Supplementary Information

Synthesis of Site-Specific DNA-Protein Conjugates and Their Effects on DNA Replication

Jung Eun Yeo,[†] Susith Wickramaratne,[†,‡] Santoshkumar Khatwani,[‡]

Yen-Chih Wang,[‡] Jeffrey Vervacke,[‡] Mark D. Distefano,[†,‡] and Natalia Y. Tretyakova[†,*]

METHODS

Materials

C8-alkyne-dT-CE phosphoramidite and all other reagents for DNA synthesis were purchased from Glen Research (Sterling, VA). Synthetic DNA oligonucleotides containing native DNA bases and nucleobase modifications were prepared by solid phase synthesis using an ABI 394 DNA synthesizer (Applied Biosystems, CA), purified by HPLC on a Synergi 4u Hydro-RP 80A column, and desalted using NAP-5 columns (GE Healthcare, NJ). T4 polynucleotide kinase was obtained from New England Biolabs (Beverly, MA). γ-^{32}P ATP was purchased from PerkinElmer Life Sciences (Waltham, MA). The unlabeled dNTPs were obtained from Omega Bio-Tek (Norcross, GA). 40% 19:1 acrylamide/bis solution was purchased from Bio-Rad (Hercules, CA). Ammonium persulfate, CH_3CN, and EDTA were obtained from Fisher (Fair Lawn, NJ). Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) was purchased from AnaSpec Inc. (Fremont, CA). NuPAGE Novex 12% Bis-Tris gels (Life Technologies, Grand Island, NY) were run in 1× NuPAGE MOPS SDS running buffer (Life Technologies, Grand Island, NY) and stained with SimplyBlue SafeStain (Life Technologies, Grand Island, NY). Trypsin was obtained from Promega (Madison, WI), and ZipTips for peptide desalting were purchased from Millipore (Billerica, MA). Sep-Pak C-18 SPE cartridges were from Waters (Milford, MA). Micro biospin-6 size exclusion columns were obtained from BioRad (Hercules, CA). Recombinant hPol η, hPol κ and hPol ι were obtained from Enzymax (Lexington, KY). All other chemicals and solvents were purchased from Sigma-Aldrich (Milwaukee, WI) and were of the highest grade available.

Oligonucleotide Synthesis

DNA 23-mer (5'-AGG GTT TTC CCA GXC ACG ACG TT-3') and 18-mer (5'-TCA TXG AAT CCT TCC CCC-3'), where X = 5-(octa-1,7-diynyl)-uracil (5-alkyne-dU), were prepared by solid phase synthesis on an ABI 394 DNA synthesizer using commercial phosphoaramidites (Glen Research, Sterling, VA). Manual coupling was employed for the incorporation of 5-alkyne-dU. Deprotection of the synthesized oligonucleotides was carried out in NH_4OH at room temperature for 2 days. The corresponding unmodified 18-mer (5'-TCA TTG AAT CCT TCC CCC-3') was prepared by standard solid phase methodology. DNA 13-mer (5'-GGG GGA AGG ATT C-3') and 9-mer (5'-GGG GGA AGG-3') were purchased from Integrated DNA Technologies (Coralville, IA). All synthetic DNA oligomers were purified by HPLC on a Synergi Hydro-RP 80A (10 x 250 mm, 4 μm) column eluted at a flow rate of 3 mL/min. HPLC solvents were 150 mM NH_4OAc (A) and 1:1 mixture of 150 mM NH_4OAc and CH_3CN (B). A linear gradient of 10% to 42% B over 60 min was used. HPLC fractions containing full length oligomers were collected, concentrated under vacuum, and desalted by size exclusion via NAP-5 columns. All synthetic DNA strands were characterized by HPLC-ESI$^-$MS and quantified by UV spectrophotometry.

