Journal of Industrial Microbiology & Biotechnology (2000) 25, 204–212

www.nature.com/jim

Development of marker-free strains of *Bacillus subtilis* capable of secreting high levels of industrial enzymes

B Widner, M Thomas, D Sternberg, D Lammon, R Behr and A Sloma

Novo Nordisk Biotech, Inc., Davis, CA 95616

Different strategies have been employed to achieve high-level expression of single-copy genes encoding secreted enzymes in *Bacillus subtilis*. A model system was developed which utilizes the *aprL* gene from *Bacillus clausii* as a reporter gene for monitoring expression levels during stationary phase. An exceptionally strong promoter was constructed by altering the nuceotide sequence in the −10 and −35 regions of the promoter for the *amyQ* gene of *Bacillus amyloliquefaciens*. In addition, two or three tandem copies of this promoter were shown to increase expression levels substantially in comparison to the monomer promoter alone. Finally, the promoter and mRNA stabilization sequences derived from the *cry3A* gene of *Bacillus thuringiensis* were used in combination with the mutant *amyQ* promoter to achieve the highest levels of *aprL* expression. These promoters were shown to be fully functional in a high-expressing *Bacillus* strain grown under industrial fermentation conditions. The ability to obtain maximum expression levels from a single copy gene now makes it feasible to construct environmentally friendly, marker-free industrial strains of *B. subtilis*. *Journal of Industrial Microbiology & Biotechnology* (2000) **25,** 204–212.

Keywords: promoter; RNA stability; gene expression; *Bacillus subtilis*

Introduction

Many *Bacillus* species are considered industrial workhorses because of their ability to secrete enzymes at very high levels [15]. In order to achieve maximum expression of a particular gene in *Bacillus* it is almost always necessary to amplify the expression cassette [4,31]. This is accomplished by integrating into the host chromosome a plasmid, which contains the expression cassette, an antibiotic resistance marker, and a site of homology, and selecting for growth in the presence of progressively higher concentrations of antibiotic. This results in the generation of directly repeated, tandem copies of the expression cassette and antibiotic resistance marker. The fact that amplification is required for optimal expression implies that mRNA is limiting when a gene is expressed from a single copy on the chromosome. During amplification, eventually a point is reached where increasing the copy number of an expression cassette no longer results in an increase in expression. Presumably, when this condition occurs, the levels of gene-specific message have saturated the posttranscriptional machinery and the bottleneck for expression now lies downstream of transcription.

The strains containing tandem repeats of the expression cassettes are actually quite stable and loss of the cassettes due to homologous recombination between the direct repeats is not a problem during fermentation [18]. Nonetheless, single-copy constructs which are not flanked by direct repeats theoretically would insure 100% stability during a fermentation. In addition, from a governmental regulatory standpoint, it is desirable to have strains that are devoid of antibiotic resistance genes. In fact, with the increasing appearance of medically relevant bacterial strains that are refractile to all antibiotics in use, there will undoubtedly be increasing pressures on industry to stop using antibiotic resistance markers altogether [22]. Technically, this is impossible to accomplish with the current strains containing tandem repeats of the expression cassettes. The resistance markers are absolutely required in order to select for amplification of the expression cassette; no alternative is currently available. Furthermore, any scheme to loop out resistance genes would unfortunately remove the expression cassettes as well. One way around this problem would be to construct strains which contain only one copy of an expression cassette yet still yield saturating levels of gene-specific mRNA. Strains of *Bacillus* which contain only a single copy of an expression cassette and no antibiotic resistance markers can be readily constructed using traditional gene-replacement techniques.

The steady-state level of a particular mRNA is determined by its rate of synthesis and its rate of degradation. The former can be increased by simply utilizing a stronger promoter. Traditionally, this has been the method of choice for overexpressing heterologous genes in the laboratory. The latter can be decreased by incorporating stabilization sequences into the message. Examples of this strategy are found in nature, and recently it has been proposed as an alternative method for controlling heterologous gene expression [9]. In particular, two reports have appeared in the last few years that describe the positive effect a strong ribosome binding site located at the 5′ end of an mRNA can have on its stability in *Bacillus* [3,17]. The mRNAs transcribed from both the A region of bacteriophage SP82 of *Bacillus subtilis* and the *cry3A* crystal protein gene of *Bacillus thuringiensis* have been shown to have a very strong ribosome binding site located near the 5′ end of the message; in the latter case, it is located approximately 125 bases upstream of the actual ribosome binding site for the *cry3A* gene. In fact, these mRNA stabilizing sequences were shown to function when fused to the 5′ end of heterologous messages as well. It has been proposed that the strong ribosome binding site somehow impedes access of RNases to the 5′ end of the transcript, thereby increasing the half-life of the message up to severalfold.

Correspondence: Dr B Widner, Novo Nordisk Biotech., Inc., Davis, CA 95616, USA.
E-mail: wwidner@nnbt.com
Received 5 January 2000; accepted 26 June 2000

In this report, we describe the application of these strategies to generate a single-copy expression cassette that is capable of producing saturating levels of mRNA in *B. subtilis*. This technology has allowed us to construct production strains which are devoid of any antibiotic resistance markers.

Materials and methods

Bacterial strains and growth conditions

Escherichia coli strains DH5α, JM101, and TOP10 (Invitrogen, Carlsbad, CA, USA) were used for cloning purposes. Expression cassettes were transformed into the low secretor *B. subtilis* PL1801*spoIIE*, a derivative of *B. subtilis* 168 (Bacillus Stock Center, Columbus, Ohio) with deletions in *amyE*, *aprE*, and *nprE*, and a Tn*917* insertion in the *spoIIE* gene, and the high secretor *B. subtilis* A164Δ5, which has deletions in the *spoIIAC*, *amyE*, *aprE*, *nprE*, and *srfAC* genes (Table 1). *B. subtilis* cells were made competent by the method of Anagnostopoulos and Spizizen [5].

Cells were grown either on LB [25] or TBAB (tryptose blood agar base, Difco, Detroit, MI, USA) plates, supplemented with chloramphenicol at 5 μg/ml, neomycin at 5 μg/ml, or ampicillin at 100 μg/ml, as appropriate. Plates were overlaid with TBAB containing 1% final concentration nonfat skim milk for the detection of protease-producing colonies. Shake flask cultures were performed in PS-1 medium containing, per liter: 10 g sucrose, 40 g soybean flour, 10 g Na_2HPO_4, 5 g Ca_2CO_3, and 0.1 ml pluronic L61 (BASF Corporation, Parsippany, NJ) as an antifoaming agent.

Plasmids

Plasmid pDG268ΔNeo is a pDG268-derived plasmid and was used for introducing *aprL* expression cassettes into the *amyE* locus of *B. subtilis* in single-copy. This plasmid has been described in detail in US patent 5,955,310 [29].

