

TECHNICAL BRIEF

Protein carbamylation: In vivo modification or in vitro artefact?

Laxmikanth Kollipara and René P. Zahedi

Leibniz-Institut für Analytische Wissenschaften – ISAS – e.V., Dortmund, Germany

Carbamylation (carbamoylation) of lysine residues and protein *N*-termini is a nonenzymatic PTM that has been related to protein ageing. In contrast to other PTM, such as phosphorylation, carbamylation can be artificially introduced during sample preparation with urea, thus affecting studies directed toward in vivo carbamylation. In aqueous solution, urea—commonly used for denaturing proteins—is in equilibrium with ammonium and isocyanate. Under alkaline conditions, the latter can react with primary amines of free *N*-termini and ϵ -amine groups of lysines to form carbamyl derivatives. Despite being a relatively slow process, which is accelerated at elevated temperatures, prolonged incubation of protein/peptide samples in urea buffers can induce undesired carbamylation, hampering not only the proteolytic digestion with trypsin and peptide identification by MS, but also interfering with stable isotope-labeling techniques such as iTRAQ, tandem mass tags, and isotope-coded protein labeling. Here, we evaluated the extent of urea-induced carbamylation under commonly used sample preparation conditions. From our results, we can deduce that carbamylation occurs in all cases involving urea, however with varying degree: e.g. carbamidomethylation in the presence of 8.0 M urea induced carbamylation of 17% of *N*-termini and 4% of Lys residues. Additionally, researching a recently published large-scale dataset revealed a high degree of urea-induced carbamylation in current proteomic samples.

Received: October 3, 2012
Revised: October 31, 2012
Accepted: November 14, 2012

Keywords:

Carbamylation / Digest / Isocyanate / PTM / Technology / Urea



Additional supporting information may be found in the online version of this article at the publisher's web-site

In spite of being a rare and not wellstudied PTM, carbamylation of proteins has been associated with severe renal and cardiovascular disorders [1–3]. In the past, detection of protein carbamylation by MS has been reported for the high-potential iron–sulfur protein isozyme I in *Ectothiorhodospira mobilis* (*N*-terminal carbamylation) as well as human lens alphaB-crystallin (Lys-carbamylation) [4, 5]. One of the major issues when analyzing in vivo carbamylation of proteins is the decisive risk for urea-induced carbamylation during sample

preparation. Being a potent chaotropic agent aiding in unfolding/solubilizing proteins that can easily be removed, in proteomics urea is often utilized during sample lysis, preparation, carbamidomethylation, and digestion. However, in aqueous solutions urea dissociates upon heating and over time, respectively. One of its degradation products is isocyanate. Depending on the pH of the solution (between 7 and 9), it covalently reacts mainly with protein *N*-termini and ϵ -amines of lysines [6], thus mimicking in vivo carbamylation. Indeed, Righetti and coworkers demonstrated that in 2DE, which employs high concentrations of urea, artificial carbamylation of proteins does not occur under electrophoretic conditions, yet can be a consequence of sample preparation [7, 8]. In contrast, Anderson and colleagues heated purified proteins for various lengths of time in urea buffer

Correspondence: René P. Zahedi, Signaling Proteomics, Leibniz-Institut für Analytische Wissenschaften, D-44227, Dortmund, Germany

E-mail: rene.zahedi@isas.de

Fax: +49-231-1392-4850

Abbreviations: **AmBic**, ammonium bicarbonate; **FDR**, false discovery rated; **PSM**, peptide spectrum matches

Colour Online: See the article online to view Fig. 1 in colour.

Table 1. Summary of the different conditions used to assess the degree of carbamylation during common sample processing procedures (3–9) when compared to controls (1–2)

Condition	Buffer	Incubation conditions	Adapted from
1	No treatment (0.1% TFA v/v)	–	–
2	8.0 M urea, 100 mM Tris-HCl (pH 8.5)	61°C, 15 h	–
3	8.0 M urea, 100 mM Tris-HCl (pH 8.5)	56°C, 30 min	[13]
4	2.0 M urea, 50 mM AmBic (pH 7.8)	37°C, 15 h	[14]
5	2.0 M urea, 100 mM Tris-HCl (pH 8.5)	37°C, 15 h	[16]
6	0.1 M urea, 50 mM AmBic (pH 7.8)	37°C, 15 h	
7	0.2 M Gu-HCl, 50 mM AmBic (pH 7.8)	37°C, 15 h	
8	8.0 M urea, 100 mM Tris-HCl (pH 8.5)	20°C, 18 h	[15]
9	8.0 M urea, 100 mM Tris-HCl (pH 7.9) and 2.0 M urea, 100 mM Tris-HCl (pH 7.9)	20°C, 18 h, and 20°C, 4 h	[15]

