Supplementary Material for

“Sequence-dependent Formation of Intrastrand Crosslink Products from the UVB Irradiation of Duplex DNA Containing a 5-Bromo-2’-deoxyuridine or 5-Bromo-2’-deoxycytidine”

by Yu Zeng and Yinsheng Wang

(Nucleic Acids Res. 2006, Vol. 34)

Table of Contents

Pages S3-S4. Structure elucidation of G^U and A^U crosslinks.

Table S1. 1H NMR chemical shift data in p.p.m. for the crosslink products.

Table S2. The peak area values used for constructing the calibration curves of LC-MS/MS quantifications of the intrastrand crosslink products.

Scheme S1. Solution-phase synthesis of d(GBrU).

Scheme S2. Solution-phase synthesis of d(ABrU).

Scheme S3. Proposed mechanism for the formation of d(A[8-5]U) and d(A[2-5]U).

Figure S1. Negative-ion ESI-MS of d(BrUG) (a), d(BrU) (b), d(CG) (c), and d(BrU). The corresponding product-ion spectra of the [M – H]– ions are shown below the MS.

Figure S2. Negative-ion ESI-MS of d(ATGGCGBrUGCTAT) (a) and d(ATGGCABrUGCTAT) (c). Their corresponding product-ion spectra are shown in (b) and (d).

Figure S3. HPLC traces for the separation of the aerobic irradiation mixtures of d(GBrU) (a) and d(ABrU) (b).

Figure S4. Product-ion spectra of the [M + H]+ ions of d(U[5-N(2)]G): MS/MS (a), MS3 (b), MS4 (c); d(G[N(2)-5]U): MS/MS (d), MS3 (e), MS4 (f).

Figure S5. Product-ion spectra of the [M - H]– ions of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-8]G) (c); d(C[5-N(2)]G) (d); d(A[8-5]U) (e); d(A[2-5]U) (f); d(G[8-5]U) (g); d(G[N(2)-5]U) (h).
Figure S6. Product-ion spectra of the [M + H]⁺ ions of d(A[2-5]U) (a) and d(A[8-5]U) (b).

Figure S7. ¹H NMR spectra (500 MHz, D₂O) of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-N(2)]G) (c); d(G[8-5]U) (d); d(G[N(2)-5]U) (e); d(A[2-5]U) (f); d(A[8-5]U) (g).

Figure S8. ¹H NMR spectra (600 MHz, DMSO-d₆) of d(C) (a); d(G) (b); d(C[5-N(2)]G) (c).

Figure S9. Portions of 2-D NOE spectra of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-N(2)]G) (c); d(G[8-5]U) (d); d(A[2-5]U) (e); d(A[8-5]U) (f), showing the correlations between aromatic and H₁ protons.

Figure S10. UV absorption spectra of: d(U[5-8]G) and d(G[8-5]U) (a); d(U[5-N(2)]G) and d(G[N(2)-5]U) (b); d(C[5-N(2)]G) (c); d(A[2-5]U) and d(A[8-5]U) (d).

Figure S11. LC-MS/MS calibration curves for the analyses of the crosslink products.

Figure S12. Positive-ion product-ion spectra averaged from peaks in SICs shown in Figure 3b&d. The sample was the enzymatic digestion products of UVB-irradiated duplex d(ATGGCGBrUGCTAT)/d(ATAGCACGCCCAT).

Figure S13. Positive-ion product-ion spectra averaged from peaks observed in the SIC shown in Figure 5b. The sample was the enzymatic digestion products of UVB-irradiated duplex d(ATGGCGBrCGCTAT)/d(ATAGCGCGCCCAT).

Figure S14. LC-MS/MS analysis of the two standard d(A^U) crosslink products (a) and the enzymatic digestion products of the UVB-irradiated duplex ODN d(ATGGCABrUGCTAT)/d(ATAGCATGCCCAT) under aerobic conditions (b). Shown are the SICs for monitoring the m/z 540→246 transition.

Figure S15. Positive-ion product-ion spectra averaged from peaks observed in the SIC shown in Figure S14b. The sample was the enzymatic digestion products of UV-irradiated duplex d(ATGGCABrUGCTAT)/d(ATAGCATGCCCAT).

Figure S16. HPLC traces for the separation of the UVB-irradiation mixture of d(ATGGCGBrCGCTAT)/d(ATAGCGCGCCCAT): (a) the irradiation was done under aerobic conditions where the solution was exposed to, but not bubbled with air; (b) the irradiation was carried out under saturated oxygen conditions where the solution was constantly bubbled with O₂ during irradiation.
Structure Elucidation of crosslink products formed from the UVB light irradiation of d(G\textsuperscript{B}U).

