AMINO ACID SEQUENCE OF HOG PANCREATIC α-AMYLASE ISOENZYME I

I. KLUH

Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 166 10 Prague 6, Czechoslovakia

Received 20 October 1981

1. Introduction

Crystalline α-amylase (EC 3.2.1.1) from hog pancreas contains two α-amylase enzyme forms designated isoenzyme I and isoenzyme II with respect to their different electrophoretic mobilities [1,2]. We have characterized both isoenzymes by determination of relative molecular mass (M_r) and amino acid composition; the C-terminal amino acid [3] and N-terminal amino acid sequence PCA-Tyr (PCA, pyrrolid-2-one-5-carboxylic acid) are identical in both isoenzymes. The M_r and amino acid composition are likewise identical within the range of experimental error. So far partial sequential data have been published on isoenzyme II obtained by the analysis of its tryptic digest [5] and partial or complete data on certain cyanogen bromide fragments of isoenzyme I [6].

Here the primary structure of isoenzyme I comprising 496 amino acid residues is reported. The primary structure of isoenzyme I of pancreatic α-amylase can be compared with the structure of α-amylase precursors from mouse pancreas, salivary glands and liver. The structures of these precursors have been derived from the nucleotide sequences of the corresponding mRNAs [7,8]. The comparison together with the data on the N-terminal sequence of α-amylase from mouse pancreas and salivary glands [9] point to the possibility of identical post-translational cleavage of α-amylase precursors in 2 animal species.

2. Materials and methods

Crystalline hog α-amylase was prepared according to [10]. Isoenzyme I was obtained by chromatography on DEAE-cellulose [3]. The CNBr cleavage of the native enzyme was performed in 70% formic acid for 24 h at 25°C. The reaction mixture was freed of the excess of CNBr and separated by gel permeation chromatography on Sephadex G-100 in 0.4% formic acid. The individual fragments were isolated by rechromatography under identical conditions or by ion-exchange chromatography on SE-Sephadex G-25 in 8 M urea. The amino acid sequence of the N-terminal parts of the fragments isolated was determined by automatic sequential analysis based on the phenylisothiocyanate method in a Beckman model 890 C amino acid sequenator. The sequences thus obtained were checked and completed by sequential analysis of peptides from tryptic and chymotryptic digests of the corresponding fragments. These analyses were effected by manual degradation using phenylisothiocyanate method [11] which was replaced by the 4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate/phenylisothiocyanate double coupling method [12] during later stages of the work.

3. Results and discussion

3.1. *Determination of amino acid sequence*

The primary structures of isoenzyme I was derived from the sequential analysis of fragments isolated from the cyanogen bromide digest of native enzyme (fig.1). The following cyanogen bromide fragments were isolated: positions 1–178, 179–202, 203–274, 203–287, 288–328, 329–339, 340–394, 395–496. Fragment 1–178 contained 3 homoserine residues and was converted into its *S*-carboxymethyl derivative by mercaptolysis and subsequent treatment with iodoacetic acid. Fragment 104–178 with one homoserine residue was isolated from this material. Fragment 203–287 contained one homoserine and one methionine residue and was obtained in minor quantity only. Fragment 395–496 did not contain homoserine, its C-terminal was leucine; this fragment therefore repre-

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.75

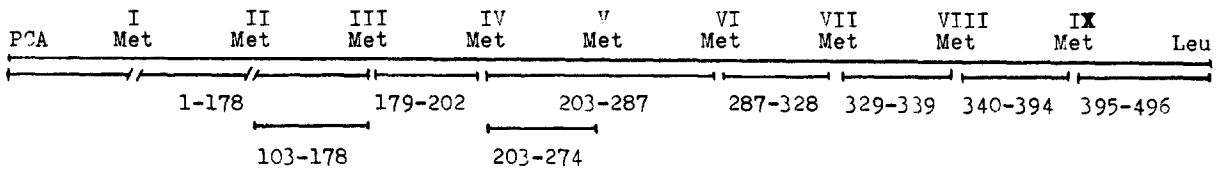

Fig.1. Schematic representation of cleavage of the native molecule of isoenzyme I of hog pancreatic α-amylase by cyanogen bromide and of the fragments formed.

