THE AMINO TERMINAL SEQUENCES OF ACID PROTEASES - HUMAN PEPSIN AND GASTRICSIN AND THE PROTEASE OF Rhizopus chinensis[1]

Patricia Sepulveda, Kenneth W. Jackson and Jordan Tang

From the Laboratory of Protein Studies, Oklahoma Medical Research Foundation and the Department of Biochemistry and Molecular Biology University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma

Received February 27,1975

Summary

The amino acid sequences near the amino termini of human pepsin (34 residues) and gastricsin (24 residues) and the acid protease from Rhizopus chinensis (27 residues) have been determined using automated Edman degradation. From these results three additional observations were made. First, two structural variants have been observed for human gastricsin and for the Rhizopus protease. Both cases are apparently genetic in origin. Second, a stretch of sequence in the Rhizopus protease, residues 14 to 26, is highly homologous to the known sequence of porcine pepsin at the region of residues 11 to 23. Third, the sequences of the NH_2-terminal region of human pepsin and gastricsin are homologous.

The homology in amino acid sequences among gastric acid proteases is now well known. The amino acid sequences of porcine pepsin (1) and bovine chymosin[2] are highly homologous over almost the entire length of the molecules (2,3). Human gastricsin (4) is homologous to human and porcine pepsin in the amino acid sequences near the carboxyl-termini (5). The structural homology between a microbial protease, penicillopepsin, and the gastric proteases has been shown by Hofmann and co-workers (6). Although the carboxyl terminal regions revealed conclusive similarity, the amino terminal regions showed only weak relatedness (6). It is not certain, therefore, whether this structural homology exists throughout the entire sequences of these two evolutionarily distant enzymes.

---

[1]This work was supported by NIH Research Grants AM-01107 and AM-06487.

[2]The name chymosin (EC 3.4.4.3) was proposed by Foltmann (15) for rennin.

MATERIALS AND METHODS

Human pepsin and gastricsin were purified as previously reported (4). The acid protease of Rhizopus chinensis was purchased from Miles Laboratories. The reagents used for the automatic Sequencer were obtained from Beckman Instruments. Polyamide layers were obtained from Gallard-Schlesinger Chemical Mfg. Corp. (a product of Cheng Chin Co., Taipei, Taiwan). Other reagents used were of the highest purity obtained commercially.

Automated Edman degradations were carried out in a Beckman 890C Sequencer (Palo Alto, Calif.). The procedures were essentially that of Edman and Begg (8). A modified program of Beckman No. 071872 (Fast Peptide-DMAA Program) was used. At the end of the period of coupling with phenylisothiocyanate and drying, a 200-second extraction with ethyl acetate was added. A double cleavage with heptafluorobutyric acid similar to that of Beckman program No. 072172C was also used. The phenylthiohydantoins (PTH-amino acids) were identified by gas-liquid chromatography (9) in a Beckman GC-65 with glass columns and SP-400 packing. The identifications were done both as PTH-amino acids and as their trimethylsilyl derivatives. Supportive thin-layer chromatography (10) was carried out selectively for some residues. Hydrolysis of PTH-amino acids was carried out in 6N HCl at 150° for 24 hr. The amino acid analysis of the hydrolyzate was carried out in a Beckman 120B amino acid analyzer using the procedure of Spackman (11).

RESULTS AND DISCUSSION

The amino terminal sequences of two human gastric proteases, pepsin and gastricsin and a microbial acid protease from Rhizopus chinensis are shown in Fig. 1. The 34-residue sequence of human pepsin was obtained from two samples, purified human pepsin and a 190-residue cyanogen bromide fragment starting at residue 12 (Fig. 1). The sequences of human gastricsin, 24 residues, and of Rhizopus protease, 25 residues, were determined using the pure enzymes.

