New Molecular Systems. Peculiarities of Structural Organization of Biopolymers

I. P. Shabalkin and P. I. Shabalkin

Translated from *Kletochnye Tehnologii v Biologii i Medicine*, No. 4, pp. 211-215, November, 2008
Original article submitted October 21, 2008.

We proposed a new model of DNA structure, which allows rearrangement of monomers in the polymer chain backbone according to mathematical laws. On the basis of the analysis of structural organization of DNA we concluded that rearrangements of monomers should also occur during the formation of the molecular structure of other polymers (RNA and proteins).

Key Words: *structure; biological polymers; rearrangements; monomers*

Successful development of progressive technologies in biology and medicine is determined, first of all, by the bases for the creation of new technological processes. This implies revision of principles and mechanisms underlying the primary structures of nucleic acids and proteins as the mean elements of living matter deciphered more than 50 years ago. According to L. Pauling model, the backbone of the polypeptide chain consists of relatively simple blocks linked with repeating structures including animo and carboxyl group residues. At this time, Watson and Crick [15] proposed a model of DNA; according to this model, the backbone of the molecule consists of monomers linked via repeating structures including a sugar (deoxyribose) and phosphoric acid residue (phosphate group, PO_4). Hence, both models imply that repeating moieties in the protein and DNA polymer chain backbones should strictly follow each other in the sequence predicted by the model. Elementary structures (blocks) in the chain backbone are bound with rigid covalent bonds, *i.e.* these macromolecules are rigid-chain polymers. The previously proposed models cast some limitation on the structure of the polypeptide chain, which do not allow compact structure of the molecule in a limited space. They also do not take into account some other parameters determining conformation of the polypeptide chain. For instance, both models do not consider the processes involving rearrangements (regrouping) of repeating moieties in the backbone of the polymer chain occurring in practically all polymers [13]. Rearrangements of monomers in the polymer chain lead to the creation of new absolutely predictable molecular biosystems meeting the requirements of optimum bending deformation of the polymer chain and the appearance of new conformations of various local chain points involved in not only specific binding to the ligands (*e.g.* drugs), but also regulating functional activity of polymers. These prerequisites determined the purposes and methods of our investigation. We intended to demonstrate the possibility and necessity of the existence of molecular systems with predicted spatial location of individual fragments of the chain, *e.g.* DNA chain. Analysis of published data drove us to the following conclusions.

First, the existence of a macromolecular polymer system, in particular, DNA is determined by stable modification of the chain backbone, which depends on the following factors:

1) internal energy of structure elements, *i.e.* capacity of these elements to form a spatial lattice;

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

0007-4888/08/1464-0526

2) increase in the symmetry of the structure, *i.e.* its tendency to bipolarity, because monopolar structure is unstable [5];

3) free energy of structure oscillations evaluated by the minimum energy expenditure for bending deformation of the chain.

Second, according to Watson—Crick model, B-form DNA is a rigid rod with limited macroscopic flexibility. Twisting of this long molecule requires extreme energy expenditures. In these rigid macromolecules, only one local conformation is practically possible [10], which does not agree with the existence of various spatial constructions of DNA: long A and B fragments, left-handed Z-form DNA, cruciform structures, loops, *etc*. These facts suggest that the "concept existing until recent time that B-DNA is a perfect double helix and its geometry does no depend on nucleotide sequence is not quite correct" [8].

Third, DNA is a co-polymer. Its backbone includes moieties of two types: sugar (deoxyribose) and phosphoric acid residue (phosphate group). These two moieties are components of a single element, a monomer of DNA macromolecule, because polymerization is the formation of a polymer chain from similar elements. Here we do not discuss the role of nitrogen bases, because they differ by the molecular compositions and structure and the amount of different bases in DNA differs in different organisms. Thus, if DNA molecule is a co-polymer, at least 4 acts of chain elongation are possible during polymerization [13]. In other words, if a single element includes a monomer (sugar-phosphate), the DNA chain can contain rearrangements, *i.e.* different types of fragments:

1) (P—S)+(S—P);
2) (P—S)+(P—S);
3) (S—P)+(S—P);
4) (S—P)+(P—S);

where S and P are sugar and phosphate, respectively.

