MUTANT T7 RNA POLYMERASE
GP1(LYS222) EXHIBITING ALTERED
PROMOTER RECOGNITION

Inventor: Richard A. Ikeda, Atlanta, Ga.
Assignee: Georgia Tech Research Corporation, Atlanta, Ga.

Appl. No.: 106,433
Filed: Aug. 13, 1993

Int. Cl. C07H 21/00; C12N 9/12; C12N 15/54
U.S. Cl. 435/172.3; 435/194; 536/23.2; 935/31; 935/41

Field of Search 435/69.1, 172.3, 320.1, 435/194, 252.33; 536/23.2; 935/31, 29, 41

References Cited


ABSTRACT

E. coli harboring the mutant plasmid pKGP-HAlmut4 and an inactive pCM-X# are chloramphenicol resistant and that the mutation responsible for the expression of CAT from the inactive pCM-X# plasmid is a G to A transition at nucleotide 664 of T7 gene 1 that converts glutamic acid (222) to lysine. This mutation expands the range of T7 promoter sequences that can be utilized by the enzyme. The mutant T7 RNA polymerase, GP1(lys222), utilizes inactive T7 promoter point mutants more efficiently than wild-type T7 RNA polymerase both in vivo and in vitro. Furthermore, the correlation of in vivo and in vitro promoter utilization suggests that the restoration of chloramphenicol resistance in the cotransformed E. coli results from the ability of GP1(lys222) to initiate transcription from T7 promoter point mutants that are normally inactive.

4 Claims, 8 Drawing Sheets
PUBLICATIONS
Shaw, W. V., Chloramphenicol Acetyltransferase from Chloramphenicol–Resistant Bacteria, Antibiotic Inactivation and Modification, p. 737–775.
FIG. 1
FIG. 3
Start T7 gene 1 (T7 3170) 

- AlwNI 753 
- BstXI 1141 

End T7 gene 1 (T7 5820) 

Mutation: ATGCTCATTGAGTCAACCGGAATGGTTAGCTTACACCGCCAAAATGCTGGC
Wild Type: MetLeuIleGluSerThrGlyMetValSerLeuHisArgGlnAsnAlaGly
Mutation: Lys

FIG. 4
1

MUTANT T7 RNA POLYMERASE GPl(LYS222) EXHIBITING ALTERED PROMOTER RECOGNITION

STATEMENT OF GOVERNMENT INTEREST

This work was supported by a FIRST grant (A124905) from the National Institutes of Allergies and Infectious Diseases. Additionally, the major equipment used in this work was part of a Biological Instrumentation Facility that was assembled with support from the National Science Foundation under Grant No. DIR-9011409. The U.S. Government may have certain rights to the invention.

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention generally relates to the alteration of a first plasmid to produce a T7 RNA polymerase capable of recognizing a T7 promoter on a second plasmid and transcribing a gene that is cloned behind the promoter resulting in changed properties of an *E. coli* in which the two plasmids are harbored. This invention specifically relates to conferring chloramphenicol ("cam") resistance to *E. coli* harboring a pKGP-HA1mut4 plasmid producing the T7 RNA polymerase GPl(lys222) and a pCM-X# plasmid, specifically those selection plasmids listed in Table I and Table II.

2. Prior Art

Bacteriophage T7 RNA polymerase, the product of T7 gene 1, is a protein produced early in T7 infection; it is a single-chain enzyme with a molecular weight close to 100,000. It appears that the basis for the selectivity of the T7 RNA polymerase is the interaction of the RNA polymerase with a relatively large promoter sequence, a sequence large enough that it is unlikely to be found by chance in any unrelated DNA. In the case of T7, the highly conserved promoter sequence appears to consist of approximately 23 continuous base pairs, which includes the start site for the RNA chain. If exact specification of even as few as 15 of these base pairs were required for initiation of chains, chance occurrence of a functional promoter would be expected less than once in a billion nucleotides of DNA.


We recently described two compatible plasmids that together can be used to determine whether a mutant T7 promoter is active or inactive in vivo [Ikeda, R. A., et al., *Biochemistry, 31*: 9073–9080 (1992); Ikeda, R. A., *J. Biol. Chem., 267*: 2517–2524 (1992)]. The first plasmid, pKGP1-l, is a pACYC177 [Chang, A. C. Y. and Cohen, S. N., *J. Bacteriol., 154*: 1141–1155 (1987)] derivative that carries T7 gene 1 (the gene encoding T7 RNA polymerase) ligated to a tac promoter [deBoer, H. A., et al., *Proc. Natl. Acad. Sci. USA, 80*: 21–25 (1983); deBoer, H. A., et al., *Promoters, Structure and Function, 462–481 (1982)], while the second plasmid, pCM-X#, is a pKK232-8 [Brosius, J., and Lupski, J. R., *Methods in Enzymology, 153*: 54–68 (1987); Brosius J., and Holy, A., *Proc. Natl. Acad. Sci. USA, 81*: 6929–6933 (1984)] derivative that carries the gene encoding CAT ligated to potential T7 promoters. pCM-X# is the general designation for this family of plasmids derived from pKK232-8. A specific plasmid within this family is designated with a letter and a number in place of X#. The following abbreviations are used throughout this specification: An, absorbance at the designated wavelength (x) in nm; amp, ampicillin; bla, β-lactamase; BSA, bovine serum albumin; CAT, chloramphenicol acetyl transferase; cam, chloramphenicol; CoA, coenzyme A; DTNB, 5,5′-dithio-bis-(2-nitrobenzoic acid); DTT, dithiothreitol; EDTA, ethylenediamine tetraacetic acid; IPTG, isopropyl-β-D-thiogalactopyranoside; kan, kanamycin; LB, Luria-Bertani (medium); NTP, nucleoside triphosphate; Tris, tris (hydroxymethyl) aminomethane; u, units. *E. coli* harboring these two plasmids are cam resistant if the pCM-X# plasmid carries an active T7 promoter and are cam sensitive if the pCM-X# plasmid carries an inactive T7 promoter. The pCM-X# plasmids that carry T7 promoter point mutants that destroy promoter activity are designated inactive pCM-X# plasmids, while pCM-X# plasmids that carry T7 promoter point mutants with moderate activity or wild-type activity are designated intermediate pCM-X# plasmids and strong pCM-X# plasmids, respectively. Point mutations that were found to inactive the T7 promoter are a Cytidine ("C") to Adenosine ("A") (plasmid pCM-P1031) or Guanosine ("G") (plasmid pCM-
chloramphenicol resistance. The selection of mutants of 65 matic representation of the two plasmid selection sys-

Methods to analyze the selection of mutants of T7 RNA polymerase that exhibit altered promoter rec- tern. Abbreviations: kan, kanamycin resistance gene; cat, chloramphenicol gene that is cloned behind the promoter. When mater activity (termed an inactive pCM-X#), the T7 RNA polymerase will not utilize the T7 promoter at -9, RNA polymerase will not utilize the T7 promoter point moters.

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

Further characterization of promoter recognition and utilization by T7 RNA polymerase would be greatly pCM-T270 (DNA sequence) pCM-T286 (DNA sequence) pCM-T297 (DNA sequence)...

SUMMARY OF SEQUENCE LISTINGS

Various DNA sequences and protein sequences are referred to throughout this specification. Following is a chart of the sequence listings and their sequence identification number as contained on pages S1–S86. Sequence ID numbers 1 and 2 correspond to the invention disclosed and claimed herein. Sequence ID numbers 3–23 correspond to other DNA sequences which are discussed herein.

SEQUENCE IDENTITY

T7 RNA polymerase GPl(Lys222) (DNA sequence of the mutant RNA polymerase) T7 RNA polymerase GPl(Lys222) (Protein sequence of the mutant RNA polymerase) pK2232-8 (DNA sequence) pCM-X# (DNA sequence) pCAT10-1 (DNA sequence) pCM-T297 (DNA sequence) pCM-P1160 (DNA sequence) pCM-T270 (DNA sequence) pCM-P1087 (DNA sequence) pCM-P1198 (DNA sequence) pCM-T286 (DNA sequence) pCM-B64 (DNA sequence) pCM-P1208 (DNA sequence) pCM-P1031 (DNA sequence) pCM-T221 (DNA sequence) Wild type T7 promoter (DNA sequence) WT (DNA sequence in Table IV) B (DNA sequence in Table IV) T (DNA sequence in Table IV) P (DNA sequence in Table IV) C (DNA sequence in Table IV) G (DNA sequence in Table IV) Primer (DNA sequence in Table IV)...

BRIEF DESCRIPTION OF THE FIGURES

FIG. 1: The Consensus T7 Promoter. The base pairs at -9, -10 and -11 have been implicated in the ability of T7 RNA polymerase to distinguish T7 and T3 promoters.

FIG. 2: Plasmids pKGP1-1 and pCM-X#. A schematic representation of the two plasmid selection systems. At altered: kan, kanamycin resistance gene; amp, ampicillin resistance gene; cat, chloramphenicol resistance gene; p15A ori, p15A origin of replication;
pBR ori, colE1 origin of replication; rnmBT1T2/5S, transcriptional terminators from the E. coli 5S rRNA gene; tac, promoter; mut T7 φ10, wild-type or mutant T7 φ10 promoter; T7 gene 1, gene encoding T7 RNA polymerase.

FIG. 3: Localization of the Mutation in pKGP-HAl-mut4. T7 gene 1 restriction fragments from pKGP-HAlmut4 were cloned into the corresponding sites of wild-type pKGP1-1. The restriction enzymes used to generate each set of clones are listed above each arrow. The newly constructed recombinant plasmids are shown on the left side of the arrow, and the T7 gene 1 restriction fragments ligated into wild-type pKGP1-1 are shown as shaded segments on the maps. The clones listed under Mutant Phenotype produced cam resistant E. coli in the cotransformation experiments. The clones listed under WT Phenotype produced cam sensitive E. coli in the cotransformation experiments.

FIG. 4: Sequencing of pMutA/B Reveals a Single Mutation. The location and identity of the mutation in pMutA/B that is responsible for producing cam resistance in the cotransformation experiments is shown. The mutant restriction fragment ligated into the T7 gene 1 region of pMutA/B is shaded. The position of the start and end of T7 gene 1 is listed with respect to T7 (above the map) and with respect to pMutA/B (below the map). The amino acid sequence shown lists the amino acid positions within T7 gene 1.

FIGS. 5-5B: Relative In Vitro (FIG. 5B) and In Vivo (FIG. 5A) Utilization of Potential T7 Promoters by GP1(lys222) and Wild-type T7 RNA Polymerase. The promoters are listed on the horizontal axis, and relative activity is represented along the vertical axis. The arrow on the right side of the graph indicates the approximate activities that differentiate the inactive, intermediate, and strong promoters. The --8 T to G mutant is classified as an intermediate promoter, while the --6 A to G mutant shows some characteristics of a strong promoter, and the --8 T to C mutant shows some characteristics of an inactive promoter. For clarity, only the top halves of the error bars are shown; however, the error bars should extend below the tops of the data columns for a distance equivalent to the value that the error bars extend above the data columns on logarithmic scale. FIG. 5A: Relative in vivo promoter utilization. The Relative In Vivo Transcriptional Activity (or Promoter Strength) is defined below in the section Methodology. The errors associated with these results were calculated by standard methods for the propagation of errors. No error bar is shown for use of the wild-type promoter by wild-type T7 RNA polymerase since the consensus T7 promoter has been defined to have a relative in vivo activity of 1.00. FIG. 5B: Relative in vitro promoter utilization. The Relative In Vitro Promoter Strength (or Promoter Strength) is defined in the section Methodology. The arrows on the right side of the graph indicate the approximate activities that differentiate the inactive, intermediate, and strong promoters. The results are the average of the two time points of at least three different samples, and the error associated with the measurements is the greater of the standard deviation observed. For clarity, only the top halves of the error bars are shown; however, the error bars should extend below the tops of the data columns for a distance equivalent to the value that the error bars extend above the data columns on logarithmic scale. No error bar is shown for use of the wild-type promoter by either wild-type T7 RNA polymerase or GP1(lys222) since the consensus T7 promoter has been defined to have a relative in vitro activity of 1.00. FIG. 6: Relative In Vitro Utilization of Bacteriophage SP6 and T3 Promoters by GP1(lys222) and Wild-type T7 RNA Polymerase. The promoters are listed on the horizontal axis, and relative activity is represented along the vertical axis. The Relative In Vitro Promoter Strength (or Promoter Strength) is defined in the section Methodology. The arrows on the right side of the graph indicate the approximate activities that differentiate the inactive, intermediate, and strong promoters. The results are the average of the two time points of at least three different samples, and the error associated with the measurements is the greater of the standard deviation observed. For clarity, only the top halves of the error bars are shown; however, the error bars should extend below the tops of the data columns for a distance equivalent to the value that the error bars extend above the data columns on logarithmic scale. No error bar is shown for use of the wild-type promoter by either wild-type T7 RNA polymerase or GP1(lys222) since the consensus T7 promoter has been defined to have a relative in vitro activity of 1.00. In these assays the concentration of T7 RNA polymerase was 80 nM while the concentration of GP1(lys222) was 40 nM. All other conditions were as previously described [Ikeda, R. A., et al., Biochemistry. 31: 9073–9080 (1992)].

FIG. 7: Comparison of the Run-off Transcripts Synthesized by T7 RNA Polymerase and GP1(lys222). The autoradiograph of a denaturing 5% acrylamide gel shows that the run-off transcripts synthesized by T7 RNA polymerase and GP1(lys222) are identical in length. Lanes 1–5 show the run-off transcripts produced by 40 nM GP1(lys222) in the presence of pCAT10-1/Ndel (0.2 µl), pCM-P1198/Ndel (2.0 µl), pCM-T270/Ndel (10.0 µl), pCM-P1208/Ndel (10.0 µl), and pLM10/PvuII (10.0 µl), respectively, while lanes 6–10 show the run-off transcripts produced by 80 nM T7 RNA polymerase in the presence of pCAT10-1/Ndel (0.2 µl), pCM-P1198/Ndel (2.0 µl), pCM-T270/Ndel (10.0 µl), pCM-P1208/Ndel (10.0 µl), and pLM10/PvuII (10.0 µl), respectively. All other conditions are described in the section Methodology. The volumes in parentheses following each template is the volume of each sample that was loaded on to the gel. Different volumes were loaded on to the gel to try to equalize the amount of each transcript run-off in each lane. The transcripts synthesized in the presence of the pCM-X# templates are approximately 2900 nucleotides long while the barely perceptible transcripts synthesized in the presence of the pLM10 templates are approximately 2370 nucleotides long. The run-off doublets seen in lanes 1–4 and 6–9 are produced by a partially effective transcriptional terminator near the ends of the pCM-X# templates.

FIG. 8: Binding of a T7 Promoter by Wild-type T7 RNA Polymerase and GP1(lys222). The autoradiograph of a 5% acrylamide gel shows the binding of wild-type T7 RNA polymerase in lanes 6–9 and GP1(lys222) in lanes 2–5 to a 5P labeled oligonucleotide that carries a T7 promoter. Lane 1–no RNA polymerase, 5 µM promoter; Lane 2–26.5 nM GP1(lys222), 5 µM promoter; Lane 3–132.7 nM GP1(lys222), 5 µM promoter; Lane 4–26.5 nM GP1(lys222), 5 µM promoter, 3 µg lambda DNA; Lane 5–132.7 nM GP1(lys222), 5 µM promoter, 3 µg lambda DNA; Lane 6–12.9 nM T7 RNA polymerase, 5 µM promoter, Lane 7–64.7 nM T7 RNA polymerase, 5 µM promoter; Lane 8–12.9 nM T7 RNA polymerase, 5 µM promoter, 3 µg lambda DNA; Lane 9–64.7 nM T7 RNA polymerase, 5 µM promoter, 3 µg lambda DNA. Complex = Position of the enzyme/promoter complexes. Promoter = Position of the unbound promoter oligo.
The ability to produce the pKGP-HA1mut4 plasmid and to use the plasmid to produce the mutant T7 RNA polymerase GP1(lys222), which will recognize the T7 promoter carried on an inactive pCM-X# plasmid, to transcribe selectively a specific gene, namely the chloramphenicol acetyl transferase gene, can serve as the basis for altering characteristics of T7 RNA polymerase to give it unique properties, such as restoring chloramphenicol resistance to the E. coli. This invention discloses a mutant T7 RNA polymerase, and a means for selecting mutant T7 RNA polymerases, which will utilize inactive T7 promoter point mutants more effectively than wild-type T7 RNA polymerase both in vivo and in vitro.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

To arrive at the mutant T7 RNA polymerase which recognizes the T7 promoter carried on an inactive pCM-X# plasmid, we devised and carried out the following process:

1. Randomly mutagenize pKGP1-1 plasmids with aqueous hydroxylamine to improve the possibility that a T7 gene 1 mutation would be responsible for the expression of the CAT gene cloned behind the inactive T7 promoter point mutant, and dialyzing the treated plasmids to remove the hydroxylamine.

2. Cotransform E. coli with the mutagenized plasmids and a mixture of seven different inactive pCM-X# plasmids and then subjecting the transformed E. coli cells to various antibiotics, including chloramphenicol, to isolate cells exhibiting chloramphenicol resistance. E. coli cells also were cotransformed with the mutagenized plasmids and pCAT10-1 to use as a comparison.

3. Select a representative pKGP1-1 plasmid, pKGP-HA1mut4, and cotransform E. coli with pKGP-HA1mut4 and pCAT10-1, pKK232-8, or each of the seven different inactive pCM-X# plasmids separately to confirm that a T7 promoter-like sequence must be present on the CAT plasmid for expression of the CAT gene in GP1(lys222) and the location and identity of the mutation responsible for altered promoter specificity.

4. Ligate restriction fragments from pKGP-HA1mut4 into a wild type pKGP1-1 plasmid to confirm that the mutation responsible for altered promoter specificity is within T7 gene 1. The results suggested that the mutation was located in the amino half of the gene.

5. Ligate the various restriction fragments from the amino half of the pKGP-HA1mut4 plasmid to localize the promoter specificity mutation.

This process confirmed that the T7 gene 1 mutation responsible for altered promoter specificity was located on the 383 bp AlwNI/BstXI restriction fragment. Sequencing of the 383 bp region using an M13mp19 clone of the pKGP1-1mut4 mutant gene 1 revealed that the only difference between the wild type T7 gene 1 and the mutant gene 1 was a G to A transition at position 664 of T7 gene 1, which changes glutamic acid (222) to lysine (222). The resulting mutant T7 RNA polymerase having the altered promoter specificity is GP1(lys222).

The selection of promoter recognition mutants of T7 RNA polymerase was accomplished by random mutagenizing pKGP1-1 with aqueous hydroxylamine, cotransforming E. coli with the mutagenized pKGP1-1 and a mixture of the seven different inactive pCM-X# plasmids, and isolating and characterizing the RNA polymerase that was present in those colonies that exhibited chloramphenicol resistance. It was established that E. coli harboring the mutant plasmid pKGP-HA1mut4 and an inactive pCM-X# are chloramphenicol resistant and that the mutation responsible for the expression of CAT from the inactive pCM-X# plasmid is a G to A transition at nucleotide 664 of T7 gene 1 that converts glutamic acid (222) to lysine (FIGS. 3 and 4).

To determine if the observed growth of the cotransformed E. coli in the presence of chloramphenicol reflects the ability of the mutant T7 RNA polymerase to utilize the T7 promoter point mutants found on the pCM-X# plasmids, in vivo and in vitro promoter utilization were measured and compared (FIG. 5). In vivo promoter utilization was determined by measuring the relative abundance of CAT in extracts of E. coli that harbored pMutA/B (the variant of pKGP1-1 that carries a single G to A transition at nucleotide 664 of T7 gene 1) and a pCM-X# plasmid, while in vitro promoter utilization was determined by measuring RNA synthesis in the presence of purified RNA polymerase and purified template. Furthermore, the location of the initiation of transcription by GP1(lys222) and T7 RNA polymerase was confirmed by comparing the lengths of the run-off transcripts synthesized by the two enzymes in the presence of linearized pCM-X# and pLM10 templates. Although the absolute magnitudes of in vivo and in vitro promoter utilization differ, the in vivo and in vitro data show the same relative trends. The mutant T7 RNA polymerase, GP1(lys222), utilizes the seven inactive T7 promoter point mutants and three intermediate T7 promoter point mutants more efficiently than wild-type T7 RNA polymerase. The correlation of the in vivo and in vitro data and the observation that GP1(lys222) and T7 RNA polymerase initiate transcription at the same location and synthesize run-off transcripts of identical length suggest that the restoration of chloramphenicol resistance in the cotransformed E. coli results from the ability of GP1(lys222) to initiate transcription from T7 promoter point mutants that are normally inactive.

The observed changes in the promoter specificity of GP1(lys222) and the location and identity of the mutation in GP1(lys222) are notable. First, the Glu to Lys substitution at amino acid 222 of T7 RNA polymerase is located near the amino-terminal domain of the enzyme, which alters promoter recognition by the mutant RNA polymerase. Similarly, it had been previously reported that a two amino acid insertion at position 222 disrupts DNA binding while preserving polymerase function [Gross et al., J. Mol. Biol., 228: 488-505 (1993)]. Apparently, regions near the aminoterminal domain of T7 RNA polymerase are involved in promoter binding. Second, the Glu to Lys substitution at amino acid 222 of T7 RNA polymerase is located near the amino-terminal domain of the enzyme (amino acids 1 to 179) and alters promoter recognition by the mutant RNA polymerase. Conversely, the inability of GP1(lys222) to utilize an SP6 promoter (FIG. 6), the inability of GP1(lys222) to specifically initiate at a T3 promoter (FIG. 7), and the ability of GP1(lys222) to specifically utilize the inactive T7 promoter point mutant indicates that promoter speci-
 specificity is not eliminated and that at a minimum GP1(Lys222) requires the context of a T7-like promoter.

Mechanistically, the Glu(222) to Lys substitution could reduce the specificity of GP1(Lys222) by at least two plausible mechanisms. First, the Lys substitution could cause a global structural change in the RNA polymerase that alters promoter binding, and second, the substitution of a positively charged amino acid side chain for a negatively charged amino acid side chain could stabilize the binding of the small negatively charged RNA initiation products within the transcription complex and increase the efficiency of utilization of a weak promoter. A third plausible mechanism seems, however, to be excluded by the observation that GP1(Lys222) binds a T7 promoter less tightly than the wild-type T7 RNA polymerase (FIG. 8). This observation suggests that it is unlikely that the positively charged lysine side chain stabilizes the binding of the RNA polymerase to the negatively charged DNA template and improves the ability of the mutant enzyme to recognize promoter-like sequences.

Mutant T7 RNA polymerases that exhibit altered promoter specificity can be isolated by screening for chloramphenicol resistance in E. coli harboring a plasmid that expresses T7 RNA polymerase (pKGP1-1) and a promoter selection vector that carries an inactive T7 promoter point mutant (an inactive pCM-X#). The mutation responsible for the altered promoter specificity of the mutant T7 RNA polymerase can be easily identified, and the effect of the mutation on promoter recognition can be measured. By isolating and characterizing mutations that alter promoter recognition by T7 RNA polymerase it should be possible to identify the regions of the RNA polymerase that can contribute to promoter recognition.

The following methodology provides additional details of the polymerase and processes of this invention. This methodology is not intended to restrict the invention to the methodology and uses described herein. In these examples, the following materials and methodology were used throughout.

1. Cell Strains. E. coli DH5 and DH5FG' were obtained from Gibco/BRL, E. coli JM101 was obtained from Stratagene, and E. coli BL21 was obtained from Novagen.

2. Chemicals. Acrylamide, agarose, ammonium persulfate, buffers, dithiothreitol, N,N,N',N'-tetramethylethylene diamine, N,N'-methylene-bis-acrylamide, and urea were electrophoresis grade. IPTG was molecular biology grade. Media was from Difeo. Antibiotics were from Sigma. DTNB and CENT A/ß-Lactamase were from Calbiochem. Hydroxylamine was from United States Biochemicals.

3. Enzymes. Restriction endonucleases, Klenow fragment of E. coli DNA polymerase I were from New England Biolabs. Calf intestine alkaline phosphatase, Sequenase Version 2.0, T4 polymerase, T4 ligase, and T4 ligase were purchased from United States Biochemical.

4. Purification. T7 RNA polymerase was purified by standard methods [Ikeda, R. A. et al., J. Biol. Chem., 267: 11322–11328 (1992)], and the mutant T7 RNA polymerase was purified by standard methods [Tabor, S. and Richardson, C. C., Proc. Natl. Acad. Sci. USA, 82: 1074–1078 (1985); Ikeda, R. A. and Richardson, C. C., J. Biol. Chem., 262: 3790–3799 (1987)] from E. coli BL21 harboring plasmids pMutA/B and pAGR3R. The T7 RNA polymerase was greater than 98% pure and was estimated to have a specific activity of 43,700 units/mg, while the mutant T7 RNA polymerase was greater than 95% pure and was estimated to have a specific activity of 14,800 units/mg [Chamberlin, M., et al., Nature, 228: 227–231 (1970)].

5. Media. LB media consists of 10.0 g of tryptone, 5.0 g of yeast extract, and 10.0 g of NaCl per liter of media; the pH of the media was adjusted to 7.5 with NaOH. SOC media consists of 20 g tryptone, 5.0 g yeast extract, 0.6 g/l NaCl, 0.5 g/l KCl, 10 mM MgCl2, 10 mM MgSO4, and 20 mM glucose; the pH of the media was adjusted to 7.5 with NaOH.

6. Nucleoside 5'-triphosphates. [2,8-3H]-ATP (50 Ci/mmol), [α-32P]-dATP (1000–1500 Ci/mmol), [α-32P]-UTP (800 Ci/mmol), and [γ-32P]-ATP (3000 Ci/mmol) were purchased from Du Pont/New England Nuclear Research Products. Ribonucleoside triphosphates were obtained from Pharmacia/LKB. dNTPs and ddNTPs were from United States Biochemicals.

7. Oligodeoxyribonucleotides. Complementary 24 base pair oligonucleotides containing the T7 promoter (5'ATTATAATGCAGACTCACTATAAGGACT3' and 3'TAAATTATGCTGATGATATCTGTA5') were purchased from Genosys Biotechnologies, Inc. M13 Sequencing Primer (−40) was purchased from New England Biolabs. Sequencing primers TEMP −34 (5'SATA GGT ACG ATT TAC3'), 171 (5'STCA A TAA AGC TGG3'), 298 (5'SCGCCA GGC TGC TAC CGG TTA3'), and 563 (5'DCGTA ACA TGC TCT CTA3') were purchased from Genosys Biotechnologies, Inc. The names of the primers indicate the location with respect to T7 gene 1 (in base pairs) that the primers will anneal. Since ligation of the pKGP-HA mut4 EcoRI/PstI fragment into the EcoRI and PstI sites of M13mp19 inserts the mutant T7 gene 1 into M13mp19 in counter clockwise direction, the noncoding (with respect to translation) strand of T7 gene 1 is the template for sequencing.

8. Plasmids. Plasmid pKKh232-8 was obtained from Pharmacia. The plasmid pKGP1-1 (an expression clone of T7 gene 1) (FIG. 2), the plasmid pCAT10-1 (a selection plasmid carrying the CAT gene cloned under the control of a wild-type T7 promoter), the plasmids pCM-T70, pCM-T286, pCM-T297, pCM-P1031, pCM-P1087, pCM-P1160, and pCM-P1208 (the plasmids carrying CAT genes cloned under the control of T7 promoter point mutants that are normally inactive in the presence of T7 RNA polymerase), and the plasmids pCM-P1198, pCM-B64, and pCM-T221 (the plasmids carrying CAT genes under the control of T7 promoter point mutants with moderate activity) have been previously described [Ikeda, R. A., et al., Biochemistry, 31: 9073–9080 (1992); Ikeda, R. A. et al., Biochemistry, 31: 20: 2517–2524 (1992)] (Table 1). Plasmid pSP64, a plasmid carrying an SP6 promoter, was obtained from Promega; plasmid pLM10, a pBR-derived carrying a T3 promoter, was a gift from Dr. William McAllister, State University of New York, Brooklyn; and plasmid pAGR3R, a plasmid carrying the lac i gene, was a gift from Dr. William Jack, New England Biolabs. All plasmids were prepared by standard methods [Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2ed (1989)]. The identities of the plasmids were confirmed by restriction mapping [Sambrook, J., et al., Molecular Cloning: A Laboratory Manual, 2ed (1989)] and by phenotypic analysis, in vivo [Ikeda, R. A., et al., Biochemistry, 31: 9073–9080 (1992); Ikeda, R. A., et al., Nucl.
**METHODOLOGY**

1. Mutagenesis of pKGPl-1 with Hydroxylamine. Mutagenesis of pKGPl-1 was performed in a 500 µl reaction containing 17-25 µg of pKGPl-1, 0.8 M hydroxylamine, and 0.1 M potassium phosphate buffer, pH 6.0. The reaction was incubated at 70 °C, and at 15, 30, 45, 60, and 75 min 100 µl samples were removed from the reaction and placed on ice. The samples were then dialyzed extensively against 75 mM CaCl$_2$ at 4 °C, and stored at -20 °C.

