HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2016 December 02.

Published in final edited form as:

Science. 2016 November 4; 354(6312): 623–626. doi:10.1126/science.aah4428.

A chemical biology route to site-specific authentic protein modifications

Aerin Yang[1], **Sura Ha**[1], **Jihye Ahn**[2], **Rira Kim**[1], **Sungyoon Kim**[1], **Younghoon Lee**[1], **Jaehoon Kim**[2], **Dieter Söll**[3,*], **Hee-Yoon Lee**[1,*], and **Hee-Sung Park**[1,*]

[1]Department of Chemistry, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea.

[2]Department of Biological Sciences, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea.

[3]Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06511, USA.

Abstract

Many essential biological processes are controlled by posttranslational protein modifications. The inability to synthetically attain the diversity enabled by these modifications limits functional studies of many proteins. We designed a three-step approach for installing authentic posttranslational modifications in recombinant proteins. We first use the established *O*-phosphoserine (Sep) orthogonal translation system to create a Sep-containing recombinant protein. The Sep residue is then dephosphorylated to dehydroalanine (Dha). Last, conjugate addition of alkyl iodides to Dha, promoted by zinc and copper, enables chemoselective carbon-carbon bond formation. To validate our approach, we produced histone H3, ubiquitin, and green fluorescent protein variants with site-specific modifications, including different methylations of H3K79. The methylated histones stimulate transcription through histone acetylation. This approach offers a powerful tool to engineer diverse designer proteins.

Posttranslational modifications (PTMs) play vital roles in expanding protein functional diversity and critically affect numerous biological processes (1). The availability of proteins with specific modifications at selected residues is essential for experimental strategies to investigate fundamental biological mechanisms. Methods to generate diverse native protein covalent modifications do not exist at present. Genetic code expansion approaches are useful in producing recombinant proteins with specific modifications (2, 3) but rely on the availability of an orthogonal transfer RNA (tRNA) synthetase•tRNA pair for acylation of a

*Corresponding author. dieter.soll@yale.edu (D.S.); hylee@kaist.ac.kr (H.-Y.L.); hspark@kaist.ac.kr (H.-S.P.).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/354/6312/623/suppl/DC1

Materials and Methods

Figs. S1 to S23

Tables S1 to S7

References (26–28)

specific noncanonical amino acid. Despite much progress, the creation of many important protein modifications (e.g., trimethyllysine) is not yet feasible. Among chemical conjugation approaches (4), Cys-based strategies have been widely applied to generate protein conjugates (5) and mimics of PTMs (6, 7). Yet the final products are PTM analogs whose value for finding unidentified properties of the natural system may be questionable (8). Thus, despite extensive efforts, synthetic approaches for many authentic PTMs are not available, because no carbon-carbon bond–forming reactions have thus far been successfully applied to protein modifications, despite the prevalence of such reactions in organic chemistry (9).

Here we propose a three-step strategy (Fig. 1A) that, in principle, is applicable to generating diverse forms of authentic and selective protein modifications. (i) The site of the intended modification is established by cotranslational *O*-phosphoserine (Sep) incorporation into a recombinant protein by using the Sep orthogonal *Escherichia coli* translation system (10, 11). (ii) Then, the Sep residue of the purified recombinant protein is converted by phosphate removal to dehydroalanine (Dha),which in turn serves as a radicalophile, enabling a bioorthogonal chemical reaction. (iii) Last, PTM moieties are directly coupled to Dha through metal-mediated conjugate additions of alkyl iodides in aqueous solution, facilitating chemoselective carbon-carbon bond formation in proteins (Fig. 1B).

To demonstrate the feasibility and versatility of our strategy, we set out to generate recombinant *Xenopus laevis* histone H3K79 with five different modifications: monomethylation, dimethylation, trimethylation, formylation, and acetylation. Reversible lysine methylation in proteins presents the most complex and dynamic modification (12). Several approaches have attempted lysine methylation (13), but the trimethylated product has proved elusive. Histone H3K79 appears to be dynamically regulated (14) and associated with diverse cellular processes; its exact role has not been fully examined because such authentically modified histones have not been generated.

