Pergamon

Tetrahedron Vol. 50, No. 5, pp. 1379-1396, 1994

0040-4020/94 $6.00+0.00

0040-4020(93)E0207-V

Probing the Underlying Basis for the Binding Specificity of Calicheamicin γ_1^I. A Molecular Dynamics Study

David R. Langley*†, Terrence W. Doyle† and David L. Beveridge‡

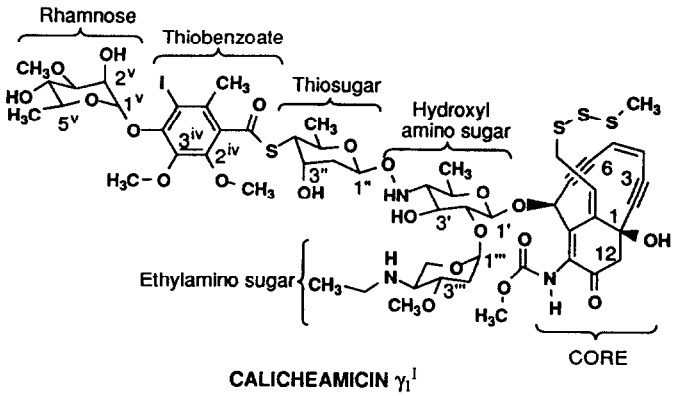
†Bristol-Myers Squibb Company, Pharmaceutical Research Institute, 5 Research Parkway, Wallingford, CT 06492-7660

‡Chemistry Department, Hall-Atwater Laboratories, Wesleyan University, Middletown, CT 06457

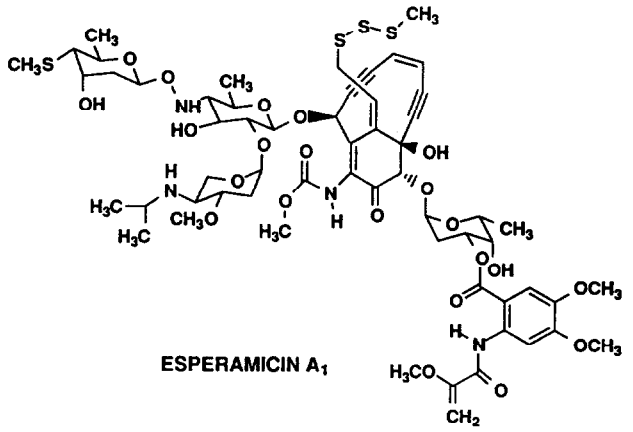
Abstract: Six different models of the DNA-calicheamicin complex were constructed to investigate the mode of binding of the rhamnose, thiobenzoate and thiosugar residues of the carbohydrate-thiobenzoate tail of calicheamicin with DNA. The stability of each model was tested in a 500 psec solvated molecular dynamics simulation. Five of the six models predicted the correct DNA hydrogen abstraction pattern. However, only one model was found to be consistent with all of the available experimental data and provided a clear rationale for the observed 5'-TCCT/3'-AGGA sequence specificity.

INTRODUCTION

Calicheamicin γ_1^I[1,2] and esperamicin A_1 [3,4] (Figure 1) are closely related enediyne containing natural products. Other members of this class include neocarzinostatin[5], dynemicin[6,7] kedarcidin[8,9] and C-1027[10,11]. This class of molecules represents some of the most potent antitumor antibiotics known to man. Their antitumor activity has been attributed to their ability to produce single and/or double strand breaks in DNA[5,12,13]. Calicheamicin is also one of the most sequence specific non-protein DNA cleaving molecules reported to date. It has a cleavage preference for TCCT sites, however, other sites such as (TCCC, TCCA, ACCT, TCCG, GCCT, CTCT, TCTC)[12] and (TTTT, TTCA and TTGT)[14] are also attacked. Upon reductive activation calicheamicin produces two carbon centered radicals at positions C3 and C6 (Figure 2). The diradical intermediate (CAL R) concomitantly cleaves each strand of the DNA by abstracting the H5" (pro *S*) from the 5'-C of the TCCT sequence and the H4' from the base on the opposite strand that is staggered in the 3'-direction by three bases from the 5'-C[15,16] (Figure 2). It should be noted that the DNA breaks always appear at the 5'-end of the sequence specific site and that the H5" DNA hydrogen is abstracted by the C6 radical of CAL R while H4' quenches the C3 radical. The regio and stereo selective transfer of DNA hydrogens to CAL R clearly demonstrates a mono-directional mode of binding. The removal of the rhamnose moiety from calicheamicin decreases its cutting efficiency but does not alter its sequence specificity[17].



CALICHEAMICIN γ_1^I



ESPERAMICIN A_1

Figure 1.

Esperamicin A_1 is comparatively much less sequence specific and causes single strand DNA breaks and has been predicted to bind simultaneously in the major and minor grooves of the DNA[18]. Esperamicin C (Figure 3), a hydrolysis product of esperamicin A_1[13], can be thought of as an analog of calicheamicin[14,19] which has a hydroxyl at C12 and thiomethyl and isopropylamino groups in place of the thiobenoate-rhamnose moieties and ethylamino group, respectively. Esperamicin C, like calicheamicin, binds in the minor groove of DNA[13,20] and causes double strand DNA breaks by simultaneously abstracting the H5" from one strand and a H4' from the opposite strand[18,21], however it is considerably less sequence specific and its cleavage pattern indicates a bidirectional mode of binding[18,21]. Collectively, these results strongly suggest that the thiobenzoate plays a major role in determining the sequence specificity and binding directionality exhibited by calicheamicin but has little or no effect on orienting the CORE within the DNA for hydrogen abstraction. There is a wealth of data characterizing the mechanism of activation of the aglycone and the process of DNA

hydrogen atom abstraction which can be used to infer the relative orientation of the CORE within the DNA cleavage sites. However, very little is known about the 3D structure of the DNA-calicheamicin complex, specifically, the molecular basis for its high sequence specificity.

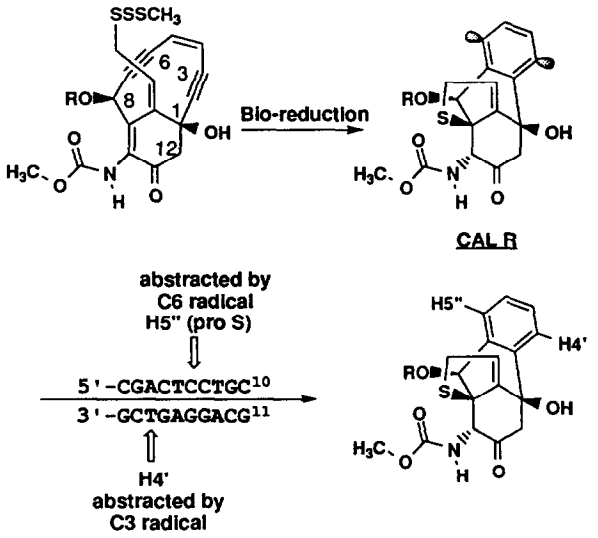


Figure 2. Mechanism of activation for calicheamicin and predicted cleavage sites within the modeled DNA sequence.

