

Development of a Positive Genetic Selection System for Inhibition of Protein Splicing Using Mycobacterial Inteins in *Escherichia coli* DNA Gyrase Subunit A

Eric Adam¹, and Francine B. Perler*

New England BioLabs, 32 Tozer Road, Beverly, MA 01915, USA

¹Present address: SYRRX, Inc., 10450 Science Drive Suite 100, San Diego, CA 92121, USA

Abstract

An intein-based positive genetic selection system was developed to study protein splicing and to provide a selection system with the potential for finding splicing inhibitors. Inteins can be novel antimicrobial targets when present in essential proteins since blocking splicing would kill the organism. For example, pathogenic mycobacteria encode inteins that interrupt DNA gyrase. The gyrase selection system exploits (1) splicing of inteins out of Gyrase A and (2) the dominant lethal effect of quinolone poisoning of DNA gyrase, which in turn blocks replication. The system was adapted for whole-cell high-throughput screening using green fluorescent protein as an automatable readout of viability. To demonstrate the efficacy of this system, mutations that blocked splicing of the *Mycobacterium xenopi* Gyrase A intein were isolated. Splicing was then assayed at a second temperature to identify inteins with a temperature-sensitive splicing phenotype. Mutations were mapped onto a structure-based sequence alignment, which led to the rational prediction of a temperature-sensitive splicing mutation. GyrA intein subdomain relationships also provided insight into intein evolution.

Introduction

Inteins are in-frame insertions that excise themselves from protein precursors while the flanking regions (exteins) are ligated to generate functional host proteins (Noren *et al.*, 2000). The intein splicing domain is formed by sequences on both sides of an intein linker or homing endonuclease domain. It has a conserved structural fold, termed the HINT module, which is shared with Hedgehog autoprocessing proteins (Duan *et al.*, 1997; Hall *et al.*, 1997; Ichiyonagi *et al.*, 2000; Klabunde *et al.*, 1998; Poland *et al.*, 2000). Homing endonuclease activity can mobilize an intein gene into the identical insertion site in homologous host protein loci lacking the intein (Gimble, 2000; Jurica

et al., 1999). The 198 amino acid (aa) *Mycobacterium xenopi* (Mxe) DNA Gyrase subunit A (GyrA) intein has a compact structure composed mainly of β -strands with 2 antiparallel helices and a disordered section in its linker region (Klabunde *et al.*, 1998).

Since the majority of the >125 inteins registered in InBase (<http://www.neb.com/neb/inteins.html>) (Perler, 2000) are found in active sites and conserved motifs (Dalgaard *et al.*, 1997; Perler, 2000; Pietrovskii, 1998), blocking splicing should prevent activation of the host enzyme. Inteins are present in proteins involved in DNA replication and recombination in *Mycobacterium tuberculosis* (DnaB helicase and RecA), *Mycobacterium leprae* (DnaB helicase, RecA and GyrA), *Mycobacterium smegmatis* (DnaB helicase) and several life-threatening opportunistic mycobacterial pathogens associated with immunodeficiency syndromes and HIV (Davis *et al.*, 1994; Fsihi *et al.*, 1996; Lamden *et al.*, 1996; Perler, 2000; Sander *et al.*, 1998; Telenti *et al.*, 1997). DNA gyrase has been an effective antimicrobial drug target, but increasing resistance to classic inhibitors has created a need for new anti-gyrase agents (Jacobs, 1999). Many mycobacteria contain an intein in GyrA, which represents an unexploited drug target that could fill this void.

Previously described genetic systems only permit positive selection for protein splicing or single splice junction cleavage and cannot be used to directly select for loss of intein function, since cells that express inactive inteins do not grow under selective pressure (Daugelat *et al.*, 1999; Derbyshire *et al.*, 1997; Nogami *et al.*, 1997; Wood *et al.*, 1999). This study describes a positive genetic selection system that can be used for the discovery of *trans*-acting GyrA intein inhibitors or for inhibiting splicing of any intein that is functional in GyrA. It is based on the well-known properties of *Escherichia coli* (Eco) GyrA, including its sensitivity to quinolone drugs such as ofloxacin.

Results

A Positive Genetic Selection System for Inhibiting Protein Splicing of Inteins in DNA Gyrase Subunit A

DNA gyrase is a heterotetramer of GyrA and GyrB subunits that functions during DNA replication as a type II topoisomerase. It catalyzes the negative supercoiling of DNA by making double-strand breaks. During this process, GyrA is covalently bound to the cleaved DNA through its active site tyrosine. Most wild type microbial gyrases are quinolone-sensitive (quinolone^s) and many fully functional, quinolone-resistant (quinolone^r) Eco *gyrA* mutants have been identified, such as

*For Correspondence. Email perler@neb.com; Tel. (+1) 978 927 5054; Fax. (+1) 978 921 1350.

