Adenine photodimerization in deoxyadenylate sequences: elucidation of the mechanism through structural studies of a major d(ApA) photoproduct

Shiv Kumar, Prakash C.Joshi+ Narain D.Sharma, Samarendra N.Bose, R.Jeremy H.Davies*, Naohito Takeda[1] and James A.McCloskey[1]
Biochemistry Division, School of Biology and Biochemistry, Queen's University, Belfast BT9 7BL, UK and [1]Departments of Medicinal Chemistry and Biochemistry, University of Utah, Salt Lake City, UT 84112, USA

Received April 3, 1991; Revised and Accepted May 8, 1991

ABSTRACT

The mechanism of the photodimerization of adjacent adenine bases on the same strand of DNA has been elucidated by determining the structure of one of the two major photoproducts that are formed by UV irradiation of the deoxydinucleoside monophosphate d(ApA). The photoproduct, denoted d(ApA)*, corresponds to a species of adenine photodimer first described by Pörschke (Pörschke, D. (1973) *J.Am.Chem.Soc.* 95, 8440–8446). From a detailed examination of its chemical and spectroscopic properties, including comparisons with the model compound *N*-cyano-*N*'-(1-methylimidazol-5-yl)formamidine, it is deduced that d(ApA)* contains a deoxyadenosine unit covalently linked through its C(8) position to C(4) of an imidazole *N*(1) deoxyribonucleoside moiety bearing an *N*-cyanoformamidino substituent at C(5). On treatment with acid, d(ApA)* is degraded with high specificity to 8-(5-aminoimidazol-4-yl)adenine whose identity has been confirmed by independent chemical synthesis. It is concluded that the primary event in adenine photodimerization entails photoaddition of the N(7)-C(8) double bond of the 5'-adenine across the C(6) and C(5) positions of the 3'-adenine. The azetidine species thus generated acts as a common precursor to both types of d(ApA) photoproduct which are formed from it by competing modes of azetidine ring fission.

INTRODUCTION

Most of the mutational changes induced by short wavelength ultraviolet (UV) radiation are initiated by photochemical damage to the nucleobases of DNA (*1*). Several species of photoproducts derived from the pyrimidine bases in DNA have been successfully isolated and characterized (*2*), notably the cyclobutane photodimers and bipyrimidine (6-4) photoadducts which are formed with quantum yields of the order 10^{-2} to 10^{-3} mole einstein^{-1}. However, progress in identifying photoproducts derived from the purine bases in DNA has been much more limited though their existence has been clearly demonstrated in a variety of gel sequencing experiments (*3,4*). At certain sites in a fragment of the human alphoid DNA sequence, the yield of purine photoproducts is reported to be comparable to that of pyrimidine dimers (*5*).

The first example of an intramolecular photoreaction involving purine bases in a DNA sequence was reported by Pörschke (*6,7*) in 1973 and concerned the photodimerization of adjacent adenine bases in oligomers and polymers of deoxyadenylic acid; this occurs with the relatively high quantum yield of 2.5×10^{-3} mole einstein^{-1} in poly(dA). Recently strong evidence for adenine photodimerization in single stranded regions of DNA has been obtained by photofootprinting (*4,8*). Pörschke (*7*), and subsequently Rahn (*9*), isolated a discrete photoproduct, denoted $d(pA)_2$*, from the UV-irradiated dinucleotide $d(pA)_2$. However, despite its potential importance in photobiology, the structure of this photoproduct has remained unresolved. Mainly on the basis of fluorescence measurements, Gasparro and Fresco (*10*) postulated that $d(pA)_2$* was generated by covalent linking of adjacent adenine residues through their respective C(8) positions but this assignment was not underpinned by definitive analytical or spectroscopic data. The experimental results described in this paper enable us to advance a quite different structure for the Pörschke photoproduct and to deduce its mode of formation.

We have investigated adenine photodimerization in the deoxydinucleoside monophosphate d(ApA) and earlier showed (*11*) that, on irradiation at 254 nm in aqueous solution, two distinct photoproducts are formed which can be separated by reversed-phase high performance liquid chromatography. One of these corresponds in its spectroscopic and electrophoretic properties to the species described by Pörschke and is designated d(ApA)*. The other photoproduct, A=A, which affords 4,6-diamino-5-guanidinopyrimidine on acid hydrolysis, is produced in

* To whom correspondence should be addressed
+ On leave from I.T.R.C., Lucknow, India

comparable yield. The spectroscopic and chemical properties of A=A are consistent (*11*) with it being formed by hydrolytic fission of an azetidine precursor resulting from photoaddition of the N(7)-C(8) double bond of the 5'-adenine across the C(6) and C(5) positions of the 3'-adenine. Here, we present new analytical and spectroscopic data for the Pörschke-type photoproduct, d(ApA)*, and show that it is degraded by acid to 8-(5-amino-imidazol-4-yl)adenine. We conclude that in d(ApA)* the 5'-adenine is linked through its C(8) position to C(4) of an imidazole nucleus bearing an *N*-cyanoformamidino substituent at C(5). It is inferred that both of the d(ApA) photoproducts are derived from the same azetidine precursor through competing modes of opening of the highly strained four-membered ring.

EXPERIMENTAL

Materials

The ammonium salt of d(ApA), 4,5,6-triaminopyrimidine sulphate, 4-hydroxymethylimidazole hydrochloride and trifluoro-acetic acid (TFA) were obtained from Sigma Chemical Company Ltd. 4-Nitro-1-(tetrahydropyran-2-yl)imidazole-5-carboxaldehyde was synthesised from 4-hydroxymethylimidazole according to Wu *et al* (*12*). *N*-Cyano-*N*^{1}-(1-methylimidazol-5-yl)formamidine was a gift from Dr C.A.Ramsden. *C*-Perdeuterioglycerol, 97+ atom%D, and *O*-perdeuterioglycerol, 98.8 atom%D, were purchased from MSD Isotopes, St.Louis, MO. *O,S*-Perdeuterio-thioglycerol was prepared by vigorously shaking a 1:2 mixture of 1-thioglycerol and D_2O, evaporating the D_2O, and repeating this procedure three additional times.

