Sequence of the mouse glandular kallikrein gene, mGK-5

Catherine C.Drinkwater and Robert I.Richards

The Howard Florey Institute of Experimental Physiology and Medicine, University of Melbourne, Parkville, Victoria 3052, Australia

Submitted October 22, 1987 Accession no.Y00500

The mouse glandular kallikrein gene, mGK-5, is expressed in salivary glands in a sexually dimorphic fashion (1). This is in contrast to the neighbouring gene, mGK-6, which encodes renal kallikrein, or kininogenase, and is expressed at equal levels in salivary glands of both sexes (1,2). Nucleotide sequence from the five exons and 607 bp of 5′ flanking DNA from mGK-5 was determined by the enzymatic method (3). mGK-5 and mGK-6 are the most closely related of the mouse glandular kallikrein genes, being 95% homologous. Exon-intron boundaries of mGK-5 (boxed) are identical to those of all other mouse glandular kallikrein genes (4). The gene also has the variant 5′-TTTAAA-3′ box typical to kallikrein promoter regions and a normal poly-adenylation signal (5, both underlined). By analogy with mGK-6 (2), the protein encoded by mGK-5 (numbers in italics, Ile-1 arrowed) has a hydrophobic signal sequence (*-24* to *-8*) followed by a zymogen peptide (*-7* to *-1*). The mature protein is 237 amino acids long and contains residues essential to an active serine protease. His-41, Asp-96 and Ser-189 form the charge-relay system (6), whilst Asp-183 is believed to direct cleavage at basic residues (7). mGK-5 expression was examined in several mouse tissues but only detected in salivary gland (not shown). Since mGK-5 and mGK-6 are regulated differently (1) their 5′ flanking regions were compared. No differences apart from single base insertions, deletions and substitutions were observed (not shown).

```
GATCCACCTGGCTCTGGCTTCAGTTCCTGGTACTGAAGATTTGCACCACTACACGCAGCTCGCAAACTACTCATAGGAACAATAAAATCAGTGATCCAGAGTTCCAAGTCCAAAAAGGAA  120
GCATTTTGGCATTATTTAAAAAAAAGGATTTGTATCTGCAGGGAGCTTGCAGCCCTCTTGGCTATCACTAGTGCAAAGAATTTGTTCTCAGTGCAACTTTCTGGGGATAATCTTGGACTT  240
GAGGTACAGGAACTTTCTTAGAACTGGGGGTTGGGAGTCACTTGCTCCTCCATCCTCCTTTAACACCAACTCACCATTGCCCCTGACCTTGCCCTCTGGCTAAGGTCCCCGGAACTCAAA  360
CAGCTCCAGAAACAGCAGGACCTGCCCCACCTAATCTCATCCCTGGGAAGGTGTCCAGGGTGACCTAGGGACTACTGCACAGCAGGTGCAACAGGGCCTCCTCTCCCACAGCCAGGAGGG  480
TGCAAGAAGGAAGTATAACTCTGTCAGAAAGGACAGAGCTTTGGTCAGCATCTGAGCGGCAACAGGACAGGGGAGGGGCTGTGGGAGAGAATGAGGGTTTTTAAAGTCTCCCTGAGGAGC  600
                                                                           MetTrpPheLeuIleLeuPheLeuAlaLeuSerLeuGlyGlyIleA
CTCAACAGCTCCAAGCTCACAGCCTGCAGCTCCTGGACAGCTGTCACCATGTGGTTCCTGATCCTGTTCCTAGCCCTGTCCCTAGGAGGGATTGGTGAGATGAGGGAAAGGGAGGGATCA  720
                                                             spAlaAlaProProValGlnSerArgIlePheGlyGlyPheAsnCysGluLysAsnSerGlnProTrpGlnValAlaValTy   19
