Volume 11 Number 15 1983 Nucleic Acids Research

Possible conformations of double-helical polynucleotides containing incorrect base pairs

Vasily P.Chuprina and Valery I.Poltev

Institute of Biological Physics, Academy of Sciences of the USSR, Pushchino, Moscow Region, USSR

Received 15 June 1983; Accepted 13 July 1983

ABSTRACT

Theoretical conformational analysis using classical potential functions has shown the possibility of incorporation of nucleotide mispairs with the bases in normal tautomeric forms into the DNA double helix. Incorrect purine-pyrimidine, purine-purine and pyrimidine-pyrimidine pairs can be incorporated into the double helix existing both in A- and B-conformations. The most energy favourable conformations of fragments containing a mispair have all the dihedral angles of the sugar-phosphate backbone within the limits characteristic of double helices consisting of Watson-Crick nucleotide pairs. Incorporation of mispairs is possible practically without the appearance of reduced interatomic contacts. Mutual position of bases in the incorporated mispair does not differ much from their position at the energy minimum of the corresponding isolated base pairs. Conformational parameters of irregular regions of double-stranded polynucleotides containing G:U, I:A, I:A*(syn) and U:C pairs are presented. Distortion of the sugar-phosphate backbone is the least upon incorporation of the G:U pair. Formation of mispairs in the processes of nucleic acid biosynthesis and spontaneous mutagenesis is discussed.

INTRODUCTION

Transfer of genetic information from one generation to another and its realization are ensured by exclusiveness of Watson-Crick adenine-thymine (uracil) and guanine-cytosine pairs. The fact that in the processes of nucleic acid biosynthesis only these pairs are usually formed is due both to the properties of nitrogen bases (coincidence of dimensions of A:U (A:T) and G:C pairs, their adjustment to the sugar-phosphate backbone in the double helix, specificity of intermolecular interactions) and to the effect of enzymes which increase accuracy of nucleic acid biosynthesis. The formation of any other base pairs at biosynthesis leads to point mutations, transitions (if guanine-thymine or adenine-cytosine pairs are formed) or transversions (if any of purine-purine or pyrimidine-pyrimidine pairs is formed). To elucidate the mechanism of transfer and realization of genetic information, the ways in which mutagenic base analogs act and the effect of chemical modification of nucleic acids on their role as carriers of genetic information,

it is necessary to construct a detailed molecular model of nucleic acid fragments containing mispairs.

The formation of G:U (G:T) and A:C pairs allowing their incorporation into the double helix without its deformation is possible when one of the bases is in a minor tautomeric form. Such a mechanism of spontaneous mutations (transitions) was proposed by Watson and Crick in 1953 (1) and has been repeatedly discussed since then. Topal and Fresco (2) applied this mechanism to transversions suggesting that they are induced by the formation of purine-purine pairs. The authors assumed that one base must be in a minor tautomeric form while the other nucleotide is in the syn-conformation. However, in two of four such pairs one base must be in the imino-lactim tautomeric form, the probability of whose formation is extremely low, i.e. in a "twice minor" form.

In papers (3-4) a molecular mechanism of all spontaneous transitions and transversions leading to mispairing of bases in normal tautomeric forms has been proposed and energy of such pairs has been calculated. The analysis of experimental data and the results of calculations led to formulation of a concept of enzyme-nucleotide recognition in the processes of nucleic acid biosynthesis and the recognizing site of the enzyme of template directed synthesis (3-5). According to this concept, the interaction of the enzyme recognizing site with structural invariants of complementary nucleotide pairs increases the probability of such pairs formation and prevents mispairing. The difference in mutual position of bases in incorrect pairs from that in A:T and G:C pairs is a constraint for incorporation of incorrect bases into the double helix at biosynthesis because of the strain in the sugar-phosphate backbone and disturbance of the interaction of nucleotide pairs with the enzyme recognizing site. This difference and the energy of interaction of bases in pairs affect the probability of such pair formation at nucleic acid biosynthesis.

Calculations (3-4) have shown that the mutual position of bases in mispairs corresponding to the energy minimum differs from that in Watson-Crick pairs but for many base pairs there are minima in which this difference is not large enough. There is a qualitative agreement between the relative probabilities of mispairing predicted by calculations (3,4) and the experimentally observed frequencies of base replacements.

Theoretical conformational analysis of regular double-helical polynucleotides has shown that there are regions of minimum energy values of intramolecular interactions corresponding to A- and B-families of nucleic acid con-

formations (6-8). Within these regions different mutual positions of the base pairs at rather close values of dihedral angles of the sugar-phosphate backbone are possible (8). On the other hand, a search for low-energy conformations of complexes of dinucleoside phosphates has shown that considerable deviations of the double-stranded structure from regularity are possible without noticeable energy increase (9). This conformational freedom permitted us to hope that incorporation of incorrect base pairs in normal tautomeric forms into the double helix is possible without any considerable distortion of the sugar-phosphate backbone.

The present study is devoted to elucidation of such a possibility and determination of conformational parameters of DNA fragments containing mispairs in normal tautomeric forms. Calculating the energy of nonbonded interactions and searching for the minimum of such energy as a function of conformational variables we found conformations of double helices consisting of Watson-Crick pairs and containing incorporated pairs G:U, I:A, I:A*(syn) and C:U. This incorporation practically does not result in reduction of interatomic contacts and the angles of the sugar-phosphate backbone remain within the limits characteristic of regular double helices. As a result, we have constructed models of nucleic acid fragments containing nucleotide mispairs. Computations confirm the earlier suggestion (3-5) that recognition of correct nucleotide pairs by the enzymes is necessary to ensure the experimentally observed high accuracy of nucleic acid biosynthesis.

METHODS

Energy of nonbonded interactions has been calculated by the method of atom-atom potential functions. Details of the method application to the problems of structure and functioning of nucleic acids and the choice of parameters of potential functions were described earlier (10,11).

Energy was calculated as a sum of electrostatic and van-der-Waals atom-atom interactions, torsional energy of rotation about sugar-phosphate backbone bonds and energy of distortion of variable bond angles.

