

10-7-2021

## **Intein Inhibitors as Novel Antimicrobials: Protein Splicing in Human Pathogens, Screening Methods, and Off-target Considerations**

Diana A. Wall  
*College of the Holy Cross*

Chunyu Wang  
*Rensselaer Polytechnic Institute*

Kenneth Mills  
*College of the Holy Cross*

Christopher W. Lennon  
*Murray State University, clennon1@murraystate.edu*

Follow this and additional works at: <https://digitalcommons.murraystate.edu/faculty>



Part of the [Biology Commons](#)

---

### **Recommended Citation**

Wall, D. A., Tarrant, S. P., Wang, C., Mills, K., & Lennon, C. W. Intein Inhibitors as Novel Antimicrobials: Protein Splicing in Human Pathogens, Screening Methods, and Off-target Considerations. *Frontiers in Molecular Biosciences*, 975.

This Journal Article is brought to you for free and open access by the Faculty Works at Murray State's Digital Commons. It has been accepted for inclusion in Faculty & Staff Research and Creative Activity by an authorized administrator of Murray State's Digital Commons. For more information, please contact [msu.digitalcommons@murraystate.edu](mailto:msu.digitalcommons@murraystate.edu).



# Intein Inhibitors as Novel Antimicrobials: Protein Splicing in Human Pathogens, Screening Methods, and Off-Target Considerations

Diana A. Wall<sup>1</sup>, Seanan P. Tarrant<sup>1</sup>, Chunyu Wang<sup>2,3</sup>, Kenneth V. Mills<sup>1\*</sup> and Christopher W. Lennon<sup>4\*</sup>

<sup>1</sup>Department of Chemistry, College of the Holy Cross, Worcester, MA, United States, <sup>2</sup>Department of Biological Sciences, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY, United States, <sup>3</sup>Department of Chemistry and Chemical Biology, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY, United States, <sup>4</sup>Department of Biological Sciences, Murray State University, Murray, KY, United States

## OPEN ACCESS

### Edited by:

Venkaiah Betapudi,  
United States Department of  
Homeland Security, United States

### Reviewed by:

Saroj Nayak,  
Indian Institute of Technology  
Bhubaneswar, India  
Sasmita Nayak,  
KIIT University, India  
Tsuyoshi Takahashi,  
Gunma University, Japan  
Neel H Shah,  
Columbia University, United States

### \*Correspondence:

Christopher W. Lennon  
clennon1@murraystate.edu  
Kenneth V. Mills  
kmills@holycross.edu

### Specialty section:

This article was submitted to  
Cellular Biochemistry,  
a section of the journal  
Frontiers in Molecular Biosciences

**Received:** 03 August 2021

**Accepted:** 24 September 2021

**Published:** 07 October 2021

### Citation:

Wall DA, Tarrant SP, Wang C, Mills KV  
and Lennon CW (2021) Intein Inhibitors  
as Novel Antimicrobials: Protein  
Splicing in Human Pathogens,  
Screening Methods, and Off-  
Target Considerations.  
Front. Mol. Biosci. 8:752824.  
doi: 10.3389/fmolb.2021.752824

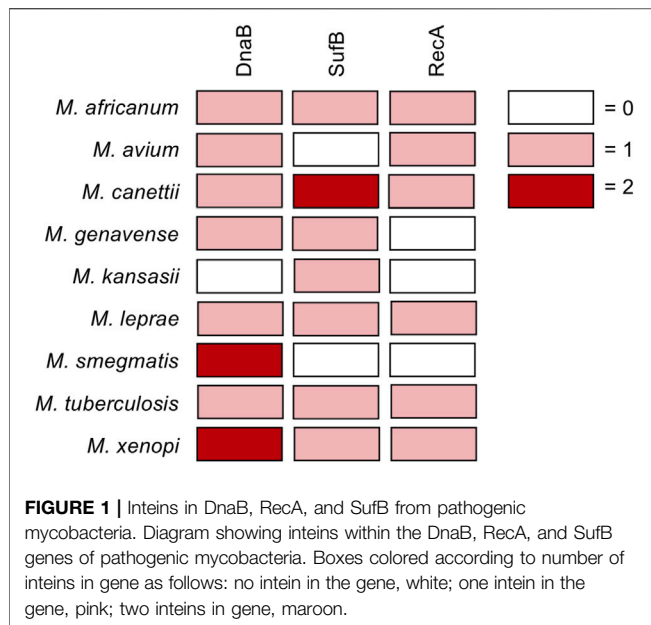
Protein splicing is a post-translational process by which an intervening polypeptide, or intein, catalyzes its own removal from the flanking polypeptides, or exteins, concomitant with extein ligation. Although inteins are highly abundant in the microbial world, including within several human pathogens, they are absent in the genomes of metazoans. As protein splicing is required to permit function of essential proteins within pathogens, inteins represent attractive antimicrobial targets. Here we review key proteins interrupted by inteins in pathogenic mycobacteria and fungi, exciting discoveries that provide proof of concept that intein activity can be inhibited and that this inhibition has an effect on the host organism's fitness, and bioanalytical methods that have been used to screen for intein activity. We also consider potential off-target inhibition of hedgehog signaling, given the similarity in structure and function of inteins and hedgehog autoprocessing domains.

**Keywords:** intein, protein splicing, novel drug target, inhibitor screening, tuberculosis, hedgehog signaling, hedgehog autoprocessing

## INTRODUCTION

Inteins are intervening polypeptides translated within host proteins that are removed through protein splicing. Although inteins are widespread in the microbial world (Novikova et al., 2016), they are not present in the genomes of metazoans. Given that several devastating human pathogens house inteins within proteins essential for survival, and intein removal is considered necessary to permit host protein function, inteins represent attractive novel drug targets (Gangopadhyay et al., 2003; Paulus, 2003; Grainger and Beggs, 2005).

Of particular interest are mycobacterial pathogens including *Mycobacterium tuberculosis*, *Mycobacterium leprae*, and non-tuberculosis mycobacteria (NTM), which house inteins within important genes such as DnaB, RecA, SufB, and others (Figure 1). Prior to the recent Covid-19 pandemic, tuberculosis was the leading cause of death from an infectious agent worldwide, claiming 1.4 million deaths in 2019 (2020b). Additionally, a number of fungal pathogens such as *Cryptococcus neoformans*, *Cryptococcus gattii*, *Aspergillus fumigatus*, *Aspergillus nidulans*, and *Histoplasma capsulatum* also harbor inteins, most often in the essential spliceosomal protein Prp8 (Green



et al., 2019). *C. neoformans* infections pose a particular threat to immunocompromised individuals such as those with HIV/AIDS, with an estimated 220,000 cases of cryptococcal meningitis occurring worldwide each year (Centers for Disease Control and Prevention, 2020).

Inteins are found in all three domains of life, as well in viruses, with the highest abundance within archaea and bacteria, with about 47 and 24% of species harboring inteins, respectively (Novikova et al., 2016). Only about 1% of the eukarya have inteins (Novikova et al., 2016). There is considerable evidence, although in some cases indirect, that supports the idea that blocking protein splicing will compromise host protein function. For example, inteins often localize to critical regions, such as the active site, of essential proteins involved in DNA replication, recombination, and repair (Novikova et al., 2014; Kelley et al., 2016; Novikova et al., 2016). If the intein is within an essential gene, as many are, inhibition of splicing should result in killing or prevention of reproduction of the host organism. However, as inteins have historically been studied in splicing reporters to study activity and not within the native extein context, there are not many examples of studies that monitor intein influence on host protein function; this is important future work needed in the field. However, some direct evidence of intein inhibition influencing host protein function does exist. One example of the link between splicing and host protein function is found in the intein-interrupted SufB protein of *M. tuberculosis*. SufB is part of the essential SUF [Fe-S] cluster biogenesis complex and interacts with SufC and SufD to form the SufBCD. Prior to splicing, the SufB intein was shown to block interactions with SufC and SufD (Huet et al., 2006). In the case of intein-interrupted RadA from the archaeon *Pyrococcus horikoshii*, which harbors an intein within the ATP-coordinating P-loop, ATPase activity is compromised prior to protein splicing (Topilina et al., 2015b). Recently, exciting work has demonstrated that selective inhibition of inteins interrupting

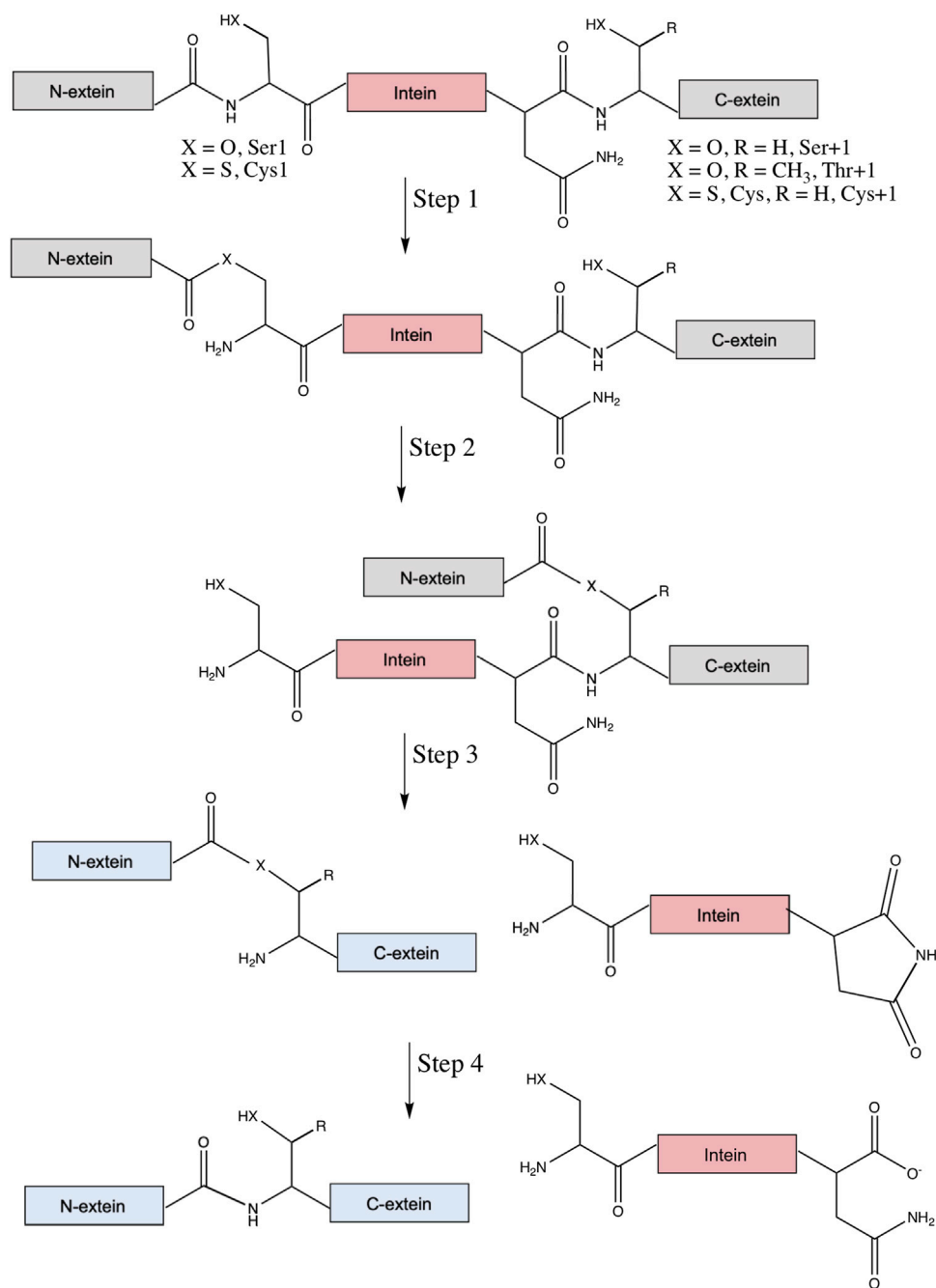
the Prp8 protein in intein-containing fungal pathogens can influence survival of the organism (discussed below; (Li et al., 2021).

Here, we will discuss inteins as potential targets to selectively kill bacterial and fungal pathogens, along with the screening assays that could be used to test the ability of small molecules to inhibit protein splicing. There is proof of concept evidence that divalent cations (in particular zinc and copper), reactive oxygen, nitrogen, and chlorine derivatives, as well as small molecules including cisplatin and diethanolamine can act as splicing inhibitors and impact the fitness of pathogenic mycobacteria and fungi (Gangopadhyay et al., 2003; Zhang et al., 2010; Zhang et al., 2011; Dasari et al., 2015; Topilina et al., 2015a; Kelley et al., 2018; Li et al., 2019; Lennon et al., 2021; Panda et al., 2021) The discovery of drug-like molecules that can successfully inhibit the splicing of proteins essential for survival of pathogens could represent new treatment options. Because inteins are absent in humans, specific intein inhibitors could have limited side effects. The recent discovery of a small molecule inhibitor of the Prp8 intein that selectively inhibits splicing without targeting proteases and inhibits the growth of intein-containing *C. gattii* and *C. neoformans* is a particularly exciting breakthrough that suggests intein inhibitors may have therapeutic benefits (Li et al., 2021).

