0026-895X/91/030299-08$03.00/0

MOLECULAR PHARMACOLOGY, 39:299-306

Selective Inhibition of Sequence-Specific Protein-DNA Interactions by Incorporation of 6-Thioguanine: Cleavage by Restriction Endonucleases

LYNN M. IWANIEC, JEFFREY J. KROLL, W. MAURICE ROETHEL, and JONATHAN MAYBAUM

Department of Pharmacology, Upjohn Center for Clinical Pharmacology, University of Michigan Medical School, Ann Arbor, Michigan 48109-0504

Received September 4, 1990; Accepted December 21, 1990

SUMMARY

Incorporation of the antileukemic agent 6-thioguanine (TG) into cellular DNA has been demonstrated to be a major determinant of its cytotoxicity. We have previously shown that complete replacement of G by TG within one DNA strand of the SV40 origin of replication can completely inhibit sequence-specific binding of the viral replication protein T antigen. The aim of the present study was to determine the effect of more selective TG substitutions on DNA-protein interactions, by utilizing the simpler base recognition sequence motifs of restriction endonucleases. In the first part of our study, we replaced G with TG in one or two of four possible sites within the duplex hexameric recognition sequence of *Bam*HI (5′-G↓GATCC-3′), by enzymatic extension of primed oligonucleotides. This extension was stalled, but not completely inhibited, at locations where insertion of consecutive TG moieties was required. Both strands of molecules containing a single substitution were cleaved by *Bam*HI at reduced rates, with the substituted strand inhibited to a greater degree. In molecules containing two substitutions, neither strand was cut by *Bam*HI. In contrast, we found that scission of these same mono- and disubstituted substrates by the less stringent isoschizomer *Mbo*I (5′-N↓GATCN-3′) was inhibited only slightly. In the second part of our study, we investigated the effect of analog substitution on scission by the type II-S enzymes *Alw*I and *Fok*I, in order to separately determine the effects of restriction site modification versus scission site modification. We found that the reactivity of these enzymes was completely abolished by TG substitution within the recognition site, whereas substitution at the scission site had no effect. Our results demonstrate that infrequent TG substitutions within symmetric DNA sequences can inhibit sequence-specific interactions in an asymmetric fashion. In addition, although previous reports have shown that TG forms a relatively weak base pair with cytosine, it appears that the inhibition of restriction endonuclease-mediated cleavage resulting from TG incorporation is a function of the sequence requirements of the protein and not a general consequence of disrupted base-pairing at the recognition locus. These data support the idea that the cytotoxic consequences of TG incorporation may be due to inhibition of sequence-specific protein-DNA interactions.

Analogs of purine and pyrimidine nucleosides and bases have a number of therapeutic applications, most prominently in the treatment of neoplasms. The underlying basis for their use in this capacity has been their disruption of DNA synthesis and function and, therefore, of cellular proliferation. Although a generalized disturbance of DNA function provides a modest degree of selectivity for neoplastic cells, recent advances in the understanding of the roles of specific DNA sequences and motifs raise the possibility that analogs could be designed that would affect cellular processes in a much more selective manner. To this end, we and others have begun to define the effects on sequence-specific DNA-protein interactions caused by incorporation of some of the analogs already known to have clinical utility.

This work was supported by National Institutes of Health-National Cancer Institute Grant R01-40260.

Although the effects of incorporated analogs are quite varied, these agents may be roughly divided into two groups, depending upon the rapidity with which their impact on DNA synthesis becomes apparent. One group, composed mainly of analogs with modifications in the sugar moiety, tends to cause relatively rapid and severe disturbances in DNA strand elongation. These disturbances may be manifested in a number of ways, including pausing, stalling, or termination of chain elongation, competitive inhibition of DNA polymerases, and depletion of nucleoside triphosphates. Agents in this category would include, e.g., the arabinosyl derivatives of various bases, dideoxy nucleosides, and acyclic or carbocyclic nucleoside analogs. Although compounds of this type have recently been shown to exhibit sequence selectivity in their actions (1–3), it seems likely that the magnitude of their immediate inhibition of DNA synthesis would overshadow any more subtle effects their incorporation might have on processes controlled by DNA-protein interactions.

ABBREVIATIONS: TG, 6-thioguanine; TdGTP, 6-thio-2′-deoxyguanosine-5′-triphosphate; dNTP, 2′-deoxynucleoside-5′-triphosphate.

In contrast, several analogs containing alterations in the base moiety are incorporated to a much greater extent than are sugar-modified analogs and often exert their effects in a more delayed manner (4–7). Our recent studies with one such analog, TG, have suggested that this clinically useful antileukemic drug may indeed exert its cytotoxic effects by selectively disrupting sequence-specific DNA-protein interactions (8). Results obtained in intact cells show that, although TG has little impact on DNA replication during the S phase in which the drug is first incorporated, the presence of this incorporated analog in the template strand during the subsequent round of replication induces drastic nonrandom DNA damage (9). Experiments using SV40-infected cells showed that the effects of TG incorporation on replication of this virus [which is thought to be a good model for chromosomal replication (10, 11)] are directed towards initiation, rather than elongation of nascent DNA strands (8). Furthermore, we observed in a cell-free system that extensive incorporation of TG abolished the sequence-specific binding of the viral protein T antigen to SV40 origin sequences (8). Because T antigen binding is an absolute requirement for initiation of origin-dependent replication in this system, these data support the idea that the major biological consequence of TG incorporation in intact cells may be disruption of sequence-specific initiation of DNA replication.

Although our previous *in vitro* studies showed that extensive TG incorporation is capable of disrupting DNA-protein interactions, the effects of infrequent incorporation events (as would normally be seen *in vivo*) are unknown. Furthermore, it is not known whether TG incorporation would disrupt protein-binding recognition sites indiscriminately or in a manner dependent upon the protein and recognition sequence involved. We have, therefore, undertaken the present study, the overall objective of which is to determine how selective insertion of TG affects DNA function in a well defined, easily accessible, model system, i.e., restriction endonuclease cleavage of duplex oligonucleotides.

In the first part of our study, we synthesized DNA molecules with TG asymmetrically substituted within or near the recognition sequence of the isoschizomeric pair *Bam*HI and *Mbo*I, and we examined the effects of TG substitution on cleavage of both the substituted and the unsubstituted strands. In the second part of our study, we examined the effects of TG substitution on the actions of *Alw*I and *Fok*I. These enzymes are among a small group that have nonpalindromic recognition sequences and for which the site of cleavage is distant from the recognition sequence. By substituting TG for G in the recognition sequence or the scission site for these enzymes, we were able to determine the relative effects of TG substitution on these two processes.