and pooled the samples to generate defined *pI* standards for 2DE [9]. Their so called “carbamylation trains” are based on the varying degree of protein carbamylation accompanied by a change in net charge, and were used for testing batches of ampholytes [10]. Angel et al. utilized carbamylation for quantitative proteomics studies by intentionally labeling peptides with light and heavy (stable isotope coded) forms of urea [11].

Here we systematically evaluated the occurrence of carbamylation under conditions often used during sample preparation and proteolytic digestion. Therefore, we generated a peptide mixture of five model proteins, namely serum albumin, beta-casein, beta-lactoglobulin, haemoglobin (all bovine), and myoglobin (horse). These proteins were separately reduced, alkylated, and subsequently digested with trypsin as previously described [12], in absence of urea. The tryptic peptides thus generated were desalted using C18 tips and eluates were pooled. The resulting peptide solution was divided into 50 pmol aliquots that were dried completely under vacuum and then incubated under different conditions, mimicking typical urea-based protocols often used in proteomics (Table 1 and supplementary information). Using a stock on the peptide instead of the protein level allowed for a better intercomparability and consequently quantitation of the extent of carbamylation, since variances induced during independent sample processing and digestion of aliquots were omitted. Moreover, this allowed for a more thorough examination of *N*-terminal carbamylation compared to the presence of only five accessible *N*-termini on the protein level. The dried peptide mixtures (50 pmol) were dissolved

in 50 μ L of freshly prepared urea buffers and subjected to different incubation conditions as summarized in Table 1. Temperature and time of incubation were selected based on common conditions used in typical top-down proteomics workflows [13–16] and each experimental setup was performed in triplicate, once with approximately 4-month-old and once with approximately 8 year-old urea.

Finally, 1 pmol of each solution was analyzed by LC-MS, using an Ultimate 3000 RSLC system coupled to an LTQ-Orbitrap XL mass spectrometer (both Thermo Scientific). Peptides were preconcentrated on a 75 μ m $\times$ 2 cm C18 trapping column for 10 min using 0.1% TFA v/v at flow rate of 20 μ L/min followed by separation on a 75 μ m $\times$ 15 cm C18 main column (both Pepmap, Thermo Scientific) with a 50-min LC gradient ranging from 4 to 40% ACN in 0.1% FA v/v at a flow rate of 300 nL/min. MS data interpretation were conducted using Proteome Discoverer 1.3 (Thermo Scientific) and Mascot 2.4 (Matrix Science). Database searches were performed in target/decoy mode against a database containing the five model protein sequences taken from Uniprot, in a yeast background taken from *saccharomyces* genome database (SGD) (in total 6723 target sequences). Trypsin was selected as enzyme and two missed cleavage sites were allowed. Carbamidomethylation of cysteine was set as fixed and carbamylation of *N*-terminus, lysine, and arginine as dynamic modifications. MS and MS/MS tolerances were set to 10 ppm and 0.5 Da, respectively, only peptide spectrum matches (PSMs) with a false discovery rate (FDR) < 1% (high confidence Peptide Validator setting) were considered.

Using our peptide mixture, we evaluated the extent of urea-induced carbamylation under following conditions: (1) control (in 0.1% TFA v/v); (2) harsh treatment (8.0 M urea, 100 mM Tris-HCl, pH 8.5; 61°C, 15 h); (3) reduction of disulfide bonds (8.0 M urea, 100 mM Tris-HCl, pH 8.5; 56°C, 30 min), different digestion conditions with trypsin at 37°C for 15 h, namely; (4) 2.0 M urea, 50 mM ammonium bicarbonate (AmBic), pH 7.8; (5) 2.0 M urea, 100 mM Tris-HCl, pH 8.5; (6) 0.1 M urea, 50 mM AmBic, pH 7.8; (7) digestion in presence of guanidine hydrochloride instead of urea (0.2 M Gu-HCl, 50 mM AmBic, pH 7.8; 37°C, 15 h); (8) Lys-C digestion overnight (8.0 M urea, 100 mM Tris-HCl, pH 8.5; 20°C, 18 h).

