Nucleic Acids Research, 1993, Vol. 21, No. 18 4281–4287

(+)-CC-1065 as a structural probe of Mu transposase-induced bending of DNA: overcoming limitations of hydroxyl-radical footprinting

Zhi-Ming Ding[1], Rasika M.Harshey and Laurence H.Hurley[1,*]
Department of Microbiology and [1]Drug Dynamics Institute, College of Pharmacy, University of Texas at Austin, Austin, TX 78712-1074, USA

Received May 27, 1993; Revised and Accepted July 23, 1993

ABSTRACT

Phage Mu transposase (A-protein) is primarily responsible for transposition of the Mu genome. The protein binds to six *att* sites, three at each end of Mu DNA. At most *att* sites interaction of a protein monomer with DNA is seen to occur over three minor and two consecutive major grooves and to result in bending up to about 90°. To probe the directionality and locus of these A-protein-induced bends, we have used the antitumor antibiotic (+)-CC-1065 as a structural probe. As a consequence of binding within the minor groove, (+)-CC-1065 is able to alkylate N3 of adenine in a sequence selective manner. This selectivity is partially determined by conformational flexibility of the DNA sequence, and the covalent adduct has a bent DNA structure in which narrowing of the minor groove has occurred. Using this drug in experiments in which either gel retardation or DNA strand breakage are used to monitor the stability of the A-protein – DNA complex or the (+)-CC-1065 alkylation sites on DNA (*att* site L3), we have demonstrated that of the three minor grooves implicated in the interaction with A-protein, the peripheral two are 'open' or accessible to drug bonding following protein binding. These drug-bonding sites very likely represent binding at at least two A-protein-induced bending sites. Significantly, the locus of bending at these sites is spaced approximately two helical turns apart, and the bending is proposed to occur by narrowing of the minor groove of DNA. The intervening minor groove between these two peripheral sites is protected from (+)-CC-1065 alkylation. The results are discussed in reference to a proposed model for overall DNA bending in the A-protein *att* L3 site complex. This study illustrates the utility of (+)-CC-1065 as a probe for protein-induced bending of DNA, as well as for interactions of minor groove DNA bending proteins with DNA which may be masked in hydroxyl radical footprinting experiments.

INTRODUCTION

Transposition of the Mu genome occurs by a series of precisely regulated events wherein Mu DNA ends are cleaved and joined to target DNA (1, 2). Although several steps in this reaction have been elucidated, the detailed mechanisms by which these are executed are unknown. The Mu A-protein initially binds to an array of six *att* sites, three at each end of Mu DNA (L1–L3 at *att* L and R1–R3 at *att* R) (3, 4). By events poorly understood, involving interaction of the A-protein with an internal enhancer element (5–7) as well as interaction of the *E.coli* HU protein with the DNA (8, 9), the A-protein eventually forms a tetrameric higher-order nucleoprotein complex responsible for executing the Mu DNA cleavage and strand transfer steps of transposition (10–12).

DNA bending, which serves an important architectural role in the assembly of complex nucleoprotein structures involved in nearly all aspects of DNA–protein transactions, including DNA packaging, transcription, replication, and recombination (13), is also involved in the formation of the Mu transposition complex. The hallmark of this transposition complex is its unusual stability (50% disruption in 4.6 M urea or at 56°C) (10). Interaction of the Mu A-tetramer with only three *att* sites is sufficient to confer this extraordinary stability to the complex (14). At each *att* site a monomer of A-protein is proposed to contact a 24-base-pair region along one face of the DNA helix and bend the DNA (4, 14). The overall DNA bend is ~80–90° at five of the strong *att* sites and ~60° at the weakest *att* site, L2, as determined by polyacrylamide gel electrophoresis. The DNA bends must play an indispensable role in achieving specific wrapping of DNA in the transposition complex (4). A knowledge of the locus and directionality of these bends is essential not only for establishing the path of DNA in the complex but also for understanding how the DNA bends facilitate protein–DNA contacts that impart maximum stability to the complex.

To gain insight into the sites at which A-protein-induced bends occur, we have used the DNA-reactive drug (+)-CC-1065 as a structural probe. (+)-CC-1065 is a potent antitumor antibiotic that binds overlapping with a 4–5 bp region within the minor