Reporter systems have been developed to detect intein activity and could be used as screening assays for potential drug leads. Here, we will discuss screening systems involving kanamycin resistance, green fluorescent protein (GFP), luciferase, CcdB/CcdA, thymidylate synthetase (TS), and  $\beta$ -galactosidase, and evaluate how they could be adapted to a drug screening system (Daugelat and Jacobs, 1999; Wood et al., 1999; Adam and Perler, 2002; Lew and Paulus, 2002; Gangopadhyay et al., 2003; Paulus, 2003; Buskirk et al., 2004; Zeidler et al., 2004; Bradley et al., 2005; Wong et al., 2005; Hiraga et al., 2009; Liang et al., 2011; Zhang et al., 2011; Chan et al., 2016; Neugebauer et al., 2017; Stevens et al., 2017; Kelley et al., 2018; Woods et al., 2020). Additionally, inteins and the metazoan Hedgehog (Hh) family of proteins are both HINT (Hedgehog-INTEin) domain proteins with a similar fold (Porter et al., 1996a; Porter et al., 1996b; Hall et al., 1997). We therefore consider Hh autoprocessing, as the cholesterification process is similar mechanistically to protein splicing. Although the Hh pathway is relatively inactive in adults outside of stem cell maintenance, inhibition of the pathway could be a problematic unintended target of a potential intein-directed drug (Rimkus et al., 2016). On the other hand, autoprocessing inhibitors could be useful as a cancer therapy in certain basal cell carcinomas or medulloblastomas (Gupta et al., 2010).

## MECHANISM OF PROTEIN SPLICING, INTEIN VARIETY, AND POSSIBLE ROLE AS ENVIRONMENTAL SENSORS

Small molecules that target intein action have so far been shown to modify the active site nucleophiles of the intein active site. In this section we briefly review the mechanism of protein splicing, and in the sections that follow we discuss in detail how these steps are affected by known intein inhibitors.



**FIGURE 2 |** Canonical mechanism of protein splicing. Protein splicing is a post-translational process where an intein (in pink) interrupts the N- and C-exteins (in gray), which usually inactivates the extein's native function. Step 1 is an amide-to-ester (or thioester) rearrangement of the peptide bond linking the N-extein and intein, facilitated by the N-terminal Ser1 or Cys1 of the intein ( $X = O$  for Ser,  $S$  for Cys). Step 2 is a trans-thioesterification reaction that transfers the N-extein from the side chain of the first intein residue to the side chain of the first C-extein residue, either Ser+1, Thr+1, or Cys+1 ( $X = O, R = H$  for Ser,  $X = S$  and  $R = H$  for Cys, and  $X = O$  and  $R = CH_3$  for Thr). Step 3 is cyclization of the C-terminal Asn (or sometimes Gln) of the intein coupled to peptide bond cleavage between the intein and C-extein. Linkage of the exteins and excision of the intein likely restores function to the exteins (now colored blue). Finally, step four converts the ester bond linking the exteins to the amide peptide bond and, in some cases, converts the C-terminal aminosuccinimide of the excised intein to Asn (or *iso*-Asn).

Key conserved residues in most inteins include an N-terminal Cys or Ser and a C-terminal Asn. Inteins also require a Cys, Ser, or Thr as the first residue following the intein. Via protein splicing,

inteins can self-catalyze their own excision from the two flanking polypeptide sequences, or exteins, to form a mature, functional host protein. The intein residues are numbered such that the

N-terminal residue is Cys1 or Ser1 and continuing downstream, with the first C-extein residue as +1 and the last N-extein residue as -1.

The canonical mechanism of protein splicing is a four-step process (Figure 2). First, the amide bond that connects the N-extein to the intein is rearranged to an ester or thioester, with the N-terminal Ser or Cys of the intein as the nucleophile. Then, transesterification via nucleophilic attack by the first residue of the C-extein (Cys, Ser, or Thr) on the ester formed in step one results in a branched intermediate, with the N-extein moved from the side chain of the first intein residue to the side chain of the first C-extein residue. These reactive nucleophiles are the target of most intein inhibitors discussed below. It will be essential that future lead compounds are selected that bind selectively to inteins. Given the long history of the development of selective protease inhibitors in therapy (Sojka et al., 2021), and that inteins promote similar chemical steps despite serving as their own substrate (Paulus, 2001), it seems possible that traditional pharmacological strategies would be useful for increasing selectivity of intein inhibitors as well.

Next, in the third step of splicing the branched ester or thioester is resolved by asparagine cyclization coupled to peptide bond cleavage between the intein and C-extein. The final step converts the ester bond linking the exteins to an amide bond, as well as hydrolysis of the C-terminal aminosuccinimide of the intein to Asn or *iso*-Asn (Mills et al., 2014; Shah and Muir, 2014). Two alternative classes of splicing also exist, with key differences reviewed elsewhere (Mills et al., 2014). Irreversible off-pathway reactions can also occur, where either the N- or C-extein is cleaved from the intein prior to ligation. With the ability of inteins to rearrange and cleave peptide bonds, they have been utilized extensively in biotechnology, a topic covered in several other comprehensive reviews (Volkman and Mootz, 2013; Shah and Muir, 2014; Topilina and Mills, 2014; Wood and Camarero, 2014; Debelouchina and Muir, 2017; Sarmiento and Camarero, 2019; Nanda et al., 2020; Romero-Casanas et al., 2020).

Inteins come in a few different varieties. Full-length inteins contain a nested homing endonuclease (HE) domain within the autoprocessing domain. The HE domain can facilitate intein homing into intein-less alleles by making a double strand break in the extein sequence that lacks the interrupting intein gene, followed by DNA repair via homologous recombination using the intein-containing allele as a template. Mini-inteins, which lack the HE domain but retain the autoprocessing domain, are also abundant. In both cases, these contiguous inteins facilitate splicing in *cis* (Mills et al., 2014). In contrast, some inteins are naturally split, with the autoprocessing domain expressed as two separate polypeptides that must first assemble before protein splicing occurs in *trans* (Volkman and Iwai, 2010; Aranko et al., 2014; Li, 2015).

Although inteins don't require any cofactor or chaperone to splice (Mills et al., 2014), the rate and accuracy of the reaction can be highly dependent on environmental conditions, including redox state, pH, salt concentration, and temperature (Belfort, 2017; Lennon and Belfort, 2017). As many inteins also are mobile genetic elements, they have traditionally been viewed as selfish

molecular parasites (Gogarten and Hilario, 2006). However, several recent studies have provided evidence that some inteins may play a beneficial role for the host organism by providing post-translational control of key enzymes in response to stress (Belfort, 2017; Lennon and Belfort, 2017).

## INTEIN PRESENCE IN HUMAN PATHOGENS AND SPLICING INHIBITION

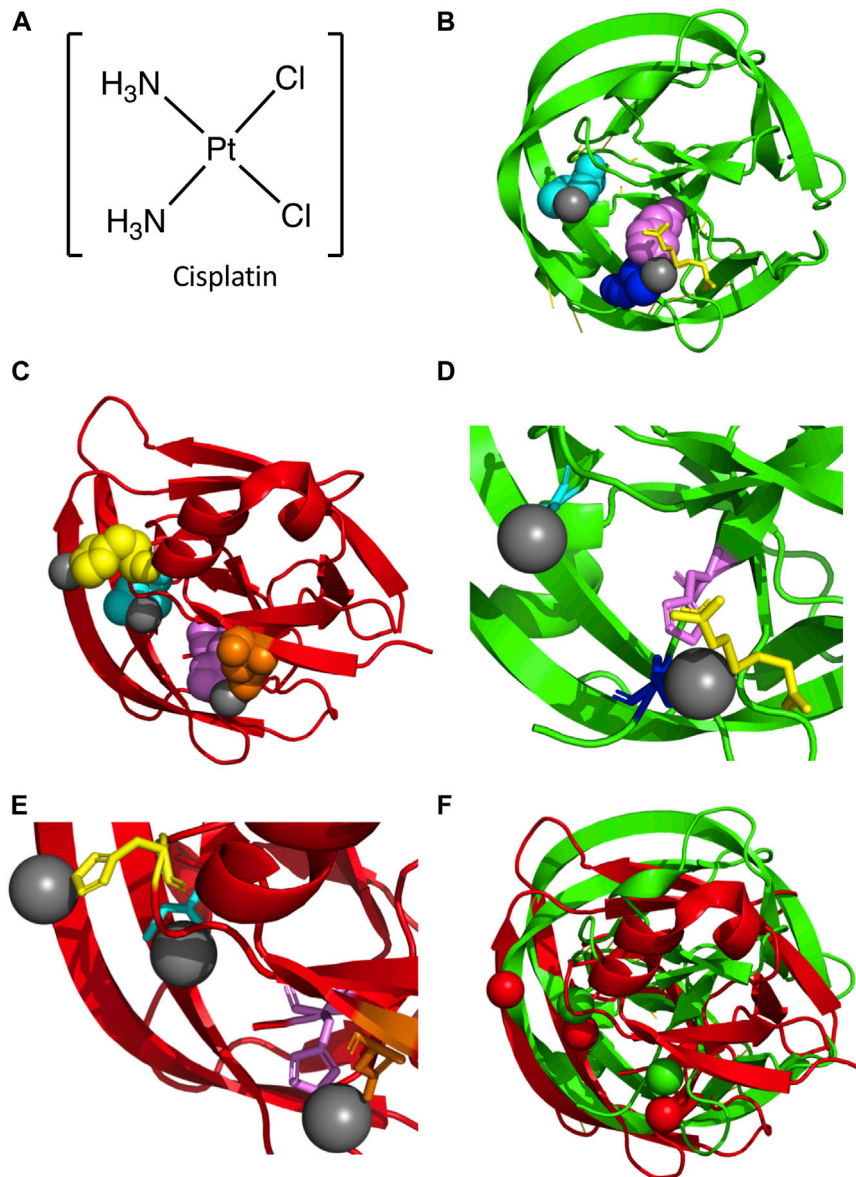
Inteins are found in the genomes of archaea, bacteria, single-celled eukaryotes, and viruses, often within proteins required for survival. Here we discuss the mycobacterial inteins within RecA, DnaB, and SufB (Figure 1A), as their splicing is important for the survival of pathogenic mycobacteria. Additionally, we consider Prp8 inteins present in fungal pathogens including *Cryptococcus neoformans*. We also discuss the chemical means by which inhibitors of protein splicing interact with their target intein.

### Mycobacterial Inteins

Mycobacteria are rod-shaped, acid-fast bacteria. Many species are human pathogens, most notably *M. tuberculosis* and *M. leprae*, as well as non-tuberculosis mycobacteria (NTM; i.e. mycobacteria other than *M. tuberculosis* or *M. leprae* that cause disease in humans). Tuberculosis remains a leading cause of death worldwide, causing an estimated 1.4 million deaths in 2019, with multidrug-resistant (MDR) strains becoming increasingly common (Paulus, 2003; World Health Organization, 2020). The World Health Organization has deemed it a priority to address MDR tuberculosis and emerging extensively drug-resistant tuberculosis (Kempker et al., 2015). In 2018, there were ~209,000 new cases of leprosy, a disease that can cause permanent damage to the skin, nerves, limbs, and eyes if untreated. Emerging drug resistance of *M. leprae* has also been detected (da Silva Rocha et al., 2012). NTM, while less common and generally considered opportunistic pathogens, cause a range of serious diseases and are increasing in incidence (Shamaei and Mirsaedi, 2021). Mycobacteria are rich in inteins, particularly in three important genes critical for the growth of *M. tuberculosis*: DnaB, SufB, and RecA (Figure 1) (Paulus, 2003). Related mycobacterial species also have inteins present in these same genes, yet the location and sequence can vary. If a small molecule is able to inhibit splicing of these proteins, the mature form of the protein will not be produced and could prevent infection. Therefore, it is of great interest and urgency to investigate inteins within these mycobacterial pathogens as targets for compounds that specifically block splicing. The potential role of inhibiting the splicing of the RecA and DnaB inteins in *M. tuberculosis* has previously been reviewed (Paulus, 2003), and these inteins have subsequently been the subject of investigation both *in vitro* as well as *in vivo* in the mycobacteria.

