Published in final edited form as:
Angew Chem Int Ed Engl. 2011 July 11; 50(29): 6511–6515. doi:10.1002/anie.201102909.

Kinetic Control of One-Pot Trans-Splicing Reactions by Using a Wild-Type and Designed Split Intein**

Neel H. Shah, **Dr. Miquel Vila-Perelló**, and **Prof. Tom W. Muir**
Laboratory of Synthetic Protein Chemistry, The Rockefeller University, 1230 York Avenue, New York, NY 10065 (USA), Fax: (+1)212-327-7358

Miquel Vila-Perelló: mvila@rockefeller.edu; Tom W. Muir: muirt@rockefeller.edu

Keywords

ion pairs; protein design; protein–protein interactions; semisynthesis

Coulombic forces play an important role in facilitating protein–protein interactions. It has been demonstrated that a strong electrostatic potential between two interacting proteins correlates with a fast rate of association and a strong binding affinity.[1] Indeed, this principle has been exploited to enhance the binding properties of engineered protein interfaces.[2] Nonetheless, little research has focused on utilizing intermolecular ion pairs to modulate specificity in protein–protein interactions.[3–5] Naturally split inteins are a potentially interesting system for engineering electrostatically driven specificity, as the formation of a catalytically competent structure requires the association of two oppositely charged protomers. In their endogenous environment, split inteins catalyze protein trans-splicing (PTS, Scheme 1) of essential gene fragments, and thus are under evolutionary pressure to associate rapidly with high fidelity and dissociate slowly post-catalysis to prevent the re-formation of unproductive complexes. Out of their native context, split inteins have seen widespread use in a number of biotechnological applications, as a result of their capacity to ligate flanking protein sequences (exteins) in trans.[6] Applications of PTS include segmental isotopic labeling of proteins for NMR spectroscopy,[7] protein immobilization,[8] the labeling of proteins with extrinsic probes,[9] protein and peptide cyclization,[10] and control of protein function.[11,12]

Despite the utility of PTS, little is known about what drives efficient association of split intein fragments. Sequence alignments of naturally split inteins show highly conserved charge segregation, with acidic residues concentrated at specific positions on the N-intein and basic residues conserved on the C-intein (N- and C-terminal protomers, Figure 1a).[13] Furthermore, a bioinformatic sequence analysis of the intein family indicates that this charge segregation is significantly more prevalent in naturally split inteins than intact ones (Figure 1c). Interestingly, when the conserved charged residues are mapped onto the structure determined by NMR spectrocopy of the wild-type DnaE intein from *Nostoc punctiforme* (Npu_{WT}), many are found to be participating in intermolecular ion pairs and ion triads (Figure 1b).[14]

**We thank Dr. P. Moyle for valuable discussions and Dr. J. Fernandez from The Rockefeller University Proteomics Resource Center for the MS/MS analysis of PARP1. This work was supported by the US National Institutes of Health (NIH grant GM086868).

Correspondence to: Miquel Vila-Perelló, mvila@rockefeller.edu; Tom W. Muir, muirt@rockefeller.edu.

While ionic interactions have previously been postulated to play a role in split intein assembly,[13,15] the involvement of electrostatic forces has not been validated experimentally. In this study, we set out to test the hypothesis that ionic interactions facilitate the association of split intein fragments and thus could be manipulated to control the relative reactivities of different N- and C-intein complexes.

To evaluate the role of specific ionic interactions for split intein splicing and specificity we screened a series of charge-swapped Npu intein mutants and tested their splicing activity in vivo. Starting from the wild-type protomers, $NpuN_{WT}$ and $NpuC_{WT}$, mutation positions were chosen based on two criteria: 1) residues should be involved in an intermolecular ion pair or triad, and 2) these residues should be moderately isolated from intramolecular ionic interactions. Based on these criteria, we identified four ion clusters comprised of one pair and three triads. For each cluster and several combinations of clusters, three variants were generated: an N-intein mutant in which native acidic residues were mutated to basic residues, a C-intein mutant in which basic residues were mutated to acidic residues, and a charge-swapped intein that combined all N- and C-intein mutations (Figure 2a). The activity of each set was analyzed using an in vivo splicing assay where the kanamycin resistance in *E. coli* was dependent on the intein splicing activity.[16] Of all the combinations tested (see the Supporting Information), the intein fragments in which all the possible ion clusters were charged swapped, $NpuN_{MUT}$ and $NpuC_{MUT}$, showed the least cross-reactivity with wild-type fragments in vivo (Figure 2b). Specifically, $NpuN_{WT}$ did not react with $NpuC_{MUT}$, $NpuN_{MUT}$ reacted minimally with $NpuC_{WT}$, and both mutant fragments displayed substantial splicing activity when combined in vivo. Importantly, the new charge-swapped intein, Npu_{MUT}, catalyzed protein splicing more efficiently than the widely used SspDnaE intein.

