A Model for the T-Antigen-Induced Structural Alteration of the SV40 Replication Origin Based upon Experiments with Specific Probes for Bent, Straight, and Unwound DNA†

Frank Xiaoguang Han and Laurence H. Hurley*

Drug Dynamics Institute, College of Pharmacy, The University of Texas at Austin, Austin, Texas 78712-1074

Received February 1, 1996; Revised Manuscript Received April 12, 1996⊗

ABSTRACT: The T-antigen-induced structural changes of the SV40 replication origin were probed with three DNA-reactive antitumor agents: (+)-CC-1065, bizelesin, and pluramycin. (+)-CC-1065 is an N3 adenine minor groove alkylating agent that selectively reacts with AT-rich DNA sequences with a bent conformation; bizelesin also reacts with the minor groove of AT-rich sequences but is selective for a straight DNA conformation. Pluramycin is an intercalative guanine alkylator whose reactivity is increased by unwinding and decreased by compression of the minor and/or major grooves of DNA. We show that while binding of T-antigen reduced the ability of (+)-CC-1065 to alkylate the AT tract in the SV40 replication origin, it did not interfere with bizelesin modification of the same sequence. These unexpected results suggest that when T-antigen binds to the SV40 origin the AT tract is in a straight DNA conformation. High-resolution DNase I footprinting experiments indicate that at least three helically in-phase T-antigen binding sites exist in the GC box region located immediately downstream of the AT tract. The binding of T-antigen enhances the reactivity of (+)-CC-1065 to the two 5′-AGTTA* (the asterisk indicates the covalent bonding site) drug modification sites in the GC box region, demonstrating that these sites are in a bent conformation. In contrast, T-antigen inhibited the reactivity of pluramycin at sequences within the GC box region that are known not to bind T-antigen. These data, in combination with the DNase I footprinting results, suggest that T-antigen binding induces a conformational change in the DNA that no longer favors pluramycin intercalation. Based on our results, we propose that T-antigen binds tightly to the upstream region of the AT tract of SV40 replication origin forming double hexamers. In the downstream region, binding of T-antigen to the helically in-phase sites in the GC box region induces DNA bending in the opposite direction of the natural AT tract bending, while simultaneously transforming the naturally bent AT tract DNA into a straight conformation.

Simian virus 40 (SV40) replication provides an excellent model for understanding eukaryotic DNA replication. The SV40 large tumor antigen, called T-antigen (T-ag), is the only viral protein required for SV40 replication *in vitro* (DeLucia et al., 1986; Li et al., 1987). T-ag has two strong binding sites near the SV40 *ori* (sites I and II) and several weaker sites, collectively termed site III, at the GC box region [Mastrangelo et al., 1985; reviewed by Stillman (1989) and Borowiec et al. (1990)]. To initiate DNA replication *in vitro* or *in vivo*, T-ag binds to the 64-bp core origin, which consists of three functional domains: a 27-bp early palindromic (EP) element, a four pentanucleotide (PEN) (5′-GAGGC) repeat region, and a 17-bp adenine/thymine (AT)-rich region (Gerard & Gluzman, 1986; Deb et al., 1986a,b, 1987; Li et al., 1987) (Figure 1). When T-ag binds to the PEN region, a double hexamer structure is formed (Mastrangelo et al., 1989; Borowiec et al., 1990; Parsons et al., 1991). The interaction, which is ATP-dependent (Borowiec & Hurwitz, 1988b; Palit & Tegtmeyer, 1987), induces the melting of the EP region (Wold et al., 1987; Dodson et al., 1987; Roberts, 1989) and distorts the AT domain, opening the core region (Dean et al., 1987a,b; Borowiec & Hurwitz, 1988a; Borowiec et al., 1991; Dean & Hurwitz, 1991). The intrinsic helicase activity of T-ag (Goetz et al., 1988) then unwinds the DNA duplex further in both directions (Dodson et al., 1987). To date, the precise mechanism by which T-ag deforms the AT region is unclear (Borowiec et al., 1990; Lin & Kowalski, 1994). To further investigate the T-ag-induced distortion of this region, we probed the T-ag−SV40 origin DNA complex with three sequence-specific structural probes, (+)-CC-1065, bizelesin, and pluramycin, that induced structural alteration in DNA previously characterized by high-field NMR.

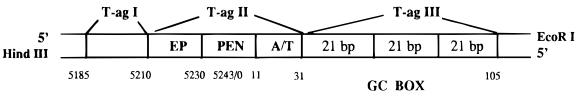

FIGURE 1: 209-bp DNA *Eco*RI−*Hin*dIII fragment from the pCAT control plasmid containing the DNA replication origin of SV40. T-ag binds at three regions: T-ag I, II, and III. T-ag II includes the early palindrome (EP) region, the four pentanucleotide (PEN) region, and the 17-bp adenine/thymine A/T-rich region; T-ag III contains three 21-bp repeats rich in GC.

(+)-CC-1065 (structure shown in Figure 2a), an extremely potent cytotoxic antitumor antibiotic produced by *Streptomyces zelensis* (Hanka et al., 1978; Reynolds et al., 1985), fits into the DNA minor groove and alkylates DNA at the N3 position of adenine (Lin & Hurley, 1990; Scahill et al., 1990). The preferred target sequences for (+)-CC-1065 are 5′-PuNTTA* and 5′-AAAAA* (the asterisk indicates co-

† Research is supported in part by The University of Texas at Austin, the Welch Foundation, and the U.S. Public Health Service (CA-49751).

* Address correspondence to this author.

⊗ Abstract published in *Advance ACS Abstracts,* June 1, 1996.

