Positive-selection vector for direct protein expression

Andreas F. Haag1,2 and Christian Ostermeier2

1Department of Biotechnology & Bioinformatics, Weihenstephan University of Applied Sciences, Freising, Germany, and 2Novartis Institutes for BioMedical Research, Basel, Switzerland

Keywords: positive selection; direct selection; direct expression; protein expression; seamless cloning

We describe the development of a novel positive-selection vector, RHP-AmpS, that is suitable for seamless cloning and high-level protein expression in Escherichia coli. In this vector, β-lactamase (Bla) was rendered nonfunctional by replacing the codon for the C-terminal amino acid of the β-lactamase gene (bla) with a stop codon. Insertion of a target gene in the correct orientation (tail to tail) results in the reconstruction of the C-terminal codon (W290) of bla. This restores the function of the gene and allows the selection of positive recombinants on agar plates containing ampicillin. To allow a high level of protein expression, this selection cassette was inserted into the T7 polymerase–based expression cassette of the Novagen pET28a expression vector. To our knowledge, this is the first example of true positive-selection cloning and direct, high-level expression from a single vector.

Introduction

As more and more genome sequences become available, researchers are presented with a multitude of new, not-yet described proteins. For many experiments, these proteins have to be overexpressed and purified to have sufficient material for analysis. In addition to cloning the full-length gene into an expression vector, multiple, truncated, and/or mutated versions of the gene often have to be cloned and expressed to obtain sufficient and soluble protein for downstream applications. However, cloning reactions are very inefficient processes and vector background from parental plasmids (template vector and cloning vector) can make it very tedious to find the right clone. In modern high-throughput structural biology laboratories, an increasing number of gene constructs need to be cloned and expressed as quickly as possible in order to characterize their protein products and interactions in a timely manner. Reducing the time and effort spent on identifying the correct clones reduces costs significantly and leads to higher throughput at the cloning interface. Various ways to address the problem of parental vector background have been described (1). However, all of these classical methods still have a considerable background, need unusual selection media, or are not suitable for direct protein expression.

To facilitate the selection of correct recombinants, a number of positive-selection vectors have been developed; in these vectors, the successful cloning of DNA fragments results in an obvious change of phenotype. The positive selection in these vectors is achieved either by the inactivation of a genetic marker (2–9) or the replacement of said marker by the target gene (10–13). For a detailed review on positive selection vectors, see Reference 14. These powerful selection strategies, however, are often only suitable for cloning; direct high-level protein expression from the cloning vector is not possible. In addition, most of the known selection systems have no active, positive selection toward the correct insert in its correct orientation and therefore incorrectly inserted fragments can result in a surviving phenotype.

Huang et al. (15) found that 43 out of the 263 amino acids that form the core protein of β-lactamase (Bla) are essential for its full function. Interestingly, the last amino acid at the C-terminus of Bla, tryptophan (W290), was one of the residues that did not tolerate any changes. We reasoned that using this observation, a positive-selection vector could be constructed containing a truncated version of bla that is missing the C-terminal codon for W290. Cloning a target gene carrying the respective sequence to reconstruct this amino acid residue at the C-terminus of Bla into the vector would result in restoring the protein’s full functionality (Figure 1). Colonies carrying the correct insert in the correct orientation would therefore be ampicillin-resistant and could be easily selected. However, no fusion protein with Bla is created, as the two genes are oriented in opposing transcriptional directions and are each separated by two stop codons, respectively.

By using this approach, we have developed the vector RHP-AmpS (rubredoxin/His6/PreScission ampicillin selection; Figure 2, GenBank accession no. FJ545754), which is capable of both positive selection and direct, high-level protein expression.

Materials and methods

PCR and primer design

All PCR reactions were performed using Phusion High-Fidelity DNA Polymerase (Finnzymes Oy, Espoo, Finland) in accordance with the manufacturer’s protocol. Primers were designed to have a melting temperature (Tm) of ≥65°C using the Finnzymes Tm calculator (www.finnzymes.com/tm_determination_old.html). Primers used for cloning with homologous recombination had a 20-bp 5’ non-anealing tail homologous to the termini of the linearized vector. For cloning into the positive selection vector RHP-AmpS, the reverse primer also contained the sequence to restore the C-terminal tryptophan of Bla and a double stop codon for selective marker and target gene, respectively.

For type IIs restriction enzyme cloning (16), the sequences homologous to the vector ends were replaced by BsaI recognition sites. The sites were designed so that digesting with BsaI would result in the loss of these recognition sequences.