To determine if the observed trend for in vivo splicing activities was a direct result of relative fragment binding affinities we developed an in vitro binding assay. Using ubiquitin (Ub) and the small ubiquitin-like modifier (SUMO) as model N- and C-extein domains, we expressed and purified Ub-NpuN and NpuC-SUMO fusion proteins bearing wild-type or mutant intein sequences. The Ub-NpuN fusions contained a cysteine to alanine mutation (C1A) to prevent trans-splicing. To measure the binding affinities, we took advantage of the single tryptophan residue (W47) in the Npu intein and measured a decrease in the intrinsic fluorescence of NpuN in the presence of increasing concentrations of NpuC-SUMO (Table 1 and see the Supporting Information). As expected, the wild-type fragments had a low-nanomolar binding affinity, consistent with previous measurements for the highly homologous SspDnaE intein.[15] The fully charged-swapped Npu_{MUT} N- and C-inteins associated with a 40-fold weaker affinity than the Npu_{WT} protomers (K_d = 118.4 nM), which may explain their slightly diminished splicing activity in vivo. Surprisingly, both wild-type/mutant hybrids had measurable binding affinities only two- to threefold weaker than $NpuN_{MUT}$ + $NpuC_{MUT}$, despite their extremely low or undetectable levels of splicing in vivo. These data indicated that while charge-swapping modulated the intein fragment affinities with the expected trend, the magnitude of these energetic effects do not fully explain the splicing selectivity observed in vivo. Notably, we observed that the $NpuN_{WT}$ + $NpuC_{MUT}$ combination showed only a 29% decrease in N-intein fluorescence upon binding, while all other fragment combinations showed a 60–70% decrease in fluorescence (see the Supporting Information). This anomalous fluorescence strongly suggests that the $NpuN_{WT}$ + $NpuC_{MUT}$ complex adopts a unique conformation which could explain its lack of splicing activity in vivo.

We next examined the relative rates of trans-splicing for Npu_{WT}, Npu_{MUT}, and their N- and C-intein combinations in vitro. Splicing kinetics at 30°C were measured by pairwise mixing of Ub-NpuN and NpuC-SUMO fusions at equimolar ratios and monitoring the formation of

the Ub-SUMO spliced product by gel electrophoresis (Table 1). Npu_{WT} showed extremely rapid splicing ($t_{1/2} \approx 1$ min) as previously observed.[17] As expected, given the low-nanomolar affinity for the wild-type fragments, this rapid rate of splicing was relatively independent of intein concentrations from 1.0 μM to 0.05 μM. Consistent with the in vivo splicing and in vitro binding experiments, Npu_{MUT} spliced slightly slower than Npu_{WT} at 1.0 μM ($t_{1/2}$ = 6.5 min), but its rate decreased roughly tenfold at intein concentrations near its K_d value. Surprisingly, the $NpuN_{MUT}$ + $NpuC_{WT}$ combination catalyzed splicing almost as quickly as $NpuN_{MUT}$ + $NpuC_{MUT}$ at high concentrations, but this difference in trans-splicing rates increased dramatically at lower concentrations, because of the weaker binding affinity for $NpuN_{MUT}$ + $NpuC_{WT}$. The $NpuN_{WT}$ + $NpuC_{MUT}$ combination showed no detectable trans-splicing at all three concentrations tested. Not only is this observation consistent with the in vivo results, it also supports the fluorescence data that suggest that this N- and C-intein bind to form a catalytically incompetent complex. The in vitro binding and splicing assays collectively demonstrate that charge complementation or repulsion can bias relative split intein binding affinities, which in turn bias relative splicing kinetics. These experiments also shed light on our in vivo splicing assays. Specifically, the degree of concentration dependence on splicing rates observed in vitro indicates that the intein concentration in vivo is probably at the low nanomolar range, at which level the in vivo and in vitro results would be consistent.

