

Protein splicing: A novel form of gene expression and paradigm for self-catalyzed protein rearrangements

Henry Paulus

*Boston Biomedical Research Institute, 20 Staniford Street, Boston, MA 02114 USA
and Dept. of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115 USA*

Abstract: Protein splicing is one of the mechanisms by which genes that are interrupted by intervening sequences can produce functional proteins. It involves the self-catalyzed excision of an internal segment from an inactive precursor protein and the ligation of the flanking N- and C-terminal segments to yield an active protein. A key reaction in protein splicing is the rearrangement of a peptide bond involving the amino group of serine or cysteine to an ester bond. Such N-O or N-S acyl shifts are also the basis of other self-catalyzed protein rearrangements, which include the cleavage of hedgehog proteins and certain amidotransferases and the formation of pyruvoyl enzymes. Although N-O or N-S acyl rearrangements are thermodynamically unfavorable, their coupling to self-catalyzed irreversible steps drives the protein rearrangements to completion. In protein splicing, these steps are intramolecular transesterification followed by asparagine cyclization and peptide bond cleavage. All steps of protein splicing are catalyzed by the intervening sequence, which is a composite protein with separate catalytic centers for protein splicing and DNA homing endonuclease activity. Experiments are in progress to study the structure and function of the catalytic center for protein splicing by the genetic dissection of the intervening sequence.

INTRODUCTION

Protein splicing is a novel form of gene expression in which an interrupted gene is translated into a protein with an intervening sequence, followed by a self-catalyzed protein processing step in which the intervening sequence is excised from the primary translation product to yield a mature, functional protein (Fig. 1). This process involves the precise excision of the intervening sequence (intein) and the ligation of the two flanking segments (N-extein and C-extein). At the present time, 36 proteins carrying putative self-splicing inteins are known (1), distributed across all three domains of life – bacteria, archaea, and eukarya. The protein splicing element, or intein, usually has a second function as a homing endonuclease, which plays a role in the insertion of its coding sequence into a host gene. In this respect, protein splicing elements resemble Group I introns (2), except that splicing occurs on the protein and not the RNA level, that splicing is catalyzed by a polypeptide rather than RNA, and that the catalytic sites for protein splicing and homing endonuclease activity reside on a single, bifunctional polypeptide. Whether protein splicing elements, like Group I introns, merely represent a clever machine for the propagation of selfish DNA or whether protein splicing plays a regulatory role in the activation of enzyme precursors is a fascinating evolutionary questions for which we don't yet have an answer. However, the observation that the mechanism of protein splicing is a paradigm for a diverse group of self-catalyzed protein rearrangements found in many organisms ranging from bacteria to higher vertebrates suggests that protein splicing may have originated as a regulatory mechanism.

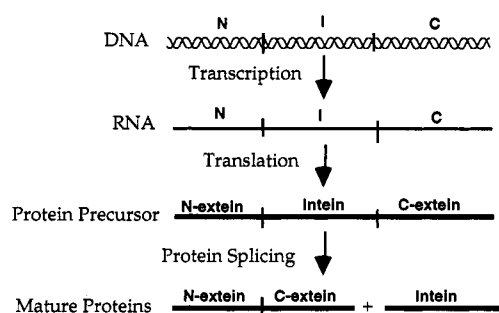


Fig. 1. Modulation of genetic information by protein splicing.

MECHANISM OF PROTEIN SPLICING

The characteristic presence of hydroxyl or thiol amino acids at the splice junctions served as an important clue in the elucidation of the mechanism of protein splicing by suggesting the involvement of ester intermediates. The study of these and other intermediates in protein splicing was made possible by the development of an *in vitro* splicing system based on the Psp Pol1 intein from the extreme thermophile, *Pyrococcus* sp. GB-D, inserted into a foreign context so as to allow expression of the unspliced precursor at low temperatures, its facile purification, and induction of splicing by raising the temperature (3). In these hybrid constructs, the maltose binding protein (MBP) was generally used as the N-extein so as to allow affinity purification of the unspliced precursor as well as the splicing intermediates on amylose resins. The observation that proteins containing an intein inserted between two foreign polypeptides can be made to splice merely by increasing temperature in the absence of any added cofactors demonstrated that all information essential for the splicing process resides in the intein and that catalysis of protein splicing by the intein requires no cofactors.

