Saudi Journal of Biological Sciences 26 (2019) 854–859

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Engineered *Ssp* DnaX inteins for protein splicing with flanking proline residues

Xiao Zhang [a,b], Xiang-Qin Liu [b,*], Qing Meng [a,*]

[a] Institute of Biological Sciences and Biotechnology, Donghua University, Shanghai, People's Republic of China
[b] Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada

ARTICLE INFO

Article history:
Received 11 June 2017
Revised 19 July 2017
Accepted 19 July 2017
Available online 22 July 2017

Keywords:
Intein
Protein splicing
Flanking prolines

ABSTRACT

Inteins are internal protein sequences capable of catalyzing a protein splicing reaction by self-excising from a precursor protein and simultaneously joining the flanking sequences with a peptide bond. Split inteins have separate pieces (N-intein and C-intein) that reassemble non-covalently to catalyze a protein *trans*-splicing reaction joining two polypeptides. Protein splicing has become increasingly useful tools in many fields of biological research and biotechnology. However, natural and engineered inteins have failed previously to function when being flanked by proline residue at the −1 or +2 positions, which limits general uses of inteins. In this study, different engineered inteins were tested. We found that engineered *Ssp* DnaX mini-intein and split inteins could carry out protein splicing with proline at the +2 positions or at both −1 and +2 positions. Under *in vivo* conditions in *E. coli* cells, the mini-intein, S1 split intein, and S11 split intein spliced efficiently, whereas the S0 split intein did not splice with proline at both −1 and +2 positions. The S1 and S11 split inteins also *trans*-spliced efficiently *in vitro* with proline at the +2 positions or at both −1 and +2 positions, but the S0 split intein *trans*-spliced inefficiently with proline at the +2 position and did not *trans*-splice with proline at both −1 and +2 positions. These findings contribute significantly to the toolbox of intein-based technologies by allowing the use of inteins in proteins having proline at the splicing point.

1. Introduction

Inteins capable of protein splicing reactions have become useful tools in many fields of biological research and biotechnology. For example, controllable inteins capable of conditional protein splicing can be used as molecular switches for controlling the functions of host proteins, in which the protein splicing activity may be triggered by temperature, ligand, or light (Zeidler et al., 2004; Buskirk et al., 2004; Mootz et al., 2003; Tyszkiewicz and Muir, 2008; Binschick et al., 2011). Protein *trans*-splicing by split inteins has been used in transgenic plants to prevent environmental escape of the transgen (Yang et al., 2003; Sun et al., 2001), in the production of cytotoxic proteins by joining non-toxic pieces (Evans et al., 1998), and in a gene therapy procedure to allow delivery of small pieces of a large therapeutic gene (Li et al., 2008). It has also been used for segmental isotope labeling of proteins in NMR studies (Busche et al., 2009), for protein interaction studies and protein localization in cells in two-hybrid procedures (Ozawa et al., 2003), and for site-specific chemical modifications (or labeling) and cyclization of proteins (Mootz, 2009; Muralidharan and Muir, 2006; Evans et al., 2000; Iwai and Pluckthun, 1999; Scott et al., 1999). Furthermore, inteins have also been modified to catalyze site-specific cleavages in proteins, which has been developed into the expressed protein ligation method for protein modifications and the IMPACT method for one-step purification and cleavage of recombinant fusion proteins (Muir, 2003; Chong et al., 1997).

Inteins were discovered in natural organisms and host proteins as an internal protein sequence that catalyzes a protein splicing reaction by self-excising from a precursor protein and simultaneously joining the N-extein and C-extein with a peptide bond (Perler et al., 1994; Saleh and Perler, 2006). The typical mechanism of protein splicing includes the following catalytic steps (Noren et al., 2000): (1) at the N-terminus of the intein, an N-S or N-O acyl rearrangement breaks the peptide bond between N-extein and N-intein and substitutes an ester bond for peptide bond; (2)

* Corresponding authors.
E-mail addresses: paul.liu@dal.ca (X.-Q. Liu), mengqing@dhu.edu.cn (Q. Meng).
Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sjbs.2017.07.010
1319-562X/

following a transesterification reaction, the N-extein is moved to the N-terminus of the C-extein forming an ester bond; (3) the peptide bond at the downstream splice junction is broken through a cyclization of the last residue Asn of intein, which separates the intein from the exteins; and (4) the ester bond between the two exteins is changed to a peptide bond through an S-N or O-N acyl rearrangement. These steps may be disrupted, if the intein is mutated or placed in some non-native host proteins, and result in site-specific cleavages at the N- or C-terminus of the intein without splicing. Most natural inteins are bi-functional inteins consisting of a splicing domain and a homing endonuclease domain. Mini-inteins, which can be natural or engineered, have only a splicing domain that is approximately 140 amino acids in size. Different inteins have low similarities in amino acid sequences, but their crystal structures (at least the splicing domain) are highly conserved. The splicing domain consisting of ~12 β-strands is shaped like a flattened disk, with the N- and C-terminus of the intein located near a centrally located catalytic pocket.

