Model of the interactions of calichemic in γ_1 with a DNA fragment from pBR322

RONALD C. HAWLEY, LAURA L. KIESSLING, AND STUART L. SCHREIBER

Department of Chemistry, Yale University, New Haven, CT 06511

Communicated by Harry H. Wasserman, October 17, 1988

ABSTRACT An analysis of the binding interactions of several DNA-drug complexes that utilize carbohydrates for DNA recognition has been undertaken. It is proposed that the carbohydrate residues function as general minor groove binding elements, and the stereochemistry of aglycone attachment sites is generally disposed to promote a right-handed helical geometry that is complementary to right-handed DNA. The constitution and stereochemistry of the DNA double-strand cleaving agent calichemicin γ_1 is consistent with this analysis. Docking experiments with computer-generated models of this drug and a dodecamer duplex that was found to serve as a calichemicin cleavage site were performed to gain insight into the origin of the drug's sequence-selective binding and cutting properties. A model is presented that provides a molecular level understanding of the double-strand cleavage patterns that result from the action of calichemicin γ_1 on DNA.

DNA binding molecules from nature have provided important insights into the factors controlling the molecular recognition of DNA (1, 2). X-ray crystallography has served as the primary vehicle for elucidating the nonbonded interactions between the host nucleic acid and guest molecules (3) that include water (4), linked heterocycles [e.g., netropsin (5), echinomycin (6)], fused aromatics [e.g., triostin A (7), daunomycin (8)], and peptides [e.g., trp repressor (9), phage 434 Cro (10)]. NMR spectroscopy (11) and molecular modeling (12) have contributed additional understanding into these systems and others that have not yielded to the demands of the former technique. Concepts such as spine of hydration binding, intercalative binding, and peptide binding motifs have been developed by these studies. Importantly, the relationship of ligand geometry and oligonucleotide conformational variability (13) to sequence-selective binding is beginning to be understood (14, 15).

A considerable number of DNA binding natural products are equipped with carbohydrate residues that are likely to serve as DNA recognition elements. A comparatively small body of structural information is available concerning the nonbonded interactions that stabilize these carbohydratenucleic acid complexes (8, 16). This paper examines the molecular interactions between a newly discovered and potent antitumor agent, calichemicin γ_1 , and a DNA binding site and offers an explanation for the relationship of drug geometry to the specific double-strand cleavage pattern elicited by this molecule.

General Features of Glycoconjugate Ligands to DNA

Several antibiotics that serve to illustrate a common geometric property of DNA binding glycoconjugates are illustrated in Fig. 1. In each instance, a rigid aglycone is attached to sugar residues (recognition elements), and in three cases by two ether linkages with trans stereochemistry. As will be

discussed, it is precisely this stereochemistry that is required to place both saccharide chains and the polycyclic core into the minor groove of the host DNA. Consider first the case of the anthracycline antibiotic A-447C (17). This compound is representative of a large class of DNA binding natural products that include close relatives such as cosmomycin D (18) and less highly glycosylated members such as daunomycin and doxorubicin (19). Although the structures of the DNA complexes of A-447C and cosmomycin B have not been examined, the crystal structure of a daunomycin-d(CpGp-TpApCpG) complex has been solved by x-ray crystallographic methods at 1.2 Å resolution (8). This structure is important as it represents the only crystallographic analysis of carbohydrate-nucleic acid nonbonded interactions. There are several striking features of the complex that are relevant to the analysis of other members of this class. Most conspicuous is the intercalative mode of binding of the aglycone in the d(CpG) sequence at the end of the helix. The α stereochemistry of the (D-ring) C-7 amino sugar and β stereochemistry of the C-9 acetyl group combine to provide the drug with a right-handed helical twist that is complementary (matched pair) to the host DNA. Indeed, daunomycin has been shown to shift the solution equilibrium of DNA conformers from the Z-form (left-handed) to the B-form (right-handed) found in the crystal structure of the complex (15). As noted by Rich, Wang, and coworkers (8), anthracyclines such as aclacinomycin A that are equipped with a β -carbo-methoxyl substituent at the C-10 position can be accommodated into the same structure without complications of steric interactions with this group and the DNA backbone that would be present with the C-10 epi (α) configuration (mismatched pair). The full complementarity of structure between A-447C (Fig. 1) and the host (B-form of DNA) is also apparent. The ether linkages that connect the intercalating aglycone to the two trisaccharide groups at C-7 and C-10 have the proper relative and absolute stereochemistry to direct the sugars into the minor groove of B-DNA. The ubiquitous (among DNA binding natural products) 1,4sugar linkage represents the connectivity pattern that maximizes the linearity of the trisaccharide and therefore allows for efficient winding into the minor groove. This structural feature can be identified in esperamicin A_1 and calichemicin γ_1 (see below).

