Supporting Information

Optical Control of DNA Helicase Function through Genetic Code Expansion

Ji Luo,[1] Muwen Kong,[2,3] Lili Liu,[3] Subhas Samanta,[1] Bennett Van Houten,[2,3] and Alexander Deiters*,[1,3]

[1]Department of Chemistry, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, United States; [2]Department of Pharmacology and Chemical Biology; [3]The University of Pittsburgh Cancer Institute, Hillman Cancer Center, Pittsburgh, Pennsylvania 15213, United States.

Cloning of expression vectors.

1) Construction of pETM11-UvrD-K37A and pETM11-UvrD-K37TAG: Plasmids were obtained by converting the AAG (Lys) codon of pETM11-wt-UvrD into an TAG and GCG (Ala) codon using primers QC1/QC2 and QC3/QC4, respectively, with a QuikChange site-directed mutagenesis kit (Agilent).

2) Construction of pBAD-wt-UvrD-PyltRNA, pBAD-UvrD-K37TAG-PyltRNA, and pBAD-UvrD-K37A-PyltRNA: A 6×HIS-tagged UvrD was generated through amplification of the UvrD gene from pETM11-UvrD (a gift from Dr. Caroline Kisker, University of Wurzburg) using primers N1/N2 and cloning into the pBAD-PyltRNA backbone,[1] creating pBAD-6HIS-UvrD-PyltRNA. The mutated UvrD-K37TAG (pETM11-UvrD-K37TAG) and UvrD-K37A (pETM11-UvrD-K37A) genes were cloned into the pBAD-PyltRNA expression vector with the *Nco*I and *Nde*I sites to creating pBAD-6HIS-UvrD-K37TAG-PyltRNA and pBAD-6HIS-UvrD-K37A-PyltRNA plasmids, respectively.

Table S1. Primer list.

Primer No.	**Sequences**
QC1	5' CGGGAGCGGGCAGCGGATAGACGCGCGTGCTGACGCAC 3'
QC2	5' GTGCGTCAGCACGCGCGTCTATCCGCTGCCCGCTCCCG 3'
QC3	5' GAGCGGGCAGCGGAGCAACGCGCGTGCTGAC 3'
QC4	5' GTCAGCACGCGCGTTGCTCCGCTGCCCGCTC 3'
N1	5' GAATTAACCATGGTGCATCATCATCATCATCATATGAATTTTTTATCGGA AAAGC 3'
N2	5' CGAATTCCCATATGTCATTATACTTTCTCAATCGGCG 3'

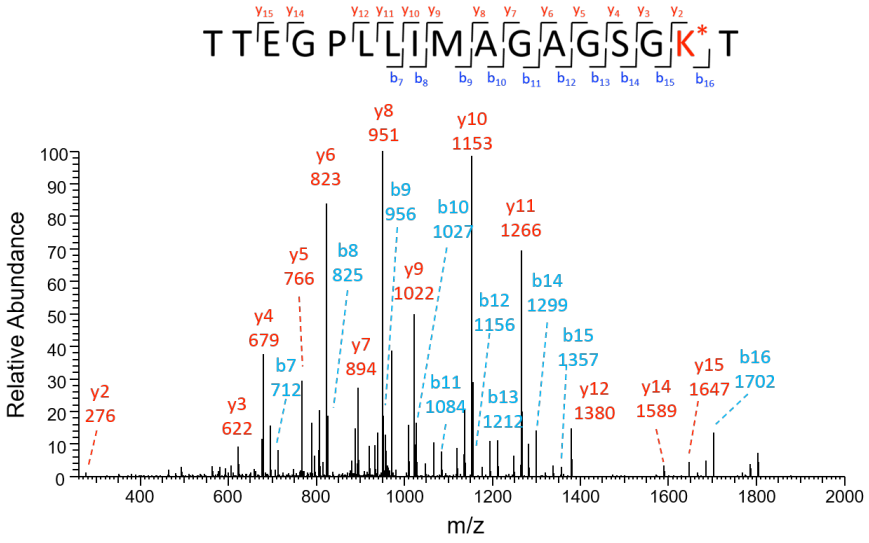
Expression and purification of caged UvrD. The plasmid pBAD-UvrD-K37TAG-PyltRNA was co-transformed with pBK-BhcKRS[2] into *E. coli* Top10 cells. A single colony was grown in LB media containing 25 μg/ml of tetracycline and 50 μg/ml of kanamycin overnight, and then 500 μL of the overnight culture was added to 25 mL LB media supplemented with 1 mM of **HCK** and 25 μg/mL of tetracycline and 50 μg/mL of kanamycin. Cells were grown at 37 °C, 250 rpm, and protein expression was induced with 0.1% arabinose when the OD_{600} reached ~0.6. After overnight expression at 37 °C, cells were pelleted at 4 °C, 5,000 g, for 10 minutes, and washed once with 10 ml of PBS. The cell pellets were re-suspended in 6 ml of phosphate lysis buffer (Qiagen; 50 mM, pH 8.0) and Triton X-100 (60 μL, 10%), gently mixed, and incubated for an hour at 4 °C. The cell mixtures were sonicated with four short burst of 30 s followed by intervals of 30 s for cooling, and the cell lysates were centrifuged at 4 °C, 13,000 g, for 10 minutes. The supernatant was transferred to a 15 mL conical tube and 100 μL of Ni-NTA resin (Qiagen) was