General methods

UV and IR spectra were recorded with Perkin-Elmer Lambda 2 and 983G instruments respectively. ^{1}H NMR spectra were acquired at 300 MHz with a General Electric QE-300 spectrometer. The residual proton signal in $(CD_3)_2SO$ at δ 2.49 served as a reference for chemical shifts; for the spectrum run in D_2O the DHO signal was assigned δ 4.80. Reversed phase high performance liquid chromatography (RP-HPLC) was carried out on columns of μBondapak C_{18} (from Millipore/Waters) with detection at 254 nm as described previously (*11*). The flow rate was 0.8 ml min^{-1} for analytical (300×3.9 mm) columns and 2 ml min^{-1} for semi-preparative (300×7.8 mm) columns. Separations were performed using isocratic elution with solvent A for 5 min then a linear gradient to a 1:1 mixture of solvents A and B after 15 min followed by isocratic elution with this solvent mix. For water-methanol gradients, solvent A was water and solvent B a 60:40 methanol:water mixture; for TFA-methanol gradients, solvent A was 0.05% aqueous TFA and solvent B a 60:40 methanol:0.05% TFA mixture.

Mass spectrometry

Fast atom bombardment (FAB) mass spectrometry of the photoproduct d(ApA)* was carried out with a VG 70-SEQ mass spectrometer and VG-250 data system. Xenon atoms were used to bombard the sample, the ion gun conditions being typically 7 kV accelerating potential and 1 mA discharge current; the source accelerating potential was 8 kV. Other FAB mass spectrometric analyses were performed on a MAT 731 mass spectrometer fitted with an Ion Tech FAB 11N ion source, with either magnetic or electric scanning at 8 kV accelerating potential. A neutral Xe beam having 6 keV energy was used for ion desorption. The same instrument was used to obtain electron ionization (EI) mass spectra taken at 70 eV ionizing energy.

Exact mass measurements were made by the peak matching technique at a resolution of 7000−10000. The values for compounds **8**, **9**, **10** and 5-amino-1-methylimidazole were determined by the SERC Mass Spectrometry Service at Swansea.

The active hydrogen content of molecules was determined by FAB mass spectrometry utilising selectively deuterated glycerol or thioglycerol as the matrix to effect hydrogen-deuterium exchange. Full details of the method have been given elsewhere (*13,14*).

Preparation of the photoproduct, d(ApA)*

Neutral aqueous solutions of d(ApA) were irradiated, at 254 nm, with a PCQ-X1 Photochemical Lamp (Ultra-Violet Products, San Gabriel, California) as described previously (*11*). Irradiation was continued until the maximum absorbance of the solution at 258 nm had decreased by 10% and the minimum absorbance at 227 nm had approximately doubled. After lyophilization, the residue was subjected to high voltage paper electrophoresis at pH 8.7. Under these conditions, d(ApA)* migrates as a dianion and is well separated from d(ApA) and the other photoproduct, A=A, which is uncharged (*11*). The crude photoproduct extracted with water from the d(ApA)* band on the paper was further purified by semi-preparative RP-HPLC using a water-methanol gradient. Fractions comprising the main peak (retention time 14.5 min) were lyophilized to give d(ApA)* as a white amorphous solid. The UV spectrum of the d(ApA)* thus prepared (λ_{max} (pH 7) 264 nm) was identical to that published by Pörschke (*7*) for d(pA)$_2$*. Typically, 4 A_{264} units of d(ApA)* (~ 100 μg) were obtained per mg of UV-irradiated d(ApA). IR (KBr disc): ν_{max} 2201 cm^{-1} (s). ^{1}H NMR ($(CD_3)_2SO$): δ 11.0 (bd s, 1H, exch. D_2O, NH), 8.58, 8.20, 8.17, 8.11 and 8.07 (5 singlets (from two geometrical isomers), 3H, 3×base C-H), 7.40 (bd s, 2H, exch. D_2O, NH_2), 6.42 (m, 1H, deoxyribose C(1')H), 5.99 (m, 1H, deoxyribose C(1')H), 5.51 (bd s, 1H, exch. D_2O, OH), 4.6−2.0 (series of complex multiplets assigned to remaining deoxyribose protons). The mass spectrum of d(ApA)* is discussed under *Results and Discussion*.

Acid hydrolysis of d(ApA)*

Milligram amounts of d(ApA)* were dissolved in 1 M HCl (500 μl) and heated in a sealed tube at 100°C for 1 h. After removing excess HCl *in vacuo*, the solid residue was redissolved in 0.05% TFA and fractionated by analytical RP-HPLC with a TFA-methanol gradient. A small amount of adenine (retention time 8.5 min) was detected but the major UV-absorbing product (retention time 18.7 min) had λ_{max} 320 and 239 nm; approximately 10 A_{320} units were recovered per mg of hydrolysed photoproduct. ^{1}H NMR ($(CD_3)_2SO/D_2O$): δ 8.40 (s, 1H), 7.43 (s, 1H). Mass spectrum (FAB, positive ion, glycerol matrix): *m/z* 217.0948 (MH^+), calc. for $C_8H_9N_8$ *m/z* 217.0950.

When d(ApA)* was similarly hydrolysed with 90% formic acid at 100°C, for 2 h, the HPLC profile revealed the formation of small amounts of adenine and the HCl hydrolysis product (both identified by UV spectra, retention times and co-chromatography) as well as a major hydrolysis product with λ_{max} 319 nm having a retention time of 22.1 min.

