

Convergent Chemical Synthesis of Histone H2B Protein for the Site-Specific Ubiquitination at Lys34**

Peter Siman, Subramanian Vedhanarayanan Karthikeyan, Miroslav Nikolov, Wolfgang Fischle, and Ashraf Brik*

Posttranslational modifications (PTMs) play a vital role in regulating the structure and dynamics of chromatin as well as DNA-driven cellular processes by covalently modifying the primary structures of the histone proteins (H2B, H2A, H3, H4) with various chemical entities.^[1] One such example is the modification of the histone H2B by ubiquitin (Ub) at multiple lysine residues. Of these, ubiquitination at Lys120 (H2K120Ub) is best studied. This PTM was suggested to modulate gene expression, trigger H3 methylation, and activate DNA damage response pathways.^[2] In contrast, little is known about the effects and functions of the other sites of ubiquitination in H2B, K34, K46, K108, and K116.^[3] Notably, the preparation of any of these ubiquitinated histone proteins for biochemical analyses has been a challenge in the field and has thus hampered a variety of studies aiming to understand the effect of ubiquitination on chromatin. Only recently, novel nonenzymatic approaches were introduced and enabled the preparation of homogeneously native H2K120Ub on a milligram scale for structural and functional analyses.^[4] These approaches proved to be successful for the incorporation of this single modification or any other modification within the short C-terminal synthetic peptide.

In order to investigate the role of ubiquitination at other sites and to study the interplay of multiple modifications, a new synthetic strategy is needed. For example, it has been reported recently that ubiquitination of H2B at Lys34 directly regulates H3K4 and K79 methylation through trans-tail crosstalk both in vitro and in cells.^[5] Hence, the preparation of such an analogue will enable studies aiming to understand how this modification regulates H3 methylation and potentially affects the structure of chromatin, as well as help to delineate mechanisms of K34 ubiquitination and deubiquitination. Here, we report our endeavors toward the total

chemical synthesis of the H2B protein, and the successful site-specific ubiquitination at Lys34 (H2K34Ub) for initial functional characterization.

If one considers the preparation of H2K34Ub semi-synthetically through the expression of a large C-terminal H2B fragment that bears an N-terminal Cys or through total chemical synthesis by applying a ligation approach with three fragments, the 57mer H2B(1–57), which bears the K34 modification, has to be prepared synthetically. Because H2B lacks any Cys residue, an Ala to Cys mutation for native chemical ligation (NCL) combined with desulfurization^[6] should be performed at Ala58 (Scheme 1). Even if the semisynthetic method were applicable, it would not allow the insertion of different PTMs along the sequence but only at the N-terminal fragment.^[7] Our initial studies with a three-fragment approach revealed that the synthesis of the 57mer peptide via Fmoc-SPPS was very difficult. We thus concluded that a total chemical synthesis of the H2B protein from four fragments is inevitable. Hence, we divided the H2B sequence, taking into consideration the position of Ala residues, to include H2B(1–20) labeled with an HA tag, H2B(21–57), H2B(58–96), and H2B(97–125). In such an approach, ubiquitination at K34 can be achieved by incorporating δ -mercaptolysine (K*)^[8] in the fragment H2B(21–57), thus enabling isopeptide native chemical ligation (ICL). We also envisioned a convergent chemical synthesis approach via a peptide hydrazide intermediate^[9] for the final assembly, first by ligating fragments H2B(1–20) and H2B(21–57), which bears a Ub, followed by ligation with H2B(58–125), which is prepared through ligation of fragments H2B(58–96) and H2B(97–125). Convergent protein synthesis offers several advantages over linear approaches and several promising strategies have been reported recently.^[10]

Initially, we examined the synthesis of unmodified H2B to set the ground for the preparation of H2B that is modified at any desired site. Peptides H2B(1–20), H2B(21–57), and H2B(58–96), which bear activated C-termini, were prepared using the N-acylurea method (Supporting Information).^[11] The N-terminal Ala in H2B(21–57) and H2B(97–125) were substituted with Cys to give peptides **3** and **1**, respectively, and in H2B(58–96) with a 1,3-thiazolidine-4-carboxy (Thz) moiety to give peptide **2**. To enable a convergent synthesis, the N-acylurea moiety in the H2B(21–57) fragment was substituted with hydrazine to form the stable C-terminal acyl hydrazide **3** (Scheme 1). Direct synthesis of this peptide on hydrazine-functionalized resin^[9] failed to give the desired product directly. Overall, the syntheses of all fragments were straightforward with good yields (30–50%; Supporting Information).

