Relationship of Solution and Protein-Bound Structures of DNA Duplexes with the Major Intrastrand Cross-Link Lesions Formed on Cisplatin Binding to DNA

Luigi G. Marzilli,* Jamil S. Saad, Zsuzsanna Kuklenyik, Kelly A. Keating, and Yinghai Xu

Contribution from the Department of Chemistry, Emory University, Atlanta, Georgia 30322

Received March 6, 2000. Revised Manuscript Received November 25, 2000

Abstract: DNA bases in the three-base-pair (3bp) region of duplexes with the two major lesions of cisplatin (*cis*-$PtCl_2(NH_3)_2$) with DNA, namely d(XG*G*) and d(XA*G*) (* = N7-platinated base), differ in their relative positions by as much as ~3.5 Å in structures in the literature. Such large differences impede drug design and assessments of the effects of protein binding on DNA structure. One recent and several past structures based on NMR-restrained molecular dynamics (RMD) differ significantly from the reported X-ray structure of an HMG-bound XG*G* 16-mer DNA duplex (Ohndorf, U.-M.; Rould, M. A.; He, Q.; Pabo, C. O.; Lippard, S. J. *Nature* **1999**, 399, 708). This 16-mer structure has several significant novel and unique features (e.g., a bp step with large positive shift and slide). Hypothesizing that novel structural features in the XG*G* or XA*G* region of duplexes elude discovery by NMR methods (especially because of the flexible nature of the 3bp region), we studied an oligomer with only G•C bp's in the XG*G*Y site by NMR methods for the first time. This 9-mer gave a 5′-G* N1H signal with a normal shift and intensity and showed clear NOE cross-peaks to C NHb and NHe. We assigned for the first time ^{13}C NMR signals of a duplex with a G*G* lesion. These data, by adding NMR-based criteria to those inherent in NOESY and COSY data, have more specifically defined the structural features that should be present in an acceptable model. In particular, our data indicated that the sugar of the X residue has an N pucker and that the G*G* cross-link should have a structure similar to the original X-ray structure of *cis*-$Pt(NH_3)_2$(d(pGpG)) (Sherman S. E.; Gibson, D.; Wang, A. H.-J.; Lippard, S. J. *J. Am. Chem. Soc.* **1988**, *110*, 7368). With these restrictions added to NOE restraints, an acceptable model was obtained only when we started our modeling with the 16-mer structural features. The new X-ray/NMR-based model accounted for the NOESY data better than NOE-based models, was very similar in structure to the 16-mer, and differed from solely NOE-based models. We conclude that all XG*G* and XA*G* (X = C or T) duplexes undoubtedly have structures similar to those of the 16-mer and our model. Thus, protein binding does not change greatly the structure of the 3bp region. The structure of this region can now be used in understanding structure−activity relationships needed in the design of new carrier ligands for improving Pt anticancer drug activity.

Introduction

Synthetic models of the two major lesions of cisplatin (*cis*-$PtCl_2(NH_3)_2$) with DNA, namely d(XG*G*Y) and d(XA*G*Y) (* = residue with an N7-platinated base; X and Y = 5′ and 3′ flanking residues), have been studied intensely.[1−17] In a review of reported NMR spectra,[4] we concluded that all duplexes reported have the same structure in the three-base-pair (3bp) 5′-XG*G*-3′ region. Since that review and after this work was submitted, yet another NMR study was reported.[17] In this recent case, the Pt-bearing G*G* strand (the Pt strand) was more purine rich than in previous NMR studies. However, allowing for the effects on the NMR spectra from the greater anisotropy of the purines, we continue to believe that the NMR data support our

* To whom correspondence should be addressed. E-mail: lmarzil@emory.edu.

(1) Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, *99*, 2467−2498.
(2) Ohndorf, U.-M.; Rould, M. A.; He, Q.; Pabo, C. O.; Lippard, S. J. *Nature* **1999**, *399*, 708−712.
(3) Yang, D.; van Boom, S. S. G. E.; Reedijk, J.; van Boom, J. H.; Wang, A. H.-J. *Biochemistry* **1995**, *34*, 12912−12920.
(4) Ano, S. O.; Kuklenyik, Z.; Marzilli, L. G. In *Cisplatin: Chemistry and Biochemistry of a Leading Anticancer Drug*; Lippert, B., Ed.; Wiley-VCH: Weinheim, 1999; pp 247−291.
(5) Coll, M.; Sherman, S. E.; Gibson, D.; Lippard, S. J.; Wang, A. H.-J. *J. Biomol. Struct. Dyn.* **1990**, *8*, 315−330.
(6) den Hartog, J. H. J.; Altona, C.; van Boom, J. H.; van der Marel, G. A.; Haasnoot, C. A. G.; Reedijk, J. *J. Am. Chem. Soc.* **1984**, *106*, 1528−1530.
(7) den Hartog, J. H. J.; Altona, C.; van Boom, J. H.; van der Marel, G. A.; Haasnoot, C. A. G.; Reedijk, J. *J. Biomol. Struct. Dyn.* **1985**, *2*, 1137−1155.
(8) Dunham, S. U.; Dunham, S. U.; Turner, C. J.; Lippard, S. J. *J. Am. Chem. Soc.* **1998**, *120*, 5395−5406.
(9) Fouchet, M.-H.; Guittet, E.; Conget, J. A. H.; Kozelka, J.; Gauthier, C.; Bret, M. L.; Zimmermann, K.; Chottard, J.-C. *J. Biol. Inorg. Chem.* **1997**, *2*, 83−92.
(10) Gelasco, A.; Lippard, S. J. *Biochemistry* **1998**, *37*, 9230−9239.
(11) Herman, F.; Kozelka, J.; Stoven, V.; Guittet, E.; Girault, J.-P.; Huynh-Dinh, T.; Igolen, J.; Lallemand, J.-Y.; Chottard, J.-C. *Eur. J. Biochem.* **1990**, *194*, 119−133.
(12) Mukundan, S., Jr.; Xu, Y.; Zon, G.; Marzilli, L. G. *J. Am. Chem. Soc.* **1991**, *113*, 3021−3027.
(13) Sherman, S. E.; Gibson, D.; Wang, A. H.-J.; Lippard, S. J. *Science* **1985**, *230*, 412−417.
(14) Sherman, S. E.; Gibson, D.; Wang, A. H.-J.; Lippard, S. J. *J. Am. Chem. Soc.* **1988**, *110*, 7368−7381.
(15) Takahara, P. M.; Rosenzweig, A. C.; Frederick, C. A.; Lippard, S. J. *Nature (London)* **1995**, *377*, 649−652.
(16) Takahara, P. M.; Frederick, C. A.; Lippard, S. J. *J. Am. Chem. Soc.* **1996**, *118*, 12309−12321.
(17) Parkinson, J. A.; Chen, Y.; Murdoch, P. d. S.; Guo, Z.; Berners-Price, S. J.; Brown, T.; Sadler, P. J. *Chem.-Eur. J.* **2000**, *6*, 3636−3644.

