Convergent Hybrid Phase Ligation Strategy for Efficient Total Synthesis of Large Proteins Demonstrated for 212-residue Linker Histone H1.2

Ziyong Z. Hong¹, Ruixuan R. Yu², Xiaoyu Zhang¹, Allison M. Webb¹, Nathaniel L. Burge², Michael G. Poirier^{1,2,3}, Jennifer J. Ottesen^{1,2}

¹Department of Chemistry & Biochemistry, ²Ohio State Biochemistry Program, ³Department of Physics, The Ohio State University, Columbus, 43210, USA.

Correspondence: Jennifer J. Ottesen (ottesen.1@osu.edu)

Abstract: Simple and efficient total chemical synthesis of large proteins remains a significant challenge. Here, we report development of a convergent hybrid phase native chemical ligation (CHP-NCL) strategy that should be generally applicable for facile preparation of large proteins. Key to the strategy is the use of sequential ligation on the solid phase for the directed assembly of ~100-residue segments from short, synthetically accessible peptide components. These segments can then be assembled via convergent solution phase ligation, exploiting *o*-aminoaniline as a chemically flexible cryptic thioester with multiple activation modalities on resin and *in situ*. We demonstrate the feasibility of our approach through the total synthesis of 212-residue linker histone H1.2 in unmodified, phosphorylated, and citrullinated forms, each from eight component peptide segments. We further demonstrate that fully synthetic H1.2 replicates the binding interactions of linker histones to intact mononucleosomes, as a proxy for the essential function of linker histones in the formation and regulation of higher order chromatin structure.

Keywords: convergent ligation • solid phase ligation • native chemical ligation • linker histone • PTMs

The ability to generate homogenous samples of proteins with specific defined modifications remains a limiting step in understanding the structure, dynamic, and functional outcomes of protein post-translational modification (Fierz & Muir, 2012). Of the many strategies that have been developed in the past decades (Chuh, Batt, & Pratt, 2016; Dadová, Galan, & Davis, 2018; Howard, Yu, Gardner, Shimko, & Ottesen, 2015; Luger et al., 2018; Simon et al., 2007), total protein synthesis has the highest potential for combination of multiple modifications throughout a protein sequence, but also represents a significant effort. Typically, this has required division of a protein into accessible segments that can be stitched together using chemoligation strategies such as native chemical ligation (Agouridas et al., 2019; Dawson, Muir, & Kent, 1994; Kent, 2009; Müller & Muir, 2015) (NCL, resulting in a native peptide bond, most commonly at a cysteine site) followed by desulfurization (allowing expanded ligation sites, including alanine) (Jin, Li, Chow, Liu, & Li, 2017; Wan & Danishefsky, 2007; Yan & Dawson, 2001). Sequential (Raibaut, Ollivier, & Melnyk, 2012; Shimko, North, Bruns, Poirier, & Ottesen, 2011), one-pot (Bang & Kent, 2004; Kamo, Hayashi, & Okamoto, 2018; Li et al., 2014; Tang et al., 2015; Zuo, Zhang, Yan, & Zheng, 2018), and convergent (Bang, Pentelute, & Kent, 2006; Durek, Torbeev, & Kent, 2007; Fang, Wang, & Liu, 2012; Qi, He, Ai, Guo, & Li, 2017) ligation strategies have enabled the synthesis of small to medium-sized proteins (<150 residues) but have been limited in scope and practical application by the size and synthetic accessibility of the component peptide segments, the chemical or kinetic control of intermediate ligation specificity, and the number of purification steps required to achieve the final product. In theory, sequential ligation carried out on the solid phase should overcome these barriers, allowing facile recombination of peptide segments while eliminating the need for intermediate

purification (Brik, Keinan, & Dawson, 2000; Canne et al., 1999; Jbara, Seenaiah, & Brik, 2014).

However, in practice, this has proven challenging – solid phase ligation approaches have been successfully demonstrated only for small proteins.

The largest protein or segment prepared by entirely solid phase ligation as documented in the Protein Chemical Synthesis Database (Agouridas, El Mahdi, Cargoët, & Melnyk, 2017) is the 125-residue histone H2B (Jbara et al., 2014). Recently, our laboratory demonstrated that for two proteins (104-residue core histone H4 and 140-residue Cenp-A), solid phase ligation strategies could provide significant improvements in yield and purity in initial ligation steps, but yields were significantly reduced after 75-100 residues were assembled in three ligations (Yu et al., 2016). For the two proteins in question, we achieved our product by carrying out additional sequential solution phase ligation steps for a hybrid phase ligation approach. A similar approach was used by Melnyk and coworkers for preparation of the 180-residue NK1-B, from an 83-residue segment prepared by solid phase ligation, and two ~50-residue synthetic segments ligated in solution (Ollivier et al., 2017). However, these hybrid approaches are inherently limited again by synthetic accessibility of the additional segments prepared by traditional approaches.

