Chemistry of Natural Compounds and Bioorganic Chemistry

Helical complexes of oligonucleotides with derivatives of oligoamides. Specificity of complementary recognition

I. A. Il'icheva and V. L. Florentiev*

V. A. Engel'gardt Institute of Molecular Biology, Russian Academy of Sciences, 32 ul. Vavilova, 117984 Moscow, Russian Federation.
Fax: +7 (095) 135 1405

The method of molecular mechanics with an AMBER 3.0 computer software package, supplemented by specially developed subprograms, was used to carry out a conformational analysis of hybrid double helices with noncanonical H-bonds composed of oligodeoxyribonucleotide and oligoamide fragments $(dT)_5 \cdot (pT)_5$ (**1**) and $(dA)_4(dT)_2(dA)_4 \cdot (dT)_3(pT)_4(dT)_3$ (**2**). The purpose of the work was to perform a pre-synthesis estimation of Watson—Crick specificity of binding specially constructed oligomers, which carry oligoamide inserts of 2-aminoethylglycine units in the center, connected to nitrogen bases through a methylenecarbonyl group. A comparison of helical parameters and potential energies in optimized structures of hybrid oligonucleotides with noncanonical T·T-pairs with the energies of analogs containing only canonical A·T-pairs showed that disruption of Watson—Crick complementarity results in a crucial distortion of hybrid double helices, which leads to their destabilization. For this reason, the expected probability of mispairing of oligonucleotides is low for the proposed analogs of oligonucleotides carrying oligoamide inserts. Hence, the synthesis of these oligonucleotides is promising for creating reagents with selective action on single-stranded oligonucleotides inside cells.

Key words: DNA, nucleic acid analogs, Watson—Crick complementarity, specific actions, sequence, conformational analysis, molecular mechanics.

The property of single- and double-stranded DNA to form stable hydrogen-bonded complexes with polyamides constructed from 2-aminoethylglycine units, which are connected with nitrogen bases through the methylenecarbonyl group (the so-called peptide nucleic acids or PNA),[1—5] continues to cause great interest to these compounds. The main reason is the possibility of creating reagents for gene-directed actions on the basis of these compounds.[6] The perspectives for their use are substantially limited by the poor permeability of cell membranes to these nucleotides. At the same time, it has been recently shown that active transport of

* Papers presented in the section "Chemistry of Natural Compounds and Bioorganic Chemistry" of this issue are based on the materials delivered at the XII Seminar on Intermolecular Interactions and Conformation of Molecules, Khar'kov, October 3—8, 1994.

Translated from *Izvestiya Akademii Nauk. Seriya Khimicheskaya,* No. 7, pp. 1364—1368, July, 1995.

1066-5285/95/4407-1316$12.50

oligonucleotides with the length of up to 20 base pairs through a membrane is possible.[7] In order to overcome the low permeability of cell membranes to oligoamides, we have proposed using them as inserts with lengths from four amide units in the oligonucleotide chain.[8] Molecular mechanical calculations of double helices of decanucleotides, carrying complementary inserts of four monomeric PNA units in one of the chains, showed that the intramolecular potential energy of such complexes appears to be significantly lower than that of decanucleotides with the same base composition but without oligoamide inserts.[8] Therefore, one should expect these complexes to have increased stability, in particular, increased melting points, as compared with the corresponding oligonucleotides. However, the problem of Watson—Crick specificity of binding of the proposed hybrid oligomers, which carry inserts, with oligonucleotides remains unresolved.

Experimental studies of PNA complexes with oligodeoxyribonucleotides indicate that two types of helices can be formed, namely, triple-stranded helices composed of two PNA chains and one oligodeoxyribonucleotide chain (stereochemistry is 2 : 1),[1—4,9] and double-stranded helices (1 : 1).[5] Triple-strained helices are always formed in cases when only homopurine-homopyrimidine chains take part in the reaction. According to the opinion of the authors in Ref. 5, the CD spectra of helical complexes of oligomers containing all four bases indicate that only double-stranded hybrids are formed. The data about the dependence of the stability of complexes on the sequence of bases suggest that Watson—Crick type hydrogen bonds exist both in triple- and double-helical complexes. Disruption of the Watson—Crick complementarity in the sequence results in a dramatic decrease in the stability of helices, which is expressed by a melting point depression by 20 °C. For the double-helical complexes proposed by us, which carry inserts of several PNA units in one of the nucleotide chains, the dependence of stability on disruption of the Watson—Crick complementarity may appear to be significantly weaker due to the apparently short length of an insert.