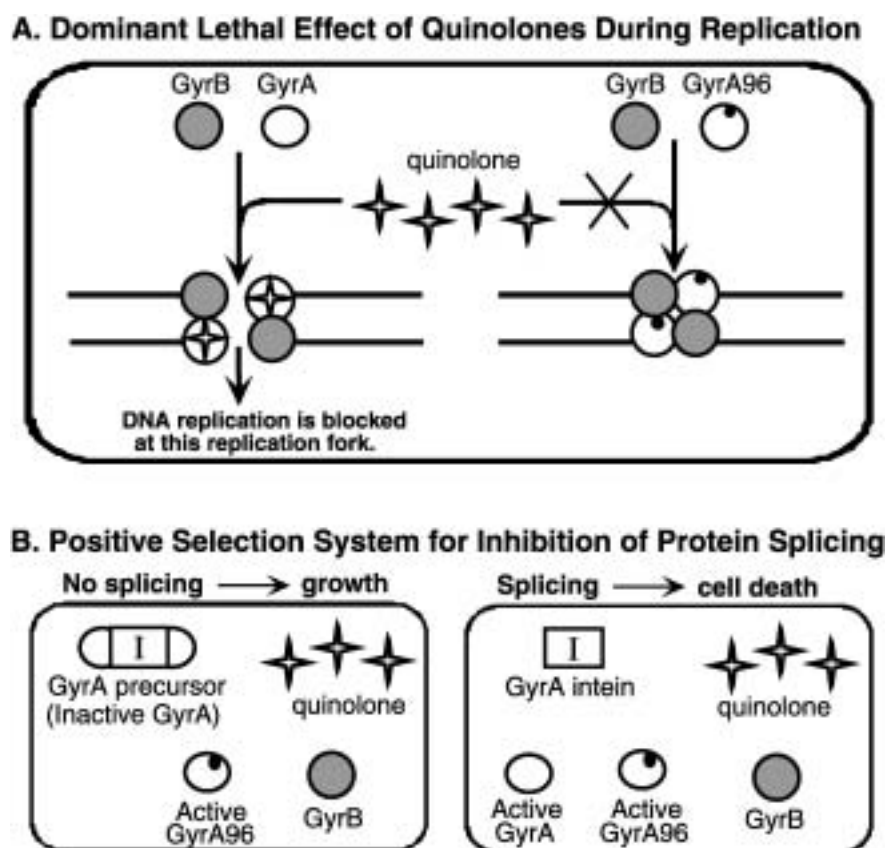


Figure 1. A positive genetic selection system for splicing of inteins in GyrA. (A) Quinolones, such as ofloxacin, bind to the wild type, quinolone^s GyrA subunit of Eco DNA gyrase. Quinolone^s GyrA confers dominant lethality, since the irreversible complex between ofloxacin and gyrase at the replication fork blocks replication, leading to cell death irrespective of the presence of quinolone^r GyrA96. (B) The GyrA positive genetic selection system consists of a merodiploid expressing both chromosomal quinolone^r GyrA96 and a plasmid-borne intein-containing (I), quinolone^s GyrA precursor. Cells die in the presence of ofloxacin if spliced quinolone^s GyrA is produced. If splicing is blocked, the quinolone^s GyrA precursor is inactive and the quinolone^r GyrA96 supports growth.

gyrA96 (Jacobs, 1999). Binding of quinolones to quinolone^s GyrA irreversibly freezes the gyrase-replication fork complex, blocking replication despite the presence of quinolone^r GyrA (Figure 1A).

A positive genetic selection system for inactive inteins was developed based on this dominant lethal phenotype: merodiploids expressing an intein-containing quinolone^s *gyrA* gene on a plasmid and an intein-free quinolone^r *gyrA96* allele on the chromosome die in the presence of quinolones unless splicing of the intein-containing GyrA precursor is blocked (Figure 1B). The chromosomal *gyrA96* gene provides active GyrA to permit growth prior to transformation with the GyrA plasmid and in the absence of splicing. To avoid any potential toxicity caused by interaction of mycobacterial and *E. coli* gyrase subunits, the Mxe GyrA intein was cloned into its homologous insertion site in Eco GyrA after Tyr122 (DSAAAMRY/TEAPLTPL). A mutated, inactive Mxe GyrA intein was also cloned into Eco GyrA to serve as a null-splicing control. All plasmid-borne Eco GyrA constructs have a His tag and their expression is isopropyl- β -D-thiogalactopyranoside (IPTG) inducible. The active Mxe GyrA intein spliced *in vivo* when present in Eco GyrA, while the inactive Mxe GyrA intein did not (Figure 2 and data not shown). Allelic

GyrA inteins from pathogenic mycobacteria (*Mycobacterium leprae*, *Mycobacterium kansasii*, *Mycobacterium malmoeense*, *Mycobacterium goodnae* and *Mycobacterium flavescens*) were cloned into the same insertion site in Eco *gyrA*. Eco GyrA precursors containing all GyrA intein alleles tested yielded >80% spliced product, except *M. leprae* (Figure 2).

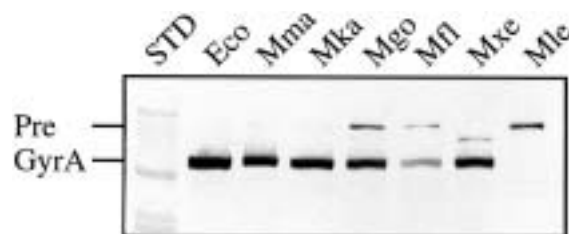


Figure 2. Splicing of GyrA inteins in Eco GyrA. Western blot analysis of mycobacterial GyrA inteins in Eco GyrA when expressed at 37°C for 2 h. Only recombinant Eco GyrA is detected with the anti-His tag sera. The Mxe GyrA mini-intein is ~222aa smaller than other GyrA inteins. GyrA intein in Eco GyrA: Eco, no intein; Mma, *M. malmoeense*; Mka, *M. kansasii*; Mgo, *M. goodnae*; Mfl, *M. flavescens*; Mle, *M. leprae*. Abbreviations: STD, size markers.

The minimum inhibitory concentration of the quinolone drug ofloxacin was determined by spotting 10^5 *E. coli* expressing plasmid-borne quinolone^s GyrA precursors on agar plates containing 0–0.5 $\mu\text{g}/\text{ml}$ ofloxacin (data not shown). The optimal ofloxacin concentration that inhibited growth of merodiploids expressing Mxe or *M. flavescens* GyrA intein fusions without inhibiting growth of cells expressing the respective inactive GyrA intein fusions was 0.3 $\mu\text{g}/\text{ml}$.

Testing the Efficacy of the GyrA Positive Selection System by Isolation of Thermosensitive-Splicing Mutants

Temperature-sensitive (*ts*) splicing variants were identified using positive genetic selection of inactive Mxe GyrA inteins mutagenized by error-prone polymerase chain reaction (PCR). Clones expressing mutated inteins that failed to splice at 37°C were able to grow on ofloxacin plates (Figure 1B). These colonies were then screened at 30°C on ofloxacin plates to identify clones that failed to grow because they produced spliced quinolone^s Eco GyrA at this lower temperature. Approximately 0.2% of a 26,000 member library displayed the *ts* splicing phenotype and 42 clones were further analyzed (Table 1). Little, if any, splicing was detected after expression at 37°C in the absence of quinolone selection, while varying amounts of spliced product were observed at 15°C (Figure 3). Small amounts of spliced product in some *ts* samples induced at 37°C in the absence of selective pressure correlated with longer sample preparation time. Single substitutions were present in 26 *ts* clones (Table 1 and Figure 4).