N-S or N-O acyl rearrangement

The Psp Pol1 intein is flanked by Ser residues at both splice junctions. In order to study the possible formation of an ester intermediate, site-directed mutagenesis was used to replace Ser at the upstream splice junction with Cys. In addition, Ser at the downstream splice junction was either deleted or replaced with Ala to assure that only the first step in protein splicing was being investigated (4,5). If this step involved an acyl rearrangement at the upstream splice junction, the *ser1cys* mutant intein should give rise to a thioester, which is subject to nucleophilic displacement by hydroxylamine or thiols. Indeed, treatment of fusion proteins of MBP and the *ser1cys* intein with neutral hydroxylamine at 37°C leads to rapid cleavage at the splice junction of the fusion protein and the production of the C-terminal hydroxamate derivative of MBP (4,5). Cleavage of the *ser1cys* mutant can also be effected by high concentrations of cysteine (5). Similar experiments were carried out with the VMA intein of *Saccharomyces cerevisiae*, which is naturally flanked by Cys residues, fused to the C-terminus of MBP and altered by modifying the downstream splice junction so as to block all but the first step in the protein splicing process (6). The unspliced precursor protein, ordinarily stable at neutral pH, undergoes rapid cleavage at the upstream splice junction in the presence of hydroxylamine or cysteine, leading to the formation of MBP derivatives with C-terminal hydroxamate or cysteine moieties. These results, supported by genetic studies that suggested an essential role for serine or cysteine residues at the upstream splice junction (5,7), constitute strong evidence that the first step in protein splicing involves the rearrangement of the peptide bond involving the serine or cysteine residue at the upstream splice junction to form an ester or thioester intermediate.

N-S or N-O acyl rearrangements in other self-catalyzed protein rearrangements

Nucleophilic attack by a Ser or Cys side chain on the adjacent peptide bond, which is seen in protein splicing, is paradigmatic of a diverse group of polypeptide rearrangements. As illustrated in Fig. 2, many types of proteins undergo self-catalyzed rearrangements at specific Ser or Cys residues to yield ester or thioester intermediates. The broad distribution and diverse biological functions of these proteins suggest that such rearrangements arose relatively early in evolution.

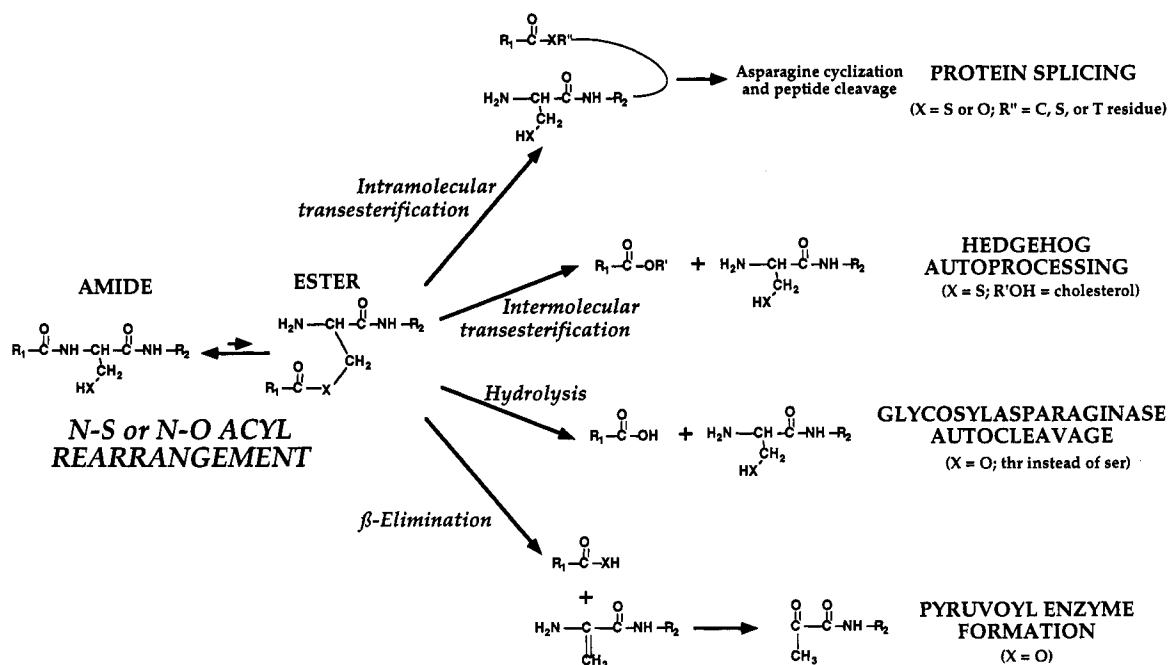


Fig. 2. Self-catalyzed protein rearrangements involving ester or thioester intermediates

