Pergamon

Bioorganic & Medicinal Chemistry, Vol. 3, No. 6, pp. 729–741, 1995

0968-0896/95 $9.50 + .00

0968-0896(95)00054-2

New Insights into Calicheamicin–DNA Interactions Derived from a Model Nucleosome System

Li Yu, Aaron A. Salzberg and Peter C. Dedon*

Division of Toxicology, 16-336B, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Abstract—Using the *Xenopus borealis* 5S RNA gene, we have identified several new features of the interaction of calicheamicin (CAL), an enediyne antitumor agent, with nucleosomal and naked DNA targets. CAL-mediated DNA damage was generally reduced by incorporation of the DNA into a nucleosome. However, in one instance, the frequency of DNA damage was enhanced in the nucleosome compared to naked DNA. This increase in CAL damage may result from bending-induced changes in the target site, while the association of histone proteins with DNA in the nucleosome may generally reduce the affinity of CAL for its targets by imposing dynamic constraints on the DNA, by altering target structure, or by steric hindrance. One implication of these observations is that new structural features created by incorporation of DNA into chromatin may produce 'hot spots' for CAL-mediated DNA damage not apparent in naked DNA studies. In a second set of experiments, the orientation of CAL at damage sites in naked 5S rDNA was determined. The results suggest that minor groove width *per se* is not a major determinant of CAL target selection. Our studies support the generality of an oligopurine recognition element, with the additional requirement that the purine tract is interrupted at the 3'-end by a pyrimidine(s). To account for these observations, we propose a model in which CAL recognizes the unique structural and dynamic features associated with the 3'-end of an oligopurine tract. Finally, we conclude that the dyad axis of pseudosymmetry of the 5S rRNA gene nucleosome cannot be determined with any degree of certainty. This places significant limitations on the interpretation of results from the study of drug–DNA interactions with reconstituted nucleosomes.

Introduction

The enediyne antitumor agent calicheamicin (CAL) has received significant attention both for its potency as a cytotoxic agent[1,2] and for the chemistry of its DNA damage.[3,4] The bulk of this attention has been focused on the reaction of CAL with isolated DNA fragments. However, the true target of DNA-damaging agents *in vivo* is the complex of proteins and DNA that comprises chromatin. In this report, we have examined the reaction of CAL with the basic element of chromatin structure, the nucleosome.

Like other members of the enediyne family, CAL produces DNA damage via a diradical intermediate that, when positioned in the minor groove, abstracts deoxyribose hydrogen atoms (Fig. 1).[3,4] The sugar radicals undergo oxygen-dependent reactions to form abasic sites or strand breaks, with the final products specific to each position in the deoxyribose. Single molecules of CAL produce predominantly double-stranded lesions, with damage sites on each strand staggered by 3 base pairs (bp) in a 3'-direction.[2,5,6]

There has been recent and significant progress in defining the elements of the CAL target recognition process. It is now generally accepted that the tetrapurine sequence selectivity displayed by CAL mainly reflects recognition of sequence-dependent local DNA conformation and dynamics rather than a direct reading of base-specific information.[2,7–17] Isotope transfer studies, hydroxyl radical footprinting, and NMR solution structures of the CAL–DNA complex all suggest that CAL binds to target sequences with its extended carbohydrate tail directed toward the 5'-end of the tetrapurine strand.[10,13,15,16,18,19] The tail, which appears to adopt a right-handed screw that complements that of the minor groove,[13,15,16,20] is responsible for the sequence-selectivity and is a major contributor to the binding energetics.[8,9,12,21]

The dynamics of CAL and its DNA binding sites may also play a role in target selection. According to NMR studies, the relatively rigid CAL carbohydrate side chain[22] is proposed to induce a widening of the minor groove along the purine·pyrimidine tract.[13,15,16] In addition, there is NMR evidence for two orientations of the aglycone at the binding site,[13] suggesting an element of hinge-like flexibility. An altered binding site structure is also supported by the apparent CAL-induced helical overwinding observed in circular dichroism studies by Ellestad and coworkers,[23] and the ability of CAL to increase the apparent negative superhelical density of plasmid DNA (L. Yu and P. Dedon, unpublished observations).

Given the importance of DNA dynamics and conformation in CAL activity, the structural perturbations associated with incorporation of DNA into chromatin add another dimension to the complexity of CAL target recognition. The nucleosome, one of the basic elements of chromatin structure, is one example

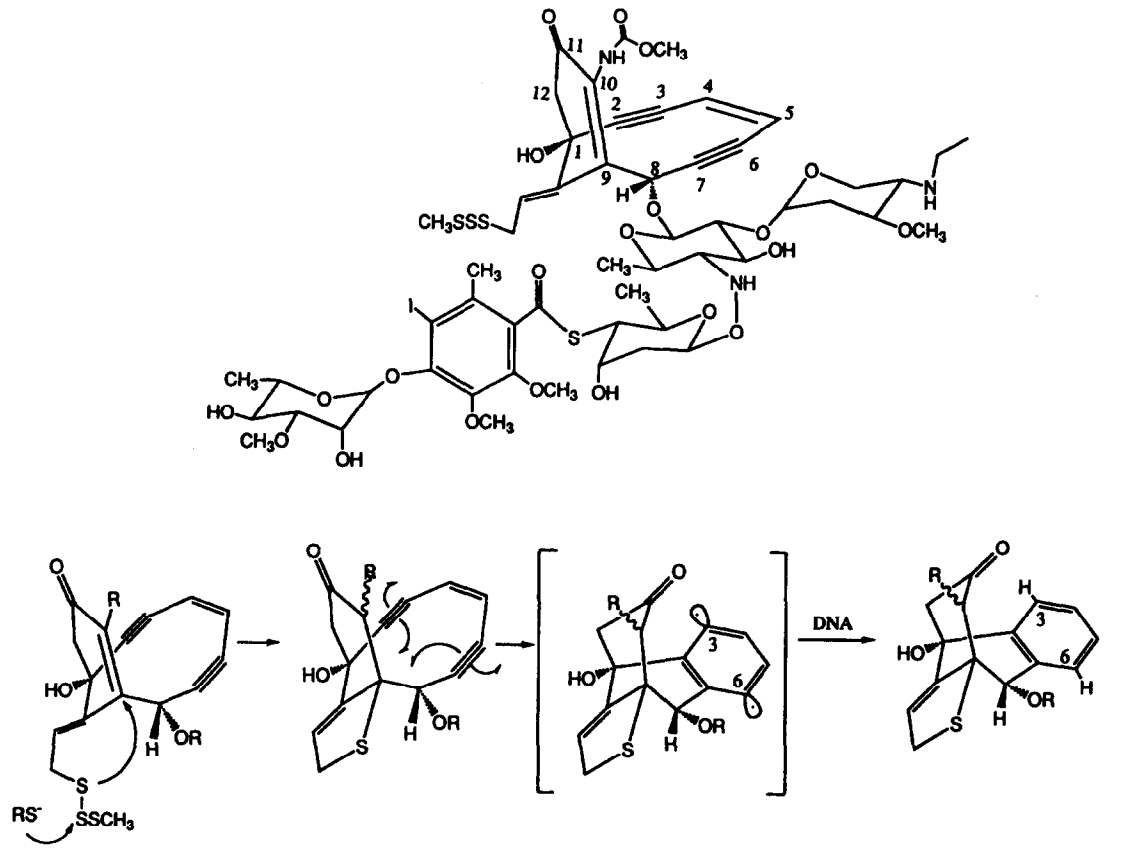

Figure 1. Calicheamicin structure and activation.

of altered DNA conformation in chromatin. The nucleosome consists of essentially two regions, core and linker.[24,25] The core is composed of ~146 base pairs (bp) of DNA wrapped ~1.8 times in a left-handed superhelix around four pairs of histone proteins, and the linker represents the 20–60 bp of DNA joining adjacent cores. In addition to the bending-induced changes in minor groove width, there are several features of DNA structure in the nucleosome core that could affect the interaction of CAL with DNA. These include: (1) both helical underwinding[26] and an S-shaped bend in the ~30 bp of core DNA surrounding a dyad axis of pseudosymmetry passing through the center of the core DNA;[27] (2) helical overwinding of the remaining ~120 bp of core DNA;[26] and (3) kinking or sharp bending of the core DNA at 4 sites symmetrically-positioned at one and four helical turns from the dyad axis.[27]

We have previously demonstrated that CAL is capable of damaging both nucleosome core and linker DNA in nuclei and isolated core particles, with cleavage sites positioned where the minor groove faces away from the histone core.[28] Unlike intercalating agents, such as the enediyne esperamicin A1,[28–30] CAL is capable of binding to the bent and dynamically-constrained DNA of the nucleosome core.

In this work, we present studies of CAL-mediated DNA damage in both nucleosomal and naked DNA. As a model nucleosome system, we have chosen the uniquely-positioned and well-characterized nucleosome that forms on the 5S rDNA of *Xenopus borealis*.[26,31] A preliminary study of CAL-mediated damage in naked 5S rDNA has been reported by Mah *et al.*[11] We have identified three important features of CAL target selection: (1) nucleosomal-induced enhancement of cleavage at a unique site; (2) no apparent relationship between minor groove width and CAL target selection; and (3) a general target consensus sequence that appears to involve disruption of oligopurine tracts by pyrimidines. Based on these and previous observations, we propose a new model for CAL damage site structure.

