A Dual Protein Expression System in *Bacillus subtilis*

Running title: Dual expression system

Annie Y. Chan,* Mei M. Chan,* Hei M. Lo,* Yun C. Leung,† and Boon L. Lim*,†

*Department of Zoology, the University of Hong Kong, Pokfulam Road, Hong Kong, China.
†Department of Applied Biology and Chemical Technology, the Polytechnic University of Hong Kong, Hunghom, Hong Kong, China.

†To whom correspondence should be addressed: Department of Zoology, the University of Hong Kong, Pokfulam Road, Hong Kong, China.

Tel: (852) 22990826
Fax: (852) 25599114
E-mail: bllim@hkuec.hku.hk

The number of figures: 5

The number of tables: 1

The number of words in the main text: 2056 words.

The GenBank accession numbers of the sequences that are reported in this paper are Z34519 and Z46268.

Keywords: Dual expression, *Bacillus subtilis*, xylanase, glucanase, prophage
ABSTRACT

We have developed a dual expression system for the simultaneous overexpression of two proteins in *Bacillus subtilis*. Two candidate genes, xylanase (*xynA*) and glucanase (*bglS*) from *B. subtilis* strain 168, which were engineered with independent Shine-Dalgarno (SD) sequences, were cloned in tandem into a transfer vector, which was then transformed into *B. subtilis* strain 1A304 (φ105MU331). The genes were under the transcriptional control of a strong promoter of a bacteriophage, φ105, where transcription was initiated upon thermal induction. Six constructs were made to compare the factors that affected the yields of the gene products. The expression level of each candidate gene was found to correspond to its position relative to the phage promoter, irrespective of the identity of the insert. The lower expression level of the second insert might have been due to limited resources for protein synthesis, a short half-life of the mRNA, or an early termination of the RNA polymerase. Curiously, gene duplications in tandem did not lead to further production.
INTRODUCTION

Many industrial enzymes are produced by recombinant technology, and usually more than one enzyme is involved in an industrial process. For example, mixtures of protease, lipase, and amylase are sold for detergent applications. We developed a dual expression system based on the Bacillus subtilis φ105 MU331 prophage system (1-3). In our system, two proteins were produced simultaneously after heat induction, and the factors that affected their yields were studied. Briefly, two genes were cloned in tandem at a position downstream from the bacteriophage φ105 promoter, with their expression controlled by the φ105 repressor protein. The cts-52 mutation in the 1A304(φ105MU331) strain renders the φ105 repressor temperature sensitive so that protein expression is induced by temperature shift. Upon thermal induction, the φ105 repressor protein falls off, the promoter is de-repressed, and transcription commences.

To examine the factors that affect the yields of the two proteins, the xynA (4) and bglS (5) genes from B. subtilis 168 were chosen as test models. As these genes are endogenous to the host strain B. subtilis 1A304 (φ105MU331), they are better candidates as test models than the other heterologous genes. The expression of the other heterologous genes could be hindered by many factors, such as codon biases, inefficient signal peptide processing, incorrect protein folding, and susceptibility to the host’s cellular proteases, etc. The gene product of xynA, endo-1,4- β-D-xylanohydrolase (EC
3.2.1.8), hydrolyses xylan, is the major constituent of hemicellulose (6). It has many applications, especially in the paper and pulp industry (6). The gene product of \textit{bglS}, \(\beta\)-(1,3)-(1,4)-glucanase (EC 3.2.1.73), catalyses hydrolysis of \(\beta\)-1,4 bonds in mixed \(\beta\)-glucan. It is particularly useful when it acts in combination with the other hydrolytic enzymes. For instance, both xylanase and glucanase are used as additives to increase the digestibility of animal feed.
MATERIALS AND METHODS

Bacterial strains and culture media

*B. subtilis* strain 168 served as the template for the glucanase and xylanase genes. 

*E. coli* Top 10 (Invitrogen, Netherlands) was used as a host strain for plasmid transformation. Luria-Bertani (LB) agar was used for the selection of both *B. subtilis* and *E. coli* strains, and was supplemented with ampicillin (100 mg/L) or chloramphenicol (25 mg/L) when necessary. BHY medium (bovine brain heart infusion, 37 g/L and yeast extract 5 g/L) was used for the *B. subtilis* cultivation.

Construction of expression vectors

The terminator of the α-amylase gene from *B. licheniformis* was cloned into the plasmid pSG1112, which is a derivative of pSG703 (Thornewell et al., 1993), to create the backbone plasmid pSGt (Fig. 1). pSGt contains a promoter and a Shine-Dalgarno (SD) sequence (a ribosome binding site) of the prophage φ105 MU331, followed by a multiple cloning site and a cat gene. The promoter region and the cat gene allow the insertion of heterologous genes into the lysogenic *B. subtilis* 1A304(φ105MU331) chromosome via homologous recombination, as described by Thornwell et al. (7).