The energy of interaction between i and j atoms was calculated according to the equation

$$U_{ij}(r) = k \cdot \frac{e_i \cdot e_j}{\varepsilon \cdot r} - A_{ij} r^{-6} + B_{ij} \cdot r^{-12}$$

where r is the interatomic distance, e_i and e_j are charges of i and j atoms, k is the coefficient depending on the chosen system of units, A_{ij} and B_{ij} are the parameters depending on the type of i and j atoms, ε is the effective

dielectric constant. All bond lengths and bond angles (except internal angles of the sugar ring and the C_3'-O-P angle) are presumed to be rigid. The parameters for calculating energy as well as bond length and bond angles are the same as in our paper dealing with regular double-helical polynucleotides (8). The search for minimum energy values was performed by the method of parallel tangents. In the minimization process electrostatic interactions only between bases in the pairs were taken into account. When the minimum was found with such an assumption, further minimization was performed taking into account electrostatic interactions between all atoms, but this did not lead to a considerable shift of the minimum. The found low energy conformations were examined for the presence of reduced interatomic contacts, then they were visualized with the help of a plotter according to the programs making use of the complex of GRAFOR subroutines. The criteria for evaluation whether an interatomic contact is reduced were taken from (12). Some interatomic distances were also considered if they corresponded to the van-der-Waals energy more than 0.4 kcal/mole though they were somewhat larger than the reduced interatomic contacts as defined in (12). Computations and visualization were done using a ES 1040 computer (Computer Research Center, Scientific Center of Biological Research, Academy of Sciences of the USSR).

Fragments of double-helical polynucleotides consisting of two subsequent Watson-Crick A:U pairs and one of the G:U, A:I, I:A* or C:U mispairs were considered. Energy of nonbonded interactions has been calculated as a function of the following variables: 1) parameters which determine the mutual position of bases in pairs (all pairs are presumed to be planar, the position of bases in canonical pairs is fixed); 2) parameters determining the position of pairs relative to the helix axis; 3) helix parameters (the distance d between neighbouring pairs along the helix axis and rotation angles τ); 4) angles χ determining mutual positions of bases and sugars in nucleosides according to Arnott (13); and 5) angles determining conformations of sugar rings (two valent and two dihedral angles for each ring). Coordinates of atoms and dihedral angles of the sugar-phosphate backbone have been determined from independent conformational parameters using an algorithm similar to that described earlier for regular double-helical polynucleotides (6).

First, to reduce the number of independent variables, calculations were performed with the assumption that the conformational parameters of all four nucleosides constituting Watson-Crick pairs and the parameters determining the position of these pairs relative to the helix axis are equivalent while the parameters of the mispair nucleosides can differ from each other and

from the corresponding parameters of canonical fragment pairs. In this case the sugar-phosphate backbone is distorted only between the incorrect and the nearest canonical pairs and the regular helix can be constructed by repeating the fragment consisting of two canonical pairs. Thus we have obtained conformations of the fragments containing the G:U pair (A- and B-forms). For other fragments the attempts to incorporate incorrect pairs lead, at such assumptions, to a considerable increase of energy and to a great deviation of the angles of the sugar-phosphate backbone from their values for the regular Watson-Crick helices. Therefore for I:A, I:A* and C:U pairs we calculated the energy as a function of more variables assuming that the parameters of nucleosides of an incorrect pair and the neighbouring correct pair differ while the parameters of two nucleosides of the next correct pair are equivalent. Then we found conformations of regular double-stranded polynucleotides which could continue these fragments from the side of canonical pairs. The parameters characterizing the position of bases in pairs, that of pairs relative to the helix axis, mutual positions of sugars and bases as well as sugar ring conformations are the same as for the nucleotide pairs situated on the boundary of irregular regions. Helical parameters d and τ of regular conformations can slightly (no more than by 0.15 Å and 0.5°) differ from the corresponding parameters of the two canonical pairs of an irregular fragment. Conformations of regular helices continuing each of irregular conformations are close to the bottom of the corresponding valley of minimum energy values, but can be less favourable by 1÷3 kcal/mole and somewhat differ by conformational parameters.

For each incorrect X:Y pair we considered two fragments dApdApdX:dYpdUpdU and dXpdApdA:dUpdUpdY containing the incorrect pair on one or another side of the successive correct pairs. One of these fragments results from incorporation of the incorrect base X during synthesis of Y-containing template and the other from incorporation of the incorrect base Y during synthesis on the X-containing template. At DNA synthesis on a template one fragment corresponds to incorporation of an incorrect base and the second one to elongation of synthesis after this incorporation.

When searching for low-energy conformations we took as the starting parameter values those for regular polynucleotides consisting of canonical A:U pairs (8). Then the bases in the incorrect pair were gradually shifted in small steps to the position corresponding to the energy minimum of interaction of isolated bases constituting this pair. At each step the mutual position of bases in the incorrect pair was fixed and energy was minimized by other independent variables. When the bases achieved the position close to that for the energy

minimum in the isolated incorrect pair, minimization was done by all independent variables. For all low-energy conformations found (except that with the C:U pair) the energy of interaction of bases in the incorrect pair is only 0.1÷0.5 kcal/mole higher than the minimum and the mutual position of bases in these pairs is similar to that at the energy minimum for isolated pairs differing by no more than 0.3 Å and 5°.

Conformations with the parameters of sugar-phosphate backbone and glycoside bond positions characteristic of regular polynucleotides consisting of A:U pairs are sterically allowed for the fragments containing coplanar uracyl and cytosine. The energy of such conformations is about 6 kcal/mole higher than for the most favourable conformations of fragments with U:C pairs, mainly owing to a worse interaction between U and C bases. The most favourable conformations of fragments with the U:C pair correspond to the mutual position of bases when the distance between U and C is about 0.5 Å larger and the interaction energy in the pair 2 kcal/mole higher than in the isolated base pair having the minimum energy. Conformations of fragments containing the U:C pair with two hydrogen bonds and the energy close to the minimum which will be discussed in the next section are only 0.5÷1.0 kcal/mole less favourable energetically.

