Unlocking HIV's Secrets: Synthesis of Histone H3K36Me3 to Investigate LEDGF Protein Binding to Modified Nucleosomes using Cryo-Electron Microscopy

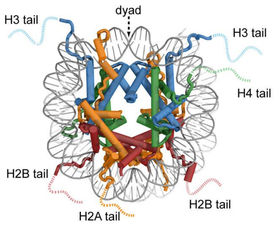
Ramesh M. Chingle[1], Samanthreddy Kedika[1], Tao Jing[2], Dmitry Lyumkis[2,*] and Terrence R. Burke, Jr. [1,*]

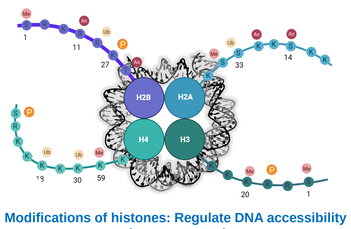
[1]*Chemical Biology Laboratory, National Cancer Institute, Frederick, Maryland, 21702, United States*

[2]*The Salk Institute for Biological Studies, Laboratory of Genetics, La Jolla, California 92037, United States*

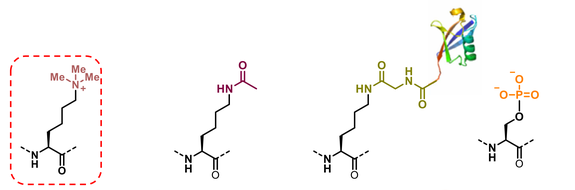
https://doi.org/10.17952/37EPS.2024.P2298

Background: Nucleosome structure and histone



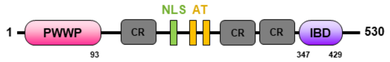


Modifications of histones: Regulate DNA accessibility and gene expression



Histone post-translational modifications (PTMs) intricately orchestrate the epigenetic landscape, profoundly influencing chromatin dynamics and functional gene expression. Among PTMs, methylation of histone lysine residues emerges as a pivotal regulator in diverse biological processes, notably gene transcription. Our research endeavors to engineer modified histones with augmented affinity for the methyl-lysine binding site within the PWWP domain of LEDGF/p75, a pivotal cellular cofactor of HIV-1 integrase (IN). Leveraging native chemical ligation (NCL), we introduce non-standard amino acid residues into histone 3.3, facilitating subsequent structural and functional investigations. This study unveils optimized protocols merging solid-phase peptide synthesis with sequential NCL, enabling the synthesis of fully synthetic variants of the methylated 135-residue protein, H3.3K36(Me)3, featuring trimethylated and dimethylated lysine at residue 36. These tailored histones are envisioned to form nucleosomes with heightened LEDGF affinity, poised for cryo-electron microscopy (cryo-EM) elucidation. Our work not only advances the synthetic accessibility of modified histones but also paves the way for unraveling intricate chromatin-LEDGF interactions, holding promise for understanding HIV-1 integration mechanisms at a molecular level.

Role of LEDGF protein in HIV integration



Domain structure of LEDGF/p75
(cellular cofactor for HIV-1 integrase)

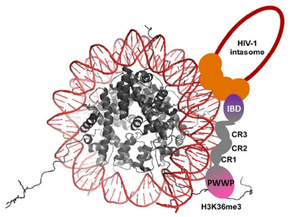


Fig 1. Representation of H3K36(Me)3-modified nucleosome bound to the PWWP domain of LEDGF.

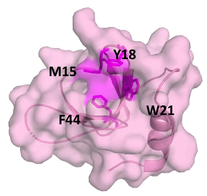


Fig 2. PWWP aromatic cage to recognize H3K36(Me)3 (cation-π interaction).

