BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Branched Intermediate Formation Is the Slowest Step in the Protein Splicing Reaction of the Ala1 KlbA Intein from *Methanococcus jannaschii*

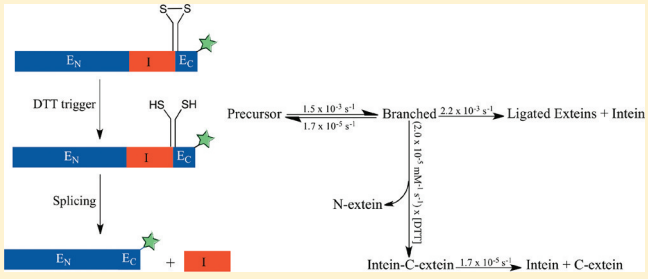
Lana Saleh,*[,†] Maurice W. Southworth,[†] Nancy Considine,[†] Colleen O'Neill,[†] Jack Benner,[†] J. Martin Bollinger, Jr.,[‡] and Francine B. Perler*[,†]

[†]New England Biolabs, Ipswich, Massachusetts 01938, United States

[‡]Department of Chemistry and Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

Supporting Information

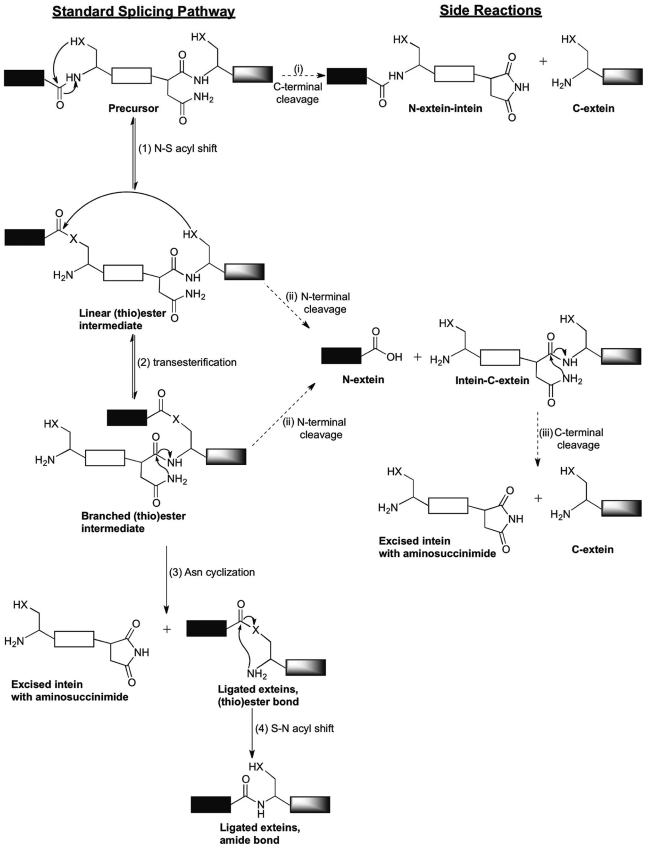
ABSTRACT: We report the first detailed investigation of the kinetics of protein splicing by the *Methanococcus jannaschii* KlbA (*Mja* KlbA) intein. This intein has an N-terminal Ala in place of the nucleophilic Cys or Ser residue that normally initiates splicing but nevertheless splices efficiently in vivo [Southworth, M. W., Benner, J., and Perler, F. B. (2000) *EMBO J. 19*, 5019−5026]. To date, the spontaneous nature of the cis splicing reaction has hindered its examination in vitro. For this reason, we constructed an *Mja* KlbA intein−mini-extein precursor using intein-mediated protein ligation and engineered a disulfide redox switch that permits initiation of the splicing reaction by the addition of a reducing agent such as dithiothreitol (DTT). A fluorescent tag at the C-terminus of the C-extein permits monitoring of the progress of the reaction. Kinetic analysis of the splicing reaction of the wild-type precursor (with no substitutions in known nucleophiles or assisting groups) at various DTT concentrations shows that formation of the branched intermediate from the precursor is reversible (forward rate constant of 1.5×10^{-3} s^{-1} and reverse rate constant of 1.7×10^{-5} s^{-1} at 42 °C), whereas the productive decay of this intermediate to form the ligated exteins is faster and occurs with a rate constant of 2.2×10^{-3} s^{-1}. This finding conflicts with reports about standard inteins, for which Asn cyclization has been assigned as the rate-determining step of the splicing reaction. Despite being the slowest step of the reaction, branched intermediate formation in the *Mja* KlbA intein is efficient in comparison with those of other intein systems. Interestingly, it also appears that this intermediate is protected against thiolysis by DTT, in contrast to other inteins. Evidence is presented in support of a tight coupling between the N-terminal and C-terminal cleavage steps, despite the fact that the C-terminal single-cleavage reaction occurs in variant *Mja* KlbA inteins in the absence of N-terminal cleavage. We posit that the splicing events in the *Mja* KlbA system are tightly coordinated by a network of intra- and interdomain noncovalent interactions, rendering its function particularly sensitive to minor disruptions in the intein or extein environments.



Inteins are intervening sequences that are post-translationally excised from precursor proteins with simultaneous splicing of flanking regions, termed the exteins, to form mature proteins.[1] Standard protein splicing is believed to occur via the mechanism summarized in Scheme 1.[2−4] All standard inteins utilize a Cys, Thr, or Ser residue at position 1 to perform an acyl rearrangement and form a (thio)ester linkage at the N-terminal splice junction in the first step of the reaction (Scheme 1, step 1).[5] Splicing is blocked upon nonconservative substitution of this residue.[5,6] Therefore, it has long been believed that noncanonical inteins, such as *Methanococcus jannaschii* KlbA (*Mja* KlbA) intein, which harbors an Ala at position 1 (Ala1), cannot undergo splicing. Nevertheless, it has been shown that the *Mja* KlbA intein splices efficiently in vivo and does so by an alternative splicing mechanism (Scheme 2).[7] In this mechanism, a nucleophilic attack by the Cys located at the N-terminus of the C-extein (Cys+1) on the peptide bond at the N-terminal splice junction occurs as the first step of the splicing reaction (Scheme 2, step 1).[7] This step results in the formation of a branched intermediate with two N-termini, one of the N-extein and another of the intein. This situation is fundamentally different from what is observed in the standard intein pathway, in which the C-extein nucleophile attacks a previously formed linear (thio)ester intermediate resulting in the formation of the branched intermediate (Scheme 1, step 2).[8] In both pathways, the branched intermediate is resolved during a transamidation step performed by the C-terminal intein residue, Asn, which results in the release of the intein (Scheme 1, step 3; Scheme 2, step 2). A spontaneous S−N or O−N acyl shift, which results in the formation of a peptide

Received: May 25, 2011
Revised: October 25, 2011
Published: October 26, 2011

Scheme 1. Standard Class 1 Intein Splicing Pathway[a]

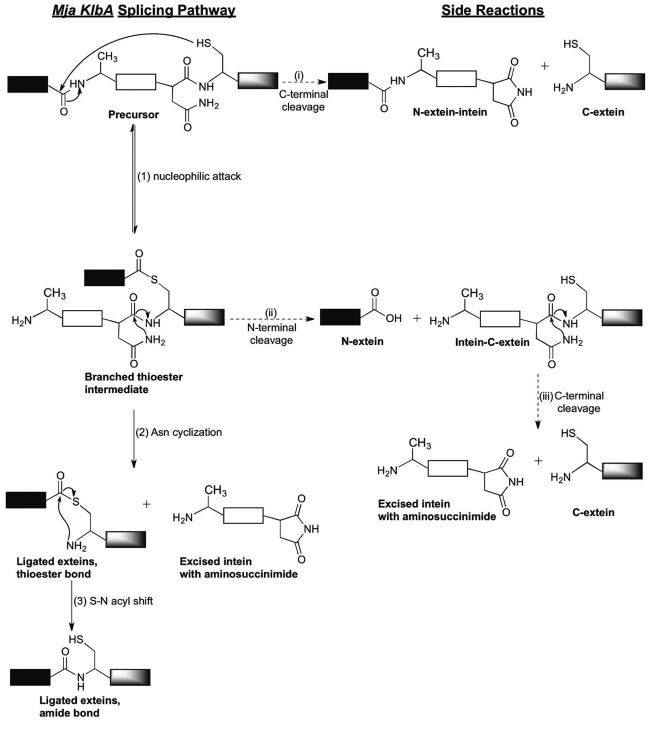


[a](1) N−S or −O acyl shift, (2) transesterification, (3) Asn cyclization, and (4) S− or O−N acyl shift. The succinimide ring in the excised intein might undergo hydrolysis to form Asn or isoAsn. Side reactions expected to occur under certain conditions are (i) and (iii) Asn cyclization resulting in C-terminal cleavage and (ii) attack of the (thio)ester by an exogenous nucleophile resulting in N-terminal cleavage. The filled rectangle represents the N-extein, the empty rectangle the intein, and the shaded rectangle the C-extein. The solid arrows indicate steps in the main pathway, while the dashed arrows indicate side reactions. X corresponds to either -SH or -OH.

bond between the N- and C-exteins, completes the reaction (Scheme 1, step 4; Scheme 2, step 3). More recently, another class of atypical inteins was identified, which splices by a third mechanism involving two branched intermediates.[9] Intein splicing mechanisms are now divided into three classes. Class 1 inteins follow the standard splicing pathway. Class 2 inteins follow the *Mja* KlbA splicing pathway. Class 3 inteins follow the two-branch intermediate splicing pathway.

Side reactions off the main splicing pathway have been detected in all classes of inteins, often as a consequence of improper coordination between various steps of the splicing mechanism resulting from substitution of catalytically important amino acid residues. The side products arise from cleavage at either or both splice junctions without concomitant ligation (Schemes 1 and 2, steps i–iii). Previous studies have shown that substitution of essential catalytic residues at one splice junction usually inhibits splicing and isolates the cleavage side reaction at the other junction. It has also been shown that nucleophiles such as DTT, hydroxylamine, and sodium 2-mercaptoethanesulfonate (MESNA) can intercept the (thio)-ester intermediates (linear and branched), resulting in the

Scheme 2. Class 2 *Mja* KlbA Intein Splicing Pathway[a]



[a](1) Nucleophilic attack by nucleophile at position 1, (2) Asn cyclization, and (3) S−N acyl shift. Side reactions expected to occur are identical to those described in the footnote of Scheme 1. The filled rectangle represents the N-extein, the empty rectangle the intein, and the shaded rectangle the C-extein. The solid arrows indicate steps in the main pathway, while the dashed arrows indicate side reactions.

formation of N-terminal cleavage products (Schemes 1 and 2, step ii).

In this study, we sought to define the kinetic details of the nonstandard *Mja* KlbA intein splicing reaction to improve our understanding of its mechanism. To this end, we designed an intein−mini-extein precursor with a redox switch and synthesized it by intein-mediated protein ligation [IPL, also known as expressed protein ligation (EPL)]. This system allowed us for the first time to dissect the kinetics of cleavage and cis splicing of this novel intein from an active precursor with no substitutions in known nucleophiles or assisting groups. The effects of pH variation and of substitution of catalytic and assisting groups on the kinetics of the overall splicing reaction and individual steps were further examined with this system. On the basis of our results, we suggest an explanation for how a direct attack on a peptide bond at the N-terminal splice site by Cys+1 is facilitated in the *Mja* KlbA intein when it has been blocked in all standard inteins studied to date.

MATERIALS AND METHODS

Materials. Bacto yeast extract and Bacto tryptone were purchased from Becton, Dickinson, and Company (Sparks, MD). Dextrose, magnesium chloride, sodium hydroxide, tris(hydroxymethyl)aminomethane (Tris), sodium chloride, and urea were purchased from J. T. Baker Chemical Co. (Phillipsburg, NJ). Isopropyl β-D-thiogalactopyranoside (IPTG), MESNA, chloramphenicol, *N*-(2-hydroxyethyl)-piperazine-*N*′-2-ethanesulfonic acid (HEPES), and Sephadex-G50 were purchased from Sigma (St. Louis, MO). Ampicillin was purchased from America Pharmaceutical Partners, Inc.

(Schaumburg, IL). Sodium dodecyl sulfate (SDS) was purchased from Bio-Rad (Hercules, CA). Oligonucleotide primers, reagents for the polymerase chain reaction (PCR), restriction enzymes, 10× CP buffer, pTWIN1 and pMAL-c2X vectors, and competent *Escherichia coli* cells (NEB TURBO and T7 Express) were obtained from New England Biolabs (NEB). The pRIL vector was purchased from Stratagene (La Jolla, CA).