Experimental

Materials

The plasmid pXP-10, containing the *Xenopus borealis* 5S RNA gene, was kindly provided by Dr Alan Wolffe (NIH/NICHHD).[32] CAL was generously provided by Dr George Ellestad (Wyeth–Ayerst Research). Citrate-preserved chicken blood was purchased from Rockland, Inc.

Labeling of 5S rDNA

Uniquely end-labeled DNA fragments were obtained from the pXP-10 construct by a combination of restriction enzyme digestions and [^{32}P] end-labeling, followed by nondenaturing polyacrylamide gel electrophoretic purification of the desired fragment.[33] Labeling of the 5'-ends with [^{32}P] was accomplished

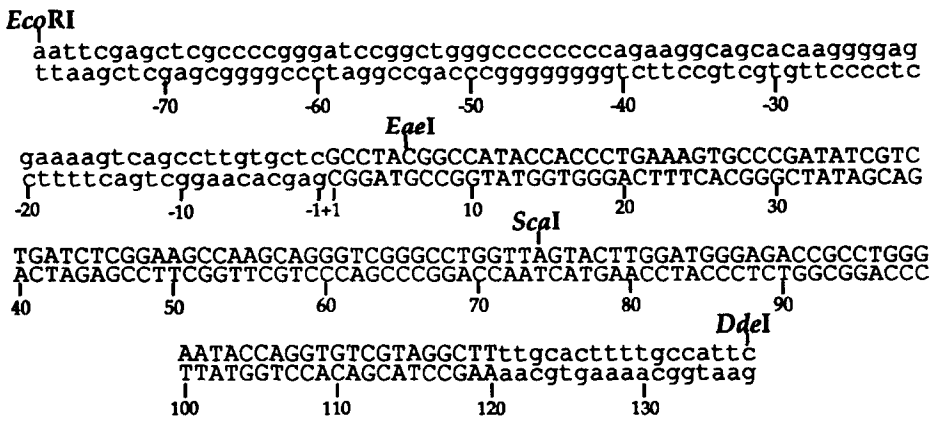

Figure 2. Sequence of the *Eco*RI/*Dde*I fragment of pXP-10 containing the 5S RNA gene. The sequence is numbered relative to the start of transcription of the 5S RNA gene (+1).

with [γ-^{32}P] ATP (Amersham) and T4 polynucleotide kinase (New England Biolabs), and 3'-ends were labeled with [α-^{32}P] ddATP (5,000 Ci/mmol, Amersham) using the Klenow fragment or terminal deoxynucleotidyl transferase (United States Biochemical).[33] The following fragments were prepared with labels at either the 3'- or 5'-ends: (1) the 215 bp *Eco*RI/*Dde*I fragment, labeled at either the *Eco*RI or *Dde*I site; (2) the 131 bp *Eae*I/*Dde*I fragment labeled at the *Eae*I site; and (3) the 68 bp *Sca*I/*Eae*I fragment, labeled at the *Sca*I site. The location of these restriction enzyme cleavage sites is shown in the context of the entire *Eco*RI/*Dde*I region of pXP-10 in Figure 2.

Reconstitution of nucleosomes on 5S rDNA

Nucleosome core particles from chicken erythrocytes were isolated as described previously.[28] The 215 bp *Eco*RI/*Dde*I fragment of pXP-10, labeled at either the 3'- or 5'-end of the *Eco*RI site or the 5'-end of the *Dde*I site, was reconstituted into nucleosomes by exchange of histone cores from isolated chicken erythrocyte core particles in urea/salt gradients, as described elsewhere.[26] The quality of the reconstitution was monitored by 0.7% agarose gel electrophoresis of nucleoprotein species, with a running buffer of 45 mM Tris, 45 mM sodium borate, 2.5 mM EDTA, pH 8.3. In all cases, more than 95% of the DNA was incorporated into nucleosomes (data not shown).

CAL damage reactions and hydroxyl radical footprinting with Fe/EDTA

Either naked or nucleosomal 215 bp *Eco*RI/*Dde*I DNA (30 or 60 μg mL^{-1}) was treated with CAL (40 or 80 nM) in 10 mM glutathione, 15 mM HEPES, 1 mM EDTA, pH 7, at 37°C for 15 min followed by ethanol precipitation. The histone proteins in nucleosomal DNA were then removed by adding SDS to 0.5% followed by extensive extraction with phenol:$CHCl_3$ (1:1) and isoamyl-alcohol:$CHCl_3$ (1:24); naked DNA was treated in the same manner. The abasic sites produced by CAL were converted to strand breaks by treatment in 0.1 M putrescine at 37°C for 1 h.[6] The damaged DNA was purified again by ethanol precipitation.

In the study of drug orientation, both 3'- and 5'-[^{32}P] labeled 68 bp *Sca*I/*Eae*I fragments were treated with CAL as above. Each reaction mixture was then divided into three aliquots, one of which was kept as a control. For 5'-[^{32}P] labeled samples, the other two aliquots were further treated for 1 h with either 0.1 M hydrazine at room temperature or 0.1 M putrescine at 37°C. For 3'-[^{32}P] labeled samples, one of the two remaining aliquots was treated twice with 50 mM sodium borohydride in 0.5 M HEPES, pH 6.5 at 0°C for 15 min, while the other aliquot was treated with 0.1 M piperidine at 90°C for 30 min followed by lyophilization. Finally, ethanol precipitation in 0.3 M sodium acetate (pH 7.0) was carried out for all DNA samples.

Hydroxyl radical damage studies were performed with naked and nucleosomal DNA as described elsewhere.[34]

Sequencing gel analysis

Equal quantities of radioactivity for all samples were loaded on either 8% (for 215 bp *Eco*RI/*Dde*I fragments) or 15% (for 68 bp *Sca*I/*Eae*I fragments) sequencing gels, along with Maxam–Gilbert chemical sequencing standards.[35] Dried gels were subjected to Phosphorimager analysis.

Results

The 5S rDNA reconstituted nucleosome

To study DNA target recognition by CAL in nucleosomes, we conducted experiments with the uniquely-positioned nucleosome that forms on the 215 bp *Eco*RI/*Dde*I sequence of plasmid pXP-10;[26] the nucleosome-forming fragment contains the 5S RNA gene and a portion of vector at the 5' end.[26] The initial series of experiments involved definition of the structure of the 5S rDNA nucleosomal DNA as a foundation for interpretation of results with CAL.

One of the major structural reference points of the nucleosome is the 'dyad'. This axis of pseudosymmetry passes roughly through the center of the ~146 bp of the

nucleosome core DNA,[24,25] where the minor groove faces away from the histone proteins.[27] The ~30 bp of DNA encompassing the dyad takes an S-shaped jog[27] and is underwound relative to B-DNA, with a helical repeat of ~10.7 bp/turn.[26] The remaining core DNA is overwound at ~10.0 bp/turn.[26] There are also four symmetrically-placed sharp bends or kinks in the core DNA located one and four helical turns away from the dyad on either side.[27]

The precise location of the dyad has not been established with certainty in nucleosomes reconstituted on the 5S rDNA-containing *Eco*RI/*Dde*I fragment.* The hydroxyl radical studies of Hayes *et al.* suggest a dyad location near the –2 position, with the sequence numbered relative to the transcription start site at +1.[26] However, Pruss and Wolffe[38] and Mah *et al.*[11] have assigned the dyad one helical turn away near positions +7 to +9. There are two possible explanations for these ambiguous assignments of the dyad: (1) the translational position of the DNA on the histone core may be sensitive to the reconstitution technique; and (2) technical limitations and a small data set prevent a precise localization.

We have attempted to define the location of the dyad using two strategies, one involving the same waveform analysis adopted by Hayes *et al.*[26] We observed a phase shift of 50–100° between sine functions fit to the hydroxyl radical cleavage patterns for sequences flanking the ~30 bp surrounding dyad assignments at –2 or +10 (data not shown). This degree of phase shift correlates well with the ~68° difference observed by Hayes *et al.*,[26] which is expected if the ~30 bp surrounding the dyad has a helical repeat ~0.2 bp/turn larger than the remaining DNA.[26] Furthermore, we found that the helical repeat of the ~30 bp surrounding either dyad position was >10.6 and that of the remaining DNA was <10.3, again consistent with the determinations of Hayes *et al.*[26] (data not shown). The other methods of assigning the dyad, micrococcal nuclease trimming or simply counting the DNase I or hydroxyl radical cleavage maxima, also led to ambiguity in the location of the dyad (data not shown).

