Nucleotide sequence of porcine liver albumin

Janet Weinstock and Graham S.Baldwin

Ludwig Institute for Cancer Research, Melbourne Tumour Biology Branch, PO Royal Melbourne Hospital, Victoria 3050, Australia

Submitted August 16, 1988 Accession no. X12422

Albumin is the major protein constituent of serum (1). A clone (A13, 364 bases) encoding the N terminal 99 amino acids of mature porcine albumin and 22 amino acids of the pre sequence was isolated from a porcine liver cDNA library (λ gt 11, Clontech, Palo Alto, CA). The probe used was a 41 base, 2-fold degenerate oligonucleotide complementary to the sequence encoding amino acids 28-41 of preproalbumin, the hybridization conditions were as described by Ullrich (2), and the filters were washed in 6 x SSC, 0.1% SDS at 42°C. An 1887 base clone (A11) encoding the remainder of the mature protein was then isolated from the same library by screening with nick-translated clone A13. The two clones were subcloned into m13 and fully sequenced in both directions by Sanger's dideoxy method using Sequenase (United States Biochemicals, Cleveland, Ohio) and a series of synthetic oligonucleotides as primers. Bases 68-1816 of the overlapped clones encode a mature protein of 583 amino acids with 76% identity to human albumin.

```
GTGGGTGACTTTTATTTCCCTTCTCTTTCTCTTCAGCTCTGCTTATTCCAGGGGTGTGTTTCGTCGAGATACATACAAGAGTGAAAATTGCTCATCGGTTTAAAGATTTGGGAGAACAATA
      10        20        30        40        50        60        70        80        90       100       110       120
TTTCAAAGGCCTAGTGCTGATTGCCTTTTCTCAGCATCTCCAGCAATGCCCATATGAAGAGCATGTGAAATTAGTGAGGGAAGTAACTGAGTTTGCAAAAACATGTGTTGCTGATGAGTC
     130       140       150       160       170       180       190       200       210       220       230       240
AGCTGAAAATTGTGACAAGTCAATTCACACTCTCTTTGGAGATAAATTATGTGCAATTCCATCCCTTCGTGAACACTATGGTGACTTGGCTGACTGCTGTGAAAAAGAAGAGCCTGAGAG
     250       260       270       280       290       300       310       320       330       340       350       360
AAACGAATGCTTCCTCCAACACAAAAATGATAACCCCGACATCCCTAAATTGAAACCAGACCCTGTTGCTTTATGCGCTGACTTCCAGGAAGATGAACAGAAGTTTTGGGGAAAATACCT
     370       380       390       400       410       420       430       440       450       460       470       480
ATATGAAATTGCCAGAAGACATCCCTATTTCTACGCCCCAGAACTCCTTTATTATGCCATTATATATAAAGATGTTTTTTCAGAATGCTGCCAAGCTGCTGATAAAGCTGCCTGCCTGTT
     490       500       510       520       530       540       550       560       570       580       590       600
ACCAAAGATTGAGCATCTGAGAGAAAAAGTACTGACTTCCGCCGCCAAACAGAGACTTAAGTGTGCCAGTATCCAAAAATTCGGAGAGAGAGCTTTCAAAGCATGGTCATTAGCTCGCCT
     610       620       630       640       650       660       670       680       690       700       710       720
GAGCCAGAGATTTCCCAAGGCTGACTTTACAGAGATTTCCAAGATAGTGACAGATCTTGCAAAAGTCCACAAGGAATGCTGCCATGGTGACCTGCTTGAATGTGCAGATGACAGGGCGGA
     730       740       750       760       770       780       790       800       810       820       830       840
TCTTGCCAAATATATATGTGAAAATCAAGACACAATCTCCACTAAACTGAAGGAATGCTGTGATAAGCCTCTGTTGGAAAAATCCCACTGCATTGCTGAGGCAAAAAGAGATGAATTGCC
     850       860       870       880       890       900       910       920       930       940       950       960
TGCAGACCTGAACCCATTAGAACATGATTTTGTTGAAGATAAGGAAGTTTGTAAAAACTATAAAGAAGCAAAGGATGTCTTCCTGGGCACGTTTTTGTATGAGTATTCAAGAAGGCACCC
     970       980       990      1000      1010      1020      1030      1040      1050      1060      1070      1080
AGACTACTCTGTCTCATTGCTGCTGAGAATTGCCAAGATATATGAAGCCACACTGGAGGACTGCTGTGCCAAAGAGGATCCTCCGGCATGCTATGCCACAGTGTTTGATAAATTTCAGCC
     1090      1100      1110      1120      1130      1140      1150      1160      1170      1180      1190      1200
TCTTGTGGATGAGCCTAAGAATTTAATCAAACAAAACTGTGAACTTTTTGAAAAACTTGGAGAGTATGGATTCCAAAATGCGCTCATAGTTCGTTACACCAAGAAAGTACCCCAAGTGTC
     1210      1220      1230      1240      1250      1260      1270      1280      1290      1300      1310      1320
AACTCCAACTCTTGTGGAGGTCGCAAGAAAACTAGGACTAGTGGGCTCTAGGTGTTGTAAGCGTCCTGAAGAAGAAAGACTGTCCTGTGCTGAAGACTATCTGTCCCTGGTCCTGAACCG
     1330      1340      1350      1360      1370      1380      1390      1400      1410      1420      1430      1440
GTTGTGCGTGTTGCACGAGAAGACACCAGTGAGCGAAAAAGTTACCAAATGCTGCACAGAGTCCTTGGTGAACAGACGGCCTTGCTTTTCTGCTCTGACACCAGACGAAACATACAAACC
     1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550      1560
CAAAGAATTTGTTGAGGGAACCTTCACCTTCCATGCAGACCTATGCACACTTCCTGAGGATGTGAAACAAATCAAGAAGCAAACTGCACTCGTTGAGTTGTTGAAACACAAGCCTCATGC
     1570      1580      1590      1600      1610      1620      1630      1640      1650      1660      1670      1680
AACAGACGAACAACTGAGAACTGTCCTGGGCAACTTTGCAGCCTTTGTACAAAAGTGCTGCGCCGCTCCTGACCATGAGGCCTGCTTTGCTGTGGAGGGTCCGAAATTTGTTATTGAAAT
     1690      1700      1710      1720      1730      1740      1750      1760      1770      1780      1790      1800
TCGAGGGATCTTAGCCTAAACAACACAGTGACAAGCATCTCAGACTACCCTGAGAATAAGAGAAAGAGAAATGAAGACCTAGACTTATCCACG
     1810      1820      1830      1840      1850      1860      1870      1880      1890
```

ACKNOWLEDGEMENTS

We thank Michael Eefting and Gavin Reid for synthesis of many oligonucleotides.

REFERENCES

1. Schultze, H.E. and Heremans, J.F. (1966) Molecular Biology of Human Proteins. Elsevier, Amsterdam. Vol. 1, pp 182-183
2. Ullrich, A., Berman, C.A., Dull, T.J., Gray, A. and Lee, J.M. (1984) EMBO J. **3**, 361-364.