JOURNAL OF BACTERIOLOGY, Jan. 1991, p. 514–520
0021-9193/91/020514-07$02.00/0

Vol. 173, No. 2

Nucleotide Sequence of *Escherichia coli katE*, Which Encodes Catalase HPII

INGEMAR VON OSSOWSKI, MICHAEL R. MULVEY, PAMELA A. LECO, ANDREW BORYS, AND PETER C. LOEWEN*

Department of Microbiology, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2

Received 6 August 1990/Accepted 14 November 1990

A 3,466-bp nucleotide sequence containing the *katE* gene of *Escherichia coli* has been determined. An open reading frame of 2,259 bp was found and was preceded by a potential ribosome-binding site. The predicted N-terminal sequence agreed with the sequence determined by direct amino acid sequencing, and the predicted direction of transcription was confirmed by expression of the gene cloned in both directions behind a T7 promoter. The start site of transcription was determined to be 127 bp upstream from the start of the open reading frame, and a potential RNA polymerase-binding site similar to a sequence preceding the *xthA* gene, which is also controlled by the KatF protein, was identified. The predicted sequence of the 753-amino-acid protein was compared with known sequences of other catalases, revealing significant similarity to the shorter catalases, including the residues in the putative active site and residues involved in heme binding.

Escherichia coli produces two catalases, a bifunctional catalase-peroxidase (hydroperoxidase I [HPI]) and a monofunctional catalase (HPII). Both enzymes have been purified and characterized to be quite different from each other and from the typical catalase, which is active as a tetramer of 65,000-Da subunits and four protoheme IX groups. HPI was found to be a tetramer of 78,000-Da subunits with just two protoheme IX groups per tetramer and with an associated broad-spectrum peroxidase activity (6). HPII was purified (7) and characterized as a hexamer of 93,000-Da subunits (18) and six heme d-like groups (5) per hexamer. The genes encoding HPI and HPII, *katG* and *katE*, respectively, are unlinked, mapping at 89.2 (21) and 37.2 (17) min, respectively. A third gene, *katF*, mapping at 59.0 min (20), was found to be required for expression of *katE* but not *katG*. The *katF* gene has been cloned (28) and sequenced (27), revealing a striking similarity between the KatF protein and known sigma transcription factors, suggesting its mechanism as a positive effector of *katE* (27), of *xthA* (32), and possibly of other genes involved in resistance to near-UV radiation (33).

The *katG* gene has been cloned (22, 42) and sequenced (41) and predicts an amino acid sequence that bears no resemblance to any of the known catalase sequences. Instead, a striking resemblance to a peroxidase from *Bacillus stearothermophilus* has been observed (23), suggesting that the bifunctional HPI is more closely related to the family of peroxidases than catalases. On the basis of its very different physical structure, it seemed unlikely that HPII would have any sequence or structural similarity to the common catalases. This paper describes the sequence analysis of *katE*, the identification of the site of transcription initiation, and a comparison of the predicted amino acid sequence of HPII with the sequences of other catalases.

MATERIALS AND METHODS

Bacterial strains and plasmids. The strains JM101 (45) and NM522 (25) were used for plasmid preparations and for the preparation of single-stranded DNA for sequencing. Cultures were grown in LB medium (26) which was supplemented with 150 μg of ampicillin per ml for transformed strains. The plasmid pAMkatE72 was prepared by ligating a 3.5-kb *Pst*I-*Cla*I fragment containing *katE* from pAMkatE6 (28) (Fig. 1) into the M13 Bluescript KS$^+$ vector (Stratagene Cloning Systems). Subclones of pAMkatE72 were prepared in the M13 Bluescript SK$^+$, SK$^-$, KS$^+$, and KS$^-$ vectors by using various combinations of restriction nucleases. The 4.0-kb *Cla*I fragment from pAMkatE6 was cloned into pT7-5 (38) in two orientations relative to the T7 promoter to generate pT7E1 and pT7E2 (Fig. 1).

Expression of *katF*. The plasmids pT7E1 and pT7E2 were transformed into UM255 (28) along with pGP1-2. The double transformants were heat shocked by transferring a cell culture from 30 to 42°C as described previously (38). After 30 min, the cells were collected and lysed by sonication. Extracts of the crude extract were separated by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (16, 44). The proteins were labeled with [^{35}S]methionine, and the labeled protein bands were localized by autoradiography (4).

DNA manipulation and sequence analysis. DNA manipulations, including restriction digests, ligations, transformations, and electrophoresis, were carried out as described by Maniatis et al. (24). DNA sequencing was by the dideoxy chain termination method of Sanger et al. (34). The following synthetic oligonucleotides were used as primers in the chain elongation reactions from the multiple cloning site into the cloned regions of the hybrid M13 plasmids: universal primer, 5′-GTAAAACGACGGCCAGT; reverse primer, 5′-AACA GCTATGACCATG; SK primer, 5′-TCTAGAACTAGTG GATC; and the KS primer, 5′-CGAGGTGGCGACGGTA TCG. The following synthetic oligonucleotides corresponding to portions of the *katE* gene were used for sequencing: 1, 5′-TTTGCGTGTATTTCATA; 2, 5′-TTCGGTATTCACAC CTT; 3, 5′-CGCTGTTGCAGGGACGT; 4, 5′-ACGTCCCT GCAACAGCG; 5, 5′-CACGTTGTGCAGAGTTT; 6, 5′-GG TAGGTTGTGCACCTG; and 7, 5′-TACCACGCGACCTT TCA.

The transcription initiation site was identified by the primer extension method (1) by using reverse transcriptase to elongate the primer 5′[^{32}P]CTGATGTGGGTTCTTTTCG

* Corresponding author.

