Investigation of Thioformacetal Modified Nucleic Acids by Quantum Mechanics, Molecular Mechanics, and Molecular Dynamics

James M. Veal* and Frank. K. Brown

Contribution from the Division of Molecular Sciences, Glaxo Research Institute, 5 Moore Drive, Research Triangle Park, North Carolina 27709

Received November 7, 1994[⊗]

Abstract: The stereoelectronic effects of replacement of the phosphodiester linker in nucleic acids by a thioformacetal linker have been investigated by computational techniques. Ab initio studies of the potential energy surfaces of thioformacetal, methyl propyl sulfide, and methyl 2-methoxy ethyl sulfide were initially conducted at the 6-31G* level of theory. These studies show that COCS and OCSC dihedral angles have strong gauche preferences (3.1 and 1.3 kcal/mol). In contrast, CCCS, CCSC, and OCCS dihedral angles are observed to have global minima which are trans conformations. The results of the ab initio studies were subsequently used to obtain potential energy force field parameters for use in molecular mechanics and molecular dynamics studies. The molecular mechanics studies were centered on the pseudorotational pathways for 3′-thio, 2′-deoxyribose and 5′-thio, 2′-deoxyribose nucleosides (dA and dT). The results of the molecular mechanics and quantum mechanics studies lead to the conclusion that when sulfur occupies the O5′ position, it is generally unfavorable due to both steric and electronic effects. However, the opposite orientation substitution, in which sulfur is incorporated at the O3′ position, is found to be detrimental in the case of a B form type nucleic acid conformation but generally well tolerated in the case of an A form type nucleic acid conformation. These results have been subsequently confirmed in part by experimental studies. A second experimental result, namely that the 3′*S*-thioformacetal linker could adopt a (+g, +g) conformation in oligomers, was investigated using thymine dimers and high temperature molecular dynamics simulations. Such a conformer is in fact found as a low energy conformation and is characterized by the backbone dihedral angles ϵ, ζ, α, β, and γ having a unique and previously unobserved +g, +g, +g, t, t pattern of torsional values.

The synthesis and characterization of modified nucleic acids is currently an area of intense interest due to the potential of these agents to exert therapeutic effects via antisense or antigene mechanisms.[1,2] Replacement of the phosphodiester group of the nucleic acid backbone with alternative groups is one attractive modification strategy. Substitution of the phosphodiester linker can confer resistance to nucleases, and, in the case of replacement by uncharged linkers, improve lipophilicity characteristics. A number of such substitutions have now been reported.[3−14]

We have investigated the properties of modified nucleic acids containing formacetal and thioformacetal linkers substituted for the normal phosphodiester linkage[15−22] (Figure 1). We are interested specifically in the effect these substitutions have on conformational properties of nucleic acid oligomers and how

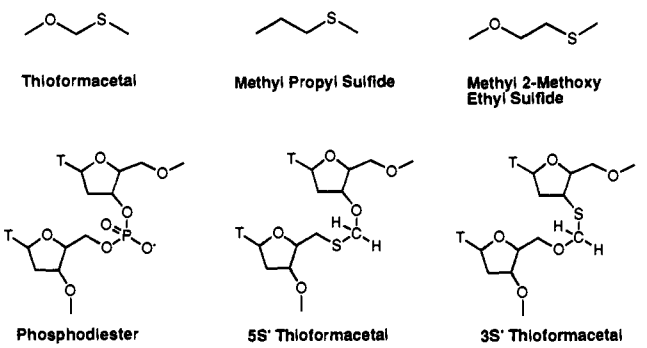


Figure 1. Top: Structures of small molecules studied by quantum mechanics and molecular mechanics. Bottom: Nucleotide dimers showing the native phosphodiester linkage and the thioformacetal linkage in either the 5′ or 3′ orientation.

those effects correlate with DNA−DNA homoduplex and DNA−RNA heteroduplex stability. Duplex stability can experimentally be monitored via thermal denaturation (Tm) measurements. NMR and computational studies on the effects

⊗ Abstract published in *Advance ACS Abstracts*, February 1, 1995.

(1) Uhlman, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 544.
(2) Zon, G. *Pharm. Res.* **1988**, *5*, 539.
(3) Eckstein, F. *Annu. Rev. Biochem.* **1985**, *54*, 367.
(4) Stec, W. J. *Acc. Chem. Res.* **1983**, *16*, 411.
(5) Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497.
(6) Loke, S. L.; Stein, C. A.; Zhang, X. H.; Mori, K.; Nakanishi, M.; Subasinghe, C.; Cohen, J. S.; Neckers, L. M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3474.
(7) Tittensor, J. R. *J. Chem. Soc. (C)* **1971**, 2656.
(8) Gait, M. J.; Jones, A. S.; Walker, R. T. *J. Chem. Soc., Perkin Trans. 1* **1974**, 1684.
(9) Ogilvie, K. K.; Cormier, J. F. *Tetrahedron Lett.* **1985**, *26*, 4159.
(10) Schneider, K. C.; Benner, S. A. *Tetrahedron Lett.* **1990**, *31*, 335.
(11) Wang, H.; Weller D. D. *Tetrahedron Lett.* **1991**, *32*, 7385.
(12) Vasseur, J.-J.; Debart, F.; Sanghvi, Y. S.; Cook, P. D. *J. Am. Chem. Soc.* **1992**, *114*, 4006.
(13) Reynolds, R. C.; Crooks, P. A.; Maddry, J. A.; Akhtar, M. S.; Montgomery, J. A.; Secrist, J. A. *J. Org. Chem.* **1992**, *57*, 2983.
(14) Huie, E. M.; Kirshenbaum, M. R.; Trainor, G. *J. Org. Chem.* **1992**, *57*, 4569.
(15) Matteucci, M.; Lin, K.-Y.; Butcher, S.; Moulds, C. *J. Am. Chem. Soc.* **1991**, *113*, 7767.
(16) Gao, X.; Brown, F. K.; Jeffs, P.; Bischofberger, N.; Lin K.-Y.; Pipe, A. J.; Noble, S. A. *Biochemistry* **1992**, *31*, 6228.
(17) Matteucci, M. *Tetrahedron Lett.* **1990**, *31*, 2385.
(18) Veeneman, G. H.; van der Marel, G. A.; van den Elst, H.; van Boom, J. H. *Tetrahedron* **1991**, *47*, 1547.
(19) Veal, J. M.; Gao, L. X.; Brown, F. K. *J. Am. Chem. Soc.* **1993**, *115*, 7139.
(20) Brown, F. K.; Veal, J. M *International J. of Quantum Chemistry: Quantum Biology Symposium* **1993**, *20*, 38.
(21) Gao, L. X.; Jeffs, P. K. *J. Biol. NMR* **1994**, *4*, 17.
(22) Jones, R. J.; Lin, K-Y.; Milligan, J. F.; Wadwani, S.; Matteuci, M. D. *J. Org. Chem.* **1993**, *58*, 2983.

