More than 80% (approximately 29 kilobase pairs) of the adenovirus serotype 2 genome was surveyed for the presence of unusual DNA conformations. Seven recombinant DNAs containing the largest HindIII fragments of AD2 DNA were analyzed for the presence of negative supercoiled-dependent S1 nuclease-sensitive sites. Four plasmids each contained a specific site of S1 nuclease sensitivity whereas the other three showed no reaction. Further investigation was focused on a plasmid containing one of the positively reacting fragments (fragment C) which contained the major late promoter at coordinate 16.4 on the genome; three serotypes (Ad2, Ad7, Ad12) were studied. Fine mapping studies revealed the S1-sensitive sites to be a small region (~6 base pairs) located at the TATA box of the major late promoter in all three cases. Other determinations (supercoil relaxation, T7 gene 3 product sensitivity, bromoacetaldehyde reactivity, anomalous gel mobility, the influence of negative superhelical density on nuclease sensitivity) led to the conclusion that the B-helix deformation was not due to a previously recognized DNA conformation (left-handed Z-DNA, cruciform, bent DNA), but may be accounted for by the homopurine-homopyrimidine nature of this region.

Adenovirus is a well characterized tumor virus which contains a linear duplex DNA genome of 36 kb (1). The entire sequence of adenovirus serotype 2 DNA has recently been established.²

Unusual DNA structures characterized to date include the following: left-handed Z-DNA, which is found principally at alternating purine-pyrimidine sequences (3-5); cruciforms which occur at inverted repeats (6-8); slipped structures, which were postulated to occur at direct repeats (9, 10); and bent DNA, detected in fragments from kinetoplast DNAs (11) and three genetic control regions (12, 13).³ Negative supercoiling at physiological densities stabilizes all of these unusual structures except bent DNA (6, 9, 15). S1 nuclease is a sensitive and convenient probe for left-handed Z-DNA, cruciforms and slipped structures (6, 9, 15).

In an effort to extend our studies on the structure-function relationship of regions of DNA, we surveyed the adenovirus genome for the presence of unusual conformations. The availability (16) of ~80% of the adenovirus serotype 2 genome, cloned as HindIII fragments in pBR322, facilitated this approach. Our efforts focused in particular at the region containing the major late promoter of adenovirus serotypes 2, 7, and 12. Godding and Russell (17) generally localized an S1 nuclease-sensitive site near the TATA box of the major late promoter of Ad2. We have employed a variety of techniques including fine mapping to determine the type of unusual DNA secondary structure responsible for this behavior. We conclude that the unusual conformation is not a cruciform but may be due to its purine-pyrimidine bias.

**EXPERIMENTAL PROCEDURES**

**DNAs and Enzymes—** S1 nuclease and restriction enzymes were from Bethesda Research Laboratories. Topoisomeric samples of plasmids were generated and characterized essentially as reported previously (18). T7 gene 3 protein (19) was a generous gift of Dr. Joseph Coleman (Yale University).

**Plasmids—** Plasmid DNA was isolated and characterized as described previously (20). Plasmids containing the HindIII fragments of adenovirus serotype 2 cloned into the unique HindIII site of pBR322 were characterized previously (16). pRW786 (previously named pJB757 A (21)) contains the 375 bp BamHI-HindIII fragment encompassing the major late promoter of adenovirus serotype 7 cloned into the BamHI-HindIII sites of pBR322. pRNA7 contains the 644-bp EcoRI-HindIII fragment encompassing the major late promoter of adenovirus serotype 12 cloned into the EcoRI-HindIII sites of pBR322; this plasmid was derived from clone Ad12 RI D which was provided to J. A. E. by Dr. Walter Doerfler (University of Cologne, Cologne, West Germany). pRW784 was constructed by digesting the clone of the HindIII C fragment of Ad2 in the HindIII site of pBR322 with SalI and XhoI to generate two fragments of 3601 and 4394 bp. The 4394-bp fragment was isolated and ligated to form pRW784, a 4384-bp plasmid containing the major late promoter of Ad2 in a 453-bp HindIII-XhoI fragment within the HindIII-SalI sites of pBR322. pRW784, -786, and -787 were characterized by restriction mapping with several enzymes in each case to confirm the expected sequences. pGF3 containing the 5'-flanking region of the chicken β-globin gene was obtained from Dr. Gary Felsenfeld (National Institutes of Health) (22). pBR16 contains sequences from the adult chicken β-globin gene as described (23) and was a generous gift of Dr. James D. Engel (Northwestern University). pHTA5'-186 contains the 5'-flanking regions of Drosophila melanogaster heat shock genes and was obtained from Dr. Andrew A. Travers (Medical Research Council, Cambridge, Great Britain) (9). pCol-Md containing a 1.25-kb fragment of the mouse α(2) collagen promoter was obtained from Dr. Benoit de Crombrugghe (National Institutes of Health) (10). pRW784, -786, and -787 were characterized by restriction mapping with several enzymes in each case to confirm the expected sequences.