10.1021/ja0007915 CCC: $20.00
Published on Web 03/01/2001

Table 1. Abbreviations and Sequences of Some Duplexes with a 1,2-d(GpG) Intrastrand Cross-Link

abbrev	sequence	ref
9-mer	d(CTCCG*G*CCT)·d(AGGCCGGAG)	this study
PtL_{HMG}[a]	d(CCTCTCTG*G*ACCTTCC)· d(GGAAGGTCCAGAGAGG)	2
PtW	d(CCTG*G*TCC)·d(GGACCAGG)	3
PtL_{NMR}	d(CCTCTG*G*TCTCC)·d(GGAGACCAGAGG)	10
PtS	d(ATACATG*G*TACATA)·d(TATGTACCATGTAT)	17

[a] The "16-mer" studied by X-ray crystallography. All other duplexes were studied by traditional NOE-based restraint refinement. The 9-mer was also studied by the X-ray/NOE-based method described here.

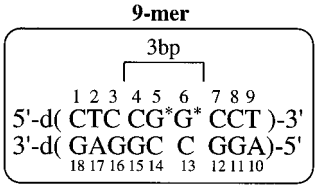
contention that all duplexes with assigned NMR signals (Table 1 and Supporting Information) must have similar structures.[3,4,7−11,17]

Our conclusion was reached on the basis of the striking similarity in a large number of unusual NMR observations reported for all these duplexes (namely, NOE cross-peak and imino proton exchange patterns, ^{1}H and ^{31}P NMR chemical shifts). These characteristic atypical NMR features are concentrated in the 3bp region. As will become evident, the relationship between the 5′-flanking base pair (X·X′ bp) and the 5′-G*·C bp (the XG* base pair step) provides the key to explaining the NMR data. The most intriguing NMR observations common to the Pt strand of all duplexes are the following: the H2′ signal of the 5′-flanking C/T residue shows a large relative upfield shift; the 5′-G* H1 exchanges with water much more rapidly than the 3′-G* H1; and, when reported,[4] the ^{31}P NMR signal from the phosphate group between the two G*'s is substantially downfield. In the non-Pt-strand, the X′ and C residues at the key XG* bp step have one characteristic and unusually weak NOE cross-peak (C H2′−X′ H8). At the G*Y bp step, the NMR data for the 3′-G*·C and the Y·Y′ bp's are characteristic of B-form DNA.

Explaining these NMR findings with a structural model has proved to be extremely difficult. Despite the similarities in the experimental data, the reported NMR-restrained molecular dynamics (RMD) structural models[3,10,17] agree only in some gross features. The molecular models obtained by various procedures differ in base stacking, backbone conformation, base step rise, the extent and nature of distortions of the $PtG*_2$ geometry, and even the number of conformers present.[3,8−11,17] We describe the $PtG*_2$ geometry using the average of the displacements of Pt out of each G*'s plane (Pt/G*).

Of particular note, the relative positions of bases in the 3bp region differ by as much as 3.5 Å in reported solution models (Supporting Information)[2,3,10,17] from those in an X-ray structure of a 16-mer bound to domain A of rat HMG1.[2] Such large differences impede drug design,[18] and the solution structures have been too uncertain to use for assessing the effects of protein binding on DNA structure.[1,2,19,20] We believe that the similarities in the ^{1}H NMR data indicate that the geometry at Pt should not differ from adduct to adduct and should be normal. We note that ^{195}Pt NMR shifts are very sensitive to geometry; ^{195}Pt NMR shifts for conformers of bis-guanine derivatives differing only by rotation about the Pt−N7 bond can vary by as much as 13 ppm.[21] The ^{195}Pt NMR shifts for relevant adducts with *cis*-$Pt(NH_3)_2$ bound to two G N7 donors are all found to be within

Chart 1



10 ppm for duplex DNA, for single-stranded oligonucleotides, and for nucleosides and nucleotides.[22−24] We believe that this similarity of ^{195}Pt NMR shifts is inconsistent with the high (∼0.7 Å) Pt/G* in some RMD NMR models compared to the ∼0.4 Å Pt/G* for small models.[5,14,25] A recent study reports a moderate (0.55 Å) Pt/G*, but the N7−Pt−N7 angle is distorted (69° vs 90°).[17]

These and other considerations convinced us that NMR methods could not define the structure well, even if the techniques are carefully applied and even if NMR methods are the best approach toward gaining solution structural information.[4] We hypothesized[4] that NMR methods face *two* difficulties in defining the solution structure. *First*, as-yet unrecognized novel structural features in the 3bp region elude discovery by NMR methods. (Indeed, a reviewer of this work suggested we mention that the NMR/modeling method is not sufficient when unusual structural features are present. Also, this reviewer suggested that the novel findings to be described below have "important general consequences" for "restraint sets" used in the NMR/modeling software program.) *Second*, the distorted and flexible nature of the 3bp region worsens the normally restricted applicability of RMD methods to DNA.

