

Protocol for High-Yield Production of Photo-Leucine-Labeled Proteins in *Escherichia coli*

Bastian Kohl, Mitchell Brüderlin, Danilo Ritz, Alexander Schmidt, and Sebastian Hiller*



Cite This: *J. Proteome Res.* 2020, 19, 3100–3108



Read Online

ACCESS |



Metrics & More



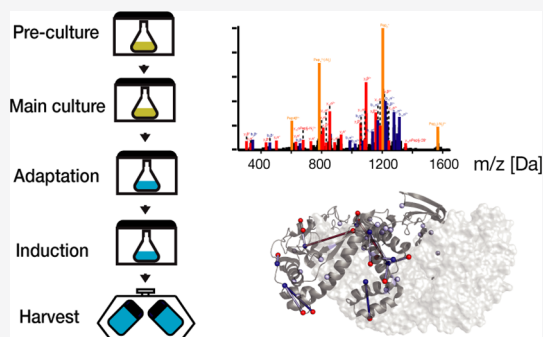
Article Recommendations



Supporting Information

ABSTRACT: UV-cross-linking mass spectrometry is an emerging technique to obtain structural information of biomacromolecules and their complexes *in vivo* and *in vitro*. In particular, certain photo-reactive amino acids (pA) such as photo-leucine (pLeu) and photo-methionine can provide unique short-distance information on the structural core regions of proteins. Here, we present a protocol for high-yield incorporation of pLeu in proteins recombinantly expressed in *Escherichia coli*. The protein of interest is expressed at high cell densities, which reduces the required amount of the pA by a factor of 10, as compared to the standard protocols, while maintaining high incorporation rates. For the two chaperones, trigger factor and SecB, up to 3 mg of pLeu-labeled protein were thus obtained from 100 mL of cell culture, with label incorporation rates of up to 34%. For trigger factor, UV-induced cross-linking leads to the identification of 12 cross-links that are in agreement with the published three-dimensional structures. The accessibility of milligram amounts of pLeu-labeled proteins at low costs will be highly useful to address structural biology questions.

KEYWORDS: photo-leucine, mass spectrometry, high-cell density culture, UV-cross-linking, molecular chaperones



1. INTRODUCTION

Cross-linking mass spectrometry (XL-MS) has a strong potential to deliver quantitative information on structures and dynamics of proteins and their complexes *in vivo* and *in vitro*.¹ In the past decade, several technological advances, in particular, the development of advanced cross-linkers and mass spectrometers have enhanced the power of XL-MS dramatically.² Typical cross-linkers consist of one or several reactive groups, such as *N*-hydroxysuccinimide ester (NHS-ester), maleimide, carbodiimide, or photo-reactive groups like diazirine, connected by a linker ranging from “zero” length to more than 30 Å.^{3–6} For structural biology, the class of “zero”-length cross-linkers such as carbonyldiimidazole (CDI) and photo-reactive amino acid (pA) analogues are particularly interesting as they have the highest spatial resolution.^{7,8} Cross-linking reactions can be performed as an endogenous or exogenous reaction. Exogenous cross-linking is performed by adding the cross-linking substance to the protein of interest. In this case, the cross-linker reacts mostly with surface accessible functional groups of the protein such as lysine, serine, threonine, aspartate, or glutamate residues.^{7,9–12} In contrast, endogenous cross-linking employs amino acids incorporated at specific sites in the protein such as cysteines and pA analogues.^{6,13} As compared to exogenous cross-linkers, endogenous cross-linkers can be incorporated in the hydrophobic core of the protein and therefore provide complementary information. Also, this makes them a powerful tool to

study hydrophobic protein–protein interactions, protein structure changes, and dynamics, as well as membrane proteins *in vivo* and *in vitro*.

Among the available approaches for photo-cross-linking,^{13–24} a highly interesting class are the diazirine analogues. A large library of up to eight different diazirine-type pA analogues has been developed, including the photo-lysine derivative DizPK.^{19,25–30} These analogues can be incorporated into proteins by the human cellular translation machinery and can be used to obtain photo-induced cross-links for proteins in cells.⁶ Additionally, studies could show that photo-methionine can be used to label recombinant proteins in *Escherichia coli*.³¹ A drawback with photo-methionine, however, is that methionine is generally a rare amino acid in proteins and therefore can provide only limited structural data.³² In contrast, photo-leucine (pLeu) is highly abundant in protein cores and thus of high interest to address structural biology questions by XL-MS.³²

Two different types of strategies have been developed to incorporate reactive amino acid analogues into proteins. The

Received: February 19, 2020

Published: May 15, 2020



first strategy incorporates the amino acid in a site-specific manner by using the amber stop-codon approach.¹⁵ While this method allows to control the site of incorporation, for each site a mutation is needed making this method highly time-consuming and costly. The second method uses the endogenous cellular translation machinery to incorporate the amino acid analogues along with the regular amino acid,⁶ leading to a statistical incorporation rate and giving simultaneous access to entire proteins. A protocol for labeling recombinantly expressed protein in *E. coli* with photo-methionine has been shown to lead to average incorporation rates of 25%.³¹ Further investigation is needed to determine whether *E. coli* expression protocols can also be adapted to pLeu. Furthermore, the published protocols report a very high consumption of photo-amino acid, which can lead to overall high costs. A possible strategy to deal with these issues are *E. coli* high cell-density cultures. This expression technique originated from the field of NMR spectroscopy, where labeling of proteins with stable isotopes such as ¹⁵N or ¹³C is routinely performed.^{33,34} High cell-density expression protocols consist of two growth phases. In the first phase, the cells are cultured in an unlabeled medium until they reach a sufficient cell density in early- to mid-log growth phase. The cells are then transferred to the labeled minimal medium under simultaneous increase in their density by a factor of 10. This effect reduces the required amount of the labeling compound accordingly while maintaining the protein yields. Importantly, labeling protocols based on high-cell density cultures do not allow label incorporation rates of close to 100% because of the dilution of the label with unlabeled biomass from the initial growth phase.^{33,34} While this feature has prevented their widespread use in solution NMR spectroscopy, it might be less of a limitation for XL-MS, where partial labeling incorporation of below 50% is sufficient.

Here, we present such a high cell-density protocol for labeling protein with pLeu in *E. coli* and demonstrate its feasibility with the two molecular chaperones: trigger factor and SecB. We determine protein yields and incorporation rates and demonstrate that the protocol can be successfully used to determine chemical cross-links in intact *E. coli* cells.