Autoprocessing of hedgehog proteins

The hedgehog proteins are signaling proteins that function in developmental patterning of all multicellular animals, from nematodes to mammals. A typical example is the hedgehog protein from *Drosophila*. The 45-kDa protein precursor is secreted and then undergoes self-catalyzed processing involving polypeptide cleavage to yield a 25-kDa fragment with cysteine at its N-terminus (Hh-C) and a 20-kDa protein (Hh-N) whose carboxyl terminus is esterified with cholesterol (8-10). The esterified Hh-N autoprocessing product, which strongly interacts with cell surfaces owing to its C-terminal cholesterol moiety, is responsible for developmental signaling (9-11).

The autocleavage of hedgehog proteins occurs adjacent to a highly conserved cysteine residue. This cysteine is followed by a 12-amino acid sequence motif resembling the conserved A domain of self-splicing proteins (see Fig. 4) (12), suggesting mechanistic similarities between protein splicing and hedgehog autoprocessing (11). The mechanism of hedgehog autoprocessing was studied using a bacterially expressed modified hedgehog protein in which a hexahistidine sequence replaced most of the Hh-N domain, thereby allowing the facile purification of the precursor protein on Ni⁺⁺ resin. Evidence for a thioester intermediate came from the observation that hydroxylamine and thiols promote cleavage of the precursor, yielding the hydroxylamine or thiol adduct of Hh-N (9,11). Replacement of His-329, which resides in a conserved domain resembling the B domain of self-splicing proteins (see Fig. 4), blocks the autoprocessing reaction (11), suggesting that the B domain may play a role in the acyl-rearrangement leading to the ester intermediate both in protein splicing and hedgehog protein autoprocessing.

Autocleavage of glycosylasparaginase precursors

Glycosylasparaginase hydrolyses AspNHGlcNAc and related glycans, and its deficiency in humans leads to aspartylglycosaminuria, a genetic disorder of glycoprotein degradation, which is relatively common, especially in Finnish populations (13). The human and *Flavobacterium meningosepticum* glycosylasparaginases have a high degree of amino acid sequence and structural homology, both being composed of α and β subunits, which are encoded by a single gene. The primary gene product is a polypeptide precursor which is converted to the α and β chains by self-catalyzed cleavage adjacent to a

threonine residue (13-15). Replacement of this threonine residue (Thr-152) in bacterial glycosylasparaginase with amino acids other than serine or cysteine prevents autocleavage, whereas replacement with cysteine or serine greatly reduces the rate of autocleavage (16). The use of *thr152ser* and *thr152cys* mutant proteins fused to the C-terminus of MBP, in conjunction with the discovery that glycine and small L-amino acids severely inhibit autocleavage, made possible the purification of glycosylasparaginase precursors by affinity chromatography and the study of the autocleavage mechanism (16). Evidence for the involvement of an ester intermediate comes from the observations that the cleavage of the *thr152cys* mutant protein is promoted by the nucleophile hydroxylamine and inhibited by the thiol-blocking reagent iodoacetamide (16).

An interesting aspect of the autocleavage of glycosylasparaginase is that the N-terminal threonine residue generated in this process is essential for enzyme activity. In this respect, glycosylasparaginase resembles a group of other amidohydrolases, termed N-terminal nucleophile hydrolases, which include glutamine PRPP amidotransferase, penicillin acylase, and the proteasome. These enzymes have similar three-dimensional structures and their activity depends on an N-terminal serine, threonine, or cysteine residue that is generated by the cleavage of a precursor protein, probably through self-catalysis (17).

Autocleavage leading to pyruvoyl enzyme formation

The activity of a large group of pyridoxyl phosphate-independent bacterial amino acid decarboxylases and reductases depends on an N-terminal pyruvate prosthetic group. Through the pioneering work of E.E. Snell and coworkers (summarized in ref. 18), it was established that the enzyme-bound pyruvate was generated through the self-catalyzed cleavage of a protein precursor. In the case of the histidine decarboxylase from *Lactobacillus* 30a, processing of the proenzyme involves cleavage of the peptide bond between Ser-81 and Ser-82, coupled to the conversion of Ser-82 to an N-terminal pyruvate residue (18). The first suggestion that this process involves an ester intermediate came from isotopic labeling experiments, which showed that ^{18}O from the side chain of Ser-82 is transferred to the Ser-81 carboxyl group in the course of the cleavage reaction (19). The putative ester intermediate could be trapped as the hydroxamate after replacing the relevant serine residue with cysteine (20,21), owing to the much higher susceptibility of thioesters than oxygen esters to attack by nitrogen nucleophiles.