Given the differences in splicing kinetics observed in vitro, we envisioned that Npu_{WT} and Npu_{MUT} could be used simultaneously to catalyze multiple trans-splicing reactions with kinetically controlled selectivity.[18,19] To test this, we developed an in vitro competition assay (Figure 3a) using our splicing assay fusion proteins and two additional fusions to the Npu_{WT} fragments bearing unique exteins—maltose binding protein (MBP) and enhanced green fluorescent protein (eGFP). As a control reaction, we mixed two $NpuN_{WT}$ fusions and two $NpuC_{WT}$ fusions bearing unique exteins at equimolar concentrations (0.5 μM) at 30°C. The formation of all four possible products was monitored by Western blotting (see the Supporting Information). As expected, all the products formed to a similar extent in the control reaction (Figure 3b).

When Ub-$NpuN_{MUT}$ and $NpuC_{MUT}$-SUMO were used in place of their wild-type counterparts, a clear bias in product formation was observed. The Npu_{WT} product (MBP-eGFP) formed rapidly, and the Npu_{MUT} product (Ub-SUMO) emerged at a slightly slower rate, consistent with our in vitro splicing assays. These products formed almost exclusively, close to 50% each, while the $NpuN_{MUT}$ + $NpuC_{WT}$ product (Ub-eGFP) accounted for less than 5% of the total product formed, and the $NpuN_{WT}$ + $NpuC_{MUT}$ product (MBP-SUMO) was not observed (Figure 3c). These results were reproducible over a range of temperatures —from 25°C to 37°C—and at higher ionic strength, thus indicating that the electrostatically driven selectivity between Npu_{WT} and Npu_{MUT} is extremely robust (see the Supporting Information).

Next, we explored the utility of these inteins for one-pot three-piece ligations of proteins (Figure 4a). Previously, orthogonal intein systems have been developed by combining naturally and artificially split inteins or by using a wild-type and linearly permuted split intein.[15,20] While these systems could catalyze three-piece ligations, the prior system showed low efficiency, and the latter could not be carried out in one pot. We envisioned that our Npu_{WT} and Npu_{MUT} pair could efficiently catalyze one-pot three-piece ligations, given the high degree of kinetic control seen in our competition assay. To test this, we designed a model system to ligate a Src-homology 3 (SH3) domain, domain B1 of protein G (GB1), and eGFP. The domains were fused to split intein sequences to generate an N-terminal fragment (N: SH3-$NpuN_{MUT}$), a middle fragment (M: $NpuC_{MUT}$-GB1-$NpuN_{WT}$), and a C-terminal fragment (C: $NpuC_{WT}$-eGFP). Importantly, we designed the middle domain flanked by the

N- and C-inteins that do not react in trans since this would preclude GB1 cyclization or oligomerization (Figure 4a). To test our three-piece ligation system, the N, M, and C fragments were mixed with a slight excess of the middle fragment. Analysis of the reaction mixture by gel electrophoresis indicated that the reaction between $NpuN_{WT}$ and $NpuC_{WT}$ (M + C) occurred extremely fast to yield one intermediate product. This intermediate was more slowly converted into the full-length three-piece-ligated product, SH3-GB1-eGFP, upon reaction with the N fragment, consistent with the kinetically controlled trans-splicing paradigm (see the Supporting Information).

Next, we sought to apply this three-piece ligation technology towards the semisynthesis of human poly(ADP-ribose) polymerase 1 (PARP1). This approximately 115 kDa enzyme, which catalyzes the transfer of ADP-ribose from nicotinamide adenine dinucleotide (NAD) onto protein side chains in the form of monomers or polymers, is involved in DNA-damage pathways and is a promising target for chemo-therapeutics.[21,22] Given its large size and its capacity to automodify itself,[23] full-length PARP1 is not readily isolable by over-expression in *E. coli*.[24,25] We envisioned that our orthogonal inteins could be used to generate full-length PARP1 by expression of its fragments separately in *E. coli* followed by in vitro three-piece ligation. Based on known domain boundaries,[26,27] we chose two ligation sites in putatively flexible regions that separated PARP1 into three fragments: the N-terminal zinc finger DNA-binding domains (PARP1-N), the central dimerization and automodification domains (PARP1M), and the C-terminal catalytic domain (PARP1-C) (Figure 4b). Importantly, each of these functional domains is required for activation and regulation of PARP1 catalysis,[28] and thus their efficient and accurate assembly is a rigorous test of our three-piece ligation system.

