Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supporting Information

Ruixuan R. Yu, Santosh K. Mahto, Kurt Justus, Mallory Alexander, Cecil J. Howard, Jennifer J. Ottesen
Department of Chemistry and Biochemistry, The Ohio State University

Table of Contents

1 General

1.1 Materials

Rink Amide MBHA resin LL (100-200 mesh, 0.36 mmol/g loading) was purchased from Novabiochem. PL-PEGA resin (300-500 μm, 0.2 mmol/g) was purchased from Varian. *N,N*-Dimethylformamide C_3H_7NO (DMF), dichloromethane CH_2Cl_2 (DCM), acetonitrile C_2H_3N (ACN), and diethyl ether $(C_2H_5)_2O$ were purchased from Fisher Scientific. *N*-methyl pyrrolidone C_5H_9NO (NMP) was purchased from AGTC Bioproducts. Piperidine $C_5H_{11}N$, 4-nitrophenyl chloroformate $ClCO_2C_6H_4NO_2$ (NPCF), 4-mercaptophenyl acetic acid $HSC_6H_4CH_2CO_2H$ (MPAA), $C_9H_{15}O_6P$ Tris(2-carboxylmethyl)phosphine (TCEP), N,N-diisopropylethylamine $C_8H_{19}N$ (DIEA), Phenyl silane C_6H_8Si, and Tetrakis(triphenylphosphine)palladium(0) ($Pd(PPh_3)_4$) were purchased from Sigma Aldrich. Fmoc protected amino acids were purchased from AAPPTec and Novabiochem. Fmoc-6-Aminohexanoic acid (Fmoc-Ahx-OH), Fmoc-L-norleucine (Fmoc-Nle-OH), and 4-dimethylaminopyridine $C_7H_{10}N_2$ (DMAP) were purchased from Novabiochem. (2-(7-aza-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate $C_{10}H_{15}F_6N_6OP$ (HATU), 2-(6-chloro-1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate $C_{11}H_{15}ClF_6N_5OP$ (HCTU), and 1-hydroxy-6-chloro-benzotriazole $C_6H_4ClN_3O$ (6-Cl-HOBt) were purchased from AAPPTec. Acetic anhydride and sodium 2-sulfanylethane sulfonate $C_2H_5NaO_3S_2$ (MESNA) was purchased from Fluka Analytical. *N,N'*-diisopropylcarbodiimide $C_7H_{14}N_2$ (DIC) was purchased from Alfa Aesar. VA-044-US $C_{12}H_{22}N_6.2HCl$ was purchased from Wako Chemicals. Ultra-pure guanidine-HCl CH_6ClN_3 (GuHCl) was purchased from MP Biomedicals. Di-Fmoc-3,4-diaminobenzoic acid $C_{37}H_{28}N_2O_6$ was purchased from Anaspec. Allyl chloroformate $C_4H_5ClO_2$ was purchased from Acros Organics. Triisopropyl silane $C_9H_{22}Si$ (TIS) was purchased from GFS Chemicals. Boc-(R)-5,5-dimethyl-1,3-thiazolidine-4-caboxylic acid (Boc-dmThz-OH) was purchased from Chem-Implex International. α-Cyano-4-hydroxycinnamic acid $C_{10}H_7NO_3$ (HCCA) was purchased from Bruker Daltonics.

1.2 RP-HPLC

Analytical reverse phase HPLC (RP-HPLC) was run on a Shimadzu or Waters instrument using an analytical column (Supelco C18 15 cm × 4.6 mm × 5 μm, flow rate 0.9 mL/min). Preparative RP-HPLC was run on a Waters instrument using a semi-preparative column (Supelco C18 25 cm × 10 mm × 10 μm, flow rate 5 mL/min), or a preparative column (Supelco C18 25 cm × 21.2 mm × 10 μm, flow rate 18 mL/min). Buffer A was 0.1 % TFA in water, and Buffer B was 1:9 water:acetonitrile, 0.1% TFA,. Gradients in the ESI are reported as the percentage Buffer B in Buffer A.

1.3 Mass spectrometry

Peptide masses were confirmed by MALDI-TOF-MS (Bruker Daltonics Microflex) using flexControl 3.3 and flexAnalysis 3.3 sofwares. HCCA was used for the matrix and Peptide Calibration Standard II (Bruker) was used for calibration. All expected masses are reported as the mass average.

1.4 UV Spectroscopy

UV spectroscopy was carried out on a NanoDrop 1000 (Thermo Scientific).

2 Peptide Preparation

2.1 Synthesis of 3-Fmoc-Dbz-OH

3,4-Diaminobenzoic acid (1 g, 6.5 mmol) was resuspended in 125 mL 1:1 CH_3CN:$NaHCO_3$. Reaction was initiated by the dropwise addition of Fmoc-OSu (2.4 g, 7.1 mmol) in 15 mL 1:1 CH_3CN:$NaHCO_3$ and proceeded for 2 h. HCl was added to a final pH of 1.0, and the mixture was filtered. Filtrate was

dissolved in 4 mL DMSO, precipitated with acidified reaction buffer, washed extensively, and dried under vacuum to yield a light gray product. Product identity and purity was validated by NMR spectroscopy.[1]

2.2 Peptide Synthesis

Peptides were synthesized using standard Fmoc-N-α protection strategies either manually (for short sequences) or on an AAPTec APEX 396 automated synthesizer, using HCTU activation and 20% piperidine deprotection, except where otherwise specified. The N-terminal residue of each peptide synthesized on Dbz or Dbz(Alloc) resin was coupled as the Boc derivative to allow for conversion to the N-acylurea, or protected as the α-acetyl derivative in the case of unmodified H4. Peptides H4-A, H4-B, and H4-C were synthesized on Rink-amide resin (Novabiochem) to which 3-Fmoc-Dbz-OH was added manually, followed by Alloc protection (overnight treatment of the resin with 250 mM allyl chloroformate in DCM with 1 eq DIEA to resin loading).

Amino acid coupling directly onto 4-Alloc-Dbz resin was accomplished using HATU activation with 1 hour coupling time, followed by acetyl capping using a cocktail containing 300 mM 6-Cl-HOBt and 300 mM Acetic anhydride in 1:9 DCM:DMF. In the H4-PepC and CpA-5 peptides, the amino acid following the HMBA linker was double-coupled as the symmetric anhydride: 10 eq Fmoc-AA-OH was dissolved in DCM. 5 equivalent of DIC was added, and incubated on ice for 30 minutes. The filtrate and 0.1 eq DMAP were added to DMF-swollen resin. The reaction was allowed to proceed for 1 hour.

Dbz(Alloc) deprotection is accomplished on resin by treatment with 0.35 eq $Pd(PPh_3)_4$ and 20 eq Phenyl silane in DCM for 45 minutes. For most peptides, Dbz is converted into Nbz by treating the resin 5 equivalent of NPCF in DCM for 30 minutes, followed by 0.5 M DIEA in DMF for 15 minutes according to literature protocols[2].

Peptides H4-PepA, H4-PepH, and CpA-2 generated multiple alternate products on Nbz conversion in DCM. However, conversion in DMF dried over molecular sieves cleanly generated formylated Nbz, as observed by the Brik Laboratory[3]. For these peptides, the resin was swollen in dry DMF, and solid NPCF was added directly to the suspension to 50 mM. This is discussed in more detail in section 3.1.

All peptides were cleaved in 95:2.5:2.5 TFA:TIS:water for 2 hours. TFA was reduced with a stream of nitrogen, and peptides were precipitated and washed with cold Et_2O, then resuspended in water and lyophilized before analysis and purification.

2.3 Base Resin for SP-NCL

All base resins were prepared manually using 0.2 mmol/g PL-PEGA resin. In order to reduce steric crowding, the loading of the resin was reduced 10-fold to 0.02 mmol/g by coupling of the resin with a mixture of 1:9 Fmoc-Gly-OH:Boc-Gly-OH. Standard coupling using Fmoc-SPPS was performed after the loading cut. The sequences of the base resins used are listed below along with their theoretical loading in methanol. Gly where the loading cut is performed is indicated in bold. The base resins are stored in methanol, and substitution assessed based on the volume for methanol-swollen beads.

H4 SP-NCL: Fmoc-Thz-Ala-Rink-**Gly**-Gly –PEGA, 0.00166 mmol/mL
CpA SP-NCL: Fmoc-Thz-Ala-Gly-Gly-Rink-**Gly**-PEGA 0.00166 mmol/mL
H4 and CpA Hybrid NCL: Fmoc-Thz-Ala-Ahx-Rink-**Gly**-PEGA 0.00133 mmol/mL

3 H4 Peptides

3.1 H4-A (acSer$_1$-Leu$_{37}$)-Nbz

ac-SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRL-Nbz

H4-A peptide was synthesized with 0.05 mmol of mFmoc-Dbz(Alloc) resin. The N-terminus was acetylated to mimic the constitutive acetylation of eukaryotic H4. Alloc deprotection was followed by Nbz conversion in DCM, which resulted in several additional products (Fig. S1 left), although sufficient desired product to proceed with purification. For further discussion, see section 3.2. The peptide was purified on semi-preparative RP-HPLC using a gradient of 15-30 % Buffer B over 40 minutes.

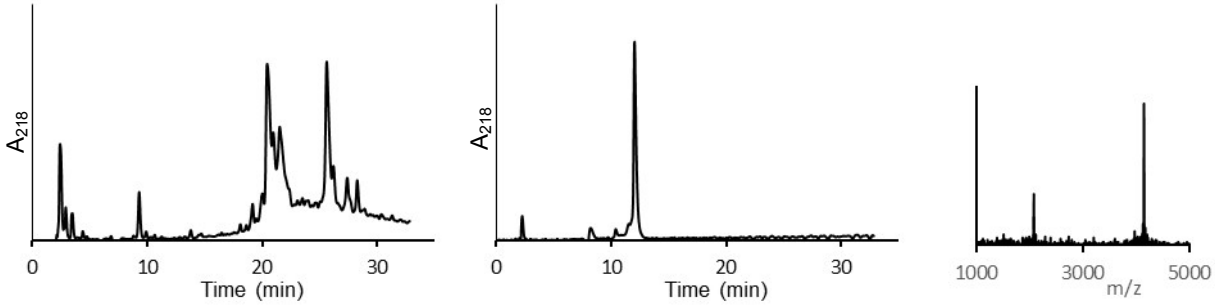


Figure S1: RP-HPLC of crude (left 15-30 % B) and purified (right 20 -35 % B) H4-A-Nbz. MALDI-TOF (right) of purified H4-A-Nbz: $[M + H]^+$ observed: *m/z* 4137, expected: *m/z* 4138.

