Photochemistry of 4-Thiothymine Derivatives in the Presence of *N*-9-Substituted-Adenine Derivatives: Formation of *N*-6-Formamidopyrimidines

Carole Saintomé,[†,‡] Pascale Clivio,*[,†] Alain Favre,[‡] and Jean-Louis Fourrey*[,†]

Institut de Chimie des Substances Naturelles, CNRS, 91198 Gif-sur-Yvette Cedex, France, and Institut Jacques Monod, CNRS-Université Paris VII, 2 Place Jussieu, 75251 Paris Cedex 05, France

Received June 26, 1997[⊗]

UV irradiation of aqueous solutions containing either 4-thiothymin-1-ylacetic acid (**1b**) and adenosine (**2a**), 4-thiothymidine (**1a**) and adenin-9-ylacetic acid (**2b**), or **1b** and **2b** led to 4,5-diamino-6-formamidopyrimidine (N-6-Fapy-Ade) derivatives as observed after irradiation of a mixture of **1a** and **2a** (*J. Am. Chem. Soc.* **1996**, *118*, 8142−8143). These new observations demonstrate that the replacement of one or both nucleoside sugar residues by a carboxymethyl group does not affect the regioselective course of the photochemical reaction. The thermal decomposition of **3a** that resulted from irradiation of **1a** in the presence of **2a**, was examined along with its behavior under mild alkaline conditions. Finally, irradiation of *N*-3-methyl-4-thiothymidine (**6a**) in the presence of adenosine gave the *N*-3-methylcytidine derivative **7**.

Introduction

Cyclobutane pyrimidine dimers and (6−4) pyrimidine pyrimidone photoproducts are recognized as the major deleterious adducts generated by UV irradiation of DNA.[1] A great number of model studies have been designed to examine their structural effect at the molecular level in nucleic acids as well as their formation and repair mechanisms.[1,2] In this context, the remarkable photochemical properties of the sulfur analogues of nucleic acid bases (4-thiouracil, 4-thiothymine) have successfully been utilized to study in detail the mechanism of formation and photoreversal of (6−4) adducts.[3,4] Purine bases are also suspected to lead, under UV radiation, to potent toxic and mutagenic lesions.[1b,5] Among the purine photoproducts, formamidopyrimidine (Fapy) derivatives have been identified recently.[6] This latter type of DNA damage, also induced under oxidative conditions (including ionizing radiations),[7,8] alkylating agents,[9] or alkaline conditions,[10] can be repaired by the formamidopyrimidine glycosylase (fpg) protein.[11] Later, during our examination of the light-induced reaction between 4-thiothymidine and adenosine, we observed the formation of a unique adduct (**3a**) having a formamidopyrimidine−adenine (Fapy-Ade) structure.[12] It occurred to us that such an easy access to this type of compound might be very helpful in the field of DNA lesion studies. Indeed, in regard to the strategy of site specific incorporation of lesions in oligonucleotides, successfully developed for cyclobutane pyrimidine dimers[13] and (6−4) pyrimidine photoproducts[14] that allowed a number of biophysical and biological studies,[1a,15] formamidopyrimidine-modified oligonucleotides could be used to examine, at the molecular level, either DNA repair mechanisms or to determine the factors responsible for the toxicity and mutagenicity of this type of adducts.

Herein, in view of elucidating the respective role of each sugar residue on the course of the reaction between 4-thiothymidine (**1a**) and adenosine (**2a**) to give **3a**, we report the photochemical behavior of 4-thiothymin-1-ylacetic acid (**1b**) in the presence of adenosine (**2a**), as well as that of 4-thiothymidine (**1a**) in the presence of adenin-9-ylacetic acid of (**2b**) and also of **1b** in the

[†] Institut de Chimie des Substances Naturelles.

[‡] Institut Jacques Monod.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1997.

(1) (a) Taylor, J.-S. *Pure Appl. Chem.* **1995**, *67*, 183−190. (b) Davies, R. J. H. *Biochem. Soc. Trans.* **1995**, *23*, 407−418. (c) Carell, T. *Chimia* **1995**, *49*, 365−373.

(2) For selected examples see: (a) Cadet, J.; Vigny, P. In *Bioorganic Photochemistry: The Photochemistry of Nucleic Acids*; Morrison, H., Ed.; Wiley and Sons: New York, 1990; Vol. 1, pp 1−272. (b) Begley, T. P. *Acc. Chem. Res.* **1994**, *27*, 394−401. (c) Fenick, D. J.; Carr, H. S.; Falvey, D. E. *J. Org. Chem.* **1995**, *60*, 624−631. (d) Carell, T.; Epple, R.; Gramlich, V. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 620−623. (e) Prakash, G.; Falvey, D. E. *J. Am. Chem. Soc.* **1995**, *117*, 11375−11376.

(3) (a) Fourrey, J.-L.; Gasche, J.; Fontaine, C.; Guittet, E.; Favre, A. *J. Chem. Soc., Chem. Commun.* **1989**, 1334−1336. (b) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *J. Am. Chem. Soc.*, **1991**, *113*, 5481−5483. (c) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *Tetrahedron Lett.* **1992**, *33*, 1615−1618. (d) Clivio, P.; Fourrey, J.-L.; Szabo, T.; Stawinski, J. *J. Org. Chem.* **1994**, *59*, 7273−7283.

(4) (a) Liu, J.; Taylor, J.-S. *J. Am. Chem. Soc.* **1996**, *118*, 3287−3288. (b) Clivio, P.; Fourrey, J.-L. *J. Chem. Soc., Chem. Commun.* **1996**, 2203−2204.

(5) (a) Duker, N. J.; Gallagher, P. E. *Photochem. Photobiol.* **1988**, *48*, 35−39. (b) Gallagher, P. E.; Duker, N. J. *Photochem. Photobiol.* **1989**, *49*, 599−605. (c) Zhao, X.; Taylor, J.-S. *Nucleic Acids Res.* **1996**, *24*, 1561−1565.

