Intein Zymogens: Conditional Assembly and Splicing of Split Inteins via Targeted Proteolysis

Josef A. Gramespacher,[†] Adam J. Stevens,[†] Duy P. Nguyen,[‡] Jason W. Chin,[‡] and Tom W. Muir*,[†]

[†]Frick Laboratory, Department of Chemistry, Princeton University, Princeton, New Jersey 08544, United States
[‡]Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, U.K.

S *Supporting Information*

ABSTRACT: Naturally split inteins have found widespread use in chemical biology due to their ability to drive the ligation of separately expressed polypeptides through a process termed protein trans-splicing (PTS). In this study, we harness PTS by rendering association of split intein fragments conditional upon the presence of a user-defined protease. We show that these intein "zymogens" can be used to create protein sensors and actuators that respond to the presence of various stimuli, including bacterial pathogens, viral infections, and light. We also show that this design strategy is compatible with several orthogonal split intein pairs, thereby opening the way to the creation of multiplexed sensor systems.

Inteins are intervening protein domains that undergo a spontaneous post-translational process known as protein splicing. During this autoprocessing event, the intein excises itself from a larger precursor polypeptide, in the process ligating the flanking polypeptides (exteins) via a peptide bond (Figure S1a).[1] Unlike the more common contiguous inteins, naturally split inteins such as DnaE from the cyanobacterium *Nostoc punciforme PCC73102* (Npu) are translated separately and must first associate to initiate the splicing reaction, so-called "protein trans splicing" (PTS) Figure S1b). Many of these split intein pairs have extremely high affinity for one another (Npu, K_D = 1.2 nM)[2] and, upon complex formation, exhibit remarkably fast splicing rates[3−6] (Npu splicing $t_{1/2}$ < 1 min),[4] making them attractive tools for a wide range of protein engineering applications.[1]

Due to the dramatic change in primary sequence that accompanies protein splicing, it has long been recognized that controlling intein activity offers a means to regulate associated protein (i.e., extein) function at the post-translational level.[7] Indeed, a number of so-called "conditional protein splicing" (CPS) systems have been reported in which the splicing event is triggered by external stimuli such as small molecules[7−11] or light.[12−16] Split inteins such as Npu are especially attractive starting points for the development of CPS methods since they support the extremely rapid ligation of two separately expressed polypeptides. While significant progress has been made in the development of CPS systems based on the naturally split inteins,[12,14,17] none of these offer modularity with respect to the type of triggers employed. Here, we describe a general strategy for controlling the association and subsequent trans splicing of naturally split inteins. Guided by the mechanism of split intein

association and folding,[2] we designed genetically encoded caged split intein pairs that are activated by targeted proteolysis, in effect creating split intein zymogens. We show that this design paradigm can be applied to several orthogonal split intein pairs and used to create CPS sensors and actuators that respond to the presence of bacterial pathogens, viral infections, and light.

Previously, we found that complementation of the Npu split intein pair is initiated by electrostatic interactions between an unstructured cationic region in the C-terminal fragment, NpuC (residues 1−13), and an unstructured anionic region in the N-terminal fragment, NpuN (residues 51−102).[2] This initial "capture" step leads to formation of a binding intermediate structure that subsequently undergoes a hydrophobic "collapse" to complete the native intein fold (Figure 1a). Based on this folding mechanism, we designed caged NpuN and NpuC constructs in which residues 51−102 of NpuN were fused to full-length NpuC (NpuC^{Cage}) and residues 1−13 of NpuC were fused to full-length NpuN (NpuN^{Cage}). We hypothesized that these cage sequences would participate in intramolecular interactions with the split intein fusion partner, effectively locking each half of the split intein in its binding intermediate structure and thus preventing fragment association and splicing (Figure 1b). As part of our design, we also included a protease cleavage sequence between each native split intein and its corresponding cage sequence. We imagined that proteolytic removal of the cage sequences would trigger split intein association and hence PTS (Figure 1b). Consequently, we refer to these as split intein zymogens, reflecting their potential to be enzymatically activated.

