AMINO ACID SEQUENCE HOMOLOGIES BETWEEN H1 AND H5 HISTONES

MAKOTO YAGUCHI, CAMILLE ROY, MICHAEL DOVE and VERNER SELIGY

Division of Biological Sciences, National Research Council, Ottawa, Ontario, Canada K1A 0R6

Received March 25,1977

SUMMARY: The amino acid sequence of the first 112 residues of goose H5 and the first 38 residues of pigeon H5 has been determined. When a gap of 12 residues is introduced in the N-terminal regions of avian H5, almost half of the residues become identical to those of H1 from various species. The region of this gap corresponds to a homologous repeating sequence in H1. This suggests a deleation or duplication in this region and that the genes for H1 and H5 histones have evolved from a common ancestral gene.

A comparison of the N-terminal amino acid sequence of H5 isolated from chicken, quail, duck, goose and pigeon showed considerable sequence variation (1). The species variability and microheterogeneity (2) of avian histone H5 are similar but not as extensive as those found in H1 histones from rabbit thymus (3, 4, 5). However, both H1 and H5 histones are quite different from H2A, H2B, H3 and H4 histones which exhibit an unusually high degree of sequence conservation (6, 7). Sautiere and coworkers determined the amino acid sequence of the first 111 residues of chicken H5 (8, 9, 10) and reported that about 15 residues adjacent to the single phenylalanine of H5 were analogous to a region corresponding to Phe-106 in rabbit thymus H1 (9, 11). Strickland et al. (12) found that about 25 residues including Phe-93 of chicken H5 were very similar to the corresponding region of H1 from rabbit (11), sea urchin (12) and trout (13). We have extended the amino acid sequence of goose H5 and pigeon H5, and have found that the sequence homologies between H1 and avian H5 are much more extensive than those reported previously (6, 7, 9, 12, 14).

ISSN 0006-291X

MATERIAL AND METHODS

The source for geese (Anser anser) and pigeons (Columbia livia) and the methods of isolation of nuclei, extraction of H5, fractionation and purification of H5 have been described (1, 15). The purity of each protein was assessed by electrophoresis at pH 2.7 and in sodium dodecylsulphate (1, 15).

The first 36 residues of goose H5 and the first 38 residues of pigeon H5 were automatically degraded with a Beckman model 890C sequenator with a 0.5 M quadrol (program No. 122974). The thiszolinone or PTH derivatives were hydrolyzed with 6N HCl in the presence or absence of 0.1% $SnCl_2$ (16) at 130°C for 20 hours, and the amino acids formed were analyzed with the Durrum D-500 amino acid analyzer. The identification of some PTH derivatives (Asp, Asn, Glu, Gln) was made by thin-layer chromatography on silica gel plates (17, 18).

Large fragments of goose H5 were produced by cyanogen bromide, N-bromosuccinimide, acetic acid (10), and Staphyloccocus aureus protease (19) which was performed at pH 4.0 for 16 hours. These fragments were isolated by gel filtration on Sephadex G-75 Superfine (1.5 cm x 260 cm) eluted with 0.01 N HCl. The N-terminal region (up to 36 residues) of these fragments were degraded automatically and identified. The amino acid sequence data provided sufficient overlaps to determine unequivocally the residues from 32 to 112. Details of the sequence determination will be described elsewhere.

RESULTS AND DISCUSSION

The amino acid sequences of the first 112 residues of goose H5 and first 38 residues of pigeon H5 were compared with the first 111 residues of chicken H5 (10) and available partial sequence data of histone H1 from rabbit thymus (11), calf thymus (5), sea urchin (12), and trout (13). This is shown in Figure 1. The identical amino acids in the corresponding positions of both H5 and H1 are confined within blocks. Homologous residues exclusive to either H5 or H1 are not indicated. With the insertion of specific gaps to improve sequence alignment, 63 of the 111-112 known residues of the H5 histones could be made to coincide to identical residues in at least one of the H1 species. The homology is probably more extensive than this because at least 8 conservative substitutions (Lys-Arg, Thr-Ser, Asp-Glu, Val-Ile, and Leu-Met) are also found between H1 and H5.

