Biopolymers, Vol. 29, pp 597–607 (1990)

Computational Analysis of the Effects of Site-Specific Phosphate Alkylation in the DNA Oligomer {d-[GGAATTCC]}$_2$

MICHELLE S. BROIDO and MIHALY MEZEI

Department of Chemistry and Center for the Study of Gene Structure and Function, Hunter College and the Graduate School of CUNY 695 Park Ave., New York, NY 10021.

SYNOPSYS

Alkylation of the sugar-phosphate backbone of DNA can result upon exposure to several potent carcinogens, inducing DNA misfunction. In order to assess the structural and energetic changes in DNA helices induced by such alkylation, we have performed AMBER-based analyses on phosphotriester containing analogues of $\{$d-[GGAATTCC]$\}_2$. Fourteen analogues of the non-alkylated oligomer were examined, each bearing a single alkylation of known stereochemistry. Results indicate that although there is minimal effect on the aromatic bases the presence of a phosphotriester disturbs the sugar-phosphate backbone in complex ways. For most analogues, total minimum energies are lower for the S_p-alkylations than for the R_p-alkylations which point directly into the major groove of the helix; however, different energetic contributions follow different, or no, trends in dependence on alkylation site and/or stereochemistry. Where data is available, experimental NMR results agree with the calculations reported here.

INTRODUCTION.

Early in the study of DNA biochemistry, details of nucleic acid structure were found to be important correlates to nucleic acid function. From that time and continuing to date, there have been numerous investigations of the interactions of alkylating agents with DNA. These range from studies of the *in vivo* physiological effects of DNA alkylation to the effects of alkylation on aspects of DNA molecular structure.[1−4] In the latter category, most of these studies have been aimed at determining the sites of alkylation, the sequence dependence of the alkylation sites, and the relative biological consequences of the different alkylation sites.[5−12] Many DNA alkylating agents alkylate almost exclusively at sites on the bases,[2−4] and for many years *base* alkylation was thought to be paramount for mutagenic effects or DNA inactivation. Within the past two decades, a number of studies have shown that phosphate alkylation, leading to phosphotriester formation, occurs with several potent carcinogens,[8−10] and for some ethylating agents there appears to be a positive correlation between the percentage of phosphate triester formed, relative to total alkylation, and carcinogenicity.[8,13]

The lethality of phosphotriester formation, as well as the site of alkylation, is highly dependent upon the nature of the alkylating group and on whether the nucleic acid sugars are ribose or deoxyribose.[8,14−19] Deoxyribophosphotriesters appear to be significantly more stable than ribotriesters and whereas methyl triesters are rarely lethal, higher alkylations are often lethal.[8,14,19,20] *In vitro* studies have shown that the absolute stereochemical configuration of a phosphotriester is also a determinant in the consequences on DNA biochemistry.[11]

Based on a variety of experimental results, the biologically significant effects of phosphate alkylation on nucleic acids are postulated to be loss of susceptibility to enzyme hydroloysis, perturbations in interactions with complementary polynucleotides as a consequence of charge neutralization, steric interference of the alkyl groups with protein-nucleic acid interaction, and changes in conformation which alter enzyme recognition sites.[13,15,21]

Van Genderen and coworkers have reported molecular mechanics calculations on methylated parallel {d-[TTTTTT]}$_2$ and antiparallel {d-[GCGCGC]}$_2$ and for each type compared the two possible methylated diastereomers.[22] The calculations were based on the AMBER force field.[23] They found that the removal of the phosphate charge upon methylation stabilized the parallel helix and that the phosphotriester torsion angles were the conformational parameters most affected by the stereochemistry of the methylation. For the antiparallel helix, it was observed that the S_P substitution resulted in a larger major groove and a smaller minor groove as compared with the unsubstituted oligomer.

We have recently begun both NMR and theoretical studies on the effects of phosphate alkylations on the solution conformation of the DNA octamer, {d-[5'G^{1}pG^{2}pA^{3}pA^{4}pT^{5}pT^{6}pC^{7}pC^{8}3']}$_2$ investigating both the effects of alkylation site and the effects of stereochemistry at a given alkylation site. We report here the results of molecular mechanics studies on the seven possible diastereomeric pairs (R_p and S_p, *vide infra*) of analogues with a single ethylphosphotriester per strand, e.g., (R_p,R_p)-{d-[G(et-p)GpApApTpTpCpC]}$_2$ and (S_p,S_p)-{d-[G(et-p)GpApApTpTpCpC]}$_2$, etc. We also report preliminary comparison of these theoretical data with relevant experimental data.

METHODS

The molecules we discuss here are all analogues of {d-[GGAATTCC]}$_2$ in which a single site on each strand of the sugar- phosphate backbone has a known modification, i.e., the replacement of a normal phosphodiester with an ethyl phosphotriester. The fourteen resulting analogues (7 internucleotide phosphates, two diastereomers at each phosphate) were energy-minimzed and the resulting structures were examined both with regard to conformation and with regard to energy. The fourteen self-complementary alkylated molecules investigated are referred to as R_p-G(et)G, S_p-G(et)G, R_p-G(et)A, S_p-G(et)A, R_p- A(et)A, S_p-A(et)A, R_p-A(et)T, S_p-A(et)T, R_p-T(et)T, S_p-T(et)T, R_p-T(et)C, S_p-T(et)C, R_p-C(et)C, and S_p-C(et)C, where X(et)Y refers to the internucleotide phosphate bearing the alkyl moiety. The absolute sterechemical designation R_p refers to the configuration wherein the ethyl group is oriented into the major groove of the helix (see Figure 1); for the S_p diastereomer the ethyl group is oriented away from the helix. Also energy minimized were the unmodified molecule (parent) and a hypothetical analogue in which the first internucleotide phosphate, G-PO$_2$-G, was given the reduced charge of a triester (*vide infra*) but no alkyl substituent, (G(no)G).