For low-energy conformations of fragments dApdApdX:dYpdUpdU and dXpdApdA:dUpdUpdY the difference in conformational parameters of the nucleoside X:Y pair is a few degrees by the angle variables and some tenths of Å by the variables characterizing shifts. The energy of the two sequences can differ by 1÷3 kcal/mole. Then the conformational parameters of nucleoside mispairs of the two fragments were made equivalent by shifting them gradually towards each other and minimizing by other variables at each step. The total energy of nonbonded interactions of each fragment increases by 1÷3 kcal/mole. Thus, we have found possible conformations of dApdApdXpdApdA: :dUpdUpdYpdUpdU, i.e. conformations of fragments of double-chain polynucleotides containing a mispair incorporated between correct pairs. These conformations will be discussed below.

RESULTS AND DISCUSSION

Conformational parameters of fragments

Characteristics of the obtained conformations of irregular DNA fragments containing mispairs are listed in Tables 1-4. It should be noted that the given values of the parameters correspond in each case to one low-energy conformation, to one point within an extended region of minimum energy values.

Table 1. Conformational parameters of DNA fragments containing mispairs (B-form)

	d(Å)	τ	D(Å)	ω	φ	ψ	θ	ξ
U:A	3.36	36.9	9.7 9.7	184 184	249 249	-70 -70	182 182	60 60
U:A	3.34	33.3	9.5 9.3	205 177	228 252	-71 -73	165 192	61 59
U:G	3.42	41.1	9.6 9.7	177 177	260 257	-72 -70	184 188	63 58
U:A	3.39	36.3	9.7 9.7	180 180	252 252	-67 -67	185 185	57 57
U:A								
U:A	3.53	39.2	9.3 9.2	187 190	241 246	-66 -74	175 180	62 63
U:A	3.43	32.4	9.8 10.2	168 210	272 226	-73 -75	200 161	53 64
I:A	3.47	35.0	9.8 10.0	213 168	225 272	-80 -73	158 197	68 55
U:A	3.51	39.2	9.3 9.2	191 187	245 241	-73 -69	179 175	62 65
U:A								
U:A	3.42	36.4	9.6 9.6	186 180	253 254	-68 -70	185 186	57 58
U:A	3.39	31.5	9.5 9.6	186 177	245 257	-74 -62	185 190	60 51
I:A	3.50	42.1	9.7 9.4	184 177	250 258	-73 -70	179 186	65 61
U:A	3.27	37.1	9.6 9.6	179 178	253 253	-67 -68	187 187	56 56
U:A								
U:A	3.34	38.3	9.4 9.4	182 185	246 247	-71 -74	181 180	64 65
U:A	3.63	38.4	9.0 8.9	216 168	224 273	-73 -76	154 201	66 55
U:C	3.58	37.6	9.1 8.8	170 222	271 220	-74 -73	200 150	53 64
U:A	3.33	38.4	9.4 9.3	186 181	246 245	-71 -71	178 180	64 65
U:A								

For each fragment the table gives four d and τ values corresponding to the distances and rotation angles between each two successive base pairs as well as eight values for all angles of the sugar-phosphate backbone and the phosphour atom - helix axis distance (D). The arrows indicate the direction of the sugar-phosphate backbone from C_3' to C_5'. For each angle the left value corresponds to the sequence dUpdUpdYpdUpdU and the right one to the sequence dApdApdXpdApdA. The dihedral angles are measured as follows: ω (C_4'-C_3'-O_3'-P), φ (C_3'-O_3'-P-O_5'), ψ(O_3'-P-O_5'-C_5'), θ(P-O_5'-C_5'-C_4'), ξ(O_5'-C_5'-C_4'-C_3').

As in the case of regular double-stranded polynucleotides (6-8), the fragments considered have extended regions of minimum energy values, i.e. valleys corresponding to A- and B-families of conformations. Since for irregular regions the number of independent conformational variables is large (more than 40 for each three successive nucleotide pairs), the whole region of minimum energy values has not been studied as this would require too many calculations. Conformations of regular double-stranded polynucleotides which could continue these fragments in both directions have been found for each conformation presented.

Fig. 1. represents poly-dA:poly-dU double helices with incorporated I:A mispairs and demonstrates similarity of conformations of regular and irregular fragments. Such resemblance was revealed for other mispairs. Figs. 2 and 3 represent irregular fragments consisting of 3 pairs (with a mispair on the

Table 2. Conformational parameters of DNA fragments containing mispairs (B-form).

	θ_1	χ	σ $O'_3-C'_3-C'_4-C'_5$	P
U:A	-3.2	127.4	140.6	155.1
U:A				
U:G	-4.1	128.2 128.2	141.2 140.5	156.3 155.1
U:A				
U:A	-4.4	128.4	139.6	155.0
U:A	-5.6	126.4	143.8	160.6
U:A	-5.2	126.9 130.2	133.2 140.7	144.9 160.0
I:A	-3.6	128.0 128.3	141.3 139.4	159.3 153.6
U:A	-4.9	129.7 127.2	140.8 134.0	158.1 146.1
U:A	-4.9	126.7	143.6	160.1
U:A	-4.2	127.4	138.8	155.4
U:A	-4.1	127.4 127.7	138.9 138.1	155.7 154.6
I:A*	-3.6	128.0 306.8	142.5 133.0	154.1 144.6
U:A	-3.3	129.2 129.2	141.3 141.4	158.6 155.6
U:A	-3.2	129.0	141.3	158.5
U:A	-4.7	126.8	143.5	159.1
U:A	-4.9	128.8 126.2	139.4 143.0	153.8 157.8
U:C	-7.8	128.5 128.4	137.1 136.7	151.0 150.2
U:A	-5.0	125.4 129.3	142.9 139.1	157.8 152.9
U:A	-4.2	127.0	143.4	159.1

Each base pair has its tilt value θ_1 (angle between the base pair plane and that perpendicular to the helix axis). Each nucleoside of incorrect and neighbouring correct pairs has its own value of angles χ (C_8(Pu) or C_6(Py)-N-C'_1-C'_2), σ and P, the pseudorotation angles of sugar (14). For the fragment with the G:U pair the parameters of the two nucleosides of each correct pair are equivalent. Conformations are the same as in Table 1.

boundary). The mutual position of bases in each conformation corresponds to the formation of two practically linear hydrogen bonds in mispairs.