To create vectors with dual protein expression capacity, the chromosome DNA
from *B. subtilis* strain 168 was used as a template for glucanase (*bglS*, GeneBank Accession no. Z46268) and xylanase (*xynA*, GeneBank Accession no. Z34519) gene amplification. The primers are listed in Table 1. Restriction sites (VspI, BamHI, and XbaI) were included in the primers to facilitate target gene insertion into cloning vector pSGt. Six plasmid vectors were created: pSGt-X, pSGt-XX, pSGt-XG, pSGt-GX, pSGt-GG and pSGt-G (Fig. 2). All of the constructs contained an upstream $\phi 105$ promoter and a downstream terminator of the $\alpha$-amylase gene from *B. licheniformis*. The plasmids pSGt-X and pSGt-G carried a single gene only, and the other four constructs contained two genes in tandem. Both genes had an upstream SD sequence (GGAGGG) that was located 7 base pairs upstream from the initiation codon ATG.

**Generation of recombinant Bacillus subtilis strains**

The construction of recombinant plasmids was manipulated in *E. coli* strain Top 10. To generate recombinant *Bacillus subtilis* strains, plasmids were extracted from *E. coli* by a kit (Quantum Prep, BIO-RAD, Hong Kong) before they were transformed into *B. subtilis*. The preparation of *B. subtilis* 1A304($\phi$105MU331) competent cells and plasmid transformation was accomplished according to the procedure that was described by Osbume et al. (8). Positive transformants were selected on agar plates that were supplemented with chloramphenical (25 ug/ml). The colonies were picked for
further PCR screening using a φ105 promoter-specific primer (sense) and an insert-specific primer (antisense) (Table 1).

*Over-expression of target protein in B. subtilis 1A304(φ105MU331)*

For the production of recombinant proteins, recombinant *B. subtilis* strains were grown in 10 ml of BHY medium with chloramphenicol in a 250 ml flask and shaken at 37 °C at 280 rpm overnight. 1 ml of the overnight culture was then transferred to 15 ml of BHY medium without antibiotics, and the culture was shaken at 37 °C at 280 rpm. After 2 hours, when OD600 reached 3.5 to 4.3, recombinant protein expression was induced by heat shock at 50 °C for 5 minutes. As the enzymes were secreted into the culture medium, 1 ml of cell-free supernatant was collected at various time points from each clone for SDS-PAGE analysis and enzyme assaying. For the OD600 measurement, the cells were diluted 10-fold with culture medium before reading OD.

*Enzyme assays*

Xylanase and glucanase activities were assayed according to Dygert’s Method (9). 0.1% (w/v) oat spelt xylan and 0.5% (w/v) barley-β-glucan were used as substrates for the xylanase and glucanase assays. Culture supernatants were diluted in 50 mM of phosphate buffer (xylanase assay), pH 6.5, or 50 mM of acetate buffer (glucanase
assay), pH 6.0. Both assays were carried out at 50 °C. The amount of reducing sugar that was released was assayed by using xylose and glucose as standards. One unit of activity was defined as the amount of enzyme that produced 1 umol of reducing sugar per minute.

Total RNA extraction and one-step RT-PCR

To confirm the presence of a fusion-transcript, a one-step RT-PCR (QIAgen OnseStep RT-PCR Kit) was performed. Clones pSGt and GX were cultured as described above. 30 minutes after induction, total RNA was collected using an RNease® Mini Kit (QIAgen). Contaminated DNA was digested and removed by treating the total RNA with components from the DNA-free kit (Ambion). The integrity of the RNA sample was monitored by gel electrophoresis and OD_{260/280} measurement. 50 ng of each RNA sample was added as a template in an individual PCR reaction, for which t1112 was employed as a control sample. A reaction mix without template RNA was included as a negative control. Three sets of one-step RT-PCR with different primer combinations were prepared. In one of these reactions, gene-specific primer C and primer F (Table 1) were added for amplification of the fusion-transcripts {bglS}xyn{A}. The second reaction was aimed at amplifying {bglS} using another pair of gene-specific primers C and D (Table 1). Xyn{A} was detected using gene-specific primers E and F.
(Table 1). Each primer was added at a final concentration of 0.6 μM. The RNA samples were reverse-transcribed at 50 °C for 30 minutes, followed by initial DNA polymerase activation by incubation at 95 °C for 15 minutes before amplification. This incubation also inactivated the reverse transcriptase. An amplification cycle was carried out by denaturation of cDNA at 94 °C for 40 seconds, annealing at 50°C for 40 seconds, then polymerization at 72 °C for 2 minutes. A final polymerization step was performed at 72 °C for 10 minutes.
RESULTS

