Published in final edited form as:

Methods Mol Biol. 2020 ; 2133: 201–219. doi:10.1007/978-1-0716-0434-2_10.

***In vivo* Histone Labeling Using Ultrafast *trans*-Splicing Inteins**

Nicholas A. Prescott^{1,2}, Yael David^{2,3,4}

¹Tri-Institutional PhD Program in Chemical Biology, New York, NY

²Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY

³Department of Pharmacology, Weill Cornell Medical College, New York, NY

⁴Department of Physiology, Biophysics and Systems Biology, Weill Cornell Medical College, New York, NY

Abstract

The development of expressed protein ligation widened the scope of questions that could be addressed by mechanistic biochemistry. Protein *trans*-splicing (PTS) relies on the same basic chemical principles, but utilizes split inteins to tracelessly ligate distinct peptide or polypeptide fragments together with native peptide bonds. Here we present a method to adapt PTS methodologies for their use in live cells, which can deliver synthetic or native histone modifications. As an example, we provide a protocol to incorporate a small molecule fluorophore into chromatinized histones. The protocol should be easily adaptable to incorporate other modifications to chromatin *in vivo*.

Keywords

split inteins; *in cellulo* labeling; synthetic biology; chromatin

1. Introduction

Expressed protein ligation has proven to be an invaluable technique by vastly expanding the scope of questions that can be asked and answered by classic *in vitro* biochemical methods (1, 2). However, this expanded tool-box has not yet managed to bridge the gap between mechanistic studies and the systems-level correlative information gained from *in vivo* and *in cellulo* experiments. In recent years, the development and expansion of effective protein *trans*-splicing (PTS) methods into cultured cells has begun to provide the opportunity to perform experiments that may distinguish causation from correlation in a temporally-controlled and chemically-defined manner.

Classical protein splicing utilizes intact inteins which are able to autoexcise themselves from a single polypeptide chain. Such inteins have been identified in species across all domains of life (3). It was later found that whole inteins could be artificially split into separate polypeptide fragments which could facilitate protein *trans*-splicing (4). Naturally-occurring

To whom correspondence should be addressed: davidshy@mskcc.org.

split inteins were first identified in cyanobacteria. In these species, the catalytic subunit of DNA Polymerase III, DnaE, is genetically encoded as two so-called extein fragments, which are spliced together by N- and C-inteins, generally termed Int^N and Int^C, respectively (5). The DnaE class of split inteins have since served as the canonical example of split inteins, though split inteins have been identified in other organisms. For all split inteins, when Int^N and Int^C bind one another, they excise themselves and tracelessly ligate their exteins (Fig. 1). Expressed protein ligation techniques were quickly adapted to capitalize on the accessibility of these split inteins, and protein *trans*-splicing became yet another powerful tool for *in vitro* biochemical experiments, including isotopic labeling for NMR structural studies, regioselective side chain modification, and mechanistic studies of protein splicing itself (6–8).

Methods to label proteins in live cells have long been sought after. These have historically included genetic fusions to effector domains such as GFP to study subcellular localization, or the FKBP/rapamycin system to chemically induce dimerization of proteins of interest. More recently, techniques utilizing the SNAP- and Halo-tags to efficiently label some proteins with certain synthetic cargoes have been reported (9, 10). While these developments minimize the remaining artefacts from labeling, they are unable to fully eliminate them.

Efforts to bring protein splicing out of the test tube and into living cells to tracelessly label proteins date back over a decade. Early traceless splicing attempts showed only moderate signs of success, owing in part to the slow splicing rates of the split inteins employed (11). Some attempts used split inteins combined with genetic engineering to conditionally induce protein splicing (12). Recently, we developed an efficient strategy utilizing ultrafast *trans*-splicing split inteins to label histones in live cells (13). This method relies on the genetic fusion of a C-terminally truncated histone and an Int^N (in this case we used the N-intein from *Anabaena variabilis*, termed Ava^N) that is expressed in cells and incorporated into chromatin, and a synthetic piece generated by solid-phase peptide synthesis (SPPS), which contains the C-intein fragment (in this case we used the C-intein from *Nostoc punctiforme*, termed Npu^C). The Npu^C-cargo fusion is then conjugated by direct disulfide bonding to the HA2-TAT cell-penetrating peptide (CPP), which is derived from the viral protein transduction domains HA2 and TAT, allowing for its timely delivery into the cells (14).

In this chapter we provide a method for the C-terminal labeling of histone H2B *in vivo* with a small molecule fluorophore and exogenous epitope tag, including Int^N construct design and transient overexpression; synthesis of synthetic cargo and CPP; delivery of the synthetic cargo to cells; and finally, analysis of splicing efficiency by Western blot and live imaging. While this methodology was originally applied towards the study of histone proteins incorporated into chromatin, the considerable modularity of this technique makes it generally applicable for many proteins in the nucleus or cytosol. No reports have been published of successful protein *trans*-splicing within mitochondria or the lumen of other organelles, though such a reaction should be possible in principle by appending the proper mitochondrial targeting sequence or signal peptide (15, 16). In addition, although we originally employed the Ava^N-Npu^C split intein pair for the site-specific manipulation of histones, other intein families can be easily used in combination as many of them are conserved in structure and will readily react with other families. In contrast, other

orthogonal pairs of split inteins (such as IMPH/GP41) could in theory be used in tandem to modify multiple histones (or other proteins) simultaneously (4). Finally, current efforts are aimed towards expanding the methodology from cultured cells to specific tissues in live animals in order to understand the causal role of certain PTMs at the systems or organismal level.