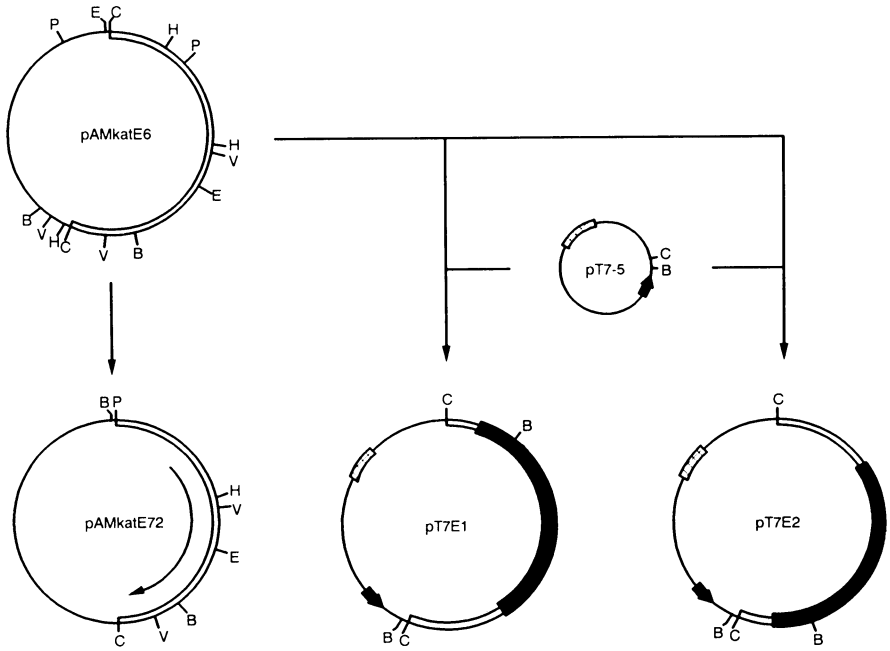

FIG. 1. Construction of various *katE*-containing plasmids. For the construction of pAMkatE72, used for sequence analysis, the plasmid pAMkatE6 (28) was cut with *Pst*I and *Cla*I and the 3.5-kb fragment containing *katE* was inserted into pM13KS$^+$, also cut with *Pst*I and *Cla*I in the multiple cloning site. For the construction of pT7E1 and pT7E2, used for determining the direction of transcription, pAMkatE6 was cut with *Cla*I and the 4.2-kb fragment was inserted in pT7-5, also cut with *Cla*I. The two *Bam*HI sites in pT7E1 and pT7E2 were used to determine the orientation of the insert relative to the T7 promoter, indicated by the large arrows in pT7-5, pT7E1, and pT7E2. The unfilled portion of the various plasmid outlines indicates the chromosomal DNA insert, and the heavy filled segments in pT7E1 and pT7E2 indicate the location of the *katE* gene. The arrow inside pAMkatE72 indicates the approximate location and direction of transcription of *katE*. The cross-hatched segments indicate the location of the *bla* gene in the pT7 plasmids. The following abbreviations were used for restriction enzymes: C, *Cla*I; B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; and V, *Eco*RV.

TTATGTTGCGA, which corresponds to the complement of bases +4 to +33 of *katE* annealed to mRNA synthesized from plasmid pAMkatE72. Total RNA was isolated as described by Gilman and Chamberlin (12), with the following procedural modifications. Cells were ruptured by extracting twice with an equal volume of redistilled phenol, and the subsequent nucleic acid precipitate was redissolved in 100 mM sodium acetate (pH 5) and 10 mM $MgSO_4$ for the DNase I and proteinase K treatments.

N-terminal sequence analysis. Catalase HPII was purified as described by Loewen and Switala (18) from strain UM255 (39) transformed with pAMkatE22 (28). The N-terminal sequence was determined by the Tripartite Microanalytical Center at the University of Victoria, Victoria, Canada.

Amino acid analysis. Samples containing 1 nmol of HPII subunit were mixed with constant boiling HCl, sealed in a glass tube, and incubated at 110°C for 24, 48, and 72 h. For cysteic acid determination, 1 nmol of HPI was oxidized for 20 h at 5°C in performic acid (prepared by mixing 0.2 ml of 30% H_2O_2 with 1.8 ml of 90% formic acid), followed by lyophilization and hydrolysis. The amino acid analyses were performed on an LKB 4151 Alpha Plus amino acid analyzer operated by the Department of Animal Science, University of Manitoba.

Nucleotide sequence accession number. The GenBank accession number for the *katE* gene is M55161.

RESULTS

Direction of transcription. Plasmids pT7E1 and pT7E2 were constructed with the *Cla*I fragment of pAMkatE72 inserted in the two possible orientations behind the T7 promoter of plasmid pT7-5 (Fig. 1). They were then transformed into a strain containing the plasmid pGP1-2, which encodes T7 RNA polymerase downstream from the lambda P_L promoter and is controlled by the heat-sensitive cI857 repressor. Heat induction of the doubly transformed strain resulted in expression of T7 RNA polymerase from pGP1-2, which in turn initiated transcription of DNA downstream from the T7 promoter on the pT7 plasmids. Of the two plasmids containing *katE*, only pT7E1 yielded a 93-kDa protein band labeled with [^{35}S]methionine (Fig. 2). This indicated that the direction of transcription was from the *Pst*I end of pAMkatE72 towards the *Cla*I end, as shown in Fig. 1.

Nucleotide sequence of *katE*. The size of the chromosomal insert containing *katE* in pAMkatE72 was 3.5 kb, of which approximately 2.5 kb was required to encode the 93,000-Da HPII subunit. The strategy for sequencing both strands of the *katE*-containing insert in pAMkatE72 is shown in Fig. 3 and resulted in the definition of a 3,466-bp sequence, extending from the *Pst*I to the *Cla*I sites (Fig. 4). A single open reading frame of 2,259 bp (821 to 3079) which encodes a protein of 753 amino acids is indicated in Fig. 4. The molecular weight of the predicted protein was 84,173 Da, smaller than the 93,000 Da observed on SDS-polyacrylamide gels (18), possibly a result of its slightly acidic character. The predicted amino acid composition of the protein is compared with the experimentally determined composition in Table 1, revealing close agreement and a slight predominance of acidic (14.2%) over basic (10.8%) residues. Analysis of the

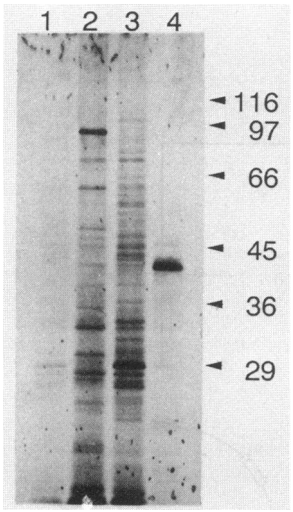

FIG. 2. Analysis of proteins expressed from the pT7 *katE*-containing plasmids to determine the direction of transcription of *katE*. Extracts from cells transformed with the following plasmids were separated on a 9% gel: pGP1-2 (lane 1), pGP1-2 and pT7E1 (lane 2), pGP1-2 and pT7E2 (lane 3), and pGP1-2 and pT7F2 (lane 4) (27). The location and sizes (10^3) of various molecular weight markers are indicated alongside the gel.

codon usage pattern suggests that *katE* is a moderately expressed gene, with codons characteristic of both weakly and strongly expressed genes being used (13). The N-terminal sequence of the protein was determined to be xxx-Ser-Gln-His-Asn-Glu-Lys-Asn-Pro-His-Gln, consistent with the predicted amino acid sequence from residues 2 to 10. The identity of the first amino acid was not determined for technical reasons.

