Spectroscopic and Thermodynamic Characterization of the Interaction of N^7-Guanyl Thioether Derivatives of d(TGCTG*CAAG) with Potential Complements†

Magnus Persmark and F. Peter Guengerich*

Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146

Received October 19, 1993; Revised Manuscript Received May 23, 1994®

ABSTRACT: The oligomer d(TGCTGCAAG) corresponds to a region of bacteriophage M13mp18 DNA where mutations have been found to be induced by *S*-(2-chloroethyl)glutathione (glutathione, GSH) [Cmarik, J. L., Humphreys, W. G., Bruner, K. L., Lloyd, R. S., Tibbetts, C., & Guengerich, F. P. (1991) *J. Biol. Chem. 267*, 6672–6679]. This oligomer was prepared with the central G replaced by *S*-(2-N^7-guanylethyl)-GSH or *N*-acetyl-*S*-(2-N^7-guanylethyl)Cys methyl ester; these derivatives were purified by HPLC and by affinity chromatography in the latter case. UV mixing and CD spectroscopy studies showed no evidence for preferred pairing of the *S*-(2-N^7-guanylethyl)GSH moiety to any base other than C. UV melting studies of duplexes were performed with complementary strands containing the normal C, as well as the three mismatches (T, A, and G), across from the adducted base. Thermal stabilities were reduced in all cases when G was replaced by either N^7-guanyl adduct; the C-containing complement was still the most stable. The reduced stability of the duplex d(TGCTG*CAAG)/d(CTTGCAGCA), where *S*-(2-N^7-guanylethyl)-GSH corresponds to G*, was characterized by an increase in $\Delta G°$ of 1.4–2.0 kcal mol^{-1} (in the range of 25–37 °C) relative to the unadducted duplex. van't Hoff analysis of concentration-dependent melting experiments indicated that the $\Delta H°$ of the duplex was actually more favorable when this adduct was introduced ($\Delta\Delta H°$ = 13 kcal mol^{-1}), but the decreased thermal stability was due to the entropic component. Similar results were observed when G* was *N*-acetyl-*S*-(2-N^7-guanylethyl)Cys methyl ester. Under the conditions used, the overall relative stabilities of the oligomeric duplexes containing various base pairs do not indicate that *S*-(2-N^7-guanylethyl)GSH would contribute to a higher frequency of T misinsertion than G. The possibility that ionization at the guanine N1 position may be involved in mutagenesis by N^7-guanyl adducts is considered.

Carcinogenic chemicals often exert genotoxic effects by irreversibly binding to DNA to alter the fidelity of replication and transcription (Miller, 1970; Heidelberger, 1975; Lawley, 1984); not only mutations but also more complex clastogenic events result from these lesions (Brusick, 1989). Even though much is known about specialized and unusual DNA structures (Blaho & Wells, 1989; Crothers *et al.*, 1990; Roberts & Crothers, 1991), there is a paucity of knowledge regarding molecular mechanisms of how mutations are introduced by the chemical modification of DNA. One of the problems in this area is that physical studies are not often done with the same sequences that show mutations. Considerable selectivity is often seen as to which positions in a DNA chain yield mutations, even when the same adduct is present at different sites (Bigger *et al.*, 1989; Richardson & Richardson, 1990; Singer & Essigmann, 1991; Cheng *et al.*, 1992). One reason that a limited number of studies has been done is that producing oligonucleotides in which only one base of a series of the same is modified may be difficult, particularly if a strategy is required in which modification must be done after the oligonucleotide has been synthesized. This has been the case with many interesting carcinogens, including aflatoxin B_1 *exo*-8,9-oxide (Gopalakrishnan *et al.*, 1989) and some of the polycyclic hydrocarbon dihydrodiol epoxides (Hruszkewycz & Dipple, 1991; Mao *et al.*, 1992; Cosman *et al.*, 1992). Separation of positional isomers of adducted oligonucleotides is usually a challenging task, and few syntheses of this type have been done. Fuchs and his associates were able to separate the three positional isomers after treating an oligonucleotide containing a run of three guanines with an *N*-hydroxyarylamine; they found that these individual derivatives had markedly different biological properties (Koehl *et al.*, 1989; Seeberg & Fuchs, 1990).

Our laboratory has been interested in the carcinogen 1,2-dibromoethane (ethylene dibromide) and its ability to produce mutations in bacterial systems. The activation of this chemical occurs through its conjugation with GSH[1] to form a half-mustard that can react with DNA to form *S*-(2-N^7-guanylethyl)GSH (Chart 1), which accounts for >95% of the DNA adducts (Ozawa & Guengerich, 1983; Inskeep & Guengerich, 1984; Koga *et al.*, 1986; Inskeep *et al.*, 1986; Peterson *et al.*, 1988; Cmarik *et al.*, 1990; Humphreys *et al.*, 1990; Kim & Guengerich, 1990). The dominance of G:C to A:T transitions in the mutations is consistent with a role for the lesion *S*-(2-

† This research was supported in part by U.S. Public Health Service Grants CA44353 and ES00267. M.P. was supported by U.S. Public Health Service Training Grant ES07028 and Fellowship ES05592.

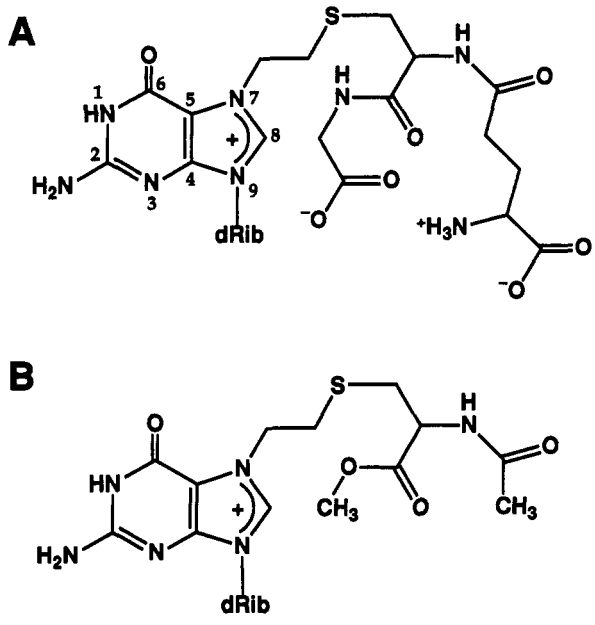
* Address correspondence to this author.

® Abstract published in *Advance ACS Abstracts,* July 1, 1994.

[1] Abbreviations: EDTA, (ethylenedinitrilo)tetraacetic acid; GSH, glutathione; G*, N^7-alkylguanine adduct; G^{GSH}, an oligonucleotide containing *S*-(2-N^7-guanylethyl)GSH; G^{Cys}, an oligonucleotide containing *N*-acetyl-*S*-(2-N^7-guanylethyl)Cys methyl ester; HPLC, high-performance liquid chromatography; T_m, melting temperature (maximum of the first derivative of the temperature-induced melting curve); t_R, retention time; Tris, tris(hydroxymethyl)aminomethane. All oligonucleotide sequences are written in the direction 5′ to 3′ unless specifically noted otherwise.

0006-2960/94/0433-8662$04.50/0

Chart 1: Structures of the Adducts
S-(2-N^7-Guanylethyl)GSH (A) and
N-Acetyl-*S*-(2-N^7-guanylethyl)Cys Methyl Ester (B)



N^7-guanylethyl)GSH, although direct proof for this event has not been obtained (Foster *et al.*, 1988; Humphreys *et al.*, 1990). Biological selectivity may be influenced considerably by both adduct and local DNA structure; for instance, the ratio of *Salmonella typhimurium* TA100 revertants to N^7-guanyl DNA adducts varied considerably among the series of substituted Cys-containing half-mustards used (Humphreys *et al.*, 1990), and within the *lacZ* α complementation region the patterns of N^7-guanyl adducts and mutants are not coordinate (Cmarik *et al.*, 1992). Thus, the nature of the sequence may be critical in determining whether or not a specific adduct will produce a mutation in a particular biological setting.

We previously prepared the oligonucleotides d(ATGCAT) and d(CATGCCT), with *S*-(2-N^7-guanylethyl)GSH substituted for guanine, and examined some of the physical characteristics (Oida *et al.*, 1991; Kim & Guengerich, 1993). The lesion was found to considerably disrupt duplex formation as judged by T_m, and there was no evidence that such a modification led to a more stable pairing to a base other than C (expected to be T on the basis of the mutation work). Nevertheless, the point can be raised that these oligonucleotides have not been examined as to whether mutations would actually occur if the sequences were modified in DNA, and it would be desirable to carry out physical studies with a sequence known to be subject to mutations. Recently, we found that different N^7-guanyl adducts [in d(CATGCCT)] vary considerably in their effects on duplex stability (Kim & Guengerich, 1993). Ezaz-Nikpay and Verdine (1992) recently reported that substitution of N^7-methylguanine:C for G:C in a self-complementary dodecamer made the $\Delta H°$ of duplex formation significantly more favorable, although $\Delta S°$ was less favorable. We were interested in the applicability of this study to our own work and to general conclusions about the effect of the positively charged imidazole ring of N^7-guanyl adducts on DNA physical parameters. In particular, the main purpose of the present study was to determine whether the base-pairing properties of *S*-(2-N^7-guanylethyl)GSH are altered in a sequence where mutations have been found.

We synthesized the *S*-(2-N^7-guanylethyl)GSH and *N*-acetyl-*S*-(2-N^7-guanylethyl)Cys methyl ester derivatives of d(TGCTG*CAAG) (modified position marked with *) (Chart 1). These modified oligonucleotides were examined by UV and CD spectroscopy. In addition, the thermodynamic parameters of duplex formation were investigated by concentration-dependent thermal denaturation with regard to both the complementary oligonucleotide and a derivative with T opposite the adduct, corresponding to the most frequently observed mutation.