0002-7863/95/1517-1873$09.00/0

of the formacetal modification on nucleic acid structure and annealing properties have been recently reported.[16,19-21] Incorporation of 3′-thioformacetal linkers into oligonucleotide duplexes, like incorporation of formacetal linkers, results in reduced Tms for DNA−DNA homoduplexes with reductions in Tm of 2−6 °C observed.[16,22] However, improved annealing properties, reflected in Tm increases of 1−5 °C, are seen in the case of DNA−RNA heteroduplexes, in which the DNA strand contains the 3′-thioformacetal modification. Replacement of the phosphodiester linker by a 5′-thioformacetal linker has a strongly detrimental effect on both oligonucleotide homoduplex and heteroduplex formation with complex formation essentially eliminated. As with the formacetal modification, computational and NMR studies were initiated in an effort to understand the stereoelectronic effects caused by the thioformacetal modifications. For the computational studies, the objective was to predict the effects prior to NMR structure determination.

We report here the results of computational studies directed at understanding the conformational properties of oligonucleotides containing thioformacetal linkers. Ab initio studies at the 6-31G* level of theory for thioformacetal (TF), methyl propyl sulfide (MPS), and methyl, 2-methoxy ethyl sulfide (MMES) are presented. Additionally force field parameters developed to reproduce the 6-31G* potential energy surfaces for these molecules are discussed. These parameters are then used to investigate by molecular mechanics the pseudorotational barriers of thymine 5′-thio, 2′-deoxyribose, and thymine 3′-thio, 2′-deoxyribose nucleosides as well as their adenine counterparts. Finally, results from a high temperature molecular dynamics simulation of a thymine dinucleotide containing a 3′-thioformacetal linker are presented. The results from these studies prestage further studies, involving solvated molecular dynamics simulations and incorporating NMR derived distance restraints of oligonucleotide duplexes containing thioformacetal linkages. However, the results presented here alone provide logical rationales for the observed effects of the 3′- and 5′-thioformacetal modifications on binding affinity of oligomers containing these substitutions. Additionally, the molecular dynamics simulation yields a novel backbone conformation for the 3′-thioformacetal linked dinucleotide.

Methods

Ab initio calculations were conducted for thioformacetal (TF), methyl propyl sulfide (MPS), and methyl 2-methoyxy ethyl sulfide (MMES) using GAUSSIAN90[23] with the 6-31G* basis set.[24,25] For TF, the potential energy surface was investigated by fixing COCS and COSC dihedral angles at various multiples of 60°. Torsional angles for either COCS and COSC were also constrained to 60° and the second COCS/OCSC torsional angle was constrained at either 0°, 30°, 60°, 90°, 120°, 150°, or 180°. The geometry of all other parameters were fully optimized. A similar protocol was followed for MPS. For MMES, the central OCCS torsion was varied in increments of 30°: all other parameters were fully relaxed with the flanking COCC and CSCC torsions in trans conformations for all calculations. In addition to the above calculations, full geometry optimization for gauche−gauche, gauche−trans, and trans−trans conformations was performed for TF and MPS as well as for gauche and trans conformations for MMES. The calculations were done on a CRAY Y-MP at the North Carolina Supercomputing Center.

Molecular mechanics calculations on TF, MPS, and MMES were done using the X-PLOR package[26] and the AMBER potential energy force field[27] on a Silicon Graphics R4400 Indigo Extreme Workstation. Charges were obtained from electrostatic potential fitting of the global minimum conformation of each molecule using the program CHELP[28] and scaling by a factor 0.8.[29,30] All existing AMBER parameters were used with the exceptions of those for the OCS angle term (taken from ref 31), COCS, OCSC, and OCCS specific dihedral terms, and the sulfur van der Waal's radius (2.1 Å was used). For the COCS, OCSC, and OCCS dihedral terms multiple potential terms were used with values as discussed in the Results section. For minimizations performed during force field parameter development, a 15 Å nonbonded cutoff and dielectric constant =1 were used in conjunction with an electrostatic shift function and van der Waals switch function. These functions act to apodize the electrostatic and van der Waals interactions at the nonbonded cutoff.[26] The molecules were minimized (Powell conjugate gradient method[32]) to a δ rms gradient of <0.001 kcal/mol·Å^2.

Nucleosides and nucleotide dimers were built using the TRIPOS modeling software and outputted as pdb files for use in conjunction with X-PLOR psf topology files (all dihedrals explicitly defined). The nucleoside structures were modified to have 5′/3′ OH or SH groups. Charges for the SH groups in the nucleosides were obtained by altering the AMBER 5′ OH/3′ OH nucleoside charges so as to approximately reproduce the differences observed between the electrostatic potential derived charges for 6-31G* optimized geometries of methanol and methyl sulfide. For the dimers containing thioformacetal linkers, the ESP derived charges for the TF portion of the dimer were patched onto the normal AMBER nucleotide charges with slight modifications so as to maintain charge neutrality.

For all molecular mechanics and molecular dynamics calculations a 15 Å nonbonded cutoff and distance dependent dielectric = $4r_{ij}$[33] were used in conjunction with van der Waals and electrostatic switch functions. Pseudorotational surfaces were investigated by using O4′-C1′-C2′-C3′ and C1′-C2′-C3′-C4′ dihedral restraints designed to restrain the deoxyribose to *P* values over the range of 0° to 180° in increments of 18° with a maximal pucker of 38°. Each nucleoside or dimer conformation was minimized to a delta rms gradient of < 0.001 kcal/mol · Å^2. Additionally, energies for unrestrained local C2′ endo and C3′ endo minima were obtained for all nucleosides and dimers.

The high temperature molecular dynamics studies of the TT 3′*S*-thioformacetal dimer were carried out using X-PLOR Verlet dynamics and simulated annealing protocols.[26] One thousand structures were obtained by simulating at 1100 °C for 0.4 ps, cooling at 400 °C for 0.2 ps, cooling at 100 °C for 0.5 ps, and finally mimizing for 2000 iterations. In each case the minimized structure served as the starting structure for the next round of dynamics. All masses were set to 100 Daltons, and a time step of 0.5 fs was used. To bias the structures toward right handed helical structures with antiglycosidic torsional values, weak square well potentials were placed on the two C4−N1−C1′−O4′ dihedrals (−140° ± 60°) and the N1(T1)−C1′(T1)−C1′(T2)−N1(T2) torsional angle (25° ± 15°). The minimized structures were then analyzed by initially discarding all structures not within 4.0 kcal/mol of the lowest energy structure found and then screening the remaining structures for those having interatomic distances consistent with duplex formation. These distances included N1−N1, N3−N3, and C5−C5 interbase distances, N1−C1′ distances, and C5′(T1)−O3′(T2), C5′(T1)−N1(T2), and N1(T1)−O3′(T2). The retained structures were then visually inspected, prior to acceptance.