2. Materials

18.2 MΩ-cm ultrapure water or equivalent should be used in the preparation of all aqueous solutions.

2.1 Solid Phase Peptide Synthesis

1. Dimethylformamide (DMF)
2. Dichloromethane (DCM)
3. Activation solution: 0.5 M HBTU (O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate), 0.5 M HOBt (Hydroxybenzotriazole) in DMF
4. ChemMatrix Resin™ with Rink Amide linker, 0.5 mmol/g loading capacity (*see* Note 1)
5. 20 mL polypropylene gravity flow reaction columns
6. 20 % Piperidine: 20% piperidine solution (v/v) in DMF
7. Standard Fmoc-SPPS amino acids, with Fmoc protecting groups on the alpha amine and the following side-chain protecting groups: Arg(Pbf), Asn(Trt), Asp(OtBu), Cys(Trt), Glu(OtBu), Gln(Trt), His(Trt), Lys(Boc), Ser(tBu), Thr(tBu), Trp(Boc) and Tyr(tBu), as well as Lys(ivDde) and Cys(StBu)
8. N-terminally Boc-protected amino acids (for final N-terminal residue of each peptide), including Boc-Cys(Npys)-OH
9. Acetic anhydride
10. Diisopropylethylamine (DIPEA)
11. 2 % Hydrazine: 2 % hydrazine (v/v) in DMF
12. Tetramethylrhodamine N-hydroxysuccinimidyl ester (TMR-SE)
13. Cleavage cocktail: 95 % trifluoroacetic acid (TFA), 2.5 % triisopropylsilane (TIS), 2.5 % water (v/v)
14. Diethyl ether
15. QSY™ 9 C₅-Maleimide (*see* Note 2)
16. Maleimide reaction buffer: 6 M Guanidine HCl (GnHCl), 1 X PBS, 1 mM EDTA, pH 7.0, in water
17. L-Cysteine

18. Tris(2-carboxyethyl)phosphine hydrochloride salt (TCEP)
19. Liquid Chromatography-Mass Spectrometry (LC-MS) system with C18 column
20. LCMS solvent A: 0.1 % formic acid (v/v) in water
21. LCMS solvent B: 90 % acetonitrile (v/v) and 0.1 % formic acid (v/v) in water
22. Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) system with C18 column
23. HPLC solvent A: 0.1 % TFA (v/v) in water
24. HPLC solvent B: 90 % acetonitrile (v/v) and 0.1 % TFA (v/v) in water
25. N,N'-Diisopropylcarbodiimide (DIC)

2.2. Directed Disulfide Conjugation of CPP to Cargo Peptide

1. Resuspension solution: 6 M GnHCl in water
2. Conjugation Buffer: 6 M GnHCl, 100 mM phosphate, pH 6.0 in water
3. Liquid Chromatography-Mass Spectrometry (LC-MS) system with C18 column
4. LCMS solvent A: 0.1 % formic acid (v/v) in water
5. LCMS solvent B: 90 % acetonitrile (v/v) and 0.1 % formic acid (v/v) in water
6. Reverse Phase-High Performance Liquid Chromatography (RP-HPLC) system with C18 column
7. HPLC solvent A: 0.1 % TFA (v/v) in water
8. HPLC solvent B: 90 % acetonitrile (v/v) and 0.1 % TFA (v/v) in water

2.3. Transfection of Cells and Peptide Delivery

1. 6-well, treated tissue culture treated polystyrene cell culture plate
2. Tissue culture CO₂ incubator
3. HEK 293T cells
4. Sterile, ventilated tissue culture hood
5. Full cell medium: Dulbecco's Modified Eagle Medium (DMEM), supplemented with 10 % fetal bovine serum (FBS), 2 mM glutamine, 500 units/mL penicillin/streptomycin antibiotic
6. 1.7 mL Eppendorf tube
7. Opti-MEM Reduced Serum Media
8. Lipofectamine 2000
9. Opti-MEM Reduced Serum Media
10. Dimethylsulfoxide (DMSO)

2.4. Analysis of Manipulated Cells

1. 1.7 mL Eppendorf tube
2. PBS: phosphate buffered saline in water
3. Iodoacetamide solution: 80 mM iodoacetamide in water
4. Hypotonic lysis buffer: 10 mM Tris, 15 mM NaCl, 1.5 mM MgCl₂, protease inhibitors, pH 7.6 in water
5. Dounce homogenizer with tight and loose pestles
6. Triton X-100
7. SDS RIPA lysis buffer: 25 mM Tris pH 8.0, 150 mM NaCl, 1 % (w/v) sodium dodecyl sulfate, 1 % (v/v) NP-40, 1 % (w/v) deoxycholate, pH 7.5 in water
8. 4X sample buffer (SB): 200 mM Tris-HCl, 280 mM sodium dodecyl sulfate, 6 mM bromophenol blue, 40% v/v glycerol, 20% v/v 2-mercaptoethanol in water
9. 12% polyacrylamide Bis-Tris gel and gel running apparatus
10. MOPS running buffer: 40 mM (N-morpholino) propanesulfonic acid (MOPS), 10 mM sodium acetate, 1 mM EDTA, pH 7.0 in water
11. Methanol
12. Transfer buffer: 25 mM Tris base, 192 mM glycine, 20% (v/v) methanol in water
13. PVDF membrane and membrane transfer apparatus
14. Bovine Serum Albumin (BSA)
15. PBST: phosphate buffered saline (PBS) supplemented with 0.1% (v/v) Tween-20 in water
16. Antibodies: rabbit- α HA, mouse- α H2B, α -rabbit800 and α -mouse680 (see Note 3 about secondary antibodies).
17. 35 mm glass bottom cell culture dish
18. Opti-MEM media
19. Leibovitz's L-15 media
20. Confocal microscope