(6) Doetsch, P. W.; Zastawny, T. H.; Martin, A. M.; Dizdaroglu, M. *Biochemistry* **1995**, *34*, 737−742.

(7) (a) Hems, G. *Nature* **1958**, *181*, 1721−1722. (b) Hems, G. *Radiation Res.* **1960**, *13*, 777−787. (c) Teoule R. *Int. J. Radiat. Biol.* **1987**, *51*, 573−589. (d) Steenken, S. *Chem. Rev. (Washington, D.C.)* **1989**, *89*, 503−520. (e) Boiteux, S.; Gajewski, E.; Laval, J. Dizdaroglu, M. *Biochemistry* **1992**, *31*, 106−110.

(8) Dizdaroglu, M. *Mutat. Res.* **1992**, *275*, 331−342.

(9) (a) Tomasz, M.; Lipman, R.; Verdine, G. L.; Nakanishi, K. *J. Am. Chem. Soc.* **1985**, *107*, 6120−6121. (b) Tomasz, M.; Lipman, R.; Lee, M. S.; Verdine, G. L.; Nakanishi, K. *Biochemistry* **1987**, *26*, 2010−2027. (c) Hertzog, P. J.; Lindsay Smith, J. R.; Garner, R. C. *Carcinogenesis* **1982**, *3*, 723−725. (d) Wood, M. L.; Lindsay Smith, J. R.; Garner, R. C. *Cancer Res.* **1988**, *48*, 5391−5396. (e) Chetsanga, C. J.; Polidori, G.; Mainwaring, M. *Cancer Res.* **1982**, *42*, 2616−2621.

(10) Lönnberg, H.; Lehikoinen, P. *J. Org. Chem.* **1984**, *49*, 4964−4969 and references cited herein.

(11) (a) Boiteux, S. *J. Photochem. Photobiol. B: Biol.* **1993**, *19*, 87−96. (b) Demple, B.; Harrison, L. *Annu. Rev. Biochem.* **1994**, *63*, 915−948. (c) Bhagwat, M.; Gerlt, J. A. *Biochemistry* **1996**, *35*, 659−665.

(12) Saintomé, C.; Clivio, P.; Favre, A.; Fourrey, J.-L.; Riche, C. *J. Am. Chem. Soc.* **1996**, *118*, 8142−8143.

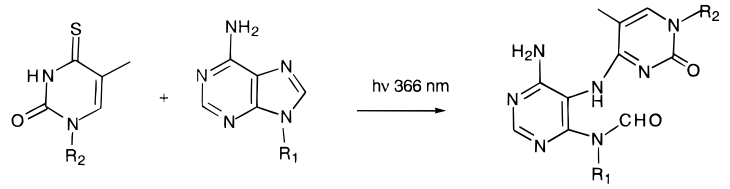
(13) Taylor, J.-S.; Brockie, I. R.; O'Day, C. L. *J. Am. Chem. Soc.* **1987**, *109*, 6735−6742.

(14) Iwai, S.; Shimizu, M.; Kamiya, H.; Ohtsuka, E. *J. Am. Chem. Soc.* **1996**, *118*, 6742−6743.

(15) (a) Taylor, J.-S. *Acc. Chem. Res.* **1994**, *27*, 76−82. (b) Fujiwara, Y.; Iwai, S. *Biochemistry* **1997**, *36*, 1544−1550.

S0022-3263(97)01162-6 CCC: $14.00

Scheme 1



1a R_1= Deoxyribosyl
1b R_1= CH_2COOH

2a R_2= Ribosyl
2b R_2=CH_2COOH

3a R_1= Ribosyl; R_2= Deoxyribosyl
3b R_1= Ribosyl; R_2= CH_2COOH
3c R_1= CH_2COOH; R_2= Deoxyribosyl
3d R_1= R_2= CH_2COOH
5 R_1= 2',3',5'-O-Triacetylribosyl;
R_2= 3',5'-O-Diacetydeoxyribosyl

presence of **2b**. In addition, we have examined the mechanism of the thermal degradation of **3a**.[12] Furthermore, in an attempt to get insight into the mechanism of formation of **3a**, we have studied the photochemical reaction between **6a**, the *N*-3-methyl derivative of **1a**, and adenosine.

Results and Discussion

Isolation of Photoproducts 3a−d and 7. Photochemical reactions were monitored by the disappearance, on the UV spectrum, of the thiocarbonyl absorption maximum (ca. 335 nm) of **1a,b** and **6a**. Photoproducts **3a−d** were obtained as unique adducts by irradiation of a 1:1 mixture of **1a** and **2a** (aqueous solution), and of mixtures of **1b** and **2a**, **1a** and **2b**, and **1b** and **2b**, respectively, as inferred from the ^{1}H NMR analysis of the crude irradiation products. For these three latter irradiation experiments, the concentration of the adenine derivatives (**2a,b**) was twice that of the 4-thiothymine derivatives (**1a,b**) in order to accelerate photoproduct formation. In the same concentration conditions, photolysis of **6a** in the presence of **2a** led to a 50:50 mixture (^{1}H NMR) of photoproducts **7** and **6b**.[16]

Photoproducts **3a** and **7** were purified by medium-pressure reverse phase (RP 18) chromatography using a gradient of acetonitrile in water. Photoproducts **3b−d** were purified by reverse phase HPLC eluting with a gradient of acetonitrile in an aqueous solution of triethylammonium acetate.

Formation of photoproduct **3a** was by far the most efficient.[12] Photoadducts **3b−d** and **7** were isolated after purification in 20−27% yield.