The data of Tables 1 and 3 suggest the existence of an unfolding effect for I:A and U:C pairs, i.e. the mean values of the helical angle between a mispair and the neighbouring canonical pairs is somewhat smaller than for regular polynucleotides. The distances between base pairs (d) in the irregular fragments are, as a rule, slightly larger, especially between a mispair and the neighbouring pairs, than the optimal distance for the conformations of the corresponding family of regular polynucleotides consisting of Watson-Crick pairs. A decrease of d and an increase of τ without an essential energy increase would be, however, observed when moving along the bottom of the corresponding valley.

The values of χ angles are within a rather narrow range. It is even nar-

Table 3. Conformational parameters of DNA fragments containing mispairs (A-form). Designations are the same as in Table 1.

	d(Å)	τ	D(Å)	ω	φ	ψ	θ	ξ
U:A	2.98	30.5	9.6 9.6	205 205	-69 -69	-74 -74	180 180	60 60
U:A	3.12	27.3	9.6 9.7	210 198	-84 -64	-77 -72	169 188	65 56
U:G	3.09	32.2	10.1 10.1	212 207	-70 -66	-75 -71	174 173	63 62
U:A	3.08	29.7	10.1 10.1	207 207	-71 -71	-77 -77	176 176	67 67
U:A								
U:A	3.28	30.1	9.6 9.6	203 201	-71 -66	-72 -73	179 182	60 58
U:A	3.35	28.6	9.7 9.8	203 227	-59 -98	-74 -69	193 149	53 69
I:A	3.46	26.2	10.0 10.0	242 202	-103 -56	-74 -74	133 189	79 57
U:A	3.25	30.0	9.8 9.8	203 203	-69 -71	-76 -71	180 177	62 60
U:A								
U:A	3.11	30.7	9.4 9.4	205 205	-70 -70	-75 -72	180 179	61 58
U:A	3.31	24.5	9.4 9.5	209 205	-73 -71	-75 -64	178 183	60 56
I:A*	3.09	35.4	9.5 9.7	210 203	-74 -61	-73 -74	171 180	65 60
U:A	3.21	30.1	9.6 9.6	203 203	-68 -69	-74 -72	180 180	60 59
U:A								
U:A	2.95	31.8	9.4 9.3	208 207	-72 -69	-75 -79	177 177	61 62
U:A	3.46	29.1	9.2 9.4	242 205	-108 -61	-64 -71	134 188	68 56
U:C	3.46	28.6	9.3 9.3	204 242	-60 -110	-73 -66	189 134	56 68
U:A	2.94	31.7	9.4 9.3	209 209	-72 -72	-78 -75	177 177	63 61
U:A								

Table 4. Conformational parameters of DNA fragments containing mispairs (A-form). Designations are the same as in Table 2.

	θ_1	χ	σ $O'_3-C'_3-C'_4-C'_5$	P
U:A	11.4	76.4	78.9	17.5
U:A				
U:G	8.6	75.7 76.2	80.2 78.2	18.2 20.0
U:A	7.1	72.5	78.4	23.9
U:A				
U:A	7.4	77.8	79.5	17.9
U:A	7.1	76.5 78.2	78.3 79.6	19.8 17.5
I:A	11.6	71.7 75.9	80.0 81.7	17.4 13.0
U:A	7.1	77.0 75.9	79.5 77.9	20.9 23.5
U:A	6.7	76.2	79.6	21.0
U:A				
U:A	9.7	76.1	79.5	15.5
U:A	9.4	76.5 76.1	79.4 79.5	15.6 15.3
I:A*	9.1	74.1 257.9	80.6 75.9	15.9 24.3
U:A	8.4	77.4 77.2	79.3 78.4	17.1 18.4
U:A	8.3	77.2	79.4	17.0
U:A				
U:A	13.1	76.0	80.1	16.2
U:A	12.7	75.6 74.8	79.0 79.0	17.3 17.4
U:C	7.9	76.6 76.6	78.6 78.5	20.2 20.4
U:A	12.9	75.0 75.7	79.4 80.2	17.1 16.2
U:A	12.6	75.1	79.9	16.7

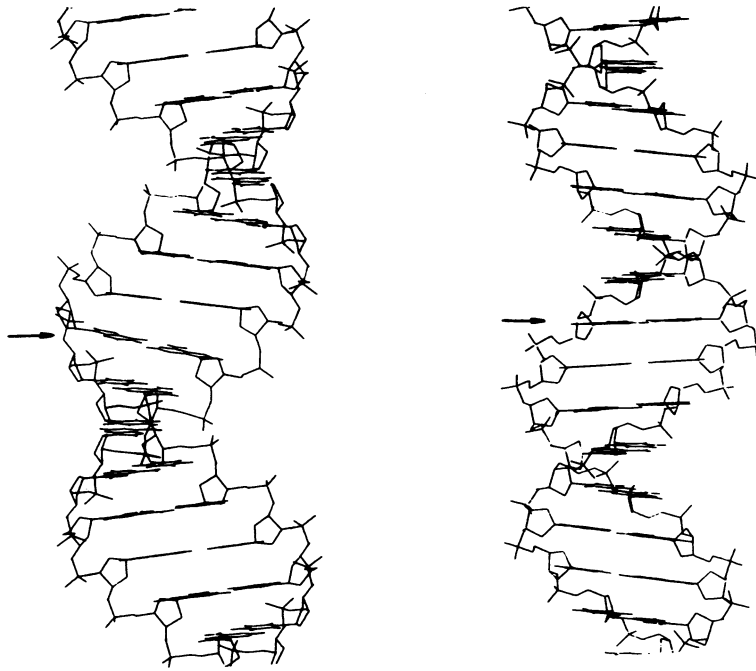

Fig. 1. Double-helical polynucleotides containing I:A mispairs. Projection on the plane of the helix axis. Left, the A-form, right, the B-form. C-H bonds in the sugar-phosphate backbone are not shown. The arrows indicate the mispair positions.

rower than that reported by us (8) for the constraints imposed on the two bases of the pair to lie in one plane. Conformations of sugar rings correspond to C_2'-endo forms for the B-family and to C_3'-endo forms for the A-family. Many fragments have a small displacement of the second of C_2' and C_3' atoms from the $C_1'O_1'C_4'$ plane in the same direction, i.e. C_2'-endo-C_3'-endo puckering for the B-family and C_3'-endo-C_2'-endo puckering for the A-family.