To address the first difficulty, we explored incorporating into our modeling several significant novel and unique features (e.g., a bp step with large positive shift and slide) of the recently reported 16-mer X-ray structure.[2] The *cis*-$Pt(NH_3)_2$(d(GpG)) moiety in this 16-mer has a small Pt/G* (∼0.35 Å),[2] a feature consistent with the ^{195}Pt NMR data. In addition, we obtained new types of NMR data, hoping that this combined approach would decrease the structural uncertainty arising from limitations of the NMR approach. To address the second difficulty, we chose a sequence with the G*G* lesion flanked by two G·C bp's in both the 5′ and 3′ flanking regions, giving a run of six G·C bp's (Chart 1).

Experimental Section

Sample Preparation and NMR Spectroscopy. Details can be found in the Supporting Information. The solution of the 9-mer in 99.998% D_2O at pH 7.0 (uncorrected) was ∼3 mM. For NMR experiments in H_2O, the solution was lyophilized and the sample dissolved in 90% H_2O/10% D_2O, and the pH was adjusted to 7.0. Most NMR experiments were run on GE GN-500 or Ω-600 spectrometers. 2D NOESY spectra in H_2O were recorded using a Varian Unity+600 spectrometer. All NMR experiments were performed at 5 °C except as indicated.

Molecular Modeling. (Further details can be found in the Supporting Information.) All modeling calculations were performed with InsightII/Discover (Biosym/MSI). Calculations were carried out in vacuo using a modified version of the AMBER force field.[26] We typically used an EM/MD/EM cycle (EM = energy minimization, MD = molecular

(18) Hambley, T. W. *Coord. Chem. Rev.* **1997**, *166*, 181−223.
(19) Dunham, S. U.; Lippard, S. J. *Biochemistry* **1997**, *36*, 11428−11436.
(20) Cohen, S. M.; Mikata, Y.; He, Q.; Lippard, S. J. *Biochemistry* **2000**, *39*, 11771−11776.
(21) Sullivan, S. T.; Ciccarese, A.; Fanizzi, F. P.; Marzilli, L. G. *Inorg. Chem.* **2001**, *40*, 455−462.
(22) Bancroft, D. P.; Lepre, C. A.; Lippard, S. J. *J. Am. Chem. Soc.* **1990**, *112*, 6860−6871.
(23) Miller, S. K.; Marzilli, L. G. *Inorg. Chem.* **1985**, *24*, 2421−2425.
(24) Kline, T. P.; Marzilli, L. G.; Live, D.; Zon, G. *J. Am. Chem. Soc.* **1989**, *111*, 7057−7068.
(25) Admiraal, G.; van der Veer, J. L.; de Graff, R. A.; den Hartog, J. H. J.; Reedijk, J. *J. Am. Chem. Soc.* **1987**, *109*, 592−594.
(26) Yao, S.; Plastaras, J. P.; Marzilli, L. G. *Inorg. Chem.* **1994**, *33*, 6061−6077.

Table 2. ^{1}H and ^{31}P NMR Chemical Shifts (ppm) of the 9-mer at 5 °C

	H8, H6	H2, H5, CH_3	H1′	H2′	H2″	H3′	H4′	H1, H3	H4b	H4e	3′−P
C1	7.85	5.80	5.76	2.27	2.56	4.61	4.07		7.93	7.01	−4.46
T2	7.68	1.63	6.15	2.32	2.60	4.89	4.27	14.00			−4.29[a]
C3	7.61	5.67	5.88	2.25	2.44	4.79	4.16		8.50	7.03	−3.91[a]
C4	7.45	5.72	5.75	1.33	2.41	4.62	4.04		8.85	7.27	−4.21
G*5	8.66		6.18	2.38	2.68	5.14	4.19	13.54			−3.08
G*6	8.39		5.65	2.36	2.45	4.73	4.22	13.43			−4.28
C7	7.58	5.42	6.08	2.23	2.51	4.81	4.20		8.39	6.84	−4.33
C8	7.66	5.72	6.06	2.25	2.48	4.80	4.16		8.58	7.19	−4.20
T9	7.56	1.69	6.25	2.27	2.27	4.55	4.04	14.03			
A10	8.01	7.81	5.97	2.48	2.66	4.84	4.19				−4.19
G11	7.85		5.58	2.72	2.75	5.00	4.37	13.08			−3.83[a]
G12	7.74		6.04	2.62	2.74	4.90	4.38	13.19			−4.10[a]
C13	7.50	5.54	5.96	2.04	2.42	4.75	4.09		8.61	6.81	−4.08
C14	7.45	5.55	5.59	1.85	2.24	4.77	4.02		8.25	6.84	−4.23[a]
G15	7.86		5.61	2.66	2.74	4.95	4.28	12.93			−4.19[a]
G16	7.77		5.54	2.62	2.73	5.01	4.34	12.89			−4.27[a]
A17	8.01	7.95	6.16	2.63	2.91	5.02	4.43				*b*
G18	7.62		5.94	2.22	2.38	4.62	4.15	13.27			

[a] Interchangeable assignments: T2pC3 for G12pC13 and G11pG12 for G15pG16/G16pA17. [b] Not assigned.

dynamics). EM included 5000 steps of steepest descents and conjugate gradient EM, or enough steps to reach a 0.1 kcal/mol maximum derivative. After a heating period of 30 ps, 200 ps MD simulation was performed at 300 K; trajectories were saved every 1 ps. The saved trajectory archive was subjected to an EM cycle as above.

We performed numerous parallel EM/MD/EM calculations by following both common modeling approaches and also unusual variations of these approaches. Our standard modeling used NOE restraints (100 and 300 ms data) and WC hydrogen-bonding restraints (based on 2D phase-sensitive 90% H_2O NOESY spectra[27]). On many occasions, we attempted to introduce other constraints to obtain models that explained the shift information in addition to the NOE data. For example, we employed restraints to move the H2′ of the 5′-flanking sugar into the shielding cone of the 5′-G* or to restrict the Pt/G* geometry to conform to the X-ray structure of *cis*-Pt$(NH_3)_2$(d(pGpG)).[14] We also used both right- and left-handed canted starting *cis*-Pt$(NH_3)_2$-(d(GpG)) models. Regardless of the approach, the many resulting 9-mer NOE-based models were similar in structure to published models. However, these models, like other NOE-based models in the literature, had unsatisfactory features (e.g., if the sugar puckers were correct, the models failed to explain the normal ^{195}Pt NMR shift found for DNA adducts).[22,23] This effort, covering many years, is much too extensive to describe; therefore, we present results in the Supporting Information for only four related and representative models of the many model types we evaluated. Our preferred model, J, was constructed from the PtL_{HMG} 16-mer structure by exchanging only the base residues to conform to the 9-mer sequence. No artificial restraints were employed to restrict the Pt geometry or to position the H2′ in the 5′-G* shielding cone; instead, only the initial NOE and WC hydrogen bond restraints were used. To relax the duplex, the structure was subjected to 500 iterations of steepest descents EM without any restraints; the structure was then subjected to 5000 steps of steepest descents EM with restraints.

