BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 6335–6347 6335
DOI: 10.1021/bi900531z

Unusual DNA−DNA Cross-Links between a Photolyase Deoxyribozyme, UV1C, and Its Bound Oligonucleotide Substrate†

Gurpreet S. Sekhon and Dipankar Sen*

Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada

Received March 28, 2009; Revised Manuscript Received May 23, 2009

ABSTRACT: UV1C is a photolyase deoxyribozyme that repairs thymine dimers in a DNA oligonucleotide substrate. We report that treatment with iodine generates specific DNA−DNA cross-links between UV1C and a bound substrate analogue, LDPs, in which a single phosphate at the photoreactivation site has been replaced with a phosphorothioate. Although iodine has been reported to generate lysine−cysteine cross-links within a protein, the formation of DNA−DNA cross-links is both unexpected and novel. We have used different mapping procedures to identify a number of bases located in loops of the G-quadruplex fold of UV1C as the sites for cross-linking with LDPs. Mutation of one cross-linking adenine, in particular, leads to a major reduction in UVIC's catalytic activity. A map of these contact cross-linking sites enables us to refine an earlier structural−topological model for the folded UV1C·LDPs complex. The surprising facility with which these novel contact cross-links can be generated between a nucleic acid enzyme and its substrate's reaction site opens up a powerful new approach to mapping the active sites of different ribozymes and deoxyribozymes as well as enabling the dissection of key contacts within RNA−protein complexes.

The UV1C deoxyribozyme is a single-stranded DNA that was isolated from an in vitro selection experiment designed to yield DNA enzymes (deoxyribozymes or DNAzymes) capable of repairing *cis*,*syn*-cyclobutane thymine dimer (CPD)[1] lesions within a DNA substrate. UV1C, which most optimally uses >300 nm ultraviolet light for its catalysis (*1*), provided the first evidence that nucleic acid enzymes are capable of carrying out photocatalytic processes. Other than light, UV1C does not require any extraneous cofactor for its catalysis of photoreactivation (*1*, *2*); however, another deoxyribozyme, Sero1C, obtained from the same in vitro selection experiment as UV1C, requires serotonin for its catalysis of thymine dimer photoreactivation (*3*). Both of these DNAzymes were selected to repair a special thymine dimer-containing single-stranded DNA substrate (TDP), which differed from a standard thymine dimer-containing DNA in lacking the phosphodiester linking the two thymidine participants in the dimer (*1*) (the sequences of TDP and all other oligonucleotides used in this study are given in Table 1). Extensive mechanistic and structural studies conducted on UV1C (*1*, *2*) revealed that when bound to its original substrate, TDP (or to a substrate analogue, LDP, a continuous piece of DNA of the same nucleotide sequence as TDP, but in which the dimerized thymines of TDP are present in undimerized form), UV1C folds to form a guanine quadruplex (G-quadruplex), which can act as a light-harvesting antenna capable of absorbing light in the 300−310 nm wavelength range. It is within this wavelength range where UV1C is most optimally active (*1*). A hypothesis to explain UV1C's catalytic activity posited that one or more photoexcited guanines within UV1C's G-quadruplex are located sufficiently proximal to the substrate's thymine dimer to efficiently transfer an electron to the dimer. This could set up chemical reactivation steps similar to those catalyzed by naturally occurring photolyase protein enzymes (*4*). To test this hypothesis, two variants of the LDP substrate analogue in which the dimerizing thymines were replaced, one at a time, with 5-iodoracil were synthesized. These iodoracil-substituted LDP oligomers, capable of forming contact cross-links with proximal DNA residues upon UV irradiation, were found to cross-link with a number of the UV1C guanines involved in forming its G-quadruplex (*2*), supporting the hypothesis about the UV1C catalytic strategy. A map of the above contact cross-links, along with truncation and deletion experiments on UV1C and the substrate, enabled the proposal of a topological model for the UV1C·LDP/TDP complex (Figure 1). The G-quadruplex was proposed to consist of two stacked G-quartets, with the participating guanines linked by two edgewise loops (L1 and L3) and a single double-chain reversal loop (L2). The guanines colored

†This work was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Grant RGPIN/105785. D.S. is an Associate of the Canadian Institutes for Advanced Research (CIFAR).

*To whom correspondence should be addressed. E-mail: sen@sfu.ca. Phone: (778) 782-4386. Fax: (778) 782-5583.

[1]Abbreviations: Pu, purine base; Ps, phosphorothioate; DNAzyme, deoxyribonucleotide enzyme; dNTP, deoxynucleotide triphosphate; ddATP, dideoxyadenosine triphosphate; PAGE, polyacrylamide gel electrophoresis; ELISA, enzyme-linked immunosorbent assay; EDTA, ethylenediaminetetraacetic acid; TBE, Tris Borate EDTA buffer; TE, Tris EDTA buffer; NEB, New England Biolabs; CPD, cyclobutane pyrimidine dimer.

Published on Web 06/10/2009

Table 1: Complete List of DNA Oligonucleotides Used in This Study

Name	Sequence [a]	Length
T1	AGG ATC TAC ATG TAT	15
T1-ATGT-TAAA	AGG ATC TAC TAA AAT	15
T2	TGT GTG CGT ACG AGT ATA TG	20
TDP	AGG ATC TAC ATG TAT^TGT GTG CGT ACG AGT ATA TG	35
TDP-ATGT-TAAA	AGG ATC TAC TAA AAT^TGT GTG CGT ACG AGT ATA TG	35
LDPs	AGG ATC TAC ATG TAT$_{Ps}$TGT GTG CGT ACG AGT ATA TG	35
LDPs-ATGT-TAAA	AGG ATC TAC TAA AAT$_{Ps}$TGT GTG CGT ACG AGT ATA TG	35
Sero1C	TAA GTC GGA TGG AGG GAT CCG TTG GCA CAT GTA GTC ACG T	40
UV1C	GGA GAA CGC GAG GCA AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
UV1C-G13I	GGA GAA CGC GAG ICA AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
UV1C-G13Pu	GGA GAA CGC GAG PCA AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
UV1C-A15Pu	GGA GAA CGC GAG GCP AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
UV1C-A15T	GGA GAA CGC GAG GCT AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
UV1C-C14T	GGA GAA CGC GAG GTA AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
UV1C-AAAT-TTTA	GGA GAA CGC GAG GCA AGG CTG GGA GTT TAG TGG ATC ACG ATT	42
UV1C-CA-TT	GGA GAA CGC GAG GTT AGG CTG GGA GAA ATG TGG ATC ACG ATT	42
Neg-Enz	TGG AGG TGA ACC AGC CCT GCA CCC TAT AGT GAG TCG TAT TAC AT	44
Primer 1	AAT CGT GAT CCA CAT TTC TC	20
Primer 2	AAT CGT GAT CCA CAT	15
Splint 1	ACG CAC ACA ATA CAT GTA	18
Splint 2	ACG CAC ACA ATT TTA GTA	18

[a] Thymine dimers linked via cyclobutane bonds without phosphodiester linkage represented by a carrot. Mutated bases colored red. Phosphorothioate linkages are represented by green Ps. I represents the hypoxanthine base. P represents the purine base.

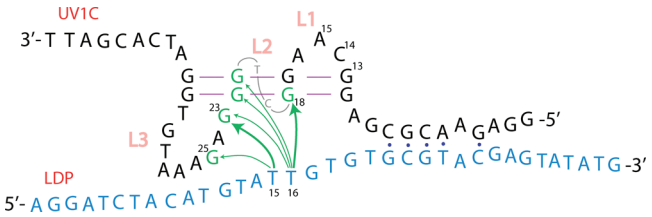


FIGURE 1: Secondary structure folding model for the UV1C·LDP DNAzyme·substrate complex. The green arrows show iodoracil-mediated photo-cross-links formed between thymines 15 and 16 (substituted individually with iodoracil) in LDP and individual bases in UV1C. Major and minor cross-links formed are represented by thick and thin arrows, respectively.

blue in Figure 1 form contact cross-links with bases T15 and T16 of LDP (the dimerizing thymines of TDP) when these latter bases are individually substituted with 5-iodoracil (*2*).

An important question that arose from the studies described above was whether UV1C is also capable of photorepairing thymine dimers within a more physiological substrate, namely, a continuous DNA single strand (the LDP oligonucleotide, but including a thymine dimer). Given the difficulty of both generating and monitoring the repair of a single thymine dimer within a 35-nucleotide DNA oligonucleotide, we explored a number of experimental schemes for monitoring such repair. One concept involved the replacement of the phosphodiester linking the dimerized thymidines with a phosphorothioate moiety, to take advantage of a phosphorothioate's susceptibility to reaction with iodoethanol, which leads to DNA strand cleavage (*5*, *6*). An analogous protocol for RNA is currently in wide use for mapping and footprinting transcripts (*7*), except that the RNA phosphorothioate is reacted with iodoethane or, more commonly, elemental iodine (*5*−*7*).

The DNA substrate analogue, LDPs, which we tested for cleavage, was LDP with undimerized thymines but included a single phosphorothioate residue within it (Table 1). Treatment of the UV1C·LDPs complex with iodoethanol, iodoethane, and iodine controls, followed by examination of the results using denaturing gel electrophoresis, revealed an unexpected low-gel mobility species that appeared repeatedly and reproducibly only in the iodine-treated UV1C·LDPs mixtures.