3. Methods

3.1 Splicing Construct Design and Cloning

The use of split inteins for *in vivo* protein semi-synthesis provides great flexibility and utility, but also added complexity in construct engineering. In designing the construct for eventual use, one must first identify a site where the protein of interest (in our case the core histone H2B) can be split into two fragments – one to be expressed within cells, and one to be chemically synthesized. These are to be joined together by preparing each piece as a fusion to one half of a split intein pair, which will rapidly and tracelessly self-excise and

ligate the fragments together upon interacting within the cell (Fig. 1). Several key considerations must be made when designing the exact splicing site for PTS, including proximity to the protein terminus and the extein-dependence of the selected split intein pair (17). In this section, we describe the process to identify an adequate splicing site within the protein of interest, which will determine the Int^N-fusion protein to be recombinantly expressed and the Int^C-cargo peptide to be chemically synthesized. This design process was shown to allow for the delivery of a multitude of native and synthetic cargoes, including monoubiquitinated H2B within intact chromatin. However, for clarity, we later illustrate the method by appending a small molecule fluorophore and epitope tag to the C-terminus of full-length histone H2B.

1. Protein *trans*-splicing is most efficient when performed on a protein region that is unstructured, since the truncation and fusion of the Int^N fragment to the C-terminus of the protein may prevent proper folding of an ordered polypeptide. Identify an unstructured region within the protein of interest. For histone H2B, the C-terminal region spanning residues Ala¹¹⁰-Lys¹²⁵ is adequate (Fig. 1a).
2. Within the identified region, any residue that can be mutated to cysteine without disrupting protein structure or function could plausibly serve as the N-terminal cysteine residue of the delivered Int^C-cargo peptide. Determine which residues within the region are good candidates to divide the two splicing fragments. For the region of H2B identified, residues Ala¹¹⁰, Ala¹¹⁷, Ser¹²³, and Ala¹²⁵ meet these criteria (*see* Note 4).
3. Once candidate residues have been identified that could be mutated to Cys, the exact splicing site must be determined. Depending on the split intein pair used, the chosen pair may exhibit a dependence of splicing kinetics on the immediately neighboring N- or C-extein residues (*see* Note 5) (4). To circumvent this, choose a splice site from the list of candidates that most closely reflects the optimal extein sequences. Alternately, if no such option is available, it may be advisable to insert an extein linker between the protein of interest and the N-intein (*see* Note 6, Table 1 for constructs used herein).
4. With the splice site now designed, use conventional cloning methods to generate a construct for mammalian cell expression of the desired protein-Int^N fusion (*see* Notes 7, 8).

3.2 Synthesis of Int^C-Cargo Peptide

In this section we provide a protocol for the synthesis of the Int^C-cargo peptide using solid phase peptide synthesis (SPPS), which includes the peptide synthesis, conjugation of small molecules, and purification of the final peptide. The construct here is designed to enable conjugation of the fluorescence quencher (QSY-maleimide) to the Int^C fragment and the fluorophore (TMR-SE) to the extein sequence (*see* Note 9). This design ensures that fluorescence signal should only be detectable upon completed intein splicing. Standard Fmoc-based SPPS is used for the synthesis of the construct that will be delivered. All peptides are synthesized with Rink Amide linkers to generate C-terminal carboxamide that will result in a near-native (but more stable) C-terminal end on the target protein. N-

terminally Boc-protected amino acids are incorporated for the final N-terminal residue of each peptide. Peptides are synthesized using manual addition of the reagents (using a stream of dry N₂ to agitate the mixture). This protocol describes synthesis on a 0.1 mmol scale using resin with a 0.5 mmol/g capacity, but it can be scaled up as needed (*see* Table 1 for peptide sequence).

