Molecular models of neocarzinostatin damage of DNA: analysis of sequence dependence in 5'GAGCG:5'CGCTC

A.Galat^{1,2*} and I.H.Goldberg¹

¹Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, MA 02115 and ²Department of Chemistry, 12 Oxford Street, Harvard University, Cambridge, MA 02138, USA

Received December 14, 1989; Revised and Accepted March 16, 1990

ABSTRACT

Model building and molecular mechanics and dynamics calculations have been performed on a number of complexes of the post-activated form of the neocarzinostatin chromophore (NCS) with the B-DNA oligomer 5'GAGCG:5'CGCTC. Stable structures with the naphthoic acid moiety intercalated at all base pairs can be constructed. The observed bistranded lesions consisting of an abasic site at the Cyt residue in AGC and a direct break at the Thy residue on the complementary strand can be explained by assuming that NCS in the (R,R) form intercalates between the Ade2-Thy9/Gua3-Cyt8 base step with its 'diradical' core oriented towards the 3'-end of the (+) strand. Sites at C5', C4' and C1' in the minor groove are within a short enough distance from the two radical centers on NCS to permit hydrogen atom abstraction and the formation of the bistranded lesions. Strand cleavage at Thy9 may occur as a single lesion if NCS is intercalated into the Gua3-Cyt8/Cyt4-Gua7 base step with its active core towards the 3'-end of the (-) strand. The results are analyzed, and the utility and limitations of this type of model building are discussed.

INTRODUCTION

Neocarzinostatin (NCS, see Fig. 1a) is a DNA-damaging antitumor antibiotic [1] composed of a highly strained diacetylene-containing chromophore [2-4] non-covalently complexed with its carrier protein. Binding of the ene-diyne bicyclic core of NCS in the minor groove of DNA has been proposed with the naphthoate moiety intercalated between the DNA base pairs and the amino sugar moiety interacting with the DNA backbone [5]. Spectrophotometric titrations have revealed that there are at least two types of binding sites for NCS in DNA, a tight one saturating at one drug molecule per four base pairs and a less preferred one with a binding site per two base pairs [6,7].

NCS induces mainly single strand breaks; most of the double strand breaks can be accounted for by the random placement of single breaks within a few nucleotides of one another on the complementary DNA strands [8-10]. Direct strand cleavage by

NCS have been shown to result from its activation by thiol to a presumed free radical form [11,12] that abstracts a hydrogen atom from the C5' of deoxyribose. In the presence of oxygen this leads to breaks having mainly a nucleoside 5'-aldehyde at the 5'-side and a phosphate at the 3'-side [11]. It has been established that NCS attacks mostly at Thy and Ade and more rarely at Cyt and Gua [13,14]. In addition to direct strand breaks, NCS produces abasic sites which can result in alkali-labile breaks [15-17]. These lesions account for up to 25% of the total strand breaks under certain conditions [10]. They occur at Cyt residues in the preferred sequence AGC and have been implicated in mutagenesis in *E.coli* [18,19]. Recently it has been found that such lesions may be generated by the abstraction of a hydrogen atom from the C1' of the Cyt residue in AGC to form 2-deoxyribonolactone with intact phosphodiester linkages [20,21]. In contrast to other strand breaks, the abasic site at Cyt is always accompanied by a direct break on the complementary strand two nucleotides to the 3'-side at Thy [22]. It is believed that the activation of NCS by thiol (or sodium borohydride) occurs via nucleophilic addition at C12 and rearrangement of the bicyclic ene-diyne core into a diradical species with radical centers at the C2 and C6 atoms [3,12] (Fig. 1b). These free radical centers are responsible for hydrogen atom abstraction from the carbons of DNA deoxyribose [23]. This suggests that the binding geometry at AGC may allow for simultaneous attack by the diradical form of the drug at the C1' of Cyt of one strand and at the C5' of Thy of the other strand. However, since the number of breaks at Thy exceed the abasic sites at the Cyt [20], more than one mechanism is likely to be involved in the attack at Thy.

The experimental conclusions outlined above raise a number of questions that can be addressed by model building and molecular mechanics and dynamics techniques. We consider possible intercalation sites which have free radical centers in position to attack the different sugar moieties. The pentamer sequence 5'd(Gua₁ Ade₂ Gua₃ Cyt₄ Gua₅(+)):5'd(Cyt₆ Gua₇ Cyt₈ Thy₉ Cyt₁₀(-)), which has been studied experimentally [20], is used to examine whether there exists a binding geometry that permits simultaneous hydrogen atom abstraction from the C5' atom of Thy₉ and the C1' atom of Cyt₄ to form a direct strand break and an abasic site, respectively. We consider NCS

* To whom correspondence should be addressed

intercalation sites on the pentamer and determine which ones can lead to double stranded attack at Thy9 and Cyt4 and which ones can attack only at Thy9. Although experiments indicate that NCS occurs naturally as the C10(R)-C11(R) stereoisomers [24], we consider both C10(R)-C11(R) and C10(S)-C11(S) in model building; the latter could presumably be synthesized and any differences between the behavior of (R,R) and (S,S) would be of interest. It appears that both (R,R) and (S,S) can be positioned to produce bistranded lesions. The naphthoic acid moiety of NCS (R,R) must be intercalated at the Ade2-Gua3 base step with the active portion oriented towards the 3'-end of the GAGCG strand. By contrast, there are several possible intercalation sites for the single strand cleavage with attack at Thy9. We examine briefly some additional sequences for which NCS cleavage data are available.

