



Cite this: DOI: 10.1039/d6sc01808d

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 3rd March 2026  
Accepted 22nd April 2026

DOI: 10.1039/d6sc01808d

rsc.li/chemical-science

## Traceless protein semisynthesis in cells using the promiscuous ultra-fast split intein NrdJ-1

Xuanjia Ye,<sup>1</sup> Joshua Sokol,<sup>1</sup> Christian Kofoed,<sup>1</sup> Anushka Dheer, Juner Zhang<sup>1</sup> and Tom W. Muir<sup>1</sup>\*

Chemistry-driven approaches that allow proteins to be manipulated in ways not permitted by standard genetics are indispensable in the basic and applied biomedical sciences. Of the various platform technologies in common use, those that exploit intein-mediated protein splicing have proven especially powerful owing to compatibility with both *in vitro* and *in vivo* applications. Here, we provide detailed biochemical characterization of the split intein, NrdJ-1. We demonstrate that the rapid splicing kinetics associated with this system are largely independent of splice junction primary sequence context. The promiscuity and high efficiency of this split intein makes it well suited to the generation of chemically modified proteins *in vitro* and in living cells, as demonstrated by the traceless generation of semisynthetic chromatin harboring various post-translational modifications. These studies highlight the utility of NrdJ-1 for traceless protein editing and suggest it is a superior choice for many intracellular biochemical applications.

### Introduction

Understanding the biochemical roles of post translational modifications (PTMs) remains an extremely challenging undertaking.<sup>1–3</sup> Unambiguous assignment of function requires the ability to introduce a given PTM site-specifically into an otherwise native protein sequence, something that cannot be achieved using standard genetic approaches. Several complementary chemistry-driven methods have been developed in recent decades that allow the incorporation of a wide range of non-canonical amino acids (ncAAs), including many PTMs, into proteins.<sup>4–6</sup> Approaches such as genetic code expansion, chemical protein (semi)synthesis and sidechain-directed conjugations have become important weapons in the biomedical research arsenal, allowing incorporation of biochemical probes and PTMs into those proteins accessible to the methods. Among these approaches, protein semisynthesis, a strategy involving the assembly of recombinant with chemically synthesized polypeptide fragments, has emerged as especially powerful for the study of PTMs due to the range of chemotypes that can be introduced at one or more discrete positions in a protein.<sup>4</sup>

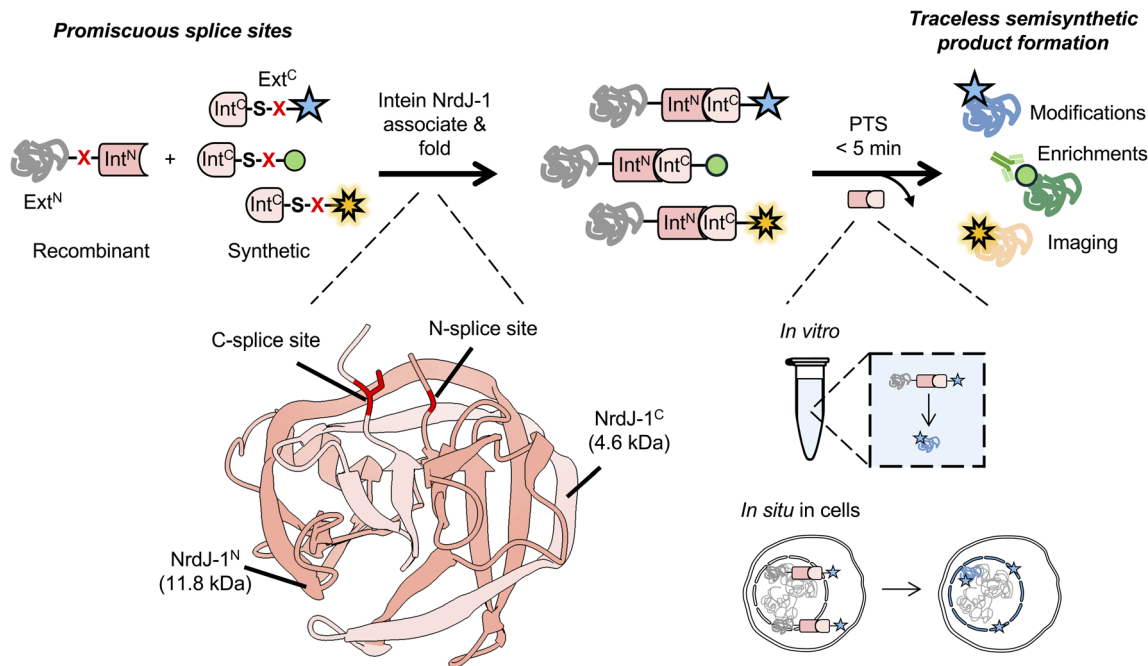
While protein semisynthesis is widely used in the *in vitro* (herein defined as test-tube biochemistry) analysis of PTMs – for example, we and others have introduced numerous modifications into histone proteins for biochemical studies<sup>7</sup> – methods such as native chemical ligation (NCL) and expressed protein ligation (EPL) cannot be directly implemented in cells due to the requirement for high reactant concentrations and the

instability of reaction intermediates under physiological conditions.<sup>8,9</sup> By contrast, chemo-enzymatic strategies offer greater flexibility in terms of compatibility with both *in vitro* and cellular applications.<sup>4</sup> Among these, split intein-mediated protein *trans*-splicing (PTS) is especially attractive due to the high specificity and autocatalytic nature of the process.<sup>4,10,11</sup> In PTS, complementary N- and C-terminal split inteins (designated Int<sup>N</sup> and Int<sup>C</sup>) associate and mediate the ligation of the polypeptides they are fused to which, by convention, are referred to as the exteins. If one of the extein sequences is a synthetic peptide then the resulting PTS reaction results in the semisynthesis of a protein (Fig. 1). Split intein-mediated PTS have been used in numerous protein engineering applications *in vitro*, in living cells and *in vivo*.<sup>10,12,13</sup>

Since PTS involves an excision-ligation reaction, the process has the potential to return a final protein product with few if any alterations to the native primary sequence. This differentiates the approach from the commonly used enzymatic ligation method based on bacterial sortases (*i.e.* sortagging) which requires short peptide sequences in both reactants and thus typically leaves behind a ‘ligation scar’ in the product.<sup>14</sup> In practice, inteins have evolved to efficiently splice in a native extein context, meaning that they often have local sequence preferences either side of the splice junction.<sup>15,16</sup> Moreover, PTS minimally requires a Cys, Ser or Thr residue immediately adjacent to the C-intein fragment which places further constraints on the choice of ligation sites in a target protein sequence.<sup>4</sup> Deviations from these optimal sequences can perturb the alignment of key residues in the intein splicing center leading to the accumulation of side-products and slower splicing kinetics.<sup>16</sup> Even in an optimal extein context, different

Department of Chemistry, Princeton University, Princeton, NJ, USA. E-mail: muir@princeton.edu





