Volume 182, number 1 FEBS 2334 March 1985

Amino acid sequence of the predominant basic protein in human seminal plasma

H. Lilja and J.-O. Jeppsson

Department of Clinical Chemistry, University of Lund, Malmö General Hospital, S-214 01 Malmö, Sweden

Received 29 October 1984

The predominant basic protein in liquefied human seminal plasma is the major degradation product of the gel-forming protein secreted by the seminal vesicles. The amino acid sequence of this basic protein is presented. The basic protein contains 52 amino acid residues. It is devoid of cysteine, methionine, tryptophan, and leucine, but contains seven histidine residues located in the NH_2-terminal half of the molecule. The calculated M_r of 5753 is in close agreement with that obtained from gel filtration in guanidine-HCl on Sephacryl S-200 ($M_r = 6000$).

Amino acid sequence *Inhibin* *Seminal plasma* *Seminal vesicle*

1. INTRODUCTION

The coagulate structure of freshly ejaculated semen is liquefied within 15 min by prostatic enzyme(s) [1]. It has been proposed that the protein predominant in seminal vesicle secretion is the structural protein of the coagulate [2], and that during liquefaction this protein is cleaved into a number of basic proteins, the major one of which has recently been purified [3]. As estimated using SDS-polyacrylamide gel electrophoresis (SDS-PAGE), the protein's M_r is about 12800. Its NH_2-terminal sequence, comprising 31 identified amino acid residues, is unique according to the Dayhoff data base [3]. Seidah et al. [4] recently presented an identical NH_2-terminal sequence of a basic peptide, also purified from human seminal plasma and consisting of about 35 amino acid residues, and which was found to inhibit secretion of rat pituitary follicle-stimulating hormone. A synthetic replica of this NH_2-terminal sequence has been shown to have a similar inhibitory capacity [5]. Here, we report the complete sequence of the predominant basic protein in human seminal plasma.

2. MATERIALS AND METHODS

2.1. *Chemicals, separation materials, and enzymes*

Sephacryl® S-200 and FPLC® system equipped with an HR 5/5 PepRPC(C_2/C_{18}) column were from Pharmacia (Uppsala); Ultropac® TSK G 2000 SW column from LKB (Bromma, Sweden); diisopropylfluorophosphate (DFP) from Sigma (St Louis, MO); trypsin treated with L-(tosylamide-2-phenyl)ethyl chloromethyl ketone (TPCK) from Worthington (Freehold, NJ). Carboxypeptidase Y was a gift to Professor J. Stenflo from Professor M. Ottesen, Carlsberg Laboratories (Copenhagen, Denmark).

2.2. *Sample, material, purifications, preparation of fragments, and COOH-terminal analysis*

The predominant basic protein was purified from pooled human seminal plasma as in [3]. Lysine-blocked protein was prepared by treatment with citraconic anhydride (30 mol/mol purified protein) according to Atassi and Habeeb [6]. Tryptic digestion of intact or lysine-blocked protein (3 mg/ml) in 0.1 M NH_4HCO_3 was carried out at 37°C for 3 h, at an enzyme:substrate ratio of

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

1:20. TPCK-trypsin was inactivated by DFP (20 mM). The digestion products of the lysine-blocked protein were kept at pH 4 for 3 h at 37°C to deblock the amino groups before separation. Fragments obtained were purified by reversed-phase chromatography (RPC) as described in the legend to fig.2. Intact protein or isolated fragments were digested with carboxypeptidase Y [7] as described in the legend to fig.3.

2.3. M_r *estimation of purified predominant basic protein*

Reduced and carboxymethylated proteins were gel filtered on a Sephacryl S-200 column (1749 × 16 mm) equilibrated in 0.25 M sodium phosphate with 6 M guanidine-HCl, pH 7.4 [8]. HPLC gel filtration in 0.1 M sodium phosphate with 6 M guanidine-HCl, pH 6.9, was performed in a TSK G 2000 SW column at a flow rate of 0.5 ml/min.