added. The mixture was incubated at 4 °C for 2 hours under mild shaking. The resin was then collected by centrifugation (1,000 g, 10 min), washed twice with 400 µL of lysis buffer, followed by two washes with 400 µL of wash buffer containing 20 mM imidazole. The protein was eluted with 300 µL of elution buffer containing 250 mM imidazole. The purified proteins were analyzed by 10% SDS PAGE and stained with Coomassie Blue.

Expression and purification of wild-type UvrD. BL21(DE3) pLysS *E. coli* competent cells (Promega) were transformed with plasmid encoding 6×His-tagged UvrD and selected for kanamycin resistance on LB agar plates overnight at 37 ºC. Colonies were picked and grown in liquid LB starter media with 25 µg/mL kanamycin overnight on shaker at 37 ºC. 2L LB media was inoculated with starter culture, grown to OD_{600} = 1 on shaker at 37 ºC, before IPTG was then added to the final concentration of 1 mM for induction at 30 ºC. Cells were harvested 2 hrs after induction and pellets stored at -80 ºC. Pellets were first thawed and roughly re-suspended in Buffer A (20 mM HEPES pH 7.5, 100 mM KCl, 3 mM $MgCl_2$, 2.5 mM BME) with 0.2 mg/ml lysozyme and protease inhibitor cocktail. Cells were then fully re-suspended in a homogenizer and lysed via sonication before centrifuged at 35,000 rpm for 45 min at 4 ºC. Supernatant, supplemented with 5 mM imidazole was loaded onto a Ni column. The column was washed with 4 column volumes of Buffer A containing 5 mM imidazole, and then eluted with 5 column volumes of Buffer A and an imidazole gradient up to 200 mM. Fractions were collected and checked using a 10% Bis-Tris gel via SDS-PAGE. Fractions of similar concentrations were pooled and dialyzed against Buffer A overnight to remove imidazole. Dialyzed sample was loaded onto a HiLoad 16/60 Superdex 200 prep grade size exclusion column (GE healthcare). Peak fractions were collected and checked for quality via SDS-PAGE. Fractions of similar quality were again combined and dialyzed overnight against Buffer A containing 50% glycerol. Final concentrations of purified UvrD were determined via SDS-PAGE and Bradford assay, comparing to BSA of known concentrations.

