CHICKEN ERYTHROCYTE HISTONE H5

IV. Sequence of the carboxy-termined half of the molecule (96 residues) and complete sequence

G. BRIAND, D. KMIECIK, P. SAUTIERE, D. WOUTERS, O. BORIE-LOY, G. BISERTE, A. MAZEN+ and M. CHAMPAGNE+

Unité 124 INSERM et Institut de Recherches sur le Cancer, Place de Verdun, BP 3567, 59020 Lille Cédex and *+Institut de Biologie Moléculaire et Cellulaire du CNRS, 15, rue Descartes, 67000 Strasbourg, France*

Received 17 January 1980

1. Introduction

The sequence of the first 111 residues and some partial sequences of the carboxy-terminal region of chicken erythrocyte histone H5 (189 residues, mol. wt 20 580) have been determined [1,2]. Most of the 66 basic residues present in the histone appeared to be accumulated in the carboxy-terminal half of the protein molecule; the elucidation of such a highly basic sequence was problematic. Specific cleavage with pepsin at the single phenylalanine residue of the protein, located at position 93 of the sequence, released the carboxy-terminal fragment of H5. The sequence of this fragment (96 residues) was determined mainly from data provided by peptides obtained by cleavage with trypsin and thermolysin. Additional information was obtained from peptides generated by cleavage with elastase. The complete sequence of the 189 amino acid residues of the chicken erythrocyte histone H5 is elucidated here.

2. Materials and methods

All materials and methods were as in [1,3], with the following additions: Histone H5 dissolved in 5% formic acid, was hydrolyzed with pepsin (Worthington) at 37°C for 2 h, using an enzyme-to-substrate ratio of 1:50. The hydrolysate was fractionated on a Sephadex G-50 (fine) column (200 × 2.5 cm) equilibrated and eluted with 0.01 M HCl. Digestion of the carboxy-terminal peptic fragment Px with thermolysin (Merck) was performed at 40°C in 0.1 M ammonium bicarbonate (pH 8.0) for 30 min or 2 h using an enzyme-to-substrate ratio of 1:100. The fragment Px was also hydrolyzed with porcine elastase (Calbiochem) at 37°C in 0.1 M ammonium bicarbonate (pH 8.0) for 30 min using an enzyme-to-substrate ratio of 1:500 or for 2 h using an enzyme-to-substrate ratio of 1:100.

Maleylation, citraconylation or succinylation of peptide Px prior to tryptic digestion were performed at 0°C in 0.1 M K_2HPO_4 (pH 9.0) with a 10 M excess of maleic, citraconic or succinic anhydride relatively to the content of α- and ϵ-amino groups. Hydrolysis of the modified peptide Px with Tos-PheCH_2Cl-treated trypsin (Worthington) was performed at 37°C in 0.1 M ammonium bicarbonate (pH 8.0) for 4 h, using an enzyme-to-substrate ratio of 1:50.

The enzymatic digests were fractionated on a Sephadex G-25 (fine) column (200 × 2.5 cm) equilibrated and eluted with 0.01 M HCl.

The fractions containing large peptides, were further fractionated on a CM-cellulose (Whatman CM-52) column (25 × 1.2 cm) equilibrated in 0.02 M ammonium acetate–0.2 M NaCl adjusted to pH 6.0 with HCl. Peptides were eluted with a linear gradient of NaCl (from 0.2–0.6 M) in 0.02 M ammonium acetate buffer (pH 6.0). The small peptides ranging from di- to hexapeptides were separated by ion-exchange chromatography on Chromobeads P (Technicon) with pyridine–formate and pyridine–acetate buffers [4].

