CHEM**BIO**CHEM
Communications

Off-Pathway-Sensitive Protein-Splicing Screening Based on a Toxin/Antitoxin System

Hannes M. Beyer and Hideo Iwaï*[a]

Protein-splicing domains are frequently used engineering tools that find application in the in vivo and in vitro ligation of protein domains. Directed evolution is among the most promising technologies used to advance this technology. However, the available screening systems for protein-splicing activity are associated with bottlenecks such as the selection of pseudo-positive clones arising from off-pathway reaction products or fragment complementation. Herein, we report a stringent screening method for protein-splicing activity in *cis* and *trans*, that exclusively selects productively splicing domains. By fusing splicing domains to an intrinsically disordered region of the antidote from the *Escherichia coli* CcdA/CcdB type II toxin/antitoxin system, we linked protein splicing to cell survival. The screen allows selecting novel *cis*- and *trans*-splicing inteins catalyzing productive highly efficient protein splicing, for example, from directed-evolution approaches or the natural intein sequence space.

Protein splicing is a naturally occurring post-translational processing of proteins, in which an intervening protein (intein), residing within a host protein, catalyzes self-excision and concomitant ligation of the N- and C-terminal flanking host protein sequences (N- and C-exteins) with a peptide bond.[1] Inteins typically insert near the active site of essential host proteins, thus requiring efficient intein self-excision to prevent the loss of fitness or possible cell death. Inteins generally provide no benefit to host organisms under normal conditions and, for this reason, are often regarded as selfish parasitic elements.

The unique autocatalytic reaction catalyzed by inteins, which is capable of ligating protein domains, has enabled diverse biotechnological applications of inteins such as protein purification, ligation, and chemical modification.[2–8] Protein splicing can not only occur in *cis* but also in *trans*, through artificially or naturally split inteins, leading to the ligation of individual polypeptides. Importantly, protein splicing tolerates various non-native extein contexts, although their junction sequences can strongly attenuate the splicing efficiency in *trans*

[a] Dr. H. M. Beyer, Dr. H. Iwaï
    Research Program in Structural Biology and Biophysics
    University of Helsinki
    Viikinkaari 1, 00014 Helsinki (Finland)
    E-mail: hideo.iwai@helsinki.fi

as well as in *cis*.[9–12] Protein *trans*-splicing (PTS) by split inteins has opened many possibilities for connecting two or more polypeptide chains with peptide bonds in vitro or in vivo. Intein-mediated protein engineering has thus become increasingly popular for chemical, biological, and synthetic biological applications. Despite strong interest in the use of inteins, there are a number of limitations associated with the split intein approach, such as ligation kinetics, junction and extein sequence dependencies, and the limited solubility of split intein fragments; these make it necessary to adapt inteins to each application.[13–15] Therefore, engineering and/or discovering "super" inteins with robust features that overcome one or more of these issues have been an ongoing quest. Directed evolution,[5,12,16–18] consensus sequence propagation,[19] rational design,[20–22] and genome mining[23] have been successfully applied to alleviating some of the issues associated with split inteins for PTS.
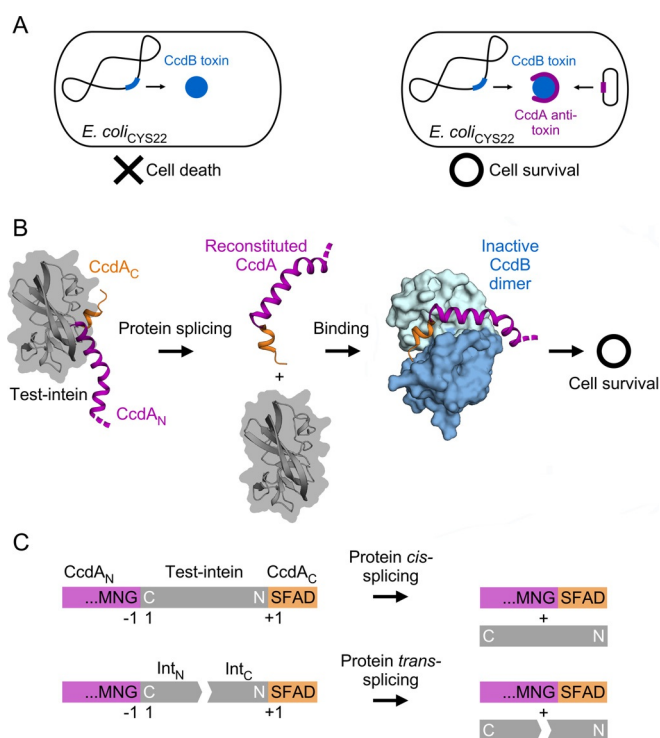
Directed evolution is a powerful method for exploring a large sequence space by linking genotypes and phenotypes in diverse manners. Kanamycin-resistant gene (Kan^R) has been successfully used to identify intein sequences with increased activity in one or more junction sequence contexts because the Kan^R system gives rise to cell growth only when active aminoglycoside-3′-phosphotransferase (3′-APH) is produced.[11,12,21,24,25] The activities of β-lactamase, β-galactosidase, and the fluorescence of GFP have also been used to monitor the activity of inteins for screening purposes.[16,26,27] However, these systems might not be sensitive enough to distinguish off-pathway cleavage reactions from productive splicing because many proteins—despite being split—can fold properly and reconstitute an active form when the two split fragments associate even without covalent connection between them; this is known as fragment complementation. To avoid this pitfall, not only has the intein to be inserted at a crucial position (e.g., near the active site) of the reporter protein, but also the reporter function must be capable of detecting that a covalent bond connects the fragments. Additionally, the system must be sensitive to branched intermediates that are unable to complete the splicing reaction. Therefore, it would be of great interest to develop some different screening methods that are exclusively sensitive to productive protein splicing.

