Calicheamicin−DNA Recognition: An Analysis of Seven Different Drug−DNA Complexes

Allison Kalben,† Santona Pal,† Amy Hamilton Andreotti,† Suzanne Walker,† David Gange,‡ Kaustav Biswas,† and Daniel Kahne*,†

Contribution from the Department of Chemistry, Princeton University, Princeton New Jersey 08544, and American Cyanamid Company, Agricultural Research Division, Princeton, New Jersey 08540

Received February 11, 2000

Abstract: A study of calicheamicin γ_1^I complexed to seven different recognition sites is presented. The recognition sites encompass a range of oligopyrimidine sites that present different topological features in the minor groove. Intermolecular NOE networks for the different calicheamicin-DNA complexes show that the drug binds in the *same mode* to each recognition site. Calicheamicin binding also induces a set of characteristic conformational changes in the DNA in each complex that maximize the complementarity of the fit between calicheamicin and the DNA. Based on an analysis of the different complexes as well as biochemical information on cleavage preferences, we propose that calicheamicin displays a shape-selective preference for pyrimidine tracts through an induced-fit mechanism. We predict that any carbohydrate that maintains the overall shape of the calicheamicin oligosaccharide will exhibit similar sequence selectivity. This hypothesis is supported by experiments on calicheamicin oligosaccharide analogues reported in the following contribution.

Introduction

A detailed analysis of the origins of DNA binding selectivity is necessary for the design of DNA binders with predictable specificity. Structural studies of various minor groove binders complexed with different DNA sequences have proven useful in understanding the principles of selectivity in minor groove recognition. In some cases, these studies have led to rationally designed novel minor groove binders. For example, structural studies of the netropsin/distamycin class of polypyrrole binders complexed to AT-rich DNA duplexes were carried out in the 1980s and provided important information on minor groove recognition.[1,2] However, the selectivity was not fully understood until Wemmer and co-workers obtained a complex of distamycin bound as a side-by-side dimer to DNA.[3] This structure provided a paradigm for the design of molecules that recognize specific minor groove sequences using networks of hydrogen bonds to contact each strand of the DNA. Dervan and co-workers have now shown how to design polyamide dimers to recognize virtually any short sequence of DNA.[4] These results demonstrate the advantages gained by structural studies on a range of drug−DNA complexes in obtaining a complete picture of the origins of binding selectivity.

Calicheamicin γ_1^I (Figure 1a) is an enediyne antitumor antibiotic that functions by cleaving DNA at pyrimidine-rich recognition sites, including TCCT, CTCT, and TTTT.[5,6] DNA cleavage is mediated by the enediyne moiety, which undergoes a rearrangement to form a reactive diradical in the presence of thiols. The cleavage selectivity is determined largely by preferential binding of the aryl tetrasaccharide moiety to pyrimidine sequences.[6−9] The minor groove at these pyrimidine sequences is topologically diverse and presents a variety of

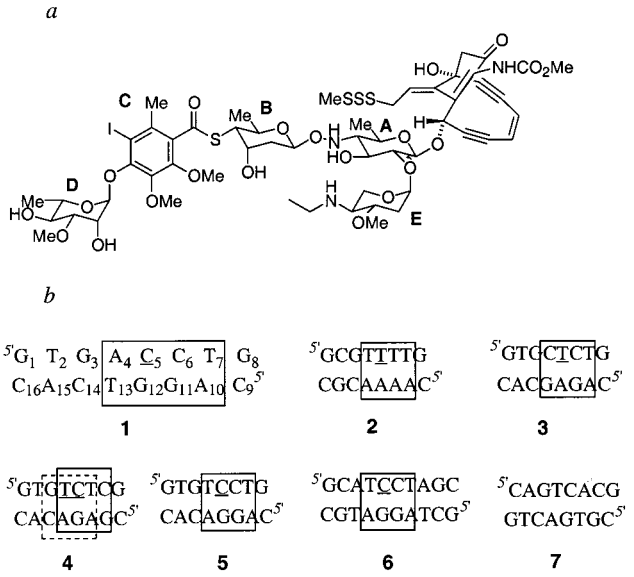

Figure 1. (a) Calicheamicin γ_1^I. (b) Sequences of the seven different DNA duplexes. The nucleotide numbers referred to in the text are shown for the ACCT duplex. The expected cleavage sites based on cleavage experiments on 10-mer duplexes containing the NMR sequences are underlined. The oligopyrimidine recognition sequences are encapsulated in black boxes. The second box in duplex **4** (dotted line) represents an overlapping binding site with a GC base pair in the third position. The duplex **7** is a "nonsense" sequence, with no putative binding site.

† Princeton University.
‡ American Cyanamid Co.

(1) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376−1380.
(2) Klevit, R. E.; Wemmer, D. E.; Reid, B. R. *Biochemistry* **1986**, *25*, 3296−3303.
(3) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723−5727.
(4) See: Urbach, A. R.; Szewczyk, J. W.; White, S.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 11621−11629 and references therein.
(5) Zein, N.; Sinha, A. M.; McGahren, W. J.; Ellestad, G. A. *Science* **1988**, *240*, 1198−1201.
(6) Walker, S.; Landovitz, R.; Ding, W.-D.; Ellestad, G. A.; Kahne, D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 4608−4612.
(7) Drak, J.; Iwasawa, N.; Danishefsky, S.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7464−7468.

10.1021/ja0005183 CCC: $19.00
Published on Web 08/18/2000

different functional group arrays. The mechanism by which calicheamicin recognizes these diverse sites is unclear.

Several studies in recent years have addressed the structural basis for the binding selectivity of calicheamicin. A model based on specific interactions has been proposed to explain calicheamicin recognition of GC-rich pyrimidine sites, particularly TCCT sites.[10,11] This model involves a critical contact between the aromatic iodine of calicheamicin and the 2-amino group of guanine. Calicheamicin also binds to TTTT sites with similar affinity, however.[6,12] Since this site does not contain any guanines, there is no iodine−amino contact. In fact, the pattern of functional groups at TTTT sites is significantly different from that at GC-rich sites. Given the differences in functional group arrays, it is not possible for calicheamicin to bind to TCCT, TTTT, CTCT, and other recognition sites by making the same set of contacts. The absence of common functional group arrays at the preferred pyrimidine-rich binding sites raises questions about any model for calicheamicin−DNA recognition based on specific directional interactions between the ligand and DNA. To rationalize the pyrimidine selectivity of calicheamicin, however, it is essential to establish first that the molecule binds in the same mode to all the specific recognition sites.

To date, three NMR studies of calicheamicin bound to GC-rich recognition sites (ACCT/TCCT) have been published.[13−16] In all of these studies, the principal emphasis was on identifying contacts between calicheamicin and the DNA and generating a model for the complex. There is substantial agreement on how calicheamicin binds to ACCT/TCCT sites, and a refined structure of a calicheamicin−TCCT complex has been reported by Patel and co-workers. However, neither this recent study nor the earlier studies adequately addressed the molecular basis for calicheamicin's pyrimidine selectivity. To do so requires having structural information on complexes that represent the range of calicheamicin recognition sites.