2.4. *Amino acid analysis*

Amino acid analysis was performed on a Beckman 119 CL analyser after hydrolysis of fragments in 6 M HCl at 110°C for 24 h [9]. Hydrolysis in 4 M methanesulphonic acid for tryptophan analysis was performed according to Inglis [10].

2.5. *Sequence determination*

Sequence determination by automatic Edman degradations was performed in a Beckman 890 C sequencer. The Beckman no.127974 Quadrol program was used and 2 mg polybrene was added to the spinning cup. PTH amino acids were identified by HPLC [11].

3. RESULTS

3.1. *NH_2-terminal sequence degradation of intact protein*

Automatic Edman degradations of the intact protein (20 and 90 nmol) allowed identification of 31 NH_2-terminal residues and separate residues to position 45 (fig.1).

3.2. *Tryptic fragments*

Tryptic fragments were purified by RPC (fig.2a). Fragments TR0–TR4 were hydrolysed and amino acid analysis showed an equal molar yield for all fragments (65% of theoretical yield).

Fig.1. Amino acid sequence of the predominant basic protein in human seminal plasma. The sequence is presented in the standard IUPAC, one-letter code for amino acid residues. The numbering of tryptic fragments (TR1–TR4) and the numbering of fragments of lysine-blocked protein (ARG1–ARG4) are explained in fig.2. (●) Residues identified by sequenator analysis; (○) residues not identified by sequenator analysis; (⟵) residues identified by carboxypeptidase Y digestion.

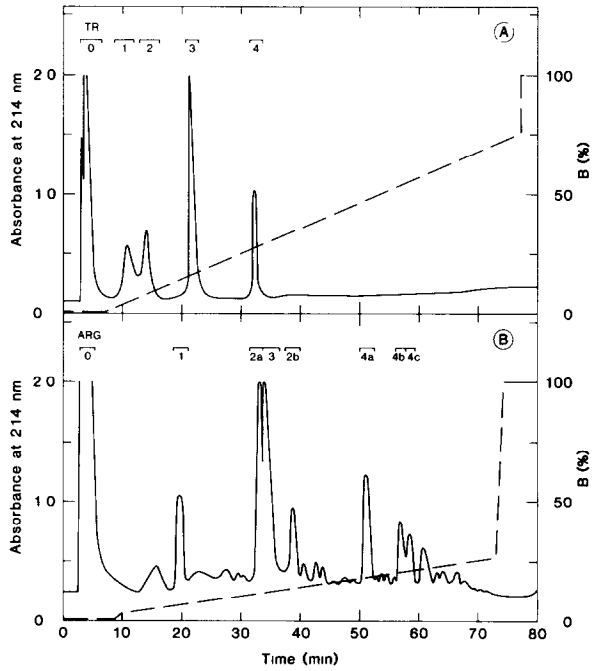

Fig.2. Reversed-phase chromatography on an HR 5/5 Pep RPC (C_2/C_{18}) column at a flow rate of 0.5 ml/min. Elution was performed with a 80 min gradient from 0 to 25% B, or from 0 to 75% B; solvent A, 0.1% (v/v) trifluoroacetic acid in water; solvent B, 80% (v/v) acetonitrile–0.1% (v/v) trifluoroacetic acid; (A) TR fragments obtained from tryptic digestion of intact protein; (B) ARG fragments obtained from tryptic digestion of lysine-blocked protein.

The molar ratios of the amino acids in TR0 corresponded to those of the residues constituting position 1–11 plus 25–30 (fig.1). Sequenator analysis of TR1, TR2, and TR4 allowed identification of all residues. Ten NH_2-terminal residues in TR3 were identified by sequenator analysis.

