Volume 15 Number 10 1987 Nucleic Acids Research

Nucleotide sequence of a human ubiquitin Ub B processed pseudogene

Rohan T.Baker and Philip G.Board

Department of Human Genetics, John Curtin School of Medical Research, Australian National University, PO Box 334, Canberra 2601, Australia
Submitted April 23, 1987 Accession no. Y00332

The human ubiquitin Ub B gene consists of 3 tandem repeats of the 76 amino acid coding unit followed by an extra cys codon, and contains a 715bp intron in its 5' non-coding flank [1]. During the isolation of this gene, two Ub B processed pseudogenes were also isolated. However, one pseudogene contained two coding unit repeats, while the other only one [1]. Here, we report the sequence of another single coding unit Ub B processed pseudogene, EHB6. Numbering is from the A of the start codon. Noted features are: start codon, stop codon, polyadenylation signal and encoded poly-A tail (underlined); flanking direct repeats (overlined); and positions corresponding to the gene CAP and polyadenylation sites (arrowheads). The latter sites denote the 5' and 3' limits of homology to the gene. The pseudogene 5' and 3' flanks are respectively 67.6% and 88.7% homologous to the corresponding gene regions [1]. The translation is shown above the sequence and differs from ubiquitin at 15 positions (*) due to single base changes (underlined bases). This pseudogene has an open reading frame but is presumably not transcribed in vivo.

```
EHB6:-175 GCCCTCTGTTCTTCATCATGTGGGCCTCACACAGCTGTAGGCATCATGAGGAACA
-120 AGTAAGAAAGAGGGAACAGAGTGGCGAATGGCCTTGTTGAATGGCGAGCGTCTGGAGGCA
 -60 TTCCAGTGGCTGAATGTGATTGGTGATCTGCAGCTTCCTCGCATCCTCCAAGAGGTCAAA
      M  Q  I  F  L* K  T  L  T  G  K  T  I  T  L  E  V  E  P  S
   1 ATGCAAATCTTCCTGAAAACCCTGACCGGCAAGACCATCACCCTGGAGGTGGAGCCAAGT
      D  I* L* Q* N  V  K  A  K  I  H* V* K  E  G  I  P  P  D  Q
  61 GACATCCTCCAAAATGTGAAGGCCAAGATCCATGTTAAAGAGGGCATCCCCCCTGACCAG
      H* S* L  I  F  V* G  K  Q  L  E  D  G  C* T  V* C* D  Y  N
 121 CACAGTCTCATCTTTGTAGGCAAGCAGTTAGAAGATGGCTGCACTGTTTGTGACTACAAC
      I  Q  K  E  S  A* L  H  L  V  L  H* L  R  G  G  Y*ter
 181 ATTCAGAAAGAGTCAGCCCTGCACCTGGTCCTCCATCTGAGGGGTGGCTATTAATTCTTC
 241 AGTCTTGCATTCGTAGTGACAAGTGATGGCACTACTCTGCACTAAAGCCATTTGCCCCAA
 301 TTTAAGTTTATAAATTACCAGTTTCGGTAATAGCTGAACCTGCTCAAAATGTTAATAAGC
 361 GTTTTGTTGCATGGTTAAAAAAAAAAAAAAGGGAATAAGACAGAAACCACAGTTCTTCATA
 421 GCTTAATCTCATCCATCATTTTTGCCATATTCCATTGATTAGAAGCAAGTCACTGGTCTA
```

REFERENCE
[1] Baker,R.T. and Board,P.G. (1987) Nucl. Acids Res. 15, 443-463