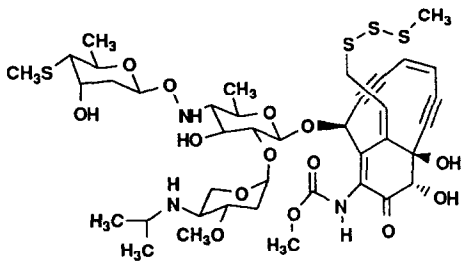


Figure 3.

Two modeling studies have been reported that describe the DNA-calicheamicin complex. However, both have been found (at least in part) to be inconsistent with the experimental evidence available prior and subsequent to their publication. The Schreiber model[22] has several areas of contention, however, it should be noted that the model correctly predicted the *S*-stereochemistry at C8 of the aglycone, providing theoretical

support to the supposition that the aglycone of calicheamicin γ_1^I had the same stereochemistry as esperamicin A_1[3,4,23]. In addition, the model was the first to suggest that the DNA cleavage that occurs via the C5'-hydrogen abstraction[12] takes place by the transfer of the H5" (pro-*S*) hydrogen from the 5'-C of the TCCT binding site to the radical generated at C6 of calicheamicin. More recently, the abstraction of the H5" (pro-*S*) hydrogen by the C6-radical has been proven by site specific deuterium labeling studies[15,16]. The most apparent problem with this model pertains to the conformation of the calicheamicin hydroxylamino sugar. The model places this sugar in a boat conformation, which is extremely unusual for a pentasubstituted hexanose whose chair conformation allows all five substituents to be in equatorial positions. The crystal structure of dihydrocalicheamicin pseudoaglycone[2] had previously shown this sugar to exist in the expected chair conformation. More recently, NMR studies[24] have confirmed the chair conformation as the preferred solution state conformation in three different solvents and at different temperatures. Although the model correctly predicts that calicheamicin abstracts the H5" from 5'-C of the TCCT containing strand it wrongly predicts the abstraction of the H1' from the cut site on the opposite strand. Finally, this model places the thiobenzoate in the minor groove so that its iodine interacts with the 2-amino group of each of the guanines in the 5'-TCCT/AGGA binding site, i.e., the N^2-I distances are within 3-4Å. The iodine nitrogen interaction was supported by the crystal structures of p-iodobenzonitrile and the 1:1 iodoform/quinoline complex. Both X-ray structures illustrate the attractive interaction between iodine and the electronegative sp^1 or sp^2 nitrogen and suggests that an aromatic iodine is electropositive. However, it is hard to imagine how the iodine of the thiobenzoate can interact with the nitrogen of the guanine 2-amino group without first interacting with at least one of it's polar hydrogens, as they extend farther off the floor of the minor groove. In the second model, reported by the Lederle group[17], calicheamicin is docked into the minor groove, but resides almost entirely to the 5'-side and out of the preferred binding sequence. While both models place the iodine of the thiobenzoate near the floor of the minor groove the two models differ in several ways. The main difference is in the conformation of the calicheamicin sugars and their geometry relative to one another. The Lederle model has all of the sugars in a chair conformation and places the 3"-hydroxyl and 2^V-hydroxyl of the thiosugar and rhamnose, respectively, near the floor of the minor groove where they hydrogen bond with the bases. In contrast, the Schreiber model has the hydroxylamino sugar in a boat conformation and has the hydroxyl groups pointing out of the minor groove into the environment surrounding the DNA. No discussion is provided in either paper that describes the possible contribution of the ethylamino sugar. However, an examination of the figures suggests that it interacts with the local environment surrounding the DNA but not directly with the DNA. Both of these models may contain valuable insights into the DNA-calicheamicin complex but because of their inconsistencies with the experimental data and each other it is difficult to know what to take from them. This prompted our study of the DNA-calicheamicin complex.

THE MODELS

Eight different models (Figure 4) were initially constructed to investigate the interactions between the thiosugar, thiobenzoate and rhamnose moieties and the minor groove of DNA. In the first four models the rhamnose moiety is docked into the DNA with its methyl group ($C6^V$) pointing into the groove and near O2 of T:8 and H2 of A:13. The first four models are; (1) the IH_in model (Figure 4) which places both the

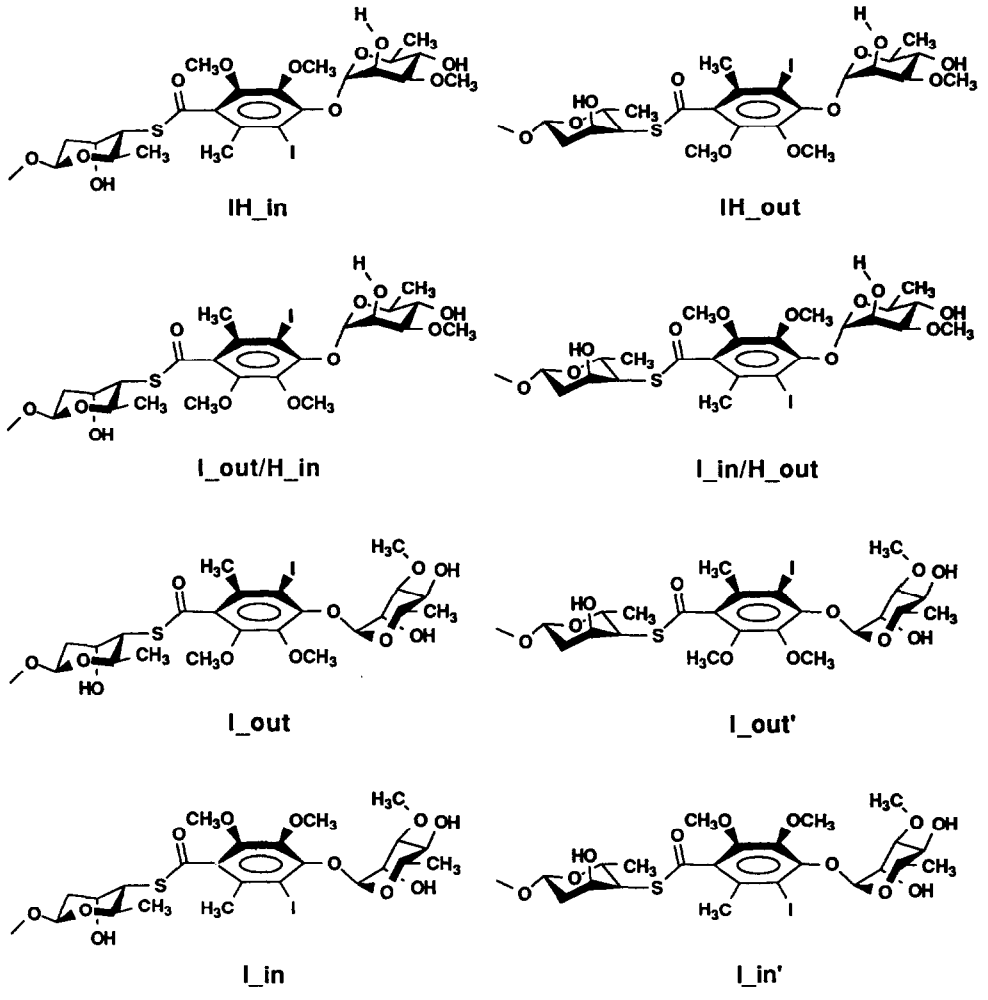


Figure 4. Eight possible models resulting from the binding perturbations of the thiosugar, thiobenzoate and rhamnose.