**Fig. 1** Schematic of NrdJ-1 mediated protein *trans*-splicing (PTS), featuring promiscuous splice sites and traceless, ultrafast ligation both *in vitro* and in cells. NrdJ-1 fragments (Int<sup>N</sup>, Int<sup>C</sup>) spontaneously associate and fold to initiate ligation of the adjoining exteins upon mixing. Minimally, a serine residue is required at the splice junction. For protein semisynthesis, a synthetic fragment bearing user-defined chemical functionalities (indicated by colored stars/circle) is ligated to a recombinant protein of interest. This approach enables installation of diverse chemical modifications, enrichment handles, or imaging probes and can be performed either *in vitro* or *in situ* within the native cellular environment. Red X: a variable residue at the splice junction. Structure of the folded NrdJ-1 intein is shown (PDB: 8UBS).

split inteins support PTS with varied kinetics. In the best cases, the half-life for PTS can be on the order of seconds, making these ‘ultrafast’ split inteins particularly appealing for protein engineering applications.<sup>4</sup>

Of the ultrafast split inteins reported to date, the NrdJ-1 system, first identified in a marine metagenomic study, is especially attractive for protein semisynthesis.<sup>17–22</sup> While most characterized ultrafast split inteins require a cysteine at the first C-extein position, NrdJ-1 utilizes a serine (Fig. S1), thereby offering expanded flexibility in selecting splice sites due to the substantially higher abundance of serine (~8–9%) compared to cysteine (~2–3%) in the human proteome.<sup>23</sup> Moreover, several recent studies have hinted that NrdJ-1 might have relaxed extein dependence,<sup>21,22,24</sup> although a systematic evaluation of this is yet to be reported. In this study, we provide a detailed characterization of NrdJ-1, including an assessment of its local extein sequence tolerance. These studies reveal NrdJ-1 to be a remarkably promiscuous split intein and provide insights into the underlying basis of its ultrafast splicing kinetics. Motivated by these studies, we show that this split intein can be used for traceless protein semisynthesis *in vitro* and in cells, with a particular focus on editing the highly modified N-terminal tail of histone H3.

## Result

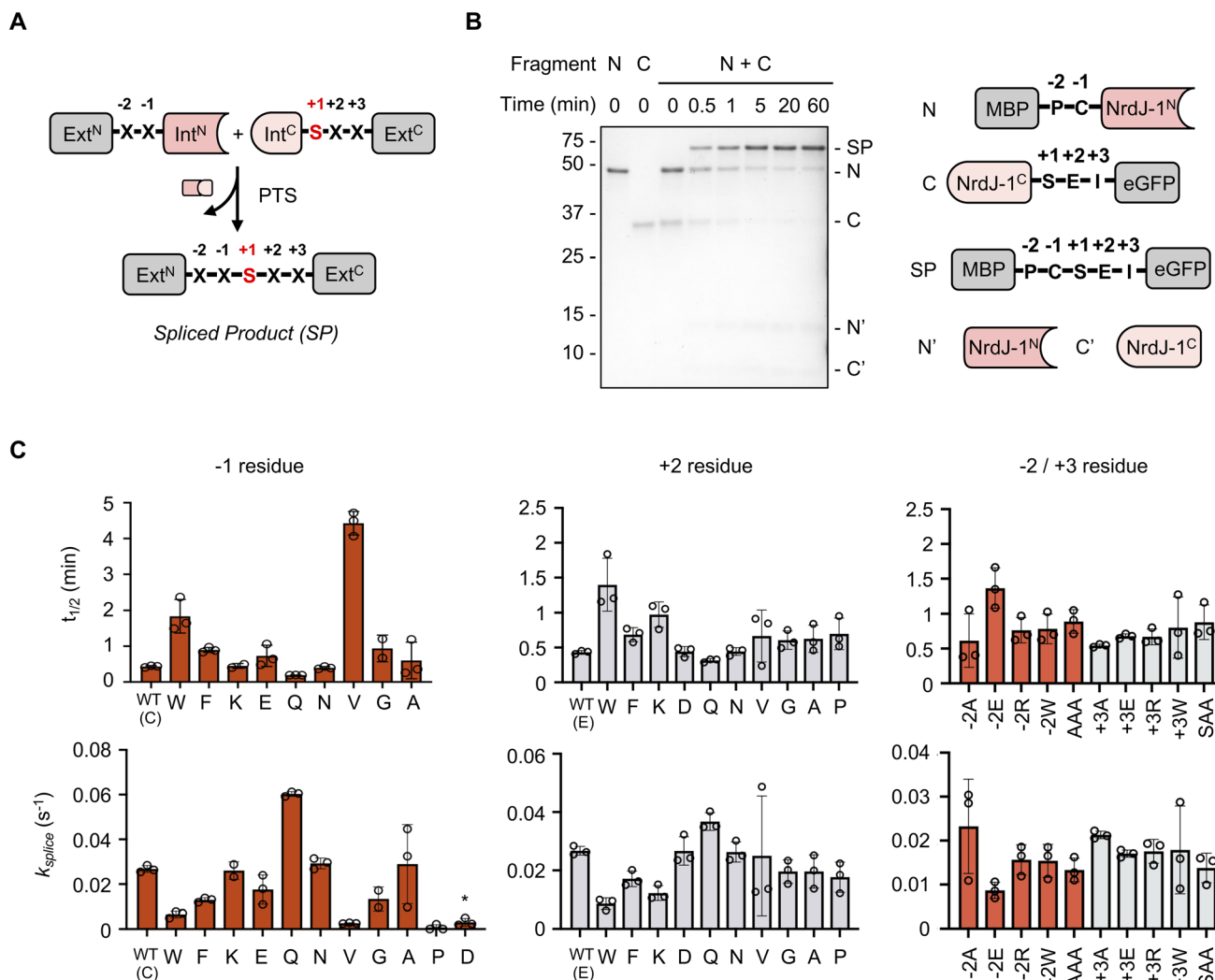
### NrdJ-1 exhibits low extein dependence

To explore the extein dependence of NrdJ-1, we first generated a series of recombinant protein constructs of the general

structure, MBP-X<sub>2</sub>X<sub>1</sub>-NrdJ-1<sup>N</sup> and NrdJ-1<sup>C</sup>-X<sub>1</sub>X<sub>2</sub>X<sub>3</sub>-eGFP, where MBP and eGFP are model N- and C-extein proteins, respectively, and X<sub>n</sub> denotes a variable amino acid residue on either side of the splicing junction (Fig. 2A). We focused on four positions known to influence intein activity: –1 and –2 on the N-extein, and +2 and +3 on the C-extein.<sup>15,16,19,25</sup> Note, the +1 position in the C-extein was kept constant as a serine, since this is required for NrdJ-1 mediated PTS. Following bacterial expression, these proteins were purified and characterized by LC-MS (Fig. S2). As a benchmark, we first determined the splicing rate of NrdJ-1 in a native extein context. The requisite proteins were combined (1 μM each) at 37 °C and the reaction progress was monitored over time by SDS-PAGE and LC-MS (Fig. 2B, S2A, S4A and SI Table 1). Under these conditions, NrdJ-1 supported rapid and highly efficient splicing with an apparent first order rate constant  $k_{\text{splice}} = 0.027 \pm 0.001 \text{ s}^{-1}$ , corresponding to a  $t_{1/2} = 26.0 \pm 1.2 \text{ s}$ , confirming its ultrafast kinetics and high catalytic efficiency under physiological conditions.<sup>17</sup>