Generation of the recombinant Bacillus strains

The PCR products of the xylanase and glucanase genes were 642 and 729 bp respectively. The identities of the six constructs that are shown in Fig. 2 were confirmed by HincII digestion, because different constructs produce different DNA fragments. The genotype of the host B. subtilis strain 1A304 (φ105MU331) is ind-125cts-52Ω (lacZ'[ClaI]-ermC-cat [NeoI])331Δ(DI:1t)trpC2metB52xon-1 SPB(S), in which an engineered temperate bacteriophage φ105MU331 has been incorporated into the bacterial genome (7). The pSGt constructs and the B. subtilis 1A304(φ105MU331) chromosome DNA contain homologous regions. The target inserts were flanked by a φ105 fragment (approximately 600 bp) and a complete cat gene in the plasmid (Fig. 1), whereas an ermC gene was flanked by the same φ105 fragment and a truncated cat' gene in the B. subtilis 1A304(φ105MU331) chromosomal DNA (7). After the plasmid was taken up by the competent B. subtilis 1A304(φ105MU331) strain in its late exponential phase, homologous double crossover recombination occurred. Therefore, the transformants should have obtained the phenotype CmRErS. Among the colonies that were picked for PCR screening, 70% of those with the CmRErS phenotype showed positive transformation. The transformants were designated as X, XX, XG, GX, GG,
and G, according to their insert identity. As the plasmid contains no replication origin for the binding of *B. subtilis* DNA polymerase, it cannot autonomously replicate unless stably integrated into the bacterial DNA.

*Overproduction of glucanase and xylanase*

After thermal induction during the late log phase, all six clones were found to have overexpressed the corresponding gene products, and the target proteins gradually increased from 1 hour to 4 hours after induction. Culture supernatant (15 ul) that was collected at 4 hours after induction was analyzed by SDS-PAGE (Fig. 3). The protein bands of 22 kDa and 25 kDa represent xylanase and glucanase.

*Enzyme activity assays*

Enzyme assays were carried out using cell free supernatant collected at 4 hours after induction. The results of the activity assays are shown in Fig. 4. It is obvious that the clones GX and XG can produce both glucanase and xylanase, but at different levels. The clone GX produced more glucanase (6.5 U/ml) than xylanase (3.0 U/ml), whereas the clone XG produced more xylanase (9.0 U/ml) than glucanase (4.4 U/ml). In contrast, the double clones XX and GG, which carried two identical genes, only secreted a similar level of enzymes when compared with the clones (X and G) that carried a single
copy gene. This might have been due to limited resources that were available for enzyme production within the host. The elevated expression of these enzymes was not derived from the endogenous xylA and bgIS genes, as the host strain B. subtilis 1A304(φ105MU331) produced negligible enzyme activities (data not shown).

Total RNA extraction and one-step RT-PCR

To confirm that the two genes within the same vector were transcribed from the same promoter, a RT-PCR was performed on the clone that carried the pSGt-GX construct. mRNA was first isolated at 30 minutes after heat induction, and various sets of PCR primers were used. As shown in lane 3 of Fig. 5, a PCR product of approximately 1300 bp was amplified by the glucanase sense primer (C) and the xylanase antisense primer (F). The PCR results also indicated that no endogenous glucanase or xylanase mRNAs were transcripted in the host strain that was transformed with the pSGt vector.
DISCUSSION

We have developed a dual expression system in *B. subtilis* that can simultaneously co-express two proteins. However, the expression levels of these two proteins vary. As shown in Fig. 4, the clone GX secreted more glucanase than xylanase, whereas the clone XG secreted more xylanase. This indicates that the expression level of the genes was related to their relative positions in the construct. The closer that the gene was to the prophage promoter, the higher was the expression level. This is unrelated to the identity of the gene or its signal peptide sequence.

