Volume 238, number 1, 105–108 FEB 06339 September 1988

Identification of a Glu > Lys substitution in the activation segment of human pepsinogen A-3 and -5 isozymogens by peptide mapping using endoproteinase Lys-C

Ruud A. Bank, Bart C. Crusius, Toon Zwiers, Stephan G.M. Meuwissen*, Fre Arwert and Jan C. Pronk

*Institute of Human Genetics and *Department of Gastroenterology, Free University, Amsterdam, The Netherlands*

Received 28 July 1988

The isozymogens PGA-3 and PGA-5 of human pepsinogen A were digested with endoproteinase Lys-C. The peptides were separated by reverse-phase HPLC. PGA-5 showed a peak strongly absorbing at 254 nm absent in PGA-3. Analysis of amino acid composition using the Pico-Tag methodology combined with DABITC-sequencing reveals the sequence Tyr-Phe-Pro-Gln-Trp-Lys (peptide 37–43 of the activation segment). This confirms a study at the DNA level by our group [16] suggesting a Glu > Lys mutation at position 43 in the activation segment of PGA-5. Furthermore, it is proposed that the number of genetic variants of PGA is higher than is actually seen by electrophoresis.

Pepsinogen; Peptide mapping; Amino acid sequence; Evolution; (Human)

1. INTRODUCTION

Human pepsinogen A (PGA), the inactive precursor of pepsin A (EC 3.4.23.1), is synthesized in the gastric mucosa. It consists of a single polypeptide chain of 373 amino acids with a molecular mass of 40300 Da [1,2]. During activation an activation segment of 47 residues is released from the N-terminus giving rise to pepsin [3,4]. After agar or polyacrylamide gel electrophoresis, PGA reveals a multibanded pattern, consisting of 1–5 isozymogens, designated PGA2–5 and 5S in decreasing order of anodal mobility [5–7]. PGA-5S is only rarely seen whereas PGA-2 is a posttranslational modification of PGA-3 [6,8]. The main PGA isozymogens, PGA-3, -4 and -5, differ by approximately one charge unit [9] as may be estimated from the electrophoretic fractionation and the order of elution from anion-exchange columns. Whitecross et al. [10] and Korsnes and Gedde-Dahl [6] have demonstrated that PGA-3 and -5 are converted to pepsins which are indistinguishable by electrophoresis. Hence charge differences between PGA-3 and -5 must be present in the activation segment. Up to now it has not been possible to identify the amino acid differences between the various PGA isozymogens by protein sequencing. In this paper we demonstrate a Glu > Lys substitution in the activation segment leading to a charge difference of two units.

Correspondence address: J.C. Pronk, Institute of Human Genetics, Free University, PO Box 7161, 1007 MC Amsterdam, The Netherlands

2. MATERIALS AND METHODS

2.1. *Materials*

Endoproteinase Lys-C from *Lysobacter enzymogenes* was obtained from Boehringer and was used without further purification. Reagents for amino acid analysis and sequencing (TFA, DABITC, PITC, TEA, heptane, ethyl acetate, *n*-butyl acetate, pyridine, constant boiling HCl) were obtained from Pierce (sequanal grade) as well as urea and 2-mercaptoethanol. Acetonitrile was an HPLC-grade product from Rathburn. Water was deionized, then purified with the Milli-Q water system. All other chemicals were at least of analytical grade.

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

2.2. *Purification of PGA-3 and PGA-5*

PGA-3 and PGA-5 have been isolated from a human stomach (phenotype BD) obtained after gastrectomy. The isozymogens were purified on a Sephadex G75 column combined with consecutive separations with FPLC using Mono-Q and Superose 12 columns as described previously [11].

2.3. *Carboxymethylation*

Reduction and *S*-carboxymethylation were performed as described by Crestfield et al. [12]. In short, 2–3 mg PGA was dissolved in 12 ml of a solution containing 8 M urea, 0.36 M Tris and 5.4 mM disodium EDTA, adjusted to pH 8.6 with HCl. After the addition of 150 μl 2-mercaptoethanol the solution was gently stirred and held under nitrogen. After 4 h 1 ml of a freshly prepared solution containing 1 M NaOH and 1.53 M iodoacetic acid was added. Alkylation was allowed to proceed under nitrogen and in the dark. 15 min after the addition of iodoacetic acid, the carboxymethylation was stopped by adding 150 μl 2-mercaptoethanol. The samples were dialysed exhaustively for at least 2 × 24 h in the dark against 0.1 M NH_4HCO_3 and freeze-dried.

