1554–1560 *Nucleic Acids Research, 1996, Vol. 24, No. 8*

The structure of d(TpA)*, the major photoproduct of thymidylyl-(3′-5′)-deoxyadenosine

Xiaodong Zhao, Sourena Nadji, Jeffrey L.-F. Kao[1] and John-Stephen Taylor*

Department of Chemistry and [1]Chemistry Department High Resolution NMR Facility, Washington University, St Louis, MO 63130, USA

Received October 3, 1995; Revised and Accepted March 5, 1996

ABSTRACT

Irradiation of the dinucleotide TpdA and TA-containing oligonucleotides and DNA produces the TA* photoproduct which was proposed to be the [2+2] cycloaddition adduct between the C5–C6 double bonds of the T and the A [Bose,S.N., Kumar,S., Davies,R.J.H., Sethi,S.K. and McCloskey,J.A. (1984) *Nucleic Acids Res.* 12, 7929–7947]. The proposed structure was based on a variety of spectroscopic and chemical degradation studies, and the assignment of a trans–syn-I stereochemistry was based on an extensive ^{1}H-NMR and molecular modeling study of the dinucleotide adduct [Koning,T.M.G., Davies,R.J.H. and Kaptein,R. (1990) *Nucleic Acids Res.* 18, 277–284]. However, a number of properties of TA* are not in accord with the originally proposed structure, and prompted a re-evaluation of the structure. To assign the ^{13}C spectrum and establish the bond connectivities of the TA* photoproduct of TpdA [d(TpA)*], ^{1}H-^{13}C heteronuclear multiple-quantum coherence (HMQC) and heteronuclear multiple bond correlation (HMBC) spectra were obtained. The ^{13}C shifts and connectivities were found to be inconsistent with the originally proposed cyclobutane ring fusion between the thymine and adenine, but could be explained by a subsequent ring-expansion reaction to give an eight-membered ring valence isomer. The new structure for the d(TpA)* resolves the inconsistencies with the originally proposed structure, and could have a stereochemistry that arises from the *anti*, *anti* glycosyl conformation found in B form DNA.

INTRODUCTION

The TA* photoproduct was first isolated from the irradiation products of thymidylyl-(3′-5′)-deoxyadenosine, d(TpA), and was proposed to be a [2+2] cycloadduct between the 5,6 double bonds of the T and the A (Fig. 1; Structure **1a**) based on spectroscopic and chemical degradation studies (1). Assignment of the trans–syn-I stereochemistry to the proposed cycloadduct was based on the analysis of ^{1}H-NMR coupling constant and NOE data which was then used to constrain molecular dynamics simulations (2). Because of the unique fluorescence properties of the acid hydrolysis product of TA*, compound **2**, (1,3) it was possible to demonstrate the formation of TA* in duplex poly(dA-dT) and calf thymus DNA irradiated with 254 nm light (4). The TA* photoproduct is of interest because of its implications for gene expression, which depends on a highly conserved TATA site in the promoter region. Recently, we have succeeded in constructing a 49mer containing a site-specific TA* photoproduct for biological studies, and in particular, *in vivo* mutagenesis studies (5). To understand properly the biological properties of this interesting photoproduct requires knowledge of its precise structure, and though the proposed structure was well supported by various spectroscopic and chemical studies, there are a number of inconsistencies. Unlike most [2+2] cycloadducts involving the 5,6-double bond of thymine, the TA* product was reported not to be photoreversed by 254 nm UV light (1). Another surprising aspect of structure **1** is that it would be stable in spite of what must be a sizeable amount of ring strain, and a number of viable pathways for ring fragmentation.

In addition to d(TpA), direct irradiation of thymidylyl-(3′,5′)-2′-deoxyinosine, d(TpI), with 254 nm has also been reported to give the analogous d(TpI)* photoproduct with a similar quantum yield (4). The far-UV irradiation of various other adenine-containing substrates, however, have been reported to lead to other ring fusions (Fig. 2). The photoadduct **3** is proposed to form between the 5,6-double bond of thymine and the 4,5-double bond of hypoxanthine when the two bases are linked by a trimethylene chain which then hydrolyzes to give **4**. Irradiation of 5-alkyldeoxyuridylyl-(3′-5′)-deoxyadenosine has been proposed to give **5**, a [2+2] adduct between the 5,6-double bond of the uridine derivative and the 7,8-double bond of adenine. Irradiation of d(ApA) gives two major products A=A and d(ApA)* which have been assigned structures **7** and **8** (6). These products were proposed to arise from fragmentation of an initial [2+2] cycloadduct between the 5,6- and 7,8-double bonds, compound **6**. Although previous spectroscopic measurements and chemical degradation studies of d(TpA)* appeared to rule out ring fusions involving anything but the 5,6-double bonds of T and A, direct spectroscopic evidence for the actual nature of the ring fusion has been lacking. Herein, we report the assignment of the ^{13}C-NMR spectrum of d(TpA)*, and the results of ^{1}H-^{13}C heteronuclear multiple-quantum coherence (HMQC) and hetero multiple bond correlated (HMBC) NMR experiments. The NMR