EXPERIMENTAL PROCEDURES

Chemicals. GSH and human semen prostatic acid phosphatase (EC 3.1.3.2) were obtained from Sigma Chemical Co. (St. Louis, MO). 1-Bromo-2-chloroethane and tetra-*n*-butylammonium hydroxide were obtained from Aldrich Chemical Co. (Milwaukee, WI). Cyanoethyl phosphoramidites were obtained from Applied Biosystems (Foster City, CA). 3′-Amino modifier controlled pore glass (CPG) was obtained from Glen Research (Sterling, VA). Affi-Gel 10 was obtained from Bio-Rad Laboratories (Richmond, CA). Octadecylsilane (C18) SepPak cartridges were from Waters-Millipore (Medford, MA).

Chromatography. HPLC was carried out with a SpectraPhysics SP8700 or SP8800 system attached to a Milton Roy UV detector (Thermo Separation Products, Piscataway, NJ). Semipreparative reverse-phase separation was performed on a Beckman Ultrasphere octadecylsilane (C18) column (10 × 250 mm, 5 μm, Beckman, San Ramon, CA) at a flow rate of 2.5 mL min^{-1}. Semipreparative ion-exchange HPLC was performed using a Phenomenex W-Porex DEAE column (10 × 250 mm, 10 μm, Phenomenex, Torrance, CA) at a flow rate of 3.0 mL min^{-1}. Analytical reverse-phase and ion-pair HPLC were performed on a Beckman Ultrasphere C18 column (4.6 × 250 mm, 5 μm) at a flow rate of 1.0 mL min^{-1}.

Oligonucleotide Synthesis and Purification. Deoxyribooligonucleotides were synthesized by standard automated solid-support chemistry on an Applied Biosystems Model 391 synthesizer using cyanoethyl phosphoramidites. Oligomers were deprotected in concentrated NH_4OH at 55 °C for 17–24 h, after which NH_3 was removed *in vacuo*. Crude oligomers were purified by semipreparative reverse-phase HPLC using a gradient of 7–20% (v/v) CH_3CN in 0.10 M $NH_4CH_3CO_2$ buffer (pH 7.0) over 40 min. Oligomers were then depleted of salt by gel filtration on a 1.5 × 30 cm column of Sephadex G10 (Pharmacia, Piscataway, NJ, equilibrated in H_2O) and stored frozen at −20 °C.

Synthesis and Purification of d(TGCTG^{GSH}CAAG). d(TGCTGCAAG) contains a central guanine residue, which has been identified as a site of mutation by *S*-(2-chloroethyl)GSH (Cmarik *et al.*, 1992). Oligomers containing an *S*-(2-N^7-guanylethyl)GSH residue at the central guanine position—d(TGCTG^{GSH}CAAG)—were prepared as follows. *S*-(2-Chloroethyl)GSH was synthesized and adducted oligomers were formed as described (Humphreys *et al.*, 1990; Oida *et al.*, 1991). Briefly, d(TGCTGCAAG) (typically 6 μmol in 3 mL of 0.50 M potassium phosphate buffer, pH 7.0) was treated with the half-mustard *S*-(2-chloroethyl)GSH (~150 mg/μmol oligomer) for 30 min at 37 °C. The solution was then cooled on ice, depleted of salt by gel filtration at 4 °C as above, and dried by lyophilization. Since the sequence TGCTGCAAG contains three guanine residues, seven N^7-guanyl adduct species are possible. The initial step in the separation of the desired d(TGCTG^{GSH}CAAG) was semipreparative reverse-phase HPLC, using a mobile phase

consisting of mixtures of the two solvents, (A) 0.10 M $NH_4CH_3CO_2$ (pH 7.0) and (B) CH_3CN. Chromatograms were developed with a gradient of time (min)/%B: 0/7; 5/7; 40/10; 45/20; 47/7. Several peaks appeared after d(TGCTGCAAG) ($t_R \sim 18$ min) and were collected. To identify the fraction containing d(TGCTG^{GSH}CAAG), each peak was collected, and an aliquot ($\sim$0.1 A_{260} unit) was treated with hot piperidine (Mattes *et al.*, 1986), lyophilized twice, and treated with 1 unit of prostatic acid phosphatase in 60 mM sodium acetate buffer (pH 5.0) containing 0.1 mM $ZnCl_2$. Fragments were analyzed by analytical ion-pair HPLC using a mobile phase consisting of mixtures of the two solvents, (A) 50 mM potassium phosphate buffer (pH 7.0) containing 4 mM tetra-*n*-butylammonium hydroxide and (B) CH_3CN, and a gradient of time (min)/%B: 0/18; 5/18; 23/30; 28/30; 30/18. Elution profiles were compared with synthetic oligonucleotides corresponding to the cleavage products of each monoadducted 9-mer:

$$TG^*CTGCAAG \rightarrow T + CTGCAAG$$

$$TGCTG^*CAAG \rightarrow TGCT + CAAG$$

$$TGCTGCAAG^* \rightarrow TGCTGCAA$$

The d(TGCTG^{GSH}CAAG) fraction [$t_R \sim 27$ min, yield $\sim$8% (i.e., 0.5 μmol from 6 μmol of starting material)] was concentrated by lyophilization and further purified by semipreparative ion-exchange HPLC using the two buffers, (A) 0.40 M $NH_4CH_3CO_2$ and (B) 1.2 M $NH_4CH_3CO_2$, both pH 7.0 and containing 30% (v/v) CH_3OH. Chromatograms were developed with a gradient of time (min)/%B: 0/5; 5/5; 22/30; 28/80; 32/5. Following the removal of salt as above, the main peak ($t_R \sim 18$ min, yield $\sim$21%) was lyophilized and analyzed for purity by ion-pair HPLC *(vide supra)*. The purity at this stage was $\sim$87%. A final purification step involved semipreparative reverse-phase HPLC, using a shallow gradient prepared by mixing (A) 0.10 M $NH_4CH_3CO_2$ (pH 7.0) containing 5% (v/v) CH_3CN and (B) 0.10 M $NH_4CH_3CO_2$ (pH 7.0) containing 50% (v/v) CH_3CN, with a gradient of time (min)/%B: 0/2; 5/2; 35/7; 40/34; 42/2. d(TGCTG^{GSH}CAAG) eluted at $t_R \sim 34$ min, and purity was analyzed by ion-pair HPLC as above. The yield was $\sim$50%, following desalting on Sephadex G10 as above.

Synthesis of d(TGCTG^{Cys}AAG). *N*-Acetyl-*S*-(2-bromoethyl)Cys methyl ester was synthesized as described (van Bladeren *et al.*, 1981), and the identity and purity of the product were established by NMR and mass spectrometry. Adducts were formed by reacting this half-mustard with d(TGCTGCAAG) (typically 4 μmol of oligomer and $\sim$15 mg of half-mustard/μmol of oligomer) in 60% (v/v) aqueous CH_3OH for 2.5 h at 0 °C (Oida *et al.*, 1991). Following the removal of CH_3OH *in vacuo* and lyophilization, the mixture of adducts was separated by semipreparative reverse-phase HPLC, as described above for the derivatives formed with *S*-(2-chloroethyl)GSH. The mobile phase consisted of mixtures of (A) 0.10 M $NH_4CH_3CO_2$ buffer (pH 7.0) containing 5% (v/v) CH_3CN and (B) CH_3CN. Samples were developed with a gradient of time (min)/%B: 0/4.2; 2.5/4.2; 30/12.6; 40/31.6; 45/31.6; 47/4.2. Peaks were collected and oligonucleotides were identified by a piperidine and phosphatase treatment as above. A well-separated peak ($t_R \sim 20$ min) contained an equimolar mixture of d(TGCTG^{Cys}CAAG) and d(TGCTGCAAG^{Cys}) (yield $\sim$22%). These two positional isomers were not separable by conventional HPLC methods.

Preparation of the Affinity Column. Our approach to the separation of the two positional isomers was based on the assumption that d(TGCTGCAAG^{Cys}) would form a more stable duplex than d(TGCTG^{Cys}CAAG) with the complement d(CTTGCAGCA). 3′-Amino linker CPG was used to introduce a primary amino group at the 3′-terminus of the partially complementary d(TTGCAGCAA) directly in the oligonucleotide synthesizer. Following deprotection *(vide supra)*, NH_2-d(TTGCAGCAA) was purified by analytical reverse-phase HPLC. The mobile phase consisted of (A) 0.10 M $NH_4CH_3CO_2$ buffer (pH 7.0) containing 5% (v/v) CH_3CN and (B) CH_3CN. Samples were developed with a gradient of time (min)/%B: 0/1.1; 5/1.1; 25/10; 27/15; 30/1.1. Following desalting and repeated lyophilization, NH_2-d(TTGCAGCAA) ($t_R \sim 18$ min) was immobilized on Affi-Gel 10, essentially as specified by the manufacturer and Ghosh and Musso (1987). Purified NH_2-d(TTGCAGCAA) (0.6 μmol, $\sim$50 A_{260} units) in 4 mL of 0.20 M sodium 4-(2-hydroxyethyl)-1-piperazineethanesulfonate buffer (pH 7.7) (or 0.10 M sodium carbonate buffer, pH 9.0) was added to 1.0 mL of washed Affi-Gel 10 in a 50-mL conical polypropylene tube. The mixture was mixed vigorously on a rotary shaker at 4 °C overnight. The gel was then transferred to a small column, washed extensively, and stored in 1.0 M NaCl at 4 °C. Coupling yields were about 35%, and it has been shown that this method leads to a high level of end-coupling (Ghosh & Musso, 1987). However, the hybridization capacity, or effective yield, was only approximately 10% of this value.