Automated Edman degradation of large peptides was performed in a Beckman 890 C Sequencer using a DMAA program (102974) in the presence of poly-

Abbreviations: Tos-PheCH_2Cl, *N*-tosylphenylalanine chloromethyl ketone; DMAA, dimethylallylamine; PTH, phenylthiohydantoin

brene [5,6]. In some cases peptides were treated with 3-sulfophenylisothiocyanate prior to Edman degradation [7]. PTH amino acids identified by high-pressure liquid chromatography on a column of C_{18} micro Bondapak (Waters Associates) [8] and by gas-chromatography as in [9]. The sequence of the small peptides was determined by the manual dansyl-Edman method [10] or by a modified Edman degradation developed in [11] using 4-*N*,*N*-dimethylaminoazobenzene-4′-isothiocyanate (DABITC, Fluka AG) in combination with phenylisothiocyanate as coupling reagent.

3. Results and discussion

When histone H5 is hydrolyzed with pepsin, the specific cleavage of the Phe$_{93}$—Arg$_{94}$ bond yields a peptide of 96 residues, designated by Px, with the following amino acid composition:

Asp_1, Thr_3, Ser_{12}, Pro_9, Gly_2, Ala_{17}, Val_3, Leu_1, Lys_{34}, Arg_{14}.

From kinetic studies with carboxypeptidases B and C performed simultaneously on peptide Px and on histone H5, peptide Px was found to be the carboxy-terminal half of the protein. The amino acid sequence of peptide Px is presented in fig.1. Detailed results which led to the above sequence will be published elsewhere. Thus the carboxy-terminal sequence of chicken histone H5 which remained to be determined [1,2], is now elucidated and with it, the complete amino acid sequence of the protein (fig.2).

Chicken erythrocyte histone H5 contains 189 residues. By amino acid analysis and structural studies of peptide Px a precise evaluation of the number of serine, lysine and arginine present in the protein could be obtained. Thus histone H5 contains 26 serine, 44 lysine and 22 arginine, instead of 25, 42 and 21, respectively, as described in [2]. Moreover the sequences Gly$_{90}$—Ser$_{91}$ and Ala$_{110}$—Lys$_{111}$ in [1], were found