The oligonucleotides were modeled by the AMBER force field[23] which represents the energy of the system by harmonic bond stretching and bending terms and trigonometric torsion terms, supplemented with Van der Waals ($1/r^{12}$ exchange repulsion and $1/r^6$ dispersion), electrostatic ($1/(\epsilon r)$) interaction between non-bonded atoms, and a special $1/r^{12} - 1/r^{10}$ term for hydrogen bonded atoms. For the electrostatic interaction, a distance dependent dielectric constant $\epsilon = 4r$ was used to model the strong screening effect of solvent water (which was not included explicitly in the calculation). All hydrogens were treated explicitly from the outset of the calculations, and a 5 nm cut-off was used for defining atomic interactions, thus ensuring that all possible binuclear interactions were considered.

The AMBER force field includes parameters for the standard nucleic acids, but not for the derivatives considered in this study. In particular, the fractional charges on the atoms, used in the electrostatic energy term, are affected by the ethylation which necessarily causes neutralization of the phosphate charge. For the phosphotriester moiety we took advantage of the partial charges calculated for the methylated case by Van Genderen et al.[22]; the charges for the ethyl group were deduced from the methyl data by requiring overall charge neutrality. The actual charges used are shown in Table I. Our approach is justified by remarking that the approximation involved in using trimethyl phosphate as the representative of the phosphotriester in the DNA backbone[22] is certain to include larger inaccuracies than our extrapolation from methyl to ethyl.

The initial configuration of oligonucleotides was obtained by the NUCGEN module of AMBER[24] in the "idealized" B conformation and counterions (Na^+) were placed near the charged phosphate groups. Subsequently, each configuration was minimized with the program, obtaining the nearest local energy minimum of the system. A conformation was considered to be at the minimum energy when the root mean square of the gradients was less than 0.005nm. The resulting parameters were then processed using the analysis module of AMBER to obtain the decomposition of the energy both into the contributions from the various functional groups and into the different types of energetic interactions. The DOCK program[25] was used to view the minimized structures on an Evans & Sutherland PS390 graphics system and to obtain various structural parameters.

Table I Partial charges used for the ethylated phosphates

O3'	O5'	02	OSE	P	C2	H2	C3	H3
-0.277	-0.277	-0.354	-0.290	0.660	0.118	0.039	0.021	0.004

Legend: O3' and O5' are the backbone ester oxygens; OSE is the ethyl ester oxygen. The atoms C2 and H2 form the methylene group and the atoms C3 and H3 form the methyl group.

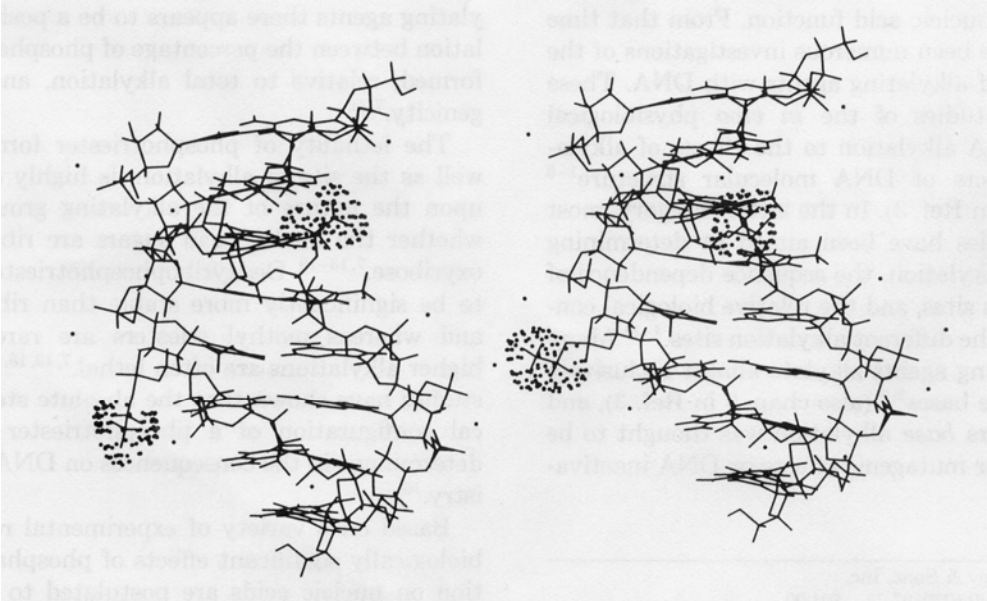

Figure 1. Energy minimized structures of R_p-T(et)T (left) and S_p-T(et)T (right); the Van der Waals radii of the ethyl moities are indicated with dotted surfaces; single dots represent counterion positions

RESULTS AND DISCUSSION

Conformational Features

In each case, the system remained in a general B-DNA conformation and the counterions stayed near their respective phosphates. The strict planarity of base pairing was lost, conforming to experimental data on oligomer duplexes.[26] Many of the "traditional" indicators of conformation were remarkably similar for all 16 molecules investigated. The helical lengths, measured as the interatomic separation of the deoxyribose C1' of G^1 to C1' of C^8 (on a given strand), ranged from 2.41 nm to 2.49 nm. All diastereomeric pairs had helical lengths which differed from each other by no more than 0.02 nm and the direction of the deviation did not appear to show any systematic dependence on the substitution site.