Results and Discussions

Quantum Mechanics Studies. Previous ab initio studies[20] for thioformacetal (TF) have shown that the gauche−gauche conformation for the central two dihedral angles (COCS and

(23) Frisch, M. J.; Head-Gordon, M.; Foresman, J. B.; Trucks, G. W.; Raghavachari, K.; Schlegel, H. B.; Robb, M.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. *Gaussian 90;* Gaussian Inc.; Pittsburgh, PA.
(24) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 203.
(25) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1983**, *77*, 3054.
(26) Brunger, A. T. *X-PLOR (version 2.11) Manual,* Yale University: New Haven, CT, 1991.
(27) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230.
(28) Chirlian, L. E.; Francl, M. M. *J. Comput. Chem.* **1987**, *8*, 894.
(29) Cox, S.; Williams, D. *J. Comput. Chem.* **1981**, *2*, 304.
(30) Weiner, S. J; Kollman, P. A; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.
(31) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(32) Powell, M. J. D. *Mathematical Programming* **1977**, *12*, 241.
(33) Orozco, M.; Laughton, C. A.; Herzyk, P.; Neidle, S. *J. Biomol. Struct. Dyn.* **1990**, *8*, 359.

Table 1. 6-31G* Calculated Energies and Geometries for Thioformacetal Conformers[a]

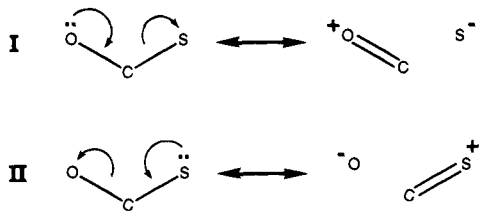
	OCSC[b]						
COCS−	180	−120	−60	0	60	120	180−
180	4.00	4.80	1.38	5.72	1.37	4.80	4.00
−120	4.88	6.21	2.64	7.00	3.65	6.64	4.88
−60	4.00	4.30	0.63	11.09	7.08	3.88	4.00
0	9.47	9.44	9.49	20.69	9.49	9.44	9.47
60	4.00	3.88	7.08	11.09	0.63	4.30	4.00
120	4.88	6.64	3.65	7.00	2.64	6.21	4.88
180	4.00	4.80	1.38	5.72	1.37	4.80	4.00

COCS/OCSC[b]							
60/0	60/30	60/60	60/64	60/90	60/120	60/150	60/180
11.09	4.89	0.63	0.58	1.92	4.30	4.64	4.00
0/60	30/60	60/60	75/60	90/60	120/60	150/60	180/60
9.49	5.28	0.63	0.03	0.55	2.64	2.36	1.37

C−O	C−S	C−O−C	O−C−S	C−S−C	COCS	OCSC	hartrees	kcal/mol[b]
1.384	1.817	115.74	114.96	99.61	75.1	63.4	−590.610 635 8[c]	0.00
1.379	1.832	115.86	111.06	99.02	80.0	189.6	−590.605 649 9[d]	3.13
1.390	1.799	113.59	110.92	99.98	180.7	64.1	−590.608 523 9[d]	1.33
1.384	1.813	113.88	106.69	98.65	179.8	180.0	−590.604 262 3[d]	4.00

[a] All values are for optimized geometries using the 6-31G* basis set. [b] Energies expressed in kcal/mol are relative to the global minimum gauche−gauche conformation (−590.610636 Hartrees). [c] Global minimum. [d] Local minima.

Scheme 1



OCSC) is preferred due to anomeric stabilization, similar to results seen for dimethoxymethane and dimethyl phosphate.[19,34-36] In this study the thioformacetal potential energy surface has been characterized in additional detail, at the 6-31G* basis set level, in the regions most relevant to nucleic acid polymer conformation, and the results are summarized in Table 1. The thioformacetal is asymmetric deriving from the distinct conformational preferences of the COCS and CSCO torsions. For the CSCO dihedral, a strong gauche preference is observed (3.1 kcal/mol) which is similar in magnitude to that observed for a COCO torsion (2.7 kcal/mol [34]). The COCS gauche preference is much weaker (1.3 kcal/mol). This reduced anomeric effect can be interpreted as resulting from a combination of factors. The van der Waal's radius of sulfur is substantially larger than that of oxygen (values of 2.10 and 1.65Å, respectively, were used in the modeling work in this study—see below) resulting in a gauche minimum of 75.1° for the COCS dihedral. This deviation from an ideal 60° likely diminishes the anomeric stabilization of the σ antibonding orbital. Additionally, in the resonance structures which can be drawn for the COCS and CSCO σ^* orbitals (Scheme 1) the positive charge for the COCS torsion resides on the oxygen (I) which would be expected to be less favorable, based on relative electronegativities, than in the CSCO case (II) where the positive charge resides on the sulfur and the negative charge is placed on the oxygen.

The asymmetry in the thioformacetal potential energy surface is significant when nucleic acid conformation is considered. Two dihedral angles, C3′−O3′−P−O5′ (ζ) and O3′−P−O5′−C5′ (α), characterize the phosphodiester linker conformation.[37] In an A form nucleic acid, characteristic of RNA and in some cases

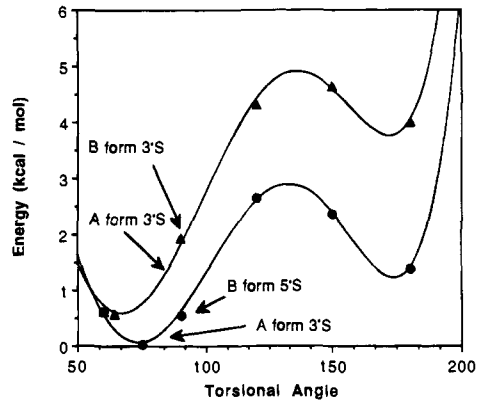


Figure 2. Thioformacetal potential energy curves at the 6-31G* basis set level. The graph is for either the COCS (●) or OCSC (▲) dihedral angle constrained to 60° with rotation about the unconstrained COCS or OCSC torsion at 30° intervals. The arrows denote the regions typically occupied by nucleic acid duplexes assuming α and ζ values of −60°, −90° for a B form duplex and −75°, −60° for an A-form duplex.