Potential control regions flanking *katE*. There is a 5-base ribosome-binding or Shine-Dalgarno sequence, AGGAG, 10 bp upstream from the ATG codon. The potential −10 and −35 regions of an RNA polymerase-binding site are indicated at 133 and 154 bp upstream from the ATG codon. These sequences are very different from the consensus −10 and −35 sequences (14) but were chosen on the basis of the

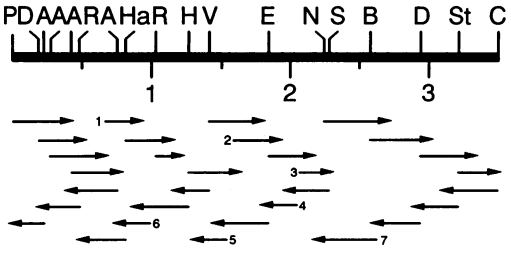

FIG. 3. Sequencing strategy for *katE*. The restriction map of the chromosomal insert in pAMkatE72 was used to generate a number of subclones of portions of the insert which were used for sequencing. The arrows indicate the direction and extent of elongation in the sequencing reactions. Where the arrow commences with a number, the sequence was initiated with a synthetic primer, and the number corresponds with the list of primers in Materials and Methods. Both strands of the insert were sequenced. In addition to the restriction enzyme abbreviations used in Fig. 1, the following abbreviations were used: D, *Dra*I; A, *Sau*3A; R, *Rsa*I; Ha, *Hae*III; N, *Nru*I; S, *Ssp*I; and St, *Stu*I.

```
                                                                      91
CTGCAGCCTTTCTTTAAAAGAGTCGAAAGCCAGGCTTTTAATATTTAAATCACCATAATTACTCTGTATTAAGTTTGTAGAAAACATCTCC
                                                                     181
CGCCTCATATTGTTAACAAAATTATTATCTCATTTAAATCTAAGTCATTTACAATATAAGTTTAAGAGCGACGCCACAGGATGAACTATC
                                                                     271
AAAAATAGCTCATCATGATTAGCAAAACTTAACCATTTTAAAATAAATAAACAATTAAAGAAAAAAGATCACTTATTTATAGCAATAGAT
                                                                     361
CGTCAAAGGCAGCTTTTTGTTACAGGTGGTTTGAATGAATGTAGCAACGAAATACAGAATTTCAGGTCATGTAACTCCCGGCAAACCGGG
                                                                     451
AGGTATGTAATCCTTACTCAGTCACTTCCCCTTCCTGGCGGATCTGATTTGCCCAACGTTGGGCAGATTCAGGCACAGTAAACGCCGGTG
                                                                     541
AGCGCAGAAATGACTCTCCCATCAGTACAAACGCAACATATTTGCCACGCAGCATCCAGACATCACGAAACGAATCCATCTTTATCGCAT
                                                                     631
GTTCTGGCGGCGCGGGTTCCGTGCGTGGGACATAGCTAATAATCTGGCGGTTTTGCTGGCGGAGCGGTTTCTTCATTACTGGCTTCACTA
                                                                     721
AACGCATATTAAAAATCAGAAAAACTGTAGTTTAGCCGATTTAGCCCCTGTACGTCCCGCTTTGCGTGTATTTCATAACACCGTTTCCAG
                                  -35                 -10       *
                                                                     811
AATAGTCTCCGAAGCGGGATCTGGCTGGTGGTCTATAGTTAGAGAGTTTTTGACCAAAACAGCGGCCCTTTCAGTAATAAATTAAGGAG
                                                                    S.D.
                                                                     901
ACGAGTTCAATGTCGCAACATAACGAAAAGAACCCACATCAGCACCAGTCACCACTACACGATTCCAGCGAAGCGAAACCGGGGATGGAC
          M  S  Q  H  N  E  K  N  P  H  Q  H  Q  S  P  L  H  D  S  S  E  A  K  P  G  M  D
                                                                     991
TCACTGGCACCTGAGGACGGCTCTCATCGTCCAGCGGCTGAACCAACACCGCCAGGTGCACAACCTACCGCCCCAGGGAGCCTGAAAGCC
 S  L  A  P  E  D  G  S  H  R  P  A  A  E  P  T  P  P  G  A  Q  P  T  A  P  G  S  L  K  A
                                                                    1081
CCTGATACGCGTAACGAAAAACTTAATTCTCTGGAAGACGTACGCAAAGGCAGTGAAAATTATGCGCTGACCACTAATCAGGGCGTGCGC
 D  T  R  N  E  K  L  N  S  L  E  D  V  R  K  G  S  E  N  Y  A  L  T  T  N  Q  G  V  R
                                                                    1171
ATCGCCGACGATCAAAACTCACTGCGTGCCGGTAGCCGTGGTCCAACGCTGCTGGAAGATTTTATTCTGCGCGAGAAAATCACCCACTTT
 A  D  D  Q  N  S  L  R  A  G  S  R  G  P  T  L  L  E  D  F  I  L  R  E  K  I  T  H  F
                                                                    1261
GACCATGAGCGCATTCCGGAACGTATTGTTCATGCACGCGGATCAGCCGCTCACGGTTATTTCCAGCCATATAAAAGCTTAAGCGATATT
 H  E  R  I  P  E  R  I  V  H  A  R  G  S  A  A  H  G  Y  F  Q  P  Y  K  S  L  S  D  I
                                                                    1351
ACCAAAGCGGATTTCCTCTCAGATCCGAACAAAATCACCCCAGTATTTGTACGTTTCTCTACCGTTCAGGGTGGTGCTGGCTCTGCTGAT
 K  A  D  F  L  S  D  P  N  K  I  T  P  V  F  V  R  F  S  T  V  Q  G  G  A  G  S  A  D
                                                                    1441
ACCGTGCGTGATATCCGTGGCTTTGCCACCAAGTTCTATACCGAAGAGGGTATTTTTGACCTCGTTGGCAATAACACGCCAATCTTCTTT
 V  R  D  I  R  G  F  A  T  K  F  Y  T  E  E  G  I  F  D  L  V  G  N  N  T  P  I  F  F
                                                                    1531
ATCCAGGATGCGCATAAATTCCCCGATTTTGTTCATGCGGTAAAACCAGAACCGCACTGGGCAATTCCACAAGGGCAAAGTGCCCACGAT
 Q  D  A  H  K  F  P  D  F  V  H  A  V  K  P  E  P  H  W  A  I  P  Q  G  Q  S  A  H  D
                                                                    1621
ACTTTCTGGGATTATGTTTCTCTGCAACCTGAAACTCTGCACAACGTGATGTGGGGCGATGTCGGATCGCGGCATCCCCCGCAGTTACCGC
