Pergamon

Tetrahedron Vol. 50, No. 5, pp. 1361-1378, 1994

0040-4020/94 $6.00+0.00

0040-4020(93)E0206-U

Features of DNA Recognition for Oriented Binding and Cleavage by Calicheamicin

Stanley C. Mah, Mary Ann Price,‡ Craig A. Townsend,* and Thomas D. Tullius*

Department of Chemistry, The Johns Hopkins University, Baltimore, MD 21218, U.S.A.

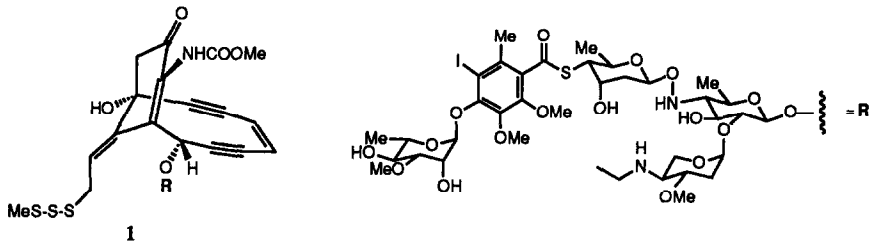
Abstract: The cleavage by calicheamicin γ_1^I of duplex DNA containing four or more contiguous adenines on one strand proceeds through C5' hydrogen abstraction at the penultimate thymidine at the 5' end of the T-tract. Cleavage of longer A-tracts (A_{16} and A_{42}) results, in addition to the primary cleavage site, in a cascade of weaker cuts that persists until the four or fifth nucleotide from the 5' end of the A-tract. Cleavage by calicheamicin γ_1^I of free DNA known to form nucleosomes shows a periodic pattern that is correlated to structural periodicities found in the DNA. We suggest that calicheamicin recognizes and cleaves sequences of DNA that have a narrow minor groove, and regions of DNA that may have a propensity for helix deformation.

The reaction of calicheamicin (CLM) γ_1^I (**1**) with restriction fragments from prokaryotic plasmid DNA has been investigated in this laboratory and elsewhere.[1-5] In contrast, the interactions of CLM with the DNA of higher organisms has received little attention. In this paper we examine the interactions of CLMγ_1^I with a DNA sequence from a eukaryote that is known to form a uniquely positioned nucleosome, and with synthetically designed duplex DNA sequences containing stretches of oligo(dA)•oligo(dT). From the results of these experiments we derive several principles of the recognition and cleavage of DNA by CLM.

From the earlier experiments on DNA from prokaryotic systems it was apparent that DNA cleavages caused by calicheamicin were surprisingly sequence-selective and chiefly double-stranded.[1-4] The sequence TCCT (and similar sequences) was noted in this early work as a particularly favorable site for cleavage.[1] Subsequently, we have used synthetic oligonucleotides containing TCCT sequences to investigate the activation of CLMγ_1^I,[6-8] and the identity of the deoxyribosyl hydrogens that are abstracted from the DNA backbone in the first step of oxidative strand scission.[9-11] These experiments show impressive specificity in

‡ Present address: Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, U.K.

the atom transfer process with respect to both the identity of the DNA hydrogens removed and the location of atom transfer on the reduced form of the drug.[10,11] CLMγ_1^I removes a hydrogen atom from C5' of a deoxyribose on one strand of the DNA,[10] and from C4' of a deoxyribose across the minor groove on the other strand.[11] It may be deduced from these experiments that CLM binds in the minor groove in a single orientation such that the carbohydrate side chain is directed to the 3'-side of the C5' abstraction site in the TCCT strand of the dodecamer used.[10,11]



1

Structure **1** - Calicheamicin γ_1^I

While these experiments defined direction, they did not demonstrate location of the oligosaccharide side chain relative to DNA, and pertained in strict terms to a single DNA sequence favorable to cleavage. Generalization of the observations in the atom transfer experiments was made in hydroxyl radical footprinting studies on CLMε, the reduced rearrangement product of CLMγ_1^I, bound to the *Nar*I-*Eco*RI restriction fragment of plasmid pUC18.[5] For the sequences examined, footprints were observed where cleavage occurred. Approximately 4-5 nucleotides on each strand of DNA are protected by bound CLMε from hydroxyl radical attack. The footprinting and cleavage data are best fit by a model in which CLM binds in an extended conformation[12] such that the oligosaccharide is oriented within and along the minor groove to the 3'-side of the site of C5'-deoxyribosyl hydrogen abstraction.[5]

Assignment of sites of C5'-abstraction by CLMγ_1^I. At relatively isolated CLMγ_1^Icleavage sites we were able to distinguish between C5' and C4' hydrogen abstraction by the drug. Assignment of the sites of C5'-abstraction by CLMγ_1^I is readily made as this reaction leaves a DNA strand with a 5'-aldehyde terminus.[1,3,13] This terminus is easily converted by base treatment and β-elimination to a DNA fragment with a 3' phosphate end. The 5'-aldehyde and the base-catalyzed elimination product are easily distinguished on DNA sequencing gels on the basis of their differing electrophoretic mobilities.[1,3,13] Similarly, in well-resolved portions of the gel it was possible to observe phosphoglycolate-containing fragments indicative of C4'-hydrogen abstraction expected for the complementary strand.[4,11] An analysis of this kind has also been used by Stubbe and Kozarich to investigate the nature of DNA cleavages by the esperimicins.[14] Application of this approach is shown in Figure 1 where the *Nar*I-*Eco*RI restriction fragment of pUC18 is alternatively end-labeled with ^{32}P at the 3'-end of one strand and then the other, so that the nature of the 5' termini of the cleavage sites may be investigated. Assignment of the sites of C5' hydrogen abstraction for a section of the *Nar*I-*Eco*RI fragment is shown in Figure 2. Thus, from three experimental points of view, namely atom transfer, hydroxyl radical