A mini-intein may be converted into a split intein by splitting the intein sequence into two separate parts (N-intein and C-intein), assuming that the separate parts can reassemble non-covalently to catalyze a protein *trans*-splicing reaction (Saleh and Perler, 2006). For protein *trans*-splicing, two recombinant polypeptides are produced separately, with N-precursor fusion protein being an N-extein fused to the N-intein, and C-precursor fusion protein being the C-intein fused to a C-extein. In *trans*-splicing, the N- and C-inteins are removed, and the N- and C-exteins are joined with a peptide bond to form the spliced protein. Different types of split inteins may be engineered by splitting an intein sequence at different places. A conventional or S0 type of split intein has the split site near the middle of the intein sequence while unconventional split inteins have the split site near the N- or C-terminus of the intein sequence, resulting in the S1 type or the S11 type split inteins, respectively (Sun et al., 2004; Appleby et al., 2009). The S1 and S11 split inteins are especially applicable for splicing synthetic peptides onto the N- or C-terminus of proteins of interest. The small N-intein (eleven amino acids) of the *Ssp* DnaB S1 split intein can be easily contained in a synthetic peptide with a small N-extein (Ludwig et al., 2009), where the N-extein may contain desired chemical label or modification to be *trans*-spliced onto the N-terminus of a target protein (C-extein). In contrast, the *Ssp* GyrB S11 split intein has an extremely small C-intein (six amino acids) that can be easily included in a synthetic peptide with a small C-extein carrying the desired chemical label or modification to be *trans*-spliced onto the C-terminus of a target protein (N-extein) (Appleby et al., 2009).

Split inteins are powerful tools for protein research and protein engineering, including site-specific chemical labeling or modification of target proteins. The S11 split intein has been used to add fluorescent groups to the C-terminus of recombinant target proteins (Volkmann and Liu, 2009), while the S1 split intein has been used in N-terminal fluorescent labeling of recombinant proteins *in vitro* (Ludwig et al., 2006) and on mammalian cell surface (Ando et al., 2007). The site-specific addition of fluorescent or isotopic labeling to proteins can be useful tools for studying protein structure-function, protein location and trafficking in cells, and biomedical imaging. Adding chemical modifications or unnatural amino acids to proteins, if done site-specifically, can be an effective way of studying the structure-function relationship of proteins. To efficiently carry out site-specific protein labeling and modification through intein-based *trans*-splicing, the following obstacle needs to be overcome. Inteins often exhibit specificities toward the extein amino acid residues flanking the intein and thus limits their general usefulness (Amitai et al., 2009; Iwai et al., 2006). In particular, inteins have not worked well when being flanked by a proline residue either at the -1 positions or at the +2 positions (Iwai et al., 2006; Mathys et al., 1999; Cheriyan et al., 2013) Directed evolution has been used to produce more general inteins, but even these improved inteins do not splice when being flanked by a proline residue at the −1 position (Appleby-Tagoe et al., 2011).

In this study, we tested different engineered inteins *Ssp* DnaX, *Ter* DnaE-3 and *Npu* DnaE to find one that can initiate protein splicing when being flanked by proline residues. It was found that various engineered *Ssp* DnaX mini-intein and split inteins could splice in model proteins when proline was present at the +2 positions or at both −1 and +2 positions.

2. Materials and methods

2.1. Plasmid construction

Plasmids pMTE$_3^{m1}$-m and pMNE^{m1}-m were constructed as follows. The genes of *Ter* DnaE-3 and *Npu* DnaE mini-inteins were amplified by PCR with the primers TE$_3$-m1for (5′-TACDCTCGA GACTTGTTTAACTTATGAGAC-3′), TE$_3$-m1rev (5′-TCA TACCGGTTGGGCAGGGAGGACAGTTGGAAGCAAT-3′) designed from pMTE$_3$-m and NE-m1for (5′-TACGCTCGAGACTTGTTTAAGCTATG-3′), NE-m1rev (5′-TCATACCGGTTGGGCAGGGAGGACAATTAGAAGC TATG -3′) from pMNE-m.

The PCR products were double digested with AgeI/XhoI, inserted into the similarly digested plasmids pMTX-m, respectively. Plasmids pMTE$_3$ ^{m2}-m and pMNE^{m2}-m were constructed using the same method, with primers changed to TE$_3$-m2for (5′-TACGCTCGAGACTTGTCCTCCCTGTTTAACTTATGAG-3′), TE$_3$-m2rev (5′-TCATACCGGTTGGGCAGTTGGAAGCAATTAACC-3′) and NE-m2for (5′-TACGCTCGAGACTTGTCCTCCCTGTTTAAGCTATG-3′), NE-m2rev (5′-TCATACCGGTTGGGCAATTAGAAGCTATGAAGC-3′).

Plasmids pMSX^{m1}-m, -S0, -S1 and -S11 were constructed by replacing the AgeI-XhoI fragment in plasmid pMTX-m with an AgeI-XhoI fragment produced by PCR amplification with primers SX-m1for (5′-TACGCTCGAGACTTGCTTAACGGGGGACTC-3′) and SX-m1rev (5′-TCATACCGGTTGGGCAGGGAGGACAGTTATGGAC TAG-3′), using pMSX-m, -S0, -S1 and -S11 as templates, respectively. Another pair of primers, SX-m2for (5′-TACGCTCGA GACTTGTCCTCCCTGCTTAACGGGG-3′) and SX-m2rev (5′-TCA TACCGGTTGGGCAGTTATGGACTAGTAAGCC-3′) was used for construction of pMSX^{m2}-m, -S0, -S1 and -S11.