S0006-2960(96)00251-6 CCC: $12.00

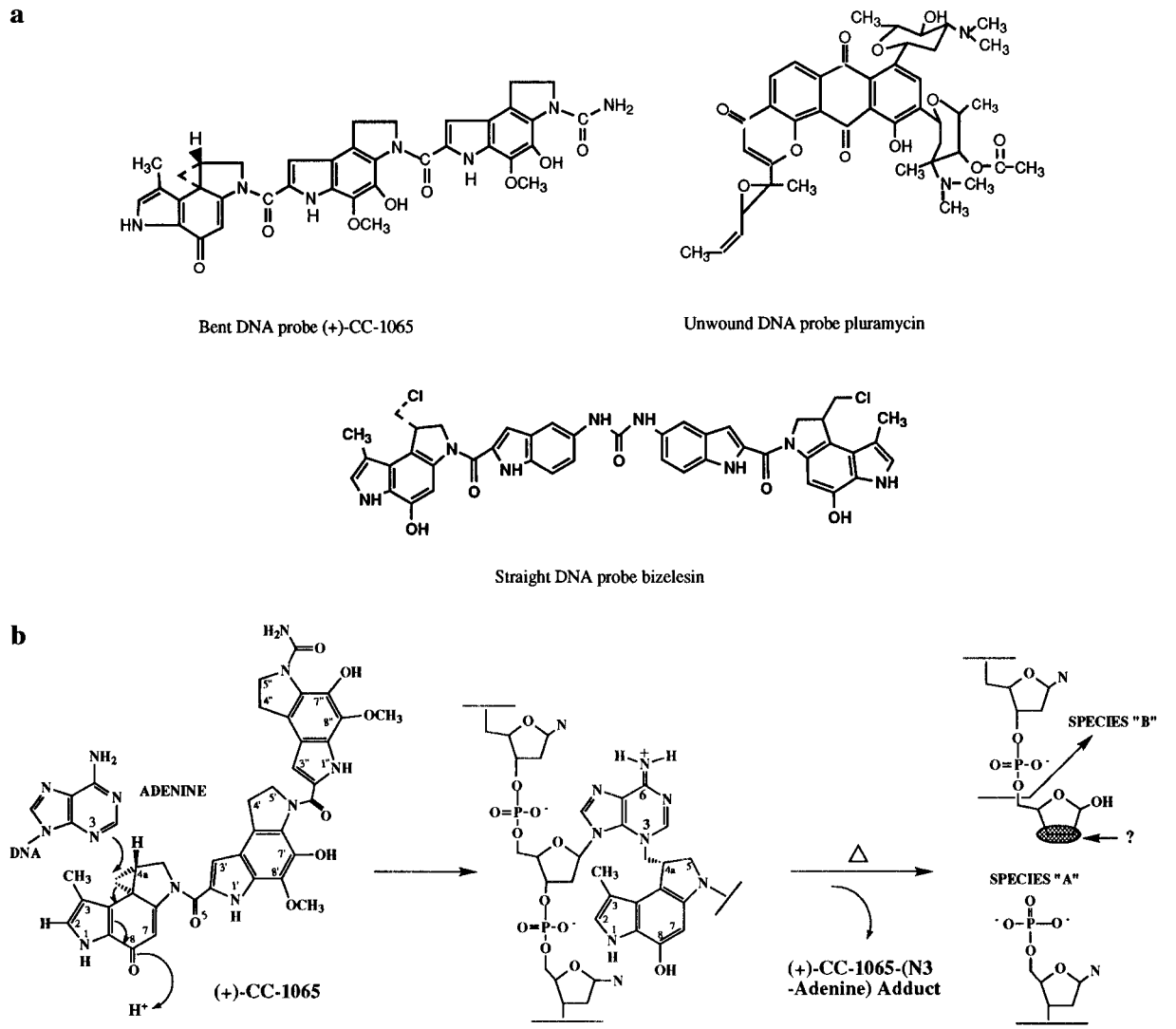

FIGURE 2: Structures of probes used and their mechanism of reaction with DNA represented by (+)-CC-1065. (a) Structures of (+)-CC-1065, pluramycin, and bizelesin. The chloromethyl is converted to the cyclopropyl in aqueous solution. (b) Reaction of (+)-CC-1065 with DNA to form the (+)-CC-1065−(N3 adenine) DNA adduct and products of heat-strand breakage of modified DNA. Species "A" is the first product produced by thermal treatment and is converted into species "B" by further thermal treatment or subsequent treatment with piperidine. For bizelesin, the reaction and strand breakage are analogous (Sun & Hurley, 1993), but for pluramycin, the reaction is with N7 of guanine (Sun et al., 1995b), and the strand breakage is analogous.

valently modified adenine) (Reynolds et al., 1985; Hurley et al., 1990). ^{1}H and ^{31}P NMR studies indicate that (+)-CC-1065 alkylation of DNA produces a bent DNA that resembles A-tracts (Lin et al., 1991, 1992). Recently, (+)-CC-1065 has been successfully used to detect the DNA bending induced by Sp1 (Sun & Hurley, 1994), *Mu* transposase (Ding et al., 1993), and IHF (Sun et al., 1995a). In this paper, we refer to (+)-CC-1065 as an adenine monoalkylator and a bent DNA probe. Bizelesin (structure shown in Figure 2a), a symmetrical dimeric analog of (+)-CC-1065, also alkylates adenine at the N3 position in the minor groove (Ding & Hurley, 1991). Bizelesin forms interstrand cross-links between adenines 6 or 7 bp apart (Ding et al., 1993; Sun & Hurley, 1993) and preferentially modifies 5′-TAATTA*/TAATTA* and the A-tract 5′-TAAAAA*/TTTTTA*, entrapping the intrinsically bent A-tract DNA in its straight form (Thompson & Hurley, 1995; Thompson et al., 1995). Accordingly, in this paper we refer to bizelesin as a minor groove adenine cross-linker and a straight DNA probe. Pluramycin (structure shown in Figure 2a), discovered by scientists at Abbott Laboratories, Chicago, IL, intercalates into DNA via a threading mechanism and alkylates N7 of guanine (Sun et al., 1995b). The drug has been used as a probe to demonstrate the TBP-induced asymmetrical unwinding of the DNA on the downstream side of the TATA box, which is required for initiation of transcription (Sun & Hurley, 1995). In this paper, we refer to pluramycin as a guanine alkylator and DNA unwinding probe.

By using these three structural probes in combination with the DNase I footprinting experiments, we have obtained evidence revealing the structural alterations that T-ag induces in the AT tract and GC box region. These conclusions are based upon the observations that T-ag protected T-ag site I completely from (+)-CC-1065 attack; the AT tract in T-ag site II showed diminished alkylation by (+)-CC-1065, but not by the cross-linker bizelesin; the PEN region and the GC box region (containing the three 21-bp repeat region) were protected from pluramycin alkylation; and there were two enhanced cleavage sites for (+)-CC-1065 within the second and third 21-bp repeat region. This surprising result strongly suggests that T-ag-induced DNA bending occurs

in the GC box region, reminiscent of the DNA bending induced by Sp1 binding to the same region (Sun & Hurley, 1994). Our high-resolution DNase I footprinting results revealed three specific T-ag binding sites within the GC box that are in-phase relative to the DNA helical turns. Based on our results, we propose that T-ag binding induces a looping structure in the GC box region in the opposite direction of the natural bendings in the AT tract. This changes the DNA conformation of the AT tract into a straight form.

MATERIALS AND METHODS

^{32}P End-Labeling of the DNA Fragment Containing the SV40 Replication Origin. The pCAT control plasmid (Promega) contains the SV40 replication origin between restriction sites *Eco*RI (position 396) and *Hin*dIII (position 605). The 209-bp DNA fragment was 5′-end-labeled at either end with [γ-^{32}P]ATP (Amersham) following standard procedures. The labeled fragment was purified on a 6% nondenaturing polyacrylamide gel and eluted with 1 mL of distilled water.

Maxam−Gilbert Sequencing of the 209-bp Fragment. The sequencing reactions were performed using the γ-^{32}P-labeled DNA fragment (ca. 200 000 cpm) with 10 μL of calf thymus DNA (1 μg/μL) in a total volume of 20 μL. For the AG reaction, 30 μL of formic acid was added and reacted for 2 min. For the TC reaction, 20 μL of hydrazine was added for 1 min. Both reactions were stopped with 10 μg of tRNA and ethanol-precipitated. The pellets were lyophilized with a speedvac and boiled in 100 μL of 10% piperidine for 30 min, dried, and then washed 3 times with 100 μL of distilled water. The samples were then dissolved in sequencing dye at 10 000 cpm/μL and preheated 3 min before loading onto a 6% polyacrylamide gel.