Plasmid pRB115 is an *E. coli/B. subtilis* shuttle vector. It was used for deleting the *cat* gene from selected *B. subtilis* strains containing a single copy of an expression cassette. It was constructed by SOE (splicing by overlap extension) [16] using plasmid pDG268 as template DNA [6]. A 520-bp upstream deletion fragment was PCR amplified using the following two primers: 5′-GTCAAGCTTAATTCTCATGTTTGACAGCTTATCATCGG-3′ and 5′-TCCATGTCGAGAGTTATATCACCCCTTGTCACTAAG-3′. The underlined nucleotides in the former were added to create a *Hin*dIII site. The underlined nucleotides in the latter represent the region of overlap between the upstream and downstream fragments. A 524-bp downstream deletion fragment was PCR amplified using the following two primers: 5′-TGATATAACTCTCGACATGGATGAGCGATGATGATATCC-3′ and 5′-GCGGCCGCGGTACCTGTGTTACACCTGTT-3′. The underlined nucleotides in the former represent the region of overlap between the upstream and downstream fragments. The underlined nucleotides in the latter were added to create an *Asp*718 site. These two fragments were joined by SOE to generate a 1023-bp cat deletion fragment; the PCR product was cloned directly into pCR2.1TOPO vector (Invitrogen, Carlsbad, CA, USA). This fragment was subcloned into the *E. coli/B. subtilis* temperature-sensitive shuttle vector pSK.sup.+/pE194 [26] as an *Asp*718/*Hin*dIII fragment to yield pRB115.

Table 1 Bacterial strains and plasmids

Bacterial strains (genotype)	Reference
B. subtilis	
PL1801 *spoIIE* (*aprE*Δ *nprE*Δ *amyE*Δ *spoIIE*::Tn*917*)	[29]
A164Δ5 (*aprE*Δ *nprE*Δ *amyE*Δ *spoIIAC*Δ *srfAC*Δ)	[26]
E. coli	
DH5α	[7]
JM101	[30]
TOP10	[14]
Plasmids	
pCR2.1-TOPO	[20]
pDG268	[6]
pDG268A Neo	[29]
RB115	This work

DNA isolation

Plasmids were grown in *E. coli* and purified using the Qiagen plasmid kit (QIAGEN, Inc., Valencia, CA, USA). Genomic DNA was isolated by the method of Pitcher *et al.* [24]. Restriction enzyme digestion, Southern hybridization, and subcloning were performed according to standard techniques [25].

Oligonucleotide synthesis, PCR amplification, and DNA sequencing

Oligonucleotides were synthesized on an Applied Biosystems 394 DNA/RNA Synthesizer. PCR amplification was performed on either an Ericomp Twinblock EZ Cycler (Ericomp, Inc., San Diego, CA) or a RoboCycler 40 (Strategene, La Jolla, CA). DNA sequencing was performed on an Applied Biosystems 373A DNA Sequencer (Applied Biosystems, Foster City, CA, USA).

PCR amplification of promoter elements and assembly of aprL expression cassettes

The *amyL* promoter (from *Bacillus licheniformis*) was PCR amplified using plasmid pDN1981 as template DNA [13,19]. The *cry3A* promoter was PCR amplified using total DNA obtained from *B. thuringiensis* var. *tenebrionis* (NB125) as template DNA [1]. The *amyQ* promoter (from *Bacillus amyloliquefaciens*) and the *aprL* gene (from *Bacillus clausii*) were both PCR amplified using plasmid pSX222 as template DNA [8,23]

The consensus *amyQ* promoter was constructed as follows: The following two oligonucleotides were annealed together and extended with Klenow fragment to generate a 68-bp double-stranded fragment containing mutations in the -10 and -35 regions of the wild-type *amyQ* promoter (highlighted in bold letters): 5′-GGAATAAAGGGGGG**TTGACA**TTATTTTACTGATATG**TATAAT**AT-3′ and 3′-AATAAAATGACTATACATATTATATTAAACATATTCTTTTACCTCGAG-5′. A second double-stranded fragment comprising the upstream region of the *amyQ* promoter was generated by PCR using the following two primers: 5′-GGCCTTAAGGGCCTGCAATCGATT-3′ and 5′-TGTCAACCCCCCTTTATTCCTT-3′. Both double-stranded DNA fragments were then fused together by traditional SOE methods [16] (SOE overlaps are underlined) to generate a mutated version of the *amyQ* promoter designated "consensus *amyQ*"; primers 5′-GGCCTTAAGGGCCTGCTGTCCAGACTGTCCGCT-3′ and 5′-GAGCTCCATTTTCTTATACAAATTATAT-3′ were used to PCR amplify the consensus *amyQ* promoter fragment.

All PCR fragments were cloned using the pCR2.1 vector and either the TA or TOPO TA cloning kit supplied by Invitrogen. All promoter fragments are flanked by an *Sfi*I or *Apa*I site at the 5′ end and a *Sac*I site at the 3′ end. The *amyQ*, consensus *amyQ*, and *amyL* promoter fragments include the −10 and −35 regions, the transcription start region, and at least 46 bases of sequence upstream of the −35 region to insure full promoter function. Likewise the *cry3A* promoter fragment was designed to contain the −10 and −35 regions, sufficient upstream sequence, and downstream sequence including the mRNA stabilization sequence. All of the *aprL* constructions referred to in this report (excluding those utilizing the consensus *amyQ* promoter) are described in US patent 5,955,310 [29]. The *aprL* expression cassettes utilizing the consensus *amyQ* promoter were constructed in a fashion identical to those described utilizing the *amyQ* promoter.

Strain constructions

The *B. subtilis* strains containing tandem repeats of the *amyL* promoter/*aprL* expression cassette were constructed as follows. First the cassette was cloned into the *Sfi*I–*Bam*HI sites of the integration plasmid pCAsub2 [26]. The resulting plasmid was then transformed into *B. subtilis* selecting for chloramphenicol resistance and integration into the *amyE* locus on the chromosome. A transformant was chosen and passaged onto TBAB plates containing progressively higher concentrations of chloramphenicol ultimately reaching levels of 80 μg/ml. This treatment results in a strain which contains several tandem repeats (≥8) of the expression cassette in the *amyE* locus of the chromosome.

All of the single-copy *aprL* expression cassettes were integrated into the *amyE* locus of *B. subtilis* strain PL1801*spoIIE* as follows: plasmid pDG268ΔNeo and derivatives harboring *aprL* expression cassettes were digested with *Sca*I to linearize the plasmids. One microgram of linearized plasmid DNA was used to transform competent PL1801*spoIIE* to chloramphenicol resistance. Theoretically, all transformants should contain a single copy of the *aprL* expression cassette in the *amyE* locus as a result of a double crossover event; this was confirmed by PCR and/or Southern blotting analysis or by confirmation of the absence of the neomycin resistance marker.

The *aprL* expression cassettes were transferred to the A164Δ5 background as follows. Chromosomal DNA was isolated from the appropriate PL1801*spoIIE* strains harboring single-copy *aprL* expression cassettes and was used to transform competent A164Δ5 to chloramphenicol resistance. Halo formation on TBAB plates overlayed with 1% final concentration of nonfat skim milk confirmed that the expression cassette was indeed present.

aprL expression in B. subtilis PL1801spoIIE

Integrants containing a single copy of each expression cassette were grown to mid-log phase in 1 ml of LB broth at 37°C. Then 100 μl of each mid-log phase culture was inoculated into 50 ml of PS-1 medium. The cultures were shaken vigorously at 37°C for 5 days. Aliquots (1 ml) were removed after 3, 4, and 5 days, centrifuged at 12,000×*g* for 2 min, and 0.5 ml of each supernatant was frozen at −20°C until all samples were taken. The frozen samples were then thawed and assayed for AprL activity using the following protocol to determine relative yields.

