Bacteriophage-Based Genetic System for Selection of Nonsplicing Inteins

Isaac K. O. Cann,[†] Kensey R. Amaya,[†] Maurice W. Southworth,
and Francine B. Perler*

New England Biolabs, Inc., Beverly, Massachusetts 01915

Received 3 September 2003/Accepted 21 January 2004

A genetic selection system that detects splicing and nonsplicing activities of inteins was developed based on the ability to rescue a T4 phage strain with a conditionally inactive DNA polymerase. This phage defect can be complemented by expression of plasmid-encoded phage RB69 DNA polymerase. Insertion of an intein gene into the active site of the RB69 DNA polymerase gene renders polymerase activity and phage viability dependent on protein splicing. The effectiveness of the system was tested by screening for thermosensitive splicing mutants. Development of genetic systems with the potential of identifying protein splicing inhibitors is a first step towards controlling proliferation of pathogenic microbes harboring inteins in essential proteins.

Over the last decade, intensive research has yielded important insights into the mechanism by which biological activity is restored to proteins invaded by inteins (4, 5, 12, 13). Unlike introns, which are removed from RNA before translation, inteins are cotranslated with the invaded protein to form a precursor polypeptide (4, 5, 13). The intein is then self-catalytically excised from the precursor, with concomitant ligation of the upstream (N extein) and downstream (C extein) flanking polypeptides (4, 5), to yield two proteins: the mature host protein and the intein. Inteins are often found in active sites and conserved motifs of proteins indispensable to cell metabolism (6, 8, 9). Inteins are therefore new antimicrobial targets in pathogens with inteins in essential proteins since splicing must occur for normal cell function. To this end, rapid methods are desired to screen for compounds capable of blocking protein splicing. Three *in vivo* methods for detection of splicing and nonsplicing intein variants were recently published (1, 3, 11).

This report describes a new genetic selection system for the identification of splicing and nonsplicing intein variants inserted into phage RB69 DNA polymerase. The method is based on growth versus lysis of *Escherichia coli* cells infected with conditionally defective T4 gp43[−] phage, which contains amber mutations in the T4 DNA polymerase gene (gene 43) that render the phage inviable in nonsuppressor strains. As a result, colony formation is observed with T4-susceptible *E. coli* strains lacking amber suppressors, such as ER2566. Plasmid-borne DNA polymerase from the closely related phage RB69 can complement this defect in T4 gp43[−] phage, resulting in cell lysis (10). This system for controlling phage viability was converted into a genetic selection system for protein splicing by in-frame insertion of an intein gene into the active site of the plasmid-encoded RB69 DNA polymerase gene (Fig. 1), rendering the RB69 DNA polymerase inactive in the absence of

protein splicing. T4 gp43[−] phage viability would then require protein splicing to produce active RB69 DNA polymerase (Fig. 2).

The *pol-2_{Tli}* intein gene portion of *Thermococcus litoralis* (Vent) DNA polymerase (7) was cloned into the homologous active-site, region I motif of the RB69 DNA polymerase gene in pCW19R (10). All methods were carried out according to the manufacturer's instructions. StuI and SacII cloning sites were introduced into the RB69 DNA polymerase gene by site-directed mutagenesis (Quikchange; Stratagene), and the resultant plasmid was designated pCW19R'. Next, an endonuclease-deficient mutant (2) of the Pol-2_{Tli} intein was amplified by PCR with Vent DNA polymerase (New England Biolabs). The forward primer, TliPol-2F (5' aaagaggccttcgtcctttatggga cAGTGTCTCCGGAGAAAGT), incorporated a StuI site (underlined) and a BspEI site (bold), and the reverse primer, TliPol-2R (5' ccagaccgaggagacataaatgctgtcagtATTGTGTAC CAG), incorporated a SacII site. Aside from the intein sequence, each primer coded for eight RB69 DNA polymerase residues from region I (lowercase), rather than the *T. litoralis* flanking residues (Fig. 1A). The Pol-2_{Tli} intein PCR product and pCWR19' DNA were digested with StuI and SacII and ligated to create pTli Pol-2/I^A containing the *pol-2_{Tli}* intein gene inserted in frame in region I of the RB69 *pol* gene (Fig. 1B). A precursor with an inactive intein (pTli Pol-2/I^{IN}) was made by generating a Ser1Ala mutation at the N-extein junction (Fig. 1C).