To construct the N-precursor fusion protein consisting of maltose binding protein (M) and N-intein, the pMSX$^{(m1/m2)}$-S0, -S1, and -S11 plasmids were double digested with NdeI/ HindIII or AflII/HindIII, and self-ligated, yielding plasmids pMSX$^{(m1/m2)}$-S$_0$ N, -S$_1$ N and -S$_{11}$ N. To construct the C-precursor fusion protein consisting of C-intein and thioredoxin (T), plasmids pMSX $^{(m1/m2)}$-S0 and -S1 were double digested with NdeI/PstI, and the isolated NdeI-PstI fragment was ligated with similarly digested plasmid vector pTTX-S$_1$ C, yielding plasmids pTSX $^{(m1/m2)}$-S$_0$ C and -S$_1$ C. For construction of pESX $^{(m1/m2)}$-S$_{11}$ C, the I$_{S11\ C}$T coding sequence amplified by PCR from pMSX $^{(m1/m2)}$-S11 with primers 5′-GGAAA AACATATGGGCTTACTAGTCCAT-3′ and 5′-CTAGCTAGCCGCCAGGT TAGCGTCGAG-3′, was digested with NdeI/NheI and cloned into similarly digested plasmid vector pESG-S$_{11}$C.

2.2. Protein expression, purification, and splicing

Plasmids pMSX $^{(m1/m2)}$-m, -S0, -S1, -S11, pMTE$_3$ $^{(m1/m2)}$-m and pMNE $^{(m1/m2)}$-m were transformed into *E. coli* DH5α cells, using a standard *E. coli* transformation protocol. The cells were grown in Luria Broth (LB) medium at 37 °C and induced by a final concentration of 0.8 mM IPTG at 25 °C overnight (Arshadullah et al., 2017). Total cellular proteins were analyzed by Western blotting, using anti-thioredoxin (anti-T) antibody (Invitrogen).

The plasmids for N-precursors expression ($MI_{S0\ N}$, $MI_{S1\ N}$ and $MI_{S11\ N}$) were transformed into *E. coli* DH5α cells, while the plasmids for C-precursors expression ($I_{S0C}T$, $I_{S1C}T$ and $I_{S11C}T$) were transformed into *E. coli* BL21 (DE3) cells. The transformed *E. coli* cells were pre-cultured in 3 ml Luria Broth (LB) medium at 37 °C overnight, the culture was transferred into 50 ml Luria Broth (LB) medium for further culturing to an OD_{600} of 0.6, a final concentration of 0.8 mM IPTG was added to induce protein expression at 25 °C overnight. After the cells were collected by centrifugation and lysed using a French Press (14,000 PSI), the cell lysate was centrifuged at 10,000*g* for 15 min to remove any insoluble materials. The soluble N-precursor proteins were purified by amylose resin according to the manufacturer's instructions (New England Biolabs). The soluble C-precursor proteins were purified by Ni–NTA resin according to the manufacturer's instructions (QIAGEN).

N-precursor proteins and C-precursor proteins were mixed in a 1:1 or 10:1 M ratio to allow *trans*-splicing *in vitro* at room temperature with 1 mM DTT added. Splicing products were visualized by Western blotting using an anti-thioredoxin (anti-T) antibody (Invitrogen).

3. Results

To find inteins that can splice with proline at -1 (immediately before the intein) and +2 (second position after the intein) positions, we tested different engineered inteins in a model system. We initially tested *Ssp* DnaX, *Ter* DnaE-3 and *Npu* DnaE mini-inteins for protein *cis*-splicing in a model fusion protein. In the fusion protein, a maltose binding protein (M) was used as the N-extein, a thioredoxin protein (T) was used as the C-extein, with the mini-intein inserted in the middle. A proline residue was placed at the +2 positions or at both −1 and +2 positions. Amino acid residue at the +1 positions (first position after the intein) must be a nucleophilic residue, which is Cys for the *Ssp* DnaX, *Ter* DnaE-3 and *Npu* DnaE inteins. Protein *cis*-splicing would produce a spliced protein (MT) consisting of the maltose binding protein (M) and the thioredoxin (T), with the intein sequence excised. The spliced protein MT could readily be detected by Western blotting using an anti-thioredoxin (anti-T) antibody, with the protein identified by its predicted size matching that of the control protein in addition to antibody recognition.

Our tests revealed that the *Ssp* DnaX mini-intein with proline at the +2 positions or at both −1 and +2 positions could *cis*-splice efficiently in the above model protein in *E. coli* cells, whereas the *Ter* DnaE-3 and *Npu* DnaE mini-inteins spliced inefficiently when proline was present at the +2 positions (Fig.1A) and almost did not splice when proline was present at both −1 and +2 positions (Fig. 1B).

As shown in Fig. 1, the precursor (MIT) had *cis*-spliced after expression in *E. coli* cells, showing a clearly visible spliced protein (MT), which was identified both by its antibody (anti-T) recognition and by its apparent size matching that of the control protein (lane 1). *Ssp* DnaX mini-intein exhibited efficient *cis*-splicing when being flanked by proline at the +2 positions or at both −1 and +2 positions, indicated by the accumulation of spliced protein MT (lane 2). No precursor remained, indicating that the splicing reaction proceeded to completion. The *Ter* DnaE-3 mini-intein spliced inefficiently (∼30%, lane 3 of Fig. 1A) with proline at the +2 positions and did not splice with proline at both −1 and +2 positions (lane 3 of Fig. 1B). The precursor protein underwent N-cleavage, indicated by the cleavage product IT. *Npu* DnaE mini-intein showed ∼30% (lane 4 of Fig. 1A) and ∼5% (lane 4 of Fig. 1B) splicing activity when being flanked by proline at the +2 positions or at both −1 and +2 positions, respectively. The band of N-cleavage product IT indicated that the precursor protein had underwent N-cleavage.