First, the Sep-containing histone H3Sep79 was made by expressing a *X. laevis* histone H3 mRNA containing a UAG codon at position 79 and a coding sequence for a C-terminal His_6 tag in *E. coli* containing the orthogonal Sep translation system [the engineered SepRS9•$tRNA^{Sep}$ pair and the evolved elongation factor (EF)–Sep21] (11). H3Sep79 was routinely obtained in good yield (~20 mg/liter of culture) (fig. S1). The incorporation of Sep at the intended position (79) was confirmed by matrix-assisted laser desorption/ ionization–time-of-flightmass spectrometry (MALDITOF MS) analysis of the purified recombinant histone (fig. S2) and its tryptic peptides (Fig. 2A and tables S1 and S2).

Because phosphoamino acids are known to be labile under alkaline conditions (15), we attempted to convert Sep to Dha by mild alkali treatment. To determine the optimal procedure, purified H3Sep79 was incubated with LiOH, $Ba(OH)_2$, or $Sr(OH)_2$ under different conditions, followed by neutralization with acetic acid and dialysis against distilled water (fig. S3). The reaction progress was analyzed by MALDI-TOFMS after in-gel trypsin digestion (tables S1 and S2). Near-complete disappearance of the Sep residue and concurrent generation of Dha took 30 min at room temperature in 40mM$Ba(OH)_2$ solution. Mass analysis of whole proteins (fig. S4) and tryptic peptides (Fig. 2A) demonstrated that

these conditions led to highly selective chemical transformation of H3Sep79 into H3Dha79 with no noticeable protein oxidation or other side reactions (16).

Dha-mediated conjugation approaches have been successfully applied to the generation of thiol-linked PTM analogs or mimics (17, 18). To produce authentic PTMs from Dha, we needed a new coupling scheme enabling chemoselective carbon-carbon bond formation in proteins (Fig. 1B). We were encouraged by reports of water-based organic reactions in which an alkene group functions as a radical acceptor (19). Knowing that Dha could function as a radical acceptor, we chose alkyl radicals that can be generated from alkyl iodides by transition metals (19). Inspired by recent metal-mediated conjugate addition reactions (20), we reasoned that organozinc species would form from the precursor halides and Zn metal. Subsequently, zinc-to-copper transmetalation would occur, generating organo copper reagents, which would lead to the formation of alkyl radical species. Last, conjugate coupling of the alkyl radical with the Dha residue of a protein would take place (Fig. 1B). To demonstrate the proposed coupling scheme, we tested first the addition of trimethyl iodide (3–iodo-*N,N,N,*-trimethylpropan-1-amine, 3; fig. S5) to H3Dha79 to generate the expected product H3K79me3. The reaction products were analyzed by Western blotting using an antibody against H3K79me3. Our initial attempts to directly use water-based organic reactions were unsuccessful. After testing a large array of conditions (described in the supplementary materials, figs. S6 to S16)—including different buffers and pH ranges, surfactants for protein stabilization, essential reagents for the metal-mediated coupling reaction (Zn metals and copper salts), and possible auxiliary reagents [e.g., tetramethylethylenediamine (TMEDA)]—we arrived at conditions that reproducibly led to the formation of carbon-carbon bonds in high yield. The deduced optimal conditions were H3Dha79 (10 μM), alkyl iodide (30 mM), Zn powder (0.4 mg), $Cu(OAc)_2$ (1 mM), Triton X-100 (2.0 weight %), and TMEDA (10 mM) in sodium acetate (pH 4.5, 0.5 M). Reactions (20 to 50 μl) were incubated at room temperature.

