Isolation and Structure of an Intrastrand Cross-Link Adduct of Mitomycin C and DNA†

Roland Bizanek,‡ Brian F. McGuinness,§,‖ Koji Nakanishi,§ and Maria Tomasz*,‡

Departments of Chemistry, Hunter College, City University of New York, New York, New York 10021, and Columbia University, New York, New York 10027

Received November 12, 1991; Revised Manuscript Received January 23, 1992

ABSTRACT: A new covalent mitomycin C–DNA adduct (**4**) was isolated from DNA exposed to reductively activated mitomycin C (MC) in vitro. The MC-treated DNA was hydrolyzed enzymatically under certain conditions, and the new adduct was isolated from the hydrolysate by HPLC. Its structure was determined by ultraviolet and circular dichroism spectroscopy and chemical and enzymatic transformations conducted on microscale. In the structure, a single 2″β,7″-diaminomitosene residue is linked bifunctionally to two guanines in the dinucleoside phosphate d(GpG). The guanines are linked at their N^2 atoms to the C1″ and C10″ positions of the mitosene, respectively. A key to the structure was a finding that removal of the mitosene from the adduct by hot piperidine yielded d(GpG); another was that the adduct was slowly converted to the known interstrand cross-link adduct **3** by snake venom diesterase and alkaline phosphatase. Adduct **4** represents an intrastrand cross-link in DNA formed by MC. Of the two possible strand-polarity isomers of **4**, **4a** in which the mitosene 1″-position is linked to the 3′-guanine of d(GpG) is designated as the proper structure, on the basis of the mechanism of the cross-linking reaction. The same adduct **4** was isolated from poly(dG)·poly(dC), synthetic oligonucleotides containing the GpG sequence, and *Micrococcus luteus* and calf thymus DNAs. The relative yields of interstrand and intrastrand cross-links (**3** and **4**) were determined under first-order kinetic conditions; an average 3.6-fold preference for the formation of **3** over that of **4** was observed. An explanation for this preference is proposed. Energy-minimized structural models were generated for the two strand-polarity isomers of the intrastrand cross-link (**4a** and **4b**), incorporated in a duplex decanucleotide. Both models indicate bending of DNA near the cross-link site. The identification and structural properties of the new, intrastrand cross-link lesion of DNA by MC are significant with respect to cancer drug development as well as structural chemistry of DNA-damaging agents.

The antitumor antibiotic mitomycin C (**1**; MC;[1] Chart I) is widely used in anticancer chemotherapy. Its mode of action is fundamentally related to its ability to bind covalently to DNA and cross-link two complementary DNA strands (Szybalski & Iyer, 1964). These effects are dependent on reductive activation of MC, known to occur in living cells as mediated by flavoenzymes (Keyes et al., 1984). Mechanistically, the initial reduction of the quinone function of MC induces a series of spontaneous transformations in the molecule, giving rise ultimately to an unstable form which functions as a monofunctional and bifunctional alkylating agent (Iyer & Szybalski, 1964; Moore, 1977; Lin et al., 1976; Tomasz & Lipman, 1981; Kohn & Zein, 1983; Pan et al., 1984; Danishefsky & Cuifolini, 1984; Egbertson & Danishefsky, 1987; Peterson & Fisher, 1986; Hoey et al., 1988; Tomasz et al., 1988a).

The major alkylation products of DNA by reduced MC have been isolated, and their structures have been determined as **2a**, **2b**, and **3** (Tomasz et al., 1986, 1987, 1988b), revealing that the alkylation is exquisitely specific: both the C1-aziridine and C10-carbamate alkylating functions of the drug react exclusively with 2-amino groups of guanines. Bisadduct **3** corresponds to the interstrand cross-link, as shown rigorously by its isolation from well-characterized cross-linked oligonucleotides (Borowy-Borowski et al., 1990a). NMR analysis of a cross-linked hexanucleotide duplex provided confirmatory evidence (Norman et al., 1990).

The evidence did not exclude the possibility, however, that some bisadduct **3** originated from *intrastrand* cross-links of DNA, formed between two adjacent guanines (Chart II). The enzymatic hydrolysis conditions used to generate the MC adducts from alkylated DNA could potentially lead to the same product **3** from both types of cross-links, and thus the existence of the intrastrand cross-link could have been overlooked.

Intrastrand cross-links by a bifunctional agent have been observed so far only in the case of platinum drugs, to our knowledge. Both *cis*-DDP and *trans*-DDP form such lesions as their major reactions with DNA, and the relationship of these adducts to platinum drug-mediated antitumor activity, cytotoxicity, mutagenicity, and intracellular processing of DNA damage is under intensive investigation (Bellon et al., 1991, and references cited therein). The bifunctional interaction of the natural product, mitomycin C, with DNA is potentially another case where intrastrand DNA cross-links may be formed (Tomasz et al., 1987). We report here that, indeed, reductively activated MC forms intrastrand cross-links in DNA between adjacent guanines in significant proportions to the previously detected interstrand cross-links. These cross-links could be readily identified by the structural

† This research was supported by National Institutes of Health Grants CA28681 (to M.T.) and GM34509 (to K.N.), Research Centers in Minority Institutions Award RR003037 from the NIH Division of Research Resources (to Hunter College), and a Postdoctoral Fellowship from Deutsche Forschungsgemeinschaft (to R.B.).

‡ Hunter College.

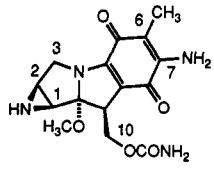
§ Columbia University.

‖ Present address: Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139.

[1] Abbreviations: MC, mitomycin C; UV, ultraviolet; CD, circular dichroism; SVD, snake venom diesterase; APase, alkaline phosphatase; P_1, nuclease P_1; M, mitosene (see definition of mitosene in footnote 2).

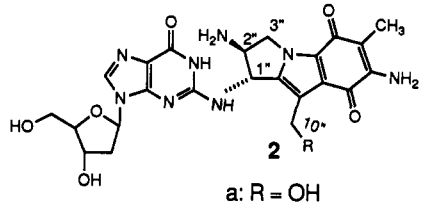
0006-2960/92/0431-3084$03.00/0

Chart I



1

MC

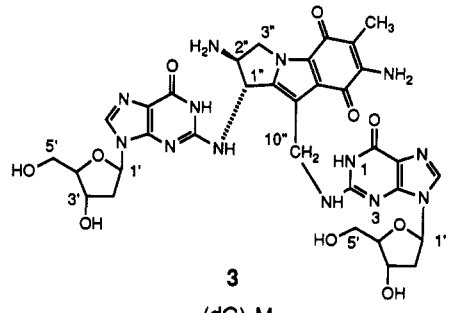


2

a: R = OH

b: R = $OCONH_2$

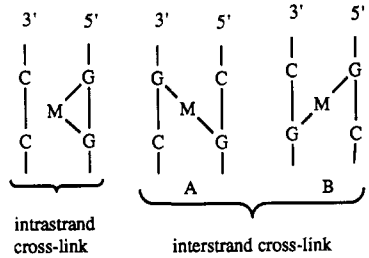
c: R = HSO_3^-



3

$(dG)_2M$

Chart II



characterization of a unique new adduct obtained from the DNA under certain hydrolytic conditions. We also aimed at determining the relative rates of formation of the interstrand and intrastrand cross-links. Furthermore, energy-minimized structural models were generated for both potential isomers of the intrastrand cross-link, incorporated in a duplex decanucleotide. The identification and structural properties of the new DNA lesion induced by MC should be of interest to investigators in the fields of cancer drug development as well as structural chemistry of DNA-damaging agents.