1. Before beginning synthesis, prepare 1 mL activation solution per coupling to be performed.
2. Weigh 200 mg of Rink Amide ChemMatrix Resin™ into reaction column, and swell in 5 mL DMF for 30 min. with gentle agitation before draining.
3. Add 3–4 mL of 20 % piperidine solution (*see* Note 10) to resin and incubate for 1 minute to pre-equilibrate resin. Drain and repeat, incubating this time for 20 min. to allow complete deprotection to take place.
4. Drain and wash the resin under continuous flow, first with 5 column volumes (CV, relative to volume of swelled resin) of DCM and then with 5 CV DMF.
5. Dissolve 5 molar eq. Fmoc-protected amino acid (relative to the resin) in 1 mL activation solution. Add 10 eq. DIPEA to amino acid solution. For residues Lys³⁹ and Cys³⁶, the orthogonal side-chain protecting groups ivDde and StBu must be used, respectively, instead of the standard Boc and Trt groups.
6. Add amino acid solution to the N-terminally deprotected resin, and allow to react for 30 minutes at room temperature, with gentle agitation.
7. Wash resin with 3 CV DMF under continuous flow.
8. Add 10 eq. acetic anhydride and 20 eq. DIPEA into 1 mL DMF, and incubate with the resin for 20 min. with agitation. Drain and repeat once.
9. Wash resin with 3 CV DMF under continuous flow.
10. Add an excess of 20 % (v/v) piperidine in DMF to remove Fmoc.
11. Drain and wash resin thoroughly with DMF to remove trace piperidine, as in step 9.
12. Repeat steps 5–11 for each amino acid coupling until the desired peptide has been completed (*see* Table 1 for peptide sequence).
13. Upon completed peptide chain assembly, deprotect the Lys³⁹ sidechain by adding 5 mL of 2 % hydrazine solution and incubating for 3 minutes. Drain, then wash resin as in step 4. Repeat entire step 5 times to ensure full deprotection of the sidechain.
14. Dissolve 2 eq. of TMR-SE in approximately 4 mL DMF, then add 10 eq. DIPEA into the solution. Add fluorophore solution to resin, and allow to react for 12 hours or overnight at room temperature with agitation. Cover sides of reaction column with foil to keep resin protected from light if necessary.
15. Drain and wash the resin thoroughly, as in step 4.

16. To cleave peptide from resin, add 20 μ L cleavage cocktail per mg of peptidyl-resin into the reaction column and allow the cleavage reaction to take place for 1.5–2 hours at room temperature with agitation.
17. Drain the peptide-containing filtrate into a 15 mL conical tube and evaporate the solution by gently blowing nitrogen over sample for 30 minutes. The column and resulting cleaved resin can then be disposed of.
18. Add cold diethyl ether to the evaporated peptide solution such that any remaining cleavage cocktail is less than 10 % of the total volume in order to precipitate the peptide. Centrifuge at 4,000 rcf for 10 min. at 4 °C. Carefully decant the supernatant, and repeat twice more by adding 10 mL of ether each time.
19. Allow the resulting peptide pellet to dry by leaving it at room temperature overnight in a chemical hood to remove all ether, protecting from light to avoid possible photobleaching of the fluorophore.
20. To alkylate the free thiol group of Cys²² with the QSY quencher, first resuspend the Int^C-cargo peptide to a concentration of 4 mM in maleimide reaction buffer (see Note 11).
21. Separately, dissolve 2 eq. QSY-maleimide in an equal volume of maleimide reaction buffer (i.e., to make an 8 mM solution). Combine peptide and maleimide solutions, and allow reaction to take place at room temperature with agitation for several hours (see Note 12, 13).
22. Quench excess maleimide by dissolving 25 eq. L-cysteine (with respect to initial maleimide concentration) into the reaction solution. Allow reaction to take place at room temperature with agitation for several hours (see Note 14).
23. Once excess maleimide is quenched, proceed to remove the StBu protecting group from Cys³⁶ by weighing out TCEP and adding directly into peptide solution to a final concentration of 90 mM.
24. Monitor deprotection periodically by LCMS and only proceed to purify the peptide when complete.
25. Purify product by semi-prep or preparative RP-HPLC (Fig. 2a) and lyophilize before storing peptide at –80 °C.

3.3 Synthesis of CPP and Directed Disulfide Conjugation to Int^C-Cargo Peptide

In this section we provide a protocol for the synthesis of the CPP (HA2-Tat) by solid phase peptide synthesis (SPPS) and subsequent disulfide-conjugation of the CPP to the Int^C-cargo peptide, as well as the purification of the final product. The CPP-mediated delivery of synthetic cargo into cells provides two great strengths when compared to other methods like viral or liposomal delivery: increased synthetic accessibility and lower levels of cytotoxicity (see Note 15) (18, 19). Crucially, the targeted disulfide conjugation of the CPP to the catalytic cysteine residue of the Int^C fragment also brings along three key advantages: protection of the catalytic residue from oxidation or other reactions prior to delivery,

traceless intracellular removal from the Int^C-cargo by reductive cleavage within the cytosol, and greater efficacy than when the CPP is fused C-terminally to the cargo peptide.

3.3.1 Synthesis of HA2-TAT Peptide—The standard SPPS protocol from section 3.2 was used to synthesize the CPP, with the following modifications (*see* Table 1 for CPP amino acid sequence):

1. When preparing to couple the Boc-Cys(Npys)-OH after the rest of the HA2-TAT peptide has been produced, instead of following the standard protocol, dissolve 5 molar eq. each of Boc-Cys(Npys)-OH, HOBt, and DIC (with respect to the resin) in 1 mL DMF, and incubate with peptidyl resin for 1 hour to couple (*see* Note 16).
2. Using the same cleavage cocktail as described previously, cleave peptide from resin for 3 hours (*see* Note 17).
3. After cleavage, drain CPP solution into a 15 mL conical tube and precipitate crude peptide with cold diethyl ether, as described previously.
4. Resuspend the pelleted peptide in a 50:50 mixture of HPLC solvents A and B and lyophilize the resuspended peptide overnight.
5. Dissolve lyophilized peptide in resuspension buffer.
6. Purify CPP by RP-HPLC (Fig. 2b, *see* Note 18).