Regardless of the method of analysis, our results were similar for dyad assignments at either –2 or +10. We are thus unable to assign the location of the dyad with any degree of certainty, and therefore assume that the dyad axis lies near either the –2 or +10 position. Given this uncertainty, we have approached the interpretation of subsequent CAL damage studies in the 5S rDNA system with caution, avoiding arguments based on the various features of nucleosome structure, such as the sharp bends present at positions ±1 and ±4 relative to the dyad.[27]

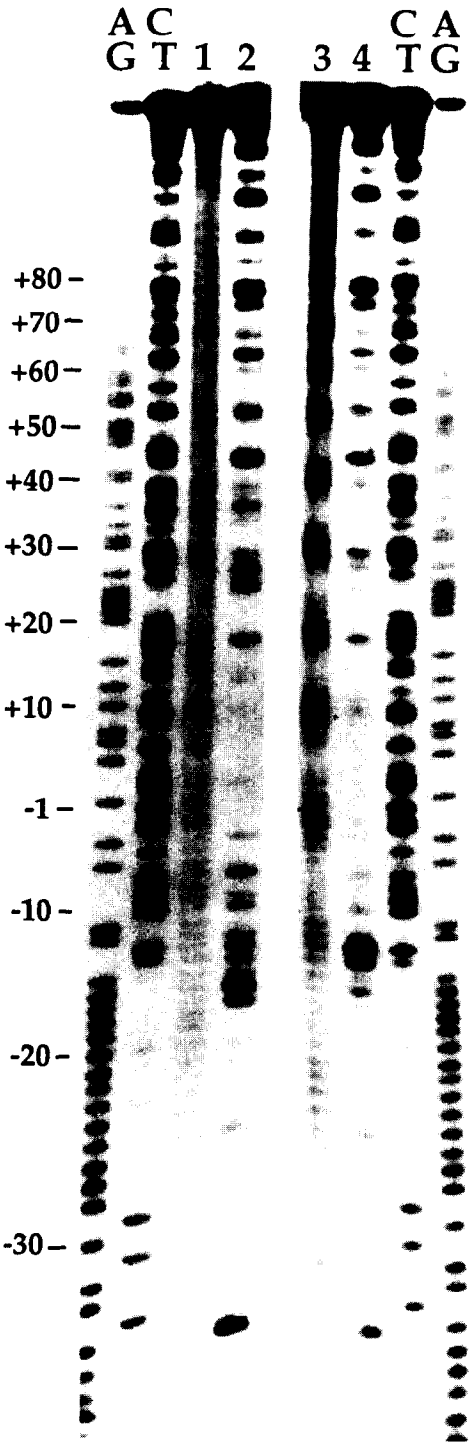

Figure 3. DNA damage produced by calicheamicin and hydroxyl radical in the noncoding strand of nucleosomal and naked 5S rDNA. Naked (lanes 1, 2) and nucleosomal (lanes 3, 4) forms of the *Eco*RI/*Dde*I fragment of plasmid pXP-10, labeled at the 5'-end of the *Eco*RI site with [^{32}P], were treated with Fe/EDTA (lanes 1, 3) or calicheamicin (lanes 2, 4; 40 μM). The DNA was resolved on an 8% sequencing gel. A/G and C/T denote Maxam–Gilbert sequencing reactions.[35] The sequence is read 5'-to-3' from bottom to top.

*The discussion of translational positioning of the nucleosome is limited to the *Eco*RI/*Dde*I fragment of pXP-10, containing the 5S rRNA gene of *X. borealis*. There are species and flanking sequence effects that may alter the translational position in other systems.[36,37] The reconstituted nucleosome studied by Rhodes[31] was prepared with the 236 bp *Hpa*II/*Dde*I fragment of the pXbs 1 construct, which differs both in length and sequence at the 5'-end of the noncoding strand from the present construction.

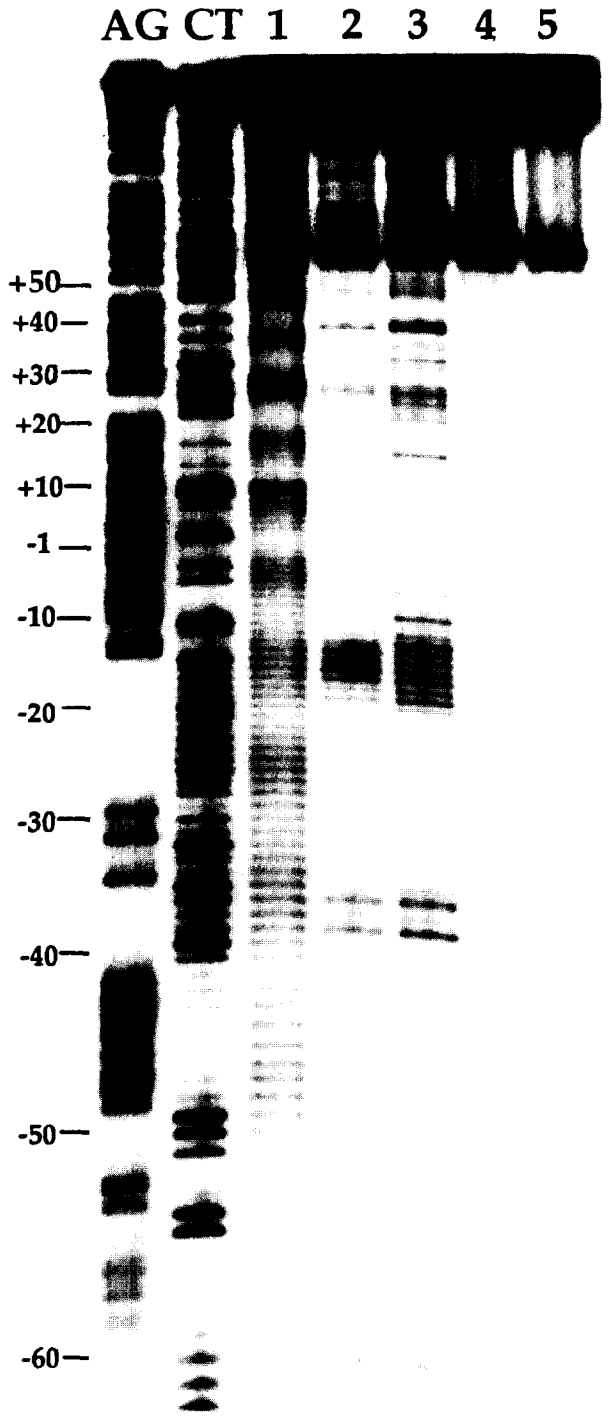

Figure 4. DNA damage produced by calicheamicin and hydroxyl radical in the coding strand of nucleosomal and naked 5S rDNA. Naked (lanes 3, 5) and nucleosomal (lanes 1, 2, 4) forms of the *Eco*RI/*Dde*I fragment of plasmid pXP-10, labeled at the 3'-end of the *Eco*RI site with [^{32}P], were treated with Fe/EDTA (lane 1) or calicheamicin (lanes 2, 3; 40 nM). The DNA was resolved on an 8% sequencing gel. A/G and C/T denote Maxam–Gilbert sequencing reactions.[35] The sequence is read 5'-to-3' from top to bottom.

Although we cannot precisely define the translational setting of the nucleosomal DNA, we observed the predominance of a single rotational orientation of the core DNA, as indicated by the hydroxyl radical cleavage patterns shown in Figures 3 and 4. While cleavage of the naked DNA by the Fe/EDTA reagent produces a seemingly monotonous pattern of fragmentation (Figs 3 and 4), a sinusoidal variation in cleavage frequency is apparent for the nucleosomal DNA (Figs 3 and 4). The cleavage maxima and minima in the nucleosomal DNA occur when the minor groove faces away from and toward the histone core, respectively.[26] This conclusion is supported by the minor groove selectivity of the Fe/EDTA reagent[34] and the correlation of the hydroxyl radical patterns with the cleavage maxima produced by the minor groove specific DNase I (data not shown).

CAL damage sites in naked and nucleosomal 5S rDNA

Our previous studies demonstrated that CAL is capable of producing double-stranded (DS) DNA damage in the core DNA of isolated 'random sequence' nucleosomes.[28] To investigate the effect of nucleosome-altered DNA structure and dynamics on CAL target recognition, we undertook studies of CAL damage in the 5S rDNA system.

The 5S rDNA was labeled with [^{32}P] at the 3'- or 5'-end of the *Eco*RI site, and the naked or nucleosomal DNA treated with 40 nM CAL in the presence of glutathione as the activating thiol. This low concentration of CAL resulted in a drug-to-DNA (bp) ratio of ~ 1: 1200, which limited the damage to essentially 'single-hit kinetics', and prevented multiple CAL binding events from significantly altering the rotational setting of the DNA in the nucleosome and exposing otherwise inaccessible damage sites to further attack.

The damaged DNA was resolved on sequencing gels, as shown in the representative autoradiographs in Figures 3 and 4. In all experiments, several gel runs were performed for each damage reaction to permit single nucleotide resolution along the entire 5S rDNA construct sequence (data not shown). The cleavage patterns were then compared by Phosphorimager analysis, as shown in Figures 5 and 6 for the gels run in Figures 3 and 4, respectively. A comparison of the CAL damage patterns in the 5S rDNA reveals a 3 bp stagger between cleavage sites on opposite strands, as expected for a minor groove selective agent; this is most apparent in Figure 8.