The assay for AprL activity was performed with casein fluorescein isothiocyanate as substrate. The casein fluorescein isothiocyanate stock solution was prepared according to Twining [27] and contained 0.5 mg of casein fluorescein isothiocyanate per 100 ml of 50 mM Tris–HCl pH 7.2 buffer. The assay reaction was initiated by the addition of 40 μl of a 1:1 v/v mixture of the stock solution and 0.25 M borate pH 9.0 buffer to 10 μl of the enzyme sample diluted in 0.25 M borate pH 9.0 buffer as appropriate. The reaction was incubated for 10 min at 37°C and then quenched by adding 150 μl of 5% trichloroacetic acid. The quenched reaction was placed at 4°C for 10 min and then centrifuged at 10,000×*g* for 2 min. A 10-μl aliquot of the supernatant was transferred to a test tube containing 2 ml of 0.5 M borate pH 9.0 buffer and mixed well. A 200-μl aliquot of this solution was transferred to a black U-bottom 96-well plate (Dynatech Laboratories, Inc., Chantilly, VA, USA) and fluorescence was measured using a Fluorolite 1000 fluorimeter (Dynatech Laboratories, Inc., Chantilly, VA, USA) using channel 3 at reference setting 1176 and a lamp voltage of 4.1 V. AprL activity was calculated by reference to a standard curve generated with an AprL standard (Novo Nordisk A/S, Bagsværd, Denmark) in the range of 1.8–9.0 NPU (Novo protease unit) per ml. The activity of the standard was determined according to Novo Analytical Method AF 220/1-GB available upon request from Novo Nordisk.

mRNA isolation

Overnight cultures (0.5 ml) of *B. subtilis* strains were added to 25 ml of LB medium and grown at 37°C with constant shaking for 6.5 h (at which time cells were well into stationary phase). At this stage, rifampicin was added (200 μg/ml) to stop mRNA synthesis and 3-ml aliquots were taken at various time points and transferred into two 1.7-ml microcentrifuge tubes. Cells were immediately pelleted and the supernatant removed. These pellets were resuspended in residual supernatant and transferred to a single tube. The cells were repelleted and resuspended in 100 μl diethyl-pyrocarbonate-treated distilled water [25]. The cells were then disrupted and the RNA extracted using the FastRNA BLUE Kit (Bio 101, Vista, CA, USA). RNA concentrations were determined by measuring the absorbance at 260 nm. All samples were diluted to a concentration of 0.2 μg/ml with diethyl-pyrocarbonate-treated distilled water and stored at −80°C.

Probe synthesis

The region extending from nucleotide position 120 to 726 (with respect to the ATG start codon) of the *aprL* gene was amplified by polymerase chain reaction (PCR) using SAVF3 (5′-GAGCAGGAAGCTGTCAGTGAG-3′) and SAVR2 (5′-GGCACTTGGCGAAGGGCTTCC-3′) primers. This PCR fragment was DIG-dUTP labeled using the PCR DIG Probe Synthesis Kit (Boehringer Mannheim, Indianapolis, IN, USA) and purified using the QIAquick Gel Extraction Kit (QIAGEN, Valencia, CA, USA).

mRNA half-life determination

RNA (1 μg per lane) was separated on gels that contained 1.0% agarose, 10.0% 10× MOPS (Amresco, Solon, OH, USA), and 16.2% formaldehyde (Sigma, St. Louis, MO) and transferred to an MSI membrane (Micron Separations, Inc., Westborough, MA, USA) in 20× SSC for 1 h at 80 mmHg using a Stratagene Posiblot pressure blotter. Samples were prehybridized at 65°C for 1 h in DIG Easy Hyb (Boehringer Mannheim, Indianapolis, IN, USA). The prehybridization solution was replaced with fresh DIG Easy Hyb and the denatured labeled probe was added. The membrane was

incubated overnight at 65°C and washed five times, each at 45°C for 10 min (twice in 2× SSC/0.1% SDS, twice in 2× SSC, and once in 0.4× SSC). The hybridized blots were probed with anti-digoxigenin-AP (Boehringer Mannheim, Indianapolis, IN, USA) and reacted with ECF substrate in detection buffer from the Signal Amplification Module by Amersham, Arlington Heights, IL, USA. The fluorescent substrate reactions were analyzed by STORM fluorescent imaging.

Three-liter fermentor experiments

aprL expression: The main tank medium contained standard salts (in g/ml: 6.5 KH_2PO_4, 4.5 Na_2HPO_4, 3.0 $(NH_4)_2SO_4$, 2.0 Na_3 citrate·$2H_2O$, 3.0 $MgSO_4 \cdot 7H_2O$), trace metals (in mg/l: 120.0 $FeSO_4 \cdot 7H_2O$, 30.0 $MnSO_4 \cdot H_2O$, 12.0 $CuSO_4 \cdot 5H_2O$, 12.0 $ZnCl_2$), vitamins (in mg/l: 4.5 thiamine HCl, 2.4 riboflavin, 12.0 nicotinic acid, 15.0 Ca pantothenate, 0.15 biotin, 1.5 folic acid, 3.0 pyridoxal HCl), and was supplemented with 20 g sucrose/l and 35 g casitone/l (casitone is a protein hydrolysate from Difco, Detroit, MI, USA.) The feed medium contained 470 g sucrose and 42 g casitone per kilogram. Fermentation temperature was controlled at 38°C and pH between 6.8 and 7.2 by titration with H_3PO_4 and NH_4OH. Feeding was started about 5 h after inoculation with a rate increase to about 13 g/l_0 h (l_0 = initial tank volume) by 12 h and then a more gradual rate increase to a maximum of about 17 g/l_0 h by 35 h that was held to the end of the fermentation (47 h). Dissolved oxygen levels were maintained at or above 25% saturation.

Results and discussion

A model system for monitoring promoter activity in B. subtilis

In order to monitor gene expression in a meaningful way it is desirable to use a reporter gene that can be readily assayed in the crude media normally used for shake flask analyses. Furthermore, the reporter gene should be relevant to the types of genes typically expressed in industrial settings, i.e., genes encoding secreted proteins. The *aprL* gene [8] (encoding the subtilisin from *B. clausii*) was chosen because the enzyme is readily secreted by *Bacillus* and an easy biochemical assay exists. Furthermore, the gene can be manipulated in *E. coli*, unlike the genes encoding α-amylases which are not tolerated very well by this organism (unpublished observations). Also, the AprL protease is a major enzyme product which is sold by the tons worldwide, so it is logical to develop promoters to express this gene with the hope that any promoters which perform well for *aprL* expression will also perform well with other genes.

