

PROMOTERS FOR GENE EXPRESSION

FIELD OF THE INVENTION

This invention relates to the art of genetic engineering as applied to bacterial hosts. More particularly, the invention relates to expression systems for production of proteins in bacterial hosts.

BACKGROUND OF THE INVENTION

The advent of recombinant DNA technology has enabled the production of various naturally occurring and synthetic proteins in organisms such as bacteria, fungi, yeast and mammalian cells. In general it involves the insertion of genes that encode a desired protein into a host organism, and utilizing the host's cellular machinery to express the gene.

Recombinant DNA technology is continually developing to achieve production of proteins in commercially acceptable yields. A limiting factor in recombinant production of proteins is the rate at which the gene encoding the desired protein is expressed. In particular, it has been found that the promoter region of a gene is critical in the transcription process of gene expression. An efficient promoter such as the *trp* promoter found in *E. coli*, binds tightly to DNA-directed RNA polymerase to initiate transcription of the gene in generating mRNA. A less efficient promoter such as the *lac* promoter binds RNA polymerase less tightly, resulting in a lower rate of mRNA generation.

The *trp* promoter has been widely used in the production of heterologous proteins because of its ability to efficiently initiate transcription. Despite its efficiency, an inherent shortcoming of the *trp* promoter is that it is not easily controlled. Specifically, the *trp* promoter is not fully repressible, i.e. it can drive transcription before the host is grown in culture to a phase appropriate for protein production. Another widely used promoter is *lac* which is less efficient than *trp*, but is more controllable.

To develop more efficient promoters, functional components of different promoters have been combined, for instance those described in United States patent US 5,362,646. In one example, portions of the bacteriophage T7 promoter A₁ (P_{A1}) were combined with two *lac* operators. Specifically, the spacer region between the so called -35 and -10 regions of the bacteriophage T7 promoter was replaced with a modified *lac* operator sequence, and to control the resulting promoter hybrid, a second *lac* operator was introduced downstream. The resulting promoter/operator system which is incorporated on the commercially available pUHE plasmids was found to initiate transcription efficiently upon induction and yet is highly repressed before induction.

Another promoter described by Tsung et al (Proc. Natl. Acad. Sci. USA, 1990, 87:5940) comprises the efficient *trp* -35 region, the -10 region from the highly efficient *lpp*P-5 promoter (a variant of *lpp* promoter) and a spacer derived from the *lac* promoter. This promoter was shown to be so highly efficient in initiating transcription as to result in cell lethality.

While various promoters have allowed improved yields of proteins in microbial hosts, there still remains a need for promoters that drive production of commercially valued proteins more efficiently.

SUMMARY OF THE INVENTION

According to an aspect of the present invention, there is provided a novel recombinant DNA construct useful for expressing a protein in a bacterial host. The construct comprises a coding region for a protein and linked operably therewith, a control region comprising a promoter having a DNA sequence selected from:

5'-TTGACAACATAAAAACTTTGTGTTATACT-3'; and
5'-TTGACACTTTATGCTTCCGGCTCGTATACT-3'.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates the nucleotide sequence and a schematic representation of an expression cassette incorporating recombinant DNA constructs in accordance with the invention;

- 5 Figures 2 and 3 illustrate nucleotide sequences of DNA constructs according to the invention comprising a promoter and operator region.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides DNA sequences useful in driving DNA expression with high efficiency in bacterial hosts such as *E. coli*. The use of expression vectors comprising these sequences provides a valuable means for achieving increased production of expressible proteins, both endogenous and heterologous. In one aspect of the invention, there is provided a novel recombinant DNA construct useful for expressing a protein in a bacterial host. The construct comprises a coding region for the protein operably linked with a control region comprising a promoter to enable expression of said protein in the host, wherein the promoter comprises a DNA sequence selected from:

5'-TTGACAACATAAAAACTTTGTGTTATACT-3'; and

5'-TTGACACTTTATGCTTCCGGCTCGTATACT-3'.

- 20 These promoters have in common a -35 region of the consensus sequence TTGACA and a -10 region of the sequence TATACT. The spacer sequence, i.e. the sequence of 18 bases which is intervening, can according to the invention be either the sequence ACATAAAAACTTTGTGT or CTTTATGCTTCCGGCTCG and is preferably
- 25 the sequence ACATAAAAACTTTGTGT. Accordingly, the invention provides in a preferred embodiment, DNA constructs in which DNA coding for a desired protein is linked operably under expression control to a promoter of the sequence 5'-TTGACAACATAAAAACTTTGTGTTATACT-3'.

