ELSEVIER

FEBS Letters 339 (1994) 113–118

FEBS 13647

FEBS LETTERS

The R-form of DNA does exist

Anna K. Shchyolkina, Edward N. Timofeev, Olga F. Borisova, Irina A. Il'icheva, Elvira E. Minyat, Elena B. Khomyakova, Vladimir L. Florentiev*

V.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilova str., 32, 117984 Moscow, Russian Federation

Received 27 December 1993

Abstract

Oligonucleotide 5'-d(CATGCTAACT)-L-d(AGTTAGCATG)-L-d(CATGCTAACT)-3' [L = $pO(CH_2CH_2O)_3p$] is shown to fold back on itself twice forming at pH 7 a sufficiently stable triplex (T_m is about 30°C) with parallel-orientated identical strands (the recombinant or R-form of DNA). Experimental evidence was obtained by studying thermal denaturation, chemical modification and binding of fluorescent probes. The stability of the R-triplex increases in the presence of divalent ions or spermidine. Its structure is characterized by a certain heterogeneity that causes the cooperativity of a triplex-to-duplex transition to decrease. On the basis of conformational modeling, the possible types of base tripling in all four triplets are proposed. The experimental data as well as the molecular mechanic calculations indicate that the stabilities of triplets in the R-triplex decrease in the order: G:C-G = A:T-A >> T:A-T > C:G-C.

Key words: DNA triplex; Recombination; Arbitrary sequence

1. Introduction

It was shown earlier that during homologous recombination promoted by RecA and some other recombinase proteins, a triple-stranded DNA intermediate was formed and persisted even after deproteinization [1–8]. It was suggested that in this triplex the third strand was hydrogen bonded and positioned in the major groove of the Watson–Crick duplex [1]. Such a structure is a new form of the triple helix which was designated as the R-form of DNA (recombinant form of DNA) [6,7]. In the case of *E. coli* RecA protein the available data suggest that the protein-decorated putative triple helix is extended and underwound by about 50% compared to B-DNA. After deproteinization, stacking was restored and the triplex collapsed. The collapsed triplex seems to be structurally related to the extended form.

Broadly speaking, the RNA triple helix was first described 35 years ago [9]. Over the past seven years, DNA triple helixes have been intensively investigated and at present their properties are well known. Their specific features include (1) the requirement for a strand containing only purines or pyrimidines, and (2) an antiparallel orientation of the third strand with respect to the identical (or similar) strand of the duplex. We call these triplexes antiparallel or 'classical' triplexes. In sharp contrast to the 'classical' triplexes, the recombinant structures must accommodate all four bases and their third strand must be parallel to the identical strand of the Watson–Crick duplex. We denote this structure as parallel or R-triplex. Until quite recently, the R-form was obtained only by deproteinization of joint molecules generated by recombination proteins. Therefore, investigation of the R-form was limited mainly to theoretical conformational computations which demonstrated that the R-form is stereochemically possible for any arbitrary sequence and positions of all four bases in the third strand are nearly isomorphic [10–12].

Our group [13] simultaneously with Churokov's group [14] earlier proposed an approach which was defined as the 'forced strand orientation'. This approach centres on compounds which consist of oligonucleotide segments coupled by a hydroxyalkyl chain of appropriate length. At a sufficiently low concentration of oligomer, when intramolecular complexes are predominantly formed, the relative orientation of strands is unambiguously determined by the oligomer structure. It is very important that the intramolecular DNA complexes are substantially more stable than those formed from separate strands [15]. This approach was successfully applied to investigate multi-stranded DNA [15–19].

In the present work, to study parallel triplexes of arbitrary nucleotide sequence (R-triplexes), we synthesized the oligomer 5'-d(CATGCTAACT)-$pO(CH_2CH_2O)_3p$-d(AGTTAGCATG)-$pO(CH_2CH_2O)_3p$-d(CA-TGCTAACT)-3' (parARB). This single-stranded oligonucleotide is able to fold back on itself twice to give an intra-

* Corresponding author. Fax: (7) (095) 135 1405.

Abbreviations: parARB, 5'-d(CATGCTAACT)-$pO(CH_2CH_2O)_3p$-d(A-GTTAGCATG)-$pO(CH_2CH_2O)_3p$-d-(CATGCTAACT)-3'.