 F  W  D  Y  V  S  L  Q  P  E  T  L  H  N  V  M  W  A  M  S  D  R  G  I  P  R  S  Y  R
                                                                    1711
ACCATGGAAGGCTTCGGTATTCACACCTTCCGCCTGATTAATGCCGAAGGGAAGGCAACGTTTGTACGTTTCCACTGGAAACCACTGGCA
 H  E  G  F  G  I  H  T  F  R  L  I  N  A  E  G  K  A  T  F  V  R  F  H  W  K  P  L  A
                                                                    1801
GGTAAAGCCTCACTCGTTTGGGATGAAGCACAAAAACTCACCGGACGTGACCCGGACTTCCACCGCCGCGAGTTGTGGGAAGCCATTGAA
 K  A  S  L  V  W  D  E  A  Q  K  L  T  G  R  D  P  D  F  H  R  R  E  L  W  E  A  I  E
                                                                    1891
GCAGGCGATTTTCCGGAATACGAACTGGGCTTCCAGTTGATTCCTGAAGAAGATGAATTCAAGTTCGACTTCGATCTTCTCGATCCAACC
 A  G  D  F  P  E  Y  E  L  G  F  Q  L  I  P  E  E  D  E  F  K  F  D  F  D  L  L  D  P  T
                                                                    1981
AAACTTATCCCGGAAGAACTGGTGCCCGTTCAGCGTGTCGGCAAAATGGTGCTCAATCGCAACCCGGATAACTTCTTTGCTGAAAACGAA
 K  L  I  P  E  E  L  V  P  V  Q  R  V  G  K  M  V  L  N  R  N  P  D  N  F  F  A  E  N  E
                                                                    2071
CAGGCGGCTTTCCATCCTGGGCATATCGTGCCGGGACTGGACTTCACCAACGATCCGCTGTTGCAGGGACGTTTGTTCTCCTATACCGAT
 Q  A  A  F  H  P  G  H  I  V  P  G  L  D  F  T  N  D  P  L  L  Q  G  R  L  F  S  Y  T  D
                                                                    2161
ACACAAATCAGTCGTCTTGGTGGGCCGAATTTCCATGAGATTCCGATTAACCGTCCGACCTGCCCTTACCATAATTTCCAGCGTGACGGC
 T  Q  I  S  R  L  G  G  P  N  F  H  E  I  P  I  N  R  P  T  C  P  Y  H  N  F  Q  R  D  G
                                                                    2251
ATGCATCGCATGGGGATCGACACTAACCCGGCGAATTACGAACCGAACTCGATTAACGATAACTGGCCGCGCGAAACACCGCCGGGGCCG
 M  H  R  M  G  I  D  T  N  P  A  N  Y  E  P  N  S  I  N  D  N  W  P  R  E  T  P  P  G  P
                                                                    2341
AAACGCGGCGGTTTTGAATCATACCAGGAGCGCGTGGAAGGCAATAAAGTTCGCGAGCGCAGCCCATCGTTTGGCGAATATTATTCCCAT
 K  R  G  G  F  E  S  Y  Q  E  R  V  E  G  N  K  V  R  E  R  S  P  S  F  G  E  Y  Y  S  H
                                                                    2431
CCGCGTCTGTTCTGGCTAAGTCAGACGCCATTTGAGCAGCGCCATATTGTCGATGGTTTCAGTTTTGAGTTAAGCAAAGTCGTTCGTCCG
 R  L  F  W  L  S  Q  T  P  F  E  Q  R  H  I  V  D  G  F  S  F  E  L  S  K  V  V  R  P
                                                                    2521
TATATTCGTGAGCGCGTTGTTGACCAGCTGGCGCATATTGATCTCACTCTGGCCCAGGCGGTGGCGAAAAATCTCGGTATCGAACTGACT
 I  R  E  R  V  V  D  Q  L  A  H  I  D  L  T  L  A  Q  A  V  A  K  N  L  G  I  E  L  T
                                                                    2611
GACGACCAGCTGAATATCACCCCACCTCCGGACGTCAACGGTCTGAAAAAGGATCCATCCTTAAGTTTGTACGCCATTCCTGACGGTGAT
 D  Q  L  N  I  T  P  P  P  D  V  N  G  L  K  K  D  P  S  L  S  L  Y  A  I  P  D  G  D
                                                                    2701
GTGAAAGGTCGCGTGGTAGCGATTTTACTTAATGATGAAGTGAGATCGGCAGACCTTCTGGCCATTCTCAAGGCGCTGAAGGCCAAAGGC
 K  G  R  V  V  A  I  L  L  N  D  E  V  R  S  A  D  L  L  A  I  L  K  A  L  K  A  K  G
                                                                    2791
GTTCATGCCAAACTGCTCTACTCCCGAATGGGTGAAGTGACTGCGGATGACGGAACGGTGTTGCCTATAGCCGCTACCTTTGCCGGTGCA
 V  H  A  K  L  L  Y  S  R  M  G  E  V  T  A  D  D  G  T  V  L  P  I  A  A  T  F  A  G  A
                                                                    2881
CCTTCGCTGACGGTCGATGCGGTCATTGTCCCTTGCGGCAATATCGCGGATATCGCTGACAACGGCGATGCCAACTACTACCTGATGGAA
 P  S  L  T  V  D  A  V  I  V  P  C  G  N  I  A  D  I  A  D  N  G  D  A  N  Y  Y  L  M  E
                                                                    2971
GCCTACAAACACCTTAAACCGATTGCGCTGGCGGGTGACGCGCGCAAGTTTAAAGCAACAATCAAGATCGCTGACCAGGGTGAAGAAGGG
 A  Y  K  H  L  K  P  I  A  L  A  G  D  A  R  K  F  K  A  T  I  K  I  A  D  Q  G  E  E  G
                                                                    3061
ATTGTGGAAGCTGACAGCGCTGACGGTAGTTTTATGGATGAACTGCTAACGCTGATGGCAGCACACCGCGTGTGGTCACGCATTCCTAAG
 I  V  E  A  D  S  A  D  G  S  F  M  D  E  L  L  T  L  M  A  A  H  R  V  W  S  R  I  P  K
                                                                    3151
ATTGACAAAATTCCTGCCTGATGGGAGCGCGCAATTGCGCCGCCTCAATGATTTACATAGTGCGCTTTGTTTATGCCGGATGCGCGTGAA
 I  D  K  I  P  A
                                                                    3241
CGCCTTATCCGGCCTACAAAACTGTGCAAATTCAATATATTGCAGGAAACACGTAGGCCTGATAAGCGAAGCCATCAGGCAGTTTTGCGT
                                                                    3331
TTGTCAGCAGTCTCAAGCGGCGGCAGTTACGCCGCCTTTGTAGGAATTAATCGCCGGATGCAAGGTTCACGCCGATCTGGCAAACATCCT
                                                                    3421
CACTTACACATCCCGATAACTCCCCAACCGATAACCACGCTGAGCGATAGCACCTTTCAACGACGCTGATGTCAACACATCCAGCTCCGT
                                                                    3466
TAAGCGTGGGAAACAGTAAGCACTCTGACGGATAGTATTATCGAT
```

