

Protein splicing of a recombinase intein induced by ssDNA and DNA damage

Christopher W. Lennon,¹ Matthew Stanger,¹ and Marlene Belfort^{1,2}

¹Department of Biological Sciences, RNA Institute, University at Albany, Albany, New York 12222, USA; ²Department of Biomedical Sciences, School of Public Health, University at Albany, Albany, New York 12201, USA

Inteins (or protein introns) autocatalytically excise themselves through protein splicing. We challenge the long-considered notion that inteins are merely molecular parasites and posit that some inteins evolved to regulate host protein function. Here we show substrate-induced and DNA damage-induced splicing, in which an archaeal recombinase RadA intein splices dramatically faster and more accurately when provided with ssDNA. This unprecedented example of intein splicing stimulation by the substrate of the invaded host protein provides compelling support in favor of inteins acting as pause buttons to arrest protein function until needed; then, an immediate activity switch is triggered, representing a new form of post-translational control.

Supplemental material is available for this article.

Received August 31, 2016; revised version accepted December 8, 2016.

Inteins (or protein introns) are microbial intervening sequences that self-splice at the protein level and invade genes at the DNA level (Mills et al. 2014; Shah and Muir 2014). Found in approximately one-quarter of bacterial and one-half of archaeal genomes (Novikova et al. 2016), inteins have been considered as purely selfish elements (Naor et al. 2016). However, 70% of inteins occur in ATP-binding proteins (Novikova et al. 2016), and splicing can be regulated by environmental stresses (Callahan et al. 2011; Topilina et al. 2015a,b; Reitter et al. 2016), suggesting that some inteins have evolved to regulate host protein function.

Some archaeal and bacterial recombinases, which catalyze homologous recombination and contain ATPase domains, harbor inteins. Such recombinases are ubiquitous in nature, being critical in DNA repair and homologous recombination (Seitz et al. 2001; Lin et al. 2006; Cox 2007; Bell and Kowalczykowski 2016). These proteins first bind ssDNA to form nucleoprotein filaments and then direct pairing with homologous dsDNA and initiate ssDNA strand invasion into the dsDNA template in an ATP-dependent manner (Seitz et al. 2001). RadA is the archaeal homolog of Rad51 in eukaryotes and RecA in prokaryotes

(Seitz et al. 1998). The RadA intein from the hyperthermophilic archaeon *Pyrococcus horikoshii* (*Pho*) provides a striking example of conditional protein splicing, in which increased temperature is required for the reaction to proceed (Topilina et al. 2015b).

Here we report that the *Pho* RadA intein-containing precursor is triggered to splice dramatically faster and more accurately when presented with ssDNA substrate. Additionally, RadA intein splicing is stimulated in *Escherichia coli* in response to DNA damage and thus intracellular ssDNA. Furthermore, RadA can bind to ssDNA prior to intein splicing, promoting the novel concept that intein-containing proteins may retain partial function. Our discovery represents an unprecedented example of intein splicing stimulation by the substrate of the invaded protein. This work also provides compelling support for inteins acting as pause buttons to arrest protein function until needed for an immediate activity switch, representing a new form of post-translational control.

Results and Discussion

Pho RadA splicing is triggered specifically by ssDNA substrate

The archaeal *Pho* RadA intein provides an impressive example of conditional protein splicing, in which the native exteins have a strong influence on reaction rate (Topilina et al. 2015b). Interactions between the intein and C-terminal extein (C-extein) inhibit splicing, and either increased temperature or solvent conditions that disrupt protein-protein interactions are required for splicing to proceed (Topilina et al. 2015b). Whereas temperature represents a general environmental condition important for growth of the hyperthermophilic *Pho*, we asked whether another layer of regulation might exist in which splicing is influenced by a factor directly related to RadA function.

The *Pho* RadA intein is within the ATPase domain, inserted directly into the P loop required for ATP binding (Topilina et al. 2015b). However, ATP does not increase the rate of splicing (Supplemental Fig. 1A). We therefore hypothesized that an environmental signal prior to ATP hydrolysis may serve to trigger splicing. Because RadA must bind ssDNA in order to initiate DNA repair and recombination and because ssDNA activates RadA ATPase activity (Seitz et al. 1998; Komori et al. 2000b; Galkin et al. 2006), we asked whether this physiologically relevant substrate might influence splicing. Indeed, M13mp18 ssDNA, a recombinase ssDNA substrate commonly used in vitro, increased the extent of *Pho* RadA intein splicing manifold (Fig. 1A). Addition of ATP with ssDNA did not stimulate splicing more than ssDNA alone (Supplemental Fig. 1A). Our discovery that ssDNA stimulates RadA intein splicing is in accord with the notion that inteins may function to limit host protein activity prior to splicing and immediately switch on when the host protein is needed by the cell. We chose to use a truncated version

[*Keywords*: conditional protein splicing; post-translational regulation; RadA]

Corresponding author: mbelfort@albany.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.289280.116>.

