The partial amino acid sequence of a non-histone chromosomal protein

John M. Walker*, J.R.B. Hastings†, E.W. Johns†

*Unit of Human Cancer Biology, London Branch, Ludwig Institute of Cancer Research, in conjunction with the Royal Marsden Hospital, London, England.

†Chester Beatty Research Institute, Institute of Cancer Research:Royal Cancer Hospital, Fulham Road, London SW3 6JB, England

W. Gaastra

Department of Biochemistry, University of Gröningen , Gröningen, The Netherlands

Received September 8,1976

Summary

The amino acid sequence of the first thirty nine residues of the non-histone chromosomal protein HMG-17 has been determined. Results presented here give a molecular weight of 11,000 for the protein. Some interesting sequence homology with the trout specific histone, histone-T,is noted.

Introduction

The non-histone proteins that are extracted from chromatin with 0.35M NaCl contain a group of proteins which are characterised by their relatively low molecular weights ($<$30,000) (1). These proteins, termed the high mobility group (HMG) proteins, have been fractionated and the main components isolated in a pure form (2). One of these proteins, HMG-17, is the subject of this paper. The amino acid analysis of HMG-17 (2) shows it to have a high content of basic residues (24.3% lysine, 4% arginine) which is similar to the histones, but unlike the histones it has a moderately high acidic amino acid composition (22.5% aspartic and glutamic residues). In addition it has a high alanine, proline and glycine content (42.5% total), but an absence of detectable amounts of cystine, methionine, isoleucine, tyrosine, phenylalanine, histidine and methyl lysine. The absorption spectrum of a solution of HMG-17 shows no absorbance in the 260 to 300 nm range and thus establishes the absence of tryptophan as well as tyrosine and phenylalanine. Results presented here suggest that HMG-17 has a molecular weight of approximately 11,000. As part of our studies on the function of this protein we are determining its primary structure. This paper describes the first thirty nine residues of HMG-17, obtained by sequenator analysis of the native

protein, together with overlap with a peptic peptide of HMG-17. Some interesting sequence homology with the trout specific histone, histone-T , is noted.

Materials and Methods

Isolation of HMG-17

HMG-17 was prepared by a modification of the method described by Sanders (3). All work was carried out at 4°C unless otherwise stated. 1.6 Kg of fresh minced calf thymus tissue was divided into four lots and each blended with 600 ml of 5% perchloric acid for 3 min at full speed (12,000 rev/min) in a domestic blender (Kenwood Kenmix). Each homogenate was centrifuged for 30 min at 2,000 x g, and then each sediment re-extracted with 300 ml of 5% perchloric acid and centrifuged. The combined supernatants were pooled, made 18% with respect to trichloroacetic acid, by the addition of the appropriate volume of 100% trichloroacetic acid, stirred for 10 min, and then centrifuged for 30 min at 2,000 x g. The combined supernatants were passed through a No. 4 sintered glass funnel and then precipitated by the addition of 6 vol. of acetone. The protein precipitate was collected by centrifuging at 2,000 x g for 10 min, washed twice with acetone/conc.HCl (200 ml/0.6 ml) and then three times with acetone before drying under vacuum. 1.0 g of material was recovered. This material was further purified by ion exchange chromatography on CM Sephadex. The sample was dissolved in 7.5 mM borate buffer pH 9.0 (10 ml), titrated to pH 8.8 and then centrifuged for 30 min at 24,000 rpm. The sample was loaded onto a CM Sephadex column (5.0 x 40 cm) equilibrated in borate buffer, followed by a sodium chloride gradient from 0.2M to 2.0M (2 x 1.6 l). The flow rate was 120 ml/hr. 20 ml fractions were collected and read at 220 nm. HMG-17 eluted at approximately 0.4M NaCl and was recovered by precipitation with 6 vol. of acetone, washed twice with acetone/0.1M HCl (6:1 v/v) and then three times with acetone before drying under vacuum. The recovered material had a single N-terminal amino acid (proline) and ran as a single band on polyacrylamide gel electrophoresis at pH 2.2. 90 mg of HMG-17 was recovered.

