

5. Designer Reagents for Mass Spectrometry-Based Proteomics: Clickable Cross-Linkers for Elucidation of Protein Structures and Interactions

ABSTRACT We present novel homobifunctional amine-reactive clickable cross-linkers (CXLs) for investigation of three-dimensional protein structures and protein-protein interactions (PPIs). CXLs afford distinct advantages over other currently available cross-linkers, including (1) their small size and cationic nature in physiological pH for good water solubility and cell permeability, (2) an alkyne group for bio-orthogonal conjugation to a biotin-azide or other affinity tags via click reaction for enrichment of cross-linked peptides, (3) a highly selective nucleophilic displacement reaction by the resultant 1,2,3-triazole yielding a reporter ion for fast screening of cross-linked peptides, (4) a readily modifiable chain length between the amine-reactive groups to act as molecular “rulers” for enhanced elucidation of structural details. Ubiquitin, a small and lysine-abundant protein, is used as a model system to demonstrate protein structural studies using CXLs. To validate the sensitivity of our approach, biotin-azide labeling and subsequent enrichment of cross-linked peptides are performed for cross-linked ubiquitin digests mixed with yeast cell lysates. Cross-linked peptides are successfully detected and identified by collision induced dissociation (CID) and electron transfer dissociation (ETD) in LTQ-FTICR and LTQ-Orbitrap mass spectrometers. The observed cross-linked lysine residues and their connectivities are consistent with the high resolution X-ray crystallography structure of ubiquitin. The application of CXLs in more complex systems (*e.g. in vivo* cross-linking) is illustrated by detection of Cul1 complexes in HEK 293 cells, confirming good water solubility and cell permeability.

KEYWORDS Chemical Cross-linkers, Cross-linked Peptide, Protein Structure, Protein-Protein Interaction, Mass Spectrometry, Copper-Catalyzed Azide-Alkyne Cycloaddition, Reporter Ion.

5.1.Introduction

Elucidation of three-dimensional structures of protein complexes and protein-protein interactions (PPIs) is one of the central goals in current biological research. Proteins bind to each other to carry out specific biological functions by forming various protein complexes.,^{225,226} On average, proteins *in vivo* act not alone but rather as part of a protein complex comprising 10 protein subunits in the cell.²²⁷ Proteolysis by poly-ubiquitination is a good example for a functional protein complex.²²⁸ The development of new reagents and methods for identification of binding partners and their interfaces is an enabling part of proteomic science.

Chemical cross-linkers have been widely employed in analysis of three-dimensional protein structures and protein-protein interactions (PPIs).^{229,230} For identification of cross-linked proteins, traditional experimental methodologies including affinity-based chromatography and Western blot have been performed. However, no detailed structural information relative to the nature of specific protein interfaces is revealed in these experiments. Full atomistic structures of isolated proteins and their complexes can be obtained from NMR spectroscopy and X-ray crystallography but these methodologies usually require large amounts of sample for analysis. Crystallization of diffraction quality protein complexes is often the bottleneck in structure determination by X-ray crystallography.

Recently, mass spectrometry (MS)-based analysis has allowed detection of binding partners and specific contacting residues in more sensitive ways.^{21,23,24,231-235} *In vitro* cross-linking and enzymatic digestion produce cross-linked peptides containing spatial information between residues reactive with the cross-linker. This topological information constrains relative distances of amino acid residues, thus aiding in the reconstruction of protein complex subunits.

For investigation of *in vivo* PPIs, protein complex immunoprecipitation (*i.e.*, co-IP or “pull-down”) is often performed to recover strongly interacting partners, such as an enzyme bound to its inhibitor. Co-IP requires the use of several antibodies to validate putative binding partners by

running successive rounds of experiments. Alternatively, affinity tags can be infused into genes of target proteins to permit efficient purification from cell lysates.²³⁶ However, many of the important signaling pathways are believed to be relayed via weak interactions that occur at the outside of strongly bound core protein complexes,²²⁶ and co-IP often fails to identify those weak binding partners. Chemical cross-linking has been performed to freeze weak interactions by forming covalent bonds, and then sample analysis is usually combined with other targeted protein purification techniques.^{213,237} Some previous reports have shown promising results by taking an *in vivo* chemical cross-linking strategy for investigation of PPIs.^{237,238}

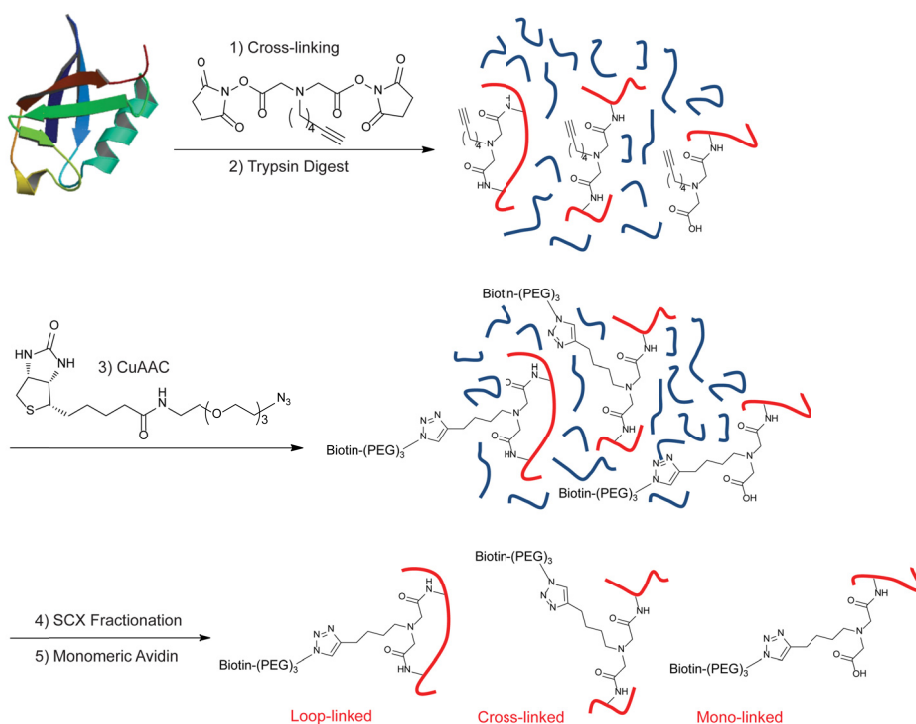
Nonetheless, most chemical cross-linking studies have been limited to purified proteins and specific targets. The formidable complexity of protein interaction networks greatly hinders identification of PPIs using cross-linking strategies at the systems level. The complexity of protein samples steeply increases upon cross-linking, roughly proportional to the half of the square of the numbers of enzymatically cleaved peptides in the cell.²³⁴ It leads to two practical problems associated with computational and experimental challenges. A protein cross-link searching program accounting for PPIs in mammalian cells on a genome-wide scale is not available, and current computational platforms are limited to just a few proteins. Computational resources required for this type of unrestricted searching algorithm are therefore substantial. A recent study done by the Aebersold group tackled this problem by introducing a new searching algorithm, xQuest, which reduces the searching space by an upstream candidate-peptide search and by isotope-coded cross-linkers.²³⁹ They successfully showed that cross-linked peptides are identified from a total *E. coli* lysate with an unrestricted database search.

For selective and sensitive detection of cross-linked peptides, functionalized chemical cross-linking reagents are required. Various designs including biotinylated²⁴⁰⁻²⁴², isotope-coded²⁴³⁻²⁴⁶, fluorophore labeled^{20,22,247}, mass-tag labeled²⁴⁸, amidinating²⁴⁹ and chromophore labeled²⁵⁰ cross-linking reagents have been reported. However, the addition of functional groups can often cause the cross-linker to become very bulky or less cell-permeable, thus not very effective for *in vivo*

cross-linking.²³⁸ To reduce the total size of the cross-linker, separation of the cross-linking step from conjugation of affinity tags is one effective strategy.

New cross-linking and enrichment strategies for separation of the cross-linking reaction from enrichment steps have recently been developed based on bio-orthogonal chemistries such as the azide-alkyne “click” cycloaddition²¹⁸⁻²²⁰ and Staudinger ligation²⁵¹ using alkyne^{252,253} or azide^{254,255} tagged cross-linkers. Azides and alkynes are not naturally found in proteins, peptides, nucleic acids or glycans. The orthogonality of azides and alkynes to biological processes (*i.e.*, competing reactions) is a significant advantage of this approach, and holds a great promise for protein cross-linking studies. Moreover, the “click” cycloaddition can be performed under aqueous conditions, allowing the enrichment of cross-linked products by conjugation of an appropriate affinity tag.

Here we report a novel clickable cross-linker (CXL) that addresses and overcomes many of the current challenges described above (**Scheme 5.1**).



Scheme 5.1

CXLs offer several advantages over previous amine-reactive cross-linkers: (1) small size, (2) good cell permeability, (3) water solubility, (4) ease of synthesis, (5) alkyne group for bio-orthogonal conjugation to a biotin-azide affinity label via copper-catalyzed azide-alkyne cycloaddition (CuAAC, or click reaction) for enrichment of cross-linked peptides by avidin-biotin chromatography, (6) a highly selective nucleophilic displacement reaction by the resultant 1,2,3-triazole in the gas phase, yielding a reporter ion at m/z 525.3 for fast screening of cross-linked peptides and (7) synthetically tunable chain length enabling the preparation of a group of CXLs to acquire enhanced structural information using various lengths as molecular “rulers”. Ubiquitin, a small and lysine-abundant protein, is chosen as our model system for three-dimensional protein structure study using a hexynyl CXL. To demonstrate the non-denaturing feature of CXLs in solution, secondary structures of cross-linked ubiquitin are monitored using circular dichroism spectrometry. After cross-linked peptides are labeled with biotin-azide by CuAAC, they are enriched by strong cation exchange fractionation (SCX) and avidin affinity chromatography. Cross-linked peptides from ubiquitin are analyzed by collisional induced dissociation (CID) and electron transfer dissociation (ETD) in LTQ-FTICR and LTQ-Orbitrap mass spectrometers. Monitoring the gas-phase reporter ion at m/z 525.3 by CID allows fast screening of results with high confidence. Highly charged cross-linked peptide ions are observed due to basic residues in CXL (*i.e.* tertiary amine and the 1,2,3-triazole), yielding high quality ETD spectra and improved sequencing. The application of CXL in more complex systems is also tested *in vivo* by cross-linking of Cul1, a ubiquitin ligase E3, in HEK293 cells, confirming good water solubility and cell permeability.