thiobenzoate iodine and 3"-hydroxyl of the thiosugar near the floor of the minor groove, (2) the IH_out model points both the iodine and 3"-hydroxyl out of the minor groove, (3) the I_out/H_in model and (4) the I_in/H_out model alternately points one group in and the other out of the minor groove. In the last four models the rhamnose moiety is docked into the DNA with its methoxy group at $C3^V$ pointing into the minor groove and its hydroxyl groups at $C2^V$ and $C4^V$ hydrogen bond with O4' of G:15 and O3' of G:9, respectively. In models (5) the I_in model and (6) I_out model (Figure 4 and 5) the iodine is alternately pointed in or out of

the minor groove as their names imply and the thiosugar is docked with its 3"-hydroxyl near the floor of the minor groove where it hydrogen bonds with N3 of A:17. In the last two models, I_in' and I_out' (Figure 4),

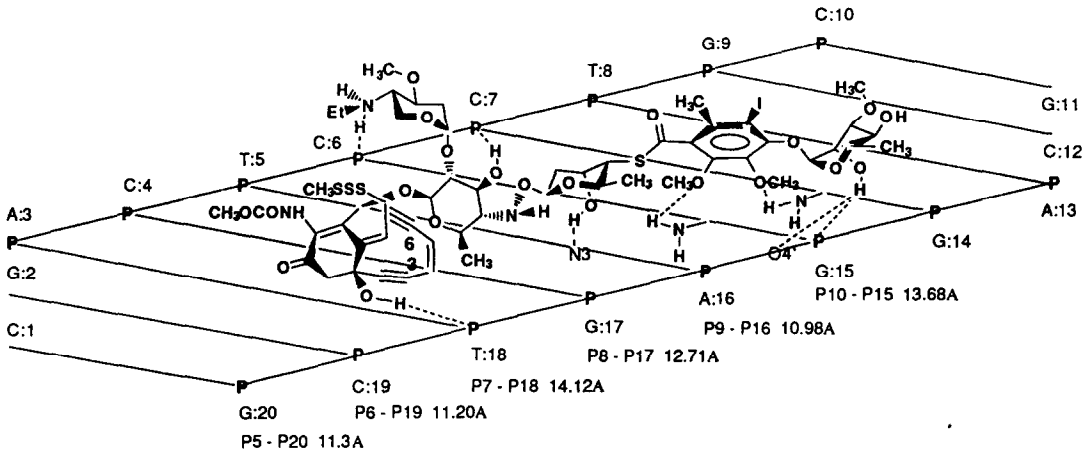


Figure 5. Graphical representation of the DNA-calicheamicin (I_out) model. The closest inter strand phosphate-phosphate distance in canonical B-form DNA is 11.46Å. Dotted lines represent hydrogen bonds.

the iodine is docked in or out of the minor groove and the 3"-hydroxyl of the thiosugar is pointing out into the environment surrounding the DNA. The last two models (I_in' and I_out') were aborted due to the dynamical instability of the IH_out and I_in/H_out models and their lack of consistency with either the X-ray crystal structure of dihydrocalicheamicin pseudoaglycon[2] or our DNA-esperamicin model[18].

The initial docking of the CORE, hydroxylamino sugar and ethylamino sugar residues was based on our DNA-esperamicin C model[18] The CORE sets in the minor groove positioning C3 and C6 (radical bearing carbons) near and directionally aligned for the abstraction of the DNA H4' of T:18 and H5" of C:6 of the d(CGACTCCTGC)·d(GCAGGAGTCG) oligonucleotide, respectively (Figure 2). Additionally, the C1-hydroxyl hydrogen bonds with O3' of T:18. The hydroxylamino sugar lies flat in the mouth of the minor groove approximately one to two water layers above the floor with its 1',4'-axis parallel with the groove floor. Its methyl group (C6') sits near the deoxyribose of T:18 and its C3'-hydroxyl hydrogen bonds with O2P of C:7. The ethylamino sugar hovers at the lip of the minor groove over the deoxyribose of C:6 and forms a hydrogen bond/salt bridge[25] with O5' and O2P of the C:6 phosphate group.

The major differences between the models arise from the interaction of the thiosugar thiobenzoate and rhamnose with the DNA binding pocket. In the IH_in, I_out/H_in, I_in and I_out models the thiosugar C3"-hydroxyl group forms a hydrogen bond with N3 of A:16 and the thiosugar C2"-methylene is in close Van der Waals (VDW) contact with the C2-methine of A:16. When C6" of the thiosugar is turned in (IH_out, I_in/H_out models) it sets near N3 of A:16, The additional steric bulk of the equatorial C6"-methyl compared to the axial C3"-hydroxyl pushes the thiosugar away from the floor of the minor groove. This increases the arch of the drug and reduces both the VDW and electrostatic interactions between the drug and the DNA and shortens the reach of the thiobenzoate and rhamnose down the minor groove by 0.3-0.5Å. In the three models

(IH_out, I_out/H_in and I_out) where the dimethoxy edge of the thiobenzoate is turn toward the minor groove floor, the thiobenzoate methoxy group at $C2^{iv}$ and $C3^{iv}$ accepts a hydrogen bond from the 2-amino group of G:15 and G:14, respectively. When the iodine is turned into the minor groove (IH_in, I_in/H_out and I_in models) the interactions between the thiobenzoate and G:14 and G:15 are quite different. The iodine sets nearly equal distance between the planes defined by the C:6-G:15 and C:7-G:14 base pairs and interacts almost exclusively with the guanine containing strand. However, the contact points between the thiobenzoate and the C-G base pairs is different due to the base pair twist angle. The closest DNA interactions with the iodine are (ascending order) NH22/G:14, H1'/G:15, N2/G:14, N3/G:15, NH22/G:15, NH21/G:14, O5'/G:15 and O5'/A:16. It should be noted that direct interaction between the iodine and N2 of G:14 or G:15 is sterically blocked by NH22 of that same group. The NH22 hydrogen by convention is the 2-amino group hydrogen not involved in Watson-Crick hydrogen bonding and points away from the floor of the minor groove. The $C6^{iv}$-methyl group of the thiobenzoate also makes contacts with the floor of the minor groove at (ascending order) NH22/G:15, H4'/T:8, H1'/A:16, H1'/T:8 and O2/C:7. All other contacts are greater then 4.0Å. The placement of rhamnose $C6^{v}$ near the floor of the minor groove allows this hexanose to sit in the center of the groove. However, when the rhamnose is docked with its $C3^{v}$-methoxy near the T:8-A:13 base-pair (I_out and I_in models) it is forced to sit closer to the lower 5'-AGGA containing strand with its 1^{v}-4^{v}-axis pointing across the minor groove.

Each complex was submitted to 515 psec of molecular dynamics at 300°K. In the first 115 psec a soft distance constraint[26] of 2.7(+0.1 or -2.7)Å with a 1kcal/mol·Å force constant was used between the CAL R radicals at C3 and C6 and the DNA H4'/T:18 and H5"/C:6, respectively, to equilibrate the complexes with experimental points of reference. It must be emphasized that only the distance between the radical center and the hydrogen that has been experimentally shown to be abstracted by that radical was constrained. The angle dependence between the radicals and the abstracted hydrogens as well as all other aspects of the equilibrated conformation resulted solely from the MD simulation and the CHARMM force field[27]. The final 400 psec of the MD runs were conducted without constraints.

