

cDNA sequence coding for human kidney catalase

G.I.Bell, R.C.Najarian, G.T.Mullenbach and R.A.Hallewell

Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA

Submitted 23 May 1986

cDNAs encoding human catalase (1) were isolated from a kidney library in  $\lambda$ gt10 by hybridization with synthetic oligonucleotides to the 5' and 3' ends of the partial cDNA sequence reported by Korneluk et al. (2). The sequence of the 2,264 bp insert of  $\lambda$ hCAT-2 is presented below. This sequence includes 70 bp of the 5'-untranslated region, the 1,581 bp coding region and 613 bp of the 3'-untranslated region. This clone does not have a polyA tract but as its 3' end is adjacent to a canonical polyA addition site, only about 20-25 bp of the 3'-untranslated region may be missing. There are several differences between our sequence and that presented by Korneluk et al. (2). In the latter, the codons for Gln 386 and Asp 388 are CAA and GAC, respectively, and there are three substitutions and a deletion of one nucleotide in the 3'-untranslated region.

```

-1      1      10
Met Ala Asp Ser Arg Asp Pro Ala Ser Asp Gln Met
TTTGCTGCTGAGGGTGGAGACCCACGAGCCGAGGCCCTCTGCA GTTCTGCACAGCAAACCGCACBCT
ATG GCT GAC AGC CGG GAT CCC GCC AGC GAC CAG ATG 106

      20      40
Gln His Trp Lys Glu Gln Arg Ala Gln Lys Ala Asp Val Leu Thr Thr Gly Ala Gly Asn Pro Val Gly Asp Lys Leu Asn Val Ile
CAG CAC TGG AAG GAG CAG CGG GCC GCG CAG AAA GCT GAT GTC CTG ACC ACT GGA GCT GGT AAC CCA GTA GGA GAC AAA CTT AAT GTT ATT 196

      50      70
Thr Val Gly Pro Arg Gly Pro Leu Leu Val Gln Asp Val Val Phe Thr Asp Glu Met Ala His Phe Asp Arg Glu Arg Ile Pro Glu Arg
ACA GTA GGG CCC CGT GGG CCC CTT GTT CAG GAT GTG GTT TTC ACT GAT GAA ATG GCT CAT TTT GAC CGA GAG AGA ATT CCT GAG AGA 286

      80      90
Val Val His Ala Lys Gly Ala Gly Ala Phe Gly Tyr Phe Glu Val Thr His Asp Ile Thr Lys Tyr Ser Lys Ala Lys Val Phe Glu His
GTT GTG CAT GCT AAA GGA GCA GGG GCC TTT GGC TAC TTT GAG GTC ACA CAT GAC ATT ACC AAA TAC TCC AAG GCA AAG ATA TTA TTT GAG CAT 376

      110      120
Ile Gly Lys Lys Thr Pro Ile Ala Val Arg Phe Ser Thr Val Ala Gly Glu Ser Gly Ser Ala Asp Thr Val Arg Asp Pro Arg Gly Phe
ATT GGA AAG AAG ACT CCC ATC GCA GTT CGG TTC TCC ACT GTT GCT GGA GAA TCG GGT TCA GCT GAC ACA GTT CGG GAC CCT CGT GGG TTT 466

      140      150
Ala Val Lys Phe Tyr Thr Glu Asp Gly Asn Trp Asp Leu Val Gly Asn Asn Thr Pro Ile Phe Phe Ile Arg Asp Pro Ile Leu Phe Pro
GCA GTG AAA TTT TAC ACA GAA GAT GGT AAC TGG GAT CTC GTT GGA AAT AAC ACC CCC ATT TTC TTC ATC AGG GAT CCC ATA TTG TTT CCA 556

      170      180
Ser Phe Ile His Ser Gln Lys Arg Asn Pro Gln Thr His Leu Lys Asp Pro Asp Met Val Trp Asp Phe Trp Ser Leu Arg Pro Glu Ser
TCT TTT ATC CAC AGC CAA AAG AGA AAT CCT CAG ACA CAT CTG AAG GAT CCG GAC ATG GTC TGG GAC TTC TGG AGC CTA CGT CCT GAG TCT 646

      200      210
Leu His Gln Val Ser Phe Leu Phe Ser Asp Arg Gly Ile Pro Asp Gly His Arg His Met Asn Gly Tyr Gly Ser His Thr Phe Lys Leu
CTG CAT CAG GTT TCT TTC TTG TTC AGT GAT CGG GGG ATT CCA GAT GGA CAT CGC CAC ATG AAT GGA TAT GGA TCA CAT ACT TTC AAG CTG 736

      230      240
Val Asn Ala Asn Gly Glu Ala Val Tyr Cys Lys Phe His Tyr Lys Thr Asp Gln Gly Ile Lys Asn Leu Ser Val Glu Asp Ala Ala Arg
