New Molecular Systems. Regularities of Supramolecular Biopolymer Structure Formation

I. P. Shabalkin and P. I. Shabalkin

Translated from *Kletochnye Tehnologii v Biologii i Medicine*, No. 4, pp. 214-218, December, 2009
Original article submitted March 3, 2009

A new model of supramolecular structure of DNA is proposed, according to which the processes caused by the monomer rearrangements in the carcass of the molecule underlie DNA compactization. Rearrangement in the DNA chain can result in the appearance of dimers of three types with different bending angles. The dimer geometry corresponds to geometrical figures (correct tetrahedron and octahedron) conforming to the volume-surface economy rule. These figures serve as the unit of a correct crystal lattice. The formation of a DNA supramolecular unit (nucleosome) is realized in accordance with the "golden ratio" rule. The minimum fragment used in assembly of the nucleosome core consists of two mini-segments: large (containing 8 monomers) and small (5 monomers). This model suggests the conformations of functional proteins, that is, the DNA matrix determines the amino acid sequence in the protein polymer chain.

Key Words: *supramolecular structure; biopolymer; dimer; nucleosome*

Correct reading of the information on the stereostructure, recorded in the primary structure of a biopolymer, promotes the development of optimal technological solutions. New scientific and technological solutions of this or that problem in biology and medicine should conform to certain mathematical regularities, underlying the precise calculation, formòxample, location of the points through which the processes of biopolymer supramolecular structure compactization and decompactization pass, or detection of sites involved in specific ligand binding and regulating the functional activity of the biopolymer.

Based on these concepts and our previous hypothesis on the formation of primary structure of DNA molecule [14], we demonstrate the biopolymer compactization modes in a limited space under conditions of minimum energy expenditures for provision of essential information for the object as exemplified by the formation of DNA supramolecular structure. As the total length of a DNA molecule in a cell can reach 2 m, there are mechanisms for contracting its linear size in nature. One of them is bending deformation of short segments in the DNA chain. This phenomenon is based on the concept according to which DNA is a flexible chain polymer [14,15]. As a result of phosphate-sugar (P–S) monomer rearrangement in the primary structure of the chain carcass, dimers with three different bond types can form: 1) [(S–P)+(S–P)]; 2) [(S–P)+(P–S)]; and 3) [(P–S)+(S–P)].

Polymerization of DNA molecule with consideration for the existence of different dimer types promotes an increase of the molecule elasticity, because the resultant fragments of the structure can have different strength of the monomer-monomer bonds [14].

Let us consider the processes of compactization of a fragment of one DNA strand as exemplified by the formation of loop structures during the formation of the nucleosome core.

Why the Nature has chosen the loop for compactization of the DNA strand? First, appearance of lateral segments (loops), distant from the main chain, leads to shortening of the linear length of DNA chain.

N. N. Blokhin Russian Oncological Center, the Russian Academy of Medical Sciences, Moscow, Russia. ***Address for correspondence:*** shabalkin@km.ru. I. P. Shabalkin

Second, it promotes cooperation of the loops [16] due to local orientation ordering of the segments which became closer to each other in space during condensation of the main chain [10]. The segments brought closer to each other stabilize the chain carcass at the expense of the formation of extra bonds. Third, transition of a chain site from linear to annular form (loop) is more advantageous energetically and for compactization of DNA structure. This closed circular structure can fold into a more complex structure vertically (along the chain carcass) and horizontally. This process leads to the formation of a variety of 3D structures.

Hence, based on the previously developed idea on the DNA chain polymerization in accordance with the mathematical regularity known as Fibonacci numerical series [13,16], let us analyze the formation of the nucleosome core, assuming that a fragment of the primary structure of one DNA strand, including 144 monomers, corresponds to the nucleosome. This value virtually does not differ from the data of other authors [3,11], who consider that the nucleosome core contains 140-147 monomers. It is noteworthy that we do not speak about nucleotides on purpose, as in order to understand the compactization process we should study the structure of just the DNA chain carcass consisting from only two units: sugar (deoxyribose) molecules and phosphoric acid residue (phosphate group). Hence, let us assume that the first fragment of the DNA chain primary structure (144 monomers long, according to the new model of DNA structure [13,15]) looks as follows:

1 2 3 4 5 6 7 8 9 10
(P–S)+(S–P)+[(S–P)+(P–S)]+[(S–P)+(P–S)+(S–P)]+[(S–P)+(P–S)+(S–P)+

11 12 13 14 15 16 17 18 19 20
(S–P)+(P–S)]+[(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)]+

21 22 23 24 25 26 27 28 29 30
[(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+

31 32 33 34 35 36 37 38 39 40
(S–P)+(S–P)+(P–S)]+[(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+

41 42 43 44 45 46 47 48 49 50
(S–P)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+

51 52 53 54 55 56 57 58 59 60
(P–S)+(S–P)+(P–S)+(S–P)]+[(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+

61 62 63 64 65 66 67 68 69 70
(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+

71 72 73 74 75 76 77 78 79 80
(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+

81 82 83 84 85 86 87 88 89 90
(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)]+[(S–P)+(P–S)+

91 92 93 94 95 96 97 98 99 100
(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(S–P)+

101 102 103 104 105 106 107 108 109 110
(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+

111 112 113 114 115 116 117 118 119 120
(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+

121 122 123 124 125 126 127 128 129 130
(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+

131 132 133 134 135 136 137 138 139 140
(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+(P–S)+(S–P)+(S–P)+(P–S)+

141 142 143 144
(S–P)+(P–S)+(S–P)]+[(S–P)+……

In order to understand, how the nucleosome core forms from a fragment of DNA strand primary structure, we have to determine the length of the minimum fragment of the DNA strand participating in the nucleosome core assembly. For this, let us analyze the division of this fragment into its two constituents in accordance with the "golden ratio" rule.

The golden ratio philosophy is based on the fact that after division of a section into two parts, the proportion between the length of the longer fragment and the entire section is equal to the proportion between the shorter and longer fragments [8]. Numerically, the equality of two proportions (golden ratio) is expressed by an irrational number ≈0.618, if m/M=M/T, where m is the lesser part, M the larger one, and T is the whole.

The process of division of a DNA chain fragment 144 monomers long can be described by stages of this fragment division on condition that each of the two different (by length) parts of the initial fragment (let us call them segments) are fragments (shorter than the initial one!) also capable of dividing in accordance with the golden ratio rule until the equality of two proportions remains valid (Fig. 1). Hence, at stage I the fragment (total) divides in point 0 into two segments, one of them consisting of 89 monomers (longer segment) and the other has 55 monomers (shorter one). According to the golden ratio, we have two equal proportions: 55/89=89/144 or 55/89≈0.617...; 89/144≈0.618... The two proportions remain equal (0.62) after this division of the initial fragment into two parts. This equality (0.62) is retained at other stages of the initial fragment division:

34/55=55/89 or 34/55≈0.618; 55/89≈0.618;
21/34=34/55 or 21/34≈0.618; 34/55≈0.618;
13/21=21/34 or 13/21≈0.619; 21/34≈0/618;
8/13=13/21 or 8/13≈0.615; 13/21≈0.619.

The process of the initial fragment division is over at stage IV (Fig. 1) by the formation of one mini-fragment consisting of 13 monomers. It consists of two mini-segments of different length: large (8 monomers) and small (5 monomers): 5/8=8/13 or 5/8≈0.625; 8/13≈0.615.

The next stages of fragmentation of the mini-fragments consisting of 8 or 5 monomers each do not correspond to the golden ratio rule, as the equality of the two proportions is violated: 3/5≠5/8 or 0.6≠0.625; 2/3≠3/5 or 0.67≠0.6.

Hence, the results of analysis indicate that the last mini-fragment, which determines the number of monomers, components of the unit participating in the nucleosome core formation; the unit is to have 5 or 8 monomers.

The structure assembled by stages (in opposite order; Fig. 1) from these mini-segments is geometrically similar to a plus (Fig. 2). In this case the chain curvature is realized at the interface of two segments in points having dimers [(S–P)+(P–S)]. The presence of two phosphate groups close to each other promotes the elasticity of the polymeric chain segment, the curvature angle formed by the two monomers being 90°. This bending deformation of the linear chain causes the formation of crystal forms, thermodynamically) stable structures [7] providing orderly packing of separate segments of the polymer.