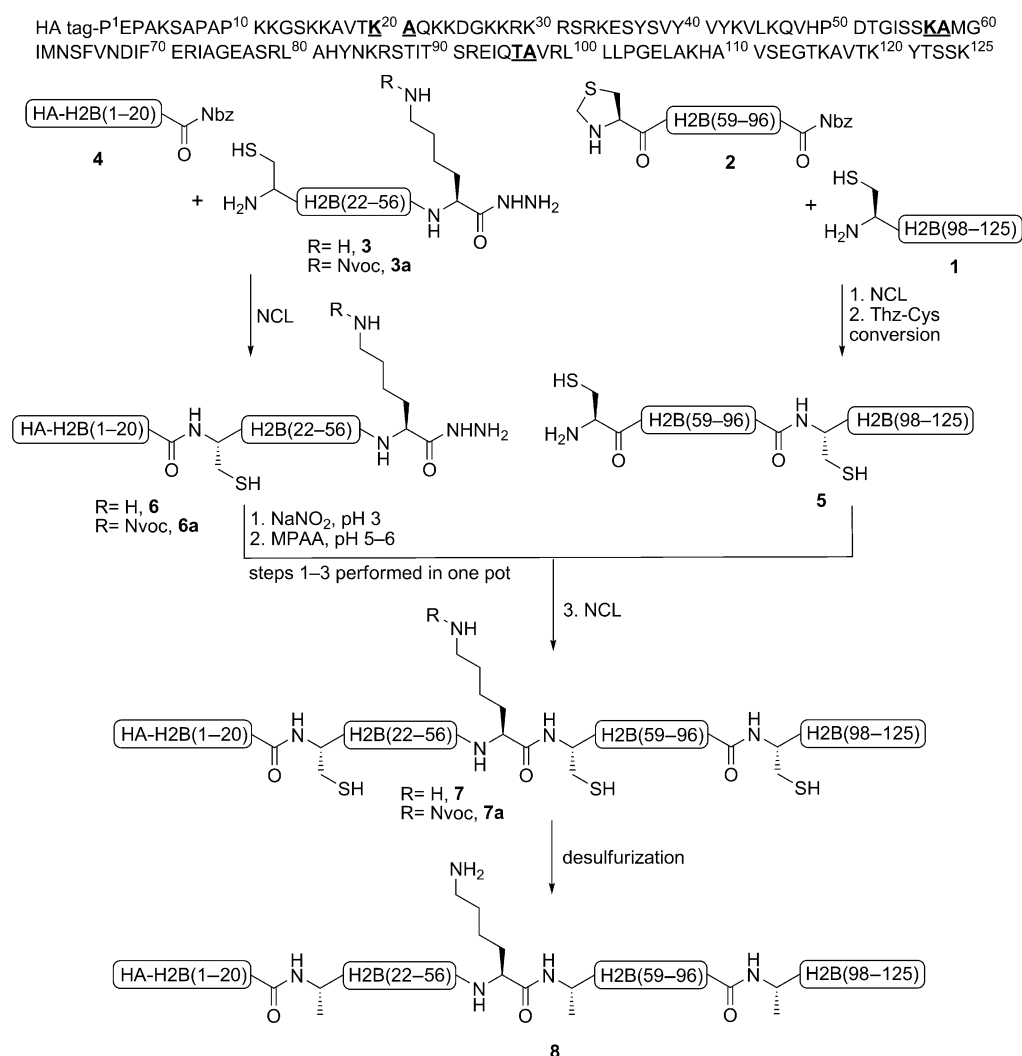
[*] P. Siman,^[†] Dr. S. V. Karthikeyan,^[†] Prof. A. Brik
Department of Chemistry
Ben Gurion University of the Negev
Beer Sheva 84105 (Israel)
E-mail: abrik@bgu.ac.il
Homepage: <http://www.bgu.ac.il/~abrik>

Dr. M. Nikolov, Dr. W. Fischle
Laboratory of Chromatin Biochemistry
Max Planck Institute for Biophysical Chemistry
Göttingen 37077 (Germany)

[†] These authors contributed equally to this work.

[**] This work was supported by the Israel Science Foundation (A.B.) and the Max Planck Society (W.F.). P.S. gratefully acknowledges the Israel Ministry of Science and Technology for financial support.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201303844>.



Scheme 1. H2B synthesis applying convergent strategy. HA tag = YPYDVPDYA.

