Isolation and Identification of Cross-Links from Formaldehyde-Treated Nucleic Acids†

Yu Fen M. Chaw,‡ L. Eric Crane, Peter Lange, and Robert Shapiro*

ABSTRACT: Cross-linked nucleosides have been isolated from formaldehyde-treated ribonucleic acid (RNA) and deoxyribonucleic acid (DNA) by using reverse-phase high-pressure liquid chromatography. Methylene-bridged products containing cytosine, as well as the purines, have been isolated. A combination of ultraviolet and nuclear magnetic resonance (NMR) spectra, pK_a values, hydrolysis, and reduction has been used to establish that the cross-links connect the amino groups of the nucleosides involved. For example, the 6-amino functions of two adenosine residues are linked by a methylene bridge to produce a compound which, when dissolved in deuterated dimethyl sulfoxide, displays 11 unsplit resonances in its proton-decoupled ^{13}C NMR spectrum. The procedure reported here is more rapid and less laborious than an earlier one recommended by us for the isolation of cross-linked products from DNA [Dubelman, S., & Shapiro, R. (1977) *Nucleic Acids Res. 4*, 1815–1827]. This new approach may be of value in the study of other types of cross-linking reactions involving nucleic acids.

Formaldehyde has been prominent as a mutagen (Auerbach et al., 1977) and carcinogen (*Nature (London)*, 1979), as well as a reagent for the chemical modification of DNA and RNA (Feldman, 1973). It undergoes two different reactions with the nucleic acids and their components. A rapid and reversible one converts –NH– groups and amino groups to $-NCH_2OH$ derivatives (Fraenkel-Conrat, 1954; Feldman, 1973). A slower reaction results in the formation of cross-links. This reaction occurs at both the monomer and polymer level (Feldman, 1973). The products persist, at neutral pH, after the formaldehyde has been removed.

A number of known genetic effects of formaldehyde are considered to be a result of the cross-linking reaction. These effects include mutations in *Escherichia coli*, recombination in yeast, lethality in both *E. coli* and yeast, and the mutations that result in *Drosophila* when formaldehyde is included in the food of the young larvae (Auerbach et al., 1977).

Despite their biological importance, the chemical structures of the cross-links have not been established, for technical reasons. To cite a recent review (Auerbach et al., 1977), "...there are still no methods available to isolate these bridges from DNA".

As part of a program to develop improved procedures in this area, we recently published a method for the isolation of the cross-links induced by nitrous acid treatment of DNA (Dubelman & Shapiro, 1977). This method was effective, but laborious, and, in part, took advantage of the ionic character of the cross-linked nucleosides involved. We have now devised a more rapid method, suitable for less polar substances, and used it to isolate formaldehyde cross-links from DNA and RNA. The cross-links involve cytosine, as well as the purines, whose presence has been previously recognized (Feldman, 1973). By a combination of chemical and spectroscopic methods, we have demonstrated that the cross-links connect exclusively through the exocyclic amino groups of the bases involved, in both RNA and DNA.

Experimental Procedures

Reagents and Spectroscopic and Chromatographic Methods. High-pressure LC[1] was performed with a Chromatronix 3500 B equipped with a programmable gradient control and either a Chromatronix Vydac ion-exchange column (AX-107, 0.21 × 150 cm; CX-107, 0.21 × 100 cm) or a Waters Associates C_{18} reverse-phase column (10 μm, 0.39 × 30 cm). Distilled in glass methanol was obtained from Burdick and Jackson. UV spectra were recorded with either a Cary Model 15 or a Varian Techtron MLC spectrophotometer. The pH was measured with a Radiometer PHM 62 meter equipped with a G2322C electrode. NMR spectra were obtained with a 60-mHz Perkin-Elmer R-20B, a Varian XL-100 (^1H and ^{13}C), and a Varian HR/NTC TT-220 located at Rockefeller University. ^1H and ^{13}C NMR spectral data were obtained at the specified frequency in dimethyl-d_6 sulfoxide unless otherwise noted, except for 220-mHz spectra which were all taken in dimethyl-d_6 sulfoxide containing up to 5% D_2O. Sonification was performed with either a Model 14 ultrasonic cleaner or a Model W185 sonifier–cell disrupter both made by Heat Systems-Ultrasonics. A small quantity of 90% enriched formaldehyde was provided, courtesy of Dr. Steve Taylor and the Research Laboratories of Du Pont. The sources for DNA and enzymes have been given previously (Dubelman & Shapiro, 1977). Ultraviolet spectra were taken in solutions buffered by HCl, acetate, Tris,[1] phosphate, or KOH at the pH values indicated. Yeast RNA (Schwarz/Mann, Orangeburg, NY) was purified by perchloric acid reprecipitation, followed by washes with ethanol and water. Formaldehyde solution, 37% by weight, was purchased from Fisher Scientific Co., Fairlawn, NJ. R_f values (TLC)[1] were obtained by using Polygram Cel 400 0.1-mm cellulose plates (Machery-Nagel and Co.) with development by solvent A (0.1 M sodium metaborate, pH 9.6–2-propanol, 1:2 v/v) or solvent B (0.1 M Na_2HPO_4 adjusted to pH 12.0 with KOH–2-propanol, 1:2 v/v).

† From the Department of Chemistry, New York University, New York, New York 10003. *Received May 6, 1980.* This research was supported by a Grant GM 20176 from the National Institutes of Health, U.S. Public Health Service.

‡ Present address: Institute of Chemistry, Academica Sinica, Peking, China.

[1] Abbreviations used: high-pressure LC, high-pressure liquid chromatography; OAc, acetate; Tris, 2-amino-2-(hydroxymethyl)-1,3-propanediol; TLC, thin-layer chromatography; Ex.D, exchangeable with D_2O. In cross-links, Ado, adenosyl, Guo, guanosyl, and Cyd, cytidyl (otherwise the free nucleoside is indicated); prefix d-, 2′-deoxyribonucleoside; Ade, adenine; Gua, guanine; Thd, thymidine.