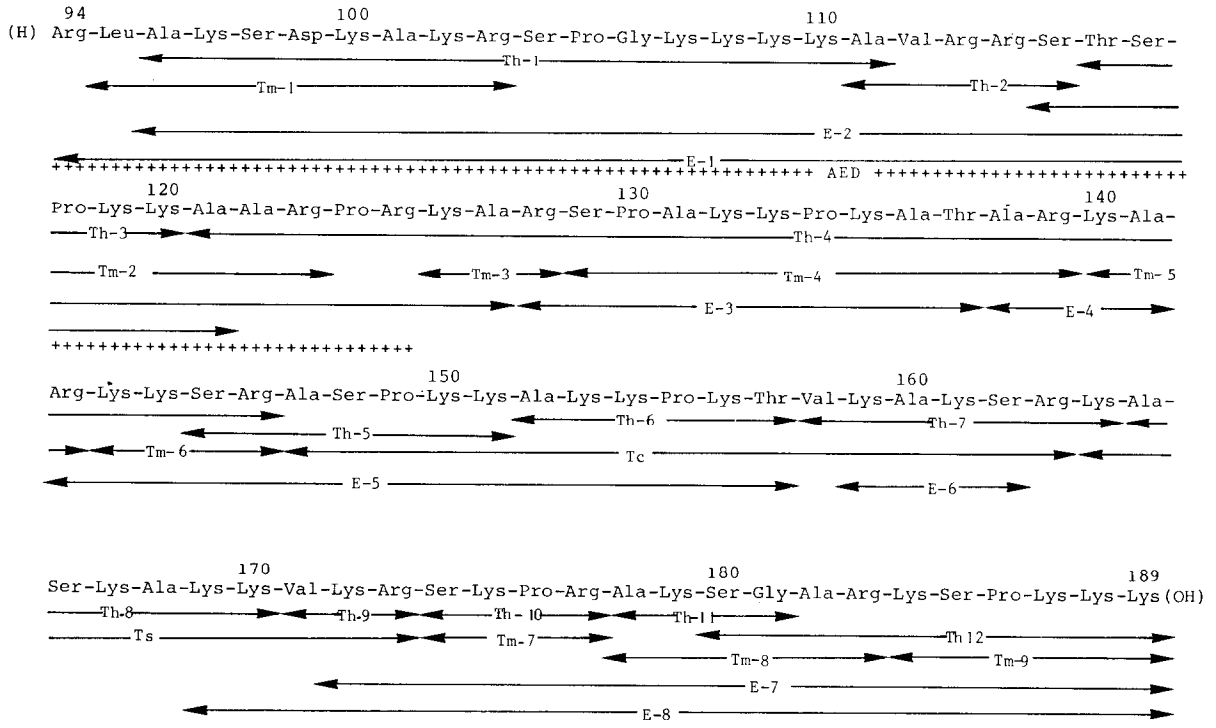

Fig.1. Carboxy-terminal sequence of chicken erythrocyte histone H5. Th, peptide from digestion with thermolysin; Tm, peptide from digestion with trypsin after maleylation; Tc, peptide from digestion with trypsin after citraconylation; Ts, peptide from digestion with trypsin after succinylation; E, peptide from digestion with elastase; AED, automated Edman degradation.

```
     1                                   10                                  20
(H) Thr-Glu-Ser-Leu-Val-Leu-Ser-Pro-Ala-Pro-Ala-Lys-Pro-Lys-Gln/Arg-Val-Lys-Ala-Ser-Arg-Arg-Ser-

                                  30                                  40
    Ala-Ser-His-Pro-Thr-Tyr-Ser-Glu-Met-Ile-Ala-Ala-Ala-Ile-Arg-Ala-Glu-Lys-Ser-Arg-Gly-Gly-

                          50                                  60
    Ser-Ser-Arg-Gln-Ser-Ile-Gln-Lys-Tyr-Ile-Lys-Ser-His-Tyr-Lys-Val-Gly-His-Asn-Ala-Asp-Leu-

                 70                                  80
    Gln-Ile-Lys-Leu-Ser-Ile-Arg-Arg-Leu-Leu-Ala-Ala-Gly-Val-Leu-Lys-Gln-Thr-Lys-Gly-Val-Gly-

         90                                  100                                 110
    Ala-Ser-Gly-Ser-Phe-Arg-Leu-Ala-Lys-Ser-Asp-Lys-Ala-Lys-Arg-Ser-Pro-Gly-Lys-Lys-Lys-Lys-

                                 120                                 130
    Ala-Val-Arg-Arg-Ser-Thr-Ser-Pro-Lys-Lys-Ala-Ala-Arg-Pro-Arg-Lys-Ala-Arg-Ser-Pro-Ala-Lys-

                          140                                 150
    Lys-Pro-Lys-Ala-Thr-Ala-Arg-Lys-Ala-Arg-Lys-Lys-Ser-Arg-Ala-Ser-Pro-Lys-Lys-Ala-Lys-Lys-

                     160                                 170
    Pro-Lys-Thr-Val-Lys-Ala-Lys-Ser-Arg-Lys-Ala-Ser-Lys-Ala-Lys-Lys-Val-Lys-Arg-Ser-Lys-Pro-

                 180                                 189
    Arg-Ala-Lys-Ser-Gly-Ala-Arg-Lys-Ser-Pro-Lys-Lys-Lys(OH)
```

Fig.2. Amino acid sequence of chicken erythrocyte histone H5.

to be Ser(90)–Gly(91) from the sequence determination of the tryptic peptide (residues 86–94) and Lys(110)–Ala(111) from the data provided by automated Edman degradation of peptide Px.