**S1 Nuclease Reactions—** To determine the sensitivity of regions of the adenovirus genome to S1 nuclease, plasmids containing fragments of the adenovirus genome (1.5 μg) at appropriate superhelical den-
Adenovirus DNA Structures

Fine Mapping—For fine mapping of the S1 nuclease-sensitive sites in pBR784, -786, and -787, the supercoiled plasmids (2 μg) (−α = 0.06) were treated with S1 nuclease (1 unit) at 37°C for 5 min. The plasmids were then digested with HindIII and the resulting fragments dephosphorylated with calf intestinal phosphatase and labeled at their 5' termini with polynucleotide kinase and [γ-32P]ATP using established procedures. The linear plasmids were then digested with a restriction enzyme that cuts the other side of the S1-sensitive site from HindIII (TaqI for pRW784, BamHI for pRW786, and EcoRI for pRW787) and the resultant fragments separated on a 4% polyacrylamide gel electrophoresis. The major late promoter containing fragments were eluted from the gel and run, alongside sequencing ladder markers, on an 8% acrylamide, 50% urea gel to determine the sites of S1 nuclease cleavage. For pRW786 and pRW787 the procedure was repeated but reversing the order of cleavage with HindIII and BamHI or EcoRI.

Bromoacetaldehyde Reactions—Bromoacetaldehyde (BAA) was prepared from bromoacetaldehyde diethylacetal as described (24). Reaction of BAA with plasmids and analysis of BAA-sensitive sites was essentially as described previously (25). Briefly, 2 μg of plasmid in 100 μl of 10 mM Tris-HCl, pH 7.6, 0.1 mM EDTA was incubated at 37°C for 2 h in the presence of 2% BAA. The DNA was recovered by precipitation and redisolved in the appropriate restriction enzyme buffer. Digestion was then carried out with the appropriate restriction endonuclease. 10 μl of S1 nuclease buffer (40 mM sodium acetate, 50 mM sodium chloride, 1 mM zinc sulfate, pH 4.6) containing 1.5 units of S1 nuclease was then added and was incubated at 37°C for 15 min. The reaction was terminated by addition of EDTA, pH 7.5, to 50 mM and the DNA fragments were separated by agarose gel electrophoresis.

Other Methods—Two-dimensional gel electrophoretic analysis of topoisomers of plasmids was carried out as described previously (26, 27). Reactions with T7 gene 3 protein were as described (28).

