatant column with 1X PBS, pH 7.5, 20 mM imidazole (4 X 5 CV, CV = column volume), 50 mM (4 X 5 CV), 100 mM (2 X 5CV), 200 mM (6 X 5 CV), 500 mM (5 X 2 CV); and dissolved pellet supernatant column with dissolving buffer with 20 mM imidazole (4 X 5 CV), 50 mM (4 X 5 CV), 500 mM (3 X 3 CV). Using 15,000 MWCO dialysis tubings, 500 mM imidazole elution fractions from the cell lysate supernatant column and dissolved pellet supernatant column were separately combined and dialyzed twice at 4 ºC for 1.5 h against 2 L of 1X PBS, 1 mM MESNa, 1 mM EDTA, pH 7.5 (for

cell lysate supernatant column), or 1 L of 10 mM Tris, 200 mM NaCl, 2 M Urea, 1 mM MESNa, pH 7.5 (for dissolved pellet supernatant column). Thiolysis reactions were performed by adding MESNa powder to each dialyzed protein to a final concentration of 100 mM after dialysis, re-adjusted pH to 7.5, and nutated for 16 h at 4 °C. Both reactions were dialyzed against 2 L of 4 M Gn·HCl, 50 mM Na_2HPO_4, pH 7.5, at 4 °C for 3 h using 3,500 MWCO dialysis tubing, then applied to a 5 mL Ni-NTA column pre-equilibrated with the same buffer. The columns were nutated at 4 °C for 30 min, then H2B(1-116)-MES α-thioester was eluted from the columns, followed by washing twice with 1 CV of the same buffer. The flow-through and washes from each column were dialyzed against 4 L of Millipore purified water at 4 °C for 12 h and then again for 3 h using 3,500 MWCO dialysis tubing. The crude H2B(1-116)-MES was flash-frozen and lyophilized to reduce liquid volume. A portion of the crude product was purified by C4 analytical RP-HPLC using a gradient of 0-73% B, 30 min. Fractions containing H2B(1-116)-MES were identified using ESI-MS. The typical yield of lyophilized H2B(1-116)-MES after dialysis was ~1 mg/L of LB growth media. Calculated *m/z* $[M+H]^+$: 12,933.8 Da; observed 12931.2 ± 0.5 Da.

2.4.5 *Overexpression and Purification of Ub(1-75)-MES*

The protocol was revised from Weller et. al [11]. pTXB1-Ub(1-75)-AvaDNAE-AAFN-His_6 was transformed into *E. Coli* BL21(DE3) cells. Cells were grown in 3 L of LB media supplemented with 100 µg/mL of Ampicillin with shaking at 240 rpm at 37 °C until the OD_{600} reached 0.4-0.6. Overexpression was induced by the addition of 0.75 mM IPTG and grown at 30 °C for a further 4 h. The cells were harvested by centrifugation at 7,000x*g* for 15 min. The cell pellet was resuspended in 2X weight of 50 mM Na_2HPO_4, 300 mM NaCl, 5 mM Imidazole, pH 8 (Lysis Buffer), lysed by pulse-sonication on ice for 15 min, and centrifuged at 20,000x*g* for 15 min at 4 °C. The cell lysate supernatant was filtered through a 0.45 micron filter and applied to 6 mL of Ni-

NTA column pre-equilibrated with lysis buffer. The cell lysate supernatant on column was nutated at 4 °C for 1 h. The column was then washed with Lysis Buffer containing increasing amounts of imidazole under 4 °C: 5 mM (5 CV), 20 mM (5 CV), 50 mM (2.5 CV), 250 mM (5 CV). Pooled fractions containing Ub(1-75)-Ava-His_6 were dialyzed against 2 L of 100 mM Na_2HPO_4, 150 mM NaCl, 1 mM EDTA, 1 mM MESNa, pH 7.2 for 1 h at 4 °C twice using 3,500 MWCO dialysis tubing. The thiolysis reaction was performed by directly adding MESNa powder to the dialyzed column eluate fractions to a final concentration of 100 mM and nutation for 18 h at 30 °C. Crude Ub(1-75)-MES was purified by C18 preparative RP-HPLC using a gradient of 30-60% B, 60 min. Fractions containing Ub(1-75)-MES were identified using ESI-MS. The typical yield of pure lyophilized Ub(1-75)-MES was ~15 mg/L of LB media. Calculated *m/z* $[M+H]^+$: 8632.8 Da; observed 8631.6 ± 0.7 Da.

2.4.6 *Synthesis of H2B(117-125) Fragment*

The peptide H(Boc)N-CVTKYTSAK-C(O)OH corresponding to the final nine residues of the H2B C-terminus with an A117C mutation was synthesized using SPPS based on an N^{α}-Fmoc protecting group strategy. Each amino acid (5.0 equivalents) was coupled to Fmoc-Lys(Boc)-Wang resin (0.4 mmol/g) using a scale of 0.1 mmol. The resin was pre-swollen in ~2 mL of 50/50 (v/v) DMF/DCM for 20 min, and the amino acids were deprotected then single-coupled sequentially at 75 °C, except for Ala being double coupled. Fmoc- deprotection was undertaken for 3 min using 4 mL of 5% (w/v) piperazine and 0.05 M HOBt to suppress racemization during peptide synthesis. Single coupling reactions were performed for 5 min using Fmoc-amino acids (0.5 mmol), and activator (0.5 mmol DIC) and coupling additive (0.5 mmol Oxyma Pure and 0.05 mmol DIEA) in DMF. Fmoc-Ala was coupled twice, each time for 10 min, and Boc-Cys was coupled for 10 min. Final deprotection was undertaken at 90 °C for 65 s. K120 was protected by Alloc group on the side-

