Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 1764–1768, April 1987
Biochemistry

Evidence for altered DNA conformations in the simian virus 40 genome: Site-specific DNA cleavage by the chiral complex Λ-tris(4,7-diphenyl-1,10-phenanthroline)cobalt(III)

(DNA secondary structure/Z-DNA/photoactivated DNA probe/promoter/enhancer)

BARBARA C. MÜLLER, ADRIENNE L. RAPHAEL, AND JACQUELINE K. BARTON*

Department of Chemistry, Columbia University, New York, NY 10027

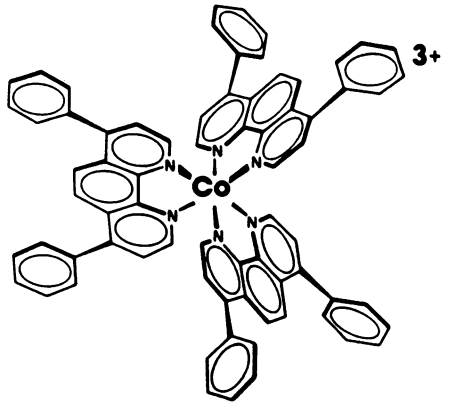
Communicated by Nicholas J. Turro, November 10, 1986 (received for review September 17, 1986)

ABSTRACT **Λ-Tris(4,7-diphenyl-1,10-phenanthroline)cobalt(III), a photoactivated DNA-cleaving agent, is a small molecular probe of DNA structure. Because of its chirality, the complex cannot bind to regular right-handed B-form DNA but exhibits site-specific cleavage along the polymer strand at conformationally distinct sites such as those in a left-handed conformation. Both coarse and higher resolution mapping experiments using the chiral cobalt complex indicate intriguing conformational variations along the simian virus 40 genome. Highly specific cleavage is evident in the enhancer and promoter blocks and in the region downstream of 3′ termini. A specific cleavage pattern borders an alternating purine/pyrimidine stretch within the enhancer, which was found earlier to bind anti-Z-DNA antibodies. Throughout the simian virus 40 genome, variations in structure delineated with the cobalt complex appear to correlate with regions important for control of gene expression.**

Conformational heterogeneity along the DNA strand is becoming increasingly apparent. Aside from the common B-DNA conformation, DNA under certain conditions may adopt other shapes, such as single-stranded loops, hairpins, cruciforms, or even the left-handed double-helical Z-form (1–7). Variations in DNA secondary structure along the strand might play a chemical role in the regulation of gene expression. The role that Z-DNA might play in the cell is still a matter of much debate (1, 8, 9). Physiologically reasonable conditions are known that stabilize and maintain the conversion from right-handed B-DNA to a left-handed Z form (10–12). Stretches of alternating purine/pyrimidine residues favor the formation of Z-DNA, and small regions of such alternation have been found to be associated with biologically interesting positions along both eukaryotic and prokaryotic genomes (13–16). Using a chemical probe that cleaves left-handed DNA sites, we found (17) an association of cleavage sites with the ends of coding regions in the bacterial plasmid pBR322. In the case of simian virus 40 (SV40) DNA, antibodies elicited against Z-form poly(dG·dC) were found to bind in particular within the regulatory region of the viral DNA near the origin of replication (16).

The detection of local variations in structure along the polymer strand requires the development and characterization of enzymatic and chemical probes that are conformation- or site-selective. Mung bean, micrococcal, and S1 nucleases appear to react preferentially at several different altered conformations (18–22). Bis(phenanthroline)copper and methidium-propyl-EDTA·iron(II) have been useful in identifying open, accessible regions along the strand (23, 24). Treatment of DNA with reagents (such as diethylpyrocarbonate, osmium tetroxide, or hydroxylamine) showing enhanced reactivity at accessible bases or sugars, when coupled to sequencing methods, has been useful in marking regions of altered conformations (25–28). Although the chemical mechanism of hyperreactivity with these reagents is not known, sequencing experiments, in detecting variations in conformation near alternating purine/pyrimidine inserts in plasmids (29) and in regions of alternation in pBR322, support findings that left-handed Z-form segments may occur along the strand.