* To whom correspondence should be addressed

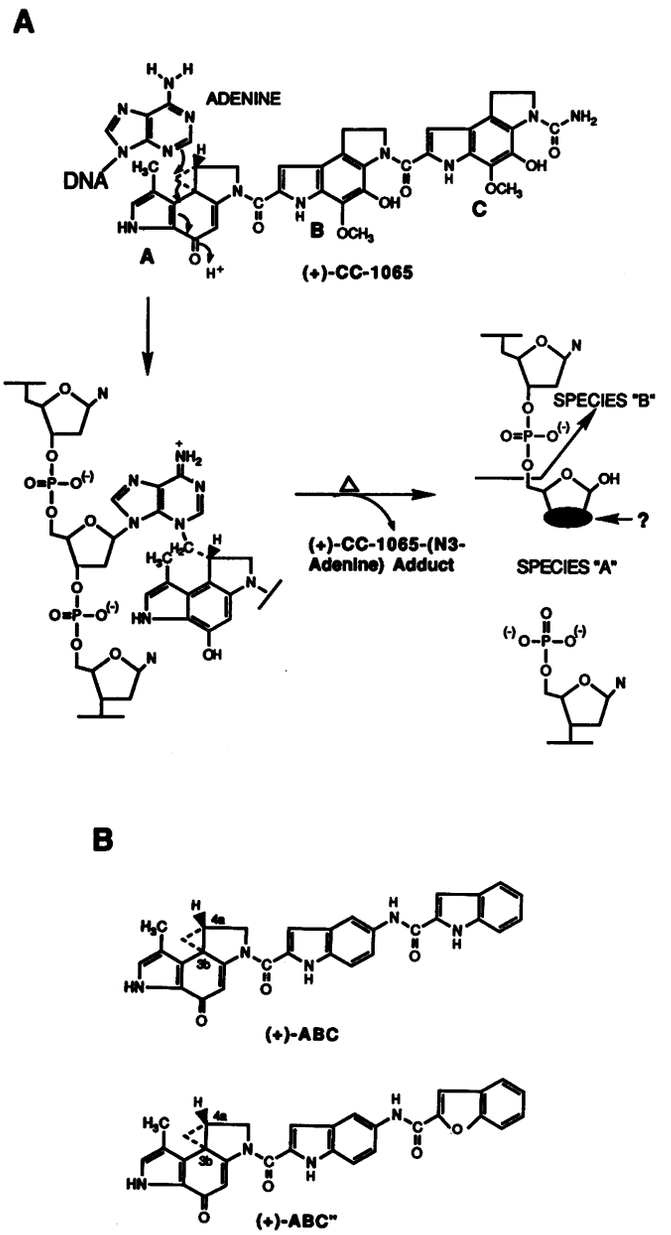

Figure 1. (A) The reaction of (+)-CC-1065 with DNA to form the (+)-CC-1065-(N3-adenine)-DNA adduct and products of heat-strand breakage of modified DNA. Species 'A' is the first product produced by thermal treatment and is converted into species 'B' by further thermal treatment or subsequent treatment with piperidine (16). (B) Structures of related drugs (+)-ABC and (+)-ABC''

groove of DNA and alkylates N3 of adenine (Figure 1A) in certain sequences (15–17). DNA flexibility has been identified as an important component of sequence recognition for covalent reaction of (+)-CC-1065 with DNA and, in particular, the propensity to form a bent DNA structure as a consequence of covalent bonding (18, 19). The (+)-CC-1065-induced bending of DNA is associated with *narrowing* of the minor groove and discontinuities on either side of the compressed region. In many respects the structural characteristics of (+)-CC-1065-induced bends are analogous to those associated with A-tract bending (20), and a 'truncated junction bend model' has been proposed for this structure (21, 22). As would be expected for a structural probe that favors minor groove narrowing and associated helix bending, A-tracts, and in particular the 3′ adenines of short A-tracts, are favored sites for (+)-CC-1065 modification (16, 21).

In this article we have determined at which sequences within the A-protein binding region of L3 at *att* L, covalent bonding by (+)-CC-1065 is compatible with protein-induced bending. Based upon these results we predict that there are at least two A-protein-induced bending sites within L3 of *att* L. For these two sites, bending is into minor grooves spaced approximately two helical turns apart. We discuss how the overall 90° bend might be accomodated at L3.

MATERIALS AND METHODS

Materials

(+)-CC-1065, (+)-ABC, and (+)-ABC″ were provided by Patrick McGovren of The Upjohn Company, Kalamazoo, MI. (γ-^{32}P) ATP was purchased from ICN, T4 polynucleotide kinase and deoxynucleotides from United States Biochemical Co., and G-25 Sephadex quick-spin columns from Boehringer. Electrophoretic reagents [acrylamide/bis(acrylamide), TEMED, and ammonium persulfate] were from Bio-Rad and X-ray film, intensifying screens, and developing chemicals from Amersham. Mu A-protein was purified as described previously (14).

Synthesis and purification of oligonucleotides

The oligonucleotides were synthesized on an automated DNA synthesizer (Applied Biosystems 381A) by the phosphoramidite method. Oligonucleotides were deprotected in concentrated ammonium hydroxide at 55°C overnight, dried under vacuum, and purified by 16% denaturing acrylamide gels. Oligonucleotides were 5′-end labeled with ^{32}P, as described before (23).

Formation of Mu transposase–DNA complexes and gel retardation assay

Mu A-protein–DNA complexes were formed in 25 μl of solution containing 0.1 pmoles of 5′-labeled DNA, 25 mM HEPES-KOH, pH 7.2, 10 mM $MgCl_2$, 70 mM NaCl, 1mM DTT, and indicated amounts of protein. The samples were incubated at 30°C for 15 min. Protein–DNA complexes were separated from free DNA by gel electrophoresis on 6% nondenaturing gels (acrylamide: bisacrylamide, 29:1) in TBE (89 mM Tris-borate, pH 8.3, 2 mM EDTA) at 4°C.

Preparation of *att* L3 selectively modified on either the (+) or (−) strand

5′-^{32}P-end-labeled single stranded DNA (top and bottom strand in Figure 2) in 25μl of distilled water was annealed with a primer (5′-CGTGTGTGTGTTTTTCGTATTT-TCAAT-3′ or 5′-ACT-TGCTTTTGTTTCATTGAA-3′, respectively). These partial duplex DNA molecules were modified with 0.28 mM (+)-CC-1065 at room temperature for three days, followed by phenol extraction to remove unbound drug. The primer was extended by adding 3 μl of 0.2 mM deoxynucleotides and 10 u of Klenow fragment. Each species of drug-modified duplex DNA was then purified by 8% gel electrophoresis.