Neocarzinostatin (NCS) chromophore is the nonpeptidyl and biologically significant component of the antitumor agent NCS (20). Goldberg (21) has shown that NCS chromophore binds in the minor groove of the B-form of DNA ($K_d \approx 10^{-6}$ M) by intercalation of the naphthoate between base pairs in A+T-rich regions (21). Note the trans stereochemistry of the alkoxy substituents (recognition elements) that are attached, with a vicinal relationship, to the bicyclic core. The absolute stereochemistry of these centers (22, 23) results in a geometry of the drug that is complementary to the right-handed DNA host (matched pair). The naphthoate serves as a point of attachment to the DNA; the intercalative mode of binding

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. \$1734 solely to indicate this fact.

Abbreviation: NCS, neocarzinostatin.



FIG. 1. Aglycone stereochemistry directs sugars into the minor groove of B-DNA: A general feature of DNA binding/damaging natural products.

orients this group parallel to the base pairs. The rest of the drug must wind its way into the minor groove, as shown by the experiments of Goldberg. Three bonds connect the bicyclic core to the naphthoate; only one of these (indicated) is easily rotated. It is well appreciated that the lowest energy rotamer of this functional grouping (ester of a secondary alcohol) prefers a conformation with the methine C-H bond nearly parallel to the carbonyl C=0 bond (24). Moreover, this conformation is found in the crystal structure of a DNAtriostin A complex in a related functional group (7). If the NCS chromophore-DNA complex were to adopt the same local conformation as that found in the triostin A-DNA complex, then the R,R-stereochemistry would be required to wind the bicyclic core into the minor groove and place the amino sugar (D-fucosamine) in a position in the minor groove that is analogous to the amino sugar of daunomycin. The S,S-enantiomorph in a similar conformation would present a mismatched (left-handed helical) geometry. The acetylenic carbon that serves as the progenitor to a DNA damaging aryl radical by a priming mechanism (involving activation with a nucleophilic thiol agent) analogous to the reaction transformation recently discussed by Myers (23) is found to be proximal to the 5' carbon of the deoxyribose of the DNA backbone. Kappen and Goldberg (25) have shown that the thiol-mediated DNA single-strand cleavage of the DNA-drug complex proceeds via the intermediacy of a 5' carbon radical.

Esperamicin A_1 is a potent antitumor agent that was recently discovered by a group at Bristol-Meyers (26). This molecule (Fig. 1) is representative of a class of DNA binding/damaging agents that function as highly efficient double-strand DNA cleaving reagents. In the context of earlier discussions concerning the stereochemistry of ether linkages of DNA recognition elements, a commonality between esperamicin A_1 and the anthracycline class and NCS becomes apparent. The *S*,*S*-stereochemistry of the two carbon centers of the bicyclic core equipped with the ether linkages are well suited for binding to the right-handed helix of B-DNA. It is seen that this absolute stereochemistry requires the enediyne portion of the bicyclic core be placed into the groove of DNA with the plane of this group oriented approximately perpendicular to the base pairs. The short length of this group allows a perpendicular arrangement that does not interfere with hydrogen bonding present in the Watson–Crick base pairs. Although the sequence selectivity of esperamicin A₁ binding to DNA has not been reported, the recently described DNA cleavage properties of the related natural product calichemicin γ_1 provide support for this model. In the following section, a model of a calichemicin γ_1 – DNA complex will be presented that provides an explanation for the unique DNA-cleaving properties of the natural product and suggests opportunities for the design of new chemotherapeutic agents and DNA cleaving reagents.