To assess the efficacy of our CPS strategy, we used recombinant expression methods to generate the two caged split inteins, each fused to model exteins (Figure S2). The purified constructs were then individually incubated with the corresponding Npu wild-type (Npu^{WT}) split intein fragment, also fused to a model extein (Figures S2 and S3). This revealed that the caging strategy was effective for NpuC, increasing the half-life of the Npu splicing reaction from less than 1 min to well over 1 h (Figure S3b). By contrast, attachment of the caging sequence to NpuN had no observable effect on the splicing rate (Figure S3d). Anticipating that our CPS system would ultimately need both intein fragments effectively inhibited, we carried out several rounds of optimization on the caged NpuN construct, ultimately arriving at a design which contained tandem copies of the cage unit, in addition to some
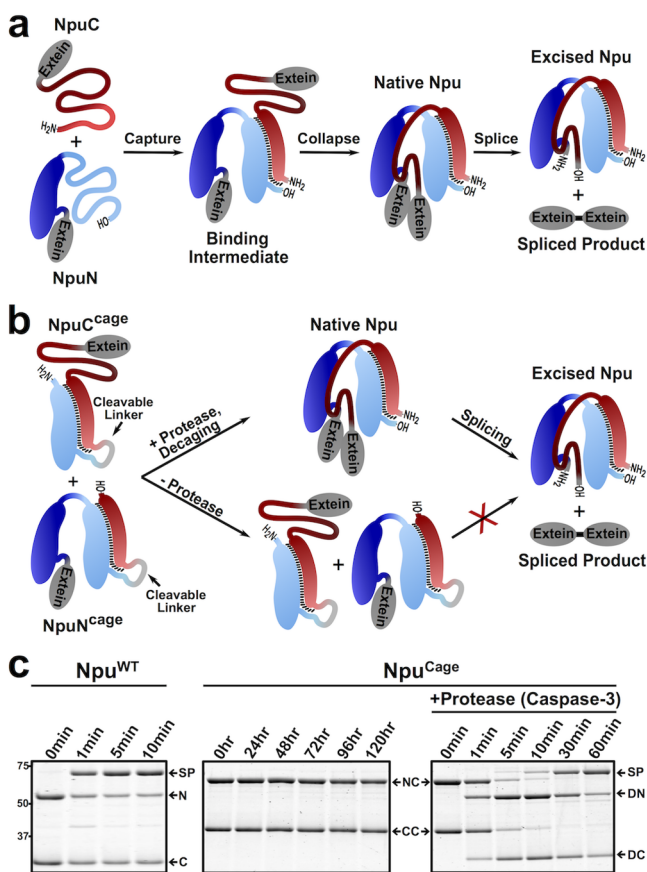
Figure 1. Design and characterization of caged intein zymogens: N = MBP-NpuN, C = NpuC-eGFP, NC = MBP-NpuN^Cage, CC = NpuC^Cage-eGFP, DN = decaged-NC, DC = decaged-CC, SP = splice product (MBP-eGFP). (a) Schematic depicting two-step mechanism by which the split fragments of Npu (NpuC and NpuN) assemble and splice. (b) Schematic depicting the design of NpuC^Cage and NpuN^Cage constructs, where addition of a protease cleaves or "decages" the inteins leading to protein splicing. (c) Coomassie-stained SDS-PAGE gels of splicing reactions (37 °C, 1 μM each intein fragment) monitored over time of either the wild-type (NpuWT) or zymogen (Npu^Cage) versions of the split Npu intein. Caspase-3 was added to a separate aliquot of Npu^Cage at time = 0 h.

point mutations at the binding interface (see Figures S2, S4, and S5 for details). Use of this optimized NpuN^Cage in PTS reactions with NpuC^WT led to a substantial reduction in the rate of product formation, analogous to what was observed with the NpuC^Cage construct (Figure S4). Importantly, our two CPS constructs, NpuN^Cage and NpuC^Cage, exhibited no affinity for one another in pull-down assays, and, critically, no splicing activity was observed between them, even after 5 days of incubation (Figures S6 and 1c). We note that this represents a remarkable level of inhibition over the native Npu split intein, which splices in seconds.[4] Consistent with our original design, addition of a protease (caspase-3) that is able to remove the cage sequence from these constructs leads to rapid onset of protein splicing ($t_{1/2} \approx 15$ min) (Figure 1c).