Next, we assessed the splicing activities of the constructs containing mutations in the variable positions at the splice junction (Fig. 2C, S3 and SI Table 1). The mutations were designed to explore the impact of altering the charge (*e.g.* Lys, Glu, Asp), size (*e.g.* Ala, *vs.* Trp/Phe) and conformation properties (Pro, Gly) of the residues on splicing efficiency. These studies revealed NrdJ-1 possesses remarkable extein tolerance, with high spliced product yields and fast kinetics observed for most mutants explored. The exceptions to this were Asp and Pro





**Fig. 2** NrdJ-1 exhibits low extein dependence. (A) Schematic showing the splicing junction of PTS. Ext<sup>N/C</sup> and Int<sup>N/C</sup> refer to the N- and C-terminal exteins and split inteins, respectively. Extein residues flanking the splice site are indicated. For NrdJ-1 the +1 position is an obligate serine (red). (B) Time course of NrdJ-1 PTS reaction featuring native extein residues at the splice junction. MBP-NrdJ-1<sup>N</sup> and NrdJ-1<sup>C</sup>-eGFP were combined (1  $\mu$ M of each reactant) in splicing buffer (pH 7.2) at 37  $^{\circ}$ C and the reaction progress was monitored by SDS-PAGE with Coomassie staining. Identity of the spliced product (SP) was confirmed by mass spectrometry (see Fig. S2). The schematic on the right depicts the starting materials and corresponding end products. (C) Impact of varying indicated N- and C-extein residues on the kinetics. Reactions were performed and monitored as in panel B, and the extent of product determined by gel densitometry analysis (details in Material and Methods). The half-lives (upper panel) and apparent first order rate constants (lower panel) for the reactions are presented. \* denotes the side product yield when Asp is at the -1 position (see Fig. S4). Error bars represent SD ( $n = 2-3$ ).

at the -1 position, which significantly compromised reaction fidelity (Fig. 2C, S3, S4B and C). Both these residues are well known to be incompatible with intein-mediated protein splicing;<sup>4,16</sup> Asp promotes premature N-extein cleavage through a cyclization/hydrolysis process (Figure S4C-F), whereas the presence of a -1 Pro disrupts the first step in splicing involving an N-to-S(O) acyl shift. Notably, Pro was found to be compatible when placed at the +2 extein position (Fig. 2C and S3B), which is not the case for many other ultrafast inteins such as the commonly used DnaE family of split inteins.<sup>25</sup> Further underlining the flexibility of the NrdJ-1 system, we found that efficient splicing was supported when multiple mutations were made at the N- (AAA) and C-splice junctions (SAA) (Fig. 2C and S3C).

Together these studies reveal NrdJ-1 to have exceptionally low extein-dependence, making it an excellent candidate for further engineering and applications.

### Mechanisms of NrdJ-1 rapid kinetics

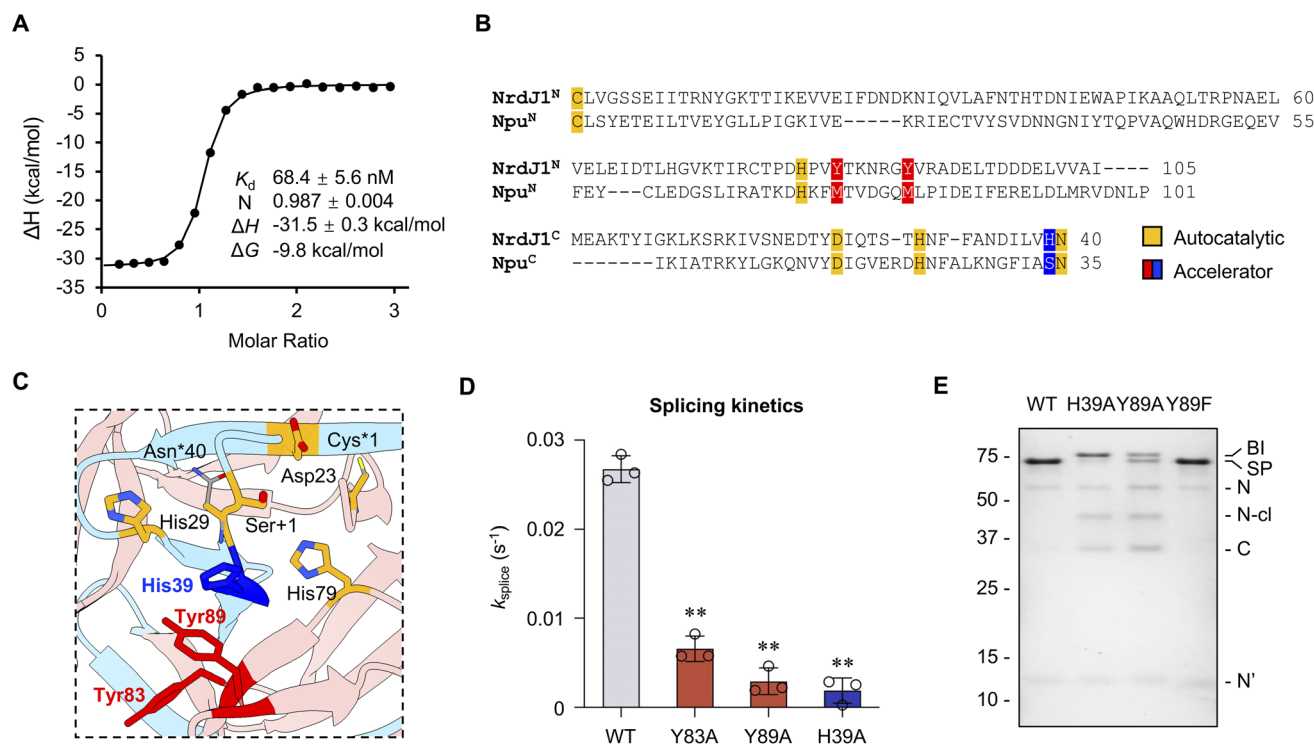
We next turned to the underlying basis of the fast-splicing kinetics associated with NrdJ-1. The first step in PTS is the association of the Int<sup>N</sup> and Int<sup>C</sup> fragments (Fig. S1). Naturally split inteins often exhibit high affinity between the fragments, a property that is desirable for cellular applications where low concentrations of the reactants are present.<sup>26</sup> Since the binding affinity of NrdJ-1 fragments has not previously been reported, we performed isothermal titration calorimetry (ITC) using MBP-



NrdJ-1<sup>N</sup> and NrdJ-1<sup>C</sup>-SUMO constructs containing mutations that block splicing chemistry but not fragment association (C1A in the N-intein and N145A in the C-intein, Fig. S5A and S5B). These fusion partners were selected to mimic the interaction environment of NrdJ-1 when appended to proteins in cellular applications, rather than measuring the affinity of the isolated intein fragments alone. For this experiment, native NPC and SEI extein sequences were retained at the splice junctions. This analysis revealed tight association between the fragments, with ~1 : 1 stoichiometry and a dissociation constant ( $K_d$ ) of 68.4 ± 5.6 nM (Fig. 3A and S5C).