The Curves, Dials and Windows (CDW) analysis package introduced by Ravishanker et al.[28] was used to analyze the MD results. The structural inferences described below are based on this analysis. A complete CDW graphical description of all the DNA parameters is available from the supplementary material.

RESULTS

During the first 115 psec the I_out/H_in model underwent the least amount of change and settled down in the RMS space near the starting point (Table 1). Like the I_out model the hydrogen bonding and VDW interactions between the drug and the DNA remained, for the most part, unchanged from the energy minimized starting point. It should be noted that this is a dynamical model and the distance between VDW contacts points fluctuate and some of the observed hydrogen bonds flicker on and off, particularly the hydrogen bond between the thiobenzoate $C2^{iv}$-methoxy and NH22 of G:15. In both the I_out/H_in and I_out models the most significant changes in the DNA occurred at the T:5-A:16 and C:6-G:15 base pair steps. The axis-junction X-displacement (AXD, Figure 5) parameter changed from 0.0Å to 0.8Å revealing a displacement of the C:6-G:15 base pair into the major groove. The shift (SHF, Figure 6) inter-base pair

parameter for the C:6-G:15 base pair also increased from 0.0Å to 1.6Å as the T:5-A:16 base pair move into the minor groove. The propeller twist of the T:5-A:16 base pair also changed and stabilized around -34.2 degrees as a result of the hydrogen bond between N3 of A:16 and the 3"-hydroxyl group of the thiosugar. The remaining DNA parameters fluctuated around values near the starting energy minimized conformations and are characteristic of B-form DNA. The changes in the DNA at the T:5-A:16 and C:6-G:15 base pairs occurred as the drug settled into the binding pocket. To accommodate the methoxy groups near the floor of the groove

Table 1. The total energy (kcal/mol) for the energy minimized solvated complex. RMSD (Å) between the starting point and structural components (drug-DNA complex, DNA only and drug only) averaged over 65.5-115 psec. The numbers in parenthesis were averaged over the last 100 psec of simulation. The average ensemble energy (kcal/mol) for the last 400 psec of the molecular dynamics simulation.

	I_out/H_in	I_in/H_out	IH_out	IH_in	I_out	I_in
Total Energy	-12553.795	-12522.611	-12533.940	-12530.278	-12576.437	-12536.100
Complex	1.432 (3.236)	1.524 (1.806)	1.893 (3.353)	1.942 (3.256)	1.765 (1.674)	1.901 (2.477)
DNA	1.366 (3.151)	1.374 (1.729)	1.762 (3.108)	1.780 (3.151)	1.795 (1.650)	1.878 (2.350)
Drug	1.356 (2.255)	1.705 (1.574)	1.789 (1.602)	1.897 (2.255)	1.215 (1.146)	1.465 (2.221)
Ensemble Energy	-8182.983	-8153.864	-8164.441	-8138.328	-8185.724	-8172.178

the C:6-G:15 base pair moved slightly into the major groove and at the same time to increase the VDW and electrostatic interactions between the T:5-A:16 base pair and the thiosugar, the T:5-A:16 base pair move slightly into the minor groove. The greatest difference in the DNA between the I_out/H_in and I_out is localized around the rhamnose moiety. In the I_out/H_in model the rhamnose sets in the center of the minor groove allowing for a narrow minor groove in this region of the DNA, however, in the I_out model the rhamnose resides closer to the lower strand and points across the minor groove which caused the minor groove to open up during the first 115 psec of the Molecular dynamics simulation.

The I_in/H_out model also settled down in the RMS space near its starting point (Table 1). However, the drug underwent considerable conformational changes. The thiobenzoate moved slightly (0.5-1Å) away from the floor of the minor groove and turned to move the iodine closer to the lower strand, away from NH22 of G:14 and approximately equal distance from N3 and O4' of G:15 and O4' of A:16. Over the same time period the thiosugar turned to lie flat in the mouth of the minor groove, some two to three water layers above the floor of the groove. From this new binding position the thiosugar C3"-hydroxyl group hydrogen bonds with O3' of C:7. The DNA on the other hand fluctuated near its energy minimized conformation, maintaining a strong B-form DNA conformation. The rhamnose remained sandwiched between the DNA backbones in the minor groove, but like the thiobenzoate and thiosugar moved away from the floor of the groove.

The IH_in model showed the largest deviation from the starting point in both the conformation of the DNA and drug (Table 1). The thiobenzoate in this model also turned its iodine away from NH22 of G:14 and toward the lower stand. However, the thiobenzoate did not move off the floor of the minor groove but instead slid down the minor groove by 1-2Å (in the 5'-direction relative to the 5'-TCCT sequence). The repositioning of the thiobenzoate in the minor groove not only caused the entire drug to move down the groove but forced

the groove to open up by bulging the lower strand around the thiobenzoate group. The thiosugar maintained its position sandwiched in the minor groove, with its C3"-hydroxyl hydrogen bonding intermittently with N3 of A:16 at first and then with O4' of G:17, and is presumably responsible for holding the thiobenzoate near the floor of the minor groove. However, as the MD progressed and the minor groove opened up around the thiobenzoate, the thiosugar finally lost its grip on the lower strand but maintained close VDW contact with the deoxyribose of C:7 in the top strand.

In the I_in model the non-bonded interactions between the CORE and hydroxylamino, thio and ethylamino sugars and the DNA binding site are for the most part unchanged from the minimized structure. However, during the equilibration phase (first 115 psec) of the molecular dynamics run the thiobenzoate turned and tilted so that its $C6^{iv}$-methyl sets deep in the minor groove close to the deoxyribose of C:7, near C1' of C:7, and the $C5^{iv}$-I bond is parallel to the floor of the minor groove. The iodine moved away from the C:7-G:15 base pair and into the plane of the C:8-G:14 base pair and is the closest calicheamicin atom to the T:9-A:13 base pair. Accompanying the relocation of the thiobenzoate is the movement of the rhamnose sugar away from the floor of the minor groove. The rhamnose settled down near the lower strand of the DNA where its $C2^{v}$-hydroxyl hydrogen bonds with O2P of A:16.

The IH_out model also settled down after developing considerable distortion in both the drug and the DNA compared to its starting point (Table 1). The thiosugar turned to lie flat in the minor groove where its C3"-hydroxyl hydrogen bonds with the thiobenzoate carbonyl and O2P of T:8. Additionally, the thiosugar flipped from a chair to a boat conformation. The thiobenzoate maintained its position in the minor groove where it continued to hydrogen bond via its methoxy groups with NH22 of G:14 and G:15. The change in the conformation and binding position of the thiosugar caused the DNA minor groove to swell open around it. This large bulge in the DNA minor groove accounts for the majority of the DNA distortion.