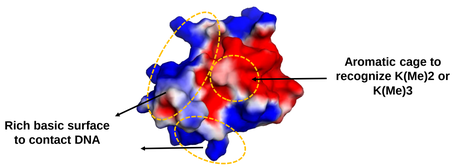
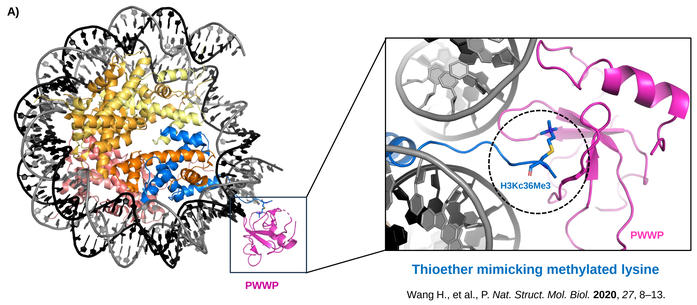


Fig 3. Rich basic surface to contact DNA (H3K36Me3 binding pocket).

cryo-EM structure of the PWWP domain with H3Kc36(Me3) modified nucleosome



Thioether mimicking methylated lysine

Wang H., et al., P. *Nat. Struct. Mol. Biol.* **2020**, *27*, 8–13.

This construct will be used to incorporate into nucleosome

To establish structural basis for interactions in between LEDGF protein (all domains) and H3K36(Me3) modified nucleosome

Improving binding affinity and selectivity

(Difficult to control the site, extent, or degree of Methylation)

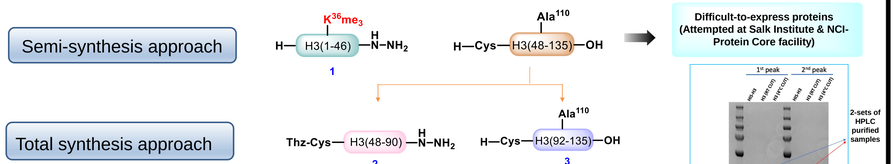
Fig 4. A) cryo-EM structure of the H3K36(Me)3-modified nucleosome bound to the PWWP domain of LEDGF. **B)** Pull-down of LEDGF from crude nuclear extract by using nucleosomes bearing wild-type (WT) H3, MLA (H3Kc36Me3), or H3K36(Me)3, H3K36(Me)3 peptide (panel b Figure adopted from Ruthenburg, et. al., *Chem. Bio. Chem.* **2014**, *15*, 2071-2075).

Goal : Integrated approach to generate H3K36(Me)3 for incorporation into nucleosomes

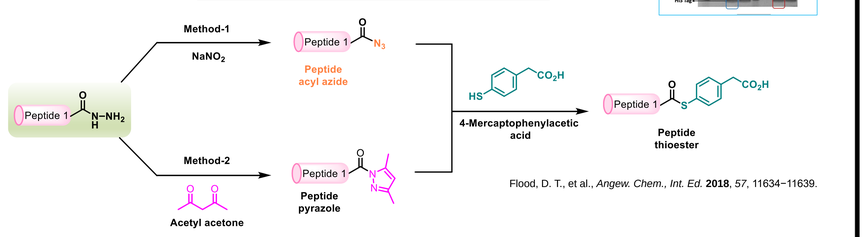
1 ARTKQTARKS TGGKAPRKQL ATKAARKSAP STGGV**K(me)**KPHR YRPGTV 70
ALRE IRRYQKSTEL LIRKLPFQRL

71 VREIAQDFKT DLRFQSAAIG 90 91 ALQEASEAYL VGLFEDTNLC AIHAKRVTIM PKDIQLARRI RGERA 135

Note: Yellow highlight = Site of Cys-to-Ala mutation; Met-to-Nle to avoid Met-oxidation. Cyan highlight = Site of Cys incorporation for native ligation, with subsequent radical desulfurization to Ala; Green highlight = Site of K36Me3 residue;

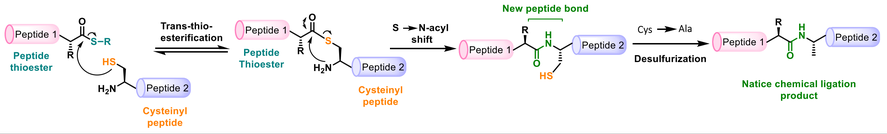


Thioesterification via hydrazide approach



Flood, D. T., et al., *Angew. Chem., Int. Ed.* **2018**, *57*, 11634−11639.