Materials and Methods

Poly(dG-dC)·poly(dG-dC), poly(dG)·poly(dC), and calf thymus DNA (type I) were purchased from Sigma (St. Louis, MO). *Micrococcus luteus* DNA was obtained from ICN Biomedicals, Costa Mesa, CA. Both DNAs were sonicated before use. Sources for all other materials used were given in a previous publication (Borowy-Borowski et al., 1990a).

HPLC. A Beckman Model 332 HPLC system, equipped with a Model 427 integrator and Model 165A absorbance detector, was used throughout. Conditions (semipreparative scale): Beckman ODS Ultrasphere column, 10 × 250 mm; eluant, 6–12% acetonitrile gradient in 0.02 M NH_4OAc, pH 5.5, in 30 min; flow rate, 3 mL/min. Analytical scale: same column, 4.5 × 250 mm; eluant, 6–18% acetonitrile gradient in 0.03 M potassium phosphate, pH 5.0, in 30 min; flow rate, 1 mL/min. Deviations from these conditions are noted when appropriate.

Identification of known adducts and nucleosides in the HPLC patterns was based upon direct comparison of elution times with those of authentic standards, chromatographed just before or after running the sample; mixed runs were carried out in cases of ambiguity.

Quantitative analysis of nucleotide-derived substances and free drug was carried out either by solution spectrophometry or by peak area analysis of HPLC patterns, as described (Borowy-Borowski et al., 1990a).

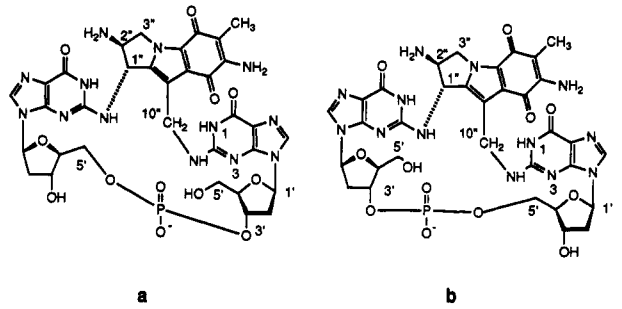
Treatment of Synthetic Polynucleotides, Oligonucleotides, and DNA with MC under Reductive Bifunctional Activating Conditions. A solution of 0.7 mM polynucleotide, etc., measured in mononucleotide equivalents, and 0.35 mM MC in 0.015 M Tris, pH 7.4, buffer was deaerated by purging with argon. $Na_2S_2O_4$ in deaerated water (22.5 mM) was added to the mixture to a final concentration of 0.53 mM, in five equal portions from a syringe under continuous purging with argon. After this treatment, the MC–polynucleotide complex was isolated as the earliest ultraviolet-absorbing fraction upon chromatography of the mixture over a Sephadex G-25 (fine) column, using 0.02 M NH_4HCO_3 as eluant, followed by lyophilyzation.

Enzymatic Digestion of the MC–Polynucleotide, –Oligonucleotide, or –DNA Complexes and Isolation of Adducts **3** *and* **4.** The complex, prepared by the method described above, was digested to the nucleoside level by the following protocol: Nuclease P_1 (0.5 unit/A_{260} unit of complex) was added to the complex (3 A_{260} units/mL) in dilute aqueous acetic acid, pH 5.0, followed by incubation for 2 h at 55 °C. The pH was adjusted to 8.2, and $MgCl_2$ was added to a 0.9 mM concentration. Addition of SVD (4.5 units/A_{260} units of complex) and a 1-h incubation at 37 °C was followed by addition of alkaline phosphatase (1.6 units/A_{260} units of complex) and 2 h of incubation at 37 °C. Adducts **3** and **4** were isolated by semipreparative HPLC of the digest. Elution times were 38 and 25 min, respectively.

Treatment of Poly(dG)·poly(dC) or Poly(dG-dC)·poly(dG-dC) with MC under Reductive Monofunctional Activating Conditions and Isolation of Monoadduct **2b.** A mixture of MC (0.33 mM), polynucleotide (0.67 mM in mononucleotide units), and PtO_2 (40 μg/mL) in 0.015 M Tris, pH 7.4, was deaerated by purging with helium. H_2 gas was bubbled for 5 min, followed by argon for 5 min and by air for 2 min. Isolation of the MC–polynucleotide complex was conducted by Sephadex G-25 chromatography as above. Enzymatic digestion of the complex and detection of **2b** by HPLC followed the protocols above. Elution time of **2b** was 15 min on the analytical scale.

Conversion of Monoadducted Polynucleotide–MC Complex to Cross-Linked Complex by $Na_2S_2O_4$ Treatment. The monoadducted complex (1 mM in mononucleotide units) in 0.015 M Tris, pH 7.4, was treated with 1.5 mM $Na_2S_2O_4$ under anaerobic conditions followed by isolation of the cross-linked complex as described previously (Tomasz et al., 1988a).

Chart III



a b

4

[d(GpG)M]

Molecular Modeling. All-atom minimizations were performed on a VAX station 3100 using the program MacroModel V3.0 (Mohamadi et al., 1990). The AMBER force field and charges (Weiner et al., 1984) were used in the absence of solvent and counterions, with dielectric R_{ij} and with cutoffs of 8 Å (van der Waals) and 13 Å (electrostatic) in the united atom mode. Minimizations were conducted by using a block-diagonal Newton Raphson routine followed by a Polak-Ribiere conjugate gradient to a root-mean-square gradient of 0.01 kJ mol^{-1} $Å^{-2}$ for all structures. In the absence of a crystallographic structure for the mitosene[2] unit, we energy-minimized 2″β,7″-diamino-1″α-hydroxymitosene by MacroModel. After the 1″-OH and 10″-$OCONH_2$ bonds were severed, this structure was placed visually in the central minor groove region of the computer-generated B-DNA decamer d(GCATGGATGC)·d(GCATCCATGC) using the Evans and Sutherland 390 graphic system. Bonds were introduced that linked the mitosene C-1″α and C-10″ positions to the guanine N^2 atoms of the central GpG sequence of the decamer to form a crude version for both isomeric decamer–mitosene adducts. These were then refined with the procedure described above. For comparison, the parent decamer duplex was also generated by the MacroModel (GROW mode) and was energy-minimized by the described procedure.