FIG. 4. Sequence of the 3,466-bp chromosomal insert in pAMkatE72 containing *katE*. The open reading frame commences at base 821, and the predicted amino acid sequence is indicated in italics underneath the DNA sequence. The transcription start site, a G at position 695, is indicated by *, and potential −10 and −35 components of the promoter are underlined and labeled upstream of the start base. A potential palindromic sequence is underlined about 60 bp after the end of the open reading frame. A potential ribosome-binding site is underlined and labeled S.D. just upstream of the open reading frame.

TABLE 1. Comparison of predicted and actual amino acid compositions of catalase HPII

Amino acid	Predicted no. (%)	Actual no. (%)[a]
Alanine	64 (8.5)	61.2 (8.1)
Arginine	44 (5.8)	41.4 (5.5)
Asparagine	33 (4.4)	—
Aspartate	58 (7.7)	95.1 (12.6)
Cysteine[b]	2 (0.3)	2.2 (0.3)
Glutamine	26 (3.5)	—
Glutamate	49 (6.5)	86.3 (11.5)
Glycine	53 (7.0)	57.0 (7.6)
Histidine	28 (3.7)	22.7 (3.0)
Isoleucine	43 (5.7)	40.7 (5.4)
Leucine	60 (8.0)	57.7 (7.7)
Lysine	37 (4.9)	36.6 (4.9)
Methionine[c]	12 (1.6)	11.5 (1.5)
Phenylalanine	40 (5.3)	31.9 (4.2)
Proline	59 (7.8)	60.0 (8.0)
Serine[d]	40 (5.3)	31.8 (4.2)
Threonine	37 (4.9)	39.0 (5.2)
Tryptophan[e]	9 (1.2)	12.5 (1.7)
Tyrosine	19 (2.5)	13.8 (1.8)
Valine	40 (5.3)	43.7 (5.8)

[a] The asparagine and glutamine values are combined with the aspartate and glutamate values, respectively, because of hydrolysis of the amides.
[b] Determined as cysteic acid following performic acid oxidation.
[c] Determined as methionine solufoxide and methionine sulfone following performic acid oxidation.
[d] Determined by extrapolation of hydrolysis time to zero to compensate for hydrolytic destruction.
[e] Determined by the method of Edelhoch (8).

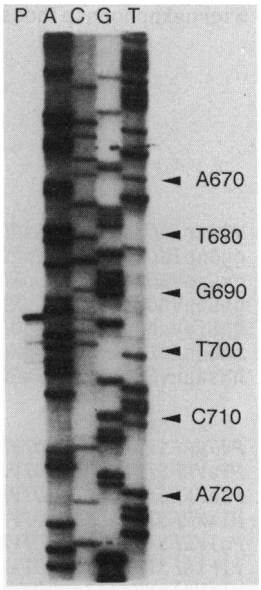

FIG. 5. Primer extension mapping of the *katE* promoter. The ^{32}P-labeled primer was elongated and electrophoresed on a 6% polyacrylamide sequencing gel in lane P alongside a sequencing ladder (lanes A, C, G, and T) generated by using the same primer. The sequence shown is of the strand complementary to that shown in Fig. 4, but the numbered bases shown on the right of the gel correspond to the sequence in Fig. 4. For the primer extension reaction, 25 μg of total RNA from NM522 carrying pAMkatE72 was used. The predominant and longest extended primer band migrated adjacent to the C complementary to G-695.

fact that the start site of transcription identified by primer extension on the mRNA is G at position 695 (Fig. 5). Furthermore, a similar sequence exists just upstream from the transcription start site in *xthA* (35) (Fig. 6), another gene controlled by the KatF protein. About 130 nucleotides beyond the putative termination codon, there is an element of twofold symmetry with a predicted stability of −9.1 kcal/mol which is followed by a run of four T's, indicative of a moderately strong transcription termination site.

Comparison of the predicted HPII sequence with other catalase sequences. A comparison of the deduced amino acid sequence of the HPII subunit with the known sequences of catalases from bovine liver, rat liver, human kidney, *Candida tropicalis*, *Saccharomyces cerevisiae*, and maize (*Zea mays*) revealed significant similarity between portions of HPII and all of the shorter catalases (Fig. 7). Because of its greater length, HPII contained long stretches at both the N terminus (57 amino acids) and the C terminus (168 amino acids) that were not similar to any parts of the shorter enzymes. However, rat liver (11), bovine liver (36), human kidney (2), *S. cerevisiae* (15), *C. tropicalis* (30), and maize (3) catalases were found to be, respectively, 51.8, 53.6, 51.2, 41.8, 49.4, and 46.9% similar (considering both identical amino acids and conservative replacements) to the remaining amino acids in the core of HPII. In fact, the degree of similarity was much greater for certain portions of the sequence; for example, residues 161 to 213 were 84.3% similar to the corresponding portion of the rat liver catalase. Generally, the similarity was greatest in the N-terminal half of the proteins and decreased towards the C-terminal end. Most significantly, the three residues identified as participating in the active site of the enzyme of the bovine liver enzyme, His-74, Ser-113, and Asn-147 (9), were conserved in the form of His-128, Ser-167, and Asn-201. Furthermore, the seven residues identified as interacting with the heme group of the bovine liver enzyme, Val-73, Thr-114, Phe-152, and Phe-160 on the distal side of the heme and Pro-335, Arg-353, and Tyr-357 on the proximal side of the heme, have as counterparts in the HPII sequence Val-127, Thr-168, Phe-206, Phe-214, Pro-393, Arg-411, and Tyr-415, respectively. Of the four amino acids involved in NADPH binding in the bovine liver enzyme (10), Arg-202, Asp-212, Lys-236, and His-304, the first three are conserved as Arg-260, Glu-270, and Lys-294, but the fourth, His-304, has been replaced by Glu-362.

DISCUSSION

The sequence of HPII deduced from the nucleotide sequence of *katE* contains a core segment that is approximately 50% similar to the complete sequences of six other catalases, considering both identical amino acids and con-