3.2 H4-A-K5ac,K12ac (Ser$_1$-Leu$_{37}$)-Nbz(formyl)-Arg

SGRG**Kac**GGKGLG**Kac**GGAKRHRKVLRDNIQGITKPAIRRL-Nbz-R

H4-A-K5ac,K12ac peptide was synthesized with 0.04 mmol of mFmoc-Dbz(Alloc)-R resin. The N-terminal Ser was added as Fmoc-Ser(tBu)-OH. After Alloc deprotection, Nbz conversion was carried out in dry DMF to generate the Nbz(formyl) derivative, and the N-terminal Fmoc was removed by treatment with 1% DBU in DMF for three minutes. The peptide was purified on semi-preparative RP-HPLC using a gradient of 10-30 % Buffer B over 40 minutes.

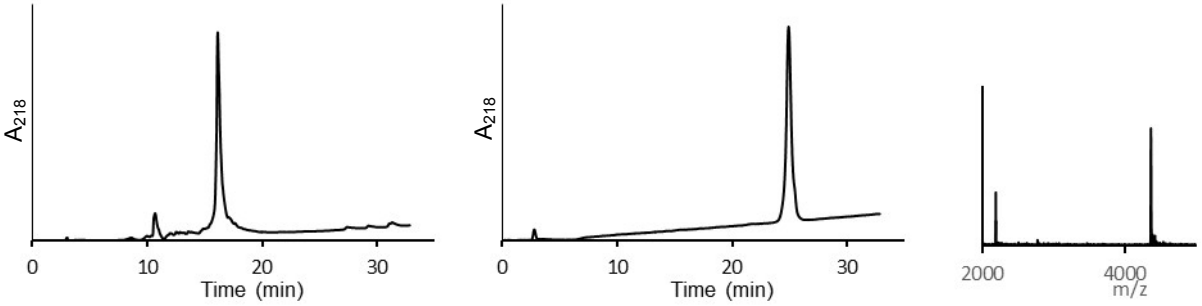


Figure S2: RP-HPLC of crude (left 0-73 % B) and purified (middle 10-30 % B) H4-A-K5ac,K12ac-Nbz(formyl)-R. MALDI-TOF (right) of purified H4-A-K5ac,K12ac-Nbz(formyl)-R: $[M + H]^+$ observed: *m/z* 4366, expected: *m/z* 4365.

3.3 Synthesis of H4-B (Thz_{38}-Gly_{56})-Nbz

Thz-RRGGVKRISGLIYEETRG-Nbz

H4-B peptide was synthesized with 0.05 mmol of mFmoc-Dbz(Alloc) resin. The peptide was purified on preparative RP-HPLC using a gradient of 20-35 % Buffer B over 40 minutes.

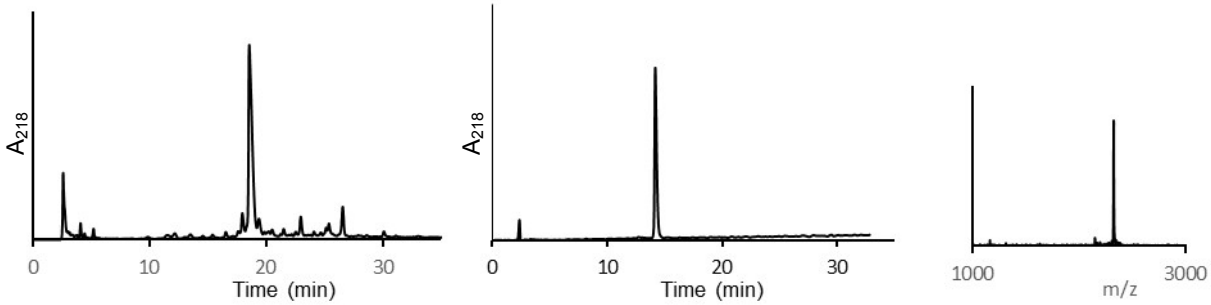


Figure S3: RP-HPLC of crude (left 15-35 % B) and purified (middle 20-35 % B) H4-B-Nbz. MALDI-TOF (right) of purified H4-B-Nbz. $[M + H]^+$ observed: *m/z* 2322, expected: *m/z* 2321.

3.4 Synthesis of H4-H (Pen_{57}-H_{75})-Nbz(formyl)-Arg

dmThz-LKVFLENVIRDAVTYTEH-Nbz(formyl)-R

Synthesis of H4-H was complicated by racemization of the C-terminal His on the diaminobenzoic acid linker. The amine of Dbz(Alloc) resin is relatively deactivated, requiring stringent coupling conditions for full loading of the first amino acid. In the case of Fmoc-His(Trt)-OH, which is notably susceptible to racemization, we suspected that racemization might be occurring under previously used HATU coupling conditions.[1] To assess this, we cleaved and analyzed the product (RP-HPLC over a 0-20% B gradient) at several early points in the synthesis of the H4-H peptide, and found that while we observed only one peak for Fmoc-His-Dbz(Alloc) and Fmoc-His-Dbz(Alloc)-R peptides, multiple peaks with the same mass were visible for Glu-His-Dbz(Alloc) and Glu-His-Dbz(Alloc)-R peptides. To confirm the identity of this peak, we generated a racemization standard by coupling Fmoc-His(Trt)-OH in the presence of significant DMAP.[4] We considered coupling conditions along two axes: percentage coupling (observable as Glu-Dbz-Arg or Glu-Dbz(Alloc)-Arg) and percentage racemization. No acceptable conditions could be found for coupling directly to Dbz(Alloc) resin. However, even unprotected Dbz resin proved problematic; acceptable racemization (<1 %) was achieved using 1.1:1:1 AA:6-Cl-HOBt :DIC in the absence of DIEA.

	Base Resin	Pre-activation / coupling time	Reagents	Coupling (%)	Racemization (%)
Standard	Dbz(Alloc)	30 min / 60 min	22:10:1 Fmoc-His(Trt)-OH:DIC:DMAP		
1	Dbz(Alloc)	3 min / 60 min	0.23 M His, 0.23 M HATU, 0.42 M DIEA	N/A	11.6
2	Dbz(Alloc)-R	None / 30 min (double)	0.28 M His, 0.25 M HATU, 0.04 M DIEA	82	4.8
3	Dbz(Alloc)-R	None / 30 min (triple)	0.28 M His, 0.25 M HATU, 0.04 M DIEA	88	6.4
4	Dbz(Alloc)-R	None / 30 min (double)	0.27 M His, 0.25 M 6-Cl-HOBt, 0.125 M DIC	25	2.3
5	Dbz-R	None / 45 min	0.11 M His, 0.1 M HCTU, 0.16 M DIEA	98	6.5
6	Dbz-R	None / 30 min	0.11 M His, 0.1 M 6-Cl-HOBt, 0.1 M DIC	98	<1

Table S1: Histidine loading conditions. Percent coupling for Dbz and Dbz(Alloc) resin is not available because the Dbz(Alloc) and Dbz species are soluble in the ether wash during peptide precipitation. Therefore unreacted species could not by observed without the Arg tag.

We used condition 6 from table for preparation of H4-H peptide on Dbz-R resin. Nbz conversion was carried out in dry DMF. The peptide was then cleaved from the resin using TFA. The peptide was purified on preparative HPLC using a gradient of 35-50 % Buffer B over 40 minutes.

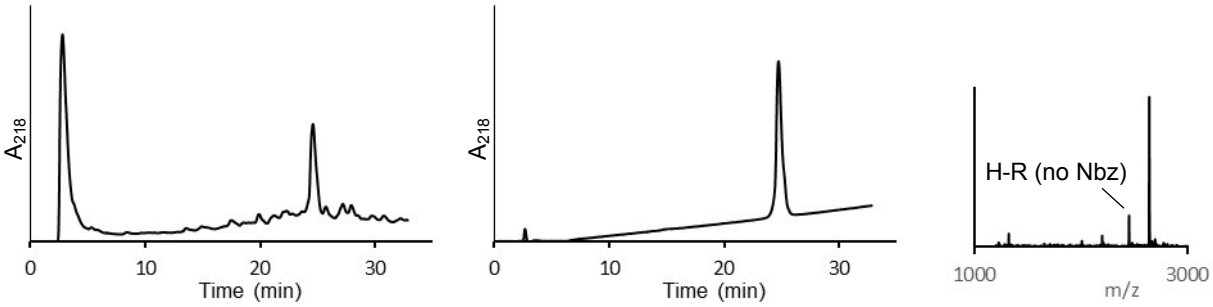


Figure S4: RP-HPLC of crude (left 30-60 % B) and purified (middle 30-60 % B) H4-H-Nbz-R. MALDI-TOF (right) of purified H4-H-Nbz-R: $[M + H]^+$ observed: *m/z* 2636, expected: *m/z* 2634. Note the species lacking Nbz was a result of incomplete coupling of Dbz when preparing the base resin. Because the species cannot convert to thioester without the Nbz linker, it does not interfere with the ligation.

3.5 Synthesis of H4-C (Thz_{76}-Gly_{102}-HMBA-Arg-Gly)-Nbz Met84Nle

Thz-KRKTVTA**Nle**DVVYALKRQGRTLYGFGG-**HMBA**-RG-Nbz

H4-C peptide was synthesized with 0.05 mmol of G-HMBA-RG resin. Nle was used in place of Met in order to avoid oxidation. The peptide was purified on preparative RP-HPLC using a gradient of 30-45 % Buffer B over 40 minutes.

