MINIREVIEW

⌘ *Author's Choice*

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 21, pp. 14498–14505, May 23, 2014

Protein Splicing: How Inteins Escape from Precursor Proteins*

Published, JBC Papers in Press, April 2, 2014, DOI 10.1074/jbc.R113.540310

Kenneth V. Mills[‡], Margaret A. Johnson[§], and Francine B. Perler[¶1]

From the [‡]Department of Chemistry, College of the Holy Cross, Worcester, Massachusetts 01610, the [§]Department of Chemistry, University of Alabama at Birmingham, Birmingham, Alabama 35294, and [¶]New England Biolabs, Inc., Ipswich, Massachusetts 01938

Inteins are nature's escape artists; they facilitate their excision from flanking polypeptides (exteins) concomitant with extein ligation to produce a mature host protein. Splicing requires sequential nucleophilic displacement reactions catalyzed by strategies similar to proteases and asparagine lyases. Inteins require precise reaction coordination rather than rapid turnover or tight substrate binding because they are single turnover enzymes with covalently linked substrates. This has allowed inteins to explore alternative mechanisms with different steps or to use different methods for activation and coordination of the steps. Pressing issues include understanding the underlying details of catalysis and how the splicing steps are controlled.

The first intein sequence was published 25 years ago (1). In those early days of gene analysis, it was hard to decipher why the *Saccharomyces cerevisiae Sce VMA1* vacuolar ATPase gene was so large. Two years later two groups showed that a section of the *VMA1* gene was absent in the mature ATPase (2, 3). They challenged existing fundamental beliefs about gene expression by predicting that an internal section of this protein was removed by protein splicing instead of RNA splicing and that a single gene encoded two stable proteins: the host protein (extein) and the intervening protein (intein) (4). All attempts to demonstrate RNA splicing failed. *In vivo* time courses suggested protein splicing (5, 6), which was definitively established as a new method of gene expression when the elusive precursor protein was isolated by cloning a *Pyrococcus* species DNA polymerase intein between two unrelated proteins, resulting in temperature-dependent splicing (7). This first example of *in vitro* splicing revealed important mechanistic insights: splicing occurs when the intein and the first C-extein residue are embedded in a heterologous host protein, inefficient splicing can result in off-pathway single splice site cleavage (Fig. 1), splicing can be controlled, and splicing proceeds through a slowly migrating branched intermediate with two N termini.

We consider inteins to be single turnover enzymes because they use the same strategies as classical enzymes to perform catalysis (8). Splicing occurs in the absence of any known cofactor, chaperone, or energy source. All that is required is proper folding of the intein in the precursor to align nucleophilic residues and residues that assist catalysis (Fig. 2), leading some to call inteins nature's escape artists (9). Because inteins utilize groups of similar nucleophiles, subtle variations in the reactivity of these amino acids require different sets of assisting residues. As a result, some residues facilitating catalysis may still be unknown.

There are three classes of inteins based on sequence signatures and splicing mechanisms (10). The standard class 1 intein splicing mechanism (Fig. 3) consists of 1) an acyl rearrangement to convert the N-terminal splice site peptide bond from an amide to a (thio)ester, 2) a transesterification to form a branched intermediate, 3) Asn cyclization resolving the branched intermediate by cleaving the C-terminal splice site, and 4) a second acyl shift to form an amide bond between the ligated extein segments (5–7, 11–16). Off-pathway cleavage occurs when coordination of the steps is perturbed by mutation or by expression between foreign exteins (Fig. 1). This may result from an increase in the cleavage rate at that splice site, a decrease in the reaction rate of another step, or interference with a mechanism-linked conformational change required to promote a downstream step.

As the number of sequenced inteins increased, sequence alignments revealed four splicing motifs termed blocks A, B, F, and G (Fig. 2) (17–20). Although not conserved in their entirety, several positions in each motif contain highly conserved groups of similar amino acids. The nucleophiles for each step are: Cys^{1} or Ser^{1} in step 1; Cys^{+1}, Ser^{+1}, or Thr^{+1} in steps 2 and 4; and the intein C-terminal $Asn^{G:7}$ in step 3 (see Fig. 2 for residue nomenclature). Known assisting residues include positions 7 and 10 in block B ($Thr^{B:7}$ and $His^{B:10}$), the intein penultimate $His^{G:6}$, and the less conserved positions 4 and 13 in block F (Fig. 2). Position F:4 is most commonly Asp, followed by Cys and then Trp, and F:13 is most commonly His (10, 20).

Inteins come in many flavors. Most inteins are large chimeras containing both a splicing domain and the same type of endonucleases that mediate intron mobility (Fig. 2) (20–22). Other inteins are naturally occurring mini-inteins that are as small as 134 residues and lack an endonuclease domain (20). Studies of both native and engineered mini-inteins helped define the intein splicing domain (20, 23–27). Intein genes may also be split between motifs B and F; however, the expressed precursor protein fragments rapidly assemble to splice in *trans* by the same mechanisms used in *cis*-splicing inteins (28–32). Both naturally occurring and engineered split inteins have found great utility in biotechnology applications (33). Intein splicing domains may have been derived from ancient enzymes because they are small and are closely related by structure, conserved motifs, and enzymatic activities to Hedgehog autoprocessing domains, which activate essential signaling proteins for metazoan development (34, 35). Inteins are also related to bacterial intein-like (BIL)[2] domains (36–39).

* This work was supported by the National Science Foundation under Grant MCB-1244089 (to K. V. M.), by a Henry Dreyfus Teacher-Scholar Award (to K. V. M.), by New England Biolabs (to F. B. P.), and by University of Alabama at Birmingham startup funding (to M. A. J.). This is the second article in the Thematic Minireview Series "Inteins."

⌘ *Author's Choice*—Final version full access.

[1] To whom correspondence should be addressed: New England Biolabs, 240 County Rd., Ipswich, MA 01938-2723. Tel.: 978-380-7326; E-mail: perler@neb.com.

[2] The abbreviations used are: BIL, bacterial intein-like; Pol, polymerase.