© 2016 Lennon et al. This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see <http://genesdev.cshlp.org/site/misc/terms.xhtml>). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

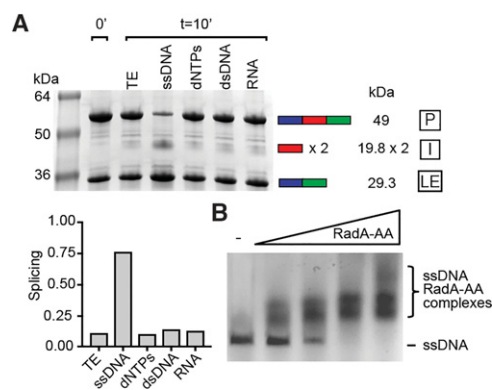


Figure 1. *Pho* RadA intein splicing is stimulated by ssDNA, which binds to the RadA precursor. (A) Specificity of RadA intein splicing stimulation. Splicing reaction mixtures were incubated for 10 min at 63°C, separated by SDS-PAGE (8%–16%), and stained with Coomassie, and the levels of precursor (P), ligated exteins (LEs), and intein (I) were measured by densitometry. The intein appears as a dimer (×2), consistent with previous observations (Oeemig et al. 2012; Aranko et al. 2013). (TE) Tris-EDTA, the storage buffer for ssDNA; (ssDNA) M13mp18 ssDNA; (dNTPs) deoxynucleotide triphosphates; (dsDNA) double-stranded plasmid DNA; (RNA) total cellular RNA. ssDNA, dNTPs, dsDNA, and RNA were all at 187.5 ng/μL. (Below) Bar graph quantitation of the gel in A. (B) Splicing-inactive RadA precursor (RadA-AA) binds to ssDNA. M13mp18 ssDNA is at 5 μM binding sites (3 bases per RadA-binding site) (Seitz et al. 1998), and RadA-AA is at 0, 2.5, 5, 10, or 20 μM (from left to right). Samples were incubated for 60 min at 63°C, separated on 1% agarose gels, and visualized using EZ Vision dye under UV light. Data in all panels are representative of at least three experiments.

of RadA lacking the N-terminal domain (residues 1–115) that leaves the ATPase intact, as this domain is capable of binding ssDNA and hydrolyzing ATP (Komori et al. 2000a). Additionally, the ATPase domain preserves all predicted interactions between the C-extein and intein required for temperature-dependent splicing (Topilina et al. 2015b), and pilot experiments showed that this region was dispensable for response to ssDNA (Supplemental Fig. 1B).

The stimulatory effect of ssDNA is specific, with deoxynucleotide triphosphates (dNTPs), dsDNA of either the same sequence as or a different sequence than M13mp18 ssDNA, or total cellular RNA having no measurable influence on splicing (Fig. 1A; Supplemental Fig. 1C). As further confirmation that ssDNA was directly responsible for the observed increase in RadA splicing, we tested whether the stimulation of splicing could be attenuated by the formation of secondary structure in the ssDNA. Addition of MgCl₂ induces base pairing in ssDNA and has been shown to prevent RecA binding and nucleoprotein filament formation (Chow and Radding 1985). MgCl₂ led to a loss in splicing stimulation by ssDNA in a manner dependent on MgCl₂ concentrations similar to those that disrupt RecA binding (Chow and Radding 1985), whereas MgCl₂ in the absence of ssDNA did not inhibit splicing (Supplemental Fig. 1D). These results suggest that increased secondary structure in the ssDNA decreases RadA splicing, although it is possible that the MgCl₂ effects observed are due to a conformational change in RadA itself. Nevertheless, the data demonstrate that splicing is specifically accelerated when the precursor is presented with ssDNA, the natural RadA substrate, which signals that RadA activity is needed by the cell.

The *Pho* RadA precursor binds to ssDNA

To determine whether ssDNA is acting directly on RadA to accelerate splicing, we asked whether the RadA precursor (P) could bind to ssDNA prior to splicing. For this, we used a splicing-inactive mutant of RadA in which the first residue of the intein, a catalytic cysteine (C1), and terminal asparagine, critical to intein release, are changed to alanines (RadA-AA), trapping RadA in the precursor form. Using electrophoretic mobility shift assays, we found that RadA-AA could bind to ssDNA in a concentration-dependent manner with an apparent K_d of ~4.5 μM (Fig. 1B; Supplemental Fig. 2). To determine whether this interaction was mediated by extein or intein residues, we also measured ssDNA binding by fully spliced RadA and by free intein. Whereas RadA-AA binds to ssDNA with a profile similar to that of spliced RadA (apparent K_d of ~3 μM), free intein binding to ssDNA could not be detected (Supplemental Fig. 2). This is the first report of an intein-containing precursor retaining any function (DNA binding), which challenges the paradigm that the inteins must splice out prior to any manifestation of host protein activity. These findings are particularly germane given that ssDNA represents a cellular signal that RadA is needed.