Cloning by homologous recombination

Cloning was performed using a protocol for homologous recombination (modified from Reference 17). The respective vectors were linearized by PCR and purified. One microliter of linearized vector and 1 μL of PCR-amplified target gene were added to 50 μL of One Shot TOP10 competent cells (Invitrogen AG, Basel, Switzerland). Cells were incubated on ice for 15 min, heat-shocked at 42°C for 45 s and recovered with 200 μL of SOC medium (1) for 1 h at 37°C, 650 rpm before plating onto Luria-Bertani.
media (1) containing the selective antibiotic. The plates were incubated overnight at 37°C and 8–16 colonies per cloning reaction were analyzed further.

Cloning by type IIs restriction enzymes
Vector and insert were digested separately with BsaI (New England Biolabs, Ipswich, MA, USA) according to the manufacturer's protocol. Ligation reactions were set up using T4 DNA ligase (New England Biolabs) and a 3:1 ratio of insert to vector according to the manufacturer’s protocol. One to five microliters of the ligation reaction were transformed into 50 μL of One Shot TOP10 competent cells (Invitrogen AG) as described above.

Amplification and cloning of the truncated bla gene
The bla gene was amplified from the pETBlue-1 vector (Novagen, Merck KGaA, Darmstadt, Germany) using primers AH02 and AH03 (Table 1). AH02 was designed to include a region 394 bp upstream of the ATG start codon to ensure that the gene’s own promoter was included. AH03 was designed to amplify bla without the sequence encoding for W290 (blaΔW290). RHP-CcdB was linearized using primers AH01 and AH10 (Table 1) and the truncated bla gene cloned into the vector by homologous recombination resulting in RHP-AmpS1 (similar to RHP-AmpS but carrying the pETBlue-1 bla promoter; Figure 2).

Amplification of Cm(R) and pET17b-bla promoter
A 104-bp fragment of the chloramphenicol acetyltransferase promoter from the Cm(R) cassette (5′-CCGCATTAG-GCACCACCAGGCTTACCTTATGCTCCGCTGATATA-\text{GTGTTGAGTTTGGATGAGT}-GATCCGCAGAGATTTTGAGTTAG-GAGCTAAGGAGCTAA-3′) was amplified using primers AH04 and AH05 (Table 1), each carrying a 20-bp non-annealing tail encoding a homologous sequence in order to replace the pETBlue-1 promoter of RHP-AmpS1. A 76-bp fragment prior to the start codon of bla from pET17b (5′-AATACATTCAAATATGTATCCGC-3′) was amplified using primers AH06 and AH07 (Table 1) as a further alternative to the Cm(R) promoter. RHP-AmpS1 was linearized using the primers AH10 and AH11 (Table 1) and the two promoter fragments were cloned using homologous recombination.

Amplification of GFP and mCherry for W290 reconstruction
The gene for GFP used for homologous recombination cloning was amplified using primers AH08 and AH09 (Table 1) and for BsaI cloning it was amplified using the primers AH12 and AH13 (Table 1). The gene for mCherry was amplified using primers AH14 and AH15 (Table 1). Both the gfp and the mCherry gene were cloned into RHP-AmpS linearized with primers AH01 and AH16 (Table 1). In addition, the gfp gene was also cloned into RHP-AmpS using type IIs restriction enzyme cloning.

Test of selection
Our test genes were cloned into the vector as described previously. After the recovery, the transformation assays were spread onto LB agar plates (1) containing the selective antibiotic. The plates were incubated overnight at 37°C and 8–16 colonies per cloning reaction were analyzed further.

**Cloning by type IIs restriction enzymes**
Vector and insert were digested separately with BsaI (New England Biolabs, Ipswich, MA, USA) according to the manufacturer’s protocol. Ligation reactions were set up using T4 DNA ligase (New England Biolabs) and a 3:1 ratio of insert to vector according to the manufacturer’s protocol. One to five microliters of the ligation reaction were transformed into 50 μL of One Shot TOP10 competent cells (Invitrogen AG) as described above.

**Amplification and cloning of the truncated bla gene**
The bla gene was amplified from the pETBlue-1 vector (Novagen, Merck KGaA, Darmstadt, Germany) using primers AH02 and AH03 (Table 1). AH02 was designed to include a region 394 bp upstream of the ATG start codon to ensure that the gene’s own promoter was included. AH03 was designed to amplify bla without the sequence encoding for W290 (blaΔW290). RHP-CcdB was linearized using primers AH01 and AH10 (Table 1) and the truncated bla gene cloned into the vector by homologous recombination resulting in RHP-AmpS1 (similar to RHP-AmpS but carrying the pETBlue-1 bla promoter; Figure 2).

