

# Construction of a Reporter Gene Vector for *Clostridium beijerinckii* using a *Clostridium* Endoglucanase Gene

Keith W. M. Quixley and Sharon J. Reid\*

Department of Microbiology, University of Cape Town, Rondebosch, 7700, Capetown, South Africa

## Abstract

A  $\beta$ -1,4-endoglucanase gene (*eglA*) cloned from *C. acetobutylicum* P262 was selected for use in the development of a reporter system for *C. beijerinckii* NCIMB 8052. The reporter plasmid, pER1, was constructed by ligating the promoterless *eglA* gene into the *B subtilis*/*Clostridium* shuttle vector, pFNK1, which can replicate and is stably maintained in *C. beijerinckii*. The expression of the endoglucanase enzyme from its own promoter was not significantly induced in cells grown in glucose, sucrose or galactose, while growth of cells in cellobiose or fructose resulted in lower levels of activity. The enzyme was efficiently secreted into the culture medium and did not remain associated with the cell in any way. A transcriptional fusion between the glutamine synthetase (*glnA*) promoter region and the promoterless *eglA* gene resulted in high levels of endoglucanase expression, which reflected an 11-fold increase in expression levels over the *eglA* promoter.

## Introduction

Metabolic engineering of bacteria with industrial application requires a selection of sophisticated genetic techniques and tools for the improvement of strains. A variety of genes have been characterised from the *Clostridium acetobutylicum*/*beijerinckii* group of bacteria, which produce acetone and butanol by anaerobic fermentation. However, knowledge of the general structure of *Clostridium* promoters and other regulatory features of gene expression is still poor. Studies on gene regulation in *C. beijerinckii* are hampered by the lack of an efficient reporter gene vector, which is maintained stably in this host. Both the *xylE* gene, encoding the xylanase enzyme from *Pseudomonas*, and the *cat* gene from the *Staphylococcus* plasmid pC194 have been used to assess the activity of the *Clostridium pasteurianum* ferredoxin promoter in the development of an expression vector (Minton *et al.*, 1993). While expression of *xylE* was relatively inefficient, high levels of *cat* expression were observed. However, chloramphenicol is rapidly inactivated by reduction in *C. beijerinckii* cells (O'Brien and Morris, 1971).

Other expression vectors have been developed for use in different *Clostridium* species. In *C. perfringens*, two

different genes have been used as reporters. The first utilised the chloramphenicol acetylase gene (*catP*) cloned from the *C. perfringens* plasmid pIP401 (Steffen and Matzura, 1989). The second reporter system uses the genes encoding the luciferase enzyme of *Vibrio fischeri*, which has been widely used as a reporter system in aerobic and facultative organisms (Stewart and Williams, 1992). The luciferase enzyme catalyses light production from FMNH<sub>2</sub> and long-chain aliphatic aldehydes, and requires oxygen for the reaction. This system has therefore not been used in obligately anaerobic bacteria, but expression of luciferase has been achieved in the relatively aerotolerant *C. perfringens* (Phillips-Jones, 1993). Production of light in *C. perfringens*, in which the expression of the *luxAB* was under the control of the *C. perfringens* alpha-toxin promoter, was shown to be reflective of luciferase expression. The light production was found to decay rapidly however, and it is possible that the use of luciferase as a reporter could be problematic in strains that are less aerotolerant such as *C. beijerinckii*.

A  $\beta$ -1,4-endoglucanase gene (*eglA*) was isolated from *C. acetobutylicum* P262 (Zappe *et al.*, 1988), although this strain had been reported to produce only weak endoglucanase activity which was induced by a small unidentified molecule present in molasses medium (Allcock and Woods, 1981). A promoterless version of this gene was used successfully as a reporter gene for investigating the regulation of glutamine synthetase (*glnA*) transcription in *Bacteroides fragilis* (Abratt *et al.*, 1993). Analysis of the sequence of the *eglA* gene had identified a putative signal sequence for secretion into the growth medium (Zappe *et al.*, 1988). In confirmation of this observation, enzyme activity in both *E. coli* and *B. fragilis* was found to be located predominantly in the periplasm (Zappe *et al.*, 1988; Abratt *et al.*, 1993).