Separation of d(TGCTG^{Cys}CAAG) and d(TGCTGCAAG^{Cys}). About 42 nmol (3.5 A_{260} units) of the mixture of the two oligonucleotides [in 1.0 mL of 50 mM potassium phosphate buffer (pH 7.0) containing 1.0 M NaCl] was loaded on the oligoaffinity column (capacity, $\sim$1.6 A_{260} units under these conditions) at 4 °C and incubated for 5 min. Unbound oligomers (flow-through fraction) were eluted by washing with 5× 1.0 mL of 1.0 M NaCl followed by 4× 1.0 mL of 0.4 M NaCl to remove weakly bound material. Following transfer of the column to room temperature and equilibration, bound oligomers were eluted with 6× 1.0 mL of H_2O. A_{260} was measured and oligomer-containing fractions were pooled and desalted on C18 SepPak cartridges. Flow-through material was pooled, lyophilized, and passed twice more over the affinity column as above, except that the amounts loaded in the second and third rounds were $\sim$6 and $\sim$10 A_{260} units, respectively. The identity and purity of adduct mixtures were analyzed by analytical ion-pair HPLC after piperidine cleavage *(vide supra)*. The mobile phase consisted of mixtures of (A) 50 mM potassium phosphate buffer (pH 7.0) containing 4 mM tetra-*n*-butylammonium hydroxide and 5% CH_3CN (v/v) and (B) CH_3CN and was developed with a gradient of time (min)/%B: 0/7.4; 2/7.4; 12/11.6; 32/27.3; 35/36.7; 37/36.7; 40/7.4. After the first pass, the resulting material contained 77% (50% yield) d(TGCTG^{Cys}CAAG), and this value increased to 88% ($\sim$66% yield) and 91% ($\sim$63% yield) after the two additional steps, respectively. A final purification step involved HPLC on an analytical reverse-phase C18 column as above using a gradient composed of mixtures of (A) 0.10 M $NH_4CH_3CO_2$ buffer (pH 7.0) containing 5% (v/v) CH_3CN and (B) CH_3CN [time (min)/%B: 0/2.2; 5/2.2; 25/13; 30/2.2]. The desired product ($t_R \sim 18$ min) constituted about 60% of the eluted material, as degradation products and other contaminants were also concentrated by this method, giving an overall yield of $\sim$3%.

Spectroscopy. UV spectra and absorbance measurements were recorded using a modified Cary-14/OLIS spectrophotometer (On-Line Instrument Systems, Bogart, GA). Temperature control was obtained through a jacketed cell holder

connected to a Neslab water bath (Neslab, Newington, NH) controlled by the OLIS computer. CD spectra were recorded using a JASCO J-720 spectropolarimeter (Japan Applied Spectroscopic Co., Tokyo). CD spectra were obtained in 50 mM potassium phosphate buffer (pH 7.0) containing 50 μM EDTA at a total strand concentration of 13 μM, at temperatures of 4, 15, and 30 °C.

Duplex Preparations. Concentrations were estimated from the following extinction coefficients (ϵ, mM^{-1} cm^{-1}), calculated according to Borer (1975): d(TGCTGCAAG), 84.7; d(CTTGCAGCA), 83.0; d(CTTGTAGCA), 86.3; d(CTTGGAGCA), 86.9; d(CTTGAAGCA), 88.8. As the effect of introducing an N^7-guanyl adduct into the sequence TG*C is unknown, we used a value of ϵ = 84.7 mM^{-1} cm^{-1} for the alkylated oligonucleotides as well, although the extinction coefficient for N^7-methylguanosine has been reported to be significantly lower than that of guanosine (Möller *et al.*, 1981). We therefore assumed that the change in the composite ϵ upon the introduction of an N^7-guanyl adduct would be negligible. Stock solutions of oligomers were stored in H_2O at −20 °C. Mixing experiments were performed to verify that the largest hypochromicity was obtained with a 1:1 ratio of d(TGCTG^{GSH}CAAG) and its complement d(CTTGCAGCA). In addition, mixing experiments served to determine pairing preferences of d(TGCTG^{GSH}CAAG) with d(CTTGCAGCA) compared with the three mismatches at the central base. These experiments were carried out at a constant 13 μM total strand concentration, while molar ratios of respective strands were varied from 1:0 to 0:1 in 50 mM potassium phosphate buffer (pH 7.0) containing 50 μM EDTA at 15 °C. Absorbance values were plotted against molar ratios at 254, 260, and 270 nm.

Thermodynamic Studies. Changes in A_{260} were measured in a Cary 14/OLIS spectrophotometer as a function of temperature for oligomer mixtures of total strand concentrations ranging from 1.63 to 104 μM in stoppered cuvettes of 2.0 or 10 mm path length, respectively. Immediately prior to analysis, samples were incubated at 37 °C for 5 min and then slowly cooled. The heating rate was ~0.85 °C min^{-1}, and melting curves were derivatized incrementally and smoothed with the software provided by the supplier. As adduct-bearing oligonucleotides had to be used in several experiments due to the low amounts available, samples were repurified by reverse-phase HPLC after use.

Thermodynamic values were obtained from the concentration dependence of temperature-induced melting using the van't Hoff method (Albergo *et al.*, 1981; Marky & Breslauer, 1987; Puglisi & Tinoco, 1989). By assuming a two-state model for melting, the equilibrium constant K for two non-self-complementary oligonucleotides can be written as

$$K = 2\theta/C_t(1-\theta)^2$$

where θ is the fraction of strands in the duplex form and C_t is the total strand concentration (Marky & Breslauer, 1987). We define T_m as the temperature where the first derivative of the melting curve reaches its maximum. It has been shown that the value of θ at this point is 0.414 (Gralla & Crothers, 1973). By assuming that the values of $\Delta S°$ and $\Delta H°$ are independent of temperature, the relationship

$$\Delta G° = \Delta H° - T\delta S° = -RT \ln K$$

can be rearranged to

$$1/T_m = (R/\Delta H°) \ln C_t + (\Delta S° - R \ln 2.356)/\Delta H°$$

where R = 1.987 cal deg^{-1} mol^{-1} = 8.314 J deg^{-1} mol^{-1}. Thus, the transition entropy $\Delta S°$ and transition enthalpy $\Delta H°$ can be obtained from a plot of ln C_t versus $1/T_m$. Each van't Hoff plot included at least 12 data points and was analyzed by linear regression. The standard deviations in Table 2 were calculated essentially as described by Turner and his associates [SantaLucia *et al.* (1991a,b) and references cited therein]. The error limits correspond to the standard deviations in the slopes (m) and y-intercepts (b) of the van't Hoff plots. The variances $\sigma^2_{\Delta H°}$, $\sigma^2_{\Delta S°}$, and $\sigma^2_{\Delta G°}$ for $\Delta H°$, $\Delta S°$, and $\Delta G°$, respectively, were calculated with the following equations, and each standard deviation was obtained by taking the square root of the variance.

$$\sigma^2_{\Delta H°} = \left(-\frac{R}{m^2}\right)^2 \sigma^2_m$$

$$\sigma^2_{\Delta S°} = R^2\left\{\frac{1}{m^2}\sigma^2_b + \left(-\frac{b}{m^2}\right)^2 \sigma^2_m - 2\left(\frac{b}{m^3}\right)\sigma^2_{bm}\right\}$$

$$\sigma^2_{\Delta G°} = \left(\frac{RT}{m}\right)^2 \sigma^2_b + \left(-\frac{R}{m^2} + \frac{RTb}{m^2}\right)^2 \sigma^2_m - 2\left(\frac{RT}{m}\right)\left(-\frac{R}{m^2} + \frac{RTb}{m^2}\right)\sigma_{bm}$$

σ^2_m, σ^2_b, and σ_{bm} correspond to the variance in the slope and the y-intercept and the covariance of these, respectively, which were calculated using standard statistical methods (Meyer, 1975).

pK_a Determination. In order to estimate the pK_a of the N^7-alkylguanine N1 proton when present in an oligonucleotide, spectra were recorded between 240 and 340 nm at pH values ranging from 5.5 to 11.5 at intervals of 0.25–0.5, in the following buffers: potassium acetate (pH 5.5); potassium phosphate (pH 5.75–7.5); Tris-HCl (pH 7.75–8.75); glycine–KOH (pH 9.0–11.5). A 200-μL aliquot of an aqueous solution of d(TGCTGCAAG) or of the mixture d(TGCTG^{Cys}CAAG)/d(TGCTGCAAG^{Cys}) (~1:1 ratio) was mixed with 200 μL of 0.20 M buffer containing 0.20 M KCl, giving a final A_{260} close to 1.0 for both oligonucleotides. In the case of the alkylated oligonucleotide, two independent sets of measurements were obtained. To minimize degradation of the N^7-guanine residue at the higher pH values, samples were prepared immediately prior to analysis. After inspection of plots of A versus pH, A_{267} values were chosen to plot titration curves. In order to compensate for pipeting and baseline errors, absorbances at the isosbestic points (245.5, 256, and 280.5 nm for unmodified 9-mer and 245.5 and 280.5 nm for alkylated 9-mer, respectively) were normalized relative to the average value, and each A_{267} reading was corrected with this factor (0.98–1.02). Standard deviations were 0.8–1.3% from the mean in all cases.