Pepsin digestion of HMG-17

HMG-17 (5 mg) was dissolved in 5% acetic acid (1 ml), pepsin solution (10 µl of 1 mg/ml soln. ex-hog stomach, from Koch Light) added and the solution incubated at 37°C for 7 hr. 10 µl of solution was then taken for N-terminal analysis, and the rest loaded directly onto chromatography paper for chromatographic purification (see below).

Thrombin digestion of HMG-17

HMG-17 (5 mg) was dissolved in 0.1M ammonium bicarbonate soln.(1 ml) and a soln. of thrombin added (50 µl, 250 NIH units ml, ex-bovine plasma from Sigma). The mixture was incubated at 37°C for 24 hr. 10 µl of the soln. was then taken for N-terminal analysis and the rest of the solution was subjected to chromatographic purification (see below).

Paper chromatography of pepsin and thrombin digests

Both digests were applied as narrow strips on sheets of Whatman No.3 MM at a concentration of 0.1 mg/cm. Papers were eluted by descending chromatography in butanol/pyridine/glacial acetic acid/water (90:60:18:72) for a period of 5 days. Peptide bands were identified by cutting 2 cm

strips from the sides of the chromatography sheets and staining them with ninhydrin. Peptide bands were eluted with 5% acetic acid and lyophilised.

N-terminal amino acid analysis

N-terminal amino acids were determined by the dansyl technique described by Hartley (4).

Amino acid analysis

Total amino acids were measured using a Rank Hilger Chromaspek amino acid analyser. Samples were hydrolysed for 24 hr in 6N HCl at 110°C. No corrections were made for hydrolytic losses.

Automatic sequence determinations

Sequence determinations were carried out on a Beckmann 890C Sequencer, using the standard DMBA buffer programme. The phenylthiohydantoin (PTH) derivatives of the released amino acids were determined by gas chromatography on a Beckmann GC45 gas chromatograph, and by thin layer chromatography. Thin-layer chromatography was carried out on silica gel plates (aluminium coated) with fluorescent indicator (Merck) in a chloroform-methanol (90:10 v/v) mixture (5). Fluorescent patterns were viewed at 254 nm. Additional information on the identity of the PTH amino acid residues was obtained by spraying the silica gel plates with a solution of ninhydrin (100 mg) and collidine (5 ml) in absolute ethanol (95 ml). After heating for 5 min at 150°C many of the PTH derivatives produced characteristic colours (6). PTH derivatives of arginine which remained in the aqueous phase after conversion were identified by staining with phenanthrenequinone (7). The chemicals used for the phenylisothiocyanate degradation were sequenal grade products from Beckmann (Palo Alto), or from Pierce Chemical Co. (Rockford). For sequence determinations on the native protein, 5 mg (~500 nmoles) of HMG-17 was used. For peptide PA (see below), 1.8 mg (~200 nmoles) was used.

Results and Discussion

Pepsin digestion

N-terminal analysis of the pepsin digest showed the presence of N-terminal serine as well as the original N-terminal proline. Paper chromatography gave two bands, one remaining at the origin (peptide PA) and one moving approximately 2 cm from the origin (peptide PB). Peptide PB had the N-terminal amino acid proline, presumably deriving from the N-terminal portion of the molecule. Peptide PA had the N-terminal amino acid serine and was used in the sequenator studies described here. The amino acid analyses of peptides PA and PB are shown in Table 1.