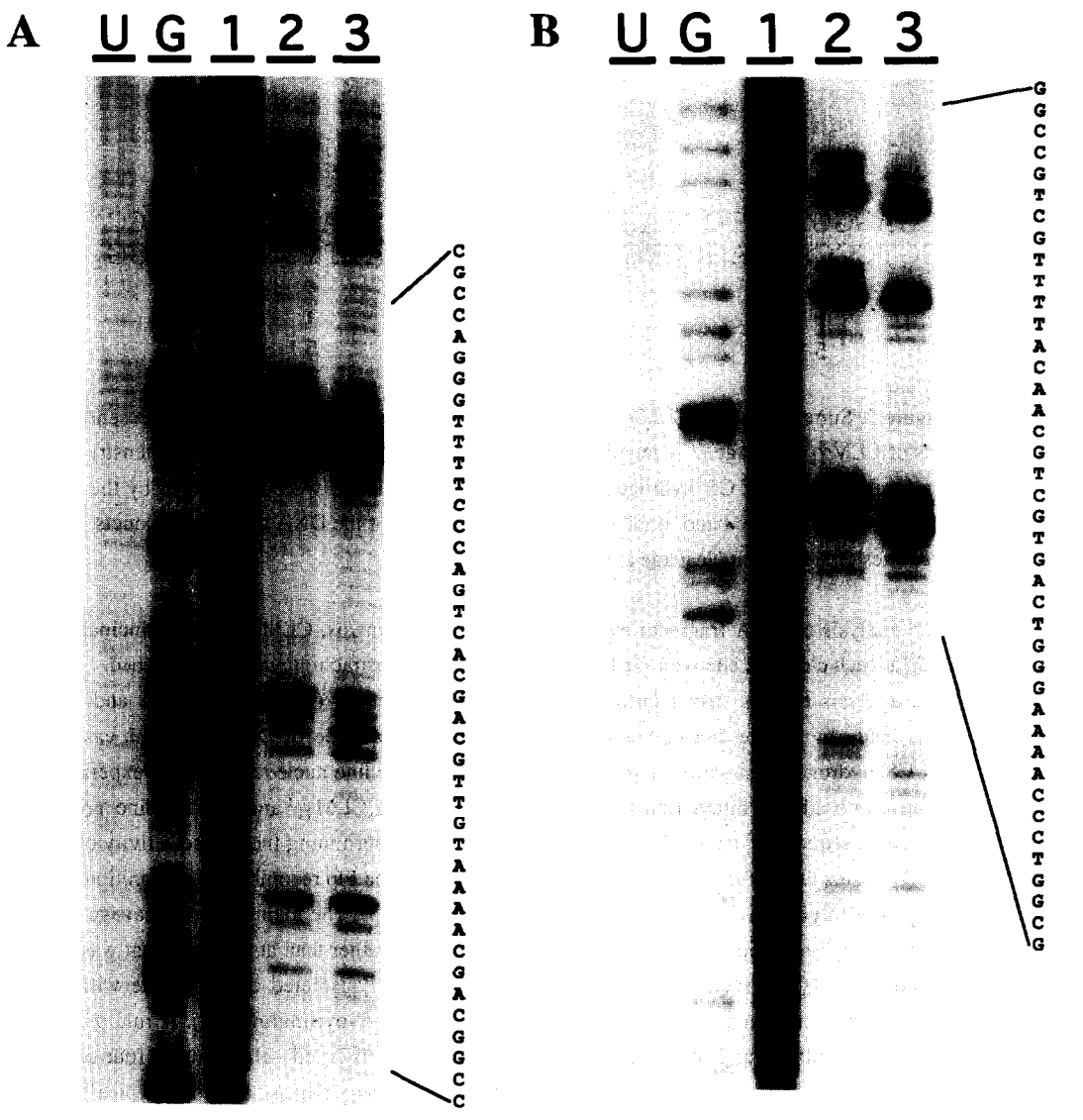


Figure 1. CLMγ_1^I cleavage of the *Eco*RI-*Nar*I DNA restriction fragment from plasmid pUC18. Shown are the autoradiographs of DNA sequencing gels on which the cleavage products were separated. A, the top strand 3'-end labeled at the *Eco*RI site (see Figure 2). B, the bottom strand 3'-end labeled at the *Nar*I site. Lanes U, uncut full-length DNA. Lanes G, products of Maxam-Gilbert guanine-specific DNA sequencing reaction. Lanes 1, products of hydroxyl radical cleavage of this DNA. This reagent attacks at every nucleotide position with no regard to sequence. Lanes 2, products of CLMγ_1^I cleavage reaction. Lanes 3, NaOH-treated DNA fragments from the CLMγ_1^I cleavage reaction. Differences in mobility observed between fragments in lanes 2 and 3 are the result of β-elimination of 5' aldehyde nucleosides. The sequences of the DNA strands are listed to the right, and are read 5' to 3' from top to bottom.

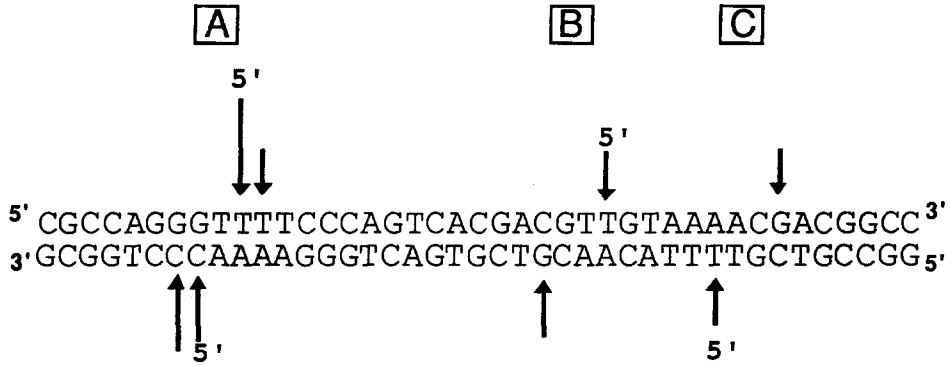


Figure 2. Summary of the CLMγ_1^I cleavage data presented in Figure 1. Arrows represent positions of CLMγ_1^I cleavage. The length of an arrow represents the relative cleavage intensity within one strand. Sites of C5' hydrogen abstraction are labeled, and were determined by the observation of a shift of the band mobility upon base-treatment of the DNA cleavage products. The three sites of significant cleavage are labeled A, B, and C.

footprinting, and analysis of DNA fragment mobility by gel electrophoresis, CLMγ_1^I binds principally in a single orientation at each cleavage site with the locus of C5'-hydrogen abstraction readily established.

The above analysis reveals strong cutting in two TTTT sequences (Figure 2; sites A and C), an observation also made by Kahne and coworkers on the identical sequence of DNA.[3] In both cases, the principal site of C5'-hydrogen abstraction is at the penultimate 5'-pyrimidine nucleotide. This is experimental evidence that the proposed key interaction between the aryliodide of CLMγ_1^I and the 2-amino group of guanines in an AGGA sequence[15] cannot be of major significance in determining the site-selectivity of DNA cleavage by CLMγ_1^I. This fact is further borne out in experiments with the Flp recombinase recognition site of yeast[16] which contains three TTCCT and a TTCTCT sequence (Figure 3). The main sites of cleavage appear again at the penultimate 5'-pyrimidine (dT) in each of these sequences, rather than at the 5' dC that was found in the earlier studies on TCCT sequences. In addition at the relatively isolated cleavage site within the TTCTCT sequence, C5' hydrogen abstraction occurs at the penultimate 5'-pyrimidine (dT). A rule of thumb emerges in these observations—and others, although far from strictly observed—that runs of four or more pyrimidines often coincide with preferred sites of double-stranded CLMγ_1^I-induced cleavage, typically with C5'-hydrogen abstraction occurring most strongly at the second nucleotide in from the 5'-terminus of the pyrimidine tract.

DNA containing tracts of adenines. NMR studies[17-19] and X-ray crystal structures[20,21] of duplex DNA containing short runs of adenines have shown the minor groove of DNA in these sequences to be narrow. In addition, hydroxyl radical cleavage experiments on bent DNA containing short A-tracts of five contiguous adenines demonstrated that the minor groove of these A-tracts is narrow compared to mixed-sequence DNA.[22] The observation that CLMγ_1^I cleaves at short oligo(dA)•oligo(dT) tracts led us to study the CLMγ_1^I cleavage pattern of DNA containing a longer oligo(dA)•oligo(dT) tract.