Peptides used in this study are summarized in Table 1 and comprise native C-extein sequences. $\mathbf{P_{1NC}}$, $\mathbf{P_{2AC}}$, $\mathbf{P_{3QC}}$, $\mathbf{P_{4NS}}$, and $\mathbf{P_{5NA}}$ were synthesized at NEB. $\mathbf{P_{1NCs}}$, $\mathbf{P_{2ACs}}$, and $\mathbf{P_{6NT}}$ were purchased from AnaSpec (San Jose, CA). Peptides $\mathbf{P_{1NC}}$ and $\mathbf{P_{1NCs}}$, and peptides $\mathbf{P_{2AC}}$ and $\mathbf{P_{2ACs}}$, respectively, are identical in the nature of their functional residues but differ in their lengths (Table 1). The pure forms of the peptides were characterized as the desired products by electrospray ionization time-of-flight mass spectrometry (ESI-TOF) using a 6210 ESI-TOF mass spectrometer with a capillary electrospray ionization source (Agilent Technologies, Santa Clara, CA).

Table 1. Peptides Used in IPL Reactions

peptide	sequence[a]
$\mathbf{P_{1NC}}$	H-**CN**CSGTLHANSADEAILRLTSPPMNVPKIMLT-Lys(5(6)-FAM)-OH
$\mathbf{P_{1NCs}}$	H-**CN**CSGTLHANSA-Lys(5(6)-FAM)-OH
$\mathbf{P_{2AC}}$	H-**CA**CSGTLHANSADEAILRLTSPPMNVPKIMLT- Lys(5(6)-FAM)-OH
$\mathbf{P_{2ACs}}$	H-**CA**CSGTLHANSA-Lys(5(6)-FAM)-OH
$\mathbf{P_{3QC}}$	H-**CQ**CSGTLHANSADEAILRLTSPPMNVPKIMLT-Lys(5(6)-FAM)-OH
$\mathbf{P_{4NS}}$	H-**CN**SSGTLHANSADEAILRLTSPPMNVPKIMLT-Lys(5(6)-FAM)-OH
$\mathbf{P_{5NA}}$	H-**CN**ASGTLHANSADEAILRLTSPPMNVPKIMLT-Lys(5(6)-FAM)-OH
$\mathbf{P_{6NT}}$	H-**CN**TSGTLHANSA-Lys(5(6)-FAM)-OH

[a]Bold residues correspond to the last two residues of the *Mja* KlbA intein. Underlined residues correspond to the first 10 or 30 C-extein residues. Lys(5(6)-FAM) is the abbreviation for Lys[5(6)-carboxyfluorescein].

Construction of the $\mathbf{MI_{166}}$ Expression Vector. The previously described *Mja* KlbA intein precursor, which encodes the native intein (168 residues) flanked by seven native N-extein residues (MNTGHDG) and six native C-extein residues (CSGTLH),[10] was used as the initial PCR template. PCR was conducted using a sense primer (5′-ATAAACTGCCAGGAATTGGGGATCGG-3′), which anneals to a sequence upstream of the *Nde*I restriction site, and an antisense primer (5′-TGTTACGCTCTTCTGCA**AGC**AG-CAAAACCTTCGTTTTTTCC-3′), which truncates the intein after residue 166. The antisense primer also introduces a substitution at codon Val166 [GTC (V) to GCT (A), triplet complement in bold] and a *Sap*I restriction site (underlined). PCR resulted in a 650 bp fragment, which was then digested with *Nde*I and *Sap*I, and ligated into the pTWIN1 vector (NEB) digested with the same enzymes. To introduce a maltose binding protein (MBP) tag at the N-terminus of the precursor preceding the *Mja* KlbA seven-amino acid native N-extein, we first cut the resultant plasmid with *Nde*I to yield a linear vector. Next, a *Hin*dIII site was introduced adjacent to the *Nde*I site by insertion of a double-stranded oligonucleotide cassette (5′-TATGCTAGGCAAGCTTGTA-3′ and 5′-TATACAAGCTTGCCTAGCA-3′). Finally, the resultant plasmid was digested with *Nde*I-*Hin*dIII and ligated to the 1201 bp *Nde*I-*Hin*dIII restriction fragment of pMAL-c2X, carrying the *malE* gene encoding MBP, to yield the desired expression vector, pTWIN1-$\mathbf{MI_{166}}$. $\mathbf{MI_{166}}$ corresponds to the MBP-tagged (N-terminus) *Mja* KlbA intein with seven native N-extein residues and with intein truncated at residue A166.

Variants of $\mathbf{MI_{166}}$. Single-site substitutions were introduced into pTWIN1-$\mathbf{MI_{166}}$ using the Phusion Site-Directed Mutagenesis Kit (NEB). The primers used are listed in Table 2. PCRs were conducted according to the manufacturer's instructions. After amplification, the parental DNA strand was digested with *Dpn*I restriction enzyme, and the digested mixture was used to transform NEB Turbo competent *E. coli* cells.

The sequences of the coding regions of all plasmid constructs prepared in this study were verified to ensure the absence of unwanted mutations. Sequencing was performed at the NEB DNA Sequencing Facility.

Expression and Purification of $\mathbf{MI_{166}}$ and Variants. *E. coli* T7 Express cells were transformed with pTWIN1-$\mathbf{MI_{166}}$ and pRIL plasmids. pRIL encodes cognate tRNAs for rare codons AGG/AGA, AUA, and CUA and confers resistance to chloramphenicol (Stratagene). Transformed cells were grown with vigorous aeration at 37 °C in LB medium containing 10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl, 1 g/L dextrose, 1 g/L $MgCl_2$ (pH 7.2), 0.25 g/L ampicillin, and 0.025 g/L chloramphenicol. Details of the growth protocol have been described previously.[10]

Ligation of $\mathbf{MI_{166}}$ into the *Nde*I and *Sap*I sites of pTWIN1 results in the fusion of the *Mycobacterium xenopi* GyrA (*Mxe* GyrA) intein to the C-terminus of the expressed protein. The chitin-binding domain present on the *Mxe* GyrA intein allows the affinity purification of the expressed protein on a chitin resin. $\mathbf{MI_{166}}$ and variant fusion proteins were purified as instructed in the IMPACT-TWIN manual (NEB). Buffer A [20 mM Tris-HCl (pH 7.5) and 500 mM NaCl] was used during sonication of harvested cells, and buffer B [20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 50 mM MESNA] was used as the elution buffer. Fractions containing the protein of interest, as judged by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE), were pooled and concentrated to ∼10 mg/mL using an ultrafiltration spin column with a YM30 membrane (Vivascience). The concentration of the protein was determined with the Bio-Rad Bradford protein assay kit standardized with BSA. The $\mathbf{MI_{166}}$ protein was flash-frozen in

Table 2. Primers Used for Construction of $\mathbf{MI_{166}}$ Variants[a]

variant	sense primer	antisense primer
T93A	GATTACGCTA**GC**CCACGACC	GGTCGTG**GGC**TAGCGTAATC
D147A	GGACACATATAT**GC**GTTAACAGTTGAAGATAATCACAC	GTGTGATTATCTTCAACTGTTAA**CGC**ATATATGTGTCC
D147E	GGACACATATAT**GAG**CTCACAGTTGAAGATAATCACAC	GTGTGATTATCTTCAACTGTGAG**CTC**ATATATGTGTCC
Y156A	GAAGATAATCACACA**GCA**ATTGCTGGAAAA	TTTTCCAGCAAT**TGC**TGTGTGATTATCTTC

[a]In bold are the codons introducing the desired substitutions. Underlined are the sequences introducing new restriction sites.

liquid N_2 and stored at −80 °C until it was used. SDS−PAGE revealed the protein to be >98% pure. A typical yield was ∼10 mg/g of cell paste.

IPL Reaction. A typical IPL reaction mixture contained, in a final volume of 160 μL, 100 mM Tris-HCl (pH 8.5), 10 mM MESNA, 50−80 μM MI_{166} protein, and 0.5 mM peptide **P**. All of the steps were performed at 4 °C unless stated otherwise. The reaction sample was incubated for ∼15 h for ligation to occur before it was applied to a Sephadex-G50 spin column for removal of MESNA. The column fractions were analyzed on a Novex 10−20% Tris-glycine polyacrylamide gel (Invitrogen, Carlsbad, CA). The concentrations of protein fractions containing the IPL product, MI_{166}-**P** (also known as the precursor), were determined using the Bradford method. These fractions were flash-frozen in liquid N_2 and stored at −80 °C until they were used.

Protein Splicing Assay. A typical protein splicing assay mixture contained, in a final volume of 540 μL, 20 mM Tris-HCl (pH 7.5), 500 mM NaCl, 0.1 μg/μL precursor, and 1, 5, or 25 mM DTT. Protein splicing was initiated by the addition of DTT after incubation of the other components of the assay mixture at 42 °C for 5 min. Aliquots (20 μL) of the assay mixture were removed at designated times, added to blue loading buffer [6% (w/v) SDS] containing 8 M buffered urea[a] [1 M Tris-HCl (pH 8.0)], and placed on dry ice until all time point samples were collected. Some of the samples were then incubated in a boiling water bath for 5 min prior to being loaded on a Novex 10−20% Tris-glycine polyacrylamide gel for analysis, whereas others were loaded directly on the gel, as illustrated in Results. Coomassie Brilliant Blue R250 staining was used to visualize protein bands.

For pH effect studies, 20 mM sodium phosphate (Na_2HPO_4/NaH_2PO_4) buffer (pH 6.0), 20 mM HEPES (pH 6.5 or 8.0), or 20 mM Tris-HCl (pH 8.5 or 9.0) was substituted for 20 mM Tris-HCl (pH 7.5) in the assay described above.

Acquisition of Data and Kinetic Analysis. Fluorescent band signals were detected using a CCD camera (ChemiDoc XRS, Bio-Rad) and quantified using Bio-Rad's Quantity One. Kinetic traces were plotted as $[I_i]$ versus time (seconds), where $[I_i] = I_i/(I_1 + I_2 + I_3 + \ldots I_n)$ and $I_1 = I^\circ_1 - I_b$, $I_2 = I^\circ_2 - I_b$, $I_3 = I^\circ_3 - I_b$, etc. (I° is the measured fluorescence intensity of the species at time t, I_b is the background intensity, and I_i is I_1, I_2, I_3, ..., or I_n).

The ratio of the sum of fluorescent intensities at time t $[\sum(I_1 + I_2 + I_3 + \ldots I_n)_t]$ to the sum of fluorescent intensities at time zero $[\sum(I_1 + I_2 + I_3 + \ldots I_n)_0]$ was determined to be ∼1, which ensures no loss of fluorescence intensity over the time course of the experiment.

Kinetic simulations were performed using KinTekSim (KinTek Corp., Austin, TX) according to the simulation models provided in the corresponding figures.

N-Terminal Sequencing. For N-terminal sequencing, protein bands from 10−20% Tris-glycine polyacrylamide gels were transferred to a PVDF membrane (Problott, Applied Biosystems Inc.) according to the procedure of Matsudaira,[11] with modifications described previously.[12] The membrane was stained with Coomassie Blue R250, and the desired bands were excised from the surrounding membrane and subjected to sequential degradation on a Procise 494 protein/peptide sequencer (Applied Biosystems Inc.).[12]

RESULTS

Rationale for the Construction of the *Mja* KlbA Intein−Mini-Extein Precursor (MI_{166}-P). The in vitro characterization of protein splicing of any intein (cis splicing or trans splicing) requires the assembly of an unspliced and active form of the intein known as the precursor and a mechanism for initiating splicing at the desired time. For example, the cis splicing *Pyrococcus* GB-D DNA polymerase (GB-D pol) intein can be purified in a precursor form when inserted into the MBP-paramyosin ΔSal gene and its splicing initiated by an increase in temperature or pH.[8] Other examples include naturally and artificially split inteins, which are isolated as two or more separate precursor fragments and splice only upon assembly of these fragments. Attempts to identify conditions that would slow the splicing of the wild-type *Mja* KlbA intein in *E. coli* and allow for the isolation of an intact precursor were not successful. As an alternative approach, we resorted to IPL to construct a functional precursor, which would maintain the splicing efficiency exhibited by the wild-type *Mja* KlbA intein in vivo. The crucial design feature is a built-in redox switch consisting of a disulfide linkage between the existant Cys +1 and an engineered neighboring Cys. The approach was inspired by previous studies employing thiol−disulfide redox switches to control other protein activities.[13−16]

The IPL system consists of two protein fragments, the recombinantly expressed N-terminal portion of the precursor with a C-terminal thioester-activated bond and a synthetic peptide fragment bearing a Cys as the first residue. Each synthetic peptide fragment encodes the last residues of the *Mja* KlbA intein, when required, followed by native *Mja* KlbA C-extein sequences. A number of *Mja* KlbA site-directed variants were prepared to determine the best site for peptide ligation (Supporting Information). Two major concerns were the positioning of the engineered Cys in the proximity of Cys+1 and the identity of the C-terminal amino acid of the expressed fragment. Previous studies showed that a C-terminal Val, Thr, and Ile often diminish ligation efficiency, whereas His, Cys, and Gly give the most efficient ligation with an N-terminal Cys peptide.[17] Assessments of ligation and splicing efficiencies showed that the best position for truncation of the *Mja* KlbA intein is Val166, which is substituted with Ala to increase the efficiency of ligation. The peptide has Ser167 as the first residue, which is substituted with Cys to permit ligation of the peptide to the expressed protein fragment and formation of a deactivating disulfide with Cys+1. As a result, the IPL system used in our study consists of a recombinant fragment, MI_{166}, and the chemically synthesized peptide, **P**. MI_{166} comprises MBP,[b] seven native N-extein residues (MNTGHDG), and the first 166 intein residues with Ala166 as the last residue. **P** contains the two most C-terminal intein residues (S167C and N168), 10 or 30 native C-extein residues (depending on the **P** used), and a fluorescein tag (FAM) (Figure 1). The V166A substitution and the S167C substitution were tested independently in a standard precursor and shown not to inhibit splicing in vivo (Tables S1 and S2 of the Supporting Information).