In order to estimate the specificity of binding of the oligonucleotide target with the proposed oligomers, in the present work we carried out a conformational analysis of double-stranded complexes with noncanonical hydrogen bonds, namely hybrid helices $(dT)_5 \cdot (pT)_5$ (**1**) and decamers composed of oligonucleotide and oligoamide fragments $(dA)_4(dT)_2(dA)_4 \cdot (dT)_3(pT)_4(dT)_3$ (**2**), which contain an insert of four oligoamide units and two noncanonical T · T-pairs in the center of the insert. Saenger[10] has revealed the possibility of the existence of anti-parallel double helices that are composed of two identical oligonucleotide chains, the thio-derivatives of oligouridines. The pyrimidine pair formed in this case has no elements of symmetry. Earlier we have carried out[11] optimization of the geometry for all possible variants of pairs of canonical bases connected with not less than two hydrogen bonds. The angle between the glycoside bonds, which is typical of the optimized geometry of the non-symmetrical T · T-pair in question, is equal to 56.2°, whereas this angle is 72.1° in the Watson—Crick A · T-pair. (Two other variants of symmetric hydrogen bonding of thymines can theoretically form helices only with a parallel orientation of the sugar-phosphate chains.) Therefore, insertion of a T · T-pair into double-stranded helices of oligo(dA) · oligo(dT) should substantially distort the homogeneity of the structure. In the case of interest, the insertion of several PNA units itself distorts the structure of the helix. Nevertheless, as we have previously shown,[8] such complexes are characterized by a noticeably lower potential energy as compared with that of oligo(dA) · oligo(dT) helices which is associated with the non-polar character of the oligoamide insert. In this work we estimated the value of additional disturbances that appear as a result of formation of noncanonial T · T-pairs in the center of the insert.

Computational methods

Conformational analysis of molecules **1** and **2** was carried out using molecular mechanical methods with an AMBER 3.0 computer software package[12] supplemented by two additional subprograms. One of the programs allows us to define the coordinates of atoms in regular helices from the coordinates of the first monomeric unit and is used to determine the starting point for optimization of the conformational energy. The other program allows us to calculate the helical parameters of the structure obtained from the optimization of the potential energy.

The potential energy optimization was carried out in two stages. The steepest descent algorithm was used in the first step (until a gradient of 0.5 kcal mol^{-1}) and the conjugate gradient algorithm (Pollak—Rybierra) was used in the second step (until a gradient of 0.01 kcal mol^{-1}).

The semiempirical quantum-chemical methods CNDO and AM1 were used for determination of effective atomic charges in oligomeric units, which could not be calculated by the AMBER 3.0 program. The former two methods were used to calculate the charges of both amide and nucleotide units. The values of charges obtained by the CNDO method were chosen for molecular mechanical calculations, since they appeared to be the closest to the charges of nucleotides used in the AMBER programs.

The coordinates of heavy atoms of the B-form of Arnott DNA[13] were used as starting coordinates since the amide and nucleotide monomeric units are isomorphous. Calculations were made for helices with anti-parallel arrangement of the 5′→3′ and N→C vectors, because, as we have shown previously,[11] such arrangement is preferable both for double-stranded hybrid helices composed of oligonucleotide and

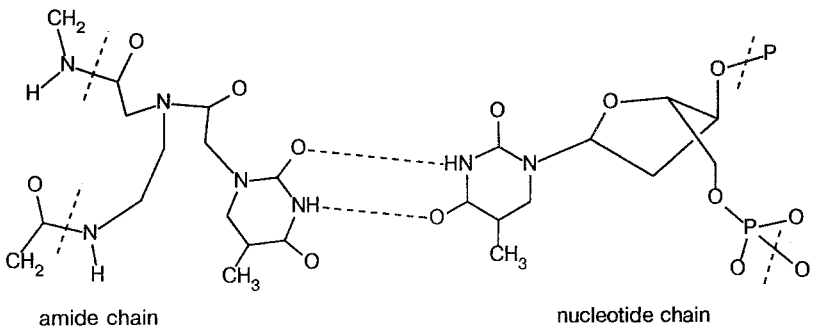

Fig. 1. Monomeric unit of a (dT) · (pT) double hybrid helix.

oligoamide chains, and for oligoamide inserts in the complexes proposed.