RESULTS

Supercil-dependent S1 Nuclease-sensitive Sites in the Adenovirus Genome—Fig. 1 shows a map of the HindIII fragments of the adenoavirus serotype 2 genome. Negatively supercoiled plasmids containing fragments A–F, and H were treated with S1 nuclease to detect unusual structural features. Fig. 2 shows agarose gel analyses for these seven plasmids; S1 nuclease-sensitive sites were located by mapping with HindIII. All seven plasmids show the presence of S1 nuclease specific bands of 3 and 1.3 kb. These bands are not present in the minus S1 nuclease control lanes, and represent cleavage by S1 nuclease at the pBR322 cruciforms (6, 7). In addition to these “cruciform specific” bands, lanes showing treatment of plasmids containing fragments A, B, C, and H contain other S1 nuclease specific bands, representing S1 nuclease specific sites in the adenoavirus DNA. These bands are of 7.1 and 0.9 kb for the plasmid containing fragment A, of 3.1 and 2.2 kb for the plasmid containing fragment B, of 3.2 kb for the plasmid containing fragment C. These and further mapping studies (with BamHI and HpaI for the fragment A containing plasmid; PstI and BgIII for the fragment B containing plasmid; XhoI for the fragment C containing plasmid; and EcoRI for the fragment H containing plasmid) confirmed the presence of S1 nuclease-sensitive sites in these four plasmids. Thus of the seven plasmids studied, four contained S1-sensitive sites. These sites occur approximately 6.0 (the region of the major late promoter), 9.4, 19.2, and 26.7 kb from the left terminal end of the adenoavirus serotype 2 genome as determined from these mapping studies. Ar, S1 nuclease-sensitive site in the region of the Ad2 major late promoter has previously been postulated as being due to the formation of a cruciform (17).

S1 Nuclease Sensitivity of the Major Late Promoters of Three Adenovirus Serotypes—Studies were focused on plasmids containing the major late promoters of three adenoavirus serotypes because of the biological importance of this region. To facilitate structural determinations on the major late promoters, small fragments encompassing this region were cloned into pBR322 for each of the three adenoavirus serotypes. pRW784 contains a 653-bp fragment from serotype 2, pRW786 contains a 375-bp fragment from serotype 7, and pRW787 contains a 664-bp fragment from serotype 12.

For pRW784, -786, and -787, an S1 nuclease-sensitive site, mapping to the region of the major late promoter, was found in each case. Fig. 3 shows the influence of negative supercoiling on the sensitivity of the site in the major late promoter in pRW787 to S1 nuclease. At a superhelical density of −0.008 (lane 1), no S1 nuclease sensitivity was observed. However, at negative superhelical densities above 0.045 (lanes 2, 3, and 4), fragments of 1196 and 1065 bp resulting from cleavage in the major late promoter were evident.

Similar studies were performed on pRW784 and pRW786 with the phage T7 gene 3 product. This enzyme has similar
properties to S₁ nuclease but shows optimum activity around neutral pH. The sites of cleavage were virtually identical to those observed with S₁ nuclease (data not shown). Thus the unusual structural feature recognized by S₁ nuclease is not generated by the lower pH of the S₁ nuclease reaction conditions.

Fig. 4 shows the influence of negative supercoiling on the S₁ nuclease sensitivities of pRW784, pRW786, and pRW787. S₁ nuclease reactions were performed as described under "Experimental Procedures"; primary data for four of the topoisomer populations of pRW787 is presented in Fig. 3. For pRW787, specific cleavage was determined from the intensity of the 1196-bp S₁ nuclease specific band, quantitated by scanning a photographic negative of the gel with a Joyce-Loebl microdensitometer. The intensity of the 396-bp sensitive site was used as a normalization factor to correct for any differences in the total amount of DNA loaded in each lane. Similar analyses were carried out for pRW784 and pRW786. Maximum specific cleavage (of around 30% cleavage as judged by the decrease in the intensity of the 2265-bp band (which contains the major late promoter S₁ nuclease-sensitive site) relative to the minus S₁ nuclease control. Closed circle, pRW786; open circle, pRW787; open triangle, pRW784.

Fine Mapping of S₁ Nuclease Cleavage Sites—Fine mapping studies were performed on the S₁ nuclease-treated plasmids to determine the initial sites of cleavage at the base pair level. Fig. 5 shows a typical analysis on pRW786; the site spans a 5-bp region as shown in Fig. 6. Similar analyses were performed on pRW784 and -787 (Fig. 6). In all three cases, the sites of cleavage are at the 5' side of the TATA box.