GTT AAT GCA AAT GGG GAG GCA GTT TAT TGC AAA TTC CAT TAT AAG ACT GAC CAG GGC ATC AAA AAC CTT TCT GTT GAA GAT GCG GCG AGA 826

      260      270
Leu Ser Gln Glu Asp Pro Asp Tyr Gly Ile Arg Asp Leu Phe Asn Ala Ile Ala Thr Gly Lys Tyr Pro Ser Trp Thr Phe Tyr Ile Gln
CTT TCC CAG GAA GAT CCT GAC TAT GGC ATC CGG GAT CTT TTT AAC GCC ATT GCC ACA GGA AAG TAC CCC TCC TGG ACT TTT TAC ATC CAG 916

      290      300
Val Met Thr Phe Asn Gln Ala Glu Thr Phe Pro Phe Asn Pro Phe Asp Leu Thr Lys Val Trp Pro His Lys Asp Tyr Pro Leu Ile Pro
GTC ATG ACA TTT AAT CAG GCA GAA ACT TTT CCA TTT AAT CCA TTC GAT CTC ACC AAG GTT TGG CCT CAC AAG GAC TAC CCT CTC ATC CCA 1006

      320      330
Val Gly Lys Leu Val Leu Asn Arg Asn Pro Val Asn Tyr Phe Ala Glu Val Glu Gln Ile Ala Phe Asp Pro Ser Asn Met Pro Pro Gly
GTT GGT AAA CTG GTC TTA AAC CGG AAT CCA GTT AAT TAC TTT GCT GAG GTT GAA CAG ATA GCC TTC GAC CCA AGC AAC ATG CCA CCT GGC 1096
    
```

```

350
Ile Glu Ala Ser Pro Asp Lys Met Leu Gln Gly Arg Leu Phe Ala Tyr Pro Asp 360
ATT GAG GCC AGT CCT GAC AAA ATG CTT CAG GGC CGC CTT TTT GCC TAT CCT GAC ACT CAC CGC CAT CGC CTG GGA CCC AAT TAT CTT CAT 1186

380
Ile Pro Val Asn Cys Pro Tyr Arg Ala Arg Val Ala Asn Tyr Gln Arg Asp Gly Pro Met Cys Met Gln Asp Asn Gln Gly Gly Ala Pro
ATA CCT GTG AAC TGT CCC TAC CGT GCT CGA GTG GCC AAC TAC CAG CGT GAT GGC CCG ATG TGC ATG CAG GAC AAT CAG GGT GGT GCT CCA 1276

410
Asn Tyr Tyr Pro Asn Ser Phe Gly Ala Pro Gln Gln Gln Pro Ser Ala Leu Gln His Ser Ile Gln Tyr Ser Gly Glu Val Arg Arg Phe
AAT TAC TAC CCC AAC AGC TTT GGT GCT CCG GAA CAA CAG CCT TCT GCC CTG GAG CAC AGC ATC CAA TAT TCT TCT GGA GAA GTG CCG AGA TTC 1366

440
Asn Thr Ala Asn Asp Asp Asn Val Thr Gln Val Arg Ala Phe Tyr Val Asn Val Leu Asn Gln Glu Gln Arg Lys Arg Leu Cys Gln Asn
AAC ACT GCC AAT GAT GAT AAC GAT ACT CAG GTG CGG GCA TTC TAT GTG AAC GTG CTG AAT GAG GAA CAG AGG AAA CGT CTG TGT GAG AAC 1456

470
Ile Ala Gly His Leu Lys Asp Ala Gln Ile Phe Ile Gln Lys Lys Ala Val Lys Asn Phe Thr Glu Val His Pro Asp Tyr Gly Ser His
ATT GCC GGC CAC CTG AAG GAT GCA CAA ATT TTC ATC CAG AAG AAA GCG GTC AAG AAC TTC ACT GAG GTC CAC CCT GAC TAC GGG AGC CAC 1546

500
Ile Gln Ala Leu Leu Asp Lys Tyr Asn Ala Glu Lys Pro Lys Asn Ala Ile His Thr Phe Val Gln Ser Gly Ser His Leu Ala Ala Arg
ATC CAG GCT CTT CTG GAC AAG TAC AAT GCT GAG AAG CCT AAG AAT GCG ATT CAC ACC TTT GTG CAG TCC GGA TCT CAC TFG GCG GCA AGG 1636

526
Glu Lys Ala Asn Leu OP
GAG AAG GCA AAT CTG TGA GGCCGGGGCCCTGCACCTGTGCAGCGAACGTTAGCGTTTCATCCGTGTAACCCCGCTCATCTGGATGAAGATTCTCCTGTGCTAGATGTGCAAAATG 1750
CAAGCTAGTGCCCTCAAATAGAGAATCCCACCTTCTATAGCAGATTGTGTAACAATTTTAATGCTATTTCCCGAGGGAAAATGAAGGTTAGGATTTAACAGCTATTTAAAAAAAAT 1870
TTGTTTTGACGGATGATTGGATTATTCATTTAAAATGATTAGAAGGCAAGTTTCTAGCTTAGAAATATGATTTTATTTGACAAAATTTGTTGAAATATGTATGTTTACATATCACCTCA 1990
TGCCCTATTATATAAAATATGGCTATAAATATATAAAAAAGAAAAGATAAAGATGATCTACTCAGAAATTTTTATTTTCTAAGGTTCTCATAGGAAAAGTACATTTAATACAGCAGTGT 2110
CATCAGAAGATACTTGAGCACCGTCAATGGCTTAATGTTTATCTCGTATAAATATGATCAAATTCATTTTTTTCACGTGGATTACATTAATGTTAATTCAGCAGTATTTACACAGAGA 2230
TCAATTTGTAATTGCTTACATTTTACAATAAAT 2264

```

REFERENCES

1. Schroeder, W.A. et al. (1982) Arch. Biochem. Biophys. 214, 422-424.
2. Korneluk, R.G. et al. (1984) J. Biol. Chem. 259, 13819-13823.