Materials and Methods

Synthesis of TdGTP. The deoxyribonucleoside triphosphate form of TG was synthesized from the starting material, thiodeoxyguanosine, as described previously (8). The finished product was analyzed chromatographically on a Dynamax 300A C-18 reverse phase column, using a linear gradient of acetonitrile (0.40%) in a buffer system that also contained 0.1 M triethylammonium acetate (pH 7.0), and was determined to contain less than 1% contaminating dGTP.

Synthesis of TG-substituted duplex DNA oligonucleotides. The general approach used here was to insert TG moieties by enzymatic extension of chemically synthesized primer-template pairs. Template sequences and primer lengths were designed so as to achieve TG substitution patterns that would allow us to test various hypotheses. Fig. 1 shows an example in which TG was incorporated into the internal G site in the *Bam*HI/*Mbo*I recognition sequence, as well as into three consecutive G sites just outside of this recognition sequence, by extension of a 28-base primer annealed to a 48-base template. Alternatively, use of a 27-base primer (the 28-mer without the 3′ G residue) allowed TG incorporation at the external G location in the *Bam*HI/*Mbo*I site, in addition to incorporation obtained with the 28-mer. A 34-base primer was used to permit incorporation into the three consecutive G loci but not within the enzyme recognition site.

Different primer-template pairs were used to create substrates for the class II-S restriction enzyme *Alw*I, in which either the recognition site or the scission site contained analog substitutions. Fig. 2 illustrates this strategy.

End-labeling was performed using T4 polynucleotide kinase (10 units/μl; Bethesda Research Laboratories) and [γ-^{32}P]ATP (300 Ci/mmol; Amersham), according to standard procedures. These end-labeled primer molecules were then annealed to the appropriate template molecules by heating at 75° for 2 min, in buffer containing 40 mM Tris, pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl, and 5 mM dithiothreitol, and slow cooling to 25°. Template molecules were present in a 2-fold molar excess over primer molecules in the annealing reaction. Annealed primer/template molecules were washed three times with TE (10 mM Tris·HCl, 1 mM EDTA, pH 8.0) by filtration through a Centricon 30 microconcentrator (Amicon Corp.).

Annealed primer/template molecules were extended with Sequenase (modified T7 polymerase version 2.0, 12 units/μl, 4 units/μg of DNA; United States Biochemicals), in the presence of 1 mM dATP, dCTP, dTTP, and either dGTP or TdGTP, for control and TG-substituted molecules, respectively. Reactions were incubated at 37° for 2.5 hr, except where indicated. The products were extracted with chloroform and precipitated with ethanol before use as restriction enzyme substrates.

High performance liquid chromatography purification of fully extended duplexes from incompletely extended duplex DNA molecules. Fully extended control and TG-substituted molecules were separated from incomplete extension products by high performance liquid chromatography, using a Nucleogen 4000-10 anion exchange column. Fractions were eluted from the column with a linear gradient of LiCl (0.1–0.5 M), in a buffer system that also contained 20 mM sodium acetate and 30% acetonitrile. The gradient was run for 60 min at 1 ml/min, with the column immersed in a 70° water bath.

Fractions were collected from the column, and aliquots were analyzed both for radioactivity (in cpm/μl), by liquid scintillation counting, and for assessment of the degree of separation of fully extended molecules from incomplete extension products, by electrophoresis on a sequencing

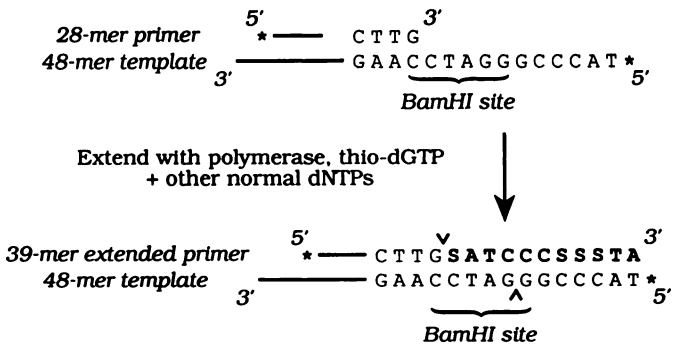

Fig. 1. Strategy for synthesis of oligonucleotides with selective TG substitutions at the *Bam*HI recognition site. In this example, a 28-base primer is annealed to a 48-base template and then extended with a polymerase and a dNTP mixture containing either dGTP or TdGTP. *Bold type* denotes the newly synthesized region of the 39-base extended primer, with TG moieties represented by an *S*. The locations of TG substitution are dictated by the length of the primer used. *Carets* show the expected cleavage sites for *Bam*HI.

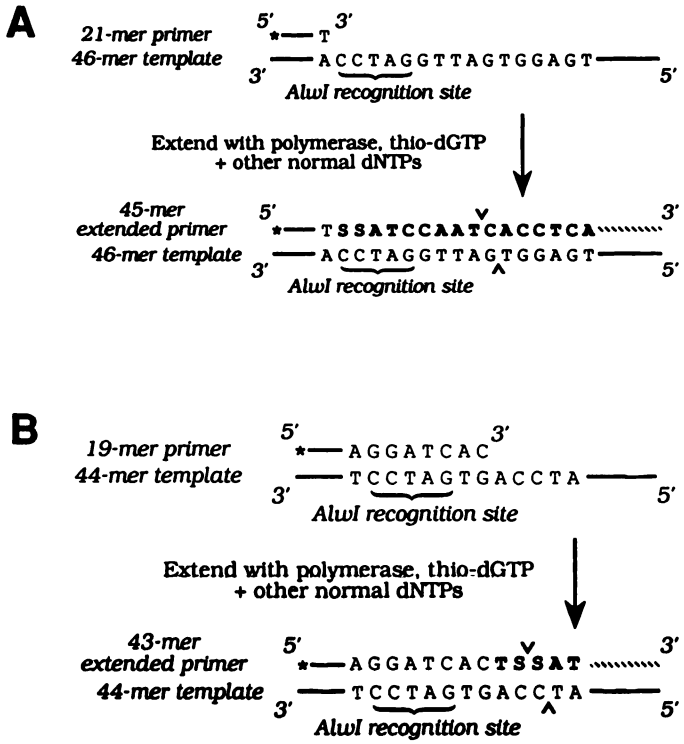

Fig. 2. Strategy for synthesis of oligonucleotides containing TG substitutions at either the recognition site (A) or the scission site (B) for *AlwI*. As in Fig. 1, the locations of substitutions are determined in part by the length of the primer used. In this case the templates used in A and B differ at the scission site, so as to prevent incorporation of TG outside of the recognition site in A. Again, newly synthesized regions are in *bold type*, TG moieties are indicated by *S*, and *carets* indicate the expected cleavage sites.