Analyses for Left-handed Z-DNA and Bent DNA—Supercoil relaxation studies provide a critical analysis for the presence and length of left-handed Z-DNA (15, 26, 27). Two-dimensional gels (26-28) were used to analyze topoisomers of pRW786. No relaxation (discontinuity) of the topoisomers was observed up to a supercoil density of -0.06 (data not shown). Thus we conclude that the structural feature present
in the major late promoter is unlikely to be left-handed Z-DNA. Bent DNA has been identified in DNA fragments from kinetoplasts and three genetic control regions (12, 13, 29). Anomalous gel mobility is a convenient method for diagnosing the presence of bent DNA fragments. The 375-bp HindIII-BamHI fragment from pRW786 was analyzed on 2% agarose at 22 °C and 6% polyacrylamide at 4, 22, and 55 °C along with markers of the 123-bp ladder and HindII fragments of pSP64 (a control plasmid) and its derivative pHK09 (30) which contains a segment of bent DNA (14). No unusual gel mobility was observed indicating the absence of bent DNA.

Bromoacetaldehyde Reactions—Bromoacetaldehyde is a sensitive probe for cruciforms (25, 31), reacting with A and C residues in the non-paired loops to give imidazole derivatives which cannot base pair (32). The location of these “wedged open” regions can be mapped with S1 nuclease after linearization of the plasmid. Fig. 7 shows the results of probing supercoiled pRW786 with BAA. No reaction of BAA at the major late promoter was observed whereas the presence of 2.7- and 1.7-kb bands corresponding to reaction at the pBR322 cruciforms acted as a positive control. As expected, no fragments corresponding to BAA reaction were observed with Sall linearized pRW786. Results similar to those found for pRW786 were found for pRW787. Hence we conclude that the S1 nuclease-sensitive sites in the major late promoters are not due to the presence of cruciforms.

To evaluate the potential for BAA to react with other structural features, we performed similar studies on pCol-Md, pGF3, pβ1BR16, and pH7D5'-186. pCol-Md and pTHD5'-186 contain S1 nuclease-sensitive direct repeats which have been postulated to form slipped structures (9, 10); since these slipped structures contain non-paired bases, they might be expected to be BAA reactive. The two other plasmids (pGF3 and pβ1BR16) contain S1 nuclease-sensitive dG-dC tracts (22, 23). In all four cases, no BAA sensitivity was observed in the inserts although the vector cruciforms were reactive, thus serving as positive controls. However, it may be noted that a recent report (33) indicates the BAA reactivity with a dG-dC tract in other plasmids under somewhat different conditions.

DISCUSSION

The determination of the entire sequence of Ad2 DNA (2) enabled a survey for unusual DNA structural features in a large mammalian viral genome. The presence of single, well-defined nuclease reactive loci in negatively supercoiled plasmids containing four of the large HindIII fragments of adenovirus but not in clones containing three other fragments indicates the rarity of these structural features. Clearly, the presence of an unusual DNA structure detected by S1 nuclease in a region of such biological activity as the major late promoter makes it especially worthy of interest. Interestingly, despite the presence of a number of regions with important biological functions in the areas of the genome contained in HindIII fragments D, E, and F, no S1 nuclease-sensitive sites were found therein. Definition of the structural features responsible for the S1 nuclease-sensitive sites in fragments A, B, and H will require further investigation (the structure in fragment C is discussed below). Attempts to correlate the presence or absence of S1 nuclease-sensitive sites with known biological functions on the genome would be highly speculative and hence will be postponed until further biochemical and genetic results are obtained.

Adenovirus has not been shown to replicate through a circular intermediate (1). Negative supercoiling was used as a structural perturbant to mimic the in vivo influences of ligands, environmental conditions, and proteins for this DNA structural study.

The nuclease-sensitive site in plasmids containing fragment C was localized at the TATA box (+3 bp) of the major late promoter. The presence of a distinct DNA structural feature at the initiation point of the major late adenovirus transcript may participate in the regulation of this important genetic locus.

