Zinc Inhibition of Protein trans-Splicing and Identification of Regions Essential for Splicing and Association of a Split Intein*

Received for publication, December 7, 2000, and in revised form, April 26, 2001
Published, JBC Papers in Press, April 30, 2001, DOI 10.1074/jbc.M011049200

Inca Ghosh‡, Luo Sun‡, and Ming-Qun Xu§
From New England Biolabs, Inc., Beverly, Massachusetts 01915

Two important aspects of protein splicing were investigated by employing the trans-splicing intein from the dnaE gene of Synechocystis sp. PCC6803. First, we demonstrated that both protein splicing and cleavage at the N-terminal splice junction were inhibited in the presence of zinc ion. The trans-splicing reaction was partially blocked at a concentration of 1–10 μM Zn2⁺ and completely inhibited at 100 μM Zn2⁺; the inhibition by zinc was reversed in the presence of ethylenediaminetetraacetic acid. We propose that inactivation of Cys160 at the C-terminal splice junction by the chelation of zinc affects both the N-S acyl rearrangement and the transesterification steps in the splicing pathway. Furthermore, in vivo and in vitro assays were established for the determination of intein residues and regions required for splicing or association between the N- and C-terminal intein halves. N-terminal truncation of the intein C-terminal segment inhibited both splicing and association activities, suggesting this region is crucial for the formation of an interface between the two intein halves. The replacement of conserved residues in blocks B and F with alanine abolished splicing but allowed for association. This is the first evidence showing that the conserved residues in block F are required for protein splicing.

Protein splicing is a post-translational processing event, which involves the self-catalyzed excision of an internal protein segment, or intein, from a protein precursor with the concomitant joining of the flanking polypeptide sequences, the exteins (1, 2). Sequence alignment reveals that an intein can be divided into three segments. The N-terminal region possesses ~120–150 amino acid residues including highly conserved blocks A and B, whereas the C-terminal region is composed of ~35–50 residues containing conserved blocks F and G. Between the two terminal regions is an optional endonuclease domain, which has been found in a majority of inteins (3, 4). In the case of protein trans-splicing, however, a functional intein is reconstituted from two non-covalently linked N- and C-terminal segments that are separately translated (5, 6). The crystal structures of inteins from a vacuolar ATPase subunit of Saccharomyces cerevisiae (PI-SceI or Sce VMA intein),¹ Mycobacterium xenopi GyrA (Mxe GyrA intein), and Pyrococcus furiosus ribonucleotide reductase (PI-PfuI) revealed that the N- and C-terminal regions form a horseshoe-shaped Hint (hedgehog and intein) domain (7–9). The intein structure contains an unusual β-fold with the splice junctions at the ends of two adjacent antiparallel β-strands, forming a catalytic pocket. The catalytic residues implicated in the splicing mechanism have been found in the conserved blocks, A and G, present at the two splice junctions (10–16). Presumably, these residues directly participate in protein splicing by three concerted nucleophilic replacements (1, 17) (Fig. 1). Other residues and regions that may be involved in assisting these catalytic reactions have yet to be identified, although a highly conserved histidine residue in block B has been found to be necessary for protein splicing (18). Furthermore, protein splicing requires the precise alignment of the two splice junctions to form the active site. Examination of the interaction between the N- and C-terminal intein halves would offer important insight into the tertiary folding that brings all the reacting groups in close proximity.

Recently, the crystal structure of a precursor protein containing the Sce VMA intein with the flanking native extein residues was solved (19). Remarkably, a zinc atom was found at the catalytic center of this splicing-deficient precursor protein. The residues contributing to the zinc coordination include a cysteine following the C-terminal splice site, which presumably receives the N-extein domain by transesterification, a critical step in the complex splicing pathway. Thus, it is of particular interest to investigate the possibility that zinc may play a modulatory role during the process of protein splicing. However, the rapid processing of intein precursors presents a major obstacle to the conduction of such an investigation (10–12).