Characterization of Photoproducts 3b−d. The UV spectrum of **3b−d** (λ_{max} ca. 295 nm) was similar to that of **3a**, immediately suggesting a 5-methylcytosine structure for these adducts. The positive FAB mass spectra of the three compounds **3b−d** displayed a quasimolecular ion at *m*/*z* 474, 458, and 400 [$(M + Na)^+$], respectively, indicating that all of them were formed by means of a condensation reaction between the 4-thiothymine and adenine derivatives implicating the replacement of the sulfur atom by an oxygen atom. From the comparison of the ^{1}H NMR spectra of **3b−d** with that of **3a**, it was clear that **3b−d** contained the same *N*-4-(4-amino-6-formamidopyrimidin-5-yl)-5-methylcytosine motif. Indeed, the ^{1}H NMR spectra of **3a−d** displayed three singlets in the downfield region corresponding to H-6, H-2, and H-8 (see Experimental Section). At room temperature, the H-8 proton signal of **3a,b** appeared as a very broad signal coalescing around 70 °C. This indicated that **3a** and **3b** consisted of a slow equilibrium mixture of various species, with respect to the NMR time scale, which could be caused by the restricted rotation of the formamido group.[9b,17] Moreover, since the signals of protons H-8 of **3c** and **3d** appeared as well-resolved singlets, it could be suggested that the ribose moiety was implicated in this restricted rotation.

^{13}C NMR data of **3b−d** supported the proposed structures. The corresponding spectra were analyzed by analogy with those of **3a** and its peracetyl derivative **5** (Table 1) for which unambiguous HMBC correlations were observed in methanol-d_4. Confirmation of the opening of the adenosine imidazole ring to give a *N*-6-formyl-4,5,6-triaminopyrimidine was established by the presence of a characteristic C-8 sp^2 carbon around 166 pm that correlates with H-1′ of ribose on the HMBC spectrum of **5**. Furthermore, chemical shifts of C-4, C-5, and C-6 of the pyrimidine nucleus of **3b−d** were in very good agreement with those of **3a** and of other 5-methylcytosine structures.[18] The 6 ppm deshielding of the signals due to C-6 of the thymine base in **3b** and **3d** compared to **3a** and **3c** was attributed to the replacement of the deoxyribose residue by a carboxymethyl group. Signal broadening was noticed at positions corresponding to C-4, C-5, and C-1′ of the adenosine part of **3a** and **3b**, and these observations corroborated the hypothesis of a restricted rotation of the formamido group.

Mechanism of Formation of 4 by Thermal Degradation of 3a. Since the formation of *N*-4-(4-amino-6-formamidopyrimidin-5-yl)-5-methylcytosine between a 4-thionopyrimidine and a purine base appeared to be quite general, **3b−d** being obtained as unique products, we decided to investigate more precisely the previously observed thermal degradation of **3a**[12] that was found to undergo both deformylation and cleavage of the *N*-ribosidic bond in neutral aqueous solution to lead to the 4,5,6-triaminopyrimidine derivative **4**. As thermally induced decomposition of purine derivatives in alkaline conditions is well documented,[10] we examined first the stability of **3a** in alkaline solution by ^{1}H NMR spectroscopy. Thus, treatment of an aqueous solution of **3a**, at room temperature, with K_2CO_3 (pD ca. 10) led within 2 h to (1) a decrease of the H-2 and H-8 signals (overlapped singlets) which was accompanied by the appearance of a signal at 8.1 ppm (attributed to H-2 of **8**) and (2) a

(16) Clivio, P.; Fourrey, J.-L.; Gasche, J.; Favre, A. *J. Chem. Soc., Perkin Trans. 1* **1992**, 2383−2388.

(17) (a) Humphreys, W. G.; Guengerich, F. P. *Chem. Res. Toxicol.* **1991**, *4*, 632−636. (b) Raoul, S.; Bardet, M.; Cadet, J. *Chem. Res. Toxicol.* **1995**, *8*, 924−933.

(18) Shaw, A. A.; Sheltlar, M. D. *J. Am. Chem. Soc.* **1990**, *112*, 7736−7742.

Table 1. ^{13}C NMR Chemical Shifts (62.90 MHz)* of 3a−d, 7 (in D_2O relative to dioxane 67.8 ppm), 5 (in CD_3OD set at 49.0 ppm), and 4 (in DMSO-d_6 set at 39.7 ppm)

	3a	5**	4	7	3b	3c	3d
C_6T	140.6	139.5	138.4	136.3	146.7	140.7	146.8
C_2T	157.9	157.6	156.1	152.3	159.0	158.0	159.0
C_4T	164.0	164.2	163.7	154.4	163.9[a]	163.3[a]	163.3[a]
C_5T	106.5	105.6	104.7	108.9	106.1	106.6	106.0
CH_3	13.6	13.5	13.9	16.6	13.4	13.6	13.3
$C_{1'}T$	87.3	87.8	85.6	87.0[a]		87.9[b]	
$C_{2'}T$	40.6	38.8	40.7	39.9		40.6	
$C_{3'}T$	71.6	75.8	70.9	71.5[b]		71.6	
$C_{4'}T$	87.9	89.7	87.4	87.7		87.3[b]	
$C_{5'}T$	62.5	65.0[a]	61.8	62.3[c]		62.3	
C_2A	157.9	157.6	155.3	153.2	157.8	157.5	157.4
C_4A	152.9, 155.6	156.6	159.8	150.9	154.5 br	156.6	156.5
C_5A	113.0, 115.0	111.1	97.1	126.1	114.6 br	110.6	110.9
C_6A	164.0	164.2	159.8	159.2	164.3[a]	164.4[a]	164.6[a]
C_8A	166.0	164.2		143.9	166.2	165.5	165.6
$C_{1'}A$	91.5, 93.9	90.8		89.7	91.6, 94.0		
$C_{2'}A$	73.1	74.0[b]		74.9	73.3		
$C_{3'}A$	70.7	71.1[b]		71.8[b]	70.8		
$C_{4'}A$	84.6	80.1		87.2[a]	84.8		
$C_{5'}A$	62.4	64.0[a]		62.3[c]	62.2		
CH_2T					54.2		54.1
CH_2A						48.7	48.8
COOH					175.9	175.6	176.0
NCH_3				32.4			