Preparation of Radiolabeled Oligodeoxynucleotides

DNA 23-mers containing 5-alkyne-dU at the 14^{th} position (5'-AGG GTT TTC CCA GXC ACG ACG TT-3') or 18-mers containing an 5-alkyne-dU at the 5^{th} position (5'-TCATXGAATCCTTCCCCC-3') (2 nmol in 17 μL of water) were incubated in 10x PNK buffer (3 μL) in the presence of T4 PNK (20 U) and γ-^{32}P ATP (30 μCi) at 37 °C for 1 h. The mixture

was heated at 65 °C for 10 min to inactivate the enzyme and passed through Illustra Microspin G25 columns (GE Healthcare, Pittsburgh, PA) to remove excess γ-^{32}P ATP.

Preparation of Azide-Containing Peptides

Peptide synthesis was carried out using an automated solid-phase peptide synthesizer (PS3, Protein Technologies Inc, Memphis, TN) employing standard Fmoc chemistry and HCTU coupling procedures. The 10-mer peptide $N_3(CH_2)_3CO$-EQKLISEEDL-NH_2 was synthesized on a Rink-amide-MBHA resin (0.1 mmol scale) by coupling each amino acid in DMF containing Fmoc-amino acid (4 eq), HCTU (4 eq), 6-Cl-HOBt (4 eq.) and N-methylmorpholine (8 eq) for 20 min. Fmoc deprotection of the peptide was performed with 20% piperidine in DMF for 5 min twice. N-terminal 4-azidobutanoic acid[43] was coupled in DMF containing 4-azidobutanoic acid (4 eq.), HCTU (4 eq.), 6-Cl-HOBt (4 eq.) and DIPEA (8 eq.) for 2 h. The peptide was simultaneously cleaved from resin and deprotected by treatment with Reagent K (0.5 g phenol, 0.5 mL H_2O, 0.5 mL thioanisole, 0.25 mL 1, 2-ethanedithiol in 10 mL trifluoroacetic acid (TFA)) for 2 h. The solution was concentrated to 2 mL by bubbling with N_2 and precipitated in Et_2O. HPLC purification of the peptide was performed using a Luna C18 (10 x 250 mm, 10 μm) column (Phenomenex, Torrance, CA) eluted with a gradient of aqueous 0.1% TFA (A) and 0.1% TFA in CH_3CN (B). Solvent composition was maintained at 0% B for 5 min, followed by a linear gradient of 0 to 70% B over 50 min at a flow rate of 5 mL/min. Calculated for $C_{55}H_{93}N_{16}O_{21}$, $[M+H]^+$ = 1313.6696, found $[M+H]^+$ = 1313.6436 from ESI^+-MS.

Synthesis of the 23-mer peptide began on a preloaded Fmoc-Glu(OtBu)-Wang resin (0.10 mmol) and the peptide chain was elongated using HCTU/N-Methylmorpholine-catalyzed, single coupling steps with protected amino acids (4 eq.) and HTCU (4 eq.) for 30 min. Following

complete chain elongation, the peptide's N-terminus was deprotected with 10% piperidine in DMF (v/v) and the presence of the resulting free amine was confirmed by ninhydrin analysis. The resin containing the peptide was washed with CH_2Cl_2, dried *in vacuo* overnight, weighed, and divided into two portions for further synthesis on a reduced scale. Using 50.0 μmol of peptide, the free amino terminus was acylated with the 4-azidobutanoic acid (13.5 mg, 100 μmol, 2 eq.) catalyzed by DIC (13 mg, 100 μmol, 2 eq.) in the presence of DIEA (8.6 μL, 5.0 μmol, 0.1 eq.) in DMF (5 mL) for 10 h. After acylation was judged complete by ninhydrin analysis, the resin bound peptide was washed thoroughly with CH_2Cl_2 and dried *in vacuo* for 4 h. The peptide was cleaved from the resin along with simultaneous side chain deprotection by treatment with Reagent K for 2 h at rt. The released peptide was collected and combined with TFA washes of the resin before precipitation of the peptide in chilled Et_2O (100 mL). The crude solid peptide was collected by centrifugation, the supernatant was removed, and the resulting pellet was washed 2 times with cold Et_2O (50 mL) repeating the centrifugation and supernatant removal steps each time. The crude peptide was purified using a semi-preparative C_{18} RP-HPLC column with detection at 220 nm and eluted with a gradient of solvent A (H_20/0.1% TFA, v/v) and solvent B (CH_3CN/0.1% TFA, v/v). The crude peptide (150 mg) was dissolved in a DMF/H_2O solution (1:5 v/v, 25 mL), applied to the column equilibrated in Solvent A, and eluted using a linear gradient of (0-70% Solvent B over 1.5 h at a flow-rate of 5 mL/min). Fractions were analyzed for purity by an analytical C_{18} RP-HPLC column employing a linear gradient (0-100% Solvent B over 60 min at a flow-rate of 1 mL/min) and detected at 220 nm. Fractions containing peptide product of at least 90% purity were pooled and concentrated by lyophilization to yield 35 mg (25% yield) of a white solid. . A small amount (< 1mg) of the resulting purified peptide was dissolved in 10 μl of 0.1% TFA/CH_3CN and diluted 1:50 in a mixture of CH_3CN/H_2O (1:1 v/v)

