Solution Structures of Psoralen Monoadducted and Cross-Linked DNA Oligomers by NMR Spectroscopy and Restrained Molecular Dynamics[†,‡]

H. Peter Spielmann,[§] Tammy J. Dwyer,[∥] John E. Hearst, and David E. Wemmer*

Structural Biology Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, California 94720

Received April 27, 1995; Revised Manuscript Received July 25, 1995[⊗]

ABSTRACT: We have used two-dimensional ^{1}H NMR spectroscopy to determine the solution structures of the 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT) furanside monoadducted (MAf) and the photoisomeric HMT interstrand cross-linked (XL) DNA oligonucleotide d(5′-GCGTACGC-3′)$_2$. The determination of the structure was based on total relaxation matrix analysis of the NOESY cross-peak intensities using the program MARDIGRAS. Improved procedures to consider the experimental "noise" in NOESY spectra during these calculations have been employed. The NOE-derived distance restraints were applied in restrained molecular dynamics calculations. Twenty final structures each were generated for both the MAf and XL from both A-form and B-form dsDNA starting structures. The root-mean-square (rms) deviations of the coordinates for the 40 structures for the MAf and XL were 1.12 and 1.10 Å, respectively. The rmsd of the MAf with respect to the XL is 2.20 Å. The local DNA structure is distorted in both adducts, with the helix unwound by 34° and 25° for the MAf and XL, respectively, and an overall helical repeat of 11 base pairs, caused by intercalation of the HMT. The MAf is a photochemical intermediate on the path to interstrand XL. Considerable local structural distortion is induced by both adducts, but the DNA returns to B-form structure within three base pairs of the damage site. There is no significant bend in the helix axis of either the MAf or the XL. We have evaluated the accuracy of the two major methods of converting NOESY data into interproton distances, the isolated spin-pair approximation (ISPA) and the complete relaxation rate matrix analysis (RMA). Both methods were evaluated by comparing the resulting calculated interproton distances generated to known covalently fixed distances in the HMT. The overall structures were evaluated by checking their agreement with biophysical evidence from non-NMR techniques. Only the modified RMA method gave correct interproton distances.

Psoralens are a class of photomutagenic (Cassier et al., 1984; Fujita, 1984; Misra & Vos, 1993; Sage et al., 1993; Yang et al., 1994) and photochemotherapeutic agents (Corash et al., 1992; Morel et al., 1992; Lin et al., 1993) that produce monoadducts and interstrand cross-links in DNA and RNA capable of blocking replication (Shi et al., 1988a,b) and transcription (Sastry & Hearst, 1991a,b). Psoralens have been used for the treatment of vitiligo, psoriasis, and other cutaneous diseases (Fitzpatrick et al., 1982; Pathak & Fitzpatrick, 1992). Thymidine is the primary nucleoside in DNA at which photoreaction occurs. The reaction takes place between the 3,4 (pyrone) or 4′,5′ (furan) double bonds of the psoralen and the 5,6 double bond in pyrimidines (Straub et al., 1981; Kanne et al., 1982a,b; Shi et al., 1990). The effects of psoralen adducts on the structure and function of DNA have been extensively studied (del Olmo et al., 1993; Sastry et al., 1993; Huang & Sancar, 1994; Islas et al., 1994; Ramaswamy & Yeung, 1994). Previous studies have shed light on the mechanism of psoralen adduct formation and their biological consequences (Van Houten et al., 1986a, 1988; Sancar & Sancar, 1988; Smith, 1988; Cheng et al., 1988a; Sladek et al., 1989; Smith et al., 1989).

4′-(Hydroxymethyl)-4,5′,8-trimethylpsoralen (HMT,[1] **1**), is a synthetic psoralen derived from the natural product 4,5′,8-trimethylpsoralen (TMP). HMT has greatly enhanced water solubility relative to TMP and forms high yields of adducts with DNA. The structures of the three main adducts between an isolated thymidine nucleoside and HMT (Cimino

[†] This work was supported in part by National Institutes of Health Grant GM-43219 (D.E.W.), by postdoctoral fellowship GM-14966 (H.P.S.), through instrumentation grants from the U.S. Department of Energy, DE FG05-86ER75281, and the National Science Foundation, DMB 86-09305 and BBS 87-20134, and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098 (D.E.W. and J.E.H.).

[‡] Coordinates for the HMT furanside monoadduct of d(GCGTACGC)$_2$ and for the HMT interstrand cross-link of d(GCGTACGC)$_2$ have been deposited in the Brookhaven Data Bank under Identifier Codes 203D and 204D.

* Author to whom correspondence should be addressed at the Department of Chemistry.

[§] Present address: Department of Biochemistry, University of Kentucky, 800 Rose St., Lexington, KY 40536-0084.

[∥] Present address: Department of Chemistry, University of San Diego, 5998 Alcala Park, San Diego, CA 92110.

[⊗] Abstract published in *Advance ACS Abstracts,* October 1, 1995.

[1] Abbreviations: AMBER, assisted model building with energy refinement; dsDNA, double-stranded DNA; EDTA, ethylenediaminetetraacetic acid; MARDIGRAS,- matrix analysis of relaxation for discerning the geometry of an aqueous structure; MOPAC, molecular orbital package; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; RANDMARDI, randomized MARDIGRAS; RMD, restrained molecular dynamics; TPPI, time-proportional phase incrementation; HMT, 4′-(hydroxymethyl)-4,5′,8-trimethylpsoralen; AMT, 4′-(aminomethyl)-4,5′,8-trimethylpsoralen; TMP, 4,5′,8-trimethylpsoralen; ISPA, isolated spin-pair approximation; COSY, correlation spectroscopy; UM, unmodified DNA oligomer 5′-GCGTACGC-3′; MAf, psoralen furanside monoadduct; XL, psoralen interstrand cross-link; DQF-COSY, double-quantum-filtered correlation spectroscopy; RMA, relaxation rate matrix analysis.

0006-2960/95/0434-12937$09.00/0

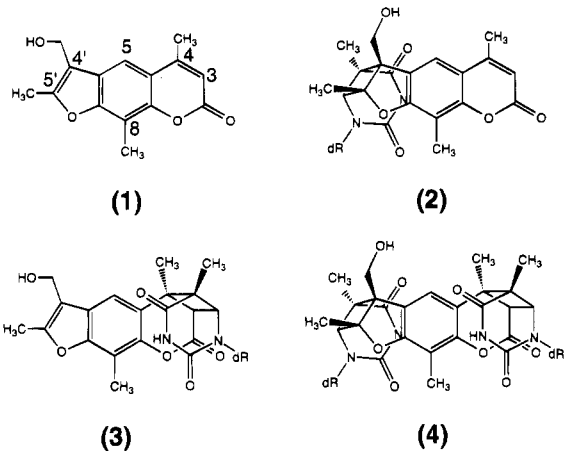

FIGURE 1: Structures of HMT and primary DNA−HMT nucleoside photoadducts: HMT (**1**), psoralen furanside monoadduct (MAf) (**2**), psoralen pyroneside monoadduct (MAp) (**3**), and psoralen interstrand cross-link (XL) (**4**).

et al., 1985) are shown in Figure 1. The photoadducts have a *cis*−*syn* stereochemistry about the cyclobutane rings determined by the noncovalent intercalation complex. The reaction of psoralen with DNA has been proposed to take place via the intercalation of the planar psoralen into double-stranded nucleic acid, followed by absorption of a long-wavelength UV photon (320−410 nm) causing the psoralen to react with pyrimidine bases to form either a furanside **2** or pyroneside monoadduct **3**. The yield of furanside monoadduct relative to pyroneside adduct in DNA depends upon the presence or absence of the psoralen 4-methyl group (Kanne et al., 1984). The predominant monoadduct formed by psoralens with a 4-methyl group is the furanside monoadduct. An interstrand cross-link **4** can be formed by absorption of a second photon by the furanside monoadduct in a site with an adjacent pyrimidine base on the opposite strand of double-stranded DNA (or RNA). Pyroneside monoadducts do not absorb light at wavelengths above 320 nm and cannot be driven to interstrand cross-links with long-wavelength UV light. All of the psoralen−DNA adducts can be photoreversed by exposure to 254 nm UV light (Cimino et al., 1986) or, in the case of either furanside monoadduct and cross-link, by treatment with base (Shi et al., 1988a; Yeung et al., 1988). The presence of a psoralen monoadduct or an interstrand cross-link stabilizes the helix to thermal denaturation relative to unmodified DNA (Shi & Hearst, 1986).

There are conflicting reports in the literature regarding the structural consequences of psoralen-induced damage to dsDNA. Flow linear dichroism provided evidence for pseudointercalative adduct geometry of the furanside monoadduct (Vigny et al., 1987). Models based on the structure of the isolated furanside psoralen−thymine monoadduct led to the proposal that a psoralen cross-link induces a sharp kink in and a large unwinding of the DNA duplex (Pearlman et al., 1985). A model derived from NMR data for the octamer d(5′-GGGTACCC-3′)$_2$ cross-linked with the psoralen AMT [4′-(aminomethyl)-4,5′,8-trimethylpsoralen] exhibited a kink at the site of photoalkylation (Tomic et al., 1987). Such results contradict studies (Sinden & Hagerman, 1984; Haran & Crothers, 1988) in which psoralen cross-links did not alter the mobility of DNA fragments in a gel shift assay, indicating that the cross-links do not induce a bend into the DNA helix axis.

Due in part to the controversy surrounding the solution structure of the XL and the lack of structural information about the MAf, we have determined the structures of the HMT furanside monoadduct of d(GCGTACGC)$_2$ (MAf) and the HMT interstrand cross-link d(GCGTACGC)$_2$ (XL) using ^{1}H NMR. We recently reported a preliminary account of the structural and functional characterization of these molecules (Spielmann et al., 1995a). NOESY spectra were used to obtain interproton distance estimates using the RANDMARDI procedure (H. Liu, H. P. Spielmann, N. B. Ulyanov, D. E. Wemmer, and T. L. James, unpublished results), a variation of the MARDIGRAS complete relaxation matrix approach (Borgias & James, 1988; Borgias et al., 1990; Schmitz et al., 1992). Distance estimates derived using this approach were then used as restraints in molecular dynamics simulations (RMD). Since many NOE contacts were observed between the HMT adducts and the surrounding DNA, the resulting structures have fairly high resolution and allowed determination of local conformational features in the DNA structure after HMT photobinding. We discuss the conformational and dynamic changes that the DNA undergoes upon adduct formation, and we also discuss the conformation of the MAf with respect to interstrand cross-link formation.

We have evaluated the accuracy of the two major methods of converting NOESY data into interproton distances, the isolated spin-pair approximation (ISPA) and complete relaxation rate matrix analysis. The HMT-modified DNAs are unusual in that there are a significant number of covalently fixed proton pairs spanning the range of distances observable by the NOE. These provide a very important test for the evaluation of the accuracy of the determined distances.

EXPERIMENTAL PROCEDURES

Materials. The DNA oligomers were synthesized on a 10-mol scale using the phosphite triester method on an Applied Biosystems Inc. (Foster City, CA) automated DNA synthesizer. The DNA oligomer samples were cleaved from the solid support and deprotected using standard methods.

We use the following numbering scheme to describe the UM, MAf, and XL:

```
5'- G1  C2  G3  T4  A5  C6  G7  C8 -3'
3'- C16 G15 C14 A13 T12 G11 C10 G9 -5'
```

Synthesis of Psoralen-Modified DNAs. The MAf and XL were synthesized and purified as described (Spielmann et al., 1992). Single-stranded MAf was formed by irradiation of an HMT + d(5′-GCGTACGC-3′)$_2$ solution with 406.7 nm light from a krypton ion laser. The resulting MAf did not absorb light at the irradiation wavelength, allowing it to accumulate in the reaction without formation of a cross link. A single isomer of the MAf was produced in the synthesis. The furanside monoadducted DNA was separated from the unmodified DNA using a multistep HPLC purification procedure. The XL was formed by irradiation of an HMT + d(5′-GCGTACGC-3′)$_2$ solution with 365 nm UV light from an argon ion laser. The major product was separated from the reaction mixture using a multistep HPLC purification procedure.

NMR Sample Preparation. (1) Unmodified DNA (UM). The purified octamer duplex was dissolved in 500 μL of 10 mM phosphate buffer (pH 7) containing 200 mM NaCl,

lyophilized several times from D_2O, and finally dissolved in 500 μL of 99.996% D atom D_2O. The resulting NMR sample was 6 mM in UM.

(2) MAf. In order to prepare a homogeneous NMR sample of the double-stranded MAf heteroduplex, the HMT monoadducted DNA strand must be hybridized with its unmodified complement. This is possible because the MAf heteroduplex is approximately 1.3 kcal/mol more stable than the unmodified DNA duplex and 3.1 kcal/mol more stable than the doubly modified DNA homoduplex (Shi & Hearst, 1986). Formation of the MAf heteroduplex required annealing one monoadducted octamer strand and one unmodified d(5′-GCGTACGC-3′)$_2$ (UM) strand. The purified, single-stranded monoadducted oligomer was taken up in 80 μL of 1 M NaCl. The UM duplex was dissolved in 200 mM NaCl/10 mM Tris-HCl-d_{11} buffer (pH 7.0), lyophilized from D_2O repeatedly, and finally taken up in 200 μL of 20 mM Tris-HCl-d_{11} buffer (pH 7.0)/400 mM NaCl and then lyophilized. Each sample was lyophilized several times from D_2O, and the octamers MAf and UM were dissolved in 400 and 70 μL of 99.996% D atom D_2O, respectively. The single-stranded octamer MAf was then titrated with the more concentrated UM to give duplex DNA. The titration was monitored by observing the appearance of two resonances assigned to the C8 protons of the adenines in the ^{1}H NMR at 600 MHz. Aliquots of the UM duplex were added; the sample was heated to approximately 80 °C in the NMR tube and then allowed to cool in order to anneal the strands. Additions of the UM duplex were made until the resonances corresponding to single-stranded MAf disappeared. The resulting sample was 8 mM in MAf heteroduplex, 200 mM in NaCl, and 10 mM in Tris-HCl-d_{11} buffer (pH 7.0).