Along with this variant of the nucleosome core compactization, one more, we think, even more significant packing of the DNA strand in the nucleosome is possible, due to the fact that it has not only dimers [(S–P)+(P–S)], but also dimers [(P–S)+(S–P)], [(S–P)+(S–P)]. Creation of the optimal compact volume structure in a limited space should meet the requirements and optimal bending deformation of the polymeric chain, and therefore, structural elements with different forms of bending deformation of the com-

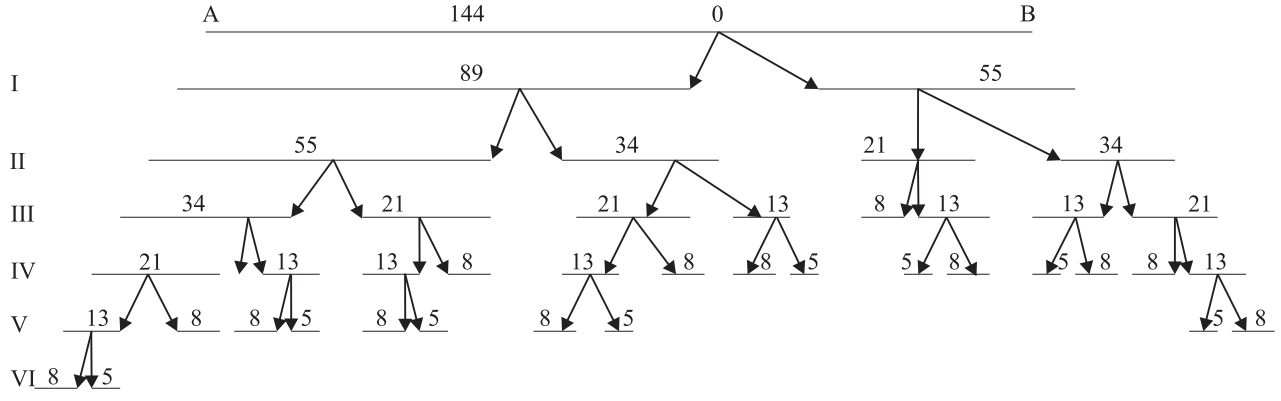

Fig. 1. Scheme of division of one DNA strand fragment (144 monomers long) into two segments according to the "golden ratio" rule. I-IV: stages of the initial fragment division to the minimum size, meeting the golden ratio requirement. Continuous lines: segments; arrow shows division of the segment from beginning to end. Arabic figures show the number of monomers.

ponents should participate in the construction of this system; in other words, a certain angle of monomer bending relatively to the other one should correspond to each of the three types of the DNA molecule dimers [15]. We think that for dimer [(P–S)+(S–P)] this angle is 60°, for dimer [(S–P)+(P–S)] 90°, and for dimer [(S–P)+(S–P)] 120°C. It is known that any compact surface tends to acquire the most thermodynamically beneficial conformation (possessing the minimum free energy [14], and hence, the first structure with the minimum volume in the DNA chain should be a correct tetrahedron constructed from 4 similar monomers, as this figure is homeomorphic to sphere [2,5]. The minimum energy is attained only on systems of this type [9]. On the other hand, a correct tetrahedron is a geometrical figure characteristic of solid crystal bodies whose basic property is the formation of a spatial crystal lattice [6]. Since the DNA molecule have to be packed into the cell core as compactly as possible,

