THE PRIMARY STRUCTURE OF NON-HISTONE CHROMOSOMAL PROTEIN HMG17 FROM CHICKEN ERYTHROCYTE NUCLEI

J. M. WALKER, C. STEARN and E. W. JOHNS

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

1. Introduction

Chromatin contains a group of non-histone proteins called the high mobility group (HMG) proteins (reviewed [1]). There are 4 main HMG proteins in calf thymus, HMG 1, 2, 14 and 17, which have all been shown to be present in isolated nucleosomes [2]. Some minor HMG protein components of calf thymus chromatin have also been characterised [3]. Attention has been focused on HMG14 and 17 because of their possible role in gene transcription [4,5]. The complete amino acid sequences of calf thymus HMG14 and 17 have been determined [6,7] and partial sequences are available for HMG1 and 2 [1,8]. HMG proteins have been shown to be present in a variety of organisms and tissues, including avian erythrocytes [9,10] trout testis and liver [11,12] and wheat and yeast [13] but little is known about the structures of the HMG proteins from these sources. One of the questions that needs to be answered concerning the HMG proteins is whether they exhibit the same extreme evolutionary stability as shown by the nucleosome core histones. As part of a programme of study on the structure of HMG proteins from various species we have now determined the complete amino acid sequence of HMG17 from chicken erythrocyte nuclei.

2. Experimental

2.1. *Isolation of HMG17*
Chicken erythrocyte HMG17 was prepared as in [14].

2.2. *Proteolytic cleavage of HMG17 and peptide purification*
HMG17, and peptide P1 derived from HMG17, were subjected to enzymatic digestion with the following enzymes:
(i) Pepsin digestion of HMG17 was carried out on 15 mg protein (5 mg/ml in 1 N acetic acid) at an enzyme:protein ratio of 1:50 (w/w) at 37°C for 4 h. The cleavage products (designated P) were loaded directly onto Whatman 3MM paper and separated by descending paper chromatography in butanol:pyridine:acetic acid:water (15:10:1:12, by vol.) for 5 days. Peptide P1 remained at the origin while peptide P2 moved ~2 cm.
(ii) Peptide P1, produced as in (i) above, was further cleaved with V8 protease (Staphylococcal protease) and trypsin. V8 protease digestion of 8 mg peptide P1 (10 mg/ml in 0.2 M ammonium bicarbonate) was carried out at an enzyme:protein ratio of 1:40 (w/w) at 37°C for 4 h. Cleavage products (designated V) were separated by paper chromatography for 3 days as for the pepsin digest. Tryptic digestion of 8 mg peptide P1 (10 mg/ml in 0.2 M ammonium bicarbonate) was carried out at an enzyme:substrate ratio of 1:50 (w/w) at 37°C for 24 h. Cleavage products (designated T) were separated by paper chromatography for 1 day, and where necessary, further purified by high-voltage paper electrophoresis at pH 6.5.

### 2.3. *Peptide sequence determinations*
Automated Edman degradations were carried out on a Beckman 890C protein sequencer using a 0.1 M quadrol buffer programme with a double-cleavage

step on each cycle essentially as in [15]. Polylbrene (5 mg) was used as a carrier and taken through 3 cycles of the Edman degradation together with 100 nmol glycyl glycine prior to each sequenator run [15]. PTH derivatives of released amino acids were determined both directly by high pressure liquid chromatography (HPLC) and indirectly, by back hydrolysis to the free amino acid. HPLC was carried out on a DuPont 830 liquid chromatogram using a Partisil PX5 ODS column (Whatman). PTH amino acids were eluted with a linear gradient of acetonitrile from 15–48% in 0.01 M sodium acetate buffer (pH 4.5) over 7 min, then holding at 48% acetonitrile for a further 4 min. Eluted PTH amino acids were identified by their A_{269}. Back hydrolysis of PTH amino acids was carried out in 65% hydriodic acid at $110°$ C for 24 h. Liberated amino acids were identified on a Rank-Hilger Chromaspek amino acid analyser.

