The Complexation of the Anticancer Drug ThioTEPA with Methylated DNA Base Guanine: Combined Ab Initio and QTAIM Investigation

ARTICLE *in* MOLECULAR INFORMATICS · FEBRUARY 2014

Impact Factor: 1.65 · DOI: 10.1002/minf.201300059

3 AUTHORS, INCLUDING:

Tymofii Nikolaienko

National Taras Shevchenko University of Kyiv

32 PUBLICATIONS **101** CITATIONS

SEE PROFILE

Leonid Bulavin

National Taras Shevchenko University of Kyiv

319 PUBLICATIONS **824** CITATIONS

SEE PROFILE

Full Paper

www.molinf.com

molecular informatics

DOI: 10.1002/minf.201300059

The Complexation of the Anticancer Drug ThioTEPA with Methylated DNA Base Guanine: Combined Ab Initio and QTAIM Investigation

Tymofii Yu. Nikolaienko,*[a] Leonid A. Bulavin,[a] and Leonid F. Sukhodub[b]

Abstract: Non-covalent complexes of methylated nitrogenous DNA base guanine (m^9Gua) with 1 to 6 molecules of anticancer drug ThioTEPA (1,1′,1″-phosphorothioyltriaziridine) have been investigated by molecular modeling techniques (molecular docking and DFT geometry optimization), ab initio wavefunction calculations and the quantum theory of atoms in molecules (QTAIM). The accuracy of complex structures predicted by standard molecular docking techniques have been assessed by comparing them with ab initio calculations, and the most important differences have been discussed. Obtained stabilization enthalpies (kcal/mol) for the m^9Gua···(ThioTEPA)$_n$ complexes with $n=1\ldots6$ have been found to be −15.6, −26.5, −38.4, −49.6, −60.5 and −69.3 respectively. The non-covalent interactions revealed by the QTAIM method have been shown to be a dominating factor responsible for the complex stability, with hydrogen bonds of NH···N type being the most important interactions in small ($n=1$ to 4) and CH···N bonds – in large ($n=5$, 6) complexes. The obtained results may help to understand ThioTEPA-DNA interactions and clarify the mechanism of the drug action.

Keywords: Anticancer drug design · DNA · Hydrogen bonds · Medicinal chemistry · Ab initio calculations

1 Introduction

The cornerstone problem of drug development is how to maximize their efficiency and to minimize side effects. On a microscopic scale both these requirements imply that we want to find a drug molecule with the highest possible selectivity in binding the target.[1,2] Hence, much effort is made towards understanding the basic physical mechanisms governing the drug–target interaction, with particular attention paid to non-covalent interactions.[2]

ThioTEPA (the IUPAC name: 1,1′,1″-phosphorothioyltriaziridine) is the anti-cancer drug invented over 60 years ago[3] which still remains important for practical application in medicine[4–8] and physiology.[9] Although this pharmaceutical compound is widely used, not only the physical properties of its molecule have not been studied carefully until recently,[10,11] but even the mechanism of its action remains vague enough.[5,8,12–15] The researchers assume that the drug acts basically by cross-linking the cancer cell DNA double helix via chemical bond between two strands, thus preventing the DNA replication process.[5,14] To stimulate further progress towards understanding the mechanisms of anticancer drug action it is important to elucidate the principal physics of interaction between the DNA or its constituents and ThioTEPA.

Experimental studies[16,17] using plasma-desorption mass-spectrometry technique with ionization by ^{252}Cf fragments revealed that, even in vacuum, ThioTEPA can form extremely stable complexes with deoxyguanosine monophosphate (dGMP), one of canonical DNA monomer units, with pronounced selectivity of binding. It was found[17] that each molecule of dGMP can bind as many as six ThioTEPA molecules, while other deoxyribonucleotides, which differ from dGMP only by the type of nitrogenous base, typically bind only one molecule of ThioTEPA. This binding specificity can be used in chemotherapy, so it is important to elucidate the mechanism of this phenomenon. Since dGMP has no evident structural preferences over other deoxyribonucleotides except for the presence of guanine as its nucleotide base, it is reasonable to assume that it is the guanine base that determines high binding affinity of dGMP with respect to ThioTEPA. This makes it important to investigate ThioTEPA binding to isolated guanine. In order to block artifacts, which may arise if a strong proton donor is present at glycosidic nitrogen atom (–NH group) of isolated guanine,

[a] *T. Y. Nikolaienko, L. A. Bulavin*
Faculty of Physics, Taras Shevchenko National University of Kyiv 4 Hlushkova ave., 01601 Kyiv, Ukraine
phone/fax: +380974524557
**e-mail: tim_mail@ukr.net*

[b] *L. F. Sukhodub*
Medical Institute, Sumy State University, Ministry of Education and Science of Ukraine
2 R. Korsakova Str., 40007 Sumy, Ukraine

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/minf.201300059.

methylated guanine (m^9Gua) seems to be a potentially better model compound.

The aim of the present paper is to reveal the basic physical factors responsible for high affinity of binding of ThioTEPA with guanine-containing DNA units. Particular attention will be paid to properties of non-covalent interactions including non-covalent bonding and cooperativity effects. Such information is necessary to advance investigations of the interactions between ThioTEPA and DNA structural units and to obtain new knowledge about their binding abilities in terms of physics, which is necessary for further rational improvement of the drug pharmaceutical properties.

2 Computation Details

The investigation of m^9Gua-ThioTEPA complexation has been performed in four steps: (i) the preliminary search for the most stable spatial configuration of the complex, (ii) the clarification of the configuration, (iii) the evaluation of stabilization energy, and (iv) the investigation of non-covalent interactions in the complexes.

Although it is well known that in gas phase (i.e., under experimental working conditions[16,17]) several tautomers of methylated guanine coexist,[18] only the 'canonical' one (i.e., the 9N-keto form) has been studied in details throughout the investigation. The reason for this is that the 7N-enol form is not able to form stable complexes even with one molecule of ThioTEPA since such a complex would have positive stabilization Gibbs energy. In the Supporting Information to the paper we present the spatial structure and stabilization energies for ThioTEPA···7N-enol-m^9Gua complex obtained by quantum-mechanical geometry optimization (Figure SF2 and Table S2) and show that the reason for its instability is the weakness of hydrogen bonds in it (see Table S3), which, in turn, is due to an 'inappropriate' spatial arrangement of proton donors and acceptors in 7N-enol tautomer. The 7H-forms of guanine were excluded a priori since they could not exist under experimental working conditions,[16,17] where no hydrogen was present at the 9-th position of guanine base.

