Solution Structure by NMR and Molecular Dynamics of a Duplex Containing a Guanine Opposite a *N*-(2-Deoxy-β-D-*erythro*-pentofuranosyl)formamide Lesion†

Corinne Maufrais,‡ G. Victor Fazakerley,‡ Jean Cadet,§ and Yves Boulard*,‡

CEA, Service de Biochimie et Génétique Moléculaire, Bat 142, CEA Saclay 91191 Gif-sur-Yvette Cedex, France, and CEA, Service de Chimie Inorganique et Biologique, Département de Recherche Fondamentale sur la Matière Condensée, F.38054 Grenoble Cedex, France

Received December 16, 1999; Revised Manuscript Received February 23, 2000

ABSTRACT: One- and two-dimensional NMR spectroscopy has been used combined with molecular dynamics to determine the fine structure of the DNA duplex 5′-d(AGGAGCCACG).d(CGTGGFTCCT) where F is the *N*-(2-deoxy-β-D-*erythro*-pentofuranosyl)formamide residue which is a ring fragmentation product of thymine. The formamide deoxyribose exists as two isomers with respect to the orientation about the peptide bond. The two isomers (trans and cis) are observed in a ratio 3:2 in solution. For both species, the oligonucleotide adopts a globally B form structure although conformational changes are observed around the mismatch site. The formamide residue, whatever the isomer, is intrahelical and can pair with the guanine on the opposite strand with one hydrogen bond. For the cis isomer, the residue adopts a syn orientation and is able to form a second hydrogen bond with the guanine on the 5′ side on the same strand. Off-resonance ROESY experiments have been used to investigate the chemical exchange observed at low temperature of the duplex. Conformational exchange has only been found for the oligonucleotide with the formamide residue in the trans conformation.

In living cells, base modifications, single, and double strand breaks are the most important damage products caused by ionizing radiation in DNA. These damages are implicated in biological processes such as mutagenesis, carcinogenesis and lethality. *N*-(2-deoxy-β-D-*erythro*-pentofuranosyl)formamide is a major product of the irradiation of thymine in aqueous solution (*1*, *2*). Hydroxyl radicals are produced by radiolysis of water molecules surrounding the biopolymer. Attack of hydroxyl radicals at the C-5 position leads to a C-6 carbon reduced radical to which O_2 is rapidly fixed. The subsequent peroxyl radical may, through several mechanisms, lead to formation of several oxidation products; among them, the formamide residue arises from the opening of the 5,6 bond and subsequent hydrolysis.

Only few studies have been reported concerning the consequences of the presence of a formamide residue in template DNA. The Klenow fragment directs the misinsertion of a guanine base opposite the formamide and Taq DNA polymerase gives rise preferentially, but not exclusively, to a deletion (*3*). It has been observed that the Klenow fragment and Taq DNA polymerase with a frequency estimated at 33 and 11%, respectively, bypass the formamide residue. Thus, the formamide residue appears to be a mutagenic lesion similar to that of an apurinic site (*4*). On the other hand, the small formamide adduct does not influence the AP endonuclease activity (*5*). Exonuclease III or endonuclease IV, the major repair AP nucleases in *Escherichia coli*, cleave double-stranded oligonucleotides containing the formamide residue, while it prevents hydrolysis of the 3′ phosphodiester bond by nuclease P1 (*6*). Oligonucleotide duplexes, which contain a formamide residue, are cleaved by the Fpg protein at the 3′- and the 5′-phosphodiester bonds of the formamide lesion and the endo III protein cleaves mainly the phosphodiester bond 3′ of the formamide residue by a hydrolysis reaction 7.

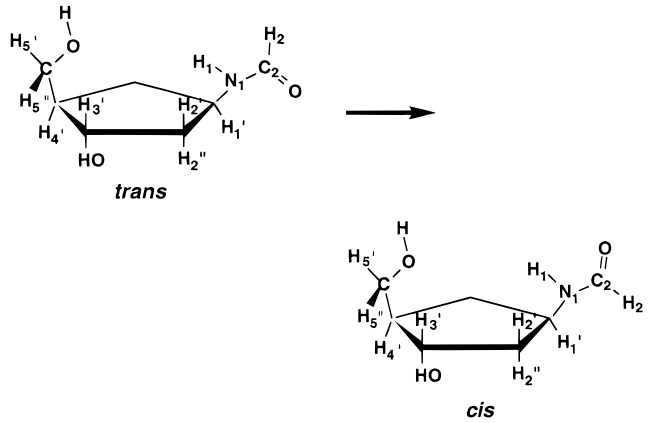

FIGURE 1: Structure of *N*-(2-deoxy-β-D-*erythro*-pentofuranosyl)-formamide (F) in the anti conformation.

The formamide lesion has lost part but not all of the coding information, and it is still able to form hydrogen bonds by the presence of one hydrogen bond acceptor and one donor, Figure 1. Two isomers, trans and cis, resulting from the slow rotation of the C−N amide bond have been characterized (*8*) and observed in NMR[1] spectra in the ratio 3:2 (*9*).

† This work was, in part, supported by funds from the Ministère de l'Education Nationale et de l'Enseignement Supérieur, ACC-SV8.

* To whom correspondence should be addressed. Phone: 01 69 08 35 84. Fax: 01 69 08 47 12. E-mail: ybou@matthieu.saclay.cea.fr.

‡ Service de Biochimie et Génétique Moléculaire.

§ Service de Chimie Inorganique et Biologique.

10.1021/bi9928849 CCC: $19.00
Published on Web 04/15/2000

In an attempt to understand why the formamide residue can act as a strong block to DNA polymerase and to determine the reasons for the preferential insertion of guanine, we report studies on the structural and dynamic properties of an oligonucleotide with a guanine facing the formamide lesion.