One notable feature of the CAL cleavage profiles in Figures 5 and 6 is the similarity between the location of damage sites in the naked and nucleosomal DNA. The frequency of cleavage at many sites in the nucleosome is reduced, while the damage near position –13 is increased in frequency in the nucleosome. However, there are no new cleavage sites that appear in the nucleosomal DNA. Thus, there are no sequence-independent, nucleosome-specific DNA conformations that become attractive sites for CAL binding. This suggests that CAL targets sequence-specific DNA structures that are altered, favorably or unfavorably, by incorporation of DNA into nucleosomes.

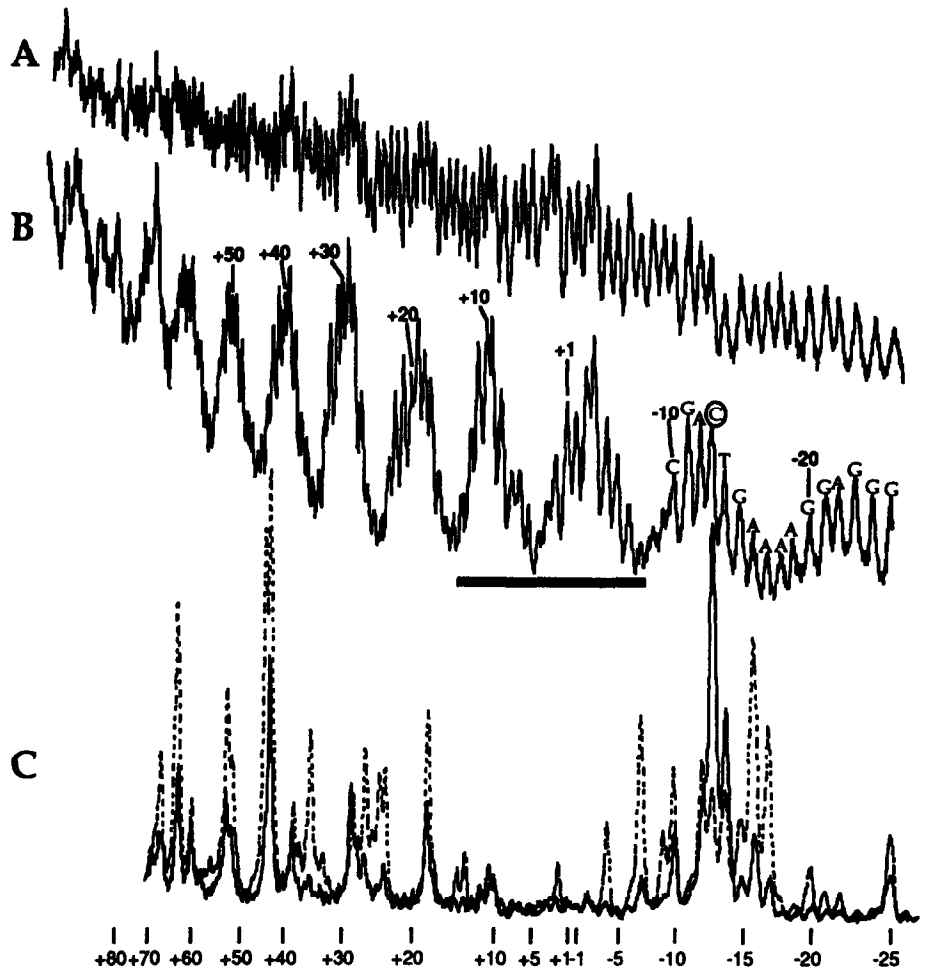

Figure 5. Calicheamicin and hydroxyl radical cleavage profiles in the noncoding strand of nucleosomal and naked 5S rDNA. The gel shown in the autoradiograph in Figure 3 was analyzed on a Molecular Dynamics Phosphorimager. (A) Hydroxyl radical cleavage of naked DNA. The solid bar indicates the possible locations of the dyad. (B) Hydroxyl radical cleavage of nucleosomal DNA. (C) Overlay of the damage patterns produced by calicheamicin in naked (dashed line) and nucleosomal DNA (solid line). The 5'-end of the strand is on the right.

Although the majority of the damage sites persist in the nucleosome, the frequency of CAL-induced damage decreases in the nucleosome compared to naked DNA at all locations, with one important exception discussed shortly. This reduction occurs even at unhindered binding sites in the minor groove facing away from the histone core, such as the damage at positions +60, +51, +29, and +18 in the noncoding strand (Figs 3 and 5). The most prominently diminished cleavage sites occur at minima of the hydroxyl radical cleavage pattern, for example positions +44, +36, +25, –7, and –17 in the noncoding strand (Figs 3 and 5). At these positions, the minor groove faces the histone proteins, which suggests that CAL binding is hindered by protein–DNA contacts. However, several CAL damage sites, in particular those at +18, +29, +39, +43, +60 and +68 in the noncoding strand (Figs 3 and 5), persist in the nucleosome in spite of their physical proximity to the histone core.

Contrary to most other sites, the frequency of CAL-induced damage between –10 and –20 in the 5S rDNA nucleosome increases 2- to 4-fold compared to naked DNA (Figs 5 and 6). This cluster of cleavage sites occurs at the 3'-end of a 13 bp homopurine·homopyrimidine tract that terminates in A_4/T_4, which has been identified as a high-frequency cleavage site for CAL damage.[11,14]

Orientation of CAL at damage sites in naked 5S rDNA

It has been proposed that minor groove width is a significant determinant in CAL binding orientation.[11] To test this hypothesis, we have established the putative binding orientation of CAL at several damage sites in relation to the local minor groove width as judged by the hydroxyl radical cleavage pattern.

The basis for these experiments is the hydroxyl radical cleavage pattern observed by Hayes *et al.* in the 5S rDNA construct.[26] Their studies revealed a sinusoidal variation in the cleavage frequency on both strands between ~ +10 and +70. Such variations in cleavage frequency have been proposed to reflect variation in the accessibility of the deoxyribose to Fe/EDTA-generated hydroxyl radical, which in turn is a direct function of minor groove width.[34] This sinusoidal variation in cleavage frequency is apparent in the dashed lines in Figure 8, which are plots of hydroxyl radical cleavage frequency along the +14 to +73 positions of the naked 5S rDNA construct; (the plot has been exaggerated

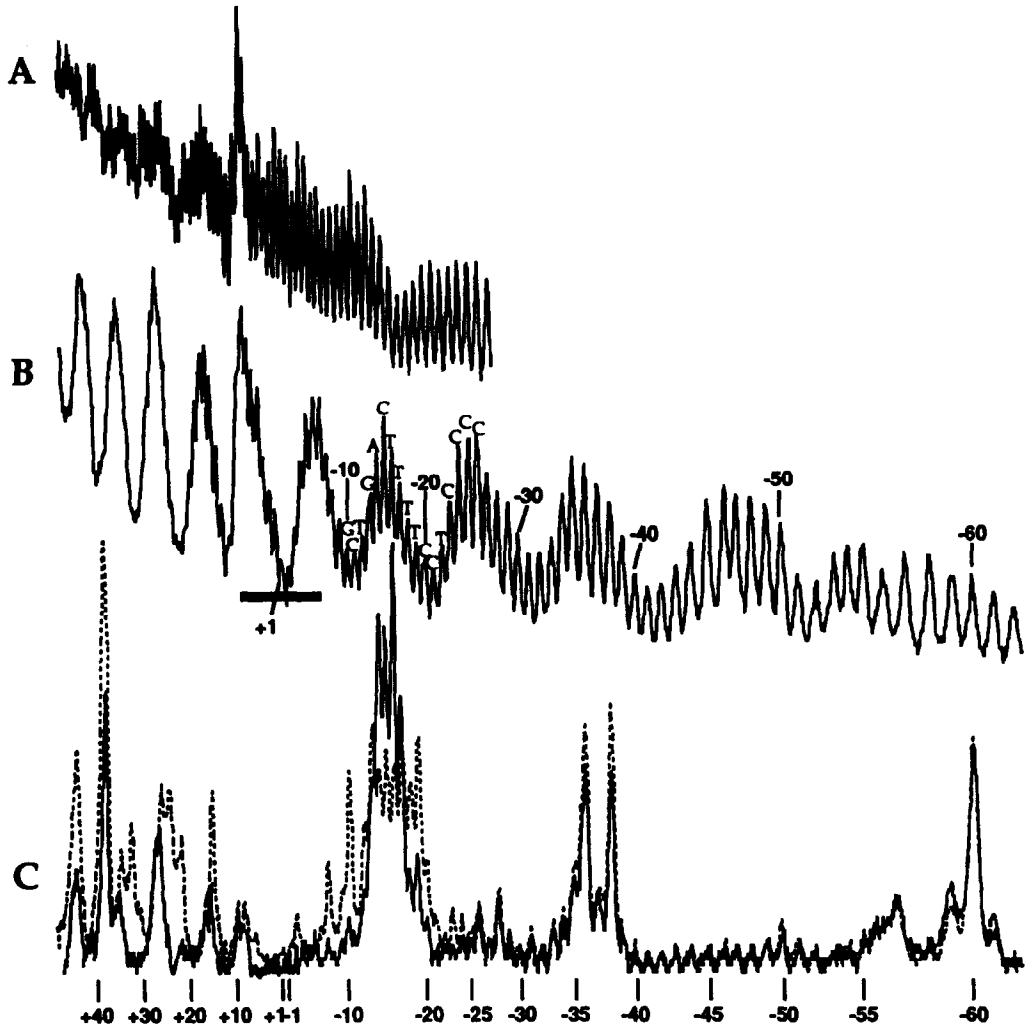

Figure 6. Calicheamicin and hydroxyl radical cleavage profiles in the coding strand of nucleosomal and naked 5S rDNA. The gel shown in the autoradiograph in Figure 4 was analyzed on a Phosphorimager. (A) Hydroxyl radical cleavage of naked DNA (data are taken from another sequencing gel also containing the samples shown in Figure 4). The solid bar indicates the possible locations of the dyad. (B) Hydroxyl radical cleavage of nucleosomal DNA. (C) Overlay of the damage patterns produced by calicheamicin in naked (dashed line) and nucleosomal DNA (solid line). The 3'-end of the strand is on the right.

