Published in final edited form as:
J Proteome Res. 2009 July ; 8(7): 3702–3711. doi:10.1021/pr900257z.

Selective Enrichment of Azide-Containing Peptides from Complex Mixtures

Merel A. Nessen, **Gertjan Kramer**, **JaapWillem Back**[†], **Jeremy M. Baskin**[††], **Linde E. J. Smeenk**, **Leo J. de Koning**, **Jan H. van Maarseveen**, **Luitzen de Jong**, **Carolyn R. Bertozzi**[††,‡,§,§§], **Henk Hiemstra**, and **Chris G. de Koster**[*]

Mass Spectrometry of Biomacromolecules, Swammerdam Institute for Life Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018WV Amsterdam, The Netherlands and Organic Synthesis, Van 't Hoff Institute for Molecular Sciences, University of Amsterdam, Nieuwe Achtergracht 129, 1018WS Amsterdam, The Netherlands.

[††]Department of Chemistry, University of California, Berkeley, CA 94720

[‡]Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720

[§]Howard Hughes Medical Institute, University of California, Berkeley, CA 94720

[§§]The Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

Abstract

A general method is described to sequester peptides containing azides from complex peptide mixtures, aimed at facilitating mass spectrometric analysis to study different aspects of proteome dynamics. The enrichment method is based on covalent capture of azide-containing peptides by the azide-reactive cyclooctyne (ARCO) resin and is demonstrated for two different applications. Enrichment of peptides derived from cytochrome *c* treated with the azide-containing cross-linker bis (succinimidyl)-3-azidomethyl glutarate (BAMG) shows several cross-link containing peptides. Sequestration of peptides derived from an *Escherichia coli* proteome, pulse labeled with the bio-orthogonal amino acid azidohomoalanine as substitute for methionine, allows identification of numerous newly synthesized proteins. Furthermore, the method is found to be very specific, as after enrichment over 87% of all peptides contain (modified) azidohomoalanine.

Keywords

Azide; azidohomoalanine; azide-alkyne cycloaddition; cross-linking; cyclooctyne; enrichment; mass spectrometry; proteomics; peptides; pulse-chase labeling

*To whom correspondence should be addressed. Chris G. de Koster, Telephone: +31-(0)20-525 5457, Fax: +31-(0)20-525 6971, Email: c.g.dekoster@uva.nl..
[†]Current address: Pepscan Therapeutics BV, Zuidersluisweg 2, 8243 RC Lelystad, The Netherlands.

Supporting information available: Mass spectra of the Cu^{I}-catalyzed 1,3-dipolar cylcoaddition of azide-containing peptides with terminal alkynes can be found in Figure 1 of Supporting Information I. A complete overview of all peptides derived from a tryptic digest of cytochrome *c* cross-linked with BAMG assigned in the MALDI-TOF spectra before and after enrichment can be found in Table 1 of Supporting Information I. Detailed experimental descriptions of determining the loading of the ARCO-resin and optimization of the enrichment procedure can be found in Supporting Information I as well. Supporting Information II contains an Excel file of all identified proteins from a labeled *E. coli* proteome before and after enrichment, using the different LC conditions as described in *Material and methods* and discussed in *Results and discussion*. Results from decoy database are included as well. This material is available free at http://pubs.acs.org.

Introduction

Profound understanding of cellular behavior requires knowledge about cellular proteome dynamics. Information on changes in protein synthesis rates, posttranslational modifications and protein-protein interactions on a proteome-wide scale upon environmental changes will reveal regulatory circuits underlying cellular adaptation. We describe here the development of a general method for targeted peptide isolation based on azide-chemistry aimed at facilitating mass spectrometric analysis of these different aspects of cellular proteome dynamics.

Many proteomic studies are focused on identification and quantification of as many proteins as possible. The combination of two dimensional gel electrophoresis [1] or, more recently, two (or three) dimensional liquid chromatographic methods [2-4] with mass spectrometry has led to the identification and quantification of up to hundreds of proteins for *Escherichia coli* [5, 6]. Isolated proteins or proteomes are subjected to digestion and subsequent mass spectrometric analysis allows identification and quantification of the proteins. However, the enormous complexity of peptide mixtures from complete proteomes implies that even with the most advanced fractionation methods and mass spectrometric equipment, identification and quantitation is often based on only a few or even single unique peptides. In studies in which the interest lays in a subset of peptides, such as in phosphoproteomics [7, 8], this might hamper identification of peptides of interest.

For the sequestration of (bio)molecules a functionality is needed that can be used to selectively modify the molecules of interest. The unique properties of the azide make it a well suited chemical group for this purpose. The azide is inert towards the biological system, though it does provide a chemical handle that can be used for selective modification, using reactions such as the Staudinger ligation [9-12], Cu^{I}-catalyzed (3+2)-cycloaddition [13, 14] and the strain-promoted (3+2) cycloaddition [15-18]. Several studies have shown the applicability of azides in biological systems [9, 12, 13, 15-23] ; to visualize glycan trafficking in the cell, azide-reactive fluorophores can be used [17, 21] and global *de novo* protein synthesis has been monitored by using azide-containing amino acids [20, 24, 25].

Studying the different aspects of proteome dynamics by mass spectrometry requires selective analysis of subsets of peptides, which are post-translationally modified, cross-linked or newly synthesized. Peptides derived from these groups have in common that they are usually rare and, therefore, have a high chance to escape detection in a background of more abundant (unmodified) peptides. Selective enrichment of rare (azide-containing) peptides from complex mixtures will aid in their identification and thereby facilitate obtaining information on proteome dynamics.

The use of azide-containing glycans [26], lipids [27] or azide-containing reactive reagents directed towards posttranslational modifications in combination with a sequestration method for azides will allow the mapping of posttranslational modification sites. Protein cross-linking can provide information on spatial distance constraints to reveal protein-protein interactions and to enable validation of 3-D structure models. The mapping of linked amino acids will be tremendously facilitated by sequestration of the rare and low abundant cross-linked peptides in protein digests, using azide-containing cross-linkers, such as the recently described bis (succinimidyl)-3-azidomethyl glutarate (BAMG; Figure 1c) [28, 29]. Furthermore, for identification of newly synthesized proteins pulse-chase experiments can be performed, using the methionine-analogue azidohomoalanine (azhal; Figure 1b) [20, 24, 25].

Here we present a solid phase purification method that highly enriches azide-containing peptides from complex peptide mixtures. In Figure 1a, the workflow of the method is given. The enrichment method makes use of the azide reactive cyclooctyne resin (ARCO-resin)

depicted in Figure 1d. Azide-containing peptides are captured on the resin via the strain-promoted (3+2) cycloaddition as has been described by Agard *et al*. [15]. This click-reaction of azides and cyclooctynes is very selective and can be carried out both under physiological conditions and in presence of different organic solvents. The covalent capturing enables stringent washing conditions to reduce nonspecific binding. A linker that can be cleaved upon reduction is included to release captured azide-peptides from the resin allowing further purification and analysis. Mass spectrometric analysis of the enriched samples reveals the identity of the sequestered peptides. We demonstrate this enrichment method to work for azide-containing peptides from different origin, including azhal-containing peptides from an azhal-labeled *E. coli* proteome, allowing identification of up to 90 (newly synthesized) proteins.

Materials and Methods

Synthesis of L-azhal

L-azhal was synthesized from L-Boc-2,4-diaminobutyric acid (Chem-Impex, Wood Dale, USA) by a diazotransfer, using triflic azide, as previously described [30].

Synthesis of cyclooctyne probe

The cyclooctyne carboxylic acid was synthesized as described by Agard *et al*. [16]. The protocol was adjusted for the workup of the cyclooctyne methyl ester, doubling the yield of this reaction step. Instead of using water, the reaction mixture was quenched with saturated NH_4Cl, extracted with EtOAc and dried over $MgSO_4$. After column chromatography (silica gel, 9:1 hexanes:EtOAc), the product was isolated as a colorless oil (50%).