MATERIALS AND METHODS

NMR Spectroscopy. The decamer containing the *N*-(2-deoxy-β-D-*erythro*-pentofuranosyl)formamide residue was synthesized as previously reported (*3*). The pair of single strands was heated to 80 °C followed by slow cooling to form the duplex. The sequence is

Strand G:	5′	A1	G2	G3	A4	G5	C6	C7	A8	C9	G10	3′
Strand F:	3′	T20	C19	C18	T17	F16	G15	G14	T13	G12	C11	5′

The duplex was 2 mM in single strand concentration dissolved in 10 mM phosphate buffer, 150 mM NaCl and 0.2 mM EDTA. Chemical shifts were measured relative to the internal ref 3-(trimethylsilyl)propionate (TSP) at 0 ppm. NMR spectra were recorded on either Bruker DRX500, DRX600, or DRX800 spectrometers and all the spectra were acquired in the phase sensitive mode (*10*). NOESY spectra were recorded with 40, 50, 60, 80, 100, or 400 ms mixing times in D_2O and 80, 100, 150, or 250 ms in H_2O at 2, 7, 14, and 21 °C. In D_2O, the residual water resonance was presaturated during the relaxation and mixing delays. In H_2O, the water signal was suppressed with the WATERGATE sequence (*11*). TOCSY experiments were recorded with 40 or 80 ms mixing times. DQF-COSY were recorded with time-proportional phase incrementation (*12*). The off-resonance ROESY experiments with adiabatic rotation, used to suppress the unwanted HOHAHA transfers and reduce the offset effects, were performed at 60, 90, or 100 ms mixing times in the phase sensitive mode at 2 and 7 °C (*13*, *14*). Two spectrum with different angles between the static and the effective field in the rotating frame were recorded. The first angle had a values of 54.7° for which all HOHAHA transfers are suppressed and then only pure dipolar cross-relaxation is observed and the second angle had a value of 35.3° for which dipolar cross relaxation between protons of the macromolecule vanish and exchange phenomena can be selectively observed. The spin-lock pulse carrier frequency was shifted by approximately 5.5 kHz from the center of the spectrum.

Molecular Modeling. All initial structures were generated from the canonical Arnott B-DNA form (*15*). A set of conformations was generated by varying the χ angle of the formamide residue inside the helix. Certain computed structures could be rapidly discarded as they fitted poorly the NMR distances derived from build-up curves (*16*). Each structure was minimized with the program SANDER (*17*), using the AMBER force field (*18*), in two steps. Minimizations were first performed with constraints on δ angles of the sugars, intra- and internucleotide NMR-derived distances, and hydrogen bonds of the two terminal base pairs. The force constants used were 50 kcal mol^{-1} rad^{-2} for the δ angle and 5 kcal mol^{-1} Å^{-2} for the NMR-derived distances and hydrogen bonds. Subsequently, the best structures were minimized without any constraints in order to relax the final conformations to a root-mean-square (rms) energy gradient of less than 0.1 kcal Å^{-1} mol^{-1}. The structures were displayed on SGI workstations using the program MORCAD (*19*). The molecular dynamic (MD) computations were performed with the AMBER program as previously described (*20*, *21*). All MD runs were computed over 500 ps and with weak restraints on all δ torsion angles (except for those of the terminal residues) to reinforce the C2′-endo conformation. Two MD runs were performed: MDI includes all NMR distance constraints and MDII without NMR constraints. Helical parameters were computed with the OCL and MORCAD programs (*19*, *22*).

RESULTS AND DISCUSSION

Nonexchangeable Protons. For the F strand alone, two species in the ratio 3:2 were observed in 1D spectra as certain resonances corresponded to less than one proton. These arise from the two isomers of the formamide as previously reported for the monomer (*9*) where the major form is the trans isomer. We have recorded a series of 1D spectra as a function of pH and temperature to determine the optimum experimental conditions. We therefore chose to first study the duplex at pH 6.8 and 21 °C.

Major Species. The nonexchangeable proton resonance assignment was obtained by analyzing a NOESY spectrum recorded at 21 °C with a 400 ms mixing time and a TOCSY spectrum with a 40 ms mixing time. Figure 2 shows interactions between the aromatic protons H6/H8 and the H1′ sugar protons of a NOESY spectrum. For the G strand, we can follow the connectivities expected for a right-handed B-DNA helix from A1 to G10 without interruption, solid line in Figure 2A. We note that the interactions for the central part of the strand are normal, suggesting that no major disruption of the helix takes place from A4 to C6, the residues which could be influenced by the formamide on the opposite strand.

For the second strand, Figure 2B, the chain of connectivities can be followed from C11 to T20 without interruption. For the central part of the duplex, we can follow the H6/H8−H1′ connectivities from G14 to T17, especially the F16H2−G15H1′, F16H2−F16H1′, and the T17H6−F16H1′ interactions. The formyl proton (H2) of the formamide at 7.16 ppm shows both intra- and interresidue NOEs as for an H6/H8 aromatic proton. The others regions of the NOESY spectrum confirm this assignment. All the interactions observed for the central part of the duplex show that F16 stacks between G15 and T17.

Minor Species. At this stage, in both the NOESY and the TOCSY spectra, unassigned cross-peaks remained which correspond to the minor species (40% of the sample). In particular, we have found in the TOCSY spectrum an additional CH5−CH6 cross-peak at 7.25/4.97 ppm and another TH6-TCH$_3$ cross-peak at 7.35/1.54 ppm (data not shown). In Figure 2A, only the chemical shifts of the residues at the center of the duplex are different between the two

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; DQF-COSY, double-quantum-filtered correlated spectroscopy; TOCSY, total correlated spectroscopy; ROESY, rotating-frame Overhauser effect spectroscopy; 1D, one-dimensional; rms, root-mean-square; MD, molecular dynamic; F, *N*-(2-deoxy-β-D-*erythro*-pentofuranosyl)formamide, formamide lesion; τ_m, mixing time.