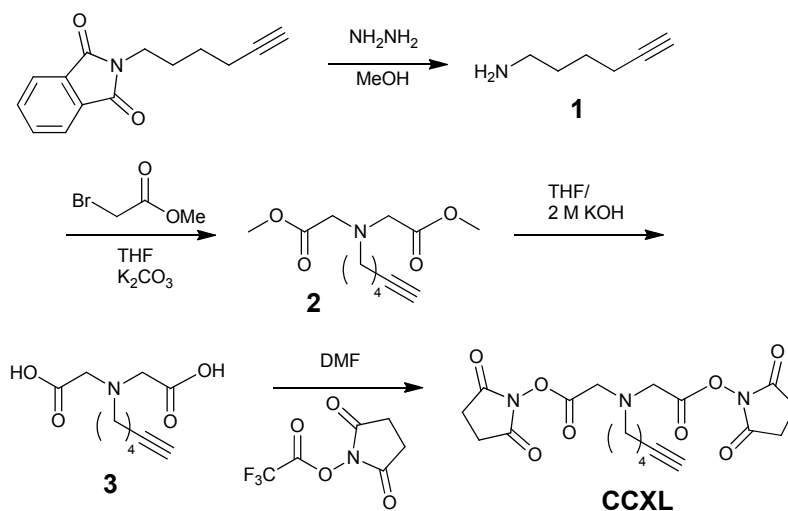
5.2. Experimental Section

5.2.1. Materials

N-(5-Hexynyl)phthalimide, anhydrous diethyl ether (Et₂O), lithium aluminum hydride (1 M in Et₂O), *N*-hydroxysuccinimide, methyl bromoacetate, trifluoroacetic anhydride, tetraethylammonium bicarbonate, tris(2-carboxyethyl)phosphine, and bovine ubiquitin were obtained from Sigma-Aldrich Co. (St. Louis, MO). The hydrophilic ligand, tris[(hydroxyethyl-triazolyl)methyl]amine (TBTA-OH) was a kind gift from the Sharpless and Fokin groups at The Scripps Research Institute. The model peptide, Ac-AAKAAAAAKAR (98% purity), was obtained from Biomer-Tech (Pleasanton, CA). Yeast cell lysate was a generous gift from the Coon group at the University of Wisconsin-Madison. Monomeric avidin resin kit was purchased from Pierce (Rockford, IL). Biotin-(PEG)₃-azide was acquired from Berry & Associates, Inc. (Dexter, MI). Sequencing grade trypsin was obtained from Promega (Madison, WI). OMIX-C18 desalting tips (100 µL capacity) were purchased from Varian, Inc. (Palo Alto, CA). Microcon YM-3K spin filter units and C18-ZipTip (10 µL scale) were purchased from Millipore (Billerica, MA). Microspin SCX columns were acquired from the Nest Group, Inc. (Southborough, MA). High Capacity Neutravidin Agarose Resin, n-dodecyl-β-D-maltoside, and SuperSignal West Dura Extended Duration chemiluminescent substrate were from Thermo Scientific (Rockford, IL). Lysyl endopeptidase (LysC) was from Wako Chemicals USA (Richmond, VA). Cell culture reagents, Flip-In T-REx 293 cells, plasmids, and monoclonal antibodies for Cull1 and Cand1 were from Invitrogen (Carlsbad, CA). Plasmid DNA containing the human Cull1 sequence was purchased from Open Biosystems (Huntsville, AL). MLN4924 was a generous gift from Millennium: The Takeda Oncology Company (Cambridge, MA). All other general chemicals for buffers were purchased from Fisher Scientific (Hampton, NH), EMD (Gibbstown, NJ), VWR International (West Chester, PA), and Sigma-Aldrich (St. Louis, MO). All chemicals mentioned above were used as received without further purification.

5.2.2. Synthesis of Clickable Cross-Linker (CXL)

Overall synthetic steps are summarized in **Scheme 5.2**.



Scheme 5.2

5.2.2.1. 6-Amino-hex-1-yne (**1**)

The amino alkyne was prepared from 2-(hex-5-ynyl)isoindoline-1,3-dione and hydrazine as described in the literature without modification.²⁵⁶ The crude product was purified by flash chromatography (silica gel, dimethylene chloride:methanol = 3:1 ~ 1:1) to yield 6-amino-hexyne as a pale greenish yellow oil. Yield: 30%. ^1H NMR spectra is reproduced as reported.²⁵⁷ ESI-MS $[\text{M}+\text{H}]^+ m/z$ 98.1.

5.2.2.2. Dimethyl 2,2'-(hex-5-ynylazanediyl)diacetate (**2**)

The 6-amino-hex-1-yne (0.3 g) was added to the stirring solution of 20 mL THF, 2 eq K_2CO_3 , and 2.4 eq methyl bromoacetate. The mixture was further stirred at room temperature for 3 h under a stream of dry N_2 . The reaction was monitored by thin layer chromatography (TLC) using hexanes:ethyl acetate (= 2:1) as the mobile phase, and the mixture was filtered after completion of the reaction. The filtrate was concentrated and purified by flash chromatography (silica gel, hexanes:ethyl acetate = 1:1). The final product, dimethyl 2,2'-(hex-5-ynylazanediyl)diacetate, was

concentrated by rotary evaporation and acquired as a transparent oil. Yield: 59%. ESI-MS $[M+H]^+$ m/z 242.1, 1H NMR ($CDCl_3$) δ 3.71 (s, 6H), 3.56 (s, 4H), 2.73 (t, 2H), 2.22 (m, 2H), 1.94 (t, 1H), 1.57 (m, 4H).

5.2.2.3. *2,2'-(hex-5-ynylazanediyl)diacetic acid (3)*

To the obtained ~1.4 g dimethyl 2,2'-(hex-5-ynylazanediyl)diacetate was added 20 mL THF and 20 mL of 2 M KOH. The mixture was stirred overnight and monitored by TLC. The organic layer was separated, and the aqueous layer was quenched by addition of 20 mL of 2 M HCl. The solvent (H_2O) was completely removed by rotary evaporation, and the resulting solid was dissolved into acetonitrile (ACN). The insoluble KCl salt was filtered, and the filtrate was concentrated by rotary evaporation. The final product of free acid was obtained as a greasy transparent oil. Yield: quantitative. ESI-MS: $[M+H]^+$ m/z 214.1. To obtain the hydrochloride salt, an additional 10 mL of 2 M HCl was added before removal of the solvent. Dimethyl formamide (DMF, 3×30 mL) was added to the resulting solid and filtered. The hydrochloride salt was obtained as a white solid after concentration under reduced pressure. Yield: quantitative.

5.2.2.4. *NHS-activated 2,2'-(hex-5-ynylazanediyl)diacetic acid (CCXL)*

N-hydroxysuccinimide trifluoroacetate was prepared by stirring *N*-hydroxysuccinimide (NHS) and 4 eq trifluoroacetic anhydride for 5 h. The mixture was concentrated under reduced pressure and further dried under high vacuum overnight. The product was obtained as a white, highly hygroscopic solid and stored in an anhydrous desiccator before use. The obtained ~1.1 g 2,2'-(hex-5-ynylazanediyl)diacetic acid was activated by 2.4 eq *N*-hydroxysuccinimide trifluoroacetate in 10 mL anhydrous DMF under a stream of dry N_2 . The mixture was stirred overnight and monitored by TLC using hexanes:ethyl acetate (= 2:1) as the mobile phase. After completion of the reaction, the mixture was concentrated to ~500 μ L by rotary evaporation and subjected to flash chromatography using hexanes:ethyl acetate (= 2:1) as the mobile phase. The

final product, NHS-activated 2,2'-(hex-5-ynylazanediy)diacetic acid (CCXL) was concentrated by rotary evaporation, and obtained as a pale yellow oil. Several 200 μ L aliquots of 50 mM stock solution dissolved in anhydrous dimethyl sulfoxide (DMSO) were prepared and stored at -80°C . The sealed stock aliquots were opened immediately before use, and NHS activation was verified by ESI-MS in 100% ACN. ESI-MS $[\text{M}+\text{H}]^{+}$ m/z 408.1, $[\text{M}+\text{Na}]^{+}$ m/z 420.0. (Note: the reaction yield can be improved by adding stoichiometric equivalent of triethylamine.)

5.2.3. Cross-Linking of a Model Peptide

A 50 μ g portion of the model peptide, Ac-AAKAAAAAKAR (98% purity), was dissolved in 50 μ L of HPLC grade H_2O . A mixture of 5 μ L CXL stock solution (10 $\mu\text{g}/\mu\text{L}$ in DMSO), 5 μ L Ac-AAKAAAAAKAR stock solution (10 $\mu\text{g}/\mu\text{L}$), and 15 μ L ACN was prepared. The mixture was allowed to react at room temperature for 1 h. The reaction was terminated by adding 5 μ L formic acid (FA). The solvent was completely removed by speed-vac, and the residue was reconstituted in 100 μ L of 0.1% FA (aq) with additional 2 μ L FA to further acidify. The resulting solution was desalted using an OMIX-C18 tip (100 μ L capacity) following the standard procedure. The cross-linked peptide (~ 50 μ g) was eluted in 100 μ L solution of 0.1% FA, 50% ACN, and 50% H_2O , and 5 μ L of the eluted cross-linked peptide solution was diluted to 5 μM by 0.1% FA, 50% ACN, and 50% H_2O , and analyzed by a LCQ ion trap mass spectrometer. The remaining cross-linked peptide solution was dried for click reaction.