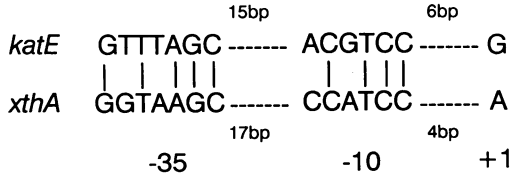

FIG. 6. Comparison of the DNA sequences upstream of the first transcribed base in *katE* and *xthA* (35). The initiation site is indicated by +1, and the −10 and −35 regions are indicated by −10 and −35. The spacing between the various sequences is also indicated.

```
HPII: msqhneknphqhqsplhdsseakpgmdslapedgshrpaaeptppgaqptapqSlkapDtrnEkLNsleDvRkGSENyaLTT
RLC:                                                        adsrDpasDqMQewkEqRapqkpdvLTT
BLC:                                                        adnrDpasDqMkhwkEqRaAqkpdvLTT
HKC:                                                        adsrDpasDqMQhwkEqRaAqkNdvLTT
SCC:                                                                   MNvfgk KeekQEkvySl
CTC:                                                                               aptfTn
ZMC:                                                     vrrSgsssslatamdpykharRapStp vLd

                                                      †*
HPII: NQGvrIADDqN SLRAGsRGPtLLDEFILrEKItHFDHERIPERIVHARGSAAHGYFQpyKSLSDIT   KAdFLsDpnKiT
RLC:  ggGnpIGDklN iMtAGpRGPlLVQDvVftDeMaHFDRERIPERVVHAKGaGAfGYFEvtH   DITrysKAkvfehigKrT
BLC:  ggGnpVGDklN SLtAGpRGPlLVQDvVftDeMaHFDRERIPERVVHAKGaGAfGYFEvtH   DITrysKAkvfehigKrT
HKC:  gaGnpVGDklN vItAGpRGPlLVQDvVftDeMaHFDRERIPERVVHAKGaGAfGYFEvtH   DITkysKAkvfehigKkT
SCC:  QNGfpyshhpyaSqysrpdGPiLLQDFhLlEnIasFDRERVPERVVHAKGgGcRleFEltdSLSDIT   yAapyqNvgykc
CTC:  sNGqpIpEpfa TqRvGqHGPlLLQDFnLiDsLaHFDRERIPERVVHAKGSGAyGvFEvtd   DITdvcaAkFLdtvgKkT
ZMC:  hNsaprvEQrQ lLtvGaRGPiLLEDlVV EKLanFDRERIPERVVHrcGasAKGFFEvtH   DITtdvadsagrtagmqa

        *†                                  *      †        †
HPII: PVFVRFSTVqGGaGSADTVRDiRGFATKFYTEEGiFDLVGNNTPIFFIqDAhKFPdFVHAvK  PEpHwaiPQgqsahDTF
RLC:  PIaVRFSTVaGesGSADTVRDpRGFAvKFYTEDGnWDLVGNNTPIFFIrDAmlFPsFIHsqKrnPQtHlkdP     Dmv
BLC:  PIaVRFSTVaGesGSADTVRDpRGFAvKFYTEDGnWDLVGNNTPIFFIrDAllFPsFIHsqKrnPQtHlkdP     Dmv
HKC:  PIaVRFSTVaGesGSADTVRDpRGFAvKFYTEDGnWDLVGNNTPIFFIrDpilFPsFIHsqKrnPQtHlkdP     Dmv
SCC:  PglVRFSTVgGesGTpDTVRDpRGvsfKFYTEwGnhDwVfNNTPVFFLrDAiKFPvFIHsqKrdPQsHlnqfQ     DTti
CTC:  rIFtRFSTVgGelGSADTaRDpRGFATKFYTEEGnlDLVyNNTPVFFIrDpsKFPhFIHtqKrnPEtHlkda     NmF
ZMC:  vIaVslncValAlryhDgspEtlGprvasvytgGnWt Vet lP slIrDGiKsghv HAIKpnPrtHigdn     wri

                     N         N                        N
HPII: WDYVSLQPETLHNVMWamSDRGIPrSYRTMEGFGIHTFRLINAEGKATFVRFHWKPlaGkaSLvwDEAqKLTGrDPDFHrR
RLC:  WDFwSLcPESLHQVtFlfSDRGIPdghRhMNGYGsHTFKLVNANGeAvYcKFHYKtdqGiknLpvEEAgRLaqeDPDYglR
BLC:  WDFwSLrPESLHQVsFlfSDRGIPdghRhMDGYGsHTFKLVNADGeAvYcKFHYKtdqGiknLsvEDAaRLaheDPDYglR
HKC:  WDFwSLrPESLHQVsFlfSDRGIPdghRhMNGYGsHTFKLVNANGeAvYcKFHYKtdqGikwLsvEDAaRLSqeDPDYgiR
SCC: yWDYLTLNPESIHQItYmfgDRGtPaSWaSMNAYsgHSFKMVNkEGKdTYVqFHvlsdtGfeTLtgDkAaeLSGshPDYnqa
CTC:  WDYLTTNeESVHQVMvlfSDRGtPaSYReMNGYsgHTYiwsNnkGewfYVqvHFisdqGikTLtnEEAgsLaGsNPDYaqe
ZMC:  lDFfShhPESLHmfsFlfdDcGIPadvRphgGsGVHTYtLVsraGtvTYVKFHWRPtcGvrSLmdDEAvaVgAq PQpRtK

                                   ↓         N
HPII: ELWEAIEAGDFPEYELgfQLIpEEDEfKFDF DLLDpTKLiPEelVPVqRVGKMVLNRNPDNFFAENEQAAFhPGHIV PGL
RLC:  DLFNAIasGNYPsWtfyiQVMtfkEaetFpF NpfDlTKVwPhkdyPLipVGKLVLNRNPaNYFAEvEQmAFdPsnMp PGI
BLC:  DLFNAIatGNYPsWtLyiQVMtfsEaeiFpF NpfDlTKVwPhgdyPLipVGKLVLNRNPvNYFAEvEQlAFdPsnMp PGI
HKC:  DLFNAIatGkYPsWtfyiQVMtfNQaetFpF NpfDlTKVwPhkdyPLipVGKLVLNRNPvNYFAEvEQiAFdPsnMp PGI
SCC:  kLFtqLQnGEkPkFNcyvQtMtpEQatKFrY sVnDlTKIwPhkefPMrkfGtItLteNvDNYFqEiEQvAFsPtntciPhI
CTC:  DLFknIaAGNYPsWtcyiQtMtEaQakeaEF sVfDlTKVwPhgkyPLrRfGKftLNeNPkNYFAEvEQAAFsPAHtV PGM
ZMC:  DLtDAIaAGNFPEWtLyiQtMdpEmEdRlQhlDpLDvTKtwPE dVPLEpVGRLVLNRNiDNFFAENEQlAFsPGlIV PGI

                ↓    ↓
HPII: DFTNDPLLQGRLFSYTDTQisRLGgPNFhEIPINRPt   CPYH                   NFQ RDG MhrMGiDTNpaNYe
RLC:  EpSpDkMLQGRLFaYpDThrhRLG PNYlQIPVN      CPYRa            rvaNYQ RDGpMcmhdnQggapNYy
BLC:  EpSpDkMLQGRLFaYpDThrhRLG PNYlQIPVN      CPYRa            rvaNYQ RDGpMcmMdnQggapNYy
HKC:  EaSpDkMLQGRLFaYpDThrhRLG PNYlhIPVN      CPYRa            rvaNYQ RDGpMcmqdnQggapNYy
SCC:  kpSNDsVLQARLFSYpDTQrhRLG aNYqQLPVNRPrnlgCPYskgdsqytaeqcpfkavNFh RDGpMsyy nfgpepNYi
CTC:  EpSaDPVLQsRLFSYaDThrhRLG tNYtQIPVN      CPvtg            avfNpQmRDGpMnvnGnlgNhpNYl
ZMC:  yYSDDkLLQtRIFSYSDTQrhRLG PNYllLPaN      aPlca            hhnNhy  DGsMnfM hrTEevDYp

HPII: PNSiNDNWpreTpPgPKrggfesyQerVeGNkVrERsPsfgEYyshPRLfwlsQtpFeq rHIVDGFsF   el sKvVRpY
RLC:  PNSfsapeqqgSalehHsqcsadvkrfnsANednvtqvrtfytkvlneeerkrlcenia nHLkDAqlF   iq rKaVKn
BLC:  PNdfsapehqpSalehRthfsgdvQrfnsANddnvtqvrtfylkvlneeqrkrlcenia gHLkDAqlF   iq kKaVKn
HKC:  PNdfgapeqqpSalehsiqysgevrrfntANddnvtqvrafyvnvlneeqrkrlcenia gHLkDAqiF   iq kKaVKn
SCC:  sslpmQtlkfknevndevsdkfkgivldiqteVsvRkqeqdQirnehiVdakiNqyYyvygispldFeqpralyekvyndeq
CTC:  asdkpiEFkqfSlqedqevwhgaatpfhwkEtpaDfkqatelWkvlkKypnqq  ehlahnvaVhAsaadapiqdr  ViaY
ZMC:  PsryDavrnapryPiPtahiagrrEktViskennfKqPgeryramdParqe rfitngs t Leipgyr   ar iRnhgh

HPII: iRE rvvDqLAH  IDl tLaQaVAKnlGIeLtDdqlnITpppdVNglkkDpslslyaipdGdvkgrvvaillndevrsadll
RLC:  ftD vhpDygAR  VQa lLdQynsqkpknaIhtyvqagShiaakgkanl
BLC:  fsD vhpEygsR  IQa lLdkyneekpkn
HKC:  ftE vhpDygsH  IQa lLdkynAekpknaIhtfvqsgShlaarEkanl
SCC:  kKlfvhnvvchackIkdpkVkkrVtqyfGLlneDlgkvIaeglgVpwepvDlegyaktwsiAsa
CTC:  ftk vhpDlgdl  IkkeiLelsprK
ZMC:  scf lcelvL R  sQf htt yepea

HPII: ailkalkakgvhakllysrmgevtaddgtvlpiaatfagapsltvdavivpcgniadiadngdanyylmeaykhlkpialag

HPII: darkfkatikiadqgeegiveadsadgsfmdelltlmaahrvwsripkidkipa
```