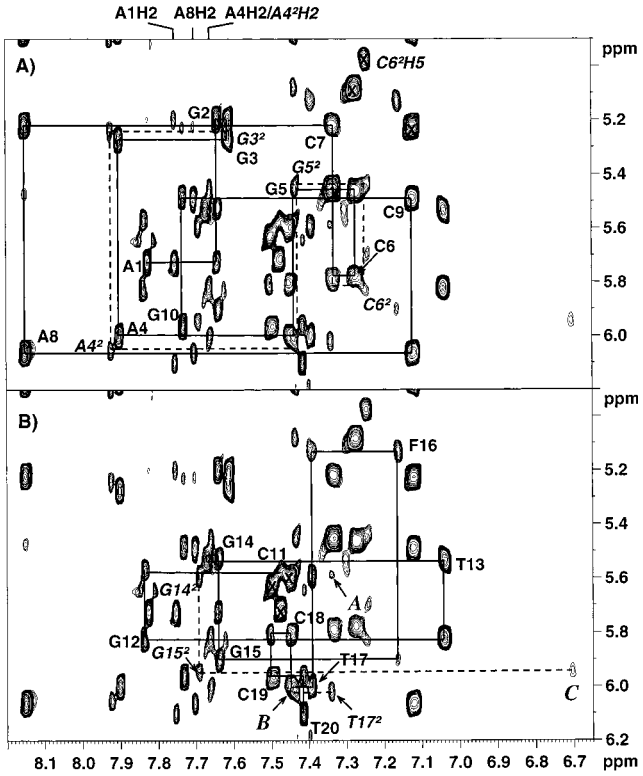

FIGURE 2: Expanded contour plot of the H6/H8−H1′/H5 region of the NOESY spectrum (400 ms mixing time) of the duplex at 21 °C in D_2O. Cross-peaks marked with an X correspond to H5−H6 interactions. A and B correspond to interactions A1 to G10 and C11 to T20, respectively. Interactions marked with a solid line correspond to the major species and those in broken lines to the minor species. Peaks A, B, and C are described in the text.

species. For A1 and G2 and from C7 to G10, the sequential assignment is the same for both species, and for these residues, the protons of the minor and major species cannot be distinguished. For the G strand, we observe two connectivity pathways from G3. We indicate the minor species in italic with an exponent of 2. The $G3^2-C6^2$ sequential interactions are clearly resolved, although the chemical shifts of the protons of each species are very similar. The $G5^2$ aromatic proton shows a weak cross-peak, visible at lower contour levels, with the additional H5 proton. Thus, the additional CH5−CH6 cross-peak at 7.25/4.97 ppm found in the TOCSY spectrum corresponds to $C6^2$ and the connectivities can be followed from $G5^2$ to C7. As for the major species, no major disruption is found in the center of the duplex.

For the F strand, we also observe splitting in the connectivity pathway on approaching the center of the duplex, from G14 to T17. The additional H6−CH_3 interaction found in the TOCSY spectrum must belong to the $T17^2$ residue. The weak $T17^2$H6−C18H5 interaction, peak A in Figure 2B, and the C18H6−$T17^2$H1′ interaction, peak B, which is better resolved when the data are more strongly filtered, confirm this assignment. The chain of connectivities is interrupted at $T17^2$, since the $T17^2$H6−$F16^2$H1′ interaction is not observed. The formyl proton of $F16^2$ at 6.71 ppm could be assigned from the interaction with the $T17^2CH_3$ resonance, which is expected since $F16^2$ is the adjacent residue in the 5′ direction. Cross-peaks between the aromatic proton of $T17^2$ and the H2′/H2″ protons of $F16^2$ confirm this assignment. The assignment can be continued from $F16^2$ to $G14^2$, with the $F16^2$H2−$G15^2$H2″ interaction (not shown). In the H6/H8−H1′ region Figure 2B, peak C corresponds to the $F16^2$-H2−$G15^2$H1′ interaction. The $G14^2$H8−T13H1′ interaction is visible when the data are less severely filtered. In conclusion, all the interactions observed for this species suggest that the formamide is intrahelical.

A series of NOESY spectra at short mixing times to measure NOE build-up curves allowed the determination of proton−proton distances. Each interproton distance measurement requires the choice of a reference (*23*). The spatial position of the formyl proton could not be directly compared to that of classical interactions of an aromatic proton; all the interproton distances for the formyl residue were thus derived by comparison with that of the fixed cytosine distance, H5−H6.

Sugar and Base Conformations. The sugar conformations of the Watson−Crick residues were determined by analyzing the relative intensities of the intraresidue cross-peaks H6/H8−H2′ compared to those of the H6/H8−H3′ in a NOESY spectrum recorded with a 60 ms mixing time. The H3′ assignments have been confirmed in a DQF-COSY spectrum. For all nonterminal sugars of the duplex, a predominantly C2′-endo conformation for both species was observed. The sugar/base orientations of the Watson−Crick residues were all determined anti by comparing the intensities of the intraresidue H6/H8−H1′ and H6/H8−H2′ cross-peaks in the same NOESY spectrum. It was, however, impossible to determine the sugar conformation and the sugar/base orientation for both isomers of the formamide. We could only determine the H2−H1′, the H2−H2′/H2″, and the H2−H3′ distances, but the spatial position of H2 is not unambiguously defined. In the DQF-COSY spectra, the necessary coupling constants could be measured (*24*), showing that the formamide sugar pucker of the major species is predominantly C2′-endo. Unfortunately, that of the minor species could not be determined because of resonance overlap.

Nonexchangeable Protons at Low Temperature. When the temperature is lowered, in the range from 7 to 2 °C, we observed interactions both in NOESY and TOCSY spectra between C6 and a new resonance at 7.87 ppm. These cross-peaks must correspond to intraresidue interactions. To confirm this hypothesis, we have recorded an off-resonance ROESY spectrum (Figure 3) with $\theta = 54.7°$, $\tau_m = 60$ ms and $T = 7$ °C, which demonstrates that this new resonance corresponds to a new C6H6 proton. Exchange cross-peaks were observed for the C6, C7, G15, and F16 residues of the major species with a second species noted with an asterisk (*). The new cross-peaks observed in both NOESY or TOCSY spectra can be assigned to intraresidue interactions or cross-peaks relayed by exchange. We observe for the nucleotide C6* the intraresidue C6*H6−C6*H2′/2″ interactions (not shown), but not the internucleotide interactions with G5 and C7* which could be expected, if C6* stacks inside the helix. Either C6* is extrahelical or the absence of these interactions is due to a dynamic effect induced by the exchange process. However, both the C6*H6 ($\delta = 7.87$ ppm) and C6*H5 ($\delta = 6.01$ ppm) resonances are strongly shifted downfield relative to that of the Watson−Crick base pair of the major species which strongly suggests that C6* is extrahelical. For the C7*, G15*, and F16* residues, only the exchange cross-peaks, but no intra- or interresidue