As a consequence of the lack of cysteine residues near the desired splice junctions, we introduced two modest mutations, S364C and T656C, to allow for split intein catalysis and fused the PARP1 fragments to our split inteins. In addition, we designed a traceless tagging strategy in which His_6 tags were placed at the free termini of every intein sequence (see the Supporting Information). These tags could be used not only to facilitate enrichment of the proteins after over-expression but also to trap any remaining starting materials, intermediates, and spliced intein fragments after a three-piece ligation. To generate full-length PARP1 we expressed the three intein-fusion fragments separately in *E. coli* and enriched the proteins over nickel columns. The semipure proteins were mixed and allowed to react at room temperature for 21 h. By following the reaction by Western blotting against PARP1-N and PARP1-C, we observed the formation of both reaction intermediates as well as full-length PARP1. The reaction mixture was passed through a nickel affinity column to trap residual starting materials, intermediates, and free intein fragments. The flow-through, containing the full-length product, was then purified by size-exclusion chromatography to yield pure, three-piece-ligated PARP1 (Figure 4c). Mass spectrometric analysis of tryptic fragments confirmed the identity of the purified product, and peptides spanning the ligation junctions were further analyzed by MS/MS to confirm their sequences (see the Supporting Information).

To determine whether our semisynthetic PARP1 was active, we conducted ADP-ribosylation assays using biotinylated-NAD and streptavidin blotting. We observed that our three-piece ligated PARP1 could catalyze automodification (Figure 4d). Importantly, this activity required activated DNA, could be inhibited by benzamide, and was stimulated by histone octamers, all of which are hallmarks of full-length PARP1 activity.[29,30] Furthermore, our enzyme could catalyze the formation of poly(ADP-ribose) chains on histones, as previously reported for endogenous PARP1 (see the Supporting Information).[31] These enzymatic data unequivocally demonstrate that our three-piece ligated PARP1 is full-length and properly folded, as the DNA-dependent activity of this

enzyme requires allosteric communication between all three segments. Consistent with this requirement, the PARP1-C fragment, bearing only the catalytic domain of PARP1, did not catalyze poly(ADP-ribosylation) of histones (see the Supporting Information).

In this study, we probed the role of intermolecular ion clusters for fragment assembly and splicing in the Npu_{WT} split intein both in vivo and in vitro. Through these experiments, we rationally designed a new split intein, Npu_{MUT}, which displays low cross-reactivity with Npu_{WT}. These orthogonal inteins were used to generate the large, full-length, active mammalian protein PARP1 through a one-pot three-piece ligation. Collectively, our results demonstrate that electrostatic interactions can engender kinetic control in a complex enzymatic system. Furthermore, these results provide insight into the molecular requirements for efficient and specific protein trans-splicing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Schreiber G, Fersht AR. Nat Struct Biol. 1996; 3:427. [PubMed: 8612072]
2. Selzer T, Albeck S, Schreiber G. Nat Struct Biol. 2000; 7:537. [PubMed: 10876236]
3. Graddis TJ, Myszka DG, Chaiken IM. Biochemistry. 1993; 32:12664. [PubMed: 8251485]
4. O'Shea EK, Lumb KJ, Kim PS. Curr Biol. 1993; 3:658. [PubMed: 15335856]
5. Gunasekaran K, Pentony M, Shen M, Garrett L, Forte C, Woodward A, Ng SB, Born T, Retter M, Manchulenko K, Sweet H, Foltz IN, Wittekind M, Yan W. J Biol Chem. 2010; 285:19637. [PubMed: 20400508]
6. Vila-Perelló M, Muir TW. Cell. 2010; 143:191. [PubMed: 20946979]
7. Yamazaki T, Otomo T, Oda N, Kyogoku Y, Uegaki K, Ito N, Ishino Y, Nakamura H. J Am Chem Soc. 1998; 120:5591.
8. Kwon Y, Coleman MA, Camarero JA. Angew Chem. 2006; 118:1758.Angew Chem Int Ed. 2006; 45:1726.
9. Giriat I, Muir TW. J Am Chem Soc. 2003; 125:7180. [PubMed: 12797783]
10. Scott CP, Abel-Santos E, Wall M, Wahnon DC, Benkovic SJ. Proc Natl Acad Sci USA. 1999; 96:13638. [PubMed: 10570125]
11. Mootz HD, Blum ES, Tyszkiewicz AB, Muir TW. J Am Chem Soc. 2003; 125:10561. [PubMed: 12940738]
12. Vila-Perelló M, Hori Y, Ribó M, Muir TW. Angew Chem. 2008; 120:7878.Angew Chem Int Ed. 2008; 47:7764.
13. Dassa B, Amitai G, Caspi J, Schueler-Furman O, Pietrokovski S. Biochemistry. 2007; 46:322. [PubMed: 17198403]
14. Oeemig JS, Aranko AS, Djupsjöbacka J, Heinämäki K, Iwaï H. FEBS Lett. 2009; 583:1451. [PubMed: 19344715]
15. Shi J, Muir TW. J Am Chem Soc. 2005; 127:6198. [PubMed: 15853324]
16. Lockless SW, Muir TW. Proc Natl Acad Sci USA. 2009; 106:10999. [PubMed: 19541616]
17. Zettler J, Schütz V, Mootz HD. FEBS Lett. 2009; 583:909. [PubMed: 19302791]
18. Zhang Z, Ollmann IR, Ye X-S, Wischnat R, Baasov T, Wong C-H. J Am Chem Soc. 1999; 121:734.
19. Bang D, Pentelute BL, Kent SBH. Angew Chem. 2006; 118:4089.Angew Chem Int Ed. 2006; 45:3985.
20. Busche AEL, Aranko AS, Talebzadeh-Farooji M, Bernhard F, Dötsch V, Iwaï H. Angew Chem. 2009; 121:6244.Angew Chem Int Ed. 2009; 48:6128.
21. Hassa PO, Haenni SS, Elser M, Hottiger MO. Microbiol Mol Biol Rev. 2006; 70:789. [PubMed: 16959969]