We report here the results of coarse and fine mapping experiments on SV40 DNA, using the chiral metal complex Λ-tris(4,7-diphenyl-1,10-phenanthroline)cobalt(III) [Λ-Co(DiP)$_3^{3+}$], a specific molecular probe for helical conformations. Cleavage experiments with SV40 DNA as substrate were conducted to define and correlate the potential regions of altered conformation with sites important to genetic expression in this well-investigated eukaryotic model (30–33). Tris(phenanthroline) complexes of cobalt(III) may be used to introduce single-strand breaks in the duplex upon photoactivation (34). The chiral metal complexes become useful probes for DNA helicity, since steric constraints prevent intercalation of the Λ isomer of diphenylphenanthroline complexes into right-handed B-form DNA (35–38). Intercalative binding of ruthenium(II) analogues to A-form helices or cleavage using the cobalt complex of single-stranded DNA are not found (17, 37). Instead, selective binding to and subsequent specific cleavage of left-handed Z-DNA by Λ-Co(DiP)$_3^{3+}$ has been demonstrated (17, 39). Strictly speaking, any stacked conformation that locally untwists or significantly unwinds the duplex, thereby increasing the dimensions of the groove, should yield specific cleavage (unwinding to the point of a left-handed twist being an extreme). The structure of Λ-Co(DiP)$_3^{3+}$, our site-specific conformational probe, is given below.



The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DiP, 4,7-diphenyl-1,10-phenanthroline; SV40, simian virus 40.

*To whom reprint requests should be addressed.

MATERIALS AND METHODS

Materials. Λ-Co(DiP)$_3$[(−)-tartrate]$_3$ was prepared as described (34). SV40 DNA and enzymes were purchased from Bethesda Research Laboratories. ^{32}P-labeled nucleotides were from ICN.

Methods. *Photocleavage of DNA with Λ-Co(DiP)$_3^{3+}$.* SV40 DNA (33 μM nucleotides) at native superhelical density (σ = −0.04) was incubated with 10 μM Λ-Co(DiP)$_3^{3+}$ in 20 mM Tris acetate, pH 7/18 mM NaCl. Irradiation was conducted at 315 ± 10 nm using a 1000-W Hg/Xe lamp (Oriel, Stamford, CT) for 8–20 min, which yields partial-to-complete conversion to nicked form II DNA, which was then precipitated with ethanol.

Coarse Mapping. A procedure analogous to that described earlier (17) was employed. After photolysis, SV40 DNA was linearized using either *Hpa* II, *Acc* I, *Eco*RI, or *Bam*HI. The linear DNA was then treated with nuclease S1 to cut opposite the nicked sites to yield small double-stranded fragments. This mixture of fragments was resolved by electrophoresis in a 1% agarose gel, and from their lengths the discrete sites of Λ-Co(DiP)$_3^{3+}$ cleavage were determined.

Fine Mapping. After photolysis, ^{32}P-end-labeled DNA fragments were prepared. The 3′ termini were labeled using [α-^{32}P]dNTPs (specific activity, 3000 Ci/mmol; 1 Ci = 37 GBq) and the large fragment of DNA polymerase I. The 5′ termini were labeled using [γ-^{32}P]dATP (specific activity, 4500 Ci/mmol) and bacteriophage T4 polynucleotide kinase. For analysis of the regulatory region, a 454-base-pair (bp) fragment spanning the origin of replication of SV40 was obtained by restriction with *Hin*fI and *Hpa* II [positions 5136–346 (33)]. A *Bam*HI–*Hin*fI fragment containing the 3′ ends of the structural genes (288 bp; positions 2533–2824) was isolated to determine the sites of photolysis by Λ-Co(DiP)$_3^{3+}$. The gel-purified fragments were then electrophoresed in denaturing (8 M urea) 6% polyacrylamide gels. To determine the position of the nicks generated by Λ-Co(DiP)$_3^{3+}$, untreated DNA was sequenced by the method of Maxam and Gilbert (40) and coelectrophoresed in adjacent lanes. For autoradiography, gels were exposed to Kodak XAR-5 films at −60°C for 1–14 days, using Cronex Lightning*Plus* intensifying screens.

RESULTS AND DISCUSSION

Coarse Map of Cleavage Sites Throughout SV40 DNA. We first localized the sites cleaved by Λ-Co(DiP)$_3^{3+}$ by a general examination of the full SV40 genome, using the coarse mapping assay. A schematic view of Λ-Co(DiP)$_3^{3+}$ recognition sites in the complete SV40 circular DNA is shown in Fig. 1. The relative band intensities vary in order of sites 4570 > 2600, 1611 > 5187, 4924, 900, 436 > 3309, 102 (broad band with many cleavage points). We find that the cleavage sites appear to correlate with alternating purine/pyrimidine stretches that border coding regions in the virus. Using simple combinatorial analysis, we found that alternating purine/pyrimidine stretches of 8–12 bp with one base out of register are significantly underrepresented in SV40, as they are in a number of other DNAs (16). However, in control regions, sequences of 10 alternating bases are overrepresented, whether considering strictly noncoding areas in SV40 or −50 bp from each coding region start and +50 bp from each coding end. Table 1 lists the major Λ-Co(DiP)$_3^{3+}$ cleavage sites obtained by the coarse assay and contiguous areas of possible biological relevance. The alternating sequences within the regions outlined in Table 1 show no base preference and are flanked by ≥4-base homopurine sequences twice as frequently as other alternating purine/pyrimidine sequences. Most strikingly, sites cleaved with the cobalt complex seem to correlate closely with noncoding control