The assembly of the peptide fragments started with the ligation of peptides **2** and **1**, followed by unmasking of the protected Cys of H2B(58–125) to furnish the desired product **5** in 40 % yield. On the other hand, the H2B(1–77) fragment **6**, which bears a C-terminal acyl hydrazide moiety, was prepared in 55 % yield through ligation of peptides **4** and **3** for 6 hours. Final ligation between fragments **6** and **5** was started by dissolving peptide **6** in GnHCl (pH 3–4), followed by oxidation with sodium nitrite.^[9] Subsequently, 4-mercapto-phenylacetic acid (MPAA)^[12] was added in several portions to furnish H2B(1–57)-MPAA. Finally, peptide **5** was added and the pH value of the solution was adjusted to around 6.5 by the addition of a solution of tris(2-carboxyethyl)phosphine hydrochloride (TCEP) in guanidine hydrochloride (Gn-HCl) to give a final peptide concentration of 1 mM. The ligation was completed after 6 hours to give the ligation product H2B(1–125) (**7**) in 25 % yield (Figure S12, Supporting Information). Peptide **7** was desulfurized using the described radical-based approach^[13] to give the final H2B protein (**8**) in 45 % yield. Notably, in the final ligation step, a significant amount of side product ($\approx 40\%$) was observed, which results from the lactamization of the C-terminus of peptide **6**. This side

product was formed during the oxidation toward the acyl azide, and the formation of the side product also progressed during thioester formation and ligation (Figure S12, Supporting Information). We reasoned that the lactamization was caused by nucleophilic attack of the ϵ -amine of the K57 side chain at the activated K57 C-terminus. Hence, we rationalized that the protection of the ϵ -amine of K57 with a 6-nitrover-atryloxycarbonyl (Nvoc) group to give **3a** should stop the lactamization in **6a** and facilitate the ligation (Scheme 1). Indeed, following this approach, the lactamization side reaction was suppressed to less than 10 %, which was due to partial deprotection of the Nvoc during the oxidation step. Interestingly, when **7a** was subjected to radical-based desulfurization conditions, we observed quantitative in situ deprotection of the Nvoc group, presumably

through a radical mechanism,^[14] to give the final product **8** (Figure 1). It should be mentioned that during ligation of peptides **3** and **4**, lactam formation was also observed, it occurred however to a lesser extent and could be suppressed significantly by adding peptide **4** to the ligation mixture in portions. Such an approach was not applicable to the final ligation, because the thioester was prepared through a hydrazide intermediate. Surprisingly, several groups have already reported ligation at Lys residues,^[15] however, such side reactions were not seen. It is plausible that the method of thioester preparation, the peptide sequence, and the ligation conditions (e.g. different thiol additives) may play a role in lactam formation. Thus, when ligating at a Lys site, extra precautions should be taken.

After the successful preparation of the unmodified H2B protein, we focused our attention on the application of the same strategy to the synthesis of H2K34Ub. The difference with this synthetic target is that an additional ligation step with Ub, which is located within peptide **3**, is required. Because of the use of thioazolidine-protected δ -mercaptolysine in this fragment, the N-terminal Cys had to be protected orthogonally with an acetamidomethyl group (Acm) to give