The calculations were preceded by obtaining the optimal conformation of a non-symmetrical T · T-pair (Fig. 1).

Results and Discussion

Table 1 summarizes the conformational and helical characteristics, as well as the values of the total potential energy of the hybrid helical duplexes **1**. For presentation of the structures, we chose the most characteristic conformational angles, *i.e.* χ(C(2)—N—C(1′)—C(2′)), the angle which determines the orientation of the sugar cycle with respect to a base, γ(O(5′)—C(5′)—C(4′)—C(3′)), the angle with respect to the internucleotide bond, and α(O(3′)—P—O(5′)—C(5′)), one of the angles of the phosphoesteric bond, as well as the characteristics of pseudorotation of the sugar cycle (phase angle P and amplitude of pseudorotation θ_m). The orientation with respect to the amide bond, the ω angle, is shown for the amide chain. The choice of these angles for characterization of the whole structure is not accidental. Earlier[11] we have shown the existence of two main types of stable conformers of DNA double helices, namely, structures where the values of the α and γ angles lie in the *gosh*$^+$-region and structures with these angles in the *trans*-region. The majority of the structures solved by X-ray structural analysis of oriented fibers, involving A- and B-forms of DNA, RNA, and conformations of chains in triple-stranded helices, is attributed to the first type. The second type of helical conformer is less abundant. Calculations of the potential energy for helices of nucleic acids with anti-parallel as well as parallel sugar-phosphate chain orientation, determine the second lowest minimum of energy in the *trans*-region. This conformation was also determined by means of X-ray crystallography in the pyrimidine chain of double-stranded hybrids of polydeoxyribonucleotides with polyribonucleotides, poly(A) · poly(dT). As follows from the conformational calculations,[11] it can play an important role in helical structures where the chains are connected with hydrogen bonds other than those of the canonical Watson—Crick type, particularly, in parallel DNA. To characterize the amide chain, it is sufficient to specify one of the two possible variants of planar orientation of substituents around the amide bond (*cis* or *trans*) in a particular structure.

As can be clearly seen from the data of Table 1, the stability is significantly lower (the energy in the minimum is higher) for double-helical hybrids where a Watson—Crick A · T-pair is changed to a noncanonical T · T-pair (for convenience, the data for $(dA)_5 \cdot (pT)_5$ and $(dA)_5 \cdot (dT)_5$ calculated with the same parameters of potential functions[8] are also given). Relaxation of the structure of double helices with noncanonical T · T-binding of chains requires significant conformational changes in the sucrose cycle. This is probably the main reason for the experimentally observed Watson—Crick specificity in the bonding of oligoamides with oligonucleotides.[1—6,9] The relative values of energies in the minima of four conformational regions of molecule **1** (tgg=ttt<ctt=cgg) are an interesting illustration of the previously discovered[11] conformational regularity, stabilization of *trans,trans*-orientation in helices with noncanonical binding of chains.

To estimate the probability of mispairing for binding of hybrid oligonucleotides carryinf oligoamide inserts with canonical oligonucleotides, we present the helical parameters at the ends and in the center of the inserts, which characterize the disruptions of homogeneity in the double helix, as well as the energies corresponding to the conformational minimum of decamers (according to the calculations for compounds **2**) (Table 2). For convenience of comparison, the same Table also contains data for decamers that carry PNA inserts in one strand but with fully complementary sequences.[8] The presence of noncanonical T · T-pairs in the oligoamide insert causes significant inhomogeneities in the helix (Fig. 2). The changes in the local helical parameters in neighboring stacking-contacts reach 10° for the helical rotation angle and 1.3 Å for the helical pitch. The helix is significantly extended at the 3′-end and is strongly compressed and twisted in the center of the insert. The intramolecular potential energy of a decamer carrying such an insert is 50—60 kcal mol^{-1} higher than the energies of decamers with fully complementary oligoamide inserts in the oligo(dT) or respective oligo(dA)