2.4. *Cleavage by endoproteinase Lys-C*

The freeze-dried samples of PGA-3 (2 mg) and PGA-5 (3 mg) were dissolved in 1.25 ml and 1.50 ml of 0.1 M NH_4HCO_3, pH 8.0, respectively. Enzymatic degradation of PGA by endoproteinase Lys-C was performed by incubation at 37°C for 2 h, at an enzyme to substrate ratio of 1:75 (w/w). After this incubation the same amount of enzyme was added, and after 22 h the digestion was stopped by addition of 6 M HCl (40 μl) and subsequently analyzed by HPLC for peptide mapping.

2.5. *High-performance peptide mapping chromatography*

Chromatography was performed with a Waters HPLC-system consisting of two 6000 A pumps, a 680 gradient controller, a U6K injector with a 2 ml sample loop and a 440 detector. Digestions were chromatographed on a column (250 × 4.6 mm i.d.) packed with Nucleosil 10C18 (Macherey-Nagel) and detected at 214 and 254 nm. Elution was performed with a linear gradient at a flow rate of 1 ml/min. The gradient began at 5% B (= 0.1% TFA in 67% acetonitrile; A = 0.1% TFA in water) and rose to 70% at 30 min. Where necessary, the column effluent was collected into tubes and aliquots were removed for amino acid compositional analysis.

2.6. *Amino acid analysis by the Pico-Tag method*

Amino acid analyses of the peptides were carried out according to the protocol developed by Bidlingmeyer et al. [13] and described in detail in the Pico-Tag amino acid analysis operator's manual [14]. The Pico-Tag method is based upon a gaseous acid hydrolysis of the peptides followed by the formation of a PTC-derivative of the amino acids via derivatization with PITC. The derivatized amino acids undergo HPLC-separation on a reversed-phase Nova-Pak C18 column (150 × 3.9 mm i.d.) which is specially selected and tested for suitability in separating PTC amino acids (Pico-Tag column).

2.7. *Amino acid sequence determination*

N-terminal analysis was carried out with the manual liquid-phase DABITC/PITC double coupling procedure as described by Chang [15].

3. RESULTS

The complete primary structure of human PGA has already been deduced from nucleotide se-

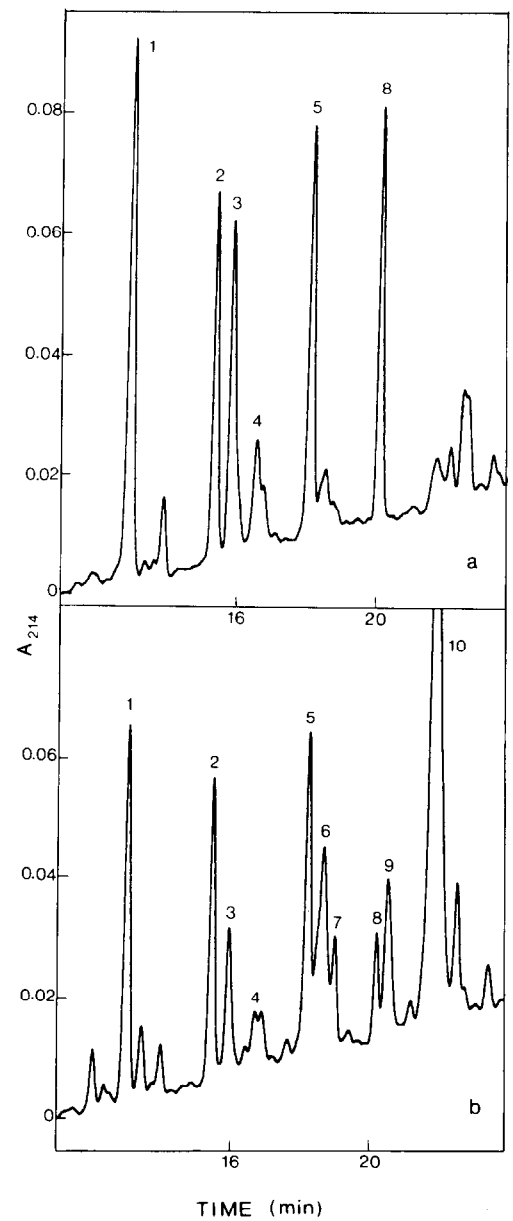

Fig.1. Separation by reversed-phase HPLC of peptides from the digests of PGA-3 (a) and PGA-5 (b) by lysylendopeptidase-C. For experimental conditions see section 2.5. Peak identification and sequences are given in fig.2. In contrast to the other peptides, peak 10 also strongly absorbed at 254 nm (not shown).