0006-2960/80/0419-5525$01.00/0

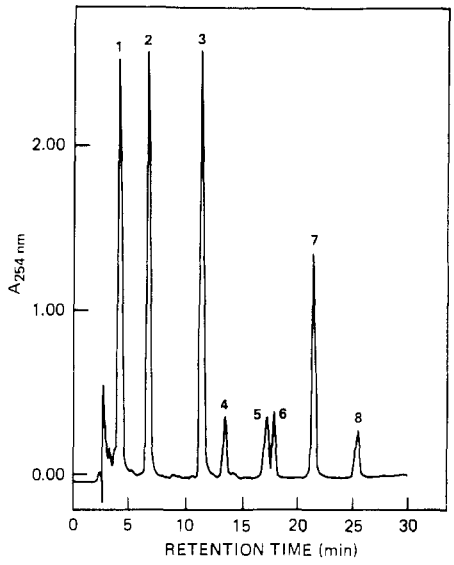


FIGURE 1: Separation of RNA hydrolysate by reverse-phase high-pressure liquid chromatography. Identity of peaks: 1, Cyd plus Urd; 2, Guo; 3, Ado; 4, Cyd-CH_2-Guo; 5, Guo-CH_2-Guo; 6, Ado-CH_2-Cyd; 7, Ado-CH_2-Guo; 8, Ado-CH_2-Ado.

Reaction of RNA with Formaldehyde. A solution containing 800 mg (16 700 A_{260} units) of purified yeast RNA and 0.2 M formaldehyde in 133 mL of 0.05 M sodium acetate–acetic acid buffer, pH 4.8, was prepared. The solution was kept at 25 °C for 18 days and then dialyzed against 3.5 L of distilled water for 150 h (the water was changed after 20, 44, 68, and 146 h). The sample was lyophilized, redissolved in 50 mL of 0.5 M KOH, and incubated at 37 °C for 20 h. The solution was neutralized with $HClO_4$ and the resulting $KClO_4$ precipitation removed by filtration. The filtrate was lyophilized and redissolved in 5 mM Tris-HCl buffer, pH 8.4.

Half of the RNA hydrolysate (8120 A_{260} units) was brought to a volume of 38 mL by the addition of 0.1 M Tris-HCl buffer, pH 8.2. Alkaline phosphatase (72 units) was added, and the solution was kept at 28 °C for 6 h. A precipitate that had formed was removed by centrifugation, and the reaction mixture was analyzed by high-pressure LC as described below.

Reaction of DNA with Formaldehyde. A solution was prepared containing 312 mg (6200 A_{260} units) of calf thymus DNA in 62 mL of 0.17 M formaldehyde and 0.04 M sodium acetate–acetic buffer, pH 4.5. The solution was stirred at 25 °C for 40 days and then dialyzed against 5 mM Tris-HCl and 1 mM NaCl (pH 7.5) for 10 days (the buffer was changed after 3 and 7 days). One-fifth of the sample in the dialysis bag was diluted to 50 mL with additional 5 mM Tris-HCl and 1 mM NaCl solution. Sufficient $MgCl_2$ was added to bring its concentration to 1 mM, and the pH was adjusted to 8.8. Deoxyribonuclease I (2 mg, 4000 units) was added, and the solution was kept at 30 °C for 16 h. Venom phosphodiesterase (1.5 mg, 30 units) and spleen phosphodiesterase were introduced, and incubation was continued for another 6 h. Finally, alkaline phosphatase (150 μL, 250 units) was added, and then the hydrolysis was allowed to proceed for another 6 h. The reaction mixture was analyzed by high-pressure LC, as described below.

Isolation of Cross-Linked Nucleosides by Reverse-Phase High-Pressure LC. In analytical runs, 25-μL samples of a hydrolysate of RNA or DNA were fractionated on a μBondapak C_{18} column (3.9 mm × 30 cm, Waters Associates, Milford, MA). A linear water–methanol gradient varying from 10 to 70% methanol (Essigman et al., 1977) was used over 40 min at ambient temperature with a pressure of 1600 psi and a flow rate of 1 mL/min. The separations achieved with typical runs with RNA and DNA hydrolysates are illustrated in Figures 1 and 2, respectively.

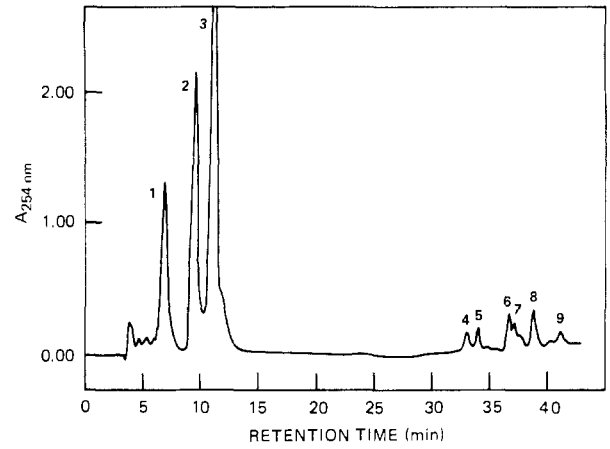


FIGURE 2: Separation of DNA hydrolysate by reverse-phase high-pressure liquid chromatography. Identity of peaks: 1, dCyd plus dThd; 2, dGuo; 3, dAdo; 4, dCyd-CH_2-dGuo; 5, dGuo-CH_2-dGuo; 6, dAdo-CH_2-dCyd; 7, unknown; 8, dAdo-CH_2-dGuo; 9, dAdo-CH_2-dAdo.

In large scale runs, a 5-mL sample loop (Waters Associates, Milford, MA) was utilized to lead 5-mL samples of solution onto the above column. Elution was conducted with 10% methanol in water for 60 min, and then the linear gradient was applied to raise the methanol concentration to 70% over 60 min. In each run with the RNA hydrolysate, ~1100 A_{260} units was applied to the column. The yields of cross-linked nucleosides were approximately 30 A_{260} units for peak 7 (see Figure 1) and 8 A_{260} units each for peaks 4, 5, 6, and 8.

In the DNA work, which was initiated at a later date than the RNA work, it was found advantageous to fractionate the hydrolysate before large-scale chromatography by using a C_{18} SEP-PAK cartridge (Waters Associates, Milford, MA). A 5-mL sample containing 270 A_{260} units of DNA hydrolysate was introduced into the cartridge, which was then eluted with 1 mL of water. Only mononucleotides were eluted (as determined by a run on the reverse-phase column). The cross-links, and some mononucleosides, were then eluted with 10 mL of methanol, while enzymes were retained in the cartridge. The methanol was evaporated, and the sample was redissolved in 5 mL of water and fractionated on the reverse-phase column, as described above for RNA. In the case of peaks 6 and 7, which partly overlapped (see Figure 2), only the noncontaminated portion of each peak was used for further work. Their purity was checked by rechromatography on the reverse-phase column. Yields were about 1 A_{260} unit for peaks 4, 5, 6, 7, and 9 and 1.5 A_{260} units for peak 8 (see Figure 2).