METHODS

We outline the methods used with emphasis on the present applications. Models were built for the two components, the B-DNA oligomer and NCS in its activated diradical form. Docking was done visually by adjusting the DNA structure to permit intercalation between different sets of base pairs and orienting the NCS to obtain what appeared to be a reasonable starting structure. This was followed by energy minimization and in some cases molecular dynamics simulations to adjust the structure and determine its stability. Finally, the resulting models were analyzed in terms of their energetics and suitability for carrying out the reactions of interest.

All calculations were performed with the CHARMM program. Polar hydrogen DNA parameters were taken from references [25,26]. Because the structure of NCS involves elements not available in standard CHARMM set, the geometry and parameters had to be determined. Since we are interested in studying binding sites that permit reaction with certain sugar carbons, we constructed not NCS itself (Fig. 1a), but rather the post-activated form (Fig. 1b). It differs from the postulated active diradical only by the presence of hydrogens at positions C2 and C6; the distance between C2 and C6 is 3.58 Å. In the models, we consider as possible reactive sites sugar carbons C1', C4', or C5' which are less than 5 Å from position C2 or C6 of NCS. Although no detailed theoretical study of the hydrogen abstraction reaction has been made, *ab initio* calculations of free radical reactions suggest that the mean heavy atom distances in the transition state is on the order 2.80 Å [27]. Since hydrogen atoms were present at the 'radical centers' in the present calculation and no attempt was made to mimic the reaction, we consider a longer distance in analyzing the simulation results.

The Cartesian coordinates of NCS were built based on the x-ray coordinates of 1-hydroxy-2-carboxynaphthoic acid [28]. The remaining parts of NCS were added using bond length and angles established from empirical and *ab initio* data [29]. Two different configurations for the C10-C11 stereocenters of NCS were utilized, namely S,S [4] and (R,R) [24]. Each of the molecules has α -D-N-methylfucosamine as the sugar moiety [30]. The force constants for the unsaturated part of NCS were obtained from available references [25,26]. Charges for the naphthoic acid residue were determined from optimized structures of two substituted benzene subfragments using Gaussian 82 [31]; the structures were optimized using a 3-21G basis set followed by single point computation with a 6-31G basis set to determine the charges from a Mulliken population analysis [29]. The charges

for the sugar and the unsaturated moieties were obtained from the MOPAC program [32].

Starting coordinates of 5'GAGCG:5'CGCTC were determined from the standard B-DNA model cylindrical polar coordinates of Arnott and Hukins [33] with the CORGEN program [34]. To intercalate a planar aromatic molecule into DNA two base pairs must be vertically separated along the Z-axis. Also, there must be an 'unwinding' of DNA at the intercalation site [35]. Molecules which intercalate into DNA with their long axis parallel to the long axis of the base pair usually require larger unwinding angle than those whose long axis is perpendicular to the base pair. The unwinding of DNA and change in the base pair separation for the intercalation site are introduced with the rotation-translation matrix

$$\begin{pmatrix} X^{\text{rot}} \\ Y^{\text{rot}} \\ Z^{\text{rot}} \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \begin{pmatrix} \cos(\Phi) & \sin(\Phi) & 0.0 \\ \sin(\Phi) & -\cos(\Phi) & 0.0 \\ 0.0 & 0.0 & \Delta_k \end{pmatrix} \quad (1)$$

where Δ_k is the vertical rise required for the intercalator and Φ is the unwinding angle. Only the DNA segment below the intercalation site is displaced. The phosphate backbone bonds of the intercalation site that are broken during such an operation are rebuilt using values of torsion angles taken from the X-ray diffraction structures of ethidium bromide-DNA complexes [35].

Docking of NCS to DNA with the active portion of NCS in the minor groove was performed with the use of the HYDRA graphics system [36]. In addition to the positioning of the intercalated aromatic moiety, the molecule has rotational freedom about six dihedral angles ($\Psi_1 - \Psi_6$; see Fig. 1). The rotation about Ψ_1 and Ψ_2 is likely to be partially restricted due to conjugation of the π -electrons of the ester group with the naphthalene ring. Moreover, the 2-hydroxyl group may form an internal hydrogen bond to the carbonyl oxygen. We explored DNA-NCS complexes which have the ester group in the plane of the naphthalene ring and also some where the group is out-of-plane.