We irradiated d(G\textsuperscript{B}U) with the Pyrex-filtered UV light and separated the irradiation mixture by HPLC (Figure S3a). Several lines of evidence supports that the 12.6- and 24.7-min fractions contain d(G[8-5]U) and d(G[N(2)-5]U), respectively. First, positive-ion ESI-MS results show that the two products have the same molecular weight (m/z 556) as the two d(U^G) crosslinks. In the product-ion spectra of the [M + H]\textsuperscript{+} ions of the crosslink products, we observed the formation of a dominant fragment ion of m/z 458, which is attributed to the loss of the familiar C\textsubscript{5}H\textsubscript{6}O\textsubscript{2} component (Figures 2d&S4d). Further fragmentation of the ion of m/z 458 shows that the ion of m/z-262 dominates the product-ion spectra (Figures 2e&S4e), and this ion is attributed to the protonated ion of the covalently bonded nucleobase portion.

Second, multi-stage MS experiments also support the structures of the two crosslinks. In this respect, we consider the crosslink products whose two nucleobases are covalently bonded in the same fashion as sequence isomers, i.e., d(U[5-8]G) and d(G[8-5]U) as well as d(U[5-N(2)]G) and d(G[N(2)-5]U). We reason that the collisional activation of the fragment ion corresponding to the crosslinked nucleobase moiety of sequence isomers will give the same product-ion spectra. In this context, we compared the product-ion spectra of the protonated ions of the crosslinked nucleobase component (that is, MS\textsuperscript{4}) of the 12.6- and the 24.7-min fractions with those of d(U[5-8]G) and d(U[5-N(2)]G). The similar MS\textsuperscript{4} observed for the 12.6-min fraction and d(U[5-8]G) (Figure 2c&f) support that the 12.6-min fraction is a sequence isomer of d(U[5-8]G), assigned as d(G[8-5]U). Along this line, the similarity between the MS\textsuperscript{4} for the 24.7-min fraction and that for d(U[5-N(2)]G) (Figure S4c&f) demonstrates that the 24.7-min fraction contains a sequence isomer of d(U[5-N(2)]G), and we assign the product as d(G[N(2)-5]U).

Third, \textsuperscript{1}H-NMR and 2-D NOE spectra furnish additional evidence for our structure assignments. We observed a singlet aromatic proton resonance (δ 7.92 ppm) in the \textsuperscript{1}H-NMR spectrum of d(G[8-5]U) and a strong cross peak between the H6 of uracil and the H\textsubscript{1} of the 3’ nucleoside in the 2-D NOE spectrum, which is consistent with the structure of the product (Figure S7d&S9d). \textsuperscript{1}H-NMR spectrum of d(G[N(2)-5]U), on the other hand, showed two singlet aromatic proton resonances, which were assigned as the H(6) of uracil (δ 8.35 ppm) and the H(8) of guanine (δ 8.13 ppm, Figure S7e & Table S1). Last, but not the least, the UV absorption spectra of the d(G[8-5]U) and d(G[N(2)-5]U) are distinctive from each other and very similar to those of d(U[5-8]G) and d(U[5-N(2)]G), respectively (Figure S10a-b), supporting again our structure assignments.

Crosslink products formed from d(A\textsuperscript{B}U)

To compare the crosslinking chemistry of duplex ODNs containing a 5’-G\textsuperscript{B}U-3’ or 5’-A\textsuperscript{B}U-3’ sequence motif, we first investigated the photochemistry of d(A\textsuperscript{B}U).

It turned out that two intranstrand crosslink products, i.e., d(A[8-5]U) and d(A[2-5]U) (structures shown in Scheme 1), could be induced from the photoirradiation (Figure S3b gives the HPLC trace for the separation of the irradiation mixture). First, positive-ion ESI-MS showed that the measured masses of these two products are consistent with their calculated masses. In addition, the product-ion spectra of the [M + H]\textsuperscript{+} ions (m/z 540) showed the formation of the ion of m/z 246 (Figure S6a-b), which is attributed to the loss of the 2-deoxyribose-phosphate backbone and strongly supports the covalent linkage between adenine and uracil.
Furthermore, we observed that the \([M - H_2O]\) fragment \((m/z 522)\) is significantly more abundant in the product-ion spectrum of the \([M + H]^+\) ion of \(d(A[8-5]U)\) than that in the corresponding MS/MS of \(d(A[2-5]U)\) (Figure S6). This result is in accordance with what we found for the two analogous crosslink products formed at AC site (Ref. S1). In this regard, the \([M - H_2O]\) fragment dominates the product-ion spectrum of the \([M + H]^+\) ion of \(d(A[8-5]C)\), whereas the \([M - H_2O]\) ion is less abundant than the protonated ion of the crosslinked nucleobase portion in the product-ion spectrum of the \([M + H]^+\) ion of \(d(A[2-5]C)\) (Ref. S1).