Table 5 summarizes the ranges within which the dihedral angles of the sugar-phosphate backbone change at incorporation of mispairs. For the B-family the table gives the limit values of these angles in a synthetic dodecanucleotide crystal (15) and for the A-family - the limit values of the corresponding angles of phenylalanine tRNA double-helical fragments in crystals of two different classes (16,17). Ranges of these angles changes as calculated by Zhurkin et al. (7) are given for comparison. The ranges of values of dihedral angles in regular double-helical polynucleotides are less extended according to other authors and our data (8) because they were estimated for a smaller

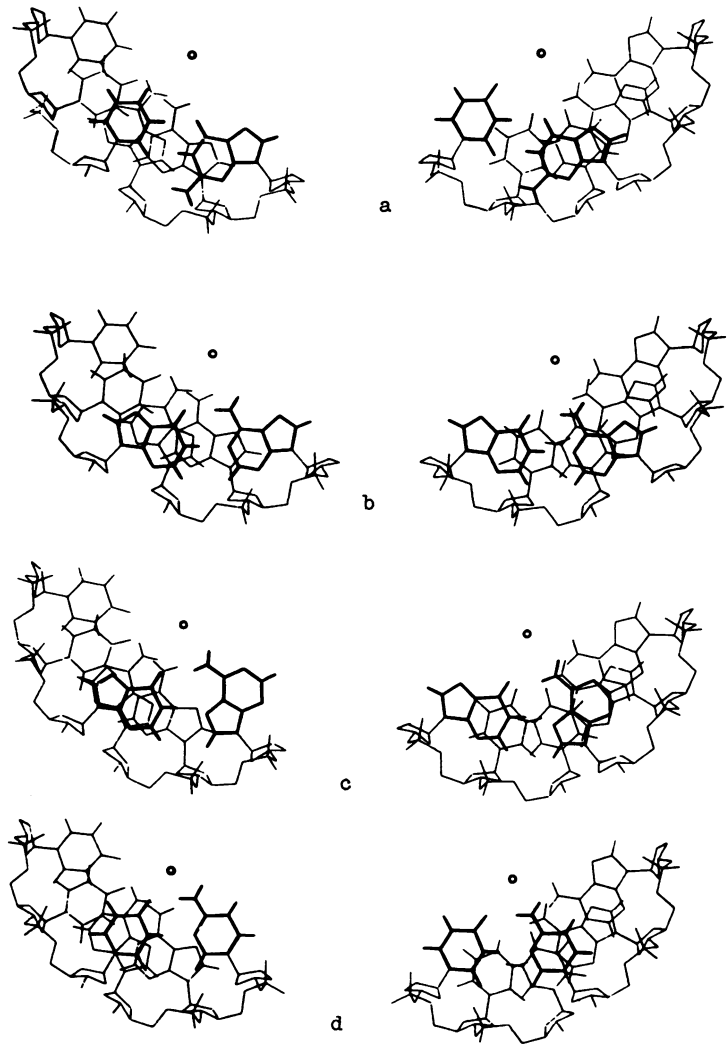

Fig. 2. Fragments of double-helical polynucleotides containing mispairs. A-form. Projection on the plane perpendicular to the helix axis. C-H bonds in the sugar-phosphate backbone are not shown. Bold lines designate the mispairs: a, G:U; b, I:A; c, I:A*; d, C:U. For each pair fragment dApdApdX:dYpdUpdU is on the left, fragment dXpdApdA: dUpdUpdY is on the right.

energy deviation from the minimum values. A comparison shows that all conformations obtained have the angles of the sugar-phosphate backbone within the limits or slightly deviating from the angles characteristic of crystals

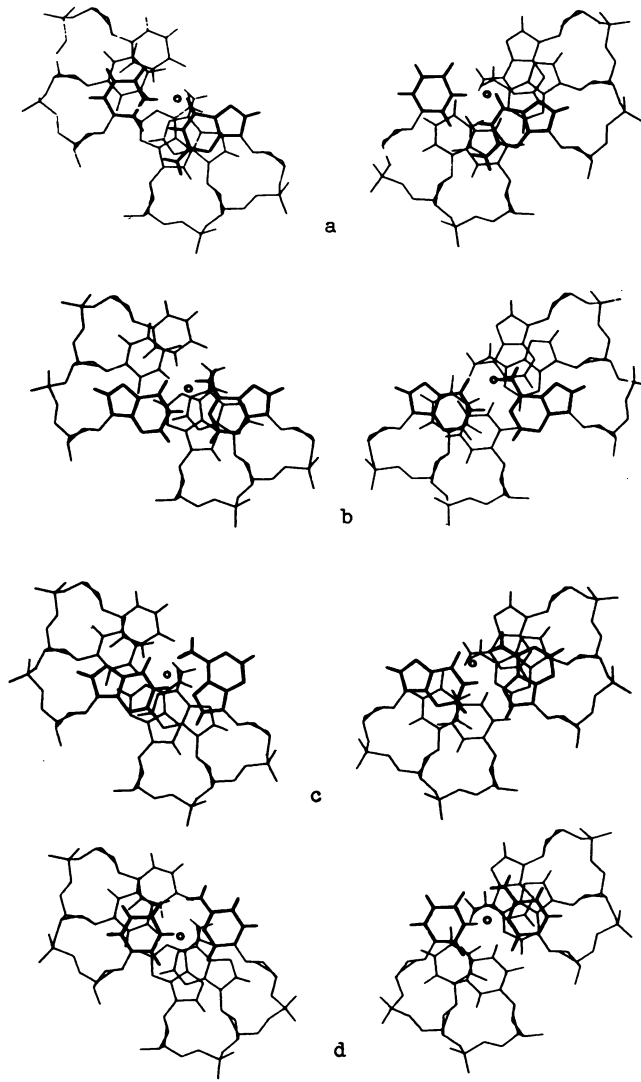

Fig. 3. Fragments of double-helical polynucleotides containing mispairs. B-form. Designations are the same as in Fig. 2.

and calculated theoretically for polynucleotides consisting of Watson-Crick pairs.