Assembly of MI166-P_{1NCs} Using IPL. MI_{166} expressed in *E. coli*, as detailed in Materials and Methods, was readily purified with a chitin affinity column using MESNA as the pTWIN vector intein−thioester cleavage reagent. The MI_{166}-MESNA analogue (Figure 1) was incubated with an 8−10-fold molar excess of P_{1NCs} for ∼15 h at 4 °C. Under these

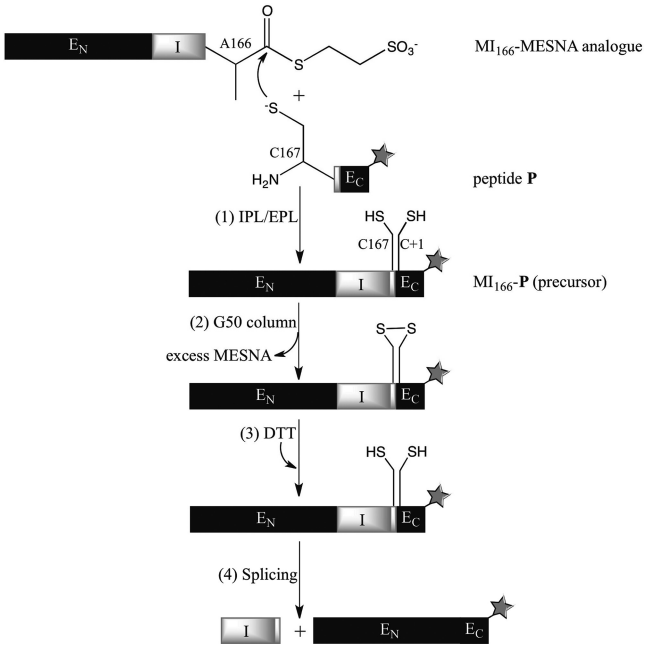


Figure 1. Disulfide redox switch that controls the triggering of the splicing reaction of MI_{166}-**P**. Expressed MI_{166} eluted off the chitin resin with MESNA, resulting in MI_{166}-MESNA. (1) This analogue is reacted with synthesized **P** in the IPL reaction at pH 8.5, yielding MI_{166}-**P**. (2) The IPL product sample is then applied to G50 resin to remove excess MESNA and form oxidized MI_{166}-**P**. (3) Addition of DTT reduces the disulfide bond of MI_{166}-**P** and initiates (4) the splicing reaction. E_N corresponds to N-extein (seven native residues) with an N-terminal MBP tag; E_C corresponds to C-extein (10 or 30 native residues), and the star corresponds to the Lys(5(6)-FAM) tag.

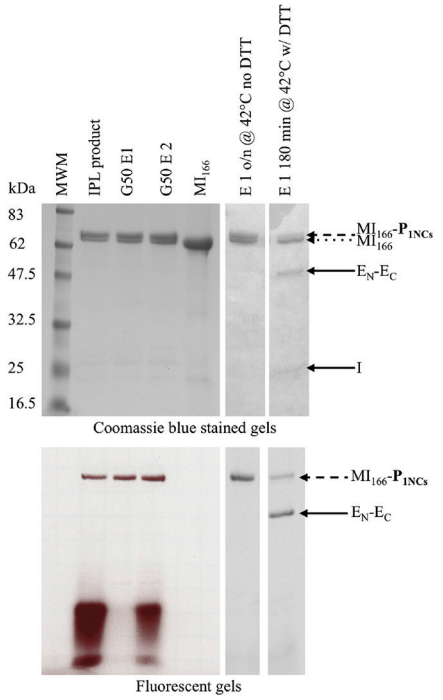


Figure 2. SDS–PAGE gel of the IPL reaction between MI_{166} and $\mathbf{P_{1NCs}}$ showing the formation of the precursor, MI_{166}-$\mathbf{P_{1NCs}}$, and demonstrating the stability of the formed precursor in the absence of DTT. From left to right: lane 1, broad range molecular mass marker; lane 2, products of the IPL reaction between MI_{166} (30 μM) and $\mathbf{P_{1NCs}}$ (0.6 mM) after incubation for 15 h at 4 °C; lane 3, IPL product desalted on a G50 column [elution fraction 1 (E1)]; lane 4, G50 elution fraction 2 (E2); lane 5, unreacted MI_{166}; lane 6, E1 after incubation for 24 h at 42 °C; lane 7, E1 after incubation for 180 min at 42 °C in the presence of 1 mM DTT. The top panel is a Coomassie blue-stained gel, and the bottom panel is the same gel visualized with a Bio-Rad CCD camera (ChemiDoc XRS). The fluorescent smear running below the 16.5 kDa marker is unreacted fluorescent peptide, $\mathbf{P1_{NCs}}$.

conditions, a strongly fluorescent major new band at an apparent molecular mass of 65.8 kDa was detected, confirming the formation of the ligated product (the splicing precursor) (Figure 2). The ligation reaction does not go to completion: not all of MI_{166} (64.3 kDa) is converted to MI_{166}-$\mathbf{P_{1NCs}}$, as seen in Figure 2. However, neither the splicing products [ligated exteins (E_N-E_C^c) (46.4 kDa) and free intein (I) (19.4 kDa)] nor the cleavage side products [MBP-N-extein (E_N^c) (44.9 kDa) and intein-C-extein (I-E_C) (20.9 kDa)] form at detectable levels, indicating that the conditions under which the ligation reaction is performed are optimal. Upon removal of excess MESNA by chromatography on G50 resin, the desired disulfide of MI_{166}-$\mathbf{P_{1NCs}}$ apparently forms (Figure 1), as the precursor is then stable until the initiating reducing agent, DTT, is added, after which the splicing reaction occurs (Figure 2). Formation of the splicing products, E_N-E_C and I, is confirmed with a denaturing gel (Figure 2, lane 7) and ESI-TOF MS (Figure S4 of the Supporting Information).

Reactions of MI_{166}-P_{1NC} and the MI166-P_{2AC} Variant. Initiation of the MI_{166}-$\mathbf{P_{1NC}}$ splicing reaction with 1 mM DTT at 42 °C results in the appearance of four new species over the course of 3 h (Figure 3A): E_N-E_C (47.9 kDa), E_N (44.9 kDa), I-E_C (23.1 kDa), and I (19.4 kDa). Only two of these species are fluorescent, E_N-E_C and I-E_C (Figure 3B). The assignment of these species is done on the basis of: (i) their estimated molecular masses determined in comparison to a broad range molecular mass marker, (ii) the products that are expected to form during the reaction of MI_{166}-$\mathbf{P_{1NC}}$, and (iii) the fluorescent properties of the species (only those species with intact E_C should fluoresce). Kinetic traces for this reaction exhibit an exponential decay for MI_{166}-$\mathbf{P_{1NC}}$ [Figure 3C (●)], an exponential rise for E_N-E_C [Figure 3C (■)], and transient behavior for I-E_C [Figure 3C (▲)].

Mixing of MI_{166}-$\mathbf{P_{2AC}}$, a variant precursor disabled in the C-terminal cleavage reaction by the substitution of N168 with A, with 1 mM DTT at 42 °C results in the formation of two new species (Figure 4A,B). These species are assigned as E_N (44.9 kDa) and I-E_C (23.1 kDa), products of the N-terminal cleavage side reaction. Species I-E_C appears to be stable [Figure 4C (■)].

Trapping of the Branched Intermediate. We were puzzled by the transient nature of I-E_C in the reaction of MI_{166}-$\mathbf{P_{1NC}}$ (Figure 3C). A possible interpretation is that the decay of I-E_C is a result of Asn cyclization, which would lead to the release of E_C and formation of free I. The detection of stable I-E_C in the reaction of MI_{166}-$\mathbf{P_{2AC}}$ strengthens this interpretation (Figure 4C). However, our attempts to simulate the kinetic traces obtained in the reaction of MI_{166}-$\mathbf{P_{1NC}}$ with such a model were unsuccessful. It was necessary to invoke the accumulation of an intermediate, which decays via two constituent pathways to form E_N-E_C and I in one pathway and I-E_C and E_N in the other (Figure 6B; discussed in the next section), to reproduce the experimental data. We suggest that this intermediate is the branched thioester intermediate, as discussed below.

dx.doi.org/10.1021/bi200810j | *Biochemistry* 2011, 50, 10576–10589

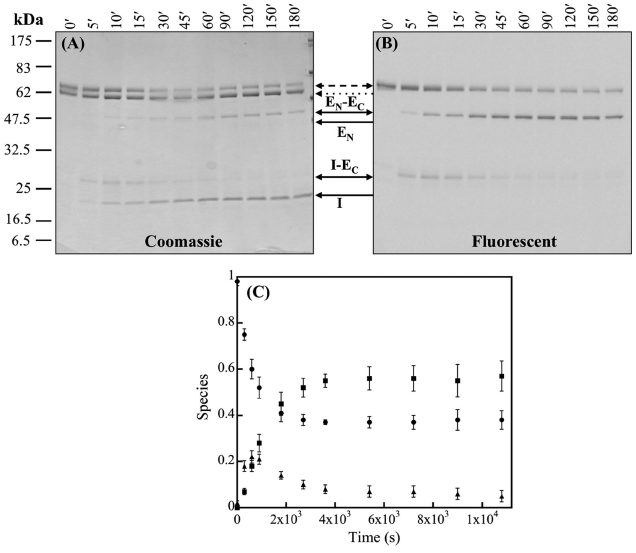


Figure 3. Time course reaction of MI_{166}-$\mathbf{P_{1NC}}$ in the presence of 1 mM DTT at 42 °C. (A) The formation of four new species, E_N-E_C (47.9 kDa), E_N (44.9 kDa), I-E_C (23.1 kDa), and I (19.4 kDa), is detected over the course of 3 h on a Coomassie blue-stained 10−20% Tris-glycine polyacrylamide gel. The molecular mass markers (kilodaltons) indicate the positions of the protein standards of the broad range prestained protein marker (NEB) on the gel. (B) E_N-E_C and I-E_C are shown to be fluorescent. The dashed and dotted arrows indicate the precursor MI_{166}-$\mathbf{P_{1NC}}$ (68.1 kDa) and unreacted MI_{166} (64.3 kDa), respectively. Data for the kinetic decay of MI_{166}-$\mathbf{P_{1NC}}$ and formation of E_N-E_C and I-E_C are shown in panel C. The circles correspond to data for MI_{166}-$\mathbf{P_{1NC}}$, the squares to data for E_N-E_C, and the triangles to data for I-E_C. The details of the protein splicing assay and quantitation of the various species are as described in Materials and Methods. Samples were quenched in blue loading buffer [6% (w/v) SDS] containing 8 M buffered urea, flash-frozen on dry ice, and then boiled for 5 min before being loaded on the gel. MI_{166}-$\mathbf{P_{1NC}}$ was used at a final concentration of 0.15 μg/μL. Error bars represent the standard error of the mean.