gel (12% polyacrylamide, 7 M urea, 50°). Fully extended molecules eluted as duplexes within 40–45 min of the onset of the LiCl gradient (0.345 M LiCl). These molecules could be readily separated from incomplete extension products, which eluted within 30–35 min of the initiation of the LiCl gradient. Fractions containing completely extended duplexes were pooled, desalted, and concentrated into TE buffer with a Centricon 30 microconcentrator. Typically, 40% of the total counts loaded onto the column were recovered as completely extended DNA duplexes, free from contamination by incomplete extension products.

Restriction endonuclease digestion of control and TG-substituted DNA duplexes. Duplex DNA molecules, with or without TG substitutions, were assayed as substrates for *Bam*HI endonuclease in reactions containing 10 mM Tris·HCl, pH 8.0, 100 mM NaCl, 5 mM $MgCl_2$, and 1 mM 2-mercaptoethanol. Reactions were initiated by the addition of enzyme (*Bam*HI, 12 units/μl; Boehringer Mannheim Biochemicals). Immediately before use, the enzyme stock was diluted in buffer containing 10 mM Tris·HCl, pH 7.4, 100 mM KCl, 1 mM EDTA, 1 mM dithiothreitol, 0.015% Triton X-100, and 50% glycerol. The final enzyme concentration in the assay was 0.05 units/μl. Reactions were incubated in a 15° water bath. Aliquots (10 μl/time point) of the reaction mixture were removed, pipetted into Eppendorf tubes containing 5 μl of stop solution (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue), and then placed immediately into a dry ice bath.

Reaction products were separated from undigested substrate molecules by electrophoresis on a denaturing (7 M urea) sequencing gel. Samples were electrophoresed at 50° for 2.5 hr. Gels were dried, and autoradiography was performed overnight (12–16 hr). The amount of product formed was quantitated by liquid scintillation counting of the excised areas of the gel corresponding to substrate and product.

For the kinetic assays, K_m and V_{max} values were determined with control unsubstituted DNA molecules as the substrate molecules. The final concentrations of substrate ranged from 5 nM (10 units of enzyme/pmol of substrate) to 100 nM (0.5 units of enzyme/pmol of substrate) in the assay. The specific activity of the substrate varied between experiments from 10,000 to 50,000 cpm/pmol. For experiments in which both the template and primer strands were 5′ end-labeled, the final assay concentration of substrate molecules was 175 nM (0.3 units of enzyme/pmol of substrate), and the specific activity ranged from 12,000 to 15,000 cpm/pmol.

The amount of product formed at each time point was calculated from the fraction of radioactivity in the substrate region converted to product, multiplied by the total pmol of oligonucleotide in the assay mixture, as follows:

$$\text{pmol of product}_x = \left[\frac{(\text{corrected cpm of product})_x}{(\text{corrected cpm of product})_x + (\text{cpm of substrate})_x}\right] \cdot (\text{total pmol of oligonucleotide})$$

where

$$\text{Corrected cpm of product}_x = \text{cpm of product}_x - \left[\frac{\text{cpm of product}_0}{\text{cpm of product}_0 + \text{cpm of substrate}_0}\right] \cdot (\text{cpm of product}_x + \text{cpm of substrate}_x)$$

In these studies, the value for background product cpm never exceeded 7.5% of the total cpm in a given lane.

For digestion by *Mbo*I, control and TG-substituted DNA molecules were synthesized as described above and digested in a buffer containing 10 mM Tris·HCl, pH 7.4, 100 mM NaCl, 10 mM $MgCl_2$, and 100 μg/ml bovine serum albumin. Assays were initiated by addition of *Mbo*I (5 units/μl; New England BioLabs), at a final concentration of 0.0166 units/μl, and were incubated at 15°. Reactions were terminated by pipetting of 10-μl aliquots into 5 μl of stop solution. For the kinetic assays, 2.56–250 nM control DNA was used. This corresponded to 6.4–0.66 units of enzyme/pmol of substrate. For assays of *Mbo*I digestion of control and TG-substituted molecules, a final concentration of 120 nM DNA (0.14 units of enzyme/pmol of substrate) was used.

Substrates for *Alw*I were digested with 0.1 unit/μl endonuclease (2 units/μl; New England BioLabs), in a reaction containing 10 mM Tris·HCl, pH 7.4, 10 mM $MgCl_2$, 10 mM 2-mercaptoethanol, and 100 μg/ml bovine serum albumin. Reactions were incubated at 20° and terminated by removal of 5-μl aliquots at the appropriate intervals and pipetting into 5 μl of stop solution.

Results

For the first part of our study, we synthesized DNA molecules with TG substituted within or near the recognition sequence of the isoschizomeric pair *Bam*HI and *Mbo*I. The recognition sequences of *Bam*HI and *Mbo*I are 5′-G↓GATCC-3′ and 5′-N↓GATCN-3′, respectively.

Extension of primer/template oligonucleotides with modified T7 DNA polymerase (Sequenase). We utilized a 48-mer oligonucleotide containing a single *Bam*HI/*Mbo*I site as a template for the synthesis of duplex DNA substrates with TG modifications in or near the recognition locus. As described in Materials and Methods and Fig. 1, primer oligonucleotides of various lengths were annealed to the template 48-mer and extended with Sequenase in the presence of TdGTP or dGTP.