To examine other DNA sequences, the parent structure (5'GAGCG:5'CGCTC) was mutated into 5'GGGCG:5'CGCCC, 5'GAICG:5'CGCTC and 5'GAGCG:5'CICTC. In each case standard bond length, bond angles and torsion angles were used.

The empirical potential energy used in model-building, energy minimization, and molecular dynamics has the form [37]

$$E = E_b + E_a + E_t + E_i + E_{\text{vdw}} + E_{\text{el}} + E_{\text{hb}} \quad (2)$$

where E_b is the bond energy, E_a is the bond angle energy, E_t is the torsion angle energy, E_i is the improper torsion angle energy, E_{vdw} is the van der Waals energy, E_{el} is the electrostatic energy, and E_{hb} is the hydrogen bond energy. With respect to E_{el} , procedures employed previously for DNA simulations in vacuum were followed; i.e., the effect of solvent was approximated by a 1/R dielectric screening function [37] and by reducing the net charge on the phosphate group to $-0.32e$ [38]. The nonbonded interactions were switched off between 9.5 Å and 10.5 Å, by use of a cubic switching function, with pairs up to a distance of 11.5 Å included in the nonbonded list. After model building, each structure was optimized with 1000 steps of the Adapted Basis Newton-Raphson minimization method (ABNR) [37]. The net energy of complexation is the sum of the energy required to distort the DNA helix, the energy required to distort the NCS molecule and the interaction energy of the two [39-41]; i.e., $\Delta E = E_{\text{complex}} - E_{\text{DNA}} - E_{\text{NCS}}$ [40,41]. The energy of DNA (E_{DNA}) and NCS (E_{NCS}) were obtained from the

Table 1. Energy and structure of NCS complexes with 5'GAGCG:5'CGCTC^a

Structure	E	ΔE	ψ_1	ψ_2	ψ_3	Sites ^b
Model-1a	-263.9	-51.9	31.2	178.0	-63.2	C2: C4'(Cyt4)4.45; C5'(Cyt4)4.79 C6: C4'(Thy9)4.38
Model-1b	-290.1	-78.1	24.9	-164.2	77.0	C2: C5'(Cyt10)4.78 C6: C4'(Cyt4)4.57; C5'(Cyt5)4.48
Model-2a	-289.4	-77.4	-3.0	34.9	109.3	C2: C1'(Cyt4)4.43; C4'(Cyt4)3.30 C6: C4'(Thy9)3.32; C5'(Thy9)3.31
Model-2b	-273.4	-61.4	32.2	-107.0	121.2	C2: C5'(Thy9)4.41 C6: C1'(Cyt4)4.69; C4'(Cyt4)4.68
Model-3a	-273.5	-61.5	2.6	125.6	-142.3	C2: C5'(Thy9)4.54
Model-3b	-259.8	-47.8	24.0	-169.0	-76.4	C2: C4'(Cyt4)4.75
Model-4a	-312.6	-99.4	23.7	179.0	129.5	C6: C1'(Cyt4)3.24; C4'(Cyt4)3.50; C5'(Cyt4)4.87
Model-4b	-268.8	-56.8	2.3	-174.0	-59.3	C2: C4'(Cyt8)4.38 C2: C4'(Cyt4)3.42; C5'(Cyt4)4.97 C1'(Cyt4)3.34 C6: C4'(Thy9)3.35; C5'(Thy9)3.11

^a Distances in Å, angles in degrees and energies in kcal/mole. The numbering corresponds to the four possible intercalation sites; e.g., 1 is between Gua₁-Cyt₁₀ and Ade₂-Thy₉. Models a and b refer to (R,R) and (S,S), respectively. Models 1 and 2 have the NCS pointing toward the 3' end of (+) strand while models 3 and 4 have it pointing toward the 5' end of (+) strand. The unwinding angle varies between 17 and 27° and the increase in length in between 2.6 and 3 Å.

^b Sites are listed if the distance from the free radical center to a sugar carbon is less than 5 Å.

optimized structures of the isolated molecules; their values are -312 and 100 kcal/mole, respectively. Since no solvent is included, the energies are only indicative of the stabilities; they permit one to determine sterically appropriate structures and exclude those that have a very high energy. However, quantitative comparisons of the similar energies cannot be used to conclude that one structure is preferred over another.

Molecular dynamics simulations of the DNA-NCS complexes were performed to examine the stability of certain models. They were started from the optimized structures. The classical equations of motions for the atoms were integrated by use of the Verlet algorithm [42] with a time interval of 0.001 ps per step. The nonbonded interaction lists were updated every 0.02 ps. Bond lengths involving hydrogen atoms were kept fixed by use of the SHAKE algorithm [43]. To minimize end effects, the terminal base-pairs were restrained with harmonic constraints equal 10 kcal/mole-Å per atom. The structure was heated from 10°K to 300°K over 5 ps followed by 5 ps of equilibration. The production portion of the dynamics simulation was then carried out. Average values of geometrical parameters and fluctuations about these averages were computed. The trajectories were displayed on an Evans and Sutherland PS300 color graphics system using HYDRA [36]. Analyses of optimized structures and dynamics trajectories were performed with the CORDAN program [44].