Synthesis of the azide reactive cyclooctyne (ARCO)-resin

250 mg of PL-DMA resin [31] (1.0 mmol/g 55-250μm (Polymer Laboratories, Shropshire, United Kingdom) was washed with 10 mL DMF, 10 mL CH_2Cl_2 and 10 mL DMF for 15 minutes each, prior use. The resin was incubated for three days with 15 mL of a 1 M cystamine solution (1 eq cystamine·2HCl, 2.5 eq DBU in DMF) in a reaction vessel at 30 °C under nitrogen. Positive coupling of cystamine to the PL-DMA beads was determined by bromo-phenol blue [32]. After removal of the cystamine, the resin was washed by 10 mL DMF, 10 mL CH_2Cl_2, 10 mL DMF and 10 mL CH_2Cl_2 for 15 minutes respectively. The linker derivatized resin (1.2 eq) was then incubated with a solution of HATU (O-(7-azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate) (1 eq), cyclooctyne carboxylic acid (1 eq) and diisopropyl-ethyl-amine (DIPEA) (2 eq) in dry DMF for 3 hours at room temperature. After a negative bromo-phenol blue test, the resin was washed as described above and dried under nitrogen. This yielded the ARCO-resin with a cyclooctyne loading of 0.25 mmol/g (see Supporting Information I for experimental details).

Optimization of enrichment procedure

Acetylated peptide (Ac-H_3) was incubated with the ARCO-resin as described below for 24, 48 and 72 hours at room temperature and 40 °C, under constant mixing. More details on the preparation of acetylated peptide and the enrichment can be found in Supporting Information I. For quantification D_3-acetylated model peptide and the H_3-acetylated peptide (before addition) were mixed in 1:1, 1:4 and 4:1 ratios. The flow-through and enriched product were mixed in a ratio of 3:1 (H_3:D_3) with the D_3-acetylated peptide, assuming either nothing had reacted (ft) or all had been captured and released (ep). Mixed samples were diluted to < 5% acetonitrile with 0.1% TFA and desalted immediately, using 2 μg capacity C18 tips (Omix, Varian Inc., Lake Forest, Canada). The desalted peptides were dried in a vacuum centrifuge and subsequently resuspended in 100 mM NH_4HCO_3 for overnight digestion with trypsin (sequencing grade, Roche, Mannheim, Germany) at 37 °C using a 1:25 (w/w) protease:protein

ratio. The digested peptide mixture was then dried in a vacuum centrifuge, resuspended in 0.1% TFA and spotted for MALDI-TOF analysis.

Preparation of peptide samples for enrichment

Photoactive Yellow Protein was produced as described in Back *et al*. [33]. Isolated protein was subjected to overnight digestion at 37 °C using a 1:50 (w/w) protease:protein ratio, with TPCK treated trypsin (Sigma-Aldrich, St. Louis, USA) in 50 mM potassium phosphate buffer pH 7.5. After digestion the peptide mixture was used for capturing experiments without any purification. A tryptic digest of BAMG-cross-linked cytochrome *c* was prepared as described by Kasper *et al*. [28]. A tryptic digest of an azhal labeled *E. coli* proteome was prepared as described by Kramer *et al*. [25]. Deviations from this protocol are described below. After washing, cells were transferred to M9 minimal medium in which methionine was replaced by 400 mg/L azhal and cells were grown for 2 hours in a shake flask at 37 °C. After tryptic digestion, reducing and alkylating reagents were removed with C18 tips, eluting in 75% acetonitrile.

Enrichment procedure

Digests (AzPYP, 50 μg (~3 nmol), 75 μL; cross-linked cytochrome *c*, 20 μg (~1.6 nmol), 60 μL; *E. coli* proteome, 400 μg, 150 μL) in a solution of 50% acetonitrile and 50% 50 mM potassium phosphate buffer pH 7.5, were added to dry beads. The mixtures were incubated with ~2 mg resin (AzPYP & cross-linked cytochrome *c*) or ~5 mg resin (*E. coli* proteome) at 40 °C for 24 hours, under constant mixing. To collect the supernatant the mixture was centrifugated. The beads were washed twice with a solution of 50% acetonitrile and 50% 50 mM potassium phosphate buffer pH 7.5, followed by washes of 15 minutes with acetonitrile; 50% acetonitrile, 50% 50 mM potassium phosphate buffer pH 7.5; 50 mM potassium phosphate buffer pH 7.5; 2 M NaCl and 50 mM potassium phosphate buffer pH 7.5. For cleavage of the disulfide linker, the beads were incubated with 5 mM TCEP in 50 mM potassium phosphate buffer pH 7.5 for 1 hour at room temperature, followed by 2.5 mM TCEP in a solution of 50% acetonitrile and 50% 50 mM potassium phosphate buffer pH 7.5 for 15 minutes. The two fractions were combined and iodoacetamide (55 mM final concentration) was added for 30 minutes at room temperature in de the dark. Reducing and alkylating reagents were removed by desalting the peptide mixture on C18 reversed phase tips.

Strong Cation Exchange chromatography

For the labeled *E. coli* proteome sample 100 μg of protein digest was loaded, for the enriched peptide mixture all recovered material was loaded. Peptide mixtures were separated on a PolySULPHOETHYL Aspartamide™ Column (2.1mm ID, 10 cm length) (PolyLC Inc., Columbia, USA) on an Ultimate HPLC system equipped with a Probot fraction collector (LC Packings, Amsterdam, The Netherlands). Elution (flow rate: 0.1 mL/min) was performed using a gradient of 30 minutes, from 0 to 250 mM KCl over the first 22.5 minutes, ending at 500 mM KCl after 30 minutes, with 20 mM potassium phosphate buffer pH 2.9, 20% acetonitrile as buffer. 30 fractions of 100 μL were collected and lyophilized; fractions showing absorbance at 214 nm were redisolved in 0.1% TFA, 5% acetonitrile and used without any further purification for LC-ESI-Q-TOF analysis as described below.

Mass spectrometry

For reflectron MALDI-TOF, 0.5 μL sample was mixed with 0.5 μL of a 10 mg/mL α-hydroxy cinnaminic acid solution in 1:1 (v:v) acetonitrile:ethanol and subsequently spotted on a MALDI target plate and dried. Spectra were recorded on a Tofspec 2EC mass spectrometer (Micromass, Wythenshawe, U.K.) provided with a 2 GHz digitizer. For analysis by the CoolToolBox

software program (see below) spectra were deconvoluted using the MaxEnt3 algorithm included in the Masslynx Proteinlynx software.

Electrospray MS and low energy collision induced dissociation (MS/MS) analyses were performed on a Q-TOF (Micromass, Whyttenshawe, UK) mass spectrometer with a Z-Spray orthogonal ESI source. For LC prior to MS, peptide mixtures were loaded on an Ultimate nano-HPLC system (LC Packings, Amsterdam, The Netherlands) and separated on a PepMap100 C_{18} reversed phase column (75 μm I.D., 25 cm length; Dionex, Sunnyvale, CA, USA). For untreated samples ~10 pmol of material was loaded, for enriched AzPYP samples ~5% of the enriched material was injected. For the SCX fractions of the proteome sample before enrichment an amount corresponding to 10-15 mAU absorption was loaded. SCX fractions of the enriched *E. coli* proteome samples were split into two and run with both a linear (gradient B) and a step gradient (gradient C; see below). After loading on a trap column (Acclaim PepMap100, 300 μm I.D., 5 mm length, Dionex, Sunnyvale, CA, USA), elution (flow rate: 0.3 μl/min) of the peptides was performed using a linear gradient (gradient B), using solvent A, 0.1% formic acid and solvent B, 50% acetonitrile with 0.1% formic acid, going from 0% to 15% B over the first 12 min, then continuing to 45% B from 12-35 min and ending at 100% B at 45 min (gradient returned back to 0% B in 2 minutes). The step-gradient (gradient C) combined a linear and a step gradient, using solvent A, 0.1% formic acid and solvent C, 95% acetonitrile with 0.1% formic acid. After loading the following gradient was used; 0-5 min from 0 to 16% C, 5-8 min from 16 to 19% C, 8-10 min 19% C, 10.2-12.2 min 22% C, 12.4-14.4 min 24% C, 14.6-16.6 min 26% C, 16.8-18.8 min 28% C, 19-21 min 31% C, 21-27 min from 31 to 100% C, 27-37 min 100% C, and 37-40 min from 100% C to 0% C. Direct infusion of the flow was supported by a Nanobore Emitter (Proxeon, Odense, Denmark). Survey scans were acquired from m/z 350–1,750. For MS/MS most intense ions were selected in a data-dependent mode, recorded from m/z 50-2,500, scan time 1.00 sec, interscan delay 0.10 sec with argon as collision gas at a pressure of 4×10^{-5} bar measured on the quadrupole pressure gauge. Fragmentation spectra were processed with the MaxEnt3 algorithm included in the Masslynx Proteinlynx software, generating pkl-files (peak list).