DNA, α is typical ~ −75° and ζ is typical −60°. Conversely in a B form nucleic acid, typical of double stranded DNA, α is usually near −60°, whereas ζ ranges from −90° to −120°.[37] When the thioformacetal substitutions are considered, O3′ is replaced by sulfur in the case of the 3′-thioformacetal, and O5′ is substituted by sulfur in the 5′-thioformacetal (Figure 1). It becomes apparent that the COCS and CSCO minima energy conformations are more compatible with an A form backbone linker conformation in the case of a 3′-thioformacetal linker, and the 5′-thioformacetal linker should prefer a more B-form-like conformation (Figure 2, Table 2). In the case of the 3′-thioformacetal, the A-form linker conformation is in fact very close to the actual global energy minimum for thioformacetal. For the 5′*S*-thioformacetal the conformational energy cost to adopt a B from conformation is similar to that calculated for dimethyl phosphate anion.[19]

Other heavy atom dihedrals which occur in the 3′- and 5′-thioformacetal nucleotides and involve sulfur are OCCS, CCCS, and CSCC. The potential energy surfaces for these dihedral terms were investigated using methyl propyl sulfide (MPS) and

(34) Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1989**, *111*, 4821.
(35) Liang, C.; Ewig,C. S.; Stouch,T. R.; Hagler, A. T. *J. Am. Chem. Soc.* **1993**, *115*, 1537.
(36) Gorenstein, D. G.; Findlay, J. B.; Luxon, B. A.; Kar, D. *J. Am. Chem. Soc.* **1977**, *99*, 3473.
(37) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

Table 2. Expected Energy Penalties for 5′- and 3′-Thioformacetal Linkers To Adopt Conformations Typical of A and B Form Nucleic Acids[a]

X[b]	Y[b]	Z[b]	B-form[c]	A-form[c]
O	PO_2	O	0.7	0.4
O	CH_2	S	0.6	1.2
S	CH_2	O	1.9	0.2

[a] Energies are expressed in kcal/mol and are relative to the global minimum gauche−gauche conformations. [b] X, Y, and Z are the O3′, PO_2, and O5′ positions, respectively, of a standard nucleic acid. [c] Values are based on an assumed B-form conformation in which the α and ζ (see text) dihedrals are −60° and −90°, respectively, and an assumed A-form confromation in which α and ζ are −75° and −60°, respectively.

methyl 2-methoxy ethyl sulfide (MMES) as representative molecules. Results for MPS (Table 3) show that the fully trans conformation is preferred by a modest 0.9 kcal/mol relative to the gauche/gauche conformation. This preference derives from both CCCS and CSCC torsions favoring the trans conformation by small amounts: the gauche−trans energy differences are 0.8 and 0.3 kcal/mol, respectively. The small energy difference observed for the CSCC gauche and trans conformations is consistent with experimentally observed and calculated values for ethyl methyl sulfide which indicate that gauche and trans conformations are approximately equal in energy.[38-40] The CSCC barrier to rotation is relatively small, 1.8 kcal/mol based on the 180°/120° value in Table 3. This lower barrier to rotation, relative to butane,[41] can partially be explained as resulting from reduced 1−4 van der Waal's repulsion in the eclipsed conformation because of the longer 1.82 Å C−S bond. Conversely, the CCCS 120° barrier is increased, compared to butane, with a value of 3.5 kcal/mol. The larger sulfur van der Waal's radius and the resulting increased 1−4 repulsion can be considered as a source for the increased barrier.

For MMES, the fully trans conformation is preferred (Table 4). Note that for the quantum mechanics studies the focus was on the central OCCS torsion and the flanking COCC and CCSC torsions were in trans conformations for all calculations. The calculated OCCS gauche−trans energy difference at the 6-31G* level of theory is 0.8 kcal/mol (Table 4) with a barrier to rotation (120°) of 3.6 kcal/mol. The local minima in the gauche region is 64.4°. Note that the 0.8 kcal/mol gauche−trans energy difference is reduced relative to the 1.4 kcal /mol gas phase preference for the trans conformation observed for dimethoxyethane (DME) at the 6-31G* level of theory.[42] This difference can be rationalized in part as deriving from reduced electrostatic repulsion in the MMES gauche conformation between the sulfur and oxygen atoms relative to the two oxygens in DME. This explanation is also consistent with MMES having a gauche conformation OCCS torsion of 64.4° versus 71.9° for the OCCO torsion in DME despite the larger van der Waal's radius of sulfur relative to oxygen.

Force Field Parameterization. The 6-31G* potential energy surfaces were subsequently used to develop force field parameters for use in molecular mechanics calculations. The general strategy was to adopt AMBER[27] bond stretch, angle bending, and wild card dihedral (e.g., x-ct-ct-x) terms unchanged and to parameterize by adjusting the sulfur van der Waal's radius and specific dihedral terms. Emphasis was placed on accurately parameterizing the regions of conformational space most

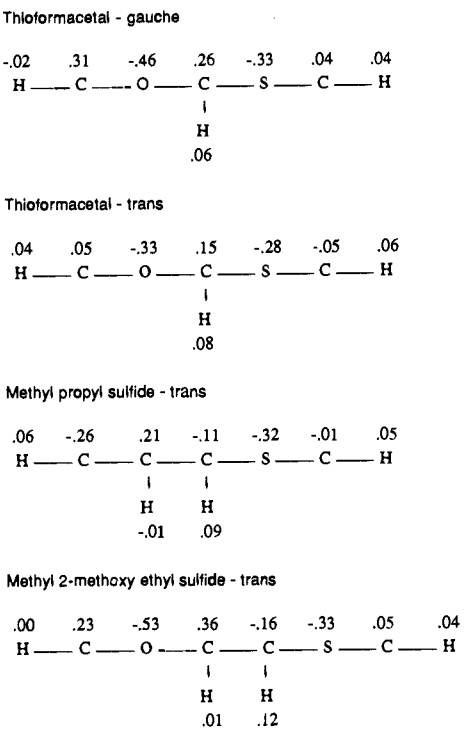


Figure 3. Electrostatic potential fitted charges for 6-31G* optimized geometries of TF, MPS, and MMES. The gauche and trans conformation ESP derived charges of TF are shown to illustrate the significant conformational effects on the calculated charges.

relevant to nucleic acid conformation, principally in the vicinity of local and global minima. Partial charges were derived from electrostatic potential (ESP)[28] fitting of the lowest energy conformations of each of the three molecules examined. These charges are summarized in Figure 3: the ESP partial charges of the trans conformation of thioformacetal are also included in Figure 3 to illustrate the conformational dependence of these charges particularly with respect to the sulfur atom. Clearly, conformational effects on charge can be significant[43-46] and complicate force field parameterization especially in high energy eclipsed conformations.

