Role of Aromacity, Ring Strain, and Stereochemistry of Selected Disulfides and Their Congeners in the Oxidative Linkage of DNA Strands at the Major Groove. A Computational Study†

Robert Izsak,‡ Balazs Jojart,§ Imre G. Csizmadia,‡,|| and Bela Viskolcz*,‡

Department of Chemistry and Chemical Informatics, University of Szeged, Szeged, Boldogasszony sgt. 6. Hungary 6725, Department of Pharmacodynamics and Biopharmacy, University of Szeged, Szeged, Eötvös u. 6. Hungary 6720, and Department of Chemistry, University of Toronto, Ontario M5S 3H6, Canada

Received April 4, 2006

Three aromatic disulfides and their two selected congeners were studied in order to see if they could be initiators or catalysts in a process that aims to tie together the two strands of a DNA double helix. Thermodynamic functions were calculated from first principle molecular computations on deoxyribose models and the disulfide congener. The gas reaction of one of the disulfides turned out to be the closest to the thermoneutral process. Solvation study caused a relatively minor energetic modification. The disulfides were subjected to docking studies, and the stabilities of the complexes obtained were further analyzed by molecular dynamics simulations. Only one of the two atropic disulfide enantiomers was the matching partner of the chiral DNA double helix, and therefore this remained in the major groove of the DNA. The other enantiomer dissociated from the major groove as it was incapable of forming a stable complex with the chiral double helix. A mechanistic scheme for the reaction has been investigated.

INTRODUCTION

One of the main causes of cancer is that the DNA is damaged. Traditionally, anticancer drugs were discovered through large-scale screening of synthetic chemicals and natural products against animal tumor systems (murine leukemias). The compounds discovered in 1950−1970 largely interact with the DNA or its precursors, inhibiting the synthesis of new genetic material or causing irreparable damage to DNA itself. These agents may be grouped as follows: (1) purin and pyrimidine biosynthesis inhibitors (pentostatin,[1] 6-mercaptopurin,[2] antifolates[3]); (2) covalant linkage between the two DNA strands (alkylating agents,[4,5] platinum analogues[6,7]); (3) topoisomerase blockers (anthracyclines)[8]; (4) protein synthesis inhibitors (L-asparaginase); and (5) microtubulus function inhibitors (Vinca alkaloids[9]). At the present time in the cancer therapy, besides these drugs, the following are also included: imatinib[10] (tyrosine kinase inhibitor), interferon-α, interleukin, hormones (adrenocortocosteroids, progestins, estrogens, etc.).

The reactions of therapeutic alkylating agents (melphalan, chlorambucil) and platinum derivatives (cisplatin and carboplatin) result in the formation of covalent linkages between the two strands or in one strand. The chemotherapeutic and cytotoxic effects are directly related to the alkylation of DNA. The nitrogen atom, at position 7 of the guanine moiety, is particularly susceptible to the formation of a covalent bond with bifunctional alkylating agents and may well present the key target of alkylation. This N atom is also included in platinum cross-links between adjacent guanines on the same DNA strand; guanine-adenine cross-links also readily form. Anthracyclines can exert their anticancer effects through several mechanisms. These compounds can intercalate the DNA, generate free radicals by virtue of their quinone groups, react with cytochrome P450, and interfere with topoisomerase II.[11]

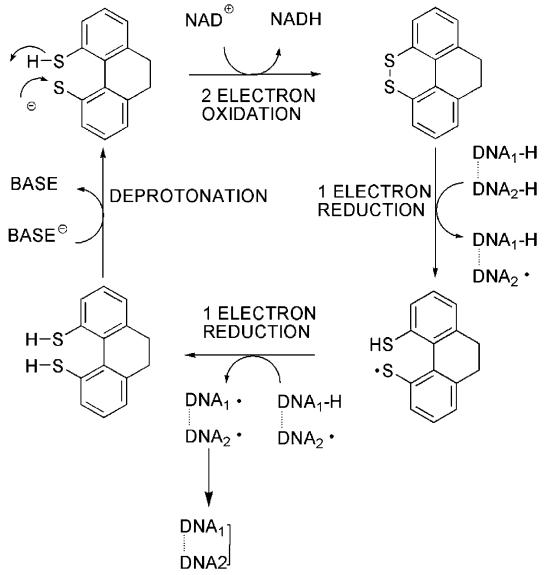


Figure 1. Inhibitor and recovery cycle of a catalytic putative drug molecule that creates cross-links in a DNA double helix by forming a covalent bond.

In this study we investigated new possible therapeutic compounds (disulfide derivatives and its congeneers) by means of computational methods. The possible reaction mechanism is based on oxidative and reductive reactions (Figure 1). It is possible to induce covalent bond formation

* Corresponding author e-mail: viskolcz@jgytf.u-szeged.hu.

† Dedicated to Professor J. Ladik, a champion of cancer theory, on the occasion of his 80th birthday.

‡ Department of Chemistry and Chemical Informatics, University of Szeged.

§ Department of Pharmacodynamics and Biopharmacy, University of Szeged.

|| University of Toronto.

10.1021/ci060123o CCC: $33.50
Published on Web 09/13/2006

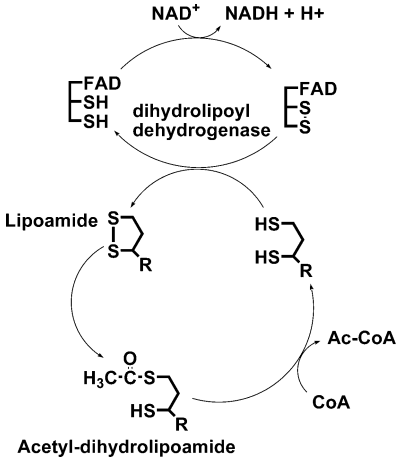


Figure 2. Lipoamide as a redox catalyst in the formation of acetyl-CoA.

between the two helices, provided that an appropriate initiator is used. One of the possible mechanisms is that a special initiator forms a radical center on the DNA helix. Before the initiator leaves the linkage position, it may form a secondary radical center on the opposite chain. The two unpaired electrons should be paired to prepare the linkage between the two chains of the DNA. The initiator should be able to be transformed very easily into biradical form in close proximity to the helix. If the initiator can be recycled biologically, it then becomes a catalyst and may continue its action several times. A recycled initator molecule would become a drug that may be used in smaller dosage. If it could be recycled 10 times, then only 10% of the required dosage would deliver 100% of the pharmaceutical effect. On the other hand, with structural modifications the compound could be interact via H-bonds with the DNA; therefore, it could act as an intercalating agent in the cancer therapy.

It should perhaps be noted that the disulfide linkage of the amide of α-lipoic acid has been observed to perform a redox catalytic effect as shown in Figure 2.[12] The mechanism suggested in Figure 1 is somewhat analogous to this process.

SCOPE

The first step in this process is to establish an ab initio study on a simplified model system, where the DNA double helix is represented by two deoxyribose-derivates. As such, these molecules modeled the pair of DNA-chains and were therefore included in the actual chemical reaction. After the thermodynamic functions of the oxidative linkage have been determined, the interaction between covalently reacting drug candidates and the DNA-chain was examined using the methods of molecular docking and dynamics.

A series of reactions was studied where heteroatomic aromatic compounds initiate H-abstractions and subsequent dimerization of monosaccharides. The general reaction scheme can be seen in Figure 3. There are altogether 18 possible reactions as **X**, **Y**, and **Z** are varied. With varying **X**, the atomic radius increases from O to Se, and the reactivity of the reaction is expected to decrease. In the case of **Y**, the ring strain and aromacity changes. Finally, **Z** can be either F or OH as both are highly electronegative, but in the case of OH, there is a possibility of H-bonds, which would affect reaction thermodynamics and perhaps the mechanism itself. The different heteroaromatic molecules will be discussed later; all molecules were numbered using the IUPAC nomenclature of phenanthrene. The aromatics studied can be considered as 4,5 substituted cases of this molecule, at least as far as numbering is concerned, which is depicted in Figure 3 (right).