RESULTS

Detection of the New Adduct **4** *in Various Substrates.* Treatment of poly(dG)·poly(dC) with bifunctionally activated MC yielded a new adduct (Figure 1a) after enzymatic digestion, later shown to have structure **4** (Chart III). In contrast, under the same conditions poly(dG-dC)·(dG-dC) yielded **3**, the known cross-link adduct (Tomasz et al., 1987) as the major product (Figure 1b). The new adduct **4** was also detected in *M. luteus* and calf thymus DNAs (Figure 1c,d) and in oligonucleotides 2 and 3 (Table I; Figure 1e,g) after their similar treatment with MC. The new adduct **4** was absent or barely detectable, while adduct **3** was increased when the enzymatic digestion of the same complexes was conducted under our formerly used conditions (Tomasz et al., 1986, 1987) which employed DNAse I instead of nuclease P_1, much larger amounts of SVD, and prolonged incubation (data not shown).

Isolation of the New Adduct **4** *in Pure Form* was accomplished using *M. luteus* DNA as substrate and scaling up the reaction and enzymatic digestion. The product was collected from semipreparative HPLC, in multiple batches, and then desalted on a Sephadex G-25 column (0.01 M NH_4HCO_3 as eluant) from which it was eluted *after* the salt-containing fraction. Typically, yields of 3 A_{260} units/10 mg of DNA were obtained. The ultraviolet spectrum was essentially identical

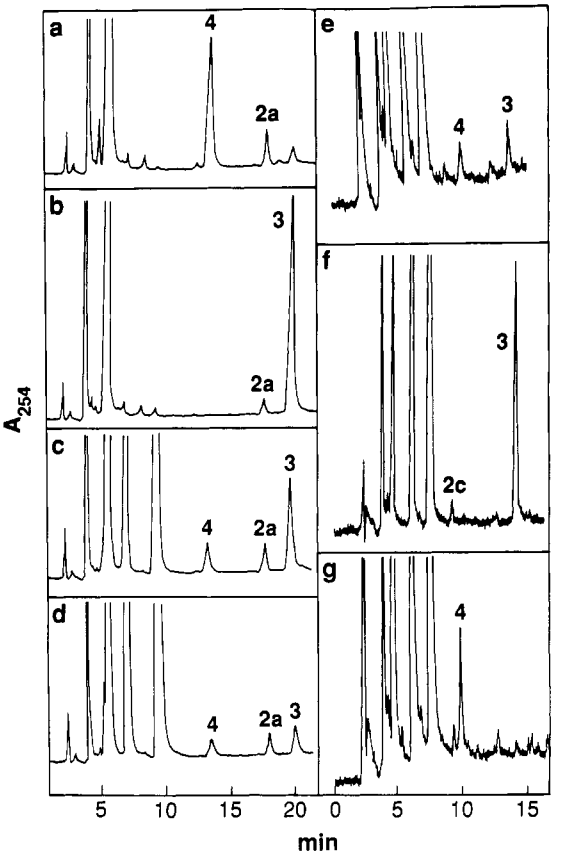


FIGURE 1: HPLC patterns from enzymatic digests of various bifunctionally activated MC–DNA and MC–oligonucleotide complexes. Analytical columns were used throughout. (a) MC–poly(dG)·poly(dC) complex; (b) MC–poly(dG-dC)·poly(dG-dC) complex; (c) MC–*M. luteus* DNA complex; (d) MC–calf thymus DNA complex; (e) MC–oligonucleotide 2 complex; (f) MC–oligonucleotide 1 complex; (g) MC–oligonucleotide 3 complex. The longer elution times observed for same substances in panels e–g are probably due to the use of a new column.

Table I: Relative Yields of Interstrand and Intrastrand Cross-Link Adducts **3** and **4** Observed in Various Substrates[a]

substrate	relative yield; **3**:**4**[a]	CpG:GpG frequency ratio	relative yields; normalized[a]
5′-ATATACGTATAT 3′-TATATGCATATA (oligonucleotide 1)[d]	4.8:1	1:1	4.8:1
5′-ATAATGGATATA 3′-TATTACCTATAT (oligonucleotide 2)[e]			
5′-TATACCG¹G²TATA 3′-ATATG²G¹CCATAT (oligonucleotide 3)	2.8:1	1:1	2.8:1
poly(dG-dC)·poly(dG-dC)[d] poly(dG)·poly(dC)[e]	1.7:1	1:1	1.7:1
M. luteus DNA	3.4:1	1.17:1[b]	2.6:1[c]
calf thymus DNA	1.7:1	0.27:1[b]	6.3:1[c]

[a] Relative yields were calculated from the ratios of areas of HPLC peaks of the adducts **3** and **4** (Figure 1). When two *separate substrates* were compared [the pair of oligonucleotides 1 and 2, or the pair poly(dG)·poly(dC) and poly(dG-dC)·poly(dG-dC)], each adduct area was first normalized to the area of dC in the same HPLC run. [b] From data in Fasman (1975). [c] By dividing the relative yield by the CpG/GpC frequency ratio. [d] To give **3**. [e] To give **4**.

[2] The term mitosene signifies indoloquinone derivatives of MC. Mitosene itself is the indoloquinone as in structure **2b**, but without substituents in the 1″-, 2″-, and 7″-positions (Webb et al., 1962). In a less rigorous sense, we use this term as the DNA-bound mitomycin unit in general.