Synthetic compounds

8-(5-Nitroimidazol-4-yl)adenine (**3**). A mixture of 4-nitro-1-(tetrahydropyran-2-yl)imidazole-5-carboxaldehyde (**2**) (300 mg,

1.3 mmol), 4,5,6-triaminopyrimidine (**1**) (125 mg, 1 mmol) and *p*-benzoquinone (108 mg, 1 mmol) was stirred and refluxed in 2-propanol (20 ml) for 15 h. The solution was filtered hot and the residual solid was washed with 2-propanol (2×5 ml). The combined filtrate was evaporated to dryness and the dark brown residue then extracted with ether to remove *p*-benzoquinone. The crude product thus obtained was further purified by preparative HPLC (Waters Prep LC 3000 system with Delta Pak C18−300 Å column (300×19 mm), flow rate 25 ml min^{-1}, detection at 254 nm). On elution with a TFA-methanol gradient compound **3** was recovered in the main peak (retention time 14.5 min) as a pale yellow solid (75 mg, 30%) with m.p. > 300°C. UV (MeOH): λ_{max} 359 nm (ϵ 5300 M^{-1} cm^{-1}) and 266 (7800). ^1H NMR ($(CD_3)_2SO/D_2O$): δ 8.42 (s, 1H), 8.04 (s, 1H). Mass spectrum (EI): *m/z* 246.0612 (M^+), calc. for $C_8H_6N_8O_2$ *m/z* 246.0614.

8-(5-Aminoimidazol-4-yl)adenine (**4**). A solution containing 8-(5-nitroimidazol-4-yl)adenine (**3**) (25 mg, 0.1 mmol) dissolved in a mixture of methanol (5 ml) and 1 M HCl (2 ml) was hydrogenated (at 3 atmospheres pressure) in the presence of 10% palladium/charcoal (6 mg) for 4 h. After filtration to remove the catalyst, the filtrate was evaporated to dryness. The residual crude product was purified to homogeneity by semi-preparative RP-HPLC utilising a TFA-methanol gradient (retention time 16.0 min). It was thus obtained as an amorphous white solid (10 mg, 45%) which darkened and decomposed above 160°C. UV (pH 1): λ_{max} 320 nm (ϵ 9400 M^{-1} cm^{-1}) and 239 (4900); (pH 7) λ_{max} 318 nm (ϵ 9600 M^{-1} cm^{-1}) and 241 (5400). ^1H NMR ($(CD_3)_2SO/D_2O$): δ 8.45 (s, 1H), 7.47 (s, 1H). Mass spectrum (FAB, positive ion, glycerol): *m/z* 217.0956 (MH^+), calc. for $C_8H_9N_8$ *m/z* 217.0950.

8-(8-Adenosyl)adenosine (**9**) *and 8-(8-adeninyl)adenine* (**10**). Acetone photosensitized irradiation of an equimolar mixture of adenosine and 8-bromoadenosine was used to prepare 8-(8-adenosyl)adenosine as described previously (*15*). Its identity was confirmed by its UV and ^1H NMR spectra (*15*) and by mass spectrometry (FAB, positive ion, thioglycerol): *m/z* 533.1833 (MH^+), calc. for $C_{20}H_{25}N_{10}O_8$ *m/z* 533.1857.

8-(8-Adeninyl)adenine was prepared by heating 8-(8-adenosyl)adenosine (1 mg) with 1 M HCl (1 ml) at 100°C for 1 h in a sealed tube. After evaporation to dryness the residual hydrolysate was dissolved in 0.05% aqueous TFA and subjected to semi-preparative RP-HPLC with a TFA-methanol gradient. The main UV-absorbing component (retention time 23.6 min) comprised 8-(8-adeninyl)adenine. UV (pH 1): λ_{max} 314 (sh), 325 and 341 nm; (pH 13) λ_{max} 322 (sh), 332 and 349 nm. Mass spectrum (FAB, positive ion, thioglycerol): *m/z* 269.1021 (MH^+), calc. for $C_{10}H_9N_{10}$ *m/z* 269.1012.

Hydrolysis of 8-(8-adenosyl)adenosine with 90% formic acid at 100°C for 2 h gave the same product in similar yield.

Properties of N-cyano-N^1-(1-methylimidazol-5-yl)formamidine (**8**). A crystalline sample of the model compound **8**, had the following spectroscopic and chemical properties. IR (KBr disc): υ_{max} 2194 cm^{-1} (s). UV (pH 1): λ_{max} 246 nm (ϵ 10900 M^{-1} cm^{-1}); (pH 7) λ_{max} 279 nm (ϵ 11800 M^{-1} cm^{-1}) and 230 (sh) (7900); (pH 13) λ_{max} 287 nm (ϵ 13600 M^{-1} cm^{-1}). Spectroscopic pKs (at 20°C): 4.8 and 8.0. ^1H NMR (D_2O, 20°C): signals observed for two components in approx. 2:1 ratio; major component δ 8.35 (s, 1H, exocyclic CH), 7.86 (s, 1H, C(2)H), 7.16 (s, 1H, C(4)H) and 3.62 (s, 3H, NCH_3), minor component singlets at δ 8.14, 8.08, 6.92 and 3.60. Mass spectrum (EI): *m/z* 149.0701 (M^+), calc. for $C_6H_7N_5$ *m/z* 149.0701.

The action of 1 M HCl (100°C, 1 h) on compound **8** was investigated by analytical RP-HPLC with a TFA-methanol gradient. The starting material (retention time 6.0 min) was converted into a single product (detected at 214 nm) with a retention time of 4.2 min. The EI mass spectrum of the isolated hydrolysis product showed *m/z* 97.0673 (M^+), calc. for $C_4H_7N_3$ *m/z* 97.0640.

