A Residue in MutY Important for Catalysis Identified by Photocross-Linking and Mass Spectrometry†

Cindy Lou Chepanoske,‡ Olga A. Lukianova, Murielle Lombard,§ Marie-Pierre Golinelli-Cohen,|| and Sheila S. David*

Department of Chemistry, University of Utah, 315 S 1400 E, Salt Lake City, Utah 84112-0850

Received August 27, 2003; Revised Manuscript Received November 11, 2003

ABSTRACT: MutY is an adenine glycosylase in the base excision repair (BER) superfamily that is involved in the repair of 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG):A and G:A mispairs in DNA. MutY contains a $[4Fe-4S]^{2+}$cluster that is part of a novel DNA binding motif, referred to as the iron−sulfur cluster loop (FCL) motif. This motif is found in a subset of members of the BER glycosylase superfamily, defining the endonuclease III-like subfamily. Site-specific cross-linking was successfully employed to investigate the DNA−protein interface of MutY. The photoreactive nucleotide 4-thiothymidine (^{4S}T) incorporated adjacent to the OG:A mismatch formed a specific cross-link between the substrate DNA and MutY. The amino acid participating in the cross-linking reaction was characterized by positive ion electrospray ionization (ESI) tandem mass spectrometry. This analysis revealed Arg 143 as the site of modification in MutY. Arg 143 and nearby Arg 147 are conserved throughout the endo III-like subfamily. Replacement of Arg 143 and Arg 147 with alanine by site-directed mutagenesis reduces adenine glycosylase activity of MutY toward OG:A and G:A mispairs. In addition, the R143A and R147A enzymes exhibit a reduced affinity for duplexes containing the substrate analogue 2′-deoxy-2′-fluoroadenosine opposite OG and G. Modeling of MutY bound to DNA using an endonuclease III-DNA complex structure shows that these two conserved arginines are located within close proximity to the DNA backbone. The insight from mass spectrometry experiments combined with functional mutagenesis results indicate that these two amino acids in the $[4Fe-4S]^{2+}$ cluster-containing subfamily play an important role in recognition of the damaged DNA substrate.

Oxidative stress has been implicated as an important contributor to cancer and the aging process (*1*−*3*). The DNA bases are particularly susceptible to modifications resulting from oxidative stress, and the oxidation potential of guanine renders it the most likely of the four DNA bases to be oxidized (*4*). Arguably, the most notorious oxidation product of guanine formed in DNA is 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG)[1] (*5*). The ability of OG to mispair with A during DNA replication results in GC → TA transversion mutations (*6*). The mutagenic potential of OG threatens all organisms; therefore, repair systems for OG are present across all phylogeny (*7, 8*). In *Escherichia coli*, the GO repair pathway is dedicated to the prevention of mutations caused by OG and relies on three enzymes, Fpg (MutM), MutY, and MutT (*9*−*11*). Fpg is a base-excision repair (BER) glycosylase that removes OG from OG:C base-pairs, while the MutY glycosylase removes the undamaged adenine from an OG:A mismatch. MutT catalyzes the hydrolysis of d(OG)-TP to prevent its misincorporation by DNA polymerases into nascent DNA strands. Recently, a direct link between inherited defects in the human homologue of MutY, MYH, and adenomatous polyposis colon cancer has been established (*12*), indicating the importance of the repair system for OG in the prevention of carcinogenesis.

The adenine glycosylase MutY has been shown to remove adenine from a variety of mispairs, most notably OG:A, G:A, and C:A mispairs (*9, 13*−*15*). However, the importance of these activities relative to the activity with OG:A substrates remains to be established. The adenine glycosylase MutY is composed of two domains that are susceptible to separation upon partial proteolytic digestion. The N-terminal domain

† This work was supported by a grant to S.S.D. from the National Institutes of Health (CA67983). Mass spectral data were acquired at the University of Utah Mass Spectrometry Facility, supported in part by N.I.H. Grant P30 CA42014.

* Corresponding author. E-mail: david@chemistry.utah.edu. Tel: (801) 585-9718. Fax: (801) 581-8433.

‡ Present address: Prolexys Pharmaceuticals, 2150 Dauntless Ave., Salt Lake City, UT 84116.

§ Present address: Groupe Biocatalyse et Chimie fine, Université de la Méditerranée, Faculté des Sciences de Luminy, Case 901, 163 Avenue de Luminy, 13288 Marseille, France.

|| Present address: Laboratoire d'Enzymologie et Biochimie Structurales, Centre National de la Recherche Scientifique, UPR 9063, 1 Avenue de la Terras, 91190 Gif sur-Yvette, France.

[1] Abbreviations: ^{4S}T, 4-thiothymidine; AP, apurinic/apyrimidinic; *B. st., Bacillus stearothermophilus;* BER, base excision repair; Endo III, *Escherichia coli* endonuclease III; ESI, electrospray ionization; F, 2′-deoxyformycin A; FA, 2′-deoxy-2′-fluoroadenosine; FCL, $[4Fe-4S]^{2+}$ cluster loop; HhH, helix−hairpin−helix; MS/MS, tandem mass spectrometry; OG, 7,8-dihydro-8-oxo-2′-deoxyguanosine; PAGE, polyacrylamide gel electrophoresis; R143A, mutated form of MutY with R143 replaced with alanine; R147A, mutated form of MutY with R147 replaced with alanine; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; TIC, total ion chromatograph.

10.1021/bi035537e CCC: $27.50
Published on Web 12/31/2003

(Met 1 to Lys 225) retains catalytic activity and exhibits a sequence and structure that place MutY within the base-excision repair superfamily of DNA glycosylases. The C-terminal domain has sequence (*16*) and structural (*17*) homology to the d(OG)TPase MutT. This similarity to MutT, coupled with kinetic experiments with substrates (*16, 18*) and binding assays with substrate analogues (*18*), suggests a role of the C-terminal domain in OG recognition and nucleotide flipping. MutY contains a $[4Fe-4S]^{2+}$cluster within the N-terminal domain, and the X-ray structure (*19*) revealed a solvent-exposed loop formed by the polypeptide chain that contains the first two cysteine ligands of the $[4Fe-4S]^{2+}$ cluster (Ile 191 to Cys 199), similar to that previously observed in endo III (*20*). This region, referred to as the iron−sulfur cluster loop (FCL) motif (*20*), has been implicated to play a role in damage recognition and repair (*20, 21*). MutY and endo III contain a unique spacing of the four cysteine ligands to the $[4Fe-4S]^{2+}$ cluster; these four conserved cysteines are also present in a subset of members of the BER superfamily defining a subfamily of endo III-like BER glycosylases.

The glycosylases of the BER pathway recognize diverse types of damaged or inappropriate bases. Many questions relating to the important features of the recognition of DNA damage by these fascinating enzymes remain unanswered. For example, the mechanisms by which glycosylases recognize the damaged and mismatched bases amidst a vast excess of undamaged bases in the genome is unclear (*22*). A variety of experiments with *E. coli* MutY are beginning to provide the initial clues of important features of the enzyme and its damaged substrate that are necessary for detection of the mismatched and damaged base pair OG:A. In particular, significant differences in the rates for adenine removal have been observed depending on whether the opposite base is OG or G (*23*). Moreover, MutY has been shown to use a processive mechanism with G:A substrates, but the high affinity of MutY for the OG:AP site precludes efficient scanning of duplexes containing multiple OG:A mismatches (*24*).

Cross-linking experiments have revealed specific amino acids of MutY that are in close proximity to the DNA substrate. Specifically, mechanism-based cross-linking of MutY via OG oxidation revealed that lysine 142 is proximal to the site of oxidative damage (*25, 26*). Interestingly, Lys 142 was implicated to be in proximity to the mispaired adenine and/or abasic site product via sodium borohydride reductive trapping of a Schiff base imine intermediate (*27, 28*). Furthermore, Lys 142 does not participate in the glycosylase reaction of MutY; however, this residue forms a trappable Schiff base intermediate that is dependent on the active form of the enzyme (*28*).

A more generic approach that has been used to study protein−DNA recognition is the incorporation of photoreactive nucleobases, such as 5-iodouracil, 5-bromouracil, 8-azidoadenine, and 4-thiouracil, into DNA or RNA to produce site-specific cross-links with the protein upon irradiation with ultraviolet light (UV) (*29*−*34*). The amino acid residues of the protein in close proximity to the modified nucleobase can then be identified by subsequent degradation and analysis of the nucleic acid-peptide adduct by mass spectrometry or nucleic acid sequencing. An advantage of this type of approach to examining damage recognition by BER glycosylases is that the photoreactive nucleotide may be incorporated in a variety of positions proximal to the damaged site and therefore potentially reveal a variety of amino acids involved in damage recognition.

In the work presented herein, we have demonstrated the specific formation of MutY-DNA cross-links using the thio-modified nucleotide 4-thiothymidine (^{4S}T) incorporated adjacent to the OG:A mismatch. Additionally, the amino acid participating in the photocross-linking reaction was identified as Arg 143 by trypsin proteolysis of the complex and nucleic acid digestion of the tryptic peptide−DNA fragment, followed by positive ion electrospray ionization tandem mass spectrometry (ESI-MS/MS). Notably, Arg 143 and nearby Arg 147 are conserved throughout the endo III-like subfamily of the BER superfamily of glycosylases. Replacement of Arg 143 and Arg 147 of MutY with alanine by site-directed mutagenesis results in reduced affinity for substrate analogues and a reduced adenine glycosylase activity. This suggests that these arginine residues play an important role in DNA recognition, and this further implicates the participation of the region nucleated by the $[4Fe-4S]^{2+}$ cluster in damaged DNA recognition.