* For convenience, numbering of the parent adenosine is maintained. ** Acetate signals: 20.4; 20.5; 20.7; 20.9 (×2) and 172.2 (×2); 172.1; 171.3; 171.2 ppm. [a−c] Interchangeable attributions within a column.

complete disappearance of the anomeric H-1′ signal of the ribose part, concomitantly with the apparition of two anomeric doublets at 5.86 ppm (J = 4.3 Hz) and 5.64 ppm (J = 5.5 Hz) (integrating for one proton) corresponding to α- and β-ribofuranosyl derivatives **8a** and **8b**, respectively.[10] The LSIMS spectra of both compounds displayed the expected quasimolecular peaks at m/z 504 ($M + Na^+$) and m/z 520 ($M + K^+$). After 28 h, two additional anomeric doublets characteristic of the two α- and β-pyranosyl derivatives **8c** and **8d** were present at 5.48 ppm (J = 3.5 Hz) and 5.35 ppm (J = 9.2 Hz).[10] After 7 days, only the signals of **8c** and **8d** remained (quasimolecular ions at m/z 504 ($M + Na^+$) and 520 ($M + K^+$) on LSIMS spectra) demonstrating that deformylation had occurred first.

We then studied the thermally induced decomposition of **3a**[12] at neutral pD using variable temperature ^1H NMR spectroscopy. Initial deformylation should lead to the above mentioned intermediates **8a**−**d**. Conversely, initial loss of the ribose residue should be accompanied by the formation of intermediate **9**.

Heating a solution of **3a** (10 mg in 500 μL of D_2O) at 90 °C for 4 h 15 min led (1) to the decrease of the signal at 8.4 ppm (overlapped H-2/H-8 singlets) which was accompanied by the progressive appearance of a singlet at 8.1 ppm and (2) to the disappearance of the anomeric H-1′ of ribose associated with the progressive apparition of a set of four anomeric proton doublets at δ 5.3 (J = 4 Hz), 5.2 (J = 2 Hz), 4.9 (J = 6 Hz), and 4.8 (J = 2 Hz) ppm, corresponding to the release of free D-ribose as its α- and β-ribofuranosyl and pyranosyl isomers.[19] Accordingly, the absence of any of the intermediates **8** during the thermal decomposition of **3a** was in favor of pathway A. However, we could not exclude pathway B due to the known very fast thermal hydrolysis of *N*-arylglycosamines analogues that proceeds via a Schiff base.[10] Thus, a mixture of **8a** and **8b** in D_2O (pD ca. 7, after careful addition of K_2CO_3) was heated for 20 min at 90 °C. The only result of this treatment, as observed by ^1H NMR spectroscopy, was the partial conversion of **8a** and **8b** to the pyranosyl forms **8c** and **8d** and not the release of D-ribose; however, this experiment did establish that **8a** and **8b** were thermally stable under neutral conditions and ruled out pathway B.

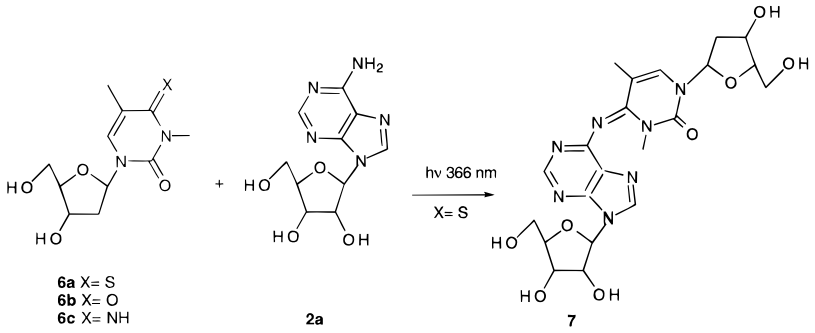
The easy and clean transformation of **3a** into **4** is reminiscent of the release of Fapy-G from ionizing radiation-treated DNA upon warming at pH 7.[20] This underlines the interest of this reaction pathway in view of mapping to eventual N7−C8 purine cross-linked sites in photolabeling experiments[21] and for designing artificial specific endonucleases.[22]

Characterization of Photoproduct 7. In an attempt to confirm the postulated mechanism leading to **3a**,[12] we have studied the photochemistry of *N*-3-methyl-4-thiothymidine (**6a**) in the presence of adenosine (**2a**), reasoning that the suggested thiaazetidine intermediate should be trapped as was the thietane intermediate in the case of Tps^4T.[3c,d] Unexpectedly, the compound which was formed in this reaction was attributed structure **7**. The UV spectrum of **7** (λ_{max} 304 nm, ϵ 14817 mol^{-1} cm^{-1} dm^3, H_2O, pH 7) indicated the loss of the thiocarbonyl chromophore. Exact mass measurement of its pseudomolecular ion (m/z 506.1975, $(M + H)^+$) in the positive mode FAB mass spectrum indicated that **7** resulted from the condensation of **6a** and **2a** followed by H_2S elimination. Inspection of the ^{13}C NMR data (Table 1) provided key structural informations. Assignment of all the carbons were made through the analysis of the 2D one