RESULTS AND DISCUSSION

Isolation and spectroscopic properties of d(ApA)*

On UV irradiation at 254 nm in neutral aqueous solution d(ApA) is converted into the two main photoproducts, A=A and d(ApA)*, with a combined quantum yield of ~ 1×10^{-3} mole $einstein^{-1}$ (*11*). These are conveniently separated from one another and from unreacted d(ApA) by high voltage paper electrophoresis at pH 8.7. The d(ApA)* extracted from the paper in water was further purified by RP-HPLC. Material isolated in this way was chromatographically homogeneous and had the same UV spectrum (Fig. 1) as reported for $d(pA)_2$* by Pörschke (*7*). It could be stored as a lyophilized solid in a sealed tube at −20°C for several months with only slight decomposition. However, as judged by RP-HPLC, d(ApA)* is somewhat unstable in solution and breakdown products could be detected after periods of several hours at room temperature.

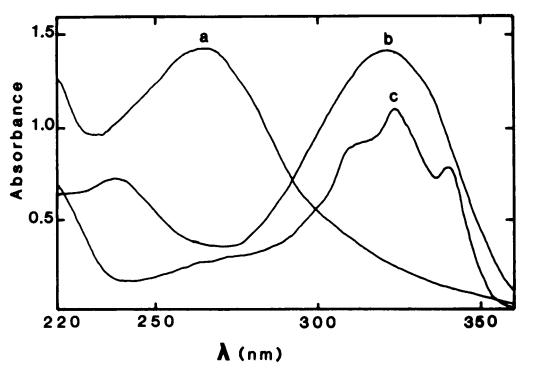


Figure 1. UV spectral profiles of: curve a, the photoproduct d(ApA)* at pH 7 (0.07 mM); curve b, 8-(5-aminoimidazol-4-yl)adenine (**4**) at pH 1 (0.15 mM); curve c, 8-(8-adeninyl)adenine (**10**) at pH 1.

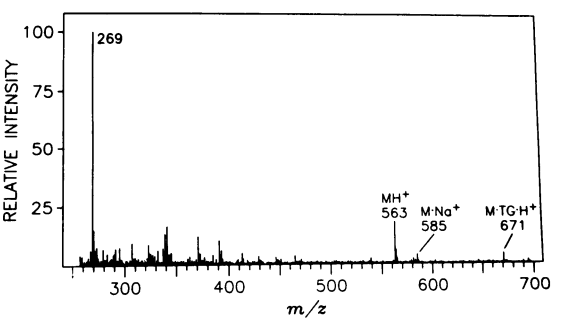


Figure 2. Upper mass range of the FAB mass spectrum of the photoproduct d(ApA)*. 1-Thioglycerol (TG) matrix; 7 keV Xe° beam.

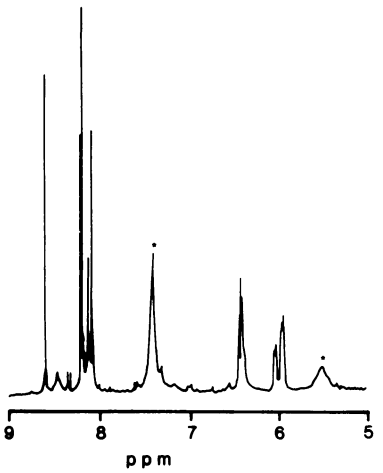
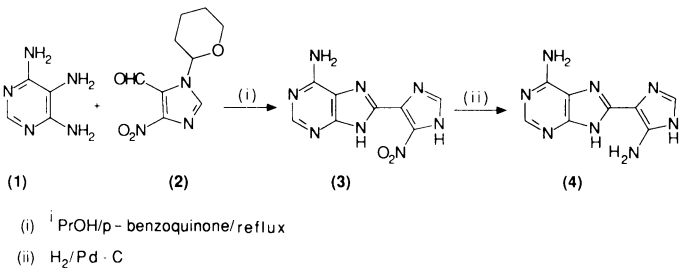


Figure 3. Section of the low field ^1H NMR spectrum of the photoproduct d(ApA)* in $(CD_3)_2SO$. Signals for D_2O-exchangeable protons are marked with an asterisk.

The molecular mass of d(ApA)* was determined (Fig. 2) as 562 by FAB mass spectrometry. High resolution measurements on the MH^+ ion (*m/z* 563.1526) support the composition $C_{20}H_{24}N_{10}O_8P$ (calculated *m/z* 563.1516) implying that d(ApA)* is derived from d(ApA) with a net loss of two hydrogen atoms. A prominent fragment ion (Fig. 2) of measured mass *m/z* 269.1014 is consistent with the composition $C_{10}H_9N_{10}$ (calculated *m/z* 269.1012) and thus represents the protonated base of d(ApA)* resulting from cleavage of both glycosidic bonds, a characteristic fragmentation mechanism in the mass spectra of base-base linked dinucleosides (*16*). This establishes that the formation of d(ApA)* from d(ApA) involves covalent coupling of the two adenine ($C_5H_5N_5$) bases to give some form of dehydrodimer. When a deuterated matrix was used to induce total deuterium exchange during FAB mass spectrometry, the MH^+ ion (*m/z* 563) shifted to *m/z* 570 and the protonated base ion (*m/z* 269) to *m/z* 275. Comparison of computer-generated isotopic patterns with the experimentally measured isotopic cluster in the mass spectrum (*14*) rigorously identified the *m/z* 570 species as d_6MD^+, hence showing that d(ApA)* contains six exchangeable hydrogen atoms. Similar analysis of the *m/z* 275 ion region established that the free heterocyclic base moiety contains five active hydrogens, of which three will remain when the base is incorporated into the intact photoproduct through two *N*-glycosidic linkages. These results are incompatible with the (8-8) coupled dehydrodimer structure previously advanced (*10*) for d(ApA)* which would contain a total of seven exchangeable hydrogens with four of them being attached to the dimeric base moiety.