Below we present a comparison of calicheamicin bound to seven different DNA duplexes using NMR and biochemical cleavage data. On the basis of these studies, we show that calicheamicin binds to all of these sequences in the same mode. Calicheamicin is thus able to sense some structural feature common to oligopyrimidine tracts. We propose that the binding of calicheamicin to these sequences is a result of a shape-selective interaction in which the DNA conformation adapts to accommodate the ligand. In the following contribution, we report binding studies with calicheamicin oligosaccharide derivatives that support this hypothesis.[17]

Results

The Complexes. We prepared complexes of calicheamicin with seven different DNA sequences (one decamer and six octamer duplexes). Six of the duplexes (Figure 1b) contained four base pair sequences that have been identified as preferred calicheamicin recognition sites. The seventh duplex (**7**, Figure 1b) was a "nonsense" duplex, chosen from a pUC19 Nde1-Acc1 restriction fragment as a sequence that is not cleaved by calicheamicin. This nonsense sequence provided information about the spectral appearance of a nonspecific complex. Of the six other duplexes, two included a TCCT site. TCCT is widely regarded as the canonical calicheamicin binding/cleavage site. One of the TCCT duplexes was an octamer (**5**, Figure 1b) and the other was a decamer that has previously been studied by Paloma et al. (**6**, Figure 1b).[14] The four other duplexes contained the sequences ACCT, TTTT, CTCT, and TCTC, respectively (**1**, **2**, **3**, and **4**, Figure 1b).[18] All these sequences were identified in early studies on calicheamicin as preferred cleavage sites.[5,6,12]

To evaluate the quality of the spectral data before attempting full assignments, we prepared complexes of calicheamicin with the various duplexes at both a 0.5:1 and a 1:1 ratio. A comparison of the 1D NMR spectra of the nonsense duplex and the ACCT duplex upon titration with calicheamicin shows that there are clear differences between nonspecific and specific complexes (Figure 2). Upon addition of 0.5 equiv of calicheamicin, the DNA resonances in the ACCT complex double but remain sharp, indicating the presence of both free and bound duplexes in slow exchange. The free DNA resonances disappear completely upon the addition of 1.0 equiv of calicheamicin. In contrast, the 1D spectrum of the nonsense duplex shows broadened lines upon the addition of 0.5 equiv of calicheamicin, and the resonances remain broad after the addition of a full equivalent of calicheamicin. The line broadening suggests exchange between different binding modes, an interpretation confirmed by analysis of the NOESY spectrum of the 1:1 complex, which shows at least three different sets of cross-peaks. We have concluded that the presence of a single set of resonance lines at a 1:1 ratio of calicheamicin to DNA represents the formation of a single, specific 1:1 complex.

The octamers containing the CTCT, TTTT, and TCCT sites showed similar behavior to the ACCT octamer upon titration with calicheamicin, indicating the formation of specific calicheamicin−DNA complexes. The TCTC duplex, in contrast, showed peak doubling even after the addition of a full equivalent of calicheamicin, and an analysis of the NOESY spectra showed two distinct sets of cross-peaks. As discussed below, the TCTC duplex was designed to contain two overlapping calicheamicin binding sites, and the molecule exchanges between them. The decamer TCCT complex also showed significant line broadening, more than could be explained by the larger size of the duplex. The line broadening indicates more than one binding mode to the duplex. Paloma et al. have proposed that the aglycon exchanges between two different orientations and have suggested that this motion is required for double-stranded cleavage. However, this motion is not observed in any of the other specific complexes, including the octamer TCCT complex and Patel's hairpin TCCT complex.[19] Therefore, we suggest that the exchange behavior observed in the decamer complex reflects the choice of flanking base pairs. The flanking base pairs may influence whether the ligand binds tightly to one particular site or "slips" in the minor groove.

The four specific complexes, which included the octamers with the ACCT, CTCT, TTTT, and TCCT recognition sites (**1**,

(8) Aiyar, J.; Danishefsky, S. J.; Crothers, D. M. *J. Am. Chem. Soc.* **1992**, *114*, 7555−7557.
(9) Nicolaou, K. C.; Tsay, S.-C.; Suzuki, T.; Joyce, G. *J. Am. Chem. Soc.* **1992**, *114*, 7555−7557.
(10) Hawley, R. C.; Kiessling, L. L.; Schreiber, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 1105−1109.
(11) Li, T.; Zeng, Z.; Estevez, V. A.; Baldenius, K. U.; Nicolaou, K. C.; Joyce, G. F. *J. Am. Chem. Soc.* **1994**, *116*, 3709−3715.
(12) Mah, S. C.; Price, M. A.; Townsend, C. A.; Tullius, T. D. *Tetrahedron* **1994**, *50*, 1361−1378.
(13) Walker, S.; Murnick, J.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 7954−7961.
(14) Paloma, L. G.; Smith, J. A.; Chazin, W. Z.; Nicolaou, K. C. *J. Am. Chem. Soc.* **1994**, *116*, 3697−3708.
(15) Ikemoto, N.; Kumar, R. A.; Ling, T.; Ellestad, G. A.; Danishefsky, S. J.; Patel, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10506−10510.
(16) Kumar, R. A.; Ikemoto, N.; Patel, D. J. *J. Mol. Biol.* **1997**, *265*, 187−201.
(17) Biswas, K.; Pal, S.; Carbeck, J. D.; Kahne, D. *J. Am. Chem. Soc.* **2000**, *122*, 8413−8420.

(18) The ACCT complex has been reported by us previously (see ref 13).

(19) We have found that calicheamicin effects double-stranded cleavage of all the specific complexes. This exchange process is therefore not a general feature of calicheamicin binding and activity.

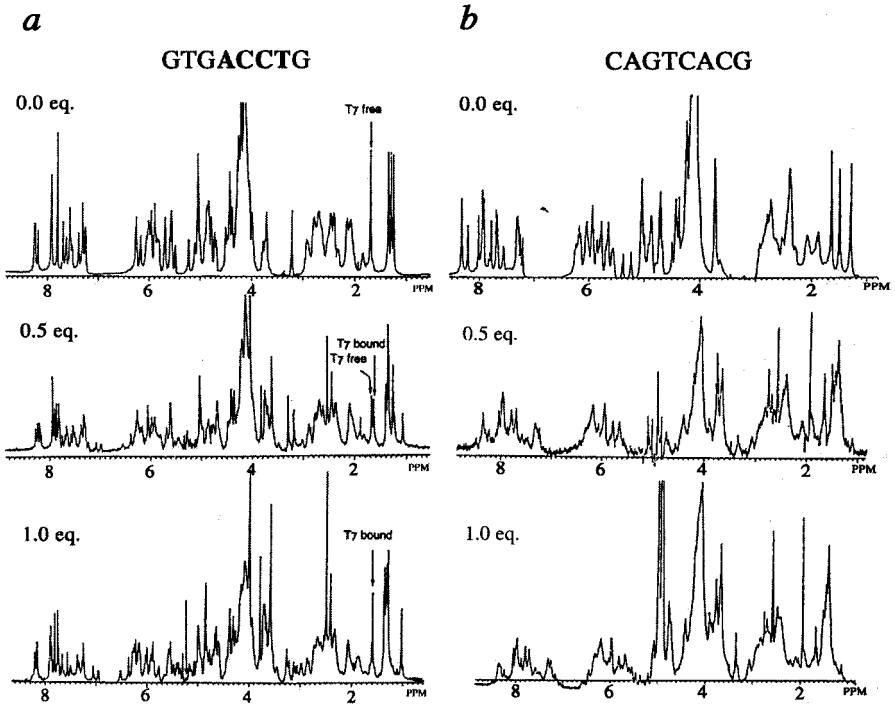

Figure 2. 1D ^{1}H NMR spectra in D_2O at 15 °C of the DNA duplexes: (a) GTGACCTG (containing the ACCT recognition site), and (b) CAGTCACG (the "nonsense" sequence), in the presence of 0, 0.5, and 1.0 equiv of calicheamicin, respectively.

2, **3**, and **5**, respectively, Figure 1b), were assigned using standard two-dimensional NMR techniques and structures were generated using restrained molecular dynamics. A comparison of these four complexes will be presented below. The other complexes will also be discussed since their behavior sheds additional light on how calicheamicin recognizes DNA.