Selective protection of DNA to drug modification upon Mu transposase binding

The A-protein–DNA complexes were formed as described above in a molar ratio of 1:30 (DNA to protein) at which DNase I

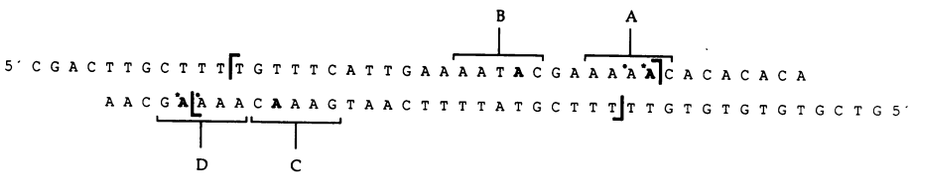

Figure 2. Sequence of oligomer encoding *att* site L3 used in this study. (+)-CC-1065 bonding sites are indicated in boldface, and the horizontal brackets indicate the drug overlap sites. The symbols (*) and (•) indicate major and minor sites for drug bonding. Vertical bold brackets indicate MPEβFe(II) footprinting data (4). A larger region (~ 30 bp), which includes drug bonding site D, is protected from DNAaseI cleavage at this site (3).

footprinting showed complete occupancy by Mu A-protein. Drug was added to the mixture in a final concentration 0.65 mM and incubated at 4°C for the indicated time. The reaction was quenched by three phenol extractions and followed by ethanol precipitation. The DNA was finally dissolved in distilled water and subjected to thermal treatment at 100°C for 45 min. (16). After evaporation of water, the DNA was redissolved in formamide dye and analyzed on 20% denaturing gels.

Interference of Mu transposase binding to DNA after partial modification with (+)-CC-1065

DNA was modified with 1.25 mM of (+)-CC-1065 for the indicated times. The reaction was quenched by three phenol extractions followed by ethanol precipitation. The partially drug-modified DNA was subjected to Mu A-protein binding and gel retardation. Protein-bound and free DNA species were excised from the gel; the DNA was extracted with distilled water and subjected to thermal treatment as before.

RESULTS

(+)-CC-1065 interacts with four 'A-tract' regions in *att* L3

The L3 *att* L site (see Figure 2) consists of approximately 24 bp, as defined by MPEβFe (II) footprinting experiments (4; bold line brackets in Figure 2). Within this region there are several AT-rich regions that are potential minor groove bending sites, as well as (+)-CC-1065 bonding sites. (+)-CC-1065 bonding within this region on both strands was determined using a thermal cleavage assay (16) in which, after incubation with drug, the duplex oligomer ^{32}P-labeled at a single 5′ end was heated at 90°C for 45 min. This results in depurination and two subsequent β-eliminations to yield two or three major products, depending upon precise conditions, of which the one found at lower molecular weight is equivalent to a Maxam and Gilbert 'A' sequencing reaction [lower molecular weight band of each pair in Figure 6A (see later)]. The position of this band relative to sequencing lanes pinpoints the adenine at which covalent modification by (+)-CC-1065 has occurred. The determined bonding sites on the upper (+) and lower (−) strands of the L3 oligomer are in accord with the established sequence specificity for (+)-CC-1065 (16, 17). For sites A and D, the major alkylation site in the A-tract is the 3′ adenine (* in Figure 2) with a minor amount of drug bonding at the adenine to the 5′ side (● in Figure 2). Because of the closeness of drug bonding C and D on the (−) strand, it is unlikely that full saturation at both sites on this strand is attained. In one of the subsequent experiments (see later), two synthetic analogs of (+)-CC-1065, (+)-ABC and (+)-ABC′′ (Figure 1B), which are structurally related to (+)-CC-1065 but have a more restricted sequence specificity (18), were used. In these cases the covalent bonding sites are the same as (+)-CC-1065, except site C on the (−) strand is less reactive to these drugs.

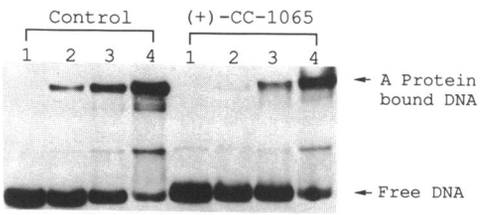

Figure 3. Effect of (+)-CC-1065 modification of *att* L3 on the binding of Mu A-protein. 5′-^{32}P-labeled L3 oligomer (Figure 2) was modified with 0.28 mM of (+)-CC-1065 at room temperature for 3 days to achieve maximum bonding. 0.1 pmole of control (unmodified) or (+)-CC-1065-modified DNA was incubated with 0, 0.5, 1, and 3 pmole of A-protein (lanes 1−4) and analyzed by the gel retardation assay, as described in Materials and Methods.