chain for subsequent Fmoc-Cys(Acm) coupling. The peptidyl resin was thoroughly washed with DMF, followed by washing 4-5 times with DCM. Then the peptidyl resin was nutated at room temperature for 1 h with 0.4 equivalent of $Pd(PPh_3)_4$ and 20 equivalents of phenylsilane (relative to initial resin loading) in 1-2 mL of DCM to remove the Alloc group. The resin was washed with DCM then DMF before transferring back to the synthesizer. Fmoc-Cys(Acm) was coupled at K120 twice, each time at 75 °C for 10 min, followed by a final deprotection. The peptidyl resin was rinsed with DMF then DCM before drying and lyophilizing. The peptidyl resin was nutated at room temperature for 3 h in 10 μL of cleavage cocktail (80% TFA, 5% phenol, 5% thioanisole, 5% H_2O, 2.5% EDT, and 2.5% TIS) per mg of peptidyl resin for peptide cleavage and deprotection. Then the peptide was precipitated using a more than 9-fold volume of ice-cold diethyl ether relative to the cleavage cocktail. The resin was further rinsed with a small amount of TFA. The precipitated peptide was washed twice with fresh ice-cold ether by decanting the ether after centrifugation at 20,000x*g* at 4 °C for 5 min. The dry peptide was purified by C18 Preparative RP-HPLC using a gradient of 10-50% B, 60 min. Fractions containing the pure H2B(117-125) fragment (A117C, K120-C(Acm)) were identified using ESI-MS (set dry gas to 7-8, and trap drive to 75). Calculated for the H2B(117-125) (A117C, K120-Alloc) intermediate *m/z* $[M+H]^+$: 1,085.5 Da; observed 1,084.6 ± 0.1 Da. Calculated for the final peptide H2B(117-125) (A117C, K120-C(Acm)) *m/z* $[M+H]^+$: 1,175.6 Da; observed 1,174.6 ± 0.2 Da.

2.5 REFERENCES

[1] Gomes, L.; Monteiro, G.; Mergulhão, F. The Impact of IPTG Induction on Plasmid Stability and Heterologous Protein Expression by Escherichia Coli Biofilms. *Int. J. Mol. Sci.* **2020**, *21* (2), 576.

[2] Kataoka, K.; Takasu, A. IPTG-Independent Autoinduction of Extracellular Matrix Proteins Using Recombinant E. Coli as the Expression Host. *Polym. J.* **2021**, *53* (2), 385–391.

[3] Pavankumar, T. L. Inteins: Localized Distribution, Gene Regulation, and Protein Engineering for Biological Applications. *Microorganisms* **2018**, *6* (1).

[4] Erny, M.; Lundqvist, M.; Rasmussen, J. H.; Ludemann-Hombourger, O.; Bihel, F.; Pawlas, J. Minimizing HCN in DIC/Oxyma-Mediated Amide Bond-Forming Reactions. *Org. Process Res. Dev.* **2020**, *24* (7), 1341–1349.

[5] Peptide Chemistry Portal. Oxima Pure/DIC peptide coupling – https://peptidechemistryportal.com/oxima-pure-dic-peptide-coupling/.

[6] Luna, O. F.; Gomez, J.; Cárdenas, C.; Albericio, F.; Marshall, S. H.; Guzmán, F. Deprotection Reagents in Fmoc Solid Phase Peptide Synthesis: Moving Away from Piperidine? *Molecules* **2016**, *21* (11), 1542.

[7] Wilson, K. R.; Sedberry, S.; Pescatore, R.; Vinton, D.; Love, B.; Ballard, S.; Wham, B. C.; Hutchison, S. K.; Williamson, E. J. Microwave-Assisted Cleavage of Alloc and Allyl Ester Protecting Groups in Solid Phase Peptide Synthesis: MW-Assisted Cleavage of Alloc and Allyl Ester Protecting Groups in SPPS. *J. Pept. Sci.* **2016**, *22* (10), 622–627.

[8] McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W. Chemically Ubiquitylated Histone H2B Stimulates HDot1L-Mediated Intranucleosomal Methylation. *Nature* **2008**, *453* (7196), 812–816.

[9] Luger, K.; Rechsteiner, T. J.; Flaus, A. J.; Waye, M. M.; Richmond, T. J. Characterization of Nucleosome Core Particles Containing Histone Proteins Made in Bacteria. *J. Mol. Biol.* **1997**, *272* (3), 301–311.

[10] Shah, N. H.; Dann, G. P.; Vila-Perelló, M.; Liu, Z.; Muir, T. W. Ultrafast Protein Splicing Is Common among Cyanobacterial Split Inteins: Implications for Protein Engineering. *J. Am. Chem. Soc.* **2012**, *134* (28), 11338–11341.

[11] Weller, C. E.; Dhall, A.; Ding, F.; Linares, E.; Whedon, S. D.; Senger, N. A.; Tyson, E. L.; Bagert, J. D.; Li, X.; Augusto, O.; Chatterjee, C. Aromatic Thiol-Mediated Cleavage of N-O Bonds Enables Chemical Ubiquitylation of Folded Proteins. *Nat. Commun.* **2016**, *7* (1), 12979.