RESULTS

Synthesis and Purification of Oligomers. Guanine N^7-adducts are sensitive to acidic and alkaline conditions, as well as to heat. This class of adducts therefore must be introduced into oligonucleotides following the completion of synthesis and deprotection. Since the target 9-mer contains three guanine residues, a mixture of seven mono-, di-, and triadducted species is expected after reaction with the half-mustard. Chromatographic separation of positional isomers of adducted oligonucleotides was previously accomplished by Koehl *et al.* (1989) in work with oligonucleotides containing acetylaminofluorene adducts. We were able to successfully separate d(TGCTG^{GSH}CAAG) by a combination of conventional HPLC procedures. The product was obtained in >95% purity

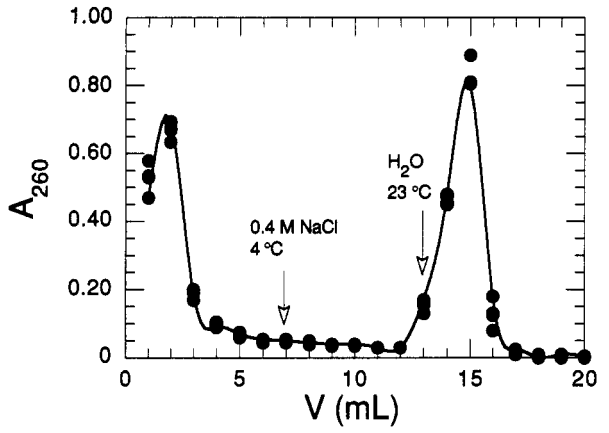


FIGURE 1: Elution of oligonucleotides from an affinity column with the sequence 5′-d(TTGCAGCAA)-NH–spacer–gel. A 1:1 mixture of d(TGCTG^{Cys}CAAG) and d(TGCTGCAAG^{Cys}) was applied in 1.0 M NaCl at 4 °C, and elution proceeded as described in the figure and Experimental Procedures. d(TGCTG^{Cys}CAAG) was obtained in the flow-through fractions, while d(TGCTGCAAG^{Cys}) was preferentially hybridized. Data from four consecutive runs are shown.

after three steps. Approximately 60 nmol (4.5 A_{260} units) was obtained from 6 μmol of d(TGCTGCAAG), corresponding to an overall yield of ~1%.

Conventional chromatographic methods proved insufficient to resolve d(TGCTG^{Cys}CCAAG) from positional isomers. d(TG^{Cys}CTGCAAG) was cleanly separated from d(TGCTG^{Cys}CAAG and d(TGCTGCAAG^{Cys}), but the latter two products eluted as one peak under a variety of HPLC methods and conditions. An attempt to remove the d(TGCTGCAAG^{Cys}) enzymatically by cleaving the unwanted oligomer with the restriction enzyme *Fnu*4HI (which has a recognition site in the unmodified oliomer when bound to its complement) was only partially successful (results not presented). When the oligonucleotide mixture of d(TGCTG^{Cys}CAAG) and d(TGCTGCAAG^{Cys}) was paired with the complementary 9-mer d(CTTGCAGCA), we did not observe a biphasic melting curve of A_{260} versus temperature (results not shown). Thus, detailed characterization of the behavior of d(TGCTG^{Cys}CAAG) in duplexes was not possible with the mixture.

Our next approach to resolve the oligomers was based on the assumption that the presence of an adduct in the central position would selectively destabilize a duplex containing a partially complementary strand. Hybridization to the immobilized oligonucleotide d(TTGCAGCAA) leaves the N^7-guanyl adduct position of d(TGCTGCAAG^{Cys}) unpaired. Thus, the duplex with G^{Cys} in the central position should be less stable than that with G^{Cys} at the end; the latter oligonucleotide was therefore expected to preferentially hybridize to the affinity column.

Under appropriate conditions (*i.e.*, high salt, low temperature, and an amount of isoform mixture far exceeding the binding capacity of the column), the affinity column was highly selective for d(TGCTGCAAG^{Cys}). With a single pass over the affinity column, the ratio of d(TGCTG^{Cys}CAAG) to d(TGCTGCAAG^{Cys}) in the bound fraction decreased from 1:1 to 1:14. However, the desired d(TGCTG^{Cys}CAAG) remained in the flow-through material, from which only a fraction of the d(TGCTGCAAG^{Cys}) had been removed. Therefore, an amount of oligonucleotide mixture approximately twice the capacity of the affinity column was applied. The flow-through fractions—enriched in d(TGCTG^{Cys}CAAG)—were collected, desalted, pooled, and applied to the regenerated column for second and third cycles. After the first pass the flow-through material contained ~77% d(TGCTG^{Cys}CAAG), and this value increased to ~88% and ~92% after two additional cycles, respectively. Thus, three cycles yielded material with a d(TGCTG^{Cys}CAAG) to d(TGCTGCAAG^{Cys}) ratio of ~12:1, which was not improved by additional cycling. The capacity of the column, stored in buffer containing 1.0 M NaCl at 4 °C, did not decrease during several months and many separations; its properties were also highly reproducible (Figure 1).

Table 1: T_m Values of d(TGCTGCAAG) and an *S*-(2-N^7-Guanylethyl)GSH Derivative Hybridized with Complements Containing the Normal C and Mismatches Opposite the Central Guanine Residue (5′-dTGCTXCAAG-3′ and 3′-dACGAYGTTC-5′)

	T_m[a] (C)	
Y	X = G	X = *S*-(2-N^7-guanylethyl)GSH
C	39.4	28.3
G	12	[b]
T	20.1	8.8
A	10	[b]

[a] The estimate of the error in T_m is ±0.75 °C (from replicate experiments). [b] Under the conditions used, transitions occurred at <6 °C, the approximate limit at which T_m values could be determined with confidence.

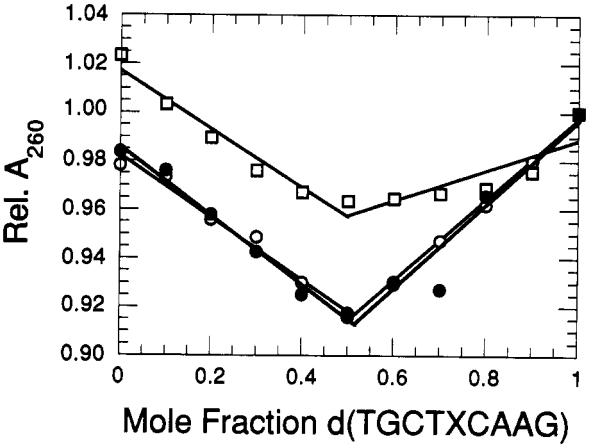


FIGURE 2: Mixing curves of d(TGCTXCAAG)/d(CTTGYAGCA). The total strand concentration was 13 μM, in 50 mM potassium phosphate buffer (pH 7.0) containing 50 μM EDTA, and the temperature was 15 °C. The lowest absorbance corresponds to the highest concentration of oligonucleotides in the duplex form. X:Y: ○, G:C; □, G:T; ●, G^{GSH}:C. The value at mole fraction 0.9 for the G^{GSH}:C sample was not measured due to an insufficient amount of the G^{GSH} oligomer.

Characterization of d(TGCTG^{GSH}CAAG) Duplexes. Initial studies with d(TGCTGCAAG) and d(TGCTG^{GSH}CAAG) duplexes were carried out with complementary 9-mers [d(CTTGNAGCA)] containing the normal C, as well as the mismatches T, G, and A across from the modified center guanine residue. The effects of introducing a mismatched base opposite G^{GSH} were evaluated by determining duplex thermal stabilities. The d(TGCTGCAAG) and d(TGCTG^{GSH}CAAG) duplexes displayed the same relative stabilities in the order G:C ≫ G:T ≫ G:G > G:A (Table 1). This order corresponds to that previously found for other G:N mismatch-containing duplexes in other sequence contexts (Aboul-ela *et al.*, 1985; Gaffney *et al.*, 1989). Introduction of the G^{GSH} moiety caused a depression of T_m by ~11 °C relative to a nonadducted G:N pair. This may be compared to the ~19 °C reduction in T_m of both G:T pairs compared to the respective G:C pairs.

The above results were confirmed by mixing experiments in which the molar ratios of two strands are varied while the overall concentration is kept constant (Figure 2). At the

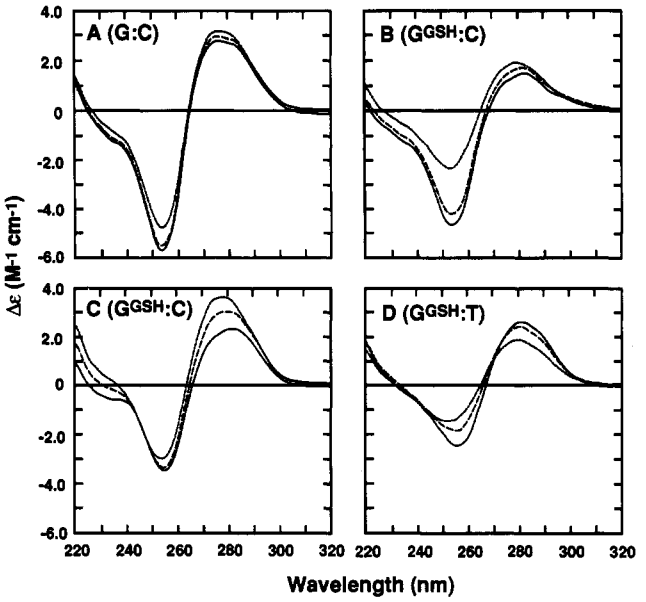


FIGURE 3: CD spectra of d(TGCTXCAAG)/d(CTTGYAGCA). Solvent conditions were as for Figure 2, and spectra were recorded at 13 μM strand concentration at 4 (- - -), 15 (– –), and 30 °C (—), respectively. $\Delta\epsilon$ values are based on the total monomeric residue concentration (117 μM). X:Y: A, G:C; B, G^{GSH}:C; C, G:T; D, G^{GSH}:T.

temperature used (15 °C), only the most stable pairs of oligomers (G:C, G^{GSH}:C, and G:T matches) showed minima indicative of duplex formation (Figure 2), whereas for the other combinations tried (G:A, G:G, and G^{GSH}:T) a straight line was observed through all points (not shown), confirming the results of melting experiments. All minima occurred at molar ratios of 1:1 for simple double-helix formation. Furthermore, in the case of the d(TGCTG^{GSH}CAAG)/d(CTTGCAGCA) mixture, the hypochromicities observed by mixing experiments were very similar to those of a d(TGCTGCAAG)/dCTTGCAGCA) mixture. This observation indicated that our assumptions regarding the extinction coefficient of G^{GSH} in oligonucleotides were valid, within the uncertainty of the experiment.