The basic technique used to obtain the initial structures of m^9Gua-ThioTEPA complexes at step (i) was molecular docking performed with the AutoDock 4.2/AutoGrid 4.2.3 software package.[19] The docking was performed with standard Lamarckian genetic algorithm.[20] At first, m^9Gua was treated as a 'macromolecule', and a single ThioTEPA molecule was considered as its flexible ligand. For this stage, the precise molecular structures of m^9Gua and ThioTEPA molecules were obtained by the quantum-chemical geometry optimization performed at the DFT B3LYP/6-31G(d,p) level of theory[21] with *tight* convergence criteria *ultrafine* integration grid and by using Gaussian 03 package.[22]

Next, the most stable configuration of the m^9Gua·ThioTEPA complex obtained in this way was used as a new macromolecule for the new docking procedure, in which the complexation of m^9Gua···ThioTEPA 'macromolecule' with additional molecule of ThioTEPA was studied. The most stable complex of (m^9Gua···ThioTEPA)+ThioTEPA found by the same docking algorithm as before was subjected to geometry optimization at the DFT B3LYP/6-31G(d,p) level of theory with *tight* convergence criteria and *ultrafine* integration grid in order to account for possible reorganization of the 'macromolecule' caused by adding new ThioTEPA molecule (see Section 3 'Results and Discussion' for more details). Such a quantum chemical geometry optimization was the essence of step (ii).

Although modern ab initio methodologies are able to account for the solvation effects (at least, by the continuum solvation models), in the present paper we consider the complexation in vacuum only. The reasons for it are the following. First, this corresponds to the experimental working conditions,[16,17] interpretation of which is our goal. Second, it is hardly possible to understand the role of the environment unless one knows the properties of an isolated system. Finally, no univocal model for the dielectric properties of the enzyme active site exist in literature.[23] Indeed, although a water-like surrounding of the DNA is usually assumed, it has been suggested[24] that the water molecules have likely been removed from the DNA polymeraze active site when the nucleotide enters it. Similar conclusion has been made by Dewar and Storch [25,26] on the basis of kinetic rather than thermodynamic arguments. The validity of vacuum approximation can also be supported by the fact that active sites of many enzymes, which surround the DNA during its functioning, are characterized by rather low dielectric response,[27] which would be unlikely for water. The mentioned reasoning made vacuum approximation widely used in studying physical properties of DNA constituents and their complexes.[28,29]

In the similar way, all the m^9Gua···$(ThioTEPA)_n$ ($n=3$, 4, 5, 6) structures were obtained by taking steps (i) and (ii) iteratively as follows: the optimized structure of m^9Gua···$(ThioTEPA)_{n-1}$ complex was used as a macromolecule in docking a new ThioTEPA molecule. Then, the most stable complex determined by docking was used as the initial structure for geometry optimization at the DFT B3LYP/6-31G(d,p) theory level, and the completely optimized structure of m^9Gua···$(ThioTEPA)_n$ complex obtained in this way was used as the macromolecule for docking the next ThioTEPA molecule. Finally, the same quantum-chemical geometry optimization was performed for m^9Gua···$(ThioTEPA)_{n+1}$ complex. The process flowchart is depicted in Figure SF1 in the Supporting Information.

It should be noted that a priori two mechanisms of complex formation can be assumed, namely the 'step-by-step' (when the next ligand binds the target *after* the previous one has already been bound) and 'one-step' ones (when all ligands approach the m^9Gua molecule at once). Although both cases are worth studying, only the former is considered throughout the work, since it seems to be closer to

experimental working conditions.[16,17] Indeed, assuming the collisions of the ligand with the target are random, an interaction of target with a single ThioTEPA molecule seems to be a much more probable event than its interaction with two or more ligand molecules.

After all the optimized structures of m^{9}Gua···(ThioTEPA)$_n$ complexes had been found, step (iii) was taken in which stabilization energies of the complexes were evaluated. In order to properly account for all the relevant intra- and intermolecular interactions, including dispersion forces, interaction energies were calculated by the 2-nd order Moller–Plesset perturbation theory (MP2)[30,31] using Dunning's cc-pVTZ basis set[32,33] and the RI-JK resolution-of-identity technique[34,35] as implemented in the ORCA quantum-chemical software package.[36] Basis set superposition errors (BSSE) were corrected by the standard counterpoise procedure of Boys and Bernardi.[37] The monomer deformation energies were also taken into account and the final electronic stabilization energy (ΔE) of complexes was calculated as follows

$$\Delta E = E^{\mathrm{DCBS}} - \sum\nolimits_i E_i^{\mathrm{DCBS,cg}} + \sum\nolimits_i \left(E_i^{\mathrm{MCBS,cg}} - E_i^{\mathrm{MCBS,og}}\right) \qquad (1)$$

where E^{DCBS} is the electron energy of the complex, $E_i^{\mathrm{DCBS,cg}}$ is the energy of an isolated i-th monomer with exactly the same geometry as in the complex, and which is calculated using the same basis set as that used to evaluate the complex total energy (i.e., with basis functions on 'ghost' atoms[37] kept), and summation over i goes over all the monomers. The two last terms represent the deformation energy, i.e., the work needed to transform the isolated monomers in their most energetically favorable conformations into the conformations which they acquire in the complex, i.e., $E_i^{\mathrm{MCBS,cg}}$ and $E_i^{\mathrm{MCBS,og}}$ stand for the energies of the i-th monomer evaluated at the 'in-complex' geometry ($E_i^{\mathrm{MCBS,cg}}$) and the optimized 'free' geometry ($E_i^{\mathrm{MCBS,og}}$) respectively (no 'ghost atoms' have been used in the basis set for these two terms).

In order to make our results usable for interpreting the ThioTEPA binding experiments,[16,17] we have also studied the thermodynamic characteristics of the binding. In particular, within the standard rigid rotor-harmonic oscillator thermochemistry model as implemented in Gaussian 03, we have found the internal energy $U=\langle E\rangle$, enthalpy $H=U+RT$ and Gibbs free energy $G=H-TS$ per mole of ideal gases of isolated m^{9}Gua and ThioTEPA molecules, as well as for a mole of ideal gases of each of m^{9}Gua···(ThioTEPA)$_n$ complexes. In the above notation E stands for the sum of translational, rotational, vibrational and the ground state electron energies with $\langle\ldots\rangle$ denoting the thermodynamical averaging over vibrational, translational and rotational states at a given temperature; $R=8.31$ J/(mol·K) is the universal gas constant and S stands for the entropy per mole of gas. Next, the corresponding binding energies have been obtained as $\Delta A = A_{\mathrm{complex}} - \Sigma A_i$, where A denotes one of E, H or G, and the subscript 'complex' or 'i' refers to their values for the whole complex or its i-th monomer respectively with proper BSSE correction included. All thermodynamic quantities have been calculated for the temperature $T=298.15$ K and pressure $p=1$ atm and, separately, for temperature $T=0$ K (the latter case will be distinguished by the subscript '0' near the quantities). The vibration spectra involved in thermochemical calculations have been obtained at the same level of theory as those used for the geometry optimization (DFT B3LYP/6-31G(d,p)). All the vibration spectra were checked for the absence of imaginary frequencies. The total processor time spent on molecular docking, geometry optimization, single point energy calculations and thermochemical calculations performed in steps (i), (ii) and (iii) was about 4446 hours.