Hydrolysis of Cross-Linked Nucleosides. A sample of 0.5–1.0 A_{260} unit of each cross-linked nucleoside was added to 6 mL of 1 N HCl, and the reaction mixture was allowed to stand for 5 h at 60 °C. The solution was then lyophilized and redissolved in water. The composition of the products was determined by chromatography on the reverse-phase column described above. Typical results, from the hydrolysis of peak 7, Figure 1, are illustrated in Figure 3. For those peaks that contained cytosine nucleosides, the above time was insufficient to complete the hydrolysis. These samples had to be redissolved in 1 N HCl, heated at 90 °C for 1 h, and then analyzed by reverse-phase chromatography. Each product of hydrolysis was identified by its characteristic retention volume and by its ultraviolet spectra at acid, neutral, and basic pH.

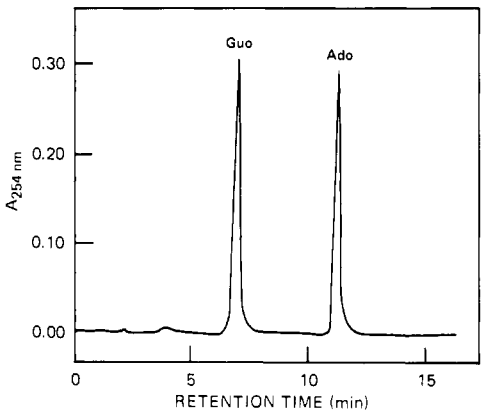


FIGURE 3: Separation of acidic hydrolysate of Ado-CH_2-Guo by reverse-phase high-pressure liquid chromatography.

Di(N^6-adenosyl)methane (Ia). Adenosine (0.81 g, 3.0 mmol), 3.7% formaldehyde solution (15 mL, 18.5 mmol), and 185 mL of 0.2 M sodium acetate buffer, pH 4.5, were combined and stirred for 16 h at 25 °C. The solution was then stirred for an additional 265 h at 45 °C. The precipitate which had formed was separated by filtration, washed with 100 mL of distilled water, and dried under vacuum at 50 °C for 16 h. The white solid obtained weighed 0.33 g (40%): mp 273–275 °C dec [lit. 277 °C (Feldman, 1973)]. The ultraviolet spectrum of the compound agreed with that reported (Feldman, 1973). The compound was homogeneous on cellulose TLC plates: solvent A, R_f = 0.73; solvent B, R_f = 0.74. It also appeared as a single peak when passed separately through three different high-pressure LC columns. The retention volumes were 25.4 mL (C_{18} reverse phase, conditions given under Isolation of Cross-Linked Nucleosides by Reverse-Phase High-Pressure LC), 21.0 mL (Vydac CX-107 cation exchange, 3.9 × 250 mm, linear gradient beginning at 10^{-3} M $(NH_4)_2SO_4$ plus 10^{-3} M NH_4OAc, pH 4.40, and ending 40 min later at 0.25 M $(NH_4)_2SO_4$ plus 10^{-3} M NH_4OAc, pH 4.60; flow rate = 1 mL/min), and 19.0 mL (Vydac AX-107 anion exchange, 3.9 × 300 mm, 0.005 M sodium phosphate, pH 10.54; flow rate = 1 mL/min). NMR (^{1}H, 60 mHz, δ): H-2 (8.48, 2 H), H-8 (8.35, 2 H), N^6-H (8.22, 2 H, Ex.D), H-1′ (5.95, 2 H, d), N^6–CH_2–N^6 (5.41, 2 H). NMR (^{1}H, 220 mHz, δ): H-2 (8.34), H-8 (8.31), H-1′ (5.88), H-2′ (4.56), H-3′ (4.15), H-4′ (3.98), H-5′ (3.65, 3.59), N^6–CH_2–N^6 (5.34). NMR (^{13}C, 25.2 mHz, δ): C-1′ (87.8), C-4′ (85.6), C-2′ (73.4), C-3′ (70.3), C-5′ (61.4), N^6–CH_2–N^6 (47.7). The remainder of the ^{13}C NMR, ultraviolet, and pK_a data are given in Tables I and II.

Di(N^6-guanosyl)methane (IIa). Guanosine (1.3 g, 4.2 mmol), 3.7% formaldehyde solution (20 mL, 24.7 mmol), and 200 mL of 0.2 M sodium acetate buffer, pH 4.5, were combined and stirred at 50 °C for 9 days. A precipitate which had formed was separated from the supernatant by removing the latter with a pipet. The precipitate was washed thoroughly with 10 mL of water at 60 °C by using a sonicator to agitate the mixture for 10 min. The resulting slurry was centrifuged and the supernatant removed with a pipet. This procedure was repeated 5 times, and the final precipitate was dried under vacuum at 50 °C for 24 h. The product weighed 0.37 g (30%): mp 250–253 °C. It was homogeneous upon cellulose TLC in solvent A, with an R_f = 0.09, and in solvent B, with an R_f = 0.15. It also appeared as a single peak when passed separately through three different high-pressure LC columns with retention volumes of 22.2 mL (C_{18} reverse phase), 9.0 mL (Vydac CX-107, conditions as specified for Ia), and 3.8 mL (Vydac AX-107, conditions as specified for Ia). The analytical sample was prepared by further drying (100 °C, 0.1 Torr, 24 h). Anal. Calcd for $C_{21}H_{26}N_{10}O_{10}$ (found): C, 43.60 (43.55); H, 4.50 (4.53); N, 24.22 (24.08). NMR (^{1}H, 60 mHz, δ): N^1-H (10.81, 2 H, Ex.D), H-8 (8.03, 2 H), N^2-H (7.15, 2 H, Ex.D), H-1′ (5.88, 2 H, d), N^2–CH_2–N^2 (4.91, 2 H). NMR (^{1}H, 220 mHz, δ): H-8 (8.00), H-1′ (5.84), H-2′ (4.47), H-3′ (4.15), H-4′ (3.96), H-5′ (3.64, 3.61), N^2–CH_2–N^2 (4.88). NMR (^{13}C, 25.2 mHz, δ): C-1′ (86.7), C-4′ (85.0), C-2′ (74.0), C-3′ (70.1), C-5′ (61.1), N^2–CH_2–N^2 (47.1). The remainder of the ^{13}C NMR, ultraviolet, and pK_a data are given in Tables I and II.