At the end of the equilibration period (0-115 psec) the two distance constraints were removed and the MD simulation continued for an additional 400 psec. The distance and angle[18,29] between the carbon centered radicals and the three closest DNA hydrogens were monitored as a function of time to determine which if any of the models would remain consistent with the experimental hydrogen abstraction data. The average distance and angle between the radical centers and the three closest DNA hydrogens with respect to each radical is presented in Tables 2 and 3, respectively. The removal of the distance constraints was accompanied by a drift in RMS space in all models (Table 1). Interestingly, only the I_out model drifted in the direction of its energy minimized starting point. The changes in the DNA for both the I_in/H_out and I_out models were subtle, originating from minor changes throughout the DNA. The most noticeable change that occurred in the I_in/H_out was the translation of the drug away from the floor and along the minor groove in the 5'-direction relative to the 5'-TCCT strand, resulting in poorer alignment between the radicals centers and the abstracted DNA hydrogens (Table 3). The largest drift in RMS space occurred in the I_out/H_in, IH_out and IH_in models. The additional distortion was visually apparent. In the I_out/H_in model the minor groove opened around the rhamnose allowing it to turn slightly in the groove so its $O4^{v}$ and $C2^{v}$-hydroxyls can hydrogen bond with NH22 of G:14 and O2 of T:8, respectively. However, during the course of the final 400 psec of free MD the rhamnose conformation relative to the DNA minor groove fluctuated between its starting conformation where the $C6^{v}$ methyl group is pointing straight into the minor groove and the conformation described above. The bulges in the minor groove of the IH_out and IH_in models that developed over the first

115 psec continued to grow once the distance constrains were removed. The bulge in the minor groove in the IH_in model remained centered around the thiobenzoate and covers approximately six base pairs. However, in the IH_out model the minor groove opens up asymmetrically around the thiosugar. This lopsided opening around the thiosugar allows the thiobenzoate to remain sandwiched between the DNA backbones but moves the backbone of T:18 away from the radical at C3 (Table 2).

Table 2. Average distance (Å) between the calicheamicin radical centers and closest three DNA hydrogens. In some of the models the C6-H1'/C:6 distance was found to be slightly closer then C6-H4'/C:5.

	I_out/H_in	I_in/H_out	IH_out	IH_in	I_out	I_in
C6-H5"/C:6	**2.79**	**2.87**	**2.73**	**3.05**	**2.65**	**2.61**
C6-H5'/C:6	3.88	3.71	3.82	3.61	3.66	3.55
C6-H4'/C:5	4.33	4.14	4.29	4.56	4.38	4.13
C3-H4'/T:18	**2.74**	**2.80**	**7.65**	**3.10**	**2.81**	**2.68**
C3-H1'/T:18	3.45	3.47	7.79	5.61	3.30	3.43
C3-H5"/C:19	3.03	3.12	8.24	3.10	3.23	3.15

Table 3. Average angle (degree) between the calicheamicin radical centers and closest three DNA hydrogens.

	I_out/H_in	I_in/H_out	IH_out	IH_in	I_out	I_in
C3-C6-H5"/C:6	**153.9**	**155.6**	**160.9**	**142.9**	**170.5**	**171.5**
C3-C6-H5'/C:6	148.4	161.2	158.7	170.3	159.1	159.6
C3-C6-H4'/C:5	127.6	163.9	165.6	146.6	111.7	108.8
C6-C3-H4'/T:18	**135.9**	**128.9**	**110.8**	**132.2**	**143.8**	**136.1**
C6-C3-H1'/T:18	129.3	137.3	107.4	128.9	134.5	129.0
C6-C3-H5"/C:19	128.6	133.9	132.7	132.0	121.3	124.4

All six models correctly predict the H5" (pro-*S*) of C:6 as the prime candidate for abstraction by the radical at C6 (Table 2 and 3). However, the prediction of which DNA hydrogen will be abstracted by the C3 radical varies between models. Only the IH_out model clearly does not correctly predict the abstraction of the H4' of T:18. The IH_in model equally predicts the H4' and H5" of T:18, however, since the H4' is tertiary it should be energetically easier to abstract. The I_out/H_in, I_in/H_out, I_in and I_out models all cleanly predict the correct abstraction of the H4' of T:18. Therefore, all but the IH_out model must be considered as possible modes of binding solely based on the hydrogen abstraction data. While the minor groove width, within the drug binding site, for all of the DNA-calicheamicin (Figure 5 and 6) and DNA-esperamicin models[18] is larger than that found in canonical B-form DNA only the I_out/H_in, IH_in and IH_out models produce large to very large bulges in the DNA around the rhamnose, the thiosugar, thiobenzoate and rhamnose and the thiosugar and CORE moiety, respectively. In the I_in/H_out model the drug interacts predominately with the DNA backbone in the mouth of the minor groove and poorly aligned for hydrogen abstraction. This leaves the I_in and I_out models as primary candidates for the DNA-calicheamicin complex.

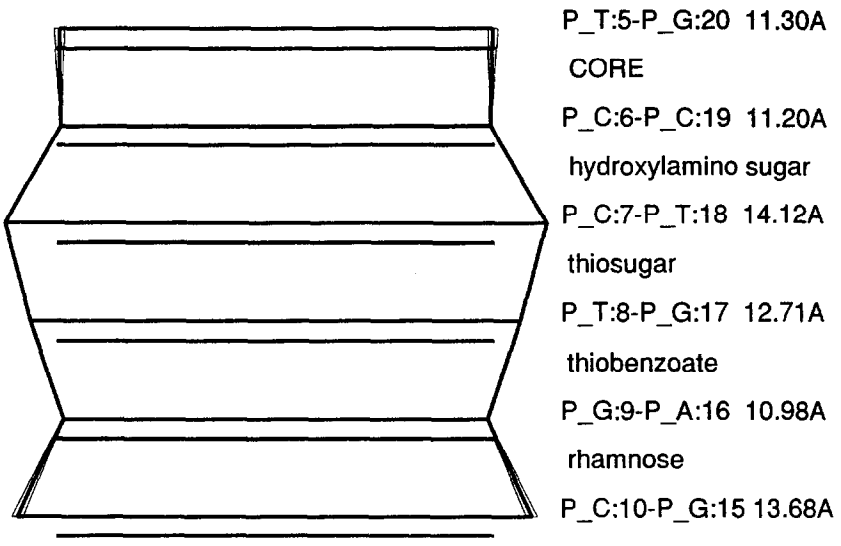


Figure 6. Graphical representation of the DNA minor groove width for the I_out model over the last 400 psec. The width is determined by monitoring the closest inter strand phosphate distances as determined from canonical B-form DNA. The lower (non-attached) horizontal bars represent the phosphate-phosphate distances (11.46Å) from canonical B-form DNA.

DISCUSSION

The molecular interactions between 5'-TCCT/3'-AGGA sequence and the calicheamicin diradical intermediate from the I_out model provide a clear and simple rationale for the observed sequence specificity (Figure 7). The thiosugar C3"-hydroxyl hydrogen bonds with N3 of the A:16 and the thiosugar C2"-methylene sits near the C2 methine of A:16 providing a steric preference against a guanine at this site. The thiobenzoate methoxy groups at $C2^{iv}$ and $C3^{iv}$ sit near the floor of the minor groove and accept a hydrogen bond from the 2-amino group of G:15 and G:14, respectively. The rhamnose binds in the minor groove in a way that places its $C3^{v}$-methoxy group approximately equal distance between the 2-carbonyl of T:8 and the C2-methine of A:13 which sterically selects for the T/A base. This binding orientation of the rhamnose sugar is further stabilized by a hydrogen bond between the $C2^{v}$-hydroxyl groups and O4' of G:15. The carbohydrate-thiobenzoate tail reads the DNA recognition elements along the floor of the 3'-AGGA strand. Since the binding site is not palindromic the recognition elements are different when viewed from different directions resulting in the observed unidirectional mode of binding. While this model provides a straight forward explanation for 5'-TCCT sequence specificity it does not clearly rationalize the high specificity for 5'-TTTT or the other high affinity sequences[14] mentioned above. The cleavage affinity studies[12,14,17] along with this modeling study suggest that different conformations of calicheamicin may recognize different sequences of DNA.