Those skilled in the art will appreciate that the promoters of the invention constitute one essential element of the components required, within the functional control region, to drive expression. Moreover, these promoters can be inserted using standard procedures into any suitable expression
5 vector which can replicate in Gram -ve or +ve bacteria. More particularly, and to form a gene expression control region, the present promoters will be incorporated with such other control elements as are typically required for that expression, including a ribosome binding site and in embodiments of the invention, an operator that functions to control promoter function.

10 These components are necessarily arranged relative to each other as required for expression to occur according to well understood principles of gene expression.

In an embodiment of the invention, the control region of the construct incorporates an operator. Operators that can be used include those
15 directly inducible by chemical inducers and those, such as *lac* which are de-repressible. Examples of suitable operators include the *E. coli* lactose, galactose, tryptophan and tetracycline operators (see Miller et al "The Operon", Cold Spring Harbour Laboratory, 1980 and Hillen et al, J. Mol. Biol., 1984, 172:185). Preferred operators are highly repressible so that
20 expression of DNA coding for the protein can be controlled. In a specific embodiment, the control region comprises the *lac* operator (fig. 1) which prevents expression from the promoter in the absence of the synthetic inducer isopropyl- β -D-thiogalactopyranoside (IPTG) and the natural inducer lactose.

25 The control region further comprises a ribosome binding site (RBS) sequence to facilitate the binding of ribosomes to the mRNA transcript and thereby initiate the translation of the RNA coding region to generate the protein. Suitable ribosome binding sites include *lac* and bacteriophage T5.

In a preferred embodiment the RBS is a sequence derived from the T5 bacteriophage RBS having the sequence 5'-ATTAAAGAGGAGAAATTAAGC-3'.

5 The control region of constructs according to the present invention are operably linked with a coding region for an endogenous or heterologous protein. The term "heterologous protein" refers to a polypeptide or protein which, although not naturally produced by the bacterial host, is expressed by this host when suitably transformed with DNA coding for the protein; genomic DNA, cDNA and synthetic DNA can be used to transform the
10 host. Proteins that may be produced using the system herein described include, but are not limited to, hormones such as parathyroid hormone (PTH), glucagon or fragments thereof and related peptides such as GLP-1 and GLP-2; growth factors such as epidermal growth factor (EGF); and lymphokines such as interleukin-6 and -8 (IL-6, -8). To facilitate isolation
15 of the authentic form of the protein, i.e. protein without an additional N-terminal Met residue, fusion proteins may also be produced which are cleaved subsequent to expression. For example, DNA encoding a protein of interest may be preceded by DNA encoding a signal peptide, such as the *E. coli* outer membrane protein ompA. In this instance the expressed
20 gene yields a fusion protein comprising a N-Met-containing ompA signal peptide which is followed by the desired protein. The signal peptide carries the fusion protein through the intermembrane of the bacterium where the signal peptide is cleaved. Other signal peptides which may be used include alkaline phosphatase, *E. coli* heat stable enterotoxin II and
25 protein A from Streptococcus. Alternatively, a fusion protein may be synthesized and cleaved in a separate procedure to yield the desired protein. For example glutathione-S-transferase (GST) may be cleaved from a desired protein with thrombin or factor Xa.

In a specific embodiment of the invention, the coding region comprises

DNA encoding human PTH, the amino acid sequence of which is described by Hendy et al (Proc. Natl. Acad. Sci. USA, 1981, 78:7365). In the examples herein described, the DNA sequence coding for PTH is coupled directly to and in reading frame with DNA coding for the ompA signal peptide.

The preferred recombinant DNA constructs illustrated in Figure 1, having the lpp or lacUV5 spacer, were produced from single-stranded oligonucleotides synthesized by the phosphoramidite method. The gel-purified strand comprising the sequences from the XhoI to EcoRI restriction site was then used as an initial PCR target and was PCR amplified into a double stranded DNA fragment using complementary single stranded DNA oligonucleotides which hybridized specifically to the ends of either of the initial oligonucleotide sequence shown or its complementary strand. Thus, the constructs are prepared using standard gene synthesis methodology, as described for example by Maniatis et al ("Molecular Cloning" Cold Spring Harbour Laboratories, 1982) and Innis et al ("PCR Protocols, A Guide to Methods and Applications").

In another aspect of the invention there is provided expression vectors useful for producing bacterial host cell transformants which incorporate a recombinant DNA construct according to the invention. DNA constructs according to the invention may be incorporated as a "cassette" into a vector, preferably a plasmid vector, by established techniques. Generally, a vector is cleaved at restriction sites that correspond with restriction sites on either end on the cassette. The cassette is then introduced by ligating the ends to the complementary cleaved sites on the vector.