The second variant corresponds to standard phosphodiester bond in the Watson—Crick model of DNA. The third variant also can occur in the Watson—Crick DNA model, because it can be presented in the inverted form: (S—P). At the same time, the first and fourth variants are not presented in this model. The fact that Watcon—Crick DNA model does not take into consideration possible rearrangements of monomers is one of the main drawbacks of this model, because different types of interactions between the monomers in the sugar-phosphate backbone of the polymer DNA chain allow identification of relative unstable segments and sites of predominant location and coordination binding of ligans and this can determine changes in the geometry of the nucleotide in DNA structure. Thus, it can be assumed with high probability that rearrangements of monomer links in DNA will increase flexibility of the molecule, because the appearing segments of the structure can have different strength degree (*i.e.* the energy required for breaking or formation of a bond between atoms in one segment differs from that in another segment). Since the phosphodiester bond P—O—C includes elements of both P—O and C—O bonds, the differences in the character of bonds between different heteroatoms in this case consist in evaluation of the properties of P—O and C—O bonds. Ionic nature of bonds can be one of these factors. It is known that bonds including more pronounced ionic component are associated with high bond energy, *i.e.* a correlation exists between ionic component and increased strength of the bond [6].

According G. Alcock data, P—O and C—O are ionic bonds by 39 and 22%, respectively [6]. Hence, evaluation of the ionic nature of the bond shows that P—O bond is more stable than C—O bond. This conclusion is also confirmed by other results [12]. For instance, the energies of P—O and C—O bonds are 342 and 332 kJ/mol. Higher energy of P—O bond attests to participation of this bond in thermal stability of the polymer. Apart from thermodynamic stability, the P—O bond is characterized by the resistance to hydrolysis. According to published data [3], the reaction of non-ionized oxygen ester with water during hydrolysis of phosphoric acid diesters proceeds via cleavage of C—O bond and only in 10% cases via cleavage of P—O bond. Other authorities [3] also showed that hydrolysis proceeds via cleavage of C—O bond in 78% cases and via cleavage of P—O bond in 22% cases.

These facts suggest that in the polymer DNA chain, the C—O bond is less stable than P—O bond. The presence of segments characterized by different strength of intermonomer bonds in the DNA chais extends the possibility of spatial DNA compactization. This suggests that DNA molecule corresponds to flexible-chain rather than to rigid-chain polymers consisting of rigid and flexible segments.

This reasoning gives the idea on the model of structural organization of DNA, but does not answer the question whether the formation of primary DNA structure in the course of chemical evolution was a random or a regular process. Bearing in mind that case is just a form of unknown regularity and that harmony of the universe is a harmony of numbers [11], we assumed that natural polymerization

of DNA satisfies a mathematic regularity known as Fibonacci sequence. Fibonacci numbers are elements of numerical sequence 1;1;2;3;5;8;13;21… (Fibonacci sequence), where each member starting from the third one is equal to the sum of two previous members [7]. Taking into account possible existence of inverted monomers [13], let us review the process of formation of primary single-strand DNA structure with the assumption that the first member in the numerical sequence (let us denote it 1^a) is a monomer where phosphate ("head") is in the first position and sugar ("tail") is in the second position: (P—S). Then the second member in the Fibonacci sequence (1^b) is an inverted monomer (S—P). Hence, each member of the Fibonacci sequence consisting of a certain number of moieties of DNA chain can be presented as follows:

1^a=(P—S);
1^b=(S—P);
$2=1^b+1^a$=[(S—P)+(P—S)];
$3=2+1^b$=[(S—P)+(P—S)+(S—P)];
$5=3+2$=[(S—P)+(P—S)+(S—P)+(S—P)+(P—S)]
etc.

Sequential binding of next member of the Fibonacci sequence to the previous ($1^a+1^b+2+3+5+8+\ldots$) yields a fragment of primary structure of DNA molecule:

(P—S)+(S—P)+[(S—P)+(P—S)]+[(S—P)+
(P—S)+(S—P)]+[(S—P)+(P—S)+(S—P)+
(S—P)+(P—S)]+[(S—P)+(P—S)+(S—P)+
(S—P)+(P—S)+(S—P)+(P—S)+(S—P)]+
[(S—P)+*etc.*

This fragment of DNA consisting of 21 moieties is characterized by the following peculiarities. First, polymerization via addition reactions is mediated by three types of bonds:

1) dimers with phosphodiester bond P—O—S, [(S—P)+(S—P)];
2) dimers with phosphodiester bond P—O—P, [(S—P)+(P—S)];
3) fragments with glycoside bonds S—O—S, [(P—S)+(S—P)];

The first variant is typical of Watson—Crick DNA model.

The second variant attests to possible presence of rigidly bound monomers in the DNA molecule, because two phosphate groups located together (second type of intermonomer bonds) considerably reduce the conformation space because of tight contact between these groups. Hence, these fragments of DNA polymer chain should have limited number of conformations [2]. The third variant is the bond between two sugars (P—S)+(S—P). Published data suggest that sugar—sugar contacts lead to the formation of DNA chain bending [2]. We can hypothesize that in these points two monomers are bound via an anhydride bond: their phosphate groups form 5',5'-pyrophosphate bridge similar to that in NAD coenzyme [4].