* To whom correspondence should be addressed

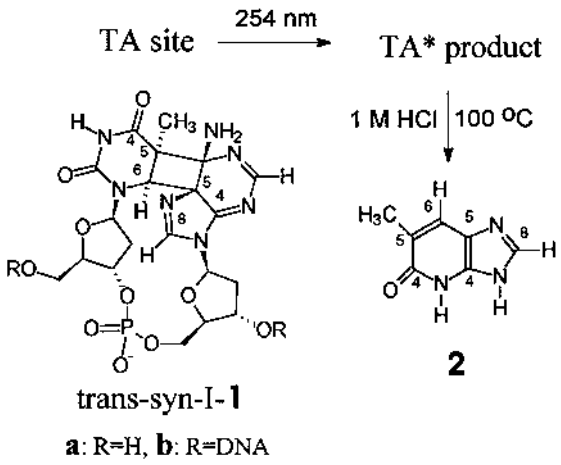

Figure 1. Originally proposed trans–syn-I structure for d(TpA)* and its acid degradation product, which is labeled according to the likely origin of the carbon atoms (2,3).

data is inconsistent with the originally proposed structure **1a** for d(TpA)*, but is consistent with an eight-membered ring valence isomer of **1a** that explains its inability to be photoreversed by 254 nm irradiation.

MATERIALS AND METHODS

Reagents and equipment

Anion-exchange HPLC was carried out on a Nucleogen DEAE 60-7 column. Reverse phase HPLC was carried out on a Dynamax C-18 column (4.6 × 250 mm for analytical work and 21.6 × 250 mm for preparative work). NMR spectra were obtained on Varian UNITY plus-500 and UNITY-600 spectrometers and the data processed on a SPARC 10 computer with VNMR software.

Preparation of d(TpA)*

TpdA, **3**, (20.3 mg, 0.036 mmol) in 20 ml double-distilled water was dispersed on two Petri dishes to give an overall exposure area of 127 cm^2 and irradiated with 2.23 mW/cm^2 254 nm light for 12 h on ice. The irradiated solution was concentrated and subjected to preparative C-18 HPLC on a Dynamax column (8 µm, 21.6 mm × 250 mm) with a 4.0 ml/min gradient of acetonitrile in 75 mM KH_2PO_4/K_2HPO_4, pH 6.8 that went from 0–10% over 30 min, followed by 20 min at 10%, and then 10–25% over 10 min. The fraction with an approximate retention time of 49 min was concentrated and then desalted on the same C-18 column by washing with 100 ml water to give 7.2 mg (36%) of d(TpA)*. The ^{1}H-NMR spectrum was the same as that reported in the literature (2); ^{31}P-NMR (121.5 MHz, D_2O, referenced to TMP), δ –1.35; MS (FAB), m/z (intensity), M+K, 594.

NMR spectroscopy of d(TpA)*

All 2D spectra were acquired in phase-sensitive mode by employing the Hypercomplex method (7). ROESY (8) spectra were recorded with a 125 and 250 ms mixing times and a 2 kHz spin-lock rf field strength. A total of 2 × 320 × 2048 data matrix with 16 scans per t1 value were collected. Gaussian line broadening and sine-bell

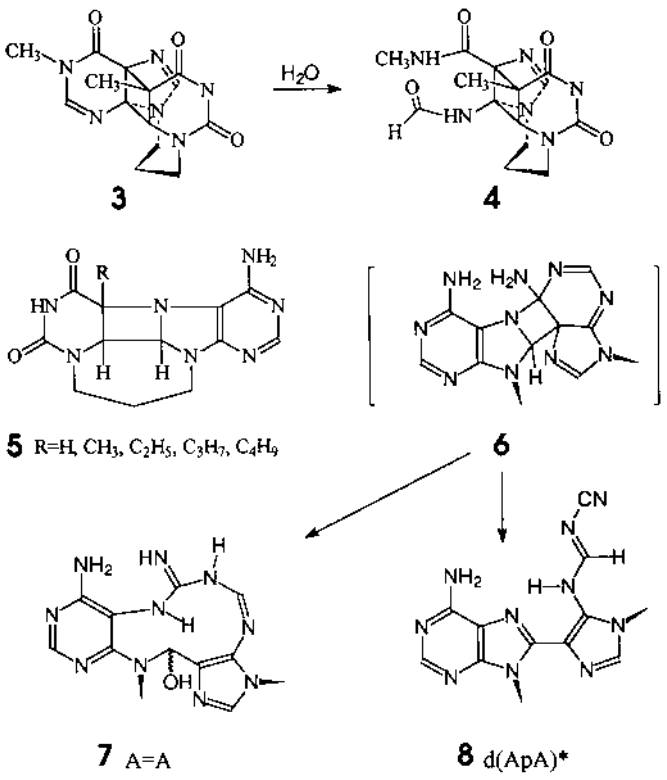

Figure 2. Proposed structures of other pyrimidine–purine and purine–purine photoadducts.