MATERIALS AND METHODS

Materials and Instrumentation. The plasmid containing the *mutY* gene, pKKYEco, and *E. coli* strains JM101 *mutY*$^-$ and GT100 *mutY*$^-$*mutM*$^-$ were kindly provided by M. L. Michaels and J. H. Miller (*11*, *35*, *36*). The 2′-deoxy-2′-fluoroadenosine (FA) phosphoramidite (*37*) was kindly provided by G. Kamilar and P. Beal (University of Utah). Standard 2′-deoxynucleotide-β-cyanoethyl phosphoramidites were purchased from Applied Biosystems, Inc. The OG and ^{4S}T phosphoramidites were purchased from Glen Research. All substrate 2′-deoxyoligonucleotides were synthesized on an Applied Biosystems, Inc. automated oligonucleotide synthesizer model 392. Substrate oligonucleotides were purified by anion exchange chromatography on a Waters AP1 DEAE 8HR column or by reverse phase chromatography on a Waters C_{18} RCM column. DNA used for PCR primers was purified using oligonucleotide purification cartridges (OPC) from Applied Biosystems, Inc. PCR reactions were carried out using a GeneAmp PCR system 2400 from Perkin-Elmer. Plasmid DNA was purified using the Wizard*Plus* MiniPrep kit from Promega. Gene sequencing was performed by automated fluorescent sequencing on a Prism 377 DNA sequencer from Applied Biosystems, Inc. at the University of Utah Medical School sequencing facility. The 5′-end labeling was performed with T4 polynucleotide kinase purchased from New England Biolabs and [γ-^{32}P] ATP from Amersham Life Sciences. Labeled oligonucleotides were purified using Probe Quant G-50 spin columns from Amersham Pharmacia. Sequencing grade trypsin, snake venom phosphodiesterase I, and alkaline phosphatase were purchased from Promega. Bovine serum albumin (BSA) and Bradford reagents were purchased from BioRad, and all PCR reagents were purchased from Roche. All other reagents were of analytical purity and purchased from Fisher or USB. Gel imaging and quantitation were performed using a Molecular Dynamics Storm 840 PhosphorImager using ImageQuant

(version 4.2a) software. Photocross-linking reactions were carried out in a Rayonet RMR-600 photoreactor from Southern New England UV Co. UV−vis spectroscopy was performed using a Hewlett-Packard 8452A diode array spectrophotometer. Electrospray mass spectrometry was performed using a C_{18} reverse phase capillary 1 mm i.d. HPLC column (GraceVydac) directly coupled to the inlet spray of an LCQ Deca ion trap spectrometer (ThermoFinnigan). Kinetic experiments were performed using a rapid-quench flow instrument (model RQF-3) from Kintek Corporation.

Site-Directed Mutagenesis, Enzyme Overexpression, and Purification. MutY and Fpg were overexpressed and purified as previously described (*18*, *38*). Site-directed mutagenesis of MutY was performed using a polymerase chain reaction (PCR) based method similar to that described previously (*35*). The mutation R143A was introduced with the following oligonucleotides: 5′-cag cgc gcc agc acg gct ttg acg tta ccg tc-3′. The R147A mutation was introduced with 5′-ctt aca gca tag cag gcc gcc agc acg cgt ttg-3′. Overexpression and purification of enzymes generated by site-directed mutagenesis did not require any changes in the standard cell culture and purification procedures used for MutY (*18*). The enzyme concentration used in cross-linking experiments was determined by the method of Bradford (*39*). The percent active enzyme concentration of R143A, R147A, and WT used in kinetic experiments and gel shift assays was calculated using the total protein concentration (Bradford) and the active enzyme concentration from the active site titration (*23*). The percent active enzyme concentration for the enzyme preparations used in this work were as follows: R143A, 55% and R147A, 18%. These percent active site concentrations for R143 are consistent with the range of values routinely observed with the WT enzyme; in the case of R147A, the percent activity was somewhat lower than seen with WT, suggesting some instability toward the purification protocol.

DNA Substrate Preparation. Oligonucleotides were prepared by automated DNA synthesis according to standard protocols. Oligonucleotides containing ^{4S}T were deprotected in concentrated NH_4OH containing 50 mM NaSH at room temperature for 24 h. DNA duplexes containing a mispair at the position X:Y were prepared after automated synthesis and ion-exchange HPLC purification of each single-stranded oligonucleotide. The following duplex was used for gel retardation and adenine glycosylase assays: (5′-CGATCATGGAGCCAC**X**AGCTCCCGTTACAG-3′)•(3′-GCTAGTACCTCGGTG**Y**TCGAGGGCAATGTC-5′), with X= 2′-deoxyguanosine (G) or 7,8-dihydro-8-oxo-2′-deoxyguanosine (OG) and Y= 2′-deoxyadenosine (A), 2′-deoxy-2′-fluoroadenosine (FA), 2′-deoxyformycin A (F), or 2′-deoxycytosine (C). The Y-containing strand was ^{32}P-5′-end-labeled and annealed to 1.5 M excess of the complementary sequence in a buffer containing 150 mM NaCl, 20 mM Tris-HCl, pH 7.6, 10 mM EDTA. Annealing was facilitated by heating the mixture to 90 °C and then slowly cooling to 25 °C for 4−6 h. For kinetic experiments, only 3−5% of the A-containing strand was 5′-end-labeled to enable accurate substrate concentration determination.

For cross-linking experiments, duplexes were used containing ^{4S}T in the following sequences:

Duplex 1:(5'CGATCATGGAGCCAC**OG**AGCTCCCGTTACAG3')•
(3'GCTAGTACCTCGGTG**A**^{4S}**T**CGAGGGCAATGTC-5');

Duplex 2: (5'-CGATCATGGAGCCAC**OG**^{4S}**T**GCTCCCGTTACAG-3')•
(3'-GCTAGTACCTCGGTG**A** ACGAGGGCAATGTC-5');

Duplex 3: (5'-CGATCATGGAGCCAA**OG**AGCTCCCGTTACAG-3')•
(3'-GCTAGTACCTCGGT^{4S}**TA**TCGAGGGCAATGTC-5');

Duplex 4: (5'-CGATCATGGAGCCA^{4S}**TOG**AGCTCCCGTTACAG-3')•
(3'-GCTAGTACCTCGGTA **A**TCGAGGGCAATGTC-5')

Analytical Scale Cross-Linking of MutY. DNA duplexes 1−4 were used for MutY photocross-linking reactions. The final reaction conditions were 10 nM DNA and 250−400 nM MutY (based on the concentration determined by Bradford). Buffer conditions were 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, and 1 mM DTT. Cross-linking reactions were incubated at 37 °C for 5 min in the dark, and 25 μL was placed onto a sheet of Parafilm, followed by direct irradiation at 256 nm for 45 min at 4 °C using a Rayonet RMR-600 reactor. Enzyme aliquots (6 μL of 400 nM MutY) were added to the reaction at 12 and 24 min and mixed with a pipet. Aliquots (5 μL) were taken directly from the cross-linking reaction and quenched in SDS loading buffer (125 mM Tris-HCl, pH 8.0, 5% SDS, 25% glycerol, 0.025% bromophenol blue). The aliquots were heated to 90 °C for 10 min prior to loading onto a 12% polyacrylamide SDS gel. After electrophoresis at 200 V for 40 min, the gel was dried and exposed to a storage phosphor screen overnight. Analytical cross-linking yields were quantitated using ImageQuant software.

Preparative Scale Cross-Linking and Isolation of the MutY-DNA Complex. The same reaction conditions were employed as for the analytical reactions, except the reaction volume of 50 mL was contained in a quartz vessel. After 12 and 24 min, 3 mL of 4−6 μM MutY was added, and the reaction was gently stirred. Each reaction was quenched by freezing at −78 °C. The experiment was repeated four to eight times to generate enough material for subsequent mass spectrometry experiments. The reaction was diluted to twice the volume with H_2O, and the cross-linked complex and unreacted protein were separated using a BioLogic FPLC system (BioRad) equipped with a HiTrap heparin affinity column (Pharmacia) using a buffer containing 20 mM sodium phosphate, pH 7.5, 0.5 mM DTT, 0.1 mM EDTA, and 20% glycerol (buffer A). The cross-linked product was collected in the flow-through of the column while the unreacted MutY eluted from the column at ∼350 mM NaCl during a linear gradient of 5−70% of 1 M NaCl in buffer A. The product fractions from each reaction were pooled and concentrated to 5−10 mL using an Amicon concentrator. The combined reactions were lyophilized to ∼1 mL aliquots containing 1−3 nmol of the MutY-DNA complex.

In-Gel Digestion. Five Eppendorf tubes containing the isolated MutY-DNA complex (∼2.5 nmol per tube) were dissolved in 250 μL of water and 50 μL of SDS loading buffer prior to heating for 10 min at 90 °C. The sample was loaded onto a 1.5 mm × 7 cm × 10 cm SDS gel in 100 μL aliquots per well and electrophoresed at 150 V for 1 h. The gel was copper-stained (BioRad) to visualize the complex.