prior to MS analysis. MS was performed using a 50 μL injection and collecting 3000 scans. Calculated for $C_{109}H_{178}N_{34}O_{38}$, $[M+2H]^{+2}$ = 1258.1345, found $[M+2H]^{+2}$ = 1258.1406 from ESI^{+}-MS.

Preparation of 6xHis-eGFP-CVIA

Green fluorescent protein genetically engineered to contain a hexa-histidine tag (6xHis-eGFP) was expressed and purified as previously described.[44,45] 6xHis-eGFP-N_3 was enzymatically prenylated on the cysteine of the C-terminal CVIA CaaX box motif using yeast farnesyl transferase.[35] Enzymatic reactions (total volume 10 mL) were carried out by incubating a solution of the protein (2 μM) and DTT (5 mM) (premixed and incubated at RT for 1 h) with $MgCl_2$ (10 mM), Tris-HCl (10 μM, pH 7.5), C10 dihydroazide (40 μM)[44] and yeast farnesyl transferase (150 nM) at 30 °C for 2 h. The reaction mixture was concentrated to ~ 500 μL by Centricon centrifugal filters (MWCO 10,000), and the excess azide was removed by size exclusion with a NAP-5 column eluted with PBS buffer (50 mM NaH_2PO4, 0.1 M NaCl, pH 7.3). The concentration of the prenylated protein was determined by measuring its absorbance at 488 nm (ε_{488} for eGFP, 55, 000).

Cross-Linking Reaction Between Azide-Functionalized Peptide and Alkyne-Containing DNA

^{32}P labeled oligodeoxynucleotides (5'-AGG GTT TTC CCA GXC ACG ACG TT-3' or 5'-TCA TXG AAT CCT TCC CCC-3', 15 pmol), where X = C8-alkyne-dU, were mixed with 1.5 nmol of azide-functionalized peptide 10-mer $N_3(CH_2)_3CO$-EQKLISEEDL-NH_2 or the 23-mer $N_3(CH_2)_3CO$-PDAQLVPGINGKAIHLVNNESSE, 1 μL of TBTA (5 mM stock in DMSO:t-BuOH 1:4), 10 μL of TCEP (5 mM stock in H_2O), 10 μL of $CuSO_4$ (5 mM stock in H_2O), and 50 mM phosphate buffer (pH 7.5), in a final reaction volume of 50 μL. The reaction was mixed

using a rotatory shaker at room temperature. After allowing the reaction to proceed for 1.5-2 h, it was quenched by adding 0.5 mM EDTA. An aliquot was radiolabeled using γ-^{32}P ATP as described previously,[46] loaded onto a 20% denaturing PAGE gel, and ran at 15 W for 1.5 h. The products were visualized using a Typhoon FLA 7000 phosphorimager (GE Healthcare, Pittsburgh, PA).