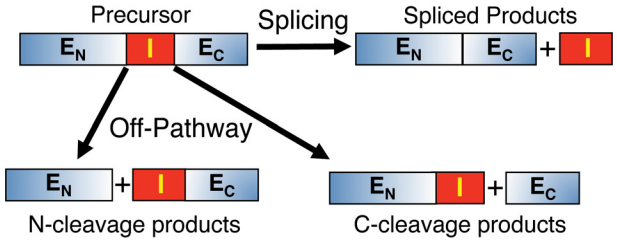

FIGURE 1. **Potential intein reactions.** Protein splicing results in ligation of the N-extein (E_N) and C-extein (E_C), as directed by the intein (*I*). When inteins are mutated or inserted in heterologous contexts, off-pathway reactions can occur resulting in N-terminal, C-terminal, or double cleavage products that are unable to splice. Off-pathway N-terminal cleavage can occur in both the linear and the branched (thio)ester intermediates. Off-pathway C-terminal cleavage occurs when cyclization of the intein C-terminal residue precedes branch intermediate formation.

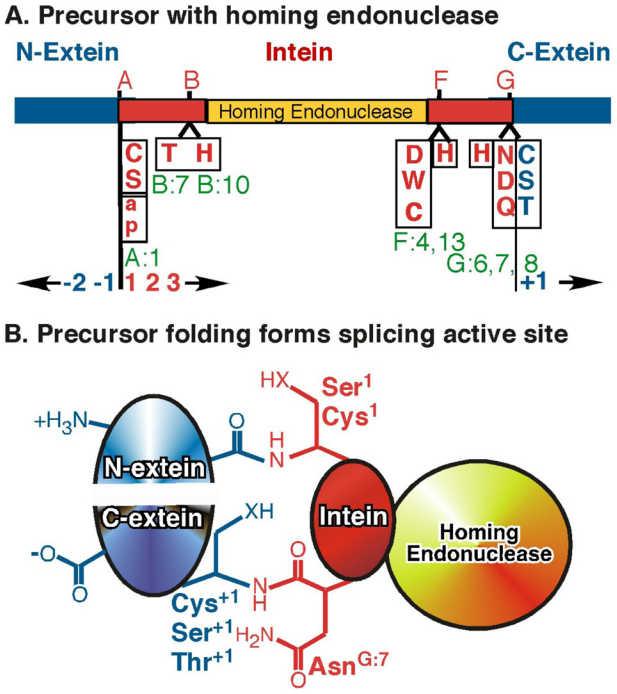

FIGURE 2. **Precursor domains and conserved motifs.** *A*, a precursor with an intein containing a homing endonuclease domain (*gold*) is depicted with intein splicing domain motifs (*red*) listed above and conserved residues that participate in catalysis listed below. Residues in intein motifs are numbered based on their position within each motif (*green*) as defined in InBase (20). Residues specific to class 2 or 3 inteins are in *lowercase*, and only a subset of residues found at A:1 is depicted. Motifs A, B, F, and G have also been called N1, N2, C2, and C1, respectively (17–20). Motifs C, D, E, and H are specific to certain homing endonucleases and are not shown. To simplify discussion of inteins in various precursors, residues in each part are numbered independently. Intein residues are numbered from the N to C terminus beginning with 1. Residues in the N- and C-exteins (*blue*) are numbered from the splice site outwards and include a *minus sign* for N-extein and a *plus sign* for C-extein residues. *B*, folding of the precursor forms the intein active site and initiates protein splicing. Homing endonuclease domains in larger inteins fold separately from the intein and extein domains. Association of extein fragments can influence precursor folding and active site architecture. *X* represents an oxygen or a sulfur atom.

Establishing a universal mechanism for all inteins at the atomic level is unlikely because the methods used to promote each rearrangement and the roles played by the assisting residues vary. Instead, we will discuss both the canonical and the alternative protein splicing mechanisms, and general strategies for catalysis and coordination of the steps. Detailed mechanism reviews are available (8, 9, 40), and intein structure, molecular dynamics, evolution, and applications are covered in companion reviews in this series (21, 33, 41).

The Class 1 Splicing Mechanism: Step 1

Conversion of amide bonds to enzyme-linked thioesters or esters, as in step 1 of protein splicing, is a common method of catalysis used by proteases and autoprocessing enzymes including glycosyltransferases and pyruvoyl enzymes (8, 9, 40). Although protein splicing employs a series of bond rearrangements rather than the bond cleavage facilitated by proteases, inteins use similar strategies to destabilize the peptide bond to favor (thio)ester formation, including catalytic bond strain, general acid/base catalysis, and an oxyanion hole to stabilize the tetrahedral intermediate.

Most residues involved in catalysis assemble near the center of the disk-shaped HINT (Hedgehog-intein) fold of the intein splicing domain (34). For example, residues near the N-terminal scissile bond include $Thr^{B:7}$ and $His^{B:10}$ in a type I β-turn and $Asp^{F:4}$ in a β-strand. Experimental data show that $Thr^{B:7}$, $His^{B:10}$, and residue F:4 affect N-terminal splice site reactions, as do flanking extein residues (13, 14, 20, 42–51). $His^{B:10}$ is the most conserved intein residue (20). Of the two inteins without $His^{B:10}$, one is a degraded pseudogene (52), and the other (the *Thermococcus kodakaraensis* Tko CDC21-1 intein) uses Lys^{58} to activate the N-terminal splice site by possibly stabilizing the initial N–S acyl shift tetrahedral intermediate (53). Lys^{58} lies outside the conserved intein motifs (53) and is one residue beyond a newly identified position (22 residues past $His^{B:10}$) that potentially activates the N-terminal nucleophile (40). N-extein residues were shown to influence the equilibrium position between amide and ester in the Sce VMA intein (44) and to affect N-terminal reactions by van der Waals contacts with *Pyrococcus horikoshii* Pho RadA intein residues (46). The *Nostoc punctiforme* Npu DnaE intein +2 C-extein residue also affects splicing, possibly by filling space at the active site to optimally align catalytic residues (48, 49).

Some inteins distort the N-terminal scissile bond generating catalytic strain to accelerate step 1. Adjacent extein residues and block B residues help form this strained local conformation, as evidenced by both structural and biochemical studies. A crystal structure of the Sce VMA intein displays bond angle distortions near the N-terminal splice junction (54). The *Mycobacterium xenopi* Mxe GyrA intein crystal structure has a *cis*-peptide bond linking the N-extein and intein, and NMR data suggest a lack of amide bond resonance that is resolved when $His^{B:10}$ is mutated (42, 55). $His^{B:10}$ is in hydrogen bond distance to the amide nitrogen of the N-terminal scissile bond in several inteins, suggesting that it plays a role in coordinating the scissile bond (54–57). A similar role was observed for $Thr^{B:7}$ in a *Synechocystis* species Ssp DnaE intein structure (58).