The splicing activity of split inteins relies on a conserved network of autocatalytic residues organized within the HINT (Hedgehog/Intein) fold<sup>10</sup> (Fig. S6A). In addition to this well-defined autocatalytic core, previous studies on the DnaE family of split inteins have identified a class of auxiliary positions, referred to as “accelerator” residues, that enhance splicing efficiency.<sup>27</sup> These residues are located in the second coordination shell relative to the autocatalytic core and have been shown to promote rapid splicing by stabilizing key

transition states or intermediates. To identify potential “accelerator” residues in NrdJ-1, we performed a multiple sequence alignment (MSA) with the well-characterized ultrafast DnaE intein, Npu.<sup>27</sup> As expected, this revealed strong conservation of the canonical autocatalytic residues, consistent with the shared structural homology of the two split inteins (Fig. 3B and S6B). Given this similarity, we hypothesized that NrdJ-1 might possess functional counterparts to the accelerator residues previously identified in Npu. Intriguingly, at the positions corresponding to the Met and Ser accelerator residues in Npu, NrdJ-1 instead contains aromatic residues (Fig. 3B). Structural mapping revealed that these residues lie adjacent to the autocatalytic core and occupy spatially analogous positions to the established accelerator sites in Npu (Fig. 3C and S6B). To test their functional roles, we introduced alanine substitutions at Tyr83 and Tyr89 in the NrdJ-1<sup>N</sup> fragment and His39 in NrdJ-1<sup>C</sup> (Fig. S7A). All three mutants displayed significant reductions in splicing kinetics relative to the wildtype intein, indicating that these residues contribute to efficient autocatalysis (Fig. 3D, S8A and B). Consistent with their spatial proximity to the



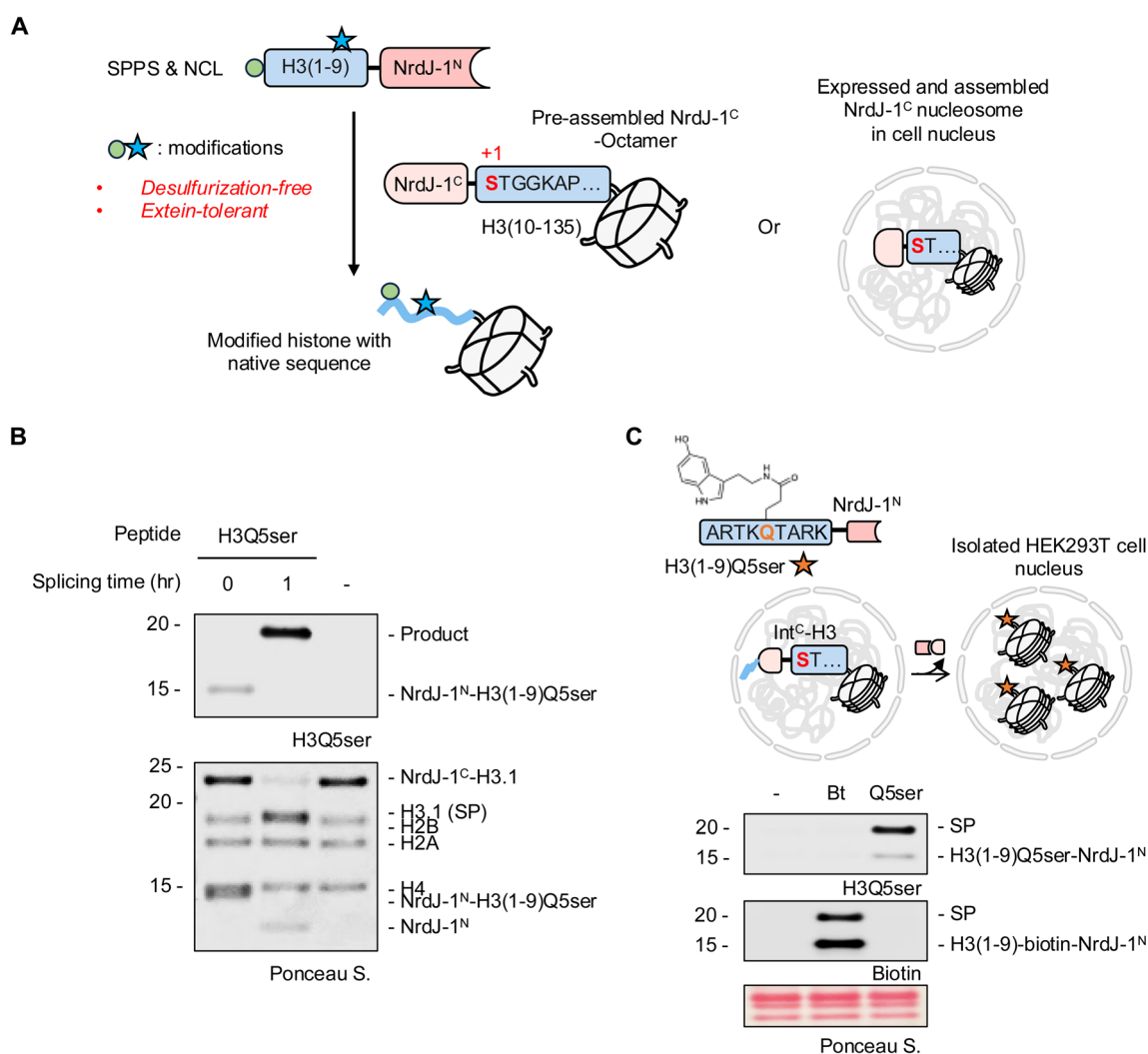
**Fig. 3** ‘Accelerator’ residues facilitate succinimide formation in the splicing reaction. (A) Thermogram and thermodynamic parameters of NrdJ-1 fragment association. Analysis was performed by ITC using MBP-NrdJ-1<sup>N</sup> and NrdJ-1<sup>C</sup>-SUMO bearing inactivating mutations (C1A and N145A).  $K_d$ : dissociation constant; N: stoichiometry;  $\Delta H$ : enthalpy change;  $\Delta G$ : Gibbs free energy change. (B) Multiple sequence alignment of the intein fragments of NrdJ-1 with a representative fast-splicing split intein, Npu. Conserved residues involved in autocatalysis and identified accelerators are highlighted. (C) Structure of the NrdJ-1 split intein showing the spatial positions of conserved autocatalytic residues (Cys1, Asp23, His29, Asn40, Ser +1, highlighted in orange) and the candidate accelerator residues (Tyr83, Tyr89 within NrdJ-1<sup>N</sup> and highlighted in red and His39 within NrdJ-1<sup>C</sup> and colored blue). The regions corresponding to the N- and C-intein fragments in the split version are colored pink and blue, respectively. The structure is generated from an engineered fused version of NrdJ-1 bearing inactivating mutations: C1A and N145A. The side chains of Cys1 and Asn145 were modeled *in silico* using Chimera rotamer simulations. PDB: 8UBS. (D) *In vitro* PTS kinetics with wildtype NrdJ-1 and corresponding accelerator mutants.  $k_{splice}$  represents first-order kinetic fits of splice product formation (see Materials and Methods in SI). Reactions were carried out as in Fig. 2 and products quantified from densitometry analysis of SDS-PAGE analyses of time courses. Error bar indicates SD (n = 3); \*\*: p ≤ 0.01. (E) PTS reactions employing wild-type or indicated mutant PTS mixtures. Reactions were quenched at 30 min and visualized by SDS-PAGE followed by Coomassie staining. BI: branched intermediate; SP: spliced product; N: MBP-NrdJ-1<sup>N</sup>; N-cl: N-terminal cleavage; C: NrdJ-1<sup>C</sup>-eGFP; N': NrdJ-1<sup>N</sup> after splicing.



autocatalytic core, alanine mutations at His39 and Tyr89 exerted a stronger impact on splicing efficiency than the Y83A mutation (Fig. 3C, D, S8A and B).