Results

NMR assignments and shifts for the 9-mer are shown in Table 2. The unusual upfield shift of the 9-mer C4 H2′ signal (1.33 ppm, ∼0.6 ppm more upfield vs ∼2.0 ppm generally found in 5′-(CG)-3′ and other pyrimidine−purine steps of B-DNA) indicates that the XG* bp step has a structural feature unique to XG*G* and XA*G* duplexes. The 5′-G*•C bp of the 9-mer has other ^{1}H NMR shifts characteristic of all previously studied XG*G*Y duplexes (e.g., the relatively upfield shift of the C NHb signal and the universally similar shift of ∼8.7 ppm for the downfield-shifted 5′-G* H8 signal).[3,4,8,10,17] *The 9-mer is representative of all other relevant duplexes*.

The intensities of some interresidue NOE cross-peaks near the binding site 5′-(C4G*5G*6C7)-3′•5′-(G12C13C14G15)-3′ differed significantly from those for residues in the rest of the duplex, 5′-(C1T2C3)-3′•5′-(G16A17G18)-3′ and 5′-(C8T9)-3′•5′-(A10G11)-3′, where the relative intensity patterns were consistent with a B-DNA conformation. One of the most significant of such unusual NOE cross-peaks is the relatively strong G*5 H8−G*6 H8 cross-peak; the estimated distance is 3.0−3.2 Å (5 Å is normal for B-DNA), consistent with coordination of Pt to these residues in a head-to-head fashion. In the H1′−H6/H8 region, the C4 H1′−G*5 H8, G*5 H1′−G*6 H8, and C13 H1′−C14 H6 cross-peaks were weaker than the equivalent cross-peaks from the flanking B-DNA regions (Supporting Information), indicating distances larger than those in normal B-DNA. This weakness was especially obvious in the shorter mixing time spectra (100 ms, not shown), where these peaks are barely above the noise level. Such weak interresidue NOE cross-peaks present in spectra of all other duplexes lead to over 1.1 Å differences in interproton distances of NOE-based models. The interresidue C14 H2′−G15 H8 NOE cross-peak was also very weak, even at long mixing time, whereas a C14 H6−G15 H8 cross-peak was not observed. The other duplexes studied have these same patterns. *These NOE cross-peak patterns for the 9-mer further support our argument*[4] *that all duplexes have the same structure with a remarkable XG* base pair step.*

The relative intensity of the H2″−H3′ and H1′−H2′ DQF−COSY cross-peak reflects the relative size of the coupling between these sugar protons and can be used to provide information about sugar conformation. Sugars with a *C2′-endo* (*S*) conformation give fairly weak H2″−H3′ and strong H1′−H2′ DQF−COSY cross-peaks, while sugars with near *C3′-endo* (N) conformations give the opposite relative intensity pattern, strong H2″−H3′ and weak H1′−H2′ cross-peaks. For the 9-mer, the DQF COSY spectrum (Supporting Information) indicated relatively strong H2″−H3′ and weak H1′−H2′ couplings for C4, G*5, and C13, consistent with sugar puckering in the N conformation range. This pattern was also observed for T9 and G18, but this is typical for 3′-terminal residues. For all the other residues, the weak H2″−H3′ and strong H1′−H2′ couplings observed suggest that the sugars have an *S* conformation. Insofar as the COSY data on other duplexes can be assessed from literature reports, we judge the COSY patterns to be consistent with a similar structure for all duplexes, with the 9-mer being

(27) Piotto, M.; Saudek, V.; Sklenár, V. *J. Biomol. NMR* **1992**, *2*, 661−665.

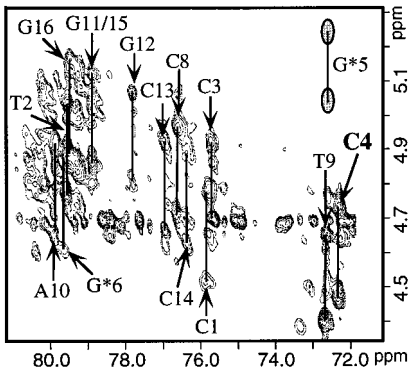


Figure 1. H3′−C3′ region of the coupled HMQC spectrum of the 9-mer at pH 7 and 30 °C. The gray doublet indicates the position of the G*5 H3′−C3′ cross-peak missing at 30 °C but found at 10 °C.

a representative duplex. However, since literature models of XG*G* and XA*G* duplexes differ in the puckers of some sugars, we turned to ^{13}C NMR spectroscopy as an additional means of defining structural features such as sugar pucker that should be present in an acceptable model.

In our previous studies on cisplatin DNA adducts containing G*pG* intrastrand cross-links,[12,28] the *C*3′ sugar carbon signal (an italic *C* indicates carbon instead of cytosine) of the 5′-G* residue was found to have a significant upfield shift that was related to the N sugar pucker for the 5′-G* sugar (determined by NOESY and DQF−COSY data). We also hoped that ^{13}C NMR shifts might allow us to assess the Pt geometry since a distorted geometry can give large differences in such shifts for G* *C*8.[28] For these reasons and in order to contribute to a very small body of reported ^{13}C chemical shifts for DNA with a bound metal species, we obtained ^{13}C NMR data for the 9-mer (Figure 1), assigned as described in the Supporting Information. The interpretation of the data is presented below.