A second strategy is to accelerate the rate at which the amide-ester equilibrium is reached by activating the N-terminal nucleophile via thiol deprotonation (experimentally detected as a lower pK_a), stabilizing the tetrahedral intermediate, and/or influencing how the tetrahedral intermediate is resolved. Likely contributors again include $Thr^{B:7}$, $His^{B:10}$, and $Asp^{F:4}$. A close look at the *Mycobacterium tuberculosis* Mtu RecA intein pro-

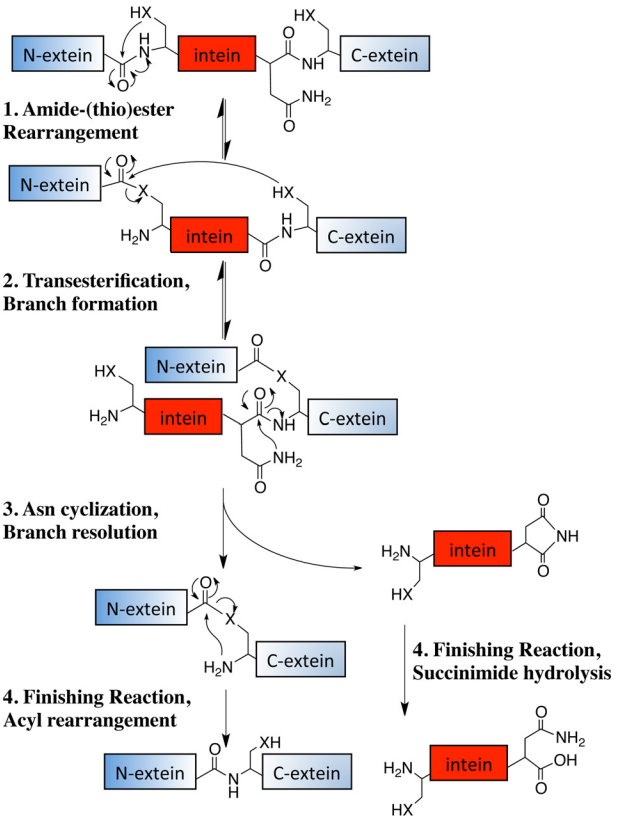

FIGURE 3. **The intein-mediated class 1 protein splicing mechanism.** Class 1 inteins with a C-terminal Asn and a Cys, Ser, or Thr at the first position in both the intein and the C-extein splice using the standard four-step protein splicing mechanism depicted in this figure. Inteins with C-terminal Glu, Gln, or Asp use this same mechanism except for Glu, Gln, or Asp cyclization in step 3, although other mechanisms are possible. Succinimide hydrolysis can also produce iso-Asn. *X* represents an oxygen or a sulfur atom. For clarity, tetrahedral intermediates and residues facilitating each step are omitted. Although the definition of an intein is the excised sequence (4), for brevity we will include the C-extein nucleophile when discussing mechanisms.

vides several lines of evidence to support these strategies. NMR and quantum mechanical/molecular mechanics studies suggest that $His^{B:10}$ may deprotonate the thiol of the N-terminal Cys^{1} to drive formation of the tetrahedral intermediate and then donate the proton to the Cys^{1} α-amino group to resolve the tetrahedral intermediate as a thioester (59). These roles are supported by changes in the pK_a of $His^{B:10}$ from neutral to acidic during splicing (60). Similarly, a Mtu RecA intein structure shows $Asp^{F:4}$ in position to hydrogen bond to the thiol of Cys^{1} to deprotonate this nucleophile (61). Moreover, the pK_a of the Mtu RecA intein $Asp^{F:4}$ is elevated and the pK_a of Cys^{1} is lower than normal, but both pK_a values return to normal when either residue is mutated (59). Taken together, these studies support a proposed proton transfer network in the Mtu RecA intein that assists deprotonation of the Cys^{1} nucleophile and the forward resolution of the tetrahedral intermediate (59–61). However, studies of the *Synechocystis* sp. PCC6803 Ssp DnaB intein with an unnatural N-terminal residue indicate that activation of this nucleophile is not essential for linear thioester formation (62), emphasizing the unique active sites and catalytic strategies utilized by individual inteins.

The Class 1 Splicing Mechanism: Step 2

The second step has proven the most challenging to study as it is difficult to isolate branched intermediates. Mutations that should result in accumulation of branched intermediates often result in decay to N-terminal cleavage products, especially when a thioester linkage is present (13). Ester-linked branched intermediate formation is reversible, which can result in accumulation of precursor rather than intermediate (7). Several studies suggest that the intein promotes step 2 by controlling the protonation state of the +1 nucleophile. For example, the pK_a of Cys^{+1} in the Mtu RecA intein is depressed to 5.8, increasing its nucleophilicity at physiological conditions (63). Furthermore, quantum mechanical simulations suggest that Cys^{+1} in the Mtu RecA intein may be deprotonated by $Asp^{F:4}$ and that this deprotonation may be driven in part to stabilize the positive charge on the α-amino group of Cys^{1} in the linear thioester intermediate (64). Step 2 is strictly coupled to step 1 in class 1 inteins, although the exact mechanism has yet to be determined (13–15, 65). It is possible that linear thioester formation removes elements that are masking the reactive thiol of the +1 residue or induces a conformational change to align active site residues for transesterification (see below).

The Class 1 Splicing Mechanism: Step 3

Evidence for the third step of splicing includes loss of C-terminal splice site cleavage after mutation of the intein C-terminal $Asn^{G:7}$ and the detection of excised inteins with C-terminal succinimide residues (11–15). The intein must catalyze Asn cyclization, because in other systems it results in side-chain deamidation rather than peptide bond cleavage (66), and computational modeling suggests very high energy barriers in non-catalyzed models of cleavage by Asn cyclization (67).

Several strategies have been proposed for enzymatic activation of step 3 including three coupled modes of catalysis: 1) $His^{F:13}$ increases the nucleophilicity of the C-terminal $Asn^{G:7}$ by deprotonation, 2) the tetrahedral intermediate is stabilized by charged $His^{F:13}$ and $His^{G:6}$ residues, and 3) the electrophilicity of the backbone amide may be increased by $His^{G:6}$ (55, 57, 67–69). Alternatively, given that C-terminal cleavage is favored at low pH (65, 70, 71), protonation of the backbone amide nitrogen of the scissile peptide bond may have precedence over deprotonation of the Asn side-chain amide (72). Separate studies suggest two other modes of catalysis: change in the local environment near the scissile bond that depends on branched ester formation (69) and destabilization of the scissile bond by a polarizable adjacent C-extein residue (73).