Resolution of the ester intermediates

The completion of the protein rearrangements summarized in Fig. 2 requires coupling the formation of the ester intermediate, characterized by an unfavorable equilibrium constant, to one or more thermodynamically favored reactions. As outlined in Fig. 2, these reactions are quite diverse and probably involve unrelated catalytic elements.

Protein splicing

This is by far the most complex of the protein rearrangements under consideration, involving three additional steps as shown in Fig. 3. The first of these is an attack on the ester by the side chain of the serine, threonine, or cysteine residue at the downstream splice junction to yield a branched ester intermediate. Evidence for a branched intermediate was obtained in the course of SDS-PAGE analysis of the products of *in vitro* splicing of chimeric proteins involving the Psp Poll intein, which revealed a slowly migrating species with the kinetic properties of a splicing intermediate (3). N-terminal sequencing showed the presence of two N-termini corresponding to those of the N-extein and the intein (3), and stability studies showed that the N-extein was linked by an alkali-labile bond, i.e. through ester linkage (22), consistent with the structure shown in Fig. 3. The equilibrium constant for the transesterification reaction that leads to the formation of the branched intermediate is about 1 if the esters involved are of the same type, i.e. thioester or oxygen ester, or about 50 if it involves the conversion of a thioester to an oxygen ester; however, a more important factor in the stabilization of the branched intermediate is that the ester bond is not immediately adjacent to a free α -amino group, thereby precluding the rapid cyclization that makes the linear ester intermediate not only thermodynamically but kinetically extremely unstable (22,23).

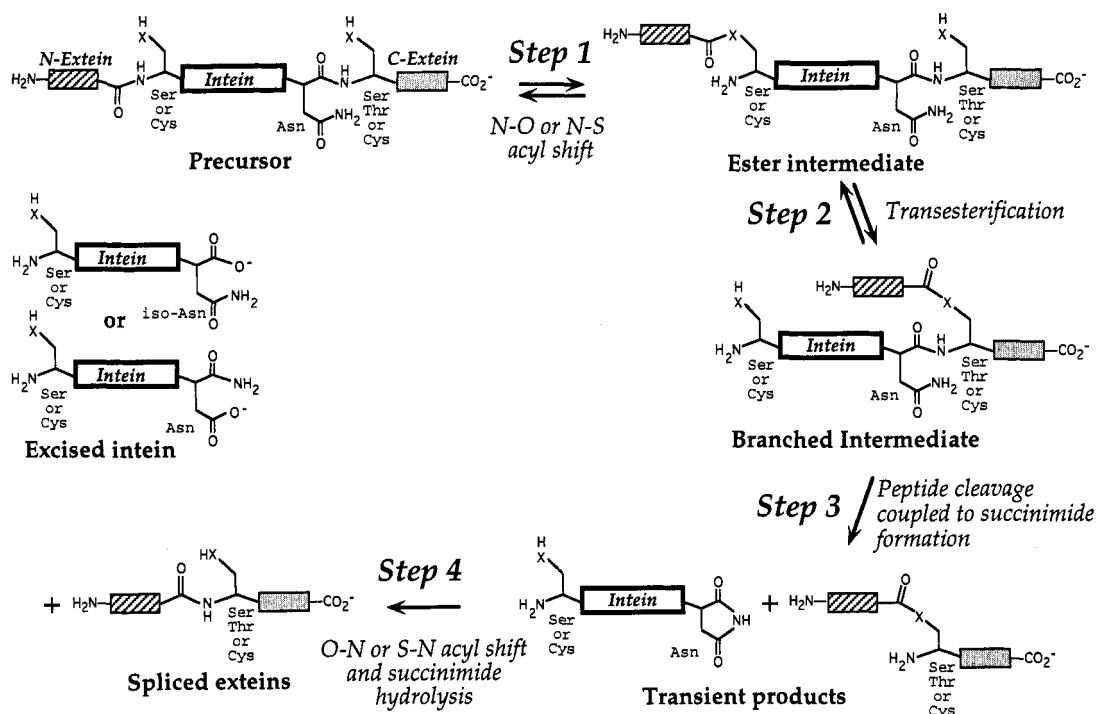


Fig. 3. Mechanism of protein splicing
(X = O or S).