We have developed a reporter gene vector using a promoterless  $\beta$ -1,4-endoglucanase gene (*eglA*) to study the regulation of different genes and operons in *C. beijerinckii*. This vector has been used to identify the promoter region of the *C. beijerinckii* glutamine synthetase (*glnA*) gene, and to demonstrate that the *glnA* promoter is a strong promoter which causes an 11-fold increase in expression of the reporter gene when compared to its own promoter.

## Results and Discussion

### Construction of Reporter Plasmid

The reporter plasmid, pER1, was based on the *Bacillus*/*Clostridium* shuttle vector pFNK1 (Mermelstein *et al.*, 1992; Table 1), which has the erythromycin (Em) resistance marker for selection. The promoterless *C. acetobutylicum* P262 *eglA* gene from pHZ117 $\Delta$ 12 (Abratt *et al.*, 1993), was ligated into the *PvuII* site of pFNK1, resulting in the 3607 bp plasmid, pER1 (Figure 1). In pER1, the *PstI*, *BamHI* and *SmaI* restriction sites are unique to the cloning site, and can therefore be used for generating promoter fusions.

Received October 18, 1999; revised November 7, 1999; accepted November 12, 1999. \*For correspondence. Email shez@molbiol.uct.ac.za; Tel. 27 21 650 3257; Fax. 27 21 689-7573.

Table 1. Strains and Plasmids

Name	Relevant Characteristics	Reference
<b>Bacterial strains</b>		
<i>B. subtilis</i> IA46	<i>recA-4 thr-5 trpC-2</i>	Bacillus Genetic Stock Centre
<i>C. acetobutylicum</i> NCP262	wild-type	Jones <i>et al.</i> (1982)
<i>C. beijerinckii</i> NCIMB 8052	wild-type	NCIMB type strain
<i>E. coli</i> JM109	$\Delta(lac-proAB) lac^F \Delta(lacZ)M15 recA1$	Yanisch-Perron <i>et al.</i> , (1985)
<b>Plasmids</b>		
pHZ117	Ap <sup>R</sup> EG <sup>+</sup> ; <i>eglA</i> in pUC19	Abratt <i>et al.</i> (1993)
pHZ117 $\Delta$ 12	Ap <sup>R</sup> , promoterless <i>eglA</i> ( <i>eglA<math>\Delta</math>P) in pUC19</i>	Abratt <i>et al.</i> (1993)
pBluescriptSK <sup>+</sup>	Ap <sup>R</sup> , T <sub>3</sub> and T <sub>7</sub> polymerase promoters	Stratagene, Ja Holla, California
pFNK1	Em <sup>R</sup>	Mermelstein <i>et al.</i> (1992)
pER1	<i>eglA<math>\Delta</math>P from pHZ117<math>\Delta</math>12 cloned in pFNK1</i>	This study
pEC1	<i>eglA</i> from pHZ117 cloned in pFNK1	This study
pESK	<i>eglA<math>\Delta</math>P cloned in pBluescriptSK<sup>+</sup></i>	This study
pCBN1	3.8 kb chromosomal insert containing the <i>glnA-glnR</i> genes from <i>C. beijerinckii</i> in pBluescriptSK <sup>+</sup>	Quixley, MSc thesis, University of Cape Town (1999)
pEGP1	<i>C. beijerinckii</i> NCIMB 8052 <i>glnA</i> promoter/ <i>eglA<math>\Delta</math>P fusion cloned in pFNK1</i>	This study

Although there are two *EcoRI* sites in the multiple cloning site, they are located close together and could be used as an alternative cloning site as deletion of the intervening sequence would have no effect on the construct. The reporter gene possesses its own ribosomal binding site and can therefore be used as a promoter probe. In order to test the functionality of the reporter plasmid, a 726bp *StyI* DNA fragment containing the *C. beijerinckii glnA* promoter was cloned upstream of the reporter gene, resulting in pEGP1.