To explore the limitations of the solid phase and hybrid approaches, we selected the 212-residue linker histone H1.2. Variants of linker histone H1 facilitate chromatin condensation by interactions with folded nucleosome core particles (Fyodorov, Zhou, Skoultschi, & Bai, 2018). They are known to be heavily post-translationally modified (Izzo & Schneider, 2016), but the functions of the modifications are not well understood, in part because of the dearth of

chemical tools for preparation of the modified proteins. As such, it is a high value target for synthetic protein chemistry but challenging to achieve due to its large size.

Initially, we attempted preparation of histone H1.2 through a simple solid phase ligation strategy from nine component peptides. However, as found in the literature studies described above, we observed significant yield loss on the solid phase after 4 ligations (~100 residues, see Supplemental S5.7).

We hypothesized that we could overcome this limitation by dividing the sequence into two “blocks”, each assembled by ligation of short, synthetically accessible peptide segments on the solid phase. Key to this strategy is use of a cryptic thioester in the N-terminal block that would be orthogonal to all required ligation chemistry, but could be selectively activated *in situ* for use in a convergent solution phase ligation. We identified peptide *o*-aminoanilides, incorporated through synthesis on the diaminobenzoic acid (Dbz) linker, as a reasonable fit for these requirements (Blanco-Canosa & Dawson, 2008; Blanco-Canosa, Nardone, Albericio, & Dawson, 2015; Mahto, Howard, Shimko, & Ottesen, 2011; Wang et al., 2015). This moiety survives peptide ligation, desulfurization, and cleavage conditions, but can be converted to thioester via a triazole intermediate in aqueous conditions [Scheme 1] (Wang et al., 2015). In addition, since residues C-terminal to the Dbz are eliminated in conversion to the active thioester, these can be exploited to anchor the peptide segment to resin. Here, we selected a reverse NCL anchoring strategy, in which a Cys incorporated through isopeptide linkage to a Lys side chain (**11**) was exploited for ligation to a thioester resin (**12**) to immobilize our peptide segment.

Using this strategy, we developed a ligation-desulfurization scheme such that native H1.2 could be prepared from ligation of 8 synthetically accessible peptide fragments [Scheme 2]. We identified 7 Ala residues at sites with reasonable predicted ligation kinetics (Haase, Rohde, & Seitz, 2008; Johnson & Kent, 2006) spaced such that each segment was < 36 residues in length. These were grouped into two segments of four fragments each. The anchor fragment for the N-terminal segment (**11**) would be synthesized with the internal aminoanilide and C-terminal isopeptide ligation handle for immobilization on the solid phase. The anchor segment for the C-terminal segment (**18**) was synthesized with an internal HMBA linker to allow generation of the native protein C-terminus after ligation (Harris & Brimble, 2010; Hossain et al., 2009; Yu et al., 2016).

Each individual peptide was synthesized with an N-terminal thiazolidine (Thz) protection strategy for the Cys, which is commonly used for chemical control of ligation (Bang & Kent, 2004). Interestingly, we found that use of Dbz as cryptic thioester required significant finesse, as no single approach was suitable for all peptides (full discussion in Supplemental S4). For all but the N-terminal anchoring peptide, N-methyl (**1**), N-alloc (**3**), or unprotected (**2**) diaminobenzoic acid strategies were used for the generation of the peptide thioester (**10**) (Blanco-Canosa & Dawson, 2008; Blanco-Canosa et al., 2015; Mahto et al., 2011; Wang et al., 2015). In brief, we found that while N-methyldiaminobenzoic acid is compatible with microwave peptide synthesis, it 1) adds complexity to the characterization of the peptide-MeDbz intermediate due to formation of the benzimidazole derivative under acidic cleavage or HPLC conditions (Bird, Silvestri, & Dawson, 2018), and 2) often shows reduced yield of the N-acylurea derivative, mostly likely due to reduced reactivity of the N-substituted amine. The N-

(alloc)diaminobenzoic acid derivative is compatible with microwave synthesis only if piperidine treatment for Fmoc removal is carried out at room temperature; heated deprotection cycles resulted in premature N-acylurea formation and peptide release from the resin (Tsuda, Uemura, Mochizuki, Nishio, & Yoshiya, 2017). However, we observe minimal product loss (~1% loss per coupling cycle) when carrying out coupling steps using microwave heating.