To determine whether protein splicing yielded active RB69 DNA polymerase that could complement the defective T4 gp43[−] phage, a cell line comprising either pTli Pol-2/I^A (Fig. 3A) or pTli Pol-2/I^{IN} (Fig. 3B) in ER2566 cells was challenged with the defective T4 gp43[−] phage (plates II) or wild-type T4 phage (plates III) at room temperature (ca. 25°C) or 37°C. Each cell line was spread at 3 × 10⁷ cells/plate on Luria-Bertani (LB)-ampicillin plates with approximately 2.5 × 10⁷ PFU of T4 gp43[−] phage. When cells were plated with freshly spread T4 gp43[−] phage, only cells expressing the active intein fusion lysed (Fig. 3A, plate II). All cells incubated without phage (plate I) and the pTli Pol-2/I^{IN} cells incubated with T4 gp43[−] phage (Fig. 3B, plate II) yielded confluent growth. All

* Corresponding author. Mailing address: New England Biolabs, Inc., 32 Tozer Rd., Beverly, MA 01915. Phone: (978) 927-5054, ext. 326. Fax: (978) 921-1350. E-mail: perler@neb.com.

[†] Present address: Laboratory of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801.

A

(i) Native sequence

GAA GCT TTC GTT CTT TAT GGT GAT ACA GAC TCT ATT TAC GTA TCT GCT GAT
 E A F V L Y G D T D S I Y V S A D
 Region I of RB69 DNA polymerase

(ii) Mutated sequence in pCW19R'

*Stu*I *Sac*II
GAG GCC TTC GTT CTT TAT GGT GAT ACA GAC TCT ATT TAC GTA TCC GCG GAT
 E A F V L Y G D T D S I Y V S A D
 Region I of RB69 DNA polymerase

B Tli Pol-2/I^A C Tli Pol-2/I^{IN}

FIG. 1. Construction of pTli Pol-2/I^A and pTli Pol-2/I^{IN}. (A) The sequence surrounding the RB69 family B DNA polymerase region I (underlined) in the native gene (i) and the mutated gene (ii) are shown with the engineered restriction enzymes sites also underlined. Schematic diagram of the RB69 DNA polymerase (white boxes, exteins) and Pol-2_{Tli} intein (gray box) precursor containing either an active intein (B) pTli Pol-2/I^A or an inactive intein (C) (pTli Pol-2/I^{IN}). Splicing is required to generate a functional RB69 DNA polymerase. Intein amino acid sequences are indicated above the precursor, and DNA polymerase sequences (exteins) are indicated below. pTli Pol-2/I^{IN} has a Ser1-to-Ala mutation that results in C-terminal cleavage in the absence of splicing.

FIG. 2. Genetic selection system for protein splicing. Lysis of *E. coli* cells by the T4 gp43⁻ phage requires complementation of the phage DNA polymerase defect. A plasmid-borne RB69 DNA polymerase containing an intein can complement the T4 gp43⁻ phage defect only if protein splicing occurs. See the text for details.

cells were lysed by wild-type T4 phage (plates III) irrespective of intein activity, thus showing that survival of pTli Pol-2/I^{IN} cells was not due to phage resistance. These results demonstrate that the Pol-2_{Tli} intein can splice in the RB69 DNA polymerase precursor to generate a functional RB69 DNA polymerase that complements the defect in T4 gp43⁻, using background levels of RB69 DNA polymerase expression and

without a requirement for overexpression of the fusion protein by isopropyl-β-D-thiogalactopyranoside induction. No complementation was observed with an inactive intein fusion or if the Pol-2_{Tli} intein was inserted into RB69 DNA polymerase with eight *T. litoralis* DNA polymerase residues flanking both sides of the intein instead of RB69 residues (data not shown).