The efficient splicing of *Ssp* DnaX mini-intein prompted us to test whether *trans*-splicing could also occur when the *Ssp* DnaX mini-intein was changed to split inteins. As shown in Fig. 2, the *Ssp* DnaX mini-intein sequence was split at either of three different sites (S0, S1, and S11) to produce the corresponding split inteins, according to the *Ssp* DnaB mini-intein whose crystal structure has been determined (Ding et al., 2003). For each split intein consisting of N-intein and C-intein, recombinant genes were constructed to express two precursor proteins. As illustrated in Fig.3A, the N-precursor protein consisted of the maltose binding protein (M) followed by the N-intein, and the C-precursor protein consisted of the C-intein followed by the thioredoxin protein (T). To detect protein *trans*-splicing activity *in vivo*, the two precursor proteins were co-expressed in the same *E. coli* cell, from a recombinant plasmid that had the two recombinant genes in an operon and after an IPTG-inducible promoter. After gene expression, total cellular proteins were analyzed by Western blotting using an anti-thioredoxin (anti-T) antibody to detect the presence of the C-precursor protein and possibly the spliced protein.

As shown in Fig. 3B and C, the spliced protein (MT) was clearly observed, indicating that the N-precursor and the C-precursor had *trans*-spliced after co-expression in *E. coli* cells. When proline was present at the +2 positions or at both −1 and +2 positions, efficient splicing was detected for *Ssp* DnaX-S1 and S11 split inteins (lanes 3 and 4). No C-precursor protein remained, indicating that the *trans*-splicing reaction was complete. However, *Ssp* DnaX-S0 split intein was found to have efficient *trans*-splicing reaction only when the +2 position was occupied by proline (lane 2 of Fig. 3B). When proline was present at both −1 and +2 positions, no spliced protein was detected, while the C-precursor protein was accumulated and its weak C-cleavage band T was clearly observed (lane 2 of Fig. 3C).

We then tested the *Ssp* DnaX split inteins for protein *trans*-splicing *in vitro*. The N- and C-precursor proteins were expressed separately in *E. coli*, purified, and mixed *in vitro* to allow protein *trans*-splicing. The results detected by Western blotting are shown in Fig. 4 and summarized in Table 1. Fig.4A shows *in vitro trans*-splicing with proline at the +2 positions. A low level of *trans*-splicing activity *in vitro* was detected for the S0 split intein at the two precursor ratios tested, as indicated by the small amount of spliced proteins (∼5%, lane 2 and ∼15%, lane 3) and the remaining precursor. Efficient protein *trans*-splicing was observed for S1 split intein (>60%, lane 4) as well as S11 split intein (>50%, lane 6), and the splicing reached near completion when the N-precursor was used at a 10-fold excess (>90%, lane 5 and >98%, lane 7). No C-cleavage of the C-precursor was observed.

Fig. 4B shows *in vitro trans*-splicing with proline at both −1 and +2 positions. For the S0 split intein, no spliced protein was observed (lanes 2 and 3), indicating a lack of *trans*-splicing activity of this intein *in vitro*. For the S1 and S11 split intein, the spliced protein was clearly observed (lanes 4–7), indicating protein *trans*-splicing *in vitro*. When the two precursors were mixed at 1:1 ratio, a significant amount of the C-precursor remained, indicating that the *trans*-splicing reaction for S1 and S11 split intein proceeded inefficiently (∼30%, lane 4 and ∼20%, lane 6). When the N-precursor was used at a 10-fold excess, very little C-precursor remained for S1 split intein, indicating that the *trans*-splicing reaction proceeded efficiently (>80%, lane 5), while the C-precursor remained for S11 split intein was almost not visible, indicating that the *trans*-splicing reaction reached near completion (>98%, lane 7). For S1 split intein, a portion of the C-precursor also underwent C-cleavage, as indicated by a presence of the cleavage product T. However, the C-precursor for S11 split intien did not undergo C-cleavage, which is different from the S1 split intein.