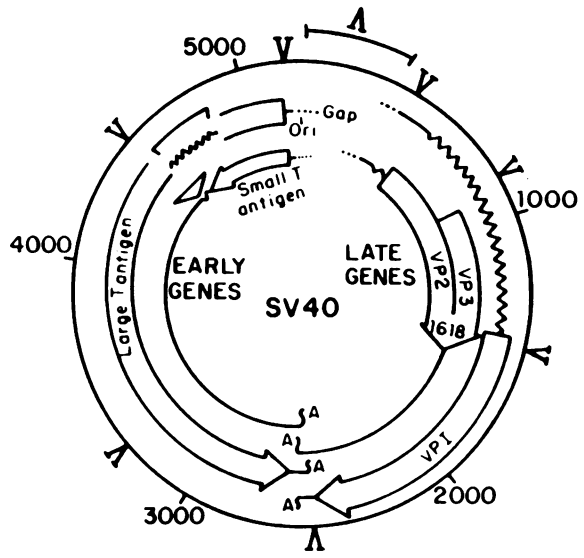


FIG. 1. Coarse map of Λ-Co(DiP)$_3^{3+}$ sites on SV40 ($-\sigma \approx 0.04$), obtained using nuclease S1 to cut opposite nicked sites (17). *Hpa* II, *Acc* I, *Eco*RI, and *Bam*HI were used for restriction cleavage. Numbers refer to position on the SV40 genome according to ref. 41. Ori, origin of replication; T antigen, tumor antigen; VP, virion protein; A, site of poly(A) addition. See Table 1.

regions in the plasmid. This finding is reminiscent of that seen earlier in the bacterial construct pBR322 (17). A majority of these recognition sites seem to occur within mRNA limits but, in fact, before or after the actual peptide-coding sequences. Indeed, the closest correlation seems to be with gene splicing sites. Perhaps we are actually detecting indirectly, on the DNA level, sequence-dependent variations in RNA structure. Extensive cleavage within the regulatory region is found, but the resolution here is too low to be informative (see below). As may be seen simply in Fig. 1,

Table 1. Coarse mapping of Λ-Co(DiP)$_3^{3+}$ cleavage in SV40

Cleavage sites*	Biological landmark	Alternating purine/pyrimidine sequence found in region†
102 ± 40	Ori (5135–5214)	126 ATGCATGC
	Enhancers (100–250)	198 ATGCATGC
436 ± 9	Late mRNA start (325)	420 TGTGTTTGT
	VP2 start (562)	
900 ± 14	VP3 start (916)	826 GTGTGAGCGC
		928 TGTATAGGC
1611 ± 75	VP2 and VP3 ends (1620)	1567 GTGCAAGTGC
2600 ± 50	Early mRNA end (2586)	2690 TGCATTCAT
	VP1 end (2593)	
	Late mRNA end (2674)	
3309 ± 60	Out-of-frame termination codon (3268–3273)	3224 TGTACTCATTCATG
4570 ± 54	T-antigen intron end (4572)	4576 ATACACAAACA
	Small T-antigen end (4639)	4611 TATACACTTA
4924 ± 48	Large T-antigen exon end (4918)	4954 ATGAGCATAT
5187 ± 140	Early mRNA start (5236)	5254 TGCATAAATA
	T-antigen starts (5163)	

Numbers refer to position along the SV40 genome according to ref. 41. See Fig. 1 legend for abbreviations.
*Error limits are derived from variations in fragment size measured in several trials and do not take the gel resolution into account.
†Out-of-register residues are underlined.

however, cleavage reactions with Λ-Co(DiP)$_3^{3+}$ in general show that conformationally distinct sites border coding regions throughout SV40 DNA.

All mapping experiments were conducted with native supercoiled SV40 DNA. Photolysis in the presence of the cobalt complex of relaxed closed circular DNA showed a reduction in single-strand scission by a factor >5. Hence both coarse and fine mapping could only be conducted with the supercoiled substrate.