Herein, we report a novel screening method for protein-splicing activity in *cis* and in *trans*, in which productive protein splicing is selected by cell growth. In this system, we exploit the CcdA/CcdB type II toxin/antitoxin system from *Escherichia coli*,[28,29] which has been used for the selection of positive clones during gene cloning (Figure 1 A).[30–34] The toxic effect of CcdB, a potent poison of the A subunit of topoisomerase II gyrase A, can be neutralized by complexation with the antitoxin

Figure 1. Toxin/antitoxin intein screen. A) The *E. coli* strain CYS22 harbors a chromosomal insertion of the *ccdB* gene, causing cell death (left). Production of the CcdA antitoxin, for example, through expression from a plasmid, neutralizes the effect of CcdB, and cells become viable (right). B) To screen functional inteins, a test intein is inserted into the C-terminal part of the CcdA antitoxin. Intein-mediated protein splicing reconstitutes potent CcdA, inactivating CcdB and causing selection through cell survival. The figure was generated based on the gp41-1 intein structure (PDB ID: 6QAZ)[38] and the partial CcdA/B complex (PDB: 3HPW).[35] C) Junction sequences resulting from the insertion of *cis*- (single polypeptide) or *trans*- (Int$_N$ and Int$_C$) splicing inteins into CcdA. The common intein residue numbering is shown.

unlikely to influence the protein-splicing reaction.[35] Inserting the intein within the CcdA antitoxin interaction interface to the CcdB toxin ensures completion of the protein-splicing reaction, thus rendering the screen sensitive to, for example, branched intermediates. Additionally, the cloning system could allow convenient construction of DNA libraries.

To test the concept of exploiting the CcdA/CcdB system, we cloned several inteins with glycine at the −1 position and SFAD as the C-terminal extein fragment into the StablyCloning vector pSTC3.1 (Figure 1 C). We took the +1 position of inteins as the first residue of SFAD, which usually is one of serine, cysteine, or threonine for most inteins. First, we selected the DnaB intein from *Nostoc punctiforme* (*Npu*DnaB) to test the screening system because *Npu*DnaB intein has serine at the +1 position (+1S) and a high tolerance of variations at the splicing junctions.[23] The plasmid containing the engineered mini-intein *Npu*DnaB$^{\Delta283}$ with SFAD inserted after residue 63 of CcdA in the pSTC3.1 vector indeed grew in *E. coli* strain CYS22 bearing the *ccdB* toxin gene in the chromosome (Table 1 and Figure 2 A, B). As a control, we introduced a C-cleaving C1G mutation at the first (catalytic) residue of *Npu*DnaB$^{\Delta283}$ intein, which resulted in no colonies when transformed into *E. coli*