At the final step (iv), the non-covalent interactions in each of the complexes have been investigated by R. Bader's quantum theory of atoms-in-molecules (QTAIM).[38] In this part of the study, the spatial distribution of the electron charge density obtained at the DFT B3LYP/6-31G(d,p) theory level was analyzed with AIMExt program from the AimAll package.[39] The presence of the bond path containing (3,-1)-type critical point which connects a pair of atoms was considered as both necessary and sufficient condition for treating this pair of atoms as 'bonded'.[40,41]

3 Results and Discussion

3.1 Geometry

The final spatial structures of the m^{9}Gua···(ThioTEPA)$_n$ ($n=1\ldots6$) complexes obtained by the quantum-chemical geometry optimization are presented in Figure 1, and their corresponding atomic coordinates suitable for 3D visualization are given in the Supporting Information (Table S1).

The effect of quantum chemical geometry optimization on complex spatial structures obtained directly from molecular docking (the structures designated as "the most stable" ones by AutoDock software) can be easily estimated from Table 1. In this table, position of each ThioTEPA molecule is characterized by two parameters: 1) $r_{\mathrm{mGua\cdots ThioTEPA}}$ – the distance between the center of masses of m^{9}Gua and the center of masses of a given ThioTEPA molecule; 2) ΔI – the distance between the centers of masses of a given ThioTEPA molecule in the initial (obtained directly from AutoDock) and quantum-chemically optimized complex structures under the condition that the m^{9}Gua spatial positions and orientations in these structures are made as close as possible in terms of root-mean-square deviations by the method described by Kabsch.[42]

The general trends are the following: first, with few exceptions, quantum-chemical geometry optimization practically does not change $r_{\mathrm{mGua\cdots ThioTEPA}}$ (the typical change does not exceed 0.5 Å, while the maximum change is 1.43 Å). On the other hand, the changes of ΔI are much more evident, and during quantum-chemical complex geometry optimization the ThioTEPA molecules are likely to 'travel' on the sur-

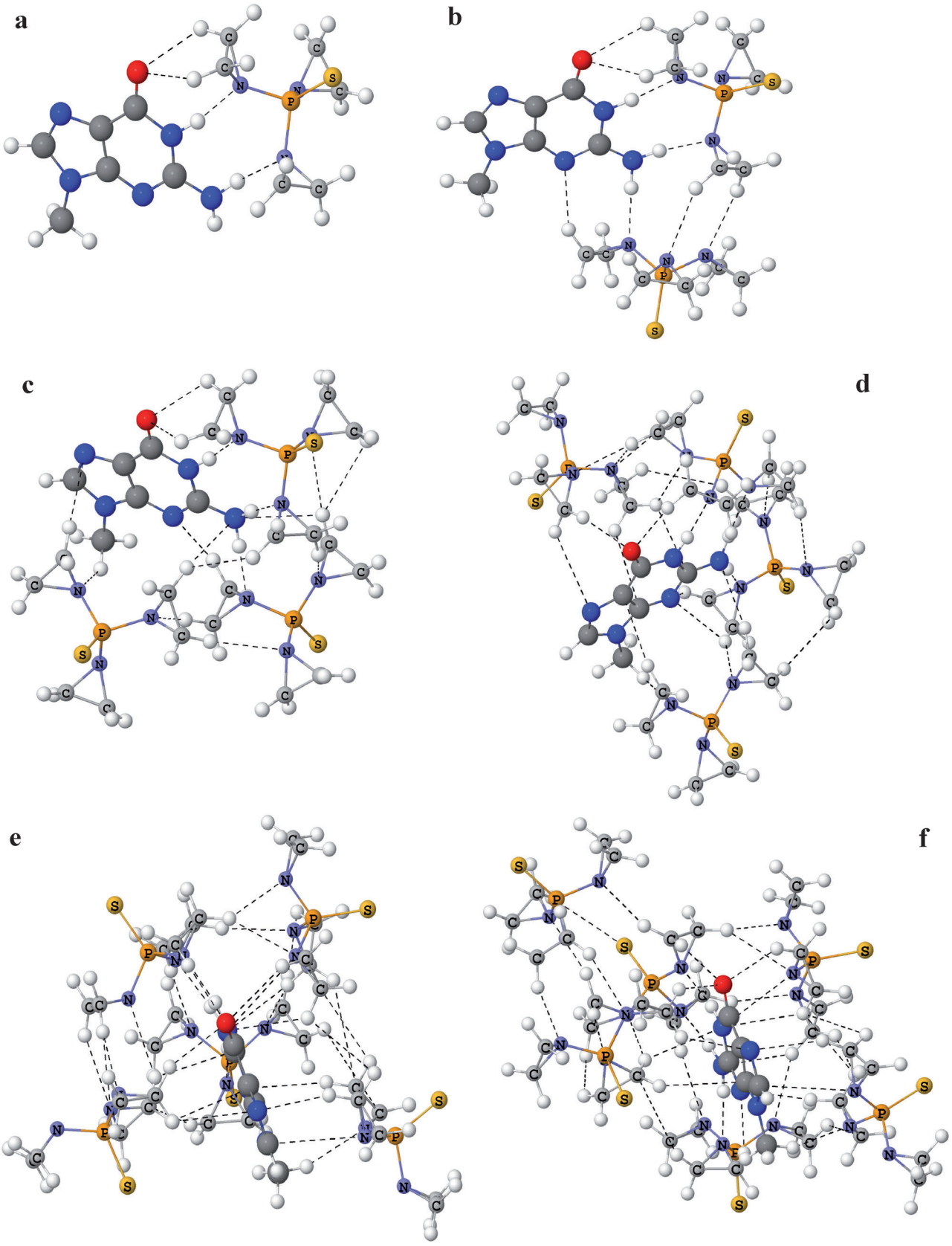

Figure 1. The spatial structure of methylated guanine-(ThioTEPA)$_n$ ($n = 1$ (a), 2 (b), 3 (c), 4 (d), 5 (e), 6 (f)) complexes obtained by the quantum-chemical calculations at the DFT B3LYP/6-31G(d,p) level of theory. The non-covalent interactions revealed by the QTAIM topological analysis of spatial electron density distributions are shown with dashed lines.

Table 1. The absolute changes in positions of ThioTEPA molecules in m9Gua···(ThioTEPA)n complex observed in the result of geometry optimization by quantum-chemical DFT method.