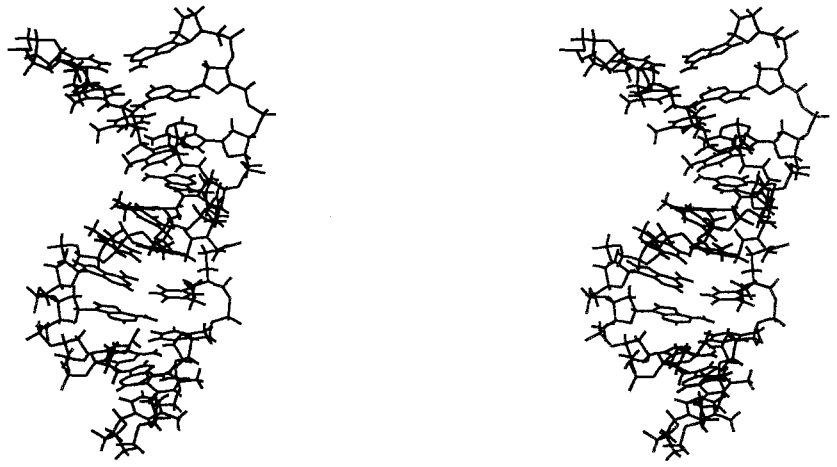

Fig. 2. Conformer corresponding to the minimum intramolecular energy of a $(dA)_4(dT)_2(dA)_4 \cdot (dT)_3(pT)_4(dT)_3$ double helix carrying an amide insert of four units with two noncanonical T · T-pairs in the center. Stereo view.

Table 1. Conformations and potential energies of hybrid helices with noncanonical T · T-pairs*

Oligomer	Type of conformation	Conformational angles/deg						Helical parameters		E_{tot}** kcal mol^{-1}
		Deoxyribonucleotide chain					$(pT)_5$ chain			
		χ	α	γ	P	θ_m	ω	d(A)	τ(°)	
$(dT)_5 \cdot (pT)_5$	tgg	133.1	286.4	63.7	139.0	39.6	178.4	2.96	42.4	−70.0
	cgg	128.9	281.9	62.5	150.0	42.1	3.4	3.2	41.0	−66.4
	ttt	95.4	139.9	176.5	69.0	41.8	172.9	3.1	40.2	−70.0
	ctt	87.7	122.1	174.5	30.0	41.4	345.0	2.9	44.0	−66.4
$(dA)_5 \cdot (pT)_5$	tgg	137.3	283.6	62.6	160.3	39.4	188.4	2.7	39.6	−80.4
	cgg	134.1	283.4	65.6	169.6	39.8	353.5	2.94	46.2	−78.8
	ttt	156.2	102.2	177.8	158.0	45.1	185.6	2.91	32.2	−76.5
	ctt	123.6	160.6	172.5	178.9	37.7	352.9	2.86	48.7	−79.2
$(dA)_5 \cdot (dT)_5$	Amide chain	96.2	287.2	65.7	80.2	41.7		3.16	35.3	−62.5
	Nucleotide chain	131.0	289.8	61.6	163.3	34.2				

* All data are assigned to helices with an anti-parallel orientation of vectors of 5′→3′ of oligonucleotide and N→C of oligoamide chains.
** The energy values refer to one nucleotide pair.