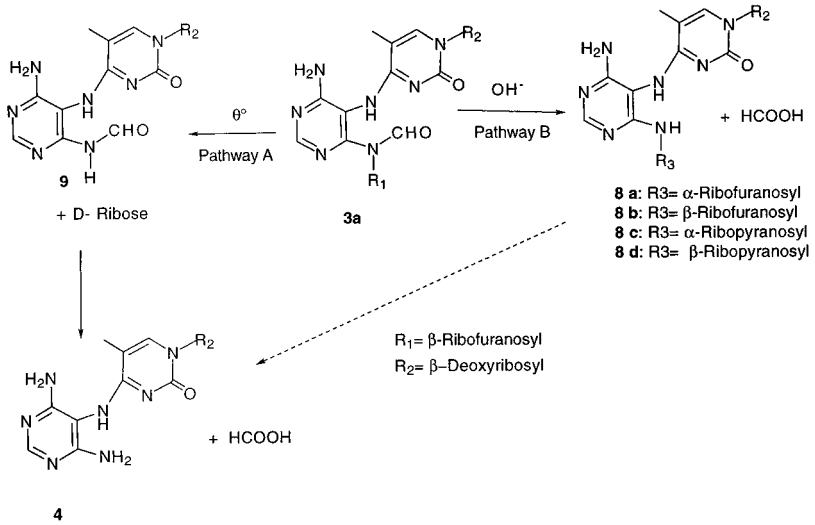
(19) Observed anomeric proton chemical shifts and coupling constants were identical with those obtained with an authentic sample of D-ribose.

(20) Hems, G. *Nature* **1960**, *186*, 710−712.
(21) Favre, A. In *Bioorganic Photochemistry: Photochemistry and the Nucleic Acids*; Morrison, H., Ed.; J. Wiley & Sons: New York, 1990; Vol. 1, pp 317−378. (b) Sontheimer, E. J. *Mol. Biol. Rep.* **1994**, *20*, 35−44. (c) Favre, A., Fourrey, J.-L. *Acc. Chem. Res.* **1995**, *28*, 375−382.
(22) (a) Sigman, D. S.; Chen, C.-H. *Annu. Rev. Biochem.* **1990**, *59*, 207−236. (b) Sigman, D. S.; Mazumder, A.; Perrin, D. M. *Chem. Rev.* **1993**, *93*, 2295−2316.

Scheme 2



Scheme 3



bond and multiple bond $^{13}C-^{1}H$ spectra. The signals of the anomeric H-1′ of each residue allowed attribution of the signals of carbons C-2 (δC 152.3 ppm) and C-6 (δC 136.3 ppm) of the pyrimidine nucleus and those of C-4 (δC 150.9 ppm) and C-8 (δC 143.8 ppm) of the adenine nucleus (^{3}J). Carbons C-4 of the pyrimidine base (δC 154.4 ppm) and C-5 of the purine base (δC 126.1 ppm) were assigned (^{3}J correlations) from protons H-6 and H-8, respectively. Confirmation of the attribution of C-4 of the pyrimidine base was obtained from long range (LR) (^{3}J) correlations with the two methyl group protons. The remaining aromatic proton H-2 allowed attribution of C-6 (δC 159.2 ppm) from its ^{3}J coupling. Bonding between the 6-amino group of the adenine with position C-4 of the pyrimidine was deduced from the carbon chemical shifts of **7**. Compared to **2a**, the chemical shifts of carbon C-6 ($\Delta\delta$ = 8.7 ppm), C-5 ($\Delta\delta$ = 7.6 ppm), C-2 ($\Delta\delta$ = 5.3 ppm), and C-4 ($\Delta\delta$ = 5.8 ppm) of the adenine part were deshielded (no significant variation was observed at the C-8 position) as expected for a substitution of the 6-amino group.[23] Finally, the high shielding of the signal due to C-4 of **6a** (192.4 ppm), now observed at 154.4 ppm (see above) in **7**, is in agreement with the proposed structure. Moreover, an authentic sample of 3,5-dimethyl-2′-deoxycytidine **6c**, prepared by treatment of **6a** with liquid ammonia, exhibited very close ^{13}C chemical shift data.

From a mechanistic point of view, we propose that photoproduct **7** results from an attack of the 6-amino group of **2a** at the electrophilic C-4 position of the excited thiocarbonyl group leading to an unstable intermediate that subsequently undergoes hydrogen sulfide elimination. Since the C-4 position of 4-mercaptopyrimidine bases is susceptible to undergo nucleophilic displacement by amines, we have confirmed that the formation of **7** was not the result of a thermal reaction (no reaction occurred in the absence of light). This postulated mode of formation of **7** might suggest another pathway, different from the previously proposed one,[12] to account for the formation of **3a**. As with **7**, **3a** could have resulted from a nucleophilic attack of the N-7 nitrogen of **2a** at C-4 of the excited thiocarbonyl group of **1a** leading to an unstable adduct that, after elimination of H_2S, would undergo imidazole ring opening as observed with N-7-alkylated purines.[9]

Conclusion

In summary, we have demonstrated that the formation of *N*-4-(4-amino-6-formamidopyrimidin-5-yl)-5-methylcytosine adducts by irradiation of N-1-substituted-4-thiothymine and N-9-substituted-adenine is a general reaction. In this process an imidazole ring fission reaction could be observed. It might be related to the action mechanism of biologically important DNA-modifying agents known to induce the formation of *N*-7-alkyl-Fapy sites.[9] It is probable that the preferred N-7;C-8 imidazole ring fission of adenosine leading to *N*-9-formyl derivatives be governed by the electron deficient nature of the N-7-substituent. Indeed, this pathway is reminiscent of the analogous ring opening of *N*-9-adenines derivatives upon

(23) Breitmaier, E.; Voelter, W. *Carbon-13 NMR Spectroscopy*; Ebel, H. F., Ed.; VCH: Weinheim, 1987.