vertically to emphasize the variation in cleavage frequency). The cleavage maxima and minima on each strand are offset by ~3 nt in a 3'-direction as a result of the minor groove specificity of Fe/EDTA damage.[34] The ~10 nt phasing of the cleavage pattern, and thus the minor groove width, is consistent with planar DNA curvature or anisotropic flexibility that may be important for positioning of the DNA in the nucleosome.[26]

The putative binding orientation of CAL at each damage site can be derived from the chemistry of the damage on each strand, due to the remarkably consistent selection of DNA hydrogen atoms by CAL.[2,6,10,14,18,19] Isotope transfer studies have revealed that 4'- and 5'-deoxyribose hydrogen atoms are removed by the C1 and C4 positions of the diradical form of CAL, respectively (Fig. 1).[18,19]

With the enediynes, abstraction of each of the 4'- and 5'-hydrogens results in a unique set of damage products, as shown in Figure 7.[4] Abstraction of the 4'-hydrogen produces either a keto-aldehyde abasic site, or a strand break with a phosphate on the 5'-end and a phosphoglycolate residue on the 3'-end (Fig. 7A). Based on precedent with BLM,[39] the keto-aldehyde abasic site reacts with hydrazine to form a strand break with a 3'-phosphopyridazine, which will cause a DNA fragment with a 5'-[^{32}P] label to migrate more slowly than the phosphate-ended fragments produced by the Maxam–Gilbert sequencing reactions (Fig. 7C).[35] Although not apparent in Figure 7C, 5'-labeled fragments containing a 3'-phosphoglycolate residue migrate ~1/4 nt faster than the 3'-phosphate-ended fragments. Also consistent with the presence of an abasic site, putrescine treatment causes an increase in 3'-phosphate-ended fragments (Fig. 7C).[40]

5'-Hydrogen abstraction results in either a strand break with 5'-nucleoside aldehyde- and 3'-phosphate-ended DNA fragments, or a strand break with both 3'- and 5'-phosphate-ended fragments (Fig. 7A). With 3'-[^{32}P]-labeled DNA substrates, the 5'-nucleoside aldehyde slows the migration of the DNA fragment by 1–4 nt

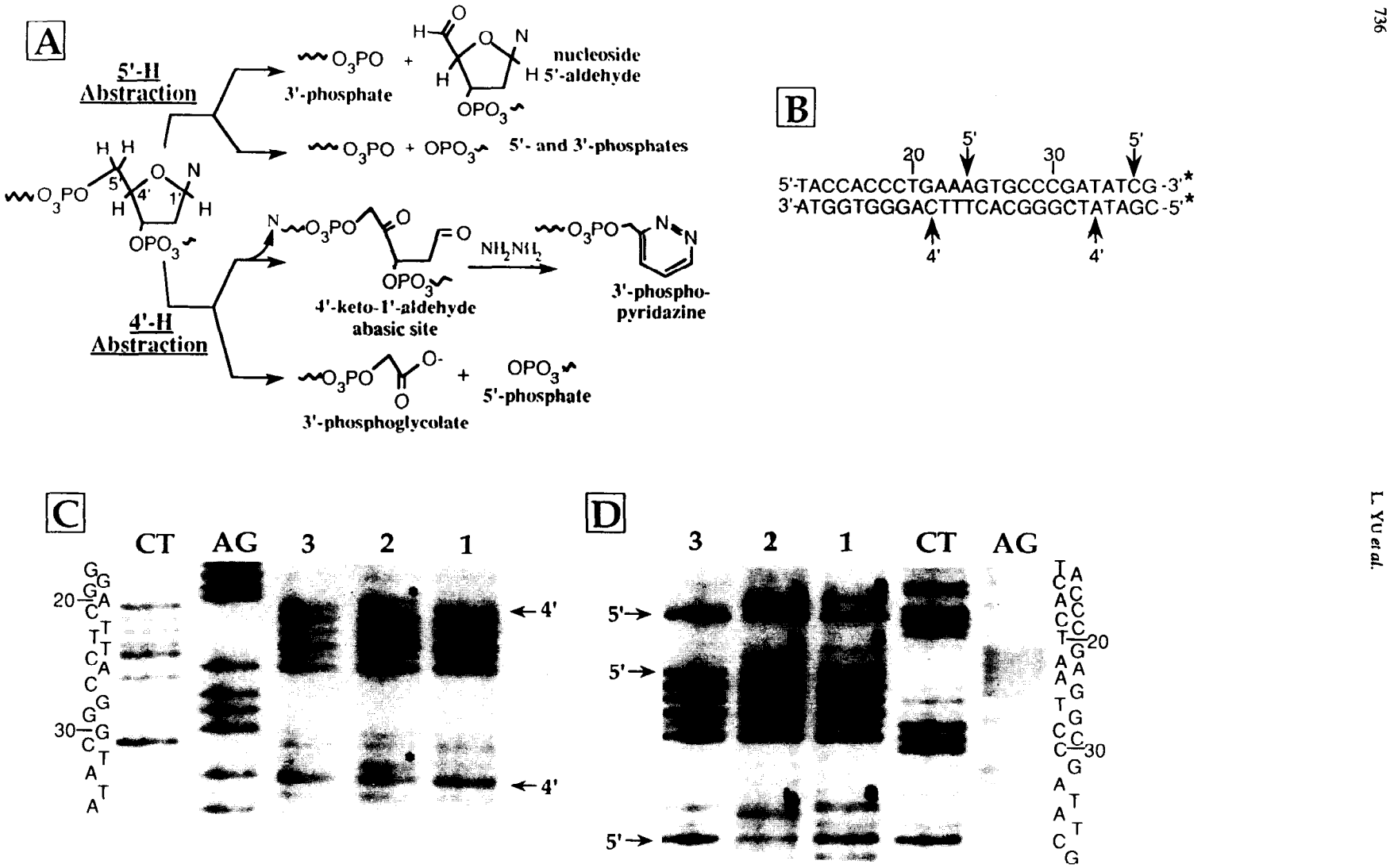

Figure 7. Examples of calicheamicin-induced DNA damage chemistry derived from sequencing gel mobility shifts. **(A)** Products resulting from the abstraction of the 4'- or 5'-hydrogen atoms from deoxyribose.[4,6] **(B)** The sequence of the +12 to +37 region of the 5S rDNA construct analyzed in this figure. Arrows denote the location of the abstracted hydrogen atoms on each strand at two CAL damage sites. Asterisks denote the locations of the 5'- and 3'-[^{32}P] end-labels at the *Sca*I cleavage site of the DNA fragments employed in the examples in panels **C** and **D**. **(C)** The 68 bp *Sca*I/*Eae*I fragment was labeled at the 5'-end of the *Sca*I site, and subjected to CAL treatment and subsequent derivatization reactions. Lane 1: CAL damage with no derivatization; lane 2: hydrazine treatment; lane 3: putrescine treatment. Asterisks denote the locations of putative 3'-phosphopyridazine-ended DNA fragments resulting from hydrazine reaction with the keto-aldehyde abasic site produced by 4'-hydrogen abstraction.[4,6] Sites of putative 4'-chemistry are denoted with arrows. CT and AG represent phosphate-ended fragments produced by Maxam–Gilbert chemical sequencing standards; the sequence is indicated in the left margin. **(D)** The 68 bp *Sca*I/*Eae*I fragment was labeled at the 3'-end of the *Sca*I site, and subjected to CAL treatment and subsequent derivatization reactions. Lane 1: CAL damage with no derivatization; lane 2: $NaBH_4$ reduction; lane 3: piperidine treatment. The putative nucleoside aldehyde-ended damage fragments are denoted by the letter **a** (above the band); and the nucleoside reduction products by the letter **b** (above the band). Sites of 5'-hydrogen abstraction are indicated with arrows. The sequence is indicated in the right margin.