Table II The C3'-O3'-P-O5', C2-OSE-O-O5' (E1), and the P-O2-C2-C3 (E2) Torsion Angles in the Minimum Energy conformation

	C3'-O3'-P'-O5'							Ethyl	
	GG	GA	AA	AT	TT	TC	CC	E1	E2
Parent	-95.8	-97.8	-100.7	-95.3	-89.3	-91.8	-91.9		
G(no)G	-93.9	-95.7	-99.1	-95.3	-89.9	-91.3	-88.9		
R_p-G(et)G	-94.3	-99.2	-99.6	-94.5	-88.9	-94.3	-95.9	-66.7	177.4
S_p-G(et)G	-94.3	-98.9	-99.6	-94.7	-88.3	-94.9	-92.3	73.3	180.1
R_p-G(et)A	-105.5	-90.8	-100.5	-94.3	-85.7	-99.2	-94.1	-66.6	178.1
S_p-G(et)A	-93.9	-114.3	-102.0	-94.8	-88.5	-91.8	-86.8	73.1	176.9
R_p-A(et)A	-95.5	-96.1	-96.7	-96.5	-90.8	-91.0	-88.8	-65.2	176.9
S_p-A(et)T	-95.1	-97.4	-116.2	-96.0	-90.4	-90.5	-85.3	72.8	179.9
R_p-A(et)T	-95.4	-99.2	-100.0	-93.1	-89.9	-94.2	-96.6	-62.9	178.0
S_p-A(et)A	-96.4	-98.4	-99.6	-107.6	-92.6	-86.4	-95.1	74.3	179.8
R_p-T(et)T	-94.0	-96.9	-98.5	-93.5	-79.0	-96.9	-93.2	-69.5	179.3
S_p-T(et)T	-95.3	-97.4	-99.2	-93.9	-87.3	-96.8	-90.5	71.7	179.1
R_p-T(et)C	-95.6	-96.3	-99.6	-94.9	-90.1	-83.5	-92.4	-65.7	177.5
R_p-T(et)C	-95.2	-96.4	-99.3	-94.0	-85.5	-120.8	-94.8	73.9	180.0
R_p-C(et)C	-94.4	-97.2	-98.8	-94.9	-86.8	-91.1	-78.8	-68.6	178.4
R_p-C(et)C	-93.1	-95.0	-98.6	-94.4	-88.6	-90.6	-121.5	74.3	179.6

Base-pairing hydrogen-bond distances were all within 0.01 nm of the analogous distances in the parent molecule, 0.19 - 0.21 nm. Similarily, both the base-base total interaction energies and the specific hydrogen-bond interaction terms showed negligible variability. These results are seemingly in contrast with the experimentally observed duplex stabilities of singly-alkylated analogues of $\{$d-[GGAATTCC]$\}_2$ which have different melting temperatures, T_m, depending upon both the site and the stereochemistry of the alkylation.[27–30] Whereas S_p-A(et)A and S_p-A(et)T melt within two degrees of the parent octamer, R_p-A(et)A and R_p-A(et)T melt at temperatures 4-6^{o}C lower.[29,30] Isopropyl-bearing triesters show similar trends; if the alkylation points into the major groove of the helix, the T_m is lower than

for the other diastereomer, and there is a pronounced difference in the melting temperatures of analogues with a given stereochemistry depending upon the site of alkylation.[27,28] Possible sources for this apparent difference in stability are mentioned below in the course of the energetic analysis since the results discussed above show that it cannot be simply explained by geometric changes in base pairing. Of course, the energies calculated are those of the alkylated duplexes and they are compared with that of the non-alkylated duplex, while T_m refers to comparison of the duplex and melted states of a given oligomer.

The most evident geometric differences between the molecules were found in the backbone torsion angles, C3'-O3'-P-O5'. These angles, for each internucleotide linkage, are found in Table II for all 16 molecules studied. With the exception of T(et)T, the S_p substitution causes a significant (10-20^o) decrease in the torsion angle, relative to that in the parent octamer, while the R_p substitution has an opposite but smaller effect. For the T(et)T substitution both diastereomers show small increases in the torsion angles. A tempting explanation for the T(et)T behaviour would be that a repulsion exists between the methyl group on the thymines and the ethyl group. However, neither A(et)T nor T(et)C exhibit this anomaly, and the interaction energies between the thymine and ethyl groups (not shown) are similar in all three substitutions in the vicinity of thymine. The thymine-thymine interactions were also compared and were found unaffected by the ethylation. Thus the explanation of the anomalous T(et)T behaviour must lie elsewhere.

The orientation of the ethyl group can be described by the O5'-P-OSE-CT2 and P-OSE-CT2-CT3 torsion angles, also shown in Table II for each alkylated molecule. For the O5'-P-OSE-CT2 angle there is a 6.6^o range between the seven R_p molecules, and a 2.6^o range between the seven S_p molecules. There is no obvious correlation between the site of alkylation and the deviation from the mean torsion angle, however, R_p-T(et)T has the most negative torsion and S_p-T(et)T has the least positive torsion for both diastereomers. A similar trend is observed for the P-OSE-CT2-CT3 angle, with somewhat smaller ranges but with the extreme values again occuring at R_p-T(et)T and S_p-T(et)T. This correlates with the anomaly in the C3'-O3'-P-O5' torsion angle exhibited by the T(et)T substitution.

A geometric feature which correlates directly with experimental data is the interaction of the ethyl moiety with neighboring bases. In a study on the G(et)A diastereomeric twins, we found that the ethyl moiety of the R_p-diastereomer exhibits NOEs with the base proton of the adenosine; no such NOEs were observed for the S_p-diastereomer.[28] Using a 0.45 nm distance cutoff to determine whether such NOEs should be observed, examination of the energy-minimzed G(et)A diastereomers shows that only the R_p would exhibit such NOEs. We observed similar results for the isopropyl-bearing analogues, A(iPr)A and A(iPr)T.[27] Although the authors of the study on the A(et)T twins did not investigate the interaction of the ethyl moiety with the base proton to the 3'-side of the phosphotriester (and the signal to noise in the published spectrum does not allow us to assess the presence or absence of this interaction), they did report weak NOEs between the ethyl protons and the H3' of the adenosine for the R_p-isomer, but not for the S_p.[29] Examination of the energy-minimized A(et)T molecules indicate that these results are as would be expected.