CuAAC with the biotin-(triethyleneglycol)-azide (biotin-(PEG)₃-azide) was performed as follows: Ten μ g of the CXL cross-linked Ac-AAKAAAAAKAR peptide were dissolved in 100 mM tetraethylammonium bicarbonate (TEAB) at pH 8.5, 250 μM hydrophilic ligand tris[(hydroxyethyl-triazolyl)methyl]amine (TBTA-OH), 2.5 mM $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 5 mM tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, and 1 mM biotin-(PEG)₃-azide in a total volume of 100 μ L containing 99% H_2O and 1% DMSO (from the TBTA-OH stock). The mixture was reacted for 2 h with gentle mixing at 37°C and quenched by 5% FA (aq). The solvent was

removed by speed-vac, and the residue was desalted by OMIX-C18 tip as described above. The eluent was properly diluted to 10 μ M with 0.1% FA, 50% ACN, and 50% H₂O and directly infused to the LCQ ion trap mass spectrometer for analysis.

5.2.4. Cross-Linking of Ubiquitin

The general reaction summary of cross-linking of ubiquitin (PDB ID: 1UBQ) is shown in **Scheme 5.1**. Twenty μ g of ubiquitin were dissolved in 200 μ L of 1X PBS (pH = 7.4) and 1.2 μ L of 50 mM CXL stock solution in DMSO was added and reacted for 30 min at room temperature. The reaction was quenched by 50 μ L of 100 mM Tris-HCl buffer (pH = 8.5) and incubated for 15 min. The cross-linked ubiquitin was concentrated to $\sim$ 30 μ L, and the buffer was exchanged to 100 mM ammonium bicarbonate at pH 8.5 using Microcon YM-3K spin filter units. The trypsin digest reaction volume was adjusted by adding 185.5 μ L of 100 mM ammonium bicarbonate buffer containing 2 M urea, and 2.5 μ L of 100 mM CaCl₂. Two μ L of 0.5 μ g/ μ L trypsin in 5 mM acetic acid (proteins:trypsin = 20:1 w/w) was added and incubated for 15-18 h at 37 °C. The reaction was terminated by addition of 5% FA (aq). The resulting tryptic digest was desalted by OMIX-C18 tip and a 1 μ g portion was injected into a nanoLC-LTQ-FTICR mass spectrometer for analysis.

Forty μ g of the cross-linked tryptic digest of ubiquitin were subjected to click reaction by combining resulting peptides from two identical cross-linking experiments. The desalted tryptic digest was dissolved in 100 mM TEAB, 250 μ M hydrophilic ligand TBTA-OH, 2.5 mM CuSO₄•5H₂O, 5 mM TCEP hydrochloride, and 1 mM biotin-(PEG)₃-azide in a total volume of 100 μ L containing 99% H₂O and 1% DMSO (from TBTA-OH stock). An additional sample was prepared by mixing 50 μ g cross-linked digest of ubiquitin with 50 μ g yeast cell lysate and subjected to similar click reaction conditions to demonstrate enrichment from a complex sample. Mixtures were reacted at 37 °C for 12 h with gentle shaking. Reactions were quenched by addition of 5% FA (aq).

Microspin SCX columns (200 μ L scale, with 50 μ L of the bed volume for SCX material) were used for removal of excessive TBTA-OH and biotin-(PEG)₃-azide. A 10 μ g portion of the peptides from the click reaction (25 μ L) was dried to completeness by speed-vac, and the residue was reconstituted with 0.5% FA, 5% ACN (aq). Microspin SCX columns were prepared by applying 4 bed volumes (200 μ L) of MeOH then H₂O, respectively. Activation of the SCX material was performed by 200 μ L of 500 mM ammonium acetate and incubated for 1 h at room temperature. After activation, the spin columns were washed by H₂O and equilibrated with 0.5% FA, 5% ACN (aq). The peptide sample solution was applied to the spin column and flushed twice to bind completely. The spin column was washed by 400 μ L of 0.5% FA, 5% ACN (aq), which corresponds to at least 8 bed volumes of the SCX material. The peptides were fractionated by 400 μ L of 50, 250, and 500 mM ammonium acetate in 0.5% FA, 25% ACN (aq), and additional 500 mM ammonium acetate solution was used for complete elution of highly charged cross-linked peptides. Each fraction was desalted by C18-ZipTip following manufacturer instructions, and eluents were dried by speed-vac. The residues were reconstituted with 5 μ L of 0.2% FA (aq) and injected to a nanoLC-LTQ-FTICR mass spectrometer for analysis.

Monomeric avidin-biotin affinity chromatography was performed using the batch style procedure according to the manufacturer manual with modifications. Peptide samples after click reaction were eluted without SCX fractionation using either 500 mM ammonium acetate in 0.5% FA, 25% ACN (aq) or 50 mM ammonium acetate in 0.1% TFA, 25% ACN (aq). The SCX eluents were dried by speed-vac and readjusted to 1X PBS at the same concentration range used in the cross-linking reaction. By incubating the mixture of the monomeric avidin resin and peptide samples at room temperature or 4 °C for 12 h under the gentle mixing, the biotin-PEG₃-azide conjugated peptides were bound to monomeric avidin. Unmodified peptides were washed away by flushing 4 bed volume capacity of PBS, 100 mM Tris-buffer (pH 7.4), 100 mM ammonium bicarbonate (pH 7.4), and water. The final products of interest were eluted by 0.4%

TFA, 50% ACN (aq). An aliquot of the eluent was analyzed by a nanoLC-LTQ-FTICR mass spectrometer.

5.2.5. *In vivo* Cross-Linking of Cul1

The applicability of CXL to *in vivo* cross-linking was evaluated with Cul1. Cul1 is a ubiquitin ligase that attaches a ubiquitin chain on target substrates for proteasome-catalyzed degradation.²¹¹ Cul1 is a prototype of the cullin ligase family, and constitutes modular ligase complexes with other binding partners. The *in vivo* cross-linking by CXL and following Western blot analysis were carried out as described previously with minor modifications.²¹² Briefly, to facilitate the purification of Cul1, a HEK 293-derived stable cell line capable of expressing tagged Cul1 upon tetracycline treatment was constructed using the T-RExTM (Tetracycline-regulated Expression) system (Invitrogen, Carlsbad, CA). The tandem tag we adopted consists of a hexa-histidine sequence and a biotinylation signal sequence.²¹³ Biotinylation is catalyzed by endogenous biotin ligases, which are present in all eukaryotic cells.²¹⁴ A specific lysine residue in the biotinylation signal sequence functions as an acceptor site for biotin *in vivo*.²¹⁵

Tagged Cul1 was induced with 1.0 µg/mL tetracycline for 4 h in experiments for *in vivo* cross-linking. Twenty four hours after induction, cells were subject to *in vivo* cross-linking by treating 0, 0.1, 0.2, 0.5, and 1.0 mM of CXL, respectively and incubated for 1h at 37 °C. After the completion of the cross-linking, cells were lysed for 30 min at 4 °C with the lysis buffer (0.050 M HEPES, pH 7.5, 0.0050 M Mg(OAc)₂, 0.070 M KOAc, 10% glycerol, and 0.4% IGEPAL CA630). The lysate was centrifuged at 16,600 g at 4 °C for 20 min and the supernatant was used for Western blot analysis.

5.2.6. Mass Spectrometry

The CXL cross-linked Ac-AAKAAAAAKAR model peptide was analyzed by a LCQ-deca XP ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The sample was directly

infused by the standard electrospray ionization unit with the constant flow at 3 $\mu\text{L}/\text{min}$. The critical instrumental parameters were set up as follows: the spraying voltage at 3.0 kV, capillary voltage at 25 V, capillary temperature at 200 $^{\circ}\text{C}$, and tube lens voltage at -25 V. Fifty scans were recorded for each spectrum.

The cross-linked peptides from ubiquitin were analyzed by a nanoflow HPLC (Waters Co.) coupled on-line via a home-built nanoelectrospray ion source to a LTQ-FTICR mass spectrometer (Thermo Fisher Scientific). Samples in 5 μL of 0.2% FA (aq) were loaded onto a C_{18} -reversed phase column (15 cm long, 100 μm inner diameter, packed in-house with Magic C_{18} -AQ 5 μm resin (Michrom Bioresources) in buffer A (2% ACN, 0.2% FA) with a flow rate of 250 nL/min for 24 min and eluted with a linear gradient from 0% to 36% buffer B (98% ACN, 0.2% FA) over 110 min, followed by 10 min at 100% buffer B, at a flow rate of 250 nL/min . The column was re-equilibrated with buffer A. Mass spectra were acquired in the positive ion mode applying data-dependent acquisition with automatic switching between survey scan and tandem mass spectrum acquisition. Samples were analyzed with a top 10 method; acquiring one FTICR survey scan in the mass range of m/z 400-1600 followed by MS/MS of the ten most intense ions in the LTQ. The target ion value in the LTQ-FTICR was 500,000 for survey scan at a resolution of 50,000 at m/z 400. Fragmentation in the LTQ was performed by CID with a target value of 5,000 ions. Selected sequenced ions were dynamically excluded for 30 s. Critical mass spectrometric parameters were: spray voltage, 2.4 kV; no sheath and auxiliary gas flow; ion transfer tube temperature, 200 $^{\circ}\text{C}$; normalized collision energy (35%) for MS/MS.