SSDI 0014-5793(94)00031-P

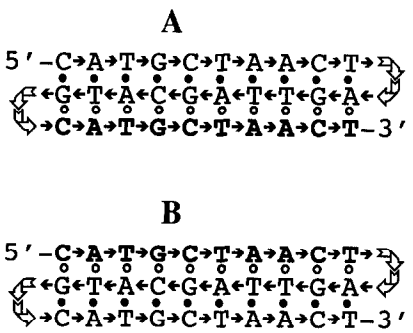

Fig. 1. Topology of two possible configurations of the intramolecular triplex of parARB. The third strand is marked by bold letters.

molecular triple helix (three-stranded clip or simple clip). Such a model structure was chosen for practical reasons: they can be synthesized by a standard method using automatic synthesizer (all segments are coupled by 3′ → 5′ linkages), and they can be sequenced and chemically modified by the routine procedures (since there is only one 5′ end). However, such a model may be folded into two different configurations (A and B in Fig. 1). We have shown convincingly that at low temperature and pH 7 parARB forms a sufficiently stable (T_m from 26 to 34°C depending on conditions) triple helix with parallel-orientated identical strands resembling the R-form of DNA. Moreover, our data indicate that the conformer A predominates in equilibrium.

2. Materials and methods

The triacontamer 5′-d(CATGCTAACT)-pO(CH_2CH_2O)$_3$p-d(AGTTAGCATG)-pO(CH_2CH_2O)$_3$p-d-(CATGCTAACT)-3′ was synthesized by the phosphite triester method on a solid-phase carrier from 3′-phosphoramidites of protected nucleotides and 1-*O*-dimethoxytrityl-10-*O*-(*N*-diisopropylaminomethoxyphosphinyl)-1,8-dihydroxy-3,6-dioxaoctane. Partly deblocked 5′-dimethoxytrityl oligonucleotide was isolated by reverse-phase HPLC. After treatment with acetic acid, the reaction product was purified by ion-exchange HPLC and finally desalted and concentrated by gel filtration.

All solutions were prepared in 0.01 M Na-phosphate buffer, pH 7.0. The oligonucleotide solutions were annealed by heating up to 90°C with subsequent slow cooling. Concentrations of oligonucleotides were determined by UV spectra, assuming the extinction coefficient, ε_{260}, to be equal to 11,300 $M^{-1} \cdot cm^{-1}$ per 1 mol of nucleotides at 90°C.

Melting curves were registered with a Beckman spectrophotometer (USA) at 260 nm in a thermostatic cuvette with the optical path length of 1 cm. The heating rate was constant and equal to 0.2°C/min.

The fluorescent polarization, P, was monitored with an AMINCO SPF 1000 (USA) spectrofluorometer in thermostatic cuvettes at temperatures from 3 to 10°C. The rotation relaxation time, ρ, was calculated by the following equation:

$$\rho = 3\tau(1/P_o - 1/3)/(1/P - 1/P_o), \quad (1)$$

where τ is the time of fluorescent quenching (τ = 25 ns at 3°C), P_o is the limit value of P at $T/\eta \to 0$ (P_o = 42%), η is the viscosity of the solution, and T is absolute temperature [16].

Sequencing was carried out by the routine procedure [20]. Chemical modification and enzymatic digestion were performed at 0–3°C by the methods previously described [21]. The autoradiograms were scanned down the lanes with a 300A Computing Densitometer (Molecular Dynamics, USA).

Conformational energy minimization was performed with the AMBER 3.0 program [22]. No water molecules were explicitly taken into account, but their effect was simulated by use of a distance-dependent dielectric constant.

3. Results

The melting curves of parARB were registered in 0.01 M Na-phosphate buffer, pH 7.0, at NaCl concentrations of 0.25 and 1.0 M, as well as in the presence of either divalent ions (0.01 M Mg^{2+} or Mn^{2+}) or 0.1 M spermidine. The obtained data are partly presented in Fig. 2. All melting curves possess an obvious biphasic character. The higher temperature transition is of a cooperative type, as would have been expected for thermal denaturation of duplexes. The lower temperature transition is low cooperative. Moreover, in the low temperature region the melting curve itself (except melting in the presence of spermidine) also demonstrates the weak biphasic character (see inset in Fig. 2B). A two-stage model is not applicable for these transitions and the Van 't Hoff thermodynamic parameters cannot be calculated. The effective melting temperatures, T_m, (Table 1) may be used as a measure of the stability of the low temperature complexes. Mono- and divalent ions increase the stability of these complexes. Although the sodium ions do not influence the cooperativity of the low temperature transition, divalent ions and especially spermidine make this transition more cooperative. Using the procedure described earlier [15], the apparent thermodynamic parameters of