The metastable branched intermediate then undergoes an essentially irreversible transformation, which involves the cyclization of the asparagine residue adjacent to the downstream splice junction and the concomitant cleavage of the peptide bond between the intein and the C-extein (Fig. 3). The C-terminal aminosuccinimide residue of the excised intein, which results from the cyclization of asparagine, was identified by mass spectrometry and comparison with synthetic model peptides, both with the *Pyrococcus* sp. GB-D intein (22,24) and the *S. cerevisiae* VMA intein (6). The conserved histidine that is adjacent to the asparagine residue may play a catalytic role in the cyclization reaction, for its replacement by other amino acids greatly retards that reaction without affecting branched intermediate formation (5,6).

The transient splicing products that are formed as a result of asparagine cyclization, i.e. the aminosuccinimide-terminated excised intein and the two exteins linked by an ester bond (Fig. 3), undergo a time-dependent rearrangement to stable molecules. The hydrolysis of the aminosuccinimide ring to asparagine or isoasparagine occurs with a half-time of 80 h at 25°C and pH 7.4 (24), whereas the ester bond linking the exteins rearranges to the thermodynamically more stable, normal peptide bond extremely rapidly, with a half-time of less than 1 min at neutral pH (23). The latter rearrangement, which is presumably uncatalyzed because it occurs after the intein has been excised, makes protein splicing essentially irreversible.

The identification of the intermediates and the sequence of steps in the protein splicing pathway was much assisted by the study of mutants with replacements of the conserved amino acid residues at the splice junctions (5,6). Replacement of the serine, threonine, or cysteine residues at the splice junctions with other hydroxyl or thiol amino acids modulates the rate of protein splicing, whereas replacement with other amino acids completely blocks the process, consistent with a role of these amino acids as nucleophiles in ester or thioester formation. Replacement of the asparagine at the downstream splice junction with aspartate or glutamine prevents splicing but has no effect on the formation of the linear or branched ester intermediate, consistent with a role in a late step in protein splicing. However, although the overall protein splicing process is blocked by the replacement of these essential amino acid residues, partial reactions, manifested by cleavage at either the upstream or the downstream splice junction, can still occur. For example, inteins with mutations affecting the amino acids at the downstream splice junctions can still undergo hydroxylamine- or thiol-dependent upstream cleavage, presumably through nucleophilic attack on the ester intermediate. On the other hand, mutant inteins with replacement of the

serine or cysteine residues at the upstream splice junction undergo cleavage at the downstream splice junction, suggesting that cyclization of the asparagine residue can occur independent of the earlier steps in the splicing process. In view of the relative autonomy of the steps in protein splicing illustrated in Fig. 3, it appears that the intein must have painstakingly evolved in its splicing context to assure that the relative rates of these steps are perfectly balanced so as to prevent abortive side reactions. It is interesting that the splicing of chimeric proteins, in which the intein is inserted into a foreign context, is usually accompanied by significant cleavage at both splice junctions (3,6), suggesting that transposition of an intein into a different environment tends to disturb the balance between the various splicing steps.

Other protein rearrangements

In the other types of protein rearrangement under consideration, the formation of the ester intermediate is followed by much simpler reactions. Hedgehog protein autoprocessing involves intermolecular transfer of the acyl moiety of the thioester intermediate to the 3β hydroxyl group of cholesterol, catalyzed by the C-terminal domain of the protein (10). In the autocleavage of glycosylasparaginase, the ester intermediate simply undergoes hydrolysis (16). In contrast, the autocleavage reaction that leads to the activation of pyruvoyl enzymes involves β -elimination at the serine residue rather than ester hydrolysis (19).

STRUCTURE AND FUNCTION OF PROTEIN SPLICING ELEMENTS

Protein splicing elements may be viewed as highly unusual enzymes. They are single polypeptide chains that catalyze three mechanistically quite unrelated reactions, are simultaneously substrate and catalyst, and have an active center that is composed of the two ends of the polypeptide chain. Another layer of complexity is added by the fact that, after excision, most inteins also function as homing endonucleases, catalyzing double-strand breaks at specific DNA sequences.

