(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 July 2002 (04.07.2002)

PCT

(10) International Publication Number
WO 02/051238 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/US01/49977

(22) International Filing Date:
21 December 2001 (21.12.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/258,291 26 December 2000 (26.12.2000) US

(71) Applicant *(for all designated States except US)*: **MONSANTO TECHNOLOGY LLC** [US/US]; HUELSKAMP, Nancy, 800 N. Lindbergh Blvd., E2NA, St. Louis, MO 63167 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BOGOSIAN, Gregg** [US/US]; 374 Herworth Court, Clarkson Valley, MO 63005 (US). **O'NEIL, Julia** [US/US]; 8747 White Avenue, St. Louis, MO 63144 (US). **AARDEMA, Noelle** [US/US]; 2529 Orange Avenue, Costa Mesa, CA 92627 (US).

(74) Agent: **BECK, George**; Monsanto Company, 800 N. Lindbergh Blvd., E2NA, St. Louis, MO 63167 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/051238 A2

(54) Title: RECOMBINANT DNA VECTORS FOR EXPRESSION OF SOMATOTROPINS

(57) Abstract: A family of somatotropin vectors with which micro-organisms, preferably bacteria such as, but not limited to E. coli, can be transformed to enable the expression of bovine somatotropin at high levels using conventional fermentation and induction conditions.

RECOMBINANT DNA VECTORS FOR EXPRESSION OF SOMATOTROPINS

Cross-Reference to Related Applications

This application claims priority under 35 U.S.C. 119(e) of Application No. 60/258,291 filed December 26, 2000, the entirety of which is herein incorporated by reference.

Background of the Invention

In mammals, growth hormones (somatotropins) are initially expressed in a pre-hormone form comprising a leader sequence which is removed during secretion of the mature, biologically active hormone from pituitary cells. As used herein, "mature" somatotropin ("ST") means having an amino acid length essentially like those of the ST secreted into the bloodstream of the animal, e.g., N-Ala-Phe-Pro- or N-Phe-Pro-bovine ST, N-Phe-Pro-human, -porcine (identical to canine) or -equine ST, etc.

Attempts to bacterially express human and bovine STs in non-secretion systems using structural genes having the sequences of their cDNAs were at first unsuccessful. Researchers succeeded after introducing silent mutations into the front end of the structural gene. P.H. Seeburg et al., attributed the original difficulties to translation being impeded by secondary structure of the mRNA corresponding to cDNA, and taught lessening such secondary structure to enable significant expression. DNA, Vol. 2, No. 1, 1983, pp.37-45; U.S. Patents 5,254,463 and 5,260,201.

It has been possible to obtain commercially viable ST expression levels using DNA containing certain silent mutations. However, the need for new vectors that are highly productive and less expensive to use has been difficult to satisfy because the relationships between the configurations and expression levels of "mature" somatotropin vectors remain poorly understood and highly unpredictable.

Summary of the Invention

Described herein is a family of ST vectors with which micro-organisms, preferably bacteria such as, but not limited to E. coli, can be transformed to enable the expression of bovine ST ("bST") at high levels using conventional fermentation and induction conditions. This family is exemplified by the 44 vectors listed in Table I which have SEQ. NOS. 1 through 44. Each sequence extends from an EcoRI site (GAATTC) before the promoter region, through the promoter, ribosome binding site ("RBS") and entire bST structural gene, ending with the translation stop codon for the structural gene.

Detailed Description of the Invention

Each of the vectors having Seq. Nos. 1-44 includes a synthetic promoter, herein designated "cpex-20", which is disclosed in PCT Appln. PCT/US00/40014 filed March 31, 2000, which claims priority of U.S. Provisional Patent Appln. 60/127,449, filed April 1, 1999. This promoter is situated between the EcoRI site (GAATTC) and the AscI site (GGCGCGCC). Various other conventional and novel promoters can be used in these vectors in lieu of cpex-20 to achieve good levels of ST expression.