In this report, an important advance in understanding the complexities of protein splicing is made possible by the manipulation of the trans-splicing intein involved in the maturation of the Synechocystis sp. DnaE protein (5, 6). The naturally occurring split intein, consisting of an N-terminal segment of 123 residues and a C-terminal segment of 36 residues, can mediate efficient splicing in trans in foreign protein contexts (20, 21). Similar to the Sce VMA intein, the sulphhydryl group of the cysteine residues presumably function as nucleophiles at both the N- and C-terminal splice junctions of the DnaE intein (Fig. 1). In this study, we took advantage of the unique property of the trans-splicing Ssp DnaE intein to explore the effect of zinc on splicing activity. This trans-splicing system also allowed us to investigate the intein elements responsible for the interaction between the N- and C-terminal halves of a protein splicing element, independent of its catalytic activity. We demonstrated the correlation between the splicing or association functions of the DnaE intein and the reconstitution of Esche-

¹ The abbreviations used are: Sce VMA intein, the intein from the 69-kDa vacuolar ATPase subunit of S. cerevisiae; ALS, E. coli acetyl-CoA synthase isofrom II; CBD, chitin binding domain; DTT, 1,4-dithiothreitol; GST, glutathione S-transferase; I₉, the 36 C-terminal amino acids of the Ssp DnaE intein; Iₖ, the 123 N-terminal amino acids of the Ssp DnaE intein; IPTG, isopropyl-β-D-thiogalactopyranoside; MBP, maltose-binding protein; Mxe GyrA intein, the intein from M. xenopi GyrA; PAGE, polyacrylamide gel electrophoresis; Ssp DnaE intein, the split intein encoded by the dnaE gene from Synechocystis sp. PCC6803.
Characterization of a Split Intein from Synechocystis sp. PCC6803

richia coli acetolactate synthase isomor II (ALS). Co-expression of the two DNA intein fragments fused to two halves of ALS rescued E. coli ER2744, which lacks an active ALS, from growth inhibition by valine (22–24). The effect of amino acid substitutions or deletions on the splicing and association activities of the two intein halves were further characterized by an in vitro assay system.

Experimental Procedures

Cells and Materials—E. coli ER2744 (flaA2 glnV44 el4- rfbD1 relA1 endA1 spoT1 thi-1 Δ(mcr-c-mrr)114::IS10 lacZΔ(T7 genl) contains a mutated ALS gene.

Primer Construction—pMEB10 is a derivative of pMEB4 (20) and expresses a fusion protein of E. coli maltose-binding protein (MBP or M) fused at its C terminus to the N-terminal 123-residue segment (Iβ) of the Ssp DnaE intein. pBEL11 expresses a three-part fusion protein containing a N-terminal chitin binding domain (CBD) from Bacillus circulans followed by the C-terminal 36-residue fragment (Iα) of the Ssp DnaE intein and T4 DNA ligase (20). Two pBEL11 variants were created by substitution of the last intein residue, Asn159, or the first DnaE intein and T4 DNA ligase (20). Two pBEL11 variants were created by linker replacement using the unique NruI and AgeI sites. The coding sequence of E. coli ALS was split into two segments and fused in-frame to the N- and C-terminal halves of the Ssp DnaE intein (Iα and Iβ) as described by Sun et al. (25). All intein fusion constructs contain five native residues at the splice junction, allowing in vitro splicing efficiencies (20, 23). pEN10 was created by the in-frame fusion of the DNA fragment encoding the N-terminal 327 amino acid residues of ALS, or ALS(N), carrying the A262 mutation (24), to the N terminus of Iα generating the ALS(N)-Iβ fusion gene. pKCE1 was constructed by fusing the DNA fragment encoding the C-terminal 221 amino acid residues of ALS, ALS(C), in frame from the C terminus of Iα, creating the Iα-ALS(C) fusion gene. Mutations in block B were introduced by site-directed mutagenesis (QuickChange site-directed mutagenesis kit, Stratagene, La Jolla, CA) using the following primers: T89A, 5′-CAGTAACTTGAGCTGCTTGCACGACCCACGC-3′; and 5′-CAGTGAATCTGATCCGAGGAAACGCTCCAGATCTCTG-3′; and H72A, 5′-GATGATCTGTATGTCGGCAACCCCTTATCGG-3′. Co-expression of MBP-IN (MIN, 5′-TATGGTTTCAGCGCATCTTTGA-3′, and 5′-TATGGTTTCAGCGCATCTTTGA-3′) with Iα and Iβ sites in pKEC3, was obtained from NCBI Structures data base (9) and used as a model for the Ssp DnaE intein. Fig. 3B was produced using Swiss-PdbViewer version 3.7b1. The structure corresponding to the region between positions Gly132 and Tyr132 of the Mxe GyrA intein is not shown.