CD spectra were recorded for d(TGCTGCAAG) and d(TGCTG^{GSH}CAAG) paired with their respective complements containing C or T at the central position at three temperatures. All spectra except that of the unmodified d(TGCTGCAAG)/d(CTTGCAGCA) duplex (T_m = 39.4 °C; Table 1) were sensitive to temperature changes in the range 4–30 °C, in accord with their reduced thermal stabilities. The CD spectra of all four duplexes were consistent with an overall B-DNA structure (Figure 3) (Gennis & Cantor, 1972; Ivanov *et al.*, 1973; Gray *et al.*, 1992). In the case of the d(TGCTGCAAG):T mismatch duplex, the amplitude of both the negative and positive signals is reduced relative to the native G:C duplex, although the two maxima are still found at ~255 and ~275 nm, respectively (Figure 3A,C). This may be a result of the slight distortion of the duplex expected from the presumed wobble G:T base pair (Figure 3C). The introduction of the *S*-(2-N^7-guanylethyl)GSH moiety caused still larger duplex perturbations, expressed mainly as a reduction in the amplitudes of the negative and positive signals at ~255 and ~280 nm, respectively (Figures 3B,D). This effect has also been observed in the case of other oligonucleotide duplexes containing this adduct (Kim & Guengerich, 1993). Thus, the presence of the N^7-alkylguanine moiety caused the maximum of the positive signal in both cases to shift toward the red by ~5 nm. The distortion of the positive signal of the N^7-alkyl G:C duplex, *i.e.*, the shift of the positive maximum to ~280 nm and the appearance of a shoulder ~300 nm (see especially Figure 3B), is very similar to the effects observed by CD upon methylation of DNA (Ramstein *et al.*, 1971; Zavriev *et al.*, 1979). Sequence changes can produce dramatic changes in the shape of the CD spectra of short oligonucleotides (El antri *et al.*, 1993). Interpretation of the effects of the *S*-(2-N^7-guanylethyl)GSH moiety on the helix conformation should therefore mainly be made between the two isomeric duplex pairs. The observed attenuation and distortion of the CD signal may be the result of a deformation of the double helix. However, a caveat is the positive charge imposed by guanyl-N^7 adduction, which would change the electronic structure of the base. This in turn may affect the dipole transition moment of this base and thus cause the observed effect without an inherent deformation of the helix *per se*.

Concentration-Dependent Melting Studies. The results presented thus far did not indicate a preferential pairing between G^{GSH} and T, which might be inferred from mutagenesis results yielding >75% G:C to A:T transitions (Cmarik *et al.*, 1992). The thermodynamic parameters of duplex stability were examined in more detail. Using the van't Hoff method, plots of $1/T_m$ versus ln C_t yield values for $\Delta H°$ and $\Delta S°$ (Albergo *et al.*, 1981; Marky & Breslauer, 1987; Puglisi & Tinoco, 1989). The properties of oligonucleotides containing mutagenic lesions have previously been studied using these methods (Gaffney & Jones, 1989; Taylor *et al.*, 1990; Bishop & Moschel, 1991; Plum *et al.*, 1992). In our case, one potential problem was the sensitivity of the guanyl-N^7 adduct to thermal depurination. In order to minimize experimental artifacts due to this effect, melting experiments were conducted with low oligonucleotide and salt concentrations, both of which lower the temperature of the helix–coil transition. In addition, the temperature increase was relatively fast (~0.85 °C min^{-1}) to avoid excessive incubation at each temperature step, although in a range acceptable for short oligonucleotide duplexes (Puglisi & Tinoco, 1989). The amount of depurination during the course of a melting experiment was negligible (≤5%) under the conditions used, as demonstrated by subsequent ion-pair HPLC (not shown).

Typical melting curves and smoothed first derivative curves are shown in Figure 4. A van't Hoff plot of all six duplex species studied is shown in Figure 5, and the corresponding thermodynamic parameters are given in Table 2. This analysis clearly showed that all G:T mismatches had unfavorable $\Delta H°$ values compared to the respective G:C pairs, whereas the $\Delta S°$ values were slightly favorable. The $\Delta G°$ values at 25 and 37 °C were 2.1–3.1 kcal mol^{-1} lower than those for the corresponding G:C pairs. Introduction of an N^7-alkyl-G:T pair reduced the duplex stability by ~4 kcal mol^{-1} relative to the normal G:C pair, corresponding to a difference in K_{eq} of about 3 orders of magnitude. The three G:C/G:T pairs (G, G^{GSH}, and G^{Cys}) are approximately isoenergetic in this temperature range; this might be of relevance for considerations regarding mutational specificity. Furthermore, the similarity of the magnitude of the $\Delta H°$ and $\Delta S°$ values for the three T mismatch pairs may indicate that neither the adduct side chain nor the positive charge significantly alters the base pairing of the G:T mismatch pair, at least at pH 7.0.

The results for the three G:C pairs (G:C, G^{GSH}:C, and G^{Cys}:C) were quite different. The $\Delta H°$ for base-pair formation with both adducted G:C pairs is greatly enhanced (Table 2). That is, the G^{GSH}:C and G^{Cys}:C pairs are stabilized relative

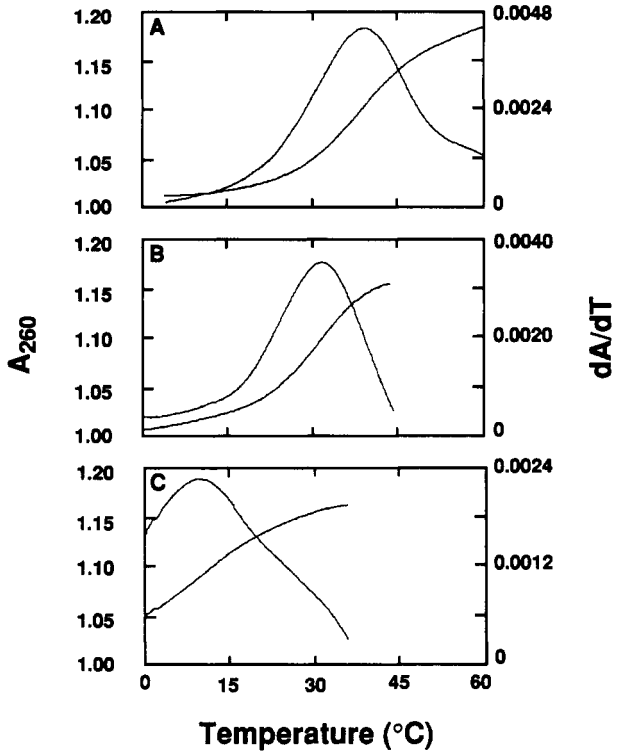


FIGURE 4: A_{260} versus temperature melting curves and first derivatives for d(TGCTXCAAG)/d(CTTGYAGCA). Solvent conditions were as for Figure 2. X:Y: A, G:C; B, G^{Cys}:C; C, G^{Cys}:T. Strand concentrations for the spectra shown were 13 μM.

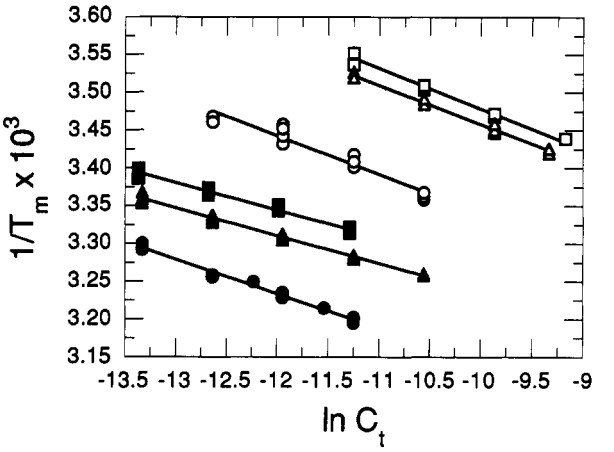


FIGURE 5: van't Hoff plots of $1/T_m$ versus ln C_t for duplex to coil transitions of d(TGCTXCAAG)/d(CTTGYAGCA). Solvent conditions were as for Figure 2. X:Y: ●, G:C; ▲, G^{Cys}:C; ■, G^{GSH}:C; ○, G:T; △, G^{Cys}:T; □, G^{GSH}:T.

to the G:C pair by ~10 kcal mol^{-1}. We initially speculated that this favorable enthalpy may be due in part to an interaction of the GSH moiety with the DNA. However, thermodynamic parameters for the oligonucleotides containing *N*-acetyl-*S*-(2-N^7-guanylethyl)Cys methyl ester and *S*-(2-N^7-guanylethyl)GSH were virtually identical (Table 2). The former adduct has considerably fewer hydrogen bond acceptor/donor groups than the latter and contains, in addition, bulky methyl groups. It thus appears more likely that the favorable transition enthalpy may be due to the positive charge imposed on the imidazole ring by N^7-guanine alkylation.

pH Titration of Oligonucleotide with N-Acetyl-S-(2-N^7-guanylethyl)Cys Methyl Ester. The ionization of the N1 imino proton of guanine may affect its base-pairing properties, and the pK_a of N^7-alkylguanine in the context of a single-stranded oligonucleotide was thus measured. Both the nonmodified and the alkylated oligomer showed distinctly different pH-dependent UV spectra (Figure 6) (the G^{Cys} oligomer was an approximately 1:1 mixture of the two positional isomers). When the absorbance values at 267 nm were plotted versus pH, a clear difference was observed (Figure 7). The spectrum of d(TGCTGCAAG) was independent of pH until ~9.5 (Figure 6A) and titrated in a single region with a pK_a of ~10.4 (Figure 7). However, changes in the spectra of the N^7-alkylguanine-containing oligonucleotide appeared at about pH 7.5 (Figure 6B). From the titration curve at 267 nm, it appears that this oligonucleotide titrates in two pH regions, with pK_a values of ~8.2 and ~10.4, respectively (Figure 7). In addition, the amplitude of the second region was significantly lower than that of the unmodified oligonucleotide. These results indicate that N^7-alkylguanine in a single-stranded oligonucleotide chain may be partially ionized, even at physiological pH.