FIG. 7. Comparison of the predicted amino acid sequence of the HPII subunit with the sequences of other catalases including rat liver (RLC) (11), bovine liver (BLC) (36), human kidney (HKC) (2), *S. cerevisiae* (SCC) (15), *C. tropicalis* (CTC) (30), and maize (ZMC) (3). Amino acids that are identical in HPII and one or more of the other catalases are capitalized and italicized. Amino acids that are the result of conservative replacements between HPII and any one of the other catalases are capitalized. Conservative amino acid replacements are defined as occurring in the following groups: (S and T), (A and G), (F, Y, and W), (H, K, and R), (D, E, N, and Q), and (I, L, M, and V). Residues predicted to be in the active site of the bovine liver enzyme (9) are indicated by *. Residues predicted to be involved in binding the heme on either the distal (↑) or proximal (↓) sides (9) are also indicated. The residues predicted to be involved in NADP binding are indicated by N.

servative replacements. Shorter segments of the sequences have even greater degrees of similarity, and all of the residues identified as playing a role in the active site of the bovine liver enzyme and in binding the protoheme group have counterparts in the HPII sequence. This similarity is surprising when the differences in physical properties between HPII and other catalases are considered. For example, HPII is active as a hexamer, whereas the normal catalases from eucaryotic and some bacterial sources are active as tetramers. HPII has a predicted subunit size of 84,118 Da, compared with approximately 65,000 Da for other catalases, and it is green rather than brown in color, a result of the heme d-like group replacing the normal protoheme IX. Because of the close similarity between portions of HPII and bovine liver catalase, for which the crystal structure has been determined (29), HPII may be useful as a model to test some of the predictions about the roles of individual amino acids in the active site and in heme binding. The greater length of HPII is contained in two sections, a 57-amino-acid segment at the N terminus and a 168-amino-acid segment at the C terminus. The role of this additional protein may be elucidated when the crystal structure of HPII (39) is completed, but it is likely that it is involved in changing the quaternary structure of the active enzyme from the normal catalase tetramer form to the hexamer form of HPII.

The HPII heme d group differs from protoheme IX in that it has two hydroxyl groups in a *cis* configuration on ring C. The amino acid residues interacting with the heme have not changed, indicating that the diol configuration does not affect these interactions. Little is known about the oxidative conversion of protoheme into the relatively uncommon HPII heme d or the process of inserting the heme into the protein. The cell is capable of making sufficient protoheme IX for the plasmid-encoded synthesis of HPI, and the conversion of protoheme IX to the HPII heme d must be sufficiently facile to allow the rapid synthesis of large amounts of active HPII. Further work is required to determine whether the protoheme-to-heme d conversion occurs prior to insertion or as part of the insertion process.

Binding of NADPH by the bovine liver enzyme involves a water molecule, which is predicted to interact with Lys-236, His-234, and Tyr-214, and His-304, which is predicted to interact with the pyrophosphate portion of NADPH. Lys-236 and His-234 have Lys-294 and His-292 as counterparts in HPII, but the binding of the water molecule would be adversely affected by the replacement of Tyr-214 with Phe-272 in HPII, which lacks the necessary phenolic OH group for interaction with water. Furthermore, His-304 in the bovine enzyme is replaced by the negatively charged Glu-362 in HPII, which should not interact favorably with the negatively charged pyrophosphate group of NADPH. These differences suggest that HPII does not bind NADPH, but this remains to be confirmed.

The sequence similarity in the −10 and −35 regions upstream from the transcription start sites in *katE* and *xthA* is consistent with the observation that transcription of both genes is controlled by the KatF protein (32), a putative σ factor (27). Other genes controlled by the KatF protein have not yet been identified, but it is unlikely that a σ factor would have evolved for the expression of just two genes, neither of which is absolutely essential for viability. A number of genes, including *appA* (encoding an acid phosphatase [40]), the genes for glycogen accumulation (31), and other less-well-defined genes (37, 43) have been identified in *E. coli* and *Salmonella typhimurium* as responding to starvation conditions. The expression of *katE*, which is turned on as the cells enter stationary phase (19), brought on by starvation for a medium component may therefore be part of a larger strategy for surviving starvation conditions controlled at least in part by *katF*.

ACKNOWLEDGMENTS

This work was supported by grant OGP0009600 (to P.C.L.), a postgraduate scholarship (to I.V.), and an undergraduate scholarship (to A.B.), all from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1989. Current protocols in molecular biology, vol. 1. Wiley Interscience, New York.
2. **Bell, G. I., R. C. Najarian, G. T. Mullenbach, and R. A. Hallewell.** 1986. cDNA sequence coding for human kidney catalase. Nucleic Acids Res. **14**:5561–5562.
3. **Bethards, L. A., R. W. Skadsen, and J. G. Scandalios.** 1987. Isolation and characterization of a cDNA clone for the *Cat2* gene in maize and its homology with other catalases. Proc. Natl. Acad. Sci. USA **84**:6830–6834.
4. **Boulnois, G. J., and K. N. Timmis.** 1984. Advanced molecular genetics. A. Puhler and K. Timmis (ed.) Springer-Verlag KG, Berlin.
5. **Chiu, J. T., P. C. Loewen, J. Switala, R. B Gennis, and R. Timkovich.** 1989. Proposed structure for the prosthetic group of the catalase HPII from *Escherichia coli*. J. Am. Chem. Soc. **111**:7046–7050.
6. **Claiborne, A., and I. Fridovich.** 1979. Purification of the *o*-dianisidine peroxidase from *Escherichia coli* B. J. Biol. Chem. **254**:4245–4252.
7. **Claiborne, A., D. P. Malinowski, and I. Fridovich.** 1979. Purification and characterization of hydroperoxidase II of *Escherichia coli* B. J. Biol. Chem. **254**:11664–11668.
8. **Edelhoch, H.** 1967. Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry **6**:1948–1954.
9. **Fita, I., and M. G. Rossmann.** 1985. The active center of catalase. J. Mol. Biol. **185**:21–37.
10. **Fita, I., and M. G. Rossmann.** 1985. The NADPH binding site on beef liver catalase. Proc. Natl. Acad. Sci. USA **82**:1604–1608.
11. **Furuta, S., H. Hayashi, M. Hijikata, S. Miyazawa, T. Osumi, and T. Hashimoto.** 1986. Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver catalase. Proc. Natl. Acad. Sci. USA **83**:313–317.
12. **Gilman, M. Z., and M. Chamberlin.** 1983. Developmental and genetic regulation of *Bacillus subtilis* genes translated by σ^{28}-