(Fig. 7D, band **a**), while reduction of the aldehyde with $NaBH_4$ produces a nucleoside that migrates slightly faster than the aldehyde (Fig. 7D, band **b**).[41] The nucleoside aldehyde can also be cleaved by treatment with piperidine, which results in 5'-phosphate-ended fragments (Fig. 7D). The presence of these chemical products has been previously established for calicheamicin.[2,6,10,14,42]

Having identified 4'- and 5'-chemistry at the CAL damage sites on each strand, we can now define the orientation of CAL at these damage sites, and correlate this orientation with the hydroxyl radical-defined minor groove width. The results of these orientation studies are presented in Figure 8. The orientation of CAL at each site is indicated by the stylized representation of the CAL diradical, with circles for the two radicals and a box for the carbohydrate side chain. The orientation and size of this structure are based on the footprinting studies of Mah *et al.*[10] and NMR structures of the CAL/DNA complex.[13,15,16]

The CAL orientations in the 5S rDNA do not yield a consistent relationship to minor groove width. At high and low frequency damage sites alike, the diradical is situated in both narrow and wide minor grooves, and the carbohydrate side chain is oriented in either direction relative to the putative groove width.

Discussion

The results from both naked and nucleosomal DNA studies have revealed several important features of CAL target recognition.

CAL binding to naked and nucleosomal DNA

With one notable exception, CAL damage is less frequent in the nucleosome than in naked DNA. The least affected sites are those that occur in the most accessible regions of the nucleosomal DNA where the minor groove faces away from the histone core. The most prominently diminished cleavage sites occur where the DNA faces the histone proteins. While the orientation studies suggest that the bulk of the drug is positioned away from the protein–DNA contacts at several of these sites, the diminution in cleavage frequency may still reflect a physical barrier posed by the histone proteins. One explanation for the persistence of CAL-induced damage at sites near the surface of the histone core may be a drug-induced alteration of the local rotational setting of the DNA on the core, as observed with echinomycin and distamycin.[43]

In general, the reduction in CAL damage likely reflects nucleosome-dependent changes in DNA structure that are unfavorable to CAL binding, such as bending-induced alterations in groove structure and dynamic

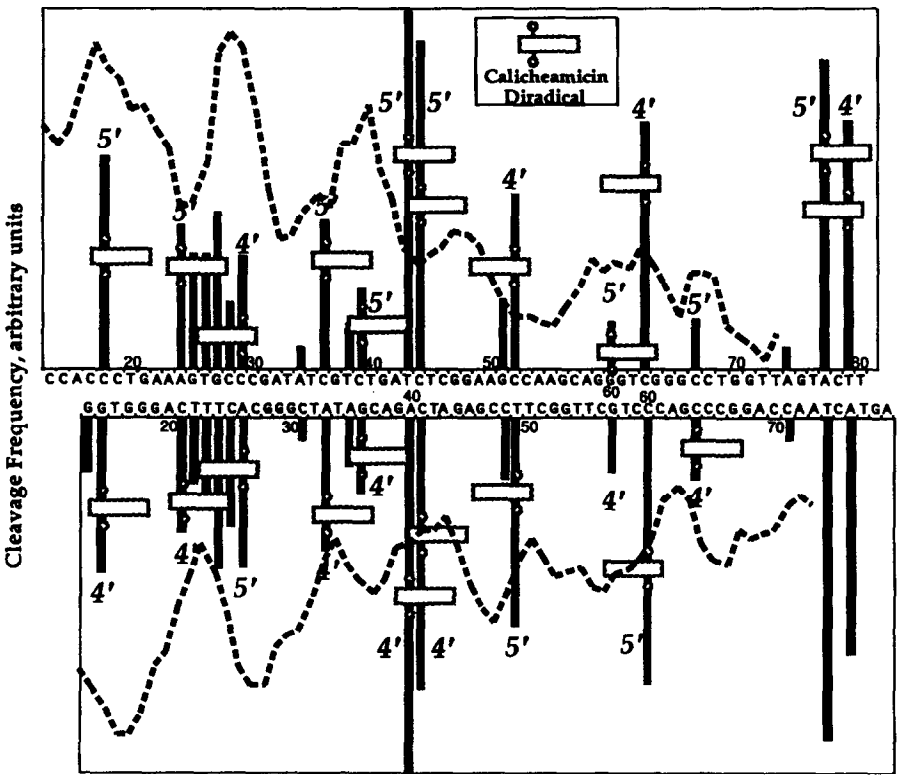

Figure 8. Putative orientation of calicheamicin at damage sites in naked 5S rDNA. The frequency of CAL damage is plotted along the sequence of the +14 and +80 region of the 5S rDNA construct. The lower (coding) strand has been offset 3 nt in a 3'-direction relative to the top strand to emphasize the relationship between CAL damage sites on each strand. The dashed line represents the hydroxyl radical cleavage profile obtained with naked 5S rDNA (see Figures 4 and 5); the y-axis scale has been expanded to emphasize the variation in cleavage frequency. The stylized representation of CAL structure, with circles for the two radicals and a box for the carbohydrate side chain, denotes the putative orientation of the drug at several damage sites, and is based on the footprinting studies[10] and an NMR structure of the CAL/DNA complex.[15]

constraints imposed by histone–DNA contacts. The observed reduction in cleavage frequency at most sites and the absence of new damage sites in the nucleosome argues that CAL targets sequence-dependent structures that cannot be mimicked solely by incorporation of DNA into nucleosomes.

There is one notable exception to the general observation of reduced cleavage in the 5S rDNA nucleosome. Between positions –10 and –20 in the noncoding strand, there is a 2- to 4-fold increase in CAL damage at the 3'-end of this 13 bp homopurine·homopyrimidine tract, which makes it the most frequently cleaved site in the nucleosome. A nucleosome-induced conformational change thus results in a more favorable binding site for CAL or more efficient cleavage by the drug. In the naked DNA, the damage sites lie in close proximity to the four adenines at the 3'-end of the purine tract, which raises the possibility that CAL is behaving as it does in longer A-tracts. Mah *et al.* observed preferential cleavage by CAL at the 3'-end of A-tracts.[11] They conclude that CAL binding is enhanced at the 3'-end either by a progressive narrowing of the minor groove or by a structural discontinuity that is hypothesized to occur near the junction of A-tracts and mixed sequence DNA.[11]

Two possible explanations for the enhancement of CAL damage are the orientation of the A_4/T_4 sequence in the nucleosome and the presence of a sharp bend in the nucleosomal DNA. The A_4/T_4 sequence is located with its narrowed minor groove facing the histone core (Fig. 5), an orientation that is consistent with the proposed role for A-rich sequences as a positioning signals in the core DNA.[44] This hypothesis is supported by the observed nonrandom location of A/T-rich sequences with their minor grooves facing the histone core,[44] and the fact that sequences containing A-tracts are curved with a compressed minor groove.[45] Additional nucleosome-induced bending of this sequence could narrow the minor groove further or possibly alter the A-tract/B-DNA junction. The unique rotational setting of this site in the nucleosome may also contribute to the observed increase in CAL damage.

Alternatively, if the dyad in our nucleosome preparations occurs at the –2 position, then the damage around –13 (noncoding strand) would occur one helical turn away from the dyad. This is the location of a sharp bend in the DNA, according to the crystallographic studies of Richmond *et al.*[27] A sharp bend alone, however, cannot be responsible for defining the damage site, since there is no similar increase in CAL damage one helical turn on the other side of the dyad (+10 or +20 in the 5S rDNA nucleosome). CAL thus appears to select a combination of both sequence-dependent conformation and nucleosome-induced structural alteration. The association of a sharp bend with the increase in CAL damage awaits more precise localization of the dyad.

Regardless of the mechanism, a more favorable binding site for CAL has been created by an interaction of proteins with DNA. This has important implications for targeting of the drug in the nucleus, since it is possible that 'hotspots' for CAL-mediated DNA damage *in vivo* are associated with chromatin-dependent DNA structural alterations.

New insights into the structure of CAL binding sites from the 5S rDNA studies

The results of CAL damage studies in both naked and nucleosomal 5S rDNA revealed several novel features of the structure of CAL targets in DNA. The orientation of CAL at several damage sites between +14 and +80 demonstrated that the minor groove width *per se* is not the primary determinant of CAL binding. This is contrary to proposed behavior of CAL observed with oligo(A) tracts[11] in which there is a defined narrowing of the minor groove at the 3'-end. It is possible that the hydroxyl radical cleavage pattern does not reflect minor groove width. However, the fact that both strands share the same ~10–11 nt periodicity of cleavage frequency (staggered in a 3'-direction) argues against this point.

The results of the orientation studies permit us to examine the DNA sequences surrounding CAL damage sites in detail. Typically, notations for CAL damage sites are organized as 6 bp sequences, with one strand containing a purine-rich sequence and two 3'-flanking nucleotides, the 3'-most being the damage site.[2,14] Based on the studies of CAL orientation at several sites between +14 and +80, we have arranged the damage site sequences in Figure 9A to conform to this notation. In virtually every case, the carbohydrate tail appears to be directed toward the 5'-end of the oligopurine strand (underlined), which contains the deoxyribose 4'-chemistry, and toward the 3'-end of the oligopyrimidine tract, which contains the 5'-chemistry.