2. MATERIALS AND METHODS

2.1. High-Cell Density Expression Protocol for pLeu Labeling

A pET-based vector encoding the protein of interest was transformed into *E. coli* BL21(DE3) lemo cells using the heat shock method. Cells were grown overnight at 37 °C, 300 rpm in an orbital shaker. By next morning, 100 mL of the Luria broth (LB) medium were inoculated with the overnight culture to a theoretical OD₆₀₀ = 0.02. Cells were grown at 37 °C, 300 rpm in an orbital shaker until an OD₆₀₀ = 0.6. Cells were pelleted at 1500g for 10 min at 37 °C and were resuspended in a 10 mL pLeu-M9 (pA-M9) labeling medium (47 mM Na₂HPO₄, 22 mM KH₂PO₄, 8.5 mM NaCl, 1× solution Q [1000×: 1.25 mM CaCl₂, 0.12 mM CoCl₂, 23 μM CuCl₂, 25.1 mM FeCl₃, 2.5 mM ZnCl₂, 1 mM H₃BO₃, 0.2 mM MnCl₂, Na₂Mo₄], 2 mM MgSO₄, 1× amino acid mix [130 mg/L alanine, 3 mg/L arginine, 14 mg/L aspartate, 10 mg/L glutamate, 12 mg/L glycine, 1 mg/L histidine, 4 mg/L isoleucine, 6 mg/L lysine, 4 mg/L phenylalanine, 4 mg/L proline, 4 mg/L serine, 5 mg/L threonine, 3 mg/L tyrosine, 6 mg/L valine], 1× vitamin mix [100×: 0.1 mg/mL D-biotin, 0.1

mg/mL choline chloride, 0.1 mg/mL folic acid, 0.2 mg/mL myo-inositol, 0.1 mg/mL nicotinamide, 0.1 mg/mL D-pantothenic acid, 0.1 mg/mL pyridoxal-HCl, 0.01 mg/mL riboflavin, 0.1 mg/mL thiamine HCl, 8.5 mg/mL NaCl], 10 g glucose/L, 1 g/L NH₄Cl), supplemented with 1 or 4 mM pLeu (Thermo Scientific). Cells were cultured for 1 h at 37 °C, 300 rpm to adapt to new medium conditions. Afterward, cells were induced with 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cell density at this point was typically OD₆₀₀ = 8–10. Expression was performed for 4 h at 37 °C, 300 rpm.

2.2. Preparation of Full Cell Tryptic Protein Extracts

To measure the incorporation rates of pLeu, non-UV irradiated cells were resuspended in lysis buffer (1% Sodium deoxycholate, 10 mM tris-[2-carboxyethyl]-phosphine (TCEP), 100 mM Tris, pH = 8.5), and cells were disrupted by ultrasonication with a Bioruptor pico system (diagenode). After cell disruption, samples were centrifuged for 10 min, 21,130g, 4 °C, and the supernatant was used for further analysis. To prepare the tryptic peptides, extracted proteins were reduced with 10 mM TCEP at 95 °C for 10 min, and after cooling to room temperature, the proteins were alkylated with 15 mM chloroacetamide for 30 min at 37 °C. Proteins were digested using sequencing-grade-modified trypsin (1/50, w/w, trypsin/protein; Promega, USA) overnight at 37 °C. After digestion, the samples were supplemented with trifluoroacetic acid (TFA) to a final concentration of 2.5%. Peptides were cleaned up using PreOmics Cartridges (PreOmics, Martinsried, Germany) following the manufacturer's instructions, dried using a vacuum concentrator, and stored at –20 °C until further analysis.

2.3. Determination of pLeu Incorporation Rates

Dried peptides were resuspended in 0.1% formic acid by sonication and subjected to liquid chromatography–mass spectrometry (LC–MS/MS) analysis using a dual pressure LTQ-Orbitrap Elite mass spectrometer fitted with an EASY-nLC 1000 (both Thermo Fisher Scientific) and a custom-made column heater set to 60 °C. Peptides were separated using a reverse-phase high-performance liquid chromatography (RP-HPLC) column (75 μm × 30 cm) packed in-house with C18 resin (ReproSil-Pur C18-AQ, 1.9 μm resin; Dr. Maisch GmbH) at a flow rate of 200 nL/min. The following gradient was used for peptide separation: from 5% B to 10% B over 5 min to 35% B over 70 min to 50% B over 15 min to 95% B over 2 min followed by 18 min at 95% B. Buffer A was 0.1% formic acid in water, and buffer B was 80% acetonitrile, 0.1% formic acid in water.

The mass spectrometer was operated in a data-dependent acquisition (DDA) mode set to obtain one high-resolution MS scan in the orbitrap at a resolution of 240,000 full width at half maximum (fwhm) (@400 *m/z*, MS) followed by MS/MS scans in the linear ion trap of the 20 most intense MS signals. The charged state screening modus was enabled to exclude unassigned and singly charged ions, and the dynamic exclusion duration was set to 45 s. The ion accumulation time was set to 300 ms (MS) and 25 ms (MS/MS). MS and MS/MS scans were acquired at a target setting of 1 × 10⁶ ions and 10,000 ions, respectively. The normalized collision energy (NCE) was set to 35%, and one micro scan was acquired for each spectrum.

Raw MS data were processed using the MaxQuant software suite version 1.6.2.3. MS/MS data were searched against the *E. coli* strain K12 Swissprot databases containing 4350 proteins

downloaded from the UniProt homepage on March 18, 2018.^{35–37} The following analysis settings were used: (i) digestion mode: Trypsin/P, (ii) max missed cleavages: 2, (iii) peptide tolerance first search 20 ppm; main search 4.5 ppm, (iv) three variable modifications (oxidation of methionine, acetylation of protein N-term and pLeu) and one fixed modification (carbamidomethylation of cysteine), (v) “include contaminants” was enabled, (vi) peptide spectrum matches false discovery rate (FDR) 1%, protein FDR 1%, and (vii) FTMS MS/MS match tolerance (20 ppm), ITMS MS/MS match tolerance (0.5 Da). From the “evidence.txt” output file, all peptides that were marked as “reverse,” “potential contaminants,” or “only identified by site” were discarded, and raw intensities for each peptide summed and extracted for further analysis.

2.4. *In Vivo* UV-Cross-Linking of *E. coli* Cells

To obtain intracellular cross-links, cells equal to 100 μ g whole cell protein were transferred into a 12-well Petri dish plate and on ice exposed in UV Stratalink 2400 with 1.25 J UV light with a 365 nm wavelength (Stratagen). Cross-linked peptides were identified by XL-MS analysis (see Section 2.6).

2.5. Western Blot Analysis

For the Western blot (WB) analysis of *in vivo* cross-linked proteins, whole cells were lysed in 5 \times Laemmli buffer and incubated for 10 min at 95 °C. The solution was centrifuged for 10 min at 21,000g and run on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (4–15% Mini-PROTEAN TGX gel, BioRad). Blotting was performed, according to the manufacturer's instruction of the BioRad Trans-Blot Turbo System. Transfer efficiency was monitored by a stained protein ladder. For the detection of overexpressed His-tagged proteins, a mouse peroxidase coupled—anti-His₆ antibody (Sigma-Aldrich) was used. Blocking was performed with bovine serum albumin. The antibody-binding step was performed for 1 h at room temperature followed by 3 \times 10 min washing steps. The blot was developed with a Lumi-Light Western blotting substrate kit (Roche) and detected with a Fusion FX system (Vilber).