(3) XL. The purified XL was dissolved in 500 μL of 10 mM phosphate buffer (pH 7) containing 200 mM NaCl, lyophilized several times from D_2O, and finally dissolved in 500 μL of 99.996% D atom D_2O. The resulting NMR sample was 6 mM in XL.

Proton NMR Studies. Phase-sensitive two-dimensional proton NOESY spectra in D_2O were recorded on a Bruker AMX 600 MHz NMR spectrometer. Proton chemical shifts were indirectly referenced using the residual HOD resonance. A standard NOESY sequence was used with the TPPI (Drobny et al., 1979) method of phase cycling. NOESY spectra were recorded for mixing times of 50, 100, 175, and 250 ms for both the MAf and XL. An additional 25 ms mixing time NOESY was recorded for the MAf. A single NOESY spectrum with a 200 ms mixing time was recorded for the UM. The data sets were collected with 512 t_1 experiments using 1024 complex data points in the t_2 dimension with a sweep width of 6024 Hz and 48 or 64 scans per t_1 increment. A delay time of 2.5 s between scans was allowed for relaxation of the magnetization. The carrier frequency was placed on the residual HOD resonance, which was presaturated during the relaxation delay and the mixing time. All data were transferred to a Silicon Graphics Personal Iris 4D/35 and processed with Felix (Biosym, San Diego, version 2.05). The data sets were apodized with a skewed sine-bell function in the t_2 dimension (1024 points, 60° phase shift, 0.7 skew). The future pattern of the interferogram in t_1 was linear predicted from points 513 to 650 by using the LPL protocol of Felix. The first point of the interferogram was multiplied by 0.56. The data were then apodized with a skewed sinebell function (650 points, 60° phase shift, 0.7 skew) and zero-filled to give a final 1K × 1K matrix. The two-dimensional NOESY spectra in H_2O were recorded with the three-pulse sequence described above, except that the third pulse was replaced by a 11 jump−return pulse optimized for maximum excitation of the imino protons (11−13 ppm). A 200 ms mixing time was used with 2048 data points in t_2 and 490 points in t_1 (zero-filled to 2K), collected for a spectral width of 13.1 kHz in both dimensions.

Pure absorption double-quantum-filtered COSY (DQF-COSY) spectra were recorded on a General Electric GN 500 MHz spectrometer. A spectral width of 4000 Hz in both t_2 and t_1 was used, with a repetition delay of 2.5 s; 64 scans were averaged per t_1 value; 512 FIDs were recorded in t_1 and 2048 points in t_2. Soft-COSY experiments were run as described (Emsley et al., 1993).

Calculation Methods. (A) Distance Constraints. Cross-peaks in the NOESY spectra arise from dipole−dipole relaxation between nuclei that are within ca. 5 Å. The intensities of these NOE cross-peaks provide information about the internuclear distance based on the initial NOE buildup relationship:

$$\mathrm{NOE} = f(\tau_c)\cdot\frac{1}{r^6} \quad (1)$$

where $f(\tau_c)$ is a function of the correlation time τ_c and represents the influence of molecular motion on the observed NOE. Given a known distance r_{kl}, an unknown distance r_{ij} can be directly calculated from the relationship:

$$r_{ij} = r_{kl}[(\mathrm{NOE}_{kl})/(\mathrm{NOE}_{ij})]^{1/6} \quad (2)$$

where NOE_{kl} and NOE_{ij} are the NOEs between atoms k and l and atoms i and j, respectively. The validity of these equations is limited. Equation 1 is valid only for two-spin relaxation. In practice, NOE cross-peaks result from multispin relaxation. Relaxation matrix methods have been developed to account for multispin effects. The $f(\tau_c)$ term in eq 1 may be canceled in eq 2 only under the assumption that the molecule tumbles as a rigid isotropic unit where all interproton vectors have the same effective correlation time. In most cases, macromolecules are not rigid bodies, and internal motions in the molecule cause different effective correlation times for different interproton vectors. This assumption leads to an estimated error of 10% in the calculated distances (Borgias & James, 1988).

The integrated volumes of symmetry-related cross-peaks above and below the diagonal were treated as independent measurements to more accurately reflect the scatter in the data. For the MAf, this gave 10 sets of integrated cross-peak intensities, two for each of the five mixing times, and for the XL, this gave 8 sets of integrated crosspeak intensities, two for each of the four mixing times. Not all cross-peaks were present at all mixing times. The spin−lattice relaxation rates for A5 H2 and A13 H2 and other representative protons were measured by the inversion recovery method. The measured intensities of the NOESY cross-peaks were scaled for the effects of partial saturation according to $A = 1/(1 - \exp(-T/T_1))$, where A is the correction factor, T is the recycle delay for the experiment, and T_1 is the longitudinal relaxation rate for the proton in question from which the magnetization originated.

(B) Data Analysis. The data were converted into interproton distances by three different methods, the two-spin approximation, the MARDIGRAS iterative relaxation matrix method, and the RANDMARDI modification of MARDIGRAS.

(1) Isolated Spin-Pair Approximation (ISPA). The volume for a given cross-peak as a function of the mixing time was fitted to a second-order polynomial function that was forced through zero time and zero intensity. The slope of the curve at zero mixing time was taken to be the buildup rate for that NOE. This was repeated for each measured cross-peak. The interproton distances were calculated using eq 2, where r_{kl} was a known distance, for example, the sugar H2′−H2″ (1.8 Å) or cytidine H5−H6 (2.52 Å).

(2) Relaxation Matrix Approach (MARDIGRAS). The MARDIGRAS method was used to interpret the NOE data (Borgias & James, 1988; Borgias et al., 1990). The distances returned by MARDIGRAS for the data sets were averaged, and standard deviations for the measurements were calculated using the program CALSD (Schmitz et al., 1992). A total of 342 experimentally measured NOE intensities were included in the MARDIGRAS calculation of the MAf data, and 327 were used in the XL. MARDIGRAS uses internal calibrations when it iteratively fits the data to a model of the molecule. In particular, it resets the cytosine H5 to H6, the thymidine 5-methyl to H6, and deoxyribose H2′ to H2″ interproton distances to the values assumed from the covalent geometries. NOE intensities for these known distances are used to normalize the NOE intensities in the model. The program does not reset fixed interproton distances within the HMT adducts because these distances were intentionally left out of the reference file. The magnetization transfer for these protons is calculated and provides a check of how well the MARDIGRAS procedure is working. Agreement between the actual and the experimentally determined distances indicates that the MARDIGRAS procedure yields more accurate interproton distances than the ISPA procedure (see discussion below).

(3) RANDMARDI Procedure. The RANDMARDI procedure (Liu et al., unpublished results) of the complete relaxation matrix analysis method MARDIGRAS was used to calculate interproton distance RMD bounds from the resulting integrated peak intensities. The RANDMARDI procedure was used to include the effects of experimental noise in the relaxation matrix calculations. RANDMARDI adds a random number from within a specified range to each input intensity used in a MARDIGRAS calculation. For the calculations performed on the HMT adducts a noise level of 1−5 times the integrated intensity of the smallest cross-peak was used as the "noise" range. The dynamic range of observed cross-peak intensities was 1000. In the RANDMARDI procedure, 30 different intensity sets from each experimental data set were generated, and MARDIGRAS calculations were performed on all of them. Resulting distances were then averaged together to form one bounds file for each data set. The bounds files from all of the mixing times were then combined into a single bounds file from which the RMD restraint file is generated. Upper and lower bounds were the average interproton distances ±1 SD calculated from all of the MARDIGRAS runs. A complete description of the RANDMARDI theory and procedure will be given elsewhere (Liu et al., unpublished results). Good agreement between the actual and the experimentally determined distances indicates that the RANDMARDI procedure yielded the most accurate interproton distances of the three methods tested (see Tables 1 and 2 and discussion below).

(C) Starting Structures. The starting models for the refinement were generated using standard Arnott parameters for A-form and B-form DNA in the Biopolymer module of InsightII (version 2.3.0) (Biosym Technologies, San Diego) and assigned potentials using the AMBER force field. The HMT moiety was created using the Builder module of InsightII, and partial charges were calculated using semiempirical AM1 parameters as implemented by the MOPAC module of InsightII. The furanside monoadduct was assembled by covalently bonding the HMT with thymidine T4 in the oligomer and then editing the atom types and connectivities to form the cyclobutane ring. For the cross-link, a second covalent modification was made between the pyrone ring of the HMT and thymidine T12. Starting models were generated by energy minimizing the MAf and XL models using only the AMBER force-field function of the Discover module (version 2.9.5) of InsightII with 100 steps of steepest descents minimization, followed by conjugate gradient refinement to an rmsd of 0.01 Å. The distance bounds generated using the RANDMARDI procedure were then applied to the models, which were minimized again to arrive at the starting structures. The five methyl groups of the psoralen and thymidines were represented as pseudoatoms at the geometric center of mass of the three protons.

(D) Restrained Molecular Dynamics. The distance restraints obtained from the RANDMARDI calculations were incorporated into an RMD procedure. The RMD and energy minimization calculations were performed using Discover. The models were displayed using InsightII. The NOE-derived distance restraints were applied to the models,

Table 1: Covalently Fixed Interproton Distances (Å) in MAf

atom pair		actual values[a]	lower bound[b]	upper bound[b]	width[b]	violations
T4 Me	T4 H6	2.66	2.57	3.00	0.43	0.00
4-Me	H5	2.72	2.42	2.64	0.22	0.08
5′-Me	T4 H6	2.85	2.17	2.58	0.41	0.27
4-Me	H3	3.02	2.77	2.94	0.17	0.08
T4 Me	5′-Me	4.11	3.97	4.86	0.89	0.00
T4 Me	H5	4.52	3.97	5.65	1.68	0.00
H3	H5	4.85	3.62	4.88	1.26	0.00
5′-Me	H5	5.60	4.45	6.82	2.37	0.00
8-Me	5′-Me	5.80	4.88	6.33	1.45	0.00

[a] Obtained from MOPAC calculations. [b] Obtained from RANDMARDI calculations and derived from the average distance ± 1 SD.

Table 2: Covalently Fixed Interproton Distances (Å) in XL

atom pair		actual values[a]	lower bound[b]	upper bound[b]	width[b]	violations
T12 H6	H3	2.35	1.89	2.19	0.30	0.16
H5	4-Me	2.67	2.57	2.91	0.34	0.00
T12 H6	T12 Me	2.70	2.61	2.96	0.35	0.00
T4 H6	T4 Me	2.74	2.66	2.90	0.24	0.00
T4 H6	5′-Me	2.88	2.43	2.64	0.21	0.24
H3	4-Me	2.89	2.71	3.33	0.62	0.00
4-Me	T12 Me	3.14	2.91	3.23	0.32	0.00
T12 H6	4-Me	3.70	3.21	4.48	1.27	0.00
H3	T12 Me	4.07	3.15	6.16	3.01	0.00
H5	T12 Me	4.18	3.44	4.50	1.06	0.00
5′-Me	T4 Me	4.46	3.84	6.41	2.57	0.00
H5	T4 -Me	4.56	3.64	6.28	2.64	0.00
H5	H3	4.71	3.30	4.96	1.66	0.00
H5	T12 H6	5.36	3.03	5.12	2.09	0.24
H5	5′-Me	5.51	3.55	6.60	3.05	0.00
H5	T4 H6	5.53	3.46	5.27	1.81	0.26
T4 Me	8-Me	5.84	5.61	6.96	1.35	0.00
T4 H6	8-Me	6.10	3.78	5.92	2.14	0.18
4-Me	T4 Me	6.83	4.15	7.12	2.97	0.00
4-Me	8-Me	6.94	6.04	7.22	1.18	0.00

[a] Obtained from MOPAC calculations. [b] Obtained from RANDMARDI calculations and derived from the average distance ± 1 SD.

followed by 20 ps of restrained molecular dynamics at 350 K for 4 ps, followed by warming to 400 K for 4 ps and then cooling to 200 K in 50 K steps of 3 ps each. The final structure was then energy minimized to a maximum derivative of 0.01 Å. The pseudoenergy term used to enforce the distance constraints was

$$E_{\text{constr}} = \begin{cases} k_1(r - r_2)^2 & \text{when } r_2 > r \\ 0 & \text{when } r_3 \geq r \geq r_2 \\ k_2(r - r_3)^2 & \text{when } r_3 > r \end{cases}$$

where r_2 and r_3 define the upper and lower bounds of the flat region of the potential, corresponding to the distance bounds derived from the cross-peak volumes, and k_1 and k_2 are force constants that could be independently assigned to each constraint. All of the NOE-derived distance constraints were assigned an upper and lower bound force constant of 50 kcal/(mol·Å^2). The width of the flat region of the potential well reflects the accuracy of a given constraint (see below). An additional 18 distance constraints were included to enforce Watson−Crick hydrogen bonding throughout the calculations. Three hydrogen bonds were included for each of the six G-C base pairs where the upper and lower margins were set to 1.7−2.1 Å. No hydrogen bond restraints were used for the A-T base pairs. Water and counterions were ignored in all calculations. Helix parameters were calculated with the program CURVES 3.1 (Lavery & Sklenar, 1988, 1989).