Data Analysis

Generated peak lists were submitted to Mascot (version 2.1, Matrix Science, London, UK) using an in-house license. Fragmentation spectra from AzPYP digests were searched against a database to which the sequence of the PYP construct was added. Analysis was carried out using following settings: TrypChymo, cleaving after K, R, Y, W, F and L except when a P follows, allowing up to 6 missed cleavages. No fixed modifications of cysteine residues were set, since the only cysteine residue in PYP is conjugated with the chromophore. Variable modifications included carbamidomethylation of cysteine; methionine replaced by azhal ($C_4H_6N_4O$, 126.054161) and methionine replaced by a modified azhal residue ($C_{24}H_{32}N_6O_3S$, 484.225660, ignore 402.196371 and 268.144987). Peptide tolerance was set at 0.5 Da, MS/MS tolerance at 0.5 Da, peptide charge at 1+. Assigned fragmentation spectra were validated manually.

For *E. coli* data, fragment ion searches were carried out using a local database of the *E. coli* K12 proteome (4506 sequences; 1426768 residues, release 14.4 04/11/2008, Uniprot consortium, http://beta.uniprot.org/). Settings were as follows: Trypsin, allowing up to 1 missed cleavage; carbamidomethylation of cysteine residues as a fixed modification, variable modifications of methionine as described for AzPYP. Peptide tolerance was set at 0.8 Da, MS/MS tolerance at 0.8 Da, peptide charge at 1+. To estimate false positive protein identification rates (FPRs), fragment ion searches were performed against a decoy database, made using the Peakhardt decoy database builder (Medizinisches Proteom Center, Bochum, Germany; http://www.medizinisches-proteom-center.de), generating a shuffled version of the *E. coli* K12

proteome. FPRs were obtained by dividing the total number of protein hits from the decoy database by the total number of protein hits from the *E. coli* K12 database and multiplying this ratio by 100 percent. The Mascot output was filtered by using: a significance threshold of $p < 0.05$, MudPiT scoring on, ion score cutoff of 25 and selection of "require bold red". For these search settings FPRs of less than 1.1% were obtained. Peptide summary reports were exported as csv-files (comma separated value) for further data-analysis in Excel (Microsoft Corporation, Redmond, USA). Exported Mascot results were filtered using the ASAP utilities add-on for Microsoft Excel (A Must in Every Office BV, Zwolle, the Netherlands; http://www.asap-utilities.com/) for modified azhal peptides. Per experiment two files were generated containing proteins identified by either all unique peptides or by unique (modified) azhal-containing peptides. To calculate labeling numbers and enrichment percentages, methionine and azhal-containing peptides with the same primary sequence were counted separately.

Identifications of cross-linked peptides (before and after enrichment) were made by analysis of the MS data using the CoolToolBox software program. This program is an upgrade of our VIRTUALMSLAB [34] software program and can perform virtual cross-link experiments, generating a mass spectrometric reference database, which can be matched with the experimental mass spectrometric data. The reference databases before enrichment consisted of i) unmodified tryptic peptides that contain up to 6 "missed" cleavages at lysine (K), ii) peptides modified with the heme group present in cytochrome *c* and iii) cross-linked, loop-linked and mono-linked peptides with BAMG, iv) a combination of ii) and iii). The reference database after enrichment is similar to the one before enrichment, except that the mass added for modification by the cross-linker (ad iii) and iv)) was increased by the cyclooctyne and the linker, adding 358.2 Da. The MALDI-TOF data was matched with a mass accuracy of at least 60 ppm.

Results and Discussion

Choice of materials

One of the most important prerequisites developing a sequestration method for (bio)molecules is selectivity. To achieve this, the (immobilized) chemical reagent to be used for binding to the desired functional group should be specific, leaving other functionalities unaffected. In addition, the solid support being used should have as little nonspecific binding as possible. Furthermore, leaching of monomers or other impurities from the matrix of the solid support should be negligible, this becomes particularly important at low analyte concentrations.

For azides several selective reactions have been described, of which in our hands the strain promoted cycloaddition with cyclooctyne was found to be the most effective. The click-reaction using terminal alkynes and Cu^{I} has been tested as well, but these conditions may cause Cu^{I}-induced oxidation of peptides, which might hamper identification (see Figure 1 in Supporting Information I). In contrast, the cyclooctyne reagents give a clean and selective reaction with azides in peptides and it was therefore chosen as reactive group for our purification method.

In our application controlled pore glass (CPG) and poly ethylene glycol (PEG) beads were found not to be suitable, showing either nonspecific binding of peptides or release of PEG molecules upon use. On the other hand, poly-dimethylacrylamide (PL-DMA), a hydrophilic resin, swells well in polar solvents, shows very little nonspecific binding and leaching, making it a good choice as a solid support.

To release sequestered peptides from the solid support, a disulfide type cleavable linker was included. Cleavage of the S-S bond can be induced under mild conditions and is very efficient.

Subsequent alkylation of the released peptides prevents any possible (re)oxidation with free thiol groups (from other released peptides) in solution.

In Figure 1d the resin is depicted, consisting of PL-DMA loaded with a disulfide cleavable linker and a cyclooctyne reactive group for capturing azide-containing peptides on the resin by the strain-promoted 1,3-dipolar cycloaddition.

Synthesis of the ARCO-resin

Synthesis of the ARCO-resin started from the commercially available poly-dimethylacrylamide (PL-DMA) resin. The cleavable S-S bond linker was introduced by reaction of cystamine with the methyl ester on the PL-DMA resin, yielding an amide bond. Coupling of the cyclooctyne to the amine of the linker was accomplished by activation of the acid moiety using the uronium/guanidinium type peptide coupling reagent HATU [35], obtaining the ARCO-resin (Figure 1d).

The ARCO-resin (yellowish gel-like beads) swells very well in a one-to-one mixture of water and acetonitrile. The amount of reactive groups was determined after reaction of the beads with Fmoc-azhal followed by cleavage of the Fmoc-group by a base to liberate the fluorenyl group which absorbs at 300 nm. An excess of Fmoc-azhal in DMF was incubated with the beads for 24 hours at room temperature. After washing and drying of the beads, the loading was determined to be 0.27 mmol/g, corresponding to a yield of 30% over three steps.

To determine the loading under the reaction conditions for azide-containing peptide enrichment the reaction was carried out in 50% acetonitrile 50% 50 mM potassium phosphate buffer pH 7.5 at room temperature and 40 °C (see *Enrichment of a model peptide*). This gave similar loading values (0.25 mmol/g) as performing the reaction in DMF, showing the reaction to be efficient in both DMF and a combination of acetonitrile and phosphate buffer. Incubation at higher temperatures shows no increase in loading, so we assume the reaction to be complete at room temperature when performed with an excess of Fmoc-azhal. Efficiency of the reductive cleavage of the S-S bond of the linker was determined using beads which were reacted with Fmoc-azhal. Incubation with different reducing agents at different concentrations and incubation times showed the reduction to be most effective using 5 mM TCEP for 1 hour at room temperature at pH 7.5, by which >90% of all Fmoc-azhal groups were cleaved from the resin.

Capture of a model peptide

To test the capture of azide-containing peptides on the resin, a model octadeca peptide (Pan016) containing one azhal residue was captured and released as described in *Material and methods* and depicted in Figure 1d. Upon cycloaddition, a triazole is formed and peptides are captured on the resin via their azide moiety. Reductive cleavage of the disulfide bond liberates a thiol which is alkylated by iodoacetamide. The mass of the released peptide obtained will increase with the mass of the cyclooctyne, linker and alkylating reagent used, resulting in an addition of 358.2 Da (see Figure 1e and 2).

Pan016 was captured from a solution of 50% acetonitrile and 50% 50 mM potassium phosphate buffer pH 7.5. The PL DMA beads were found to be handled best in 50% acetonitrile and furthermore the addition of acetonitrile reduces possible hydrophobic interactions of the peptides. To optimize the capture conditions further, the model peptide was captured at room temperature and 40 °C for 24, 48 and 72 hours.

Capture and release of the model peptide yields one single product after cleavage, with the expected modification, as is shown in Figure 2c. In the depicted MALDI-TOF spectra, a shift in mass can be observed after capture and release of the peptide, corresponding to the mass of

the linker, 358.2 Da. In addition, the typical N_2-loss and subsequent 1 or 2 proton uptake observed for azide-containing peptides, is not observed any more after modification (see Figure 2) [28]. Fragmentation of the product shows as well that the azidohomoalanine residue has been modified (data not shown). This shows the reaction to be effective and specific and allows further exploration of the method.