2. Screening for Mutant T7 RNA Polymerases with Altered Promoter Specificity. Competent E. coli JM101, 200 µl were cotransformed with 4 µl of the hydroxylamine treated pKGPl-1 (approx. 40 ng) and either 4 µl of a mixture of all of the inactive pCM-X# plasmids where the concentration of each individual inactive pCM-X# is 4 ng/µl or 40 ng of pCAT-10. SOC, 0.9 µl was added to the cells, and the culture was grown at 37 °C for one hour. LB, 1.1 ml, containing 100 µg/ml kan, 100 µg/ml kan and 60 µg/ml cam was then added to the SOC culture, and the cells were grown at 37 °C for an additional 4 hrs. Subsequently, equal volumes of the transformation, 200 µl, were spread on three different types of LB-agar plates containing either (1) 50 µg/ml kan and 50 µg/ml amp, (2) 50 µg/ml kan, 50 µg/ml amp, and 50 µg/ml cam, or (3) 50 µg/ml kan, 50 µg/ml amp, and 1.0 mM IPTG. The plates were incubated at 37 °C for 16 hrs, and the number of colonies on each plate was counted to determine which restriction fragment confers the altered promoter specificity.

3. Isolation of pKGPl-1 Plasmids Carrying Potential Mutant of T7 RNA Polymerase. Colonies that grew on the kan/amp/cam plates were transferred to 4.0 ml of LB containing 50 µg/ml kan, 50 µg/ml amp, and 30 µg/ml cam and were grown overnight at 37 °C. The plasmid DNA was isolated from the overnight cultures and the presence of pKGPl-1 and a pCM-X# plasmid was confirmed by restriction analysis (EcoRI plus PstI, data not shown). To separate the pKGPl-1 plasmid from the pCM-X# plasmid, the DNA isolated from a cam resistant colony was either run on a low melting agarose gel and the band corresponding to pKGPl-1 was isolated from the gel and used to transform E. coli JM101, or the DNA from the resistant colony was digested with PvuII for 5 hours at 37 °C (pKGPl-1 contains no PvuII sites, and the pCM-X# plasmids contains 3 PvuII sites, complete digestion of the pCM-X# plasmid was confirmed by gel analysis), and 3 µl of the digest was used to transform E. coli JM101. In either case the transformed cells were then plated on LB-agar plates containing (1) 50 µg/ml kan, (2) 50 µg/ml kan and 50 µg/ml amp, and (3) 50 µg/ml amp. The absence of colonies on the kan/amp and amp plates was used to confirm the removal of the pCM-X# plasmids. Colonies from the kan plates were then grown overnight in 4.0 ml of LB containing 50 µg/ml kan, and the plasmid DNA was isolated; the identity and purity of the isolated pKGPl-1 plasmid was confirmed by restriction analysis (EcoRI plus PstI).

To determine the promoter specificity of the T7 RNA polymerase mutants, E. coli JM101 was cotransformed with 2 µl of the isolated pKGPl-1 DNA and each of the following plasmids: 20-50 ng of each of the inactive pCM-X# plasmids separately, 40 ng of pCAT-10, and 40 ng of pKK232-8. SOC, 0.9 ml was added to the cells, and the culture was grown at 37 °C for one hour. Equal volumes of the transformations, 200 µl, were spread on three different types of LB-agar plates containing either (1) 50 µg/ml kan and 50 µg/ml amp, (2) 50 µg/ml kan, 50 µg/ml amp, and 50 µg/ml cam., or (3) 50 µg/ml kan, 50 µg/ml amp, and 1.0 mM IPTG. The plates were incubated at 37 °C for 16 hrs, and the number of colonies on each plate was counted to determine which restriction fragment confers the altered promoter specificity.

A final pair of clones was generated by cleaving wild type pKGPl-1 and the pMutE/B clone with AlwNI, and ligating a wild type restriction fragment to a pMutE/B and B/F. E. coli JM101 was then cotransformed with each clone and either a mixture of all seven inactive pCM-X# plasmids or pCM-P1031. The cells were then spread on LB-agar plates containing (1) 50 µg/ml kan and 50 µg/ml amp, (2) 50 µg/ml kan, 50 µg/ml amp, and 30 µg/ml cam, or (3) 50 µg/ml kan, 50 µg/ml amp, and 1.0 mM IPTG. The plates were incubated at 37 °C for 16 hrs, and the number of colonies on each plate was counted to determine which restriction fragment confers the altered promoter specificity.

5. Preparation of Cell Extracts. Extracts of E. coli JM101 harboring either pKGPl-1 and pCM-X# or...
mutant pKGPl-1 and pCM-X# were prepared [Ikeda, R. A., et al., Biochemistry, 31: 9073–9080 (1992)]. Extracts were always prepared immediately prior to use.


8. Determination of Relative Promoter Strength In Vivo. To determine relative promoter strength in vivo, the specific CAT activity of a sample was first divided by its specific bla activity to give a relative abundance ratio. The relative abundance ratio was then normalized by division by the relative abundance ratio measured for extracts made from E. coli containing a selection plasmid carrying a wild type T7 promoter (pCAT10-1).

This defines the relative, in vivo, strength of a wild type T7 RNA polymerase. The specific activities of identical T7 RNA polymerases. The specific activities of identical T7 RNA polymerases under other conditions were used.

9. Measurement of the Specific Activities of the T7 RNA Polymerases. The specific activities of GP1(lys222) and T7 RNA polymerase were determined under the conditions specified in the text, different DNAs were used in each case. As noted in the text, different DNAs were used (1.5 µg of pCAT10-1 or 1.5 µg of pCAT10-1/NdeI) and different buffer conditions were used (50 mM Tris-HCl, pH 8, 10 mM MgCl₂, 60 mM NaCl, and 2.5 mM spermidine).


11. Confirmation of the Site of Initiation of Transcription. To confirm that GP1(lys222) specifically initiates transcription at the T7 promoter carried on the various pCM-X# plasmids, run-off transcripts were examined. Run-off transcripts were produced in 60 µl reactions containing 50 mM Tris-HCl, pH 7.8, 60 mM NaCl, 2.5 mM spermidine, 10 mM MgCl₂, 1 mM dithiothreitol, 125 µM UTP, 400 µM of each of the other three rNTPs, 15 µCi [α-32P]-UTP, 30 µg/ml bovine serum albumin, 8 mM promoter/plasmid (Either pCAT10-1/NdeI, pCM-P1198/NdeI, pCM-T270/NdeI, pCM-P1208/NdeI or pLM10/PvuII), and 80 nM T7 RNA polymerase or 40 nM GP1(lys222). The reactions were equilibrated at 37° C, and RNA synthesis was initiated by addition of T7 RNA polymerase. At 60 min the reactions were stopped by the addition of an equal volume of loading buffer (90% formamide, 10 mM Tris-HCl, pH 7.8, 0.1% xylene cyanol, 0.1% bromophenol blue). The samples were heated to 90° C for 2 min, cooled on ice, and loaded on a 5% acrylamide (30 to 1 acrylamide to bisacrylamide), 50% urea, denaturing gel. The samples were then electrophoresed for 12 hours at 250 V. After electrophoresis the gel was fixed by soaking in an aqueous solution containing 10% methanol and 10% acetic acid. The gel was then dried and visualized by autoradiography.

12. Estimation of Promoter Binding. The ability of T7 RNA polymerase and GP1(lys222) to bind a T7 promoter was determined by a gel-retardation method [Muller, D. K., et al., Biochemistry, 27: 5763–5771 (1988)]. The promoter containing oligonucleotide was identical to the one used by Muller et al., but the 25 µl binding reactions contained 10 mM potassium phosphate, pH 7.8, 1 mM EDTA, 20 mM NaCl, 10% glycerol, 5 µM promoter (approx. 3.9×10⁷ cpm) or 2 µM promoter (1.57×10⁷ cpm), T7 RNA polymerase or GP1(lys222), and in some cases a nonspecific competitor DNA (3 µg of lambda DNA). After a 10 min incubation at 25° C, the samples were loaded onto a pre-electrophoresed 5% acrylamide gel (30 to 1 acrylamide to bisacrylamide) and electrophoresed for 50 min at 12 watts in 45 mM Tris-Borate and 1 mM EDTA. The gel was then fixed by soaking in an aqueous solution containing 10% methanol and 10% acetic acid, dried, and visualized by autoradiography. The promoter/polymerase complexes were then cut out of the gel, and the amount of promoter contained in the complexes was measured by liquid scintillation.

**SELECTION AND CHARACTERIZATION OF MUTANT T7 RNA POLYMERASE GP1(LYS222)**

The methodology described above was carried out to select and characterize a mutant T7 RNA polymerase to reconstitute chloramphenicol resistance to E. coli. Referring to the appended Figs., the results of carrying out the methodology described above to select and characterize the pKGPl-HAlmut4 plasmid to be paired with an inactive pCM-X# plasmid, resulting in chloramphenicol resistance in E. coli, are detailed below.

1. Construction of the Promoter Assay Plasmids

A 50 nucleotide, double stranded, DNA fragment containing a wild type T7 Class III promoter or a mutated T7 promoter was synthesized by a combination of
chemical and enzymatic methods [Schneider, T. D. and Stormo, G. D., *Nucl. Acids Res.*, 17: 659-674 (1989)]. The primer oligo was annealed to each of the six oligodeoxribonucleotides, WT, B, T, P, C, and G (Table IV). Large fragment of *E. coli* DNA polymerase I and dNTP's were used to extend the primer and synthesize the DNA strand complementary to the oligonucleotides WT, B, T, P, C, and G. The double stranded 50 nucleotide DNA fragments were then cleaved with HindIII and ligated between the HindIII and SmaI restriction sites of pKK232-8. (When the B oligonucleotide was cleaved with HindIII and ligated into pKK232-8 it appears that a single A was lost from the 3' end of the oligonucleotide in most of the pCM-B# clones. This A deletion does not affect the sequence of the promoter and does not affect the behavior of the clones.). The ligation mixtures were transformed into competent *E. coli* DH5a, and the transformations were grown on LB plates containing 50 µg/ml ampicillin.

For the WT oligo, ten independent colonies were chosen, and plasmid DNA was isolated from overnight cultures grown at 37° C. in 4 ml of LB containing 50 µg/ml ampicillin. Insertion of the T7 φ10 promoter into plasmid pKK232-8 was confirmed by Asel restriction maps of the plasmids. The cloning places a wild type T7 φ10 promoter just upstream of the chloramphenicol acetyl transferase gene of pKK232-8. This new plasmid is designated pCAT10-1 (FIG. 2).

For the B, T, P, C, and G oligodeoxribonucleotides, 400, 350, 329, 10, and 10 independent colonies, respectively, were chosen from the transformations. Plasmid DNA was isolated from each transformant from overnight cultures grown at 37° C. in 4 ml of LB containing 50 µg/ml ampicillin. Insertion of oligo B into pKK232-8 was confirmed by HindIII and Psil restriction maps of the plasmids. Insertion of oligo T was confirmed by Scal restriction maps, and insertion of oligos P, C, and G were confirmed by Scal or Asel restriction maps. This restriction mapping identified 252 B clones, 138 T clones, 121 P clones, 8 C clones, and 8 G clones with apparent insertions of the respective oligodeoxribonucleotides. These clones are designated pCM-B#, pCM-T#, pCM-P#, pCM-C#, and pCM-G#, respectively (FIG. 2), where # is the general designation for the number assigned to the isolated clone. In this specification the entire family of clones is referred to as pCM-X#, where X refers to the B, T, P, C, and G clones, collectively.

2. Construction of pKGP1-1

To screen the promoter assay plasmids for in vivo T7 promoter activity it is necessary to express T7 RNA polymerase in the cell strain used for screening. For this purpose we chose to construct a p15A origin plasmid that expresses T7 RNA polymerase from the tac promoter.

Plasmid pGPI-5 (2.0 µg) [Tabor, S. and Richardson, C. C., *Proc. Natl. Acad. Sci USA*, 82: 1074-1078 (1985)] was linearized by cleavage with BglII, and the linearized plasmid DNA was incubated at 37° C. with 0.3 units of Bal31 to delete the lambda P7 promoter from the DNA. Samples (0.4 µg) were removed from the reaction after 5, 10, 20, and 40 minutes of incubation, and the ends of the Bal31-treated DNA were repaired with large fragment of *E. coli* DNA polymerase I in the presence of excess dNTP's. BamHI linkers were ligated to the DNA samples, the ligation mixtures were recut with BamHI, and the BamHI digested DNA was analyzed by gel electrophoresis on a 1% low molten point agarose gel. The 2.8 kilobase fragments from the 10 and 20 min. samples were cut out of the gel, and the fragments were isolated on an Elutip-d column. These 2.8 kilobase BamHI fragments that contain T7 gene 1 were ligated into the BamHI site of pUC19, and competent *E. coli* DH5a were transformed with the two ligation reactions. The transformations were grown on LB plates containing 50 µg/ml ampicillin, 20 µg/ml X-gal, and 0.5 mM IPTG. Ten white colonies from each of the two ligations were picked for further analysis. Plasmid DNA was isolated from 4.0 ml overnight cultures (37° C. and 50 µg/ml ampicillin in LB) of each of the twenty colonies, and the lengths of the DNA fragments inserted into the pUC vector were analyzed by restriction mapping with AccI, EcoRI, BamHI, BamHI & ScaI, EcoRI, HindIII, and AccI & KpnI. One clone contained an insert with 337 bases removed from the BgIII end of the BgIII/BamHI fragment of pGPI-5. This clone was designated pGPI-20B, and was chosen for further manipulation.

The 2.5 kilobase EcoRI/HindIII fragment from pGPI-20B was isolated on a 1% low melting point agarose gel and ligated into the EcoRI/HindIII site of plasmid pKK223-3. This places T7 gene 1 immediately downstream of the tac promoter of plasmid pKK223-5. This plasmid was designated pKK-gpl-1. Plasmid pKK-gp-1 was cut with NaeI, and BglII linkers were ligated onto the NaeI cleavage sites. The ligation mix was recut with BglII and BgIII, and the 4362 base pair Bgll/BglII(Nael) fragment from pKK-gp-1 was ligated to the 2892 base pair BamHI/BglII(Nael) fragment from pACYC177 [Chang, A. C. Y. and Cohen, S. N., *J. Bacteriol.*, 134: 1141-1156 (1978); Rose, R. E., *Nucl. Acids Res.*, 16: 356 (1988)]. This ligates the tac promoter-T7 gene 1 fusion to the p15A origin of pACYC177 and generates a plasmid that also expresses both ampicillin and kanamycin resistance. To inactivate the ampicillin resistance gene, the p15A/gene 1 plasmid was recut with BglII and ScaI, repaired with T4 DNA polymerase and excess dNTP's, and recircularized with T4 DNA ligase. This deletes 360 bases from the middle of the ampicillin resistance gene and produces plasmid pKGP1-1 (FIG. 2). In *E. coli* JM101 plasmid pKGP1-1 efficiently produces active T7 RNA polymerase in the presence of IPTG.

3. Mutagenesis of pKGP1-1 and Selection of Possible Mutants of T7 Gene 1

To improve the possibility that a T7 gene 1 mutation would be responsible for the expression of the CAT gene cloned behind the inactive T7 promoter point mutant, pKGP1-1 was exposed to aqueous hydroxylation at 70° C. for 15, 30, 45, 60 and 75 min. The treated plasmid samples were then dialyzed to remove the hydroxylation, and *E. coli* JM101 was transformed with the treated pKGP1-1 and either a mixture of the seven inactive pCM-X# plasmids or pCAT10-1 (Table I and FIG. 2). After the transformed cells were allowed to recover in liquid culture in the absence of antibiotics, amp, kan, and cam were added to the liquid media, and the entire culture was incubated for 4 hrs at 37° C. This step amplifies the abundance of those mutants that are cam resistant, and simplifies the isolation of mutants that might occur very infrequently. The transformed cells were then plated on LB-agar containing either (1) 50 µg/ml kan and 50 µg/ml amp (kan/amp), or (2) 50 µg/ml kan, 50 µg/ml amp, and 30 µg/ml cam (kan/amp/cam). It was observed that the pKGP1-1 sample that had been incubated with hydroxylation for 60 min...
yielded approximately 200 colonies on the kan/amp and kan/amp/cam plates when cotransformed with pCAT10-1, and 100 colonies on the kan/amp plate and 26 colonies on the kan/amp/cam plate when cotransformed with the mixture of the seven different pCM-X# plasmids that carry the seven inactive T7 promoter point mutants. This suggested that the CAT gene on at least one of the pCM-X# plasmids was expressed in the presence of the hydroxylamine treated pKGPl-1. Due to the early imposition of cam selection the 26 mutant colonies do not reflect the frequency of mutation.

To confirm that the cam resistance observed with *E. coli* harboring inactive pCM-X# plasmids required the presence of a hydroxylamine treated pKGPl-1 plasmid, mutant pKGPl-1 plasmid DNA was isolated from the 26 colonies that grew on the kan/amp/cam plate, and fresh *E. coli* JM101 was cotransformed with the mixture of the seven inactive pCM-X# plasmids and each of the different mutant pKGPl-1 plasmids. All of the isolated mutant pKGPl-1 plasmids allowed *E. coli* to grow on kan/amp/cam plates in the presence of the mixture of inactive pCM-X# plasmids. Since the 26 isolates behaved identically in this assay, one representative isolate (pKGPl-HA1mut4) was chosen for further characterization.

*E. coli* JM101 was cotransformed with pKGPl-HA1mut4 and pCAT10-1, pKK232-8, or each of the seven different inactive pCM-X# plasmids separately. As shown in Table 2, all seven inactive pCM-X# plasmids and pCAT10-1 allowed *E. coli* to grow on kan/amp/cam plates in the presence of pKGPl-HA1mut4, while *E. coli* harboring pKK232-8 and pKGPl-HA1mut4 were not resistant to chloramphenicol. This demonstrated that a T7 promoter-like sequence must be present on the CAT plasmid for expression of the CAT gene in the presence of pKGPl-HA1mut4.

4. Identification of a Mutation That Alters the Specificity of T7 RNA Polymerase

To demonstrate that the mutation(s) that alters the apparent promoter specificity of the T7 RNA polymerase encoded on pKGPl-HA1mut4 is within T7 gene 1, restriction fragments from pKGPl-HA1mut4 were ligated into a wild type pKGPl-1 plasmid, and the new plasmids were tested for their ability to confer chloramphenicol resistance to *E. coli* harboring the inactive promoter selection plasmid pCM-P1031. The first restriction fragments that were cleaved from pKGPl-HA1mut4 and ligated into wild type pKGPl-1 were the T7 gene 1 EcoRI/HpaI and HpaI/PstI restriction fragments. On pKGPl-HA1mut4 and on pKGPl-1 EcoRI cleaves between the tac promoter and the ATG of gene 1, HpaI cleaves near the middle of gene 1, and PstI cleaves just after the stop codon of gene 1. This allows the amino and carboxyl halves of gene 1 to be independently ligated into a wild type pKGPl-1 and generates the new plasmids pMutE/H and pMutH/P that contain the amino and carboxyl halves of the gene 1 from pKGPl-HA1mut4, respectively (FIG. 3).

*E. coli* JM101 was cotransformed with pCM-P1031 (Table 1) and either pMutE/H or pMutH/P, and the transformations were spread on kan/amp, kan/amp/ IPTG, and kan/amp/cam plates. Only cells containing pMutE/H and pCM-P1031 were able to grow in the presence of chloramphenicol. This confirmed that the mutation responsible for altered promoter specificity was within gene 1 and suggested that the mutation was located in the amino half of the gene.

The same strategy was used to construct and test the plasmids shown in FIG. 3, and altered promoter specificity always segregated with the amino terminal restriction fragment. This allowed us to localize the promoter specificity mutation to the 832 bp EcoRI/BstXI restriction fragment cloned into pMutE/B; however, further localization of the promoter specificity mutation required manipulation of pMutE/B.

The restriction enzyme AlwNI cleaves pMutE/B and pKGPl-1 twice, between the EcoRI and BstXI sites within gene 1 and near the origin of replication of the plasmids. By ligating the two AlwNI fragments of pMutE/B to the two reciprocal AlwNI fragments of pKGPl-1, the clones pMutE/A and pMutA/B are generated (FIG. 3), and the EcoRI/BstXI fragment of pMutE/B is further subdivided. When pMutE/A and pMutA/B were tested for altered promoter specificity, only pMutA/B conferred chloramphenicol resistance to *E. coli* harboring pCM-P1031. In addition, *E. coli* JM101 cotransformed with pMutA/B and pCAT10-1, pKK232-8, or each of the seven different inactive pCM-X# separately showed the same growth characteristics on kan/amp, kan/amp/cam, and kan/amp/IPTG plates as was observed with pKGPl-HA1mut4 (data not shown). This suggested that the T7 gene 1 mutation responsible for altered promoter specificity was located on the 383 bp AlwNI/BstXI restriction fragment.

Finally, sequencing of the 383 bp region using an M13mp19 clone of the pKGPl-HA1mut4 mutant gene 1 revealed that the only difference between the wild type T7 gene 1 and the mutant gene 1 was a G to A transition at position 664 of T7 gene 1 (FIG. 4). Since the phenotypic assays done with the mutant gene 1 subclones showed that the mutation affecting the specificity of the mutant T7 RNA polymerase was located between positions 753 and 1143 of pKGPl-HA1mut4, the mutation at 1105 of pKGPl-HA1mut4 (position 664 in relation to T7 gene 1) changes glutamic acid 222 to lysine is probably responsible for the altered promoter specificity of the mutant RNA polymerase (GP1[lys222]).

5. The In Vivo and In Vitro Specificity of GP1[lys222]

In previous work with the compatible plasmids pKGPl-1 and pCM-X#, equivalent bacterial growth on kan/amp and kan/amp/cam plates accompanied by no growth on kan/amp/IPTG plates indicated that the pCM-X# plasmid carried a strong T7 promoter [Ikeda, R. A., et al., *Biochemistry*, 31: 9073–9080 (1992); Ikeda, R. A., et al., *Nucleic Acids. Res.*, 20: 2517–2524 (1992)]. With pKGPl-HA1mut4 and pMutA/B, the seven inactive pCM-X# plasmids and pCAT10-1 all showed equivalent bacterial growth on kan/amp and kan/amp/cam plates and no growth on kan/amp/IPTG plates (Table 2), but control experiments also showed that *E. coli* JM101 harboring pMutA/B would not grow on amp/IPTG plates (data not shown). Apparently, the lack of growth on kan/amp/IPTG plates is due to the overproduction of the mutant T7 RNA polymerase and is not an indication of a strong promoter on the pCM-X# selection plasmid. The mutant RNA polymerase may be toxic or the mutant RNA polymerase may utilize cryptic T7-like promoters within the cell that express proteins that kill the host *E. coli*. Although the toxicity of the overproduced mutant T7 RNA polymerase makes it impossible to judge the efficiency of promoter usage in plating experiments, it does not seem to interfere with the selection of possible promoter recognition mutants of T7 RNA polymerase in the absence of IPTG. However, to estimate how
efficiently GP1(lys222) RNA polymerase utilizes point mutants of T7 promoters it is necessary to directly measure promoter activity in vivo and in vitro.

We have previously shown that in vivo usage of the potential T7 promoters carried on the pCM-X# plasmids can be estimated by measuring CAT activity relative to β-lactamase activity in extracts of *E. coli* harboring pKGP1-1 and pCM-X# plasmids. [Ikeda, R. A., et al., *Biochemistry*, 31: 9073-9080 (1992); Ikeda, R. A., et al., *Nucl. Acids Res.*, 20: 2517-2524 (1992)]. From these measurements, if in vivo usage of the wild-type T7 promoter (pCAT10-1) by wild-type T7 RNA polymerase (pKGP1-1) is defined as 1.0, then in vivo usage of the inactive T7 promoter point mutants by wild-type T7 RNA polymerase ranges from 0.005±0.03 to 0.01±0.005, and in vivo usage of the three intermediate strength T7 promoter point mutants ranges from 0.04±0.018 to 0.40±0.18 (FIG. 5A). In contrast, while in vivo usage of the wild-type T7 promoter (pCAT10-1) by GP1(lys222) is comparable to usage of the same promoter by wild-type T7 RNA polymerase, in vivo usage of the inactive T7 promoter point mutants by GP1(lys222) ranges from 0.024±0.109 to 0.22±0.12, and in vivo usage of the three intermediate T7 promoter point mutants ranges from 0.22±0.048 to 0.91±0.39 (FIG. 5A). This showed that GP1(lys222) uses the inactive and intermediate T7 promoter point mutants 5 to 25 times and 2 to 6 times more efficiently than wild-type T7 RNA polymerase, respectively (FIG. 5A).

To confirm the in vivo measurements, promoter usage was also measured in vitro. Although T7 RNA polymerase and GP1(lys222) exhibit different absolute activities in these in vitro reactions (Table 3) relative comparisons are informative. If in vitro usage of the wild-type T7 promoter (pCAT10-1) by both T7 RNA polymerase and GP1(lys222) is defined as 1.0, the in vitro usage of the inactive and intermediate T7 promoter point mutants by T7 RNA polymerase ranges from 0.005±0.005 to 0.018±0.005 and 0.005±0.005 to 0.36±0.03 (FIG. 5B), respectively, while the in vitro usage of the inactive and intermediate T7 promoter point mutants by GP1(lys222) ranges from 0.014±0.005 to 0.05±0.01 and 0.04±0.01 to 0.81±0.03, respectively. Although the absolute magnitudes of in vivo and in vitro promoter usage differ, the trends noted for in vivo promoter usage are almost duplicated in the in vitro measurements; however, some differences are noted. For example, while the in vitro data shows that GP1(lys222) uses the intermediate T7 promoter point mutants 2 to 8 times more efficiently than T7 RNA polymerase, the data also shows that the mutant enzyme uses the inactive T7 promoter point mutants only 2 to 6 times more efficiently than the wild-type RNA polymerase. This difference in the in vivo and in vitro data is probably due to the difficulties encountered in measuring the low in vitro usage of the inactive T7 promoter point mutants by T7 RNA polymerase. Since background is difficult to subtract from these measurements, it is likely that in vitro usage of the inactive T7 promoter point mutants by T7 RNA polymerase is underestimated and that the relative increase in the efficiency of in vitro usage of these promoters by GP1(lys222) is underestimated.

6. Utilization of Homologous Phage Promoters

The ability of GP1(lys222) to utilize the seven different T7 promoter point mutants that are not utilized by wild-type T7 RNA polymerase showed that the mutant T7 RNA polymerase is less specific than the wild-type enzyme, but these measurements do not define a limit to the sequence variations accommodated by GP1(lys222).

To determine if GP1(lys222) still requires the context of a T7 promoter, transcription from plasmids containing no T7 promoter, an SP6 promoter, or a T3 promoter was measured in vitro. FIG. 6 shows that little or no RNA is produced in the presence of templates carrying no promoter or an SP6 promoter GP1(lys222) and that a T3 promoter is poorly utilized by either wild-type T7 RNA polymerase (activity=0.009±0.003) or GP1(lys222) (activity=0.027±0.008). However, GP1(lys222) uses a T3 promoter more efficiently than wild-type T7 RNA polymerase, and the in vitro utilization of the T3 promoter by GP1(lys222) is comparable to the in vitro utilization of a number of the inactive T7 promoter point mutants. Nevertheless, this result may not actually reflect specific initiation at the T3 promoter of pLM10 since few specific transcripts are seen with GP1(lys222) in the run-off assays shown below. The in vitro activity of GP1(lys222) on pLM10 may simply reflect nonspecific initiation.

7. Initiation of Transcription Occurs at the Potential T7 Promoters

To show that transcription initiates at the potential T7 promoters on the pCM-X# plasmids run-off transcripts produced by T7 RNA polymerase and GP1(lys222) were compared. A plasmid carrying a wild-type T7 promoter (pCAT10-1), a plasmid carrying an intermediate strength T7 promoter point mutant (pCM-P1198), two plasmids carrying inactive T7 promoter point mutants (pCM-T270 and pCM-P1208), and a plasmid carrying a T3 promoter (pLM10) were linearized by cleaving the plasmids with either NdeI (the pCM-X# plasmids) or PvuII (pLM10). Cleavage of the pCM-X# plasmids with NdeI places the potential T7 promoters 2905 nucleotides from the end of each template, while cleavage of pLM10 with PvuII places the T3 promoter 2371 nucleotides from the end of its template. The linearized plasmids were then used in transcription reactions containing either 80 nM T7 RNA polymerase or 40 nM GP1(lys222). FIG. 7 shows that the lengths of the run-off transcripts produced by T7 RNA polymerase are identical to the lengths of the run-off transcripts produced GP1(lys222). In addition, specific run-off transcripts are easily observed in reactions containing T7-like promoters, while specific run-off transcript are barely detectable in reactions containing a T3 promoter (pLM10/PvuII). The run-off doublts seen in the pCM-X# lanes of FIG. 7 are due to the presence of a sequence near the end of the pCM-X# templates that acts as a terminator of T7 transcription. This would seem to indicate that transcription by GP1(lys222) is promoter dependent and that the promoter must resemble a T7 promoter.