2.6. XL-MS Protocol for Analysis of *E. coli* Chaperone Trigger Factor

2.6.1. Preparation, Purification, and Enrichment of UV-Cross-Linked Tryptic Peptides. *E. coli* cells equal to 300 μ g whole cell protein were cross-linked, as described in Section 2.4. Afterward, the cells were disrupted by ultrasonication, and the trigger factor with an N-terminal His-tag was purified via Ni-NTA chromatography. The purified protein was reduced at 95 °C for 10 min with 10 mM TCEP, and after cooling to room temperature, the protein was alkylated with 15 mM chloroacetamide for 30 min at 37 °C. For the digest, we used sequencing-grade-modified trypsin in a 1/50 ratio (w/w, trypsin/protein; Promega, USA) overnight at 37 °C. After digestion, the samples were supplemented with TFA to a final concentration of 2.5%. Peptides were cleaned up using C18 columns (Silica C18, MacroSpin column, The Nest Group Inc.) following the manufacturer's instructions, dried using a vacuum concentrator, and stored at –20 °C until further analysis. In the next step, cross-linked peptides were enriched by the size-exclusion protocol. For this, the purified peptides were resuspended in size-exclusion buffer (70% water/30% acetonitrile/0.1% TFA), sonicated for 1 min, and incubated at 25 °C for 10 min to be dissolved. Afterward, the peptide

mixture was injected onto the Superdex 30 increase (3.2/300, GE healthcare) column to enrich cross-linked peptides. Peptides were eluted with 40 min isocratic elution at a flow rate of 70 μ L/min, and fractions of 100 μ L were collected. Fractions were dried again by a vacuum concentrator and stored at –20 °C until further analysis.

2.6.2. Cross-Linking MS. Dried peptides were resuspended in 0.1% formic acid by sonication and subjected to LC–MS/MS analysis using an Orbitrap Fusion Lumos Tribrid mass spectrometer fitted with an EASY-nLC 1200 (both Thermo Fisher Scientific) and a custom-made column heater set to 60 °C. Peptides were separated using a RP-HPLC column (75 μ m \times 36 cm) packed in-house with C18 resin (ReproSil-Pur C18-AQ, 1.9 μ m resin; Dr. Maisch GmbH) at a flow rate of 200 nL/min. For DDA-mode, peptides were separated by a gradient from 5% B to 12% B over 5 min to 35% B over 40 min to 50% B over 15 min to 95% B over 2 min followed by 18 min at 95% B. For targeted analysis, peptides were separated by a gradient from 5% B to 65% B over 60 min to 95% B over 2 min followed by 18 min at 95% B. Buffer A was 0.1% formic acid in water, and buffer B was 80% acetonitrile, 0.1% formic acid in water.

For DDA, the mass spectrometer was operated with a cycle time of 3 s between MS scans. Each MS scan was acquired in the Orbitrap at a resolution of 120,000 fwhm (@200 m/z) and a scan range from 250 to 1500 m/z followed by MS/MS scans of the most intense precursors in the Orbitrap at a resolution of 30'000 fwhm, with an isolation window of the quadrupole set to 1.4 m/z . Maximum ion injection time was set to 50 ms (MS) and 54 ms (MS/MS) with an automated gain control (AGC) target of 1×10^6 and 1×10^5 , respectively. Only peptides with the charge state 2–6 were included in the analysis. Monoisotopic precursor selection was set to peptide, and the intensity threshold was set to 5×10^3 . Peptides were fragmented by collision-induced dissociation (CID) with collision energy set to 29%, and one micro scan was acquired for each spectrum. The dynamic exclusion duration was set to 30 s. Raw data were peak picked and converted into mzXML/mgf-format with MSConvert 3.0. Analysis of cross-link MS/MS data was performed with software MeroX 1.6.6 and in-house python scripts.³⁸ Identifications with a MeroX score above 70 were selected for targeted analysis. Out of these 19 targets, three were randomly selected to be measured twice to probe the instrumentation stability.

For targeted analysis, an inclusion list of 22 masses of cross-linked peptides was thus used (Table S1). One MS scan was followed by 22 MS/MS scans of the targeted cross-linked peptides. MS scans were acquired in the Orbitrap at a resolution of 120,000 fwhm (@200 m/z), a scan range from 250 to 1500 m/z , a maximum injection time of 50 ms, and an AGC target of 1×10^6 . MS/MS scans were acquired in the Orbitrap at a resolution of 60,000 fwhm, with an isolation window of the quadrupole set to 0.4 m/z , a scan range from 150 to 1700 m/z , a maximum injection time of 118 ms, and an AGC target of 1×10^6 . Peptides were fragmented by higher energy collisional dissociation (HCD) with collision energy set to NCE: 35% or CID (NCE: 29%) or stepped HCD (NCE: 25–35%). Raw data were peak picked and converted into mzXML/mgf-format with MSConvert 3.0. Analysis of XL-MS data was performed with software MeroX 1.6.6 and in-house python scripts.³⁸ To identify cross-linked peptide pairs, DDA and parallel reaction monitoring (PRM) data were searched for cross-links between pLeu and aspartate and/or glutamate.

During the search, a variable modification for all leucine residues from leucine to pLeu was set with a maximum of two modification of this type per peptide. The MS/MS spectra were deisotoped by MeroX. For the search, the precursor precision was set to 5.0 ppm. The fragment ion precision was set to 10.0 ppm. A maximum of two neutral losses of identified fragments were allowed during the search. Initially, an intensity-based prescoring with an intensity threshold of 10% was applied to the data. To increase cross-link identification confidence, the MS/MS spectra were additionally searched for specific reporter ions (RISE mode). They are produced by the cleavage of the cross-link between pLeu and aspartate/glutamate during fragmentation. A decoy database made of shuffled protein sequences but keeping protease sites constant was used. For FDR analysis, the cut off was set to 5%. For PRM analysis, cross-links identified during DDA runs with a MeroX score above 70 were included. Finally, only cross-links identified again in PRM runs with a MeroX score above 95 were included into the results.

3. RESULTS AND DISCUSSION

3.1. High-Cell Density Expression Protocol for pLeu Labeling

As a starting point to establish protein expression in *E. coli* with pLeu-labeling, we determined the toxicity of pLeu to BL21(DE3) lemo expression cells. The cells grew exponentially in the standard M9 minimal medium, reaching saturating after a few hours (Figure S1). The M9 medium does contain glucose as the sole carbon source, such the bacteria have to rely on their biogenesis to synthesize all 20 proteinogenic amino acids themselves. Strikingly, pLeu is highly toxic when added to this medium, such that already at concentrations of 0.5 mM, it completely prevented growth of the cells (Figure S1). These toxic effects were however significantly reduced by the addition of a background concentration of 0.1 mM leucine. Under these conditions, the cells were found to grow at the same rate up to pLeu concentrations of 2 mM and tolerated pLeu concentrations of up to 8 mM (Figure S1). We, therefore, aimed at medium conditions without pLeu for the initial growth phase and of 4–8 mM pLeu during expression.