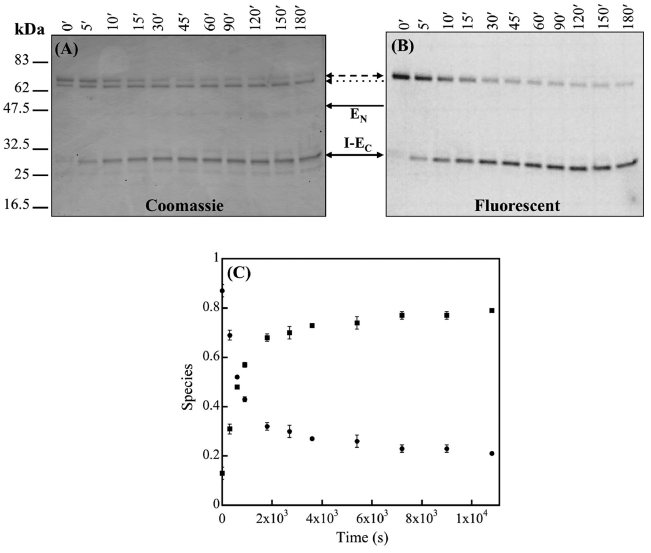


Figure 4. Time course reaction of MI_{166}-$\mathbf{P_{2AC}}$ in the presence of 1 mM DTT at 42 °C. (A) Only two products, E_N (44.9 kDa) and I-E_C (23.1 kDa), are formed over the course of 3 h, as observed on a Coomassie blue-stained 10−20% Tris-glycine polyacrylamide gel. The molecular mass markers (kilodaltons) indicate the positions of the protein standards of the broad range prestained protein marker (NEB) on the gel. (B) Two species, MI_{166}-$\mathbf{P_{2AC}}$ and I-E_C, are observed on a fluorescent gel. The dashed and dotted arrows correspond to the precursor, MI_{166}-$\mathbf{P_{2AC}}$ (68.0 kDa), and unreacted MI_{166} (64.3 kDa), respectively. Data for the kinetic decay of MI_{166}-$\mathbf{P_{2AC}}$ and formation of I-E_C are shown in panel C. The circles correspond to data for MI_{166}-$\mathbf{P_{2AC}}$ and the squares to data for I-E_C. The details of the protein splicing assay and quantitation of the various species are as described in Materials and Methods. Samples were quenched in blue loading buffer [6% (w/v) SDS] containing 8 M buffered urea, flash-frozen on dry ice, and then boiled for 5 min before being loaded on the gel. MI_{166}-$\mathbf{P_{2AC}}$ was used at a final concentration of 0.1 μg/μL. Error bars represent the standard error of the mean.

Studies of the in vivo splicing reaction of the *Mja* KlbA N168A intein variant suggest that the branched intermediate does accumulate. In an attempt to capture the branched intermediate in vitro, we omitted the heat quench step from the analysis of the MI_{166}-$\mathbf{P_{2AC}}$ reaction, anticipating that heating might result in the hydrolysis of the trapped branched thioester intermediate.[18] Interestingly, a new band at an apparent molecular mass comparable to that of the precursor accumulates, whereas species I-E_C is hardly detected under the new quenching conditions (Figure 5). Sequential Edman degradation of the new species detected the presence of two residues in each cycle: one residue corresponds to the N-terminus of MBP, and the other is equivalent to that of the intein. This result, in addition to the fact that this species is fluorescent, leads us to conclude that the species formed when heat quenching is omitted is the branched intermediate. On the other hand, sequential Edman degradation of the protein in the band assigned as I-E_C revealed that the first 10 residues do indeed correspond to the sequence of the intein consistent with this assignment.

In the MI_{166}-$\mathbf{P_{1NC}}$ reaction, the species observed on a denaturing gel are equivalent, regardless of the quenching procedure (data not shown). The branched intermediate is not observed, by contrast to the MI_{166}-$\mathbf{P_{2AC}}$ reaction. However, kinetic simulations of the MI_{166}-$\mathbf{P_{1NC}}$ time course, as discussed in detail in the next section, point to the accumulation of the

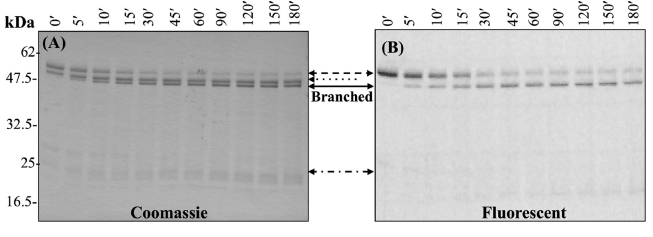


Figure 5. Time points from the reaction of MI_{166}-$\mathbf{P_{2AC}}$ in the presence of 1 mM DTT at 42 °C, which were quenched in blue loading buffer, containing 8 M buffered urea, and were loaded on the gel without prior boiling, demonstrate the accumulation of the branched intermediate on (A) a Coomassie blue-stained gel and (B) a fluorescent gel. The dashed and dotted arrows correspond to the precursor, MI_{166}-$\mathbf{P_{2AC}}$ (68.0 kDa), and unreacted MI_{166} (64.3 kDa), respectively. The double-headed arrow with both dashed and dotted symbols points to the supposed position of I-E_C. The details of the protein splicing assay and quantitation of the various species are as described in Materials and Methods. MI_{166}-$\mathbf{P_{2AC}}$ was used at a final concentration of 0.1 μg/μL.

branched intermediate during the reaction. Therefore, we conclude that the branched intermediate in the MI_{166}-$\mathbf{P_{1NC}}$ reaction decomposes upon quenching, hydrolyzing to form I-E_C and E_N.

Kinetic Simulations of the MI_{166}-P_{1NC} and MI_{166}-P_{2AC} Reactions. Simulation of the kinetics of the MI_{166}-$\mathbf{P_{1NC}}$

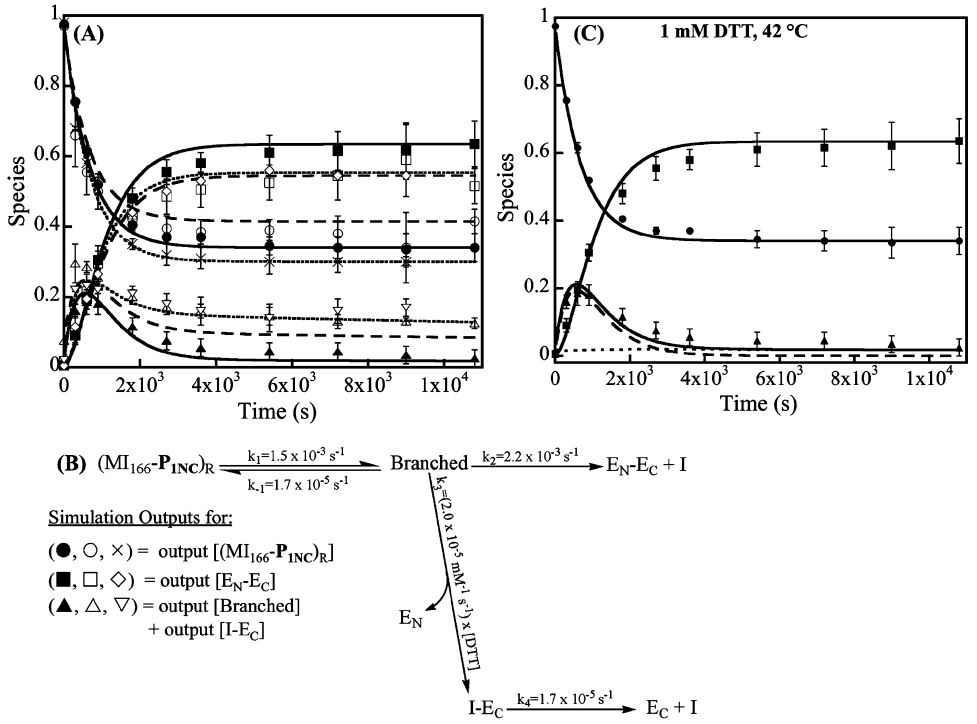


Figure 6. Kinetic simulations of data (symbols) obtained in the reaction MI_{166}-$\mathbf{P_{1NC}}$ at 42 °C in the presence of (A) 1 (—), 5 (---), and 25 mM (···) DTT, according to the kinetic mechanism described in panel B. Symbols represent the experimental data for the precursor MI_{166}-$\mathbf{P_{1NC}}$ (●, ○, and ×), E_N (■, □, and ◇), and I-E_C (▲, △, and ▽). Experimental details are identical to those described in the legend of Figure 3. (B) Kinetic simulation model used to fit the experimental data. The percentage values used in the simulation model for $(MI_{166}$-$\mathbf{P_{1NC}})_R$ are 64, 56, and 68% in the presence of 1, 5, and 25 mM DTT, respectively. (C) Simulation output for the transient species, represented as a solid line fitting the triangles, which is the sum of the simulation outputs for Branched (---) and I-E_C (···). Error bars represent the standard error of the mean.

reaction at varying DTT concentrations (1, 5, and 25 mM) was used to uncover the mechanistic details of the reaction (Figure 6A). The mechanism for which simulated traces best agreed with the experimental data is summarized in Figure 6B. The precursor is a heterogeneous mixture of two forms: unreactive, $(MI_{166}$-$\mathbf{P_{1NC}})_U$, and reactive, $(MI_{166}$-$\mathbf{P_{1NC}})_R$. Formation of the branched intermediate (Branched) from the reactive precursor is reversible with a forward rate constant (k_1) of 1.5×10^{-3} s^{-1} and a reverse rate constant (k_{-1}) of 1.7×10^{-5} s^{-1}. It decays via two pathways. The first is productive and results in the formation of the spliced products, E_N-E_C and I. This pathway is relatively fast, with an apparent first-order rate constant (k_2) of 2.2×10^{-3} s^{-1}; the second pathway is the N-terminal cleavage side reaction. It has a pseudo first-order rate constant (k_3) of $(2.0 \times 10^{-5}$ mM^{-1} $s^{-1}) \times$ [DTT] and is slower than the productive pathway at the DTT concentration used in our experiments. The N-terminal cleavage product, I-E_C, further undergoes Asn cyclization, yielding E_C and I with a k_4 of 1.7×10^{-5} s^{-1}. The simulated trace for the transient species is the sum of the concentrations of Branched and IE_C (Figure 6C, triangles, solid line), consistent with our deduction that Branched hydrolyzes upon quenching. The transient behavior is attributable to the accumulation of Branched (Figure 6C, dashed line) rather than IE_C (Figure 6C, dotted line).

Kinetic simulations of the MI_{166}-$\mathbf{P_{2AC}}$ reaction, performed under the same conditions as for the reaction of MI_{166}-$\mathbf{P_{1NC}}$, reveal interesting differences between the two mechanisms (Figure 7). First, precursor structural isomers, $(MI_{166}$-$\mathbf{P_{2AC}})_I$ and $(MI_{166}$-$\mathbf{P_{2AC}})_{II}$, are both reactive, with $(MI_{166}$-$\mathbf{P_{2AC}})_I$ decaying to form $(MI_{166}$-$\mathbf{P_{2AC}})_{II}$ with a k_0 of 1.7×10^{-5} s^{-1} (Figure 7B); second, Branched decays reversibly into another intermediate species with a reverse rate constant ($k_{-2} = 3.3 \times 10^{-2}$ s^{-1}) that is considerably greater than the forward rate constant [$k_2 = (6.7 \times 10^{-4}$ mM^{-1} $s^{-1}) \times$ [DTT]] at the DTT concentrations interrogated (Figure 7B). The nature of this species is unclear, but it is proposed to be an intermediate formed between MI_{166}-$\mathbf{P_{2AC}}$ Branched and DTT, a species predicted not to form in the reaction of MI_{166}-$\mathbf{P_{1NC}}$. We designate this intermediate (Branched·DTT) (Figure S1 of the Supporting Information). (Branched·DTT) undergoes N-terminal cleavage at a k_3 of 1.7×10^{-4} s^{-1}, which is 10-fold greater than the rate constant for N-terminal cleavage in the MI_{166}-$\mathbf{P_{1NC}}$ reaction at 1 mM DTT. As expected for this variant, no C-terminal cleavage products are detected.