Some insight into the structural relationship between the protein splicing and homing endonuclease functions of inteins can be gleaned by comparing conserved amino acid sequence motifs of a typical intein (the RecA intein from *Mycobacterium tuberculosis*) (25) with those from an intein lacking a homing endonuclease domain (the DnaB intein from *Porphyria purpurea*) (26) and a homing endonuclease encoded by a Group I intron (from *Desulfurococcus mobilis*) (27) that bears no relationship to protein splicing. As shown in Fig. 4, the *D. mobilis* homing endonuclease and the central portion of the *M. tuberculosis* intein have in common two conserved sequence motifs (C and E in the terminology of Pietrokovski (28), also known as LAGLI-DADG motifs), whereas the non-homing *P. purpurea* intein and the peripheral portions of the *M. tuberculosis* intein share four conserved sequence motifs (A, B, F, and G in Pietrokovski's terminology (28)). This comparison suggests that inteins are essentially chimeric proteins, with the peripheral segments organized into a domain for the catalysis of protein splicing and the central portion into a separate homing endonuclease domain.

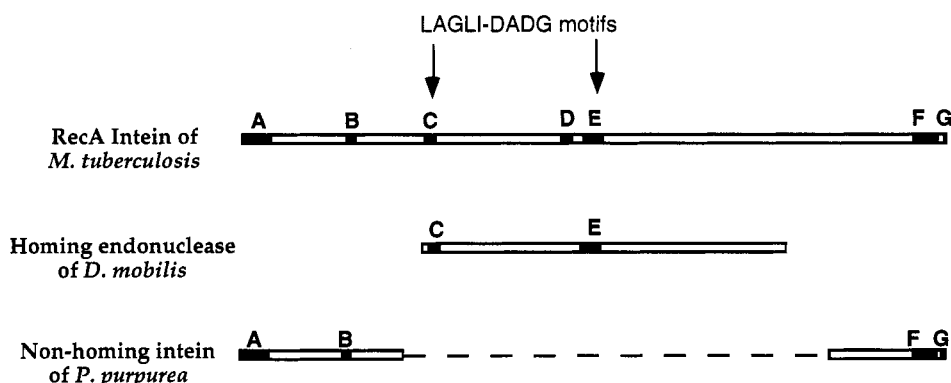


Fig. 4. Alignment of the RecA intein of *Mycobacterium tuberculosis*, the homing endonuclease of *Desulfurococcus mobilis*, and the DnaB intein of *Porphyria purpurea*. The major consensus sequences described by Pietrokovski (28) are indicated by capital letters.

Genetic Dissection of the *M. tuberculosis* Intein

In order to test the hypothesis that inteins consist of two independent catalytic centers, one specializing in protein splicing and the other in DNA recognition and cleavage, we have begun to dissect the *M. tuberculosis* RecA intein by genetic means. As our experimental system, the intein (U) was inserted between maltose binding protein (M) as the N-extein and the chitin binding domain from *Bacillus circulans* (B) as the C-extein, allowing characterization of precursors and splicing products by affinity chromatography on amylose and chitin resins (29). To determine whether any segment of the homing endonuclease domain is essential for protein splicing, the central portion of the intein gene encoding 170 amino acid residues, including both LAGLI-DADG homing endonuclease motifs, was replaced by a short linker region, and the ability to undergo protein splicing upon expression in *E. coli* was determined (K. Shingledecker, S.-q. Jiang & H. Paulus, unpublished). As shown in Fig. 5, both the wild-type chimeric precursor protein (MUB) and the corresponding deletion protein lacking the entire homing endonuclease domain underwent efficient splicing, as indicated by the accumulation of the fused exteins, (MB). This observation establishes unambiguously that the protein splicing active center is separate from the homing endonuclease and constitutes an autonomous structural domain.