22. Martin SA, Lord CJ, Ashworth A. Curr Opin Genet Dev. 2008; 18:80. [PubMed: 18343102]
23. Tao Z, Gao P, Liu H-w. J Am Chem Soc. 2009; 131:14258. [PubMed: 19764761]
24. Knight MI, Chambers PJ. Protein Expression Purif. 2001; 23:453.
25. Gagnon SN, Desnoyers S. Mol Cell Biochem. 2003; 243:15. [PubMed: 12619884]
26. Tao Z, Gao P, Hoffman DW, Liu H-W. Biochemistry. 2008; 47:5804. [PubMed: 18452307]
27. Lilyestrom W, van der Woerd MJ, Clark N, Luger K. J Mol Biol. 2010; 395:983. [PubMed: 19962992]
28. Simonin F, Höfferer L, Panzeter PL, Muller S, de Murcia G, Althaus FR. J Biol Chem. 1993; 268:13454. [PubMed: 8390463]
29. Pinnola A, Naumova N, Shah M, Tulin AV. J Biol Chem. 2007; 282:32511. [PubMed: 17827147]
30. Southan G, Szabo C. Curr Med Chem. 2003; 10:321. [PubMed: 12570705]
31. Messner S, Altmeyer M, Zhao H, Pozivil A, Roschitzki B, Gehrig P, Rutishauser D, Huang D, Caflisch A, Hottiger MO. Nucleic Acids Res. 2010; 38:6350. [PubMed: 20525793]