RESULTS AND DISCUSSION

A primary objective of the present model building and analysis was to find stable intercalation sites for NCS in the 5'GAGCG:5'CGCTC pentamer for which the carbons 2 or 6 of the diradical moiety of NCS are close to C1' of Cyt4 or C5' of Thy9. In addition, we wished to examine alternative sites for NCS and to determine the proximity of its biradical moiety to other sugar carbons of the pentamer. Only complexes docked to the minor groove were considered since the hydrogens of C1' and C4' are not accessible from the major groove; there is indirect experimental evidence of minor groove binding [6,7]. NCS can be docked to the minor groove with various orientations of the naphthoate moiety relative to the long axis of the base pairs. The

angle between the long axis of the naphthalene ring and the long axis of the base in the intercalation site is between 25° and 45°; other angles require considerable distortion of the DNA. The unsaturated part is aligned along the minor groove with ψ_3 in the neighborhood of 110°; i.e., the unsaturated part is approximately perpendicular to the naphthoate plane. Unwinding angles in the range 21 to 26° were obtained and a vertical lengthening in the range 2.6 to 3.05 Å was found. Experimental estimates correspond to 21° for the unwinding angle and 3.3 Å for the vertical lengthening [6]. The vertical lengthening is calculated as the projection on the helix axis of the mean distance between the C1' of the sugar moiety of the intercalation site bases; it is somewhat smaller than the experimental value due to a slight compression of the DNA model [38]. It is possible to find intercalation sites of reasonable energy between all base pairs with the unsaturated moiety making satisfactory contacts in the minor groove for both the (R,R) and (S,S) stereoisomers. The interaction energy is due to hydrogen bonds and electrostatic and van der Waals interactions. In addition to the contribution from the intercalated group on the order of -45 kcal van der Waals and -10 kcal electrostatic, there is also a van der Waals interaction between the unsaturated core of NCS and the minor groove atoms (on the order of -20 kcal/mole) and there are electrostatic interactions between the sugar of NCS and the backbone (on the order of -18 kcal/mole). Structures intercalated at Gua1-Cyt10/Ade2-Thy9 and Ade2-Thy9/Gua3-Cyt8 have the radical moiety pointed toward the 3' end of the (+) strand, while those at Gua3-Cyt8/Cyt4-Gua7 and Cyt4-Gua7/Gua5-Cyt6 have it oriented toward the 5' end of the (+) strand. These orientations give the strongest interaction between NCS and DNA and position the radical sites for attack on any sugar carbon.

A series of structures are listed in Table 1. From the Table it is clear that all the complexes have reasonable ΔE values. The variation of ΔE among the complexes of 25 kcal/mole is not sufficient to use it to say that one structure is more stable than another in solution. Although ψ_1 is near zero degrees in most cases (there are deviations of up to about 30°), ψ_2 tends to have a greater distortion from planarity; data for related structures indicate the range of ψ_1 values does not lead to a large energy penalty [45,46]. Of the various binding sites, only that involving

Figure 1. Neocarzinostatin (NCS) chromophore: a) chromophore itself with variable torsion angles indicated (ψ_1 = CN2-CN1-CN11-O11; ψ_2 = CN1-CN11-O11-C11; ψ_3 = CN11-O11-C11-C10, where CN1, CN2 are naphthoate atoms (C1 and C2) while C11 and O11 are the carbonyl and ester oxygen atoms); b) post-activated form with hydrogens at the radical sites C2 and C6 of the hydroindacene ring. Based on the structure proposed by Myers [12] from NMR and MS data of Hensens et al., [3].

the Ade2-Thy9/Gua3-Cyt8 base step (model-2a) yielded a low energy structure for (R,R) with an orientation that could permit attack at both Thy9 and Cyt4; the (S,S) stereoisomer at the same site can also cause bistranded lesions. The site at the Cyt4-Gua7/Gua5-Cyt6 base step with NCS in the (S,S) configuration pointing toward the 5' end of the (+) strand can also lead to bistranded lesions (model-4b). The distance between C5' of Thy9 and C6 of NCS is 3.11 Å, whereas the distance between C2 of NCS and C1' of Cyt4 is 3.34 Å. The (R,R) molecule in the same intercalation site appears incapable of causing cleavage at Thy9 because the distance between the C2 atom of NCS and the C5' of Thy9 is rather long (5.4 Å). Other intercalation sites lead to a structure which can give rise to single strand break at Thy9 or at other bases but not at both Thy9 and Cyt4 (see Table 1 and below).