With optimized reaction conditions in hand, we attempted to synthesize H3K79me1, H3K79me2, and H3K79me3 by incubating H3Dha79 with three different methyl iodides: monomethyl iodide (3–iodo-*N*-methylpropan-1-amine, **1**), dimethyl iodide (3–iodo-*N,N*-dimethylpropan-1-amine, **2**), or trimethyl iodide (Fig. 2C and fig. S5). MALDI-TOF MS analysis revealed the disappearance of Dha and the concomitant generation of methylated lysine residues (Fig. 2A). The coupling reactions were highly selective and efficient (normally >80%), as demonstrated by mass analyses of tryptic peptides (Fig. 2A and table S2) and whole proteins (fig. S17), and had a good recovery yield (between 50 and 70%). Selective and differential lysine methylations were also demonstrated by Western blot analysis using antibodies against H3K79me1, H3K79me2, and H3K79me3 (Fig. 2B). To determine whether other PTM moieties could be coupled onto H3Dha79, we synthesized formyl iodide [*N*-(3–iodopropyl)formamide, **4**] and acetyl iodide [*N*-(3–iodopropyl)-acetamide, **5**; fig. S5] and used them for producing histones H3K79N^{ε}-formyl and H3K79N^{ε}-acetyl, respectively, as verified by MALDITOF MS analysis (fig. S18).

To demonstrate whether our three-step synthesis can be used for modification of other proteins, we set out to generate ubiquitin variants with site-specific modifications. Eight ubiquitin variants with various lysine modifications (acetylation and differential

methylations) at position 33 or 48 were efficiently generated (fig. S19 and tables S3 to S5). Also, we found that different alkyl iodides (iodoethane, **6**; 2-iodopropane, **7**; 2-iodo-2-methylpropane, **8**; 1-iodobutane, **9**; iodocyclopentane, **10**; 3-iodopronionic acid, **11**; and *tert*-butyl-iodobutoxydimetylsilane, **12**) (fig. S5)were efficiently conjugated to Ub33Dha (fig. S20 and table S4). In particular, coupling of dansyl iodide (5–{dimethylamino}-*N*-{2-[(3–iodopropyl)amino]ethyl}naphthalene-1-sulfonamide, **13**; fig. S5) onto Ub33Dha led to site-specific fluorescent dye labeling via carbon-carbon bond formation (figs. S20 and S21 and table S4). Last, site-specific Lys^{204} trimethylation of green fluorescence protein (GFP) by our three-step synthetic route led to the desired product GFPK204me3 but with lower conversion efficiency (~20%) and recovery yield (~30%) compared with histone H3 and ubiquitin (fig. S22 and tables S6 and S7). Thus, the structural context within the protein may affect the ease of modification. These data show that our approach is applicable to a variety of proteins. Because each protein behaves differently with the diverse chemical reagents, individualized optimization should lead to efficient implementation of our synthetic strategy with other proteins.

Next, we performed biochemical assays with the modified histones generated by our three-step approach to examine whether they are fully functional in the biological context. Genome-wide analyses of chromatin has revealed that H3K79 methylation is enriched in actively transcribing regions (14), but its detailed role is not yet well understood. A transcription assay using in vitro assembled chromatin templates and a highly purified RNA polymerase II transcription apparatus provides a biochemically defined system to study the functions of individual histone modifications. The modified histones H3K79me1, H3K79me2, and H3K79me3 were first assembled with H2A, H2B, and H4 to form histone octamers. The methylated histones showed no noticeable difference from the unmodified (intact) histone during octamer assembly and nucleosome reconstitution (fig S23). Recombinant chromatin templates were reassembled with the histone octamers and a p53ML plasmid and were applied to p53 (activator)–and p300 (coactivator)–dependent in vitro transcription assays (Fig. 3A) (21). Intact chromatin lacking methylation showed a low level of transcription only in the presence of both the activator and coactivator, as expected (Fig. 3B). Notably, the levels of basal (activator- and/or coactivator-independent) transcription from chromatins with all three methylation states were greatly enhanced (Fig. 3B). Also, activator- or coactivator-dependent transcription from H3K79-methylated chromatin was elevated (Fig. 3B), demonstrating a direct stimulatory effect of H3K79 methylation on chromatin transcription. More importantly, we also found that H3K79 methylation increased histone acetylation mediated by p300 (Fig. 3C). Histone acetylation was differentially affected by the level of H3K79 methylation (Fig. 3C). Thus, H3K79 methylation is indeed positively associated with transcription activation through p300-mediated histone acetylation, which is differently affected by the various methylation states (Fig. 3D). These results illustrate that themethylated histones are fully active and, depending on the modification level, functionally distinct, demonstrating the utility and gentle nature of our three-step approach.