Calichemicin γ_1 : An Efficient Double-Strand Cleaving Agent

An unusual mechanism for the DNA cleaving properties of esperamicin A₁ (26) and calichemicin γ_1 (27, 28) has been proposed that is supported by the chemistry of these compounds. These natural products were suggested to first form a complex with the DNA target. The enediyne function present within the core structure (Fig. 2, 1) undergoes coupling [Bergman reaction (29)] to the benzene diyl 2. Hydrogen atom abstraction occurs at both strands of the DNA duplex to provide the reduction product 3 and a DNA diradical that undergoes reaction with dioxygen, ultimately leading to double-strand breakage. Note that direct cyclization of 1 would result in a highly strained bridgehead olefin contained within a bicyclo[3.3.1]nonane skeleton. The natural products 1 utilize a priming mechanism to effect DNA damage. It has been proposed (26, 28) that a bioreductive activation of 1 proceeds by reaction of a bionucleophile (30) with the trisulfide. The resultant thiolate is utilized to convert the neighboring trigonal center (at the bridgehead) to a tetrahedral center by way of a Michael addition reaction. Such an action would be expected to lower the activation barrier of the Bergman reaction.

The unusual structural properties of natural products such as 3 and NCS chromophore, which likely operates in a similar



FIG. 2. Proposed cleavage mechanism.

manner (22, 23), suggest (i) a role for diradicals as DNA cleaving reagents and (ii) priming mechanisms for the timely production of these high-energy species. To achieve the goal of nonnatural agent/reagent design, consideration must also be given to the geometric requirements for effective DNA recognition. These natural products offer insight into this aspect of the problem as well. Consider the case of calichemicin γ_1 .

A Calichemicin γ_1 Binding Model

A summary of the results of recent investigations of the DNA cleaving properties of calichemicin γ_1 toward a restriction fragment from pBR322 is provided in Fig. 3 (31). Since calichemicin is equipped with its own cleaving function (cyclic enediyne) this work represents a type of affinity cleavage experiment. Several striking results from these studies are (i) calichemicin promotes efficient double-strand cleavage of DNA. (ii) specific cleavage sites are found with 3-base-pair separation between strands, (iii) 3' asymmetry is observed that suggests minor groove binding, (iv) sequence selective binding at the 4-base-pair recognition level; the most efficient binding site contains the sequence d(Ap-GpGpA) [or, equivalently, d(TpCpCpT)], and (v) cleavage occurs at the 5' cytidine of a TCCT sequence and two nucleotides away from the 3' adenine (in the 3' direction) of the AGGA complement. In addition, from the electrophoretic properties of cleaved fragments and their derivatives (e.g., after treatment with NaBH₄) Zein et al. (31) inferred that hydrogen atom abstraction takes place at the C-5' position of the cytidine group. The resultant products of this reaction are the corresponding 5'-aldehyde and carboxylic acid derivatives of the cytidine nucleotide. The site of oxidation at the N_2 site on the opposite strand was not identified, although the data suggest that the major site of hydrogen atom abstraction (and oxidation) is the C-1' position. The electrophoretic



properties of the cleavage fragments are reminiscent of those reported by Sigman (32) in his cleavage studies of DNA that utilize the reagent copper phenanthroline. In this work, Sigman has shown that oxidation at the C-1' site results in the formation of 5-methylene-2-furanone from an intermediate 2'-deoxyribonolactone derivative (abasic intermediate) with liberation of DNA fragments that are terminated in 3'- and 5'-phosphomonoesters (33). The 3'-terminal phosphoglycolate that derives from C-4' hydrogen atom abstraction has increased mobility relative to the markers used in Maxam-Gilbert sequencing (34). In fact, the high-resolution gel electrophoretic mobilities of the products derived from oxidation at each of the accessible sites of a deoxyribose are distinct from one another (33-37).