RESULTS

Spectral Analysis and Assignment of Exchangeable DNA Protons. Melting studies were performed to determine the stability of the helix in the adducted DNAs. One-dimensional NMR spectra in H_2O were acquired at various temperatures while observing the imino proton resonances (Figure 2). Sharp imino resonances imply protection from exchange due to hydrogen bonds. The imino proton resonances were assigned by 1D NOE experiments and confirmed by NOESY experiments performed in H_2O.

(1) MAf. A single species was observed when the adducted strand of the MAf was titrated with UM in the NMR to form heteroduplex DNA. This confirms the observation by Shi and Hearst (1986) that the presence of a psoralen monoadduct stabilizes a duplex. It is commonly expected that intercalation imposes a large shift (~0.5−1 ppm) (Assa-Munt et al., 1985a,b) on imino proton resonances adjacent to the binding site. The T4 imino proton resonance is sharp and shifted upfield approximately 2 ppm relative to the T4 imino of the free DNA. The larger than expected upfield shift is a consequence of the formation of the cyclobutane linkage between the psoralen and T4. A similar large upfield shift was observed for the imino proton resonances of a thymine dimer in a duplex DNA (Taylor et al., 1990). In the crystal structure of the isolated MAf nucleoside, the imino proton is directly over the center of the psoralen benzene ring (Peckler et al., 1985). The upfield shift is likely due in part to a ring current effect. The T12 imino is slightly broader and resonates at 13.2 ppm. The G1 and G9 imino resonances broaden out with increasing temperature and become unobservable at 35 °C. The remaining six imino resonances all broaden with increasing temperature in a homogeneous fashion. The T4 imino resonance remains the sharpest signal in the spectrum even

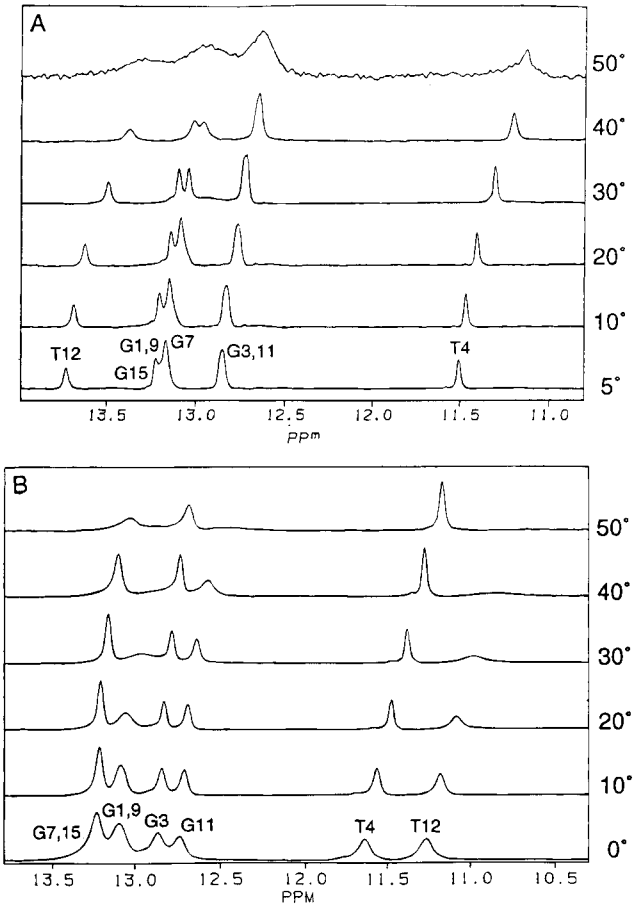

FIGURE 2: Temperature series of one-dimensional NMR spectra of the imino proton resonances of the (A) MAf and (B) XL acquired in H_2O.

at 50 °C. These data indicate that normal hydrogen bonds are formed in the MAf. The decrease in the signal-to-noise ratio of the higher temperature spectra in Figure 2A results from broadening due to melting of the DNA duplex.

(2) XL. The imino spectrum of the XL shows more symmetry than the imino spectrum of the MAf (Figure 2B). The imino resonances of both T4 and T12 in the XL are shifted upfield by 2 ppm relative to free DNA as a consequence of forming the cyclobutane rings of the adduct. There is a marked asymmetry in the stability of the helix on the two sides of the adduct in the HMT cross-linked d(GCGTACGC)$_2$. The two base pairs on the pyroneside of the adduct are destabilized relative to both normal DNA and the furanside base pairs. The T12 imino begins to broaden at 20 °C and becomes unobservable at 40 °C. The adjacent G11 imino resonance begins to broaden at 40 °C and is unobservable at 50 °C. The T4 imino proton resonance remains sharp even at 50 °C. The G3 imino resonance is sharp and only shows broadening above 50 °C. Chemical shifts indicate that the DNA returns to a normal state three base pairs from the site of adduction on both sides of the psoralen since at that point they are equivalent to UM. We interpret these data to mean that a hydrogen bond with enhanced stability is formed between T4 and A13, while a weak or nonexistent hydrogen bond is formed between T12 and A5. The hydrogen bonds for G3 are stabilized relative to those for G7 and G15. We ascribe a normal hydrogen-bonding pattern to the G3-C14 , G11-C6, G7-C10 and the G15-C2 base pairs. Similar results were observed for the AMT crosslinked oligomer d(GGGTACCC)$_2$ (Tomic et al., 1987). Hence, the localized destabilization of hydrogen

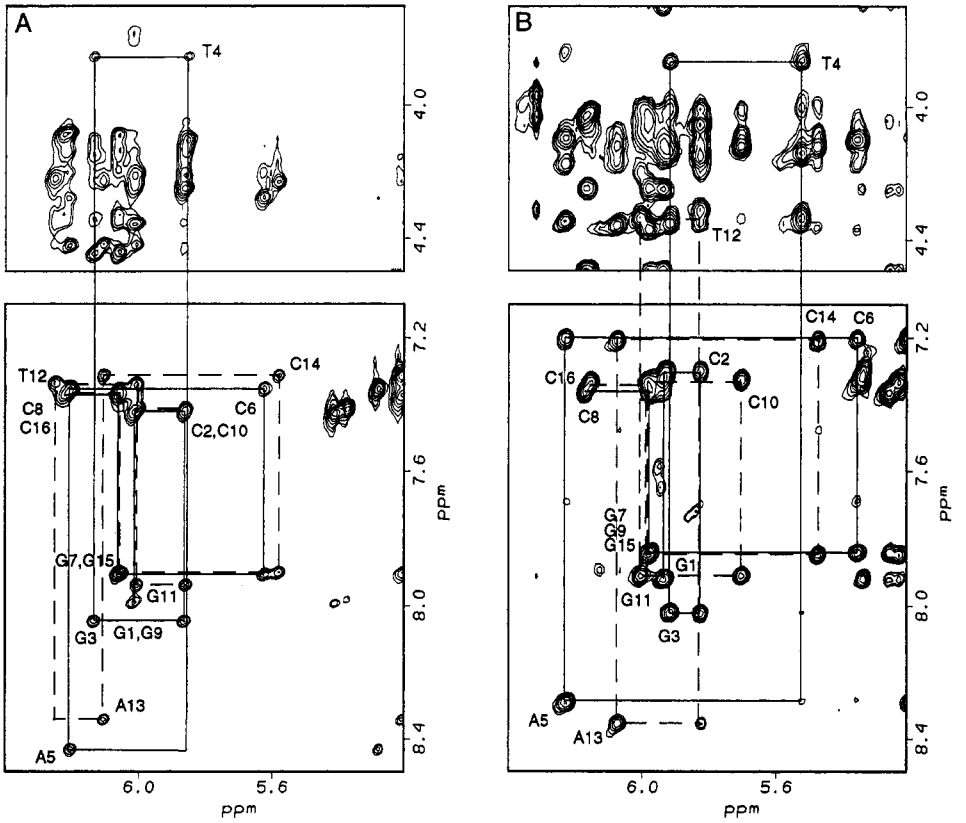

FIGURE 3: Deoxyribose H1′ to aromatic part of the NOESY spectrum (250 ms mixing time) of the (A) MAf and (B) XL. The sequential H1′−H6/H8−H1′ connectivity pathway for the furanside strand is indicated by a solid line, and the connectivity pathway for the pyroneside strand is indicated by a dotted line. The interrupted connectivities at the 5′-T4pA5-3′ base pair step due to intercalation are indicated.

Table 3: Chemical Shift Values of the HMT MAf and XL of d(GCGTACGC)$_2$ at 25 °C[a]

	H8/H6		H5/Me/H2		H1′		H2′R		H2′S		H3′		H4′		H5′S		H5′R			HMT	
	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL		MAf	XL
G1	7.99	7.91			6.01	5.93	2.85	2.73	2.66	2.21	4.93	4.82	4.35	4.24		3.69		3.69	H3	4.60	2.20
C2	7.43	7.30	5.41	5.33	5.86	5.70	2.44	2.21	2.04	1.82	4.94	4.80							4-Me	1.88	1.37
G3	8.05	7.91			6.14	6.00	2.85	2.64	2.71	2.64	5.07	4.97	4.45	4.33	4.14	4.00	4.24	4.10	H5	6.87	6.32
T4	3.86	4.31	1.27	0.93	5.85	5.82	2.27	2.33	2.33	2.07	4.63	4.58		4.31					H4′	4.13	3.95
A5	8.43	8.34	7.86	7.47	6.20	6.07	2.97	2.83	2.83	2.56	5.00	4.85	4.42	4.32	4.04	4.13	4.13	4.13	H4″	4.05	3.89
C6	7.36	7.20	5.28	5.10	5.63	5.47	2.37	2.22	2.01	1.87	4.83	4.97	4.20	4.34	4.09	4.10	4.28	4.09	5′-Me	1.69	1.53
G7	7.90	7.84			6.05	5.97	2.79	2.70	2.65	2.62	5.05	4.97	4.45	4.36	4.28	4.00	4.19	4.09	8-Me	1.12	0.35
C8	7.37	7.35	5.19	5.24	6.21	6.15	2.79	2.59	2.30	2.59	4.58	4.97		4.49		4.03		4.25			
G9	7.99	7.91			6.01	5.93	2.85	2.75	2.65	2.21	4.92	4.82	4.35	4.24		3.69		3.69			
C10	7.42	7.30	5.37	5.33	5.86	5.70	2.46	2.28	2.10	1.84	4.94	4.85									
G11	7.94	8.02			6.00	5.91	2.76	2.86	2.63	2.62	5.07	4.83	4.42	4.33	4.14	3.99	4.24	4.13			
T12	7.33	3.86	1.58	1.12	6.25	5.52	2.64	2.10	2.43	2.10	5.03	4.33	4.22	4.32							
A13	8.34	8.29	7.36	7.68	6.10	6.22	2.83	2.88	2.52	2.83	4.85	4.94	4.42	4.33	4.22	3.86	4.33	4.08			
C14	7.31	7.20	5.21	5.20	5.58	5.35	2.36	2.22	2.00	1.92	4.83	4.86	4.16	4.34	4.23	4.10		4.09			
G15	7.90	7.84			6.05	5.97	2.79	2.70	2.64	2.61	5.05	4.97	4.44	4.36	4.23	4.00	4.11	4.10			
C16	7.37	7.34	5.20	5.24	6.21	6.15	2.79	2.59	2.30	2.59	4.58	4.97		4.49		4.03		4.25			

[a] The values are given relative to the HOD signal at 4.78 ppm.

bonding on the pyrone face of the cross-link appears to be a general result for psoralen adducts and not a function of the particular psoralen used.

Assignment of Nonexchangeable DNA Protons. The NOESY spectra for the MAf in Figure 3A and for the XL in Figure 3B exhibit the characteristic features of dsDNA sequential connectivities from aromatic H6/H8 protons to both intra- and interresidue H1′ and H2′/H2″. However, the sequential connectivities are interrupted in both molecules at the 5′-TpA-3′ base pair step. This is clear evidence of intercalation by the HMT adduct in both the MAf and XL. Figure 3 also includes a few of the observed cross-peaks between the HMT adducts and the oligonucleotide. Aromatic (H6, H5, H8, and H2) and most of the deoxyribose proton (H1′, H2′, H2″, H3′, H4′, and H5′/H5″) resonances of dsDNA in the MAf and XL were assigned in the conventional way (Hare et al., 1983; Wüthrich, 1986). In normal B-form DNA, the H5′ proton of the deoxyribose has a close contact with the H1′ of the base on its 5′ face. It is possible to distinguish some of these cross-peaks from other cross-peaks in the same region of the spectrum in the short mixing time NOESY spectrum. The proton resonance assignments for the MAf and XL are given in Table 3.