### Recombinase A Inhibition

One possible mode of mycobacterial inhibition is targeting the intein in DNA Recombinase A (RecA), a protein necessary for the repair and maintenance of DNA in mycobacteria. The *M. tuberculosis* RecA intein is one of the best studied inteins, used



**FIGURE 3 |** Platinum-bound Intein Structures. **(A)** Structure of cisplatin. **(B)** Structure of the *M. tuberculosis* RecA intein bound to two Pt(II) ions (PDB510A; (Chan et al., 2016). Protein backbone in red and Pt(II) in gray. Key residues are colored and represented in space fill, with Cys1 (cyan), His439 (magenta), and Cys+1 (blue) shown. Additionally, TCEP is shown in yellow. **(C)** Structure of the *C. gattii* Prp8 intein in complex with three Pt(II) ions (PDB6MYL; (Li et al., 2019). Protein backbone in red and Pt(II) in gray. Residues important in coordination of Pt(II) are shown as space fill and colored, with Cys1 (cyan), His62 (yellow), Asp94 (orange), and His169 (magenta) shown. **(D)** Zoom in of *M. tuberculosis* RecA intein co-crystal structure bound to two Pt(II). Colored as in **(B)**, with Cys1, His439, and Cys+1 represented as sticks. **(E)** Zoom in of *C. gattii* Prp8 intein structure in complex with three Pt(II). Colored as in panel C, with Cys1, His62, Asp94, and His169 represented as sticks. **(F)** Alignment of the *M. tuberculosis* RecA intein complex with two Pt(II) and the *C. gattii* Prp8 intein bound to three Pt(II). The *M. tuberculosis* RecA intein protein backbone and Pt(II) are colored green, while the *C. gattii* Prp8 intein and Pt(II) are colored red.

in multiple mechanistic studies and several engineering applications (Davis et al., 1991; Davis et al., 1992; Mills et al., 1998; Shingledecker et al., 1998; Frischkorn et al., 2000; Saves et al., 2000; Shingledecker et al., 2000; Hiraga et al., 2005; Van Roey et al., 2007; Zhang et al., 2009; Du et al., 2010; Zhang et al., 2010). It has been found that copper and zinc ions inhibit protein splicing. Copper (II) ions inhibit splicing of a minimized RecA intein by binding to the catalytic Cys1 and block B His residue,

and may have a redox effect (Zhang et al., 2010). Zinc ions can also inhibit splicing of the full length and minimized RecA cis-splicing inteins as well as a *trans*-splicing variant (Mills and Paulus, 2001; Zhang et al., 2010). Zinc ions have been shown to bind to the C-terminal active site residues of the *Synechocystis* sp. PCC6803 DnaE and *S. cerevisiae* PI-SceI inteins (Poland et al., 2000; Sun et al., 2005). Cisplatin inhibits the splicing of the RecA intein by binding to Cys1, and other platinum-containing

compounds are also able to inhibit splicing *in vitro* (Figure 3A) (Zhang et al., 2011). Interestingly, over-expression of the intein was able to reverse the growth inhibition of *Mycobacterium bovis* by cisplatin, suggesting direct binding *in vivo* (Zhang et al., 2011). The crystal structure of a genetically minimized *M. tuberculosis* RecA intein engineered to lack its endonuclease domain ( $\Delta$ HE domain variant) in complex with cisplatin shows two dissociated platinum ions (Pt(II)) bound to residues required for intein catalysis (Chan et al., 2016) (Figures 3B, D). One Pt(II) binds directly to the catalytic Cys1 residue, with the other coordinated by the highly-conserved penultimate histidine (H439), Cys+1, and the reducing agent (tris(2-carboxyethyl)phosphine (TCEP) (Figures 3B, D). As these Pt(II) are coordinated directly by residues necessary for protein splicing, the mechanism of inhibition is relatively straightforward.

While these compounds can serve as effective inhibitors of *M. tuberculosis* growth through the inhibition of protein splicing, they do so most effectively at micromolar concentrations that may not be suitable for *in vivo* applications, making them more attractive as lead compounds (El-Kareh and Secomb, 2003). Additionally, cisplatin has significant side effects when used as a chemotherapeutic due to its ability to induce DNA damage; such side effects would likely not be tolerated for an antibiotic except perhaps for cancer patients also infected with tuberculosis (Dasari et al., 2015). However, this proof-of-concept inhibition could serve to inspire the development of selective inhibitors that target this intein.

## DnaB Inhibition

DnaB is an essential, intein-containing helicase required for replication. DnaB is also the most common intein-housing gene in bacteria (Novikova et al., 2016). In mycobacterial species housing DnaB inteins, some species contain one intein, while others contain two inteins. *Mycobacterium smegmatis* is a species with two DnaB inteins. The first intein, DnaBi1, is a mini-intein localized to the p-loop of the ATPase domain that splices by an alternative class three mechanism (Kelley et al., 2018). The second intein, DnaBi2, is a full-length, HE domain-containing intein. Based on intein sequence similarity and insertion site, DnaBi1 is similar to an intein from *M. leprae*, while DnaBi2 is similar to an intein from *M. tuberculosis* (Kelley et al., 2018). In *M. smegmatis*, DnaBi1 splicing *in vitro* is reversibly inhibited in oxidizing environments induced by high micromolar concentrations of hydrogen peroxide (Kelley et al., 2018). DnaBi1 splicing inhibition by hydrogen peroxide results in the formation of an intramolecular disulfide bond between the catalytic nucleophile (Cys118) and a structurally-adjacent cysteine (Cys48). Additionally, this intein binds zinc with apparent low micromolar affinity, which also reversibly blocks splicing (Woods et al., 2020). The crystal structure of DnaBi1 bound to zinc shows the ion coordinated by the Cys118, presumably blocking the ability of this catalytic nucleophile to initiate the splicing reaction, as well as with V119 and Y128 (Woods et al., 2020). The *M. leprae* DnaB intein also has been shown to be reversibly inhibited by zinc (Woods et al., 2020). Conversely, DnaBi2 splices quickly without being affected by oxidation.

Importantly, in addition to *in vitro* results, DnaBi1 splicing inhibition by hydrogen peroxide and zinc has been observed within *M. smegmatis* cells using a kanamycin intein splicing reporter (described below) (Kelley et al., 2018; Woods et al., 2020). These results are intriguing, both from the perspective of testing potential antibacterial agents, but also for understanding potential intein-based stress responses as they demonstrate that splicing inhibition by a small molecule can prevent extein function in mycobacteria. Interestingly, pathogenic mycobacteria encounter and can withstand hydrogen peroxide and metal stress when encountering mammalian innate immune cells such as neutrophils (Zahrt and Deretic, 2002; Wagner et al., 2005; Neyrolles et al., 2015; Weiss and Schaible, 2015; Gao et al., 2018).

## SufB Inhibition

Another intein of interest interrupts the Fe-S cluster assembly protein B, or SufB. SufB is part of a larger protein complex that functions to repair oxygen-labile iron-sulfur clusters under oxidative stress. As it is highly sensitive to oxidation and nitrosylation, this intein may help regulate SufB function. Splicing of the SufB intein is essential for *M. tuberculosis* and the assembly of the SufBCD complex, an essential system for iron-sulfur cluster assembly in mycobacteria (Huet et al., 2006). The intein Cys1 and Cys+1 residues have low pKa values, which enhances their nucleophilicity and makes them sensitive to oxidation, and thus can prevent splicing in oxidative environments (Topilina et al., 2015a). Environments containing oxidative elements such as hydrogen peroxide and diethanolamine inhibit SufB splicing, resulting in precursor accumulation or irreversible attenuation through N-terminal cleavage. Such oxidative inhibition leads to the formation of a disulfide bond between the catalytic Cys1 and Cys+1 residues, which inhibits normal splicing activity (Topilina et al., 2015a).

## Fungal Inteins

Several human fungal pathogens, such as *Cryptococcus* species, *Ajellomyces capsulatus*, and *Aspergillus* species, contain inteins, most often in the essential spliceosomal Prp8 protein (Green et al., 2019). These fungal species cause pulmonary infections in humans such as cryptococcosis, histoplasmosis, and aspergillosis, respectively, which pose a particular threat to immunosuppressed individuals (Bongomin et al., 2017).

## Pre-mRNA Processing Factor 8 Inhibition

Pre-mRNA processing factor 8 (Prp8) is a highly conserved nuclear protein ranging in size from 230 to 270 kDa. Prp8 has been identified as an essential mRNA splicing factor involved early in the mRNA splicing process (Grainger and Beggs, 2005). Prp8 is located in the catalytic core of the spliceosome complex, where pre-mRNA splicing occurs. In the spliceosome, Prp8 is a component of the U5 small nuclear ribonucleoprotein (sNRP) and U5•U4/U6 tri-sNRP and is believed to act as a cofactor in RNA catalysis.

Over 100 Prp8 inteins have been identified. There are both HE-containing inteins and mini-inteins, varying from 150 to over 800 residues (Butler et al., 2006). Many of the inteins occupy the

same insertion site, with a total of seven insertion sites identified (Green et al., 2019). Phylogenetic analysis of these inteins suggests that the first inteins were introduced into Prp8 over a billion years ago, with the most recent being introduced approximately 400 million years ago. Additionally, the inteins are phylogenetically clustered both by insertion site and by organism, suggesting that multiple, independent invasion events occurred throughout the evolution of this protein (Green et al., 2019).

The Prp8 intein is an exciting potential drug target given the essential role for this protein in RNA splicing (Liu and Yang, 2004). Divalent metals such as zinc and copper ions can inhibit the splicing of the *C. neoformans* intein. Though they do so in different ways, they both function by interfering with the catalytic N-terminal Cys1 of the intein. Copper ions reversibly oxidize the thiol of Cys1. Zinc ions have two intein binding sites, one at the C-terminal asparagine residue and the other at Cys1; binding at the Asn site may be tighter (Green et al., 2019). Similarly, cisplatin can bond with Cys1 and interfere with its activity (Li et al., 2019). The Prp8 intein of *C. neoformans* can be targeted by cisplatin, as with the *M. tuberculosis* RecA intein, which provides protection to mice from infection (Li et al., 2019). The cisplatin binds similarly at Cys1 in the Prp8 intein as it does in the RecA intein; however, a second cisplatin that binds at a downstream conserved active-site His residue may be more tightly coordinated by a second ligand (Li et al., 2019). The co-crystal structure of the *C. gattii* Prp8 intein with cisplatin (Li et al., 2019) has three dissociated Pt(II) molecules, as opposed to two Pt(II) for *M. tuberculosis* RecA (Figure 3C). As with *M. tuberculosis* RecA, one Pt(II) is bound to the catalytic Cys1 residue. Additionally, another Pt(II) is coordinated by the conserved penultimate histidine (H169) similar to *M. tuberculosis* RecA (Figures 3C, F). Unlike *M. tuberculosis* RecA, the *C. gattii* Prp8 intein binds a third Pt(II) at a non-conserved histidine (His62) (Figures 3C, F). Overall comparison with *M. tuberculosis* RecA shows considerable similarities in the overall alignment of the two structures (Figure 3D). In particular, the coordination of Pt(II) at both N- and C-terminal residues required for intein activity (Figure 3D). Furthermore, as with *M. smegmatis* DnaB1 (Woods et al., 2020), the inhibitor is bound directly by residues strictly required for protein splicing, explaining the mechanism of inhibition.

Recently, a landmark study on the use of inhibitors that target Prp8 protein splicing as a strategy to selectively kill pathogenic fungi was published (Li et al., 2021). In this work, small molecules identified from a library that inhibit the splicing of the Prp8 intein *in vitro*, in *E. coli* and in *C. neoformans*, reduced the growth of intein-containing *C. neoformans* and *C. gattii*, but not of the intein-less *Candida albicans* (Li et al., 2021). The inhibition is likely via a covalent modification to Cys1 of the Prp8 intein. This is a particularly compelling development in the exploration of inteins as novel antimicrobials, strongly supporting the viability of this strategy.

## SCREENING FOR INTEIN ACTIVITY

Selective small molecule inhibitors of protein splicing could be effective tools against several human pathogens. Ideally, a system

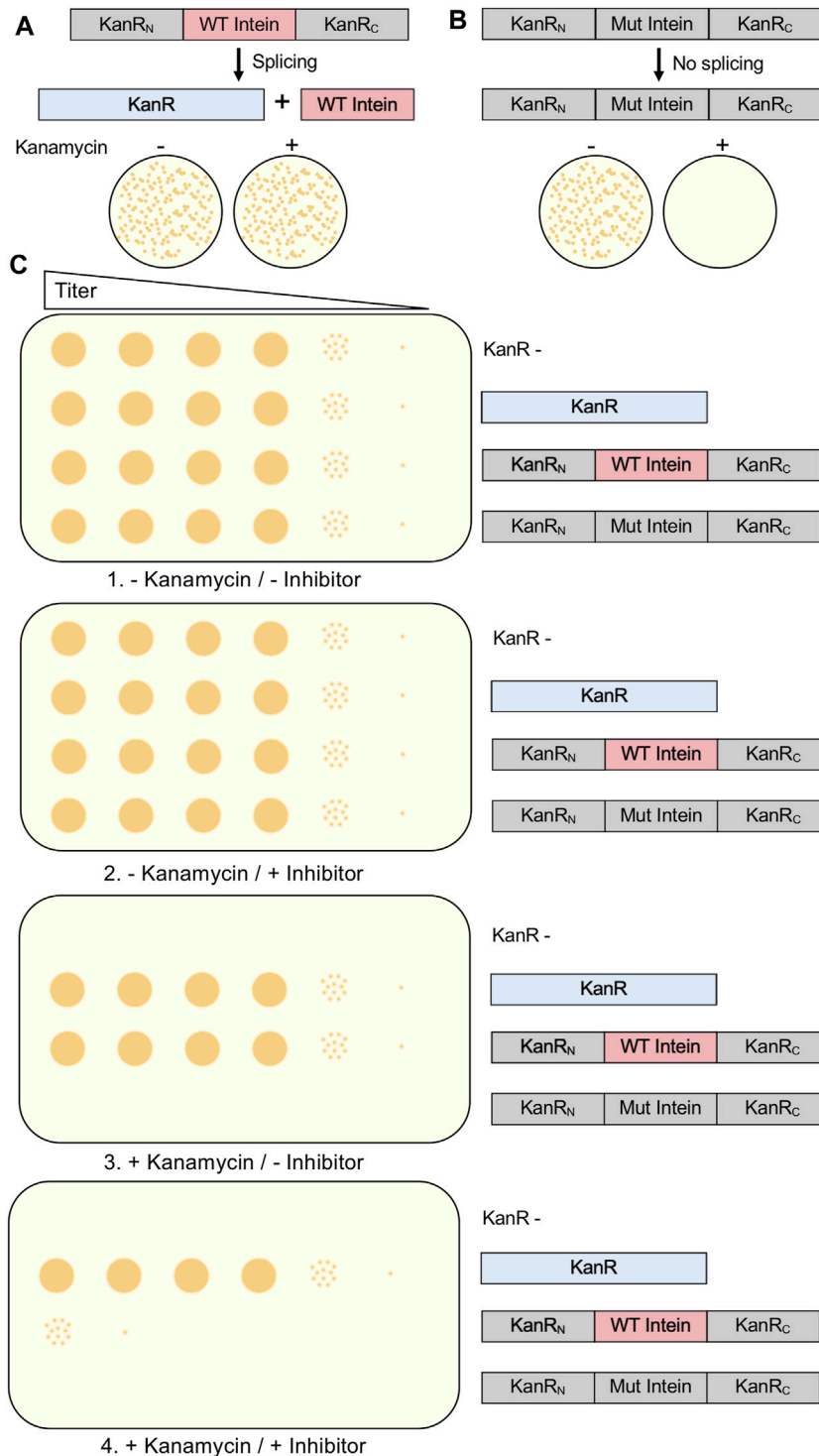
to screen for selective inhibitors would be high-throughput and could assay against inhibition of off-site targets or general cytotoxicity. Although many existing drugs are covalent mechanism-based inhibitors, it would be important to select for drugs specific to the intein rather than simply to a reactive Cys or Ser to prevent off-site toxicity. Selectable systems have been employed in directed evolution experiments to tune intein function, and split exteins have been used as reporters for splicing activity (Shah and Stevens, 2020). Several of these systems could be repurposed for drug screening. Below, we describe systems based on kanamycin resistance, GFP fluorescence and luciferase activity, CcdB/CcdA activity, thymidylate synthase activity, and  $\beta$ -galactosidase colorimetry.