The amino acids in histone H1 are asymmetrically

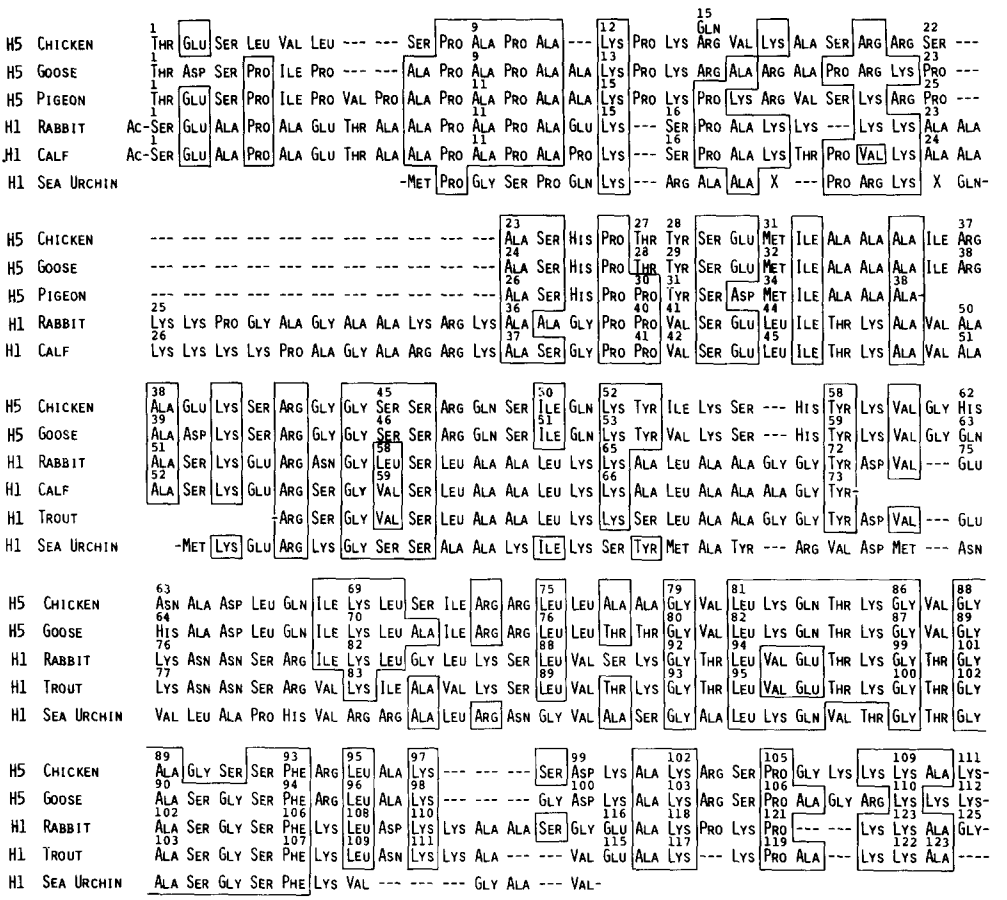

Figure 1. Comparison of partial amino acid sequences of Chicken erythrocyte H5 (ref. 10), goose erythrocyte H5, pigeon erythrocyte H5, rabbit thymus RTL-3 H1 (ref. 11), calf thymus CTL-1 H1 (ref. 5), trout testis H1 (Ref. 13), and sea urchin H1 (ref. 12).

distributed and the polypeptide chain can be divided into three regions (11, 20). Conformational studies on large fragments of calf thymus H1 (21, 22, 23) and chicken H5 (24) showed that H1 and H5 are also very similar with the possible exception of the N-terminal region. As shown in Table 1, avian H5 proteins can also be divided into three corresponding regions (10, 24). The N-terminal regions of both H5 and H1 histones have relatively high proline contents and are free of aromatic residues. On

Table 1. Distribution of amino acid residues among the three regions of H1 and H5 histones