Information regarding the relative rate of exchange of the imino protons with solvent can be obtained by comparing the relative intensities of their EXSY cross-peaks with water (Supporting Information). A larger EXSY cross-peak implies more accessibility of that proton to the solvent. In a 50 ms mixing time non-Watergate NOESY spectrum (which showed more intense imino-H_2O EXSY peaks than the Watergate NOESY spectrum), the order of decreasing intensity of the exchange peaks was G*5 > G18 > G*6 > T2 ∼ G11 > G15 > G12 > G16. The order of the exchange rate can be rationalized if fraying occurs near the ends of the duplex and the Pt binding site. G16 H1, which is the second bp from both the nearest terminal bp and the 5′-G*•C bp, has the most slowly exchanging NH.

The NOE cross-peaks from G*5 H1 to C14 H4b, G*6 H1, and G15 H1 were very weak, even in the short (25 ms) mixing time NOESY spectrum. At the same time, the G*5 H1 signal showed fast decay of its diagonal peak and build-up of its cross-peak with the H_2O signal, consistent with fast proton exchange with water. At 25 ms mixing time, the cross-peak of G*5 H1 to water was even slightly more intense, and the diagonal peak of G*5 H1 was even slightly weaker than those for the terminal residue, G18, indicating that the G*5 H1 exchanged even faster than the terminal G18 H1. In addition, as a function of temperature, the G*5 H1 of the terminal G18 H1 signals showed similar broadening (Figure 2), a result also indicative of fast proton exchange with water. Therefore, the weakness of the NOE cross-peaks from G*5 H1 is clearly due to the fast exchange of G*5 H1 rather than its long distances to other protons. As mentioned above, for all XG*G* duplexes, the 5′-G* N1H exchanges relatively readily. Thus, the 9-mer is representative of all such duplexes. We note that we observe a G* N1H signal with almost full intensity and with a normal shift for a fully base-paired G and that all NMR spectral indicators of structure are similar for all duplexes. These observations provide compelling evidence either that the dynamic process leading to exchange does not have an intermediate or, more likely, that any intermediate in the exchange pathway is present in relatively low abundance for all duplexes at low temperature.

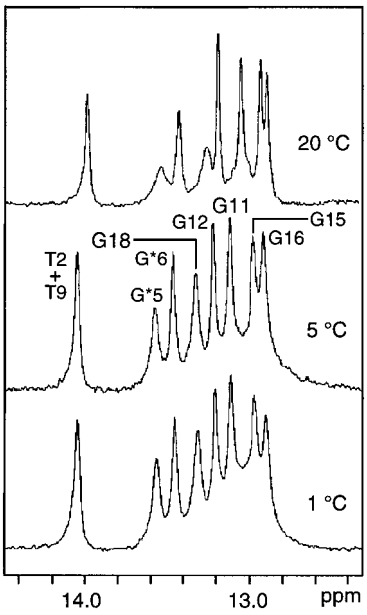


Figure 2. Imino ^{1}H NMR spectral region of the 9-mer in H_2O at various temperatures. Note the essentially normal intensity of the G* N1H signal at 5 °C and especially at 1 °C.

Table 3. Selected Parameters Comparing the 3bp Region of Various Duplexes

	XG* step		G*G* step			
abbrev	shift (Å)	slide (Å)	shift (Å)	slide (Å)	rise (Å)	Pt/G* (Å)
HMG-Bound Structure						
PtL_{HMG}	1.36	1.62	−0.44	−0.64	7.41	0.33
X-ray/RMD Solution Model						
9-mer (model J)	1.19	1.47	−0.19	−0.64	7.44	0.46
RMD Solution Models						
PtS	−0.25	−0.58	2.02	−2.52	4.56	0.55[a]
PtW	−0.74	0.58	0.75	−0.63	4.58	0.71
PtL_{NMR}	0.04	−0.23	−0.27	−0.48	5.35	0.77
9-mer (model D)	0.10	1.72	0.51	−0.93	4.77	0.65

[a] This value is small due to a distorted N7−Pt−N7 angle.

Discussion

We present in Supporting Information details of numerous structural parameters for five 9-mer models as well as for literature duplex structures. We focus here on structural parameters for the three published NOE-based models, for the X-ray/NMR-based 9-mer model "J", and for the 16-mer X-ray structure used to define the latter solution model. As can be seen from the few parameters listed in Table 3, some differences between structures are large, reaching ∼2 Å for the slides and shifts and ∼3 Å for the G*G* bp step rise. We discuss other differences below and note that many are given in the Support-

(28) Iwamoto, M.; Mukundan, S., Jr.; Marzilli, L. G. *J. Am. Chem. Soc.* **1994**, *116*, 6238−6244.

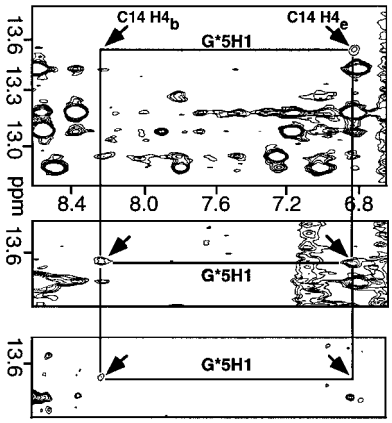


Figure 3. G N1H to C NH_2 region of the Watergate NOESY spectrum of the 9-mer in H_2O at pH 7 and 5 °C at 190 ms (top), 50 ms (middle), and 25 ms (bottom) mixing times.

ing Information, including important differences in sugar pucker. Furthermore, NOE-based models have a relatively large distortion of the Pt geometry (distortion of the N7−Pt−N7 bond angle or Pt/G* ≥ 0.6 Å). How do the 9-mer ^{13}C NMR signals assigned here for the first time for a duplex with an XG*G* lesion help define the required features?

