Nucleotide sequence of a cDNA coding for the small subunit of human calcium-dependent protease

Shigeo Ohno, Yasufumi Emori and Koichi Suzuki

Department of Molecular Biology, The Tokyo Metropolitan Institute of Medical Science, Honkomagome, Bunkyo-ku, Tokyo 113, Japan

Submitted 10 June 1986

A near-full-length cDNA coding for the small subunit of human calcium-dependent protease was isolated from a human spleen cDNA library. A comparison of the encoded amino acid sequence with those reported for rabbit and porcine[1,2] revealed only 3% differences between any two combinations of the three sequences. The human protein is composed of 268 amino acids whereas those of rabbit and porcine are composed of 266 amino acids. The differences are found in the lengths of the two glycine stretches(aa 10-20 and 37-56). Numbers of glycine residues are 11 and 20 for human, 10 and 19 for rabbit and 11 and 18 for porcine. Conservation of the amino acid sequence of the small subunit of calcium-dependent protease suggests the multiplicity of the target for this calcium binding protein. Homology of the nucleotide sequences between human and rabbit cDNAs are 73%, 92% and 61% for the 5' untranslated region, coding region and 3' untranslated region, respectively. A part of the 5' untranslated region(-125--52) is highly conserved(91%) in human and rabbit. This region may have some role in the gene expression.

```
   -158
   GCGGGCGACACGAGGGCCGCGGTGCAGTGTCCGACCCGAGAGTTGCGGCCTGAGTC
-102
ACCGGCCCCGCCCTCCGGAGCCGGACGCTCCGGGAGGCCCGGGAGCGGCAGTGGAACCGA
-42
CTCCCAGAACTCCGGACGTGTGCGGCGCAGTGAGTCGCAGCCATGTTCCTGGTTAACTCG
                                             M  F  L  V  N  S
19                                                          6
TTCTTGAAGGGCGGCGGCGGCGGCGGCGGGGGAGGCGGGGGCCTGGGTGGGGGCCTGGGA
F  L  K  G  G  G  G  G  G  G  G  G  G  G  L  G  G  G  L  G
79                                                         26
AATGTGCTTGGAGGCCTGATCAGCGGGGCCGGGGGCGGCGGCGGCGGCGGCGGCGGCGGC
N  V  L  G  G  L  I  S  G  A  G  G  G  G  G  G  G  G  G  G
139                                                        46
GGCGGTGGTGGAGGCGGCGGTGGCGGTGGAACGGCCATGCGCATCCTAGGCGGAGTCATC
G  G  G  G  G  G  G  G  G  G  T  A  M  R  I  L  G  G  V  I
199                                                        66
AGCGCCATCAGCGAGGCGGCTGCGCAGTACAACCCGGAGCCCCCGCCCCCACGCACACAT
S  A  I  S  E  A  A  A  Q  Y  N  P  E  P  P  P  P  R  T  H
259                                                        86
TACTCCAACATTGAGGCCAACGAGAGTGAGGAGGTCCGGCAGTTCCGGAGACTCTTTGCC
Y  S  N  I  E  A  N  E  S  E  E  V  R  Q  F  R  R  L  F  A
319                                                       106
CAGCTGGCTGGAGATGACATGGAGGTCAGCGCCACAGAACTCATGAACATTCTCAATAAG
Q  L  A  G  D  D  M  E  V  S  A  T  E  L  M  N  I  L  N  K
379                                                       126
GTTGTGACACGACACCCTGATCTGAAGACTGATGGTTTTGGCATTGACACATGTCGCAGC
V  V  T  R  H  P  D  L  K  T  D  G  F  G  I  D  T  C  R  S
439                                                       146
ATGGTGGCCGTGATGGATAGCGACACCACAGGCAAGCTGGGCTTTGAGGAATTCAAGTAC
M  V  A  V  M  D  S  D  T  T  G  K  L  G  F  E  E  F  K  Y
                                                          166

499
TTGTGGAACAACATCAAAAGGTGGCAGGCCATATACAAACAGTTCGACACTGACCGATCA
L  W  N  N  I  K  R  W  Q  A  I  Y  K  Q  F  D  T  D  R  S
559                                                       186
GGGACCATTTGCAGTAGTGAACTCCCAGGTGCCTTTGAGGCAGCAGGGTTCCACCTGAAT
G  T  I  C  S  S  E  L  P  G  A  F  E  A  A  G  F  H  L  N
619                                                       206
GAGCATCTCTATAACATGATCATCCGACGCTACTCAGATGAAAGTGGGAACATGGATTTT
E  H  L  Y  N  M  I  I  R  R  Y  S  D  E  S  G  N  M  D  F
679                                                       226
GACAACTTCATCAGCTGCTTGGTCAGGCTGGACGCCATGTTCCGTGCCTTCAAATCTCTT
D  N  F  I  S  C  L  V  R  L  D  A  M  F  R  A  F  K  S  L
739                                                       246
GACAAAGATGGCACTGGACAAATCCAGGTGAACATCCAGGAGTGGCTGCAGCTGACTATG
D  K  D  G  T  G  Q  I  Q  V  N  I  Q  E  W  L  Q  L  T  M
799                                                       266
TATTCCTGAACTGGAGCCCCAGACCCGCCCCCTCACCGCCTTGCTATAGGAGTCACCTGG
Y  S
859
AGCCTCGGTCTCTCCCAGGGCCGATCCTGTCTGCAGTCACATCTTTGTGGGGCCTGCTGA
919
CCCACAAGCTTTTGTTCTCTCAGTACTTGTTACCCAGCTTCTCAACATGCAGGGCCCAAT
979
TTGCCCTGCCTGGAGTTCCCCCTGGCTCTAGGACACTCTAACAAGCTCTGTCCACGGGTC
1039
TCCCCATTCCCACCAGGCCCTGCACACACCCACTCCGTAACCTCTCCCCTGTACCTGTGC
1099
CAAGCCTAGCACTTGTGATGCCTCCATGCCCCGAGGGCCTCTCTCAGTTCTGGGAGGATG
1159
ACTCCAGTCCCTGCACGCCCTGGCACACCCTTCACGGTTGCTACCCAGGCGGCCAAGCTC
1219
CAGACCGTGCCAGACCCAGGTGCCCCAGTGCCTTTGTCTATATTCTGCTCCCAGCCTGCC
1279
AGGCCAGGAGGAAATAAACATGCCCCAGTTGCTGATCTCTAA
```

1. Sakihama,T. et al. (1985) Proc.Natl.Acad.Sci.USA,82,6075-6079.
2. Emori,Y. et al. (1986) J.Biol.Chem., in press.