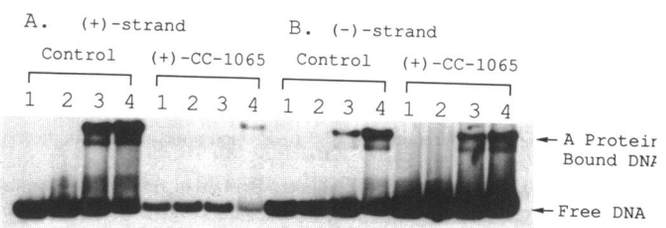

Figure 4. Effect of (+)-CC-1065 modification of the upper and lower strands of *att* L3 on Mu A-protein binding efficiency. L3 oligomer duplex molecules were 5′-^{32}P-single-end labeled, and the (+) strand (**A**) and (−) strand (**B**) selectively modified with (+)-CC-1065 in two separate experiments as described in Materials and Methods. In both control (unmodified duplex) and in (+)-CC-1065-modified experiments for both (+) and (−) strands, the duplex was incubated with 0, 0.5, 1, and 3 pmole of A-protein (lanes 1−4, respectively). Following incubation, the efficiency of A-protein binding was determined using the gel retardation assay, as described in Materials and Methods.

Mu A-protein still binds to *att* L3 after covalent modification by (+)-CC-1065

In a preliminary experiment, the effect of covalent modification of the L3 oligomer on A-protein binding was determined. A gel retardation assay (Figure 3) was used to evaluate the relative affinity of Mu A-protein for site L3 in the absence and presence of drug modification. The results show that while there is a modest amount of inhibition, the A-protein still appears to bind to a subpopulation of the drug-modified oligomers, suggesting that drug occupancy of certain of the minor grooves with the

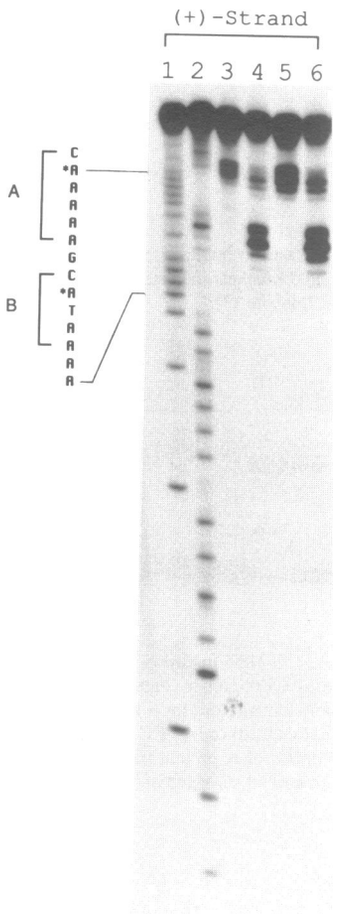

Figure 5. Binding of A-protein to *att* L3 (+) strand after partial modification with (+)-CC-1065. The DNA was 5′-^{32}P-single-end labeled on the (+) strand and incubated with (+)-CC-1065 for limited periods of time (5 or 10 min). Protein−DNA complexes were separated from non-protein complexed DNA by gel electrophoresis as described in Materials and Methods. Lanes 1 and 2 are Maxam−Gilbert purine- and pyrimidine-specific seqencing reactions, respectively. Lanes 3 and 4 are thermal treatment of A-protein-bound DNA and free DNA, respectively, following a 5-min incubation with (+)-CC-1065. Lanes 5 and 6 are thermal treatment of A-protein-bound DNA and free DNA, respectively, after a 10-min incubation with (+)-CC-1065. The drug-bonding sites (A, B), sequences, and drug-covalently-modified adenine (asterisk) are indicated on the left of the gel. In lanes 4 and 6, the multiple strand breakage sites are expected, based upon the thermal cleavage pattern shown in Figure 1A (species A and B). The two lower molecular weight bands are due to minor alkylation at 5′AAAA*TA, whereas the upper pair are due to alkylation at site B.

associated drug-enhanced bending of the corresponding A-tracts is tolerated. Based upon these preliminary but encouraging results, a series of experiments were carried out to determine in more detail how (+)-CC-1065 occupancy of individual minor grooves and associated drug-induced bending might be compatible with A-protein binding to DNA.

(+)-CC-1065 modification of the lower strand of *att* L3 does not affect A-protein binding efficiency

In order to address the question of how (+)-CC-1065 modification of either the (+) or (−) strand of the L3 oligomer effected A-protein binding, two different (+)-CC-1065 oligomer constructions in which either the upper or lower strand was modified with (+)-CC-1065 to achieve maximum bonding at each of the sites were prepared. The gel retardation assay was then used to evaluate the relative binding affinity of each (+)-CC-1065-modified or unmodified duplex to A-protein (Figure 4). While in the control experiments (A and B in Figure 4) the gel retardation assay showed a concentration dependency of Mu A-protein binding to DNA, only in the case of the lower strand (B in Figure 4) did A-protein bind efficiently to the (+)-CC-1065-modified L3 site. Drug modification of the upper strand, on the other hand, was detrimental to protein binding.

Mu A-protein excludes an L3 oligomer which is (+)-CC-1065-modified at site B, but is still able to bind to one that is modified at site A on the upper strand

In order to determine if one or both of the two drug bonding sites on the (+) strand was inhibitory to A-protein binding, partially drug-modified oligomers were prepared and incubated with A-protein prior to separation of A-protein−oligomer complexes from non-protein complexed oligomer. Partial modification of the L3 DNA was carried out by limited-time (5 or 10 min) incubation with (+)-CC-1065 with either the (+) or (−) strand ^{32}P labeled. The site of modification and extent of bonding of (+)-CC-1065 was measured using the thermal cleavage assay (16). The results for the upper strand are shown in Figure 5. Mu A-protein excludes oligomers modified at site B, but is still able to bind to oligomers that are modified at site A at the 3′ terminal adenine (compare lanes 3 with 4 and lanes 5 with 6 which show drug modification in bound vs free DNA respectively at two different drug modification times). Within the drug-bonding site A, (+)-CC-1065 modification at the *terminal* 3′ adenine (* in Figure 2) is the preferred site for A-protein binding rather than modification at the adenine to the 5′ side (compare pattern of cleavage in bound vs free DNA for this site; the middle and lower molecular weight bands represent the complete reaction products corresponding to either the 3′ adenine or its 5′ neighbor.)