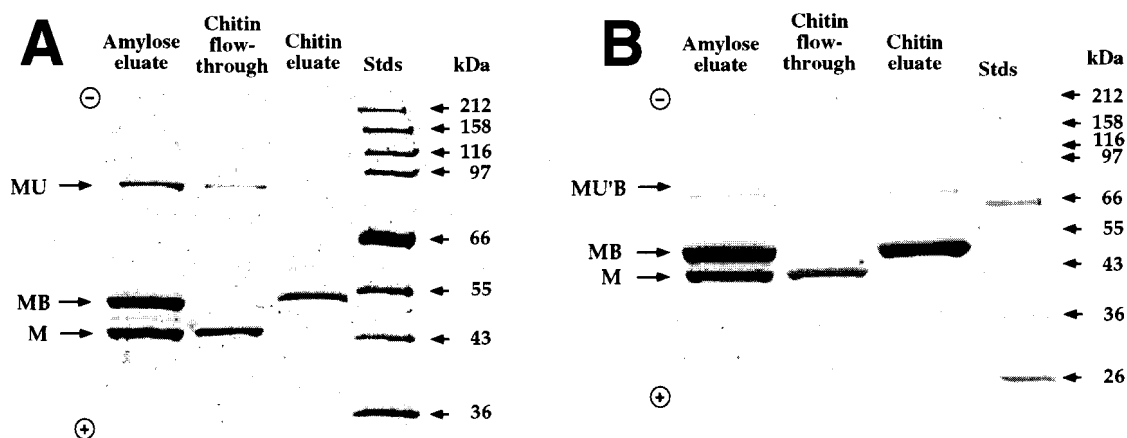


Fig. 5. Effect of deletion of the homing endonuclease domain on the ability of the *M. tuberculosis* RecA intein to undergo protein splicing. The coding sequence of the RecA intein (U) was inserted in-frame between that of the maltose binding protein (M) and the chitin binding domain (B), analogous to the constructs involving the PI-SceI intein described earlier (29), and expressed in *E. coli* at 25°C. Cell extracts were subjected to successive cycles of affinity chromatography on amylose and chitin resin as described (29) and the column eluates were monitored by SDS-PAGE, followed by protein staining with Coomassie blue. *A*, Analysis of a construct carrying the intact RecA intein; *B*, analysis of a construct in which amino acid residues #105 - 276 of the RecA intein were replaced by the linker sequence AGTSGSGIRR.

As a first step in the detailed study of the enzymology of protein splicing, we have dissected the intein into two segments that can mediate splicing in a cooperative manner. This was accomplished by inserting tandem in-frame translation termination and initiation codons into the homing endonuclease segment of the intein coding region, so that expression in *E. coli* should yield two proteins, consisting of the N-terminal portion of the intein fused to the C-terminus of the maltose binding protein (MU') and the C-terminal portion of the intein fused to the N-terminus of the chitin binding domain (U''B) (K. Shingledecker, S.-q. Jiang & H. Paulus, unpublished). Analysis of the expression products in *E. coli* by successive binding to amylose and chitin resins revealed not only the presence of the precursor protein (MU') but also spliced product (MB), suggesting that *trans*-splicing had occurred between the separately synthesized N- and C-terminal segments of the fusion protein (Fig. 6). Similar results were obtained with N- and C-terminal segments of the intein from which the entire homing endonuclease segment had been deleted (data not shown). This demonstrates that a functional protein splicing center can be reconstituted *in vivo* by the association of separately synthesized polypeptide chains. Experiments are now in progress to reconstitute active protein splicing elements *in vitro* from the purified N- and C-terminal fragments. The success of such experiments should open the way for studying the interaction of the two halves of the protein splicing active center so as to define the structural elements and catalytic groups essential for protein splicing. It should also provide a useful tool for ligating any two polypeptides by cloning them as fusion proteins with the two halves of an intein.

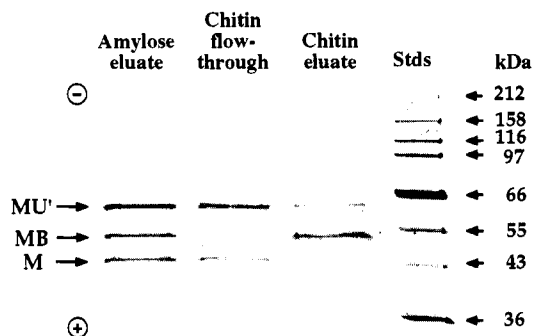


Fig. 6. *Trans*-splicing of proteins linked to separate N- and C-terminal fragments of the *M. tuberculosis* RecA intein. The coding sequence of the RecA intein (U) was inserted in-frame between that of the maltose binding protein (M) and the chitin binding domain (B) as described in the legend to Fig. 5, except that the intein was interrupted by tandem translation stop and start codons inserted in-frame, together with a ribosome binding site, between residues #105 and 109 of the RecA intein. Cell extracts were analysed by affinity chromatography followed by SDS-PAGE as in Fig. 5.