Fig. 2A shows the (R,R) structure intercalated at the Ade2-Thy9/Gua3-Cyt8 base step (model-2a) obtained from energy minimization of the model. Since the free radical centers

Figure 2. Stereodrawings of optimized structures of complexes 5'GAGCG:5'CGCTC + post-activated form of NCS which mimic simultaneous hydrogen atom abstraction from the C1' of Cyt4 and the C5' of Thy9, model-2a (A), and model-2a with glutathione attached to the C12 of NCS (B); the C2 and C6 of NCS are labeled with a dot.

in the 'activated' form of NCS have a non-symmetric position (see Fig. 1b), (R,R) and (S,S) stereoisomers are non-equivalent in terms of formation of bistranded lesions. The distance between the C2 of NCS and C1' of Cyt4 is 4.43 Å and that between C6 of NCS and C5' of Thy9 is 3.31 Å in model-2a (R,R). It is of interest that the distance between C2 of NCS and C4' of Cyt4 is 3.30 Å; other antibiotics such as bleomycin are known to abstract at C4' [47]. In model-2b (S,S) the distance between C2 of NCS and C5' of Thy9 is 4.41 Å, that between C6 of NCS and C1' of Cyt4 is 4.69 Å, and that between C6 of NCS and C4' of Cyt4 is 4.68 Å. Model-2a permits methyl thioglycolate or glutathione addition at the C12 position of NCS without steric hindrance to the NCS-DNA complexes (see Fig. 2B). The model is compatible with NMR data showing that thiol attacks the C12 of NCS in trans configuration to the naphthoate group [48]. In the structure with glutathione attached to C12 (model-2a with glutathione—Fig. 2B) the distance between the C2 and the C1' of Cyt4 is shorter by 0.34 Å due to the fact that glutathione by interacting with the backbone of the (+)strand forces the NCS hydroindacene ring deeper into the minor groove; at the same time the C4' of Cyt4 is 0.70 Å further away from the C2 of NCS, whereas the distance between the C1' of Cyt4 and C2 decreased to 4.10 Å. This result is in agreement with experiments showing that abasic site formation at Cyt of AGC is significantly better with glutathione than with other simpler thiols, such as 2-mercaptoethanol or dithiothreitol [18,20].

There are a number of intercalation sites which can give rise to hydrogen atom abstraction from the C5' of Thy9 as a single lesion (Table 1). A plausible structure is the complex where NCS in the (R,R) form is intercalated at the Gua3-Cyt8/Cyt4-Gua7 base step (model-3a); the distances between the carbons of

Figure 3. Fluctuations of the distances between the C1', C4' and C5' atoms of Cyt4 and the C2 atom of NCS (A–C) and between the C1', C4' and C5' of Thy9 and the C6 of NCS (D–F), model-2a. There are 1200 structures presented in the figure which were generated during 30 ps dynamics simulation. The average distances between these atoms are:

- (A) C1' (Cyt4)–C2 4.43 Å (f 0.24)
- (B) C4' (Cyt4)–C2 3.30 Å (f 0.16)
- (C) C5' (Cyt4)–C2 4.14 Å (f 0.23)
- (D) C1' (Thy9)–C6 5.24 Å (f 0.20)
- (E) C4' (Thy9)–C6 3.45 Å (f 0.22)
- (F) C5' (Thy9)–C6 3.30 Å (f 0.19)

deoxyribose of Cyt4 and the free radical centers on NCS are greater than 9.0 Å. The sugar moiety of NCS interacts with the sugar moiety of Cyt4. The distance between C2 of NCS and C5' of Thy9 is 4.54 Å due to van der Waals repulsion between the 2-NH₂ of Gua3 and the reactive core of NCS. The C6 of NCS is closest to C1' of Gua3 (5.97 Å) and to C4' of Gua3 (6.08 Å). If the active core is rotated about Ψ_3 torsion until it is almost perpendicular to the helix axis, then the distance between the C6 and C5' of Thy9 is 3.47 Å. The distance between the C2 and C5' is 6.51 Å and that of C6 and C4' is 4.73 Å. In this model the radical center at C6 would be expected to abstract the

hydrogen atom from C5', whereas in the preceding one the abstracting site would be at C2. Another molecular model for single strand lesions has been proposed by Chen et al., [49] in which the NCS bicyclic core is perpendicular to the minor groove of DNA. This type of NCS positioning was found to be energetically and sterically less favorable in comparison to the above models in which NCS bicyclic core lies in the minor groove. Only (S,S) NCS can cleave Thy9 if intercalated at the Cyt4-Gua7/Gua5-Cyt6 base step (model-4b).