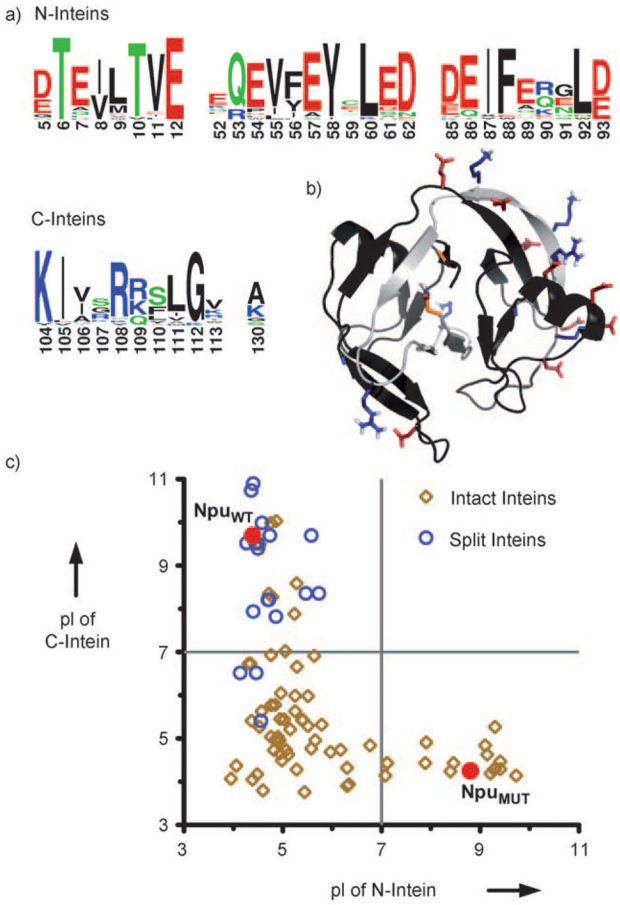


Figure 1.

Charge segregation in split inteins. a) Sequence logo based on a sequence alignment of 24 naturally split inteins highlighting the highly conserved charged residues. b) Structure of Npu_{WT} determined by NMR spectroscopy (PDB code: 2KEQ) highlighting intermolecular electrostatic interactions. $NpuN_{WT}$ and $NpuC_{WT}$ are shown in black and gray, respectively. Acidic residues are shown in red, basic residues are shown in blue, and terminal catalytic residues are shown in orange. c) Comparison of the calculated isoelectric points (pI) for the N- and C-terminal fragments of 24 naturally split inteins, the N- and C-terminal sequences of 76 intact inteins, and the mutant split intein (Npu_{MUT}) engineered in this study. Npu_{WT} and Npu_{MUT} are both indicated on the plot as red circles.

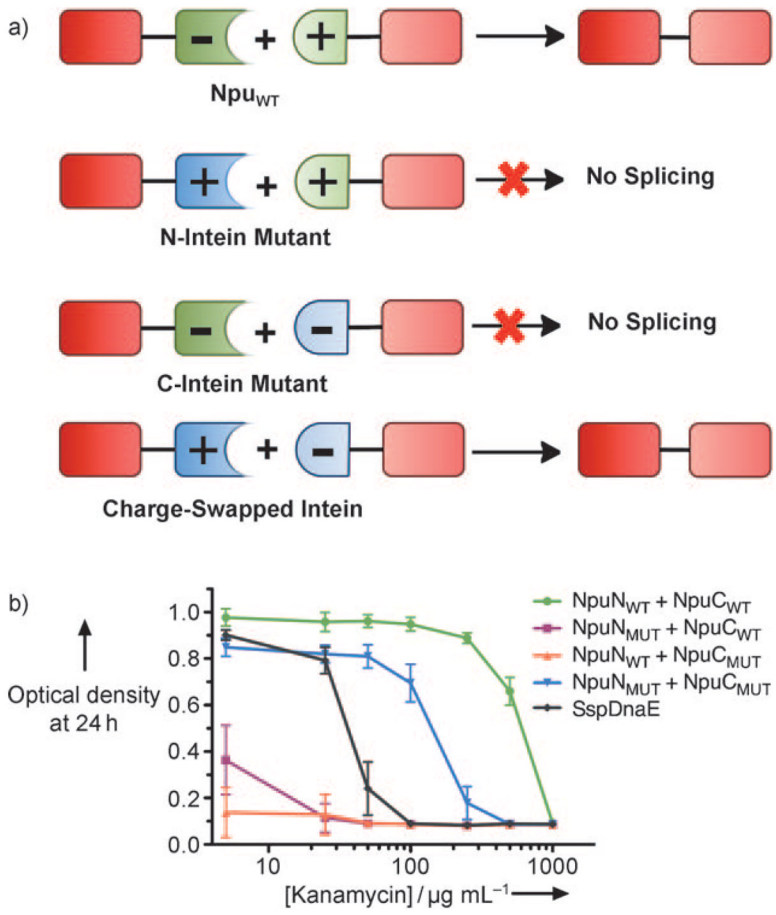


Figure 2.

Rational design strategy for charge-swapped inteins. a) Scheme depicting the desired specificity in charge-swapped inteins. b) Dose–response curves for Npu_{WT} and Npu_{MUT} fragment combinations in an in vivo intein activity-coupled kanamycin resistance assay. Mutation clusters in Npu_{MUT} are (E7K, K130E), (E52K, E54K, K113E), (D85K, E89R, R108E), and (E61K, E91K, K104E).