DNase I Footprinting. To perform DNase I footprinting experiments, 0.1 ng of γ-^{32}P-labeled DNA fragment was incubated with various amounts of T-ag (Molecular Biology Resources) for 40 min at 37 °C in replication buffer (4 mM ATP, 30 mM HEPES, pH 7.5, 7 mM $MgCl_2$, 40 mM creatine phosphate, 1 mM DTT, 20 μg/mL creatine kinase, 0.1% BSA, and 20% glycerol) in a total reaction volume of 10 μL. After incubation, 2 μL of 0.1 unit/μL of DNase I (Boehringer Mannheim) was added to the reaction solution for 90 s, and the reaction was stopped with 30 μL of DNase I stopping buffer (5 mM EDTA, 0.25 μg/μL calf thymus DNA). The samples were then extracted with phenol solution, precipitated with ethanol, dissolved in sequencing dye, and subjected to electrophoresis in a 6% denaturing polyacrylamide gel. For high-resolution DNase I footprinting, a 60 cm long gel electrophoresis apparatus (Jordan Electrophoresis) was used.

Scanning and Digitization of the DNase I Footprinting Gel. The original gel of the DNase I footprinting on the (−) strand of the SV40 promoter DNA (shown in Figure 3b) was scanned using an Epson scanner ES-1200C. The acquired image was processed using ImageQuaNT software from Molecular Dynamics Inc. Lanes 3, 5, and 10, corresponding to T-ag levels of 0.0, 0.1, and 1.0 μg, respectively, were digitized, and the DNase I footprinting regions of T-ag IV and V are presented in Figure 3c.

Kinetic Assays and Strand Breakage Reaction. A 0.1 ng sample of γ-^{32}P-labeled DNA fragment was incubated with or without 0.1 μg of T-ag in a total reaction volume of 20 μL at 37 °C for 40 min. Two microliters of a 0.2 μM solution of either (+)-CC-1065, bizelesin (gifts from The Upjohn Co.), or pluramycin (gift from Abbott Laboratories) was added and incubated for various periods of time, as indicated in the legend for Figure 4. Reactions were quenched with 10 μg of calf thymus DNA, and distilled water was added to a final volume of 100 μL. For the strand breakage reaction (mechanism shown in Figure 2b), the samples were boiled for 5 min before and after the addition of 10 μL of 10% piperidine and then subjected to phenol extraction and ethanol precipitation. The DNA pellets were washed with 70% ethanol, lyophilized, and dissolved in alkaline sequencing dye solution at 2000 cpm/μL.

RESULTS

DNase I Footprinting Experiments Suggest That T-ag Binding Sites within the GC Box Region Are In-Phase Relative to the DNA Helical Turns. Formation of the DNA−T-ag complex is a prerequisite for the initiation of SV40 DNA replication. DNase I footprinting experiments were performed on the 209-bp *Eco*RI−*Hin*dIII DNA fragment to detect the formation of the DNA−T-ag complex at the SV40 replication origin. Results indicated that with 0.05 μg of T-ag, T-ag site I is completely protected and the PEN region in T-ag site II is partially protected (Figure 3a, lane 4). This result was expected because T-ag site I has a 20−30-fold higher affinity for T-ag than T-ag II (Borowiec & Hurwitz, 1988b). As the level of T-ag is increased to 0.1 μg, the protected region expands to the EP region (Figure 3a, lane 5). Further increases in T-ag concentration lead to a broadened footprint pattern covering the entire DNA fragment (Figure 3a, lanes 6 through 13), consistent with previously published data (Borowiec & Hurwitz, 1988a; Parsons et al., 1991). These data confirm the formation of the SV40 origin−T-ag binding complex. However, the structure of the AT tract upon T-ag binding could not be predicted because the AT tract is insensitive to DNase I cleavage both with and without T-ag (Figure 3a).

High-resolution footprinting was used to determine the T-ag binding sites in the GC box region. In this experiment, the same 209-bp DNA fragment was used except that it was labeled at the 5′-end of the *Eco*RI site and the DNase I-produced products were electrophoresed through a longer gel. The footprinting patterns for T-ag sites I and II are consistent with the previously published results (Borowiec & Hurwitz, 1988a,b) and those shown in Figure 3a. However, downstream of the AT tract region, a protected region in the T-ag III′ site was also detected (Figure 3b, lane 5). The protected sequence extends from A33 to A40, as shown in Figure 3d. Two additional protected regions, in T-ag sites IV and V, were detected within the second and third 21-bp repeats (Figure 3b, lanes 5 through 13). The protected regions correspond to sequences C77 to C80 and C98 to C101, which are separated by 21 bases (2 helical turns). T-ag site IV is 44 bases (4 helical turns) away from the T-ag site III′. These data indicate that the three T-ag binding sites are on the same face of the helical cylinder.

T-ag Binding Induces DNA Conformational Alterations in the GC Box Region and the AT Tract. To gain further insight into the structural changes induced by T-ag binding to the SV40 origin, a time-dependent kinetics assay involving