function were used in weighting t2 and t1 dimensions, respectively. After two-dimensional Fourier transformation, the spectra were obtained as 2k × 2k data matrices, and were phase and baseline corrected in both dimensions. TOCSY spectra (9) were acquired with a 125 ms mixing time and were processed as for the ROESY spectra. NOESY (10,11) spectra were obtained from 2 × 320 × 2048 data matrices with 16 scans per t1 value. The spectral data were acquired at 298 K with 250, 500 and 750 ms mixing times, but the 750 ms mixing time gave the best spectrum. The time domain data were zero filled to yield 2k × 2k spectral data matrices and were processed in the similar way as the 2D ROESY spectrum. The proton-detected heteronuclear multiple quantum coherence (HMQC) (12,13) spectrum was recorded on a Unity plus-500 spectrometer using a 0.35 s ^{1}H-^{13}C nulling period (Fig. 3). The 90° ^{1}H pulse width was 7.8 µs and the 90° ^{13}C pulse width was 14 µs. The proton spectral width was set to 4200 Hz and carbon spectral width to 12 500 Hz. A 2 × 200 × 2048 data matrix with 64 scans per t1 value was collected. Gaussian line broadening was used in weighting both t2 and t1 dimensions. Two-dimensional Fourier transformation resulted in spectra with 512 × 2048 data points, which were phase and baseline corrected in both dimensions. Heteronuclear multiple bond coherence (HMBC) spectra were recorded in a similar way to the HMQC experiment except that delays of 35, 55, 75 and 95 ms were employed to observe the multiple-bond correlations. Selective HMBC (14) was employed to resolve the long-range heteronuclear correlations of the two very closely spaced resonances, TC4 and AC6. The spectrum was recorded on a Varian Unity 600 spectrometer using two 90° ^{13}C Gaussian pulses of 1.5 ms and a ^{13}C spectral width of 1358 Hz to

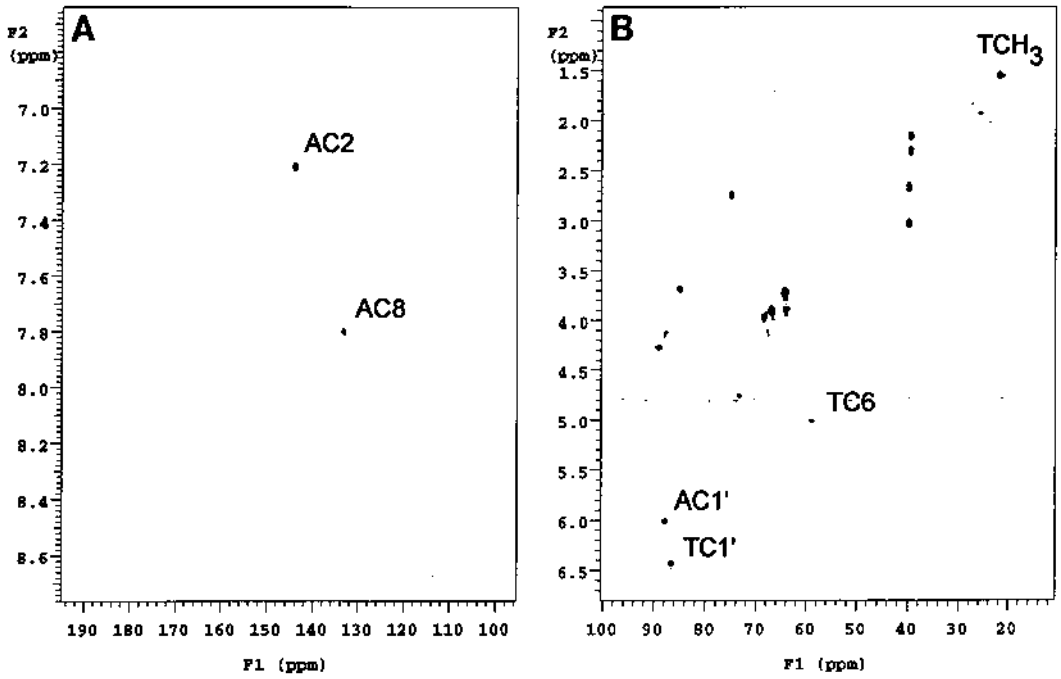

Figure 3. HMQC spectrum of d(TpA)* showing the single bond correlations between the protons and (**A**) the unsaturated carbons, and (**B**) the saturated carbons. Only the correlations to the base carbons, and the C1′ carbons are labeled.

cover the two resonances in question. A 2 × 256 × 2048 data matrix with 64 scans per t1 value was used and zero filled to 1024 × 2048 data points.

Chemical shift calculations

^{13}C chemical shift calculations were carried out with the C-13 NMR module of ChemWindow Version 3.1.3 which is an implementation of the C13Shift program (15). This program utilizes linear additivity rules and an extensive parameter set (16,17) to predict ^{13}C shifts with a mean deviation of –0.29 p.p.m. and a standard deviation of 5.5 p.p.m. for 97% of 160 000 shifts that were analyzed (15).

Molecular modeling

Torsion angle constraints were derived from the reported coupling constant data (2) as has been previously described for the (6–4), Dewar and trans–syn-II photoproducts of TpT (18,19). The newly derived torsion angles were used along with the previously derived interproton distance ranges as constraints in molecular mechanics minimizations with SYBYL. The starting geometry for the thymine–adenine ring system of d(TpA)* was obtained by first carrying out the cycloaddition between the 5,6-double bonds of T and A from a B form conformation in which the T was in either an *anti* or *syn* conformation. The AC5–AC6 bonds of the resulting trans–syn-I-**1a** and cis–syn-**1a** adducts were then broken and the resulting structures minimized with torsion angles and distance ranges imposed as constraints. Semi-empirical AM1 molecular orbital calculations were carried out with Hyperchem Version 2.0.