In this paper, we report that these high-molecular weight species are iodine-generated intermolecular DNA−DNA cross-linked products formed between the DNAzyme and the LDPs substrate. We take advantage of these unexpected cross-links to map nucleotide positions within the UV1C photolyase DNAzyme that localize proximally to the dimer-forming regions of the LDPs substrate analogue. Having done so, we are able to postulate a refined topological model for the UV1C·LDPs/LDP/TDP complex. More generally, we introduce the phosphorothioate cross-linking technique as a powerful tool for structure probing studies with biopolymer complexes.

EXPERIMENTAL PROCEDURES

Oligonucleotides. All DNA oligonucleotides were synthesized by the University Core DNA Services of the University of Calgary (Calgary, AB). T4 polynucleotide kinase was purchased from Invitrogen, terminal transferase from New England Biolabs, and Klenow DNA polymerase from Roche. Deoxy- and dideoxynucleoside triphosphates were purchased from Fermentas. [γ-^{32}P]ATP and [α-^{32}P]ddATP radionucleotides were obtained from PerkinElmer Life and Analytical Sciences. Table 1 lists all the oligonucleotides used in this study. All oligonucleotides were size-purified by denaturing polyacrylamide gel electrophoresis (PAGE) prior to their use.

Electrophoresis and Sequencing Procedures. Denaturing polyacrylamide gel electrophoresis was conducted in 12% gels containing a 19:1 bisacrylamide/acrylamide mixture and 50 mM TBE running buffer. Samples were heated in denaturing loading dye at 94 °C for 2 min prior to electrophoresis. Isolated bands were eluted using crush and soak in TE buffer overnight at 4 °C.

The 5′ radiolabeling was performed in 25 μL reaction mixtures with 4 pmol of DNA, 5 μL of 5× forward reaction buffer (Invitrogen), 5 μCi of [γ-^{32}P]ATP, and 10 units of T4 polynucleotide kinase. The 3′ labeling was conducted in 10 μL reaction mixtures containing 4 pmol of DNA, 1 μL of 10× NE Buffer 4 (NEB), 1 μL of 10× $CoCl_2$ (NEB), 20 units of terminal

transferase, and 2 μCi of [α-^{32}P]ddATP. Both reaction mixtures were incubated at 37 °C for 30 min followed by ethanol precipitation and PAGE purification.

Primer extension was completed in 10 μL reaction mixtures containing 1 μM template DNA, 1 μL of 10× Filling Buffer (Roche), and 1 μL of labeled primer. Reaction mixtures were heated to 94 °C and cooled to room temperature. Each dNTP at 1 μM and 1 unit of Klenow DNA polymerase were then added to initiate the reaction, which was allowed to proceed at 37 °C for 10 min. Denaturing loading buffer (6 μL) was added, and samples were purified via PAGE.

Sequencing Ladders. Dideoxy sequencing was completed using a modified version of the original protocol (*8*). In 25 μL, reaction ratios of 6:1 ddCTP:dCTP, 10:1 ddTTP:dTTP, 10:1 ddATP:dATP, and 5:1 ddGTP:dGTP were used. Reaction mixtures were incubated for 2 min at 37 °C and reactions stopped with denaturing loading dye.

Guanine-specific chemical sequencing ladders were created using the procedures developed by Maxam and Gilbert (*9*); 32 μL reaction mixtures contained 3 μM tRNA, 63 mM cacodylate buffer (pH 7.5), 1 μM ^{32}P-labeled DNA, and 0.6% DMS (fresh) to initiate the reaction. Methylation was stopped after 20 min with 8 μL of DMS stop solution (1.5 M NaOAc and 1 M β-mercaptoethanol).

Thymine-specific chemical cleavage was completed in a 40 μL reaction mixture with 10 μM tRNA, 100 mM Tris (pH 8.0), 1 μL of ^{32}P-labeled DNA, and 1 mM $KMnO_4$ to initiate the reaction (*10*). The $KMnO_4$ was quenched after 2 min with 1% (v/v) allyl alcohol.

Cytosine-specific chemical sequencing was completed in a 30 μL reaction mixture containing 3 μM tRNA, 1 μL of ^{32}P-labeled DNA, and 2.7 M $NH_2OH \cdot HCl$ (pH 6.0) to initiate the reaction (*10*). The reaction was stopped after 40 min by ethanol precipitation. All chemical sequencing reaction mixtures were ethanol precipitated ($^1/_{10}$ volume of 3 M NaOAc and 2.5× anhydrous ethanol) followed by piperidine cleavage.

Cross-Linking Procedure and Analysis. Phosphorothioate cross-linking was conducted with 25 μL reaction mixtures containing each of the phosphorothioate-containing substrates at 1 μM as well as the unmodified or mutated DNAzymes (or other control DNAs), in 20 mM Tris (pH 7.4) and 200 mM NaCl. These solutions were heated to 94 °C and then cooled to room temperature for folding of the DNA. Cross-linking was initiated by adding a final I_2 concentration of 50 μM (dissolved in EtOH), which made up the final reaction volume of 25 μL. Samples were quickly vortexed and spun down to allow reactions to proceed. After 2 min, reactions were stopped via ethanol precipitation or direct addition of denaturing dye. All experiments were performed under natural benchtop conditions (i.e., ambient light and room temperature not varying beyond 21−22 °C) unless specific reactions were performed under conditions of "no light" to show the negligible effects of ambient light on cross-linking.

Fe-EDTA hydroxyl radical cleavage was conducted in 25 μL reaction mixtures containing 2.5 μL of a solution containing 1 mM $Fe(NH_4)_2(SO_4)_2$ and 2 mM EDTA and 1.0 μL of a solution containing 50 mM ascorbic acid. Following mixing, the reaction was initiated by the addition of 2.5 μL of a 3% H_2O_2 stock (final H_2O_2 concentration of 0.3%). The reaction was stopped after 10 min via addition of 1 μL of 25 mM thiourea (final thiourea concentration of 1 mM). All reagents were freshly made.

Both cross-linked and non-cross-linked DNA products were treated with piperidine following the same procedures used for generating Maxam−Gilbert sequencing ladders. Samples were incubated at 90 °C for 30 min in 100 μL of a 10% (v/v) piperidine solution in water. The treated DNA solutions were then vacuumed to dryness, redissolved with 100 μL of ddH_2O, and vacuumed again.

Autoradiography was conducted on Fuji phosphor screens and scanned in a Molecular Dynamics Typhoon 9410 Variable Mode Imager. Quantitative analysis was performed using Molecular Dynamics ImageQuant.

Single-Turnover TDP Photoreversal. Thymine dimer substrates were synthesized following methods previously published (*1*). Reaction mixtures were prepared with a final volume of 60 μL consisting of 2.1 μM enzyme, 20 nM ^{32}P-5′-end-labeled substrate, 20 mM Tris (pH 7.4), and 200 mM NaCl. Enzymes were folded by being heated and cooled from 94 to 22 °C in the heating block of a thermal cycler. Samples were irradiated in a glycogen-blocked ELISA plate on a Fotodyne transilluminator (3.4×10^{-9} einsteins min^{-1} illumination). Aliquots (2 μL) were removed at designated time points and added to equal volumes of denaturing dye to be run on denaturing polyacrylamide gels (described above). The resulting gels were imaged by autoradiography and quantitated as described above.

RESULTS

Iodine-Mediated UV1C·LDPs Cross-Linking. Figure 2A shows the results of treatment of UV1C, bound to 5′-^{32}P-labeled LDPs, with iodine, iodoethane, and iodoethanol. One can see that only incubation in iodine (in as little as 30 μM) generates two closely spaced, retarded electrophoretic mobility bands (labeled, provisionally, as "cross-linked products"), whereas treatment with concentrations even as high as 1 M iodoethane or 1 M iodoethanol does not generate comparable bands. It is also notable that in parallel with the formation of these products, a significantly high level of cleavage of the LDPs oligonucleotide can be observed at the phosphorothioate site, to give cleaved products.

Figure 2B shows the specificity of cross-linking between the UV1C DNAzyme and LDPs. The 5′-^{32}P-labeled LDPs was incubated with stoichiometric amounts of a cDNA splint; with the deoxyribozyme Sero1C, which binds to LDPs; and with Neg-Enz, a DNA oligonucleotide that is the same length as UV1C but which does not interact with LDPs. The different mixtures were incubated with and without iodine (also, either illuminated or not illuminated with light). These represent experiments that were conducted in the dark; however, the presence of light makes no difference to the results (data not shown). It can be seen that the retarded gel mobility (cross-linked) bands appear only when UV1C and LDPs are incubated together in the presence of I_2 (Figure 2B, right); no comparable bands are seen in the absence of iodine treatment (Figure 2B, left). Also their formation is specific for UV1C, as no slow mobility products can be seen in the reaction mixtures containing the complementary splint, Neg-Enz, or, indeed, in the absence of any extraneous DNA. A very minor high-molecular weight band is seen to form when LDPs forms a complex with Sero1C. In this paper, we report exclusively on the properties of the high-yield products formed by the UV1C·LDPs complex.