3. Results and discussion

Residues 1–38 had been determined by sequenator

analysis of the native protein [14]. Cleavage of HMG17 with pepsin gave 2 peptides (P1 and P2), the amino acid analyses of which are given in table 1. Peptide P2 represents residues 1–27. Peptide P1 is residue 28 to the end of the molecule. Sequenator analysis of P1 allowed the determination of residues 28–65. V8 protease cleavage of peptide P1 gave 3 peptides (V1–V3) which are numbered in increasing chromatographic mobility. Peptide V1 was residues 50–83, cleavage at glutamic acid residues at position 64 and 80 not occuring. Sequenator analysis of peptide V1 gave the total peptide sequence, although some uncertainty existed over the last 5 or 6 residues. These residues were confirmed by the isolation of a tryptic peptide (T1) from V1 which gave the sequence:

Thr–Asn–Gln–Ala–Glu–Lys
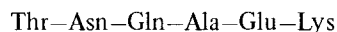
confirming residues 76–81. Further extensive cleavage of peptide V1 with V8 protease (1:50 enzyme:proteins, 24 h) resulted in the liberation of the peptide Lys–Ala–Glu from the C-terminus of V1, confirming

Table 1
Amino acid composition and N-terminal analyses of the major pepsin and V8
protease peptides obtained from HMG17

Amino acid	P1		P2		V1	V2	V3
Asp	6.4	(7)	2.8 (3)		5.9 (6)	–	0.8 (1)
Thr	1.3	(1)	1.0 (1)		1.3 (1)	–	–
Ser	2.1	(2)	1.0 (1)		– (0)	1.9 (2)	–
Glu	7.0	(7)	2.8 (3)		5.0 (5)	2.2 (2)	–
Pro	9.8	(9)	2.3 (2)		2.2 (2)	6.7 (7)	–
Gly	6.7	(7)	2.1 (2)		5.1 (5)	–	2.1 (2)
Ala	11.7	(12)	3.1 (3)		6.3 (6)	4.0 (4)	2.1 (2)
Cys	–	(0)	– (0)		–	–	–
Val	1.0	(1)	1.0 (1)		0.9 (1)	–	–
Met	–	(0)	– (0)		–	–	–
Ile	–	(0)	– (0)		–	–	–
Leu	–	(0)	0.9 (1)		–	–	–
Tyr	–	(0)	– (0)		–	–	–
Phe	–	(0)	– (0)		–	–	–
His	–	(0)	– (0)		–	–	–
Lys	16.7	(16)	6.1 (6)		8.3 (8)	7.0 (7)	1.1 (1)
Arg	–	(0)	3.8 (4)		–	–	–
N-terminus	Ser		Pro		Lys	Ser	Gly
Total residues	62		27		34	22	6
Position in sequence	28–89		1–27		50–83	28–49	84–88

The composition of each peptide is given as the molar ratios of the amino acids.
No corrections were made for hydrolytic losses. Figures in parenthesis are the
number of residues found

residues 81–83. Peptide V2 was residues 28–49, the total sequence of which confirmed part of the N-terminal sequence of P1. Peptide V3 was the only peptide not to have C-terminal glutamic acid and was placed at the C-terminus of the molecule. Peptides produced by V8 protease cleavage of peptide P1 were therefore aligned in the order V2-V1-V3. The amino acid analyses of peptides V1–V3 are shown in table 1.

The complete amino acid sequence of chicken erythrocyte HMG17 is shown in fig.1. Comparison with the published sequence of calf thymus HMG17 shows only 5 differences between the 2 molecules. These are (calf thymus residue first) Ala ↔ Thr (res. 9), Pro ↔ Ala (res. 44), Ala ↔ Pro (res. 45), Gly ↔ Ser (res. 48) and Asp ↔ Glu (res. 64). These changes are all essentially conservative changes (the changes at positions 44 and 45 are simply an inversion of the sequence Ala–Pro) and do not affect the overall architecture of the molecule. There is, therefore, a 5% sequence variation in the HMG17 protein sequence for the 2 species. A sequence variation of 5% represents little evolutionary variation when compared, for