Fine Mapping in SV40 DNA in the Regulatory Region. We next examined the regulatory region of SV40 DNA in more detail because of the apparent structural variety and biological importance of this region (30–32), because anti-Z-DNA antibodies and Z-DNA-binding proteins bind specifically to this region (8, 16), and because the coarse mapping had revealed extensive cleavage of this region with Λ-Co(DiP)$_3^{3+}$. This area contains the major control elements for the virus, including transcriptional promoters and enhancers (Fig. 2*a*). We have observed specific single-strand scission by the chiral probe within the transcriptional enhancer, in the promoter block, and in the vicinity of the origin of replication.

Fig. 2*b* shows autoradiographs of gels containing fragments cleaved by Λ-Co(DiP)$_3^{3+}$. Cleavage with the complex is shown for both strands, which were labeled at the *Hin*fI site on the 3′ and 5′ termini, respectively. Adjacent lanes show the sequencing reactions that mark guanine residues on both strands. In the two 72-bp repeats of the enhancer, three major sites of reaction with Λ-Co(DiP)$_3^{3+}$ are found, at positions 127/130, 180–183, and 203 of the 5′ strand and positions 131/133, 190–194, and 214 of the 3′ strand. Discrete cleavage at single interbase positions is evident. There is a distinct asymmetry in the cleavage of the two strands; the cleavage in the enhancer region is shifted for the 5′ strand to lower nucleotide numbers.

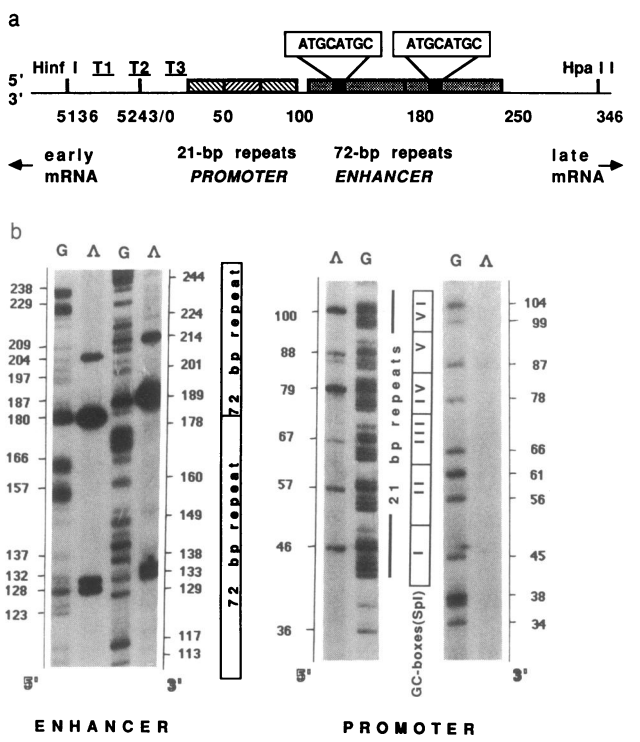


FIG. 2. Fine mapping of the regulatory gap. (*a*) Schematic view of the regulatory gap of SV40, showing positions of the 72-bp repeats of the enhancer, the promoter block consisting of the 21-bp repeats, and the origin of replication with the three binding sites for large T antigen. The two stretches of alternating purine/pyrimidine sequence within the enhancer region are given. (*b*) Autoradiographs of denaturing polyacrylamide gels resolving the products of photolysis with Λ-Co(DiP)$_3^{3+}$ of negatively supercoiled SV40 molecules (lanes Λ). Lanes G contain bands marking guanine residues of a Maxam–Gilbert sequencing reaction. Numbers correspond to positions in the genome relative to the *Bgl* I site (33). Both strands of the *Hin*fI–*Hpa* II fragment were labeled at the *Hin*fI site. Note the offset of bands relative to each other, with a shift for the 5′ late strand toward lower nucleotide numbers within the 72-bp repeats; no significant cleavage is observed for the 3′ strand within the 21-bp repeats, but the 5′ strand shows cleavage in each of the six "GC boxes."

Fig. 3 reveals in detail the cleavage pattern about the 8-bp stretch 127–134 with a shorter 5′ strand and longer 3′ strand. The asymmetry in cleavage pattern to the 5′ side is consistent with access of the complex from the major groove of the helix. Whether the conformation at this site were left- or right-handed, the same asymmetric cleavage pattern would result, given that the complex binds from the major groove. Tris(diphenylphenanthroline) and phenanthroline complexes have been shown to bind by intercalation into the major groove of B-DNA (37). Importantly, the cleavage sites obtained here (127/late, 133/early) flank the 8-bp alternating purine/pyrimidine sequence suggested earlier (16) to adopt a left-handed conformation based upon the binding of anti-Z-DNA antibodies. This specific cleavage pattern by Λ-Co(DiP)$_3^{3+}$ therefore supports a left-handed conformation at the bound site. From this pattern of cleavage, it is tempting to speculate on how the molecule may reside in its site. The "inner" cuts at 130/131 could possibly reflect reaction of an intercalated ligand, whereas the two "outer" cuts (127/133) could reflect reactions of the nonintercalated ligands aligned along the wide and shallow major groove of the helix.