Orientation of Calicheamicin in the Minor Groove. The position and orientation of calicheamicin relative to the recognition sequence in each of the four well-behaved complexes was established using intermolecular NOEs and chemical shift data. Calicheamicin binds in the same position and orientation to all four of the specific octamer complexes. The position of calicheamicin in the minor groove is illustrated for the ACCT complex in Figure 3. These two views demonstrate the orientation of the individual residues in calicheamicin with respect to each other and to the minor groove. The enediyne moiety of the aglycon faces into the groove, while the E ring protrudes out of the groove in an orientation that positions the protonated ethylamino group near the phosphate oxygens of cytosine 5. The A ring is oriented so that the C6 methyl group faces up and in toward the minor groove, making a contact with the adenine 15 H4′ proton. The B ring is oriented with the C3 hydroxyl facing in and up toward the purine strand and the C6 methyl group facing out toward the solvent. There are contacts between B ring hydrogens and cytosines 5 and 6. The C ring is oriented with the methyl and iodine substituents pointing toward the floor of the minor groove, the methyl group contacting cytosine 6. The methoxy substituents are exposed to the solvent, and the two methoxy groups point in opposite directions. The methoxy *meta* to the D ring glycosidic linkage contacts the backbone H4′ proton of thymine 13. The D ring C2 hydroxyl faces in toward the groove, while the C6 methyl faces the solvent. The D ring anomeric hydrogen makes a contact with thymine 13, while the C3 methoxy contacts guanine 12. These same contacts are observed in all the specific drug−DNA complexes, and are also consistent with previously reported data by us for the ACCT complex[13] and Patel and co-workers in their NMR study of a calicheamicin complex with a hairpin duplex containing the TCCT site.[16] The specific ligand−DNA NOEs common to the four well-behaved complexes are listed in Table 1. No intermolecular NOEs to the E ring or the aglycon could be identified, although their positions could be inferred based on the other NOEs. The similarity in the NOE networks in each complex indicates that the position of calicheamicin relative to the recognition sequence is similar in all the complexes.

Chemical shift changes also provide sensitive information on the position of calicheamicin in the groove. There are significant changes in the chemical shifts of the H1′, H3′, and H4′ protons of several ribose sugars on binding calicheamicin. The pattern and the magnitude of these shifts are similar in all the complexes. An especially dramatic chemical shift change ($\Delta\partial \sim -1.1$ to -1.3 ppm) is observed for the H4′ proton of nucleotide 13 in all of the complexes (Figure 4a). This upfield shift is due to the shielding effect of the aromatic C ring of calicheamicin. Since aromatic shielding effects fall off rapidly with distance from the center of the aromatic ring,[20] the remarkably similar shift changes for the H4′ resonance in all of the complexes indicate an essentially identical position of calicheamicin in the groove with respect to the recognition sequence.

The NOE and chemical shift data show that calicheamicin binds in the same position and orientation in all four well-behaved complexes. Overlays of 10 structures generated by simulated annealing for the ACCT and TTTT complexes show a similar mode of binding to both these DNA sequences, with comparable contacts between the drug and the recognition sites (Figure 5). The localized binding observed for the four specific complexes is not simply a function of the size of the duplex. An octamer duplex is sufficient to include at least six binding sites, if the duplex is read from both ends. The behavior of the nonsense complex shows that when there is not a strongly preferred binding site, calicheamicin exchanges between several available sites. Hence, the localized binding in the well-behaved

(20) Giessner-Prettre, C.; Pullman, B. *J. Theor. Biol.* **1970**, *27*, 87−95.

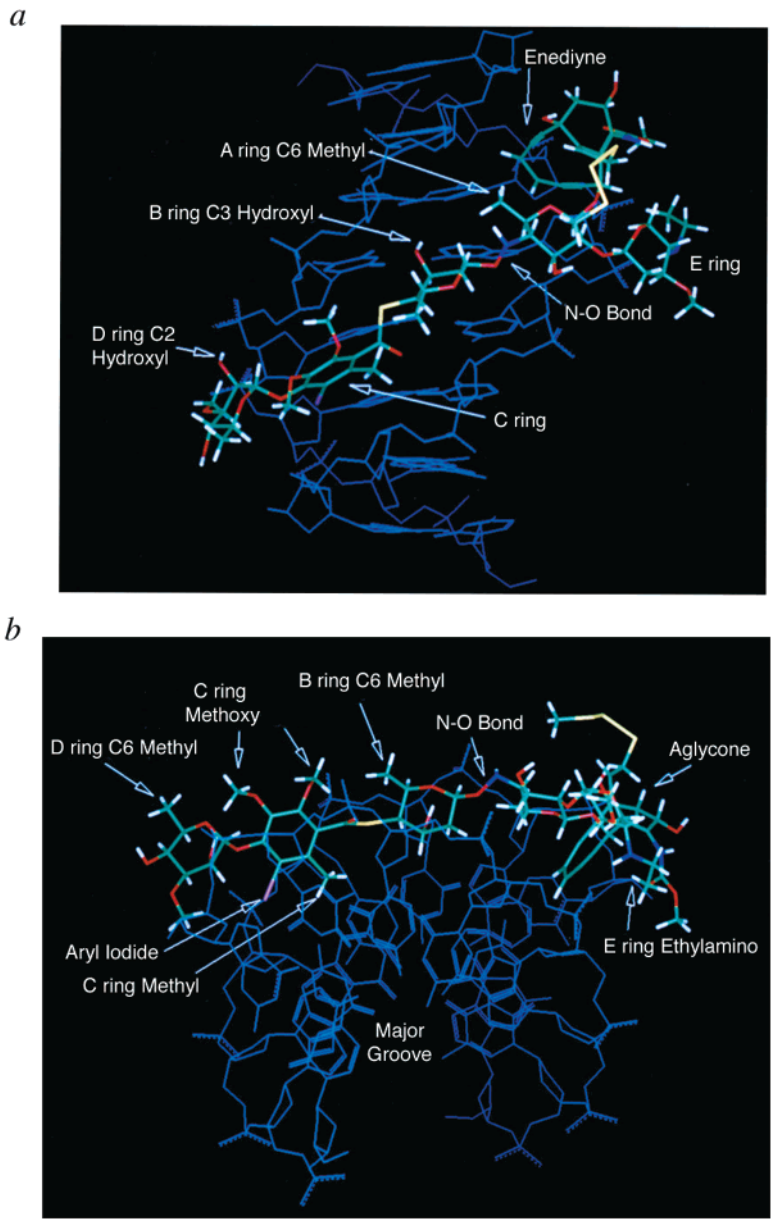

Figure 3. Two views of the calicheamicin−ACCT complex generated by a simulated annealing procedure. Salient features of calicheamicin are labeled, indicating the orientation of individual residues with respect to each other and to the minor groove. (a) Front view, with the pyrimidine strand to the right and below the oligosaccharide, (b) top view, demonstrating the inherent curvature of the molecule.

Table 1. Intermolecular NOEs Common to the TCCT, ACCT, CTCT, and TTTT Calicheamicin Complexes[a]

DNA protons	calicheamicin protons	DNA protons	calicheamicin protons
N5 H1′	B H2 (eq)	N12 H4′	C −OCH_3
N6 H1′	B H4	N13 H4′	D −H1
N6 H4′	B H4	N13 H4′	C −OCH_3 (m)
N6 H4′	B H6	N13 H5′	D −H1
N6 H1′	C −CH_3	N13 H5″	D −H1
		N15 H4′	A −H6

[a] "m" refers to the meta relationship of the methoxy group to the D ring glycosidic linkage. "eq." refers to the equatorial H2 proton of the B ring.

complexes reflects calicheamicin's strong selectivity for pyrimidine-rich sequences. This is notable because there are striking differences in the topology of the minor groove at these four sites. Any proposal for how selectivity is achieved must account for how calicheamicin can bind in the same position and orientation to sites that present very different functionality.

Conformational Changes upon Complexation. The free and bound conformations of calicheamicin and the DNA were analyzed for each of the octamer duplexes. The conformation of calicheamicin does not appear to change significantly upon binding to any of the DNA duplexes as indicated by the sugar coupling constants as well as inter- and intraresidue NOEs.[13,21]

In contrast, there is NMR evidence that distinctive conformational changes occur in the DNA upon binding calicheamicin. The deoxyribose sugars in DNA are one of the best indicators of local conformation. B form DNA is characterized by a C2′-*endo* sugar pucker in which the C2′ atom is displaced above the plane formed by C1′, O1′, and the C4′ deoxyribose ring atoms. The vicinal coupling constants for a C2′-*endo* conformation are large for H1′−H2′ and H2′−H3′, smaller for H1′−H2″ and H3′−H4′, and extremely small for H2″−H3′. Although small couplings are difficult to measure accurately, it is possible

(21) Walker, S.; Valentine, K. G.; Kahne, D. *J. Am. Chem. Soc.* **1990**, *112*, 6428−6429.