ThioTEPA molecule number	rmGua···ThioTEPA (Å)		ΔI (Å)
	Docking	DFT	
m^9Gua···$(ThioTEPA)_1$			
1	5.60	6.09	1.16
m^9Gua···$(ThioTEPA)_2$			
1	5.60	6.13	**1.93**
2	5.20	6.63	**4.12**
m^9Gua···$(ThioTEPA)_3$			
1	6.13	6.11	0.08
2	6.64	6.71	1.49
3	3.88	5.30	**2.09**
m^9Gua···$(ThioTEPA)_4$			
1	4.23	5.59	**2.26**
2	6.12	5.88	1.42
3	6.71	6.73	0.82
4	5.30	5.30	0.36
m^9Gua···$(ThioTEPA)_5$			
1	5.90	6.18	0.82
2	5.59	5.38	0.98
3	5.89	5.87	0.32
4	6.73	6.81	0.83
5	5.30	5.61	**2.78**
m^9Gua···$(ThioTEPA)_6$			
1	7.67	9.07	1.67
2	6.19	6.16	0.21
3	5.38	5.54	1.37
4	5.87	5.94	0.13
5	6.81	6.78	0.43
6	5.62	5.57	0.53

face of the imaginary sphere centered at the m^9Gua molecule. At the same time, it can be easily seen that the larger the complex, the smaller are the typical values of ΔI. The most substantial change in ThioTEPA position is observed for m^9Gua···$(ThioTEPA)_2$ complex with $\Delta I \approx 4$ Å (see Figure 2), while in all the other cases ΔI values are about the length of typical chemical bonds, therefore the overall complex geometry changes can be considered inessential.

To sum up, the standard molecular docking has been found to perform quite well in predicting the geometry of non-covalently bound m^9Gua···ThioTEPA complexes. No dramatic changes in complex structures are observed after their geometry optimization by quantum-chemical DFT method. Both the distance from ThioTEPA to m^9Gua and the relative position of ThioTEPA molecule obtained by molecular docking have an accuracy of about the typical chemical bond length (~1…2 Å).

3.2 Energetics and Thermodynamics

The stabilization energies (both purely electronic and those accounting for internal molecular degrees of freedom) for all the complexes were evaluated as described in Section 2 'Computation Details' and the results are presented in Table 2. It should be noted that BSSE correction value equals from 22 to 25% of the total stabilization energy, and increases with the increase in complex size, while the sum of deformation energies of complex fragments is only 2 to 9% and decreases with the increase in complex size.

Among all the complexes, the complex with four ThioTEPA molecules has the lowest Gibbs free stabilization energy. At the same time, the overall electron stabilization energy as well as the stabilization enthalpy increases gradually

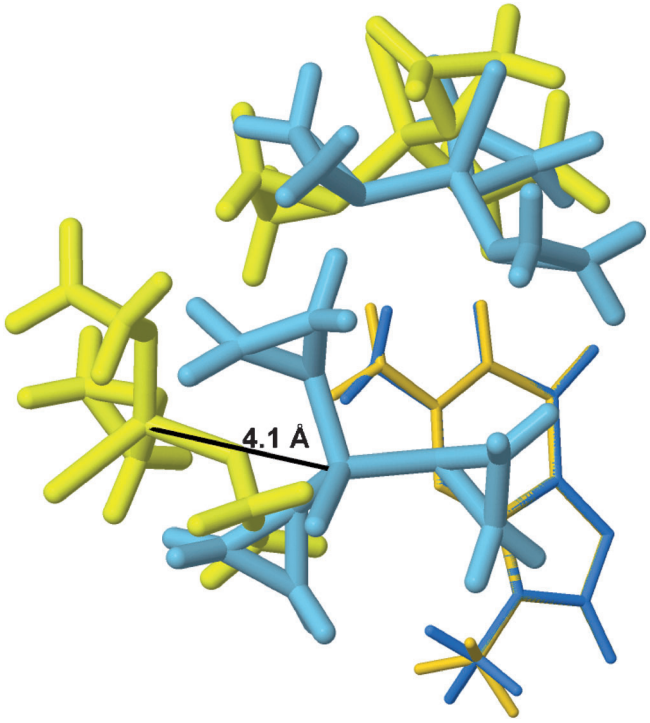

Figure 2. The spatial structures m^9Gua···$(ThioTEPA)_2$ complexes obtained directly from the molecular docking (blue) and by additional quantum chemical geometry optimization (yellow).

Table 2. The stabilization electronic energy (ΔE), internal energy (ΔU, ΔU_0), enthalpy (ΔH) and Gibbs free energies (ΔG) of m^9Gua···$(ThioTEPA)_n$ complexes with $n = 1$–6 (theoretical calculation at the RI-MP2/cc-pVTZ // DFT B3LYP/6-31G(d,p) theory level)

n	ΔE	ΔU_0	ΔU	ΔH	ΔG
1	−17.1	−16.0	−15.0	−15.6	−4.5
2	−29.6	−27.5	−25.4	−26.5	−4.3
3	−42.9	−40.1	−36.6	−38.4	−5.4
4	−55.3	−52.1	−47.3	−49.6	−6.7
5	−67.5	−63.7	−57.5	−60.5	−6.4
6	−77.6	−73.3	−65.7	−69.3	−4.6

with the increase of complex size. Moreover, the electronic stabilization energy, enthalpy and internal energy per one ThioTEPA molecule are rather close in values for all the complexes, and although they slightly decrease with the increase of the complex size (for example, ΔU per ThioTEPA molecule ranges from −15 kcal/mol in $n=1$ complex to about −11 kcal/mol in $n=6$ complex) showing little anti-cooperativity of binding, the main reason of lower Gibbs stabilization energy for large complexes should be attributed to entropic contribution. Therefore, we can claim that this is the principal factor which determines the overall complex stability. This finding is not surprising since the larger the complex, the greater number of collective low-frequency vibrations appear in its vibration spectrum. At $T=298.15$ K, the average value of $T\Delta S$ contribution per one ThioTEPA molecule is about +11 kcal/mol and is approximately constant for all complexes.

In contrast to entropic contribution, the stabilizing contributions into Gibbs free energy of the complex, are due to the internal energy, viz. the electron energy. In the next section we shall discuss the origin of complex stabilization in more details.

3.3 Stabilizing Interactions

The interatomic interactions in the complexes under study were identified by using QTAIM approach. As already mentioned in Section 2 'Computation Details', two atoms are considered to be 'bonded' if they are connected with the bond path[38,40] containing a (3,-1)-type critical point. To quantify the strength of the interaction, the electron charge density and its laplacian calculated at the bond critical point are routinely used.[38]

Table 3. The geometrical and physical properties of non-covalent interactions in mGua···(ThioTEPA)$_n$ complex obtained by quantum theory of atoms in molecules. ρ^{cp} is the spatial electron charge density evaluated at the bond critical point and $\Delta\rho^{cp}$ is the value of its laplacian (the trace of Hessian matrix) evaluated at the same point of space.