Table 1

Sample	Apparent melting temperature (°C)*	
	Triplex-to-duplex transition	Duplex-to-open strand transition
Without salts or spermidine	24.5	61.8
With 0.25 M NaCl	26.9	77.3
With 1.0 M NaCl	30.7	82.8
With 0.01 M $MgCl_2$	30.8	67.9
With 0.01 M $MnCl_2$	30.7	67.5
With 0.1 M spermidine	36.6	69.3

*Error did not exceed ± 0.5°C

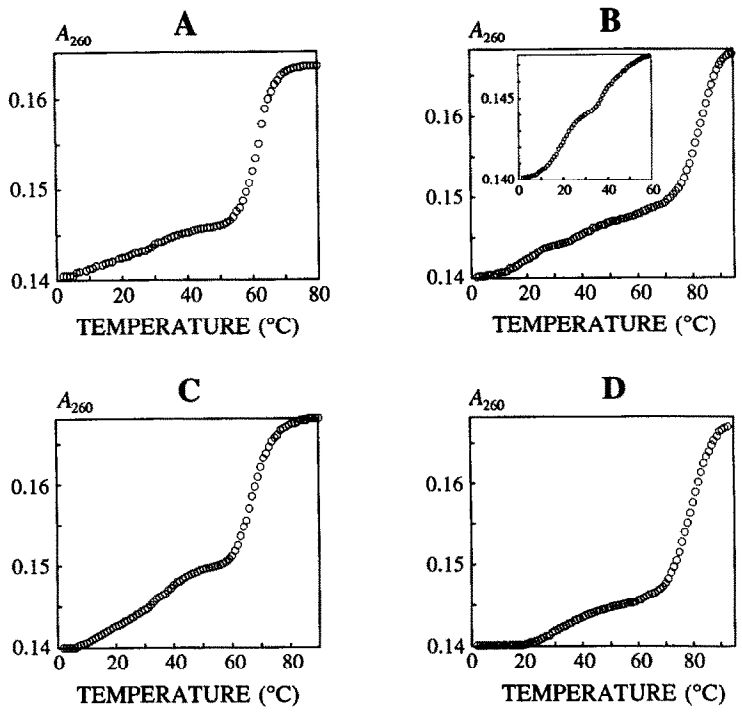

Fig. 2. The melting curves of parARB in 0.01 M Na-phosphate buffer, pH 7, (A) and in the presence of 1 M NaCl (B), 0.01 M $MnCl_2$ (C) or 0.1 M spermidine (D). Concentrations of oligonucleotide were $1–1.7 \cdot 10^{-5}$ M of nucleotides.

the structural transitions were calculated for the melting curve in the presence of spermidine. For the low temperature transition $\Delta H^\circ = 32.2 \pm 4$ kcal/mol, $\Delta S^\circ = 104.6 \pm 12$ cal $\cdot$ mol$^{-1} \cdot$ K^{-1}, for the higher temperature transition $\Delta H^\circ = 49.9 \pm 2$ kcal/mol, $\Delta S^\circ = 141.6 \pm 6$ cal $\cdot$ mol$^{-1} \cdot$ K^{-1}. The parameters for the low temperature transition indicate its lowest limit.

The modification reaction of thymines by OsO_4 and enzymatic digestion with nuclease S1 were carried out at 0–3°C. These data together with the sequencing results according to Maxam–Gilbert are presented in Fig. 3. The data indicate that all three oligonucleotide segments are protected from chemical modification and enzymatic digestion. The 3′-terminal oligonucleotide segment is less protected compared to the two other units. Modification of the terminal nucleotides in all segments is more pronounced. The results of enzymatic digestion of the 3′-terminal oligonucleotide segment are of extreme interest (Fig. 3D). The pyrimidines are found to be more easily excised than purines, and the cytidine is excised more readily than the thymine.