**Amplification of Cm(R) and pET17b-bla promoter**
A 104-bp fragment of the chloramphenicol acetyltransferase promoter from the Cm(R) cassette (5′-CCGCATTAG-GCACCACCAGGCTTACCTTATGCTCCGCTGATATA-\text{GTGTTGAGTTTGGATGAGT}-GATCCGCAGAGATTTTGAGTTAG-GAGCTAAGGAGCTAA-3′) was amplified using primers AH04 and AH05 (Table 1), each carrying a 20-bp non-annealing tail encoding a homologous sequence in order to replace the pETBlue-1 promoter of RHP-AmpS1. A 76-bp fragment prior to the start codon of bla from pET17b (5′-AATACATTCAAATATGTATCCGC-3′) was amplified using primers AH06 and AH07 (Table 1) as a further alternative to the Cm(R) promoter. RHP-AmpS1 was linearized using the primers AH10 and AH11 (Table 1) and the two promoter fragments were cloned using homologous recombination.

**Amplification of GFP and mCherry for W290 reconstruction**
The gene for GFP used for homologous recombination cloning was amplified using primers AH08 and AH09 (Table 1) and for BsaI cloning it was amplified using the primers AH12 and AH13 (Table 1). The gene for mCherry was amplified using primers AH14 and AH15 (Table 1). Both the gfp and the mCherry gene were cloned into RHP-AmpS linearized with primers AH01 and AH16 (Table 1). In addition, the gfp gene was also cloned into RHP-AmpS using type IIs restriction enzyme cloning.

**Test of selection**
Our test genes were cloned into the vector as described previously. After the recovery, the transformation assays were spread onto LB agar plates (1) with various concentra-
To test the different promoters, we choose a variety of ampicillin concentrations ranging 0–1500 μg/mL. Between 8 and 48 colonies were picked for each reaction and the correct insertion of the target gene was first verified by analytical PCR over the whole expression cassette. The sizes of the PCR fragments were then analyzed by gel electrophoresis and reactions that resulted in fragments of the correct length were sequenced.

**Expression, purification and quantification of the target proteins**

To test the expression of the two fluorescent proteins, we used the auto-inducing media described by W. Studier (18). Pre-cultures from one single colony of transformed BL21 (DE3) cells were set up the day before in MDG non-inducing medium (18). They were incubated over night at 37°C and 180 rpm. One-hundred microliters of these pre-cultures were used to inoculate another 5 mL of MDG medium (18), and grown to an OD<sub>600</sub> of ∼3. Five times 50 mL of ZYM-5052 medium (18) were inoculated with these to reach a starting OD<sub>600</sub> for the expression culture of 0.01–0.02. To maintain the selective pressure on the cells, 50 μg/mL kanamycin was added to all cultures. Each of the five cultures contained different concentrations of ampicillin to test the different promoters. The vectors either contained (I) or were missing (E) an insert to reconstitute bla. Combinations suitable for positive selection are highlighted in gray.

**Figure 2. Construction of the ampicillin-positive–selection vector.** (A) RHP-CcdB starting vector containing a chloramphenicol resistance marker Cm(R) and the gene for the CcdB. (B) RHP-GFP, expression control vector derived from (A) with GFP cloned into the multiple cloning site (MCS), replacing the Cm(R)-CcdB cassette. (C) RHP-AmpS (GenBank accession no. FJ545754), suggested vector for positive-selection cloning and high-level protein expression. The vector is derived from RHP-CcdB by replacing the Cm(R)-CcdB cassette with bla<sub>1290</sub> opposing the transcriptional direction of the T7 polymerase to prevent expression of the resistance with the target protein (for details refer to Figure 1). (D) RHP-GFP-AmpR expression construct with an active β-lactamase and GFP as expressed protein. All vectors contain a kanamycin resistance gene (KanR) to maintain the selective pressure during fermentation. Expression in all vectors is under the control of a T7-lac regulatory system followed by the ribosomal binding site (RBS), a rubredoxin and His<sub>6</sub>-tag and PreScission protease site. The transcription is initiated from the start codon following the RBS.

**Table 2. Growth of E. coli Harboring Different Cloning Vectors**

<table>
<thead>
<tr>
<th>Ampicillin concentration [μg/mL]</th>
<th>0</th>
<th>20</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Promoter origin</td>
<td>E</td>
<td>I</td>
<td>E</td>
<td>E</td>
<td>E</td>
<td>E</td>
</tr>
<tr>
<td>pETBlue-1 bla</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>pET17b bla</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

bla was under the control of varying promoters. The vectors either contained (I) or were missing (E) an insert to reconstitute bla. Combinations suitable for positive selection are highlighted in gray.

