Rapid Paper

(Pages 253–255)

The Isolation, Characterization and Partial Sequence of a Peptide Rich in Glutamic Acid and Aspartic Acid (HGA-2 Peptide) from Calf Thymus Non-Histone Chromosomal Protein HMG 2

COMPARISON WITH A SIMILAR PEPTIDE (HGA-1 PEPTIDE) FROM CALF THYMUS
NON-HISTONE CHROMOSOMAL PROTEIN HMG 1

By JOHN M. WALKER, KEITH GOODERHAM and ERNEST W. JOHNS
*Chester Beatty Research Institute, Institute of Cancer Research: Royal Cancer Hospital,
Fulham Road, London SW3 6JB, U.K.*

(Received 11 January 1979)

A 41-residue peptide (HGA-2) containing a continuous sequence of 35 glutamic and aspartic residues was isolated from non-histone chromosomal protein HMG 2. This highly acidic peptide is compared with a similar peptide (HGA-1) isolated from non-histone chromosomal protein HMG 1.

Chromatin contains a group of non-histone proteins called the HMG (high-mobility-group) proteins (Johns *et al.*, 1975). There are four main HMG proteins in thymus, HMG 1, HMG 2, HMG 14 and HMG 17, which have all been shown to be present in isolated nucleosomes (Goodwin *et al.*, 1977a). All four of these proteins have been isolated in a pure form from both pig and calf thymus (Goodwin *et al.*, 1975, 1977b; Sanders, 1975), and an HMG protein from trout testis, HMG-T, has also been isolated (Watson *et al.*, 1977). We have determined the complete amino acid sequences of proteins HMG 17 (Walker *et al.*, 1977) and HMG 14 (J. M. Walker, unpublished work), and as a further part of our characterization of the high-mobility-group proteins we are determining the amino acid sequences of proteins HMG 1 and HMG 2. One of the most interesting features of both proteins HMG 1 and HMG 2 is the fact that over 50% of their amino acid residues are charged. Like the histones, 25% of their residues are basic. However, unlike the histones the proteins also contain 30% acidic amino acids. Because of the quantities of both these proteins present in the nucleus (approx. 10^6 molecules of each protein), we consider that, like the histones, proteins HMG 1 and HMG 2 are structural proteins, possibly involved in the higher ordered structure of the nucleosomes, and are probably not involved in specific gene control. We have shown protein HMG 1 to have both a basic *N*-terminal region and a central hydrophobic region (Walker *et al.*, 1978b), which is similar to the situation found with the four nucleosome 'core' histones. However, during our work on protein HMG 1 we have isolated a most unusual peptide from the *C*-terminal region of this protein,

which provides HMG 1 with a structural feature unlike that found in any of the histones. This peptide contains a continuous amino acid sequence of 41 glutamic and aspartic residues, and because of the high glutamic and aspartic content of this peptide we refer to it as HGA-1 peptide (Walker *et al.*, 1978a). We now report the isolation and characterization of a similar, but not identical, peptide (HGA-2) from calf thymus non-histone chromosomal protein HMG 2.

Experimental

Protein HMG 2 was prepared from calf thymus as described previously (Johns *et al.*, 1975).

Amino acid analysis

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

Tryptic digestion of protein HMG 2

Protein HMG 2 (10 mg) was dissolved in 0.2M-NH₄HCO₃ solution (5 mg/ml) and incubated with trypsin (1:50, w/w) at 37°C for 48 h.

Isolation of the HGA-2 peptide

The trypsin-digested protein HMG 2 was loaded directly as a thin strip on to sheets of Whatman 3 MM chromatography paper (200 µg of protein/cm) and submitted to continuous descending chromatography in butanol/pyridine/acetic acid/water (15:10:1:2, by vol.) for 5 days. During this time all the

peptides from the tryptic digest moved considerably down the chromatography sheet, with the exception of the HGA-2 peptide, which only moved 3 cm from the point of origin. This peptide was eluted with 5% (v/v) acetic acid and freeze-dried.

Identification of *N*-terminal amino acids

N-Terminal amino acids were determined by the dansyl method as described previously (Walker *et al.*, 1977).

Sequenator analysis of the peptide [produced by tryptic cleavage of succinylated (3-carboxypropionylated) protein HMG 2] that contains the HGA-2 sequence

Sequenator analysis of this peptide was carried out on 200 μ mol of peptide by using a Beckman 890C sequenator with the standard DMAA (dimethylallylamine buffer) peptide programme. Amino acid phenylthiohydantoin were identified both qualitatively by t.l.c. and quantitatively by back hydrolysis to the free amino acid, as described previously (Walker *et al.*, 1978b). A repetitive yield of 94% was obtained for the run. The sequence given in Fig. 2 for the HGA-2 sequence represents residues 20–34 in this peptide. Because of the repetitive nature of the sequence in this region, unambiguous identification of residues beyond this point was not possible.