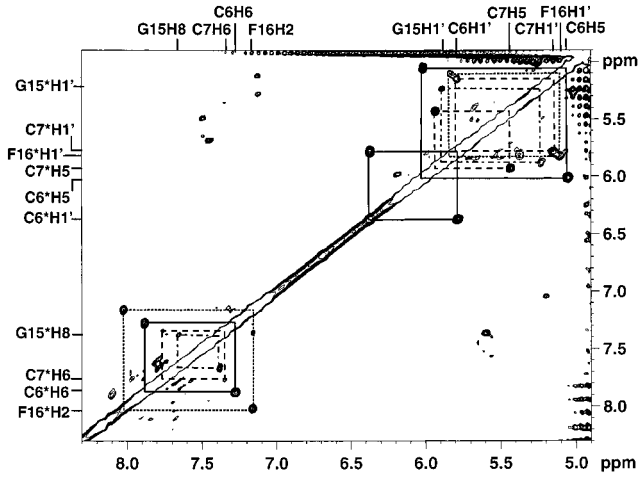

FIGURE 3: Expanded contour plot of the H6/H8/H1′/H5−H6/H8/H1′/H5 region of the ROESY off-resonance spectrum (90 ms mixing time) of the duplex at 7 °C in D_2O. The angle between the static and the effective field in the rotating frame is of 54.7°. The chemical shift of the protons of the previously assigned major species are given on the horizontal axis and the corresponding chemical shift of the species arising from chemical exchange on the vertical axis.

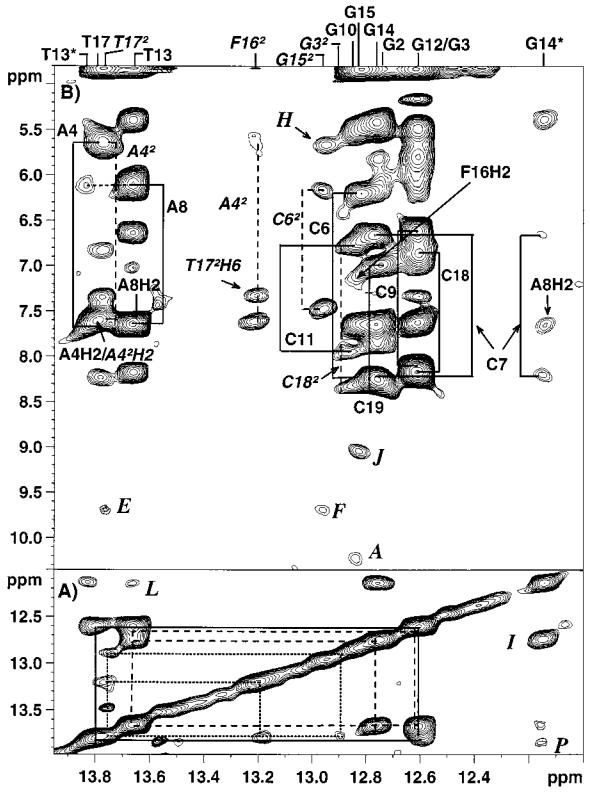

FIGURE 4: Expanded contour plots of a NOESY spectrum (150 ms mixing time) recorded at 2 °C in H_2O. (A) (Recorded at 600 MHz) Interactions between the imino protons. (B) (Recorded at 500 MHz) Interactions between the imino protons and the amino/H2/H5 protons. The pairs of protons of the amino groups of cytosine and adenine are connected by a continuous line for the major species and those of the minor species by dashed lines. Labeled cross-peaks are described in the text.

interactions are observed making the determination of their conformation impossible. We note that no exchange cross-peaks are observed for A4, G5, G14, and T17, the residues which could be influenced by the conformational change of the corresponding residues, F16 and C6, on the opposite strand.

Exchangeable Protons. One-dimensional NMR spectra were recorded between pH 5 and 8 and at different temperatures between 1 and 20 °C. Apart from the expected line width changes, the spectra showed no significant differences. The best resolution was obtained at 2 °C and pH 5.5. Under these conditions, we observe several resonances between 12.00 and 14.00 ppm where we would expect to find the thymine and guanine imino protons and three relatively broad resonances between 8.8 and 10.3 ppm. Integration of all these resonances show the existence of at least two species in solution.

Major Species. The resonances have been assigned from analysis of a NOESY spectrum recorded in 90% H_2O and 10% D_2O at 2 °C with 60 and 250 ms mixing times. Two regions are shown: imino−imino (Figure 4A) and imino-amino/H2/H5 interactions (Figure 4B). We have assigned the T13 (13.66 ppm) and T17 (13.79 ppm) imino protons of the major species from the NH−CH_3 intraresidue NOEs (not shown). Strong cross-peaks with A8H2 at 7.60 ppm and A4H2 at 7.58 ppm, respectively, confirm this assignment, Figure 4B. In Figure 4A, T13NH shows interactions with two G imino protons at 12.60 and 12.75 ppm, which must correspond to G12 and G14. The relative assignment is given by the observed interactions with C7H5 and C9H5. From the imino−imino and imino−amino/H2/H5 inter- and intra-base pair NOEs the complete assignment for all the Watson−Crick base pairs follows, with the exception of that A1·T20 for which the imino proton is not observed. All the interactions reported are typical of B DNA. The G15 imino proton shows NOEs with the amino protons of C6 on the opposite strand, with the previously assigned F16H2 proton and a broad resonance at 10.15 ppm, peak A.

At this stage, only G5NH, the amide proton of F16 and a possible amino resonance of G5 are not assigned for the major species. To assign the resonance at 10.15 ppm, we have taken two observations into consideration. This resonance does not correspond to a nonexchangeable proton. It must correspond to an imino proton but its chemical shift is not characteristic for an imino proton hydrogen bonded with a nitrogen, but rather with an oxygen atom. This imino proton shows NOEs with the nonexchangeable A4H2 and F16H2 protons of the major species and a resonance at 5.53 ppm, peak B in Figure 5. The resonance at 5.53 ppm does not correspond to any nonexchangeable proton or one of the amino protons already assigned. It must arise from a non-hydrogen-bonded G amino group in rapid rotation on a proton NMR time scale. This is only possible for an amino group of a guanine which is not involved in a Watson−Crick base pair (*25*). It must correspond to G5. We can thus unambiguously assign the 10.15 ppm resonance to the G5 imino proton. In Figure 5, we observe a very strong interaction between the F16H2 proton and a resonance at 8.24 ppm, peak C, and this resonance also shows an NOE with the G15H1′ proton, peak D. The resonance at 8.24 ppm does not correspond to a nonexchangeable proton and it is not an amino resonance because for the major species all of them are already assigned. Moreover, the chemical shift at 8.24 ppm could not correspond to an imino proton of guanine but is compatible with that of the amide proton of formamide

