COMPARISON OF THE AMINO TERMINAL SEQUENCES OF BOVINE, DOGFISH, AND LUNGFISH TRYPSINOGENS

Mark A. HERMODSON, Ross W. TYE, Gerald R. REECK, Hans NEURATH and Kenneth A. WALSH
Department of Biochemistry, University of Washington, Seattle, Washington 98105, USA

Received 12 March 1971

1. Introduction

Trypsin and trypsin-like enzymes have been isolated from phyla and species which extend over the whole range of evolution from bacteria to mammals [1]. Purified trypsinogen has been obtained from a more limited number of species. In fact, there are only two lower vertebrates from which this zymogen has been isolated viz. the dogfish (*Squalus acanthias*) [2] and the African lungfish (*Protopterus aethiopicus*) [3]. In all known cases, the activation of mammalian trypsinogens involves the cleavage of a lysyl bond in the invariant amino terminal sequence

X–Asp–Asp–Asp–Asp–Lys–Ile–Y ,

with the concomitant libertation of an activation peptide having the general structure X–$(Asp)_4$–Lys [4]. The availability of purified dogfish and lungfish trypsinogen has enabled us to determine the amino terminal sequences of these zymogens and to compare them to those of bovine trypsinogen and trypsin. Extensive homology between bovine and lungfish trypsinogens has been previously documented [2]. The present investigation is part of a general study of the evolution and phylogenetic variation of proteolytic enzymes.

2. Materials and methods

Dogfish trypsinogen and trypsin were prepared from freshly collected pancreas glands. The crude trypsinogen fraction obtained by chromatography on DEAE-cellulose at pH 8 was further purified by chromatography on SE-Sephadex at pH 3 in the presence of 4 M urea [2, 5]. African lungfish trypsinogen was prepared from aqueous extracts of the acetone powder of pancreatic tissue. The last trypsinogen fraction to emerge from the initial chromatography on DE-52 cellulose at pH 8.0 [3] was subjected to gel filtration on Sephadex G-100 and to chromatography on DE-52 cellulose, both at pH 8.0 [6].

Amino-terminal sequences were determined in the Beckman Sequencer using a modification of the procedure of Edman and Begg [7]. A volatile buffer system was employed and dithiol (1, 4-butanedithiol) was added to the chlorobutane to stabilize the sequenator products [8]. The PTH-amino acids were identified by gas chromatography with the exceptions of PTH-histidine and PTH-arginine which were identified by spot tests on paper [9, 10*].

3. Results and discussion

The present data show extensive homology among the first 20 amino acid residues of the three trypsinogens (table 1). Residues 5–11 (using the numbering system of bovine trypsinogen) are identical in all three cases, as are residues 13 and 19. The present sequences are of particular interest since they include the presumed site of bond cleavage (residue 6–7), the corresponding activation peptide, and the amino-terminal region of the trypsins. Indeed in all documented cases, including the trypsinogens of the cow, goat, horse, sheep, and pig [1], activation involves cleavage of the Lys–Ile bond in the amino-terminal sequence of the zymogen,

X–$(Asp)_4$–Lys–Ile .

It may be assumed therefore that the same bond is

* A method using phenanthrenequinone was used.

North-Holland Publishing Company – Amsterdam

Table 1

Amino-terminal sequences of bovine, *dogfish, and lungfish trypsinogens**.

	Activation Peptide ←					Trypsin →														
	1				5					10					15					20
Bovine	Val	Asp	Asp	Asp	Asp	Lys	Ile	Val	Gly	Gly	Tyr	Thr	Cys	Gly	Ala	Asn	Thr	Val	Pro	Tyr-
Dogfish	Pro	Asp	Asp	Asp	Asp	Lys	Ile	Val	Gly	Gly	Tyr	Glu	Cys	Pro	Lys	His	Ala	Ala	Pro	Trp-
Lungfish	Phe	Pro	Ile	Glu	Glu	Asp	Lys	Ile	Val	Gly	Gly	Tyr	Glu	Cys	Pro	Lys	His	-X-	Val	Pro	Trp-

* K.A. Walsh and H. Neurath, Proc. Natl. Acad. Sci. U.S. 52 (1964) 884.