Chapter 3. SEMISYNTHESIS OF H2BK120UB

3.1 INTRODUCTION

Protein semisynthesis by Native Chemical Ligation (NCL) was first reported in 1994 by Dawson et al [1]. In NCL, a native peptide bond between two fragments is formed by the presence of a C-terminal α-thioester on one fragment and an N-terminal Cys residue on the other fragment. These are the only pre-requisites for the reaction. In the course of the reaction, a reversible transthioesterification is undertaken by the N-terminal Cys residue of one fragement with the α-thioester of the other fragment to form a thioester-linked full-length protein intermediate. This is followed by a sponataneous intramolecular *S,N*-acyl shift to form the native amide bond between the two fragments (Figure 3.1). The C-terminal α-thioesters could be produced via intein thiolysis as I have demonstrated, or by SPPS. The N-terminal Cys containing fragment can be produced by SPPS or by expression in *E. coli*. The detailed protocols for generating the H2B(1-116)-MES α-thioester and H2B(117-125) (A117C, K120-C(Acm)) are provided in Chapter 2. In this chapter, H2B(1-116)-MES and the synthetic H2B(117-125) fragment were first ligated to result in a full-length H2B. After removing the Acm protective group at K120-Cys(Acm), Ub(1-75)-MES was then ligated to the full-length H2B to generate H2BK120ub.

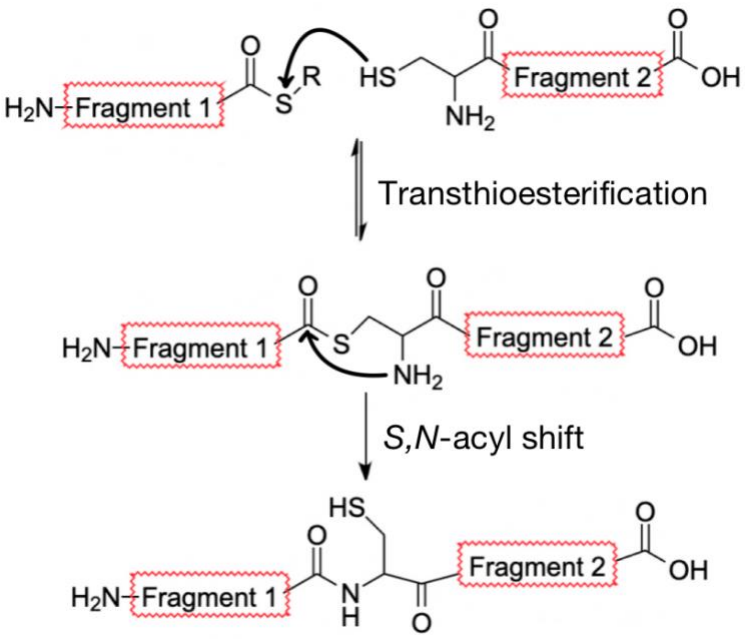

Figure 3.1. **Schematic of native chemical ligation (NCL).** The N-terminal Cys residue on Fragment 2 attacks the C-terminal α-thioester in Fragment 1 to form a thioester intermediate, followed by an *S*,*N*-acyl shift to form the native amide bond between the two fragments.

3.2 Results and Discussion

3.2.1 *H2BK120ub Production*

Since H2B(1-116)-MES **1** was not further purified due to losses during HPLC, its purity was estimated by comparisons on 15% SDS-PAGE with known concentrations of Myoglobin as a reference protein of similar size (Figure 3.2A, myoglobin 17,000 g/mol and **1** 12,933.8 g/mol). Once the gel was fully destained, the area under curve (AUC) of each band on the gel could be obtained in ImageJ. The linear regression of AUC vs. [Myoglobin] could then be fitted (Figure 3.2B), and be used to estimate the purity of crude **1**, as well the total quantity of **1** in any batch of H2B(1-116)-MES α-thioester. However, very precise pipetting and the appropriate concentrations of **1** for direct comparison with Myoglobin were required to obtain a good estimate of

concentration. For the future, a lower initial concentration of **1** could be used (such as 5.5 mg/mL) during SDS-PAGE comparisons in order to avoid band intensity saturation. Since the yield of purified **2** and **4** were very high, generating **3** using excess amounts of **1** and **2** and checking its yield could be another approach to estimate the quantity of pure **1** in a mixture containing the thioester and free acid forms. The total chemical synthesis of H2BK120ub (**5**) is challenging because of the ~50-residue upper limit of conventional SPPS [1]. Hence, we pursued native chemical ligation (NCL) for the production of **5** (Figure 3.3). The native peptide bond between unprotected **1** and **2** was formed by the presence of a C-terminal α-thioester in **1** and an N-terminal Cys residue in peptide **2**. Additional MESNa was added as a thiol catalyst for transthioesterification, and TCEP was used as a disulfide reductant. Additionally, fresh TCEP was also required prior to all C4 RP-HPLC purifications during the H2BK120ub semisynthesis process to avoid MESNa thiol-adduct formation [2]. The RP-HPLC chromatogram and ESI-MS spectrum of pure **3** are shown in Figure 3.4. Radical-based desulfurization [3] was then performed using VA-044 as the radical initiator, TCEP, and *t*BuSH as a hydrogen source to convert the unprotected N-terminal Cys residue in **3** into the wild-type Ala found in H2B. The RP-HPLC chromatogram and ESI-MS spectrum of pure desulfurized H2B(1-125) are shown in Figure 3.5. Palladium-mediated Cys-Acm deprotection [4] to unmask the Cys ligation site at K120 was performed using $PdCl_2$ and the reaction quenched by DTT before the final ligation with **4**. During this step, all the excess palladium should be precipitated otherwise later reactions would be inhibited. Full precipitation could be achieved by adding additional DTT, dropping the pH to ~6, having longer reaction time and/or centrifuging time, and flash-freezing the reaction, keeping it frozen for ~5 min, and thawing slowly. After removing the precipitated Pd by centrifugation and dialysis to remove excess DTT, the final ligation of the deprotected peptide intermediate with **4** was

undertaken. 2-mercaptophenylacetic acid (MPAA) was used as a thiol catalyst for transthioesterification and TCEP as the disulfide reductant, to yield **5**. The RP-HPLC chromatogram and ESI-MS spectrum of pure **5** are shown in Figure 3.6A and B.