Our initial experiments were performed to determine the optimal conditions for complete extension of the primer/template oligonucleotide pairs with Sequenase. The autoradiogram in Fig. 3 illustrates the results from an extension of the 27-mer

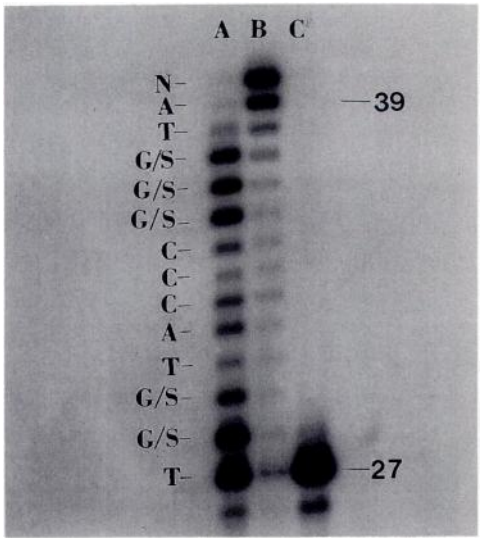

Fig. 3. Extension of primed oligonucleotides with Sequenase. In this example, a 27-base primer was 5′ end-labeled and annealed to the 48-base template, in a 1:2 primer to template ratio. Extensions were carried out in the presence of 1 mM dCTP, dTTP, dATP, and TdGTP (*lane A*), dGTP (*lane B*), or no G nucleotide (*lane C*). The samples were incubated for 45 min at 37° and then extracted and analyzed by electrophoresis and autoradiography. The sequence of the extended strand is shown on the *left*, with TG represented by *S*.

primer/48-mer template pair conducted in the presence of 1 mM TdGTP (Fig. 3, *lane A*), dGTP (Fig. 3, *lane B*), or no G (Fig. 3, *lane C*), with the normal complement of dNTPs. Fig. 3, *lane B*, shows that, in the presence of control dNTPs, the majority of the end-labeled 27-mer was utilized for extension by Sequenase after 45 min of incubation, as evidenced by the small amount of intensity remaining at the original 27-mer position on the autoradiogram. Analysis of the extension products revealed that, in addition to the anticipated 39-mer product, the polymerase also catalyzed the addition of one nucleotide to the 3′ end of the fully extended 39-mer oligonucleotide, in what appeared to be a non-template-dependent manner. That is, the replication products consisted primarily of a mixture of the predicted 39-mer and a 40-mer after 45 min of incubation with Sequenase and control dNTPs. In addition, the 0 min time point of Fig. 4A shows that, after a longer incubation period (90 min), the replication products were predominantly 40- and 41-mers.

When dGTP was replaced with TdGTP (Fig. 3, *lane A*) in the polymerization mixture, extension of the primer molecules was less efficient. The majority of the incompletely elongated DNA strands accumulated at regions where TG had been added to the growing strand. That is, the polymerase appeared to "stall" after the addition of TG to primer at positions 28 and 29, which are the two TGs within the *Bam*HI/*Mbo*I site, and at positions 35, 36, and 37, which are three contiguous TGs outside the recognition locus. When these primer/template molecules were incubated for a longer period of time (90 min) with Sequenase, they were completely extended to the predicted 39-base length (Fig. 4C, 0 min), with little or no extension beyond the end of the template.

Because the primary intention of our study was to synthesize duplex DNA molecules with specific TG substitutions, we did not characterize this phenomenon any further. We reasoned that, because the extra nucleotide(s) was (were) added several bases from the actual restriction site and because this non-template-directed addition reaction has been observed previously by Clark (12) with a variety of other prokaryotic and eukaryotic polymerases, the 40- or 41-base oligonucleotides would be equivalent to the 39-base pair oligonucleotides as substrates for the restriction endonucleases. This expectation was supported by the apparently parallel rate of disappearance of the 39-, 40-, and 41-base molecules during enzymatic digestion (see below).

***Bam*HI digestion of control unsubstituted DNA.** Fig. 4A shows a representative autoradiogram of *Bam*HI-catalyzed digestion of these molecules (0.25 μM DNA). In this study, molecules were 5′ labeled with [γ-^{32}P]ATP on the extended strand only. From the autoradiogram, we observed that the endonuclease catalyzed the hydrolysis of the correct phosphodiester bond within the ^{32}P-labeled extended strand to produce a 28-mer product and that the 40- and 41-base oligonucleotides appear to be digested to the same extent as the 39-base oligonucleotide. Digestion of the control extended strand was 50% completed after 15 min at 15° and 90% completed by 60 min. Endonuclease cleavage reactions were conducted at tempera-

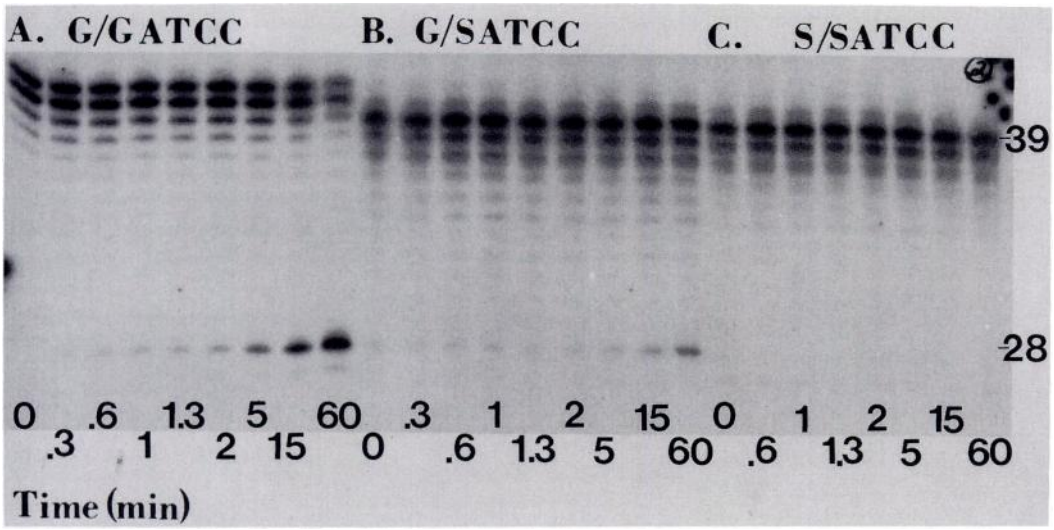

Fig. 4. Digestion of control and TG-substituted oligonucleotides by *Bam*HI. Oligonucleotides containing zero (G/GATCC) (A), one (G/SATCC) (B), or two (S/SATCC) (C), TG substitutions within the *Bam*HI recognition site were digested with that enzyme for the indicated time periods, removed, and processed for electrophoretic analysis. Markers on the *right* indicate the positions of 28- and 39-base oligomers, which correspond to the nominal products and substrates, respectively, for this reaction. In this experiment only the extended strand was labeled.

tures below the optimal temperature suggested by the supplier (37°) because our initial studies demonstrated that at 37° digestion of control DNA proceeded too rapidly for us to obtain accurate kinetic data.