Bond[a]	Count	XY (Å)		< AHB (deg.)		ρ^{cp} (10^{-2} a.u.)		$\Delta\rho^{cp}$ (10^{-2} a.u.)	
		min	max	min	max	min	max	min	max
$n=1$									
C–H···O	2	3.195	3.371	112.7	126.8	0.56	1.12	2.26	3.72
N–H···N	2	2.967	3.084	159.4	169.0	2.25	3.09	5.46	7.33
$n=2$									
C–H···O	2	3.264	3.307	118.2	121.6	0.73	0.86	2.71	3.07
C–H···N	3	3.476	3.796	142.2	158.5	0.52	1.14	1.64	3.02
N–H···N	3	2.977	3.057	166.2	174.9	2.52	2.98	6.08	7.11
$n=3$									
C–H···H–C	2	4.398	4.822	116.7	139.6	0.08	0.19	0.30	0.65
C–H···O	2	3.299	3.324	119.1	121.1	0.72	0.79	2.66	2.85
N–H···N	3	2.967	3.079	168.7	173.9	2.46	3.08	5.88	7.35
C–H···N	9	3.469	4.715	106.7	175.2	0.08	1.05	0.35	3.02
$n=4$									
C–H···H–C	2	4.415	5.008	109.7	135.5	0.07	0.19	0.23	0.60
C–H···O	2	3.321	3.390	143.5	154.5	1.15	1.20	3.46	3.59
C–H···C	3	3.689	4.022	120.6	153.9	0.26	0.49	0.89	1.56
N–H···N	3	2.975	3.041	173.8	175.7	2.67	3.12	6.36	7.31
C–H···N	12	3.478	4.545	113.0	174.9	0.07	1.00	0.32	2.72
$n=5$									
C···N	1	3.482	3.482	–	–	0.43	0.43	1.38	1.38
N···N	1	3.493	3.493	–	–	0.42	0.42	1.39	1.39
C–H···H–C	2	4.385	4.624	99.7	113.7	0.08	0.08	0.28	0.31
C–H···O	3	3.300	3.682	140.2	150.7	0.49	1.22	1.79	3.71
N–H···N	3	2.998	3.045	167.3	176.7	2.60	2.89	6.36	6.72
C–H···C	5	3.591	5.554	115.7	164.8	0.01	0.37	0.04	1.46
C–H···N	17	3.475	3.877	113.3	174.5	0.35	1.23	1.21	3.29
$n=6$									
C···N	1	3.559	3.559	–	–	0.38	0.38	1.27	1.27
N···N	1	3.504	3.504	–	–	0.41	0.41	1.35	1.35
C–H···S	1	4.170	4.170	135.7	135.7	0.33	0.33	1.00	1.00
C–H···C	3	4.011	5.493	128.8	155.8	0.01	0.28	0.06	0.87
C–H···O	3	3.332	3.545	136.8	142.9	0.70	1.15	2.37	3.52
N–H···N	3	2.996	3.047	166.9	176.5	2.58	2.94	6.32	6.99
C–H···N	19	3.474	3.974	120.7	175.0	0.27	1.25	0.93	3.34
C–H···H–C	4	4.095	4.432	100.0	140.8	0.12	0.20	0.41	0.67

[a] A···B denotes van der Waals contact, X–H···B are hydrogen bonds and X–H···H–Y stands for dihydrogen bonds.

This approach can reveal both covalent (i.e., chemical bonds) and non-covalent interactions (hydrogen bonds, van der Waals contacts etc.). However, since the process of complex formation requires no changes in the molecule chemical structures, we shall investigate only non-covalent interactions in this section.

Whenever two atoms, A and B, are bound with a non-covalent interaction, one can classify this interaction as follows: van der Waals contact (if none of the atoms A and B is a hydrogen atom), hydrogen bond (if only one of the atoms A and B is a hydrogen atom) or dihydrogen bond (if both atoms, A and B, are hydrogen atoms). It is convenient to designate these types of interactions as A···B, X–A···B and X–A···B–Y respectively, where X and Y denote other atoms, covalently bound to A and/or B. We shall use this simple classification hereinafter.

The network of non-covalent interactions revealed with QTAIM approach in all the complexes under study is presented in Figure 1 by dashed lines. It can be easily observed that the total number of non-covalent bonds increases rapidly with the increase in the complex size, being equal to 4, 8, 16, 22, 32, 35 for $n = 1 \ldots 6$ complexes respectively. Hence, the average number of non-covalent bonds per one ThioTEPA molecule equals 5.2.

Table 3 contains detailed information about non-covalent bond types, their geometrical and physical properties for each complex. It follows from these data that hydrogen bonding (mainly NH···N or CH···B bond, where B is the nitrogen or oxygen atom) is the major type of non-covalent interactions revealed in the complexes.

It should be noted that there is a pronounced correlation ($R > 0.99$) between the total number N of non-covalent bonds of all types in the complex and the complex electronic stabilization energy ΔE (Figure 3). The linear interpolation predicts the contribution of 1.8 kcal/mol per each non-covalent bond. The fact that extrapolation of this dependence to $N = 0$ gives $|\Delta E|$ about 12 kcal/mol, indicates that, apart from non-covalent 'bonds', non-bonding interactions (such as long-range electrostatics, for example) also contribute into the complex stabilization. On the other hand, such an extrapolation assumes that all the types of non-covalent bonds have the same energy, which should not be accepted a priori.

In any case, we can conclude that non-covalent bonds are clearly the main physical factor responsible for the complex stabilization. In order to estimate the relative role of each separate type of these bonds, we have estimated their energy by the method of Espinosa–Molins–Lecomte (EML).[43] According to this approach, the energy of individual non-covalent bond can be estimated as follows[43]

$$E_{\mathrm{Esp}} = -\frac{1}{2} \cdot a_0^3 \cdot V^{\mathrm{el}}(\vec{r}_{\mathrm{cp}}) \qquad (2)$$

where a_0 is the Bohr radius (equals 1 in atomic units) and V^{el} is the volumetric density of the virial of the forces exerted on electrons, which is evaluated at the bond critical point on the bond path corresponding to the given interaction.[43,44]

Although EML method has several limitations (in particular, in predicting exact bonding energy of 'strong' interactions – see discussions in Nikolaienko et al.[41] and Mata et al.[44]), it can serve as a reasonable indicator for the bond strength. Indeed, as can be concluded from Figure 4, the complex electronic stabilization energy ΔE correlates with the sum of energies of individual non-covalent bonds in each complex estimated by EML method. The root-mean-square deviation of the corresponding linear fit is only 1.5 kcal/mol, which, along with high correlation coefficient

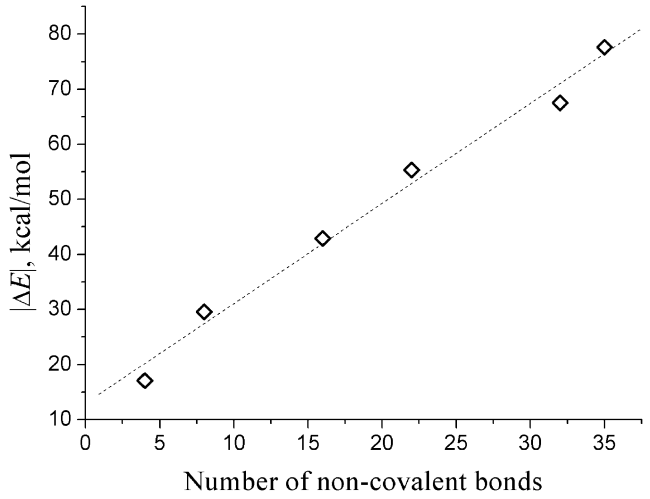

Figure 3. The correlation between the complex stabilization energy and the total number of non-covalent bonds of all types in the complex.