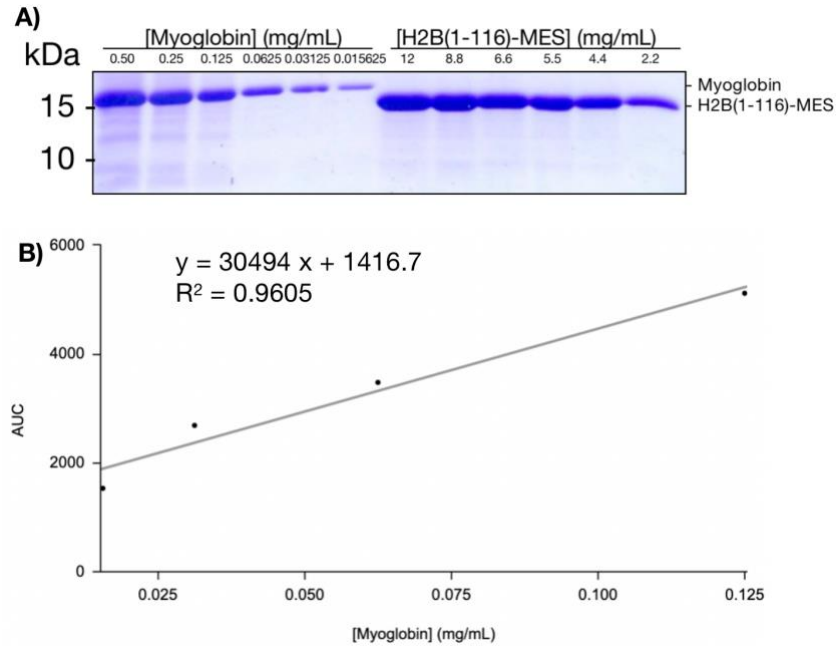

Figure 3.2. **Crude H2B(1-116)-MES 1 quantification using myoglobin reference.** (A) Coomassie-stained 15% SDS-PAGE of serially-diluted myoglobin standards (left) and crude **1** (right); (B) Linear regression of area under curve (AUC) vs. [Myoglobin] (mg/mL) for estimating the purity of crude **1**.

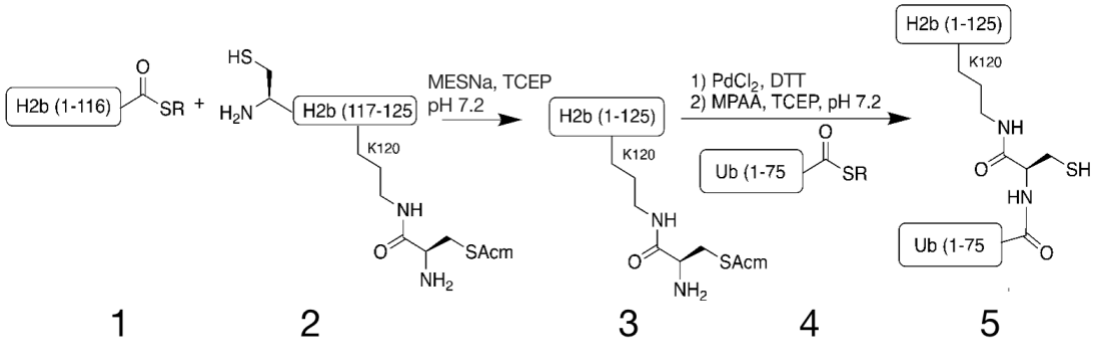

Figure 3.3. **Semisynthesis of ubiquitinated H2B.** All reactions were performed using native chemical ligation (NCL) buffer (6 M Gn·HCl, 200 mM Na_2HPO_4, pH 7.3). H2B(1-116)-MES, **1**, was ligated to the synthetic H2B(117-125) fragment, **2**, with MESNa and TCEP additives at pH 7.2 to yield H2B(1-125), **3**. Desulfurization of the Cys was undertaken using TCEP, VA-044, and

t-BuSH, followed by the removal of Acetamidomethyl (Acm) with $PdCl_2$ and DTT. The intermediate was then ligated to Ub(1-75)-MES, **4**, with MPAA and TCEP additives at pH 7.2 to yield native H2BK120ub, **5**.

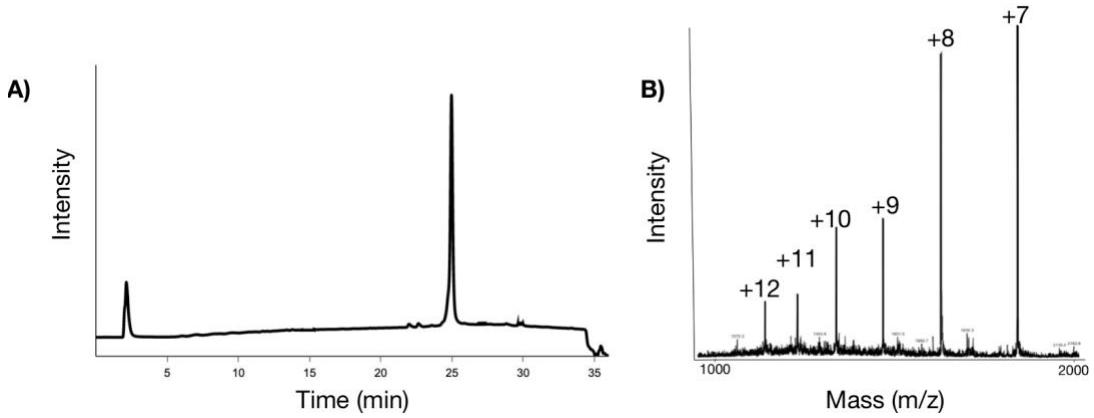

Figure 3.4. **RP-HPLC Purification of H2B(1-125) 3.** (A) C4 analytical RP-HPLC chromatogram of purified **3** with a gradient of 35-65% B at 45 °C, 30 min; (B) ESI-MS of purified **3**.