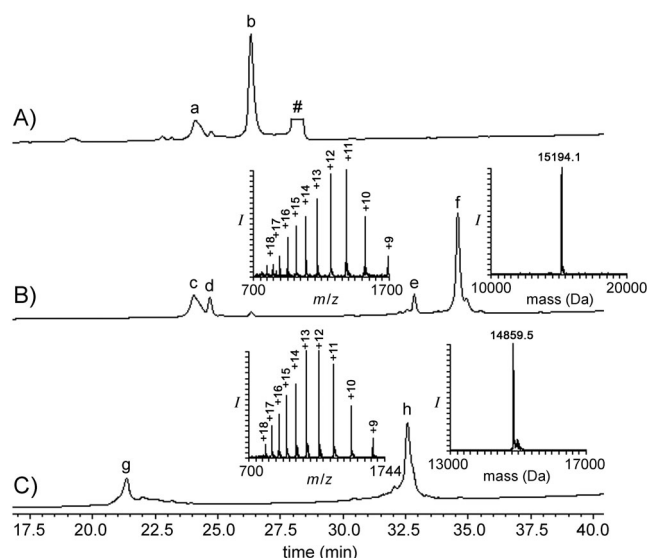
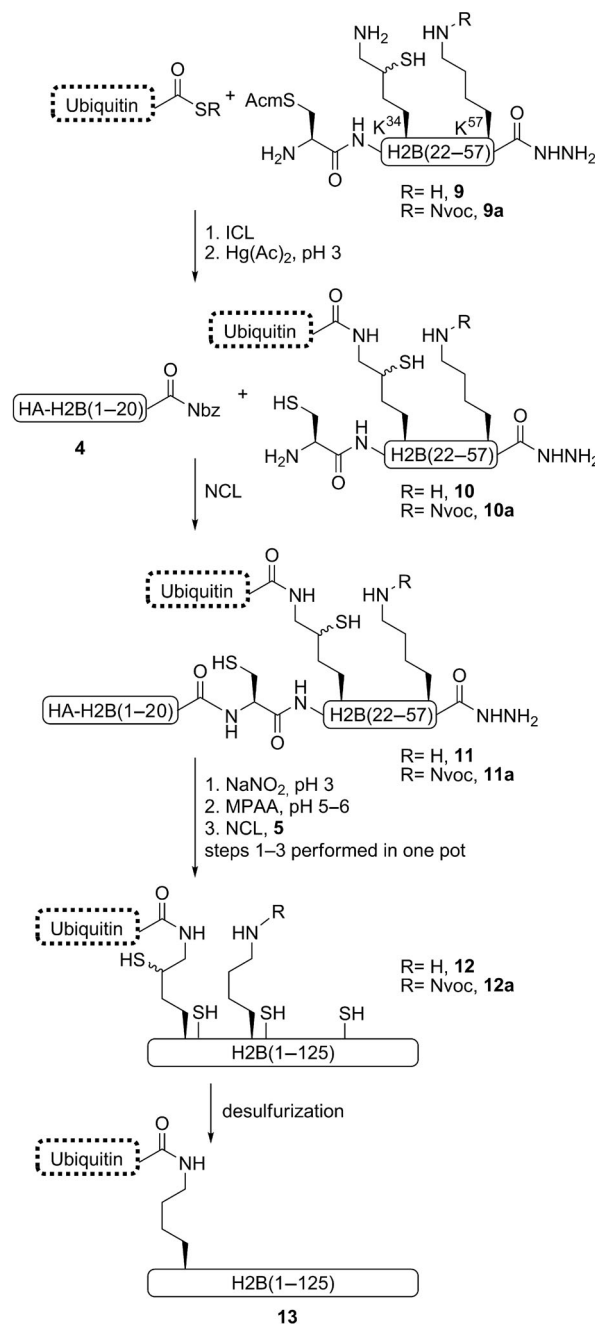


Figure 1. Ligation of **6a** with **5** followed by in situ desulfurization and Nvoc deprotection. A) Oxidation of **6a** followed by MPAA treatment. a: H2B(1–57) with lactam formation (m/z obsd 7425.0 Da, calcd 7425.6 Da); b: thioester **6a** (m/z obsd 7832.0 Da, calcd 7831.6 Da); #: oxidized MPAA. B) Ligation of **6a** with **5** after 5 h. c: H2B(1–57) with lactam formation (m/z obsd 7425.0 Da, calcd 7425.6 Da); d: hydrolysis of **6a** (m/z obsd 7682.2 Da, calcd 7681.6 Da); e: **5** (m/z obsd 7530.0 Da, calcd 7530.7 Da); f: ligation product **7a** (m/z obsd 15194.1 Da, calcd 15194.3 Da). C) Desulfurization and Nvoc deprotection of H2B(1–125) (**7a**). g: no identified mass; h: H2B protein (**8**, m/z obsd 14859.5 Da, calcd 14858.1 Da).