Since model-2a is of primary interest, because it involves (R,R) NCS and is in position to form bistranded lesions, the optimized

Table 2. Energy and structure of NCS complexed with mutants of 5'GAGCG:5'CTCGC^a

Sequence	E	ΔE	ψ_1	ψ_2	ψ_3
Model-2c	-260.9	-74.7	34.5	176.5	107.3
Model-2d	-264.7	-78.7	34.5	177.5	108.6
Model-2e	-316.1	-78.2	33.5	178.8	104.4

^a See Table 1 for definitions. All models considered use the (R,R) form of NCS in the Ade2-Thy9/Gua3-Cyt8 base step. Mutants are described in the text.

structure was subjected to a 30 ps dynamics simulation. During this period the structure was stable and all of the geometrical determinants of the B-DNA helix remained within normal values; e.g., the distance between the C1' atoms across the helix varied from 10.14 to 10.37 Å, while the global helical twist (Ω) and rise (D_z) showed small fluctuations about the optimized values (e.g., for Ade2-Gua3 the values are $28.5 \pm 2.4^\circ$ and 6.37 ± 0.14 Å, respectively). The sugar puckers are typical of a B-DNA structure; i.e., Ade2, Gua3, Cyt4 and Cyt8 are C3'-exo, whereas Gua7 and Thy9 are C2'-endo [50]. The sugar moiety of NCS forms hydrogen bonds to the phosphate backbone of Cyt10. For the abstraction reaction, the fluctuations of distances between the free radical centers of the post-activated form of NCS (C6 and C2) and the C5' of Thy9 and C1' of Cyt4 are of importance. The trajectory results are shown in Fig. 3. The distances C5-C6 and C1-C2 fluctuate between 3.0 and 4.5 Å; the average distances are 3.30 Å and 4.43 Å, respectively. After 1000 steps of ABNR minimization on the average dynamics structure, the distances between C5'-C6 is 3.37 Å and C1'-C2 is 4.60 Å. There is also a short distance (3.24 Å) between the C2 of NCS and the C4' of Cyt4; thus, it is theoretically possible to have hydrogen abstraction from the C4' of Cyt4.

Other complexes were examined by molecular dynamics simulations and the general behavior was similar to that of model-2a. For model-3a, which is positioned to permit hydrogen atom abstraction from the C5' of Thy9, the distance between C2 of NCS and C5' of Thy9 varies between 4.20 and 5.45 Å with an average value of 5.27 Å.

It has been observed that the alkali-induced cleavage at Cyt4 decreases in the sequence 5'GAICG:5'CGCTC (model-2c) relative to 5'GAGCG:5'CGCTC, whereas it increases if the Gua opposite Cyt4 is replaced by inosine; i.e., in the sequence 5'GAGCG:5'CICTC [20], (model-2d). The interaction energies for each of these sequences with NCS intercalated at the same site as in model-2a are given in Table 2; the structures are very similar to model-2a and both models (2c and 2d) have ΔE values similar to that of model-2a. Also, if the Ade-Thy base pair is mutated into a Gua-Cyt base pair (model-2e), the interaction energy is comparable to that for model-2a. Thus, the differences in cleavage of the mutant DNA cannot be due to the small differences in the interaction energy, but some geometrical variations must exist in these oligonucleotides in aqueous solution.

CONCLUSIONS

A number of intercalation sites for NCS in the B-DNA pentamer 5'GAGCG:5'CGCTC have been investigated by model building, energy minimization, and molecular dynamics. Given certain experimental data (in this case, that bistranded lesions can occur at C5' of Thy9 and C1' of Cyt4), it was relatively easy to find a low energy structure (i.e., with little DNA or NCS distortion and a sizeable interaction energy for the complex) that brings the radical centers of NCS into position to attack sugar carbons

at the known damage sites (intercalation at the Ade2-Thy9/Gua3-Cyt8 base step). Analogous results have been obtained by simple model building (without examination of the energetics) for other DNA cleaving drugs [51,52]. At the same time we have found intercalation sites, with corresponding structural and energetic properties, that can give rise to only single strand cleavage at Thy9 (see Table 1) or at other bases, i.e., Ade2 at the Gua1-Cyt10/Ade2-Thy9 base step and Gua3 at the Gua3-Cyt8/Cyt4-Gua7 base step.

The difficulty with such results is that their predictive capability is rather limited. Although sterically or energetically disallowed structures can be distinguished by the present methods, it is not possible to extrapolate such results to differences in the free energies of binding [39]. Even if one could rank the complexes in terms of their association constants, a much more detailed analysis of the hydrogen abstraction reaction would be required to estimate the relative reaction rates at the various sites. Thus, the best that one can hope to achieve by the calculations of the type reported here is to demonstrate that certain cleavage sites can be reached by the radical centers when NCS intercalates between certain bases and, of equal interest, that there appear to be no intercalation sites with satisfactory structures for a complex that permits NCS to cleave at certain bases; in the present study all cleavage sites can be reached from some intercalation site. It may be significant that in model-2a the distance between C2 and C4' (3.30 Å) is less than that between C2 and C1' (4.43 Å), suggesting C4' hydrogen abstraction as a possible reaction at Cyt4, raising the possibility that the alkali-labile C4' hydroxylation product, such as is formed by bleomycin [47], may also be formed by NCS.