example, with the variation of 28% that exists between the sequence of α globin from calf and chicken. However, this sequence variation is greater than that observed for the nucleosome core histones. The sequence of H3 differs in only one position in 135 residues between the chicken and calf proteins [16], and the sequence of calf histone H2B differs from that of chicken histone H2B at 4 positions in 125 residues [17]. Comparable data (chicken versus calf) for histones H4 and H2A are not available, but it is known that the evolutionary rate of histone H3 is similar to that of histone H4, and that of histone H2B similar to that of histone H2A. It therefore appears that HMG17 does not exhibit the degree of sequence conservation that is known to exist for the 4 nucleosome core histones.

During the course of our studies we have isolated HMG17 from chicken thymus (unpublished). The amino acid analysis of this protein is identical to that of the chicken erythrocyte protein, and the sequence of the first 35 residues of the chicken thymus protein is identical to that of the erythrocyte protein. In

20
Pro-Lys-Arg-Lys-Ala-Glu-Gly-Asp-Thr-Lys-Gly-Asp-Lys-Ala-Lys-Val-Lys-Asp-Glu-Pro-
←————————————————————————P2———————————————————————————

40
Gln-Arg-Arg-Ser-Ala-Arg-Leu-Ser-Ala-Lys-Pro-Ala-Pro-Pro-Lys-Pro-Glu-Pro-Lys-Pro-
————————P2————————→ ←————————————P1————————————
←————————————————V2————————— 60
Lys-Lys-Ala-Ala-Pro-Lys-Lys-Ser-Glu-Lys-Val-Pro-Lys-Gly-Lys-Lys-Gly-Lys-Ala-Asp-
————————————————————P1————————————————————

————————V2———————————→ ←——————————V1—————————— 80
Ala-Gly-Lys-Glu-Gly-Asn-Asn-Pro-Ala-Glu-Asn-Gly-Asp-Ala-Lys-Thr-Asn-Gln-Ala-Glu-
————————P1———————— — — — —
—————————————————————————V1—————————————————————
Lys-Ala-Glu-Gly-Ala-Gly-Asp-Ala-Lys
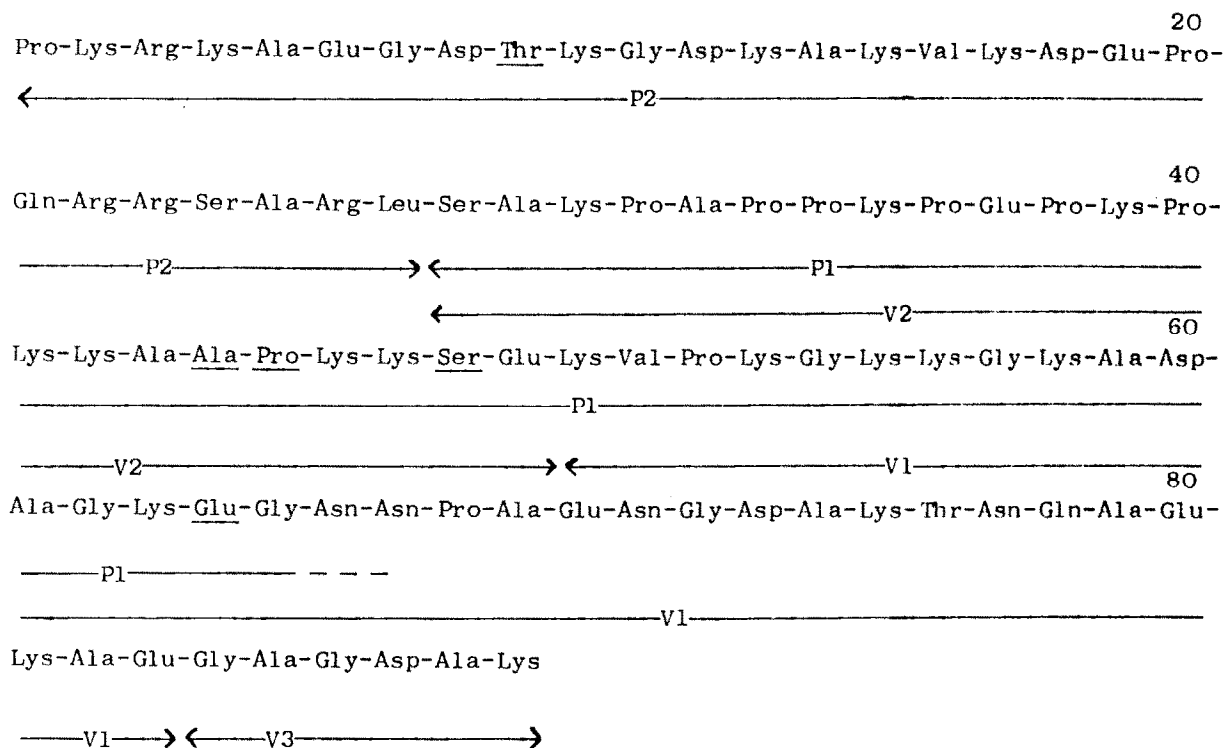
————V1———→ ←————V3—————————————→