RESULTS AND DISCUSSION

Assignment of the ^{13}C-NMR spectrum and ^{1}H-^{13}C connectivities

The assignment of the ^{13}C-NMR spectrum of d(TpA)* was facilitated by the previously reported proton NMR assignment based upon COSY and NOESY spectra (2) which was confirmed by TOCSY, ROESY and NOESY spectroscopy. To assign the ^{13}C spectrum and obtain bond connectivities ^{1}H-^{13}C HMQC and HMBC NMR spectra were obtained. The strong HMQC crosspeaks (Fig. 3) enabled assignment of all carbons directly attached to protons and included all of the sugar carbons (Table 1). The only base carbons that could be assigned by HMQC were the two aromatic carbons bearing non-exchangeable protons of the adenine subunit and the C6 and methyl carbon signals of the thymine subunit.

To complete the assignment of the ^{13}C NMR spectrum and to establish the multiple bond proton–carbon connectivities in d(TpA)*, an HMBC spectrum was obtained (Fig. 4). Coupling from the modified bases to the deoxyriboses was only observed between TH1′ and a quaternary unsaturated carbon signal at 156.5 p.p.m. allowing it to be identified as TC2 (Fig. 4D). The ^{13}C-NMR spectrum of d(TpA)* has two very close downfield quaternary unsaturated carbon peaks at δ173.9 and δ173.6 of which the lower field signal is coupled to TCH_3 (Fig. 4A) and the higher field signal is coupled to TCH_3, TH6 and AH2 (Fig. 4A, C and D) allowing them to be assigned to AC6 and TC4 respectively. Correlations between a quaternary unsaturated carbon with a shift of 128.5 p.p.m. and both the AH2 and AH8 protons allow it to be assigned to AC4 (Fig. 4D). Correlations between a quaternary unsaturated carbon with a shift of 127.6 p.p.m.

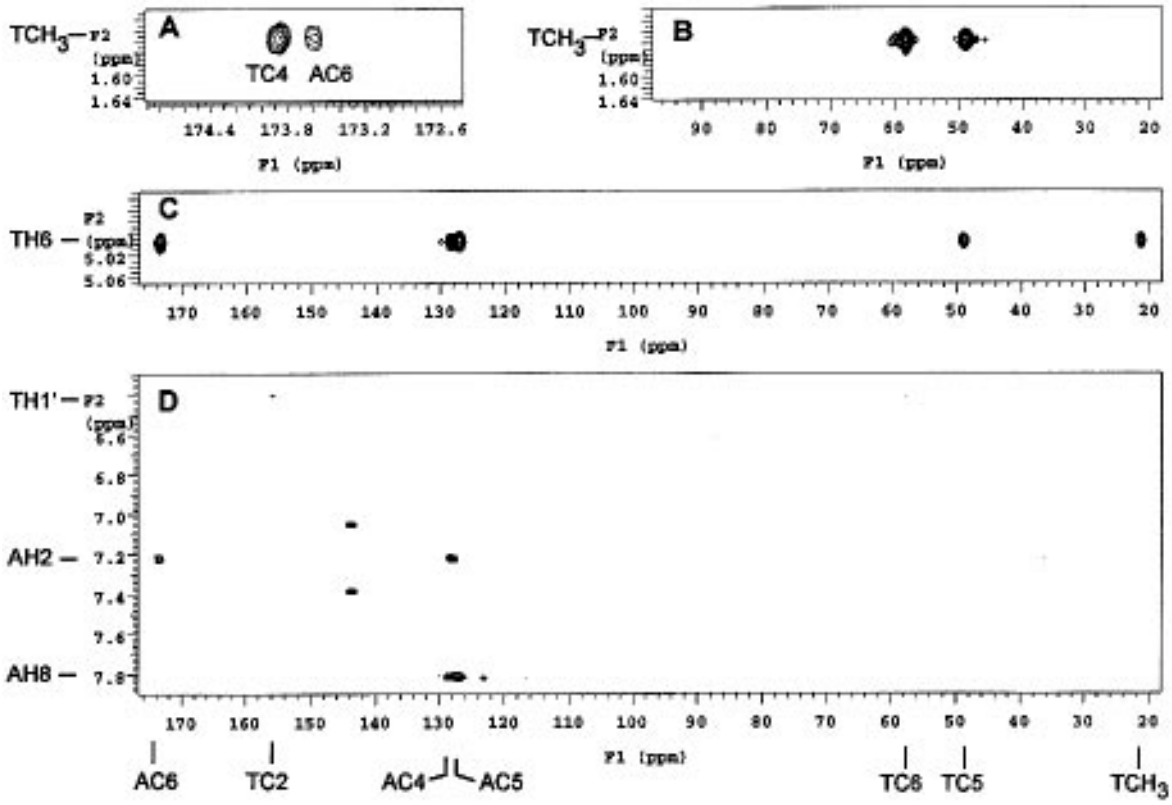

Figure 4. Portions of the selective (**A**) and full (**B–D**) MHz HMBC spectra of d(TpA)* that show correlations between the hydrogens and the carbons of the bases.

and TH6 and AH8 (Fig. 4C and D) allow it to be assigned to AC5. The remaining quaternary saturated carbon signal a 49.4 p.p.m. could then be assigned to TC5 by way of correlations with TH6 and TCH_3 (Fig. 4C and D).