3.3.2 Conjugation of CPP to Cargo Peptide

1. In degassed conjugation buffer, separately resuspend pure HA2-TAT to a concentration of 4.5 mM and cargo peptide to a concentration of 5 mM in equal volumes (*see* Note 19). Check that pH remains at 6.0 for both, and adjust with HCl or NaOH if necessary.
2. Combine Int^C-cargo and CPP to final concentrations of 2.25 mM and 1.5 mM, respectively.
3. Allow components to react for at room temperature. Reaction completion can be monitored every 30 minutes by LCMS, and typically is achieved within 1 hour.
4. Purify resulting conjugate from unreacted materials by RP-HPLC (Fig. 2c).

3.4 Expression of Histone-Int^N Fusion and Int^C-Cargo Peptide Delivery

In this section we provide a protocol for the transient transfection and overexpression of the histone-Int^N fusion protein, as well as delivery of the synthetic Int^C-cargo peptide (*see* Note 20). Stability and functionality of the Int^N-fusion protein should be tested to uncouple any effects of the fusion protein from the effects of downstream delivery or experimental conditions. While we provide a method based on transient, lipofectamine-mediated transfection protocol, other reagent-based or even viral transduction methods might be employed to produce a cell line stably expressing the Int^N-fusion.

3.4.1 Transient Transfection of Histone-Int^N Fusion—For a single well in a 6-well flat bottom plate of $\sim 2 \times 10^6$ cells, our transfection protocol is as follows (*see* Note 21):

1. Grow HEK 293T cells at 37 °C in 5 % atmospheric CO₂ in incubator to ~ 70 % confluence in full cell medium (*see* Note 22).
2. In tissue culture hood, combine 400 ng of plasmid DNA for mammalian expression of the Int^N-fusion protein, as designed Section 3.1, with 80 μ L of Opti-MEM media. Separately, combine 3 μ L of lipofectamine with 80 μ L DMEM. Incubate at room temperature for 15 minutes.
3. Combine DNA-media mixture with lipofectamine-media mixture, pipetting gently to mix, and incubate for 30 min. at room temperature.
4. Aspirate media off of cells and add 1 mL of full media. Add the transfection mixture prepared in the previous step to cells in a dropwise manner. Gently mix the media by moving the plate side to side.
5. After 6–12 hours of incubation at 37 °C in a CO₂ controlled incubator, aspirate media and supplement with 2 mL of fresh full media.
6. Allow transfected cells to grow for an additional 24 hours before proceeding to delivery of synthetic cargo peptide (Fig. 3a, *see* Notes 23, 24).

3.4.2 Int^C-Cargo Peptide Delivery

1. Prepare peptide for delivery by dissolving pure, lyophilized peptide stocks to a concentration of 1 mM in DMSO. This concentrated stock solution can be flash frozen and reused if any excess remains unused (*see* Note 25).
2. Dilute stock peptide solutions in Opti-MEM cell medium to a final volume of 500 μ L and to the necessary delivery concentration, typically around 1–10 μ M (*see* Note 26, 27).
3. Aspirate media off Int^N-expressing cells, then wash them with 2 mL Opti-MEM. Aspirate media once more.
4. Gently add peptide-Opti-MEM solution to cells by pipetting onto walls of the well and then spread by gently moving plate side to side. Incubate at 37 °C for 1 hour.
5. Supplement cells with an equal volume of full cell culture medium and incubate at 37 °C for an additional 1 to 4 hours. (Fig. 3b).

3.5 Analysis of Manipulated Cells

In this section, we provide two methods to analyze successful protein *trans*-splicing in cells: live-cell confocal microscopy and western immunoblotting analysis (Fig. 3 and 4). Following cargo peptide delivery (step 5 in section 3.4.2), cells are ready for analysis as described below (*see* Note 28), or subsequent experimentation (*see* Note 29).

3.5.1 Cell Fractionation and Western Blot Analysis of Splicing

1. Wash cells with 1 mL cold PBS, then aspirate away the buffer.
2. Resuspend cells in 1 mL cold PBS supplemented with 100 μ L of iodoacetamide solution (*see* Note 30).
3. Transfer cells to a 1.7 mL Eppendorf tube, and pellet by spinning at 400 rcf for 5 min. at 4 °C in a tabletop microcentrifuge. Discard the supernatant.
4. Resuspend cell pellet in 1 mL of hypotonic lysis buffer and lyse with 10 strokes of a loose pestle Dounce homogenizer.
5. Centrifuge at 400 rcf for 5 min. at 4 °C to pellet nuclei, and remove the supernatant, which contains the cytosolic fraction.
6. Resuspend the nuclear pellet in 1 mL of hypotonic lysis buffer supplemented with 1 % (v/v) Triton X-100. Lyse with 10 strokes of a tight pestle Dounce homogenizer.
7. Centrifuge at 10,000 rcf for 10 min. at 4 °C. Remove supernatant, which contains the nucleoplasm fraction.
8. Resuspend chromatin pellet in 300 μ L of SDS-RIPA lysis buffer and then sonicate for 5 s at 35 % amplitude with a rod sonicator.
9. In separate Eppendorf tubes, combine 5 μ L of 4X sample buffer with 15 μ L of the prepared chromatin fractions from non-transfected, transfected but non-delivered, and transfected and delivered cells. Heat at 95 °C for 5 min.
10. Load samples onto a 12 % polyacrylamide bis-tris gel, and run at a constant voltage of 150 V for 1 hour in MOPS running buffer.
11. When the gel run is over, transfer protein contents to a PVDF membrane at a constant voltage of 85 V for 75 min. at 4 °C using standard gel/membrane transfer apparatus (*see* Note 31).
12. When transfer is complete, prepare blocking solution by dissolving 0.75 g BSA in 15 mL PBS. Incubate membrane in this solution for 1 hour at room temperature.
13. After blocking, wash the membrane three times in 15 mL PBST for 5 min.
14. Blot the membrane with primary antibodies diluted into 15 mL PBST according to manufacturer instructions for 1 hour at room temperature or overnight at 4 °C (*see* Table 2 for antibodies).
15. After incubation with the primary antibody, wash the membrane as in step 13.
16. Blot the membrane with secondary antibodies diluted in 15 mL PBST according to manufacturer instructions for 1 hour at room temperature (*see* Table 2 for antibodies).
17. Image the blot according to manufacturer instructions for the secondary antibodies used.