The 9-mer G* *C*8 signals have shifts (Supporting Information) that are similar to each other and to those found for small molecules,[12] suggesting a small-molecule-like structure of the PtG^*_2 moiety.[28] Models universally agree on the pucker of some lesion sugars (5′-G* is N; 3′-G* is S) but disagree on several other sugars, including the 5′-flanking X residue (Supporting Information).[4] The upfield *C*3′ shift (Figure 1 and Supporting Information) and DQF−COSY data leave no doubt that the 9-mer X residue (C4) has an N sugar.[12] The pucker of the C sugar in the 3′-G*•C bp is also not clearly defined by published NMR studies, being assigned N or S in various NMR studies (Supporting Information).[4,11] For this residue (C13) in the 9-mer, the data suggest an unusual pucker since the ^1H NMR data favor an N pucker but the *C*3′ ^{13}C NMR shift is in the S pucker range.[12] In summary, these ^{13}C NMR results suggest that the geometry at the Pt should be nearly normal for a G*G* cross-link, that the X residue (C4) has an N sugar, and that the sugar of the C residue of the 3′-G*•C bp (C13) has an unusual pucker.

The 9-mer, designed with only G•C flanking bp's, allows observation of the 5′-G* N1H signal of normal shift and intensity (Figure 2) and clear NOE cross-peaks to C NHb and NHe (Figure 3). The latter two signals are well separated and have essentially normal shifts for a G•C bp. The properties indicate a well-formed 5′-G*•C bp (as found in the 16-mer);[2] however, the relatively small NOE cross-peak intensities indicate that the bp is very dynamic. The difficulty in observing the 5′-G* N1H signal (thereby creating an NOE "blind spot") is a major impediment in NMR studies; we hypothesized[4] that the absence of this signal when X and Y are part of A•T bp's[3,10] is due to extensive DNA "breathing", not to structural differences between duplexes with A•T and those with G•C flanking bp's. In a recently reported third study[17] of a TG*G*T duplex, the 5′-G* N1H signal, albeit weak, could indeed be observed. The shift and the weak NOE cross-peaks of this 5′-G* N1H signal to the NH_2 signals of the complementary C are very similar to the same properties of the corresponding signal of the 9-mer. These results, published after our work was submitted, confirm our conclusion[4] that sequence does not influence the overall structure but does influence dynamic properties. The results for the NH signals indicated that the 9-mer should have a well-formed 5′-G*•C bp.

Prior to the 16-mer report,[2] we had conducted extensive NOE-based RMD studies incorporating restraints guided by our new types of NMR data in addition to the NOE data. We also believed that the Pt geometry should be normal. However, all our efforts led to refined models with high (∼18%) *R*-factors (illustrated with the four NOE-based models, A to D, in Supporting Information); no model based solely on NMR experimental data accounted fully for the NMR data or improved significantly on apparently incorrect structural features in reported RMD NOE-based models. For example, the models could not explain the cross-peak intensity pattern of the G15 H8 signal to the C14 signals. This pattern, which is significantly different from that in B DNA, is found in all duplexes, including the duplex with the purine-rich Pt strand. For the 9-mer, this finding suggests that distances from G15 H8 to the C14 protons (except for H3′) are longer than in B-form DNA (Supporting Information). The distance patterns are H3′ ∼ H6 > H2′ > H2″ in B-form DNA and H6 > H2′ ∼ H3′ > H2″ in the G*G* duplexes. The 16-mer, in addition to having a less distorted Pt geometry, has the latter distance relationship.[2]

Since it was clear to us that the 16-mer had a structure[2] closer to the structure suggested by the NMR data than any NOE-based model from this or any other laboratory, we began a new round of modeling starting with the 16-mer features. The *R*-factor dropped dramatically (∼12%), and the resulting X-ray/NMR-based model, J, was consistent with all the NMR parameters and gave improvements in the fit of a number of cross-peaks (Supporting Information). In particular, the Pt/G*, the sugar puckers, the base pairing, and the C14−G15 distances of model J all account well for the NMR data. Another satisfying feature of model J is that the backbone torsion angles in the 3bp region are in the normal range, except for only one angle. In contrast, typical NOE-based models have from four to nine abnormal angles. In model J, the C13 residue has an unusual sugar pucker, consistent with the atypical NMR properties of this sugar. Also, the X H2′ shift is explained well only by the 16-mer and model J. Previous models failed to account for this shift or had an N pucker for one sugar[11] indicated by NOESY/COSY data to have an S pucker. Key features are retained in model J by starting with the 16-mer structure. These features, which are not present in the NOE-based models, are the large positive shift and slide of the XG* bp step and the large departure from the B-form position of the C complementary to this G* (C14 in the 9-mer). In model J and the 16-mer structures,[2] cross-link formation dramatically moves the center of the ring of this C base by 4.4 Å (Figure 4). In the NOE-based models, this C base position is more similar to those of B-form DNA than in the 16-mer or model J. In one model, PtS, the base position differs from that in B-form DNA by only 1.8 Å,[17] but it differs very significantly (3.5 Å) from that in model J or the 16-mer (Figure 4).

The sugar pucker of the Y′ residue in the 3′-flanking Y•Y′ bp is N in the 16-mer and S in the 9-mer and in all solution models. Also, the X residue pucker is more N in model J than in the 16-mer. Thus, the structures of the 16-mer and model J do differ. Nevertheless, model J is structurally very similar to the 16-mer (Figures 4 and 5, and Supporting Information). In fact, the geometry of the d(pG*pG*) moiety in model J (and the 16-mer) gives a very good fit to the very first X-ray structure of a cross-linked model, *cis*-$Pt(NH_3)_2$(d(pGpG)).[14] Thus, the 16-mer and the X-ray/NMR-based model J suggest that the distortion is concentrated mainly in the DNA, whereas NOE-based modeling suggests much less distortion in the DNA and somewhat more distortion in the Pt moiety.