In a similar experiment on the (−) strand, Mu A-protein was seen to bind equally well to unmodified or (+)-CC-1065 modified DNA (data not shown), in accord with previous results.

Prebinding of Mu A-protein to *att* L3 preferentially protects site B on the (+) strand from covalent modification by (+)-CC-1065 and increases the reactivity of site D to modification with (+)-ABC and (+)-ABC″

Drug protection experiments were carried out next, in order to explore the accessibility of (+)-CC-1065 to protein-bound DNA. Both the (+) and (−) strands of the oligomer were ^{32}P-labeled in separate experiments, annealed to complementary unlabeled strands, preincubated with Mu A-protein, and then reacted with (+)-CC-1065. The oligomers were then subjected to thermal treatment to reveal the drug bonding sites (Figure 6A). Of the four drug bonding sites (A−D) on the (+) and (−) strands, Mu A-protein binding selectively protects site B from (+)-CC-1065 modification (compare lanes 5 and 6 in each case). Overall, this result suggests that three of the four drug bonding sites remain open to modification by (+)-CC-1065 in the presence of A-protein. Site B, protected from drug modification, lies in the central minor groove protected by the protein (4). This result is in complete agreement with earlier results in which only drug modification of the B site on the (+) strand significantly inhibited A-protein binding.

Drug protection experiments using two synthetic analogs of (+)-CC-1065, (+)-ABC and (+)-ABC″, which selectively bind

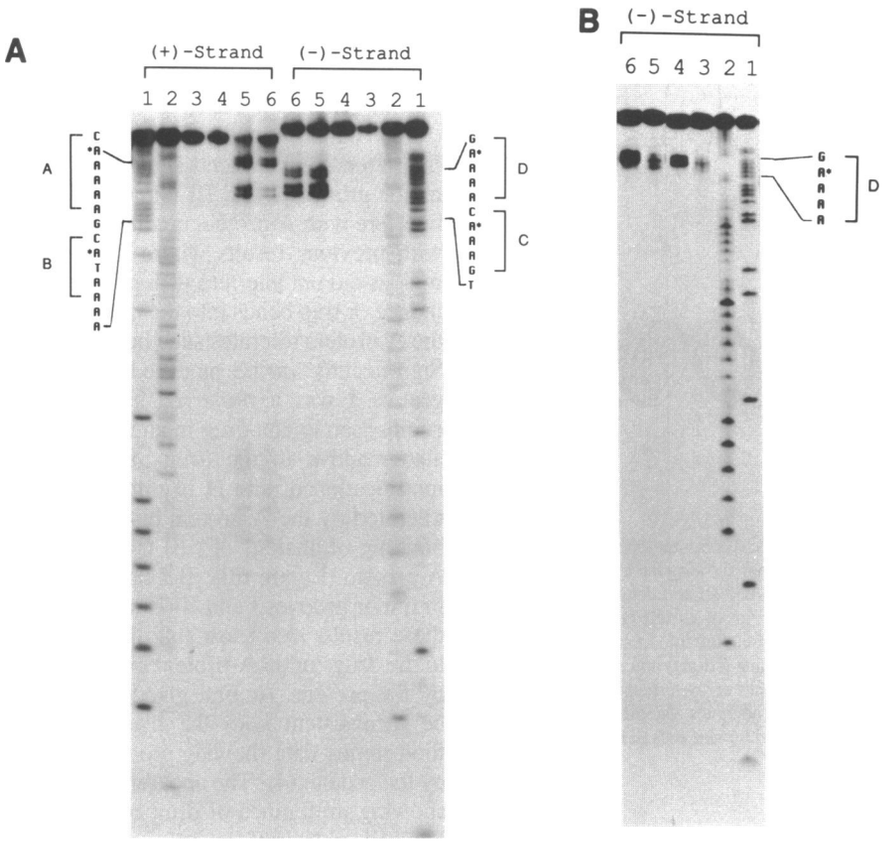

Figure 6. Effect of Mu A-protein binding to *att* L3 oligomer on (+)-CC-1065, (+)-ABC, and (+)-ABC'' alkylation. **(A)** 5'-^{32}P-labeled upper (+) and lower (−) strand duplex DNA molecules were prepared as described in Materials and Methods. In each case [(+) and (−) strands], lanes 1 and 2 are Maxam-Gilbert purine- and pyrimidine-specific sequencing reactions. Lane 3 is thermal treatment of control (unmodified) DNA, while lane 4 is thermal treatment of control DNA in the presence of Mu A-protein. Lanes 5 and 6 are thermal treatment of DNA modified with (+)-CC-1065 for 10 min in the absence and presence of A-protein, respectively. The bonding sites (A−D), sequences, and the covalently modified adenine (asterisk) are indicated at the right and left of the (−) and (+) strands. **(B)** As for (A) except that only the (−) strand was labeled and (+)-ABC and (+)-ABC'' were used in place of (+)-CC-1065. Lanes 3 and 4 are (+)-ABC and lanes 5 and 6 are (+)-ABC'' incubated for 5 min in the presence of A-protein (lanes 4 and 6) and in the absence of A-proteins (lanes 3 and 5). For this experiment, the sequence outside the *att* L3 site was slightly different to that shown in Figure 2, but this did not effect the (+)-CC-1065 alkylation pattern or A-protein binding.