## Aminoglycoside Phosphotransferase

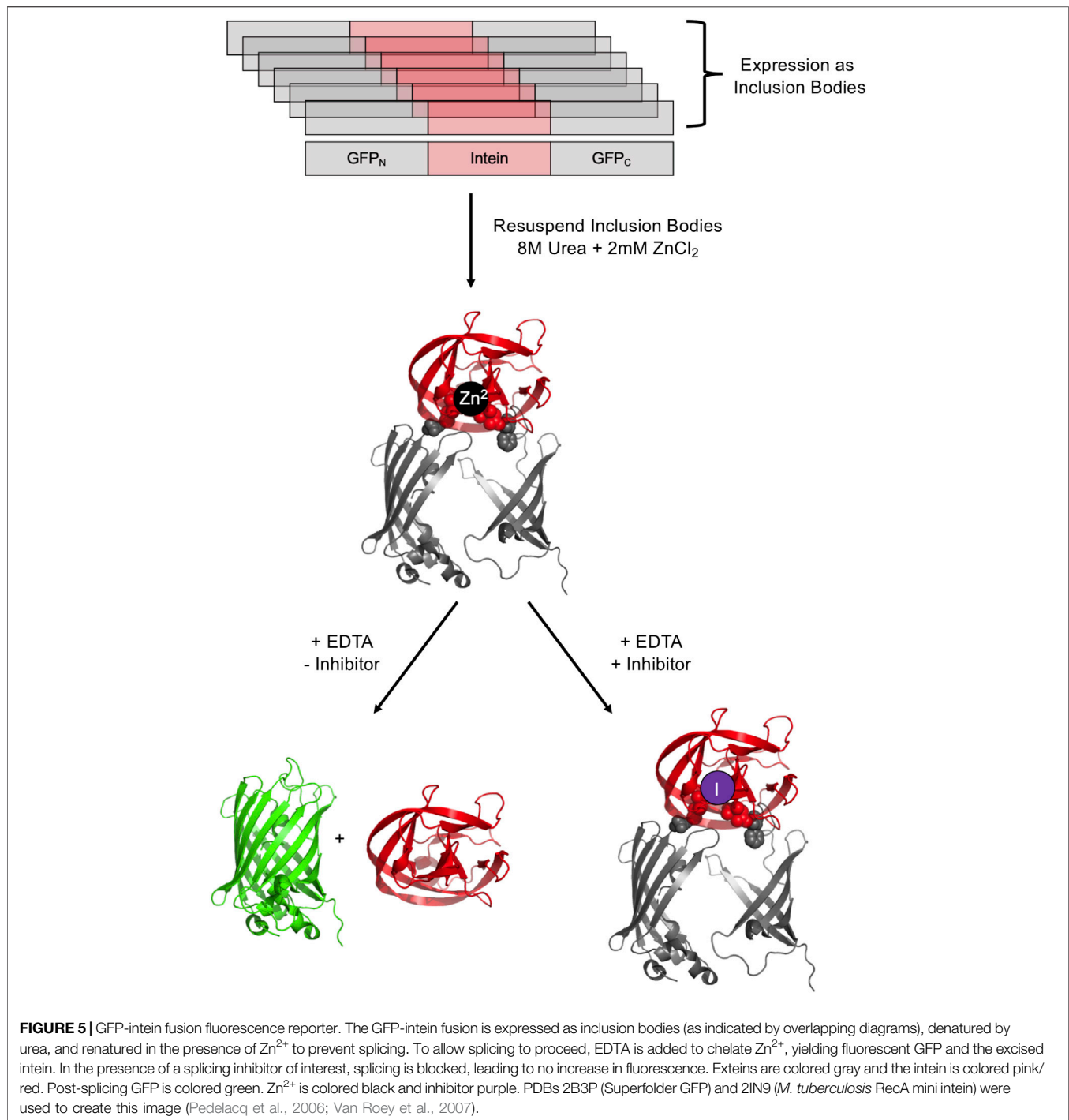
One method used to evaluate intein splicing and therefore serve as a screen for small molecule inhibitors is the use of the kanamycin resistance protein interrupted by an intein. Kanamycin A is an aminoglycoside antibiotic that inhibits translation and is widely used in molecular biology. Bacterial cells able to grow on media containing the antibiotic require a mature aminoglycoside phosphotransferase (KanR). Experimental cells are transformed to contain the KanR gene interrupted by an intein. Typically, the positive control is the uninterrupted KanR gene without an intein, as no splicing is required to produce resistance to kanamycin. When designing a kanamycin intein splicing reporter for the purpose of evaluating splicing inhibitors, it is crucial that protein splicing be required to provide resistance (Figure 4A). Somewhat surprisingly, when splicing active and inactive versions of the *M. smegmatis* DnaB1 intein were inserted at 16 positions throughout the KanR protein, five displayed splicing-dependent resistance, eight displayed splicing-independent resistance, and three provided no resistance (Woods et al., 2020). Therefore, to ensure that kanamycin resistance cannot be provided by a KanR protein that is still interrupted by an intein, which would result in false positives when selecting for intein activity or potentially false negatives when searching for splicing inhibitors, a splicing inactive intein must also be tested at the same position within the KanR gene (Kelley et al., 2018; Woods et al., 2020; Woods et al., 2021). As expected, proximity to the active site correlated with the requirement of splicing for resistance, with all *M. smegmatis* DnaB1-KanR splicing-dependent resistance fusions were within 14 residues and 16 Å of the active site (Woods et al., 2020). In addition to ensuring that the inhibition of splicing would lead to kanamycin sensitivity by blocking splicing, the potential splicing inhibitor must be tested for cytotoxicity independent of splicing and lack of kanamycin resistance (Kelley et al., 2018; Woods et al., 2020) (Figure 4B). Ideally, the intein-KanR fusion should display reduced survival relative to an uninterrupted KanR in the presence of both a potential inhibitor and kanamycin (Figure 4B).

A significant advantage of the kanamycin intein splicing reporter is that it has been shown that it can be used in the native intein host, in this case *M. smegmatis*, so long as the host is sensitive to kanamycin. This system has been used in *M. smegmatis* to detect inhibition of the DnaB1 intein by





**FIGURE 4 |** Kanamycin Intein Splicing Reporter. For a splicing-dependent reporter, KanR interrupted with a catalytically active intein is resistant to kanamycin (**A**), while KanR interrupted with a catalytically inactive intein that cannot undergo splicing is not resistant to kanamycin (**B**). In (**C**), we show the experimental design to test the specific inhibition of protein splicing using a kanamycin-intein splicing reporter. In addition to cells containing the splicing-dependent KanR-intein fusion plasmid, control plasmids include KanR without an intein, a splicing inactive KanR-intein fusion, and kanamycin cells lacking a KanR-containing plasmid. In this design, serial dilutions of cells should be grown under four conditions: 1) without kanamycin and without splicing inhibitor; 2) without kanamycin and with splicing inhibitor; 3) with kanamycin and without splicing inhibitor; and 4) with kanamycin and with splicing inhibitor. Growth of the kanamycin sensitive and splicing inactive KanR-intein fusion strains should only occur in the absence kanamycin. Presence of the splicing inhibitor in the absence of kanamycin should not decrease growth of any strain. If inhibition of protein splicing occurs, the splicing-dependent KanR-intein fusion will display reduced survival compared to KanR without an intein. Interrupted, inactive exteins are colored gray, inteins are colored pink if active and gray if inactivated by mutation, and active KanR is colored blue.



hydrogen peroxide and zinc (Kelley et al., 2018; Woods et al., 2020; Woods et al., 2021), both stressors mycobacteria can encounter during challenge from the innate immune system (Zahrt and Deretic, 2002; Wagner et al., 2005; Neyrolles et al., 2015; Weiss and Schaible, 2015; Gao et al., 2018). Although *M. smegmatis* is generally non-pathogenic, and results may not translate exactly to mycobacteria that harbor conserved inteins like *M. tuberculosis* and *M. leprae*, it is useful as proxy for these pathogens. Importantly, this general approach could be adapted

to any intein-containing pathogen, as long as the pathogen is sensitive to kanamycin and appropriate expression vectors are available. On the other hand, one disadvantage to this screening system is the possibility of missing a lead compound that could serve as an inhibitor because it cannot penetrate the cells. Additionally, adaptation of the system to a high-throughput liquid culture format, rather than agar plates, could significantly increase the speed at which potential inhibitors can be tested.

It should be noted that the KanR-intein fusion design has been used in applications other than to search for inhibitors. One of the first uses of a kanamycin screening system was to identify open reading frames of the pathogen *Haemophilus influenzae* (Daugelat and Jacobs, 1999). In another application, a kanamycin selection system was used to make splicing conditional on addition of a small molecule, 4-hydroxytamoxifen (Buskirk et al., 2004). Finally, a KanR system has been used to improve splicing of the naturally split *Nostoc punctiforme* DnaE intein through directed evolution (Lockless and Muir, 2009).

## Green Fluorescence Protein and Luciferase

It is possible to screen for intein activity by making the fluorescence of green fluorescent protein (GFP) conditional on protein splicing either in *cis* and in *trans*. GFP fluorescence is useful for small molecule inhibitor discovery as it is amenable to high throughput screening and does not require potential lead compounds to penetrate a cell. The full-length *M. tuberculosis* RecA intein was used to interrupt GFP after residue 129, based on the observation that peptide insertions adjacent to this residue result in GFP expression into inclusion bodies in *E. coli* (Figure 5) (Gangopadhyay et al., 2003). An N-terminal His tag allows for purification of intein-interrupted GFP via immobilized metal affinity chromatography, followed by denaturation in 8M urea and renaturation into a folding buffer. The addition of zinc ions during refolding prevents intein activity as measured by GFP fluorescence, and this inhibition can be readily reversed by the addition of EDTA (Gangopadhyay et al., 2003). *In vitro* screening may be facilitated by mixing potential inhibitors during the refolding process. GFP fluorescence was used in a directed evolution experiment to measure splicing efficiency of evolved versions of the *M. tuberculosis* RecA intein (Yuen et al., 2006). The *Synechocystis* sp. PCC6803 DnaE split intein was used to reconstitute GFP in an assay to detect protein-protein interactions; the intein was reported to facilitate some GFP fluorescence without other interacting partners driving the association and hence this could serve as a means to select inhibitors for DnaE split intein reassociation (Ozawa et al., 2001).

GFP fluorescence as a screening system was used to explore the effect of platinum-containing compounds on the splicing of a minimized *M. tuberculosis* RecA intein (Chan et al., 2016). These compounds, including cisplatin, Ptffbz, Zeise's salt, and phenanthriplatin, were added to a stabilized precursor GFP-RecA fusion protein in concentrations up to 80  $\mu$ M. Once splicing was initiated with EDTA, the resulting fluorescence was observed, and it was found that increased concentrations of the Pt-containing compounds led to decreased fluorescence, suggesting that the compounds could successfully inhibit RecA protein splicing (Chan et al., 2016). This supported previous work suggesting that cisplatin can be toxic to *M. tuberculosis*, with intein over-expression decreasing the toxicity, indicating direct binding of the intein to cisplatin *in vivo* (Zhang et al., 2011). GFP screening also was used to investigate the influence of mutations of non-catalytic sites on RecA intein activity (Hiraga et al., 2009).