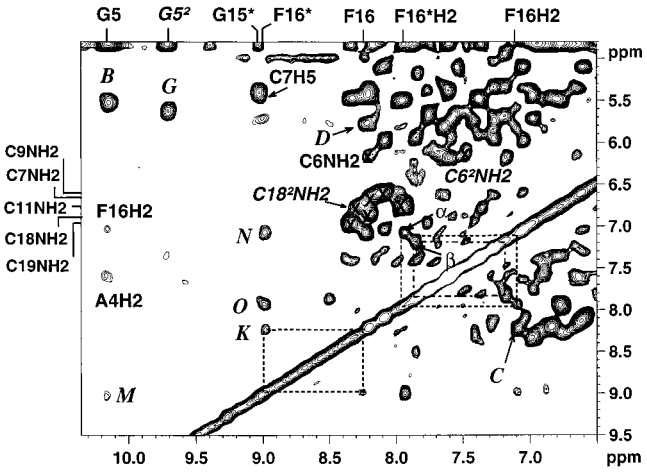

FIGURE 5: Expanded contour plot of a NOESY spectrum (150 ms mixing time) recorded at 500 MHz at 2 °C in H_2O. Peaks α and *β* are intraresidue ROESY interactions between the nonexchangeable protons for F16H2 and C6H6. The cross-peaks marked with an X correspond to intra amino−amino interactions. Labeled cross-peaks are described in the text.

(*9*) and is assigned as such. All of these observations in H_2O confirm that G15 and the formamide are intrahelical.

Minor Species. The unassigned peaks must arise from the minor species or from chemical exchange. We will describe the pure NOE interactions and, below, those involving chemical exchange observed in ROESY spectra. As described for the major species the assignments for the Watson−Crick base pairs including those adjacent to the mismatch is straightforward and the chemical shift differences are small. We will focus on interactions involving $G5^2$ and $F16^2$. We have identified the $G5^2$ imino proton at 9.71 ppm, which shows NOEs with the imino protons of $T17^2$ and $G15^2$, peaks E and F in Figure 4B, demonstrating that it lies inside the helix. It also shows a NOE at 5.63 ppm, peak G in Figure 5, assigned to its own amino group. This amino group also shows an NOE with $G15^2$NH, peak H in Figure 4B. The $F16^2$NH proton is now observed at 13.22 ppm, shifted well downfield from that of the major species, and shows NOEs with $A4^2$NH2 and $T17^2$H6. Finally, we observe a weak interaction between $F16^2$H2 and $G15^2$NH. Both $G5^2$ and $F16^2$ must lie within the helix but the $F16^2$ amide proton must now be hydrogen bonded to nitrogen, while the $G5^2$NH proton does not form a hydrogen bond with nitrogen.

Direct Chemical Exchange and Exchange Relayed NOE. An off-resonance ROESY spectrum in 90% H_2O and 10% D_2O at 2 °C with a 60 ms mixing time was recorded (not shown) in order to assign the remaining cross-peaks observed in the NOESY spectra. We observe the exchange cross-peaks (labeled α and *β* in Figure 5) of F16H2 and C6H6 already seen in D_2O and three others with exchangeable protons of the major species, peaks I (Figure 4A) and J (Figure 4B), the imino protons of G14 and G15, respectively, and peak K (Figure 5) that of the amide proton of F16. The G14NH and G14*NH chemical shifts are both characteristic of imino protons bound to nitrogen. The G15NH chemical shift is characteristic of an imino proton hydrogen bonded to a nitrogen and the G15*NH chemical shift corresponds to an imino proton hydrogen bonded to an oxygen atom.

In the NOESY spectrum, Figure 4B, we observe exchange relayed interactions between G14*NH and the H5 and NH2 protons of C7 and the A8H2 proton. We also observe in Figure 4A a NOE cross-peak, relayed by exchange, between G14* and T13, peak L. The T13*NH shows interactions with both A8H2 and A8NH2 protons. In Figure 5, NOE cross-peaks are relayed by exchange between the imino protons of G5 and G15*, peak M, and between the F16* amide proton and the F16H2 proton, peak N. We observe two direct NOEs within the new species. The F16* amide proton shows a cross-peak with its own H2 proton, peak O in Figure 5, and one between G14* and T13*, peak P in Figure 4A. We have only observed exchange or exchange relayed cross-peaks for protons of the major species.

^{31}P NMR Spectra. To search for additional structural information about the backbone chain, ^{31}P NMR spectra of the duplex at different temperatures and pH were recorded. The ^{31}P chemical shift dispersion is always small, ca. 0.9 ppm, with poor resolution and is in the normal chemical shift range for B_I conformations.

Molecular Modeling Studies. The rotation of the amide bond of the formamide leads to trans and cis isomers. In a previous NMR study (*9*) it has been shown that the isomer ratio is 3:2 and that the major species is the trans isomer. In this NMR study, where the formamide is incorporated in a DNA 10-mer duplex, we observe two species for the formamide in solution at 21 °C with the same ratio. The major species is trans, and the formamide of both species (cis and trans) is intrahelical. The sugar conformation is predominantly C2′-endo for the trans isomer but cannot be directly determined for the cis isomer, whereas all the other sugar puckers are observed to be C2′-endo. Both base pairs adjacent to G5·F16 are in a Watson−Crick conformation for the two species. In H_2O, the G5NH proton for both species is hydrogen bonded to an oxygen atom and its amino protons are not involved in strong hydrogen bonding because they are observed in H_2O as a single resonance. Thus, there is only one possibility to form a hydrogen bond between the formamide and the guanine on the opposite strand for both isomers, G5NH···F16O2. However, for the cis isomer, we have observed that the amide proton of the formamide is hydrogen bonded to a nitrogen atom. This hydrogen bonding could only occur between the formamide and the adjacent N3 of G15.