An oligonucleotide primer, complementary to nucleotides 115 to 134 of the *eglA* coding region, was used to sequence from within the gene into the promoter region to which it had been fused. The sequence of the junction regions of the fusion between the *glnA* promoter and the *eglA* coding region present on plasmid pEGP1 is shown in Figure 2. The 726 bp *StyI* fragment carrying the promoter region of the *C. beijerinckii glnA* was cloned into the *EcoRI* site upstream of the reporter gene. The fusion carries 480

bp of sequence upstream of the translational start of *glnA*, including all four transcriptional start sites (t<sub>1</sub>-t<sub>4</sub>) identified by primer extension experiments (data not shown). This construct does contain the ribosomal binding site and 240 bp of the *glnA* coding region, although an amber (UAA) stop codon has been introduced at the 5' end of the truncated *glnA* gene. As there appeared to be no structures present which would be interpreted by the RNA polymerase as a termination signal, the presence of the truncated *glnA* should not affect transcription from the *glnA* promoter.

#### Plate Assay for Endoglucanase Activity

Plasmids were introduced into *C. beijerinckii* NCIMB 8052 by electroporation and were selected on CMC plates containing Em (10 µg/ml). After overnight incubation, the plates were stained with Congo Red and areas in which the CMC had been digested showed clear zones indicating endoglucanase activity. *C. beijerinckii* transformants carrying the control plasmid pFNK1 showed no zones of clearing around the colonies, while transformants carrying pEC1 showed substantial CMC digestion after overnight incubation. *C. beijerinckii*(pER1), the promoterless construct, produced slight zones of CMC degradation but these were markedly smaller than those produced by *C. beijerinckii*(pEC1). This indicates that although the promoter has been eliminated there is still measurable expression of *eglA*, possibly due to transcription from an unidentified minor promoter downstream of the major *eglA* promoter. It is unlikely that the activity observed is due to expression of an endogenous gene with endoglucanase activity as *C. beijerinckii*(pFNK1) showed no activity on plate assays. A similar observation was made in *B. fragilis*, where a background level of endoglucanase activity was consistently measured in the presence of the promoterless gene (Abratt *et al.*, 1993).

#### Localisation of Enzyme Activity

To determine the location of the product of the cloned *eglA* gene, endoglucanase activity was tested in different cellular fractions, using the p-nitrophenyl cellobioside assay of Deshpande *et al.* (1984). Mid- and late-exponential cultures of *C. beijerinckii*(pFNK1), *C. beijerinckii*(pER1) and *C. beijerinckii*(pEC1) were grown in Clostridial Basal



Figure 1. Map of the reporter vector, pER1, showing the relative positions of the promoterless endoglucanase (*eglA $\Delta$ P) and Em<sup>R</sup> genes, and the unique restriction sites which could be used for cloning promoter fragments. Approximate positions of restriction sites in the coding region of *eglA* are shown.*