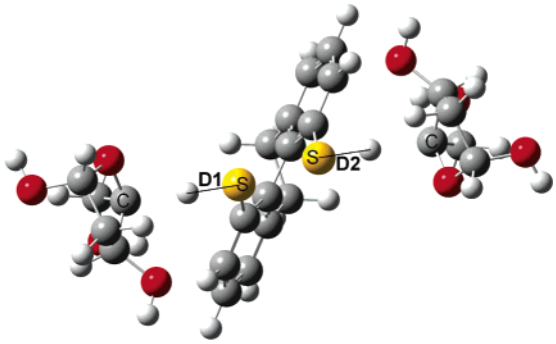


Figure 4. Compound **B**, **X** = S and **Z** = OH as a model system for reaction mechanistic calculations. During the rigid scan D1 and D2 distances were calculated, and the H atoms here are in an intermediate position between the atoms S and C. The S−C distances were kept constant at approximately 3.1 Å.

Beyond the ab initio studies, molecular mechanics methods were also used. With these latter methods, it was possible to study the interaction between the whole DNA and the drug candidates. With the aid of molecular docking we made an attempt to identify the most likely binding cavity in the DNA major groove. Finally, thereafter using explicit water molecular dynamics method the stability of the DNA-drug candidate complexes were investigated.

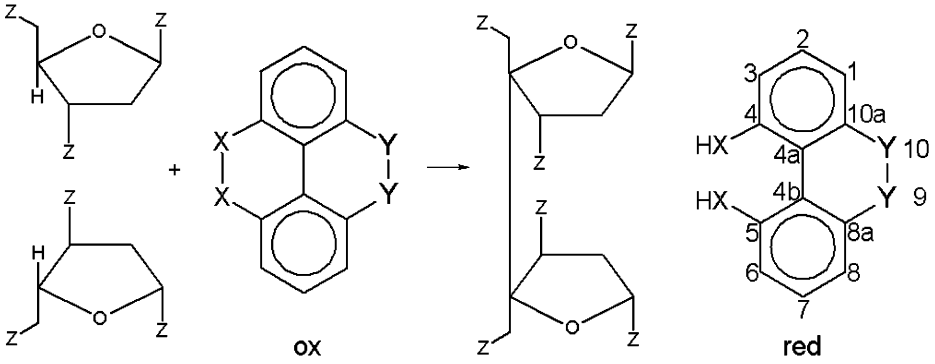


Figure 3. The general reaction scheme to model DNA dimerization via double H-abstraction. **X** = O, S, or Se; −(**Y**−**Y**)− = −(H H)− (**A**), −(CH_2−CH_2)− (**B**), or −(CH=CH)− (**C**); **Z** = F or OH.

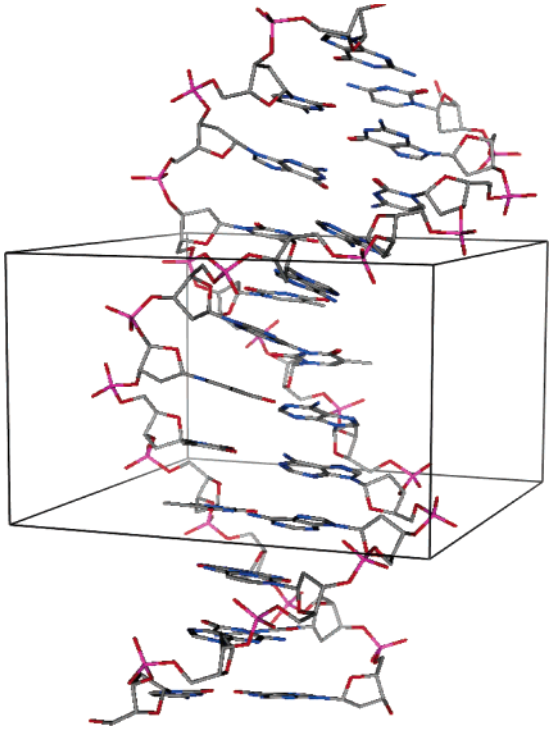


Figure 5. The position and size (60 × 60 × 40 grid points) of the docking box used in the docking procedure; the lattice point distances within this box were set to 0.325 Å apart from each other.

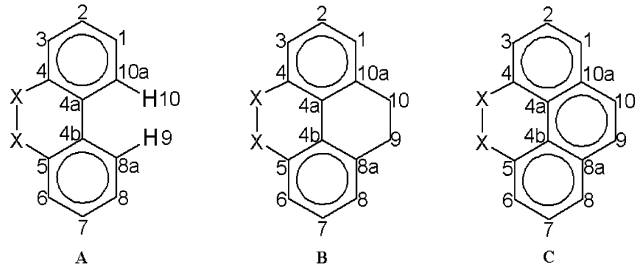


Figure 6. Studied aromatic structures and their numbering schema.

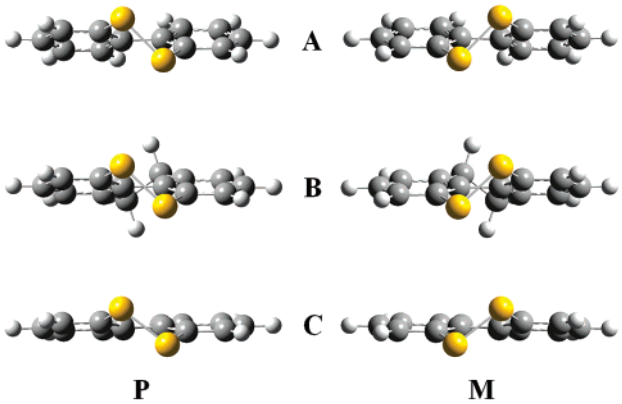


Figure 7. Enantiomer pairs of **A**, **B**, and **C** oxidized variant containing sulfur. P stands for plus and M stands for minus designation of helical chirality.