Gel Electrophoresis Purification of DNA-Protein Cross-Links

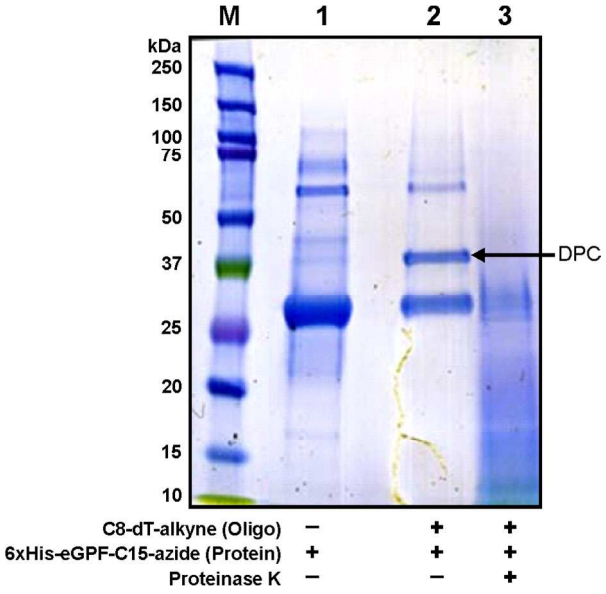
DNA-protein cross-linking reaction mixtures were desalted by Micro biospin-6 size exclusion columns (BioRad, Hercules, CA), while DNA-peptide reaction mixtures were desalted using Sep-pak C-18 SPE cartridges (Waters, Milford, MA). The resulting solutions were loaded onto 15% or 20% (w/v) denaturing PAGE gels containing 7 M urea. DPC bands were visualized by SimplyBlue SafeStain, excised, and the DPCs were extracted using the flextube gel elution system (IBI Scientific, IA) following the manufacturer's instructions. DNA-peptide conjugates were extracted from the gel by using a freeze-thaw method. Gel pieces were suspended in 1xTE buffer, and subjected to 10 cycles of subsequent freezing in dry ice-ethanol and allowed to thaw at rt. The gel bands were then incubated at 37 °C for 2 days. Following gel purification, DNA-protein cross-links were desalted using Micro biospin-6 columns, while DNA-peptide conjugates were desalted by SPE on Sep-pak C-18 cartridges.

Mass Spectrometry characterization of synthetic DNA-peptide cross-links

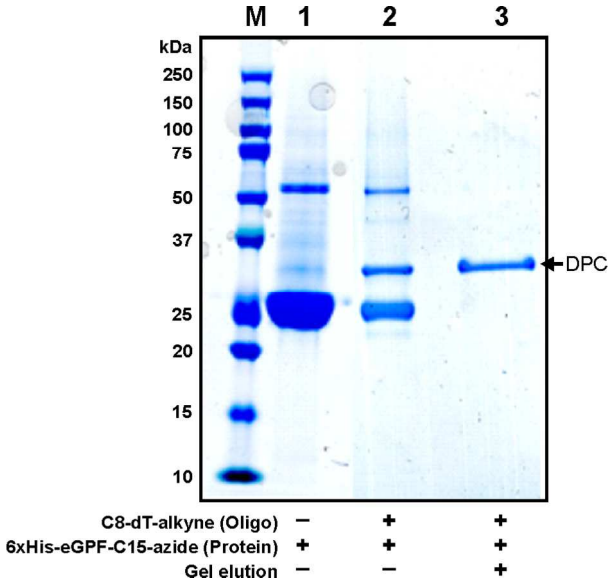
DNA-peptide conjugates, 5'-AGG GTT TTC CCA GXC ACG ACG TT-3' containing a covalent crosslink from C-5 position of dU to the N-terminus of the 10-mer peptide EQKLISEEDL, were generated by copper-catalyzed click chemistry as described above (*Scheme*