3.5.2 Live Cell Imaging

1. Rather than growing in a flat-bottom 6-well plate, grow cells on a glass bottom 35 mm dish, carrying out the above-described protocol (3.4.2).
2. After cargo delivery, wash cells with opti-MEM medium.
3. **Optional:** Some microscopes may be sensitive to the color of the cell medium. If so, wash once more with Leibovitz L-15 medium.
4. Immediately after washing, take cells for imaging. We used a confocal microscope to take multiple cross-sectional images per condition. Non-transfected cells should show no TMR fluorescence, as the splicing reaction will be unable to take place in the absence of the Int^N fusion protein. TMR fluorescence within the nucleus would be indicative of successful splicing in transfected and delivered cells.

4. Notes

1. We typically prefer to use the ChemMatrix ResinTM because it has been shown to work well for long or complicated peptides (20).
2. We have used the QSYTM series of fluorescence quenchers due to their relatively broad absorption spectra, though any compatible fluorophore/dark quencher pair should theoretically work with this procedure.
3. We prefer to use fluorescent secondary antibodies, which allow for multiplexing and imaging of multiple components concurrently, but chemiluminescent antibodies should also work in this context.
4. If an endogenous cysteine residue is present in the protein of interest N-terminal to the site of manipulation, this may serve as the ideal splicing site since it would retrosynthetically come from the N-terminal Cys of the delivered peptide. If no Cys residue is available, serine, threonine, or alanine residues may be replaced with the necessary Cys of the delivered peptide to best mimic the native protein.
5. Generally, PTS has been performed using naturally occurring split inteins. However, recent work from the Muir lab has described engineered split inteins with faster splicing kinetics and significantly lower dependence on extein residues (17, 21). Application of these engineered species towards PTS may be another way to circumvent the introduction of exogenous sequences to the final spliced product.
6. If none of the identified residues appears better than the others, multiple constructs may need to be tested empirically. If opting to append an extein linker to the native sequence of the protein being studied, we suggest running a series of control experiments to verify that the linker does not disrupt protein function or stability.
7. Any cloning method should be applicable to generate the Int^N-fusion construct.

8. As this DNA construct will be used for transfection into mammalian cells, it will be needed in high purities and at concentrations of $\sim 1 \mu\text{g}/\mu\text{L}$.
9. Although we here provide the protocol to couple the QSY-maleimide and TMR-SE, many different cargoes have been incorporated using this strategy, including synthetic amino acids with pre-defined native or non-hydrolysable PTMs and larger peptide epitopes for immunoblotting. Delivery of larger cargoes, including an entire ubiquitinated lysine residue, has also been reported, though this was most successful in isolated nuclei rather than whole cells.
10. If the sequence to be synthesized contains Asp or Asn residues followed by Gly, Ala, or Ser, it is possible that the acid or amide sidechains may cyclize with the peptide backbone upon treatment with a strong base like piperidine. If such a motif is present, use a solution of 20 % piperidine (v/v) and 0.1 M HOBt in DMF to deprotect.
11. We suggest first carrying out an analytical test reaction on a small, 0.1 mg scale, and then scaling up to as large as 10 mg Int^C peptide.
12. Verify that the combined peptide-maleimide solution is still at pH 7.0. If it is not, adjust accordingly by adding either NaOH or HCl solutions.
13. Completion of maleimide alkylation should be monitored hourly by LCMS until all unreacted peptide is consumed. If need be, reaction can be stored overnight at -20°C and resumed the following morning.
14. As with maleimide alkylation of the Int^C peptide, monitor reaction completion by LCMS. Only proceed when all unreacted maleimide is consumed.
15. Although we use HA2-TAT as our preferred cell-penetrating peptide, other CPPs should work successfully in principle. For a thorough review of common CPPs and their applications, see reference 15.
16. The coupling of the Boc-Cys(Npys)-OH can be sluggish. We suggest monitoring periodically by ninhydrin test until completion.
17. The poly-Arg(Pbf) stretch in the C-terminal region of this peptide can reduce cleavage efficiency, which is why cleavage time is extended.
18. When purifying HA2-TAT by RP-HPLC, we suggest maintaining the temperature at or below 50°C to avoid split peaks arising, likely due to possible conformational changes in the peptide.
19. We recommend performing this reaction on as large a scale as possible, as we have found smaller-scale reactions tend to have lower yields of the desired conjugation product.
20. In order to verify that splicing will proceed successfully on chromatin, we suggest carrying out a preliminary *in nucleo* splicing experiments by isolating nuclei from transfected cells by hypotonic lysis, adding the intein cargo (without CPP), and monitoring splicing by Western blot, taking time points up to 1 hour post-delivery (Figure 3). This is an important exercise to determine if any

splicing difficulties might occur with the designed construct, and uncouple them from any cellular delivery issues.