Reduced interatomic contacts

Analysis of the short interatomic distances for conformations whose characteristics are given in Tables 1-4 shows that deviation of the double

Table 5. Ranges for dihedral angles of the sugar-phosphate backbone.

ω $C_4'-C_3'-O_3'-P$	φ $C_3'-O_3'-P-O_5'$	ψ $O_3'-P-O_5'-C_5'$	θ $P-O_5'-C_5'-C_4'$	ξ $O_5'-C_5'-C_4'-C_3'$	
		B-form			
168 ÷ 222	220 ÷ 273	-62 ÷ -80	150 ÷ 201	51 ÷ 68	This paper
171 ÷ 260	150 ÷ 274	-51 ÷ -82	139 ÷ 190	40 ÷ 66	(3)
170 ÷ 240	200 ÷ 260	-35 ÷ -85	135 ÷ 200	25 ÷ 75	(7)
		A-form			
198 ÷ 242	-56 ÷ -111	-64 ÷ -79	133 ÷ 193	53 ÷ 78	This paper
191 ÷ 233	-61 ÷ -83	-65 ÷ -99	164 ÷ 186	50 ÷ 75	(16)
162 ÷ 271	-27 ÷ -125	-30 ÷ -95	119 ÷ 200	41 ÷ 56	(17)
175 ÷ 240	-40 ÷ -100	-40 ÷ -110	140 ÷ 215	30 ÷ 80	(7)

helix from regularity produced by incorporation of a mispair leads to the appearance of a rather small number of new reduced contacts.

Regular double helices consisting of A:U planar pairs obtained from our calculations have interatomic contacts which can be considered as reduced. For the A-family of conformations these are contacts within nucleoside O_1' with C_6 of pyrimidine (down to 2.64 Å) and C_8 of purine (down to 2.82 Å), C_2' with C_2 of pyrimidine (down to 3.16 Å), H_2' with N_1 of pyrimidine and N_9 of purine (down to 2.55 Å) and H_1' with C_2 of pyrimidine (down to 2.59 Å) as well as contacts $P...H_3'$ (down to 2.75 Å). For conformations found without the assumption that the pair is planar, i.e. introducing an additional degree of freedom corresponding to a twist of bases in the pair these distances increase. Incorporation of the G:U pair practically does not lead to the appearance of the new and reduction of the existing contacts in regular polynucleotides. At incorporation of other mispairs these contacts may become sometimes reduced (only by 0.01÷0.03 Å) and a few new contacts may appear. At incorporation of U:C and I:A pairs $P...H_5'$ reduced contacts (down to 2.62 Å) appear in the sugar-phosphate backbone. Besides, incorporation of the I:A pair results in contacts $O_1'...O_5'$ (2.64 Å), $O_1'...C_3'$ (2.89 Å) and $O_1'...H_3'$ (2.41 Å). Incorporation of the U:C pair leads to contacts of H_2' atoms of sugar rings of the incorrect nucleotide pair with C_6 of uracil (2.54 Å) and with C_8 of adenine (2.55 Å) of neighbouring pairs. At incorporation of the I:A* pairs contacts within the nucleoside in the syn-conformation are formed: $N_3...O_1'$ (2.61 Å), $N_3...O_5'$ (2.65 Å) and $C_4...O_1'$ (2.76 Å). These contacts may be related due to the fact

that the values of the χ angle for the obtained conformations do not correspond to the optimal mutual position of adenine and sugar in the syn region (just as for the B-family conformations obtained by us). Incorporation of a more favourable syn-nucleotide conformation into a double-helical polynucleotide requires a shift of mispair bases from coplanarity. Incorporation of mispairs into polyribonucleotides was not performed; however, replacement of deoxyribose by ribose in the obtained conformations of the A-family gives only the following contacts within the sugar-phosphate backbone: $C'_4 \ldots O'_2$ (down to 2.8 Å) and $O'_2 \ldots H'_5$ (down to 2.24 Å).

For the B-family of conformations of regular polydeoxynucleotides consisting of planar A:U pairs there are reduced contacts of uracil atoms with deoxyribose atoms $C'_2 \ldots C_6$ (down to 3.03 Å), $H'_1 \ldots C_2$ (down to 2.47 Å) and $H'_1 \ldots O_2$ (down to 2.24 Å) as well as contacts $C'_5 \ldots H'_3$ (down to 2.59 Å), $P \ldots C'_2$ (down to 3.24 Å) and that of phosphate oxygen with C'_3 (down to 2.81 Å). Just as for the A-family, these distances are not smaller than in the models constructed from X-ray data (18) and increase with the addition of a degree of freedom corresponding to the twist. Incorporation of the G:U pair almost does not reduce these contacts and only contacts $P \ldots H'_5$ (down to 2.76 Å) and $P \ldots H'_3$ (2.75 Å) are newly formed. For conformations of fragments containing I:A, I:A* and C:U pairs the $P \ldots H'_2$ contact (down to 2.68 Å) is formed, contacts $P \ldots C'_2$ can be reduced to 3.17 Å and all other contacts observed for poly-dA:poly-dU can be reduced by no more than 0.04 Å. Besides, for conformations of the fragment with the I:A pair contacts $P \ldots H'_3$ (down to 2.72 Å) and $P \ldots H'_5$ (down to 2.71 Å) are formed in the sugar-phosphate backbone as well as $H'_2 \ldots C_6$ of uracil (2.59 Å) and the N_3 atom of purines of nucleotide mispair with O'_1 of neighbouring pairs (dowm to 2.75 Å). Conformations of fragment with a C:U pair have contacts $P \ldots H'_3$ (down to 2.68 Å), $P \ldots H'_5$ (down to 2.68 Å), $H'_1 \ldots O_2$ of cytosine (2.2 Å), $H'_1 \ldots C_2$ of cytosine (2.5 Å) and O'_1 of sugars of the nucleotide mispair with N_3 of adenine (2.76 Å) and O_2 of uracyl of neighbouring pairs (2.67 Å). At incorporation of the I:A* pair contacts of N_3 of adenine of the mispair with H'_2 of two sugar rings (2.41 Å and 2.5 Å) and that of C_2 of this adenine with H'_2 of the neighbouring nucleotide (2.5 Å) are formed. Minimization of energy for polynucleotides containing thymine was not performed, but substitution of uracil for thymine in the conformations obtained shows that reduced contacts of the thymine CH_3-groups with C'_2 (down to 3.04 Å) and H'_2 (down to 2.5 Å) are formed. The amino group of guanine in the obtained conformations of fragments containing I:A and I:A* pairs does not create additional steric constraints. In all