In the I_in model[30,31] (Figure 8), the non-bonded interactions between the CORE and hydroxylamino, thio and ethylamino sugars of calicheamicin and the DNA binding site are for the most part unchanged from the I_out model. However, during the molecular dynamics run the thiobenzoate turned and tilted so that its $C6^{iv}$-methyl sets deep in the minor groove close to the deoxyribose of C:7, near C1' of C:7, and the $C5^{iv}$-I bond is parallel to the floor of the minor groove. This moved the iodine away from the C:7-G:15 base pair and into the plane of the C:8-G:14 base pair. The iodine is the closest calicheamicin atom to the T:9-A:13 base pair. The rhamnose also moved away from the floor of the minor groove and settled down near the lower strand of the DNA where its $C2^{v}$-hydroxyl hydrogen bonds with O2P of A:16. The reorganization of the thiobenzoate is presumably due to steric crowding and electrostatic repulsion between the 2-amino groups of G:17 and G:18 and the $C6^{iv}$-methyl and C^{iv}-iodine of the thiobenzoate. The crystal structures of p-iodobenzonitrile and the 1:1 iodoform/quinoline complex illustrate the attractive interaction between iodine and the electronegative sp^1 or sp^2 nitrogen[22] and suggests that an aromatic iodine has a positive partial charge. The partial charge on the iodine of iodiobenzene is predicted to be +0.02, +0.12 and +0.12 by PM3, MNDO and MNDO/3 (MOPAC 6.0)[32,33], respectively. While this model proved to be stable it does not provide an explanation for the 5'-TCCT sequence specificity beyond the shape complimentary between the DNA helix and groove structure and the drug. The lack of the exocyclic C2-amino group on adenine reduces both the steric hindrance[34] and electrostatic potential[35] for A-T base pairs near the floor of the DNA minor groove. The above results suggest that the structural and electrostatic characteristics of the 5'-TTTT binding site[14] are more suitable for the binding of calicheamicin when its iodine is turned into and near the floor of the minor groove.

The I_out model suggests that when the appropriate edge of the thiosugar, thiobenzoate and rhamnose sets near the floor of the minor groove of the 5'-TCCT/3'-AGGA they act as anchors to the 3'-A, GG and A bases of the preferred binding/cleavage site, respectively. However, this part of the carbohydrate-thiobenzoate tail does not appear to effect the conformation or binding orientation of the CORE, hydroxylamino sugar and ethylamino sugar. This is apparent from the lack of perturbation to the predicted cleavage pattern exhibited by the different models (Tables 2 and 3). These results suggest the thiobenzoate is the DNA anchor. It determines 50% of the preferred binding site and may be partly responsible for holding the thiosugar and rhamnose in place for the determination of the remainder of the binding domain.

The CORE, hydroxylamino sugar and ethylamino sugar residues of calicheamicin, like esperamicin C[18,20], esperamicin D[20] and calicheamicin T[14], bind with the DNA almost exclusively via the DNA backbone. The orthogonal relationship between the thiosugar and the hydroxylamino sugar is very similar to that found in the crystal structure of the dihydrocalicheamicin psuedoaglycone[2] and our esperamicin models[18]. The NO-

Figure 7. Stereoview of the average structure from the free molecular dynamics simulation on the I_out model. The coloring scheme for the DNA is blue for carbon, light green for nitrogen, yellow for oxygen and amber for phosphate. The coloring scheme for calicheamicin is green for carbon, red for oxygen, blue for nitrogen and yellow for sulfur. The calicheamicin radical centers at C3 (lower) and C6 (upper) are violet. H4'/T:18 is red (lower strand), H5' and H5" of C:6 (upper strand) and C:19 (lower strand) are violet. All polar hydrogens are white. Non-polar hydrogens were omitted for the sake of clarity. The white dotted lines depict hydrogen bonds.

Figure 8. Stereoview of the average structure from the free molecular dynamics simulation on the I_in model. The coloring scheme is the same as in Figure 7.

glycosidic link between the hydroxylamino and thiomethyl sugars appears to serve two important roles. It functions as an extended linker which permits the thiosugar, thiobenzoate and rhamnose fragment to read the DNA minor groove (imparting sequence specificity) and at the same time allows the CORE to be positioned for the abstraction of DNA hydrogens. Additionally, the NO-glycosidic link permits the conformation of the hydroxylamino-thiomethyl disaccharide to adopt an energy minimum in which the two residues are orthogonal to one another. This minimum is stabilized by the 4'NH and C1'''O5''' anti-parallel dipole[36]. The hydroxylamino sugar lies flat in the mouth of the minor groove, one to two water layers above the floor of the groove. Its C3'-hydroxyl group hydrogen bonds with O2P of C:7, while its methyl group (C6') sets nearly equidistant from C4' of G:17 and C5' of T:18 and points at the T:18 phosphate group. The unusual binding orientation of the hydroxylamino sugar may be partially responsible for the wider minor groove (note P_C:7-P_T:18 distance, Figure 5 and 6) which in turn may be needed to allow the esperamicin and calicheamicin COREs to achieve binding positions in the minor groove that are suitable for simultaneous hydrogen abstraction from each strand of the DNA.

The CORE and hydroxylamino sugar (in both the calicheamicin and esperamicin models) float one to two water layers above the floor of the minor groove and only directly interact with the DNA backbone. The spine of hydration that runs between the floor of the minor groove and the CORE and hydroxylamino sugar interacts with the drug and DNA in a specific manner that stabilizes the complex and may assist in the positioning of the CORE for highly efficient DNA hydrogen abstraction. The nature of the spine of hydration that runs along the minor groove of DNA has been shown to be sequence dependent[37,38]. The modeling studies along with the cleavage affinity experiments suggest that the sequence specificity observed for the esperamicins and calicheamicins is partly due to the drugs' ability to read the DNA groove structure or inducible structure[14] as well as the spine of hydration that runs along the minor groove floor[18]. DNA footprinting studies carried out on the aryl tetrasaccharide portion of calicheamicin[39,40] supports this hypothesis. The removal of the bulky CORE from calicheamicin should allow the hydroxylamino sugar to slip deeper into the minor groove of DNA sequences with "normal" or energetically non-inducible wider grooves. This should not change the sequence specificity, which is predominately selected for by the aryl tetrasaccharide tail, but should allow additional sequences to be occupied.

The IH_in and IH_out models produced local widening of the minor grooves. In the IH_in model the bulge in the groove width is centered around the thiobenzoate. The groove distortion does not reach to the CORE and the model correctly predicts the experimentally observed hydrogen abstraction pattern. However, in the IH_out model the bulge in the minor groove starts around the thiosugar and extends past the CORE. The local swelling of the groove causes the 3'-AGGA strand to move away from the C3 radical. For this reason the IH_out model predicts only a single strand hydrogen abstraction of the H5"/C:6 by the C6 radical. Similar results were obtained when oligonucleotides containing either the 5'-TCCT/3'-AGGA calicheamicin hot spot or A-G base pair mismatches were treated with calicheamicin[17]. A-G base pair mismatches are thought to cause local widening of the minor groove with out adversely affecting the overall DNA conformation[41]. When the base pair mismatch was introduced in place of the 3'-C-G base pair (i.e., 5'-TCAT/3'-AGGA) no affect on calicheamicin's ability to cause double strand breaks was observed. However, when the mismatch was move to the 5'-C-G site (5'-TACT/3'-AGGA) only the top strand was cleaved. The

base pair mismatch experiments and this modeling study suggests that the CORE, hydroxylamino sugar and ethylamino sugar bind more tightly to the 5'-TCCT (upper) strand of the preferred binding/cleavage site.