	N-Terminal Region					Central Region			C-Terminal Half			
	H5			H1		H5		H1	H5		H1	
Position	Chicken 1-27	Goose 1-28	Pigeon 1-30	Rabbit 1-40	Calf 1-41	Chicken 28-29	Goose 29-100	Rabbit 41-115	Chicken 100-198	Goose 101-198	Rabbit 116-223	Trout 116-212
Asp	0	1	0	0	0	2	3	2	0	0	0	0
Asn	0	0	0	0	0	1	0	3	0	0	0	0
Glu	1	0	1	3	2	2	1	5	0	0	0	0
Gln	0-1	0	0	0	0	4	5	0	0	0	0	0
Thr	2	2	1	1	2	1	3	3	3	3	3	2
Ser	5	2	3	2	3	10	8	9	9	11	3	3
Pro	4	8	10	7	9	0	0	0	11	11	17	12
Gly	0	0	0	3	2	7	8	8	3	2	5	0
Ala	4	7	5	14	12	9	8	10	18	18	37	34
Val	2	0	2	0	1	3	4	6	3	1	2	4
Met	0	0	0	0	0	1	1	0	0	0	0	0
Ile	0	1	1	0	0	6	5	2	0	0	0	0
Leu	2	0	0	0	0	6	6	10	0	1	0	0
Tyr	0	0	0	0	0	3	3	1	0	0	0	0
Phe	0	0	0	0	0	1	1	1	0	0	0	0
Lys	3	3	4	9	8	8	8	13	38	38	41	42
His	1	1	1	0	0	2	2	0	0	0	0	0
Arg	2-3	3	3	1	2	6	6	2	14	13	0	0
Total	27	28	30	40	41	72	72	75	99	98	108	97

Residues are calculated from sequence and amino acid composition data given in the following references: chicken erythrocyte H5 (ref. 10), rabbit thymus RTL-3 H1 (ref. 11), calf thymus CTL-1 H1 (ref. 5), and trout testis H1 (ref. 13).

the other hand, the central regions are free of proline but contain both tyrosine and phenylalanine. Furthermore, the C-terminal halves of chicken H5 and goose H5 histones are highly basic and free of acidic and of aromatic residues as found in histone H1 proteins from rabbit (11) and trout (13).

The size of the N-terminal regions of avian H5 is significantly smaller than rabbit and calf H1 (Table 1). To align the sequence Ala-Ser-His-Pro-Pro in residues 26-30 of pigeon H5 to Ala-Ser-Gly-Pro-Pro of 36-40 of calf H1 a gap of 12 residues was made between Pro-25 and Ala-23 of pigeon H5 and the corresponding residues of other avian H5 histones. This consideration was partly based on the fact that the similarities between residues 13-23 and 24-35 of rabbit H1 (see below) could be accounted for by an internal duplication of genes for residues 13-23. Since the N-terminal region of chicken H5 has

```
                13                                          25
Rabbit H1   Ala-Glu-Lys-Ser-Pro-Ala-Lys-Lys-----Lys-Lys-Ala
                24                                          35
Rabbit H1   Ala-Lys-Lys-Pro-Gly-Ala-Gly-Ala-Ala-Lys-Arg-Lys
```

only 4 proline residues whereas that of calf H1 has 9 and H5 of goose and pigeon 8 and 10 respectively, the conformation of the N-terminal region of calf thymus H1 may be expected to be more similar to that of goose and pigeon H5 than to that of chicken H5 (24).

Fasman et al. (25) have pointed out that the first 14 residues of rabbit and calf H1 are free of basic residues. The corresponding regions from chicken, goose and pigeon H5 are also free of basic residues (Table 1). Furthermore the sequence of Ala-Pro-Ala-Pro-Ala found in residues 7-10 of goose H5 and 9-13 of pigeon H5 is identical to 9-13 of rabbit and calf H1.

The total number of residues and the amino acid compsoitions of the central regions of H5 and H1 are very similar

and the high degree of sequence homologies between chicken H5 to other H1 found by other workers (9, 12) is also observed between goose H5 and other H1. This region is the most hydrophobic and its conservation may be extremely important for function. The complete amino acid sequences of the C-terminal region of avian H5 is not yet available. The number of identical residues (5-6) homologous to both H5 and H1 is known in only the first 10 residues of this region. However, considerable sequence homology is expected judging from the similarities of amino acid compositions of the remaining C-terminal regions of H5 and H1.