Many of the sites in Figure 9A conform to the previously observed tetrapurine/tetrapyrimidine consensus. As indicated in other studies involving CAL binding to AGGA sequences,[9,20] the presence of a guanine at the second position of a majority of the underlined sequences suggests the possibility of an interaction of the exocyclic N2 with the iodine substituent of the aromatic ring of CAL. However, G at the second purine position is sufficient but not necessary for CAL recognition, as is apparent with AAAA sequences[14] and several sequences in the present study (Fig. 9A). Curiously, there were no cases of CAL-induced damage involving a run of four or more guanosines, or with guanosine at both the first and fourth positions of a tetrapurine consensus (e.g., at the –50 terminus of the G8 run in the coding strand; at the 3'-C in GAAGCC near +50 in the noncoding strand). All other combinations of A and G appear to be acceptable. It is thus possible that the N2/iodine interaction plays a role in modulating CAL binding at particular target sequences. An appropriately positioned G may enhance the CAL/DNA interaction while the presence of the

guanine N2 at certain other positions may act as a physical barrier to CAL binding or may unfavorably alter the purine-tract structure. The relative contributions of the N2/iodine interaction can only be judged in the context of a single DNA sequence containing a variety of similar target sites, including AAAA and AGGA.

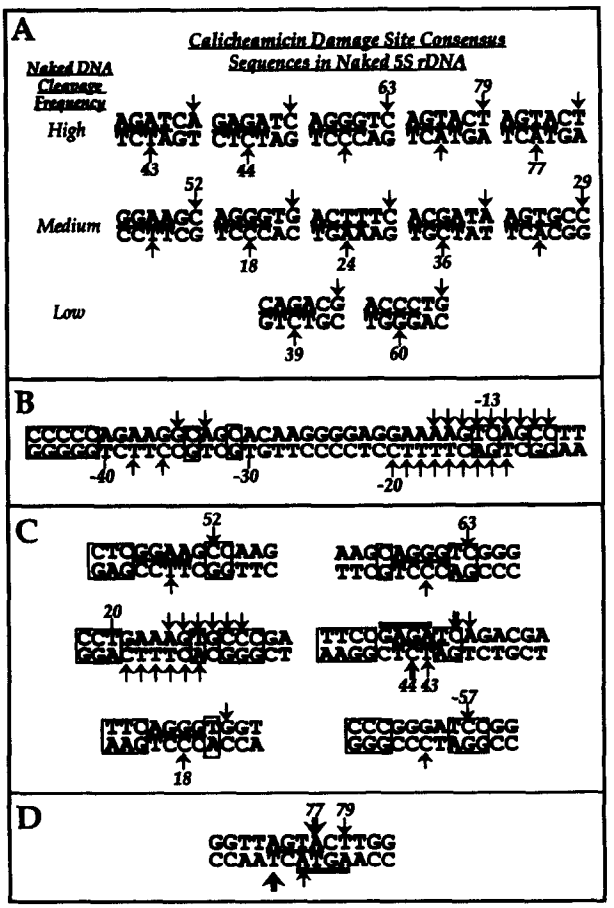

Figure 9. Analysis of CAL damage sites in 5S rDNA suggests that the drug recognizes structural discontinuities at junctions between oligopurine tracts and pyrimidine sequences. In all cases, numbering indicates sequence position in the top (noncoding) strand of the 5S rDNA construct, and arrows denote CAL damage sites on each strand. (A) Damage site sequences presented in the conventional context of oligopurine tracts, and grouped according to cleavage frequency (Fig. 8). Consensus sequences are oriented to place oligopurine tracts (underlined) in top strand, with the carbohydrate tail of the drug directed toward the 3'-end of the bottom (oligopyrimidine) strand. (B) The relationship between the long run of purines and CAL damage sites in the region –45 to –7. Putative drug orientation at damage sites in this sequence has not been defined. (C) DNA sequences surrounding major damage sites have been extended, with boxes encompassing the disruptions of the oligopurine sequences. Sequences are oriented with consensus tetrapurine in top strand. Thick and thin bars denote sequences for which binding orientation is known; thick arrows denote damage sites occurring with tetrapurine consensus marked with thick bar. (D) Features of CAL damage at an AGTACT sequence proposed to possess a structural discontinuity at the AG/CT steps.

While many cleavage sites do involve a tetrapurine sequence, more than half of the most frequently cleaved sites do not (Fig. 9A). This raises questions about the role of the oligopurine sequence in CAL target recognition. Perhaps the most telling effect of DNA sequence context occurs in the long run of purines located between –7 and –28. The 5'-end of this run is not damaged by CAL, yet contains what should be an appropriate mixture of adenines and guanines. Damage only occurs at the 3'-end of the purine tract where the sequence is interrupted by pyrimidines. This is similar to CAL-induced damage in A-tracts: the highest frequency cleavage sites are located at the 3'-end of the A-tract at the junction with mixed sequence DNA.[11]

The features of A-tract structure proposed to account for this phenomenon are either the narrowing of the minor groove at the 3'-end or the behavior of the discontinuity at the junction between A_n and mixed sequence DNA.[11] While the locus of bending associated with A_n sequences in phase with the helical repeat is still a contentious issue,[46,47] the A-tract appears to be relatively straight with bending occurring in flanking mixed sequence DNA or at the A-tract/B-DNA junction.[46,47] Hurley and coworkers also noted a similar selectivity for 3'-ends of A-tracts with the antitumor antibiotic CC-1065.[48–50] These observations support the view that CAL and CC-1065 may recognize the unique properties of an A-tract/mixed sequence junction.

We propose a similar rationale for the observed CAL targeting of the 3'-end of other purine tracts. Like A-tract DNA, there are significant base stacking interactions in a run of purines, which appears to result in an A-DNA-like sliding of the bases away from the helix axis.[51] This may account for the torsional and bending stiffness proposed for many oligopurine sequences.[52,53] The disruption of this base stacking at a purine–pyrimidine step could result in a structural discontinuity, such as either a kink or hinge at the junction or the expression of bending in the neighboring sequence.[46,47] CAL target sites might thus be viewed as runs of purines with a structural discontinuity introduced by 3'-pyrimidines.

In this light, the CAL target sequences are re-examined in Figure 9C, where the sequences associated with major cleavage sites have been extended beyond the tetrapurine motif. It is apparent that in all cases the major damage sites consist of runs of four or more purines flanked, on at least the 3'-side, by one or more pyrimidines. In almost every case, the drug is oriented with the diradical positioned near the 3'-pyrimidine interruption and the tail directed toward the 5'-end of the purine run (Fig. 9C). Also notable is the frequent occurrence of the CA/TG step, and to a lesser extent TA/TA, at the 3'-ends of the oligopurine tracts (Fig. 9). McNamara *et al.* have proposed that these steps are prone to transient kinking.[54] All of these features are consistent with a model for CAL targeting of purine tracts with structural discontinuity at the 3'-end.

CAL-induced damage at the AGTACT sequence around +76 presents a situation that may again reflect the importance of an unusual DNA structure in CAL target selection. The site sustains a high level of CAL-induced damage, and consists of the palindromic AGTACT

motif, which may account for the two opposing CAL orientations at the two symmetric cleavage sites (Figs 8 and 9D). Studies performed by Hurley and coworkers with the AGTT·AACT sequence indicate the presence of an altered helical structure, possibly a kink, at the A-G·C-T step.[50] The presence of TA and TG steps at the 5'- and 3'-ends of the tAGTACTtg sequence (Fig. 9D) also introduces another possible structural dislocation.[54] The proximity of the diradical to these altered structures and their 3'-location in the CAL consensus target are again consistent with the importance of a structural discontinuity in the CAL binding site.

These observations lead us to propose a model for CAL interaction with its target sequences. In the induced fit model of Kahne and coworkers,[15] based on a drug-induced distortion of the pyrimidine strand, flexibility of the purine·pyrimidine tract is the main determinant of drug binding. While this phenomenon may be essential to the target recognition process, we propose that the structural perturbation lying 3' to the purine tract, at or beyond the purine–pyrimidine junction, may represent the major element of flexibility that serves as an important determinant of drug binding. Indeed, preliminary gel mobility and DNA circle-closure studies suggest that CAL damage sites are flexible in the absence of bound drug and that CAL either fixes a curved structure or induces bending of the sequence (A. Salzberg and P. Dedon, unpublished observations). Experiments to test this model are currently underway.

Conclusions

1. The location of the nucleosome dyad in the 5S rDNA system has, to date, not been precisely defined. This prevents any conclusions being drawn regarding the effects of the major structural features of the nucleosome on target recognition by DNA-interactive agents.

2. With one important exception, CAL target selection is generally hampered by incorporation of DNA into a nucleosome. Alterations of CAL targets in the nucleosome include dynamic constraint imposed by protein–DNA contacts, structural alteration of targets, or steric hindrance posed by the histone proteins.