As expected, when treated with harsh conditions (2), nearly all identified peptides were carbamylated (99% *N*-termini, 90% Lys) (Fig. 1A). Under conditions often used for reduction [13] (3), we found nearly 13% of peptides carbamylated at their *N*-terminus whereas only 4% of lysine side chains were modified. This might be attributed to higher reactivity of primary amines toward isocyanate compared to the ϵ -amino group on lysines at basic pH [17]. In many protocols, after lysis (and sometimes initial Lys-C digestion) urea concentration is lowered to 2.0 M to keep proteins (or polypeptides) unfolded and solubilized during digest. Under this condition, we found nearly the same number of carbamylated peptides for both buffers tested i.e. (4) AmBic [14] (21% *N*-termini, 2.2% Lys) and (5) Tris-HCl [16] (23% *N*-termini, 1.2% Lys), suggesting

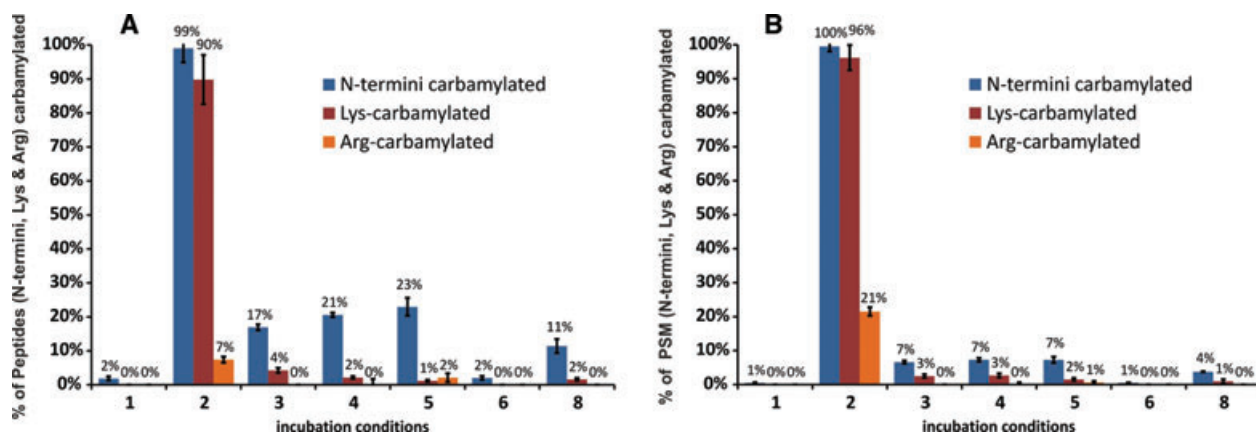


Figure 1. Share of carbamylated peptides (A) and PSM (B) identified after different incubation conditions as summarized in Table 1.

that one-fifth of *N*-termini and ~2% of Lys residues could be carbamylated during overnight incubation at 37°C in presence of 2.0 M urea. We therefore evaluated a further digestion buffer containing only 0.1 M urea (6). Here, 2% of *N*-termini and 0% of Lys residues were modified. In addition, we evaluated a buffer based on Gu-HCl as an alternative chaotropic agent; as expected carbamylation could not be detected due to the absence of urea, though, primary amines are potentially prone to guanidinylation in presence of Gu-HCl. However, overnight incubation at 37°C with Gu-HCl (7) did not result in detectable guanidinylation of *N*-termini and Lys side chains. Finally, (8) we evaluated conditions used for Lys-C digestion [15], incubating our peptide mixture in 8.0 M urea for 18 h at 20°C, thus detecting 11% of *N*-termini and 2% of Lys residues carbamylated.

In principle when prepared freshly, we could detect no differences between the 4-month- and the 8-year-old urea, since its degradation only occurs in aqueous solution.