Repetition of the experiments described above using the oligonucleotide LDP instead of LDPs (LDP lacks the single phosphorothioate functionality present within LDPs) fails to

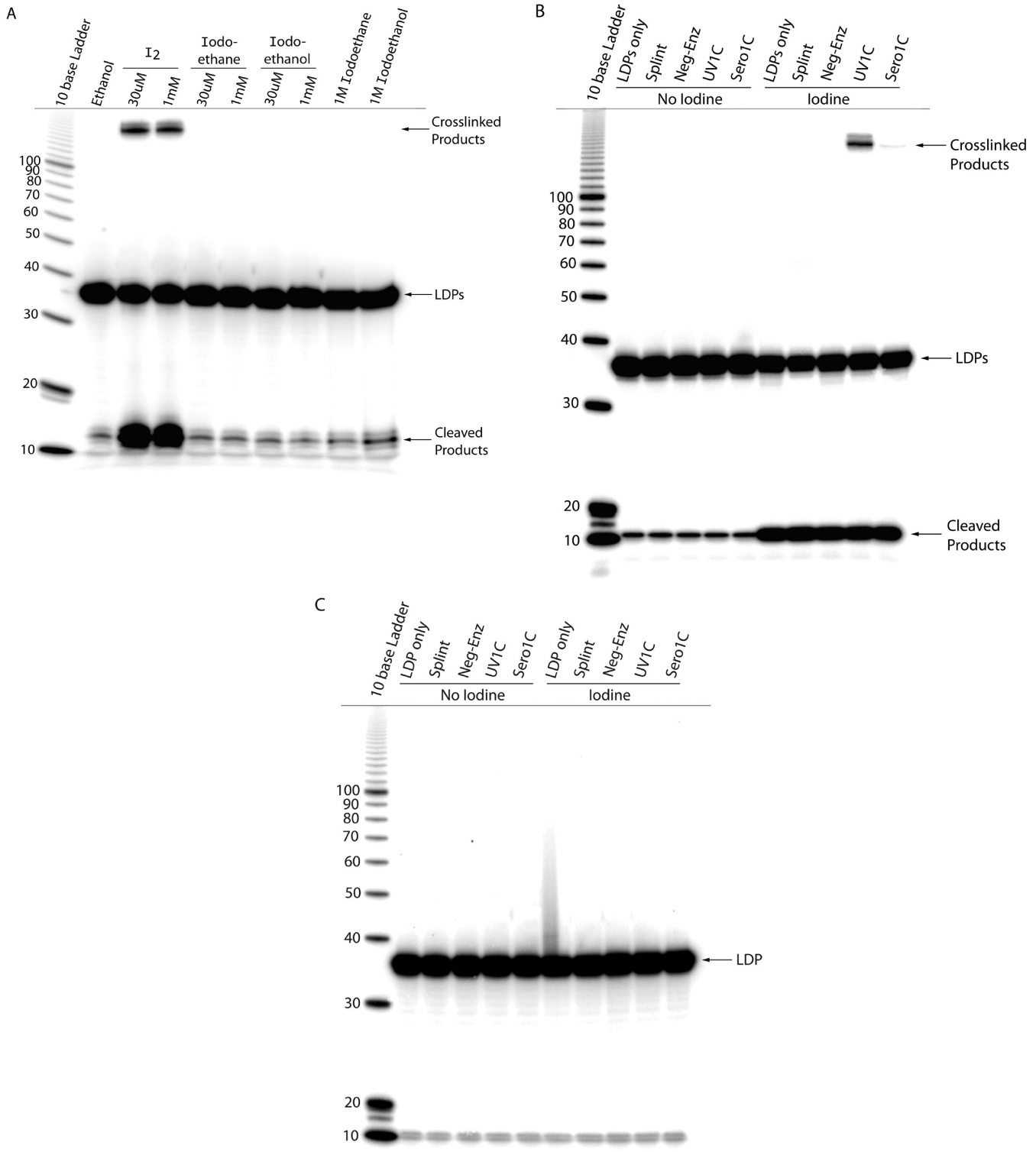


FIGURE 2: Formation of cross-linked products upon treatment of the UV1C·LDPs complex with iodine. (A) Denaturing gel that shows the treatment of the UV1C·LDPs complex with iodine, iodoethane, and iodoethanol. The lane marked Ethanol shows a negative control, in which the UV1C·LDPs complex was not subjected to any of the compounds mentioned above. (B) Denaturing gel showing the formation or lack of formation of cross-linked products from equimolar mixtures of LDPs and a complementary splint oligonucleotide; an LDPs nonbinding oligonucleotide, Neg-Enz; UV1C; and the serotonin-dependent DNAzyme, Sero1C. (C) Denaturing gel showing the same incubations shown in panel B, except that all incubations were conducted with the LDP oligonucleotide rather than LDPs.

produce any comparable species of high molecular weight (Figure 2C).

We can now suggest that the low-gel mobility species (above) are, indeed, some kind of cross-linked products. If so, it is important to define whether they contain UV1C or whether they are merely dimers of the phosphorothioate-containing substrate, LDPs. To address this point, reciprocally labeled cross-linking experiments in which assembled and iodine-treated UV1C·LDPs complexes containing either (a) 5′-^{32}P-radiolabeled UV1C and unlabeled LDPs or (b) unlabeled UV1C and 5′-^{32}P-radiolabeled LDPs were conducted. The data shown in Figure 3 confirm that the major low-gel mobility products are indeed cross-linked

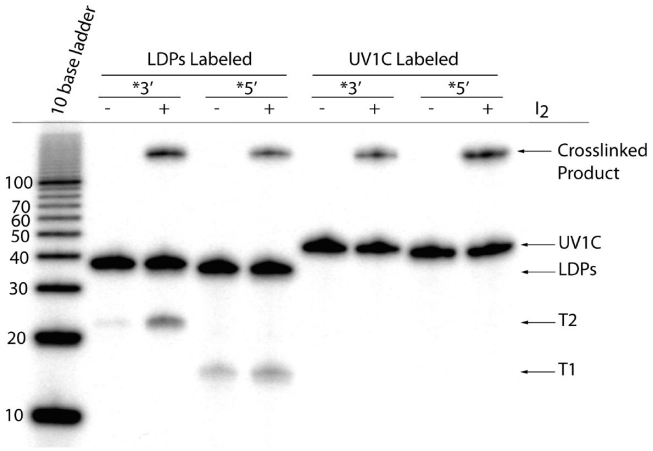


FIGURE 3: Cross-labeling experiment with the aim of investigating the constituents of the major cross-linked species formed by iodine treatment of UV1C·LDPs. The denaturing gel shows the products formed from the UV1C·LDPs complex, when the latter is made up of only one of two constituent oligomers radiolabeled at either its 5′ end or its 3′ end. T1 and T2 refer to the fragment oligonucleotides produced upon cleavage of the LDPs oligonucleotide at its phosphorothioate site.

species containing both the substrate, LDPs, and the DNAzyme, UVIC. These experiments cleanly eliminate the possibility that the products are a dimer of the LDPs substrate, perhaps linked by disulfide bonds. More critically, the production of the same low-gel mobility bands, regardless of whether it is the 3′ or 5′ end of either UV1C or LDPs that is radiolabeled, confirms that the cross-linked species contain full-length (unbroken) versions of both the UV1C and LDPs oligonucleotides.

Mapping the Sites of Cross-Linking. Thiolated nucleobases such as 6-thiouracil and 6-thioguanine have widely been used in protein and nucleic acid light-mediated cross-linking studies as well as between two nucleic acids [a notable recent example is the photo-cross-linking of the 8−17 DNAzyme to its bound substrate (*11*)]. The cross-links formed by such thiolated bases are typically prone to facile piperidine-mediated deglycosylation and concomitant DNA cleavage at the cross-linking site (enabling accurate mapping of such sites). Preliminary treatment of the major iodine-generated cross-linked species obtained from UV1C·LDPs species, however, showed that this cross-linked species was poorly cleaved by piperidine. This result suggests that a phosphate from UV1C is not the likely cross-linking partner for the LDPs phosphorothioate, given that the resulting phosphate−phosphorothioate linkage would undoubtedly be exceedingly labile to base.

To map the cross-linking site or sites in UV1C, we first used hydroxyl radicals generated by a Fenton reaction to footprint the cross-linked species (*12*). If a 5′-radiolabeled UV1C was used to generate the cross-links, the hydroxyl radical-mediated partial degradation of the species should give rise to a continuous nucleotide ladder that stretches from the 5′ end of UV1C to the site of cross-linking (at which point there should be an abrupt termination of the ladder). Likewise, 3′ labeling of UV1C would generate a ladder stretching from the 3′ end of UV1C to the point of cross-linking. Figure 4A shows the data generated from the major cross-linked species formed from 5′-^{32}P-labeled UV1C and unlabeled LDPs, and Figure 4B shows the analogous data obtained from 3′-labeled UV1C and unlabeled LDPs. Both gels identify a cytosine, C14, as the site for the cross-linking in UV1C, although the gel shown in Figure 4A does show a faint shadow at A15, the next base in the UV1C sequence, raising the possibility of the latter's participation in cross-linking to some degree.