Borohydride Reduction of Di(N^6-adenosyl)methane (Ia) and Di(N^2-guanosyl)methane (IIa). A sample of 1 mg of Ia and another of 2 mg of IIa were each dissolved separately in 0.25 M KOH (3 mL). Sodium borohydride was added (50 mg to Ia and 100 mg to IIa), and the mixtures were stirred at 60 °C. At intervals, aliquots were withdrawn, quenched by the addition of glacial acetic acid, drop by drop, until bubbling ceased, and analyzed by cation exchange high-pressure LC in the manner specified for analysis of Ia. After 20 h, Ia had been completely converted to adenosine and N^6-methyladenosine in a ratio of 5:1. After 40 h, 80% of IIa had reacted and had been converted into guanosine and N^2-methylguanosine in a ratio of 3:1 (Scheme I). The products were identified by comparison of their elution volumes and their ultraviolet spectra at various pH values with data obtained by using authentic samples.

N^6-Adenosyl-N^2-guanosylmethane (IIIa). Guanosine (1.03 g, 3.6 mmol), adenosine (0.32 g, 1.2 mmol), 3.7% formaldehyde (20 mL, 24.7 mmol), and 175 mL of 0.2 M sodium acetate buffer, pH 4.5, were stirred together at 50 °C for 9 days. The initially cloudy solution became clear after 1 day, and no precipitate developed during the following 8 days. The volume of the solution was reduced to 70 mL by rotary evaporation under vacuum to yield a viscous liquid. The concentrated reaction mixture was treated with an additional portion of 3.7% formaldehyde (15 mL, 18.5 mmol) and stirred at 45 °C for 5 more days to afford a small amount of white precipitate. The precipitate was filtered, washed with 100 mL of water, and dried under vacuum at 50 °C for 24 h to afford 20 mg of white powder. Portions of this crude product were dissolved in aqueous dimethyl sulfoxide and subjected to cation-exchange and reverse-phase high-pressure LC under standard conditons. The mixture was resolved into three components upon each column. Two of them were identified as Ia and IIa by comparison of ultraviolet spectra, pK_a, and retention volume with authentic samples. The third product, intermediate in retention volume between Ia and IIa (20.0 mL, reverse phase, conditions as specified for Ia), was identical in its high-pressure LC retention volumes, spectral properties, and behavior upon hydrolysis with the RNA product, Ado-CH_2-Guo (IIIa). NMR (^{1}H, 220 mHz, δ): (Ado) H-2 (8.37), H-8 (8.34), H-1′ (5.89), H-2′ (4.56), H-3′ (4.15), H-4′ (3.98, H-5′ (3.64, 3.60); N^6–CH_2–N^2 (5.06); (Guo) H-8 (7.99), H-1′ (5.83), H-2′ (4.49), H-3′ (4.15), H-4′ (3.98), H-5′ (3.64, 3.60).

N^6-Adenosyl-N^4-cytidylmethane (IVa) and N^2-Guanosyl-N^4-cytidylmethane (Va). Two mixtures, one with cytidine (0.5 g, 2 mmol) and adenosine (0.2 g, 0.67 mmol) and the other with cytidine (0.74 g, 3 mmol) and guanosine (0.29 g, 1 mmol), were separately combined with 8-mL portions of a 3.7% formaldehyde solution (9.9 mmol) and 10-mL portions of 0.2 M sodium acetate buffer, pH 4.5. Each reaction was stirred for 168 h at 68 °C without noticeable precipitation. Each reaction proved to be a complex mixture upon analysis by reverse-phase high-pressure LC. The peak with retention

volumes corresponding to the RNA product Ado-CH_2-Cyd was collected by repeated injection of 25-μL aliquots of the first reaction mixture. Similarly, the peak corresponding to the RNA product Cyd-CH_2-Guo was collected from the second reaction mixture. The nucleoside products were identical with those isolated from RNA in their ultraviolet spectra at various pH values and elution volumes as well as their ^{1}H NMR spectra. Analytical data for Ado-CH_2-Cyd (IVa): NMR (^{1}H, 220 mHz, δ) (Ado) H-2 (8.46), H-8 (8.33), H-1′ (5.93), H-2′ (4.60), H-3′ (4.16), H-4′(3.98), H-5′ (3.70, 3.57); N^6–CH_2–N^4 (5.07); (Cyd) H-6 (7.94), H-5 (5.84), H-1′ (5.79), H-2′, -3′, -4′ (3.83–3.98), H-5′ (3.58, 3.47). Other data are supplied in Tables I and II. Analytical data for Cyd-CH_2-Guo (Va): NMR (^{1}H, 220 mHz, δ) (Guo) H-8 (7.99), H-1′ (5.83), H-2′ (4.49), H-3′ (4.17), H-4′ (3.94), H-5′ (3.62); N^2–CH_2–N^4 (4.95); (Cyd) H-6 (7.95), H-5 (5.83), H-1′ (5.83), H-2′, -3′, -4′ (3.83–3.98), H-5′ (3.62). Other data are supplied in Tables I and II.

Results

Both RNA and DNA were allowed to react with formaldehyde for a number of days. Formaldehyde was removed, and labile reaction products were allowed to decompose during an extended dialysis period. This dialysis would also serve to remove any products formed from contaminating nucleotides. The RNA was then hydrolyzed to the nucleoside level by treatment with alkali, followed by alkaline phosphatase. Initial attempts to degrade the RNA to the nucleoside level by enzymatic means alone failed, because the cross-links inhibited the digestion. More success was obtained in the DNA series, where the extent of modification was less. Hydrolysis of the formaldehyde-treated DNA to nucleosides was achieved by treatment with substantial amounts of deoxyribonuclease I, venom phosphodiesterase, spleen phosphodiesterase, and alkaline phosphatase. The resulting nucleoside mixtures were then fractionated directly by reverse-phase high-pressure LC. This procedure, in one step, separated conventional nucleosides from cross-linked ones and resolved the latter into five substances in the RNA series and six in the DNA series. The combined yield of cross-linked nucleosides, in terms of products isolated, was ~6% of total nucleosides with RNA and 2% with DNA.