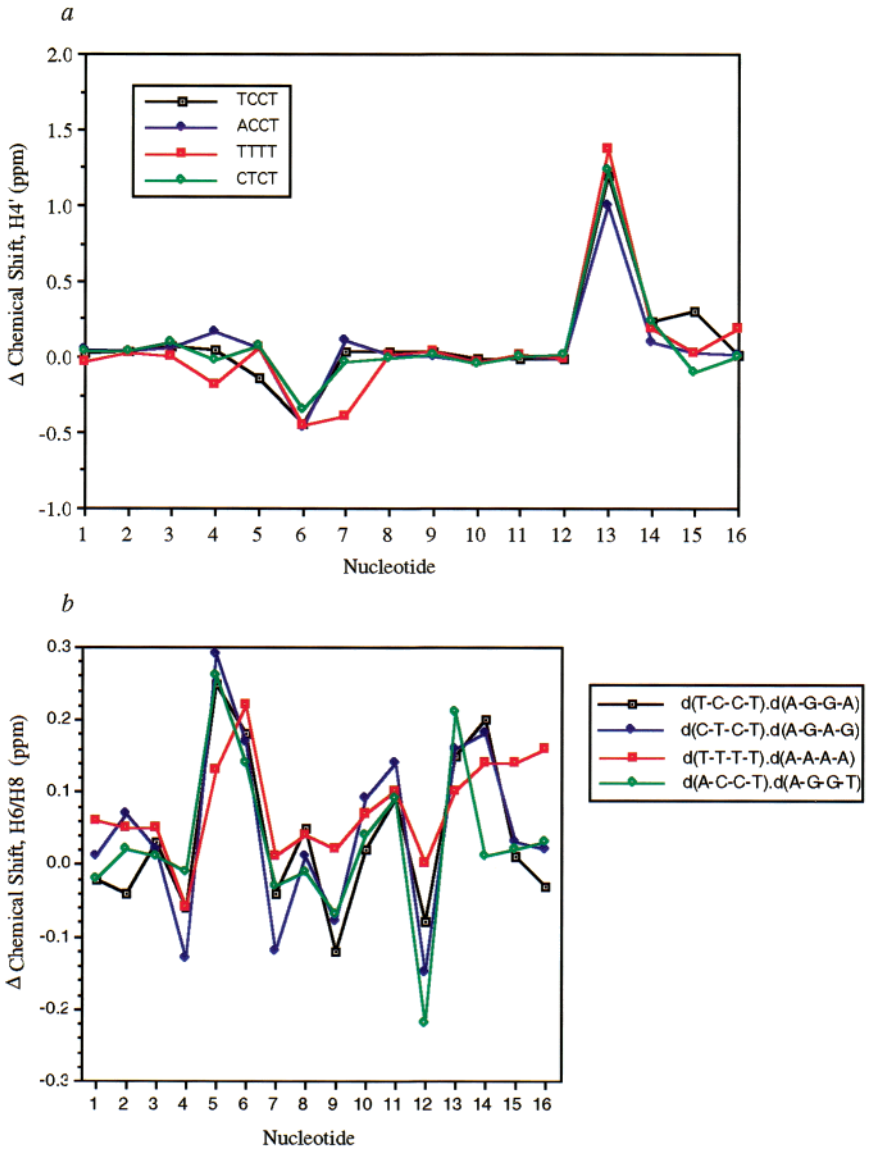

Figure 4. (a) Chemical shift changes ($\delta_{free} - \delta_{bound}$) for the H4′ proton for each nucleotide (1−16) for all four calicheamicin−DNA complexes. Note the large chemical shift change (>1 ppm) that is localized to nucleotide 13 for each recognition site. (b) Chemical shift changes for the H6/H8 protons that occur upon complexation of calicheamicin. The direction of the shifts are similar for each complex although the magnitudes differ.

to draw conclusions about the sugar pucker on the basis of the relative sizes of the couplings in each ribose sugar, which are evident from inspection of the cross-peaks.[22] The deoxyribose torsion angles, which are reflected in the vicinal coupling constants, can be described by a one-dimensional parameter called the pseudorotation angle (P).[23] The pseudorotation angle for a C2′-*endo* conformation is centered around 162° (144° < P < 180°). The pseudorotation angle ranges can be converted into dihedral angle restraints.[24] We have used COSY cross-peak patterns to estimate the range of pseudorotation angles for each residue. Because chemical exchange at ligand−DNA interfaces (e.g., due to dissociation) does have some effect on the line widths, and thus on the cross-peak intensities, the estimated ranges were deliberately broad.

With one exception, described below, the internal sugars of the free DNA duplexes display the cross-peak pattern expected for B form DNA. Upon complexation with calicheamicin, however, specific deoxyribose sugars (N5, N6, N14, and N15) undergo conformational changes. The changes are manifested in small couplings between H1′ and H2′ and large couplings between H1′ and H2″ and between H3′ and H4′, as deduced from the COSY cross-peaks (Figure 6a). The changes in nucleotide 6 are especially pronounced in the ACCT and CTCT complexes, which both show distinct cross-peaks between H2″ and H3′ (Figure 6b) in addition to aberrant couplings for other correlations. The H2″−H3′ cross-peak is only observed in sugars having a pseudorotation angle of less than 90°. The pseudorotation angles of each ribose sugar in the octamer ACCT, TCCT, CTCT, and TTTT complexes, estimated from the relative couplings reflected in the COSY cross-peaks, are similar, showing that the different recognition sites undergo a characteristic set of conformational adaptations upon binding calicheamicin (Figure 7).[25] Patel and co-workers have likewise reported conformational changes in some of the ribose sugars in a TCCT-containing hairpin duplex upon binding calicheamicin.[16]

(22) Wuethrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.

(23) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1973**, *95*, 2333−2344.

(24) de Leeuw, H. P. M.; Haasnoot, C. A. G.; Altona, C. *Isr. J. Chem.* **1980**, *20*, 108−126.

(25) Pal, S. Ph.D. Thesis, Princeton University, 1997.