to just site D on the (−) strand, are shown in Figure 6B. The results show that preincubation with Mu A-protein *increased* the reactivity of this site to drug modification (compare lanes 3 and 5 with lanes 4 and 6). It is not obvious why there is an apparent difference between (+)-CC-1065 and the two analogs, (+)-ABC and (+)-ABC'', in this regard (possibly due to inherent differences in drug reactivity at sites C and D). Irrespective of the exact reasons for these differences, it is clear that Mu A-protein predisposed the conformation of site D to covalent modification by (+)-ABC and (+)-ABC''. On the (+) strand, while both (+)-ABC and (+)-ABC'' bind to site A, both drugs were preferentially inhibited in binding to site B by preincubation with Mu A-protein, a result similar to the one obtained with (+)-CC-1065 above (data not shown).

DISCUSSION

The Mu transposase is a sequence-specific DNA binding and bending protein (3, 4, 14). DNA recognition is unusual in that, unlike most DNA binding proteins, the Mu A-protein appears to recognize a rather large DNA sequence (24 bp) with no apparent dyad symmetry and binds to it as a monomer. MPEβFe (II) and hydroxyl-radical footprints at five of the six *att* sites show protection of one face of the DNA helix over three consecutive minor and two major grooves (4). Binding specificity is achieved primarily through major groove contacts. Although several runs of adenines occur in the *att* sites, no essential minor groove contacts were identified at most sites in the earlier study (4). Since four or more adenines in runs (A-tracts) impart curvature to DNA if repeated in phase with the DNA helical repeat (22), it was speculated that the presence of these adenine runs might impart conformational flexibility to the *att* sites, thus facilitating binding and bending of the DNA by Mu A-protein.

(+)-CC-1065 and its structural analogs (+)-ABC and (+)-ABC'' are minor groove reactive compounds that covalently modify DNA through reaction with N3 of adenine (15, 16). Only adenines in certain sequences, such as 5' PuNTTA* and 5' AAAAA* (* is covalently modified adenine), are reactive towards (+)-CC-1065 (16), and one of the mechanisms for sequence recognition is proposed to involve a sequence dependent conformational flexibility (18,19). A common structural consequence of (+)-CC-1065 covalent bonding to adenines in DNA is a bent DNA structure in which the bending occurs into the minor groove (23). Studies using gel electrophoresis, hydroxyl-radical footprinting, and high-field NMR have revealed that the structural origin of the (+)-CC-1065-induced bend is quite similar to that associated with A-tracts (20, 21), and a truncated junction bend model (21) has been proposed. (+)-

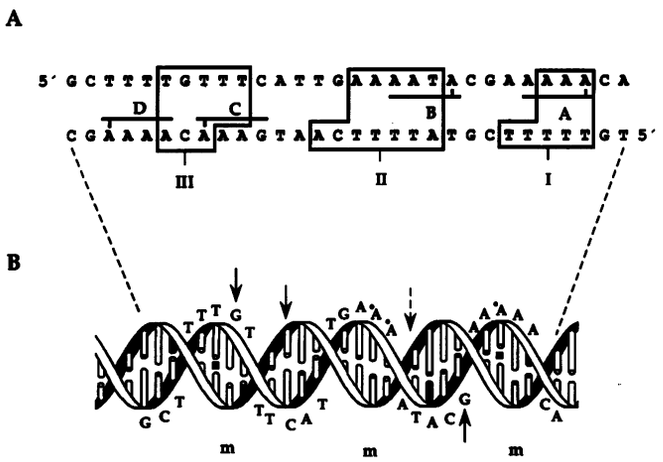

Figure 7. Summary of A-protein (4) and (+)-CC-1065 interactions with the *att* L3 site. (A) The duplex sequence of *att* L3 oligomer showing the hydroxyl-radical protection regions covering minor grooves I, II, and III and drug bonding sites (A–D) indicating the drug overlap in the minor grooves. (B) Ribbon diagram of the same region showing the proposed loci for bending into minor (m) grooves I and III by *solid squares* within base pairs. Guanine (strong) and adenine (weak) protection sites (4) are shown as *solid* and *broken* arrows, respectively. Drug bonding sites are shown as *filled bases*, and adenines showing methylation protection but not interference (4) are marked with *solid dots*. The sequence of the upper (+) strand is shown.