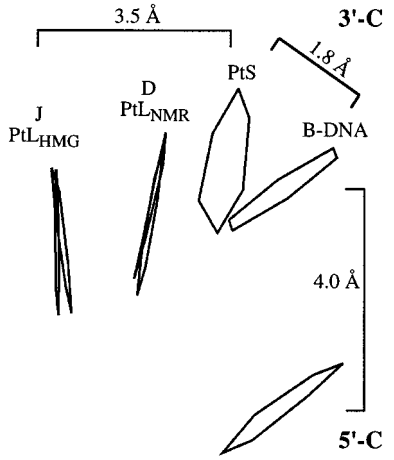


Figure 4. Base separation (between the center of the six-membered ring of the C bases in the G*•C bp's of the 9-mer models J and D (C13 and C14) compared to the separation in PtL$_{HMG}$ (16-mer),[2] PtL$_{NMR}$,[10] PtS,[17] and B-DNA (cf. Table 1 for complete sequences). (The figure was constructed by superimposing the 5′-C bases (in the 3′-G*•C bp).)

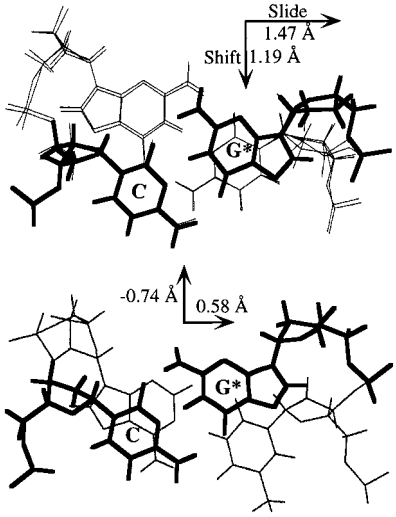


Figure 5. Structures of XG* regions: PtW (bottom) and model J superimposed on the starting 9-mer structure based on the PtL$_{HMG}$ X-ray structure (top). Bold lines indicate the G*•C bp of the J (top) and PtW (bottom) models. This 9-mer starting structure is identical to the X-ray structure with the A and T nucleobases exchanged with G and C nucleobases. The figure illustrates that restrained refinement of the starting 9-mer does not cause large changes in structure in this XG* region.

The most important result of our work is that the duplex in solution has a structure very similar to that of the 16-mer bound to HMG.[2] To gain acceptance of the correctness of this X-ray/NMR-based result, we must expose the reasons that NOE-based methodology led to a different conclusion than the X-ray/NMR method employed here. The RMD method depends primarily on interproton distances ("contacts") based on the observed NOE cross-peak intensities. In general, short contacts position the residues. Long distances carry little value in the method. These are blind spots; we noted the blind spot due to the exchange of the 5′-G* N1H. Numerous other, but more subtle, blind spots exist for protons (such as the G NH_2) with an undetectable signal. The contacts for the XG*G* and XA*G* 3bp region calculated from NOE-based models differ, but not for many contacts and often not so dramatically. Thus, for the C base differing in position by 3.5 Å between models (Figure 4), the distance of H6 or H5 of this C to the H6 of the adjacent C

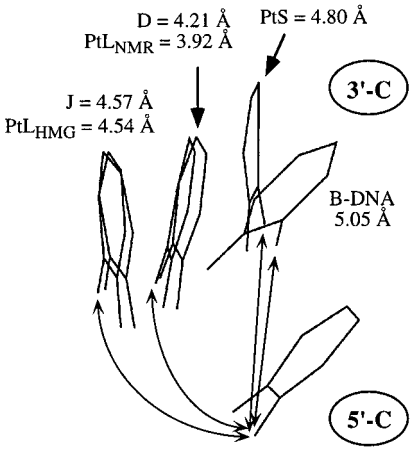


Figure 6. H6−H6 distances in various models and in the 16-mer for the C bases in the G*•C bp's. (The figure was constructed by superimposing the 5′-C bases (in the 3′-G*•C bp.)

differs by only ∼0.3 Å between models (Figure 6 and Supporting Information). This situation arises because the modeling was directed at optimizing interproton distances. It is clear that such contacts are, in many cases, not sensitive to the positions of the bases. Since the backbone is not well defined by NOE data, modeling directed at explaining the observed NOE cross-peaks has led to some abnormal backbone torsion angles. However, there is no evidence for abnormal backbones. As mentioned, the ^{31}P NMR chemical shifts, which can change upon distortion of the backbone, are normal except for the G*pG* signal.

When we abandoned the NOE-based modeling effort for the X-ray/NMR method, we were surprised at how well the NOE data were fit. The better *R*-factor of model J results from a general overall improvement of the fit of the NOE cross-peaks. As mentioned, all the NOE-based models for the duplexes in Table 1 fit the NOE data for that duplex reasonably well, but a few contacts are not explained well by the published models or our NOE-based ones. For the non-Pt strand, one contact between the C residues mentioned above does differ between models. The C13 H3′−C14 H5 distance is well below 4 Å in model J but above 4 Å in all other models. The distances between the purine H8 of the 5′-flanking base pair and the protons of the sugar residue of the adjacent C (complementary to the 5′-G*) were mentioned above. For example, the weak C H2′-purine H8 NOE cross-peak (C14 H2′−G15 H8 for the 9-mer) indicates that this contact is longer by ∼1 Å than those suggested by NOE-based models. A long contact distance (4.5 Å) is possible only when the XG* bp step has a large positive slide and shift and when the base complementary to this G* is moved substantially from the B-DNA position, as in model J and the 16-mer.[2] For the Pt strand, even fits of contacts from the bp's flanking the 3bp region are improved. For example, only model J fits the C3 H2″−C4 H5 and C3 H3′−C3 H6 distances well. Since the NMR data are similar for all relevant duplexes, model J accounts well for all reported NMR data.