After thermal induction, the \(\phi\)105 prophage DNA is excised from the *B. subtilis* host chromosome and starts DNA replication for gene expression. Thereafter, a DNAase targeting host DNA is produced. As the DNase accumulates and destroys the host DNA, the host is unable to supply sufficient amino acids and energy for recombinant protein synthesis. As the 5’ ends of mRNA molecules are first transcribed, the gene in close proximity to the \(\phi\)105 promoter could be translated before the mRNA synthesis is completed, which will result in a higher expression level. When the second gene that is located at the 3’ end of the mRNA molecule is available for translation, the preceding gene has already depleted a certain amount of limited resources. Thereafter, the gene that is located at the 3’ end of the mRNA shares the depleting resources with the gene that is located at the 5’ end. In addition, RNA polymerase has a certain
probability of terminating before the completion of the message, thereby favoring the
expression of promoter-proximal genes as compared to those further downstream from
the promoter.

The dual expression system may be applied to enhance the level of protein
secretion in *Bacillus subtilis*. The high level expression of secretory proteins can
saturate the secretion apparatus in *B. subtilis*. For example, in *B. subtilis* strains that
overexpress α-amylase, amylase precursors with uncleaved signal peptide were found
to accumulate either in the cytoplasm or associate with the cytoplasmic membrane (10,
11). The secretory protein usually carries a signal peptide that is recognized and
processed by a signal peptidase on the cell membrane. After the nascent polypeptide is
translocated across the cytoplasmic membrane, post-translational folding of the
polypeptide is required for the generation of a functional protein. The correct folding of
many secreted proteins is not spontaneous, but is dependent on assisting factors such as
chaperones or foldases. (12-14). To date, only one extra-cytoplasmic protein in the
gram-positive bacterium *B. subtilis*, Prs A, is known to be involved in protein secretion.
Prs A is a 33 kDa protein that is anchored on the outer surface of the cytoplasmic
membrane. The overexpression of Prs A was shown to enhance the secretion of
α-amylase in *B. subtilis* (15) by approximately 5-fold. Therefore, the co-expression of a
target protein with PrsA using the dual expression system that is described in this article
may enhance the secretion of the target protein.

\textit{B. subtilis} has an excellent secretory ability, which is partly due to its abundance in signal peptides. There are six signal peptide genes in the chromosome DNA of \textit{B. subtilis}, including five type I signal peptidases (SPase I) and one type II signal peptidase (SPase II). These two types of signal peptidases are believed to be responsible for the cleavage of the signal peptides of secretory proteins and lipoproteins respectively (16, 17). Even though the five type I signal peptidases, SipS, SipT, SipU, SipV, and Sip W, have various degrees of sequence homology, it is believed that they have different substrate specificities. In addition, their transcription levels are controlled by different mechanisms. While SipU and SipV are constitutively transcribed at a low level, the transcriptions of SipS and SipT are more active in the post-exponential growth phase and exponential growth phase respectively (18). In the future, we intend to study the effects of incorporating these signal peptide genes into the dual expression system to test whether any of them can enhance the secretion of the protein of interest.
Acknowledgement

We are grateful to Professor J. Errington for providing the plasmid pSG1112.
FIGURE LEGENDS

Fig. 1. The plasmid map of pSGt. Heterologous genes were inserted into the multicloning sites (MCS) of the vector. The vector carried a pBR 322 replication origin for *E. coli*, the promoter and the ribosome binding site (SD) of prophage Φ-105 MU331 (Φ-105), the transcription terminator of *B. licheniformis* α-amylase (term), a cat gene for chloramphenical resistance, and a bla gene for ampicillin resistance.

FIG. 2. Construction map of the six constructs. The thick solid line represents the sequence that was derived from the vector pSGt. The gray box represents the Φ-105 promoter; the black circle ‘SD’ represents Shine Dalgarno sequence GGAGGG; and the checkered box represents the terminator of the *B. licheniformis* α-amylase gene, ACGGATTCCCTGAAGGAAATCCGT.

FIG. 3. SDS-PAGE of the culture supernatant that was collected 4 hours after induction.

The left-most lane contains the molecular weight markers. Lanes 1 to 6 were collected from clones X, XX, XG, GX, GG, and G. Lane 7 is the culture supernatant from the cells that were transformed with the vector only. 15 ul samples were loaded into each lane.

FIG. 4. Comparison of the xylanase and glucanase activities that were secreted by
various constructs.

FIG. 5. RT-PCR products from clones pSGt and GX.

Lanes 1 to 3, 4 to 6, and 7 to 9 were amplified by primers C/F, C/D, and E/F respectively.