CONCLUSION

Six models of the DNA-calicheamicin complex were constructed and the stability of each model was tested using solvated molecular dynamics. Five of the models produced results that are consistent with the DNA hydrogen abstraction data. However, only the I_out model provided a clear rationale for the high sequence specificity and unidirectional mode of binding exhibited by calicheamicin at 5'-TCCT sites. This modeling study suggests that the interactions between the DNA and the calicheamicin CORE, hydroxylamino sugar and ethyamino sugar residues are independent of the remainder of the molecule and that it is the interaction of these fragments with the DNA backbone that positions the CORE within the DNA for efficient double stranded DNA cleavage. The thiosugar, thiobenzoate and rhamnose moieties, on the other hand, read the DNA minor groove and determine the directionality and sequence specificity. The models also suggest that thiobenzoate may have an edge sequence specificity, i.e., in the 5'-TCCT binding site the dimethoxy edge appears to be preferred near the floor of the minor groove but in 5'-TTTT sites the iodine edge may be preferred.

METHODS

Computational Details. The modeling studies were conducted with MSI's[42] Quanta/CHARMM software (version 3.1) and force field[27] running on an IBM RS/6000 workstation. The oligonucleotide and the drug were treated as electrically screened charged systems. The phosphate and ammonium were screened to -0.32 and +0.32, respectively, according to Manning counterion condensation theory[43] and a continuous dielectric of one was used. The shift/vshift functions were used to smooth long range interaction to zero at 9.0Å with a 10.0Å of cutlist[29] in both the energy minimization (EM) and molecular dynamics (MD). The hydrogen bond potential was used. The potential was smoothed to zero between 4.0-5.0Å and 130-110° with a cutlist of 5.5Å and 90°, respectively. The Verlet algorithm[44] was used to calculate the classical equations of motion for the atoms and the X-H bonds were fixed using the SHAKE algorithm[45] during MD. All in vacuum calculations were performed without cutoffs using a distance-dependent dielectric. All EM studies were minimized to a RMS gradient force of ≤ 0.100 with either a steepest descent (SD)[46] or an adopted-basis Newton Raphson (ABNR) minimizer[47] unless otherwise stated.

The CHARMM detailed dynamics protocol described in our previous work[18,29] was used unchanged.

Model Building. The d(CGACTCCTGC)·d(GCAGGAGTCG) sequence[12] was constructed in the B-DNA form using Quanta's nucleic acid builder. The oligonucleotide was minimized in vacuum with 50 steps of SD minimization to locate a local minimum in the vicinity of the starting structure. The minimized structure was then used in the DNA-calicheamicin docking studies.

Calicheamicin R was constructed by making the appropriate modification to our esperamicin C model[18]. The different diradical forms of calicheamicin were then docked with the DNA in a way that positioned C3 and C6 near H4' of thymine 18 (H4'/T:18) and H5" of cytosine 6 (H5"/C:6), respectively. The continuous

energy facility in Quanta was used in the docking procedure to find the optimal starting point for EM. The DNA was constrained and the individual complexes were minimized with 50 steps of ABNR minimization, followed by 50 steps of unconstrained SD then 500 steps of ABNR minimization prior to solvation. Each DNA-drug 10-mer complex comprising 20 nucleic acid and 6 calicheamicin R residues was solvated in a droplet of water containing 814 pre-equilibrated TIPS3[48] water molecules. The complexes were then minimized in a two step process to relax the solute/solvent interface. First solute was constrained and the system was minimized with 200 steps of SD followed by 1000 steps of ABNR minimization. Then the constraints were removed and the system was minimized with an additional 200 steps of SD followed by 1000 steps of ABNR. This procedure was repeated for each system producing six unique complexes.

ACKNOWLEDGMENTS

We would like to thank Dr. G. Ravishanker for many useful discussions on molecular dynamics tool chest (MDTC) analysis of the results. This work was supported by grants to DLB from the National Institutes of Health (GM-37909) and from the Bristol-Myers Squibb Company via a Connecticut Cooperative High Technology Research and Development Grant.

REFERENCES

1. Lee, M. D.; Dunne, T. S.; Siegal, M. M.; Chang, C. C.; Morton, G. O.; Borders, D. B. *J. Am. Chem. Soc.* **1987**, *109*, 3464-3466.
2. Lee, M. D.; Dunne, T. S.; Chang, C. C.; Ellestad, G. A.; Siegel, M. M.; Morton, G. O.; McGahren, W. J.; Borders, D. B. *J. Am. Chem. Soc.* **1987**, *109*, 3466-3468.
3. Golik, J.; Clardy, J.; Dubay, G.; Groenewold, G.; Kawaguchi, H.; Konishi, M.; Krishnan, B.; Ohkuma, H.; Saitoh, K. I.; Doyle, T. W. *J. Am. Chem. Soc.* **1987**, *109*, 3461-3462.
4. Golik, J.; Dubay, G.; Groenewold, G.; Kawaguchi, H.; Konishi, M.; Krishnan, B.; Ohkuma, H.; Saitoh, K. I.; Doyle, T. W. *J. Am. Chem. Soc.* **1987**, *109*, 3462-3464.
5. Goldberg, I. H. *Acc. Chem. Res.* **1991**, *24*, 191-198.
6. Konishi, M.; Ohkuma, H.; Matsumoto, K.; Tsuno, T.; Kamei, H.; Miyaki, T.; Oki, T.; Kawaguchi, H.; VanDuyne, G. D.; Clardy, J. *J. Antibiot.* **1989**, *42*, 1449-1452.
7. Konishi, M.; Ohkuma, H.; Tsuno, T.; Oki, T.; VanDuyne, G. D.; Clardy, J. *J. Am. Chem. Soc.* **1990**, *112*, 3715-3716.
8. Leet, J. E.; Schroeder, D. R.; Hofstead, S. J.; Golik, J.; Colson, K. L.; Huang, S.; Klohr, S. E.; Doyle, T. W.; Matson, J. A. *J. Am. Chem. Soc.* **1992**, *114*, 7946-7948.
9. Leet, J. E.; Schroeder, D. R.; Langley, D. R.; Hofstead, S. J.; Golik, J.; Colson, K. L.; Huang, S.; Klohr, S. E.; Lee, M. S.; Doyle, T. W.; Matson, J. A. *J. Am. Chem. Soc.* **1993**, *115*, 8432-8443.
10. Minami, Y.; Yoshida, K.-I.; Azuma, R.; Saeki, M.; Otani, T. *Tetrahedron Lett.* **1993**, *34*, 2633-2636.
11. Yoshida, K.-I.; Minami, Y.; Azuma, R.; Saeki, M.; Otani, T. *Tetrahedron Lett.* **1993**, *34*, 2637-2640.
12. Zein, N.; Sinha, A. M.; McGahren, W. J.; Ellestad, G. A. *Science* **1988**, *240*, 1198-1201.

