Volume 16 Number 22 1988 Nucleic Acids Research

Sequence of mGK-11, a mouse glandular kallikrein gene

Catherine C.Drinkwater and Robert I.Richards

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

Submitted October 24, 1988 Accession nos X13215–X13218 (incl.)

Like most members of the mouse glandular kallikrein gene family (for review see 1), expression of mGK-11 shows sexual dimorphism in the salivary glands, the level in males being about 30-fold greater than that in females. Nucleotide sequence from the five exons and 625 bp of the 5' flanking region was determined using the enzymatic method[2]. The exon-intron boundaries of mGK-11 (boxed) are identical to those of all other mouse glandular kallikrein genes (see 1, 3 and references therein). The gene also has the variant 5'-TTTAAA-3' box typical to kallikrein promoter regions and a normal polyadenylation signal (both underlined). By analogy with known glandular kallikreins, the mGK-11 gene product has a hydrophobic signal sequence (*-24* to *-8*, numbering relative to Ile-1 of the mature protein) followed by a seven residue zymogen peptide. The mature protein is 237 amino acids long and contains the three residues believed necessary for serine protease activity (His-41, Asp-96 and Ser-189) as well as Asp-183 which is thought to direct cleavage at basic residues, giving trypsin-like specificity. Expression of mGK-11 was examined in several mouse tissues by Northern blot analysis, but could only be detected in salivary glands (not shown).

```
TTTAAAATCACACAGATCCACCTGGCTCTGGCTTCAAAATGTGGTATTAAAGGTGTGAACCACTACACCCAGCTCCCAAACTTCTCATAGAAACAATAAAATCAGTGACACAGAGTTCAA   120
AGTGCAAAAAGGCAAGAACTTTGGCTTTCTCAAAAAAATTAATTGTATATACAGTGAACTGGCAGCCGTCTCATCTATCACTGGTGAAAAAAATCTTGTTCTAACTGCTACTTTCTTGGG   240
ATAATCTTGGACTTGAGGTACAGGAACTTCCTTAGAACTGGGGGTTGGGAGTCACTTGCTCCTCCATCCTCCTTTAACACCAACTCACCATTGTCCCTGACCTTGCCATCTGGCTAAGGT   360
CCCCGGAACTCAAACAGCTCCAGGAACAACAGCACTGCCCAACCCAGTTTCATCCCTAGGAAGGTATCCAGGGTGAACTAGGGCCTACTGTACAGCGGGCGCAACGAGGCCCTCCAATTC   480
CACAGCCAGGAGGGTGCAAGCAGGGAGTCCACCTCTGTCAGCAAGGGCAGGGGCTTAGTCAGCATCTGGGTGGCAACAGGACAGGGGAGGGGCTGTGGGGAGAATGAGTGTATTTAAAAT   600
                                                                                         -20
                                                                               MetTrpPheLeuIleLeuPheLeuAlaLeuSerLeuGlyGlyIleA
GTCCTGAAGAGCCTCAAAAGCTCCAAGCTCACCTCCTGCAGCTCCTGGACACCTGTTACCATGTGGTTCCTGATCCTGTTCCTAGCCCTCTCCCTAGGAGGATTG[GT]GAGATGAGGGAA   720

AGGGAGGGGTCACAGCCCTAACTCTCCTGCTGCCTTTGACCC---1.8kb---GATCCCAGCAGTGCTTAACTGTGTCCATCAAGGGTTGACCTCCCTACATGTGCCAACTCTCCCCTCC   840
                           *                1                                                                20
                     spAlaAlaProProValGlnSerArgIleValGlyGlyPheAsnCysGluLysAsnSerGlnProTrpHisValAlaValTyrArgTyrAsnLysTyrIle
CCCATTGCTCCTGTTGC[AG]ATGCTGCACCTCCTGTCCAGTCTCGAATTGTTGGAGGATTTAACTGTGAGAAGAATTCCCAACCCTGGCATGTGGCTGTGTACCGCTACAACAAATATATA   960