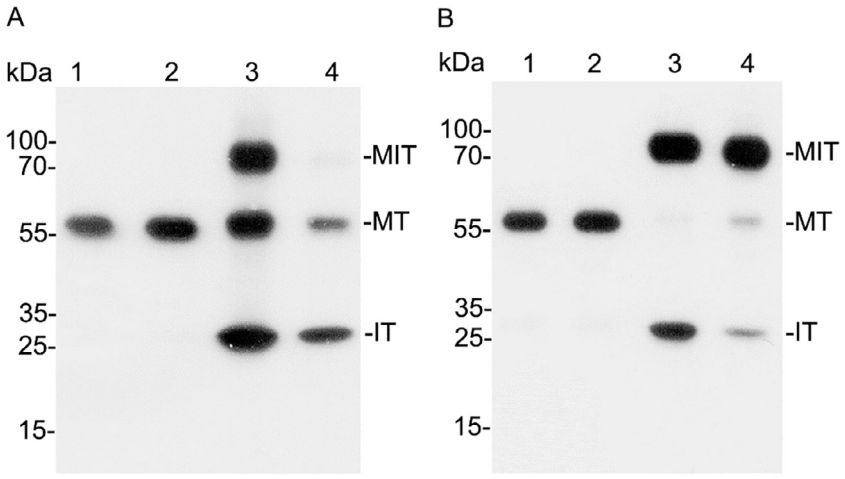

Fig. 1. Protein *cis*-splicing mediated by mini-inteins *in vivo*. The total cellular proteins were analyzed after the precursor proteins were expressed in *E. coli* cells. Protein bands were visualized by Western blotting using an anti-T antibody. The precursor protein is a fusion protein consisting of a maltose binding protein (M) used as the N-extein, a thioredoxin protein (T) used as the C-extein and the mini-intein (I) inserted in the middle. The spliced product MT is marked with "-", together with the precursor protein MIT and cleavage protein IT. (A) Protein *cis*-splicing mediated by mini-inteins with proline at the + 2 positions. Lane1, MT protein used as a size control. Lanes 2–4, protein *cis*-splicing mediated by *Ssp* DnaX, *Ter* DnaE-3 and *Npu* DnaE mini-inteins, respectively. (B) Protein *cis*-splicing mediated by mini-inteins with proline at both −1 and +2 positions. Lane1, MT protein used as a size control. Lanes 2–4, protein *cis*-splicing mediated by *Ssp* DnaX, *Ter* DnaE-3 and *Npu* DnaE mini-inteins, respectively.

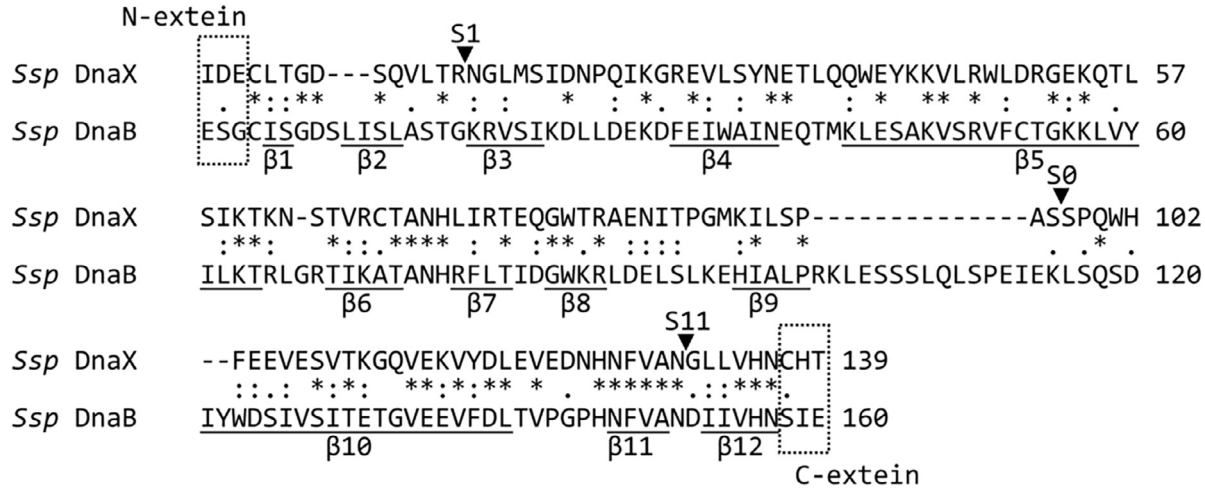

Fig. 2. Amino acid sequence of the *Ssp* DnaX mini-intein. The sequence is aligned with the *Ssp* DnaB mini-intein sequence. Symbols: '-'in the sequence represents gaps introduced to optimize the alignment; '*', ":"and '.' between sequences mark positions of identical, conserved and similar amino acids, respectively. In the *Ssp* DnaX sequence, the S0, S1, S11 split site are marked with arrowheads. In the *Ssp* DnaB sequence, the 12 β-strands (β1 to β12) are underlined.

4. Discussion

In this study, we found that *Ssp* DnaX mini-intein and split inteins could carry out protein splicing efficiently when being flanked by proline residue either at the +2 positions or at both −1 and +2 positions. This finding contrasts with the previous finding that inteins did not work well when being flanked by proline residue either at the -1 positions or at the +2 positions (Iwai et al., 2006; Mathys et al., 1999; Appleby-Tagoe et al., 2011). Previously the engineered *Ssp* DnaB mini-intein did not work well with proline either at the −1 or +2 positions. The natural split intein *Ssp* DnaE did not work well with proline at the +2 positions. Even the improved *Ssp* DnaB mini-intein after directed evolution for general uses spliced inefficiently with proline either at the −1 or +2 positions. And the *Ter* DnaE-3 and *Npu* DnaE mini-inteins used in this study spliced inefficiently or did not splice. Here, under *in vivo* conditions in *E. coli* cells, the mini-intein and the three split inteins of *Ssp* DnaX all spliced efficiently when being flanked by proline at the +2 positions. With proline at both −1 and +2 positions, the mini-intein, S1 split intein, and the S11 split intein spliced efficiently, whereas the S0 split intein did not splice. Under *in vitro* conditions, when the N-precursor was excessive, the C-precursor protein underwent splicing reaction thoroughly, promoting the splicing reaction to completion and improving the splicing efficiency. When the N-precursor was used at a 10-fold excess, the S1 and S11 split inteins *trans*-spliced efficiently with proline at the +2 positions or at both −1 and +2 positions *in vitro*, but the S0 split intein *trans*-spliced inefficiently with proline at the +2 position and did not *trans*-splice with proline at both −1 and +2 positions. Moreover, the splicing was more difficult with proline at both -1 and +2 positions than that with proline only at the +2 positions, with an exception that the S11 split intein showing comparable splicing activity (>98%) under this two flanking proline conditions.