We find a different cleavage pattern in the second 72-bp repeat of the enhancer. In the closed circular DNA, the conformation seems to differ over these two regions of identical sequences. Here we observe the strongest bands at positions 180 and 190 for the late and early strand, respectively. However, these intense bands are not consistently observed over different samples and have appeared in one DNA fragment irradiated in the absence of Λ-Co(DiP)$_3^{3+}$. Perhaps small variations in torsional strain that appear with the different DNA preparations account for this variability. The asymmetric cleavage pattern at the two other strong sites in this region, at positions 203 (late) and 213 (early), again reveals the equivalent of 5′ protruding ends. However, the cleavage site is not identical to that of the first 72-bp repeat; the identical 8-bp stretch 5′ ATGCATGC 3′ runs from

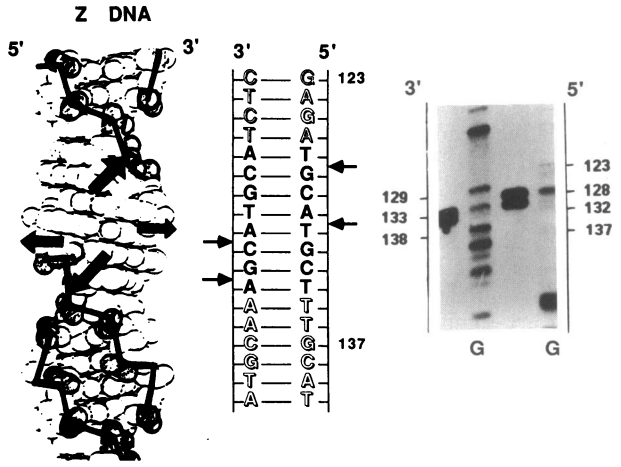


FIG. 3. Model illustrating the topology associated with cleavage by Λ-Co(DiP)$_3^{3+}$ in the SV40 enhancer block containing the 8-bp stretch of alternating purines and pyrimidines. Base pairs composing the binding and cleavage site are shaded. Arrows near the sequence indicate positions of strand scission in the 3′ and 5′ strand as derived from the autoradiograph shown at right. Arrows on the model show corresponding positions along the helix.

positions 198 to 204. A shift in the reactive site for the chiral molecule appears here, toward the right (late) 7-bp alternating purine/pyrimidine flanking sequences. This shift may be related to the strand breakage at 180/190; when the 180/190 bands are not apparent, cleavage bands at positions 195 and 203, for late and early strand respectively, are found (data not shown); this pattern may also reflect an increase in the size of the left-handed segment, extending out from the alternating purine/pyrimidine stretch with possible hyperreactivity at the junctions. Experiments to analyze cleavage as a function of supercoiling will be needed to sort this out. We have in various experiments also observed additional weak cleavage sites at 139, 150, 162, 172, 183, and 240 only for the late strand. Some of these positions seem to correspond with regions sensitive to nuclease digestion (42) and activated by drug binding (41). Experiments with the cobalt complexes overall support the notion that particular sites in this region are in some respects conformationally sensitive and reactive (43).

The invariant locations at which cleavage with Λ-Co(DiP)$_3^{3+}$ is correlated on early and late strands are the two segments within the enhancer 72-bp repeats, at positions 124–133 and in the vicinity of positions 196–205. These are two of the three locations proposed for Z-DNA formation within the SV40 regulatory region on the basis of anti-Z-DNA antibody studies (16). Λ-Co(DiP)$_3^{3+}$, a molecular probe distinct from an antibody, appears to recognize the same structure along the strand as does the antibody. A third antibody binding site, at 258–265 (358–232), was proposed on the basis of binding of the antibody to the 232–358 restriction fragment. We have not observed significant cleavage in this region.