Most of the NMR results (intra- and interresidue interactions, sugar puckers, and hydrogen bonding) are characteristic of a B-DNA structure. Therefore, the initial structures were obtained by incorporating the formamide in the trans or in the cis conformation into a classical Arnott structure with all sugar puckers in C2′-endo conformation. A third duplex in which the formamide sugar pucker of the minor species (cis) is C3′-endo was also constructed. The NMR data does not unambiguously define the spatial position of the formamide, thus to determine the sugar/formamide orientation of the isomers, we have generated for the three duplexes families of structures. The glycosidic torsion angle (*χ*) was varied through the range from the anti to the syn conformation in steps of 15° yielding 24 structures for each duplex. The NMR distances obtained between the formyl proton of the formamide and the sugar atoms (H2−H1′, H2−H2′/H2″, and H2−H3′) gives the sugar/formamide orientation. Most of the generated structures can be rejected for one of two reasons: there is no global qualitative agreement with the NMR data or there is no possibility to form a hydrogen bond

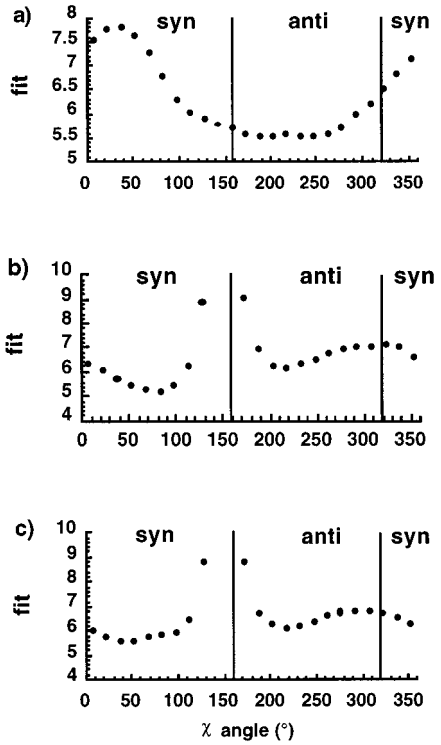

FIGURE 6: Plot of the fits from model building versus χ angle for the duplex containing (a) the trans isomer of F16, (b) the cis isomer which has a C2′-endo conformation and (c) the cis isomer with a C3′-endo conformation.

between the formamide and a base either on the same or on the opposite strand.

Major Species. The fit of all generated structures where the formamide is in a trans conformation is shown in Figure 6, panel a. Good quantitative agreement with the NMR data for χ angles of the formamide in the anti range (172−263°) is observed. Only distance constraints which furnish structural information were used, such as sugar puckers, base orientation, and internucleotide distances. We have obtained seven possible models in good agreement with the NMR distance constraints from the NOE build-up curves. There is a good hydrogen bond between the G5 imino proton and the F16 oxygen atom. The seven retained constructions were minimized as described in Materials and Methods. For all the minimized structures the formamide χ angle has a value of ca. 225° and the interproton distances are very close to the NMR-derived distances. The minimized structure with the best fit and the lowest energy, Figure 7A, was used as the starting structure for MD runs of 500 ps. We have carried out two MD runs with NMR distance constraints: MDI with a reinforcement (5 kcal mol^{-1} Å^{-2}) of the G5NH···F16O2 hydrogen bond and MDII without. We observe similar time-averaged results during the two MD runs. The mean structure is described in Table 1 in terms of average torsion angles for the three central base pairs. The angles are close to their classical B DNA values and sugar puckers are in the C2′-endo conformation. The length of the G5NH···F16O2 hydrogen bond observed during MDII as a function of time fluctuates, 2.19 ± 0.4 Å. This compares with a value of 2.1 ± 0.2 Å for normal nonterminal base pairs. During the MD runs, we observed that only one hydrogen bond can stabilize the G5·F16 base pair. The G5 amino group is never involved in a hydrogen bond, in agreement with our NMR results.

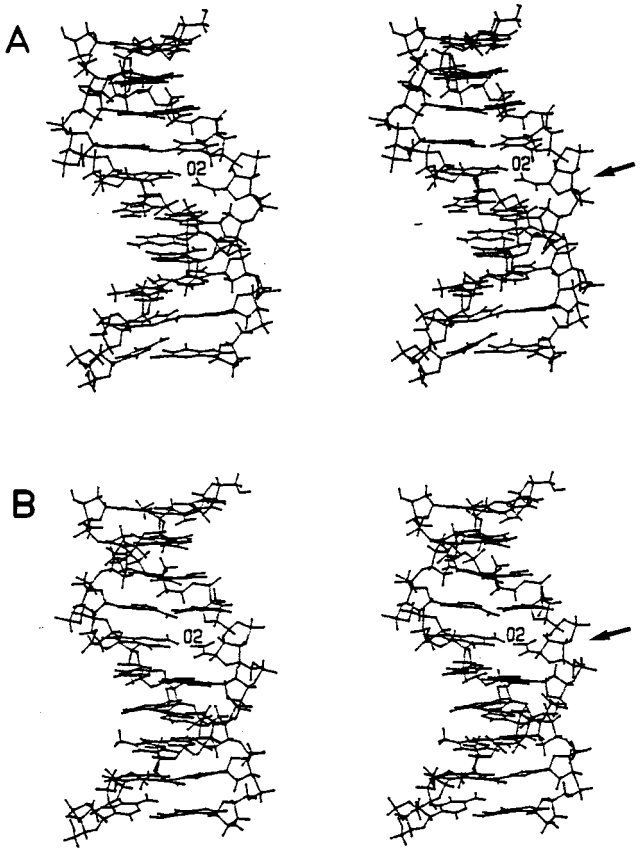

FIGURE 7: Stereoscopic view of the duplex containing (A) the trans isomer of F16 after minimization and (B) the cis isomer of F16 took during the MD runs. The arrow indicates the formamide.