CcdA; this rejuvenates CcdB-gyrase A complexes through allosteric regulation.[35] Importantly, the intrinsically disordered nature of the C-terminal domain of CcdA (residues 42–72) is critical for the antitoxin mechanism. Splitting CcdA at the C-terminal domain is thus detrimental to the antitoxin activity. The StabyCloning system (Delphi Genetics) ingeniously takes advantage of a vector containing a fragment of the *ccdA* gene that lacks the sequence encoding residues 64–72. A 14 bp DNA fragment in a primer for gene cloning adds an additional sequence for serine-phenylalanine-alanine-aspartic residue (SFAD) upon successful gene insertion. This will produce a functional CcdA protein and neutralize the CcdB toxin encoded in the chromosomally inserted *ccdB* gene of *E. coli* strain CYS22. Hence, StabyCloning restricts growth to cells bearing a successfully ligated plasmid with the designed 14 bp DNA sequence encoding SFAD.[36,37] We thought that the short length of SFAD and intrinsically disordered nature of CcdA would be ideal for monitoring protein-splicing activity without being affected by the extein substrates (Figure 1 B, C).[35] The SFAD peptide is unlikely to complement the activity of CcdA, thereby reducing pseudo-positives resulting from off-pathway cleavages. Furthermore, the three-dimensional structure of CcdA suggests that specific conformations of the extein sequences would be

Table 1. Survival CFUs obtained from assaying intein *cis* splicing.[a]		
Intein	CFUs × 10³	Plasmid
*Npu*DnaB$^{\Delta283}$	16.30 ± 1.18	pBHSTC97
*Npu*DnaB$^{\Delta283}$ (C1G, inactive)	0.00 ± 0.00	pHBSTC159
gp41-1	6.10 ± 0.56	pHYSTC299
gp41-1 (C1G, inactive)	0.00 ± 0.00	pHBSTC147
gp41-1 (N125A, inactive)	0.00 ± 0.00	pHBSTC181
*Npu*DnaE (+1C)	0.00 ± 0.00	pBHSTC216
*Npu*DnaE (+1S)	0.01 ± 0.00	pBHSTC214
[a] Averaged CFU numbers with one standard deviation obtained from three replicates are shown except for tests on the *Npu*DnaE intein, for which four replicates were averaged. For plasmids see Table S1.		

CYS22, thus supporting the idea that the insertion of an intein at this position abolishes the antitoxin activity (Figure 2 B). We were intrigued to test the widely used naturally split DnaE intein from *N. puntiforme* (*Npu*DnaE) having +1Cys, because *Npu*DnaE intein has a high activity with phenylalanine at the +2 position.[10] However, subjecting the engineered *cis*-splicing *Npu*DnaE intein with +1Cys did not result in any growth of *E. coli* CYS22; this suggests that serine within SFAD might be crucial for the antitoxin activity (Table 1 and Figure 2 C). Next, we tested *Npu*DnaE intein with +1Ser, which can splice in some contexts but is also known to produce cleaved products.[12,25,39] The cells transformed with the vector bearing *Npu*DnaE intein with +1Ser gave no or very few colonies (Table 1 and Figure 2 D). This result backs up the concept that the CcdA/CcdB system does not produce pseudo-positive colonies due to off-pathway cleavages.

Another interesting intein to test with this system is the naturally split gp41-1 intein, which has very high ligation kinetics and was previously identified from metagenomics sequen-
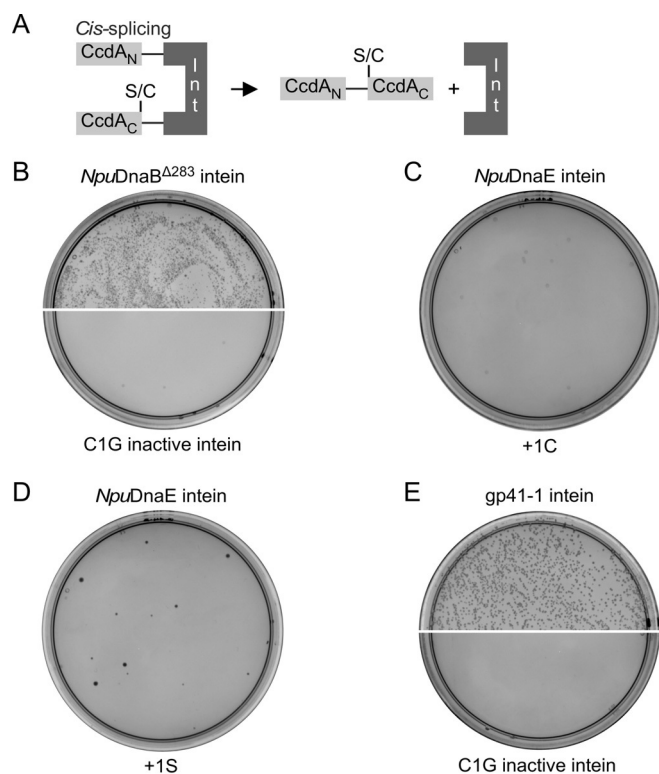
Figure 2. Cell survival and death caused by functional and inactive *cis*-splicing inteins. A) Schematic *cis*-splicing of inteins with CcdA$_N$ and CcdA$_C$ exteins. The location of residue position $+1$ (S/C) within CcdA$_C$ is shown. B) Survival of *E. coli* CYC22 cells expressing *cis*-splicing *Npu*DnaB$^{\Delta283}$ intein inserted into the CcdA antitoxin (top) and cell death after introducing the inactivating C1G mutation (bottom). C) Cell survival test by using the *cis*-splicing *Npu*DnaE intein with $+1$C C-terminal splice junction (CcdA$_C$ = CFAD amino acid sequence). D) Cell survival test of the *Npu*DnaE intein with $+1$S (CcdA$_C$ = SFAD amino acid sequence). E) Experiment as in (B) but using *cis*-splicing gp41-1 intein.