Therefore, the protocol to produce recombinant proteins with pLeu-labeling at high yields consists of two growth phases, a culturing growth phase in the nonlabeled medium until early-logarithmic growth, followed by a transfer to the pLeu-labeled medium in which high-cell density conditions are thus reached, and the expression is then induced (Figure 1). In the first step, the cells are cultured in 100 mL of an unlabeled LB medium until an optical density of $OD_{600nm} = 0.6$ is reached. At this point, the cells are still in exponential growth. They are pelleted by centrifugation and resuspended in a 10-fold reduced volume of the M9 medium, supplemented with pLeu, resulting in an initial $OD_{600} = 8.0$ at time of induction. It is crucial that this transfer occurs under mild conditions to ensure cell viability, indicating centrifugation at low centrifugal forces and a temperature of 37 °C and resuspension of the cell pellet by shaking instead of pipetting. Furthermore, because of the high-cell density experienced in the second step, baffled flasks not filled to more than 1/5 of their volume are recommended to ensure sufficient oxygen supply. Anerobic growth conditions would reduce protein expression rates drastically. After the transfer, the cells are adapted to a new medium by growing them for 1 h. Then, protein expression is

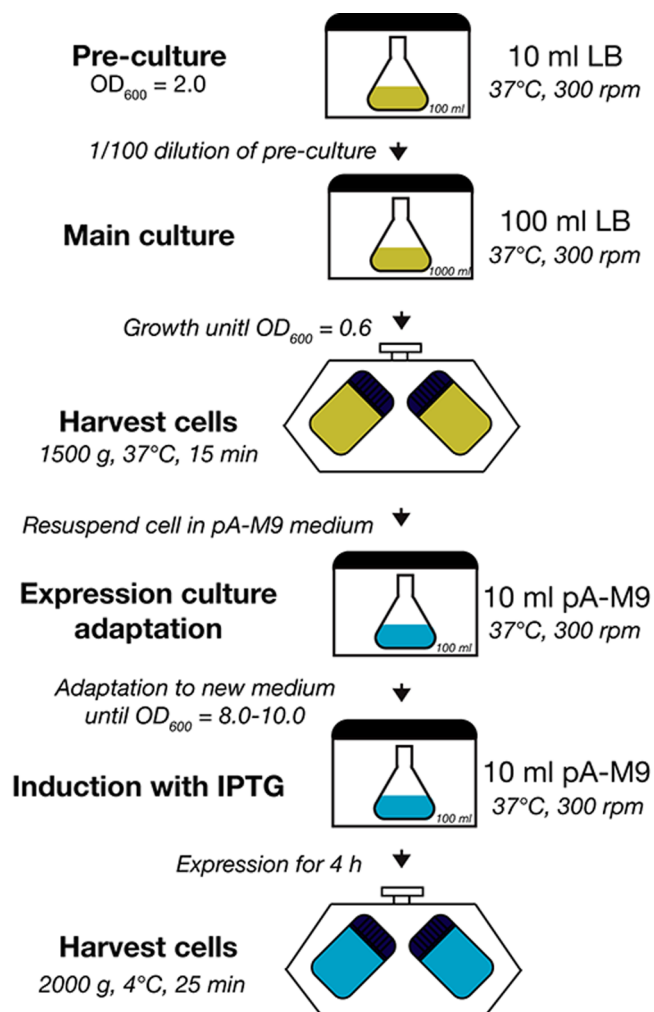


Figure 1. Flowchart for production of recombinant proteins in high-cell density *E. coli* culture, with incorporation of a pA. See the text for all details.

induced by adding the inducer IPTG for the pET-vector expression system (Figure S2).

3.2. Application to Two Chaperone Proteins

The applicability of the protocol was tested with the two *E. coli* chaperones: trigger factor and SecB. An efficient labeling protocol should feature sufficient incorporation rates and high overall yields under simultaneous cost-efficiency. Importantly, the protocol needs to balance between too low incorporation rates that would lead to weak detection sensitivity for the final cross-links and too high incorporation rates that would not be tolerated by *E. coli* and that could lead to protein misfolding because of structural differences between the amino acid and the photo-reactive analogue. Therefore, we aimed at label incorporation rates of the order of 20–30%. In the first step, we validated the successful pLeu-incorporation into the overexpressed proteins: SecB and trigger factor. An example is given by the mass spectrum and fragment ion mass spectrum for the peptide ³⁴⁵VVG³⁵⁶LLLG³⁵⁶EVIR of trigger factor (Figure 2A,B). For this peptide, MS mass peaks for the unmodified (m/z : 633.91 Da, charge: 2⁺) and modified peptide (m/z : 639.9 charge: 2⁺) were detected with a mass shift of 11.98 Da, which is exactly the mass shift for a pLeu-labeled peptide with the charge state of two (Figure 2A). The example fragment ion