Proline is special among the 20 amino acids, it is known to cause abrupt turns (a kind of disruption) in protein secondary

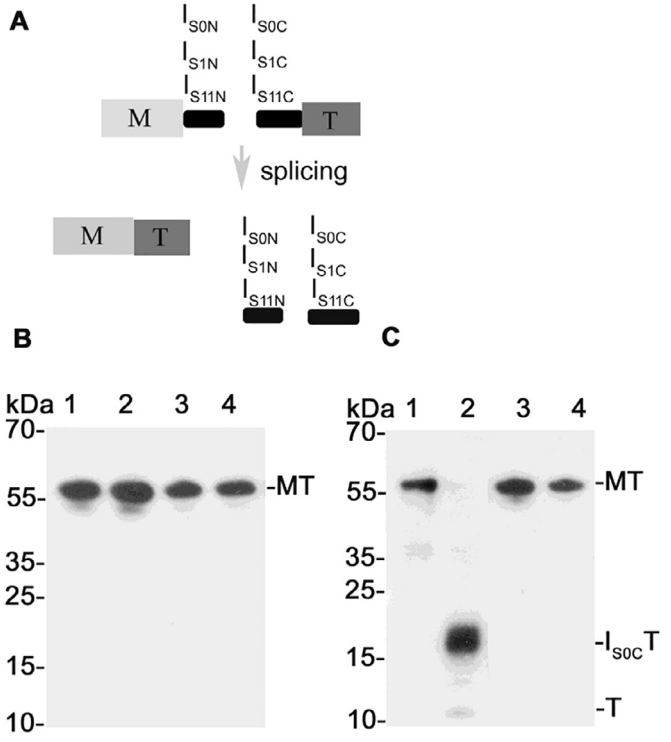

Fig. 3. Protein *trans*-splicing mediated by *Ssp* DnaX split inteins *in vivo*. (A) Schematic illustration of the protein *trans*-splicing reaction. The N-precursor proteins (MI_{S0N}, MI_{S1N} and MI_{S11N}) consist of a maltose binding protein (M) and N-inteins (I_{S0N}, I_{S1N} and I_{S11N}); the C-precursor proteins ($I_{S0C}T$, $I_{S1C}T$ and $I_{S11C}T$) consist of C-inteins (I_{S0C}, I_{S1C} and I_{S11C}) and a thioredoxin protein (T). (B) Western blotting analysis of protein *trans*-splicing. The spliced protein MT is marked with "-", together with the C-precursor $I_{S0C}T$ and cleavage protein T. The total cellular proteins were analyzed after the N-precursor and C-precursor were co-expressed in the same *E. coli* cell. Protein bands were visualized by Western blotting using an anti-T antibody. Lane1, MT protein used as a size control. Lanes 2–4, protein *trans*-splicing mediated by *Ssp* DnaX-S0, S1 and S11 split inteins, respectively, with proline at the +2 positions. (C) Western blotting analysis of protein *trans*-splicing. Lane1, MT protein used as a size control. Lanes 2–4, protein *trans*-splicing mediated by *Ssp* DnaX-S0, S1 and S11 split inteins, respectively, with proline at both −1 and +2 positions.

Table 1
Summary of Splicing Activities of split inteins with flanking proline residues.

intein	+2		−1/+2	
	N:C = 1:1	N:C = 10:1	N:C = 1:1	N:C = 10:1
S0	+	+	–	–
S1	+++	+++++	++	++++
S11	+++	++++++	++	++++++

For each split intein listed on the left, the position of proline residue is shown on the top as +2 and −1/+2. It means intein was being flanked by proline at the +2 positions or at both −1 and +2 positions. The N:C = 1:1 and 10:1 means the molar ratio of N-precursor and C-precursor. The splicing efficiency was estimated as percentage of C-precursor that had been converted into the spliced protein MT (see Fig. 3) and presented as ++++++ for >98%, +++++ for 90–95%, ++++ for 80–90%, +++ for 50–65%, ++for 20–35%, + for 5–15% and – for 0 (not detectable).

structures (alpha helix, beta-strands), because its side chain is cyclic with the alpha-amino group. Therefore, proline at the −1 or +2 positions of inteins may cause 'disruption' at the active site and prevent splicing. A likely explanation for *Ssp* DnaX intein splicing when being flanked by proline is that the *Ssp* DnaXmini-intein and S1, S11 split inteins have flexible structure comparing with other inteins to avoid 'disruption', catalyzing the protein splicing with N- and C-terminus of the intein located near a centrally located catalytic pocket. Whereas the S0 split intein is not that flexible like S1 and S11 split inteins. In addition, probably the structure of S11 split intein was more flexible than S1 split intein, because it is showing comparable splicing activity *in vitro* with proline at the +2 positions or at both −1 and +2 positions when the N-precursor was used at a 10-fold excess.

Inteins often exhibit specificities toward the neighboring amino acids and thus limits their application in host proteins. In a non-native host protein, the inserting site of intein must avoid being flanked by proline, especially with proline at −1 or +2 positions as previously reported, because the proline residue is near the splice sites and influence the intein's catalytic center. Our findings suggested that the application of intein could not be limited by proline residue; the intein could catalyze efficient splicing reaction when being flanked by proline, facilitating a more general use of