DISCUSSION

The oligonucleotide used in the present study, d(TGCTGCAAG), represents a mutational site in the phage M13mp18 (Cmarik *et al.*, 1992) and was chosen in order to extend the studies of the 1,2-dibromoethane-derived DNA adduct *S*-(2-N^7-guanylethyl)GSH to a biologically relevant context target. It has been shown that DNA sequence context strongly affects both the degree of DNA damage and the ensuing mutation rates in the case of several carcinogen–DNA adducts (Kohn *et al.*, 1987; Warpehoski & Hurley, 1988; Bigger *et al.*, 1989; Singer *et al.*, 1989; Wurdeman *et al.*, 1989; Richardson & Richardson, 1990; Singer & Essigmann, 1991; Cheng *et al.*, 1992). Reaction of d(TGCTGCAAG) with *S*-(2-chloroethyl)GSH can result in three monoadduct positional isomers, three diadduct species, one triadduct species, and minor adduct species such as *S*-(2-N^1-adenylethyl)GSH (Kim *et al.*, 1990).

Table 2: Thermodynamic Data for Duplex Formation by 5′-(dTGCTXCAAG)-3′/3′-d(ACGAYGTTC)-5′

X:Y	T_m[a] (°C)	$\Delta H°$ (kcal mol^{-1})	$\Delta S°$ (cal mol^{-1} deg^{-1})	$\Delta G°$ (25 °C) (kcal mol^{-1})	$\Delta G°$ (37 °C) (kcal mol^{-1})	$\Delta\Delta G°_C$ (25 °C)[b] (kcal mol^{-1})	$\Delta\Delta G°_T$ (25 °C)[c] (kcal mol^{-1})
G:C	39.4	−43.9 ± 1.42	−116.3 ± 4.62	−9.2 ± 0.05	−7.8		−2.6
G^{GSH}:C	28.3	−55.1 ± 2.55	−158.7 ± 8.55	−7.8 ± 0.60	−5.8	1.4 ± 0.65	−2.8
G^{Cys}:C	31.5	−54.2 ± 1.52	−153.7 ± 5.05	−8.3 ± 0.02	−6.5	0.9 ± 0.07	−3.1
G:T	20.1	−39.0 ± 3.36	−108.6 ± 11.5	−6.6 ± 0.08	−5.3	2.6 ± 0.13	
1 M NaCl[d]	31.9	−43.6	−119.0	−7.6			
G^{GSH}:T	8.8	−37.3 ± 1.60	−108.2 ± 3.28	−5.0 ± 0.04	−3.7	4.2 ± 0.09	
G^{Cys}:T	10.8	−38.9 ± 1.66	−112.8 ± 5.54	−5.2 ± 0.06	−3.9	4.0 ± 0.11	
1 M NaCl[d]	18.2	−45.1	−132.4	−5.6			

[a] Measured at 13.0 μM total strand concentration, except for G^{GSH}:C where the concentration was 12.5 μM, in 50 mM potassium phosphate buffer (pH 7.0) containing 50 μM EDTA. The estimated precision is ±0.75 °C. Values shown are averages of three experiments. [b] Free energy difference relative to the duplex containing the nonmodified G:C pair. [c] Free energy difference of a G:C pair relative to its corresponding G:T pair. [d] Measurements were obtained in 50 mM potassium phosphate buffer (pH 7.0) containing 1 M NaCl and 50 μM EDTA.

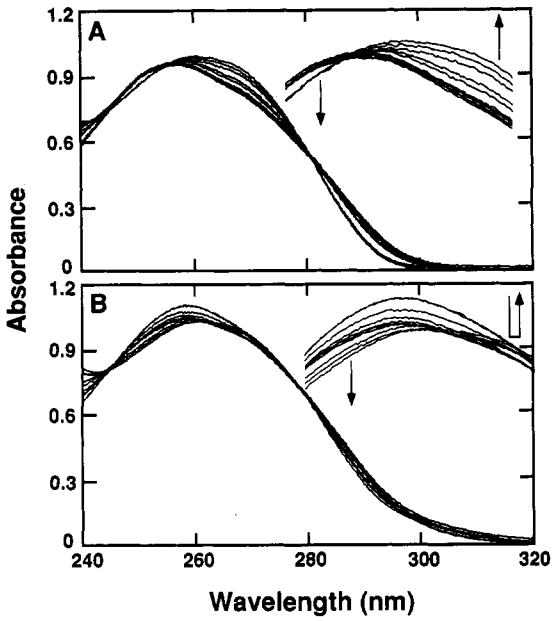


FIGURE 6: Absorption spectra of (A) d(TGCTGCAAG) (pH 6.5, 7.0, 7.5, 8.5, 9.0, 9.75, 10.0, 10.25, 10.75, 11.0, and 11.5) and (B) a mixture of d(TGCTG^{Cys}CAAG) and d(TGCTGCAAG^{Cys}) (pH 6.5, 7.5, 7.75, 8.0, 8.5, 9.0, 9.5, 10.0, 10.25, 10.5, and 11.0) versus pH (pH values of spectra are in parentheses; not all spectra are shown). Spectra were recorded in 0.10 M buffers containing 0.10 M KCl at intervals of 0.25–0.5 pH unit. Insets show expansions of the regions between 250 and 270 nm. Arrows indicate the changes in absorbance with increasing pH.

From this complex mixture, d(TGCTG^{GSH}CAAG) was isolated in three HPLC steps. Separation of oligonucleotide adduct positional isomers has been reported only in a few cases (Koehl *et al.*, 1989; Mao *et al.*, 1992) and remains a considerable challenge. Thus, in the case of the reaction products obtained with *N*-acetyl-*S*-(2-chloroethyl)Cys methyl ester, conventional chromatography was unsuccessful in separating the two positional isomers TGCTG^{Cys}CAAG and TGCTGCAAG^{Cys}.

It was, however, possible to achieve a relatively high degree of purification (~92%) of the desired positional isomer, TGCTG^{Cys}CAAG, by the use of oligonucleotide-based affinity chromatography. Clearly, this purity is insufficient for most biological applications (*e.g.*, site-specific mutagenesis), but appears acceptable for the analysis of spectral and thermodynamic properties; a 10% oligonucleotide impurity introduced an error in $\Delta H°$ of less than 3% for RNA duplex stability (Kierzek *et al.*, 1986). Oligomer-based affinity chromatography is based on the destabilizing effect of a specific base modification. The degree of purity attainable is thus related to the differential thermodynamic stability between the duplex containing the modified and the nonmodified bases, respectively, and may be affected by kinetic factors as well. Similar strategies have been utilized to differentially hybridize unmodified oligonucleotides (Nelson & Liu, 1989; Pei *et al.*, 1991); however, this appears to be the first time differential hybridization has been used for the purification of adduct-bearing oligonucleotides.

UV melting and CD experiments showed that the G^{GSH}:T pair is considerably less stable and appears to distort the helix more than the G^{GSH}:C pair. However, a relatively enhanced thermodynamic stability of G^{GSH}:T compared to G:T might be a reasonable explanation for G:C to A:T transition mutations. In order to investigate this possibility, the concentration-dependent thermal denaturation of six oligonucleotide duplexes was analyzed by the van't Hoff method,

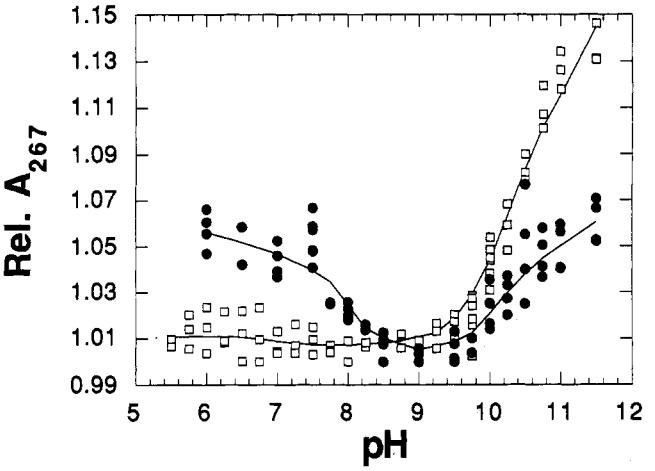


FIGURE 7: Plot of A_{267} versus pH values for d(TGCTGCAAG) (●, ■) and d(TGCTG^{Cys}CAAG)/d(TGCTGCAAG^{Cys}) (○, □). Circles and squares refer to values normalized to absorbances at the isosbestic points at 245.5 and 280.5 nm, respectively (Figure 6). In the case of d(TGCTG^{Cys}CAAG)/d(TGCTGCAAG^{Cys}), data from two independent measurements are shown.

yielding thermodynamic data for the helix–coil transitions. The duplex containing an unmodified G:C pair is the most stable at 25–37 °C; at a temperature close to 0 °C, the $\Delta G°$ values of the two N^7-alkyl-G:C-containing duplexes are identical to that of the native G:C duplex (within 0.3 kcal mol^{-1}). In the physiologically relevant range of 25–37 °C, the G^{GSH}:T and G^{Cys}:T mispair-containing duplexes were clearly the least stable, with $\Delta G°$ values ~4 kcal mol^{-1} higher (less favorable) than the that of G:C duplex. An interesting result was that for G^{GSH}:C and G^{Cys}:C the transition enthalpy $\Delta H°$ was considerably more favorable (~11 kcal mol^{-1}) than for the G:C pair. This favorable enthalpy was offset by an unfavorable entropy (~39 cal mol^{-1} deg^{-1}). Our results are in accord with a previous study, which reported the same trend for a self-complementary oligonucleotide containing two N^7-methylguanine:C pairs (Ezaz-Nikpay & Verdine, 1992), and also with a theoretical calculation on the effect of guanine methylation upon G:C pairing (Abdulnur & Flurry, 1976). These results suggest that the favorable enthalpy and unfavorable entropy largely result from the positive charge imposed on the imidazole ring. It has previously been reported that stacking interactions are enhanced by guanine N^7-alkylation or protonation (Jordan & Sostman, 1973; Ishida *et al.*, 1988). If the increased transition enthalpy is also due to improved stacking interactions in our study, nearest-neighbors may affect this phenomenon; in the methyl adduct study, the 5′ and 3′ neighbors were C and A, respectively (Ezaz-Nikpay & Verdine, 1992), whereas they were T and C, respectively, in this case.