Lanes 3, 6, and 9 demonstrated the presence of transcripts _bglSxynA_, _BglS_, and _xynA_ in the clone GX. No PCR products were shown in lanes 1, 4, and 7 in the absence of template RNA, or in lanes 2, 5, and 8 where mRNA from clone pSGt was used as a template. Molecular weight markers M₁ (1 kb Molecular weight marker, Gibco BRL) and M₂ (GeneRuler 100 bp DNA ladder plus, MBI) were employed.
Table 1: Sequences of oligonucleotides

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’→ 3’)</th>
<th>R. E. site</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>GAAACATTAATTGTGTTAAGTTTTAAAAAGAAATTCTTAG</td>
<td>Vsp I</td>
</tr>
<tr>
<td>B</td>
<td>TAGGATCCGCTACCCCTGAT TAAGGATG</td>
<td>BamH I</td>
</tr>
<tr>
<td>C</td>
<td>AATCGATTATGCCTTATCTGAAACGAGT</td>
<td>Vsp I</td>
</tr>
<tr>
<td>D</td>
<td>TAGGATCCCAGCAGGGTTCTTTTACAC</td>
<td>BamH I</td>
</tr>
<tr>
<td>E</td>
<td>ATGGATCCGGGGGTAACATATGTGTTAAGTTTTAAAAAG</td>
<td>BamH I</td>
</tr>
<tr>
<td>F</td>
<td>ACTCTAGATTACCACACTGTTACGTTAGA</td>
<td>Xba I</td>
</tr>
<tr>
<td>G</td>
<td>ATGGATCCGGGGGCCCATAATGCTTTATCTGAAAC</td>
<td>BamH I</td>
</tr>
<tr>
<td>H</td>
<td>AGTCTAGAAGCAGCAGGTTCTTTTCACA</td>
<td>Xba I</td>
</tr>
<tr>
<td>Phi105S</td>
<td>CACCGATAAT AAGGTGGGAG</td>
<td>Nil</td>
</tr>
</tbody>
</table>

Restriction sites are underlined. Initiation codon ATG is shown in bold type.
SD sequences (GGAGGG) are shown in italics.
REFERENCES


Fig. 1
Gene B

Gene A

NdeI/VspI (ATG)

Bam HI

Xba I

Sal I

φ 105 promoter

SD sequence

 terminator of the α-amylase gene from *Bacillus licheniformis*

<table>
<thead>
<tr>
<th>Vector</th>
<th>Gene A</th>
<th>Primers</th>
<th>Gene B</th>
<th>Primers</th>
</tr>
</thead>
<tbody>
<tr>
<td>pSGt-X</td>
<td>Xylanase</td>
<td>A/B</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td>pSGt-XX</td>
<td>Xylanase</td>
<td>A/B</td>
<td>Xylanase</td>
<td>E/F</td>
</tr>
<tr>
<td>pSGt-XG</td>
<td>Xylanase</td>
<td>A/B</td>
<td>Glucanase</td>
<td>G/H</td>
</tr>
<tr>
<td>pSGt-GX</td>
<td>Glucanase</td>
<td>C/D</td>
<td>Xylanase</td>
<td>E/F</td>
</tr>
<tr>
<td>pSGt-GG</td>
<td>Glucanase</td>
<td>C/D</td>
<td>Glucanase</td>
<td>G/H</td>
</tr>
<tr>
<td>pSGt-G</td>
<td>Glucanase</td>
<td>C/D</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
</tbody>
</table>

N.A. Not Applicable.
Fig. 3
<table>
<thead>
<tr>
<th>Clone</th>
<th>Xylanase Activity (E.U/ml)</th>
<th>Fold*</th>
<th>Glucanase activity (E.U/ml)</th>
<th>Fold*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>12.1</td>
<td>86.1X</td>
<td>0.06</td>
<td>0.8X</td>
</tr>
<tr>
<td>XX</td>
<td>11.7</td>
<td>83.6X</td>
<td>0.08</td>
<td>1.0X</td>
</tr>
<tr>
<td>XG</td>
<td>6.5</td>
<td>46.1X</td>
<td>3.0</td>
<td>37.3X</td>
</tr>
<tr>
<td>GX</td>
<td>4.4</td>
<td>31.1X</td>
<td>9.0</td>
<td>113.1X</td>
</tr>
<tr>
<td>GG</td>
<td>0.08</td>
<td>0.6X</td>
<td>12.4</td>
<td>155.1X</td>
</tr>
<tr>
<td>G</td>
<td>0.12</td>
<td>0.9X</td>
<td>11.9</td>
<td>148.4X</td>
</tr>
<tr>
<td>pSGt</td>
<td>0.14</td>
<td>1.0X</td>
<td>0.08</td>
<td>1.0X</td>
</tr>
</tbody>
</table>

E.U.: Enzyme Unit, defined as the amount of enzyme producing 1umol reducing sugar per minute.
*Number of fold refers to the enzyme activities compared with that of the vector clone (pSGt).
Figure 5.