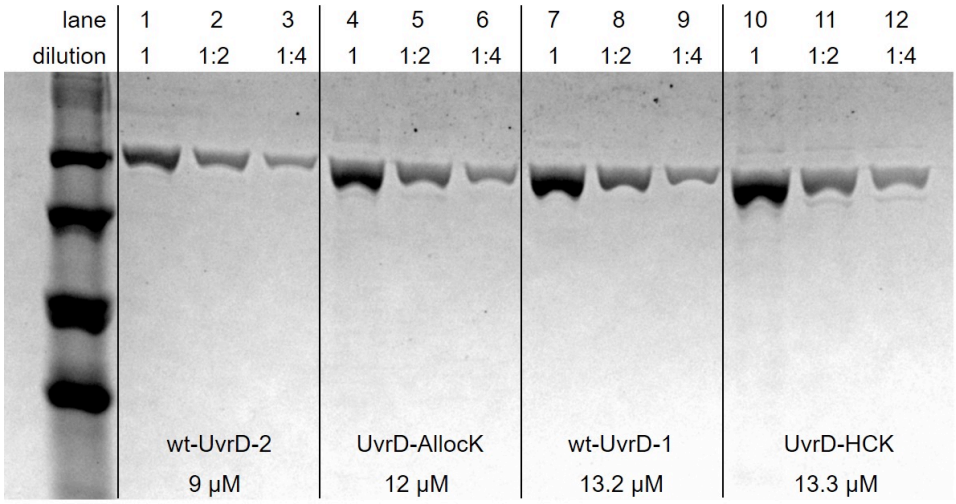
LC-MS/MS sequencing confirms site-specific incorporation of HCK into UvrD. Caged UvrD purified from *E. coli* containing the UvrD-K37TAG and BHCKRS/$tRNA_{CUA}$ plasmids was analyzed by SDS-PAGE and MS/MS sequencing (Supporting information Figure S1).
1) In gel trypsin digestion: In gel trypsin digestion was carried out as previously described.[3] Excised gel bands were washed with HPLC water and destained with 50% acetonitrile (ACN)/25 mM ammonium bicarbonate until no visible staining. Gel pieces were dehydrated with 100% ACN, reduced with 10 mM dithiothreitol (DTT) at 56°C for 1 hr, followed by alkylation with 55 mM iodoacetamide (IAA) at room temperature for 45 min in the dark. Gel pieces were then again dehydrated with 100% ACN to remove excess DTT and IAA, and rehydrated with 20 ng/µl trypsin/25 mM ammonium bicarbonate and digested overnight at 37°C. The resultant tryptic peptides were extracted with 70% ACN/5% formic acid, vacuum dried and re-constituted in18 µl 0.1% formic acid.
2) Tandem mass spectrometry: Proteolytic peptides from in gel trypsin digestion were analyzed by a nanoflow reverse-phased liquid chromatography tandem mass spectrometry (LC-MS/MS). Tryptic peptides were loaded onto a C18 column (PicoChip™column packed with 10.5cm Reprosil C18 3µm120Å chromatography media with a 75 µm ID column and a 15 µm tip, New Objective, Inc., Woburn, MA) using a Dionex HPLC system (Dionex Ultimate 3000, ThermoFisher Scientific, San Jose, CA) operated with a double-split system (Personal communication with Dr. Steve Gygi from Department of Cell Biology, Harvard Medical School) to provide an in-column nano-flow rate (~300 nl/min). Mobile phases used were 0.1% formic acid for A and 0.1% formic acid in acetonitrile for B. Peptides were eluted off the column using a

52 min gradient (2-40% B in 42 min, 40-95% B in 1min, 95% B for 1 min, 2% B for 8 min) and injected into a linear ion trap MS (LTQ-XL, ThermoFisher Scientific) through electrospray.
The LTQ XL was operated in a date-dependent MS/MS mode in which each full MS spectrum [acquired at 30000 automatic gain control (AGC) target, 50ms maximum ion accumulation time, precursor ion selection range of m/z 300 to1800] was followed by MS/MS scans of the 5 most abundant molecular ions determined from full MS scan (acquired based on the setting of 1000 signal threshold, 10000 AGC target, 100ms maximum accumulation time, 2.0 Da isolation width, 30ms activation time and 35% normalized collision energy). Dynamic exclusion was enabled to minimize redundant selection of peptides previously selected for CID.
3) Peptide identification by database search: MS/MS spectra were searched using MASCOT search engine (Version 2.4.0, Matrix Science Ltd) against the UniProt E. Coli proteome database with the inclusion of amino acid sequence for UvrD-**HCK** and myoglobin-APIK. The following modifications were used: static modification of cysteine (carboxyamidomethylation, +57.05 Da), variable modifications of methionine (oxidation, +15.99Da), protein N-terminus (acetylation, +42.01) and lysine (hydroxycoumarin modification, +218.17). The mass tolerance was set at 1.4 Da for the precursor ions and 0.8 Da for the fragment ions. Peptide identifications were filtered using PeptideProphet™ and ProteinProphet® algorithms with a protein threshold cutoff of 99% and peptide threshold cutoff of 90% implemented in Scaffold™ (Proteome Software, Portland, Oregon, USA).



Supporting Figure S1. MS/MS analysis of UvrD-**HCK** expressed in Top10 cells and purified by Ni-NTA column. It confirms the site-specific incorporation of **HCK** into UvrD. (K* represents **HCK**.)