After reviving the principles of organization of the primary structure of one chain (leading chain) of DNA molecule, we should discuss the peculiarities of the second (lagging) chain of the molecule. The first member of Fibonacci sequence ($1^{à}$) for the lagging DNA chain is (S—P), the second (1^b) is (P—S):

1^a=(S—P);
1^b=(P—S);
$2=1^b+1^a$=[(P—S)+(S—P)];
$3=2+1^b$=[(P—S)+(S—P)+(P—S)];
$5=3+2$=[(P—S)+(S—P)+(P—S)+
(P—S)+(S—P)] *etc*.

This yields a fragment of primary structure of the lagging DNA chain consisting of 21 monomers:

(S—P)+(P—S)+[(P—S)+(S—P)]+[(P—S)+
(S—P)+(P—S)]+[(P—S)+(S—P)+(P—S)+
(P—S)+(S—P)]+[(P—S)+(S—P)+(P—S)+
(P—S)+(S—P)+(P—S)+(S—P)+(P—S)]+
[(P—S)+*etc.*

During the study of the primary DNA structure, a question arose about the size of a minimum fragment of single DNA strand that can be a moiety participating in the formation of a supramolecular structure of the molecule. To this end, we analyzed division of a linear segment into two parts according to the golden proportion rule. The golden section principle is based on the fundamental property: the ratio of the greater part to the total segment is equal to the ratio of the smaller part to the greater part [7]. In numerical terms, the equality of these two ratios (golden proportion) is expressed by an irrational number ≈0.618…, i.e. m/M=M/T (where m and M are the smaller and greater parts and T is the total).

Since the backbone of the primary structure of DNA molecule is a linear structure consisting of monomer moieties of the same length, we can select a mini-fragment that is the first building block of DNA molecule. Imagine that we divide a single-strand DNA fragment consisting of 21 monomers.

The process of section of DNA fragment can be divided into steps, when each of two unequal parts of the initial fragment (let us call them segments) is also a fragment (but shorter than the initial one!), which can be divided according to the golden proportion principle until the equality of two proportions is true (Fig. 1).

Step I. The total fragment is divided in point O into two unequal parts (segments) according to the "golden proportion" principle: m/M=M/T, *i.e.* the smaller segment includes 8 monomers, the greater one includes 13 monomers (8/13=13/21, or 8/13≈0.615; 13/21≈0.619). This section of the initial fragment into two parts retains the equality of the two ratios to the number 0.62. Thus, the quantitative difference between the two constituents of the initial fragments at this step is described by the distribution 1:1.

Step II. Section of each segment of the initial fragment yields four new segments, the constituents of the initial fragment. They are distributed in the following way: the smallest segment consists of 3 monomers, the two longer segments include 5 monomers, and the longest segment contains 8 monomers. This difference by the length between the segments is expressed as the ratio 1:2:1. It should be noted that in contrast to step I, when the constituents of the total segments had only quantitative differences, at this stage we see qualitative differences between the segments derived from the initial fragment. For instance, two segments formed during step II do not differ by the length (each includes 5 monomers), but have qualitative differences: one of them is a smaller segment resulting from section of a 13-monomer fragment, while the other is a greater segment appearing after section of a 8-monomer fragment (Fig. 1).

At **step III** the process of section of the initial fragment into constituent segments is practically completed, because during **step IV** this process eventuates in the formation of just two segments from one mini-fragment: one of them includes 2 monomers and the other contains 3 monomers.

Since the last segment containing 3 monomers cannot be divided into unequal parts in accordance with the golden proportion principle [1/2=0.5≠2/3=0.67], the last mini-fragment that will determine what number of monomers constitutes the initial DNA fragment should consist of at least 5 monomers dividable according to the golden proportion principle: 2/3=3/5.

It should be noted that if the selected fragment consists of greater number of monomers (*e.g.* 34), other distributions will be obtained. 1:3:3:1 *etc.* However, all these distributions will lead us to the mini-fragment containing at least 5 monomers. This suggests that division of the initial long fragment into smaller segments according to the golden proportion principle is described by a mathematic regularity known as Pascal triangle [1,9]:

	1
Step I	1 : 1
Step II	1 : 2 : 1
Step III	1 : 3 : 3 : 1
	etc.

Moreover, this regularity, as is seen from step II, can reflect qualitative differences between quantitatively identical elements of the same nature.