Encouraged by these initial results, we next asked whether the same caging strategy used on Npu could be extended to create zymogens of other split inteins. We focused on the GP41-1, GP41-8, and NrdJ-1 split inteins, since they all exhibit ultrafast splicing rates comparable to Npu.[6] Moreover, they are complementary to Npu in terms of the local extein environments they prefer to splice from (i.e., those residues

immediately flanking the intein), thus, making them attractive candidates for expanding our CPS toolbox. We aligned each intein sequence to Npu and used NpuC^Cage and an early iteration of NpuN^cage as templates to design corresponding IntC^Cage and IntN^Cage pairs (Figures S2 and S7). Similar to Npu, the caged versions of GP41-8, GP41-1, and NrdJ-1 were completely inactive until treated with a protease (caspase-3) that removes the cage sequences, whereupon efficient PTS ensued (Figures 2a and S8). The ease with which this caging strategy, originally developed for Npu, could be applied to other naturally split inteins reveals a remarkable level of robustness in our basic design and suggests a common folding mechanism between these separate families of split inteins.

We next sought to test two key features of our CPS platform, namely the exchangeable nature of the triggering feature and the capacity to generate a functional protein as output. To explore the first of these parameters, we replaced the caspase cut site in the original Npu zymogen with a cleavage sequence for the NS3/4A protease from the hepatitis C virus (Figure S2). As in the original example, we observed protease-dependent PTS (Figure 2b). The exchangeability of the trigger feature was also confirmed in the context of the GP41-8, GP41-1, and NrdJ-1 inteins—these were converted into intein zymogens that could be activated by human rhinovirus 3C protease, tobacco etch virus (TEV) protease, and thrombin, respectively (Figures 2b and S2). The efficiency of splicing varied somewhat depending on the intein-protease pair used, for example, Npu splicing triggered by the NS3/4A protease was less efficient than that triggered by Cas3. Notably, by installing unique cleavage sites in each of the four orthogonal inteins, we could selectively activate a desired PTS reaction in a split intein mixture (Figure S9). This illustrates the orthogonality of the inteins and highlights the potential of this system for multiplexed type applications.

Turning to the question of functional output, we generated an Npu intein zymogen pair designed to sense the bacterial pathogen *Staphylococcus aureus* and generate a pathogen-specific antibacterial agent as a response. These constructs contained cleavage sequences for the *S. aureus* secreted protease ScpA and were fused to a peptidoglycan degrading domain (Lys16) and a cell wall binding domain (SH3B), which possess potent anti-staphylococcal activity when linked[18] (Figures S2 and S10a). In initial studies, we confirmed that PTS activity was dependent on the presence of log phase *S. aureus* cultures (Figure S10b). Gratifyingly, the splicing reaction between the exogenously added caged constructs was directly coupled to a 99.9% reduction in bacterial colony forming units, confirming the ability of the CPS system to act as an actuator (Figure S10c).