5.2.7. Circular Dichroism Spectrometry

The CXL cross-linked ubiquitin samples were analyzed by an Aviv Model 62A DS Circular Dichroism (CD) Spectrometer (Lakewood, NJ) at room temperature. The critical instrumental parameters included: acquisition range = 200 to 300 nm, step size = 1.00 nm, averaging time = 5 s, bandwidth = 1 nm, and path length = 1 mm. One scan was recorded for each spectrum. For

CCXL cross-linked samples, 1 to 4 μL of 50 mM CCXL stock solution in DMSO was added to 196 to 199 μL of 0.1 mg/mL ubiquitin in 1X PBS pH 7.4 to result in a final concentration of 0.25 to 1 mM CCXL in 200 μL . CD spectra were collected upon addition of CCXL ($t = 0$) and after 30 min of cross-linking reaction ($t = 0.5$ h).

5.2.8. xQuest Search

The raw files from the LTQ-FTICR mass spectrometer were converted to Mascot generic format (MGF) files using ReAdW4Mascot (version 20090305a, available from the National Institute of Standards and Technology at http://peptide.nist.gov/software/ReAdW4Mascot2_20090305a.zip), and all but the 150 most intense fragment ions were filtered out. The cross-linked peptide search was performed using xQuest (<http://prottools.ethz.ch/orinner/public/htdocs/xquest>).²³⁹ The database search parameters were as follows: 0.5 Da fragment ion mass tolerance; 0.3 Da common fragment ion mass tolerance; 10 ppm precursor ion mass tolerance; trypsin enzyme specificity (up to two missed cleavages); fixed carbamidomethyl (57.02146 Da) modification of cysteine; and variable modifications of methionine oxidation (15.99491 Da), cross-linked primary amines (177.07898 Da), mono-linked dead ends (195.08954 Da), cross-linked and biotin-(PEG)₃-azide clicked primary amines (621.295568 Da), mono-linked and biotin-(PEG)₃-azide clicked primary amines (639.305584 Da), and reporter ion (525.28537 Da). Both reporter ion-filtered and non-filtered MGF files by xQuest search were tested, but identical results were reported. Reduced MS/MS scans for xQuest search resulted in faster searching for the reporter ion-filtered MGF files.

5.2.9. X-ray Crystal Structure Analysis

Human ubiquitin (PDB ID: 1UBQ. Note that the sequences of human and bovine ubiquitin are identical.) structure was analyzed by UCSF Chimera (version 1.5.3rc).²⁵⁸ For each pair of cross-

linked lysine residue, the distances between the alpha carbons (C_{α} – C_{α}) and between nitrogens of ϵ -amines (NZ–NZ) were measured.

5.3. Results and Discussion

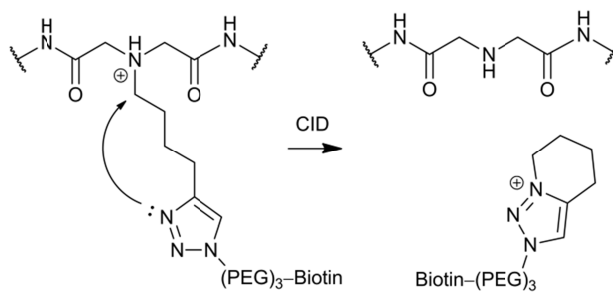
Before presenting our development to address the problems described above, it is worthy to review general aspects of chemical cross-linking strategies used in this study and structures of peptide products resulting from cross-linking. Structurally, there are three types of cross-linked peptides: intermolecular cross-linked, intramolecular cross-linked and mono-linked peptides (**Scheme 5.1**). Intermolecular cross-linked peptides (*i.e.*, cross-linking between proteins or between two residues within a protein), herein referred to as just “*cross-linked*” peptides for convenience, only exist sub-stoichiometrically compared to other *loop-linked* (*i.e.*, intramolecular modified peptides), *mono-linked* (*i.e.*, “dead-end” modified peptides) and dominantly unmodified linear peptides (**Scheme 5.1**). Among these species, only cross-linked peptides provide meaningful information for PPIs.

In this study, a novel clickable cross-linker (CXL) that incorporates an alkyne tag to enable enrichment of cross-linked peptides after coupling to an affinity tag via CuAAC is designed, synthesized, and evaluated. To demonstrate the efficiency of the new reagent, cross-linking reactions are first monitored with a model peptide, Ac-AAKAAAAAKAR (loop-linking), and the resulting cross-linked peptide is subject to CID to investigate their gas-phase fragmentation patterns. Second, cross-linking of ubiquitin as a model protein is examined by Circular Dichroism (CD) spectrometry and subject to tryptic digestion, followed by analysis of the resulting cross-linked peptides using both CID and ETD. Further evaluation of the enrichment capability of CXL is performed using CuAAC to attach biotin-(PEG)₃-azide, followed by biotin-avidin affinity chromatography to separate and identify cross-linked peptides from a digested yeast cell lysate.

Lastly, HEK293 cells are cross-linked *in vivo* and Western blot analysis of Cul1 is performed to monitor the cell permeability and water solubility of CXL.

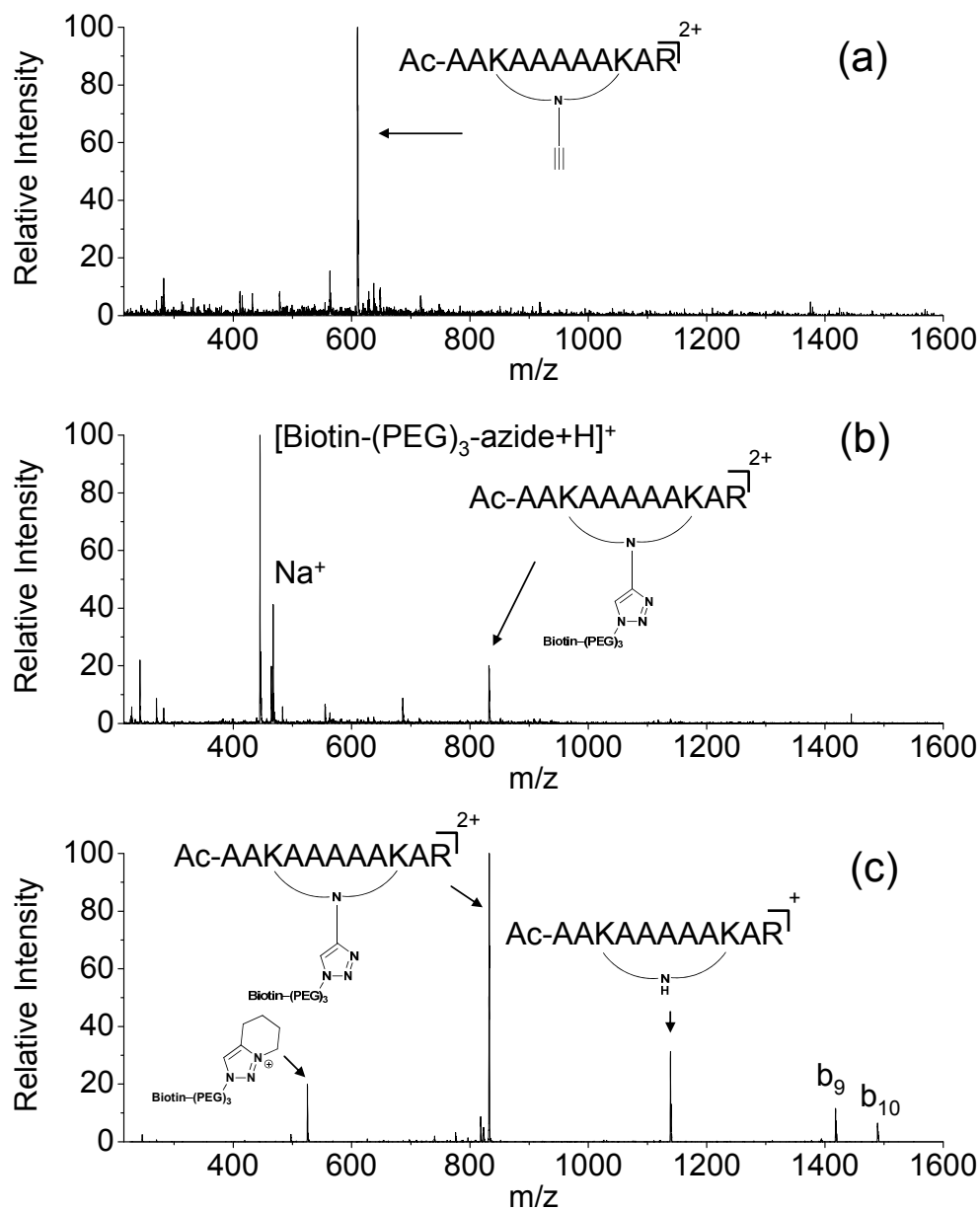
5.3.1. Model Peptide Cross-Linking

ESI-MS and CID spectra of the cross-linked model peptide Ac-AAKAAAAAKAR are shown in **Figure 5.1**. The two lysine residues in the model peptide Ac-AAKAAAAAKAR are cross-linked by CXL (m/z 609, **Figure 5.1a**). Protonation sites are expected to be the arginine side chain and the central tertiary amine in the cross-linker. The 1,2,3-triazole product from conjugation of biotin-(PEG)₃-azide via CuAAC corresponds to the doubly charged ion at m/z 832 in **Figure 5.1b**. No precursor ion (m/z 609) is observed, indicating quantitative conversion via CuAAC (**Figure 5.1b**). CID of the biotin-(PEG)₃-azide conjugated peptide dication yields two backbone fragments along with the reporter ion at m/z 525.3. This product ion is generated by a nucleophilic attack forming a six membered ring as depicted in **Scheme 5.3**.