An open question concerns the diastereose-lectivity of our coupling reaction. It is known that radical and thiol conjugate additions to Dha lead to epimeric mixtures (22–24); a

definitive analysis of the epimeric ratio will require high-resolution crystallography of the protein. However, the steric context of the local protein conformation may substantively affect the final diastereomeric ratio of the products. The fact that our synthetic proteins H3K79me1, H3K79me2, and H3K79me3 were well recognized by antibodies, and that these modified histones could assemble into octamers with biological in vitro activity, underscores the utility of our approach.

We anticipate that with well-tailored alkyl iodides [e.g., (25)], our approach can be extended to produce designer proteins with diverse forms of chemical modifications (e.g., glycosylated amino acids and phosphotyrosine). Such efforts will markedly expand the available chemical diversity in proteins and facilitate the study of many PTM-mediated biological processes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank M. Simon and S. Miller for stimulating discussions. This work was supported by grants from the National Research Foundation of Korea (KAIST Systems Healthcare, 2011-0020322 and 2014M3A6A4075060 to H.-S.P. and 2012M3A9B4027956 and 2012M3A9C6049937 to J.K.) and from the U.S. National Institute for General Medical Sciences (GM22854-42 to D.S.). All data are available in the supplementary materials. H.-S.P., H.-Y.L., A.Y., and S.H. are inventors on patent application no. 62/383,245, submitted by the Korea Advanced Institute of Science and Technology, which covers protein chemical modifications mediated by carbon-carbon bond formation. H.-S.P., H.-Y.L., and A.Y. conceived and designed the study; H.-S.P. and A.Y. performed and analyzed all experiments, except chemical synthesis and in vitro transcription assays, which were conducted by S.H., R.K, Y.L., J.A., and J.K; H.-S.P., H.-Y.L., and D.S. supervised the project; and H.-S.P. and D.S. wrote the manuscript.

REFERENCES AND NOTES

1. Walsh CT, Garneau-Tsodikova S, Gatto GJ Jr. Angew. Chem. Int. Ed. Engl. 2005; 44:7342–7372. [PubMed: 16267872]
2. Lang K, Chin JW. Chem. Rev. 2014; 114:4764–4806. [PubMed: 24655057]
3. Liu CC, Schultz PG. Annu. Rev. Biochem. 2010; 79:413–444. [PubMed: 20307192]
4. Spicer CD, Davis BG. Nat. Commun. 2014; 5:4740. [PubMed: 25190082]
5. Vinogradova EV, Zhang C, Spokoyny AM, Pentelute BL, Buchwald SL. Nature. 2015; 526:687–691. [PubMed: 26511579]
6. Simon MD, et al. Cell. 2007; 128:1003–1012. [PubMed: 17350582]
7. van Kasteren SI, et al. Nature. 2007; 446:1105–1109. [PubMed: 17460675]
8. Nguyen DP, Garcia Alai MM, Virdee S, Chin JW. Chem. Biol. 2010; 17:1072–1076. [PubMed: 21035729]
9. Hart DJ. Science. 1984; 223:883–887. [PubMed: 17781610]
10. Park HS, et al. Science. 2011; 333:1151–1154. [PubMed: 21868676]
11. Lee S, et al. Angew. Chem. Int. Ed. Engl. 2013; 52:5771–5775. [PubMed: 23533151]
12. Greer EL, Shi Y. Nat. Rev. Genet. 2012; 13:343–357. [PubMed: 22473383]
13. Liu WR, Wang YS, Wan W. Mol. Biosyst. 2011; 7:38–47. [PubMed: 21088799]
14. Nguyen AT, Zhang Y. Genes Dev. 2011; 25:1345–1358. [PubMed: 21724828]
15. Oda Y, Nagasu T, Chait BT. Nat. Biotechnol. 2001; 19:379–382. [PubMed: 11283599]
16. Wang J, Schiller SM, Schultz PG. Angew. Chem. Int. Ed. Engl. 2007; 46:6849–6851. [PubMed: 17685371]