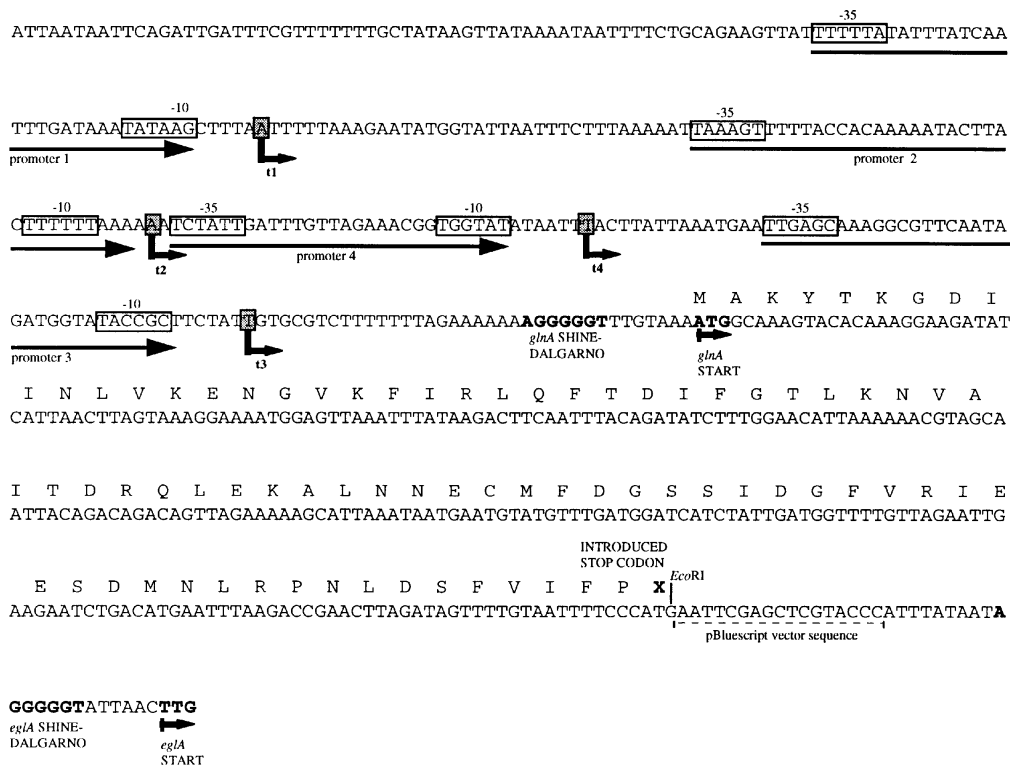


Figure 2. Sequence of the *C. beijerinckii* NCIMB 8052 *glnA* promoter region and the *egIA* reporter gene. The positions of Shine-Dalgarno sequences (bold), start codons (bold), transcriptional start sites (t<sub>1</sub>-t<sub>4</sub>; grey boxes), proposed promoter regions (bold arrows) and consensus sequences (boxed) are illustrated. Also illustrated is the amino acid sequence of the truncated 5' terminus of the *glnA* structural region. PBluescriptSK<sup>+</sup> vector sequence is underlined with a dotted line.

Medium (CBM; Jones *et al.*, 1982). Cells were harvested, lysed by sonication and fractionated into soluble and insoluble cellular fractions (Table 2). Very low activity could be detected in the soluble and insoluble fractions of the cell lysates from all three constructs. Although small amounts of endoglucanase activity (<1 μmol pNP/min/OD unit) could be detected in the extracellular culture medium of *C. beijerinckii*(pFNK1) and *C. beijerinckii*(pER1), that of the *egIA* gene expressed from its own promoter (pEC1) was shown to contain relatively high levels of activity (17.3 μmol pNP/min/culture OD unit) at mid exponential phase and 30.52 μmol pNP/min/culture OD unit at late exponential phase. Qualitatively, the results observed for the culture supernatant were reflective of those observed in the plate assays, with the exception of *C. beijerinckii* (pER1), the promoterless *egIA*. Although this strain produced small

zones of CMC digestion on plates, only very low activity could be measured after growth in liquid medium. The most probable explanation for this is that the zones on CMC plates are due to the local accumulation of endoglucanase resulting from very low levels of expression of the promoterless construct during the 18h incubation period. In contrast, the endoglucanase activity secreted into the liquid culture medium would be diluted to very low levels. It is therefore apparent that plate assays are at best a crude qualitative measure of endoglucanase activity. From the results presented here, it can be concluded that the endoglucanase activity produced by the cloned *egIA* gene from *C. acetobutylicum* P262 is exported from the host *C. beijerinckii* 8052 cells into the medium, and does not remain associated with the cell in any way.