To probe the mechanistic role of the putative “accelerator” residues, we further analyzed the reaction products obtained when using alanine mutants. SDS-PAGE analysis of the reaction mixtures revealed that the H39A and Y89A variants accumulated a distinct high molecular-weight species not observed in the wildtype intein (Fig. 3E and S8B). Subsequent analysis by RP-HPLC and mass spectrometry identified this species as the branched intermediate (BI) wherein the N-extein is linked to the +1 serine through an ester linkage (Fig. S1A, S7B and S8C). This intermediate is on the canonical reaction pathway for protein

splicing and its resolution through succinimide formation is typically the rate-limiting step in the process.<sup>28</sup> The accumulation of BI suggests a defect in this step. Notably, replacement of Tyr89 with a phenylalanine did not result in BI accumulation (Fig. 3E, S8D and E), suggesting that the presence of an aromatic sidechain at this position is critical for the resolution of this key splicing intermediate. Inspection of the crystal structure of NrdJ-1 reveals that Tyr89 packs against both Tyr83 and His39 (Fig. S6B). Importantly, His39 is immediately adjacent to Asn40 in NrdJ-1<sup>C</sup>, the residue that directly participates in succinimide formation. Previously, we have shown that even minor perturbations to the positioning of the Asn side chain perturbs this reaction.<sup>16,28</sup> Thus, it seems reasonable to propose



**Fig. 4** Traceless installation of H3Q5ser into chromatin using NrdJ-1. (A) Schematic of histone H3 semisynthesis using NrdJ-1-mediated PTS. NrdJ-1<sup>N</sup> fused to a synthetic PTM-bearing H3 peptide fragment (in the example shown residues 1–9) is reacted with the remainder of H3 (residues 10–135) fused to the complementary NrdJ-1<sup>C</sup> intein. PTS leads to the traceless semisynthesis of modified H3, exploiting the ability of NrdJ-1 to utilize native serine residues. In principle, this workflow can either work *in vitro* on pre-assembled histone octamers or in isolated cell nuclei containing endogenously expressed chromatin-associated NrdJ-1<sup>C</sup> histone fusion. (B) *In vitro* PTS reaction employing H3(1–9, Q5ser)-NrdJ-1<sup>N</sup> and NrdJ-1<sup>C</sup>-H3(10–135) containing histone octamers. Reaction mixture was analyzed at indicated timepoint by western blot (top) and SDS-PAGE (bottom). (C) Top: schematic of the experimental design. The chemically modified substrate H3(1–9)-NrdJ-1<sup>N</sup> containing either a biotin handle (Bt, serve as splicing control) or Q5ser was incubated with isolated HEK293T nuclei transiently expressing Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135). Protein *trans*-splicing excises the N-terminal Flag-H3(1–28)-NrdJ-1<sup>C</sup> and ligates the modified H3(1–9) fragment to H3(10–135), installing biotin or Q5ser on chromatin; (Bottom): western blot analysis of nuclear extracts after 1 h incubation at 37 °C. SP, spliced product.



that disruption of the Tyr83-Tyr89-His39 network would interfere with this step in the splicing reaction. More generally, tuning the activity of split inteins through manipulation of 'accelerator residues' was key to the recent development of a protein editing technology that involves the activity of matched pairs of orthogonal split inteins.<sup>29</sup> We expect that the above insights into the NrdJ-1 splicing are likely to be useful for its integration into this protein engineering manifold.

### Traceless serotonylation on histone H3 tails

The biochemical features of NrdJ-1, its efficiency and broad extein tolerance, motivated us to evaluate its utility across multiple contexts. First, we sought to perform *in vitro* protein semisynthesis involving installation of PTMs with the view to subsequently extend the workflow to cellular applications. Histone H3 represents an ideal target for such efforts: its N-terminal tail is densely modified with regulatory PTMs that influence chromatin structure and gene expression.<sup>2</sup> Importantly, the native H3 tail lacks cysteine, rendering cysteine-dependent PTS strategies non-traceless (Fig. S9). Although cysteines can be introduced and later converted to alanine *via* radical desulfurization, this approach is incompatible with many nCAAs and several PTMs. By contrast, the first 35 residues of H3 contain three native serine residues (S10, S28, and S31) that can serve directly as splice junctions for PTM conjugation by PTS (Fig. 4A and S9). Thus, NrdJ-1 offers a means to achieve scarless ligation at native serine sites, eliminating the need for desulfurization and substantially expanding the positional and chemical flexibility of histone semisynthesis.

Protein serotonylation involves the covalent attachment of serotonin, an essential neurotransmitter and hormone, to glutamine sidechains and is mediated by the enzyme transglutaminase 2 (TG2).<sup>30</sup> Although discovered several decades ago, the diverse biological roles of protein serotonylation have only recently begun to emerge.<sup>31–34</sup> In particular, histone serotonylation at H3Q5 (H3Q5ser) is enriched in the gut and brain and has been implicated in establishing a permissive transcriptional state in serotonergic neurons.<sup>32</sup> While access to homogeneous preparations of H3Q5ser would greatly facilitate ongoing studies into the precise biochemical functions of this PTM, the incompatibility of serotonin with standard NCL/desulfurization workflows prevents the generation of truly authentic versions of this modified protein.<sup>31</sup> In principle, NrdJ-1-mediated semisynthesis provides a solution to this problem since PTS can exploit a native serine in the H3 N-terminal tail as a potential splice junction, thereby obviating the need for the problematic desulfurization step.

While histone H3 contains multiple native serine residues, we selected Ser10 as the splice junction for synthetic convenience (Fig. 4A and S9). H3(1–9) peptides bearing Q5ser were prepared by solid phase peptide synthesis (SPPS), converted to thioesters, and ligated to NrdJ-1<sup>N</sup> *via* NCL (Fig. S10). The cysteine at position 76 in NrdJ-1<sup>N</sup>, previously shown to be dispensable for catalysis<sup>22</sup> (Fig. S11), was mutated to alanine to minimize disulfide formation; this variant was used for all subsequent semisynthesis unless otherwise noted. Rather than

performing the PTS reaction on the isolated H3 fragments, we decided to explore whether the reaction could be performed at the level of histone octamers since this would, in principle, streamline the manufacture of chromatin for downstream *in vitro* studies. The purified NrdJ-1<sup>C</sup>-H3(10–135) construct was combined with recombinant histones H2A, H2B and H4 and then folded into octamers using established protocols (Fig. S12A–C). Combining these purified histone octamers with H3(1–9, Q5ser)-NrdJ-1<sup>N</sup> led to the efficient generation of the desired modified histone H3 as indicated by western blotting and LC-MS (Fig. 4B and S12D). The success of the reaction is especially notable since the splicing buffer it was performed in contained 2 M NaCl, which is required for the stability of the histone octamer complex in the absence of DNA. The tolerance of the NrdJ-1 to such conditions further underlines the utility of this split intein system for protein semisynthesis.