Optimization and kinetics of capture/release of a model peptide

Quantification of the capture and release of the model peptide was done mass spectrometrically, using isotopically labeled Pan016 as an internal standard. For these experiments, Pan016 was labeled by treatment with D_6-acetic anhydride providing Ac(D_3)-Pan016. The reference peptide was mixed in a 1:3 ratio with i) the flow through, ii) the sequestered Ac(H_3)-Pan016 and as a control with iii) Ac(H_3)-Pan016. In Figure 3 an overview of the quantification for the different coupling experiments tested is given. It shows that at room temperature the reaction proceeds well but relatively slowly. After 24 hours, most of the peptide is still detected in the flow-through, and after 72 hours of incubation around 40% of the model peptide is captured. Rising the incubation temperature to 40 °C significantly increases the rate of the reaction (Figure 3b) and after 24 hours practically all peptide was captured. The yield of the enrichment at 40 °C is near 50% and doesn't increase after prolonged incubation.

Enrichment of azhal-containing peptides derived from a single protein

To show the efficiency of the enrichment of azide-containing peptides from more complex peptide mixtures, a digest of an azide-containing single protein, the photoactive yellow protein (PYP), was incubated with the ARCO-resin for 24h at 40 °C. PYP contains six methionine residues, evenly distributed in the protein, and provides after tryptic digestion differently sized peptides, with and without methionine residues.

Mass spectrometric analysis shows an enormous enrichment of the azhal peptides as is depicted in Figure 4. Before capturing, 11 out of the 22 peptides (50%) detected contain an azhal residue. After capture/release this amount is increased to 18 out of 19, an enrichment of 95 %. With MALDI-TOF even the N-terminal peptide consisting of only two amino acids (MR), which could not detected before is observed. As is shown in Figure 4, four out of the six azhal residues are detected before enriching the sample. After the purification all six azhal residues are retrieved, modified with the cyclooctyne and the linker. Thus, apart from an increase in the number of peptides found to contain (modified) azhal, the enrichment allows us to detect all azhal-residues as well.

The only non-azhal peptide identified in the enriched peptide mixture contains the cysteine residue, normally conjugated to the chromophore. Before enrichment this peptide is detected with and without chromophore (no alkylation was carried out for tryptic digestion). After sequestration this peptide was observed only with an alkylated cysteine residue, suggesting that disulfide exchange occurred, selectively capturing thiol containing species. This shows the need to alkylate all (free) thiols before incubation with the resin, to prevent covalent binding to the resin via a disulfide exchange reaction.

Enrichment of azide-containing peptides from *in vitro* BAMG-cross-linked cytochrome *c*

Apart from peptides with azide-containing amino acids, our purification method is applicable to sequester peptides modified via chemical cross-linking, using cross-linkers containing an azide in the spacer such as BAMG (see Figure 1c). This amine-reactive bis(succinimidyl) ester is used to covalently link primary amines in a protein or protein complex, which are found in the lysine residues and the N-terminus. From the linked amines, distances can be inferred restricted by the length of the spacer of the cross-linker. Hereby models of proteins can be validated and refined and interaction sites of protein complexes can be determined [36-38].

Two major limitations hamper mass spectrometric cross-link analysis of large proteins or mixtures of proteins. The first is the exponential increase of entries in calculated databases of cross-linked peptides with an increase in protein complexity [28, 34, 39, 40]. This can easily give rise to unmanageable large numbers of candidate cross-linked species corresponding to a given measured mass. The second major limitation is the low abundance of cross-linked peptides, leading to weak mass signals hampering identification by MS/MS.

The azide-group in BAMG renders cross-linked peptides scissile in a way that, for a given measured mass, only one candidate for a cross-link can be derived from the database, thereby circumventing the first limitation [28, 29]. Sequestering BAMG-linked peptides on the ARCO resin in a parallel approach is expected to meet the second limitation of cross-link analysis, which in the future will allow analysis of protein complexes.

In this study cytochrome *c* was cross-linked *in vitro* by BAMG as a model system. After tryptic digestion of the cross-linked protein, MALDI-TOF spectra were recorded and subsequently analyzed using CoolToolBox (VMSL). A mixture of unmodified and several (putative) BAMG-containing peptides is observed in the spectrum before enrichment (Figure 5a). Mass spectrometric analysis of the enriched cross-linked peptide mixture shows, apart from a single unmodified peptide, only cross-linker-containing peptides with the expected modification (Figure 5b).

Different types of cross-link containing peptides are expected: i) loop-links or type 1 cross-links, a cross-link within one peptide ii) cross-links or type 2 cross-links, a cross-link between two different peptides and iii) mono-links or type 3 cross-links, modification of peptides by a partially hydrolyzed cross-linker [41, 42]. Of these, the first two give structural information on distances and can be used to refine models. The last type of cross-link gives information on the accessibility of the protein; only lysines available on the surface can be modified.

Before enrichment of the cross-linked cytochrome *c* digest, two cross-links, six loop-links and two mono-links were observed in the MALDI-TOF spectrum. Out of these ten cross-link containing peptides, five can be found again in the spectrum after enrichment. In total, 13 cross-link containing peptides were assigned in the MALDI-TOF spectrum of the peptide mixture after enrichment. These comprise two cross-linked peptides, which have not been observed before, eight loop-links and three mono-links. In Table 1 of Supporting Information I, a complete overview of all peptides assigned in the MALDI-TOF spectra before and after enrichment can be found.

The only non-BAMG containing peptide observed is the heme-containing peptide present in cytochrome *c*. Because of the hydrophobic character of the porphyrin ring an interaction of the heme with the cyclooctyne on the ARCO-resin is likely to occur. Only after reductive cleavage of the S-S linker, this heme-peptide together with the cyclooctyne is released from the resin.

The selective purification of several BAMG-containing peptides shows the applicability of the method to azide-containing biomolecules other than azhal-containing peptides. Combination of the sequestration with (2D)LC-MS/MS might reveal many more cross-links and more complex protein systems can be subjected to this enrichment method, thereby elucidating interaction partners and binding sites.

Enrichment of azhal-containing peptides derived from an *in vivo* labeled *Escherichia coli* proteome

We next demonstrate selective sequestering from complex peptide mixtures of azhal-containing peptides obtained after *in vivo* labeling of proteins synthesized by *E. coli*. For these

experiments, *E. coli* was grown with azhal for two hours as described in *Material and methods*. The obtained enriched peptide mixture was subsequently analyzed by LC-MS/MS.

First analyses showed that singly charged ions with m/z 359.18, 717.35 and 1075.52, dominated the spectra in LC-MS/MS. These signals emerged from the unreacted cyclooctyne moiety in single form (m/z 359.18), and as a dimer $[2M+H]^+$ (m/z 717.35) or a trimer $[3M+H]^+$ (m/z 1075.52). Apart from obscuring detection, the unreacted cyclooctyne reduced the performance of the C18 column enormously.

Removal of the unreacted cyclooctyne, which is uncharged at low pH, was accomplished by inclusion of a strong cation-exchange (SCX) chromatographic step in the protocol. For the reversed phase LC experiments it was observed that the enriched peptide mixtures show a delay in elution and elute within a relatively short time using standard gradients. The modification of the peptides by the cyclooctyne and the linker makes the peptides more hydrophobic and chromatographic behavior is dominated by the cyclooctyne group. Changing the gradient from linear to a combination of a linear and a step-gradient (see *Materials and methods* for details) improved the mass spectrometric analysis enormously; many more ions were selected for fragmentation and the number of peptides identified by Mascot is increased as well (see for comparison Supporting information II).

Yield of enrichment of E. coli derived azide-containing peptides—Based on UV absorption of SCX separation after the purification, roughly 8 μg of peptides was recovered from the azhal-labeled *E. coli* proteome after enrichment. To estimate the yield corresponding to this amount we take into account that before enrichment 17.1% of the unique peptides contain methionine, while 6.9% contain azhal (see Table 1). From this observation a maximum of 27.6 μg of enriched azhal peptides was expected, given that the experiment used 400 μg of azhal-labeled proteome. Below we show that 87.6% of the (8 μg) recovered material consists of azhal-containing peptides. This results in a yield of the enrichment for this experiment of 25.3%, which is half of the percentage found for the quantification experiments described earlier. In connection to this, it should be noted that peptides after modification may suffer from poor solubility in aqueous media. High acetonitrile concentrations were needed to elute the enriched peptide mixtures from the reversed phase column, already suggesting the hydrophobic character. In addition, it is known that the biotin-cyclooctyne conjugate has a poor solubility in water [16]. To overcome possible solubility problems the azacyclooctyne recently reported by Sletten and Bertozzi might be used instead [43], as it has a more hydrophilic character. Additionally, it possesses a slightly higher reactivity compared to the cyclooctyne used here.