The first objective was to develop a system for measuring promoter activity accurately. We utilized plasmid pDG268 which was already developed for monitoring *lacZ* expression in *B. subtilis* [6]. This particular plasmid contains the *lacZ* gene with a polylinker located just upstream of the coding sequence and a choramphenicol resistance marker located upstream of the polylinker. All of this sequence is flanked by the 5′ and 3′ portions of the *amyE* gene in order to introduce a single-copy of

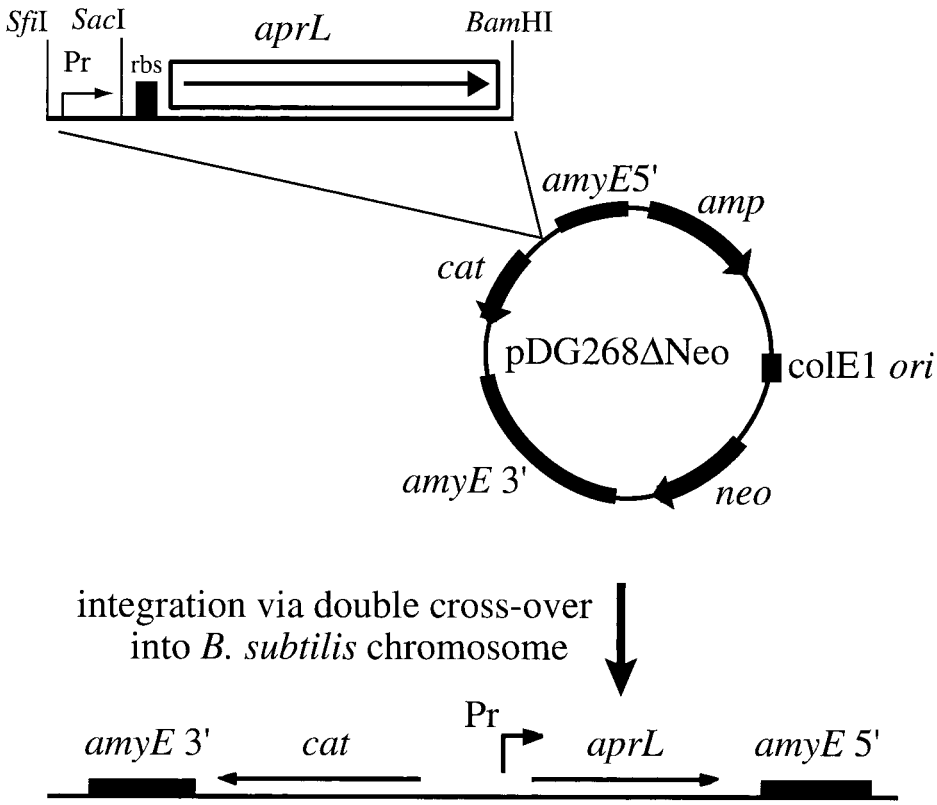

Figure 1 Strategy for introducing single copies of various promoters expressing the *aprL* gene into the *amyE* locus of *B. subtilis*. Pr, promoter; rbs, *aprL* ribosome binding site; *aprL*, the gene encoding the subtilisin from *B. clausii*; *amp*, the gene encoding beta-lactamase; *cat*, the gene encoding chloramphenicol acetyl transferase; *neo*, the gene encoding neomycin phosphotransferase; and *amyE* 5′ and *amyE* 3′ designate the 5′ and 3′ regions of the *amyE* locus encoding the alpha amylase of *B. subtilis* (sites of homology for homologous recombination).

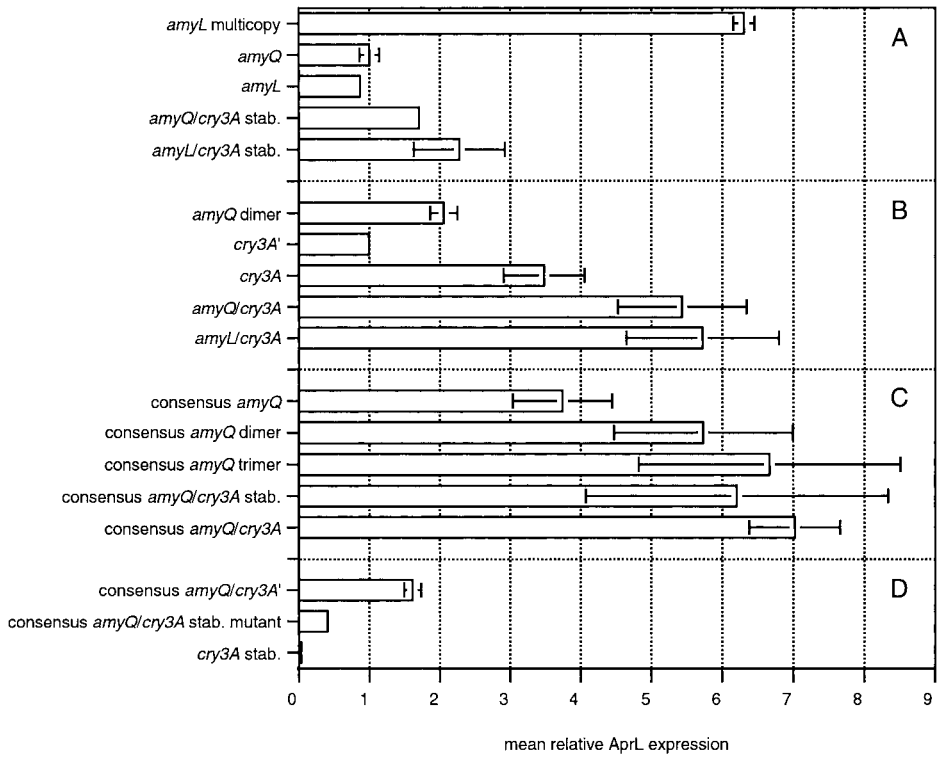

Figure 2 Relative AprL expression from various promoters. *B. subtilis* PL1801 *spoIIE* integrants with *aprL* expression cassettes inserted in single copy at the *amyE* locus were subjected to analysis in shake flasks cultures. Also tested was a PL1801 *spoIIE* integrant containing tandem repeats of the *amyL* promoter/*aprL* expression cassette (*amyL* multicopy). Expression levels are reported relative to that from the *amyQ* promoter/*aprL* cassette, which has been set arbitrarily at 1.

the *lacZ* expression cassette into the *amyE* locus via double crossover. In order to modify this plasmid to more adequately fulfill our needs, the *lacZ* gene was replaced with a custom polylinker into which various *aprL* expression cassettes were cloned. In addition, a neomycin resistance marker was introduced outside of the *amyE* fragments in order to distinguish double *versus* single crossovers; this plasmid was designated pDG268ΔNeo (Figure 1).

Several *aprL* expression cassettes were assembled using plasmid pDG268ΔNeo. These were all integrated as single copies into the *amyE* locus of *B. subtilis* strain PL1801*spoIIE* (deleted in *amyE*, *aprE*, and *nprE*). Integrants were grown in shake flasks and assayed for AprL activity as an indicator of promoter strength. This particular *B. subtilis* strain was chosen because of its low protease and amylase background and its ability to become highly competent.

