CALF THYMUS HISTONE III: SEQUENCES OF THE AMINO- AND CARBOXYL-TERMINAL REGIONS AND OF THE REGIONS CONTAINING LYSYL RESIDUES MODIFIED BY ACETYLATION AND METHYLATION*

Robert J. DeLange, Emil L. Smith, and James Bonner

Department of Biological Chemistry, UCLA School of Medicine, and Molecular Biology Institute, University of California, Los Angeles California 90024
and
Division of Biology, California Institute of Technology, Pasadena, California 91109

Received July 20, 1970

SUMMARY: Sequence studies on the tryptic peptides from maleylated calf thymus histone III showed one site of ε-N-methylation in the sequence Arg-Lys $(CH_3)_{0-2}$-Ala-Ser-Pro-Ala-Thr-Gly-Gly-Val-Lys-Lys-Pro-His-Arg, and two sites of ε-N-acetylation in the sequences Arg-Lys-Ser-Thr-Gly-Gly-Lys(Ac)-Ala-Pro-Arg and Arg-Lys-Gln-Leu-Ala-Thr-Lys(Ac)-Ala-Ala-Arg. The NH_2-terminal sequence of histone III is probably Ala-Arg, and the sequence of 20 residues at the COOH-terminus has been established. There is little or no similarity when these regions of calf thymus histones III and IV are compared.

It has been demonstrated that certain amino acid side-chains in histones are modified by acetylation (1,2), methylation (3-7), and phosphorylation (8,9). Thus, ε-N-acetyllysine (1,2), ε-N-monomethyllysine (3,4), ε-N-dimethyllysine (4), ε-N-trimethyllysine (5), 3-methylhistidine (6), ω-N-monomethylarginine (7), α-N-methyl, guanidinomethylated arginine (7), and phosphoserine (8,9) have been reported as components of pure or mixed histones. These modifications are thought to be important to the proposed function of histones in genetic regulation (10-12). The positions of methylated and acetylated lysyl residues in the sequences of calf thymus and pea seedling histones IV were established in this laboratory by DeLange et al. (1,13,14). We now report the sequences of peptides containing these derivatives in calf thymus histone III, as well as the COOH-terminal and NH_2-terminal sequences of this histone.

Calf thymus histone III was prepared as previously described (15). Tryptic hydrolysis of the maleylated histone (14 μmoles) and fractionation of the resulting peptides on columns of Sephadex G-25 and G-50 and by paper electro-

* This work was aided by Grants GM 11061 and GM 13762 from the National Institute of General Medical Sciences, U. S. Public Health Service.

phoresis and chromatography were performed as previously described for histone IV (13,14). Treatment of the non-oxidized, unmaleylated histone (1 μmole) with cyanogen bromide yielded 2 fractions on Sephadex G-50, by the method previously described (16). Amino acid analysis, the Edman degradation, and enzymic hydrolysates were performed as previously described (13,14). Analysis for methylated lysines was made by the method of DeLange, Glazer, and Smith (17).

RESULTS AND DISCUSSION

The sequences of 8 of the 20 tryptic peptides from histone III are given in Table I. The NH_2-terminal residue of histone III was determined as alanine in agreement with the results of Fambrough and Bonner (15) and of Johns et al. (18). Since Peptide T(m)-IVc was the only tryptic peptide isolated with NH_2-terminal alanine, it must be the NH_2-terminal peptide, unless this peptide was not isolated. Since the other peptides do not have NH_2-terminal alanine, they

TABLE I

AMINO ACID SEQUENCES OF TRYPTIC PEPTIDES FROM MALEYLATED CALF THYMUS HISTONE III

Peptide	Sequence
T(m)-IVc	Ala-Arg
T(m)-Ia	Lys$(CH_3)_{0-2}$[a]-Ala-Ser-Pro-Ala-Thr-Gly-Gly-Val-Lys-Lys-Pro-His-Arg
T(m)-IIa	Lys-Ser-Thr-Gly-Gly-Lys(Ac)[b]-Ala-Pro-Arg
T(m)-IIb1	Lys-Gln-Leu-Ala-Thr-Lys-Ala-Ala-Arg
T(m)-IIb2	Lys-Gln-Leu-Ala-Thr-Lys(Ac)[b]-Ala-Ala-Arg
T(m)-IVa	Ile-Arg
T(m)-IVb	Gly-Glu-Arg-Ala
T(m)-IIIa	Ile-Arg-Gly-Glu-Arg

[a] $Lys(CH_3)_{0-2}$ may be lysine, ε-N-monomethyllysine, or ε-N-dimethyllysine (see text).