Sequences determined in this work:
Human Pepsin
Human gastricsin
Acid protease of R. chinensis
Sequences previously reported:
Porcine pepsin (1)
Bovine chymosin (2)
Penicillopepsin (6)
Bovine pepsin (16)
Highly conservative residues
Val
Pro
Asp/Asn Tyr
Asp
Glu Tyr Phe Gly
Ile/Val
Ile Gly Thr Pro

The amino terminal sequences of human pepsin and gastricsin are obviously homologous even though many changes are found in the first 15 residues. This is in contrast to the comparison of the carboxyl terminal regions of the enzymes where few different residues were found (5). From the alignment of the first 8 residues of these two sequences (Fig. 1), it is interesting to note that the amino terminus of gastricsin is one residue longer than pepsin. The amino terminus of chymosin is two residues longer than that of pepsin (12) when the amino terminal sequences of the two enzymes are aligned (Fig. 1). An interesting aspect of the gastricsin sequence is the presence of two amino terminal residues, a major serine residue and a minor valine residue - about 20% of total. Three separate gastricsin preparations gave the same results when analyzed in the Sequencer. Minor contamination with human pepsin was excluded because no other residue corresponding to the pepsin sequence was found in the next several steps of Edman degradation. Although a possibility of a cleavage at a specific position in some gastricsin molecule cannot be completely discounted, such an event is not probable. Therefore, we believe that there are two genetic variants of human gastricsin. Since the enzyme was purified from pooled samples of human gastric juice (7), the origin of these structural variants cannot be certain at this time.

In the sequence of the Rhizopus protease, residue 15 gave rise to both isoleucine and valine. This observation was confirmed in another experiment. Fig. 2 shows the

Fig. 1. The NH_2-terminal sequence of human pepsin, gastricsin and the acid protease from Rhizopus chinensis. These new sequences are shown on the first three lines. The signs under each residue indicate the method(s) of positive identification of PTH-amino acids: ⟶ , gas-liquid chromatography; ⇢ , thin-layer chromatography, ⟹ , both chromatographic methods; — , no identification; and *, hydrolysis and amino acid analysis. The residue numbers are assigned separately for each enzyme, starting at NH_2-termini. The alignments of the sequences and the gaps are arbitrarily chosen to create maximum homology. Four previously published NH_2-terminal sequences of acid proteases are included for comparison. Homology between the sequences of the protease of R. chinensis and porcine pepsin are indicated as vertical links: solid lines, identical residues, and broken lines, residues whose differences could be the result of a single-step mutation. The two residues on the left of NH_2-terminus of porcine pepsin (-Ala-Leu, residues -2 and -1) are part of the pepsinogen molecule (1,17).

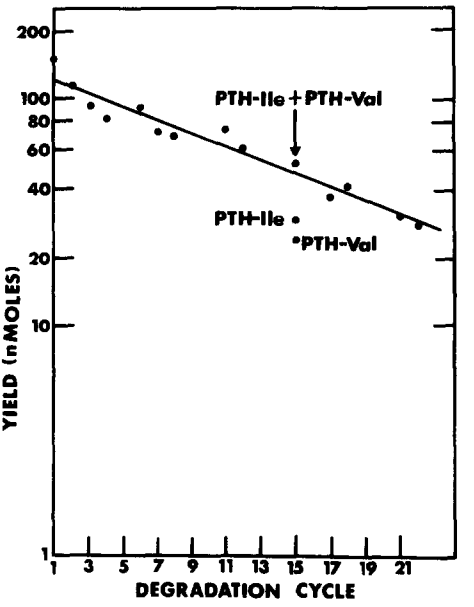

Fig. 2. The quantitative yield of PTH-amino acids of the hydrophobic residues from the automatic degradation of the NH_2-terminal region of the acid protease of Rhizopus chinensis. At residue 15, the yield of PTH-Val or PTH-Ile is apparently low. The combined yields of the two is comparable to the yield of the other residues. The data were calculated from the peak heights of the gas-liquid chromatography, uncorrected for recovery losses.

