Volume 16 Number 23 1988 Nucleic Acids Research

Nucleotide sequence of a cDNA encoding the rat p53 nuclear oncoprotein

Thierry Soussi*, Claude Caron de Fromentel, Claude Breugnot and Evelyne May

Unité d'Oncologie Moléculaire, IRSC, 7 rue Guy Mocquet, 94802 Villejuif Cedex, France

Submitted October 10, 1988 Accession no. X13058

We have isolated clones which contain rat p53 cDNA from a rat Py T21 cell cDNA library (1). Three cDNA clones hybridize with a probe corresponding to the mouse p53 cDNA (2) and all of them were shown to derive from a single mRNA species. The cDNA contains an open reading frame encoding a polypeptide of 390 amino acids. The predicted amino acid sequence reveales 88% homology with mouse p53 (2), 78% with human p53 (3) and 50 % with <u>Xenopus laevis</u> p53 (4). The 5' non coding region of the cDNA is not complete but overlaps perfectly with the corresponding region deduced from the genomic sequence of the first exon (5). The 3' untranslated region is 82% homologous with the same region of the mouse p53 cDNA.

```
GGGTCTGCTG GGATTGGGAC TTTCCCCTCC CACGTGCTCA CACTGGCTAA AGTTCTGAAG CTCCAGTTCA TTGGGACTTT
CCTTGCTATA GGTAGCGACT ACAGTTAGGG GGTACCTGGC ATCCGGGCCC GCACCCTACT TCCCAGCAGG GTGTCACGCT
CCCCTGAAGACTGG
::::::::::::::
cccctgaagactggataactgtc ATG GAG GAT TCA CAG TCG GAT ATG AGC ATC GAG CTC CCT CTG AGT CAG GAG    74
                         M   E   D   S   Q   S   D   M   S   I   E   L   P   L   S   Q   E
ACA TTT TCA TGC TTA TGG AAA CTT CTT CCT CCA GAT GAT ATT CTG CCC ACC ACA GCG ACA GGG TCA CCT   143
 T   F   S   C   L   W   K   L   L   P   P   D   D   I   L   P   T   T   A   T   G   S   P
AAT TCC ATG GAA GAT CTG TTC CTG CCC CAG GAT GTT GCA GAG TTG TTA GAA GGC CCA GAG GAA GCC CTC   212
 N   S   M   E   D   L   F   L   P   Q   D   V   A   E   L   L   E   G   P   E   E   A   L
CAA GTG TCA GCT CCT GCA GCA CAG GAA CCT GGA ACT GAG GCC CCT GCA CCC GTG GCC CCT GCT TCA GCT   281
 Q   V   S   A   P   A   A   Q   E   P   G   T   E   A   P   A   P   V   A   P   A   S   A
ACA CCG TGG CCT CTG TCA TCT TCC GTC CCT TCT CAA AAA ACT TAC CAA GGC AAC TAT GGC TTC CAC CTG   350
 T   P   W   P   L   S   S   S   V   P   S   Q   K   T   Y   Q   G   N   Y   G   F   H   L
GGC TTC CTG CAG TCA GGG ACA GCC AAG TCT GTT ATG TGC ACG TAC TCA ATT TCC CTC AAT AAG CTG TTC   419
 G   F   L   Q   S   G   T   A   K   S   V   M   C   T   Y   S   I   S   L   N   K   L   F
TGC CAG CTG GCG AAG ACA TGC CCT GTG CAG TTG TGG GTC ACC TCC ACA CCT CCA CCT GGT ACC CGT GTC   488
 C   Q   L   A   K   T   C   P   V   Q   L   W   V   T   S   T   P   P   P   G   T   R   V
CGT GCC ATG GCC ATC TAC AAG AAG TCA CAA CAC ATG ACT GAG GTC GTG AGA CGC TGC CCC CAC CAT GAG   557
 R   A   M   A   I   Y   K   K   S   Q   H   M   T   E   V   V   R   R   C   P   H   H   E
CGT TGC TCT GAT GGT GAC GGC CTG GCT CCT CCC CAA CAT CTT ATC CGG GTG GAA GGA AAT CCG TAT GCT   626