17. Guo J, Wang J, Lee JS, Schultz PG. Angew. Chem. Int. Ed. Engl. 2008; 47:6399–6401. [PubMed: 18624319]
18. Chalker JM, Lercher L, Rose NR, Schofield CJ, Davis BG. Angew. Chem. Int. Ed. Engl. 2012; 51:1835–1839. [PubMed: 22247073]
19. Postigo A, Nudelman NS. Coord. Chem. Rev. 2011; 255:2991–3030.
20. Lipshutz BH, Huang S, Leong WW, Zhong G, Isley NA. J. Am. Chem. Soc. 2012; 134:19985–19988. [PubMed: 23190029]
21. Kim J, Guermah M, Roeder RG. Cell. 2010; 140:491–503. [PubMed: 20178742]
22. Crich D, Davies JW. Tetrahedron. 1989; 45:5641–5654.
23. Zhu Y, van der Donk WA. Org. Lett. 2001; 3:1189–1192. [PubMed: 11348191]
24. Müller MM, Muir TW. Chem. Rev. 2015; 115:2296–2349. [PubMed: 25330018]
25. Garcia Fernandez JM, Mellet CO, Pradera Adrian MA, Mota JF. Carbohydr. Res. 1992; 216:21–32.

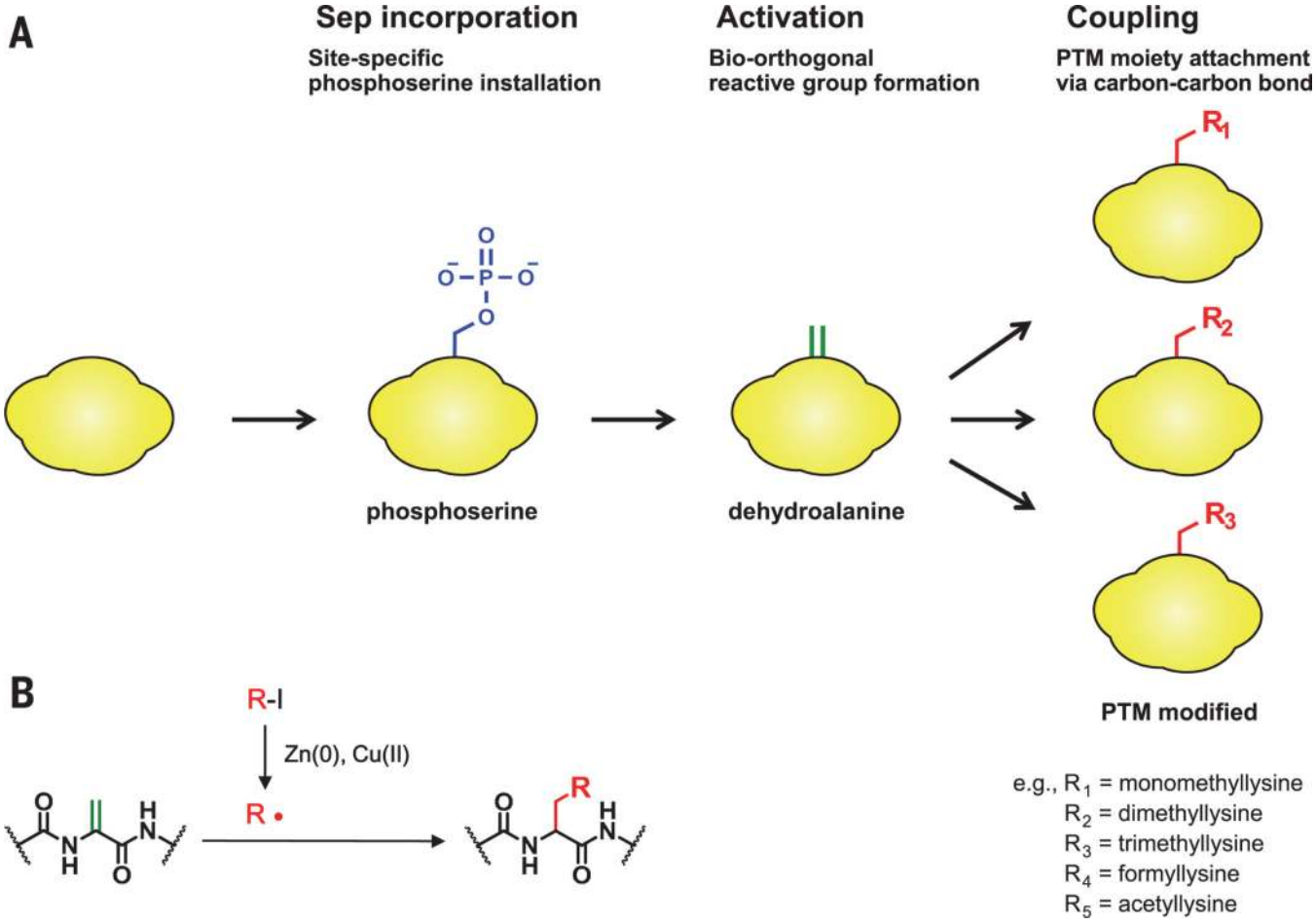

Fig. 1. Approach for protein chemical modifications

(**A**) Schematic representation of the three-step synthesis of proteins with authentic PTMs.

(**B**) Our coupling scheme enables chemoselective carbon-carbon bond formation in proteins. R, PTM moiety.

