Volume 211, number 1, 59–63 FEB 04360 January 1987

The primary structure of histone H3 from the nematode *Caenorhabditis elegans*

Jacques R. Vanfleteren, Stefaan M. Van Bun* and Jozef J. Van Beeumen*

Laboratorium voor Morfologie en Systematiek der Dieren and **Laboratorium voor Microbiologie en Microbiële Genetica, Rijksuniversiteit Gent, Ledeganckstraat 35, B-9000 Gent, Belgium*

Received 17 November 1986

The complete amino acid sequence of histone H3 (135 residues) from the nematode *Caenorhabditis elegans* has been established. Microheterogeneity occurs at positions 96 and 100 of the chain. The sequences of the nematode H3 isoforms are very similar to the major chain of calf thymus H3 with which they show 4 substitutions in total. The major variant has cysteine in position 96. This is the first report of cysteine in this position in H3 from non-mammalian tissue. An exceptional methylation site has been detected at position 79. Various other sites of secondary modification are of a conservative nature.

Amino acid sequence; Histone H3; (*Caenorhabditis elegans*)

1. INTRODUCTION

The existence of histone isoforms has been well documented [1–3]. The variants are coexpressed in the same species in a cell-cycle-dependent manner or a stage-specific way. The sea urchin histones H1, H2A and H2B, for example, are expressed in a stringently programmed fashion during embryogenesis [2,4]. Here we report the complete primary structure of histone H3 extracted from nematode tissue. Microheterogeneity appears to exist at two positions. The experimental results should facilitate further study of the expression of these isoforms during development and ageing of *Caenorhabditis elegans*.

Correspondence address: J.R. Vanfleteren, Laboratorium voor Morfologie en Systematiek der Dieren, Rijksuniversiteit Gent, Ledeganckstraat 35, B-9000 Gent, Belgium

Abbreviations: AUC, acetic acid/urea/cetyltrimethylammonium bromide; AUT, acetic acid/urea/Triton X-100; DPTU, diphenylthiourea; HPLC, high-performance liquid chromatography; PTH, phenylthiohydantoin; TEMED, *N,N,N′,N′*-tetramethylethylenediamine

2. MATERIALS AND METHODS

2.1. Caenorhabditis elegans *H3*

A dauer defective strain DR27 daf-17(m27) (provided by the *Caenorhabditis* Genetics Center, which is supported by contract no. N01-AG-9-2113 between the National Institutes of Health of the US and the Curators of the University of Missouri) of *C. elegans* was used in this study. The procedures for nematode growth and histone preparation have been outlined previously [5]. Size-exclusion chromatography in acetate-bisulfide buffer (pH 5.6) as recommended by Van der Westhuyzen and Von Holt [6] yielded a fraction that was enriched in the arginine-rich histones but our attempts to purify H3 by means of conventional chromatographic procedures failed. Histone H3 was then finally purified by preparative SDS-PAGE (15% gel). Protein bands were located by short staining with Coomassie R-250. They were excised and eluted with 70% HCOOH for 24 h at room temperature [7]. SDS and Coomassie blue

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

were removed from the dried (Savant Speed Vac Concentrator) sample with a mixture containing equal volumes of ethanol and methanol to which H_2SO_4 was added to a final concentration of 0.01 N. The protein pellet was then washed twice with ethanol/methanol (1:1) to remove the acid and dried.

2.2. *Separation of peptides*

Approx. 200 nmol intact protein was carboxymethylated following the procedure of Alfageme et al. [8]. After termination of the reaction the protein was precipitated with 20% trichloroacetic acid. The precipitate was subsequently washed with 20% trichloroacetic acid, then with acetone containing 0.05 N HCl, twice with acetone only and dried under reduced pressure. The pellet was dissolved in 50 mM ammonium acetate (pH 4) and digested with *Staphylococcus aureus* V8 proteinase (EC 3.4.19; Miles, England) overnight at 37°C using an enzyme to substrate ratio of 1:60. A second aliquot of fresh enzyme was added and digestion was continued for another 20 h. The resulting peptides were separated on a Vydac C-4 reversed-phase column (4 × 250 mm) in 0.1% trifluoroacetic acid using a linear (0.65% per min) acetonitrile gradient.

Partial acid hydrolysis was performed on 50 nmol intact H3. The protein was dissolved in 0.5 ml of 2% HCOOH and heated at 106°C for 4 h [9]. The peptides were separated on a μBondapak C-18 reversed-phase column (4 × 150 mm) in 0.2% trifluoroacetic acid using a linear (0.65% per min) gradient from 0 to 60% acetonitrile, 20% isopropanol.