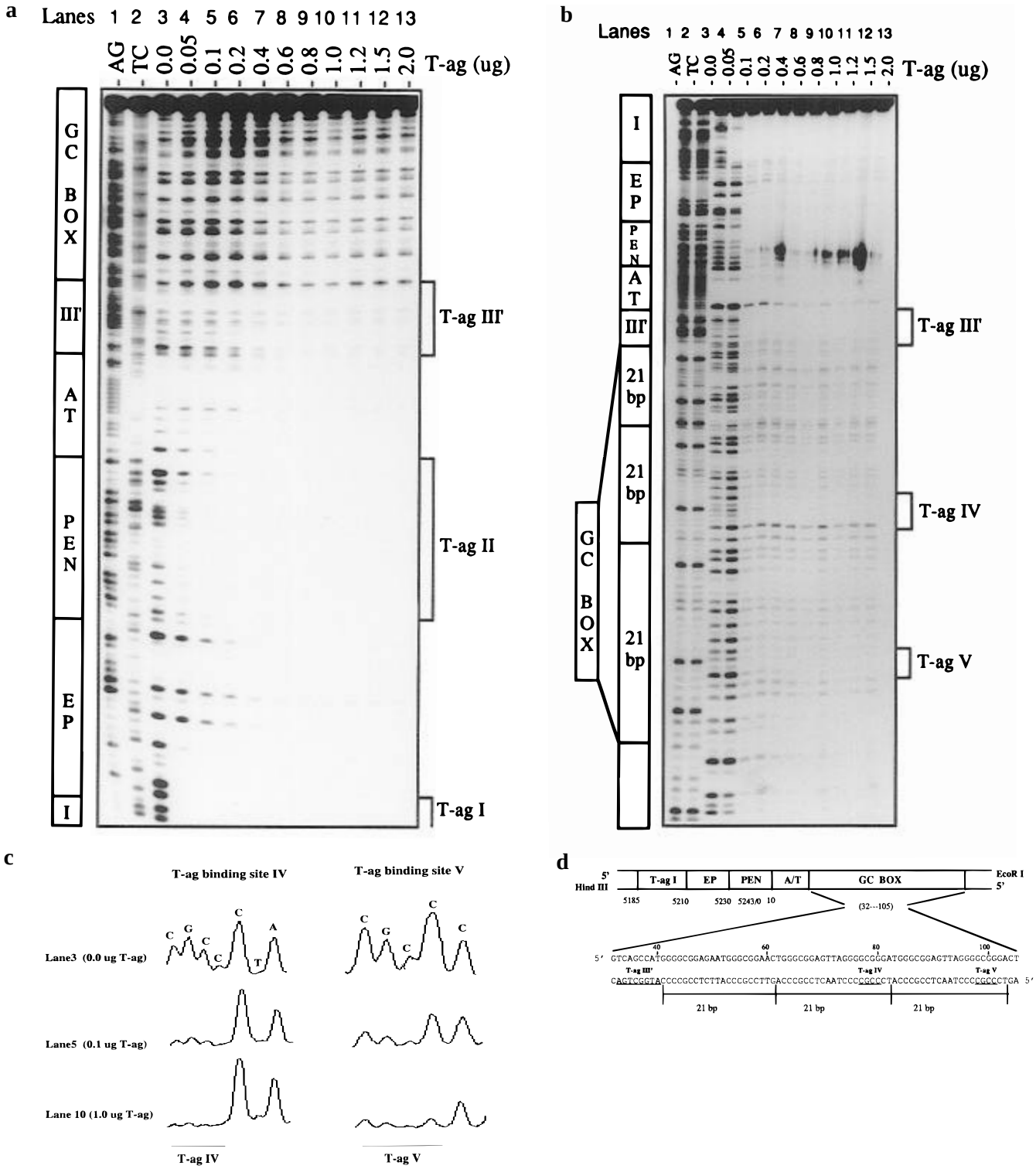

FIGURE 3: DNase I footprinting of the SV40 replication origin. (a) DNase I footprinting of the (+) strand was performed as described under Materials and Methods using the 209-bp *Eco*RI−*Hin*dIII fragment labeled at the 5′-end of the *Hin*dIII overhang. Lanes 1 and 2 contain the Maxam−Gilbert sequencing reactions for AG and TC. Lane 3 is the control without T-ag. In lanes 4−13, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, and 2.0 μg of T-ag were used, respectively, in DNase I footprinting. (b) High-resolution DNase I footprinting of the (−) strand was performed using the same 209-bp DNA fragment but labeled at the 5′-end of the *Eco*RI site. The lane assignments are the same as in (a). The random dark blotches in the PEN region are full-length double-stranded DNA due to insufficient heat denaturing before loading the gel. (c) Scanning representation of the DNase I footprinting on the (−) strand in the T-ag IV and V regions. (d) Schematic illustration of the DNase I footprinting results on the GC box and flanking region. T-ag III′ (protected regions 32−40) is adjacent to the AT tract; T-ag sites IV and V (protected regions 77−80 and 98−101) are within the second and third 21-bp repeat region.

(+)-CC-1065 was performed. As previously mentioned, (+)-CC-1065 is a sequence-specific minor groove bending probe that has a preference for bent DNA. In the experiment, (+)-CC-1065 was added to the DNA or DNA−T-ag complex for increasing periods of time, as indicated in the legend of Figure 4a. A strand breakage assay was performed to determine the site and degree of drug modification of the DNA fragment, which is indicated by a band in the

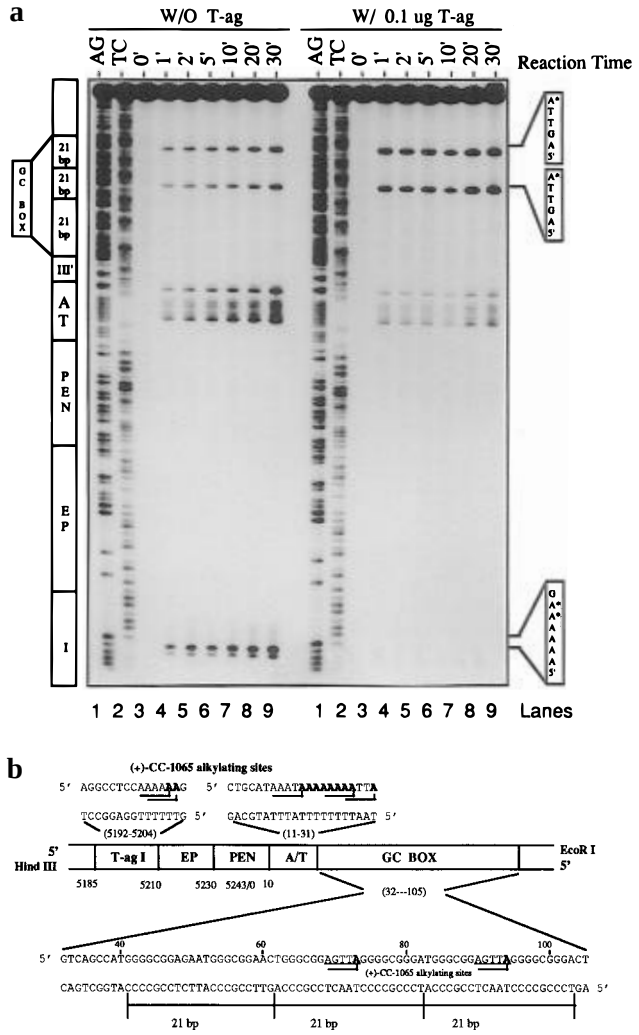

FIGURE 4: Reaction of SV40 replication origin with (+)-CC-1065. (a) DNA alone (left panel) or DNA−protein complex (right panel) formed at 0.1 μg of T-ag was probed with (+)-CC-1065 at a final concentration of 0.2 μM as described. Lanes 1 and 2 contain Maxam−Gilbert sequencing reaction for AG and TC. Lanes 3−9 correspond to the reaction times with drug for 0, 1, 2, 5, 10, 20, and 30 min, respectively. (b) Schematic illustrations of the (+)-CC-1065 alkylation sites in the T-ag I, AT tract, and the GC box region. DNA sequences with boldface letters indicate the targeted adenines, and the solid lines indicate the direction of (+)-CC-1065 orientation in the DNA minor groove.

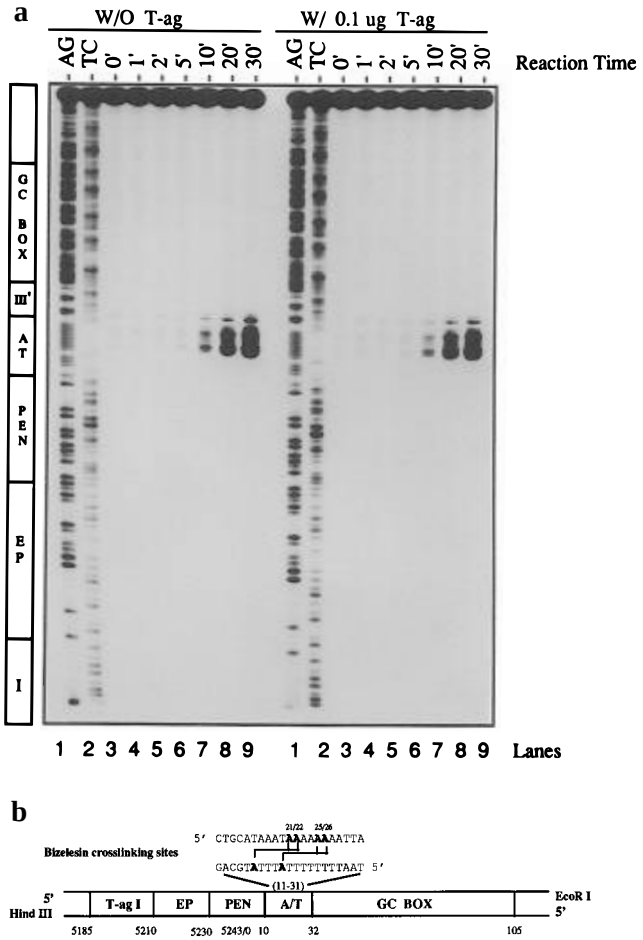

FIGURE 5: Reaction of SV40 replication origin with bizelesin. (a) DNA or DNA−protein complex formed in the presence of 0.1 μg of T-ag was probed with bizelesin as described. Lane assignments are the same as in the legend of Figure 4a. (b) Schematic illustration of the potential bizelesin cross-linking sites at the AT tract. Since bizelesin cross-links adenines six or seven bases apart on different DNA strands, there are only four possible cross-linking combinations within the AT tract sequence. Boldfaced adenines indicate the cross-linking sites. A_{16} on the bottom strand can cross-link with A_{21} (six base) or A_{22} (seven base), and A_{20} can cross-link with A_{25} (six base) or A_{26} (seven base).