Energetic features

The analysis module of the AMBER package decomposes the total energy in ways that enable the tracing of the source of overall energetic differences to specific contributions and to the identity of differences in various contributions which might not be obvious due to cancellation upon summation. For the purpose of the analysis, we defined as groups the individual bases, the sugars, the PO_4 unit, the combined counterions, and the ethyl group.

Table III Total Energies and Their Partition to Contributions from Various Interaction Types

	$E_{\rm tot}$	$E_{\rm vdW}$	$E_{\rm el}$	$E_{\rm HB}$	$E_{\rm bond}$	$E_{\rm ang}$	$E_{\rm dih}$	$E_{\rm 14vdW}$	$E_{\rm 12el}$
Parent	-310.0	-252.7	-34.0	-11.4	4.7	54.0	144.0	82.1	-296.3
G(no)G	-289.2	-244.9	-13.6	-11.3	4.6	54.7	141.8	82.3	-291.8
$R_{\rm p}$-G(et)G	-278.8	-258.7	-13.7	-11.4	4.7	59.5	150.1	82.7	-292.0
$S_{\rm p}$-G(et)G	-281.6	-258.8	-13.3	-11.5	4.7	59.3	147.5	82.7	-292.1
$R_{\rm p}$-G(et)A	-277.9	-257.6	-14.0	-11.2	4.7	59.3	150.5	82.3	-292.0
$S_{\rm p}$-G(et)A	-279.2	-256.4	-15.1	-11.3	4.7	58.1	149.7	82.8	-291.7
$R_{\rm p}$-A(et)A	-274.7	-256.2	-15.1	-11.4	4.7	54.4	153.5	82.9	-291.5
$S_{\rm p}$-A(et)T	-280.1	-256.9	-14.7	-11.4	4.7	58.3	149.0	82.6	-291.7
$R_{\rm p}$-A(et)T	-271.6	-253.5	-14.6	-11.5	4.9	57.6	154.5	82.6	-291.6
$S_{\rm p}$-A(et)A	-273.9	-252.9	-15.1	-11.3	4.8	56.5	152.9	82.8	-291.6
$R_{\rm p}$-T(et)T	-279.2	-258.0	-14.1	-11.3	4.7	59.2	149.7	82.5	-292.0
$S_{\rm p}$-T(et)T	-280.0	-257.2	-14.0	-11.4	4.7	58.5	148.9	82.6	-291.9
$R_{\rm p}$-T(et)C	-278.0	-257.8	-14.0	-11.3	4.7	57.8	151.5	82.9	-291.9
$R_{\rm p}$-T(et)C	-277.0	-254.1	-14.6	-11.3	4.8	58.9	148.8	82.3	-291.8
$R_{\rm p}$-C(et)C	-279.2	-258.2	-14.4	-11.4	4.7	58.1	151.1	82.8	-291.7
$R_{\rm p}$-C(et)C	-278.5	-255.6	-14.8	-11.2	4.8	59.2	148.2	82.6	-291.8

$E_{\rm tot}$ is the total minimized energy; $E_{\rm vdW}$ is the sum of all $6-12$ terms; $E_{\rm el}$ is the sum of all $1/(\epsilon r)$ terms; $E_{\rm HB}$ is the sum of the special hydrogen-bondend terms; $E_{\rm bond}$ is the sum of all bond stretching terms; $E_{\rm ang}$ is the sum of all bond angle bending terms; $E_{\rm dih}$ is the sum of all torsion contributions; $E_{\rm 14vdW}$ is the sum of all $6-12$ terms that are separated by only two other atoms; $E_{\rm 12el}$ is similar sum for the electrostatic interactions.

Table III shows the calculated minimum energies for the 16 systems studied with a breakdown into contributions from the terms in the potential function discussed in Sec. II. For all but two of the modification sites, the calculated total minimum energies are lower for the $S_{\rm p}$ substituted oligomers than for the $R_{\rm p}$ substitutions. The $S_{\rm p}$ analogues of G(et)G, G(et)A, A(et)A, A(et)T and T(et)T are more stable than their diastereomeric twins by 0.8 - 4.4 kcal/mol while the oligomers substituted at the "last" two phosphates, T(et)C and C(et)C, show a difference in the reverse direction by 0.8 - 1.0 kcal/mol. The relative stability indicated by the majority of the substitutions thus appears to conform to the experimentally observed stability difference of the duplex oligomers.[22–25] This is not surprising, given that the $R_{\rm p}$ substitution points into the major groove of the helix and would be expected to be disruptive.

The breakdown of the total energies of the ethylated molecules (Table III) shows that the variations in the electrostatic, hydrogen bond, bond stretching terms, as well as the third neighbour terms (labeled 14) are significantly less than are the variations in the total energy. Variations in the Van der Waals terms (6-12 terms), in the angle bending terms, and in the torsion terms are of comparable magnitude to the variations in the total energy. Experimental data indicate that the primary source of lowered T_m is steric perturbation induced by the alkyl group.[28,29,31] The calculations presented here report that this steric interaction between the alkyl group and the DNA introduces a small but significant conformational change in the backbone that was observed both in the C3'-O3'-P-O5' torsion angles and in the torsion contribution to the total energy.