** There is tentative evidence that the residue marked "X" in the lungfish zymogen is Thr. Identification of Cys at residue 13 of lungfish trypsinogen should be considered tentative. Lungfish trypsinogen gave only single residues after each degradation and was therefore judged to be better than 95% pure. Dogfish trypsinogen was contaminated with dogfish trypsin and hence the sequence after residue 7 was determined on trypsin which was at least 85% pure.

split during the activation of dogfish and lungfish trypsinogens and that the resulting enzymes have the identical amino terminal sequence,

Ile–Val–Gly–Gly–Tyr–.

Perhaps of greatest interest is the sequence corresponding to the activation peptide of lungfish trypsinogen. Whereas the tetraaspartyl sequence heretofore has been considered to be an invariant and characteristic attribute of trypsinogens, in lungfish trypsinogen one of the aspartyl residues is replaced by isoleucine and two others by glutamic acid. Since the activation of all three trypsinogens is similarly dependent on the presence of calcium ions [5, 6, 11], it follows that only three negatively-charged residues are maximally involved in the calcium effect and that glutamic acid can serve for aspartic acid in two of these positions. These conclusions are in accord with model studies on the effect of calcium ions on the rate of tryptic cleavage of the Lys_5–Ile_6 bond [12] and with the abolishment of calcium requirements by amidation of the aspartyl residues of the activation peptide [11].

The present comparative analysis is a further documentation of the utility of sequenator analysis in probing related proteins for sequence homology [13, 14]. The method is particularly effective for zymogens of proteases since in all known cases, activation involves the cleavage of a peptide bond in the amino-terminal region of the protein [4].

Acknowledgements

This work has been supported in part by research grants from the National Institutes of Health (GM 15731) and the American Cancer Society (P–79). G.R.R. has been a recipient of a predoctoral fellowship of the National Science Foundation; R.W.T. has been supported by a departmental training grant of the National Institutes of Health (GM 00052).

References

[1] K.A. Walsh, in: Methods in Enzymology, Vol. XIX, eds. G. Perlmann and L. Lorand (Academic Press, New York, (1970) p. 41.

[2] R.A. Bradshaw, H. Neurath, R.W. Tye, K.A. Walsh and W.P. Winter, Nature 226 (1970) 237.

[3] G.R. Reeck, W.P. Winter and H. Neurath, Biochemistry 9 (1970) 1398.
[4] H. Neurath, R.A. Bradshaw and R. Arnon, in: Structure-Function Relationships of Proteolytic Enzymes, eds. P. Desnuelle, H. Neurath and M. Ottesen (Munksgaard, Copenhagen, 1970) p. 113.
[5] R.W. Tye, Ph. D. Thesis, University of Washington (1971).
[6] G.R. Reeck, Ph. D. Thesis, University of Washington (1971).
[7] P. Edman and G. Begg, European J. Biochem. 1 (1967) 80.
[8] M.A. Hermodson, L.H. Ericsson and K.A. Walsh, Federation Proc. 29 (1970) 728.
[9] F. Sanger and H. Tuppy, Biochem. J. 49 (1951) 463.
[10] J. Ohms, Beckam Instruments, Palo Alto, California, personal communication.
[11] T.M. Randhakrishnan, K.A. Walsh and H. Neurath, Biochemistry 8 (1969) 4020.
[12] J.P. Abita, M. Delaage, M. Lazdunski and J. Savrda, European J. Biochem. 8 (1969) 314.
[13] H.D. Niall and P. Edman, Nature (London) 216 (1967) 262.
[14] H. Rochat, C. Rochat, C. Kupeyan, F. Miranda, S. Lissitzky and P. Edman, FEBS Letters 10 (1970) 349.