N-7-acylation.[24] In connection with this regioselective ring fission between N-7 and C-8, it is important to note that, although the majority of N-7-alkylated purine derivatives led to *N*-5-formamidopyrimidines, N-7-adducts of aflatoxin B1 can lead also to *N*-6-formamidopyrimidines.[9c,d] Interestingly, structurally related DNA adducts, involving the N-7 position of purines, give rise to divergent biological responses.[25] Consequently, it can be suggested that derivatives of type **3** could be useful for studying the structure−activity relationship of such adducts together with the enzymatic excision mechanism of the Fpg protein[11] and the enzymatic reclosure of opened imidazole rings of purine.[26]

Experimental Section

Chemicals. 4-Thiothymidine (**1a**), 4-thiothymin-1-ylacetic acid (**1b**), and *N*-3-methyl-4-thiothymidine (**6a**) were prepared according to literature procedures.[27,28] Adenin-9-ylacetic acid (**2b**)[29] was prepared by saponification of its ethyl ester.[30] ^{1}H NMR chemical shifts (δ) are reported relative to residual solvent traces (CD_3OD: 3.3 ppm, DMSO-d_6: 2.6 ppm D_2O: 4.8 ppm). ^{13}C chemical shifts (δ) are reported relative to solvent (DMSO-d_6: 39.7 ppm, CD_3OD: 49.0 ppm) or for spectra recorded in D_2O to external dioxane (67.8 ppm). HMBC experiments were optimized to suppress ^1J C−H coupling of 200 Hz and recorded at 276 K for **3a**, 253 K for **5** and 300 K for **7**. FAB HRMS were performed by the Service Central d'Analyze du CNRS (Lyon, France).

Irradiation Conditions. Irradiation experiments were performed under continuous nitrogen bubbling in aqueous solution. To obtain **3a**, an equimolecular solution (140 mL) of adenosine and 4-thiothymidine (1.2 mmol) in a Pyrex tube (17 × 3.5 cm i.d.) was kept at 5 °C and irradiated for 234 h using an Original Hanau Quarzlampen Fluotest-Forte ref 5261. To obtain **3b**, a 50 mL aqueous solution of adenosine (0.26 mmol) and 4-thiothymin-1-ylacetic acid (0.13 mmol) was placed in a cylindrical Pyrex flask (2.7 × 6 cm i.d.) and irradiated for 20 to 26 h at a 4.5 cm distance from a superpressure 350 W lamp equipped with a filter to cutoff wavelengths below 310 nm. The system was thermostated at 4 °C. To obtain **3c**: A 50 mL aqueous solution of adenin-9-ylacetic acid (**2b**) (0.25 mmol) and 4-thiothymidine (0.12 mmol) (addition of an aliquot of Na_2CO_3 allowed dissolution of **2b**) was irradiated for 14 to 20 h using the conditions reported for **3b**. To obtain **3d** a 50 mL aqueous solution of adenin-9-ylacetic acid (**2b**) (0.25 mmol) and 4-thiothymin-1-ylacetic acid (**1b**) (0.12 mmol) (addition of an aliquot of Na_2CO_3 allowed dissolution of **2b**) was irradiated for 12−18 h using the conditions reported for **3b**. To obtain **7**, a 50 mL aqueous solution of adenosine (0.49 mmol) and *N*-3-methyl-4-thiothymidine (**6**) (0.24 mmol) was irradiated for 106 h using the conditions reported for **3b**.

Progress of these reactions was monitored by UV (disappearance of the thiocarbonyl absorption at ca. 335 nm, appearance of a new maximun at ca. 295 nm for **3a**−**d**) or ^{1}H NMR spectroscopy. After completion of the reaction, the irradiated solutions were lyophilized or concentrated under high vacuum below 20 °C and stored at −18 °C prior purification.

General Purification Procedure. Photoproducts **3a** and **7** were purified on a medium pressure (400 mbar) RP-18 column (40−63 μm; 14 × 3 cm i.d. for **3a** and 12 × 3 cm i.d. for **7**). Fractions (4 mL) were collected. The eluant system for **3a** was water (80 mL) then water−acetonitrile 1% (400 mL); 2% (500 mL); 5% (100 mL); 10% (200 mL). Fractions containing **1a** (229/246), **2a** (96/228), and **3a** (37/88) were concentrated then lyophilized leading to **3a** (60% yield), **2a** (33%), and **1a** (24%). The eluant system for **7** was water (100 mL) and then water−acetonitrile 1% (50 mL); 4% (100 mL); 6% (100 mL); 8% (100 mL), 12% (100 mL), 14% (100 mL). Fractions (4 mL) containing **7** (134/141) were concentrated and then lyophilized (22% yield).

The other photoproducts **3b**−**d** were purified by HPLC using a C18 (6 μm, 60 Å) cartridge (25 × 100 mm) with a flow rate of 8 mL/min. A photodiode array detector was employed. Fractions containing the products were lyophilized or concentrated under high vacuum below 20 °C and then desalted on a medium pressure (400 mbar) RP-18 column (40−63 μm; 9 × 2.5 cm i.d. for **3b,c** and 9 × 1.5 cm i.d. for **3d**). Fractions (2 mL) were collected, and elution was performed with H_2O from fractions 1 to 17 and then 1% CH_3CN for **3b,c** and H_2O for **3d**. Fractions 18/39 contained **3b**, fractions 16/42 **3c** and 1/4 **3d**.

Photoproduct **3b**. A 40 min linear gradient of 0−12% acetonitrile in 0.05 M triethylammonium acetate (pH 6.0) was used. Fractions containing **3b** were concentrated and desalted (27% yield).

Photoproducts **3c** and **3d**. An isochratic 0.01 M triethylammonium acetate (pH 6.0) solution was used for 30 min, and then a 10 min (**3c**) or 30 min (**3d**) linear gradient of 0−12% acetonitrile in 0.01 M triethylammonium acetate (pH 6.0) was used. The fractions containing the photoproducts were concentrated and desalted leading to **3c** (24% yield) and **3d** (20% yield).