Table 1. Geometry Data at the B3LYP/6-31G(d) Level of Theory, Gaseous Phase

	A ox			**B** ox			**C** ox		
oxidized X	O	S	Se	O	S	Se	O	S	Se
R(X−X)	1.481	2.100	2.347	1.488	2.095	2.338	1.490	2.092	2.336
R(X-4)	1.382	1.790	1.925	1.384	1.790	1.924	1.382	1.786	1.920
R(4−4a)	1.406	1.415	1.413	1.398	1.414	1.414	1.407	1.427	1.427
R(4a-4b)	1.469	1.488	1.492	1.453	1.489	1.494	1.424	1.460	1.466
R(4a-10a)	1.399	1.406	1.407	1.399	1.415	1.417	1.412	1.429	1.432
R(9−10)	2.333	2.201	2.275	1.545	1.532	1.532	1.369	1.358	1.357
R(1−2)	1.396	1.395	1.394	1.398	1.393	1.392	1.384	1.378	1.377
A(X-4−4a)	119.2	120.3	120.4	118.9	120.9	121.4	118.9	121.4	122.3
A(4−4a-4b)	116.4	123.2	124.3	117.6	123.8	124.8	118.4	124.5	125.6
A(4−4a-10a)	117.8	117.1	117.1	120.7	117.6	117.3	120.5	116.7	116.1
A(10a-4a-4b)	125.8	119.7	118.7	121.7	118.6	117.9	121.0	118.7	118.3
D(X-4−4a-4b)	3.742	−0.777	−2.466	5.123	4.232	3.962	8.450	11.28	11.36
D(4−4a-4b-5)	18.36	32.71	40.01	15.07	24.89	30.16	9.305	13.69	18.45
D(10a-4a-4b-8a)	19.24	32.18	38.43	14.60	22.81	26.70	4.076	8.822	11.80
D(X-4−4a-3)	−175.0	−177.7	179.9	−174.1	−175.5	−177.3	−171.9	−171.2	−172.9
D(4−4a-4b-10a)	179.6	−179.7	−179.2	−179.8	−179.0	−178.3	−177.4	−177.6	−176.7
	A re			**B** re			**C** re		
reduced **X**	O	S	Se	O	S	Se	O	S	Se
R(X−X)	2.787	4.249	3.497	2.621	3.477	3.420	2.583	2.967	3.081
R(X-4)	1.365	1.786	1.933	1.386	1.796	1.927	1.398	1.813	1.944
R(4−4a)	1.417	1.412	1.408	1.409	1.413	1.410	1.421	1.429	1.425
R(4a-4b)	1.490	1.497	1.494	1.494	1.489	1.495	1.470	1.464	1.462
R(4a-10a)	1.409	1.405	1.402	1.420	1.418	1.419	1.435	1.434	1.431
R(9−10)	2.521	3.397	2.889	1.533	1.541	1.541	1.356	1.356	1.357
R(1−2)	1.398	1.396	1.393	1.393	1.395	1.396	1.382	1.379	1.379
A(X-4−4a)	123.9	123.9	121.4	118.3	120.1	121.0	121.7	123.1	123.2
A(4−4a-4b)	123.0	122.7	122.9	125.6	125.0	125.4	126.1	126.0	125.7
A(4−4a-10a)	117.7	118.4	118.2	116.0	117.7	117.8	114.9	116.2	116.4
A(10a-4a-4b)	119.2	118.9	118.9	118.4	117.1	116.7	119.0	117.7	117.8
D(X-4−4a-4b)	−2.728	−1.304	−2.167	4.610	11.44	12.57	7.602	16.20	17.22
D(4−4a-4b-5)	52.48	89.49	69.45	35.67	48.34	47.62	23.34	29.18	30.55
D(10a-4a-4b-8a)	47.28	85.80	64.71	28.39	37.82	37.36	14.99	20.53	21.09
D(X-4−4a-3)	−177.4	−179.8	178.7	−178.2	−173.5	−171.0	−179.2	−171.6	−170.4
D(4−4a-4b-10a)	−175.8	−178.2	−177.6	−177.6	−174.5	−174.9	−176.5	−175.7	−175.3

carbohydrates **Z**	R(mon-mon)
F	1.557
OH	1.589

Table 2. Geometry Data at the B3LYP/6-31G(d) Level of Theory, Aqueous Solution

	DDT ox			HDT ox			PDT ox		
oxidized **X**	O	S	Se	O	S	Se	O	S	Se
R(X−X)	1.485	2.102	2.351	1.493	2.097	2.343	1.496	2.095	2.340
R(X-4)	1.388	1.793	1.928	1.390	1.792	1.927	1.387	1.788	1.923
R(4−4a)	1.406	1.415	1.414	1.398	1.415	1.414	1.407	1.428	1.428
R(4a-4b)	1.470	1.489	1.492	1.454	1.489	1.494	1.425	1.461	1.466
R(4a-10a)	1.400	1.408	1.408	1.400	1.416	1.418	1.413	1.431	1.433
R(9−10)	2.331	2.198	2.273	1.545	1.532	1.532	1.370	1.360	1.358
R(1−2)	1.397	1.396	1.395	1.399	1.394	1.394	1.385	1.380	1.379
A(X-4−4a)	119.0	120.3	120.4	118.5	120.8	121.5	118.5	121.4	122.3
A(4−4a-4b)	116.5	123.3	124.3	117.8	123.9	124.9	118.6	124.7	125.7
A(4−4a-10a)	117.7	117.0	117.0	120.6	117.5	117.2	120.4	116.6	116.0
A(10a-4a-4b)	125.8	119.7	118.7	121.6	118.6	117.9	121.0	118.7	118.2
D(X-4−4a-4b)	3.975	−0.7411	−2.552	5.343	4.168	3.875	8.697	11.23	11.27
D(4−4a-4b-5)	18.29	32.53	40.00	15.06	24.84	30.16	9.345	13.53	18.41
D(10a-4a-4b-8a)	19.45	32.07	38.30	14.92	22.86	26.72	4.365	8.922	11.86
D(X-4−4a-3)	−175.1	−177.9	179.8	−174.2	−175.6	−177.4	−171.9	−171.3	−172.9
D(4−4a-4b-10a)	179.4	−179.8	−179.1	−179.9	−179.0	−178.3	−177.5	−177.7	−176.7

	DDT re			HDT re			PDT re		
reduced **X**	O	S	Se	O	S	Se	O	S	Se
R(X−X)	2.697	4.353	3.532	2.597	3.470	3.421	2.557	3.004	3.104
R(X-4)	1.368	1.792	1.935	1.379	1.806	1.930	1.398	1.814	1.946
R(4−4a)	1.417	1.411	1.408	1.412	1.419	1.411	1.424	1.430	1.425
R(4a-4b)	1.491	1.498	1.495	1.495	1.490	1.495	1.472	1.465	1.462
R(4a-10a)	1.409	1.406	1.403	1.420	1.418	1.419	1.437	1.435	1.432
R(9−10)	2.407	3.418	2.921	1.534	1.541	1.541	1.356	1.357	1.358
R(1−2)	1.398	1.397	1.395	1.396	1.394	1.397	1.382	1.381	1.381
A(X-4−4a)	123.3	124.2	121.2	118.1	124.3	120.9	121.9	122.9	122.9
A(4−4a-4b)	123.4	123.2	122.9	125.2	125.4	125.4	126.6	126.1	125.8
A(4−4a-10a)	117.5	118.2	118.1	116.5	117.8	117.7	114.8	116.1	116.3
A(10a-4a-4b)	118.9	118.5	119.0	118.2	116.6	116.7	118.6	117.6	117.8
D(X-4−4a-4b)	−3.107	−1.796	−2.190	5.142	16.61	12.89	7.978	16.44	17.70
D(4−4a-4b-5)	48.16	92.14	70.52	35.40	48.14	47.39	21.61	29.82	30.57
D(10a-4a-4b-8a)	43.38	86.33	65.81	28.26	37.73	37.35	14.18	20.94	21.37
D(X-4−4a-3)	−177.1	−179.2	178.6	−178.0	−167.5	−171.1	−179.1	−171.2	−170.1
D(4−4a-4b-10a)	−175.4	−177.2	−177.6	−177.7	−175.0	−175.0	−177.4	−175.5	−175.4

carbohydrates **Z**	R(mon-mon)
F	1.558
OH	1.589

METHODS

Ab Initio Methods. Initial geometry optimizations were performed at the HF/3-21G level of theory, and final geometry optimizations with frequency calculations were carried out at the B3LYP/6-31G(d) level of theory using the Gaussian03 program package.[13] The latter geometries were reoptimized on the same level using PCM water solvation model.

To better understand the reaction mechanism, two potential energy surfaces were generated on the UHF/3-21G* level of theory for singlet and triplet state H abstractions for compound **B** and X = S. These were the result of rigid scans in which the initial geometry was loosely optimized at the same level with the distance between the S atom and the C atom of the deoxyribose kept frozen (approximately 3.1 Å, similar to that in DNA). In the rigid scans the energy was plotted as a function of the two H atoms' distance from the corresponding S atom (see D1 and D2 in Figure 4). To ensure convergence the scf=xqc option of Gaussian was used requiring the quadratic convergence method of energy calculation in case the regular method fails, making the calculation much longer and resulting in the choice of the relatively low level of theory.