Figure 4 shows comparisons of the molecular mechanics derived potential energy surfaces to those obtained from the 6−31 G* calculations. The parameters used to obtained the molecular mechanics energies are summarized in Table 5. Use of a 2.1 Å van der Waal's radius for sulfur, wild card threefold torsional terms taken from AMBER, and twofold dihedral terms (onefold in the case of OCCS) was sufficient to parameterize the dihedral terms of interest. Rms deviations of the molecular mechanics and all quantum mechanics derived values (Tables 1, 3, and 4) are 1.0, 0.7, and 0.1 kcal/mol for TF, MPS, and MMES, respectively but the values for TF and MPS are substantially increased by differences involving high energy conformations. As can be seen from Figure 4, the potential energy surface is steep in these regions, especially for TF, and given that one set of partial charges is employed for parameterization of all conformations, high accuracy force field parameters are difficult to obtain for the high energy conformations. When only the rms deviations for conformations less than 4.0 kcal/mol above the quantum mechanics derived global minima are considered, values of 0.5, 0.3, and 0.1 kcal/mol for TF, MPS, and MMES, respectively, are obtained. Further

(38) Hayashi, M.; Shimanouchi, T.; Mizushima, S. *J. Chem. Phys.* **1957**, *26*, 608.
(39) Sakakibara, M.; Matsuura, H.; Harada, I.; Shimanouchi, T. *Bull. Chem. Soc. Jpn.* **1977**, *50*, 111.
(40) Allinger, N. L.; Quinn, M.; Rahman, M.; Chen, K. *J. Phys. Org. Chem.* **1991**, *4*, 647.
(41) Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1988**, *110*, 8029.
(42) Murcko, M. A.; DiPaola, R. A. *J. Am. Chem. Soc.* **1992**, *114*, 10010.
(43) Williams, D. E. *Biopolymers* **1990**, *29*, 1367.
(44) Reynolds, C. A; Essex, J. W.; Richards, W. G. *Chem. Phys. Lett.* **1992**, *199*, 257.
(45) Stouch, T. R.; Williams, D. E. *J. Comput. Chem.* **1992**, *13*, 622.
(46) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.

Table 3. 6-31G* Calculated Energies and Geometries for Methyl Propyl Sulfide Conformers[a]

	CCSC[b]						
CCCS	−180	−120	−60	0	60	120	180−
180	0.00	1.85	0.59	4.49	0.59	1.85	0.00
−120	3.47	5.96	4.04	9.08	4.13	5.94	3.47
−60	0.96	2.44	1.18	8.38	4.96	2.65	0.96
0	6.29	8.41	8.89	17.43	8.89	8.41	6.29
60	0.96	2.65	4.96	8.38	1.18	2.44	0.96
120	3.47	5.94	4.13	9.08	4.04	5.96	3.47
180	0.00	1.85	0.59	4.49	0.59	1.85	0.00

C−C	C−S	C−C−C	C−C−S	C−S−C	CCCS	CCSC	hartrees	kcal/mol[b]
1.528	1.818	111.94	110.78	100.11	180.0	180.0	−554.805 337[c]	0.00
1.531	1.821	114.41	112.04	99.82	67.6	178.0	−554.804 050[d]	0.81
1.528	1.818	111.87	115.36	101.58	178.1	72.7	−554.804 856[d]	0.30
1.531	1.820	114.13	116.26	101.24	66.5	74.1	−554.803 910[d]	0.90

[a] All values are for optimized geometries using the 6-31G* basis set. [b] Energies expressed in kcal/mol are relative to the global minimum trans−trans conformation. [c] Global minimum. [d] Local minima.

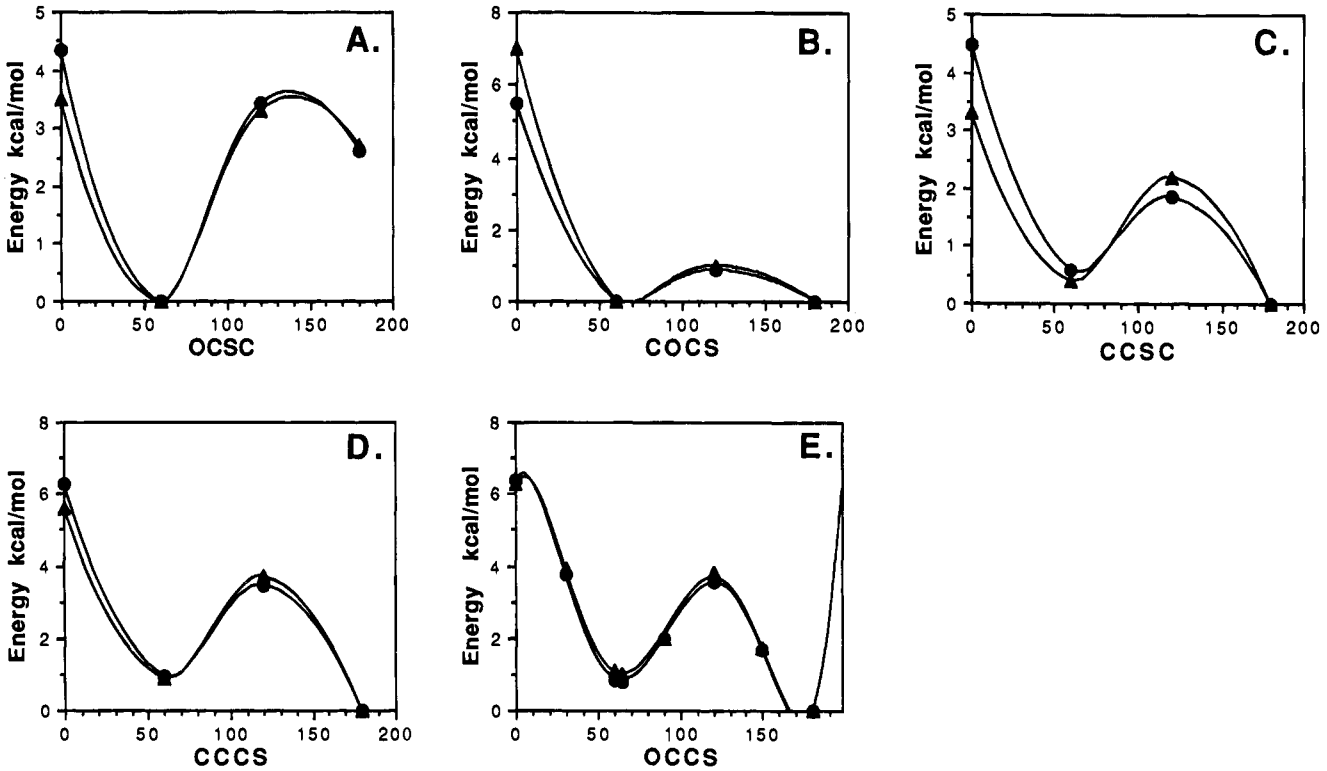


Figure 4. A comparison of the ab initio quantum mechanics (●) and force field (▲) potential energy surfaces for TF (A) and (B), MPS (C) and (D), and MMES (E). All heavy atom dihedral angles except for the one indicated on the abscissa were constrained to 180°. Values on the curves are taken relative to the mimimum on the curve and not overall global minima for the different molecules. See Table 5 and the text for force field parameter information.