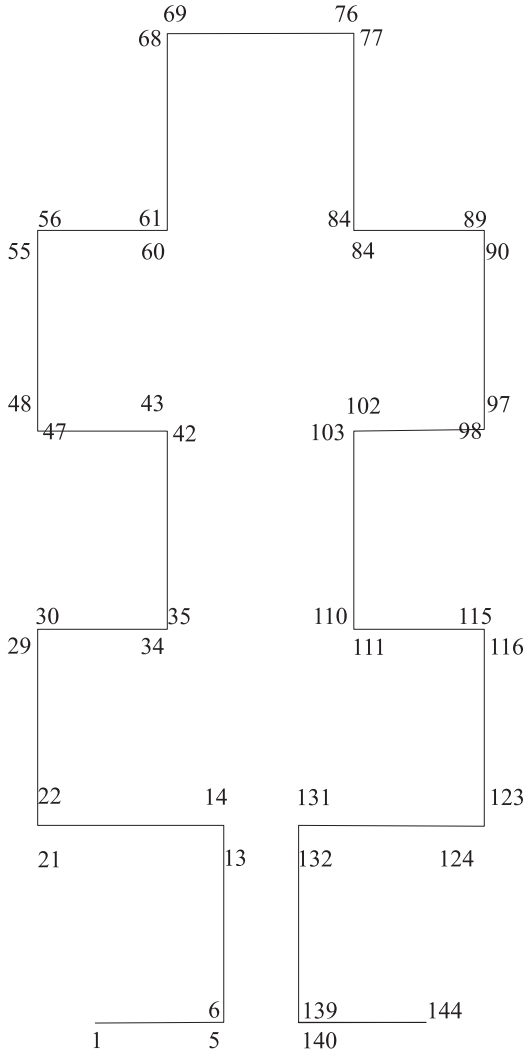

Fig. 2. Scheme of DNA strand nucleosomal loop formation in a two-dimensional space. Arabic figures: orderly numbers of monomers.

ideally we can assume that DNA is a crystal with fragments of a crystal lattice in its structure, with a correct tetrahedron as its obligatory component. This fragment is to strive to the structure of a correct crystal lattice, meeting the requirement of minimum surface-volume economy. However, it is impossible to construct a correct crystal lattice exclusively from tetrahedrons [1]. This lattice, one of whose elements is a correct tetrahedron, can be constructed only from correct tetrahedrons and correct octahedrons with equal sides. Tetrahedrons are constructed on the sides of adjacent octahedrons, while the sides of tetrahedrons are used for construction of new octahedrons, *etc.* [1]. Hence, we conclude that the elementary structure of the unit of a fragment of a correct crystal lattice of a DNA molecule is a structure including correct geometrical figures (tetrahedron and octahedron), constituted from similar components of the DNA macromolecule. Since the monomer order and alternation mode in the chain determine the primary structure of a DNA macromolecule, different pattern of the monomer location should promote the appearance of fragments with different conformations in the polymer. For illustration let us trace the formation of the first least structural unit, participating in the construction of the nucleosome core, a mini-segment consisting of 5 first monomers:

1 2 3 4 5
(P–S)+(S–P)+(S–P)+(P–S)+(S–P)+…

The order of the monomers location in the analyzed segment indicates the formation of dimers of all three possible types. In the first dimer [(P–S)+(S–P)] monomers *1* and *2* form an angle of 60°. The second dimer [(S–P)+(P–S)] is a result of monomers *3* and *4* binding at an angle of 90°. The third dimer [(S–P)+(S–P)] forms as a result of monomers *2* and *3* binding at an angle of 120°. It is involved as a binding component in the formation of the first and second dimers into one structural unit.

Let us note one more possible role of bending deformation of dimers: different bending angles in different dimer types induces modification of the area occupied by the dimer in space, which causes selective positioning of the nucleotides in the chain, depending on the stereometrical parameters of the nitrogen base. Since purines occupy greater volume in space than pyrimidines, combination of the maximally bulky structure with a structure of lesser volume can promote more compact packing of the macrostructure units in space. Hence, the presence of this or that nucleotide in the DNA chain depends on the geometry of its nitrogen base, which, in turn, determines the need in appearance of another nucleotide with nitrogen base, complementary to the base of this nucleotide (in other words, complementarity reflects compactization of DNA structure).