Docking Studies. Docking studies were carried out using the AutoDock 3.05 software.[14] The initial structure of the DNA was taken from the PDB database (accession number: 102D). After removing the ligand (propamidine) and water used in crystallization, the input files were generated with AutoDockTools.[15]

During the docking procedure, DNA was kept rigid, and rigid-body movement (translation and quaternation) was allowed for the ligand at geometries obtained from the ab initio optimization. The grid box ($60 \times 60 \times 40$ grid points) was centered on the DNA (Figure 5), and the lattice point distance was set to 0.375 Å.

Lennard-Jones potentials 12−10 and 12−6 were used to model H-bonds and van der Waals interactions, respectively. In calculations of the electrostatic grid map, the distance-dependent dielectric constant of Mehler and Solmayer was utilized.[16] Using the distance-dependent form in the grid calculation the screening effect of the solvent was taken into consideration. During the docking procedure, the Lamarckian Genetic Algorithm with the pseudo-Solis and Wets method was used with 250 individuals in the population. The stopping criterion was defined by the total number of energy evaluations, which was set to 1×10^7. The translation step was set to 0.5 Å/step, and in quaternation step 1.0 degree/step was used. The docking procedure was performed 50 times.

Molecular Dynamics Simulation on the Complexes Obtained. The explicit water molecular dynamics simulation with particle-mesh Ewald (PME) electrostatic summation was carried out using the AMBER8 program package.[17] The periodic box of the system consists of the DNA, the ligand, 22 Na+ counterions, and 3841 TIP3P water molecules (12 329 atoms). The ff99 force field[18] and General Amber Force field was used for the DNA and the ligand, respectively. The charges of the ligand were calculated using the

Table 3. Direct Energetic Data in Hartrees: HF Energy at 0 K, B3LYP Energy at 0 K, ZPE Corrected B3LYP Energy, Energy, Enthalpy, and Gibbs Free Energy at $T = 298.15$ K

	X	E(0,HF)	E(0)	E(ZPE)	E(T)	H(T)	G(T)
				a. Gaseous Phase			
	O	−605.327093	−612.4505684	−612.282937	−612.273338	−612.272394	−612.318053
A ox	S	−1247.64382	−1258.480559	−1258.316932	−1258.306352	−1258.305408	−1258.352921
	Se	−5233.525874	−5260.882117	−5260.719729	−5260.708461	−5260.707517	−5260.75767
	O	−681.815634	−689.8830466	−689.679416	−689.668623	−689.667678	−689.71502
B ox	S	−1324.132961	−1335.915308	−1335.71558	−1335.703853	−1335.702909	−1335.753339
	Se	−5310.013477	−5338.315593	−5338.116996	−5338.104604	−5338.103659	−5338.156666
	O	−680.652242	−688.6801159	−688.499781	−688.489518	−688.488574	−688.535411
C ox	S	−1322.965669	−1334.710524	−1334.534179	−1334.522924	−1334.52198	−1334.571438
	Se	−5308.844671	−5337.10903	−5336.933867	−5336.921921	−5336.920976	−5336.973113
	O	−606.579915	−613.7400494	−613.549464	−613.538521	−613.537577	−613.586097
A re	S	−1248.791005	−1259.672077	−1259.492676	−1259.479881	−1259.478937	−1259.532702
	Se	−5234.652095	−5262.054234	−5261.87729	−5261.863945	−5261.863	−5261.919973
	O	−683.065492	−691.1709569	−690.944055	−690.93219	−690.931246	−690.981324
B re	S	−1325.263093	−1337.091982	−1336.876237	−1336.862721	−1336.861777	−1336.915952
	Se	−5311.120906	−5339.477737	−5339.263911	−5339.250068	−5339.249124	−5339.305233
	O	−681.895279	−689.9617354	−689.758212	−689.746795	−689.745851	−689.795133
C re	S	−1324.090439	−1335.878656	−1335.686712	−1335.673429	−1335.672485	−1335.72666
	Se	−5309.94875	−5338.267622	−5338.078021	−5338.064281	−5338.063337	−5338.119821
mono	F	−563.496828	−569.4698028	−569.34662	−569.33843	−569.337486	−569.381033
	OH	−491.842002	−497.4014062	−497.24238	−497.232984	−497.23204	−497.276946
di	F	−1137.736611	−1137.736611	−1137.510202	−1137.494011	−1137.493067	−1137.554698
	OH	−993.6097398	−993.6097398	−993.30967	−993.291804	−993.29086	−993.354042
				b. Aqueous Solution			
	O	−605.327093	−612.4614743	−612.295108	−612.285488	−612.284544	−612.330225
A ox	S	−1247.64382	−1258.491108	−1258.328668	−1258.318049	−1258.317105	−1258.364685
	Se	−5233.525874	−5260.892618	−5260.73141	−5260.720094	−5260.71915	−5260.769409
	O	−681.815634	−689.8930704	−689.690691	−689.679869	−689.678924	−689.726957
B ox	S	−1324.132961	−1335.925116	−1335.726534	−1335.714766	−1335.713822	−1335.764324
	Se	−5310.013477	−5338.325387	−5338.127924	−5338.115483	−5338.114539	−5338.167648
	O	−680.652242	−688.691124	−688.51214	−688.50185	−688.500906	−688.54777
C ox	S	−1322.965669	−1334.721622	−1334.546512	−1334.53522	−1334.534276	−1334.583788
	Se	−5308.844671	−5337.120167	−5336.946219	−5336.934225	−5336.93328	−5336.984864
	O	−606.579915	−613.7577985	−613.569481	−613.558691	−613.557747	−613.605995
A re	S	−1248.791005	−1259.685169	−1259.50786	−1259.494848	−1259.493904	−1259.549829
	Se	−5234.652095	−5262.064963	−5261.889344	−5261.876771	−5261.875827	−5261.930212
	O	−683.065492	−691.1921316	−690.967507	−690.955689	−690.954745	−691.004734
B re	S	−1325.263093	−1337.106855	−1336.892927	−1336.879394	−1336.87845	−1336.932661
	Se	−5311.120906	−5339.488408	−5339.275549	−5339.261672	−5339.260728	−5339.316872
	O	−681.895279	−689.9815732	−689.780834	−689.769367	−689.768423	−689.818197
C re	S	−1324.090439	−1335.896478	−1335.707102	−1335.693614	−1335.69267	−1335.747469
	Se	−5309.94875	−5338.278932	−5338.090799	−5338.077754	−5338.07681	−5338.131412
mono	F	−563.496828	−569.47943	−569.357201	−569.34902	−569.348076	−569.391178
	OH	−491.842002	−497.4312792	−497.275353	−497.265883	−497.264939	−497.310159
di	F	−1137.736611	−1137.751288	−1137.526396	−1137.510174	−1137.50923	−1137.570268
	OH	−993.6097398	−993.6563935	−993.36133	−993.3433	−993.342356	−993.405936

multiconformer, multiorientation RESP calculation implemented in the R.E.D. II software,[19] with the Gaussian03 quantum mechanics program.[13]

The initial complex was first minimized with the DNA and ligand restrained (force constant set to 500 kcal/mol$Å^2$) in 2000 cycles. After this minimization, the whole system was set free to move, and the complex was minimized in 3000 cycles. The system was heated from 0 to 300 K in 20 ps with a time step of 0.0005 ps (positional restraints for the DNA and ligand atoms, with a 10 kcal/mol$Å^2$ force constant and constant volume molecular dynamics) and equilibrated for 2 ns at 300 K (constant pressure molecular dynamics). The time step was set to 0.002 ps, and the SHAKE algorithm was used to constrain all bonds involving hydrogens.