Earlier we used the rotational relaxation time (Eqn. 1)

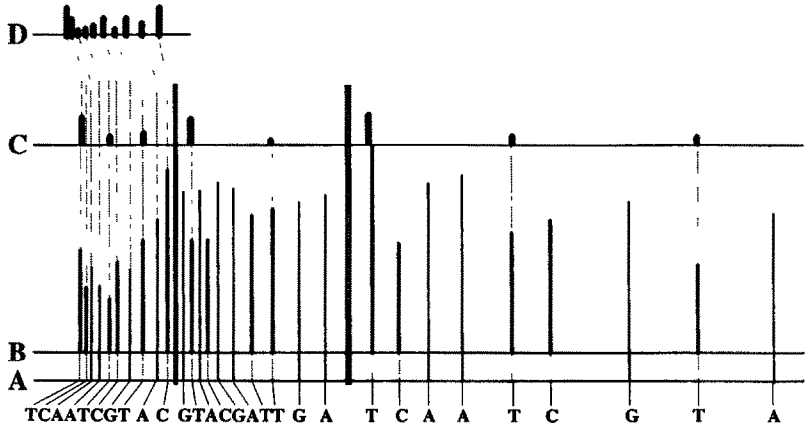

Fig. 3. Densitograms of A+G (A) and C+T (B) squencings, $OsO_4 \cdot$ py modification (C) and S1 nucleas degistion (D) (only the pattern for 3′-segment is shown; two other segments are degested to a lesser extent). The shaded bars indicate the boundaries of oligonucleotide segments. The peaks of densitogram are presented as bars. The height of a bar is equal to the integrated peack area.

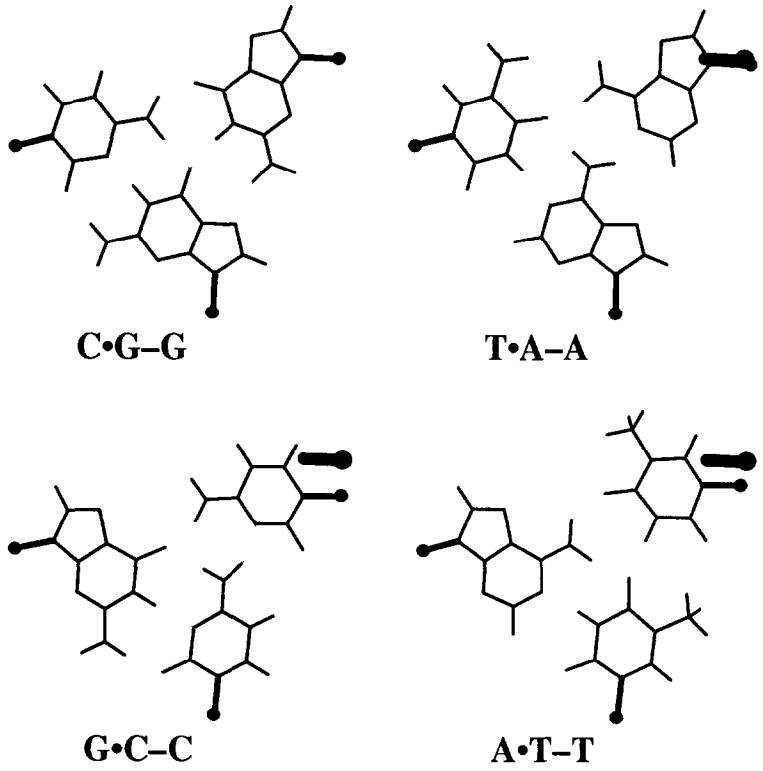

Fig. 4. Scheme of hydrogen bonding in four possible triplexes. Thick lines indicate the location of glycosidic bonds; the very thick line indicates the position of glycosidic bond of third G of C:G-G triplets.

for determining the strandedness of multistranded short oligonucleotides [16]. Since for the triplexes studied the ratio of the helix height to its diameter is less than 1.5, the sphere approximation may be used. In this case the hydrodynamic volume of the molecule is

$$V = kT\rho/3\eta, \quad (2)$$

where k is the Boltzmann constant. Relative hydrodynamic volume, V, (or relative mass M) may be calculated as the ratio of rotational relaxation times of the compounds under study and a reference compound:

$$\rho/\rho_o = V/V_o \cong M/M_o, \quad (3)$$

where ρ_o, V_o and M_o belong to the d(CCGGATCCGG)$_2$ duplex chosen as a reference compound. At 3°C the ratio for parARB appeared to be 1.45 ± 0.01.