The efficiency of the system was further tested by looking for thermosensitive splicing variants generated by error-prone PCR. Mutagenized intein genes were cloned into the wild-type RB69 DNA polymerase gene, and the resultant transformants were grown on master plates in the absence of phage. The pol-2_{Tli} intein gene was mutagenized by incorporation of dPTP and 8-oxo-dGTP, which stimulate transitions and transversions during PCR amplification (14). The primers used were TliPol-2F and TliPol-2R, and the PCR protocol was as described previously (1), except that the concentration of each nucleotide analog was 5 μM. PCR products were purified, digested with StuI and SacII, and cloned into StuI/SacII-digested pTli Pol-2/I^A. The ligation mix was used to electroporate *E. coli* ER2566 cells, which were then spread on LB-ampicillin plates and incubated at 37°C. Two different primary screening methods were used. In the first method, transformants were streaked on LB-ampicillin plates without phage (master plate) and on two LB-ampicillin plates with T4 gp43⁻ phage (3.5 × 10⁸ PFU/plate): one for incubation at 37°C and the other at room temperature. The second primary screen involved replica plating a master plate of transformants with velvet onto two LB-ampicillin plates with T4 gp43⁻ phage and incubating as described for the streak method. Colonies corresponding to thermosensitive variants were picked from the master plate, and their plasmids were extracted for a secondary screen with fresh *E. coli* ER2566 to confirm the phenotype, using the streak method. Cell lines with pTli Pol-2/I^A and pCW19R (wild-type RB69 DNA polymerase construct) served as positive controls, and those with pTli Pol-2/I^{IN} provided a negative control. All DNA constructs were sequenced by the DNA sequencing facility of New England Biolabs.

Inteins with a temperature-sensitive splicing phenotype are characterized by a failure to splice at elevated temperatures permitting colony formation, while splicing at lower temperatures yields active RB69 DNA polymerase and cell lysis. Four out of 3,550 clones and 8 out of 6,000 clones screened by the streak and replica plating methods, respectively, displayed a temperature-sensitive phenotype. To eliminate the possibility of phage-resistant clones, the phenotype was confirmed by transforming plasmids from these temperature-sensitive strains into fresh cells (Fig. 4). The specific mutations present in eight selected temperature-sensitive clones are listed in Table 1. Approximately 30% of all clones grew at both temperatures, indicating that they contained mutations leading to inactive inteins.

A screening system that can select for both active and inactive inteins was developed. Experiments with control cells expressing active or inactive inteins indicate that complementation of the DNA polymerase defect in T4 gp43⁻ requires a functional intein. This study also demonstrates that the Pol-2_{Tli} intein from a hyperthermophilic archaeon can splice in the mesophilic phage RB69 DNA polymerase at temperatures as low as ~25°C. The system was sensitive enough to isolate 12 temperature-sensitive clones out of 10,000.

FIG. 3. Complementation of the DNA polymerase defect in T4 gp43⁻ phage. Cells expressing active intein fusions (pTli Pol-2/I^A) (A) or inactive intein fusions (pTli Pol-2/I^{IN}) (B) were challenged with no phage (plates I), T4 gp43⁻ phage (plates II), or wild-type T4 phage (plates III) at the indicated temperatures.

FIG. 4. Secondary screening of genetically selected clones expressing a temperature-sensitive splicing phenotype. *E. coli* was transformed with plasmids isolated from clones expressing the temperature-sensitive splicing phenotype in the primary screen, and transformants were retested by streaking on plates without phage, with T4 gp43⁻ phage, or with wild-type T4 phage at the indicated temperatures. pCW19R expressing RB69 DNA polymerase (streak a), pTli Pol-2/I^A (streak b), or pTli Pol-2/I^{IN} (streak c) controls were streaked on the top row. Eight temperature-sensitive splicing mutants (Table 1) were streaked in the second (TS1, TS2, TS3, and TS4, from streak method) and third (TS5, TS6, TS7, and TS8, from replica plating method) rows.

No essential intein nucleophile or known assisting residue was mutated, which was not unexpected since temperature-sensitive phenotypes are often the result of minor structural distortions (1) and since six N-terminal and four C-terminal intein residues were fixed in the PCR primers. The high percentage of clones that failed to splice at both temperatures was probably due to a high level of mutation, as indicated in the sequenced temperature-sensitive variants, which all contained multiple substitutions, except for TS6. None of the temperature-sensitive mutations was resolved, as the objective for the thermosensitivity experiment was to determine the capacity of the selection system to detect nonsplicing intein variants, which is a prerequisite before using the system to test for intein inhibitors.