Steady state kinetic analyses of the *Bam*HI-catalyzed cleavage of control duplex molecules showed that cleavage by the endonuclease was saturable and obeyed Michaelis-Menten kinetics. A Lineweaver-Burke plot of these data revealed a K_m value for the enzyme of 35 ± 3 nM and a V_{max} of 40 ± 0.53 nmol/min/unit of enzyme. George *et al.* (13) have reported a K_m value of 3.6 nM for the enzyme, using SV40 DNA as substrate. The higher value that we obtained with our 39-base pair substrate is consistent with the observation that the magnitude of the K_m for restriction endonucleases generally increases with decreasing substrate length (14).

***Bam*HI digestion of TG-substituted DNA.** Fig. 4, B and C, also shows the effects of a single TG substitution at the internal G position or of a double TG substitution at both G positions of the *Bam*HI locus. Fig. 4B shows that the single TG substitution did not appear to affect the specificity of the enzyme for scission at the correct phosphodiester bond, because the ^{32}P-labeled extended strand was cleaved to the predicted length of 28 bases. However, when compared with control substrates, the single TG substitution appeared to decrease the rate of cleavage of the DNA by *Bam*HI. Even after 60 min of incubation with *Bam*HI, the production of the 28-mer product from these substrates remained decreased, compared with control (compare Fig. 4A, 60 min, and Fig. 4B, 60 min). The effect of substituting TG for G at both G sites within the *Bam*HI locus was more pronounced (Fig. 4C). In these studies, the doubly substituted DNA did not appear to be digested even after 60 min of incubation with *Bam*HI.

From the initial digestion experiments described above, it appears that the extent of TG substitution within the *Bam*HI locus has differential effects on substrate reactivity. To further characterize these substitution effects, duplex DNA substrates were synthesized as described above but with the template unextended strand also 5′ labeled with ^{32}P. In this way, the effects of the TG substitutions on *Bam*HI digestion of both the extended and the unextended strands could be monitored within the same experiment. Control and TG-substituted oligonucleotides (0.175 μM) were digested with *Bam*HI, and the amount of product that was formed was quantitated as described above and in Materials and Methods. Fig. 5 shows the digestion of each strand (the 39-base extended and the 48-base template) of the different duplex DNA substrates. Product formation is expressed as a percentage of the total substrate molecules in the assay. For the control substrates, digestion of the template unsubstituted strand paralleled the digestion of the extended unsubstituted strand throughout the time course of the assay. Digestion of both strands reached 50% within 10 min of incubation and was 90% complete within an incubation time of 40 min.

From the digestion profile of the TG-substituted molecules, it appears that the single TG substitution altered the ability of *Bam*HI to catalyze the digestion of not only the substituted extended strand but also the unsubstituted template strand. After 10 min of incubation, when digestion of control extended strands reached 50%, cleavage of the strand with a single TG substitution within the recognition locus was only 15% of total; it reached 50% after 60 min. In comparison, the template

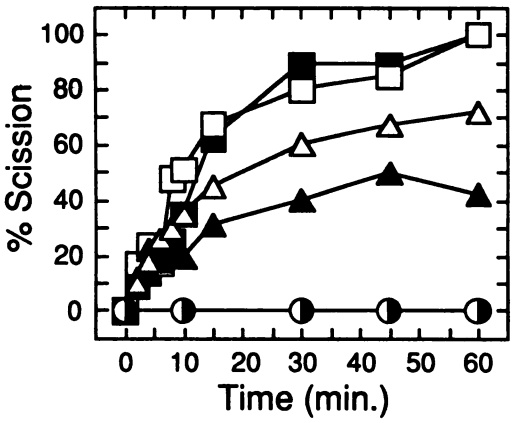

Fig. 5. Effect of TG substitutions on the digestion of oligonucleotides by *Bam*HI. Oligonucleotides with zero (■, □), one (▲, △), or two (●, ○) substitutions of TG for G within one strand of the *Bam*HI recognition site were synthesized as described in Materials and Methods. These substrates (175 nM) were incubated with *Bam*HI endonuclease at 15°, and aliquots were removed at the intervals indicated and analyzed by electrophoresis. The data represent the percentage of starting material converted to expected scission products at each time point, both for the extended strand containing the substitutions (■, ▲, ●) and for the template unsubstituted strand (□, △, ○) of each duplex. The results from one of three comparable experiments are presented.

unsubstituted strand of these substrates was digested more efficiently than the substituted strand. Cleavage of template strands reached 40% after 10 min and 75% after 60 min. The results from Fig. 4C and from Fig. 5 indicate that substitution of both Gs within one strand of the *Bam*HI locus completely abolished digestion of both the substituted and unsubstituted strands.

Assay of *Mbo*I digestion of TG-containing oligonucleotides. Control DNA (2.6–250 nM) was digested with 0.0166 units/μl *Mbo*I, and initial rates of enzyme hydrolysis were determined. Michaelis-Menten analysis of the *Mbo*I-digested DNA revealed a K_m of 12 nM and a V_{max} of 184.4 ± 14 nmol/min/unit of enzyme. Hayakawa *et al.* (15) have reported a K_m of 380 nM for their decanucleotide substrate. TG-substituted DNA was digested with *Mbo*I at substrate concentrations 10-fold above the K_m for the enzyme (100 nM). We found that TG substitution within the *Mbo*I locus did not affect the specificity of the enzyme for cleavage at the phosphodiester bond (data not shown). Reaction products were quantitated, and the time course of digestion is presented in Fig. 6. Again, as in the case of *Bam*HI, TG substitution appeared to impair *Mbo*I digestion of these oligonucleotides. However, the effects of a single substitution versus a double substitution within the recognition locus were not as pronounced with *Mbo*I as with the endonuclease. Single TG-substituted molecules were 50% digested after 25 min, compared with control molecules, which were 50% digested after only 10 min. Digestion of oligonucleotides with two TG substitutions within the recognition locus was also impaired, compared with control. However, in contrast to the results with *Bam*HI, cleavage of these substrates was not completely abolished. Digestion of the DNA reached a level as high as 50% after 40 min of incubation with *Mbo*I.