Scheme 5.3

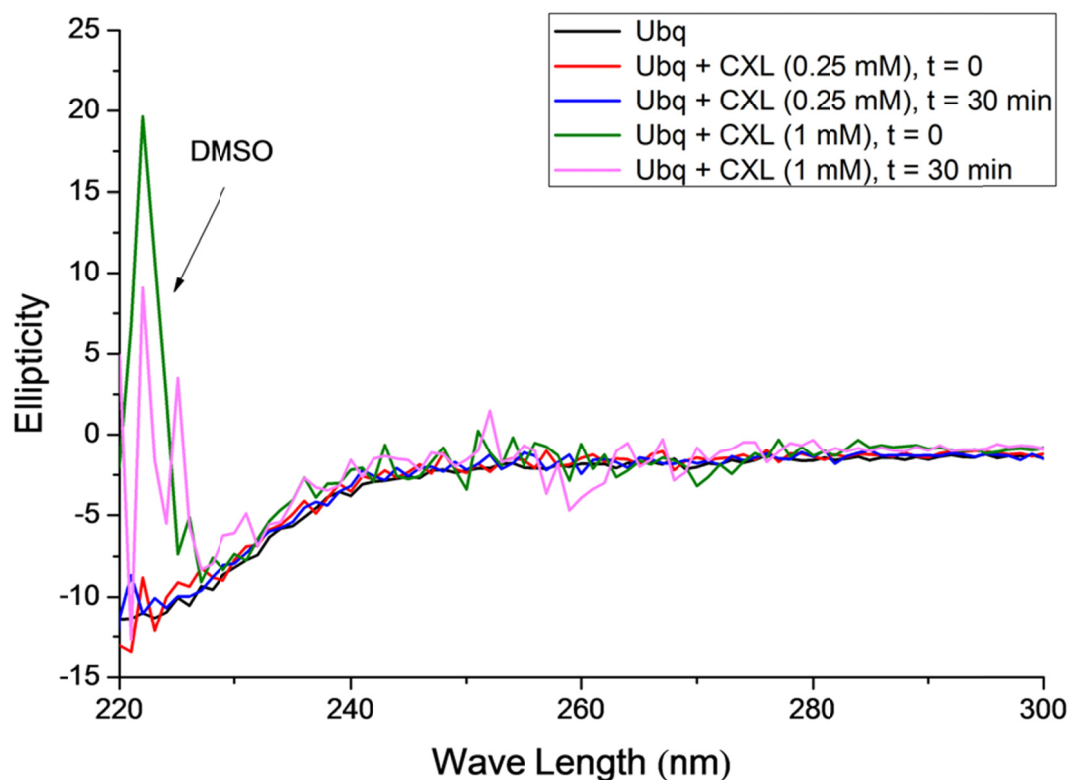
It is thus shown that the resulting b- and y-type ions generated by CID of cross-linked peptides allow the efficient sequencing of cross-linked peptide chains without ambiguity. In this regard, it is important that the fragmentation pathway of the CIT reporter ion occurs in competition with those of other b- and y-type ions, yielding rich sequence information.

Figure 5.1 ESI-MS and CID spectra of the cross-linked, clicked model peptide

The ESI-MS spectra of the cross-linked (a), and clicked (b) Ac-AAKAAAAAKAR peptide. The CID spectrum (c) of the cross-linked and clicked Ac-AAKAAAAAKAR peptide. The reporter ion at m/z 525.3 is observed along with backbone fragments in the CID spectrum. Note that in b, CuAAC with biotin-(PEG)₃-azide proceeded almost quantitatively, showing no cross-linked precursor peptide ion at m/z 609.

5.3.2. Structural Analysis by Circular Dichroism

Circular dichroism (CD) spectrometry is a very useful analytical technique for quick investigation of secondary structure, folding, and binding properties of peptides and proteins.²⁵⁹ Cross-linking reactions with proteins should be efficient without structural perturbation or denaturation under the working concentration of the cross-linker. To probe the structural effect, ubiquitin cross-linked by CXL is examined by CD spectrometry. The CD spectra of native ubiquitin were previously recorded, yielding the secondary structure analysis of 6% α -helix, 10% β -sheet and 84% random structures, which can be observed as increasingly negative ellipticity over the range of 225 to 240 nm.²⁶⁰ In this work, the CD spectra were acquired in the presence of 0.25 or 1 mM CCXL ($t = 0$) and after a duration of 30 min ($t = 30$ min). The CD spectra of CCXL cross-linked ubiquitin samples are depicted in **Figure 5.2**, showing no significant change in various conditions. These results suggest that the secondary structure is essentially unchanged by the addition of CXL or its cross-linking reaction. Therefore, the contribution of CCXL on ubiquitin secondary structure seems to be minimal or undetectable using CD spectrometry. After this confirmation, we proceeded to analyze cross-linked peptides obtained by tryptic digestion of cross-linked ubiquitin using mass spectrometers.

Figure 5.2 CD spectra of CXL cross-linked ubiquitin

The CD spectra of ubiquitin cross-linked by various concentrations and reaction times of CXL. No significant change is observed among ubiquitin samples that are native (black), cross-linked at 0.25 mM (red) or 1 mM (green) of CXL and incubated for 30 min after the initiation of cross-linking reaction (blue for 0.25 mM and pink for 1 mM CXL). Abundant peaks from 220 to 230 nm result from the increased DMSO portion in 1 mM CXL cross-linked ubiquitin samples (from 1% to 4%). The CD signal fluctuation in 1 mM CXL experiments is mainly caused by the light scattering with the increased concentrations of small molecules such as cross-linkers and DMSO.

5.3.3. Ubiquitin Cross-Linking

Ubiquitin cross-linked peptides and their conjugates with biotin-(PEG)₃-azide were analyzed by a nanoLC-LTQ-FTICR mass spectrometer, and the resulting LC/MS and MS/MS profiles were searched against xQuest. All types of loop-linked, mono-linked and cross-linked peptides (**Scheme 5.1**) that are found before click reaction are summarized in **Tables 5.1** and **5.2**. Modified lysines in mono-linked peptides indicate solvent accessible residues. Loop-linked peptides are generated by cross-linking of two lysine residues that are closely positioned and where there are no unmodified lysine or arginine residues for tryptic cleavage between them. Cross-linked peptides have one or more unmodified lysine or arginine residues between cross-linked lysine residues, yielding two peptide chains covalently joined by the cross-linker. These intermolecular cross-linked peptides constrain the topology of surface exposed lysine residues, which is of particular interest for the construction of low resolution three-dimensional protein structures.

Table 5.1 Mono- and Loop-linked Peptides from Ubiquitin

Sequence	Before Click	After Click & SCX Fractionation	Reporter Ion
¹ MQIFV ⁶ K [^] TLTG ¹¹ K	Yes, 2+, 3+	Yes, 2+, 3+, 250 mM	Yes
⁷ TLTG ¹¹ K [^] TITLEVEPSDTIENV ²⁷ K	Yes, 2+, 3+, 4+	Yes, 3+, 4+, 250 mM	Yes
¹² TITLEVEPSDTIENV ²⁷ K [^] A ²⁹ K	Yes, 3+	Yes, 3+, 250 mM	Yes
²⁸ A ²⁹ K [^] IQD ³³ K	Yes, 2+	No	N/A
²⁸ A ²⁹ K [^] IQDKEGIPPDQQ ⁴² R	Yes, 3+	No	N/A
²⁸ AKIQD ³³ K [^] EGIPPDQQ ⁴² R	Yes, 3+	Yes, 4+, 250 mM	Yes
³⁰ IQD ³³ K [^] EGIPPDQQ ⁴² R	Yes, 2+, 3+	Yes, 2+, 3+, 250 mM	Yes
⁴³ LIFAG ⁴⁸ K [^] QLEDG ⁵⁴ R	Yes, 2+, 3+	Yes, 2+, 3+, 4+, 50, 250 mM	Yes
⁴³ LIFAG ⁴⁸ K [^] QLEDGRTLSDYNIQ ⁶³ K	Yes, 2+, 3+, 4+	Yes, 3+, 4+, 250 mM	Yes
⁵⁵ TLSDYNIQ ⁶³ K [^] ESTLHLVL ⁷² R	Yes, 2+, 3+	Yes, 3+, 4+, 250 mM	Yes
¹ MQIFV ⁶ K [^] TLTG ¹¹ K [^]	Yes, 2+, 3+	Yes, 2+, 3+, 250 mM	Yes
²⁸ A ²⁹ K [^] IQD ³³ K [^] EGIPPDQQ ⁴² R	Yes, 2+, 3+	Yes, 3+, 250 mM	Yes

[^]: linked residues. Superscript numbers are the residue numbers of amino acids of ubiquitin.

