Cloning and Expression of a cDNA Coding for Catalase from Zebrafish (*Danio rerio*)

Chuian-Fu Ken,[†,‡] Chi-Tsai Lin,[‡] Jen-Leih Wu,[‡,§] and Jei-Fu Shaw*[,‡,||]

Graduate Institute of Life Sciences, National Defense Medical Center, Taipei, Taiwan, Institute of Marine Biotechnology, National Taiwan Ocean University, Keelung, Taiwan, and Institutes of Zoology and Botany, Academia Sinica, Nankang, Taipei, Taiwan

A full-length complementary DNA (cDNA) clone encoding a catalase was amplified by the rapid amplication of cDNA ends−polymerase chain reaction (RACE−PCR) technique from zebrafish (*Danio rerio*) mRNA. Nucleotide sequence analysis of this cDNA clone revealed that it comprised a complete open reading frame coding for 526 amino acid residues and that it had a molecular mass of 59 654 Da. The deduced amino acid sequence showed high similarity with the sequences of catalase from swine (86.9%), mouse (85.8%), rat (85%), human (83.7%), fruit fly (75.6%), nematode (71.1%), and yeast (58.6%). The amino acid residues for secondary structures are apparently conserved as they are present in other mammal species. Furthermore, the coding region of zebrafish catalase was introduced into an expression vector, pET-20b(+), and transformed into *Escherichia coli* expression host BL21(DE3)pLysS. A 60-kDa active catalase protein was expressed and detected by Coomassie blue staining as well as activity staining on polyacrylamide gel followed electrophoresis.

Keywords: *Zebrafish; Danio rerio; molecular cloning; cDNA; catalase; expression; Escherichia coli; RACE−PCR; pET-20b(+)*

INTRODUCTION

Catalase (H_2O_2:H_2O_2 oxidoreductase, E.C. 1.11.6) is one of the important antioxidant enzymes that catalyze the decomposition of hydrogen peroxide to oxygen and water. It is found in all aerobic organisms, and it protects cells against the toxic effects of hydrogen peroxide. There are 74 catalase sequences which have been reported, including 29 bacterial, 8 fungal, 30 plant, but only 7 animal sequences (Klotz et al., 1997). The reported animal catalase sequences include those of the nematode (Henkle-Duehrsen, 1994, unpublished, accession number X82175), fruit fry (Orr et al., 1990), bovine (Schroeder et al., 1982), rat (Furuta et al., 1986), mouse (Shaffer et al., 1990), swine (Lin et al., 1997), and human (Quan et al., 1986). Most mammalian catalase is located within the peroxisome and forms a homotetramer structure; each subunit contains a single ferriprotoporphorin IX group, and the molecular mass is about 60 kDa (Deisseroth and Dounce, 1970). There is no report on the cloning and characterization of cDNA for fish catalase.

In fish, environmental pollution may enhance oxidative stress and thus disturb the natural antioxidant enzyme system (Mather-Mihaich and Diguilio, 1986; Radi and Marcovics, 1988; Nicholls et al., 1989; St. Dimitrova et al., 1994). Several comparative studies have demonstrated the basal antioxidant enzyme activities in various fish species (Aksnes and Njaa, 1981; Wdzieczak et al., 1982; Witas et al., 1984). Recently, antioxidant enzymes including catalase were demonstrated to be useful biomarkers of environmental pollutants which cause oxidative stress in fish (Pedrajas et al., 1995; Roche and Boge, 1996).

The maximum life span of a fly can be extended by increasing both catalase and superoxide dismutase (SOD) activities using a transgenic technique (Orr and Sohal, 1994). These results confirm the free radical theory (Harman, 1956) that completely removed reactive oxygen species (ROS) can reduce damage and prolong the life span in the fly. Zebrafish is a model fish for testing if the same theory is applicable to teleosts, primitive vertebrates. The first step is to clone and characterize the cDNAs of both SOD and catalase from zebrafish.

Previously, we have cloned a Cu/Zn-SOD cDNA from zebrafish (Ken et al., 1998). Here we report the cloning of the full coding region of catalase cDNA with its 3′-untranslated region. In addition, the coding region of catalase cDNA from zebrafish was introduced into an expression vector, pET-20b(+), and transformed into *Escherichia coli* BL21(DE3)pLysS. An active recombinant catalase was detected.

MATERIALS AND METHODS

Zebrafish (*Danio rerio*) were raised in freshwater for 2 weeks before use in cloning experiments.

Preparation of mRNA and cDNA Synthesis. Total mRNA and poly(A)$^+$ RNA were extracted as described in our previous work (Ken et al., 1998). Double-strand blunted cDNA was synthesized using a kit (cDNA synthesis module RNP 1256) from Amersham (Little Chalfont, Buckinghamshire, U.K.).

Subcloning and DNA Sequence Analysis. One microgram of blunted cDNA was ligated with 30 pmol of Marathon

* Address correspondence to this author at the Institute of Botany, Academia Sinica. Telephone: 886-2-27821258 (Ext. 108). Fax: 886-2-27821605. E-mail: boplshaw@gate.sinica.edu.tw.

[†] National Defense Medical Center.

[‡] National Taiwan Ocean University.

[§] Institute of Zoology, Academia Sinica.

[||] Institute of Botany, Academia Sinica.