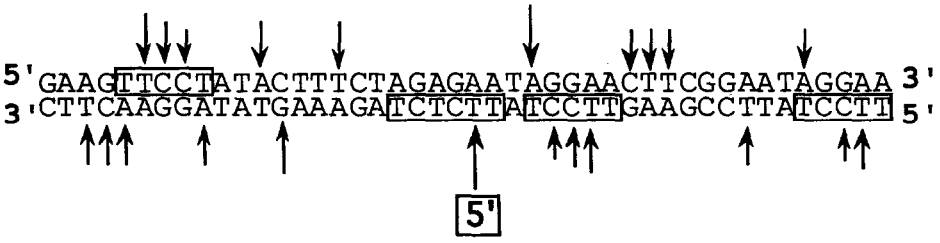


Figure 3. Summary of the CLMγ_1^I cleavage data for a DNA molecule containing several TCCT sequences. CLMγ_1^I cleavage reactions were performed on the *Eco*RI-*Hin*dIII restriction fragment from plasmid pJ3 with the duplex DNA singly end-labeled in turn in four separate experiments. Assignment of the presence of a 5' aldehyde nucleoside resulting from C5' hydrogen abstraction (boxed 5') was made by the observation of a two-nucleotide shift in apparent cleavage position (compared to Maxam-Gilbert DNA sequencing size markers) depending on which end of the DNA strand was labeled. No shift of the apparent cleavage position when comparing the results of 3' vs. 5' data on the same strand indicated the absence of C5' hydrogen abstraction. This method can be used in lieu of or in addition to β-elimination of the 5' aldehyde nucleoside, *via* alkali treatment, to identify the position of C5' hydrogen abstraction. Unfortunately, cleavage sites within the three TTCCT sequences (boxed) overlap with cleavage sites within adjacent CTTC sequences on the complementary strand, making it difficult to determine the sites of C5' hydrogen abstraction in these sequences.

Zein et al. have suggested the width of the minor groove as a possible determinant of the sequence selective cleavage of DNA by CLMγ_1^I.[23] As a test of this hypothesis, we performed CLMγ_1^I cleavage experiments on DNA containing tracts of A_{16}•T_{16}. We expected that CLMγ_1^I would recognize and cleave a region of the A-tract where the width of the minor groove provides an optimal fit for the diradical intermediate. Based on hydroxyl radical cleavage experiments on A_{16}•T_{16} (discussed below), we found that the minor groove of A_{16}•T_{16} presents a graduated range of widths.

Structural studies of long A-tracts. The hydroxyl radical cleavage patterns of longer A-tracts (A_{16} and A_{42}) indicate that fluctuations in the width of the minor groove exist at the 5' and the 3' end of these A-tracts. Figure 4 shows the hydroxyl radical cleavage pattern of DNA containing an A_{16}•T_{16} tract. An abrupt decrease in hydroxyl radical cutting frequency on the A-strand is observed from the flanking sequence to the first adenine at the 5' end (Figure 4A). Progressing into the A-tract, a gradual decrease to a minimum level in cutting frequency occurs from the first adenine to the fourth adenine. The cutting frequency remains at this minimum level throughout most of the A-tract. At the 3' end of the A_{16}-tract, the cutting frequency gradually returns to a maximum level over approximately four nucleotides. This change begins with the 3' terminal adenine, and proceeds through the sequences flanking the 3' end.

A similar, but not identical, hydroxyl radical cleavage pattern is observed for the T_{16}-strand (Figure 4B). At the 3' end of the T-tract (corresponding to the 5' end of the A-tract), an abrupt change in the cutting frequency is observed from the sequence flanking the T-tract to the first thymine where the cutting frequency

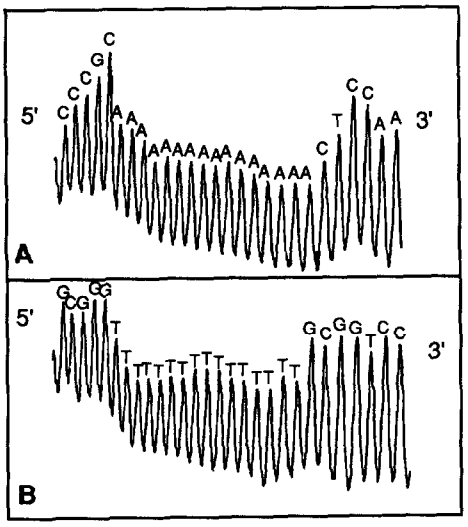


Figure 4. Hydroxyl radical cleavage of DNA containing an A_{16}•T_{16} tract. Shown are the densitometric scans of the autoradiograph of the DNA sequencing gel on which the cleavage products were separated. A, A_{16}-containing strand. B, T_{16}-containing strand.

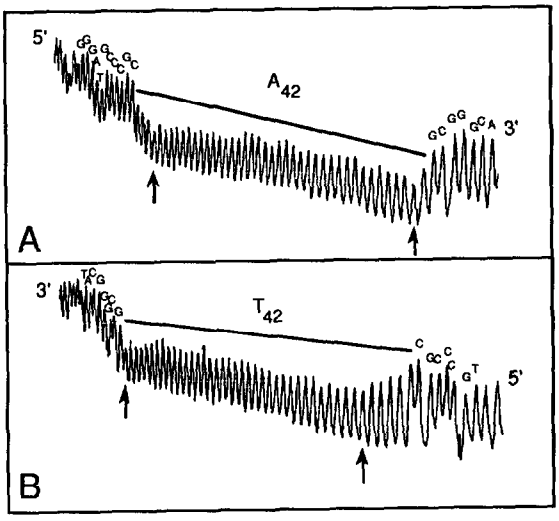


Figure 5. Hydroxyl radical cleavage of DNA containing an A_{42}•T_{42} tract. Shown are the densitometric scans of the autoradiograph of the DNA sequencing gel on which the cleavage products were separated. A, A_{42}-containing strand. B, T_{42}-containing strand. The bars mark the A_{42} and T_{42} tracts. The arrows mark the beginning and the end of poly(dA)•poly(dT)-like structure.

reaches its minimum level as observed for the central region of the A-tract. The cutting frequency remains at the minimum until the 5' end of the T-tract where it returns to a maximum level over two thymines.

The hydroxyl radical cleavage pattern of DNA containing an A_{42}-tract is shown in Figure 5. The patterns are similar to the patterns observed for the A_{16}-tract, although with a few subtle differences. Fluctuations in the hydroxyl radical cutting frequencies are seen at the 5' and the 3' ends, with the central region of the A-tract cut at a constant minimum level. At the 5' end of the A-tract, a gradual transition from the maximum to the minimum cutting frequency begins with the terminal adenine and proceeds through four adenines, as observed for A_{16}. At the 3' end of the A-tract, the transition from the minimum to the maximum cutting frequency begins with the penultimate adenine, unlike in A_{16} where the last adenine marks the minimum, and occurs over approximately four nucleotides. For the complementary region at the 5' end of the T-tract, this transition occurs over approximately six nucleotides, compared to three observed for T_{16}. Overall, the hydroxyl radical cleavage pattern of the T_{42}-strand is offset from that of the A_{42}-strand by approximately three nucleotides to the 3' direction. This offset is indicative that the cleavage experiment is monitoring a feature of the minor groove.[22,24] We proposed earlier[22] that the reason for the decreased cleavage by hydroxyl radical within A-tracts is due to a narrow minor groove in these sequences.

We conclude that regions of the A_{16}-tract and the A_{42}-tract where the hydroxyl radical cleavage frequencies are at a minimum level have adopted a structure similar to that of poly(dA)•poly(dT). A model for the structure of poly(dA)•poly(dT), proposed on the basis of X-ray fiber diffraction data, has as a key feature a double helix with a decidedly narrow minor groove compared to B-form DNA.[25] In addition, our work has shown that there is a smooth transition between the normal width of the minor groove of sequences flanking the A_{16}-tract and the A_{42}-tract and the narrow minor groove exhibited by these A-tracts. This is similar to what was observed previously for shorter A-tracts.[17-22]

Cleavage by CLMγ_1^I of long A-tracts. Figure 6 shows the CLMγ_1^I cleavage pattern of DNA containing the A_{16}-tract with either the A_{16}-strand or the T_{16}-strand radiolabeled. A major site of cleavage is observed two nucleotides from the 3' end of the A-tract and, on the other strand, at the penultimate thymidine at the 5' end of the T-tract. Based on an analysis of the CLMγ_1^I cleavage products (as discussed above), we find that abstraction of a C5' hydrogen occurs at the penultimate thymidine which, according to our model, would position the drug in the direction of the A-tract at that cleavage site. In addition, and in striking contrast to other CLMγ_1^I cleavage sites in which at most a couple of nucleotides are strongly cleaved, a number of weaker cuts also occur which continue towards the 5' end of the A-tract. This set of cuts ends approximately five nucleotides from the 5' end of the A-tract. A summary of the data is provided in Figure 7.