$His^{F:13}$ and $His^{G:6}$ are not required for Asn cyclization in all inteins (28, 56, 74, 75). Mutation of $His^{F:13}$ in a class 2 intein had no effect (76), and ~5% of functional inteins have an alternate G:6 residue (20, 28, 56, 74, 75). Splicing can be enhanced by "reverting" back to $His^{G:6}$ in some inteins, whereas a $His^{G:6}$ actually impairs splicing in other inteins (28, 74, 75, 77). These differences may reflect different positions along the evolutionary path to overcoming loss of $His^{G:6}$.

Some inteins lacking $Asn^{G:7}$ have similar residues (Asp and Gln) that can undergo cyclization to cleave the C-terminal splice site (20, 71, 78–80). For both the *Pyrococcus abyssi* and

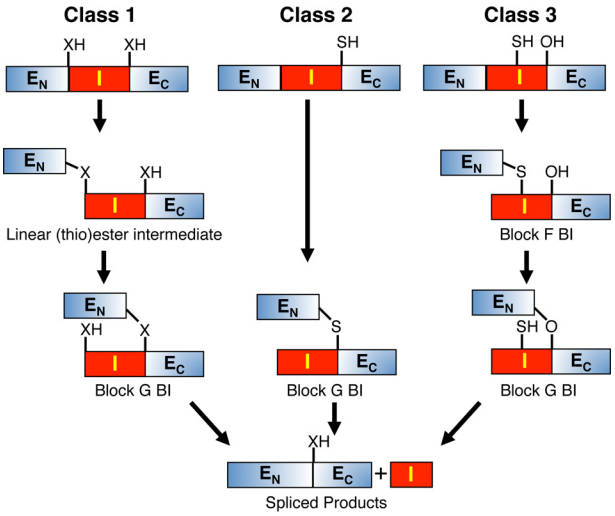

FIGURE 4. **Variations in splicing mechanisms.** Inteins missing the standard N-terminal nucleophile use various strategies to get to the same block G branched intermediate formed after step 2 in class 1 inteins. Class 2 inteins form the block G branched intermediate after direct attack on the amide bond at the N-terminal splice site by Cys^{+1}. Class 3 inteins first form a block F branched intermediate with $Cys^{F:4}$ as the branch point and then transfer the N-extein to the +1 residue to form the block G branched intermediate. Once the block G branched intermediate is formed, class 2 and class 3 inteins follow the same steps (3 and 4) to complete splicing as in class 1 inteins. Abbreviations used are: E_N, N-extein; E_C, C-extein; *I*, intein; *BI*, branched intermediate; *X*, an oxygen or a sulfur atom.

the *Methanoculleus marisnigri* Pol II inteins, splicing with a C-terminal Gln is slow, but is improved with substitution to Asn (71, 78, 79). On the other hand, the *Chilo* iridescent virus ribonucleotide reductase (CIV RNR) intein can splice with a native C-terminal Gln more efficiently than with Asn (80). As in the case of inteins lacking $His^{G:6}$, it is likely that these variant inteins represent different stages in evolving optimal activity after an initial mutation removed a catalytically important residue. Retaining a slow or inefficient step 3 may not be detrimental when it does not lead to off-pathway N-terminal cleavage.

The Class 1 Splicing Mechanism: Step 4

Step 4 consists of two finishing steps, neither of which is necessarily catalyzed by the intein. The intein C-terminal aminosuccinimide is slowly hydrolyzed to Asn or iso-Asn (11–13), and the (thio)ester linking the extein segments reverts to the amide. Experiments with model peptides demonstrate that the rate of conversion from a (thio)ester to an amide is faster than the overall rate of splicing (81). This final acyl shift is thermodynamically favorable and is not influenced by the presence of the intein (11, 69).

Variant Splicing Mechanisms: Class 2 Inteins, Class 3 Inteins, and BILs

The robustness of intein-mediated protein splicing is illustrated by the array of acceptable modifications to the standard four-step mechanism. BILs lack the C-extein +1 nucleophile and are therefore unable to form the block G branched intermediate (36–38). Both class 2 and class 3 inteins can still splice, although they lack a Ser^1 or Cys^1 nucleophile and are thus unable to form the linear (thio)ester intermediate (Fig. 4).

To date, all class 2 inteins are orthologs of the *Methanococcus jannaschii Mja KlbA* intein (20, 77). They all have $Ser^{G:6}$ instead of the more common $His^{G:6}$ and Ala^1. Class 2 inteins bypass the first step of splicing with Cys^{+1} directly attacking the N-terminal splice site amide bond, resulting in the same block G branched intermediate as in class 1 inteins. Thereafter, they follow the standard splicing pathway (76, 77). How class 2 inteins activate the N-terminal splice site for direct attack by Cys^{+1} and why class 1 inteins cannot (13–15, 65) remains to be fully determined. A possible explanation comes from the NMR structure of the Mja KlbA intein where a slight widening of its active site as compared with class 1 inteins allows the Cys^{+1} nucleophile to approach the N-terminal splice site without formation of a linear (thio)ester intermediate (76). The same three residues ($Thr^{B:7}$, $His^{B:10}$, $Asp^{F:4}$) that activate class 1 N-terminal splice sites are also required in class 2 inteins (76, 77). Mutation of $His^{B:10}$ and $Asp^{F:4}$ block splicing and drastically reduce both N-terminal and C-terminal cleavage (76, 77). $Thr^{B:7}$ and $His^{B:10}$ are positioned near the backbone nitrogen of Ala^1, and modeling of an active conformation showed $Asp^{F:4}$ hydrogen bonding to the Cys^{+1} thiol to possibly activate it by deprotonation (76).

Class 3 inteins have a remarkable mechanism that includes two branched intermediates (Fig. 4) (10, 82, 83). Cys is conserved at position F:4 in all class 3 inteins. It directly attacks the N-terminal splice site amide bond, resulting in the N-extein linked by a thioester to $Cys^{F:4}$, yielding a block F branched intermediate. Next, the N-extein is transferred to the side chain of Cys^{+1} to form a standard block G branched intermediate. Tori *et al.* (10) hypothesized that the position of $Cys^{F:4}$ in the intein active site allows it to substitute for the loss of the intein N-terminal nucleophile, in conjunction with two other positions that are conserved in all class 3 inteins. Monophyletic class 3 inteins appear to have arisen in a phage gene and spread to helicase genes in numerous organisms (52, 82). Thus the evidence suggests that both class 2 and class 3 inteins arose from single events.