Table 1. Mutations present in *ts* Mxe GyrA inteins

Clone	Mutation	Clone	Mutation
1	E155G/I190T	24	V61M
2	D5G/K36R/V64A	25	R160Q
3	I97TK121E	26	T72A
4	E123G/Y129C/D154G/N192S	27	F189L
5	V45A	28	I21T
7	L140S	29	S53P
8	V81A	30	V86M
9	A172V	32	D186G
10	N31S	33	F189L
11	E96G/R119H	34	I97T
12	E65K	35	G42S/S108G/E141G
13	K92E	36	I97T/Q106R
14	K92E	37	L140S/H143R
15	T184M	38	I21V/L89P
16	N192H	39	I97T/K121R
17	G193R	40	I21T
19	R17S/V64A/I152T	41	G159R
20	R17S/S29G/K36R/V45G	42	L136P
21	L93P/E123G	43	L136P
		44	V169A
		45	Y162G/V174A
		46	D5G

Clones were selected for a *ts* splicing phenotype. The clone number is followed by the mutations present in each intein, using the single letter amino acid code.

Prediction of a Thermosensitive-Splicing Mutation Based on Intein Structure

The Mxe GyrA intein splicing domain consists of 2 subdomains that are related by a pseudo-twofold axis of symmetry (Hall *et al.*, 1997; Klabunde *et al.*, 1998). Three pairs of *ts* mutations were present in 3 structurally superimposable backbone positions in the Mxe GyrA intein subdomains: Ile21Thr and Ile97Thr; Ser53-Pro and Ala172Val; Glu65Lys and Thr184Met (Figure 4 and 5). The structural and sequence similarities between Mxe GyrA intein subdomains and the presence of these symmetric *ts* mutations suggested that it might be possible to use *ts* splicing mutations in one subdomain to predict mutations with similar phenotypes in the other subdomain. To test this hypothesis, an analogous mutation to Thr72Ala was made at Thr3, the equivalent position in the N-terminal subdomain. As predicted, the engineered Thr3Ala mutant had typical *ts* behavior indistinguishable from the genetically selected Thr72Ala mutant (Figure 3).

Development of a High-Throughput Selection System for Inhibition of Splicing in Eco GyrA

High-throughput whole-cell screens (HTS) of combinatorial chemical or natural product libraries would facilitate the search for potential protein splicing inhibitors. The GyrA positive selection system was modified for HTS by including green fluorescent protein (GFP) on a compatible plasmid to facilitate optical readout of cell growth in automated systems (Misteli *et al.*, 1997), although any convenient readout would be equally applicable. Several parameters were analyzed using the Mxe GyrA intein in a 96-well microtiter plate format (Figure 6A and data not shown). Parameters tested included ofloxacin concentration (0 to 0.4 $\mu\text{g}/\text{ml}$), cell density of microtiter plate inoculum (cultures were serially diluted 2-fold from 1:2 to 1:4096), induction conditions (0.4–2 mM IPTG either at or 2 h prior to inoculating microtiter plates) and growth temperature (30–37°C). To achieve a high differential fluorescence between active and inactive inteins in a reasonable time period (24 h), induction and culture densities had to be balanced because both GFP and GyrA precursors are induced with IPTG, and GFP is stable after lysis. Induction prior to dilution and dispensing into microtiter plates improved the bactericidal effect of ofloxacin without raising background fluorescence. Higher cell densities resulted in high background fluorescence, while lower densities required more time to reach a significant fluorescence value. At 37°C, all IPTG concentrations yielded similar results. At 30°C or below, cell growth was too slow to yield acceptable fluorescence values in 24 h.

Optimum HTS conditions consisted of: (a) diluting fresh overnight cultures 10-fold, (b) incubating at 37°C to mid log ($\text{OD}_{600} = 0.5$), (c) adding 0.4 mM IPTG and growing for 2 h at 37°C (to $\text{OD}_{600} = 2$), (d) dispensing into microtiter plates 200 μl of an ~ 4096 -fold dilution of the induced culture plus 0.2–0.3 $\mu\text{g}/\text{ml}$ ofloxacin and 0.4 mM IPTG, and (e) shaking the microtiter plates for 24 h at 37°C while periodically measuring growth by fluorescence.



Figure 3. Western blot analysis of *tsMxe GyrA* inteins in *Eco GyrA* (in the absence of selection) when expressed at 37°C for 2 hours or after the induced cells were shifted to 15°C overnight. All clones were genetically selected, except for Thr3Ala (T3A), which was rationally predicted based on its similar spatial location to the genetically selected Thr72Ala (T72A) mutation in the C-subdomain. See Table 1 for a list of the mutations present in each clone. Abbreviations: pre, precursor; GyrA, spliced *Eco GyrA*.

No growth of cells expressing the active Mxe GyrA intein fusion occurred at ofloxacin concentrations ≥ 0.2 $\mu\text{g/ml}$ (Figure 6A). At 0.2 $\mu\text{g/ml}$ ofloxacin, there was a 9-fold difference in fluorescence between clones expressing active vs. inactive Mxe GyrA inteins. High concentrations of ofloxacin are known to be toxic to all cells, including those cells expressing the inactive Mxe GyrA intein fusion (Figure 6A). A 7 to 10-fold decrease in fluorescence was observed after 24 h with cells expressing Mxe, *M. leprae*, *M. kansasii*, *M. malmoense*, *M. goodnae* and *M. flavescens GyrA* intein fusions in cultures containing 0.25–0.3 $\mu\text{g/ml}$ ofloxacin compared to cultures without ofloxacin (Figure 6B).