We next tested the prediction that RadA splicing acceleration by ssDNA is specific to native exteins. For this, we used a previously developed splicing reporter with foreign exteins, called MIG (MBP-intein-GFP) (Topilina et al. 2015a,b). In the MIG construct, the intein with short native exteins (10 residues) is flanked on the N terminus by maltose-binding protein (MBP) and on the C terminus by green fluorescent protein (GFP). In agreement with our observation that free inteins do not bind ssDNA (rather that the native exteins are responsible), we found that ssDNA had no influence on *Pho* RadA intein splicing in the MIG extein context (Supplemental Fig. 3).

Pho RadA splicing is faster and more accurate in the presence of ssDNA substrate

To determine the degree of *Pho* RadA intein splicing acceleration by ssDNA, we performed time-course experiments to follow the reaction kinetics, measuring the fraction spliced after various incubation times at 63°C in the presence or absence of ssDNA (Fig. 2A,B). We observed an initial rate increase of >46-fold in the presence of ssDNA. This dramatic acceleration of splicing is exaggerated when the RadA concentration is lowered relative to ssDNA (to twofold to threefold greater rate increases at one-fourth the RadA concentration), indicating a competition for ssDNA-binding sites (Supplemental Fig. 4A).

Splicing has the potential to generate irreversible off-pathway reactions, the products of which currently have no known function (Topilina et al. 2015a). For example, when N-terminal cleavage occurs, the reaction produces free N-extein and intein joined to C-extein (I-C), with less ligated extein (LE); namely, functional RadA. In addition to stimulating rate, ssDNA also improved the accuracy of *Pho* RadA intein splicing, with more on-pathway processing of RadA precursor to LE and less off-pathway I-C production (Figs. 2A, 3A,B). Indeed, the accuracy of splicing with ssDNA approaches the estimated theoretical limit of 1.67 new LE (Fig. 3A, dashed line), which was calculated from the amount of precursors at the start of the time course (see the Materials and Methods). In contrast, only one out of every two to three precursors forms

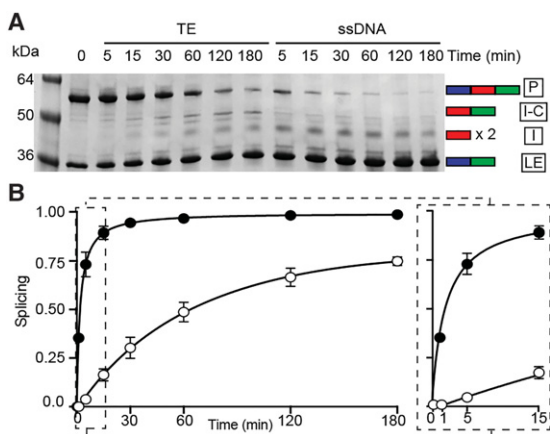


Figure 2. *Pho* RadA intein splicing is faster in the presence of ssDNA. (A) Time course after ssDNA addition. The experiment was conducted as in Figure 1A with extended incubation times. In the absence of ssDNA, the off-pathway splicing product intein–C-terminus (I-C) appears over time. (B) Kinetics of splicing. Measurements were in the presence (filled circles) or absence (open circles) of ssDNA, with calculations as described in the Materials and Methods. An expansion of the graph to display the first 15 min of the splicing reaction is at the right. Initial rates are increased 46-fold with ssDNA (3.54×10^{-1} per min vs. 7.60×10^{-3} per min). Error was calculated as the standard deviation from three independent experiments.

LE without ssDNA. At the end of a 3-h splicing time course in which >80% of the starting precursor had processed, we observed approximately threefold more new LE formation with ssDNA than without (Fig. 3B). Finally, in the absence of ssDNA, there is the formation of ~14-fold more off-pathway I-C product, the C-terminal fragment resulting from N-terminal cleavage (Fig. 3B).

Pho RadA splicing is promoted by intracellular DNA damage in *E. coli*

We next sought to determine whether the effects of ssDNA on *Pho* RadA intein splicing rate in vitro were observable in a cellular context. In this case, *E. coli* is a valuable proxy for *Pho* for in vivo splicing studies, since substrate ssDNA is biochemically the same in all organisms. Additionally, we found that acceleration of *Pho* RadA intein splicing by ssDNA was robust across a wide temperature range, including 37°C (Supplemental Fig. 4B), the optimum growth temperature of *E. coli*. To determine whether *Pho* RadA intein splicing can be influenced by intracellular ssDNA, we treated *E. coli* with nalidixic acid (Nal), a gyrase inhibitor that leads to ssDNA accumulation (Kohiyama et al. 2013), or UV. Because induction of the SOS response generates the ssDNA acted on by recombinases during repair (Bell and Kowalczykowski 2016), we examined splicing in SOS⁺ and their isogenic SOS⁻ strain backgrounds (Supplemental Table 3). To this end, we used a deletion of the *recA* recombinase (*recA*⁻) and a noncleavable variant of the *lexA* repressor (*lexA3*), both of which render the SOS response noninducible. We observed up to 10-fold more splicing within the first hour after Nal treatment in the SOS⁺ strains compared with the SOS⁻ strains (Fig. 4A,B; Supplemental Fig. 5). Importantly, this dramatic splicing response to newly produced ssDNA was observed only in SOS⁺ strains in response to Nal

treatment and not in *recA*⁻ or *lexA3* cells. UV light exposure, another means to damage DNA, activates SOS and thus produces ssDNA in vivo, which leads to an approximately threefold increase in splicing when comparing SOS⁺ with SOS⁻ strains (Supplemental Fig. 5). There was slightly more splicing in SOS⁺ strains even in the absence of Nal (Supplemental Fig. 5) but to a much lower degree than following Nal treatment, which is not surprising given the constant background of DNA damage. We also found that DNA copurifies with RadA only after Nal treatment (Fig. 4C), strongly suggesting a direct effect on splicing caused by RadA–DNA interaction under conditions of DNA damage.