The ^1H NMR spectrum (Fig. 3) of d(ApA)* showed two sets of signals for the anomeric deoxyribose protons thus confirming the integrity of the glycosidic bonds in the photoproduct. Three exchangeable hydrogens on the base moiety could be accounted for by a low field resonance at δ 11.0 for one NH proton and another singlet at δ 7.40 typical of an adenine amino group. The predicted three non-exchangeable hydrogens on the heterocyclic base moiety unexpectedly gave rise to a cluster of five singlet peaks in the aromatic region of the spectrum; this feature is discussed below, together with the significance of an infrared absorption band at 2201 cm^{-1}.



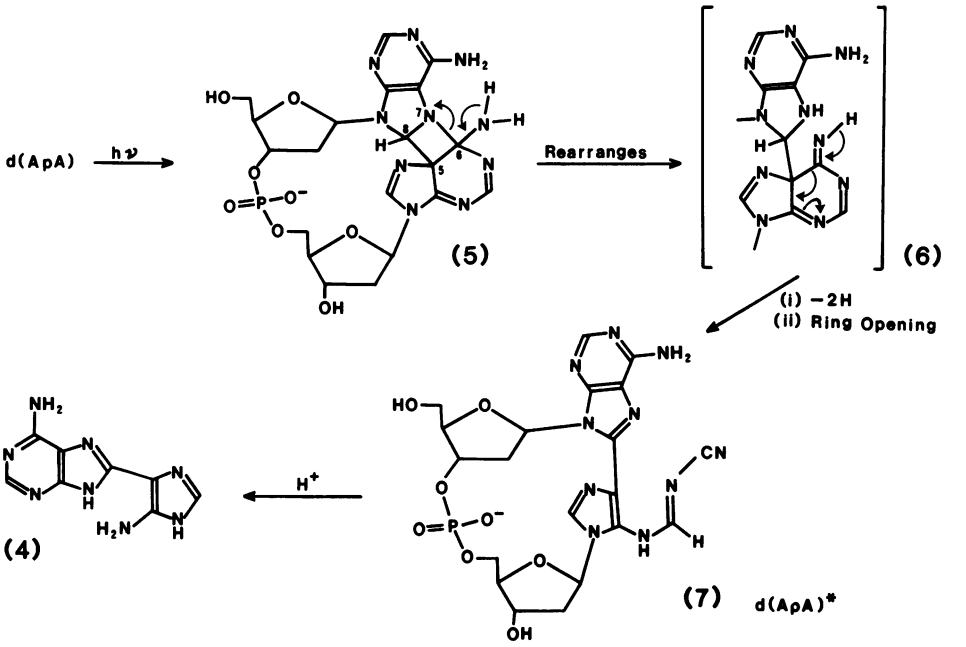
Scheme 1. Synthetic route for preparation of 8-(5-aminoimidazol-5-yl)adenine (**4**).

Acid degradation of d(ApA)*

Further information concerning the structure of the heterocyclic base moiety of d(ApA)* was sought by attempting to liberate it from the deoxyribose-phosphate backbone by the action of acid. Since 1 M HCl, at 100°C for 1 h, degrades the photoproduct A=A specifically to 4,6-diamino-5-guanidinopyrimidine, the same conditions were investigated with d(ApA)*. This treatment converted d(ApA)* in high yield (>60%) into a heterocyclic base with an absorption maximum at 320 nm (Fig. 1); a small quantity of adenine was also produced. The main hydrolysis product was purified to homogeneity by RP-HPLC and determined, by mass spectrometry, to have the elemental composition $C_8H_8N_8$ which requires a total of nine rings and double bonds. Six of the eight hydrogen atoms are readily exchanged by deuterium, as determined by the FAB mass spectrometry exchange method; the remaining two hydrogens gave rise to singlet proton resonances in the aromatic region of the ^1H NMR spectrum.

Tentative identification of the liberated heterocyclic base as 8-(5-aminoimidazol-4-yl)adenine, **4**, was confirmed by chemical synthesis (Scheme 1). Condensation of 4,5,6-triaminopyrimidine (**1**) with 4-nitro-1-(tetrahydropyran-2-yl)imidazole-5-carboxaldehyde (**2**) in the presence of *p*-benzoquinone as oxidant afforded 8-(5-nitroimidazol-4-yl)adenine (**3**), the tetrahydropyranyl protecting group being lost during the reaction. The symmetrical structure of the triaminopyrimidine ensured that only a single cyclised product could be formed. Subsequent reduction of the nitro compound, **3**, by catalytic hydrogenation yielded **4**. The synthetic 8-(5-aminoimidazol-4-yl)adenine was identical with the acid hydrolysis product derived from d(ApA)* as judged by UV spectra at pH 1 and pH 7, ^1H NMR spectra and co-elution on RP-HPLC. Also, when dissolved in concentrated ammonia solution at room temperature, they both exhibited the same pattern of slow spectroscopic changes (over the course of ~ 1 h) leading to the loss of absorption in the region of 320 nm and generation of a new maximum at 274 nm.