The high degree of amino acid sequence homology and the other structural similarities noted here suggest that the genes for H1 and H5 histones could have evolved from a common ancestral gene and that the internal gene duplication proposed in the N-terminal and C-terminal (13) regions of H1 may be responsible for its larger size in comparison to H5. The remaining major difference between H1 and H5 is the serine, alanine and arginine contents (26) which are largely located in the C-terminal region. This difference may distinguish the functional aspects of these proteins.

(Issued as N.R.C.C. No. 15845)

REFERENCES

1. Seligy, V., Roy, C., Dove, M., and Yaguchi, M. (1976) Biochem. Biophys. Res. Commun. 71, 196-202.
2. Greenaway, P. J., and Murray, K. (1971) Nature New Biology 229, 233-238.
3. Bustin, M., and Cole, R. D. (1968) J. Biol. Chem. 243, 4500-4505.
4. Langan, T. A., Rall, S. C., and Cole, R. D. (1971) J. Biol. Chem. 246, 1942-1944.
5. Rall, S. C., and Cole, R. D. (1971) J. Biol. Chem. 246, 7175-7190.
6. Elgin, S. C., and Weintraub, H. (1975) Annu. Rev. Biochem. 44, 725-774.
7. Delange, R. J., and Smith, E. L. (1975) In: The Structure and Function of Chromatin, pp. 59-70. Ciba Foundation Symp. 28. Associ. Sci. Publisher, Amsterdam.

8. Garel, A., Mazen, A., Champagne, M., Sautiere, P., Kmiecik, D., Loy, O., Biserte, G. (1975) FEBS Lett. 50, 195-199.
9. Sautiere, P., Kmiecik, D., Loy, O., Briand, G., Biserte, G., Garel, A., and Champagne, M. (1975) FEBS Lett. 50, 200-203.
10. Sautiere, P., Briand, G., Kmiecik, D., Loy, O., Biserte, G., Garel, A., and Champagne, M. (1976) FEBS Lett. 63, 164-166.
11. Jones, G. M. T., Rall, C., and Cole, R. D. (1974) J. Biol. Chem. 249, 2548-2553, and Cole, R. D. (personal communication cited by reference 6).
12. Strickland, W. N., Schaller, H., Strickland, M., and Van Holt, C. (1976) FEBS Lett. 66, 322-327.
13. Dixon, G. H., Candido, E. P. M., Honda, B. M., Louie, A. J., MacLeod, A. R., and Sung, M. T. (1975) In: The Structure and Function of Chromatin. pp 229-258. Ciba Foundation Symp. 28. Associ. Sci. Publishers, Amsterdam, and Dixon, G. H. (personal communication cited by reference 12).
14. Temussi, P. A. (1975) J. Theor. Biol. 50, 25-33.
15. Tobin, R. S., and Seligy, V. L. (1975) J. Biol. Chem. 250, 358-364.
16. Mendez, E., and Lai, S. Y. (1975) Analyt. Biochem. 68, 47-53.
17. Wittmann-Liebold, B., Geissler, A. W., Marzinzig, E. (1975) J. Supramol. Struc. 3, 426-447.
18. Chen, R. (1976) Hoppe-Seyler's Z. Physiol. Chem. 357, 873-886.
19. Houmard, J., and Drapeau, G. R. (1972) Proc. Nat. Acad. Sci. US 69, 3506-3509.
20. Bustin, M., Rall, S. C., Stellwagen, R. H., and Cole, R. D. (1969) Science, 163, 391-393.
21. Bradbury, E. M., Cary, P. D., Chapman, G. E., Crane-Robinson, C., Danby, S. E., Rattle, H. W. E., Boublik, M., Palau, J., Aviles, F. J. (1975) Eur. J. Biochem. 52, 605-613.
22. Bradbury, E. M., Chapman, G. E., Danby, S. E., Hartman, P. G., and Riches, P. L. (1975) Eur. J. Biochem. 57, 512-528.
23. Chapman, G. E., Hartman, P. G., and Bradbury, E. M. (1976) Eur. J. Biochem. 61, 69-75.
24. 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.
25. Fasman, G. D., Chou, P. Y., and Adler, A. J. (1976) Biophysical J. 16, 1201-1238.
26. Hnilica, L. S. (1972) In: The Structure and Biological Functions of Histones. pp. 3-45. The Chemical Rubber Co. Cleveland.