Calculation of protein concentrations. Concentrations of wt-UvrD, UvrD-**HCK**, and UvrD-K37A were determined by SDS-PAGE, ChemiDoc imaging, and ImageJ analysis (Supporting Figure S2).

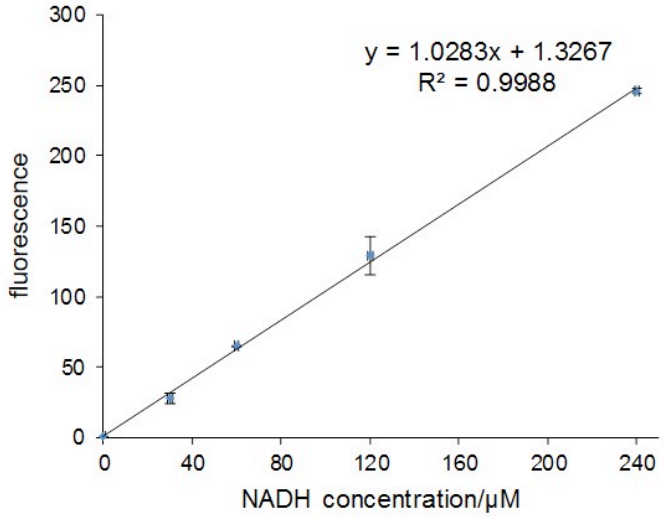


Supporting Figure S2. SDS-PAGE analysis of wild-type UvrD (wt-UvrD), inactive UvrD-K37A mutant (UvrD-K37A), and photocaged UvrD (UvrD-**HCK**) expressed in *E. coli*. The gel was stained with Coomassie blue and imaged by ChemiDoc (BioRad). The proteins were loaded in a serial dilution (1, 1:2, and 1:4) to determine the concentration. The protein concentrations were calculated using ImageJ and normalized to that of the original wild-type UvrD (13.2 µM, wt-UvrD-1).

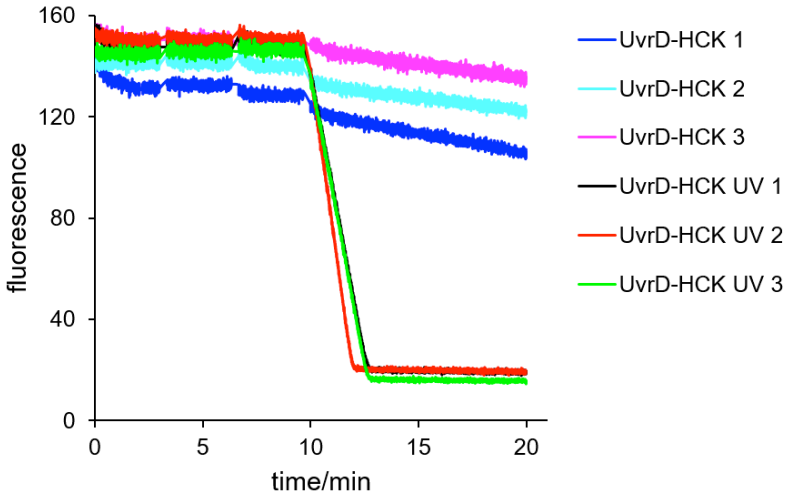
Table S2. Oligo list for DNA substrates.

Substrates	**Sequences**
5FAM/39/15t	5' /56-FAM/AGTGCGCAAGCTTGTCAAGGCACTTTTTTTTTTTTTTT 3'
3Dab/39/20t	5' TTTTTTTTTTTTTTTTTTTTCTTGACAAGCTTGCGCACT/3Dab/ 3'
39/20t	5' TTTTTTTTTTTTTTTTTTTTCTTGACAAGCTTGCGCACT 3'
FL16	5' CTGCAGCGAG/iFluorT/CCATG 3'
UD16	5' CTGCAGCGAGTCCATG 3'
disp2	5' CATGGACTCGCTGCAG 3'