Using a similar design by which interruption with an intein blocks reporter function, the luciferase from *Renilla reniformis*

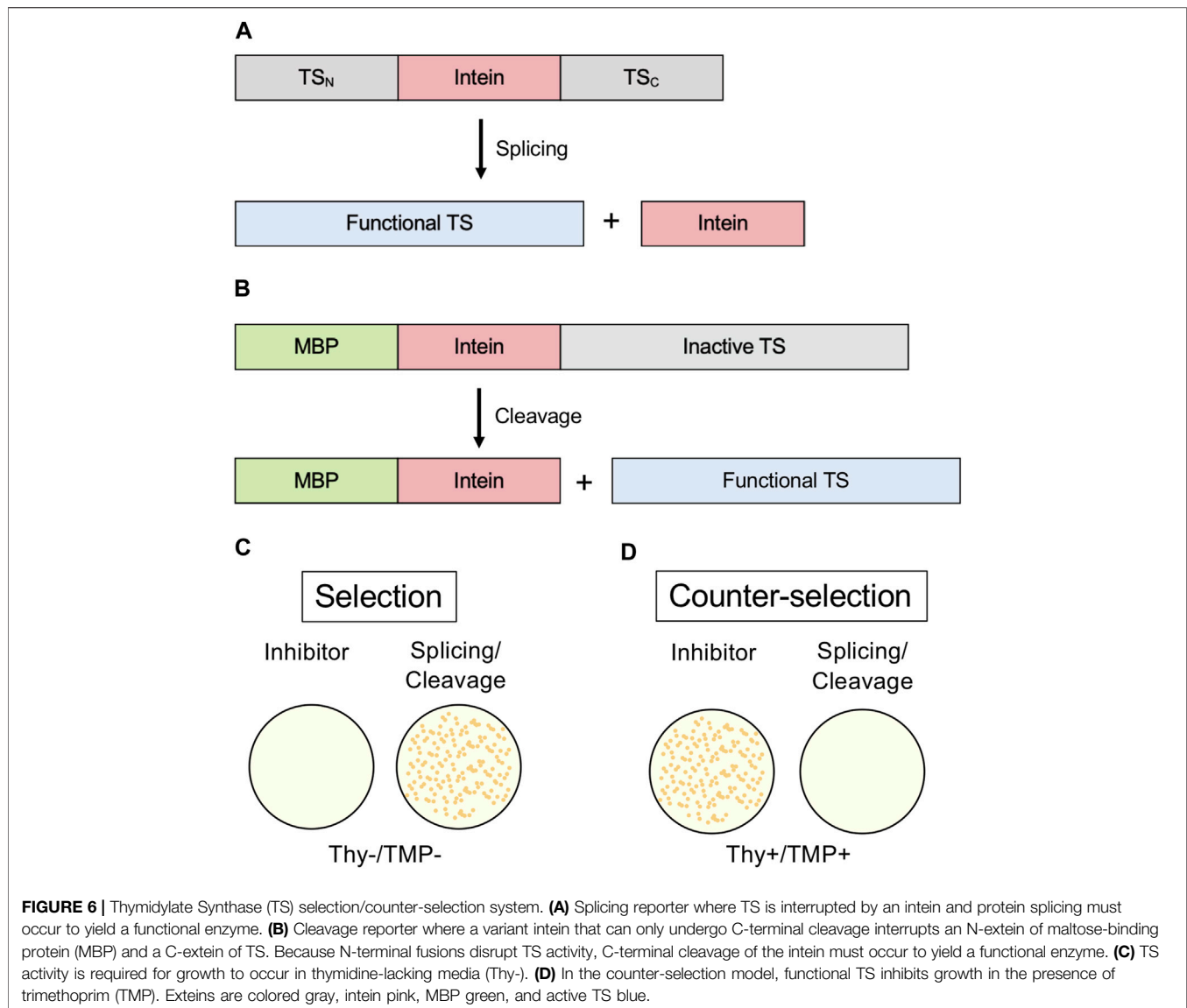
was split between residues 229 and 230 with the Prp8 intein from *C. gattii* (Li et al., 2019). The fusion protein was isolated at pH 9, where splicing was inefficient, and then incubated with or without inhibitor at pH 8, with a decrease in fluorescence measuring splicing inhibition, in this case with cisplatin (Li et al., 2019). Both this split luciferase system, as well as a split GFP system, was used to measure the splicing of the *C. neoformans* Prp8 intein (Li et al., 2021). GFP fluorescence was dependent on TCEP, showing that splicing and thus fluorescence was dependent on the intein active site Cys being present as a free thiol. The system also allowed for discovery of small molecule inhibitors of the intein, as discussed above (Li et al., 2021).

## CcdA/B Toxin-Antitoxin

Another in-cell screening system utilizes the *E. coli* toxin CcdB, which targets the topoisomerase GyrA (Bernard and Couturier, 1992). The full length *M. tuberculosis* RecA intein was used to interrupt CcdB on both a low and high copy number plasmid, such that cell growth is coupled to splicing inhibition. The intein-ccdB fusion was used to determine splicing efficiency as modulated by mutations to flanking extein residues (Lew and Paulus, 2002). Using plasmids with variable copy number and/or expression levels using a tunable promoter could vary the stringency used in different rounds of small molecule screening. A similar screening system utilizes the *E. coli* antitoxin CcdA, which acts as the CcdB antitoxin, forming an inactivated complex. In this method, an intein-less CcdB is expressed by an *E. coli* strain, and the intein is instead inserted into the CcdA protein in a region that directly interacts with CcdB, allowing CcdA to only inactivate CcdB if protein splicing occurs. This system worked with the *cis*-splicing *Nostoc punctiforme* DnaB intein and the naturally split *trans*-splicing cyanobacterial gp41-1 intein (Beyer and Iwai, 2019). The CcdA system can serve as an *in vivo* counter-selection method for the CcdB model, as effective intein inhibitors would cause cell death rather than permit cell growth.

## Thymidylate Synthase

An intein-thymidylate synthase (TS) fusion represents another *in vivo* system that may be used to monitor protein splicing. *E. coli* in the absence of thymidine in media cannot grow without functional TS; therefore, TS can serve as an indicator of splicing when it is interrupted by an intein (Wood et al., 1999). *E. coli* that lack their native TS can be transformed with a plasmid coding for TS with an intein inserted. If the transformed bacteria are grown in thymidine-lacking media, growth will not occur without protein splicing, as the unspliced intein-TS fusion protein is nonfunctional (Figure 6A). This system conveniently has a natural negative control, as growth in thymidine-rich media would not be splicing-dependent, and an inhibitor should show little effect on cell growth if its activity is intein-specific. Another useful aspect of this system is that N-terminal fusions inactivate TS (Figure 6B). Therefore, intein-mediated splicing and C-terminal cleavage can also be monitored by this system (Figure 6C) (Wood et al., 1999). This system was used to isolate an activating splicing mutant of a minimized version of the *M. tuberculosis* RecA intein



and to select a mutant that facilitates pH-dependent C-terminal cleavage (Wood et al., 1999).

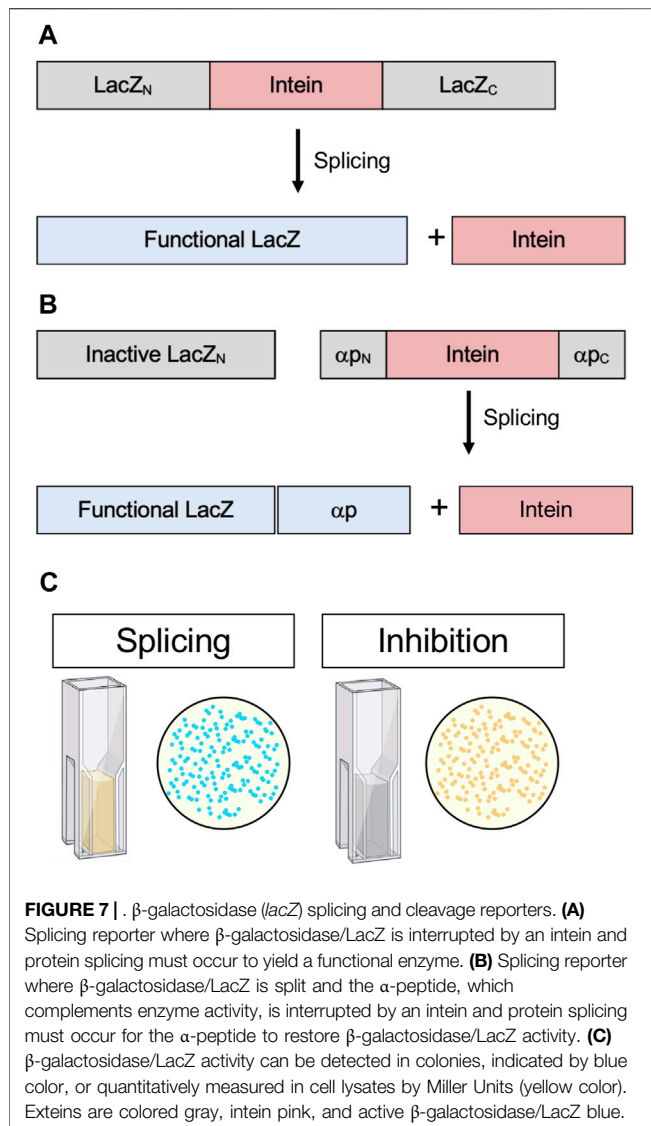
The TS system also has been used to design an intein-based system by which the miniaturized *M. tuberculosis* RecA intein was interrupted by a thyroid hormone receptor, making splicing and TS activity sensitive to the addition of thyroid hormone (Skretas and Wood, 2005). A TS-splicing system, in a modified format, also has been used to screen libraries of *de novo* protein sequences for frameshifts or stop codons. Insertion of gene sequences upstream of the intein-TS fusion that are out of frame prevent expression of the intein-TS fusion and hence growth on thymidine-free media (Bradley et al., 2005).

Screening for intein activity by conditional TS activity also has the advantage of a counter-selection system, using trimethoprim (TMP) (Figure 6D) (Wong et al., 2005). TMP inhibits dihydrofolate reductase and therefore prevents the production of tetrahydrofolate (THF). If TS is active, THF is depleted. In

thymidine-rich media, the cells don't require active TS for thymidine and can grow, but active TS is toxic in the presence of TMP. This model of counter selection allows for screening of inhibitors of splicing or C-terminal cleavage.

### **β-galactosidase**

β-galactosidase is a widely employed *E. coli* reporter. Functional β-galactosidase cleaves X-gal into 5-bromo-4-chloro-3-hydroxyindole, which spontaneously dimerizes and oxidizes to form the bright blue pigment 5,5'-dibromo-4,4'-dichloro-indigo. Active β-galactosidase in *E. coli* results in blue colonies, whereas inactive β-galactosidase results in white colonies. In a commonly used technique β-galactosidase is split, and insertions into the α-peptide, which complements enzyme activity, can be used for blue-white screening (Green and Sambrook, 2019). If colonies are white, an insertion has been made in the α-peptide and β-galactosidase is nonfunctional. Additionally, β-galactosidase



activity can be assayed quantitatively from cell lysates by measuring Miller Units, another colorimetric assay that measures the conversion of *o*-nitrophenyl-β-D-galactoside to galactose and *o*-nitrophenol, the latter which can be measured spectroscopically (Smale, 2010). Modified blue-white screening has been used by interrupting *E. coli* β-galactosidase with an intein. For example, this was used to analyze the splicing activity of an estrogen-sensitive version of the *S. cerevisiae* vacuolar ATPase (VMA) intein in the presence of estrogen or synthetic analogs (Figure 7A). Rather than inhibition, this system tested for activation of the intein (Liang et al., 2011).