The fact that direct breaks at the Thy, which is base-paired with the Ade of AGC, exceed the alkali-labile lesions at the abasic site (Cyt of AGC) [20] can be explained in two ways: 1) A single binding mode (model-2a) produces both types of lesions, but abasic site formation is less efficient, possibly for geometric reasons or possibly because the carbon-centered radical at C1' is inherently less readily oxidized to form the irreversible lesion or is more readily repaired by thiol scavengers; 2) The single strand lesion (direct break) and the bistranded lesions are due to two different kinds of drug binding; e.g., model-3a and model-2a, where NCS intercalates between Gua3-Cyt8/Cyt4-Gua7 and between Ade2-Thy9/Gua3-Cyt8 base steps, respectively. To distinguish these two possibilities concerning the origin of the bistranded lesions in the AGC sequence, separately deuteriated C5' of Thy9 and C1' of Cyt4 can be used, i.e., in model-2a one of the hydrogen atoms on the C5' of Thy9 should be abstracted by the C6, whereas the H1' on Cyt4 should be abstracted by the C2 of NCS. By contrast, in model-3a only abstraction from the C5' of Thy9 by the C2 should be observed.

Another antitumor antibiotic, calicheamicin, causes sequence-dependent double-strand breaks which are separated by three bases [52,53]. This antibiotic is a non-intercalative agent, and

its activated form has symmetrically centered diradicals on the benzene core (2.88 Å apart). Hence, these free radicals abstract hydrogens which are separated by a shorter distance. For standard B-DNA [33], a shorter distance occurs between C5' and C1' on different strands separated by two or three base pairs (on the order of 8.340 and 8.347 Å, respectively); those separated by one base pair are generally farther apart (11.06 Å). The activated form of NCS has the radical centers separated by a longer distance (3.58 Å). Such a geometry permits hydrogen abstraction from the C5' and C1' separated by a longer distance than those for calicheamicin. Since NCS intercalates into DNA and elongates it by nearly one full residue (see above), it abstracts hydrogen atoms from the C5' and C1' atoms of residue on opposite strands which are separated by two base pairs rather than three.

It is hoped that the present study will help in designing agents that cleave DNA by mechanisms similar to that of NCS and make workers in the field aware of the possibilities and limitations of model-building results.

ACKNOWLEDGEMENTS

This work was supported in part by a U.S. Public Health Service Research Grants CA 44257 and GM12573 (I.H.G) from the National Institutes of Health and an award from Merck and Company (A.G.). We thank Prof. Martin Karplus and B. Tidor for support and helpful discussions. We are indebted to CRAY Research, Mendota Heights, Minnesota, and to the John von Neumann Center in Princeton, New Jersey, for computer time on CRAY-S1 and Cyber 205, respectively.

