

Nucleotide sequence of porcine liver transferrin

Graham S. Baldwin and Janet Weinstock

Ludwig Institute for Cancer Research, Melbourne Tumour Biology Branch, PO Royal Melbourne Hospital, Victoria 3050, Australia
 Submitted July 26, 1988

Accession no. X12386

Transferrin is the major iron-binding glycoprotein of serum (1). A 966 base clone (Tf6) encoding the N-terminal 321 amino acids of the mature form of porcine transferrin was isolated from a porcine liver cDNA library (λ gt11, Clontech, Palo Alto, CA) by screening with a 38 base, 8-fold degenerate oligonucleotide complementary to the sequence encoding the N-terminal 13 amino acids. The conditions were as described by Ullrich (2), except that the filters were washed in 6 x SSC, 0.1% SDS at 42°C. A 1591 base clone (Tf5) encoding the remainder of the mature protein was then isolated from the same library by screening with nick-translated clone Tf6. 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. Differences between the two clones were noted at positions 668 (G in 5, A in 6), 788 (T in 5, C in 6) and 925 (A in 5, G in 6). Only the latter change is reflected in the amino acid sequence, where it results in substitution of R for K. Bases 3-2090 of the overlapped clones encode a mature protein of 696 amino acids with 70% identity to human transferrin.

```
CGTGC GCCAAAAAAGTCTGAAGATGGTGCACCATTTCAAATCAGGAGGCCAATAAGTGTCCAGTTTCGGTGA AAAACATGAGTAAAGCCGTTAAGAATGGTCTCTGTCTCTCTGTGTGA
10 20 30 40 50 60 70 80 90 100 110 120
AAAAATCCTCTTACTCGGATTCATCAAGCCATCAGGATAAAGAAGCAGATGCTGTGACACTGGATGCGGGTTTGGTATTTGAGGCAGGCTGGCCCTTACAACCTGAAGCCTGTTG
130 140 150 160 170 180 190 200 210 220 230 240
TGGCAGAGTTCTATGGGCAGAAAGATAATCCACAAACCCATTATTACGGCGTGGCCGGTGGTGAAGAAGGGCAGCAACTTCCAGTGGAAACAGCTCCAGAGCAAGAGGCTCTGCCACACAG
250 260 270 280 290 300 310 320 330 340 350 360
GCCTCGGCGCTTCTGCTGATGGATCCCATGGGATTTACTTTAGATCAACTGCCTGAGCCACGGAAACCTATTGAGAAAGCAGTGGCCAGTTTCTTCCAGACGCTGCGTCCCTC
370 380 390 400 410 420 430 440 450 460 470 480
GTGCGGATCCGGTGAATCTCCCAAACTGTGCCAACAGTGTGGGAAAGGGGTGAAAGTGTGCTGCTCCAACTGATGACCTACTTGGCTACGGGGTGCCTCAATTGCCTGA
490 500 510 520 530 540 550 560 570 580 590 600
AGGAGATGCTGGGAGCGTGCCTTTTGTCAAGCACTCAACAGTATTGGAGAACTGCTGTACAGGGGACAGGGACAGTACGAGCTCTCTGAGGACCAACACAGGAGGCGCTGGG
610 620 630 640 650 660 670 680 690 700 710 720
ATGACTACGAGAACTGCTACTGGCAGGTCCTTCCACAGCCGCTGGTGGCCGAGCGTGGATGTCAGGAGACTCAATCTGGAGCTTCTCAACAGGCCAGGAGATTTGGCA
730 740 750 760 770 780 790 800 810 820 830 840
GAGATAATCTCCAGACTCCAGCTTTTTCAGCTCTCTCACTGGGAAGACTGCTGTTTAAGGATTCGCAATGGTCTCTTAAGATTCCTCTTAAGATGGACTCTCCCTGTACTCTGG
850 860 870 880 890 900 910 920 930 940 950 960
GGTATCAGTATGCTACTGCTCTCGGAATCTAAGAGAGAAATAGCCCTGATAGCTCAAGAATGAATGCAAGAAGTGGAGTGGTGGCATAGGCCACAGGAGAGCCAAAAATGCC
970 980 990 1000 1010 1020 1030 1040 1050 1060 1070 1080
ATGCGTGGAGCATAAACAGTGGTGGGAAAATAGATGTGTATCCCGAGAAAACCCGAGACTGTATGCCAAGATCGTGAAGAGAGGGCTGATGCCATGACTTGGATGGAGCTACA
1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200
TCTACATTCGCGGCAAGTGTGGTGTGGTGTCTGCTGGCAGAGACTACAAAATCAGGGTGA AAAACGTGTGTA AACACACAGAGAAAGGGATTTGGCTGTGGCTGTGTTAAGAAAT
1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320
CATCAGTCTGTACTCACTGGAACAATCTAAAGGGCAGAGTCTGCCACACTGCATAGATAGAATGCTGGCTGGAACATCCCATGGGCTGCTCAACAAGAATTAACAGTT
1330 1340 1350 1360 1370 1380 1390 1400 1410 1420 1430 1440
GTAATTTGATCAGTTTTTCGGTGAAGGCTGTGCCCTGGTCTCAGCGAAAATCCAGTCTCTGTGCTGTGTATCGGCTCGGAAGCTGGCCCTGGAAGGGAAATGCTTAGCCAACACC
1450 1460 1470 1480 1490 1500 1510 1520 1530 1540 1550 1560
ATGAGAGATACTATGGCTACACGGGGCTTTCAGGTGTCTGGTTGAGAAGGGAGATGTGGCCTTTTGAAGGATCAGGTTGTCCAGCAGAACTGACGGAAAAAACAAGGATGATGGG
1570 1580 1590 1600 1610 1620 1630 1640 1650 1660 1670 1680
CAAAAGTCTGAAGCAGATGGATTTTGAAGTGTGTCGAGAACGGCCAGGGAGCTGGGATAATGCTGAGAACTGCCACCTGGCCAGAGCCCAACCAATGCTGTGATGACAGGG
1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800
ATGATAAAGTACCTGTGTTGCTGAGGAGTTACTTAAACAGCAGGCTCAATTTGGAAGACATGAACCGACTGCTAAGCAGCTTTTGCATGTTCAAGTCAAAATACCAAGGACCTCCTCT
1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920
TCAGGGATGACACCAATGTTTGGCCAGAGTTGGCAAACTACGTATGAATCTACTTAGGACCGACTACATCAGGGCTGGTCAACTGAGAAAAATGCTCAACTCAAAACTCCTGG
1930 1940 1950 1960 1970 1980 1990 2000 2010 2020 2030 2040
AAGCCTGTACTTCCACAGTGC1AAAAATCCAAGATGGAGACTACCACCTGATGGAGACGGGAGCTCACGGGACCCATAAGTTCCTCTGGTGTGCTGGCCTGAGTGGTTGGTGGTC
2050 2060 2070 2080 2090 2100 2110 2120 2130 2140 2150 2160
CTTACACTTCTGGTGGCCACTCTCAGGACAAAAAATAAAAAATATTATCAATTTTGTCTTTTAAAAAATAAAAAAAAAAAAAAAAAAAGCCG
2170 2180 2190 2200 2210 2220 2230 2240 2250
```

REFERENCES

1. Aisen, P. and Listowsky, I. (1980) *Ann. Rev. Biochem.* **49**, 357-393.
2. Ullrich, A., Berman, C.A., Dull, T.J., Gray, A. and Lee, J.M. (1984) *EMBO J.* **3**, 361-364.