Seq. Nos. 1-17 have an identical RBS, herein designated "dps", that is native to E. coli. Seq. Nos. 18-29 share a novel synthetic RBS which is herein designated "L49". Seq. Nos. 30-42 share a second novel synthetic RBS herein designated "L437". Seq. Nos. 43-44 share a third novel synthetic RBS herein designated "R806". Each RBS is situated between the AscI site and the ATG translation start codon of the bST structural gene. The AscI site begins at coordinate 76 of each sequence. The ATG start codon begins at approximately coordinate 109; this can vary slightly due to the RBSs employed having slightly different lengths.

All of the vectors having SEQ. NOS. 1-44 have structural genes that differ from cDNA encoding "mature" bST by containing various arrays of silent changes within the nucleotides encoding the first 16 codons of bST. As shown in Figure I, vector pXT709 (SEQ. NO. 16) differs from pXT674 (SEQ. NO. 8) by having an additional 12 codons within the structural gene changed from those of cDNA, 11 to E. coli-preferred codons and a twelfth to eliminate an undesirable restriction site, by changing the translation stop codon to two tandem stop codons, and by substituting a different terminator which has two lacUV5 promoters in tandem. pXT703 (SEQ. NO. 15) and pXT747 (SEQ. NO. 17) differ from pXT601 (SEQ. NO. 2) and pXT686 (SEQ. NO. 13), respectively, in the same way.

Each of the vectors having SEQ. NOS. 1 through 44 have expressed mature bST at unoptimized levels of at least about 3.5 grams/liter ("gm/l"), most of them above 4.5 gm/l, and some at 7 gm/l or higher, using the fermentation conditions described in Bogosian et al., J. Biol. Chem., vol. 264, pp.531-39, January 5, 1989, except that 50 ppm of nalidixic acid was used as the inducer instead of indole acrylic acid.

The expression levels of mature ST vectors are very unpredictable from published teachings, e.g., P.H. Seeburg et al. In fact, the vectors assigned SEQ. NOS. 1 through 17, 30 through 35, and 37 through 42 have been found, using calculations as recommended in Seeburg et al., to have mRNA secondary structure greater than, rather than lessened from that of bST cDNA.

We claim:

1. A vector useful for expressing mature bovine somatotropin in E. coli and comprising a DNA sequence selected from SEQ. NOS. 1 through 44.