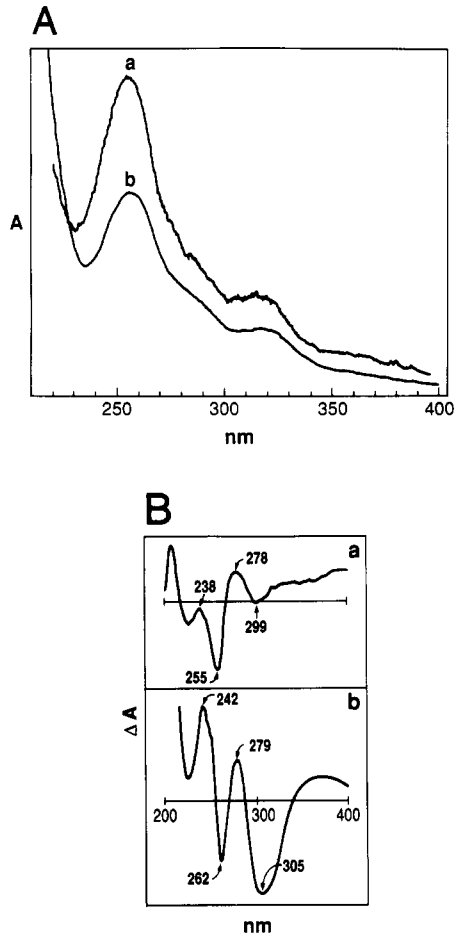


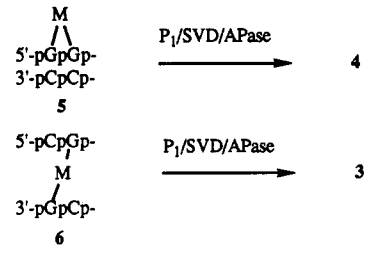
FIGURE 2: Ultraviolet and circular dichroism spectra of adducts **3** and **4**. (A) Ultraviolet spectra in 0.01 M potassium phosphate buffer, pH 7.0, of (a) adduct **3** and (b) adduct **4**, measured by using a Cary 14 spectrophotometer. (B) Circular dichroism spectra in water of (a) adduct **4** and (b) adduct **3**, measured by a Jasco J500A spectropolarimeter.

to that of **3** (Figure 2A). The ultraviolet circular dichroism spectrum is highly complex and resembles that of bisadduct **3**, rather than that of monoadducts **2a** and **2b** (Figure 2B; Tomasz et al., 1987). The visible CD spectrum indicates a negative minimum at 540 nm, characteristic of 1″α substitution stereochemistry (data not shown; Tomasz et al., 1984).

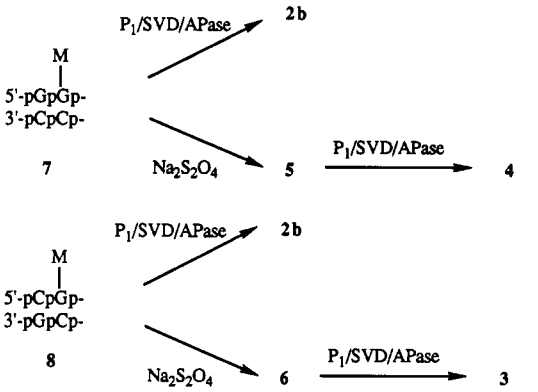
Chemical Conversions To Prove the Structure of **4** (*Scheme II*). Treatment of both poly(dG)·poly(dC) and poly(dG-dC)·poly(dG-dC) with monofunctionally activated MC yielded, after enzymatic digestion, the known monoadduct **2b** (Figure 3a,b; Tomasz et al., 1986). However, when the two monoadducted complexes were first reactivated briefly with $Na_2S_2O_4$, subsequent enzymatic digestion indicated the adducts **4** and **3**, respectively, as major products (Figure 3c,d). (The minor peaks in Figure 3c are **2a** and **3**.) This proves that the new adduct **4** is derived from the monoadduct **2b** by activation of the C10″ function of **2b** by $Na_2S_2O_4$ in analogy to the known **2b** → **3** transformation in poly(dG-dC)·poly(dG-dC). Furthermore, it is clear that the **2b** → **4** transformation is dependent on the poly(dG)·poly(dC) sequence. (It should be noted that *unbound* **2b** gives a bisulfite adduct, **2c**, upon $Na_2S_2O_4$ treatment; McGuinness et al., 1991).

Enzymatic hydrolysis of the isolated adduct **4** to adduct **3** was observed to occur very slowly, upon incubation with SVD and alkaline phosphatase (approximately 25% conversion after 2 days; Figure 4). This transformation proves that in **4** both C1″ and C10″ are linked to N^2 of a dG residue. It does not exclude, however, that one (or more) extra nucleotide is present

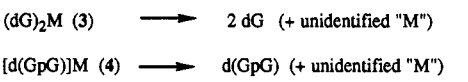
Scheme I



Scheme II



Scheme III



in **4**, between the two cross-linked dGs, as in some *cis*-DDP and *trans*-DDP intrastrand cross-links (Eastman, 1986; Fichtinger-Schepman et al., 1985). Proof of *direct* phosphodiester linkage between the two dGs of adduct **4** was provided by the transformation of **4** upon hot piperidine treatment into d(GpG), identified by coelution with the authentic d(GpG) standard (Figure 5). In the comparative control experiment, adduct **3** gave dG (Figure 5), as reported previously (Borowy-Borowski et al., 1990a). The results are summarized in Scheme III.

In summary, the identity of **4** was rigorously proven by the following set of results: (i) The UV and CD spectra are characteristic of M–$(dG)_2$ adducts (Tomasz et al., 1987). In particular, the A_{254}:A_{312} ratio in the spectrum of the known bisadduct **3** is identical to that of **4** (2.98 and 3.00, respectively). This is characteristic of a bisadduct; the A_{254}:A_{312} ratio in monoadducts (M–dG; **2a** and **2b**) is lower (2.5–2.8). (ii) Formation of **4** is dependent on the presence of at least two adjacent guanines in oligo- or polynucleotides (Figure 1). (iii) The enzymatic and chemical conversions from or into known compounds (Schemes I–III) together verify the structure of **4**. This obviates the need for mass spectral confirmation, which defied all attempts in the case of this elusive substance.

None of the above data, however, allowed us to distinguish between the two isomeric forms of **4**, i.e., **4a** and **4b**. Isomer **4a** is favored by the mechanism of the formation of MC cross-links (see Discussion).

Relative yields of intrastrand and interstrand cross-link adducts **3** *and* **4** in various substrates were determined from areas of corresponding HPLC peaks, as illustrated in Figure 1. The results are given in Table I. The reactions involving the first five substrates in the table were designed to compare relative yields when CG·CG sites and GG·CC sites were present at the same initial concentration. Oligonucleotides 1 and 2 and poly(dG)·poly(dC) and poly(dG-dC)·poly(dG-dC) represented two such pairs, while oligonucleotide 3 had both

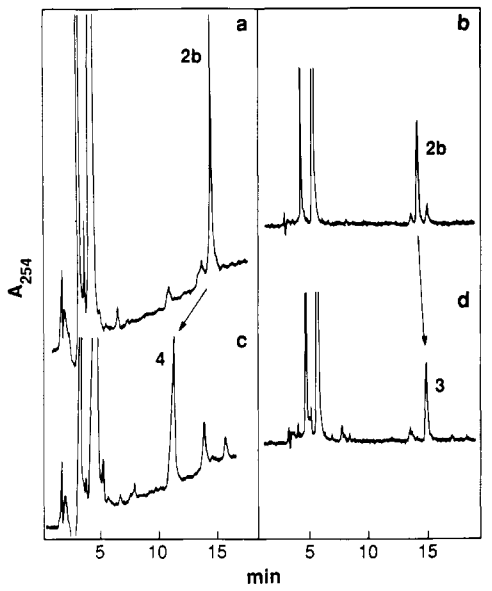


FIGURE 3: Conversion of monofunctionally bound MC residues to cross-link adducts upon reactivation by $Na_2S_2O_4$ in MC–poly(dG)·poly(dC) and MC–poly(dG-dC)·poly(dG-dC) complexes. HPLC patterns from enzymatic digestion of (a) **2b**–poly(dG)·poly(dC) complex before and (b) after $Na_2S_2O_4$ reactivation; (c) **2b**–poly(dG-dC)·poly(dG-dC) complex before and (d) after $Na_2S_2O_4$ reactivation. The arrows indicate conversion of the adducts as seen by HPLC peak shifts.