There are at least two classes of BILs (38). Type A BILs have C-terminal His-Asn residues like inteins and can splice, although cleavage products dominate; type B BILs lack similarity to intein block G and catalyze splice site cleavage reactions uncoupled to splicing (38). The proposed mechanism for type A BIL splicing involves formation of a thioester bond at the BIL N terminus (intein step 1) and cleavage at the BIL C terminus by Asn cyclization (intein step 3). The free amino group on the C-terminal fragment attacks the N-terminal thioester bond to ligate the fragments flanking the BIL (36–38). The *Magnetospirillum magnetotacticum* BIL did not splice until Tyr^{+1} was mutated to Cys (39), suggesting that it is still tuned to act like an intein. It is likely that BILs arose in the distant past from mutated inteins or from a common ancestor of inteins.

Regulation of Splicing by Mechanism-linked Conformational Changes and Kinetic Rates

Although the basic steps in protein splicing were elucidated in the 1990s, we still lack a consensus for how they are coordinated. Two basic processes are invoked: 1) conformational changes triggered by a preceding step result in formation of a robust active site for the next step and 2) differences in kinetic

rates for each step ensure correct reaction order. Conformational changes may be as simple as fixing different rotamer positions, or they may involve larger movements. Evidence for conformational control can be inferred from the absolute coupling of N-terminal and C-terminal reactions observed in some inteins where C-terminal cleavage only occurs if preceding steps have been completed (14, 59, 69, 84, 85).

The most common argument for larger scale movement in intein active sites comes from intein structures. Only the Sce VMA intein (86) and the Pho RadA intein (46) structures have distances between the C-extein nucleophile and N-terminal scissile bond that are directly compatible with catalysis (3.8 Å). This distance is much larger (~8 Å) in all other intein structures to date and requires a conformational change for catalysis (41). A conformational shift was also proposed in the class 2 Mja KlbA intein where a rearrangement of $Ser^{G:6}$, $Asn^{G:7}$, and Cys^{+1} (G:8) backbone torsional angles could enable a close approach of the Cys^{+1} nucleophile to the N-terminal scissile bond (76). It remains to be determined why inteins display such an open active site and whether it represents a true conformation or an artifact of experimental conditions that prevent splicing, including mutations to active site residues and differences in extein sequence or length.

Movement of side chains during splicing can coordinate the reaction by the gain or loss of hydrogen bonds and changes in van der Waals packing interactions to align catalytic residues. For example, structures of the Pho RadA intein suggest that $Asp^{F:4}$ hydrogen-bonds to $Asn^{G:7}$, preventing Asn cyclization until branched intermediate formation causes reorientation of the $Asp^{F:4}$ side chain (46). Another example involves coupling of N- and C-terminal cleavage in the Ssp DnaE intein, which is proposed to be due to Tyr^{-1} preventing proper orientation of $Arg^{B:11}$ until formation of the linear and/or branched thioester intermediate results in movement of the Tyr^{-1} side chain, allowing the $Arg^{B:11}$ side chain to reorient and assist Asn cyclization (87). In the Mxe GyrA intein, NMR data show that chemical or conformational changes in the branched intermediate stimulates Asn cyclization (69).

Kinetic data can provide further insight into how inteins control the steps of splicing. Asn cyclization is the slowest step for most inteins studied to date, including the Pab Pol II intein (71), the split Ssp DnaE intein (32, 49, 88), and the Mxe GyrA intein (69). In the Pab Pol II intein, substitution of $Gln^{G:7}$ with the more common $Asn^{G:7}$ accelerated C-terminal cleavage by 20-fold and the overall splicing reaction by 3-fold. The naturally split Npu and Ssp DnaE inteins have been extensively investigated as model systems for intein kinetics because it is easy to initiate reactions by mixing fragments (29, 32, 48, 49, 87–90). Ssp DnaE intein studies demonstrate that association between the fragments is not rate determining (32). Whereas Asn cyclization is the slow step for the Ssp DnaE intein, all steps occur with similar rates in the Npu DnaE intein (32, 49). Although the Ssp DnaE intein splices with overall rates similar to standard inteins, the Npu DnaE intein splices very rapidly, with a half-life of 1 min or less (30, 89). Recently discovered split inteins from metagenomic samples can splice even more rapidly (91).

The class 2 Mja KlbA intein was studied using a semisynthetic intein precursor that could be induced to splice with a redox switch (92). Branched intermediate formation was comparable with the rate of Asn cyclization, and a clear rate-limiting step was not identified.

In all of these inteins, the rates of Asn cyclization are comparable or slower than preceding steps, ensuring that Asn cyclization will not precede extein ligation, which occurs during branched intermediate formation. Kinetic control complements coordination strategies involving conformational changes. This was shown experimentally in the Mxe GyrA intein, where the rate of C-terminal cleavage increased 10-fold when a branched intermediate was present (69).

Conditional Protein Splicing

Inteins have evolved to tightly regulate the steps of splicing. This is essential as inteins interrupt highly conserved domains of proteins important to their host organisms, including DNA polymerases and helicases (93). However, no evidence has been discovered for a physiologically relevant role for conditional protein splicing. This suggests that modern inteins are likely molecular parasites and that efficient, traceless splicing is essential for their maintenance in the host genome. However, inteins can be engineered to be sensitive to changes in light, pH, temperature, or redox state and to be responsive to the addition of small molecules (33, 94). Even unmodified inteins can be controlled under specific conditions. For example, inteins from thermophilic organisms display temperature-dependent splicing in heterologous precursors (7, 78, 95), both *cis*-splicing and *trans*-splicing inteins are sensitive to inhibition by divalent cations (87, 88, 96–98), and disulfide bonds involving active site Cys residues sensitize splicing to cellular oxidation state (79, 99, 100).

Conclusions

Remaining mechanistic challenges include deciphering how reactions are coordinated and illuminating the diverse ways that inteins promote catalysis. Going forward, detailed studies of catalytic mechanisms, intein kinetics, and structures must occur in the context of native host exteins, which will distinguish between physiologically significant observations and those that may be artifacts of heterologous model systems. Furthermore, detailed studies of multiple inteins will determine whether catalytic strategies are universal or specific to a subset of inteins.

The plethora of reactions performed by HINT domain proteins highlights the robust and flexible nature of catalysis when rapid turnover and substrate binding are not required. This allows for survival of mutated inteins as long as compensatory residues are present to permit a low level of splicing and provides time for the intein to evolve into a more efficient enzyme by testing new catalytic strategies. Thus the flexibility of inteins, BILs, and Hedgehog proteins provides a blueprint for modifying enzyme activity by varying nucleophiles and strategies to activate these nucleophiles.