The bands across the gel were excised and destained as per manufacturer's protocol in six Eppendorf tubes, and an additional gel control slice as large as the protein band was also excised. An in-gel digest procedure was employed as described (*40*). After crushing the gel slices with an ethanol-rinsed spatula, 700−800 μL of buffer/acetonitrile solution (50% 25 mM NH_4HCO_3, pH 8: 50% acetonitrile) was added to cover the gel pieces, and the tubes were incubated with shaking for 10 min at room temperature. The supernatant was removed; this step was repeated 3 times until the wash was colorless. The pieces were dried for 40 min in vacuo, then 25 mM NH_4HCO_3, pH 8 (∼800 μL), was added to each tube of dried gel pieces. Sequencing grade trypsin (Promega) was freshly diluted to 0.5 mg/mL with dilution buffer (provided by the manufacturer), and 2 μg was added to each tube. The tubes were incubated with shaking at 30 °C overnight, and 1.5 μg more of trypsin was added after 12 h. The protease/MutY ratio was 1:20 after the second addition of trypsin. The reaction was quenched after 40 h by adding 50 μL of 1% TFA. The supernatant was removed, and 600 μL of 60% acetonitrile/0.01% TFA was added to each tube, followed by shaking at 30 °C for 20 min. The supernatant was removed and combined with the first fraction, and this step was repeated two more times to extract the peptides. The fragments were dried in vacuo.

Nucleotide Digestion. The nucleotide digestion reactions were performed analogously to that described previously (*41*). Lyophilized snake venom phosphodiesterase I (Promega) (SVP) was resuspended in dilution buffer (110 mM Tris-HCl, pH 8.9, 110 mM NaCl, 15 mM $MgCl_2$, and 50% glycerol in H_2O) to a concentration of 1 mg/mL. The cross-linked peptide (∼240 pmol) was diluted with H_2O in a buffer containing 350 mM Tris-HCl, pH 7.5 and 15 μM $MgCl_2$. Diluted SVP (16 μL) was added for a total volume of 200 μL. The reaction was incubated for 16 h at 37 °C, and the SVP was inactivated by heating at 90 °C for 5 min. Buffer (20 μL) provided by the manufacturer for the alkaline phosphatase (Promega) was added after the reaction cooled. Alkaline phosphatase was freshly diluted (1:20) with alkaline phosphatase buffer, and 8 μL was added to the SVP-digested reaction mixture followed by incubation at 37 °C for 30 min. The reaction was dried in vacuo. As a control, 240 pmol of duplex 1 was subjected to identical conditions.

Mass Spectrometry Analysis. LC-MS experiments were performed on a Finnigan LCQ Deca ion trap mass spectrophotometer equipped with a nebulization-assisted electrospray ionization (ESI) source. The buffer gas inside the ion trap was ultrapure helium gas (40 psi). The microcapillary HPLC system was equilibrated at 40 μL/min with 0.1% formic acid and was directly coupled to the mass spectrometer electrospray interface. The digested cross-linked product (30−90 pmol) was diluted to 10 μL with 10% acetonitrile/0.1% TFA followed by centrifugation, and the sample was injected onto the column. The peptides were eluted with acetonitrile containing 0.08% formic acid with a gradient of 0−90% in 90 min. Typical conditions utilized a capillary voltage of 5 V at 275 °C. The activation time for MS/MS was 60−80 ms. Xcalibur software was used to operate the instrument and analyze the data. Using LC-MS only, the identification of a possible cross-linked species was facilitated by comparison of the spectra resulting from the chromatograph containing WT peptides, in addition to the sample containing unmodified digested DNA. The sample for the MS/MS experiments was from the same reaction that gave the spectrum illustrating the parent ion of interest.

Equilibrium-Dissociation Constant (K_d) Determination. Gel retardation assays (*42*) were performed as described previously (*43*). In the case of reactions with substrate duplexes containing OG, the final concentration of the duplex was 10 pM. With all other duplexes, the final concentration of the duplex was 20 pM. Dissociation constants, K_d, were determined by fitting the data from a storage phosphor autoradiogram (percent bound substrate vs log[MutY]) with the equation for one-site ligand binding using Grafit version 4.06 (Erithacus). The reported K_d values are corrected for the activity of the enzyme as determined by the active site titration method (*23*). In addition, each reported value is an average of at least three separate experiments.

Adenine Glycosylase Assays. The adenine glycosylase activity was monitored under conditions of single- and multiple-turnover as described in detail previously (*18*, *23*). In the case where the rate of excision was too fast for a manual assay, a KinTek Rapid Quench instrument was employed. Briefly, reactions were conducted at 37 °C regulated by a circulator with a digital controller (VWR). Samples (24 μL) containing 40 nM DNA, 40 mM Tris-HCl, pH 7.5, 20 mM EDTA, and 60 mM NaCl were rapidly mixed with equal amounts of samples containing 80−160 nM MutY, 25 nM nonspecific DNA duplex, 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, and 20% glycerol. At various time points raging from 0.8 s to 5 min, reactions were quenched with 0.5 M NaOH. Further sample handling was conducted in the same fashion as in our manual assays. The multiple-turnover experiments were performed in buffer solutions containing 20 nM OG:A or G:A duplex DNA and the appropriate amount of active MutY necessary to provide a burst amplitude corresponding to 5−10% product formation. The single-turnover experiments were performed by incubation of 20 nM of duplex DNA with 30 nM of active protein. For all kinetic experiments, the samples were analyzed by 15% denaturing PAGE. The gels were exposed to a storage phosphor screen for at least 6 h. The resulting storage phosphor autoradiogram was quantified using ImageQuant V4.2a. The data were fit using GraFit 4.0 with the relevant equations (*18*, *23*) to extract values for the k_2 and k_3 rate constants.

RESULTS

Photocross-Linking to Probe the DNA-MutY Interface. MutY is highly specific for duplex DNA containing an OG:A mispair and forms a tight DNA−protein complex as indicated by measured equilibrium dissociation constants (100−400 pM) with substrate analogues (*43*−*45*). Because of the absence of a structure of MutY bound to a mismatch-containing duplex, there is limited knowledge about contacts of MutY with specific nucleotides and the structure of the DNA duplex in the MutY-DNA complex. Thus, the positioning of a photoreactive nucleotide will likely have a dramatic influence on the extent of MutY-DNA cross-linking. To determine optimal positioning of the photoreactive probe, the photoreactive nucleotide 4-thiothymidine (^{4S}T) was incorporated at a variety of positions adjacent to the OG:A mismatch in a 30-bp duplex (Figures 1 and 2) via automated DNA synthesis.

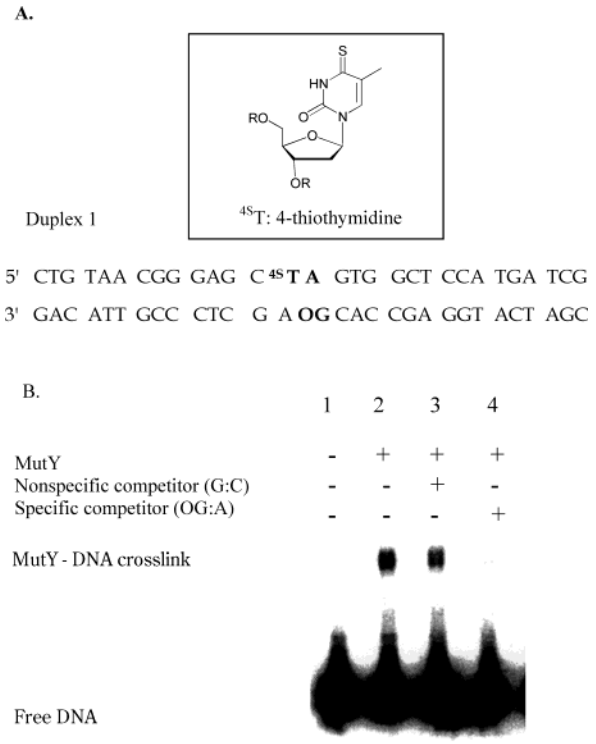

FIGURE 1: (A) Photoreactive nucleotide 4-thiothymidine (^{4S}T). The 30-base pair substrate used (duplex 1) contains ^{4S}T on the 5′-side of the OG:A mispair. (B) SDS−PAGE analysis of competition assay. The duplex and enzyme used in each assay are denoted at the top. 10 nM DNA and 250 nM MutY in the presence or absence of competitor DNA were incubated at 37 °C for 5′ followed by irradiation at 256 nm for 40′. Lane 1, control cross-linking reaction without MutY; lane 2, reaction without competitor DNA; lane 3, reaction with a nonspecific competitor duplex added; and lane 4, reaction with a specific competitor duplex added.

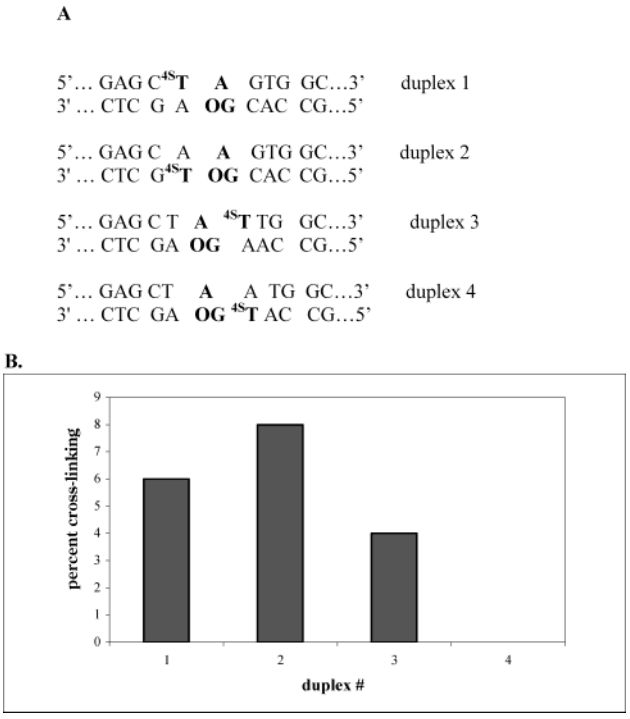

FIGURE 2: (A) Location of the ^{4S}T relative to the OG:A mismatch in DNA duplexes used in cross-linking experiments with MutY. (B) A graphical representation of the percent cross-linking efficiency for each duplex.

