4602 *Nucleic Acids Research, Vol. 18, No. 15*

Complete primary structure of lamb preprochymosin deduced from cDNA

Jože Pungerčar, Borut Štrukelj, Franc Gubenšek, Vito Turk and Igor Kregar
Department of Biochemistry, Jočef Stefan Institute, Jamova 39, 61111 Ljubljana, Yugoslavia

Submitted June 27, 1990 EMBL accession no. X53037

A cDNA coding for lamb preprochymosin was isolated from an abomasum cDNA library successively screened with two oligonucleotide probes corresponding to the 5′- and 3′-ends of the coding region for calf prochymosin (1−3). Similar to the situation for the calf enzyme, a 16 amino acid residue signal peptide (underlined) is followed by 42 residues of the proenzyme region. The mature lamb chymosin begins with glycine at position 59. The coding nucleotide and deduced amino acid sequences of lamb preprochymosin show about 95% and 94% similarity to calf preprochymosin, respectively. Both aspartic acid residues in the active site at positions 92 and 274 are conserved. Lamb chymosin described here can be denoted as the B form because of the presence of glycine at position 302 (*cf.* 4).

ACKNOWLEDGEMENTS

We thank Roman Jerala for synthesis of the C-terminal probe. This work was supported by the grants of the Research Community of Slovenia, Federal Secretariat for Development and Krka Pharmaceutical Works.

REFERENCES

1. Harris,T.J.R., Lowe,P.A., Lyons,A., Thomas,P.G., Eaton,M.A.W., Millican,T.A., Patel,T.P., Bose,C.C., Carey,N.H. and Doel,M.T. (1982) *Nucl. Acids Res.* **10**, 2177−2187.
2. Moir,D., Mao,J.-I., Schumm,J.W., Vovis,G.F., Alford,B.L. and Taunton-Rigby,A. (1982) *Gene* **19**, 127−138.
3. Nishimori,K., Kawaguchi,Y., Hidaka,M., Uozumi,T. and Beppu,T. (1982) *J. Biochem.* **91**, 1085−1088.
4. Foltmann,B., Pedersen,V.B., Kauffman,D. and Wybrandt,G. (1979) *J. Biol. Chem.* **254**, 8447−8456.

```
                                                                                  CCCAGATCCAAG    -1

ATGAGGTGTCTTGTGGTGCTACTTGCTGTCTTTGCTCTCTCCCAGGGCGCTGAGATCACCAGGATCCCACTGTACAAAGGCAAGCCTCTG    90
 M  R  C  L  V  V  L  L  A  V  F  A  L  S  Q  G  A  E  I  T  R  I  P  L  Y  K  G  K  P  L     30

AGGAAGGCACTGAAGGAGCGTGGGCTTCTGGAGGACTTCCTGCAGAAACAGCAATATGGCGTCAGCAGCGAGTACTCCGGCTTTGGGGAG   180
 R  K  A  L  K  E  R  G  L  L  E  D  F  L  Q  K  Q  Q  Y  G  V  S  S  E  Y  S  G  F  G  E     60

GTGGCCAGTGTGCCCTTGACCAATTACCTGGATAGTCAGTACTTTGGGAAGATCTACCTCGGGACCCCGCCCCAGGAGTTCACCGTGCTG   270
 V  A  S  V  P  L  T  N  Y  L  D  S  Q  Y  F  G  K  I  Y  L  G  T  P  P  Q  E  F  T  V  L     90

TTTGACACCGGCTCCTCTGACTTCTGGGTACCCTCTATCTACTGCAAGAGCAATGCCTGCAAAAACCACCAGCGCTTCGACCCAAGAAAG   360
 F  D  T  G  S  S  D  F  W  V  P  S  I  Y  C  K  S  N  A  C  K  N  H  Q  R  F  D  P  R  K    120

TCGTCCACCTTCCAGAACCTGGGCAAGCCCCTGTCTATCCGCTATGGGACGGGCAGCATGCAGGGCATCCTGGGCTACGACACCGTCACT   450
 S  S  T  F  Q  N  L  G  K  P  L  S  I  R  Y  G  T  G  S  M  Q  G  I  L  G  Y  D  T  V  T    150

GTCTCCAACATTGTGGACATCCAGCAGACAGTAGGCCTGAGCACCCAGGAGCCTGGGGATGTCTTCACCTATGCCGAGTTCGACGGGATC   540
 V  S  N  I  V  D  I  Q  Q  T  V  G  L  S  T  Q  E  P  G  D  V  F  T  Y  A  E  F  D  G  I    180

CTGGGGATGGCCTACCCCTCGCTCGCCTCAGAGTACTCGGTGCCCGTGTTTGACAACATGATGGACAGGCGCCTGGTGGCCCAGGACCTG   630
 L  G  M  A  Y  P  S  L  A  S  E  Y  S  V  P  V  F  D  N  M  M  D  R  R  L  V  A  Q  D  L    210

TTCTCGGTTTACATGGACAGGAGTGGCCAGGGGAGCATGCTCACACTGGGGGCCATCGACCCGTCCTACTACACAGGGTCCCTCCACTGG   720
 F  S  V  Y  M  D  R  S  G  Q  G  S  M  L  T  L  G  A  I  D  P  S  Y  Y  T  G  S  L  H  W    240

GTGCCCGTGACGCTGCAGAAGTACTGGCAGTTCACCGTGGACAGTGTCACCATCAGCGGTGCGGTTGTGGCCTGTGAGGGTGGCTGTCAG   810
 V  P  V  T  L  Q  K  Y  W  Q  F  T  V  D  S  V  T  I  S  G  A  V  V  A  C  E  G  G  C  Q    270

GCCATCCTGGACACGGGCACCTCCAAGCTGGTCGGGCCCAGCAGCGACATCCTCAACATCCAGCAGGCCATTGGAGCCACACAGAACCAG   900
 A  I  L  D  T  G  T  S  K  L  V  G  P  S  S  D  I  L  N  I  Q  Q  A  I  G  A  T  Q  N  Q    300

TATGGCGAGTTTGACATCGACTGCGACAGCTTGAGCAGCATGCCCACTGTGGTCTTTGAGATCAATGGCAAAATGTACCCACTGACCCCC   990
 Y  G  E  F  D  I  D  C  D  S  L  S  S  M  P  T  V  V  F  E  I  N  G  K  M  Y  P  L  T  P    330

TACGCCTATACCAGCCAGGAGGAGGGCTTCTGCACCAGTGGCTTCCAGGGTGAAAATCATTCCCATCAATGGATCCTGGGGGATGTTTTC  1080
 Y  A  Y  T  S  Q  E  E  G  F  C  T  S  G  F  Q  G  E  N  H  S  H  Q  W  I  L  G  D  V  F    360

ATCCGAGAGTATTACAGCGTCTTTGACAGGGCCAACAACCTCGTGGGGCTGGCCAAAGCCATCTGATCACATCGCTGACCAAGAACCTCA  1170
 I  R  E  Y  Y  S  V  F  D  R  A  N  N  L  V  G  L  A  K  A  I  *                            381

CTGTCCCCACACACCTGCACATACACACGCACACGTGTACGTGAGCACACGTGTGCACACACAGATGAGGTTTCCAGACAGATGATTCTC  1260

AATAAATGTTGTCTTTCTGCAAAAAAAAAAAAAAAAAAAAAAAAA                                             1305
```