Fig.1. The amino acid sequence of chicken erythrocyte HMG17. Positions which differ from the calf thymus sequence are underlined.

particular residue 9 is threonine, which differs from that of the calf sequence. It seems reasonable to assume, therefore, that the sequence of the chicken erythrocyte protein is representative of HMG17 from chicken tissue in general and is not peculiar to the erythrocyte.

Acknowledgements

References

[1] Walker, J. M., Goodwin, G. H., Smith, B. J. and Johns, E. W. (1980) in: Comprehensive Biochemistry, vol. 19B (Neuberger, A. ed) Elsevier/North-Holland, Amsterdam, New York.

[2] Goodwin, G. H., Woodhead, L. and Johns, E. W. (1977) FEBS Lett. 73, 85–88.

[3] Goodwin, G. H., Brown, E., Walker, J. M. and Johns, E. W. (1980) Biochim. Biophys. Acta in press.

[4] Levy, W. B. and Dixon, G. H. (1978) Nucleic Acid Res. 5, 4155–4163.

[5] Weisbrod, S. and Weintraub, H. (1979) Proc. Natl. Acad. Sci. USA 76, 630–634.

[6] Walker, J. M., Goodwin, G. H. and Johns, E. W. (1979) FEBS Lett. 100, 394–398.

[7] Walker, J. M., Hastings, J. R. B. and Johns, E. W. (1977) Eur. J. Biochem. 76, 461–468.

[8] Walker, J. M., Gooderham, K. and Johns, E. W. (1979b) Biochem. J. 181, 659–665.

[9] Rabbani, A., Goodwin, G. H. and Johns, E. W. (1978) Biochem. Biophys. Res. Commun. 81, 351–358.

[10] Sterner, R., Boffa, L. C. and Vidali, G. (1978) J. Biol. Chem. 253, 3830–3836.

[11] Watson, D. C., Peters, E. H. and Dixon, G. H. (1977) Eur. J. Biochem. 74, 53–60.

[12] Rabbani, A., Goodwin, G. H., Walker, J. M., Brown, E. and Johns, E. W. (1980) 109, 294–298.

[13] Spiker, S., Mardian, J. K. W. and Isenberg, I. (1978) Biochem. Biophys. Res. Commun. 82, 129–135.

[14] Walker, J. M. and Johns, E. W. (1980) Biochem. J. 185, 383–386.

[15] Hunkapiller, M. W. and Hood, L. E. (1978) Biochemistry 17, 2124–2129.

[16] Brandt, W. F. and Von Holt, C. (1974) Eur. J. Biochem. 46, 419–429.

[17] Van Holden, P., Strickland, W. N., Brandt, W. F. and Von Holt, C. (1978) Biochim. Biophys. Acta 533, 278–281.