As noted above, the modified G:C pairs and the modified G:T pairs have rather similar $\Delta\Delta G°$ values at both 25 and 37 °C (Table 2). If the thermodynamic parameters of an oligomeric duplex in solution correspond in general to the conditions in the polymerase active site, an N^7-alkylguanine would not be expected to give significantly higher levels of T misinsertion than a normal G. One model of polymerase fidelity suggests that the specificity of base-pair formation is enhanced in the polymerase active site by entropy suppression (Petruska *et al.*, 1986, 1988). A tight active site is considered to exclude H_2O, thereby reducing entropy differences between base pairs and increasing enthalpy differences to the extent that $\Delta\Delta G° \approx \Delta\Delta H°$ (Petruska *et al.*, 1986, 1988). If this model is applicable in the case of N^7-alkylguanine adducts,

insertion of C compared to T across from G^{GSH} would actually be even more favorable than insertion of C opposite G. If these conditions apply, then guanine N^7-alkyl derivatives should not be mutagenic. However, thermodynamic parameters of adducted base-pair stabilities in solution may be a weak measure of mutagenic potential. For instance, it has been shown for several adducts that the thermodynamic parameters and thermal stabilities of adduct-bearing oligonucleotide duplexes do not clearly correlate to the mutations observed. Studies on O^6-alkylguanine and O^4-alkylthymine have shown that the view that base-pair strength is a main factor in replication specificity is overly simplistic (Gaffney & Jones, 1989; Lakshman *et al.*, 1992; Plum *et al.*, 1992). It has been proposed instead that the physical similarity between an adduct and an unmodified base or an alignment of a base pair most closely resembling the geometry of that of a canonical Watson–Crick pair is the basis of mutagenicity by these and perhaps other adducts (Swann, 1990).

In the context of the working hypothesis that these N^7-guanyl adducts are mutagenic and cause the observed G:C to A:T transitions, it is of interest to compare two models for G:T and N^7-alkylguanine:T mispairs. The structure of the unmodified G:T mispair has been studied in some detail and shown to adopt a wobble structure, with marginal distortion of the overall helix structure (Hunter *et al.*, 1986; Kennard, 1987; Shibata *et al.*, 1991). A major structural deviation from Watson–Crick pairs is the shift of T into the major groove in order to bring the O2 atom of thymine into hydrogen-bonding distance of the N1 of guanine, leading to a loss of the characteristic symmetry (Hunter *et al.*, 1986; Kennard, 1987).

Another model for G:T mispairing involves ionized bases (Lawley & Brookes, 1961, 1962; Sowers *et al.*, 1987). In the case of N^7-alkylguanines, ionization at N1 occurs at physiological pH as a result of the reduction in the pK_a of N1 from ~9 in unmodified guanosine to ~7 in alkylated nucleosides (Brookes & Lawley, 1961; Lawley & Brookes, 1963). Such an ionized base pair retains Watson–Crick symmetry and is proposed to be stabilized by two hydrogen bonds between N1 of guanine and N3 of thymine and between N2 of guanine and O2 of thymine, respectively (Brookes & Lawley, 1961). A wobble G:T mispair involving ionized guanine would only be stabilized by one hydrogen bond and thus is unlikely.

The question then arises as to the extent of N1 deprotonation of N^7-alkylguanines in polynucleotides, both in the single- and double-stranded states. As mentioned, the pK_a of N^7-alkylguanine, as well as guanylic acid derivatives, is found to be lowered by ~2 units relative to the nonalkylated base (Brookes & Lawley, 1961; Lawley & Brookes, 1963; Michelson & Pochon, 1965; Hendler *et al.*, 1970). The pK_a values in both single- and double-stranded polynucleotides have, however, been reported to be ≥9.5, suggesting that the protonated form is stabilized in polynucleotides (Michelson & Pochon, 1966; Hendler *et al.*, 1970). These reports are in partial conflict with a later study, which reported that N^7-alkylated guanine changes from the keto to the zwitterionic form upon melting (Mansy & Peticolas, 1976). The involvement of ionization at guanine N1 has been used to rationalize the mutagenesis of the bulky N^7-alkylguanine adduct ICR191 (Sahasrabudhe *et al.*, 1990). These authors suggested that as the strands are separated at the temporarily stalled replication fork, the N^7-alkylguanine adduct may lose the N1 proton to facilitate N^7-alkylguanine:T mispairing (Sahasrabudhe *et al.*, 1990). These and other results may suggest that the mutagenic properties of some N^7-alkylguanyl adducts result from a combination of (transient) blockage of DNA polymerases (Refolo *et al.*, 1985; Sambamurti *et al.*, 1988; Sahasrabudhe *et al.*, 1990; Gray *et al.*, 1991) and guanine N1 deprotonation. In this scenario it is conceivable that G:T misincorporation is facilitated by an overall Watson–Crick conformation. As pointed out previously, if an ionized base is highly efficient in directing misincorporations, the degree of ionization could be quite low while still giving mutation rates significantly higher than background (Goodman *et al.*, 1993). In a closely related case, it was found that 5-bromouracil and 5-fluorouracil (5-halo-U) form mispairs with G, the conformations of which are pH-dependent (Sowers *et al.*, 1988, 1989). At high pH, 5-halo-U was ionized and formed a G mispair with Watson–Crick symmetry, whereas a wobble structure was observed at lower pH; the pK_a for the conformational transition was ~8.5. Two laboratories have also shown that the extent of misincorporation opposite 5-halo-U is enhanced significantly as the pH is raised, supporting a model in which ionization of 5-halo-U is primarily responsible for mutagenesis by these bases (Driggers & Beattie, 1988; Yu *et al.*, 1993), as originally proposed by Lawley and Brookes (1962). Indeed, it has been proposed that ionized base pairs may contribute significantly to mispairing and mutagenesis (Sowers *et al.*, 1987; Goodman *et al.*, 1993), in preference over the model involving disfavored tautomers (Topal & Fresco, 1976).

While this study does not provide a definitive basis for putative mutagenesis by G^{GSH}, it has established some of the physical properties imparted on the guanine residue by this and other N^7 adducts. The work clearly distinguishes between enthalpic and entropic contributions to the stabilities of some of the base pairs of interest. The extent of ionization of N^7-alkylguanines in DNA and, ultimately, in the DNA polymerase active site remains to be determined. In addition, the concomitant effects of pH on the thermodynamic parameters of N^7-alkyl-G:T pairs and DNA polymerase misinsertion frequencies will need to be investigated. Such information is probably needed to further conclusions regarding mechanisms of mutagenesis by this type of compound. Studies of replication fidelities *in vitro* with the N^7-alkylguanyl adducts placed at specific sites in oligonucleotides such as the one utilized here should provide valuable information in this respect. At this time, we cannot exclude the hypothesis that another, minor (guanine) adduct is the major mutagenic lesion, and this possibility must be considered as well.

ACKNOWLEDGMENT

We thank M. V. Martin for her help in the statistical analysis of the thermodynamic data.

REFERENCES

Abdulnur, S. F., & Flurry, R. L., Jr. (1976) *Nature 264*, 369–370.

Aboul-ela, F., Koh, D., Tinoco, I., Jr., & Martin, F. H. (1985) *Nucleic Acids Res. 13*, 4811–4824.

Albergo, D. D., Marky, L. A., Breslauer, K. J., & Turner, D. H. (1981) *Biochemistry 20*, 1409–1413.

Bigger, C. A. H., Strandberg, J., Yagi, H., Jerina, D. M., & Dipple, A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2291–2295.

Bishop, R. E., & Moschel, R. C. (1991) *Chem. Res. Toxicol. 4*, 647–654.

Blaho, J. A., & Wells, R. D. (1989) *Prog. Nucleic Acid Res. Mol. Biol. 37*, 107–126.

Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology*, 3rd ed. (Fasman, G. D., Ed.) p 589, CRC Press, Cleveland, OH.