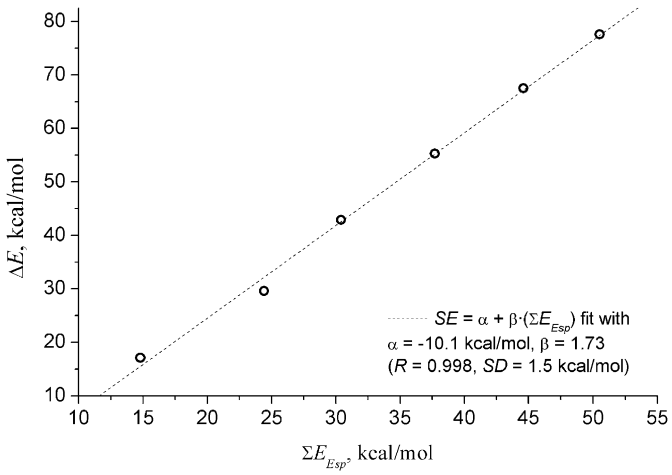

Figure 4. Correlation between the complex stabilization energy (ΔE) and the sum of energies of individual non-covalent bonds (ΣE_{Esp}) estimated with the Espinosa–Molins–Lecomte method.

$R = 0.998$, supports the reliability of EML method for comparing *relative* strength of bonds.

This finding allows us to point out those types of non-covalent bonds which make the dominant contribution into the complex stabilization. For this purpose, we have calculated the relative contribution p_{type} of all the bonds of a given type into the total stabilization energy of each complex by the simple relation

$$p_{type} = \frac{\sum_{i \in type} \varepsilon_i}{\sum_{j \in \text{all bonds in complex}} \varepsilon_j} \quad (3)$$

where the 'type' of the bond is determined by the types of atoms A and B for van der Waals contacts (A···B), X and B for hydrogen bonds (X–H···B) or X and Y for dihydrogen bonds (X–H···H–Y), ε_i is the energy of the individual bond obtained by EML formula, the summation over i in numerator goes over all the bonds of a given type in the complex, while the summation over j in denominator goes over all the bonds in the complex. The p_{type} values obtained in this way are given in Table. 4.

These data imply that the stability of small complexes ($n = 1 \ldots 4$) is maintained by the hydrogen bonds of NH···N and CH···O type, while in bigger complexes ($n = 5$, 6) hydrogen bonds of CH···N type become the most significant binding interaction, with NH···N and CH···O bonds sharing the second and the third ranks.

Hence, it may be concluded that, all in all, C–H···O, N–H···N and C–H···N hydrogen bonds contribute about 93…100% in the total stabilization energy. The contribution from van der Waals contacts and dihydrogen bonds as well as C–H···S and C–H···C bonds is negligible (not more than 5% in total).

The leading role of hydrogen bonds in stabilizing m^{9}Gua···ThioTEPA complexes motivates us to pay more attention to their geometrical and physical properties. For XH···B bonds these properties are the distance XB between proton donor (X) and proton acceptor (B) as well as XHB angle, which can serve as a measure of bond 'straightness'. Both parameters are shown on Figure 5 for all types of hydrogen bonds in all the complexes.

Among all the hydrogen bonds that have been revealed, the number of CH···N bonds is the largest, which makes it possible to analyze their statistical properties. First of all, it can be concluded that CN distance in all of CH···N bonds except two (i.e. 97%) lies in a relatively narrow range of 3.70 Å ± 0.30 Å, while the distribution of CHN angle is much wider (see Figure 5). This is likely due to the fact that the energy of CH···N bonds depends strongly on the interatomic distance (especially on H···N – see Figure 6, a), whereas there is no clear correlation between the binding energy and CHN angle (see Figure 6, b). Hence, CHN angle is a rather flexible structural parameter which can freely 'adjust' itself to the complex geometry.

Table 4. The relative contribution (in %) of all the bonds of a given type into the overall stabilization energy of mGua···(ThioTEPA)$_n$ complexes.

Bond type	Bond count	Energetic contribution p_{type} (%)
$n = 1$		
N–H···N	2	76.6
C–H···O	2	23.4
$n = 2$		
N–H···N	3	70.3
C–H···N	3	16.7
C–H···O	2	13.0
$n = 3$		
N–H···N	3	56.9
C–H···N	9	32.2
C–H···O	2	9.8
C–H···H–C	2	1.1
$n = 4$		
N–H···N	3	48.0
C–H···N	12	33.6
C–H···O	2	13.4
C–H···C	3	4.2
C–H···H–C	2	0.8
$n = 5$		
C–H···N	17	42.7
N–H···N	3	39.0
C–H···O	3	11.6
C–H···C	5	2.7
N···N	1	1.9
C···N	1	1.6
C–H···H–C	2	0.5
$n = 6$		
C–H···N	19	48.0
N–H···N	3	34.7
C–H···O	3	10.9
N···N	1	1.7
C–H···H–C	4	1.5
C···N	1	1.3
C–H···C	3	1.1
C–H···S	1	0.8

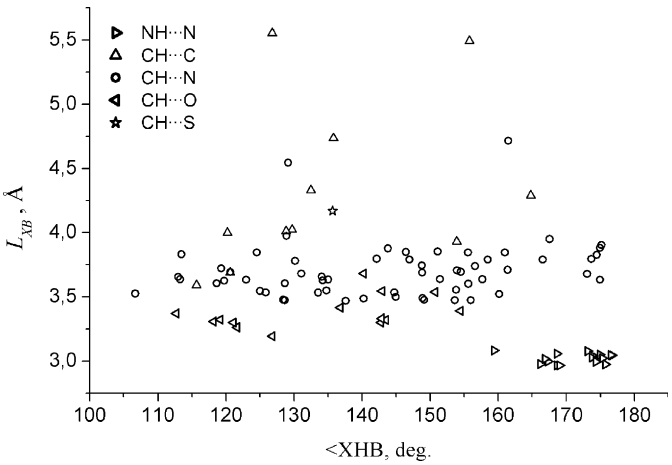

Figure 5. The geometrical characteristics of hydrogen bonds (XH···B) in the angle (XHB)-distance (XB) frame.