structures calculated the distance between hydrogen atoms is no less than 2 Å.

The contacts considered above characterize conformations of fragments consisting of five nucleotide pairs obtained by adjustment of low-energy conformations of two fragments consisting of three nucleotide pairs. In these low-energy conformations the contacts mentioned above are not so close or are even absent. It should be noted that many reduced interatomic contacts pertain to type 1...4, i.e. contacts of atoms divided by three bonds and their small reduction as compared to contacts of the corresponding type of atoms in crystals is admissible. Besides, a number of reduced contacts are those with hydrogen atoms whose small deviations do not induce a considerable energy increase. An increase of the degrees of freedom and sometimes a search for another conformation within the limits of minimum energy values can yield other conformations which are no less favourable energetically and have more distant interatomic contacts.

Incorporation of mispairs and accuracy of nucleic acid biosynthesis

Incorporation of mispairs in normal tautomeric forms into DNA as a mechanism of spontaneous mutations has been also considered by Rein et al. (19,20). It has been shown (20) that purine-purine (syn) pairs with bases in ordinary tautomeric forms are energetically more favourable than those with a base in a rare form proposed by Topal and Fresco (2). It has been noted, however, that purine nucleosides in the syn-conformation cannot be incorporated into B-DNA because of too short interatomic contacts (down to 0.5 Å). Assuming that the mutual position of bases in DNA is fixed, the authors of paper (19) could adjust the sugar-phosphate backbone only for pairs U:U and G:U, but even then they obtained a too large increase of energy of the sugar-phosphate backbone. As a result, the authors made an erroneous conclusion that the small probability of formation of such pairs in DNA synthesis is completely determined by the interactions of nucleic acid components, i.e. without the participation of enzymes of template synthesis.

The results of our calculations show that incorporation of mispairs is practically possible without the appearance of reduced interatomic contacts and without a considerable distortion of the sugar-phosphate backbone. In other words, the sugar-phospahte backbone possesses a sufficient flexibility to allow incorporation into the double helix of nucleotide pairs having the dimensions (i.e. distances between C_1' atoms) of about 2 Å larger (I:A pair) and 2 Å smaller (U:C pairs) than the canonical Watson-Crick pairs as well as incorporation of pairs in which one base is shifted relative to the other

by about 2 Å along the dyad axis (G:U pair). Almost all distortions of the backbone occur in the part connecting the mispair with the neighbouring correct pairs.

The analysis of calculations shows that G:U pairs are most easily incorporated into the double helix. Their incorporation does not involve the formation of new reduced contacts, requires a small increase of energy of nonbonded interactions and, as seen from Tables 1 and 3, causes a smaller than for other pairs deviation of angles of the sugar-phosphate backbone from their values for double-stranded polynucleotides consisting of A:U pairs. It should be noted that these deviations could be smaller if calculations for the G:U pairs took into account more degrees of freedom as done for the fragments containing other mispairs (see METHODS). X-ray data (16, 17) have shown the existence of a G:U wobble pair in the helical tRNA fragment. IR and Raman spectroscopy were used to demonstrate the formation of such pairs at complexation of poly(A,G) with poly(U) (21) and their formation in oligonucleotide complexes was shown by NMR (22). Recently the formation of double helices by selfcomplementary deoxydodecanucleotide with two wobble G:T pairs was demonstrated by nuclear magnetic resonance and differential scanning calorimetry technique (23). The incorporation of G:T mispair results in changes in base pair overlaps and phosphodiester torsion angle changes at the mismatch site only. This finding is in agreement with our prediction.

Calculations show that the incorporation of U:G, I:A and I:A* pairs changes the energy of nonbonded interactions of the fragment by no more than 3 kcal/mole as compared to the energy of fragments consisting of A:U pairs and having a conformation belonging to the same family. This result as well as those mentioned above indicate that mainly such pairs (and not those with rare tautomers) are most likely formed at biosynthesis of nucleic acids as the probability of rare tautomeric forms is no more than 10^{-4}. The experimentally observed decrease of probability of mispair formation in template synthesis seems to be due to enzymes which increase the accuracy. As mentioned in the INTRODUCTION, according to the concept of enzyme-nucleotide recognition (3-5) this increase is achieved owing to interaction of the enzyme recognizing site with structural invariants of complementary nucleotide pairs. These invariants are the position and conformation of the sugar-phosphate backbone, position of N_3 atoms of purines and O_2 of pyrimidines. According to our calculations, the formation of mispairs is possible when the sugar-phosphate backbone conformation changes no more than it is possible for polynucleotides consisting of Watson-Crick pairs;

therefore, to increase the accuracy of biosynthesis, the enzyme recognizing site should interact with nitrogen bases. Interaction only with the sugar-phosphate backbone is not sufficient since for fragments with G:U, I:A, C:U and I:A* pairs it is impossible to clearly differentiate the backbone conformation and the position of phosphate groups from the corresponding characteristics of polynucleotides consisting of Watson-Crick pairs.

