Nucleotide sequence of a cDNA encoding the monkey cellular phosphoprotein p53

Pascal Rigaudy and Walter Eckhart

Molecular Biology and Virology Laboratory, The Salk Institute, PO Box 85800, San Diego, CA 92138-9216, USA

Submitted September 5, 1989

EMBL accession no. X16384

We report here the nucleotide sequence of a monkey p53 cDNA isolated from BSC-1 African green monkey epithelial cells. A BSC-1 Okayama-Berg library (1) was screened with a human p53 cDNA (2). The longest clone isolated contained an open reading frame of 379 amino acids homologous to human p53. However this cDNA lacked an apparent initiation codon. The 5' end of the monkey p53 cDNA was then cloned from BSC-1 RNA using reverse transcription combined with polymerase chain reaction (PCR). A monkey specific oligonucleotide was used to prime reverse transcription reactions and then to amplify the cDNAs by PCR, in combination with an oligonucleotide homologous to the human p53 5' untranslated region (2). The full length sequence of the monkey p53 cDNA is 1400 base pairs and contains a single open reading frame which encodes a protein of 393 amino acids. The predicted amino acid sequence is 95% identical to the human protein.

```

GGGGCTCCGGGGGACACTTGGCGTCCGGGCTGGAAGCGTGCTTTCCAAGACGGTGACACGCTTCCTTGAGGAT 72
      M E E P Q S D P S I E P P L S Q E T F
TGGCAGCAGACTGCTTTAGGGTCACTGCAATGGAGGAGCGCGAGTCAGATCCAGCATOGAGCGGCTCTGAGTCAGGAACATT 159
S D L W K L L P E N N V L S P L P S Q A V D D L M L S
TCAGACCTATGGAACACTTCTCTGAAAACAAAGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
P D D L A Q W L T E D P G P D E A P R M S E A A P H M
COGGATGATCTTCACAATGGTTAACTGAAGACCCAGGTCCAGATGAAGCTCCAGATGTCAGAGGCTGCTGCTGCTGCTGCT 321
A P T P A A P T P A A P A P A P A P S W P L S S S V P S Q
GCCCCACACAGCAGCTCTACACGGGGGCGGCTGCAACAGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 402
K T Y H G S Y G F R L G F L H S G T A K S V T C T Y S
AAAACTACCAAGCAGCTACGGTTTGGCTCTGGGCTCTGCTGCAATCTGGAACAGCAAGCTCTGCTGCTGCTGCTGCTGCTGCT 483
P D L N K M F C Q L A K T C P V Q L W V D S T P P P G
CTGACTCAACAGATGTTTGGCAGCTGGCGAAGCACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
S R V R A M A I Y K Q S Q H M T E V V R R C P H H E R
AGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
C S D S D G L A P P Q H L I R V E G N L R V E Y S D D
TGCTCAGACAGGATGGCTGCGGCTCTGCTGCAATCTTATCGAGTGAAGGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726
R N T F R H S V V V P Y E P P E V G S D C T T I H Y N
AGAAACACTTTTGGACATAGTGTGGTGGTGGCTATGAGCGGCTGAGGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 807
Y M C N S S C M G G M N R R P I L T I I T L E D S S G
TACATGTGTAAACGTTCTGCTGATGGGGGGCATGAACCGGAGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 888
N L L G R N S F E V R V C A C P G R D R R T E E E N F
AATCTACTGGGAGGAGCACTTTGAGGTGGAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
R K K G E P C H E L P P G S T K R A L P N N T S S S P
CGCAAGAAAGGGGAGCTTTGCAACAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCT 1050
Q P K K K P L D G E Y F T L Q I R G R E R F E M F R E
CAGCAAGAAAGAAACCACTGGATGGAGATATTTCAACCTTCAGATGCGGGGGGCTGAGCGCTTGGAGATGTTTGGAGAG 1131
L N E A L E L K D A Q A G K E P A G S R A H S S H L K
CTGATGAGGCTTGGACTCAAGGATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 1212
S K K G Q S T S R H K K F M F K T E G P D S D *
TCCAAGAGGGGCAATCTCACTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCT 1296
CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1398

```

References:

- 1-Hanks, S.K., Armour, R., Baldwin, J.H., Maldonado, F., Spiess, J., and Holley, R.W. (1988) *Proc.Natl.Acad.Sci.USA*, **85**, 79-82.
- 2-Harlow, E., Williamson, N.M., Ralston, R., Helfman, D.M., and Adams, T.E. (1985) *Mol.Cell.Biol.*, **5**, 1601-1610.