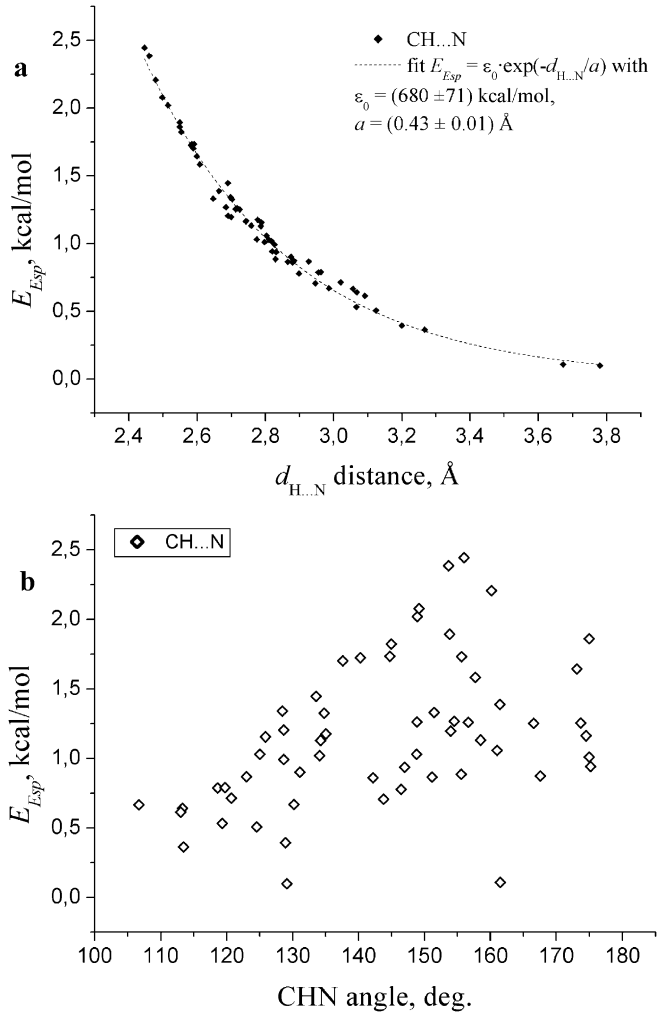

Figure 6. The dependence of CH···N bond energy E_{Esp} estimated by Espinosa–Molins–Lecomte method on H···N distance (a) and on CHN angle (b) for CH···N hydrogen bonds found in all the complexes.

While the CH···N bond energy E_{Esp} as the function of H···N distance d can be accurately approximated with a smooth curve $E_{Esp} = \varepsilon_0 \cdot e^{-d/a}$ giving the root-mean-square error (*RMSE*) of only 0.06 kcal/mol, the *RMSE* of the same fitting for E_{Esp} as a function of CN distance is 6 times higher (0.36 kcal/mol). As for the numerical values of the fitting coefficients, $\varepsilon_0 = 680$ kcal/mol has the order of magnitude of one Hartree (627.5 kcal/mol), while $a = 0.43$ Å has the order of the Bohr radius (0.53 Å), both being at the natural scales of corresponding atomic properties.

The distinctive feature of NH···N bonds, which are as important for small complexes as CH···N bonds for the bigger complexes, is that their geometrical properties are concentrated in a relatively narrow range: (3.03 ± 0.06) Å for N–N distance, (2.03 ± 0.08) Å for H···N distance and 168° ± 9° for NHN angle. Thus, the geometry of all the bonds of this type is really 'linear'.

As for the physical properties of hydrogen bonds, the electron charge density at corresponding bond critical points covers by almost three orders of magnitude broad range: from $\sim 7 \times 10^{-5}$ a.u. (here a.u. = e/a_0^3) in the weakest CH···C bond found in m^9Gua···(ThioTEPA)$_5$ complex to $3.1 \cdot 10^{-2}$ a.u. in the strongest NH···N bond found in m^9Gua···(ThioTEPA)$_4$ complex (see Table 3).

4 Conclusions

Non-covalent binding of methylated nitrogenous DNA base guanine (m^9Gua) to ThioTEPA anticancer drug has been studied by using molecular docking, DFT, ab initio and QTAIM methodologies. The energetically most favorable structures of m^9Gua···(ThioTEPA)$_n$ complexes with $n = 1$ to 6 found by molecular docking have been compared with the structures obtained by DFT geometry optimization. It is shown that the standard molecular docking techniques are able to predict spatial positions of ThioTEPA molecules with respect to m^9Gua within the accuracy of about a typical covalent bond length: the average difference in the ThioTEPA center-of-masses positions is 1.2 Å; in m^9Gua···(ThioTEPA)$_2$ complex, however, it reaches 4.1 Å. Complex binding Gibbs free energies (ΔG) obtained by quantum chemical calculations indicate that the complex with $n = 4$ ThioTEPA molecules is the most stable one ($\Delta G = -6.7$ kcal/mol), while the complexes with $n = 1$ and $n = 6$ are slightly less favorable (for both ΔG is about −4.5 kcal/mol).

The complex stabilization enthalpy per one ThioTEPA molecule is rather close in all the complexes (from −15.6 kcal/mol in $n = 1$ complex to −11.5 kcal/mol in $n = 6$ complex) and decreases inconsiderably with the increase of the complex size, thus revealing a small anticooperativity effect in binding. The complex binding energy grows almost linearly as the complex size increases, and so does the complex entropy: at $T = 298$ K the value of $T\Delta S$ part of Gibbs free energy per one ThioTEPA molecule is almost constant in all complexes (ranges from +10.8 to +11.1 kcal/mol). Hence, further analysis of entropic contribution into Gibbs free energies of the complexes (probably, beyond the standard rigid rotator-harmonic oscillator approximation) may be desirable in order to calculate accurately the binding energies at room temperature and determine the most stable complex.

The major stabilizing interaction responsible for complexation has been shown to be C–H···O, N–H···N and C–H···N hydrogen bonds, which together contribute about 93…100% of total stabilization energy. Other types of bonds found in the complexes, namely C–H···S and C–H···C, C···N and C···C van der Waals contacts, and C–H···H–C dihydrogen bonds, have been found to make a negligibly small (< 5%) total contribution into the complex binding energy. As for the C–H···N hydrogen bonds, which are the most sig-

nificant binding interaction in large ($n=5$, 6) complexes, their energy has been found to depend strongly on H···N distance and to be almost insensitive to CHN angle, which makes the latter an easily 'adjustable' degree of freedom during the complex formation. The distinctive feature of NH···N bonds, which are as important for small ($n=1$ to 4) complexes as CH···N bonds for larger complexes, is that their geometrical properties are concentrated in the narrow range: 3.03 ± 0.06 Å for NN distance, and $168^\circ\pm9^\circ$ for NHN angle.

The obtained results may be useful in further modeling the chemical reactions between the ThioTEPA and DNA constituents and getting a deeper insight in the mechanism of ThioTEPA drug action.

Acknowledgements

The authors highly appreciate the access to high-performance computational resources provided by the *Boholubov Institute for Theoretical Physics of National Academy of Sciences of Ukraine*. We also encourage kind help of *Dr. Roman Zhurakivsky* in setting up the calculations and express our gratitude to *Mrs. Ludmila Rudkovska* for her valuable assistance in the manuscript preparation.