Histone H5 is characterized by an asymmetrical distribution of the hydrophobic and basic residues similar to that found in trout histone H1 [12]. The amino terminal half of histone H5 contains most of the hydrophobic residues and all the aromatic residues of the protein and has been shown to take up a globular conformation at ionic strength >0.1 [2,13]. The globular region is, in fact, from residues 22–100 [14]. The carboxy-terminal half, which corresponds exactly to the fragment Px (residues 94–189) is strongly basic. Lysine and arginine account together for 50% of the total residues in the fragment Px. The remainder of the residues is almost exclusively composed of alanine, serine and proline. Such a composition is not expected to support an organized structure. As a matter of fact no secondary or tertiary structure was shown by circular dichroism and nuclear magnetic resonance studies, between residues 100–189 [2]. However the distribution of the proline residues within the carboxy-terminal region of histone H5 delimits two long basic sequences (fig.2) (the first between proline-134 and proline-149, the second between proline-155 and proline-176) which could take up a helical conformation when the positive charges are neutralized by the phosphate groups of DNA. The fact that serine-145 and serine-166 have been found phosphorylated in vitro by a rat pancreatic kinase (A. Martinage, P. S., unpublished) supports the idea that these two sequences are likely privileged sites for the binding of histone H5 to DNA, through electrostatic interactions.

On the other hand, when the sequence of chicken histone H5 is aligned for maximum homology with

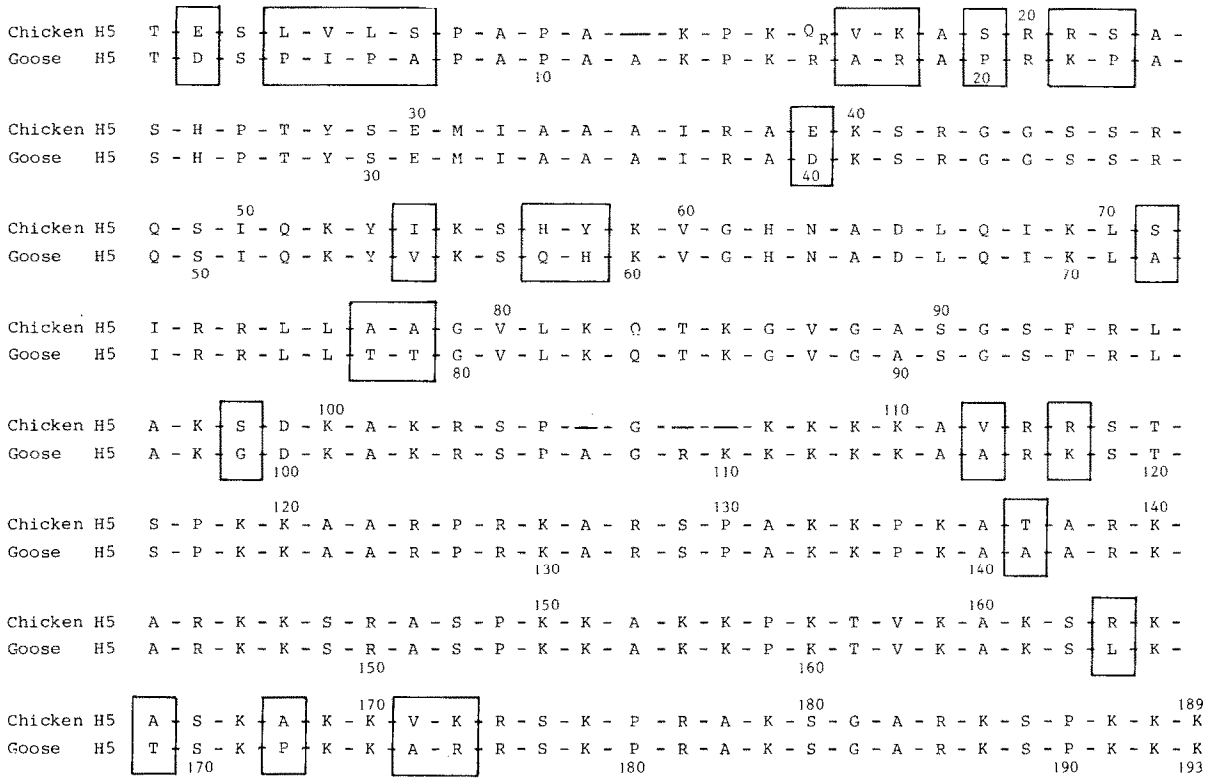

Fig.3. Comparison of amino acid sequences of chicken erythrocyte histone H5 and goose erythrocyte histone H5 [15]. The two sequences shown with the one-letter code (see J. Biol. Chem. (1979) 254, 7), are aligned for maximum homology. Amino acid changes are indicated by boxes and deletions by (——).