Intercalation Site of HMT. The d(GCGTACGC)$_2$ duplex (UM) has only eight spectroscopically distinct residues due to its self-complementary nature. The 2-fold symmetry of the UM duplex is broken by HMT adduct formation in the MAf and XL. In the MAf, the HMT is covalently attached

Table 4: Distribution of NMR-Derived Distance Restraints Used in RMD Calculations of the MAf[a]

intraresidue				interresidue				HMT−DNA		
	intrasugar	base sugar	total		residue−residue	base adduct−DNA	total			intraadduct
G1	5	4	9	G1−C2	8		8	G3−HMT	4	
C2	5	3	8	C2−G3	10		10	T4−HMT	7	2[b]
G3	5	7	12	G3−T4	4	9	13	A5−HMT	7	
T4	5	8	13	T4−A5	1	5	6			
A5	14	8	22	A5−C6	10		10	G11−HMT	1	
C6	7	8	15	C6−G7	7		7	T12−HMT	10	
G7	3	7	10	G7−C8	6		6	A13−HMT	4	
C8	1	4	5							
G9	5	5	10	G9−C10	7		7	HMT−HMT		5[b]
C10	4	5	9	C10−G11	8		8			
G11	3	6	9	G11−T12	7		7			
T12	5	7	12	T12−A13	6		6			
A13	7	7	14	A13−C14	12		12			
C14	6	6	12	C14−G15	7		7			
G15	3	7	10	G15−C16	6		6			
C16	1	4	5							
total	79	96	**175**	total	99	14	**113**	total	**33**	**7**[b]

[a] The restraints were derived from RANDMARDI calculations. The 342 NOE connectivities gave a total of 328 usable distance restraints. Eighteen distance restraints ensuring Watson−Crick base pairing were also included in the calculations. [b] Restraints to the HMT 4′-hydroxymethyl protons.

to T4 through a cyclobutane ring formed between the 5,6 double bond of the thymidine and the 4′,5′ double bond of the psoralen. The HMT in the XL is attached to the DNA on the furanside by a cyclobutane ring to T4 and on the pyroneside by a cyclobutane ring to T12. The strand-specific assignments of the unsymmetrical modified DNAs were resolved by observation of NOE connectivities between the psoralen protons and the thymidine H6 and between the psoralen methyl protons and the adenosine H8 protons. Although the regular sequential path in NOESY spectra is interrupted at the intercalation site, there is a sequential path from the 5′ end to the 3′ end of each strand of the dsDNA by including connectivities from protons on the oligonucleotide to protons on the HMT. It is also possible to "walk" across the two strands through the psoralen. There is only a single possible path of each type, conclusively defining the intercalation site. These results indicate that the HMT is intercalated in the central 5'-TpA-3' site in both adducted molecules.

Formation of the MAf causes the H6 proton of the T4 thymidine to undergo a dramatic upfield shift from 7.3 to 3.9 ppm. The H6 protons of T4 and T12 in the XL undergo upfield shifts from 7.3 to 3.9 and 4.4 ppm, respectively.

In both molecules, the corresponding resonances for the terminal pseudo-symmetry-related G1-C16 and C8-G9 base pairs were degenerate, and most of the corresponding pseudo-symmetry-related resonances for the C2-G15 and G7-C10 base pairs were also degenerate.

Assignment of the HMT Protons. NOEs from the HMT H5 proton to the 4-methyl group and to the HMT H3 protons of the psoralen assigned the pyrone ring protons. In the NOESY spectra of both the MAf and XL there are cross-peaks from HMT H5 to the diastereotopic 4′-hydroxymethyl protons on the furan ring, the T4-methyl and spin diffusion mediated cross-peaks to both the HMT 5′-methyl protons and the HMT 8-methyl protons.

Intermolecular HMT−DNA Contacts. It was possible to assign many NOEs between the HMT and DNA protons in the adducts. There were 47 NOEs from the adduct to the surrounding bases and sugar residues in the MAf and 49 in the XL. Cross-strand NOEs are uncommon in regular DNA, arising only from AH2 to AH2 and H1′ crosspeaks in suitable sequences. There are 15 cross-strand NOEs from the psoralen to the opposite strand bases and sugar residues in the MAf. The presence of the cross-strand NOEs confirms the intercalated nature of the MAf and tightly constrains the conformation of the damaged region of the DNA. The HMT is localized relative to the unmodified strand by NOEs from the 4-methyl group to T12 H6, to A13 H8, and to the T12 methyl group. The HMT 8-methyl shows prominent NOEs to both the A5 H2 and A13 H2. NOEs between the 4′-hydroxymethyl protons and T4 H6 confirm the presence of the cyclobutane ring and its *cis−syn* geometry. In the XL, the covalent geometry of the adduct imposes a fixed pattern of NOEs that arise from the substituents on the pyroneside cyclobutane ring in addition to the furanside resonance pattern seen in the MAf. The covalent structure of the cross-link fixes the orientation and position of T4 relative to T12 and bridges the helix. The protons of the T4−HMT−T12 adduct give rise to many NOEs to the surrounding bases, and the orientation of the interstrand adduct in the helix is tightly constrained relative to the other bases. The pattern of NOEs found in the XL is similar to that for the MAf. However, the intensities of corresponding cross-peaks are different in the two molecules. This reflects the different interproton distances and geometries in the MAf and XL.

Structure Calculation. More than 400 NOE cross-peaks were observed in the NOESY spectrum for both the MAf and XL obtained with a mixing time of 250 ms. Some of them resulted predominantly from spin diffusion and were consequently not observed in NOESY spectra with shorter mixing times. The accuracy of the integration of some cross-peaks was hampered by spectral overlap. A summary of observed NOEs used in the structure determination are included in Tables 4 and 5.

Distance Restraints. In both the MAf and XL there are well-resolved direct and spin diffusion mediated NOEs among the psoralen protons. There is no ambiguity in the distances separating these protons due to the rigid covalent geometry of the psoralen adducts. The NOE intensities of cross-peaks arising from protons that are known to be a fixed distance apart were used to check the accuracy of the calculated distances. The HMT MAf and XL systems

Table 5: Distribution of NMR-Derived Distance Restraints Used in RMD Calculations of the XL[a]

intraresidue				interresidue				HMT−DNA		
	intrasugar	base sugar	total		residue−residue	base adduct−DNA	total			intraadduct
G1	3	4	7	G1−C2	5		5	G3−HMT	4	
C2	5	4	9	C2−G3	11		11	T4−HMT	7	2[b]
G3	6	5	12	G3−T4	5	9	14	A5−HMT	7	
T4	3	5	9	T4−A5	7		7			
A5	8	8	16	A5−C6	12		12	G11−HMT	4	
C6	2	4	7	C6−G7	5		5	T12−HMT	8	
G7	3	6	9	G7−C8	10		10	A13−HMT	4	
C8	2	4	6							
G9	3	4	7	G9−C10	5		5	HMT−HMT		6[b]
C10	5	6	11	C10−G11	6		6			
G11	0	5	5	G11−T12	0	4	4			
T12	5	7	13	T12−A13	4	1	5			
A13	4	6	11	A13−C14	12		12			
C14	2	5	7	C14−G15	5		5			
G15	3	7	10	G15−C16	10		10			
C16	2	4	6							
				T4−A13	1		1			
				C10−T12		1	1			
total	56	84	**140**	total	98	15	**113**	total	34	**8**[b]

[a] The restraints were derived from RANDMARDI calculations. The 327 NOE connectivities gave a total of 295 usable distance restraints. Eighteen distance restraints ensuring Watson−Crick base pairing were also included in the calculations. [b] Restraints to the HMT 4′-hydroxymethyl protons.

contains covalently determined distances that span the range of observable NOE contacts from 2.35 to 6.94 Å. The HMT−thymidine adduct has a variety of single protons and methyl groups that cross-relax each other. These psoralen protons are also cross-relaxed by the nearby protons in the DNA that are not rigidly connected to the adduct. For the MAf, there were 67 NOEs involving the HMT-T4 base unit protons, 52 NOEs from the HMT-T4 base unit to the surrounding DNA, and 15 intra-HMT−T4 NOEs. For the XL, a total of 90 NOEs involved the T12−HMT−T4 adduct, 61 NOEs from the T12−HMT−T4 unit to the surrounding DNA, and 29 intra-T12−HMT−T4 NOEs.

(1) ISPA. This method failed to produce a model for the MAf or the XL that was able to fit the observed NOE intensities with the covalentely fixed distances in the HMT because it does not account for spin diffusion. Some of the distances calculated were up to 1.3 Å too short. A number of cross-peaks in the MAf and XL NOESY spectra become visible only at the longer (175 or 250 ms) mixing times resulting from magnetization transfer due to spin diffusion. The assumption that NOEs result only from direct magnetization transfer is appropriate only for NOEs recorded with very short mixing times where signal to noise is poor. The use of the two-spin approximation is inappropriate for commonly used mixing times in cases where there is spin diffusion. Failing to compensate for spin diffusion in a molecule can lead to artificially short distances derived from the intensities measured from cross-peaks in the spectra.

(2) MARDIGRAS. The data were reinterpreted using MARDIGRAS, developed by James and co-workers to give interproton distances that account for spin diffusion. The distance bounds returned by the procedure were better than ISPA but also did not correctly fit all the observed NOE intensities to the psoralen covalent geometry. The actual distances for the MAf fell within the range of the MARDIGRAS upper and lower distance limits in five of the nine measurements. For the other four measurements, the calculated distance was between 0.19 and 0.49 Å too short or too long. Results for the XL were similar. MARDIGRAS has no provision to account for noise in the spectrum or for fluxional behavior of the molecule.

(3) RANDMARDI. The data were most successfully interpreted using the RANDMARDI modification of the MARDIGRAS procedure. Some of the NOE cross-peaks observed in the NOESY spectrum obtained with a mixing time of 250 ms resulted predominantly from spin diffusion and were consequently not observed in NOESY spectra with shorter mixing times. The absence of a measured cross-peak at some of the mixing times was taken into account during the statistical analysis of the results generated by the RANDMARDI procedure (see above). Cross-peak integrals that corresponded to fixed distances in the DNA were not converted into distance restraints for use in the RMD simulations. *MAf.* The integrated intensities from 342 NOESY cross-peaks for the MAf were used in the total relaxation matrix analysis. Ten sets of measured NOE intensities were converted to distance restraints using the RANDMARDI procedure. The calculations returned 328 interproton distances that are not covalently fixed in either the HMT or DNA in the MAf (Table 4). *XL.* The integrated intensities from 327 NOESY cross-peaks for the XL were used in the total relaxation matrix analysis. Eight sets of measured NOE intensities were converted to distance restraints using the RANDMARDI procedure. The calculations returned 295 interproton distances that are not covalently fixed in either the HMT or DNA in the XL (Table 5).

The reliability of these distance restraints was examined by comparing the covalently fixed distances in HMT−thymine adducts to the distances obtained from the RANDMARDI calculations (Tables 1 and 2). The almost complete agreement between the actual and the experimentally determined distances indicates that the RANDMARDI procedure yields accurate interproton distances. The MARDIGRAS/RANDMARDI calculations were performed with various values of the correlation time for the complex in the range 2−5 ns. The results were fairly insensitive to the actual value, and a correlation time of 3.75 ns was chosen for the final calculations.

NOEs to the diastereotopic HMT 4′-hydroxymethyl group protons were included in the RMA calculation. The distances from these protons to others in the MAf and XL are highly dependent upon the rotational conformation of the group. Because it was not possible to stereospecifically assign them, they were not included in the comparison of the actual and experimentally derived distances. Uncertainties in the distance constraints arise from internal motion in the molecule that averages the intensity of the NOE, integration errors, and the incomplete experimental description of the relaxation matrix.

Generation of Interproton Restraints. The bounds for use in RMD simulations were determined by combining the results from all of the individual MARDIGRAS calculations performed during the RANDMARDI procedure into one set. The distance bounds were calculated individually for each proton pair corresponding to a NOESY cross-peak included in the RANDMARDI calculations. Each individual restraint was generated from the average distance calculated for that proton pair from the MARDIGRAS calculations, ±1 SD calculated for that proton pair. Tables 4 and 5 give a summary of the distance restraints used in the RMD calculations. For the MAf the range of differences between the upper and lower bounds for the 328 NOE-derived restraint bounds used (cf. Table 4) was between 0.12 and 2.96 Å with an average flat well potential width of 1.18 Å. For the XL, the range of differences between the upper and lower bounds for the 295 NOE-derived restraint bounds used (cf. Table 5) was between 0.16 and 3.67 Å with an average flat well potential width of 1.53 Å. This range between the upper and lower bounds represents the accuracy of the restraints calculated from the NOESY data. Weak NOESY cross-peaks that arose primarily from spin diffusion generally had the greatest difference between their upper and lower bounds.

Deoxyribose Ring Conformations. The DQF-COSY spectra were acquired in an attempt to get *J*-coupling information to determine the pseudorotation angle of the deoxyribose residues (Ernst et al., 1987; Gochin et al., 1990). In B-form DNA the deoxyribose adopts a predominantly $C_{2'}$-*endo* conformation where the H2′ and H2″ protons may be distinguished from one another by their coupling patterns with H1′ in a DQF-COSY spectrum. The deoxyribose H1′ to H2′ cross-peak is characterized by a strong active coupling and a weak passive coupling, while the H1′ to H2″ crosspeak is characterized by a more complex pattern. The assignment of H2′ and H2″ protons from DQF-COSY coupling patterns must be made with caution because the patterns are a function of the sugar pucker. The H1′ proton is closer to the H2″ proton than to the H2′ proton in all conformations accessible to the sugar. A comparison of the relative intensities of the H1′ to H2′ and the H1′ to H2″ NOESY cross-peaks allowed unambiguous assignment of all H2′ and H2″ protons. Interpretation of the coupling patterns observed for the H1′ to H2′ and H2″ cross-peaks in the MAf and XL was difficult due to broad lines in the DQF-COSY spectra which led to cancellation of the antiphase multiplets and chemical shift degeneracy for many of the deoxyribose ring protons that further obscured the coupling patterns. High-resolution selective-COSY methods were employed in an attempt to extract backbone angle constraints for use in the modeling. It was shown that the deoxyribose sugars of the residues in and around the HMT binding site in both the MAf and the XL are in a complex mixture of rapidly equilibrating conformations (Emsley et al., 1993; Spielmann et al., 1995a). The biological implications of these complex conformational states are described elsewhere (Spielmann et al., 1995a,b).