In Vivo Reconstitution of Active ALS by Plate Assays—E. coli ER2744 cells were transformed with the appropriate plasmids and plated on agar plates containing M9 minimal medium containing 2 μg/ml thiamine, 2 μM MgSO4, 0.1 mM CaCl2, 0.2% glucose, 50 μg/ml kanamycin, 100 μg/ml ampicillin, 0.3 mM IPTG, and 100 μg/ml valine. The plates were incubated at 30 °C for 48–72 h.

Detection of Spliced Product by Western Blot Analysis—A single bacterial colony was inoculated in LB medium supplemented with 100 μg/ml ampicillin and 50 μg/ml kanamycin for 4 h at 37 °C. Protein expression was induced with 0.3 mM IPTG for 16 h at 30 °C. Whole cell lysates were resolved by SDS-PAGE. Western blots with the antibodies against the N- or C-terminal fragments of ALS were performed as described previously (23).

In Vitro Assays for Splicing and Association Activities—The chimeric proteins MBP-IN and IC-GST, were mixed in vitro to assess the effect of mutations on splicing and the interaction of the two intein halves. The MBP-IN proteins were expressed and purified on amyllose. The IC-GST fusion protein, expressed from pEG3 or its mutant derivative, was purified from the cell lysate by glutathione-Sepharose4B resin (Amersham Pharmacia Biotech). The purified MBP-IN protein was added to Iα-GST-bound resin. After a 45-min incubation at 4 °C, the column was washed with at least 20 column volumes of Buffer A. Samples of glutathione-Sepharose resin taken when MBP-IN was taken on the resin (t0), after incubation of 45 min, and after the resin was thoroughly washed. The samples were incubated at 4 °C for 24 h before being analyzed by Coomassie Blue-stained SDS-PAGE.

Results

Zinc Inhibits Protein trans-Splicing—The effect of zinc on trans-splicing of the Ssp DnaE intein was examined by employing the in vitro trans-splicing system described previously (20) (Fig. 1). This system allowed the splicing reaction to occur between two bacterially expressed proteins, MBP-IN (M, 57 kDa) and CBD-Iα-T4 DNA ligase (BI CL, 69 kDa), to yield the spliced product MBP-T4 DNA ligase (ML, 100 kDa). Analysis of the reaction by SDS-PAGE revealed that the spliced product, ML, was not observed when zinc ion is present at 1 mM concentration (Fig. 2A), suggesting protein trans-splicing is inhibited. This inhibitory effect by zinc ion could be reversed by including 5 mM EDTA (metal chelating agent). The presence of either 1 mM Ca2+ or 1 mM Mg2+ appeared to have no significant effect on trans-splicing. These data indicated that the inhibitory effect of zinc ion on trans-splicing is not simply due to the presence of a divalent cation, but is specific to zinc. This inhibitory effect was observed for presence of either zinc acetate or zinc chloride (data not shown). The effect of zinc ion on cleavage at the N-terminal splice junction, the initial step in the splicing pathway, was examined in the presence of 5 mM DTT. DTT-induced N-terminal cleavage is due to nucleophilic attack by thiol at thioster bond formed at Cys1 (10, 15). In the absence of Zn2+, both the spliced product ML and the N-
terminal cleavage product (MBP or M, 42 kDa) were generated (Fig. 2A, lane 8). In the presence of zinc, however, DTT induced N-terminal cleavage was abolished since no N-extein, M, was observed (lane 9).

To further investigate the inhibitory effect of zinc on trans-splicing, MBP-IN and CBD-IC-T4 DNA ligase were incubated at 4 °C for 24 h in various concentrations of zinc (Fig. 2B). The Coomassie Blue-stained SDS-PAGE showed that the effect of zinc on trans-splicing is concentration-dependent. The presence of 1 or 10 μM Zn^{2+}, at 2-20-fold in excess of protein concentration, resulted in partial inhibition of trans-splicing (lanes 5 and 6). The amount of the spliced product, ML, decreased by ~85% in the presence of 10 μM Zn^{2+}, when compared with the yield in the absence of zinc (lane 4). However, in the presence of 100 μM zinc ion, no ML was observed (lane 7), indicating that trans-splicing was inhibited.