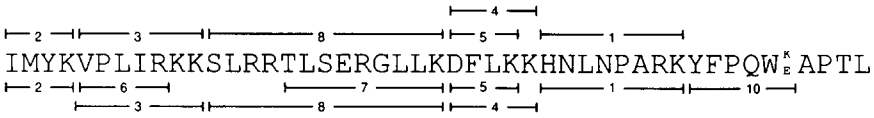

Fig.2. Peak identification and sequence of the major fragments generated by digestion of PGA-3 (upper part) and PGA-5 (lower part) by lysylendopeptidase-C according to the analysis of PTC amino acids combined with DABITC-sequencing. Fragments are numbered according to their elution position on HPLC (see fig.1). The complete amino acid sequence of the activation segment of human pepsinogen A is shown here, according to data in references [1,16].

quences [1]. These data have shown that lysine residues are exclusively located in the activation segment. Detailed analysis of the PGA gene complex at the DNA level by our group has revealed a nucleotide substitution giving rise to a Glu > Lys substitution of the 43rd amino acid residue of the activation segment [16]. It has been postulated that these residues are present in PGA-3 and PGA-5, respectively. On the basis of these previous DNA sequence analyses, cleavage of PGA-3 and PGA-5 by endoproteinase Lys-C should generate 5 and 6 peptides, respectively. These fragments, covering a large part of the total sequence of the activation segment of PGA, were in fact identified in the peptide maps of PGA-3 and PGA-5 (fig.1). The assignment of each peak, deduced by the amino acid compositional analysis combined with partial N-terminal analysis by DABITC-sequencing, is given in fig.2. Analysis of the peptide maps showed that cleavage had occurred primarily on the C-terminal side of lysine residues. The number of peaks however exceeds the expected number of theoretical peptides because of irregular or incomplete cleavage, largely due to Lys-Lys sequences (see fig.2).

4. DISCUSSION

PGA is a highly polymorphic protein, not only in man but also in other mammals, e.g. monkeys, bears and rabbits [17–19]. For example, in rabbits 6 different PGA proteins are sequenced, five of them having a different amino acid sequence of the activation segment [19]. In humans, up to 4 primary gene products are found after polyacrylamide gel electrophoresis named PGA-3, -4, -5 and -5S. Table 1 shows the existence of 4 different activation segments. Sequence II belongs to PGA-3 whereas sequence III belongs to PGA-5 ([16] and this study). According to Evers et al. [16,20], PGA-4 has also sequence II: it differs from PGA-3 in having 4 amino acid substitutions in the pepsin part. According to Foltmann [9], the preparation sequenced by Foltmann and Axelsen [4] (= sequence I) consisted mainly of PGA-3. Thus several components are present in the electrophoretic zone called PGA-3 indicating that the number of genetic variants is higher than is actually seen by electrophoresis. The sequence of PGA-5S is unknown nor is it known to which PGA-isozymogen sequence IV belongs.

There is evidence that the amino acid exchange rate is higher in the activation peptide than in the pepsin part [2,21]. This is of interest in view of the molecular evolution of pepsinogens and useful for the molecular taxonomy of the animals whose pepsinogen is sequenced. Because of the fact that lysine residues are almost completely restricted to the activation segment of PGA, the peptide mapping technique presented here will be useful in analyzing genetic variants of PGA. Because of its reasonably high specificity endoproteinase Lys-C generates only a few fragments thus giving rise to simple peptide maps which are easy to interpret. In addition, lysine is often involved in substitutions between PGA proteins of different animals (fig.3)

Table 1

The different amino acid sequences of the activation segments of human pepsinogen A according to the literature

Sequence	Amino acid position				Isozymogen	Reference(s)
	18	20	42	43		
I	Ile	Asp	ND	ND	PGA-3	4
II	Ser	Arg	Trp	Glu	PGA-3,4	1,16,20
III	Ser	Arg	Trp	Lys	PGA-5	16,20
IV	Ser	Arg	Arg	Lys	PGA-?	3

In this report we could confirm sequence II for PGA-3 and sequence III for PGA-5. ND, not determined

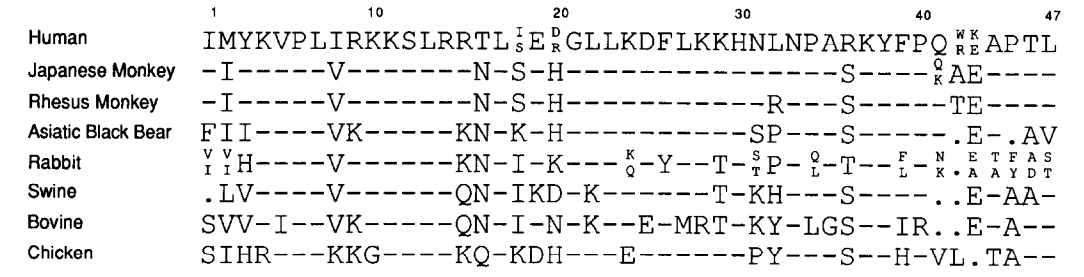

Fig.3. Comparison of the amino acid sequence of the activation segments of human pepsinogen A [1,3,4,16] with those of Japanese monkey (*Macaca fuscata*) [2,22], Rhesus monkey (*Macaca mulatta*) [21], Asiatic black bear (*Selenarctos thibetanus japonicus*) [18], rabbit [19], porcine [23,24,25], bovine [26] and chicken [27] pepsinogens A. The amino acids are expressed in the single letter code recommended by the IUPAC-IUB Commission on Biochemical Nomenclature. When 2 types of residues were identified at one position, both are shown (small letter codes). Residue numbers are given based on the sequences of human and monkey pepsinogen A. The residues common to human pepsinogen A are shown by dashes. The dots indicate gaps to maximize homology.

thus allowing intra- and interindividual difference studies.