The position, breadth, and magnitude of the sigmoidal transitions observed for the influence of supercoiling on nuclease sensitivity in the major late promoters (Fig. 4) are similar to those observed previously for left-handed Z-DNA and cruciforms (8, 34, 35), and indeed the nuclease-sensitive site in the major late promoters may, in principle, be explained by the presence of Z-DNA, a cruciform, or a slipped structure. Indeed, several groups (17, 36, 37) have suggested that cruciforms may exist in the vicinity of the major late promoter and may be involved in the regulation of transcription. However, prior studies (5, 15, 26, 34) showed that two nuclease-sensitive sites (at each B-Z junction) are found for left-handed Z-DNA, which has generally been associated with alternating purine-pyrimidine sequences, whereas a single site is observed (6–8, 31, 35) in the center of an inverted repeat for a cruciform, or in a loop of a direct repeat for a slipped structure (9, 10). Interestingly, inspection of the sequences around the S1 nuclease sites at the TATA boxes for the three serotypes reveals the absence of these features.

Further evidence that the nuclease sensitivity in the major late promoters is not due to formation of left-handed Z-DNA is provided by the supercoil relaxation studies, since no relaxation was observed. In principle, this technique will also detect cruciforms but it should be noted that the pBR322 cruciforms

![Fig. 7. Reaction of pRW786 with bromoacetaldehyde.](image-url)
have not been demonstrated by this method although their presence is well established by other techniques; this analysis has not been applied to slipped structures, bent DNA, or homopurine-homopyrimidine regions.

Furthermore, it is unlikely that the structure at the region around the TATA box is bent, since S1 nuclease sensitivity was not found for other cloned bent fragments from kinetoplasts (29, 43), and no anomalous gel mobility was observed for fragments containing this region. Also, the sequence features generally associated with bent DNA (37, 38) are not present.

Further evidence that the unique structural feature in the TATA box locus is not a cruciform is provided by its lack of BAA reactivity. The inability of slipped structures and homopurine-homopyrimidine sequences in negatively supercoiled plasmids to react with BAA was determined. Interestingly, no BAA reactivity was observed despite the presumed presence of non-paired bases in the slipped structures. The lack of BAA reactivity also argues against the region around the major late promoter being especially thermally labile. Parenthetically, numerous studies have indicated that S1 nuclease sensitivity should not necessarily be equated with single strandedness (22, 23, 39–41).

In summary, the structure of the feature located at the major late promoter is uncertain. However, it should be noted that a 25-bp region surrounding the TATA box is 75–80% (depending on serotype) homopurine-homopyrimidine sequence. The S1 nuclease sensitivity of these types of sequences have unusual biochemical and physical properties (14, and reviewed in Ref. 2).

REFERENCES

1. Doerfler, W. (ed) (1983) Current Topics in Microbiology and Immunology, Vols. 109–111 (The Molecular Biology of the Adenoviruses) Springer-Verlag, Berlin

2. Wells, R. D., Blakesley, R. W., Burd, J. F., Chan, H. W., Dodgson, J. B., Hardie, S. C., Horne, G. T., Jensen, K. F., Larson, J., Nes, I. F., Selsing, E., and Wartell, R. M. (1977) Crit. Rev. Biochem. 4, 305–340

3. Wells, R. D., Brennan, R., Chapman, K. A., Goodman, T. C., Hart, P. A., Hillen, W., Kellogg, D. R., Kilpatrick, M. W., Klein, R. D., Klysk, J., Lambert, P. F., Larson, J. E., Miglietta, J. J., Neuendorf, S. K., O’Connor, T. R., Singleton, C. K., Stirdivant, S. T., Veneziale, C. M., Wartell, R. M., and Zacharias, W. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 77–84

4. Rich, A. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 1–12

5. Wells, R. D., Erlanger, B. F., Gray, H. B., Hanau, L. H., Jovin, T. M., Kilpatrick, M. W., Klysk, J., Larson, J. E., Martin, J. C., Miglietta, J. J., Singleton, C. K., Stirdivant, S. M., Veneziale, C. M., Wartell, R. M., Wei, C. F., and Zaring, D. (1983) in Genom Res. ULCA Symp. 8, 3–18

6. Panayotatos, N., and Wells, R. D. (1981) Nature 289, 466–470

7. Lilley, D. M. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 6468–6472

8. Singleton, C. K. (1982) J. Biol. Chem. 256, 7651–7666

9. Mace, H. A. F., Pelham, H. R. B., and Travers, A. A. (1983) Nature 304, 555–557

10. McKeon, C., Schmidt, A., and de Crombrugghe, B. (1984) J. Biol. Chem. 259, 6636–6640

11. Marini, J. C., Levene, S. D., Crothers, D. M., and Englund, P. T. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 7664–7668