Table 1. Observed single (s) and multiple (m) bond ^{13}C-^{1}H correlations of the base carbons and hydrogens d(TpA)* and their assignment to structure **9a**

δ ^{13}C / (δ ^{1}H)	1.57 (TCH_3)	5.01 (TH6)	6.40 (TH1′)	6.01 (AH1′)	7.25 (AH2)	7.81 (AH8)
173.9 (TC4)	m (3)	*m (3)				
173.6 (AC6)	m (3)	m (3)			m (3)	
156.5 (TC2)		*m (3)	m (3)			
144.2 (AC2)					s	
133.4 (AC8)				*m (3)		s
128.5 (AC4)		m (3)		*m (3)	m (3)	m (3)
127.6 (AC5)		m (2)				m (3)
87.5 (AC1′)				s		*m (3)
86.2 (TC1′)		*m (3)	s			
58.7 (TC6)	m (3)	s	*m (3)			
49.4 (TC5)	m (2)	m (2)				
21.8 (TCH_3)	s	m (3)				

The number of multiple bonds is given in parentheses. An asterisk indicates that a crosspeak was not observed.

Structure and mechanism of formation of d(TpA)*

The presence of only three saturated carbon signals (δ 21.8, 49.4 and 58.7) that are assignable to the base portion of d(TpA)* is inconsistent with the originally proposed structure for d(TpA)* (2,3) which would have five such signals. This data rules out any [2+2] adducts between the 5,6-double bond of the thymine ring and any double bond of the adenine ring, but suggests, that the d(TpA)* structure may be the result of a fragmentation of an initial [2+2] adduct. Of the three [2+2] pyrimidine–purine adducts typified by structures **1**, **3** and **5**, only **1** appears to be able to fragment to a structure containing only two saturated carbons that is also consistent with all of the other available data (Fig. 5, structure **9**). A similar fragmentation has been proposed for the formation of A=A, **7**, from the [2+2] adduct of d(ApA), **6** (6) (Fig. 2). Structure **9a** has the same molecular mass as **1a**, and because the crucial bond connections are preserved, structure **9a** would lead to the same acid hydrolysis product, **2** (Fig. 1). Structure **9a** would also explain why d(TpA)* is not photoreversed by 254 nm irradiation, as it is not a [2+2] adduct. The increase in pK of d(TpA)* relative to TpdA (from 3.5 to 5.5) (1,3) is also consistent with the formation of the more basic amidino functional group present in **9a**.

Structure **9a** not only has empirically calculated ^{13}C chemical shifts that fit the experimental data reasonably well (Table 2), but all the assigned multiple bond correlations take place through three or less bonds (Table 1). Some three bond correlations were not observed with a delay of 55 ms (noted by asterisks) and were probably absent because of small coupling constants and short relaxation times (20). Attempts to detect the missing three bond

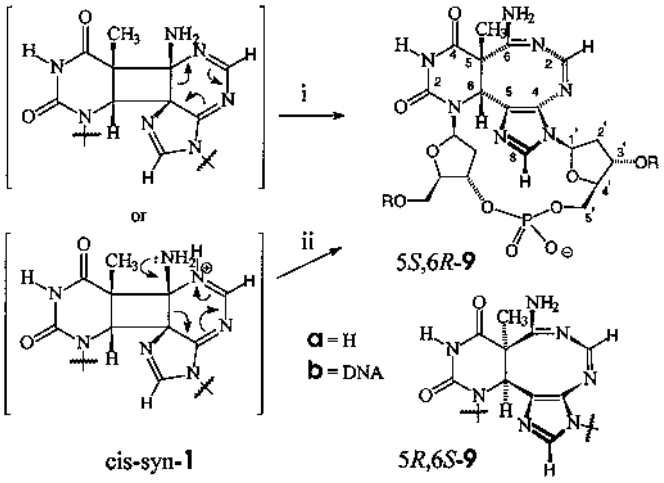

Figure 5. Proposed structure **9a** for the d(TpA)* product, and electrocyclic (path i) and acid-catalyzed (path ii) mechanisms for the formation of the 5*S*,6*R* stereoisomer of **9** from the cis–syn [2+2] adduct **1**. The 5*R*,6*S* stereoisomer would result from the trans–syn-I-**1** intermediate shown in Figure 1.

correlations expected for structure **9a** by using delays of 35, 75 and 95 ms were unsuccessful and resulted in fewer observable multiple bond correlations.

Table 2. ^{13}C NMR assignments of the d(TpA)* photoproduct to structures **1a** and **9a** along with calculated ^{13}C chemical shifts

	Tp ring			pdA ring		
	Assigned	Calcd for **1a**	Calcd for **9a**	Assigned	Calcd for **1a**	Calcd for **9a**
C1′	86.2	75	74	87.5	72	73
C2′	39.4	33	33	39.8	36	36
C3′	74.7	67	67	73.0	64	65
C4′	84.6	79	79	88.7	79	79
C5′	63.9	61	61	66.6	64	64
C2	156.5	163	163	144.2	163	163
C4	173.9	176	176	128.5	**164**	122
C5	49.4	56	56	127.6	**79**	122
C6	58.7	44	40	173.6	**78**	164
CH_3/C8	21.8	8	6	133.4	**163**	136

Key differences are in bold type. The ^{13}C NMR shifts are referenced to 3-(trimethylsilyl)-l-propanesulfonic acid, sodium salt.

Valence tautomerization of the presumed intermediate [2+2] adduct to the cyclooctatriene ring system could take place by either a thermally allowed disrotatory Cope reaction (Fig. 5, path i) or by a general acid base catalyzed process (Fig. 5, path ii). The electrocylic pathway is well precedented by the observation that 1,3,5-cyclooctatriene is in equilibrium with the bicyclo [4.2.0] isomer (21). Unlike the carbocyclic analog, which contains ~15% of the bicyclic compound at equilibrium, formation of **9** from **1** may be driven by aromatic stabilization of the imidazole ring. Semi-empirical molecular orbital calculations with AM1 (22) on the substructure containing only the C5 and C6 carbons of the T as methylenes and all the atoms of the A, suggest that the eight-membered ring form is favored by ~26 kcal/mol.