Model of *Pho* RadA splicing acceleration by ssDNA

Structural studies of RadA and Rad51 have identified regions of the protein involved in ssDNA binding; namely, loop 1 (L1) and L2 (Conway et al. 2004; Wu et al. 2004; Chen et al. 2007, 2008). Our attention turned to a residue adjacent to L2, R503, a highly conserved amino acid in both archaeal RadA and eukaryotic Rad51 (Prasad et al. 2006) that was previously identified as contacting intein residues (Supplemental Fig. 6A,B; Topilina et al. 2015b). Substitution of the corresponding residue in human Rad51 (R310) to alanine decreases affinity for ssDNA (Prasad et al. 2006). We therefore tested the extent to which *Pho* RadA splicing could be accelerated by ssDNA in R503A as well as the affinity of the mutant for ssDNA. As opposed to the 46-fold effect with wild-type RadA, we found that the initial rate of *Pho* RadA-R503A intein splicing was only 11-fold faster in the presence of ssDNA (Supplemental Fig. 6C,D). This approximately fourfold reduction in splicing stimulation is in close agreement with an approximately threefold lower affinity of RadA-R503A for ssDNA (apparent K_d of ~9 μM) (Supplemental Fig. 2), consistent with ssDNA binding to RadA being

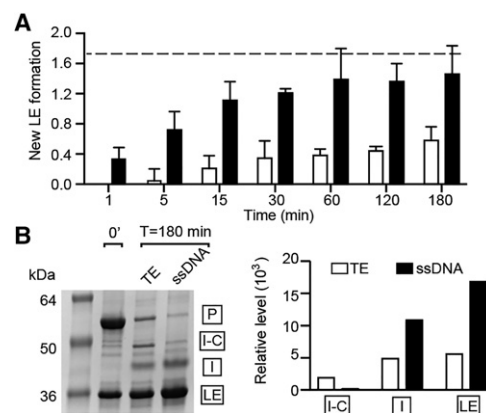


Figure 3. *Pho* RadA intein splicing is more accurate in the presence of ssDNA. (A) New LE formation with and without ssDNA. New LE was determined as the amount of LE for each time point, minus the amount of LE at time 0. The dashed line indicates the estimated maximum possible new LE formation from the starting amount of precursors. Error was calculated as the standard deviation from three independent experiments. For A and B, open bars are in the absence of ssDNA and closed bars are in the presence of ssDNA. (B) The accuracy of splicing after 3 h in the presence and absence of ssDNA, where total splicing exceeds 80% in both cases, and off-pathway I-C appears without ssDNA. The bar graph of the gel shows on-pathway (I and LE) and off-pathway (I-C) products.

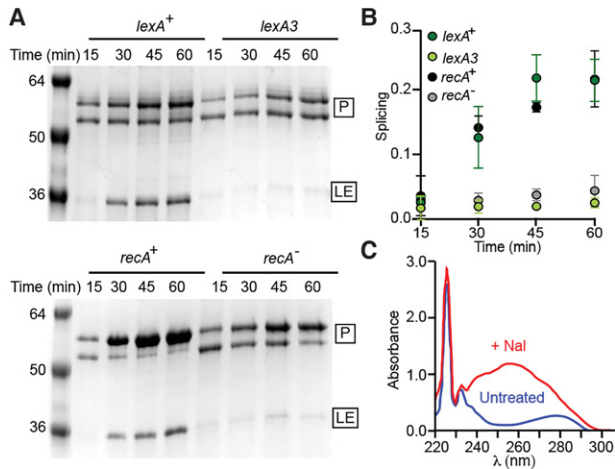


Figure 4. DNA damage in *E. coli* leads to increased *Pho* RadA intein splicing. (A) Splicing after DNA damage. RadA was expressed in *E. coli* at 37°C in SOS⁺ (*recA*⁺ and *lexA*⁺) and their SOS⁻ counterparts (*recA*⁻ and *lexA*³) in the presence of 200 μg/mL Nal. The precursor (P) and LE were isolated by his tag-mediated pull-down, separated by SDS-PAGE, and stained with Coomassie, and the amounts of precursor and LE were determined by densitometry. The precursor often resolves as two bands when expressed at 37°C (Topilina et al. 2015b). (B) Quantitation of splicing. Plots represent three independent experiments, with error as the standard deviation. (C) DNA copurifies with RadA only when expressed after Nal treatment. UV spectra following pull-down of equal amounts of RadA protein expressed in the presence (red) or absence (blue) of Nal.

responsible for the effect on splicing. We propose that, in the presence of DNA damage leading to ssDNA, *Pho* RadA precursor binds directly to ssDNA, which promotes dramatically faster and more accurate protein splicing, possibly by breaking the same intein-C-extein interactions controlled by temperature (Fig. 5). Splicing leads to ATPase activation, which is also greatly stimulated by ssDNA (Seitz et al. 1998; Komori et al. 2000b; Galkin et al. 2006), producing fully functional RadA for downstream recombination.