In fact, the mechanism of the nucleosome core fragment assembly is a process repeating (in a certain sequence) all 3 types of dimers, which suggests the formation of a compact structure, that is, the short fragments of the chain with different shape of bending deformation can be packed most compactly. The minisegment of DNA chain carcass including 5 monomers little resembles the unit of a correct crystal lattice constituted (as we have mentioned above) from correct tetrahedron and octahedron. It is more likely a fragment of a broken line assembled from similar components bending at different angles. However, this structure is a unit of a correct crystal lattice, as the monomers constituting it form dimers with angles corresponding to the angles of polygonal geometrical figures of this unit. The fact that it has "vacant" spaces indicates that initially it was similar to a common broken line, but later (in the course of biochemical evolution) all these spaces were presumably filled, and the structural unit of the crystal lattice was fully realized. The "vacant" spaces are most likely the zones in which the DNA repeats are located, serving for improving the stability of the biological system, as the more repeats a system contains, the more reliable it is.

The role of repeats seems to consist in the formation of stable single-strand DNA helix. The first step in this process is the formation of the first pair from two monomers, a dimer. However, the existence of this pair alone is not beneficial, because it is associated with an increase in free energy of the system. Addition of a new monomer leading to appearance of a trimer is the optimal variant, because this process reduces free energy of the system. This structure can be considered as a helix turn formed from 3 similar monomers. This step of the helix (let us call it a tetrahedral helix) is equal to the length of one monomer. It can be presented as a thread coiled around a trihedral prism. As we mentioned, a correct crystal lattice is to include also an octahedron, and it is therefore logical to admit that the second element of the correct crystal lattice, containing repeats, is a structure representing a helix turn consisting already of 4 similar monomers. This step of the helix (let us call it octahedral) is also equal to the monomer length, but the strand is coiled around a sort of a tetrahedral prism. Presumably, at the modern stage of evolution of living nature, only the tetrahedral helix turns are used, while octahedral helix functions no longer. In other words, the information recorded in the octahedral helix is though reproduced, but is actually removed during splicing. It is most likely that the octahedral helix is a structure from which the four-letter codons can be read [17,18]. All this means that the DNA structure is to contain the units of a correct crystal lattice.

REFERENCES

1. T. Varga, *Mathematics* [in Russian], Moscow (1978).
2. M. Gardner, *Mathematical Novels* [in Russian], Moscow (1974).
3. I. B. Zbarskii, *Organization of Cell Nucleus* [in Russian], Moscow (1988).
4. S. A. Lavrov and M. V. Kibanov, *Uspekhi Biol. Khim.*, **47**, 53-88 (2007).
5. Massi J. Stolling, *Algebraic Topology* [in Russian], Moscow (1977).
6. V. V. Nardov, *Practical Manual of Geometrical Crystallography* [in Russian], Leningrad (1974).
7. G. Olcock, *Heterocyclic Compounds and Polymers Based on Them* [in Russian], Moscow (1970).
8. A. Renyi, *The Trilogy of Mathematics* [in Russian], Moscow (1980).
9. R. Feinman, R. Layton, and M. Sands, *Feinman's Lectures in Physics* [in Russian], Moscow (1977), Vol. 7.
10. A. R. Khokhlov, *Liquid Crystal Polymers* [in Russian], Moscow (1988), pp. 32-38.
11. Yu. S. Chentsov, *Introduction in Cell Biology* [in Russian], Moscow (2004).
12. I. P. Shabalkin, *Vopr. Onkol.*, **42**, No. 6, 86-90 (1996).
13. I. P. Shabalkin and P. I. Shabalkin, *Aktual'n. Probl. Sovr. Nauki*, No. 3, 269-271 (2006).
14. I. P. Shabalkin and P. I. Shabalkin, *Kletochn. Tekhnol. Biol. Med.*, No. 4, 211-215 (2008).
15. V. P. Shibaev, *Liquid Crystal Polymers* [in Russian], Moscow (1988), pp. 190-191.
16. *Polymer Encyclopedia* [in Russian], Moscow (1977), Vol. 2, p. 580.
17. M. C. Thomson, J. L. Macfarlane, C. T. Beagley, and D. R. Wolstenholme, *Nucleic Acids Res.*, **22**, No. 25, 5745-5752 (1994).
18. R. Yoshinaga, T. Kakehi, Y. Shima, *et al.*, *Ibid.*, **5**, No. 23, 4830-4834 (1997).