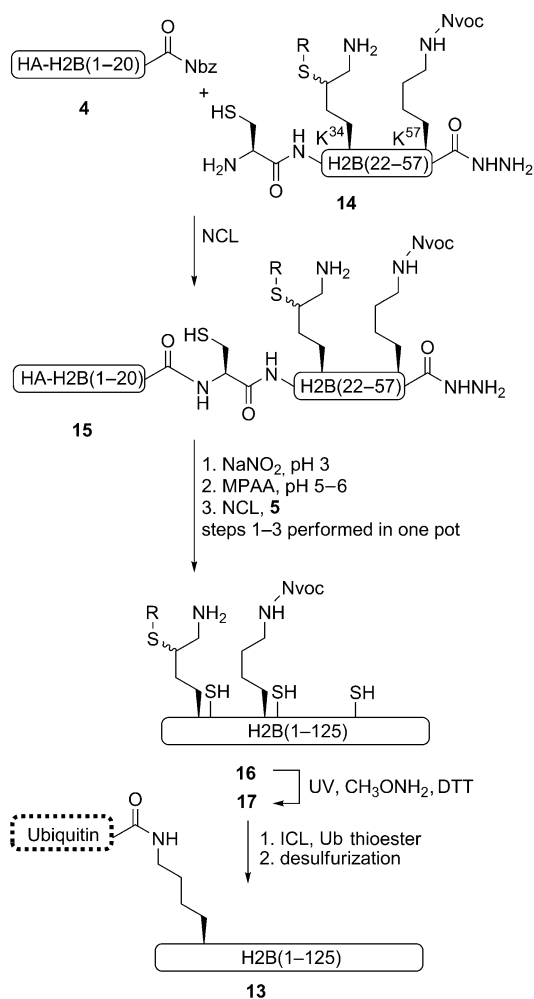
Cys(Acm)-[H2B(20–57)]K34K* (**9a**), and allow sequential ligation after the ubiquitination step (Scheme 2). With all the fragments in hand, the synthesis started with the ligation of **9a** with the Ub thioester followed by Acm deprotection to give ubiquitinated Cys-[H2B(20–57)]K34K* (**10a**). This intermediate was ligated with peptide **4** to furnish ubiquitinated H2B(1–57)-NHNH₂ (**11a**). Contrary to our results with unmodified H2B, the oxidation of the acyl hydrazide in **11a** did not give the desired thioester intermediate, but produced an unidentified mixture. To our surprise, the unprotected K57 fragment **11** was partly oxidized and resulted in 60% conversion to the desired thioester product along with 40% lactamization and hydrolysis side products (Figure S17A, Supporting Information). As observed previously, the lactam formation, which progressed during the course of the reaction to reach about 50% (Figure S17B and C, Supporting Information), resulted in a very low yield ($\approx 15\%$) of the ligation product **12**. The ligation product was then desulfurized to give the final product H2K34Ub (**13**) in 40% yield after 8 hours (Figure 3A). All attempts to increase the yield were unsuccessful. Obviously and for unknown reasons, the presence of Ub in **11** (or **11a**) affected the oxidation step. Consequently, we reasoned that the ICL step with Ub should be postponed to after the oxidation–ligation step when using the Nvoc protecting group, which might prevent lactam formation.

We envisioned the assembly of full-length H2B bearing orthogonally protected δ -mercaptolysine at position 34. Hence, the H2B(21–57) fragment **14** was prepared with an unprotected Cys at the N-terminus, K57 protected with Nvoc,



Scheme 2. Synthesis of H2K34Ub by applying ubiquitination, followed by the full assembly of H2B.

and δ -mercaptolysine protected with the photolabile *o*-nitrobenzyl^[16] protecting group (Scheme 3). The ligation between peptides **14** and **4** proceeded in a straightforward manner, giving the ligation product **15** in 50% yield. The oxidation of peptide **15**, which bears an acyl hydrazide, also proceeded smoothly with around 90% conversion, followed by further ligation with peptide **5** to furnish the full-length H2B(1–125) (**16**), which bears both photolabile protecting groups, in 35% yield (Figure 2B). The removal of these groups in the presence of methoxylamine and 1,6-dithiothreitol (DTT), was complete within 2 hours, giving the desired product [H2B(1–125)]K34K* (**17**) in 30% yield (Scheme 3). This relatively low yield is due to the oxidation of the two Met



Scheme 3. Synthesis of H2K34Ub by applying the synthesis of H2B, followed by site-specific ubiquitination. R = *o*-nitrobenzyl.

residues ($\approx 15\%$) that are present in the sequence, which was difficult to avoid (Figure 2C). Additional oxidation of these Met residues was also observed during the ligation of **17** with the Ub thioester, resulting in 15% yield (Figure 2D). Desulfurization of the ligation product furnished the final product of H2K34Ub (**13**; Figure 3A). Although the oxidation of Met residues reduced the yield of the synthesis, we believe that their substitution with norleucine will significantly increase the yield of this synthetic strategy toward this construct and others analogues.