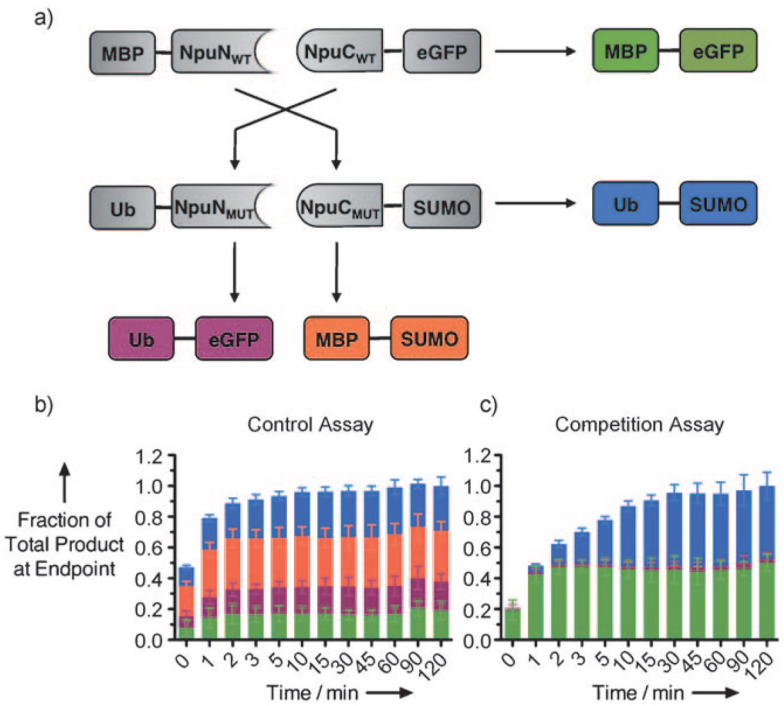


Figure 3.

Competition assays with Npu$_{WT}$ and Npu$_{MUT}$. a) Scheme showing four possible splicing reactions in the competition assay. b) Control reaction with wild-type intein fragments. c) Competition reaction with wild-type and mutant fragments.