Discussion

An intein-based positive genetic selection system was developed to study protein splicing and inhibition of splicing. It provides a basis for future studies to identify *trans*-acting molecules that can control splicing. Fully active *gyrA* inteins were cloned into the homologous insertion site of quinolone^s *Eco gyrA*, linking splicing to a selectable growth phenotype: in the presence of ofloxacin, splicing led to cell death, while inhibiting splicing permitted cell growth. As an initial test, the system was used to select for mutations that inhibited splicing at 37°C. To identify *ts* splicing mutants, these same clones were then screened for inteins that were active at 30°C. In the absence of selection, some *ts* splicing mutants could rapidly splice if not stringently maintained at the non-permissive temperature, as indicated by the small amount of spliced product detected from a few clones induced at 37°C. The selection system was then adapted to HTS by simply adding GFP, a growth readout amenable to automation. HTS was sensitive and robust, yielding positive results even with the *M. leprae GyrA* intein. Although the *M. leprae GyrA* intein failed to produce detectable spliced product in Western blots, enough quinolone^s *GyrA* was made to block replication and cell growth. This reflects an intrinsic advantage of cell-

based assays because cells usually require far fewer molecules of functional enzyme than *in vitro* assays, compensating for poor expression of active protein or inefficient splicing. Moreover, cell-based assays better mimic the biological role of target proteins and lead compounds as compared to *in vitro* activity or binding assays (Persidis, 1998). In the future, the HTS could be used to rapidly screen compound libraries with very little cell preparation time.

The high frequency of *ts* splicing mutations suggests that this phenotype can result from a variety of structural distortions. Since mutant precursors were not degraded and since they were able to splice when shifted to permissive temperatures, it is likely that these mutations induced localized or minor changes in intein structure. Only 2 mutations (Thr72 and Ser53) were in amino acids that facilitate splicing by hydrogen bonding to splice junction residues to properly align the atoms and generate an electrophilic center at the scissile bond (Kawasaki *et al.*, 1997; Klabunde *et al.*, 1998; Noren *et al.*, 2000; Poland *et al.*, 2000). Neither residue is conserved in all inteins (Perler, 2000). Six *ts* mutations were found in the Mxe GyrA intein linker region (residues 109–160), including 2 in the disordered loop (residues 112–129). This suggests that the linker region is important for folding, structure or activity of the Mxe GyrA intein and may explain why some central deletions result in inactive inteins or *ts* splicing (Derbyshire *et al.*, 1997; Southworth *et al.*, 1998). Numerous mutations were selected in structural elements that are near the catalytic residues in intein Blocks B and G (β -strands B6, B7, B8, the C-terminal half of B5, and, B11 and its adjacent loops). Three pairs of *ts* mutations were observed in structurally superimposable backbone positions in each of the Mxe GyrA intein subdomains. The conservation of Mxe GyrA intein subdomain structure and sequence led to the prediction of a *ts* mutation at Thr3 based on the selected mutation at Thr72. This rationally designed mutation exhibited the predicted *ts* phenotype.

The structure-based alignment revealed an unusually high degree of sequence identity between the two

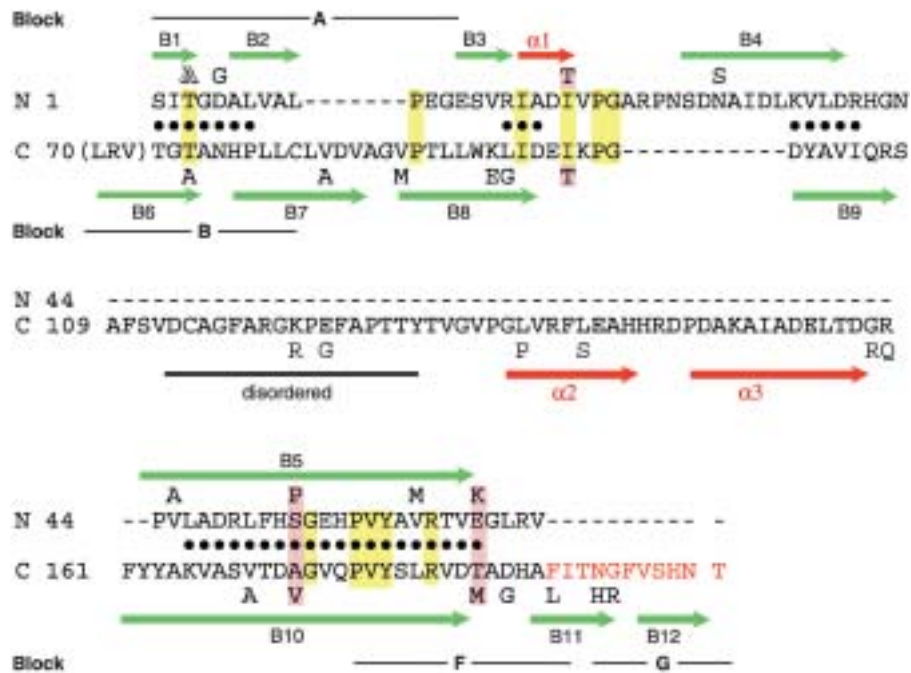


Figure 4. Structure-based sequence alignment of the *Mxe GyrA* intein showing subdomain similarities and *ts* mutations. The amino and carboxy-terminal subdomains are marked with an 'N' or 'C', respectively. Conserved intein motifs (Blocks A, B, F, and G) are indicated. Residues are present in both subdomains at 58 positions and 35 pairs can be superimposed with a RMS deviation of 1.50Å for all backbone atoms (black dots). Identical amino acids in both subdomains are highlighted in yellow. Green and red arrows depict β -strands and α -helices, respectively. Dashes represent gaps. The residues in parenthesis preceding the C-subdomain are part of the N-subdomain. The Polypeptide Ligation Region is in red. *Ts* mutations from clones with single substitutions are indicated above or below the mutated residue and mutations in structurally superimposable backbone