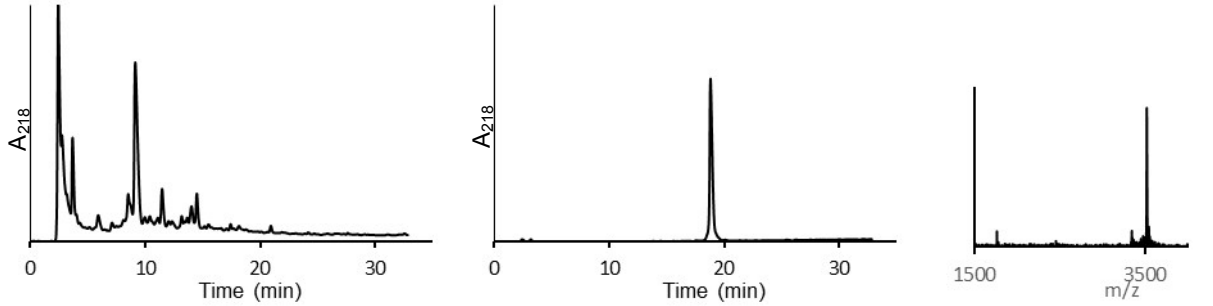


Figure S5: RP-HPLC of crude (left 30-45 % B) and purified (middle 25-40 % B) H4-C-Nbz. MALDI-TOF (right) of purified H4-C-Nbz. H4-C-Nbz: $[M + H]^+$ observed: *m/z* 3517, expected: *m/z* 3520.

3.6 Synthesis of H4-C-K91ac (Thz_{76}-Gly_{102}-HMBA-Arg-Gly)-Nbz

Thz-KRKTVTA**Nle**DVVYAL**Kac**RQGRTLYGFGG-**HMBA**-RG-Nbz

H4-C-K91ac peptide was synthesized with 0.05 mmol of Gly-HMBA-Arg-Gly-Dbz(Alloc) resin. The peptide was purified by RP-HPLC using a gradient of 30-45% Buffer B over 40 minutes.

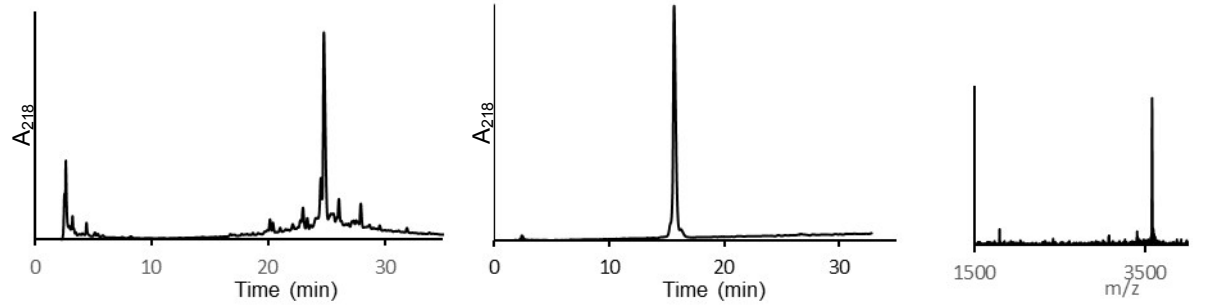


Figure S6: RP-HPLC of crude (left 0-73 % B) and purified (middle 30-40 % B) H4-C-K91ac-Nbz. MALDI-TOF (right) of purified H4-C-K91ac-Nbz: $[M + H]^+$ observed: *m/z* 3579, expected: *m/z* 3581.

4 Preparation of CENP-A (CpA) Peptides

4.1 Synthesis of CpA-1 (Gly_2-Gly_{34})-Nbz

GPRRRSRKPEAPRRRSPSPTPTPGPSRRGPSLG-Nbz

CpA1 peptide was synthesized with 0.05 mmol of mFmoc-Dbz(Alloc) resin. The N-terminal Gly was added as Fmoc-Gly-OH. After alloc deprotection and Nbz conversion, Fmoc was removed using 1% DBU before cleavage from the resin with TFA. The peptide was purified on preparative RP-HPLC using a gradient of 12-25% Buffer B over 40 minutes.

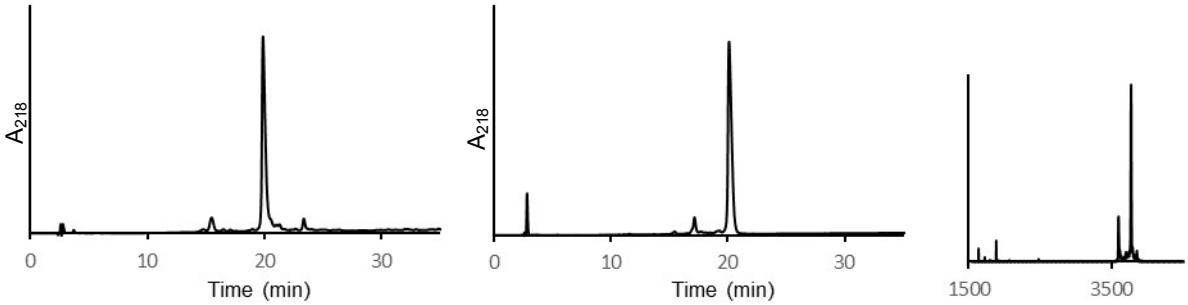


Figure S7: RP-HPLC of crude (left 12-30 % B) and purified (middle 12-30 % B) CpA-1-Nbz. MALDI-TOF (right) of purified CpA-1-Nbz. CpA-1-Nbz: $[M + H]^+$ observed: *m/z* 3765, expected: *m/z* 3765.

4.2 Synthesis of CpA-2 (Thz_{35}-Leu_{70})-Nbz (formyl)

Thz-SSHQHSRRRQGWLKEIRKLQKSTHLLIRKLPFSRL-Nbz(formyl)

CpA2 peptide was synthesized with 0.05 mmol of mFmoc-Dbz(Alloc) resin with the automated synthesizer. The N-terminal residue was added as Boc-Thz-OH. Alloc was removed. NPCF chloroformate addition was first carried out in dry DCM followed by treatment with DIEA/DMF to generate the Nbz, but with significant additional side products resulting in minimal yield (Fig. S8A). Several reaction conditions were assessed, but NPCF addition in DMF dried with molecular sieves resulted in the highest yields (Fig. S8B). The peptide was purified by RP-HPLC using a gradient of 30-60% Buffer B over 40 minutes.

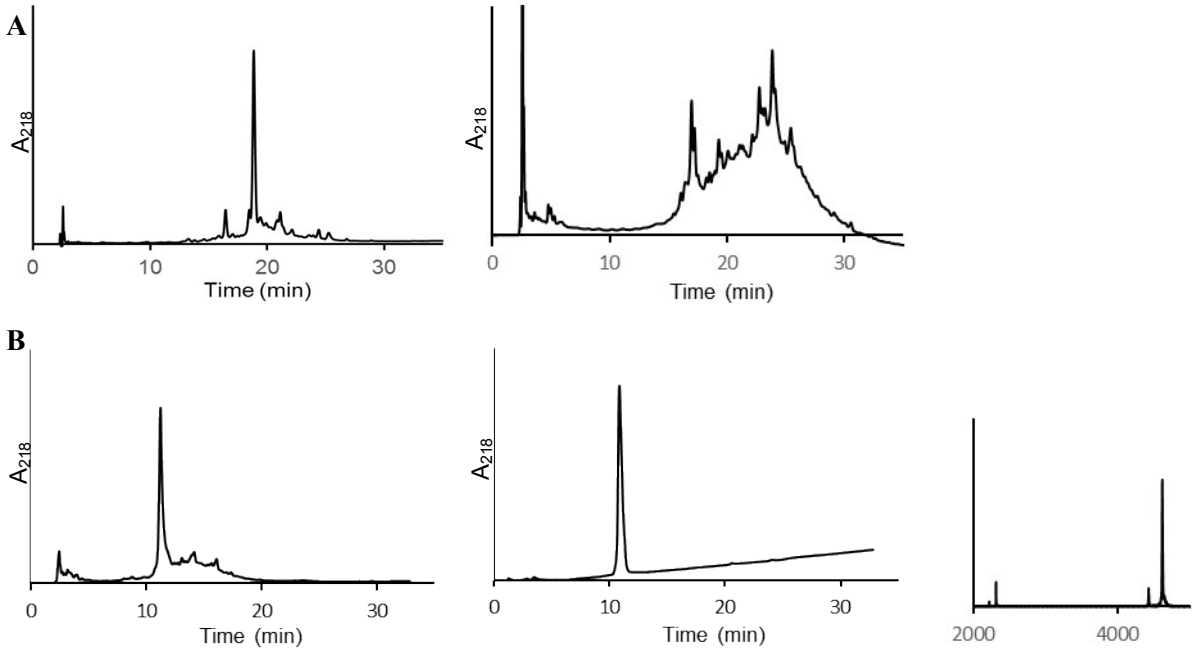


Figure S8: (A) RP-HPLC of CpA-2-Dbz(Alloc) (left 30-60 % B) and Nbz conversion using DCM (right 30-60 % B). (B) RP-HPLC of CpA-2-Nbz conversion using DMF (left). RP-HPLC (middle) and MALDI-TOF (right) of purified CpA-2-(formyl)Nbz: $[M + H]^+$ observed: *m/z* 4625, expected: *m/z*

4.3 Synthesis of CpA-3 (Thz_{71}-Ala_{97})-Nbz-Arg-Arg

Thz-REISVKFTRGVDFNWQAQALLALQEA-Nbz-RR

CpA-3 peptide was synthesized with 0.05 mmol of mFmoc-Dbz(Alloc)-RR resin with the automated synthesizer. The two Arg tags were added for improved solubility. The N-terminal amino acid was added as Boc-Thz-OH. The peptide was purified by RP-HPLC using a gradient of 30-45% Buffer B over 40 minutes. Purified peptide contains a species lacking the Nbz linker (CpA3-RR).