Cleavage by the cobalt complex in the promoter block of the regulatory region is qualitatively different from that found in the enhancer block. In the 21-bp repeats of the promoter block, there appears to be a marked strand preference, and the asymmetry seen in the enhancer is lost. Strong cutting is found for the 5′ strand within each of the six "GC boxes" at 5′ GGG↓CGG 3′; only weak cleavage is visible at the complementary motif of the 3′ strand. Intensities of strand scission among the GC boxes of the 5′ strand vary and, compared to each other, are found to be strong for boxes IV and VI, medium for I and II, and weaker for III and V. The sites of weakest cleavage correspond with those that show higher binding affinities for the transcription factor Sp1, and sites more sensitive to Λ-Co(DiP)$_3^{3+}$ correlate with boxes of minor influence for Sp1 binding (31). It is difficult to speculate about the structure of this sequence, but certainly the structures in these GC boxes must differ from those recognized by the cobalt complex in the enhancer region; the cleavage pattern differs greatly from that of the enhancer, where both strands are cut with equal intensity and asymmetrically. These differing cleavage patterns point to the presence in the pure closed circular DNA of a family of conformationally distinct sites.

Studies by others (44–47) have indicated structural heterogeneity in the vicinity of the origin of replication. We observe resistance to *Bgl* I, which cuts at the origin, after photolysis in the presence of Λ-Co(DiP)$_3^{3+}$. We have also observed, in some samples, bands that correspond to cleavage at or near the origin of replication (positions 5218, 5220, 5243, and 6 for the 5′ strand and 5222, 5223, 7, and 8 for the 3′ strand).

Overall, our results indicate that the regulatory region of SV40 displays a remarkable conformational heterogeneity. In particular, cleavage with Λ-Co(DiP)$_3^{3+}$ suggests that a left-handed site may exist within the domain essential for enhancer activity, a domain that binds cellular factors. A different conformation is seen through the cobalt cleavage experiments in the promoter block, at sites that bind specific transcription factors. Perhaps these conformations recognized by the cobalt complex are also essential to the recognition and regulation by proteins.

Fine Mapping in SV40 DNA Near the 3′ Termini. Fig. 4 shows the results of higher resolution Λ-Co(DiP)$_3^{3+}$ mapping of the 288-bp *Hin*fI-*Bam*HI (2537–2825) fragment. In contrast to the regulatory region, this region of SV40 is cleaved by Λ-Co(DiP)$_3^{3+}$ at only a few locations and with high specificity. On the 5′ strand, major cleavage sites are at positions 2591 and 2717. The 3′ strand is cleaved less overall; two major sites are at positions 2687 and 2643, and weaker bands appear at positions 2694 and 2705. Our certainty is ±2 bp for this region.

What do these major sites have in common? The actual strand scissions all occur at or near GT sites, but not all GT sites are cleaved. While it is possible that Λ-Co(DiP)$_3^{3+}$ might bind preferentially to this sequence, these results may also suggest that these GT sites are particularly reactive. A recent NMR study (48) indicated a local maximum of the imino proton exchange rate at GTG sequences in several DNA regulatory regions; it was suggested that these frequently occurring sequences may serve as markers for protein interaction. On the 3′ strand, cobalt cleavage sites border 8-bp alternating purine/pyrimidine segments [with an in-register mistake (49)]. Further, the strongest cleavage sites, at positions 2591 and 2717 on the 5′ strand, consist of a 7-bp alternating purine/pyrimidine segment with one base out of register (boxed) next to at least 4 bp of a homopurine stretch (50). The cleavage pattern is, in addition, different on the two strands; Λ-Co(DiP)$_3^{3+}$ cleavage occurs on one strand but at neither the opposite nor the symmetric position on the other strand. We cannot analyze the detailed polarity for strand cleavage in this region, because the resolution of the gels is