To confirm that C14, and possibly A15, is indeed the site of cross-linking, point mutants of UV1C were synthesized in which C14, A15, and G13 (a base that has been shown to participate in the formation of the guanine quadruplex fold of UV1C) (*1*, *2*) were mutated. Thus, G13 was changed to hypoxanthine (in the oligonucleotide UV1C-G13I) or to purine (in UV1C-G13Pu); A15 was mutated to purine (in UV1C-A15Pu) or to thymine (in UV1C-A15T), and, finally, C14 was mutated to thymine (in UV1C-C14T). These various mutant oligonucleotides were tested for their ability to cross-link to LDPs within their respective substrate−DNAzyme complexes.

Figure 5A shows a denaturing gel highlighting the various cross-linked bands formed by the different complexes. As seen in Figure 2A, two cross-linked species of close electrophoretic mobility (labeled X1 and X2 in Figure 5A) are formed from the unmutated UV1C·LDPs complex. Curiously, all the mutated versions of UV1C mentioned above give cross-linked products (at 5−15% yields) whose gel mobilities, for the most part, approximate those of X1, X2, or both. To map the cross-linking sites within these products, hot piperidine-mediated strand cleavage was used, even though the level of cleavage is relatively low. The laddering seen in the appropriate lanes shows nonspecific cleavage of the DNA caused by the piperidine treatment, at all nucleotides but mostly intensely at guanosines.

A key question was whether such piperidine-mediated mapping produces results consistent with the earlier hydroxyl radical mapping. Piperidine cleavage was therefore conducted on the major (X2) and minor (X1) cross-linked products from the UV1C·LDPs complex (with UV1C 5′-radiolabeled), and the strand cleavage results are shown in Figure 5B. The location of the X2 cross-link is found, as with the earlier hydroxyl radical mapping (Figure 4), to be at C14 and A15. The minor X1 species, however, shows slightly ambiguous data. Whereas G25 appears to be the 3′-most band showing in piperidine-treated X1, G23 is the most unambiguously visible 3′-most site of cleavage. This may suggest that both G23 and G25 are participants in the cross-linking.

The cross-linking sites within the different UV1C mutants were then examined, starting with the mutant (UV1C-A15Pu) in which A15 is replaced with a purine and a mutant (in UV1C-A15T) in which A15 is replaced with a thymine. These mutations were made to eliminate the potentially nucleophilic 6-amino functionality of adenine (as in the purine mutation), and the N1, N3, and N7 positions (the thymine mutation). Figure 5A shows that both mutants produce cross-linked products with gel mobilities comparable to that of the major (X2) cross-linked product of unmutated UV1C. Figure 5C shows that purine substitution yields a strong nonspecific piperidine cleavage at the purine site within the un-cross-linked and cross-linked complexes. In the cross-linked species, piperidine cleavage of the 5′-labeled UV1C strand shows a discontinuity of background cleavage bands up to position 15; however, the nonspecific lability of the DNA strand at the purine, perhaps from the non-cross-linked species (indicated by a green arrow), prevents a definitive identification of the purine (as opposed to the adjacent C14) being the cross-linking site. In the lane showing piperidine treatment of the cross-linked species, however, breakdown at the purine results in a reduced mobility band (shown with a blue arrow in Figure 5C) that runs faster than the unbroken cross-linked species itself. This could only be possible if the

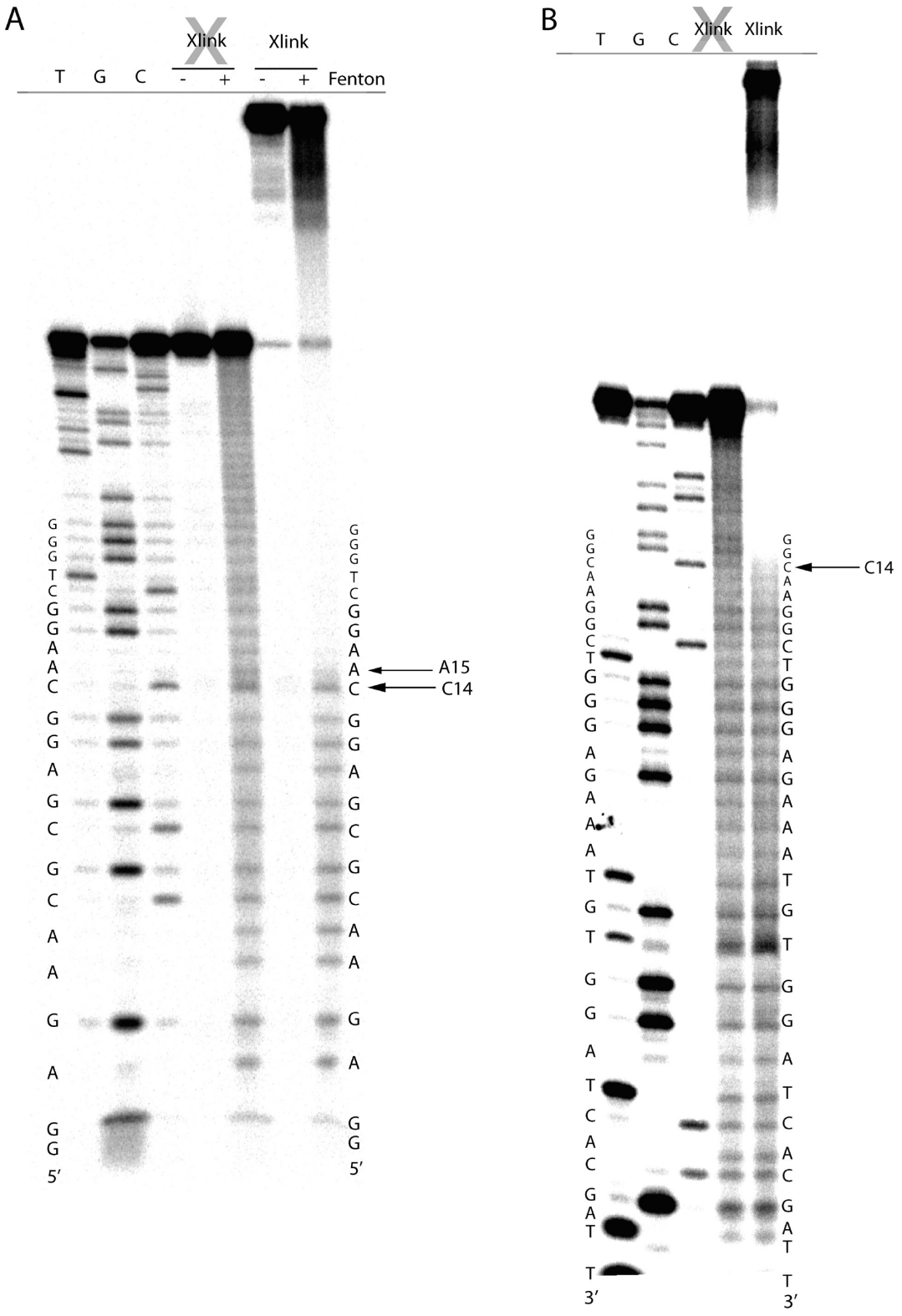
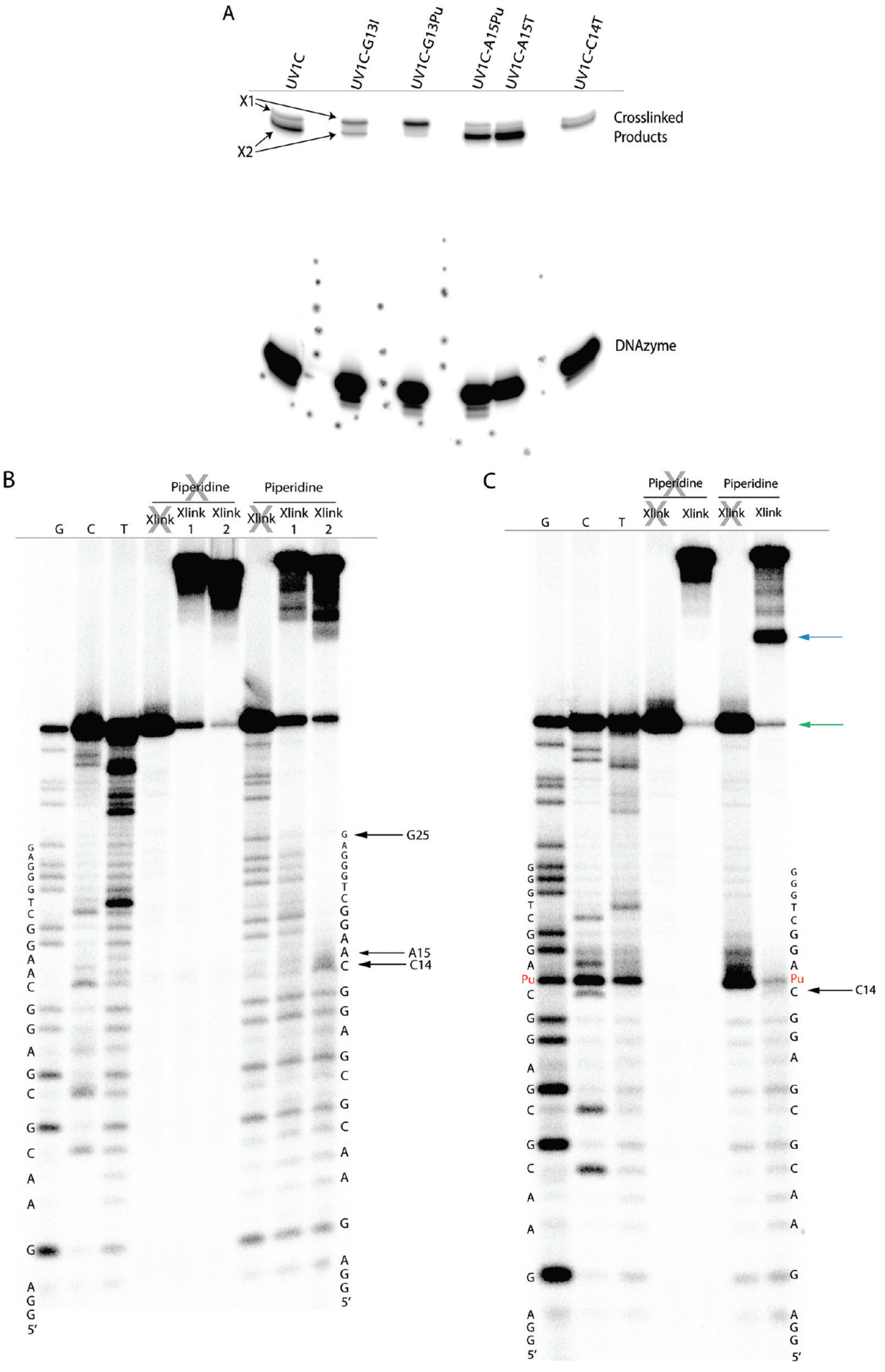