2A) and isolated using 15% or 20% (w/v) denaturing polyacrylamide gels containing 7 M urea. Following elution from the gel using the freeze-thaw method, the cross-links were desalted by SPE on Sep-pak C-18 cartridges. The DNA component of the conjugate was digested to nucleosides with PDE I (120 mU), PDE II (105 mU), DNase (35 U), and alkaline phosphatase (22 U) at 37 °C overnight in 10 mM Tris-HCl/15 mM $MgCl_2$ (pH 7) buffer. The resulting dU-peptide conjugates were dried under vacuum and desalted with C18 ZipTips. Samples were dissolved in 0.1% acetic acid (25 μL), and 5-8 μL of this solution was used for MS analysis. NanoLC-nanospray-MS/MS analysis was conducted using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Waltham, MA) as described above for DNA-protein cross-links, with the exception that HPLC was conducted using a gradient of 0.1% formic acid in H_2O (A) and CH_3CN (B). Mass spectrometry analysis of DNA-peptide crosslinks was performed using a FTMS mass analyzer with a resolution of 30,000 and a scan range of 300 to 2000 in the full scan mode using an isolation width of 1.0 m/z, 35% normalized CID collision energy. Peptide MS/MS spectra were collected using an isolation width of 3.0 m/z, 40% normalized HCD collision energy with a resolution of 7500 and a scan range of 50 to 2000 m/z.

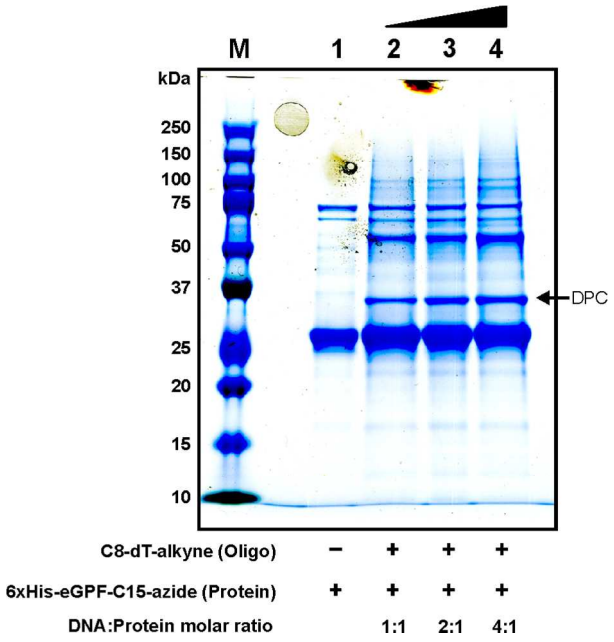
S-1. Formation of DNA-Protein cross-link (DPC) upon copper catalyzed cycloaddition reaction between His-eGFP-C15-N_3 and 23-mer oligonucleotide (5'-AGG GTT TTC CCA G***C8-alkyne-dT***C ACG ACG TT-3'). Samples were resolved on a 12% (w/v) SDS-PAGE gel and visualized by SimplyBlue staining for protein. M: protein marker; Lane 1: His-eGFP-C15-N_3; Lane 2: alkyne containing oligonucleotide and His-eGFP-C15-N_3 in the presence of Cu catalyst; Lane 3: A CuAAC reaction mixture from lane 2 subjected to proteinase K digestion.



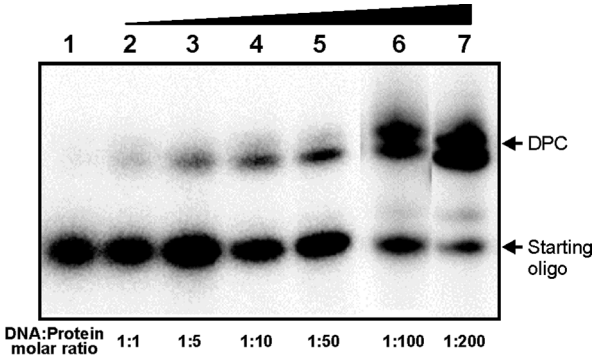
S-2. Purification of DPCs by gel electrophoresis – electroelution. DPCs were generated by CuAAC of His-eGFP-C15-N_3 and 23-mer oligonucleotide (5'-AGG GTT TTC CCA G***C8-alkyne-dT***C ACG ACG TT-3'). The reaction mixture was desalted by Micro Biospin-6 column, loaded onto a 12% (w/v) SDS-PAGE gel, and the gel band containing the DPC was excised and applied to electroelution. M: protein marker; Lane 1: His-eGFP-C15-N_3; Lane 2: a CuAAC reaction mixture containing His-eGFP-C15-N_3 and alkyne containing oligonucleotide in the presence of Cu catalyst; Lane 3: a purified sample via electroelution method.