Therefore, two types of CLMγ_1^I cleavage can be distinguished in the A_{16}-tract. The major CLMγ_1^I cleavage site observed at the 3' end is similar in its intensity and its singular nature to other more isolated CLMγ_1^I cleavage sites seen in DNA restriction fragments (for example, see Figure 1). However, the uniform pattern of cleavage that trails into the A-tract from the 3' end is unusual. These cleavages are somewhat less intense than the major cut at the 3' end and other strongly cleaved but more isolated sites. This pattern is consistent with the drug recognizing and cleaving at several overlapping A_4•T_4 sites within the A_{16}-tract, whose structures probably closely resemble one another. This region is likely characterized by high propeller twist that optimizes base-stacking leading to a narrow minor groove.[20,21,26]

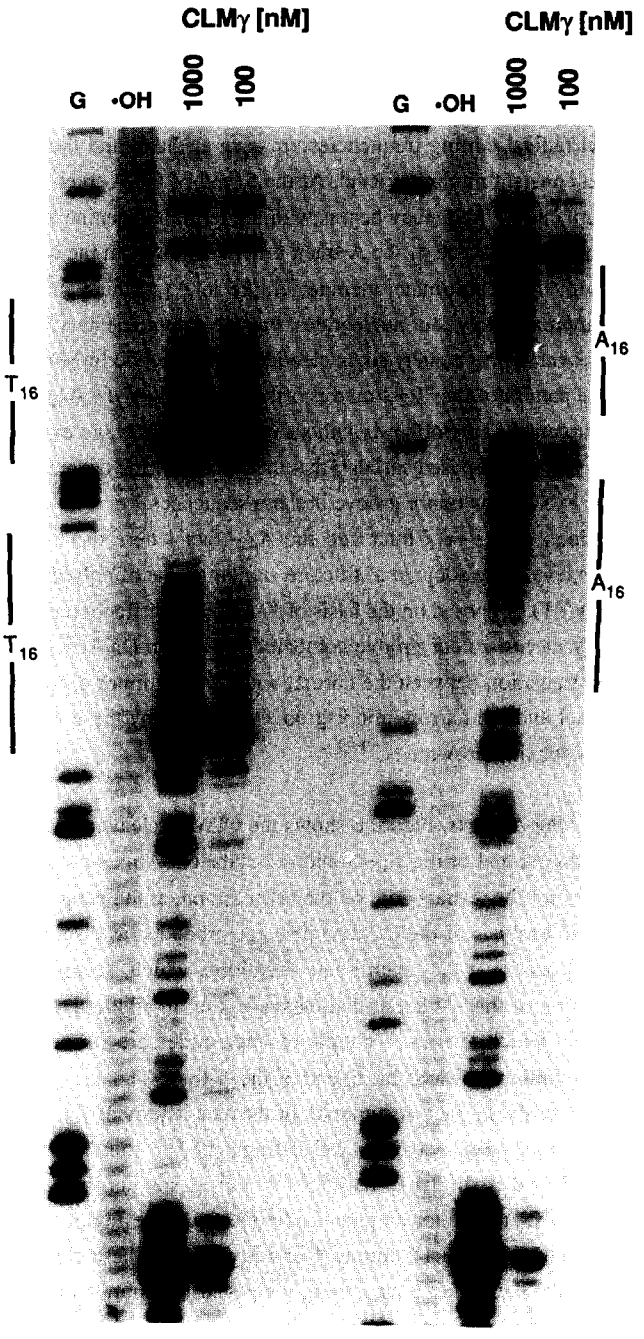


Figure 6. CLMγ_1^I cleavage of DNA containing two A_{16}•T_{16} tracts. Shown is the autoradiograph of a DNA sequencing gel on which the cleavage products of the two A_{16}•T_{16} tracts were separated. The DNA was radiolabeled at the 5' ends, so no 5' aldehyde nucleosides are present. The sequence is read 3' to 5' from top to bottom.

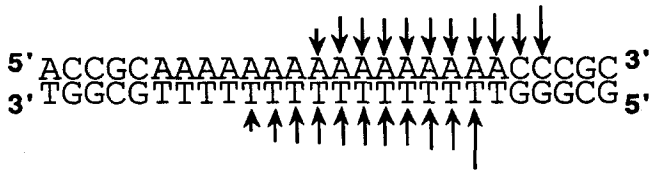


Figure 7. Summary of the CLMγ_1^I cleavage data shown in Figure 6 for one of the A_{16}•T_{16} tracts. The lengths of the arrows represent the relative cleavage intensities within one strand.

Our results on the cleavage of an A_{16}-tract by CLMγ_1^I raised two questions. (1) Does the cascade of weaker cleavage sites terminate close to the 5' end of the A-tract due to structural features at that end? Or (2) do the cleavage sites end after approximately ten nucleotides from the 3' end of the A-tract, regardless of the length of the A-tract? We examined the CLMγ_1^I cleavage pattern of DNA containing an A_{42}-tract to answer these questions. Figure 8 shows that the major CLMγ_1^I cleavage site at the 3' end of the A_{42}-tract is identical in position to the major cleavage site observed for the A_{16}-tract. In addition, a series of cuts extends into the A_{42}-tract, as observed for the A_{16}-tract. However, these cuts extend well beyond the middle of the A_{42}-tract, and terminate at the fourth or fifth nucleotide from the 5' end of the longer A-tract. The CLMγ_1^I cleavage sites on the T-strand terminate closer to the 5'-most thymine. This offset between the two strands is similarly indicative of cleavage across the minor groove.[24]

The results of the CLMγ_1^I cleavage experiments on two long A-tracts show that the drug recognizes an asymmetry inherent in the structure of these A-tracts that favors strong cutting by CLMγ_1^I at the 3' end. Several studies on the structure of A-tracts have demonstrated that the 3' end of these tracts has a structure that deviates from canonical B-DNA. Experiments using diethylpyrocarbonate (DEPC) as a chemical probe of the structure of DNA showed very high reactivity with the penultimate adenine of some A-tracts.[27] This adenine is base-paired with the thymine that is cleaved most avidly by CLMγ_1^I in the A-tracts studied here. It was suggested based on the DEPC experiments that high roll and/or tilt angles occur at the 3' end of an A-tract, providing a possible structural basis for the major CLMγ_1^I cleavage site that we observe at the 3' end. Investigation of the interactions between the antitumor antibiotic CC-1065 and A-tract-containing DNA showed that the 3' end of an A-tract is the site of covalent linkage of CC-1065 with DNA.[28] The stereochemistry of the CC-1065/DNA adduct requires that the drug is directed toward the A-tract, similar to what we see for CLMγ_1^I. Furthermore, experiments examining the bending of DNA caused by A-tracts showed that phasing the 3' ends of A-tracts with the helical repeat of DNA led to greater bending than phasing the 5' ends.[29] This result suggested that the 3' end of an A-tract contributes more to the overall bending of the DNA than the 5' end.