The total "group energies" (i.e. the sum of all interaction energies of each atom in a group) of each sugar, each phosphate, and each base were compared for all 16 molecules considered. The variation in the sugars was generally larger than in the bases. The group energy of the bases changed only 2-5% as a result of the different substitutions, and the magnitudes of the changes relative to the parent octamer were in the order G < A < C < T; the sugars showed variations from 21% to 47% relative to the parent, in the order of G < C < A $\sim$ T. These variations reflect the relative rigidities of the aromatic bases and the sugars and may suggest that guanosine is the most rigid base-sugar complex and thymidine the most flexible.

The intragroup energies (i.e. interactions between the atoms in a given group only) generally show very little variation from compound to compound. This can be taken as an indication of the success of the calculation in maintaining the integrity of the chemical units of the macromolecule. In all cases the intramolecular energy of the ethyl group was found to be 0.1 kcal/mol. Introduction of the ethyl group increases the intramolecular energy of the phosphate group from 0.1 kcal/mol to 1.9 kcal/mol indicating that the ethyl groups have a destabilizing effect on the phosphates. It is not evident as to the cause of this effect; it is not due to the charge neutralization (or alteration in atomic charges, Table I) of the phosphate group since the group energy of the neutralized but not ethylated phosphate (G(no)G) remained 0.1kcal/mol. The range of intramolecular energies of the sugars and bases is very narrow; the extreme values differ from each other by less than 2%. This is in marked contrast with the much wider range of intergroup interactions observed for these groups.

The immediate consequence of phosphotriester formation is a local charge neutralization of the sugar-phosphate backbone. In the calculations, seven counterions were placed along the strand of the parent octamer and six along each modified strand. Comparison of the placement of the counterions subsequent to minimization showed that the counterion placement at the charged phosphates was essentially indistinguishable in all molecules. The contributions of the counterions to both phosphate and sugar group energies did differ, however, depending upon the stereochemisry of the alkylation. Some sequence specificity was also found (*vide infra*).

Table IV Electrostatic Group Interactons for Phosphate with and without Counterion contribution: Differences Between the Alkylated Octamers and Nonalkylated $\{$d-[CGAATTCC]$\}_2$[a]

	GpG		GpA		ApA		ApT		TpT		TpC		CpC	
	TES[b]	-CI[c]	TES	-CI	TES	-CI	TES	-CI	TES	-CI	TES	-CI	TES	-CI
G(no)G	-13.5	2.5	[d]	2.6		0.9								
R_p-G(et)G	-12.9	2.4	[d]	2.6		0.8								
S_p-G(et)G	-12.9	2.4	[d]	2.6		0.7								
R_p-G(et)A		2.4	-12.7	3.6	[d]	2.4				0.8		0.7		
S_p-G(et)A		2.5	-12.7	3.7	[d]	2.3								
R_p-A(et)A		0.9		2.4	-12.6	4.0	0.7	2.5		0.9				0.8
S_p-A(et)A		0.8		2.4	-12.5	4.0	0.7	2.5		0.7				0.7
R_p-A(et)T				0.8		2.4	-12.3	4.7	0.8	3.0		1.6		1.6
S_p-A(et)T				0.8		2.3	-12.2	4.8	0.8	2.8		1.4		1.8
R_p-T(et)T						0.9		2.5	-12.2	5.0	1.0	3.4		1.3
S_p-T(et)T						0.8		2.4	-12.2	5.1	1.1	3.4		1.2
R_p-T(et)C								1.1	0.7	3.0	-12.4	4.8	0.7	2.4
S_p-T(et)C								1.1	1.0	3.4	-12.2	5.1	2.8	0.8
R_p-C(et)C						0.9		1.3		1.5		2.4	-12.7	3.4
S_p-C(et)C						0.8		1.3		1.4		2.5	-12.5	3.6
Parent[a]	-14.1	3.0	-13.7	5.0	-13.6	5.6	-13.3	7.0	-13.4	6.5	13.2	6.5	-13.8	4.6

[a]Numbers in table are $E_{\text{parent}} - E_{\text{modified}}$; only absolute values > 0.6 kcal (kT) are shown; the parent values are included to provide information as to the sign of the energy terms.
[b]TES: total electrostatic intergroup energy, i.e., the sum of all intergroup electrostatic energies.
[c]-CI: TES minus the sum of the electrostatic interactions of the specified group with the group of counterions.
[d]Values $0.5 - 0.6$ kcal discussed in text.

The deviation of total electrostatic energy for each phosphate group of a given modified octamer from that in the parent octamer, in all 15 modified molecules, is found in Table IV. Also found in that table are the deviations of these total group electrostatic energies with the electrostatic contribution from the phosphate-counterion interaction subtracted. Examination of the data for the total electrostatic energy shows that at the site of modification there is an ∼12.5 kcal/mol electrostatic destabilization relative to the parent octamer due to the ethylation, with very little dependence on either the site or the stereochemistry of the triester. The larger destabilization of G(no)G suggests that the presence of the ethyl group provides a small stabilizing influence relative to the effects of pure charge neutralization. There appears to be a slight stabilization of the phosphodiester to the 5' side of the phosphotriester in the alkylated molecules, again suggesting that the presence of the phosphotriester disrupts some (unidentified) unfavorable electrostatic interaction. The counterion-phosphate interaction is ∼15.8 kcal/mol less favorable for a triester than for the corresponding diester (in Table IV, the difference between the columns labeled TES and -CI), indicating that there are 2 – 5 kcal/mol of favorable electrostatic interaction not involving the counterions at the triester

site. Further, this stabilizing factor appears to be greatest towards the center of the chain; this may arise from a sequence specificity to the interaction, or it may simply be due to decreased proximity from strand ends. Also, this stabilizing factor is more delocalized than the immediacy of the counterion electrostatic interaction, as reflected in the (symmetrical) graduated effect on either side of the neutralized phosphate.