sents the C-terminal fragment of the molecule. The sequence of residues 1–102 was elucidated by analysis of overlapping fragments isolated from the tryptic and chymotryptic digest of fragment 1–178. Fragment 275–287 was not isolated. The amino acid sequence accounting for this region was obtained by the analysis of tryptic peptide 274–282, obtained from the minority fragment 203–287. The difference in the amino acid composition of fragments 203–287 and 203–274 and the existence of peptide 274–282 show that the positions 283–287 are occupied by glycines and leucine. Even though we did not obtain proof of the presence of tryptophan in this region we assume that, in the light of the similarity between isoenzyme I and the precursors of mouse amylases, the structures in this region are identical. The numbering of amino acid residues in fig.2 is based on the same assumption. The cyanogen bromide fragments were arranged on the basis of the knowledge of tryptic peptides containing methionine residues isolated from isoenzyme II [5]. Peptides K7 and K11 were subjected to repeated sequential analysis and their structures collected as follows: K7 His–Met–Trp–Pro–Gly–Asp–Ile–Lys, K11 Met–Ser–Tyr–Leu–Lys. The complete identity in the primary structures of isoenzyme I and isoenzyme II in regions around the methionine residues permitted the methionine peptides isolated from isoenzyme II to be used as links of the fragments of isoenzyme I. The primary structure of isoenzyme I is shown in fig.2.

3.2. *Structural relations between isoenzyme I of hog pancreatic α-amylase and the precursors of mouse amylases*

The knowledge of the N-terminal sequence of mouse pancreatic α-amylase [9] permits us to align the structure of the part of the precursor corresponding to the active enzyme, with the structure of isoenzyme I of hog amylase (fig.2). Altogether 86% of amino acid residues occupy the same positions in both structures. Considerable similarities exist also between hog pancreatic amylase and mouse amylases from salivary glands and liver. The comparison in fig.2 shows that pancreatic amylases of both species differ in length by 3 amino acid residues. The insertion of these 3 residues into the hog pancreatic amylases is identical to the insertion of the same residues in the same position observed with mouse salivary and liver amylases. In view of the similarities existing, a similar post-translational cleavage of all mouse α-amylase precursors and of hog amylase precursor is also highly probable. This cleavage involves in all cases the scission of the peptide bond at the amino side of glutamine which is subsequently cyclized to pyrrolidonecarboxylic acid.

