THE COMPLETE AMINO ACID SEQUENCE OF HISTONE F_3 FROM CHICKEN ERYTHROCYTES

Wolf F. BRANDT and Claus VON HOLT
Department of Biochemistry, University of Cape Town, Rondebosch, Cape Town, South Africa

Received 10 April 1972

1. Introduction

We wish to report the complete sequences of the 136 amino acids of the only cysteine containing histone isolated from chicken erythrocytes. For the first time, to our knowledge, all the positions of the amino acids, in the course of establishing the primary structure of a protein, have been assigned without the use of overlapping sequences. To achieve this, specific chemical cleavages rather than enzymatic degradations were chosen and applied, first to the original protein chain, and subsequently to the generated polypeptides to yield sets of not more than 3 peptides in any single cleavage. Their relative position in the protein or polypeptides became evident after comparison of the N- and C-terminal amino acid in the cleavage products and the uncleaved starting material.

The simplicity of the peptide mixture after each cleavage, resulting in easy separation of the peptides, together with the highly efficient Edman degradations of automatic sequencing, allowed us to perform a rapid and relatively non-laborious primary structure determination of histone F_3.

2. Experimental

We isolated F_3 histone and purified it through a series of gel filtration steps as reported previously [1], making use of the fact that it is the only cysteine containing histone and can thus undergo controlled dimerization. Gel electrophoresis, dansylation, cyanogen bromide (CNBr) cleavage on the performic acid oxidized [2] protein and amino acid analysis were carried out as reported previously [1] but using thioglycollic acid instead of phenol, to prevent oxidative losses during protein hydrolysis.

Tyrosine containing peptides were cleaved with *N*-bromosuccinimide (NBS) [3] at room temp. by adding sequently approx. 1/10 of the theoretical amount, and monitoring the spectrum until no further change occurred before the next addition of NBS. Titration was considered to be complete when addition of NBS caused no further increase in the 260 nm reading. Aspartic acid containing peptides were cleaved in dilute acid [4].

Instead of the Quadrol [5] and dimethylallylamine [6] buffers generally used, completely aldehyde-free dimethylaminopropyne buffer [7] was used in the Edman degradation.

The protein and peptide programs, as provided by the manufacturer of the Beckman Model 980 sequencer, were modified, and together with the dimethylaminopropyne buffer, resulted in higher yields, cleaner chromatographic backgrounds and lower losses due to peptide extraction.

PTH and TMS–PTH amino acids were identified by gas chromatography (GC) [8] and amino acid analysis. Ethyl acetate soluble and insoluble derivatives from each step were pooled and hydrolysed to the free amino acids in 6 N HCl–1% thioglycollic acid for 24 hr at 130°. Amino acid analysis was performed on a Model 116 modified to allow the detection of 3 nmoles amino acid. Full account of experimental details will be published elsewhere.

3. Results and discussion

The electrophoretic microheterogeneity of F_3 histone, apparent on gel electrophoresis, using the method of Panyim and Chalkley [9] (not evident using the method of Bonner et al. [10]), is due to acetylation of lysine residues since, on incubation with calf thymus histone deacetylase [1], the two slower bands disappear. In addition, acetic acid was identified by GC [1]. The amino acid composition of the protein was found to be: Lys_{12}, $Lys\text{-}\epsilon\text{-}(Me)_1$, His_2, Arg_{18}, Asp_5, Thr_{10}, Ser_6, Glu_{16}, Pro_6, Gly_7, Ala_{18}, Cys_1, Val_6, Met_2, Ile_7, Leu_{12}, Tyr_3, Phe_4, $(NH_3)_{10}$ [1].

The modified degradation procedure allowed us to carry the sequence determination on the uncleaved protein through 48 steps compared with 26 steps by the standard method. The subsequent approach to the complete sequence determination is given in fig. 1a. By inspecting the amino acid composition of the protein or peptides, a cleavage method was chosen which yielded not more than 3 fragments. Their relative position became obvious from the knowledge of the C-terminal amino acid (the point of cleavage) and the establishment of the N-terminal residue. After CNBr cleavage [1] peptides CN-3 and CN-2 were sequenced 15 and 29 steps, respectively. Assignments of the ultimate amino acids were aided by the knowledge of the amino acid composition of the peptide and the C-terminal residue. Peptide CN-1 contained two tyrosine residues and was thus cleaved with NBS. Peptides CN-1, NB-2 and CN-1,NB-1 (1) were sequenced 12 and 36 steps, respectively, yielding the complete sequence of the protein.