***Alw*I digestion of control and TG-substituted oligonucleotides.** Because the sites for recognition and scission are coincident for both *Bam*HI and *Mbo*I, it is difficult to determine the effect of TG substitution on either process individually. In order to be able to determine the individual contribution TG

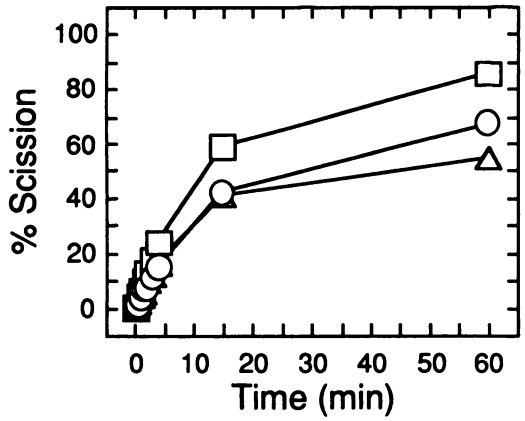

Fig. 6. Effect of TG substitutions on the digestion of oligonucleotides by *Mbo*I. Oligonucleotides with zero (□), one (△), or two (○) substitutions of TG for G within one strand of the *Mbo*I recognition site were synthesized as described in Materials and Methods. These substrates (120 nM) were incubated with *Mbo*I at 15°, and aliquots were removed at the indicated time points and analyzed by electrophoresis. The data shown represent the extent of scission of the extended analog-containing strand. Results from one of three comparable experiments are shown.

substitution may have on phosphodiester bond cleavage and on DNA-protein recognition mechanisms, we constructed synthetic molecules that were substrates for the nonpalindromic enzymes *Alw*I and *Fok*I, with the recognition sequences 5′-GGATC(N_4)↓-3′ and 5′-GGATC(N_9)↓-3′, respectively. The overall scheme is described in Materials and Methods and Fig. 2.

Control and TG-substituted oligonucleotides representing recognition site or scission site modifications were digested with *Alw*I, and the products were quantitated and expressed as a percentage of total substrate in the assay (Fig. 7). When DNA molecules with TG substitutions within the recognition sequence were digested with *Alw*I, there was a marked reduction in the formation of the 31- and 17-base products from the 45-base substrate strand, compared with control. After 60 min, 87% of the control 45-mer substrate strand was cleaved to its respective products. Only 0.6% of the TG-substituted extended strand was digested after 60 min. In contrast, when DNA substrates with TG substituted in the scission sequence were digested with *Alw*I, the amount of strand cleavage was comparable to that of control. The percentage of substrate molecules converted to product was 87% and 88% for control and TG-substituted molecules, respectively. Comparable results were obtained with *Fok*I, which is also a type II-S endonuclease (data not shown). That is, TG substitution within the recognition site completely prevented enzymatic cleavage, whereas the analog had no effect when inserted on both sides of the scission site.

Discussion

Construction of TG-containing oligonucleotides. Although oligonucleotides containing a variety of modified bases have been chemically synthesized, the inclusion of TG moieties by chemical synthesis has been problematic, apparently due to the lability of the thiol group. We know of only one instance in which such a synthesis has been reported (16) and, because the method used was somewhat time consuming and required expertise not available to us, we chose to employ an enzymatic approach toward the production of oligonucleotides containing selective patterns of TG substitution.

The technique is straightforward and relies on the ability of a DNA polymerase to extend the 3′ hydroxyl end of an annealed primer molecule with the deoxyribonucleotide form of TG and the normal complement of dNTPs. We chose to utilize the polymerase activity of modified T7 DNA polymerase (Sequenase), because this enzyme has been demonstrated to have a high rate of nucleotide incorporation as well as a high degree of processivity (17). It is interesting to note that extension by Sequenase was apparently stalled at particular sites in the reactions containing TdGTP. These stalls were pronounced at regions of the template coding for successive Gs (Fig. 3A). Other investigators have found that the arabinosyl nucleosides such as (ara-C)1-β-D-arabinofuranosylcytosine and (ara-A), 1-β-D-arabinofuranosyladenine as well as the 2-halogenated dATP analogs, can affect the elongation of the growing DNA strand (1, 2, 18, 19). In these studies, accumulation of DNA strands at sequences coding for the insertion of contiguous analog molecules was demonstrated. Our results suggest that the substitution of TdGTP for dGTP can also alter the reactivity of the 3′ terminus of the growing DNA strand, such that elongation is impaired.

The ability of Sequenase to carry out the addition of a nucleotide to the 3′ end of the completely extended primer oligonucleotide was an unexpected phenomenon. Clark (12) has reported the non-template-dependent addition of deoxyribonucleotides onto duplex oligonucleotides carried out by DNA polymerases from a variety of prokaryotic and eukaryotic sources. The results of our oligonucleotide extensions suggest that Sequenase may also have this property. Alternatively, it is known that DNA polymerases such as the Klenow fragment of

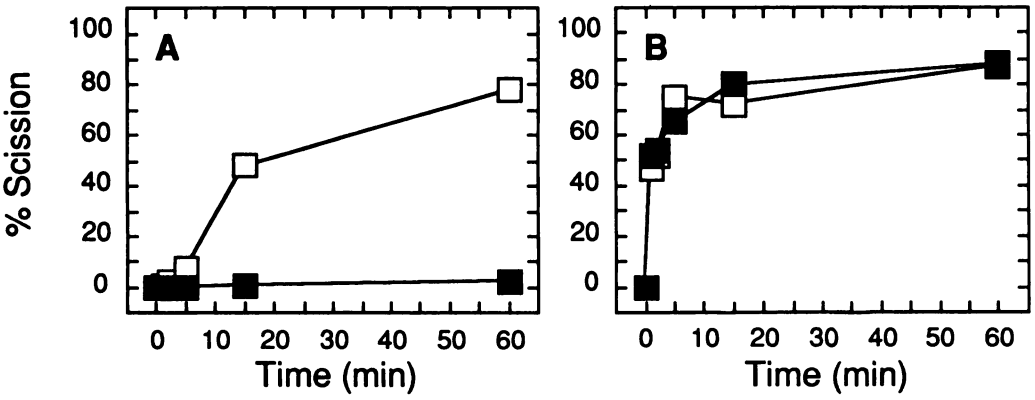

Fig. 7. Effect of TG substitution on scission by *Alw*I. Substrates for *Alw*I were constructed to contain TG substitutions either within the recognition sequence (A) or flanking the scission site (B), as described in Materials and Methods. Control (□) or TG-containing (■) oligonucleotides (80 nM) were digested with *Alw*I at 20° for the indicated times, at which point aliquots were removed and analyzed by electrophoresis. Data shown represent the extent of scission of the substituted strand, from one of three similar experiments.

DNA polymerase I have the capacity to extend a growing strand of DNA via a "snapback" mechanism (20). However, we feel that this alternative explanation is unlikely, because snapback synthesis typically results in the addition of greater than one or two nucleotides to the extended oligonucleotide primer.