Fluorescence-based coupled ATPase activity assays. This ATPase assay design is based on the reaction of the conversion of phosphoenolpyruvate (PEP) to pyruvate by pyruvate kinase (PK), which is then coupled to the conversion of pyruvate to lactate by lactate dehydrogenase (LDH).[4] In the first reaction, one ATP molecule is hydrolyzed by the ATPase to ADP, and then converted back to ATP by PK. In the second reaction, one NADH molecule is oxidized into a NAD^+ by LDH. A master mix containing 20 U/ml PK, 20 U/ml LDH, 2 mM PEP, 100 µM NADH in 1× reaction buffer (20 mM HEPES pH 7.5, 50 mM KCl, 3 mM $MgCl_2$, 5 mM DTT) was prepared in a 400 µl quartz cuvette (Agilent). The concentration of NADH is monitored throughout the experiment by a Cary Eclipse Fluorescence Spectrophotometer (Agilent) via excitation at 384 ± 5 nm and emission at 461 ± 5 nm with a PMT voltage of 700 V at 37 ºC. Data collection started with only the master mix in the absence of DNA, protein, or ATP. Fluorescence intensity was allowed to stabilize before 100 nM of ssDNA (39/20t, Table S2), 25 nM of protein, and 1 mM of ATP (final concentrations) were added sequentially. Typically, additions in each step were timed five minutes apart to allow the signal to stabilize again. ATPase assays were performed for wild-type UvrD, photocaged UvrD, decaged UvrD, and inactive UvrD-K37A.



Supporting Figure S3. NADH standard curve (λ_{ex} = 384 ± 5 nm, λ_{em} = 461 ± 5 nm). Δ1.028 unit of fluorescence intensity equals to Δ1 μM NADH (ATP). Error bars represent standard deviations from three independent experiments.



Supporting Figure S4. Fluorescence-based coupled ATPase activity assay. Time-course fluorescence measurement of NADH consumption driven by UvrD-**HCK** in the absence and presence of UV treatment (three independent experiments).

Based on the standard curve and fitting curve established in Figure S3, the ATPase activities were calculated as shown in Table S3.

Table S3. ATP turnover number of UvrDs in ATPase assays.

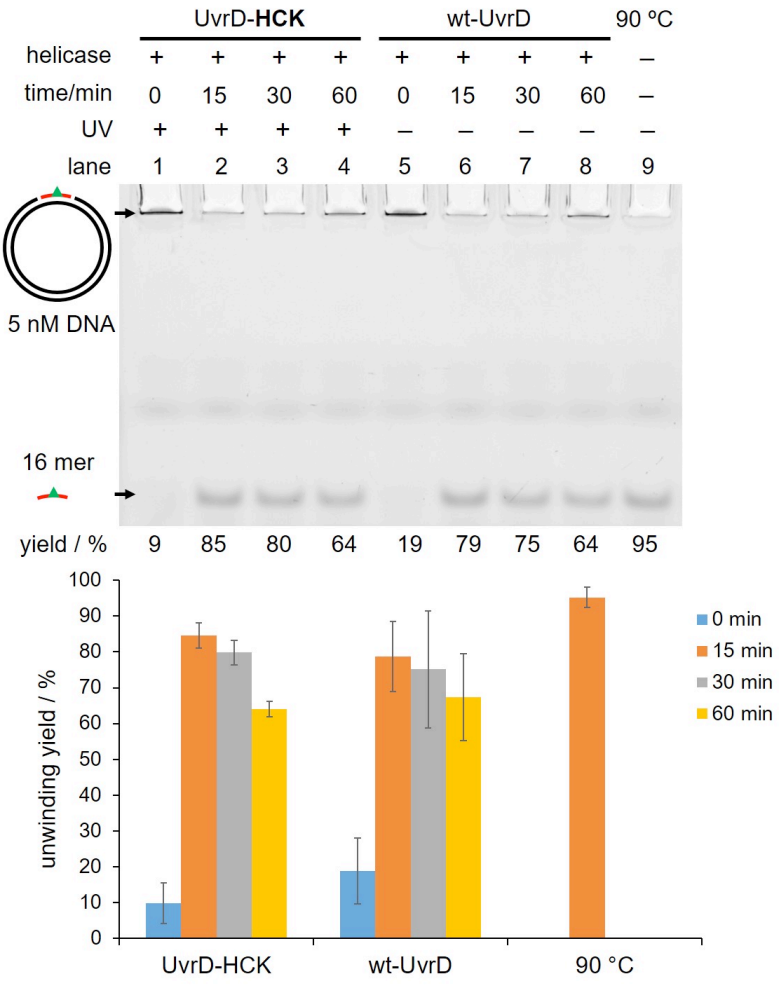
Protein	ATP turnover number* (s^{-1}, mean ± s.d.)
wt-UvrD	46.0 ± 20.9
UvrD-K37A	2.6 ± 1.3
UvrD-**HCK** (+UV)	41.2 ± 14.6
UvrD-**HCK** (–UV)	2.0 ± 0.5

*Error bars represent standard deviations (s.d.) from three independent experiments.