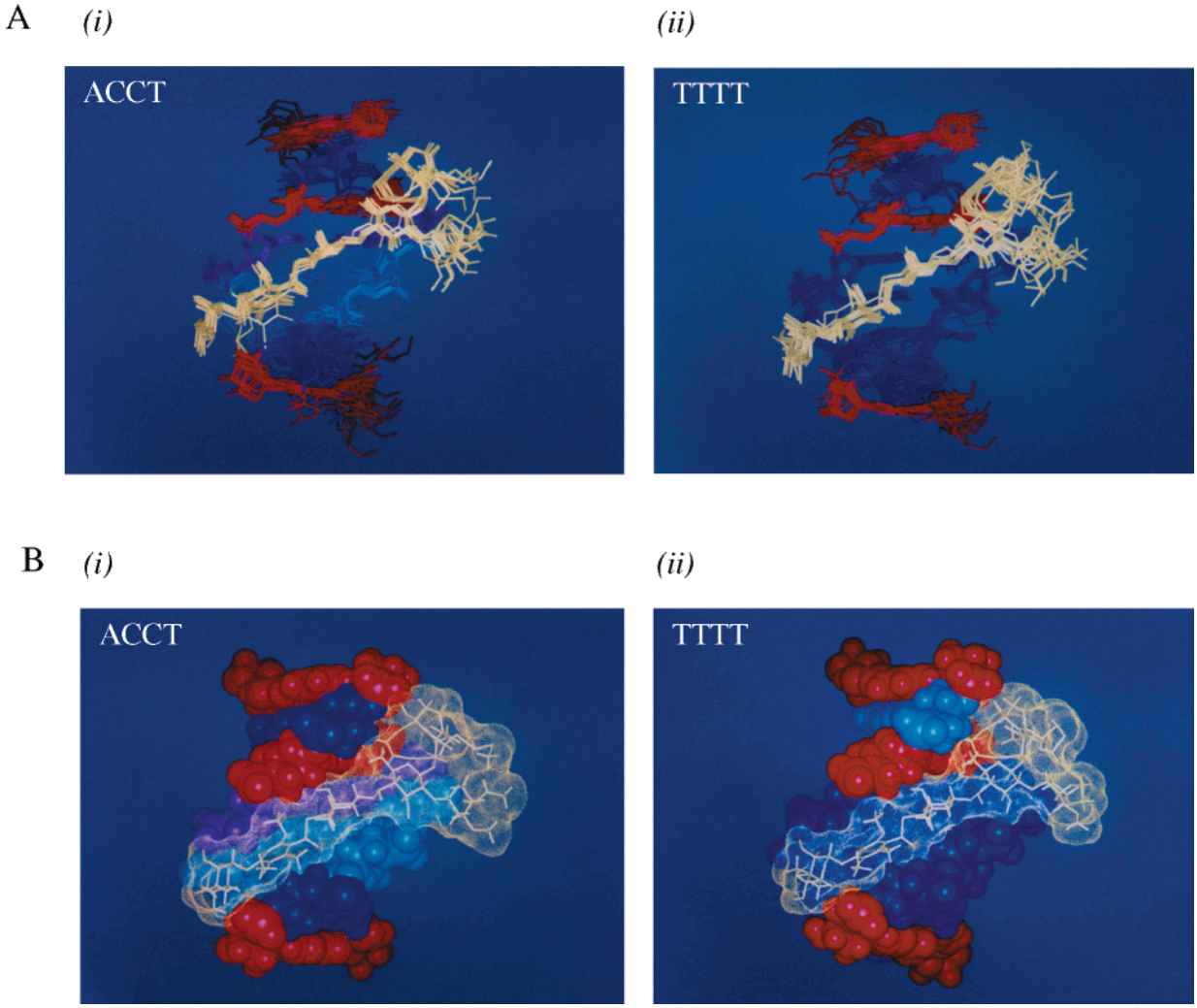

Figure 5. (A) Overlays of 10 structures generated by simulated annealing for the (i) calicheamicin−ACCT and (ii) calicheamicin−TTTT complexes. Calculated RMSDs for the complexes, excluding the terminal base-pairs and sodium atoms, were 1.106 Å for ACCT and 1.135 Å for TTTT, respectively. (B) Structures of the (i) ACCT and (ii) TTTT complexes. The GC base pairs are in red, the CG base pairs in blue, the AT base pairs are purple and the TA base pairs are in dark blue (the first base in each pair refers to the pyrimidine-rich strand). Calicheamicin is in yellow.

Additional evidence that calicheamicin induces a characteristic conformation in the DNA upon binding comes from a comparison of the free and bound conformations of the CTCT duplex. In the free duplex, the C4 ribose sugar has a pseudorotation angle below 90°, and the C6 sugar has a pseudorotation angle between 144 and 180°, which is characteristic of a C2′-*endo* conformation. In the complex, the C4 sugar flips into a C2′-*endo* conformation, with a pseudorotation angle between 144 and 180°, while the C6 sugar adopts an angle below 90°. Hence, the CTCT duplex actually changes from a conformation containing an "anomalous" sugar pucker at nucleotide 4 to a conformation with a "normal" C2′-*endo* sugar pucker at nucleotide 4 but distorted sugars at positions 5, 6, and 15, just as in the other complexes. Thus, despite significant differences in the base pair composition and in the specific conformations of the free duplexes, calicheamicin binding induces similar conformations in the DNA upon binding.

Chemical shift changes provide more evidence for characteristic conformational changes in the DNA upon calicheamicin binding. The H8/H6 resonances, which are located in the major groove close to the phosphate backbone, are sensitive to changes in backbone torsion angles and shifts in the bases. The pattern of the H6/H8 shift changes for all four specific complexes is very similar, which indicates similar conformational changes (Figure 4b). The characteristic conformational changes of the DNA, together with the absence of conformational changes in calicheamicin, suggest that the DNA adapts to accommodate the drug.

Intermolecular Contacts. Both Paloma et al.[14] and Patel[15,16] have attributed the selectivity of calicheamicin for TCCT sequences to specific contacts between functional groups on calicheamicin and the DNA. To compare the calicheamicin−DNA contacts in a range of pyrimidine recognition sites, we used a simulated annealing procedure to generate structures of the ACCT, CTCT, and TTTT calicheamicin−DNA complexes. This effort required developing AMBER-compatible force field parameters to treat the N−O bond in the calicheamicin oligosaccharide.[26,27] Homans had previously developed AMBER-compatible parameters for carbohydrates that were used in the simulations.[28,29] The structures generated permit us to identify hydrogen bonds and other contacts that are not immediately evident from the NMR data.

(26) Walker, S.; Yang, D.; Kahne, D.; Gange, D. *J. Am. Chem. Soc.* **1991**, *113*, 4716−4717.
(27) Walker, S.; Gange, D.; Gupta, V.; Kahne, D. *J. Am. Chem. Soc.* **1994**, *116*, 3197−3206.
(28) Homans, S. W. *Biochemistry* **1990**, *29*, 9110−9118.

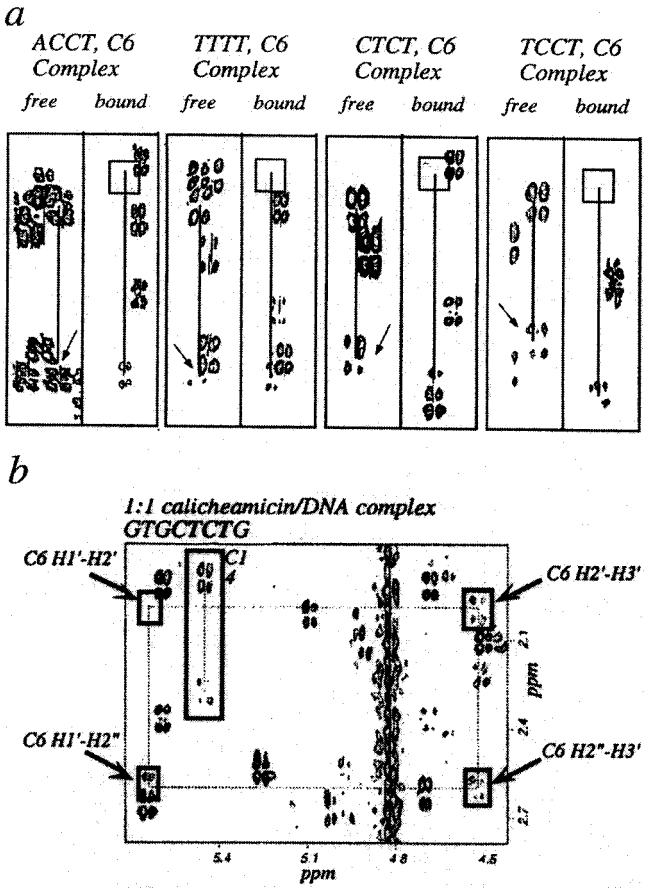

Figure 6. (a) Portions of the two-dimensional DQF-COSY spectra for the four well-behaved complexes. For each recognition site the H1′−H2′ and the H1′−H2″ COSY cross-peaks are shown for nucleotide 6 in both the free and bound samples. The H1′−H2′ peaks are boxed and the H1′−H2″ peaks are indicated with arrows. In all cases, the H1′−H2″ COSY cross-peak intensity of bound DNA is larger than the H1′−H2′ cross-peak intensity of free DNA. This indicates a shift away from the B-form DNA upon binding of calicheamicin. (b) Expansion of the DQF-COSY plot of the 1:1 complex of calicheamicin bound to the CTCT containing duplex. The peaks of interest are boxed and labeled. The existence of a COSY cross-peak between the H2″ and H3′ protons of C6 is indicative of further distortion away from B form DNA. It should be noted that the TTTT complex does not show this peak suggesting that the required distortion for this recognition site may be less. The H1−H2′/H2″ cross-peaks corresponding to C14 are boxed and represent the COSY cross-peak pattern observed for standard B form DNA.