We reasoned, based on the Sce VMA intein structure, that the inhibition of splicing and N-terminal cleavage is caused by the binding of the zinc ion to Cys^{160}. This chelation will inhibit transesterification, thereby affecting the equilibrium between
Characterization of a Split Intein from Synechocystis sp. PCC6803

Figure 3. Sequence alignment and a model structure of the Ssp DnaE intein. A, alignment of the amino acid sequences of the Ssp DnaE intein with the Mxe GyrA intein and Sce VMA intein plus their first C-extein residue (3, 4, 5, 8, 9). Amino acid residues are numbered forward from the N-terminal splice junction. The conserved intein sequence blocks (A, B, F, and G) are labeled above the sequences. The secondary structure assignments for β strands (β1-β10) and α helices (α1-α2) for the Mxe GyrA intein are underlined. The endonuclease domain (256 amino acids) and DNA recognition region (240-541) of the Sce VMA intein were not included in the diagram. The mutated amino acid residues are boxed. B, a model structure of the Ssp DnaE intein based on the x-ray structure of the Mxe GyrA intein (9).

Ribbon drawing shows the region that corresponds to the C-terminal segment of DnaE intein (red) and the region that corresponds to the N-terminal segment of DnaE intein (green). The C-terminal 23 amino acid residues of Ic, corresponding to the flexible linker region of the Mxe GyrA intein, is not shown. N and C termini of the intein are labeled N and C, respectively. Also labeled are the predicted amino acid sequences from residue 95 to 100 (IN, overlapping β9 strand) in white and the residues from residue 127 to 130 (IC, overlapping β10 strand) in blue.
The highly conserved intein blocks B in \( I_N \) and F in \( I_C \) on the production of an active ALS. The residues Thr\(^{69}\) and His\(^{72}\) in block B and Phe\(^{139}\), Asp\(^{140}\), His\(^{147}\), Asn\(^{148}\), and Phe\(^{149}\) in block F were substituted with Ala. For all these mutants, co-expression with the wild-type fusion protein permitted the growth of ER2744 cells in the presence of valine (Fig. 4C).

**Analysis of Splicing Activities by Immunoblotting**—Reconstitution of functional ALS could be due to either protein trans-splicing or association between the two intein segments. Western blot analysis was carried out to detect if a spliced ALS product (59 kDa) was produced as a result of co-expression of both intein fusion constructs. As shown in Fig. 5, co-expression of the fusion proteins possessing the wild-type intein segments permitted protein trans-splicing to occur, generating a 59-kDa product, which reacted with antibodies against the ALS N-terminal fragment or the C-terminal fragment. A nonspecific 61-kDa protein cross-reacted with anti-ALS(N) antibody. The species corresponding to the full-length ALS product was also observed when 3–23 residues were deleted from the C terminus of \( I_N \) (\( I_N^{Δ3–23} \)) or 1 and 2 residues (\( I_C^{Δ1} \) and \( I_C^{Δ2} \)) were deleted from the N-terminal region of \( I_C \) (Fig. 5, A–C). These results are in agreement with the plate assay data showing that these mutants rescued the host cells from the growth inhibition by valine. A minor discrepancy was observed in the \( I_N^{Δ3} \) mutant, which exhibited splicing activity (Fig. 5C) but only partial growth in the plate assay (Fig. 4A); the reason for this difference is not clear. The spliced ALS product was not detected when a 29- or 45-amino acid deletion was introduced at the C terminus of \( I_N \) or when 6–12 residues were removed from the N-terminal region of \( I_C \). Furthermore, the substitutions in blocks B and F abolished trans-splicing, since no ALS product was detected using the anti-ALS(C) antibody (Fig. 5D), suggesting that these conserved residues are required for trans-splicing.