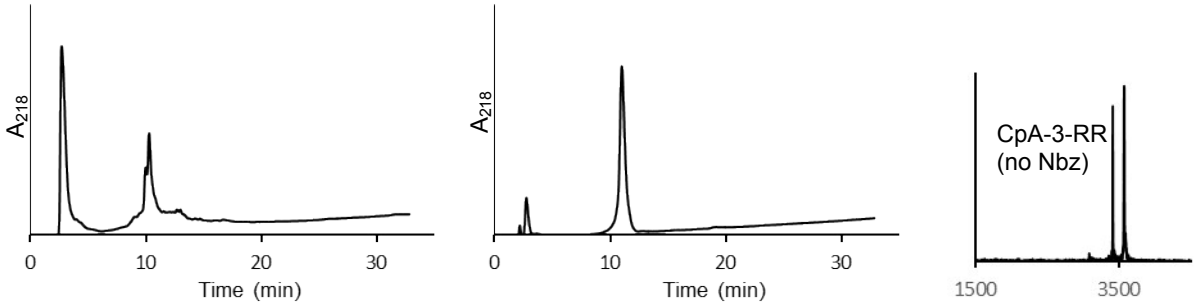


Figure S9: RP-HPLC of crude (left 30-60 % B) and purified (middle 30-60 % B) CpA-3-Nbz-RR. MALDI-TOF (right) of purified CpA-3-Nbz-RR: $[M + H]^+$ observed: *m/z* 3577, expected: *m/z* 3578.

4.4 Synthesis of CpA-4 (Thz_{98}-H_{115})-Nbz-R

Thz-EAFLVHLFEDAYLLTLH-Nbz-R

CpA4 peptide was synthesized with 0.05 mmol of Fmoc-His-Dbz-R resin, prepared by the methods described for H4-PepH. Of note, Nbz conversion also appears to be somewhat problematic for this peptide. However, sufficient product was generated to move forward with ligation.

The peptide was purified by RP-HPLC using a gradient of 35-60% Buffer B over 40 minutes. Again, species lacking the Nbz linker is observed because of a faulty base resin was used. We roughly estimated 50% of each species from RP-HPLC. A subsequent synthesis of this peptide was cleaner, but this data describes the peptide batched used in this work.

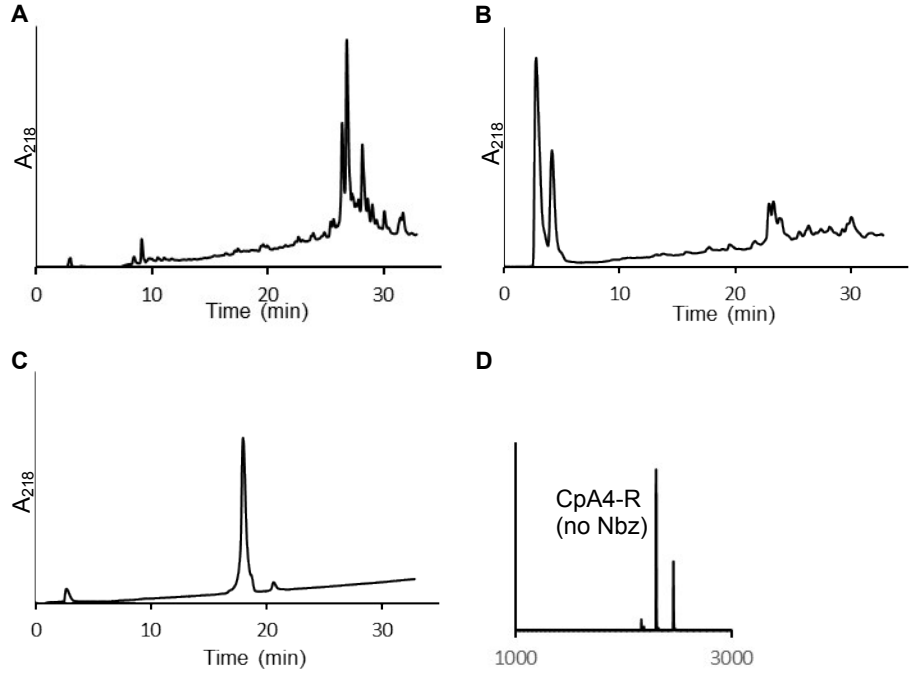


Figure S10: RP-HPLC of CpA-4-Dbz-R (A) 0-73 % B and RP-HPLC of CpA-4-Nbz-R (B) 30-60 % B. RP-HPLC (C) 35-65 % B and MALDI-TOF (D) of purified CpA-4-Nbz-R: $[M + H]^+$ observed: *m/z* 2463, expected: *m/z* 2462.

4.5 Synthesis of CpA-5-K124ac (Thz_{116}-Gly_{14}-HMBA-Arg-Gly)-Nbz

Thz-GRVTLFP**Kac**DVQLARRIRGLEEGLG-**HMBA**-RG-Nbz

CpA-5 peptide was synthesized with 0.05 mmol of Gly-HMBA-Arg-Gly-Dbz(Alloc) resin. The peptide was purified by RP-HPLC using a gradient of 30-55 % Buffer B over 40 minutes.

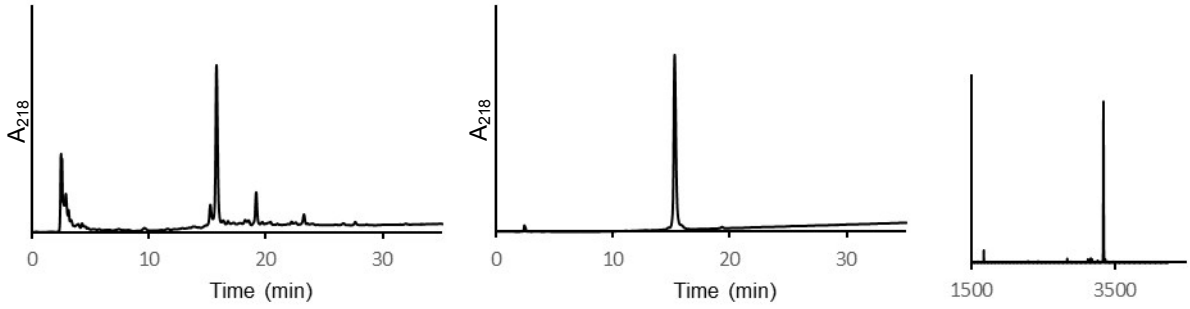


Figure S11: RP-HPLC of crude (left 30-55 % B) and purified (middle 30-55 % B) CpA-5-Nbz. MALDI-TOF (right) of purified CpA-5-Nbz: $[M + H]^+$ observed: *m/z* 3343, expected: *m/z* 3345.

5 H4 Synthesis using Solid Phase Native Chemical Ligation (SP-NCL)

Buffers:
Wash Buffer: 0.1 M Phosphate, 6 M GuHCl, 0.5 M NaCl, pH 7
Deprotection Buffer: 0.1 M Phosphate, 0.4 M Methoxyamine, 6 M GuHCl, 0.5 M NaCl, pH 4
Ligation Buffer: 0.1 M Phosphate, 0.05 M MPAA, 6 M GuHCl, 0.5 M NaCl, pH 7

Synthesis note: after addition of the C_0 peptide, resin should not be stored in methanol due to susceptibility of the HMBA linker to methanolysis. All ligations and deprotections, whether on solid-phase or in solution, are conducted at room temperature unless otherwise indicated.

5.1 Fmoc Deprotection of Base Resin

Base resin Fmoc-Thz-Ala-Rink-Gly-Gly-PEGA (section 2.3) was swelled in DMF for 20 minutes, and treated with three repetitions of 20% Piperidine in NMP for 5 minutes. The resin was then washed with 5 column volumes of DMF, followed with 5 columns volumes of methanol, and then with 5 column volumes of water. The resin was nutated in water for 10 minutes. The resin was then flow-washed with 3 column volumes of Wash Buffer, and nutated in Wash Buffer for 5 minutes. The flow-wash/nutation step was repeated 3 more times.

5.2 Thz Deprotection on the Solid Phase

Thz ring opening to generate the 1,2-aminothiol for ligation was accomplished by treatment with 0.4 M methoxylamine in Deprotection until complete as assessed by MALDI-TOF MS of microcleavages. Buffer was replaced at least three times per deprotection cycle, with reaction proceeding for a total of 6 hours (thiazolidine opening to Cys) or 12 hours (dimethyl-thiazolidine opening to penicillamine).

5.3 Solid-Phase Native Chemical Ligation

For a typical ligation cycle, after deprotection, the flow-wash/nutation step was repeated at least three times with Wash Buffer. To ensure a reactive ligation handle, the resin was equilibrated with Wash Buffer + 10 mM TCEP for 10 minutes and then rinsed with H4 wash buffer + 50 mM MPAA. Peptide was dissolved in Ligation Buffer + 20 mM TCEP as described below.

Ligations of H4-A, H4-H, and H4-C peptides were carried out with 2 mM peptide and allowed to proceed overnight. Peptide H4-B, which is added to a sterically hindered penicillamine, was ligated at 4 mM for 20 hours. Microcleavage was performed by taking a small sample of the resin and rinsing the resin extensively with Wash Buffer and then water. The sample is lyophilized briefly to remove most of the water. For cleavage at the Rink linker, the resin is incubated with TFA for 30 minutes. The sample is filtered and diluted before analysis by RP-HPLC and MALDI-TOF MS. For cleavage at the HMBA linker, the resin is incubated with 0.1 M NaOH for 15 minutes. The sample is filtered and neutralized with HCl before analysis by RP-HPLC and MALDI-TOF MS

Ligations were typically assessed in two ways: by RP-HPLC of the ligation reaction supernatant after the overnight coupling to assess remaining unreacted peptide, and by analysis of the product after cleavage. Only the product data is shown here.