CC-1065 modification of A-tracts occurs predominantly at the 3′ adenine, most probably due to the special reactivity of the bp at the 3′ junction site of the A-tract. As a structural consequence of covalent modification of an A_5-tract, the magnitude of bending into the minor groove is increased from about 14° to about 25°, and the locus of bending is moved about 0.5 bps to the 3′ of the center of the A-tract (21). Because of the well-characterized and unique structure of the (+)-CC-1065-entrapped/induced bend in DNA and its structural similarity to A-tract bending, this drug is potentially a very useful probe for monitoring protein-induced DNA bending, which involves A-tracts or other sequences that may bend into the minor groove. In principle, (+)-CC-1065 modification of A-tracts that are the site of protein-induced bending should be tolerated or even enhanced sites for protein binding, provided the drug does not disrupt important protein–DNA contacts. Conversely, protein-induced bending into the minor groove of DNA, which takes advantage of intrinsically bent sequences such as A-tracts, should still be (+)-CC-1065 covalent modification sites in the presence of protein. On the other hand, if protein binding to a potential (+)-CC-1065 bonding sequence prevents bending of the DNA into the minor groove or sterically occludes drug binding within the minor groove, drug bonding will be inhibited. Likewise, it is also expected that (+)-CC-1065 modification of a sequence that is not the site of a protein-induced bend into the minor groove will be detrimental to protein binding. Based upon the sequence of *att* L3 site, which has several potential A-tract bending sites, and the observation that binding of A-protein results in a 90° overall bend (14), a series of experiments was designed to use (+)-CC-1065 and its structural analogs as structural probes to determine the location of sites of A-protein-induced minor groove bending within this sequence.

We have shown in this study that (+)-CC-1065 covalently bonds to four sites on *att* L3 DNA. One can see (Figure 7A) that on naked DNA these four drug bonding sites (A–D) interact within or partially overlap each of the three minor grooves (labeled I–III) that show protection by Mu A-protein to hydroxyl-radical attack (4). These covalent modification sites are shown (filled bases) on the ribbon model of L3 in Figure 7B. Interference experiments show that (+)-CC-1065 bonding within minor grooves I and III (sites A, C, D) *does not* significantly interfere with A-protein binding (Figures 4 and 5), in agreement with previous results where methylation of adenines in these regions did not interfere with protein binding (4). Since the drugs induce or trap bends into the minor groove, these data show that the A-protein tolerates such bends over minor grooves I and III. Significantly, *att* L3 modified with drug at site A within minor groove I was a *preferred* site for binding of A-protein, in comparison to non-drug-modified *att* L3 (Figure 5). This site was also reactive to the drug after A-protein binding. So also, modification of site D in minor groove III was not only well tolerated by the A-protein but this site was a *preferred* site for bonding of analogs of (+)-CC-1065, subsequent to binding of A-protein (Figure 6B). Besides showing that drug modification in minor grooves I and III is not detrimental to protein binding, these results also show that the two minor grooves are 'open' to the drugs upon A-protein binding and therefore not protected by the protein. At first glance, the latter inference appears to be inconsistent with the hydroxyl-radical and MPEβFe (II) footprinting data showing protection of these two minor grooves by the protein (4). The apparent paradox is explained by the data showing preference of drug bonding to sites in minor grooves I and III, either before (for site A) or after (for site D) A-protein binding, which suggest that these sites are bent or compressed into the minor groove by the A-protein. Hydroxyl-radical attack is known to be strongly influenced by the width of the minor groove (24, 25). Very likely, such narrowing also prevents MPE, which intercalates via the minor groove (26), from entering this region, thus rationalizing the apparent conflicting data.

A quite different situation occurs in the central minor groove II, where drug bonding strongly interferes with subsequent protein binding. After A-protein binding, site B was the only site unavailable for drug bonding (Figure 6A). This suggests that either Mu A-protein binds within the minor groove to sterically preclude drug access to its bonding site or that the protein prevents minor groove compression, which is required for stabilizing the transition state for the covalent reaction with (+)-CC-1065. The results of the previous hydroxyl-radical footprinting experiments are only consistent with the first interpretation (4). We note that methylation of one adenine in this region (Figure 7B, broken arrow) showed weak interference at *att* L3, while two other adenines (Figure 7B, solid dots) were protected from methylation at all *att* sites (4). These drug interference and protection data are consistent with partial penetration of minor groove II by A-protein.

The DNA bending sites within minor grooves I and III are phased approximately two helical turns apart, if we assume the locus of bending at minor grooves I and III is at the center of bonding site A and at the center of bonding sites C and D, respectively (see Figure 7B). Independent support of this assumption is provided by examining the hydroxyl-radical protection data (4), where the maxima of protection profiles in minor grooves I and III are very close to those deduced from the drug bonding data. Assuming that binding of protein does not significantly change the pitch of the DNA helix, these two flanking minor groove bends could be sufficient to produce a

cumulative bend of up to about 60° towards the protein. Mu A-protein produces an overall 90° bending of L3; therefore, an additional 30° needs to be accounted for, exclusive of that associated with minor grooves I and III. One possible explanation is that at least part of this additional 30° bend occurs into minor groove II, which is phased approximately equally between minor grooves I and III. However, other possible explanations include bending into the major groove on either side of minor groove II, which would also produce additive bending with minor grooves I and III. We note that only a 60° bend has been observed at *att* L2 (14), which is missing the two important guanines in the left major groove and shows a hydroxyl radical footprint only over minor grooves I and II (4). The loss of a 30° bend into minor groove III at L2, is consistent with our data.