Cleavage of TG-substituted oligonucleotides. One way in which incorporated analogs might affect DNA-protein interactions is by altering higher order structure in the neighborhood of the insertion. Previous studies have shown that the strength of the TG/C base pair is considerably weaker than that of the naturally occurring G/C pair, as determined by melting temperatures of homoribopolymers (21) and by calculation of the association constant of TG-containing oligodeoxynucleotides (16). In the latter case, it was reported that the difference between the two species was more than 100-fold. In light of these findings, it seemed possible that disruption of base-pairing by a few consecutive TG substitutions within or adjacent to a recognition site might be severe enough to affect a wide variety of sequence-specific DNA-protein interactions indiscriminately. We also considered the possibility that perturbations in helix structure caused by multiple TG moieties incorporated next to (rather than within) a recognition sequence might have an effect on the interaction of that sequence with proteins, especially because it has been shown that variations in vicinal sequences of normal nucleotides can influence restriction endonuclease cleavage (22).

Neither of these circumstances appeared to apply to the enzyme-substrate systems studied here. In the case of both *Bam*HI and *Mbo*I, a single substitution within the recognition site decreased, but did not obliterate, enzyme activity. Also, the presence of three consecutive TG moieties very near the *Bam*HI site (positions 35–37 of the extended primer strand) had no detectable effect on cleavage by this enzyme. It, therefore, seems unlikely that disruption of local structure by TG, merely as a consequence of weakened base pairing, has a major effect on these enzyme activities. It is possible, however, that poor base pairing may contribute to the stalling of polymerase-catalyzed primer extension and to the non-template-directed nucleotide addition, which were discussed above.

It appears that primary recognition sequence is more critical in determining the effects of TG incorporation on the activity of these two enzymes, as evidenced by the results obtained with the disubstituted substrates. Whereas *Mbo*I cleaved the disubstituted substrate about as well as the monosubstituted one, *Bam*HI activity was completely eliminated by the addition of a second TG substitution at the external G position. Although the external G locus of the *Bam*HI site is not part of the canonical recognition sequence for *Mbo*I, insertion of N^6-methyladenine at this position can completely prevent *Mbo*I-catalyzed cleavage of the substituted strand (23). This is in sharp contrast to the minimal effects of TG insertion at that position (i.e., the relatively minor difference between the mono- and disubstituted species shown in Fig. 6), even though incorporation of either N^6-methyladenine or TG each represents an alteration of a base-pairing site in the major groove. We interpret this result to indicate that the contrasting responses of *Bam*HI and *Mbo*I to TG substitution at both G loci is not a result of indiscriminate acceptance by *Mbo*I of any modified base at the external position, but rather the result of TG incorporation exerting a selective inhibition of *Bam*HI as a result of fine details of the interactions of the enzyme with the oligonucleotides. This finding is also consistent with the proposal that the external G position is one of the most important contact sites for *Bam*HI (13).

Effects of TG substitution at unsubstituted locations. Specific binding of proteins to DNA often involves the association of multimeric protein subunits with symmetrical DNA sequences. This is the case for many restriction enzymes, including *Bam*HI (24), as well as for the binding of T antigen to SV40 replication origin (25). Because the extent of TG incorporation in intact cells is typically no greater than a few percent (6, 8, 26), the occurrence of more than one TG molecule within a given target sequence will probably be infrequent. For this reason, we wanted to determine whether a single TG substitution has the potential to affect not only the subunit with which it interacts directly, but also the other subunit(s).

Although many studies have described the consequences of analog incorporation into restriction enzyme recognition sites, few have examined the effects of asymmetric analog insertion on the opposite (unsubstituted) strand. One group reported that, when methylated derivatives were used, the rate of cleavage of the substituted strand could be either much greater or much less than that of the unsubstituted strand, depending on the enzyme and substrate used (23). In the case of *Bam*HI, we show that the effects of asymmetric TG substitution are grossly similar on both strands. Although this finding may not necessarily apply to other DNA-protein interactions, it at least demonstrates the potential for TG substitutions to exert effects in remote locations of the same binding domain.

Finally, the experiments performed using the type II-S enzymes *Alw*I and *Fok*I address the possibility that some portion of the effects of TG on restriction endonuclease cleavage may be due to direct inhibition of the phosphodiester bond cleavage, rather than to alteration of the recognition site. This prospect seemed remote, because there are no known sequence requirements for such bond cleavage and because many enzymes do not even require the DNA to be double-stranded in order to be cut (27). As expected, we found that TG-induced inhibition of scission by these enzymes was only related to substitution within the recognition site.

Overview. Before the present study, we had demonstrated that complete replacement of G by TG within the SV40 origin eliminated specific T antigen binding. Although that situation was the simplest to test at first, it is not representative of the extent of TG incorporation in intact cells, which typically does not exceed a few percent. Furthermore, because TG derivatives are sufficiently similar to their guanine counterparts to be metabolized by several highly specific enzymes (28), it was not clear that a single TG moiety within a multibase recognition sequence would be likely to produce a significant alteration of the protein binding of that sequence. Using restriction endonuclease-catalyzed strand scission as a model system, we show here that single TG substitutions are, in fact, capable of inhibiting such processes, that two consecutive TG substitutions can completely abolish activity (*Bam*HI and *Alw*I), and that the magnitude of the effect can vary greatly from protein to protein (*Bam*HI versus *Mbo*I).

The question remains of whether the frequency with which TG is incorporated into DNA is consistent with the idea that disruption of replication initiation (or some other process depending on sequence-specific DNA-protein interactions, e.g., transcriptional control) is a critical consequence of this incor-

poration. The number of origins containing a significant number of "hits" by TG is presumably related to their size, their average G content, and the extent of TG incorporation necessary to alter their functionality. Although the probability that any individual origin will be affected by TG incorporation is low, the total population of origins is relatively high (10^4 to 10^5/eukaryotic cell). Assuming 10^4/cell, disruption of 1% of replication origins would lead to 100 potentially harmful events. Although the consequences of disabling a single replication origin are unknown, it is thought that one unrepaired double-strand break can kill a cell. It is, therefore, not inconceivable that an average of 1–2% replacement of G by TG could result in a lethal insult through this mechanism. In order to more directly evaluate the possibility that replication origins are an important target for TG *in vivo*, we intend to follow up the findings presented here by determining the frequency and location of TG substitutions necessary to alter activity in the SV40 *in vitro* replication system (29).