The general features of the binding model are illustrated in Fig. 4. First, note that the 1,4-hydroxylamino-linked disaccharide, the 4-substituted thioester sugar, and the p-substituted phenyl ring combine to present a linear array of sugar groups that are reminiscent of the trisaccharide found within A-447C. (i) The 4-base-pair recognition is due largely to interactions of the polysaccharide and aromatic moieties of calichemicin with the host DNA. This is symbolized by the oval, labeled as an AGGA recognition element, in Fig. 4. (ii) The attachment of the bicyclic core to the AGGA recognition element by the ether linkage with the S-stereochemistry (38) provides the drug (in an approximately perpendicular binding mode) with a right-handed helical structure that is a stereochemical match for the minor groove of the B-DNA host. In a parallel binding mode, the S-stereochemistry would result in a mismatched combination (left-handed helical structure of the drug). (iii) The cyclic enediyne serves as a nondiffusible double-strand cleaving function toward duplex DNA. The stereoelectronic requirements for hydrogen atom abstraction (see below) combine with the above structural features of the complex to result in the experimentally observed DNA cleavage pattern.

Docking experiments with calichemic γ_1 were performed with MacroModel on a VAX 11-750 and an E&S PS390 graphics terminal and utilized with $d(N^{568}-N^{579})$ cleavage site [d(CCATTCCTTGCG)] on the BamHI/Sal I restriction fragment of pBR322 that was used in the affinity-cleavage experiments. A B-form input structure of the oligonucleotide duplex was minimized for 2000 iterations with the AMBER force field. The preminimized drug was then manually docked, followed by minimization with the MM2 forcefield to an rms gradient of <0.10 utilizing the substructure minimization mode of the modeling program MacroModel. This allowed movement of all DNA atoms within 4 Å of any atoms of the drug, with distance-dependent constraints. The quality of fit of the bicyclic core was found to be highly dependent on its location within the minor groove. When the olefin of the enediyne was placed in a region that straddled 2 base pairs (Fig. 5), energy minimization led to rapid convergence with little displacement from the input structure. When the core was vertically displaced from this region (e.g., the olefin of the enediyne is raised up to the level of a base pair), energy



FIG. 4. DNA binding model.



FIG. 5. (a) Computer-generated energy minimized model of calichemicin γ_1 (yellow)–DNA (orange) complex. Hydrogen atoms that are abstracted by the activated calichemicin biradical are blue. (b) View of calichemicin γ_1 bicyclic core (yellow)–DNA (orange) interactions at double-strand cleavage site. (c) View of proposed interactions of hexasubstituted aromatic of calichemicin γ_1 with GpG dinucleotide within the calichemicin binding site. Nitrogen (from C-2 amino of guanine)–iodine distances are calculated to be within 3–4 Å.

minimization led to poor convergence with a large displacement of the drug out of the minor groove. The general topographical features of the minor groove provide a complementarity of fit only when the bicyclic core is bound in the region of space between 2 base pairs. The geometries of the drug and DNA are such that when the cleaving function is placed between the base pairs associated with the $d(N^1pN^2)$ dinucleotide (Fig. 3), the polysaccharide/aryl moiety, when wound into the minor groove, spans the entire AGGA (equivalent to TCCT) recognition site and only that site.

What are the consequences of this placement with regard to the double-strand cleavage of DNA? To answer this question, the stereoelectronic requirements for hydrogen atom transfer to the enediyne must be considered. Only those "on-line" hydrogen atoms of the DNA donor that are in the plane formed by the enediyne (and the subsequently formed benzene diyl) are well suited for transfer to the intermediate drug diradical. The transfer of "off-line" hydrogen atoms would result in a transition state structure that suffers from partial C—H bonds that are bent relative to the product geometry (C—H bond in the plane of the benzene ring). The energetically favored on-line geometry for hydrogen atom transfer leads precisely to the 3-base-pair separation (Fig. 5) that has been observed experimentally (31).