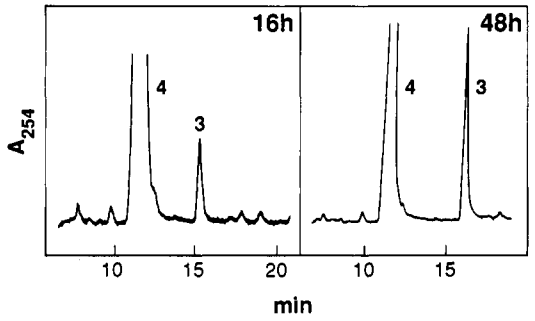


FIGURE 4: HPLC pattern from prolonged enzymatic treatment of the intrastrand cross-link adduct **4**. (a) Treatment with SVD and alkaline phosphatase for 16 h. (b) Same as (a), and then incubation continued until 64 h total; fresh enzymes were added at 16 and 40 h.

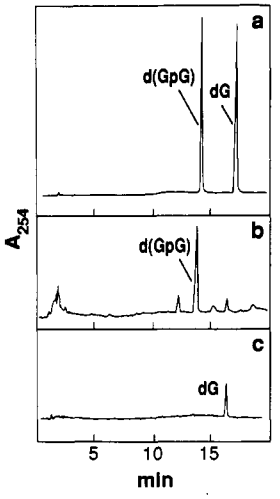
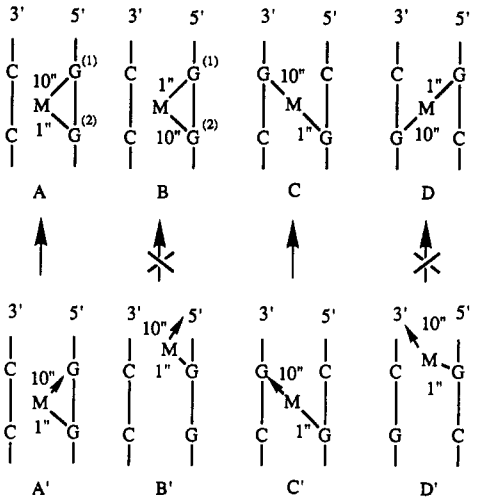


FIGURE 5: HPLC pattern and identification of the products formed upon treatment of adduct **4** with hot piperidine. (a) Mixture of d(GpG) and dG standards. (b) Reaction mixture of adduct **4** treated with hot piperidine. (c) Reaction mixture of adduct **3** treated with hot piperidine. HPLC conditions: 0–30% methanol gradient in H_2O, in 30 min; flow rate, 1 mL/min; analytical column.

Scheme IV



sites within the same molecule. In each case, cross-linking of guanines at CG·CG sites which yields interstrand cross-link **3** was favored significantly over that at GG·CC sites, yielding intrastrand cross-link **4**: 4.8-, 2.8-, and 1.7-fold preferences, were observed, respectively. The two natural DNAs also exhibited similar preference. After the relative yields of the two adducts were normalized to the relative CpG/GpG frequencies of the DNA, 2.6- and 6.3-fold relative preferences were obtained for interstrand over intrastrand cross-links for *M. luteus* and calf thymus DNAs, respectively.

The values of relative yields are related to the relative rates of formation of **3** and **4**, since first-order kinetics has been shown previously for this type of reaction (i.e., $Na_2S_2O_4$ as reducing agent, large excess of MC over oligonucleotide, etc.) (Teng et al., 1989; Borowy-Borowski et al., 1990b).

Molecular Models of **4a** *and* **4b** *Incorporated in the Two Central G Residues of the Decamer Duplex* [*d*(*GCATGGATGC*)]·[*d*(*GCATCCATGC*)] (**4a**–*Decamer and* **4b**–*Decamer*; *Figure 6*). The structures represent the two possible orientations of the mitomycin in the cross-linked segment of the decamers. (The simple diagrams A and B in Scheme IV point out succinctly the difference between the covalent structures of **4a**–decamer and **4b**–decamer.) The total energies obtained after minimizing the structures of **4a**–decamer, **4b**–decamer, and decamer itself were −1013.4, −1005.0, and −1009.3 kcal/mol, respectively. Both cross-linked decamers exhibit widening of the minor groove at the cross-linked site: **4a**–decamer at the TGG·CCA, and **4b**–decamer at the GAT·ATC segments. This effect is accompanied by constriction of the minor groove on both sides of these segments. Typically, 14.0–16.0-Å widths were measured at the widened region, and 11.0–11.6 Å in narrowed regions, compared to a constant width of 12.0–13.5 Å in the control decamer. Sugar–base orientations are all anti, and all base pairs and base stacking appear preserved in both drug-substituted models. Qualitatively, these latter features have been observed in all previous models of MC *interstrand cross-links* in oligomers (Rao et al., 1986; Tomasz et al., 1987; Teng et al., 1989; Arora et al., 1990),[4] as well as in the experimentally determined model of a CpG-cross-linked duplex hexamer (Norman et al., 1990). The orientation of MC in **4a**–decamer is characterized by the placement of the 2″-NH_3(+) group at the bottom of the minor groove, contacting both the O^2 of C15 and N^3 of A7 through H-bonds, while the indoloquinone chromophore, for the most part, is situated in the 5′-direction from the cross-link, asymmetrically disposed in the groove, in van der Waals contact with the backbone of the *opposite strand*. Two