The cross-linked products were identified initially on the basis of their hydrolysis products in acid as summarized in Table I. The same pattern of products was obtained from both RNA and DNA, symmetrical cross-links involving only adenine and guanine and unsymmetrical ones containing adenine and guanine, adenine and cytosine, and guanine and cytosine. An additional unidentified product was observed in the DNA series. The ultraviolet spectra at several pH values and pK_a values of the corresponding products from the RNA and DNA series were almost identical (Table I). We therefore concentrated our remaining work on the more accessible RNA products and assigned the DNA structures by analogy.

In order to confirm the composition of the cross-links and to obtain additional amounts for characterization, the direct preparation of each was attempted, by reaction of the appropriate nucleoside(s) with formaldehyde. Feldman (1973) had obtained products which he thought to be cross-links on reaction of adenosine and of guanosine with formaldehyde. Chemical analysis of the adenosine product had demonstrated the presence of 2 mol of nucleoside to 1 of formaldehyde. The guanosine product had not been characterized.

The workup of the reactions of adenosine and of guanosine with formaldehyde was simplified, because of spontaneous precipitation of the products. The adenosine cross-link had an ultraviolet spectrum identical with that reported by Feldman (1973). The elution volumes of the adenosine and guanosine cross-links on our reverse-phase high-pressure LC column were identical with those of the corresponding products isolated from RNA. Analysis of the formaldehyde reaction mixtures containing two different nucleosides was more difficult. A complex mixture of products was formed. The known elution position of the RNA product upon reverse-phase high-pressure LC was used as a guide to select the proper peak from the nucleoside reaction mixture. That peak alone was collected by repeated injection of the concentrated reaction liquor onto the reverse-phase column. The quantities that could be obtained conveniently in this manner were of the same magnitude as those obtained from RNA. In each case, the RNA product had the same spectroscopic and pK_a properties as the peak of the same mobility isolated from the reaction mixture of the appropriate nucleosides.

Table I: Cross-Linked Nucleosides Isolated from RNA and DNA after Formaldehyde Treatment

compound	UV max[a] (nm)	hydrolysis products
AdoCH_2Ado	277 (0.7), 272 (7, 12.5)	Ado
GuoCH_2Guo	262 (0.8), 257 (6.9), 262–267.5 (12.4)	Guo
AdoCH_2Cyd	280 (0.7), 272 (6.8, 11.6)	Ado, Cyd
AdoCH_2Guo	271 (0.7), 267 (6.2), 270 (12.7)	Ado, Guo
CydCH_2Guo	260 sh,[b] 281 (0.8), 258, 273 (6.8), 272 (12.1)	Cyd, Guo
dAdoCH_2dAdo	277 (0.8), 272.5 (6.7, 12.1)	Ade
dGuoCH_2dGuo	262 (0.3), 257 (6.3), 262–267.5 (11.5)	Gua
dAdoCH_2dCyd	280 (0.9), 272 (7.0, 12.0)	Ade, dCyd
dAdoCH_2dGuo	271 (1.1), 267 (6.2), 270 (11.7)	Ade, Gua
dCydCH_2dGuo	259 sh, 281 (0.9), 260, 273 (6.2), 273 (11.7)	dCyd, Gua
unknown[c]	268, 275 sh[b] (0.9), 268, 280 sh[b] (6.2, 11.7)	*d*

[a] Numbers in parentheses correspond to pH values. [b] sh = shoulder. [c] peak 7, Figure 2. [d] Available samples were insufficient for determination of the hydrolysis products.

Our initial efforts to determine the position of covalent attachment of the cross-links concentrated on the adenosine and guanosine products, which were available in quantity. These products decomposed slowly to the parent nucleosides at pH values up to 9 but were stable indefinitely at pH 13 and 25 °C. The proton NMR spectra of Ia and IIa resembled those of adenosine and guanosine, respectively, save that one amino proton was missing per nucleoside residue and a new peak had appeared. The spectra could best be reconciled with the microanalyses by assuming that the new peak, derived from formaldehyde, represented two protons. Each peak in both spectra then also represented protons from two nucleoside moieties, with the exceptions noted, suggesting that the two pairs of nucleoside residues were each linked, in a symmetrical fashion, through their exocyclic amino functions.

The C_2 symmetry of cross-links Ia and IIa was further confirmed by their ^{13}C NMR spectra. Each spectrum showed 11 unsplit resonances, with the broad high-field signal assigned to the CH_2 bridge. The correctness of this assignment in the case of Ia was subsequently confirmed by the ^{13}C NMR spectra of two samples prepared with ^{13}C-enriched formaldehyde.

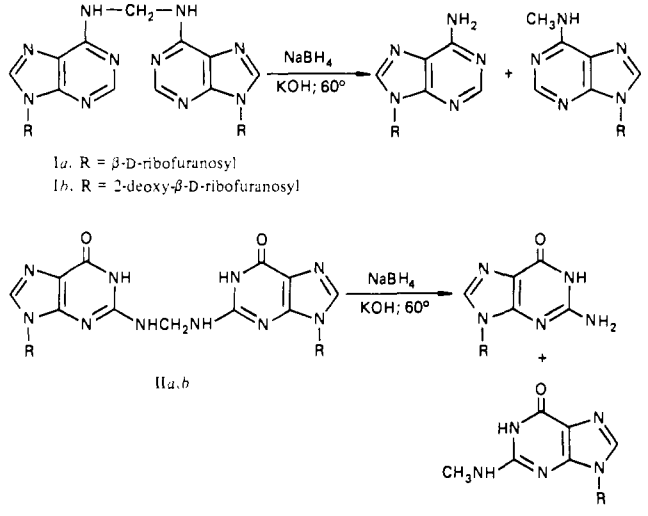
The ^{13}C NMR spectra also provided confirmation of the position of methylene attachment. Comparison of the chemical shifts of the purine ring carbons of Ia with those of the two model compounds N^6-methyladenosine and 1-methyladenosine