With all peptides in hand, we proceeded to assess the solid phase ligation for the two major protein blocks. To generate the N-terminal block **26**, we first prepared thioester resin **12**. Rink amide PEGA resin was treated with diglycolic anhydride. This was further treated with thiophenol in presence of DIC and DMAP to obtain resin thiophenol thioester **12**. Anchor peptide **11** was added, and native chemical ligation afforded the linked product. A key step was a capping step in which the resin was treated with an excess of Cys in ligation buffer to deactivate any remaining thioester resin, to avoid potential crossreactivity in subsequent ligation steps.

Treatment with methoxylamine afforded the “ring-opened” Cys product **13** suitable for use in subsequent ligation cycles (**14-16**). For documentation of these initial syntheses, ligation progress was also monitored by RP-HPLC (Fig. 3) and SDS-PAGE (see ESI); however, these analyses consume significantly more material than mass spectrometric analysis, and are less sensitive than MALDI-TOF for detection of unreacted material. For preparative purposes, we recommend eliminating these steps and restricting analysis to MALDI-TOF MS. In all, we carried out 4 replicated assembly of the N-terminal segment **16** to assess the robustness of the approach; one with unmodified segments, and three with citrullinated Arg53 to generate **16cit**.

Typically, desulfurization is carried out as the last step in a convergent ligation scheme.

However, desulfurization is complicated with large numbers of thiols in a protein (c.f. **27**). We noted that the thiols in the N-terminal segment **16** are not required for subsequent ligation steps. We therefore explored desulfurization on the solid phase, which we hypothesized could simplify the final desulfurization step (**27** to **28**). Across the four preparations, yields after four ligation/desulfurization cycles ranged from 50-72% by weight, in sufficient purity to use directly in solution phase ligations.

For preparation of the C-terminal segment, we followed the protocols developed for the synthesis of H4 and CENP-A (Yu et al., 2016). PEGA-PL resin was used for solid support. Fmoc-Thz-Ala-Ahx-Tyr-Lys-Gly handle sequence (sufficient for detection by RP- HPLC and MALDI-TOF MS) was introduced on the rink amide PEGA resin. Immediately prior to the first ligation, the Fmoc was removed by treatment with 20% piperidine in DMF. Treatment with methoxylamine then afforded the N-terminal cysteine, to enable the attachment of H1-(Thz₁₈₉-Lys₂₁₂)-HMBA-RG-MeNbz/Nbz through native chemical ligation. After ligation, resin was washed and the cycle was repeated for the remaining rounds of SP-NCL. To assess the robustness of our approach, we repeated this synthesis 3 times for unmodified peptide segments. Across the three syntheses, the yields ranged from 77-81% by weight, and the crude product was typically sufficiently pure to be used without purification for the final solution phase ligation (Fig. 4-D).

We then prepared C-terminal segment with Ser172ph as described above. Of note, this modified variant is the one synthesis for which we assessed an alternate rapid palladium-based protocol for deprotection of the thiazolidine (Jbara, Maity, Seenaiiah, & Brik, 2016). Given the lower purity observed for this product ([Fig. 4-H] compared to [Fig. 4-D]), we recommend the

methoxyamine protocol on the solid phase. However, the crude product (73% yield) was still of sufficient purity to use without further purification.

With both large protein segments in hand, we proceeded to the final solution phase ligation. The N-terminal peptide **17** is activated by treatment of sodium nitrite to form the triazole **25**, for conversion to thioester **26** *in situ* for the following ligation. In initial ligation tests, we carried out thioester conversion of the N-terminal followed by addition of the C-terminal segment in ligation buffer (6M Gdn, 0.1M phosphate, 50mM MPAA). However, with ligation carried out at pH 7.3, we found significant hydrolysis of the triazole in competition with thioester formation to reduce our overall yield. We probed this process using the model peptide [H1.2(Lys₈₉-Gly₉₉)-Dbz-Gly-NH₂], and found minimal hydrolysis over 6 hours at pH 6. We also assessed whether the process could be simplified by pre-mixing **17** with **24** prior to activation (Fang et al., 2011; Wang et al., 2015), rather than sequential addition of **24** to converted **25**. We found no significant differences in ligation kinetics or yield for this considerably simpler experimental protocol. We therefore converged to ligation via activation of the mixed components, followed by addition of ligation buffer to reach pH ~6.5 for ligation over 4 hours at room temperature.