Table 4. 6-31G* Calculated Energies and Geometries for Methyl 2-Methoxy Ethyl Sulfide Conformers (COCCSC)[a]

O−C	C−C	C−S	O−C−C	C−C−S	OCCS	hartrees	kcal/mol[b]
1.394	1.520	1.817	107.47	110.38	180.0	−629.645 226[c]	0.00
1.392	1.518	1.818	109.26	111.57	64.4	−629.643 930[d]	0.81

OCCS	hartrees	kcal/mol[b]	OCCS	hartrees	kcal/mol[b]
0	−629.635 048		90	−629.642 030	2.01
30	−629.639 235	3.76	120	−629.639 529	3.56
60	−629.643 849	0.86	150	−629.642 586	1.66

[a] All values are for optimized geometries using the 6-31G* basis set. Values are for rotations about the central OCCS torsion with the flanking COCC and CCSC torsions having trans geometries. [b] Energies expressed in kcal/mol are relative to the global minimum trans conformation. [c] Global minimum. [d] Local minimum.

improvement for the TF rms deviation is obtained if only conformers 3.0 kcal/mol or less above the global deviation are considered with a rms value of 0.3 kcal/mol obtained. Again the objective was to parameterize the dihedral terms of interest in the regions most relevant to biomolecule conformation: the high energy eclipsed conformations are not expected to be substantially populated and thus the strategy for parameterization is viewed as successful based especially on the latter rms values.

Molecular Mechanics Derived Pseudorotation Potential Energy Surfaces for Modified Nucleosides. The goal in developing the force field parameters described above was to evaluate the effect of incorporation of thioformacetal linkers

Table 5. Force Field Parameters Used for Molecular Mechanics Calculations[a]

dihedral	Vn/2 (kcal/mol)	periodicity	phase (deg)	Vn/2 (kcal/mol)	periodicity	phase (deg)
COCS	0.38	3	0	1.2	2	0
CSCO	0.33	3	0	0.8	2	0
CCCS	0.14	3	0			
CCSC	0.33	3	0			
OCCS	0.14	3	0	1.2	1	180

[a] Bond stretch and angle bend terms from the AMBER force field were used for all calculations. An equilibrium value of 110.5° and a force constant of 75 kcal/mol-deg^2 were used for the OCS angle term.[31] AMBER van der Waal's terms were also used with the exception of sulfur for which r and and e values of 2.1 Å and 0.20 kcal/mol were used, respectively.

Table 6. Molecular Mechanics Derived Conformational Energies for Thio-Substituted Nucleosides[a]

nucleoside	C2′ endo energy (P)[b,c]	C3′ endo energy (P)[c]	O4′ endo energy (P)[d]
dA	0.0 (146)	0.5 (17)	1.2 (72)
3′*S* dA	0.1 (160)	0.0 (17)	1.0 (90)
5′*S* dA	0.0 (135)	0.8 (33)	1.1 (72)
T	0.0 (148)	0.9 (28)	1.3 (72)
3′*S* T	0.0 (157)	0.3 (27)	0.9 (90)
5′*S* T	0.0 (135)	1.1 (44)	1.2 (72)

[a] Values are for deoxyadenosine (dA), thymidine (T), and their thio derivatives in which sulfur is substituted for the 3′ oxygen (3*S*) or 5′ oxygen (5*S*). [b] Shown are relative energies with pseudorotation values in kcal/mol in parentheses. [c] Global and local minima for fully relaxed conformations in the C2′ endo and C3′ endo conformational regions. [d] Approximate barriers to pseudorotation obtained by constraining v2 and v1 so as to obtain pseudorotational values of 72° and 90° with a maximal pucker of 38°.

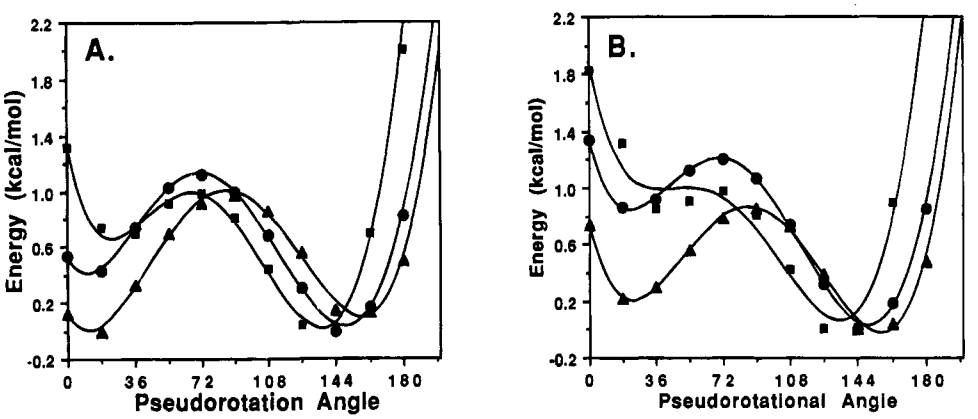


Figure 5. Comparison of pseudorotational energy surfaces for deoxyadenosine (A) and deoxythymidine (B) nucleosides. (●) indicates the curves for the native nucleosides, (▲) indicates the 3′*S* nucleoside curves, and (■) indicates the 5′*S* nucleoside curves.

on nucleic acid conformation. This incorporation results in substitution of a sulfur atom for an oxygen at either the O3′ (3′*S*-thioformacetal) or O5′ (5′*S*-thioformacetal) location. Initially, we wished to examine the effects of these substitutions on the pseudorotation pathway[47] of simple deoxyadenosine and deoxythymidine nucleosides with 3′/5′ OH (SH) groups. This was investigated by restraining $\nu 1$ (04′−C1′−C2′−C3′) and $\nu 2$ (C1′−C2′−C3′−C4′) to various values designed to span the pseudorotation pathway from P values of 0° to 180° with a maximal pucker of 38°. The results are summarized in Table 6 and illustrated in Figure 5. For unmodified nucleosides, as previously reported,[27] C2′ endo type (P ≈ 156°) conformations, typical of B form nucleic acid conformations, are preferred (Table 6) by 0.5 (dA) and 0.9 kcal/mol (dT) over their C3′ endo type (P ≈ 18°) counterparts which are usually observed for A form nucleic acid conformations.[37] Barriers to pseudorotation via the O4′ endo conformation are calculated to be 1.2−1.3 kcal/mol.