Table 1: Average Torsion Angles for the 500 ps MD Run for the Three Central Base Pairs of the Duplex with the Formamide in the trans Conformation

residue	α	β	γ	δ	ε	ξ	χ
B-DNA	g$^-$	t	g$^+$		t	g$^-$	
A4	−73	171	59	139	171	−111	−112
G5	−65	169	53	146	173	−108	−113
C6	−68	172	57	139	173	−104	−118
G15	−67	171	63	157	166	−118	−111
F16	−71	167	59	143	168	−106	−135
T17	−45	160	12	157	171	−118	−97

Minor Species. For the minor species (cis), two series of models have been studied by constraining the angle δ. In the first series (A), all nucleotides are C2′-endo ($\delta = 144°$), while for the second series (B), only the formamide is C3′-endo ($\delta = 82°$). For each series, we have generated 24 structures by varying the χ angle. We have calculated the fit to the NMR data (144 structural distance constraints) for all these structures, as shown in Figure 6, panels b and c. For both series, the structures which have the best quantitative agreement with the NMR data, all have the formamide in a syn conformation where the χ angle varies from 53 to 113° for series A, panel b, and from 23 to 98° for series B, panel c. The variation observed between the fit of the best structures of the two series is small and not significant to determine unambiguously the sugar conformation of the formamide. By varying the χ angle, we obtain for the series A and B five and six structures, respectively, which are in good agreement with the NMR data. In particular, we observe that the G5NH···F16O2 hydrogen bonding is conserved, and

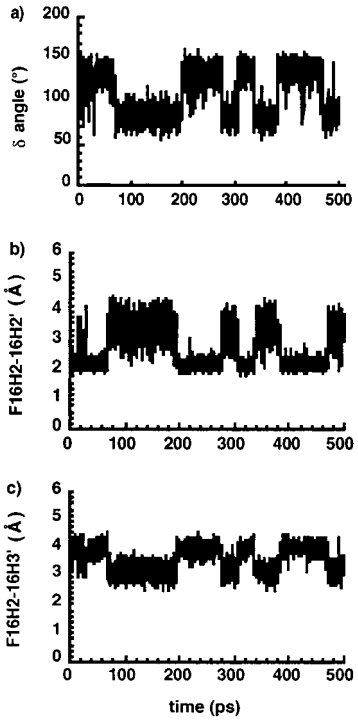

FIGURE 8: Torsion angles δ (a) and interproton distances H2−H2′ (b), H2−H3′ (c), for the cis isomer of F16 residue as a function of time.

it is possible to form a hydrogen bond between the F16NH and the G15N3. Energy minimization has been carried out for these models without constraining either of the G5NH···F16O2 and F16NH···G15N3 hydrogen bonds. For all the structures retained in each series, the formamide χ angle does not change significantly: it is ca. 62° for the series A and ca. 45° for the series B. During energy minimization, the G5NH···F16O2 and the F16NH···G15N3 hydrogen bonds are conserved.

The structures of each series fitting best the NMR data were used as the starting conformations for the MD runs with all NMR distance constraints and with a weak constraint on the torsion angle δ for all the sugars except F16. The F16O2···G5NH hydrogen bond was weakly constrained at 2 Å while the F16NH···G15N3 hydrogen bond was not. The two starting structures show no significant differences in terms of interproton distances and torsion angles. Thus, we will focus our description on the duplex with the formamide in the C2′-endo conformation as the starting point (Figure 7B). All sugar puckers were found in the C2′-endo conformation except for that of F16 which is in equilibrium between the C2′- and C3′-endo conformations. This is clearly seen in Figure 8: when a nucleotide changes from the C2′-endo to the C3′-endo conformation, the torsion angle δ decreases from ca. 144° to ca. 82°, Figure 8a. In Figure 8, panels b and c, the intraresidue distance F16H2−F16H2′ is short when F16H2−F16H3′ is long and vice versa and follows the variations of the sugar pucker of F16. The torsion angles of the three central base pairs are given in Table 2. The backbone angles of the three central base pairs are close to their classical B-DNA values except for a deviation in one or both of the ϵ and ξ angles for the formamide residue. The values of the angles show that the formation of the F16NH···G15N3 hydrogen bond gives rise to a modification

Table 2: Average Torsion Angles for the 500 ps MD Run for the Three Central Base Pairs of the Duplex with the Formamide in the cis Conformation[a]

residue	α	*β*	*γ*	*δ*	*ϵ*	*ξ*	*χ*
B-DNA	g^-	t	g^+		t	g^-	
A4	−71	170	61	137	172	−108	−112
G5	−66	171	59	146	173	−109	−110
C6	−69	171	59	134	169	−109	−119
C2′-*endo*							
G15	*−106*	170	*170*	164	165	−121	−127
F16	*−77*	166	59	*144*	167	*−76*	*59*
T17	*−51*	171	*28*	156	171	−121	−99
C3′-*endo*							
G15	*−106*	170	*170*	164	165	−121	−127
F16	*−77*	166	59	*82*	*130*	*−76*	*26*
T17	*−51*	171	*28*	156	171	−121	−99

[a] Angles are given for structures where the F16 deoxyribose is either in a stable C2′-endo or C3′-endo conformation. Numbers in italics correspond to torsion angles which differ significantly from B-DNA angles.

of the backbone structure relative to classical B-DNA. In B-DNA duplexes, the BI conformation with ϵ (t) and ξ (g^-) is the most common backbone conformation compared to BII where ϵ and ξ are g^- and t, respectively. Results obtained for the ϵ and ξ angles of F16 indicate that the backbone is not in a BII conformation but remains close to the BI conformation. Torsional angle changes in the deoxyribose phosphate backbone lead to variations in the ^{31}P chemical shifts of individual phosphates in duplex oligonucleotides (*26*, *27*). However, we do not observe a shifted peak in the ^{31}P spectra which would be expected for a such modification. This is probably due to a sensitivity problem caused by the weak population (40%) of the formamide residue in the cis conformation in the sample. The major difference is observed for the α and γ angles of G15 and T17. The variation of these values for G15 could be explained by the formation of a hydrogen bond with F16. The T17 angles are disturbed but remain close to the B-DNA values. The variation from the C2′-endo to the C3′-endo conformation does not significantly change the global energy of the system but only leads to a local perturbation of the duplex geometry and does not disrupt the hydrogen bonding of the adjacent base pairs. Moreover, whatever the sugar conformation, the second possible hydrogen bond, F16NH···G15N3, can be formed.