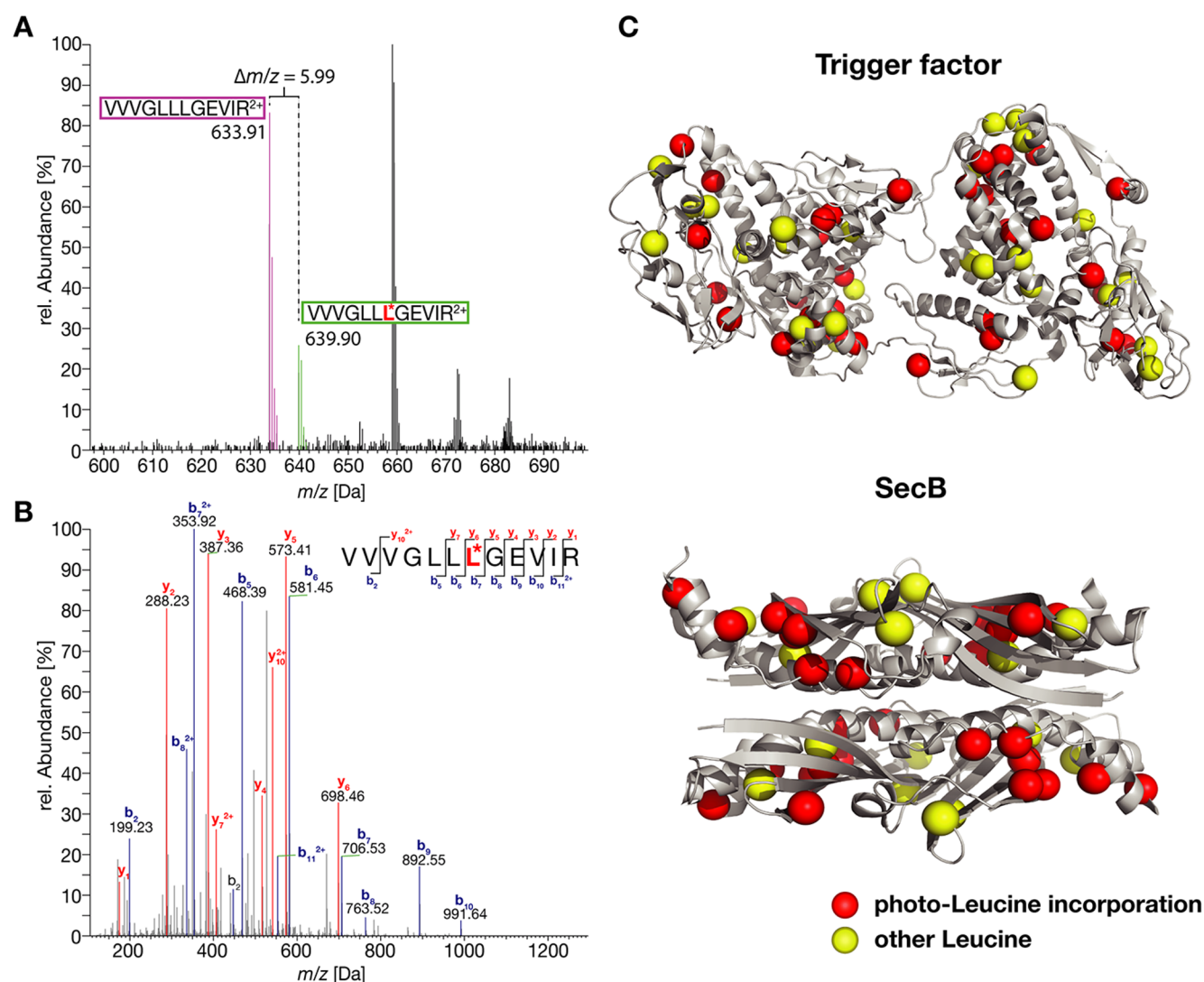


Figure 2. MS of pLeu-labeled chaperones. (A) Example MS-spectrum of trigger factor, showing the unlabeled and pLeu-labeled peptide 345–356 of the trigger factor (VVVGLLLGEVIR) in magenta and green, respectively. (B) Product ion mass spectrum of the pLeu-labeled peptide from (A) (green peak). Ions of b-series and y-series are labeled in blue and red, respectively. (C) 3D structures of the *E. coli* chaperones: trigger factor (39, SOWI) and SecB (41, IQYN). Leucine residues, for which pLeu incorporation was detected, are shown as red spheres and all other leucine residues as yellow spheres.

Table 1. Incorporation Rates of pLeu into All Detected *E. coli* Proteins and for the Overexpressed Proteins (SecB and Trigger Factor)

	SecB (1 mM pLeu) ^a	SecB (4 mM pLeu) ^a	trigger factor (4 mM pLeu) ^a
pLeu/Leu in proteome [%]	0.16 ± 0.1	2.0 ± 0.2	1.9 ± 0.4
pLeu/Leu in protein of interest [%]	7.6 ± 1.9	34.9 ± 5.3	16.4 ± 5.1

^apLeu concentration used in the expression culture.

mass spectrum shows specific b- and y-ion peaks clearly indicating the incorporation of pLeu into the peptide at position 7 (Figure 2B).

Subsequently, pLeu incorporation rates were determined for either the full *E. coli* proteome or the overexpressed protein, based on the raw intensity values of MS/MS data of tryptic digests of the expression culture cells (Table 1). The overall incorporation rates of pLeu into the *E. coli* proteome are 2.0%

on an average. For the overexpressed proteins (trigger factor and SecB), average incorporation rates of 16.4 ± 5.1 and $34.9 \pm 5.3\%$, respectively, were observed at 4 mM pLeu concentration in the medium (Figures S3 and S4). The difference in these values is readily rationalized by the fact that the bulk of cellular biomass is synthesized already during the growth phase of the cell in the unlabeled LB medium, while the production of recombinant protein occurs only after induction by the T7 promoter in the pLeu-labeled medium. Consequently, the pLeu incorporation rates into the overexpressed proteins are higher. Overall, for trigger factor, 17 of the 31 leucine residues (54%) and for SecB, 9 of the 12 leucine residues (75%) were detected with a pLeu incorporation in MS/MS sequencing (Figure 2C). These ratios are satisfactory because it is not expected that all peptides are observed because of under sampling and different peptide ionization efficiencies. Under the tested conditions, our high-cell density protocol produced up to 3 mg of recombinant protein per 100 mL of the labeling medium. In summary, we obtained