Photoproduct **3a**: UV (H_2O) see ref 12; ^{1}H NMR (300 MHz; D_2O) δ 8.40 (brs, 1H, H-8), 8.38 (s, 1H, H-2), 7.81 (s, 1H, H-6), 6.26 (t, 1H, J = 6.5 Hz, H-3′ T), 5.50 (brs, 1H, H-2′ A), 4.45 (m, 2H, H-2′ A, H-3′ T), 4.06 (m, 2H, H-3′ A, H-4′ T), 4.00−3.40 (m, 5H, H-5′/H-5″ A, H-5′/H-5″ T, H-4′ A), 2.37 (m, 2H, H-2′/H-2″ A), 2.09 (s, 3H, CH_3); ^{13}C NMR (62.90 MHz; D_2O) see Table 1; HRMS (FAB) calcd for $C_{20}H_{27}N_7O_9Na$ $(M + Na)^+$ 532.1770, found 532.1790. Anal. Calcd for $C_{20}H_{27}N_7O_9 \cdot 2H_2O$: C, 44.03; H, 5.73; N, 17.97. Found: C, 43.56; H, 5.55; N, 17.57.

Photoproduct **3b**: UV (H_2O) λ_{max} 294 nm; ^{1}H NMR (300 MHz; D_2O) δ 8.49 (br s, 1H, H-8), 8.34 (s, 1H, H-2), 7.47 (s, 1H, H-6), 5.52 (d, 1H, J = 4.2 Hz, H-1′), 4.40 (br m, 1H, H-2′), 4.30 (m, 2H, N-1 CH_2), 4.06 and 3.90 (br m, 2H, H-3′ and H-4′), 3.42−3.80 (m, 2H, H-5′/H-5″), 2.04 (s, 3H, CH_3); FAB-MS (positive mode) m/z 452 $[(M + H)^+]$, 474 $[(M + Na)^+]$; ^{13}C NMR (62.90 MHz; D_2O) see Table 1.

Photoproduct **3c**: UV (H_2O) λ_{max} 295 nm; ^{1}H NMR (250 MHz; D_2O) δ 8.37 (br s, 1H, H-8*), 8.33 (s, 1H, H-2*), 7.83 (s, 1H, H-6), 6.28 (t, 1H, J = 6.3 Hz, H-1′), 4.47 (m, 1H, H-3′), 4.25 (s, 2H, N-9 CH_2), 4.07 (m, 1H, H-4′) 3.94−3.73 (m, 2H, H-5′/H-5″), 2.57−2.25 (m, 2H, H-2′/H-2″), 2.10 (s, 3H, CH_3) *: interchangeable attributions; FAB-MS (positive mode) m/z 458 $[(M + Na)^+]$, 480 $[(M - H + 2Na)^+]$; ^{13}C NMR (62.90 MHz; D_2O) see Table 1.

Photoproduct **3d**: UV (H_2O) λ_{max} 298 nm; ^{1}H NMR (250 MHz; D_2O) δ 8.37 (s, 1H, H-8*), 8.31 (s, 1H, H-2*), 4.36 and 4.25 (2s, 4H, N-1-CH_2 and N-9-CH_2), 2.07 (s, 3H, CH_3) *: interchangeable attributions; FAB-MS (positive mode) m/z 400 $[(M + Na)^+]$; ^{13}C NMR (62.90 MHz; D_2O) see Table 1.

Thermic Degradation of Photoproduct 3a. An aqueous solution of **3a** (88 mg, 0.173 mmol) in 1.15 mL of H_2O was heated to 90 °C for 4 h 40 min. The solution was concentrated to dryness and then purified by RP HPLC on a column identical with that used to purify photoproducts **3b**−**d**. An isochratic 0.01 M triethylammonium acetate (pH 6.0) solution was used for 20 min, and then a 5 min linear gradient of 0−12% acetonitrile in 0.01 M triethylammonium acetate (pH 6.0) was used. Fractions containing **4** were concentrated and desalted leading to 35 mg of product (58% yield). UV (H_2O) λ_{max} 284 nm, ϵ 8942 mol^{-1} cm^{-1} dm^3; ^{1}H NMR (400 MHz;

(24) (a) Altman, J.; Ben-Ishai, D. *J. Heterocycl. Chem*. **1968**, *5*, 679−682. (b) Leonard, N. J.; McDonald, J. J.; Henderson, R. E. L.; Reichmann, M. E. *Biochemistry* **1971**, *10*, 3335−3342.

(25) Baertschi, S. W.; Raney, K. D.; Shimada, T.; Harris, T. M.; Guengerich, F. P. *Chem. Res. Toxicol.* **1989**, *2*, 114−122.

(26) Chetsanga, C.; Grigorian, C. *Proc. Natl. Acad. Sci., U.S.A.* **1985**, *82*, 633−637.

(27) Fox, J. J.; Van Praag, D.; Wempen, I.; Doerr, I. L.; Cheong, L.; Knoll, J. E.; Eidinoff, M. L.; Bendich, A.; Brown, G. B. *J. Am. Chem. Soc.* **1959**, *81*, 178−183.

(28) Saintomé, C.; Clivio, P.; Fourrey, J.-L.; Woisard, A.; Favre, A. *Tetrahedron Lett.* **1994**, *35*, 873−876.

(29) Jones, A. S.; Lewis, P.; Withers, S. F. *Tetrahedron* **1973**, *29*, 2293−2296.

(30) Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Peterson, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem*. **1994**, *59*, 5767−5773.

DMSO-d_6) δ 7.92 (s, 1H, N-7-H), 7.85 (s, 1H, H-2), 7.76 (s, 1H, H-6), 6.26 (t, 1H, J = 6.7 Hz, H-1′), 6.06 (br s, 2H, NH_2), 4.32 (m, 1H, H-3′), 3.84 (m, 1H, H-4′), 3.66 (m, 2H, H-5′/H5″), 2.15 (m, 1H, H-2′), 2.09 (s, 3H, CH_3), 2.03 (m, 1H, H-2″); FAB MS (positive mode) m/z 350 [$(M + H)^+$]; ^{13}C NMR (62.90 MHz; DMSO-d_6) see Table 1.