REFERENCES

1. Shih, C. K., Wagner, R., Feinstein, S., Kanik-Ennulat, C., and Neff, N. (1988) A dominant trifluoperazine resistance gene from *Saccharomyces cerevisiae* has homology with F_0F_1 ATP synthase and confers calcium-

sensitive growth. *Mol. Cell. Biol.* **8,** 3094–3103
2. Hirata, R., Ohsumk, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. (1990) Molecular structure of a gene, *VMA1*, encoding the catalytic subunit of H^+-translocating adenosine triphosphatase from vacuolar membranes of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **265,** 6726–6733
3. Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebl, M., and Stevens, T. H. (1990) Protein splicing converts the yeast TFP1 gene product to the 69-kD subunit of the vacuolar H^+-adenosine triphosphatase. *Science* **250,** 651–657
4. 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
5. Davis, E. O., Jenner, P. J., Brooks, P. C., Colston, M. J., and Sedgwick, S. G. (1992) Protein splicing in the maturation of *M. tuberculosis* recA protein: a mechanism for tolerating a novel class of intervening sequence. *Cell* **71,** 201–210
6. Hodges, R. A., Perler, F. B., Noren, C. J., and Jack, W. E. (1992) Protein splicing removes intervening sequences in an archaea DNA polymerase. *Nucleic Acids Res.* **20,** 6153–6157
7. 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
8. Paulus, H. (2001) Inteins as enzymes. *Bioorg. Chem.* **29,** 119–129
9. Evans, T. C., Jr., and Xu, M. Q. (1999) Intein-mediated protein ligation: harnessing nature's escape artists. *Biopolymers* **51,** 333–342
10. 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
11. Xu, M. Q., Comb, D. G., Paulus, H., Noren, C. J., Shao, Y., and Perler, F. B. (1994) Protein splicing: an analysis of the branched intermediate and its resolution by succinimide formation. *EMBO J.* **13,** 5517–5522
12. Shao, Y., Xu, M. Q., and Paulus, H. (1995) Protein splicing: characterization of the aminosuccinimide residue at the carboxyl terminus of the excised intervening sequence. *Biochemistry* **34,** 10844–10850
13. 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
14. Xu, M. Q., and Perler, F. B. (1996) The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15,** 5146–5153
15. 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
16. Shao, Y., Xu, M. Q., and Paulus, H. (1996) Protein splicing: evidence for an N–O acyl rearrangement as the initial step in the splicing process. *Biochemistry* **35,** 3810–3815
17. Pietrokovski, S. (2001) Intein spread and extinction in evolution. *Trends Genet.* **17,** 465–472
18. Perler, F. B., Olsen, G. J., and Adam, E. (1997) Compilation and analysis of intein sequences. *Nucleic Acids Res.* **25,** 1087–1093
19. Pietrokovski, S. (1994) Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins. *Protein Sci.* **3,** 2340–2350
20. Perler, F. B. (2002) InBase: the Intein Database. *Nucleic Acids Res.* **30,** 383–384
21. Novikova, O., Topilina, N., and Belfort, M. (2014) Enigmatic distribution, evolution, and function of inteins. *J. Biol. Chem.* **289,** 14490–14497
22. Gimble, F. S., and Thorner, J. (1992) Homing of a DNA endonuclease gene by meiotic gene conversion in *Saccharomyces cerevisiae*. *Nature* **357,** 301–306
23. Derbyshire, V., Wood, D. W., Wu, W., Dansereau, J. T., Dalgaard, J. Z., and Belfort, M. (1997) Genetic definition of a protein-splicing domain: functional mini-inteins support structure predictions and a model for intein evolution. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 11466–11471
24. Telenti, A., Southworth, M., Alcaide, F., Daugelat, S., Jacobs, W. R., Jr., and Perler, F. B. (1997) The *Mycobacterium xenopi* GyrA protein splicing element: characterization of a minimal intein. *J Bacteriol.* **179,** 6378–6382
25. 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
26. Hiraga, K., Derbyshire, V., Dansereau, J. T., Van Roey, P., and Belfort, M. (2005) Minimization and stabilization of the *Mycobacterium tuberculosis recA* intein. *J. Mol. Biol.* **354,** 916–926
27. Elleuche, S., Döring, K., and Pöggeler, S. (2008) Minimization of a eukaryotic mini-intein. *Biochem. Biophys. Res. Commun.* **366,** 239–243
28. Wu, H., Hu, Z., and Liu, X. Q. (1998) Protein *trans*-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. *Proc. Natl. Acad. Sci. U.S.A.* **95,** 9226–9231
29. Iwai, H., Züger, S., Jin, J., and Tam, P. H. (2006) Highly efficient protein *trans*-splicing by a naturally split DnaE intein from *Nostoc punctiforme*. *FEBS Lett.* **580,** 1853–1858
30. Zettler, J., Schütz, 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
31. Caspi, J., Amitai, G., Belenkiy, O., and Pietrokovski, S. (2003) Distribution of split DnaE inteins in Cyanobacteria. *Mol. Microbiol.* **50,** 1569–1577
32. 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
33. Wood, D. W., and Camarero, J. A. (2014) Intein applications: from protein purification and labeling to metabolic control methods. *J. Biol. Chem.* **289,** 14512–14519
34. Hall, T. M., Porter, J. A., Young, K. E., Koonin, E. V., Beachy, P. A., and Leahy, D. J. (1997) Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell* **91,** 85–97
35. Porter, J. A., von Kessler, D. P., Ekker, S. C., Young, K. E., Lee, J. J., Moses, K., and Beachy, P. A. (1995) The product of *hedgehog* autoproteolytic cleavage active in local and long-range signalling. *Nature* **374,** 363–366
36. Dassa, B., Haviv, H., Amitai, G., and Pietrokovski, S. (2004) Protein splicing and auto-cleavage of bacterial intein-like domains lacking a C′-flanking nucleophilic residue. *J. Biol. Chem.* **279,** 32001–32007
37. Aranko, A. S., Oeemig, J. S., and Iwaï, H. (2013) Structural basis for protein *trans*-splicing by a bacterial intein-like domain: protein ligation without nucleophilic side chains. *FEBS J.* **280,** 3256–3269
38. Amitai, G., Belenkiy, O., Dassa, B., Shainskaya, A., and Pietrokovski, S. (2003) Distribution and function of new bacterial intein-like protein domains. *Mol. Microbiol.* **47,** 61–73
39. Southworth, M. W., Yin, J., and Perler, F. B. (2004) Rescue of protein splicing activity from a *Magnetospirillum magnetotacticum* intein-like element. *Biochem. Soc. Trans.* **32,** 250–254
40. Volkmann, G., and Mootz, H. D. (2013) Recent Progress in intein research: from mechanism to directed evolution and applications. *Cell. Mol. Life Sci.* **70,** 1185–1206
41. Eryilmaz, E., Shah, N. E., Muir, T. W., and Cowburn, D. (2014) Structural and dynamical features of inteins and implications on protein splicing. *J. Biol. Chem.* **289,** 14506–14511
42. 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
43. Kawasaki, M., Nogami, S., Satow, Y., Ohya, Y., and Anraku, Y. (1997) Identification of three core regions essential for protein splicing of the yeast Vma1 protozyme: a random mutagenesis study of the entire *VMA1*-derived endonuclease sequence. *J. Biol. Chem.* **272,** 15668–15674
44. 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
45. Amitai, G., Callahan, B. P., Stanger, M. J., Belfort, G., and Belfort, M. (2009) Modulation of intein activity by its neighboring extein substrates.