10.1021/jf990838+ CCC: $19.00
Published on Web 05/20/2000

Table 1. Primer Sequences Used for Cloning and Sequencing the Catalase Gene of Zebrafish (*D. rerio*)[a]

CA-1	5′-GC(ATC)TTCGG(ATC)TACTTCGAGGT(GTC)AC-3′
CA-2	5′-AG(ATG)GGGTAGTC(TC)TTGTG(ATG)GGCCA-3′
CA-5	5′-TGTACTGCAAGTTCCACTACAAGAC-3′
CA-6	5′-CCTGGATGTAGAAGGTCCAGGATGG-3′
CA-7	5′-GGAATCCATTTGATTTGACCAAGG-3′
CA-8	5′-CGGTGTCGTCTTTCCAACATGCTC-3′
CA-9	5′-CTACCCCAACAGCTTCAGTGCTCC-3′
CA-10	5′-CCGCTCTCGGTCAAAATGGGCC-3′
5′ up-stream primer	5′-CCGGAATTCGATGGCAGACGACAGAG-3′
3′ down-stream primer	5′-CCGCTCGAGCATCTTAGAAGCTGCAGCCACAG-3′
AP-1	5′-CCATCCTAATACGACTCACTATAGGGC-3′
M13 forward primer[b]	5′-GTAAAACGACGGCCAG-3′
M13 reverse primer[b]	5′-CAGGAAACAGCTATGAC-3′
T7 promoter primer[b]	5′-TTAATACGACTCACTATAGGG-3′
T7 terminator primer[b]	5′-CAATAACGAGTCGCCACCG-3′

[a] CA-1 and CA-2 are degenerate primers. AP-1 refers to the commercial Marathon adaptor 1 primer (Clontech, Palo Alto, CA). [b] Primers used for sequencing only.

cDNA adaptor (Clontech, Palo Alto, CA) at 4 °C for 16 h. When compared to the rat (Furuta et al., 1986), mouse (Shaffer et al., 1990), and human (Quan et al., 1986) catalase sequences with the Genetics Computer Group (GCG) program (Wisconsin Package Version 10.0, Madison, WI), two conserved sequences (81-AFGYFEVT-88, 303-WPHGDYPL-310) were selected for synthesis of the degenerate primers, CA-1 and CA-2. Using 0.05 μg of the ligated cDNA as a template, 2.5 units of AmpliTaq Gold DNA polymerase (Perkin-Elmer Applied Biosystems, Norwalk, CT) and 10 pmol of each of the two degenerate primers were added. One 0.7-kb cDNA was amplified by the polymerase chain reaction (PCR) technique (25 cycles of 94 °C for 1 min, 45 °C for 1 min, and 72 °C for 1 min). The 0.7-kb cDNA was subcloned into the pGEM-T cloning vector (Promega, Madison, WI) using *E. coli* JM109 as the host. The nucleotide sequence was determined in both directions by using an ABI PRISM 377-96 DNA autosequencer (ABI PRISM bigdye terminator cycle sequencing ready reaction kit, Perkin-Elmer Applied Biosystems). According to the 0.7-kb cDNA sequence, we synthesized a CA-7 primer (5′ GGA ATC CAT TTG ATT TGA CCA AGG 3′) for 3′-end extension and a CA-8 primer (5′ CCT GGA TGT AGA AGG TCC AGG ATG G 3′) for 5′-end extension. Using 0.05 μg of the ligated cDNA as a template, 10 pmol of Clontech adaptor primer 1 (AP1) and 10 pmol of CA-7 primer were added to perform PCR (25 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min). Again using 0.05 μg of the ligated cDNA as a template, 10 pmol of AP1 primer and 10 pmol of CA-8 primer were added to perform PCR (25 cycles of 94 °C for 1 min, 50 °C for 1 min, and 72 °C for 1 min). One 1.3-kb cDNA (3′ RACE−PCR product) and one 0.7-kb cDNA (5′ RACE−PCR product) were amplified. Both 1.3-kb and 0.7-kb PCR fragments were subcloned into the pCR2.1-TOPO TA cloning vector (Invitrogen, Carlsbad, CA). Sequence analysis revealed that the 1.3-kb 3′ RACE−PCR product contained a catalase C-terminal peptide sequence and its 3′ untranslated region. The 0.7-kb 5′ RACE−PCR product contained a partial catalase N-terminal peptide sequence but lacked the translation initiation codon; we therefore synthesized a CA-10 primer (5′ CCG CTC TCG GTC AAA ATG GGC C 3′) for 5′-end cloning. Using 0.05 μg of the ligated cDNA as a template, 10 pmol of AP1 primer and 10 pmol of CA-10 primer were added to perform PCR (25 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s). One 0.3-kb PCR product contained the complete catalase N-terminal sequence, and its 5′ untranslated region was amplified. Using these 5′ and 3′ RACE−PCR products as templates, one 2.3-kb fused full-length cDNA of catalase was created by the RACE−PCR technique (EMBL accession number AJ007505). All sequences of primers used are shown in Table 1.