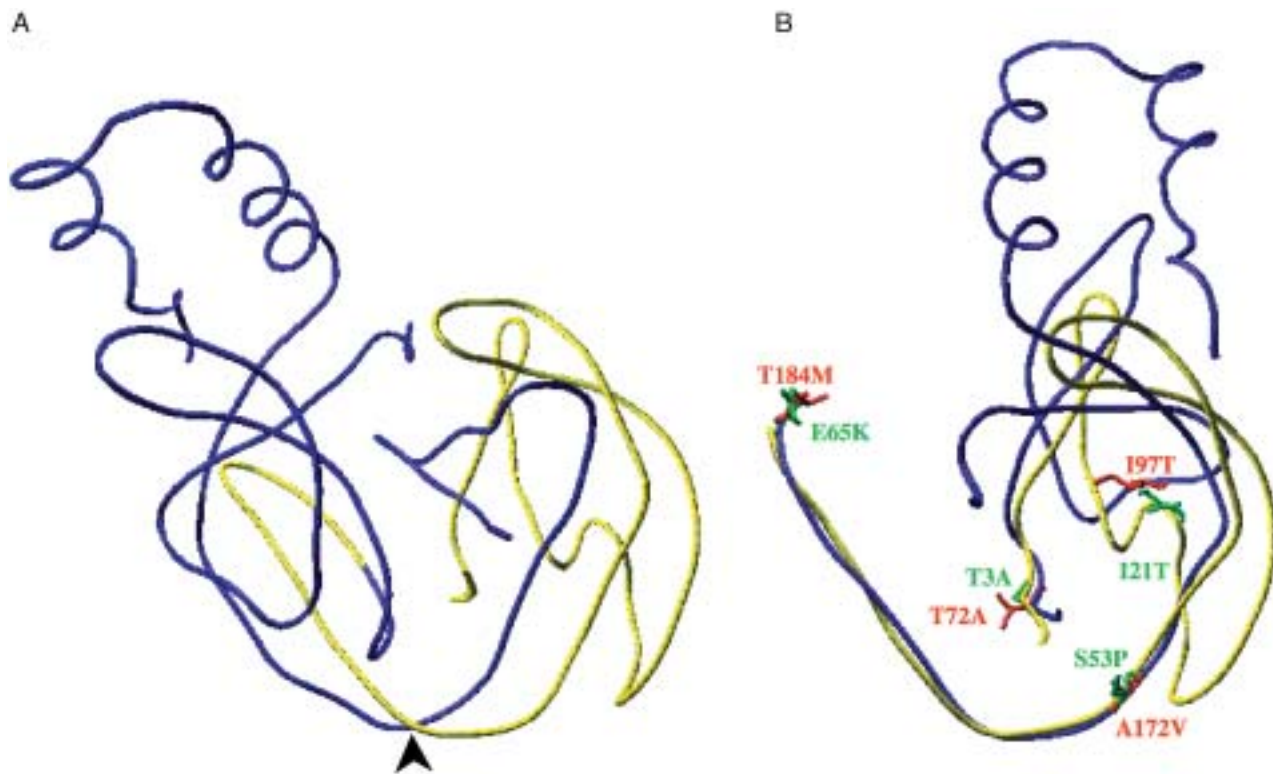


Figure 5. *Mxe GyrA* intein structure. (A) The backbone trace of the *Mxe GyrA* intein with the N-subdomain in yellow (residues 1-65) and the C-subdomain in blue (residues 70-184). The pseudo 2-fold axis of symmetry is indicated by the arrowhead and the N-terminus of the intein by 'N'. (B) Superposition of the two *Mxe GyrA* intein subdomains after rotation around the pseudo 2-fold axis of symmetry. *Ts* mutations in structurally homologous backbone positions in the 2 subdomains are indicated.