Table II: Comparison of Purine Ring ^{13}C Chemical Shifts in CH_2-Bridged Diadenosine and Diguanosine with Model Nucleosides[a,b]

compound	C-2	C-4	C-5	C-6	C-8
[N^6-MeAdo]	152.1	148.0	119.5	154.8	139.4
$\Delta\delta$ [c]	−0.1	+1.0	−0.2	−0.7	+0.9
Ado-CH_2-Ado	152.0	149.0	119.3	153.9	140.3
$\Delta\delta$ [c]	+3.8	+7.7	−3.1	−0.2	+2.6
[N^1-MeAdo]	148.2	141.3	122.4	154.1	137.7
[N^2-MeGuo]	152.9	150.7	116.6	156.6	136.0
$\Delta\delta$ [c]	−2.0	−1.6	−0.4	−1.3	−0.7
Guo-CH_2-Guo	150.9	149.1	116.2	155.3	135.3
$\Delta\delta$ [c]	−2.1	+1.7	+2.1	+1.0	+1.3
[N^1-MeGuo]	152.1	147.4	114.1	154.3	134.0

[a] All values were obtained with a Varian XL-100 NMR spectrometer in the FT Mode. The solvent in all cases was dimethyl-d_6 sulfoxide with a drop of D_2O added. [b] All chemical shifts are downfield from internal tetramethylsilane with an estimated error of ±0.2 ppm. [c] Values of $\Delta\delta$ were obtained by subtraction of the chemical shifts of the model compounds (in brackets) from the corresponding shifts of the CH_2-bridged dinucleosides.

Scheme I

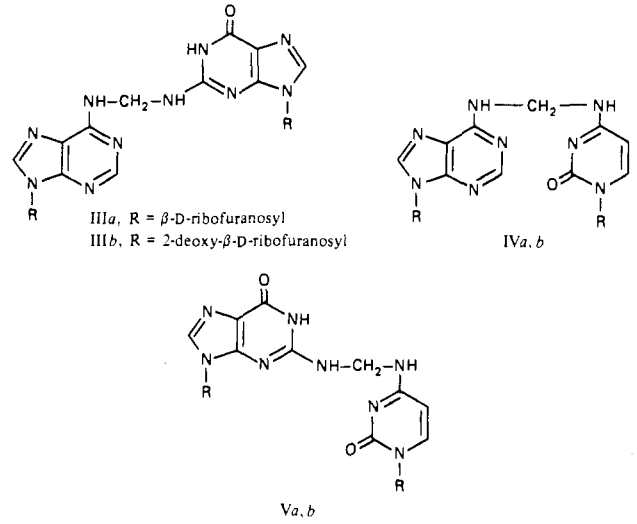


demonstrated that the spectrum of Ia more closely resembled the former than the latter (Table II). A similar study conducted with IIa, N^2-methylguanosine, and 1-methylguanosine again favored exocyclic amino group substitution, although the comparison was less decisive in this case.

Additional evidence supporting amino group attachment of the cross-links was provided by pK_a data. The ultraviolet spectrum of the adenosine cross-link Ia varied only in the pH range 1.3–4.5, indicating dissociations in that area. If cross-link attachments were at N-1 or N-3, ultraviolet spectral shifts corresponding to pK_a values would have been observed in the pH range 6–8 (Singer, 1975). The guanosine cross-link, on the other hand, exhibited an ultraviolet spectrum which varied over the pH ranges 1–4 and 8–11. The spectral change in alkali is consistent with the dissociation of the N-1 proton of the guanosine ring, whose pK_a is normally in that area. On the other hand, guanosines with 1-, O^6-, 3-, or 7-substitution would not be expected to show a dissociation in that range (Singer, 1972; P. Cope and R. Shapiro, unpublished data).

Further confirmation of the structures of Ia and IIa was provided by chemical reduction. When the adenosine and guanosine cross-links were heated with sodium borohydride in basic solution, they afforded mixtures of adenosine and N^6-methyladenosine and of guanosine and N^2-methylguanosine, respectively (Scheme I). The unsubstituted nucleoside was obtained in greater amount in each case, which

Chart I



suggests that hydrolysis accompanied the expected reduction.

Structures of Unsymmetrical Cross-Links IIIa–Va. Three products had been isolated from the RNA hydrolysates which, upon digestion with acid, gave two different nucleosides. They were also prepared from mixtures of the appropriate nucleosides with formaldehyde. Since these products were only available in submilligram quantities, less characterization was performed on them than was done for the symmetrical cross-links Ia and IIa. The unsymmetrical ones were assigned the amino-substituted structures IIIa, IVa, and Va (Chart I) on the following basis. (1) There was an analogy to Ia and IIa, whose structures were securely established. (2) Their ^1H NMR spectra were in accord with the proposed structures; they showed the requisite number of protons bound to carbon. Two-proton signals were observed at shifts close to one another, and to the values established for the methylene bridges in Ia and IIa. (3) The pK_a data confirmed one structure (IIIa) but required us to resort to an alternative technique, difference spectra, to resolve ambiguities for the two other structures. The dissociations of IVa and Va in the range pH 7–11 could be assigned to the 1 proton of guanosine or to the dissociation of an exocyclic amino proton from the cation of a 1-substituted adenosine in IVa or a 3-substituted cytidine in Va. The ultraviolet spectra of IIa, IVa, Va, and the model compound N^2-methylguanosine, taken at pH 11, were subtracted from those of the same compounds taken at pH 6.8. A series of four similar curves were obtained, with maxima near 246 and 290 nm and minima near 274 nm. The similarities of these curves (and their nonresemblance to curves derived from 1-methyladenosine or 3-methylcytidine) strongly implied that the same process was involved in the dissociation of IIa, IVa, and Va in alkali. This process would be the ionization of a proton from the 1 position of guanosine.

Deoxyribonucleoside Series of Products. A series of cross-links was isolated from formaldehyde-treated DNA which corresponded exactly, in order of elution from the reverse-phase high-pressure LC column and in nucleoside composition, to the cross-links isolated from RNA (an additional uncharacterized product was also isolated from DNA). The DNA-derived products had ultraviolet spectra, difference spectra, and pK_a behavior which were essentially identical with those of the corresponding cross-links isolated from RNA. On this basis, the structures Ib–Vb have been assigned, tentatively, to the DNA-derived cross-links.

The great similarity in the properties of these two series of products, despite the differing procedures used for hydrolysis

of the parent nucleic acids, makes it unlikely that the RNA products were the result of a Dimroth rearrangement or a similar process catalyzed by alkali.