Two nucleophiles are required for *Pho* RadA splicing: the first residue of the intein (the catalytic cysteine [C1]) and the first residue of the C-extein (a threonine [T+1]) (Mills et al. 2014). Nucleophilic attack by C1 on the preceding peptide bond to form a thioester (Fig. 5, shaded box, encircled 1) is followed by nucleophilic attack by T+1 on the thioester (Fig. 5, shaded box, encircled 2). Next, the intein is released, and LEs are formed (Mills et al. 2014). Although 177 residues separate R503 from the T+1 residue T325, these two residues are predicted to be within 4 Å in three-dimensional space in a RadA precursor model (Supplemental Fig. 5B; Topilina et al. 2015b). This proximity suggests a structural connection between ssDNA binding and splicing catalysis within the ATP-binding P loop, where the intein resides. Not only was R503 proposed to interact with the intein (Topilina et al. 2015b), but a *Pho* RadA-R503A substitution caused slower splicing than wild type in the absence of ssDNA (Topilina et al. 2015b), suggesting that perturbations in the ssDNA-binding region of RadA has consequences for splicing rate.

Interestingly, T+1 plays pivotal roles in both splicing catalysis and ATPase activity. This threonine not only is the nucleophile required for the second step of splicing (Mills et al. 2014) but also coordinates a catalytic Mg²⁺ during ATP hydrolysis (Wu et al. 2004). T+1 may therefore

serve as a relay center, coordinating ssDNA binding with splicing through proximity to R503, followed by Mg²⁺ binding for ATP hydrolysis (Fig. 5). Consistent with this role for T+1 is the appearance of off-pathway I-C in the absence of ssDNA (Figs. 2A, 3B), which can result from sub-optimal nucleophilic attack by T+1 in the second step of splicing after initial thioester formation by C1.

Splicing activation as instantaneous post-translational control

Although DNA damage has been shown to induce RadA expression in most archaea (Rolfsmeier et al. 2010), in the close *Pho* relative *Pyrococcus furiosus*, RadA is not induced by several DNA-damaging agents, suggesting a post-translational regulatory mechanism to control activity (Komori et al. 2000b). It is unknown whether *Pho* shares this trait with *P. furiosus*, but, in any event, intein splicing in response to ssDNA represents an immediate mechanism to post-translationally activate RadA, providing the cell a much faster avenue than de novo transcription and translation to produce functional RadA under stress in response to ssDNA. Thus, the *Pho* RadA intein can lock the protein in an ATPase-inactive state until ssDNA is available, as the ATPase cannot function prior to splicing (Topilina et al. 2015b). Whereas other in vivo factors undoubtedly play roles in RadA binding to ssDNA,

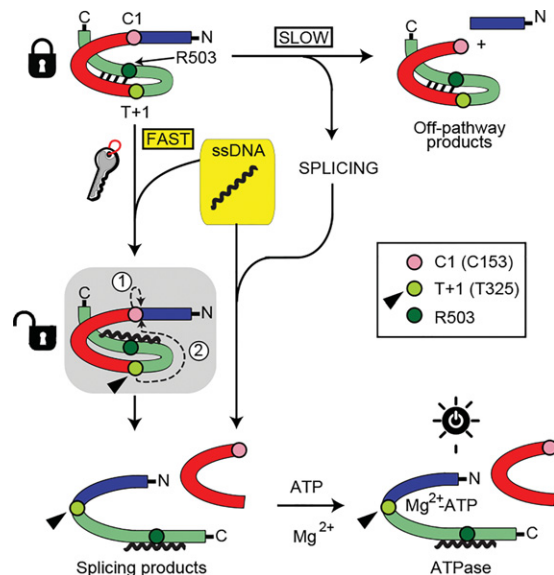


Figure 5. Model for RadA splicing acceleration by ssDNA. The intein is red with C1 highlighted, and exteins are blue and green with T+1 (T325) and R503 highlighted, using colors as in Supplemental Figure 5B. Interactions between the intein and C-extein control RadA splicing, as shown by thermoregulation (Topilina et al. 2015b). In the presence of substrate ssDNA, binding to the C-extein by ssDNA is proposed to disrupt C-extein-intein interactions, promoting fast, accurate splicing. Without ssDNA, splicing is slow, with prominent off-pathway product formation. Splicing is depicted in the shaded box, with steps 1 and 2 (circled) described in the text. After splicing, ATP binds for hydrolysis. T+1 (arrowhead) may act as a relay center to regulate splicing and ATPase activity in response to ssDNA. Binding of ssDNA to the R503 region primes T+1 for nucleophilic attack of the thioester bond between N-extein and intein. Following splicing, T+1 is free to assist in Mg²⁺-ATP binding, turning on ATPase activity.

as has been shown for the better-studied bacterial RecA (Cox 2007; Bell and Kowalczykowski 2016) and eukaryotic Rad51 (San Filippo et al. 2008; Krejci et al. 2012), our work clearly demonstrates that ssDNA alone has a tremendous influence on the rate and accuracy of RadA splicing.

