3. V. D. Solov'ev, Vestn, Akad. Med. Nauk SSSR, No. 7, 27 (1976).
4. V. D. Solov'ev, N. B. Varshaver, and N. E. Gulevich, Vopr. Virusol., No. 5, 580 (1963).
5. V. D. Solov'ev and N. E. Gulevich, in: Symposia of the 9th International Congress of Microbiology [in Russian], Moscow (1966), p. 407.
6. V. D. Solov'ev and Ya. E. Khesin, Vestn, Akad. Med. Nauk SSSR, No. 10, 20 (1970).
7. Ya. E. Khesin, A. M. Amchenkova, N. V. Karazhas, et al., Tsitologiya, No. 11, 1382 (1972).
8. B. Carrit and P. Goldfarb, Nature, 264, 556 (1976).
9. P. Couillin, A. Bouè, N. Van Cong, et al., Cytogenetics, 16, 111 (1976).
10. B. Dutrillaux and J. Lejeune, C. R. Acad. Sci. (Paris), 272, 2638 (1971).
11. "Human gene mapping. 3," Cytogenetics, 16, 1 (1976).
12. Ya. E. Khesin, A. M. Amchenkova, and G. P. Sovjetova, J. Gen. Virol., 24, 17 (1974).
13. T. Kusano, R. Wang, R. Pollack, et al., J. Virol., 5, 682 (1970).
14. D. A. Miller, O. J. Miller, V. G. Dev. et al., Cell, 1, 167 (1974).
15. M. Seabright, Lancet, 2, 971 (1971).

GEOMETRIC COMPLEMENTATION OF THE PRIMARY MOLECULAR STRUCTURES OF HISTONES H2A, H2B, H3, AND H4 AND SOME POSSIBLE CONSQUENCES OF THIS PHENOMENON

D. M. Spitkovskii

UDC 577.112.823

Conformity between the geometries of arrangement of extended sequences of basic and nonpositively-charged amino acids along the polypeptide chains of histone H2A, H2B, H3, and H4 molecules, forming the protein skeleton of the nucleosomes, discovered for the first time, are described in this paper. The number of histone packing variants in the tetramer satisfying the conditions of complementation discovered is 10^4, and in a histone octamer 10^8 respectively. It is postulated that the structural heterogeneity of the nucleosomes may have functional significance and that the choice of packing variant of their protein skeleton can take place by mechanisms of allosteric regulation or of the primary structure of that part of the DNA molecule which is a component of the nucleosome.
KEY WORDS: *nucleosome; histones; code; regulation; complementation.*

In recent years substantial progress has been made in the elucidation of the structure of the elementary deoxyribonucleoprotein (DNP) fibril which lies at the basis of organization of the eukaryote chromosome. This fibril has been shown to be a chain of repeating DNP subunits (nucleosomes) with a definite number of DNA molecules between them. Yet information at present available provides no grounds for the determination of the specific functions of the nucleosomes. They are ascribed mainly a structural role (DNA packing). This suggestion is based on at least two theses: a) the constancy of composition of the proteins in nucleosomes and b) the structure of nucleosomes is independent of the primary structure of their DNA components.

However, these theses do not take into account a further possibility, to be described below, which is a consequence of relationships discovered by the writer between the geometries of arrangement of extended sequences of nonpositively-charged amino-acid residues along the polypeptide chains of four types of histones participating in the formation of the protein skeleton of nucleosomes.

Laboratory of Molecular Biology, Institute of Medical Genetics, Academy of Medical Sciences of the USSR, Moscow. (Presented by Academician of the Academy of Medical Sciences of the USSR S. S. Debov.) Translated from Byulleten' Éksperimental'noi Biologii i Meditsiny, Vol. 86, No. 11, pp. 579-583, November, 1978. Original article submitted April 12, 1978.

 0007-4888/78/8611-1496$07.50

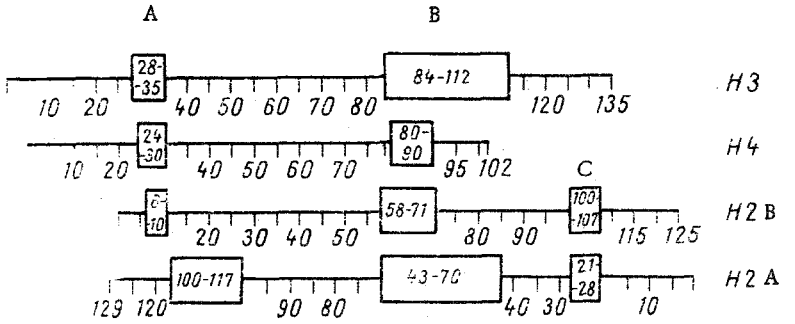

Fig. 1. Complementation between blocks (A, B, and C) with no basic amino-acid residues in their composition in primary structures of histones H3, H4, H2B, and H2A. Numbers denote serial numbers of amino-acid residues in primary structures of corresponding histones. Counting starts at NH end of molecule, which for all histones except H2A is located on the left of the figure. Scale: one division corresponds to 5 amino-acid residues. Letters A, B, and C denote corresponding matching blocks of histones H2A, H2B, H3, and H4. Serial numbers of amino-acid residues corresponding to beginning and end of block are inscribed inside it.