The tryptic core was generated by *p*-tosyl-L-phenylalanine-chloromethyl ketone-treated trypsin (EC 3.4.21.4) digestion (2 h, 37°C) of approx. 50 nmol protein dissolved in 50 mM ammonium bicarbonate buffer (pH 8.2). The precipitating material was pelleted, extensively washed with buffer and dried.

2.3. *Gel electrophoresis*

Histone isoforms were resolved by two-dimensional electrophoresis. The first dimension gels (0.8 mm thick) were composite gels in which the histones were first separated in a 7 cm long, 12% polyacrylamide (0.08% bisacrylamide) gel portion containing 5% acetic acid, 0.5% Triton X-100 and 7.5 M urea before entering the 12 cm long second portion of the gel, where the concentration of urea was lowered to 6 M. Photopolymerization was accomplished to avoid the need for extensive pre-electrophoresis. It was based on the TEMED (0.5%)-riboflavin (0.0002%) redox system and initiated by illumination of the cast gel for 1 h with an ultraviolet lamp emitting light at 366 nm.

The second gel dimension was either a 15% polyacrylamide/1 M acetic acid/6 M urea gel with 0.15% cetyltrimethylammonium bromide in the upper buffer [10] or a 15% SDS-PAGE gel according to Laemmli [11] run in 0.05 M Tris, 0.384 M glycine, 0.1% SDS.

The gels were stained with Coomassie G-250 in 2% phosphoric acid, 6% ammonium sulfate according to Neuhoff et al. [12].

To avoid problems such as proteolysis, oxidation and loss of modification that are usually associated with long purification procedures we preferred to work with a partly purified histone sample. Nuclei were prepared following a simplified quick procedure [13] and extracted with 2 M NaCl containing 20 μg/ml aprotinin and 1% thiodiglycol. Non-histones and DNA were precipitated with 0.35 N HCl and removed by centrifugation. All these operations were performed at 2°C. The histones were then collected by precipitation with 20% trichloroacetic acid, washed twice with 20% trichloroacetic acid containing 1% thiodiglycol, 1% 2-mercaptoethanol and dried.

2.4. *Automated sequence analysis*

Automated Edman degradation was carried out by using an Applied Biosystems 470 gas-phase sequenator. The PTH-amino acids were analyzed by reversed-phase HPLC on a 5 μm IBM cyanopropyl column (4.6 × 250 mm) essentially following the procedure of Hunkapiller and Hood [14] but with 3.75% tetrahydrofuran in the acetate buffer in order to differentiate diphenylthiourea from PTH-Met.

3. RESULTS AND DISCUSSION

The complete amino acid sequence of histone H3 is given in fig.1 along with the fragmentation and alignment of the peptides that have led to this proposal. P is a polypeptide obtained during the