Results and Discussion

Peptide HGA-1 may be prepared from protein HMG 1 by the identical procedure as described for peptide HGA-2, when the HGA-1 peptide moves 1 cm from the point of origin during 5 days chromatography. Our previous report on the isolation of peptide HGA-1 (Walker *et al.*, 1978a) required the

prior isolation of the *C*-terminal CNBr peptide for isolation of peptide HGA-1. The method detailed here therefore offers a more rapid and convenient method for isolating both peptides HGA-1 and HGA-2 directly from the parent molecule;

The HGA-2 peptide as isolated by paper chromatography has a single *N*-terminal amino acid, lysine, as determined by the dansyl method. Amino acid analysis of peptide HGA-2 showed the presence of four amino acids, glutamic acid, aspartic acid, lysine and proline, in the proportions Lys₃, Glu₂₉, Asp₇, Pro₂. Comparison of this amino acid analysis with that obtained previously for peptide HGA-1 (Walker *et al.*, 1978a) is shown in Fig. 1. Both peptides contain the same number of glutamic residues (29 residues each), but HGA-2 has a lower aspartic content (7 residues) than HGA-1 (12 residues). The most significant difference between the two peptides, however, is the presence of two proline residues in HGA-2 which are absent from HGA-1. The presence of these proline residues should confer a significant structural difference between the two peptides, since a proline residue will necessarily induce a bend in the polypeptide chain. We have previously reported the amino acid sequence of the first 12 residues of peptide HGA-1 (Walker *et al.*, 1978a). We have also recently isolated, by tryptic cleavage of succinylated protein HMG 2, a peptide (approx. 80 residues) containing 55% aspartic and glutamic residues (J. M. Walker, unpublished work). This peptide contains the HGA-2 peptide and we have determined, by sequenator analysis of this peptide, the sequence of the residues before peptide HGA-2 as well as the first 16 residues of peptide HGA-2. The sequence of the latter is shown in Fig. 2 together with the *N*-terminal sequence of peptide HGA-1. The results show that the HGA region in both peptides is preceded by a group of at least three basic amino acids (the sequence of peptide HGA-1 would be preceded by a basic amino acid, since the peptide was produced by tryptic cleavage). The HGA peptide cannot be the *C*-terminal tryptic peptide of protein HMG 2, since we have previously shown the *C*-terminal sequence of protein HMG-2 to

HGA-1	Glu ₂₉ Asp ₁₂ Lys ₃
HGA-2	Glu ₂₉ Asp ₇ Lys ₃ Pro ₂

Fig. 1. Amino acid compositions of the HGA peptides isolated from proteins HMG 1 and HMG 2

HGA-1	Lys-Lys-Glu-Glu-Glu-Glu-Asp-Glu-Glu-Asp-Asp-Glu- . . (Glu ₂₂ , Asp ₉ , Lys ₁)
HGA-2	Lys-Lys-Asn-Glu-Pro-Glu-Asp-Glu-Glu-Glu-Glu-Glu-Glu-Asp- (Glu ₂₀ , Asp ₄ , Pro ₁ , Lys ₁)

Fig. 2. Comparison of the amino acid sequences of the *N*-terminal regions of the HGA peptides from proteins HMG 1 and HMG 2

be Phe-Ala-Lys (Walker *et al.*, 1976), and neither phenylalanine nor alanine is present in the analysis of peptide HGA-2. Therefore, since both peptides are produced by tryptic cleavage, they must necessarily both have a lysine residue at the C-terminal. In view of the high proportion of glutamic acid in both peptides it is not surprising that there is sequence homology between the glutamic residues in the N-terminal regions of the HGA peptides. One of the two proline residues in peptide HGA-2 is located in the N-terminal part of this molecule. The position of the second proline residue can only be determined when the complete sequence of peptide HGA-2 is known. It seems likely that, as for peptide HGA-1, all the aspartic and glutamic residues in peptide HGA-2 are present as the acids and not amides (with the exception of the single asparagine residue detected towards the N-terminal of the HGA-2 sequence; see Fig. 2). We have previously shown that only eight of the aspartic and glutamic residues in protein HMG 2 are present as amides (Walker *et al.*, 1976) and we have already found this number of amides in peptides from other regions within the HMG 2 sequence. Also, the sequenator run on the succinylated peptide was taken to 25 residues beyond the start of the HGA-2 peptide, and although an unambiguous sequence of aspartic and glutamic residues was not possible throughout this region, all acidic residues and no amides were detected by t.l.c. on all of these cycles.

The HGA region in both proteins HMG 1 and HMG 2 must obviously play an important role in the function of these proteins, but the exact structural significance of this HGA region has yet to be determined. At present our sequence results for the

high-mobility-group proteins represents a large part of the available sequence data for mammalian non-histone chromosomal proteins. It will therefore be interesting to see if this feature of a highly negatively charged region, which we have now shown to be present in two chromosomal proteins, proves to be a common feature of non-histone chromosomal proteins.

Financial support for J. M. W. and E. W. J. from the Medical Research Council is gratefully acknowledged. K. G. acknowledges the Institute of Cancer Research for a postgraduate studentship. We thank Dr. J. R. B. Hastings for carrying out the amino acid analyses.

References

- Goodwin, G. H., Nicolas, R. H. & Johns, E. W. (1975) *Biochim. Biophys. Acta* **405**, 280–291
- Goodwin, G. H., Woodhead, L. & Johns, E. W. (1977a) *FEBS Lett.* **73**, 85–88
- Goodwin, G. H., Rabbani, A., Nicolas, R. H. & Johns, E. W. (1977b) *FEBS Lett.* **80**, 413–416
- Johns, E. W., Goodwin, G. H., Walker, J. M. & Sanders, C. (1975) *The Structure and Function of Chromatin: Ciba Found. Symp.* **28**, 95–108
- Sanders, C. (1975) Ph.D. Thesis, University of London
- Walker, J. M., Goodwin, G. H. & Johns, E. W. (1976) *Eur. J. Biochem.* **62**, 461–469
- Walker, J. M., Hastings, J. R. B. & Johns, E. W. (1977) *Eur. J. Biochem.* **76**, 461–468
- Walker, J. M., Hastings, J. R. B. & Johns, E. W. (1978a) *Nature (London)* **271**, 281–282
- Walker, J. M., Parker, B. P. & Johns, E. W. (1978b) *Int. J. Pept. Protein Res.* **12**, 269–276
- Watson, P. C., Peters, E. H. & Dixon, G. H. (1977) *Eur. J. Biochem.* **74**, 53–60