FIGURE 4: Use of the Fenton reaction-generated hydroxyl radicals to map the major cross-linking site in UV1C. Denaturing gels show Fenton-generated ladders from the UV1C·LDPs complex in which (A) the UV1C is 5′-^{32}P-labeled and LDPs is unlabeled and (B) the UV1C is 3′-^{32}P-labeled and LDPs is unlabeled.

purine residue is located 3′ to the site of cross-linking. This suggests that the primary site of cross-linking in UV1C-A15Pu lies at C14. Nonetheless, the fact that a faint purine band appears at position 15 indicates that cross-linking to the purine base cannot be ruled out.

Figure 5D shows that the site of cross-linking in the UV1C-A15T mutant continues to be at C14, as the intensity of the shadow band seen at position 15 is significantly reduced, when compared with that of the wild-type cross-link. Conversely, the mapping of the UV1C-C14T cross-link (Figure 5E) shows that this mutation shifts the site of cross-linking from position 14 to position 15 predominantly.

Cumulatively, the data presented above indicate that both C14 and A15 positions in UV1C are bona fide cross-linking sites to


A
UV1C
UV1C-G13I
UV1C-G13Pu
UV1C-A15Pu
UV1C-A15T
UV1C-C14T
X1
X2
Crosslinked Products
DNAzyme
B
Piperidine
Piperidine
Xlink
Xlink 1
Xlink 2
G C T
G25
A15
C14
C
Piperidine
Piperidine
Xlink
Xlink
G C T
C14

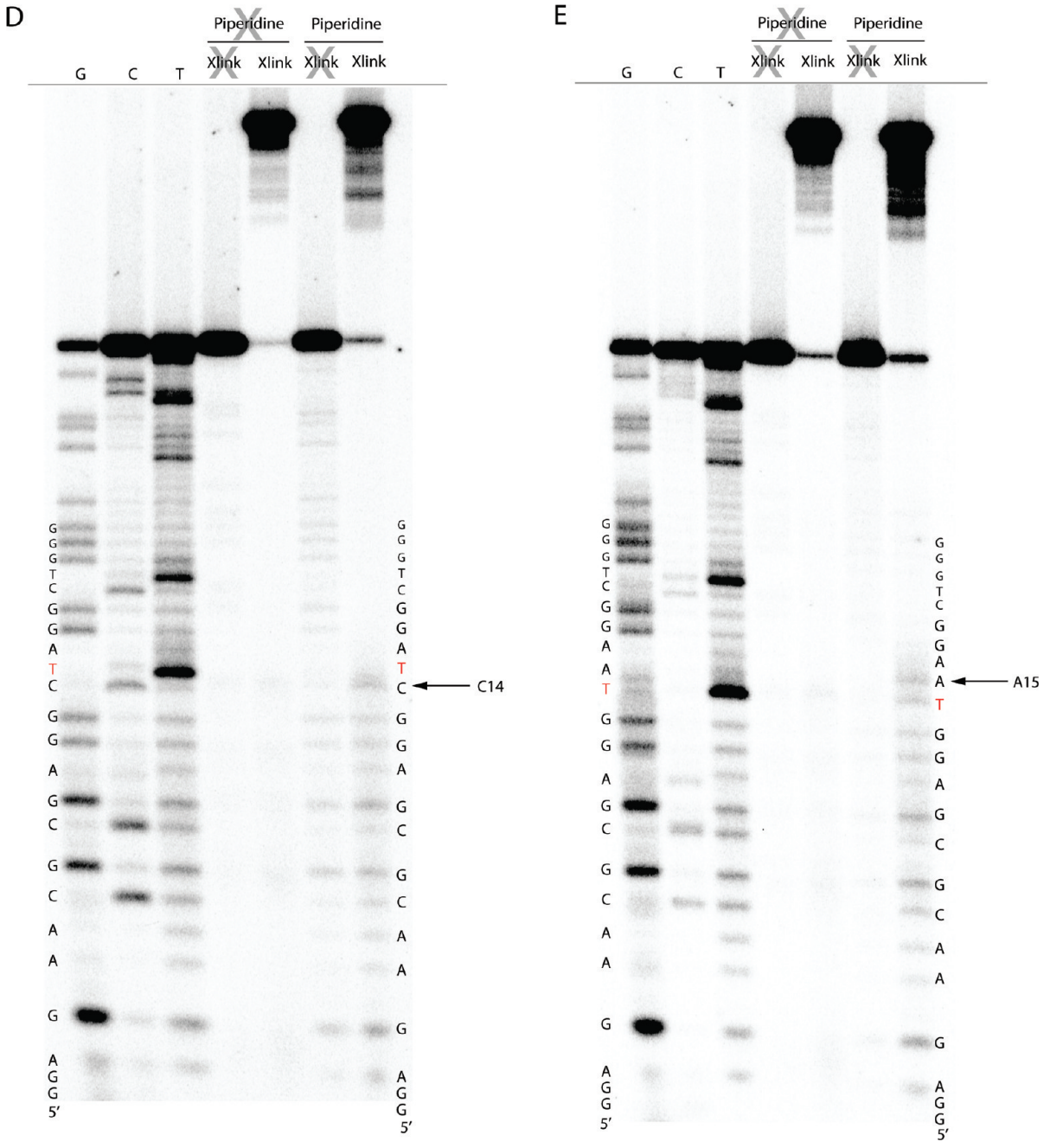


FIGURE 5: (A) Denaturing gel showing the formation of cross-linked species formed by LDPs complexed to UV1C, and to five UV1C point mutants. The two cross-linked products formed by the UV1C·LDPs complex are labeled as X1 and X2. (B−E) Denaturing gel showing piperidine-generated sequence ladders to pinpoint the site of cross-linking in 5′-^{32}P-labeled UV1C (B), with UV1C-A15Pu (C), with UV1C-A15T (D), and with UV1C-C14T (E).

the substrate's phosphorothioate. Elimination of one site (C14) nevertheless retains cross-linking to the other site (A15).

It is curious that the two mutations to residue G13 that we have examined (UV1C-G13I and UV1C-G13Pu) continue to generate cross-links with LDPs (Figure 5A). G13 has been shown to participate in the required G-quadruplex fold of UV1C, and mutation of this residue to purine (in UV1C-G13Pu) should disrupt the G-quadruplex and perturb the overall fold of the UV1C·LDPs complex. However, the persistence of cross-linking in this mutant suggests that even though the G-quadruplex itself may be compromised, the substrate and the DNAzyme continue to bind together, enabling the cross-linking of one or more DNAzyme residues to the substrate. It is notable in this respect that the cross-links formed by the G13Pu mutant run slower than the X2 species formed by unmodified UV1C (Figure 5A). The G13I mutant, however, generates bands with mobilities comparable to those of the X1 and X2 species. We have shown previously that mutation of G13 to hypoxanthine (in UV1C-G13I) will maintain the G-quadruplex fold while weakening it (*2*). Piperidine cleavage of the different cross-linked species formed by UV1C-G13Pu and UV1C-G13I shows that the UV1C-G13Pu mutant cross-links generate broad DNA cleavage patterns, with no single dominant site(s) of cross-linking to LDPs (data not shown). By contrast, the X2-like species formed by UV1C-G13I show cross-linking at C14, while the X1 species do not show a clear cross-linking site (analogous to the situation with UV1C-G13Pu).