Proc. Natl. Acad. Sci. U.S.A. **106,** 11005–11010
46. Oeemig, J. S., Zhou, D., Kajander, T., Wlodawer, A., and Iwaï, H. (2012) NMR and crystal structures of the *Pyrococcus horikoshii* RadA intein guide: a strategy for engineering a highly efficient and promiscuous intein. *J. Mol. Biol.* **421,** 85–99
47. Nogami, S., Satow, Y., Ohya, Y., and Anraku, Y. (1997) Probing novel elements for protein splicing in the yeast Vma1 protozyme: a study of replacement mutagenesis and intragenic suppression. *Genetics* **147,** 73–85
48. Cheriyan, M., Pedamallu, C. S., Tori, K., and Perler, F. (2013) Faster protein splicing with the *Nostoc punctiforme* DnaE intein using non-native extein residues. *J. Biol. Chem.* **288,** 6202–6211
49. Shah, N. H., Eryilmaz, E., Cowburn, D., and Muir, T. W. (2013) Extein residues play an intimate role in the rate-limiting step of protein *trans*-splicing. *J. Am. Chem. Soc.* **135,** 5839–5847
50. Pearl, E. J., Bokor, A. A., Butler, M. I., Poulter, R. T., and Wilbanks, S. M. (2007) Preceding hydrophobic and β-branched amino acids attenuate splicing by the CnePRP8 intein. *Biochim. Biophys. Acta* **1774,** 995–1001
51. Pearl, E. J., Tyndall, J. D., Poulter, R. T., and Wilbanks, S. M. (2007) Sequence requirements for splicing by the Cne PRP8 intein. *FEBS Lett.* **581,** 3000–3004
52. Tori, K., and Perler, F. B. (2011) The Arthrobacter species FB24 Arth_1007 (DnaB) intein is a pseudogene. *PLoS One* **6,** e26361
53. Tori, K., Cheriyan, M., Pedamallu, C. S., Contreras, M. A., and Perler, F. B. (2012) The *Thermococcus kodakaraensis* Tko CDC21–1 intein activates its N-terminal splice junction in the absence of a conserved histidine by a compensatory mechanism. *Biochemistry* **51,** 2496–2505
54. Poland, B. W., Xu, M. Q., and Quiocho, F. A. (2000) Structural insights into the protein splicing mechanism of PI-SceI. *J. Biol. Chem.* **275,** 16408–16413
55. 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
56. 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
57. Ding, Y., Xu, M. Q., Ghosh, I., Chen, X., Ferrandon, S., Lesage, G., and Rao, Z. (2003) Crystal structure of a mini-intein reveals a conserved catalytic module involved in side chain cyclization of asparagine during protein splicing. *J. Biol. Chem.* **278,** 39133–39142
58. Dearden, A. K., Callahan, B., Roey, P. V., Li, Z., Kumar, U., Belfort, M., and Nayak, S. K. (2013) A conserved threonine spring-loads precursor for intein splicing. *Protein Sci.* **22,** 557–563
59. Du, Z., Zheng, Y., Patterson, M., Liu, Y., and Wang, C. (2011) pK_a coupling at the intein active site: implications for the coordination mechanism of protein splicing with a conserved aspartate. *J. Am. Chem. Soc.* **133,** 10275–10282
60. Du, Z., Shemella, P. T., Liu, Y., McCallum, S. A., Pereira, B., Nayak, S. K., Belfort, G., Belfort, M., and Wang, C. (2009) Highly conserved histidine plays a dual catalytic role in protein splicing: a pK_a shift mechanism. *J. Am. Chem. Soc.* **131,** 11581–11589
61. 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
62. Schwarzer, D., Ludwig, C., Thiel, I. V., and Mootz, H. D. (2012) Probing intein-catalyzed thioester formation by unnatural amino acid substitutions in the active site. *Biochemistry* **51,** 233–242
63. Shingledecker, K., Jiang, S. q., and Paulus, H. (2000) Reactivity of the cysteine residues in the protein splicing active center of the *Mycobacterium tuberculosis* RecA intein. *Arch Biochem. Biophys.* **375,** 138–144
64. Pereira, B., Shemella, P. T., Amitai, G., Belfort, G., Nayak, S. K., and Belfort, M. (2011) Spontaneous proton transfer to a conserved intein residue determines on-pathway protein splicing. *J. Mol. Biol.* **406,** 430–442
65. Wood, D. W., Wu, W., Belfort, G., Derbyshire, V., and Belfort, M. (1999) A genetic system yields self-cleaving inteins for bioseparations. *Nat. Biotechnol.* **17,** 889–892
66. Brennan, T. V., and Clarke, S. (1995) Effect of adjacent histidine and cysteine residues on the spontaneous degradation of asparaginyl- and aspartyl-containing peptides. *Int. J. Pept. Protein Res.* **45,** 547–553
67. Mujika, J. I., Lopez, X., and Mulholland, A. J. (2009) Modeling protein splicing: reaction pathway for C-terminal splice and intein scission. *J. Phys. Chem. B* **113,** 5607–5616
68. Mujika, J. I., Lopez, X., and Mulholland, A. J. (2012) Mechanism of C-terminal intein cleavage in protein splicing from QM/MM molecular dynamics simulations. *Org. Biomol. Chem.* **10,** 1207–1218
69. 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
70. Wood, D. W., Derbyshire, V., Wu, W., Chartrain, M., Belfort, M., and Belfort, G. (2000) Optimized single-step affinity purification with a self-cleaving intein applied to human acidic fibroblast growth factor. *Biotechnol. Prog.* **16,** 1055–1063
71. 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
72. Shemella, P., Pereira, B., Zhang, Y., Van Roey, P., Belfort, G., Garde, S., and Nayak, S. K. (2007) Mechanism for intein C-terminal cleavage: a proposal from quantum mechanical calculations. *Biophys. J.* **92,** 847–853
73. Shemella, P. T., Topilina, N. I., Soga, I., Pereira, B., Belfort, G., Belfort, M., and Nayak, S. K. (2011) Electronic structure of neighboring extein residue modulates intein C-terminal cleavage activity. *Biophys. J.* **100,** 2217–2225
74. Chen, L., Benner, J., and Perler, F. B. (2000) Protein splicing in the absence of an intein penultimate histidine. *J. Biol. Chem.* **275,** 20431–20435
75. Wang, S., and Liu, X. Q. (1997) Identification of an unusual intein in chloroplast ClpP protease of *Chlamydomonas eugametos*. *J. Biol. Chem.* **272,** 11869–11873
76. Johnson, M. A., Southworth, M. W., Herrmann, T., Brace, L., Perler, F. B., and Wüthrich, K. (2007) NMR Structure of a KlbA intein precursor from *Methanococcus jannaschii*. *Protein Sci.* **16,** 1316–1328
77. 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
78. 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
79. Nicastri, M. C., Xega, K., Li, L., Xie, J., Wang, C., Linhardt, R. J., Reitter, J. N., and Mills, K. V. (2013) Internal disulfide bond acts as a switch for intein activity. *Biochemistry* **52,** 5920–5927
80. Amitai, G., Dassa, B., and Pietrokovski, S. (2004) Protein splicing of inteins with atypical glutamine and aspartate C-terminal residues. *J. Biol. Chem.* **279,** 3121–3131
81. Shao, Y., and Paulus, H. (1997) Protein splicing: estimation of the rate of O–N and S–N acyl rearrangements, the last step of the splicing process. *J. Pept. Res.* **50,** 193–198
82. Tori, K., and Perler, F. B. (2011) Expanding the definition of class 3 inteins and their proposed phage origin. *J. Bacteriol.* **193,** 2035–2041
83. Brace, L. E., Southworth, M. W., Tori, K., Cushing, M. L., and Perler, F. (2010) The *Deinococcus radiodurans* Snf2 intein caught in the act: detection of the Class 3 intein signature block F branched intermediate. *Protein Sci.* **19,** 1525–1533
84. O'Brien, K. M., Schufreider, A. K., McGill, M. A., O'Brien, K. M., Reitter, J. N., and Mills, K. V. (2010) Mechanism of protein splicing of the *Pyrococcus abyssi* Lon protease intein. *Biochem. Biophys. Res. Commun.* **403,** 457–461
85. Volkmann, G., and Liu, X. Q. (2011) Intein lacking conserved C-terminal motif G retains controllable N-cleavage activity. *FEBS J.* **278,** 3431–3446
86. Mizutani, R., Anraku, Y., and Satow, Y. (2004) Protein splicing of yeast VMA1-derived endonuclease via thiazolidine intermediates. *J. Synchrotron Radiat.* **11,** 109–112
87. Sun, P., Ye, S., Ferrandon, S., Evans, T. C., Xu, M. Q., and Rao, Z. (2005) Crystal structures of an intein from the split *dnaE* gene of *Synechocystis* sp. PCC6803 reveal the catalytic model without the penultimate histidine

