PARTIAL STRUCTURAL ANALYSIS OF A HIGHLY BASIC LOW MOLECULAR WEIGHT PROTEIN FROM RAT TESTIS

W.S. Kistler, C. Noyes, and R.L. Heinrikson

The Ben May Laboratory for Cancer Research and Department of Biochemistry
University of Chicago, Chicago, Illinois 60637

Received January 28, 1974

Summary. Automated Edman degradation of a testis-specific basic protein isolated from the rat gave the following NH_2-terminal sequence of amino acids: H_2N-Ser-Thr-Ser-Arg-Lys-Leu-Lys-Thr-His-Gly-Met-Arg-Arg-Gly-Lys-Asx-Arg-Ala-Pro-His-Lys-Gly-Val---.
5 10 15 20
Cleavage of the native protein with cyanogen bromide produced two fragments which were purified by gel filtration. Amino acid analysis of the smaller fragment revealed it to be the NH_2-terminal undecapeptide resulting from cleavage at Met_{11}. The partial sequence analysis of the intact protein coupled with compositional analyses of these cyanogen bromide peptides indicate that the basic testis protein contains 24 basic amino acids and a single methionine in a sequence of 54 amino acids.

A highly basic and organ-specific protein has recently been isolated from the testis of the rat (1). It is associated with the presence of haploid cell types in the germinal epithelium but could not be extracted by a variety of techniques from either epididymal spermatozoa or whole epididymes (1, unpublished observations). The rapid *in vitro* incorporation of labeled amino acids into this protein by teased seminiferous tubules, its distinctive amino acid composition, and other lines of evidence made it appear unlikely that it represented a breakdown product of any known histone (1). To establish this protein more exactly as a molecular entity and to investigate more rigorously its possible relationship to any known chromosomal proteins, we have embarked on a study of its primary structure. The following results provide nearly half of the total sequence from the NH_2-terminus. In addition, the compositional analyses of two fragments resulting from the cleavage of the intact protein with cyanogen bromide indicate that the molecular weight of the polypeptide chain of the protein is 6,200, a value half of that suggested earlier (1).

EXPERIMENTAL PROCEDURES

Purification of the testis-specific basic protein. The procedure described previously (1) was modified slightly to reduce exposure of the protein to strong acids. Specifically,

pooled fractions from chromatography on Sephadex C-25 were desalted by batchwise passage through a column (2.5 x 30 cm) of Bio Gel P-2 equilibrated and developed with 1 M acetic acid. Fractions containing protein were lyophilized, and the residue was dissolved in a small volume of 1 M acetic acid. The solution was then applied to a column (2.6 x 90 cm) of Sephadex G-75 equilibrated and developed with 1 M acetic acid. Fractions of 6 ml were collected, and those containing the single large peak of protein were combined and lyophilized. The final preparation was homogeneous as judged by polyacrylamide gel electrophoresis under the conditions described previously (1).

Cleavage with cyanogen bromide. To a solution containing approximately 1 mg of the purified protein in 1.9 ml of 70% formic acid was added 0.1 ml of a freshly prepared aqueous solution of 2% cyanogen bromide. The mixture was allowed to stand at room temperature in the dark for 24 hrs after which time it was concentrated under vacuum, diluted with 3 ml of water, and lyophilized.

The dried material was dissolved in a small volume of 0.01 M acetic acid and a portion set aside for examination by gel electrophoresis. The remainder was subjected to gel filtration through a column (0.9 x 57 cm) of Sephadex G-75 equilibrated and developed with 0.01 N HCl. Fractions of approximately 1.7 ml were collected and their absorbance at both 230 nm and 275 nm determined. Fractions containing material absorbing light at either of these wavelengths were appropriately pooled and lyophilized (Fig. 1).

Automated Edman degradation and identification of liberated phenylthiohydantoin amino acids. The amino acid sequence of residues at the NH_2-terminus of the intact protein was derived by use of a Beckman protein-peptide Sequencer (model 890-B). The following programs supplied by the manufacturer were used: Cycles 1 to 12, Dimethylallylamine protein program 050771; Cycles 13 to 23, Dimethylallylamine peptide program 090872.

The phenylthiohydantoin amino acids liberated after each cycle of the degradation were identified and quantitated as such or as the trimethylsilyl derivatives by gas chromatography (2) on a Beckman GC-45 unit. Those residues that could not be positively identified by this technique were back converted to the amino acid as described previously (3), with the resulting amino acid identified on the amino acid analyser.

Miscellaneous procedures. The concentration of solutions of the purified protein was determined by the amino acid content of a representative sample, or by assuming from the tyrosine content of the protein (4), that a 1mg/ml solution in dilute acetic acid has an absorbance of 0.48 at 275 nm.