Conformational analysis

To obtain models for the d(TpA)* photoproduct, torsion angle and distance range constrained molecular mechanics energy minimizations were carried out. The torsion angle constraints were derived from the reported coupling constant data (2), according to methods that we have used to analyze the (6–4), Dewar and trans–syn-II photoproducts of TpT (18,19,23). Our evaluation of the coupling constant data according to unrestricted one and two state models is in general agreement with that previously reported (2) (Table 3). Unlike the previous study, which concluded that the deoxyribose ring of the T could be described by only 55% of a C3′-endo-C4′-exo conformation (P = 42, Φ = 35), we conclude that this is the major if not exclusive conformation (P = 30, Φ = 35 RMS deviation of 0.55 Hz). In accord with the previous study, we conclude that the pdA sugar is indeed represented by 50% of a C3′-exo conformation, and in addition, that the remaining conformers can be represented by a single C2′-exo conformer (RMS deviation of 0.32 Hz for the two state model). We are also in agreement with the analysis of the backbone angles, except for the previously reported value of 152° for ε. The coupling constant of 7.4 Hz corresponds to an HCOP torsion angle of ±30°, which corresponds to ε values of 210° (–150°) and 270° (–90°), of which –150° is more reasonable.

Table 3. Conformational parameters (in degrees) of d(TpA)* used to constrain modeling calculations

	Tp	pdA (50%)	pdA′ (50%)
β	–	180	180
γ	none	60	60
δ	+85	+158	+96
ε	–150	–	–
P,Φ_M	30, 35	198, 40	342, 40
	C3′-endo	C3′-exo	C2′-exo
τ0,τ1,τ2,τ3,τ4	–8, –15, 30, –35, 26	–1,23,–39,38,–24	23,–39,38,–24,–1

Molecular models of d(TpA)*

Interproton distance ranges for constrained molecular modeling of the various stereoisomers of structure **9a** were taken from the original NMR study of d(TpA)* (2). With one exception, distances for NOE's that were not reported were confirmed not to be present by analysis of both 125 and 250 ms ROESY and 250, 500 and 750 ms NOESY spectra, and taken to be >3.5 Å. The one exception was an extremely weak crosspeak between AH8 and TH3′ that was observed in both halves of the 750 ms NOESY spectrum (Fig. 6) and appears to have been incorrectly assigned as being between AH8 and AH2″ in the original study (2). This crosspeak could not have arisen indirectly as no other NOEs were observed between AH8 and any of the other protons of the T subunit, and suggested that AH8 and TH3′ were within 3.5 Å of each other. The presence of this crosspeak and crosspeaks between AH8 and AH2′ and none between AH8 and AH1′ indicated that the A was in an *anti* glycosyl conformation. Such a product would have most likely arisen from fragmentation of an

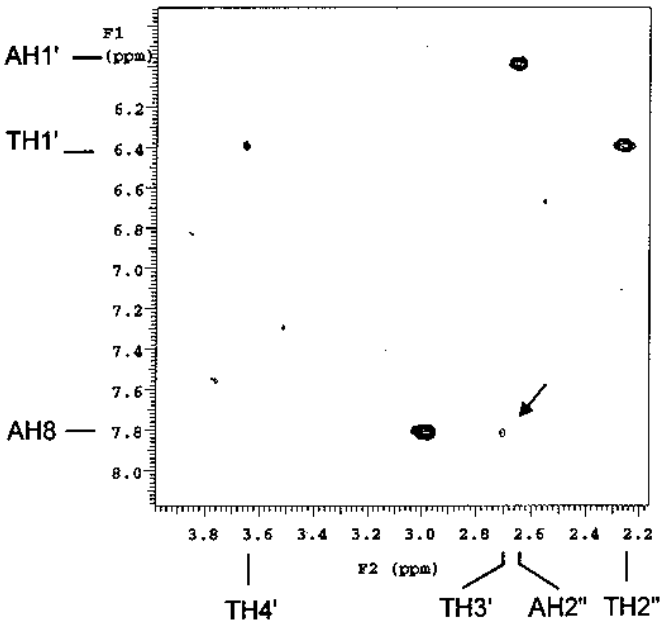

Figure 6. Portion of the 500 MHz 750 ms NOESY of d(TpA)* showing the weak but detectable correlation between AH8 and TH3′ identified with the arrow.

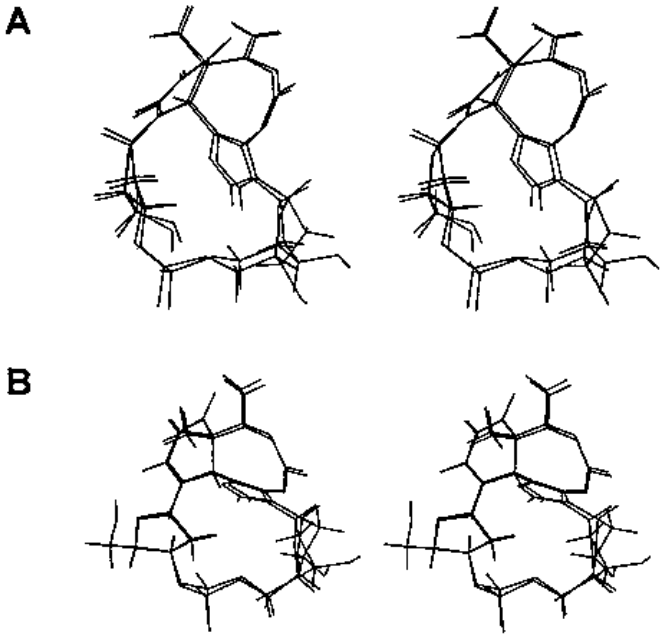

Figure 7. Stereoprojections of molecular models of the (A) 5*S*,6*R* and (B) 5*R*,6*S* isomers of **9a** that are consistent with the NMR data for d(TpA)*.