ces.[13,40] The gp41-1 intein also has serine at the $+1$ position as the natural extein sequence, suitable for the CcdA/CcdB screening system. First, we generated a *cis*-splicing gp41-1 intein by fusing the N- and C-terminal split fragments.[38] While the control with the inactive gp41-1 intein carrying the C1G mutation gave no colonies, the plasmid with the active gp41-1 intein with SFAD as the C-extein also survived when transformed in *E. coli* CYS22, thus confirming that the screening method can be used for different inteins with $+1$Ser as the natural extein residue (Table 1 and Figure 2E). The lower number of colony-forming units (CFUs) observed for the g41-1 intein could be attributed to the lower tolerance of alterations within the junction sequence compared to the *Npu*DnaB mini-intein, although the number of CFUs might not always correlate with productive splicing of the inserted intein. We also tested another control by mutating the last asparagine residue of the gp41-1 intein to alanine (N125A). This mutation halts the reaction at the branched intermediate where the extein moieties are already covalently connected by an ester bond but cyclization of the asparagine residue cannot take place to release the intein.[12] This control did not produce surviving

CFUs, thereby corroborating the hypothesis that the system is sensitive to branched intermediates not having completed the splicing reaction. CcdB complexation by CcdA requires excision of the intein from the interaction interface and peptide ligation of split CcdA fragments (Table 1 and Figure 1B).

Inteins are single-turnover enzymes that carry their substrates on the same polypeptide chain. Therefore, the general concept of k_{on}/k_{off} of the substrates is not valid. In the case of protein *trans*-splicing, the association and folding of the two split intein fragments is a critical step. For the two fragments of the naturally split *Npu*DnaE intein, a binding affinity of 1.2 nM was reported.[41] The high affinity of the split intein is advantageous for bringing two foreign extein domains together, allowing the ligation of two globular domains fused to the split fragments. However, when the split fragments are inserted into a globular enzyme such as 3′-APH for *trans*-splicing, the split enzyme fragments can associate together even if there is no contribution from the split intein fragments. In other words, the extein sequences could force split intein fragment association. Thus, Kan^R-based screening likely exhibits inaccurate *trans*- and *cis*-splicing kinetics. Although this is not an issue for the directed evolution of naturally split inteins with a high affinity between Int$_N$ and Int$_C$, the split Kan^R system could be problematic when evolving split intein fragments with no or very low affinity, thus limiting their use in different extein contexts.