### Endoglucanase Expression in *C. beijerinckii*

Expression of the *egIA* gene in *C. acetobutylicum* P262 was previously reported to be inducible rather than constitutive (Allcock *et al.*, 1981). Although *C. beijerinckii* NCIMB 8052 cannot utilise cellulose as a substrate (Minton *et al.*, 1993), it is possible that a homologue of the *egIA* gene does exist in *C. beijerinckii* as common gene arrangements have been previously reported for these two strains. The effect of various carbohydrate sources on the expression of the reporter under the control its own promoter was therefore studied in the heterologous host, *C. beijerinckii*. Cultures of *C. beijerinckii* NCIMB 8052 harbouring pFNK1, pER1 or pEC1 were grown in CBM with glucose, sucrose, galactose, fructose or cellobiose

Table 2. Localisation of the Endoglucanase Activity in Log-Phase Cells of *C. beijerinckii* Containing Different Plasmid Constructs

Plasmid construct	%Total Enzyme Activity		
	Cytoplasm	Membrane	Extracellular
pFNK1	0	0	100 <sup>a</sup>
pER1	0	0	100 <sup>a</sup>
pEC1	10.5	0	89.5 <sup>b</sup>

<sup>a</sup>Total enzyme activity in extracellular fraction: <1 μmol pNP/min/OD unit  
<sup>b</sup>Total enzyme activity in extracellular fraction: 17.3 μmol pNP/min/OD unit

as carbohydrate sources. Growth of the various strains was monitored in each medium and endoglucanase activity was assayed in the culture supernatant at mid- and late-exponential phase and in stationary phase. In all cases, *C. beijerinckii* (pFNK1) and *C. beijerinckii*(pER1) produced very little, if any, reporter activity at any stage of growth (results not shown). This confirms that under the conditions of these experiments, there was no significant induction of endogenous endoglucanase activity by any of the tested carbohydrates. Expression of the cloned *eglA* gene under the control of its own promoter (pEC1) was assayed in response to the same carbohydrate sources. In all media, endoglucanase activity in mid-exponential phase was relatively low, varying between 6.8 and 15.7  $\mu\text{mol pNP/min/OD}$  with progressively higher levels as growth continued. This pattern could be interpreted as repression of *eglA* expression under conditions of carbohydrate excess, with the repression being lifted as the easily metabolised carbohydrates were exhausted. Alternatively, this pattern could represent accumulation of activity as a result of constitutive *eglA* expression. The enzyme was observed to be relatively stable under the culture conditions employed, and it is likely that the activity measured during stationary phase represented the sum total of all enzyme produced during growth, rather than that being produced at the time of the assay. Growth on cellobiose or fructose resulted in lower levels of endoglucanase activity than the other carbohydrate tested, although both cellobiose and fructose are readily utilised by *C. beijerinckii*.

The promoter from the *C. beijerinckii glnA* gene encoding the glutamine synthetase enzyme, was cloned into pER1 to test endoglucanase expression from a heterologous promoter. Reporter expression was assayed in the extracellular medium from cells containing the different plasmid constructs grown in CBM with Em (10  $\mu\text{g/ml}$ ) to mid- and late-exponential phase. In contrast to the expression from the native *eglA* promoter, expression from the *glnA* promoter resulted in high level expression of the reporter gene, showing 11-fold and 8-fold increases over the *eglA* promoter in mid- and late- exponential phase respectively (Figure 3).

## Conclusion

The *eglA* of *C. acetobutylicum* P262 is an attractive candidate for a reporter gene for *C. beijerinckii* NCIMB 8052 for several reasons. It has been shown to be expressed in different bacterial strains, notably *E. coli* (Zappe *et al.*, 1988) as well as the anaerobic, Gram-negative organism, *B. fragilis* (Abratt *et al.*, 1993). The codon composition would therefore be expected to allow efficient expression in other clostridial isolates. The activity of the cloned enzyme has been well characterised in these heterologous hosts (Zappe *et al.*, 1988; Abratt *et al.*, 1993) and well-established, relatively easy and inexpensive colorimetric assays are available (Desphande, *et al.*, 1984). We have demonstrated that the *eglA* gene was expressed and efficiently secreted by *C. beijerinckii* NCIMB 8052, and that endogenous endoglucanase activity which may have interfered with the interpretation of the assays was not detected under the experimental conditions tested. Furthermore, qualitative assessment of endoglucanase activity can be performed using a plate assay, which is useful when screening large numbers of constructs expressing the reporter gene.