This work provides an unprecedented example of extein interactions with substrate regulating protein splicing, which is especially compelling given the relevance of ssDNA to recombinase function and ATPase stimulation. These observations strongly support the notion that inteins can function as post-translational regulatory switches. Such control could not only benefit the host but also be advantageous to the persistence of the intein. The regulatory potential of the domesticated intein would be selected for, and the intein would be maintained, which would have particular importance for the RadA intein because it lacks a homing endonuclease. The RadA intein thereby adds to the growing list of invasive genetic elements co-opted by the host to serve useful functions, particularly under conditions of stress—in this case, DNA damage.

This work demonstrates a novel mechanism of recombinase control, adding an intriguing new example to the rich tapestry of recombinase regulation (Seitz et al. 2001; Cox 2007; San Filippo et al. 2008; Krejci et al. 2012; Bell and Kowalczykowski 2016). DNA binding by RadA precursors is also of significance beyond our present study, given our discovery that host proteins may retain some activities prior to splicing. Whether the precursor can form higher-order structures—such as the typical RadA non-DNA-bound heptameric rings (Shin et al. 2003) and ssDNA-bound nucleoprotein filaments (Wu et al. 2004)—remains an open question that may be relevant to the many intein-containing proteins that form higher-order structures (Novikova et al. 2016). Regardless, post-translational protein activation by intein splicing in response to factors pertinent to host protein function, such as substrates or binding partners, represents a new and immediate means for controlling intein-containing protein function in microorganisms.

Materials and methods

Plasmid and strain construction

N-terminal deletion (codons 1–115) plasmids p Δ N-RadAi and p Δ N-RadAi-AA were constructed from plasmids pFL-RadAi and pFL-RadAi-AA (Topilina et al. 2015b), respectively, by inverse PCR (primers 3557/3558). Following amplification, PCR products were digested with BamHI and ligated using T4 DNA ligase. Plasmid p Δ N-RadAi-R503A was constructed using the QuikChange Lightning site-directed mutagenesis kit (Agilent) with primers 3674/3675. To construct pXI-INT-RadA, pFL-RadAi was used to amplify the intein (primers 4172/4173), and pXI Δ Alh (Hiraga et al. 2005) was EcoRI- and HindIII-treated and assembled with In-Fusion HD (Clontech). AB1157lexA3(DE3) was made by P1 transduction of the lexA3zja::Tn10 allele from AB1157lexA3 into AB1157(DE3). Oligonucleotides (Supplemental Table 1), plasmids (Supplemental Table 2), and strains (Supplemental Table 3) are listed.

Protein expression and purification

Δ N-RadAi, Δ N-RadAi-R503A, and Δ N-RadAi-AA were expressed from pET45b(+) with an N-terminal his tag in BL21(DE3) or ArcticExpress (DE3). The RadA intein was expressed from the pXI-INT-RadA with in-frame chitin-binding domain (CBD) in JM109. Cells were grown in LB broth (Miller) at 37°C to ~0.4 OD at 600 nM, temperature was reduced to 16°C,

and 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was added, and cells were grown for 16–20 h and centrifuged at 4000g for 20 min.

After RadA expression, pellets were resuspended in Ni²⁺ buffer A (20 mM Tris-HCl at pH 8.5, 500 mM NaCl, 30 mM imidazole, 10% glycerol), lysed by sonication, and centrifuged at 20,700g for 20 min. Soluble protein was applied to HisTrap HP resin (GE Healthcare) in Ni²⁺ buffer A, washed with 20 column volumes of Ni²⁺ buffer A, and eluted with Ni²⁺ buffer B (20 mM Tris-HCl at pH 8.5, 500 mM NaCl, 500 mM imidazole). Proteins were concentrated to 4 mg/mL and exchanged into RadA storage buffer (RSB; 20 mM Tris-HCl at pH 8.5, 200 mM NaCl, 10% glycerol) by diafiltration using a 10-kDa concentrator (Amicon) or dialysis.

After expression of CBD-RadA intein, pellets were resuspended in chitin-binding buffer (CBB; 20 mM Tris-HCl at pH 8.0, 200 mM NaCl, 0.1 mM EDTA), lysed/centrifuged as for RadA above, applied to chitin resin (New England Biolabs), washed with 15 column volumes of CBB, washed with 5 column volumes of CBB with 400 mM NaCl, and eluted by on-column intein cleavage with 75 mM dithiothreitol in CBB for 72 h at 4°C. Intein was concentrated/exchanged into RSB. The MIG reporter was expressed from pMIG-RadAi as described (Topilina et al. 2015b).