RESULTS AND DISCUSSION

Evaluation of Ab Initio Results. As it was mentioned earlier, several reactions were studied in this work. In Figure 3, it is shown that there are three possibilities as far as the basic aromatic structure is concerned depending on the nature of **Y**. The different structures are shown in Figure 6. If $-(\mathbf{Y}-Y)- = -(CH{=}CH)-$, then the basic structure is phenanthrene, which will be referred to as structure **C**. With $-(CH_2-CH_2)-$, the molecule is partially hydrogenized, and it will be designated as structure **B**. Finally, the −(H H)− does not have a middle ring and will be referred to as structure **A**. These compounds have two variants: the oxidized (ox) on the left and the reduced (red) form on the right of Figure 3.

Note that these aromatic compounds have helical chirality, which is not important when energetics are calculated, but is essential during docking since DNA is also a chiral structure. In Figure 7, all the enantiomer pairs are shown with the oxidized forms of **A**, **B**, and **C** structures containing sulfur. Enantiomers will be marked in square brackets after the given structure if essential to the discussion (**C[M]**, **A-[P]**, etc.).

Table 4. Calculated Changes of Energy, Enthalpy, Gibbs Free Energy (kJ/mol), and Entropy (J/mol·K) for the Redox Reaction Described in Figure 3, Computed at the B3LYP/6-31G(d) Level of Theory, $T = 298.15$ K

X	Y	Z	ΔE(T)	ΔH(T)	ΔG(T)	ΔS(T)
			a. Gaseous Phase			
O	**A**	F	−216.17	−218.65	−159.30	−199.03
S	**A**	F	24.47	21.99	72.43	−169.17
Se	**A**	F	71.85	69.37	118.32	−164.17
O	**B**	F	−211.93	−214.41	−154.74	−200.13
S	**B**	F	62.96	60.48	117.50	−191.25
Se	**B**	F	98.15	95.67	154.38	−196.91
O	**C**	F	−195.41	−197.89	−137.46	−202.70
S	**C**	F	84.92	82.44	136.91	−182.69
Se	**C**	F	106.30	103.82	159.26	−185.95
O	**A**	OH	−238.97	−241.45	−179.04	−209.31
S	**A**	OH	1.67	−0.81	52.69	−179.45
Se	**A**	OH	49.04	46.57	98.58	−174.45
O	**B**	OH	−234.73	−237.21	−174.47	−210.41
S	**B**	OH	40.16	37.68	97.77	−201.52
Se	**B**	OH	75.35	72.87	134.64	−207.19
O	**C**	OH	−218.21	−220.69	−157.19	−212.97
S	**C**	OH	62.12	59.64	117.17	−192.97
Se	**C**	OH	83.50	81.02	139.52	−196.22
			b. Aqueous Solution			
O	**A**	F	−224.05	−226.53	−167.20	−199.01
S	**A**	F	29.06	26.58	70.74	−148.13
Se	**A**	F	81.89	79.41	134.65	−185.28
O	**B**	F	−230.92	−233.40	−172.47	−204.39
S	**B**	F	61.01	58.53	114.87	−188.95
Se	**B**	F	109.42	106.94	165.05	−194.88
O	**C**	F	−209.12	−211.60	−153.17	−195.99
S	**C**	F	77.38	74.90	127.09	−175.05
Se	**C**	F	116.41	113.93	172.08	−195.03
O	**A**	OH	−222.48	−224.96	−161.17	−213.92
S	**A**	OH	30.63	28.15	76.76	−163.04
Se	**A**	OH	83.46	80.98	140.67	−200.19
O	**B**	OH	−229.35	−231.83	−166.44	−219.30
S	**B**	OH	62.59	60.11	120.89	−203.87
Se	**B**	OH	111.00	108.52	171.07	−209.80
O	**C**	OH	−207.55	−210.03	−147.15	−210.90
S	**C**	OH	78.95	76.48	133.12	−189.97
Se	**C**	OH	117.98	115.50	178.10	−209.95

The numerical results of the first principle molecular computations can be found in Tables 1−4. Tables 1 and 2 show the geometric data for the B3LYP/6-31G(d) levels of theory, in gaseous phase and in aqueous solution, respectively. Calculations at the HF/3-21G level may be found in the Supporting Information in Table 7. These tables contain data for primarily the aromatic compounds and the length of the bond connecting the two monomers in the carbohydrate dimers.

Table 3 contains all the direct energetic data: HF energies, B3LYP energies, and thermodynamic functions. Table 4 includes the calculated changes of thermodynamic functions in the reactions. Part a. of these tables contain data for the gaseous phase while part b. for aqueous solution.

The changes in reaction enthalpy (Table 4, Figure 8) indicate that within the same aromatic structure, the reactions proceed from being exothermic (in the case of O) to more and more endothermic (i.e. the thermodynamic reactivity decreases ongoing from oxygen to selenium). It is also observed that the reactions with carbohydrates with F substituents have higher reaction enthalpies than those with containing OH. This occurs because of the H-bonds formed in the OH containing dimer. Thus, the fluoro congener of deoxyribose is not a good model in the gaseous phase. However, in the aqueous solution the OH and F models become about equally good (see Table 4 and Figure 9). This finding has far reaching consequences in carbohydrate ring

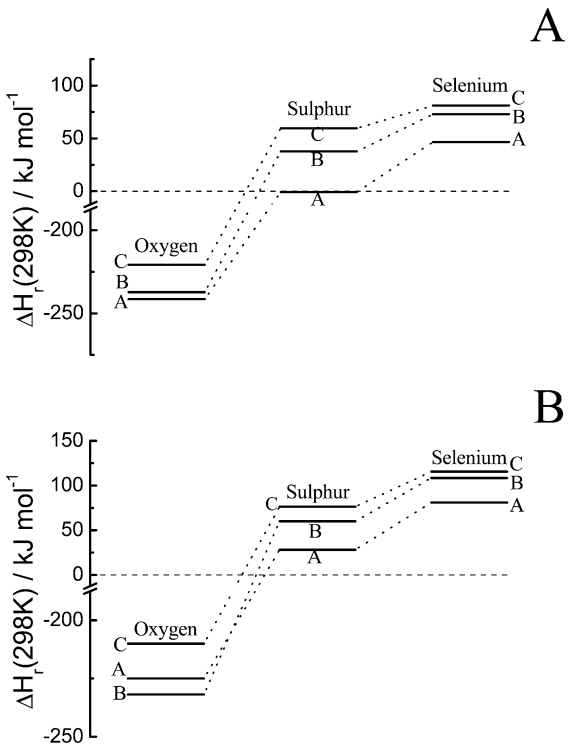


Figure 8. Enthalpy changes during reactions when **Z** = OH. A is in the gaseous phase, and B is in the aqueous solution.

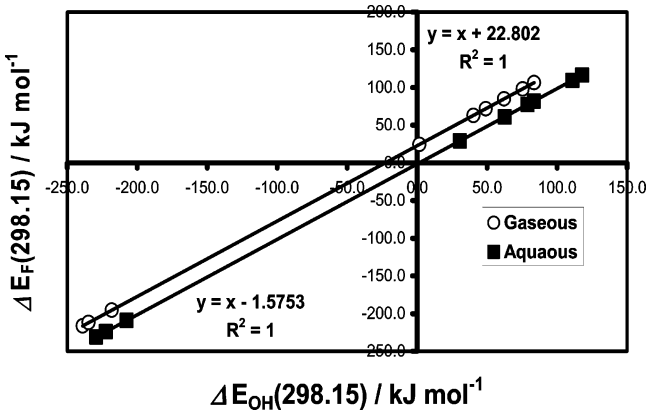


Figure 9. Correlation of relative energies of fluorinated carbohydrate model (ΔE_F) with hydroxyl containing carbohydrates (ΔE_{OH}) in the gaseous and in the aqueous phase.

conformation analysis if all −OH groups are replaced by fluorine in the preoptimization process. If different aromatic structures are compared, the values for structures **A**, **B**, and **C** form a gradatient; **A** is the lowest and **C** is the highest for all three heteroatoms (O, S, Se). This latter fact and also the reactivity discussed above are at least partially related to aromacity and ring strain energetics.