4. Discussion

4.1. Evidence for the existence of the triplex with parallel-orientated identical strands

The biphasic character of the melting curves testifies to the formation of a multistranded structure at low temperatures. The measured relative mass of such a complex unambiguously indicates that this structure is an intramolecular triplex. At low concentrations, ethidium bromide intercalates only between the hydrogen-bonded duplex or triplex layers. In the case of a compound consisting of oligonucleotide units bound by flexible hinges, the method reveals the hydrodynamic volume or the mass of the paired or tripled units of this compound (for example, the relative mass of a hairpin with dangling segment is close to 1). It can be concluded therefore that in parARB the three oligonucleotide segments are rigidly bound to each other. Additional evidence for the triplex formation is provided independently by the chemical modification and enzymatic degradation of the low temperature complex. Due to the parARB primary structure, such a triplex must necessarily contain parallel-orientated identical strands (R-triplex). Thus the obtained data indicate conclusively that a triplex has been formed with identical strands orientated in parallel, similar to the R-form of DNA.

4.2. Some features of the triplex with parallel-orientation of identical strands

The R-triplex of parARB is less stable than 'classical' triplexes. Even in the presence of spermidine, the enthalpy change due to the binding of the third strand to the duplex is reduced almost by half as compared to the enthalpy of the Watson–Crick duplex formation. According to our previous observation [15], intermolecular triplexes melt at about 30°C lower than corresponding

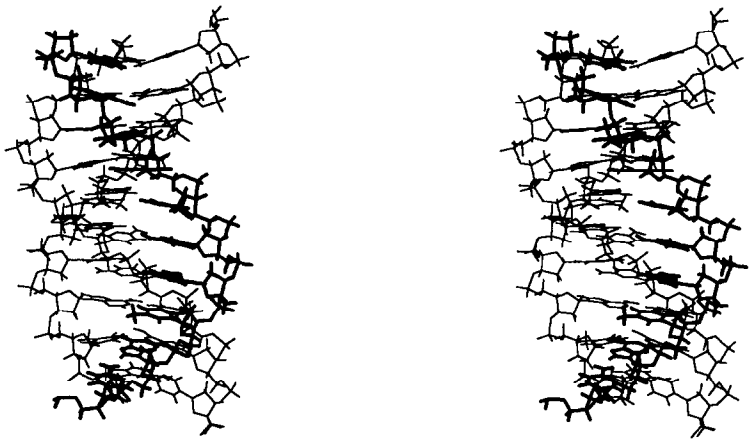

Fig. 5. Stereoview of the optimal structure of the three-stranded helix of parARB. The third strand is marked by thick sticks.

intramolecular ones, therefore it should be expected that the T_m of parallel triplexes formed from separate strands will not be far from 0°C. It is no wonder that attempts to reconstruct the R-triplex from separate strands has ended in failure [1]. The surprisingly high stability of the deproteinization product of joint molecules generated by recombinase proteins [1] can be explained by an extremely slow denaturation of the R-triplex.

The melting curves indicate that the thermal denaturation of the R-triplex proceeds via a number of intermediate stages. The data on enzymatic digestion show that first the hydrogen bonds break in the dipyrimidine–purine triplets, and then these bonds break in the dipurine–pyrimidine triplets. This accounts for the biphasic mode of the melting curves in the region of the triplex melting. The conformational computation correlates well with the experimental data. Conformational analysis shows that the formation of the R-form is stereochemically possible for the sequence investigated. By optimizing the geometry the schemes of hydrogen bonding in the four triplets are proposed (Fig. 4). A significant feature of the proposed triplets is that the third base is bound not with one individual base but with both bases of the Watson–Crick pair. The triplets appeared to be relatively isomorphous. The experimentally observed stability of the triplets: G:C-G = A:T-A>>T:A-T>C:G-C correlates with the extent of difference between a triplet and the most stable G:C-G triplet. It is of importance that our calculations show that the absence of the N7...HN2 hydrogen bond in G:C-G should not influence dramatically the triplex stability. It can explain why substitution of the N7 atom of guanine scarcely effects the R-triplex formation [23].

The oligonucleotide studied can form two isomeric R-triplexes (Fig. 1). The available data on the chemical and enzymatic degradation suggest that the A configuration predominates in the equilibrium mixture.

The calculated conformation of the R-triplex of parARB is shown in Fig. 5.

Acknowledgements: We thank Dr. V.B. Zhurkin for interesting and useful discussions. This work was partially supported by Russian Foundation for Fundamental Investigations and National Program 'Novel Methods of Biotechnology'.