It is known that spontaneous transitions occur more often than spontaneous transversions. This fact correlates well with our conclusion that among all mispairs it is the G:U pair that is the most easily formed in the double helix, thus leading to transitions.

Our calculations do not give enough ground to choose between I:A and I:A* pairs as an intermediate leading to transversions. The energy of nonbonded interactions in fragments containing these pairs is approximately the same, though its components can differ. Incorporation of the I:A* pair leads to small distortions of the sugar-phosphate backbone, but the syn-nucleoside has a conformation 1-2 kcal/mole less favourable than the optimal syn-conformation. Consideration of more degrees of freedom would probably allow one to determine which pair is more likely to be formed.

Conformations of fragments with the C:U pair (3 nucleotide pairs) have the energy of nonbonded interactions of about 8 kcal/mole higher than for the fragments with other pairs considered, which belong to the same family of conformations. This unfavourability is due to a smaller (by the absolute value) energy of interaction in the mispair and between pairs and to the increase of energy of the base-sugar interaction when purine is replaced by pyrimidine. It has been shown recently that frequencies of formation of pyrimidine-pyrimidine pairs in DNA synthesis are considerably lower than those of purine-pyrimidine and purine-purine formation (24). The results of our calculations help to clarify this experimental fact.

Conformations of the sugar-phosphate backbone considerably differing from A- and B-conformations seem to be possible for double-stranded polynucleotides containing mispairs as well as for regular helices with Watson-Crick pairs. In this study we consider only conformations with the sugar-phosphate backbone angles lying in the same region as for A- and B-conformations of nucleic acids. Therefore our results cannot be compared directly with the experimental data on interaction of two tRNAs having mutually complementary or almost complementary anticodons (25). The authors of paper (25) suggest a considerable change of dihedral angles of the sugar-phosphate backbone upon formation of pyrimidine-pyrimidine pairs.

Realization of such conformations in the processes of nucleic acid biosynthesis would probably be considerably hampered by the enzymes of biosynthesis.

ACKNOWLEDGEMENTS

The authors express their gratitude to Drs. V. I. Bruskov, G. G. Malenkov, V. I. Ivanov and V. B. Zhurkin for fruitful discussions.

REFERENCES

1. Watson, J.D. and Crick, F.H.C. (1953) Nature 171, no. 4361, 963-964.
2. Topal, M.D. and Fresco, J.R. (1976) Nature 263, 285-289.
3. Poltev, V.I. and Bruskov, V.I. (1977) Molekul. Biol. 11, 661-670.
4. Poltev, V.I. and Bruskov, V.I. (1978) J. Theor. Biol. 70, 69-83.
5. Bruskov, V.I. and Poltev, V.I. (1974) Dokl. Akad. Nauk SSSR 219, 231-234.
6. Khutorsky, V.E. and Poltev, V.I. (1976) Biofizika 21, 201-206.
7. Zhurkin, V.B., Lysov, Yu.P. and Ivanov, V.I. (1978) Biopolymers 17, 377-412.
8. Chuprina, V.P., Khutorsky, V.E. and Poltev, V.I. (1981) Studia biophysica 85, 81-88.
9. Poltev, V.I., Milova, L.A., Zhorov, B.S. and Govyrin, B.A. (1981) Biopolymers 20, 1-15.
10. Poltev, V.I. (1979) Int. J. Quant. Chem. 16, 863-868.
11. Zhurkin, V.B., Poltev, V.I. and Florentiev, V.L. (1980) Molekul. Biol. 14, 1116-1130.
12. Zefirov, Yu.V. and Zorky, P.M. (1976) Zh. Strukt. Khimii 17, 994-998.
13. Arnott, S. (1970) Progr. Biophys. Mol. Biol. 21, 265-319.
14. Altona, C. and Sundaralingam, M. (1972) J. Amer. Chem. Soc. 94, 8205-8212.
15. Drew, H.R., Wing, R.M., Takano, T., Broka, C., Tanaka, S., Itakura, K. and Dickerson, R.E. (1981) Proc. Natl. Acad. Sci. USA 78, 2179-2183.
16. Hingerty, B., Brown, R.S. and Jack, A. (1978) J. Mol. Biol. 124, 523-534.
17. Quigley, G.J., Seeman, N.C., Wang, A.H.-J., Suddath, F.L. amd Rich, A. (1975) Nucl. Acids Res. 2, 2329-2341.
18. Arnott, S., Smith, P.J. and Chandrasekharan, R. (1975) in Handbook of Biochemistry and Molecular Biology, Eastman, C.D. Ed. 414-422, CRC-Press.
19. Rein, R., Coeckelenbergh, Y. and Egan, J.T. (1975) Int. J. Quant. Chem.: Quant. Biol. Symp., no.2, 145-153.
20. Garduno, R., Rein, R. , Egan, J.T., Coeckelenbergh, Y. and MacElroy, R.D. (1977) Int. J. Quant. Chem.: Quant. Biol. Symp., no.4, 197-204.
21. Ackermann, Th., Gramlich, V., Klump, H., Knable, Th., Schmid, E.D., Seliger, H. and Stulz, J. (1979) Biophys. Chem. 10, 231-238.
22. Romaniuk, P.J., Hughes, D.W., Gregoire, R.J., Bell, R.A. and Neilson, T. (1979) Biochemistry 18, 5109-5116.
23. Patel, D.J., Kozlowski, S.A., Marky, L.A., Rice, J.A., Broka, C., Dallas, J., Itakura, K. and Breslauer, K.J. (1982) Biochemistry 21, 437-444.
24. Fersht, A.R. and Knill-Jones, J.W. (1981) Proc. Natl. Acad. Sci. USA 78, 4251-4255.
25. Grosjean, H.J., Henau, S. and Crothers, D.M. (1978) Proc. Natl. Acad. Sci. USA 75, 610-614.