Evaluation of the Docked Structures. During docking studies, only the sulfur containing oxidized structures were examined, since these are the most favorable for our purposes as they are the closest to being thermoneutral in their reaction. The structures obtained from the rigid docking were cluster-analyzed using the estimated ΔG_{bind} as a descriptor with 0.5 RMS tolerance. The lowest energy structures were selected from the most loaded population. The possible S−H distances (Figures 10 (**C[M]**) and 11 (**C[P]**)) were measured and presented in Table 5 after docking.

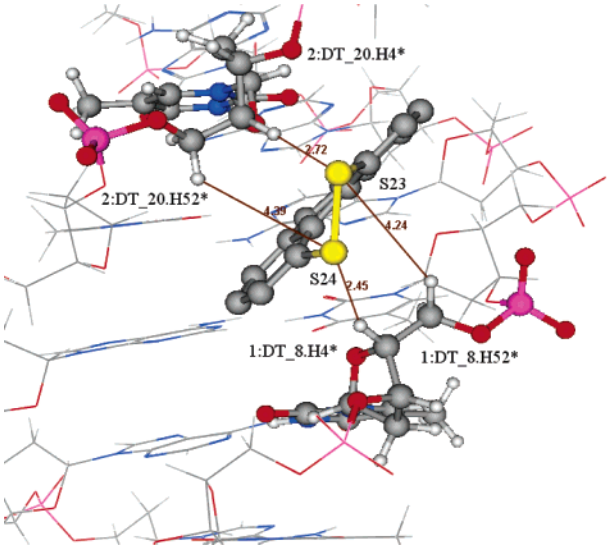


Figure 10. The structure and the measured distances of complex **C[M]** as obtained from docking calculations.

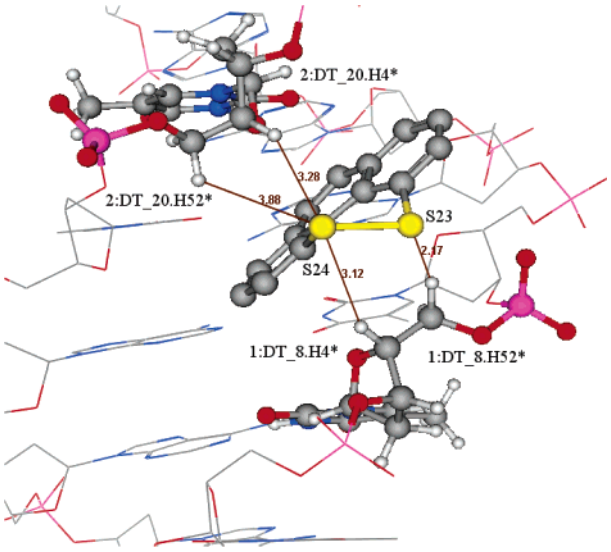


Figure 11. The structure and the measured distances of complex **C[P]** as obtained from docking calculations.

In two cases, **B[M]** and **C[M]**, the lowest distances were measured between the S23−2:DT_20.H4* (2.82 and 2.72 Å) and S24−1:DT_8.H4* (2.27 and 2.45 Å) atoms. These results are in good agreement with the fact that the most reactive hydrogens are those that are attached to the C4 carbon atoms in deoxyribose. For remaining enantiomers another possible active conformation was determined.

Although two possible active conformations have been described, we have to keep in mind that during the docking procedure the DNA was kept rigid, and for the ligand only rigid body movement was allowed. The evolved complexes could change, when flexibility is taken into account; therefore, we performed explicit water molecular dynamics simulation for both complexes. The two enantiomers of compound C was selected, particular because in the case of **C[M]** 84% of the docked structures was determined at the same target molecule cavity, in comparison to the enantiomer **C[P]** which was only 6%.

After minimization, the measured distances were not changed much; however, during the molecular dynamics simulation, the two complexes equilibrated in two different ways. Changes in the previously determined distances during the whole simulation are depicted in Figures 12 (**C[M]**) and 13 (**C[P]**).

Table 5. Measured Distances in Å between S and Nonpolar H Atoms

	noc[a]	distances between atoms in DNA	S23	S24
A[P]	8	1:DT_8.H4*		3.7
		1:DT_8.H52*	2.76	
		2:DT_20.H4*	3.23	
		2:DT_20.H52*		2.53
A[M]	9	1:DT_8.H4*		4.03
		1:DT_8.H52*	5.69	
		2:DT_20.H4*	4.14	
		2:DT_20.H52*		3.91
B[P]	7	1:DT_8.H4*		4.31
		1:DT_8.H52*	3.24	
		2:DT_20.H4*	3.09	
		2:DT_20.H52*		3.24
B[M]	6	1:DT_8.H4*		2.27
		1:DT_8.H52*	4.54	
		2:DT_20.H4*	2.82	
		2:DT_20.H52*		4.24
C[P]	3	1:DT_8.H4*		3.12
		1:DT_8.H52*	2.17	
		2:DT_20.H4*		3.28
		2:DT_20.H52*		3.88
C[M]	42	1:DT_8.H4*		2.45
		1:DT_8.H52*	4.24	
		2:DT_20.H4*	2.72	
		2:DT_20.H52*		4.39

[a] noc − number of conformers in the most loaded population.

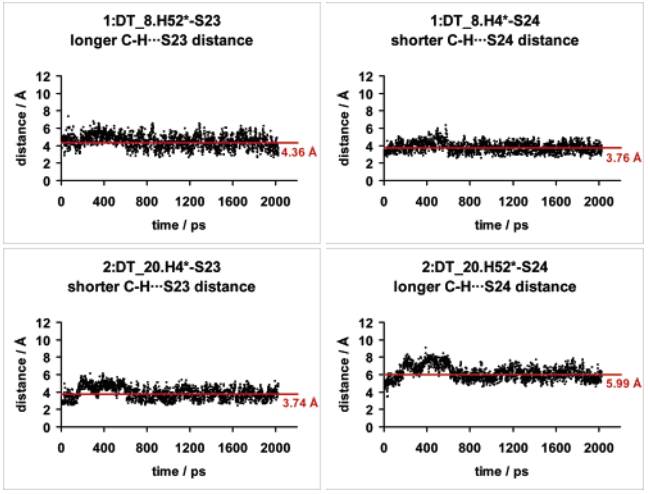


Figure 12. Changes to the measured distances of **C[M]** (the well fitted enantiomer). The horizontal lines mark the average distances calculated from the last 1300 steps.

Based on these results, we can conclude that in case of **C[M]**, a more stable complex could evolve, as compared to **C[P]**, during the examined time interval. The distance range in the case of **C[M]** is a well-defined range, contrary to this, in case of **C[M]** the distances fluctuate in a greater range, and from the initial values it equilibrated to a +4 Å greater distance (Table 6).

It should be emphasized that the **C[P]** ligand was rotated 90° from the initial position (Figure 14). In contrast, the ligand was retained in the groove of the DNA for **C[M]** (Figure 15).

Mechanistic Considerations. According to Figure 1 in each of the hydrogen abstraction steps two odd electrons are

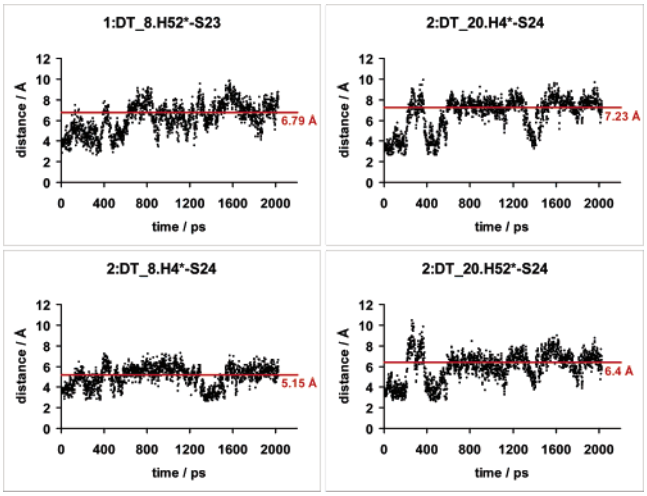


Figure 13. Changes to the measured distances of **C[P]** (the badly fitted enantiomer). The horizontal lines mark the average distances calculated from the last 1300 steps.