To initiate photocross-linking, MutY was incubated in 20−30-fold excess with ^{32}P-labeled duplex 1, and the reaction was subject to UV light in a photoreactor at different exposure times. Although thionucleobases absorb light at wavelengths between 330 and 360 nm (*30*), MutY did not form a covalent complex with duplex 1 at 350 nm during irradiation times of up to 2 h. Cross-linked species were formed in a 5−10% yield only at 256 nm. The inability to photocross-link at 350 nm may be related to inner filter effects from the broad absorption features centered at 410 nm arising from thiolate-to-Fe(III) charge-transfer transitions of the $[4Fe-4S]^{2+}$ cluster present in MutY. To test this idea, a duplex containing ^{4S}T was prepared adjacent to an OG:C mismatch, which is the primary substrate for Fpg, a BER glycosylase that does not contain a $[4Fe-4S]^{2+}$ cluster (*7*). Fpg was incubated with the photoreactive OG:C-containing duplex and subjected to different irradiation wavelengths and compared to the results with MutY. ^{4S}T-containing duplex DNA forms a covalent complex with Fpg at 350 nm with comparable efficiency to the reaction with MutY but not at 256 nm (data not shown). Conversely, MutY formed a DNA cross-link with duplex 1 only during irradiation at 256 nm (Figure 1B), and cross-link formation is dependent upon ^{4S}T incorporation into the substrate.

A competition assay was performed to determine if the DNA-MutY complex formed by photocross-linking represents a specific recognition event. While MutY has a strong affinity for a duplex containing an OG:A mispair (K_d < 0.200 nM), it does not exhibit as high affinity to an undamaged duplex of similar sequence (K_d ∼150 nM). Thus, unlabeled DNA duplexes containing a central OG:A mispair or a duplex without a mispair (central G:C) can act as specific or nonspecific competitors for the cross-linking reaction. MutY was incubated with ^{32}P-labeled duplex 1 followed by the addition of a 20-fold excess of unlabeled specific or nonspecific competitor duplex. The autoradiogram of the competition reactions after SDS−PAGE is shown in Figure 1B. In the presence of a nonspecific competitor, the efficiency of the cross-linking reaction was not affected. However, addition of excess unlabeled OG:A-containing duplex diminished the cross-linking reaction, presumably by occupying the specific binding sites of MutY. This demonstrates that the interaction of MutY with the DNA is specific, and the cross-linking reaction forms a covalent site on the protein that is proximal to a site important for recognition and DNA binding or catalysis since the location of the photoreactive nucleotide is adjacent to the mismatch. Notably, since the reaction with MutY and an OG:A-containing substrate under these conditions is very fast (*23*), cross-linking with duplex 1 is to the OG:(AP site) product of the reaction. PAGE analysis of cross-linked samples that were treated with base to cleave any AP site product indicated full conversion to product within the incubation time prior to initiating photocross-linking (20 min).

The effect of varying the position of ^{4S}T relative to the mispair (Figure 2A) on the extent of cross-link formation is shown in Figure 2B. Interestingly, there is no covalent complex formed between duplex 4 and MutY, suggesting that ^{4S}T in this position (on the 5′-side of OG) is not sufficiently close to any amino acid of the enzyme to react in the photochemically mediated process. There is a 2-fold

difference in the cross-linking efficiency between duplexes 1−3, and it is anticipated that each position could yield a covalent adduct with MutY via different amino acids.

On the basis of the relatively high amount of MutY-DNA cross-link formed with duplex 1, we focused on this particular duplex for further analysis. The substitution of ^{4S}T for T at the position neighboring the mismatch was expected to cause minimal distortion to the DNA helix and the subsequent interaction of MutY with the mismatch containing duplex. However, to determine if there were any effects due to sulfur substitution in the major groove, the adenine glycosylase activity of MutY was evaluated with duplex 1. Under single-turnover conditions using rapid-quench methods, duplex 1 was converted to the AP site product with the same efficiency ($k_2 = 12 \pm 1$ min^{-1}) as the corresponding unmodified OG:A-containing substrate ($k_2 = 12 \pm 2$ min^{-1}). Additionally, DNA binding experiments were performed with an analogous duplex containing ^{4S}T on the 5′-side of the substrate analogue 2′-deoxyformycin A (F). The 2′-deoxyformycin-β-cyanoethyl phosphoramidite was prepared as described previously and incorporated adjacent to ^{4S}T in a 30-nucleotide strand (duplex 1F) (*46*). The affinity of MutY for duplex 1F ($K_d = 0.14 \pm 0.08$ nM) is similar to that obtained for the corresponding OG:F-containing duplex lacking the ^{4S}T nucleotide ($K_d = 0.28 \pm 0.05$ nM). These results show that the incorporation of ^{4S}T has not altered the ability of MutY to recognize an OG:A mispair and remove the mispaired adenine.

Preparative Cross-Linking Experiments. The cross-linking reaction with duplex 1 was scaled up to yield a sufficient amount of covalent complex for enzymatic digestion and subsequent analysis by mass spectrometry and peptide identification by MS/MS. The preparative scale cross-linking experiments contained the same protein and DNA concentrations used for the analytical reactions in a volume of 50 mL, keeping the same conditions to ensure specific formation of a DNA−protein cross-link. The reaction was repeated six to eight times to generate enough material for LC-MS/MS experiments. A schematic representation of the cross-linking reaction and the route to simplify its identification is illustrated in Figure 3.

The cross-linking yield was low (5−10%); thus, unreacted protein and DNA were removed from the covalent complex using heparin affinity chromatography to increase the possibility of detecting the cross-linked peptide among the numerous unmodified peptides. Indeed, it was expected that the detection of the cross-linked peptide would be facilitated by chromatographic separation; however, it is possible for an unmodified peptide and a peptide cross-linked to a nucleotide to coelute or have minimal reverse-phase separation (*47*). In anticipation of sample loss from a variety of sources, preparation of samples for LC-MS/MS analysis was performed on 3−5 nmol of the isolated, unlabeled DNA-MutY complex. The sample was then subject to proteolysis by in-gel digestion procedures (*40*). Our lab and others have observed that sites normally susceptible to enzymatic cleavage are inefficiently processed in the presence of the covalent attachment to DNA (*27*). Thus, the DNA-MutY complex was digested with trypsin for 40 h. Following quenching of the proteolysis reaction and isolation of the peptides from the gel, the sample was digested with snake venom phosphodiesterase and alkaline phosphatase to reduce the size of the attached oligonucleotide fragment and provide free hydroxyl ends.

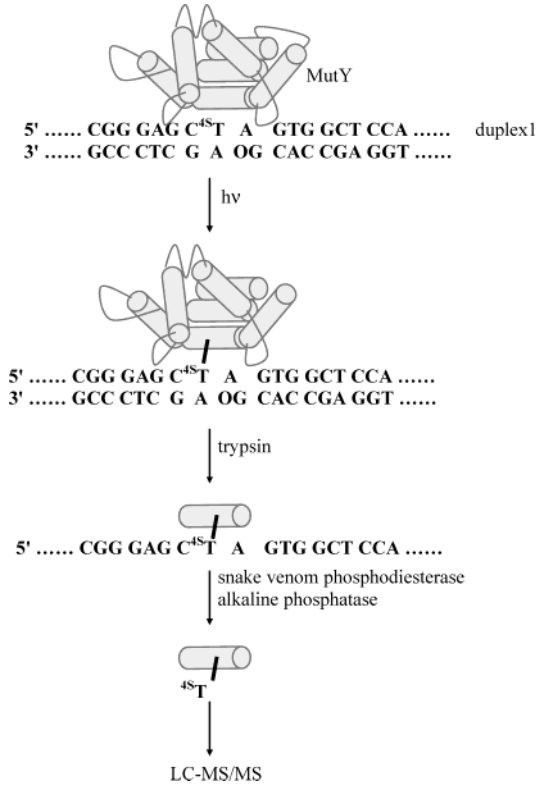

FIGURE 3: Photochemical cross-linking scheme with a ^{4S}T-containing DNA substrate. The unreacted MutY was separated from the reaction by heparin affinity chromatography. The next step includes in-gel digestion of MutY and recovery of the peptides. Additionally, the DNA was digested with snake venom phosphodiesterase and alkaline phosphatase.

LC-MS/MS Analysis. The digested sample was split into three portions for various types of mass spectral characterization. Approximately 30−90 pmol of the peptide mixture was analyzed using a C_{18} reverse phase capillary HPLC column directly coupled to the inlet spray of an ion-trap mass spectrometer operating in the positive mode. The corresponding spectra from the total ion chromatogram were dissected to find an ion corresponding to the peptide−nucleotide complex. On the basis of the proposed radical mechanism for ^{4S}T-mediated cross-linking (*48*), it was not expected that a net change would occur in the molecular weight of the covalent complex. A portion of the spectrum at the retention time of 57.9 min identified a unique doubly charged ion and its corresponding singly charged species. These ions *m*/*z* 582.5 and 1163.5 identified in this region correspond to a mass of 1162.5 and are unique from the WT peptide ions, nucleotide fragments, or other species from the multiple enzymes used in the reactions (Figure 4).