Native chemical ligation reaction mechanism



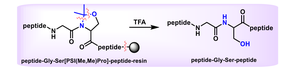
Results: Three segment C-to-N synthesis strategy for H3K36(Me3)

1 ARTKQTARKS TGGKAPRKQL ATKAARKSAP STGGV**K(me)**KPHR YRPGTV 70
ALRE IRRYQKSTEL LIRKLPFQRL

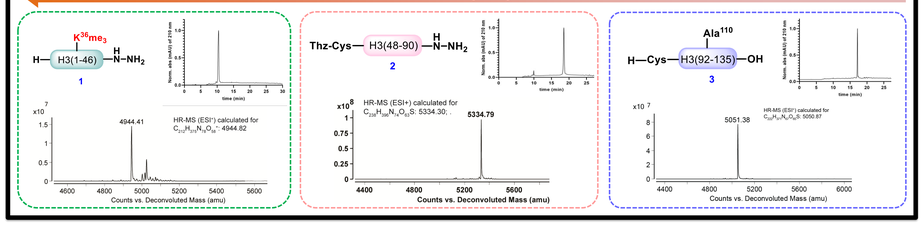
71 VREIAQDFKT DLRFQSAAIG 90 91 ALQEASEAYL VGLFEDTNLC AIHAKRVTIM PKDIQLARRI RGERA 135

Developed synthetic methodology using Automated peptide Synthesizer

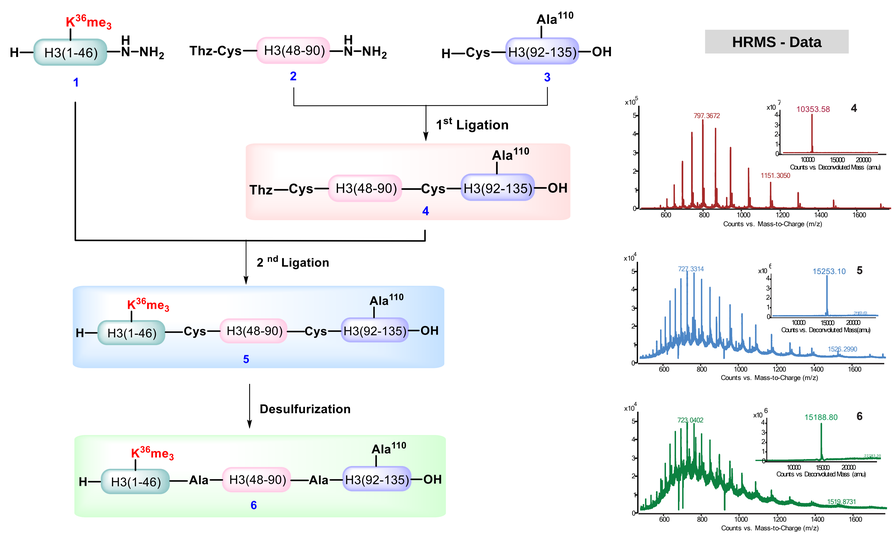
1. All constructs were build using Microwave assisted - 2 min each coupling time
2. **Pseudo-prolines** used to overcome intermolecular aggregation, improving synthesis efficiency



Difficulty in synthesis



Sequential native chemical ligation



Conclusion and Future Perspective

- Optimized robust methodology for total synthesis of H3K36(Me)3 with an excellent overall yield.
- Native chemical ligation (NCL) technique used for stitching together peptide fragments.
- The final 135-mer constructs will be used to prepare nucleosomes for cryo-EM studies.

Acknowledgements

CBL (Chemical Biology Laboratory)
- Terrence R. Burke Jr.*
- Xue Zhi Zhao
- Steve Smith
- Pankaj Mahajan
- Samanthreddy Kedika
- Anand Kulkarni

NCI Biophysics Resource
- Sergey G. Tarasov
- Marzena Dyba

NCI Administration
- Patricia Beard
- Debbie Perone

Salk Institute
- Dmitry Lyumkis*
- Tao Jing