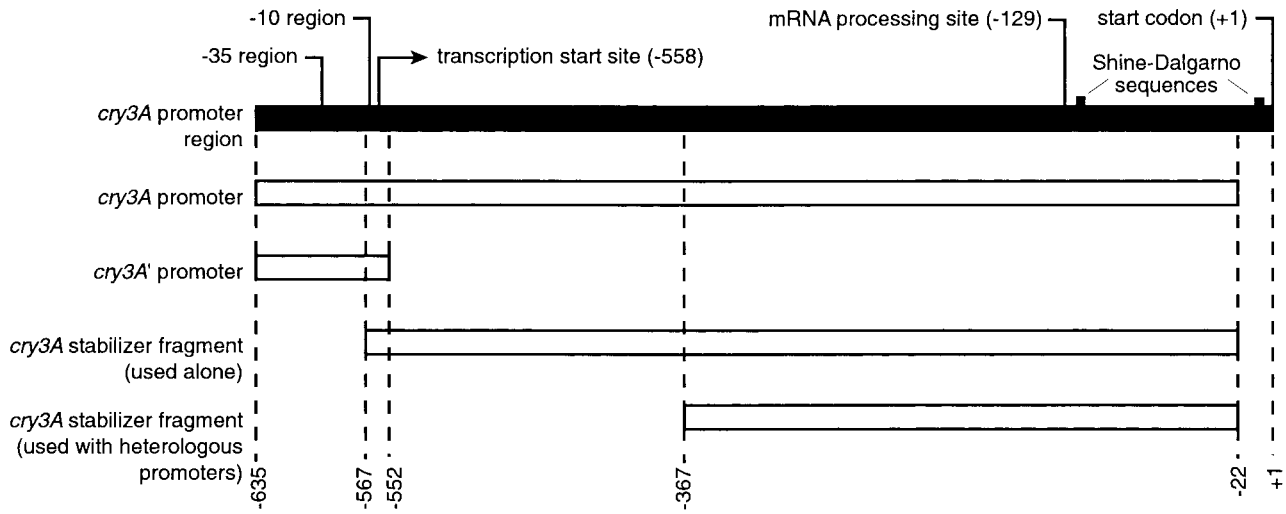

Figure 3 Organization of the *cry3A* promoter region [3]. Positions of the transcription start site, mRNA processing site, and DNA fragments used in *aprL* expression cassettes are indicated relative to the native *cry3A* initiation codon.

The cry3A mRNA stabilizer sequence increases gene expression

The promoters initially tested were the promoter for the *amyQ* gene of *B. amyloliquefaciens* [23] and the promoter for the *amyL* gene of *B. licheniformis* [13]. Both of these promoters function quite well in a *B. subtilis* background (unpublished results) and therefore were considered good examples of how strong promoters should perform. Shake flask results showed that the *amyQ* and *amyL* promoters function equally well in *B. subtilis* (Figure 2A). In comparison to a strain which contained several tandem repeats of the *amyL* promoter/*aprL* expression cassette integrated into the chromosome, expression levels from both strains containing single copies of the expression cassette were down approximately 6-fold. Thus, there needed to be a considerable improvement over these promoters in order to achieve the goal of maximal expression fron a single-copy gene on the chromosome.

The *cry3A* promoter drives the expression of the *cry3A* crystal protein gene to very high levels in its native host *B. thuringiensis*. A recent study has shown that the 5′ untranslated region of the *cry3A* mRNA confers enhanced message stability [3]. The general architecture of the *cry3A* promoter is shown in Figure 3. The −10 and −35 regions are located approximately 570 bp upstream of the coding region for the *cry3A* gene. The initial transcript is quickly processed to generate a shorter, very stable mRNA with the 5′ end mapping approximately 129 bp upstream of the ATG start codon for the gene. This untranslated leader contains an extra ribosome binding site near the 5′ end of the message which confers considerable stability to the *cry3A* mRNA.

The fact that the *cry3A* stabilizing sequence has a substantial effect on expression levels is an intriguing observation. We decided to test whether placing this stabilizer sequence downstream of our heterologous promoters would lead to an increase in gene expression. We cloned a 350-bp fragment of the *cry3A* sequence which includes the presumed processing/stabilizing sequence (Figure 3) downstream of the *amyQ* and *amyL* promoters to yield the *amyQ/cry3A* stab and *amyL/cry3A* stab promoters, respectively. Transcription from these promoters is predicted to result in a short-lived mRNA which is subsequently processed to yield a shorter, stable mRNA with the stabilizer sequence located at the very 5′ end of the message. Indeed, inserting the mRNA stabilizer sequence downstream of these promoters did increase *aprL* expression approximately 2-fold for each promoter as determined by analysis of shake flask cultures (Figure 2A).

Two promoters are better than one

Another alternative to amplifying entire expression cassettes to achieve mRNA saturation would be to express a single gene from multiple promoters arranged in tandem. To test this concept, we constructed an *amyQ* dimer promoter. The *amyQ* monomer promoter is a 113-bp fragment containing the −10 and −35 regions, the transcription start site, and approximately 46 bases upstream of the −35 region. This small fragment has full promoter activity and was ligated in a head-to-tail fashion with a second