Fragmentation behavior of enriched azide-containing peptides—Fragmentation of the peptide backbone is not affected by the tag on the azhal residue and Mascot analysis is not hampered. Modified azhal-containing peptides give regular MS/MS spectra, from which sequence information can readily be derived. In Figure 6, a typical fragmentation spectra of a peptide enriched from a labeled *E. coli* proteome is shown. All characteristic y”-type ions can be assigned as indicated in the spectrum. In addition, fragmentation of modified peptides yields elimination of the linker including the triazole. Both the remaining mother ion and a reporter ion at m/z 402.2 can be observed. Also the y”-type ion series of the peptide that lack the triazole can be observed in the spectrum. Further fragmentation of the reporter ion gives rise to cleavage of the amide bond, leaving a benzoylium ion at m/z 268.1. These observations open possibilities to use the reporter ion(s) in identification of peptides. However, in order to be able to use these ions, the frequency and mechanism of this fragmentation reaction should be studied into more detail.

For protein identification using Mascot, peptide scores will be influenced by this cleavage of the linker; less peaks can be assigned in the fragmentation spectra. This can be found back in

the data presented here as well; somewhat less fragmentation spectra can be assigned to peptides after enrichment, compared to before enrichment. Before enrichment, 55% of all queries can be assigned; after enrichment, 40% is assigned. However, the enormous enrichment achieved makes up for this difference.

Efficiency of enrichment of E. coli derived azide-containing peptides—Before mass spectrometric analysis of the peptide mixtures, digests were separated by SCX. In total 30 fractions were collected for peptide mixtures before and after enrichment. For the peptide mixture after enrichment, 17 sequential fractions showed absorbance at 214 nm. The first three fractions collected were left out of the subsequent analysis, as they contained predominantly the unreacted cyclooctyne, which bound strongly to the reversed phase column. The remaining 14 collected fractions were split into two and run with either a linear (up to 50% acetonitrile) or a step-gradient (up to 95% acetonitrile) on a LC-Q-TOF and fragmentation spectra were recorded. The data of the runs was combined for each gradient separately and analyzed using Mascot.

For the peptide digest before enrichment, 14 SCX fractions analogous to the ones of the enriched peptide mixture were run with a linear gradient (up to 50% acetonitrile) and after combining all data, the fragmentation spectra were analyzed using Mascot. In Supporting Information II a complete overview is given of all peptides and proteins identified.

Mascot analysis of the SCX fractions of the proteome before enrichment yielded 290 proteins, of which 43 could be identified by azhal-containing peptides, corresponding to 14.8% of all proteins (see Table 1). As discussed above, SCX fractions of the captured and released peptide mixture contained less material, which is reflected in the number of proteins identified. With the linear standard gradient in total 59 proteins were found of which 58 solely by azhal-containing peptides, the step-gradient yielded 87 and 78 proteins respectively. Combining the results of the two runs gives a total number of 89 identified proteins solely by the (modified) azhal peptides, which is 90.8% of all proteins found by Mascot (see Table 1).

To determine the efficiency of the enrichment on the peptide level unique peptides identified by Mascot were listed before and after sequestration (complete lists can be found in Supporting Information II). Before purification, out of the 809 unique peptides used for identification, only 56 contain azhal, corresponding to an occurrence of 6.9%. After enrichment, with the two gradient runs combined, a total of 145 unique peptides can be identified by Mascot analysis (Table 1), of which 127 peptides contain modified azhal. This is an enrichment of 87.6% on the peptide level, which is an enormous improvement.

To summarize, our selective capture/release and LC-MS/MS analysis of an azide-labeled *E. coli* proteome shows an enrichment of >87% of azide-containing species on both the peptide and protein level. For comparison, before enrichment only 14.8% of all proteins are identified by azhal peptides and of all peptides not more than 6.9% contain azhal. This shows the power of this solid phase purification method and the applicability to selectively enrich rare azide-containing peptides.

Comparison with other sequestration methods

Other purification methods for azide-containing peptides being described, include a method described by Dieterich *et al*., denoted as Bioorthogonal noncanonical amino acid tagging (BONCAT) [20]. In this approach azhal is used for identification of newly synthesized proteins in mammalian cells, in combination with deuterated leucine. Purification of labeled proteins is accomplished by conjugation of biotin to azhal-containing proteins via the Cu^{I} catalyzed cyclo-addition, after which the tagged proteins are sequestered on a Neutravidin affinity resin. On resin digestion yields a peptide mixture in which both deuterated leucine and azhal-

containing peptides represent newly synthesized proteins, resulting in more peptides per protein to be used for identification and quantification compared to the use of a single (labeled) amino acid. However, the method is elaborate and contains several critical steps as indicated by the authors [24]. Our method comprises only a few non-critical steps and the solid phase enrichment by covalent attachment allows stringent washing, minimizing nonspecific binding.

A second purification method for azide-containing peptides, recently published by our group, makes use of combined fractional diagonal chromatography (COFRADIC) in combination with an azide-specific cleavage reaction [25, 28, 33]. In this way over 500 newly synthesized proteins in *E. coli* could be identified in a 15 minute labeling time frame. However, the diagonal chromatographic approach is laborious and has a low specificity; less than 60% of the identified peptides using Mascot contain an azhal-residue. In contrast, the method described here increases the content of azhal-containing peptides to 87%.

Other solid phase enrichment methods for peptides described in literature that make use of covalent binding are limited in number. One example is the solid-phase isotope tagging method described by Zhou *et al*. [44], directed towards cysteinyl peptides. A thiol reactive group is attached via a photocleavable linker to glass beads and an ICAT-based isotope tag has been included for quantification. Galactose-induced changes in protein abundance in yeast were studied and with this enrichment method almost exclusively cysteinyl peptides were recovered, allowing identification and quantification of 82 proteins. A second example is the recently described method by Foettinger *et al*. [45], in which tryptophan containing peptides are captured on hydrazine beads. With this strategy on average 25 proteins could be identified by tryptophan-containing peptides in a single run and an enrichment of 86% is achieved. This is comparable to our method, in which 87% of all peptides identified contain (modified) azhal.

Future perspectives

In addition to the azide-containing biomolecules described in this paper, the method can be applied to other (bio)molecules as well. As mentioned in the introduction, glycans and lipids labeled with azides have been used as metabolic labels in bioorthogonal chemical reporter strategies. But apart from visualization, the azide can be used as well to profile posttranslational modifications on the proteome level. For example, the use of the azido analogues of glycans will allow selective enrichment of glycosylated peptides and thereby mapping of the glycoproteome. Other posttranslational modifications can be mapped as well, using azide-containing reagents directed towards modifications such as phosphorylation.

Conclusion

In conclusion, an enrichment method has been developed that allows selective enrichment for azide-containing peptides from complex peptide mixtures. The use of azhal makes the method suitable to determine protein synthesis rates. In addition, we have shown sequestration of peptides modified by a bifunctional cross-linker with an azide-conjugated spacer, showing the wide applicability of the method. This makes it possible to address questions on the different aspects of proteome dynamics as described in the introduction. Furthermore, apart from enriching for azide-containing peptides, the observation of reporter ions during fragmentation after enrichment by the ARCO-resin gives opportunities for advanced searches as well. Analysis of complex peptides mixtures after enrichment may benefit from improved solubility in aqueous media when using a more hydrophilic cycloctyne as reactive group on the resin to sequester azide-containing peptides.

Future challenges will be the application of the method to study changes in protein synthesis by growing *E. coli* under different conditions, to determine protein synthesis rates by pulse-labeling experiments with azhal. Furthermore, this method will be used for other azide-

containing species; peptides modified with the azide-containing cross-linker BAMG have been successfully captured to and released from the ARCO-resin, opening possibilities to map protein-protein interactions in more complex protein systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

The authors thank prof. dr. Bőck and dr. Thanbichler for providing *E. coli* strain MTD123. M.A.N. was supported by an ECHO-project grant of the Netherlands Organization for Scientific Research (NWO), C.R.B. acknowledges support from the National Institutes of Health (GM058867) and J.M.B. was supported by National Science Foundation and National Defense Science and Engineering predoctoral fellowships.