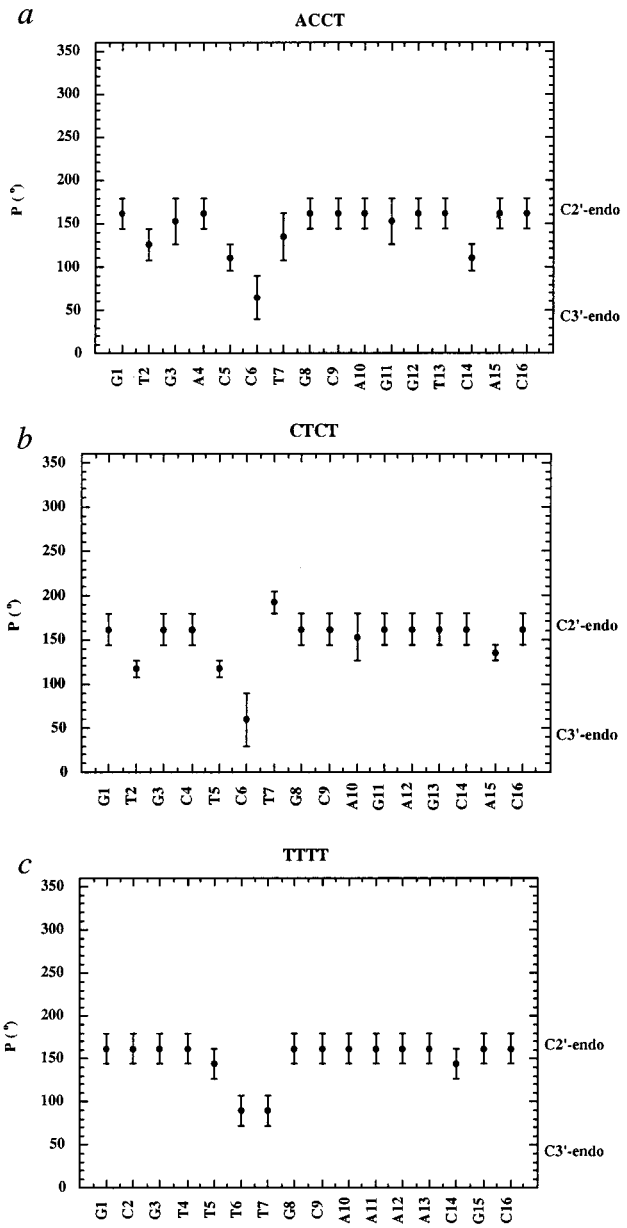

Figure 7. Distribution of the sugar pucker pseudorotation angle P for (a) the ACCT, (b) the CTCT, and (c) the TTTT complexes with calicheamicin.

The models show that there are several hydrogen bonds from calicheamicin to the DNA which are present in all of the complexes. For example, all of the complexes show hydrogen bonds between the A ring C3 hydroxyl and the phosphate at N5−N6, between the D ring C2 hydroxyl and the phosphate at N12−N13, and between the aglycon C8 hydroxyl and the N16 phosphate. In addition, the positively charged ethylamino group of the E ring interacts with the N5 phosphate in all the complexes. These contacts to the phosphate backbone help orient and anchor calicheamicin in the minor groove (see Figure 3).

In addition to the phosphate contacts that are conserved in all of the complexes, there are sequence-dependent contacts between calicheamicin and certain recognition sequences that may play a role in affinity. For example, there is a hydrogen bond between the B ring C3 hydroxyl in calicheamicin and the G13 amino group in the CTCT sequence. There is also an interaction between the iodo substituent on calicheamicin and the 5′ guanine amino group in the CTCT, TCCT, and ACCT complexes. These and other sequence-dependent interactions may contribute to binding to particular sites; however, since these interactions are only present in some complexes, they cannot explain the underlying pyrimidine selectivity.

Oligopyrimidine versus Guanine Amino Selectivity. Calicheamicin prefers pyrimidine sequences, but some studies have suggested a contact from the aryl iodide to a guanine amino group contributes significantly to the binding affinity (and hence, selectivity).[10,11] To gain more insight into the relative importance of these factors in binding site selection, we prepared a complex of calicheamicin with the duplex d(GT**GTCTC**G)−d(CGAGACAC), which was designed to contain two overlapping binding sites, TCTC and GTCT. The TCTC site contains

(29) Although Homans and others have argued that the torsional preferences about CCOC glycosidic bonds (the ψ angle in carbohydrate nomenclature) can be adequately simulated without explicitly including torsional terms, we have obtained better results with explicit parameters that produce a better fit between the molecular mechanics torsional profile for $CH_3CH_2OCH_3$ and experimental results.

a tract of four pyrimidines with a TA base pair at the position that contacts the aryl−iodo substituent. The GTCT site has a GC base pair at this position, but contains only three pyrimidines.

The GTCTC complex exhibited two sharp sets of proton resonances in a ∼3:1 ratio, representing two binding modes in slow exchange. Assignment was possible only for the major binding mode. Several key NOEs showed that the major binding mode corresponds to complexation of the TCTC site. The minor binding mode could not be adequately characterized by NMR due to its relatively low concentration and overlap problems. However, cleavage studies on a decamer duplex, d(**GTGTC-TCG**GC)−d(GC**CGAGACAC**) containing the octamer sequence studied by NMR show that calicheamicin cleaves at both the GTCT and TCTC sites. A longer duplex, containing an ATCTC sequence, also showed two cleavage sites, corresponding to the overlapping ATCT and TCTC binding sites. Thus, the cleavage data suggests that the minor binding mode observed in the NMR complex is calicheamicin bound to the GTCT site. Hence, in this competition between an all-pyrimidine site and a site presenting a GC base pair positioned appropriately to satisfy the guanine−amino interaction, calicheamicin shows a preference for the all-pyrimidine site.

Discussion

The Structural Basis for Selectivity. Calicheamicin cleaves DNA selectively at pyrimidine-rich sequences. NMR studies, footprinting, and cleavage inhibition analyses strongly support the hypothesis that the cleavage preferences reflect thermodynamic binding affinities.[8,9,11,13−16,30] TCCT is widely regarded as the "best" recognition site, and numerous studies have been carried out on this site in an attempt to rationalize TCCT-selective binding. However, calicheamicin also cleaves a variety of other sequences, including ACCT, CTCT, and TTTT. Furthermore, the available evidence suggests that the affinity of calicheamicin for these other sites is comparable to its affinity for TCCT sites. For example, in a Sal I−Bam H1 restriction fragment from pBR322, calicheamicin cleaves only two sites strongly—TCCT and ACCT.[5] Moreover, in a synthetic duplex containing only two pyrimidine sequences, TCCT and TTTT, calicheamicin cleaves at both of these sites but not elsewhere. There is a difference in the extent of cleavage of the two sites, with the TCCT site preferred, but Townsend and Tullius have estimated that the difference in affinity for the two sites is only ∼0.8 kcal/mol.[30] Experiments reported in the following paper show an even smaller difference for the binding of the calicheamicin oligosaccharide to TCCT and TTTT sites.[17] In contrast, the difference in affinity of calicheamicin for TTTT sites and mixed sequences of DNA—that is, sequences that do not contain at least three pyrimidines—is significantly greater than the difference between TCCT and TTTT.[31] Hence, in attempting to rationalize the binding selectivity of calicheamicin, it is important to consider not only how it binds to TCCT sites, but also how it recognizes seemingly disparate pyrimidine-rich sequences.

It is possible for a molecule to recognize DNA using different binding modes.[3] Therefore, we thought it necessary to compare the structures of a set of different calicheamicin−DNA complexes. We found that calicheamicin forms specific complexes with a range of DNA duplexes containing pyrimidine-rich recognition sites. The specific complexes are characterized by a single predominant set of relatively sharp resonance lines at a 1:1 calicheamicin:DNA ratio. The less specific complexes, in comparison, exhibit a range of different behaviors. In the complex with the GTCTC sequence, there were two sets of cross-peaks in the 2D spectra and the resonance lines in the 1D spectrum were broadened. In a duplex that did not contain any putative recognition sites, at least three different binding modes could be observed in the 2D spectra. In the complex with the TCCT-containing decamer, the resonance lines were unusually broad and there were some exchange cross-peaks and aberrant cross-peak intensities in the 2D spectra, indicating considerable motion at the binding interface (Figure 8). Although this duplex contains a preferred recognition site, and there are some indications that calicheamicin binds in the vicinity of the recognition site, we did not analyze the complex in detail because the quality of the NMR data is compromised by exchange.[32] Instead, we analyzed the four complexes displaying the hallmarks of high specificity in order to determine if calicheamicin uses the same binding mode in recognizing the different sequences. Our results show that calicheamicin binds in exactly the same orientation and makes the same set of contacts—with minor exceptions related to base pair composition—to all four sequences. Therefore, pyrimidine-rich sequences must share some structural features which calicheamicin is capable of sensing and exploiting in choosing its binding sites.