Next, we examined the functionality of the CPS system in mammalian cells. We designed intein zymogens that could be utilized to generate a splicing-dependent enhanced green fluorescent protein (eGFP) (Figures S11 and S14). This allowed us to test whether our CPS system can be used as a sensor for viral infection (Figure 3a). Initial studies revealed that NpuN^Cage was poorly expressed in HEK-293T cells, a problem that was ultimately traced to two residues that had been mutated during the initial optimization process (E7K and E61K)—fortunately reverting these to their native glutamates largely rescued expression to WT levels (Figures S12 and S14). The TEV cleavage sequence was embedded within NpuN^Cage and NpuC^Cage and the resulting split intein zymogen pair fused to the respective N- and C-terminal fragments of an eGFP
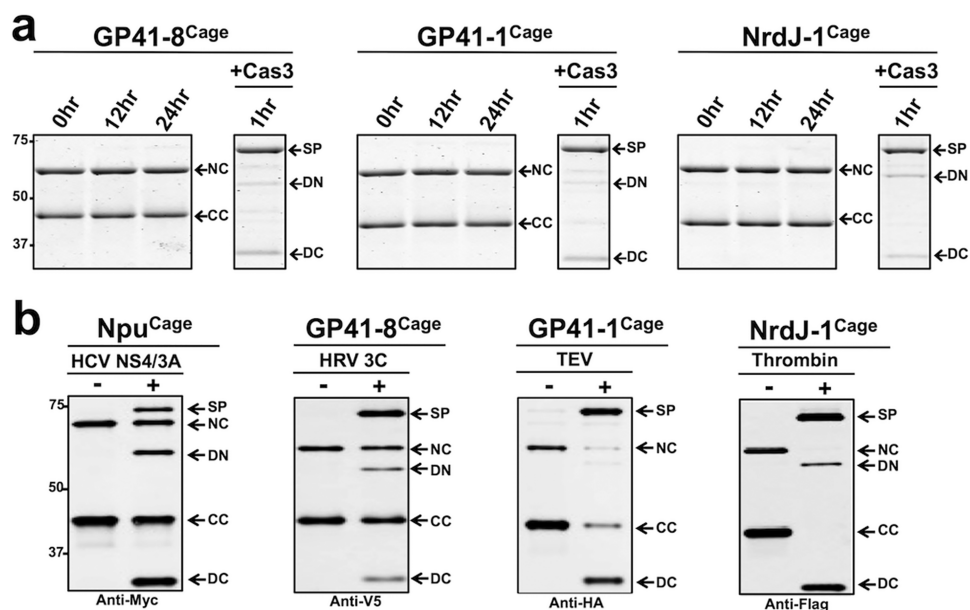
Figure 2. Splicing reactions of orthogonal intein zymogens: NC = MBP-IntN^{Cage}, CC = IntC^{Cage}-eGFP, DN = decaged-NC, DC = decaged-CC, SP = splice product (MBP-eGFP). (a) Coomassie-stained SDS-PAGE gels showing in vitro splicing reactions of caged versions of GP41-8, GP41-1, and NrdJ-1 split inteins in the absence and presence of protease trigger, caspase-3 (Cas3). Each intein zymogen contains a Cas3 cut site and for each reaction, Cas3 was added in a separate aliquot at time = 0 h. (b) Western blots showing in vitro splicing reactions of caged GP41-8, GP41-1, NrdJ-1, and Npu split inteins with indicated protease cleavage sequences in the absence and presence of protease triggers, carried out for 1 h at 25 °C and 1 μM of each indicated intein zymogen pair.
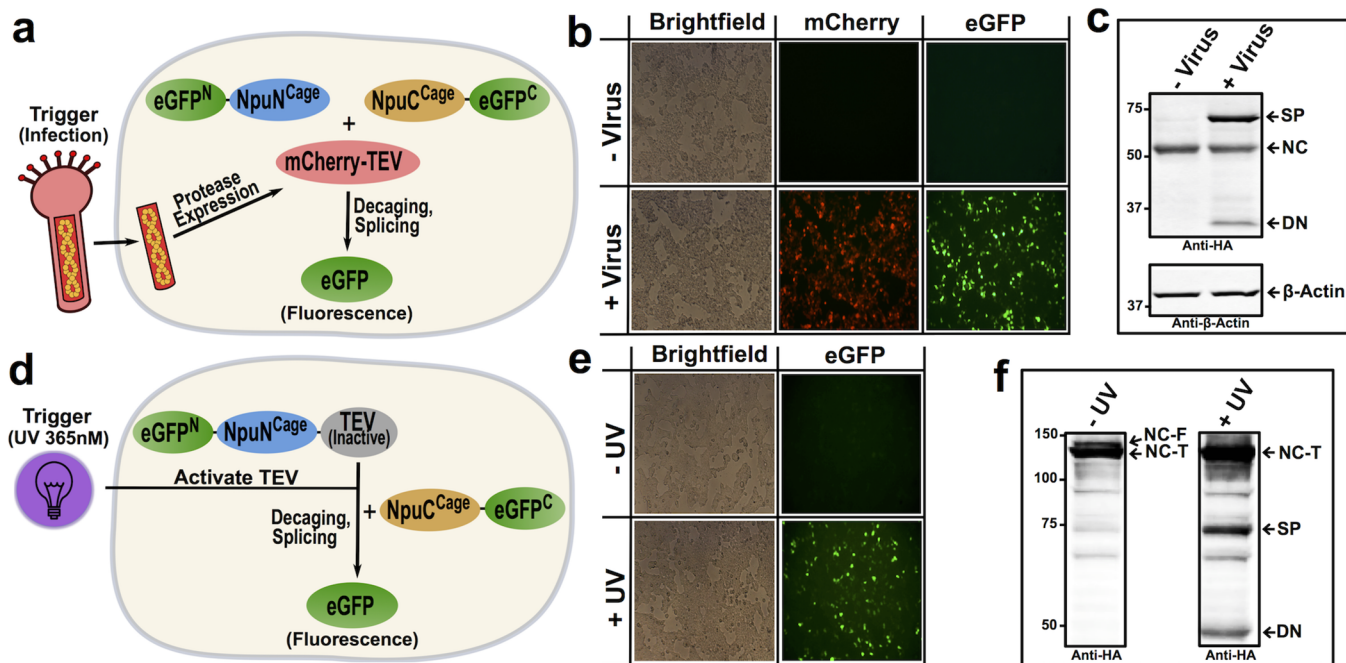