Interestingly, MI_{166}-$\mathbf{P_{1NC}}$ and MI_{166}-$\mathbf{P_{2AC}}$ Branched are both less reactive to thiolysis by DTT, as evident from the extracted rate constants (Figures 6B and 7B), than has been previously observed for standard inteins such as the *Sce* VMA intein,[19] *Ssp* DnaE intein,[20] and *Pab* PolII intein.[21,22] This explains the productive formation of spliced product in the former and the stability of Branched in the latter, despite the presence of relatively high concentrations of DTT, which have been shown to intercept the (thio)ester intermediates in the splicing pathways of other inteins.[19−21,23]

Effect of Peptide Length on the Kinetics of MI_{166}-$\mathbf{P_{1NC}}$ and MI_{166}-$\mathbf{P_{2AC}}$ Reactions. The lengths of peptides $\mathbf{P_{1NC}}$ and $\mathbf{P_{2AC}}$ were reduced from 32 amino acids to 12 amino acids ($\mathbf{P_{1NCs}}$ and $\mathbf{P_{2ACs}}$) (Table 1) to study the effect of peptide length on the assigned rate constants of the various steps of the MI_{166}-$\mathbf{P_{1NC}}$ and MI_{166}-$\mathbf{P_{2AC}}$ reactions. The results obtained indicate that precursors with a longer E_C exhibit higher K_{eq} values in the equilibrium step between the reactive precursor

dx.doi.org/10.1021/bi200810j | *Biochemistry* 2011, 50, 10576−10589

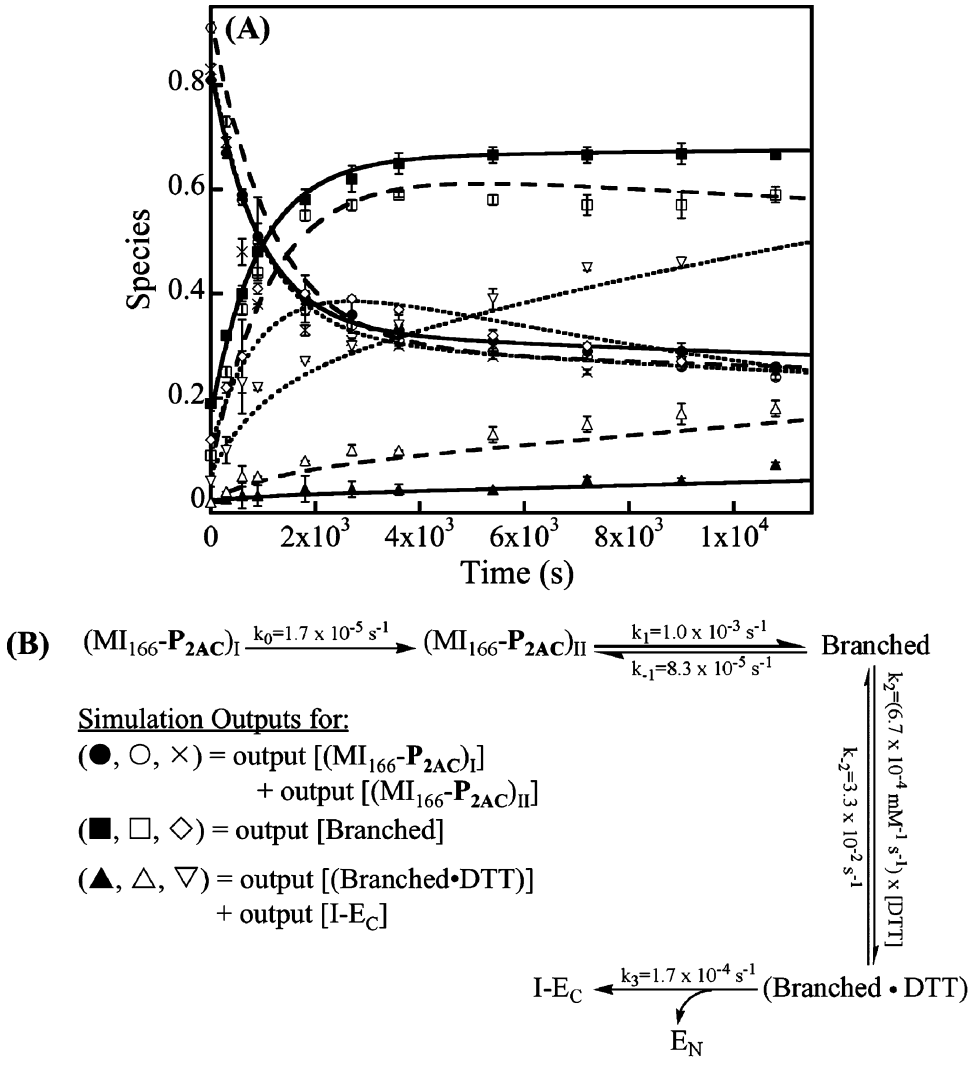


Figure 7. (A) Kinetic simulations of the reaction of MI_{166}-$\mathbf{P_{2AC}}$ in the presence of DTT at 42 °C and (B) kinetic simulation model used to fit the experimental data. Experimental details are identical to those described in the legend of Figure 4. Simulations of the reaction of MI_{166}-$\mathbf{P_{2AC}}$ in the presence of 1, 5, and 25 mM DTT at 42 °C are represented as solid, dashed, and dotted, respectively. Symbols represent the experimental data for the precursor MI_{166}-$\mathbf{P_{2AC}}$ (●, ○, and ×), Branched (■, □, and ◇), and I-E_C (▲, △, and ▽). The percentage values used in the simulation model for $(MI_{166}$-$\mathbf{P_{2AC}})_I$ are 27, 24, and 27% and for $(MI_{166}$-$\mathbf{P_{2AC}})_{II}$ are 56, 69, and 57% in presence of 1, 5, and 25 mM DTT, respectively. Error bars represent the standard error of the mean.

and Branched (Figure S2 of the Supporting Information). It was not necessary to vary other rate constants to accommodate the data for the reactions of the precursors formed with either $\mathbf{P_{1NC}}$ or $\mathbf{P_{2AC}}$.

Kinetic Analysis of the Reaction of MI_{166}-$\mathbf{P_{3QC}}$. Replacement of the N168 residue with Q prevents C-terminal cleavage in vivo. In vitro, reaction of MI_{166}-$\mathbf{P_{3QC}}$ in the presence of 1 mM DTT at 42 °C results in the formation of three new species (Figure 8). The first species starts with the first 10 amino acids of MBP, as shown by N-terminal sequencing, and is not fluorescent. It further runs at the same position as E_N, observed in the reactions of MI_{166}-$\mathbf{P_{1NC}}$ and MI_{166}-$\mathbf{P_{2AC}}$, and is therefore identified as E_N (44.9 kDa), the product of the N-terminal cleavage side reaction. Each of the two other species has the N-terminus of the intein and the C-terminus of E_C (as designated by their fluorescence properties), which indicates that both are I-E_C products (I-E_C I and I-E_C II). It is unclear why these species exhibit distinct mobilities on a denaturing gel.

Kinetic simulations reveal that the MI_{166}-$\mathbf{P_{3QC}}$ precursor is heterogeneous in nature, with two different conformations that are stable and undergo their own decay pathways (Figure 9). In pathway I, $(MI_{166}$-$\mathbf{P_{3QC}})_I$ decays reversibly to form a branched intermediate (Branched I), which is then subject to rapid N-terminal cleavage by DTT with a k_2 of $(1.0 \times 10^{-2}$ mM^{-1} $s^{-1})$ × [DTT] to form I-E_C I. In pathway II, a branched intermediate (Branched II) is formed reversibly from $(MI_{166}$-$\mathbf{P_{3QC}})_{II}$, with a K_{eq} that is slightly larger than that observed for the formation of Branched I. Branched II is then subject to thiolysis to form (Branched II·DTT) reversibly [k_2 = $(1.3 \times 10^{-3}$ mM^{-1} $s^{-1}) \times$ [DTT], and $k_{-2} = 1.2 \times 10^{-3}$ s^{-1}]. In the last step, (Branched II·DTT) undergoes N-terminal cleavage to form I-E_C II with a k_3 of 1.7×10^{-4} s^{-1}. According to the simulations, I-E_C I and I-E_C II form in parallel. This deduction offers an explanation for the detection of two different forms of I-E_C on a denaturing gel.

Intriguingly, the intermediates in pathway I are analogous of those in the MI_{166}-$\mathbf{P_{1NC}}$ reaction, whereas the intermediates in pathway II are analogous of those in the MI_{166}-$\mathbf{P_{2AC}}$ reaction. However, the kinetic parameters indicate that Branched I and Branched II are more reactive toward DTT in comparison with

dx.doi.org/10.1021/bi200810j | Biochemistry 2011, 50, 10576–10589

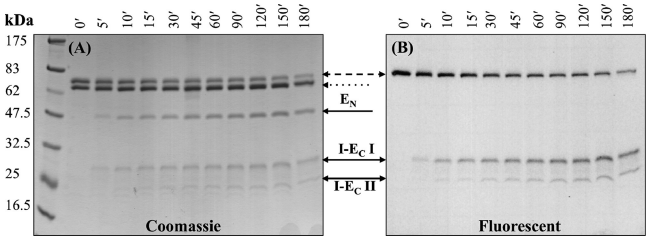


Figure 8. Time course reaction of MI_{166}-$\mathbf{P_{3QC}}$ in the presence of 1 mM DTT at 42 °C. (A) The species formed, over the course of 3 h, are E_N (44.9 kDa), I-E_C I (~29.0 kDa), and I-E_C II (23.0 kDa), as observed on a Coomassie blue-stained 10−20% Tris-glycine polyacrylamide gel. (B) Species MI_{166}-$\mathbf{P_{3QC}}$, I-E_C I, and I-E_C II are observed on a fluorescent gel. The molecular mass markers (kilodaltons) indicate the positions of the protein standards of the broad range prestained protein marker (NEB) on the gel. The dashed and dotted arrows correspond to the precursor, MI_{166}-$\mathbf{P_{3QC}}$ (68.0 kDa), and unreacted MI_{166} (64.3 kDa), respectively. The details of the protein splicing assay and quantitation of the various species are as described in Materials and Methods. Samples were quenched in blue loading buffer, containing 8 M buffered urea, and were loaded on the gel without prior boiling. MI_{166}-$\mathbf{P_{3QC}}$ was used at a final concentration of 0.1 μg/μL.

the cognate intermediates in the MI_{166}-$\mathbf{P_{1NC}}$ and MI_{166}-$\mathbf{P_{2AC}}$ reactions. This increased reactivity with DTT can explain why more I-E_C is formed in the MI_{166}-$\mathbf{P_{3QC}}$ reaction than in the MI_{166}-$\mathbf{P_{2AC}}$ reaction at matching DTT concentrations.

Splicing Reactions of Thr93, Tyr156, Asp147, and Cys +1 Variants. Variants with substitutions of residues implicated in the splicing reaction of the *Mja* KlbA intein have been examined in this study. The Cys+1 variants, MI_{166}-$\mathbf{P_{4NS}}$, MI_{166}-$\mathbf{P_{5NA}}$, and MI_{166}-$\mathbf{P_{6NT}}$, which are expected to inhibit N-terminal cleavage and result in C-terminal cleavage products,[7] showed no evidence of the formation of C-terminal cleavage products in vitro (Figure S3 of the Supporting Information). Only ~4% of the precursor decays after 20 h at 42 °C to yield what might be the products of the slow intrinsic hydrolysis reaction of the scissile peptide (the bands' identities were confirmed by N-terminal sequencing as E_N and I-E_C). Other variants examined include MI_{166}/T93A-$\mathbf{P_{1NCs}}$, MI_{166}/Y156A-$\mathbf{P_{1NCs}}$, and MI_{166}/D147A(E)-$\mathbf{P_{1NCs}}$. In vitro, the reaction of MI_{166}/T93A-$\mathbf{P_{1NCs}}$ in the presence of 1 mM DTT at 42 °C showed the formation of the splicing products but at a very diminished rate. N-terminal cleavage products were also detected in the reaction (Figure 10A,B). This is consistent with the role of block B Thr93, which is a member of the catalytic triad (Thr93, His96, and Cys +1) thought to facilitate reactions at the N-terminal splice junction.[10] It is also consistent with in vivo reports in which replacement of Thr93 with Ala is shown to slow the splicing reaction of *Mja* KlbA intein without disrupting it.[7] Kinetic simulation of this reaction at 42 °C shows that only 33% of the precursor is reactive. The rate constants for both N- and C-terminal catalytic cleavage events are much lower than those observed for MI_{166}-$\mathbf{P_{1NCs}}$. Rate constants of the side reactions remain the same (Figure 11A). Another variant tested was MI_{166}/Y156A-$\mathbf{P_{1NCs}}$. Residue Tyr156 of block F is highly conserved among inteins as either Phe or Tyr. The NMR structure of the *Mja* KlbA intein shows that Tyr156 is centrally located in the active site.[10] Mutation studies suggest that the role performed by Tyr156 is structural and not catalytic.[10] In vivo, replacement of Tyr156 with Phe does not affect the splicing efficiency of the intein, while replacement with Ala results in equal amounts of precursor, N-terminal single-cleavage products, and free intein.[10] Interestingly, no C-terminal cleavage is observed in the reaction of MI_{166}/Y156A-$\mathbf{P_{1NCs}}$ in the presence of 1 mM DTT at 42 °C in vitro. The only