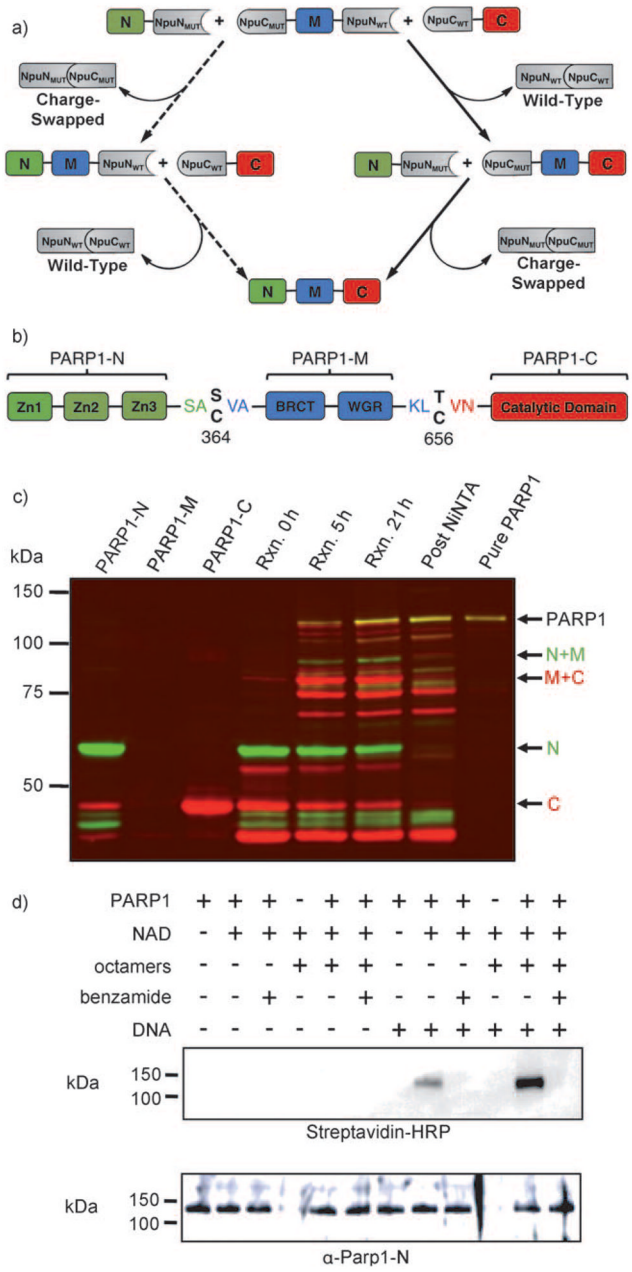
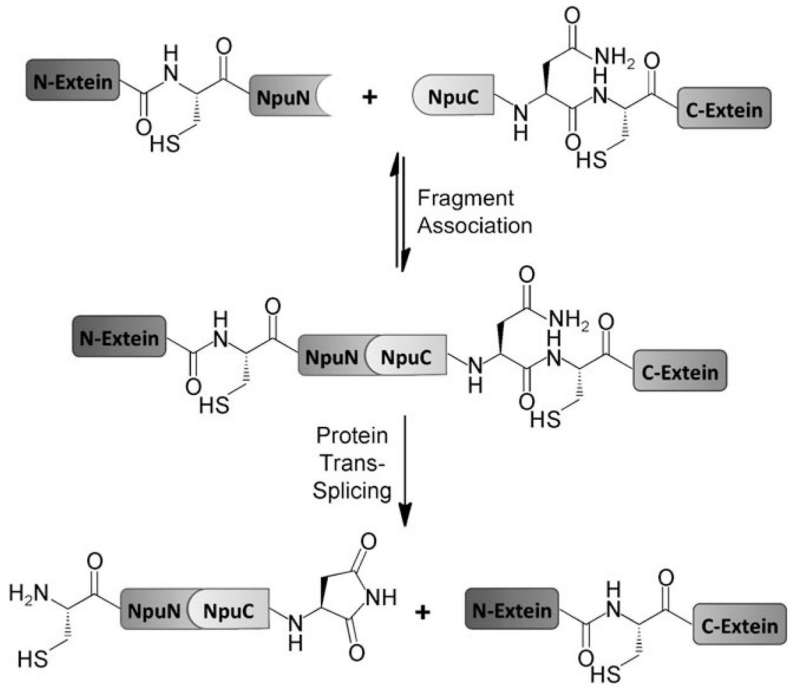


Figure 4.

Three-piece ligation of human PARP1. a) Scheme depicting a three-piece ligation using Npu_{WT} and Npu_{MUT}. The major reaction pathway is indicated by solid arrows, and the minor reaction pathway is indicated by dashed arrows. b) Domain organization and splice junctions in PARP1. c) Western blot of PARP1 three-piece ligation and purification. Color scheme: green bands (α-PARP1-N), red bands (α-PARP1-C), and yellow bands (α-PARP1-N + α-PARP1-C). Un-annotated bands are identified in Figure S20 in the Supporting Information. d) Biotinylated-NAD blot showing automodification of PARP1 (streptavidin-HRP) and PARP1 blot loading control (α-PARP1-N).



Scheme 1.
Protein trans-splicing.

Table 1

In vitro characterization of Npu_{WT} and Npu_{MUT}.

Fragment pair		K_d [nM][a]	Initial $t_{1/2}$ of trans-splicing [min][a,b]		
NpuN	**NpuC**		**1.0 μM**	**0.1 μM**	**0.05 μM**
WT	WT	2.9 ± 1.8	0.9 ± 0.1	1.1 ± 0.2	2.4 ± 0.4
MUT	WT	211.6 ± 21.0	8.2 ± 0.4	71.0 ± 4.9	350.3± 41.2
WT	MUT	281.8 ± 51.5	n/a[c]	n/a[c]	n/a[c]
MUT	MUT	118.4 ± 11.4	6.5 ± 0.4	43.7 ± 2.1	103.0± 8.7

[a] Error values indicate the standard error from a global fit of three replicate data sets.

[b] Initial half-lives are calculated from a second-order rate constant.

[c] No detectable trans-splicing was observed for the $NpuN_{WT}$ + $NpuC_{MUT}$ combination.

n/a = not active.

Angew Chem Int Ed Engl. Author manuscript; available in PMC 2012 July 11.