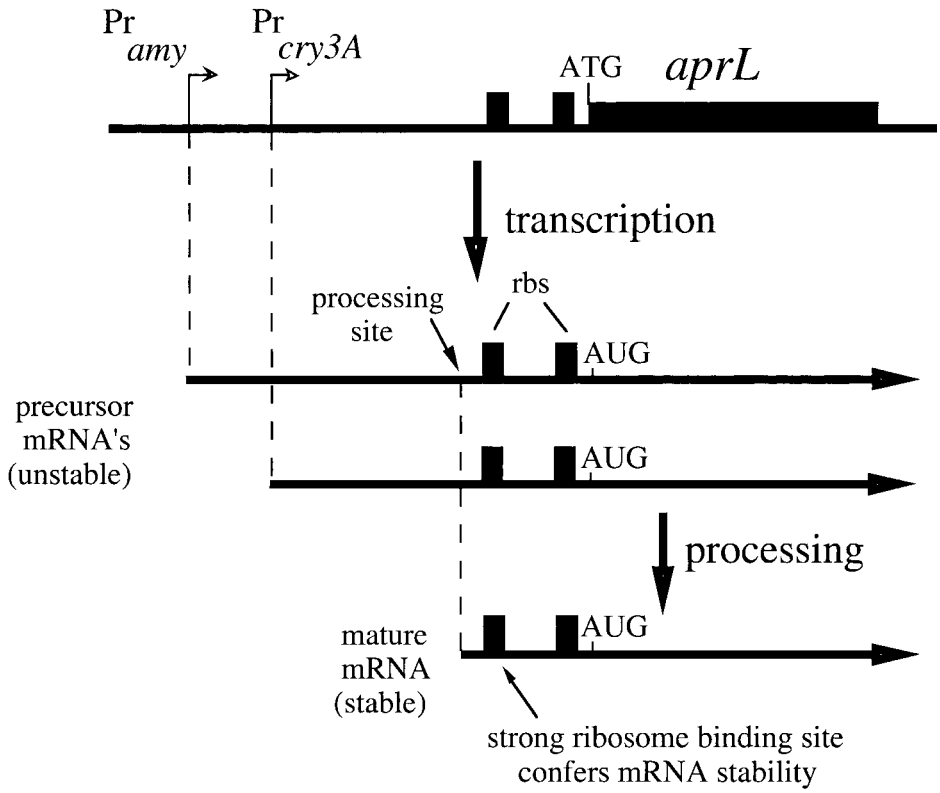

Figure 4 The tandem promoter/*cry3A* stabilizer model. Long precursor mRNAs are sythesized from two tandem promoters which are quickly processed to yield a discrete population of mature, stable mRNA. The shorter mRNA is more stable due to the presence of the extra *cry3A* ribosome binding site (rbs) at the 5′ end of the mRNA. Pr_{amy}, promoter for gene encoding an alpha-amylase (either *amyQ*, consensus *amyQ*, or *amyL*); Pr_{cry3A}, promoter for the gene encoding the cry3A crystal protein; ATG, the start codon for the *aprL* gene; AUG, the translation start codon for the *aprL* mRNA.

copy of the promoter to construct the dimer promoter. Presumably, this promoter should initiate transcription from two start sites, thereby leading to an increase in *aprL* message levels. This prediction appears to hold true; expression of *aprL* from the dimer promoter was approximately 2-fold higher in comparison to the monomer promoter (Figure 2B). Whether both of the promoters in the tandem functioned equally well was not determined.

We then tested how multiple promoters would perform in combination with the *cry3A* stabilizer sequence. The *amyQ* and *amyL* promoters were placed directly upstream of the *cry3A* promoter (including the mRNA stabilizing sequences) to yield *amyQ/cry3A* and *amyL/cry3A* tandem promoters. The *cry3A* promoter is a σ^A-type promoter, unlike other *cry* gene promoters which are known to be under sporulation control, and is therefore expected to function in the *spoIIE* background of the host strain [2]. It was predicted that expression from both promoters in these tandem arrays would lead to production of unstable transcripts due to the presence of a long untranslated region at the 5′ end of the messages. However, both populations of transcripts would contain the mRNA processing/stabilizing sequence and therefore should be processed to yield a shorter, stable message (Figure 4). Theoretically, this should lead to higher levels of gene expression. Comparing expression levels from the *amyQ, amyQ/cry3A* stab, *cry3A*, *cry3A′* (no stabilizer), and *amyQ/cry3A* tandem promoters, it is clear that the tandem promoter is superior to all of the other promoters; there is at least a 5-fold increase when compared to the *amyQ* or *cry3A′* promoters alone, a 3-fold increase compared to *amyQ/cry3A* stab, and a 1.5-fold increase compared to the *cry3A* promoter (Figure 2B). Similarly, the *amyL/cry3A* tandem promoter gives higher expression than the *amyL*, *amyL/cry3A* stab, *cry3A*, and *cry3A′* monomer promoters.

A consensus amyQ promoter

The *amyQ* and *amyL* promoters are transcribed by the σ^A-containing form of RNA polymerase, and both of these promoters deviate somewhat from the established consensus promoter sequence for this particular σ factor. Strong promoters transcribed by the σ^A holoenzyme in *B. subtilis* are known to contain the following −35 and −10 sequences: TTGACA and TATAAT,

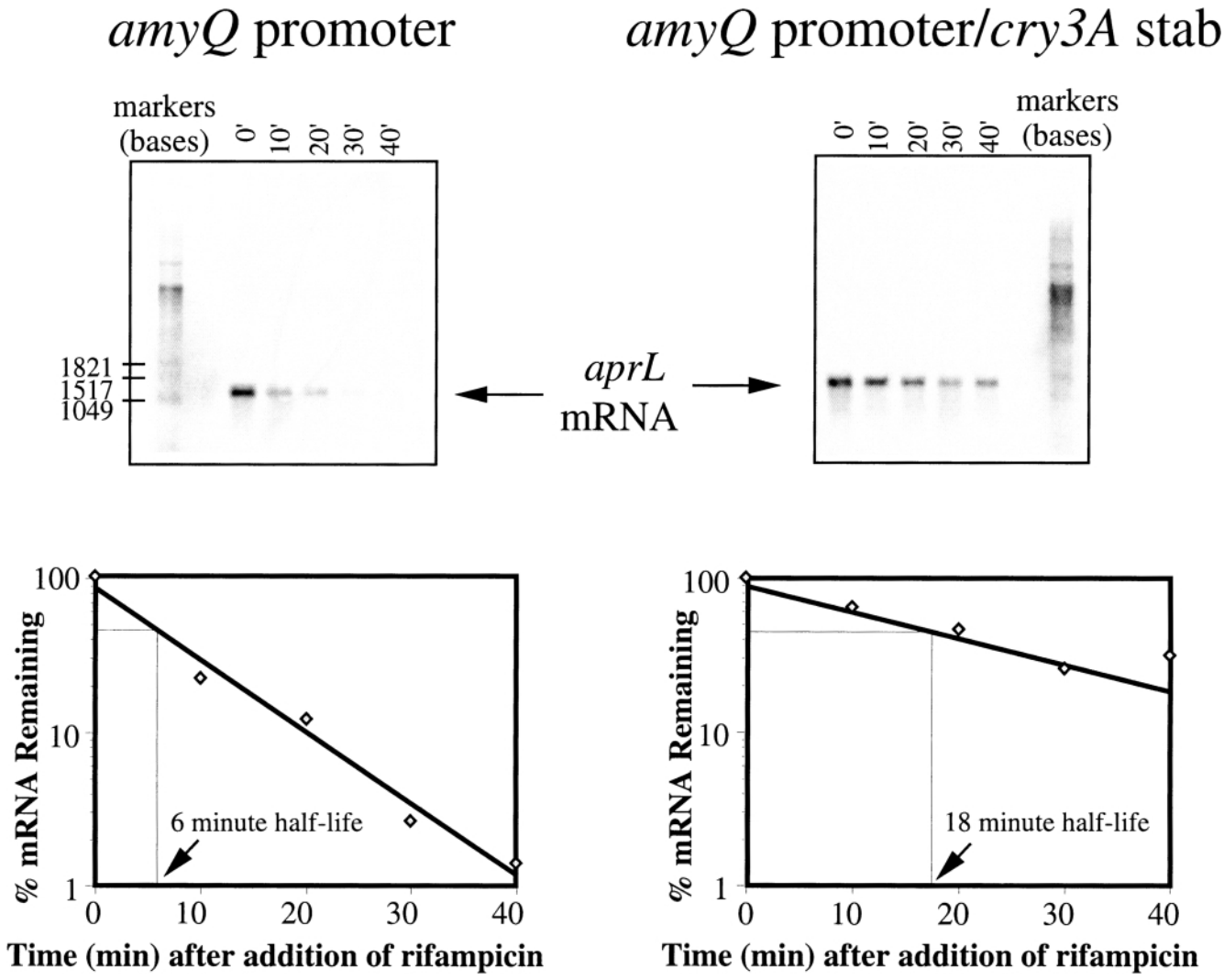

Figure 5 mRNA half-life determinations for *aprL* transcripts originating from the *amyQ* promoter (left panels) and the *amyQ* promoter including the *cry3A* stabilizer sequence (right panels). Top: Northern blot analyses showing *aprL* mRNA isolated at 0, 10, 20, 30, and 40 min after rifampicin addition. Bottom: a semilog plot of percent *aprL* mRNA remaining *versus* time (after rifampicin addition).

respectively; deviation from this consensus typically results in reduced expression [21]. In fact, altering the *B. subtilis amyE* promoter to conform to the consensus sequences resulted in a 56-fold increase in promoter activity [28]. A consensus *amyQ* promoter was constructed to test whether conforming more closely to the σ^A consensus sequence would lead to increased expression in *B. subtilis*. Results were quite encouraging: the consensus *amyQ* promoter gave higher expression than the wild-type *amyQ* promoter by a factor of four (Figure 2C).