Ultraviolet Spectra of Cross-Links. This datum has been included in Table I. In some cases, the spectra of the cross-links are similar to those of appropriate model compounds. For example, the spectrum of Ia at neutral pH closely resembles that of N^6-methyladenosine in shape, with a shift of 5 nm to longer wavelength. The spectrum of Ia in acid, however, is different in shape from that of its model and has its maximum 12 nm higher. It can be concluded that ultraviolet spectra, taken alone, are an unreliable indicator of structure for cross-links of this type. It is important to remember this because structural assignments made in the past, for example, the cross-link involving alkylation of guanosine by diepoxybutane (Brookes & Lawley, 1961), have relied on ultraviolet spectra to establish the position of substitution.

Discussion

Chemical agents that covalently cross-link nucleic acids are important as mutagens, carcinogens, antibiotics, antitumor drugs, and toxic substances (Dubelman & Shapiro, 1977). Determination of the chemical structures involved in the cross-links is an important first step in the path to understanding the mechanism of action of these agents. In view of the numerous technical difficulties involved in the isolation of cross-links and in the determination of structure of samples available in microgram quantities, it is not surprising that only a few of these structures have been fully characterized to date.

In isolating and characterizing the cross-links induced by formaldehyde from RNA and DNA, we have made extensive use of reverse-phase high-pressure LC. This technique has had a number of recent applications to biochemistry (Karger & Giesen, 1978). In the nucleic acid field, it has been particularly useful in separating adducts of nucleosides with polycyclic aromatic carcinogens from unmodified nucleosides (Jeffrey et al., 1976; Essigman et al., 1977). A number of studies have also been made on the use of the technique to fractionate normal components of nucleic acids (Hartwick & Brown, 1976). In general, nucleosides have eluted in order of increasing lipophilic character: cytidine, then uridine, guanosine, and adenosine. The same order has been followed in our cross-link separation, with the substances eluting as expected on the basis of their composition. The most striking effect is the much greater retardation of a cross-link such as guanosine–CH_2–guanosine, in comparison to the monomer, guanosine. This extra degree of lipophilicity cannot be attributed to the methylene group alone, since N^2-methylguanosine, in our system, elutes in a position between the monomer and dimer, but closer to the monomer. It is apparently the total area of lipophilic surface in a molecule that governs its affinity for the reverse-phase column, rather than the ratio of lipophilic and hydrophilic surfaces, which would be similar for monomer and dimer. In accord with this, it has been noted in separations of protected synthetic oligonucleotides on reverse-phase columns that oligomers of longer chain length elute later than shorter ones of the same composition (Fritz et al., 1978).

The procedure reported here has a number of advantages over those used in previous cross-link separations. An earlier worker had reported the isolation alkali and cross-linked nucleotides from RNA, after formaldehyde treatment (Feldman, 1967). The RNA was hydrolyzed to nucleotides in alkali and then fractionated by conventional anion-exchange chromatography. The method was lengthy and had the disadvantage that three nucleotide peaks were produced for each type of cross-link, representing 2′ plus 2′, 2′ plus 3′, and 3′ plus 3′ isomers. No products containing cytosine were reported. The peaks were eluted with formic acid, which could have led to some decomposition of the products.

Our new procedure is also preferable, for the case of the formaldehyde cross-links, to the one we developed to isolate cross-links induced in DNA by nitrous acid (Dubelman & Shapiro, 1977). The earlier method was more laborious and time consuming, because it involved hydrolysis to the nucleotide level, fractionation of nucleotides and oligonucleotides on DEAE-Sephadex, and desalting, prior to dephosphorylation to afford nucleosides. Some loss of product was the inevitable result of the handling of material involved in these three steps. In addition, further loss occurred when we attempted to apply the procedure to the formaldehyde cross-links, since some decomposition of the products took place under the conditions of the DEAE-Sephadex column (65 °C, 72 h). (The formaldehyde-derived cross-links were difficult to handle generally, because they were somewhat unstable in acid and Ia and IIa were quite insoluble in neutral solution.) The final step in the nitrous acid workup was a fractionation of the nucleoside products by anion-exchange high-pressure LC. It worked well in that case because the products were ionic at neutral pH. The same procedure failed to resolve the less polar formaldehyde cross-links. The reverse-phase column not only separated the formaldehyde products but also furnished them in a fully volatile medium. The anion-exchange separation, in the nitrous acid case, left traces of salt, which were a hindrance in later studies by mass spectrometry (Shapiro et al., 1977).

For the above reasons, we would recommend that our current procedure be tried first in future attempts at cross-link isolation, except in those cases where the end product of digestion is charged. Such cases would occur if enzymatic hydrolysis was blocked at the level of short oligonucleotides or if the nucleoside products were ionic at neutral pH. Charged products of these types would be hydrophilic in character and would not be retarded by the reverse-phase column. In such instances, the methods of our earlier nitrous acid study (Dubelman & Shapiro, 1977) may be tried.

Our results indicate that cytosine participates in the cross-linking of nucleic acids by formaldehyde. It had been claimed previously that only the purines were involved (Auerbach et al., 1977; Feldman, 1967) although an alternate deduction had been made at an earlier stage (Fraenkel-Conrat, 1954). We have also unequivocally demonstrated the presence of cross-links in formaldehyde-treated DNA. Their existence has been hitherto postulated on the basis of indirect evidence, such as the existence of a fraction of DNA, after formaldehyde treatment, that resists denaturation (Freifelder & Davison, 1963; Collins & Guild, 1968; Van Der Ebb et al., 1969).

It is of interest that RNA and DNA, which differ so much in their secondary structure, afford the same type of cross-links upon reaction with formaldehyde. In particular, in the case of DNA, some of the cross-links involve two sites, one of which is in the minor groove in native DNA (the guanine amino group) and another which is exposed in the major groove (the adenine or cytosine amino group). Since such products could not form readily in native DNA, it is likely that cross-link formation in DNA is preceded by other interactions with formaldehyde which lead to local denaturation of the DNA (McGhee & Von Hippel, 1977).

References

Auerbach, C., Moutschen-Dahmen, M., & Moutschen, J. (1977) *Mutat. Res. 39*, 317–362.

