Evidence for the existence of a mutated enkephalin sequence in preproenkephalin

Amos Bairoch

Biochimie Médicale, Centre Médical Universitaire, Rue Lombard 21, 1211 Genève 4, Switzerland

Received 10 May 1982

Preproenkephalin *Nucleotide sequence* *Molecular evolution* *Computer simulation* *Precursor processing*

A number of papers have recently appeared presenting the mRNA sequence for bovine [1,2] and human [3] preproenkephalin. The mRNA codes for a precursor protein containing four copies of Met-enkephalin, one of Leu-enkephalin and one copy each of two different carboxyl extended Met-enkephalins (Met-enkephalin-Arg^6-Phe^7 and Met-enkephalin-Arg^6-Gly^7-Leu^8). These seven regions of the sequence are all flanked on both side by paired basic residues (except at the carboxyl terminus of preproenkephalin). The relative positions of the enkephalin peptides in the bovine precursor sequence are shown in fig.1.

There is a rather large region coding for no visible enkephalin-like sequence between residue 140 in the bovine sequence (equivalent to position 143 in the human sequence) and 180 (184 in man). However the nucleotide sequence in part of this region – that between residues 163 (166 in man) and 171 (174 in man) – shows a striking similarity with those coding for the enkephalin regions (fig.2).

There are 7 (9 in man) nucleotide differences between this sequence and that coding for a Met-enkephalin and only 5 (7 in man) when compared to that for a Leu-enkephalin. In addition there are 3 (4 in man) differences in the region analogous to that of the flanking basic residues. Overall the region in question is closer to an enkephalin peptide sequence in the bovine than in the human precursor. Compared with a Leu-enkephalin this region shows 70% identity in the bovine molecule and 59% in the human molecule. A computer simulation of 10 000 random nucleotide sequences of 27 residues, taking into account the codon usage in the molecule, indicates that the random occurrence of a fit as good as or better than the one actually observed is 4.7% for the human and 0.2% for the bovine molecule.

I propose that originally this part of the molecule probably coded for an active enkephalin peptide. The divergence of the sequence could have had its origin in a mutation in one of the flanking

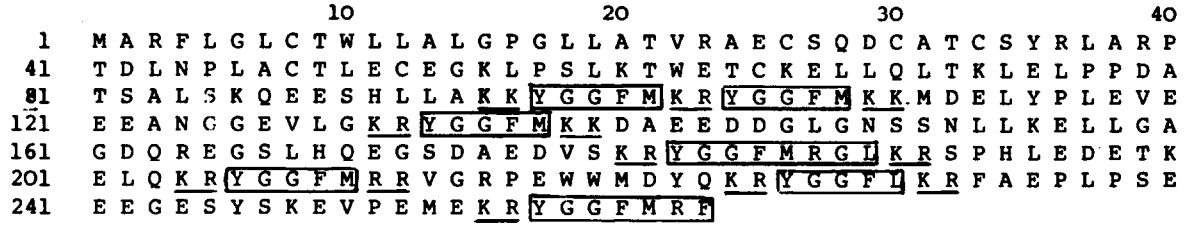

Fig.1. Amino-acid sequence of bovine preproenkephalin. The enkephalin sequences are boxed and the flanking basic residues are underlined.

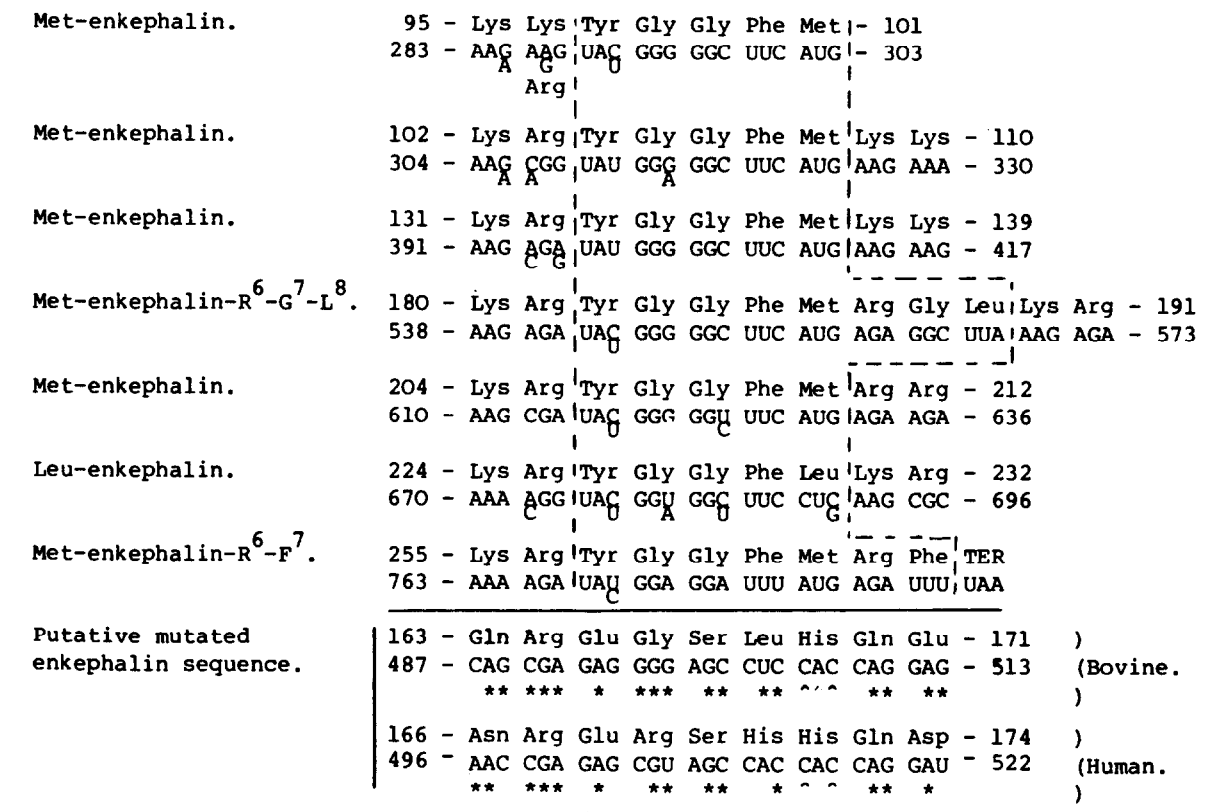

Fig.2. Comparison of the nucleotide sequences of the enkephalin regions and of the putative mutated enkephalin sequence. (*): Nucleotide common to a Met and a Leu-enkephalin. (^): Nucleotide common to only a Leu-enkephalin. For the enkephalin sequences the first two lines represent the amino-acid and the nucleotide sequence of bovine preproenkephalin. The substitutions in the human molecule are shown underneath.

basic residues, which would have prevented the peptide from undergoing normal processing. The peptide sequence was then free to mutate to its present structure.

ACKNOWLEDGEMENTS

I thank Dr R.E. Offord an Dr K. Rose for their comments on the manuscript and the Fond National Suisse de la Recherche Scientifique for computing time.

REFERENCES

[1] Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Hirose, T., Inayama, S., Nakanishi, S. and Numa, S. (1982) Nature 295, 202–206.
[2] Gubler, U., Seeburg, P., Hoffman, B.J., Gage, L.P. and Udenfriend, S. (1982) Nature 295, 206–209.
[3] Comb, M., Seeburg, P.H., Adelman, J.A., Eiden, L. and Herbert, E. (1982) Nature 295, 663–666.