Table V Total Group Interactons for Phosphate with and without Counterion contribution: Differences Between the Alkylated Octamers and Nonalkylated $\{$d-[CGAATTCC]$\}_2$[a]

	GpG		GpA		ApA		ApT		TpT		TpC		CpC	
	TGI[b]	-CI[c]	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI
G(no)G	-12.5	2.6		2.5		0.9						0.7		
R_p-G(et)G	-16.5	-0.6	1.0	3.0		0.9						0.7		0.7
S_p-G(et)G	-15.0	-2.1	1.1	3.0		0.9						0.7		
R_p-G(et)A		2.4	-16.5	-1.1	0.8	2.7				1.5				0.8
S_p-G(et)A		2.5	-14.4	0.8	2.9		0.7					1.6		
R_p-A(et)A		1.0		2.4	-16.7	-0.9	0.9	2.8		0.8				0.7
S_p-A(et)A		1.1		2.5	-14.7	1.1	1.1	2.9		0.7				
R_p-A(et)T				0.9		2.6	-16.4	—	0.9	3.0		1.7		1.8
S_p-A(et)T				0.8		2.3	-12.2	4.8	0.8	2.8		1.4		1.8
R_p-T(et)T						0.9		2.4	-14.5	1.8	1.8	4.1		1.6
S_p-T(et)T						0.9		2.4	-13.9	2.5	2.0	4.2		1.3
R_p-T(et)C								1.2		1.7	-15.9	—	0.9	2.9
S_p-T(et)C								1.1		3.8	-13.8	2.6	1.5	3.5
R_p-C(et)C						0.9		1.2		2.1		2.5	-15.6	—
S_p-C(et)C						0.8		1.3		1.4		2.8	-14.5	0.8
Parent[a]	-11.8	4.4	-11.1	6.7	-11.3	7.1	-10.9	8.2	-10.5	8.8	10.7	8.3	-11.3	6.3

[a]Numbers in table are $E_{\text{parent}} - E_{\text{modified}}$; only absolute values > 0.6 kcal (kT) are shown; the parent values are included to provide information as to the sign of the energy terms.
[b]TGI: total intergroup energy, i.e., the sum of all intergroup interaction energies.
[c]-CI: TGI minus the sum all interactions of the given group with the group of counterions.

Table V contains data on the total energy of interaction of the phosphate group with all other groups in the molecule. Again, the data are listed as deviations from the parent octamer, and are tabulated with and without the contributions of the phosphate-counterion interactions. Comparison of all data in both Tables IV and V indicates that although most of the destablization at the site of the alkylation comes from a loss of the stereochemically-independent favorable phosphate-counterion electrostatic interaction, additional instabilities caused by the ethyl group are greater for the R_p-modification than for the S_p, by 0.7-2.0kcal/mol. This difference reflects the trend in the melting behaviour of the experimentally investigated oligomers. Additionally, as the site of the phosphotriester moves down the strand, from the 5' end towards the 3' end, the non-counterion contribution to the

total energy remains more destabilizing for the R_p-alkylations than for the S_p, but these contributions do become lower in energy than those in the unmodified octamer.

Table VI Electrostatic Group interactions for Sugars with and without Counterion contribution: Differences Between the Alkylated Octamers and Nonalkylated $\{$d-[CGAATTCC]$\}_2$[a]

	G[1]		G[2]		A[3]		A[4]		T[5]		T[6]		C[7]		C[8]	
	TES[b]	-CI[c]	TES	-CI	TES	-CI	TES	-CI	TES	-CI	TES	-CI	TES	-CI	TES	-CI
G(no)G	—	-1.5	—	-1.5												
R_p-G(et)G	—	-1.6	-0.7	-1.6												
S_p-G(et)G	—	-1.5	-0.7	-1.7												
R_p-G(et)A				-1.6	-0.7	-1.7										
S_p-G(et)A				-1.5	-0.7	-1.7										
R_p-A(et)A						-1.6	-0.8	-1.7								
S_p-A(et)T						-1.5	-0.8	-1.7								
R_p-A(et)T								-1.6	-0.8	-1.7						
S_p-A(et)A								-1.5	-0.9	-1.8						
R_p-T(et)T										-1.6	-0.9	-1.9				
S_p-T(et)T										-1.5	-1.0	-1.9				
R_p-T(et)C												-1.7	-0.7	-1.6		
R_p-T(et)C												-1.9	-0.9	-1.9		
R_p-C(et)C														-1.4	-0.7	-1.6
R_p-C(et)C														-1.6	-0.8	-1.7
Parent[a]	-0.9	-3.1	-1.8	-4.8	-2.0	-5.3	-2.1	-5.5	-2.7	-6.2	-2.6	-6.1	-2.1	-5.4	-1.7	-4.0
Parent[d]	2.2		3.9		3.2		3.4		3.5		3.5		3.3		2.3	

[a]Numbers in table are $E_{\text{parent}} - E_{\text{modified}}$; only absolute values > 0.6 kcal (kT) are shown; the parent values are included to provide information as to the sign of the energy terms.

[b]TES: total electrostatic intergroup energy, i.e., the sum of all intergroup energies.

[c]-CI: TES minus the sum of the electrostatic interactions of the specified group with the group of counterions.

[d]Electrostatic counterion interaction energy with parent sugar groups.