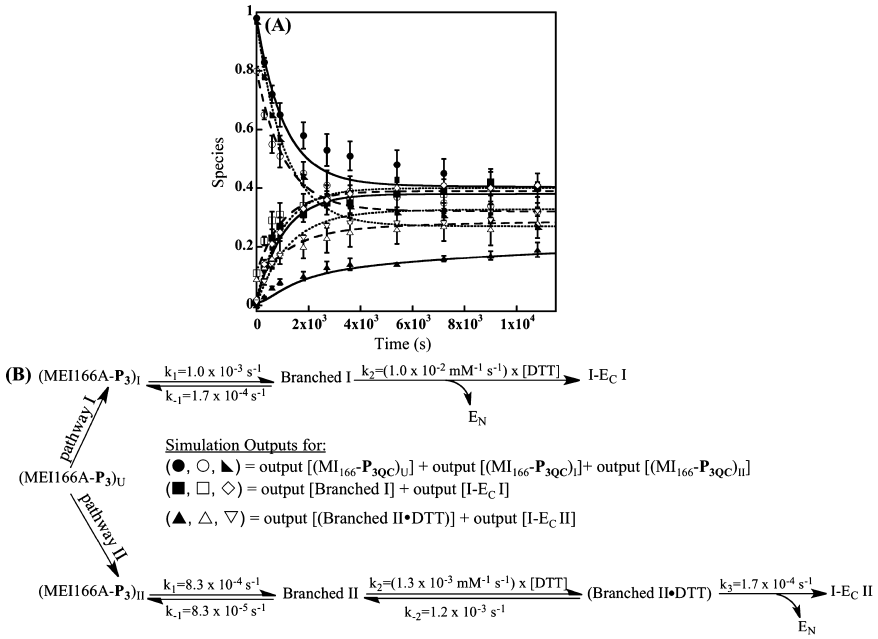


Figure 9. (A) Kinetic simulations of the reaction of MI_{166}-$\mathbf{P_{3QC}}$ in the presence of DTT at 42 °C and (B) kinetic simulation model used to fit the experimental data. Experimental details are identical to those described in the legend of Figure 8. Solid, dashed, and dotted lines correspond to the simulation results of the experimental kinetic data for the reaction of MI_{166}-$\mathbf{P_{3QC}}$ in the presence of 1, 5, and 25 mM DTT, respectively, at 42 °C. Symbols represent the experimental data for the precursor MI_{166}-$\mathbf{P_{3QC}}$ (●, ○, and ×), I-E_C I (■, □, and ◇), and I-E_C II (▲, △, and ▽). The percentage values used in the simulation model for $(MI_{166}$-$\mathbf{P_{3QC}})_I$ are 37, 28, and 38% and for $(MI_{166}$-$\mathbf{P_{3QC}})_{II}$ are 21, 20, and 32% in the reaction in the presence of 1, 5, and 25 mM DTT, respectively. Error bars represent the standard error of the mean.

dx.doi.org/10.1021/bi200810j | *Biochemistry* 2011, 50, 10576−10589

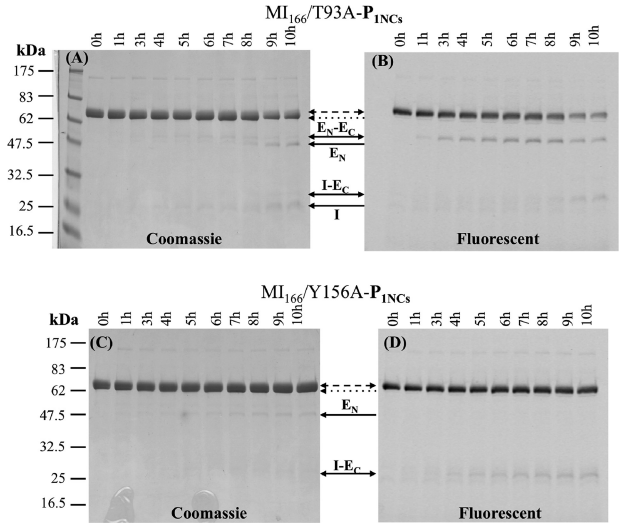


Figure 10. Time course reactions of MI_{166}/T93A-$\mathbf{P}_{1NCs}$ (A and B) and MI_{166}/Y156A-$\mathbf{P}_{1NCs}$ (C and D) in the presence of 1 mM DTT at 42 °C. (A) Formation of E_N-E_C (46.4 kDa), E_N (44.9 kDa), I-E_C (20.9 kDa), and I (19.4 kDa) detected over the course of 10 h on a Coomassie blue-stained 10−20% Tris-glycine polyacrylamide gel. (B) MI_{166}/T93A-$\mathbf{P}_{1NCs}$, E_N-E_C, and I-E_C are fluorescent. (C) Only two products, E_N (44.9 kDa) and I-E_C (20.9 kDa), are formed over the course of 9 h, as observed on a Coomassie blue-stained 10−20% Tris-glycine polyacrylamide gel. (D) MI_{166}/Y156A-$\mathbf{P}_{1NCs}$ and I-E_C are fluorescent. The dashed arrows correspond to the precursors, MI_{166}/T93A-$\mathbf{P}_{1NCs}$ (65.8 kDa) and MI_{166}/Y156A-$\mathbf{P}_{1NCs}$ (65.7 kDa). The dotted arrow corresponds to unreacted MI_{166} (64.3 kDa). The molecular mass markers (kilodaltons) indicate the positions of the protein standards of the broad range prestained protein marker (NEB) on the gel. MI_{166}/T93A-$\mathbf{P}_{1NCs}$ and MI_{166}/Y156A-$\mathbf{P}_{1NCs}$ were used at final concentrations of 0.1 and 0.08 μg/μL, respectively.

products detected are those of N-terminal cleavage by DTT (Figure 10C,D). Simulation provides evidence that this substitution renders the protein mostly inactive (only 10% of the Y156A precursor is reactive) and results in the slow production of I-E_C (Figure 11B). Last but not least, substitution of block F Asp147 with Glu or Ala [MI_{166}/D147E(A)-$\mathbf{P}_{1NCs}$] was examined in vitro in the presence of 1 mM DTT at 42 °C. Asp147 is believed to play a role in Branched formation as well as Asn cyclization/C-terminal cleavage.[10] In vivo, replacement of Asp147 with Glu results in predominant formation of the C-terminal cleavage product, while its replacement with Ala inhibits both N- and C-terminal cleavage reactions.[10] In this study, both of these substitutions inhibit N- and C-terminal cleavage events in the reaction of MI_{166}/D147E(A)-$\mathbf{P}_{1NCs}$ (data not shown).

Despite the fact that evidence of C-terminal single-cleavage reactions, in the absence of N-terminal cleavage, has been observed for mutant *Mja* KlbA inteins in vivo, results from our in vitro studies of T93, D147, and C+1 variants support tight coupling between the C-terminal and N-terminal cleavage steps in the *Mja* KlbA intein−mini-extein precursor.

Effect of pH on the Kinetics of the MI_{166}-$\mathbf{P}_{1NC}$ Reaction. As an additional probe of the splicing mechanisms, the pH was varied in the reaction of MI_{166}-$\mathbf{P}_{1NCs}$ in the presence of 1 mM DTT at 42 °C. The effects on the rate constants of the various steps of the MI_{166}-$\mathbf{P}_{1NCs}$ reaction are summarized in Table 3. The results indicate that physiological pH (7.0−7.5) gives the most efficient formation of the spliced product (maximal values for $K_{eq} = k_1/k_{-1}$ and k_2 are obtained in this range). A bell-shaped effect of pH on k_1 is also observed. This profile most likely reflects general acid−base catalysis during the generation of the branched intermediate. High pH disfavors Branched in its equilibrium with the precursor. This observation can be explained by the greater stability of amides relative to (thio)esters at elevated pH.[24,25] k_2 decreases at pH >7.5, implying a decreased efficiency of Asn cyclization, possibly because of inhibition of the protonation of the amide N of the scissile bond during succinimide formation or deprotonation or misalignment of a yet to be identified residue that assists Asn cyclization. Similar observations for the decreased efficiency of splicing at high pH have been reported in various other inteins, such as the *Pyrococcus* DNA polymerase intein,[8] *Mycobacterium tuberculosis* RecA intein (*Mtu* RecA),[14] DnaE split intein from *Synechocystis* sp. PCC6803 (*Ssp* DnaE),[20] and *Pyrococcus abyssi* PolII intein (*Pab* PolII).[21] Conversely, the increase in k_3 at pH >7.5 is explained by an increase in the nucleophilicity of DTT upon thiol deprotonation. This has also been observed for the *Ssp* DnaE intein.[20] pH variation appears to have no effect on k_4.

DISCUSSION

A fundamental unanswered question in the noncanonical splicing mechanism of the *Mja* KlbA intein has been how Ala1 inteins overcome the barrier to direct attack of a distal Cys residue on the scissile amide bond, when the equivalent attack apparently cannot occur in standard inteins with a C1A substitution. In an effort to address this question, we previously determined the structure of the *Mja* KlbA intein by NMR spectroscopy.[10,26] The structure suggests that a widening in the active site of the *Mja* KlbA intein as compared to standard inteins, such as *Thermococcus kodakaraensis* Pol-2 intein,[27] *Saccharomyces cerevisiae* VMA intein (*Sce* VMA),[28] and *Mxe* GyrA intein,[29] could allow the C+1 nucleophile to access the −1 carbonyl group. Steric hindrance blocks such attack at the N-terminal bond in standard inteins. In this study, we have performed a kinetic dissection of the cis splicing mechanism of the *Mja* KlbA intein, utilizing an intein−mini-extein system with a disulfide redox switch, to shed additional light on the various steps of this reaction. Intriguingly, our results reveal that the rate constants for formation of the branched intermediate and its resolution to spliced products in the MEI_{166}-$\mathbf{P}_{1NC}$ reaction are almost indistinguishable (Figure 6 and Table 4). This observation contrasts with what has previously been observed for standard inteins (cis and trans splicing), in which C-terminal cleavage (Asn or Gln cyclization) has been identified as the slowest step of the splicing reaction (a summary of rate constants from various intein systems is presented in Table 4). In *Pab* PolII intein, C-terminal cleavage is 10-fold slower than N-terminal cleavage.[21,22] In the *Ssp* DnaE intein, it is ~7-fold slower.[20] A recent study by the Muir group directly monitoring the formation of the succinimide product from a semisynthetic branched intermediate presents strong evidence that favors intein-succinimide formation being the slowest step in the *Mxe* GyrA intein splicing mechanism.[30] Another study by the same group reports rate constants for branched intermediate formation being significantly greater than the rate constants of spliced product formation in the trans splicing reactions of *Ssp* DnaE and the chimeric split intein, composed of the N-terminal *Nostoc punctiforme* DnaE intein and C-terminal *Ssp* intein (Npu*), both inteins with model N- and C-exteins and a C+1S substitution.[31]

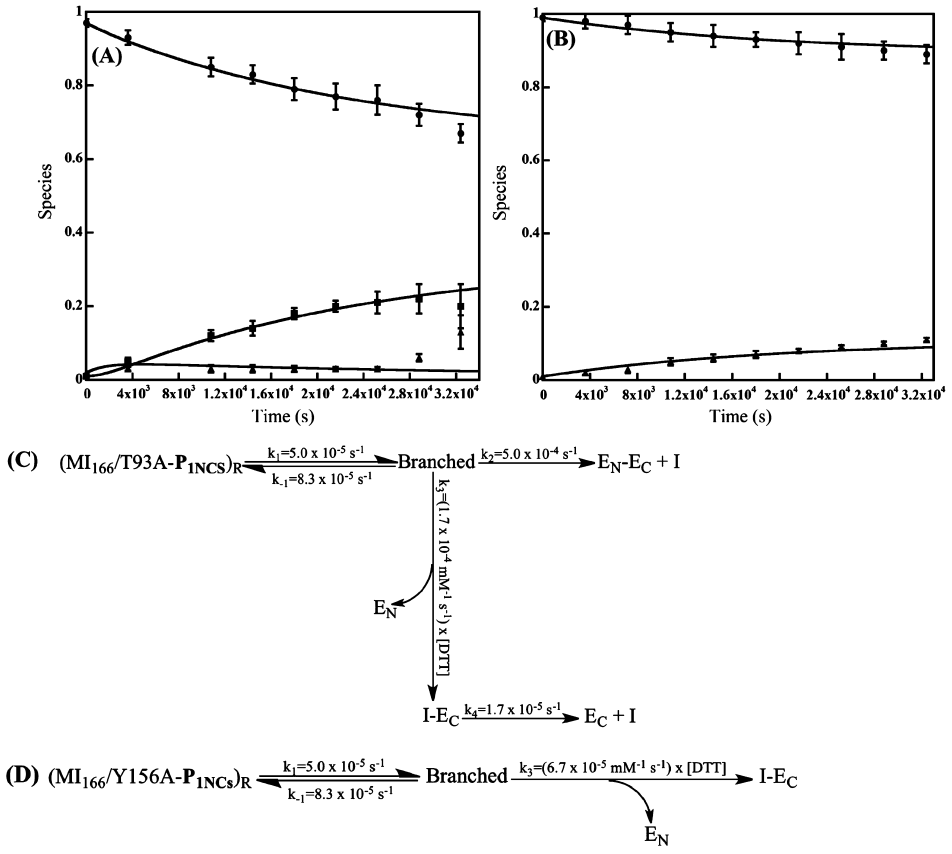
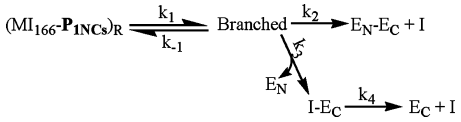