By analysis of all possible molecular weights of tryptic fragments plus ^{4S}T-containing nucleotides, the initial identification of the modified peptide was determined. The initial data suggested that the tryptic peptide (K)RVLAR, residues 143−147 in MutY, was covalently attached to a dinucleotide (C^{4S}T) from the ions at *m*/*z* 582.5 and 1163.5, corresponding to the doubly and singly charged species of the complex, respectively (±1.0 Da). We suggest that steric interference due to the attachment of the peptide to the ^{4S}T nucleotide

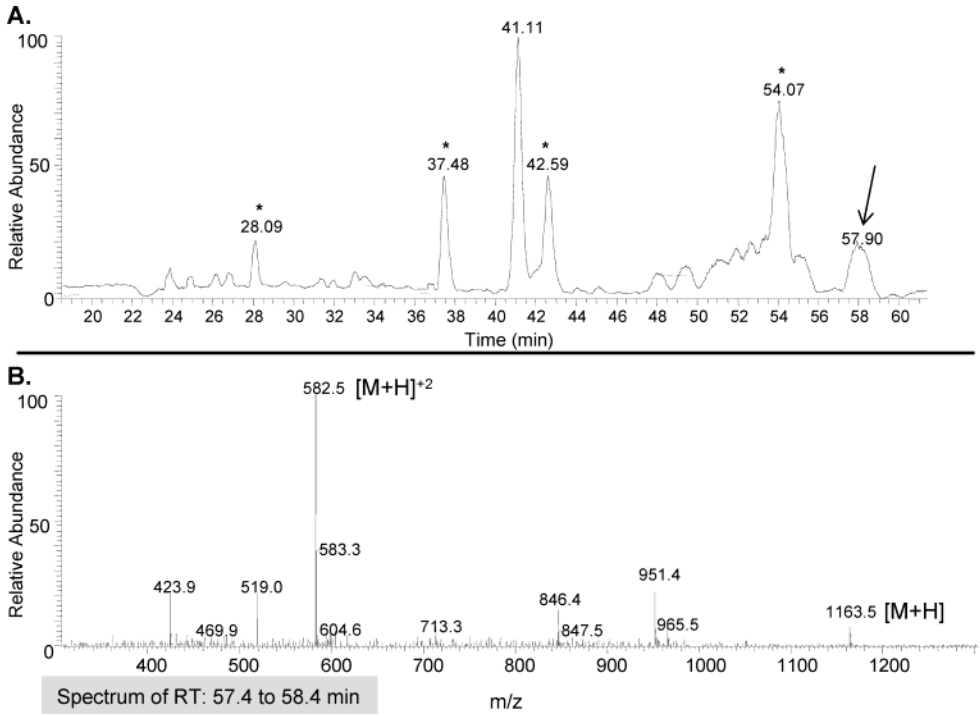

FIGURE 4: LC-ESI-MS of the MutY peptide−nucleotide mixture. (A) An expanded region of the total ion chromatograph (TIC). The asterisks correspond to other MutY tryptic fragments (28.08 min, aa 236−241; 37.48 min, aa 130−142; 42.59 min, aa 31−38; and 54.07 min, trypsin autolysis fragment). The peak at 41.1 min could not be identified. The arrow corresponds to the mass spectrum illustrated in panel B. The ions at *m*/*z* 582.5 and 1163.5 are not present in the chromatograph of the WT sample or the enzymatically digested DNA.

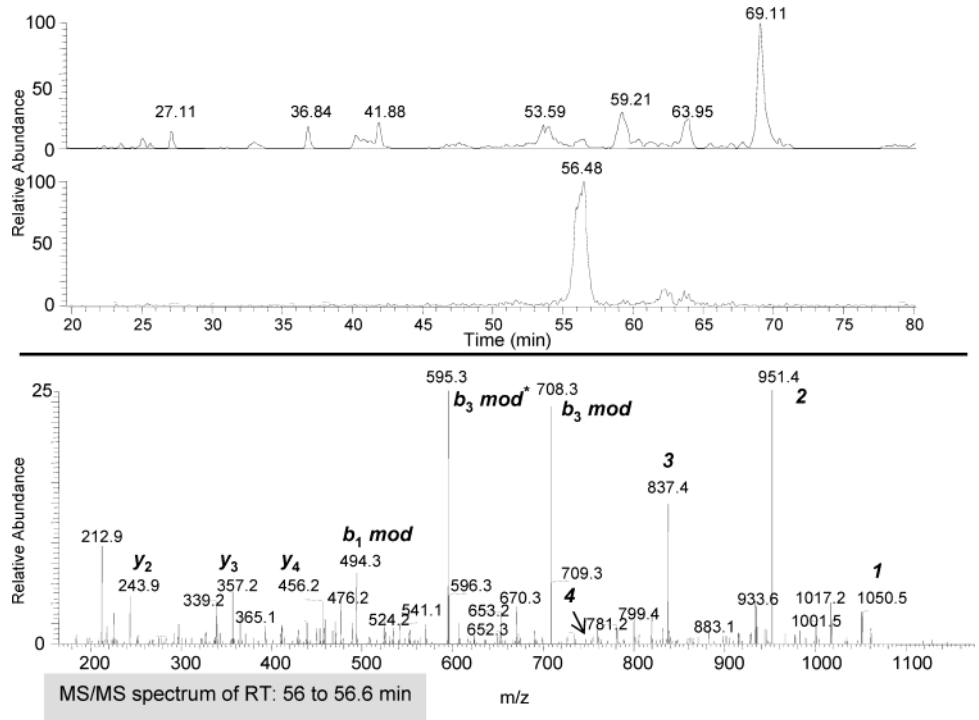

FIGURE 5: LC-ESI-MS/MS of the MutY peptide−nucleotide mixture. (Top panel) Total ion chromatograph for all scan ranges between *m*/*z* = 165−1920. The second panel is the TIC of the MS/MS experiment with collision energy focused on *m*/*z* = 582.5. The bottom panel shows the full MS/MS spectrum of the doubly charged parent ion *m*/*z* = 582.5. The predicted *b* and *y* fragment ions are indicated as well as fragments with modifications (*b*- and *y*-*mod*). The proposed structures of ions 1−4 are the resulting dissociation products of the dinucleotide attached to the peptide.

may result in inefficient reaction by the otherwise vigorous exonuclease, snake venom phosphodiesterase. Additionally, the tryptic peptide RVLAR normally elutes at 52 min from a digestion mixture of unmodified MutY under similar conditions, and this peptide is not present in the chromatogram of the cross-linked MutY, consistent with this peptide assignment. This ion, *m*/*z* 582.5, was the candidate for MS/MS analysis with the remaining cross-linked peptide sample (Figure 5). Indeed, the LC-MS analysis did not reveal any other reasonable ions corresponding to a molecular weight of a tryptic peptide in MutY with a mono-, di-, or trinucleotide.

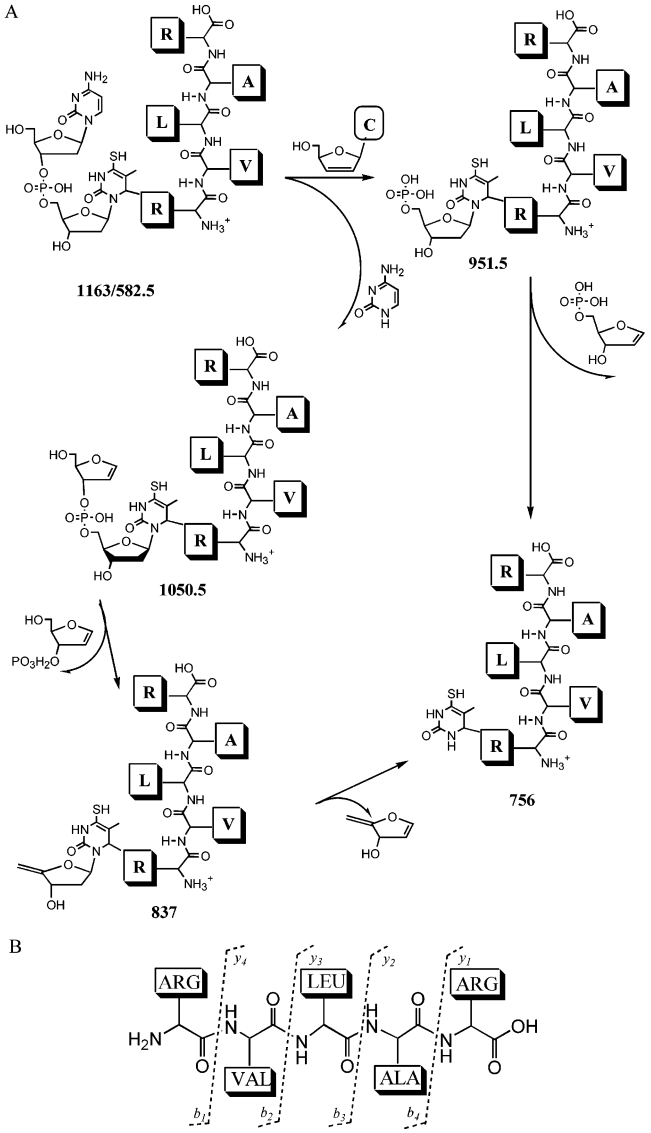

FIGURE 6: (A) Proposed structures for the fragmentation of the MutY peptide-dinucleotide based on the spectrum from the LC-ESI(+) MS/MS analysis. (B) Predicted fragmentation of the MutY tryptic peptide RVLAR.