autoradiogram where the DNA is cleaved. The results show that in the absence of T-ag (left panel, Figure 4a), (+)-CC-1065 targets T-ag site I, the AT tract of T-ag site II, and the two 5′-AGTTA* sequences in the second and third 21-bp repeat regions (lanes 4−9). In the presence of 0.1 μg of T-ag (right panel, lanes 4−9), no (+)-CC-1065-induced cleavage occurred at T-ag site I (bottom), reduced cleavage occurred at the 17-bp AT tract (middle), and enhanced cleavage occurred at the two 5′-AGTTA* sites (top) in the GC box region that correspond to A73 and A94 in Figure 4b. Because (+)-CC-1065 is the probe that favors bent DNA with a narrowed minor groove, the enhanced cleavage in the GC box region strongly suggests that the DNA minor grooves at the two sites are compressed and the DNA is in a bent conformation. Conversely, the reduced cleavage at the AT tract by (+)-CC-1065 may be due to a T-ag-induced DNA conformational transition, such as straightening of the DNA, which is unfavorable for (+)-CC-1065 modification. Our results, however, cannot exclude the possibility that the reduced cleavage in the AT tract by (+)-CC-1065 is the consequence of steric hindrance by T-ag.

T-ag Binding Does Not Affect Bizelesin Bonding to the AT Tract, Confirming That the AT Tract Is a Straight DNA. To further investigate the reduced cleavage of the AT tract by (+)-CC-1065 in the presence of T-ag, we used bizelesin as the probe. As mentioned earlier, bizelesin targets DNA sequences similar to those of (+)-CC-1065, but preferentially cross-links AT tracts that are in the straight conformation. Figure 5a shows the results of a bizelesin kinetics experiment similar to that performed with (+)-CC-1065. As in the previous experiment, a strand breakage assay was performed to determine the site and degree of drug modification. With (right panel) or without (left panel) 0.1 μg of T-ag, the eight adenines in the AT tract reacted with bizelesin, indicating that the presence of T-ag did not diminish the reactivity of each adenine with bizelesin. In lane 7 of both panels, the cleavage patterns of alternative weaker and stronger double bands could be clearly resolved. The two darker double bands in the eight adenine tracts correspond to A_{21}, A_{22} and A_{25}, A_{26}, whereas the two weaker double bands correspond

to A_{23}, A_{24} and A_{27}, A_{28} (Figure 5b). These patterns are consistent with the cross-linking pattern of bizelesin (shown in Figure 5b), in which adenines 6 or 7 bases apart are cross-linked in the minor groove of DNA. Since bizelesin is a straight DNA probe, these results indicate that in the presence of 0.1 μg of T-ag the DNA minor groove in the AT tract exists as straight DNA and is still accessible to drug modification. In combination with the results from the (+)-CC-1065 kinetic experiment, these data indicate that the reduced alkylation in the AT tract by (+)-CC-1065 is not due to steric occlusion by T-ag, but instead is most likely due to freezing or trapping out of the straight DNA conformation that is only reactive toward bizelesin.

Decreased Pluramycin Reactivity in the GC Box Region Indicates That T-ag Binding Induces DNA Groove Compression. To gain further insight into the enhanced DNA cleavage by (+)-CC-1065 at the two 5′-AGTTA* sites in the GC box region, the T-ag−SV40 origin complex was probed with pluramycin, a DNA intercalator and guanine alkylator (structure shown in Figure 2a). In order to intercalate into DNA, pluramycin must reside in both major and minor grooves with its bulky sugar side chains. An unwound DNA structure favors the intercalation, whereas a compressed minor and/or major groove(s) resulting from DNA bending will inhibit this process.

In the experiment, the DNA from the SV40 origin of replication was probed with pluramycin in the presence and absence of T-ag. As in the previous two experiments, a strand breakage assay was performed to detect the drug modification sites. In the absence of T-ag, strong pluramycin reactivity was observed in the PEN and GC box regions, weak pluramycin reactivity in the EP region, and no reactivity in both the AT tract and T-ag site I (Figure 6a, left panel). This pattern reflects the relative guanine content of respective sequences.

In the presence of T-ag (Figure 6b, right panel), the PEN region was fully protected from pluramycin modification. This result is consistent with the DNase I footprinting data, which indicate that T-ag completely occupies the PEN region at the 0.1 μg level (Figure 3a,b, lanes 5). In the GC box region, the pluramycin reactivity was diminished not only at T-ag binding sites III′, IV, and V but also at sites in this region where T-ag does not bind (top). These results suggest that at the binding sites T-ag can inhibit the pluramycin cleavage reaction in the GC box region by direct steric blocking, and at the nonbinding sites it must inhibit the pluramycin alkylation by indirect conformational alteration of the DNA. These results, in combination with our DNase I footprinting data (Figure 3b), confirm that T-ag not only binds to T-ag sites III′, IV, and V but also causes DNA groove compression in the GC box region.

An interesting observation in the pluramycin experiment is that with T-ag, an enhanced cleavage site by pluramycin was seen at the EP region for the 30-min reaction (Figure 6a, lane 9). This is presumably because T-ag can unwind the EP region (Dodson et al., 1987; Wold et al., 1987; Roberts, 1989), which produces a favorable unwound DNA structure for pluramycin intercalation.

DISCUSSION

Using DNase I footprinting techniques in combination with the DNA sequence-selective covalently bonding drugs (+)-CC-1065, bizelesin, and pluramycin, we have established significant new insights into how T-ag interacts with the SV40 replication origin. Our results indicate that T-ag has three in-phase binding sites in the GC box region, and that T-ag binding to the DNA causes minor groove compression in the GC box region and enhanced (+)-CC-1065 cleavage at the two 5′-AGTTA* sites. Our data also indicate that upon T-ag binding the AT tract is entrapped in its straight DNA conformation. The rationale behind these conclusions is addressed in the following sections, and a model for T-ag binding to the origin of replication is proposed.