Brookes, P., N & Lawley, P. D. (1961) *J. Chem. Soc.*, 3923.
Collins, C. J., & Guild, W. R. (1968) *Biochim. Biophys. Acta 157*, 107–113.
Dubelman, S., & Shapiro, R. (1977) *Nucleic Acids Res. 4*, 1815–1827
Essigman, J. M., Croy, R. G., Nadzan, A. M., Busby, W. F., Jr., Reinhold, V. N., Buchi, G., & Wogan, G. N. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 1870–1874.
Feldman, M. Ya. (1967) *Biochim. Biophys. Acta 149*, 20–34.
Feldman, M. Ya. (1973) *Prog. Nucleic Acid Res. Mol. Biol. 13*, 1–47.
Fraenkel-Conrat, H. (1954) *Biochim. Biophys. Acta 15*, 307–309.
Freifelder, D., & Davison, P. F. (1963) *Biophys. J. 3*, 49–63.
Fritz, H. J., Belagaje, R., Brown, E. L., Fritz, R. N., Jones, R. A., Lees, R. G., & Khorana, H. J. (1978) *Biochemistry 17*, 1257–1267.
Hartwick, R. A., & Brown, P. R. (1976) *J. Chromatogr. 126*, 679–691.
Jeffrey, A. M., Blobstein, S. H., Weinstein, I. B., & Harvey, R. G. (1976) *Anal. Biochem. 73*, 378–385.
Karger, B. L., & Giesen, R. W. (1978) *Anal. Chem. 50*, 1048A–1059A.
McGhee, J. D., & Von Hippel, P. H. (1977) *Biochemistry 16*, 3276.
Nature (London) (1979) *281*, 625.
Shapiro, R., Dubelman, S., Feinberg, A. M., Crain, P. F., & McCloskey, J. A. (1977) *J. Am. Chem. Soc. 99*, 302–303.
Singer, B. (1972) *Biochemistry 11*, 3939.
Singer, B. (1975) in *CRC Handbook of Biochemistry and Molecular Biology* (Tanuan, G. D., Ed.) 3rd ed., Vol. I, p 409, CRC Press, Cleveland, OH.
Van Der Eb, A. J., Van Kesteren, L. W., & Van Bruggen, E. F. (1969) *Biochim. Biophys. Acta 182*, 530–541.

Fluorescence of Terbium Ion–Nucleic Acid Complexes: A Sensitive Specific Probe for Unpaired Residues in Nucleic Acids†

Michael D. Topal and Jacques R. Fresco*

ABSTRACT: The interaction of the lanthanide cation Tb^{3+} with the phosphate moieties of non-hydrogen-bonded residues of nucleic acids has been shown to result in substantial enhancement of the fluorescence of this cation. The excitation spectrum for this fluorescence is characteristic of the base moiety of the residue to which the Tb^{3+} is bound, while the emission spectrum is characteristic of the cation itself. The intensity of the fluorescence enhancement, however, is dependent upon the base of the ligand moiety, with G inducing the strongest enhancement, C and T rather less, and A very little. Base-paired residues of nucleic acids induce *no* such fluorescence enhancement, even though the cation is more tightly bound to double helical regions than to residues in single strands. The enhancement of Tb^{3+} fluorescence upon binding to non-hydrogen-bonded residues therefore provides a highly specific conformational probe for such residues. This probe has been exploited successfully for the purpose of analyzing the kinetics of reassociation of DNAs (C_0t analysis) and as a specific stain for single-strand DNA bands on polyacrylamide gels.

The rare-earth lanthanide Tb^{3+} (terbium ion) exhibits a low level of intrinsic fluorescence that is enhanced many-fold upon interaction with GMP (Formosa, 1973). This behavior is consistent with findings that there is general enhancement of the fluorescence of rare-earth cations by energy transfer from excited ligands chelating the ion (Weissman, 1942, 1950; Yuster & Weissman, 1949). The site of Tb^{3+} binding to nucleic acid monomers appears to be their negatively charged phosphate moiety, as has been shown for other lanthanide ions (Barry et al., 1971), which is to be expected in view of the generally observed strong affinity of trivalent cations for the nucleic acid backbone [e.g., Karpel et al. (1975, 1980)]. We have therefore explored the possibility that the concentration of charge in the nucleic acid backbone might result in greater enhancement of the fluorescent emission of Tb^{3+}. Since the nature of the base and its involvement in polynucleotide secondary structure should influence the transfer of energy to the bound Tb^{3+}, we also examined whether there is any secondary structural and/or base specificity to the enhanced Tb^{3+} fluorescence. In fact, all three effects were observed: Tb^{3+} bound to unpaired residues in nucleic acid chains exhibits much greater fluorescence than when bound to free nucleotides in solution; the amount of fluorescence enhancement depends on the identity of the residue; Tb^{3+} bound to *base-paired* residues shows no fluorescence enhancement. These properties enable the use of Tb^{3+} as a sensitive and specific stain for unpaired nucleic acid residues in solution and on gels. In this respect, the fluorescent behavior of Tb^{3+} on binding to nucleic acids complements that of ethidium bromide, which exhibits strong fluorescence on binding to double-stranded DNA but only weak fluorescence on binding to single strands (LePecq & Paoletti, 1967). A preliminary report of this work has been made (Topal & Fresco, 1979).

Materials and Methods

Materials. $TbCl_3$ (D. F. Goldsmith Chemical and Metal Corp.) was used without further purification. Cacodylic acid was recrystallized from hot 50% ethanol. Guanosine and nucleotides (Sigma Chemical Co.) were used without further purification.

† From the Department of Biochemical Sciences, Princeton University, Princeton, New Jersey 08544 (J.R.F.), and the Department of Pathology and the Cancer Research Center, University of North Carolina Medical School, Chapel Hill, North Carolina 27514 (M.D.T.). *Received February 14, 1980.* This work was supported by a National Cancer Institute Postdoctoral Fellowship to M.D.T. at Princeton University, Grant IN-15-S from the American Cancer Society and Grant P30 CA 16086 from the National Institutes of Health to M.D.T. at the University of North Carolina, and Grant GM 07654 from the National Institutes of Health and Grant GB 35595 from the National Science Foundation to J.R.F. at Princeton University. This is paper 17 in a series entitled "Polynucleotides", of which the previous paper is Karpel et al. (1980).

0006-2960/80/0419-5531$01.00/0