In vitro splicing assays

One part *Pho* Δ N-RadAi (86 μ M in RSB) was diluted with either three parts Tris-EDTA (TE; 10 mM Tris-HCl at pH 8.0, 1 mM EDTA) or M13mp18 ssDNA in 250 ng/ μ L TE (New England Biolabs). dNTPs, pKW08-Lx dsDNA (Addgene), or M13mp18 RFI dsDNA (New England Biolabs) and cellular RNA were added at the same concentration as ssDNA, as was done when testing splicing of the *Pho* RadA intein in MIG. MgCl₂ was added as indicated. ATP was added at 2 mM. For testing FL-RadAi response to ssDNA, FL-RadAi in RSB was mixed with an equal volume of either TE or M13mp18 ssDNA.

Splicing time courses were initiated by incubation at the temperatures indicated, and samples were taken at various times, mixed with SDS loading dye but not boiled, separated using 8%–16% nonreducing Mini-Protean TGX SDS-PAGE gels (Bio-Rad), stained with Coomassie, scanned, and analyzed using ImageJ (<https://imagej.nih.gov/ij>). Splicing was calculated as $[(LE + I/LE + I + I-C + P)_{tX} - (LE + I/LE + I + I-C + P)_{t0}] / [1 - (LE + I/LE + I + I-C + P)_{t0}]$, where LE is ligated exteins, P is precursor, I is intein, I-C is intein-C-extein, t0 is time zero, and tX is a time at a given point. New LEs was calculated by $(LE_{tX} - LE_{t0}) / LE_{t0}$, where LE_{tX} is LEs at a given point and LE_{t0} is LEs at time 0. The maximum new LEs was estimated as P_{t0} / LE_{t0} , which was 1.67 for Δ N-RadAi, and is the amount of LEs that could be made if all precursors at t0 were converted to LEs while accounting for the LEs present at t0.

DNA binding

Δ N-RadAi, Δ N-RadAi-AA, Δ N-RadAi-R503A, and free intein were incubated with M13mp18 ssDNA for 1 h at 63°C in four parts TE with one part RSB. An equal volume of 83% glycerol and 2 \times EZ Vision (Amresco) was added and run on 1% agarose gels in 1 \times Tris-acetate-EDTA at the ssDNA and protein concentrations indicated. ssDNA was visualized under UV light. Apparent K_d values were calculated using Prism7 (GraphPad), accounting for ligand depletion.

In vivo splicing following DNA damage

BL21(DE3) and BLR(DE3) (*recA*⁻) or AB1157(DE3) and AB1157lexA3(DE3) were used as isogenic pairs except for the *recA* and *lexA* alleles (Supplemental Table 3). Cells with p Δ N-RadAi were grown at 37°C to ~0.4 OD at 600 nM and split in two for each strain. To one culture, 200 μ g/mL Nal and 0.3 mM IPTG were added. Nal was freshly prepared at 30 mg/mL in 0.3 M NaOH. To the other culture, 0.3 mM IPTG and an equivalent volume of 0.3M NaOH were added. Cultures were grown at 37°C, and cells were harvested by centrifugation at the times indicated and lysed with B-PER (Thermo) plus 12.5 U/mL DNase I. Soluble His-tagged proteins were isolated with HisPur Ni-NTA magnetic beads (Thermo), washed with Ni²⁺ buffer A, eluted with Ni²⁺ buffer B, and analyzed using 12% SDS-PAGE gels as above. Splicing was calculated as $LE / (LE + P)$. When treated with UV, the dose was 0.6 J/cm². DNA copurification was analyzed using UV spectroscopy after 2 h of Nal treatment.

Acknowledgments

We thank Carol Lyn Piazza and Natalya Topilina for construction of p Δ N-RadAi and p Δ N-RadAi-AA, Dorie Smith for construction of ABL157lexA3 (DE3), and Rebecca McCarthy for manuscript help. This work was supported by National Institutes of Health grants GM39422 and GM44844 to M.B., and F32GM121000 to C.W.L.