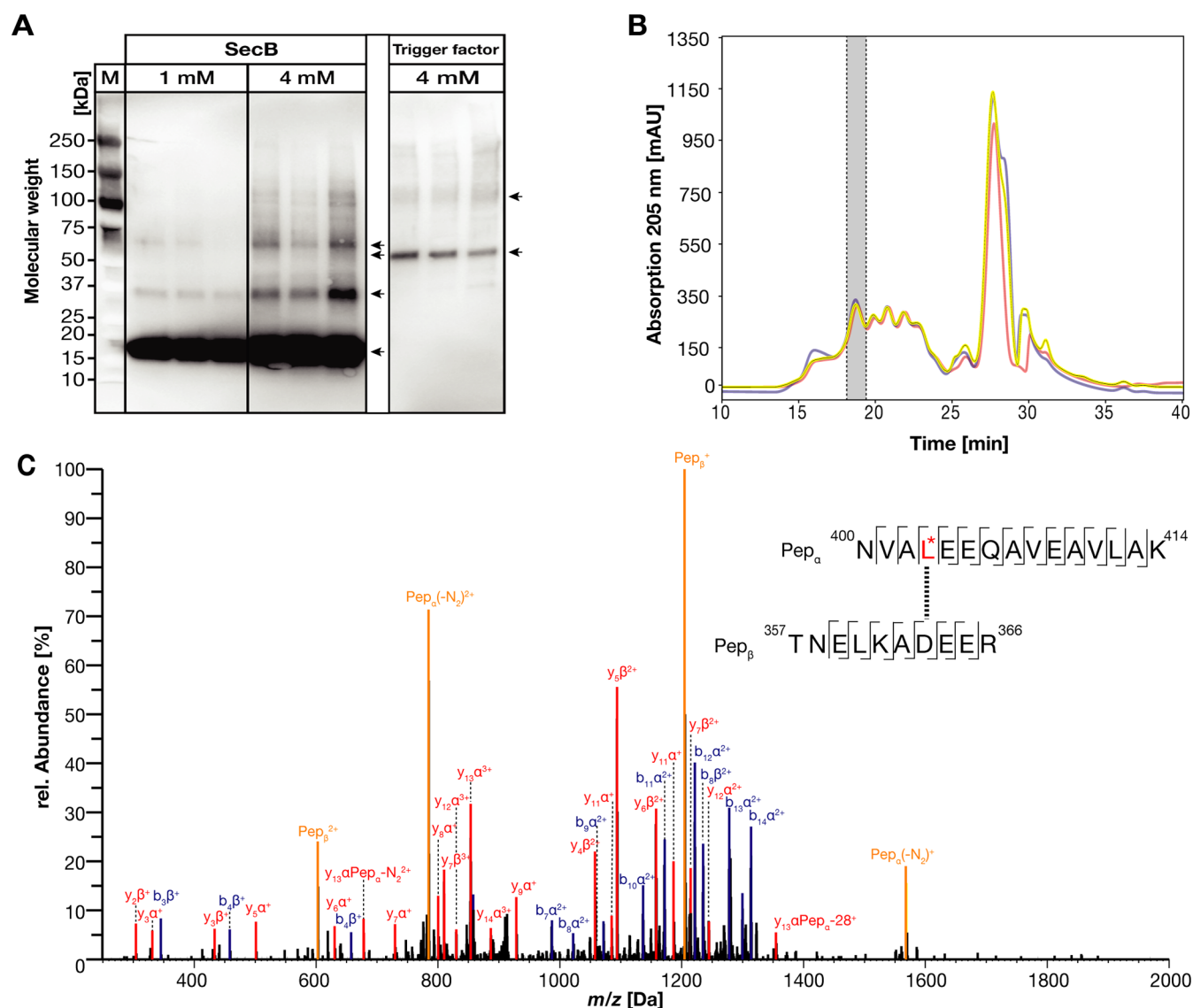


Figure 3. In-cell cross-linking of recombinantly expressed pLeu-labeled proteins in *E. coli*. (A) Anti-His WB of in-cell UV-cross-linked proteins after purification. (B) SEC elution chromatogram of in-cell UV-cross-linked purified trigger factor after purification and tryptic digestion. Colors red, blue, and yellow represent three biological triplicates. The gray-shaded fraction has been selected for MS/MS identification run in panel C. (C) Fragment ion mass spectrum (CID fragmentation: 29%) of a cross-linked peptide pair of the trigger factor. Ions of b-series and y-series are labeled in blue and red, respectively. Cross-linker cleavage-induced peptide fragments are labeled in orange.

incorporation rates of at least 16.4% for both tested proteins at a 4 mM pLeu concentration. These values are comparable to the published rates for photo-methionine in *E. coli*.³¹

Next, we explored the effect of varying the pLeu concentration in the medium. A reduction of the pLeu concentration from 4 to 1 mM in the expression medium results in a reduced target-specific incorporation rate of 7.6 ± 1.9 for SecB and an accordingly reduced proteome incorporation rate of $0.16 \pm 0.1\%$ (Table 1). In contrast, an increase in the pLeu concentration to 8 mM did not result in further increased incorporation (Table S2). Notably, at a concentration of 10 mM pLeu, incorporation of pLeu completely broke down, which is in full agreement with the toxicity test and presumably to the toxic effects of pLeu under these conditions (Table S2, Figure S1). Overall, based on these data, we recommend a pLeu concentration in the medium of 4 mM, which results in sufficient incorporation for UV XL-MS experiments, while keeping the consumed amount of pLeu low.

In the next step, to validate the efficiency of the protocol, we did UV-cross-linking of each of the overexpressed chaperones in cells and analyzed with WB (Figure 3A). For each of the two proteins, we observed the formation of high-molecular complexes, corresponding to the cross-linked species. For the chaperone SecB, a known tetramer, the monomeric, dimeric, trimeric, and tetrameric species were detected on WB. For the trigger factor, a known dimer, the monomeric and dimeric species and additional species were detected (Figure 3A). Then, we performed a complete XL-MS cross-linking experiment with the *E. coli* chaperone trigger factor. Bacterial cells with the overexpressed trigger factor at a cellular concentration of presumably more than 100 μM were irradiated with UV light (350 nm) for 10 min, and afterward, the trigger factor was purified via His-tag purification. To identify cross-links, a tryptic digest protocol was used. We enriched cross-linked peptides by performing a size-exclusion purification and measured an earlier elution fraction on an Orbitrap Lumos

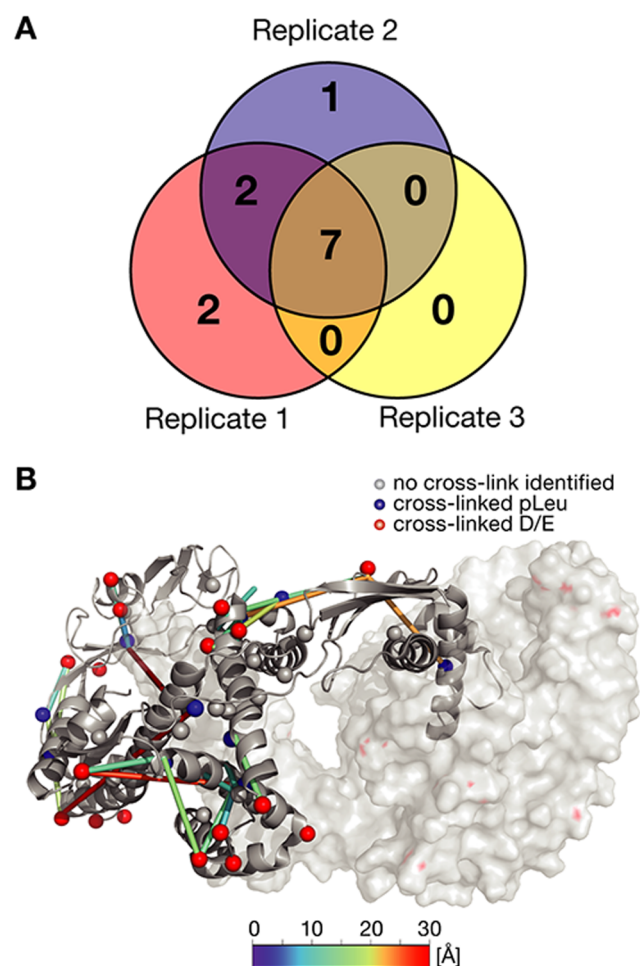


Figure 4. In-cell cross-linking of the chaperone trigger factor. (A) Venn-Diagram of identified cross-links in UV-cross-linked [pLeu]-trigger factor from three biological replicates. (B) 3D structure of the trigger factor (40, PDB 6D6S). Identified unique cross-links are shown as lines and color coded, according to the C^β – C^β distance. The spheres represent the C^β -atoms of cross-linked residues.

Fusion MS to identify pLeu cross-links (Figure 3B). Cross-links were identified in the first step by MeroX 1.6.6.6 software package and validated by manual inspection. Because of the low abundance of the cross-linked peptide pairs we performed,

based on DDA-run-identified possible cross-links, a targeted PRM method including 19 possible cross-link candidates was used for further validation. Finally, 12 out of 19 cross-links identified by the DDA-run could be validated by these runs and were included into the results (Figures 3, S5–S16). For example, the product ion mass spectrum of the cross-link D363/E364 to L403 (m/z : 924.478, charge: 3+) contains high number of b- and y-ions for peptide A and B (Figure 3C). As described before, the UV-cross-linking induced reaction of the diazirine with the carboxyl group of aspartate or glutamate results in an ester bond which can be cleaved in the MS-fragmentation chamber using the CID/HCD method.¹⁸ This results in the formation of highly specific reporter ions in the CID/HCD spectrum of cross-linked peptides. In our MS/MS data, we could identify the $\text{Pep}_\alpha(-\text{N}_2)$ -ion of peptide A which shows a mass loss of 28 Da, which corresponds to the loss of two nitrogen atoms (Figure 4C). Additionally, we observed the reporter ion Pep_β , which corresponds to the mass of the free B peptide after cleavage of the ester bond (Figure 4C). In total, we could identify 12 cross-links out of 3 biological replicates (Figure 4A and Table 2). Seven cross-links could be identified in all replicates. Out of the 12 cross-links, 9 were identified in at least 2 of the 3 replicates. Thereby, most of the identified cross-links are located between a pLeu and an aspartate or glutamate. For 7 out of the 12 cross-links, we observed the highly specific reporter ions. The lack of these ions for the other ones is likely the result of a noncleavable cross-link between pLeu with an amino acid containing no carboxyl group.