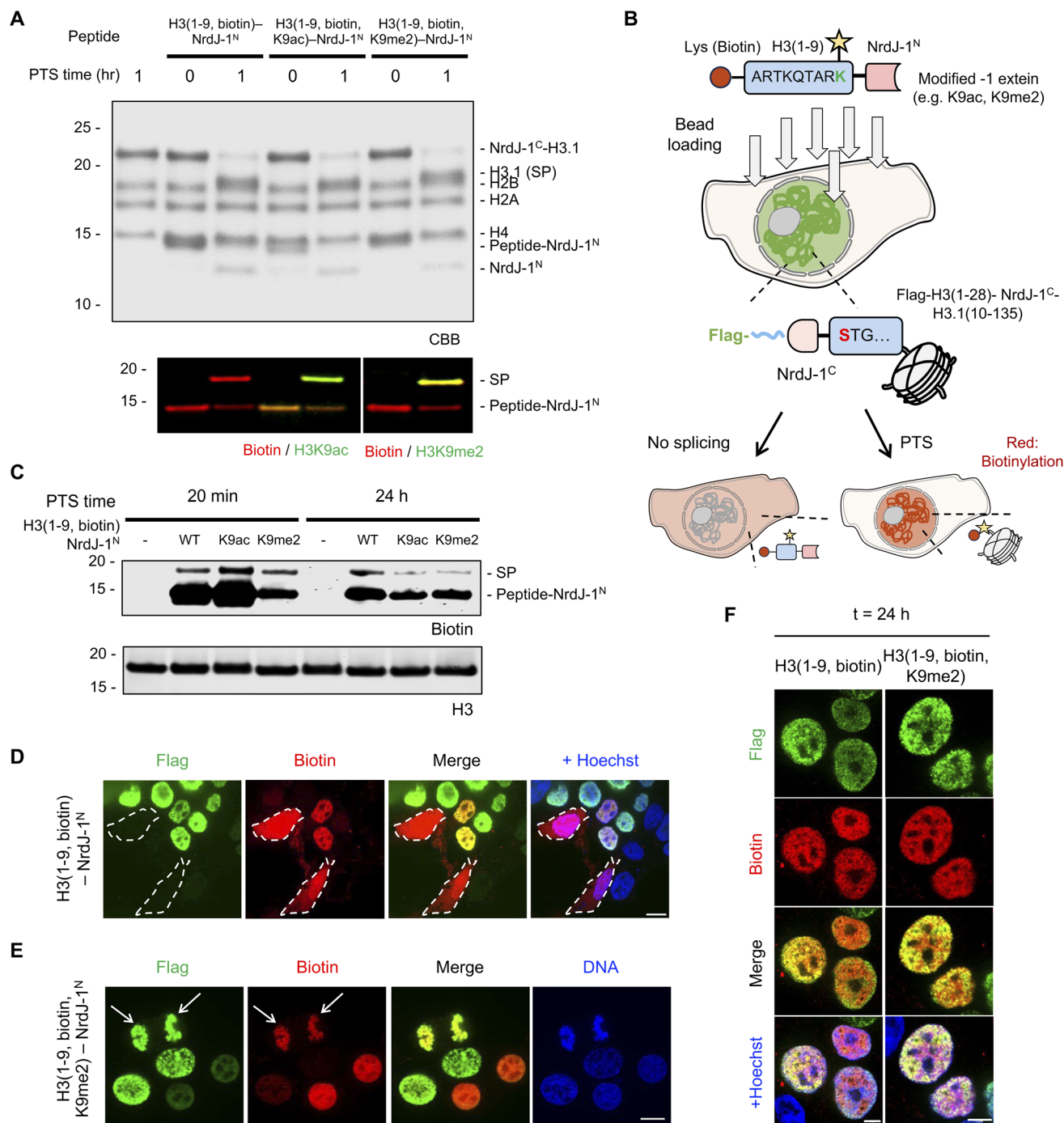
To explore whether NrdJ-1-mediated histone editing could be conducted in a native chromatin setting, we expressed a histone H3.1 split intein fusion, Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135), in HEK293T cells. The H3(1–28) segment was included in this construct to enhance protein expression and chromatin incorporation.<sup>25</sup> Nuclei from these cells were isolated using established procedures<sup>35</sup> and then treated with either a H3(1–9)-NrdJ-1<sup>N</sup> fusion containing either an N-terminal biotin handle (Fig. S13A), or the previously described H3(1–9, Q5ser)-NrdJ-1<sup>N</sup> fusion. Following incubation for 1 h, the *in nucleo* PTS reactions were analyzed by western blotting (Fig. 4C) which revealed the generation of the expected semisynthetic histones. Note, HEK293T cells do not contain endogenous H3Q5ser levels, providing a clean background for assessing site-specific editing in this example. The success of the experiment represents the first traceless semisynthesis of a modified histone in a native chromatin context, an advance that will serve as a foundation for future studies gearing towards better understanding the biochemical functions of this PTM.

### Traceless histone H3 semisynthesis in live cells

Having established NrdJ-1-mediated installation of H3Q5 serotonylation *in vitro* and *in nucleo*, we next sought to (i) explore the scope of histone PTMs accessible to the system and (ii) apply this strategy in live cells to enable temporally controlled synthesis of PTM-modified histones. To this end, we selected three modifications, biotin, H3K9ac and H3K9me2, to evaluate the generality of the approach (Fig. S13A–C). In addition to expanding beyond endogenous PTMs, biotin provides a versatile biochemical handle for visualization and affinity enrichment. H3K9me2 is an important mark for genome organization, whereas H3K9ac is associated with active transcription;<sup>36–38</sup> both modifications play key roles in chromatin regulation. Importantly, the Lys9 residue occupies the –1 position relative to the splice junction, providing an additional testbed for extein tolerance (Fig. S13D). To the best of our knowledge, PTS with a chemically modified –1 extein has not been previously demonstrated.