Structure Convergence. A total of 346 restraints were used for the MAf, and a total of 313 restraints were used for the XL including the hydrogen-bonding restraints. Twenty final structures each were generated for both the MAf and XL starting with dsDNA in B-form and A-form. All the structures converged to the same conformation. The root-mean-square (rms) deviations of the coordinates of the 40 structures for the MAf and XL were 1.12 and 1.10 Å, respectively. The rmsd of the MAf with respect to the XL is 2.20 Å. The resulting total violations of the MAf summed to 11.3 Å for the A-form starting structure and to 12.6 Å for the B-form starting structure and for the XL summed to 8.4 Å for the A-form starting structure and to 9.3 Å for the B-form starting structure. Seven violations greater than 0.25 Å but less than 0.5 Å were found in the structures of the MAf derived from A-form coordinates, six for the MAf derived from B-form coordinates, five for the XL derived from A-form coordinates, and eight for the XL derived from B-form coordinates. The majority of these residual errors are localized near the HMT 4′-hydroxymethyl group and arise from the influence of the apparent conformational heterogeneity of that group and to the conformational heterogeneity of the sugars on the RMA. Different views of the calculated structures are given in Figures 4−6.

DISCUSSION

Description of the Structures. Qualitative analysis of the pattern of NOEs between the psoralen and the surrounding bases and sugars indicates that the HMT is intercalated into the helix in both the MAf and XL. In both the MAf and the XL the 8-position of the psoralen is in the minor groove, and the 4′-, 5′-, 3-, 4-, and 5-positions of the psoralen are in the major groove. Furthermore, the adducted base(s) still adopt(s) a Watson−Crick type of base-pairing geometry. The enhanced stability of the MAf relative to the UM is due to strong stacking interactions between the psoralen and the surrounding bases, in addition to the continued presence of normal Watson−Crick hydrogen bonds. As a consequence of the covalent addition of the psoralen in the MAf, the long axis of the HMT is nearly parallel to the T4-A13 axis. In the MAf there are four nonadducted bases in the immediate vicinity of the HMT adduct (G3, A5, T12, and A13) in addition to the photoadducted thymidine.

NMR-derived restraints are short range in nature, and some helical parameters are better defined by the experimental data than others. In addition, the anisotropic distribution of observable interproton contacts by NMR in free dsDNA (largely confined to the ribbon of the sugar−phosphate backbone) defines some helical parameters poorly. In the case of the MAf and XL, there are a large number of NMR-derived restraints in the center of the helix due to the presence of the HMT. These HMT−DNA contacts are distributed fairly isotropically and therefore lock down the conformation for the central four base pairs of both molecules.

(1) MAf. The adenine bases are stacked coplanar with the psoralen in the MAf (Figure 4A). The six-membered ring of A5 is stacked onto the HMT 8-methyl group, and the adenine 6-amino group is stacked directly on top of the central benzene ring of HMT. The pyrimidine ring of T12 in the MAf is centered over the H3 of the HMT, and there

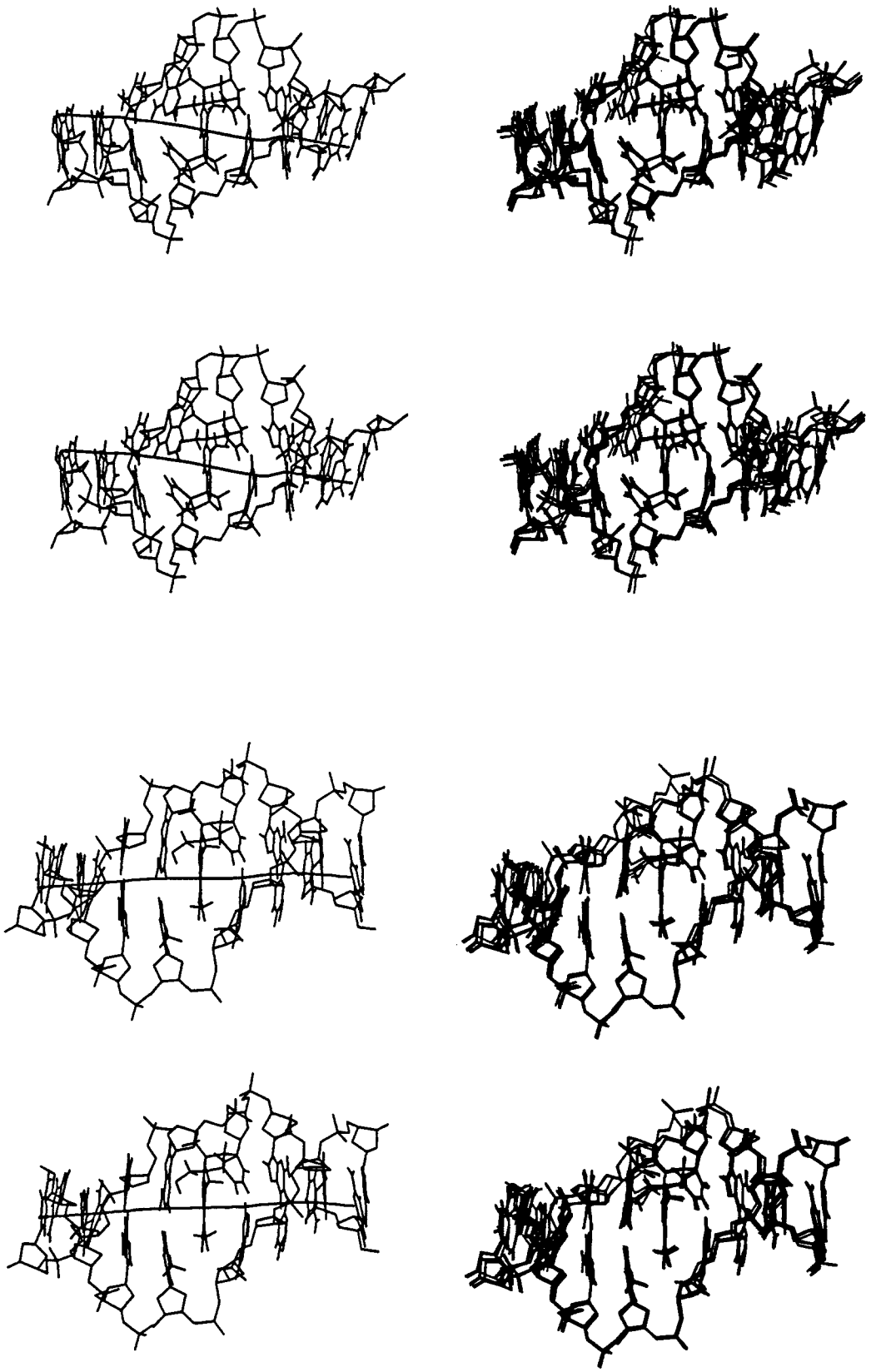

FIGURE 4: Stereoview of the stick plot of the structure of the (A, left) MAf and (B, right) XL. Deoxyribose protons have been omitted for clarity. (Top) One of the calculated structures showing the calculated global helix axis. (Bottom) Superposition of 10 of the 40 structures obtained by RMD calculations.

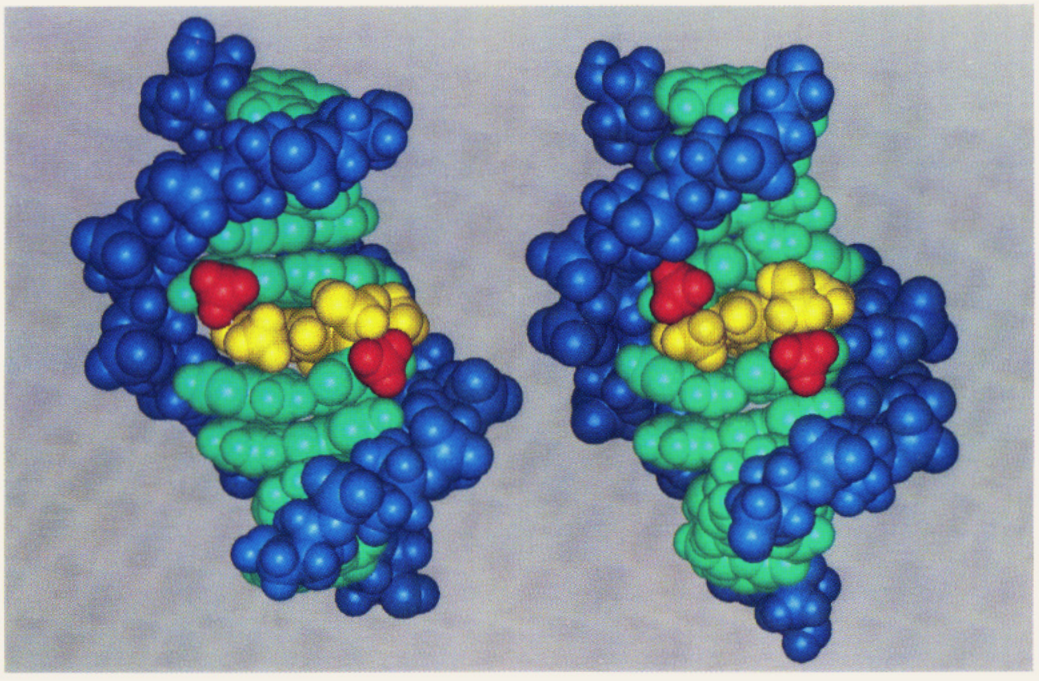

FIGURE 5: Space-filling models of the structure of the (A, left) MAf and (B, right) XL viewed into the major groove showing the solvent-accessible edge of the HMT between the T-A base pairs. HMT is in yellow, and the thymine methyl groups are in red. The local overwinding of the helix at the HMT in the XL is evident. The T4 and T12 5-methyl groups are conspicuously closer together in the XL than in the MAf. It is also evident that the psoralen has been pulled across the helix toward T12 in the XL.

is good van der Waals contact between the 4-methyl group of the HMT with the 5-methyl group of the thymidine.

(2) XL. As a consequence of XL formation, A5 is displaced relative to its position in the MAf (Figure 4B). In the XL, the six-membered ring of A5 still makes contact with the 8-methyl group, but the 6-amino group is now positioned over the furan ring of the HMT. There is nearly complete overlap of the six-membered ring of A13 and the pyrone ring of the HMT in the MAf. This stacking arrangement is retained upon conversion of the MAf to the XL. T12 is in contact with the 4-methyl and positioned over the H3 substituent of the HMT pyrone ring in the MAf. T12 is located directly over the pyrone ring when the MAf is converted to XL.

The stacking between T4 and G3 is perturbed in the MAf and XL by the upward protrusion of the adducted base into G3. In addition, the interaction between G11 and T12 in the XL is distorted. Cyclobutane ring formation changes the hybridization of the thymidine 5 and 6 carbons from sp^2 to sp^3, forcing the 5-methyl group (5-Me) and 6-H of the adducted thymidine away from the HMT with respect to the pyrimidine ring. The adjacent guanine stacks onto the surface formed by the adducted thymidine 5-Me, 6-H, and pyrimidine ring that is much less tilted than the pyrimidine ring alone.

There is a 1 μs relaxation delay between the absorption of the photon that induces monoadduct formation and the ability of the MAf to then undergo further chemistry to form the cross-link (Johnston et al., 1977). Absorption of a photon by the MAf chromophore is not prohibited; rather, it is necessary for the system to complete a conformational change caused by the MAf cyclobutane ring formation before the reaction can occur. A significant local reorganization in the conformation of the DNA is required to accommodate the conversion of the MAf to the XL (Spielmann et al., 1995a). The MAf structure presented here is presumed to be the post-1 μs relaxed structure that is competent to absorb a photon and form a cross-link. Indeed, irradiation of the MAf heteroduplex with 366 nm light results in quantitative transformation to the cross-link (Spielmann et al., 1992). The 3.47 Å separation of T12 C-5 and MAf C-4 and the 3.35 Å separation of T12 C-6 and MAf C-3 change to 1.54 Å each upon XL formation. The MAf−T4 adduct moves 1.5 Å into the major groove relative to the helix axis. T12 rolls down toward the HMT when the pyroneside cyclobutane ring is formed, but it is not displaced relative to the helix axis. The HMT shifts into the major groove and across the helix axis toward T12 upon cross-link formation. A5 moves 1.0 Å toward the major groove upon cross-link formation and rolls slightly into the minor groove because of a steric clash between the exocyclic amino group and the 4′-hydroxymethyl group of the HMT. A13 moves slightly into the minor groove when the 4-methyl group of the HMT clashes with the five-membered ring of A13 upon cross-link formation.