The observed cross-links are mostly in good agreement with the published structures of trigger factor protein (Figure 4B).^{39,40} Of the 12 unique cross-links, 1 is in the range of 5–10 Å, 8 in the range of 10–20 Å, and 3 above 20 Å. The observed cross-links that did not match the published ground state crystal structures of the trigger factor are likely the result of protein dynamics. The trigger factor has been shown to adopt a highly dynamic dimer structure, which is interconverting between multiple conformations.⁴⁰ While the current data clearly undersample the conformationally accessible space of the trigger factor, it is an exciting possibility that future application could deliver more comprehensive cross-link data to describe the conformational space better. In conclusion, we could show that the developed high-cell density protocol is

Table 2. Identified UV-Cross-Links in [pLeu]-Labeled Trigger Factor^{a,b}

position 1	position 2	m/z	charge	consecutive ^c
E241/242	L245/248	523.0392	4 ⁺	×
E292	L187	549.5358	4 ⁺	
L330	E96	597.3156	4 ⁺	
L32	E96	605.5688	4 ⁺	
E259/260/264	L248	674.1022	4 ⁺	×
E363/364/365	L403	729.3705	3 ⁺	
L294	D260/E264	731.3788	3 ⁺	
E419	L403	748.0840	3 ⁺	×
E199/200	L187	791.7123	3 ⁺	×
L332	E335/338/339	601.6490	3 ⁺	×
L351	E359	618.7073	3 ⁺	×
E359/363/364/365	L403	924.4778	3 ⁺	

^aA MeroX score above 95 (MeroX version 1.6.6.6), corresponding to an FDR <1%. ^bIndividual mass spectra, as shown in Figures S5–S16. ^cCross-links between two tryptic peptides that are sequential in the protein primary sequence.

able to label recombinant proteins with pLeu, and UV-cross-links can be observed and identified.

4. CONCLUSIONS

In this work, we have presented a highly efficient protocol for labeling recombinantly expressed proteins with pLeu in *E. coli*. The use of a high-cell density culturing approach reduces the amount of material enormously by keeping the protein yield and incorporation rates high. This protocol thus overcomes key issues of standard labeling protocols, including high costs for bigger expression cultures or low protein yields by using small culture sizes. Our protocol allows to produce proteins with different levels of pLeu incorporation rates for a broad range of XL-MS-based structural biology experiments. We could show that our protocol is able to produce *in vivo* cross-linking data of *E. coli* chaperones: trigger factor and SecB.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00105>.

PRM inclusion list, concentration dependence of pLeu incorporation, toxicity of pLeu for *E. coli*, overexpression of SecB, mass spectra of pLeu incorporation, and product ion mass spectra of identified cross-linked peptide pairs (PDF)

All identified proteins by MaxQuant (XLSX)

All identified pLeu modification sites by MaxQuant (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Sebastian Hiller – Biozentrum, University of Basel, 4056 Basel, Switzerland; orcid.org/0000-0002-6709-4684;
Email: sebastian.hiller@unibas.ch

Authors

Bastian Kohl – Biozentrum, University of Basel, 4056 Basel, Switzerland

Mitchell Brüderlin – Biozentrum, University of Basel, 4056 Basel, Switzerland

Danilo Ritz – Biozentrum, University of Basel, 4056 Basel, Switzerland

Alexander Schmidt – Biozentrum, University of Basel, 4056 Basel, Switzerland; orcid.org/0000-0002-3149-2381

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acs.jproteome.0c00105>

Author Contributions

B.K. and S.H. designed the study. B.K., M.B., D.R., and A.S. performed the experiments. All authors analyzed the data. B.K. and S.H. wrote the manuscript with input from all authors. All authors have read the final version of the manuscript and approved its content.

Notes

The authors declare no competing financial interest.

The mass spectrometry proteomics raw and result files data have been deposited to the ProteomeXchange Consortium via the PRIDE⁴² partner repository with the dataset identifier PXD017584.

■ ACKNOWLEDGMENTS

This work was supported from the Swiss National Science Foundation (grant 185388 to S.H.).

■ ABBREVIATIONS

AGC, automated gain control; CDI, carbonyldiimidazole; CID, collision-induced dissociation; DDA, data-dependent acquisition; FDR, false discovery rate; FT-MS, Fourier transformation mass spectrometry; HCD, high-energy collision-induced dissociation; IPTG, isopropyl- β -D-thiogalactopyranoside; IT-MS, ion trap mass spectrometry; LB, Luria broth; LC-MS, liquid chromatography MS; MS, mass spectrometry; NCE, normalized collision energy; NHS-ester, N-hydroxysuccinimide ester; NMR, nuclear magnetic resonance; OD, optical density; pA, photo-reactive amino acid; pLeu, photo-leucine; PSM, peptide spectrum matches; SDS, sodium dodecyl sulfate; TCEP, tris-[2-carboxyethyl]-phosphine; WB, western blot; XL-MS, cross-linking mass spectrometry

■ REFERENCES

- (1) Chen, Z. A.; Rappsilber, J. Quantitative cross-linking/mass spectrometry to elucidate structural changes in proteins and their complexes. *Nat. Protoc.* **2019**, *14*, 171–201.
- (2) Sinz, A. Cross-linking/mass spectrometry for studying protein structures and protein-protein interactions: where are we now and where should we go from here? *Angew. Chem., Int. Ed.* **2018**, *57*, 6390–6396.
- (3) Bragg, P. D.; Hou, C. Subunit composition, function, and spatial arrangement in the Ca²⁺- and Mg²⁺-activated adenosine triphosphatases of *Escherichia coli* and *Salmonella typhimurium*. *Arch. Biochem. Biophys.* **1975**, *167*, 311–321.
- (4) Cox, G. W.; Mathieson, B. J.; Giardina, S. L.; Varesio, L. Characterization of IL-2 receptor expression and function on murine macrophages. *J. Immunol.* **1990**, *145*, 1719–1726.
- (5) Partis, M. D.; Griffiths, D. G.; Roberts, G. C.; Beechey, R. B. Cross-linking of protein by ω -maleimido alkanoyl-N-hydroxysuccinimide esters. *J. Protein Chem.* **1983**, *2*, 263–277.
- (6) Suchanek, M.; Radzikowska, A.; Thiele, C. Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells. *Nat. Methods* **2005**, *2*, 261–268.
- (7) Hage, C.; Iacobucci, C.; Rehkamp, A.; Arlt, C.; Sinz, A. The first zero-length mass spectrometry-cleavable cross-linker for protein structure analysis. *Angew. Chem., Int. Ed.* **2017**, *56*, 14551–14555.
- (8) Mouradov, D.; King, G.; Ross, I. L.; Forwood, J. K.; Hume, D. A.; Sinz, A.; Martin, J. L.; Kobe, B.; Huber, T. Protein structure determination using a combination of cross-linking, mass spectrometry, and molecular modeling. *Methods Mol. Biol.* **2008**, *426*, 459–474.
- (9) Grabarek, Z.; Gergely, J. Zero-length crosslinking procedure with the use of active esters. *Anal. Biochem.* **1990**, *185*, 131–135.
- (10) Kao, A.; Chiu, C.-I.; Vellucci, D.; Yang, Y.; Patel, V. R.; Guan, S.; Randall, A.; Baldi, P.; Rychnovsky, S. D.; Huang, L. Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell. Proteomics* **2011**, *10*, M110.002212.
- (11) Liu, F.; Rijkers, D. T. S.; Post, H.; Heck, A. J. R. Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* **2015**, *12*, 1179.
- (12) Müller, M. Q.; Dreier, F.; Ihling, C. H.; Schafer, M.; Sinz, A. Cleavable cross-linker for protein structure analysis: reliable identification of cross-linking products by tandem MS. *Anal. Chem.* **2010**, *82*, 6958–6968.
- (13) Lee, K. J.; Kang, D.; Park, H. S. Site-specific labeling of proteins using unnatural amino acids. *Mol. Cells* **2019**, *42*, 386–396.
- (14) Bayley, H. *Photogenerated Reagents in Biochemistry and Molecular Biology*; Elsevier: Amsterdam, 1983.