To begin analyzing the biochemical properties of H2K34Ub, we reconstituted histone octamers together with recombinant histones H2A, H3, and H4.^[17] No difference in quality and efficiency of association compared to unmodified H2B could be observed in gel-filtration analysis (Figure 3B). We then assembled H2K34Ub containing octamers into dodecamer nucleosomal arrays by the classical salt-dialysis method using the $12 \times 200 \text{ bp} \times 601$ DNA sequence template.^[18] Like octamers that contain recombinant unmodified H2B or tagged H2B-6XHis, H2K34Ub-containing material assembled into chromatin as indicated by an upward shift in native agarose gel electrophoresis analysis (Figure 3C). Nucleosomal digestion with the unspecific DNase and micro-

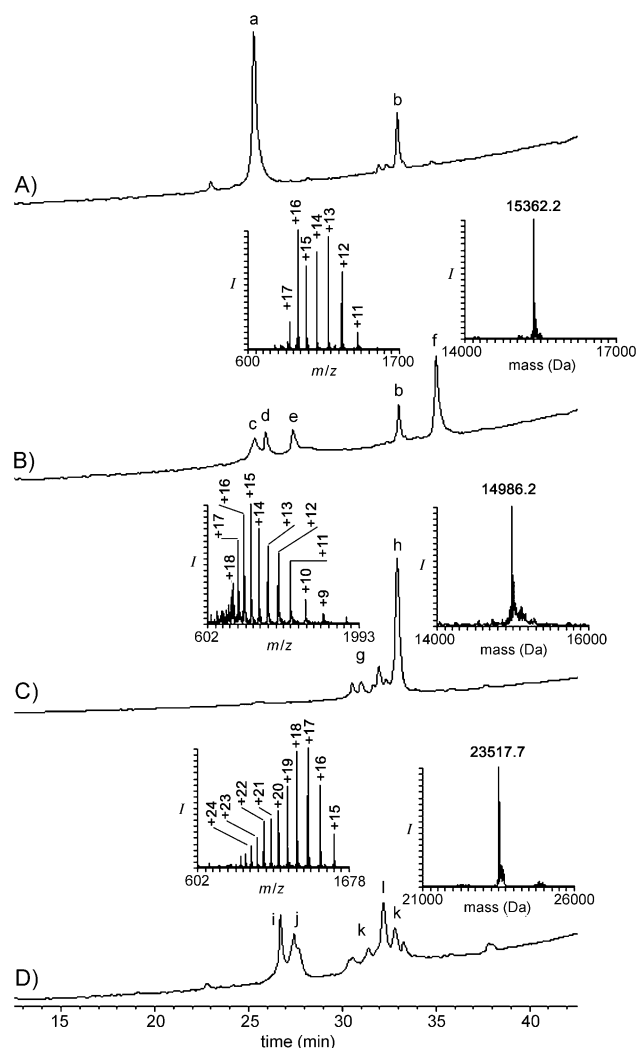


Figure 2. Acyl hydrazide oxidation and ligation of **14**, with **5**, followed by UV irradiation then ligation with the Ub thioester. A) Reaction at $t=0$, a: **14** (m/z obsd 7862.9 Da, calcd 7863.6 Da); b: **5** (m/z obsd 7530.0 Da, calcd 7530.7 Da). B) oxidation and ligation of peptide **15** with **5** after 5 h. c: lactamization of **15**; d: **15** hydrolysis; e: **15** bearing thioester (m/z obsd 7998.0 Da, calcd 7998.6 Da); f: peptide **5** (m/z obsd 7530.0 Da, calcd 7530.7 Da); g: ligation product **16** (m/z obsd 15362.2 Da, calcd 15360.3 Da). C) Nvoc and *o*-nitrobenzyl deprotection using UV irradiation at 365 nm for 2 h. g: Met oxidation in peptide **17**; h: deprotection product **17** with free K^+ (m/z obsd 14986.2 Da, calcd 14986.3 Da). D) Ligation of **17** with the Ub thioester after 8 h. i: hydrolyzed Ub (m/z obsd 8546.2 Da, calcd 8546.8 Da); j: Ub thioester (m/z obsd 8696.9 Da, calcd 8696.8 Da); k: oxidized ligation products; l: ligation product **12** (m/z obsd 23517.7 Da, calcd 23515.1 Da).

coccal nuclease verified reconstitution of regularly spaced nucleosomes on the DNA (Figure 3D). From these results, we concluded that ubiquitination of H2B at Lys34 does not interfere with the assembly of nucleosomes.