Acetylation of 3a. Photoproduct **3a** (260 mg, 0.51 mmol) was dissolved in anhydrous pyridine (9 mL), and acetic anhydride (2.5 mL) was added. The solution was stirred overnight at rt with exclusion of moisture. The solution was chilled to 0 °C, and methanol (2.5 mL) was added. The solution was then concentrated to dryness below 20 °C under high vacuum and purified by silica gel chromatography using a gradient of methanol in CH_2Cl_2 affording 127 mg of the penta-*O*-acetyl product **5** (34.6% yield). UV (CH_3OH) λ_{max} (nm) 291; ^{1}H NMR (300 MHz; CD_3OD) δ 8.42 (br s, 1H, H-8), 8.25 (s, 1H, H-2), 7.65 (s, 1H, H-6), 6.17 (t, 1H, J = 6.8 Hz, H-1′ T), 5.67−5.36 (m, 3H, H-1′ A, H-2′ A, H-3′ A), 5.23 (m, 1H, H-3′ T), 4.47−3.98 (m, 6H, H-5′/H-5″ A, H-5′/H-5″ T, H-4′ A, H-4′ T), 2.52 (m, 1H, H-2′ T), 2.24 (m, 1H, H-2″ T), 2.10−2.08−2.02−2.00 (4s, 18H, 5 OAc, CH_3); FAB MS (positive mode) m/z 742 [$(M + Na)^+$]; ^{13}C NMR (62.90 MHz; CD_3OD) see Table 1.

Photoproduct 7: UV (H_2O, pH 7) λ_{max} 304 nm, ϵ 14817 mol^{-1} cm^{-1} dm^3; ^{1}H NMR (250 MHz; D_2O) δ 8.53 (s, 1H, H-2), 8.47 (s, 1H, H-8), 7.41 (s, 1H, H-6), 6.32 (t, 1H, J = 6.4 Hz, H-1′ T), 6.12 (d, 1H, J = 6.0 Hz, H-1′ A), 4.83 (m, 1H, H-2′ A), 4.45 (m, 2H, H-3′ T, H-3′ A), 4.29 (m, 1H, H-4′ A), 4.02 (m, 1H, H-4′ T), 3.96−3.69 (m, 4H, H-5′/H-5″ A, H-5′/H-5″ T), 3.49 (s, 3H, NCH_3), 2.39 (m, 2H, H-2′/H-2″ T), 1.34 (s, 3H, CH_3); ^{13}C NMR (62.90 MHz; D_2O) see Table 1; HRMS (FAB) calcd for $C_{21}H_{28}N_7O_8$ $(M + H)^+$ 506.1999, found 506.1975.

N-3,5-Dimethyl-2′-deoxycytidine (6c). *N*-3-Methyl-4-thiothymidine (**6a**) (50 mg, 0.18 mmol) was dissolved in liquid ammonia (1 mL), and the solution was heated to 60 °C for 39 h in a steel container. After evaporation of ammonia, the residue was dissolved in methanol and concentrated to dryness leading to 24 mg of **6c** (51% yield). An analytical sample was obtained after purification by chromatography on a silica gel column eluted with a gradient of methanol in dichloromethane (5 to 40%). UV (H_2O, pH 7) λ_{max} (nm) 268, ϵ 1430 mol^{-1} cm^{-1} dm^3; ^{1}H NMR (300 MHz; D_2O) δ 7.69 (s, 1H, H-6), 6.30 (t, 1H, J = 6.4 Hz, H-1′), 4.48 (m, 1H, H-3′), 4.08 (m, 1H, H-4′), 3.85 (m, 2H, H-5′/H-5″), 3.48 (s, 3H, NCH_3), 2.42 (m, 2H, H-2′/H-2″), 2.06 (s, 3H, CH_3); ^{13}C NMR (75.47 MHz; D_2O) δ 161.1 (C4), 150.9 (C2), 136.3 (C6), 106.7 (C5), 88.0 (C1′*), 87.8 (C4′*), 71.4 (C3′), 62.2 (C5′), 40.1 (C2′), 31.4 (NCH_3), 14.1 (CH_3) *: interchangeable attributions; HRMS (CI, CH_4) calcd for $C_{11}H_{18}N_3O_4$ $(M + H)^+$ 256.1262, found 256.1258.

Adenin-9-ylacetic Acid (2b). Ethyl adenin-9-yl acetate[30] (994 mg, 4.5 mmmol) was dissolved in MeOH (25 mL) and cooled to 0 °C. A 2 N aqueous NaOH solution (25 mL) was added. After stirring 30 min at 0 °C, the aqueous solution was brought to pH 1 by addition of a 4 N aqueous HCl solution. After filtration, washing with water, and drying, 822 mg of **2b** was obtained (95% yield). For ^{1}H NMR data see ref 29. ^{13}C NMR (62.90, D_2O) δ 174.9 (COOH), 155.9 (C6), 153.0 (C2), 149.4 (C4), 143.5 (C8), 118.5 (C5), 47.6 (CH_2); CI-MS m/z 194 [$(M + H)^+$].

Acknowledgment. Thanks are due to M.-T. Martin for recording the HMBC spectra and M.-T. Adeline and J. Hamon for their assistance in HPLC purifications. We also acknowledge C. Girard, P. Varenne and O. Laprévote for recording or interpreting some mass spectra. We thank D. Guillaume for recording variable temperature ^{1}H NMR experiments and for unvaluable discussions.

Supporting Information Available: Copies of ^{1}H NMR and ^{13}C NMR spectra of compounds **3a**−**d**, **4**, **5**, **6c**, and **7** (17 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering information.

JO971162P