Although the magnitudes of the changes are not large, several trends are apparent when the sulfur substitutions are considered. The most significant observation is that in the case of the 3′*S* substitution, the preference for a B-form nucleoside conformation is abolished and the barrier to pseudorotation is diminished. Conversely, for the 5′*S* substitution, the B-form conformational preference is enhanced for both the dA and dT nucleoside derivatives. These changes can be interpreted most readily as arising from steric interactions between the substituted sulfur and heavy atoms in the deoxyribose ring. In a B-form conformation, O3′ is gauche with respect to C1′ and O4′, and the O3′−C1′ and O3′−O4′ distances are observed to be ≈3.0 and ≈3.2 Å, respectively. Substitution of sulfur at the 3′ position, as a result, leads to significant steric crowding. This crowding can be alleviated by pseudorotation to a more A-form like conformation: the unmodified nucleoside distances are ≈3.6 Å for both C1′ to O3′ and O4′ to O3′.

For the 5′*S*-thio substitution, the molecular mechanics calculations indicate an enhanced preference for the C2′ endo conformational region. The minimized unmodified nucleoside for both C2′ endo and C3′ endo structures have close contact distances between O5′ and both O4′ (2.9 Å) and C3′ (3.0 Å). These close contacts arise in part because the structures were minimized with the γ torsion (O5′−C5′−C4′−C3′) in a positive gauche conformation, the rationale for this choice being that this is the experimentally observed conformation in both A-form and B-form unmodified nucleic acids.[37] The C2′ endo conformation is apparently better able to accommodate the 5′*S* substitution by through compensating conformational changes. However, the overall preferred conformation is with the 5′*S* trans to C3′, i.e., with the γ angle trans (dE ≈ 0.6 kcal/mol by molecular mechanics calculations), and, as a result, the standard low energy nucleoside conformations do not as readily accommodate the 5′*S* substitution.

For the 5′*S*-thioformacetal linker, additional steric contacts

(47) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.

Table 7. Nucleic Acid Parameters for Possible Conformers of the 3′*S*-Thioformacetal Modified TT Dimer as Generated by High Temperature Molecular Dynamics[a]

conformer	χ_1[b]	P_1	δ_1	ϵ	ζ	α	β	g	χ_2	P_2	δ_2
A[c]	−150	142	124	−178	−79	−68	−180	62	−124	132	126
B[d]	−130	81	79	48	54	79	176	−176	−129	140	132
C	−117	84	79	55	60	−174	−176	59	−123	55	77
D	−105	91	83	72	80	173	63	167	−148	170	150

[a]Parameters are for the TT dimer T-3′*S*-CH_2-5′O-T as described in the results section. [b] α, β, γ, δ, ϵ, ζ, and χ are the standard torsional angles describing nucleic acid conformation and *P* is the pseudorotation parameter describing the deoxyribose ring conformation.[47] The subscripts 1 and 2 denote the 5′ T and 3′T, respectively, and all other parameters refer to the linker region between the two bases. [c] Conformation A is that which is typically observed for nucleic acids. [d] Conformation B is consistent with NMR experimental data[21] and is characterized by α, ζ being +g, +g rather than the typical −g, −g.

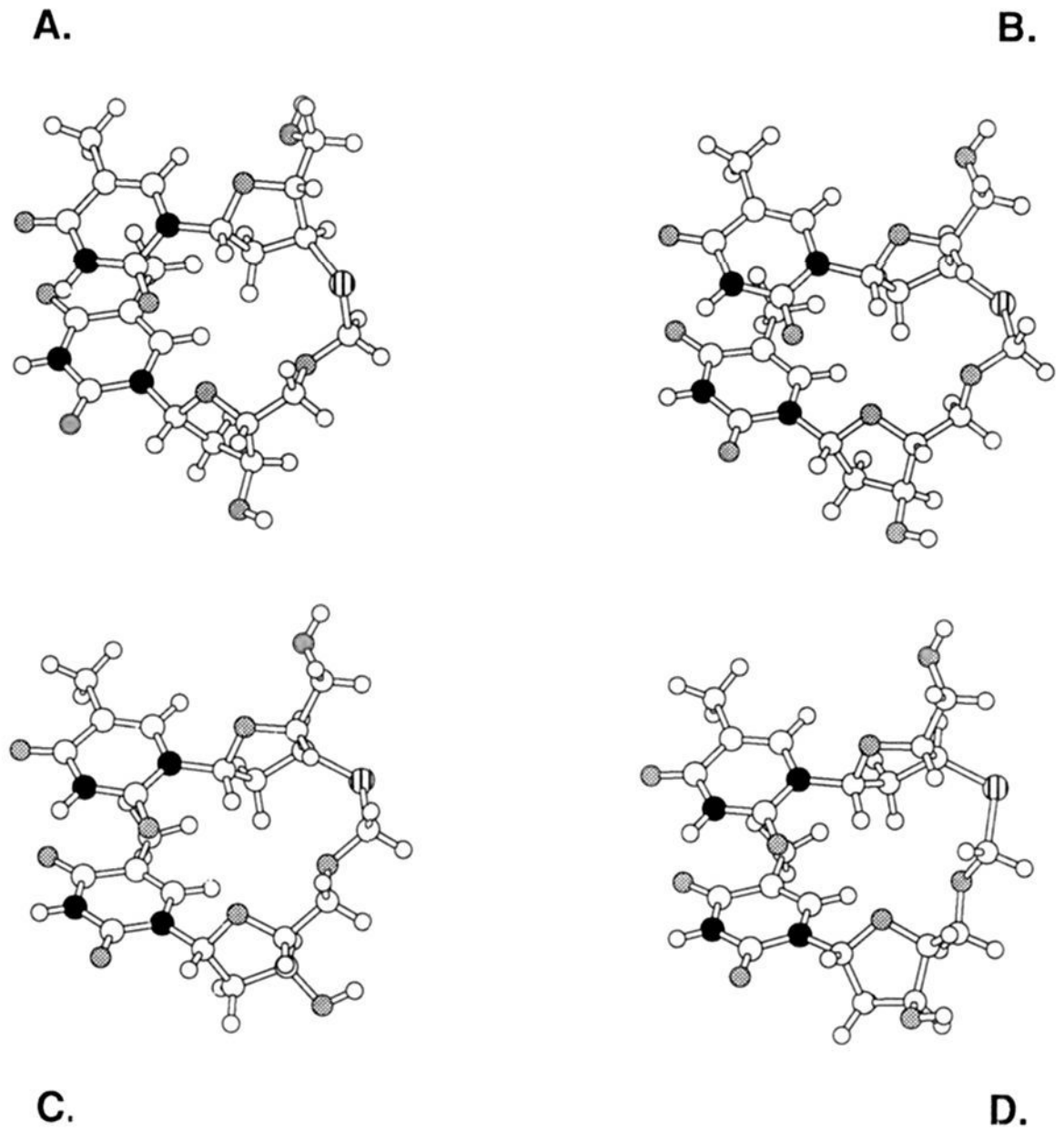


Figure 6. Structures obtained from high temperature molecular dynamics simulations of 3′*S*-thioformacetal modified TT dimers. Atom shading is carbons and hydrogens, white; nitrogens, black; oxygens, grey; sulfur, vertical lines. The structures shown are those found which are most likely to be compatible within the context of an oligomer duplex. Structure (A) is typical of native oligonucleotides. Structure (B) is the low energy conformation consistent with NMR experimental data. Structures (C) and (D) have conformations compatible with duplex formation, but they have higher relative potential energies than (A) or (B).