One disadvantage of the PTS approach is the size of split inteins. The naturally split *Npu*DnaE and gp41-1 inteins have 36–37-residue C-terminal fragments—long enough to discourage chemical synthesis. Because of this, several groups, including our group, have been trying to engineer or discover shorter split intein fragments. We discovered that the six-residue C-terminal fragment of the *Npu*DnaE intein is sufficient for *trans*-splicing, albeit at a poorer efficiency than the original split intein.[42] Next, we looked at whether the CcdA/CcdB toxin system could be directly useful for the directed evolution of such shorter intein fragments. The Kan^R system has only been used for the directed evolution of *cis*-splicing and a natural split intein, presumably because of the above-mentioned reasons. Due to the requirement of $+1$Ser, we tested a few artificially split inteins derived from the *cis*-splicing gp41-1 intein plasmid by inserting a DNA sequence containing a stop codon for the N-terminal fragment and a start codon for the C-terminal fragment separated by a ribosome-binding site. When split at the natural site located 37 residues from the C terminus of the gp41-1 intein, the fragments (gp41-1$_{\Delta C37}$/gp41-1$_{C37}$) clearly resulted in cell survival (Table 2 and Figure 3). However, the newly engineered split gp41-1 intein split at position C14 (gp41-1$_{\Delta C14}$/gp41-1$_{C14}$), which is split before 14 residues from the C terminus of the gp41-1 intein, did not allow cell growth, thus suggesting that this split site did not produce sufficient active CcdA to neutralize the CcdB toxin (Table 2 and Figure 3C). Similarly, gp41-1$_{\Delta C7}$/gp41-1$_{C7}$, which is split in front of the C-terminal seven residues of *cis*-splicing gp41-1 intein, did not give rise to positives when using the CcdA/CcdB system (Table 2 and Figure 3C). These observations were mainly consistent with *trans*-splicing results of these split inteins when

Table 2. Survival CFUs obtained from assaying intein *trans*-splicing.[a]		
Split intein pair	CFUs×10³	Plasmid
gp41-1$_{\Delta C37}$/gp41-1$_{C37}$	4.03 ± 0.30	pBHSTC104
gp41-1$_{\Delta C37}$/gp41-1$_{C37}$ (C1G, inactive)	0.00 ± 0.00	pHBSTC146
gp41-1$_{\Delta C14}$/gp41-1$_{C14}$	0.01 ± 0.00	pBHSTC120
gp41-1$_{\Delta C7}$/gp41-1$_{C7}$	0.01 ± 0.00	pBHSTC111
gp41-1$_{N12}$/gp41-1$_{\Delta N12}$	0.00 ± 0.00	pBHSTC204
*Npu*DnaB$^{\Delta 283}_{\Delta C39}$/*Npu*DnaB$^{\Delta 283}_{C39}$	0.00 ± 0.00	pHBSTC160
[a] Averaged CFU numbers with one standard deviation obtained from three replicates are shown. For plasmids see Table S1.		
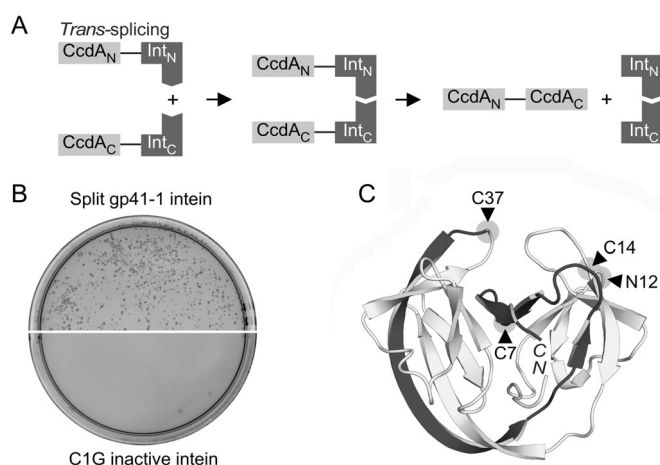
Figure 3. Cell survival and death caused by functional and inactive *trans*-splicing inteins. A) Schematic intein *trans*-splicing with CcdA$_N$ and CcdA$_C$ exteins. B) Test for *E. coli* CYC22 cell survival mediated by CcdA synthesis with *trans*-splicing gp41-1 intein split at the natural position C37 as shown in (C). An inactive variant carrying the C1G mutation is shown at the bottom. C) Structure of the *cis*-splicing gp41-1 intein variant (PDB ID: 6QAZ) highlighting the tested split positions N12, C37, C14, and C7.[38] The natural N-terminal fragment is shown in white, the C-terminal part in gray.