The ligation product **27** was dialyzed extensively to remove MPAA, which can quench phosphine-mediated desulfurization (Rohde, Schmalisch, Harpaz, Diezmann, & Seitz, 2011). The final full-length protein was purified by RP-HPLC after desulfurization. Although the ligation reaction proceeds to near completion (Fig 5B, C, D), purification yields are typically poor for

histone proteins – which is one driving factor for the development of minimal purification approaches such as solid phase NCL. A 12% isolated yield (assessed by in-gel quantitation against a commercial H1.0 standard) was obtained for the full-length H1.2 after ligation, desulfurization, and RP-HPLC purification. Ligations were characterized by SDS-PAGE (Fig. 5D lane 6) and mass spectrometry (Fig 5F, Supplemental S5.7).

The primary native function of linker histones, including variants H1.0 and H1.2, is to interact with nucleosomes to form a chromatosome, which facilitates chromatin compaction and may alter nucleosome dynamics and local protein interactions (Bernier et al., 2015; Christophorou et al., 2014; Fyodorov et al., 2018; Liao & Mizzen, 2016; Lopez et al., 2015). Recent cryo-EM structural determination provides a consensus structure for the folded core of the H1.0 protein bound to the nucleosome (Bednar et al., 2017), depicted in Fig. 6E, although the majority of the H1 sequence remains relatively unstructured. In order to assess the functionality of our synthetic proteins, we carried out electrophoretic mobility shift assays (EMSA) to compare the binding of commercial unmodified H1.0 (NEB), our synthetic unmodified H1.2, H1.2-R53cit, and H1.2-S172ph to purified mono-nucleosomes (Figure 6). In each case, we observe that the mononucleosome band begins to shift at 5-10 nM with similar changes in electrophoretic mobility indicating that the unmodified and modified synthetic H1.2 form chromatosomes similarly to recombinant H1.0. At higher concentrations (≥ 30 nM), the nucleosomes do not migrate into the gel, which is consistent with the formation of larger aggregates mediated by H1. EMSA is not expected to be sufficiently sensitive to detect quantitative differences in H1.2 function due to individual modifications. H1 binds nucleosomes at picomolar affinities (Yue, Fang, Wei, Hayes, & Lee, 2016) and has a very slow dissociation rate that is great than 10

minutes (White, Hieb, & Luger, 2016). So, in this EMSA assay, H1 binds stoichiometrically; however, these results clearly demonstrate that the synthetic proteins carry out their expected functions, and that the modifications do not reduce the ability of H1 to form a chromosome. These syntheses therefore enable future studies to determine how H1 modifications, individually or in combination, may impact nucleosome dynamics or macromolecular interactions, for example transcription factor binding to DNA sites within the linker histone-bound nucleosome (Brehove et al., 2015).

In summary, we have developed a practical method to allow the benefits of solid phase protein ligation to be extended to the total synthesis of large proteins by exploiting diaminobenzoic acid (peptide O-aminoanilides) as a cryptic thioester. This allows the traceless assembly of ~100-residue protein segments from small, synthetically accessible peptides. We demonstrated its feasibility by the preparation of unmodified, R53cit, and S172ph modified linker histone H1.2 from 8 component peptides with only 1 purification step, enabling the first characterization of homogeneously modified linker histone. We demonstrated its robustness by carrying out solid phase ligation cycles 7 times, with yields ranging from 50-81%, each time with sufficient purity to use in subsequent ligations. While we have demonstrated this for a 212-residue protein assembled from two “blocks” with only one purification step, we hypothesize that the method could logically be extended to one-pot, sequential, or convergent ligation of 3-5 “blocks” for the facile assembly of very large proteins.

Acknowledgements

This work was supported by the National Institutes of Health (GM121966). We thank Prof. Vicki Wysocki (OSU) and Jing Yan (Wysocki Lab, OSU) for carrying out the high-resolution MS in Fig. 5E,F.