The computer-generated models in Fig. 5 were constructed from docking and energy minimization experiments that followed the aforementioned guidelines. This complex represents a local energy minimum that is suggested to be related to the actual geometry of a "preprimed" calichemicin-DNA complex. As the priming/Bergman/hydrogen abstraction reactions proceed along their respective reaction coordinates, geometric perturbations will undoubtedly occur; nevertheless, the indicated complex illuminates the relationship of the drug's binding selectivity to the site of double-strand cleavage. On inspection of this model, the double-strand cleavage patterns that result from the action of calichemicin y_1 on DNA can be understood at the molecular level. The requirement for on-line hydrogen atom abstraction from this complex leads to an understanding of (i) the 3-base-pair separation between cuts of the two strands, (ii) the relationship of the sequence selective binding to the cleavage pattern, and (iii) the site of hydrogen atom abstraction within the target deoxyribose sugars (C-5' position of cytidine, C-1' position of the N^2 nucleotide). Concerning the latter point, the model cannot clearly distinguish this path from an alternative process that is apparently not observed. The (pro-S) C-5' hydrogen of the N³ nucleotide appears available for transfer to the enediyne and such a reaction would result in double-strand cleavage with 4-base-pair separation. Apparently, a subtle geometric factor results in the lower activation barrier for the C-1 hydrogen transfer process.





Chemistry: Hawley et al.

Nevertheless, the location of only three plausible sites (two are observed) of hydrogen atom abstraction provides considerable support for the model when the numerous candidate sites on the host oligonucleotide are considered.

The most significant feature of the model with regard to the sequence selective binding of the drug concerns the positioning of the hexasubstituted aromatic ring. As shown in Fig. 5, the aryl group lies approximately parallel to the base pairs and between the d(CpC)·d(GpG) tetrad. Of the cleavage sites noted by the Lederle workers, this tetrad is encountered with high frequency (31). This finding supports the hypothesis that the iodo substituent on the aromatic ring of the drug forms a favorable interaction with the N² amino substituents of the two guanines of the d(GpG) dinucleotide. The large and polarizable iodo substituent is ideally equipped to function in this bridging manner. To gather evidence for the existence of an attractive force between the nitrogen and iodine substituents, a search of the Cambridge Crystallographic Data Base (CCDB) for structures with an intermolecular N-I distance of <5 Å was performed (39). Two of many examples of the noncovalent bonding interactions between these atoms in the solid state are provided in Fig. 6 (40, 41). In each instance, it is seen that the interatomic distance is considerably shorter than the sum of their van der Waals radii. These structures provide support for the existence of an attractive (bonding) interaction between these atoms and the suggestion that the glycosylated aromatic ring in the natural product serves as a GpG recognition element.

The authors wish to express their gratitude to Prof. D. M. Crothers (Department of Chemistry, Yale University) for stimulating discussions on this topic. We thank Prof. W. C. Still for providing us with a copy of the program MacroModel. Financial support from the National Institute of General Medical Sciences and the National Science Foundation is gratefully acknowledged.

- 1. Remers, W. A. (1985) J. Nat. Prod. 48, 173–192.
- 2. Dervan, P. B. (1986) Science 232, 464–471.
- 3. Saenger, W. (1984) Principles of Nucleic Acid Structure (Springer, New York), pp. 253-282.
- 4. Prive, G. G., Heinemann, U., Chandrasegaran, S., Kan, L.-S. & Dickerson, R. E. (1987) Science 287, 498-504.
- Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R. E. (1985) Proc. Natl. Acad. Sci. USA 82, 1376–1380.
- Ughetto, G., Wang, A. H.-J., Quigley, G. J., Pjura, P., van der Marel, G. A., van Boom, J. H. & Rich, A. (1985) Nucleic Acids Res. 13, 2305-2323.
- Wang, A. H.-J., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G., van Boom, J. H. & Rich, A. (1984) Science 225, 1115-1120.
- Wang, A. H.-J., Ughetto, G., Quigley, G. J. & Rich, A. (1987) Biochemistry 26, 1152-1163.
- Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F. & Sigler, P. B. (1988) Nature (London) 335, 321–329.
- Wolberger, C., Dong, Y., Ptashne, M. & Harrison, S. C. (1988) Nature (London) 335, 789-795.