Cross-Linking Site Identification Using Primer Extension. In addition to piperidine cleavage, we have also mapped the

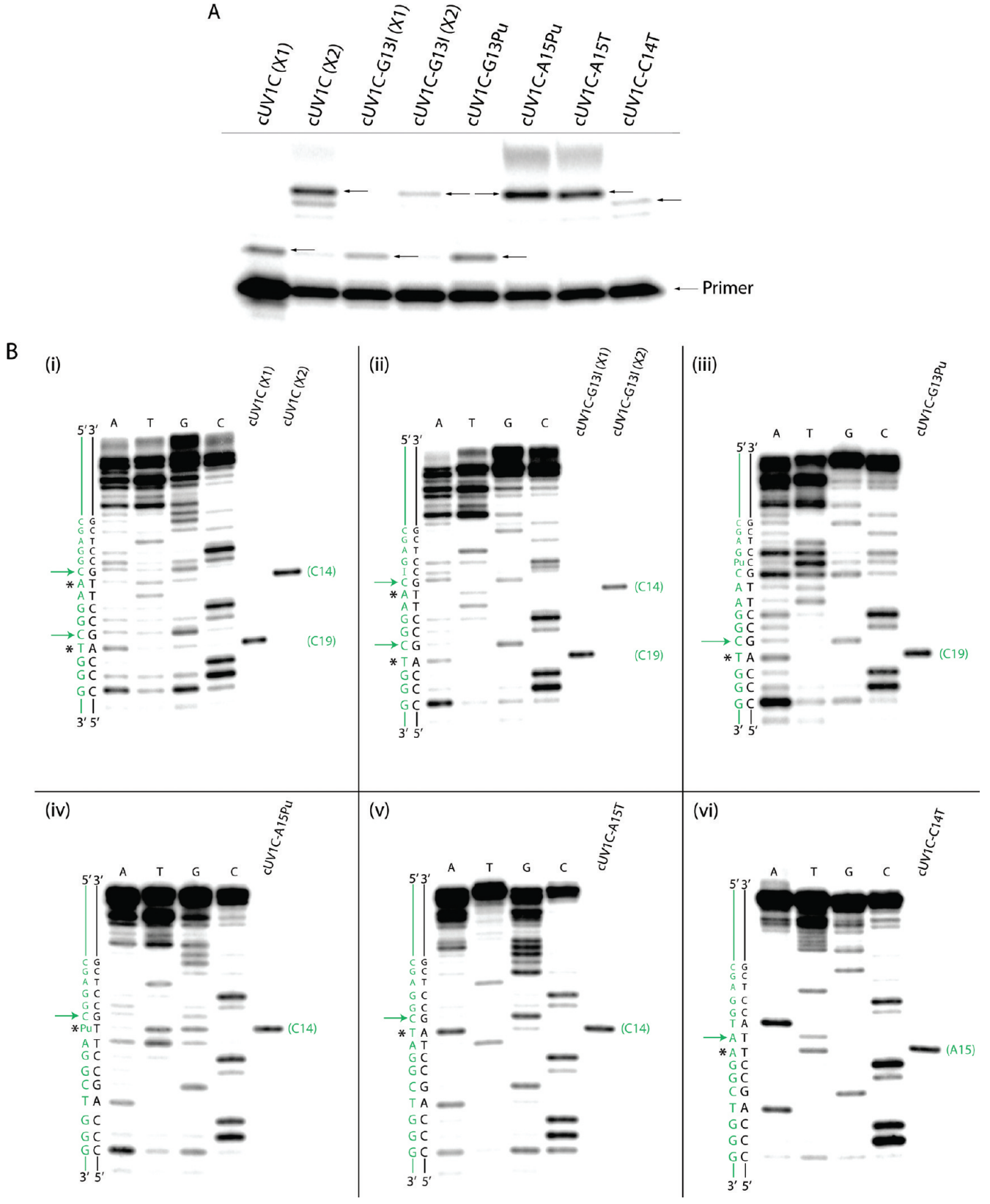


FIGURE 6: (A) Denaturing gel showing primer extension products obtained from cross-linked species formed, variously, by the complexes formed by LDPs with unmutated as well as mutated UV1C oligonucleotides. (B) Full-sized primer extension products shown with horizontal arrows in panel A were extracted for accurate sizing against reference dideoxy sequencing ladders. Panels i−vi show these data. The sizing ladders show the sequence of the complement of UV1C, written in black alongside. The actual sequence of UV1C is written in green. The base in the UV1C sequence shown with an asterisk is complementary to the 3′-most base of the primer extension product. The actual site of cross-linking in UV1C is shown with a green arrow in each case, and its identity is indicated in parentheses adjacent to the primer extension band in question.

cross-linking sites in UV1C, and its mutants, by primer extension using a 20-nucleotide DNA primer that is complementary to the 3′-most 20 nucleotides of UV1C. The expectation is that primer extension will proceed up to the 3′-most cross-linking site within

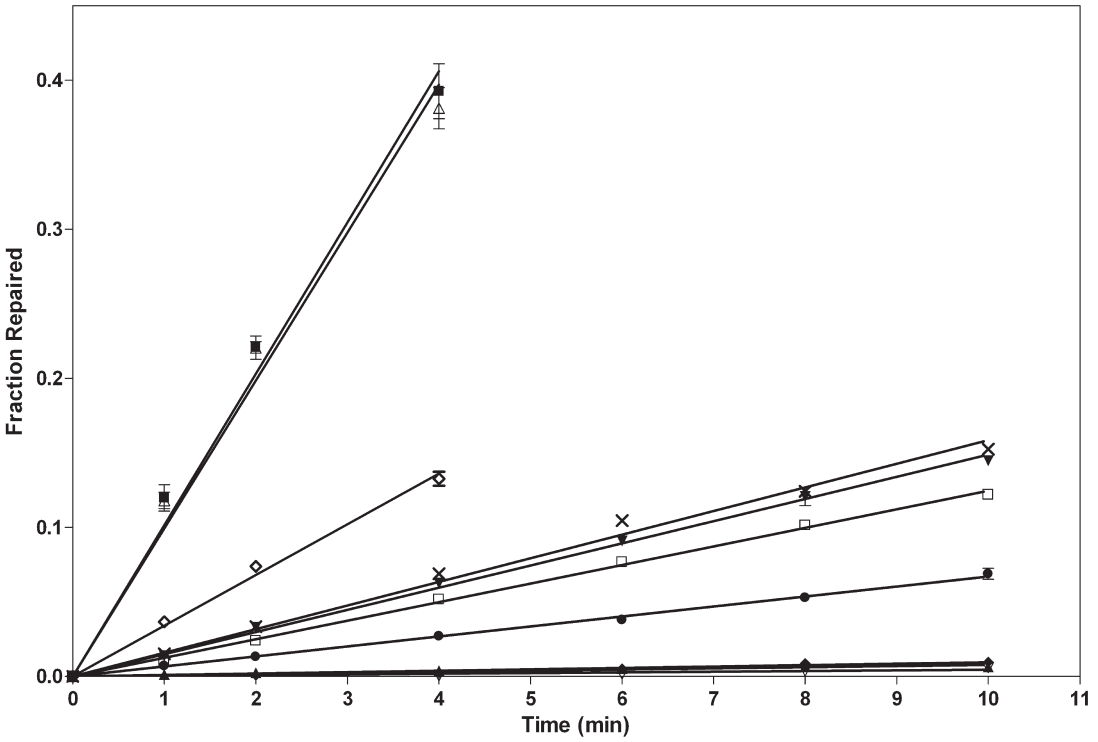


FIGURE 7: Initial rates of photoreactivation of the thymine dimer-containing TDP substrate, catalyzed under single-turnover conditions, by UV1C (■), Neg-Enz (▲), G13I (▼), G13Pu (◆), A15Pu (●), A15T (□), C14T (△), AAAT (▽), and CA-TT (×). DNAzyme-AAAT repair of the TDP-ATGT-TAAA substrate (◊).

a given molecule of UV1C and then come to a halt. Figure 6A shows the primer extension data obtained from the various cross-linked species formed by UV1C and its mutants. An immediate observation is that the X1 and X1-like cross-linked species all generate notably shorter primer extension products than the X2 and X2-like species.

The full-sized primer extension products (shown with horizontal arrows in Figure 6A) are extracted from the gel for accurate sizing in comparison with reference dideoxy sequencing ladders. Figure 6B (panels i−vi) shows these data (note that the sizing ladders used here report the sequence of the complement of UV1C, written in black alongside; the actual sequence of UV1C is written in green beside it). The UV1C base(s) complementary to the 3′ termini of the various primer extension products (shown with an asterisk, Figure 6B) can be identified by reading along the UV1C (or mutant) sequence written in green. The actual site of cross-linking is at position N-1 (marked with a green arrow) relative to the nucleotide marked with an asterisk. Figure 6B (panel i) shows in the case of the unmodified UV1C DNAzyme, cross-link X1 localizes to C19 whereas X2 localizes to C14. These results corroborate the data obtained from piperidine cleavage for the major, X2 product, which had localized to C14 (Figure 5B). In the case of X1, the piperidine data had shown its cross-linking site to be at G23/G25 (not at C19). However, a key feature of the primer extension method, using primer 1, is the fact that the longest extension product obtained will refer to the 5′-most cross-linking site on UV1C, whereas the piperidine cleavage studies (conducted with 5′-end-labeled UV1C) would highlight the 3′-most cross-linking site on UV1C. To see if G23 and/or G25 could directly be identified as cross-linking sites within the UV1C·LDPs complex, using primer extension, we employed a shorter, 15-nucleotide, primer (primer 2) which nevertheless generated products indicating cross-linking at C14 for X2 and C19 for X1 (data not shown). These results are essentially identical to those obtained using the longer primer 1. While the underlying causes of this result are not immediately apparent, a possible explanation may be an unusually complex and inaccessible secondary/tertiary structure within those cross-linked products of the UV1C·LDPs complex in which G23 and G25 participate.