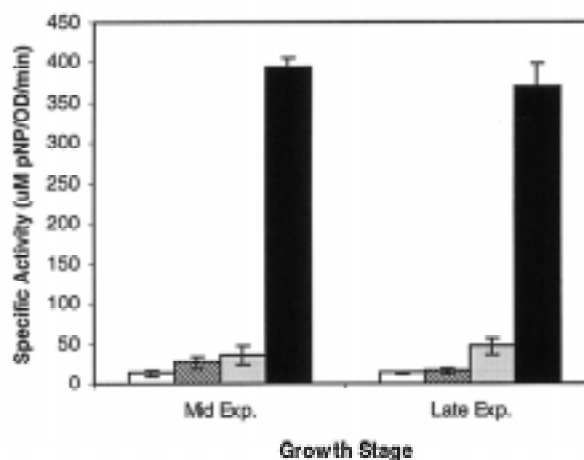


Figure 3. Expression of different reporter gene constructs at mid- and late-exponential growth phase. Endoglucanase activity was measured in supernatant of cultures grown in CBM, and is expressed as  $\mu\text{mol p-nitrophenol released/min/culture OD unit}$ . All values represent the average of three independent experiments. Strains tested were as follows: (open bars) *C. beijerinckii*(pFNK1); (diagonally-striped bars) *C. beijerinckii*(pER1); (stippled bars) *C. beijerinckii*(pEC1); (filled bars) *C. beijerinckii*(pEGP1).

To date very little is known about *Clostridium* promoters despite earlier studies on this subject (Graves and Rabinowitz, 1986; Minton *et al.*, 1993). Regulatory mechanisms controlling genes also differ vastly from one bacterium to another, and promoter fusion analysis of different genes and operons can contribute useful information in this regard. The development of a promoterless reporter plasmid based on the *C. acetobutylicum eglA* gene is therefore a useful tool in the qualitative and quantitative analysis of gene expression in *C. beijerinckii*. The plasmid on which the reporter vector is based is a *Bacillus/Clostridium* shuttle vector, which is stably maintained in *C. beijerinckii* and does not show the segregational instability which is typical of *E. coli/Clostridium* shuttle vectors. In a preliminary attempt to demonstrate the use of this vector as a reporter plasmid, a transcriptional fusion was created by ligation of the *glnA* promoter upstream of the promoterless *eglA* gene. Very efficient endoglucanase expression and secretion was obtained under the control of this heterologous promoter, resulting in as much as 11-fold increase in expression levels over the native *eglA* promoter. The reporter system therefore shows great potential for use as a tool for investigating transcriptional regulation and the structure of promoters in *Clostridium*.

## Experimental Procedures

### Plasmid Construction

The *EcoRI-PstI* fragment of pHZ117 $\Delta$ 12 (Table 1), containing the promoterless *eglA* gene (*eglA $\Delta$ P), was blunted-ended with T4 DNA polymerase and cloned into the *PvuII* site of pFNK1, resulting in pER1 (Figure 1). To generate the control plasmid, pEC1, the *EcoRI-PstI* fragment of pHZ117, which carries the coding region and functional promoter of *eglA* was cloned in the same manner. The promoter region of the *C. beijerinckii glnA* gene was ligated to *eglA $\Delta$ P resulting in plasmid pEGP1 by the following means: The plasmid pHZ117 $\Delta$ 12 was digested with *PstI* and the resulting cohesive ends removed with T4 polymerase. The linearised plasmid was then cut with *EcoRI* and the 1.4 kb fragment carrying the *eglA $\Delta$ P gene was cloned into the *SmaI-EcoRI* sites of pBluescriptSK<sup>+</sup>, creating pESK (Table 1). The 726bp *StyI* fragment of pCBN1 containing the promoter region of the *glnA* gene, was inserted into the blunted *EcoRI* site of pESK, upstream***