References

[1] D. C. Swinney, *Nat. Rev. Drug Discov.* **2004**, *3*, 801 – 808.
[2] Y. Kawasaki, E. Freire, *Drug Discov. Today* **2011**, *16*, 985 – 990.
[3] M. P. Sykes, D. A. Karnofsky, F. S. Philips, J. H. Burchenal, *Cancer* **1953**, *6*, 142 – 148.
[4] S. Nussbaumer, P. Bonnabry, J.-L. Veuthey, S. Fleury-Souverain, *Talanta* **2011**, *85*, 2265 – 2289.
[5] M. J. van Maanen, C. J. M. Smeets, J. H. Beijnen, *Cancer Treat. Rev.* **2000**, *26*, 257 – 268.
[6] E. van der Wall, J. H. Beijnen, S. Rodenhuis, *Cancer Treat. Rev.* **1995**, *21*, 105 – 132.
[7] S. Rodenhuis, A. Westermann, M. J. Holtkamp, W. J. Nooijen, J. W. Baars, E. van der Wall, I. C. Slaper-Cortenbach, J. H. Schornagel, *J. Clin. Oncol.* **1996**, *14*, 1473 – 1483.
[8] M. J. van Maanen, I. M. Tijhof, J. M. A. Damen, C. Versluis, J. J. K. van den Bosch, A. J. R. Heck, S. Rodenhuis, J. H. A. Beijnen, *Cancer Res.* **1999**, *59*, 4720 – 4724.
[9] J. Xu, Q. Wang, *J. Insect. Physiol.* **2010**, *56*, 102 – 106.
[10] D. Kheffache, O. Ouamerali, *J. Mol. Model.* **2010**, *16*, 1383 – 1390.
[11] H. Torabifard, A. Fattahi, *Struct. Chem.* **2013**, *24*, 1 – 11.
[12] S. M. Musser, S.-S. Pan, M. J. Egorin, D. J. Kyle, P. S. Gallery, *Chem. Res. Toxicol.* **1992**, *5*, 95 – 99.
[13] N. A. Cohen, M. J. Egorin, S. W. Snyder, B. Ashar, B. E. Wietharn, S.-S. Pan, D. D. Ross, J. Hilton, *Cancer Res.* **1991**, *51*, 4360 – 4366.
[14] Z. H. Siddik, in *The Cancer Handbook* (Ed: M. R. Alison), Wiley, Chichester, **2002**, pp. 1 – 16.
[15] H. Torabifard, A. Fattahi, *J. Mol. Model.* **2012**, *18(8)*, 3563 – 3576.
[16] L. F. Sukhodub, *Biopolimery i kletka (Kiev)* **1991**, *7*, 15 – 32.
[17] L. F. Sukhodub, *Mass Spectrom. Rev.* **1995**, *14*, 253 – 254.
[18] a) M. S. de Vries, P. Hobza, *Annu. Rev. Phys. Chem.* **2007**, *58*, 585 – 612; b) K. Seefeld, R. Brause, T. Häber, K. Kleinermanns, *J. Phys. Chem. A* **2007**, *111*, 6217 – 6221; c) E. Nir, C. Janzen, P. Imhof, K. Kleinermanns, M. S. de Vries, *J. Chem. Phys.* **2001**, *115*, 4604 – 4611; d) W. Chin, M. Mons, I. Dimicoli, F. Piuzzi, B. Tardivel, M. Elhanine, *Eur. Phys J. D*, **2002**, *20*, 347 – 355.
[19] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *16*, 2785 – 2791.
[20] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639 – 1662.
[21] a) T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, *Ukr. Biokhim. Zh.* **2010**, *82*(6), 76; b) T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, *Ukr. Biokhim. Zh.* **2011**, *83*(4), 16; c) T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, O. O. Missura, *Ukr. Biokhim. Zh.* **2011**, *83*(1), 54; d) T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, *Biopolym. Cell* **2011**, *27*(1), 74; e) T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, *J. Biomol. Struct. Dyn.* **2011**, *29*(3), 563; f) T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, *Biopolym. Cell* **2011**, *27*(4), 291.
[22] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, *Gaussian 03, Revision E.01*, Gaussian, Inc., Wallingford, CT, **2004**.
[23] P. A. Kollman, *Curr. Opin. Struct. Biol.* **1992**, *2*, 765 – 771.
[24] J. Petruska, L. Sowers, M. Goodman, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 1559 – 1562.
[25] M. J. S. Dewar, D. Storch, *Proc. Natl. Acad. Sci. USA* **1985** *82*, 2225 – 2229.
[26] M. J. S. Dewar, *Enzyme* **1986**, *36*, 8 – 20.
[27] E. L. Mertz, L. I. Krishtalik, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2081 – 2086.
[28] M. S. Ahmadi, A. Fattahi, *J. Theor. Comput. Chem.* **2010**, *9*, 585 – 609.
[29] O. O. Brovarets', Y. P. Yurenko, I. Ya. Dubey, D. M. Hovorun, *J. Biomol. Struct. Dyn.* **2012**, *29(6)*, 1101 – 1109.
[30] C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618 – 622.
[31] D. Cremer, *WIREs Comput. Mol. Sci.* **2011**, *1*, 509 – 530.
[32] T. H. Dunning Jr., *J. Chem. Phys.* **1989**, *90*, 1007 – 1023.
[33] D. E. Woon, T. H. Dunning Jr., *J. Chem. Phys.* **1993**, *98*, 1358 – 1371.
[34] S. Kossmann, F. Neese, *Chem. Phys. Lett.* **2009**, *481*, 240 – 243.
[35] F. Weigend, *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285 – 4291.
[36] F. Neese, *WIREs Comput. Mol. Sci.* **2012**, *2*, 73 – 78.
[37] S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553 – 566.
[38] R. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, Oxford, **1994**.
[39] T. A. Keith, *AIMAll (Version 08.11.29)*, TK Gristmill Software, Overland Park KS, USA, **2012**, aim.tkgristmill.com.

[40] R. F. W. Bader, *J. Phys. Chem. A* **1998**, *102*, 7314 – 7323.
[41] T. Yu. Nikolaienko, L. A. Bulavin, D. M. Hovorun, *Phys. Chem. Chem. Phys.* **2012**, *14*, 7441 – 7447.
[42] W. Kabsch, *Acta Cryst.* **1976**, *A32*, 922 – 923.
[43] E. Espinosa, E. Molins, C. Lecomte, *Chem. Phys. Lett.* **1998**, *285*, 170 – 173.
[44] I. Mata, I. Alkorta, E. Espinosa, E. Molins, *Chem. Phys. Lett.* **2011**, *507*, 185 – 189.

Received: April 2, 2013
Accepted: November 12, 2013
Published online: February 12, 2014