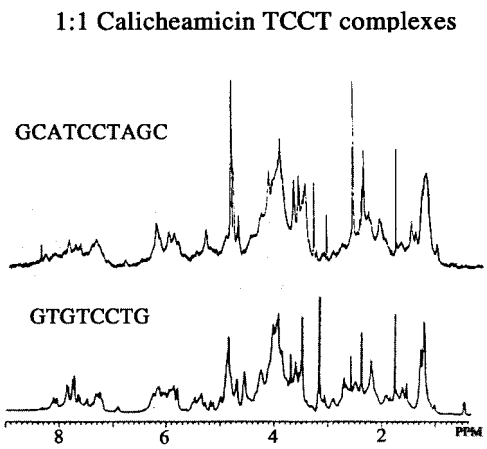

Figure 8. 1D ^{1}H NMR spectra of calicheamicin bound to two TCCT containing duplexes with different flanking sequences. The 10-mer duplex is identical to that used in the Paloma et al. study (see text). The 8-mer duplex has a flanking sequence similar to those used for the other recognition sites in this study. Both complexes were prepared in an identical manner and the spectra were acquired at 10 °C.

Two other groups have carried out NMR studies of calicheamicin bound to DNA. Both studies were of calicheamicin bound to TCCT sites. The binding mode identified by these groups to TCCT sites is similar to the binding mode reported here and also in a previous report. The authors, however, concluded that the sequence selectivity was due to a set of specific contacts made between calicheamicin and the TCCT sites.[14−16] Some of the contacts reported to be critical for selectivity are found only in TCCT and closely related sites. These contacts may contribute to the 0.8 kcal/mol difference in binding between TCCT and TTTT sites,[30] but they do not explain why calicheamicin binds preferentially to pyrimidine-

(30) Chatterjee, M.; Mah, S. C.; Tullius, T. D.; Townsend, C. A. *J. Am. Chem. Soc.* **1995**, *117*, 8074−8082.

(31) Calicheamicin does not cleave mixed DNA sequences, as observed in numerous gel-electrophoresis experiments. We observed no binding of the native oligosaccharide to mixed sequences using capillary electrophoresis (see ref 17).

(32) This complex was previously analyzed by Paloma et al. (ref 14) Because the exchange process made it impossible to assign the spectra at room temperature, these authors made their assignments at 45 °C, a temperature at which conformational averaging simplified the spectra.

rich sequences over other sequences. Our aim has been to rationalize the pyrimidine-selective binding of calicheamicin. Therefore, in analyzing the four complexes, we looked both for similarities in functional groups that would distinguish pyrimidine-rich sites from all other sites, and for any other characteristics that might explain the unusual selectivity. The only specific functionality that is conserved across pyrimidine tracts are the pyrimidine carbonyls that line the floor of the minor groove. Neither we nor others have been able to detect hydrogen bonding or other contacts to these groups.

One feature that we did observe in all four specific complexes, however, was a precise conformation in the DNA caused by calicheamicin binding. In the absence of calicheamicin, all four duplexes have a B form conformation, and all of the ribose sugars except one have a C2′-*endo* conformation. The one ribose sugar that has a different conformation is the ribose at nucleotide 4 in the CTCT duplex. This ribose has a pseudorotation angle below 90°. When calicheamicin binds to DNA, this ribose adopts a C2′-*endo* conformation as do most of the other ribose sugars in all four duplexes. However, there are three nucleotides in all four complexes that show a detectable change away from a C2′-*endo* conformation upon binding calicheamicin. These nucleotides include N5, N6, and N14/15. The largest conformational change occurs in the ribose sugar at nucleotide 6. Therefore, although the DNA duplexes have different free conformations, they all converge to a similar conformation upon binding calicheamicin. We have concluded that calicheamicin induces this conformation. We propose that calicheamicin can only bind at sites where it can be accommodated by a set of conformational changes in the DNA that create a complementary binding surface. Evidently, pyrimidine-rich sequences share this ability to adapt to fit calicheamicin.

Conformational changes similar to those we observed were also reported by Patel and co-workers, who have generated a high-resolution structure of calicheamicin bound to a hairpin duplex containing a TCCT site.[15,16] Hence, at least two different TCCT sequences show the same conformational changes upon DNA binding. Paloma et al. did not report any conformational changes in their TCCT-containing decamer duplex upon binding calicheamicin, but as we have discovered, the line broadening in the Paloma compelx is so severe that it is not possible to evaluate the conformations of the sugars using 2D data. Hence, we can conclude that for all complexes that are sufficiently well-behaved to obtain high quality spectra, the same characteristic conformational changes are observed.

Our results suggest that calicheamicin senses DNA conformation and that the pyrimidine selectivity of calicheamicin results from a shape-selective interaction in which the pyrimidine-rich sequences adapt to the shape of the calicheamicin oligosaccharide.[33] We do not know exactly what makes pyrimidine sequences able to adapt better than other sites. The Townsend and Dedon groups have proposed that calicheamicin recognizes a hinge-like flexibility that is peculiar to 3′-junctions of oligopurine runs.[12,37,38] This explanation does not shed light on why calicheamicin binds in the same mode and with almost identical affinity to TCCT and ACCT sites. Thus, there may also be differences within runs of pyrimidines/purines that provide them with an ability to adapt to the shape of calicheamicin.

Conclusions

Previous studies have shown that calicheamicin is a relatively rigid molecule, with a specific three-dimensional shape.[21,27,39] Here, using a set of DNA sequences with a range of different static conformations and functionalities, we have shown that calicheamicin binds in the same orientation and induces similar conformational changes in the various DNA sequences. The conformation of calicheamicin itself does not change detectably. On the basis of these observations, we propose that calicheamicin's ability to bind to a range of different oliopyrimidine sequences is based on shape-dependent DNA recognition by the rigid drug. While specific contacts can help calicheamicin discriminate between the various oligopyrimidine sequences, they cannot explain the preference for these sequences over mixed sequences of DNA. If this hypothesis is correct, then any structural scaffold that preserves the overall shape of calicheamicin should be able to recognize oligopyrimidine sequences by exploiting their ability to distort appropriately. Conversely, changes that alter the nature of individual contacts but do not significantly change the overall shape of the molecule would not be expected to have a significant effect on binding selectivity. These predictions are tested in the following paper, and the results support the induced-fit hypothesis for calicheamicin binding selectivity.[17]

Experimental Section

Preparation of Calicheamicin−DNA Complexes. Purified calicheamicin γ_1 was a gift from Dr. George Ellestad (Wyeth-Ayerst Research). The 1:1 complexes of calicheamicin and the seven different DNA duplexes studied (Figure 1b) were prepared as described previously.[13] The purified DNA oligonucleotides were synthesized on a 10 μmol scale by the Princeton Synthesis Facility. Following dialysis to remove TFA salts, the strands were lyophilized and dissolved in 0.22 mL of NMR buffer (10 mM sodium phosphate, pH 7.0/70 mM NaCl/0.05 mM EDTA). Absorbances were measured from the calculated extinction coefficients.[40] Equimolar amounts of the two corresponding strands were mixed, and the volume of each sample was brought to 0.5 mL with NMR buffer. The concentration of each DNA duplex in 0.5 mL was 3.3 mM. After annealing and repeated lyophilization from D_2O, the samples were dissolved in 1.0 mL of D_2O and 1.0 equiv of calicheamicin was added to each in ~0.5 mL of CD_3-OD. The individual mixtures were transferred to amber NMR tubes, where the volume was reduced to 0.5 mL by evaporation under argon.