In 1976, Rahn (*9*) reported that treatment of the Pörschke photoproduct derived from the dinucleotide $d(pA)_2$ with 90% formic acid, at 100° for 2 h, gave rise to a substance absorbing maximally at 310 nm which migrated as a fluorescent spot on paper chromatography. We obtained similar results when d(ApA)* was hydrolysed in this fashion. Examination of the hydrolysate by RP-HPLC revealed a single major component plus lesser amounts of adenine and compound **4**. When the major component was heated with 1 M HCl, at 100°C for 1 h, it was, as judged by RP-HPLC, converted almost entirely into **4**; conversely, heating **4** with formic acid reversed the process. On this evidence, we conclude that the formic acid hydrolysis product described by Rahn is an *N*-formyl derivative of **4**. Although



Scheme 2. Mechanism proposed to account for formation of the Pörschke photoproduct d(ApA)* on UV irradiation of d(ApA). For clarity, the deoxyribose-phosphate backbone has been omitted from structure **6**.

fluorescent on paper, it emits very weakly in solution; compound **4**, by comparison, is non-fluorescent.

Structure assignment

It is clear from its mass and elemental composition that the acid hydrolysis product **4** does not constitute the intact heterocyclic base moiety of d(ApA)*. Nevertheless, the structure of **4** implies that the conversion of d(ApA) into d(ApA)* entails covalent bond formation between C(8) of one adenine and C(5) of the other. The net loss of two carbon and two nitrogen atoms during acid hydrolysis evidently occurs from the C(5) substituted adenine nucleus whose structure was elucidated as follows.

Experimental and theoretical studies on 5-methyladenine suggest (*17*) that a 5-substituted adenine nucleus will be unstable and tend to undergo ring opening to an isomeric *N*-cyanoformamidino-imidazole derivative as illustrated in Scheme 2 for the transformation of **6** into **7**. To evaluate this possibility the properties of d(ApA)* were compared with those of the model compound *N*-cyano-*N*^{1}-(1-methylimidazol-5-yl)formamidine, **8**, whose synthesis was reported recently (*18*).

A key observation is that the infrared spectrum (Fig. 4) of d(ApA)* exhibits a strong absorption peak at 2201 cm^{-1} attributable to a cyano group. A corresponding peak occurs at 2194 cm^{-1} in the infrared spectrum of **8** confirming its existence as the cyano tautomer. The ^{1}H NMR spectrum of **8** shows signals for two distinct components which, in common with oximes and imines (*19*), arise from equilibration between the *syn* and *anti* conformers of the cyano group about the azomethine bond (*20*). This phenomenon can account for the fact that the signals for the three non-exchangeable protons in the heterocyclic base moiety of d(ApA)* are resolved into five singlet peaks in the aromatic region of its ^{1}H NMR spectrum (Fig. 3). The two protons attached to the *N*-cyanoformamidino-imidazole portion would each be expected to give rise to two peaks while the signal for the remote adenine C(2) proton is likely to be

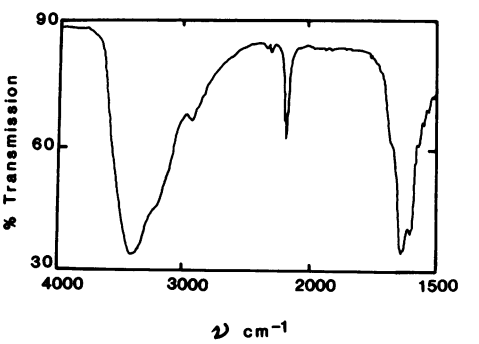
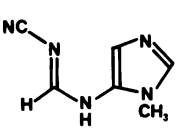
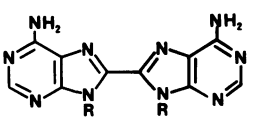


Figure 4. IR spectrum of the photoproduct d(ApA)* between 1500 and 4000 cm^{-1}.



(8)



(9) R = β-D-ribofuranosyl

(10) R = H

unaffected. The shape of the signal for the anomeric proton at higher field (δ 5.99) is also suggestive of it comprising two closely spaced multiplets. Compound **8** exhibits a spectroscopic pK ~8.0 which was shown by paper electrophoresis to correspond to dissociation of a proton from the neutral molecule to give the monoanion. A similar ionization step could explain the otherwise curious observation that d(ApA)* migrates effectively as a dianion on electrophoresis at pH 8.7 (*7,11*). Furthermore, the UV spectral changes associated with this ionization in **8** qualitatively resemble those reported for the equivalent titration of d(pA)$_2$* by Pörschke (*7*). Finally, in keeping with the formation of **4** from

d(ApA)*, treatment of **8** with 1 M HCl, at 100°C for 1 h, degrades it to 5-amino-1-methylimidazole with the loss of two carbon and two nitrogen atoms.