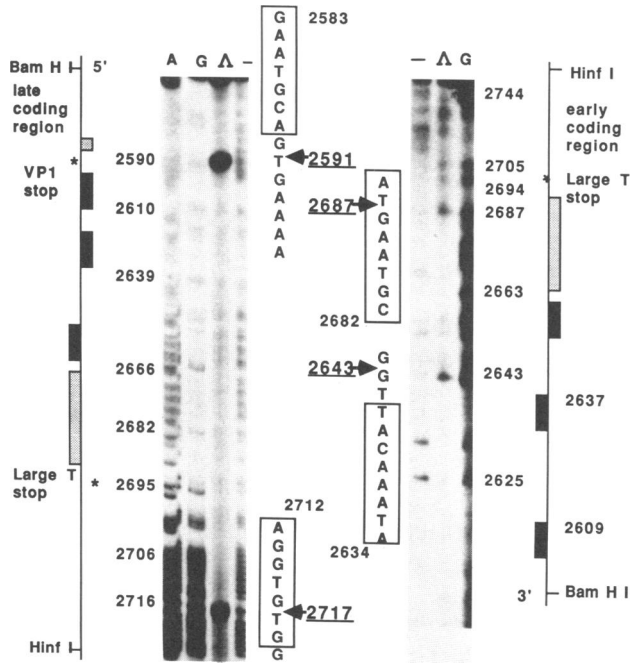


FIG. 4. Fine mapping of Λ-Co(DiP)$_3^{3+}$ cleavage sites near the 3′ termini of genes in SV40. Autoradiographs of sequencing gels show sites of strand scission with Λ-Co(DiP)$_3^{3+}$ on both strands of the region 2533–2750 (lanes Λ). Lanes G and A show bands marking guanine and adenine residues of Maxam–Gilbert sequencing reactions. Control samples irradiated with light without the addition of Λ-Co(DiP)$_3^{3+}$ are shown in lanes –. Alternating purine/pyrimidine sequences with a base out of register are boxed. Strand orientation and biological landmarks are indicated in diagrams at left and right. Large T, large tumor antigen; black boxes, AAUAAA polyadenylylation signals; stippled boxes, poly(A) tails; asterisks, UAA termination codons.

not suitable. There is certainly no extended sequence homology common to these sites.

Distinct conformations recognized by the cobalt complex again appear to correlate with the ends of coding regions. The strong 2591 cleavage is positioned directly at the termination codon for the late genes. On the 3′ strand, the 2643 scission occurs between polyadenylylation consensus sequences (AATAAA) and the termination codons for both early and late genes. An unusual conformation at this location could conceivably influence both early- and late-transcript processing. In fact, Λ-Co(DiP)$_3^{3+}$ cleavage sites flank the polyadenylylation sites in both directions. The 2687 site occurs in an area that was found to be essential for correct formation of late mRNA (51–53), and the site showed no clear sequence requirement. Hence this conformationally distinct site recognized by the cobalt complex corresponds to a known site downstream of a gene that is necessary for its proper expression.

Conclusions. Λ-Co(DiP)$_3^{3+}$ targets a family of conformationally distinct sites along SV40 DNA. With this simple chiral probe, site-specific cleavage along the strand may be obtained. The conformationally distinct sites cleaved by Λ-Co(DiP)$_3^{3+}$ show a high correlation with sites important as control elements along the genome and, in fact, flank protein-coding regions. Just as we see that local DNA conformations are clearly recognized and specifically bound by the simple coordination complex, perhaps also the secondary structure of DNA plays an active part in the recognition or release of transcriptional protein factors, indeed in the regulation of expression. Molecular probes targeted to specific structures along the strand may be powerful tools in finding clues to the biological roles of DNA secondary structures.

We are grateful for the generous support of the National Institutes of Health (GM33309), the National Foundation for Cancer Research, and (to B.C.M.) the Deutscher Akademischer Austauschdienst and Verband der Chemischen Industrie.