For amino acid analysis samples were hydrolysed with redistilled constant-boiling

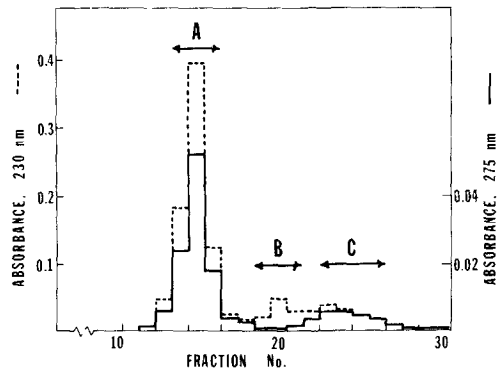

Fig. 1. Gel filtration on Sephadex G-75 of the fragments produced by treatment of the testis-specific basic protein with cyanogen bromide. The fractions indicated A, B, and C were pooled and lyophilized.

HCl for 24 hrs in vacuo at 110 C and the hydrolysates analysed on a Durrum 500 analyser.

RESULTS AND DISCUSSION

The amino acid composition of the testis-specific basic protein purified by the modified procedure reported herein agrees closely with that published earlier (1) and is given in Table 1. The minimal composition, assuming single residues of proline, methionine, and valine, is 54 amino acids. The protein contains no glutamic acid, glutamine, cysteine, isoleucine, tryptophan, or phenylalanine, and basic amino acids comprise 44 moles per cent of the total composition.

Automated Edman degradation (5,6) of 2.6 mg (414 nmoles) of the intact protein was performed as described above (Table 2). Unambiguous identification of the phenylthiohydantoin amino acids liberated at each step could be made until the 23rd cycle; thereafter gaps appeared in the sequence, and the overlap problem prevented confident assignments. Serine was the only NH_2-terminal amino acid identified, and there was no evidence of heterogeneity of the sample at any step of the sequence analysis.

As a first step in the dissection of the intact protein into smaller fragments, cleavage at methionine residues was effected by treatment with cyanogen bromide (7). Electrophoretic examination of a sample of the cleaved protein on gels of 20% acrylamide at pH 4.5 (system 1 of ref 1), revealed, in addition to traces of material indistinguishable from the intact protein, two more rapidly migrating species (mobility 1.1 and 1.7 times that of the intact protein respectively). The staining by amido schwarz of the more rapidly migrating fragment was very unstable, and it characteristically disappeared during the course of the gradual destaining of the gels by diffusion.

Table 1

Amino acid composition of intact testis-specific protein and of fragments resulting from cleavage with cyanogen bromide

Amino Acid	Intact protein		Cyanogen bromide fraction B		Cyanogen bromide fraction A	
Aspartic acid	4.01	(4)	0		3.91	(4)
Threonine	1.97	(2)	1.83	(2)	0.02	
Serine	6.63	(8)	1.64	(2)	5.11	(6)
Proline	0.82	(1)	0		0.90	(1)
Glycine	6.15	(6)	1.04	(1)	5.14	(5)
Alanine	2.14	(2)	0		2.05	(2)
Valine	0.98	(1)	0		0.94	(1)
Methionine	1.00	(1)	0[a]	(1)	0	
Leucine	3.00	(3)	1.00	(1)	2.00	(2)
Tyrosine	2.05	(2)	0		2.07	(2)
Histidine	2.80	(3)	1.01	(1)	1.92	(2)
Lysine	10.08	(10)	2.12	(2)	8.70	(8)
Arginine	11.04	(11)	1.10	(1)	10.30	(10)
		(54)		(11)		(43)

Values are molar ratios with leucine assigned the value 3.00 for the intact protein, the value 1.00 for cyanogen bromide fraction B, and the value 2.00 for cyanogen bromide fraction A. The values for the intact protein represent the averages of the analyses performed on three separate hydrolysates. The values for the cyanogen bromide fractions are the results in each case of a single analysis. The total amount of leucine recovered in the hydrolysate of cyanogen bromide fraction B was 24 nmoles, and in the hydrolysate of cyanogen bromide fraction A, 54 nmoles. The extrapolated values for serine have been corrected for an assumed hydrolytic loss of approximately 15%.
a) A mixture of homoserine and homoserine lactone sufficient to account for one residue of methionine was recovered.

The remainder of the treated protein was subjected to gel filtration (Fig. 1). The lyophilized material from each of the pooled fractions, designated A, B, and C in Fig. 1, was then dissolved in a small volume of 0.01 M acetic acid. Equal portions of each frac-

Table 2

Amino terminal sequence derived by automated Edman degradation of the testis-specific basic protein of the rat

Residue Identified by:	NH_2-Ser	Thr	Ser	Arg	Lys	Leu	Lys	Thr	His	Gly	Met	Arg-
Gas Chromatography	165	+	185		+	143	+	+		66	59	
Back Conversion				115	57		66		78			80

	Arg	Gly	Lys	Asx	Arg	Ala	Pro	His	Lys	Gly	Val---
Gas Chromatography	49					45	25				10
Back Conversion	38		39	45	30			12	10	9	7

The amount in nmoles of each residue recovered is indicated except for those residues where only a qualitative identification (+) could be made.

tion were hydrolysed while a sample of the remaining material was examined by electrophoresis on acrylamide gels as before. Fraction A was found to contain the peptide migrating slightly faster than the intact protein as well as traces of material with mobility identical to that of the intact protein. Fraction B was found to contain only the peptide responsible for the single band with very rapid mobility. Fraction C contained nothing that gave rise to a stained band after electrophoresis.