using the B1 domain of the *Streptococcus* sp. IgG binding protein G (GB1) as exteins, although it cannot be directly compared with the CcdA/CcdB system due to the different extein sequences (Figure S1 A and B in the Supporting Information). We also tested the N-terminal split gp41-1 intein (gp41-1$_{N12}$/gp41-1$_{\Delta N12}$) in the CcdA/CcdB system with similar results (Table 2 and Figure 3 C). The short fragment of the disordered region of CcdA does not have any affinity for the remaining CcdA fragment. We speculate that the pairs of gp41-1$_{\Delta C14}$/gp41-1$_{C14}$, gp41-1$_{\Delta C7}$/gp41-1$_{C7}$, and gp41-1$_{N12}$/gp41-1$_{\Delta N12}$ have no or only very weak affinity for each other, thereby producing little or no active CcdA in the cell. Generating a split variant of the engineered *Npu*DnaB$^{\Delta 283}$ mini-intein (*Npu*DnaB$^{\Delta 283}_{\Delta C39}$/ *Npu*DnaB$^{\Delta 283}_{C39}$) also did not result in cell survival (Table 2). Artificially splitting *cis* splicing inteins often results in low solubility, thus causing the activity to fall below the detection threshold.[20] Thus, the CcdA/CcdB system is a stringent system for screening productive protein splicing.

Encouraged by the stringency of the CcdA/CcdB reporter system, we next attempted to evolve a C7 version of the gp41-1 split intein (Table 2, Figure S1) with the goal of increasing the binding affinity between the fragments. We generated a library with a diversity of 5×10^6 by using the pSTC3.1 vector system and degenerated oligonucleotides to introduce codon variation for four out of the seven possible positions within the C7 C-terminal split fragment (gp41-1$_{C7-4aa}$, Table S1). Upon transformation of the library into *E. coli* CYC22 cells, we were indeed able to isolate surviving clones (Figure S2). Among those, most clones carried unclear sequence information within the mutated region, possibly resulting from the propagation of a plasmid mixture. However, we repeatedly isolated a clone propagating the seven-residue sequence NKRSGHN evolved from the original NDILTHN sequence of the C-terminal intein fragment. To our surprise, this isolated peptide sequence did not enhance the splicing efficiency in our model system (Figure S1 C). We think that the small library size is probably insufficient to cover a sequence space containing an improved active C7 peptide, if existing at all, and that the isolated peptide sequence neutralizes the CcdB toxin through a hitherto unknown mechanism, thus resulting in the relatively high background of this library (30 colonies per ng). We found that the neutralization of CcdB likely occurs on the protein level because altering the codons used in the isolated sequence also led to cell survival ($>10\,000$ CFUs, Table S1).

Overcoming the limitations of currently applied protein-splicing screens suitable for engineering enzymes by directed evolution will open novel intein design approaches. However, one disadvantage of the CcdA/CcdB system compared to the Kan^R screening method is the lack of systematic CcdB toxicity control (e.g., adjusting the kanamycin concentration when using Kan^R). On the other hand, stringent screening by the CcdA/CcdB method could easily complement other screening methods using different reporter systems because of the simple commercial cloning kit. The sequential combination of different screening methods might be particularly useful for finding "super" inteins. For example, a preselected pool from the Kan^R system could be subjected to the more stringent CcdA/CcdB system. As the StabyCloning system was developed as a cloning kit, one can simultaneously combine cloning of inteins with testing their protein-splicing activity.

To date, more than 1500 intein genes have been identified from sequence databases; however, protein splicing of only a small fraction of these inteins has been tested experimentally. Cloning and conventional SDS-PAGE analysis to check protein splicing for each intein from various genomes can be time consuming and labor intensive. The CcdA/CcdB system described here could provide accurate screening of active inteins with $+1$Ser from available genomic sequences, which tolerate the variation of the extein sequences, thereby allowing us to explore the intein sequence space more efficiently.

In summary, we have reported a simple survival screening method that exclusively selects productive protein-splicing activity. This method, derived from a commercial cloning kit, makes use of the intrinsically disordered part of the bacterial antitoxin CcdA to neutralize the toxicity of CcdB. The short

extein sequence minimizes potential conformational interferences of the exteins on the protein-splicing efficiency, such as the contribution or inhibition of the fragment association, thereby distinguishing productive splicing from off-pathway cleavages. This new tool for exploring the intein sequence space will complement existing approaches and accelerate the hunt for "super" inteins, which might be used more universally with various targets in synthetic and chemical biology.