Recombinant DNA Preparation and Transformation. According to the 2.3-kb catalase cDNA sequence, a 5′ upstream primer (5′ CCG GAA TTC GAT GGC AGA CGA CAG AG 3′) and a 3′ downstream primer (5′ CCG CTC GAG CAT CTT AGA AGC TGC AGC CAC AG 3′) were synthesized. Using 0.05 μg of zebrafish-blunted cDNA as a template, 10 pmol each of the 5′ upstream and 3′ downstream primers was added; a 1.6-kb fragment containing the full coding region of catalase was amplified by PCR, ligated with pCR 2.1, and then transformed into the *E. coli* TOPO 10 host. A positive clone was selected by hybridization with ^{32}P-labeled catalase cDNA as a probe, and plasmid DNA was prepared. A suitable amount of plasmid DNA was digested with *Eco*RI and *Xho*I and then run on 0.8% agarose gel. A 1.6-kb insert DNA containing the *Eco*RI and *Xho*I sites was recovered and ligated with the pET-20b(+) expression vector that was pretreated with *Eco*RI and *Xho*I. Then the ligated DNA was transformed into the *E. coli* expression host, BL21(DE3)pLysS (Novagen, Madison, WI). A transformed clone was selected by hybridization with ^{32}P-labeled catalase cDNA as a probe.

Culture and Enzyme Extraction. The transformed *E. coli* cells were grown at 37 °C in 3.0 mL of Luria−Bertani medium containing 50 μg/mL ampicillin until A_{600} reached 0.9. Isopropyl β-D-thiogalactopyranoside (IPTG) was added to a concentration of 1 mM. The culture was incubated at 37 °C for 4 h by shaking at 120 rpm, and the bacterial cells were then harvested by centrifugation at 12000*g* for 5 min. The cells were suspended in 0.1 mL of 50 mM phosphate buffer (pH 7.0) containing 0.1% glycerol and 0.01 g of glass beads and then vortexed for 5 min and centrifuged at 9000*g* for 5 min. The enzyme extract contained active catalase.

Enzyme Activity Assay by Spectrophotometer and SDS−PAGE Analysis. Catalase activity was measured by the decrease in H_2O_2 concentration at room temperature. The decomposition of H_2O_2 was recorded with absorbance at 240 nm at time intervals of 3 min in a cuvette containing 50 mM phosphate buffer (pH 7.0) and 10 mM H_2O_2. The H_2O_2 concentration was determined using $E_{240} = 40\ M^{-1}\ cm^{-1}$ (Aebi, 1984), where one unit of enzyme activity was defined as the amount of enzyme decomposed by 1 μmol of H_2O_2 per minute. Catalase activity staining assay was performed as described by Wayne and Diaz (1986).

A 15-μL aliquot containing 7.8 μg of protein of the extracted enzyme with the sample buffer was heated for 5 min and then subjected to electrophoresis on a 10% polyacrylamide gel with SDS (Laemmli, 1970).

Protein Concentration Measurement. Protein concentration was determined by a kit (Bio-Rad Protein Assay) from BIO-RAD (Hercules, CA).