initial [2+2] adduct **1a** in which the A was also in an *anti* glycosyl conformation. The presence of a crosspeak between TH6 and only TH1′ of all the sugar protons suggested that it could be in either the *syn* or *anti* glycosyl conformation and likewise could have arisen from a [2+2] adduct **1a** in which the T was also in either an *anti* or *syn* conformation.

Thus, only two stereoisomers of **9a** were considered for further evaluation, the 5*S*,6*R*-**9a** structure which results from fragmentation of cis–syn-**1a**, and the 5*R*,6*S*-**9a** structure which results from fragmentation of trans–syn-I-**1a**. Distance ranges and torsion angles derived from the NOE and coupling constant data were used to constrain energy minimization of models of both the 5*S*,6*R* and 5*R*,6*S* stereoisomers of structure **9a**, each with two different puckers for the deoxyribose subunit of the A (Fig. 7). The 1,3-diazacyclooctatriene subunit of the resulting models had the expected tub shape known for cyclooctatriene and heteroatom derivatives (24) in which the π bonds are almost orthogonal to each. The poorly conjugated nature of the π-orbital system is consistent with the absence of any absorption peak >240 nm (1). Recently, there was a report of a substituted pyridine that photodimerizes to give a [2+2] adduct which then undergoes a Cope rearrangement followed by bridgehead oxidation to give a bicylo[6.4.0] ring system (25). The X-ray crystal structure of the oxidized product has a tub conformation with nearly orthogonal double bonds in the cyclooctatetraene ring system.

Stereochemical assignment

The angle and distance constrained models for the 5*S*,6*R* and 5*R*,6*S* isomers of structure **9a** fit all the available NMR data quite well (Table 4). The molecular models, however, suggest that the 5*S*,6*R* stereoisomer might be distinguishable from the 5*R*,6*S* isomer by the presence of an NOE between TH6 and TH3′. Such a NOE was not reported in the previous NMR study of d(TpA)* (2) or in the present study but an extremely weak crosspeak was observed between these two protons in the NOESY spectrum of d(GTATTA*TG), but not in d(GTA*TTATG).

Table 4. Comparison of selected interproton distances in Å corresponding to observed and unobserved NOE's in d(TpA)* for the models of the 5*S*,6*R* and 5*R*,6*S* stereoisomers of **9a** derived from distance and angle constrained molecular mechanics calculations

1st H	2nd H	Exp[a]	5*S*,6*R*-a	5*S*,6*R*-b	Avg[b]	5*R*,6*S*-a	5*R*,6*S*-b	Avg[b]
TH6	TH1′	<3.5	3.3	3.3	3.3	2.6	2.5	2.5
	TH2′	>3.5	4.3	4.3	4.3	3.0	3.0	3.0
	TH2″	>3.5	4.9	4.9	4.9	3.5	3.5	3.5
	TH3′	>3.5[c]	4.0	4.0	4.0	4.7	4.8	4.7
AH8	TH6	>3.5[c]	4.7	4.6	4.6	5.4	5.4	5.4
	TH3′	<3.5[d]	3.1	3.3	3.2	3.4	3.8	3.6
	AH1′	>3.5	3.6	3.8	3.7	3.8	3.9	3.8
	AH2′	<2.5	2.7	2.5	2.6	2.3	2.8	2.5
	AH2″	>3.5[d]	3.1	4.0	3.4	3.3	4.3	3.6
	AH3′	>3.5[c]	4.9	2.6	2.9	4.8	2.6	2.9

[a]Distance ranges reported in the previous study (2).
[b]The average corresponds to an r^{-6} weighted average distance.
[c]Distances with an asterisk correspond to NOEs that were not specifically described in the previous study and were confirmed to be absent in the present study.
[d]Misassigned in the previous study (see text).

Evidence that the d(TpA)* could have the 5*S*,6*R*-**9a** structure, comes from an explanation for the highly unusual chemical shift of TH3′. The TH3′ signal appears at 2.73 p.p.m. in d(TpA)*, compared with 4.59 p.p.m. in TpdA (26), and appears at 2.68 p.p.m. in d(GTATTA*TG) compared with 4.93 p.p.m. in the parent oligonucleotide (5). The authors of the previous NMR study of d(TpA)* (2) attributed the large upfield shift of TH3′ to shielding by the C2 carbonyl of the thymine in a syn glycosyl conformation, which could also explain the upfield shift in the 5*R*,6*S* stereoisomer of **9a**. The large upfield shift of TH3′ is also

consistent with the 5*S*,6*R* stereoisomer of **9a**, as TH3′ is situated directly under the imidazole ring and therefore in its shielding region (Fig. 7).