2. A vector of Claim 1 having a DNA sequence selected from SEQ. NOS. 1 through 17, 30 through 35 and 37 through 42.

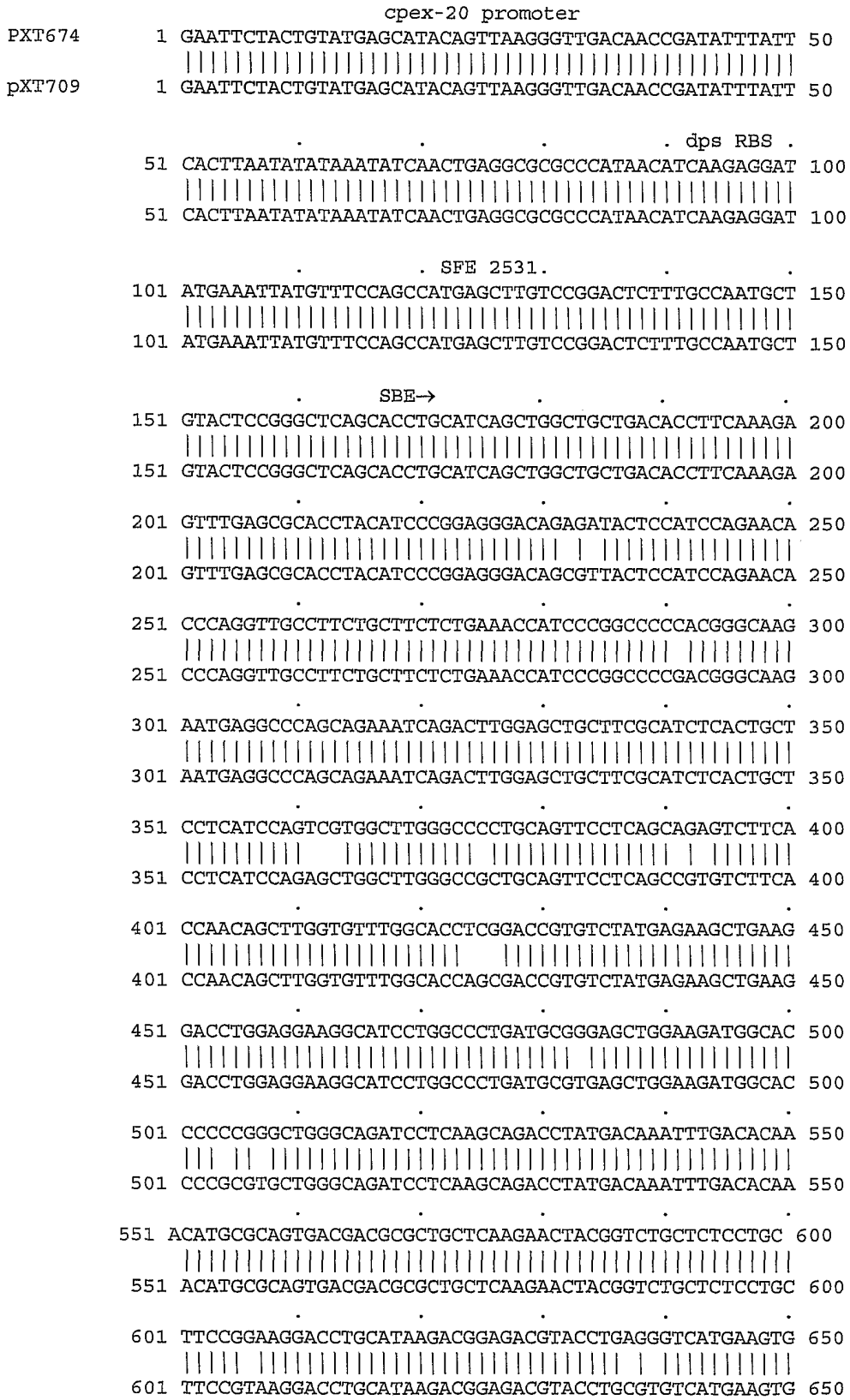
3. A vector of Claim 1 having a DNA sequence essentially like that of pXT709 (SEQ. NO. 16).

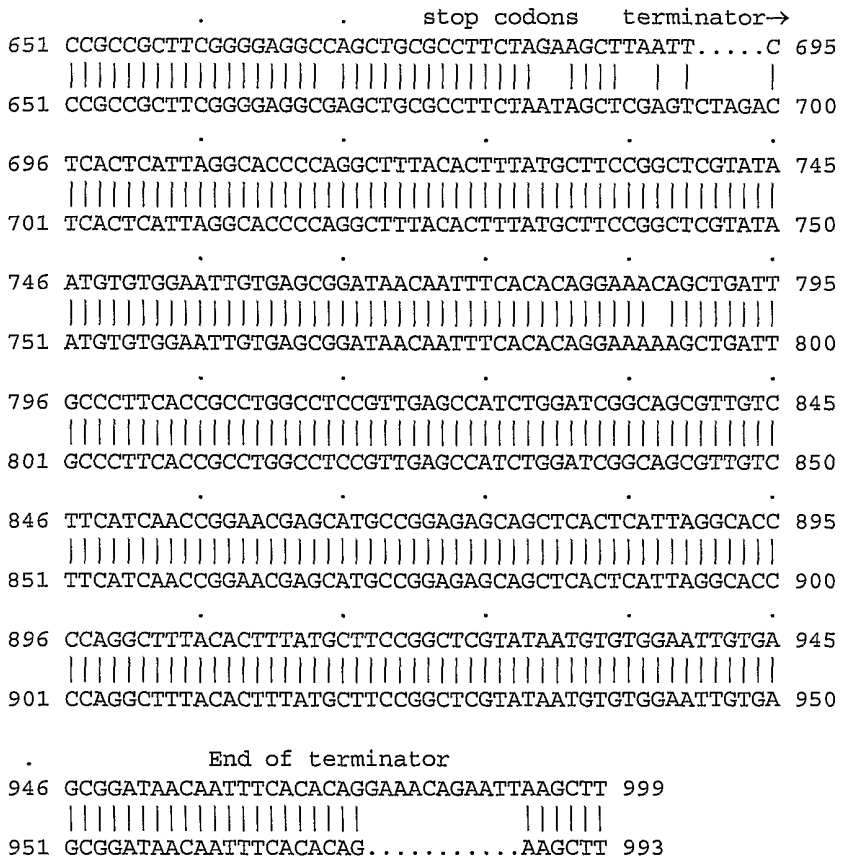
4. E. coli transformed with a vector of Claim 1, 2 or 3.