Consistent with our earlier findings, NrdJ-1 readily accommodated these sequence contexts both in pre-assembled





**Fig. 5** Semisynthesis of modified histone H3 in live cells. (A) *In vitro* PTS reactions on histone octamers containing NrdJ-1<sup>C</sup>-H3.1(10–135). Purified octamers were reacted with either H3(1–9, biotin)-NrdJ-1<sup>N</sup>, H3(1–9, biotin, K9ac)-NrdJ-1<sup>N</sup> or H3(1–9, biotin, K9me2)-NrdJ-1<sup>N</sup>. Reaction mixtures were analyzed at indicated timepoint by western blot (bottom) and SDS-PAGE (top). SP, spliced product. (B) Schematic of live-cell PTS via bead loading of modified extein substrates. Cells lacking Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135) expression display diffuse biotin signal (red) distributed throughout the nucleus and cytosol after bead loading. In contrast, cells expressing Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135) exhibit nuclear colocalization of Flag (green) and biotin (red) as a result of productive chromatin-associated PTS. (C) Live cell PTS reactions monitored by western blotting. Transfected HEK293T cells were bead loaded with indicated H3(1–9)-NrdJ-1<sup>N</sup> constructs and then analyzed for biotin content by streptavidin blotting at indicated timepoints. (D and E) Confocal microscopy of HEK293T cells transiently expressing Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135). Biotinylated H3(1–9)-NrdJ-1<sup>N</sup> (D) or biotinylated H3(1–9, K9me2)-NrdJ-1<sup>N</sup> (E) peptides were introduced into the cultured cells by bead loading 24 h post-transfection. Cells were fixed 20 min post addition of peptides. Immunofluorescence staining was performed with anti-Flag (green) and neutravidin (red). In panel D, two non-transfected cells are outlined with dashed lines. DNA is stained with Hoechst (blue). Arrows indicate a cell in anaphase. Scale bar: 10  $\mu$ m. (F) Confocal microscopy of HEK293T cells expressing Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135). Biotinylated H3(1–9)-NrdJ-1<sup>N</sup> or biotinylated H3(1–9, K9me2)-NrdJ-1<sup>N</sup> peptides were introduced into cultured cells by bead loading 24 h post-transfection. Immunostaining was performed as described in panel C but after 24 h. Scale bar: 5  $\mu$ m.



octamers and in isolated nuclei expressing the NrdJ-1<sup>C</sup>-H3.1(10–135) fusion (Fig. 5A and S14). These results further demonstrate the flexibility of NrdJ-1 with respect to the choice of residues at the splice junction. We next sought to extend this approach to living cells. To deliver the synthetic intein fragments, we adopted a previously reported bead-loading method in which transient mechanical disruption of the cell membrane enables uptake of macromolecules (Fig. S15A).<sup>39,40</sup> We estimate that ~10–15% of cells take up an exogenously added eGFP protein following bead-loading (Fig. S15B). As an initial test of this peptide delivery approach, we transiently expressed an mCherry–Lamin A–NrdJ-1<sup>N</sup> construct in HEK293T cells and then introduced purified NrdJ-1<sup>C</sup>-eGFP *via* bead-loading (Fig. S15C). Gratifyingly, we observed localized GFP signal at the nuclear periphery only in those cells expressing the Lamin A–NrdJ-1<sup>N</sup> fusion, whereas no GFP signal was detected in the negative controls without protein or bead loading processing. (Fig. S15D and S15E).

Encouraged by these results, we delivered modified H3(1–9)-NrdJ-1<sup>N</sup> conjugates bearing PTMs into cells expressing the complementary NrdJ-1<sup>C</sup>-histone fusion and evaluated splicing after 20 min incubation using the biotin detection handle (Fig. 5B). Western blotting revealed successful PTS-mediated semisynthesis had occurred upon delivery of the H3(1–9, biotin)-NrdJ-1<sup>N</sup>, H3(1–9, biotin, K9ac)-NrdJ-1<sup>N</sup> and H3(1–9, biotin, K9me2)-NrdJ-1<sup>N</sup> (Fig. 5C). Moreover, immunofluorescence revealed robust nuclear signals only in those cells expressing Flag-H3(1–28)-NrdJ-1<sup>C</sup>-H3.1(10–135), both for the biotinylated substrate and for the substrate bearing biotin/H3K9me2 dual modifications, whereas no biotin signal was detected in the negative controls (Fig. 5D, E, S16A and B). Notably, we also detected strong biotin enrichment on mitotic chromatin (*e.g.* cells undergoing anaphase in Fig. 5D and S16B), suggesting that modifications installed on chromatin using PTS might be maintained through cell division. To this point, the biotin signal remained detectable 24 h after splicing (Fig. 5C–F and S16C). Overall, these findings establish NrdJ-1-mediated PTS as a practical strategy for modifying endogenous proteins and chromatin in living cells while preserving native primary sequence.

## Conclusions

The discovery of ultrafast split inteins has greatly expanded the scope of PTS in the protein engineering area. Here, we focused on one such split intein, NrdJ-1, focusing on its ability to support protein semisynthesis without the need to introduce facilitating mutations in the target sequence, as often the case when using split inteins.<sup>4,22,29</sup> Our studies show that the exceptional splicing efficiency of NrdJ-1 is largely independent of local extein sequences. This along with its use of a serine at the +1 extein position makes NrdJ-1 well suited to the traceless installation of ncAAs and PTMs into target proteins. As a demonstration of this, we used NrdJ-1 to install various modifications into histone H3 sequence both *in vitro* and in cells. The ability to edit native chromatin without the need to make mutations in the primary sequence opens new avenues

for studying chromatin function. In particular, the rapid kinetics of NrdJ-1 PTS should allow for temporally resolved studies of PTM dynamics, such as monitoring the appearance, disappearance, turnover or redistribution of specific modifications over defined time windows. Importantly, split inteins from different families share only structural homology but not primary sequence identity, so that they do not cross-react.<sup>19,41</sup> This inherent orthogonality enables multiplexed applications in which multiple engineered split inteins, including NrdJ-1, can be used in combination without interference. The current studies, by surveying the scope of NrdJ-1 PTS, are expected to accelerate such efforts.

## Author contributions

X. Y. and T. W. M. conceived the project, analyzed the data and wrote the manuscript. Experiments were performed by X. Y. with the help of J. S., C. K., A. D. and J. Z. T. W. M. supervised the project.

## Conflicts of interest

TWM is a founder and scientific advisory board member of SpliceBio and is a consultant for Merck.

## Data availability

All data are included in the main figures or supplementary information (SI). References cited in the SI<sup>42–44</sup> have been listed at the bottom of the main article's reference list. Supplementary information: detailed materials and methods. See DOI: <https://doi.org/10.1039/d6sc01808d>.

## Acknowledgements

We thank member of the Muir lab for valuable discussions. This work was supported by the National Institutes of Health (NIH-GMS grant R01 GM086868) and by funds from the Ludwig Institute for Cancer Research and the Princeton Catalysis Initiative. CK was supported by an EMBO postdoctoral fellowship (ALTF 1189-2020). JZ is a Damon Runyon Fellow supported by the Damon Runyon Cancer Research Foundation (DRG-2506-23).

## References

- 1 C. T. Walsh, S. Garneau-Tsodikova and G. J. Gatto, *Angew. Chem., Int. Ed.*, 2005, **44**, 7342–7372.
- 2 G. Millán-Zambrano, A. Burton, A. J. Bannister and R. Schneider, *Nat. Rev. Genet.*, 2022, **23**, 563–580.
- 3 L. Murgas, G. Pollastri, E. Riquelme, M. Sáez and A. J. M. Martin, *Briefings Bioinf.*, 2025, **26**(1), bbae638.
- 4 R. E. Thompson and T. W. Muir, *Chem. Rev.*, 2020, **120**, 3051–3126.
- 5 J. W. Chin, *Nature*, 2017, **550**, 53–60.
- 6 P. G. Isenegger and B. G. Davis, *J. Am. Chem. Soc.*, 2019, **141**, 8005–8013.