also become significant when TT dimers are considered. For a canonical B form type conformation with the deoxyriboses in C2′ endo conformations, several short interatomic distances are observed which are not readily alleviated by conformational adjustments during minimization. In particular the *S*5′ distances to C6 and H6 atoms of the same thymine residue and the C1′, C2′, H1′, and H2′ atoms of the the adjacent thymine base (5′-side) are all significantly less than the optimum van der Waal's contact distances. These sterics clashes are not significant in the case of an O5′ atom and the difference in van der Waal's radii for sulfur and oxygen, thus, becomes a key factor in considering the suitability of thioformacetal substitutions.

The results of the quantum mechanics and molecular mechanics calculations described above yield a consistent structural picture which is predictive of the conformational behavior of the modified nucleic acids. First, the asymmetry of the thioformacetal conformational preferences when incorporated into a nucleic acid linker lead to the prediction that an A-form type backbone should be preferred for the 3′*S*-thioformacetal linker such that the COCS and CSCO torsional angles can be readily accommodated near their −75° and −65° respective minima. Conversely, the 5′*S*-thioformacetal incorporation is more compatible with a B-form like structure. The expected B-form conformation has an energy difference, relative to the global minimum, similar to that calculated for dimethyl phosphate. These preferences are reinforced by the molecular mechanics calculated pseudorotational profiles, in which the normal preference for a C2′ endo conformation is removed in the case of the 3′*S*-thio substitution. Again, the converse is true for the 5′*S* substitution in which the apparent preference for a C2′ endo conformation, typical of B-form nucleic acids, is enhanced. However, the C2′ endo conformational is still not a low energy one due to steric interactions between the *S*5′ and both O4′ and C3′. The conclusions can thus be made: that the

3′*S*-thioformacetal modified nucleotide is an acceptable antisense modification with respect to annealing affinity and that this modification should drive the local nucleic acid conformation away from a B-form type structure and toward an intermediate or fully A-form like structure. The 5′*S*-thioformacetal modification is rationalized as being likely unacceptable principally for steric reasons with respect to a B form type duplex and for both steric and electronic reasons with respect to an A form type duplex. The thioformacetal computational results are fully consistent with the experimental thermal denaturation results, and the predicted effects for the 3′*S*-thioformacetal substitution on deoxyribose conformation have subsequently been confirmed by NMR experimental data.[21]

High Temperature Dynamics of a TT Dimer. In a final effort to examine conformational effects caused by the 3′*S*-thioformacetal substitution, high temperature molecular dynamics simulations of a thymine dimer incorporating the 3′*S*-thioformacetal linker were conducted. The objective was to access what backbone conformations were low in potential energy while being compatible with duplex formation. This investigation was conducted in part because of the experimental indications that the thioformacetal conformation in an oligomer studied by NMR spectroscopy[21] was in an alternate conformation rather the standard −g/−g conformation typically observed for nucleic acids. One thousand structures were generated during the course of the simulation, which was additionally biased toward duplex compatible structures by restraining the two chi torsional angles to the anticonformational region of space. These conformations were then screened for those having distances and vectors compatible with duplex formation. The results are informative. The standard deoxyribose−phosphate backbone pattern for the DNA torsional angles α, β, γ, δ, ϵ, and ζ is in fact found as one of the predicted low energy conformations, i.e., $\alpha = -g$, $\beta = t$, $\gamma = +g$, $\delta = t(g)$, $\epsilon = t$, and $\zeta = -g$ (Table 7, Figure 6a). However, alternate conformations are also observed: in particular a low energy +g/+g thioformacetal linker conformation, consistent with NMR spectroscopy data,[21] is observed to be readily accommodated within the context of a nucleic acid conformation suitable for duplex formation, provided there are concomitant changes to the other backbone torsional angles.

Specifically, +g, +g, +g, t, and t conformations for dihedral angles ϵ, ζ, α, β, and γ, respectively lead to a thymine dimer conformation suitable for duplex formation (Table 7, Figure 6b). In this unique conformation, every backbone torsional parameter, with the exception of β, is significantly altered from a conventional DNA backbone conformation. One feature of this alternative conformation is that the ϵ torsion, C4′−C3′−*S*3′−CTL where CTL is the linker methylene carbon, is in a partially eclipsed conformation with $\epsilon \approx 45°$. This partially eclipsed conformation is in part accomodated due to the longer C−S bond length and may explain why this conformation is experimentally observed for the 3′ *S*-thioformacetal linker, but not for a native phosphodiester linkages. Additionally, the 5′T deoxyribose conformation is pushed into an O4′ conformational state, consistent with both predictions deriving from the potential energy pseudorotational profiles described above and experimental data. The larger van der Waal's radius of the sulfur atom is also well tolerated in this structure, with the sulfur lone pairs directed toward solution.

In addition to the two conformations described above, two other dimer conformations were observed which were higher in potential energy but which also appeared to have some potential for duplex formation (Table 7, Figure 6c,d). These conformations are characterized by ϵ, ζ, and α being in +g, +g, and t conformations, respectively, and by by β and γ dihedrals being either t, +g (Figure 6c) or +g, t (Figure 6d). It is interesting to note that in all four of the observed low energy conformers, there are three gauche torsions and two trans torsions, indicating that this profile is most suitable for a favorable stacking arrangement which is amenable to duplex formation. Like the alternate conformer B, these conformers also have the 5′T deoxyribose in and O4′ endo conformation.

Conclusions

Ab initio studies have been conducted for thioformacetal, methyl propyl sulfide and methyl 2-methoxy ethyl sulfide. The results of these studies provide significant evidence toward explaining the experimentally observed effects of substitution of a thioformacetal linker for a phosphodiester linker in a nucleic acid oligomer. Most significantly, the strong preference of the thioformacetal for a gauche−gauche conformation is most readily accommodated into an A form type nucleic acid backbone in the case of a 3′*S*-thioformacetal linker. The results of the ab initio studies have been used to develop force field parameters for use in molecular dynamics simulations. High temperature molecular dynamics simulations of a TT dimer in turn have provided additoinal explanation for the effects of the 3′*S*-thioformacetal linker. The results support a previously unobserved conformation for the nucleic acid backbone in the region of the thioformacetal. This unique conformation may be exploited in the future for design of novel nucleic acid monomers with improved antisense/antigene properties.

JA943606U