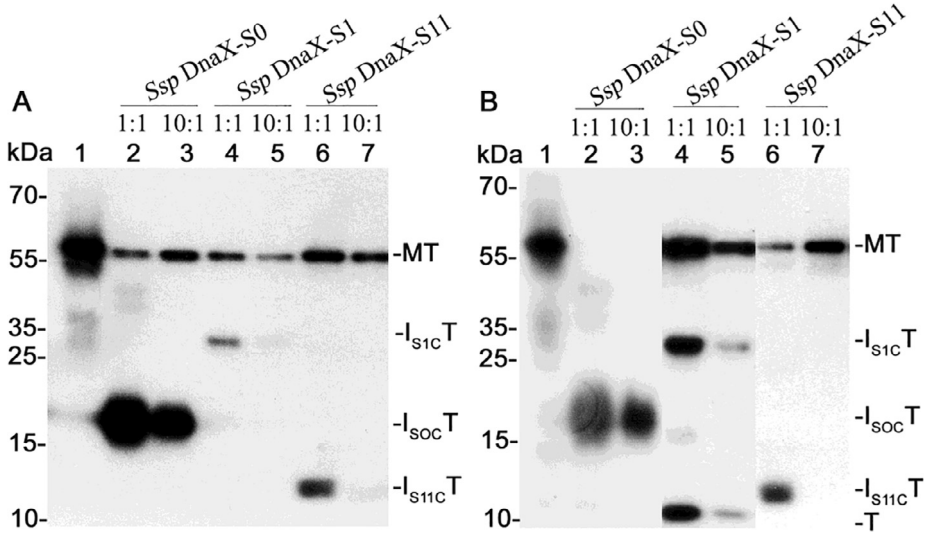

Fig. 4. *Trans*-splicing mediated by split inteins *in vitro*. After the N-precursor proteins (MI_{S0N}, MI_{S1N} and MI_{S11N}) and the C-precursor proteins ($I_{S0C}T$, $I_{S1C}T$ and $I_{S11C}T$) were separately expressed in *E. coli* cells, purified, they were mixed in a 1:1 or 10:1 M ratio *in vitro* to allow *trans*-splicing. The spliced protein MT is marked with "-", together with the C-precursor protein and cleavage protein T. Protein bands were visualized by Western blotting using an anti-T antibody. (A) *Trans*-splicing mediated by split inteins with proline at the +2 positions. Lane1, MT protein used as a size control. Lanes 2–7, protein *trans*-splicing with N: C being 1:1 and 10:1 mediated by the *Ssp* DnaX-S0, S1 and S11 split inteins, respectively. (B) *Trans*-splicing mediated by split inteins with proline at both −1 and +2 positions. Lane1, MT protein used as a size control. Lanes 2–7, protein *trans*-splicing with N: C being 1:1 and 10:1 mediated by the *Ssp* DnaX-S0, S1 and S11 split inteins, respectively.

intein. Overall, these findings contribute significantly to the toolbox of intein-based technologies because they potentially allow protein splicing to be carried out in proteins having proline at the intein insertion site.

Acknowledgments

This work was supported by a grant to X.-Q.L. from National Sciences and Engineering Research Council Grant of Canada, by research grants to Q.M. from Shanghai Science and Technology Commission project (NO. 14521100700), Shanghai Science and Technology Commission of International Cooperation project (NO. 14520720200), the State Bureau of Foreign Experts "Advanced Foreign Experts project" (NO. GDW20143100071), Universities Characteristic Project of the Ministry of Education of China (NO. TS2011DHDX025), the National Natural Science Foundation of China (NO. 31570721) and by grants to Y. Lin from National Natural Science Foundation of China (No. 31470836), Shanghai Natural Science Foundation (No. 13ZR1400600).

Author Contributions

Conceived and designed the experiments: XZ XQL QM. Performed the experiments: XZ. Analyzed the data: XZ XQL QM. Contributed reagents/materials/analysis tools: XQL QM. Wrote the paper: XZ.

References

Amitai, G., Callahan, B.P., Stanger, M.J., Belfort, G., Belfort, M., 2009. Modulation of intein activity by its neighboring extein substrates. Proc. Natl. Acad. Sci. USA 106 (27), 11005–11010.

Ando, T., Tsukiji, S., Tanaka, T., Nagamune, T., 2007. Construction of a small-molecule-integrated semisynthetic split intein for in vivo protein ligation. Chemical communications (Cambridge, England). (47), 4995–7.

Appleby, J.H., Zhou, K., Volkmann, G., Liu, X.Q., 2009. Novel split intein for trans-splicing synthetic peptide onto C terminus of protein. J. Biol. Chem. 284 (10), 6194–6199.

Appleby-Tagoe, J.H., Thiel, I.V., Wang, Y., Wang, Y., Mootz, H.D., Liu, X.Q., 2011. Highly efficient and more general cis- and trans-splicing inteins through sequential directed evolution. The J. Biol. Chem. 286 (39), 34440–34447.

Arshadullah, M., Suhaib, M., Baber, R., Usama, M., Zaman, B.U., Mahmood, I.A., Hyder, S.I., 2017. Growth of Chenopodium quiona Wild under Naturally Salt Affected Soils. Malaysian J. Sustain. Agr. 1 (1), 01–03.

Binschik, J., Zettler, J., Mootz, H.D., 2011. Photocontrol of protein activity mediated by the cleavage reaction of a split intein. Angewandte Chemie (International ed in English). 20;50 (14), 3249–3252.

Busche, A.E., Aranko, A.S., Talebzadeh-Farooji, M., Bernhard, F., Dotsch, V., Iwai, H., 2009. Segmental isotopic labeling of a central domain in a multidomain protein by protein trans-splicing using only one robust DnaE intein. Angew. Chem. Int. Ed. Engl. 48 (33), 6128–6131.

Buskirk, A.R., Ong, Y.C., Gartner, Z.J., Liu, D.R., 2004. Directed evolution of ligand dependence: small-molecule-activated protein splicing. Proc. Natl. Acad. Sci. USA 101 (29), 10505–10510.