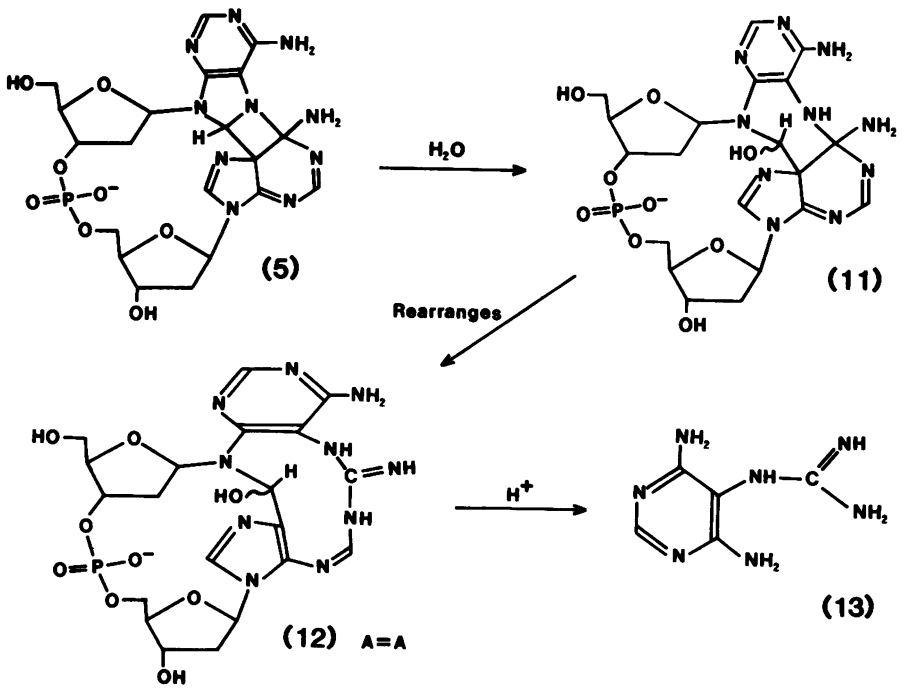
Accordingly, there is convincing evidence to suggest that d(ApA)* contains an *N*-cyanoformamidino-imidazole function. All the foregoing chemical and spectroscopic information can then be satisfactorily accommodated if d(ApA)* is assigned structure **7** (Scheme 2) in which the 5'-adenine of d(ApA) is linked through its C(8) position to C(4) of the substituted imidazole ring derived from the 3'-adenine. The choice of this polarity is discussed below. The proposed structure **7** is compatible with the intense UV absorption of d(ApA)* which is a distinctive property of the Pörschke-type of photoproduct (*7*). However, steric constraints imposed by the deoxyribose-phosphate backbone must prevent the adenine and imidazole rings from being coplanar because the absorption spectrum of d(ApA)* is approximately the sum of its individual chromophores. By contrast, in compound **4** the absorption maximum of 320 nm is indicative of extended conjugation between the two rings.

At this point it is appropriate to comment further on the earlier proposal by Gasparro and Fresco (*10*) that the Pörschke photoproduct is formed by covalent linking of adjacent adenine bases through their respective C(8) positions to give a dehydrodimer. This assignment was based largely on the observation that solutions of UV-irradiated deoxyadenylates possess fluorescence characteristics resembling those reported (*15*) for the (8-8) coupled biribonucleoside 8-(8-adenosyl) adenosine (**9**). In this respect, it is noteworthy that when UV-iradiated d(ApA) is fractionated on paper strips by electrophoresis the d(ApA)* photoproduct appears as a dark band when viewed under UV light *i.e.* it is not appreciably fluorescent. However, minor photodecomposition products, giving rise to bands with a bluish fluorescence, are also evident and it is likely that they constitute the species monitored by Gasparro and Fresco.

To rigorously exclude (8-8) coupling of the adenine bases in d(ApA)*, acid hydrolysates of the photoproduct were screened by RP-HPLC for the presence of 8-(8-adeninyl)adenine (**10**) which should have been liberated under the experimental conditions. The latter material, **10**, was prepared by hydrolysis of the biribonucleoside **9** with either HCl or formic acid and characterised by high resolution mass spectrometry. No traces of **10** were detected in HCl or formic acid hydrolysates of d(ApA)*; its chromatographic retention time and UV absorption spectrum (Fig. 1) were quite distinct from those of the major HCl hydrolysis product **4**. Hence, an (8-8) linkage of the bases in d(ApA)* can be discounted though Gasparro and Fresco (*10*) were partially correct in identifying the photoproduct as a dehydrodimer involving bond formation at one (but not both) of the C(8) positions.

Mechanistic implications

An important deduction arising from the structure assigned to d(ApA)* is that its formation can now be rationalised in terms of the same singlet-state photochemical process that we previously invoked (*11*) to account for formation of the other d(ApA) photoproduct, A=A. A [2+2] photocycloaddition reaction is assumed to occur linking the N(7) and C(8) atoms of the 5'-adenine in d(ApA) to the C(6) and C(5) atoms of the 3'-adenine. With this polarity, the reacting centres will be favourably juxtaposed for cycloaddition when the two bases adopt their normal stacked right-handed helical conformation in solution (*21*); there is also precedent (*22*) for a similar cyclisation involving C(5) and C(6) of the 3'-adenine during formation of the intramolecular photoadduct of d(TpA). The primary photochemical step generates an unstable azetidine species **5** which acts as a common precursor to both d(ApA) photoproducts. Its decomposition by hydrolytic fission (*11*) leads to A=A (Scheme 3) while a competing reaction pathway, shown in



Scheme 3. Proposed mechanism of formation of the photoproduct A=A from the common azetidine precursor (**5**).

Scheme 2, yields d(ApA)*. We propose that a prototropic shift causes fission of the highly strained azetidine ring producing the 7,8-dihydroadenine species **6** as an unstable intermediate. This mode of azetidine ring fission resembles that which occurs during the formation of bipyrimidine (6-4) photoadducts involving a 3'-cytosine base (*23*). The isolated photoproduct, d(ApA)*, is then derived from **6** by rapid oxidation of the reduced purine ring to generate a C(8)-substituted deoxyadenosine moiety accompanied by rearrangement of the 5-substituted adenine nucleus as discussed above. A unified and self-consistent explanation of photodimerization in oligodeoxyadenylate sequences is thus achieved.

Finally, it should be noted that, in view of the propensity for a 5-substituted adenine nucleus to undergo rearrangement (*17*), we consider that the structure **11** originally assigned (*11*) to A=A should be revised to the ring opened form **12** (Scheme 3). The latter incorporates a substituted guanidino function which more satisfactorily accounts for A=A having no net charge at neutral pH and for its ready hydrolysis by acid to 4,6-diamino-5-guanidinopyrimidine (**13**).