Thus, we conclude that the d(TpA)* product could arise from an *anti*, *anti* glycosyl conformation of the two nucleotides, as would be found in B DNA, and as was originally proposed for the [2+2] adduct (3). Only one stereoisomer of the TA* product of d(GTATTATG) appears to be produced by 254 nm irradiation, and has an ^{1}H-NMR spectrum that is very similar to that of d(TpA)*, suggesting that it also has the same stereochemistry (5). This fact, coupled with the fact that the TA* product is also produced in double-strand DNA, in which the *anti* glycosyl conformation predominates, further supports the notion that TA* has the 5*S*,6*R* stereochemistry. Such a stereochemical assignment would be more consistent with the fact that the cis–syn dimer, which also arises from an *anti*, *anti* glycosyl conformation, is the major product of TpT, and TT in both single- and double-strand DNA, while the trans–syn-I isomer, which arises from the *syn*, *anti* conformation, is a very minor product (23,27). Though all the available data for d(TpA)* is consistent with structure **9a**, a crystal structure will be required to unambiguously determine its stereochemistry and conformation, and attempts to crystallize d(TpA)* are in progress.

ACKNOWLEDGEMENTS

This work was supported by NIH Grant R37 CA40463. The assistance of the Washington University High Resolution NMR Facility, funded in part through NIH Biomedical Research Support Shared Instrument Grants RR-02004, RR-05018 and RR-07155 is also gratefully acknowledged.

REFERENCES

1 Bose,S.N., Davies,R.J.H., Sethi,S.K. and McCloskey,J.A. (1983) *Science* **220**, 723–725.
2 Koning,T.M.G., Davies,R.J.H. and Kaptein,R. (1990) *Nucleic Acids Res.* **18**, 277–284.
3 Bose,S.N., Kumar,S., Davies,R.J.H., Sethi,S.K. and McCloskey,J.A. (1984) *Nucleic Acids Res.* **12**, 7929–7947.
4 Kumar,S. and Davies,R.J.H. (1987) *Photochem. Photobiol.* **45**, 571–574.
5 Zhao,X., Kao,J.L.F. and Taylor,J.-S. (1995) *Biochemistry* **34**, 1386–1392.
6 Kumar,S., Joshi,P.C., Sharma,N.D., Bose,S.N., Davies,R.J.H., Takeda,N. and McCloskey,J.A. (1991) *Nucleic Acids Res.* **19**, 2841–2847.
7 States,D.J., Haberkorn,R.A. and Ruben,D.J. (1982) *J. Magn. Reson.* **48**, 286–292.
8 Kessler,H., Griesinger,C., Kerssebaum,R., Wagner,K. and Ernst,R.R. (1987) *J. Am. Chem. Soc.* **109**, 607–609.
9 Bax,A. and Davis,D.G. (1985) *J. Magn. Reson.* **65**, 355–360.
10 Kumar,A., Ernst,R.R. and Wuthrich,K. (1980) *Biochem. Biophys. Res. Commun.* **95**, 1–6.
11 Macura,S., Huang,Y., Suter,D. and Ernst,R.R. (1981) *J. Magn. Reson.* **43**, 259–281.
12 Bax,A. and Subramanian,S. (1986) *J. Magn. Reson.* **67**, 565–569.
13 Muller,L. (1979) *J. Am. Chem. Soc.* **101**, 4481.
14 Kessler,H., Schmieder,P., Köck,M. and Kurz,M. (1990) *J. Magn. Reson.* **88**, 615–618.
15 Pretsch,E., Fürst,A., Badertscher,M., Bürgin,R. and Munk,M.E. (1992) *J. Inf. Comput. Sci.* **32**, 291–295.
16 Fürst,A., Robien,W. and Pretsch,E. (1990) *Anal. Chim. Acta* **233**, 213–222.
17 Pretsch,E., Fürst,A. and Robien,W. (1991) *Anal. Chim. Acta* **248**, 415–428.
18 Taylor,J.-S., Garrett,D.S. and Cohrs,M.P. (1988) *Biochemistry* **27**, 7206–7215.
19 Taylor,J.-S., Garrett,D.S. and Wang,M.J. (1988) *Biopolymers* **27**, 1571–1593.
20 Summers,M.F., Marzilli,L.G. and Bax,A. (1986) *J. Am. Chem. Soc.* **108**, 4285–4294.
21 Cope,A.C., Haven,A.C., Ramp,F.L. and Trumbull,E.R. (1952) *J. Am. Chem. Soc.* **74**, 4867–4871.
22 Dewar,M.J.S., Zoebisch,E.G., Healy,E.F. and Stewart,J.J.P. (1985) *J. Am. Chem. Soc.* **107**, 3902–3909.
23 Kao,J.L.-F., Nadji,S. and Taylor,J.-S. (1993) *Chem. Res. Toxicol.* **6**, 561–567.
24 Paquette,L.A. (1971) *Angew. Chem., Int. Ed. Engl.* **10**, 11–20.
25 Sakamoto,M., Takahashi,M., Kimura,M., Fujihira,M., Fujita,T., Iida,I., Nishio,T. and Watanabe,S. (1994) *J. Org. Chem.* **59**, 5117–5119.
26 Cheng,D.M. and Sarma,R.H. (1977) *J. Am. Chem. Soc.* **99**, 7333–7348.
27 Patrick,M.H. and Rahn,R.O. (1976) In Wang,S.Y. (ed.), *Photochemistry and Photobiology of Nucleic Acids.* Academic Press, New York, NY, Vol. II, pp. 35–95.