Thrombin digest

N-terminal analysis of the thrombin digest showed the presence of mainly N-terminal serine and proline, but traces of a number of other N-terminal amino acids were also present. Paper chromatography of this digest

Table 1 Amino acid compositions of HMG-17 and peptides isolated from pepsin and thrombin cleavages

Amino acid	HMG-17	PA		PB		TB	
	moles %	moles %	moles per mole	moles %	moles per mole	moles%	moles per mole
Aspartic acid	12.0	11.4	7.8	11.6	3.2	12.8	2.9
Threonine	1.2	1.3	0.9	-	-		-
Serine	2.3	1.3	0.9	3.3	0.9		-
Glutamic acid	10.5	8.7	6.0	10.7	2.9	12.1	2.8
Proline	12.9	14.4	9.9	6.5	1.8	8.8	2.0
Glycine	11.2	12.5	8.7	6.7	1.8	8.8	2.0
Alanine	18.4	19.0	13.0	16.0	4.3	13.8	3.1
Valine	2.0	1.8	1.2	3.7	1.0	3.8	0.9
Cystine	-	-	-	-	-	-	-
Methionine	-	-	-	-	-	-	-
Isoleucine	-	-	-	-	-	-	-
Leucine	1.0	-	-	3.2	0.9	-	-
Tyrosine	-	-	-	-	-	-	-
Phenylalanine	-	-	-	-	-	-	-
Lysine	24.3	29.5	20.3	22.5	6.1	26.5	6.0
Histidine	-	-	-	-	-	-	-
Arginine	4.1	-	-	15.5	4.3	13.3	3.0
Total residues			69		27		23
N-terminal	Proline	Serine		Proline		Proline	

VAL 5 10 15
PRO-LYS-ARG-LYS-ALA-GLU-GLY-ASP-ALA-LYS-GLY-ASP-LYS-ALA-LYS-VAL-LYS-
ALA
20 25 30
ASP-GLU-PRO-GLN-ARG-ARG-SER-ALA-ARG-LEU-SER-ALA-LYS-PRO-ALA-PRO-PRO-
35 39
LYS-PRO-GLU-PRO-LYS-(ASX_8, GLX_5, THR, PRO_5, GLY_9, ALA_{11}, VAL, LYS_{17})

Fig. 1. Partial amino acid sequence of HMG-17

showed one band remaining at the origin (peptide TA) and one band moving about 2 cm from the origin (peptide TB). Peptide TB had the N-terminal amino acid proline, presumably deriving from the N-terminal portion of the molecule. Peptide TA had mainly the N-terminal amino acid serine but traces of other N-terminal amino acids were present. Peptide TA was not studied further. The amino acid analysis of peptide TB is shown in Table 1.

Sequenator results

A sequence of 34 residues was obtained from the sequence studies on native HMG-17. A further sequence of 13 residues was obtained from the sequenator studies on peptide PA. These two peptides overlapped to give the first thirty nine residues on HMG-17 shown in fig. 1. Peptide PA started at residue 28, thereby unambiguously overlapping the 34 residues obtained from the native HMG-17 by 7 residues. The amino acid composition of the remaining portion of the molecule shown in fig. 1 was obtained by difference from the amino acid analysis of peptide PA. The sequence data, together with the peptide analysis suggests that HMG-17 contains about 96 amino acid residues. At one position, residue 3, microheterogeneity was observed. Although arginine was obviously the main residue at this position, valine and alanine were also detected at this position. This observation, made on the sequenator run on native HMG-17, was confirmed by manual Edman degradation of peptide PB. It was possible to determine the first nine residues by the manual procedure, and again at position 3 both alanine and valine were detected in addition to arginine. All other positions showed a single amino acid residue.

Peptic cleavage obviously occurred C-terminal to the leucine residue at position 27. The sequence of the first 27 residues is in excellent agreement with the amino acid analysis of peptide PB. From the amino acid analysis of peptide TB is can be seen that thrombin cleavage occurred at the arginine-serine linkage at residue 23. The presence of mainly N-terminal serine in peptide TA agrees with this cleavage, although

```
                          5                  10                  15
HMG-17      PRO-LYS-ARG-LYS-ALA-GLU-GLY-ASP-ALA-LYS-GLY-ASP-LYS-ALA-LYS-VAL-LYS-

Histone-T   PRO-LYS-ARG-LYS-SER-ALA-THR-LYS-GLY-
                          5

                    20                  25                  30                  35
HMG-17      ASP-GLU-PRO-GLN-ARG-ARG-SER-ALA-ARG-LEU-SER-ALA-LYS-PRO-ALA-PRO-PRO-LYS-PRO-

Histone-T   ASP-GLU-PRO-ALA-ARG-ARG-SER-ALA-ARG-LEU-SER-GLY-ARG-PRO-VAL-PRO-LYS-PRO-ALA-
            10                  15                  20                  25
```

Fig. 2. Sequence homology between HMG-17 and Histone-T

the presence of traces of other N-terminal amino acids in peptide TA suggests that there were other minor cleavage positions. Thrombin cleavage at arginine-serine bonds has been previously reported (8).