\(^1\)H-NMR and 2-D NOE spectra offer additional evidence about the structures of the crosslink products. There are two singlet aromatic proton resonances in the \(^1\)H-NMR spectra of both products (Figure S7f-g), and we observed a strong correlation peak between the H(6) of uracil and the H\(_1'\) of the 3’ nucleoside in the 2-D NOE spectra of both products (Figure S9e-f). The 2-D NOE spectrum of \(d(A[8-5]U)\) also exhibits a strong correlation peak between the H(2) of adenine and the H\(_5'\) of the 5’ nucleoside (Figure S9f). Furthermore, we observed marked downfield displacement of the H\(_2\) proton of the 2’-deoxyuridine in \(d(A[8-5]U)\), but not in \(d(A[2-5]U)\) (Figure S7f-g, Table S1). This observation is consistent with what we found for the two corresponding \(d(A^C)\) crosslinks, where we found similar significant downfield shift of H\(_2\) proton of 2’-deoxycytidine in \(d(A[8-5]C)\), but not in \(d(A[2-5]C)\) (Ref. S1).

The third piece of evidence lies in the UV absorption spectra of crosslink products (Figure S10d). The absorption maximum of \(d(A[2-5]U)\) occurs at a longer wavelength than that of \(d(A[8-5]U)\). In addition, \(d(A[2-5]U)\) exhibits a broader absorption peak than \(d(A[8-5]U)\). Similar phenomena have been observed in the UV absorption spectra of \(d(A[2-5]C)\) and \(d(A[8-5]C)\) (Ref. S1).

Formation of crosslink products from the UVB irradiation of \(d(ATGGCA^BrUGCTAT)/d(ATAGCATGCCAT)\)

After having established the structures of the two \(d(A^U)\) products, we next investigated whether these products can form in duplex DNA. To this end, we irradiated duplex \(3\) (Table 1), digested the mixture with the same four enzymes, and analyzed the digestion mixture by LC-MS/MS. The SIC for the \(m/z 540\to 246\) transition from the analysis of two standard \(d(A^U)\) showed two peaks at 17.8 and 29.3 min, corresponding to \(d(A[2-5]U)\) and \(d(A[8-5]U)\), respectively (Figure S14a). The SIC for the analysis of the digestion mixture showed two fractions with similar elution time (i.e., 18.4 and 29.4 min) as the two standards (Figure S14b). Moreover, the product-ion spectra of the products agree well with those of standards (Figure S6&S14). These results allow us to conclude that both \(A[2-5]U\) and \(A[8-5]U\) are generated from the UV irradiation of duplex \(3\).

Reference:

S1. Hong, H. and Wang, Y. (2005) Formation of intrastrand cross-link products between cytosine and adenine from UV irradiation of \(d(BrCA)\) and duplex DNA containing a 5-bromocytosine. J. Am. Chem. Soc., 127, 13969-13977.
Table S1. $^1$H NMR chemical shift data in p.p.m. for the crosslink products (25°C, 500 MHz, D$_2$O).