We then constructed tandem and *cry3A*-based promoters similar to those described above using the stronger consensus *amyQ* promoter instead of the wild-type *amyQ* promoter. A consensus *amyQ* trimer promoter was also constructed that is similar to the dimer promoter except that there are three promoter elements oriented in a head-to-tail fashion instead of just two. Basically, the trend with this new set of promoters remained the same but the overall expression levels were significantly higher (Figure 2C). In particular, expression from the consensus *amyQ* trimer promoter, the consensus *amyQ/cry3A* tandem promoter, and the consensus *amyQ/cry3A* stab promoter were comparable to levels obtained from the strain containing tandem repeats of the *aprL* expression cassette. We have reached our goal of achieving saturating levels of mRNA from a single-copy gene on the chromosome.

The cry3A stabilizer increases the half-life of the aprL mRNA

The consensus *amyQ/cry3A* tandem promoter was analyzed in detail to determine which component parts actually contribute toward enhanced expression levels. Clearly, the sequence downstream of the *cry3A* promoter which includes the processing/stabilizing sequence is very important for enhanced expression; if this region is deleted (consensus *amyQ/cry3A′*) or if point mutations are introduced into the stabilizing RBS sequence (consensus *amyQ/cry3A* stab mutant), expression of *aprL* drops precipitously (Figure 2D). Also, the stabilizer sequence by itself (*cry3A* stab) has no apparent promoter activity. Interestingly, deleting the *cry3A* promoter and leaving the downstream processing/stabilizing sequence intact does not reduce expression (consensus *amyQ/cry3A* stab). Most likely, the consensus *amyQ/cry3A* tandem and the consensus *amyQ/cry3A* stab promoters are both producing mRNA levels at or close to saturation, making it difficult to detect any further contribution from the *cry3A* promoter itself.

The promoters containing the *cry3A* stabilizer sequence are expected to produce *aprL* transcripts of at least 1800 bases in length. However, these transcripts are expected to be rather short-lived (perhaps even undetectable) and processed to yield a population of mRNA of a single, discrete size of approximately 1300 bases. In addition, this processed mRNA should have a longer half-life in comparison to an *aprL* transcript that does not contain the *cry3A* stabilization sequence. To test these predictions, total mRNA was isolated from *B. subtilis* cells expressing *aprL* from the *amyQ* and the *amyQ/cry3A* stab promoters at various time points after the addition of rifampicin. Transcript size and the rate of mRNA decay were determined by Northern blot analyses (Figure 5). The strain lacking the *cry3A* stabilizing sequence (*amyQ*) showed a half-life of approximately 6 min. With the addition of the *cry3A* stabilization sequence downstream of this promoter, there was an increase in half-life to 18 min. The *aprL* transcript was estimated to be approximately 1300 bases in length with no hint of a larger transcript being present. Apparently the larger transcript is very short-lived and processed as predicted. These results confirm that the presence of the *cry3A* stabilization sequence results in a decrease in the rate of mRNA degradation.

Protein secretion under fermentation conditions

Even within the same species, some strains of *Bacillus* are capable of secreting significantly higher amounts of protein than others. The common *B. subtilis* strain 168 (from which PL1801*spoIIE* is derived) is considered a low secretor and not really useful in an industrial setting. However, strains of *B. subtilis* have been identified which are high secretors and are used routinely in industry for producing enzyme products. Therefore, the single-copy *aprL* constructs utilizing the consensus *amyQ* and consensus *amyQ/cry3A* tandem promoters were transferred to the chromosome of a high-secreting *B. subtilis* strain, A164Δ5 [26]. In addition, these strains were grown in a 3-l fermentor, rather than a shake flask, in order to more closely mimic the growth conditions found in large industrial fermentors. The results were quite encouraging: the consensus *amyQ/cry3A* tandem promoter appears to drive the expression of *aprL* to a level comparable to the strain containing tandem repeats of the *aprL* expression cassette (Figure 6), indicating that saturating levels of mRNA are being achieved with only a single-copy gene on the chromosome.

It is crucial that these promoters function to express other genes in single-copy as well, i.e., produce saturating levels of gene-specific mRNA. Therefore, two other genes were tested. The *amyM* gene from *Bacillus stearothermophilis* [12] (recently reclassified as a separate species, *Bacillus* sp. TS-25) and the pullulanase gene from *Bacillus deramificans* [11] were transcriptionally fused to the

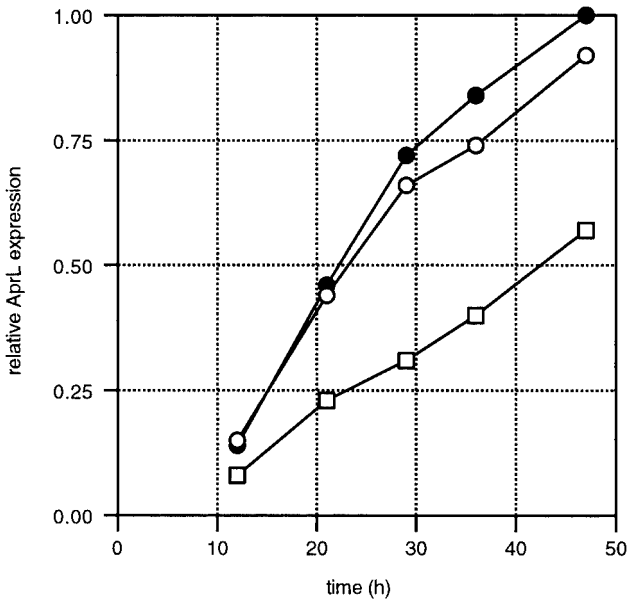

Figure 6 AprL expression profiles in 3-l fermentors. *B. subtilis* A164Δ5 integrants with *aprL* expression cassettes inserted at the *amyE* locus were grown in 3-l fermentation tanks as described in Materials and methods. AprL activity in the supernatants was determined at various time points by colorimetric assay and plotted relative to the maximum titer. The *aprL* gene was integrated in single copy under control of the consensus *amyQ* promoter (□) or consensus *amyQ/cry3A* tandem promoter (○) or in multicopy under control of the *amyL* promoter (●).

consensus *amyQ/cry3A* stab promoter and introduced as single copies into the A164Δ5 background. Both strains were tested in 3-l fermentors and expression levels were shown to be equal to those obtained from strains containing several tandem repeats of these expression cassettes (data not shown). Thus, it appears likely that these promoters will be useful for expressing many different genes in *B. subtilis*. Finally, the chloramphenicol marker was deleted from several of these high-expressing strains containing single copies of the expression cassettes by traditional gene replacement techniques [10] using plasmid pRB115. As expected, the expression levels were identical to the isogenic marked strains (data not shown).