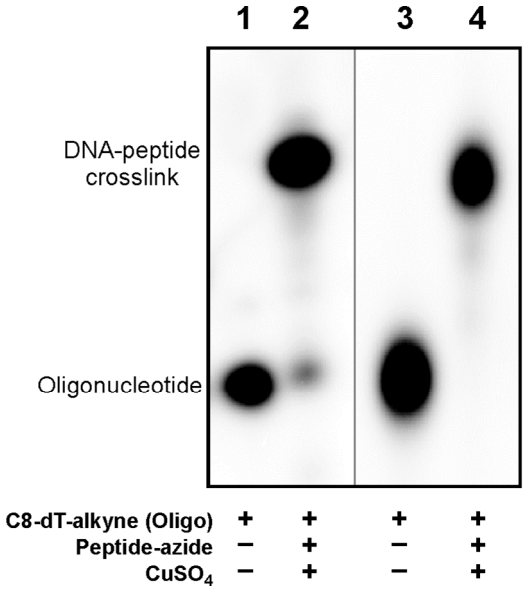
S-3. Variation of DNA-protein crosslink yields with increased protein:DNA molar ratio. DPCs were generated by CuAAC of His-eGFP-C15-N_3 and 23-mer (5'-AGG GTT TTC CCA G***C8-alkyne-dT***C ACG ACG TT-3') by increasing the molar equivalents of DNA to protein from 1:1 to 4:1. The reaction mixtures were resolved on a 12% (w/v) SDS-PAGE gel, and unreacted protein and DPCs were visualized by SimplyBlue staining for protein. The DPC yield increased with increasing amount of the DNA.



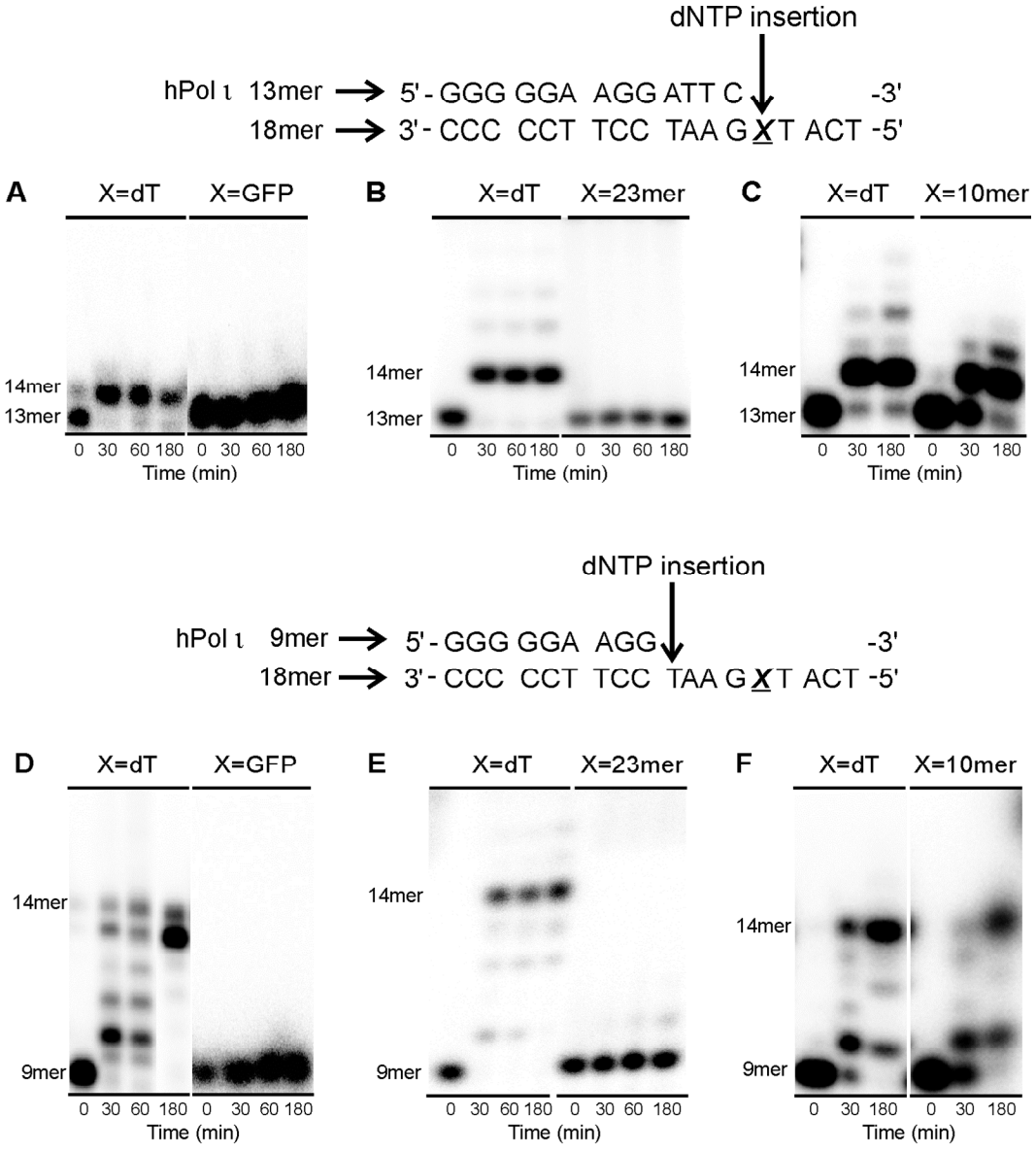
S-4. Variation of DNA-peptide crosslink yields with increased DNA:peptide molar ratio. DPCs were generated by copper-catalyzed cycloaddition of 10mer peptide-N_3 ($N_3(CH_2)_3$COEQKLISEEDLNH_2) and 5'-^{32}P-23-mer (5'-AGG GTT TTC CCA G***C8-alkyne-dT***C ACG ACG TT-3') by increasing the molar equivalents of DNA to peptide from 1:1 to 1:200. The reaction mixtures were resolved on a 20% (w/v) denaturing PAGE gel, and DNA and DPCs were visualized by phosphorimaging analysis. The DPC yield increased with increasing amount of the peptide.