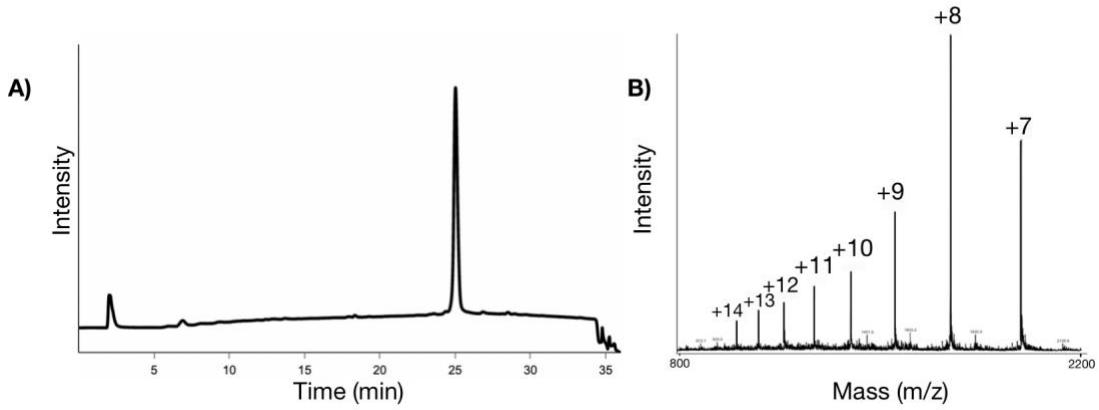

Figure 3.5. **RP-HPLC Purification of Desulfurized H2B(1-125).** (A) C4 analytical RP-HPLC chromatogram of purified desulfurized H2B(1-125) with a gradient of 35-65% B, 30 min; (B) ESI-MS of purified desulfurized H2B(1-125).

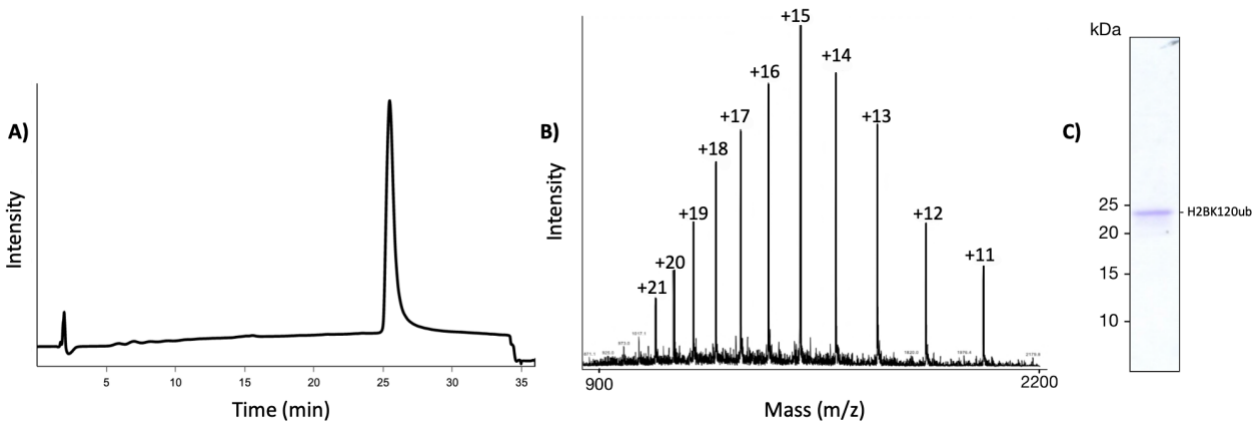

Figure 3.6. **Purification of H2BK120ub 5.** (A) C4 analytical RP-HPLC chromatogram of purified desulfurized H2B(1-125) with the gradient of 35-65% B at 45 ºC, 30 min; (B) ESI-MS of purified **5**; (C) Coomassie-stained 15% SDS-PAGE of purified **5**.

3.2.2 *Octamer and Mononucleosome Assembly*

After the purity of H2BK120ub **5** was verified (Figure 3.6C), the histone octamer containing wild-type (wt) H2A, **5**, H3(C110A), and H4 could be assembled by combining equimolar amounts of the four core histones in a guanidine unfolding buffer, then dialyzing against high-salt refolding buffer in the presence of DTT as a reducing agent. The final octamer was purified by FPLC using size-exclusion chromatography (Superdex S200), and the formation of the octamer was verified by 15% SDS-PAGE (Figure 3.7). The calculated concentration of the octamer was 3.922 µM, which was used for mononucleosome reconstitution by serial dilution with double-stranded DNA. A series of small-scale mononucleosome assemblies were performed by dilution transfer with the H2Bub octamer and a 147 bp Widom 601 nucleosome positioning sequence (concentration 5.98 µM) using protease inhibitor (PMSF) containing buffers [5]. An octamer/DNA molar ratio of 2.6 was verified to be the optimal saturation ratio to obtain high quality mononucleosomes (Figure 3.8).