Thus, we have achieved our goal of developing promoters which are capable of expressing single-copy genes in *B. subtilis* to levels comparable to strains containing several tandem repeats of an expression cassette. To our knowledge, this is the first report of utilizing mRNA stabilization sequences to increase expression levels of an industrial enzyme. We hope that this technology will be applicable to other *Bacillus* species. Given the fact that the mRNA stabilizer was originally derived from the distantly related *B. thuringiensis*, it seems plausible that the technology should function in other species as well. If this holds true, it is conceivable that in the near future all industrial fermentation strains of *Bacillus* can be constructed to be more environmentally friendly and marker-free.

Acknowledgements

We thank Donna Moyer and Debbie Yaver for technical advice. Also, the technical support group at Novo Nordisk Biotech contributed greatly to this work, in particular Liz Golightly and Sheryl Bernauer in DNA chemistry and Carin Morris in protein chemistry.

References

1 Adams LF, S Mathewes, P O'Hara, A Petersen and H Gürtler. 1994. Elucidation of the mechanism of CryIIIA overproduction in a mutagenized strain of *Bacillus thuringiensis* var. *tenebrionis*. *Mol Microbiol* 14: 381–389.
2 Agaise H and D Lereclus. 1994. Expression in *Bacillus subtilis* of the *Bacillus thuringiensis cryIIIA* toxin gene is not dependent on a sporulation-specific sigma factor and is increased in a spo0A mutant. *J Bacteriol* 176: 4734–4741.
3 Agaise H and D Lereclus. 1996. STAB-SD: a Shine–Dalgarno sequence in the 5′ untranslated region is a determinant of mRNA stability. *Mol Microbiol* 20: 633–643.
4 Albertini AM and A Galizzi. 1985. Amplification of chromosomal region in *Bacillus subtilis*. *J Bacteriol* 163: 1203–1211.
5 Anagnostopoulos C and J Spizizen. 1961. Requirements for transformation of *Bacillus subtilis*. *J Bacteriol* 81: 741–746.
6 Antoniewski C, B Savelli and P Stragier. 1990. The *spoIIJ* gene, which regulates early developmental steps in *Bacillus subtilis*, belongs to a class of environmentally responsive genes. *J Bacteriol* 172: 86–93.
7 Bethesda Research Laboratories. 1986. BRL pUC host: *E. coli* DH5α competent cells. *Bethesda Res Lab Focus* 8: 9.
8 Branner S, S Hastrup, N Eriksen, P Lindegaard, OH Olsen, E Casteleijn, MR Egmond, J Haverkamp, W Musters and J de Vlieg. 1996. Subtilisin mutants. United States Patent: 5,482,849.
9 Carrier TA and JD Keasling. 1997. Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. *Biotechnol Prog* 13: 699–708.
10 Chak KF, H DeLencastre, HM Liu and PJ Piggot. 1982. Facile *in vivo* transfer of mutations between the *Bacillus subtilis* chromosome and a plasmid harbouring homologous DNA. *J Gen Microbiol* 128: 2813–2816.
11 Deweer P and A Amory. 1998. Pullulanase. United States Patent: 5,721,127.
12 Diderichsen B and L Christiansen. 1988. Cloning of a maltogenic alpha-amylase from *Bacillus stearothermophilus*. *FEMS Microbiol Lett* 56: 53–60.
13 Diderichsen BK and S Jørgensen. 1997. Bacillus promoter derived from a variant of a *Bacillus licheniformis* α-amylase promoter. United States Patent: 5,698,415.
14 Grant SG, H Hessee, FR Bloom and D Hanahan. 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-resistant mutants. *Proc Natl Acad Sci USA* 87: 4645–4649.
15 Harwood CR. 1992. *Bacillus subtilis* and its relatives: molecular biological and industrial workhorses. *Trends Biotechnol* 10: 247–56.
16 Horton R, H Hunt, S Ho, J Pullen and L Pease. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77: 61–68.
17 Hue KH, SD Cohen and DH Bechhofer. 1995. A polypurine sequence that acts as a 5′ mRNA stabilizer in *Bacillus subtilis*. *J Bacteriol* 177: 3465–3471.
18 Janniere L, B Niaudet, E Pierre and SD Ehrlich. 1985. Stable gene amplification in the chromosome of *Bacillus subtilis*. *Gene* 40: 47–55.
19 Jørgensen PL, CK Hansen, GB Poulsen and B Diderichsen. 1990. *In vivo* genetic engineering: homologous recombination as a tool for plasmid construction. *Gene* 96: 37–41.
20 Mead DA, NK Pey, C Hermastadt, RA Marcil and LM Smith. 1991. A universal method for the direct cloning of PCR amplified nucleic acid. *Bio/Technology* 9: 657–663.
21 Moran CP Jr, N Lang, SFJ LeGrice, G Lee and M Stephens. 1982. Nucleotide sequences that signal the initiation of transcription and translation in *Bacillus subtilis*. *Mol Gen Genet* 186: 339–346.
22 Neu HC. 1992. The crisis in antibiotic resistance. *Science* 257: 1064–1072.
23 Palva I, RF Pettersson, N Kalkkinen, P Lehtovaara, M Sarvas, H Soederlund, K Takkinen and L Kaariainen. 1981. Nucleotide sequence of the promoter and NH_2-terminal signal peptide region of the alpha-amylase gene from *Bacillus amyloliquefaciens*. *Gene* 15: 43–51.
24 Pitcher DG, NA Saunders and RJ Owen. 1989. Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett Appl Microbiol* 8: 151–156.
25 Sambrook J, E Fritsch and T Maniatis. 1989. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26 Sloma A and L Christianson. 1999. Nucleic acids encoding a polypeptide having protease activity. United States Patent: 5,891,701.
27 Twining SS. 1984. Fluorescein isothiocyanate-labeled casein assay for proteolytic enzymes. *Anal Biochem* 143: 30–34.
28 Voskuil MI, K Voepel and G Chambliss. 1995. The −16 region, a vital sequence for the utilization of a promoter in *Bacillus subtilis* and *Escherichia coli*. *Mol Microbiol* 17: 271–279.
29 Widner W, M Thomas and A Sloma. 1999. Methods for producing a polypeptide in a *Bacillus* cell. United States Patent: 5,955,310.
30 Yanisch-Perron C, J Viera and J Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33: 103–119.
31 Young M. 1984. Gene amplification in *Bacillus subtilis*. *J Gen Microbiol* 130: 1913–1921.