References:

- Agouridas, V., El Mahdi, O., Cargoët, M., & Melnyk, O. (2017). A statistical view of protein chemical synthesis using NCL and extended methodologies. *Bioorganic and Medicinal Chemistry*, 25(18), 4938–4945. <https://doi.org/10.1016/j.bmc.2017.05.050>
- Agouridas, V., El Mahdi, O., Diemer, V., Cargoët, M., Monbaliu, J.-C. M., & Melnyk, O. (2019). Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chemical Reviews*, acs.chemrev.8b00712. <https://doi.org/10.1021/acs.chemrev.8b00712>
- Bang, D., & Kent, S. B. H. (2004). A one-pot total synthesis of crambin. *Angewandte Chemie - International Edition*, 43(19), 2534–2538. <https://doi.org/10.1002/anie.200353540>
- Bang, D., Pentelute, B. L., & Kent, S. B. H. (2006). Kinetically controlled ligation for the convergent chemical synthesis of proteins. *Angewandte Chemie - International Edition*, 45(24), 3985–3988. <https://doi.org/10.1109/ECTC.2014.6897331>
- Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A. R., Papai, G., Reymer, A., ... Dimitrov, S. (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. *Molecular Cell*, 66(3), 384-397.e8. <https://doi.org/10.1016/j.molcel.2017.04.012>
- Bernier, M., Luo, Y., Nwokelo, K. C., Goodwin, M., Dreher, S. J., Zhang, P., ... Poirier, M. G. (2015). Linker histone H1 and H3K56 acetylation are antagonistic regulators of nucleosome dynamics. *Nature Communications*, 6, 1–10. <https://doi.org/10.1038/ncomms10152>
- Bird, M. J., Silvestri, A. P., & Dawson, P. E. (2018). Expedient on-resin synthesis of peptidic benzimidazoles. *Bioorganic and Medicinal Chemistry Letters*, 28(16), 2679–2681. <https://doi.org/10.1016/j.bmcl.2018.04.062>
- Blanco-Canosa, J. B., & Dawson, P. E. (2008). An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angewandte Chemie - International Edition*, 47(36), 6851–6855. <https://doi.org/10.1002/anie.200705471>
- Blanco-Canosa, J. B., Nardone, B., Albericio, F., & Dawson, P. E. (2015). Chemical Protein Synthesis Using a Second-Generation N-Acylurea Linker for the Preparation of Peptide-Thioester Precursors. *Journal of the American Chemical Society*, 137(22), 7197–7209.

<https://doi.org/10.1021/jacs.5b03504>

- Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., ... Poirier, M. G. (2015). Histone core phosphorylation regulates DNA accessibility. *Journal of Biological Chemistry*, 290(37), 22612–22621. <https://doi.org/10.1074/jbc.M115.661363>
- Brik, A., Keinan, E., & Dawson, P. E. (2000). Protein synthesis by solid-phase chemical ligation using a safety catch linker. *Journal of Organic Chemistry*, 65(12), 3829–3835. <https://doi.org/10.1021/jo000346s>
- Canne, L. E., Botti, P., Simon, R. J., Chen, Y., Dennis, E. A., & Kent, S. B. H. (1999). Chemical protein synthesis by solid phase ligation of unprotected peptide segments. *Journal of the American Chemical Society*, 121(38), 8720–8727. <https://doi.org/10.1021/ja9836287>
- Christophorou, M. A., Castelo-Branco, G., Halley-Stott, R. P., Oliveira, C. S., Loos, R., Radziskeuskaya, A., ... Kouzarides, T. (2014). Citrullination regulates pluripotency and histone H1 binding to chromatin. *Nature*, 507(7490), 104–108. <https://doi.org/10.1038/nature12942>
- Chuh, K. N., Batt, A. R., & Pratt, M. R. (2016). Chemical Methods for Encoding and Decoding of Posttranslational Modifications. *Cell Chemical Biology*, 23(1), 86–107. <https://doi.org/10.1016/j.chembiol.2015.11.006>
- Dadová, J., Galan, S. R., & Davis, B. G. (2018). Synthesis of modified proteins via functionalization of dehydroalanine. *Current Opinion in Chemical Biology*, 46, 71–81. <https://doi.org/10.1016/j.cbpa.2018.05.022>
- Dawson, P. E., Muir, T. W., & Kent, S. B. H. (1994). Synthesis of Proteins by Native Chemical Ligation. *Science*, 266(5186), 776–779.
- Durek, T., Torbeev, V. Y., & Kent, S. B. H. (2007). Convergent chemical synthesis and high-resolution x-ray structure of human lysozyme. *Proceedings of the National Academy of Sciences of the United States of America*, 104(12), 4846–4851. <https://doi.org/10.1073/pnas.0610630104>
- Fang, G. M., Li, Y. M., Shen, F., Huang, Y. C., Li, J. Bin, Lin, Y., ... Liu, L. (2011). Protein chemical synthesis by ligation of peptide hydrazides. *Angewandte Chemie - International Edition*, 50(33), 7645–7649. <https://doi.org/10.1002/anie.201100996>
- Fang, G. M., Wang, J. X., & Liu, L. (2012). Convergent chemical synthesis of proteins by ligation of peptide hydrazides. *Angewandte Chemie - International Edition*, 51(41), 10347–10350. <https://doi.org/10.1002/anie.201203843>
- Fierz, B., & Muir, T. W. (2012). Chromatin as an expansive canvas for chemical biology. *Nature Chemical Biology*, 8(5), 417–427. <https://doi.org/10.1038/nchembio.938>
- Fyodorov, D. V., Zhou, B. R., Skoultchi, A. I., & Bai, Y. (2018). Emerging roles of linker histones in regulating chromatin structure and function. *Nature Reviews Molecular Cell Biology*, 19(3), 192–206. <https://doi.org/10.1038/nrm.2017.94>