Acknowledgements: The authors wish to thank Thea A. Abrahami de Melverda and Marianne Tolk for their excellent technical assistance and Dr Rune R. Frants for carefully reading this manuscript.

REFERENCES

[1] Sogawa, K., Fuji-Kuriyama, Y., Mizukami, Y., Ichihara, Y. and Takahashi, K. (1983) J. Biol. Chem. 258, 5306–5311.
[2] Kageyama, T. and Takahashi, K. (1986) J. Biol. Chem. 261, 4395–4405.
[3] Kageyama, T. and Takahashi, K. (1980) J. Biochem. 88, 571–582.
[4] Foltmann, B. and Axelsen, N.H. (1980) Fed. Proc. 60, 271–280.
[5] Sammloff, I.M. (1969) Gastroenterology 57, 659–669.
[6] Korsnes, L. and Gedde-Dahl, T. (1980) Ann. Hum. Genet. 43, 199–212.
[7] Taggart, R.T., Yu, P.L., Karn, R.C., Conneally, P.M. and Merritt, A.D. (1978) Cytogenet. Cell Genet. 22, 335–340.
[8] Defize, J., Pals, G., Pronk, J.C., Frants, R.R., Rimmelzwaan, G., Westerveld, B.D. and Eriksson, A.W. (1985) Scand. J. Clin. Lab. Invest. 45, 649–655.
[9] Foltmann, B. (1985) in: Pepsinogens in Man. Clinical and Genetic Advances (Kreuning, J. et al. eds) pp.1–13, Alan R. Liss, New York.
[10] Whitecross, D.P., Armstrong, C., Clarke, A.D. and Piper, D.W. (1973) Gut 14, 850–855.
[11] Pals, G., Westerveld, B.D., Defize, J., Pronk, J.C. and Eriksson, A.W. (1985) in: Proceedings First FPLC-Symposium, pp.63–66, Vonk, Zeist.
[12] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622.
[13] Bidlingmeyer, B.A., Cohen, S.A. and Tarvin, T.L. (1984) J. Chromatogr. 336, 93–104.
[14] Pico-Tag Amino Acid Analysis System, Operator's Manual No. 88140. Waters Assoc., Milford, MA, June 1984.
[15] Chang, J.-Y. (1983) Methods Enzymol. 91, 455–466.
[16] Evers, M.P.J., Zelle, B., Peeper, D.S., Mager, W.H., Planta, R.J., Eriksson, A.W. and Frants, R.R. (1987) Hum. Genet. 77, 182–187.
[17] Kageyama, T. and Takahashi, K. (1980) J. Biochem. 88, 635–645.
[18] Kageyama, T., Moriyama, A. and Takahashi, K. (1983) J. Biochem. 94, 1557–1567.
[19] Kageyama, T. and Takahashi, K. (1984) Eur. J. Biochem. 141, 261–269.
[20] Evers, M.P.J., Zelle, B., Bebelman, J.P., Van Beusechem, V., Kraakman, L., Pronk, J.C., Mager, W.H., Planta, R.J., Eriksson, A.W. and Frants, R.R., in preparation.
[21] Evers, M.P.J., Zelle, B., Bebelman, J.P., Pronk, J.C., Mager, W.H., Planta, R.J., Eriksson, A.W. and Frants, R.R. (1988) Gene 65, 179–185.
[22] Kageyama, T. and Takahashi, K. (1980) J. Biochem. 88, 9–16.
[23] Ong, E.B. and Perlmann, G.E. (1968) J. Biol. Chem. 243, 6104–6109.
[24] Pedersen, V.B. and Foltmann, B. (1973) FEBS Lett. 35, 255–256.
[25] Stepanov, V.M., Baratova, L.A., Pugacheva, I.B., Belyanova, L.P. and Timokhina, E.A. (1973) Biochem. Biophys. Res. Commun. 54, 1164–1170.
[26] Harboe, M., Andersen, P.M., Foltmann, B., Kay, J. and Kassell, B. (1974) J. Biol. Chem. 249, 4487–4494.
[27] Baudys, M. and Kostka, V. (1983) Eur. J. Biochem. 136, 89–99.