3.3. *Fragments of lysine-blocked protein*

Fragments obtained by tryptic digestion of citraconylated protein were purified by RPC (fig.2b). ARG0 contained no amino acids. ARG1–ARG4 fragments were obtained in equal molar yields (50% of theoretic yield). Amino acid analysis of ARG1, ARG2a, ARG2b, and ARG3 allowed their identification (fig.1). Sequenator analysis of ARG4a (15 nmol) allowed identification of all residues except those at positions 13 and 20. The coinciding results from amino acid analysis of ARG4a, ARG4b, and ARG4c were consistent with the presence of proline and serine in these positions (fig.1). Carboxypeptidase Y digestion liberated tyrosine as the COOH-terminal residue from ARG4a, 4b, and 4c.

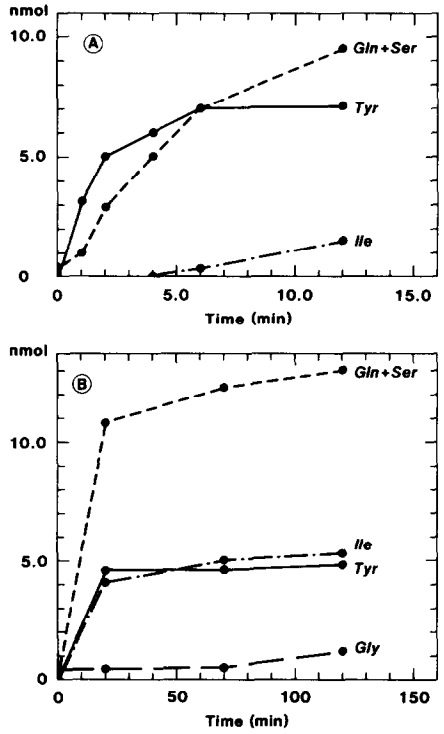

Fig.3. Liberated amino acids after digestion of intact protein with carboxypeptidase Y at an enzyme:substrate ratio of 1:30 or 2:30. Digestion was performed in 0.1 M pyridine-acetate, pH 5.5, at 37°C. Reaction was stopped by adding equal volume of 6% (w/v) sulphosalicylic acid to the vial which was then immediately placed in boiling water and then centrifuged. The separated supernatant was analysed on a Beckman 6300 amino acid analyser; (A) 8 nmol of intact protein digested for 0–12 min; (B) 6 nmol of protein digested for 0–120 min.

3.4. *COOH-terminal analysis of intact protein*

COOH-terminal residues were identified by following the release of amino acids during carboxypeptidase Y digestion of intact protein (fig.3a and b).

3.5. M_r *estimation of purified predominant basic protein*

Gel filtration in 6 M guanidine-HCl on Sephacryl S-200 gave the purified protein an M_r of 6000 (not shown). The protein eluted, however, in two distinct peaks from a TSKG 2000 SW column equilibrated in 6 M guanidine-HCl (not shown).

Table 1

Amino acid composition of the purified predominant basic protein

Amino acid	nmol/ sample	Residue/ mol	Sequence
Aspartic acid	229.9	6.3	6
Threonine	38.6	1.1	1
Serine	198.1	5.4	6
Glutamic acid	228.6	6.2	6
Proline	76.7	2.1	2
Glycine	298.3	8.1	8
Alanine	39.9	1.1	1
Valine	48.4	1.3	2
Isoleucine	67.6	1.8	2
Tyrosine	34.2	0.9	1
Phenylalanine	41.2	1.1	1
Lysine	223.5	6.1	6
Histidine	269.2	7.3	7
Arginine	116.5	3.2	3
Total	1910.7		52

430 μg of intact protein was hydrolysed in 4 M methanesulphonic acid at 110°C for 24 h. Hydrolysate was analysed on the Beckman 6300 analyser. No tryptophan was found in the sample

The first peak eluted with the same retention as carboxymethylated cytochrome *c* (M_r = 12800), and the second peak eluted with the same retention as carboxymethylated aprotinin (M_r = 6400).