8. Promoter Binding by GP1(lys222)

A two amino acid insertion at position 222 of T7 RNA polymerase has been previously reported to disrupt promoter binding without affecting polymerase function [Gross, L., et al., *J. Mol. Biol.*, 228: 488-505 (1993)]; however, the opposite phenomenon, tighter promoter binding, would be a plausible mechanism that might explain the ability of GP1(lys222) to recognize an expanded range of T7 promoter-like sequences. The positively charged lysine side chain could stabilize the binding of the RNA polymerase to the negatively charged DNA template and improve the ability of the mutant enzyme to recognize promoter-like sequences. To test whether the expanded specificity of
GP1(lys222) is due to stabilization of the promoter/enzyme complex, the binding of T7 RNA polymerase and GP1(lys222) to an oligonucleotide containing a T7 promoter was measured by gel retardation. FIG. 8 shows that GP1(lys222) is still capable of forming promoter specific complexes, but that the affinity of the mutant enzyme for a T7 promoter is diminished. The data obtained from two different trials suggests that the promoter binding affinity of GP1(lys222) (1.4 × 10^6 M\(^{-1}\) ± 5 × 10^5 M\(^{-1}\)) is twenty times less than the promoter binding of wild-type T7 RNA polymerase (2.8 × 10^6 M\(^{-1}\) ± 2.2 × 10^5 M\(^{-1}\)). In addition, it is also observed that a GP1(lys222)/T7 promoter complex is more easily disrupted by nonspecific competitor DNA (FIG. 8, lanes 4 and 5) than a wild-type T7 RNA polymerase/T7 promoter complex (FIG. 8, lanes 8 and 9). The susceptibility of GP1(lys222)/promoter complexes to disruption by nonspecific DNA could either be due to the decreased stability of the specific enzyme/promoter complex or to an increased affinity for nonspecific DNA. Whatever the case, the Glu to Lys substitution in GP1(lys222) produces an enzyme with reduced affinity for T7 promoters and an ability to utilize an expanded range of T7 promoter-like sequences.

The above detailed description of the preferred embodiment of this invention is for illustrative purposes and is not meant to limit the spirit or scope of this invention, or its equivalents, as defined in the appended claims.

### TABLE I

<table>
<thead>
<tr>
<th>Selection Plasmid</th>
<th>Mutation Carried by the Promoter</th>
<th>Phenotypic Promoter Strength</th>
</tr>
</thead>
<tbody>
<tr>
<td>pKK232-8</td>
<td>None</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCATI0-1 Wild Type T7</td>
<td>None</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-T297 Point Mutant</td>
<td>-11G to T</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-P1160 Point Mutant</td>
<td>-8C to G</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-T270 Point Mutant</td>
<td>-9C to A</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-P1087 Point Mutant</td>
<td>-9C to T</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-P1198 Point Mutant</td>
<td>-8T to G</td>
<td>Intermediate</td>
</tr>
<tr>
<td>pCM-T286 Point Mutant</td>
<td>-8T to A</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-B64 Point Mutant</td>
<td>-8T to C</td>
<td>Intermediate</td>
</tr>
<tr>
<td>pCM-P1208 Point Mutant</td>
<td>-7C to G</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-P1031 Point Mutant</td>
<td>-7C to A</td>
<td>Inactive</td>
</tr>
<tr>
<td>pCM-T251 Point Mutant</td>
<td>-6A to G</td>
<td>Intermediate</td>
</tr>
</tbody>
</table>

### TABLE III

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Specific Activity (u/mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7 RNA polymerase</td>
<td>14,800</td>
</tr>
<tr>
<td>pCATI0-1</td>
<td>2,500</td>
</tr>
<tr>
<td>pCATI0-1/Ndel</td>
<td>2,200</td>
</tr>
<tr>
<td>T7</td>
<td>4,370</td>
</tr>
<tr>
<td>pCATI0-1</td>
<td>18,900</td>
</tr>
<tr>
<td>pCATI0-1/Ndel</td>
<td>15,100</td>
</tr>
</tbody>
</table>

### TABLE II

Phenotypic Characterization of pKGP-HAlmut4

<table>
<thead>
<tr>
<th>Promoter Selection Plasmid</th>
<th>Mutation in the T7 Promoter</th>
<th>T7 RNA Polymerase Plasmid</th>
<th>Kan/ Amp/ Cam</th>
<th>IPTG</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCATI0-1</td>
<td>None</td>
<td>pKGP-1</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>Mix of all 7 All 7 Inactive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pK232-8 No Promoter</td>
<td>None</td>
<td>pKGP-1</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>pCM-T270 -9C to A</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>pCM-T286 -8T to A</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>pCM-T297 -11G to T</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>pCM-P1031 -9C to T</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>pCM-P1087 -9C to T</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>pCM-P1160 -9C to G</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
<tr>
<td>pCM-P1208 -7C to G</td>
<td>pKGP-HAlmut4</td>
<td>200</td>
<td>0</td>
<td>200</td>
</tr>
</tbody>
</table>

### TABLE IV

<table>
<thead>
<tr>
<th>Designation</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
<tr>
<td>B</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
<tr>
<td>F</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
<tr>
<td>I</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
<tr>
<td>C</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
<tr>
<td>G</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
<tr>
<td>Primer</td>
<td>5'-CTGATTCGAAATTTATACGAGCACATCTATTAGGAGGAAACTTGTTACCA-3'</td>
</tr>
</tbody>
</table>

Oligonucleotide B, C, G, and Primer were purchased from National Biosciences. Oligo T was purchased from New England Biolabs, and Oligo WT was a gift from Dr. Keith McKenney, Center for Advanced Research in Biotechnology. In oligos B and T, the lower case italics are used to designate those bases in the oligonucleotide that are a mixture of 96.1% of the indicated base and 1.3% of each of the other 3 bases. In oligo P, the lower case italics are used to designate those bases in the oligonucleotide that are a mixture of 67.0% of the indicated base and 11% of each of the other 3 bases. The mutation in oligo C is an A to C (lower case italics) point mutation at -10, and the mutation in oligo G is an A to G (lower case italics) point mutation at -10. The primer oligo is complementary to the 3' end of oligos WT, B, P, T, C, and G.
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(iii) NUMBER OF SEQUENCES: 23

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2652 Base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Other nucleic acid

(A) DESCRIPTION: Molecule sequenced is a clone of the T7 genomic DNA that spans T7 RNA polymerase

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Bacteriophage T7
(B) STRAIN: Wild-type
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: Not applicable
(B) CLONE: pKGP-HAtmut4

(viii) POSITION IN GENOME:

(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:

(A) NAME/KEY: T7 RNA Polymerase GP1(lys222)
(B) LOCATION: 1 to 2652
(C) IDENTIFICATION METHOD: By expressing and characterizing the protein encoded by the gene.

(D) OTHER INFORMATION: The glu to lys substitution at residue 222 alters promoter recognition by the T7 RNA polymerase

(x) PUBLICATION INFORMATION:

(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warsamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER: 

(i) RELEVANT RESIDUES IN SEQ ID NO: Nucleotides 1 to 2652 encode the entire T7 RNA polymerase GP1(lys222); however, the difference between GP1(lys222) and wild-type T7 RNA polymerase is a G to A substitution at nucleotide 664

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG AAC ACG ATT AAC ATC GCT GCT TCT TCT GAC GAC ATC GAA CTG
Met Asn Thr Ile Asn Ile Ala Lys Asn Asp Phe Ser Asp Ile Glu Leu

GCT GCT ATC CCG TCT AAC ACT CTG GCT GAC CAT TAC GGT GAG CCT TTA
Ala Ala Ile Pro Phe Asn Thr Leu Ala Asp His Tyr Gly Glu Arg Leu

10 15 20 25 30
<table>
<thead>
<tr>
<th>Position</th>
<th>Amino Acid 1</th>
<th>Amino Acid 2</th>
<th>Amino Acid 3</th>
<th>Amino Acid 4</th>
<th>Amino Acid 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-20</td>
<td>Ala</td>
<td>Arg</td>
<td>Glu</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td>21-30</td>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Lys</td>
<td>Met</td>
</tr>
<tr>
<td>31-40</td>
<td>Glu</td>
<td>Asp</td>
<td>Glu</td>
<td>Lys</td>
<td>Gly</td>
</tr>
<tr>
<td>41-50</td>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>51-60</td>
<td>Ala</td>
<td>Arg</td>
<td>Lys</td>
<td>Met</td>
<td>Gly</td>
</tr>
<tr>
<td>61-70</td>
<td>Ala</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Lys</td>
</tr>
<tr>
<td>71-80</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>81-90</td>
<td>Ile</td>
<td>Ala</td>
<td>Asp</td>
<td>Asp</td>
<td>Trp</td>
</tr>
<tr>
<td>91-100</td>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Gly</td>
<td>Ala</td>
</tr>
<tr>
<td>101-110</td>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>111-120</td>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>121-130</td>
<td>Ala</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Ser</td>
</tr>
<tr>
<td>131-140</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>141-150</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>151-160</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>161-170</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>171-180</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>181-190</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>191-200</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>201-210</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>211-220</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>221-230</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>231-240</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>241-250</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>251-260</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>261-270</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>271-280</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>281-290</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>291-300</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>301-310</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>311-320</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>321-330</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>331-340</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>341-350</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td>351-360</td>
<td>Met</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>355</td>
<td>360</td>
<td>365</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATG AAT CCT GAG CTC ACC GCG TGG AAA CGT GCT GCC GCT GGT GTG</td>
<td>Met Asn Pro Glu Ala Leu Thr Ala Trp Lys Arg Ala Ala Ala Ala Val</td>
<td>1152</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAC CGC AAG GAC AAG GCT CGC AAG TCT CGC GTT ATC AGC GGT TTC</td>
<td>Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe</td>
<td>1200</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATG CTT GAG CAA GCC AAT AAG TTT CCT AAC CAT AAG GCC ATC TGG TTC</td>
<td>Met Leu Glu Gln Ala Asn Lys Phe Ala Asn His Lys Ala Ile Trp Phe</td>
<td>1248</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCT TAC AAC ATG GAC TGG CGC GGT CGT GTT TAC GCT GTG TCA ATG TTC</td>
<td>Pro Tyr Asn Met Asp Trp Arg Gly Arg Val Tyr Ala Val Ser Met Phe</td>
<td>1296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC CCG CAA GGT AAC GAT ATG ACC AAA GGA CTG CTG AAA TAC CAC GTG</td>
<td>Asn Pro Glu Gly Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys</td>
<td>1344</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GGT AAA CCA ATC TGG AAG GAA GGT TAC TAC TGG CTG AAA ATC CAC GTG</td>
<td>Gly Lys Pro Ile Gly Lys Gly Tyr Tyr Trp Leu Lys His Gly</td>
<td>1392</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCA AAC TCT CGG GGT GTG CAT AAG GTT CCG TTC CCT GAG CGC ATC AAG</td>
<td>Ala Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys</td>
<td>1440</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTC ATT GAG GAA AAC CAC GAG AAC ATC ATG GTG GTC AAG CTA GCT CCA</td>
<td>Phe Ile Glu Gln Ala Asp Met Thr Ala Cys Ile Met Tyr Pro</td>
<td>1488</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CTG GAG AAC ACT TGG TGG GCT GAG CAA GAT TCT CCT CTC GTG TTC CTT</td>
<td>Leu Glu Asn Thr Trp Trp Ala Glu Asp Ser Pro Phe Cys Phe Leu</td>
<td>1536</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCG TTC TCC CTC CGG TGT GAG TAC GCT TGG GGA GTA CAG CAC CAC GCC CGT AGC TAT</td>
<td>Ala Phe Cys Phe Cys Phe Gly Glu Gln His His Gly Leu Ser Tyr</td>
<td>1584</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAC TGG CTC CTT CCG CGT CCC TTT GAC GGG GGT CTC TGG TGC TTC</td>
<td>Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Glu</td>
<td>1632</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAC TCC TCC CCC CGT GCC TTT GAC GGG GGT CTC TGG TGC TTC</td>
<td>His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Arg Arg Ala Val Asn</td>
<td>1680</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTG GCT TCC CGC ATG GAG ACC GTT CAG GAC ATC AAC GAT GAA ACC GTT</td>
<td>Leu Leu Pro Ser Glu Thr Val Gln Asp Ile Tyr Gly Ile Val Ala Lys</td>
<td>1728</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAA GTC AAC GAG ATT CTA CAA GCA GAC GCA ATC AAT GGG ACC GAT AAC</td>
<td>Lys Val Asn Glu Ile Glu Ala Asn Ala Ile Asn Thr Asn Gly His</td>
<td>1776</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAA GTA GTT ACC GTG ACC GAT</td>
<td>Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Gly Lys</td>
<td>1824</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTC AAC CGT GCC ACC AAC</td>
<td>Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Trp Leu Ala Tyr Gly</td>
<td>1872</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GTC AAG CGT GCC ACC AAC GCA CTG GCT CAA TGG CTG GCC GCT TAC GTG</td>
<td>Val Lys Leu Gly Trp Ala Leu Ala Alan Trp Leu Ala Tyr Gly</td>
<td>1920</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCC AAA GAG TTC GCC TGC TTC CTT CAG CAA CAA GTC GCT GAA GAT</td>
<td>Ser Lys Glu Phc Arg Gin Gln Val Leu Glu Asp Thr Ile Gln</td>
<td>1968</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCA GCT ATT GAT TCC GCC AAG GTG CTG ATG TCC ACT CAG CGG AAT CAG</td>
<td>Pro Ala Ile Asp Gly Lys Lys Met Thr Gin Pro Asn Gin</td>
<td>2016</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCT GGT GAA TAC ATG GCT AAG CTG ATT TGG GAA TCT GTG AGC GGT ACG</td>
<td>Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr</td>
<td>2064</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
GTG GTA GCT GCG GAT GAA GCA ATG AAC TGG CTT AAG TCT GCT GCT AAG
Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys
690 695
CTG CTG GCT GCT GAG GTC AAA GAT AAG AAG ACT GGA GAG ATT CTT CGC
Leu Leu Ala Ala Glu Val Lys Asp Lys Thr Gly Glu Ile Leu Arg
705 710 715 720
AAG CTT TGC GCT GTO CAT TGG GTA ACT CCT GAT GGT TTC CTT GTC
Lys Arg Cys Ala Val His Thr Val Thr Pro Asp Gly Pro Val Trp
725 730 735
CAG GAA TAC AAG AAG CCT ATT CAG GCT GCT GAG GTC AAA GAT AAG AAG
Glu Ala Met
690
Asn Trp Leu Lys
745
ATT GAT GCA CAC AAA CAG GAG TCT GGT ATC GCT CCT AAC TTT GTA CAC
Ile Asp Ala His Lys Glu Ser Gly Ile Ala Pro Asn Phe Val His
770 775 780
AAG TAC GGA ATC GAA GCT CCC GCT AAA GAT AGC GAG
Lys Ala Asp Asp Lys
805
ATT CCG GAC GAT GCT GCC TTC GTC GAG TCT GGT ATC GCT CCT AAC TTT GTA CAC
Ile Asp Ala His Lys Glu Ser Gly Ile Ala Pro Asn Phe Val His
770 775 780
AGC CAA GAC AGC CAC CTT CGT AAG ACT GTA TGG GCA CAC GAG
Ser Glu Asp Gly Ser His Leu Arg Lys
785 790 795 800
TTC GCT GAC CAG TAT GAG TCT TGG GTA CTA GCT GAT TCT AAC TTT GTA CAC
Phe Ala Asp Leu His Glu Ser Gly Ile Ala Pro Asn Phe Val His
830 835 840
CGC GCT GAT CAG GAT GCT GCC TTC GTC GAG TCT GGT ATC GCT CCT AAC TTT GTA CAC
Pro Ala Lys Glu Asp Leu Asp Leu Arg Asp Ile Leu Glu Ser Asp Phe
870 875 880
CGC TTC GCC TAA 2652
Ala Phe Ala
883
(2) INFORMATION FOR SEQ ID NO:2:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 883 Amino Acids
(B) TYPE: Amino Acids
(C) STRANDEDNESS: Not Applicable
(D) TOPOLOGY: Not Applicable

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Entire protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bacteriophage T7
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pKGP-HAmut4
(ii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITE: Not applicable

(i) FEATURE:
(A) NAME/KEY: T7 RNA Polymerase GP1(lys222)
(B) LOCATION: 1 to 883
(C) IDENTIFICATION METHOD: By expressing and characterizing the protein encoded by the gene.
(D) OTHER INFORMATION: The glu to lys substitution at residue 222 alters promoter recognition by the T7 RNA polymerase

(i) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Wamsamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:

(k) RELEVANT RESIDUES IN SEQ ID NO: Amino Acids 1 to 883 encode the entire T7 RNA polymerase GP1(lys222); however, the difference between GP1(lys222) and wild-type T7 RNA polymerase is a Glu to Lys substitution at residue 222.

(i) SEQUENCE DESCRIPTION: SEQ ID NO:2

<table>
<thead>
<tr>
<th>Met</th>
<th>Asn</th>
<th>Thr</th>
<th>Ile</th>
<th>Ala</th>
<th>Lys</th>
<th>Asn</th>
<th>Asp</th>
<th>Phe</th>
<th>Ser</th>
<th>Asp</th>
<th>Ile</th>
<th>Glu</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td>Ile</td>
<td>Pro</td>
<td>Phe</td>
<td>Asn</td>
<td>Thr</td>
<td>Leu</td>
<td>Ala</td>
<td>Asp</td>
<td>His</td>
<td>Tyr</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20</td>
<td>15</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Arg</td>
<td>Glu</td>
<td>Gln</td>
<td>Leu</td>
<td>Ala</td>
<td>Leu</td>
<td>Glu</td>
<td>Arg</td>
<td>Ser</td>
<td>Tyr</td>
<td>Glu</td>
<td>Met</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35</td>
<td>40</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Arg</td>
<td>Lys</td>
<td>Met</td>
<td>Phe</td>
<td>Glu</td>
<td>Arg</td>
<td>Glu</td>
<td>Leu</td>
<td>Lys</td>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>55</td>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asp</td>
<td>Asn</td>
<td>Ala</td>
<td>Ala</td>
<td>Lys</td>
<td>Pro</td>
<td>Leu</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>65</td>
<td></td>
<td>70</td>
<td>75</td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Ile</td>
<td>Ala</td>
<td>Arg</td>
<td>Ile</td>
<td>Asn</td>
<td>Asp</td>
<td>Trp</td>
<td>Phe</td>
<td>Glu</td>
<td>Val</td>
<td>Lys</td>
<td>Ala</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>90</td>
<td></td>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Lys</td>
<td>Arg</td>
<td>Pro</td>
<td>Thr</td>
<td>Ala</td>
<td>Phe</td>
<td>Glu</td>
<td>Phe</td>
<td>Glu</td>
<td>Lys</td>
<td>Ala</td>
<td>Lys</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Val</td>
<td>Ala</td>
<td>Tyr</td>
<td>Ile</td>
<td>Thr</td>
<td>Ile</td>
<td>Thr</td>
<td>Thr</td>
<td>Leu</td>
<td>Ala</td>
<td>Cys</td>
<td>Leu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>115</td>
<td>120</td>
<td>125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asp</td>
<td>Asn</td>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
<td>Glu</td>
<td>Ala</td>
<td>Val</td>
<td>Ala</td>
<td>Ser</td>
<td>Ala</td>
<td>Ile</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>135</td>
<td></td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Glu</td>
<td>Asp</td>
<td>Glu</td>
<td>Ala</td>
<td>Arg</td>
<td>Phe</td>
<td>Glu</td>
<td>Arg</td>
<td>Ile</td>
<td>Asp</td>
<td>Leu</td>
<td>Ala</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>150</td>
</tr>
<tr>
<td>His</td>
<td>Phe</td>
<td>Lys</td>
<td>Asn</td>
<td>Val</td>
<td>Glu</td>
<td>Glu</td>
<td>Lys</td>
<td>Asn</td>
<td>Lys</td>
<td>Arg</td>
<td>Val</td>
<td>Gly</td>
<td>His</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>170</td>
</tr>
<tr>
<td>Val</td>
<td>Tyr</td>
<td>Lys</td>
<td>Ala</td>
<td>Phe</td>
<td>Met</td>
<td>Gin</td>
<td>Val</td>
<td>Val</td>
<td>Glu</td>
<td>Ala</td>
<td>Asp</td>
<td>Met</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>180</td>
<td>185</td>
<td>190</td>
</tr>
<tr>
<td>Lys</td>
<td>Gly</td>
<td>Leu</td>
<td>Leu</td>
<td>Gly</td>
<td>Glu</td>
<td>Ala</td>
<td>Trp</td>
<td>Ser</td>
<td>Ser</td>
<td>Trp</td>
<td>His</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>195</td>
</tr>
<tr>
<td>Ser</td>
<td>Ile</td>
<td>His</td>
<td>Val</td>
<td>Gly</td>
<td>Val</td>
<td>Arg</td>
<td>Cys</td>
<td>Ile</td>
<td>Glu</td>
<td>Met</td>
<td>Leu</td>
<td>Ile</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>210</td>
</tr>
<tr>
<td>Gly</td>
<td>Met</td>
<td>Val</td>
<td>Ser</td>
<td>Leu</td>
<td>His</td>
<td>Arg</td>
<td>Gin</td>
<td>Asn</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Val</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>225</td>
</tr>
<tr>
<td>Ser</td>
<td>Glu</td>
<td>Thr</td>
<td>Ile</td>
<td>Glu</td>
<td>Leu</td>
<td>Ala</td>
<td>Pro</td>
<td>Glu</td>
<td>Tyr</td>
<td>Ala</td>
<td>Glu</td>
<td>Ala</td>
<td>Ile</td>
</tr>
</tbody>
</table>

SER     VAL     GLY     VAL     VAL     CYS     ILE     GLU     MET     LEU     ILE     LYS     VAL     THR     LYS     VAL     VAL     GLU     VAL     VAL     GLY     GLU     ASP     SER     VAL     ILE     GLU     ALA     PRO     GLU     TYR     ALA     GLU     ALA     ILE     ALA     THR     LYS     VAL     VAL     GLU     VAL     MET     LEU     LYS     VAL     VAL     GLU     VAL     VAL     GLY     GLU     ASP     SER     GLU     THR     ILE     GLU     LEU     ALA     PRO     GLU     TYR     ALA     GLU     ALA     ILE     ALA     THR     LYS     VAL     VAL     GLU     VAL     VAL     GLY     GLU     ASP     SER     GLU     THR     ILE     GLU     LEU     ALA     PRO     GLU     TYR     ALA     GLU     ALA     ILE     ALA     THR     LYS     VAL     VAL     GLU     VAL     VAL     GLY     GLU     ASP     SER     GLU     THR     ILE     GLU     LEU     ALA     PRO     GLU     TYR     ALA     GLU     ALA     ILE     ALA     THR     LYS     VAL     VAL     GLU     VAL     VAL     GLY     GLU     ASP     SER     GLU     THR     ILE     GLU     LEU     ALA     PRO     GLU     TYR     ALA     GLU     ALA     ILE     ALA     THR     LYS     VAL     VAL     GLU     VAL     VAL     GLY     GLU     ASP     SER     GLU     THR     ILE     GLU     LEU     ALA     PRO     GLU     TYR     ALA     GLU     ALA     ILE     ALA     THR 5,385,834
Arg Ala Gly Ala Leu Ala Gly Ile Ser Pro Met Phe Gin Pro Cys Val
245
250
255
Val Pro Pro Lys Pro Trp Thr Gly Ile Thr Gly Gly Gly Tyr Trp Ala
260
265
270
Asn Gly Arg Arg Pro Leu Ala Leu Val Arg Thr His Ser Lys Lys Ala
275
280
285
Leu Met Arg Tyr Glu Asp Val Tyr Met Pro Glu Val Tyr Lys Ala Ile
290
295
300
Asn Ile Ala Gin Asn Thr Ala Trp Lys Ile Asn Lys Lys Val Leu Ala
305
310
315
320
Val Ala Asn Val Ile Thr Lys Trp Lys His Cys Pro Val Glu Asp Ile
325
330
335
340
345
350
355
360
365
Pro Ala Ile Glu Arg Glu Glu Leu Pro Met Lys Pro Glu Asp Ile Asp
370
375
380
Tyr Arg Lys Asp Lys Ala Arg Lys Ser Arg Arg Ile Ser Leu Glu Phe
385
390
395
400
Met Leu Glu Gin Ala Gin Ala Gin Ala Gin Asp Met Thr Lys Gly Leu Leu Thr Leu Ala Lys
405
410
415
420
425
430
435
440
445
Gly Lys Pro Ile Gly Lys Glu Gly Tyr Tyr Trp Leu Lys Ile His Gly
446
450
455
460
Leu Asn Cys Ala Gly Val Asp Lys Val Pro Phe Pro Glu Arg Ile Lys
465
470
475
480
Phe Ile Glu Glu Asn His Glu Asn Ile Met Ala Cys Ala Lys Ser Pro
485
490
495
500
505
510
Leu Glu Asn Thr Trp Trp Ala Glu Gin Asp Ser Pro Phe Cys Phe Leu
515
520
525
Ala Phe Cys Phe Glu Tyr Ala Gly Val Gin His His Gly Leu Ser Tyr
530
535
540
Asn Cys Ser Leu Pro Leu Ala Phe Asp Gly Ser Cys Ser Gly Ile Gin
545
550
555
560
His Phe Ser Ala Met Leu Arg Asp Glu Val Gly Gly Arg Ala Val Asn
565
570
575
Lys Val Asn Gin Ile Leu Gin Ala Asp Ala Ile Asn Gly Thr Asp Asn
580
585
590
Glu Val Val Thr Val Thr Asp Glu Asn Thr Gly Glu Ile Ser Glu Lys
595
600
605
Val Lys Leu Gly Thr Lys Ala Leu Ala Gly Gin Trp Leu Ala Tyr Gly
610
615
620
Val Thr Arg Ser Val Thr Lys Arg Ser Val Met Thr Leu Ala Tyr Gly
625
630
635
640
Ser Lys Glu Phe Gly Phe Arg Gin Gin Val Leu Gin Asp Thr Ile Gin
645
650
655
Pro Ala Ile Asp Ser Gly Lys Gly Leu Met Phe Thr Gin Pro Asn Gin
660
665
670
675
Ala Ala Gly Tyr Met Ala Lys Leu Ile Trp Glu Ser Val Ser Val Thr
680
685
Val Val Ala Ala Val Glu Ala Met Asn Trp Leu Lys Ser Ala Ala Lys 690 695 700
Leu Leu Ala Ala Glu Val Lys Asp Lys Thr Gly Glu Ile Leu Arg 705 710 715 720
Lys Arg Cys Ala Val His Trp Val Thr Pro Asp Gly Phe Pro Val Trp 723 730
Gln Glu Tyr Lys Lys Pro Ile Gln Thr Arg Leu Asn Leu Met Phe Leu 740 745 750
Gly Gln Phe Arg Leu Gln Pro Thr Ile Asn Thr Asn Lys Asp Ser Glu 755 760 765
Ile Asp Ala His Lys Gln Glu Ser Gly Ile Ala Pro Asn Phe Val His 770 775 780
Ser Gln Asp Gly Ser His Leu Arg Lys Thr Val Val Trp Ala His Gln 785 790 795 800
Lys Tyr Gln Ile Glu Ser Phe Ala Leu Ile His Asp Ser Phe Gly Thr 805 810 815
Ile Pro Ala Asp Ala Ala Asn Leu Phe Lys Ala Val Arg Glu Thr Met 820 825 830
Val Asp Thr Tyr Glu Ser Cys Asp Val Leu Ala Asp Phe Tyr Asp Gln 835 840 845
Phe Ala Asp Gln Leu His Glu Ser Gln Leu Asp Lys Met Pro Ala Leu 850 855 860
Pro Ala Lys Gln Asn Leu Asn Leu Arg Asp Ile Leu Gln Ser Asp Phe 865 870 875 880
Ala Phe Ala *** 883

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5096 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pKK232-8

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Cloning polylinker located between nucleotides 177 and 212 and in front of a promoterless chloramphenicol acetyl transferase

(CAT) gene.
(B) LOCATION: 177 to 212
**IDENTIFICATION METHOD:** The sequence was provided by Pharmacia Biotech, Oct. 1989

**OTHER INFORMATION:**

**PUBLICATION INFORMATION:**

(A) AUTHORS: Brosius, J. and Lupski, J. R.