References

- Aranko AS, Oeemig JS, Kajander T, Iwai H. 2013. Intermolecular domain swapping induces intein-mediated protein alternative splicing. *Nat Chem Biol* **9**: 616–622.
- Bell JC, Kowalczykowski SC. 2016. RecA: regulation and mechanism of a molecular search engine. *Trends Biochem Sci* **41**: 491–507.
- Callahan BP, Topilina NI, Stanger MJ, Van Roey P, 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.
- Chen LT, Ko TP, Chang YC, Lin KA, Chang CS, Wang AH, Wang TF. 2007. Crystal structure of the left-handed archaeal RadA helical filament: identification of a functional motif for controlling quaternary structures and enzymatic functions of RecA family proteins. *Nucleic Acids Res* **35**: 1787–1801.
- Chen Z, Yang H, Pavletich NP. 2008. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* **453**: 489–484.
- Chow SA, Radding CM. 1985. Ionic inhibition of formation of RecA nucleoprotein networks blocks homologous pairing. *Proc Natl Acad Sci* **82**: 5646–5650.
- Conway AB, Lynch TW, Zhang Y, Fortin GS, Fung CW, Symington LS, Rice PA. 2004. Crystal structure of a Rad51 filament. *Nat Struct Mol Biol* **11**: 791–796.
- Cox MM. 2007. Regulation of bacterial RecA protein function. *Crit Rev Biochem Mol Biol* **42**: 41–63.
- Galkin VE, Wu Y, Zhang XP, Qian X, He Y, Yu X, Heyer WD, Luo Y, Egelman EH. 2006. The Rad51/RadA N-terminal domain activates nucleoprotein filament ATPase activity. *Structure* **14**: 983–992.
- Hiraga K, Derbyshire V, Dansereau JT, Van Roey P, Belfort M. 2005. Minimization and stabilization of the *Mycobacterium tuberculosis* recA intein. *J Mol Biol* **354**: 916–926.
- Kohiyama M, Contremoulins V, Baudin X. 2013. Trashing of single-stranded DNA generated during processing of arrested replication fork in *E. coli*. *J Mol Biol* **425**: 4837–4844.
- Komori K, Miyata T, Daiyasu H, Toh H, Shinagawa H, Ishino aY. 2000a. Domain analysis of an archaeal RadA protein for the strand exchange activity. *J Biol Chem* **275**: 33791–33797.
- Komori K, Miyata T, DiRuggiero J, Holley-Shanks R, Hayashi I, Cann IK, Mayanagi K, Shinagawa H, Ishino Y. 2000b. Both RadA and RadB are involved in homologous recombination in *Pyrococcus furiosus*. *J Biol Chem* **275**: 33782–33790.
- Krejci L, Altmannova V, Spirek M, Zhao X. 2012. Homologous recombination and its regulation. *Nucleic Acids Res* **40**: 5795–5818.
- Lin Z, Kong H, Nei M, Ma H. 2006. Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc Natl Acad Sci* **103**: 10328–10333.
- Mills KV, Johnson MA, Perler FB. 2014. Protein splicing: how inteins escape from precursor proteins. *J Biol Chem* **289**: 14498–14505.
- Naor A, Altman-Price N, Soucy SM, Green AG, Mitiagin Y, Turgeman-Grott I, Davidovich N, Gogarten JP, Gophna U. 2016. Impact of a homing intein on recombination frequency and organismal fitness. *Proc Natl Acad Sci* **113**: E4654–E4661.
- Novikova O, Jayachandran P, Kelley DS, Morton Z, Merwin S, Topilina NI, Belfort M. 2016. Intein clustering suggests functional importance in different domains of life. *Mol Biol Evol* **33**: 783–799.
- Oeemig JS, Zhou D, Kajander T, Wlodawer A, Iwai 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.
- Prasad TK, Yeykal CC, Greene EC. 2006. Visualizing the assembly of human Rad51 filaments on double-stranded DNA. *J Mol Biol* **363**: 713–728.
- Reitter JN, Cousin CE, Nicastrì MC, Jaramillo MV, Mills KV. 2016. Salt-dependent conditional protein splicing of an intein from *Halobacterium salinarum*. *Biochemistry* **55**: 1279–1282.
- Rolfmeier ML, Laughery MF, Haseltine CA. 2010. Repair of DNA double-strand breaks following UV damage in three *Sulfolobus solfataricus* strains. *J Bacteriol* **192**: 4954–4962.
- San Filippo J, Sung P, Klein H. 2008. Mechanism of eukaryotic homologous recombination. *Annu Rev Biochem* **77**: 229–257.
- Seitz EM, Brockman JP, Sandler SJ, Clark AJ, Kowalczykowski SC. 1998. RadA protein is an archaeal RecA protein homolog that catalyzes DNA strand exchange. *Genes Dev* **12**: 1248–1253.
- Seitz EM, Haseltine CA, Kowalczykowski SC. 2001. DNA recombination and repair in the archaea. *Adv Appl Microbiol* **50**: 101–169.
- Shah NH, Muir TW. 2014. Inteins: nature's gift to protein chemists. *Chem Sci* **5**: 446–461.
- Shin DS, Pellegrini L, Daniels DS, Yelent B, Craig L, Bates D, Yu DS, Shivji MK, Hitomi C, Arvai AS, et al. 2003. Full-length archaeal Rad51 structure and mutants: mechanisms for RAD51 assembly and control by BRCA2. *EMBO J* **22**: 4566–4576.
- Topilina NI, Green CM, Jayachandran P, Kelley DS, Stanger MJ, Piazza CL, Nayak S, Belfort M. 2015a. SufB intein of *Mycobacterium tuberculosis* as a sensor for oxidative and nitrosative stresses. *Proc Natl Acad Sci* **112**: 10348–10353.
- Topilina NI, Novikova O, Stanger M, Banavali NK, Belfort M. 2015b. Post-translational environmental switch of RadA activity by extein-intein interactions in protein splicing. *Nucleic Acids Res* **43**: 6631–6648.
- Wu Y, He Y, Moya IA, Qian X, Luo Y. 2004. Crystal structure of archaeal recombinase RADA: a snapshot of its extended conformation. *Mol Cell* **15**: 423–435.