Ions 1−4 are consistent with the fragmentation of the dinucleotide covalently attached to the peptide RVLAR. The proposed dissociation pattern (Figure 6A) is analogous to the observed elimination reactions by ESI(−) MS/MS sequencing of oligonucleotides (*41*, *49*). Ion 1 ($m/z = 1050$) results from the depyrimidation of C by elimination, and 2 ($m/z = 951.4$) represents the molecular mass of ^{4S}T-nucleotide attached to RVLAR. Loss of the deoxyribose phosphate from 1 by δ-elimination would produce ion 3 ($-NH_3$) ($m/z = 837.4$). Finally, 4 is the $[M + H]^+$ corresponding to the ^{4S}T base attached to RVLAR, the product of 1,2 elimination of the base.

The ions of the predicted fragmentation (*50*) for RVLAR are illustrated in Figure 6B, many of which are observed in the MS/MS spectrum. The *y*-ions, (y_2-y_4) are clearly observed; however, y_1 (173.1) is not within the *m*/*z* scanning range of the experiment. The only observed *b*-ions include the calculated fragment ions plus ^{4S}T- nucleotide appendages. The ion b_3 mod ($m/z = 708.3$) is b_3 with the addition of a ^{4S}T-nucleotide appendage ($b_3 + 336.2 = 707$). In addition,

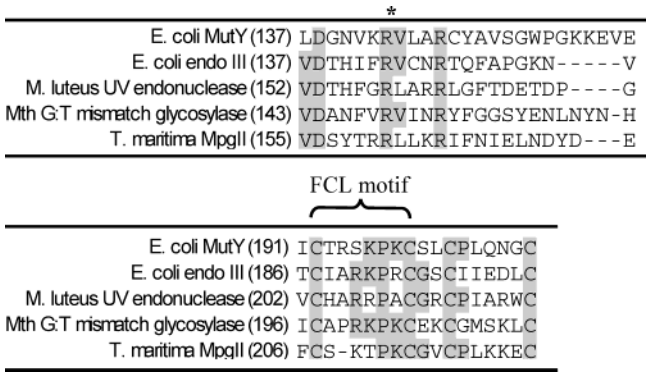

FIGURE 7: Partial sequence alignment of the endo III-like subfamily of the BER superfamily illustrating the new motif, R(V/L)XXR. In addition, the alignment of the FCL motif is shown. Shaded regions represent similar or identical amino acids. The arginine identified via cross-linking is highlighted with an asterisk.

b_3 is also modified with the ^{4S}T-reduced-sugar fragment ($b_3 + 224.3 = 594$). The ion b_1 mod ($m/z = 494.3$) is also modified by the addition of a ^{4S}T-nucleotide appendage. The modifications to the b_3 fragment implicate that Arg 143, Val 144, or Leu 145 could bear the covalent attachment. The modification to the b_1 fragment, however, provides evidence that Arg 143 participated in the cross-linking reaction. This reasoning is further supported by the *y*-ion fragmentions that do not have modifications, indicated by each predicted *m*/*z* ion in the spectrum. Thus, LC-MS/MS revealed that the specific photocross-linking reaction between MutY and ^{4S}T-containing DNA proceeded via Arg 143, and this amino acid is positioned proximal to the OG:A mispair.

Functional Consequences of Replacement of Arg 143 and Arg 147 with Alanine. Cross-linking of the ^{4S}T nucleotide to arginine 143 reveals its proximal location near the ^{4S}T nucleotide and the OG:A mismatch. This placement of Arg 143 near the scene for glycosylase action suggests a potentially important role for Arg 143 in recognition and repair of OG/G:A mispairs by MutY. A BLAST sequence alignment of the BER superfamily revealed that this arginine, as well as a second closely spaced arginine (Arg 147), are strictly conserved in the BER superfamily glycosylases containing a $[4Fe-4S]^{2+}$ cluster, forming a consensus sequence R(V/L)XXR (Figure 7). Replacements of arginines 143 and 147 with alanine by site-directed mutagenesis generated R143A and R147A mutant forms. The modified enzymes were overexpressed and purified in a manner analogous to that of the WT enzyme (*18*). Cross-linking reactions were carried out with R143A using similar conditions to that of WT reactions. The cross-linking yield with the R143A mutant was approximately one-half of the reaction yield obtained from cross-linking with WT MutY (2−3%, data not shown); however, this indicates that the mutant may form a cross-link via a different amino acid. In instances where the cross-linking agent is specific for a particular amino acid (e.g., 5-bromouridine and 5-iodoridine), alanine mutagenesis usually results in abolition of the cross-linking reaction (*29*, *51*). However, ^{4S}T cross-linking appears to be less specific for particular amino acids (*48*); therefore, the decreased cross-linking with R143A may be due to diminished binding affinity of the mutant enzyme.

The R143A enzyme was further characterized by its ability to bind to DNA containing the substrate analogue 2′-deoxy-

Table 1: Dissociation Constants (K_d) and Kinetic Constants (k_2) for WT and R143A for 30 Base-Pair DNA Duplexes[a]

	K_d (nM)			k_2 (min^{-1})		k_3 (min^{-1})	
enzyme	OG:FA[b]	G:FA[b]	G:C[b]	OG:A[b]	G:A[b]	OG:A[b]	G:A[b]
WT[c]	0.12 ± 0.05	5.8 ± 0.6	150 ± 60	12 ± 2	1.80 ± 0.05	0.003 ± 0.001	0.03 ± 0.006
R143A	2.0 ± 0.8	32 ± 15	90 ± 20	12 ± 1	0.31 ± 0.08	0.002 ± 0.001	0.019 ± 0.006
R147A	12 ± 2	44 ± 7	nd[d]	8.1 ± 0.3	0.32 ± 0.05	0.002 ± 0.001	0.003 ± 0.002

[a] Dissociation constants were measured at 25 °C, while the kinetic parameters k_2 and k_3 were measured at 37 °C. Errors reported are the standard deviation of the average. [b] Designates central base-pair in 30 bp duplex (sequence 1 lacking ^{4S}T). [c] K_d values listed are from Chepanoske et al. (*43*). [d] Not determined.

2′-fluoroadenosine (FA) opposite G and OG, as well as nonspecific DNA. Nondenaturing gel retardation experiments were used to determine relative dissociation constants (K_d) (*42*). The results for each duplex are listed in Table 1 and are compared to K_d values previously reported for the WT using similar reaction conditions (*43*). The affinity of R143A for duplexes containing OG:FA ($K_d = 2.0 \pm 0.8$ nM) was decreased approximately 15-fold relative to the WT enzyme. In addition, R143A also exhibited a 5-fold decrease in affinity for the corresponding G:FA-containing duplex ($K_d = 32 \pm 15$ nM), while its affinity for nonspecific DNA remained within error of the WT. These results are similar to the DNA binding capabilities observed for mutated forms of MutY in which positively charged residues (lysine and arginine) within the FCL and the active site cleft were replaced with alanine (*21*).

Previous work in our laboratory has shown that MutY displays biphasic kinetic behavior due to slow release of the DNA product, and based on this behavior, the following minimal kinetic scheme (Scheme 1) has been used to analyze the glycosylase activity of MutY.

Scheme 1

$$E + (DNA)_S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E\bullet(DNA)_S \xrightarrow{k_2} E\bullet(DNA)_P \xrightarrow{k_3} (DNA)_P + E$$

Using single-turnover (STO) experiments ([MutY] > [DNA]) and multiple-turnover experiments ([MutY] < [DNA]), the rate constant k_2, which includes the step involving chemistry, and the rate constant k_3, which is characteristic of the product release step, may be measured (*23*). The R143A enzyme exhibited biphasic behavior similar to the WT enzyme allowing for a similar analysis to be used. The observed rate constants (k_2 and k_3) of the reaction of each enzyme with both OG:A- and G:A-containing substrates revealed subtle differences between R143A and the WT enzyme (Table 1). In particular, the R143A enzyme exhibited a 5-fold reduced rate for adenine removal from a G:A substrate. However, somewhat surprisingly, under single-turnover conditions, the reaction of R143A with an OG:A substrate was similar to that for the WT enzyme (Table 1). The Arg-to-Ala mutations do not appear to affect product release rates (k_3) with OG:A or G:A substrates (Table 1) or decrease the stability of the enzyme. Taken together with its identification in the cross-linking reaction, the binding and kinetic data suggest that Arg 143 is involved in the recognition and removal of adenine in both mispairs by MutY.