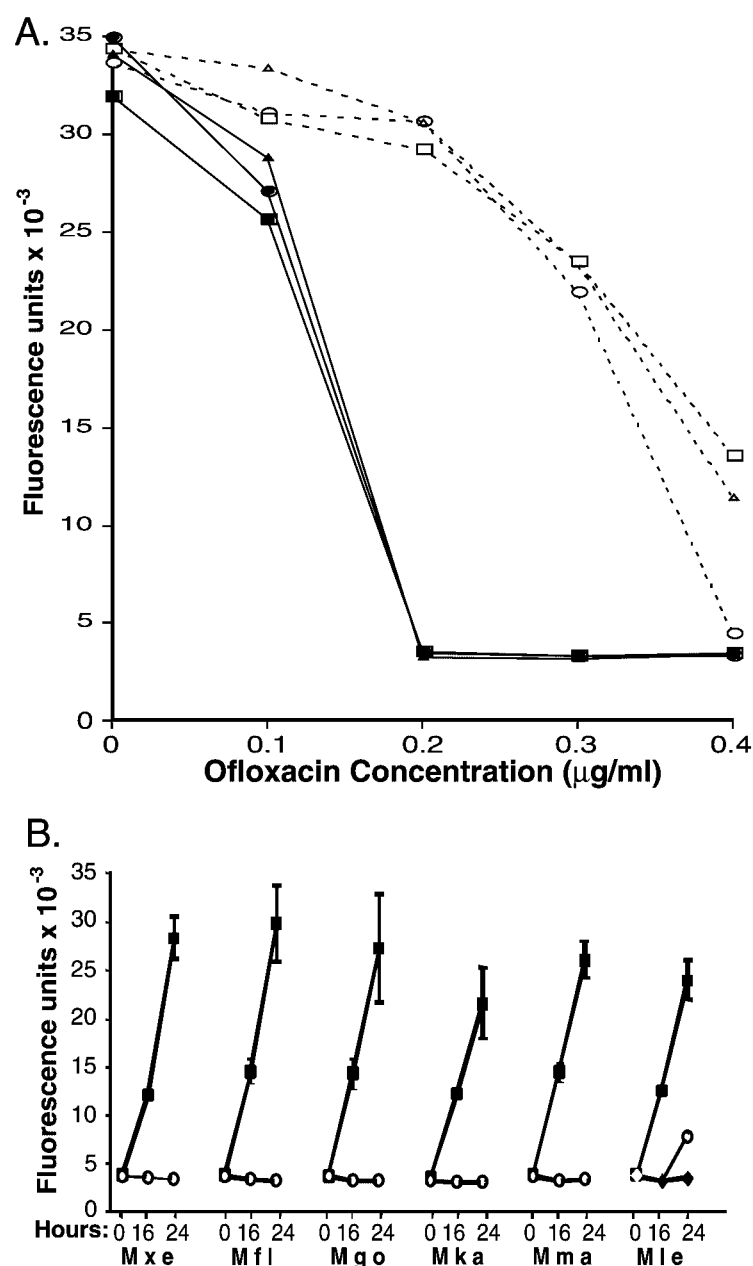


Figure 6. (A) Splicing of the *Mxe* GyrA intein in a high-throughput format. Growth of *E. coli* in a 96-well microtiter plate HTS format was monitored using fluorescence of GFP as a growth readout at 3 initial dilutions of inoculum (1:1024, squares; 1:2048, triangles; 1:4096, circles). Fluorescence at time zero was 4000 ± 100 for all samples. Fluorescence after 24 h of growth in the microtiter plates was plotted vs. ofloxacin concentration using active (black lines and symbols) or inactive (dashed lines, open symbols) *Mxe* GyrA inteins in *Eco* GyrA. (B) Splicing of various mycobacterial GyrA inteins in *Eco* GyrA assayed by HTS. Fluorescence during 24 h of growth was plotted at 0 (black squares), 0.25 (white circles) or 0.3 (black diamonds) $\mu\text{g/ml}$ ofloxacin with clones expressing different mycobacterial GyrA inteins present in *Eco* GyrA. Results of 3 independent experiments were averaged. GyrA intein in *Eco* GyrA: Mxe, *M. xenopi*; Mfl, *M. flavescens*; Mgo, *M. goodii*; Mka, *M. kansasii*; Mma, *M. malmoense*; Mle, *M. leprae*.

Mxe GyrA intein subdomains: of 58 positions with residues present in both subdomains, 60% can be superimposed with a RMS deviation of 1.50 \AA for all backbone atoms and, if strand B9 is aligned 2aa to the left, 24% are the same amino acid. With this degree of conservation, it was clear that 2 regions are present only in the C-subdomain. The first is residues 109–160,

which encode the linker region. In other inteins (including all other GyrA inteins), a homing endonuclease domain is present in this position. This suggests that the homing endonuclease domain was inserted after duplication of the subdomains. It is likely that inteins have both acquired and lost homing endonuclease domains during evolution (Figure 7). Insertion of mobile

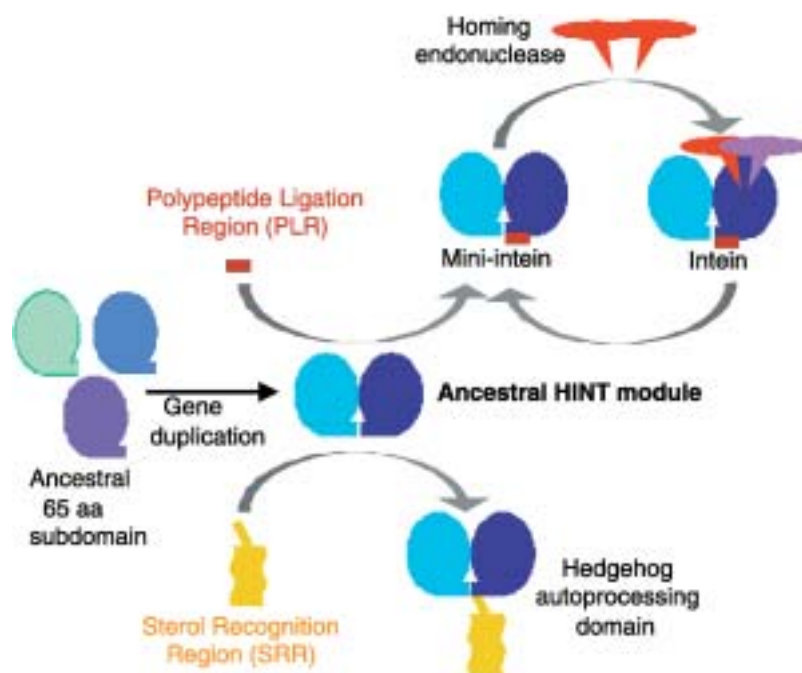


Figure 7. Evolution of inteins and Hedgehog-like autoprocessing proteins. Schematic drawing illustrating the events that may have occurred during evolution of Hedgehog protein autoprocessing domains and inteins. Prior to duplication, the ~65aa module may have functioned as a dimer. For clarity, polypeptides at the N-terminus of the HINT module and molecules to be ligated to these target polypeptides are not depicted.

homing endonuclease genes into intein genes would afford the endonuclease a safe refuge, since splicing would preserve host gene function.

The second region specific to the C-subdomain is an 11aa C-terminal extension, which positions the C-extein for ligation to the N-extein, and encodes the catalytic residues responsible for ligation of the exteins and release of the intein from the C-extein. This region is reminiscent of the Sterol Recognition Region (SRR) at the C-terminus of the Hedgehog autoprocessing domain that is required for ligation of cholesterol to the Hedgehog signaling domain. We suggest naming this C-terminal intein region the Polypeptide Ligation Region or PLR, since it provides protein splicing capabilities to the HINT module. The presence of different additions C-terminal to the ancestral HINT module in inteins and hedgehog autoprocessing proteins, led to a modification of HINT module evolution that was initially proposed by Hall *et al.* (1997). The progenitor HINT module may have been an ancient enzyme, capable of forming a reactive (thio)ester bond at the C-terminus of a fused target polypeptide. This linkage could then be directly attacked by numerous types of nucleophiles present in polypeptides or other molecules, resulting in ligation of the attacking moiety to the C-terminus of the target polypeptide and release from the HINT module. Initially, this ligation event may have occurred *in trans* with randomly associating molecules and could have been an early means of generating larger proteins or post-translational modifications (Perler, 1999; Pietrowski, 2001). Subsequently, residues C-terminal to

the core HINT module were added for selection and alignment of the attacking molecule (Figure 7). *Caenorhabditis elegans* has several genes with HINT domains linked to various SRR modules and N-terminal domains of unknown function (Aspöck *et al.*, 1999; Hall *et al.*, 1997). It is also possible that PLR and SRR regions evolved in steps, with residues first being added C-terminal to the core HINT module followed by further evolution into functional PLR and SRR regions.