Additional cleavages at aspartyl residues in peptide CN-1,NB-1 (1) and at tyrosyl residues in peptide CN-2 were performed to yield the peptides expected from the sequence of the parent polypeptides.

The distribution of basic amino acids (fig. 1b) is similar to that of histone F_{2al} and F_{2b} [11]. The N-terminal and C-terminal regions are very basic. They are separated by a stretch of 32 non-basic residues which contains the cysteine residue. More than half of all basic amino acids occur in pairs (fig. 1b). The N-terminal half of the protein contains 4 out of the

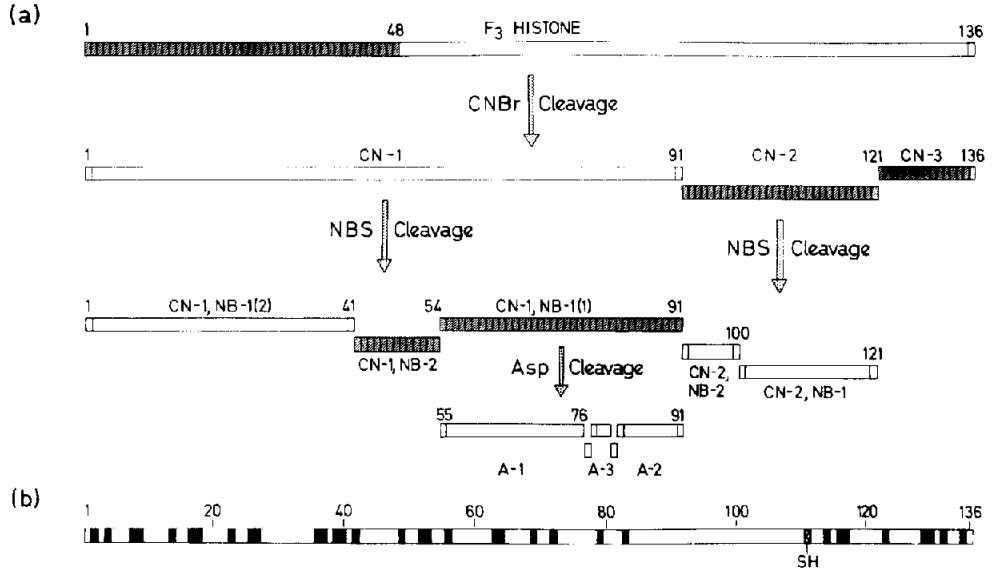

Fig. 1. a) Sequence determination of F_3 histone. □: amino acids positioned by various manual methods. ▬: uninterrupted automatic sequence determination. Nomenclature of polypeptides: CN denotes a cyanogen bromide cleavage fragment and the number its order of elution with 0.01 N HCl from Sephadex G-100 column [1]. NB denotes a *N*-bromosuccinimide cleavage fragment. The first number after NB refers to the order of elution from a Sephadex G-50 column with 0.01 N HCl. The number in brackets denotes the order of elution from a CMC-column using a NaCl gradient from 0 to 1.0 M in 0.05 M acetate buffer pH 7.5. A: denotes peptides derived from cleavage of fragment CN-1,NB-1 (1) at aspartyl residues and the number the elution order with 0.01 N HCl from Sephadex G-25. b) Distribution of basic amino acid in histone F_3.