that of goose histone H5 [15], 30 differences can be observed (fig.3). Of these differences, 14 correspond to non-conservative amino acid replacements and 4 are deletions. The changes occur mainly in the amino-terminal and carboxy-terminal sequences of the histones. In the amino-terminal sequence (residues 1–22) 1 deletion and 10 substitutions of which 5 are not conservative are found. In the carboxy-terminal sequence (residues 100–189) 9 substitutions and 3 deletions take place. The globular regions (residues 23–99) show an extensive homology; besides the few conservative changes observed, the replacement of the sequence His_{62}–Asn_{63} of chicken H5 by the sequence Gln_{63}–His_{64} in goose H5, must be quoted.

The conservation of the sequence of the globular region in avian histone H5 might be related to a fundamental function, e.g., protein–protein interaction, in the chromatin.

Acknowledgements

The skillful technical assistance of Mrs D. Bühr and Miss J. Dunand (preparation of histone 115) and of Mrs D. Belaïche, M. J. Dupire and of Miss A. Hemez (structural studies of histone H5) is gratefully acknowledged. We thank Professor M. Daune (IBMC, Strasbourg) for his continuous interest and helpful advice. The work was supported by grants ATP Chromatine no. 2875 and LA 268 of the Centre National de la Recherche Scientifique and by grant no. 77.5.048.2 of the Institut National de la Santé et de la Recherche Médicale.

References

[1] Sautière, P., Briand, G., Kmiecik, D., Loy, O., Biserte, G., Garel, A. and Champagne, M. (1976) FEBS Lett. 63, 164–166.

[2] Crane-Robinson, C., Briand, G., Sautière, P. and Champagne, M. (1977) Biochim. Biophys. Acta 493, 283–292.
[3] Garel, A., Mazen, A., Champagne, M., Sautière, P., Kmiecik, D., Loy, O. and Biserte, G. (1975) FEBS Lett. 50, 195–199.
[4] Wouters, D., Sautière, P. and Biserte, G. (1978) Eur. J. Biochem. 90, 231–239.
[5] Tarr, G. E., Beecher, J. F., Bell, M. and McKeaw, D. (1978) Anal. Biochem. 84, 622–627.
[6] Hunkapiller, M. W. and Hood, L. E. (1978) Biochemistry 17, 2124–2133.
[7] Dwulet, F. E. and Gurd, F. R. N. (1976) Anal. Biochem. 76, 530–538.
[8] Hermann, J., Titani, K., Ericsson, L. H., Wade, R. D., Neurath, H. and Walsh, K. A. (1978) Biochemistry 17, 5672–5679.
[9] Pisano, J. J. and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5597–5607.
[10] Gray, W. R. and Hartley, B. S. (1963) Biochem. J. 89, 379–380.
[11] Chang, J. Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.
[12] McLeod, A. R., Wong, N. C. W. and Dixon, G. H. (1977) Eur. J. Biochem. 78, 281–291.
[13] Crane-Robinson, C., Danby, S. E., Bradbury, E. M., Garel, A., Kovacs, A. M., Champagne, M. and Daune, M. (1976) Eur. J. Biochem. 67, 379–388.
[14] Chapman, G. E., Aviles, F. J., Crane-Robinson, C. and Bradbury, E. M. (1978) Eur. J. Biochem. 90, 287–296.
[15] Yaguchi, M., Roy, C. and Seligy, V. L. (1979) Biochem. Biophys. Res. Commun. 90, 1400–1406.