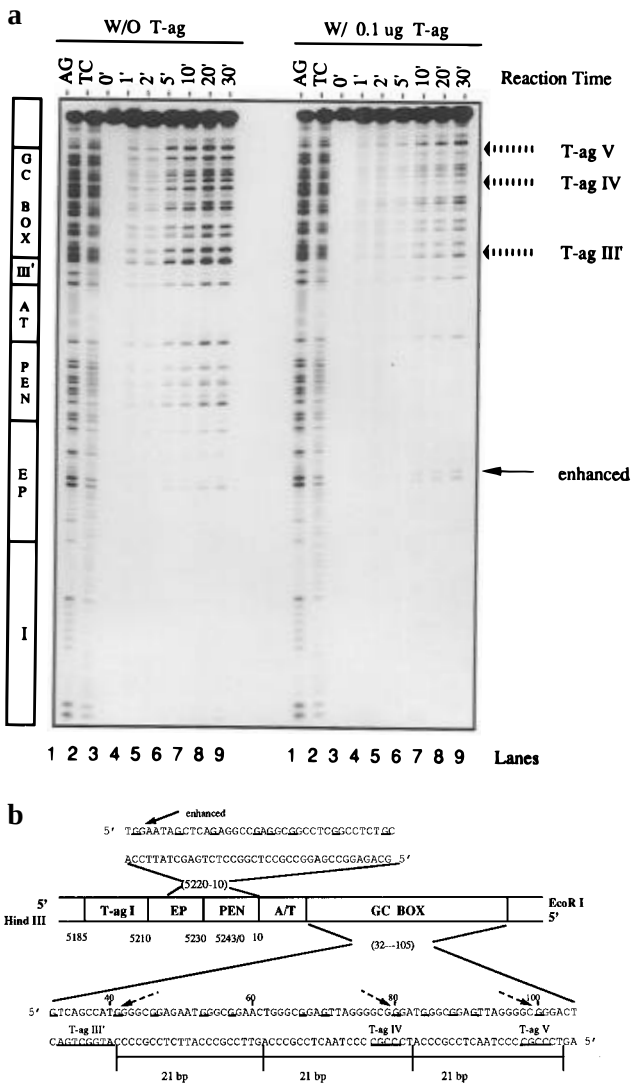

FIGURE 6: Reaction of SV40 replication origin with pluramycin. (a) DNA or DNA−protein complex formed in the presence of 0.1 μg of T-ag was probed with pluramycin at a final concentration of 0.2 μM as described. Lane assignments are the same as in the legend of Figure 4a. In the right panel, the dashed arrows indicate guanines at the T-ag binding sites T-ag III′, IV, and V that have reduced pluramycin reactivity. The solid arrow indicates the guanine that has enhanced reactivity with pluramycin in the presence of T-ag. (b) Schematic representation of the pluramycin alkylating sites in the EP, PEN, and GC box regions. The underlined guanines indicate the pluramycin cleavage sites in the absence of T-ag. The dashed arrows in the GC box region indicate sites that show strong inhibition in the presence of T-ag. In the EP region, the solid arrow at position 5222 indicates the guanine that shows enhanced cleavage by pluramycin in the presence of T-ag.

The Diminished Reactivity of (+)-CC-1065 at the AT Tract upon T-ag Binding Is a Result of a Conformational Change Rather Than Steric Occlusions. In the presence of 0.1 μg

of T-ag (Figure 4a), the reactivity of the AT tract to (+)-CC-1065 is dramatically reduced. There are at least three possible explanations for the observed reduction of the (+)-CC-1065 reactivity in this region. (1) T-ag binding renders the minor groove of AT tract DNA inaccessible to (+)-CC-1065 due to steric hindrance; (2) T-ag unwinds the AT tract, changing the DNA conformation so that it is unsuitable for reaction with (+)-CC-1065, which requires a compressed minor groove characteristic of bent DNA; and (3) T-ag does not bind directly to the AT tract but traps it in its straight conformation by binding to the flanking regions (PEN, III′, IV, and V). This interaction would prevent the AT tract from adopting the bent DNA structure required for (+)-CC-1065 alkylation. Experiments with bizelesin show that the reduced reactivity of (+)-CC-1065 at the AT tract was not due to steric occlusion, because the reactivity of bizelesin with this region was not diminished by preincubation with T-ag. Furthermore, the results of the bizelesin experiment suggest that the AT tract was not unwound, as the 6- or 7-base cross-linking patterns of bizelesin remain essentially identical in both the presence and absence of T-ag. If unwinding of the region had occurred, the distance between the adenines on the opposite strands would change and no bizelesin cross-linking could take place. Experiments from our laboratory recently demonstrated that bizelesin entraps the *straight* conformation of an A-tract (Thompson & Hurley, 1995; Thompson et al., 1995). In light of this evidence, we propose that the AT tract in the presence of T-ag adopts the straight DNA conformation, which is also consistent with the inhibition of (+)-CC-1065 reactivity at this site and with the unchanged reactivity of the region to bizelesin.

The Enhanced (+)-CC-1065 Cleavage at the Two 5′-AGTTA Sites and the In-Phase Binding of T-ag to the GC Box Region Suggest That a DNA Looping Structure Is Induced*. The DNase I footprinting results indicate that T-ag has only three binding sites, III′, IV, and V, in the GC box region, and these sites are in-phase with respect to one another (Figure 3d). Our pluramycin experiments show that upon T-ag binding, DNA cleavage by pluramycin is strongly reduced in the entire GC box region both at the T-ag binding sites III′, IV, and V and at the non-T-ag binding sites. Since T-ag does not bind to this latter region, this suggests that the pluramycin inhibition is not due to steric accessibility. Thus, most likely, when T-ag binds to sites III′, IV, and V, this non-T-ag binding region undergoes a conformational change, such as DNA bending, and this is responsible for the inhibition of pluramycin modification of the region. Based on these observations, and since in the presence of ATP as many as six T-ag molecules can aggregate even in the absence of DNA (Mastrangelo et al., 1989), we propose that T-ag induces DNA bending or looping through protein oligomerization at the in-phase DNA binding sites III′, IV, and V. The data presented here provide some support for this proposal. If T-ag induces DNA looping by in-phase oligomerization, causing groove compression at the inward side of the bent DNA, pluramycin would not be able to traverse the major and minor grooves and alkylate the GC box region. The results of the (+)-CC-1065 kinetic experiments further support our conclusion that T-ag induces DNA bending and looping in this region. The data from these experiments (Figure 4a,b) indicate that the two 5′-AGTTA* sequences show strongly *increased* (+)-CC-1065 reactivity in the presence of T-ag. Since the DNase I footprinting assays on the GC box region (Figure 3a,b,d) showed that T-ag does not contact the two sequences at the low concentration (0.1 μg), this result suggests that the enhanced reactivity of (+)-CC-1065 must be achieved indirectly from T-ag binding at regions that are adjacent to the two 5′-AGTTA* sites (T-ag III′, IV, and V sites), possibly by producing minor groove compression as a consequence of DNA bending in the GC box region. A careful examination of the DNA sequence of this region shows that the two 5′-AGTTA* sites are two helical turns (21 bp) apart (Figure 3d), indicating that the sites of T-ag induced bending are in-phase, which would permit the formation of an overall looping structure.