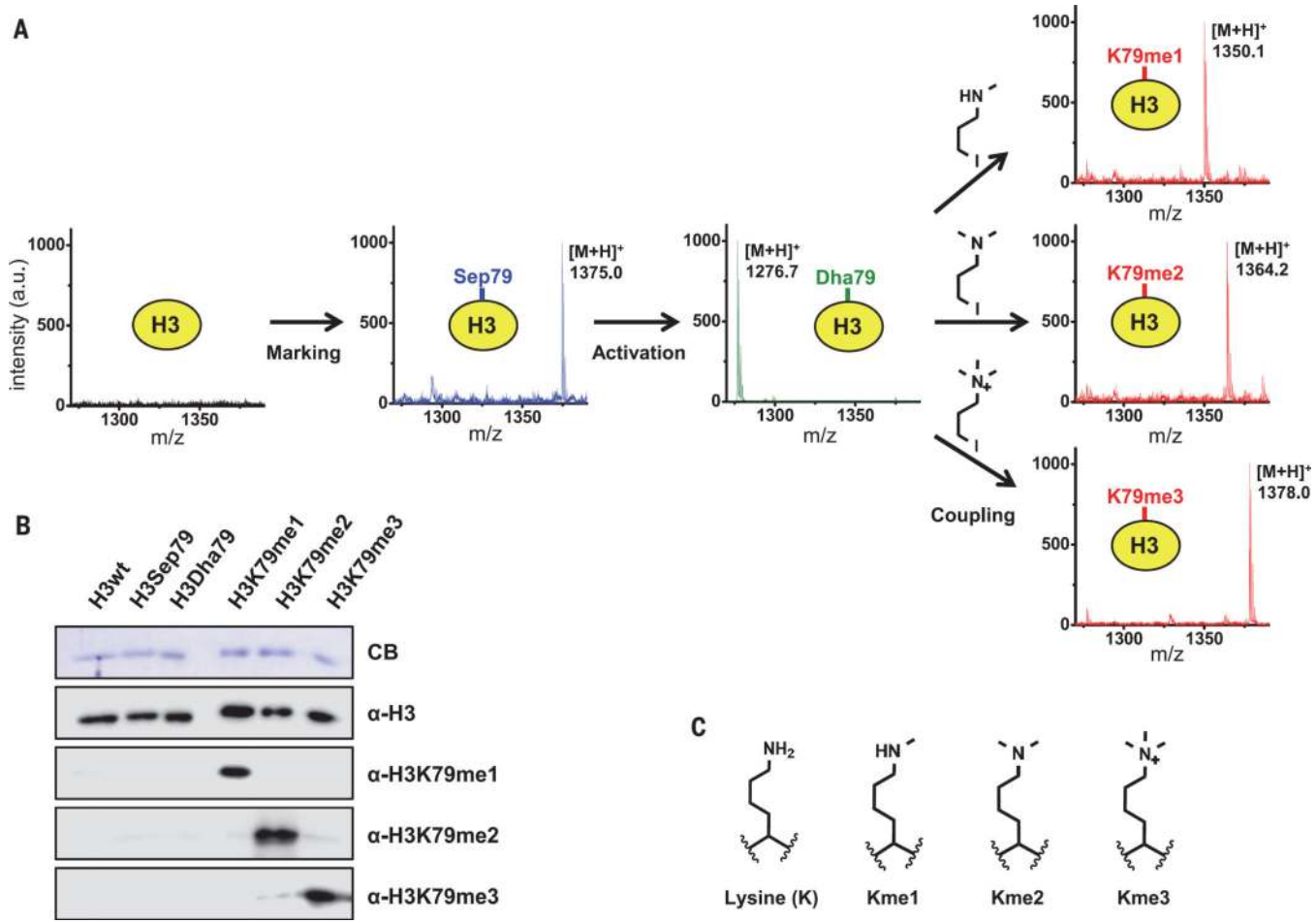

Fig. 2. Synthesis of proteins with selective differential methylations

(**A**) MALDI-TOF MS analysis of proteins after trypsin digestion. Formation of H3Sep79 is evidenced by the Sep-containing peptide (Sep79, blue).Synthesis of H3Dha79 generates a new Dha-containing peptide (Dha79, green) but eliminates the Sep-containing peptide. Coupling of the methyl iodides onto H3Dha79 produces new methylated lysine-containing peptides (K79me1, K79me2, and K79me3, red). *m/z*, mass/charge ratio; $[M+H]^+$, monoisotopic mass; a.u., arbitrary units. (**B**) Western blot analysis of the modified proteins, using antibodies (α) against H3K79me1, H3K79me2, and H3K79me3. CB, Coomassie blue. (**C**) Chemical structures of lysine and differentially methylated lysine residues.