                                          40
CysGlyGlyValLeuLeuAspArgAsnTrpValLeuThrAlaAlaHisCysHisValSe
TGCGGGGGAGTCCTGTTGGACCGTAACTGGCTTCTCACAGCTGCCCACTGCCATGTCAG[GT]AAGTAAGGGTGGAGACAGGAAAGCAGGGTGGAAGCCAGAGATC---0.7kb---CCCCC  1080
                                                                                                              60
                                                                          rGlnTyrAsnValTrpLeuGlyLysThrLysLeuPheGlnArgGluProS
TCTTCTGACTCTGACCACTCACCCTGATTCTCTCAATCCCACTCCCTCCATCTCCTGCCTTGTCCCCT[AG]CCAGTATAATGTTTGGCTGGGCAAAACAAAGCTATTCCAACGTGAACCCT  1200
                                                   80                                                     100
erAlaGlnHisArgMetValSerLysSerPheProHisProAspTyrAsnMetSerLeuLeuIleIleHisAsnProGluProGluAspAspGluSerAsnAspLeuMetLeuLeuArgL
CTGCTCAGCACCGAATGGTCAGCAAAAGCTTCCCCCACCCTGACTACAACATGAGCCTCCTCATAATCCACAACCCAGAACCTGAGGACGATGAGAGCAATGACCTGATGTTACTACGCC  1320
                                                  120                                                     140
euSerGluProAlaAspIleThrAspAlaValLysProIleAlaLeuProThrGluGluProLysLeuGlySerThrCysLeuValSerGlyTrpGlySerIleThrProThrLysP
TCAGCGAGCCAGCTGACATCACAGATGCTGTGAAGCCCATCGCCCTGCCCACTGAGGAGCCCAAGCTGGGGAGCACATGCCTAGTCTCAGGCTGGGGCAGCATTACACCCACGAAAT[GT]G  1440

                                                                                            heGlnThrProAspAspLeuGlnCys
AGTCTTCTCCAAGCAACACAGCTCGGTGTGGGGGAGGGACAGAGGGGTCTGGATTTTGGTCCTGCTCACCACCCACTTTTCCCCTGCTCCAC[AG]TCCAAACCCCAGATGATCTCCAGTGT  1560
                               160                                                        180
ValSerIleLysLeuLeuProAsnGluValCysValLysAsnHisAsnGlnLysValThrAspValMetLeuCysAlaGlyGluMetGlyGlyGlyLysAspThrCysLys
GTGTCCATCAAGCTCCTGCCTAATGAGGTCTGTGTCAAAAACCACAATCAAAAGGTGACAGATGTCATGTTGTGTGCAGGAGAGATGGGTGGAGGCAAAGATACTTGCAAG[GT]GAGACAG  1680

                                                                                                     GlyAspSerGlyGly
ACCCTCCCTACAGTGAGGTGAAGGGCTGAAAGAGAGAACTGAGG---0.3kb---GAATCCAACTGCCAGGTGCTTTCTTGCTCCTGACCCTGTCCTTCTTTC[AG]GGTGACTCAGGAGGC  1800
                         200                                                       220
ProLeuIleCysAspGlyValLeuHisGlyIleThrAlaTrpGlyProIleProCysGlyLysProAsnThrProGlyValTyrThrLysLeuIleLysPheThrAsnTrpIleLysAsp
CCACTGATCTGTGATGGTGTTCTCCATGGTATCACAGCATGGGGCCCTATCCCATGTGGCAAACCCAATACGCCGGGAGTCTACACAAAACTTATTAAGTTTACCAACTGGATAAAAGAC  1920

ThrMetAlaLysAsnPro***                                                                                                    237
ACTATGGCAAAAAACCCCTGAGTGTCACATTATCTGCTGTTCTCAATAAAATACACCATGAAACAAATGAGTTCAAGGTCTGACATTCTCTGTCCTACTGGAGTGCAGGACCAAATCAAG  2040
```

References: (1) Drinkwater, C.C., Evans, B.A. & Richards, R.I. (1988) *Trends Biochem. Sci.*, 13, 169-172. (2) Sanger, F., Nicklen, S. & Coulsen, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, 74, 5463-5467. (3) Evans, B.A., Drinkwater, C.C. & Richards, R.I. (1987) *J. Biol. Chem.*, 262, 8027-8034.