C-intron A-CTGTCCTCCCCCATTGCCCCTATGCAGATGCTGCACCTCCAGTCCAATCACGAATTTTTGGAGGATTTAACTGTGAGAAGAACTCCCAACCCTGGCAAGTGGCTGTGTA  840
rArgPheThrLysTyrGlnCysGlyGlyValLeuLeuAsnAlaAsnTrpValLeuThrAlaAlaHisCysHisAsnAs                                                45
CCGCTTCACCAAATATCAATGCGGGGGAGTCCTGCTGAACGCCAACTGGGTTCTCACTGCTGCCCACTGCCATAATGAGTGAGTAAGGGTGGAAACAGGAAAGCA-intron B-CGCTC  960
                                      pLysTyrGlnValTrpLeuGlyLysAsnAsnPhePheGluAspGluProSerAlaGlnHisArgLeuValSerLysAlaIleProHisProAspPheA   78
TATCGCCTTCCATTGCCCCCAGCAAGTACCAGGTTTGGCTGGGCAAAAACAACTTTTTTGAGGATGAACCCTCTGCCCAACACCGGCTTGTCAGCAAAGCCATCCCTCACCCTGACTTCA  1080
snMetSerLeuLeuAsnGluHisThrProGlnProGluAspAspTyrSerAsnAspLeuMetLeuLeuArgLeuLysLysProAlaAspIleThrAspValValLysProIleAspLeuP   118
ACATGAGCCTCCTGAATGAGCACACCCCACAACCTGAGGACGACTACAGCAATGACCTGATGCTGCTCCGCCTCAAAAAGCCTGCTGACATCACAGATGTTGTGAAGCCCATCGACCTGC  1200
roThrGluGluProLysLeuGlySerThrCysLeuAlaSerGlyTrpGlySerIleThrProValIleT                                                         141
CCACTGAGGAGCCCAAGCTGGGGAGCACATGCCTAGCCTCAGGCTGGGGCAGCATTACACCCGTCATATGTGGTCTTCTCCAAGCAACAGAGGGTGGCCGTGAGGGGCACAGGAGACTG  1320
                                                                yrGluProAlaAspAspLeuGlnCysValAsnPheLysLeuLeuProAsnGluAspCysValLysAlaHisIleGluL   167
AGTTGACCTCTGCTCACCACACTCTTGCCTCCTGATTCATAGACGAACCCGCAGATGATCTCCAGTGTGTGAACTTCAAGCTCCTGCCTAATGAGGACTGTGTCAAAGCCCACATAGAGA  1440
ysValThrAspValMetLeuCysAlaGlyAspMetAspGlyGlyLysAspThrCysMet                                                                   186
AGGTGACAGATGTTATGTTGTGTGCAGGAGATATGGATGGAGGCAAAGACACTTGTATGGTGAGAAGCCCCTTCTTGCAGTGAGGT-intron D-CTTACTCCTGTACCTGTCCTCTTT  1560
   GlyAspSerGlyGlyProLeuIleCysAspGlyValLeuHisGlyIleThrSerTrpGlyProSerProCysGlyLysProAsnValProGlyIleTyrThrLysLeuIleLysPhe   225
CAGGGTGACTCAGGAGGCCCACTGATCTGTGATGGTGTTCTCCACGGTATCACATCATGGGGCCCTAGCCCTTGCGGTAAACCCAATGTGCCAGGTATCTACACCAAACTTATTAAGTTT  1680
AsnSerTrpIleLysAspThrIleAlaLysAsnAla***                                                                                       237
AACTCTTGGATAAAAGACACTATAGCTAAAAATGCCTGAATATCACATTGTCCCATGTTCTCAATAAAGCCCACCATGCAGCAAATCGAGTTCAAGTTCTGACATTCTCTGGTCTTCTGG  1800
TGTGCAGTACCTGTCATGTTCACTCAGCTAAGGTTAGGACTTCCAAAGTTGATCAGCAAGGGACAGGAGTACATGTACTGTTGGCTACAGAGGTGGGTTGAGACAAATCGAGACAAAAGG  1920
CTCAGACCATGTCAGGGATGCCACTTGAGTCTCCTCTGGCAAGGAAAATGTTCAGATC
```

REFERENCES

(1) van Leeuwen, B.H. *et al* (1987) *EMBO J.*, *6*, 1705-1713 (2) van Leeuwen, B.H. *et al* (1986) *J. Biol. Chem.*, *261*, 5529-5535 (3) Sanger, F. *et al* (1977) *Proc. Natl. Acad. Sci. USA*, *74*, 5463-546 (4) Evans, B.A. *et al* (1987) *J. Biol. Chem 262*, 8027-8034 (5) Proudfoot, N.J. & Brownlee, G.G. (1976) *Nature*, *263*, 211-214 (6) Blow, D.M. *et al* (1969) *Nature*, *221*, 337-340 (7) Krieger, M. *et al* (1974) *J. Mol. Biol.*, *83*, 209-230