From the sequence data presented here and the amino acid composition of the rest of the molecule it could appear that there is no irregular distribution of charge in the molecule. This is in contrast to the non-histone chromosomal proteins HMG-1 and HMG-2 where we have previously shown the charge distribution to be highly irregular (9). However, one interesting feature of the sequence presented here is the presence of 5 proline residues between residues 30 and 39. This high number of proline residues might explain why a sequence of only 13 residues could be obtained with peptide PA, since prolyl residues are known to couple with some difficulty in the Edman degradation. It is interesting to note that the four arginine residues in the molecule are all located in the first 26 minutes.

Comparison with the known sequences of the mammalian histones shows only one short sequence homology. The tetrapeptide ARG-ARG-SER-ALA at position 22 to 25 is present in chicken erythrocyte histone V at position 20 to 23. However, comparison with the published N-terminal sequence of the trout specific histone, histone-T, (10), shows considerable sequence homology. Fig. 2 compares the first 28 residues of histone-T with the first 36 residues of HMG-17. It can be seen that the first four residues are the same, and then by aligning residue 18 of HMG-17 with residue 10 of histone-T the following 16 residues can be seen to differ in only four positions. Of the four differences that exist, three of them are conservative differences; alanine (residue 29 in HMG-17) to glycine, lysine (residue 30) to arginine and alanine (residue 32) to valine, the fourth difference, glutamine (residue 21) to alanine is non-conservative. This homologous region of the two proteins presumably indicates the presence of the same primordial gene in the genes for both proteins. It is inter-

esting to note that histone-T shows similarities with HMG-17 in that it also does not contain the amino acids, cystine, methionine, isoleucine, tyrosine, phenylalanine, histidine or tryptophan.

Acknowledgements

We wish to thank Mrs. R. Smith and Miss S.J. Pelly for excellent technical assistance. Financial support from the Medical Research Council (J.R.B.H. and E.W.J.), the Ludwig Institute of Cancer Research (J.M.W.) and the Netherlands Organization for the Advancement of Pure Research (W.G.) are gratefully acknowledged. J.M.W. wishes to thank the European Molecular Biology Organization for a short-term EMBO fellowship.

References

1. Johns, E.W., Goodwin, G.H., Walker, J.M. and Sanders, C. (1975) in: The Structure and Function of Chromatin, Ciba Foundation Symposium, No.28, pp 95-108, Associated Scientific Publishers, Amsterdam.
2. Goodwin, G.H., Nicolas, R.H. and Johns, E.W. (1975) Biochim.Biophys.Acta 405, 280-291
3. Sanders, C. Ph.D. Thesis, London 1975
4. Hartley, B.S. (1970) Biochem.J. 119, 805-822
5. Wittmann-Liebold, B. (1973) Hoppe-Seyler's Z.Physiol.Chem. 354, 1415-1431
6. Roseau, G. and Pantel, P. (1969) J.Chromatog. 44, 392-395
7. Yamada, S. and Ilano, H.A. (1966) Biochim.Biophys.Acta 130, 538-570
8. Graf, L., Barat, E., Borvendeg, J., Hermann,I. and Patthy, A. (1976) Eur.J.Biochem. 64, No. 2, 333-340
9. Walker, J.M., Shooter, K.V., Goodwin, G.H. and Johns, E.W. (1976) Biochem.Biophys.Res.Commun. Vol.70, No. 1, 88-93
10. Huntley, G.H. and Dixon, G.H. (1972) J.Biol.Chem. 247, 4916-4919