Acquisition of NMR Data. Two-dimensional proton NMR experiments were carried out on a JEOL GSX/GX 500-MHz spectrometer and processed as described previously.[13] Phase-sensitive NOESY (90 ms and 200 ms at 15 °C) and DQF-COSY (at 21 °C) experiments in D_2O using a presaturation pulse on the HOD signal were used to assign all non-exchangeable protons in the free DNA and the complex. DNA labile protons were assigned by a NOESY (300 ms at 15 °C) in 90% H_2O/10% D_2O using a jump return sequence for water suppression. The sweep width was 5000−11000 Hz, and the pulse delay was 2 s. Spectra were acquired using 2048 or 4096 complex points in the t_2 dimension and 256−350 data points in the t_1 dimension with 64−128 scans per data point. Following acquisition, the data transferred to a

(33) Several experimental results from other laboratories, including the Sugiura,[34] Ellestad,[35] and Danishefsky and Crothers laboratories,[36] also suggest that calicheamicin recognizes common conformational features first, and only secondarily discriminates between pyrimidine sequences based on specific contacts.

(34) Uesugi, M.; Sugiura, Y. *Biochemistry* **1993**, *32*, 4622−4627.

(35) Krishnamurthy, G.; Ding, W.-D.; O'Brien, L.; Ellestad, G. A. *Tetrahedron* **1994**, *50*, 1341−1349.

(36) Sissi, C.; Aiyar, J.; Boyer, S.; Depew, K.; Danishefsky, S.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10643−10648.

(37) Kuduvalli, P. N.; Townsend, C. A.; Tullius, T. D. *Biochemistry* **1995**, *34*, 3899−3906.

(38) Yu, L.; Salzberg, A. A.; Dedon, P. C. *Bioorg. Med. Chem.* **1995**, *3*, 729−741.

(39) Walker, S.; Yang, D.; Kahne, D. *J. Am. Chem. Soc.* **1991**, *113*, 4716−4717.

(40) Borer, P. N. *CRC Handbook of Biochemistry and Molecular Biology; Nucleic Acids*; Fasman, G. D., Ed.; CRC Press: Cleveland, 1975; Vol. I, p 589.

Silicon Graphics workstation and processed using the FELIX program (Biosym Technologies, San Diego, CA). The resonances in all of the calicheamicin−DNA complexes were assigned using standard procedures.[22,41]

Restrained Molecular Dynamics. Molecular dynamic simulations of the calicheamicin−DNA complexes were performed to produce sets of three-dimensional structures that were consistent with distance and dihedral restraints determined by the experimental NMR data (Discover 3.0 program, Biosym Technologies, Inc., San Diego, CA). Relative NOESY cross-peak intensities obtained at 90 and 200 ms mixing times (at 15 °C) were used to classify cross-peaks into four distance categories using the fixed distance cytosine H5−H6 cross-peaks as a reference. Four distance ranges were used, 1.8−2.5 Å for strong intramolecular NOEs, 1.8−3.5 Å for strong intermolecular NOEs, 2.51−3.5 Å for medium NOEs, and 3.51−5.0 Å for weak NOEs. Restraints were also included for the hydrogen bond distances between bases, in agreement with the NMR data showing that all but the terminal imino protons can be identified in both complexes.[42] Loose distance restraints (2.0−4.0 Å) from R3 and R6 to the putative hydrogen atom abstraction sites (N5H5′ and N15H4′) were also included. Relative DQF-COSY cross-peak intensities of the scalar coupled sugar protons were used to obtain a range of pseudorotational phase angles for each deoxyribose using the relationship between 3-bond vicinal J coupling constants and pseudorotational phase angle determined by the modified Karplus equation.[43] The range of psuedorotational phase angles were used to calculate boundaries for the dihedral angles of the deoxyribose rings.[23] Using these distance and dihedral restraints, molecular dynamics simulations were performed using a starting structure of B form DNA in a vacuum with calicheamicin positioned 10 Å outside the minor groove and with fourteen fixed positively charged sodium ions to minimize electrostatic repulsion. Partial charges for DNA are known and partial charges for calicheamicin were previously calculated by computing Mulliken charges using the MNDO method of MOPAC (Biosym Technologies, Inc., San Diego, CA). The AMBER force field with Homan's parameters for carbohydrates[28] was modified to include parameters for the unusual glycosidic linkages in calicheamicin (the N−O bond) determined by experimental data and ab initio calculations, as described previously.[26,27] The starting structure of the complex was energy minimized using a conjugate gradient algorithm until the maximum derivative was less than 0.1 kcal/Å. A distance-dependent dielectric constant of 1r and a nonbonded cutoff distance of 30 Å were used. The force constant for the distance restraints was 10 kcal mol^{-1} Å^2 and for the torsion angle restraints was 40 kcal mol^{-1} rad^{-2}. The simulations used the AMBER force field, so 1−4 parameters were scaled by 0.5 as recommended. The structures were minimized to a gradient of 0.1, and then equilibrated for 0.3 ps at 200 K. Following the equilibration, the restraints were applied, initially scaled by a factor of 0.001. The scale factor was increased to 1.0 over 10 ps of dynamics. The simulation continued with full restraints for 0.3 ps at 200 K, and then the temperature was reduced to 0 K over 2.5 ps. Ten structures were generated for each complex, and the agreement between the structures was found to be good, particularly in the central portion of the complexes where there are several intermolecular NOEs.

Cleavage Studies. Oligonucleotides d(GT**GTCTC**GGC) and d(GCT-GATG**ATCTC**AGACTGCC) were 5′-end labeled in a 10 μL reaction mixture containing 1−5 pmol of oligonucleotide, 5−7 times excess of [γ-^{32}P] ATP (3000 Ci/mmol), kinase buffer (10 mM $MgCl_2$; 70 mM Tris·HCl, pH 7.6; 5 mM DTT), and 5−10 units of T4 polynucleotide kinase, which were incubated at 37 °C for 1 h. The enzyme was denatured by heating the reaction mixture at 90 °C for 5 min. The reaction mixture was mixed with 100 μL of 1× TE buffer (10 mM Tris·HCl, 1 mM EDTA, pH 7.2), and excess ATP was removed on a G25 sephadex spin column. The 5′-labeled oligomer (1 pmol) and 20× excess of the "cold" complementary strand were hybridized in a volume of 100 μL containing 10 mM Tris·HCl, pH 7.6, which was incubated at 90 °C for 5 min and slow-cooled to 25 °C over 90 min. A typical cleavage experiment contained less than 1 fmol of the labeled duplex, 10% calicheamicin γ_1 in DMSO, 0.4 mg/mL salmon testes DNA in 20 mM NaCl and 30 mM Tris-HCl buffer (pH 7.5). Calicheamicin was allowed to incubate for 1 h at 37 °C, and cleavage was initiated by adding β-mercaptoethanol to a final concentration of 100 mM. The cleavage products were treated with piperidine or putrescine at 95 °C for 30 min and separated by 19% PAGE under denaturing conditions. The autoradiograms were scanned and quantified using a densitometer.

Acknowledgment. This work was supported by Grants GM 53066 and GM 42733 from the National Institutes of Health. K.B acknowledges a graduate fellowship from Bristol-Myers Squibb. We thank Drs. George Ellestad and Wei-Dong Ding of Lederle Laboratories for generously providing us with calicheamicin.

JA0005183

(41) Hare, D. H.; Wemmer, D. E.; Chou, S.-H.; Drobny, G.; Reid, B. R. *J. Mol. Biol.* **1983**, *171*, 319−336.

(42) Distances for the hydrogen bonding restraints were based on crystal structure analysis, see: Saenger, W. *Principles of Nucleic Acis Structure*; Springer-Verlag: New York, 1984; p 123. In addition, loose backbone torsion angle restraints were included because the NMR data indicate that the DNA maintains a right-handed conformation, see: Schmitz, U.; Sethson, I.; Egan W. M.; James, T. L. *J. Mol. Biol.* **1992**, *227*, 510−531.

(43) van de Van, F. J. M.; Hilbers, C. W. *Eur. J. Biochem.* **1988**, *178*, 1−38.