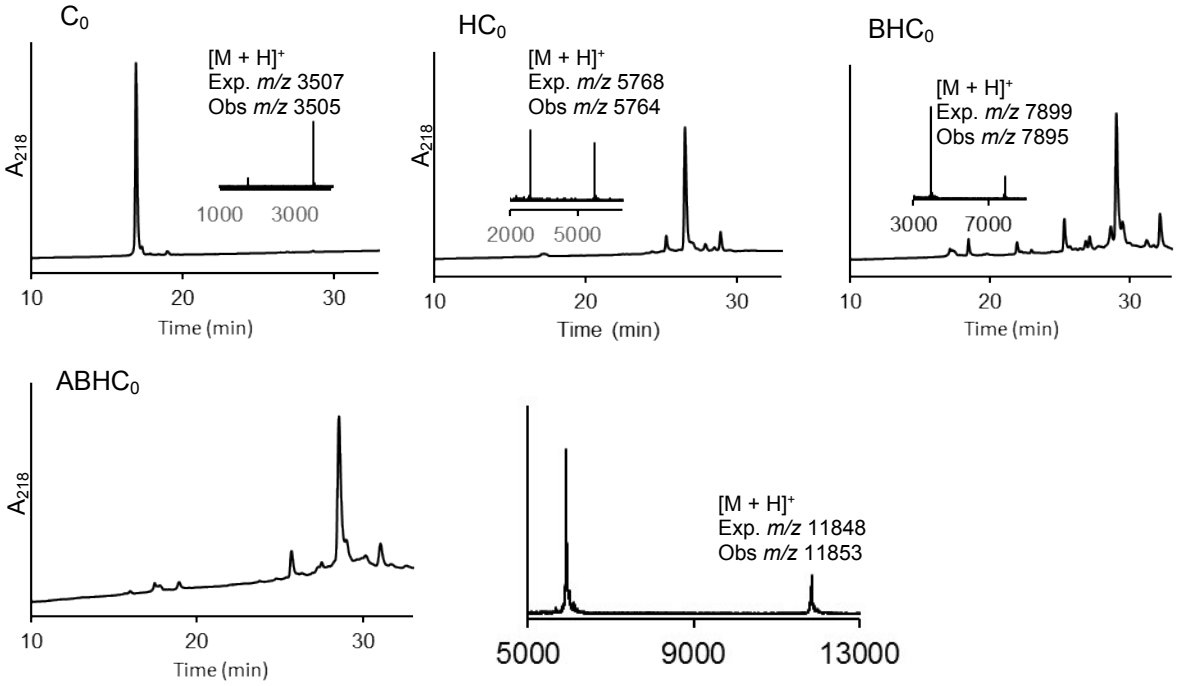


Figure S12: RP-HPLC and MALDI-TOF MS of Ligation intermediates (top 25-70 % B). RP-HPLC and MALDI-TOF MS of full length ac-H4-$ABHC_0$ (bottom 25-70 % B).

5.4 Solid-Phase Desulfurization

Desulfurization was accomplished on the solid phase in 5 M GuHCl, 0.5 M NaCl, 0.1 M phosphate, 350 mM TCEP, 75 mM MESNA, 10 mM VA-044, pH 7.4, 42 °C until complete as assessed by MALDI-TOF MS (6 hours) after cleavage from the resin by treatment with TFA.

5.5 Final product release: NaOH Cleavage at the HMBA linker

Final product H4 was released from the resin by treatment with 0.1 M NaOH for 15 minutes. The solution was neutralized by addition of HCl and eluted. The resin was extracted with TFA to recover maximal product. The product was then analyzed by RP-HPLC. Insufficient product was generated to calculate a yield.

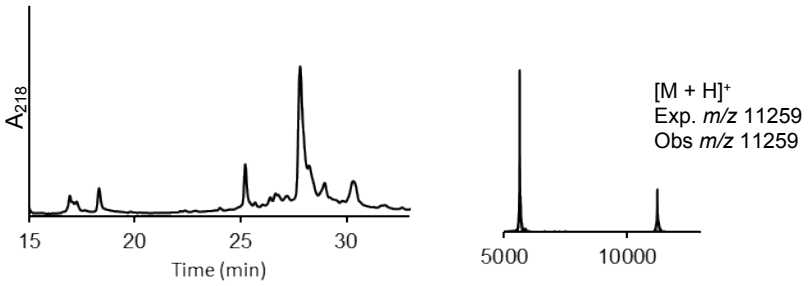


Figure S13: RP-HPLC of crude (left 25-70 % B) and MALDI-TOF MS of the product peak (right) for desulfurized ac-H4.

6 CpA Synthesis by SP-NCL

6.1 CpA-Synthesis by SP-NCL

150 uL of the base resin Thz-Ala-Gly-Gly-Rink-Gly-PEGA (section 2.3) was used for CpA SP-NCL. Based on the theoretical loading of 0.00166 mmol/mL, the scale was 0.25 µmol. Ligation was carried out using the general procedures described in Section 5 for SP-NCL of ac-H4 by SP-NCL, with the specific conditions described in Table S2 at each step. At each step, small resin samples were cleaved by rinsing the resin extensively with Wash Buffer and water, draining, and adding neat TFA; residual water in the beads remained as scavenger. Intermediate products were assessed by RP-HPLC and MALDI-TOF MS. After addition of CpA-2, insufficient product was cleaved to analyze by these approaches, and product was not observed even after addition of CpA-1. Resin bead sample was boiled in SDS loading buffer (4% SDS, 20% glycerol, 0.004% bromophenol blue, 0.125 M Tris-HCl, 100 mM DTT) and analyzed by SDS-PAGE (Fig. S14). Recombinant His-tag CpA was also ran on the gel for comparison.

Ligation Round	Peptide	MW (g/mol)	Peptide (mg)	Volume (mL)	Concentration (mM)	Molar Equivalent	Time (h)
1	CpA-5-124ac	3345	3	0.7	1.2	2.8	15
2	CpA-4	2462	2.5	0.7	1.2	2.3	17
3	CpA-3	3577	2.5	0.7	1.0	2.3	20
4	CpA-2	4628	2.5	0.5	1.1	2.6	12
5	CpA-1	3765	3	0.7	1.6	3.8	20

Table S2: Ligation conditions for SP-NCL of CpA. Molar equivalents to base resin are calculated from the theoretical scale of 0.25 µmol. Peptide concentration is calculated from the total volume of the reaction, which is the combined volume of the resin and buffer.

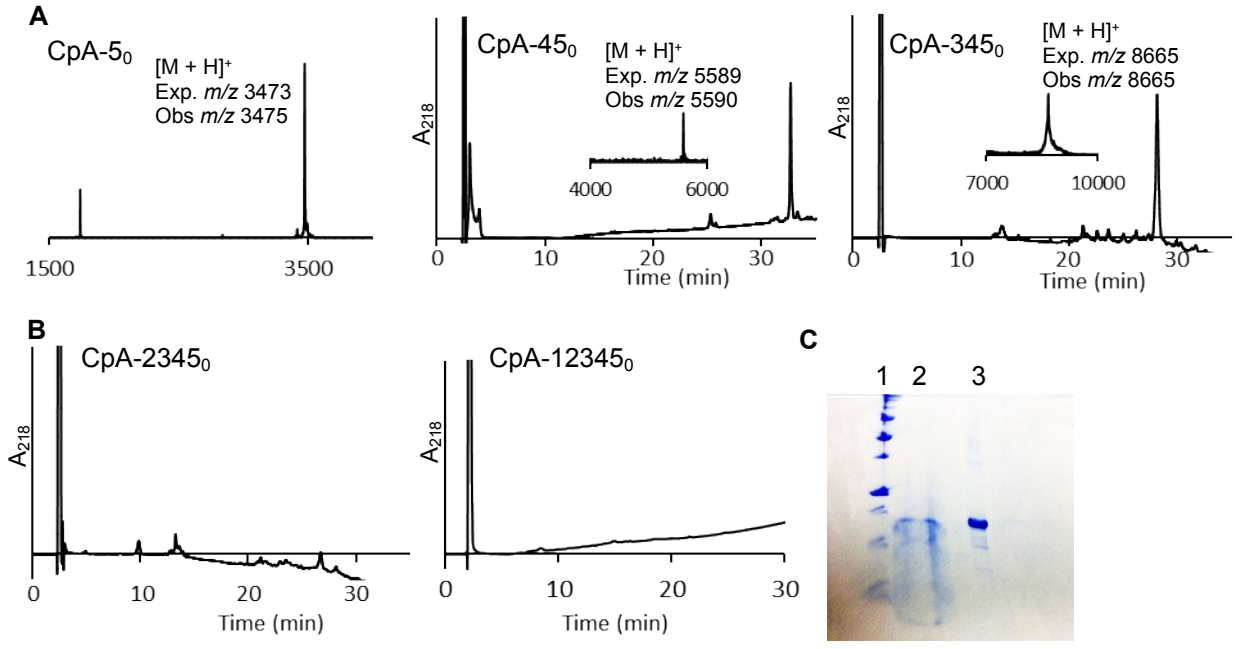


Figure S14: (A) MALDI-TOF MS of CpA-5$_0$ ligation intermediate and RP-HPLC and MALDI-TOF MS of CpA-45$_0$ (0-73 % B) and CpA-345$_0$ (30-90 % B) intermediates. (B) RP-HPLC 30-90 % B of CpA-2345$_0$ and CpA-12345$_0$. (C) SDS-PAGE of CpA12345: (lane 1) molecular weight standard, (lane 2) CpA-12345 resin boiled in SDS buffer, (lane 3) recombinant His-tagged CpA

6.2 CpA-1 Test Ligation for SP-NCL

0.5 mM CpA-1 peptide was added to 80 uL base resin in Hybrid Ligation Buffer (section 7) and allowed to react overnight. Resin was washed extensively, rinsed with water, and cleaved at the Rink linker by addition of TFA. Product was analyzed by RP-HPLC and MALDI-TOF MS; residual unreacted peptide handle from the base resin was not observed. The cleavage yield was 0.15 mg, which was calculated by quantitative ninhydrin test of cleaved CpA-1 peptide against a standard curve generated using L-Alanine. Absorbance was measured using UV Vis spectroscopy at 570 nm wavelength. With a theoretical yield of 0.4 mg calculated based on resin loading, the recovered yield was 37%.

6.3 CpA-2 Test Ligation for SP-NCL

0.5 mM CpA-2 peptide was added to 80 µL base resin in Hybrid Ligation Buffer (section 7) and allowed to react overnight. Resin was washed extensively, rinsed with water, and cleaved at the Rink linker by addition of TFA. Product was analyzed by RP-HPLC and MALDI-TOF MS; residual unreacted peptide handle from the base resin was not observed. Peptide was resuspended in 6M GuHCl, 0.1M sodium phosphate, pH 7 and quantitated by UV (ε_{280nm} = 5930 $M^{-1}cm^{-1}$) for a yield of 0.15 mg. With a theoretical yield of 0.5 mg based on resin loading, the recovered yield was 30%.

The CpA-2 test ligation was repeated as described using an additional 150 µL base resin. Product was again resuspended and quantitated by UV for a measured yield of 0.53 mg, which corresponds to a recovered yield of 56%.