**Detection of Intein Fragment Association by an in Vitro Assay**—The observation that some intein mutants failed to mediate trans-splicing but rescued ER2744 host cells from growth inhibition by valine suggested that interaction of the two intein segments may be sufficient for reconstitution of ALS activity without the formation of a spliced product. An in vitro assay was established to determine whether the mutated intein N or C fragments of the DnaE intein interact and if this interaction permitted trans-splicing in vitro. To facilitate the purification of the wild-type or mutated intein segments, the ALS N-terminal coding sequence was replaced by the DNA sequence encoding E. coli MBP (or M) (28) and the ALS C-terminal coding sequence was substituted with the DNA sequence encoding GST (or G) (25). The MBP-\( I_N \) (\( M_{IN} \)) fusion proteins were purified and then added to the purified and glutathione resin-bound \( I_C\)-GST (\( I_C\)-G), as summarized in Table I. After a 45-min incubation at 4 °C, the glutathione resin was washed thoroughly and samples of glutathione-Sepharose were taken and examined by SDS-PAGE. In the control experiment, \( M_{IN} \) and \( I_C\) possessing wild-type intein segments, interacted (Fig. 6A, compare lanes 2 and 3) and mediated trans-splicing, producing MBP-GST (MG) (lanes 2–4). The chaperone protein DnaK, which co-purified with \( I_C\)-G, ran at a molecular mass of ~72 kDa on SDS-PAGE and was present in all the \( I_C\)-G samples. Alanine substitutions in blocks B and F abolished trans-splicing in vitro (Fig. 6), in agreement with Western blot analysis, which showed a lack of in vitro spliced product (Fig. 5D). These mutations, however, did not appear to affect the interaction between \( I_N \) and \( I_C \) since the \( M_{IN} \) species remained bound to the \( I_C\)-G protein column after a thorough wash (Fig. 6, A and B, lanes 2 and 3). N-terminal cleavage activity, as indicated by the appearance of a 42-kDa product, M, was detected in all the
Characterization of a Split Intein from Synechocystis sp. PCC6803

Inhibition of Protein trans-Splicing by Zinc Ion—

**FIG. 5.** Immunoblot analysis of protein trans-splicing activity of various intein mutants. E. coli cells were co-transformed with two compatible plasmids and induced with 0.3 mM IPTG for 12 h at 30 °C to express ALS(N)-IN and IC-ALS(C) fusion proteins. One of the fusions carries a wild-type intein sequence, and the other contains a mutated intein segment, as indicated above each lane. Cell lysates from uninduced cells (lane 1), cells induced to express the full-length ALS (59 kDa, lane 2), or both fusion proteins carrying the wild-type intein sequence (lane 3) were used as controls. Whole cell lysates were resolved by SDS-PAGE. Immunoblotting was performed with antibodies against ALS N-terminal peptide in A or ALS C-terminal peptide in B–D. Protein trans-splicing generated a 59-kDa product, which reacted with both antibodies. A band of ~61 kDa observed in A is an unknown protein that cross-reacts with the anti-ALS(N) serum.

block B mutants except for the F139A mutant (Fig. 6B). Interestingly, the T69A or H72A mutants permitted DTT-induced cleavage at the N-terminal splice site (Fig. 6A, middle and right panels, lane 4). The N148A mutant exhibited splicing activity at 30 °C but not at 4 °C (Fig. 6B, compare lanes 3 and 4), as confirmed by immunoblotting with antibodies against MBP or GST (data not shown). The data suggest that the association of the intein fragments account for the in vivo ALS activity observed in the block B and F mutants. The removal of only 3 residues from the N-terminal region of IC had no effect on interaction between IN and IC or on splicing and generated the spliced product, MG (Fig. 6C). However, deletion of a 10-amino acid region of IC abolished trans-splicing as well as its association with IN, as indicated by the absence of the spliced product MG (Fig. 6C, IC Δ10, lane 2) and a dramatic decrease in the amount of MIu following the wash step (lanes 3 and 4). Similar results were observed when 6 or 8 residues were removed from IC (data not shown). The data also reveal that deletions of up to 29 residues from the C-terminal region of IN still allowed the association of IN and IC, whereas the 45-residue deletion in IN reduced the intein fragment association (Table I and Fig. 6C).