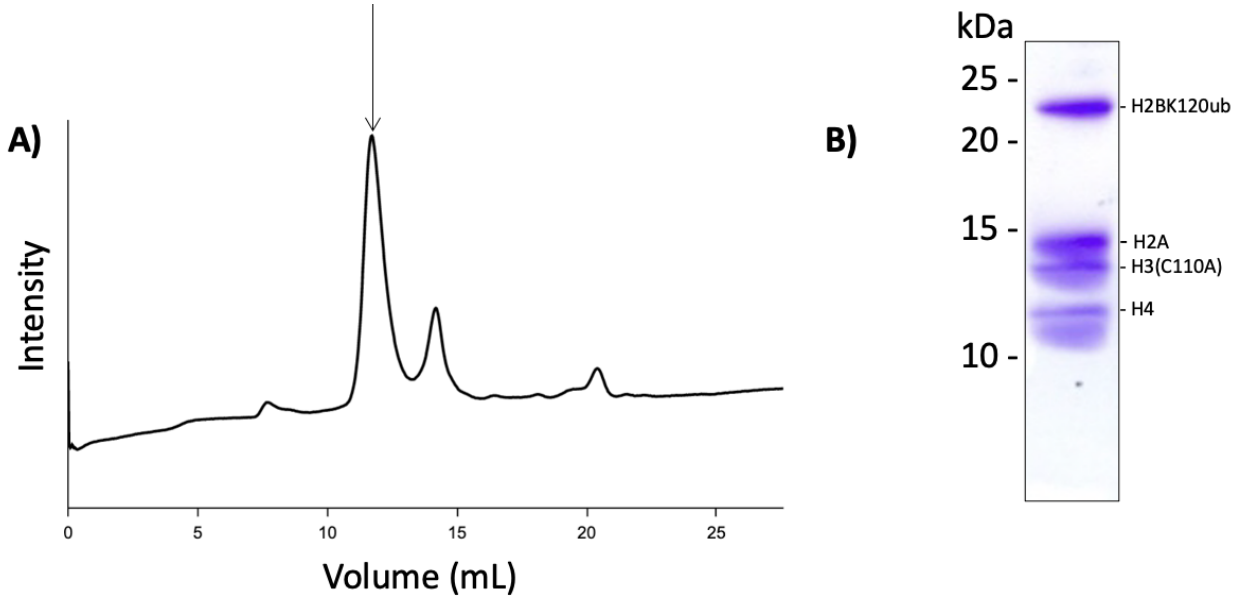

Figure 3.7. **FPLC purification of the H2BK120ub-containing octamer.** (A) Superdex 200 Increase 10/300 GL size exclusion FPLC chromatogram of the octamer, flow rate 0.4 mL/min; (B) Coomassie-stained 15% SDS-PAGE gel of octamers containing H2BK120ub.

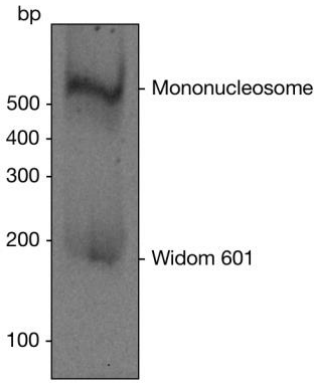

Figure 3.8. **SYBR-safe stained 5% TBE gel of H2Bub containing mononucleosomes.**

3.3 CONCLUSION AND OUTLOOK

The semisynthesis of H2BK120ub was achieved via NCL. A full-length H2B intermediate was first ligated using H2B(1-116)-MES (C-terminal α-thioester) and the synthesized H2B(117-125) fragment with an A117C mutation (N-terminal Cys residue) and Cys(Acm) attached to the epsilon-amine of K120. The desired H2BK120ub was generated after ligating the Acm-deprotected full-length H2B intermediate (side-chain Cys residue on K120) and Ub(1-75)-MES

(C-terminal α-thioester). Since RP-HPLC purification was performed on the products after each step, and equipment error occured during the initial H2BK120ub semi-preparative RP-HPLC purification resulting in some contamination of the sample, the final yield of H2BK120ub was limited to 2% based on the amount of starting H2B(1-116)-MES α-thioester and over 4 steps. Nevertheless, milligram amounts of pure H2BK120ub were produced and could be used for the formation of octamers with wt H2A, H3(C110A) and H4, and the assembly of mononucleosome using ~147 bp Widom 601. An octamer to DNA molar ratio of 2.6 to 1.0 was chosen as the optimal saturation ratio for mononucleosome formation, hence, large scale assembly of mononucleosomes could be performed with this ratio. The resulting mononucleosomes would be tested in various biochemical assays for the crosstalk among core histones in the future. For instance, hDot1L assay for H3K79 methylation which is associated with mixed lineage leukemia (MLL) [6, 7] and hSet1 assay for H3K4 methylation which is responsible for colon cancer [8]. Because both hDot1L and hSet1 require ubiquitylated mononucleosome substrates for activation, the successful synthesis of H2Bub is critical for future biochemical assays in our lab.

3.4 EXPERIMENTAL PROCEDURES

3.4.1 *General methods*

Slide-A-Lyzer dialysis cassettes were purchased from Pierce (Rockford, IL). All chemical reagents were purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Pittsburgh, PA). Analytical reversed-phase HPLC (RP-HPLC) was performed on a Varian (Palo Alto, CA) ProStar HPLC with a Grace-Vydac C4 (150 x 4.6 mm) at 1 mL/min using 0.1% TFA in water (A) and 90% acetonitrile, 0.1% TFA in water (B) as the mobile phases. Semi-preparative RP-HPLC was performed using a Grace-Vydac (Deerfield, IL) C4 column (250 x 10 mm) at 3.5 mL/min. Typical

analytical gradients were 0-73% B over 30 min. Mass spectra were generated on a Bruker Esquire ESI-MS instrument (Billerica, MA). Size-exclusion chromatography was performed on an AKTA FPLC system (GE Healthcare, Little Chalfont, UK) equipped with a P-920 pump and UPC-900 monitor using a Superdex 200 Increase 10/300 GL size exclusion column (24 mL volume). Mononucleosome gels were visualized using a GE Typhoon FLA 9000 Biomolecular Imager (GE).