12. Bossi, L., and Smith, D. M. (1984) Cell 39, 643–652

13. Wu, H. M., and Crothers, D. M. (1984) Nature 308, 509–513

14. Wells, R. D., Larson, J. E., Gant, R. C., Shortle, B. E., and Cantor, C. R. (1970) J. Mol. Biol. 45, 465–477

15. Singleton, C. K., Kilpatrick, J., Stirdivant, S. M., and Wells, R. D. (1982) Nature 299, 312–316

16. Hu, S. L., and Manley, J. L. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 820–824

17. Goding, C. R., and Russell, W. C. (1983) Nucleic Acids Res. 11, 3136

18. Singleton, C. K., and Wells, R. D. (1982) Anal. Biochem. 122, 253–257

19. Osterman, H. L., and Coleman, J. E. (1981) Biochemistry 20, 4884–4890

20. Wells, R. D., Miglietta, J. J., Klysk, J., Larson, J. E., Stirdivant, S. M., and Zacharias, W. (1982) J. Biol. Chem. 257, 10166–10171

21. Engler, J. A., Chow, L. T., and Broker, T. R. (1981) Gene (Amst.) 13, 133–140

22. Nickol, J. M., and Felsenfeld, G. (1983) Cell 35, 467–477

23. Dolan, M., Dodgson, J. B., and Engel, J. D. (1983) J. Biol. Chem. 258, 3983–3990

24. Kohwi-Shigematsu, T., Gelinis, R., and Weintraub, H. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 4389–4392

25. Kang, D. S., and Wells, R. D. (1986) J. Biol. Chem. 260, 7782–7790

26. Kilpatrick, M. W., Klysk, J., Singleton, C. K., Zarling, D. A., Jovin, T. M., Hanau, L. H., Erlanger, B. F., and Wells, R. D. (1984) J. Biol. Chem. 259, 11302–11309

27. Wang, J. C., Peck, L. J., and Becherer, K. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 85–92

28. Singleton, C. K., Kilpatrick, M. W., and Wells, R. D. (1984) J. Biol. Chem. 259, 1963–1967

29. Marini, J. C., Effron, D. N., Goodman, T. C., Singleton, C. K., Wells, R. D., Wartell, R. M., and Englund, P. T. (1984) J. Biol. Chem. 259, 8974–8979

30. Krause, H. M., and Higgins, N. P. (1986) J. Biol. Chem. 261, 3744–3752

31. Lilley, D. M. (1983) Nucleic Acids Res. 11, 3097–3113

32. Kayasuga-Mikado, K., Hashimoto, T., Negishi, T., Negishi, K., and Hayatsu, H. (1980) Chem. Pharm. Bull. (Tokyo) 28, 932–938

33. Kohwi-Shigematsu, T., and Kohwi, Y. (1985) Cell 43, 199–206

34. Singleton, C. K., Klysk, J., Stirdivant, S. M., and Wells, R. D. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 2447–2451

35. Singleton, C. K. (1983) J. Biol. Chem. 258, 7661–7668

36. Baker, C. C., and Ziff, E. (1980) Cold Spring Harbor Symp. Quant. Biol. 44, 415–428

37. Trifinov, E. N. (1985) Crit. Rev. Biochem. 19, 98–106

38. Hagerman, P. J. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 4652–4656

39. Huo, H., Lund, E., and Dahlberg, J. E. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 7288–7292

40. Schon, E., Evans, T., Welsh, J., and Efstratiadis, A. (1983) Cell 35, 837–848

41. Cantor, C. R., and Efstratiadis, A. (1984) Nucleic Acids Res. 12, 8055–8071

42. Margot, J. B., and Hardison, R. C. (1985) J. Mol. Biol. 194, 195–210

43. Kitchin, P. A., Klein, V. A., Ryan, K. A., Gann, K. L., Rauch, C. A., Kang, D. S., Wells, R. D., and Englund, P. T. (1986) J. Biol. Chem. 261, 11302–11309