- 11. Searle, M. S., Hall, J. G., Denny, W. A. & Wakelin, L. P. G. (1988) *Biochemistry* 27, 4340-4349.
- 12. Rao, S. N., Singh, U. C. & Kollman, P. A. (1986) J. Med. Chem. 29, 2484-2492.
- Peticolas, W. L., Wang, Y. & Thomas, G. A. (1988) Proc. Natl. Acad. Sci. USA 85, 2579–2583.
- Coll, M., Frederick, C. A., Wang, A. H.-J. & Rich, A. (1987) Proc. Natl. Acad. Sci. USA 84, 8385–8389.
- Chaires, J. B., Fox, K. R., Herrera, J. E., Britt, M. & Waring, M. J. (1987) *Biochemistry* 26, 227-236.
- Keniry, M. A., Brown, S. C., Berman, E. & Shafer, R. H. (1987) Biochemistry 26, 1058-1067.
- Shimosaka, A., Hayakawa, Y., Nakagawa, M., Furihata, K., Seto, H. & Otake, N. (1987) J. Antibiot. 50, 116–121.
- Ando, T., Hirayama, K., Takahashi, R., Horino, I., Etoh, Y., Morioka, H., Shibai, H. & Murai, A. (1985) *Agric. Biol. Chem.* 49, 259-262.
- 19. Crooke, S. T. & Reich, S. D. (1980) Anthracyclines: Current Status and New Developments (Academic, New York).
- Edo, K., Mizugake, M., Koide, Y., Seto, H., Furihata, K., Otake, N. & Ishida, N. (1985) *Tetrahedron Lett.* 26, 331-334.
- Goldberg, I. H. (1987) in Mechanisms of DNA Damage and Repair, eds. Simic, M. G., Grossman, L. & Upton, A. C. (Plenum, New York), pp. 231-243.
- 22. Myers, A. G., Proteau, P. J. & Handel, T. M. (1988) J. Am. Chem. Soc. 110, 7212-7214.
- 23. Myers, A. G. (1987) Tetrahedron Lett. 28, 4493-4496.
- 24. Schweizer, W. B. & Dunitz, J. D. (1982) Helv. Chim. Acta 65, 1547–1554.
- 25. Kappen, L. S. & Goldberg, I. H. (1985) Nucleic Acids Res. 13, 1637–1648.
- Golik, J., Dubay, G., Groenewold, G., Kawaguchi, H., Konishi, M., Krishnan, B., Ohkuma, H., Saitoh, K. & Doyle, T. W. (1987) J. Am. Chem. Soc. 109, 3462-3464.
- Lee, M. D., Dunne, T. S., Siegel, M. M., Chang, C. C., Morton, G. O. & Borders, D. B. (1987) J. Am. Chem. Soc. 109, 3464–3466.
- Lee, M. D., Dunne, T. S., Chang, C. C., Ellestad, G. A., Siegel, M. M., Morton, G. O., McGahren, W. J. & Borders, D. B. (1987) J. Am. Chem. Soc. 109, 3466–3468.
- Lockhart, T. P., Comita, P. B. & Bergman, R. C. (1981) J. Am. Chem. Soc. 103, 4082–4090.
- Kappen, L. S., Chen, C. & Goldberg, I. H. (1988) Biochemistry 27, 4331-4340.
- Zein, N., Sinha, A. M., McGahren, W. J. & Ellestad, G. A. (1988) Science 240, 1198–1201.
- 32. Sigman, D. S. (1986) Acc. Chem. Res. 19, 180-186.
- 33. Goyne, T. E. & Sigman, D. S. (1987) J. Am. Chem. Soc. 109, 2846-2848.
- Takeshita, M., Grollman, A. P., Ohtsubo, E. & Ohtsubo, H. (1978) Proc. Natl. Acad. Sci. USA 75, 5983-5987.
- 35. Stubbe, J. & Kozarich, J. W. (1987) Chem. Rev. 87, 1107-1136.
- Dizdaroglu, M., Schulte-Frohlinde, D. & von Sonntag, C. (1977) Z. Naturforsch. C 32, 1021–1022.
- Kappen, L. S. & Goldberg, I. H. (1983) Biochemistry 22, 4872– 4878.
- 38. Kende, A. S. & Smith, C. A. (1988) Tetrahedron Lett. 29, 4217-4220.
- 39. Bent, H. A. (1968) Chem. Rev. 68, 587-648.
- 40. Bjorvatten, T. & Hassel, O. (1962) Acta Chem. Scand. 16, 249-251.
- 41. Schlempter, E. O. & Britton, D. (1965) Acta Crystallogr. 18, 419-421.