Author Reprints

Make the most of your hard work by ordering reprints of your article published in BioTechniques. Reprints are an inexpensive and easy way to distribute your findings to students and colleagues.

Corporate Reprints

Leverage BioTechniques, the most powerful brand in the market. Reprints help support your sales effort by utilizing articles that spotlight your brand/products to educate customers at meetings and industry trade events.

More info:

www.BioTechniques.com/reprints
ampicillin (0–500 μg/mL) to see whether the presence of the second antibiotic had any effects on cell growth and expression levels. The OD_{600} was monitored during the whole experiment at an hourly rate.

After the cultures had reached stationary phase, cells were harvested and the His-tagged proteins purified from the pellets using the His GraviTrap Kit (Cat. no. 11–0033–99; GE Healthcare Europe, Otelfingen, Switzerland). The pellets were frozen and thawed twice and then re-suspended in His-GraviTrap binding buffer (20 mM sodium phosphate, 500 mM NaCl, 20 mM imidazole, pH 7.4). Fifty microliters of protease inhibitor cocktail (Cat. no. P8849; Sigma-Aldrich, Buchs SG, Switzerland) and 20 μL Lysonase (Novagen, Merck KGaA) were added. The solution was incubated for 30 min at room temperature and then returned to ice. For further cell disruption, the cells were ultrasonicated with 5 cycles of 10 bursts each (Branson Sonifer 250; Branson Ultrasonics, Geneva, Switzerland). After each sonication cycle the cells were returned to ice for 30 s. To clear the lysates, the suspensions were centrifuged at 39,000×g at 4°C for 20 min. The supernatants were purified on His GraviTrap columns according to the manufacturer’s protocol.

The protein concentration was determined by HPLC on a HP1100 HPLC system (Hewlett-Packard, Palo Alto, CA, USA). A POROS R1 10-μm stainless-steel column (2.1 mm × 100 mm, 0.3 mL; Applied Biosystems) was used. Thirty microliters of sample were run at 80°C, 20 MPa and a flow rate of 0.8 mL/min on a buffer gradient [Solvent A: MilliQ H₂O + 0.1% trifluoro acetate (TFA) (w/v), Solvent B: 90% acetonitrile (w/v) + 0.08% TFA (w/v)]. For the calculation of protein concentration, the HPLC system was calibrated using bovine serum albumin (BSA) as protein standard, thus relating all peak areas to the area measured for a defined BSA concentration.

Results and discussion

We set out to construct a vector that combines a powerful selection system and high-level protein expression. To achieve a high level of protein expression we chose to use our in-house RHP-CcdB vector (Figure 2), a modified pET28 vector (19–22) that has an additional rubredoxin tag (23), as well as the well-known Cm(R)-CcdB cassette containing a chloramphenicol-resistance gene [Cm(R), encoding for chloramphenicol acetyltransferase] and the gene coding for CcdB, a toxic protein that can be used as a selective marker (12,24).

In this vector, T7 polymerase is under the control of the lac repressor, which enables us to use auto-inducing media for protein expression. (For a detailed review of this system see Reference 18.)

We started by inserting a truncated version of bla (TEM-1 β-lactamase) into our original vector, thereby replacing the Cm(R)-CcdB selection cassette. The truncated bla_{Δ290} was amplified from pETBlue-1 (Novagen, Merck KGaA). The orientation of bla_{Δ290} in the new selection vector opposed the transcription direction of the T7 polymerase, thus preventing unintentional transcription of the resistance gene by the T7 polymerase.