RESULTS AND DISCUSSION

The complete nucleotide sequence of zebrafish catalase cDNA is shown in Figure 1. The deduced amino acid sequence of catalase comprises a complete open reading frame coding for 526 amino acid residues from the first ATG codon to the TGA stop codon. The 5′ untranslated region consists of 98 nucleotides. There is a TGA stop codon (−6 to −4 bp) located before the first ATG start codon (+1 to +3) that indicates the start site for amino acid synthesis. The 3′ untranslated region contains 592 nucleotides. A consensus polyadenylylation

```
-98                                                        CCTTAATAACCAACAACCCTCCAGACAGAC
-68      CCCGAGTCCAGTGGTTTGGATCCGCAGTAGTTTTGCCTCGTGTTTGTCACACTGAAGAGCTGAATA
+1    ATG GCA GAC GAC AGA GAA AAG TCG ACG GAT CAG ATG AAA CTG TGG AAG GAG GGT
  1    M   A   D   D   R   E   K   S   T   D   Q   M   K   L   W   K   E   G
 55   CGC GGT TCC CAG CGT CCG GAT GTC CTG ACC ACT GGA GCC GGA GTC CCG ATA GGG
 19    R   G   S   Q   R   P   D   V   L   T   T   G   A   G   V   P   I   G
109   GAC AAG TTA AAT GCG ATG ACG GCA GGG CCT CGT GGG CCA TTG CTG GTC CAG GAT
 37    D   K   L   N   A   M   T   A   G   P   R   G   P   L   L   V   Q   D
163   GTG GTG TTT ACT GAT GAA ATG GCC CAT TTT GAC CGA GAG CGG ATA CCA GAG AGA
 55    V   V   F   T   D   E   M   A   H   F   D   R   E   R   I   P   E   R
217   GTC GTG CAT GCT AAA GGA GCA GGA GCG TTT GGC TAC TTC GAA GTC ACT CAC GAC
 73    V   V   H   A   K   G   A   G   A   F   G   Y   F   E   V   T   H   D
271   ATC ACG CGC TAC TCC AAA GCC AAA GTG TTT GAG CAT GTT GGA AAG ACG ACA CCG
 91    I   T   R   Y   S   K   A   K   V   F   E   H   V   G   K   T   T   P
325   ATC GTT GTC CGC TTT TCC ACA GTG GCG GGT GAG GCT GGG TCN CCA GAT ACT GTC
109   I   V   V   R   F   S   T   V   A   G   E   A   G   S   P   D   T   V
379   CGA GAT CCT CGT GGT TTC GCA GTG AAG TTC TAC ACC GAT GAG GGC AAC TGG GAT
127   R   D   P   R   G   F   A   V   K   F   Y   T   D   E   G   N   W   D
433   CTT ACA GGA AAC AAC ACC CCC ATC TTC TTT ATC AGG GAT ACG CTT CTG TTT CCG
145   L   T   G   N   N   T   P   I   F   F   I   R   D   T   L   L   F   P
487   TCT TTC ATC CAC TCT CAG AAG CGC AAT CCG CAG ACT CAC CTG AAG GAT CCG GAC
163   S   F   I   H   S   Q   K   R   N   P   Q   T   H   L   K   D   P   D
541   ATG GTT TGG GAT TTT TGG AGC TTG CGT CCT GAA TCG CTG CAC CAG GTG TCT TTC
181   M   V   W   D   F   W   S   L   R   P   E   S   L   H   Q   V   S   F
595   CTG TTC AGT GAT CGA GGG ATT CCC GAT GGC TAC CGT CAT ATG AAC GGA TAC GGA
199   L   F   S   D   R   G   I   P   D   G   Y   R   H   M   N   G   Y   G
649   TCG CAC ACC TTT AAA CTG GTC AAT GCT CAG GGC CAG CCG GTG TAC TGC AAG TTC
217   S   H   T   F   K   L   V   N   A   Q   G   Q   P   V   Y   C   K   F
703   CAC TAC AAG ACT AAT CAG GGC ATT AAG AAT ATT CCT GTT GAA GAA GCG GAT CGT
235   H   Y   K   T   N   Q   G   I   K   N   I   P   V   E   E   A   D   R
757   CTG GCT GCC ACT GAT CCG GAT TAC TCT ATC AGA GAC CTT TAC AAC GCT ATC GCC
253   L   A   A   T   D   P   D   Y   S   I   R   D   L   Y   N   A   I   A
811   AAT GGC AAC TTC CCA TCC TGG ACC TTC TAC ATC CAG GTT ATG ACC TTT GAG CAG
271   N   G   N   F   P   S   W   T   F   Y   I   Q   V   M   T   F   E   Q
865   GCT GAA AAC TGG AAG TGG AAT CCA TTT GAT TTG ACC AAG GTC TGG TCC CAT AAA
289   A   E   N   W   K   W   N   P   F   D   L   T   K   V   W   S   H   K
919   GAG TTC CCT CTG ATT CCT GTA GGG CGT TTT GTG TTG AAC CGA AAC CCC GTT AAC
307   E   F   P   L   I   P   V   G   R   F   V   L   N   R   N   P   V   N
973   TAT TTC GCG GAG GTT GAG CAG CTG GCG TTT GAT CCC AGT AAC ATG CCA CCC GGC
325   Y   F   A   E   V   E   Q   L   A   F   D   P   S   N   M   P   P   G
1027  ATT GAG CCC AGC CCA GAC AAG ATG CTG CAG GGG CGT CTT TTC TCC TAC CCA GAC
343   I   E   P   S   P   D   K   M   L   Q   G   R   L   F   S   Y   P   D
1081  ACA CAT CGC CAT CGA CTC GGA GCT AAT TAC CTC CAA CTA CCA GTC AAC TGC CCG
361   T   H   R   H   R   L   G   A   N   Y   L   Q   L   P   V   N   C   P
1135  TAC CGC ACC CGT GTG GCA AAC TAT CAG AGA GAT GGA CCC ATG TGC ATG CAT GAC
379   Y   R   T   R   V   A   N   Y   Q   R   D   G   P   M   C   M   H   D
1189  AAC CAG GGT GGA GCT CCA AAC TAC TAC CCC AAC AGC TTC AGT GCT CCT GAC GTC
397   N   Q   G   G   A   P   N   Y   Y   P   N   S   F   S   A   P   D   V
1243  CAG CCA CGC TTC CTT GAG TCA AAG TGT AAA GTG TCT CCT GAT GTG GCC CGA TAC
415   Q   P   R   F   L   E   S   K   C   K   V   S   P   D   V   A   R   Y
1297  AAC AGT GCA GAC GAT GAC AAC GTG ACC CAG GTG CGC ACG TTC TTC ACT CAG GTG
433   N   S   A   D   D   D   N   V   T   Q   V   R   T   F   F   T   Q   V
1351  CTA AAT GAA GCC GAG AGA GAG CGT CTG TGC CAA AAT ATG GCC GGG CAT CTG AAA
451   L   N   E   A   E   R   E   R   L   C   Q   N   M   A   G   H   L   K
1405  GGA GCT CAA CTC TTC ATC CAG AAA CGC ATG GTG CAA AAC CTG ATG GCT GTT CAC
469   G   A   Q   L   F   I   Q   K   R   M   V   Q   N   L   M   A   V   H
1459  TCT GAT TAT GGC AAC CGA GTT CAG GCT CTC CTG GAT AAA CAC AAT GCT GAA GGG
487   S   D   Y   G   N   R   V   Q   A   L   L   D   K   H   N   A   E   G
1513  AAA AAG AAC ACT GTT CAT GTT TAT TCA CGT GGT GGA GCG TCT GCT GTG GCT GCA
505   K   K   N   T   V   H   V   Y   S   R   G   G   A   S   A   V   A   A
1567  GCT TCT AAG ATG TGA TTACAAAGTGACGTCCTACTACACACACACACACACACACACACACACA
523   A   S   K   M Stop
1631  CACACACACACAGATTCATTCTCTTATAGAGTGTGATATAGTGTTATAATGCCTAATTCATTCCATAT
1699  ATGGTTTAGAAATACAGACATATCTAATATGGCCAATTCACAACACAAGTGATCTTAGCAAATGCAAC
1767  ACTGACTCCTTAATCTTGATTTATGCCTTCAAGCAAAATGGGGGCCTTTGCATACAAAGATTTGAGCT
1835  AATCCAAACTAAATCAGCCTATTTCAGTATCTTTCTGCATTTTGATTTCTGTGAAATGTAGAATAGCC
1903  AATAATAATCTGAAAAGCAAAGCAACTGTTTAGTTTAATGCTGTATGACACTTGATTTTTTTTTTTTC
1971  TCCACAATTTCTAAAAACCTTGGCTACTTATGCAGTTTTAAGCTGAATTAAGATGGAATGAATGTTT
2039  GCCGTCCTTTCTGCCTTGACTGAATATAATTCGGTTCTCATAATATGAATAACTCGTTTTGTCATTTT
2107  GTAATTAATAAATGATAATATTTAAAAAAAAAAAAAAAGAAAAAAAAAAAGCGGCCGCTGAATTCT
2174
```