Similarly, cross-linking loci at C14 and C19 are seen for all of the mutant UV1C oligonucleotides (Figure 6B, panels ii−vi), with the notable exception of the UV1C-A15T mutant, where the cross-linking site is at A15 [again, corroborating the earlier piperidine results (Figure 5E)].

Kinetics of Photoreactivation by the Mutant DNAzymes. To test whether the various UV1C mutants examined for cross-linking are actually catalytically active, the initial rates of TDP photoreactivation by these mutant DNAzymes were measured. Figure 7 plots the fractions of the substrate repaired as a function of time, under single-turnover conditions. Interestingly, only the C14T mutant repairs thymine dimers at the same rate as the unmodified UV1C. As described above, our mapping studies have shown C14 to be the most prominent cross-linking site in UV1C, yet mutation of this cytosine to a thymine does not impair catalytic function. The findings on the cross-linking base, A15, however, are particularly interesting in that mutation of this base to either purine or thymine sharply reduces the catalytic ability of UV1C, decreasing the rate of photoreactivation to only 7% (A15Pu) and 14% (A15T) of the photoreactivation rate of unmodified UV1C.

A hypothesis to explain the data given above is the following: since C14 is the most prominent site for cross-linking, it is likely located very close (perhaps the closest UV1C nucleobase) to the substrate's phosphorothioate. If so, it would necessarily be located a few angstroms from the thymine bases that make up the dimer within TDP. A15, a lesser site for cross-linking to the phosphorothioate, is likely to be farther from the

phosphorothioate than C14 but may be positioned closer to the dimerized thymine bases. The sharp reduction in the catalytic activity of the A15 mutants raises the possibility that this base contributes in some way to the chemistry of catalysis. Though the strong cross-linking observed with C14 suggests its close proximal positioning to the phosphorothioate, that is not a definite conclusion. Since the precise chemistry of the cross-linking is yet to be determined, the possibility that perhaps the greater observed yield of C14 cross-linking may be more due to a favorable chemistry than a proximal location cannot be ruled out.

In contrast, substitution of G13 with a purine completely abolishes photoreactivation activity. This is understandable since G13 has been identified in prior studies as a participant in the G-quadruplex fold of UV1C (*1*, *2*). Mutation of G13 undoubtedly disrupts the two-quartet quadruplex, and since the latter likely acts both as the light-harvesting antenna and as the electron source for the photoreactivation reaction, it is understandable that the G13Pu mutant should be catalytically inactive. This last finding is also supported by our cross-link mapping data, which could not assign a definitive cross-linking site within UV1C-G13Pu. The G13I mutant, by contrast, maintains approximately 17% of the unmodified UV1C's catalytic activity. Inosines are known to participate in quartets with guanosines (*13*) (albeit more weakly), and the G to I mutation is therefore less likely to compromise the formation of the G-quadruplex than the guanine to purine mutation (*2*).

We now tested the effect of mutating both C14 and A15 to thymines. This mutant oligomer, UV1C-CA-TT, gives rise to a nonspecific cross-link (data not shown), and its catalytic activity is reduced to the level of the UV1C-A15T mutant (Figure 7). This result is consistent with our findings (above) that the mutation of C14 to T does not substantially affect the catalytic proficiency of the mutant whereas mutation of A15 to the plain purine or to thymine does.

Identification of a Second Substrate-Binding Arm in UV1C. Our earlier studies (*2*) had proposed that the 5′ arm of the UV1C may behave uniquely as the substrate-binding arm via base pairing with TDP/LDP (Figure 1). Because the extent of postulated Watson−Crick base pairing between the two seemed relatively small, it appeared likely that additional, perhaps tertiary, interactions were necessary to stabilize the binding of the DNAzyme and the substrate (*2*). Inspection of the new cross-linking data reported here, however, and a concomitant reexamination of the UV1C G-quadruplex's topology raise the possibility that the AAATG element within loop 3 [L3 (Figure 1)] of the quadruplex might be available for additional base pairing with the substrate. To test this idea, we mutated the AAATG element to TTTAG (in the DNAzyme construct UV1C-ATGT-TAAA). This mutation results in a complete knockout of the mutant DNAzyme's cross-linking (Figure 8) as well as catalytic activity (Figure 7) with the respective natural substrates. The use of "rescue mutation" substrates (LDPs-ATGT-TAAA and TDP-ATGT-TAAA), which would reestablish base pairing with the L3 loop of the mutant DNAzyme, provides striking data. The combination of the L3 loop mutated DNAzyme (UV1C-ATGT-TAAA) with the TDP-ATGT-TAAA substrate restores 30% of the thymine dimer reactivation activity with respect to the original DNAzyme (Figure 7). Also, binding of the mutated DNAzyme to the LDPs-ATGT-TAAA substrate reestablishes cross-linking (Figure 8). Cumulatively, these results confirm the plausibility of L3 serving as a second substrate recognition and binding motif within the UV1C DNAzyme.

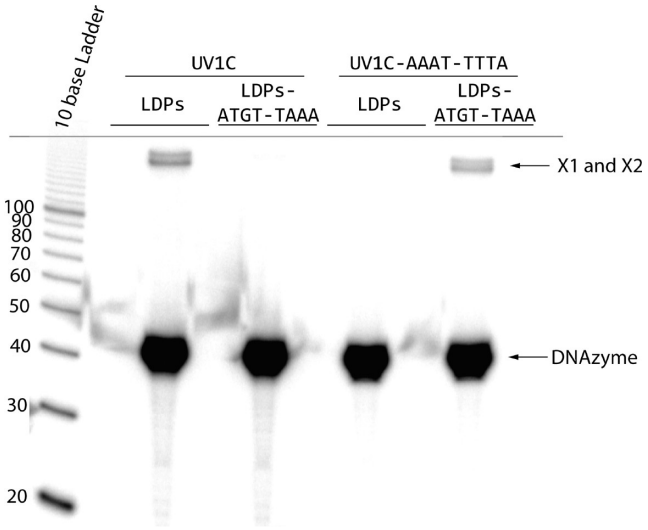


FIGURE 8: Denaturing gel showing the knockout of cross-linking of the natural LDPs substrate with a mutated UV1C DNAzyme (UV1C-AAAT-TTTA). Also shown is the rescue of cross-linking of the mutant DNAzyme when a substrate containing a compensating mutation (LDPs-ATGT-TAAA) replaces LDPs.

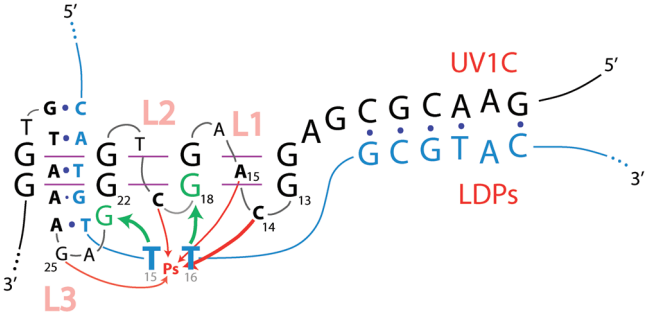


FIGURE 9: Refined model for the UV1C·LDPs/LDP/TDP DNAzyme·substrate complex. All three loops of the G-quadruplex fold of UV1C are postulated to be propeller or double-chain reversal loops. The green arrows show the two major sites of iodoracil-mediated cross-linking (*2*), whereas the red arrows indicate the UV1C bases found in this study to participate in cross-linking with the substrate's phosphorothioate functionality. The AAATG sequence within loop 3 of UV1C is shown as base pairing to the substrate.

A Refined Model for the Folding of the UV1C·LDPs Complex. Consideration of the new cross-linking data described here, along with the likely substrate binding role of the AAATG element within loop 3 of UV1C's G-quadruplex, enables us to improve the preexisting structural model (shown in Figure 1) for the UV1C·LDPs complex. Figure 9 shows the refined model. The new cross-linking sites that we have established within UV1C map to C14, A15, C19, and G23/G25. Therefore, it is reasonable to hypothesize that in the tertiary structure of the UV1C·LDPs complex, these residues (located in different loops of UV1C's G-quadruplex) cluster in the proximity of the substrate's phosphorothioate functionality. In the prior topological model of the DNAzyme (Figure 1), loop 1 had been postulated to be distal to the thymine dimer on the substrate. Moreover, as mentioned above, base pairing interactions had been proposed for only the 5′ end of UV1C with the 3′ end of LDP. In our refined model of the UV1C·LDPs complex (Figure 9), L1 is transformed from an "edgewise" loop to a double-chain reversal or "propeller" loop (*14*, *15*). This topological change enables

C14 to be positioned within contact range of the substrate's phosphorothioate. Both L1 and L2 are thus predicted to be propeller loops. On the basis of the available data, we cannot be certain about loop 3 (a large loop, containing the putatively substrate-binding AAATG motif), but we hypothesize that this, too, is a propeller loop, leading to a wholly parallel-stranded orientation of the intramolecular G-quadruplex. Such intramolecular G-quadruplexes with all-parallel strand orientations have been reported in the literature (*16*).