Table 2. Local structure and potential energy of double-helical decanucleotides carrying tetraamide inserts, with noncanonical** and canonical* complementarity

Oligomer type	Local parameters of non-homogeneity of helices						Energy kcal mol^{-1}
	$d_{5'}$ /Å	$\tau_{5'}$ /deg	d_{pna} /Å	τ_{pna} /deg	$d_{3'}$ /Å	$\tau_{3'}$ /deg	
2	3.1	31.4	2.8	41.5	4.1	31.9	−592.9
3	3.3	37.1	3.2	34.1	3.0	34.5	−643.3
4	3.9	33.5	3.1	26.7	3.0	27.2	−655.5

* Oligomer **2** $(dA)_4(dT)_2(dA)_4 \cdot (dT)_3(pT)_4(dT)_3$ containing two noncanonical T · T-pairs in the oligoamide insert.
** Oligomers **3** $(dA)_{10} \cdot (dT)_3(pT)_4(dT)_3$ and **4** $(dT)_{10} \cdot (dA)_3(pA)_4(dA)_3$ without noncanonical pairs.[8]

chains. Moreover, the potential energy of the helix with T · T-pairs in the center of the amide insert (−593 kcal mol^{-1}) appears to be even higher than that of the canonical double helix $(dA)_{10} \cdot (dT)_{10}$ (−625 kcal mol^{-1}).[8] As a consequence, the probability of errors in the recognition of oligonucleotides with the proposed specially constructed oligomers carrying oligoamide inserts may be rather low. According to our calculations, the contribution of the intramolecular enthalpy to the probability of improper T · T-binding can be estimated as $\exp(-32000/RT) = \exp(-53)$.

The authors are grateful to the Russian Foundation for Basic Research (Grant No. 94-04-11836) and the International Science Foundation (Grant No. M6X000) for financial support of this study.

References

1. P. E. Nielsen, M. Egholm, R. H. Berg, and O. Buchardt, *Science*, 1991, **254**, 1497.
2. M. Egholm, O. Buchardt, P. E. Nielsen, and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 1895.
3. M. Egholm, P. E. Nielsen, O. Buchardt, and R. H. Berg, *J. Am. Chem. Soc.*, 1992, **114**, 9677.
4. M. Egholm, C. Behrens, L. Christensen, R. H. Berg, P. E. Nielsen, and O. Buchardt, *J. Chem Soc., Chem. Commun.*, 1993, 800.
5. M. Egholm, O. Buchardt, L. Christensen, G. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, and P. E. Nielsen, *Nature*, 1993, **365**, 566.
6. J. C. Hanvey, N. J. Peffer, J. E. Bisi, S. A. Thomson, R. Cadilla, J. A. Josey, D. J. Ricca, C. F. Hassman, M. A. Bonham, K. G. Au, S. G. Carter, D. A. Bruckenstein, A. L. Boyd, S. A. Noble, and L. E. Babiss, *Science,* 1992, **258,** 1481.
7. S. L. Loke, C. A. Stein, X. H. Zhang, K. Mori, M. Nakanishi, C. Subasinghe, J. S. Cohen, and L. M. Neckers, *Proc. Natl. Acad. Sci. USA,* 1989, **86**, 3474.
8. I. A. Il'icheva, E. V. Terechova, S. Yu. Tsybenko, A. G. Kunitzin, and V. L. Florentiev, *Int. J. Quant. Chem.*, Quantum Biology Symposium, 1994, **21**, 157.
9. P. E. Nielsen, M. Egholm, R. H. Berg, and O. Buchardt, *Nucl. Acids Res.*, 1993, **21**, 197.
10. W. Saenger, *Principles of Nucleic Acid Structure*, Springer-Verlag, New York, Berlin, Heidelberg, Tokyo, 1984.
11. I. A. Il'ichova, Yu. P. Lysov, A. A. Chernyi, A. K. Shchyolkina, B. P. Gottikh, and V. L. Florentiev, *J. Biomol. Struct. & Dynam.*, 1990, **7**, 879.
12. U. C. Singh, P. K. Weiner, J. W. Caldwell, and P. A. Kollman, AMBER (1986) (UCSF), version 3.0, Department of Pharmaceutical Chemistry, University of California: San Francisco.
13. S. Arnott, P. J. C. Smith, and R. Chandrasekaran, in *Handbook of Biochemistry and Molecular Biology*, Eds. G. D. Fasman, Nucleic Acids, 18901 Granwood Parkway, Cleveland, Ohio 44128, 1976, 411.

Received December 26, 1994