The primary structure of rat α 2 μ globulin-related protein

Yuen-Ling Chan, Veronica Paz and Ira G.Wool*

Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA

Submitted October 26, 1988

Accession no. X13295

The amino acid sequence of a protein was deduced from the sequence of nucleotides in a rat recombinant cDNA isolated fortuitously whilst screening a library with synthetic oligonucleotides that designated a sequence in ribosomal protein P1. The open reading frame is 594 nucleotides in length and encodes a protein containing 198 residues and having a molecular weight of 22,560. A search (Gene-Bank) revealed that this protein is related to rat α 2u globulin (1); in an alignment there were 70 identities out of 135 possible matches with 55 breaks (93 if one scores conservative substitutions). The RELATE score for the comparison is 8.36 SD and the ALIGN score is 7.84. Rat α 2u globulin is synthesized in the liver, secreted into the serum, and excreted in the urine of males; the synthesis of the protein, whose function is not known, is affected by hormones and by other factors (1 and references therein). In the α 2u globulin-related protein, 13 of the first 16 NH₂-terminal residues are hydrophobic suggesting this is a signal sequence for secretion. The mRNA for the α 2u globulin-related protein was determined by Northern hybridization to be 1 kb in length and to be increased 10-fold in concentration in the liver of male rats 18 h after hepatectomy; the amount of mRNA was the same in liver and prostate of male rats but lower in the liver of females. Southern hybridization indicated that the α 2u globulin-related gene is present in a single copy. (Supported by NIH grant GM-21769.)

```

30
CTC TTC CTC CTC CCG CAC ACA TCG GAC CTA GTA GCT GCT GAA ACC

60
ATG GGC CTG GGT GTC CTG TGT CTG GGC CTT GTC CTG CTT GGG GTC CTG CAG AGG CAG GCC CAG CAG TCA ACT CAG AAC TTG ATC OCT GGC
MET GLY LEU GLY VAL LEU CYS LEU ALA LEU VAL LEU LEU GLY VAL LEU GLN ARG GLN ALA GLN ASP SER THR GLN ASN LEU ILE PRO ALA

100
150
CCA CCT CTG ATC AGT GTG CCC CTG CAG CCA GGC TTC TGG ACC GAA CCG TTC CAG GGC AGG TGG TTC GTT GTC GGC CTG GCA GCG AAT GCG
PRO PRO LEU ILE SER VAL PRO LEU GLN PRO GLY PHE THR THR GLU ARG PHE GLN GLY ARG TYR PHE VAL VAL GLY LEU ALA ALA ASN ALA

200
240
GTC CAG AAA GAA AGA CAA AOC CAC TTT ACC ATG TAC AGC ACC ATC TAT GAG CTA CAG GAA GAC AAT AGC TAC AAC GTC ACT TCC ATC CTC
VAL GLN LYS GLU ARG GLN SER ARG PHE THR MET TYR SER THR ILE TYR GLU LEU GLN GLU ASP ASN SER TYR ASN VAL THR SER ILE LEU

300
330
GTC AAG GGC CAG GGC TGT CAC TAC TGG ATC AGA ACA TTT GTC CCA AGC TCC AAG CCG GGC CAG TTC ACC CTG GGG AAT ATT CAC AGC TAC
VAL ARG GLY GLN GLY CYS ARG TYR TYR ILE ARG THR PHE VAL PRO SER SER ARG PRO GLY GLN PHE THR LEU GLY ASN ILE HIS SER TYR

400
420
CCT CAG ATA CAG ACC TAC GAT GTS CAA GTG GCC GAC ACT GAC TAC GAC CAG TTT GGC ATG GTA TTT TTC CAG AAG ACC TCT GAA AAC AAA
PRO GLN ILE GLN SER TYR ASP VAL GLN VAL ALA ASP THR ASP TYR ASP GLN PHE ALA MET VAL PHE PHE GLN LYS THR SER GLU ASN LYS

500
540
CAG TAC TTC AAA CTC ACC CTG TAC GGA AGA ACC AAG GGG CTG TCC GAT GAA CTG AAG GAG CGA TTC GTC AGC TTT GGC AAG TCT CTG GGC
GLN TYR PHE LYS VAL THR LEU TYR GLY ARG THR LYS GLY LEU SER ASP GLU LEU LYS GLU ARG PHE VAL SER PHE ALA LYS SER LEU GLY

600
630
CTC AAG GAT AAC AAC ATC GTT TTC GTT CCC ACC GAC CAA TCC ATT GAC AAC TGA ACA GAC GGT GAG CCG GCG TGA CTG GGA TOT GCA
LEU LYS ASP ASN ASN ILE VAL PHE SER VAL PRO THR ASP GLN CYS ILE ASP ASN END

700
720
GTG GGC TGA TGG TTC AGG TCC CAC CTG TCT GTC TGC CCG TCC ATC TTT CCT GTT GGC AGA GAA TCA CCT GGC TGC CCC ACC ACC CAT GAT

780
810
TCC ACC AAG CAT CTG ATC CCT CTT ATT TGA TCA GCT CTC CCC ATC CAC CTG TOT TAA CCG CCG ACC AAC GGG CTC CCG CTT TCT GCT

870
GAA TAA ACA CAT GTC CCC AAA

```

*To whom correspondence should be addressed

1. Dolan, K.P., Unterman, R., McLaughlin, M., Nakhasi, H.L., Lynch, K.R., and Feigelson, P. (1982) *J. Biol. Chem.* 257, 13527-13534