In summary, we have developed a novel approach to the chemical synthesis of the unmodified H2B protein and its Lys34-ubiquitinated analogue by combining state of the art chemical-ligation methods. Despite several challenges throughout the synthesis, our approaches should, after further improvements, enable the large-scale production of H2B that

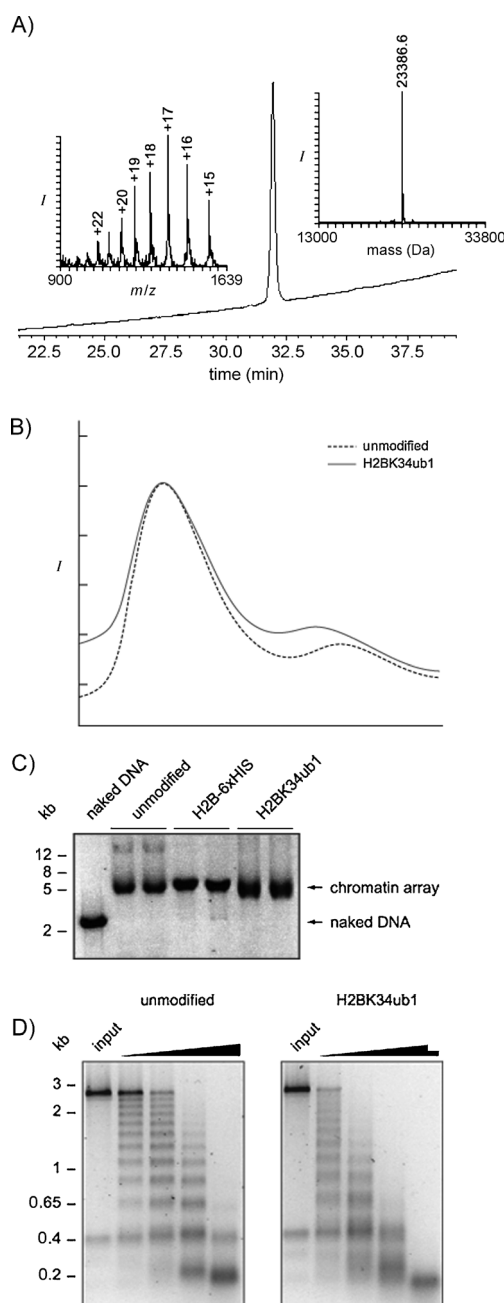


Figure 3. A) Analytical HPLC-MS of the purified synthetic H2K34Ub (**13**, m/z obsd 23386.6 Da, calcd 23386.9 Da). B) Size-exclusion chromatography (Superdex-200) of reconstituted histone octamers. Absorbance signal at 280 nm shows a main peak representing histone octamer (left) and a smaller peak corresponding to H2A/H2B dimer (right). C) The indicated chromatin templates and naked DNA were analyzed by native agarose gel electrophoresis (0.6% TAE buffer) and stained with ethidium bromide. Position of DNA size marker is indicated on the left. D) Micrococcal nuclease digestion assay (time points at 10 s, 30 s, 2 min and 30 min) of reconstituted chromatin. DNA from individual reactions was purified and analyzed on 1% TBE-agarose gel stained with SYBR Gold. Position of DNA size marker is indicated on the left. TAE buffer = buffer containing tris(hydroxymethyl)aminomethane (Tris), acetic acid, and ethylenediaminetetraacetate (EDTA).

is modified at essentiality any site. Together with the initial functional characterization, our work paves the road for

further investigation of the role of Lys34 ubiquitination on the binding to chromatin.