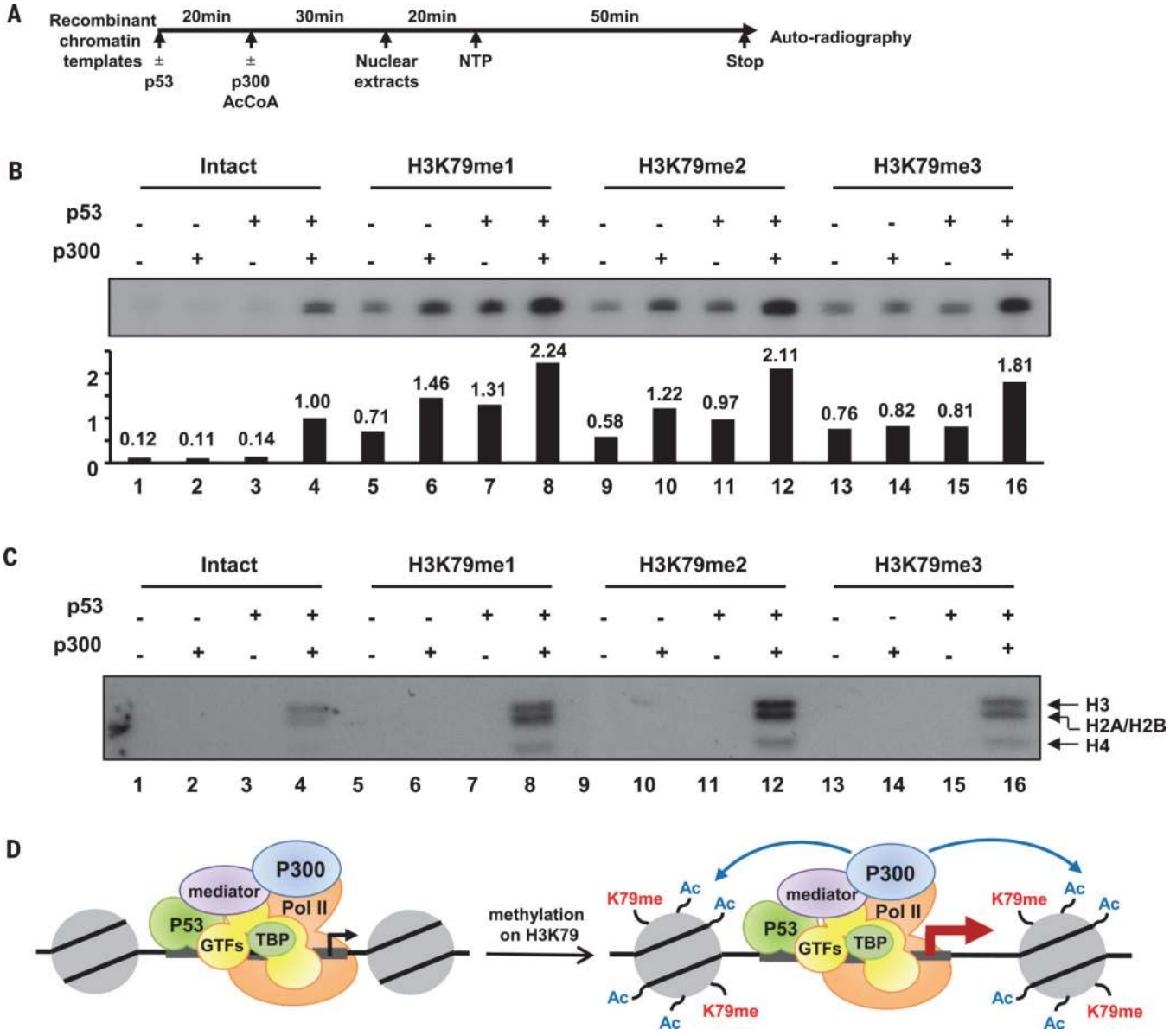

Fig. 3. Effects of H3K79 methylation on chromatin transcription

(**A**) Schematic of the standard in vitro transcription assay. AcCoA, acetyl coenzyme A; NTP, nucleoside triphosphate. (**B**) Effect of methylated H3K79 on chromatin transcription (*x* axis, lane number; *y* axis, relative transcription activity. (**C**) Effect of methylated H3K79 on p300-mediated chromatin acetylation. Histone acetylation status was monitored by fluorography. (**D**) Schematic representation of transcriptional activation by H3K79 methylation. Pol, polymerase; GTF, general transcription factor; TBP, TATA-binding protein.