- Haase, C., Rohde, H., & Seitz, O. (2008). Native chemical ligation at valine. *Angewandte Chemie - International Edition*, 47(36), 6807–6810. <https://doi.org/10.1002/anie.200801590>
- Harris, P. W. R., & Brimble, M. A. (2010). Toward the total chemical synthesis of the cancer protein NY-ESO-1. *Biopolymers*, 94(4), 542–550. <https://doi.org/10.1002/bip.21351>
- Hossain, M. A., Belgi, A., Lin, F., Zhang, S., Shabanpoor, F., Chan, L., ... Wade, J. D. (2009). Use of a temporary “solubilizing” peptide tag for the Fmoc solid-phase synthesis of human insulin glargine via use of regioselective disulfide bond formation. *Bioconjugate Chemistry*, 20(7), 1390–1396. <https://doi.org/10.1021/bc900181a>
- Howard, C. J., Yu, R. R., Gardner, M. L., Shimko, J. C., & Ottesen, J. J. (2015). Chemical and biological tools for the preparation of modified histone proteins. *Topics in Current Chemistry*, 363, 193–226. https://doi.org/10.1007/128_2015_629
- Izzo, A., & Schneider, R. (2016). The role of linker histone H1 modifications in the regulation of gene expression and chromatin dynamics. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1859(3), 486–495. <https://doi.org/10.1016/j.bbagr.2015.09.003>
- Jbara, M., Maity, S. K., Seenaiah, M., & Brik, A. (2016). Palladium Mediated Rapid Deprotection of N-Terminal Cysteine under Native Chemical Ligation Conditions for the Efficient Preparation of Synthetically Challenging Proteins. *Journal of the American Chemical Society*, 138(15), 5069–5075. <https://doi.org/10.1021/jacs.5b13580>
- Jbara, M., Seenaiah, M., & Brik, A. (2014). Solid phase chemical ligation employing a rink amide linker for the synthesis of histone H2B protein. *Chemical Communications*, 50(83), 12534–12537. <https://doi.org/10.1039/c4cc06499b>
- Jin, K., Li, T., Chow, H. Y., Liu, H., & Li, X. (2017). P–B Desulfurization: An Enabling Method for Protein Chemical Synthesis and Site-Specific Deuteration. *Angewandte Chemie - International Edition*, 56(46), 14607–14611. <https://doi.org/10.1002/anie.201709097>
- Johnson, E. C. B., & Kent, S. B. H. (2006). *J.Am.Chem.Soc.2006, Kent, chemical ligation, thiol catalyst.pdf*. (6), 6640–6646.
- Kamo, N., Hayashi, G., & Okamoto, A. (2018). Triple Function of 4-Mercaptophenylacetic Acid Promotes One-Pot Multiple Peptide Ligation. *Angewandte Chemie - International Edition*, 8904, 16533–16537. <https://doi.org/10.1002/anie.201809765>
- Kent, S. B. H. (2009). Total chemical synthesis of proteins. *Chemical Society Reviews*, 38(2), 338–351. <https://doi.org/10.1039/b700141j>
- Li, J., Li, Y., Liu, L., Li, Y., He, Q., & Li, H. (2014). One-pot native chemical ligation of peptide hydrazides enables total synthesis of modified histones. *Organic and Biomolecular Chemistry*, 12(29), 5435–5441. <https://doi.org/10.1039/c4ob00715h>
- Liao, R., & Mizzen, C. A. (2016). Interphase H1 phosphorylation: Regulation and functions in chromatin. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1859(3), 476–485. <https://doi.org/10.1016/j.bbagr.2015.11.012>