References

1. O'Farrell PH. J. Biol. Chem 1975;250:4007–4021. [PubMed: 236308]
2. Yates JR, Speicher S, Griffin PR, Hunkapiller T. Anal. Biochem 1993;214:397–408. [PubMed: 8109726]
3. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR, Garvik BM, Yates JR. Nat. Biotechnol 1999;17:676–682. [PubMed: 10404161]
4. Wolters DA, Washburn MP, Yates JR. Anal. Chem 2001;73:5683–5690. [PubMed: 11774908]
5. Corbin RW, Paliy O, Yang F, Shabanowitz J, Platt M, Lyons CE, Root K, McAuliffe J, Jordan MI, Kustu S, Soupene E, Hunt DF. Proc. Natl. Acad. Sci. U.S.A 2003;100:9232–9237. [PubMed: 12878731]
6. Maillet I, Berndt P, Malo C, Rodriguez S, Brunisholz RA, Pragai Z, Arnold S, Langen H, Wyss M. Proteomics 2007;7:1097–1106. [PubMed: 17366475]
7. van der Veken P, Dirksen EHC, Ruijter E, Elgersma RC, Heck AJR, Rijkers DTS, Slijper M, Liskamp RMJ. ChemBioChem 2005;6:2271–2280. [PubMed: 16254931]
8. Zhou HL, Watts JD, Aebersold R. Nat. Biotechnol 2001;19:375–378. [PubMed: 11283598]
9. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. Proc. Natl. Acad. Sci. U.S.A 2002;99:19–24. [PubMed: 11752401]
10. Kohn M, Breinbauer R. Angew. Chem., Int. Ed 2004;43:3106–3116.
11. Saxon E, Armstrong JI, Bertozzi CR. Org Lett 2000;2:2141–2143. [PubMed: 10891251]
12. Saxon E, Bertozzi CR. Science 2000;287:2007–2010. [PubMed: 10720325]
13. Link AJ, Tirrell DA. J. Am. Chem. Soc 2003;125:11164–11165. [PubMed: 16220915]
14. Rostovtsev VV, Green LG, Fokin VV, Sharpless KB. Angew. Chem., Int. Ed 2002;41:2596–2599.
15. Agard NJ, Prescher JA, Bertozzi CR. J. Am. Chem. Soc 2004;126:15046–15047. [PubMed: 15547999]
16. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. ACS Chem. Biol 2006;1:644–648. [PubMed: 17175580]
17. Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, Lo A, Codelli JA, Bertozzit CR. Proc. Natl. Acad. Sci. U.S.A 2007;104:16793–16797. [PubMed: 17942682]
18. Codelli JA, Baskin JM, Agard NJ, Berozzi CR. J. Am. Chem. Soc 2008;130:11486–11493. [PubMed: 18680289]
19. Baskin JM, Bertozzi CR. QSAR Comb. Sci 2007;26:1211–1219.
20. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM. Proc. Natl. Acad. Sci. U.S.A 2006;103:9482–9487. [PubMed: 16769897]
21. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. Science 2008;320:664–667. [PubMed: 18451302]
22. Strable E, Prasuhn DE, Udit AK, Brown S, Link AJ, Ngo JT, Lander G, Quispe J, Potter CS, Carragher B, Tirrell DA, Finn MG. Bioconjugate Chem 2008;19:866–875.

23. van Hest JCM, Kiick KL, Tirrell DA. J. Am. Chem. Soc 2000;122:1282–1288.
24. Dieterich DC, Lee JJ, Link AJ, Graumann J, Tirrell DA, Schuman EM. Nat. Protoc 2007;2:532–540. [PubMed: 17406607]
25. Kramer G, Springer RR, Back JW, Dekker HL, Nessen MA, Van Maarseveen JH, De Koning LJ, Hellingwerf KJ, De Jong L, De Koster CG. Mol. Cell. Proteomics. March 25;2009 paper in press2009
26. Laughlin ST, Bertozzi CR. Proc. Natl. Acad. Sci. U.S.A 2009;106:12–17. [PubMed: 19104067]
27. Prescher JA, Bertozzi CR. Nat. Chem. Biol 2005;1:13–21. [PubMed: 16407987]
28. Kasper PT, Back JW, Vitale M, Hartog AF, Roseboom W, de Koning LJ, van Maarseveen JH, Muijsers AO, de Koster CG, de Jong L. ChemBioChem 2007;8:1281–1292. [PubMed: 17600791]
29. De Jong, L.; Kasper, PT.; Back, JW.; van Maarseveen, JH.; de Koster, CG. Novel cross-linkers for obtaining structure information on molecule complexes. patent WO/2007/035080. 2007.
30. Lundquist JT, Pelletier JC. Org. Lett 2001;3:781–783. [PubMed: 11259061]
31. Arshady R. Colloid Polym. Sci 1990;268:948–958.
32. Gaggini F, Porcheddu A, Reginato G, Rodriquez M, Taddei M. J. Comb. Chem 2004;6:805–810. [PubMed: 15360217]
33. Back JW, David O, Kramer G, Masson G, Kasper PT, de Koning LJ, de Jong L, van Maarseveen JH, de Koster CG. Angew. Chem., Int. Ed 2005;44:7946–7950.
34. de Koning LJ, Kasper PT, Back JW, Nessen MA, Vanrobaeys F, Van Beeumen J, Gherardi E, de Koster CG, de Jong L. FEBS J 2006;273:281–291. [PubMed: 16403016]
35. Carpino LA. J. Am. Chem. Soc 1993;115:4397–4398.
36. Back JW, de Jong L, Muijsers AO, de Koster CG. J. Mol. Biol 2003;331:303–313. [PubMed: 12888339]
37. Sinz A. J. Mass Spectrom 2003;38:1225–1237. [PubMed: 14696200]
38. Lee YJ. Mol. BioSyst 2008;4:816–823. [PubMed: 18633483]
39. Maiolica A, Cittaro D, Borsotti D, Sennels L, Ciferri C, Tarricone C, Musacchio A, Rappsilber J. Mol. Cell. Proteomics 2007;6:2200–2211. [PubMed: 17921176]
40. Rinner O, Seebacher J, Walzthoeni T, Mueller LN, Beck M, Schmidt A, Mueller M, Aebersold R. Nat. Methods 2008;5:315–318. [PubMed: 18327264]
41. Schilling B, Row RH, Gibson BW, Guo X, Young MM. J. Am. Soc. Mass Spectrom 2003;14:834–850. [PubMed: 12892908]
42. Seebacher J, Mallick P, Zhang N, Eddes JS, Aebersold R, Gelb MH. J. Proteome Res 2006;5:2270–2282. [PubMed: 16944939]
43. Sletten EM, Bertozzi CR. Org. Lett 2008;10:3097–3099. [PubMed: 18549231]
44. Zhou HL, Ranish JA, Watts JD, Aebersold R. Nat. Biotechnol 2002;20:512–515. [PubMed: 11981568]
45. Foettinger A, Leitner A, Lindner W. J. Proteome Res 2007;6:3827–3834. [PubMed: 17655347]