B) TITLE: 

C) JOURNAL: Methods in Enzymology

D) VOLUME: 153

E) ISSUE: 

F) PAGES: 54-68

G) DATE: 1987

**RELEVANT RESIDUES IN SEQ ID NO: 1 to 5096**

**SEQUENCE DESCRIPTION:**

<table>
<thead>
<tr>
<th>Seq ID No</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-5096</td>
<td>TTCCCAAGCCA TCAAAATAAC CGAAAAAGGCTC AGTCCGAAGGA CTGGGCCTTT</td>
</tr>
<tr>
<td>5097-1000</td>
<td>CTTTTTATCT GTTTTTGTC GGTGAACGCT CTCTGATGTA GACAAATCC</td>
</tr>
<tr>
<td>1001-1500</td>
<td>CAGGACGGCC GCCATAAACT GCCAGGGAATT CCGTCCAGCT</td>
</tr>
<tr>
<td>1501-2000</td>
<td>CGACCCAAAGC TTTGAGTATGA CAAATCCGGC GAGCTTCGAC GAGATTTTCA</td>
</tr>
<tr>
<td>2001-2500</td>
<td>GCCGGGAGCG GATTTGAACGT TTGCGAAGCA AGGCTTCGATGAT</td>
</tr>
<tr>
<td>2501-3000</td>
<td>AGGCCGAGCC AACTGAAAC GTTTTCATCG CTCTGGAGTG AATACCAGTA</td>
</tr>
<tr>
<td>3001-3500</td>
<td>GTCATGAGTC AAATCCTTCG CCCCCGTTTT CACCATGGGC AAATATTATA</td>
</tr>
<tr>
<td>3501-4000</td>
<td>AAAGACGGTG ACGGTATGAG CACCTTGCTT TTTTTGAGGCA TTTCAGTCAG</td>
</tr>
<tr>
<td>4001-4500</td>
<td>CTTGCTCAATG TACCTATAAC CAGACCGTTC AGCTGGATAT TACGGCCTTT</td>
</tr>
<tr>
<td>4501-5000</td>
<td>TTAAAGACCG TAAAGAAAAAA TAAAGACAAG TTTATCTCCG CCTTTATCA</td>
</tr>
<tr>
<td>5001-5500</td>
<td>CATTCTGCG GCCCTGATGA ATGCTCATCC GGAATTCCGT ATGGCAATGA</td>
</tr>
<tr>
<td>5501-6000</td>
<td>AAGACGGTGA GCTGGTGATA TGGGATAGTG TTCACCCTTG TTACACCGTT</td>
</tr>
<tr>
<td>6001-6500</td>
<td>TTCCATGAGC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>6501-7000</td>
<td>CGTTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>7001-7500</td>
<td>CTTTCTGGCA CAGTCATGGA ACATATATTT CCAAAATCAT CACCATGGGC</td>
</tr>
<tr>
<td>7501-8000</td>
<td>GATTTTCTAC ACAATCATTTG CCAAATGACGT GCTTTTCAAC</td>
</tr>
<tr>
<td>8001-8500</td>
<td>GTTGTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>8501-9000</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>9001-9500</td>
<td>CGTTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>9501-10000</td>
<td>CTTTCTGGCA CAGTCATGGA ACATATATTT CCAAAATCAT CACCATGGGC</td>
</tr>
<tr>
<td>10001-10500</td>
<td>GATTTTCTAC ACAATCATTTG CCAAATGACGT GCTTTTCAAC</td>
</tr>
<tr>
<td>10501-11000</td>
<td>GTTGTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>11001-11500</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>11501-12000</td>
<td>CGTTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>12001-12500</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>12501-13000</td>
<td>CGTTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>13001-13500</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>13501-14000</td>
<td>CGTTTTATCT GTTTTTGTC GGTGAACGCT CTCTGAGTATG AATACCAGGA</td>
</tr>
<tr>
<td>14001-14500</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>14501-15000</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>15001-15500</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
<tr>
<td>15501-16000</td>
<td>GCCCTAAGAC AAATGAAAC GTTTTCATCG CTCTGAGATG AATACCAGGA</td>
</tr>
</tbody>
</table>
CTACAAGCCA TCCCCCACA GATACGGTAA ACTAGCTTGG TTTTGCATC 1650
AGGAAAGGAC CGTGTTTGGTT GGATAGAGAA AGATTTTCAG CCGATGATAC 1700
ATTAACATCG AAGCAGAAG CGGTCGTAATA AAAAAAGATT TTGCTGGGCG 1750
CAGTACGGCC GGGTCTCCAC ACCTACCCAT GCCGAATCAG GAAATGGAAAC 1800
GCCGTAAGGC CGATGTGTAGT GTGGGCTTCT CCACTGAGCG AGTAGGGAAC 1850
TGCCAGGCTAT CAATTACACAA GAAAACTGCA TCGAAAGAAC TGGGCTTTTC 1900
GTTCTTCTTC TGTTGGCTGC GGGATGAGAT AGATTTTGAG CTGATATCTT 2100
ACAAAGCATC CCCCCACAGA AATTAAGCAG AAGGCCATAC 2050
TGACGGATAGG CTTTTGTCGG TTTCTACAAA CTCTTCCTGT GGTATATCTC 2300
AACGGTTGAC GACCTGAGCA AACAACATGAA TGGTCTCAGG TTTCCGTGTT 2350
CTCTCTGCTG TGAGGAGCTG AGCTACGCGG CCGTTGGCAG TATCTGATTTA 2400
GCGAGAATGA TACCGGATAC CGGAGGAAAC AGGAAGCGAC TGCTGCTGCA 2500
TATTAACCAGC CGGCTGCAAG TGATCTGAGT CCTTGACATAG 2550
CGCATCCTATA CCGGCAATGTG TTTACCTCCA CAACAGTTCA GTAACCCGGC 2600
ATGCTCATCA TCGTAAACCAG GTATGCGTAG AGATCTCTCTC 2650
GTATCATTACC CCCCCGATACAA AGAAAATCACC CTTAACAGGG AGGCCATCAAG 2700
TGACCAAAAAC GGCANAAAAC GCTGCTAACAG TGCGGGGCTT TATCAGGAAA 2750
CAGACATATT AAGCTCGGGA GAAGCTCAAC GAGCTGATCGG CGCGTGCTTT 2800
GCTGCTCGGC CCGTTGCTGT GATGAGCTGT AAAACCTGCT ACACATCAGG 2850
CTCCGCGCCA GGGCAAGGAT TTGCTGGTCA GTGACGCTGCA 2900
AGCCCGTCAG GGGCGTCGAG CCGGGTGTT GGCGGGTGTCG GGCGCAGCCA 2950
TGACCCGATC AGCTAGCCGT AGGCAGAGTC ATACTGCTGC ACCTGCTCAG 3000
CATCAGGCAA GATTGCTACTG AGAGTGCACC ATATGCGGTG TGAAATACCG 3050
CACAGAGTGA CAGGGAGAAA ATACCGCATC AGGGCTGCTT CCGTCTGCTT 3100
GCTCAGCTAC TCGTCTGCGT CCGTCTGCTG CGTCGCGCGA GCCTGATGAC 3150
CTCCTCCTACT GGGATATCTA GGGTTATCCA CAGAATCAGG GGATAACCGA 3200
GGAAAGAACA TGAGGAGAAA AGGGCAAGGA AAGCCAGGGA ACCGGAAAAA 3250
GGCGGCGTTG CTGGGCTTTTT TCGTAACTCC GCCGGCGCCG ACGAGCAATC 3300
ACAAAAATCG AGCTCAAGGT CAGGGGTCAG CAAAACCGAC AGGACTATAA 3350
AGATACGCGG TGGTCTCCCC TGGAAAGCTC CCGTGCGGCT CTCTGTCTCC 3400
GACCCTGCCG CTTACCAGAT ACCTGCTGCG TTTCCTCCCT CGGTGGAGCG 3450
TGCGCTTCTA CTTAATGCTA CCGTGTAGGT ATCTCAAGGT TGTTAGGCT 3500
GGTTGTCTCA AGCTGGCGGT TGGGGCTGAA CCCCCCGTTCC AGCGGCCAGG 3550
CTGCCGCTTA TCGGTAACGT ATCGTCATGA GTCGACTGTT AGAAGACAGG 3600
ACTTATCGGC ACTGGCAGCA GACCCTGCTA ACAAGGATAGC AAGAGCAGCC 3650
TATGTAAGCG GTGTCACAGA GTTCTTGAAG TGGTGGCCTA ACTACGCTTA 3700
CAGTAAGAAG ACAAGTTGGT GTATCTGTCG CTCTGCTGAAG CTTATACAGT 3750
TCGGAAAAAG AGTTGGAAG TCCTGCTGGA GGATGAGTGG CAGATTACAG 3800
AGCCGGTGTG TTTTGTGTGG AAAAGCGAGC ATCAGGAGTGA ACCAAGGAG 3850
ATCTCAAGAA GATCCTTATG TCTTTTCTAC GGGTCTGAC GCTCGCTGGA 3900
AGCGAAACCT ACGTTAAGGG TTATTTGCTCA TGAAGGATTG AAAAAGGATC 3950
TTACCCCTGA TCTTTTAAAA TTAAAAATGA AGTTTTAATT CAATCTAAG 4000
TATATATCGG TAAAACCTTG TGTACAGTTA CAAATGCTTA ATCAGTGAGG 4050
CACCTATCTC AGCGTATCTG TTATTTGCTT CATCCATAGT TGCTGACTC 4100
CCCTGCTGTG AGATAACATG GATAGGAGGA GGTCTACAGT CTGAGGGCCAG 4150
TGCTCAATGT ATACCGCGAG ACCCAGCCTC ACGGCCTCAG GATTATAACG 4200
CAAAACCCA GCCAGCCGGA AGGCGCGCCAG GCAGAAGTGG TCCTGCAACT 4250
TAGCTCCTTC GCTCCTCGCA TCTTGGTCTG AAGTAATTGT TGCACAGTGT 4300
TATCGTAAGAT GCCTTCTTGT GACTGTTGAG TACTCAACCA AGTTAGCTTG 4350
AAGATAGTGT ATGCCGCCGAG CCAATGGGC CAGACTGGCAG TGGATGGCCAG 4400
ATATACCCGC GCCACATAGC AGAAGTTTAA AAGTGCCTAT CATTGAAGAA 4450
CTGTCTCTGG GCCAAAGACTT CTCAAAGGAC ATCAGGCTGT CTGATCGACG 4500
TTGATGTCAG CACCATCGTC AGTACAGGCT TATCTCAAGA AGTGGATGTG 4550
TACCAGGCT TCGTGTGGGC CGCTGAGGAA GAAATTGTTG GACCACAAGA 4600
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4650
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4700
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4750
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4800
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4850
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4900
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 4950
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 5000
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 5050
AACCTATAAA AATTGCTCA TCACCACGTC ACCATGCTC TCTTCTGTTT 5100

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 5110 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: double
   (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
   (A) ORGANISM: Not applicable
   (B) STRAIN: Not applicable
   (C) INDIVIDUAL ISOLATE: Not applicable
   (D) DEVELOPMENTAL STAGE: Not applicable
   (E) HAPLOTYPE: Not applicable
IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-X#(null)

POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

FEATURE:
(A) NAME/KEY: Potential T7 promoter located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (C.A.T) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The general clone was constructed from pCK232-8 and randomly mutagenized oligos. The potential promoters inserted were then sequenced, and each plasmid was given a different CM-mirical assignment.
(D) OTHER INFORMATION: If the potential T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase then CAT can be expressed from the plasmid. The lower case bases between positions 198 and 220 indicate that there is a probability that there is at least one mutation in this generalized sequence. See the definition of pCM-X copyright (c) 1990, Microsoft Corp.

PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 510

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
TTCCCAGGCA TCAAAATAAA CGAAAGGCTC AGTCGAAAGA CTGGGCCTTT
CGTTTTATCT GTTGTGGTGC GGTGAACGCT CTCCTGAGTA GGACAAATCC
GCCGGGAGCG GATTTGAACG TTGCGAAGCA ACGGCCCGGA GGGTGGCGGG
CAGGACGCCC GCCATAAACT GCCAGGGAA T TCCCCTRR W W YYGAAA T ta a 2
T7 Promoter Positions
\[
\begin{align*}
\text{TACGACGAGTATT} & \quad \text{AGAACCAATGC} \\
\text{ATTTAGAGCATA} & \quad \text{ATCCCAATCG} \\
\text{GAGCTGGAGT} & \quad \text{ATCCGAAGT} \\
\text{CGTATGGCT} & \quad \text{TGCAGCTTT} \\
\text{TTACACATAT} & \quad \text{ATTCGCAAGA} \\
\end{align*}
\]
```

RELEVANT RESIDUES IN SEQ ID NO: 1 to 510
GCTTTATCAG AAGCCAGACA TTAACGCTTC TGGAGAAAAT CAACGAGCTG 2750
GACGCAGGATG AACAGGCAAG CATCTGCTGA TCGCTTCACG ACCACGCTGA 2800
TAGAGCTTAC CGCAAGCTGCC TCGCCGCTTT CGGTGATGAC GGTGAAAAAC 2850
TCTGACACAT GCAGCTCCCG GAGCAGGTCA CAGGTGTGCT GTAAGCGGAT 2900
GCCGAGGCGA GACAAAGGGCC TCAAGGCGCG TCAGCGGGTG TTGGCGGGTG 2950
TCGGGGCCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3000
CCGGGGCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3050
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3100
GCGAGGGCTA TGAGCTCAGC AATACGGCTGA CAGCTCGCTG TGCGGGCCG 3150
CAGGGGACTAA CAAGACAAAA AATGTTTAAA CAGTTTAAAAA ACAAGGGCCA 3200
AGGAAAGCTTAA AAAAGGGCCCT GTTCTGGGCG CCCTTCCACG ACCAGGCTGA 3250
ACGACAGACT ATTAAAGAAT CAGGAGTTTT CCGCTTGCAGG TCGGGGCGCA 3300
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3350
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3400
CCGGGGCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3450
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3500
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3550
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3600
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3650
CCGGGGCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3700
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3750
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3800
CCGGGGCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3850
GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 3900
CCGGGGCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3950
TGACGTGCA GTGAAGCAGAA GGAGGATTTTT GCAGCTTGCAGG TCGGGGCGCA 4000
TATCAAAAG GATCTCCCC CACTCTTAATT AACCCGCGCC GACGATCTGC 4050
AAATCTCTAC AGGCTTCTAC CCTAGCTGATC TCCAACCTTAA TTATCCTTTA 4100
TGAGCGGCTG TGAGCTTCTTT CTTGGTGGC TGGTTTTTTC GCCAGCTCCT 4150
CCATCTGCC GCAGGCTGGA AATGACAGCC CGAGAGCAGC GCTCAGCCGC 4200
TCCAGATTCA TGAGCAATCA ACCAGGCCAC CGGAAGGGCC GACGCAGAAG 4250
GTGCTGACTC AACATTTATCC GCCTCTCATC AATTTTTAAT TGTTACCGGT 4300
GAAGCTAGAG TAGGATTCTC CGGTGTTTTC ATGATGTGCTG TACCTGATTC 4350
CATTGTCGCA GCAGATGTTG ATGTCAGTTT GCAGTTCGCC TCCGGGGCAG 4400
TGAGCTCCTG TCTCGGAACT TCGAGGCCAT TCCAAGCTTG CAGCGGCGCC 4450
GGCTCAACCA CGGCTATAAT CCGCTCACAC TAGCAGGGCA AAATACGGCTG 4500
TTAATCAGT GCCGCTGACC TCTCGACTAT GCAGTGTCCC CTATGCAGTA 4550
GATTGAGGCT TACCTCCGCT GTGATAGAAG TTGTTACACG GAAACGGCTT 4600
ACCAGCTGAA TGATGCTCAGG GCACGAGGTT GCTTTGCGCC 4650
GCCGGCAGCA CCGGATAAAA CCGGAGGAAC TACATCCAGG GATCCTTACC 4700
TACATCAGGG AAAACGTTCT CCGGGGGCAA AACTCTCAAG GATCCTTACC 4750
(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No
(iv) ANTI-SENSE: No
(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYP: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCAT10-1

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: T7 promoter located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232·8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: The T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase; consequently, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993

(ix) SEQUENCE DESCRIPTION: SEQ ID NO:5:

5,385,834

CTTGTAGAT CCAOTCGAT GAATCCTCAT CTTGCACCCA ACTGATCTTC AGATCAGAT
AGCATCTTTT ACTTTACCC ACGTTTCTGG GTGAGCAAAA ACAGGAAAGC
AAAATGCGGC AAAAGGGGAA TCAAGGGAAGA CACGAAATT GCAGAATTCTC
ACATCTTTT TTATTTCAATA TTATGAAAGC ATTTATCAGG GTTATGTCT
CTGAGCAGGA TACATATTG AATTGATTGA GAAATAAATCAAAGGTTGGG
TTGGCGCGAC ATTTCCGGCA AAAGTCGCCAC CTGACGTCTA AGAACCATT
ATTATCATGA CATAAACCTA TAAAAATAGG CGTATCACGA TCTTCAAGAA

TCTTCAAGAA 5'110
TTCCAGGCA TCAATAAAA CGAAAGGCTC AGTCGAAAGA CTGGGCTTTT 50
CGTTTTATCT GTGTGTTTGT CGTGAAAAGT CTCTCTAGTA GGAATAATCC 100
GCCGGAGCG GATTGAACG TGCTGGAGCA ACGGCCGGA GGGGCGCGGG 150
CGAGACGCCC GCCATAAATC GCCAAGGGAAT TGCCCTGAAAT TCGAAATTAA 200
* T7 Promoter Positions
  15
TACGACTCAG TATAAGGAGA AAGCTGAGT AGGCACAAATC CGCCGAGCTT 250
  .10 -0.5 +1 +5
CGACAGAGATT TTTAGGAGCT AAGGAAGCTA AAATGGAGAA AAAAACTACT 300
GGATATACCA CGGTGATATA ATCCCAATCG CATCGTAAGA AACATTTTA 350
GGCATTTAGAG TGTTTGGCTC ATGGAAACG GGTTGTATAT ATCAGGGAAT 400
ATATTAGGGAT ATGAAGGGAT ATTCAGTGCT AAGGAAAGAT AAGGGCTAAAA 450
CCGCCCTTTA TTTACTATAT ATCCCGGCCG ATGAGGCTCT ATCTGGAATT 500
CGTATTAGGCA ATGAAGAGCG GGGAGCTGGA GATATTGGAAT AGTGTTCACC 550
CTTGTATACC CGTTTTCCAT GAGCAAACTG AAACGTTTTT ATCGCTCTGG 600
AGTGAATACC ACCGAGATTG TCCGCAAGTT TTTCAGCTCT 700
AGAATATTTT TTTGCTCTCA GCCAATCCCT TGGTTGAGTT TACAGCTATT 750
GATTATAGTG TGGCAATAT ATGGGACACG TTGGCGCTCT ATGAGGCTATC 800
GGCAGAATAT TATACGCAAG GGCAGAAGGT ATCTTGAGGA AAATGGAGAA 850
AGGTTGATACG TGCTGTTCCT GTGAGGCTCT ATGTTGAGAA AATGCTATT 900
GAATTACCA ACCTGCTGGA TGAATGGGCA GGCAGGGCGT AATTTCCTTTA 950
AGCGATTTAT TGCGGCCCTT AAAGCGCTGGA TTCTAGCTAC 1000
GAATATACCA CGCGAGACG TCCCTGCTCT CGGCGCTCTTC GAGAATGTA 1050
GATCTACCC ACGTACAAAT CAGAAATGGG GAGGCTCTAC AATCTACGTA 1100
AATTGGCCAA CCACACCCCT GCACAGACG GGCTGAGATT TCCAGTTGCA 1150
CAAGAGCCCG GCACGCGGCA TCCTGCGCTT TTTGGAGGAG AAGGAAAGAT 1200
TTTCACTCTC GATACAGATTA AAATCAGAAAC AGGAAAGGCTT GCTGATAAAAC 1250
AGAAATTGGC TGGCGGCTGAAG GGGCGGTGTC TCCCACTGTTT GAGAAACTTT 1300
AACTCAGAAG TGGAAAGGCGC TACGCGCGAT CAGTGTGTAG TGGTCTCCTA 1350
TGCGAGATGA GGGGAACTGCG AAGCATCAAA TAAAGGAGGA GGGGACTGTG 1400
AAAGACTTGGA CTTTTCTCTTT TAGTCTTTTG TGGTCGGCTGA AGCTCCTCTC 1450
GAGTACAGAG CCGTCGCGCCAT GAGGGGATTT GAGGGTTGCA AAGGAAACGG 1500
CGGAAGGGTT GCAGGCCAGGA GCAGGCCCACT AAACGTGCAAG GCATCAAAAT 1550
AAGAGGAGCG CCATCGCTAC GGATGGGCTT TTTGGCTTTC TACAAACTTT 1600
TCCTGTCAGAT ATACTCTAAA GCCATCTCCC CAGCATCAGG TAAACTACTG 1650
TCTGAAGTTA CATCGAGAAA GCGCGGATTG TGGCAGTAGA GAGAAGATTT 1700
TCAGCTGGTAT ACGAGATAAA TGAGAAGGCA GAGCGGCTCT GATAAAACAG 1750
AAATTCTTCG TGGAGGTAGC GCGCGGATTG TGGCAGTAGA GAGAAGATTT 1800
CTCAAGAAGT AAACACCCCT CGACGGCTTT TGGTGTATGG TCTTCCATGG 1850
CGGAGGTAGC GAACTGCAAG CGATCAAAAT AAAGGAAAGG TCTGAGCGGA 1900
AGACTGGGCC TTTGCTTCTT TCTTGTGTCT TGGTGTGGAAC GCTCTCTCTG 1950
GTAGGACAAA TCCGCCGGGA GCCGATTTGA AGCTTGCGAA GCAACGGCCC 2000
GGAGGGTGCC GGGCAGGACG CCCGCCATAA ACTGCCAGGC ATCAAATTA 2050
GCAGAAGGCC ATCTGAGCAG ATGGGCTTTT TGCCCTCTCA AAAACTCTTC 2100
CTGTCGTCAT ATCTACGACG CATC CCCCCCA CAGATACGGT AAATACGCT 2150
GGTGTTTCA GTGTTTTCGA GTCTGTGGAA GCAACGGCCC 2200
AAACCTACA TCTGTATTTAA GGAACGCTCG GCATGGACCC TAACTAGTTT 2250
TTCTCTGTTC GTCCAGGCTC CTAACCGCCA GCAACGGCCC 2300
TCCAGTTTGC ATGCTGTGGT CTAACCGCCA GCAACGGCCC 2350
ACGGAGGATG CAACTGAGCA AAAG GCGCCCTT AACATGCGCC 2400
GCTTTTACAG AAATGAGCTC TGGGAAACGT GACGAGCTCG 2450
GACCGGATG AACAGGAGAG CATCTGTTGA TCCTTCAGCG ACCACGGCTA 2500
TGAGGGAATC TGCCAGCAGT GAGGCTTGTC AGATGCGGAT 2550
TCTGACCACT CAAGTGTCGG CAAGATGCGG CCACTGACCA 2600
CTCTGCTTTC ATGCTCAGCA TACCACCCAT GCCAGGAGAT 2650
ACGGAGGCTG CATGTCGGCG TGGCTGTAGT CAGATAGCCG 2700
GCTTTATCAG AACCAGATCG TTTAGGCTTC TGAGGAAACGT CAACTGAGCA 2750
GACCGGATG AACAGGAGAG CATCTGTTGA TCCTTCAGCG ACCACGGCTA 2800
TGAGGGAATC TGCCAGCAGT GAGGCTTGTC AGATGCGGAT 2850
TCTGACCACT CAAGTGTCGG CAAGATGCGG CCACTGACCA 2900
TGGGGGCGCA GCCCATGGCC AGTCGGTGAA TTGATGCAGT CAGATAGCCG 2950
TCTGACCACT CAAGTGTCGG CAAGATGCGG CCACTGACCA 3000
GCTTTATCAG AACCAGATCG TTTAGGCTTC TGAGGAAACGT CAACTGAGCA 3050
GACCCGGCA TCCAGAGGCTG TGGCTGTAGT CAGATAGCCG 3100
TCTTCCAGGG TCTCCGCGTAC ATGCTGAGCG TGGCTGTAGT CAGATAGCCG 3150
GCTTTATCAG AACCAGATCG TTTAGGCTTC TGAGGAAACGT CAACTGAGCA 3200
GACCGGATG AACAGGAGAG CATCTGTTGA TCCTTCAGCG ACCACGGCTA 3250
TCTGACCACT CAAGTGTCGG CAAGATGCGG CCACTGACCA 3300
CCTTCCAGGAG CATCAGACCA GCGCCGCGC TGGGTGAAGT CAGATAGCCG 3350
GACTAGGACT ATAAAGATAC CAAGCGTTTC CCCCTGGGAAT CTCCTCCTGT 3400
GGCTTATAAG TTCCACAGCT GCCGCTTATGC AGATAGCTGT CCTCCCTGT 3450
GCCCTGGAGA AGGTGTCGAG GTTCTCTAAG GCCTGTGAGT CAGATAGCCG 3500
GTGGGTGGAT CAGCTGTGGT CTCAGGCGCG CAGGCTGTGAGT CAGATAGCCG 3550
GTTCTGGGAT CAGCTGTGGT CTCAGGCGCG CAGGCTGTGAGT CAGATAGCCG 3600
GCGAAGAGAC GTGCTGACGT CAGCTGTGAGT CAGATAGCCG 3650
TTACGAGGTC GAGGTATGGA CGGTTGGAAG TCTCCAGGAG CATCAGACCA 3700
AAACCAGGCC TGGTGACCGG GTGGTTTTTG TGGGCAAGA GCAGATAGCCG 3750
GAAGCGGCTG ACCCGTGGGT CTTATCCGGT AATATCTGTC TTAATGCAGT 3800
AAAACGAGG TGGTACCGGT GTGGTTTTTG TGGGCAAGA GCAGATAGCCG 3850
GCAAGAAAAG ACAGATCTCA GAGAGATCTT CATACGTGCTT CAGATAGCCG 3900
TGACCGGTCG TGGAGAGAAA ACTGAGGTGT CAGATACGGT AAATACGCT 3950
TATCAAAAG GATCTTCGAT GTGCTTCTTA AAAATAAAA ATGAGGTTTT 4000
AAATCAATCT AAAGTATATA TGAGTAACCT TGCTTGACAA GTTACCAATG 4050
CTTAATCAGT GAAGCCACCTA TCTAGGAGAT CTGCTATTT GCTTCACTCCA 4100
TAGTTGCTG  ACTCCCGGTC GTGTAGATAA CTAAGCAGAC GGAGGCTGTA 4150
CCATCGTGGC CAAGTGCAGC AAATGATACG CAGAACCCAC GCTCACCAGGC 4200
TCCAGATTGA TCAGCAGTAA ACCAGCAGCC CGAAAGGGCC GACGCGAAGAA 4250
GGTGTCGTCG AACTTTATCC GCTCCATGCC AGTCTATTTA TTGGTCGCAAG 4300
GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT ATGTGGCGCA AGCTTGGTGC 4350
CATTGCTGCA GGATCGTGG TGTCAAGGTC GTCTTGTGTT ATGGCTTCTC 4400
TCAGCTCCGG TTCCACAGGA TCAAAGCCAG TTACATGATC CCCCATGTGG 4450
TGCAAAAAAO CGTTAGGCTC TTTGCTGTCT CGTACGTTG TGCAAGAATT 4500
GTTGGGGCAG CAAGTGTACT CCACTGGTTAT GGCAGCAGTG CATAATTCTC 4550
TTAAGTCTCAT GCACGCGTGTA AGTGTGTTTT CTGTGACGTTG TGAAGACTCA 4600
ACCAAGTCTA TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC 4650
GGCGTCAGCA CCGGATAATA CCCGGGCAAG TAGCAGAAGC TAAAGATGTC 4700
TCATGATTG AAAAGTGTCT CCGGCGGAAG AACCTCTCAAAG GATCTTACG 4750
CTGGTGAGAT CCGATCGTG AGAACACTTC GTGTACCCAC GTGCGACCCA ACTGACTCTC 4800
AGCATCTTTT ACTTTCACCA CGGGTTCGGT GTGACGAAAA ACAGGAAGGC 4850
AAAAATCGGG AAAAAGGGA ATAGGCGGCA CACGGAAATT TGGAATACTC 4900
ATACCTTTCC TTTTCAATA TTATGGAAGC ATTTTACAGG GTTATTTCT 4950
CAGCAAGGGA TAGATATTTA AATGTATTTA GAAAAATATA CAAATAGGGG 5000
TTGCCGCGCA ATTTCCGGCA AAAAGTGCCAC CTGACGCTCT AGAAACCATT 5050
ATTATCAGA CATTAACCTA TAAAATAGGG GTATACCGA GGCCCTTTCG 5100
TCTTCAAGAA 5110

(2) INFORMATION FOR SEQ ID NO:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 5110 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: double
   (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
   (A) ORGANISM: Not applicable
   (B) STRAIN: Not applicable
   (C) INDIVIDUAL ISOLATE: Not applicable
   (D) DEVELOPMENTAL STAGE: Not applicable
   (E) HAPLOTYPE: Not applicable
   (F) TISSUE TYPE: Not applicable
   (G) CELL TYPE: Not applicable
   (H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
   (A) LIBRARY: Not applicable
   (B) CLONE: pCM-T297

(viii) POSITION IN GENOME:
   (A) CHROMOSOME/SEGMENT: Not applicable
   (B) MAP POSITION: Not applicable
   (C) UNITS: Not applicable
FEATURE:

(A) NAME/KEY: Inactive T7 φ10 promoter mutant (-11G to T)
located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyltransferase (CAT) gene.

(B) LOCATION: 198 to 220

(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.

(D) OTHER INFORMATION: If the mutant T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase or a mutant T7 RNA polymerase, CAT can be expressed from the plasmid.

PUBLIC INFORMATION:

(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.

(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters

(C) JOURNAL: Biochemistry

(D) VOLUME: 32

(E) ISSUE: 35

(F) PAGES: 9115-9124

(G) DATE: Sept 7, 1993

(H) DOCUMENT NUMBER: 57

(I) FILING DATE: 58

(J) PUBLICATION DATE: 53

(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110

SEQUENCE DESCRIPTION: SEQ ID NO: 6

TTCCAGGCAA TCAAAATAAA CGAAAAGGCTC AGTCAAAAGA CTGGGCTTCTT 50
CGTTTTATCT GTGTTTTGTC GTGAAACGCT CTCCTGAGTA GGACAAATCC 100
GCCGGGAGCG GATTGGAACG TTGCGAAGCA ACGGCCCGGA GGGTGGCGGG 150
CAGGACCCC GCCATAAACT GCCAGGGAAT TCCCCTAGTA CTGAAATTAA 200

T7 Promoter Positions - 15

TACTACTCAÇ TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCCTT 250

-10 -5 +1 +5

CGACGAGATT TTTACGGAGCT AAGGAAGCTA AAATGGAGAA AAAAATCACT 300
GGATATACCA CCTGGTACTTC CTTGGGTACA AATTCCTTGA GATAGCTTGG 350
GGCAATTCGCC TagsGGATCA TATATTTTGA CCGATTAAGA AAAATCACT 400
ATATCAGGTC TTTTTTTAAG ACCGATTAAGA AAAATCACT 450
CCCGCCTTTTA TTTACATTCT TGGCCGCTTG ATGAATGCTC ATCCGGGAAT 500
CGTATGGGCA ATGAAGGAGC GTGAGCTTGG ATGTTTACCA 550
CCGAGTTACGT TTTTTTTTTAAG ACCGATTAAGA AAAATCACT 600

TGGAATATCG TACGGGCTTT TGGTACTTGA ATGGGAGGCTA AAATTCCTTGA 650

GGGAAATATCG TATACGCGGCA ATGGGAGGCTA AAATTCCTTGA 700

AGCATATATG TTAGGAAA ACTTGTCCTCA TTGTTTTTTAAG ACCGATTAAGA 750

TGTTTACGAA TACGGGCTTT TGGTACTTGA ATGGGAGGCTA AAATTCCTTGA 800

GGGAAATATCG TATACGCGGCA ATGGGAGGCTA AAATTCCTTGA 850

AGGTATCGTATG TACGGGCTTT TGGTACTTGA ATGGGAGGCTA AAATTCCTTGA 900

GGGAAATATCG TATACGCGGCA ATGGGAGGCTA AAATTCCTTGA 950

AGGCGTTATG TACGGGCTTT TGGTACTTGA ATGGGAGGCTA AAATTCCTTGA 1000

TATAAGGCA ATGAATTACG GAGGCAATCC TTTTCCTTGA GAGGCAATCC 1050

GTTTACCAATGCC TCGATTCCG ATGGGAGGCTA AAATTCCTTGA 1100
GCGAGCGGTA TCACTTCAGT CAAGGCCGGT ATACAGTTA TCCACAGAAT 3200
CAGGGGATAA CGCAAGGAAA AACATGTOAG CAAAAAGCCA GCAAAAGGCC 3250
AGAACCCTTA AAAAGGCCGC TTGGCTGGCG TTTTCCATA GGCTCCGCCA 3300
CCCTGAGGAG CATCACAAAA ATCGACGCTG TGGCGAAACC 3350
CGACAGGACT ATAAAGATAC CAGGCGTTTG CCCCTGGAAG CTCCCTCGTG 3400
GTGGGTGTGA GTGCTTTGTC TCAAGCTGGT GCTGTGTGCA GAAACCCTCC 3450
GTCAGCAGCC ACCGTCGCCG CTTATCCGCT AACATCGTCT TGGATCCCAA 3500
CCCGTTAAGA CACGACTTAT CGCACTGGGC AGCAAGCAGT GGTAAAGAGA 3550
TTAGCAGAGC GAGGTATGTA GGGGTGTTGA TAGATTTCTG GAAAGTTGCG 3600
CCTAACATCG GCTACATAG AAGGACAGTA TGTTAGTATC GGCTCTGCTG 3650
GAAGCCTGTA ACCTTTGGCA AAAAGTTGAG TAGCTCTTGA GGCTGAAACC 3700
AAACGACCAG TGGTATGCGT GGGTGGTTTG TTGCAAGAG CAAGTACAGG 3750
CCAGAGAAAA AGGAATCTCA AGAAGATCTC TGGATTTGCA GTCATGAGAT 3800
TGACGCTCAG TGGAGCGGAA ACTCACTGTA ATGGATTTTG GTCAAGTGGG 3850
TATCAAAAAAG GATCTTTACC TAGATTTTCA TAAATAAAA AGTAAAGGTT 3900
AAACTAATCT AAAAACTTAT CGCACTTCA GCCATGATCA GTTACAATTG 4000
CTTAATCAGT GAAGGACCTA TCTCACGAGT CTGCTCTTAT CTGACATCCA 4100
TAGCTCGTGT ACCTCCCGTC GTGTAGTAAA CTACGTTAGC GGAGGCTTTA 4150
CCATCTGGCC CAGGCTGCTC AGTCTATTCC TAGTGGTTGG CTGCTCTGCTG 4200
GTGGTGCGC TAAGCTTGTG CTGCTTTCGC AGGTGATTTG GTGACTTCTC 4250
GAAGATAGAG CAACTCGCTG CCCCAGGCCA AGGATTTAGG TATGTTGCCG 4300
CATTGCTGCA GGGATCTGTT GTCTCACGTT GCTGCTTGGGT TTAGCTTCTC 4350
GAAGCTGAGG TAAAGTGGTC GCAGCTTAGA GTGTTGCTTG GACGCTGGTC 4400
TGGATGCTGA AAGCTTTACCC TGGCGATATC CCAAGGCGGA GCTGATTTTA 4450
TACATATCTA ACGAGATGCA GTGCACTTGC ACCGACGTGA ATGGCTTCTG 4500
AGCGCTGCAA GTGGATGCTG CAGATGATCA TGGATTTCTG GTACGAGTGC 4550
TTACTTTCTC GCATCCTCTA AGAATGCTAA CACGATCTTCA TGGAACACTA 4600
ACCAAGTCTT TCTGAGTACG GTGATGCGG CAGGAGGCTT GCTCTTGCCC 4650
GGGGTCAAAC CGGCCGTTTT CTTGTGATGT CAGCAAGCCT CTTGTGCTGC 4700
TCACTCTGGA AAAACGTTCT TCGGGCGCAA AACTCTCAAG ATGAATCTTG 4750
CTGTGCTAGT CAGAGGTCTG GTACCCCTCT ACGTACCTGC GCTACCTCTC 4800
AGCATTTTTT ACCCTTCATG GCGTTTCCGG GAGCTGACAA ACGCAAGGCG 4850
AAAAATGCCGC AAAAAAGGGA ATAGAGGGCA CACGAAATCG TTGATACCT 4900
ATACATTCTC TTTTTCAATA TATATGAAAG ATATTTCTCA GTGATGCTCT 4950
CATGACGGAG TACTATTGG ATAGTTTTA GAAATAAAA CAATAGGGG 5000
TTCGCGCAGC ATTTCCCGGA AAAAGTGCCAC TGGCACTTGA AGAACCATT 5050
ATTATCATGA CATTAACTCTA TAAAAATAGG CAGTACAGGA GCCCTTTCG 5100
TCTTCACAGAA 5110

(2) INFORMATION FOR SEQ ID NO:7:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-P1160

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Inactive T7 φ10 promoter mutant (-9C to G)
located between nucleotides 198 and 220 and in front of
a promoterless chloramphenicol acetyl transferase (CAT)
gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from
pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: If the mutant T7 promoter between
positions 198 and 220 is recognizable by T7 RNA
polymerase or a mutant T7 RNA polymerase, CAT can
be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7
RNA Polymerase that Recognizes an Expanded Range
of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7
TTCCAGGCGA TCAAATAAAA CGAAAGGCTC AGTCGAAAGA CTGGGCCTTT 50
CGTTTTATCT GTGTTTTGTC GGTGAACGCT CTCTGAGTA GGACAAATCC 100
GCCGGAGCGC GATTTGAGC TGGCAGAGCA AGCGGGCGGA GGGTGCCGGG 150
CAGGACGCCC GCCATAAAGT GCCAGGGATA TCCCCCGAGTA CTGAATATTA 200
T7 Promoter Positions

TACGAGTCAC TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 250

CAACGAGGATT TCCAGGAGCT AGGAAGGATT AAATGAGGAA AAAAATCAGT 300
GGATATACCA CCGTTGATAT ATCCCAATCG CATCGTAAAG AACATTTTGA 3
GGCATTTCAG TCAAGTGCTC AATGTACCTA TAAACAGACC GTTCAGCTTG 4
ATATTACGCC CTTTTTAAAG ACCGTAAGGA AAAATAAGCA CAAGGGTAT 5
CCCCCATTTA TTTACACTCT TGCCGCTCTG ATGAAGCTCT ATCAGGAAAT 6
CGATAGGCA ATGAAGACCG GTGAGCTGGT GATATGGGAT AGTGTTCACC 7
CTTTTACAC GCTTTACATC GAGCAGGACT AACAGTTTTT ATGCTGCTGG 8
TGCTGGGCTG TACGTTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTTAG 9
AGAATATGT CTTGGCTCTCA GCAAATCCCC GCTGTAGATT CATGGGCATC 10
GGGCAATAT TATACCAATT GGCCAAACTTC TACCACCCCGTT TTTACCATC 11
AGGTTCACCA TCCGCTGCTG GATGCTGGCC ATGTCGCTGC GCTTTAAGGC 12
AATCGGCAA CCAACCCCTT TAGAAGAATACT TCCATCCGCT CGGCGACTCC 13
CAGCGCCGCA ACACGCAGCG GCTGGCGCCTG TGGCTGGCAG CTGAGGCGCG 14
TTTCAGCCGT AATGACCCCT TCCCCCACCTG GCAGGATGGG AATGCGCTCC 15
AATCGGCAGA AATGGGCGAG GAGGCTCCCG GTGCTGGGTT TTAGGTTTTG 16
CAGATAGGGA GCTGCGCCTG GAGATGGGAT TTAGGCGGTG CACACGACCC 17
TGGCGGTACG TCATGCGGGA AATGCGCTTC TGGCGGTGGA GTGAGGGGTA 18
CGGAGGGTG TGGCGGCTGA CAGGAGCTAG GAAAGCGGCG TGGCGGTGGA 19
TGGCTGGGCT ATGGGCTGGC TAGGAGGCTG TGGCGGTGGA CAGGAGGCGT 20
GGGATTGGA GAGGAGGGTG TGGCGGTGGA GAGGAGGGTG TGGCGGTGGA 21
AACTGACGCA AATGCAGGCA GAGGAGGGTG TGGCGGTGGA GAGGAGGGTG 22
AGGTTTGTTC TGGCAGGCTG GTGCGGCTGA CAGGAGGCGT GAGGAGGGTG 23
AAACTGCAGA TATCGAGAAG GTGAGGCTGG TGGGCTGGCG CTGGCTGGCG 24
GGGGTGATTG GAGGAGGGTG TGGCGGTGGA GAGGAGGGTG TGGCGGTGGA 25
TCGCTGGCAT ATCAGATAGA CAGAAGCCAG TGGCGGTGGA GAGGAGGGTG 26
GGGATTGGA GAGGAGGGTG TGGCGGTGGA GAGGAGGGTG TGGCGGTGGA 27
AACTGACGCA AATGCAGGCA GAGGAGGGTG TGGCGGTGGA GAGGAGGGTG 28
TCGTTTCCG TGTTCCGAAA AGTCTGGAAA CGCGGAAGTC AGCGCCCTGC 2400
ACCATTATGT TCCGGATCTG CATCGCAAGA TGCTGCTGGC TACCCCTGCG 2450
AACACCTACA TCTGTATTAA CGAAGGCTG TGATGACCC TGAAGGTATT 2500
TTTCTGGGTC CCGCCGGAGT CATACCGCAA GGGTTTACC CCCACAACGT 2550
TCCAGTAAACC GGGCAGTTC ATCATCAGTA ACCGCTATCG TGAGCCTCAT 2600
CTTCTGGTTC ATCGGTATCA TTACCCCAT GAACAGAAAT TCCCCCTTAC 2650
ACGAAGGGCAT CAAGTGAAGA AACAGAAAAA AACGCCTCCC AACATGGGCA 2700
GCTTTATCGA AGGCCAGACA TTAACGCTTC TGGAGAACT CAACGAGCTG 2750
GACGCAGAGA AGCAAGCCAG CATCTGTGAA TGCTGCTGGC TACCCTGTGG 2800
ACCCCTCTGG CCGCCGATCC CATACCGCAA GTTGTTTACC CTCACAACGT 2850
TCCAGTAACC GAGCGGCTCA GCTCTGGACC GCCGCTTACC GGATACCTGT CCGCCTTTCT 3000
GCTTTATCGA AGGCCAGACA TTAACGCTTC TGGAGAACT CAACGAGCTG 3050
GCTTTATCGA AGGCCAGACA TTAACGCTTC TGGAGAACT CAACGAGCTG 3100
GACGCAGAGA AGCAAGCCAG CATCTGTGAA TGCTGCTGGC TACCCTGTGG 3150
ACCCCTCTGG CCGCCGATCC CATACCGCAA GTTGTTTACC CTCACAACGT 3200
TCCAGTAACC GAGCGGCTCA GCTCTGGACC GCCGCTTACC GGATACCTGT CCGCCTTTCT 3300
GCTTTATCGA AGGCCAGACA TTAACGCTTC TGGAGAACT CAACGAGCTG 3350
GCTTTATCGA AGGCCAGACA TTAACGCTTC TGGAGAACT CAACGAGCTG 3400
GCTTTATCGA AGGCCAGACA TTAACGCTTC TGGAGAACT CAACGAGCTG 3450
CCTTCTGGGA AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3500
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3550
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3600
CCTTCTGGGA AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3650
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3700
CCTTCTGGGA AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3750
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3800
CCTTCTGGGA AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3850
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3900
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 3950
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4000
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4050
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4100
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4150
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4200
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4250
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4300
GTTCTGTGGC ACCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA 4350
CATTGCTGCA GCCATGTGGG TGTCAGCGCTG TGCAGTTGTG ATGGCTTTCAT 4400
TCAGCTCCGG TTCCCAACGA TCAAGGCGAG TTACATGATC CCCCATGTTG 4500
TGCAAAAAAG CGGTGACTGT CTTCGGTCCT CGGATCGTTG TCAGAAAGTA 4500
GGTTGCCGCA GTGTATGCGG CG ACCGAGTT GCTCTTGCCC 4500
TTAAGTTCTC TCCGGTTGGT TGCAGATCGG TGAGTACTCA 4600
ACCAATGTCAT TCTGAGAATA GTGTATGCGG CG ACCGAGTT GCTCTTGCCC 4600
GGGTCACCAA CGGGTACTAC CCAGCCGACA TACGAGAACT TAAAAAGTGC 4700
TGTATTGG AAAAGCTTCTT TCGGGCGAA AACTCTCAAG GATCTTACCG 4750
CTGTGTGAGAT CCAGTGCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC 4800
AGCATTTTT ACTTTCACCA GGCTTTCTGG GTGAGAAAAA ACAAGAAGGC 4850
AAAATGGGGG AAAAAAAAGGA ATAAAGGCGA CACGGAAATG TTGAATTCTG 4900
ATACTCTTCC TTTTTCATA TTATTTGAGC ATTATCAAGTT GTTATGCTT 4950
CATGACGGCA TACATATTGG AATTGATTTT GAAAAATAAAA CAAAATAGGG 5000
TTTCGCGCAC ATTTCCCACA AAAGTGCCAC CTGACGTCTA 5100
ATTATCATGA CATTACCTA TAAAAATAGG COTATACGA GGCCTTTCG 5150
TCTTCAAGAA 5110

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(iii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-T270

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Inactive T7 +10 promoter mutant (-9C to A) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: If the mutant T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase or a mutant T7 RNA polymerase, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters

C) JOURNAL: Biochemistry

D) VOLUME: 32

E) ISSUE: 35

F) PAGES: 9115-9124

G) DATE: Sept 7, 1993

H) DOCUMENT NUMBER:

I) FILING DATE:

J) PUBLICATION DATE:

K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110

( i x ) SEQUENCE DESCRIPTION: SEQ ID NO:

TTCCCAGGCA TCAATATAAA GGAAGGCTC AGTCGAAGGA CTGGGCGCTT 50

CGTTTTATCT GTTGGTGTGC GGTGAACGCT TCCCTAGTA GGAAAAATCC 100

GCCGGGACGC GATTTGAACG TTGCGAAGCA ACGGCCCGGA GGGTGGCGGG 150

CAGGACGCCC GCCATAAACT GCCAGGGAAT TCCCCTAGTA CTGAAATTAA 200

T7 Promoter Positions

TACGAATCAC TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 250

-10 -5 +1 +5

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 300

GGATATCACT CCGTGGATAT ATCCCAATCG CATCGTAAAG AACATTTTGA 350

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 400

ATATATAC TTATGTTAAG ACCCAAAAGA AAAAATAGCA CAAAGTTTAT 450

GGGCGCCTTA TCTGATCTTC GGCCTCCCTG ATGAAATGCT ATCGGAATT 500

TGTGGCGTGT TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 550

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 600

GGGCGCCTTA TCTGATCTTC GGCCTCCCTG ATGAAATGCT ATCGGAATT 650

ATGTTTACAC CCTCTTAAAT GACGAAAATC AAGGTTTCCG AATGCTTATT 700

GGGCGCCTTA TCTGATCTTC GGCCTCCCTG ATGAAATGCT ATCGGAATT 750

AGATGATGTT TCGCTCTCGA CTGCTGGATG TGGTGAGTTT CACCAGTTTT 800

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 850

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 900

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 950

AATGCAATAC ATGACTGCGA GCAATGCGAT GGCACGCGCC 1000

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 1500

TACGAATCAC TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 2000

-10 -5 +1 +5

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 2500

GGATATCACT CCGTGGATAT ATCCCAATCG CATCGTAAAG AACATTTTGA 3000

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 3500

ATATATAC TTATGTTAAG ACCCAAAAGA AAAAATAGCA CAAAGTTTAT 4000

GGGCGCCTTA TCTGATCTTC GGCCTCCCTG ATGAAATGCT ATCGGAATT 4500

TGTGGCGTGT TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 5000

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 5500

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 6000

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 6500

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 7000

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 7500

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 8000

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 8500

GGGAACTTGA TACGGTGAAA ACCTGGCCTA TTTCCCTAAA GGGTTTATTG 9000

AGGAAATGTT TTTCGTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 9500

AATGCAATAC ATGACTGCGA GCAATGCGAT GGCACGCGCC 10000

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 10500

TACGAATCAC TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 11000

-10 -5 +1 +5

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 11500

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 12000

TACGAATCAC TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 12500

-10 -5 +1 +5

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 13000

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 13500

TACGAATCAC TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 14000

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 14500

CGACAGGATTT TACAGGAAGCT AAGGAAGCTA AAAATGGAGAA AAAAATCACT 15000
CCGGAGGGTG GCGGCGAGGA CGGCGCCCAT AAACGCCAG GCATCAAATT 1550
AAACGAGAAG CCACTCTGAC GGATGCGCTT TTTGCGTTTC TACAAAATCT 1600
TCCTTGCTGTC ATATCTACAA GCCATCCCCCC CACAGATACG GTAAACTAGC 1650
CTGTTTGGCA CATCAGGAAA GCAGCTGTTT TGGCGAGATTAA 1700
TCAGGTGAT GTAGTTTATTTA TCAGAAACGA GAAGCGGCTG ATGAAAACAG 1750
AAATTGTCCGTC GCGGCGAAGGT CGCGTTGCTC CCAACCTGACC CCAATCGGAAG 1800
CTCAAGACTGG CAAAGCCTGTA GCAGCGATGG TAGTGTGGGT TCTCCCATG 1850
CGAGAGTGAA GAACTGCCAG GCAATTAAATA AAACGAAAGG CTCAGTCGAA 1900
AGACTGGGCC TTTGTTTTTA TCTGTTTGT GTCGTTGAAAC GCTCTCCTGA 1950
GTAAGAAAGA TGCACCCGGAA GCGGATTTGA GCAACGGCCC 2000
GGAGGGTGAC GGGGACAGAC CGGCGCCATAA ACTGCCAGGC ATCAAATTAA 2050
GCAAAGAGGC ATCCTGACCG ATGCGCTTTTT GCAGTGTCTTA CAAAATCTTC 2100
CTGCTCTCAT ATCTACAAAGC CATCCTCCCA CAGATACGGT AAAAGCAGCT 2150
GCTTTTCCGA TCAAGAAAGG CTAACGCCAC CTGATGTCAGG TCGCGCAGAG 2200
TGGCAGATGT GTGGCTCCTGC TCGGGGAAGGC CCGGCTTGAAG CTGGCGGGGT 2250
TGCCCTACTG TTATGACAGAA TGAATCAGCG ATACGACGGC GAAGCTGAGA 2300
CGACTGAGCG TCGAAGAAGCT CTGGGAGGCT AGCACAACGA TGAATGGCTC 2350
CTGTTTGCAC AGTCTGAAAGG CAGGGATCAGG TGGATACGAA 2400
ACCATTATGT TCGGTGATCTG CATGCGAGGA TGCTGTGCGG TACACCTGGA 2450
AACACCTACA TCTGTATTTAA GAAAGCCTGG TGTAGGAAAAC 2500
TTCTCTGGTC CCGCGCCATC ATACCCCGCA GTGGTTTACCG CTCACAAGGT 2550
TCCAGTAACC GGCTGAGTCC ATCATCAGTA ACCGGATACG TGAGCATCTC 2600
CTTCTGGTTC ATCGTATCAG TACCACCATG GAACAGAAAT TCCCCCCCC 2650
ACGGAGCATC CAAGTGACCA AACAGGAAAA AACGCGCCTT AACATGGCCC 2700
GCCTTACAGC TACACAGGATG TGGAGAAGAT CAAACGCGCT 2750
GACGGGAGCGA CAAAGCAGCA ATCGTGGGAA TTGTTGTTAC CTCACAAGGT 2800
TGGACTTCCG CAGAGCGGTC ATACCACGGCA GGTACCCAGC GCTACCGAGA 2850
TCTGACACTG CAGTGCCCTC AGACACCGTC ACGTGTCTTG GTAACCGGAT 2900
GCCGGGAGCA GACAGACCCCCCC GACGAGGCGG TCGCGGGGTG 2950
TGCGGTGCGA GCCATGAGCTG CAGTATCAGG GATGTATCGI 3000
GCTAACTAT GCCAGATCATG AGCAGATTTG ACTGAGAGTG CACCATATCG 3050
GGGTGAAATT ACCGACAGAG TGCGTAAAGAA GAAATCTGCC CATCAGGAGC 3100
TCTTGCGCTT TCCGCGCTCTG TGACCTGGCG CGCTGGTCGTG TCGGCTCGCG 3150
GCGAGCTGTA TCAAGCTCAGC CAAAGGCGGT AATACGGTCT TCCAACAGAT 3200
CAAGGGATAA CCGAAGGAAA AACATGTGAG CAAAAGGCCA GCAAAAGGCC 3250
AGAAACGCTA AAAAAGCCGCG GTTGTGCGGG CTTTTCCATC GGTCTGCCCC 3300
CCCTGACGAG CATCACAATA ATCGACGCTG AAGTCAAGGG TGGCGAACC 3350
CGACAGGCTA AATGAAGTAC CAGCGCTTCC GCCTGGAGAG CTCCTCGCTG 3400
CGCTCTGCTG TCCGGACCTC GCGGCTCCTG AGATACCTG GCTCGCTTCT 3450
CCTTGGGAGA AGCGGTCGGG TTTCTCAATG CTCAAGCTG AGGTCTTCTC 3500
GTTTGCTGTA GTCCGTCGTC TCCAAAGCTG GCTGCTGCGA CGAACCCCCC 3550
(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 5110 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: double
   (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYP: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELL: Not applicable

(viii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-1087

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Inactive T7 φ10 promoter mutant (-9C to T) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter"was sequenced.
(D) OTHER INFORMATION: If the mutant T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase or a mutant T7 RNA polymerase, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110

(ix) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
TTCCAGGCA TCAAAATAAA CGAAAAGGCTC AGTCGAAAGA CTGGGCCTTT 50
CGTTTTATCT GTGTGTGTGC GGTAAGCCTG CTCTGTAGTA GGACAAATCC 100
GCCGGGAAGC GATTTGAACG TTGGCAAGCA AGGCCCAGGA GGGGCGCCGG 150
CAGGACGCCC GCCATAAACG GCCAGGGAAT TCCCCTAGTA CTGAAATTAA 200
T7 Promoter Positions 250
   * * 5
   -10 -5 +1 +5
 CGACGAGATT TTCCAGGAGT AAGGAAGCTA AAATGGAGAA AAAAATCACG 300
GGATATACCA CGGGTATATG ATCCCAATCG TATGGTAAAG AACATTTTGA 350
GGCATTTTAC TCCATTGGTC AATGATACCT TAACCCAGGCG GTCAAACTGG 400
ATATTGGCA CTTCTTTTAA ACCTGAAAGA AAAATAAGCA AAAGTTTTAT 450
CGGCCCCCTT TTACCATTCT GCAGCCCGCT ATGAGTGCTC ATCCGAATTG 500
CGATATGCGA ATGGAAGCAG ATGAGCTGGG ATGTTTGCAC 550
CTTGTACACT GGTGTGCCC TACAAAGCTG AAACCAGCTT ATCAGCTT GG 600
AGTGAAATCC ACGGAGATT TCCCCGCTG ATCCGTATTTGGTAA 650
TGTGCGCTGT TACGGTAGAA AACCTGCTA TTTCCCATCA GGTATTTTGGT 700
```
<table>
<thead>
<tr>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGAATATGTT TTTCGTCTCA GCAAATCCCT GGGTGAGTTT CACCAGTTTT</td>
</tr>
<tr>
<td>GATTTAACGG TGCGCAATAT GGCAACCTTC TCCGGCCCCG TTTTCACCAAT</td>
</tr>
<tr>
<td>GGGCAATATAT TATACGCAAG GCCAAAGGTT GCTGATGCCG CTGCGGATTTC</td>
</tr>
<tr>
<td>AGGTTCTCATCA TGCGCTGCTGT GATGCGTTCC ATGCGGCAAG AATTCTTTAAT</td>
</tr>
<tr>
<td>GAATTTACAC AGTACTGCGA TGAGTGCGAG GCGGGGCTGT AATTTTTTTTA</td>
</tr>
<tr>
<td>AGGCAGTCTAT TGTCGCCCTT AAAAACTGGTT TGGATGCTCT GAATAGGTTA</td>
</tr>
<tr>
<td>TAAATAAGCGG ATGAATGGCA GAAATTCTGC GAGGCAGCCT ACCTGATCTATG</td>
</tr>
<tr>
<td>GATTCCACCACT TCCAGAAATT GGAGCCAAAT ATTCTTTGCG GAGAACGTG</td>
</tr>
<tr>
<td>AATGCGGACA CCAAACCTTTG GCAAGACATA TCCATCCTGT CGGCACATCCT</td>
</tr>
<tr>
<td>CAAGCAGCGCC TGCCGGCGCA TCTCTGGCTGT TTTTGCGGAT GAGAGAAGAT</td>
</tr>
<tr>
<td>TTTCACTCTG ATACAGATTA AAATCAACAG CAGAAGCGGT CTGATAAAAAC</td>
</tr>
<tr>
<td>AAGAATTCGCC TGCCGGGCGCA TCTCGGCTGT TTTGGCGGAT GAGAGAAGAT</td>
</tr>
<tr>
<td>GGCAAATAT TATACGCAAG GCGACAAGGT GCTGATGCCG CTGGCGATTC</td>
</tr>
<tr>
<td>AGGTTCATCA TGCCGTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT</td>
</tr>
<tr>
<td>GAATTACAAC CGTACTGCGA TGAGTGCGAG GCGGGGCTGT AATTTTTTTTA</td>
</tr>
<tr>
<td>AGGCAGTCTAT TTGAGCTCTT AAAAACTGGTT TGGATGCTCT GAATAGGTTA</td>
</tr>
<tr>
<td>CCGGAGGGTG GGCAGGCGGA GCGCCGCAT AAAACTGCAAG GCATCAAATT</td>
</tr>
<tr>
<td>AAGCAGAGGA CCATCCTGAC GCCAAATCTT GGAGCAAATT GAACGTTGCG</td>
</tr>
<tr>
<td>TCTGCTCTGC ATACGACATC GACTCTTCCC CACAGATACG ATCAAATGCA</td>
</tr>
<tr>
<td>CTCGTTTCTG CATCAAGAAA GCACTGTATTTC TGCGGAGTGA GAGAGATTTT</td>
</tr>
<tr>
<td>TCGCCTGTAT ATACAGATTA AAATCAACAG CAGAAGCGGT CTGATAAAAAC</td>
</tr>
<tr>
<td>AATTTTCCTG CGCCGAGTAG GCGGCTGCTTC CACAGCCGAG CACAGCAGCGA</td>
</tr>
<tr>
<td>CTCAGAAAGTG AAACCGCGTA GCGCGATGGA TAGTGGTGCG GCTTCCCCATG</td>
</tr>
<tr>
<td>CGAGAGTAGG GAATCTGCGA GCATCAAATA AAACGAAAGG TCCAGTCGAAG</td>
</tr>
<tr>
<td>AGACTGCGCC TTTGTTTATT TCTGTGTGTTT GTCGCGTACT ACTAAAACAG</td>
</tr>
<tr>
<td>GTAGAAGAAA CACCCCGGGA GCGGGATTGA CGTTTGGGCG GAAACGGCAGC</td>
</tr>
<tr>
<td>GGAGGGTGGC GGGAGGGAGG CCGCCGCTAA ACTGCGGAGC ATCAAATGAA</td>
</tr>
<tr>
<td>GCAGAAGGCC ATCTCTCGCA AGCACTGATTTC TCGGTTTCTCA CAAACTTCTT</td>
</tr>
<tr>
<td>CTCCTGCTAT ATCTACAAGC CATCCCCCG AAGATACGCT AAACCTACGCT</td>
</tr>
<tr>
<td>CTTTTTTCGA CAGAGAAACC AGTCTGCGAG GGTGGCTGTC TGGCAACGGG</td>
</tr>
<tr>
<td>TCCGATCATG CTGCTCTCTG TCGTGACTGG CCGCGCTAAG CTGCGGGGCT</td>
</tr>
<tr>
<td>TCCGGCCATT GTGATCTCGA TGATCTGACTG GCGAGAAAGC TCGGCGCTGG</td>
</tr>
<tr>
<td>CGACTGCTGC TGGAAAACTT CTGCGACCTG ACAGCAACAT TGAATGCTCT</td>
</tr>
<tr>
<td>TCGCTCTCGC ATACGACATC CACAGACATT GAACGCTCAAT TGGCGTTTCTA</td>
</tr>
<tr>
<td>TGCTTACCTG GTATGCGAGCC ATCGTGAGCC GAGTCAGAAC AAGATGCTAG</td>
</tr>
<tr>
<td>ACAACTACCA TACTGATTAAA CAAAGGCGTG GCAGACGATT GCGAGGCTTG</td>
</tr>
<tr>
<td>TTCTCTGAGT TCCGAGCTGAG CATCGCAAGA GTGCGTTGTTT CACCTCTCTTCA</td>
</tr>
<tr>
<td>ACGAATTCAG GCAGCGGCGA TACATGCTCTC TCGGAGAAACT GAGAGACTCT</td>
</tr>
<tr>
<td>TCCAGATTTA TCCAGTCTCTG GCGAGATC GCAGGCAGCT GCAGCGGCGA</td>
</tr>
<tr>
<td>ACTCCGTTTC ATCGCGCTAT CCACACCAAT GAGAGAATG TCCCCCTTAC</td>
</tr>
<tr>
<td>AGGGAGCGAT CAACTACTTG AAAGAGAAAG AAAGGCCTCT AAGATGCGGGC</td>
</tr>
<tr>
<td>GCTTTATCAG AAGGCGACCA TAAACGTCTG TGGCGAGACT GAGGAGCTTG</td>
</tr>
<tr>
<td>81</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>GACGCAGATG AACAAGCCAGA CATCTGTGAA TCGCTTCAG ACCACGCTGA</strong></td>
</tr>
<tr>
<td><strong>TGAQCTTTAC CGACGCTGCC TCGCGCGGTTC CGGTGATGAC GGTGAAAACC</strong></td>
</tr>
<tr>
<td><strong>TCTGACACAT CGACGCTCCCC GAGACGGTCA CAGCTTTGCT GTAAACGGGAT</strong></td>
</tr>
<tr>
<td><strong>GCCGGAGGCA GACAAACCGGC TAAGCAGGGG TGAGCGGATG TGAGCGGGATG</strong></td>
</tr>
<tr>
<td><strong>TGGGGGCAGCA GCCATGACCC AGTCACGTTAG CGATAGCGGA GTGTATACTG</strong></td>
</tr>
<tr>
<td><strong>50</strong></td>
</tr>
<tr>
<td><strong>GCTTAACATAT GGGCGCATGAG AGCAGATTTG ACTGAGATGG CACCATATGC</strong></td>
</tr>
<tr>
<td><strong>GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG</strong></td>
</tr>
<tr>
<td><strong>2900</strong></td>
</tr>
<tr>
<td><strong>GACGCGGATG TAAGCGCGGT AATACGTTAA TCCACAGAAT</strong></td>
</tr>
<tr>
<td><strong>CAGGGGATAA CGACGGAAAG AACATGTTAG CAAAGGGC</strong></td>
</tr>
<tr>
<td><strong>AGGAAACCGTA AAAAAAGGCG GTGGCTGGCG TTTTTCCATA GCCTCAGCCC</strong></td>
</tr>
<tr>
<td><strong>CCCTGACGAG CATCACAAGAA ATCGAGCGATG TGAGGAAACCC</strong></td>
</tr>
<tr>
<td><strong>CCGACAGACT ATAAAAGATAC CAGCGGTTTT CCCTGGGAAG CTCCCTGGCT</strong></td>
</tr>
<tr>
<td><strong>CCCTTCGGGA AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA</strong></td>
</tr>
<tr>
<td><strong>3550</strong></td>
</tr>
<tr>
<td><strong>GTTCAQCCCG ACCGCTGCCG CTATACCGGT TTGAGCTCCAA</strong></td>
</tr>
<tr>
<td><strong>CAGGGGATAA CGACGGAAAG AACATGTTAG CAAAGGGC</strong></td>
</tr>
<tr>
<td><strong>AAACCCCGGC TGAQAACGCT GTGTTTTTGT TTTGCAAGCA GCAAGTACC</strong></td>
</tr>
<tr>
<td><strong>3800</strong></td>
</tr>
<tr>
<td><strong>GCCGAGTTA CAGACCTATAT GCCCATGCTGG AGCAGGGACC</strong></td>
</tr>
<tr>
<td><strong>TTAGCAGAGC GAGGTATGTA GGCCTGGCTTA CAGAGTTTCT GAAATGAG</strong></td>
</tr>
<tr>
<td><strong>TATCAAAGAG GATCTTACC TAGATCTTTT TAAATTAAAA ATGAAGTTTT</strong></td>
</tr>
<tr>
<td><strong>AAATCTACTC AAAAGATATA TGAAGAAACCT TGGCTTGCTA ATGCTAT</strong></td>
</tr>
<tr>
<td><strong>CTTATCAGT GAGGCCAGCT GTCAGCGGAT CTGGGTATTT CTGGTAC</strong></td>
</tr>
<tr>
<td><strong>4150</strong></td>
</tr>
<tr>
<td><strong>TACATCGGCT TCCCCCGGAT GTGTTTACCT CTACGAGTAC GCGAG</strong></td>
</tr>
<tr>
<td><strong>TCCAGGATTA CAAGGCTCTTA ACCAGCTGCA CGCAGGGGCA GAGGC</strong></td>
</tr>
<tr>
<td><strong>4350</strong></td>
</tr>
<tr>
<td><strong>GGTGCTTCGC AACTTATCCCC GCCTCATATCC AGTCTATATA TTGTGGCCGG</strong></td>
</tr>
<tr>
<td><strong>4400</strong></td>
</tr>
<tr>
<td><strong>TAAGCTACG CAGACGAGTAT GGGCGTTGCTA ACTGTGT</strong></td>
</tr>
<tr>
<td><strong>4500</strong></td>
</tr>
<tr>
<td><strong>TACCTGAGCC ATGGAATTTG AGGGATTTTG GTCATGAGAT</strong></td>
</tr>
<tr>
<td><strong>TTACTGACT CATGAGTTTG GGGCGGTTG TAAGCTATCTA</strong></td>
</tr>
<tr>
<td><strong>ACCGGATCAT TGTGAGAATG ATGAGTTGG GGTGAGTTG GCTCGGG</strong></td>
</tr>
<tr>
<td><strong>4800</strong></td>
</tr>
</tbody>
</table>
AGCATCTTT TACTTCAACCA GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC 4850
AAAATGGCGC AAAAAAGGGA ATAAGGCCGA CACGGAAATG TTGGATACCT 4900
ATACTCTTTCC TTTTTCAATA TTATGGAAGC ATTTATCGAG GTTATTGCT 4950
CATGACGCGA TACATATTTT GAAATAATTA GAAATAGGGG 5000
TTCCGCGCAG AATTCCCAGA AAAGTGCCCA CTGACGTCTA AAAGACCATT 5050
ATTATCATG AATTAACTTA TAAAAATAGG GCTATCACGA GCCCCCTTCG 5100
TCTTCAAGAA 5110