Figure 11. Kinetic simulations of (A) MI_{166}/T93A-$\mathbf{P}_{1NCs}$ and (B) MI_{166}/Y156A-$\mathbf{P}_{1NCs}$ in the presence of 1 mM DTT at 42 °C. The circles are kinetic data points obtained for the precursor, squares for E_N-E_C, and triangles for I-E_C. The kinetic model used for the simulation of the variant reactions is identical to that used for MI_{166}-$\mathbf{P}_{1NC}$, described in the legend of Figure 6B, with variation in rate constants. (C) The percentage value used in the simulation model for $(MI_{166}/T93A\text{-}\mathbf{P}_{1NCs})_R$ was 33%. (D) The percentage value used in the simulation model for $(MI_{166}/Y156A\text{-}\mathbf{P}_{1NCs})_R$ was 10%. The details of the protein splicing assays and quantitation of the various species are as described in Materials and Methods. Samples were quenched in blue loading buffer [6% (w/v) SDS] containing 8 M buffered urea and loaded on a 10−20% Tris-glycine gel without prior boiling. MI_{166}/T93A-$\mathbf{P}_{1NCs}$ and MI_{166}/Y156A-$\mathbf{P}_{1NCs}$ were used at a final concentration of 0.1 μg/μL. Error bars represent the standard error of the mean.

Table 3. Effect of pH Variation on Rate Constants (s^{-1}) in the Splicing Reaction of MI_{166}-P_{1NCs} at 42 °C



	pH 6.0	pH 6.5	pH 7.0	pH 7.5[a]	pH 8.0	pH 8.5	pH 9.0
k_1	3.3×10^{-4}	5.0×10^{-4}	6.7×10^{-4}	6.7×10^{-4}	8.3×10^{-4}	8.3×10^{-4}	6.7×10^{-4}
k_{-1}	1.7×10^{-4}	1.7×10^{-4}	1.7×10^{-4}	1.7×10^{-4}	5.0×10^{-4}	5.0×10^{-4}	5.0×10^{-4}
k_2	1.7×10^{-3}	2.0×10^{-3}	1.7×10^{-3}	2.2×10^{-3}	1.5×10^{-3}	1.2×10^{-3}	8.3×10^{-4}
k_3	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}	3.3×10^{-4}	6.7×10^{-4}	6.7×10^{-4}
k_4	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}	1.7×10^{-5}

[a]The time course and kinetic simulation of the splicing reaction of MI_{166}-$\mathbf{P}_{1NCs}$ at pH 7.5 and 42 °C are presented in panels A and C of Figure S2 of the Supporting Information, respectively.

It is plausible that the diminished rate at which the branched intermediate forms in the context of the splicing reaction of the *Mja* KlbA intein might be attributed to the greater entropy of activation relative to the entropy of attack by the neighboring Cys in the standard inteins, which adds to the enthalpic cost for conversion of an amide to a (thio)ester. Despite this discrepancy between the *Mja* KlbA intein and the standard inteins, the value of the rate constant for formation of the branched intermediate in the *Mja* KlbA intein is comparable to the rate constant for formation of the linear (thio)ester intermediate in the *Mxe* GyrA intein and is 10-fold greater than the values of the rate constants for formation of the branched intermediates in the *Ssp* DnaE and Npu* inteins (Table 4). This situation reflects a surprising efficiency of the *Mja* KlbA intein in directing the attack by the distant C+1 nucleophile on the scissile peptide bond, a step standard inteins are incapable of performing in the absence of the N-terminal nucleophile. It is conceivable that such efficiency is the result of a highly ordered conformation for the active form of the *Mja* KlbA precursor, in which the spatial orientation of C+1 relative to the scissile amide bond mimics that of C1 in standard inteins. Another possibility would be a strained geometry of the active

Table 4. Summary of Rate Constants for the Different Steps in the Splicing Reactions of Various Intein Systems

intein	linear	branched	splicing[a]	N-terminal cleavage[b]	C-terminal cleavage	T (°C)	ref
Pab PolII (cis)	nd[c]	nd[c]	9.3×10^{-6} s^{-1}	~1.0×10^{-4} s^{-1}	1.2×10^{-5} s^{-1}	60	21, 22
Sce VMA (artificial/trans)[d]	nd[c]	nd[c]	0.9–2.0×10^{-3} s^{-1}	nd[c]	1.7×10^{-3} s^{-1}	25	34
Sce VMA (cis)	nd[c]	nd[c]		1.9×10^{-3} s^{-1}	nd[c]	23	19
Ssp DnaB (artificial/trans)	nd[c]	nd[c]	9.9×10^{-4} s^{-1}	nd[c]	nd[c]	25	34
Ssp DnaE (trans)	nd[c]	nd[c]	0.7–3.3×10^{-4} s^{-1}	~1.3×10^{-3} s^{-1}	1.9×10^{-4} s^{-1}	23	16, 18, 20, 35
Ssp DnaE (trans)	nd[c]	8.9×10^{-5} s^{-1}	nd[c]	nd[c]	nd[c]	30	31
Npu DnaE (trans)	nd[c]	nd[c]	1.1×10^{-2} s^{-1}	nd[c]	nd[c]	37	36
Npu*[e]	nd[c]	1.1×10^{-4} s^{-1}	3.9×10^{-5} s^{-1}	nd[c]	nd[c]	30	31
Mxe GyrA (cis)	1.0×10^{-3} s^{-1}	nd[c]	2.0×10^{-5} s^{-1}[f] 2.0×10^{-3} s^{-1}[g]	nd[c]	nd[c]	25	30, 37
Mxe GyrA (cis)	nd[c]	nd[c]	nd[c]	3.6×10^{-2} s^{-1}	nd[c]	23	23
Mja KlbA	–	(k_1 = 1.5×10^{-3} s^{-1}; k_{-1} = 1.7×10^{-5} s^{-1})[h]	2.2×10^{-3} s^{-1}	(2.0×10^{-2} M^{-1} s^{-1}) × [DTT]	1.7×10^{-5} s^{-1}[i]	42	this work

[a]Rate constant for the overall reaction except in the cases of *Mja* KlbA and Npu* where it is the rate constant for the production of ligated exteins from a branched intermediate. For *Mxe* GyrA, see footnotes f and g. [b]DTT-induced N-terminal cleavage. [c]Not determined. [d]Artificially split trans-acting intein. [e]A chimeric split intein composed of the N-terminal *Npu* DnaE intein and the C-terminal *Ssp* intein. [f]Rate constant for succinimide formation. [g]Rate constant for the O–N acyl shift. [h]k_1 is the forward rate constant and k_{-1} the reverse rate constant in the reversible step for Branched formation. [i]Rate constant for formation of the double-cleavage product.

site, which would lead to ground-state destabilization of the precursor and perhaps compensate for the energetically costly rearrangement of an amide bond to a thioester.

Several observations made in this study support the idea that the *Mja* KlbA intein has evolved to tightly control its active site for effective splicing via a complex network of hydrogen bonds involving not only catalytic but also second- and third-sphere residues. Results showing that K_{eq} describing the formation of the branched intermediate is diminished with a shorter E_C imply that splicing is driven forward not only by intradomain protein contacts (residues within the intein active site) but also by interdomain protein contacts (residues in the N- and C-terminal extein domains and the intein domain), which probably result in transition-state stabilization for branched intermediate formation. Furthermore, larger extein domains are probably better at excluding water or thiol molecules from the active site of the intein, which would explain why more off-pathway products are usually observed with heterologous exteins. More insight into the role of interdomain contacts in the overall splicing reaction is gained by the in vitro evidence presented for tight coupling between the N- and C-termini. Substitution of T93, which is involved in N-terminal activation, with A substantially decreases the rates of both branched formation and Asn cyclization. Substitution of C+1 or D147 (shown by an NMR structural model to be hydrogen-bonded via its side chain to Sγ of C+1[10]) with any residue abrogates both N- and C-terminal cleavage events. In light of these observations, and in consideration of the absence of an absolute requirement for N-terminal cleavage to precede C-terminal cleavage in vivo, we suggest the presence of a regulated cross communication between the two termini in the form of a network of shared bonds. The inconsistency of the in vitro and in vivo results could be explained by slight differences in the folding pattern of the intein, which would again indicate sensitivity in the described network to minor disruptions in bonding or folding patterns. Other intein systems, such as the *Ssp* DnaE intein, the *Mtu* RecA intein, and the *Sce* VMA intein, exhibit tight coordination between the N- and C-terminal cleavage events that is explained in terms of a conformational change during N-terminal cleavage, which optimally positions residues for C-terminal cleavage.[20,32,33] Furthermore, it has been observed that residues in the N- and C-terminal motifs, flanking the catalytic center, impose a dramatic effect on the splicing reaction at both N- and C-terminal junctions in these same inteins.[28,31] A similar explanation could be extended to the *Mja* KlbA intein, where local conformational changes could be triggered during or after branched intermediate formation to assist in positioning of essential residues for C-terminal cleavage.

The sensitivity of the *Mja* KlbA intein environment to changes in its catalytic and assisting residues is underscored by (i) the existence of structural isomers of the MI_{166}-$\mathbf{P_{1NC}}$, MI_{166}-$\mathbf{P_{2AC}}$, and MI_{166}-$\mathbf{P_{3QC}}$ precursors and (ii) their varying reactivities, as predicted by kinetic simulations. This sensitivity is clearly emphasized in the MI_{166}-$\mathbf{P_{3QC}}$ reaction, in which the folding properties of the precursor result in two parallel splicing trajectories. Moreover, a disparity in the reactivity of the branched intermediate formed in theses reactions toward DTT most likely reflects subtle conformational differences in these species. MI_{166}-$\mathbf{P_{3QC}}$ Branched I and II are both susceptible to thiolysis, whereas the MI_{166}-$\mathbf{P_{1NC}}$ and MI_{166}-$\mathbf{P_{2AC}}$ Branched species are much less so. It is quite interesting, though, that Branched I (thought to mimic MI_{166}-$\mathbf{P_{1NC}}$ Branched) undergoes faster thiolysis than Branched II (thought to mimic MI_{166}-$\mathbf{P_{2AC}}$ Branched). This greater susceptibility would indicate a more open active site for the former species and perhaps explain the stability of Branched in MI_{166}-$\mathbf{P_{2AC}}$. Furthermore, it is intriguing that DTT-induced N-terminal cleavage leads solely to the formation of the N-terminal cleavage side reaction products in various standard inteins such as the *Sce* VMA intein,[19] *Ssp* DnaE intein,[20] and *Pab* PolII intein,[21,22] while in the case of the *Mja* KlbA intein, the presence of even 50 mM DTT still leads mostly to the formation of the splicing products, ligated exteins and intein. The values of the N-terminal cleavage rate constants for the various inteins are comparable when compared at the same temperature in the presence of 50 mM DTT (Table 4). However, the rate for formation of the succinimide-intein product from the branched

intermediate is what dictates the fate of the reaction products in the presence of DTT in either case. In the studied standard inteins, Asn/Gln cyclization is the slowest step of the reaction and is ~10-fold slower than N-terminal cleavage (Table 4). As a result, the reaction is directed toward formation of N-terminal cleavage products in the presence of high DTT concentrations. On the other hand, in the *Mja* KlbA intein, Asn cyclization occurs at a rate comparable to that of branched formation and that of N-terminal cleavage in the presence of 50 mM DTT. This results in the bifurcation of the reaction in two pathways and allows the formation of the splicing products as well as N-terminal cleavage products.