Experimental Section

Molecular cloning and origins of plasmids: All plasmids and oligonucleotides used and designed in this study are described in Table S1. Plasmids based on the pSTC3.1 vector system (Delphi Genetics) were used to produce intein-inserted CcdA antitoxin controlled by the *mob* promoter with ampilicin resistance and pBR322 replication origin.[37]

Protein production, purification, and analytics: All recombinant proteins were produced in strains of *E. coli*. Expression details are given in Table S2. Plasmids were transformed into chemically competent ER2566 (NEB) or CYS22 (GE-STC1-22, lot # STC1914, Delphi Genetics) strains of *E. coli*. For transformation into CYS22, plasmid DNA (20 ng) was mixed with competent cells (20 µL). After 30 min of incubation on ice, cells were subjected to a 45 s heat shock at 42 °C. Following 2 min of incubation on ice, cells were recovered after the addition of the regeneration medium provided (180 µL; P_UNQUNE26_796/1, lot # STC1914, Delphi Genetics) for 1 h at 37 °C with shaking. Unless indicated otherwise, the suspension (100 µL) was subsequently spread on LB-agar plates containing ampicillin (100 µg mL^{-1}). For control plasmids encoding the C1G intein mutation, all the sample (200 µL) was spread. The plates were incubated at 37 °C overnight until CFUs became visible; they were then counted using Fiji (ImageJ, version 2.0.0-rc-68).[43] CFUs per 200 µL bacteria suspension are given. Expression cultures (5 mL) were harvested by centrifugation (4500 g) for 10 min. Cells were subsequently lysed using B-PER bacterial protein extraction reagent (400 µL, Thermo Scientific) according to the manufacturer's instructions, and the proteins were IMAC-purified with Ni^{2+}-NTA spin columns (Qiagen). Proteins were eluted in elution buffer (100 µL; 50 mM NaH$_2$PO$_4$ pH 8.0, 300 mM NaCl, 250 mM imidazole). Purified proteins were separated by SDS-PAGE (16.5%) and visualized by staining with Coomassie Brilliant Blue. Plate and gel images were acquired using a Gel DocEZ Imager (Bio-Rad). Structure figures were generated with PyMOL (The PyMOL Molecular Graphics System, version 1.8.6.0, Schrödinger, LLC).