Because of the relatively high share of urea-induced carbamylation under Lys-C conditions, we evaluated the conditions often used during two-step digestion procedures (ix) involving a primary digestion in 8.0 M urea (such as 8), followed by 4-fold dilution and digestion with trypsin at 20°C for 4 h [15]. Therefore, we prepared lysates from human fibroblasts with 0.1% w/v Rapigest SF (Waters) and conducted a two-step digestion with Lys-C and trypsin as detailed in the Supplementary Information and Table 1, once in the presence of urea (urea-based), and once in 0.1% (w/v) Rapigest (urea-free). Upon database search of the LC-MS/MS data, we found nearly the same number of PSMs in both cases: 12 982 in urea-based and 12 321 in urea-free samples (Supplementary Information Table 1).

In accordance with Wiśniewski et al. [15], we could detect carbamylation on less than 0.3% of the identified peptides after 22 h of incubation, which is indeed below the applied false discovery rate of ~1% and thus not significant. Moreover, we could detect no differences between the urea-based and urea-free samples. While this is in contrast to the results

above, it can be most likely ascribed to two major reasons. First, undersampling generally encountered during the mass spectrometric analysis of complex proteomes [18] will reduce the number of carbamylated peptides, since for each of the 100 000s of peptides in the complex digest, only a few percent of the present copies are actually carbamylated. Second, *N*-termini have a higher reactivity for carbamylation than Lys residues—this can cause an overestimation of carbamylated peptides in the already predigested reference sample in comparison to the fibroblast digest, in which new *N*-termini still have to be generated during digestion.

Additionally, we evaluated the occurrence of urea-induced carbamylation by researching a recently published dataset that identified several 1000 proteins from *Escherichia coli*, prepared using urea, in a single 6 h LC-MS run [19]. Using the same search parameters (without Arg carbamylation since its occurrence can be neglected) as above and the Uniprot database with taxonomy set to *E. coli* (July 30th 2012, 22 925 target sequences), we could detect that at 1% FDR, 27.4% of the identified peptides (2597 out of 9481), and 25.7% of the PSMs (4103 out of 15 948) were indeed carbamylated, preferentially on the *N*-termini (~22% compared to ~7% of Lys residues).

Summarizing, our data clearly demonstrate that when analyzing *in vivo* carbamylation the usage of urea has to be completely avoided in all sample processing steps, with *N*-termini being the most susceptible, followed by Lys residues, whereas carbamylation of Arg residues does not occur under commonly used sample processing conditions. Furthermore it is highly recommendable to avoid conditions that trigger protein carbamylation in proteomic studies that (i) are based on chemical labeling of amines such as iTRAQ [20] and ICPL [21], or (ii) target protein *N*-termini such as TAILS [22] and *N*-terminal COFRADIC [23]. It should be kept in mind, that some carbamylated *N*-termini are almost isobaric to certain sequence tags, such as carbamyl-Ala and Gly-Gly (both 114.04293 Da), potentially giving rise to false-positive identifications. However, under ion-trap CID conditions,

many—but not all—carbamylated peptides show a clear neutral loss of isocyanic acid (43.0058 Da) that can be potentially used as a marker signal.

We thank Dr. Andreas Roos (Universitätsklinikum Aachen, Germany) for providing fibroblast cell pellets. The financial support by the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen, and by the German Research Foundation (DFG ZA 639/1-1) is gratefully acknowledged.

The authors have declared no conflict of interest.