RNA polymerase. Cell **35**:285–293.
13. **Grosjean, H., and W. Fiers.** 1982. Preferential codon usage in prokaryotic genes: the optimal codon-anticodon interaction energy and the selective codon usage in efficiently expressed genes. Gene **18**:199–209.
14. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of the *E. coli* promoter sequences. Nucleic Acids Res. **15**:2343–2361.
15. **Hartig, A., and H. Ruis.** 1986. Nucleotide sequence of the *Saccharomyces cerevisiae CTT1* gene and deduced amino-acid sequence of yeast catalase T. Eur. J. Biochem. **160**:487–490.
16. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
17. **Loewen, P. C.** 1984. Isolation of catalase-deficient *Escherichia coli* mutants and genetic mapping of *katE*, a locus that affects catalase activity. J. Bacteriol. **157**:622–626.
18. **Loewen, P. C., and J. Switala.** 1986. Purification and characterization of catalase HPII from *Escherichia coli* K12. J. Biol. Chem. **64**:638–646.
19. **Loewen, P. C., J. Switala, and B. L. Triggs-Raine.** 1985. Catalases HPI and HPII in *Escherichia coli* are induced independently. Arch. Biochem. Biophys. **243**:144–149.
20. **Loewen, P. C., and B. L. Triggs.** 1984. Genetic mapping of *katF*, a locus that with *katE* affects the synthesis of a second catalase species in *Escherichia coli*. J. Bacteriol. **160**:668–675.
21. **Loewen, P. C., B. L. Triggs, C. S. George, and B. Hrabarchuk.** 1985. Genetic mapping of *katG*, a locus that affects synthesis of the bifunctional catalase-peroxidase hydroperoxidase I in *Escherichia coli*. J. Bacteriol. **162**:661–667.
22. **Loewen, P. C., B. L. Triggs, G. R. Klassen, and J. H. Weiner.** 1983. Identification and physical characterization of a ColE1 hybrid plasmid containing a catalase gene of *Escherichia coli*. Can. J. Biochem. Cell Biol. **61**:1315–1321.
23. **Loprasert, S., S. Negoro, and H. Okada.** 1989. Cloning, nucleotide sequence, and expression in *Escherichia coli* of the *Bacillus stearothermophilus* peroxidase gene (*perA*). J. Bacteriol. **171**:4871–4875.
24. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Mead, D. A., E. S. Skorupa, and B. Kemper.** 1985. Single-stranded DNA SP6 promoter plasmids for engineering mutant RNAs and proteins: synthesis of a 'stretched' preparathyroid hormone. Nucleic Acids Res. **13**:1103–1118.
26. **Miller, J. H.** 1974. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Mulvey, M. R., and P. C. Loewen.** 1989. Nucleotide sequence of *katF* of *Escherichia coli* suggests KatF protein is a novel σ transcription factor. Nucleic Acids Res. **17**:9979–9991.
28. **Mulvey, M. R., P. A. Sorby, B. L. Triggs-Raine, and P. C. Loewen.** 1988. Cloning and physical characterization of *katE* and *katF* required for catalase HPI expression in *Escherichia coli*. Gene **73**:337–345.
29. **Murthy, M. R. N., T. J. Reid III, A. Sicignano, N. Tanaka, and M. G. Rossman.** 1981. Structure of beef liver catalase. J. Mol. Biol. **152**:465–499.
30. **Okada, H., M. Ueda, T. Sugaya, H. Atomi, S. Mozaffar, T. Hishida, Y. Teranishi, K. Okazaki, T. Takechi, T. Kamiryo, and A. Tanaka.** 1987. Catalase gene of the yeast *Candida tropicalis*. Sequence analysis and comparison with peroxisomal and cytosolic catalases from other sources. Eur. J. Biochem. **170**:105–110.
31. **Preiss, J.** 1984. Bacterial glycogen synthesis and its regulation. Annu. Rev. Microbiol. **38**:419–458.
32. **Sak, B. D., A. Eisenstark, and D. Touati.** 1989. Proc. Natl. Acad. Sci. USA **86**:3271–3275.
33. **Sammartano, L. J., R. W. Tuveson, and R. Davenport.** 1986. Control of sensitivity to inactivation by H_2O_2 and broad-spectrum near-UV radiation by the *Escherichia coli katF* locus. J. Bacteriol. **168**:13–21.
34. **Sanger, F. S., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
35. **Saporito, S. M., B. J. Smith-White, and R. P. Cunningham.** 1988. Nucleotide sequence of the *xthA* gene of *Escherichia coli* K-12. J. Bacteriol. **170**:4542–4547.
36. **Schroeder, W. A., J. R. Shelton, J. B. Shelton, B. Robberson, G. Apell, R. S. Fang, and J. B. Ventura.** 1982. The complete amino acid sequence of bovine liver catalase and the partial sequence of bovine erythrocyte catalase. Arch. Biochem. Biophys. **214**:397–421.
37. **Spector, M. P., Y. K. Park, S. Tirgari, T. Gonzalez, and J. W. Foster.** 1988. Identification and characterization of starvation-regulated genetic loci in *Salmonella typhimurium* by using Mu d-directed *lacZ* operon fusions. J. Bacteriol. **170**:345–351.
38. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074–1078.
39. **Tormo, J., I. Fita, J. Switala, and P. C. Loewen.** 1990. Crystallization and preliminary X-ray diffraction analysis of catalase HPII from *Escherichia coli*. J. Mol. Biol. **213**:219–220.
40. **Touati, E., and A. Danchin.** 1987. Cloning and characterization of the pH 2.5 acid phosphatase gene, *appA*: cyclic AMP mediated negative regulation. Mol. Gen. Genet. **208**:499–505.
41. **Triggs-Raine, B. L., B. W. Doble, M. R. Mulvey, P. A. Sorby, and P. C. Loewen.** 1988. Nucleotide sequence of *katG*, encoding catalase HPI of *Escherichia coli*. J. Bacteriol. **170**:4415–4419.
42. **Triggs-Raine, B. L., and P. C. Loewen.** 1987. Physical characterization of *katG*, encoding catalase HPI of *Escherichia coli*. Gene **52**:121–128.
43. **Wanner, B. L., and R. McSharry.** 1982. Phosphate-controlled gene expression in *Escherichia coli* using Mu dl-directed *lacZ* fusions. J. Mol. Biol. **158**:347–363.
44. **Weber, K., and M. Osborn.** 1969. The reliability of molecular weight determinations by dodecyl sulfate polyacrylamide gel electrophoresis. J. Biol. Chem. **244**:4406–4412.
45. **Yanisch-Perron, C., C. J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.