 R   C   S   D   G   D   G   L   A   P   P   Q   H   L   I   R   V   E   G   N   P   Y   A
GAG TAT CTG GAC GAC AGG CAG ACT TTT CGG CAC AGC GTG GTG GTA CCG TAT GAG CCA CCT GAG GTC GGC   695
 E   Y   L   D   D   R   Q   T   F   R   H   S   V   V   V   P   Y   E   P   P   E   V   G
TCC GAC TAT ACC ACT ATC CAC TAC AAG TAC ATG TGC AAC AGC TCC TGC ATG GGG GGC ATG AAC CGC CGG   764
 S   D   Y   T   T   I   H   Y   K   Y   M   C   N   S   S   C   M   G   G   M   N   R   R
CCC ATC CTT ACC ATC ATC ACG CTG GAA GAC TCC AGT GGG AAT CTT CTG GGA CGG GAC AGC TTT GAG GTT   833
 P   I   L   T   I   I   T   L   E   D   S   S   G   N   L   L   G   R   D   S   F   E   V
CGT GTT TGT GCC TGT CCT GGG AGA GAC CGT CGG ACA GAG GAA GAA AAT TTC CGC AAA AAA GAA GAG CAT   902
 R   V   C   A   C   P   G   R   D   R   R   T   E   E   E   N   F   R   K   K   E   E   H
TGC CCG GAG CTG CCC CCA GGG AGT GCA AAG AGA GCA CTG CCC ACC AGC ACA AGC TCC TCT CCC CAG CAA   971
 C   P   E   L   P   P   G   S   A   K   R   A   L   P   T   S   T   S   S   S   P   Q   Q
AAG AAA AAA CCA CTC GAT GGA GAA TAT TTC ACC CTT AAG ATC CGT GGG CGT GAG CGC TTC GAG ATG TTC  1040
 K   K   K   P   L   D   G   E   Y   F   T   L   K   I   R   G   R   E   R   F   E   M   F
CGA GAG CTG AAT GAG GCC TTG GAA TTA AAG GAT GCC CGT GCT GCC GAG GAG TCA GGA GAC AGC AGG GCT  1109
 R   E   L   N   E   A   L   E   L   K   D   A   R   A   A   E   E   S   G   D   S   R   A
CAC TCC AGC TAC CCG AAG ACC AAG AAG GGC CAG TCT ACG TCC CGC CAT AAA AAA CCA ATG ATC AAG AAA  1178
 H   S   S   Y   P   K   T   K   K   G   Q   S   T   S   R   H   K   K   P   M   I   K   K
GTG GGG CCT GAC TCA GAC TGAcagcctctgcatcctgtccccatcaccagcctccccgtcccctcctttcttgccatttatgact 1264
 V   G   P   D   S   D   *
ttagggcttgttatgagagctgacaagacaatgctagtcccttcactgcctttttttaccttgtagatagtactcggccccctctatgcaaa 1356
ctggttcctggcccagattggggaatgggttggtagttgctgggtctctgctggtccagcgaaatcctatccggtcagttgttggacctggc 1448
acctacagtgaaatttcaccccaccccaccgcctgtaagattctatcttgggccctcatacgatctgtatcctccaggacccatttcctcca 1540
ctctgcaaagcctgtctgcatttatccatcccccacccctctccctctttttatctcttttatatatccaatttcttatttacaa        1627
```

Fig: Bold DNA sequence corresponds to the first exon deduced from genomic sequence (5).

*To whom correspondence should be addressed

References

1 - Matrisian, L. M., Glaichenhaus, N., Gesnel, M. C. and Breathnach, R. (1985). EMBO J., 4, 1435-1440.

2 - Jenkins, J., Rudge, K., Redmonds, S. and Wade-Evans A. (1984) Nucleic Acids Res., 12, 5609-5626.

3 - Zakut-Houri, R., Bienz-Tadmor, B., Givol, D. and Oren, M. (1985). EMBO J., 4, 1251-1255.

4 - Soussi, T., Caron de Fromentel, C., Méchali, M., May, P. and Kress M. (1987). Oncogene, 1, 71-78.

5 - Bienz-Tadmor, B., Zakut-Houri, R., Libresço, S., Givol, D. and Oren M. (1985). EMBO J., 4, 3209-3213