In [4] we attempted to find evidence of the presence of glutamine at various stages of preparation of hog pancreatic α-amylase. We never observed N-terminal glutamine and this leads us to conclude that pyrrolidonecarboxylic acid is not an artificial product of preparation of the enzyme. Hence the cyclization of glutamine follows after the post-translational cleavage or is a part of it. The relatively low degree of homology in the immediate neighborhood of the N-terminal sequence and especially the replacement of aspartic acid in position 5 by an uncharged alanine residue show that the immediate neighborhood of the peptide chain does not affect significantly the equilibrium of the spontaneous closure of the pyrrolidonecarboxylic acid. In this respect it should be noted that there is a possibility that the post-translational system includes the unique enzyme responsible for the conversion of glutamine to pyrrolidonecarboxylic acid. The formation of pyrrolidonecarboxylic acid during protein biosynthesis has been reviewed in [13]. The fact that α-amylases, which are the key metabolic enzymes with a considerable degree of structural conservation, belong to the great number of proteins terminated with pyrrolidonecarboxylic acid should stimulate a more profound investigation of the biological significance of N-terminal pyrrolidonecarboxylic acid in proteins.

```
 1                                  10                                  20
PCA-Tyr-Ala-Pro-Gln-Thr-Gln-Ser-Gly-Arg-Thr-Ala-Ile-Val-His-Leu-Phe-Glu-Trp-Arg-Trp-Val-Asp-Ile-Ala-
Gln     Asp     His     Ser Asp

                  30                                  40                                  50
Lys-Glu-Cys-Glu-Arg-Tyr-Leu-Gly-Pro-Lys-Gly-Phe-Gly-Gly-Val-Gln-Val-Ser-Pro-Pro-Asn-Glu-Asn-Val-Val-
                           Ala

                                     60                                  70
Thr-Gly-Asn-Pro-Ser-Arg-Pro-Trp-Trp-Glu-Arg-Tyr-Gln-Pro-Val-Ser-Tyr-Lys-Leu-Cys-Thr-Arg-Ser-Gly-Asn-
Val His                                                 Ile             Ile

                 80                                  90                                  100
Glu-Asp-Glu-Phe-Arg-Asp-Met-Val-Thr-Arg-Cys-Asn-Asn-Val-Gly-Val-Arg-Ile-Tyr-Val-Asp-Ala-Val-Ile-Asn-

                                    110                                 120
His-Met-Cys-Gly-Ser-Gly-Ala-Ala-Ala-Gly-Thr-Gly-Thr-Thr-Cys-Gly-Ser-Tyr-Cys-Asn-Pro-Gly-Asn-Arg-Glu-
                Ala     Asn Pro             Ser

                 130                                 140                                 150
Phe-Pro-Ala-Val-Pro-Tyr-Ser-Ala-Trp-Asp-Phe-Asn-Cys-Asn-Gly-Lys-Lys-Thr-Ala-Ser-Gly-Gly-Ile-Glu-Ser-
                                              Asp     Lys Cys Asn----------Gly Glu     Asp Asn

                                    160                                 170
Tyr-Asn-Asp-Pro-Tyr-Gln-Val-Arg-Asp-Gly-Cys-Gln-Val-Leu-Leu-Leu-Asp-Leu-Ala-Leu-Glu-Lys-Asp-Tyr-Val-
            Ala                     Cys Arg Leu Thr Gly

                 180                                 190                                 200
Arg-Ser-Met-Ile-Ala-Asp-Tyr-Leu-Asn-Lys-Leu-Ile-Asp-Ile-Gly-Val-Ala-Gly-Phe-Arg-Ile-Asp-Ala-Ser-Lys-
    Thr Lys Val             Met     His                                             Leu         Ala

                                    210                                 220
His-Met-Trp-Pro-Gly-Asp-Ile-Lys-Ala-Val-Leu-Asp-Lys-Leu-His-Asn-Leu-Asn-Thr-Asn-Trp-Phe-Pro-Ala-Gly-
                Arg                                                             Lys         Ser Gln

                 230                                 240                                 250
Ser-Arg-Pro-Phe-Ile-Phe-Gln-Glu-Val-Ile-Asp-Leu-Gly-Gly-Glu-Ala-Ile-Lys-Gly-Ser-Glu-Tyr-Phe-Ser-Asn-
                                                                                                    Gly

251                                 260                                 270
Gly-Arg-Val-Thr-Glu-Phe-Lys-Tyr-Gly-Ala-Lys-Leu-Gly-Thr-Val-Val-Arg-Lys-Trp-Ser-Gly-Glu-Lys-Met-Ser-
                                                        Ile             Asn

                 280                                 290                                 300
Tyr-Leu-Lys-Gly-Pro-Leu-Lys(Gly,Trp,Gly,Leu)Met-Pro-Ser-Asp-Arg-Ala-Leu-Val-Phe-Val-Asp-Asn-His-Asp-
            Asn-Trp-Gly-Glu-Gly-Trp-Gly-Leu-Val

                                    310                                 320
Asn-Gln-Arg-Gly-His-Gly-Ala-Gly-Gly-Ala-Ser-Ile-Leu-Thr-Phe-Trp-Asp-Ala-Arg-Leu-Tyr-Lys-Val-Ala-Val-
                                    Ser                                 Met         Met

                 330                                 340                                 350
Gly-Phe-Met-Leu-Ala-His-Pro-Tyr-Gly-Phe-Thr-Arg-Val-Met-Ser-Ser-Tyr-Arg-Trp-Ala-Arg-Asn-Phe-Val-Asn-
                                                                        Asn                 Gln

                                    360                                 370
Gly-Gln-Asp-Val-Asn-Asp-Trp-Ile-Gly-Pro-Pro-Asn-Asn-Asn-Gly-Val-Ile-Lys-Glu-Val-Thr-Ile-Asn-Ala-Asp-
    Lys                                                         Thr

                 380                                 390                                 400
Thr-Thr-Cys-Gly-Asn-Asp-Trp-Val-Cys-Glu-His-Arg-Trp-Arg-Gln-Ile-Arg-Asn-Met-Val-Trp-Phe-Arg-Asn-Val-
Ser

                                    410                                 420
Val-Asp-Gly-Gln-Pro-Phe-Ala-Asn-Trp-Trp-Asp-Asn-Gly-Ser-Asn-Gln-Val-Ala-Phe-Gly-Arg-Gly-Asn-Arg-Gly-
    Asn                                     Asn                         Ser

                 430                                 440                                 450
Phe-Ile-Val-Phe-Asn-Asp-Asp-Asp-Gln-Leu-Trp-Ser-Gly-Thr-Leu-Gln-Thr-Gly-Leu-Pro-Gly-Gly-Thr-Tyr-Cys-
                    Asn         Trp Ala Leu     Ala                                 Ala

                                    460                                 470
Asp-Val-Ile-Ser-Gly-Asp-Lys-Val-Gly-Asn-Ser-Cys-Thr-Gly-Ile-Lys-Val-Asn-Val-Ser-Ser-Asp-Gly-Thr-Ala-
                                Asp Gly Asn             Leu Arg             Gly             Lys

                 480                                 490                 496
Gln-Phe-Ser-Ile-Ser-Asn-Ser-Ala-Glu-Asp-Pro-Phe-Ile-Ala-Ile-His-Ala-Gln-Ser-Lys-Leu
His                                                             Asp
```