Brookes, P., & Lawley, P. D. (1961) *J. Chem. Soc.*, 3923–3928.
Brusick, D. (1989) in *Principles and Methods of Toxicology*, 2nd ed. (Hayes, A. W., Ed.) pp 407–434, Raven Press, New York.
Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., & Loeb, L. A. (1992) *J. Biol. Chem. 267*, 166–172.
Cmarik, J. L., Inskeep, P. B., Meyer, D. J., Meredith, M. J., Ketterer, B., & Guengerich, F. P. (1990) *Cancer Res. 50*, 2747–2752.
Cmarik, J. L., Humphreys, W. G., Bruner, K. L., Lloyd, R. S., Tibbetts, C., & Guengerich, F. P. (1992) *J. Biol. Chem. 267*, 6672–6679.
Cosman, M., de los Santos, C., Fiala, R., Hingerty, B. E., Singh, S. B., Ibanez, V., Margulis, L. A., Live, D., Geacintov, N. E., Broyde, S., & Patel, D. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1914–1918.
Crothers, D. M., Haran, T. E., & Nadeau, J. G. (1990) *J. Biol. Chem. 265*, 7093–7096.
Dosanjh, M. K., Galeros, G., Goodman, M. F., & Singer, B. (1991) *Biochemistry 30*, 15595–11599.
Driggers, P. H., & Beattie, K. L. (1988) *Biochemistry 27*, 1729–1735.
El antri, S., Bittoun, P., Mauffret, O., Monnot, M., Convert, O., Lescot, E., & Fermandjian, S. (1993) *Biochemistry 32*, 7079–7088.
Ezaz-Nikpay, K., & Verdine, G. L. (1992) *J. Am. Chem. Soc. 114*, 6562–6564.
Foster, P. L., Wilkinson, W. G., Miller, J. K., Sullivan, A. D., & Barnes, W. M. (1988) *Mutat. Res. 194*, 171–181.
Gaffney, B. L., & Jones, R. A. (1989) *Biochemistry 28*, 5881–5889.
Gennis, R. B., & Cantor, C. R. (1972) *J. Mol. Biol. 65*, 381–399.
Ghosh, S. S., & Musso, G. F. (1987) *Nucleic Acids Res. 15*, 5353–5372.
Goodman, M. F., Creighton, S. R., Bloom, L. B., & Petruska, J. (1993) *Crit. Rev. Biochem. Mol. Biol. 28*, 83–126.
Gopalakrishnan, S., Stone, M. P., & Harris, T. M. (1989) *J. Am. Chem. Soc. 111*, 7232–7239.
Gralla, J., & Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301–319.
Gray, D. M., Ratliff, R. L., & Vaughan, M. R. (1992) *Methods Enzymol. 211*, 389–406.
Heidelberger, C. (1975) *Annu. Rev. Biochem. 44*, 79–121.
Hendler, S., Fürer, E., & Srinivasan, P. R. (1970) *Biochemistry 9*, 4141–4153.
Hruszkewycz, A. M., & Dipple, A. (1991) *Carcinogenesis 12*, 2185–2187.
Humphreys, W. G., Kim, D.-H., Cmarik, J. L., Shimada, T., & Guengerich, F. P. (1990) *Biochemistry 29*, 10342–10350.
Hunter, W. N., Kneale, G., Brown, T., Rabinovich, D., & Kennard, O. (1986) *J. Mol. Biol. 190*, 605–618.
Inskeep, P. B., & Guengerich, F. P. (1984) *Carcinogenesis 5*, 805–808.
Inskeep, P. B., Koga, N., Cmarik, J. L., & Guengerich, F. P. (1986) *Cancer Res. 46*, 2839–2844.
Ishida, T., Doi, M., Ueda, H., Inoue, M., & Scheldrick, G. M. (1988) *J. Am. Chem. Soc. 110*, 2286–2294.
Ivanov, V. I., Minchenkova, L. E., Schyolkina, A. K., & Poletayev, A. I. (1973) *Biopolymers 12*, 89–110.
Jordan, F., & Sostman, H. D. (1973) *J. Am. Chem. Soc. 95*, 6544–6554.
Kalnik, M. W., Li, B. F. L., Swann, P. F., & Patel, D. J. (1989a) *Biochemistry 28*, 6170–6181.
Kalnik, M. W., Li, B. F. L., Swann, P. F., & Patel, D. J. (1989b) *Biochemistry 28*, 6182–6192.
Kennard, O. (1987) in *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 1, pp 25–52, Springer-Verlag, Berlin.
Kierzek, R., Caruthers, M. H., Longfellow, C. E., Swinton, D., Turner, D. H., & Freier, S. M. (1986) *Biochemistry 25*, 7840–7846.
Kim, D.-H., & Guengerich, F. P. (1990) *Carcinogenesis 11*, 419–424.
Kim, M.-S., & Guengerich, F. P. (1993) *Chem. Res. Toxicol. 6*, 900–905.
Kim, D.-H., Humphreys, W. G., & Guengerich, F. P. (1990) *Chem. Res. Toxicol. 3*, 587–594.
Koehl, P., Burnouf, D., & Fuchs, R. P. P. (1989) *J. Mol. Biol. 207*, 355–364.
Koga, N., Inskeep, P. B., Harris, T. M., & Guengerich, F. P. (1986) *Biochemistry 25*, 2192–2198.
Kohn, K. W., Hartley, J. A., & Mattes, W. B. (1987) *Nucleic Acids Res. 15*, 10531–10548.
Lakshman, M. K., Sayer, J. M, & Jerina, D. M. (1992) *J. Org. Chem. 57*, 3438–3443.
Lawley, P. D. (1984) in *Chemical Carcinogens*, 2nd ed. (Searle, C. E., Ed.) pp 324–484, American Chemical Society, Washington, D.C.
Lawley, P. D., & Brookes, P. (1961) *Nature 192*, 1081–8210.
Lawley, P. D., & Brookes, P. (1962) *J. Mol. Biol. 4*, 216–219.
Lawley, P. D., & Brookes, P. (1963) *Biochem. J. 89*, 127–138.
Mansy, S., & Peticolas, W. L. (1976) *Biochemistry 15*, 2650–2655.
Mao, B., Margulis, L. A., Li, B., Ibanez, V., Lee, H., Harvey, R. G., & Geacintov, N. E. (1992) *Chem. Res. Toxicol. 5*, 773–778.
Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601–1620.
Mattes, W. B., Hartley, J. A., & Kohn, K. W. (1986) *Biochim. Biophys. Acta 868*, 71–76.
Meyer, S. L. (1975) *Data Analysis for Scientists and Engineers*, John Wiley and Sons, Inc., New York.
Michelson, A. M., & Pochon, F. (1966) *Biochim. Biophys. Acta 114*, 469–480.
Miller, J. A. (1970) *Cancer Res. 30*, 559–576.
Möller, A., Nordheim, A., Nichols, S. R., & Rich, A. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4777–4781.
Nelson, P. S., & Liu, E. (1989) *Nucleic Acids Res. 17*, 7187–7194.
Oida, T., Humphreys, W. G., & Guengerich, F. P. (1991) *Biochemistry 30*, 10513–10522.
Ozawa, N., & Guengerich, F. P. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5266–5270.
Pei, D., Ulrich, H. D., & Schultz, P. G. (1991) *Science 253*, 1408–1411.
Peterson, L. A., Harris, T. M., & Guengerich, F. P. (1988) *J. Am. Chem. Soc. 110*, 3284–3291.
Petruska, J., Sowers, L. C., & Goodman, M. F. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1559–1562.
Petruska, J., Goodman, M. F., Boosalis, M. S., Sowers, L. C., Cheong, C., & Tinoco, I., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6252–6256.
Plum, G. E., Grollman, P., Johnson, F., & Breslauer, K. J. (1992) *Biochemistry 31*, 12096–12102.
Puglisi, J. D., & Tinoco, I., Jr. (1989) *Methods Enzymol. 180*, 304–325.
Ramstein, J., Hélène, C., & Leng, M. (1971) *Eur. J. Biochem. 21*, 125–136.
Refolo, L. M., Conley, M. P., Sambamurti, K., Jacobsen, J. S., & Humayun, M. Z. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3096–3100.
Richardson, F. C., & Richardson, K. K. (1990) *Mutat. Res. 233*, 127–138.
Roberts, R. W., & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9397–9401.
Sahasrabudhe, S. R., Luo, X., & Humayun, M. Z. (1990) *Biochemistry 29*, 10899–10905.
Sambamurti, K., Callahan, J., Luo, X., Perkins, C. P., Jacobsen, J. S., & Humayun, M. Z. (1988) *Genetics 120*, 863–873.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991a) *J. Am. Chem. Soc. 113*, 4313–4322.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1991b) *Biochemistry 30*, 8242–8251.

Seeberg, E., & Fuchs, R. P. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 191–194.
Shibata, M., Zielinski, T. J., & Rein, R. (1991) *Biopolymers 31*, 211–232.
Singer, B., & Essigmann, J. M. (1991) *Carcinogenesis 12*, 949–955.
Singer, B., Chavez, F., Goodman, M. F., Essigmann, J. M., & Dosanjh, M. K. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8271–8274.
Sowers, L. C., Saw, B. R., Veigl, M. L., & Sedwick, W. D. (1987) *Mutat. Res. 177*, 201–218.
Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., & Fazakerley, G. V. (1988) *J. Biol. Chem. 263*, 14794–14801.
Sowers, L. C., Goodman, M. F., Eritja, R., Kaplan, B., & Fazakerley, G. V. (1989) *J. Mol. Biol. 205*, 437–447.
Swann, P. F. (1990) *Mutat. Res. 233*, 81–94.
Taylor, J. S., Garrett, D. S., Brockie, I. R., Svoboda, D. L., & Telser, J. (1990) *Biochemistry 29*, 8858–8866.
Topal, M. D., & Fresco, J. R. (1976) *Nature 263*, 285–289.
van Bladeren, P. J., Breimer, D. D., Rotteveel-Smijs, G. M. T., de Knijff, P., Mohn, G. R., van Meeteren-Wälchli, B., & van der Gen, A. (1981) *Carcinogenesis 2*, 499–505.
Warpehoski, M. A., & Hurley, L. H. (1988) *Chem. Res. Toxicol. 1*, 315–333.
Wurdeman, R. L., Church, K. M., & Gold, B. (1989) *J. Am. Chem. Soc. 111*, 6408–6412.
Yu, H., Eritja, R., Bloom, L. B., & Goodman, M. F. (1993) *J. Biol. Chem. 268*, 15935–15943.
Zavriev, S. K., Minchenkova, L. E., Vorlíckovâ, M., Kolchinsky, A. M., Volkenstein, M. V., & Ivanov, V. I. (1979) *Biochim. Biophys. Acta 564*, 212–224.