The identification of Arg 143 via the cross-linking reaction brought to our attention the highly conserved R(V/L)XXR motif; therefore, we also investigated the functional consequences of replacement of Arg 147 with alanine. A comparison of the K_d values and the kinetic parameters of R147A relative to R143A and the WT enzyme was quite revealing. Indeed, the R147A exhibited significantly decreased affinity for both the OG:FA-containing and the G:FA-containing duplexes. The loss of affinity for the OG:FA duplex by the R147A enzyme relative to the WT enzyme is approximately 100-fold. This indicates a significant loss of discrimination between OG and G for the R147A enzyme relative to the WT enzyme. The hampered activity of R147A is also apparent in the measured kinetic parameters. The rate of adenine removal from the G:A substrate by the R147A enzyme is significantly reduced (approximately 6-fold relative to WT), similar in magnitude to that observed with the R143A enzyme. However, we also observe a decrease in the intrinsic rate of adenine removal (k_2) with the OG:A substrate with the R147A enzyme, which we did not observe with the R143A enzyme. This is consistent with the more drastic effect observed in the K_d value. Of particular note is that in our analysis of many mutated forms of MutY, we rarely detect a significant decrease in the rate of processing of OG:A substrates under single-turnover conditions at 37 °C (*52*). We believe that this is a consequence of the high concentrations of enzyme and DNA used in the single-turnover analysis that compensates for any deficiencies due to inefficent mismatch recognition. The R147A enzyme was also unusual in that it exhibited a decreased rate for product release (k_3) with the G:A substrate. We ruled out a possible artificially low number due to instability of the enzyme by evaluating the active-site enzyme concentration under the incubation conditions of the assay. These results indicated that the R147A enzyme is stable to the assay conditions (data not shown). Thus, the R147A mutation results in significant changes in the properties of the enzyme, and perhaps this residue is involved in a conformational change necessary for efficient recognition of both OG and A. Indeed, this mutation results in unusual changes in not only mismatch recognition and adenine removal but also product release.

DISCUSSION

Many fundamental biological processes rely on nucleic acid−protein assemblies, including DNA repair. It is possible to gain a better understanding of these processes by supplementing the information known about the structure and function of the individual components of a system with details of the way each participant interacts with one another at various stages of the process. Insight into the mechanisms of DNA repair has utilized an assortment of structural and biochemical techniques, including cross-linking. In the absence of a high-resolution structure of a repair protein bound to damaged DNA, as is the case for MutY, valuable information can be obtained about the spatial relationship

of each partner using cross-linking methods. Importantly, various types of cross-linking reactions have provided multiple avenues for determining the amino acid residues of MutY that are important for DNA recognition and catalysis. Herein, using ^{4S}T-mediated cross-linking, we identified an important amino acid that had not been previously implicated as important for damage recognition and repair by MutY.

This work illustrates the usefulness of the commercially available 4-thio-T- nucleotide in DNA protein cross-linking to identify amino acids involved in the protein−DNA interface. Interestingly, we observed that the extent of cross-linking is highly dependent on the placement of the ^{4S}T nucleotide relative to the OG:A mismatch in the duplex. In all cases, the ^{4S}T nucleotide was positioned next to either the OG or A in the mismatch; however, in the case of duplex 4, which contained the ^{4S}T flanking the 5′-side of the OG, no covalent cross-link was detected. This is consistent with considerable distortion of the DNA proximal to the mismatch upon binding of MutY. Indeed, we have previously observed that a G residue on the 3′-side of the adenine in the same 30-base pair duplex used in these experiments was hypersensitive to reaction with DMS (*43*). In addition, fluorescence experiments indicate that MutY may use a double-flip mechanism and extrude out both A and OG into base-specific pockets (*53*). Identification of residues involved in cross-linking to ^{4S}T in duplexes 2 and 3 may provide further insight into the features of OG:A mismatch recognition by MutY.

Arginine 143 is strictly conserved in the BER superfamily of glycosylases containing a $[4Fe-4S]^{2+}$ cluster in a consensus sequence R(V/L)XXR (Figure 7). The X-ray crystal structures of two enzymes in this subfamily, *E. coli* endo III and MutY, reveal that this region is included in the C-terminal end of an α-helix perpendicular to the HhH motif. These two residues are adjacent to the $[4Fe-4S]^{2+}$ cluster, positioned near the iron−sulfur cluster loop (FCL) motif (*19*, *20*). The structure of endo III suggests that Arg 143 and Arg 147 (same sequence numbers in MutY) help stabilize the FCL of endo III by hydrogen bonding interactions (*20*). Specifically, Arg 143 NϵH of endo III donates a hydrogen bond to the backbone carbonyl of the first coordinating cysteine (Cys 187) in the FCL. The structure of the catalytic domain of MutY also reveals a close proximity of Arg143 NϵH to the backbone carbonyl of Cys 192 (2.75 Å), indicating a similar interaction.

Inspection of the structure of the catalytic domain of MutY suggests that residues surrounding the $[4Fe-4S]^{2+}$ cluster in MutY are not exceptionally close to the active site containing catalytic residues (e.g., Asp 138); however, previous results illustrate that the positively charged residues in the FCL, namely, Lys 198, are intimately involved in DNA damage recognition and repair (*21*). It is likely that an ensemble of interactions between MutY and oxidatively damaged DNA are necessary for efficient recognition and removal of adenine. In the absence of an X-ray crystal structure of the full-length MutY in complex with duplex DNA, it may be difficult to predict interactions by residues of MutY that are not located in the implicated active site. Importantly, there is a recent structure of a covalently trapped complex between *B. st.* endo III and a DNA product intermediate (*54*). Grafting of the structure of the N-terminal domain of MutY in the place of the *B. st.* endo III in the complex structure (Figure 8A) indicates that both Arg 143 and Arg 147 lie in close proximity to the DNA backbone (2.3 Å). In addition, Arg 143 is in reasonable proximity (∼7 Å) to the base in the position corresponding to ^{4S}T in this modeled structure (Figure 8A). Thus, the modeled structure is consistent with an important role for Arg 143 and Arg 147 in damaged DNA recognition. It is somewhat surprising that Arg 143 cross-links to the ^{4S}T base considering its close interaction with the phosphate backbone in the endo III−DNA structure. It is possible that this amino acid is positioned closer to the DNA base in the MutY-DNA structure. Notably, in the *B. st.* endo III−DNA structure, the FCL motif is not well-defined; thus, there may be some differences in this region of the structure.

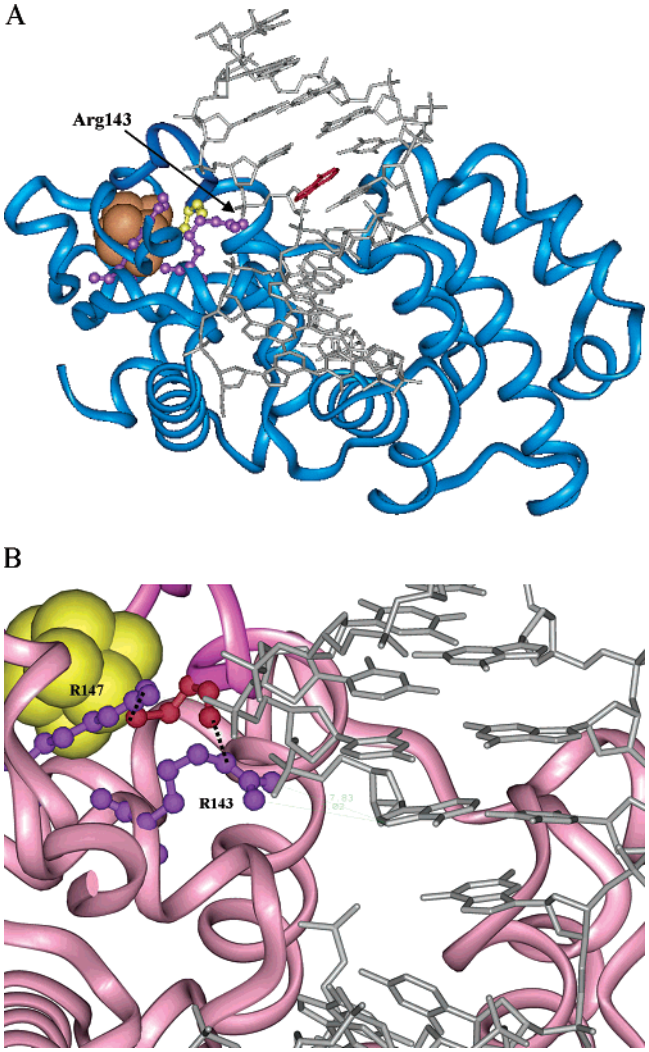

FIGURE 8: (A) Model of truncated form of MutY bound to DNA based on trapped endonuclease III-DNA structure (*54*). The base in the endo III structure corresponding to the position of ^{4S}T is red. The two arginines (143 and 147) are purple, while Cys 192 is yellow. The $[4Fe-4S]^{2+}$ cluster is gold. (B) A detailed view of the region surrounding the $[4Fe-4S]^{2+}$cluster (yellow) and the interaction of Arg 143 and Arg 147 (purple) with Cys 192 (red). For simplification, only Cys 192 is shown ligating an iron in the cluster.

In the endo III−DNA structure, both Arg 147 and Arg 143 make hydrogen-bonding contacts with a ligating cysteine (Cys 192) of the $[4Fe-4S]^{2+}$ cluster (Figure 8B). Inspection of the structures of the catalytic domain of MutY alone and in the presence of an adenine base revealed a change in the

hydrogen-bonding interaction of Arg 143 with Cys 192. In the structure of MutY alone, Arg 143 NϵH provides the hydrogen for an H-bond to the backbone carbonyl of Cys 192, and this interaction is analogous to that observed in endo III. In the presence of the adenine base, the side chain of Arg 143 has moved such that hydrogen bonding takes place between the cysteine carbonyl and Arg 143 NηH. Significant differences in position or hydrogen-bonding interactions of other amino acid side chains between the free and the bound structures were not observed. This subtle change could represent the initiation of a larger, global conformational change in MutY that also includes residues in the FCL. Thus, a disruption of the network of interactions in this region may be responsible, in part, for the reduced glycosylase activity of the R143A and R147A enzymes as compared to the WT. Additionally, results from site-directed mutagenesis of the cysteine ligands to the cluster, specifically with the C192H enzyme, suggest the importance of Cys 192 in MutY (*35*). Using an in vivo complementation assay, the C192H enzyme was shown to be unable to prevent DNA mutations in vivo. This result led to the proposal that coordination of the histidine at this position to the cluster may modify the FCL and therefore interfere with substrate recognition.