7 H4-K5ac,K12ac,K91ac Synthesis by Hybrid Solid-Solution Phase NCL

Buffers:
Hybrid Wash Buffer: 0.1 M Phosphate, 6 M GuHCl, pH 7
Hybrid Deprotection Buffer: 0.1 M Phosphate, 0.4 M Methoxyamine, 6 M GuHCl, pH 4
Hybrid Ligation Buffer: 0.1 M Phosphate, 0.05 M MPAA, 6 M GuHCl, pH 7

7.1 Fmoc Deprotection of Base Resin

500 µL of the base resin Fmoc-Thz-Ala-Ahx-Rink-Gly-PEGA (section 2.3), swelled in methanol was measured in a 1.2 mL bed volume Bio-Spin Chromatography column (Bio-Rad). Based on the theoretical loading of 0.00133 mmol/mL, the scale was 0.665 µmol. The resin was treated with three repetitions of 20% Piperidine in NMP for 5 minutes. The resin was then washed with 5 column volumes of DMF, followed with 5 columns volumes of methanol, and then with 5 column volumes of water. The resin was nutated in water for 10 minutes. The resin was then flow-washed with 3 column volumes of Hybrid Wash Buffer, and nutated in Hybrid Wash Buffer for 5 minutes. The flow-wash/nutation step was repeated 3 more times.

7.2 Thz Deprotection on Solid-Phase

Conversion of thiazolidine or dimethylthiazolidine to cysteine or penicillamine was carried out using the same procedure as H4 SP-NCL (section 5.2), but using the Hybrid Deprotection Buffer instead of the Deprotection Buffer. For future procedures, analysis of deprotection with microcleavage is recommended in order to avoid incomplete deprotection as observed with this trial. Analysis of deprotection with minimal product loss is carried out as follows: 5-10 beads are washed with water, then cleaved in a minimal volume of TFA for 30 minutes with residual water acting as scavenger. MALDI-TOF MS analysis is used to assess reaction completion.

7.3 Solid-Phase Ligation of BHC_0

After deprotection the resin was washed and appropriate peptide was added using the procedures described in section 5, except Hybrid Wash Buffer and Hybrid Ligation Buffer were used instead. Immediately prior to ligation, the resin was nutated for 10 minutes with Hybrid Wash Buffer + 20 mM TCEP in order to reduce the reactive thiol. The resin was then flow-washed with one column volume of Hybrid Ligation Buffer, and nutated for 5 minutes to ensure complete buffer exchange. The buffer was then drained from the resin using vacuum. The peptide Nbz was dissolved in Hybrid Ligation Buffer, and the solution was added to the resin. TCEP was added to the resin to a final concentration of 20 mM. While reaction is typically complete after ~5 hours, the cycles in Table S3 were used to correspond to natural human sleep-work cycles in a 24 hour diurnal cycle – we expect that total reaction time could be reduced in an automated process.

Ligation Round	Peptide	MW (g/mol)	Peptide (mg)	Volume (mL)	Concentration (mM)	Molar Equivalent	Time (h)
1	H4-C-Nbz K91ac	3520	6.5	1.5	1.2	2.8	14
2	H4-H-Nbz-R	2580	4	1.3	1.2	2.3	15
	H4-H-Nbz-R	2580	4	1.6	1.0	2.3	18
3	H4-B-Nbz	2321	4	1.6	1.1	2.6	10
	H4-B-Nbz	2321	6	1.6	1.6	3.8	20

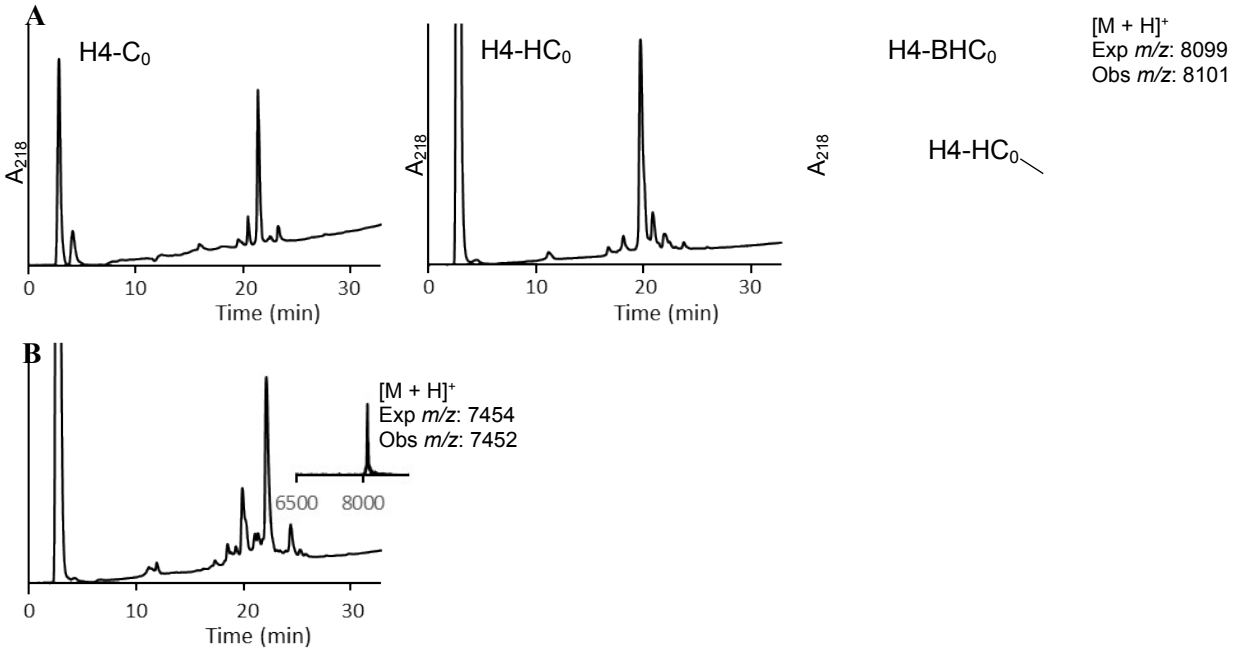
Table S3: Ligation conditions for SP-NCL of BHC_0. Molar equivalents to base resin are calculated from the theoretical scale of 0.665 µmol. Peptide concentration is calculated from the total volume of the reaction, which is the combined volume of the resin and buffer.

To assess the progress of ligation, approximately 1% of the resin was cleaved to monitor the progress of the ligation as follows: Resin sample was washed with Wash Buffer and then with water. Resin is then treated with TFA for 30 minutes; residual water in the the resin sample acts as a scavenger. The supernatant was collected by filtration, and the sample was concentrated using N_2 flow. 7:3 water:ACN was added to the sample in order to dilute the TFA, and the sample was analyzed by HPLC and MALDI.

Ligations with H4-H peptide and H4-B peptide were incomplete after the first cycle. The resin was washed and fresh peptide added for a double ligation cycle. After ligation was complete, the resin was washed, and the deprotection carried out as in section 6.2 for the next ligation cycle.

7.4 Cleavage from resin at the HMBA linker

The resin was washed (flow-wash/nutation) 3 times with Hybrid Wash Buffer, 1 time in Hybrid Wash Buffer + 10 mM TCEP, followed by 4 repetitions of flow-wash/nutation with water. The resin was then lyophilized. The peptide-resin was resuspended in 0.1 M NaOH, and the reaction was allowed to proceed for 30 minutes. The supernatant was collected via filtration. To neutralize the solution, equal volume of 0.1 M HCl was added to the resin, and combined with the supernatant. The resin was washed 3 times with TFA, and the filtrate was collected in a separate vessel. After analysis, the TFA wash was determined to contain significant H4-BHC product (in which HMBA had been cleaved). The combined sample from the NaOH cleavage, HCl neutralization, and TFA wash was lyophilized to obtain the crude H4-BHC peptide. The NaCl generated by the neutralization was removed by washing with water. The pellet was dried to give a crude yield of 4.8 mg (97% yield).



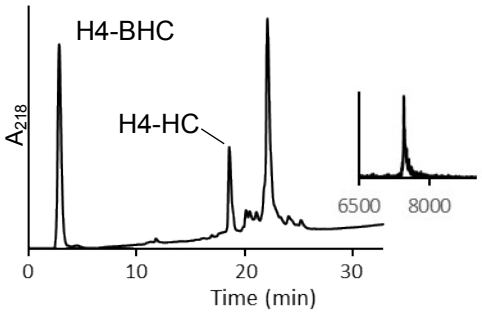


Figure S15: (A) SP-NCL of H4-BHC-K91ac as monitored by RP-HPLC 30-70 % B (0-73 % was used for H4-C_0). H4-HC side product was a result of incomplete deprotection of dmThz. (B) RP-HPLC 30-70 % B of deprotected H4-BHC.

7.5 Final Ligation of H4-A to H4-BHC in Solution

2 mg of H4-BHC-K91ac (quantitated by UV, ε_{280nm} = 7300 $cm^{-1}M^{-1}$ in 6 M GuHCl, 0.1 M sodium phosphate, pH 7) resuspended in 300 µL of Hybrid Ligation Buffer was added to 2 mg of H4-PepA-K5ac,K12ac-Nbz(formyl). TCEP was added to a final concentration of 20 mM, resulting in a final concentration of 0.9 mM H4-BHC, and 1.6 mM H4-A-K5acK,12ac-Nbz-R. The ligation was allowed to proceed for 16 hours, and the reaction was monitored by SDS-PAGE and MALDI-TOF MS.