References

1. Townsend, A. J., and Y.-C. Cheng. Sequence-specific effects of ara-5-aza-CTP and ara-CTP on DNA synthesis by purified human DNA polymerase *in vitro:* visualization of chain elongation on a defined template. *Mol. Pharmacol.* **32:**330–339 (1987).
2. Ohno, Y., D. Spriggs, A. Matsukage, T. Ohno, and D. Kufe. Sequence-specific inhibition of DNA strand elongation by incorporation of 9-β-D-arabinofuranosyladenine. *Cancer Res.* **49:**2077–2081 (1989).
3. Mikita, T., and G. P. Beardsley. Functional consequences of the arabinosylcytosine structural lesion in DNA. *Biochemistry* **27:**4698–4705 (1988).
4. Goz, B. The effects of incorporation of 5-halogenated deoxyuridines into the DNA of eukaryotic cells. *Pharmacol. Rev.* **29:**249–272 (1977).
5. Wotring, L. L., and J. L. Roti Roti. Thioguanine-induced S and G2 blocks and their significance to the mechanism of cytotoxicity. *Cancer Res.* **40:**1458–1462 (1980).
6. Tidd, D. M., and A. R. Paterson. A biochemical mechanism for the delayed cytotoxic reaction of 6-mercaptopurine. *Cancer Res.* **34:**738–746 (1974).
7. Jones, P. A. Altering gene expression with 5-azacytidine. *Cell* **40:**485–486 (1985).
8. Maybaum, J., A. N. Bainnson, W. M. Roethel, S. Ajmera, L. M. Iwaniec, D. R. TerBush, and J. J. Kroll. Effects of incorporation of 6-thioguanine into SV40 DNA. *Mol. Pharmacol.* **32:**606–614 (1987).
9. Maybaum, J., and H. G. Mandel. Unilateral chromatid damage: a new basis for 6-thioguanine cytotoxicity. *Cancer Res.* **43:**3852–3856 (1983).
10. Challberg, M. D., and T. J. Kelly. Animal virus DNA replication. *Annu. Rev. Biochem.* **58:**671–717 (1989).
11. Stillman, B. Initiation of eukaryotic DNA replication *in vitro*. *Annu. Rev. Cell. Biol.* **5:**197–245 (1989).
12. Clark, J. M. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16:**9677–9686 (1988).
13. George, J., G. Nardone, and J. G. Chirikjian. Sequence-specific *Bam*HI endonuclease: the proposed role of arginine residues in substrate binding and recognition. *J. Biol. Chem.* **260:**14387–14392 (1985).
14. Brennan, C. A., M. D. Van Cleve, and R. I. Gumport. The effects of base analogue substitutions on the cleavage by the *Eco*RI restriction endonuclease of octadeoxyribonucleotides containing modified *Eco*RI recognition sequences. *J. Biol. Chem.* **261:**7270–7278 (1986).
15. Hayakawa, T., A. Ono, and T. Ueda. Synthesis of decadeoxyribonucleotides containing 5-modified uracils and their interactions with restriction endonucleases *Bgl*II, *Sau*3AI and *Mbo*I (Nucleosides and Nucleotides 82[1]). *Nucleic Acids Res.* **16:**4761–4776 (1988).
16. Rappaport, H. P. The 6-thioguanine/5-methyl-2-pyrimidinone base pair. *Nucleic Acids Res.* **16:**7253–7267 (1988).
17. Tabor, S., and C. C. Richardson. DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84:**4767–4771 (1987).
18. Parker, W. B., A. R. Bapat, J.-X. Shen, A. J. Townsend, and Y.-C. Cheng. Interaction of 2-halogenated dATP analogs (F, Cl, and Br) with human DNA polymerases, DNA primase, and ribonucleotide reductase. *Mol. Pharmacol.* **34:**485–491 (1988).
19. Reid, R., E. C. Mar, E. S. Huang, and M. D. Topal. Insertion and extension of acyclic, dideoxy, and ara nucleotides by herpesviridae, human *alpha* and human *beta* polymerases: a unique inhibition mechanism for 9-(1,3-dihydroxy-2-propoxymethyl)guanine triphosphate. *J. Biol. Chem.* **263:**3898–3904 (1988).
20. Maniatis, T., E. F. Fritsch, and J. Sambrook. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1982).
21. Amarnath, V., and A. D. Broom. Polyribonucleotides containing thiopurines: synthesis and properties of poly(6-thioguanylic acid). *Biochemistry* **15:**4386–4389 (1976).
22. Modrich, P., and R. J. Roberts. Type-II restriction and modification enzymes, in *Nucleases* (S. M. Linn and R. J. Roberts, eds.), Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 109–154 (1982).
23. Ono, A., and T. Ueda. Synthesis of decadeoxyribonucleotides containing *N*6-methyladenine, *N*4-methylcytosine, and 5-methylcytosine: recognition and cleavage by restriction endonucleases (Nucleosides and Nucleotides 74). *Nucleic Acids Res.* **15:**219–232 (1987).
24. Smith, L. A., and J. G. Chirikjian. Purification and characterization of the sequence-specific endonuclease *Bam*HI. *J. Biol. Chem.* **254:**1003–1006 (1979).
25. Mastrangelo, I. A., P. V. Hough, J. S. Wall, M. Dodson, F. B. Dean, and J. Hurwitz. ATP-dependent assembly of double hexamers of SV40 T antigen at the viral origin of DNA replication. *Nature (Lond.)* **338:**658–662 (1989).
26. Christie, N. T., S. Drake, R. E. Meyn, and J. A. Nelson. 6-Thioguanine-induced DNA damage as a determinant of cytotoxicity in cultured Chinese hamster ovary cells. *Cancer Res.* **44:**3665–3671 (1984).
27. Nishigaki, K., Y. Kaneko, H. Wakuda, Y. Husimi, and T. Tanaka. Type II restriction endonucleases cleave single-stranded DNAs in general. *Nucleic Acids Res.* **13:**5747–5760 (1985).
28. Paterson, A. R. P., and D. M. Tidd. 6-Thiopurines. *Handb. Exp. Pharmacol.* 383–403 (1975).
29. Ishimi, Y., A. Claude, P. Bullock, and J. Hurwitz. Complete enzymatic synthesis of DNA containing the SV40 origin of replication. *J. Biol. Chem.* **263:**19723–19733 (1988).

Send reprint requests to: Dr. Jonathan Maybaum, 4302A Upjohn Center, University of Michigan Medical School, Ann Arbor, MI, 48109-0504