3. The observed increase in CAL damage at one site in the 5S rDNA nucleosome suggests the presence of other potential 'hot spots' for DNA damage in chromatin not apparent in naked DNA studies.

4. The results of drug orientation studies in naked 5S rDNA suggest that minor groove width *per se* is not a major determinant of CAL target selection.

5. Our studies support a model in which CAL recognizes the combined structural and dynamic properties of the 3'-end of an oligopurine tract.

Acknowledgments

The authors gratefully acknowledge the contributions of the following individuals: Punam Mathur for her skillful technical assistance and critical reading of the manuscript; Bill LaMarr, Jinghai Xu, Ken Moore, Stephanie Kong and Marcus Ware for critical review of the manuscript; Jeff Hayes and Tom Tullius for many helpful discussions; Alan Wolffe for providing the pXP-10 construct; George Ellestad for providing calicheamicin and for communicating the circular dichroism results prior to publication; and John Essigmann for the use of the Molecular Dynamics Phosphorimager and the Beckman ultracentrifuge. This work was supported by funding from the following sources: PHS/NIH (CA57633 and CA64524, PCD; Toxicology Training Grant ES07020, AAS), an MIT Sloan Basic Grant, and the Samuel A. Goldblith Career Development Professorship (PCD).

References

1. Zhao, B.; Konno, S.; Wu, J. M.; Oronsky, A. L. *Cancer Lett.* **1990**, *50*, 141.

2. Zein, N.; Sinha, A. M.; McGahren, W. J.; Ellestad, G. A. *Science* **1988**, *240*, 1198.

3. Lee, M. D.; Ellestad, G. A.; Borders, D. B. *Acc. Chem. Res.* **1991**, *24*, 235.

4. Dedon, P. C.; Goldberg, I. H. *Chem. Res. Tox.* **1992**, *5*, 311.

5. Kishikawa, H.; Jiang, Y.-P.; Goodisman, J.; Dabrowiak, J. C. *J. Am. Chem. Soc.* **1991**, *113*, 5434.

6. Dedon, P. C.; Salzberg, A. A.; Xu, J. *Biochemistry* **1993**, *32*, 3617.

7. Ding, W.-d.; Ellestad, G. A. *J. Am. Chem. Soc.* **1991**, *113*, 6617.

8. Drak, J.; Iwasawa, N.; Danishefsky, S.; Crothers, D. M. *Proc. Natl Acad. Sci. U.S.A.* **1991**, *88*, 7464.

9. Li, T.; Zeng, Z.; Estevez, V. A.; Baldenius, K. U.; Nicolaou, K. C.; Joyce, G. F. *J. Am. Chem. Soc.* **1994**, *116*, 3709.

10. Mah, S. C.; Townsend, C. A.; Tullius, T. D. *Biochemistry* **1994**, *33*, 614.

11. Mah, S. C.; Price, M. A.; Townsend, C. A.; Tullius, T. D. *Tetrahedron* **1994**, *50*, 1361.

12. Nicolaou, K. C.; Tsay, S.-C.; Suzuki, T.; Joyce, G. F. *J. Am. Chem. Soc.* **1992**, *114*, 7555.

13. Paloma, L. G.; Smith, J. A.; Chazin, W. J.; Nicolaou, K. C. *J. Am. Chem. Soc.* **1994**, *116*, 3697.

14. Walker, S.; Landovitz, R.; Ding, W. D.; Ellestad, G. E.; Kahne, D. *Proc. Natl Acad. Sci. U.S.A.* **1992**, *89*, 4608.

15. Walker, S.; Murnick, J.; Kahne, D. J. *J. Am. Chem. Soc.* **1993**, *115*, 7954.

16. Walker, S. L.; Andreotti, A. H.; Kahne, D. E. *Tetrahedron* **1994**, *50*, 1351.

17. Zein, N.; McGahren, W. J.; Morton, G. O.; Ashcroft, J.; Ellestad, G. A. *J. Am. Chem. Soc.* **1989**, *111*, 6888.

18. 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.

19. Hangeland, J. J.; De Voss, J. J.; Heath, J. A.; Townsend, C. A.; Ding, W.-d.; Ashcroft, J. S.; Ellestad, G. A. *J. Am. Chem. Soc.* **1992**, *114*, 9200.

20. Hawley, R. C.; Kiessling, L. L.; Schreiber, S. L. *Proc. Natl Acad. Sci. U.S.A.* **1989**, *86*, 1105.

21. Aiyar, J.; Danishefsky, S. J.; Crothers, D. M. *J. Am. Chem. Soc.* **1992**, *114*, 7552.

22. Walker, S.; Valentine, K. G.; Kahne, D. J. *J. Am. Chem. Soc.* **1990**, *112*, 6429.

23. Krishnamurthy, G.; Ding, W.-d.; O'Brien, L.; Ellestad, G. A. *Tetrahedron* **1994**, *50*, 1341.

24. Wolffe, A. *Chromatin Structure and Function*; Academic Press; San Diego, 1992.

25. van Holde, K. E. *Chromatin*; Springer-Verlag; New York, 1989.

26. Hayes, J. J.; Tullius, T. D.; Wolffe, A. P. *Proc. Natl Acad. Sci. U.S.A.* **1990**, *87*, 7405.

27. Richmond, T. J.; Finch, J. T.; Rushton, B.; Rhodes, D.; Klug, A. *Nature* **1984**, *311*, 532.

28. Yu, L.; Goldberg, I. H.; Dedon, P. C. *J. Biol. Chem.* **1994**, *269*, 4144.

29. Ikemoto, N.; Kumar, R. A.; Dedon, P.; Danishefsky, S. J.; Patel, D. J. *J. Am. Chem. Soc.* **1994**, *116*, 9387.

30. Yu, L.; Golik, J.; Harrison, R.; Dedon, P. *J. Am. Chem. Soc.* **1994**, *116*, 9733.

31. Rhodes, D. *EMBO J.* **1985**, *4*, 3473.

32. Wolffe, A. P.; Jordan, E.; Brown, D. D. *Cell* **1986**, *44*, 381.

33. Ausubel, F. M.; Brent, R.; Kingston, R. E.; Moore, D. D.; Seidman, J. G.; Smith, J. A.; Struhl, K. *Current Protocols in Molecular Biology*; John Wiley and Sons; New York, 1989.

34. Price, M. A.; Tullius, T. D. *Meth. Enzymol.* **1992**, *212*, 194.

35. Maxam, A.; Gilbert, W. *Meth. Enzymol.* **1980**, *65*, 499.

36. Lee, D. Y.; Hayes, J. J.; Pruss, D.; Wolffe, A. P. *Cell* **1993**, *72*, 73.

37. Gottesfeld, J. M. *Mol. Cell. Biol.* **1987**, *7*, 1612.

38. Pruss, D.; Wolffe, A. P. *Biochemistry* **1993**, *32*, 6810.

39. Sugiyama, H.; Xu, C.; Murugesan, N.; Hecht, S. M.; van der Marel, G. A.; van Boom, J. H. *Biochemistry* **1988**, *27*, 58.

40. Lindahl, T.; Andersson, A. *Biochemistry* **1972**, *11*, 3618.

41. Dedon, P. C.; Goldberg, I. H. *J. Biol. Chem.* **1990**, *265*, 14713.

42. Zein, N.; Poncin, M.; Nilakantan, R.; Ellestad, G. A. *Science* **1989**, *244*, 697.

43. Low, C. M. L.; Drew, H. R.; Waring, M. J. *Nucleic Acids Res.* **1986**, *14*, 6785.

44. Travers, A. A. *Trends in Biochem. Sci.* **1987**, *12*, 108.

45. Crothers, D. M.; Haran, T. E.; Nadeau, J. G. *J. Biol. Chem.* **1990**, *265*, 7093.

46. Grzeskowiak, K.; Goodsell, D. S.; Kaczor-Grzeskowiak, M.; Cascio, D.; Dickerson, R. E. *Biochemistry* **1993**, *32*, 8923.

47. Koo, H.-S.; Drak, J.; Rice, J. A.; Crothers, D. M. *Biochemistry* **1990**, *29*, 4227.

48. Lin, C. H.; Sun, D.; Hurley, L. H. *Chem. Res. Toxicol.* **1991**, *4*, 21.

49. Lin, C. H.; Hill, G. C.; Hurley, L. H. *Chem. Res. Toxicol.* **1992**, *5*, 167.

50. Sun, D.; Lin, C. H.; Hurley, L. H. *Biochemistry* **1993**, *32*, 4487.

51. McCall, M.; Brown, T.; Hunter, W. N.; Kennard, O. *Nature* **1986**, *322*, 661.

52. Hogan, M.; LeGrange, J.; Austin, B. *Nature* **1983**, *304*, 752.

53. Dickerson, R. E.; Drew, H. R. *J. Mol. Biol.* **1981**, *149*, 761.

54. McNamara, P. T.; Bolshoy, A.; Trifinov, E. N.; Harrington, R. E. *J. Biomolec. Struc. Dyn.* **1990**, *8*, 529.

(*Received in U.S.A.* 1 *November* 1994; *accepted* 5 *December* 1994)