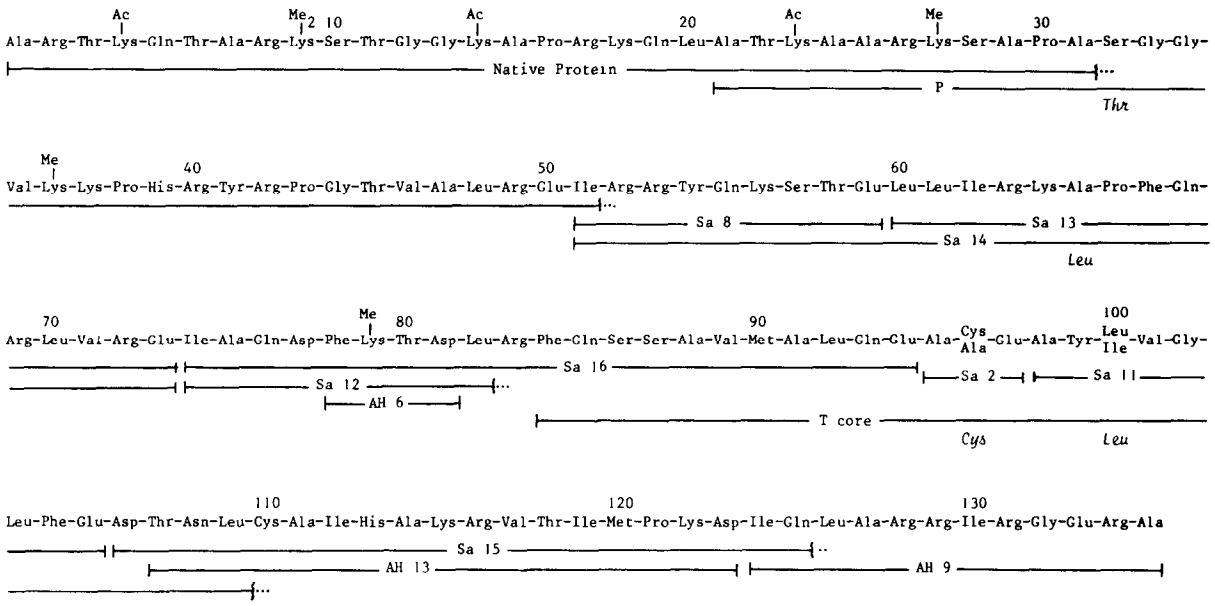

Fig.1. The primary structure of histone H3 from *C. elegans*. Automated sequence analysis was carried out on the native protein, an endogenous proteolytic fragment P and various peptides generated from H3: Sa, *S. aureus* V8 cleavage; AH, mild acid hydrolysis; T core, the insoluble peptide fraction produced by trypsin. The numbering refers to the relative position of elution of the peptides during reversed-phase HPLC. The sequencing of some peptides was deliberately discontinued before reaching the C-terminus (⟨···). Types of modification: Ac, acetylation; Me, 1 methyl group; Me_2, 2 methyl groups. Only sequence positions that differ from calf thymus H3 [17] are indicated in script below the amino acid chain.

preparative electrophoresis of the arginine-rich histones in SDS gels where it ran about halfway between intact histone H3 and H4. Sequence analysis of this material showed that it was derived from intact H3 by cleavage at the C-terminal side of Leu_{20}.

The polypeptide chain is 135 residues long. Microheterogeneity was observed at positions 96 and 100. Isoleucine appeared to be the minor residue at position 100, yet still comprised between 20 and 40% of the total. Both alanine and carboxymethylcysteine were observed at position 96 but it was not possible to quantitate their relative amounts due to the rapid decay of PTH-carboxymethylcysteine.

Electrophoresis of whole histone in the presence of Triton X-100 resolved 3 putative H3 polypeptides (fig.2). Those labeled H3.1 and H3.3 comigrate with calf thymus H3 in SDS-PAGE. We assume that H3.1 contains cysteine at position 96 and that H3.3 is a minor variant which has alanine in this position. Evidence exists that cysteine indeed binds Triton X-100 quite avidly, whereas alanine does so much less. For instance, the two isoforms H3.1 and H3.2 of the mouse differ only in a Cys/Ser substitution at position 96 and their separation in Triton gels is similar to that shown here [10]. Also, the mobility of *C. elegans* histone H4 is also considerably lower when compared to calf thymus H4 as a result of a single Thr/Cys substitution at position 73 [15]. On the other hand, H1 which generally comprises very high amounts of alanine (usually up to 25%) does not bind Triton at all and is found on the diagonal.

The identity of the polypeptide labeled H3.2 is less clear. It comigrates with H3.1 in the AUC gel and displays the same pattern (trailing spots) of secondary modification. However, SDS-PAGE reveals that this polypeptide is about 700–900 Da smaller (fig.2). We assume that this material is derived from H3.1 by proteolytic cleavage of 6–8 amino acid residues despite our efforts to avoid proteolysis during histone preparation. We have observed (not shown) that this fraction is virtually

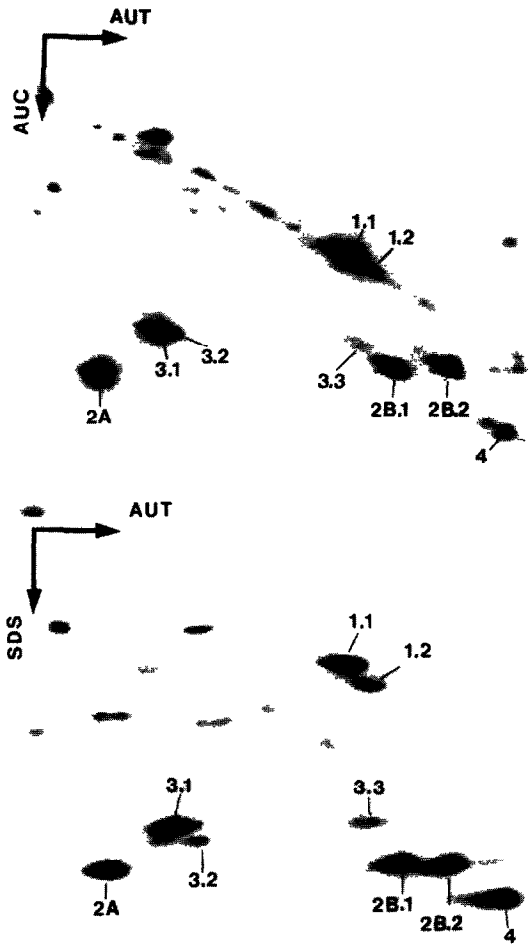

Fig.2. Two-dimensional electrophoresis of the histones of *C. elegans*. The first dimension was acetic acid/urea/Triton X-100 (AUT) and the second either acetic acid/urea/cetyltrimethylammonium bromide (AUC) or SDS-PAGE. Putative isoforms are indicated by a number which follows the main fraction, separated by a period.