The development of a positive selection system for inhibition of protein splicing is only the first step towards development of new drugs against mycobacterial infection. The GyrA selection system can be used with any intein that (a) splices at any insertion site in GyrA and (b) inactivates GyrA in the absence of splicing. For example, the *Methanococcus jannaschii* KlbA intein was cloned into Eco GyrA, yielding a small amount of spliced product in Western blots and inhibition of cell growth by ofloxacin in the HTS system (Maurice Southworth and F.B.P., unpublished). In principle, the GyrA selection system can be used with any functional combination of host cell and host GyrA protein. It is also possible that compounds that inhibit splicing of one intein will have a broader ability to inhibit splicing of other inteins. Such *trans*-acting inhibitors could serve as lead compounds for treatment of any pathogenic microbe containing an intein in an essential protein, as well as in other biotechnology applications. For example, inhibitors could control splicing *in vivo*, serving as conditional knockout reagents or allowing expression of cytotoxic proteins rendered harmless by intein insertion.

Experimental Procedures

Cloning and Expression

The Eco *gyrA* gene was amplified by PCR using primers 5'-GATAGGCTAGCGATGAGCGACCTTGCAGAG and 5'-TGAAGCA-ATTGAATTATTCTTCTTGGCTCG. Silent mutations were made to introduce unique NotI and BlnI sites surrounding Tyr122, generating pGyrA. The Mxe GyrA intein was amplified by PCR using primers XF (5'-CGACCCGCGCGCCGCAATGCGTTATTGCATCACGGGAG) and XR (5'-GCCAAAGGCGCTAAGCGGATTTCCGTGTTGTGGCT-GTGTGTGGCTGACGAACCCG), and cloned into pGyrA using NotI and BlnI. An inactive Mxe GyrA intein was cloned as above with primers containing mutations in catalytic residues (Cys1Ala and Asn198Ala). GyrA inteins from *M. leprae*, *M. goodsonae*, *M. malmoense*, *M. flavescens* and *M. kansasii* were similarly cloned into Eco *gyrA* in a pACYC184 derivative of pGyrA, as were active and inactive Mxe GyrA inteins. The GyrA fusions in pGyrA and its derivatives are under control of an IPTG inducible pTac promoter and have His tag at the C-terminus of the GyrA extein. All clones and PCR products were sequenced by the NEB core facility. Western blots used soluble lysates of cultures grown at 37°C to OD₆₀₀ = 0.5 and then induced at 37°C with 1 mM IPTG for 2 h. Western blots, including Prestained Protein Markers (NEB, Beverly, MA), were probed with anti-carboxy terminal His tag sera (Invitrogen, Carlsbad, CA).

Selection and Expression of Mxe GyrA Intein ts Splicing Mutants

Error-prone PCR reactions contained 1x Taq buffer, 1.5 mM MgCl₂, 200 μM dNTPs, 500 nM XF and XR primers, 50 ng DNA, 50 μM dPTP (Amersham, Piscataway, NJ), 50 μM 8-oxo-dGTP (Amersham) and 125 U/ml Taq polymerase (Promega, Madison, WI) in 20 μl (Zaccolo *et al.*, 1996). PCR mixtures were cycled at 94°C, 1 min; 45°C, 1 min; 72°C, 30 sec for 2 cycles and products digested with DpnI. A second PCR reaction contained 1x Taq buffer, 1.5 mM MgCl₂, 200 μM dNTPs, 500 nM primers XF and 5'-CATGGGCAGCTCTTCTTAAGCGGATTTTC, 1 μl of DpnI digested PCR product from the first reaction and 125 U/ml Taq polymerase for 30 cycles. Only intein sequences were mutated during error-prone PCR since primers included intein -1 and +1 codons. NotI/SapI digested PCR products were cloned in NotI/BlnI digested pGyrA. pGyrA derivatives were transformed into XL1-Blue (*gyrA96*, Stratagene, La Jolla, CA), plated on LB agar, 0.3 μg/ml ofloxacin (Sigma, St. Louis, MO), 100 μg/ml ampicillin and incubated at 37°C overnight. Clones were replica plated with velvet on fresh drug plates and incubated at 30°C for 3 days. Positives were retested after transformation into fresh cells. Samples for Western blot analysis were prepared from liquid cultures (in the absence of quinolone selection) as above or with an additional 15°C overnight incubation.

High-Throughput Screening

ER2818 (*gyrA96*, NEB) was transformed with pGFP (Clontech, Palo Alto, CA) and pACYC184 derivatives of pGyrA plasmids. Both GFP and GyrA fusions were under control of IPTG inducible pTac promoters. Media included 100 μg/ml ampicillin and 30 μg/ml chloramphenicol to maintain both plasmids. Plates were taped with plastic covers for 96-well plates (QIAGEN, Valencia, CA) that had been pierced to allow gas exchange and incubated with shaking (250 rpm) at 37°C. Growth was assayed by fluorescence in cultures excited at 400 nm and recorded at 500 nm in a VICTOR² V multilabel counter (Wallac, Turku, Finland) at 0, 16 and 24 h.

Structural Analysis of the Mxe GyrA Intein

The sequence alignment was based on Hall *et al.* (Hall *et al.*, 1997). The structural alignment of the Mxe GyrA intein subdomains and estimation of RMS deviations were performed using SwissPdbViewer 3.72b (Guex *et al.*, 1997) with Mxe GyrA intein coordinates (Brookhaven PDB accession number 1AM2) (Klabunde *et al.*, 1998). The alignment was manually refined at gaps to maximize sequence identity.