Table 6. Descriptive Statistics of the Distances Using the Last 1300 ps of the Trajectory

	C[P]			
	1:DT_8. H52*-S23	1:DT_8. H4*-S24	2:DT_20. H4*-S24	2:DT_20. H52*-S24
MIN	3.48	2.65	3.29	3.28
MAX	9.89	7.32	9.71	9.03
MEAN ± SD	6.79 ± 1.09	5.15 ± 0.97	7.23 ± 1.07	6.40 ± 0.94
	C[M]			
	1:DT_8. H52*-S23	2:DT_20. H4*-S23	1:DT_8. H4*-S24	2:DT_20. H52*-S24
MIN	2.64	2.45	2.56	4.60
MAX	6.58	5.44	5.04	7.86
MEAN± S D	4.36 ± 0.75	3.74 ± 0.60	3.76 ± 0.49	5.99 ± 0.60

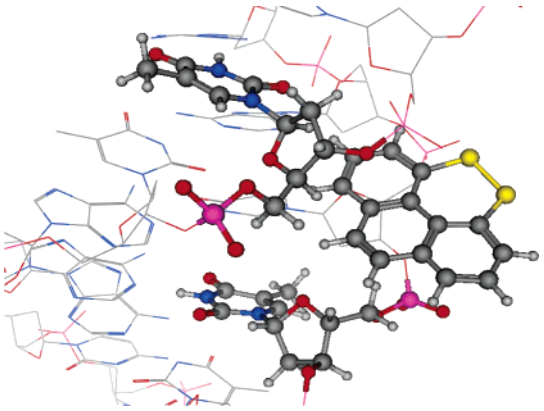


Figure 14. The position of the ligand in complex **C[P]** taken from the last frame of the dynamics simulation.

created from the breaking of electron pairs. Such two odd electrons may be formed, initially, as of singlet multiplicity, but a triplet combination usually possesses a considerably longer lifetime. Consequently some intersystem crossing (ISC) between singlet and triplet potential energy manifolds will necessarily be part of the mechanism. Figure 16 summarizes a putative reaction scheme in which two stepwise (sm_1 and sm_2) and one concerted (cm) mechanisms are shown.

The singlet and triplet potential energy surfaces are shown in Figure 17 (parts A and B, respectively). The difference

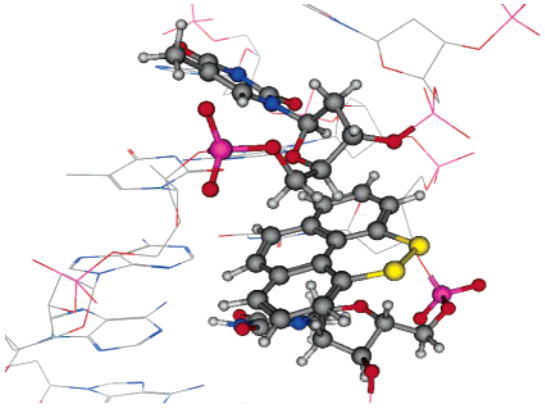


Figure 15. The position of the ligand in complex **C[M]**, taken from the last frame of the dynamics simulation.

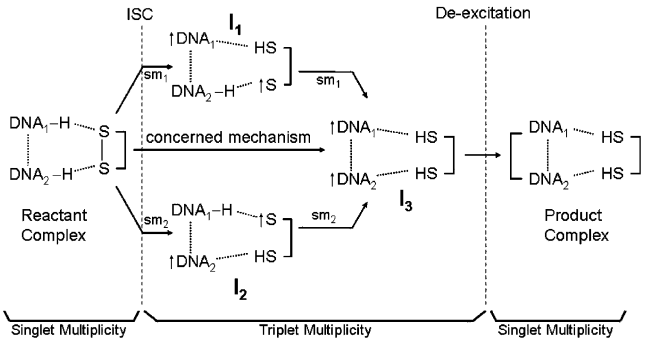


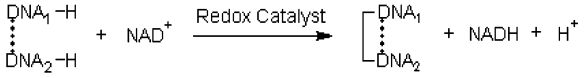
Figure 16. Two stepwise mechanisms (sm_1 and sm_2) and one concerted mechanism (cm) of the proposed double hydrogen abstraction from the two helices of DNA. The S→T intersystem crossing (ISC) and the T→S de-excitations are indicated by vertical dashed lines. The three intermediates (I_1, I_2, I_3) are all of triplet multiplicity.

of the above two surfaces are shown in Figure 17C. Here the solid line marked by 0 is the line of the crossing of the two surfaces. Intersystem crossing may be expected to occur in the vicinity of this 0 line. The crossover occurs at different places for the two stepwise (sm_1 and sm_2) and the concerted (cm) mechanism.

Indeed, in Figure 17B we can recognize three minima: I_1, I_2, and I_3; the former two are associated with the two stepwise mechanisms (sm_1 and sm_2), while the last one corresponds to the concerted mechanism (cm).

CONCLUSIONS

The thermochemistry of oxidative coupling of two DNA strands were modeled at several levels of theory. On this basis, it may be concluded that redox catalysts may be found which will enhance the formation of a covalent linkages between the two strands of DNA bases:



Such a redox catalyst would be a potential anticancer drug candidate as a cell division inhibitor.

ACKNOWLEDGMENT

This research was supported by the National Science Fund of Hungary (OTKA T046861 and F037648). One of the authors (I.G.C.) wishes to thank the Ministry of Education