5. A process for making mature bovine somatotropin comprising fermenting an E. coli of Claim 4 and inducing expression of the somatotropin.

6. DNA that is functional as an E. coli ribosome binding site, comprising the sequence identified herein as L49, L437 or R806, or a complement thereof.

Figure I


cpex-20 promoter
PXT674 1 GAATTCTACTGTATGAGCATACAGTTAAGGGTTGACAACCGATATTTATT 50
pXT709 1 GAATTCTACTGTATGAGCATACAGTTAAGGGTTGACAACCGATATTTATT 50
dps RBS
51 CACTTAATATATAAATATCAACTGAGGCGCGCCCATAACATCAAGAGGAT 100
51 CACTTAATATATAAATATCAACTGAGGCGCGCCCATAACATCAAGAGGAT 100
SFE 2531.
101 ATGAAATTATGTTTCCAGCCATGAGCTTGTCCGGACTCTTTGCCAATGCT 150
101 ATGAAATTATGTTTCCAGCCATGAGCTTGTCCGGACTCTTTGCCAATGCT 150
SBE→
151 GTACTCCGGGCTCAGCACCTGCATCAGCTGGCTGCTGACACCTTCAAAGA 200
151 GTACTCCGGGCTCAGCACCTGCATCAGCTGGCTGCTGACACCTTCAAAGA 200
201 GTTTGAGCGCACCTACATCCCGGAGGGACAGAGATACTCCATCCAGAACA 250
201 GTTTGAGCGCACCTACATCCCGGAGGGACAGCGTTACTCCATCCAGAACA 250
251 CCCAGGTTGCCTTCTGCTTCTCTGAAACCATCCCGGCCCCCACGGGCAAG 300
251 CCCAGGTTGCCTTCTGCTTCTCTGAAACCATCCCGGCCCCGACGGGCAAG 300
301 AATGAGGCCCAGCAGAAATCAGACTTGGAGCTGCTTCGCATCTCACTGCT 350
301 AATGAGGCCCAGCAGAAATCAGACTTGGAGCTGCTTCGCATCTCACTGCT 350
351 CCTCATCCAGTCGTGGCTTGGGCCCCTGCAGTTCCTCAGCAGAGTCTTCA 400
351 CCTCATCCAGAGCTGGCTTGGGCCGCTGCAGTTCCTCAGCCGTGTCTTCA 400
401 CCAACAGCTTGGTGTTTGGCACCTCGGACCGTGTCTATGAGAAGCTGAAG 450
401 CCAACAGCTTGGTGTTTGGCACCAGCGACCGTGTCTATGAGAAGCTGAAG 450
451 GACCTGGAGGAAGGCATCCTGGCCCTGATGCGGGAGCTGGAAGATGGCAC 500
451 GACCTGGAGGAAGGCATCCTGGCCCTGATGCGTGAGCTGGAAGATGGCAC 500
501 CCCCCGGGCTGGGCAGATCCTCAAGCAGACCTATGACAAATTTGACACAA 550
501 CCCGCGTGCTGGGCAGATCCTCAAGCAGACCTATGACAAATTTGACACAA 550
551 ACATGCGCAGTGACGACGCGCTGCTCAAGAACTACGGTCTGCTCTCCTGC 600
551 ACATGCGCAGTGACGACGCGCTGCTCAAGAACTACGGTCTGCTCTCCTGC 600
601 TTCCGGAAGGACCTGCATAAGACGGAGACGTACCTGAGGGTCATGAAGTG 650
601 TTCCGTAAGGACCTGCATAAGACGGAGACGTACCTGCGTGTCATGAAGTG 650


stop codons terminator→
End of terminator