In summary, site-specific photochemical cross-linking was successfully employed to investigate the interaction of MutY and substrate DNA. This is the first example of site-specific photocross-linking to a BER glycosylase, and the identification of Arg 143 illustrated the powerful technique of LC-coupled tandem mass spectrometry to reveal a peptide modification from a complex mixture. This result is complementary to those derived from mechanism-based cross-linking of MutY via oxidation of OG, suggesting that Lys 142 is proximal to the site of oxidative damage in the mispair. The successful identification of the cross-link using duplex 1 has paved the way for future cross-linking experiments with substrate and product analogues, in addition to cross-linking at other sites and/or with other photoreactive nucleotides. This approach will contribute to identifying various residues important for recognition and catalysis by MutY. Finally, the cross-linking results obtained herein, combined with the mutational analysis of Arg 143 and Arg 147, suggest a novel role in catalysis for these two arginines within this conserved arginine fork motif formed by the conserved sequence R(V/L)XXR.

ACKNOWLEDGMENT

The authors thank Dr. Vajira Nanayakkara of the University of Utah mass spectrometry facility for his assistance and helpful discussions. We also thank Chris Fromme and Greg Verdine for providing the coordinates of the *B. st.* Endo III-DNA structure prior to publication. We would also like to thank Dr. Nikolas Chmiel for his help in generating the MutY-DNA model using the Endo III-DNA structure.

REFERENCES

1. de Boer, J., Andressoo, J. O., de Wit, J., Huijmans, J., Beems, R. B., van Steeg, H., Weeda, G., van der Horst, G. T. J., van Leeuwen, W., Themmen, A. P. N., Merdaji, M., and Hoeijmakers, J. H. J. (2002) *Science 296*, 1276−1279.
2. Farr, S. B., and Kogoma, T. (1991) *Microbiol. Rev. 55*, 561−585.
3. Newcomb, T. G., and Loeb, L. A. (1998) in *DNA Damage and Repair* (Nickoloff, J. A., and Hoekstra, M. F., Eds.) pp 65−84, Humana Press, Inc., Totowa, NJ.
4. Steenken, S., and Jovanovic, S. (1997) *J. Am. Chem. Soc. 119*, 617−618.
5. Dizdaroglu, M. (1991) *Free Rad. Biol. Med. 10*, 225−242.
6. Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) *Nature 349*, 431−434.
7. David, S. S., and Williams, S. D. (1998) *Chem. Rev. 98*, 1221−1261.
8. Miller, J. H. (1998) in *DNA Damage and Repair* (Nickoloff, J. A., and Hoekstra, M. F., Eds.) pp 97−105, Humana Press, Inc., Totowa, NJ.
9. Michaels, M. L., Tchou, J., Grollman, A. P., and Miller, J. H. (1992) *Biochemistry 31*, 10964−10968.
10. Michaels, M. L., and Miller, J. H. (1992) *J. Bacteriol. 174*, 6321−6325.
11. Michaels, M. L., Cruz, C., Grollman, A. P., and Miller, J. H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7022−7025.
12. Al-Tassan, N., Chmiel, N. H., Maynard, J., Fleming, N., Livingston, A. L., Williams, G. T., Hodges, A. K., Rhodri-Davies, D., David, S. S., Sampson, J. R., and Cheadle, J. P. (2002) *Nature Gen. 30*, 227−232.
13. Au, K. G., Clark, S., Miller, J. H., and Modrich, P. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8877−8881.
14. Radicella, J. P., Clark, E. A., and Fox, M. S. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9674−9678.
15. Tsai-Wu, J.-J., Liu, H.-F., and Lu, A.-L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 8779−8783.
16. Noll, D. M., Gogos, A., Granek, J. A., and Clarke, N. D. (1999) *Biochemistry 38*, 6374−6379.
17. Volk, D. E., House, P. G., Thiviyanathan, V., Luxon, B. A., Zhang, S., Lloyd, R. S., and Gorenstein, D. G. (2000) *Biochemistry 39*, 7331−7336.
18. Chmiel, N. H., Golinelli, M.-P., Francis, A. W., and David, S. S. (2001) *Nucleic Acids Res. 29*, 553−564.
19. Guan, Y., Manuel, R. C., Arvai, A. S., Parikh, S. S., Mol, C. D., Miller, J. H., Lloyd, R. S., and Tainer, J. A. (1998) *Nat. Struct. Biol. 5*, 1058−1064.
20. Thayer, M. M., Ahern, H., Xing, D., Cunningham, R. P., and Tainer, J. A. (1995) *EMBO J. 14*, 4108−4120.
21. Chepanoske, C. L., Golinelli, M.-P., Williams, S. D., and David, S. S. (2000) *Arch. Biochem. Biophys. 380*, 11−19.
22. Verdine, G. L., and Bruner, S. D. (1997) *Chem. Biol. 4*, 329−334.
23. Porello, S. L., Leyes, A. E., and David, S. S. (1998) *Biochemistry 37*, 14756−14764.
24. Francis, A. W., and David, S. S. (2002) *Biochemistry 42*, 801−810.
25. Hickerson, R. P., Chepanoske, C. L., Williams, S. D., David, S. S., and Burrows, C. J. (1999) *J. Am. Chem Soc. 121*, 9901−9902.
26. Boon, E. M., Pope, M. A., Williams, S. D., David, S. S., and Barton, J. K. (2002) *Biochemistry 41*, 8464−8470.
27. Zharkov, D. O., and Grollman, A. P. (1998) *Biochemistry 37*, 12384−12394.
28. Williams, S. D., and David, S. S. (1999) *Biochemistry 38*, 15417−15424.
29. Meisenheimer, K. M., and Koch, T. H. (1997) *Crit. Rev. Biochem. Mol. Biol. 32*, 101−140.
30. Favre, A., Saintome, C., Fourrey, J. L., Clivio, P., and Laugaa, P. (1998) *Photochem. Photobiol. 42*, 109−142.
31. Kubareva, E. A., Volkov, E. M., Vinogradova, N. L., Kanevsky, I. A., Oretskaya, T. S., Kuznetsova, S. A., Brevnov, M. G., Gromova, E. S., Nevinsky, G. A., and Shabarova, Z. A. (1995) *Gene 157*, 167−171.
32. Wang, Z., Huq, I., and Rana, T. M. (1999) *Bioconjugate Chem. 10*, 512−519.
33. Liu, J., Fan, Q. R., Sodeoka, M., Lane, W. S., and Verdine, G. L. (1994) *Chem. Biol. 1*, 47−55.
34. Wong, D. L., and Reich, N. O. (2000) *Biochemistry 39*, 15410−15417.
35. Golinelli, M.-P., Chmiel, N. H., and David, S. S. (1999) *Biochemistry 38*, 6997−7007.
36. Cupples, C. G., and Miller, J. H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5345−5349.
37. Yi-Brunozzi, H.-Y., Easterwood, L., Kamilar, G., and Beal, P. B. (1999) *Nucleic Acids Res. 27*, 2912−2917.
38. Leipold, M. D., Muller, J. G., Burrows, C. J., and David, S. S. (2000) *Biochemistry 39*, 14984−14992.

39. Bradford, M. M. (1976) *Anal. Chem. 72*, 248−254.
40. Hellman, U., Wernstedt, C., Gonez, J., and Heldin, C.-H. (1995) *Anal. Biochem. 224*, 451−455.
41. Golden, M. C., Resing, K. A., Collins, B. D., Willis, M. C., and Koch, T. H. (1999) *Protein Sci. 8*, 2806−2812.
42. Carey, J. (1991) *Methods Enzymol. 208*, 103−117.
43. Chepanoske, C. L., Porello, S. P., Fujiwara, T., Sugiyama, H., and David, S. S. (1999) *Nucleic Acids Res. 27*, 3197−3204.
44. Porello, S. L., Williams, S. D., Kuhn, H., Michaels, M. L., and David, S. S. (1996) *J. Am. Chem. Soc. 118*, 10684−10692.
45. Chepanoske, C. L., Langelier, C. R., Chmiel, N. H., and David, S. S. (2000) *Org. Lett. 2*, 1341−1344.
46. Kuhn, H., Smith, D. P., and David, S. S. (1995) *J. Org. Chem. 60*, 7094−7095.
47. Connor, D. A., Fallick, A. M., and Shetlar, M. D. (1998) *Photochem. Photobiol. 68*, 1−8.
48. Favre, A. (1990) *Bioorganic Photochemistry*, Vol. 1, John Wiley and Sons, New York.
49. McLuckey, S. A., and Habibi-Goudarzi, S. (1993) *J. Am. Chem. Soc. 115*, 12085−12095.
50. Biemann, K. (1990) *Methods. Enzmol. 193*, 886−887.
51. Kubareva, E. A., Thole, H., Karyagina, A. S., Oretskaya, T. S., Pingoud, A., and Pingoud, V. (2000) *Nucleic Acids Res. 28*, 1085−1091.
52. Chmiel, N. H., Livingston, A. L., and David, S. S. (2003) *J. Mol. Biol. 327*, 431−443.
53. Bernards, A. S., Miller, J. K., Bao, K. K., and Wong, I. (2002) *J. Biol. Chem. 277*, 20960−20964.
54. Fromme, J. C., and Verdine, G. L. (2003) *EMBO J. 22*, 3461−3471.

BI035537E