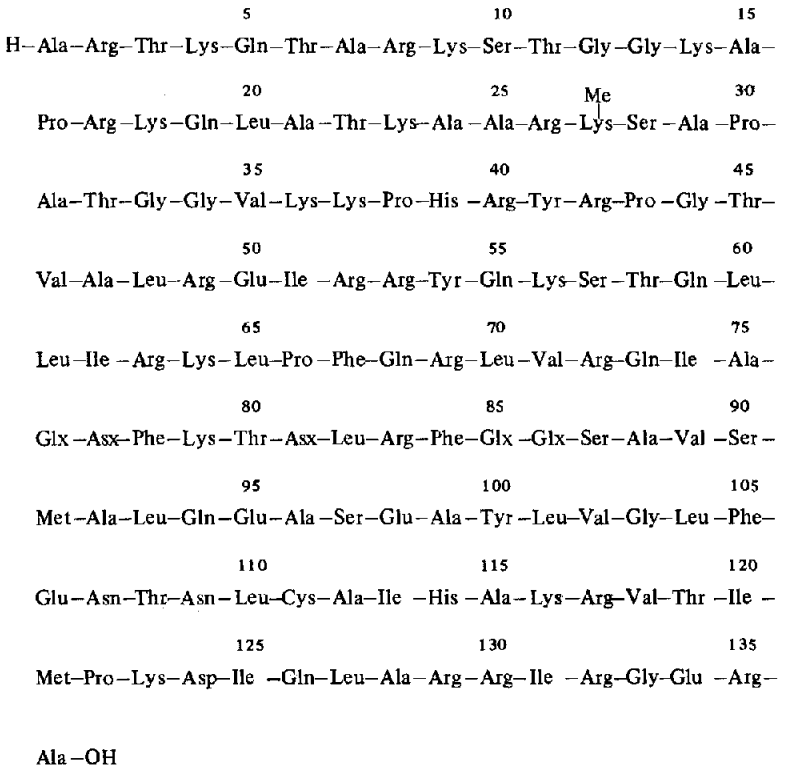

Fig. 2. Complete amino acid sequence of histone F_3.

5 proline residues and twice as many hydroxy amino acids as the C-terminal half. All the methyl and acetyl groups are located in the N-terminal region. The significance of these distributions as far as protein–protein and protein–DNA interactions are concerned remains to be elucidated.

Lysine 27 is completely methylated. Lysines in the first 37 residues are partially acetylated since peptides CN-1 and CN-1,NB-1(2) shows the same electrophoretic microheterogeneity as histone F_3.

All the tryptic peptides of calf thymus histone F_3 sequenced by Delange et al. [12] can be aligned in the N-terminal region. In chicken erythrocytes and calf thymus the same lysine residue (residue 27) is methylated and probably the same lysines are partially acetylated (residues 14 and 23 [12]).

As expected, only few differences occur between calf and chicken histone F_3. Out of the 60 residues that can be compared at this stage [12, 13] only one inversion of residue 28 and 29 is evident.

Sequence similarities between histone F_{2al} (IV) and F_3 exist. Comparing various stretches to the ancient histone IV A dodecapeptide [14] show that residues 1–6, 7–17, 28–40, 41–49 and 128–136 correspond to 83%, 75%, 88%, 81% and 85% unaltered ancient histone RNA bases, compared with 43% for random sequences.

Acknowledgements

We wish to thank Margaret Morgan for skillful technical assistance. This work was supported by grants from the C.S.I.R. (R.S.A.) and the University of Cape Town Research Committee. One of the authors (W.F.B.) held a postgraduate C.S.I.R. bursary.

References

[1] W.F. Brandt and C. von Holt, FEBS Letters 14 (1971) 338.
[2] C.H.W. Hirs, J. Biol. Chem. 219 (1952) 611.
[3] L.K. Ramachandran and B. Witkop, in: Methods in Enzymology, Vol. 11, ed. C.H.W. Hirs (Academic Press, New York, 1967) p. 283.
[4] J. Schultz, in: Methods in Enzymology, Vol. 11, ed. C.H.W. Hirs (Academic Press, New York, 1967) p. 255.
[5] P. Edman and G. Begg, European J. Biochem. 1 (1967) 80.
[6] H.D. Niall, H. Penhasi, P. Gilbert, R.C. Meyers, F.G. Williams and J.T. Potts, Federation Proc. 28 (1969) 661.
[7] G. Braunitzer and B. Schrank, Hoppe-Seyler's Z. Phys. Chem. 351 (1970) 417.
[8] J.J. Pisano and T.J. Bronzert, J. Biol. Chem. 244 (1969) 5597.
[9] S. Panyim and R. Chalkley, Arch. Biochim. Biophys. 130 (1969) 337.
[10] J. Bonner, R. Chalkley, M. Dahmus, D. Fambrough, F. Funimura, P.C. Huang, S. Huberman, R. Jensen, K. Marushige, H. Ohlenbusch and B.M. Olivera, in: Methods in Enzymology, Vol. 12b, eds. L. Grossman and K. Moldave (Academic Press, New York, 1968) p. 3.
[11] D.M.P. Phillips, in: Histones and Nucleohistones, ed. D.M.P. Phillips (Plenum Press, London and New York, 1971) p. 55.
[12] R.J. Delange, E.L. Smith and J. Bonner, Biochem. Biophys. Res. Commun. 4 (1970) 989.
[13] M.O.J. Olson, J. Jordan and H. Busch, Biochem. Biophys. Res. Commun. 46 (1972) 50.
[14] K. Bauer, Int. J. Protein Research 3 (1971) 165.