Surface Accessibility of the Adducts. A space-filling model of the structure of the MAf is shown in Figures 5A and 6A. The view of the major groove (Figure 5A) shows the solvent-accessible edge of the HMT between the T-A base pairs. The 5′-methyl group is unobscured by any of the functional groups on the DNA. The 4′-hydroxymethyl group of HMT is clearly in a sterically unhindered position, whereas the 5-position is sterically crowded by the psoralen 4′-hydroxymethyl group and the 4-methyl group, but access is not obscured by the DNA. The 5-position of HMT is directly over the T4−O4 to A13−N6 hydrogen bond. The pyrone ring 4-methyl group shows van der Waals contact to the T12 methyl group. The pyrone H3 of the psoralen is buried in a pocket bounded by the five-membered ring of A13, the T12 base, and the sugars from T12 and A13. Pyridopsoralens have a pyridine ring fused to the 3,4 double bond of the pyrone ring, and these compounds form monoadducts with DNA exclusively (Moysan et al., 1993). The steric requirements for a furanside monoadduct of this type are greater than those of HMT. Based on our model, there would

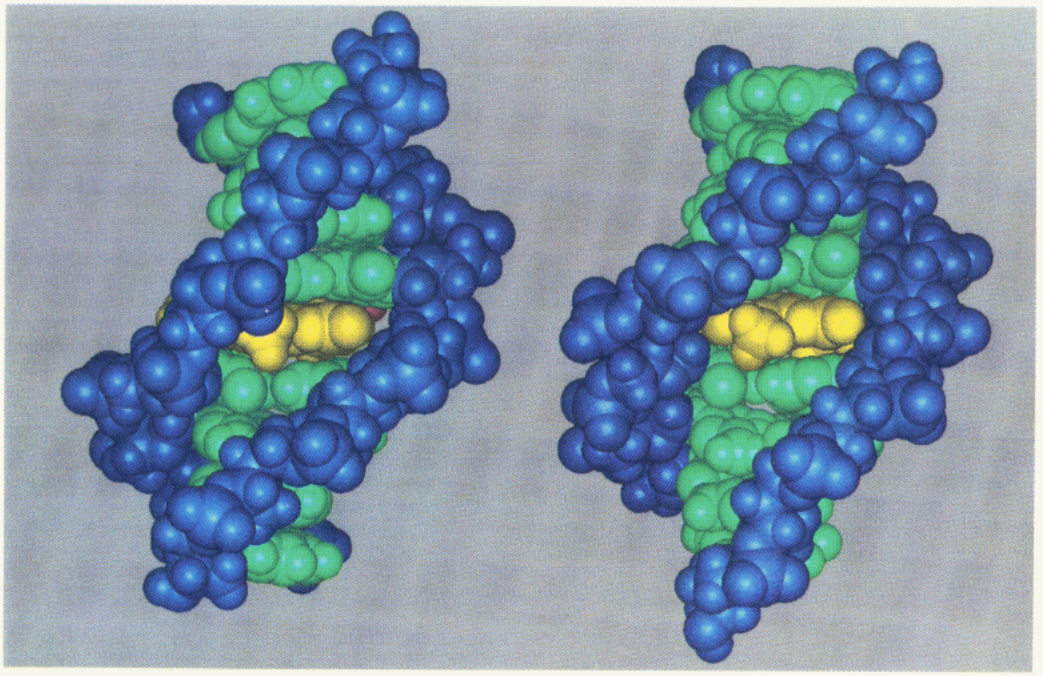

FIGURE 6: Space-filling models of the structure of the (A, left) MAf and (B, right) XL viewed into the minor groove showing the solvent-accessible edge of HMT between the T-A base pairs. HMT is yellow.

have to be a significant reorganization of the MAf structure for the HMT adduct to accommodate such a large substituent at the 3-position. DNA is flexible and has increased dynamics at the site of the photoadduct formation.

The view of the minor groove (Figure 6A) shows the edge of the HMT between the T-A base pairs. The 8-methyl group makes a close van der Waals contact with the O4′ of A5. The ester oxygens of the pyrone ring are also solvent accessible. The view of the minor groove shows how the 8-position of the HMT, the pyrone ring carbonyl, and O1 make up part of the floor of the minor groove. It also shows that access to the furan ring in the minor groove is obscured by the T4 and A5 sugar residues. The buckle introduced into the T4-A13 base pair by the photoaddition of T4 to the HMT is quite clear. Unwinding of the DNA helix at the TpA step and an elongation of the distance between the adenines from 3.3 to 6 Å to accommodate the intercalated adduct are apparent.

Figures 5B and 6B show similar space-filling models of the structure of the XL. The solvent-accessible functional groups in both the major and minor groove are qualitatively the same as those in the MAf. In the major groove view of the XL (Figure 5B) the overwinding of the helix in the vicinity of the HMT is evident. The T4 5-methyl and T12 5-methyl groups are conspicuously closer together in the XL than they are in the MAf. It is also evident that the psoralen has been pulled across the helix toward T12 in the XL. The view of the minor groove (Figure 6B) shows the edge of the HMT between the TA base pairs. The 8-methyl group is no longer in van der Waals contact with the O4′ of A5 because the psoralen has been shifted toward T12 and displaced toward the major groove.

A large number of synthetic psoralens have been produced that bear substituents on the 4′-, 5-, and 8-positions and that retain the ability to form both monoadducts and crosslinks (Cimino et al., 1985; Bisagni, 1992). The structures in Figures 5 and 6 show that functionalization of the 4′- and 5-positions does not introduce unfavorable steric interactions between the psoralen and the surrounding DNA in the monoadduct. Based on the MAf and XL structures, both the 5′- and the 4-positions of the psoralen could be substituted with a group larger than methyl without adverse steric interaction with the surrounding nucleic acid. The 8-methyl group is in the minor groove in a somewhat more sterically congested site.

Helix Parameters. There are three major categories of helical parameters: axis−base pair, intra-base pair, and inter-base pair parameters (Dickerson et al., 1989). These parameters are discussed for the MAf and XL with respect to the calculated global helix axis below. However, the parameters reported for the terminal base pairs of the MAf and XL are probably less reliable because of dynamic processes and a lower density of restraints and are included only for completeness. The DNA returns to B-form geometry three base pairs away from the site of damage. Adjustments in the helix geometry of the three base pairs on either side of HMT compensate for the distortions introduced by the psoralen damage in the MAf and XL. These perturbations away from canonical B-form DNA are described by the helix parameters shown in Figure 7.

Global Helix Axis. The global helix axis is straight in the XL, and the MAf shows only a small 8° bend (Figure 4). The bend in the helix axis was calculated by taking the inner product of the vectors describing the axis segments between the C2-G15−G3-C14 and the C6-G11−G7-C10 base pair steps. The final conformation of the terminal base pairs was considered to be suspect due to their unresolved dynamic motions. The MAf helix axis can be divided into three segments, the first running through the first four base pairs, the second traversing the T4-A13−A5-T12 base pair step, and the third running through the last four base pairs of the octamer. The first and third segments are almost parallel with each other, while the second describes a slip dislocation of the two helical segments caused by the HMT photoadduct. The dislocation in the helix axis in the MAf is probably due to the favorable stacking of the A5 onto the 8-methyl group of the HMT and the steric interaction of the 4′-hydroxymethyl group with A5.

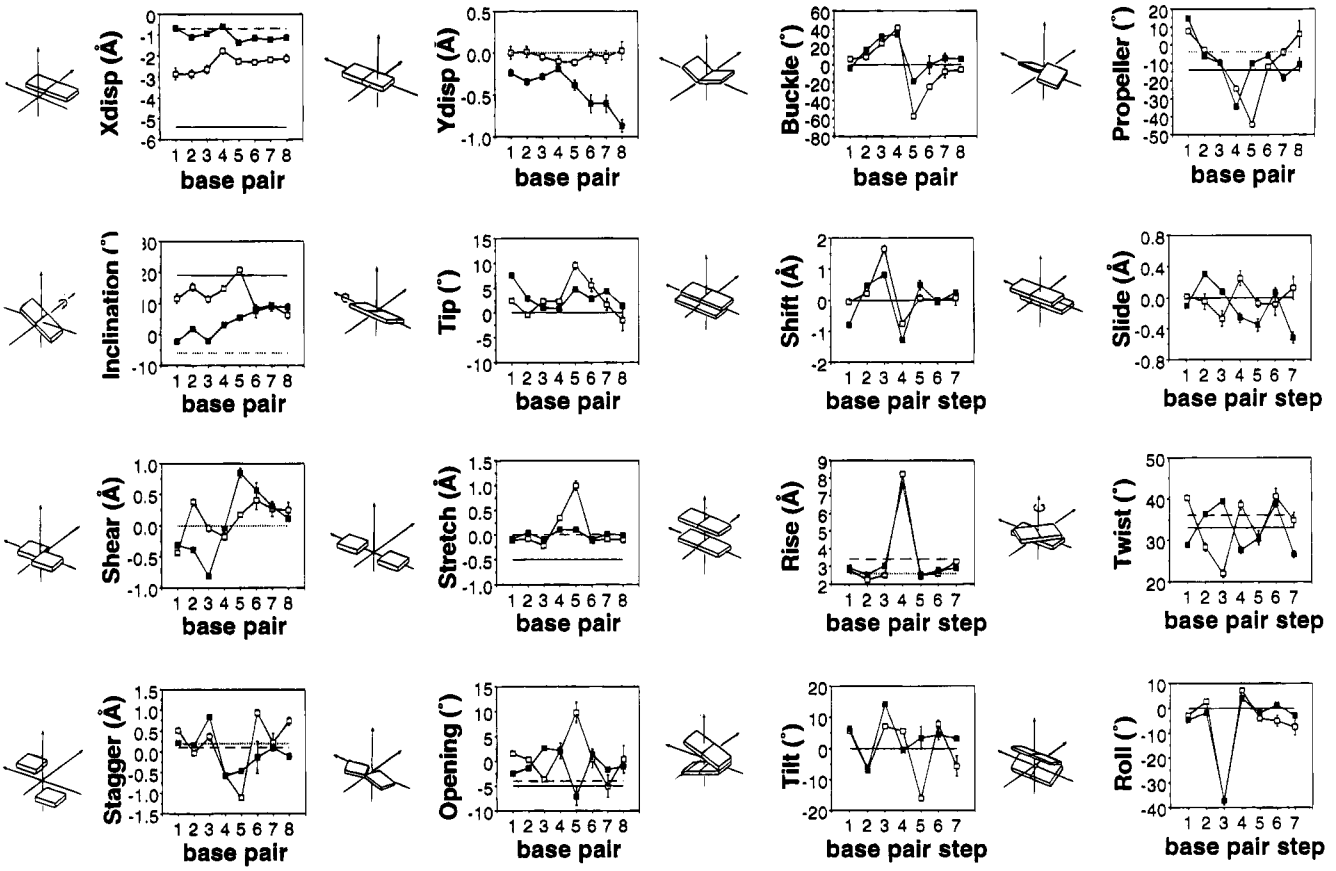

FIGURE 7: Helical parameters for the MAf (solid squares) and XL (open squares) calculated using CURVES and compared with canonical A-DNA (•••) and B-DNA (- - -).

The MAf helix axis is not kinked by the adduct because A5 and A13 remain coplanar with the HMT. G3 of the MAf is rolled slightly into the major groove contributing to the small bend into the major groove. We believe that bends of this magnitude cannot be accurately determined with current NMR methods and are difficult to measure with any technique. The presence or absence of a small bend induced by the MAf might be more reliably measured by other techniques, such as gel mobility assays, that are beyond the scope of the present work. The formation of the XL introduces a large buckle into the T4-A13 and T12-A5 base pairs. Cross-link formation can be accommodated without kinking because A5 and A13 remain coplanar with the HMT. G3 stacks onto the 5-Me and modified ring of T4, and G11 stacks onto the 5-Me and modified ring of T12 coplanar with the HMT in the XL (Figure 4). The unbent axis of the XL is in agreement with gel mobility shift data (Sinden & Hagerman, 1984; Haran & Crothers, 1988).

Axis−Base Pair Parameters. *X*-axis displacement, *Y*-axis displacement, inclination, and tip describe how the base pairs are positioned in the helix relative to the global helix axis. Formation of the XL causes a contraction across the major groove and widens and makes shallower the minor groove at the site of adduct formation. The distance in the XL between the C5 carbons of T4 and T12 is 6.6 Å, while in normal A-form DNA this distance is 9.7 Å. In addition, the C5 atoms in the XL are separated by 3 Å measured along the helix axis, comparable to 2.6 Å in A-form and 3.4 Å in B-form DNA due to the two cyclobutane rings. A-form DNA has a deep major groove and a wide, shallow minor groove, while B-form DNA has a relatively narrow deep minor groove and a wide major groove. This results in the appearance of A-form-like helix parameters for the XL and to a lesser extent MAf as the DNA alters its conformation to accommodate the HMT photoadducts.

Y-Axis Displacement. The *Y*-axis displacement in the cross-link is that found in the normal B- and A-form DNA. There are significant negative *Y*-axis displacements in the MAf. The global axis is displaced toward the adducted strand and becomes increasingly more negative 3′ to the adduct. The HMT−T monoadduct does not introduce the same amount of A-form-like character as does the T−HMT−T cross-link into the helix. The A-form-like character is asymmetrically distributed between the strands in the MAf, with the HMT damaged strand being more A-form like and the unmodified strand more B-form like.

Intra-Base Pair Parameters. The calculated parameters for shear, stretch, stagger, and opening are influenced but not completely determined by the restraints used to enforce base pair hydrogen bonding. For the base pairs in and around the photodamage site, these parameters are determined by the HMT−DNA base NOEs because no hydrogen-bonding restraints were included for the thymidine−adenosine base pairs in the RMD procedure.

Shear. For the MAf, the shear changes from positive to negative across the TpA step, compensating for the intercalation of the HMT and returning to essentially B-form values at the ends of the helix. There appears to be no net shear introduced into the MAf-adducted helix. This was also observed in the structure of the bisintercalator TOTO−DNA complex (Spielmann et al., 1995c). The covalent geometry imposed on the cross-linking site in the XL does not show this behavior. The values start out negative and only show a small negative shear immediately 5′ to the intercalation site. The hydrogen-bonding edges of the two adducted thymidines are rotated relative to each other into the major

groove. Large negative shear at the G3-C14 and positive shear at the A5-T12 base pairs in the XL would severely disrupt the hydrogen-bonding interactions between these base pairs.

Stretch. The stretch in the MAf is essentially that of B-form DNA, indicating that normal hydrogen-bonding interactions are present for the two A-T base pairs. No explicit hydrogen-bonding restraints for the A-T base pairs were used in the refinement. This is in agreement with the temperature dependence of observed imino proton exchange behavior. The XL shows a large positive stretch for the A5-T12 base pair and a smaller positive stretch for the T4-A13 base pair. This is a result of the rotation of the hydrogen-bonding edge of T12 into the major groove. The T12 H3 proton is much more solvent accessible in this conformation, in agreement with the temperature dependence of observed imino proton exchange (Patel et al., 1982).

Stagger. Structurally, the changes in stagger away from that of normal DNA appear to be due to the stacking of the flanking guanine residues onto the photodamaged thymidines. Although the guanine stacks onto the surface formed by the folded thymidine, this surface is not exactly coplanar with the adenine base paired with the thymidine. This offset is manifested by a compensatory change in stagger when going from the damaged base to the base 5′ to the adduct.