In summary, we have constructed an intein−mini-extein system, which closely mimics the native *Mja* KlbA precursor, with seven native N-extein residues and 30 native C-extein residues, and a triggerable redox switch. Our system has permitted a detailed kinetic dissection of the splicing mechanism of this intein with the assignment of rate constants for every step involved in the splicing reaction, a task unprecedented not only for noncanonical inteins such as the *Mja* KlbA intein but also for standard inteins. Importantly, our kinetic data provide a number of "firsts" that distinguish the *Mja* KlbA noncanonical intein from all studied standard inteins. (1) The rate constant for formation of the branched intermediate is comparable if not slightly less than the rate constant of Asn cyclization. (2) Despite the slow nature of formation of a branched intermediate in the splicing reaction of the *Mja* KlbA intein, it still occurs efficiently at a rate equivalent to that of linear (thio)ester intermediate formation and ~10-fold faster than branched intermediate formation in standard inteins. (3) The branched intermediate is not susceptible to thiolysis by DTT, as are the linear (thio)ester and branched intermediates in standard inteins, and this stability is ensured by comparable rates of formation of branched and succinimide-intein adducts, which guarantees that the reaction proceeds toward the productive formation of the splicing products. Therefore, although the exact molecular mechanism that allows the *Mja* KlbA intein to overcome the barrier present in standard inteins that prevents attack on an amide bond at the intein N-terminus is yet to be determined, the kinetics of the reaction described herein demonstrate that splicing in this Ala1 intein is not limited by kinetic parameters.

■ ASSOCIATED CONTENT

Supporting Information

Determination of the site for peptide ligation in the *Mja* KlbA intein−mini-extein system (text and Tables S1 and S2) and proposed structures of Branched·DTT, kinetic simulations of the MI_{166}-$\mathbf{P_{1NCs}}$ and MI_{166}-$\mathbf{P_{2ACs}}$ reactions, in vitro products for the MI_{166}-$\mathbf{P_{6NT}}$ reaction, and mass spectrometric analysis of the products of the splicing reaction of MI_{166}-$\mathbf{P_{1NCs}}$ (Figures S1−S4, respectively). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*F.B.P.: 240 County Rd., Ipswich, MA 01908; phone, (978) 380-7326; fax, (978) 921-1350; e-mail, perler@neb.com. L.S.: 240 County Rd., Ipswich, MA 01908; phone, (978) 380-7446; fax, (978) 921-1350; e-mail, saleh@neb.com.

Funding

This work was supported by New England Biolabs.

■ ACKNOWLEDGMENTS

We thank Dr. Catherine L. Madinger for performing ESI-TOF mass spectrometric analysis on the purified peptides. We thank Shelley Cushing for conducting N-terminal sequencing. We also thank Dr. Christopher J. Noren for his scientific input. We thank Dr. Don Comb for support and encouragement.

■ ABBREVIATIONS

Branched, branched intermediate; E_N, seven native N-extein residues with an N-terminal MBP tag; E_C, 10 or 30 native C-extein residues; E_N-E_C, ligated exteins; EPL, expressed protein ligation; ESI-TOF MS, electrospray ionization time-of-flight mass spectrometry; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; I, free intein; I-E_C, intein-C-extein; IPTG, isopropyl *β*-D-thiogalactopyranoside; *k*, rate constant; K_{eq}, equilibrium constant; Lys(5(6)-FAM), Lys[5(6)-carboxyfluorescein]; MBP, maltose binding protein; MI_{166}, MBP-tagged (N-terminus) *Mja* KlbA intein with seven native N-extein residues and with intein truncated at residue Ala166; MI_{166}-**P**, precursor resulting from ligation of MI_{166} and **P**; $(MI_{166}\text{-}\mathbf{P_{1NC}})_U$, unreactive form of MI_{166}-$\mathbf{P_{1NC}}$; $(MI166\text{-}\mathbf{P_{1NC}})_R$, reactive form of MI_{166}-$\mathbf{P_{1NC}}$; MESNA, sodium 2-mercaptoethanesulfonate; NEB, New England Biolabs; PCR, polymerase chain reaction; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; Tris, tris(hydroxymethyl)-aminomethane.

■ ADDITIONAL NOTES

[a]No splicing activity was detected for any of the assembled precursors in the presence of 6 M buffered urea.
[b]The addition of an MBP tag at the N-terminus distinguishes the N-terminal cleavage product and ligated exteins on an SDS−PAGE gel.
[c]E_N corresponds to the seven native N-extein residues with an N-terminal MBP tag.

■ REFERENCES

(1) Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., Noren, C. J., Thorner, J., and Belfort, M. (1994) Protein splicing elements: Inteins and exteins—a definition of terms and recommended nomenclature. *Nucleic Acids Res. 22*, 1125−1127.
(2) Paulus, H. (2000) Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem. 69*, 447−496.
(3) Perler, F. B. (2005) Protein splicing mechanisms and applications. *IUBMB Life 57*, 469−476.
(4) Saleh, L., and Perler, F. B. (2006) Protein splicing in cis and in trans. *Chem. Rec. 6*, 183−193.
(5) Xu, M. Q., and Perler, F. B. (1996) The mechanism of protein splicing and its modulation by mutation. *EMBO J. 15*, 5146−5153.
(6) Cooper, A. A., Chen, Y. J., Lindorfer, M. A., and Stevens, T. H. (1993) Protein splicing of the yeast TFP1 intervening protein sequence: A model for self-excision. *EMBO J. 12*, 2575−2583.
(7) Southworth, M. W., Benner, J., and Perler, F. B. (2000) An alternative protein splicing mechanism for inteins lacking an N-terminal nucleophile. *EMBO J. 19*, 5019−5026.
(8) Xu, M. Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. (1993) In vitro protein splicing of purified precursor and the identification of a branched intermediate. *Cell 75*, 1371−1377.
(9) Tori, K., Dassa, B., Johnson, M. A., Southworth, M. W., Brace, L. E., Ishino, Y., Pietrokovski, S., and Perler, F. B. (2010) Splicing of the mycobacteriophage Bethlehem DnaB intein: Identification of a new mechanistic class of inteins that contain an obligate block F nucleophile. *J. Biol. Chem. 285*, 2515−2526.

(10) Johnson, M. A., Southworth, M. W., Herrmann, T., Brace, L., Perler, F. B., and Wuthrich, K. (2007) NMR structure of a KlbA intein precursor from *Methanococcus jannaschii*. *Protein Sci. 16*, 1316−1328.
(11) Matsudaira, P. (1987) Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. *J. Biol. Chem. 262*, 10035−10038.
(12) Waite-Rees, P. A., Keating, C. J., Moran, L. S., Slatko, B. E., Hornstra, L. J., and Benner, J. S. (1991) Characterization and expression of the *Escherichia coli* Mrr restriction system. *J. Bacteriol. 173*, 5207−5219.
(13) Cui, C., Zhao, W., Chen, J., Wang, J., and Li, Q. (2006) Elimination of in vivo cleavage between target protein and intein in the intein-mediated protein purification systems. *Protein Expression Purif. 50*, 74−81.
(14) Mills, K. V., Lew, B. M., Jiang, S., and Paulus, H. (1998) Protein splicing in trans by purified N- and C-terminal fragments of the *Mycobacterium tuberculosis* RecA intein. *Proc. Natl. Acad. Sci. U.S.A. 95*, 3543−3548.
(15) Posey, K. L., and Gimble, F. S. (2002) Insertion of a reversible redox switch into a rare-cutting DNA endonuclease. *Biochemistry 41*, 2184−2190.
(16) Shi, J., and Muir, T. W. (2005) Development of a tandem protein trans-splicing system based on native and engineered split inteins. *J. Am. Chem. Soc. 127*, 6198−6206.
(17) Dawson, P. E., and Kent, S. B. (2000) Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem. 69*, 923−960.
(18) Nichols, N. M., Benner, J. S., Martin, D. D., and Evans, T. C. Jr. (2003) Zinc ion effects on individual *Ssp* DnaE intein splicing steps: Regulating pathway progression. *Biochemistry 42*, 5301−5311.
(19) Chong, S., Williams, K. S., Wotkowicz, C., and Xu, M. Q. (1998) Modulation of protein splicing of the *Saccharomyces cerevisiae* vacuolar membrane ATPase intein. *J. Biol. Chem. 273*, 10567−10577.
(20) Martin, D. D., Xu, M. Q., and Evans, T. C. Jr. (2001) Characterization of a naturally occurring trans-splicing intein from *Synechocystis* sp. PCC6803. *Biochemistry 40*, 1393−1402.
(21) Mills, K. V., Dorval, D. M., and Lewandowski, K. T. (2005) Kinetic analysis of the individual steps of protein splicing for the *Pyrococcus abyssi* PolII intein. *J. Biol. Chem. 280*, 2714−2720.
(22) Mills, K. V., Manning, J. S., Garcia, A. M., and Wuerdeman, L. A. (2004) Protein splicing of a *Pyrococcus abyssi* intein with a C-terminal glutamine. *J. Biol. Chem. 279*, 20685−20691.
(23) Southworth, M. W., Amaya, K., Evans, T. C., Xu, M. Q., and Perler, F. B. (1999) Purification of proteins fused to either the amino or carboxy terminus of the *Mycobacterium xenopi* gyrase A intein. *Biotechniques 27*, 110−114, , 116, 118−120.
(24) Bruice, T. C., and Benkovic, S. J. (1966) *Bioorganic Mechanisms* (Breslow, R., and Karplus, M., Eds.) Vol. *1*, W. A. Benjamin, Inc., New York.
(25) Iwai, K., and Ando, T. (1967) N-O acyl rearrangement. *Methods Enzymol. 11*, 263−282.
(26) Johnson, M. A., Southworth, M. W., Perler, F. B., and Wuthrich, K. (2007) NMR assignment of a KlbA intein precursor from *Methanococcus jannaschii*. *Biomol. NMR Assignments 1*, 19−21.
(27) Matsumura, H., Takahashi, H., Inoue, T., Yamamoto, T., Hashimoto, H., Nishioka, M., Fujiwara, S., Takagi, M., Imanaka, T., and Kai, Y. (2006) Crystal structure of intein homing endonuclease II encoded in DNA polymerase gene from hyperthermophilic archaeon *Thermococcus kodakaraensis* strain KOD1. *Proteins 63*, 711−715.
(28) Mizutani, R., Nogami, S., Kawasaki, M., Ohya, Y., Anraku, Y., and Satow, Y. (2002) Protein-splicing reaction via a thiazolidine intermediate: Crystal structure of the VMA1-derived endonuclease bearing the N- and C-terminal propeptides. *J. Mol. Biol. 316*, 919−929.
(29) Klabunde, T., Sharma, S., Telenti, A., Jacobs, W. R. Jr., and Sacchettini, J. C. (1998) Crystal structure of GyrA intein from *Mycobacterium xenopi* reveals structural basis of protein splicing. *Nat. Struct. Biol. 5*, 31−36.
(30) Frutos, S., Goger, M., Giovani, B., Cowburn, D., and Muir, T. W. (2010) Branched intermediate formation stimulates peptide bond cleavage in protein splicing. *Nat. Chem. Biol. 6*, 527−533.
(31) Lockless, S. W., and Muir, T. W. (2009) Traceless protein splicing utilizing evolved split inteins. *Proc. Natl. Acad. Sci. U.S.A. 106*, 10999−11004.
(32) Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M. Q. (1996) Protein splicing involving the *Saccharomyces cerevisiae* VMA intein. The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J. Biol. Chem. 271*, 22159−22168.
(33) Van Roey, P., Pereira, B., Li, Z., Hiraga, K., Belfort, M., and Derbyshire, V. (2007) Crystallographic and Mutational Studies of *Mycobacterium tuberculosis* recA Mini-inteins Suggest a Pivotal Role for a Highly Conserved Aspartate Residue. *J. Mol. Biol. 367*, 162−173.
(34) Brenzel, S., Kurpiers, T., and Mootz, H. D. (2006) Engineering Artificially Split Inteins for Applications in Protein Chemistry: Biochemical Characterization of the Split *Ssp* DnaB Intein and Comparison to the Split *Sce* VMA Intein. *Biochemistry 45*, 1571−1578.
(35) Nichols, N. M., and Evans, T. C. Jr. (2004) Mutational analysis of protein splicing, cleavage, and self-association reactions mediated by the naturally split *Ssp* DnaE intein. *Biochemistry 43*, 10265−10276.
(36) Zettler, J., Schutz, V., and Mootz, H. D. (2009) The naturally split *Npu* DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett. 583*, 909−914.
(37) Romanelli, A., Shekhtman, A., Cowburn, D., and Muir, T. W. (2004) Semisynthesis of a segmental isotopically labeled protein splicing precursor: NMR evidence for an unusual peptide bond at the N-extein-intein junction. *Proc. Natl. Acad. Sci. U.S.A. 101*, 6397−6402.