Thus, the following conclusions can be made: since the order and way of binding of monomers determine the primary structure of DNA macromolecule, different character of monomer positioning in the chain leads to the appearance of fragments with different conformation within the polymer. The proposed DNA model also implies that the minimum fragment in each strand involved into

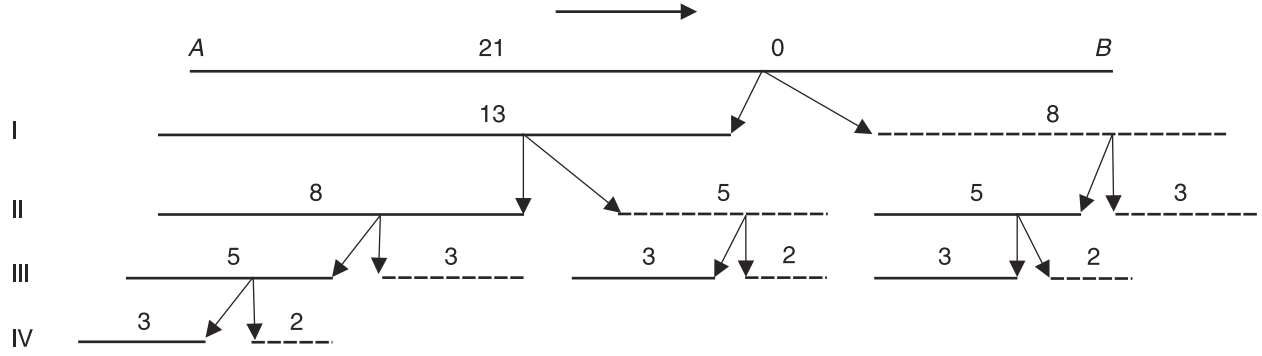

Fig. 1. Steps of division of fragment AB into constituent segments according to the golden proportion principle. I-IV: steps of division of the initial fragment to a minimum size meeting the principle of golden proportion. Solid line: greater segment, interrupted line: smaller segment at different stages of division of the initial fragment. Numbers correspond to the number of monomers (8; 13 …). Arrow: division of the segment in the start-end direction.

the formation of a supramolecular DNA structure calculated according to the golden proportion principle should include 5 monomers.

Thus, the model of structural organization of DNA molecule proposed by us implies the possibility of rearrangement of monomers in the molecule backbone determined by mathematical laws. Taking into account the fact that transcription is the synthesis of matrix RNA (mRNA) carrying information encoded on the transcribed DNA locus, while mRNA during translation is used as the template determining the sequence of amino acids in the synthesized protein polymer, mRNA and protein molecules in general reproduce the order and way of binding of monomers in the DNA backbone. Rearrangements of monomers should be present in the primary structure (backbone) of both mRNA and proteins. For instance, if we denote amino group (—NH—) in the peptide bond (CO—NH) of a protein as "A", while the double bond C=O as "C", the backbone of the primary structure of this protein chain can contain dimers with 4 different types of rearrangements: (A—C)+(C—A), (A—C)+(A—C), (C—A)+(C—A), (C—A)+(A—C). Hence, rearrangements in the backbone of DNA, mRNA, and proteins occur during polymerization of these biopolymers, *i.e.* this process is a universal phenomenon.

REFERENCES

1. N. Ya. Vilenkin, *Popular Combinatorics* [in Russian], Moscow (1975).
2. W. Zenger, *Principles of Structural Organization of Nucleic Acids* [Russian translation], Moscow (1987).
3. K. Ingold, *Theoretical Bases of Organic Chemistry* [Russian translation], Moscow (1973), pp. 978-983.
4. A. Leninger, *Biochemistry* [Russian translation], Moscow (1976), pp. 294-300.
5. N. N. Moiseev, *Universum. Information. Community*. [in Russian], Moscow (2001).
6. G. Alcock, *Heterocyclic Compounds and Polymers on the Basis of Heterocyclic Compounds* [Russian translation], Moscow (1970), pp. 45-46.
7. A. Renyi, *Trilogy about Mathematics* [Russian translation], Moscow (1980).
8. M. Singer and P. Berg, *Genes and Genomes* [Russian translation], Moscow (1998), Vol. 1, p. 44.
9. A. P. Stakhov, *Metaphysics* [in Russian], XXI Century [in Russian], Moscow (2006), pp. 174-215.
10. A. P. Khokhlov, *Liquid-Crystalline Polymers* [in Russian], Moscow (1988), pp. 32-38.
11. I. Sh. Shevelev, *Golden Section* [in Russian], Moscow (1990), pp. 6-128.
12. *Encyclopaedia of Polymers* [in Russian], Moscow (1977), Vol. 1, p. 104.
13. *Ibid.,* Vol. 2, pp. 101-102.
14. *Ibid.,* Vol. 3, pp. 532.
15. J. D. Watson and F. H. C.Crick, *Nature*, **171**, 737-738 (1953).