The central region of longer A-tracts was characterized using the hydroxyl radical cleavage method, and was shown to have a structure similar to that of poly(dA)•poly(dT), which contains a narrow minor grove. This entire region of poly(dA)•poly(dT)-like structure is recognized and cleaved by CLMγ_1^I in a uniform manner, albeit with the intensity of the cuts diminished relative to the major cut at the 3' end of the A-tracts studied.

These results allow us to propose a model for the recognition and cleavage of DNA by CLMγ_1^I. We see that CLMγ_1^I cleaves DNA that has a narrow minor groove. Since many studies have suggested that cleavage

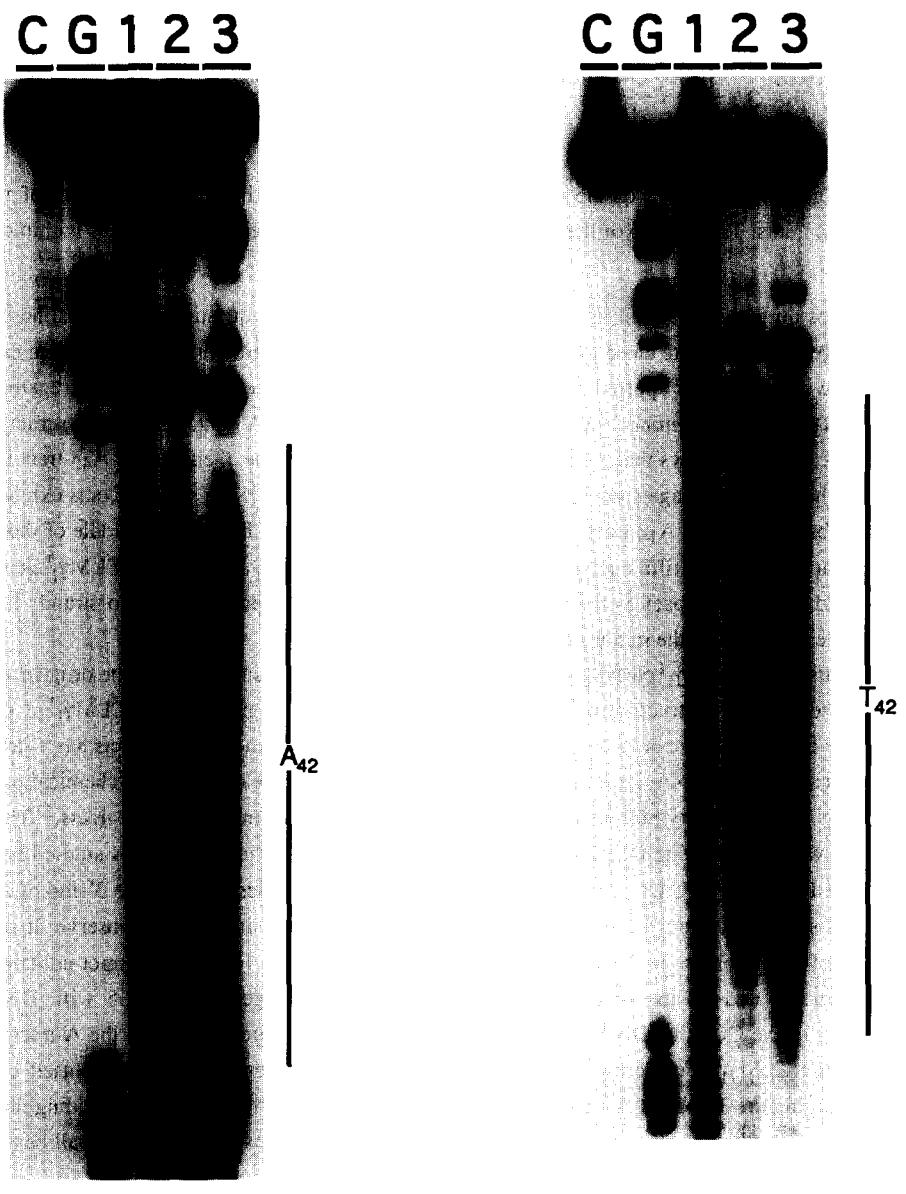


Figure 8. CLM$\gamma_1{}^I$ cleavage of DNA containing an A_{42}•T_{42} tract. Shown are the autoradiographs of DNA sequencing gels on which the cleavage products of the A_{42}•T_{42} tract were separated. The two strands of DNA were 3'-end labeled. The A_{42} tract and the T_{42} tract are marked by the bar to the right of the gel. Lanes C, uncut DNA. Lanes G, products of Maxam-Gilbert G-specific DNA sequencing size markers. Lanes 1, products of hydroxyl radical cleavage reaction. Lanes 2, products of CLM$\gamma_1{}^I$ cleavage of DNA. Lanes 3, NaOH-treated DNA fragments from the CLM$\gamma_1{}^I$ cleavage reaction. The sequences are read 5' to 3' from top to bottom.

of DNA by CLMγ_1^I is mediated mainly by the binding of the drug to DNA,[3,5,30-32] the sequence-dependent structure of DNA appears to dictate to some degree where CLMγ_1^I can bind and cleave. In addition to the series of cleavage sites in the central region of long A-tracts, a more intense cleavage site is seen at the 3' end of A-tracts. This region of the DNA is characterized by a structural discontinuity, which presumably enables CLMγ_1^I to bind to that structure and cleave strongly at one site. Furthermore, the orientation of CLMγ_1^I at that cleavage site positions the drug so that it extends into the A-tract. This suggests that, in addition to a structural discontinuity in DNA, the narrow minor groove of an A-tract may be contributing to the strong binding and intense cleavage at the 3' end.

Thus, several lines of evidence indicate that the junctions between A-tracts and the mixed-sequence DNA flanking the A-tracts at the 5' and 3' ends have different structures. We find that the junction between the 3' end of an A-tract and the flanking B-DNA structure provides a favorable site for CLMγ_1^I cleavage. We conclude that these structural differences, due perhaps to high roll or tilt between the DNA base pairs, and minor groove width, influence the positions of cleavage by CLMγ_1^I.

Cleavage by CLMγ_1^I of Xenopus nucleosomal DNA. The CLMγ_1^I cleavage pattern of DNA containing tracts of oligo(dA)•oligo(dT) demonstrated that the drug recognizes and cleaves sequences of DNA that are bent and have unusual structural features deviating from canonical B-DNA. We next studied the CLMγ_1^I cleavage pattern of the 5S ribosomal RNA gene from *Xenopus borealis*, which has been shown to associate with histones to form a translationally and rotationally uniquely positioned nucleosome.[33,34]

DNA in eukaryotes is bound to histone proteins to form a nucleoprotein complex, chromatin, which consists of basic repeating units called nucleosomes. The nucleosome consists of 146 base pairs of DNA wrapped twice around an octamer of histone proteins in a left-handed superhelix (for a comprehensive review, see ref. 35). It is now well-established that the sequence of nucleosomal DNA facilitates the periodic structural deformations necessary for bending around the histone octamer.[36] DNA is bent more in the nucleosome than it is in sequences containing phased A-tracts. Approximately 75 base pairs of DNA make one complete superhelical turn in the nucleosome, while circles of DNA with phased A-tracts can reach a minimum circumference of only 120 base pairs.[37]

Therefore, based on the unusual CLMγ_1^I cleavage patterns we observed for bent DNA containing A-tracts, we suspected that study of the CLMγ_1^I cleavage of the nucleosomal DNA from the 5S ribosomal RNA gene would add further insight into the structural basis for the recognition and cleavage of DNA by CLMγ_1^I. Another motivation for these experiments is that eukaryotic nucleosomal DNA would be the probable therapeutic target for CLM, so how the drug recognizes and cleaves this DNA is of great interest.