Figure 3. Fluorescent response to viral infection or photo-activation. (a,d) Schematics depicting TEV-activated split eGFP intein zymogen system in mammalian cells with either viral infection (a−c) or photolysis (d−f) acting as CPS triggers. Note that, for clarity, additional domains fused to eGFP fragments are not shown. For more details on constructs used, see Supporting Information. (b) Images of HEK 293T cells expressing the TEV-activated split eGFP intein zymogen constructs either with (bottom) or without (top) baculoviral infection. Images were taken following an overnight incubation with baculovirus. (c) Top: Western blot analysis (anti-HA) of cells from images in (b). SP = splice product, NC = HA-eGFP(1−64)-NpuN^{Cage}, DN = decaged-NC. Bottom: anti-β-actin loading control. (e) Images of HEK 293T cells expressing split eGFP intein zymogen system and photo-caged TEV with (bottom) or without (top) 1 min UV (365 nM) irradiation. Images were taken 6 h post irradiation. (f) Western (anti-HA) analysis of cells from images in (e). NC-F = HA-eGFP(1−64)-NpuN^{Cage}-Tev-(C151TAG) full translation product, NC-T = truncated translation product of NC-F, SP = splice product, DN = decaged-NC.

reporter (Figures 3a and S14). As expected, no protein splicing was observed when these CPS constructs were co-expressed in

HEK-293T cells. By contrast, exposure of these cells to a baculovirus strain encoding mCherry-TEV resulted in robust

generation of eGFP splice product (Figure 3b,c). This result demonstrates that our system can sense the presence of virally encoded proteases and, by extension, viral infection.

Inspired by previous TEV activation technologies,[19] we next sought to add an additional layer of control to our system by utilizing amber codon suppression to incorporate a photo-caged cysteine into the protease active site (Figures 3d and S13). In principle, this would enable photo-control of fragment association in the context of the TEV-activated intein zymogen system. Such a system would complement an existing optogenetic tool based on incorporation, by amber suppression, of a caged cysteine residue directly into an artificially fused version of Npu.[16] With this in mind, we designed a construct in which a mutant TEV protease was fused to the C-terminus of NpuN^{Cage} (NpuN^{Cage}-TEV, Figure S14). HEK-293T cells harboring the requisite orthogonal tRNA/tRNA synthetase pair for the unnatural amino acid S-[1-(4′,5′-(methylenedioxy)-2′-nitrophenyl)ethyl]cysteine[19] were transfected with NpuC^{Cage} and NpuN^{Cage}-TEV constructs fused to the split eGFP pair as N- and C-exteins (Figures 3d and S14). Consistent with our design, we observed irradiation-dependent generation of spliced eGFP in the presence of the unnatural amino acid (Figure 3e,f). The success of this light-activated split intein zymogen system is expected to expand the range of protein targets amenable to optogenetic control, due to the known promiscuity of the Npu split intein.[20] Moreover, we note that placing the photo-caged residue in the protease, rather than the intein itself, could help mitigate one of the drawbacks sometimes associated with amber suppression, namely lower expression yields caused by premature chain termination. The amplification effect inherent to the enzymatic trigger, in principle, means that a small amount of the full-length protease should be sufficient to support efficient photo-uncaging.