**TABLE I**

| Mutants | In vivo activity | In vitro activity |
|---------|-----------------|------------------|
|         | Splicing | ALS activity | Splicing | Association |
| WT      | + | + | + | + |
| I1,Δ3   | + | + | ND | ND |
| I1,Δ9   | + | + | ND | ND |
| I1,Δ16  | + | + | + | + |
| I1,Δ23  | + | + | + | + |
| I1,Δ29  | − | − | − | − |
| I1,Δ45  | − | − | − | − |
| L1,Δ1   | + | + | ND | ND |
| L1,Δ2   | + | + | + | + |
| L1,Δ5   | + | − | + | + |
| L1,Δ6   | + | − | + | + |
| L1,Δ8   | + | − | + | + |
| L1,Δ10  | + | − | + | + |
| L1,Δ12  | + | − | + | + |
| T69A    | + | + | + | + |
| H72A    | + | + | + | + |
| F139A   | + | + | + | + |
| D140A   | + | + | + | + |
| H147A   | + | + | + | + |
| N148A   | + | + | + | + |
| F149A   | + | + | + | + |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
| ND ND ND ND ND |
Characterization of a Split Intein from Synechocystis sp. PCC6803

B. B. K. Lam, Y. Lin, L. E. S. Koh, F. Q. C. Loo, T. K. S. Lee, J. H. Yang, C. S. Yang

Glu residue at the equivalent position (Glu80) in block B, which are involved in the coordination of zinc in the x-ray structure of the Sce VMA intein (19). It is likely that residues at different positions are involved in the coordination of a zinc atom in the Ssp DnaE intein, as proposed by Mills and Paulus (30). Further understanding of the mode of zinc binding relies on a precise crystal structure of the Ssp DnaE intein.

Inhibition of N-terminal Cleavage by Zinc Ion—The work presented also demonstrates that the presence of zinc ions blocks cleavage of the peptide bond at the N-terminal splice site (Fig. 2A). It is possible that chelation of a zinc atom by Cys160 may exclude the binding of DTT, thereby inhibiting its nucleophilic attack on the thioester bond. On the other hand, inactivation of Cys160 by zinc chelation could affect the acyl rearrangement involving Cys1, thereby shifting the equilibrium of amide and thioester to favor amide conformation (Fig. 1). This scenario is supported by the observation that cleavage at the N-terminal splice junction can be inhibited by zinc in the mutant carrying the N159A substitution, which blocked splicing but permitted N-terminal cleavage, presumably by hydrolysis of the thioester bond (Fig. 2C). Furthermore, substitution of Cys160 with an Ala residue resulted in no N-terminal cleavage, with or without DTT (Fig. 2D). It has been observed previously that the Cys residue at the C-terminal splice junction of the M. tuberculosis RecA intein has a low apparent pKa, which would facilitate the occurrence of transesterification by attacking the thioester bond involving Cys1 (32). Thus, Cys160 appears to play an important role in driving the first splicing step and shifting the equilibrium of amide and thioester.

Mapping the Regions Required for Intein Association—Protein splicing is presumably facilitated by inter- or intramolecular recognition between the N- and C-terminal regions of an intein, which may be separated by an endonuclease domain of more than 300 amino acid residues. For cis-splicing inteins, the tertiary folding occurs via intramolecular interactions of the intein sequence elements. The presence of a distinct endonuclease domain or a flexible linker may facilitate the formation of substructures and the folding process. The reconstitution of the split Ssp DnaE intein, however, requires a high affinity intermolecular interaction (33). The work presented here shows that the association function of the two intein segments is separable from its catalytic activity. trans-Splicing activity, however, appears to correlate with the interaction between the two intein halves, signifying that the interaction of the two fragments is necessary for the formation of the correct tertiary structure required for splicing. The data suggest that the N-terminal region of Lc, adjacent to the conserved block F and part of β10, is essential for intein association. Intein C-terminal fragments carrying 6–12-amino acid deletions are not capable of complementing the growth defect of ER2744 host cells or splicing in vitro or in vitro (Table I and Figs. 4–6). The substantially lower affinity of Lc possessing a deletion, for the wild-type Ic could be due to the disruption of the three-stranded β-sheets (β10:β5:β6). The C-terminal 23 residues of Ic appear not to be required for either association or splicing, probably because they constitute a flexible linker region (Fig. 3A). For the Mxe GyrA intein, the linker region forms a disordered loop and two α-helices extended from the β-core and does not appear to interact with the active site (9). It has yet to be determined whether the DnaE intein possesses a linker region with substructures corresponding to the α-helices of the GyrA intein (9). However, the DnaE intein sequence upstream of this region appears to align with the amino acid residues of the GyrA intein that participate in formation of the β2 strand, which appears to interact with the intein C-terminal segment, and therefore may be important for the formation and stabili-
zation of the active site. Indeed, in the ALS fusion constructs further truncations of this region, as shown in the I_{69,45} mutant, abolished splicing and failed to produce active ALS. The reduced interaction in vitro between I_{69,45} and wild-type I_{C} suggests that the region involved in formation of the β7 and β8 strands is important for both splicing and intein association (Fig. 6C). The 29-residue deletion, presumably disrupting only part of β9 strand, inhibited splicing but had little effect on association in vitro; however, partial growth was observed in the plate assay. The intein splicing and association activities could be affected by the extein context, which may result in the discrepancy between the in vivo and the in vitro data. These ALS intein fusion proteins could have suboptimal interactions, whereas MBP and GST, are effective at promoting the solubility of proteins to which they fused and probably interact proficiently (34).