|                        | H$_1$ | H$_2$ | H$_3$ | H$_4$ | H$_5$ | H$_5''$ | H$_6$ | H$_8$ | H$_2$ |
|------------------------|-------|-------|-------|-------|-------|---------|-------|-------|-------|
| d(U[5-N(2)]G)(Ura)     | 6.49  | 2.91  | 2.85  | 4.78  | 4.51  | 3.79    | 3.79  | 8.29  | --    |
| d(U[5-N(2)]G)(Gua)     | 6.28  | 2.99  | 2.60  | 4.43  | 4.35  | 4.25    | 3.88  | --    | 8.13  |
| d(U[5-8]G)(Ura)        | 6.00  | 3.72  | 2.25  | 4.65  | 4.35  | 4.21    | 4.10  | 8.15  | --    |
| d(U[5-8]G)(Gua)        | 6.51  | 2.80  | 2.80  | 4.90  | 4.00  | 3.82    | 3.65  | --    | --    |
| d(C[5-N(2)]G)(Cyt)     | 6.41  | 2.77  | 2.68  | 4.64  | 4.02  | 3.80    | 3.80  | 8.06  | --    |
| d(C[5-N(2)]G)(Gua)     | 6.23  | 2.96  | 2.54  | 4.40  | 4.28  | 4.15    | 3.93  | --    | 8.11  |
| d(G[N(2)-5]U)(Ura)     | 6.69  | 2.78  | 2.75  | 4.85  | 3.98  | 3.62    | 3.60  | 8.35  | --    |
| d(G[N(2)-5]U)(Gua)     | 6.57  | 2.90  | 2.46  | 4.47  | 4.19  | 4.09    | 4.06  | --    | 8.13  |
| d(G[8-5]U)(Ura)        | 6.21  | 3.95  | 2.68  | 4.97  | 4.38  | 4.20    | 3.67  | 7.92  | --    |
| d(G[8-5]U)(Gua)        | 6.50  | 2.98  | 2.56  | 4.53  | 4.06  | 4.26    | 3.80  | --    | --    |
| d(A[8-5]U)(Ura)        | 6.48  | 3.87  | 2.80  | 4.64  | 4.28  | 3.51    | 3.48  | 8.40  | --    |
| d(A[8-5]U)(Ade)        | 6.34  | 2.97  | 2.63  | 5.09  | 4.44  | 4.20    | 3.84  | --    | 9.24  |
| d(A[2-5]U)(Ura)        | 6.38  | 2.59  | 2.48  | 4.54  | 4.17  | 4.09    | 3.68  | 7.97  | --    |
| d(A[2-5]U)(Ade)        | 6.54  | 3.05  | 2.78  | 5.05  | 4.41  | 3.97    | 4.20  | --    | 8.37  |
Table S2. The peak area values used for constructing the calibration curves of LC-MS/MS quantifications of the eight crosslink products. The peak area values correspond to the peaks observed in selected-ion chromatogram (SICs) for monitoring the transitions indicated in the first row. The amount of standard crosslink products injected are listed in the first column.

| pmol  | Peak area ($\times 10^4$) | \(m/z\) 556–458 | \(m/z\) 556–538 | pmol  | Peak area ($\times 10^4$) | \(m/z\) 556–538 |
|-------|-------------------------|-----------------|-----------------|-------|-------------------------|-----------------|
| 50    | 90.2                    | 38.7            | 25              | 200   | 184                     | 98.1            | 100             | 130             |
| 200   | 871                     | 553             | 200             | 25    | 25.6                     |                 |
| 800   | 3.23E3                  | 1.65E3          | 400             | 792   |                          |                 |
| 1600  | 9.86E3                  | 4.57E3          | 800             | 2.04E3|                          |                 |

| pmol  | Peak area ($\times 10^4$) | \(m/z\) 555–261, 457 |
|-------|-------------------------|------------------------|
| 50    | 251                     | 125                    | 250                  |
| 125   | 942                     | 672                    | 1.29E3               |
| 250   | 1.23E3                  | 715                    | 3.54E3               |
| 500   | 2.98E3                  | 1.53E3                 | 5.34E3               |
| 1000  | 5.50E3                  | 2.77E3                 | 5.34E3               |

| pmol  | Peak area ($\times 10^4$) | \(m/z\) 540–246, 442 | pmol  | Peak area ($\times 10^4$) | \(m/z\) 540–246, 442 |
|-------|-------------------------|-----------------------|-------|-------------------------|-----------------------|
| 46    | 26.8                    | 50                    | 30.8             |
| 184   | 125                     | 200                   | 181             |
| 460   | 383                     | 400                   | 423             |
| 1380  | 1.13E3                  | 800                   | 1.08E3           |
| 1840  | 1.30E3                  | 1600                  | 1.19E3           |
Scheme S1. Solution-phase synthesis of d(GBrU).

Reagents: a: (i) 4,5-Dicyanoimidazole/DMF, (ii) t-BuOOH in nonane. b: 80% CH₃COOH. c: NH₃/MeOH.
Scheme S2. Solution-phase synthesis of d(A\textsuperscript{Br}U).

Reagents: a: (i) 4,5-Dicyanoimidazole/DMF, (ii) t-BuOOH in nonane. b: 80% CH\textsubscript{3}COOH. c: NH\textsubscript{3}/MeOH.
**Scheme S3.** Proposed mechanism for the formation of d(A[8-5]U) and d(A[2-5]U).