and the mechanism of zinc ion inhibition of protein splicing. *J. Mol. Biol.* **353,** 1093–1105
88. 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
89. Shah, N. H., Dann, G. P., Vila-Perelló, M., Liu, Z., and Muir, T. W. (2012) Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. *J. Am. Chem. Soc.* **134,** 11338–11341
90. 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
91. Carvajal-Vallejos, P., Pallissé, R., Mootz, H. D., and Schmidt, S. R. (2012) Unprecedented rates and efficiencies revealed for new natural split inteins from metagenomic sources. *J. Biol. Chem.* **287,** 28686–28696
92. Saleh, L., Southworth, M. W., Considine, N., O'Neill, C., Benner, J., Bollinger, J. M., Jr., and Perler, F. B. (2011) Branched intermediate formation is the slowest step in the protein splicing reaction of the Ala1 KlbA intein from *Methanococcus jannaschii*. *Biochemistry* **50,** 10576–10589
93. Dalgaard, J. Z., Moser, M. J., Hughey, R., and Mian, I. S. (1997) Statistical modeling, phylogenetic analysis and structure prediction of a protein splicing domain common to inteins and Hedgehog proteins. *J. Comput. Biol.* **4,** 193–214
94. Mootz, H. D. (2009) Split inteins as versatile tools for protein semisynthesis. *Chembiochem* **10,** 2579–2589
95. Cambon-Bonavita, M. A., Schmitt, P., Zieger, M., Flaman, J. M., Lesongeur, F., Raguénès, G., Bindel, D., Frisch, N., Lakkis, Z., Dupret, D., Barbier, G., and Quérellou, J. (2000) Cloning, expression, and characterization of DNA polymerase I from the hyperthermophilic archaea *Thermococcus fumicolans*. *Extremophiles* **4,** 215–225
96. Ghosh, I., Sun, L., and Xu, M. Q. (2001) Zinc Inhibition of protein *trans*-splicing and identification of regions essential for splicing and association of a split intein. *J. Biol. Chem.* **276,** 24051–24058
97. Mills, K. V., and Paulus, H. (2001) Reversible inhibition of protein splicing by zinc ion. *J. Biol. Chem.* **276,** 10832–10838
98. Zhang, L., Xiao, N., Pan, Y., Zheng, Y., Pan, Z., Luo, Z., Xu, X., and Liu, Y. (2010) Binding and inhibition of copper ions to RecA inteins from *Mycobacterium tuberculosis*. *Chemistry* **16,** 4297–4306
99. Callahan, B. P., Topilina, N. I., Stanger, M. J., Van Roey, P., and Belfort, M. (2011) Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nat. Struct. Mol. Biol.* **18,** 630–633
100. Chen, W., Li, L., Du, Z., Liu, J., Reitter, J. N., Mills, K. V., Linhardt, R. J., and Wang, C. (2012) Intramolecular disulfide bond between catalytic cysteines in an intein precursor. *J. Am. Chem. Soc.* **134,** 2500–2503