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(iii) MOLECULE TYPE: other nucleic acid

(iv) HYPOTHETICAL: No

(v) ANTI-SENSE: No

(vi) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable

(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(viii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-P1198

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Intermediate T7 φ10 promoter mutant (4T to G)
located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220

(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: The mutant T7 promoter between positions 198 and 220 is weakly recognized by T7 RNA polymerase; consequently, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters

(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER:

(J) PUBLICATION DATE:

(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110
SEQUENCE DESCRIPTION:

<table>
<thead>
<tr>
<th>SEQ ID</th>
<th>NO: IO</th>
<th>TTCCCAGGCA TCAATAAAAA CGAAAGGCTC AGTCGAAAGA CTOGGCCTTT 50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>CGTTTATCT GTTTGTTCGT GGTGAAGGCT CTCTCGGAGTA GGAACAAATCC 100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GCCGGGAGGG GATTTGAAGC CTGTCGACCGA AGCGCAGCAGA GGTGTCGCGG 150</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CAGGACGCC GCAATAAAC TGGAGGGAAT TCCCTCAGTGA GTGAATATTA 200</td>
</tr>
</tbody>
</table>

- T7 Promoter Positions

| TACGACGCAC TATAGGGAGA AAGCTTGAGT AGGACAAATCGGCGACGCTT 250 |

-10 5 +1 +5 |

| CGACACGATT TACGAGGCTC AAGGAAAGCTA AAATGGAAGGGAAAAATCACC 300 |
| GATATACCA CGCTTTGATAT ATCCCAATACG CATCGTAAAG AAACATTTCGA 350 |
| GCAATTTTCAG CAACTGGCTCT ATATGACCCTA TAACCAGACC GTCAAGCTGG 400 |
| ATATATCGGC CTCTTTAAAG ACCGTAAAGAGA AAAAAATAAGCGAATGTTTAT 450 |
| CCGGCTTTTA TTAACATCTT GGGTCGCTGG ACAGATGAGT ATCCGGAATT 500 |
| AGTAAATACC CTGACAGATTT CGGCGAGTCTT ATCACATATA ATTCGCAAGA 650 |
| TGTGCGCTGT TACGTCGAAA ACCTCGCCTTA TTTCCCTAAA GGTGTTATTG 700 |
| AGAATATGTT TTTCTCTCTCA GCCAATCCCT GGCTGATTTG CACCATTTTTA 750 |
| GATTTAAAACG TGCCCAATAG GGCAACCTTC TCCGCCCTGGG TTTCCACAT 800 |
| GGGCAAAATAT TATAGCAGA GGCGCAAGGG TCTGATCGCCT CGGCGATTT 850 |
| AGGTTCATCA TGCCCTTCTG CATGCGTCCCT ATGCGCGGAGCG AATGCTTAAAT 900 |
| GAATTACAC TGTACTGCGA TGAGTCGAGA GTGGCGGCGGCT AATTTTTTGA 950 |
| AGGCAAGTTAT TGCTGCGGCTG AAAACGCGCTG TGCTACGCTTG GAATAGGCTA 1000 |
| TAATAAGCAG ATCAGGACAG GAGCGCGGCTG CTCATCGAGA CAGCGGGGTT 1050 |
| GATTCACAC TCCAGAAATT GAGGCGCAATC AATTCTTCCTG GAGAATGCTG 1100 |
| AATGCGCAA CAAAACCTGA GCAAGAACAATA TCCATCGGCT CCGCCATCTC 1150 |
| CAGCAAGCCGG AGCGCGCGCA TCTCGGCTGT TTTGCGGGAG GAGAGAGGAT 1200 |
| TTTCCGCTCT ATACAGATTA AATCGAGAAG CGGACGGCGGC CTCGCTAAAG 1250 |
| AGAAATTGCC TGCCCGGGCTG AGCGCGGCTG TCCACCTCTG GCCATCGGCG 1300 |
| AACTCAGAAG TGAAAGCGCCG TAGGCGCGAGT GTGAATGTGG GCCTCCTCCA 1350 |
| TGCCGAGGTAGT GGGCGAATCG GAGCATCAA AAAAAAGAAA GGCTACGTCG 1400 |
| AAAGACTGCG CTTTGCTGTT TATCTGTGTT TTGCTGCGGT AGCTCTCTCT 1450 |
| GAGTAGGCGA AATCCGCGCC GAGCGGTATT GAACTGCTGG GAAACAGCGC 1500 |
| CCGGAGGCTG GCGGCGAGGA CGCCGGCAGA AAACCTCGAAG GCAATCAAAT 1550 |
| AAGCAGAAGG CCACTGCTGC GATGCCGCTC TTTGCGGTTC ATACAACTCG 1600 |
| TCCCTGCCTCAC TATCTCACAG CCATCCTCCC CACAGATACG GTAATACTACG 1650 |
| CTCTGTTTGG CATCAGGAAA GCAGCTGTTT TGCCGATGTA GAGGAATATT 1700 |
| TCAGGCCTGAT ACAGATTAAA TCAGAACGCA GAAGCGGTCTG GTAAACACAG 1750 |
| AATTTGCGCTG GCGGCGTAGG GCGGCTGCTG CCACCTGACC CATGCGGAA 1800 |
| TCGGAAAGTG AACGCGCGCTA GCAGGGATGG TATGCTGGGG TCCTCCTCATG 1850 |
| CGAGATGAGG GAACTGCCAG GCAATCAAATA AAACGAAAGG CTCACTGCGA 1900 |
| AGACTGGGCCC TTTCGTTTTA TCTGTTGTTT TGCAGTTGAAC GCTTCCTCTGA | 1950 |
|----------------|-----------------|-----|
| GTAGGACAAA TCCGCCGAGGA GCGGATTTGA ACGTTGCAAA GCAAACGGCCC | 2000 |
| GGAGGGTGGC GGGCAAGACG CCGGCATAAA ACTGCCAGGC ATCAAATTTA | 2050 |
| GCAGAAGGCC ATCCTGAGCG AGGCCCTTTT TGCCATTCTA CAAACTCTCTC | 2100 |
| CTGTGCTCAAT ATCTACAAGC CATCCCCCCA CAGATACGC ATAACTAGCT | 2150 |
| CTGTTTGGCA TCAGGAAAGG AGTCGGGCAG CTGTGGTCTT TGGCCACGGG | 2200 |
| TGCCCATGAT CGTGTCCTTG TCAGTTGAGG CCGGCTTACG TGCCGGGGTT | 2250 |
| TGCCCTTACTG GTTAGCAGAAA TGAATCAGGC ATACCGGCAG GAACGCGGTA | 2300 |
| CGACTGCTGCT GCACAACAGTG TCTGGACCTG AGCAGAAACCA TGAATGCTCT | 2350 |
| TGCGTTTCCG TCTTTCTGAA ACGTCTGAAA GCAGCGCTGC ACAGGCACTG | 2400 |
| ACCATTATGTT TCCGGATCTG CATCGCAAGGA TGCTTGCTGC TACCCTGTG | 2450 |
| AACCACCAACA TCTGTATTAA CAGAAGCGCTG GCATTGACCC TGAGTGATTT | 2500 |
| TCGGTTTCCGA TCAGGAAAGG AGTCGGGCAG CTGTGGTCTT TGGCCACGGG | 2550 |
| GGAGGAGGTTA CAGGAGGTTA TGGAGAAACT CACGGACATT CACCATATGC | 2600 |
| TCTCGTTTTC ATCGGTATCA TTACCCCCAT GAACAGAAAT TCCCCCTTAC | 2650 |
| ACCGGAGGCT CAAGTGACCA AACAGGAAAA AACCGCCCTT AACATGGCCC | 2700 |
| GCTTTATCCA AAGGCAGACA TTAACGCTTC TGGAGAAAACT CACAGAGCGT | 2750 |
| GACCGGGATG AACAGCAGAT CATCGTGAAAA GCAGTGTGAGG CGTGTGTGGA | 2800 |
| TGGCTTGTTA CCGGGCGTAC TACACCGCCA GTTGTACACT ATCCACAGCT | 2850 |
| TCTGACACAT GCGACTTGGG AGACGCTGGT CAGTGTGACTC TGAAGCGGAT | 2900 |
| GCCGGAGGCA GACAAGCCCG TCAAGGCGCG TACAGGCGTG TGCCGCGGTT | 2950 |
| TGCGGCGCAG ACGATCAGAA TCGTACGCTG CAGCTCGCTG TTCGGCTGCG | 3000 |
| GCTTTATCAC AAGGAGGATG TTACGCCGCT TCGGTTACGC CACATGGCCC | 3050 |
| GGGGTGAATAT ACGGGAGTTA TCGTAAAGGA AAATACCGC CATCAGGCAG | 3100 |
| TCTTCGGCTT CTCGCTTACC GCACTCGCTG GTCCCTGTGG TGCTGCTGAG | 3150 |
| GCGAGCGTGA TCGACTCCGG CAAAGGCCGT AAATCGCTTT TCCACAGAAT | 3200 |
| CAGGGGATAA CGACAGAAAG AACATGTTAG CAAAGGCCGA CAAAAGGCC | 3250 |
| AGAGAAGCTA AAAAGGGGCGC TCTGCTGGCC TTTTCCATGA GGCTCCGGCC | 3300 |
| CCGTGACGAG CATCACAAAA ATCGACGCTC AAGTCAAGGG TCGGAAAAAC | 3350 |
| CGCACAGAGT ATAAAGATAC CAGGCTTTTC CCGGTTGAAG CTCCCTGCTG | 3400 |
| CGCTCTCTCTT TCCGACCCCT CCGGCTTACC GGATACCTGT CGCCCTTTCT | 3450 |
| CCGTCTGGAAG ATGCCTGGGCT TTTCTCAATG CTACGCTGGT AGGATCTCTA | 3500 |
| GTTCGGTGTTA GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA GGTATCTGTA | 3550 |
| TGAGCTTTAC CGCAGCTGCC TCGGCGGGT CAGCGGCCATG CTGGCCGCTG | 3600 |
| GCCGGTAAGA CACGACTTAT GCACACTGGC AGCGACACTT GTAAACAGGA | 3650 |
| AGAAGAGCTA AAAAGGGGCGC TCTGCTGGCC TTTTCCCTAG CACATGGCCC | 3700 |
| CCGGTAGACA AACGACTTAT GCACACTGGC AGCGACACTT GTAAACAGGA | 3750 |
| TGACGGCTCA CGTACAGGTA AAGGACTGTA TTTGTTATCT GCAGCTCTGCT | 3800 |
| GAAGCAGGGT ATCCCTGGGA AAGAAGTGGG TAGCCTCCTGA TCCGCGCAAC | 3850 |
| AAACCACCGG TGGTAGCGGT GTTTTGTGTT TGGAAACGCA GCAAGTTAGC | 3900 |
| CGCAAAAAA AAGGATCTCA AGAAGATCTT TTGCTTTTGT CTACCGGGTC | 3950 |
| TGACGGCTCA CGTACAGGTA AAGGACTGTA TTTGTTATCT GCAGCTCTGCT | 4000 |
TATCAAAAAG GATCTTCACC TATGATCTTT TAAATTTAAA ATGAAAGTTT 4000
AAATCAAATCT AAAQTATAA TGAAATAACT TGATCTGACA GTTACCAATG 4050
CTTATCAGT GAGGCCACTA TTCTAAGCGT CTGTCTATTT CGTTCCATCA 4100
TAGTTCCTGTG ACTCCCCGTC GTGTAGATAA CCATCTGGCC CCAGTGCTGC 4150
TAGTTGCCTG ACTCCCCGTC GTGTAGATAA CCATCTGGCC CCAGTGCTGC 4200
GCATCTGGCC CCAGTGCTGC AATGATACCG TCCAGATTTA TCAGCAATAA 4250
ACCATCTGGCC CCAGTGCTGC AATGATACCG TCCAGATTTA TCAGCAATAA 4300
GAAGCTAGAG TAAAGTAGTTC GCCAGTTAAT CATTGCTGCA GGCATCGTGG 4350
TGCAAAAAAG CGTGATACGA CTTCTGCTCT CGCATCCTGA TCTGCTACTG 4400
TGCAAAAAAG CGTGATACGA CTTCTGCTCT CGCATCCTGA TCTGCTACTG 4450
CCAGCTGCA GGCATCGTGG TCGAAGTTAA CAGTATTTAT GTTTTGCCCG 4500
CCAGCTGCA GGCATCGTGG TCGAAGTTAA CAGTATTTAT GTTTTGCCCG 4550
CCAGCTGCA GGCATCGTGG TCGAAGTTAA CAGTATTTAT GTTTTGCCCG 4600
TCAGCTGGCG AACTCTTGTT GCCTCCATCC GCCTCCATCC GAAGCTAGAG 4650
TCAGCTGGCG AACTCTTGTT GCCTCCATCC GCCTCCATCC GAAGCTAGAG 4700
TCAGCTGGCG AACTCTTGTT GCCTCCATCC GCCTCCATCC GAAGCTAGAG 4750
TCAGCTGGCG AACTCTTGTT GCCTCCATCC GCCTCCATCC GAAGCTAGAG 4800
TCAGCTGGCG AACTCTTGTT GCCTCCATCC GCCTCCATCC GAAGCTAGAG 4850
GGCAATCGTG GGAACGATCT TGCTCTGCTCT TAAATTTAAA ATGAAAGTTT 4900
TGCAAAAAAG CGTGATACGA CTTCTGCTCT CGCATCCTGA TCTGCTACTG 4950
TGCAAAAAAG CGTGATACGA CTTCTGCTCT CGCATCCTGA TCTGCTACTG 5000
TGCAAAAAAG CGTGATACGA CTTCTGCTCT CGCATCCTGA TCTGCTACTG 5050
TGCAAAAAAG CGTGATACGA CTTCTGCTCT CGCATCCTGA TCTGCTACTG 5100
TCTTCAAGAA 5110

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(iii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-T286

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable

90
(C) UNITS: Not applicable

(i) FEATURE:
(A) NAME/KEY: Inactive T7 φ10 promoter mutant (-4T to A) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: If the mutant T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase or a mutant T7 RNA polymerase, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER: 5,385,834
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110