Table 5.2 Cross-linked Peptides from Ubiquitin

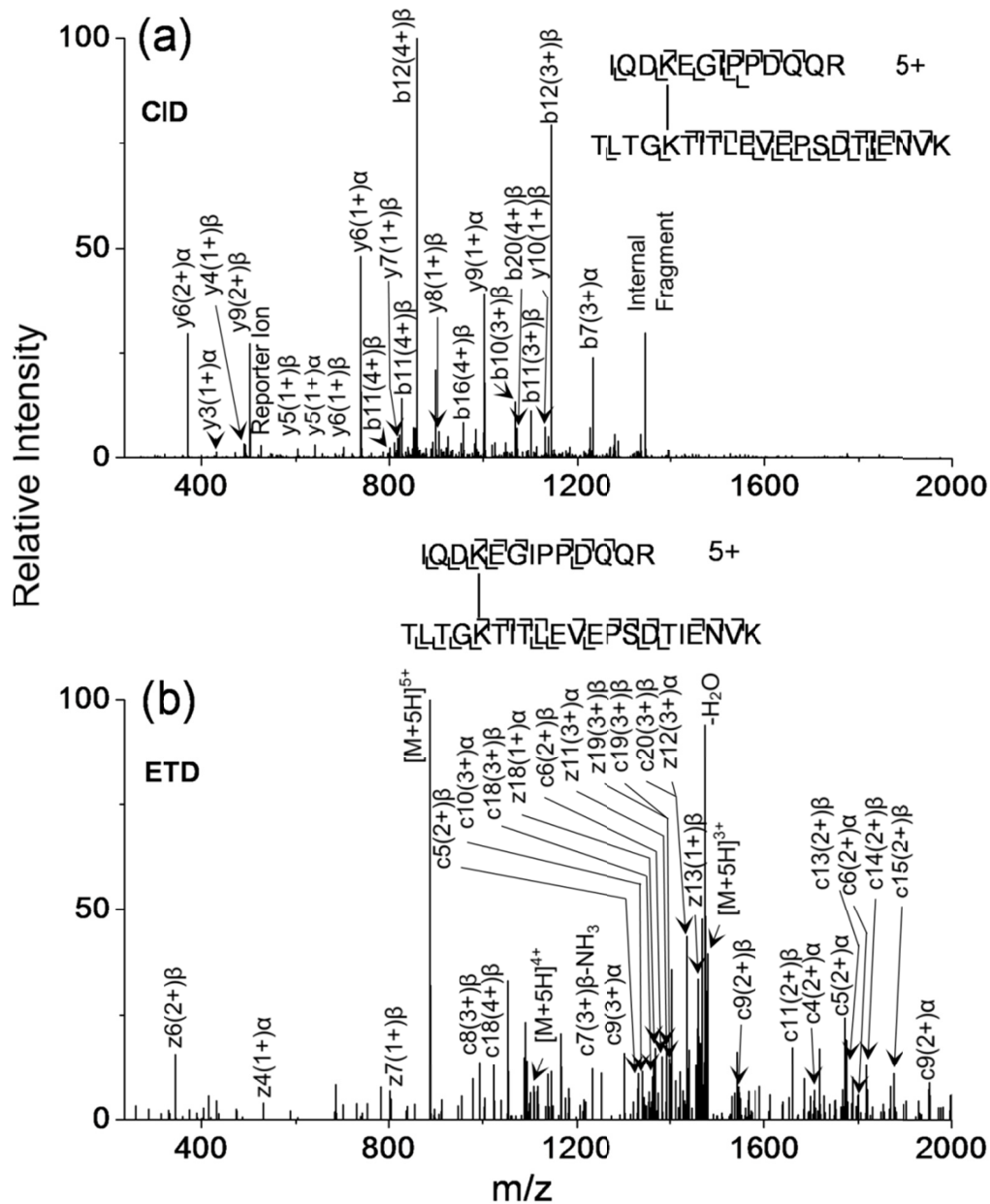
chain	chain	Before Click	After Click & SCX Fractionation	Reporter Ion	Avidin Enrichment	C _{α} distance (Å)	NZ distance (Å)
⁴³ LIFAG ⁴⁸ K [^] QLEDG ⁵⁴ R	¹ MQIFV ⁶ K [^] TLTG ¹¹ K	Yes, 3+, 4+, 5+	No	N/A	No	15.14	17.82
¹ MQIFV ⁶ K [^] TLTG ¹¹ K	⁵⁵ TLSDYNIQ ⁶³ K [^] ESTLHLVL ⁷² R	Yes, 5+, 6+	No	N/A	No	15.01	20.31
³⁰ IQD ³³ K [^] EGIPPDQQ ⁴² R	⁷ TLTG ¹¹ K [^] TITLEVEPSDTIENV ²⁷ K	Yes, 4+, 5+	Yes, 4+, 5+, 250 mM, 1st 500 mM, 2nd 500 mM	Yes	Yes, 4+, 5+	12.85	7.15
⁴³ LIFAG ⁴⁸ K [^] QLEDG ⁵⁴ R	⁴³ LIFAG ⁴⁸ K [^] QLEDG ⁵⁴ R	Yes, 3+, 4+, 5+	Yes, 5+, 2nd 500 mM	Yes	Yes, 5+	N/A	N/A
⁴³ LIFAG ⁴⁸ K [^] QLEDG ⁵⁴ R	⁵⁵ TLSDYNIQ ⁶³ K [^] ESTLHLVL ⁷² R	Yes, 6+	No	N/A	No	17.89	19.85
²⁸ A ²⁹ K [^] IQD ³³ K	³⁰ IQD ³³ K [^] EGIPPDQQ ⁴² R	No	No	Yes	Yes, 5+	6.24	9.09

^: cross-linked residues. Superscript numbers are the residue numbers of amino acids of ubiquitin.

5.3.4. CID and ETD of Cross-Linked Peptides

Fragmentation of cross-linked peptides is investigated by CID and ETD. As an example, the CID spectrum of the 5+ charged $^{30}\text{IQD}^{33}\text{K}^{\text{EGIPPDQQ}}^{42}\text{R}-^7\text{TLTG}^{11}\text{K}^{\text{TITLEVEPSDTIENV}}^{27}\text{K}$ ion is shown in **Figure 5.3**. Abundant b- and y-type ions are generated and cover many sequences in each peptide chain. The diagnostic reporter ion is found at m/z 525.3 without any ambiguity, and it does not overlap with other backbone fragments. ETD of the 5+ charged peptide ion also produces many c- and z-type ions due to its high charge state (**Figure 5.3**). It should be noted that the charge-reduced molecular ion species are less abundant compared to the ETD spectra of the previously reported cross-linker by Chowdhury *et al.* (Click-enabled Linker for Interacting Proteins, or CLIP).²⁵² With CLIP, the nitro group (NO_2) is inserted for water solubility, and neutral loss of NO_2 by CID can be used as a diagnostic peak. In ETD, the high electron affinity of the nitro group can initially trap a transferred electron, followed by proton transfer.²⁶¹ The resulting nitronic radical stabilizes the charge-reduced species and prevents further fragmentation to form c- and z-type ions, yielding abundant charge-reduced species. With CXLs, there are no occurrences of specific chemical bonds or residues that can generate unexpected neutral losses or stable charge-reduced species in ETD, so efficient electron-based dissociation processes dominate. Generally, CuAAC with biotin-(PEG)₃-azide increases the charge states of peptides (**Tables 5.1** and **5.2**). The tertiary amine and 1,2,3-triazole ring moderately increase the overall proton affinity for all types of cross-linked peptides. This feature apparently provides highly charged precursor ions for ETD with augmented fragmentation yields.

Note that all clicked peptides yield the reporter ion at m/z 525.3 by CID. Pre-screening of MS/MS scans by this particular ion provides identical searching hits as compared to the non-filtered MS/MS scan searching for every type of mono-, loop- and cross-linked peptides in *all experiments* performed in this work. Prefiltering of MS/MS scans can significantly reduce the requirement of computational resources, especially for systems level database searching, by

Figure 5.3 CID and ETD of the 5+ charged cross-linked peptide

CID (a) and ETD (b) of the 5+ charged cross-linked peptide, $IQD^{\wedge}KEGIPPDQQR$ - $TLTG^{\wedge}KTITLEVEPSDTIENVK$ where $^{\wedge}K$ is the cross-linked residue. (a) Although the molecular weight (4431 Da) of the cross-linked peptide is relatively large to be fragmented by CID, good coverage of b- and y-type ions is observed. The reporter ion at m/z 525.3 is also observed with acceptable yield. (b) The high charge state of the cross-linked peptide yields abundant c- and z-type ions in the ETD spectrum. This shows potential for the general usage of CID and ETD for efficient sequencing of cross-linked peptides.

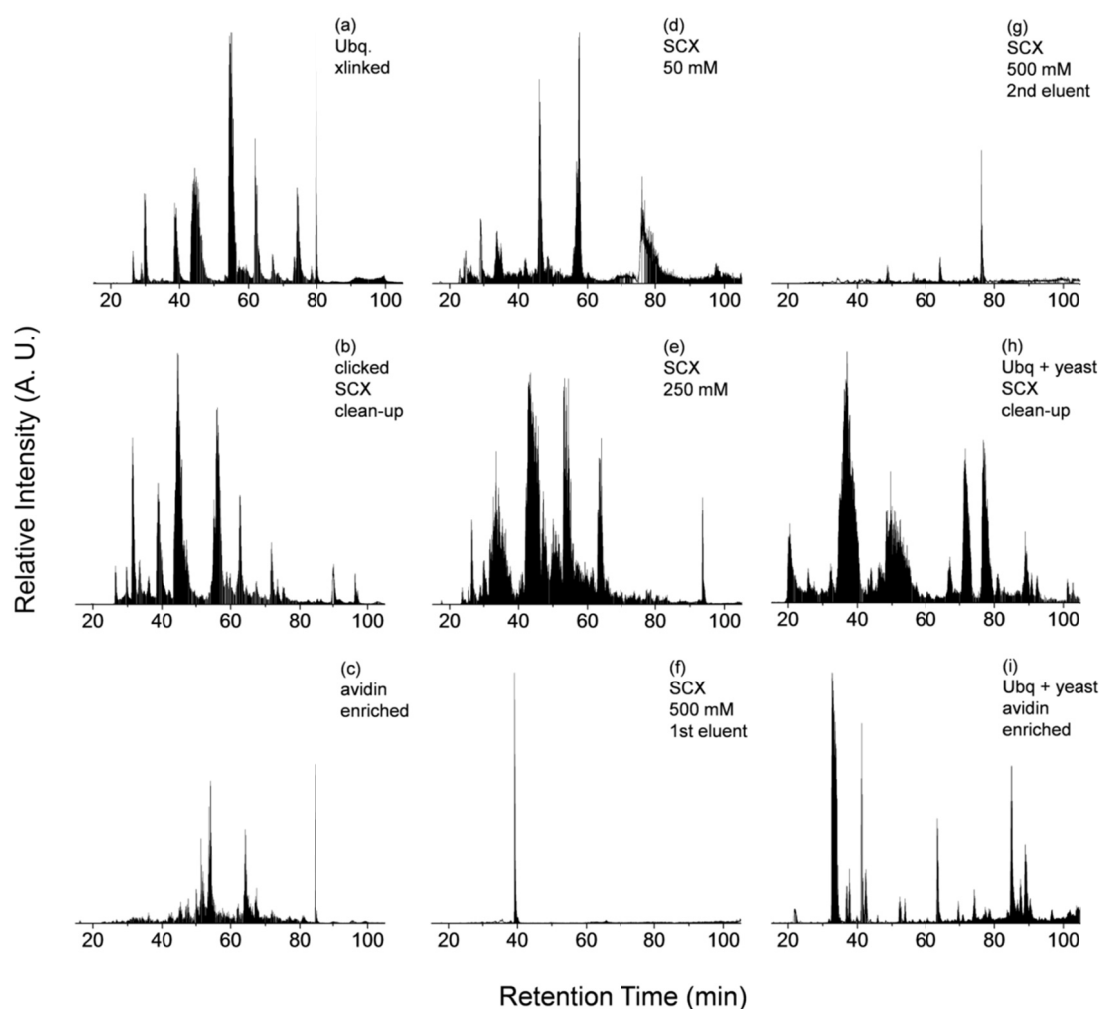
cutting down the number of candidate MS/MS spectra. This robust diagnostic reporter ion offers a technical improvement over most of the previously reported cross-linkers and affinity enrichment schemes.