[b] Lys(Ac) = ε-N-acetyllysine

must have originated from other places in the molecule, and each must be preceded by an arginyl residue, because of the limited action of trypsin on maleylated proteins (13,14,19). All peptides except the COOH-terminal peptide (T(m)-IVb) were proved to have COOH-terminal arginine (by hydrolysis with carboxypeptidase B) in accord with this specificity.

Of the 14 lysyl residues present in the 20 tryptic peptides, only one was present as the ε-N-methylated derivative. This residue is the NH_2-terminal residue of Peptide T(m)-Ia, which also contains 2 unmodified lysyl residues. The sequence of this peptide was established with the aid of trypsin hydrolysis (cleaved the Lys-Lys bond) and papain hydrolysis of the larger tryptic peptide (cleaved the Gly-Gly bond). Analysis of the larger tryptic peptide for methylated lysine by the method previously described (17) showed the presence of both ε-N-monomethyllysine and ε-N-dimethyllysine in the ratio of 0.84 to 1.00. The stoichiometry indicated that as many as one-fourth of the peptide molecules contained unmodified NH_2-terminal lysine. The only similarity in the sites of methylation in calf thymus histones III and IV (13,14) that is evident from this study is the presence of an arginyl residue NH_2-terminal to the methylated lysyl residue. However, since the sequence Arg-Lys occurs at least 3 other times in histone III and once more in histone IV (13,14) with no methylation of these lysyl residues, some additional structural feature must be responsible for the specificity of the sites of methylation.

Of the 14 lysyl residues present in the 20 tryptic peptides, only 2 were found to be ε-N-acetylated. These 2 lysyl residues were present in Peptides T(m)-IIa and T(m)-IIb2, both possessing NH_2-terminal lysine (not acetylated) and a residue of ε-N-acetyllysine 3 residues removed from the COOH-terminal arginyl residue (Table I). This same relationship of the ε-N-acetyllysyl and arginyl residues was present in calf thymus histone IV (13,14) and in one of the 2 sites of acetylation in pea seedling histone IV (14) (Lys(Ac)-Arg-His Arg), but the other site of acetylation in the pea histone IV (14) does not have this relationship to an arginyl residue. It is of interest that Sung and Dixon (20)

have reported the acetylation of lysyl Residues 5, 8, 12, and 16 in trout testis histone IV. Thus, multiple sites of acetylation have been demonstrated for histone IV from 2 species (pea and trout) (14,20) and for histone III from calf thymus (this work). Since only certain lysyl residues are acetylated, the exact requirements for the specificity of acetylation remain to be elucidated. Peptide T(m)-IIb was isolated in both the acetylated (IIb2) and the unacetylated (IIb1) forms (Table I) in yields (18% and 24%, respectively) which indicated that of every 7 molecules of histone III, 3 were acetylated at Residue 6 of Peptide IIb. The unacetylated form of Peptide IIa was never isolated, but the low yield (10%) suggested that all of the histone III molecules were not acetylated at Residue 6 of this peptide.

Since Peptide T(m)-IVb was the only tryptic peptide lacking COOH-terminal arginine, it must be the COOH-terminal peptide. This peptide had the COOH-terminal sequence Arg-Ala (Table I) in agreement with the COOH-terminal sequence of histone III reported by Phillips and Simson (21). Peptides T(m)-IVa, T(m)-IVb, and T(m)-IIIa are obviously all from the COOH-terminal region of the molecule (see Fig. 1).