- 7 N. Hananya, S. Koren and T. W. Muir, *Nat. Rev. Genet.*, 2024, **25**, 255–271.
- 8 P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776–779.
- 9 T. W. Muir, D. Sondhi and P. A. Cole, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6705–6710.
- 10 N. H. Shah and T. W. Muir, *Chem. Sci.*, 2013, **5**, 446–461.
- 11 I. Giriat and T. W. Muir, *J. Am. Chem. Soc.*, 2003, **125**, 7180–7181.
- 12 D. W. Wood, M. Belfort and C. W. Lennon, *Front. Microbiol.*, 2023, **14**, 1305848.
- 13 S. Anastassov, M. Filo and M. Khammash, *Biotechnol. Adv.*, 2024, **73**, 108349.
- 14 J. M. Antos, M. C. Truttmann and H. L. Ploegh, *Curr. Opin. Struct. Biol.*, 2016, **38**, 111–118.
- 15 A. J. Burton, M. Haugbro, E. Parisi and T. W. Muir, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 12041–12049.
- 16 N. H. Shah, E. Eryilmaz, D. Cowburn and T. W. Muir, *J. Am. Chem. Soc.*, 2013, **135**, 5839–5847.
- 17 P. Carvajal-Vallejos, R. Pallissé, H. D. Mootz and S. R. Schmidt, *J. Biol. Chem.*, 2012, **287**, 28686–28696.
- 18 B. Dassa, N. London, B. L. Stoddard, O. Schueler-Furman and S. Pietrokovski, *Nucleic Acids Res.*, 2009, **37**, 2560.
- 19 H. Tasfaout, C. L. Halbert, T. S. McMillen, J. M. Allen, T. R. Reyes, G. V. Flint, D. Grimm, S. D. Hauschka, M. Regnier and J. S. Chamberlain, *Nature*, 2024, **632**, 192–200.
- 20 F. Pinto, E. L. Thornton and B. Wang, *Nat. Commun.*, 2020, **11**, 1–15.
- 21 G. Lee and T. W. Muir, *Nat. Chem. Biol.*, 2024, **20**, 1353–1360.
- 22 C. Kofoed, G. Erkaló, N. E. S. Tay, X. Ye, Y. Lin and T. W. Muir, *Nature*, 2025, **645**, 793–800.
- 23 D. Brüne, M. A. Andrade-Navarro and P. Mier, *BMC Res. Notes*, 2018, **11**, 117.
- 24 A. Romero-Casañas, A. García-Lizarribar, J. Castro, M. Vilanova, A. Benito and M. Ribó, *Protein Sci.*, 2024, **33**, e5070.
- 25 A. J. Stevens, G. Sekar, N. H. Shah, A. Z. Mostafavi, D. Cowburn and T. W. Muir, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 8538–8543.
- 26 N. H. Shah, M. Vila-Perelló and T. W. Muir, *Angew Chem. Int. Ed. Engl.*, 2011, **50**, 6511.
- 27 A. J. Stevens, Z. Z. Brown, N. H. Shah, G. Sekar, D. Cowburn and T. W. Muir, *J. Am. Chem. Soc.*, 2016, **138**, 2162–2165.
- 28 S. Frutos, M. Goger, B. Giovani, D. Cowburn and T. W. Muir, *Nat. Chem. Biol.*, 2010, **6**, 527–533.
- 29 Y. Hua, N. E. S. Tay, X. Ye, J. A. Owen, H. Liu, R. E. Thompson and T. W. Muir, *Science*, 2025, **388**, 68–74.
- 30 M. Bader, *Front. Mol. Neurosci.*, 2019, **12**, 288.
- 31 S. Zhao, K. N. Chuh, B. Zhang, B. E. Dul, R. E. Thompson, L. A. Farrelly, X. Liu, N. Xu, Y. Xue, R. G. Roeder, I. Maze, T. W. Muir and H. Li, *Proc. Natl. Acad. Sci. U. S. A.*, 2021, **118**, e2016742118.
- 32 L. A. Farrelly, R. E. Thompson, S. Zhao, A. E. Lepack, Y. Lyu, N. V. Bhanu, B. Zhang, Y. H. E. Loh, A. Ramakrishnan, K. C. Vadodaria, K. J. Heard, G. Erikson, T. Nakadai, R. M. Bastle, B. J. Lukasak, H. Zebroski, N. Alenina, M. Bader, O. Berton, R. G. Roeder, H. Molina, F. H. Gage, L. Shen, B. A. Garcia, H. Li, T. W. Muir and I. Maze, *Nature*, 2019, **567**, 535–539.
- 33 J. C. Chan, N. Alenina, A. M. Cunningham, A. Ramakrishnan, L. Shen, M. Bader and I. Maze, *J. Mol. Biol.*, 2024, **436**, 168454.
- 34 B. J. Lukasak, M. M. Mitchener, L. Kong, B. E. Dul, C. D. Lazarus, A. Ramakrishnan, J. Ni, L. Shen, I. Maze and T. W. Muir, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2208672119.
- 35 A. J. Burton, M. Haugbro, L. A. Gates, J. D. Bagert, C. D. Allis and T. W. Muir, *Nat. Chem.*, 2020, **12**, 520–527.
- 36 K. Fukuda, C. Shimura, H. Miura, A. Tanigawa, T. Suzuki, N. Dohmae, I. Hiratani and Y. Shinkai, *Commun. Biol.*, 2021, **4**, 571.
- 37 A. Poleshko, C. L. Smith, S. C. Nguyen, P. Sivaramakrishnan, K. G. Wong, J. I. Murray, M. Lakadamyali, E. F. Joyce, R. Jain and J. A. Epstein, *Elife*, 2019, e49278.
- 38 L. A. Gates, J. Shi, A. D. Rohira, Q. Feng, B. Zhu, M. T. Bedford, C. A. Sagum, S. Y. Jung, J. Qin, M. J. Tsai, S. Y. Tsai, W. Li, C. E. Foulds and B. W. O'Malley, *J. Biol. Chem.*, 2017, **292**, 14456.
- 39 P. L. McNeil and E. Warder, *J. Cell Sci.*, 1987, **88**, 669–678.
- 40 C. A. Cialek, G. Galindo, A. L. Koch, M. N. Saxton and T. J. Stasevich, *J. Vis. Exp.*, 2021, **172**, DOI: [10.3791/62559](https://doi.org/10.3791/62559).
- 41 J. A. Gamespacher, A. J. Stevens, D. P. Nguyen, J. W. Chin and T. W. Muir, *J. Am. Chem. Soc.*, 2017, **139**, 8074–8077.
- 42 G. P. Dann, G. P. Liszczak, J. D. Bagert, M. M. Müller, U. T. T. Nguyen, F. Wojcik, Z. Z. Brown, J. Bos, T. Panchenko, R. Pihl, S. B. Pollock, K. L. Diehl, C. D. Allis and T. W. Muir, *Nature*, 2017, **548**, 607–611.
- 43 N. H. Shah and A. J. Stevens, *Methods Mol. Biol.*, 2020, **2133**, 31–54.
- 44 D. T. Flood, J. C. J. Hintzen, M. J. Bird, P. A. Cistrone, J. S. Chen and P. E. Dawson, *Angew. Chem., Int. Ed.*, 2018, **57**, 11634–11639.