Figures 9 and 10 show the hydroxyl radical cleavage pattern of the *Xenopus* 5S RNA gene reconstituted on the histone core, compared with the cleavage pattern of CLMγ_1^I on the same DNA sequence not bound to the histone proteins. Several features are interesting to note. (1) A strikingly periodic CLMγ_1^I cleavage pattern on the free nucleosomal DNA partly correlates with the sites within the nucleosome that are protected from hydroxyl radical cleavage, most notably from positions +30 to +70. (2) The region from approximately position +10 to -10, around the center of dyad symmetry of the nucleosome, is not cleaved by CLMγ_1^I. (3) An extended homopyrimidine•homopurine tract consisting of fourteen base pairs is cleaved by CLMγ_1^I extensively and predominately at the 5' end of the pyrimidine (corresponding to the 3' end of the purine)

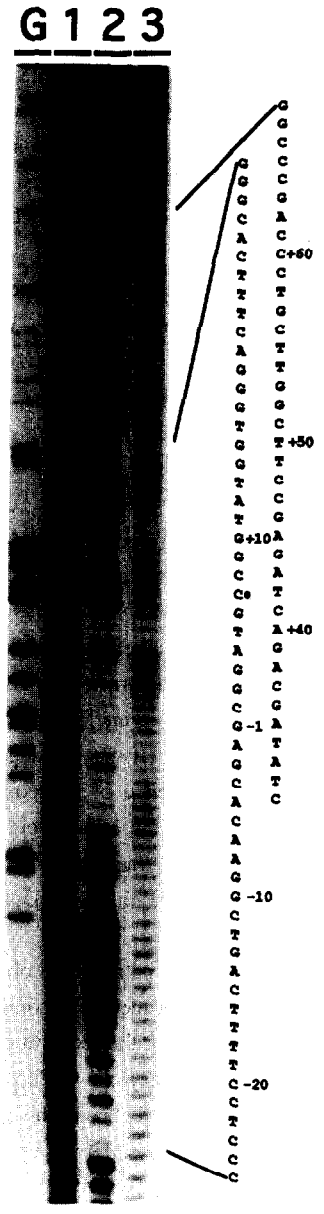


Figure 9. CLMγ$_1^I$ cleavage of the 5S RNA gene from *Xenopus*, and hydroxyl radical footprinting of the 5S DNA from *Xenopus* reconstituted into a nucleosome. Shown is the autoradiograph of a DNA sequencing gel on which the CLMγ$_1^I$ cleavage products and the hydroxyl radical cleavage products were separated. Lane 1, products of hydroxyl radical cleavage of the free 5S RNA gene from *Xenopus*. Lane 2, CLMγ$_1^I$ cleavage of the free 5S RNA gene. Lane 3, hydroxyl radical footprinting of the 5S RNA gene reconstituted into a histone core particle. The sequence is read 5' to 3' from top to bottom. The numbers indicate the positions relative to the start of transcription of the 5S gene. The dot at position +7 marks the approximate position of the center of dyad symmetry of the nucleosome.

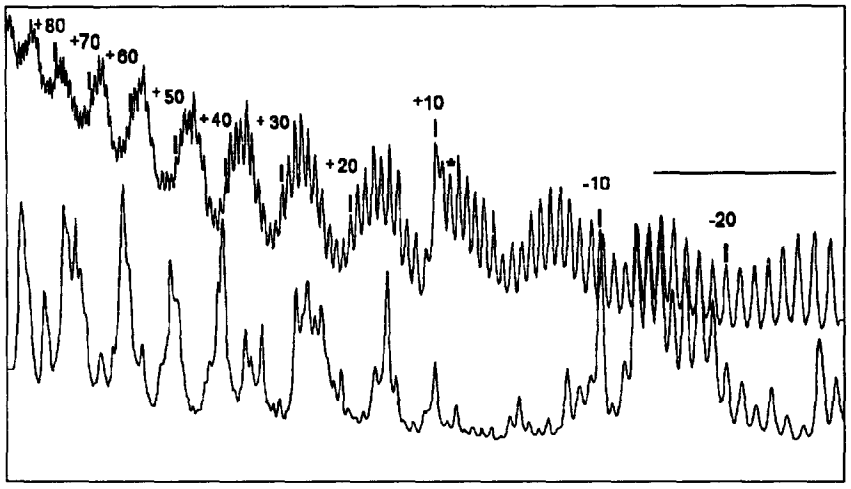


Figure 10. Densitometric scans of the autoradiograph shown in Figure 9. The top scan represents the hydroxyl radical cleavage pattern of the 5S RNA gene reconstituted onto nucleosomes. The data can be found in Figure 9, lane 3. The bottom scan represents the $CLM\gamma_1^I$ cleavage pattern of the 5S RNA gene free of histone proteins (lane 2). The scans are read 5' to 3' from left to right. The approximate position of the center of dyad symmetry of the nucleosome is indicated by an asterisk. The bar over the region from position -15 to -27 marks an oligopurine•oligopyrimidine tract consisting of 14 base pairs. Because of a gel-running artifact, two nucleotides are represented by one peak at position +10.

strand, similar to the $CLM\gamma_1^I$ cleavage pattern observed for A-tracts. Examination of the cleavage sites of $CLM\gamma_1^I$ in nucleosomal DNA shows that most consist of homopurine•homopyrimidine tracts. However, the stretches of pyrimidines are not located entirely on one strand, but rather are distributed between the two strands.

Structure of the 5S RNA gene. Analysis of the helical periodicity of the DNA *in the 5S nucleosome* by hydroxyl radical footprinting[38] showed that the central region around the dyad (positions -17 to +15) has an average helical twist of 10.7 bp per turn.[33] In striking contrast, the helical periodicity of the flanking sequences (positions +14 to +72 and -17 to -75) was found to be different—10.05 bp per turn.[33] The helical periodicity of *free 5S DNA* was found by similar methods to be 10.49 bp per turn. These results showed that the central core region of the 5S RNA gene, when reconstituted onto a histone octamer, is underwound relative to its state as free DNA, and that the flanking sequences are overwound upon nucleosome formation.

The highly periodic part of the $CLM\gamma_1^I$ cleavage pattern of the 5S gene is coincident with the portion of the DNA that becomes overwound upon nucleosome formation (Figure 10). More specifically, between positions +30 and +70, $CLM\gamma_1^I$ cleavage occur at positions in the DNA helix that would be oriented inwards facing the histone core in the nucleosome. Barton and coworkers have examined the structure of the 5S RNA

gene of *Xenopus* using metal complexes that bind to and cleave DNA.[39] Their results suggest that the regions of the 5S DNA that we find to be cleaved by CLMγ_1^I have a more open major groove than canonical B-DNA, which would suggest that the minor groove is more narrow. Sequencing of DNA from nucleosomes demonstrated that A/T-rich regions tend to have the minor groove directed inwards towards the histone core, while G/C-rich regions tend to have the minor groove facing outwards away from the histone core.[40,41] This may reflect the preference of A/T-rich DNA for having a narrow minor groove, or the ability of the minor groove in A/T-rich regions to be compressed.

Sites of distamycin binding to the 5S gene[42] were also found to correlate with the structure of the DNA in the 5S nucleosome.[34] Those sequences bound by distamycin (which binds in the minor groove)[21] corresponded well with the DNA sequences facing inward toward the nucleosome core particle. Thus, our results considered in the light of previous work lead to the conclusion that CLMγ_1^I recognizes DNA sequences that have narrow minor grooves, sequences that contain structural discontinuities deviating from canonical B-form DNA, and sequences that more readily undergo conformational changes to allow compression of the minor groove.