Effect of Mutations in Blocks B and F on Splicing and Association Activities—The data from the in vivo functional screen and in vitro analysis indicate that the conserved residues in blocks B and F are crucial for splicing but not for association of the two intein fragments. Substitution of Thr^{69} or His^{72} in block B with an alanine residue inhibited splicing but still permitted interaction between the two intein segments. These mutants failed to generate the spliced ALS proteins (Fig. 5D) but were capable of complementing the valine-sensitive growth defect of ER2744 strain (Fig. 4C), suggesting that active ALS can be reconstituted by the association of the two intein halves. The conserved Thr^{69} and His^{72} in block B are implicated to participate in the acyl rearrangement, based on their positions in the active site of the Mxe GyrA intein or the Sce VMA intein (9, 19). In this study acyl rearrangement, however, does not appear to absolutely depend on Thr^{69} or His^{72}, since the T69A and H72A mutant proteins exhibited DTT-induced cleavage activity (Fig. 6). It has been shown previously that mutation of the conserved His in block B of the Sce VMA intein abolished splicing activity in E. coli, but the single mutant was still capable of producing an active VMA protein in yeast (18). Thus, single amino acid substitutions may not completely block splicing as interactions from other intein residues may be sufficient to catalyze the reaction.

Our study provides the first experimental evidence that block F region is essential for splicing, but not association (Table I and Figs. 5 and 6). Crystallographic study of the Sce VMA intein reveals that several residues, including Ile^{434} and Phe^{444} in block F, participate in the formation of a hydrophobic surface at the back of the intein C-terminal asparagine residue. Alanyl replacement of DnaE intein Phe^{439} and Phe^{449}, corresponding to Ile^{434} and Phe^{444} of the Sce VMA intein (Fig. 3) inhibited splicing, supporting the hypothesis that they may participate in the cyclization of asparagine (19). Ser^{379} of the Mxe GyrA intein and Thr^{355} of the Sce VMA intein, at the equivalent position of the DnaE intein residue Asp^{146}, have been implicated in assisting the cyclization of intein C-terminal asparagine (9, 19). Although DnaE intein residue His^{147} is highly conserved among inteins (20 out of 24 eubacterial inteins), its role has not been elucidated. However, the presence of a His, at the equivalent position, in the vicinity of the active pocket of the Sce VMA intein suggests that it may play an important role in splicing (19).

In conclusion, we have characterized several important aspects of protein splicing. Zinc inhibition, together with the reversibility of EDTA, could provide a novel approach to modulate protein splicing and peptide bond cleavage. The in vivo functional screen along with an in vitro assay helped to dissect the regions of the intein essential for splicing and interaction between the two intein halves. Furthermore, a functional enzyme can be produced either by protein trans-splicing or by the association of two intein fragments (23, 35). Therefore, the methods demonstrated here should open up new avenues in the application of self-splicing inteins to express and modify proteins.

Acknowledgments—We thank Thomas C. Evans, Jr., Donald Comb, Florante Quiocho, Bradford Poland, Eric Adam, Lixin Chen, and Fana Mersha for valuable discussion and Emma Naylor and Sebastien Ferrandon for technical assistance.