13. Long, B. H.; Golik, J.; Forenza, S.; Ward, B.; Rehfuss, R.; Dabrowiak, J. C.; Catino, J. J.; Musial, S. T.; Brookshire, K. W.; Doyle, T. W. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 2-6.
14. Walker, S.; Landovitz, R.; Ding, W.-D.; Ellestad, G. A.; Kahne, D. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 4608-4612.
15. De Voss, J. J.; Townsend, C. A.; Ding, D.-W.; Morton, G. O.; Ellestad, G. A.; Zein, N.; Tabor, A. B.; Schreiber, S. L. *J. Am. Chem. Soc.* **1990**, *112*, 9669-9670.
16. Hangeland, J. J.; De Voss, J. J.; Heath, J. A.; Townsend, C. A.; Ding, W.; Ashcroft, J. S.; Ellestad, G. A. *J. Am. Chem. Soc.* **1992**, *114*, 9200-9202.
17. Zein, N.; Poncin, M.; Nilakatan, R.; Ellestad, G. A. *Science* **1989**, *244*, 697-699.
18. Langley, D. R.; Golik, J.; Krishnan, B.; Doyle, T. W.; Beveridge, D. L. *J. Am. Chem. Soc.* **1993**, *115*, 0000.
19. Drak, J.; Iwasawa, N.; Danishefsky, D.; Crothers, D. M. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 7464-7468.
20. Sugiura, Y.; Uesawa, Y.; Takahashi, Y.; Kuwahara, J.; Golik, J.; Doyle, T. W. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 7672-7676.
21. Christner, D. F.; Frank, B. L.; Kozarich, J. W.; Stubbe, J.; Golik, J.; Doyle, T. W.; Rosenberg, I. E.; Krishnan, B. *J. Am. Chem. Soc.* **1992**, *114*, 8763-8767.
22. Hawley, R. C.; Kiessling, L. L.; Schrieber, S. L. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 1105-1109.
23. Golik, J.; Krishnan, B.; Doyle, T. W.; VanDuyne, G.; Clardy, J. *Tetrah. Letters* **1992**, *41*, 6049-6052.
24. Walker, S.; Valentine, K. G.; Kahne, D. *J. Am. Chem. Soc.* **1990**, *112*, 6428-6429.
25. Chatterjee, M.; Cramer, K. D.; Townsend, C. A. *J. Am. Chem. Soc.* **1993**, *115*, 3374-3375.
26. Clore, G. M.; Gronenborn, A. M.; Brunger, A. T.; Karplus, M. *J. Mol. Biol.* **1985**, *186*, 435-455.
27. Momany, F. A.; Rone, R. *J.Comp. Chem.* **1992**, *13*, 888-900.
28. Ravishanker, G.; Swaminathan, S.; Beveridge, D. L.; Lavery, R.; Sklenar, H. *J. Biomol. Struct. Dyn.* **1989**, *6*, 669-699.
29. Langley, D. R.; Doyle, T. W.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 4395-4403.
30. Walker, S.; Murnick, J.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 7954-7961.
31. Note added in proof:

Table 4. Interresidue NOEs (predicted from ref 30) and distances (from the 355.5 psec snapshot) in the calicheamicin-DNA complex. (a) simple rotation of the methoxyl group reduces this distance to less then 3.0Å.

residue 1	proton	residue 2	proton	NOE	distance (Å)
CORE	8	hydroxyl amino	1'	m	3.60
	5		6'	w	5.29
	4		6'	w	4.97
	8	ethylamino	5a'''	m	3.33
	8		5e'''	m	3.69
hydroxyl amino	5'	thiosugar	1''	w	3.22
	6'		1''	m	2.98
	2'	ethylamino	1'''	s	2.25
thiosugar	3''	thiobenzoate	$7'^{V}$	w	3.61
thiobenzoate	$2'^{V}$-OCH_3	rhamnose	1^{V}	w	5.84[a]
	$2'^{V}$-OCH_3		6^{V}	m	2.40

Table 5. Intermolecular NOEs (predicted from ref 30) and distances (from the 355.5 psec snapshot) in the calicheamicin-DNA complex. (a) simple rotation of the methoxyl group reduces this distance to less then 3.0Å.

calicheamicin γ_1^I residue		DNA residue	H1'	H4'	H5', H5"	distance (Å)
hydroxyl amino	1'	C:6		w		2.17
	3'	C:6	w			4.33
	6'	C:6		s		3.30
thiosugar	2a"	C:6	w			3.25
	2e"	C:6	w			3.28
	3"	C:6	w			3.31
	4"	C:6	w	m		4.33, 2.33
	6"	C:6		m		4.52
thiobenzoate	$7'^V$	C:6	s			2.73
	$2'^V$-OCH_3	A:17		m		4.70[a]
rhamnose	1^V	A:17		w	m	4.76, 2.82
	2^V	G:16		w		3.07
	3^V-OCH_3	G:16		m		5.91[a]

Recent NMR studies (see ref 30) have shown that the calicheamicin-DNA complex containg the 5'-ACCT recognition sequence has a conformation simular to the I_in model. Tables 4 and 5 show a comparision between the NOEs (predicted from ref 30) and the distances measured from the 355.5 psec snapshot from the I_in model. While the calculated distances agree relatively well with the NOE data the calculations did not predict the the observed "shift away from B-form DNA" at the 5'-CC step.

32. Steward, J. J. P. *MOPAC, A Semi-Empirical Molecular Orbital Program, QCPE.* **1983**, 455.
33. Seiler, F. J. *MOPAC 6.0* **1990**,
34. Hurley, L. H.; Leslie Boyd, F. *Ann. Rep. Med. Chem.* **1987**, *22*, 259-268.
35. Lavery, R.; Pullman, B.; Corbin, S. *Nucl. Acids Res.* **1981**, *9*, 6539.
36. Walker, S.; Yang, D.; Kahne, D. *J. Am. Chem. Soc.* **1991**, *113*, 4716-4717.
37. Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535-556.
38. Kopka, M. L.; Fratini, A. V.; Drew, H. R.; Dickerson, R. E. *J. Mol. Biol.* **1983**, *163*, 129-146.
39. Aiyar, J.; Danishefsky, S. J.; Crothers, D. M. *J. Am. Chem. Soc.* **1992**, *114*, 7552-7554.
40. Nicolaou, K. C.; Tsay, S.-C.; Suzuki, T.; Joyce, G. F. *J. Am. Chem. Soc.* **1992**, *114*, 7555-7557.
41. Prive', G. G.; Heinemann, U.; Chandrasegaran, S.; Kan, L.-S.; Kopka, M. L.; Dickerson, R. E. *Science* **1987**, *238*, 498-504.
42. Molecular Simulations, 16 New England Executive Park, Burlington, MA 01803,USA.
43. Manning, G. S. *Quart. Rev. Biophys.* **1978** *11*, 179-246.
44. Verlet, L. *Phys. Rev.* **1967**, *159*, 98-105.
45. Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. *C. J. Comput. Phys.* **1977**, *23*, 327-341.
46. Hestenes, M.; Stieffel, E. ; Washington, DC, 1952.
47. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187-217.
48. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. V.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 929-935.

(*Received* 12 *October* 1993; *accepted* 24 *November* 1993)