References

[1] Hsieh, P., Camerini-Otero, C.S. and Camerini-Otero, R.D. (1990) Genes Dev. 4, 1951–1963.
[2] Umlauf, S.W., Cox, M.M. and Inman, R.B. (1990) J. Biol. Chem. 265, 16898–16912.
[3] Kiyama, R. and Camerini-Otero, R.D. (1991) Proc. Nat. Acad. Sci. USA 88, 450–454.
[4] Rao, B.J., Iwang, B. and Dutreix, M. (1991) Biochimie 73, 363–370.
[5] Rao, B.J., Chiu, S.K. and Radding, C.M. (1993) J. Mol. Biol. 229, 328–343.
[6] Kim, M.G., Hsieh, P., Camerini-Otero, C.S., Raghunathan, G., Jernigan, R.L., Zhurkin, V.B. and Camerini-Otero, R.D. (1992) Biophys. J. 6, A481.
[7] Kim, M.G., Camerini-Otero, C.S., Raghunathan, G., Jernigan, R.L., Zhurkin, V.B. and Camerini-Otero, R.D. (1992) FASEB J. 6, A481.
[8] Kim, M.K., Zhurkin, V.B., Jernigan, R.L. and Camerini-Otero, R.D. (1993) Biophys. J. 64, A237.
[9] Felsenfeld, G., Davies, D.R. and Rich, A. (1957) J. Am. Chem. Soc. 79, 2023–2024.
[10] Zhurkin, V.B., Raghunathan, G., Camerini-Otero, R.D. and Jernigan, R.L. (1990) J. Biomol. Struct. Dynam. 8, A268.
[11] Zhurkin, V.B., Raghunathan, G., Camerini-Otero, R.D. and Jernigan, R.L. (1992) Biophys. J. 61, A482.
[12] Zhurkin, V.B., Raghunathan, G., Camerini-Otero, R.D. and Jernigan, R.L. (1992) FASEB J. 6, A482.
[13] Shchyolkina, A.K., Lysov, Yu.P., Il'icheva, I.A., Chernyi, A.A., Golova, Yu.B., Chernov, B.K., Gottikh, B.P. and Florentiev, V.L. (1989) Dokl. Acad. Sci. USSR 304, 476–480.
[14] Churikov, N.A., Chernov, B.K., Golova, Yu.B. and Nechipurenko, Yu.D. (1988) Dokl. Acad. Sci. USSR 303, 1254–1258.
[15] Shchyolkina, A.K., Mamaeva, O.K., Borisova, O.K., Lysov, Yu.P., Timofeev, E.N., Il'icheva, I.A., Gottikh, B.P. and Florentiev, V.L. (1994) Antisence. Res. Dev., in press.
[16] Borisova, O.F., Golova, Yu.B., Gottikh, B.P., Zibrov, A.S., Il'icheva, I.A., Lysov, Yu.P., Mamayeva, O.K., Chernov, B.K., Chernyi, A.A., Shchyolkina, A.K. and Florentiev, V.L. (1991) J. Biomol. Struct. Dynam. 8, 1187–1210.

[17] Shchyolkina, A.K., Mamayeva, O.K., Borisova, O.F., Lysov, Yu.P., Timofeev, E.N., Il'icheva, I.A., Gottikh, B.P. and Florentiev, V.L. (1992) Mol. Biol. (Russ.) 26, 1314–1326.
[18] Borisova, O.F., Shchyolkina, A.K., Timofeev, E.N., and Florentiev, V.L. (1992) FEBS Lett. 306, 140–142.
[19] Durand, M., Peloille, S., Thuong, N.T. and Maurizot, J.C. (1992) Biochemistry 31, 9197–9204.
[20] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–559.
[21] Gogoshidze, M.Z., Minyat, E.E., Gorin, A.A., Demchuk, E.Ya., Farutin, V.A. and Ivanov, V.I. (1992) Mol. Biol. (Russ.). 26, 1263–1273.
[22] Singh, U.C., Weiner, P.K., Caldwell, J.W. and Kollman, P.A. (1986) AMBER (UCSF), version 3.0, Department of Pharmaceutical Chemistry, University of California: San Francisco.
[23] Jain, S.K., Inman, R.B. and Cox, M.M. (1992) J. Biol. Chem. 267, 4215–4222.