21. If cells are to be used for live imaging microscopy, carry out the transfection in the appropriate glass-bottom plate rather than transfecting in standard cell culture dishes and re-seeding Int^N-fusion expressing cells after transfection or delivery.
22. We utilize HEK 293T cells in this example, but the method has also been applied successfully in HeLa, U2OS, and LF2 cell lines (13).
23. In the case of histones, the cells must be given at least a full cell cycle to express the histone-intein fusion and subsequently incorporate it into chromatin. However, for most other cellular proteins, expression times prior to delivery need not last so long.
24. Before proceeding with delivery, control experiments should be carried out to verify that the truncated intein-fused protein of interest is not deleterious to cellular homeostasis. We suggest, at a minimum, cell viability assays. In the case of histone H2B, because histones regulate all DNA-templated processes, we also suggest functional assays such as qPCR, RNA-seq or ChIP-based experiments. For other proteins, signaling or biochemical experiments may be necessary.
25. This can be performed at any scale, since the DMSO stock can be frozen and stored, but we generally dissolve approximately 5 mg of the conjugated CPP-Int^C cargo peptide at a time.
26. When diluting the peptide DMSO stock into cell media, do so gradually by adding media into the stock incrementally, or else the peptide may crash out of solution.
27. The optimal peptide concentrations for delivery may need to be determined empirically, as it can vary depending on the split intein pair in use and extein dependence.
28. Several analytical methods can be used to verify successful splicing depending on the cargo delivered, including: Western Blot, immunofluorescence, live-cell imaging, or (co-)immunoprecipitation.
29. The end-goal or final experimental protocol is likely to vary depending on the protein being studied. In the case of our histone delivery experiments, ChIP-based methods like ChIP-seq or ChIP-qPCR, or transcriptomics experiments (i.e., RNA-seq) are likely subsequent experiments.
30. Subsequent wash steps immediately after delivery should include iodoacetamide to rapidly alkylate the catalytic cysteine residues of any remaining unreacted intein.
31. At this point, in-gel or in-membrane fluorescence of TMR could be detected by imaging without proceeding further through the western blotting procedure.

Acknowledgements

This work was supported by the Josie Robertson Foundation, the Pershing Square Sohn Cancer Alliance and the CCSG core grant P30 CA008748 and the NIH CEBRA award # DA044767 to Y. D. N.A.P. is supported by the NIH T32 GM115327-Tan chemistry-biology interface training grant and the National Science Foundation Graduate Research Fellowship Grant Number 2017239554.

References

1. Muir TW, Sondhi D, and Cole PA (1998) Expressed protein ligation: a general method for protein engineering. *Proc Natl Acad Sci U S A* 95:6705–10 [PubMed: 9618476]
2. Vila-Perelló M and Muir TW (2010) Biological Applications of Protein Splicing. *Cell* 143:191–200 [PubMed: 20946979]
3. Perler FB (2002) InBase: the Intein Database. *Nucleic Acids Res* 30:383–384 [PubMed: 11752343]
4. Shah NH, Dann GP, Vila-Perelló M, et al. (2012) Ultrafast protein splicing is common among cyanobacterial split inteins: Implications for protein engineering. *J Am Chem Soc* 134:11338–11341 [PubMed: 22734434]
5. Shah NH and Muir TW (2011) Split Inteins: Nature's Protein Ligases. *Isr J Chem* 51:854–861 [PubMed: 23620603]
6. Mootz HD (2009) Split Inteins as Versatile Tools for Protein Semisynthesis. *10:2579–2589*
7. Vila-Perelló M, Liu Z, Shah NH, et al. (2013) Streamlined expressed protein ligation using split inteins. *J Am Chem Soc* 135:286–292 [PubMed: 23265282]
8. Shah NH and Muir TW (2014) Inteins: nature's gift to protein chemists. *Chem Sci* 5:446–461 [PubMed: 24634716]
9. Juillerat A, Gronemeyer T, Keppler A, et al. (2003) Directed Evolution of O6-Alkylguanine-DNA Alkyltransferase for Efficient Labeling of Fusion Proteins with Small Molecules In Vivo. *Chem Biol* 10:313–317 [PubMed: 12725859]
10. Los GV, Encell LP, McDougall MG, et al. (2008) HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* 3:373–382 [PubMed: 18533659]
11. Giriat I and Muir TW (2003) Protein Semi-Synthesis in Living Cells. *J Am Chem Soc* 125:7180–7181 [PubMed: 12797783]
12. Mootz HD, Blum ES, Tyszkiewicz AB, et al. (2003) Conditional Protein Splicing: A New Tool to Control Protein Structure and Function in Vitro and in Vivo. *J Am Chem Soc* 125:10561–10569 [PubMed: 12940738]
13. David Y, Vila-Perelló M, Verma S, et al. (2015) Chemical tagging and customizing of cellular chromatin states using ultrafast trans-splicing inteins. *Nat Chem* 7:394–402 [PubMed: 25901817]
14. Wadia JS, Stan RV, and Dowdy SF (2004) Transducible TAT-HA fusogenic peptide enhances escape of TAT-fusion proteins after lipid raft macropinocytosis. *Nat Med* 10:310–315 [PubMed: 14770178]
15. Hoye AT, Davoren JE, Wipf P, et al. (2008) Targeting Mitochondria. *Acc Chem Res* 41:87–97 [PubMed: 18193822]
16. Yu H-C, Lu M-C, Li C, et al. (2013) Targeted Delivery of an Antigenic Peptide to the Endoplasmic Reticulum: Application for Development of a Peptide Therapy for Ankylosing Spondylitis. *PLoS One* 8:e77451 [PubMed: 24155957]
17. Stevens AJ, Sekar G, Shah NH, et al. (2017) A promiscuous split intein with expanded protein engineering applications. *Proc Natl Acad Sci* 114:8538–8543 [PubMed: 28739907]
18. Suhorutsenko J, Oskolkov N, Arukuusk P, et al. (2011) Cell-Penetrating Peptides, PepFects, Show No Evidence of Toxicity and Immunogenicity In Vitro and In Vivo. *Bioconjug Chem* 22:2255–2262 [PubMed: 21978260]
19. Guidotti G, Brambilla L, and Rossi D (2017) Cell-Penetrating Peptides: From Basic Research to Clinics. *Trends Pharmacol Sci* 38:406–424 [PubMed: 28209404]
20. García-Martín F, Quintanar-Audelo M, García-Ramos Y, et al. (2006) ChemMatrix, a Poly(ethylene glycol)-Based Support for the Solid-Phase Synthesis of Complex Peptides. *J Comb Chem* 8:213–220 [PubMed: 16529516]