(i) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
TTCCAGGCA TCAAATAAAA CGAAGAGCTC AGTCGAAAGA CTGGGCCTTT 50
CGTTTTATCT GTTGTGTGTC GGTTGAvGCT CTCCTGAGTA GGACAAATCC 100
GCCGGGAGCG GATTTGAACG TTGCGAAGCA ACGGCCCGGA GGGTGGCGGG 150
CAGGACCCCC GCCATAAACT GCCAGGGAAT TCCCCTAGTA CTGAAATTAA 200
   T7 Promoter Positions       +15
TACCACACAC TATAGGGGAA AGGCTTGAAT AGGACAAATC CGCCGAGCTT 250
   -10   -5     +1   +5
CGACGAGATT TTCAGGAGCT AAGGAAGCTA AAATGGAGAA AAAATCACT 300
GGATATCCCA CCGTGGATAT ATCCCAATCG CATGTAAGG AACCTTTTGA 350
GGCATTTCCAG TACGTTGCTC AATGTACCTA TAACCAGACC GTTCAGCTGG 400
ATATTACGCG CTGTGGAGGAA AACCTGGGCC CGACCCTTAT ATTCGCAAGA 450
CCGGCCTTTA TCCACATTCT TGCCCGGTCT ATGAATGCTC ATCCGGAATT 500
CCGTATGGCA ATGAAAGACG GTGAAGCTGG GATATGGGAT AGTGTTCACC 550
CTTGTTACAC CGTTTCCATG GAGCAGACAG GCGACGCAAAA TACCCTTATG 600
AGTGAAATTAC ACGACGATTT CCGGCACTT ATCAGCTTG CGGCGGTGTC 650
AGGAATCTTG GAGCTTGGAT GATGTAAGG ATGGGCTACC TCCAGAAATG 700
GATGGTTTAGT TACCAGGGA AGCTGGCAAT ATGGGAGCTT ATCCGGAATT 750
CCGTATGGCA ATGAAAGACG GTGAAGCTGG GATATGGGAT AGTGTTCACC 800
CTTGTTACAC CGTTTCCATG GAGCAGACAG GCGACGCAAAA TACCCTTATG 850
AGTGAAATTAC ACGACGATTT CCGGCACTT ATCAGCTTG CGGCGGTGTC 900
GATGGTTTAGT TACCAGGGA AGCTGGCAAT ATGGGAGCTT ATCCGGAATT 950
CCGTATGGCA ATGAAAGACG GTGAAGCTGG GATATGGGAT AGTGTTCACC 1000
CTTGTTACAC CGTTTCCATG GAGCAGACAG GCGACGCAAAA TACCCTTATG 1050
AGTGAAATTAC ACGACGATTT CCGGCACTT ATCAGCTTG CGGCGGTGTC 1100
```
AATGCGCAAA CCAACCCCTTG GCAGAACATA TCCATCGCGT CCGCCATCTC 1150
CAGCAGCCGC AGCCCGCAGCA TCTCGGCTGT TTTGCCGGAT GAGAGAGAT 1200
TTTCAGCCTG ATACAGATTA AATCAGAAGC CAGAAGGCGT CTGATAAAAC 1250
AGAATTTGCG TGCCGGCAGT AGCGCGGTGG TCCCCCTGA CCCATGCCG 1300
AACACGAGAG TGAACGCAGC TAGCGGCAGG GTGAAGTGCG GTCTCCCCCA 1350
TGGAAAGACTG GGAGAACGAA CAGCGGCGCA TCTCGGCTGT TTTGGCGGAT GAGAGAAGAT 1400
AAAGACTGCG CTTCTGCTTT TATCTGTTGT TTTGCGTGGA AGCTCTCCCT 1450
GAGTAGGAAC AATCCGCGAG GAGCGGATTG GAACCTGGCG AAGCAACGGC 1500
CGGAGGAGTT GCGGCGAGGA CGCCCGCAT A AACCTGACAG GCATCAAAAT 1550
AACACGAGAG CCAATGCGAC GGAATGCTTG TGGCCGATGA GAGAGATTT 1600
TCTGTCTGTC ATATCTACAA GCCATCCCCC CACAGATACG GTAAAATAGC 1650
CTCTTTTGAAT CAGATGAAGA AGCGATGCCGA AGCGATCCGA ATCAAAATTA 1700
AACGCAAAGC TATCGTAGG GCGGTGGGCG TTCCACCTGA CCCCATGCCG 1800
CTCAGAAAGT G AAACCGGTA GCGCGGATGG TAAGTGGGCG TCTCCCCATG 1850
CGAGAGTAGG AACAGGCGGA GCCAAGGGAT GGCTCTCCCT GTGACTCCCTG 1900
AGACTTGGCC TTTGTTTTTA TCTTTGTTTT GTTTGTTTAC GCTTCGCTGA 1950
GTAAGAAAGA TCCGCCCGGA GCGAGTTTGA ACGTGCGGAA GCAAGGGGCC 2000
GGAGGCTTGGC GCCGAAGACG CCGCCCATAA ATCTGCAGGC ATCACAATTA 2050
GCAAAAGGCC ATCTCGACGG ATGGCTTCCCA GAGCTTGGCA CCAATGCCGA 2100
CTGTCGCTAT ATCTACAGAC CAGCTCCCCA CAGATACGTT AAACCTAGCCT 2150
GTTTCTTACA TCAGAAGAAG AGTCGGCGGA CTTGGTTCGC TTGGCAAGGG 2200
TGCCATAGT GGTTCTCCTG TTCTGGAGGA CCCGGCTAGG CGTGATCGGG 2250
TGCTTATCGT GTAGCAGAAG TGAAATCAAGC GATACGCGCG GAACTGCGTG 2300
CGACTGCTGC TCAGAGAACG TGCTCAGCTG AGCAACAAAC TGAATGCTCT 2350
TGCTTTTCCG TGTTTTGCAA GCTGTGGCAA CCCGGAATGC AGCGCCCTGC 2400
ACCATTAGTG TGGGATCTGC CATCGCGAGA GCTTGCTTCG TACCTGGCTG 2450
AACACCTACA TCTGTATTAAG CAAGGGCGTG GCAATGACCC TGAGTCTTAT 2500
TTCTCTGCTC CGGCGCAGTC CATAAAAGCAA GTTTGTTTAC CTCACAAGCT 2550
TCCATGAAGA GGGCGATGTC ATCATGAAGA ACCGATACGA TGAAGATCCT 2600
CTCTCGTTTC ATCAGGTATCA TTACCCCAAT GAAAGAAAT TCCCCCTTAC 2650
AACGAGGCTA CAAAGGACGC AAACAGCTCT AACAGGGGCA AAACAGCTCT 2700
GCTTTATCAG AAGCCAGACA TTAACGCTTT CGAGAAACCT CAAGCGGCTG 2750
GAAGCGATAG CAAAGCAGCA CATCGTGGAA CGTCTTGCGA ACCACGCTGA 2800
TGAGCTTTAC CGAGCTGCGC TCAGCGGCTT CGTGTAGGAC GGTGAAAGCC 2850
TCTGACACAT CGCACTCGAGC GAGGAGGTCA CAGCTTGTTC GTAAAGCGGAT 2900
GCCGCGAGCA GAAAGCGGCG TCAAGGCGCG TGAGCGGTGG TTGGCGGCTG 2950
TGGGGGCGCA GCGATCAGAC AGTCAGCGTA CAGATGCGGA GTGTATACGT 3000
GCTTAACTAT CGCGCATCAG AGCAGATTGT ACTGAAGGTC CACCATTGCG 3050
GTGGTGAAGAT ACCCGACAGA TGGTAAAGGA GAAATACCG CATTAGGGGC 3100
TCTTCGGCTT CTCCTGCTAC TGAATCGCTG CGTCTGGGCG TTTGGCGCTG 3150
-continued

GCGAGCGGTA TCAGCTCACT CAAGAGCGGTT ATATACGGTTA TCCACAGAAT 3200
CAAGGGGATAA CGCAGGAAGG AACATATGAG CAAAAGGCGCA GCAAAAGGCC 3250
AGGAAGCCGTA AAAAGGGAGG TTGTTCTGGGCT TTTTCTCATATAAACGGATCC 3300
CCTGACAGC ATCACAACAA ATCAGCAGCTC AATCAGAG CTGCCAGAAACC 3350
CGACAGGACT ATAAAGATAC CAGGGTTCCT CCCCTGGAAG TCCCTCTGTG 3400
CGCTCTCCGG TTTCCAGCCT GCGGCTTACCC GAGATACCTGT TGCCGTTTCT 3450
CCTTCGGAAG ACGCTGAGCTG TCTCTCAATG TCTACGCTGT AAGGTACTCTA 3500
GGTGCGTGTA GGTGCTCCTG TCCAGCTGGG GGTGCTGTCAG CGAAAGCCC 3550
GGTTCGCTGTA GGGAGGCGC TTTTGCTGGC TTTTTCCATA GGCTCCGCCC 3600
CCCTGACGAG CATCACAAAA ATCAGCGCTC AAGTCAGAGG TGGCGAAACC 3650
CGACAGGACT ATAAAGATAC CAGGGTTCCT CCCCTGGAAG TCCCTCTGTG 3700
CCTAACTACG GTTACACTAG AAGGAGCTTA TTTTGGTCAT CTGCCCTGCT 3750
GAAGCCAGTT ACCTCCGGAA AAGAGTTCGG TACGTCTTGA TCCGGCAAAC 3800
AAAACCAGGC TGGTACGGGT GCTTTTCTTTG TTTTCAAGCA CATGATACAG 3850
CGCAGAAAAA AAGAGATCCCT TTGATCTTCT CATCGGGGTC 3900
TGATGCTAGG TGAAAGCTGAT ACTACGCTTA AGGGATTGGT GTCATGAGAT 3950
TATCAAAAAG GATCTGCTCC TAGATCTTCT TAAATTTAAA ATGAAAGTTTT 4000
AAATCAACTCT AAAGATATATA TGGTAAAAACT TGGTCCGACA TTACACATAG 4050
CTTAAATGATG GAGGAGCCTA TCTTCAGGAT CTGTCTATTT TCTTTGACGC 4100
TAATTGCGCTG ACTCCTCCCTG TGCCTACAGA ATACGGCTAG 4150
CCATCTGCCC CCAGTGCTGCA ATATACGGCT CGAGACCACCC CTCACCGGCCG 4200
TCAAGAATTCT TCAAGAATATA ACCAGGCAAAG GAGAGCCAGA 4250
GTGGCTCTGC AAATATATCCC CCTTCEATAAG ATGTATTAAAT TTTGTCGCCCC 4300
GAAGCTTAGAG TAAATGATTAA GCCAGGATATA ATTTGGGCTG ACCAGGATTTG 4350
CATTACGTCA GGAGCTCGTTG GTGGTCAGCTC GTCGTTGGGT ATGGCTTCAT 4400
TACGTCGGTG TTTCCACGGA TCAAGGGGAG TTTATCAGTAT CTCATGACTCA 4450
TGCAAAAAAG CGGGTACCTCA TTCCGGCGCT CAGCTGCAGT TTTGGTATCT 4500
GTTGGCCGCA GTTTTGATCA CATCGTCATG ACCAGGACTAG GATTTTCTTC 4550
TTACTGTCC ATCAGTGAGT TGATGCTGTT TGTTGACCTA CTGAAATGTA 4600
ACCAAGTCAT TCTTACGAGA GTCGGTACGGG CAAGAGGCTA CAGCACAACC 4650
GGCGTCAACA CCGGATAATA CGCAGCCACA TAGCAAAAATTT TAAAGTGTA 4700
TCTATTTGGT AAAAGGTCGT CCCGGGAGAA AAGACGCTGG CACATATATTG 4750
CTTGATGATG CAGTCCGTC GTTACCCACT CCTGACCACT ACTGATCTCC 4800
AGCATTTGTT TCTTTCTACCA GCCGGTCTTG GTGACACAAAA AGACAGAAGGC 4850
AAAATGACC CGAAAAAGGA ATAAGGCGCA CACGGAAAAAT TGGATTACCTC 4900
ATACTTTCCT TTTTATTTATA TATTTGCAAGG ATTTTTATGCT 4950
CATGAGCGAG TACATATTTT AATATTTTAA GAAAAATTTAA AAAAATAGGG 5000
TCCCGGCCAT AATTTCACCC AAAAAAGGCG GTCAGCCTCAG AAGAACCATT 5050
ATTATCATGA CATTAACCTA TAAAAATAGG COTATCACGA GGCCTTCG 5100
TCTTCAAGAA 5110
(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-B64

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Intermediate T7 φ0 promoter mutant (φS to C) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: The mutant T7 promoter between positions 198 and 220 is weakly recognized by T7 RNA polymerase; consequently, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Wardleman, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993

(ix) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTCCCAAGGCA TCAAATAAAA CGAAAGGCTC AGTCGAAAGA CTGGGCCTTT 50
CGTTTTATCT GTTGTTTGTC GGTGAACGCT CTCCTGAGTA GGACAAATCC 100
GCCGGGAGCG GATTTGAACG TTGCGAAGCA ACGGCCCGGA GGGTGGCGGG 150
CAGGACGCC GCCATAAACT GCCAGGGAAT TCCCCTGAAT TCGAAATTAA 200
T7 Promoter Positions

TAGGACCAC TATAGGGAGA AAGCTTGAGT AGGACAATCG CGCCGAGCTT 250

-10 -5 +1 +5
CGACGAGATT TTCAGGAGCT AAGGAAGCTA AAAATTCCTA CAAAATCTACT 300
GATATACCA CGTGTGATAT ATCCCCAATCG CATCOTAAG AACAATTTGA 350
GCGATTTTCG TCAGTTGCTC AATGTACCA TAAACCAGACC GCTCAAGTCG 400
ATATATGGCT GCTTTTTAAG ACCGTAAGAA AAAAAATCAG CAAATTTTAT 450
CCGCTTCTTA TTTCACTTCT TTCGCCGCTG ATGAAATGCTC ATCGGAATT 500
CGATTGGGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTCCAC 550
CTTTTACACC GGTTTTCATG GAGAAACTG AAAAGTTCCT ATCGCTCTGG 600
AGTGAATACC ACGACGAGTTT GCGGCCGTTT ATCCACATT ATTCGCGAAG 650
TGTTATGATC TACGCGAAAG ACGTGGCCTA TGGCCCTAAA GGCTTTATTG 700
AGGAATGTTT TTTCGCTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 750
GATTATAGCG TGGCCATATT GCACAACTTC TTGCCCCCGT TTTTACCAT 800
GCGAAATAT TATACGCAG GCGACACGTT GCTGATGCCTC CCGGCGATT 850
AGTTTTATCA TCAGCTCTGT ATGGGCTTCC ATGTCGGCAG AAATGTCAAT 900
GAATTACAAC AGTACTGCGA TGAGGCGCCTC ATGTCGGTAC CACGAGATTT 950
GGTCGGTGTT TACGGTGAAA ACCTGGCGCTA TCTCCCTAAA GGGTTTATTG 1000
GAAATATGTT TTTCGCTCTCA GCCAATCCCT GGGTGAGTTT CACCAGTTTT 1050
GATCCACAC TCCAAGAATT GGAGCCAATC AATTCTCGG GAGAACTGTG 1100
AATCGCAGAA CCAATCTCTCG GCAGAACATA TCCATCGCTG CCGCCATCTC 1150
CAGCGGCAGC AGCGAGCGCA TCTCGGCTGT TTTGGCGGAT GAGAGAAGAT 1200
TTTCAGCCCT ATACAGATTAA TACGACGCTG CAGAGAAGGT GGGCTTCGTT 1250
AGGACTGGCT GCAGGCGAGT CCGCGCCTG GCGCCATGCC 1300
AACGCAAGAG TAAAGAAGGT GCAGGCGAGT CCGCGCCTG GCGCCATGCC 1350
TGCGAGAGTA GGGACTGCCC AGGCACTAAA AAAAAACGAAA GCCTCGATTG 1400
AAAGACTGCG TTTTCTGTTT TATCTGTTGT TTGTGCTGGTA ACGCTCCCTT 1450
GAGTGAAGCA AATCCGGCGC GCAGGGAATT GACCGCGGAG CACGCGGAAG 1500
CCGAGGGTGC GCAGCCGCGC GCAGCGGCTC AAGCTGACCA GCATCAAAT 1550
AAGCAGAAAG CCATCTCGAC GGAATGCGCTT TTTGGCTTTT TGGCCCTAA 1600
TCTGTGCTCT ATATCTCAGA GCCATCCCT CAAGATCGTA GTAAACTCCG 1650
CTCGTTTGGT CATCGAGAAA GCAGCTGTTT TGGCCGATGA AGAGAGATT 1700
TCACGCTGAT ACAGATTTAA TCAGAAGCAG GAAGCGGCTC TTGATAAAAC 1750
AATTCTGCTG GCAGCCGCTG CCGTGTTGTC CCACCTGACC CCACTGCGAA 1800
CTCAGAAGTG AAAAACTCGTA GCGCCGATGG TCGTGGGGTG TCTCCCCATA 1850
CGAGAAGTGG GAAGCTGAGGG CAGAATCGTA AGAAGGCGCC 1900
AGACTGGGCC TTTCTGGTTT TCTGTGGTTT GTCTGGTGGAA GCTCCCCATA 1950
GTAGGACAAAG TCCGCGCGGA GCAGGATTTGA AGGTTGCGAA GCAACGGCCC 2000
GGAGGGTGTC GGGCGAGAGG GCCGCCATCA ACGTCGCGAC ATCAAATATTAA 2050
GACAGAACG ATCCGACTGG ATGGCCCTTGT TGGTCTCCTG AAAAAATGTT 2100
CAGGCAGATG ACTATGAGCA CATCCCGCTA GAGATACAGG ATATGAGTCT 2150
CTGTTTTACA TGAGAAAGGC AGTCGCGGAC GCTGGTCCTT TGCGACGCGG 2200
TGCGCTGATACT GTGTCTCCTG GTGCTGGAAG CCGGCTAGG CTGCGGCGGT 2250
TGCGTTCATT GTTACGAGAA TQAATCACGG ATACGCCGAG GAACTCGAAG 2300
CGACTGCTGC TGCAGAAACGT CTGCAGCCTG AGCAACAAAC GTAATGGTCT 2350
TCGGTTTCCG TGTTTGTGTA AAGCTCGAAAA CGCGGGAGTC AGCGCCCTGC 2400
ACCAATTATGT TCAGGGATCTG CATGCGAGCA TGCTGCTGGC TACCCTGTGG 2450
AACACCTACA TCTGGATTAA CGAAGGCGCT GAATGGACCC TGAGGTAGTT 2500
TCTCTGGATG CCGCGGATCTG CATGCGAGCA TGCTGCTGGC TACCCTGTGG 2600
CTCTCGTTTC ATCCGATTA GTCCTTGACT CTCACAACGT 2550
ACGAGGCAAT CAAATGCAAA AAGACGGCTT CAACTAGGCA 2700
CTTTATCAG AGGAGCAATA TATAAGCTTC TGGAAAAAAT CGACAGCTTG 2750
GAAGAGGTAG AAGACGGCTT CAAATGCAAA AAGACGGCTT CAACTAGGCA 2800
CTTGAGCTTAC CGAGGCTGCC TCGCGGATT TGGTAGTAC GGTGGAAACC 2850
TCTGAGCATG CGACTGTGCG TCGAGGCTTC GAGAGCTTGG GGTGGAAACC 2900
GCCGAGGACA GACAGGCGCT TACAGGCTTC TACAGGCTTC TACAGGCTTC 2950
TCGGCGCGCA GCGCATGTCA CACGACGTCG CGTCCTGCTG TGGGCTGCTG 3000
GCTTTACTAT AGCGGACTAG AGGAGCATTG ACTCAAGCTG CACATATGCA 3050
GGTGTTAAAT ACCGACAGAA CATCCGATAC TCCGCTCTAG CACGACGTCG 3100
TCTTCCGATT CTTCCGATT CCGGTATTCC GGTATACGT CACGACGTCG 3150
GCAGAGGTAG TACGAGAACT CAAATGCAAA AAGACGGCTT CAACTAGGCA 3200
AGGAACGGTA AAAAGGCCGC TTTTCTCCAT CAGGGCATGC CCAAGCTTTA 3250
CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC 3300
TGAGTTGGTA GAAGTGGTTG TGGGCTGCTG TGGGCTGCTG TGGGCTGCTG 3350
CGCTACTGGG AAGCCGCTGG TTTCATCAGT TCCGCTCTAG CACGACGTCG 3400
CCCTCGGATA AAGCCGCTGG TTTCATCAGT TCCGCTCTAG CACGACGTCG 3450
GGCTTCCGGA AGCGCCGGCG TTTCTCAATG CTCAGGCTTG AGGTATCTCA 3500
GGTCCGTTCA GTCGTCGCT GCAGGGCATG GCCAGCTTTA TGGGCTGCTG 3550
GCTTCAGCGC ACCTGGCGTC TCTATTCGCT AACTATGCTG TGGGCTGCTG 3600
CCGTCGTTGA GTTAGCTGAT TCCAGGGCTG GTCTGCTGCA CGAAGCCCGC 3650
TAGCACCGCT ACCGCTAGCT TTATCCGGG TACGCTGCTG TGGGCTGCTG 3700
CCTAACATCG GCTACAGTAT AAGGAGCTTA TTTGTATCT CCGCTCTGCT 3750
GAAGGACAGTT ACCCTCGGAA AAAAGCTGTTG TAGCTCCTGA TCCGCCAACC 3800
AAACCACCGC TGGAAGCGGT GGTGTTTTTG TTTGCAAGCA GACATATCAG 3850
CGAGAAAAA AGAGAACTTC AAGAAGATCTT TGGATCTTTT CCTAGGGGCT 3900
TGAGCTCAGT TGGGAAGAAA ACTCAAGGTT AAAGGTATTTT TGGTAAGGAT 3950
TATCAAAAAA GATCTTCCAC TAGATCCTTT TAAATTTAAA ATGAAGTTTT 4000
AAATCATCT AAAATATATA TGAGTTAAAT TGCTGACAGA GTACCAATAG 4050
CTTAATCGTG GAGGCACCTA TCTCAGGCTG CGTCTATTT TGTTGCTCCA 4100
TAGTGTGCTG ACTCCGCGTC GTGATGATAA CTACGATAGC GGAGGGCTTA 4150
CCATGTGCAG AACATGGCGC AATGATACCG CGAGCCACC CGTACCGGGC 4200
TGCCATATTC TCAAGCAATA ACCAGCCAGC CGGAAGGGCC GAGGCAAGAA 4250
GTGGTCCTGC AACTTACCAT GCCTTCCATCC AGTCAATTAA TGTTGCGCAG 4300
GAAGCTAGAG TAAATCTATCGGCCAATT AAGTTTCGCCA ACCTGTTGTC 4350
INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-Pl208

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Inactive T7 d10 promoter mutant (-7C to G) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyltransferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: If the mutant T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase or a mutant T7 RNA polymerase, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:

TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters

JOURNAL: Biochemistry

VOLUME: 32

ISSUE: 35

PAGES: 9115-9124

DATE: Sept. 7, 1993

SEQUENCE DESCRIPTION:

SEQ ID NO: 1 to 5110

T7 Promoter Positions

RELEVANT RESIDUES IN SEQ ID NO: 1 to 5110
CCGGAGGGTG GCGGCCAGGA GCGCCCACCAT AACTGCCCGG GCATCAAATT 1550
AAAGCAAGAG CCTCTGAGCA GGATGCGCTT TTTCGCTT TCACAAACCT 1600
TCCTGTCTGC ATATCTCAAA GCCATCTCCCC CACAGATACG GTAAAATACG 1650
CTGGTTTTGG CATCAGAAAA GCAGGATTGG TGCGGATGGA GGAAGATT 1700
TCAGCCCTAG ATACGATTAA TCGAGACGCA GAAGCGGTCT GATAAAAACAG 1750
AATTTCGCTG GCCGCGATAG CCGGCGTGTCC CACCTGACC CCATGCCGAA 1800
CTCAAGGGGT AAAAGCGGTA GCGCCGATGG TAGTGGGTTG TCTCCCATG 1850
CGAGAGTAAG GAACTGCAAG GCATCAAATA AAAAGGGGTA AGTAGCTGAA 1900
AGACTGCGCC TTTGCTTTTTA TCTGTGTATT GTCGTOAGAC GCTCTCTGGA 1950
GTAGGCACAA GCCGCGGGGA GCGGATGGGA GCAACGGGCC 2000
GAGAGGGGTC GCCGCGAAGGC CCCGCCATTA ACTGGCACGG ATCAAATTAA 2050
GCAGAAAGGC ATACCTGCGGG AGTGCTCTTTTA TGCGTTTCTA AAAACTCTTC 2100
CTGTCGTCAT ATCTACAAGC CATCCCCCA CAGATACGGT AAACTAGCCT 2150
GTTTTTGCG TCGAGAAAGGC AGTCGCCGAG CAGTTGGTCC TGGGCAAGGG 2200
TGCGCATGAT CTGGCTCTTG TCGTGTAGGA CCGGGCTAGG CTTCCCGGGG 2250
TGCTTACTCT GTTAAATTTCA AATACCGGAA GAACTGGAAG 2300
CGACTGCGCC TGCAAAAGGT CGTGAGCCTG AGCAACAACA TGAATGGTCT 2350
TCGTTTCTGA TGTTTGTGAA AGTCTGGAAA GCAGCAGTGG GCAGGCTGAA 2400
ACCATTATGT CCTCGATCCTG CATCAGCAGGA TGCTCTGCTG TACCCCGGGT 2450
AACACCTACA TCTGTATTAA CGAAGCGCTG GCATTGACCC TGAGTGATTT 2500
TTCTCTGCTCC CGCCGCGTAC TATACCCGCA GTTTTTATCC TCACCAGGCT 2550
TCCAOTAAGC GGGGCTGGTT CATATCCAGA ACCGCTATCG AGAACACGCT 2600
CTCTGCTTTC ATCGCTATCA TTACCCCCAT GAACAGAAAAT TCCCCCCTAC 2650
ACGGAGGCAAT AAGTGGAACCA AACAGGGA TAAATCCCCC ACCGAGCCC 2700
GCTTTATCA GAGCCAGACA TAAACGCTTC TGGAGAAACT CAACCAGCT 2750
GACGCGTGATG AACAGCCGA AATCTGCGGG TCGCCCGAGA CCCAGGCTG 2800
TGAGCTTTAC CGCACTGGCC TCGCGCGTTT CGGTGATGAC GGTGAAACC 2850
TCTGACACAT GCAGCGTCGC ACCGCGCCCT ACAAGGGATG GAATGATCTG 2900
GCCGGAGCA GACCAACGCC GGCTGGCCCT CGAGCGGCTG TGGGCGGGT 2950
TCGGGGGGCA GCGCTGACAGG AGTCAGCTAG CTGACTGAGGT TCTGCTGTG 3000
GCTTTACTAT GCGGTCTATCG ATACGCTAGCC ATACGCTAAG GCACTGGTCT 3050
GCTCCGTTTC GAGCCTCTGAG CAGGATATTG GATGCAATGG CACCATATGC 3100
CTCTGCGTCT TCTCAGCTAC GACATCGCTG GCTCCCGCTG TGGCCTGCG 3150
GCGAGGGTGA TCAGCTCCAG CAAGAGCCGA CAAATGCGAT CACCAGTGCC 3200
CAGAAGAGCA AAAAGGCCGG TGGCGAGGGA TCTCCGCTCA GCCGCGGCC 3250
CCCTGACAGA CATCAGAAAA ATCAGAGTGG AACAGGCGCA GCAAAAGGCC 3300
AGGAAGGCTA AAAAGGCCGC GTTCTGGGAC TTTTCTTCCAG GCAGCGGCC 3350
CCCTGACGAG CATCACAAAA ATCAGACGTC AAGTGCAAGG TGCGGAAACC 3400
CGAGGACTA ATAAAGATAC CAGGCGTTTTCC CACCAGGAGA GTCCTGCTG 3450
GCGTCTCCTG TTTGCGACCT GCCGCTTACC GAGATACCGT CGGCCCTTCT 3500
CCCTGCGGGA AGCGTGCGGC TTTCATCAAG TCTACGTGACT AGATATCTCA 3550
GTTGCGGTGT AOTCGTCTGC TCTAAGCTGG GCTGCTGCTG CAAACCGCCC 3600

-continued
(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-P1031

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Inactive T7 φ10 promoter mutant (-7C to A) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pKK232-8 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: If the mutant T7 promoter between positions 198 and 220 is recognizable by T7 RNA polymerase or a mutant T7 RNA polymerase, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warsamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 510

(ix) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
TTCCCCAGGCG TCAAAATAAA CGAAGAGGCTC AGTCAAAAGA CTGGGCCTTT 50
CGTTTTATCT GTTTTTTGTC GGTGAACGCT CTCTAGAGTA GGACAAATCC 100
GCCGGGAGGG AGTTTGAACG TTGCGAAGCA ACGGCCCGGA GGGTGGCGGG 150
CAGGACGCCC GCCTAAACT GCCAGGGAAT TCCCCTAGTA CTGAAATTAA 200
```

```
T7 Promoter Positions - 15

TACGACTAAG TATAGGGAGA AAGCTTGAGT AGGACAAATC CGCCGAGCTT 250
* * +1 +5
```

```
CGACGAGATT TTTTCCAGCT AAGGAAGCTA AAATGGAGAA AAAAAATCCT 300
GGTATACCA CCGTTCTAT ATCCCAATCG CATGCTAAA ACGATTGTA 350
GGCATTTTAC TCGGTGCTC AATGTACCTA TAACCAGACC TGGTGTCGAG 400
ATATTACGCG CTTTTAAAAG ACCGTTAAA AAAAAAAGCA CAGTTTTTAT 450
GGCCTTAATTT TCTCCATCGT GCCTGACC CTGAGTACCTA AATGGGAGAT 500
CCGTATGCGG ATGAAAGGTA AAGTGACCTG TATGCTGCGGT AATGGGAGTG 550
CTTTTACTAC CGTTTTTCAT GCACAAACTG AAACGTATTT ATCCGCAGAG 600
AGTGAATACC ACGACGATTT CCCGCAAGTT CTACACATAT ATCCGCAAAG 650
```
ACGGAGGTG TACGOTGAAAA ACCTGGCCTA TTCCCCCTAAA GGTTTATTG 700
AGAATATGTT TTTGCTCCTA GCCAATCCCC GGTTGAGTTT CACCAAGTTT 750
GATTTAAACG TGGCCAATAT GGACAATCCT TTTCCCCCGG TTTTCACTAT 800
GGGACAATAT TATACGCAAG GCGCAAGGTT GCTGATGGCG CTGGCCATTC 850
AGGTTCATCA TGGCGTCTGT GATGCGTTCG ATGTCGGCAG AATGCTTAAT 900
GAATTACAAAT ATGACTGCTG TGAATGGCAAG GCGGAGGGGT AATTTTTTAA 950
TAAATAAGCGG ATGAATGGCA GAAATTCGCT GAGGCGGCAC TCTGACACG 1050
GATTCACCAC TCAAGAATT GGAGCCAAATG AATTCTTGGG GAGAATCTTG 1100
AATGCCGCAA CCAAACCTCAG GCAGAAACATA TCCATCGCOT CGGCCATCTC 1150
CAGCAGCCGC AGCCGAGGCA CTCGCGTCTG TTTGGCGGAT GAGAGAAGAT 1200
TTTCAGGCTG TATACGCAAG GCGACAAGGT GCTGATGCCG CTGGCGATTC 1250
AGGTTCATCA TGCCGTCTGT GATGGCTTCC ATGTCGGCAG AATGCTTAAT 1300
GAGCAGGGGA TGGCGGAGGT TGGCCAGTTT GAGGAGGGGT AATTTTTTAA 1350
GATTCACCAC TCAAGAATT GGAGCCAAATG AATTCTTGGG GAGAATCTTG 1400
AAAGACTGCGG CTTCTTCGTT TATCTGTGTT TGTGCGTGTGA ACGCTCTCCT 1450
GAGTAGGACA AATCCCAGGG GAGGAGGGGT AATTTTTTAA 1500
CCGGAGGGTG GCAGGGCCAG GCAGGAGGTG TGGCCAGTTT GAGGAGGGGT 1550
AACAGAGAGG CCAATCTGAC GGTGCCCCTA TTTGGGTTTT TACCAACTCT 1600
TCTGTTTCTA ATATCTACAA GCCATCCCCC CACAGATCGC GTCATAACAC 1650
CTCCAGGCTG TATATCTAAA GCCATCCCCC GCAAGAGACG GAGCCGCTGA 1700
AAATGCGCAGA ATGAATGGGA GAAATTCGCT GAGGCGGCAC TCTGACACG 1750
CTCAAGACAG AATAGGATTT TCAAGACAGA GCAGGAGGTG TGGCCAGTTT 1800
CGAGAGAGG GCAGGAGGTG TGGCCAGTTT GAGGAGGGGT AATTTTTTAA 1850
GTAGGACAAAA GCCGAGGCTG ACGTGGCAGG GCTGTTGGTT TACCAACTCT 1900
AAATGCGCAGA ATGAATGGGA GAAATTCGCT GAGGCGGCAC TCTGACACG 1950
CTGCTGTGTC ATATCTACAA GCCATCCCCC GCAAGAAGACG GAGCCGCTGA 2000
TCCGCTGCG TATATCTAAA GCCATCCCCC GCAAGAGACG GAGCCGCTGA 2050
ACCCGAGCGG GCCGAGGCTG ACGTGGCAGG GCTGTTGGTT TACCAACTCT 2100
CTTGCGTCAT ATATCTACAA GCCATCCCCC GCAAGAAGACG GAGCCGCTGA 2150
CTCCTGCGAG GCAGGAGGTG TGGCCAGTTT GAGGAGGGGT AATTTTTTAA 2200
GAGACTGCGG TTTGCTGCTA TCTGTGCTGT GTCAGCAGGG TCGCTCTGCG 2250
GCTGGGAGGC ATATCTACAA GCCATCCCCC GCAAGAAGACG GAGCCGCTGA 2300
GCTGGGAGGC ATATCTACAA GCCATCCCCC GCAAGAAGACG GAGCCGCTGA 2350
CTCCTGCGAG GCAGGAGGTG TGGCCAGTTT GAGGAGGGGT AATTTTTTAA 2400
ACCCGAGCGG GCCGAGGCTG ACGTGGCAGG GCTGTTGGTT TACCAACTCT 2450
AACACCTGAC AATATCGGCA GACGAGGCTG ACGTGGCAGG GCTGTTGGTT 2500
TTCTGCTGCTG TATATCTAAA GCCATCCCCC GCAAGAAGACG GAGCCGCTGA 2550
TACGATGCGC ATATCTACAA GCCATCCCCC GCAAGAAGACG GAGCCGCTGA 2600
ACCCGAGCGG GCCGAGGCTG ACGTGGCAGG GCTGTTGGTT TACCAACTCT 2650
ACCCGAGCGG GCCGAGGCTG ACGTGGCAGG GCTGTTGGTT TACCAACTCT 2700
GCTTTATCAG AAGCCAGACA TTAACGCTTC TGAAAGAAC TCAACGAGCTG 2750
GACCGGAAGT GAAACGGAGA CATCTGTGAA TCGCTTCAG ACCACGCTGA 2800
TGAGCTTATC CGAAGCTGCC TCGCGCCTTT CGGTGATGAC GGTGAATACC 2850
TCTGACACAT GCACTCCCG GAGACGTCGA CAGCTTTGCT GTAACGGGAT 2900
GCCGGGAGCA GACAAATGCC TCAAGGCGAC TCAACGAGCTG 2950
TGCTGAGCTAACAGGCA GAGCATATCT ACTGAGAGTG CACCATATGC 3000
GCCGGAAGTA TCAAGCTCAG TCGTGATGAC GGTGAAAACC 3050
TGTTGAAAT TACGCACAGA TCCTGAAGGA GAAATACCGT CACAGGGC 3100
CTTCCGCTT CACTCAGCTAG TCGACTGCTG TCGCGGTCG 3150
GGAGGGTGTA TCAAGCTCAG TAAAGGCGGT TAAAGGCGGT 3200
CAAGGGATAA CCCAGGAGAA AATCTGAGA CAAAAGGCA GCAAAAGGCC 3250
AGGCAAGGTA AAAACGCGGC GTGCTGAGTC TTCTTTTCCA GCTGCGGCCC 3300
CCCCGTAAGG CAGCACCGAT ATGAAGGACTT TAAAGGCGGT 3350
TGCGAGATCC ATAAAGATAC CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG 3400
GCCGGGAGCA GACAAGGCCG TCGGCGCGCG TCGGCGGGTG 3450
TCGGGGCGCA GCCATGACCC AGTCACGTAG CGATAGCGGA GTGTATACTG 3500
GCTTAACTAT ACGCCATCA TGCAAGCTCGTG CTGCGGGGTTG 3550
GACGCGGCCG AGGTGAGGTG ACCGAGGTTA TCCACAGAAT 3600
GGATGTGTAA ACCGCGGTTC TAAAGGCGGT GAAATACCGT 3650
TGTTGCGTGA GTGTTGCGTG TCCAAAGCTTG GGTAGTGGCA GCAAAAGGCC 3700
CTTACCGCCG AGCGGCCTAC TTATCCCGGT AAAATACCGT TCGAAGGAA 3750
CTTACCGCCG AGCCGCTGCTC GCTGCTCTCTT GCAGTGGGCTA 3800
AAACCACCGC TGGTAGCAGT GGTCTTTGGT TTTCAAGAGC GCAGTGGGCTA 3850
CGAGGAAAAA AAGAGTCTCA AGAAGATCCT TATGTCTTTT CTCAGGCGGT 3900
TGAGGGTGTA GTGTTGCGTG TCCAAAGCTTG GGTAGTGGCA GCAAAAGGCC 3950
TTAATACGTTA TCTGAGATAGG GATTTGGTGT GAAATACCGT 4000
CTTAAATGAT GAGGTAGTGA GCCGCGTCA CAGAGTTCTT GAAGTGGTGG 4050
TTCTCCGCTT CACTCAGCTAG TCGACTGCTG TCGCGGTCG 4100
GGAGGGTGTA TCAAGCTCAG TAAAGGCGGT TAAAGGCGGT 4150
CCCTTGGGGA AGCGCGGTTC TTTCTCAATG TCTCATCAGT AGGTATCTCA 4200
CCCTTGGGGA AGCAGTTGAC TCCGCGGTTC TTTCTCAATG TCTCATCAGT 4250
CTTACCGCCG AGCGGCCTAC TTATCCCGGT AAAATACCGT TCGAAGGAA 4300
GACGCGGCCG AGGTGAGGTG ACCGAGGTTA TCCACAGAAT 4350
GCCGGAAGTA AAGAACCCCC AAAATACCGT TCGAAGGAA 4400
GCCGGAAGTA AAGAACCCCC AAAATACCGT TCGAAGGAA 4450
CTTACCGCCG AGCGGCCTAC TTATCCCGGT AAAATACCGT TCGAAGGAA 4500
GAGGTAGTGA GCCGCGTCA CAGAGTTCTT GAAGTGGTGG 4550
CTTACCGCCG AGCGGCCTAC TTATCCCGGT AAAATACCGT TCGAAGGAA 4600
GCAGGAAAAA AAGAGTCTCA AGAAGATCCT TATGTCTTTT CTCAGGCGGT 4650
GGAGGGTGTA GTGTTGCGTG TCCAAAGCTTG GGTAGTGGCA GCAAAAGGCC 4700
CTTACCGCCG AGCGGCCTAC TTATCCCGGT AAAATACCGT TCGAAGGAA 4750
(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5110 base pairs
(B) TYPE: nucleic acid
(C) STRANDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: pCM-T221