3.6. *Complete amino acid sequence of the predominant basic protein in seminal plasma*

The amino acid residues identified by NH_2-terminal analysis in intact protein provide considerable overlap into ARG4. The position of TR4 as the COOH-terminal fragment is well substantiated. The heterogeneity of ARG4 fragments might be explained by deamidation of glutamine residues. All peptides isolated are consistent with the sequence deduced (fig.1). Amino acid analysis of the intact protein after hydrolysis in 4 M methanesulphonic acid (table 1) agrees closely with the sequence presented. The protein contains 52 amino acid residues, from which its M_r is calculated to be 5753.

4. DISCUSSION

The calculated M_r (5753) of the predominant basic protein is in close agreement with that obtained by gel filtration on Sephacryl S-200 in 6 M guanidine-HCl. Analysis with SDS-PAGE has previously given the protein an M_r of 12800, and HPLC gel filtration in 0.1 M ammonium acetate at pH 5.5 an M_r of 15000 [3]. Upon HPLC gel filtration in 6 M guanidine-HCl, the purified protein eluted in two peaks, the first of which eluted at a position corresponding to twice the M_r of the protein sequence. These data suggest the existence of a dimeric form of the purified protein which is remarkably stable under denaturing conditions.

According to the Protein Sequence Database of the Protein Identification Resource (Washington, DC), the sequenced protein has a 47% identity with a histidine-rich portion of the bovine type I precursor of high-M_r kininogen in an overlap of 17 amino acid residues. In an overlap of 42 residues, the predominant basic protein has a 21% identity with the bovine type II precursor of high-M_r kininogen. Furthermore the protein has a 30% identity with human histone 2A in an overlap of 30 residues, and a 34% identity with DNA-dependent RNA polymerase (EC 2.7.7.6) in an overlap of 29 residues. The predominant basic protein in seminal plasma is a major degradation product of a high-M_r precursor secreted by the seminal vesicles [3], and is very labile in seminal plasma [12]. Seidah et al. [4] have purified and partly characterised an inhibin-like basic peptide, whose NH_2-terminal sequence is identical with that presented here, and which is considered to be a degradation product of the predominant basic protein in seminal plasma.

ACKNOWLEDGEMENTS

Our thanks are due to Professor Carl-Bertil Laurell for many invaluable discussions. This research was supported by grants from the Swedish Medical Research Council (project no.B83-03X-00581-19C) and from the Faculty of Medicine, University of Lund, Sweden.

REFERENCES

[1] Tauber, P.F. and Zaneveld, L.J.D. (1976) in: Human Semen and Fertility Regulation (Hafez, E.D.E. ed.) pp.153–166, Mosby, St. Louis.
[2] Lilja, H. and Laurell, C.-B. (1984) Scand. J. Clin. Lab. Invest. 44, 447–452.
[3] Lilja, H., Laurell, C.-B. and Jeppsson, J.-O. (1984) Scand. J. Clin. Lab. Invest. 44, 439–446.
[4] Seidah, N.G., Ramasharma, K., Sairam, M.R. and Chrétien, M. (1984) FEBS Lett. 167, 98–102.
[5] Ramasharma, K., Sairam, M.R., Seidah, N.G., Chrétien, M., Manjunath, P., Schiller, P.W., Yamashiro, D. and Choh Hao Li (1984) Science 223, 1199–1201.
[6] Atassi, M.Z. and Habeeb, A.F.S.A. (1972) Methods Enzymol. 25, 546–553.
[7] Martin, B., Svendsen, I. and Ottesen, M. (1977) Carlsberg Res. Commun. 42, 99–102.
[8] Mann, K.G. and Fish, W.W. (1972) Methods Enzymol. 26, 28–42.
[9] Spackman, D.H., Stein, W.H. and Moore, S. (1958) Anal. Chem. 30, 1190–1206.
[10] Inglis, A.S. (1983) Methods Enzymol. 91, 26–36.
[11] Fohlman, J., Rask, L. and Peterson, P.A. (1980) Anal. Biochem. 106, 22–26.
[12] Lilja, H. and Weiber, H. (1984) Scand. J. Clin. Lab. Invest. 44, 433–438.