Cheriyan, M., Pedamallu, C.S., Tori, K., Perler, F., 2013. Faster protein splicing with the Nostoc punctiforme DnaE intein using non-native extein residues. J. Biol. Chem. 288 (9), 6202–6211.

Chong, S., Mersha, F.B., Comb, D.G., Scott, M.E., Landry, D., Vence, L.M., 1997. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. Gene 192 (2), 271–281.

Ding, Y., Xu, M.Q., Ghosh, I., Chen, X., Ferrandon, S., Lesage, G., 2003. Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing. J. Biol. Chem. 278 (40), 39133–39142.

Evans Jr., T.C., Benner, J., Xu, M.Q., 1998. Semi synthesis of cytotoxic proteins using a modified protein splicing element. Protein Sci.: a Publication of the Protein Soc. 7 (11), 2256–2264.

Evans Jr., T.C., Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., 2000. Protein trans-splicing and cyclization by a naturally split intein from the dnaE gene of Synechocystis species PCC6803. J. Biol. Chem. 275 (13), 9091–9094.

Iwai, H., Pluckthun, A., 1999. Circular beta-lactamase: stability enhancement by cyclizing the backbone. FEBS Lett. 459 (2), 166–172.

Iwai, H., Zuger, S., Jin, J., Tam, P.H., 2006. Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostoc punctiforme. FEBS Lett. 580 (7), 1853–1858.

Li, J., Sun, W., Wang, B., Xiao, X., Liu, X.Q., 2008. Protein trans-splicing as a means for viral vector-mediated in vivo gene therapy. Hum. Gene Ther. 19 (9), 958–964.

Ludwig, C., Pfeiff, M., Linne, U., Mootz, H.D., 2006. Ligation of a synthetic peptide to the N terminus of a recombinant protein using semisynthetic protein trans-splicing. Angew. Chem. Int. Ed. Engl. 45 (31), 5218–5221.

Ludwig, C., Schwarzer, D., Zettler, J., Garbe, D., Janning, P., Czeslik, C., 2009. Semi synthesis of proteins using split inteins. Methods Enzymol. 462, 77–96.

Mathys, S., Evans, T.C., Chute, I.C., Wu, H., Chong, S., Benner, J., 1999. Characterization of a self-splicing mini-intein and its conversion into autocatalytic N- and C-terminal cleavage elements: facile production of protein building blocks for protein ligation. Gene 231 (1–2), 1–13.

Mootz, H.D., 2009. Split inteins as versatile tools for protein semisynthesis. ChemBioChem 10 (16), 2579–2589.

Mootz, H.D., Blum, E.S., Tyszkiewicz, A.B., Muir, T.W., 2003. Conditional protein splicing: a new tool to control protein structure and function in vitro and in vivo. J. Am. Chem. Soc. 125 (35), 10561–10569.

Muir, T.W., 2003. Semi synthesis of proteins by expressed protein ligation. Annu. Rev. Biochem. 72, 249–289.

Muralidharan, V., Muir, T.W., 2006. Protein ligation: an enabling technology for the biophysical analysis of proteins. Nat. Methods 3 (6), 429–438.

Noren, C.J., Wang, J., Perler, F.B., 2000. Dissecting the chemistry of protein splicing and its applications. Angew. Chem. Int. Ed. Engl. 39 (3), 450–466.

Ozawa, T., Sako, Y., Sato, M., Kitamura, T., Umezawa, Y., 2003. A genetic approach to identifying mitochondrial proteins. Nat. Biotechnol. 21 (3), 287–293.

Perler, F.B., Davis, E.O., Dean, G.E., Gimble, F.S., Jack, W.E., Neff, N., 1994. Protein splicing elements: inteins and exteins–a definition of terms and recommended nomenclature. Nucleic Acids Res. 22 (7), 1125–1127.

Saleh, L., Perler, F.B., 2006. Protein splicing in cis and in trans. Chemical record (New York, NY). 6(4), 183–93.

Scott, C.P., Abel-Santos, E., Wall, M., Wahnon, D.C., Benkovic, S.J., 1999. Production of cyclic peptides and proteins in vivo. Proc. Natl. Acad. Sci. USA 96 (24), 13638–13643.

Sun, L., Ghosh, I., Paulus, H., Xu, M.Q., 2001. Protein trans-splicing to produce herbicide-resistant acetolactate synthase. Appl. Environ. Microbiol. 67 (3), 1025–1029.

Sun, W., Yang, J., Liu, X.Q., 2004. Synthetic two-piece and three-piece split inteins for protein trans-splicing. J. Biol. Chem. 279 (34), 35281–35286.

Tyszkiewicz, A.B., Muir, T.W., 2008. Activation of protein splicing with light in yeast. Nat. Methods 5 (4), 303–305.

Volkmann, G., Liu, X.Q., 2009. Protein C-terminal labeling and biotinylating using synthetic peptide and split-intein. PLoS ONE 4 (12), e8381.

Yang, J., Fox, G.C., Henry-Smith Jr., T.V., 2003. Intein-mediated assembly of a functional beta-glucuronidase in transgenic plants. Proc. Natl. Acad. Sci. USA 100 (6), 3513–3518.

Zeidler, M.P., Tan, C., Bellaiche, Y., Cherry, S., Hader, S., Gayko, U., 2004. Temperature-sensitive control of protein activity by conditionally splicing inteins. Nature Biotechnol. 22 (7), 871–876.