5.3.5. Sample Clean-up Following Click Reaction

CuAAC is a widely used bioconjugation reaction. However, seamless integration of CuAAC into the downstream proteomics workflows can be challenging due to the persistence of residual chemical reagents such as Cu^{2+} ions, ligands (*e.g.*, TBTA), and coupling reagents (azide or alkyne).²⁵⁵ Those impurities often adversely impact the ionization efficiency of target peptides. Effective sample clean-up procedure after CuAAC is therefore very important for successful sample analysis by MS.

For this purpose, SCX, followed by C18 desalting, is the adopted clean-up procedure for the removal of non-ionic species following CuAAC. The hydrophobic TBTA reagent also has been replaced by its hydrophilic counterpart, TBTA-OH. Poor binding of TBTA-OH to the C18 matrix would result in its effective removal via C18 desalting column. Copper ions can be removed during the same desalting step.

Figure 5.4 depicts LC-MS profiles of ubiquitin cross-linked peptide samples following CuAAC. TBTA-OH is not detected in any LC-MS profile, confirming its successful removal. Excess biotin-(PEG)₃-azide molecules are mostly removed by SCX clean-up due to its poor interaction with SCX matrix (**Figure 5.4**), but not completely eliminated. The protonated biotin-(PEG)₃-azide (m/z 445.2) is eluted along with peptides (Retention Time around 55.17 min in **Figure 5.4b** and 46.22 min in **Figure 5.4d**), but the mass to charge ratio does not overlap with those of other cross-linked peptides, ensuring no disturbance on the sequencing. Based on its ion signal, the residual amount of biotin-(PEG)₃-azide after SCX clean-up seems to be less significant, and appears not to suppress peptide ionization (**Figure 5.4b**). For complete removal of the residual azide affinity tags, cleavable biotin tags can be employed along with streptavidin

Figure 5.4 LC-MS total ion current (TIC) chromatograms of cross-linked peptides

LC-MS TIC chromatograms of (a) cross-linked ubiquitin digest, (b) cross-linked, biotin-azide clicked ubiquitin digest after SCX clean-up without fractionation, (c) cross-linked, biotin-azide clicked, avidin enriched ubiquitin digest, (d) 50, (e) 250, (f) 1st 500, and (g) 2nd 500 mM NH_4OAc SCX eluents of cross-linked, biotin-azide clicked ubiquitin digests, (h) biotin-azide clicked 1:1 mixture of cross-linked ubiquitin and yeast cell lysate by weight after SCX clean-up without fractionation, and (i) avidin-enriched 1:1 mixture of cross-linked ubiquitin and yeast cell lysate.

magnetic resin, which suffer less nonspecific bindings.²⁶² This may allow us to perform much clean elution via chemical cleavage of affinity tags.

5.3.6. Peptide Fractionation by SCX

Highly charged species tend to more strongly bind to SCX matrix. Therefore, SCX can be used for sample fractionation by discriminating the charge states of analytes. Primary sample fractionation by SCX was demonstrated for sensitive detection of cross-linked peptides from abundant linear peptides by Rinner *et al.*²³⁹ In this report, peptide fractionation of ubiquitin cross-linked peptides was performed by sequential increase of the salt concentration during the SCX elution step. **Figure 5.4d-g** show LC-MS total ion current (TIC) chromatograms of differentially eluted ubiquitin cross-linked peptide samples by applying salt gradients, 50 mM, 250 mM, 1st 500 mM and 2nd 500 mM ammonium acetate, 0.5% FA, respectively (**Figure 5.4d-g**), or direct elution using 500 mM ammonium acetate, 0.5% FA with no fractionation (**Figure 5.4b**) during the SCX elution steps. As summarized in **Tables 5.1** and **5.2**, highly charged cross-linked peptides are eluted in the high concentration region. However, some of the cross-linked peptides are also co-eluted with other linear peptides at 250 mM salt concentration (**Table 5.2**, **Figure 5.4e**). Thus, marginal separation of cross-linked peptides is achieved by SCX fractionation. For further optimization of the separation, (an) additional elution step(s) using intermediate salt concentrations between 50 to 250 mM can be performed. Due to the low complexity of the ubiquitin cross-linked sample, SCX fractionation by itself is sufficient for separation and identification of cross-linked peptides from other linear peptides (**Figure 5.4d-g**, **Table 5.1** and **5.2**). As pointed out by Rinner *et al.*,²³⁹ however, the ultimate test for enrichment capability is found in the application of CXLs to higher complex systems such as *in vivo* cross-linking in eukaryotic cells.

5.3.7. Avidin Affinity Chromatography

The general design of CXLs and their capacity for bio-orthogonal incorporation of an affinity tag via CuAAC are validated by a biotin-avidin enrichment strategy. Here, monomeric avidin affinity chromatography was employed for enrichment of cross-linked peptides from the simple ubiquitin cross-linked sample prepared without SCX fractionation (**Figure 5.4c**). Cross-linked peptides enriched by avidin affinity chromatography are summarized in **Table 5.2**. Two cross-linked peptides $^{30}\text{IQD}^{33}\text{K}^{\wedge}\text{EGIPPDQQ}^{42}\text{R}-^7\text{TLTG}^{11}\text{K}^{\wedge}\text{TITLEVEPSDTIENV}^{27}\text{K}$, and the homodimer of $^{43}\text{LIFAG}^{48}\text{K}^{\wedge}\text{QLEDG}^{54}\text{R}$ that are observed in SCX fractionation are also detected by avidin affinity chromatography (**Table 5.2**), confirming the value of SCX fractionation as a separation technique. Interestingly, one additional cross-linked peptide, $^{28}\text{A}^{29}\text{K}^{\wedge}\text{IQD}^{33}\text{K}-^{30}\text{IQD}^{33}\text{K}^{\wedge}\text{EGIPPDQQ}^{42}\text{R}$ is identified only by avidin affinity chromatography. This result demonstrates the sensitivity of affinity-based enrichment of cross-linked peptides, maximizing their detection.

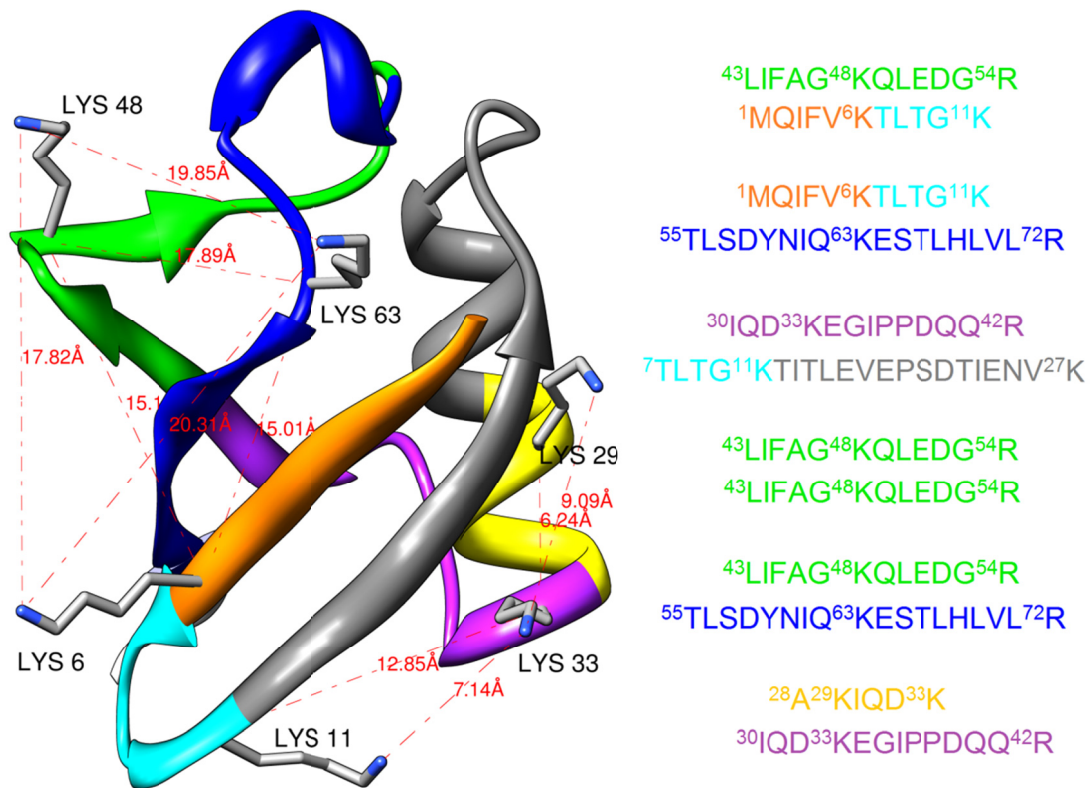
For an extreme test, a highly complex peptide sample prepared by mixing equal amounts of yeast cell lysates and ubiquitin cross-linked peptides by weight, respectively, was subjected to avidin affinity chromatography. **Figure 5.4h** and **i** depict LC-MS TIC chromatograms of the samples from SCX clean-up (**h**), and avidin enrichment (**i**) of this highly complex peptide mixture. Yeast peptides are mostly eliminated during avidin capture step, and only few are detected after enrichment. Other unmodified ubiquitin peptides that present no biotin tag are also mostly removed. The majority of the peaks in the LC-MS TIC chromatograms are singly charged impurities introduced after avidin affinity chromatography.²⁶³ The two identified cross-linked peptides $^{30}\text{IQD}^{33}\text{K}^{\wedge}\text{EGIPPDQQ}^{42}\text{R}-^7\text{TLTG}^{11}\text{K}^{\wedge}\text{TITLEVEPSDTIENV}^{27}\text{K}$, and the homodimer of $^{43}\text{LIFAG}^{48}\text{K}^{\wedge}\text{QLEDG}^{54}\text{R}$ are reproduced those that are detected in the absence of yeast cell lysate (**Table 5.2**). This proves that the click reaction for labeling of an affinity tag is efficient, and that SCX combined with avidin affinity chromatography is an effective workflow for enrichment of cross-linked peptides even in extremely complicated environments.