![Scheme S3](image-url)
Figure S1. Negative-ion ESI-MS of d\(^{\text{Br}}\)UG (a), d(G\(^{\text{Br}}\))U (b), d(B\(^{\text{Br}}\))CG (c), and d(A\(^{\text{Br}}\))U (d). The product-ion spectra of the [M – H]\(^-\) ions are shown underneath the MS.
Figure S2. Negative-ion ESI-MS of d(ATGGCGBrUGCTAT) (a) and d(ATGGCABrUGCTAT) (c). Their corresponding product-ion spectra are shown in (b) and (d).
Figure S3. HPLC traces for the separation of the aerobic irradiation mixtures of d(G^{Br}U) (a) and d(A^{Br}U) (b).
Figure S4. Product-ion spectra of the [M + H]⁺ ions of d(U[5-N(2)]G): MS/MS (a), MS³ (b), MS⁴ (c); d(G[N(2)-5]U): MS/MS (d), MS³ (e), MS⁴ (f). The relative collisional energies were 30%.
Figure S5. Product-ion spectra of the [M - H]⁻ ions of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-8]G) (c); d(C[5-N(2)]G) (d); d(A[8-5]U) (e); d(A[2-5]U) (f); d(G[8-5]U) (g); d(G[N(2)-5]U) (h). The relative collisional energies were 30%.
Figure S6. Product-ion spectra of the [M + H]+ ions of d(A[2-5]U) (a) and d(A[8-5]U) (b). The relative collisional energies were 30%.
Figure S7. $^1$H NMR spectra (500 MHz, D$_2$O) of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-N(2)]G) (c); d(G[8-5]U) (d); d(G[N(2)-5]U) (e); d(A[2-5]U) (f); d(A[8-5]U) (g).
Figure S7. $^1$H NMR spectra (500 MHz, D$_2$O) of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-N(2)]G) (c); d(G[8-5]U) (d); d(G[N(2)-5]U) (e); d(A[2-5]U) (f); d(A[8-5]U) (g).
Figure S8. $^1$H NMR spectra (600 MHz, DMSO-$_d_6$) of dC (a); dG (b); d(C[5-N(2)]G) (c).
Figure S9. Portions of 2-D NOE spectra of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-N(2)]G) (c); d(G[8-5]U) (d); d(A[2-5]U) (e); d(A[8-5]U) (f), showing the correlations between aromatic and H$_1'$ protons.
Figure S9. Portions of 2-D NOE spectra of d(U[5-8]G) (a); d(U[5-N(2)]G) (b); d(C[5-N(2)]G) (c); d(G[8-5]U) (d); d(A[2-5]U) (e); d(A[8-5]U) (f), showing the correlations between aromatic and $H_1'$ protons.
Figure S10. UV absorption spectra of: $d(U[5-8]G)$ and $d(G[8-5]U)$ (a); $d(U[5-N(2)]G)$ and $d(G[N(2)-5]U)$ (b); $d(C[5-N(2)]G)$ (c); $d(A[2-5]U)$ and $d(A[8-5]U)$ (d).
Figure S11. LC-MS/MS calibration curves for the analyses of the eight crosslink products.
Figure S12. Positive-ion product-ion spectra averaged from peaks shown up in SICs shown in Figure 3b&d. The sample was the enzymatic digestion products of UVB-irradiated duplex d(ATGGCG<sup>B</sup>UGCTAT)/d(ATAGCACGCCAT).
Figure S13. Positive-ion product-ions spectra averaged from peaks observed in the SIC shown in Figure 5b. The sample was the enzymatic digestion products of UVB-irradiated duplex d(ATGGCG^BrCGCTAT)/d(ATAGCGCGCCAT).
Figure S14. LC-MS/MS analysis of the two standard d(A^U) crosslink products (a) and the enzymatic digestion products of the UVB-irradiated duplex ODN d(ATGGCA^BrUGCTAT)/d(ATAGCATGCCAT) under aerobic conditions (b). Shown are the SICs for monitoring the m/z 540→246 transition.
Figure S15. Positive-ion product-ion spectra averaged from peaks observed in the SIC shown in Figure S14b. The sample was the enzymatic digestion products of UVB-irradiated duplex d(ATGGCA^BrUGCTAT)/d(ATAGCATGCCAT).
Figure S16. HPLC traces for the separation of the UVB-irradiation mixture of d(ATGGCG\textsuperscript{Br}CGCTAT)/d(ATAGCGCGCCAT): (a) the irradiation was done under aerobic conditions where the solution was exposed to, but not bubbled with air; (b) the irradiation was carried out under saturated oxygen conditions where the solution was constantly bubbled with O\textsubscript{2} during irradiation.