Although bacteriophage-based vectors can be used, plasmid vectors are preferred such as the pBR322 and pUC series of plasmids. Once incorporated into a suitable vector, the resulting plasmid may be amplified

in a host to provide amounts sufficient for subsequent cloning work. It will be appreciated that DNA coding for the selected protein is conveniently incorporated on the plasmid with multiple cloning sites provided thereon, using standard cloning/ligation methods. Also, a plasmid will necessarily incorporate an origin of replication and most desirably will incorporate a marker such as the ampicillin or tetracycline resistance genes to allow the selection of transformed cells. It will also be appreciated that translational stop codons in all three reading frames and a correctly positioned transcriptional termination region will be required for satisfactory expression of the desired protein. Suitable transcriptional terminators include the transcriptional terminators of the E. coli trpA, thr, his and phe genes.

Once DNA coding for the desired protein is incorporated on the vector, a selected bacterial host is transformed therewith using standard calcium chloride mediated transformation techniques. Suitable bacterial hosts include gram negative bacteria such as E. coli and Salmonella. Preferably the host is a commercially available E. coli strain and most preferably JM101 and derivatives thereof.

When the controlling region of the DNA construct comprises the *lac* operator, as described in more detail hereinafter, the transformed host strain should be capable of expressing, and desirably in some cases over-producing, the *lacI* product so that promoter function and hence expression of the protein, can be regulated. The need for *lacI* overproduction by some transformants can be met, according to one embodiment of the invention, by using hosts that already harbour the *lacI^q* gene responsible for overproduction of the *lacI* product. *LacI* over-producing strains of E. coli that may be employed as hosts include the JM series of strains available from Clontech Laboratories Inc., California, USA. Specific host strains suitable for use include JM101, JM105 and JM107.

Alternatively, *lacI* overproduction in the transformant may be met by incorporating the *lacI*^q gene on vectors of the invention. Since, in this situation, the over-production of *lacI* is mediated by the vector, any of a variety of commercially available bacterial host strains may be employed, including *E. coli* strains DH1, RR1, C600, CMK603 and EB505. The *lacI*^q gene to be incorporated on the vector may be obtained as a 1.2 kb HindIII fragment of plasmid pMMB22 (described by Bagdasarian et al (Gene, 1983, 26:273) and then incorporated non-disruptively at any site on the plasmid vector.

10 To enhance the stability of inheritance of vectors, in particular plasmids, from the strain originally transformed to its progeny, a partition element (*par*) functional in *E. coli* may also be incorporated on the vector. One such *par* element may be liberated from pSC101 as a 380bp HincII/AvaI fragment and then cloned into a suitable site on the vector.

15 Following transformation, bacterial hosts harbouring the expression vector are cultured in a culturing medium most appropriate for the selected host. For *E. coli*, LB broth or 2YT medium (yeast extract/tryptone) can be used to culture those strains herein preferred. Selective pressure for plasmid transformants should be maintained by providing a cytotoxic agent which kills the untransformed host strain. For example, a transformant with a plasmid harbouring the gene for tetracycline resistance should be cultured in medium containing tetracycline. Medium concentrations of tetracycline around 5-15 µg/mL are suitable.

25 The promoter on the construct is preferably regulatable through binding of a repressor molecule to an operator located adjacent to the promoter in the control region. In a preferred embodiment, the *lacI* gene product binds to a *lac* operator located adjacent the promoter. In this instance, binding of *lacI* product represses the promoter, lowering expression levels of coding

DNA under its control. To raise expression levels, the chemical IPTG (isopropyl- β -D-thiogalactopyranoside), which binds the *lacI* and derepresses the promoter, is added to the culture medium to derepress the promoter and induce expression. Suitably, IPTG is added to the culture medium when the cells have reached mid log growth phase.

To determine the optimum density to which cultures should be grown to realize maximum yield of the desired protein, trials can be conducted and protein levels assayed in a time-course experiment. In general, reasonable yields of protein may be recovered once cells reach mid log phase, although greater amounts of protein can be expected to accumulate within about 4-5 hours thereafter.

The desired protein can be purified by techniques established in the art as being appropriate for that protein. In a specific embodiment of the invention, expressed PTH is excreted beyond the periplasmic space and into the culture medium where it is recovered directly. When protein is excreted, the spent medium can be isolated using biochemical techniques that reflect the nature of the protein in terms of its molecular size, net charge, isoelectric point, etc. The medium may be first concentrated, for example by lyophilization. Further, when antibodies are available or a natural ligand for the protein is available, affinity columns may be used. Specific embodiments of the invention are now exemplified with reference to the drawings.

EXAMPLE 1

In its mature form, PTH is an 84-amino acid peptide that acts in humans to raise blood calcium and modulate bone formation. DNA coding for human PTH was synthesized using the established techniques and according to the amino acid sequence published by Hendy et al., *supra*.

and incorporated within the constructs described below, as shown in Figure 1.