Buckle. The large deviations in the buckle from the normal A- and B-DNA value of 0° are a direct consequence of the intercalation and photoadduction of the HMT. Unlike the perpendicular intercalators where the DNA buckles to maximize van der Waals contacts with the chromophores (Williams et al., 1992), the A5 and A13 in the adduct site stack coplanar with the HMT and are normal to the helix axis, maximizing van der Waals contacts with the chromophore. The buckling is primarily a consequence of the covalent nature of the thymidine adduct geometry. Because the adjacent guanine bases stack on the "platform" of the adducted thymidines, there is no propensity for the adduct to introduce a bend into the helix or add additional buckle to the binding site (Spielmann et al., 1995a). The binding pocket for the XL adduct is defined by the surfaces of G3, A5, G11, and A13, with the two thymidines being an integral part of the intercalating agent. The symmetry of the distortions introduced in the DNA by MAf and XL formation is evident in the relative sign and distribution of the buckle across the HMT photobinding site. It is also evident that the effect of the adduct formation on the buckle rapidly diminishes a few base pairs distal to the site of damage in both the MAf and XL. The buckle in T12 of the MAf appears to be a direct effect of stacking of that base with the HMT. The buckling is a mechanism by which the unadducted thymidine increases its van der Waals interactions with the intercalated monoadduct which does not reach all of the way across the helix (Williams & Gao, 1992). Buckling has been observed in the DNA bases of perpendicular intercalator−DNA complexes by both X-ray crystal structures and NMR solution structures (Wang et al., 1984; Addess et al., 1993; Spielmann et al., 1995c). The HMT monoadduct is similar to a parallel intercalator such as ethidium, but the MAf does not reach all the way across the helix. The T12-A5 base pair in the MAf buckles to create a cavity that optimizes the interactions between the psoralen and the surrounding DNA. Although the plane of the pyrimidine ring of T4 is at an angle of between 44° and 53° with respect to the helix axis, the base 5 and 6 carbons have converted from planar sp^2 hybrids to strained and tetrahedral sp^3 centers in the cyclobutane ring. This forces the T4 5-methyl group and the T4 H6 up into G3. To maintain stacking, G3 H8 makes a contact with the "folded thymidine" T4 at the 5-methyl, O2, and N3H. In doing so, the effect of the dramatic thymidine buckle on the helix axis is greatly diminished.

Propeller Twist. A large propeller twist is introduced into the T4-A13 base pair in the MAf because the cyclobutane ring formation twists T4 relative to the helix axis and A5 and A13 remain essentially coplanar with the HMT tricyclic nucleus. This effect is duplicated for the T4-A13 and A5-T12 base pairs of the XL. The effect of the distorted thymidines on the propeller twist is quite localized, with the remaining base pairs in the helixes having essentially normal values.

Inter-Base Pair Parameters. The roll, slide, and shift in regular A-DNA and B-DNA conformations are essentially zero. These parameters show large and correlated changes from the equilibrium values in the MAf and XL structures.

Slide. The slide of the T4-A5 base pair step changes sign in response to the formation of the interstrand cross-link, reflecting the contraction of the two strands toward each other. This change in the sign and magnitude of the slide induced by the conversion of MAf to XL seems to be propagated to each base pair step of the two molecules.

Rise. The rise values for corresponding base pair steps in the MAf and XL almost exactly mirror each other. The rise for all of the base pair steps except for the 5′-TpA-3′ is close to that of A-form DNA. The T4-A5 base pair step has a rise of 7.6 Å in the MAf and 8.4 Å in the XL, reflecting the intercalation and structural deformation introduced by the HMT photoadduct. These values are somewhat larger than that induced by simple intercalation. This is due to the cyclobutane ring formation which pushes the adducted thymidines away from coplanarity with the rest of the bases in the helix.

Twist. The helix is unwound by 34° in the MAf and by 25° in the XL. This agrees well with the unwinding angle for the XL of 28 ± 4° determined by a DNA supercoiling assay (Wiesehahn & Hearst, 1978). The helix has an overall helical repeat of 11 base pairs in both the MAf and the XL, consistent with intercalation of the HMT. The distribution of the unwinding along the molecule is different in the MAf and the XL. The MAf helix is underwound at the T4-A5 base step as a result of the insertion of the psoralen. The underwinding in the MAf is distributed over the T4-A5 and A5-C6 steps. The TA base step in the XL is overwound to 39° due to the covalent geometry imposed on the DNA by the HMT cross-link. The overwinding at the XL T4-A5 step is compensated for by the under-winding of the C2-G3, G3-T4 and A5-C6 steps. The destabilization of the helix on the pyroneside of the adduct is due, in part, to the change in helix twist at the TpA step brought about by the covalent connection of the two DNA strands.

Roll. The roll values for corresponding base pair steps in the MAf and XL almost exactly mirror each other. Positive roll opens the angle between base pairs toward the minor groove, resulting in a bend into the major groove, while a negative roll results in a bend into the minor groove. Except for the G3-T4 base pair step, there are only small deviations in the roll for both molecules away from canonical A- and B-form DNA values. The large −37° roll found for the G3-T4 base pair step in both the MAf and the XL is a result of

Table 6: Backbone Angles (deg) for the HMT MAf and XL of d(GCGTACGC)$_2$

	α		β		γ		δ		ε		ζ		χ	
residue	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf	XL	MAf
G1	285 ± 1	302	177 ± 3	189			145 ± 1	145 ± 3	186 ± 1	184	263 ± 1	262	243 ± 1	223 ± 1
C2	294 ± 1	296	174 ± 1	197	59 ± 2	50	109 ± 2	141	183 ± 1	173	280 ± 1	269	245 ± 1	263
G3	270 ± 1	295 ± 1	50 ± 1	153 ± 1	40 ± 2	39	135	141	233 ± 3	182 ± 1	284	259 ± 1	260 ± 1	274
T4	59 ± 2	59 ± 5	240 ± 1	157 ± 2	176	78 ± 1	123 ± 1	91 ± 1	290 ± 1	72 ± 1	248 ± 1	79	214 ± 1	236 ± 2
A5	75 ± 2	113 ± 98	178	139 ± 34	231 ± 1	186 ± 1	109	145 ± 3	60 ± 1	65 ± 46	62 ± 2	101 ± 68	228	263 ± 3
C6	289 ± 2	294 ± 2	184 ± 2	184 ± 2	182 ± 1	143 ± 61	135 ± 1	152 ± 17	185	173 ± 3	265 ± 1	269 ± 1	237 ± 1	261 ± 11
G7	304 ± 1	278	190	170 ± 1	40 ± 2	56 ± 1	128 ± 2	140 ± 1	170 ± 2	184	255	265 ± 1	251 ± 2	271 ± 1
C8					46 ± 1	50	148 ± 1	141 ± 1					242 ± 1	264
G9	305 ± 30	295	228 ± 65	166 ± 1			151 ± 7	147 ± 2	250 ± 59	185	247 ± 23	275 ± 1	245 ± 4	248 ± 2
C10	303 ± 3	289 ± 2	196	218 ± 2	128 ± 69	52 ± 1	104 ± 7	125 ± 1	182 ± 5	177 ± 1	276 ± 1	264	233 ± 2	243 ± 1
G11	286 ± 1	274 ± 2	170	169 ± 1	43	32 ± 1	145 ± 1	147	184 ± 1	183	273 ± 1	269	258 ± 1	280 ± 1
T12	62 ± 2	297	187 ± 3	154 ± 1	59 ± 1	45	97 ± 5	84	71 ± 1	169 ± 1	78 ± 1	300	262 ± 1	239 ± 1
A13	222 ± 82	53	100 ± 69	187	186 ± 3	134 ± 1	120 ± 10	146	201 ± 47	202	259 ± 60	244	244 ± 5	237
C14	273 ± 8	290	188 ± 3	186	173 ± 50	258	147 ± 14	147	190 ± 4	178	276 ± 3	263	231 ± 7	239
G15	309 ± 1	279	179 ± 2	173	33 ± 1	59	140	140	177 ± 1	180	237 ± 2	265	260 ± 1	268
C16					50	50	148	141 ± 1					240 ± 1	262 ± 1
B-form	314		214		36		156		155		264		262	
A-form	285		208		45		83		178		313		206	

the G3 stacking onto the platform of T4 and not on the highly propeller twisted and buckled six-membered ring of the adducted thymidine. This effect is not observed for the pyroneside adduct of the XL because the hydrogen-bonding edge of T12 is rotated into the major groove by the formation of the pyroneside cyclobutane ring. There is no overall bend introduced into the helixes of these adducted molecules because the roll and the propeller twist compensate for each other.

Backbone Angles. The backbone angles for both the MAf and the XL are reported in Table 6. Most of the backbone angles are not directly constrained by the NOESY data. There is also evidence of conformational heterogeneity in the sugar residues in and around the adduct site. The calculated values represent averages of different conformations and are therefore suspect and only reported for completeness. However, qualitative trends can be extracted from the data. The adjustments that the helix makes to accommodate the adducts must be absorbed by the backbone in the region of the damage. The sugar pucker alone cannot take care of it all; therefore we see large changes in the equilibrium values for the backbone torsion angles that correlate with the damage. We do not have direct evidence for these alterations in geometry, but during RMD, they do sample other conformations and settle to these reported values upon annealing and minimization.

Comparison with Previous Cross-Link Models. There has been much interest in the structure of the interstrand cross-link. Molecular modeling based on the structure of the isolated thymidine−psoralen−thymidine adduct suggested that the DNA helix axis was bent by as much as 70° into the major groove (Peckler et al., 1982; Kim et al., 1983; Pearlman et al., 1985). The computer model was based on the assumption that the helix on either side of the adduct propagated normal to the plane of the ring of the adducted thymidines, not from the surface formed by the 5-methyl, O2, and N3H of the folded thymidine. These assumptions were carried into the NMR-derived model of the AMT cross-link of d(GGGTACCC)$_2$ (Tomic et al., 1987), resulting in a reported bend of 53° into the major groove with a 56° unwinding of the helix for that system. These results contradict the unwinding angle of 28 ± 4° found in a supercoiling unwinding assay (Wiesehahn & Hearst, 1978) and observation that psoralen cross-links did not alter the mobility of DNA fragments in a gel shift assay, indicating that psoralen cross-links do not induce a bend into the DNA helix axis (Sinden & Hagerman, 1984; Haran & Crothers, 1988).

There are at least two technical factors which contributed to the problems with the AMT XL structure that are unrelated to the differences in the psoralen and DNA sequence. First, the contribution of spin diffusion to the NOE intensities was not accounted for in the AMT XL system. In the present work we have used the RANDMARDI modification of the relaxation matrix analysis program MARDIGRAS to obtain accurate interproton distances from the NOE data. The current calculations show that there are important spin diffusion pathways in both the MAf and the XL molecules. Significant bending of the helix and highly strained, unrealistic distortions of the psoralen and thymidine bases were observed in initial models of the HMT XL and HMT MAf generated from experimental NOE data using distance restraints derived by the isolated spin-pair approximation (ISPA). Indirect magnetization transfer makes significant contributions to the peak intensities in these psoralen−DNA complexes. Second, modern simulated annealing (SA) and restrained molecular dynamics (RMD) protocols were unavailable when the AMT XL structure was reported. The previously reported structure was refined from a bent starting model, introducing a conformational bias. Attempts to refine the AMT XL structure from a straight B-form DNA starting model without SA or RMD protocols led to strained conformations with high energies and significant distance restraint violations. The modern algorithms used in the present work are much more efficient at overcoming local energy barriers encountered during the structure refinement. Bent DNA models derived from NMR data can be artifacts due to the higher density of NOEs in the major groove and the anisotropic character of DNA. Incorrect NOE-derived distance restraints can be partially satisfied by bending the DNA.

Implications for DNA Repair. The MAf and XL are recognized by the (A)BC excinuclease and serve as excellent models to study the structural and dynamic motifs that cellular repair enzyme systems may recognize (Sancar & Sancar, 1988; Cheng et al., 1988b, 1991). The structural differences between the MAf and the XL allow examination of common structural or dynamic features that may be recognized by the enzyme system. Understanding how the (A)BC excinuclease recognizes psoralen damage to DNA

will assist in the elucidation of the mechanism of lesion recognition. Excision repair systems efficiently recognize and remove many structurally diverse DNA damaging agents while the wide range of normal DNA microstructures triggers a response that is 3−5 orders of magnitude lower in rate (Sancar & Sancar, 1988; Lin & Sancar, 1992; Huang et al., 1994). Recognition of damaged DNA by the (A)BC excinuclease repair system is initiated by the binding of the $UvrA_2B_1$ complex to the damage site. The $UvrA_2B_1$ protein complex, utilizing energy provided by its ATPase activity, sharply bends the DNA into the major groove at the 11th phosphodiester bond 5′ to the lesion, leaving the UvrB component tightly bound to the DNA (Shi et al., 1992). The "molecular matchmaker" UvrA then dissociates from the complex (Sancar & Hearst, 1993). The matchmaking mechanism uncouples the damage recognition event by UvrA from the subsequent excision steps by UvrB and UvrC. In the absence of UvrA, UvrB appears to have no affinity for DNA. Recognition of DNA damage by these systems has been proposed to take place by the detection of structural changes and/or the alteration of the dynamic behavior of the DNA (Sancar & Sancar, 1988; Lin & Sancar, 1989; Pu et al., 1989). It has been proposed that damage causes local instability in the helix and that the (A)BC excision repair system recognizes this excessive or unusual deformability (Williams & Gao, 1992). The conformations of the UM sugars have been compared with those of MAf and XL, and the HMT damage destabilizes the preference for $C_{2'}$-*endo* sugar conformation of the four nucleotides T4, A5, T12, and A13. The MAf or XL permits them to adopt a larger range of rapidly interconverting conformers with a significantly altered population distribution with respect to unmodified DNA (Emsley et al., 1993; Spielmann et al., 1995a). These changes in the equilibrium distribution of the deoxyribose conformations and the implied coupled "crankshaft motions" of the phosphate backbone detected for the MAf and XL may be recognized as the signal that positions the DNA repair system at the location of the DNA damage.