semilogarithmic plot of the yield of PTH-amino acids of hydrophobic residues in the first 22 steps of the degradation. Assuming position 15 is occupied by a single amino acid residue, the yield of PTH-Val or PTH-Ile was considerably below the expected yield, as shown by the decreasing line in Fig. 2. However, the combined yields of the two PTH-amino acids had the expected yield. These results suggest that the two amino acid residues at position 15 are derived from two structural variants of the Rhizopus enzyme. Graham, Sodek and Hofmann (13) had separated the commercially crystalline Rhizopus protease into two isoelectrically distinct fractions. Although the amino terminal 9 residues of these two fractions were identical, the amino acid compositions showed slight differences. Their results are consistant with our observations reported in this communication.

As shown in Fig. 1, a strong homology exists between the regions in porcine pepsin (residue 11-24) and in the Rhizopus protease (residue 14-27). A homology was also observed with the corresponding regions in penicillopepsin and the Rhizopus protease. These observations suggest a new alignment between penicillopepsin and pepsin (Fig. 1) which is different from that previously proposed (6).

Existing evidence now suggests that the microbial acid proteases and the gastric proteases share homologous sequences throughout their entire molecules. First, the size of the microbial enzymes is almost identical to the 327 residues in porcine pepsin (1). Second, the high degree of similarity in the primary structure of these two groups of enzymes is now known for the amino terminal region (above) and the carboxyl terminal region (6). Third, several microbial acid proteases, including the Rhizopus enzyme and penicillopepsin, are inactivated by two active-site directed, specific pepsin inhibitors (see ref. 6 for review). Amino acid sequences similar to the active site sequences of pepsin have been obtained from awamorin of Aspergillus awamori (14) and penicillopepsin (6). The close homology of microbial and gastric acid proteases also predicts a high degree of similarity between the three-dimensional structures of the gastric and microbial acid proteases.

ACKNOWLEDGEMENTS

We wish to thank Paul Lanier for amino acid analyses, Alex Jacocks and Rick Bynum for assisting in the enzyme purifications.

REFERENCES

1. Tang, J., Sepulveda, P., Marciniszyn, J., Jr., Chen, K., Huang, W-Y., Tao, N., Liu, D. and Lanier, J. P. (1973) Proc. Nat. Acad. Sci. U.S.A. 70, 3437-3439.
2. Foltmann, B. and Hartley, B. S. (1967) Biochem. J. 104, 1064-1074.
3. Tang, J. and Hartley, B. S. (1970) Biochem. J. 118, 611-623.
4. Tang, J., Wolf, S., Caputto, R. and Trucco, R. E. (1959) J. Biol. Chem. 234, 1174-1178.

5. Huang, W-Y. and Tang, J. (1970) J. Biol. Chem. 245, 2189-2193.
6. Hofmann, T. (1974) in Advances in Chemistry Series, No. 136, Food Related Enzymes pp. 146-185, Am. Chem. Soc.
7. Tang, J. (1970) Methods of Enzymol. 19, 406-421.
8. Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80-91.
9. Pisano, J. J. and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5597-5607.
10. Summers, M. R., Smythers, G. W., and Oroszlan, S. (1973) Anal. Biochem. 53, 624-628.
11. Spackman, D. H. (1963) Fed. Proc. 22, 244.
12. Tang, J. (1970) Biochem. Biophys. Res. Commun. 41, 697-703.
13. Graham, J. E. S., Sodek, J. and Hofmann, T. (1973) Can. J. Biochem. 51, 789-796.
14. Kovaleva, G. G., Shimanskaya, M. O. and Stepanov, V. M. (1972) Biochem. Biophys. Res. Commun. 49, 1075-1081.
15. Foltmann, B. (1970) Methods in Enzymol. 19, 421-436.
16. Harboe, M. and Foltmann, B. (1973) FEBS Letters 34, 311-314.
17. Pedersen, V. B. and Foltmann, B. (1973) FEBS Letters 35, 255-256.