DISCUSSION

The unexpected formation of DNA−DNA cross-linked products between UV1C and LDPs within the UV1C·LDPs complex is consistent with the notion that within the complex, elements of the UV1C DNAzyme are positioned close enough to the sulfur moiety of LDPs to enable the formation of a cross-link between them in the presence of iodine. The precise chemistry of the cross-linking, however, is not clear at this stage. Normally, the addition of iodine to a phosphorothioate-containing oligomer might be expected to generate a sulfenyl iodide intermediate (*17*, *18*). Since the UV1C cytosine, adenine, and guanine bases involved in cross-linking all contain exocyclic amino groups, it is possible that they are indeed the nucleophiles attacking the postulated phosphorosulfenyl iodide. However, attack by ring nitrogens within these nucleobases, which may be seen in the case of the Pu mutation, cannot be ruled out. Also, the precise locus of nucleophilic attack is not yet clear. In principle, attack could occur to the sulfur, leading to the formation of an N−S−P linkage and the departure of the iodide anion; alternatively, attack could occur to the phosphorus, leading to the formation of a phosphoramidate and departure of the IS^- anion. Possible reaction schemes are summarized in Figure S1 of the Supporting Information. Future experiments will aim to clarify the chemical identity or identities of the cross-links.

One extremely interesting feature of these phosphorothioate-generated cross-links is that the incorporation of a phosphorothioate into a DNA or RNA strand naturally produces a chiral center. Separation of DNA or RNA populations containing either a pure R_p or a pure S_p phosphorothioate diastereomer (*19*) would enable a further refinement of the kinds of structural information that the cross-links are capable of providing. Thus, exploration of stereospecific effects on the patterns of cross-linking within DNAzyme·substrate complexes will remain an important priority for future experiments.

On the basis of our mapping of the observed cross-links, we put forth a refined model for the folding of the UV1C·LDPs complex. The centerpiece of this structure is the catalytically key G-quadruplex formed by UV1C, for which an all-parallel strand orientation is now proposed. G-Quadruplexes of different strand orientations typically display characteristically different circular dichroism spectra (*14*, *15*). In principle, CD spectroscopy might be usable for verifying our refined model for the UV1C·LDPs complex. However, unlike DNAs that fold to form pure G-quadruplexes (with no other major secondary structures present), the UV1C·LDPs complex incorporates diverse structural elements besides the G-quadruplex and is likely to display a complex CD spectrum. Another possible approach toward verifying our structural model might be to investigate whether specific guanines participating in the G-quadruplex adopt a *syn* or *anti* glycosidic conformation. It is known that an all-parallel strand orientation in a quadruplex [such as that proposed here (Figure 9)] would imply *anti* orientations for all of the quadruplex guanosines, whereas antiparallel quadruplexes would incorporate mixtures of *syn* and *anti* guanosines (*13*). It has been shown that riboguanosine nucleotides intrinsically prefer (relative to deoxyriboguanosines) the *anti* glycosidic orientation, whereas 8-bromodeoxyriboguanosine nucleotides strongly prefer the *syn* conformation (*17*, *18*). Future experiments will systematically substitute stretches of the guanines participating in quadruplex formation with the nucleoside analogues described above, to test our model for the complex topology of UV1C·LDPs.

Here, we have also supplied evidence of the existence of a second substrate-binding arm in UV1C. In conjunction with the first substrate-binding arm, the new model for the enzyme·substrate complex places the dimer-forming thymines directly at the base of the G-quadruplex. The identification of a second substrate-binding arm will enable us, in the future, to manipulate the sequences of both hybridizing arms. Development of DNAzymes capable of targeting DNA substrates of divergent sequences may now be possible. Importantly, also, the base composition of these interacting sequences could be altered to enhance the stability of the enzyme·substrate complex, and thereby to enhance the likelihood of growing diffractable crystals. NMR spectroscopy and/or a high-resolution crystal structure would undoubtedly provide deep insights into the workings of the UV1C DNAzyme.

The phosphorothioate cross-linking method is in itself an unexpected finding, one that should enable a new kind of investigation of the active sites of diverse ribozymes and DNAzymes, natural and artificial, as well as enable the probing a variety of RNA·protein complexes. The literature describes one documented instance of formation of an intraprotein iodine-dependent cross-link between a cysteine and lysine residue (*20*). The extension of such a methodology to nucleic acids (as well as, potentially, to nucleic acid·protein complexes) opens up its application to a variety of biologically relevant systems as a means of harvesting significant structural data. Systems that potentially lend themselves to immediate investigation include RNA-cleaving DNAzymes and ribozymes (*21*, *22*), in which the facile incorporation of a phosphorothioate group at the substrate's scissile site will, as shown here, yield powerful insights into their active sites.

ACKNOWLEDGMENT

We are grateful to the members of the Sen Laboratory, Andrew Bennet, Fritz Eckstein, Joe Piccirilli, and Hani Zaher, for valuable discussions.

SUPPORTING INFORMATION AVAILABLE

Possible cross-link formation chemistry (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Chinnapen, D. J., and Sen, D. (2004) A deoxyribozyme that harnesses light to repair thymine dimers in DNA. *Proc. Natl. Acad. Sci. U.S.A. 101*, 65–69.
2. Chinnapen, D. J., and Sen, D. (2007) Towards elucidation of the mechanism of UV1C, a deoxyribozyme with photolyase activity. *J. Mol. Biol. 365*, 1326–1336.
3. Thorne, R. E., Chinnapen, D. J., Sekhon, G. S., and Sen, D. (2009) A deoxyribozyme, Sero1C, uses light and serotonin to repair diverse pyrimidine dimers in DNA. *J. Mol. Biol. 388*, 21–29.

4. Sancar, A. (2003) Structure and function of DNA photolyase and cryptochrome blue-light photoreceptors. *Chem. Rev. 103*, 2203–2237.
5. Schatz, D., Leberman, R., and Eckstein, F. (1991) Interaction of *Escherichia coli* tRNA(Ser) with its cognate aminoacyl-tRNA synthetase as determined by footprinting with phosphorothioate-containing tRNA transcripts. *Proc. Natl. Acad. Sci. U.S.A. 88*, 6132–6136.
6. Gish, G., and Eckstein, F. (1988) DNA and RNA sequence determination based on phosphorothioate chemistry. *Science 240*, 1520–1522.
7. Cochran, J. C., and Strobel, S. A. (2004) Probing RNA structure and function by nucleotide analog interference mapping. Current Protocols in Nucleic Acid Chemistry (Wiley and Sons, Hoboken, N.J.), Chapter 6, Unit 6.9.
8. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A. 12*, 5463–5467.
9. Maxam, A. M., and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.
10. Rubin, M. C., and Schmid, C. W. (1980) Pyrimidine-specific chemical reactions useful for DNA sequencing. *Nucleic Acids Res. 8*, 4613–4620.
11. Liu, Y., and Sen, D. (2008) A contact photo-cross-linking investigation of the active site of the 8−17 deoxyribozyme. *J. Mol. Biol. 81*, 845–859.
12. Tullius, T. D., and Greenbaum, J. A. (2005) Mapping nucleic acid structure by hydroxyl radical cleavage. *Curr. Opin. Chem. Biol. 9*, 127–134.
13. Williamson, J. R., Raghuraman, M. K., and Cech, T. R. (1989) Monovalent cation-induced structure of telomeric DNA: The G-quartet model. *Cell 59*, 871–880.
14. Phan, A. T., Kuryavyi, V., and Patel, D. J. (2006) DNA architecture: From G to Z. *Curr. Opin. Struct. Biol. 16*, 288–298.
15. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K., and Neidle, S. (2006) Quadruplex DNA: Sequence, topology, and structure. *Nucleic Acids Res. 34*, 5402–5415.
16. Parkinson, G. N., Lee, M. P., and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature 417*, 876–880.
17. Fraenkel-Conrat, H. (1955) The reaction of tobacco mosaic virus with iodine. *J. Biol. Chem. 217*, 373–381.
18. Field, L., and White, J. E. (1973) A Stable Solid Model for Protein Sulfenyl Iodides. *Proc. Natl. Acad. Sci. U.S.A. 70*, 328–330.
19. Connolly, B. A., Potter, B. V., Eckstein, F., Pingoud, A., and Grotjahn, L. (1984) Synthesis and characterization of an octanucleotide containing the EcoR1 recognition sequence with a phosphorothioate group at the cleavage site. *Biochemistry 23*, 3443–3453.
20. Nguitragool, W., and Miller, C. (2007) Inaugural Article: CLC Cl/H^+ transporters constrained by covalent cross-linking. *Proc. Natl. Acad. Sci. U.S.A. 104*, 20659–20665.
21. Doudna, J. A., and Cech, T. R. (2002) The chemical repertoire of natural ribozymes. *Nature 418*, 222–228.
22. Baum, D. A., and Silverman, S. K. (2008) Deoxyribozymes: Useful DNA catalysts in vitro and in vivo. *Cell. Mol. Life Sci. 65*, 2156–2174.