The Proposed Model for T-ag-Mediated Structural Changes in the SV40 Replication Origin. In light of our results, we propose a model for how T-ag interacts with its target SV40 DNA origin region (Figure 7). In the absence of T-ag, the AT tract of the SV40 origin adopts a naturally bent DNA structure, typical of A-tract DNA. Upon T-ag binding during DNA replication, site II is occupied by T-ag double hexamers. Within the GC box region, T-ag binds at sites III′, IV, and V. Due to the flexibility of DNA sequences in the GC box, T-ag molecules that bind on the same side of the DNA helices are brought close to one another and aggregate into a trimer structure that bends the flexible DNA into a looping structure. The two 5′-AGTTA* sites are located at sites on the inside of the loop where the DNA helix is bent toward the minor groove. This DNA minor groove compression facilitates the formation of the transition state for the covalent reaction of (+)-CC-1065 with the 3′-adenine base within the sequence 5′-AGTTA*. The AT tract, which is located between the upstream T-ag double hexamer and the downstream DNA looping structure, is entrapped in its straight DNA structure by T-ag interactions on either side of the AT tract.

Agreement of the Proposed Model with Existing Biochemical and Genetic Observations. For over two decades, SV40 replication has been extensively investigated, with a large body of work centering on the domain structure of the SV40 replication origin and on the T-ag mediated DNA replication initiation. Much of the evidence is in accord with the model (Figure 7) we have proposed. A study on the interactions between T-ag and the SV40 origin (Borowiec et al., 1991) using chemical probes dimethyl sulfate (DMS) and potassium permanganate ($KMnO_4$) demonstrated that in the presence of ATP, T-ag induces structural changes in the EP region and the AT tract. In the presence of T-ag, the cytosine residues in the EP domains are methylated by DMS, indicating that the DNA is melted in that region. This evidence is in agreement with our conclusion that in the presence of T-ag, increased pluramycin reactivity occurs in the EP region due to unwinding of the DNA (Figure 6a,b). Borowiec and Hurwitz also determined that the AT tract was not heavily methylated by DMS, even though it was extensively modified by $KMnO_4$, indicating that no significant melting occurred in this region; hence, in the presence of T-ag, the AT tract undergoes a structural distortion but remains double-stranded (Borowiec & Hurwitz, 1988b). In a kinetics experiment using $KMnO_4$ as the probe, Borowiec and co-workers (Borowiec et al., 1991) showed that the earliest structural transition induced by T-ag started at the distal portion of the AT tract, away from the central PEN region. In agreement with these results, our (+)-CC-1065

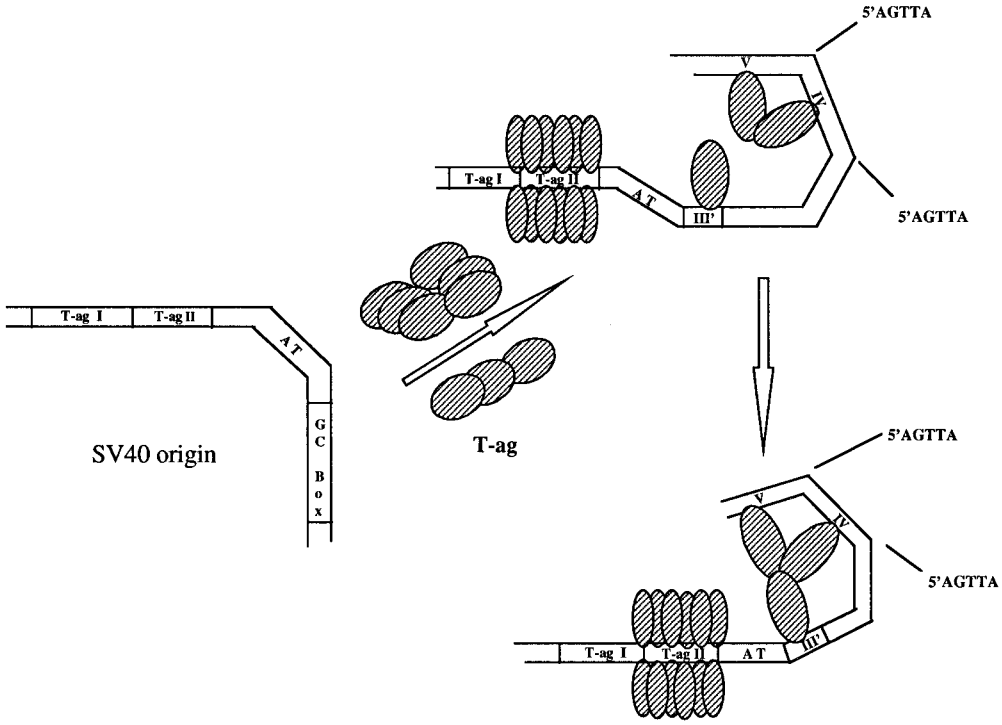

FIGURE 7: Proposed model for T-ag-induced DNA structural changes. The AT tract at the SV40 replication origin is naturally bent in the absence of T-ag. Upon T-ag binding, T-ag sites I and II are occupied (only binding at site II is shown). T-ag also binds weakly to the helically in-phase T-ag III′, IV, and V sites, leading to T-ag oligomerization that induces DNA looping at the GC box region in the opposite direction of the natural AT tract bending. The formation of the DNA looping causes the AT tract to adopt a straight conformation. The two 5′-AGTTA* sites located in the kink regions are indicated.

kinetics experiment data confirmed that the AT tract was deformed by T-ag, and the bizelesin data indicated that the AT tract region was still in the double-stranded DNA form, but was trapped as the straight conformation rather than the bent form. An important observation by Parsons and coworkers (Parsons et al., 1990) that the AT tract is not changed from its bent structure in the absence of the central GAGGC element supports our contention that trapping of the AT tract in its straight conformation requires both the upstream double hexamer structure and the downstream DNA looping structure. Relevant to T-ag DNA looping in the GC box region, one of us has previously used (+)-CC-1065 as a structural probe to show that the GC box region is very flexible in structure and can be trapped as a looping structure by in-phase binding of six Sp1 molecules to this region (Sun & Hurley, 1994). It is conceivable, therefore, that T-ag could also bend this highly flexible region in a similar manner.

Our model is also supported by previously reported genetic evidence. Experiments involving site-directed mutation of the AT region indicate that numerous single-base substitutions are detrimental to SV40 DNA replication (Gerard & Gluzman, 1986; Deb et al., 1987) and that insertions larger than a single base into the AT tract severely inhibit DNA replication (Parsons & Tegtmeyer, 1992). In addition, by studying the replication-competent phenotypic revertants of the replication-defective SV40 virus generated from the AT tract deletion mutation, Gerard and Gulzman (1986) demonstrated that the proper alignment of the 21-bp repeats with the shortened AT stretch within the core region was crucial for DNA replication. Furthermore, work by Delucia and coworkers (Delucia et al., 1986) indicates that deletion of the 21-bp repeat region reduced replication to 50−70% of wild-type levels. These results illustrate that the AT tract in the SV40 replication origin is precisely engineered in terms of DNA length, sequence, and distance between the GC box region. These critical features of the SV40 replication origin are also considered in our model. The correct sequence and length of the AT tract may ensure a flexible conformational transition from bent to straight DNA; the proper alignment of the AT tract with the 21-bp repeat region ensures that the looping in the GC box region is in the opposite direction of the AT tract bending and both the upstream PEN region and the downstream 21-bp repeat region are needed to distort the AT tract. Gerard and Gulzman (1986) proposed that a direct interaction occurs either between the DNA sequences of the 21-bp repeats and the AT-rich region or between protein factors that recognize these sequences. Likewise, our model suggests that a protein-mediated interaction occurs between the 21-bp repeats and the AT tract.