- Lopez, R., Sarg, B., Lindner, H., Bartolom??, S., Ponte, I., Suau, P., & Roque, A. (2015). Linker histone partial phosphorylation: Effects on secondary structure and chromatin condensation. *Nucleic Acids Research*, 43(9), 4463–4476. <https://doi.org/10.1093/nar/gkv304>
- Luger, K., Rechsteiner, T. J., Richmond, T. J., Ollivier, N., Vicogne, J., Vallin, A., ... Brown, D. T. (2018). Studies of biochemical crosstalk in chromatin with semisynthetic histones. *Chemical Society Reviews*, 19(1), 2296–2349. <https://doi.org/10.1038/nrm.2017.94>
- Mahto, S. K., Howard, C. J., Shimko, J. C., & Ottesen, J. J. (2011). A reversible protection strategy to improve Fmoc-SPPS of peptide thioesters by the N-acylurea approach. *ChemBioChem*, 12(16), 2488–2494. <https://doi.org/10.1002/cbic.201100472>
- Müller, M. M., & Muir, T. W. (2015). Histones: At the crossroads of peptide and protein chemistry. *Chemical Reviews*, 115(6), 2296–2349. <https://doi.org/10.1021/cr5003529>
- Ollivier, N., Desmet, R., Drobecq, H., Blanpain, A., Boll, E., Leclercq, B., ... Melnyk, O. (2017). A simple and traceless solid phase method simplifies the assembly of large peptides and the access to challenging proteins. *Chemical Science*, 8(8), 5362–5370. <https://doi.org/10.1039/c7sc01912b>
- Qi, Y. K., He, Q. Q., Ai, H. S., Guo, J., & Li, J. Bin. (2017). The convergent chemical synthesis of histone H3 protein for site-specific acetylation at Lys56 and ubiquitination at Lys122. *Chemical Communications*, 53(29), 4148–4151. <https://doi.org/10.1039/c7cc01721a>
- Raibaut, L., Ollivier, N., & Melnyk, O. (2012). Sequential native peptide ligation strategies for total chemical protein synthesis. *Chemical Society Reviews*, 41(21), 7001–7015. <https://doi.org/10.1039/c2cs35147a>
- Rohde, H., Schmalisch, J., Harpaz, Z., Diezmann, F., & Seitz, O. (2011). Ascorbate as an Alternative to Thiol Additives in Native Chemical Ligation. *ChemBioChem*, 12(9), 1396–1400. <https://doi.org/10.1002/cbic.201100179>
- Shimko, J. C., North, J. A., Bruns, A. N., Poirier, M. G., & Ottesen, J. J. (2011). Preparation of fully synthetic histone H3 reveals that acetyl-lysine 56 facilitates protein binding within nucleosomes. *Journal of Molecular Biology*, 408(2), 187–204. <https://doi.org/10.1016/j.jmb.2011.01.003>
- Simon, M. D., Chu, F., Racki, L. R., de la Cruz, C. C., Burlingame, A. L., Panning, B., ... Shokat, K. M. (2007). The Site-Specific Installation of Methyl-Lysine Analogs into Recombinant Histones. *Cell*, 128(5), 1003–1012. <https://doi.org/10.1016/j.cell.2006.12.041>
- Tang, S., Si, Y. Y., Wang, Z. P., Mei, K. R., Chen, X., Cheng, J. Y., ... Liu, L. (2015). An efficient one-pot four-segment condensation method for protein chemical synthesis. *Angewandte Chemie - International Edition*, 54(19), 5713–5717. <https://doi.org/10.1002/anie.201500051>
- Tsuda, S., Uemura, T., Mochizuki, M., Nishio, H., & Yoshiya, T. (2017). Preparation of Peptide o-Aminoanilides Using a Modified Dawson’s Linker for Microwave-Assisted Peptide Synthesis.

Synlett, 28(15), 1956–1960. <https://doi.org/10.1055/s-0036-1588862>

Wan, Q., & Danishefsky, S. J. (2007). Free-radical-based, specific desulfurization of cysteine: A powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angewandte Chemie - International Edition*, 46(48), 9248–9252.