REFERENCES

- Goldberg, I.H. (1987) *Free Radical Biol. Med.* **3**, 41–54.
- Napier, M.A., Holmquist, B., Strydom, D.J. and Goldberg, I.H. (1979) *Biochem. Biophys. Res. Commun.* **89**, 635–642.
- Hensens, O.D., Dewey, R.S., Liesch, J.M., Napier, M.A., Reamer, R.A., Smith, J.-L., Albers-Schonberg, G. and Goldberg, I.H. (1983) *Biochem. Biophys. Res. Commun.* **113**, 538–547.
- Edo, K., Mizugaki, M., Koide, Y., Seto, M., Furihata, K., Otake, N., and Ishida, M. (1985) *Tetrahedron Lett.* **26**, 331–334.
- Napier, M.A. and Goldberg, I.H. (1983) *Mol. Pharm.* **23**, 500–510.
- Povirk, L.F., Dattagupta, N., Warf, B.C. and Goldberg, I.H. (1981) *Biochemistry* **20**, 4007–4014.
- Dasgupta, D. and Goldberg, I.H. (1985) *Biochemistry* **24**, 6913–6920.
- Poon, R., Beerman, T.A. and Goldberg, I.H. (1977) *Biochemistry* **16**, 486–493.
- Hatayama, T. and Goldberg, I.H. (1979) *Biochim. Biophys. Acta* **563**, 59–71.
- Povirk, L.F. and Houlgrave, C.W. (1988) *Biochemistry* **27**, 3850–3857.
- Kappen, L.S. and Goldberg, I.H. (1985) *Nucleic Acid Res.* **13**, 1637–1648.
- Myers, A.G. (1987) *Tetrahedron Letters* **28**, 4493–4496.
- Takeshita, M., Kappen, L.S., Grollman, A.P., Eisenberg, M., and Goldberg, I.H. (1981) *Biochemistry* **20**, 7599–7606.
- Galat, A. and Goldberg, I.H. (1987) *Comp. Appl. Biosci.* **3**, 333–338.
- Bose, K.K., Tatsumi, K. and Strauss, B.S. (1980) *Biochemistry* **19**, 4761–4766.
- Kappen, L.S. and Goldberg, I.H. (1983) *Biochemistry* **22**, 4872–4878.
- Boye, E., Kohnlein, W. and Skarstad, K. (1984) *Nucleic Acid Res.* **12**, 8281–8291.
- Povirk, L.F. and Goldberg, I.H. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3182–3186.
- Povirk, L.F. and Goldberg, I.H. (1986) *Nucleic Acid Res.* **14**, 1417–1426.
- Kappen, L.S., Chen, C.Q. and Goldberg, I.H. (1988) *Biochemistry* **27**, 4331–4340.
- Kappen, L.S. and Goldberg, I.H. (1989) *Biochemistry*, **28**, 1027–1032.
- Povirk, L.F., Houlgrave, C.W. and Han, Y. (1988) *J. Biol. Chem.* **263**, 19263–19266.
- Chin, D.-H., Zeng, C., Costello, C.E. and Goldberg, I.H. (1988) *Biochemistry* **27**, 8106–8114.
- Myers, A.G., Proteau, P.J. and Handel, T.M. (1988) *J. Am. Chem. Soc.* **110**, 7212–7214.
- Rogers, D.W., Dagdagan, O.A., and Allinger, N.L. (1979) *J. Am. Chem. Soc.* **101**, 671–676.
- Nilsson, L., and Karplus, M. (1986) *J. Comput. Chem.* **7**, 591–616.
- Truong, T.N., Truhlar, D.G., Baldridge, K.K., Gordon, M.S. and Streckler, R. (1989) *J. Chem. Phys.* **90**, 7137–7142.
- Gupta, M.P. and Makata, A.P. (1976) *Cryst. Struct. Comm.* **5**, 553–556.
- Hehre, W.J., Radom, L., Schleyer, P. von R. and Pople, J. (1986) *Ab Initio under Molecular Orbital Theory*, J. Wiley and Sons, NY.
- Edo, K., Akiyama, Y., Koide, Y. and Ishida, N. (1986) *J. Antibiotics* **39**, 1615–1619.
- Frish, M.J., DeFrees, D.J., Raghavachari, R.A., Whiteside, R.A., Schlegel, H.B., Fluder, E.M., and Pople, J.A. (1982) *Gaussian 82: Department of Chemistry, Carnegie-Mellon University, Pittsburgh, PA.*
- Stewart, M.J. (1985) *MOPAC, QCPE Quantum Chemistry Program Exchange.*
- Arnott, S. and Hukins, D. W. L. (1972) *Biochem. Biophys. Res. Commun.* **47**, 1504–1507.
- Galat, A. (1987) *CORGEN, A Fortran-77 program for generation of standard and non-standard DNA helices.*
- Tsai, C. C., Jain, S. C. and Sobel, H. M. (1977) *J. Mol. Biol.* **114**, 301–315.
- HYDRA (Harvard/York Drawing Program), developed by R. Hubbard of York University.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983) *J. Comput. Chem.* **4**, 187–217.
- Tidor, B., Irikura, K.K., Brooks, B.R. and Karplus, M. (1983) *J. Biomol. Struct. Dyn.* **1**, 231–252.
- Pettitt, B.M., and Karplus, M. (1986) In Burgen, A.S.V., Roberts, G.C.K. and Tute, M.S. (eds.), *Molecular Graphics and Drug Design*. Elsevier Science Publishers, pp. 75–113.
- Platzer, K.E.B., Momany, F.A., and Scheraga, H.A. (1972) *Int. J. Peptide Protein Res.* **4**, 201–219.
- Wipff, G., Dearing, A., Werner, P.K., Blaney, J.M., and Kollman, P.A. (1983) *J. Am. Chem. Soc.* **105**, 997–1005.
- Verlet, L. (1967) *Phys. Rev.* **159**, 98–105.
- Ryckaert, J.P., Cicotti, G., and Berendsen, H.J.C. (1977) *J. Comput. Phys.* **23**, 327–337.
- Galat, A. (1987) *CORDAN, A Fortran-77 program for analysis of dynamics trajectories and optimized structures of DNA and DNA-drug complexes.*
- Trotter, J. (1961) *Acta Cryst.* **13**, 732–736.
- Tanaka, H., Kuroda, H., Marusawa, H., Hatanaka, H., Kino, T., Goto, T., Hashimoto, M. and Taga, T. (1987) *J. Am. Chem. Soc.* **109**, 5031–5033.
- Wu, J.C., Kozarich, J.W. and Stubbe, J. (1983) *J. Biol. Chem.* **258**, 4694–4697.
- Hensens, O.D. and Goldberg, I.H. (1989) *J. Antibiot.* **42**, 761–768.
- Chen, K.-X., Gresh, N. and Pullman, B. (1987) *Nucleic Acid Res.* **15**, 2175–2189.
- Dickerson, R.E. (1983) *J. Mol. Biol.* **166**, 419–442.
- Hawley, R.C., Kiessling, L.L., and Schreiber, S.L. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 1105–1109.
- Zein, N., Poncin, M., Nilakantan, R., and Ellestad, G.A. (1989) *Science* **244**, 697–699.
- Zein, N., Sinha, A.M., McGahren, W.J. and Ellestad, G.A. (1988) *Science* **240**, 1198–1201.