Preferred recombinant DNA constructs incorporating promoter #1 and #2 as well as reference promoters #3 and #4, illustrated in table 1 and figure 1, to which reference is now made were produced from a single-stranded oligonucleotide synthesized by the phosphoramidite method. The gel-purified strand comprising the sequences from the XhoI to EcoRI restriction site was then used as a initial PCR target and was PCR amplified into a double stranded DNA fragment using complementary single stranded DNA oligonucleotides which hybridized specifically to the ends of either the initial oligonucleotide sequence shown or its complementary strand. Thus, the constructs are prepared using standard gene synthesis methodology, as described for example by Maniatis ("Molecular Cloning" Cold Spring Harbour Laboratories, 1982) and Innis et al ("PCR Protocols, A Guide to Methods and Applications").

The constructs were then cloned into a pUC18 derived plasmid which confers tetracycline resistance in place of ampicillin resistance. A JM101 derived *E. coli* host strain was then transfected according established techniques (see Maniatus et al "Molecular Cloning", Cold Spring Harbour Laboratory, 1982)

EXAMPLE 2 Expression of Transformed host

The transformants containing the PTH vectors were cultured overnight at 30°C in 2YT broth containing 0.5% glucose and tetracycline and then inoculated into fresh medium of the same composition, with continued culturing at 30°C until reaching mid log phase. Cultures were then induced (1mM IPTG) at 1 hour growth intervals, aliquots of culture were withdrawn and fractionated to produce samples of culture medium to identify excreted PTH products using a standard Allegro assay. The

results of these assays are provided in Table 1 below:

Table 1

#	Promoter			RBS	Max PTH (mg/L) 6-8 hrs
	-35 region	spacer	-10 region		
1	trp	<i>lpp</i>	<i>lpp</i> P-5	T5 <i>lac</i>	245 148
2	trp	<i>lac</i> UV5	<i>lpp</i> P-5	T5 <i>lac</i>	121 10
3	trp	<i>lac</i> O	<i>lpp</i> P-5	T5 <i>lac</i>	50 5
4	T7	<i>lac</i> O	T7	T5 <i>lac</i>	100 10

Results of the Allegro assay indicate that the promoters incorporating the 18 bp *lpp* and *lac*UV5 sequences facilitate enhanced levels of heterologous PTH protein. Promoter #1 and #2 compare favourably to promoter #3 wherein the spacer was substituted with a modified *lac* operator sequence (*lac*O); and promoter #4 wherein the -35 and -10 regions are from bacteriophage T7 and the spacer is *lac*O promoter. Studies with the same promoters expressing the gene coding for chloramphenicol acetyl transferase (CAT) showed similar results of enhanced expression for promoters #1 and #2. Also, it was noted that each of the promoters studied exhibited enhanced protein yield when combined with the bacteriophage T5 RBS in comparison to the *lac* derived RBS.



WE CLAIM

1. A recombinant DNA construct useful for expressing a protein in a bacterial host, the construct comprising a coding region for the protein operably linked with a control region comprising a promoter to enable expression of said protein in the host, wherein the promoter region comprises a DNA sequence selected from:
- 5'-TTGACAACATAAAAACTTTGTGTTATACT-3'; and
5'-TTGACACTTTATGCTTCCGGCTCGTATACT-3'.
2. A recombinant DNA construct according to claim 1, wherein said promoter region comprises the DNA sequence
5'-TTGACAACATAAAAACTTTGTGTTATACT-3'.
3. A recombinant DNA construct according to claim 1, wherein said bacterial host is *E. coli*.
4. A recombinant DNA construct according to claim 3, wherein said control region further comprises an operator for regulating expression of the protein.
5. A recombinant DNA construct according to claim 4, wherein the operator is the *lac* operator.
6. A recombinant DNA construct according to claim 5, wherein the protein is human PTH.
7. A recombinant DNA construct according to claim 6, further comprising DNA coding for the OmpA signal peptide in reading frame with the coding region to enable secretion of the protein to the periplasm of the host.





WO 96/36721

PCT/IB96/00462

13

8. A recombinant DNA construct according to claim 3, wherein the control region comprises a T5 ribosome binding site having the DNA sequence **ATTAAAGAGGAGAAATTAAGC**.
- 5 9. A recombinant DNA construct according to claim 3, having a sequence as defined in figure 2.
10. A recombinant DNA construct according to claim 3, having a sequence as defined in figure 3.
- 10 11. A vector useful for producing bacterial host cell transformants comprising a recombinant DNA construct according to any preceding claim.

This Sheet Has Been Replaced