Another version of blue-white screening was used to identify active intein mutants and characterize the conditions favorable to splicing. Modified and artificially split versions of the *Synechocystis* sp. PCC6803 DnaB intein were inserted into two different sites in the β-galactosidase α-peptide whereby complementation of enzyme activity was dependent on splicing (Figure 7B). Subsequently, this system were used to

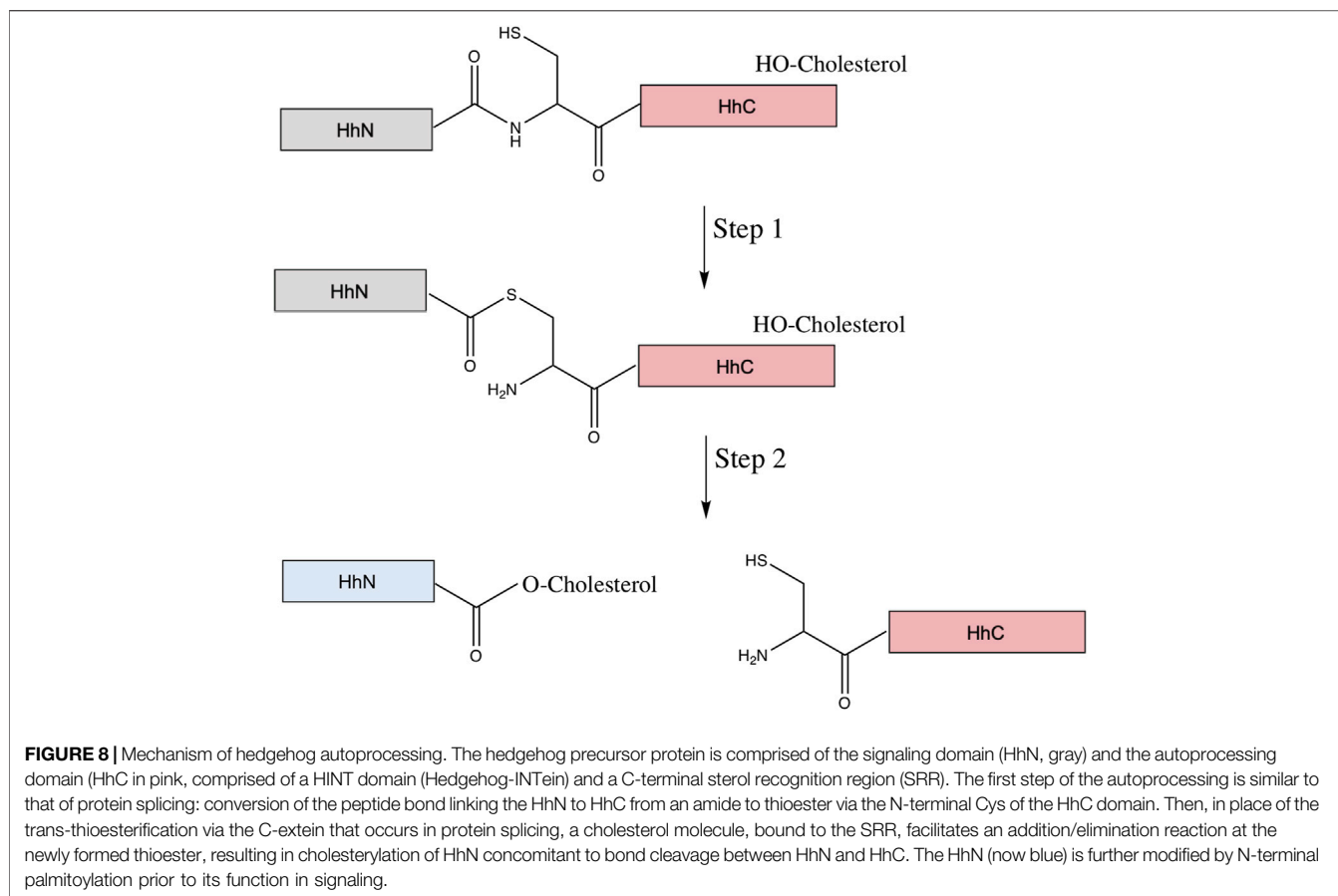
test the influence of active site mutations (Neugebauer et al., 2017). This time-tested assay could serve as a powerful way to screen for *in vivo* intein inhibition in a scalable model that does not rely on cytotoxicity as an output, as well as provide the option for quantitative measurements via Miller Units (Figure 7C).

## HEDGEHOG AUTOPROCESSING AND OFF-TARGET CONSIDERATIONS

The Hedgehog (Hh) signaling pathway is required for cell differentiation and proliferation during animal development. In adults, Hh signaling is mostly turned off but is involved in adult tissue regeneration and stem cell maintenance. Abnormal Hh activation is implicated in many types of cancer (Sigafoos et al., 2021). Because the structure and mechanism of inteins and the Hh family of proteins share so many features (Hall et al., 1997), it is important to consider the potential effect of intein inhibitors on Hh signaling, including Hh autoprocessing.

Hh ligands are produced from a unique self-catalyzed process, called Hh autoprocessing. The Hh ligand is not a direct product of translation; rather, it comes from an Hh precursor protein, which is subject to two post-translational modifications: cholesterylation at the C-terminus of the Hh ligand during Hh autoprocessing, and subsequently palmitoylation at the N-terminus (Callahan and Wang, 2015). Due to increased hydrophobicity, these post-translational modifications help establish the Hh gradients necessary for Hh signaling in development, and abnormalities in the post-translational modifications can lead to severe congenital diseases, including holoprosencephaly (HPE) (Roessler et al., 1997; Incardona and Roelink, 2000).

The Hh precursor consists of two domains: the N-terminal domain, HhN, and the C-terminal domain, HhC. HhN is the signaling domain, or Hh ligand. The HhC is the autoprocessing domain, consisting of the HINT domain and the sterol-recognition region (SRR). The HINT region shares many structural and functional similarities to the protein splicing of inteins (Xie et al., 2016). Hh autoprocessing consists of two steps (Figure 8). The first step is an N-S acyl shift, an amide to thioester rearrangement of the peptide bond linking HhN and HhC, identical to the first step of protein splicing. The second step is a transesterification using cholesterol as the nucleophile, resulting in cleavage of the thioester bond linking HhN and HhC and the cholesterylation of the C-terminus of HhN (Callahan and Wang, 2015; Xie et al., 2016). The precise means by which the SRR binds and activates cholesterol is unclear, but it is possible that the SRR may tether the entire Hh to a cholesterol-rich membrane, facilitating the recruitment of the cholesterol directly to the HhN active site, that it extracts cholesterol from the membrane and moves it to the Hh active site, or that it provides a hydrophobic conduit (Mafi et al., 2021). Next, the cholesterylated HhN is further lipidized by the hedgehog acyl transferase, which covalently links palmitic acid to the HhN N-terminus (Pepinsky et al., 1998). Once both lipids are bonded to HhN, the newly formed Hh ligand can be properly transported and used for downstream signaling. The cleaved



HhC, on the other hand, is likely degraded by the proteasome (Briscoe and Théron, 2013).

A mature N-terminally palmitoylated and C-terminally cholesterylated Hh ligand is then secreted into the extracellular space with the help of Dispatched-1 and Scube2. Once in the extracellular space, Scube2 proteins act as chaperones and bind to the Hh ligand during its transfer from producing to receiving cells. The ligand binds to Patched-1 on the receiving cell surface, which then releases its inhibition of smoothened (SMO), activating glioma-associated oncogene homolog (GLI) transcription factors to trigger subsequent signaling events such as development and patterning (Qi and Li, 2020).

As both inteins and Hh family proteins have HINT domains, they share conserved residues. The first is the catalytic Cys1 residue, the nucleophile in the amide-thioester rearrangement. Two other conserved residues, a Thr and His in the TXXH motif of “block B” of the HINT domain, catalyze the thioester formation during autoprocessing. Two additional conserved HhC residues, an Asp and a downstream Cys (D46 in C143 *Drosophila* HINT), do not have direct homologous residues in intein sequences. The downstream Cys is capable of forming an internal disulfide bond with Cys1, suggesting a role in proper folding (Chen et al., 2011). The conserved Asp likely coordinates the binding of cholesterol to the SRR with the first step of autoprocessing, and activates the hydroxyl group of cholesterol to attack the thioester (Hall et al., 1997; Xie et al., 2016).

Surprisingly, substitution of this Asp with His, a nominal charge reversal substitution (D46H in *Drosophila* HINT), can allow for a wider array of molecules, including coprostanol and epicoprostanol, to serve as the sterol substrate (Zhao et al., 2019). This is likely due to the more favorable, less spatially restrictive electrostatic stabilization of transition state provided by the H46 variant.

Hh autoprocessing can be inhibited by zinc and, similarly to protein splicing, zinc ions can bind to residues vital for autoprocessing to inhibit the function. For *Drosophila* Hh, micromolar zinc concentrations can inhibit autoprocessing, potentially by coordinating Cys1 and the essential His and Asp residues. Prostate, lung, and ovarian cancers are partially characterized by high levels of Hh ligand and low zinc levels, and zinc inhibition of autoprocessing might suggest a link between these observations (Xie et al., 2016). Given that inteins and Hh proteins share zinc-mediated inhibition, some intein-inhibitors may also inhibit Hh autoprocessing, and selective intein inhibitors may be useful. Hh autoprocessing also can be inhibited by selective alkylation of the downstream Cys residue. As measured by a FRET-based cholesterolysis assay, novel compound ST044643 covalently modifies the Cys to inhibit cholesterolysis (Owen et al., 2015). Compounds that activate hydrolysis of the scissile peptide bond between the signaling domain and the autoprocessing domain have recently been described (Smith et al., 2020). These so-called

HhC activator compounds redirect the cholesterylation reaction toward hydrolysis, even in the presence of sterol, in a non-covalent mechanism termed paracatalysis. This is exciting, as it suggests that similar paracatalytic compounds may be discovered for inteins, diverting splicing to side reactions that prevent ligation of the flanking exons. However, this also cautions us that such compounds might be cross-reactive with Hh autoprocessing domains.

Because of the shared HINT domain between intein and Hh autoprocessing, it is important to test if an intein inhibitor is also an inhibitor for Hh autoprocessing. Then, intein inhibitors can be classified into two categories: *selective intein inhibitors* that do not affect Hh autoprocessing and *nonselective inhibitors*, such as Zn ion, that inhibit both intein splicing and Hh autoprocessing. Both types of inhibitors may be useful in clinical settings. For example, if a patient has both Hh ligand-dependent cancer and tuberculosis infection, a nonselective intein inhibitor may be effective against both the cancer and infection, respectively. By inhibiting Hh autoprocessing, the drug will reduce Hh ligand, curbing cancer growth. By inhibiting intein splicing, key DNA enzymes in *M. tuberculosis*, such as the intein-interrupted RecA, cannot be generated, curbing the tuberculosis infection. Indeed, many cancer patients are immunocompromised and more susceptible to TB infection, thus such a clinical scenario can indeed arise. Selective inhibitors will be useful for cases where active Hh signaling needs to be preserved during the application of intein inhibitor, for example in treating TB during pregnancy. To distinguish between selective intein inhibitors and those that also target Hh autoprocessing, screens could be designed in tandem. For example, by combining one of the intein activity reporters described above with a FRET-based cholesterolysis assay for Hh activity (Owen et al., 2015), selective intein inhibitors could be identified.

## CONCLUSION

Numerous pathogenic bacteria and fungi, including *M. tuberculosis* and *C. neoformans*, harbor inteins within essential genes. Here, we argue that blocking protein splicing, and thus compromising the function of these

proteins, provides a potential avenue to kill these devastating pathogens. Decades of intein research has demonstrated that protein splicing can be modulated by numerous environmental factors (Belfort, 2017; Lennon and Belfort, 2017). Given the overall abundance of inteins in the microbial world, presence of inteins in essential genes of bacterial and fungal pathogens, and absence of inteins in metazoans, inteins represent attractive drug targets. While the FDA-approved chemotherapeutic cisplatin has been shown to block protein splicing of inteins from both *M. tuberculosis* and *C. neoformans* (Zhang et al., 2011; Chan et al., 2016; Li et al., 2019), it may not be an appropriate treatment option given off-target DNA damage. However, cisplatin and other splicing inhibitors (Li et al., 2021) demonstrate the proof of concept that splicing inhibitors can be used to kill pathogens and may serve as lead compounds to future drug development. As tuberculosis is typically the leading cause of death worldwide from infectious disease, and multidrug-resistance is becoming increasingly common, there is an urgent need for novel drug targets. The next stages of the development of intein inhibitors as drugs include the validation of inteins as important drug targets by clear demonstration that protein splicing is essential to pathogen survival and/or growth, and attention to developing more complex inhibitors that are specific to inteins rather than closely related hedgehog proteins or somewhat related enzymes such as proteases with reactive active site nucleophiles.

## AUTHOR CONTRIBUTIONS

DW, ST, and KM wrote the original draft of the article. CW, KM, and CL edited and revised the manuscript. CL and KM constructed the figures.

## FUNDING

We acknowledge support from the National Institutes of Health (R01 CA206592 (CW), 1R15GM132817-01 (KM), KY INBRE P20GM103436, and R15GM143662 (CL)) and the Camille and Henry Dreyfus Foundation (KM).