3.4.2 *H2BK120ub Production*

As shown in Table 3.1, MESNa and TECP-HCl were dissolved in a calculated volume of NCL buffer (final protein concentration of 0.5-1.0 mM), and the pH was re-adjusted to 7.2. **1** and **2** were then dissolved in this reaction mixture. The reaction was re-adjusted to pH 7.2 and incubated overnight (16-18 h). The completed NCL reaction product was purified by C4 Semi-preparative RP-HPLC using a gradient of 35-65% B at 45 ºC for 45 min. Fractions containing H2B(1-125) were identified using ESI-MS. Calculated *m/z* $[M+H]^+$: 13,965 Da; observed 13,964.5 ± 1.5 Da.

As shown in Table 3.2, **3** was dissolved in calculated volume of NCL buffer with 71 mg/mL of TCEP-HCl (final protein concentration 0.5-1.0 mM), then VA-044 was dissolved into the reaction mixture and incubated at 37 ºC for 15 min, followed by *t*BuSH addition and incubation at 37 ºC for 3-4 h (to a maximum of 18 h). The completed reaction was purified by C4 Semi-preparative RP-HPLC using a gradient of 35-65% B, 45 min. Fractions containing Desulfurized H2B(1-125) were identified using ESI-MS. Calculated *m/z* $[M+H]^+$: 13,933 Da; observed 13,932.3 ± 1.3 Da.

As shown in Table 3.3, desulfurized H2B(1-125) was dissolved in a calculated volume of NCL buffer with $PdCl_2$, and the reaction mixture was incubated at 37 ºC for 2 h. DTT was dissolved into the reaction and the precipitated Pd was removed by centrifugation at 20,000x*g*. The reaction

supernatant was dialyzed against 500 mL of the NCL buffer at room temperature overnight (16-18 h) using a 3,500 MWCO dialysis cassette. MPAA and TCEP-HCl were dissolved in a minimum volume of NCL buffer, and the mixture was re-adjusted to pH 7.2. **4** and the dialyzed reaction were then mixed into the reaction followed by pH re-adjusting to 7.2 (final protein concentration 0.5-1.0 mM), and incubation at 30 ºC overnight (16-18 h). The completed reaction was purified by C4 Analytical RP-HPLC using gradient of 35-65% B at 45 ºC, 30 min. Fractions containing native H2BK120ub were identified using ESI-MS. Calculated *m/z* $[M+H]^+$: 22,350.6 Da; observed 22,352 ± 1.7 Da.

Table 3.1. Ligation of H2B(1-116)-MES **1** and synthesized H2B(117-125) Fragment **2**.

Reagent	MW (g/mol)	Eq
H2B(1-116)-MES **1**	12933.8	1
H2B(117-125) Fragment **2**	1174.7	3
MESNa	142.2	100
TCEP HCl	286.65	30

Table 3.2. Desulfurization of H2B(1-125) **3**.

Reagent	MW (g/mol)	Eq
H2B(1-125) **3**	13965	1
TCEP HCl	286.65	71 mg/mL
VA-044	323.33	50
*t*BuSH	-	10% v/v

Table 3.3. Acm Removal of **3** and the ligation of Ub(1-75)-MES **4**.

Reagent	MW (g/mol)	Eq
Desulfurized H2B(1-125)	13933	1
$PdCl_2$	177.33	20
DTT	154.24	100
MPAA	168.21	100
TCEP HCl	286.65	50
Ub(1-75)-MES **4**	8631.8	3

3.4.3 *Histone Assembly*

Each histone was dissolved in 7 M Gn·HCl, 20 mM Tris, pH 7.5 at room temperature (unfolding buffer) separately to a concentration of 4 mg/mL on ice. Three $A_{280\,nm}$ measurements were obtained to calculate histone concentrations, and the histones were combined in equimolar amounts (Table 3.4A) then diluted into 1 mg/mL (Table 3.4B). The reaction was then dialyzed against 1 L of freshly made 2 M NaCl, 10 mM Tris, 1 mM EDTA, pH 7.5 at 4 °C (refolding buffer) first for 3 h, then overnight (12-16 h), and finally for 2 h with 2 mM DTT addition (minimum dialysis time of the first and the last was 2 h) using a 3,500 MWCO dialysis cassette. The dialyzed octamers were then purified by Superdex 200 Increase 10/300 GL size-exclusion FPLC, flow rate 0.4 mL/min using refolding buffer with 2 mM DTT as mobile phase. Pooled FPLC fractions containing the octamers were concentrated using a 5,000 MWCO PES Vivaspin 500 to a final volume of ~150 µL. Three A_{280nm} measurements were obtained to calculate the concentration of the octamer (Table 3.4C), followed by 10% v/v glycerol addition to the concentrated octamers and flash-freezing for storage at -80 °C.

Table 3.4. Template of octamer formation calculation.

(A) Calculation of histone concentration

Histone	MW (g/mol)	ε (M^{-1} cm^{-1})	A280 Avg.	Dilution Factor	Conc. (M)	Conc. (µM)	mg/mL
H2A	13960.2	4470	0.164	5	1.838E-04	183.818	2.566
H2BK120ub	22350.6	8940	0.108	5	6.022E-05	60.216	1.346
H3(C110A)	15224.7	4470	0.182	5	2.032E-04	203.207	3.094
H4	11236.0	5960	0.205	5	1.717E-04	171.700	1.929
Octamer	125543.0	47680					

(B) Calculation of octamer assembly

Histone	mg	Target µmol	mL of Solution	
H2A	0.084	0.006	0.03264	
H2BK120ub	0.134	0.006	0.09964	
H3(C110A)	0.091	0.006	0.02953	
H4	0.067	0.006	0.03494	
Total	0.377	0.024	0.19675	
Total Volume (mL)	0.377	Buffer Addition (mL)	0.17988	For ~1mg/mL Solution

(C) Calculation of concentrated octamer

A280 Avg.	Conc. (M)	Conc. (µM)	mg/mL
0.187	3.922E-06	3.922	0.492

A280 Avg. = average of three A280 measurements; Dilution Factor = $Volume_f$ / $Volume_i$; conc. (M) = (A280 Avg. x Dilution Factor) / ε *l*; mg/mL = conc. (M) x MW (g/mol); mg = MW (g/mol) x Target µmol / 1000; mL of solution = mg / (mg/mL).