Figure 1. Complete nucleotide sequence of zebrafish catalase cDNA and the deduced amino acid sequence. Numbers to the left refer to the catalase cDNA nucleotide sequence and its deduced amino acid residues. A stop codon before the first methionine is shown by an open box. The consensus polyadenylation signal and polyadenylation sequence are underlined. The Stop denotes the stop signal.

signal, AATAAA (2113−2118 bp), is located 11 nucleotides upstream of the poly(A) tail. Long repeats of the AC unit are found between nucleotides 1603 and 1642, and this region was identified as the binding site for redox-sensitive catalase RNA-binding protein (CAT−BP) of rat lung extract (Clerch, 1995).

Figure 2 shows the multiple sequence alignment involving six different species in GenBank. The amino acid side chains that interact with the heme molecule are conserved among the catalases from the species shown in Figure 2 (Murthy et al., 1981). The secondary structures of the α-helixes (α) and *β*-sheets (*β*) with high degrees of conservation across species are underlined on the basis of mouse catalase cDNA (Reimer et al., 1994), except the ninth *β* region (*β*9, between amino acids 414 and 420).The motif analysis was performed

```
                1                  α1                                      50
Zebrafish       MADDREKSTD QMKLWKEGRG SQRPDVLTTG AGVPIGDKLN AMTAGPRGPL
Mouse           .S.S.DPAS. ...Q...Q.A ......... G.N....... I....S....
Rat             ...S.DPAS. ...Q...Q.A P.K....... G.N....... I.........
Human           ...S.DPAS. ..QH...Q.A A.KA...... ..N.V..... VI.V......
Fly             ~MAG.DAASN .LIDY.NSQT VS~.GAI... N.A...I.DA SQ.V.....I
C.elegant       ~~MPNDP.DN .L.TY..~~T YPK.Q.I..S N.A..YS.TA VL...R...M
               51                  α2                  β1                 100
Zebrafish       LVQDVVFTDE MAHFDRERIP ERVVHAKGAG AFGYFEVTHD ITRYSKAKVF
Mouse           .......... .......... .......... .......... ..........
Rat             .......... .......... .......... .......... ..........
Human           .......... .......... .......... .......... ..K.......
Fly             .L...N.L.. .S........ .......... .......... ..Q.CA..I.
C.elegant       .M....YM.. .......... .......... .H........ .SK.C..DI.
              101                 β2                  β3                 β4 150
Zebrafish       EHVGKTTPIV VRFSTVAGEA GSPDTVRDPR GFAVKFYTDE GNWDLTGNNT
Mouse           ..I..R...A .........S ..A..... ........ED .....V....
Rat             ..I..R...A .........S ..A..... ........ED .....V....
Human           ..I..K...A .........S ..A..... ........ED .....V....
Fly             DK.K.R..LA ......G..S ..A..A.... ........ED .V...V....
C.elegant       NK...Q..LL I....G..S ..A..A.... ...I....E. .....V....
              151                 α3                  α4                 α5 200
Zebrafish       PIFFIRDTLL FPSFIHSQKR NPQTHLKDPD MVWDFWSLRP ESLHQVSFLF
Mouse           .......AI. .......... .......... ........C. ......T...
Rat             .......AM. .......... .......... ........C. ......T...
Human           .......PI. .......... .......... .......... ..........
Fly             .V.....PI. ......T... .......... .F...LT... ..A...CI..
C.elegant       .......PIH S.N...T... .........N .IF....H.. .A....M...
              201                 β5                  β6                 250
Zebrafish       SDRGIPDGYR HMNGYGSHTF KLVNAQGQPV YCKFHYKTNQ GIKNIPVEEA
Mouse           ........H. .......... .....D.EA. ........D. ....L..G..
Rat             ........H. .......... .....N.EA. ........D. ....L.....
Human           ........H. .......... .....N.EA. ........D. ....LS..D.
Fly             ....T....C .......... ..I..K.E.I .A...F..D. ....LD.KT.
C.elegant       ....L..... .......... .M..KD.KAI .V...F.PT. .V..LT..K.
              251                 β7                                      300
Zebrafish       DRLAATDPDY SIRDLYNAIA NGNFPSWTFY IQVMTFEQAE NWKWNPFDLT
Mouse           G...QE.... GL...F.... ...Y...... ......KE.. TFPF......
Rat             G...QE.... GL...F.... S..Y...... ......KE.. TFPF......
Human           A..SQE.... G....F.... T.KY...... ......N... TFPF......
Fly             .Q..S..... .......R.K TCK.....M. .....Y...K KF.Y....V.
C.elegant       GQ..SS.... .....F...E K.D..V.KMF .......... K.EF....V.
              301                 β8                  α8                 350
Zebrafish       KVWSHKEFPL IPVGRFVLNR NPVNYFAEVE QLAFDPSNMP PGIEPSPDKM
Mouse           ...P..DY.. ....KL...K .......... .M........ ..........
Rat             ...P..DY.. ....KL.... ..A....... .M........ ..........
Human           ...P..DY.. ....KL.... .......... .I........ ....A.....
Fly             ....Q..Y.. ....KM..D. ..K....... .I..S.AHLV ..V.......
C.elegant       ...P.GDY.. .E..KM.... ..R....... .S..C.AHIV ....F.....
              351                 α9                                      400
Zebrafish       LQGRLFSYPD THRHRLGANY LQLPVNCPYR TRVANYQRDG PMCMHDNQGG
Mouse           ......A... .......P.. ..I..... A......... ..........
Rat             ......A... .......P.. ..I..... A......... ..........
Human           ......A... .......P.. .HI..... A......... ....Q.....
Fly             .H......S. .......P.. ..I......K VKIE.F.... A.NVT...D.
C.elegant       ....I...T. ..F....P.. I....... S.AH.T.... A.A~Y...QH
              401                 β9                  α10 447
Zebrafish       APNYYPNSFS APDV~~QPRF LESKCKVSPD VARYNSAD~D DNVTQVRTFF
Mouse           .......... ..EQ~~.RSA ..HSVQCAV. .K.F...N~E .........Y
Rat             .......... ..EQ~~.GSA ..HHSQC.A. .K.F...N~E .........Y
Human           .........G ..EQ~~..SA ..HSIQY.GE .R.F.T.N~. ........A.Y
Fly             ....F....N G.QECPRA.A .S.C.P.TG. .Y..S.G.TE ..FG..TD.W
C.elegant       ...FF....N YGK~~TR.DV KDTTFPATG. .D..E.G.~. N.YD.P.Q.W
              448                 α11                 α12                 α13  497
Zebrafish       TQVLNEAERE RLCQNMAGHL KGAQLFIQKR MVQNLMAVHS DYGNRVQALL
Mouse           .K....E..K ...E.I.... .D.......K A.K.FTD..P ...A.I....
Rat             .K....E..K ...E.I.N.. .D......RK A.K.FTD..P ...A......
Human           VN....EQ.K ...E.I.... .D..I....K A.K.FTE..P ...SHI....
Fly             VH..DKCAKK ..V..I.... SN.SQ.L.E. A.K.FTQ..A .F.RMLTEE.
C.elegant       EK..DTGA.. .M...F..P. GECHD..I.G .IDHFSK..P .F.A..K..I
              498                 526
Zebrafish       DKHNAEGKKN TVHVYSRGGA SAVAAASKM* ~
Mouse           ..Y...KP.. AI.T.TQA.S HMA.KGKANL *
Rat             .QY.SQKP.. AI.T.VQA.S HIA.KGKANL *
Human           ..Y...KP.. AI.TFVQS.S HLA.REKANL *
Fly             NLAKSSKF*~ ~~~~~~~~~~ ~~~~~~~~~~ ~
C.elegant       Q.QARSHI*~ ~~~~~~~~~~ ~~~~~~~~~~ ~
```