Fig.2. Comparison of the structure of hog pancreatic α-amylase, isoenzyme I with a part of the peptide chain of the precursor of mouse pancreatic α-amylase [7]. The primary structure of hog pancreatic α-amylase, isoenzyme I is shown on the first line. The structure of the precursor of mouse pancreatic α-amylase is shown from amino acid residue no. 16. The amino acid residues of pancreatic α-amylase are designated by numbers above the symbols. PCA = pyrrolid-2-one-5-carboxylic acid.

Acknowledgements

The author thanks Dr Kostka, head of the department, for constant interest, Dr K. Grüner for running the automatic version of the phenylisothiocyanate method, Mrs V. Lišková for technical help and Mr Zbrožek for the amino acid analysis.

References

[1] Marchis-Mouren, G. and Paséro, L. (1967) Biochim. Biophys. Acta 140, 366–368.
[2] Kluh, I. and Sorm, F. (1968) Abst. 4th FEBS Meet. Prague 1968, p. 632.
[3] Kluh, I. (1979) Coll. Czech. Chem. Commun. 44, 288–293.
[4] Kluh, I. (1979) Coll. Czech. Chem. Commun. 44, 145–147.
[5] Meloun, B., Kluh, I. and Morávek, L. (1980) Coll. Czech. Chem. Commun. 45, 2572–2582.
[6] Paséro, L., Abadie, B., Chicheportiche, Y., Mazzei, Y., Moinier, D., Bizzozero, J. P., Fougereau, M. and Marchis-Mouren, G. (1981) Biochimie 63, 71–76.
[7] Hagenbüchle, O., Bovey, R. and Young, R. A. (1980) Cell 21, 179–187.
[8] Hagenbüchle, O., Tosi, M., Schibler, U., Bovey, R. and Wellauer, P. K. (1981) Nature 289, 643–646.
[9] Karn, R. C., Petersen, T. E., Hjorth, J. P., Nielsen, J. T. and Roepstorff, P. (1981) FEBS Lett. 126, 292–296.
[10] Fischer, E. H. and Stein, E. A. (1954) Arch. Sci. 7, 9–161.
[11] Edman, P. (1970) in: Protein Sequence Determination (Needleman, S. B. ed) p. 211, Springer, Berlin, New York.
[12] Chang, J. Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[13] Abraham, G. N. and Podell, D. N. (1981) Mol. Cell. Biochem. 38, 181–190.