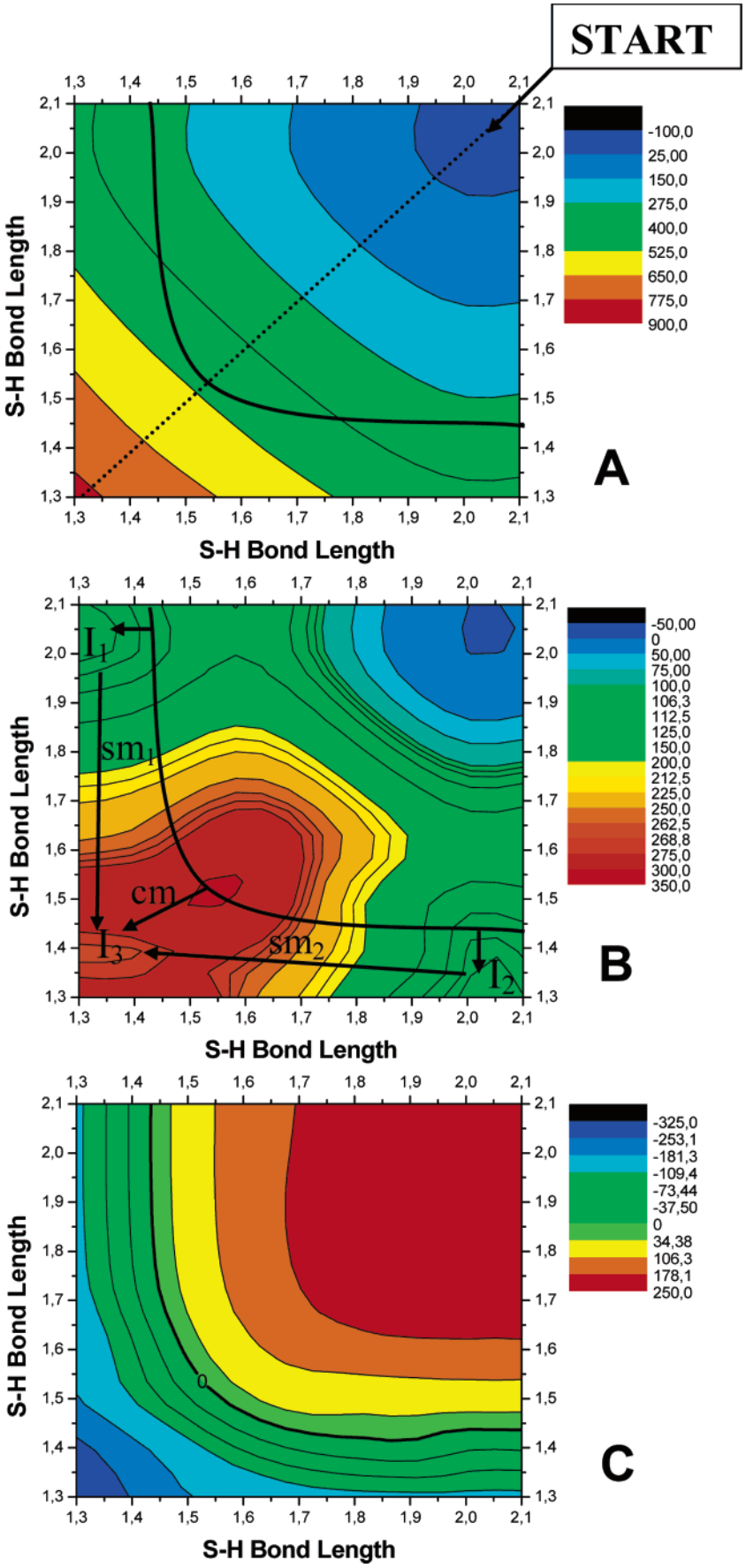


Figure 17. Potential energy surfaces, of singlet state (A), triplet state reaction (B), and their difference (C, triplet minus singlet) in kJ/mol. The solid line marked by 0 is the line of intersystem crossing (ISC). The bond length is measured in Å. The three intermediates ($I_1 \approx I_2 < I_3$) are all of triplet multiplicity.

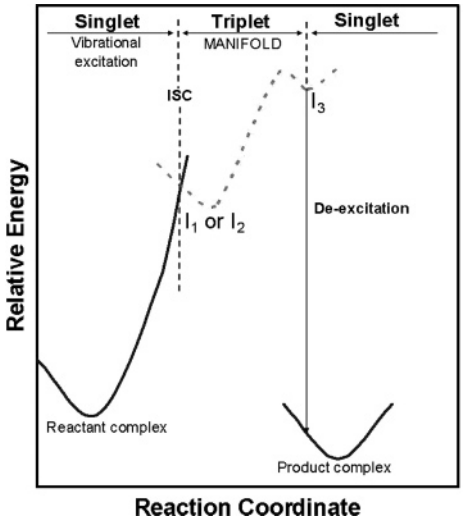


Figure 18. A schematic energy profile for the putative stepwise mechanism of the double hydrogen abstraction. The intersystem crossing (ISC) and the de-excitation are marked by vertical dashed lines. The calculation was carried out at the UHF/3-21G* level of theory.

for a Szent-Györgyi Visiting Professorship. One of the authors (B.J.) is grateful to Centenarium Foundation of Gedeon Richter Ltd. for support. We acknowledge the help of the University of Szeged (MU-00094/2002) for the donation of computational time. The authors would also like to express their gratitude to M. Labadi for technical support and Suzanne Lau for helpful discussions.

Supporting Information Available: Geometry data at the HF/3-21G level of theory (Table 7). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES AND NOTES

(1) Rafel, M.; Cervantes, F.; Beltraìn, J. M.; Zuazu, F.; Nieto, L. H.; Rayoìn, C.; Talavera, J. G.; Montserrat, E. Deoxycoformycin in the treatment of patients with hairy cell leukemia: Results of a Spanish collaborative study of 80 patients. *Cancer* **2000**, *88*, 352−357.

(2) Relling, M. V.; Hancock, M. L.; Boyett, J. M.; Pui, C. H.; Evans, W. E. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. *J. Nat. Cancer Inst.* **1999**, *91*, 2001−2008.

(3) Farber, S.; Diamond, L. K.; Mercer, R. D.; Sylvester, R. F.; Wolf, V. A. Tissue-specific Expression of Functional Isoforms of Mouse Folylpoly-γ-glutamate Synthetase: A Basis for Targeting Folate Antimetabolites. *J. Med. Chem.* **1948**, *238*, 787−793.

(4) Sunters, A.; Springerc, C. J.; Bagshawea, K. D.; Souhamib, R. L.; Hartley, J. A. The cytotoxicity, DNA crosslinking ability and DNA sequence selectivity of the aniline mustards melphalan, chlorambucil and 4-[bis(2-chloroethyl)amino] benzoic acid. *Biochem. Pharmacol.* **1992**, *44*, 59−64.

(5) Wang, Y. D.; Dziegielewski, J.; Wurtz, N. R.; Dziegielewska, B.; Dervan, P. B.; Beerman, T. A. DNA crosslinking and biological activity of a hairpin polyamide-chlorambucil conjugate. *Nucl. Acids Res.* **2003**, 31, 1208−1215.

(6) Milkevitch, M.; Storrie, H.; Brauns, E., Brewer, K. J.; Shirley, B. W. A New Class of Supramolecular, Mixed-Metal DNA-Binding Agents: The Interaction of RuII, PtII and OsII, PtII Bimetallic Complexes with DNA. *Inorg. Chem.* **1997**, *36*, 4534−4538.

(7) McGregor, T. D.; Hegmans, A.; Káspárková, J.; Neplechová, K.; Nováková, O.; Penázová, H.; Vrána, O.; Brabec, V.; Farrell, N. A comparison of DNA binding profiles of dinuclear platinum compounds with polyamine linkers and the trinuclear platinum phase II clinical agent BBR3464. *J. Biol. Inorg. Chem.* **2002**, *7*, 397−404.

(8) Gewirtz, D. A. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem. Pharmacol.* **1999**, *57*, 727−741.

(9) Budman, D. R. Vinorelbine (Navelbine): A third-generation vinca alkaloid. *Cancer Invest.* **1997**, *15*, 475−490.

(10) Druker, B. J.; Lydon, N. B.Lessons learned from the development of an Abl tyrosine kinase inhibitor for chronic myelogenous leukaemia. *J. Clin. Invest.* **2000**, *105*, 3−7.

(11) Gewirtz, D. A. A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin. *Biochem. Pharmacol.* **1999**, *57*, 727−741.

(12) Voet, D.; Voet, J. G. In *Biochemistry*; Wiley: New York, 1990; Chapter 19, pp 509−514.

(13) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, revision B01*; Gaussian, Inc.: Pittsburgh, PA, 2003.

(14) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662.

(15) Huey, R.; Sanner, M. F. Copyright: Sanner, M. TSRI 2000.

(16) Mehler, E. L.; Solmajer, T. Electrostatic effects in proteins: comparison of dielectric and charge models. *Protein Eng.* **1991**, *4*, 903−910.

(17) Case, D. A.; Darden, T. A.; Cheatham, T. E., III.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. *AMBER 8*; University of California: San Francisco, CA, 2004.

(18) Wang, J.; Cieplak, P.; Kollman, P. A. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(19) Zaffran, T.; Cieplak, P.; Dupradeau, F.-Y. *RESP ESP charge Derive (RED) II*; The Scripps Reasearch Institute: La Jolla, CA, U.S.A., 2005.

CI060123O