1. Rich, A., Nordheim, A. & Wang, A. H.-J. (1984) *Annu. Rev. Biochem.* **53**, 791–846.
2. Kmiec, E. B., Anjelides, K. J. & Holloman, W. K. (1985) *Cell* **40**, 139–145.
3. West, S. C. & Korner, A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6445–6449.
4. Costlow, N. A., Simon, J. A. & Lis, J. T. (1985) *Nature (London)* **313**, 147–149.
5. Schon, E., Evans, T., Welsh, J. & Efstratiadis, A. (1983) *Cell* **35**, 837–848.
6. Lilley, D. M. J. (1986) *Nature (London)* **320**, 487–488.
7. Larsen, A. & Weintraub, H. (1982) *Cell* **29**, 609–622.
8. Azorin, F. & Rich, A. (1985) *Cell* **41**, 365–374.
9. Peck, L. J. & Wang, J. C. (1985) *Cell* **40**, 129–137.
10. Pohl, F. M. & Jovin, T. M. (1972) *J. Mol. Biol.* **57**, 375–396.
11. Ellison, M. J., Feigon, J., Kelleher, R. J., III, Wang, A.H.-J., Habener, J. F. & Rich, A. (1986) *Biochemistry* **25**, 3648–3655.
12. Peck, L. J., Nordheim, A., Rich, A. & Wang, J. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 4560–4564.
13. Haniford, D. B. & Pulleyblank, D. E. (1983) *Nature (London)* **302**, 632–634.
14. Walmsley, R. M., Szostak, J. W. & Petes, T. D. (1983) *Nature (London)* **302**, 84–86.
15. Zarling, D. A., Arndt-Jovin, D. J., Robert-Nicoud, M., McIntosh, L. P., Thomae, R. & Jovin, T. M. (1984) *J. Mol. Biol.* **176**, 369–415.
16. Nordheim, A. & Rich, A. (1983) *Nature (London)* **303**, 674–679.
17. Barton, J. K. & Raphael, A. L. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6460–6464.
18. McCutchan, T. F., Hansen, J. L., Dame, J. B. & Mullins, J. A. (1984) *Science* **225**, 625–628.
19. Sheflin, L. G. & Kowalski, D. (1985) *Nucleic Acids Res.* **13**, 6137–6154.
20. Keene, M. A. & Elgin, S. C. R. (1984) *Cell* **36**, 121–129.
21. Cartwright, I. L. & Elgin, S. C. R. (1982) *Nucleic Acids Res.* **10**, 5835–5852.
22. Pulleyblank, D. E., Haniford, D. B. & Morgan, R. A. (1985) *Cell* **42**, 271–280.
23. Spassky, A. & Sigman, D. S. (1985) *Biochemistry* **24**, 8050–8056.
24. Cartwright, I. L., Hertzberg, R. P., Dervan, P. B. & Elgin, S. C. R. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3213–3217.
25. Johnston, B. A. & Rich, A. (1985) *Cell* **42**, 713–724.
26. Herr, W. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8009–8013.
27. Nordheim, A. & Runkel, L. (1986) *J. Mol. Biol.* **189**, 487–501.
28. Lilley, D. M. J. & Palacek, E. (1984) *EMBO J.* **3**, 1187–1192.
29. Peck, L. J. & Wang, J. C. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 6206–6210.
30. Cereghini, S., Herbomel, P., Jouanneau, J., Saragosti, S., Katinka, M., Bourachot, B., de Crombrugghe, B. & Yaniv, M. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 935–944.
31. Gidoni, D., Kadonaga, F. T., Barrera-Saldana, H., Takanishi, K., Chambon, P. & Tjian, R. (1985) *Science* **230**, 511–517.
32. Takanishi, K., Vigneron, M., Matthes, H., Wildeman, A., Zenke, M. & Chambon, P. (1985) *Nature (London)* **319**, 121–126.
33. Tooze, J. (1981) *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed.
34. Barton, J. K. & Raphael, A. L. (1984) *J. Am. Chem. Soc.* **106**, 2466–2468.
35. Barton, J. K. (1986) *Science* **233**, 727–734.
36. Kumar, C. V., Barton, J. K. & Turro, N. J. (1985) *J. Am. Chem. Soc.* **107**, 5518–5523.
37. Barton, J. K., Goldberg, J. M., Kumar, C. V. & Turro, N. J. (1986) *J. Am. Chem. Soc.* **108**, 2081–2086.
38. Barton, J. K., Danishefsky, A. T. & Goldberg, J. M. (1984) *J. Am. Chem. Soc.* **106**, 2172–2176.
39. Barton, J. K., Basile, L. A., Danishefsky, A. T. & Alexandrescu, A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1961–1965.
40. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
41. Reynolds, V. L., Molineux, I. J., Kaplan, D. J., Swenson, D. M. & Hurley, L. H. (1985) *Biochemistry* **24**, 6228–6237.
42. Saragosti, S., Cereghini, S. & Yaniv, M. (1982) *J. Mol. Biol.* **160**, 133–146.
43. Nussiniv, R., Shapiro, B., Lipkin, L. E. & Maizel, J. V., Jr. (1984) *J. Mol. Biol.* **177**, 591–607.
44. Weaver, D. T. & DePhamphilis, M. L. (1984) *J. Mol. Biol.* **180**, 961–986.
45. Gerard, R. D., Montelone, B. A., Walter, C. F., Innis, J. W. & Scott, W. A. (1985) *Mol. Cell. Biol.* **5**, 52–58.
46. Hsu, M. T. (1985) *Virology* **143**, 617–621.
47. DeLucia, A. L., Deb, S., Partin, K. & Tegtmeyer, P. (1986) *J. Virol.* **57**, 138–144.
48. Lu, P., Cheung, S. & Donlan, M. (1985) *Adv. Biophys.* **20**, 153–175.
49. Wang, A. H.-J., Gessner, R. V., Van der Marel, G. A., van Boom, J. H. & Rich, A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3611–3615.
50. Cantor, C. R. & Efstratiadis, A. (1984) *Nucleic Acids Res.* **12**, 8059–8072.
51. Conway, L. & Wickens, M. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3949–3953.
52. Fitzgerald, M. & Shenk, T. (1981) *Cell* **24**, 251–260.
53. Grass, D. S. & Manley, D. L. (1986) *J. Virol.* **57**, 129–137.