21. Stevens AJ, Brown ZZ, Shah NH, et al. (2016) Design of a Split Intein with Exceptional Protein Splicing Activity. *J Am Chem Soc* 138:2162–2165 [PubMed: 26854538]

Figure 1.

Synthetic peptide delivery to chromatinized histones. A) Schematic of intein splicing reaction showing (1) initial split intein recognition, (2) the intein-mediated S to N acyl shift, and (3) transthioesterification followed by splicing resolution to yield spliced exteins, in this case histone H2B and a peptide containing an HA epitope and small molecule fluorophore, with a native peptide bond. B) Delivery of CPP-cargo peptide into cells containing histone-Int^N fusion proteins. Upon entry to the cytosol, reductive cleavage separates the CPP from the cargo peptide, which proceeds to diffuse into the nucleus, where ultrafast *trans*-splicing occurs after recognition of Int^N by Int^C. Upon splicing, the dark fluorescence quencher is separated from the fluorophore, allowing fluorescence to be detected.

Figure 2.
Representative analytical RP-HPLC chromatograms for synthetic cargo peptide (a), HA2-TAT CPP (b), and disulfide-conjugated CPP-cargo peptide (c).

Figure 3.

Western blot analysis showing time-course of protein *trans*-splicing. Lanes 1–4: non-transfected sample shows no product formation over time; lanes 5–8: splicing product forms rapidly over course of reaction, concurrent with depletion of free cargo peptide; lanes 9–12: cells transfected with catalytically inactive Int^N fusion protein (C1A) show no formation of splicing product.

Non-Transfected

Transfected

Figure 4.

Live-cell imaging of non-transfected and transfected cells delivered with TMR-carrying CPP-cargo peptide. Top, fluorescent (left) and bright-field (right) images of non-transfected cells after cargo delivery. Bottom, representative fluorescent images of transfected cells after cargo delivery.

Table 1.

Amino acid sequence of transfected protein and synthetic peptides

Component	Amino Acid Sequence
H2B-Ava^N- Flag Fusion	PEPAKSAPAPKKGSKKAVTKAQKKGGKKRKRSRKESYSIYVYKVLKQV HPDTGISSKAMGIMNSFVNDIFERIAGEASRLAHYNKRSTITSREIQTAVR LLLPGELKHAVSEGTKAITKYTSSKGGKFAEYCLSYDTEVLTVEYGFVPI GEIVDKGIECSVFSIDSNIGIVYTQPIAQWHHRGKQEVFEYCLEDGSIKAT KDHKFMTQDGKMLPIDEIFEQELDLLQVKGLPEGDYKDDDDK
Npu^C-Cargo Peptide	IKIATRKYLGKQNVYDIGVERC ^(QSY) HNFALKNGFIASNCFNK ^(TMR) YPY DVPDYA
HA2-TAT CPP	C ^(Npys) GLFEAIAEFIENGWEGLIEGWYGGRRKRRQRRR

Table 2.

Antibodies used for western immunoblotting analysis

Probe Order	Host	Epitope	Dilution Factor
Primary	Mouse	Anti-Flag	1:1,000
Primary	Rabbit	Anti-HA	1:1,000
Secondary	Goat	Anti-Mouse-680	1:15,000
Secondary	Goat	Anti-Rabbit-800	1:15,000