This rather complex behavior of the phosphotriester groups is mimicked in the behavior of the ^{31}P chemical shifts of both ethyl-phosphotriesters and isopropyl-triesters. Although the suggestion appears in the literature that the absolute stereochemistry at a phosphotriester can be determined by comparison of the relative chemical shift positions of the ^{31}P resonance for the two diasteromers, with the inward-pointing alkyl group being more upfield shifted than the outward-pointing group, we have found that this is not universally the case.[28] The studies reported here should prove helpful in determining the reasons behind the observed ^{31}P chemical shift dispersions.

As stated earlier, the sugar group energies are quite variable when a phosphotriester is present. Less favorable total electrostatic interactions of the sugar groups are partially

compensated for by less positive sugar-counterion electrostatic interactions in the modified oligomers than in the parent. These total electrostatic interactions are of sugars to both the 5'- and 3'-side of the phosphotriester and are essentially independent of site and stereochemistry (Table VI). Most of the unfavorable electrostatic energy of the sugars can be traced to increased energy of interaction of the phosphotriesters with the sugar groups to either side (data not shown). The total energies of the sugar groups adjacent to the phosphotriesters are increased by ∼1 kcal/mol in addition to the increase in electrostatic energy (compare Tables VI and VII, columns "TGI"). These increases are somewhat greater for the R_p-isomers than for the S_p-twins, and do not appear in G(no)G. There is very little destabilization (total energy) difference between the sugars to the 5'-side of the phosphotriester as compared to those to the 3'-side of the phosphotriester, with the exception of the sugars on either side of the R_p-T(et)T triester. The T^6 sugar group has an additional kcal of destabilizing, non-counterion contribution to the total energy; as with other unusual features of the T(et)T oligomers, it was not possible to identify a single source of this energy. Again, interactions with the counterions are partially compensatory for the increase in total energies of the sugar groups adjacent to the modification site. The sugar-counterion interactions (electrostatic and total energy) are relatively independent of stereochemistry and are slightly more favorable when the sugar is to the 5'-side of the phosphotriester (Tables VI, VII and additional data from AMBER analysis).

CONCLUSIONS

Whereas this study does not address the role charge neutralization plays in modulating the interactions of phosphate- alkylated DNAs with drugs and/or proteins, this study does indicate that charge neutralization *by itself* does not lead to significant structural perturbations at the site of alkylation. The reduced charge and the presence of an alkyl moiety do effect both the structure and the net stability of an alkyl-bearing oligomer in rather complex ways. As might be expected, electrostatic contributions to phosphate and sugar group perturbations are relatively insensitive to site and stereochemistry of alkylation, whereas total energies of these groups show more sensitivity to these two parameters. Where a trend is discernible in component energies, the R_p diastereomer (in which the alkyl moiety points into the major groove of the helix) is less stable than the S_p isomer (in which the alkyl moiety points away from the helix), although this does not hold for all total energies. These findings are, for the most part, consistent with NMR studies of phosphate-alkylated analogues of {d-[GGAATTCC]}$_2$.

In closing, it is important to stress that the calculations described here are only a first step in understanding the behaviour of these molecules in aqueous media. Future work should include waters molecules explicitly; in our results the counterions completely "condense" on the helix while the theory of Manning[32,33] predicts that at least 25% should be at some distance from the phosphates. Furthermore, due to the lack of the explicit solvent molecules, hydrophobic interactions (that may be especially important for ethylations next to a thymine) were not included in our calculations. Also, minimization in these calculations corresponds to the state of the system at 0K; accurate characterization at physiological temperatures requires minimizing the free energy of the system. Unfortunately, the inclusion of the solvent itself increases the computational tasks by orders of magnitude (even if the counterions are immobilized) and the calculation of free energy with a fully solvated system

is currently only feasible at selected configurations at large computational expense[34]. Implementation of all these steps thus await further theoretical and computational developments. Nevertheless, our calculations show that even with the level of approximation used here, valuable information can be obtained.

Table VII Total Group Interactions for Sugar with and without Counterion contribution: Differences Between the Alkylated Octamers and Nonalkylated $\{$d-[CGAATTCC]$\}_2$[a]

	G[1]		G[2]		A[3]		A[4]		T[5]		T[6]		C[7]		C[8]	
	TGI[b]	-CI[c]	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI	TGI	-CI
G(no)G	—	-1.4	—	-1.4												
R_p-G(et)G	-1.9	-2.9	-1.7	-2.6												
S_p-G(et)G	-1.5	-2.9	-1.0	-1.8												
R_p-G(et)A			-1.7	-2.8	-1.9	-2.7										
S_p-G(et)A			-1.4	-2.4	-1.4	-2.2										
R_p-A(et)A					-2.2	-3.2	-2.1	-3.0								
S_p-A(et)T					-1.1	-2.1	-1.3	-2.2								
R_p-A(et)T							-2.2	-3.4	-2.3	-3.2						
S_p-A(et)A							-1.7	-2.9	-1.6	-2.5						
R_p-T(et)T									-1.0	-2.2	-2.0	-3.0				
S_p-T(et)T									-1.2	-2.4	-1.1	-2.1				
R_p-T(et)C											-1.4	-2.7	-1.8	-2.6		
R_p-T(et)C											-1.4	-2.7	-1.4	-2.3		
R_p-C(et)C													-0.7	-1.7	-2.0	-2.8
R_p-C(et)C													—	-2.2	-1.5	-2.3
Parent[a]	-0.2	-2.2	-3.0	-5.9	-3.2	-6.3	-3.7	-7.1	-2.6	-6.1	-1.6	-5.1	-1.3	-4.5	-1.4	-3.6
Parent[d]	2.2		2.9		3.2		3.4		3.5		3.5		3.3		2.3	

[a]Numbers in table are $E_{\text{parent}} - E_{\text{modified}}$; only absolute values > 0.6 kcal (kT) are shown; the parent values are included to provide information as to the sign of the energy terms.