Figure 2. Optimal alignment of amino acid sequences of catalase among six species in GenBank: zebrafish, this study (EMBL no. AJ007505); mouse (X52108); rat (M11670); human (X04076); fruit fly (X52286); and *C. elegans* (X82175). Numbers refer to amino acid residues of zebrafish. A dot refers to the identities of zebrafish. A dash denotes a deletion. Residues coordinating for secondary structure including α-helixes (α) and *β*-sheets (*β*) are underlined on the basis of mouse catalase cDNA (Reimer et al., 1994). The amino acids with direct interaction with the heme molecule are shown by open boxes. A potential N-glycosylation site, NVTQ, is shown by a gray-shaded box. A unique cAMP- and cGMP-dependent protein kinase phosphorylation site, KKNT, is shown by a wavy line. The predicted peroxisome targeting signal, SKM, is shown by double underlining.

as described by Bairoch et al. (1997). The catalase proximal heme−ligand signature (354−362) and active site signature (64−80) are highly conserved. There is a potential N-glycosylation site NVTQ (439−442) just the same as in mouse catalase. The predicted peroxisome targeting signal, SKM (524−526), was different when compared with the targeting signal, ANL (525−527), of

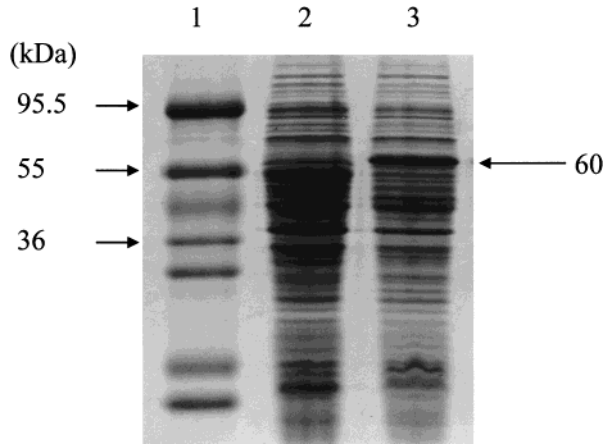

Figure 3. SDS−PAGE of the enzyme extract from the transformant carrying recombinant zebrafish catalase cDNA with Coomassie blue staining. Two aliquots of 15 μL of enzyme extract containing 7.8 μg of protein with the sample buffer were heated for 5 min and then subjected to a 10% SDS−PAGE followed by Coomassie blue staining. Lanes: 1, protein molecular mass marker; 2, BL21(DE3)pLysS carrying pET-20b(+) as control; 3, BL21(DE3)pLysS carrying recombinant catalase cDNA. An arrow denotes the expression protein of zebrafish catalase.