We propose that the A-protein interacts within major grooves positioned on either side of minor groove II, explaining the methylation protection and interference data (4) showing the importance of three guanines in the two major grooves. If the interaction of A-protein within these major grooves is proposed to cause compression of minor grooves I and III and corresponding bending into these grooves, then it seems likely that A-protein penetration into the major groove on the back side of the helix is required. The strong methylation interference in the major groove behind site III is consistent with this proposal, but unfortunately the absense of a similar reporter guanine at site I does not provide data on this site. In addition to minor groove interactions, Mu A-protein is proposed to straddle minor groove II with limited penetration, possibly producing a bending of this minor groove. Although there is no obvious dyad symmetry in the *att* L3 site, the sequence ladder and ribbon diagram (Figure 7A and B) do show elements of symmetry around the minor groove II. Both flanking major grooves have some equivalent methylation protection sites, and the minor groove compression sites I and III are symmetrically opposed, as are the drug bonding sites. This model for bending at site L3 can be extended to most *att* sites. At site R1, however, we note, that a GC-rich sequence is present at minor groove III (4).

In summary, we have exploited the known DNA bonding and bending properties of (+)-CC-1065 and structurally related compounds to identify the probable sites of two Mu A-protein-induced DNA bends into minor grooves I and III at *att* L3. A third bending site, possibly into the minor groove II, is proposed. Overall, these bends would likely result in bending the DNA towards the protein. Our study highlights the usefulness of these antitumor drugs as probes for DNA–protein interactions and especially for protein-induced DNA bends. While their utility is limited by the availability of (+)-CC-1065 bending sites within the protein recognition sequence and protein-induced bending into the minor groove, these drugs should serve as general tools for overcoming limitations of hydroxyl radical footprinting by discriminating protein protection sites from minor groove DNA bending sites as shown here for the Mu transposase.

ACKNOWLEDGEMENTS

This research was supported by grants from the National Institutes of Health (GM 33247) to Rasika Harshey and from the Public Health Service (CA-49751), the Welch Foundation, and the Burroughs Wellcome Scholars Program to Laurence Hurley. We are grateful to David M.Bishop for editorial assistance and preparation of the manuscript.

REFERENCES

1. Symonds, N., Toussaint, A., Van de Putte, P., and Howe, M. M. (eds.) (1987) *Phage Mu*, Cold Springs Harbor Labs, New York.
2. Mizuuchi, M. (1992) *Ann. Rev. Biochem.*, **61**, 1011–1051.
3. Craigie, R., Mizuuchi, M., and Mizuuchi, K. (1984) *Cell*, **39**, 387–394.
4. Zou, A., Leung, P. C., and Harshey, R. M. (1991) *J. Biol. Chem.*, **266**, 20476–20482.
5. Leung, P. C., Teplow, D. B., and Harshey, R. M. (1989) *Nature*, **338**, 656–658.
6. Mizuuchi, M., and Mizuuchi, K. (1989) *Cell*, **58**, 399–408.
7. Allison, R. G., and Chaconas, G. (1992) *J. Biol. Chem.*, **267**, 19963–19970.
8. Craigie, R., Arndt-Jovin, D. J., and Mizuuchi, K. (1985) *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 7570–7574.
9. Lavoie, B. D., and Chaconas, G. (1990) *J. Biol. Chem.*, **265**, 1623–1627.
10. Surette, M. G., Buch, S. J., and Chaconas, G (1987) *Cell*, **49**, 253–262.
11. Lavoie, B. D., Chan, B. S., Allison, R. G., and Chaconas, G. (1991) *EMBO J.*, **10**, 3051–3059.
12. Mizuuchi, M., Baker, T., and Mizuuchi, M. (1992) *Cell*, **70**, 303–311.
13. Travers, A. A. (1990) *Cell*, **60**, 177–180.
14. Kuo, C.-F., Zou, A., Jayaram, M., Getzoff, E. D., and Harshey, R. M. (1991) *EMBO J.*, **10**, 1585–1591.
15. Hurley, L. H., Reynolds, V. L., Swenson, D. H., and Scahill, T. (1984) *Science*, **226**, 843–844.
16. Reynolds, V. L., Molineux, I. J., Kaplan, D., Swenson, D. H., and Hurley, L. H. (1985) *Biochemistry*, **24**, 6228–6237.
17. Hurley, L. H., Warpehoski, M. A., Lee, C.-S., McGovren, J. P., Scahill, T. A., Kelly, K. C., Wicnienski, N. A., Gebhard, I., and Bradford, V. S. (1990) *J. Am. Chem. Soc.*, **112**, 4633–4649.
18. Hurley, L. H., Lee, C.-S., McGovern, J. P., Mitchell, M., Warpehoski, M. A., Kelly, R. C., and Aristoff, P. A. (1988) *Biochemistry*, **27**, 3886–3892.
19. Warpehoski, M. A., and Hurley, L. H. (1988) *Chem. Res. Toxicol.*, **1**, 315–333.
20. Lin, C. H., Sun, D., and Hurley, L. H. (1991) *Chem. Res. Toxicol.*, **4**, 21–26.
21. Sun, D., Lin, C. H., and Hurley, L. H. (1993) *Biochemistry*, **32**, 4487–4495.
22. Koo, H. S., Wu, H. M., and Crothers, D. M. (1986) *Nature*, **320**, 501–506.
23. Lee, C.-S., Sun, D., Kizu, R., and Hurley, L. H. (1991) *Chem. Res. Toxicol.*, **4**, 203–213.
24. Burkhoff, A. M., and Tullius, T. D. (1987) *Cell*, **48**, 935–943.
25. Burkhoff, A. M., and Tullius, T. D. (1988) *Nature*, **331**, 455–457.
26. Hertzberg, R. P., and Dervan, P. B. (1984) *Biochemistry*, 23, 3934–3945.