To test whether our positive-selection vector was suitable for the detection of positive recombinants, we cloned the gene for the green fluorescent protein (GFP) into the vector. No colonies grew at 100 μg/mL ampicillin in cloning reactions using the vector that contained the bla promoter from pETBlue-1, and we therefore reasoned that this ampicillin concentration might be too high to select for our construct. In response, we repeated the cloning reactions and selected on LB agar plates containing different concentrations of the antibiotic. In this experiment, the highest ampicillin concentration at which colonies grew was 25 μg/mL. These colonies were sequenced and we found that all of them contained the correct insert in the correct orientation. However, we were not confident that positive recombinants could be selected at this concentration with a high enough efficiency. Since pETBlue-1 is a high–copy-number plasmid (pUC origin), we reasoned that the promoter used for bla might be too weak for our expression plasmid, which is a low– to medium–copy-number vector. Therefore, to find an optimal window to select for positive recombinants, we replaced the promoter region from pETBlue-1 with either the 104-bp–long promoter of chloramphenicol acetyltransferase or a 76-bp fragment upstream of the bla gene from pET17b, respectively. We then cloned the gfp gene into the two new vectors with the different promoters. Although we observed a higher number of colonies, both vectors gave rise to a considerable number of false-positive colonies. This provided evidence that the truncated β-lactamase still had a high-enough residual activity to overcome the standard selective 100 μg/mL ampicillin concentration. Consequently, we determined the minimal inhibitory concentration (MIC) of ampicillin in the different positive selection vectors containing the bla promoter from pETBlue-1, the bla promoter from pET17b, or the promoter of the Cm(R) cassette. For this, we determined the highest concentration of ampicillin on which the respective vectors with (complete bla) or without (truncated bla) the gfp gene would grow. The vectors containing the pETBlue-1 promoter did not have any residual resistance to the antibiotic, whereas the plasmids harboring the chloramphenicol acetyltransferase promoter were able to confer resistance up to 400 μg/mL ampicillin with the empty vector and the truncated bla. The reconstructed bla for this construct had an MIC that was higher than the tested range. For the empty pET17b bla promoter construct, colonies were able to grow up to 200 μg/mL ampicillin. Here, the reconstructed bla construct also grew to concentrations higher than tested (see overview in Table 2).

We therefore determined that the construct containing the bla promoter from pET17b was giving us a suitable window for the selection of positive recombinants. We reasoned that a concentration of 300 μg/mL was high enough to ensure that no false positive colonies would be able to grow. Using these selective conditions we performed a number of cloning reactions using either gfp (25,26) or mCherry (27) as target genes. Both analytical PCR and sequencing of the PCR products showed that vectors contained the target genes in their correct orientation, reconstructing W290 of the bla gene. In our current experience (more than 20 cloning experiments, 4–16 colonies per experiment sequenced), we have not seen any false-positive colonies. However, the appearance of false-positive colonies cannot be ruled out completely; cloning of aberrant PCR products that reconstruct W290 of the β-lactamase are theoretically possible, as are plasmid rearrangements. However, these events are occurring at a very low frequency and should therefore not interfere with the intended applications.

Since we set out to construct a positive-selection vector that is also suitable for direct and high-level protein expression, we assessed the vector expression capacities by expressing the two fluorescent proteins mCherry and GFP. Because the ampicillin-positive–selection cassette was located inside the protein expression cassette, we also had to investigate whether it would have any negative influence on the expression levels. As a reference, we also cloned the two genes into the original RHP-CcdB vector, which is missing the ampicillin selection cassette but is otherwise identical (Figure 2). We did not observe any negative effects of the ampicillin selection cassette on cell growth compared with the control vector.
RHP-CcdB (data not shown). The presence of ampicillin during expression did not affect cell growth or expression levels.

The mCherry construct was expressed at similar levels as the reference (~250 μg/mL). However, the GFP construct was expressed significantly higher in the positive selection vector than in the control (50 μg/mL and 6 μg/mL, respectively). This could be due to the longer mRNA sequence transcribed in these constructs. As the selection cassette was inserted in opposite transcription direction between the T7 promoter and T7 terminator, the reverse complementary sequence for bla was also transcribed by the T7 polymerase. This additional mRNA could protect the actual gene of interest from degradation by 3′-RNases.

The mCherry positive selection construct was grown also in a 1-L fermenter in an in-house modified, rich auto-induction medium. We were able to purify an amount of 1.4 g of the rubredoxin-tagged mCherry protein.

In this study, we have used homologous recombination (17) and type II restriction enzymes (16) for cloning inserts into the positive selection vector. Unfortunately the C-terminal end of the truncated bla does not allow using classic restriction enzyme cloning for the target gene insertion. However, other cloning methods such as ligation-independent cloning (28) are well suited for this.

We set out to construct a vector for positive-selection and direct, high-level protein expression. We have constructed a vector, RHP-AmpS, that provides a simple way to achieve highly efficient selection of positive recombinants (Figures 1 and 2). Moreover, this vector was shown to be suitable for the expression of high levels of protein. To our knowledge, there is currently no vector available combining these features as efficiently as this one does.

Acknowledgments

We thank Roger M. Benoit, Anke Blechschmidt, Daniela Scherer-Becker, and Ramon N. Wilhelm for their helpful advice and inspiring discussions. We also thank Armand Wanner and Pascale Rieder for carrying out all the DNA sequencing reactions. We are also grateful to David J. Allocco for reviewing this manuscript.

The authors declare no competing interests.

References


Received 18 September 2008; accepted 5 January 2009.

Address correspondence to Christian Ostermeier, Novartis Pharma AG, Forum 1, Novartis Campus, CH-4056 Basel, Switzerland. email: christian.ostermeier@novartis.com