The primary structure of the histones concerned is now known and the distribution of the basic and hydrophobic amino acids in these proteins has been shown to be on the block principle [2-6, 8]. According to the authors just cited, the following regions can be distinguished in histone molecules. In histone H2A: an NH_2-terminal region from the 1st to the 42nd, a middle region from the 43rd to the 70th, and a COOH-terminal region from the 71st to the 129th amino-acid residue; in histone H2B: an NH_2-region from the 1st to the 34th, a middle region from the 35th to the 79th, and a COOH-terminal region from the 80th to the 125th amino-acid residue; in histone H3: an NH_2-terminal region from the 1st to the 53rd and a COOH-terminal region from the 54th to the 135th amino-acid residue; finally, in histone H4: an NH_2-terminal region from the 1st to the 45th and a COOH-terminal region from the 46th to the 102nd amino-acid residue. Within the limits of each of the above-mentioned regions of the molecule, blocks with the longest sequences without basic amino-acid residues are distinguished. The principles governing the distribution of these blocks between histones H2A, H2B, H3 and H4 are given below. As Fig. 1 shows, given a matching arrangement of the histones, the extended sequences not containing basic amino acids in their composition and constituting regions of hydrophobic binding, match one another sufficiently closely (blocks A, B, and C). Under these circumstances the regions between the blocks, common to all four histones examined, are similar in size. For instance, between blocks A and B in the molecules of histones H3, H4, H2B, and H2A there are 48, 49, and 29 amino-acid residues respectively. Some blocks in the molecules of the histones studied are also similar in size. For instance, in the A blocks in histones H3 and H4 and in the C blocks in histones H2B and H2A there are 8, 7, 8, and 8 amino-acid residues respectively. The B blocks in histones H3 and H2A contain 29 and 28 amino-acid residues respectively. The parameters of conformity indicated above, together with some others, are clearly shown in Fig. 1.

The agreement observed between the distribution of the blocks is of considerable interest for these blocks are found in proteins forming a single system (a tetramer). It is therefore unlikely that such a distribution is random, and has no meaning attached to it. To investigate this problem let us attempt to analyze what other geometric consequences may arise as a result of the above-mentioned distribution of blocks and the numerical values of the amino-acid residues contained both in the blocks and between them. For this purpose, let us examine first the possible geometric congruence between molecules of histones H2A and H2B only. Let us arrange these proteins so that the NH_2-ends of the two molecules are on the left side (Fig. 2a, b, c). It follows from Fig. 2a that the number of amino-acid residues (14) between blocks C and B in histone H2A corresponds exactly to the number of amino-

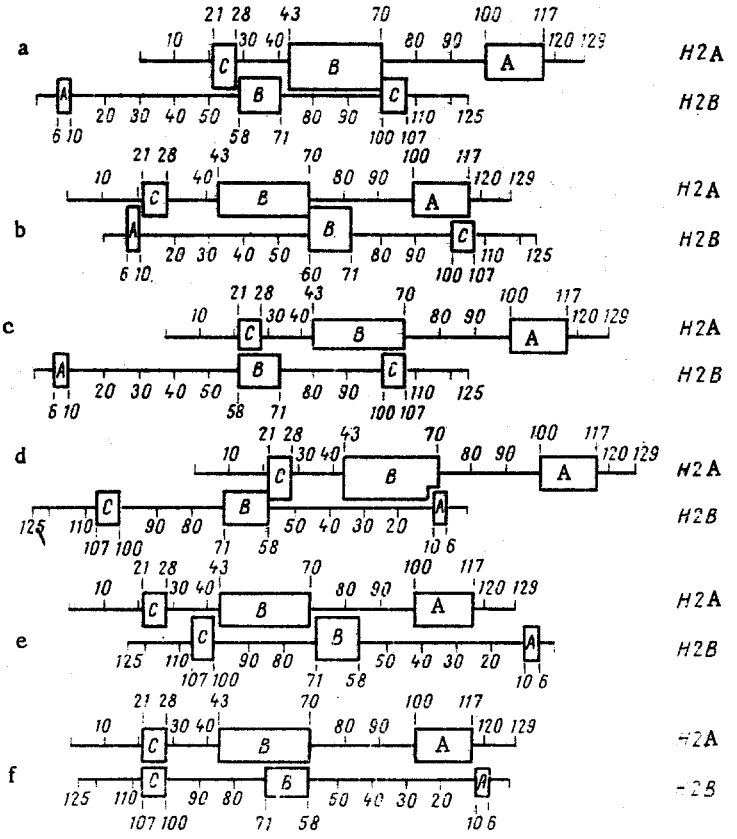

Fig. 2. Variants of geometric complementation between block A, B, and C in primary structures of histones H2A and H2B. Variants a, b, and c: NH_2 ends of histones H2A and H2B on left of figure. Variants d, e, and f: NH_2 ends of histones H2A and H2B located on left and right respectively.

amino-acid residues (14) in block B of histone H2B, and the number of amino-acid residues (28) between blocks B and C in histone H2B agrees remarkably with the size of block B (28 amino acids) in histone H2A, so that these proteins are complementary. Other variants of congruence between these two histones are given in Fig. 2b and c.