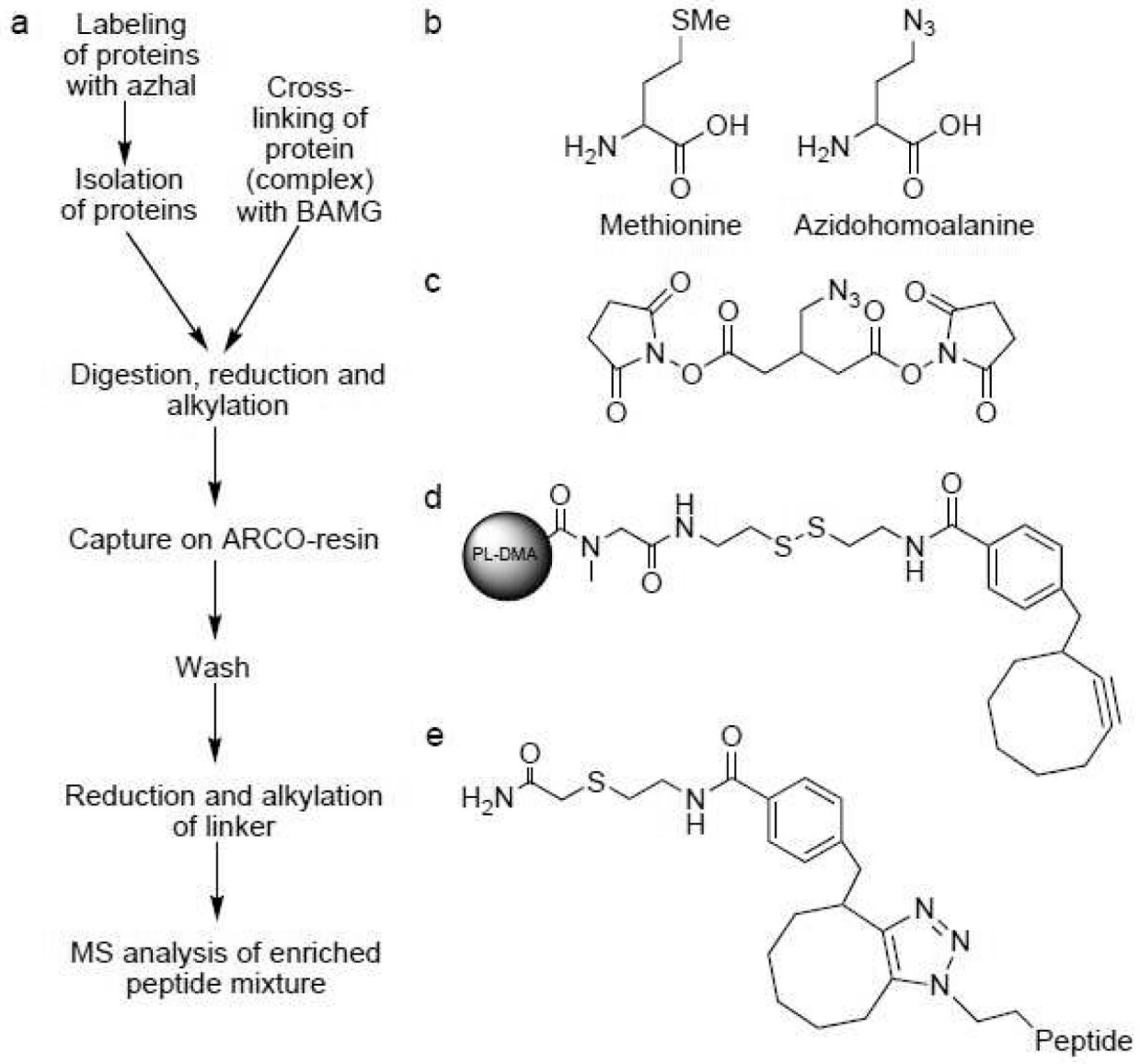

Figure 1. Overview of protocol and reagents for enrichment of azide-containing peptides

To label newly synthesized proteins, methionine auxotrophic *E. coli* K12 cells are grown on a medium where methionine has been replaced by azidohomoalanine, incorporating this non-canonical amino acid into proteins. To obtain spatial distance information, proteins are cross-linked with bis(succinimidyl)-3-azidomethyl glutarate (BAMG). After isolation or cross-linking, proteins are digested and the obtained peptide mixture is incubated with the ARCO-resin. After cleavage from the resin, enriched peptides are analysed by LC-MS/MS. **a**, Workflow of the enrichment method. **b**, Azidohomoalanine (azhal) is a methionine analogue, containing an azide group instead of a thiomethoxy group. **c**, Structure of the azide-containing cross-linker BAMG. **d**, ARCO-resin, consists of a poly-dimethylacrylamide solid support, a disulfide as cleavable linker and a cyclooctyne as reactive group towards azides. Via the strain-promoted (3+2) cycloaddition azide-containing peptides are captured on the resin. **e**, Product after enrichment, peptides are modified with the cyclooctyne and linker, adding 358.2 Da in mass.

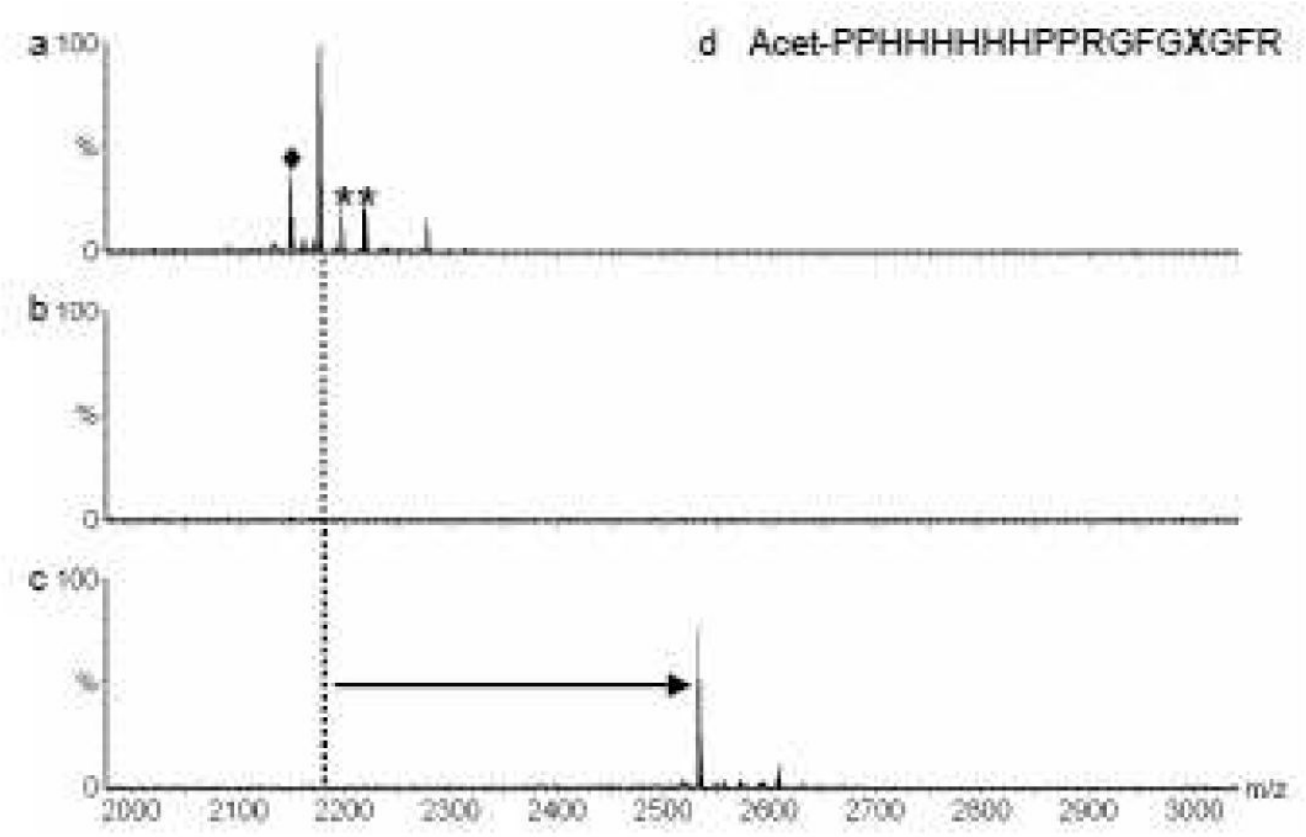

Figure 2. Enrichment of a single model peptide shows the selective reaction and release of the peptide with the ARCO-resin

A model peptide containing a single azhal residue was incubated for 24 hours at 40 °C with the ARCO-resin. **a**, Peptide before addition to the ARCO-resin. **b**, Flow through **(ft)** after 24h incubation at 40 °C **c**, Enriched peptide (**ep)**. **d**, Sequence of model peptide used to test and optimize the enrichment method. **X**: azhal. N-terminus is acetylated. ◆: loss of N_2 and uptake of 2 H. *: salt adducts.