Table 2. Comparison (Percent Identity, Percent Similarity) of Amino Acid Sequences for Catalase of Zebrafish and Other Organisms

organism	EMBL no.	species name	identity (%)	similarity (%)
zebrafish	AJ007505	*Danio rerio*		
swine	D89812	*Sus scrofa*	80.9	86.9
mouse	X52108	*Mus musculus*	80.6	85.8
rat	M11670	*Rattus norvegicus*	79.7	85.0
human	X04076	*Homo sapiens*	77.7	83.7
fruit fly	X52286	*Drosophila melanogaster*	69.2	75.6
C. elegans	X82175	*Caenorhabditis elegans*	64.2	71.7
yeast	X04625	*Saccharomyces cerevisiae*	50.8	58.6

mouse catalase. The unique cAMP- and cGMP-dependent protein kinase phosphorylation site, KKNT (505−508), is shown.

Table 1 lists the primer sequences that were used for cloning and sequencing of the catalase gene. Table 2 shows higher similarity with the amino acid sequence of catalase from mammals such as swine (86.9%), mouse (85.8%), rat (85%), and human (83.7%) than with catalase sequences from other organisms such as fruit fly (75.6%), nematode (71.1%), and yeast (58.6%). These comparisons were done by the PILEUP program developed by the University of Wisconsin Genetics Computer Group.

Figure 3 presents a predominant protein band at 60 kDa corresponding to the expected molecular mass for the catalase subunit detected (lane 3, denoted by an arrow) in SDS−PAGE. This confirms the expression of zebrafish recombinant catalase cDNA in *E. coli* BL21-(DE3)pLysS. Quantitating the protein bands by scanning with a densitometer, the amount of zebrafish catalase was estimated at about 5% of total protein. Enzyme activity assay of the crude extract by the spectrophotometric method showed that the expressed recombinant protein was active catalase with an estimated specific activity of about 3160 units/mg.

As shown in Figure 4, an extra activity band is evident (lane 2, denoted by an arrow); this also confirms

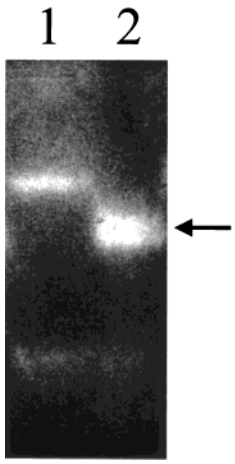

Figure 4. Activity staining of the enzyme extract of recombinant zebrafish catalase. An aliquot of 15 μL of enzyme extract containing 7.8 μg of protein was subjected to 10% native PAGE followed by activity staining. Lanes: 1, BL21-(DE3)pLysS carrying pET-20b(+) as control; 2, BL21(DE3)-pLysS carrying recombinant catalase cDNA. An arrow denotes the active protein of zebrafish catalase.

that recombinant zebrafish catalase was expressed in *E. coli*.

Xenobiotics which cause oxidative stress and repress the antioxidant enzyme system have been demonstrated in various fish species (Pedrajas et al., 1995; Roche and Boge, 1996). No report has mentioned to use the recombinant DNA technique for increasing a fish's resistance to oxidative stress. Previously, we expressed active zebrafish SOD in *E. coli* (Ken et al., 1998), and here we also express active zebrafish catalase in *E. coli*. Furthermore, we can easily purify large amounts of active enzyme by using a His-taq affinity column for further enzymatic or application studies.

CONCLUSIONS

Full-length cDNA encoding a catalase from zebrafish was cloned by the RACE−PCR technique. This clone comprises a complete open reading frame coding for 526 amino acid residues. The coding region was introduced into an expression vector, pET-20b(+), and transformed into *E. coli* BL21(DE3)pLysS. The expression of the active recombinant catalase cDNA clone was shown by SDS−PAGE and activity assay.

LITERATURE CITED

Aebi, H. Catalase in vitro. *Methods Enzymol.* **1984**, *105*, 121−126.

Aksnes, A.; Njaa, L. R. Catalase, glutathione peroxidase and superoxide dismutase in different fish species. *Comp. Biochem. Physiol.* **1981**, *69B*, 893−896.

Bairoch, A.; Bucher, P.; Hofmann, K. The PROSITE database, its status in 1997. *Nucleic Acids Res.* **1997**, *25*, 217−221.

Clerch, L. B. A 3′ untranslated region of catalase mRNA composed of a stem-loop and dinucleotide repeat elements binds a 69-kDa redox-sensitive protein. *Arch. Biochem. Biophys.* **1995**, *317*, 267−274.

Deisseroth, A.; Dounce, A. L. Catalase: physical and chemical properties, mechanism of catalysis, and physiological role. *Physiol. Rev.* **1970**, *50*, 319−375.

