

## The primary structure of rat ribosomal protein S14

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

Department of Biochemistry and Molecular Biology, The University of Chicago, 920 East 58th Street, Chicago, IL 60637, USA

Submitted October 10, 1989

EMBL accession no. X15040

The amino acid sequence of rat ribosomal protein S14 was deduced from the sequence of nucleotides in a recombinant cDNA (pS14-9) which was isolated from a library using an homologous human S14 cDNA (1) as a probe. The open reading frame has 456 nucleotides and encodes a protein containing 151 amino acids and having a molecular weight of 16,248. Rat S14 has 30 basic residues (20%) and 16 acidic ones (11%). There is only one conservative change in the amino acid sequence (asp to glu at position 8) between rat and human S14; there are 37 nucleotide differences (7.6%) of which 34 are in the third position of codons. Rat S14 is related to the following ribosomal proteins (the numbers in parentheses are the RELATE scores in S.D. units): human S14 (51.9); Chinese hamster S14 (63.9); *Drosophila melanogaster* S14 (46.5); yeast rp59 (41.9); *Escherichia coli* S11 (11.1); a number of chloroplast S11 proteins (~8.0); *Bacillus stearothermophilus* S11 (9.2); *Halobacterium marismortui* S19 (21.1). Only the last has not been reported before. (Supported by NIH grant GM 21769).

```

-30          +1          30
CCT TCC GGT GGA GGA GTC TGG AGA CGA CGT TCA GAA ATG GCA CCT CCG AAG GGG AAG GAC AAG AAG GAA GAA CAG GTC ATC AGC CTT GGA
MET ALA PRO ARG LYS GLY LYS ASP LYS LYS LYS LYS GLU GLU GLN VAL ILE SER LEU GLY
1
60          90          120
CCT CAG GTG GCT GAA GGA GAG AAT GTA TTT GGT GTC TGC CAC ATC TTT GCA TCC TTC AAT GAT ACC TTT GTC CAT GTT ACT GAT CTT TCT
PRO GLN VAL ALA GLU GLY GLU ASN VAL PHE GLY VAL CYS HIS ILE PHE ALA SER PHE ASN ASP THR PHE VAL HIS VAL THR ASP LEU SER
20          30          40
150          180          210
GGC AAG GAA ACC ATT TGC CGA GTA ACT GGT GGA ATG AAG GTG AAG GCT GAC CGA GAT GAG TCC TCT CCG TAT GCA GGC ATG TTG GCT GGC
GLY LYS GLU THR ILE CYS ARG VAL THR GLY MET LYS VAL LYS ALA ASP ARG ASP GLU SER SER PRO TYR ALA ALA MET LEU ALA ALA
50          60          70
240          270          300
CAG GAT GTG GCC CAG AGG TGC AAG GAA CTG GGC ATC ACT GGC CTG CAT ATC AAA CTC CGG GCC ACA GGA GGA AAC AGG ACC AAG ACC CCT
GLN ASP VAL ALA GLN ARG CYS LYS GLU LEU GLY ILE THR ALA LEU HIS ILE LYS LEU ARG ALA THR GLY GLY ASN ARG THR LYS THR PRO
80          90          100
330          360          390
GGA CCT GGA GCC CAG TCA GCC CTC AGA GCT CTT GCT GCG TCT GGG ATG AAG ATT GGG CGG ATT GAG GAT GTC ACC CCG ATC CCG TCT GAC
GLY PRO GLY ALA GLN SER ALA LEU ARG ALA LEU ALA ARG SER GLY MET LYS ILE GLY ARG ILE GLU ASP VAL THR PRO ILE PRO SER ASP
110          120          130
420          450
AGC ACT CGA AGG AAG GGT GGT CGT CGG GGT CGC CGT CTG TGA
SER THR ARG ARG LYS GLY GLY ARG ARG GLY ARG ARG LEU END
140          150

```

\*To whom correspondence should be addressed

## REFERENCES

1. Rhoads, D.D., Dixit, A., and Roufa, D.J. (1986) Mol. Cell. Biol. 6, 2774-2783