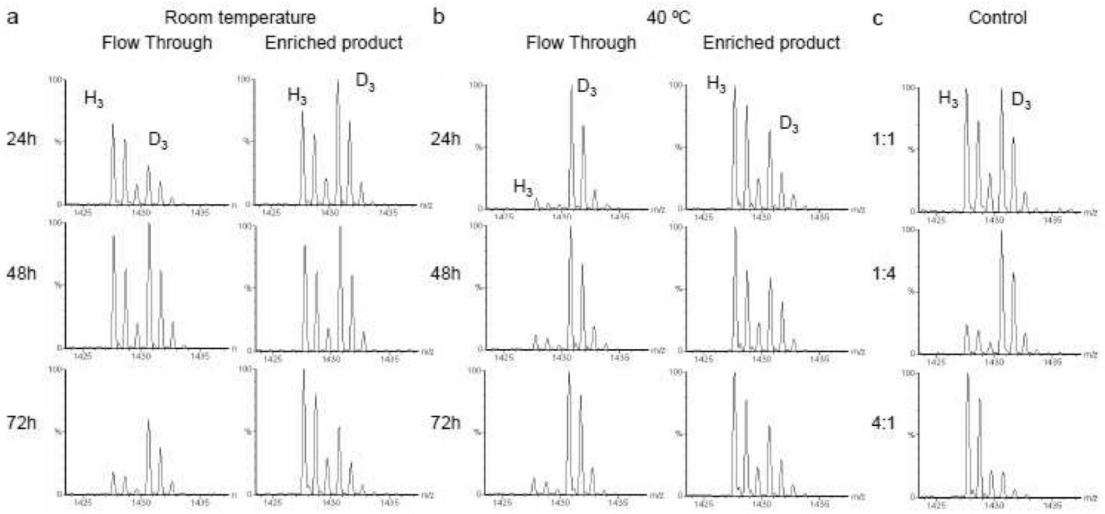

Figure 3. Quantification of enrichment of a model peptide at room temperature and 40 °C for different incubation times

MALDI-TOF spectra of quantification experiments. **a**, Enrichment experiment at room temperature. **b**, Enrichment experiment at 40 °C. **c**, Control. H_3: Ac(H_3) Pan016, flow through/ enriched peptide. D_3: Ac(D_3)-Pan016, reference peptide.

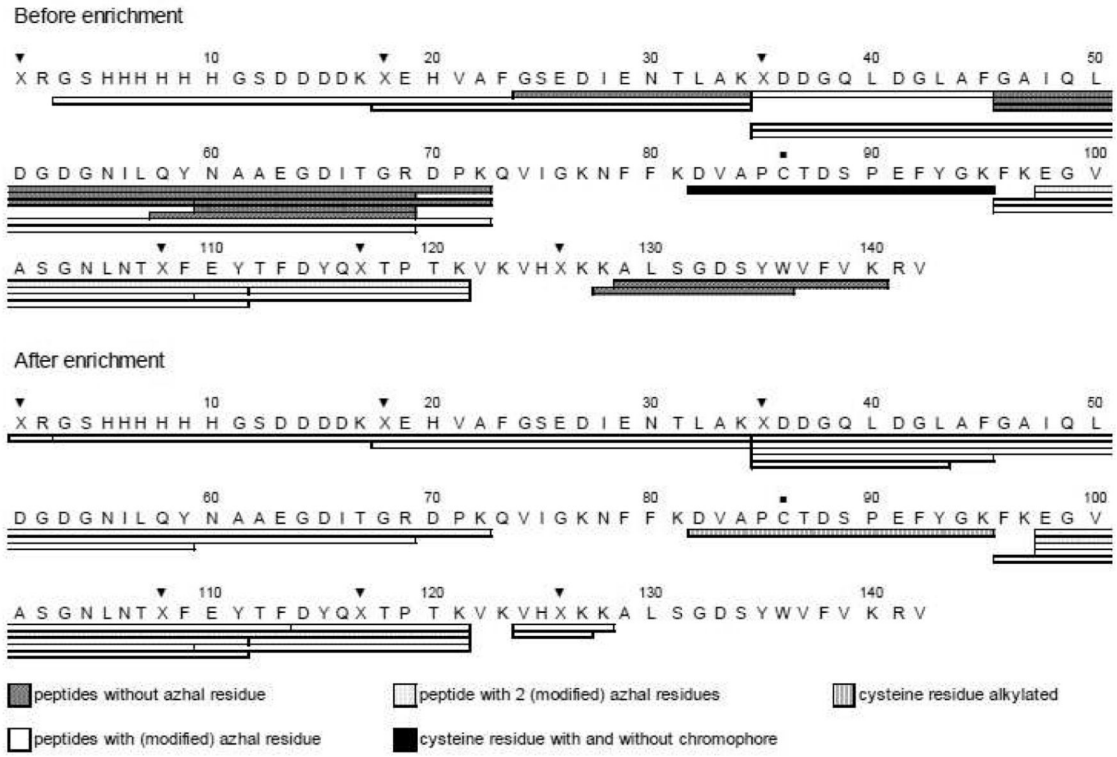

Figure 4. Sequence coverage of labeled Photoactive Yellow Protein (AzPYP) before and after enrichment

In this figure the peptides identified by both MALDI-TOF and LC-Q-TOF-MS/MS are depicted in the sequence of PYP, a protein consisting of 142 amino acids. It contains six methionine residues, which have been replaced by azhal (**X**, marked in figure by ▼). The cysteine residue has been marked by ■.

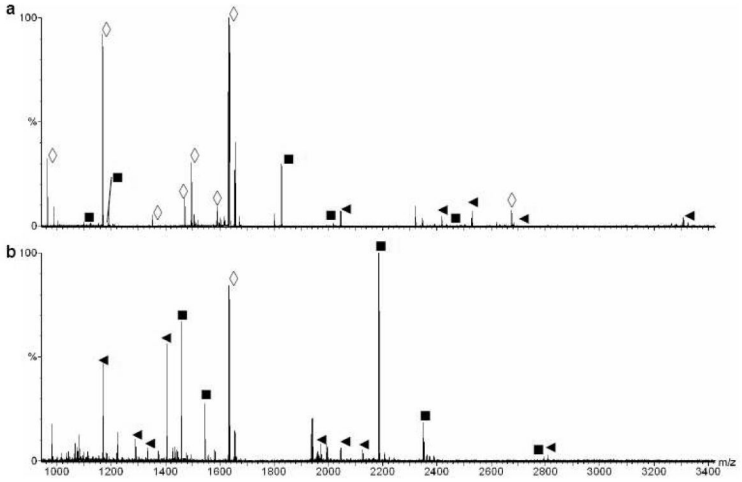

Figure 5. Enrichment of cross-linked peptides from cytochrome *c* cross-linked with BAMG

MALDI-TOF spectra of the tryptic digest of cytochrome *c* cross-linked with BAMG **a**, before and **b**, after enrichment. ◇: Unmodified peptide. ◀: Cross-link containing peptide. ■: Cross-link containing peptide observed before and after enrichment.

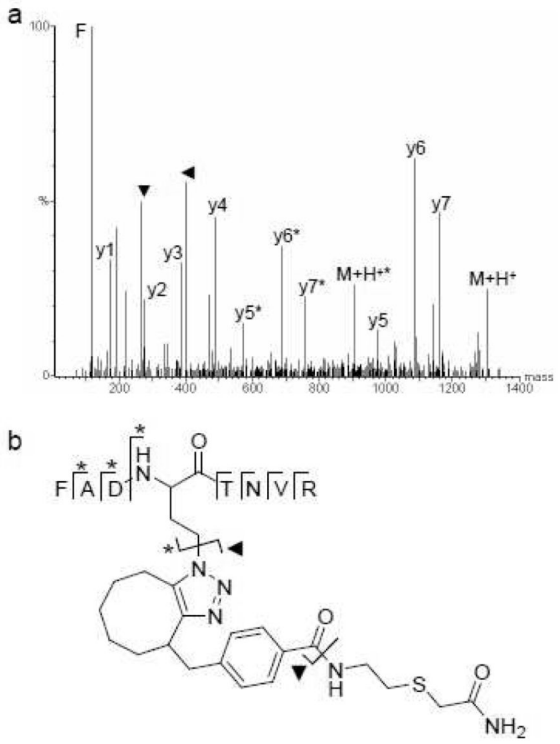

Figure 6. Fragmentation spectrum shows the fragmentation behavior of modified peptide after sequestration

a, Fragmentation spectrum of a peptide belonging to 5, 10-methylenentetrahydrofolate reductase enriched from a labeled *E.coli* proteome is shown. FADX(mod)TNVR at m/z 1306.44. **b**, Overview of observed and assigned fragments in peptide sequence. y*: loss of the cyclooctyne with the linker. ◀: cyclooctyne with linker with m/z 402.2. ▼: further fragmentation of the cyclooctyne gives rise to a benzoylium ion with 268.1 Da.

Table 1

Number of unique (methionine- and azhal-containing) peptides and proteins identified before and after enrichment of an azhal-labeled *E. coli* proteome

	Number of identified unique peptides[a]					Number of identified proteins				
	Total	Methionine & Ahzal		Azhal		Total	Methionine & Ahzal		Azhal	
			%		%			%		%
Before	809	195	24.1	56	6.9	290	109	52.2	43	14.8
After	145	n.a.	n.a.	127	87.6	98	n.a.	n.a.	89	90.8

[a] Methionine- and azhal-containing peptides with the same primary sequence are counted separately