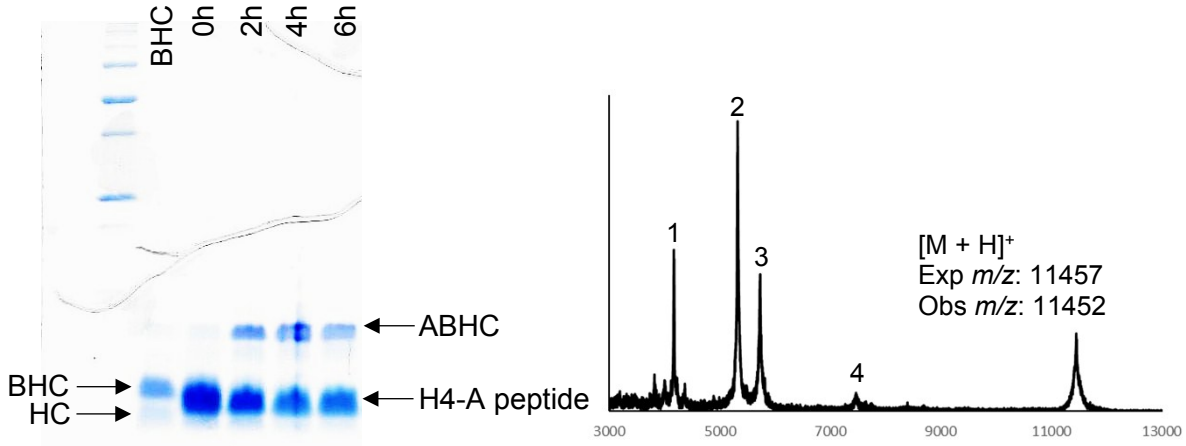


Figure S16: SDS-PAGE of Solution Phase Ligation (left) and MALDI-TOF MS of ligation (right) 1: H4-A-MPAA 2: H4-HC 3: H4-ABHC-K5ac,K12ac,K91ac (+2H) 4: H4-BHC 5: H4-ABHC (+1H)

7.6 H4 ABHC Desulfurization

Before desulfurization, the ligation reaction was dialyzed against Hybrid Wash Buffer for 6 hours in order to remove MPAA. TCEP and MESNA were added to the dialyzed sample to make the final concentration as follows: 0.1 M Phosphate, 6 M GuHCl, 75 mM MESNA, 250 mM TCEP, pH 7. The sample was sparged with argon for 30 minutes. To start desulfurization, VA-044-US was added to a final concentration of 10 mM. The sample was incubated in a 45°C water bath for 16 hours and assessed by MALDI-TOF MS for completion.

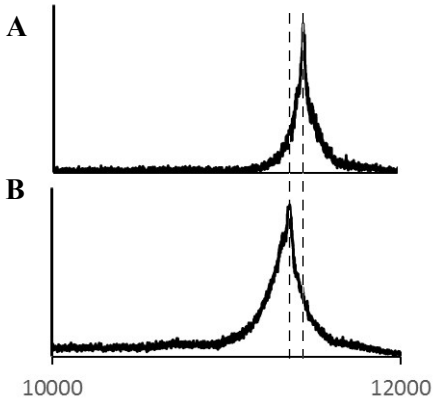


Figure S17: MALDI-TOF MS of H4 ABHC before (top) and after (bottom) desulfurization.
Before: $[M + H]^+$ observed: *m/z* 11352, expected: *m/z* 11457.
After: $[M + H]^+$ observed: *m/z* 11364, expected: *m/z* 11361.

7.7 Purification

The desulfurized sample was brought to 30% acetonitrile and purified by RP-HPLC using a gradient of 30-70% Buffer B over 50 minutes. Fractions were assessed by MALDI and SDS-PAGE, combined, and lyophilized. 0.5 mg was obtained by dry weight. The isolated yield calculated from the 2 mg of crude cleavage sample was 17%. Considering the 97% crude yield for SP-NCL, this provides an overall yield of 16%.

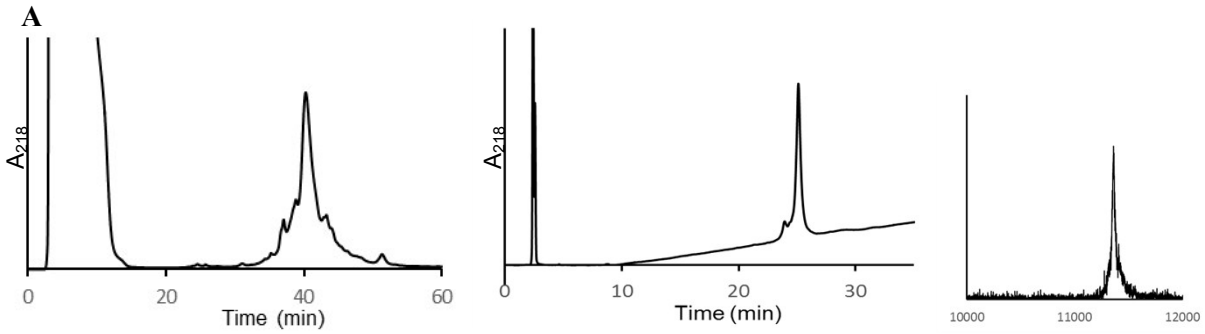


Figure S18: RP-HPLC 30-70 % B of H4 ABHC purification (left). RP-HPLC 30-70 % B (middle) and MALDI-TOF MS (right) of purified H4-ABHC. H4 K5,K12,K91ac: $[M + H]^+$ observed: *m/z* 11361, expected: *m/z* 11361.

8 CpA K124ac Synthesis using Hybrid Solution-Solid Phase NCL

The base resin Fmoc-Thz-Ala-Gly-Gly-Rink-Gly-PEGA was used in the first trial of CpA345 synthesis. In the second trial, 0.5 mL of the base resin Fmoc-Thz-Ala-Ahx-Rink-Gly-PEGA (same as H4 hybrid ligation) was used. CpA-345_0 refers to the peptide prepared in the second trial unless otherwise noted, since the only the CpA-345_0 prepared in second trial was successful in synthesizing the full length CpA. Based on the theoretical loading of 0.00133 mmol/mL, the loading was 0.665 μmol. The protocol for Fmoc deprotection, and Thz deprotection and the buffers used are identical to that of H4 hybrid ligation (see sections 7.1, 7.2, and 7.3).

8.1 Solid-Phase Ligation of CpA-345_0

Ligation was performed with 1 mM concentration of peptide Nbz in Ligation Buffer. The ligation was allowed to proceed for at least 6 hours. After the first ligation, the resin was washed at least once with Wash Buffer containing 20 mM TCEP. Analysis was done by taking 0.2% of the total resin mixture. The resin was washed extensively with Hybrid Wash Buffer and water and treated with TFA for 30 minutes. If the ligation was incomplete as determined by HPLC, the resin was washed with Hybrid Wash Buffer and the ligation was performed a second time with the appropriate peptide Nbz.

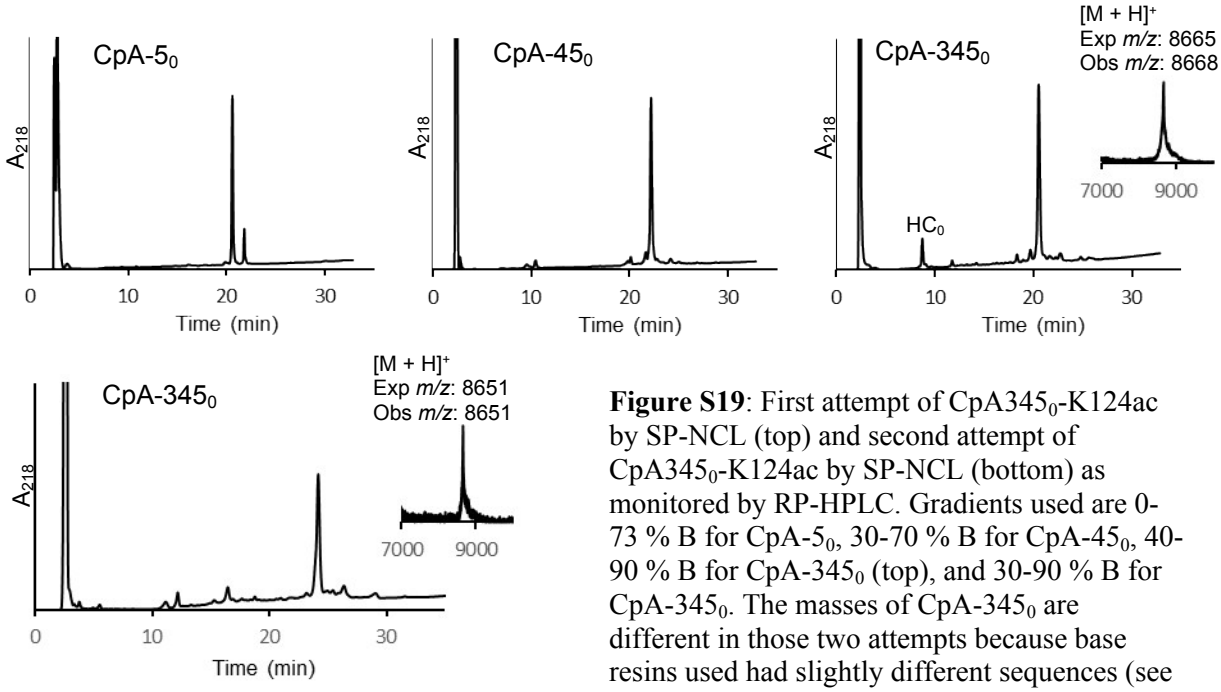


Figure S19: First attempt of CpA345_0-K124ac by SP-NCL (top) and second attempt of CpA345_0-K124ac by SP-NCL (bottom) as monitored by RP-HPLC. Gradients used are 0-73 % B for CpA-5_0, 30-70 % B for CpA-45_0, 40-90 % B for CpA-345_0 (top), and 30-90 % B for CpA-345_0. The masses of CpA-345_0 are different in those two attempts because base resins used had slightly different sequences (see section 2.3)

8.2 CpA-345$_0$ Cleavage with TFA

The peptide was treated with methoxyamine for 6 hours, and the resin was washed thoroughly with Wash Buffer containing 10 mM TCEP, and then with water. After lyophilization, the resin was treated with a mixture containing 95:2.5:2.5 TFA:H_2O:TIS for 1 hour. The solution was then filtered, and the sample was concentrated using a flow of N_2. The sample was diluted using 7:3 water:ACN mixture and lyophilized. In this representative synthesis, we obtained 6 mg with a crude yield by weight of 104% based on the theoretical loading of the base resin.