5.3.8. Validation of Cross-Linked Residues

The ubiquitin lysine residues cross-linked by CXL are compared to the known X-ray crystal structure in **Figure 5.5** (PDB ID: 1UBQ). The observed lysine pairs are all located within 20 Å. This is consistent with previous reports^{252,264,265} performed by other chemical cross-linkers that display a range of chain lengths similar to CXL. For example, in the ³⁰IQD³³K[^]EGIPPDQQ⁴²R–⁷TLTG¹¹K[^]TITLEVEPSDTIENV²⁷K cross-linked peptide, the distance between alpha carbons in each lysine residue (K11 and K33) is 12.85 Å. The maximum length of the cross-linker in all-trans conformation is ~6.6 Å which is shorter by 6.25 Å. However, the distance between NZ atoms in the side-chains of lysine residues is only 7.15 Å. Considering the flexibility of the lysine side chain and thermal motions in proteins, the observation of cross-linking between K11 and K33 residues is reasonable. The homodimer of the ⁴³LIFAG⁴⁸K[^]QLEDG⁵⁴R peptide is also detected as in the previous report, indicating the formation of native ubiquitin homodimers in solution.²⁵²

By employing chemical reactions and subsequent sample treatments, less abundant cross-linked peptides can be lost (**Table 5.2**). Compared to missing cross-linked peptides, three cross-linked peptides detected after avidin enrichment have more closely positioned lysine pairs. Therefore, the result can be rationalized by the relationship between the spatial orientation of the lysine pairs (*i.e.*, opportunity for cross-linking) and the resulting copy number of the cross-linked peptides, and is also consistent with the crystal structure. Therefore, no further enrichment by clicking an affinity tag is necessary for the low complexity samples (*e.g.*, *in vitro* cross-linking of protein complexes that include dozens of known proteins) and SCX fractionation would be a still good choice for separation of cross-linked peptides without any sacrifice for sensitivity.

Figure 5.5 Ubiquitin X-ray crystal structure versus cross-linked peptides



Ubiquitin X-ray crystal structure (PDB: 1UBQ), presented in ribbon diagram and color coded to highlight the peptide chains that are cross-linked by CXL. Lysine residues are shown explicitly, and C α -C α and NZ-NZ distances (in Å) for the experimentally observed cross-linked peptides are denoted in red.

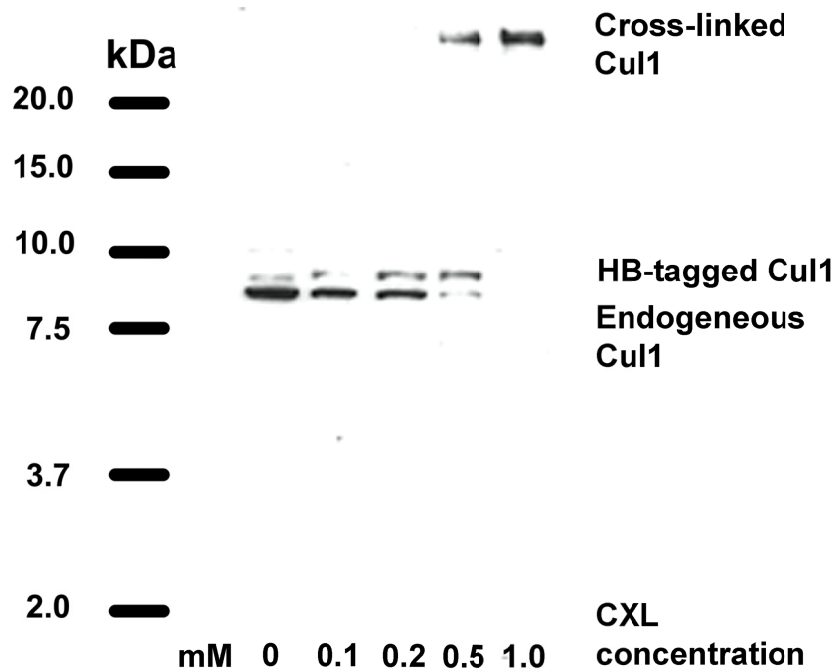
5.3.9. *In vivo* Cross-Linking of HEK 293

To test the cell permeability and water solubility of CXL, *in vivo* cross-linking of HEK 293 cells, followed by Western blot analysis of Cul1 were performed. Cul1 is a ubiquitin ligase that attaches a ubiquitin chain on target substrates for proteasome-catalyzed degradation.²¹¹ Cul1 is a prototype of the cullin ligase family, and constitutes modular ligase complexes with other binding partners. During the design of CXL, we considered two features, the cationic nature of CXL in physiological pH and the small size, for efficient penetration of cross-linkers into cell membranes, which is a widely employed strategy in synthesis of drug delivery carriers using cationic polymers.²⁶⁶ It should be firstly confirmed that CXL in working concentration ranges has no or minimal cytotoxic activity on cells. It was found that treatment of CXL up to 1 mM did not induce any significant cell toxicity and no visual change that may be caused by the entanglement of cell debris was observed. Secondly, if cytoplasmic Cul1 protein is cross-linked by CXL, the postulated cell permeability and water solubility from the design of CXL can be verified. **Figure 5.6** depicts the Western blot analysis of cross-linked Cul1 samples acquired from *in vivo* cross-linking of HEK 293 cells. From the observation of the higher molecular weight band, it is clear that Cul1 is cross-linked by CXL in 0.5~1.0 mM ranges. It also indicates that CXL is cell-permeable and soluble in PBS. In summary, we found that CXL is compatible with the cellular environment for *in vivo* cross-linking.

5.3.10. Application for Complex Systems

The encouraging results observed in this work from both *in vitro* and *in vivo* cross-linking experiments hold a promise for future application in systems level cross-linking experiments. As observed from CID and ETD of cross-linked peptides originated from ubiquitin cross-linking, tandem mass spectrometric analyses can provide useful information for sequencing of cross-linked peptides and identification of protein binding partners. Especially, the reporter ion

Figure 5.6 *In vivo* cross-linking of HEK 293 cells



In vivo cross-linking of HEK 293 and Western blot analysis. From 0 to 1.0 mM of CXL were applied to the cross-linking of Cul1. The observation of higher molecular weight bands over 20.0 kDa in 0.5 and 1.0 mM CXL lanes indicate successful cross-linking of cytoplasmic Cul1.

observed in the CID spectra is very useful for the reduction of the number of the MS/MS spectra that are subject to the data base searching. This feature would be particularly invaluable for the systems level study by saving the requirement of computational resources. The fragmentation pathway used in the formation of the reporter ion is actually universal when the 1,2,3-triazole ring is positioned through four methylene linker to the tertiary amine residue regardless of the structures of the attached affinity tags. Therefore, the highly selective reporter ion can be generated in a *mass-tunable* way by inserting various affinity tags. Cross-linkers with various chain lengths, and isotope-coded cleavable affinity tags²⁶² can be also prepared with the current synthetic scheme. By combining these versatile features, it is possible to explore future applications of CXLs to obtain more information about protein structure and PPIs that are targeted for more complex systems.

5.4.Conclusions

In this work, a highly versatile clickable cross-linker (CXL) that overcomes many limitations of the previous functionalized cross-linkers is developed for selective and sensitive identification of cross-linked peptides from complex mixtures. CXL is designed to have good solubility, cell permeability and an alkyne functional group for attachment of an affinity tag via CuAAC. Several advantages of CXLs have been validated using model peptides, a model protein, ubiquitin, and a real biological system using HEK293 cells. Especially, the application of CXL for analysis of complex systems is demonstrated by successful fractionation and enrichment of cross-linked peptides, and MS/MS scan filtering using the reporter ion at m/z 525.3. With these promising results, we are exploring the possibilities in application of CXLs in more complex systems.

5.5.Acknowledgements

This work was supported by the National Science Foundation through grant CHE-0416381 and the Beckman Institute at California Institute of Technology. Computational resources for density functional theory results were kindly provided by the Materials and Process Simulation Center at California Institute of Technology. We are also grateful to Bert Lai for technical assistance with the CD experiments. C.H.S. acknowledges a fellowship from the Kwanjeong Educational Foundation. J.E.L. was supported by the Ruth L. Kirschstein NRSA fellowship from the NIH (CA138126).