Since Peptide BrCN-3 lacks homoserine (or its lactone), it must be the COOH-terminal cyanogen bromide peptide. Its composition (see Fig. 1) is the same as the combined compositions of Peptides T(m)-IVa and T(m)-IVb plus arginine (which was isolated from the tryptic hydrolysate) and the COOH-terminal region of Peptide T(m)-IIc. Since Peptide T(m)-IIc must be preceded by an arginyl residue, Peptide BrCN-3 establishes the sequence of 20 residues at the

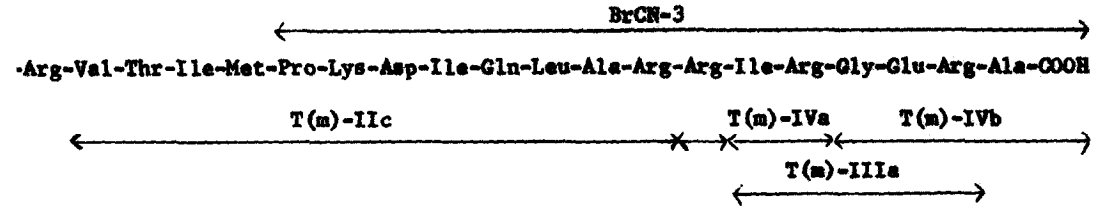

Fig. 1. The COOH-terminal 20 residues of calf thymus Histone III.

COOH terminus. The only striking similarity in the COOH-terminal sequences of calf thymus histones III and IV is the presence of the sequence Val-Thr-Ile-Met in histone III and Val-Thr-Ala-Met in histone IV (13,14).

The details of these studies and the studies on the other peptides will be presented in a later communication.

REFERENCES

1. DeLange, R. J., Smith E. L., Fambrough, D. M., and Bonner, J., Proc. Nat. Acad. Sci. U.S.A. 61, 1145 (1968).
2. Gershey, E. L., Vidali, G., and Allfrey, V. G., J. Biol. Chem. 243, 5018 (1968).
3. Murray, K., Biochemistry 3, 10 (1964).
4. Kim, S., and Paik, W. K., J. Biol. Chem. 240, 4629 (1965).
5. Hempel, K., Lange, H. W., and Birkofer, L., Naturwissenschaften 55, 37 (1968).
6. Gershey, E. L., Haslett, G. W., Vidali, G., and Allfrey, V. G., J. Biol. Chem. 244, 4871 (1969).
7. Paik, W. K., and Kim, S., Arch. Biochem. Biophys. 134, 632 (1969).
8. Ord, M. G., and Stocken, L. A., Biochem. J. 98, 5P (1966).
9. Kleinsmith, L. J., Allfrey, V. G., and Mirsky, A. E., Proc. Nat. Acad. Sci. U.S.A., 55, 1182 (1966).
10. Hnilica, L. S., Progr. Nucl. Acid Res. Mol. Biol. 7, 25 (1967).
11. Stellwagen, R. H., and Cole, R. D., Ann. Rev. Biochem. 38, 951 (1969).
12. Elgin, S. C. R., Froehner, S. C., Smart, J. E., and Bonner, J., Adv. Cell Mol. Biol. 1, in press (1970).
13. DeLange R. J., Fambrough, D. M., Smith, E. L., and Bonner, J., J. Biol. Chem. 244, 319 (1969).
14. DeLange, R. J., Fambrough, D. M. Smith, E. L., and Bonner, J., J. Biol. Chem. 244, 5669 (1969).
15. Fambrough, D. M., and Bonner, J., J. Biol. Chem. 243, 4434 (1968).
16. DeLange, R. J., Fambrough, D. M., Smith, E. L., and Bonner, J., J. Biol. Chem. 243, 5906 (1968).
17. DeLange, R. J., Glaser, A. N., and Smith, E. L., J. Biol. Chem. 245, 3325 (1970).
18. Johns, E. W., Phillips, D. M. P., Simson, P., and Butler, J. A. V., Biochem. J. 77, 631 (1960).
19. Butler, P. G., Harris, J. I., Hartley, B. S., and Leberman, R., Biochem. J. 103, 78P (1967).
20. Sung, M., and Dixon, G. H., Proc. Nat. Acad. Sci. U.S.A., in press (1970).
21. Philips, D. M. P., and Simson, P., Biochim, Biophys. Acta, 181, 154 (1969).