Compositional analyses of the fractions separated by gel filtration from the cyanogen bromide treated protein indicated that the molecule was indeed cleaved into just two fragments, each of which was recovered in approximately equimolar amounts (Table 1). The composition of fraction B clearly identifies it as the NH_2-terminal fragment generated by cleavage at the methionine residue already located at position 11. The composition and recovery of the fragment in fraction A complements fraction B exactly to generate the composition of the intact protein. As fraction C was found to contain only trace amounts of any amino acid it appears clear that the entire threonine and methionine content of the intact protein is contained in the 11 residue NH_2-terminal fragment. Thus the composition of the intact protein must reflect a polypeptide chain length of a total of 54 residues with a calculated mass of 6,219 daltons.

Previously (1), it was reported that the electrophoretic mobility of the intact protein in polyacrylamide gels containing sodium dodecyl sulfate (SDS) was consistent with a mass of about 13,000 daltons. The current work indicates that the molecular weight of the intact polypeptide chain is about half that amount. This discrepancy might conceivably result from the dimerization of the polypeptide chain, even in the presence of high concentrations of SDS. A more likely explanation is that the number of basic residues in this protein is sufficiently great to render the net negative charge of the SDS-protein complex unusually low for its size.

The sequence of the testis-specific basic protein thus far determined coupled with the composition of the remainder of the protein indicates that it is very unlikely to be derived from degradation of any of the principal somatic histones, even though it is little more than half the size of the smallest of them (8,9,10,11). While it should be emphasized that the basic protein described is not detectable in epididymal spermatozoa (1), it is intriguing that it first appears in the maturing testis after meiosis is completed in most tubules and coincident with the development of the haploid gametes to the stage of nuclear condensation (1). In view of the unpublished observation that the basic protein is associated with chromatin, it is tempting to speculate that it may in some manner be involved with the replacement of somatic histones by the arginine, cyst(e)ine rich protein finally found associated with the DNA in the head pieces of mature spermatozoa from several mammalian species (1,12,13).

ACKNOWLEDGEMENTS

W.S.K. gratefully acknowledges the encouragement and hospitality of Dr. H.G. Williams-Ashman during the course of this work, and we would like to thank Mary E. Geroch for expert assistance in preparing the pure protein. This work was supported by grants HD-04592, HD-07110, and CA-05246 of the U.S. Public Health Service and grant GB-29098 from the U.S. National Science Foundation. R.L.H. was supported by Career Research Development Award 5K05 GM-11164 of the U.S. Public Health Service.

REFERENCES

1) Kistler, W.S., Geroch, M.E., and Williams-Ashman, H.G. (1973) J. Biol. Chem. 248, 4532-4543.
2) Pisano, J.J., and Bronzert, T.J. (1969) J. Biol. Chem. 244, 5597-5607.
3) Heinrikson, R.L., Sterner, R., Noyes, C., Cooperman, B.S., and Bruckmann, R.H. (1973) J. Biol. Chem. 248, 2521-2528.

4) Edelhoch, H. (1967) Biochemistry 6, 1948-1954.
5) Edman, P. (1950) Acta Chem. Scand. 4, 277-283.
6) Edman, P., and Begg, G. (1967) Eur. J. Biochem. 1, 80-91.
7) Gross, E., and Witkop, B. (1962) J. Biol. Chem. 237, 1856-1860.
8) DeLange, R.J., and Smith, E.L. (1971) Ann. Rev. Biochem. 40, 279-314.
9) Greenaway, P.J. (1971) Biochem. J. 124, 319-325.
10) DeLange, R.J., Hooper, J.A., and Smith, E.L. (1973) J. Biol. Chem. 248, 3261-3274.
11) Yeoman, L.C., Olson, J.O.J., Sugano, N., Jordan, J.J., Taylor, C.W., Starbuck, W.C., and Busch, H. (1972) J. Biol. Chem. 247, 6018-6023.
12) Coelingh, J.P., Monfoort, C.H., Rozijn, T.H., Gevers Leuven, J.A., Schiphof, R., Steyn-Parvé, E.P., Braunitzer, G., Schrank, B., and Ruhfus, A. (1972) Biochim. Biophys. Acta 285, 1-14.
13) Monfoort, C.H., Schiphof, R., Rozijn, T.H., and Steyn-Parvé, E.P. (1973) Biochim. Biophys. Acta 322, 173-177.