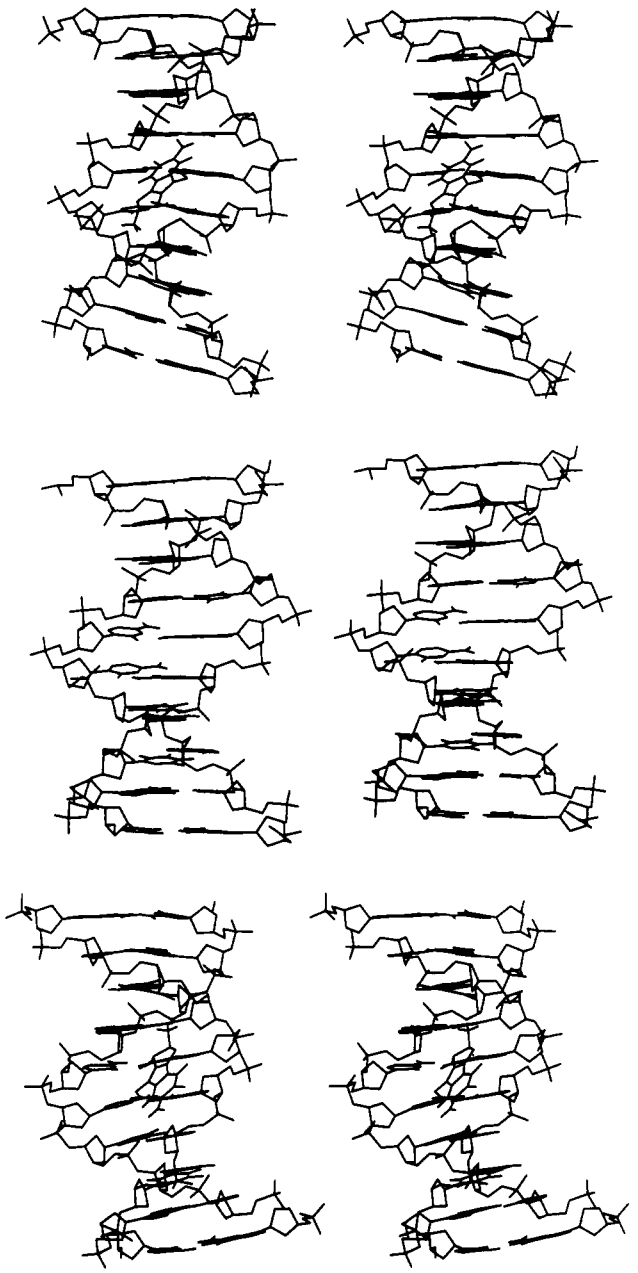


FIGURE 6: Stereomodels of (top) **4a**–decamer, (middle) unmodified decamer, and (bottom) **4b**–decamer, from molecular modeling. **4a**–decamer has adduct **4a** incorporated in place of the two central G residues G5 and G6 in the decamer duplex [d(G1-C2-A3-T4-G5-G6-A7-T8-G9-C10)]·[d(G11-C12-A13-T14-C15-C16-A17-T18-G19-C20)]. **4b**–decamer has adduct **4b** incorporated similarly in the same.

H-bonds to this backbone, by the 7″-NH_2 hydrogens are also apparent. This drug orientation is quite similar to that in the experimental *interstrand cross-linked* hexamer structure and, for that matter, to the previous computed models of the CpG cross-link. A new feature, never observed with the interstrand cross-link models, however, is that both **4a**– and **4b**–decamers are bent. Visual inspection indicates that the direction of bending is toward the minor groove. The bending effect visually appears to originate from backbone distortion and constriction of the groove between the G6·C15 and A7·T14 base pairs in **4a**–decamer and from the same between the G5·C16 and A17·T4 base pairs in the **4b**–decamer. The distances between the two cross-linked N^2 atoms of G5 and G6 are 3.1 Å in both **4a**– and **4b**–decamers and 4.3 Å in the decamer itself.

Discussion

The crucial characteristic of the intrastrand cross-link adduct **4** which distinguishes it from the interstrand cross-link **3** is the presence of a phosphodiester group between the two guanine residues, which apparently survives digestion by SVD (Scheme I). This is analogous to the resistance of the phosphate in *cis*-DDP intrastrand cross-links to nuclease P_1, distinguishing them from interstrand cross-links which are fully cleaved (Eastman, 1986; Fichtinger-Schepman et al., 1985). Nuclease P_1 does not behave analogously toward MC adducts, however, since it does not cleave the 3′-phosphodiester of *any* MC monoadduct or cross-link (Tomasz et al., 1986; Borowy-Borowski et al., 1990a). SVD, on the other hand, readily cleaves the phosphodiester bonds of the adducts shown in structures **6**, **7**, and **8**, but it cleaves only very slowly that of the intrastrand adduct in structure **5**. The reason for this may be the constrained, cyclic structure of **4**, which is likely to be unfavorable as a substrate to SVD-catalyzed cleavage. The enzymatic digestion protocol of the present work was optimized for the selective release of adducts **2a**, **2b**, and **3** as nucleosides and adduct **4** as dinucleoside phosphate. In our previous reports on MC–DNA adducts, **4** was not detected because it was hydrolyzed to **3** under the different, more forceful enzymatic digestion conditions.[3]

Adduct **4** was obtained only in very small quantities; not more than 200 μg was available at any one time. Its only sources were DNA and synthetic oligo- and polynucleotides, substituted by MC at very low overall levels (≤0.02 drug: nucleotide ratio). At higher levels of substitution, this compound disappeared and **2c**, a bisulfite adduct, artifact of the $Na_2S_2O_4$/MC activation system, was formed instead (McGuinness et al., 1991).

The direct cross-linking of d(GpG) could not be achieved, in contrast to the facile cross-linking of this substrate by cisplatin (Girault et al., 1982). This failure may be due to the fast decay of the activated form of MC in the $Na_2S_2O_4$ activation system (Cera et al., 1989; McGuinness et al., 1991). Only duplex oligonucleotides and polynucleotides but not low molecular weight guanine derivatives are reactive enough to be cross-linked by MC under these conditions (Tomasz et al., 1974).

Unfortunately, the powerful methods utilizing NMR, FT-IR, CD, and mass spectroscopies, which were developed in our laboratories for structural determination of MC–DNA adducts (Verdine & Nakanishi, 1985; Tomasz et al., 1987) could not be successfully applied to adduct **4**, due most likely to our failure to obtain a peracetylated derivative of this compound; such derivatives were key to NMR and MS analysis of previous MC adducts (Tomasz et al., 1987). Therefore, we utilized chemical and enzymatic transformations, monitored by HPLC, at a microgram scale to successfully identify adduct **4**. The data did not distinguish, however, between the two isomeric forms **4a** and **4b**, which differ in the orientation of the mitosene with respect to dinucleotide chain polarity (see the simple diagram of their DNA-incorporated forms A and B in Scheme IV). The fact that **4** could not be separated into two components by HPLC strongly suggests that it is a single isomer. The following indicates that this isomer is adduct **4a**.