<https://doi.org/10.1002/anie.200704195>

Wang, J. X., Fang, G. M., He, Y., Qu, D. L., Yu, M., Hong, Z. Y., & Liu, L. (2015). Peptide o-aminoanilides as crypto-thioesters for protein chemical synthesis. *Angewandte Chemie - International Edition*, 54(7), 2194–2198. <https://doi.org/10.1002/anie.201408078>

White, A. E., Hieb, A. R., & Luger, K. (2016). A quantitative investigation of linker histone interactions with nucleosomes and chromatin. *Scientific Reports*, 6, 1–14. <https://doi.org/10.1038/srep19122>

Yan, L. Z., & Dawson, P. E. (2001). Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *Journal of the American Chemical Society*, 123(4), 526–533. <https://doi.org/10.1021/ja003265m>

Yu, R. R., Mahto, S. K., Justus, K., Alexander, M. M., Howard, C. J., & Ottesen, J. J. (2016). Hybrid phase ligation for efficient synthesis of histone proteins. *Organic and Biomolecular Chemistry*, 14(9), 2603–2607. <https://doi.org/10.1039/c5ob02195b>

Yue, H., Fang, H., Wei, S., Hayes, J. J., & Lee, T. H. (2016). Single-Molecule Studies of the Linker Histone H1 Binding to DNA and the Nucleosome. *Biochemistry*, 55(14), 2069–2077. <https://doi.org/10.1021/acs.biochem.5b01247>

Zuo, C., Zhang, B., Yan, B., & Zheng, J.-S. (2018). One-pot multi-segment condensation strategies for chemical protein synthesis. *Organic & Biomolecular Chemistry*. <https://doi.org/10.1039/C8OB02610F>

Figure 1. Different modalities for diaminobenzoic acid linkers as cryptic thioesters in peptide synthesis and ligation. Decision points for synthesis include the protection scheme for the linker (1-3) and the activation approach (7,9), all of which can result in generation of an active thioester 10. In general, on-resin generation of the N-acylurea derivative 7 is most suitable for generation of standard peptide components. However, in situ generation of the triazole 9 in aqueous solution can be exploited in our convergent ligation scheme.

Figure 2. Convergent hybrid ligation scheme for preparation of linker histone H1.2.

Figure 3. Analytical RP-HPLC analysis of crude products for representative ligations on the solid support for each intermediate product required to generate unmodified N-terminal segment **17** and citrullinated **17cit**. Each product is cleaved at the Rink linker and includes the Dbz linker required for activation to generate thioester **26** or **26cit**. All peptide identities were confirmed by MALDI-TOF MS (see supplemental information). Of note, the side peak in panel C was identified as a thiol adduct that was removed by treatment with reducing agent. Products were obtained in 72% (D) and 50% (H) yields by weight.

Figure 4. Analytical RP-HPLC analysis of crude products for representative ligations on the solid support for each intermediate product required to generate unmodified C-terminal segment **23** and phosphorylated **23ph**. Each product is cleaved at the HMBA linker (light blue in Fig. 2). All product identities were confirmed by MALDI-TOF MS (see supplemental information). Ligation cycles to generate E-H assessed an alternate deprotection scheme that was less successful. Products were obtained in 77% (D) and 73% (H) yields by weight.

Figure 5. Analysis of solution phase ligation to generate full-length H1.2 **28**. RP-HPLC of A) pre-mixed 17cit and 24 prior to activation and ligation, B) 4 hours timepoint for ligation, C) ligation mixture after desulfurization. Peak identities were confirmed by MALDI-TOF MS (see supplemental information). D) SDS-PAGE analysis of ligation; Lanes 1: MWM; Lane 2, mixture prior to activation; Lane 3, 0 hour prior to pH shift to initiate ligation; Lane 4, 30 minutes timepoint; Lane 5, 2 hours timepoint for ligation; Lane 6, H1.2cit purified by RP-HPLC after desulfurization, taken from a separate gel. E) Calculated isotopic distribution for H1.2 **28**, F) High resolution deconvoluted ESI-MS characterization of purified, synthetic H1.2 **28**, unmodified, taken on a Thermo Q Exactive EMR Orbitrap instrument.

Figure 6. Characterization of H1 binding to nucleosomes. Left: EMSA showing A) H1.0, B) H1.2 **28**, C) H1.2-R53cit (**28cit**), D) H1.2-S172ph **28ph** binding to nucleosomes. Nucleosome band begins to shift at ~5nM H1 and is fully shifted at ~10nM H1. At 30 nM H1, nucleosomes do not migrate into the gel, demonstrating that they are forming H1-mediated aggregates. Right: Structured domain of H1 bound to nucleosome, from 5NL0.pdb^[50]