Acknowledgements

We thank M. Sibley and E. Raleigh for generating *gyrA96* strains, M. Southworth for data analysis and use of his unpublished data, S. Cole for mycobacterial GyrA intein DNA, and P. Forterre, W. Jack, S. Smith, I. Cann, E. Raleigh, G. Tzertzinis and M. Southworth for helpful discussions. We thank Don Comb for support and encouragement.

References

- Aspöck, G., Kagoshima, H., Niklaus, G., and Burglin, T.R. 1999. *Caenorhabditis elegans* has scores of hedgehog-Related genes: sequence and expression analysis. *Genome Res.* 9: 909–923.
- Dalgaard, J.Z., Moser, M.J., Hughey, R., and Mian, I.S. 1997. Statistical modeling, phylogenetic analysis and structure prediction of a protein splicing domain common to inteins and hedgehog proteins. *J. Comput. Biol.* 4: 193–214.
- 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: 644–653.
- Davis, E.O., Thangaraj, J.S., Brooks, P.C., and Colston, M.J. 1994. Evidence of selection for protein introns in the RecAs of pathogenic mycobacteria. *EMBO J.* 13: 699–703.
- Derbyshire, V., Wood, D.W., Wu, W., Dansereau, J.T., Dalgaard, J.Z., and Belfort, M. 1997. Genetic definition of a protein-splicing domain: functional mini- inteins support structure predictions and a model for intein evolution. *Proc. Natl. Acad. Sci. USA.* 94: 11466–11471.
- Duan, X., Gimble, F.S., and Quijcho, F.A. 1997. Crystal structure of PI-SceI, a homing endonuclease with protein splicing activity. *Cell.* 89: 555–564.
- Fsihi, H., Vincent, V., and Cole, S.T. 1996. Homing events in the *gyrA* gene of some mycobacteria. *Proc. Natl. Acad. Sci. USA.* 93: 3410–3415.
- Gimble, F.S. 2000. Invasion of a multitude of genetic niches by mobile endonuclease genes. *FEMS Microbiol Lett.* 185: 99–107.
- Guex, N., and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis.* 18: 2714–2723.
- Hall, T.M., Porter, J.A., Young, K.E., Koonin, E.V., Beachy, P.A., and Leahy, D.J. 1997. Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell.* 91: 85–97.
- Ichiyanagi, K., Ishino, Y., Ariyoshi, M., Komori, K., and Morikawa, K. 2000. Crystal structure of an archaeal intein-encoded homing endonuclease PI- Pful. *J. Mol. Biol.* 300: 889–901.
- Jacobs, M.R. 1999. Activity of quinolones against mycobacteria. *Drugs.* 58: 19–22.
- Jurica, M.S., and Stoddard, B.L. 1999. Homing endonucleases: structure, function and evolution. *Cell. Mol. Life Sci.* 55: 1304–1326.
- Kawasaki, M., Nogami, S., Satow, Y., Ohya, Y., and Anraku, Y. 1997. Identification of three core regions essential for protein splicing of the yeast Vma1 protozyme. A random mutagenesis study of the entire Vma1- derived endonuclease sequence. *J. Biol. Chem.* 272: 15668–15674.
- Klabunde, T., Sharma, S., Telenti, A., Jacobs, W.R., Jr., and Sacchettini, J.C. 1998. Crystal structure of GyrA intein from Mycobacterium xenopi reveals structural basis of protein splicing. *Nat. Struct. Biol.* 5: 31–36.
- Lamden, K., Watson, J.M., Knerer, G., Ryan, M.J., and Jenkins, P.A. 1996. Opportunist mycobacteria in England and Wales: 1982 to 1994. *Commun. Dis. Rep. CDR Rev.* 6: R147–R151.
- Misteli, T., and Spector, D.L. 1997. Applications of the green fluorescent protein in cell biology and biotechnology. *Nat. Biotechnol.* 15: 961–964.
- Nogami, S., Satow, Y., Ohya, Y., and Anraku, Y. 1997. Probing novel elements for protein splicing in the yeast Vma1 protozyme: a study of replacement mutagenesis and intragenic suppression. *Genetics.* 147: 73–85.
- Noren, C.J., Wang, J., and Perler, F.B. 2000. Dissecting the chemistry of protein splicing and its applications. *Angew. Chem. Int. Ed. Engl.* 39: 450–466.
- Petrokovski, S. 2001. Intein spread and extinction in evolution. *Trends Genet.* 17: 465–472.
- Perler, F.B. 1999. A natural example of protein trans-splicing. *Trends in Biol. Sci.* 24: 209–210.
- Perler, F.B. 2000. InBase, the Intein Database. *Nucleic Acids Res.* 28: 344–345.
- Persidis, A. 1998. Lead validation platforms. *Nat. Biotechnol.* 16: 100–101.
- Petrokovski, S. 1998. Modular organization of inteins and C-terminal autocatalytic domains. *Protein Sci.* 7: 64–71.
- 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: 16408–16413.

- Sander, P., Alcaide, F., Richter, I., Frischkorn, K., Tortoli, E., Springer, B., Telenti, A., and Bottger, E.C. 1998. Inteins in mycobacterial GyrA are a taxonomic character. *Microbiology*. 144: 589–591.
- Southworth, M.W., Adam, E., Panne, D., Byer, R., Kautz, R., and Perler, F.B. 1998. Control of protein splicing by intein fragment reassembly. *Embo J.* 17: 918–926.
- Telenti, A., Southworth, M., Alcaide, F., Daugelat, S., Jacobs, W.R., Jr., and Perler, F.B. 1997. The *Mycobacterium xenopi* GyrA protein splicing element: characterization of a minimal intein. *J. Bacteriol.* 179: 6378–6382.
- Wood, D.W., Wu, W., Belfort, G., Derbyshire, V., and Belfort, M. 1999. A genetic system yields self-cleaving inteins for bioseparations. *Nat. Biotechnol.* 17: 889–892.
- Zaccolo, M., Williams, D.M., Brown, D.M., and Gherardi, E. 1996. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* 255: 589–603.