absent in histones prepared from nematode eggs (which contain all stages of embryonic development) following exactly the same preparation procedure. Polypeptide H3.2 might therefore result from proteolytic processing in vivo. A similar physiologically regulated process has definitely been proven to occur in *Tetrahymena* [16].

The Leu_{100}/Ile heterogeneity is probably not resolved on any of these gels.

PTH analysis of the residues at positions 4, 14 and 23 showed that both lysine and ϵ-acetyllysine were present. This is in good agreement with the modification pattern seen in AUT/AUC gels where at least 2 modified forms of H3.1 and H3.2 are observed in addition to the unmodified molecules. The PTH derivatives of ϵ-acetyllysine coelute with PTH-Ala on our chromatograms which made it impossible to measure the extent of acetylation at the respective sites more accurately. Small amounts of PTH-ϵ-acetyllysine would not be detected for the same reason.

A substantial part of lysine residues 9 carry two methyl groups. Here also it was not possible to estimate the extent of dimethylation more accurately since PTH-ϵ-dimethyllysine coelutes with DPTU in our chromatographic system. Monomethylated lysine residues were found at positions 27 (approx. 10%), 36 (approx. 25%) and 79 (approx. 50%).

Lys_4 is generally not acetylated in mammals and the modification of Lys_{79} reported here is unique, as far as we know. The other sites of modification are of a conservative nature. All isoforms taken together it can be concluded that nematode H3 differs from the calf histone in no more than 4 positions [17]. All these replacements are conservative in nature.

ACKNOWLEDGEMENTS

We wish to thank Dr P. Sautère (Lille, France) for providing the PTH derivatives of the most common modified amino acids. This work was supported by grant no.3.0071.84 from the Fund for Medical Scientific Research awarded to J.V. and J.V.B. The first author acknowledges a senior Research Associate Career with the National Fund for Scientific Research.

REFERENCES

[1] Franklin, S.G. and Zweidler, A. (1977) Nature 266, 273–275.
[2] Brandt, W.F., Strickland, W.N., Strickland, M., Carlisle, L., Woods, D. and Von Holt, C. (1979) Eur. J. Biochem. 94, 1–10.
[3] Wu, R.S., Panusz, H.T., Hatch, C.L. and Bonner, W.M. (1986) CRC Crit. Rev. Biochem. 20, 201–263.
[4] Cohen, L.H., Newrock, K.M. and Zweidler, A. (1975) Science 190, 994–997.

[5] Vanfleteren, J.R., Van Bun, S.M., Delcambe, L.L. and Van Beeumen, J.J. (1986) Biochem. J. 235, 769–773.
[6] Van der Westhuyzen, D. and Von Holt, C. (1971) FEBS Lett. 14, 333–337.
[7] Tsugita, A., Sasada, S., Van den Broek, R. and Scheffler, J.J. (1982) Eur. J. Biochem. 124, 171–176.
[8] Alfageme, C.R., Zweidler, A., Mahowald, A. and Cohen, L.H. (1974) J. Biol. Chem. 249, 3729–3736.
[9] Inglis, A.S. (1983) Methods Enzymol. 91, 324–332.
[10] Bonner, W.M., West, M.H.P. and Stedman, J.D. (1980) Eur. J. Biochem. 109, 17–23.
[11] Laemmli, U.K. (1970) Nature 227, 680–685.
[12] Neuhoff, V., Stamm, R. and Eibl, H. (1985) Electrophoresis 6, 427–448.
[13] Meheus, L. and Vanfleteren, J.R. (1986) Mech. Age. Dev. 34, 23–34.
[14] Hunkapiller, M.W. and Hood, L.E. (1983) Methods Enzymol. 91, 254–259.
[15] Vanfleteren, J.R., Van Bun, S.M. and Van Beeumen, J. (1986) Comp. Biochem. Physiol., in press.
[16] Allis, C.D., Bowen, J.K., Abraham, G.N., Glover, C.V.C. and Gorovsky, M.A. (1980) Cell 20, 55–64.
[17] DeLange, R.J., Hooper, J.A. and Smith, E.L. (1972) Proc. Natl. Acad. Sci. USA 69, 882–884.