[b]TGI: total group energy. i.e., the sum of all intergroup energies.

[c]-CI: TGI minus the sum of all interactions of the given group with the group of counterions.

[d]Total counterion energy with parent sugar groups.

Support for this work was provided by an RCMI grant from NIH # RR-03037 and NIH grant # CA 46713 (MSB). Prof G.J. Quigley is thanked for his help both with the AMBER program and for many discussions and Paula A. Longo for her patient operation of the picture system. We are also grateful to Prof. P. Kollman for making the AMBER program package available to us and to Dr. H. Berman for the DOCK program.

REFERENCES

1. Razin, A., Cedar, H., and Riggs, A.D., in "DNA Methylation, Biochemistry and Biological Significance", Chapter 1, Razin, A., Cedar, H., and Riggs, A.D., Eds., Springer-Verlag, New York, 1984.

2. Lawley, P.D., in "Progress in Nucleic Acid Research and Molecular Biology, Volume 5", Chapter 2, Davidson, J.N. and Cohn, W.E., Eds., Academic Press, New York, 1966.

3. Singer, B. and Grunberger, D., "Molecular Biology of Mutagens and Carcinogens", Plenum Press, New York, 1983.

4.*Ibid.*, Chapter 9.

5. Lett, J.T., Parkins, G.M., and Alexander, P., (1962) *Archives Biochem. Biophys.*, **97** 80-93.

6. Rhaese, H.-J. and Freese, E., (1969) *Biochim. Biophys. Acta*, **190** 418-432.

7. Bannon, P. and Verly, W., (1972) *Eur. J. Biochem.*, **31**, 103- 111.

8. Sun, L. and Singer, B., (1975) *Biochemistry*, **14**, 1795-1802.

9. Jensen, D.E. and Reed, D.J., (1978) *Biochemistry*, **17**, 5098- 5107.

10. Jensen, D.E., (1978) *Biochemistry*, **17**, 5108-5113.

11. Miller, P.S., Chandrasegaran, S., Dow, D.L., Pulford, S.M., Kan, L.-S., (1982) *Biochemistry*, **21**, 5468-5474.

12. Briscoe, W.T. and Cotter, L.-E., (1984) *Chem.-Biol. Interactions* **52**, 103-110.

13. Weinfeld, M. and Livingston, D.C., (1986) *Biochemistry*, **25**, 5083-5091.

14. Swenson, D.H. and Lawley, P.D.(1978), *Biochem. J.*, **171**, 575- 587.

15. Brennan, R.G., Kondon, N.S., and Sundaralingam, M., (1984) *J. Am. Chem. Soc.*, (1984) **106**, 5671-5676.

16.Singer, B., Sun, L., and Fraenkel-Conrat, H., (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 2232-2236.

17. Abbondandolo, A., Dogliotti, E., Lohman, P.H.M., and Berends, F., (1982) *Mutation Research* **92**, 361-377.

18. Shooter, K.V., Howse, R., and Merrifield, R.K., (1974) *Biochem. J.*, **137**, 313-317.

19. Singer, B., Spengler, S., and Bodell, W.J., (1981) *Carcinogenesis*, **2**, 1069-1073.

20. Miller, P.S., Barrett, J.C., and Ts'o, P.O.P., Biochemistry (1974) **13**, 4887-4895.

21. Stec, W.J., Zon, G., Gallo, K., and Byrd, R.A., (1985) *Biochem, Biophys. Res. Commun.*, **26**, 2191-2194.

22. Van Genderen, M.H.P., Koole, L.H., Aagaard, O.L., Van Lare, C.E.J. & Buck, H.M., (1987) *Biopolymers*, **26**, 1447-1461.

23. Weiner, S.J., Kollman, P.A, Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr. & Weiner, P.K., (1984) *J. Am. Chem. Soc.*, **106**, 765-784.

24. Weiner, P.K. & Kollman, P.A., (1981) *J. Comp. Chem.*, **2**, 287- 303.

25. The DOCK program was developed under the direction of H. Berman at the Fox Chase Cancer Center (Philadelphia) by R. Stodola, F. Manion, W.P. Wood and S. Beckman.

26. Dickerson, R.E., Drew, H.R., Conner, B.N., Wing, R.M., Fratini, A.V., & Kopka, M.L., (1982) *Science*, **216**, 475-485.

27. Lawrence, D.P., Chen, W., Zon, G., Stec, W.J., Uznanski, B, & Broido, M.S., (1987) *J. Biomol Struct. Dyn*, **4**, 757-781.

28. Broido, M.S., Lawrence, D.P., Chen, W., Zon, G., (1987) in "Structure & Expression, Vol. 2: DNA and Its Drug Complexes", Sarma, R.H., & Sarma, M.H., eds., Adenine Press.

29. Summers, M.F., Powell, C., Wilson, W.D., & Zon, G., (1986) *Nucleic Acid Res.*, **14**, 7421-7436.

30. Cahill, S.M. & Broido, M.S., to be published.

31. LaPlanche, L.A., James, T.L., Powell, C., Wilson, W.D., Uznanski, B., Stec, W.J., Summers, M.F., & Zon, G., (1986) *Nucleic Acids Res.*, **14**, 7421-7436.

32. Manning, G.S. (1969) *J. Chem. Phys.*, **51**, 924-933; Manning, G.S., (1981) *J. Phys. Chem.*, **85**, 870-877.

33. Record, Jr., M.T. & Lohman, T.M., (1978) *Biopolymers*, **17**, 159-166.

34. Mezei, M. & Beveridge, D.L., (1986) *Ann. Acad. Sci N.Y.*, **482**, 1-23.

Received November 30, 1988
Accepted April 12,, 1989