In summary, using results from (+)-CC-1065, bizelesin, and pluramycin experiments as well as the DNase I footprinting data, we propose a model to explain previous observations concerning SV40 replication. This model exhibits similarities to the proposed *E. coli* DNA replication model in that a repeated DNA sequence that is downstream of the AT-rich region of replication origin wraps around the multimeric complex of replication initiator proteins (Bramhill & Kornberg, 1988). However, our model is more focused on the structural changes of the AT tract induced by the replication initiation protein. We believe that T-ag binding to the PEN region and the GC box is the reason for the straightening of the AT tract and therefore might facilitate the unwinding of the AT tract. Our model indicates that the GC box region plays an important role in SV40 replication, as has been proved in many previous studies, even though it is not essential for SV40 DNA replication [Dean et al., 1987b; reviewed by Stillman (1989) and Borowiec et al. (1990)]. This model helps to explain the

dramatic decrease of DNA replication efficiency in the GC box deletion mutants of SV40 origin containing plasmid. How the straightened AT facilitates the unwinding and whether it is a common feature in DNA replication of different systems remain to be determined. The final point is that this series of drug modification experiments unveils an interesting aspect of (+)-CC-1065, bizelesin, and pluramycin. The three compounds are not only potent antitumor agents but also excellent DNA structural probes that are highly selective in DNA sequence and conformation. Used properly, these structural probes can be applied to other DNA−protein interaction systems.

ACKNOWLEDGMENT

We thank Yan Kwok for technical assistance, Doug Henderson for critical reading of the manuscript, and David Bishop for proofreading and editing.

REFERENCES

Borowiec, J. A., & Hurwitz, J. (1988a) *EMBO J. 7*, 3149−3158.
Borowiec, J. A., & Hurwitz, J. (1988b) *Proc. Natl. Acad. Sci. U.S.A. 85*, 64−68.
Borowiec, J. A., Dean, F. B., Bullock, P. A., & Hurwitz, J. (1990) *Cell 60*, 181−184.
Borowiec, J. A., Dean, F. B., & Hurwitz, J. (1991) *J. Virol. 65*, 1228−1235.
Bramhill, C., & Kornberg, A. (1988) *Cell 52*, 743−747.
Dean, F. B., & Hurwitz, J. (1991) *J. Biol. Chem. 266*, 5062−5071.
Dean, F. B., Bullock, P., Murakami, Y., Wobbe, C. R., Weissbach, L., & Hurwitz, J. (1987a) *Proc. Natl. Acad. Sci. U.S.A. 84*, 16−20.
Dean, F. B., Borowiec, J. A., Ishimi, Y., Deb, S., Tegtmeyer, P., & Hurwitz, J. (1987b) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8267−8271.
Deb, S., Delucia, A. L., Baur, C.-P., Koff, A., & Tegtmeyer, P. (1986a) *Mol. Cell. Biol. 6*, 1663−1670.
Deb, S., DeLucia, A. L., Koff, A., Tsui, S., & Tegtmeyer, P. (1986b) *Mol. Cell. Biol. 6*, 4578−4584.
Deb, S., Tsui, S., Koff, A., DeLucia, A. L., Parsons, R., & Tegtmeyer, P. (1987) *J. Virol. 61*, 2143−2149.
DeLucia, A. L., Deb, S., Partin, K., & Tegtmeyer, P. (1986) *J. Virol. 57*, 138−144.
Ding, Z.-M., & Hurley, L. H. (1991) *Anti-Cancer Drug Des. 6*, 427−452.
Ding, Z.-M., Harshey, R. M., & Hurley, L. H. (1993) *Nucleic Acids Res. 21*, 4281−4287.
Dodson, M., Dean, F. B., Bullock, P., Harrison, E., & Hurwitz, J. (1987) *Science 238*, 964−967.
Gerard, R., & Gluzman, Y. (1986) *Mol. Cell. Biol. 6*, 4570−4577.
Goetz, G. S., Dean, F. B., Hurwitz, J., & Matson, S. W. (1988) *J. Biol. Chem. 263*, 383−392.
Hanka, J. J., Dietz, A., Gerpheide, S. A., Kuentzil, S. L., & Martin, D. G. (1978) *J. Antibiot. 31*, 1211−1217.
Hurley, L. H., Warpehoski, M. A., Lee, C.-S., McGovren, J. P., Scahill, T. A., Kelly, K. C., Wicnienski, N. A., Gebhard, I., & Bradford, V. S. (1990) *J. Am. Chem. Soc. 112*, 4633−4649.
Li, J. J., Peden, K. W., Dixon, R. A. F., & Kelly, T. (1987) *Mol. Cell. Biol. 6*, 1117−1128.
Lin, C. H., & Hurley, L. H. (1990) *Biochemistry 29*, 9503−9507.
Lin, C. H., Sun, D., & Hurley, L. H. (1991) *Chem. Res. Toxicol. 4*, 21−26.
Lin, C. H., Hill, G. C., & Hurley, L. H. (1992) *Chem. Res. Toxicol. 5*, 167−182.
Lin, S., & Kowalski, D. (1994) *J. Mol. Biol. 235*, 496−507.
Mastrangelo, I. A., Hough, P. V. C., Wilson, V. G., Wall, J. S., Hainfeld, J. F., & Tegtmeyer, P. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3626−3630.
Mastrangelo, I. A., Hough, P. V. C., Wall, J. S., Dodson, M., Dean, F. B., & Hurwitz, J. (1989) *Nature 338*, 658−662.
Palit, S., & Tegtmeyer, P. (1987) *J. Virol. 61*, 3649−3654.
Parsons. R., & Tegtmeyer, P. (1992) *J. Virol. 66*, 1933−1942.
Parsons, R., Anderson, M. E., & Tegtmeyer, P. (1990) *J. Virol. 64*, 509−518.
Parsons, R. E., Stenger, J. E., Ray, S., Welker, R., Anderson, M. E., & Tegtmeyer, P. (1991) *J. Virol. 65*, 2798−2806.
Reynolds, V. L., Molineux, I. J., Kaplan, D. J., Swenson, D. H., & Hurley, L. H. (1985) *Biochemistry 24*, 6228−6237.
Roberts, J. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 3939−3943.
Scahill, T. A., Jensen, R. M., Swenson, D. H., Hatzenbuhler, N. T., Petzold, G. L., Wierenga, W., & Brahme, N. D. (1990) *Biochemistry 29*, 2852−2860.
Stillman, B. (1989) *Annu. Rev. Cell. Biol. 5*, 197−245.
Sun, D., & Hurley, L. H. (1993) *J. Am. Chem. Soc. 6*, 889−894.
Sun, D., & Hurley, L. H. (1994) *Biochemistry 33*, 9578−9587.
Sun, D., & Hurley, L. H. (1995) *Chem. Biol. 2*, 457−469.
Sun, D., Harshey, R. M., & Hurley, L. H. (1995a) *Biochemistry* (submitted for publication).
Sun, D., Hansen, M., & Hurley, L. H. (1995b) *J. Am. Chem. Soc. 115*, 5925−5933.
Thompson, A. S., & Hurley, L. H. (1995) *J. Mol. Biol. 252*, 86−101.
Thompson, A. S., Sun, D., & Hurley, L. H. (1995) *J. Am. Chem. Soc. 117*, 2371−2372.
Wold, M. S., Li, J. J., & Kelly, T. J. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3643−3647.

BI960251D