S-5. Denaturing PAGE analysis of DNA-peptide cross-links generated using 24mer peptide ($N_3(CH_2)_3CO$-PDAQLVPGINGKAIHLVNNESSE) and 5'-^{32}P-23-mer (5'-AGG GTT TTC CCA G***C8-alkyne-dU***C ACG ACG TT-3') by click chemistry. Samples were resolved by 15% (w/v) denaturing PAGE and visualized by phosphorimaging analysis. Lane 1: alkyne containing oligonucleotide; Lane 2: alkyne containing oligonucleotide and peptide-N_3 in the presence of Cu; Lane 3: alkyne containing oligonucleotide; Lane 4: gel purified DNA-peptide conjugate from lane 2.



S-6. Extension of 5'-^{32}P-labeled primers annealed to template strands containing DNA-protein and DNA-peptide conjugates by hPol ι under standing start (A-C) and running start (D-F) conditions. The primers were annealed to the 18-mer templates containing unmodified dT, dT-GFP, dT-23-mer peptide, or dT-10mer peptide. The resulting primer-template complexes (40 nM) were incubated at 37 °C in the presence of hPol ι (80 nM). Primer extension was initiated by the addition of all four dNTPs (500 μM) and quenched at indicated time points. The extension products were separated by 20% (w/v) denaturing PAGE and visualized by phosphorimaging analysis.



S-7. Polymerase bypass of DNA-protein conjugates by hPol ν under standing start (A) and running start (B) conditions. The primers were annealed to the 18-mer templates containing unmodified dT and dT-GFP. The resulting primer-template complexes (40 nM) were incubated at 37 °C in the presence of hPol ν (160 nM). Reactions were started by the addition of all four dNTPs (500 μM) and quenched at indicated time points. The extension products were resolved by 20% (w/v) denaturing PAGE and visualized by phosphorimaging analysis.