8.3 Solution-Phase Ligation

3 mg of ring opened CpA345$_0$ was dissolved in Hybrid Ligation Buffer + 20 mM TCEP. 2 mg of CpA-2-Nbz(formyl) which corresponded to a molar equivalent of 1.2 was added to the solution. The final concentrations of CpA-345$_0$ and CpA-2 were 1.2 mM and 1.4 mM, respectively. The ligation was monitored by SDS-PAGE and MADLI-TOF. After 6 hours, methoxyamine was added to the solution to a final concentration of 0.4 M. The pH was adjusted to 4, and ring opening was allowed to proceed for 2 hours. Full conversion of thiazolidine to cysteine was confirmed with MALDI-TOF MS, and the sample was dialyzed against Hybrid Wash Buffer in order to remove the methoxyamine. This step was necessary since methoxyamine was found to react with both CpA-1-Nbz and CpA-1-MPAA.

Before the last ligation, NaOH was added to the sample to a final concentration of 0.1 M for 1 hour in order to cleave the HMBA linker. The pH of this solution was 10. After the cleavage, MPAA was added to a final concentration of 50 mM, which brought the pH back down considerably. The pH was further adjusted to 7 with HCl. 2 mg of CpA-1-Nbz which corresponds to a molar equivalent of 1.5 was added to the solution. TCEP was then added to a final concentration of 20 mM. The final concentration of CpA-1-Nbz was 0.88 mM. The ligation was allowed to proceed for 16 hours.

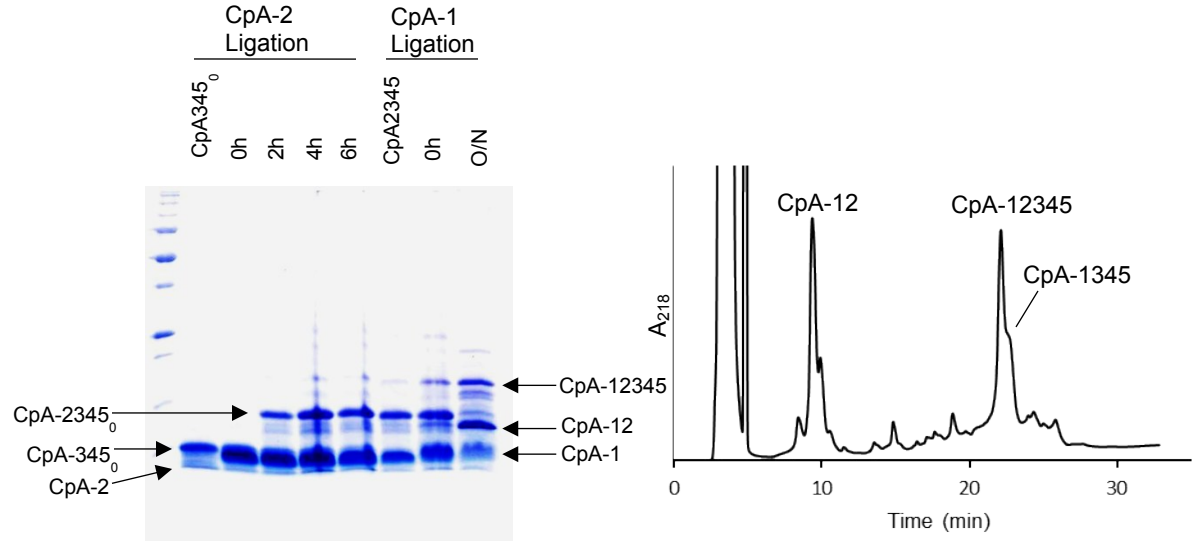


Figure S20: SDS-PAGE of Solution Phase Ligation (left) and RP-HPLC 30-90 % B of CpA-12345 (right)

8.4 CpA-12345 Desulfurization

The ligation reaction was again dialyzed against 6 M GuHCl, 0.1 M Phosphate, pH 7 in order to remove the MPAA before desulfurization. TCEP was added to the dialyzed sample to a final concentration of 250 mM, and MESNA was added to a final concentration of 75 mM. The sample was then sparged with Argon gas for 40 minutes. Desulfurization was initiated with the addition of VA-044-US to a final concentration of 10 mM, and the reaction was incubated in a 42 °C water bath overnight. Desulfurization was confirmed by MALDI-TOF

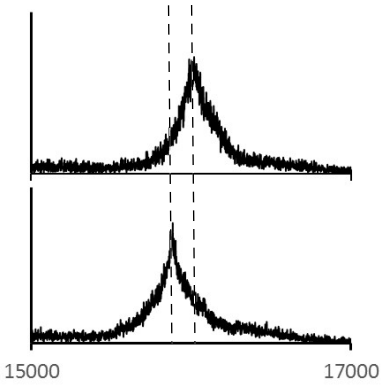


Figure S21: MALDI-TOF MS of CpA12345 before (above) and after (below) desulfurization.
Before: $[M + H]^+$ observed: *m/z* 16012, expected: *m/z* 16010.
After: $[M + H]^+$ observed: *m/z* 15883, expected: *m/z* 15882.

8.5 Purification

Full length CpA12345 was purified using the analytical HPLC using a gradient of 40-90% Buffer B over 50 minutes. Product was assessed by MALDI-TOF and SDS-PAGE. 400 µg of CpA12345 (with <5 % CpA1345, which purifies away in refolding as described below) was obtained. Of note, CpA-1345 is over-represented in the MALDI-TOF MS, but can be more accurately quantified by RP-HPLC and SDS-PAGE. This corresponds to an overall yield of 7% based on the theoretical resin loading for the first ligation step.

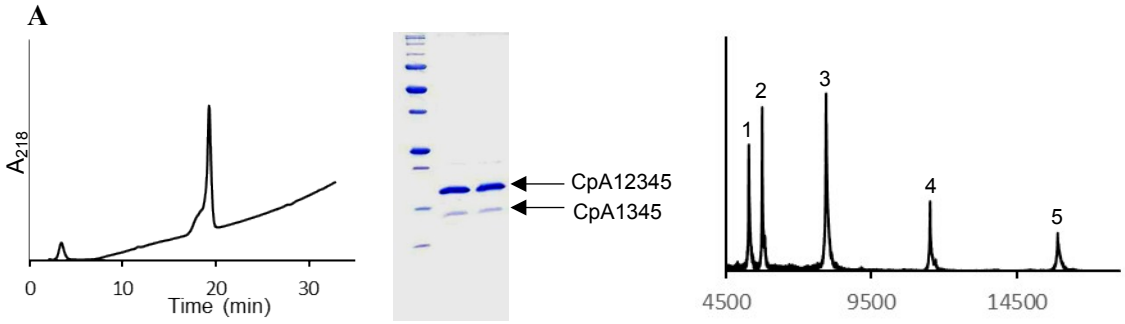


Figure S22: (A) RP-HPLC 40-90 % B (left), SDS-PAGE (middle), and MALDI-TOF MS(right) of purified CpA-12345.
Expected and Observed *m/z*:
1 CpA- K124ac: $[M + 3H]^{+3}$ Exp. *m/z* 5295, Obs. *m/z* 5297
2 CpA-1345: $[M + 2H]^{+2}$ Exp. *m/z* 5755, Obs. *m/z* 5755
3 CpA-K124ac: $[M + 2H]^{+2}$ Exp. *m/z* 7941, Obs. *m/z* 7944
4 CpA-1345 $[M + H]^+$ Exp. *m/z* 11509, Obs. *m/z* 11510
5 CpA-K124ac $[M + H]^+$ Exp. *m/z* 15882, Obs. *m/z* 15885

9 Refolding histone complexes with synthetic histones:

9.1 Refolding histone tetramer with H4-K5ac,K12ac,K91ac

To ensure complete desulfurization and to assess the optimal order of reaction steps for future syntheses, H4-K5ac,K12ac,K91ac was refolded directly out of desulfurization conditions : 100 µg of lyophilized protein was dissolved in 100 µL of sparged desulfurization buffer (0.1 M Phosphate, 6 M GuHCl, 75 mM MESNA, 300 mM TCEP, pH 7). VA-044-US was added to a final concentration of 10 mM, and the reaction was allowed to proceed for 4 hours at 44°C. Protein quality was assessed by MALDI-TOF MS before proceeding (Fig. S23) Equimolar recombinant H3-C110A was added directly to this mixture and placed into a dialysis button for double dialysis refolding against 25 mM Tris, 1 mM EDTA, 2 M NaCl, pH 7.5.[5] Refolding directly from desulfurization provided equivalent recovery of tetramer compared to representative refoldings from recombinant histones using standard procedures.

After dialysis, the sample was purified over a GE Healthcare Superdex 20/300 SEC column in 25 mM Tris, 1 mM EDTA, 2 M NaCl, pH 7.5 and assessed by SDS-PAGE (Fig. 4)

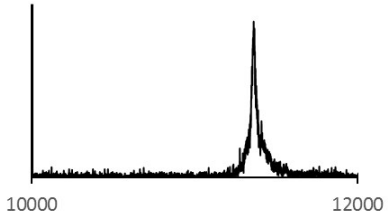


Figure S23: MALDI-TOF MS of H4 after second desulfurization $[M + H]^+$ observed: *m/z* 11361, expected: *m/z* 11361.

9.2 Refolding Histone octamer with CpA-C75S K124ac

Histone octamer was refolded by standard procedures[5]: equimolar histones were resuspended in 25 mM Tris, 1 mM EDTA, 6 M Gdn-HCl, pH 7.5 and double-dialyzed extensively against 25 mM Tris, 1 mM EDTA, 2 M NaCl, pH 7.5. Octamer was purified over a GE Healthcare Superdex 20/300 SEC column and analyzed by SDS-PAGE (Fig. 4).

1. S. K. Mahto, C. J. Howard, J. C. Shimko and J. J. Ottesen, *ChemBioChem*, 2011, **12**, 2488-2494.
2. J. B. Blanco-Canosa and P. E. Dawson, *Angewandte Chemie International Edition*, 2008, **47**, 6851-6855.
3. P. Siman, O. Blatt, T. Moyal, T. Danieli, M. Lebendiker, H. A. Lashuel, A. Friedler and A. Brik, *Chembiochem*, 2011, **12**, 1097-1104.
4. Y. X. Han, F. Albericio and G. Barany, *Journal of Organic Chemistry*, 1997, **62**, 4307-4312.
5. K. Luger, T. J. Rechsteiner and T. J. Richmond, *Methods in Enzymology*, 1999, **304**, 1-19.