The DNA repair system in this model recognizes changes in the equilibrium behavior of the DNA surrounding the adduct, not specific structural features induced by the damage (Lin & Sancar, 1989; Pu et al., 1989). Other DNA oligomers covalently modified with drugs that are recognized by cellular repair systems and have radically different structures from the MAf and the XL show evidence of altered helix dynamics by NMR (O'Handley et al., 1993). We hypothesize that the repair enzymes target lesions by detecting the conformational flexibility of the sugar−phosphate backbone induced by DNA damaging agents. There are numerous amino acid contacts to the backbone phosphates observed in high-resolution X-ray crystal structures of protein−DNA complexes (Beamer & Pabo, 1992). The flexibility of the phosphate esters in the recognition and formation of protein−DNA contacts has been proposed as a mechanism for discrimination of different operator sites by the lac repressor (Karslake et al., 1992; Botuyan et al., 1993). Flexibility of the backbone is a more general case of localized sequence-specific conformational variations in DNA that are potentially recognized by proteins. We propose that the backbone of undamaged DNA has a certain stiffness (firmness) that is ignored by the $UvrA_2B_1$ system and that introduction of a region of greater conformational freedom (softness) into the DNA backbone at the adduct is recognized as damage by the $UvrA_2B_1$ complex. Determination of the precise backbone dynamic modes that are changed by lesion formation will require other experiments such as the measurement of order parameters and further analysis of homonuclear and heteronuclear coupling constants for the sugar residues by NMR.

CONCLUSION

We have presented clear evidence of a single HMT adduct intercalated in the central 5′-TpA-3′ site in the MAf and XL. We have shown that neither the MAf nor the XL show unusual bending as a result of adduct formation. The RANDMARDI procedure yields structures of molecules in solution that are more consistent with magnetization transfer theory and agree better with results from other physical techniques. A better understanding of the fundamental processes that allow photobinding of organic molecules to DNA will aid in the design of new and improved nucleic acid specific agents that interfere with DNA metabolism. High-resolution structures and dynamics information of the various reaction products between psoralen and nucleic acids are required to understand the detailed mechanism of the photoreactions and the biological function of the adducts. The ability of psoralen adducts to stop replication, transcription, and other DNA metabolic activities such as the activity of restriction endonucleases is the same activity that is exploited by a number of chemotherapeutic agents.

ACKNOWLEDGMENT

We thank Professor T. L. James and Dr. H. Liu, University of California, San Francisco, for providing the MARDIGRAS and the RANDMARDI programs and for helpful guidelines.

REFERENCES

Addess, K. J., Sinsheimer, J. S., & Feigon, J. (1993) *Biochemistry 32*, 2498−2508.
Assa-Munt, N., Leupin, W., Denny, W. A., & Kearns, D. R. (1985a) *Biochemistry 24*, 1441−1449.
Assa-Munt, N., Leupin, W., Denny, W. A., & Kearns, D. R. (1985b) *Biochemistry 24*, 1449−1460.
Beamer, L. J., & Pabo, C. O. (1992) *J. Mol. Biol. 227*, 177−196.
Bisagni, E. (1992) *J. Photochem. Photobiol. B 14*, 23−46.
Borgias, B. A., & James, T. L. (1988) *J. Magn. Reson. 79*, 493.
Borgias, M., Gochin, M., Kerwood, D. J., & James, T. L. (1990) *Prog. NMR Spectrosc. 22*, 83.
Botuyan, M. V., Keire, D. A., Kroen, C., & Gorenstein, D. G. (1993) *Biochemistry 32*, 6863−6874.
Cassier, C., Chanet, R., & Moustacchi, E. (1984) *Photochem. Photobiol. 39*, 799−803.
Cheng, S., Van Houten, B., Gamper, H., Sancar, A., & Hearst, J. E. (1988a) Use of Triple-Stranded DNA Complexes to Study Crosslink Repair, in *Mechanisms and Consequences of DNA Damage Processing*, pp 105−113, Alan R. Liss, New York.
Cheng, S., Van Houten, B., Gamper, H., Sancar, A., & Hearst, J. E. (1988b) *Journal of Biological Chemistry 263*, 15110−15117.
Cheng, S., Sancar, A., & Hearst, J. E. (1991) *Nucleic Acids Research 19*, 657−663.
Cimino, G. D., Gamper, H., Isaacs, S., & Hearst, J. E. (1985) *Annu. Rev. Biochem. 54*, 1151−1193.
Cimino, G. D., Shi, Y., & Hearst, J. E. (1986) *Biochemistry 25*, 3013−3020.
Corash, L., Lin, L., & Wiesehahn, G. (1992) *Blood Cells 18*, 57−74.
del Olmo, M. L., Sogo, J. M., Franco, L., & Perez-Ortin, J. E. (1993) *Yeast 9*, 1229−40.
Dickerson, R. E., Bansal, M., Calladine, C. R., Diekermann, S., Hunter, W. N., Kennard, O., Lavery, R., Nelson, H. J. C., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., vonKitzing, E., Wang, A.-H., & Zhurkin, V. B. (1989) *EMBO J. 8*, 1−4.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D., & Wemmer, D. E. (1979) *Faraday Div. Chem. Soc. Symp. 13*, 49.
Emsley, L., Dwyer, T. J., Spielmann, H. P., & Wemmer, D. E. (1993) *J. Am. Chem. Soc. 115*, 7765−7771.
Ernst, R. R., Bodenhausen, G., & Wokaun, A. (1987) in *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, England.
Fitzpatrick, J. B., Stern, R. S., & Parrish, J. A. (1982) in *Psoriasis—Proceedings of the Third International Symposium*, pp 149−156, Grune & Stratton, New York.
Fujita, H. (1984) *Photochem. Photobiol. 39*, 835−839.
Gochin, M., Zon, G., & James, T. L. (1990) *Biochemistry. 29*, 11161−11171.
Haran, T. E., & Crothers, D. M. (1988) *Biochemistry 27*, 6967−6971.
Hare, D. R., Wemmer, D. E., Chou, S.-H., Drobny, G., & Reid, B. R. (1983) *J. Mol. Biol. 171*, 319−336.
Huang, J. C., & Sancar, A. (1994) *J. Biol. Chem. 269*, 19034−19040.
Islas, A. L., Baker, F. J., & Hanawalt, P. C. (1994) *Biochemistry 33*, 10794−10799.
Johnston, B. H., Johnson, M. A., Moore, C. B., & Hearst, J. E. (1977) *Science 197*, 906−908.
Kanne, D., Straub, K., Hearst, J. E., & Rapoport, H. (1982a) *J. Am. Chem. Soc. 104*, 6754−6764.
Kanne, D., Straub, K., Rapoport, H., & Hearst, J. E. (1982b) *Biochemistry 21*, 861−871.
Kanne, D., Rapoport, H., & Hearst, J. E. (1984) *J. Med. Chem. 27*, 531−534.
Karslake, C., Botuyan, M. V., & Gorenstein, D. G. (1992) *Biochemistry 31*, 1849−1858.
Kim, S. H., Peckler, S., Grave, B., Kanne, D., Rapoport, H., & Hearst, J. E. (1983) *Cold Spring Harbor Symp. Quant. Biol. 47*, 361−365.
Lavery, R., & Sklenar, H. (1988) *J. Biomol. Struct. Dyn. 6*, 63−91.
Lavery, R., & Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655−667.
Lin, J. J., Sancar, A. (1989) *Biochemistry 28*, 7979−7984.
Lin, J. J., Sancar, A. (1992) *Mol. Microbiol. 6*, 2219−2224.
Lin, L., Londe, H., Hanson, C. V., Wiesehahn, G., Issacs, S., Cimino, G., & Corash, L. (1993) *Blood 82*, 292−297.
Misra, R. R., & Vos, J. M. (1993) *Mol. Cell Biol. 13*, 1002−1012.
Morel, P., Lin, L., Wiesehahn, G., & Corash, L. (1992) *Blood Cells 18*, 27−42.
Moysan, A., Voituriez, L., Cadet, J., & Vigny, P. (1993) *J. Photochem. Photobiol. B 17*, 263−271.
O'Handley, S. F., Sanford, D. G., Xu, R., Lester, C. C., Hingerty, B. E., Broyde, S., & Krugh, T. R. (1993) *Biochemistry 32*, 2481−2497.
Patel, D. J., Pardi, A., & Itakura, K. (1982) *Science 216*, 581−590.
Pathak, M. A., & Fitzpatrick, T. B. (1992) *J. Photochem. Photobiol. B 14*, 3−22.
Pearlman, D. A., Holbrook, S. R., Pirkle, D. H., & Kim, S.-H. (1985) *Science 227*, 1304−1308.
Peckler, S., Graves, B., Kanne, D., Rapoport, H., Hearst, J. E., & Kim, S. H. (1982) *J. Mol. Biol. 162*, 157−172.
Pu, W. T., Kahn, R., Munn, M. M., Rupp, W. D. (1989) *J. Biol. Chem, 264*, 20697−20704.
Ramaswamy, M., & Yeung, A. T. (1994) *J. Biol. Chem. 269*, 485−492.
Sage, E., Drobetsky, E. A., & Moustacchi, E. (1993) *EMBO J. 12*, 397−402.
Sancar, A., & Sancar, G. B. (1988) *Annu. Rev. Biochem. 57*, 29−67.
Sancar, A., & Hearst, J. E. (1993) *Science 259*, 1415−1420.
Sastry, S. S., & Hearst, J. E. (1991a) *J. Mol. Biol. 221*, 1091−1110.
Sastry, S. S., & Hearst, J. E. (1991b) *J. Mol. Biol. 221*, 1111−1125.
Sastry, S. S., Spielmann, H. P., Hoang, Q. S., Phillips, A. M., Sancar, A., & Hearst, J. E. (1993) *Biochemistry 32*, 5526−5538.
Schmitz, U., Sethson, I., Egan, W. M., & James, T. L. (1992) *J. Mol. Biol. 227*, 510−531.
Shi, Q., Thresher, R., Sancar, A., & Griffith, J. (1992) *J. Mol. Biol. 226*, 425−432.
Shi, Y., & Hearst, J. E. (1986) *Biochemistry 25*, 5895−5902.
Shi, Y., Spielmann, H. P., & Hearst, J. E. (1988a) *Biochemistry 27*, 5174−5178.
Shi, Y., Gamper, H., Houten, B. V., & Hearst, J. E. (1988b) *J. Mol. Biol. 199*, 277−293.
Shi, Y., Lipson, S., Chi, D. Y., Spielmann, H. P., Monforte, J., & Hearst, J. E. (1990) Applications of Psoralens as Probes of Nucleic Acid Structure and Function, in *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*, 1st ed., pp 341−378, John Wiley and Sons, New York.
Sinden, R. R., & Hagerman, P. J. (1984) *Biochemistry 23*, 6299−6303.
Sladek, F. M., Munn, M. M., Rupp, W. D., & Howard-Flanders, P. (1989) *J. Biol. Chem. 264*, 6755−6765.
Smith, C. A. (1988) Repair of DNA Containing Furocoumarin Adducts, in *Psoralen DNA Photobiology*, pp 87−116, CRC Press, Inc., Boca Raton, FL.
Smith, K. C., Wang, T.-C. V., & Sharma, R. C. (1989) *BioEssays 10*, 12−16.
Spielmann, H. P., Sastry, S. S., & Hearst, J. E. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4514−4518.
Spielmann, H. P., Dwyer, T. J., Sastry, S. S., Hearst, J. E., & Wemmer, D. E. (1995a) *Proc. Natl. Acad. Sci. U.S.A. 92*, 2345−2349.
Spielmann, H. P., Chi, D.-Y., Hunt, N. G., Klein, M. P., & Hearst, J. E. (1995b) *Biochemistry* (in press).
Spielmann, H. P., Wemmer, D. E., & Jacobsen, J. P. (1995c) *Biochemistry 34*, 8542−8553.
Straub, K., Kanne, D., Hearst, J. E., & Rapoport, H. (1981) *J. Am. Chem. Soc. 103*, 2347−2355.
Taylor, J. S., Garrett, D. S., Brockie, I. R., Svoboda, D. L., & Telser, J. (1990) *Biochemistry 29*, 8858−8866.
Tomic, M., Wemmer, D., & Kim, S. H. (1987) *Science 238*, 1722−1725.
Van Houten, B., Gamper, H., Holbrook, S., Sancar, A., & Hearst, J. E. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8077−8081.
Van Houten, B., Gamper, H., Sancar, A., & Hearst, J. E. (1988) *J. Biol. Chem. 263*, 16553−16560.
Vigny, P., Blais, J., Ibanez, V., & Geacintov, N. E. (1987) *Photochem. Photobiol. 45*, 601−607.
Wang, A.-H., Ughetto, G., Quigley, G. J., Hakoshima, T., van der Marel, G. A., van Boom, J. H., & Rich, A. (1984) *Science 225*, 1115−1121.
Wiesehahn, G. P., & Hearst, J. E. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 2703−2707.
Williams, L. D., & Gao, Q. (1992) *Biochemistry 31*, 4315−4324.
Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Yang, S. C., Lin, J. G., Chiou, C. C., Chen, L. Y., & Yang, J. L. (1994) *Carcinogenesis 15*, 201−207.
Yeung, A. T., Jones, B., & Chu, C. T. (1988) *Biochemistry 27*, 3204−3210.

BI950950A