REFERENCES
1. Paulus, H. (2000) Annu. Rev. Biochem. 69, 447–496
2. 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) Nucleic Acids Res. 22, 1125–1127
3. Perler, F. B., Olsen, G. J., and Adam, E. (1997) Nucleic Acids Res. 25, 1087–1093
4. Pietrokowski, S. (1998) Protein Sci. 7, 64–71
5. Wu, H., Hu, Z., and Liu, X. Q. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 9226–9231
6. Gorbalenya, A. E. (1998) Nucleic Acids Res. 26, 1741–1748
7. Duan, X., Gimble, F. S., and Quiocho, F. A. (1997) Cell 89, 555–564
8. Ichiyanagi, K., Ishino, Y., Aiyoshi, M., Komori, K., and Morikawa, K. (2000) J. Mol. Biol. 300, 899–911
9. Klambt, T., Sharma, S., Telenzi, A., Jacobs, W. R., Jr., and Sazcchetti, J. C. (1998) Nat. Struct. Biol. 5, 31–36
10. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M.-Q. (1996) J. Biol. Chem. 271, 22187–22190
11. Cooper, A. A., Chen, Y.-J., Linderfer, M. A., and Stevens, T. H. (1993) EMBO J. 12, 2575–2582
12. Hirata, R., and Anraku, Y. (1992) Biochem. Biophys. Res. Commun. 188, 40–47
13. Xu, M.-Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. (1993) Cell 75, 1371–1377
14. Xu, M.-Q., Comb, D. G., Paulus, H., Noren, C. J., Shao, Y., and Perler, F. B. (1994) EMBO J. 13, 5517–5522
15. Xu, M. Q., and Perler, F. B. (1996) EMBO J. 15, 5146–5153
16. Shao, Y., Xu, M.-Q., and Paulus, H. (1996) Biochemistry 35, 3810–3815
17. Perler, F. B., Xu, M. Q., and Paulus, H. (1997) Curr. Opin. Chem. Biol. 1, 292–299
18. Kawasaki, M., Nogami, S., Satow, Y., Ohya, Y., and Anraku, Y. (1997) J. Biol. Chem. 272, 15668–15674
19. Poland, B. W., Xu, M.-Q., and Quiocho, F. A. (2000) J. Biol. Chem. 275, 16408–16413
20. Evans, T. C., Martin, D., Kolly, R., Panne, D., Sun, L., Ghosh, I., Chen, L., Benner, J., Liu, X.-Q., and Xu, M.-Q. (2000) J. Biol. Chem. 275, 8909–8914
21. Scott, C. P., Abel-Santos, E., Wall, M., Wahnon, D. C., and Benkovic, S. J. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 13638–13643
22. LaRossa, R. A., and Schloss, J. V. (1984) J. Biol. Chem. 259, 8753–8757
23. Sun, L., Ghosh, I., Paulus, H., and Xu, M.-Q. (2001) Appl. Environ. Microbiol. 67, 1025–1029
24. Yadav, N., McDevitt, E. R., Benard, S., and Falco, C. S. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 4418–4422
25. Smith, D. B., and Johnson, K. S. (1988) Gene (Amst.) 76, 31–40
26. Dalley, P. E., and Cronan, J. E., Jr. (1986) J. Bacteriol. 165, 453–460
27. Hill, C. M., and Duggleby, R. G. (1998) Biochem. J. 335, 653–661
28. Guan, D., Li, P., Riggs, P. D., and Inouye, H. (1988) Biochem. J. 259, 653–661
29. Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990) J. Biol. Chem. 265, 6726–6733
30. Mills, K. V., and Paulus, H. (2001) J. Biol. Chem. 276, 10832–10838
31. Kurn, D. L., Li, P., Biggs, P. D., and Inouye, H. (1988) Gene (Amst.) 76, 21–30
32. Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990) J. Biol. Chem. 265, 6726–6733
33. Martin, D. D., Xu, M.-Q., and Evans, T. C. (2001) Biochemistry 40, 1393–1402
34. Kapust, R. B., and Waugh, D. S. (1999) Protein Sci. 8, 1668–1674
35. Chen, L., Pradhan, S., and Evans, T. C. (2001) Gene (Amst.) 263, 39–48