REFERENCES

1. Perler, F. B., Olsen, G. J., and Adam, E. *Nucl. Acids Res.* **25**, 1087-1093 (1997).
2. Lambowitz, A. M., and Belfort, M. *Annu. Rev. Biochem.* **62**, 587-622 (1993).
3. Xu, M.-Q., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Perler, F. B. *Cell* **75**, 1371-1377 (1993).
4. Shao, Y., Xu, M.-Q., and Paulus, H. *Biochemistry* **35**, 3810-3815 (1996).
5. Xu, M.-Q., and Perler, F. B. *EMBO J.* **15**, 5146-5153 (1996).
6. Chong, S., Shao, Y., Paulus, H., Benner, J., Perler, F. B., and Xu, M. Q. *J. Biol. Chem.* **271**, 22159-22168 (1996).
7. Cooper, A. A., Chen, Y.-J., Lindorfer, M. A., and Stevens, T. H. *EMBO J.* **12**, 2575-2583 (1993).
8. Lee, J. J., Ekker, S. C., Kessler, D. P., Porter, J. A., Sun, B. I., and Beachy, P. A. *Science* **266**, 1528-1537 (1994).
9. Porter, J. A., Ekker, S. C., Park, W. J., Kessler, D. P. v., Young, K. E., Chen, C. H., Ma, Y., Woods, A. S., Cotter, R. J., Koonin, E. V., and Beachy, P. A. *Cell* **86**, 21-34 (1996).
10. Porter, J. A., Young, K. E., and Beachy, P. A. *Science* **274**, 255-259 (1996).
11. Porter, J. A., Kessler, D. P., Ekker, S. C., Young, K. E., Lee, J. J., Moses, K., and Beachy, P. A. *Nature* **374**, 363-366 (1995).
12. Koonin, E. V. *Trends Biochem. Sci.* **20**, 141-142 (1995).
13. Tikkanen, R., Riikonen, A., Oinonen, C., Rouvinen, J., and Peltonen, L. *EMBO J.* **15**, 2954-2960 (1996).
14. Tarentino, A. L., Quinones, G., Hauer, C. R., Changchien, L. M., and Plummer, T. H. *Arch. Biochem. Biophys.* **316**, 399-406 (1995).
15. Guan, C., Li, P., Riggs, P. D., and Inouye, H. *Gene* **67**, 21-30 (1987).
16. Guan, C., Cui, T., Rao, V., Liao, W., Benner, J., Lin, C. L., and Comb, D. *J. Biol. Chem.* **271**, 1732-1737 (1996).
17. Brannigan, J. A., Dodson, G., Duggleby, H. J., Moody, P. C. E., Smith, J. L., Tomchick, D. R., and Murzin, A. G. *Nature* **278**, 416-419 (1995).
18. van Poelje, P. D., and Snell, E. E. *Annu. Rev. Biochem.* **59**, 29-59 (1990).
19. Recsei, P. A., Huynh, Q. K., and Snell, E. E. *Proc. Natl. Acad. Sci. USA* **80**, 973-977 (1983).
20. van Poelje, P. D. *Ph.D. dissertation*, University of Texas at Austin (1988).
21. Vanderslice, P., Copeland, W. C., and Robertus, J. D. *J. Biol. Chem.* **263**, 10583-10586 (1988).
22. Xu, M.-Q., Comb, D. G., Paulus, H., Noren, C. J., Shao, Y., and Perler, F. B. *EMBO J.* **13**, 5517-5522 (1994).
23. Shao, Y., and Paulus, H. *J. Peptide Res.*, in press (1997).
24. Shao, Y., Xu, M.-Q., and Paulus, H. *Biochemistry* **34**, 10844-10850 (1995).
25. Davis, E. O., Sedgwick, S. G., and Colston, M. J. *J. Bacteriol.* **173**, 5653-5662 (1991).
26. Reith, M. E., and Munholland, J. *Plant Mol. Biol. Rep.* **13**, 333-335 (1995).
27. Dalgaard, J. Z., Garrett, R. A., and Belfort, M. *Proc. Natl. Acad. Sci. USA* **90**, 5414-5417 (1993).
28. Pietrokovski, S. *Protein Science* **3**, 2340-2350 (1994).
29. Chong, S., Mersha, F. B., Comb, D. G., Scott, M. E., Landry, D., Vence, L. M., Perler, F. B., Benner, J., Kucera, R., Hirvonen, C. A., Pelletier, J. J., Paulus, H., and Xu, M. Q. *Gene*, in press (1997).