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Intermediate T7 φ10 promoter mutant (-6A to G) located between nucleotides 198 and 220 and in front of a promoterless chloramphenicol acetyl transferase (CAT) gene.
(B) LOCATION: 198 to 220
(C) IDENTIFICATION METHOD: The clone was constructed from pJK222-5 and the inserted "promoter" was sequenced.
(D) OTHER INFORMATION: The mutant T7 promoter between positions 198 and 220 is weakly recognized by T7 RNA polymerase; consequently, CAT can be expressed from the plasmid.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Chang, L.L., and Warshamana, G.S.
(B) TITLE: Selection and Characterization of a Mutant T7 RNA Polymerase that Recognizes an Expanded Range of T7-like Promoters
(C) JOURNAL: Biochemistry
(D) VOLUME: 32
(E) ISSUE: 35
(F) PAGES: 9115-9124
(G) DATE: Sept. 7, 1993
(H) DOCUMENT NUMBER: 1 to 5110
(ix) SEQUENCE DESCRIPTION: SEQ ID NO:IS:

TTCCCAAGCCA TCAAAATAAA CGAAAGGCTC AGTCGAAGAG CTGGGCTTTT 50
CGTTTTATCT GTGTGGGTC GTGGAACGCT CTCTGAGTAT GGAACAATCC 100
GCCGGGAGCG GATTTGAAAG TGGCAGAAAG ACGGCCCCGA GGCGGGGGG 150
CAGGAGGGC GCCATAAAAT GCGAGGGAAT TCCCCCATGA CTGAAATTTA 200

T7 Promoter Positions

TACGACTGCA TATAGGGAAG AAGCTTGGAT AGGAGAATTC CGGCGAAGTT 250

-10 +1 +5

CGACGAAAGCTT TCGAGGAAGT AAGGAAGCTA AAATGGGAAAA AAAAAATCACT 300
GATATACCG CCGTTGATAT AATACCGCTA GGGTCAAGTT GATATGCACT 350
GCCGGGAGCG GATTTGAAAG TGGCAGAAAG ACGGCCCCGA GGCGGGGGG 400
ATATTACGCG CTTTGGGAAAG ACCTGGGATTC CAAATTTTAT 450
CCGGCCTTTA TCCACATTCT TGGCAGGCTT ATGAATGCTT ATCCGGAATT 500
CCGTATACGCA ATGAAAGACG GTGAGCTGGT GATATGGGAT AGTGTTCACC 550
CTTGTTCACGC GGGCAATACT TACGACTGCA TATAGGGAAG AAGCTTGGAT 600

AGTGAATACCG CACGAGAGTT TGGCAGGCTA AAGGAAGCTA AAATGGGAAAA AAAAAATCACT 650

TGTGCCGCTG TATCGTGAAA ACCTGGGCTT TTTTCAAAA GGGTTTATTG 700
AGATATACCG CCGTTGATAT AATACCGCTA GGGTCAAGTT GATATGCACT 750
GATTTAAAGC TGGCAGGCTA AAGGAAGCTA AAATGGGAAAA AAAAAATCACT 800

GGCCAAATAT TATACCGCAG CCGCGAGTCT TATACCAAGAG ATACGAGTTT 850
AGGTTCATCA TGGCGTCTGT GATGCGTTTC ATGCGGCGGT AATGCTATAT 900

GAATTACAGG TGGCAGGCTA AAGGAAGCTA AAATGGGAAAA AAAAAATCACT 950

AGGCAGTTAT TGGTGTTTCT TAAACCCCCTG TGGGCGCTTG TACGAGATT 1000

GATTCAAAGC TCAAGAATTT GAGGAAAATC AATTTCCAGC GAGACTGTAT 1050

AATGGCGAAA CCAACCCCTG GCAGAACATA CCGACCCCTG TACGAGATT 1100

CAGCAGGCGC ACGCGCGCCG TGTGCGGCTG TGGCAGGTGG GAGAAGATAT 1150

TTTCAGCCA CAGCAGGAGGGGGTTG TATACCCAGC AATACCGCTA GGGTCAAGTT GATATGCACT 1200

AGAATTTGCC TGGCGGCACTG ACAGCCTTTTT AATACCGCTA GGGTCAAGTT GATATGCACT 1250

AACCATCGA AAAAACCCCGT TAGCGTGGTGG GGGTCGCGCA CTGATCCAGC 1300

TGCGAGATT GGGGCGGCGTG TATACCGCTA GGGTCAAGTT GATATGCACT 1350

-10 +1 +5

GGTAGAGCTA GGGGATTTTT TCTTCTGTGT TCTGTGTGTA ATACGAGTTT 1400

AAAGACTGGA CACGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1450

AAGACAGGTT GGGGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1500

CGGGAGGGTT GGGGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1550

AAAGACTGGA CACGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1600

CTTTTGTGG TAGAAATGACTGTC TATACCGCTA GGGTCAAGTT GATATGCACT 1650

AAAGACTGGA CACGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1700

CTGAGGATT GGGGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1750

AAAGACTGGA CACGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1800

CTGAGGATT GGGGCGGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1850

CGAGAGTATT AAGCTGCGC TATACCGCTA GGGTCAAGTT GATATGCACT 1900
<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AGACTGGGCC</td>
<td>TTTCGTTTTA</td>
<td>TCTGTTTGT</td>
<td>TCTGCTTAAG</td>
<td>GCTCTCCTGA</td>
</tr>
<tr>
<td>GTAGGACAAAA</td>
<td>TCCGCCCGGAA</td>
<td>GCGGAGTTTA</td>
<td>AGCTTGCGAA</td>
<td>GCAACGCGCC</td>
</tr>
<tr>
<td>GGAGGGTGGC</td>
<td>GGGCAGACGG</td>
<td>CCGCCCATAA</td>
<td>ACTGCCAGGT</td>
<td>ATCAAATTAA</td>
</tr>
<tr>
<td>GCAAGAGGCC</td>
<td>ATCTCAGGGA</td>
<td>ATGGGTTTTT</td>
<td>TGGTTTTCTA</td>
<td>CAAACCTTC</td>
</tr>
<tr>
<td>CTGTCGTCAT</td>
<td>ATCTACACGC</td>
<td>CATCCCCCCA</td>
<td>CAGATACGGT</td>
<td>AAACCTAGCT</td>
</tr>
<tr>
<td>CTTTTTTGCA</td>
<td>TCAAGGAAAAC</td>
<td>AGTCCGGCAG</td>
<td>CTGGTGTTCC</td>
<td>TGGCCACGGG</td>
</tr>
<tr>
<td>TGCGCATGAT</td>
<td>CGOTCACTTG</td>
<td>TCCTTAAGGA</td>
<td>CCGCGCTAGG</td>
<td>CTGCGGCGG</td>
</tr>
<tr>
<td>TGCTTACTG</td>
<td>GTGAGCAAGAA</td>
<td>TGAATCAGG</td>
<td>ATCCCGCCTTG</td>
<td>ATCCCTGAG</td>
</tr>
<tr>
<td>TGCGTTTCCG</td>
<td>TTGTTCTGAA</td>
<td>AGCTCGGAAG</td>
<td>CCGGAGAATC</td>
<td>AGGGCCCTGC</td>
</tr>
<tr>
<td>ACCATTATGT</td>
<td>TCGGGATCTG</td>
<td>CATCGCAAGA</td>
<td>TGCTGCTGGC</td>
<td>TACCCTGCTG</td>
</tr>
<tr>
<td>AACACCTACA</td>
<td>TCTGTATTTA</td>
<td>CAGAAGGCGT</td>
<td>CTGATTGACG</td>
<td>TAGGTTATCT</td>
</tr>
<tr>
<td>TTCTCTGTCG</td>
<td>CCGCGCATCT</td>
<td>CATACCGCCA</td>
<td>GTTGGTTTAC</td>
<td>CTCAAACGCT</td>
</tr>
<tr>
<td>TCAACGATTC</td>
<td>ACAGGATTTA</td>
<td>CCTGCTAGTA</td>
<td>CCAAACTGGT</td>
<td>ACCGATCCC</td>
</tr>
<tr>
<td>CTCTGTTTTT</td>
<td>ATCGTGATCA</td>
<td>TGAAACCTTC</td>
<td>GAACGAGAAT</td>
<td>TCCCCCTTAC</td>
</tr>
<tr>
<td>AGGCGGCATT</td>
<td>CAAAGGCGCA</td>
<td>TGAATTTGGA</td>
<td>CAAACCGCGT</td>
<td>ACCGACGCTA</td>
</tr>
<tr>
<td>GCTTTATCAG</td>
<td>AAGCCAGACA</td>
<td>TAAAGCCTTC</td>
<td>TGGAGAAAAT</td>
<td>CACCGAGCTG</td>
</tr>
<tr>
<td>GACGCGCGTA</td>
<td>AACAGCCAGA</td>
<td>CATCTGCGA</td>
<td>TCGTCTACG</td>
<td>ACAGCCTGA</td>
</tr>
<tr>
<td>TGAOCTTTAC</td>
<td>CGCAGCTGCC</td>
<td>TCCGCCTTTT</td>
<td>CGGTAGATAC</td>
<td>GTGAAACCC</td>
</tr>
<tr>
<td>TCTGACACAT</td>
<td>GCAOCTCCAA</td>
<td>GAGACGGTCA</td>
<td>CAGCTTTGTG</td>
<td>GTAAACGGAT</td>
</tr>
<tr>
<td>GCCGGGAGCA</td>
<td>GAAAGCCCGC</td>
<td>TCAGGGCGCG</td>
<td>TCGGGCGG</td>
<td>TTGCGG</td>
</tr>
<tr>
<td>TCAGGGCCCA</td>
<td>GCCATGGACC</td>
<td>AGTCACTGGT</td>
<td>CGATTGGC</td>
<td>GTGATACGG</td>
</tr>
<tr>
<td>GCTTAACCTG</td>
<td>GCAGGACTCAG</td>
<td>ACAGAGATGT</td>
<td>ACTGAGAAGT</td>
<td>CACCATTATG</td>
</tr>
<tr>
<td>GGTGGTAAT</td>
<td>AACAGGAGACA</td>
<td>TGGCTAAGGA</td>
<td>GAATAACCG</td>
<td>CATCGGCCGC</td>
</tr>
<tr>
<td>TCTTCCGCTT</td>
<td>CCTGCCCTAC</td>
<td>TGATCCTGCT</td>
<td>TGGTTGCTGG</td>
<td>TCGTGCTGG</td>
</tr>
<tr>
<td>GCCAGCCGTA</td>
<td>TCAAGCTCAG</td>
<td>CAAAGGCCG</td>
<td>ATACCTGGT</td>
<td>TCAACAGAA</td>
</tr>
<tr>
<td>CAGGGGATAA</td>
<td>CGCGAGGAAA</td>
<td>AACAAGTGGG</td>
<td>CAAAAGGCCA</td>
<td>GCAAAAGGCC</td>
</tr>
<tr>
<td>AGGAACCGTA</td>
<td>AAAAGCCGGA</td>
<td>TGAAGCCTGC</td>
<td>AGCTGGGCGG</td>
<td>TTTTCCCCA</td>
</tr>
<tr>
<td>CCCTGACGAG</td>
<td>CATCACAAAC</td>
<td>ATGACAGCCTG</td>
<td>AAGTCAAGAC</td>
<td>TGCGGAAAC</td>
</tr>
<tr>
<td>GCAGCCGACT</td>
<td>ATAAAGGATAC</td>
<td>CAGCGGTTCG</td>
<td>CCGGCTCCAA</td>
<td>CTCCCTGCTG</td>
</tr>
<tr>
<td>CGCTCTCCCTG</td>
<td>TTCCGACCTC</td>
<td>GCCGCTTACC</td>
<td>GGATACGGT</td>
<td>CGGCGTTTCT</td>
</tr>
<tr>
<td>CCCCCTCGGA</td>
<td>AAGCTGGGCGG</td>
<td>TTTTCTCAAG</td>
<td>TCTACGCTG</td>
<td>AGCTATCTC</td>
</tr>
<tr>
<td>GTTCGGGTGTA</td>
<td>GGGTCTGGCG</td>
<td>TCAAGGCTTG</td>
<td>GCTGTGTGCA</td>
<td>GGAAACCCCG</td>
</tr>
<tr>
<td>GTTACGCCGG</td>
<td>ACCGCCCTCG</td>
<td>CTATGGTGGT</td>
<td>AACCTGGT</td>
<td>TCAACGCTAA</td>
</tr>
<tr>
<td>CCGGCTAAGA</td>
<td>CAGCAGCTTAG</td>
<td>CCGACCTGCC</td>
<td>AGCAAGCCACT</td>
<td>GCTAACAGGA</td>
</tr>
<tr>
<td>TTAGCAGAACG</td>
<td>GAGGTATGAT</td>
<td>GGGGGTGTCA</td>
<td>CAGAGTCTT</td>
<td>GCAGATGCTG</td>
</tr>
<tr>
<td>CTTAACTCAG</td>
<td>GCTACACTAG</td>
<td>AAGGACAGTA</td>
<td>TGGTGATCAC</td>
<td>GCGCTGCTG</td>
</tr>
<tr>
<td>GAAGCCAGTT</td>
<td>ACCCTTCGGA</td>
<td>AAAAGGTTGG</td>
<td>TAGCTCTTTA</td>
<td>TCGGCGAACC</td>
</tr>
<tr>
<td>AAACCACCCGC</td>
<td>TGATACGGGT</td>
<td>GTTTTTTTT</td>
<td>TTTGCAAGCA</td>
<td>GCAGATTAGC</td>
</tr>
<tr>
<td>GCGAGAAAAA</td>
<td>AAGAGATCCA</td>
<td>AGAAGATCCTA</td>
<td>TGATCTTTTT</td>
<td>CTAGGGGTTC</td>
</tr>
<tr>
<td>TGACGCTCAG</td>
<td>TGAGACGAAA</td>
<td>ACTCACGTTA</td>
<td>AGGATTTTAT</td>
<td>GTCATGGAGT</td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 Base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Other nucleic acid
(A) DESCRIPTION: Consensus sequence of a T7 RNA polymerase promoter

(iii) HYPOTHETICAL: No
(iv) ANTI-SENSE: No
(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Bacteriophage T7
(B) STRAIN: Wild-type
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: Not Applicable
(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Consensus T7 RNA polymerase promoter
(B) LOCATION: 1 to 23
(C) IDENTIFICATION METHOD: Sequence comparison of the 17 different T7 promoters.
(D) OTHER INFORMATION: The consensus sequence recognized by T7 RNA polymerase.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Dunn, J. J., and Studier, F. W.
(B) TITLE: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements.
(C) JOURNAL: Journal of Molecular Biology
(D) VOLUME: 166
(E) ISSUE: 
(F) PAGES: 477-533
(G) DATE: 
(H) DOCUMENT NUMBER: 
(I) FILING DATE: 
(J) PUBLICATION DATE: 
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

TAATACGACT CACTATAGGG AGA 23

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 Bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single Stranded
(D) TOPOLOGY: Linear

(iii) MOLECULE TYPE: Other nucleic acid
(A) DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCAT-10-1. This oligo is designated WT.

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: Not applicable

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCAT-10-1. This oligo is designated WT.
(B) LOCATION: 1 to 23
(C) IDENTIFICATION METHOD: Synthesized
(D) OTHER INFORMATION: The oligo contains a consensus T7 RNA polymerase promoter.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warshamana, G.S.
(B) TITLE: T7 promoter contacts essential for promoter activity in vivo
CTGAATTCCGA AATTAAATACG ACTCACTATA GGGAGAAGC TTGGTACCAG 50

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 Bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single Stranded
(D) TOPOLOGY: Linear

(iii) MOLECULE TYPE: Other nucleic acid
(A) DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCM-B® (c) 1990, Microsoft Corp designated B.

(iii) HYPOTHETICAL: No
(iv) ANTI-SENSE: No
(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: Not applicable

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCM-B® (c) 1990, Microsoft Corp
(B) LOCATION: 1 to 50
(C) IDENTIFICATION METHOD: Synthesized
(D) OTHER INFORMATION: The oligo contains a randomly mutagenized T7 RNA polymerase promoter. The bases listed in lower cases have a 96.1% probability of being the base listed and a 3.9% probability of being one of the other three possible bases.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warshamana, G.S.
(B) TITLE: T7 promoter contacts essential for promoter activity in vivo
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 20
(E) ISSUE: 20
(F) PAGES: 2517-2524
(G) DATE: May 25, 1992
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 50

(ix) SEQUENCE DESCRIPTION: SEQ ID NO:18:
CTGAATTCCGAAATTAATACGACTCACTATAGGAGAAGC TTGGTACCAG 50
INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 Bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single Stranded
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Other nucleic acid
(A) DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCM-T#copyright (c) 1990, Microsoft Corp designated T.

(iii) HYPOTHETICAL: No
(iv) ANTI-SENSE: No
(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYP: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: Not applicable

(viii) POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

(ix) FEATURE:
(A) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCM-T#copyright (c) 1990, Microsoft Corp designated T.
(B) LOCATION: 1 to 50
(C) IDENTIFICATION METHOD: Synthesized
(D) OTHER INFORMATION: The oligo contains a randomly mutagenized T7 RNA polymerase promoter. The bases listed in lower cases have a 96.1% probability of being the base listed and a 3.9% probability of being one of the other three possible bases.

(x) PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warshamana, G.S.
(B) TITLE: T7 promoter contacts essential for promoter activity in vivo
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 20
(E) ISSUE: 20
(F) PAGES: 2517-2524
(G) DATE: May 25, 1992
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 50

(x) SEQUENCE DESCRIPTION: SEQ ID NO:19:
CTAGTACTga aattaatacg actcactata gggagaAAGC TTGGTACCAG 50

INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 Bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single Stranded
(D) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Other nucleic acid
(A) DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCM-P#copyright (c) 1990, Microsoft Corp designated P.
(i) HYPOTHETICAL: No
(ii) ANTI-SENSE: No
(iii) FRAGMENT TYPE: Not applicable

ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: Not applicable

POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

FEATURE:
(A) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCM-P® (© 1990, Microsoft Corp
(B) LOCATION: 1 to 50
(C) IDENTIFICATION METHOD: Synthesized
(D) OTHER INFORMATION: The oligo contains a T7 promoter that is randomly mutagenized at promoter positions 9 through 7. The bases listed in lower cases have a 67% probability of being the base listed and a 33% probability of being one of the other three possible bases.

PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warshamana, G.S.
(B) TITLE: T7 promoter contacts essential for promoter activity in vivo
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 20
(E) ISSUE: 20
(F) PAGES: 2517-2524
(G) DATE: May 25, 1992
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 50

SEQUENCE DESCRIPTION: SEQ ID NO:20:

CTAGTACTGA AATTAATACG ActcACTATA GGGAGAAAGC TTGGTACCAG 50

INFORMATION FOR SEQ ID NO:21:

SEQUENCE CHARACTERISTICS:
(A) LENGTH: 50 Bases
(B) TYPE: Nucleic Acid
(C) STRANDNESS: Single Stranded
(D) TOPOLOGY: Linear

MOLECULE TYPE: Other nucleic acid

DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCM-P® (© 1990, Microsoft Corp designated C.

HYPOTHETICAL: No

ANTI-SENSE: No

FRAGMENT TYPE: Not applicable

ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
133

-continued

( G ) CELL TYPE: Not applicable
( H ) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
( A ) LIBRARY: Not applicable
( B ) CLONE: Not applicable

(viii) POSITION IN GENOME:
( A ) CHROMOSOME/SEGMENT: Not applicable
( B ) MAP POSITION: Not applicable
( C ) UNITS: Not applicable

(ix) FEATURE:
( A ) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCM-C® (c) 1990, Microsoft Corp
( B ) LOCATION: 1 to 50
( C ) IDENTIFICATION METHOD: Synthesized
( D ) OTHER INFORMATION: The oligo contains a T7 RNA polymerase promoter that carries an A to C mutation at position -10 of the promoter. This base is shown as a lower case letter.

(x) PUBLICATION INFORMATION:
( A ) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warchamana, G.S.
( B ) TITLE: T7 promoter contacts essential for promoter activity in vivo
( C ) JOURNAL: Nucleic Acids Research
( D ) VOLUME: 20
( E ) ISSUE: 20
( F ) PAGES: 2517-2524
( G ) DATE: May 25, 1992
( H ) DOCUMENT NUMBER:
( I ) FILING DATE:
( J ) PUBLICATION DATE:
( K ) RELEVANT RESIDUES IN SEQ ID NO: 1 to 50

(x) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTAGTACTGA AATTAATACG cCTCACTATA GGGAGAAAGC TTGGTACCAG 50

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
( A ) LENGTH: 50 Bases
( B ) TYPE: Nucleic Acid
( C ) STRANDEDNESS: Single Stranded
( D ) TOPOLOGY: Linear

(ii) MOLECULE TYPE: Other nucleic acid
( A ) DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCM-C® (c) 1990, Microsoft Corp designated G.

(iii) HYPOTHETICAL: No

(iv) ANTI-SENSE: No

(v) FRAGMENT TYPE: Not applicable

(vi) ORIGINAL SOURCE:
( A ) ORGANISM: Not applicable
( B ) STRAIN: Not applicable
( C ) INDIVIDUAL ISOLATE: Not applicable
( D ) DEVELOPMENTAL STAGE: Not applicable
( E ) HAPLOTYPE: Not applicable
( F ) TISSUE TYPE: Not applicable
( G ) CELL TYPE: Not applicable
( H ) ORGANELLE: Not applicable

(vii) IMMEDIATE SOURCE:
( A ) LIBRARY: Not applicable
( B ) CLONE: Not applicable

(viii) POSITION IN GENOME:
( A ) CHROMOSOME/SEGMENT: Not applicable
( B ) MAP POSITION: Not applicable
( C ) UNITS: Not applicable

(ix) FEATURE:
( A ) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCM-C® (c) 1990, Microsoft Corp
The oligo contains a T7 RNA polymerase promoter that carries an A to G mutation as position -10 of the promoter. This base is shown as a lower case letter.

PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warshamana, G.S.
(B) TITLE: T7 promoter contacts essential for promoter activity in vivo
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 20
(E) ISSUE: 20
(F) PAGES: 2517-2524
(G) DATE: May 25, 1992
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES IN SEQ ID NO: 1 to 50

SEQUENCE DESCRIPTION: SEQ ID NO:22:
CTAGTACTGA AATTAATAC gCTCACTATA GGGAGAAAGC TTGGTACCAG

INFORMATION FOR SEQ ID NO:23:
SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 Bases
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single Stranded
(D) TOPOLOGY: Linear

MOLECULE TYPE: Other nucleic acid
(A) DESCRIPTION: Synthetic deoxyoligonucleotide used in the construction of pCM-X© 1990, Microsoft Corp designated "Primer."

HYPOTHETICAL: No
ANTI-SENSE: No
FRAGMENT TYPE: Not applicable

ORIGINAL SOURCE:
(A) ORGANISM: Not applicable
(B) STRAIN: Not applicable
(C) INDIVIDUAL ISOLATE: Not applicable
(D) DEVELOPMENTAL STAGE: Not applicable
(E) HAPLOTYPE: Not applicable
(F) TISSUE TYPE: Not applicable
(G) CELL TYPE: Not applicable
(H) ORGANELLE: Not applicable

IMMEDIATE SOURCE:
(A) LIBRARY: Not applicable
(B) CLONE: Not applicable

POSITION IN GENOME:
(A) CHROMOSOME/SEGMENT: Not applicable
(B) MAP POSITION: Not applicable
(C) UNITS: Not applicable

FEATURE:
(A) NAME/KEY: Synthetic deoxyoligonucleotide used in the construction of pCM-X© 1990, Microsoft Corp designated "Primer."
(B) LOCATION: 1 to 14
(C) IDENTIFICATION METHOD: Synthesized
(D) OTHER INFORMATION: The "Primer" oligo is complementary to the 14 nucleotides at the 3'end of oligos WT, B, T, P, C, and G, and is used for priming the synthesis of a DNA strand complementary to the WT, B, T, P, C, and G oligos.

PUBLICATION INFORMATION:
(A) AUTHORS: Ikeda, R.A., Ligman, C.M., and Warshamana, G.S.
(B) TITLE: T7 promoter contacts essential for promoter activity in vivo
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 20
What is claimed is:

1. An isolated and purified mutant T7 RNA polymerase that carries the substitution of lysine for glutamic acid at amino acid 222.

2. The mutant T7 RNA polymerase as claimed in claim 1, having the ability to recognize and utilize point mutations of wild-type T7 promoter sequences.

3. The mutant T7 RNA polymerase as claimed in claim 2, wherein said point mutations of wild-type T7 promoter sequences are selected from the group consisting of all those promoter sequences recognized and utilized by a wild-type T7 RNA polymerase, and the following substitutions of the T7 promoter: cytidine to adenosine at -7, cytidine to guanosine at -7, thymidine to adenosine at -8, cytidine to adenosine at -9, cytidine to thymidine at -9, cytidine to guanosine at -9, and guanosine to thymidine at -11.

4. The mutant T7 RNA polymerase as claimed in claim 3, wherein said point mutations of wild-type T7 promoter sequences are selected from the group consisting of the following substitutions of the wild-type T7 promoter: cytidine to adenosine at -7, cytidine to guanosine at -7, thymidine to adenosine at -8, cytidine to adenosine at -9, cytidine to thymidine at -9, cytidine to guanosine at -9, and guanosine to thymidine at -11.

* * * * *

CTGGTACCAA GCTT 14

5,385,834