It is known that the interstrand cross-linking mechanism involves two distinct steps: formation of monoadduct **2b**,

[3] DNase I instead of nuclease P_1 was employed as endonuclease, as was a much higher quantity of SVD and a much longer incubation period (24 h as compared to 2 h in the present protocol). We suspect that DNase I itself may *preferentially* cleave the phosphate in **4** when **4** is incorporated in the polynucleotide.

followed by attack on its C10″ position by a second guanine (Tomasz et al., 1988a; Borowy-Borowski et al., 1990b). It was also shown by modeling studies that this monoadduct fits well in the minor groove in only one, unique orientation, namely, one in which its C10″ position points in the 5′-direction from the C1″-linked guanine (Tomasz et al., 1986; Remers et al., 1988; Millard et al., 1990; Arora et al., 1990).[4] It is obvious from the models that the second, cross-linking step can only occur if the fixed, oriented monoadduct finds a guanine near its C10″ activated position. This condition is met at the CpG, but not at the GpC step in duplex DNA (Scheme IV; diagrams C′ → C and D′ → D). Experimental results on the sequence specificity of the interstrand cross-link are fully consistent with this prediction. It has been shown recently in three independent laboratories that of the two potential interstrand cross-links, at CpG and GpC (see Scheme IV), only the CpG cross-link is formed in DNA (Teng et al., 1989; Weidner et al., 1989; Borowy-Borowski et al., 1990b). Since the same two-step mechanism via the monoadduct **2b** as intermediate is shown here to apply also to the GpG intrastrand cross-link (Figure 3; Scheme II), it follows that the oriented monoadduct could only lead to intrastrand cross-link isomer A and never to cross-link isomer B (Scheme IV; A′ → A and B′ → B). Isomer A corresponds to adduct **4a** and therefore **4a** is most likely to be the correct structure for the cross-link adduct. The molecular modeling results of **4a**–decamer and **4b**–decamer are in themselves not decisive on this issue. **4a**–decamer is favored by a total energy difference of only −8 kcal/mol over **4b**–decamer. Both Crothers and co-workers (Teng et al., 1989) and Arora et al. (1990) obtained similarly small values for ΔE_{total} for the analogous CpG/GpC cross-link isomer pair (C and D in Scheme IV), indicating that the observed exclusive formation of the CpG cross-link is not due to its enhanced stability over the GpC cross-link isomer but rather to a "kinetic factor," which is clearly the orientation of the monoadduct. The intrastrand GpG cross-link isomerism is completely analogous to this case.

It is now apparent that both interstrand and intrastrand cross-links by MC are formed in DNA, at the target sites CpG and GpG, respectively, while a third potential site, GpC, is not cross-linkable. Although the intrastrand cross-link is reported here to be formed in cell-free systems, preliminary evidence indicates that both interstrand and intrastrand cross-link adducts **3** and **4** were also formed in EMT6 mouse mammary tumor cells, treated with radiolabeled MC (R. Bizanek, unpublished observation). Adduct **3** has been isolated previously from rat liver DNA (Tomasz et al., 1987).

The preferential formation of interstrand cross-links over intrastrand ones, seen by adduct yield analysis (Table I), is based most likely on preferential *monoadduct* formation (the requisite intermediate for cross-linking) at the 5′-CG sequence over that at 5′-GG.[5] The strong influence of the 5′-neighboring base on the sequence specificity of MC monoadduct formation in DNA was first reported by Li and Kohn (1991), who noted that only guanines at 5′-CG and 5′-GG sequences had appreciable reactivity. We showed recently (Kumar et al., 1992) that monoadduct formation in duplex dodecanucleotide models proceeded with a 2.4 times higher yield at 5′-CGT than at 5′-GGT. A similar monoadduct yield ratio (2.2:1) was observed at 5′-CGA vs 5′-GGA sequences. The 3′ base in such NGN triplets was shown to have a modest modulating influence on the yield, which may explain the *variation* of the observed **3**:**4** adduct yield ratios in the different substrates in Table I. It is also possible that guanines at certain longer $(CG)_n$ sequences in DNA have enhanced reactivity (Teng et al., 1989). This would further vary the ratios from those observed with oligonucleotides of defined sequence.

Molecular Modeling. The observed bending of both intrastrand cross-linked decamer models is a notable effect of the mitomycin intrastrand cross-link. It is significant that neither the MC monoadduct nor the MC interstrand cross-link models showed bending of DNA, using the same AMBER force field, docking, and minimization methodologies for the computations (Rao et al., 1986; Tomasz et al., 1987; Remers et al., 1988; Teng et al., 1989; Millard et al., 1990; Arora et al., 1990). It is likely, therefore, that the bent intrastrand cross-link model is a consequence of the covalent constraint unique to this MC–DNA complex: the N^2 atoms of two adjacent guanines are pinched together by the cross-link, moving them from an average 4.3-Å distance in B-DNA to 3.1 Å in the cross-linked structures, as required by the fixed distance between the C1″ and C10″ bonding sites of the mitosene. This constraint is likely to cause greater distortions than that required in the interstrand cross-link case. In the latter, the GN^2–GN^2 distance between the two strands in average B-DNA is only 3.6 Å, and therefore it is more easily distorted to 3.1 Å by the cross-link. What is more, the covalent constraint is *not confined to one strand* as it is in the present case. The model of **4a**–decamer indicates that the localized covalent constraint in one strand combined with the tendency to maintain the usual set of energetically favorable mitosene–DNA binding interactions in the other strand [2″-NH_3(+) H-bonds to both strands; indoloquinone van der Waals contacts and H-bonds to the other strand's backbone; see Results and the above references] gives rise to a bent DNA structure. Bending of DNA due to certain covalent modifications was reported. Recent examples are pyrimidine dimers (Husain et al., 1988) *cis*-DDP and *trans*-DDP intrastrand cross-links (Rice et al., 1988), and the CC-1065–DNA complex (Lee et al., 1991). Interestingly, bending caused specifically by intrastrand cross-link adducts but not by a monoadduct was observed in the platinum drug series (Bellon & Lippard, 1990). The structural factors determining this effect are unique to each case. However, the plausible hypothesis of DNA bending by the intrastrand MC cross-link, based on molecular modeling alone, awaits critical testing. This should be possible by applying the elegant and powerful experimental probe for drug-induced DNA bending, described by Crothers, Lippard, and their co-workers (Rice et al., 1988).

The mitomycin–DNA intrastrand cross-link represents only the second known example in which a bifunctional agent links two bases together in the same strand. The first, prototypical agents are the antitumor platinum drugs, *cis*- and *trans*-DDP. Assigning biological effects to a DNA adduct of a specific type among several others and correlating them with the specific structural characteristics of the adduct is a major objective in platinum drug research (Lepre et al., 1990), as well as in the field of DNA damage in general. The discovery of the MC intrastrand cross-link, and its characterization and easy analytical distinction from the interstrand cross-link and other MC adducts, presents an opportunity for similar studies of mitomycin-induced DNA damage.