## REFERENCES

- Adam, E., and Perler, F. B. (2002). Development of a Positive Genetic Selection System for Inhibition of Protein Splicing Using Mycobacterial Inteins in *Escherichia coli* DNA Gyrase Subunit A. *J. Mol. Microbiol. Biotechnol.* 4 (5), 479–487.
- Aranko, A. S., Wlodawer, A., and Iwai, H. (2014). Nature's Recipe for Splitting Inteins. *Protein Eng. Des. Selection* 27 (8), 263–271. doi:10.1093/protein/gzu028
- Belfort, M. (2017). Mobile Self-Splicing Introns and Inteins as Environmental Sensors. *Curr. Opin. Microbiol.* 38, 51–58. doi:10.1016/j.mib.2017.04.003
- Bernard, P., and Couturier, M. (1992). Cell Killing by the F Plasmid CcdB Protein Involves Poisoning of DNA-Topoisomerase II Complexes. *J. Mol. Biol.* 226 (3), 735–745. doi:10.1016/0022-2836(92)90629-x
- Beyer, H. M., and Iwai, H. (2019). Off-Pathway-Sensitive Protein-Splicing Screening Based on a Toxin/Antitoxin System. *Chembiochem* 20 (15), 1933–1938. doi:10.1002/cbic.201900139
- Bongomin, F., Gago, S., Oladele, R., and Denning, D. (2017). Global and Multi-National Prevalence of Fungal Diseases-Estimate Precision. *JoF* 3 (4), 57. doi:10.3390/jof3040057
- Bradley, L. H., Kleiner, R. E., Wang, A. F., Hecht, M. H., and Wood, D. W. (2005). An Intein-Based Genetic Selection Allows the Construction of a High-Quality Library of Binary Patterned De Novo Protein Sequences. *Protein Eng. Des. Selection : PEDS* 18 (4), 201–207. doi:10.1093/protein/gzi020
- Briscoe, J., and Théron, P. P. (2013). The Mechanisms of Hedgehog Signalling and its Roles in Development and Disease. *Nat. Rev. Mol. Cel Biol* 14 (7), 416–429. doi:10.1038/nrm3598

- Buskirk, A. R., Ong, Y.-C., Gartner, Z. J., and Liu, D. R. (2004). Directed Evolution of Ligand Dependence: Small-Molecule-Activated Protein Splicing. *Proc. Natl. Acad. Sci.* 101 (29), 10505–10510. doi:10.1073/pnas.0402762101
- Butler, M. I., Gray, J., Goodwin, T. J., and Poulter, R. T. (2006). The Distribution and Evolutionary History of the PRP8 Intein. *BMC Evol. Biol.* 6, 42. doi:10.1186/1471-2148-6-42
- Callahan, B., and Wang, C. (2015). Hedgehog Cholesterololysis: Specialized Gatekeeper to Oncogenic Signaling. *Cancers* 7 (4), 2037–2053. doi:10.3390/cancers7040875
- Chan, H., Pearson, C. S., Green, C. M., Li, Z., Zhang, J., Belfort, G., et al. (2016). Exploring Intein Inhibition by Platinum Compounds as an Antimicrobial Strategy. *J. Biol. Chem.* 291 (43), 22661–22670. doi:10.1074/jbc.M116.747824
- Chen, X., Tukachinsky, H., Huang, C.-H., Jao, C., Chu, Y.-R., Tang, H.-Y., et al. (2011). Processing and Turnover of the Hedgehog Protein in the Endoplasmic Reticulum. *J. Cel. Biol.* 192 (5), 825–838. doi:10.1083/jcb.201008090
- Centers for Disease Control and Prevention (2020). *C. neoformans* Infection Statistics. Available at: <https://www.cdc.gov/fungal/diseases/cryptococcosis-neoformans/statistics.html> (Accessed October 14, 2020).
- Dasari, S. R., Velma, V., Yedjou, C. G., and Tchounwou, P. B. (2015). Preclinical Assessment of Low Doses of Cisplatin in the Management of Acute Promyelocytic Leukemia. *Int. J. Cancer Res. Mol. Mech.* 1 (3), 1–6. doi:10.16966/2381-3318.113
- Daugelat, S., and Jacobs, W. R., Jr. (1999). The *Mycobacterium tuberculosis* recA Intein Can Be Used in an ORFTRAP to Select for Open reading Frames. *Protein Sci.* 8 (3), 644–653. doi:10.1110/ps.8.3.644
- 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 (2), 201–210. doi:10.1016/0092-8674(92)90349-h
- Davis, E. O., Sedgwick, S. G., and Colston, M. J. (1991). Novel Structure of the recA Locus of *Mycobacterium tuberculosis* Implies Processing of the Gene Product. *J. Bacteriol.* 173 (18), 5653–5662. doi:10.1128/jb.173.18.5653-5662.1991
- Debelouchina, G. T., and Muir, T. W. (2017). A Molecular Engineering Toolbox for the Structural Biologist. *Quart. Rev. Biophys.* 50, e7. doi:10.1017/S0033583517000051
- Du, Z., Liu, Y., Ban, D., Lopez, M. M., Belfort, M., and Wang, C. (2010). Backbone Dynamics and Global Effects of an Activating Mutation in Minimized Mtu RecA Inteins. *J. Mol. Biol.* 400 (4), 755–767. doi:10.1016/j.jmb.2010.05.044
- El-Kareh, A. W., and Secomb, T. W. (2003). A Mathematical Model for Cisplatin Cellular Pharmacodynamics. *Neoplasia* 5 (2), 161–169. doi:10.1016/s1476-5586(03)80008-8
- Frischkorn, K., Springer, B., Böttger, E. C., Davis, E. O., Colston, M. J., and Sander, P. (2000). *In Vivo* splicing and Functional Characterization of *Mycobacterium leprae* RecA. *J. Bacteriol.* 182 (12), 3590–3592. doi:10.1128/jb.182.12.3590-3592.2000
- Gangopadhyay, J. P., Jiang, S.-q., and Paulus, H. (2003). An *In Vitro* Screening System for Protein Splicing Inhibitors Based on green Fluorescent Protein as an Indicator. *Anal. Chem.* 75 (10), 2456–2462. doi:10.1021/ac020756b
- Gao, H., Dai, W., Zhao, L., Min, J., and Wang, F. (2018). The Role of Zinc and Zinc Homeostasis in Macrophage Function. *J. Immunol. Res.* 2018, 1–11. doi:10.1155/2018/6872621
- Gogarten, J. P., and Hilario, E. (2006). Inteins, Introns, and Homing Endonucleases: Recent Revelations about the Life Cycle of Parasitic Genetic Elements. *BMC Evol. Biol.* 6, 94. doi:10.1186/1471-2148-6-94
- Grainger, R. J., and Beggs, J. D. (2005). Prp8 Protein: at the Heart of the Spliceosome. *Rna* 11 (5), 533–557. doi:10.1261/rna.2220705
- Green, C. M., Li, Z., Smith, A. D., Novikova, O., Bacot-Davis, V. R., Gao, F., et al. (2019). Spliceosomal Prp8 Intein at the Crossroads of Protein and RNA Splicing. *PLoS Biol.* 17 (10), e3000104. doi:10.1371/journal.pbio.3000104
- Green, M. R., and Sambrook, J. (2019). Screening Bacterial Colonies Using X-Gal and IPTG:  $\alpha$ -Complementation. *Cold Spring Harb Protoc.* 2019 (12), pdb.prot101329. doi:10.1101/pdb.prot101329
- Gupta, S., Takebe, N., and Lorusso, P. (2010). Review: Targeting the Hedgehog Pathway in Cancer. *Ther. Adv. Med. Oncol.* 2 (4), 237–250. doi:10.1177/1758834010366430
- Hall, T. M. T., 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 (1), 85–97. doi:10.1016/s0092-8674(01)80011-8
- 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 (4), 916–926. doi:10.1016/j.jmb.2005.09.088
- Hiraga, K., Soga, I., Dansereau, J. T., Pereira, B., Derbyshire, V., Du, Z., et al. (2009). Selection and Structure of Hyperactive Inteins: Peripheral Changes Relayed to the Catalytic center. *J. Mol. Biol.* 393 (5), 1106–1117. doi:10.1016/j.jmb.2009.08.074
- Huet, G., Castaing, J.-P., Fournier, D., Daffe', M., and Saves, I. (2006). Protein Splicing of SufB Is Crucial for the Functionality of the *Mycobacterium tuberculosis* SUF Machinery. *J. Bacteriol.* 188 (9), 3412–3414. doi:10.1128/JB.188.9.3412-3414.2006
- Incardona, J. P., and Roelink, H. (2000). The Role of Cholesterol in Shh Signaling and Teratogen-Induced Holoprosencephaly. *Cmls, Cel. Mol. Life Sci.* 57 (12), 1709–1719. doi:10.1007/pl00000653
- Kelley, D. S., Lennon, C. W., Belfort, M., SEA-PHAGES and Novikova, O. (2016). Mycobacteriophages as Incubators for Intein Dissemination and Evolution. *MBio* 7 (5), e01537–16. doi:10.1128/mBio.01537-16
- Kelley, D. S., Lennon, C. W., Li, Z., Miller, M. R., Banavali, N. K., Li, H., et al. (2018). Mycobacterial DnaB Helicase Intein as Oxidative Stress Sensor. *Nat. Commun.* 9 (1), 4363. doi:10.1038/s41467-018-06554-x
- Kempker, R. R., Kipiani, M., Mirtskhulava, V., Tukvadze, N., Magee, M. J., and Blumberg, H. M. (2015). Acquired Drug Resistance in Mycobacterium Tuberculosis and Poor Outcomes Among Patients with Multidrug-Resistant Tuberculosis. *Emerg. Infect. Dis.* 21 (6), 992–1001. doi:10.3201/eid2106.141873
- Lennon, C. W., and Belfort, M. (2017). Inteins. *Curr. Biol.* 27 (6), R204–R206. doi:10.1016/j.cub.2017.01.016
- Lennon, C. W., Wahl, D., Goetz, J. R., and Weinberger, J., II (2021). Reactive Chlorine Species Reversibly Inhibit DnaB Protein Splicing in Mycobacteria. *Microbiol. Spectr.* 9 (2), e00301–e00321. doi:10.1128/Spectrum.00301-21
- Lew, B. M., and Paulus, H. (2002). An *In Vivo* Screening System against Protein Splicing Useful for the Isolation of Non-splicing Mutants or Inhibitors of the RecA Intein of *Mycobacterium tuberculosis*. *Gene* 282 (1–2), 169–177. doi:10.1016/s0378-1119(01)00836-8
- Li, Y. (2015). Split-inteins and Their Bioapplications. *Biotechnol. Lett.* 37 (11), 2121–2137. doi:10.1007/s10529-015-1905-2
- Li, Z., Fu, B., Green, C. M., Liu, B., Zhang, J., Lang, Y., et al. (2019). Cisplatin Protects Mice from challenge of Cryptococcus Neoformans by Targeting the Prp8 Intein. *Emerg. Microbes Infect.* 8 (1), 895–908. doi:10.1080/22221751.2019.1625727
- Li, Z., Tharappel, A. M., Xu, J., Lang, Y., Green, C. M., Zhang, J., et al. (2021). Small-Molecule Inhibitors for the Prp8 Intein as Antifungal Agents. *Proc. Natl. Acad. Sci. USA* 118 (2), e2008815118. doi:10.1073/pnas.2008815118
- Liang, R., Zhou, J., and Liu, J. (2011). Construction of a Bacterial Assay for Estrogen Detection Based on an Estrogen-Sensitive Intein. *Appl. Environ. Microbiol.* 77 (7), 2488–2495. doi:10.1128/AEM.02336-10
- Liu, X.-Q., and Yang, J. (2004). Prp8 Intein in Fungal Pathogens: Target for Potential Antifungal Drugs. *FEBS Lett.* 572 (1–3), 46–50. doi:10.1016/j.febslet.2004.07.016
- Lockless, S. W., and Muir, T. W. (2009). Traceless Protein Splicing Utilizing Evolved Split Inteins. *Proc. Natl. Acad. Sci.* 106 (27), 10999–11004. doi:10.1073/pnas.0902964106
- Mafi, A., Purohit, R., Vielmas, E., Lauinger, A. R., Lam, B., Cheng, Y.-S., et al. (2021). Hedgehog Proteins Create a Dynamic Cholesterol Interface. *PLoS One* 16 (2), e0246814. doi:10.1371/journal.pone.0246814
- Mills, K. V., Johnson, M. A., and Perler, F. B. (2014). Protein Splicing: How Inteins Escape from Precursor Proteins. *J. Biol. Chem.* 289 (21), 14498–14505. doi:10.1074/jbc.R113.540310
- Mills, K. V., Lew, B. M., Jiang, S.-q., 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.* 95 (7), 3543–3548. doi:10.1073/pnas.95.7.3543
- Mills, K. V., and Paulus, H. (2001). Reversible Inhibition of Protein Splicing by Zinc Ion. *J. Biol. Chem.* 276 (14), 10832–10838. doi:10.1074/jbc.M011149200
- Nanda, A., Nasker, S. S., Mehra, A., Panda, S., and Nayak, S. (2020). Inteins in Science: Evolution to Application. *Microorganisms* 8 (12), 2004. doi:10.3390/microorganisms8122004