Despite the extraordinarily high binding affinity and rapid splicing kinetics of ultrafast split inteins, the CPS strategy reported herein effectively cages the association of the split intein fragments while allowing for a diverse set of proteases to trigger the splicing reaction. Applications of these intein zymogens, such as the conditional killing of *S. aureus* and sensing of viral infection in mammalian cells, illustrate this system's ability to detect and respond to biological processes involving proteases. Moreover, coupling the caged constructs to a photo-activatable TEV protease adds an additional layer of spatiotemporal control. We envision that the development of orthogonally caged systems, each able to respond to a unique trigger, will expand the range of targets amenable to CPS and enable one-pot reactions where multiple separate splicing reactions can be controlled.

■ ASSOCIATED CONTENT

S Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b02618.

> Full methods and experimental data, including Figures S1−S14 (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*muir@princeton.edu

ORCID

Tom W. Muir: 0000-0001-9635-0344

Notes

■ REFERENCES

(1) Shah, N. H.; Muir, T. W. *Chem. Sci.* **2014**, *5*, 446.
(2) Shah, N. H.; Muir, T. W.; et al. *J. Am. Chem. Soc.* **2013**, *135*, 18673.
(3) Iwai, H.; Züger, S.; Jin, J.; Tam, P. H. *FEBS Lett.* **2006**, *580*, 1853.
(4) Zettler, J.; Schütz, V.; Mootz, H. D. *FEBS Lett.* **2009**, *583*, 909.
(5) Shah, N. H.; Dann, G. P.; Vila-Perello, M.; Liu, Z.; Muir, T. W. *J. Am. Chem. Soc.* **2012**, *134*, 11338.
(6) Carvajal-Vallejos, P.; Pallisse, R.; Mootz, H. D.; Schmidt, S. R. *J. Biol. Chem.* **2012**, *287*, 28686.
(7) Mootz, H. D.; Muir, T. W. *J. Am. Chem. Soc.* **2002**, *124*, 9044.
(8) Buskirk, A. R.; Ong, Y. C.; Gartner, Z. J.; Liu, D. R. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 10505.
(9) Yuen, C. M.; Rodda, S. J.; Vokes, S. A.; McMahon, A. P.; Liu, D. R. *J. Am. Chem. Soc.* **2006**, *128*, 8939.
(10) Schwartz, E. C.; Saez, L.; Young, M. W.; Muir, T. W. *Nat. Chem. Biol.* **2007**, *3*, 50.
(11) Peck, S. H.; Chen, I.; Liu, D. R. *Chem. Biol.* **2011**, *18*, 619.
(12) Vila-Perelló, M.; Hori, Y.; Ribó, M.; Muir, T. W. *Angew. Chem., Int. Ed.* **2008**, *47*, 7764.
(13) Tyszkiewicz, A. B.; Muir, T. W. *Nat. Methods* **2008**, *5*, 303.
(14) Böcker, J. K.; Friedel, K.; Matern, J. C.; Bachmann, A. L.; Mootz, H. D. *Angew. Chem., Int. Ed.* **2015**, *54*, 2116.
(15) Jung, D.; Sato, K.; Min, K.; Shigenaga, A.; Jung, J.; Otaka, A.; Kwon, Y. *Chem. Commun.* **2015**, *51*, 9670.
(16) Ren, W.; Ji, A.; Ai, H. W. *J. Am. Chem. Soc.* **2015**, *137*, 2155.
(17) Wong, S.; Mosabbir, A. A.; Truong, K. *PLoS One* **2015**, *10*, e0135965.
(18) Paul, V. D. *BMC Microbiol.* **2011**, *11*, 226.
(19) Nguyen, D. P.; Mahesh, M.; Elsässer, S. J.; Hancock, S. M.; Uttamapinant, C.; Chin, J. W. *J. Am. Chem. Soc.* **2014**, *136*, 2240.
(20) Cheriyan, M.; Pedamallu, C. S.; Tori, K.; Perler, F. *J. Biol. Chem.* **2013**, *288*, 6202.