These results show that the system is amenable for testing chemical compound or natural product libraries in high-throughput microtiter plate formats. The replica plating method for primary screening will enable screening of large numbers of clones in experiments aimed at isolating expressed peptides that inhibit protein splicing or for exploring residues

involved in intein structure and function as described elsewhere (11). It is anticipated that compounds capable of inhibiting protein splicing in this context will function also as inhibitors of inteins in their natural insertion sites. While this selection system could be a very robust screening method, development of resistance by *E. coli* to T4 phage infection could be a setback. Introduction of positive and negative controls (Fig. 4, streaks a, b, and c), as in the secondary screening step, should minimize selection for false-positives.

We thank D. G. Comb for discussions and financial support.
We thank J. D. Karam for providing the T4 phage and the RB69 DNA polymerase construct (pCW19R).

TABLE 1. Amino acid changes in inteins showing temperature-sensitive phenotypes	
Clone	Intein mutation(s)
TS1.....	D23N, N135S
TS2.....	N95S, D205G, K218R, R334K
TS3.....	L46P, R108K, Y241N, P263S, W318R, K359E
TS4.....	E86K, S172P, K203R, F353L
TS5.....	N128D, F221V, A272V
TS6.....	S123A
TS7.....	K15R, E111G, G150V, E166G, S212G, E236V
TS8.....	N78D, D273G, F381L

REFERENCES

1. Adam, E., and F. B. Perler. 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**:479–487.
2. Hodges, R. A., F. B. Perler, C. J. Noren, and W. E. Jack. 1992. Protein splicing removes intervening sequences in an archaea DNA polymerase. *Nucleic Acids Res.* **20**:6153–6157.
3. Lew, B. M., and H. Paulus. 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**:169–177.
4. Noren, C. J., J. Wang., and F. B. Perler. 2000. Dissecting the chemistry of protein splicing and its applications. *Agnew. Chem. Int. Ed.* **39**:450–466.
5. Paulus, H. 2000. Protein splicing and related forms of protein autoprocesing. *Annu. Rev. Biochem.* **69**:447–496.
6. Perler, F. B. 2002. InBase: the intein database. *Nucleic Acids Res.* **30**:383–384.
7. Perler, F. B., D. G. Comb, W. E. Jack, L. S. Moran, B. Qiang, R. B. Kucera, J. Benner, B. E. Slatko, D. O. Nwankwo, S. K. Hempstead, C. K. S. Carlow, and H. Jannasch. 1992. Intervening sequences in an Archaea DNA polymerase gene. *Proc. Natl. Acad. Sci. USA* **89**:5577–5581.
8. Saves, I., L.-A. Lewis, F. Westrelin, R. Warren, M. Daffe, and J.-M. Masson. 2002. Specificities and functions of the *recA* and *pps1* intein genes of *Mycobacterium tuberculosis* and application for diagnosis of tuberculosis. *J. Clin. Microbiol.* **40**:943–950.

9. Telenti, A., M. Southworth, F. Alcaide, S. Daugelat, W. R. Jacobs, and F. B. Perler. 1997. The *Mycobacterium xenopi* GyrA protein splicing element: characterization of a minimal intein. *J. Bacteriol.* **179**:6378–6382.
10. Wang, C. C., A. Pavlov, and J. D. Karam. 1997. Evolution of RNA-binding specificity in T4 DNA polymerase. *J. Biol. Chem.* **272**:17703–17710.
11. Wood, D. W., W. Wu, G. Belfort, V. Derbyshire, and M. Belfort. 1999. A genetic system yields self-cleaving inteins for bioseparation. *Nat. Biotechnol.* **17**:889–892.
12. Xu, M.-Q., and F. B. Perler. 1996. The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **15**:5146–5153.
13. Xu, M.-Q., M. W. Southworth, F. B. Mersha, L. J. Honstra, and F. B. Perler. 1993. *In vitro* protein splicing of purified precursor and the identification of a branched intermediate. *Cell* **75**:1371–1377.
14. Zaccolo, M., D. M. Williams, D. M. Brown, and E. Gherardi. 1996. An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J. Mol. Biol.* **255**:589–603.