Fluorescence-based helicase activity assay. A master mix containing 5 nM of quenched Y-shaped double-stranded substrate in 1× reaction buffer (20 mM HEPES, pH 7.5, 50 mM KCl, 3 mM $MgCl_2$, 5 mM DTT) was prepared in a 96-well black plate (Greiner) at 200 µL per well. Fluorescence intensity was monitored over time with excitation at 485 ± 5 nm and emission at 520 ± 5 nm at 37 ºC. Proteins (50 nM, final concentration) and ATP (3 mM) were added to the mix sequentially. The helicase assays were carried out for wild-type UvrD, photocaged UvrD, and decaged UvrD.

Construction of a nicked plasmid with defined lesion. Plasmid pLL7 was derived from plasmid pSCW01[5] by QuikChange site-directed mutagenesis kit (Agilent), such that only two out of the four original Nt.BstNBI nick sites remain, 16 bases apart. The plasmid was then extracted from transformed *E. coli* culture with QIAGEN Plasmid Maxi Kit (12163). Nicking of purified plasmid pLL7 was carried out at the final concentration of 400 ng/µl, in the presence of 0.8 U/µl Nt.BstNBI (NEB) and 50-fold molar excess of displacer oligo (disp2, Table S2) for 4 h at 55 °C. Nt.BstNBI was inactivated by heating at 85 °C for 20 min. The reaction was allowed to slowly cool down to room temperature, generating plasmids with a 16-base gap. Gapped plasmids were purified by adding equal volume of 2× PEG solution (26% PEG-8000, 20 mM $MgCl_2$) followed by centrifugation at 16,000 *g* for 1 h at 4 °C. The precipitated DNA pellet was washed with 500 µl of 70% ethanol, spun down at 16,000 *g* for 15 min at 4 °C, and re-dissolved in deionized water. Purified gapped plasmids (550 ng/µl) were annealed with a 3-fold molar excess of a 16mer containing a site-specific fluorescein-modified thymine, FL16 (Table S2) in NEBuffer 2.1. The efficiency of the gapping and annealing reactions was confirmed by restriction enzyme digestion within the 16 base regions. Excess 16mer oligos with defined lesion were removed by purification with a QIAquick PCR Purification Kit (Qiagen).

Gel-based helicase activity assay. Gel-based assays were carried out with UvrD and FL16-containing pLL7 plasmids as prepared above. 10 µL reactions containing 5 nM DNA, 3 mM ATP, and 50 nM protein were incubated at 37 °C for different time periods. Two controls were used in these gel-based experiments: (1) DNA substrate only as a negative control, and (2) heat denatured substrate (90 °C for 3 min) as a positive control. All controls and reactions of 10 µL volumes were stopped by the addition of 10 µL of 2× Stop Buffer (500 nM UD16, 10% glycerol, 0.1% bromophenol blue, 40 mM EDTA in 1× reaction buffer). 10 µL of each tube was loaded into lanes of an 8% TBE gel, run at 200 V for 40 min in 1× TBE, imaged by a GE Typhoon 9400 scanner, and quantified by ImageJ.



Supporting Figure S5. Representative gel-based helicase activity assays of wild-type and photocaged UvrD. Bands were analyzed using ImageJ and error bars represent standard deviations from three independent experiments.

References

(1) Luo, J.; Uprety, R.; Naro, Y.; Chou, C.; Nguyen, D. P.; Chin, J. W.; Deiters, A. *J Am Chem Soc* **2014**, *136*, 15551.

(2) Gautier, A.; Nguyen, D. P.; Lusic, H.; An, W.; Deiters, A.; Chin, J. W. *J Am Chem Soc* **2010**, *132*, 4086.

(3) Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. *Nat Protoc* **2006**, *1*, 2856.

(4) Takashi, R.; Putnam, S. *Anal Biochem* **1979**, *92*, 375.

(5) Geng, H.; Du, C.; Chen, S.; Salerno, V.; Manfredi, C.; Hsieh, P. *Anal Biochem* **2011**, *413*, 179.