Conclusion

The mobility of DNA fragments on DNA sequencing gels with and without alkali treatment can be used to assign the sites of C5'-hydrogen abstraction by CLMγ_1^I activated for DNA cleavage, as is now well-established.[1,3,5,13] The single orientation per site characteristically seen for the drug is fully in keeping with earlier atom transfer experiments[10,11] and hydroxyl radical footprinting studies.[5] The suggestion that interaction between the aryliodide of CLMγ_1^I and the 2-amino group of guanine residues is the principal determinant of sequence selectivity in cleavage[15] is shown to be of relatively lesser significance. This must be especially so where several sequences with four or more A/T base pairs were found to be efficient DNA cleavage sites. While cutting by CLMγ_1^I is normally quite discrete, being limited to one or a few nucleotides per site, long runs of dT display a large number of cuts. In these and other sequences of four or more pyrimidines, the major cut frequently occurs at the penultimate 5'-pyrimidine. Finally, those sequences of overwound DNA most intimately associated with protein contacts in nucleosome formation correlate to sites of CLMγ_1^I cleavage in the free DNA, while those portions that become underwound in the nucleosome are not recognized and cleaved by CLMγ_1^I.

CLMγ_1^I recognizes and cleaves a variety of nucleotide sequences. Therefore the sequence selectivity of its cleavages cannot arise from specific interactions in the minor groove. The fact that the recently discovered and structurally distinct natural product, kedarcidin, cleaves DNA[43] (albeit mainly in a single-stranded manner) at substantially similar sites as CLMγ_1^I argues that something other than DNA sequence is recognized by these drugs to confer binding and cleavage selectivity. The strong correlation of drug cleavage to nucleosomal DNA structure in the *Xenopus* 5S RNA gene suggests that those dynamic features of DNA that make possible reduced helical periodicity are also responsible for favorable interaction with CLMγ_1^I. The deformation required to achieve DNA winding onto histones may also accommodate efficient binding by CLMγ_1^I and involve common groove narrowing, helix bending or kinking to account for the high sequence-selectivity of double strand cleavage by CLMγ_1^I.

Experimental

Plasmids

Plasmids were purified using a Sephacryl 500 (S-500) size-exclusion column attached to a Pharmacia FPLC (Fast Protein Liquid Chromatography) system. The buffer, consisting of 20 mM Tris•HCl, 2.5 mM EDTA, 0.2 M NaCl (pH 8), was run over the column at a constant flow rate of 0.8 ml/min. Supercoiled plasmid was separated from nicked plasmid on an RPC-5 ion-exchange column.

Plasmids pMP31 and pMP32 each have two A_{16}•T_{16} tracts cloned into the *Eco*RI site of pUC18 in different orientations.[44] Plasmid pJAT42,[45] containing an A_{42}•T_{42} tract inserted into pJW300, was a generous gift from Dr. Stephan Diekmann. Plasmid pJ3,[16] containing the Flp binding site, was kindly provided by Dr. Makkuni Jayaram. Plasmid pXP-10,[33] containing the 5S RNA gene, was a generous donation from Dr. Alan Wolffe.

Radiolabeling of restriction fragments from plasmid DNA

pUC18. The 215 bp *Eco*RI-*Nar*I restriction fragment was labeled with ^{32}P at the 3' end of either the *Eco*RI site or the *Nar*I site using standard procedures.[46]

pJ3. The *Eco*RI-*Hin*dIII fragment was singly end-labeled at all four ends in turn using standard procedures.[46]

pMP31 and pMP32. The A-containing strand of plasmid pMP31 was 5'-radiolabeled at the *Hin*dIII site using standard procedures.[46] The 282 bp *Hin*dIII-*Pvu*II restriction fragment, containing the region of interest, was separated from the 94 bp *Hin*dIII-*Pvu*II fragment by non-denaturing polyacrylamide gel electrophoresis. The identical labeling scheme used for pMP32 will end-label the T-containing strand.

pJAT42. The *Hin*dIII-*Eco*RI fragment was 3' end-labeled at the *Hin*dIII site to label the A-containing strand, and at the *Eco*RI site to label the T-containing strand. Standard protocols were used for labeling the *Eco*RI site,[46] and modified slightly for labeling the *Hin*dIII site. A radioactive [α-^{32}P]dGTP nucleotide was used to label the *Hin*dIII site to prevent incorporation of radioactivity at the *Eco*RI site. Labeling the *Eco*RI site with [α-^{32}P]dATP did not label the *Hin*dIII site, which is cut only after the labeling reaction.

pXP-10. The 212 bp *Eco*RI-*Dde*I fragment was 3' end-labeled at the *Eco*RI site.[33]

CLMγ_1^I cleavage and base-treatment of resulting fragments

CLMγ_1^I cleavage was performed as described by Zein et al.[1] with the following two modifications. (1) CLMγ_1^I, at a final concentration of 100 nM, was used. (2) The reactions were performed at room temperature for 30 min.

The DNA was then precipitated with sodium acetate and ethanol.[46] To treat DNA fragments with base after drying, 100 µl of 1M NaOH was added to the sample which was then placed at 90 °C for 5 minutes.[1] The DNA was again precipitated with sodium acetate and ethanol.

Hydroxyl radical cleavage of DNA

Hydroxyl radical cleavage of free DNA was performed as described by Dixon et al.[47]

Reconstitution of the nucleosome and hydroxyl radical footprinting

Chicken erythrocyte histones were reconstituted by salt/urea dialysis onto a radiolabeled DNA fragment to form nucleosomes.[48-50] Hydroxyl radical footprinting of the nucleosome was performed as described for protein/DNA complexes by Dixon et al.[47]

Sequencing gel analysis

DNA was prepared for sequencing gel electrophoresis by dissolving the samples in a buffer containing deionized formamide and dyes. Samples were heated at 90 °C for 5 minutes to denature the duplex DNA, and stored on ice until loaded onto a denaturing polyacrylamide gel (0.35 cm in thickness; Hoefer Pokerface DNA sequencing gel apparatus). After electrophoresis, the gel was dried and subjected to autoradiography.

Acknowledgments

We thank the Lederle Laboratories (American Cyanamid) for providing us with calicheamicin, Dr. Alan Wolffe for supplying us with histone proteins, and Dr. Sarah Morse and Peter Kebbekus for the synthesis of DNA oligonucleotides. We are grateful to Dr. Jeff Hayes for help with the reconstitution of nucleosomes, and for many stimulating discussions. M. A. P. was a predoctoral fellow of the National Science Foundation. This work was funded by PHS grant CA54421. We gratefully acknowledge the use of densitometry instrumentation maintained by the Institute for Biophysical Research on Macromolecular Assemblies at Johns Hopkins, which is supported by an NSF Biological Research Centers Award (DIR-8721059) and by a grant from the W. M. Keck Foundation.