Received: May 5, 2013

Published online: June 21, 2013

Keywords: chromatin · H2B protein · peptide ligation · proteins · ubiquitination

- a) T. Suganuma, J. L. Workman, *Cell* **2008**, *135*, 604–607; b) D. E. Wright, C. Y. Wang, C. F. Kao, *Front. Biosci.* **2012**, *17*, 1051–1078.
- a) J. Kim, S. B. Hake, R. G. Roeder, *Mol. Cell* **2005**, *20*, 759–770; b) R. N. Larabee, S. M. Fuchs, B. D. Strahl, *Genes Dev.* **2007**, *21*, 737–743.
- R. Y. Tweedie-Cullen, J. M. Reck, I. M. Mansuy, *J. Proteome Res.* **2009**, *8*, 4966–4982.
- a) R. K. McGinty, J. Kim, C. Chatterjee, R. G. Roeder, T. W. Muir, *Nature* **2008**, *453*, 812–816; b) M. Haj-Yahya, N. Eltarteer, S. Ohayon, E. Shema, E. Kotler, M. Oren, A. Brik, *Angew. Chem.* **2012**, *124*, 11703–11707; *Angew. Chem. Int. Ed.* **2012**, *51*, 11535–11539; c) G. Fuchs, E. Shema, R. Vesterman, E. Kotler, Z. Wolchinsky, S. Wilder, L. Golomb, A. Pribluda, F. Zhang, M. Haj-Yahya, E. Feldmesser, A. Brik, X. C. Yu, J. Hanna, D. Aberdam, E. Domany, M. Oren, *Mol. Cell* **2012**, *46*, 662–673.
- a) L. Wu, B. M. Zee, Y. Wang, B. A. Garcia, Y. Dou, *Mol. Cell* **2011**, *43*, 132–144; b) Z.-W. Sun, C. D. Allis, *Nature* **2002**, *418*, 104–108.
- L. Z. Yan, P. E. Dawson, *J. Am. Chem. Soc.* **2001**, *123*, 526–533.
- K. P. Chiang, M. S. Jensen, R. K. McGinty, T. W. Muir, *ChemBioChem* **2009**, *10*, 2182–2187.
- K. S. Ajish Kumar, M. Haj-Yahya, D. Olschewski, H. A. Lashuel, A. Brik, *Angew. Chem.* **2009**, *121*, 8234–8238; *Angew. Chem. Int. Ed.* **2009**, *48*, 8090–8094.
- a) G.-M. Fang, Y.-M. Li, F. Shen, Y.-C. Huang, J.-B. Li, Y. Lin, H.-K. Cui, L. Liu, *Angew. Chem.* **2011**, *123*, 7787–7791; *Angew. Chem. Int. Ed.* **2011**, *50*, 7645–7649; b) J.-S. Zheng, S. Tang, Y. Guo, H.-N. Chang, L. Liu, *ChemBioChem* **2012**, *13*, 542–546.
- a) J. B. Hendrickson, *J. Am. Chem. Soc.* **1977**, *99*, 5439–5450; b) D. Bang, B. L. Pentelute, S. B. H. Kent, *Angew. Chem.* **2006**, *118*, 4089–4092; *Angew. Chem. Int. Ed.* **2006**, *45*, 3985–3988; c) J. Y. Lee, D. Bang, *Biopolymers* **2010**, *94*, 441–447; d) L. Raibaut, N. Ollivier, O. Melnyk, *Chem. Soc. Rev.* **2012**, *41*, 7001–7015; e) N. Ollivier, J. Vicogne, A. Vallin, H. Drobacq, R. Desmet, O. El Mahdi, B. Leclercq, G. Goormachtigh, V. Fafeur, O. Melnyk, *Angew. Chem.* **2012**, *124*, 213–217; *Angew. Chem. Int. Ed.* **2012**, *51*, 209–213; f) J. Li, S. Dong, S. D. Townsend, T. Dean, T. J. Gardella, S. J. Danishefsky, *Angew. Chem.* **2012**, *124*, 12429–12433; *Angew. Chem. Int. Ed.* **2012**, *51*, 12263–12267.
- J. B. Blanco-Canosa, P. E. Dawson, *Angew. Chem.* **2008**, *120*, 6957–6961; *Angew. Chem. Int. Ed.* **2008**, *47*, 6851–6855.
- E. C. B. Johnson, S. B. H. Kent, *J. Am. Chem. Soc.* **2006**, *128*, 6640–6646.
- Q. Wan, S. J. Danishefsky, *Angew. Chem.* **2007**, *119*, 9408–9412; *Angew. Chem. Int. Ed.* **2007**, *46*, 9248–9252.
- For the mechanism of Nvoc and *o*-nitrobenzyl removal see: A. P. Pelliccioli, J. Wirz, *Photochem. Photobiol. Sci.* **2002**, *1*, 441–458.
- a) J. Offer, C. N. C. Boddy, P. E. Dawson, *J. Am. Chem. Soc.* **2002**, *124*, 4642–4646; b) P. Wang, L. P. Miranda, *Int. J. Pept. Res. Ther.* **2005**, *11*, 117–123.
- M. Haj-Yahya, K. S. Ajish Kumar, L. A. Erlich, A. Brik, *Biopolymers* **2010**, *94*, 504–510.
- K. Luger, T. Rechsteiner, T. Richmond in *Chromatin Protocols*, Vol. 119 (Ed.: P. Becker), Humana, Totowa, NJ, **1999**, pp. 1–16.
- P. T. Lowary, J. Widom, *J. Mol. Biol.* **1998**, *276*, 19–42.