3.4.4 *Mononucleosome Assembly*

Octamer and Widom *601* double-stranded DNA were thawed on ice, and 10 µL reactions containing the octamer and DNA were gently mixed in PCR tubes at the Octamer/DNA ratios using Table 3.5A as the template to a final NaCl concentration of 2 M. The reactions were incubated at 37 ºC for 15 min, then transferred to 30 ºC to perform serial dilutions by gently mixing with 3.3 µL, 6,7 µL, 5 µL, 3.6 µL, 4.7 µL, 6.7 µL, 10 µL, 30 µL, and 20 µL of 10 mM HEPES, 1 mM EDTA, 0.5 mM PMSF, pH 7.9 at room temperature (initial dilution buffer), with an incubation interval of 15 min. The reactions were eventually diluted with 100 µL of 10 mM Tris, 1 mM EDTA, 0.5 mM PMSF, 20% v/v Glycerol, pH 7.5 at room temperature (final dilution buffer), and incubated at 30 ºC for 15 min (Table 3.5B). The mononucleosome could be stored at 4 ºC (short-term) or -80 ºC after flash-freezing (long-term).

Table 3.5. Template of mononucleosome formation calculation and reaction time log.

(A) Calculation of mononucleosome assembly

Octamer/DNA	2.60	H_2O (µL)	0.350
$[Octamer]_i$ (µM)	3.922	5 M NaCl (µL)	1.348
$[Octamer]_f$ (µM)	2.60	DNA (µL)	1.672
$[DNA]_i$ (µM)	5.98	Octamer (µL)	6.629
$[DNA]_f$ (µM)	1.00	Total Volume (µL)	10

(B) Mononucleosome assembly reaction tracking

Time (min)	Volume Added (µL)	Final Volume (µL)	[NaCl] (M)	Temperature (°C)
-15	0	10	2.00	37
0	3.3	13.3	1.50	30
15	6.7	20	1.00	30
30	5	25	0.80	30
45	3.6	28.6	0.70	30
60	4.7	33.3	0.60	30
75	6.7	40	0.50	30
90	10	50	0.40	30
105	30	80	0.25	30
120	20	100	0.20	30
135	100	200	0.10	30

$[Octamer]_f$ = (Octamer/DNA) x $[DNA]_f$; volume of Octamer (µL) = ($[Octamer]_f$ x 10 µL) / $[Octamer]_i$; volume of 5 M NaCl (µL)= (10 µL - Octamer) x 2/5 (2 M NaCl was included in the Octamer solution); volume of H_2O (µL) = (10 µg - sum of other 3 portions).

3.5 REFERENCES

[1] Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. Synthesis of Proteins by Native Chemical Ligation. *Science* **1994**, *266* (5186), 776–779.

[2] Mitchell, N. J.; Sayers, J.; Kulkarni, S. S.; Clayton, D.; Goldys, A. M.; Ripoll-Rozada, J.; Barbosa Pereira, P. J.; Chan, B.; Radom, L.; Payne, R. J. Accelerated Protein Synthesis via One-Pot Ligation-Deselenization Chemistry. *Chem* **2017**, *2* (5), 703–715.

[3] Thompson, R. E.; Liu, X.; Alonso-García, N.; Pereira, P. J. B.; Jolliffe, K. A.; Payne, R. J. Trifluoroethanethiol: An Additive for Efficient One-Pot Peptide Ligation-Desulfurization Chemistry. *J. Am. Chem. Soc.* **2014**, *136* (23), 8161–8164.

[4] Jbara, M.; Laps, S.; Morgan, M.; Kamnesky, G.; Mann, G.; Wolberger, C.; Brik, A. Palladium Prompted On-Demand Cysteine Chemistry for the Synthesis of Challenging and Uniquely Modified Proteins. *Nat. Commun.* **2018**, *9* (1).

[5] Jing, Y.; Ding, D.; Tian, G.; Kwan, K. C. J.; Liu, Z.; Ishibashi, T.; Li, X. D. Semisynthesis of Site-Specifically Succinylated Histone Reveals That Succinylation Regulates Nucleosome Unwrapping Rate and DNA Accessibility. *Nucleic Acids Res.* **2020**, *48* (17), 9538–9549.

[6] McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W. Chemically Ubiquitylated Histone H2B Stimulates HDot1L-Mediated Intranucleosomal Methylation. *Nature* **2008**, *453* (7196), 812–816.

[7] Banday, S.; Farooq, Z.; Ganai, S. A.; Altaf, M. Therapeutic Strategies against HDOT1L as a Potential Drug Target in MLL-Rearranged Leukemias. *Clin. Epigenetics* **2020**, *12* (1), 73.

[8] Dinarvand, M.; Kiani, M.; Mirzazadeh, F.; Esmaeili, A.; Mirzaie, Z.; Soleimani, M.; Dinarvand, R.; Atyabi, F. Oral Delivery of Nanoparticles Containing Anticancer SN38 and HSET1 Antisense for Dual Therapy of Colon Cancer. *Int. J. Biol. Macromol.* **2015**, *78*, 112–121.