If the molecule of histone H2B is turned through 180°C, so that the NH_2 end of the histone H2A molecule and the COOH end of the histone H2B molecule are on the left side of Fig. 2, three further positions of geometric congruence between the blocks will be noted (Fig. 2d, e, f). Altogether six positions with geometric congruence between the molecules of histones H2A and H2B can thus be identified. In the same way it is easy to show that the number of positions of geometric congruence between particular amino-acid sequences is as follows: 2 between histones H3 and H4, 6 between H4 and H2A, 4 between H4 and H2B, 6 between H3 and H2A, and 6 between H3 and H2B. It is also easy to show that histone H1 plays no part in this geometric mosaic.

The total number of combinations of geometric congruence between the four histones analyzed is thus of the order of 10^4. Since we do not know whether the two tetramers in the octamer forming the protein skeleton of the nucleosomes in the above plan are identical, the total number of histone packing variants in the nucleosome may be potentially 10^8.

It should be noted that if not only the blocks examined above, but all nonpositively-charged blocks in the histone molecules are taken into account (i.e., 2 additional blocks for histone H2B, 2 for H3, and 2 for H4), the general picture of geometric complementation becomes still more clearly defined, although more complicated. In this case a definite num-

ber of new methods of fitting the blocks together appears, but the total number of variants of geometric congruences between the blocks is not significantly changed.

It can be concluded from the facts described above that on the basis of the same four types of histones analyzed it is potentially possible to obtain a considerable number of tetramers, identical in chemical composition but differing from each other in their methods of organization. It is of course realized that in the actual situation ability of histones to fit together depends largely on their secondary structure, the fine details of which are not yet known. However, there is no doubt that the secondary structure is a distinctive projection of the primary structure and transmits its particular features from a different angle.

On the basis of the material described above a fundamentally new idea can be put forward on the structural, but not chemical, heterogeneity of the nucleosomes, based on structural differences of one of their components, namely histone tetramers. Naturally the different tetramers may have different functional loads. For instance, differences in the combination of sequences of amino-acid residues outside the hydrophobic centers of the nucleosomes, arising as a result of different methods of organization of the tetramers, may be responsible for essential differences in their specific properties, by forming regions which recognize regulatory agents. Unfortunately we do not know how many of the total number of variants of nucleosomes mentioned above — 10^4 or 10^8 — are in fact realized and what determines the choice of organization of the nucleosomes. In principle the initiators of choice could be: chemical modification of histones and (or) a change in the ionic or dielectric parameters of the medium. Both the above types of initiators can formally undertake the choice of geometric complementation by mechanisms of allosteric regulation, i.e., through appropriate changes in the conformation of histone protomers during their binding with ligands. Besides others, certain discrete sites of acetylation or methylation in the above molecules can perform the function of allosteric centers. In the plan we are examining the concept of allosteric regulation can be applied not only to the enzymic, but also the structural function of the molecules. Finally, it should be pointed out that another candidate for the role of initiator of choice of variant of the organization of histone tetramers in the nucleosome may be the secondary structure and sequence of DNA bases within the nucleosome. This last hypothesis is particularly interesting because the numbers of structural combinations of histones in the tetramer or octamer given above become comparable with the number of combinations of triplets in the DNA region that constitutes an element of the nucleosome.

It must be emphasized that some workers [1, 7] have already pointed out that more than 30% of the amino-acid residues of histone H4 have identical arrangements with histones H2A and H2B or similar arrangements along the polypeptide chains; in their opinion this is evidence of the evolution of one original histone. Meanwhile the congruence established by the writer between the blocks, some variants of which are shown in Fig. 2, can be taken to indicate that evolutionary selection of the histone sequences took place in the direction of their possible participation in the formation of an elementary chromatin formation, identical in composition but polymorphic in structure, namely the nucleosome.

It must therefore be emphasized once again that the complementation discovered in the investigation described above between corresponding elements of the primary molecular structures of histones H2A, H2B, H3, and H4 can substantially modify our views on the role of these proteins. Combinations between molecules of the four types of histones may perhaps be comparable in importance with combinations between the nucleotides of DNA, i.e., they may play the role of a unique regulating code.

LITERATURE CITED

1. K. Bauer, Biochim. Biophys. Acta, 400, 80 (1975).
2. R. J. De Lange et al., Proc. Nat. Acad. Sci. USA, 69, 882 (1972).
3. R. J. De Lange et al., Proc. Nat. Acad. Sci. USA, 61, 1145 (1968).
4. K. Iwai et al., Nature, 226, 1056 (1970).
5. Y. Ogawa et al., J. Biol. Chem., 244, 4387 (1969).
6. P. Sautière et al., Cr. Acad. Sci. (Paris), D-274, 1422 (1972).
7. P. A. Temussi, J. Theoret. Biol., 50, 25 (1975).
8. L. C. Yeoman et al., J. Biol. Chem., 247, 6018 (1972).