References and Notes

1. Zein, N.; Sinha, A M.; McGahren, W. J.; Ellestad, G. A. *Science.* **1988**, *240*, 1198-1201.
2. Kishikawa, H.; Jiang, Y.-P.; Goodisman, J.; Dabrowiak, J. C. *J. Am. Chem. Soc.* **1991**, *113*, 5434-5440.
3. Walker, S.; Landovitz, R.; Ding, W. D.; Ellestad, G. A.; Kahne, D. *Proc. Natl. Acad. Sci. USA.* **1992**, *89*, 4608-4612.
4. Dedon, P. C.; Salzberg A. A.; Jinghai, X. *Biochemistry.* **1993**, *32*, 3617-3622.
5. Mah, S. C.; Townsend, C. A.; Tullius, T. D. *Biochemistry.* **1993**, in press.
6. De Voss, J. J.; Hangeland, J. J.; Townsend, C. A. *J. Am. Chem. Soc.* **1990**, *112*, 4554-4556.
7. Cramer, K. D.; Townsend, C. A. *Tetrahedron Lett.* **1991**, *32*, 4635-4638.
8. Chatterjee, M.; Cramer, K. D.; Townsend, C. A. *J. Am. Chem. Soc.* **1993**, *115*, 3374-3375.
9. Zein, N.; McGahren, W. J.; Morton, G. O.; Ashcroft, J.; Ellestad, G. A. *J. Am. Chem. Soc.* **1989**, *111*, 6888-6890.
10. De Voss, J. J.; Townsend, C. A.; Ding, W.-D.; Morton, G. O.; Ellestad, G. A.; Zein, N.; Tabor, A. B.; Schreiber, S. L. *J. Am. Chem. Soc.* **1990**, *112*, 9669-9670.
11. Hangeland, J. J.; De Voss, J. J.; Heath, J. A.; Townsend, C. A.; Ding, W.-D.; Ashcroft, J.; Ellestad, G. A. *J. Am. Chem. Soc.* **1992**, *114*, 9200-9202.
12. Walker, S.; Valentine, K. G.; Kahne, D. *J. Am. Chem. Soc.* **1990**, *112*, 6428-6429.
13. Kappen, L. S., & Goldberg, I. H. *Biochemistry.* **1983**, *22*, 4872-4881.

14. Christner, D. F.; Frank, B. L.; Kozarich, J. W.; Stubbe, J.; Golik, J.; Doyle, T. W.; Rosenberg, I. E.; Krishnan, B. *J. Am. Chem. Soc.* **1992**, *114*, 8763-8767.
15. Hawley, R. C.; Kiessling, L. L.; Schreiber, S. L. *Proc. Natl. Acad. Sci. USA.* **1989**, *86*, 1105-1109.
16. Jayaram, M. *Proc. Natl. Acad. Sci. USA.* **1985**, *82*, 5875.
17. Behling, R. W.; Kearns, D. R. *Biochemistry.* **1986**, *25*, 3335-3346.
18. Kintanar, A.; Klevit, R. E.; Reid, B. R. *Nucl. Acids Res.* **1987**, *15*, 5845-5862.
19. Katahira, M.; Sugeta, H.; Kyogoku, Y.; Fujii, S.; Rujisawa, R.; Tomita, K.-I. *Nucl. Acids Res.* **1988**, *16*, 8619-8632.
20. Nelson, H. C. M.; Finch, J. T.; Luisi, B. F.; Klug, A. *Nature.* **1987**, *330*, 221-226.
21. Coll, M.; Frederick, C. A.; Wang, A. H.-J.; Rich, A. *Proc. Natl. Acad. Sci. USA.* **1987**, *84*, 8385-8389.
22. Burkhoff, A. M.; Tullius, T. D. *Cell.* **1987**, *48*, 935-943.
23. Zein, N.; Poncin, M.; Nilakantan, R.; Ellestad, G. A. *Science.* **1989**, *244*, 697-699.
24. Drew, H. R.; Travers, A. A. *Cell.* **1984**, *37*, 491-502.
25. Alexeev, D.; Lipanov, A.; Skuratovskii, I. *Nature.* **1987**, *325*, 821-823.
26. Baikalov, I.; Grzeskowiak, K.;Yanagi, K.; Quintana, J.; Dickerson, R. E. *J. Mol. Biol.* **1993**, *231*, 768-784.
27. McCarthy, J. G.; Williams, L. D.; Rich, A. *Biochemistry.* **1990**, *29*, 6071-6081.
28. Reynolds, V. L.; Molineux, I. J.; Kaplan, D. J.; Swenson, D. H.; Hurley, L. H. *Biochemistry.* **1985**, *24*, 6228-6237.
29. Koo, H.-S.; Wu, H.-M.; Crothers, D. M. *Nature.* **1986**, *320*, 501-506.
30. Drak, J.; Iwasawa, N.; Danishefsky, S.; Crothers, D. M. *Proc. Natl. Acad. Sci. USA.* **1991**, *88*, 7464-7468.
31. Aiyar, J.; Danishefsky, S. J.; Crothers, D. M. *J. Am. Chem. Soc.* **1992**, *114*, 7552-7554.
32. Nicolaou, K. C.; Tsay, S.-C.; Suzuki, T.; Joyce, G. F. *J. Am. Chem. Soc.* **1992**, *114*, 7555-7557.
33. Hayes, J. J.; Tullius, T. D.; Wolffe, A. P. *Proc. Natl. Acad. Sci. USA.* **1990**, *87*, 7405-7409.
34. Rhodes, D. *EMBO J.* **1985**, *4*, 3473-3482.
35. van Holde, K.E. *Chromatin;* Springer-Verlag: New York. 1989.
36. Travers, A. A. *Annu. Rev. Biochem.* **1989**, *58*, 427-452.
37. Ulanovsky, L.; Bodner, M.; Trifonov, E. N.; Choder, M. *Proc. Natl. Acad. Sci. USA.* **1986**, *83*, 862-866.
38. Tullius, T. D.; Dombroski, B. A. *Science.* **1985**, *230*, 679-681.
39. Huber, P. W.; Morii, T.; Mei, H.-Y.; Barton, J. K. *Proc. Natl. Acad. Sci. USA.* **1991**, *88*, 10801-10805.
40. Drew, H. R.; Travers, A. A. *J. Mol. Biol.* **1985**, *186*, 773-709.
41. Drew, H. R.; Travers, A. A. *J. Mol. Biol.* **1986**, *191*, 659-675.
42. Churchill, M. E. A., Hayes, J. J., & Tullius, T. *Biochemistry.* **1990**, *29*, 6043-6050.
43. Zein, N.; Colson, K. L.; Leet, J. E.; Schroeder, D. R.; Solomon, W.; Doyle, T. W.; Casazza, A. M. *Proc. Natl. Acad. Sci. USA.* **1993**, *90*, 2822-2826.
44. Price, M. A. *Structural Studies of Adenine Tracts in DNA: Influence of Length and Sequence Context,* The Johns Hopkins University 1992.
45. Diekmann, S. *FEBS Lett.* **1986**, *195*, 53-56.

46. Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY. 1989.
47. Dixon, W. J.; Hayes, J. J.; Levin J. R.; Weidner, M. F.; Dombroski, B. A.; Tullius, T. D. *Meth. Enzymol.* **1991**, *208*, 380-413.
48. Camerini-Otero, R. D.; Sollner-Webb, B.; Felsenfeld, G. *Cell.* **1976**, *8*, 333-347.
49. Simon, R. H.; Felsenfeld, G. *Nucleic Acids Res.* **1979**, *6*, 689-696.
50. Stein, A. *J. Mol. Biol.* **1979**, *130*, 103-134.

Note added in proof: Walker and Kahne have reported NMR spectroscopic results recently (*J. Am. Chem. Soc.* **1993**, *115*, 7954-7961) for an octamer containing the sequence ACCT in which significant deformations of the oligonucleotide were observed on calicheamicin γ_1^I binding. The authors suggest that the specificity of binding may correlate to the ability of certain sequences to distort to accomodate the drug.

(*Received* 7 *October* 1993; *accepted* 24 *November* 1993)