Library construction and screening: Libraries based on the pSTC3.1 vector system (Delphi Genetics) were prepared as described in Table S1 and transformed into *E. coli* DH5α cells (Thermo Fisher). 5×10^6 CFUs were pooled, and plasmid DNA was isolated by using HiSpeed Plasmid Maxi preparation (Qiagen). Library DNA (200 ng) was transformed into *E. coli* CYS22 cells (Delphi Genetics, 100 µL), as described above.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] H. Paulus, *Annu. Rev. Biochem.* **2000**, *69*, 447–496.
[2] C. Sarmiento, J. A. Camarero, *Curr. Protein Pept. Sci.* **2019**, *20*, 408–424.
[3] D. W. Wood, J. A. Camarero, *J. Biol. Chem.* **2014**, *289*, 14512–9.
[4] P. Szweda, R. Pladzyk, R. Kotlowski, J. Kur, *Protein Expression Purif.* **2001**, *22*, 467–471.
[5] I. V. Thiel, G. Volkmann, S. Pietrokovski, H. D. Mootz, *Angew. Chem. Int. Ed.* **2014**, *53*, 1306–1310; *Angew. Chem.* **2014**, *126*, 1330–1334.
[6] T. C. Evans, J. Benner, M. Q. Xu, *Protein Sci.* **1998**, *7*, 2256–2264.
[7] S. Züger, H. Iwai, *Nat. Biotechnol.* **2005**, *23*, 736–740.
[8] M. Muona, A. S. Aranko, V. Raulinaitis, H. Iwaï, *Nat. Protoc.* **2010**, *5*, 574–587.
[9] A. A. Cooper, Y. J. Chen, M. A. Lindorfer, T. H. Stevens, *EMBO J.* **1993**, *12*, 2575–2583.
[10] H. Iwai, S. Züger, J. Jin, P.-H. Tam, *FEBS Lett.* **2006**, *580*, 1853–1858.
[11] M. Cheriyan, C. S. Pedamallu, K. Tori, F. Perler, *J. Biol. Chem.* **2013**, *288*, 6202–6211.
[12] S. W. Lockless, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10999–11004.
[13] P. Carvajal-Vallejos, R. Pallissé, H. D. Mootz, S. R. Schmidt, *J. Biol. Chem.* **2012**, *287*, 28686–28696.
[14] G. Amitai, B. P. Callahan, M. J. Stanger, G. Belfort, M. Belfort, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 11005–11010.
[15] A. S. Aranko, A. Wlodawer, H. Iwaï, *Protein Eng. Des. Sel.* **2014**, *27*, 263–271.
[16] S. H. Peck, I. Chen, D. R. Liu, *Chem. Biol.* **2011**, *18*, 619–630.
[17] A. R. Buskirk, Y.-C. Ong, Z. J. Gartner, D. R. Liu, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10505–10510.
[18] M. P. Zeidler, C. Tan, Y. Bellaiche, S. Cherry, S. Häder, U. Gayko, N. Perrimon, *Nat. Biotechnol.* **2004**, *22*, 871–876.
[19] A. J. Stevens, Z. Z. Brown, N. H. Shah, G. Sekar, D. Cowburn, T. W. Muir, *J. Am. Chem. Soc.* **2016**, *138*, 2162–2165.
[20] A. S. Aranko, J. S. Oeemig, D. Zhou, T. Kajander, A. Wlodawer, H. Iwaï, *Mol. Biosyst.* **2014**, *10*, 1023–1034.
[21] A. J. Stevens, G. Sekar, N. H. Shah, A. Z. Mostafavi, D. Cowburn, T. W. Muir, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 8538–8543.
[22] J. A. Gramespacher, A. J. Stevens, D. P. Nguyen, J. W. Chin, T. W. Muir, *J. Am. Chem. Soc.* **2017**, *139*, 8074–8077.
[23] S. Elliä, J. M. Jurvansuu, H. Iwaï, *FEBS Lett.* **2011**, *585*, 3471–3477.
[24] J. H. Appleby-Tagoe, I. V. Thiel, Y. Wang, Y. Wang, H. D. Mootz, X.-Q. Liu, *J. Biol. Chem.* **2011**, *286*, 34440–34447.
[25] M. Cheriyan, S.-H. Chan, F. Perler, *J. Mol. Biol.* **2014**, *426*, 4018–4029.
[26] J. S. Kwon, J. Bal, H. M. Hwang, J.-Y. Kim, *J. Microbiol.* **2008**, *46*, 456–461.
[27] M. Neugebauer, J. K. Böcker, J. C. J. Matern, S. Pietrokovski, H. D. Mootz, *Biol. Chem.* **2017**, *398*, 57–67.
[28] E. M. Bahassi, M. H. O'Dea, N. Allali, J. Messens, M. Gellert, M. Couturier, *J. Biol. Chem.* **1999**, *274*, 10936–10944.
[29] T. Madl, L. Van Melderen, N. Mine, M. Respondek, M. Oberer, W. Keller, L. Khatai, K. Zangger, *J. Mol. Biol.* **2006**, *364*, 170–185.
[30] M. Bukowski, A. Rojowska, B. Wladyka, *Acta Biochim. Pol.* **2011**, *58*, 1–9.
[31] S. J. Unterholzner, B. Poppenberger, W. Rozhon, *Mobile Genet. Elem.* **2013**, *3*, e26219.
[32] P. Bernard, *BioTechniques* **1996**, *21*, 320–323.
[33] P. Bernard, P. Gabant, E. M. Bahassi, M. Couturier, *Gene* **1994**, *148*, 71–74.
[34] X. Liang, L. Peng, C.-H. Baek, F. Katzen, *BioTechniques* **2013**, *55*, 265–268.
[35] N. De Jonge, A. Garcia-Pino, L. Buts, S. Haesaerts, D. Charlier, K. Zangger, L. Wyns, H. De Greve, R. Loris, *Mol. Cell* **2009**, *35*, 154–163.
[36] D. Stieber, P. Gabant, C. Szpirer, *BioTechniques* **2008**, *45*, 344–346.
[37] C. Y. Szpirer, M. C. Milinkovitch, *BioTechniques* **2005**, *38*, 775–781.

[38] H. M. Beyer, K. M. Mikula, M. Li, A. Wlodawer, H. Iwaï, *bioRxiv* **2019**, DOI: https://doi.org/10.1101/546465.

[39] J. Zettler, V. Schütz, H. D. Mootz, *FEBS Lett.* **2009**, *583*, 909 – 914.

[40] B. Dassa, N. London, B. L. Stoddard, O. Schueler-Furman, S. Pietrokovski, *Nucleic Acids Res.* **2009**, *37*, 2560 – 2573.

[41] N. H. Shah, E. Eryilmaz, D. Cowburn, T. W. Muir, *J. Am. Chem. Soc.* **2013**, *135*, 18673 – 18681.

[42] J. S. Oeemig, A. S. Aranko, J. Djupsjöbacka, K. Heinämäki, H. Iwaï, *FEBS Lett.* **2009**, *583*, 1451 – 1456.

[43] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676 – 682.