- (15) Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.
- (16) Fancy, D. A.; Kodadek, T. Chemistry for the analysis of protein-protein interactions: rapid and efficient cross-linking triggered by long wavelength light. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6020–6024.
- (17) Farrell, I. S.; Toroney, R.; Hazen, J. L.; Mehl, R. A.; Chin, J. W. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nat. Methods* **2005**, *2*, 377–384.
- (18) Iacobucci, C.; Götze, M.; Piotrowski, C.; Arlt, C.; Rehkamp, A.; Ihling, C.; Hage, C.; Sinz, A. Carboxyl-photo-reactive ms-cleavable cross-linkers: unveiling a hidden aspect of diazirine-based reagents. *Anal. Chem.* **2018**, *90*, 2805–2809.
- (19) Ge, S.-S.; Chen, B.; Wu, Y.-Y.; Long, Q.-S.; Zhao, Y.-L.; Wang, P.-Y.; Yang, S. Current advances of carbene-mediated photoaffinity labeling in medicinal chemistry. *RSC Adv.* **2018**, *8*, 29428–29454.
- (20) Müller, F.; Graziadei, A.; Rappsilber, J. Quantitative photo-crosslinking mass spectrometry revealing protein structure response to environmental changes. *Anal. Chem.* **2019**, *91*, 9041–9048.
- (21) Yanagisawa, T.; Hino, N.; Iraha, F.; Mukai, T.; Sakamoto, K.; Yokoyama, S. Wide-range protein photo-crosslinking achieved by a genetically encoded *N*ε-(benzyloxycarbonyl)lysine derivative with a diazirinyl moiety. *Mol. Biosyst.* **2012**, *8*, 1131–1135.
- (22) Yang, T.; Li, X.-M.; Bao, X.; Fung, Y. M. E.; Li, X. D. Photo-lysine captures proteins that bind lysine post-translational modifications. *Nat. Chem. Biol.* **2016**, *12*, 70–72.
- (23) Ziebell, M. R.; Nirthanan, S.; Husain, S. S.; Miller, K. W.; Cohen, J. B. Identification of binding sites in the nicotinic acetylcholine receptor for [³H]azetomidate, a photoactivatable general anesthetic. *J. Biol. Chem.* **2004**, *279*, 17640–17649.
- (24) Zhang, S.; He, D.; Lin, Z.; Yang, Y.; Song, H.; Chen, P. R. Conditional chaperone-client interactions revealed by genetically encoded photo-cross-linkers. *Acc. Chem. Res.* **2017**, *50*, 1184–1192.
- (25) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of p-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **2002**, *124*, 9026–9027.
- (26) Hino, N.; Oyama, M.; Sato, A.; Mukai, T.; Iraha, F.; Hayashi, A.; Kozuka-Hata, H.; Yamamoto, T.; Yokoyama, S.; Sakamoto, K. Genetic incorporation of a photo-crosslinkable amino acid reveals novel protein complexes with GRB2 in mammalian cells. *J. Mol. Biol.* **2011**, *406*, 343–353.
- (27) Lin, S.; Zhang, Z.; Xu, H.; Li, L.; Chen, S.; Li, J.; Hao, Z.; Chen, P. R. Site-specific incorporation of photo-cross-linker and bioorthogonal amino acids into enteric bacterial pathogens. *J. Am. Chem. Soc.* **2011**, *133*, 20581–20587.
- (28) Tanaka, Y.; Bond, M. R.; Kohler, J. J. Photocrosslinkers illuminate interactions in living cells. *Mol. Biosyst.* **2008**, *4*, 473–480.
- (29) Yanagisawa, T.; Ishii, R.; Fukunaga, R.; Kobayashi, T.; Sakamoto, K.; Yokoyama, S. Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode *N*ε-(o-Azidobenzyloxycarbonyl) lysine for Site-Specific Protein Modification. *Chem. Biol.* **2008**, *15*, 1187–1197.
- (30) Zhang, M.; Lin, S.; Song, X.; Liu, J.; Fu, Y.; Ge, X.; Fu, X.; Chang, Z.; Chen, P. R. A genetically incorporated crosslinker reveals chaperone cooperation in acid resistance. *Nat. Chem. Biol.* **2011**, *7*, 671–677.
- (31) Piotrowski, C.; Ihling, C. H.; Sinz, A. Extending the cross-linking/mass spectrometry strategy: Facile incorporation of photo-activatable amino acids into the model protein calmodulin in *Escherichia coli* cells. *Methods* **2015**, *89*, 121–127.
- (32) Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. *Molecular Cell Biology*, 4th ed.; W. H. Freeman, 2000.
- (33) Marley, J.; Lu, M.; Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **2001**, *20*, 71–75.
- (34) Reilly, D.; Fairbrother, W. J. A novel isotope labeling protocol for bacterially expressed proteins. *J. Biomol. NMR* **1994**, *4*, 459–462.
- (35) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **2008**, *26*, 1367–1372.
- (36) The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **2019**, *47*, D506–D515.
- (37) Tyanova, S.; Temu, T.; Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **2016**, *11*, 2301–2319.
- (38) Götze, M.; Pettelkau, J.; Fritzsche, R.; Ihling, C. H.; Schäfer, M.; Sinz, A. Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 83–97.
- (39) Morgado, L.; Burmann, B. M.; Sharpe, T.; Mazur, A.; Hiller, S. The dynamic dimer structure of the chaperone Trigger Factor. *Nat. Commun.* **2017**, *8*, 1992.
- (40) Saio, T.; Kawagoe, S.; Ishimori, K.; Kalodimos, C. G. Oligomerization of a molecular chaperone modulates its activity. *eLife* **2018**, *7*, No. e35731.
- (41) Dekker, C.; de Kruijff, B.; Gros, P. Crystal structure of SecB from *Escherichia coli*. *J. Struct. Biol.* **2003**, *144*, 313–319.
- (42) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yilmaz, S.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaino, J. A. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **2019**, *47*, D442–D450.