Conclusions

Cross-link formation creates larger departures from B-DNA structure than previously recognized for the XG*G* 3bp region of duplexes in solution. In contrast, the structure of the PtG*$_2$ moiety is distorted relatively little compared to the structure found many years ago for G*G* small single-strand oligomers.[14,25] The new NMR data for the 9-mer, the recently reported data for a third TG*G*T duplex with a purine-rich

Pt-strand,[17] and previously reported data for duplexes (including one with a XA*G* lesion) support the earlier proposal[4] that the DNA distortions are similar. For all duplexes with an adjacent intrastrand cross-link lesion in which the X residue is C or T, we now conclude that the new unusual base-pair step with large positive shift and slide found by Lippard and co-workers[2] is present and that the C or T residue complementary to the 5′-G* or 5′-A* is dramatically repositioned by cross-link formation, as revealed in the 16-mer structure by Lippard and co-workers.[2]

Since the structural features of the XG*G* region in the 16-mer X-ray structure[2] and the 9-mer X-ray/NMR-based model J are closely related, either can serve in the design of new carrier ligands for improving Pt anticancer drug activity and in the analysis of structure−activity relationships. It must be noted, however, that the "backside" of the carrier ligand can influence interactions of the 3bp region with proteins and enzymes.[29] The 9-mer X-ray/NMR-based model J and the 16-mer structures clearly indicate at best a weak hydrogen bond[30] between the phosphate group (XpG*) and the ammine group cis to it. It is more likely that the small size of the hydrogen atom, rather than its H-bonding ability, is the key feature of the drug. Indeed, steric clashes are evident when the NH groups are substituted by bulky alkyl groups, which project out toward the DNA in the 16-mer and model J structures. Complexes with such carrier ligands are generally less active than complexes with primary sp^3 amine donors.[18,31,32] Thus, the design of active compounds may require carrier ligands that avoid clashes with the DNA. Complexes with both amines replaced by an sp^2 N-heterocycle are active.[33] Indeed, replacement of either NH_3 with an sp^2 N-heterocycle such as pyridine or 2-picoline (2-pic) in model J does not lead to steric clashes. The complex *cis*-Pt(2-pic)(NH_3)Cl_2 is active and is a candidate for clinical use.[34−36] NMR data obtained for the PtS analogue with one NH_3 replaced by 2-pic were consistent with a model derived by slightly modifying the PtS NOE-based model.[37] Our computer docking analysis indicates that even without modification, model J accounts for these NMR data. Also, model J suggests that the 2-pic atropisomer indicated by the NMR data for the PtS analogue is preferred over the other atropisomer.

In our recent studies using retro models to assess the importance of the carrier ligand NH groups in influencing the conformation of small nucleic acid fragments,[38−41] evidence is mounting that the unique feature of the NH group is its small size, not its hydrogen-bonding ability. Thus, these small-molecule studies are consistent with the 16-mer structure and the 9-mer X-ray/NMR-based model J.

The structural similarity of the duplexes examined by NMR methods and listed in a Supporting Information table suggests that the structures of the free and protein-bound DNA are not very sequence dependent, leading to the reasonable hypothesis that the sequence dependence of protein binding[19,20] arises mainly from sequence-dependent differences in protein−DNA contacts. We also conclude that the additional structural changes in the 3bp region accompanying DNA binding to protein, at least to the HMG protein, are much less severe than the initial large changes in DNA structure caused by the intrastrand cross-link formation. The most complete NMR studies have C or T adjacent to the G* residues, and recent data suggest that duplexes with either one G* or, more particularly, both G*'s flanked by a purine may have a different structure.[42]

Acknowledgment. This work was supported by NIH Grant GM 29222.

Supporting Information Available: Details of sample preparation, 1D and 2D NMR experiments and assignments; molecular modeling of models A−D with *R*-factors; comparison of structural parameters of the 9-mer models and some previously studied duplexes; ^{13}C NMR shifts of the 9-mer; 2D NOESY, DQF−COSY, and ^{31}P NMR spectra of the 9-mer; stereoviews of model J (with PDB file); superimposed bases of representative models; tables of improved NOE cross-peaks; and the N7−Pt−N7 angles of representative models (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0007915

(29) Vaisman, A.; Lim, S. E.; Patrick, S. M.; Copeland, W. C.; Hinkle, D. C.; Turchi, J. J.; Chaney, S. G. *Biochemistry* **1999**, *38*, 11026−11039.
(30) Jeffrey, G. A. In *An Introduction to hydrogen bonding*; Jeffrey, G. A., Ed.; Oxford University Press: New York, 1997; pp 11−32.
(31) Cleare, M. J.; Hoeschele, J. D. *Bioinorg. Chem.* **1973**, *2*, 187−218.
(32) Hirano, T.; Inagaki, K.; Fukai, T.; Alink, M.; Nakahara, H.; Kidani, Y. *Chem. Pharm. Bull.* **1990**, *38*, 2850−2852.
(33) Bloemink, M. J.; Heetebrij, R. J.; Ireland, J.; Deacon, G. B.; Reedijk, J. *J. Biol. Inorg. Chem.* **1996**, *1*, 278−283.
(34) Raynaud, F. I.; Boxall, F. E.; Goddard, P. M.; Valenti, M.; Jones, M.; Murrer, B. A.; Abrams, M.; Kelland, L. R. *Clin. Cancer Res.* **1997**, *3*, 22063-2074.
(35) Kelland, L. R.; Sharp, S. Y.; O'Neill, C. F.; Raynaud, F. I.; Beale, P. J.; Judson, I. R. *J. Inorg. Biochem.* **1999**, *77*, 111−115.
(36) Wong, E.; Giandomenico, C. M. *Chem. Rev.* **1999**, *99*, 2451−2466.
(37) Yu, C.; Parkinson, J. A.; Guo, Z.; Brown, T.; Sadler, P. J. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 2060−2063.
(38) Ano, S. O.; Intini, F. P.; Natile, G.; Marzilli, L. G. *Inorg. Chem.* **1999**, *38*, 2989−2999.
(39) Wong, H. C.; Shinozuka, K.; Natile, G.; Marzilli, L. G. *Inorg. Chim. Acta* **2000**, *297*, 36−46.
(40) Carlone, M.; Fanizzi, F. P.; Intini, F. P.; Margiotta, N.; Marzilli, L. G.; Natile, G. *Inorg. Chem.* **2000**, *39*, 634−641.
(41) Sullivan, S. T.; Ciccarese, A.; Fanizzi, F. P.; Marzilli, L. G. *Inorg. Chem.* **2000**, *39*, 836−842.
(42) Pilch, D. S.; Dunham, S. U.; Jamieson, E. R.; Lippard, S. J.; Breslauer, K. J. *J. Mol. Biol.* **2000**, *296*, 803−812.