Furuta, S.; Hayashi, H.; Hijikata, M.; Miyazawa, S.; Osumi, T.; Hashimoto, T. Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liver catalase. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 313−317.

Harman, D. Aging: a theory based on free radical and radiation chemistry. *J. Gerontol.* **1956**, *11*, 298−300.

Ken, C. F.; Shaw, J. F.; Wu, J. L.; Lin, C. T. Molecular cloning of a cDNA coding for copper/zinc-superoxide dismutase from zebrafish and its expression in *Escherichia coli*. *J. Agric. Food Chem*. **1998**, *46*, 2863−2867.

Klotz, M. G.; Klassen, G. R.; Loewen, P. C. Phylogenetic relationships among prokaryotic and eukaryotic catalases. *Mol. Biol. Evol.* **1997**, *14*, 951−958.

Laemmli, U. K. Cleavage of structure protein during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227*, 680−685.

Lin, Z. H.; Wang, Y. F.; Sarai, A.; Yasue, H. Swine catalase deduced from cDNA and localization of the catalase gene on swine chromosome 2p16-p15[1]. *Biochem. Genet.* **1997**, *35*, 297−302.

Mather-Mihaich, E.; Diguilio, R. T. Antioxidant enzyme activities and malondialdehyde, glutathione and methemoglobin concentration in channel catfish exposed to DFF and *N*-butylmercaptan. *Comp. Biochem. Physiol.* **1986**, *85C*, 427−432.

Muthy, M. R. N.; Reid, T. J.; Sicignano, A.; Tanaka, N.; Rossmann, M. G. Structure of beef liver catalase. *J. Mol. Biol.* **1981**, *152*, 465−499.

Nicholls, D. M.; Teichert-Kuliszewska, K.; Girgis, G. R. Effect of chronic mercuric chloride exposure on liver and muscle enzymes in fish. *Comp. Biochem. Physiol.* **1989**, *94C*, 265−270.

Orr, E. C.; Bewley, G. C.; Orr, W. C. cDNA and deduced amino acid sequence of *Drosophila* catalase. *Nucleic Acids Res.* **1990**, *18*, 3663.

Orr, W. C.; Sohal, R. S. Extension of life-span by overexpression of superoxide dismutase and catalase in *Drosophila melanogaster*. *Science* **1994**, *263*, 1128−1130.

Pedrajas, J. R.; Peinado, J.; Lopez-Barea, J. Oxidative stress in fish exposed to model xenobiotics. Oxidatively modified forms of Cu, Zn-superoxide dismutase as potential biomarkers. *Chem.-Biol. Interact.* **1995**, *98*, 267−282.

Quan, F.; Korneluk, R. G.; Tropak, M. B.; Gravel, R. A. Isolation and characterization of the human catalase gene. *Nucleic Acids Res.* **1986**, *14*, 5321−5335.

Radi, A. A. R.; Marcovics, B. Effects of metal ions on the antioxidant enzyme activities, protein contents and lipid peroxidation of carp tissues. *Comp. Biochem. Physiol.* **1988**, *90C*, 69−72.

Reimer, D. L.; Bailley, J.; Singh, S. M. Complete cDNA and 5′ genomic sequences and multilevel regulation of the mouse catalase gene. *Genomics* **1994**, *21*, 325−336.

Roche, H.; Boge, G. Fish blood parameters as a potential tool for identification of stress caused by environmental factors and chemical intoxication. *Mar. Environ. Res.* **1996**, *41*, 27−43.

Schroeder, W. A.; Shelton, J. R.; Shelton, J. B.; Robberson, B.; Apell, G.; Fang, R. S.; Bonaventura, J. The complete amino acid sequence of bovine liver catalase and the partial sequence of bovine erythrocyte catalase. *Arch. Biochem. Biophys.* **1982**, *214*, 397−421.

Shaffer, J. B.; Preston, K. E.; Shepard, B. A. Nucleotide and deduced amino acid sequences of mouse catalase: molecular analysis of a low activity mutant. *Nucleic Acids Res.* **1990**, *18*, 4941.

St. Dimitrova, M.; Tishinova, V.; Velcheva, V. Combined effect of zinc and lead on the hepatic superoxide dismutase-catalase system in carp (*Cyprinus carpio*). *Comp. Biochem. Physiol.* **1994**, *108C*, 43−46.

Wayne, L. G.; Diaz, G. A. A double staining method for differentiating between two classes of mycobacterial catalase in polyacrylamide electrophoresis gels. *Anal. Biochem.* **1986**, *157*, 89−92.

Wdzieczak, J.; Zalesna, G.; Wujec, A.; Peres, G. Comparative studies on superoxide dismutase, catalase and peroxidase levels in erthrocytes and livers of different freshwater and marine fish species. *Comp. Biochem. Physiol.* **1982**, *73B*, 361−365.

Witas, H.; Gabryelak, T.; Matkovics, B. Comparative studies on superoxide dismutase and catalase in livers of fish and other Antarctic vertebrates. *Comp. Biochem. Physiol.* **1984**, *77C*, 409−411.

Received for review July 27, 1999. Revised manuscript received March 10, 2000. Accepted March 22, 2000. This work was supported by the National Science Council of the Republic of China under Grant NSC 88-2313-B-019-017 to C.-T.L. and by the Council of Agriculture, Executive Yuan, under Grant 88-BT-2.1-FID-01(1-7) to C.-T.L.

JF990838+