Concluding comments

The overall quantum yield for adenine photodimerization in DNA can be determined as the sum of the individual quantum yields for formation of the isolated photoproducts A=A and d(ApA)*. Although these species are produced in roughly equal amounts in UV-irradiated d(ApA) a different photoproduct ratio may pertain following adenine dimerization in other sequence environments and/or double stranded DNA. The photodimerization reaction is very sensitive to the conformational state of the polynucleotide chain. Thus, it has an absolute requirement for a DNA backbone linking the bases (*24*) and is known from photofootprinting (*4,8*) and other studies (*9,25*) to be strongly quenched by base pairing.

We recently assayed the A=A content of UV-irradiated polynucleotides and DNA (*25*) by using RP-HPLC to recover the 4,6-diamino-5-guanidinopyrimidine produced after acid hydrolysis. A similar assay can now be devised for d(ApA)* by taking advantage of its acid hydrolysis, under the same conditions, to give **4** which can also be detected and isolated by RP-HPLC. The availability of synthetic material for use as a carrier will facilitate sensitive detection of **4** when samples containing radiolabelled adenine are analysed. By applying both types of assay it will be possible to obtain reliable quantum yields for adenine photodimerization in DNA and also to monitor excision of the resultant photoproducts by repair enzymes. This information should provide the basis for a realistic assessment of the biological importance of adenine photodimerization.

Although adjacent adenine bases on an RNA backbone (*e.g.* in polyriboadenylic acid) do not undergo photodimerization (*6,24*), efficient and specific photocrosslinking has recently been observed (*26*) between two non-contiguous adenosines (A57 and A95) in a shortened version of the *Tetrahymena* ribozyme. It will be interesting to determine whether this reaction is related mechanistically to the type of photodimerization process discussed above.

ACKNOWLEDGEMENTS

We thank Dr C.A.Ramsden (Dagenham Research Centre, Rhone-Poulenc Ltd.) for kindly providing us with a sample of *N*-cyano-N^1-(1-methylimidazol-5-yl)formamidine (**8**). This work was supported by grants GR/D 40609 and GR/E 48060 from the Science and Engineering Research Council to R.J.H.D. and by grant GM 29812 from the National Institutes of Health to J.A. McC.

REFERENCES

1. Harm, W. (1980) Biological Effects of Ultraviolet Radiation. Cambridge University Press. Cambridge.
2. Cadet, J. and Vigny, P. (1990) in Morrison, H. (ed.), Bioorganic Photochemistry. John Wiley & Sons, New York, Vol. 1, pp. 1–272, and references therein.
3. Duker, N.J. and Gallagher, P.E. (1988) *Photochem. Photobiol.*, **48**, 35–39.
4. Becker, M.M. and Wang, Z. (1989) *J. Mol. Biol.*, **210**, 429–438.
5. Gallagher, P.E. and Duker, N.J. (1989) *Photochem. Photobiol.*, **49**, 599–605.
6. Pörschke, D. (1973) *Proc. Natl. Acad. Sci. USA*, **70**, 2683–2686.
7. Pörschke, D. (1973) *J. Am. Chem. Soc.*, **95**, 8440–8446.
8. Becker, M.M., Lesser, D., Kurpiewski, M., Baranger, A. and Jen-Jacobsen, L. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6247–6251.
9. Rahn, R.O. (1976) *Nucleic Acids Res.*, **3**, 879–890.
10. Gasparro, F.P. and Fresco, J.R. (1986) *Nucleic Acids Res.*, **14**, 4239–4251.
11. Kumar, S., Sharma, N.D., Davies, R.J.H., Phillipson, D.W. and McCloskey, J.A. (1987) *Nucleic Acids Res.*, **15**, 1199–1216.
12. Wu, D.C.J., Cheer, C.J., Panzica, R.P. and Abushanab, E. (1982) *J. Org. Chem.*, **47**, 2661–2663.
13. Sethi, S.K., Smith, D.L. and McCloskey, J.A. (1983) *Biochem. Biophys. Res. Commun.*, **112**, 126–131.
14. Verma, S., Pomerantz, S.D., Sethi, S.K. and McCloskey, J.A. (1987) *Anal. Chem.*, **58**, 2898–2902.
15. Joshi, P.C. and Davies, R.J.H. (1981) *J. Chem. Res. (S)*, 227; *J. Chem. Res. (M)*, 2701–2732.
16. Sethi, S.K., Gelijkens, C.F., McCloskey, J.A., Shapiro, R. and Brandes, B.A. (1983) *Biomed. Mass Spectrom.*, **10**, 665–670.
17. Hosmane, R.S., Bakthavachalam, V. and Leonard, N.J. (1982) *J. Am. Chem. Soc.*, **104**, 235–239.
18. Al-Shaar, A.H., Gilmour, D.W., Lythgoe, D.J., McClenaghan, I. and Ramsden, C.A. (1989) *J. C. S. Chem. Commun.*, 551–552.
19. McCarty, C.G. (1970) In Patai, S. (ed.), The Chemistry of the Carbon-Nitrogen Double Bond. John Wiley & Sons, New York, pp. 363–464.
20. Lythgoe, D. (1987) Ph.D. thesis, CNAA, UK.
21. Olsthoorn, C.S.M., Bostelaar, L.J., van Boom, J.H. and Altona, C. (1980) *Eur. J. Biochem.*, **112**, 95–110.
22. Koning, T.M.G., Davies, R.J.H. and Kaptein, R. (1990) *Nucleic Acids Res.* **18**, 277–284.
23. Wang, S.Y. (1976) in Wang, S.Y. (ed.), Photochemistry and Photobiology of Nucleic Acids. Academic Press, New York, Vol. I, pp. 295–356.
24. Pörschke, D. (1975) *Nucleic Acids Res.*, **2**, 1621–1627.
25. Sharma, N.D. and Davies, R.J.H. (1989) *J. Photochem. Photobiol. B: Biol.*, **3**, 247–269.
26. Downs, W.D. and Cech, T.R. (1990) *Biochemistry*, **29**, 5605–5613.