- Neugebauer, M., Böcker, J. K., Matern, J. C. J., Pietrovovski, S., and Mootz, H. D. (2017). Development of a Screening System for Inteins Active in Protein Splicing Based on Intein Insertion into the LacZα-Peptide. *Biol. Chem.* 398 (1), 57–67. doi:10.1515/hsz-2016-0229
- Neyrolles, O., Wolschendorf, F., Mitra, A., and Niederweis, M. (2015). Mycobacteria, Metals, and the Macrophage. *Immunol. Rev.* 264 (1), 249–263. doi:10.1111/imir.12265
- Novikova, O., Jayachandran, P., Kelley, D. S., Morton, Z., Merwin, S., Topilina, N. I., et al. (2016). Intein Clustering Suggests Functional Importance in Different Domains of Life. *Mol. Biol. Evol.* 33 (3), 783–799. doi:10.1093/molbev/msv271
- Novikova, O., Topilina, N., and Belfort, M. (2014). Enigmatic Distribution, Evolution, and Function of Inteins. *J. Biol. Chem.* 289 (21), 14490–14497. doi:10.1074/jbc.R114.548255
- Owen, T. S., Ngoje, G., Lageman, T. J., Bordeau, B. M., Belfort, M., and Callahan, B. P. (2015). Förster Resonance Energy Transfer-Based Cholesterolysis Assay Identifies a Novel Hedgehog Inhibitor. *Anal. Biochem.* 488, 1–5. doi:10.1016/j.ab.2015.06.021
- Ozawa, T., Takeuchi, M., Kaihara, A., Sato, M., and Umezawa, Y. (2001). Protein Splicing-Based Reconstitution of Split Green Fluorescent Protein for Monitoring Protein–Protein Interactions in Bacteria: Improved Sensitivity and Reduced Screening Time. *Anal. Chem.* 73 (24), 5866–5874. doi:10.1021/ac101717k
- Panda, S., Nanda, A., Nasker, S. S., Sen, D., Mehra, A., and Nayak, S. (2021). Metal Effect on Intein Splicing: A Review. *Biochimie* 185, 53–67. doi:10.1016/j.biochi.2021.03.006
- Paulus, H. (2001). Inteins as Enzyme. *Bioorg. Chem.* 29 (3), 119–129. doi:10.1006/bioo.2001.1203
- Paulus, H. (2003). Inteins as Targets for Potential Antimycobacterial Drugs. *Front. Biosci.* 8, s1157–1165. doi:10.2741/1195
- Pédélecq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and Characterization of a Superfolder green Fluorescent Protein. *Nat. Biotechnol.* 24 (1), 79–88. doi:10.1038/nbt1172
- Pepinsky, R. B., Zeng, C., Wen, D., Rayhorn, P., Baker, D. P., Williams, K. P., et al. (1998). Identification of a Palmitic Acid-Modified Form of Human Sonic Hedgehog. *J. Biol. Chem.* 273 (22), 14037–14045. doi:10.1074/jbc.273.22.14037
- Poland, B. W., Xu, M.-Q., and Quijcho, F. A. (2000). Structural Insights into the Protein Splicing Mechanism of PI-SceI. *J. Biol. Chem.* 275 (22), 16408–16413. doi:10.1074/jbc.275.22.16408
- Porter, J. A., Ekker, S. C., Park, W.-J., von Kessler, D. P., Young, K. E., Chen, C.-H., et al. (1996a). Hedgehog Patterning Activity: Role of a Lipophilic Modification Mediated by the Carboxy-Terminal Autoprocessing Domain. *Cell* 86 (1), 21–34. doi:10.1016/s0092-8674(00)80074-4
- Porter, J. A., Young, K. E., and Beachy, P. A. (1996b). Cholesterol Modification of Hedgehog Signaling Proteins in Animal Development. *Science* 274 (5285), 255–259. doi:10.1126/science.274.5285.255
- Qi, X., and Li, X. (2020). Mechanistic Insights into the Generation and Transduction of Hedgehog Signaling. *Trends Biochem. Sci.* 45 (5), 397–410. doi:10.1016/j.tibs.2020.01.006
- Rimkus, T., Carpenter, R., Qasem, S., Chan, M., and Lo, H.-W. (2016). Targeting the Sonic Hedgehog Signaling Pathway: Review of Smoothed and GLI Inhibitors. *Cancers* 8 (2), 22. doi:10.3390/cancers8020022
- Rocha, A. d. S., Cunha, M. d. G., Diniz, L. M., Salgado, C., Aires, M. A. P., Nery, J. A., et al. (2012). Drug and Multidrug Resistance Among *Mycobacterium leprae* Isolates from Brazilian Relapsed Leprosy Patients. *J. Clin. Microbiol.* 50 (6), 1912–1917. doi:10.1128/jcm.06561-11
- Roessler, E., Belloni, E., Gaudenz, K., Vargas, F., Scherer, S. W., Tsui, L.-C., et al. (1997). Mutations in the C-Terminal Domain of Sonic Hedgehog Cause Holoprosencephaly. *Hum. Mol. Genet.* 6 (11), 1847–1853. doi:10.1093/hmg/6.11.1847
- Romero-Casañas, A., Gordo, V., Castro, J., and Ribó, M. (2020). Protein Splicing: From the Foundations to the Development of Biotechnological Applications. *Methods Mol. Biol.* 2133, 15–29. doi:10.1007/978-1-0716-0434-2\_2
- Sarmiento, C., and Camarero, J. A. (2019). Biotechnological Applications of Protein Splicing. *Cpps* 20 (5), 408–424. doi:10.2174/1389203720666190208110416
- Saves, I., Lanéelle, M. A., Daffé, M., and Masson, J. M. (2000). Inteins Invading Mycobacterial RecA Proteins. *FEBS Lett.* 480 (2–3), 221–225. doi:10.1016/s0014-5793(00)01944-x
- Shah, N. H., and Muir, T. W. (2014). Inteins: Nature’s Gift to Protein Chemists. *Chem. Sci.* 5 (1), 446–461. doi:10.1039/C3SC52951G
- Shah, N. H., and Stevens, A. J. (2020). Identification, Characterization, and Optimization of Split Inteins. *Methods Mol. Biol.* 2133, 31–54. doi:10.1007/978-1-0716-0434-2\_3
- Shamaei, M., and Mirsaeidi, M. (2021). Nontuberculous Mycobacteria, Macrophages, and Host Innate Immune Response. *Infect. Immun.* 89, IAI0081220. doi:10.1128/IAI.00812-20
- Shingledecker, K., Jiang, S.-q., and Paulus, H. (1998). Molecular Dissection of the *Mycobacterium tuberculosis* RecA Intein: Design of a Minimal Intein and of a Trans-splicing System Involving Two Intein Fragments. *Gene* 207 (2), 187–195. doi:10.1016/s0378-1119(97)00624-0
- 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 (1), 138–144. doi:10.1006/abbi.1999.1645
- Sigafoos, A. N., Paradise, B. D., and Fernandez-Zapico, M. E. (2021). Hedgehog/GLI Signaling Pathway: Transduction, Regulation, and Implications for Disease. *Cancers* 13 (14), 3410. doi:10.3390/cancers13143410
- Skretas, G., and Wood, D. W. (2005). Regulation of Protein Activity with Small-Molecule-Controlled Inteins. *Protein Sci.* 14 (2), 523–532. doi:10.1110/ps.04996905
- Smale, S. T. (2010). Beta-Galactosidase Assay. *Cold Spring Harb Protoc.* 2010 (5), pdb.prot5423. doi:10.1101/pdb.prot5423
- Smith, C. J., Wagner, A. G., Stagnitta, R. T., Xu, Z., Pezzullo, J. L., Giner, J.-L., et al. (2020). Subverting Hedgehog Protein Autoprocessing by Chemical Induction of Paracatalysis. *Biochemistry* 59 (6), 736–741. doi:10.1021/acs.biochem.0c00013
- Sojka, D., Šnebergerová, P., and Robbertse, L. (2021). Protease Inhibition-An Established Strategy to Combat Infectious Diseases. *Ijms* 22 (11), 5762. doi:10.3390/ijms22115762
- Stevens, A. J., Sekar, G., Shah, N. H., Mostafavi, A. Z., Cowburn, D., and Muir, T. W. (2017). A Promiscuous Split Intein with Expanded Protein Engineering Applications. *Proc. Natl. Acad. Sci. USA* 114 (32), 8538–8543. doi:10.1073/pnas.1701083114
- 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 (5), 1093–1105. doi:10.1016/j.jmb.2005.09.039
- Topilina, N. I., Green, C. M., Jayachandran, P., Kelley, D. S., Stanger, M. J., Piazza, C. L., et al. (2015a). SufB Intein of *Mycobacterium tuberculosis* as a Sensor for Oxidative and Nitrosative Stresses. *Proc. Natl. Acad. Sci. USA* 112 (33), 10348–10353. doi:10.1073/pnas.1512777112
- Topilina, N. I., and Mills, K. V. (2014). Recent Advances in *In Vivo* Applications of Intein-Mediated Protein Splicing. *Mobile DNA* 5 (1), 5. doi:10.1186/1759-8753-5-5
- Topilina, N. I., Novikova, O., Stanger, M., Banavali, N. K., and Belfort, M. (2015b). Post-translational Environmental Switch of RadA Activity by Extein-Intein Interactions in Protein Splicing. *Nucleic Acids Res.* 43 (13), 6631–6648. doi:10.1093/nar/gkv612
- 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 (1), 162–173. doi:10.1016/j.jmb.2006.12.050
- Volkman, G., and Iwai, H. (2010). Protein Trans-splicing and its Use in Structural Biology: Opportunities and Limitations. *Mol. Biosyst.* 6 (11), 2110–2121. doi:10.1039/c0mb00034e
- Volkman, G., and Mootz, H. D. (2013). Recent Progress in Intein Research: from Mechanism to Directed Evolution and Applications. *Cell. Mol. Life Sci.* 70 (7), 1185–1206. doi:10.1007/s00018-012-1120-4
- Wagner, D., Maser, J., Lai, B., Cai, Z., Barry, C. E., 3rd, Höner zu Bentrup, K., et al. (2005). Elemental Analysis of *Mycobacterium Avium*-, *Mycobacterium Tuberculosis*-, and *Mycobacterium Smegmatis*-Containing Phagosomes Indicates Pathogen-Induced Microenvironments within the Host Cell’s Endosomal System. *J. Immunol.* 174 (3), 1491–1500. doi:10.4049/jimmunol.174.3.1491

- Weiss, G., and Schaible, U. E. (2015). Macrophage Defense Mechanisms against Intracellular Bacteria. *Immunol. Rev.* 264 (1), 182–203. doi:10.1111/imr.12266
- Wong, Q. N. Y., Ng, V. C., Lin, M. C., Kung, H. F., Chan, D., and Huang, J. D. (2005). Efficient and Seamless DNA Recombineering Using a Thymidylate Synthase A Selection System in *Escherichia coli*. *Nucleic Acids Res.* 33 (6), e59. doi:10.1093/nar/gni059
- Wood, D. W., and Camarero, J. A. (2014). Intein Applications: from Protein Purification and Labeling to Metabolic Control Methods. *J. Biol. Chem.* 289 (21), 14512–14519. doi:10.1074/jbc.R114.552653
- 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 (9), 889–892. doi:10.1038/12879
- Woods, D., LeSassier, D. S., Egbunam, I., and Lennon, C. W. (2021). Construction and Quantitation of a Selectable Protein Splicing Sensor Using Gibson Assembly and Spot Titters. *Curr. Protoc.* 1 (3), e82. doi:10.1002/cpz1.82
- Woods, D., Vangaveti, S., Egbunam, I., Sweeney, A. M., Li, Z., Bacot-Davis, V., et al. (2020). Conditional DnaB Protein Splicing Is Reversibly Inhibited by Zinc in *Mycobacteria*. *mBio* 11 (4). doi:10.1128/mBio.01403-20
- World Health Organization (2020). *Global Tuberculosis Report 2020*. Geneva: World Health Organization).
- Xie, J., Owen, T., Xia, K., Callahan, B., and Wang, C. (2016). A Single Aspartate Coordinates Two Catalytic Steps in Hedgehog Autoprocessing. *J. Am. Chem. Soc.* 138 (34), 10806–10809. doi:10.1021/jacs.6b06928
- Yuen, C. M., Rodda, S. J., Vokes, S. A., McMahon, A. P., and Liu, D. R. (2006). Control of Transcription Factor Activity and Osteoblast Differentiation in Mammalian Cells Using an Evolved Small-Molecule-Dependent Intein. *J. Am. Chem. Soc.* 128 (27), 8939–8946. doi:10.1021/ja062980e
- Zahrt, T. C., and Deretic, V. (2002). Reactive Nitrogen and Oxygen Intermediates and Bacterial Defenses: Unusual Adaptations in *Mycobacterium Tuberculosis*. *Antioxid. Redox Signal.* 4 (1), 141–159. doi:10.1089/152308602753625924
- Zeidler, M. P., Tan, C., Bellaiche, Y., Cherry, S., Häder, S., Gayko, U., et al. (2004). Temperature-sensitive Control of Protein Activity by Conditionally Splicing Inteins. *Nat. Biotechnol.* 22 (7), 871–876. doi:10.1038/nbt979
- Zhang, L., Xiao, N., Pan, Y., Zheng, Y., Pan, Z., Luo, Z., et al. (2010). Binding and Inhibition of Copper Ions to RecA Inteins from *Mycobacterium Tuberculosis*. *Chem. Eur. J.* 16 (14), 4297–4306. doi:10.1002/chem.200903584
- Zhang, L., Zheng, Y., Callahan, B., Belfort, M., and Liu, Y. (2011). Cisplatin Inhibits Protein Splicing, Suggesting Inteins as Therapeutic Targets in *Mycobacteria*. *J. Biol. Chem.* 286 (2), 1277–1282. doi:10.1074/jbc.M110.171124
- Zhang, L., Zheng, Y., Xi, Z., Luo, Z., Xu, X., Wang, C., et al. (2009). Metal Ions Binding to recA Inteins from *Mycobacterium tuberculosis*. *Mol. Biosyst.* 5 (6), 644–650. doi:10.1039/b903144h
- Zhao, J., Ciulla, D. A., Xie, J., Wagner, A. G., Castillo, D. A., Zwarycz, A. S., et al. (2019). General Base Swap Preserves Activity and Expands Substrate Tolerance in Hedgehog Autoprocessing. *J. Am. Chem. Soc.* 141 (46), 18380–18384. doi:10.1021/jacs.9b08914

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Wall, Tarrant, Wang, Mills and Lennon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.