ACKNOWLEDGMENTS

We thank Dr. D. M. Vyas of Bristol-Myers Squibb Co., Wallingford, CT, for a generous gift of mitomycin C. We

[4] The last reference is a study of *N*-methylmitomycin C (porfiromycin).

[5] The reader is reminded that monoadduct formation is considered at the 3′-G of 5′-GpG, since, due to the orientation of monoadduct, only that monoadduct will give rise to the cross-link (cf. Scheme IV).

acknowledge helpful discussions about molecular modeling with Dr. Gary J. Quigley, Hunter College.

References

Arora, S. K., Cox, M. B., & Arjunan, P. (1990) *J. Med. Chem. 33*, 3000–3008.

Bellon, S. F., & Lippard, G. J. (1990) *Biophys. Chem. 35*, 179–188.

Bellon, S. F., Coleman, J. H., & Lippard, S. J. (1991) *Biochemistry 30*, 8026–8035.

Borowy-Borowski, H., Lipman, R., Chowdary, D., & Tomasz, M. (1990a) *Biochemistry 29*, 2992–2999.

Borowy-Borowski, H., Lipman, R., & Tomasz, M. (1990b) *Biochemistry 29*, 2999–3004.

Cera, C., Egbertson, M., Teng, S. P., Crothers, D. M., & Danishefsky, S. J. (1989) *Biochemistry 28*, 5665–5669.

Danishefsky, S. J., & Ciufolini, M. (1984) *J. Am. Chem. Soc. 106*, 6424.

Eastman, A. (1986) *Biochemistry 25*, 3192–3215.

Egbertson, M., & Danishefsky, S. J. (1987) *J. Am. Chem. Soc. 109*, 2204–2205.

Fasman, G. D. (1975) *Handbook of Biochemistry and Molecular Biology*, 3rd ed., Vol. I, p 589, CRC Press, Cleveland, OH.

Fichtinger-Schepman, A. M. J., Van der Veer, J. L., den Hartog, J. H. J., Lohman, P. H. M., & Reedijk, J. (1985) *Biochemistry 24*, 707.

Girault, J. P., Chottard, G., Lallemand, J. Y., & Chottard, J. C. (1982) *Biochemistry 21*, 1352–1356.

Hoey, B. M., Butler, J., & Swallow, A. J. (1988) *Biochemistry 27*, 2608–2614.

Husain, I., Griffith, J., & Sancar, A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2558–2562.

Iyer, V. N., & Szybalski, W. (1964) *Science 145*, 55–58.

Keyes, S. R., Rockwell, S., & Sartorelli, A. C. (1984) *Cancer Res. 45*, 3642.

Kohn, H., & Zein, N. (1983) *J. Am. Chem. Soc. 105*, 4105–4106.

Kumar, S., Lipman, R., & Tomasz, M. (1992) *Biochemistry 31*, 1399–1407.

Lee, C.-S., Sun, D., Kizu, R., & Hurley, L. H. (1991) *Chem. Res. Toxicol 4*, 203–213.

Lepre, C. A., Chassot, L., Costello, C. E., & Lippard, S. J. (1990) *Biochemistry 29*, 811.

Li, V., & Kohn, H. (1991) *J. Am. Chem. Soc. 113*, 275–283.

Lin, A. J., Cosby, L. A., & Sartorelli, A. C. (1976) *ACS Symp. Ser. 30*, 71–80.

Mc Guinness, B. F., Lipman, R., Nakanishi, K., & Tomasz, M. (1991) *J. Org. Chem. 56*, 4826–4829.

Millard, J. T., Weidner, M. F., Raucher, S., & Hopkins, P. B. (1990) *J. Am. Chem. Soc. 112*, 3637–3641.

Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Canfield, C., Chang, G., Hendrickson, Th., Still, W. C. (1990) *J. Comput. Chem. 11*, 440–467.

Moore, H. W. (1977) *Science 197*, 527–532.

Norman, D., Live, D., Sastry, M., Lipman, R., Hingerty, B. E., Tomasz, M., Broyde, S., & Patel, D. J. (1990) *Biochemistry 29*, 2861–2876.

Pan, S.-S., Andrews, P. A., Glover, C. J., & Bachur, N. R. (1984) *J. Biol. Chem. 259*, 959–966.

Peterson, D. M., & Fisher, J. (1986) *Biochemistry 25*, 4077–4084.

Rao, N. S., Singh, U. C., & Kollman, P. A. (1986) *J. Am. Chem. Soc. 108*, 2058–2068.

Remers, W. A., Rao, S. N., Wunz, T. P., & Kollman, P. A. (1988) *J. Med. Chem. 31*, 1612–1620.

Rice, J. A., Crothers, D. M., Pinto, A. L., & Lippard, S. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4158–4161.

Szybalski, W., & Iyer, V. N. (1964) *Fed. Proc. 23*, 946–957.

Teng, S. P., Woodson, S. A., & Crothers, D. M. (1989) *Biochemistry 28*, 3901–3907.

Tomasz, M., & Lipman, R. (1981) *Biochemistry 20*, 5056–5061.

Tomasz, M., Mercado, C. M., Olson, J., & Chatterjie, N. (1974) *Biochemistry 13*, 4878–4887.

Tomasz, M., Jung, M., Verdine, G. L., & Nakanishi, K. (1984) *J. Am. Chem. Soc. 106*, 7367–7370.

Tomasz, M., Chowdary, C., Lipman, R., Shimotakahara, S., Veiro, D., Walker, V., & Verdine, G. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6702–6706.

Tomasz, M., Lipman, R., Chowdary, D., Pawlak, J., Verdine, G. L., & Nakanishi, K. (1987) *Science 235*, 1204–1208.

Tomasz, M., Chawla, A. K., & Lipman, R. (1988a) *Biochemistry 27*, 3182–3187.

Tomasz, M., Lipman, R., McGuinness, B. F., & Nakanishi, K. (1988b) *J. Am. Chem. Soc. 110*, 5892–5896.

Verdine, G. L., & Nakanishi, K. (1985) *J. Am. Chem. Soc. 107*, 6118–6120.

Webb, J. S., Cosulich, D. B., Mowat, J. H., Patrick, J. B., Broschard, R. W., Meyer, W. E., Williams, R. P., Wolf, C. F., Fulmor, W., Pidacks, C., & Lancaster, J. E. (1962) *J. Am. Chem. Soc. 84*, 3185–3186.

Weidner, M. F., Millard, J. T., & Hopkins, P. B. (1989) *J. Am. Chem. Soc. 111*, 9270–9272.

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

Zein, N., & Kohn, H. (1987) *J. Am. Chem. Soc. 109*, 1576.