CONCLUSION

In this paper, we report a NMR and MD study of a DNA duplex containing a guanine facing the formamide lesion. We have determined by NMR that two conformations of formamide exist in solution. They correspond to the cis and trans isomers where the trans isomer is the major species. For both isomers, the formamide is stacked inside the helix, with a single hydrogen bond between the imino proton of G5 and the carbonyl group of the formamide, which induces only minor structural changes for the backbone conformation. For the cis isomer, the formation of one hydrogen bond with the guanine on the opposite strand is only possible if the formamide is in the syn conformation. In the cis conformation, it is possible to form an additional hydrogen bond between F16NH and the adjacent G15N3 without deforming the structure of a B-DNA helix. The MD results show that the structure with the cis isomer corresponds to a conformational equilibrium of the deoxyribose pucker of form-

amide. This conformational change of the deoxyribose does not lead to an important modification of the structure of the duplex. The energetic terms of both isomers are very similar.

We have previously shown that, if a urea lesion is incorporated in this sequence, no detailed NMR analysis could be performed (*28*) as the system gave rise to very broad resonances. In the present study, the two isomers could be identified in the NMR spectra. We also observe at low temperature a third species which is in exchange with the trans isomer, for which we have only limited conformational information. Its concentration is low and certain protons are in slow and others are in fast exchange. However, the chemical shifts of C6* strongly suggest that this base is extrahelical, as is the case for F16*. The limited number of NOEs observed means that other models cannot be excluded.

If a guanine is incorporated in a duplex structure opposite an AP site (*23*), an intrahelical form has been observed for both the guanine and the deoxyribose at low temperature. The mobility of the guanine residue increases rapidly with the temperature and is in a "melted out state" before denaturation of the duplex. The structures observed here with formamide are quite similar. Differences between the formamide and the AP site are observed on increasing the temperature as the formamide and the guanine are always stacked inside the helix until global melting. This can be explained by the presence of a hydrogen bond which stabilizes the intrahelical species because stacking energy interactions with the formamide side chain are probably weak.

The presence of an abasic lesion on a DNA template blocks the progression of DNA polymerase. When the abasic site occurs in a duplex, relative stacking interactions determine whether the structure will be extended or collapsed. The preferential incorporation of a guanine opposite the formamide lesion could be explained by the formation of a hydrogen bond of the imino proton with the formamide and the strong interactions between G and its neighbors on the G strand.

ACKNOWLEDGMENT

We thank Marc Le Bret for the kind gift of the programs OCL and MORCAD. We thank Dr. Wilhelm Guschlbauer for helpful discussions and comments on the manuscript.

REFERENCES

1. Téoule, R. (1987) *Int. J. Radiat. Biol. 51*, 573−589.
2. Cadet, J., Berger, M., Douki, T., and Ravanat, J. L. (1997) *Rev. Physiol. Biochem. Pharmacol. 131*, 1−87.
3. Guy, A., Duplaa, A. M., Ulrich, J., and Téoule, R. (1991) *Nucleic Acids Res. 19*, 5815−5820.
4. Loeb, L. A., and Preston, B. D. (1986) *Annu. Rev. Genet. 20*, 201−230.
5. Shida, T., Noda, M., and Sekiguchi, J. (1995) *Nucleic Acids Res. Symp. Ser. 34*, 87−88.
6. Falcone, J. S., and Box, H. C. (1997) *Biochim. Biophys. Acta 1337*, 267−275.
7. Bourdat, A. G., Gasparutto, D., and Cadet, J. (1999) *Nucleic Acids Res. 27*, 1015−1024.
8. Aurine, J., LaPlanche, L. A., and Rogers, M. T. (1964) *J. Am. Chem. Soc. 86*, 337−341.
9. Cadet, J., Nardin, R., Voituriez, L., Remin, M., and Hruska, F. E. (1981) *Can. J. Chem. 59*, 3313−3318.
10. Bodenhausen, G., Kogler, H., and Ernst, R. R. (1984) *J. Magn. Reson. 58*, 370−388.
11. Piotto, M., Saudek, V., and Sklénar, V. (1992) *J. Biomol. NMR 2*, 661−665.
12. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
13. Desvaux, H., Berthault, P., Birlirakis, N., and Goldman, M. (1993) *C. R. Acad. Sci. (Paris) 317*, 19−25.
14. Desvaux, H., Berthault, P., Birlirakis, N., and Goldman, M. (1994) *J. Magn. Reson. 108*, 219−229.
15. Arnott, S., Smith, P. J. C., and Chandrasekaran, R. (1976) *Atomic coordinates and molecular conformations for DNA-DNA, RNA-DNA and DNA-RNA helices*, Vol. 2, 3rd ed., CRC Press, Cleveland, OH.
16. Cuniasse, P., Sowers, L. C., Kaplan, B., Goodman, M. F., Cognet, J. A. H., Le Bret, M., and Fazakerley, G. V. (1989) *Biochemistry 28*, 2018−2026.
17. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1995) *AMBER 4.1*, University of California, San Francisco.
18. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, G. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
19. Le Bret, M., Gabarro-Arpa, J., Gilbert, J. C., and Lemarechal, C. (1991) *J. Chem. Phys. 88*, 2489−2496.
20. Gervais, V., Cognet, J. A. H., Guy, A., Cadet, J., Téoule, R., and Fazakerley, G. V. (1998) *Biochemistry 37*, 1083−1093.
21. Boulard, Y., Cognet, J. A. H., Gabarro-Arpa, J., Le Bret, M., Carbonnaux, C., and Fazakerley, G. V. (1995) *J. Mol. Biol. 246*, 194−208.
22. Gabarro-Arpa, J., Cognet, J. A. H., and Le Bret, M. (1992) *J. Mol. Graphics 10*, 166−173.
23. Cuniasse, P., Sowers, L. C., Eritja, R., Kaplan, B., Goodman, M. F., Cognet, J. A. H., and Le Bret, M. (1987) *Nucleic Acids Res. 15*, 8003−8022.
24. Schmitz, U., Zon, G., and James, T. L. (1990) *Biochemistry 29*, 2357−2368.
25. Faibis, V., Cognet, J. A. H., Boulard, Y., Sowers, L. C., and Fazakerley, G. V. (1996) *Biochemistry 35*, 14452−14464.
26. Gorenstein, D. G. (1992) *Methods Enzymol. 211*, 254−286.
27. Roontga, V. A., Jones, C. R., and Gorenstein, D. G. (1990) *Biochemistry 29*, 5245−5258.
28. Gervais, V., Guy, A., Téoule, R., and Fazakerley, G. V. (1992) *Nucleic Acids Res. 20*, 6455−6460.

BI9928849