References

- Wang, Z., Nicholls, S. J., Rodriguez, E. R., Kummu, O. et al., Protein carbamylation links inflammation, smoking, uremia and atherogenesis. *Nat. Med.* 2007, 13, 1176–1184.
- Ok, E., Basnakian, A. G., Apostolov, E. O., Barri, Y. M., Shah, S. V., Carbamylated low-density lipoprotein induces death of endothelial cells: a link to atherosclerosis in patients with kidney disease. *Kidney Int.* 2005, 68, 173–178.
- Apostolov, E. O., Ray, D., Savenka, A. V., Shah, S. V., Basnakian, A. G., Chronic uremia stimulates LDL carbamylation and atherosclerosis. *J. Am. Soc. Nephrol.* 2010, 21, 1852–1857.
- Van Driessche, G., Vandenbergh, I., Jacquemotte, F., Devreese, B., Van Beeumen, J. J., Mass spectrometric identification of in vivo carbamylation of the amino terminus of *Ectothiorhodospira mobilis* high-potential iron-sulfur protein, isozyme 1. *J. Mass Spectrom.* 2002, 37, 858–866.
- Lapko, V. N., Smith, D. L., Smith, J. B., In vivo carbamylation and acetylation of water-soluble human lens alphaB-crystallin lysine 92. *Protein Sci.* 2001, 10, 1130–1136.
- Stark, G. R., Stein, W. H., Moore, S., Reactions of the Cyanate Present in Aqueous Urea with Amino Acids and Proteins. *J. Biol. Chem.* 1960, 235, 3177–3181.
- Righetti, P. G., Real and imaginary artefacts in proteome analysis via two-dimensional maps. *J. Chromatogr. B* 2006, 841, 14–22.
- McCarthy, J., Hopwood, F., Oxley, D., Laver, M. et al., Carbamylation of proteins in 2-D f-myth or reality? *J. Proteome Res.* 2003, 2, 239–242.
- Anderson, N. L., Hickman, B. J., Analytical techniques for cell fractions XXIV. Isoelectric point standards for two-dimensional electrophoresis. *Anal. Biochem.* 1979, 93, 312–320.
- Tollaksen, S. L., Edwards, J. J., Anderson, N. G., The use of carbamylated charge standards for testing batches of ampholytes used in two-dimensional electrophoresis. *Electrophoresis* 1981, 2, 155–160.
- Angel, P. M., Orlando, R., Quantitative carbamylation as a stable isotopic labeling method for comparative proteomics. *Rapid Commun. Mass Spectrom.* 2007, 21, 1623–1634.
- Burkhardt, J. M., Schumbrutski, C., Wortelkamp, S., Sickmann, A., Zahedi, R. P., Systematic and quantitative comparison of digest efficiency and specificity reveals the impact of trypsin quality on MS-based proteomics. *J. Proteomics* 2012, 75, 1454–1462.
- Choksawangkarn, W., Edwards, N., Wang, Y., Gutierrez, P., Fenselau, C., Comparative study of workflows optimized for in-gel, in-solution, and on-filter proteolysis in the analysis of plasma membrane proteins. *J. Proteome Res.* 2012, 11, 3030–3034.
- Helbig, A. O., Gauci, S., Raijmakers, R., van Breukelen, B. et al., Profiling of N-acetylated protein termini provides in-depth insights into the N-terminal nature of the proteome. *Mol. Cell. Proteomics* 2010, 9, 928–939.
- Wisniewski, J. R., Zougman, A., Nagaraj, N., Mann, M., Universal sample preparation method for proteome analysis. *Nat. Methods* 2009, 6, 359–362.
- Yang, B., Wu, Y.-J., Zhu, M., Fan, S.-B. et al., Identification of cross-linked peptides from complex samples. *Nat Methods* 2012, 9, 904–906.
- Stark, G. R., Smyth, D. G., The use of cyanate for the determination of NH₂-terminal Residues in proteins. *J. Biol. Chem.* 1963, 238, 214–226.
- Desiere, F., Deutsch, E., Nesvizhskii, A., Mallick, P. et al., Integration with the human genome of peptide sequences obtained by high-throughput mass spectrometry. *Genome Biol.* 2005, 6, R9.
- Cristobal, A., Hennrich, M. L., Giansanti, P., Goerdal, S. S. et al., In-house construction of a UHPLC system enabling the identification of over 4000 protein groups in a single analysis. *Analyst* 2012, 137, 3541–3548.
- Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B. et al., Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 2004, 3, 1154–1169.
- Schmidt, A., Kellermann, J., Lottspeich, F., A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 2005, 5, 4–15.
- Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A. et al., Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat. Biotechnol.* 2010, 28, 281–288.
- Gevaert, K., Goethals, M., Martens, L., Van Damme, J. et al., Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat Biotechnol.* 2003, 21, 566–569.

Supplementary information

Materials

Following materials were purchased from Sigma-Aldrich, Steinheim, Germany: bovine serum albumin, bovine beta lactoglobulin, bovine beta casein, bovine haemoglobin, horse myoglobin, ammonium hydrogen carbonate (NH_4HCO_3), guanidine hydrochloride (GuHCl), iodoacetamide (IAA) type B (4 month old) urea, anhydrous magnesium chloride (MgCl_2) and trypsin. Type A urea (8 year old) and Tris were obtained from Applichem Biochemica, Darmstadt, Germany. Dithiothreitol (DTT) and Lys-C were bought from Roche Diagnostics, Mannheim, Germany. Lyophilized surfactant RapiGestTM SF was supplied by Waters Corporation (Milford, MA, USA). Benzonase[®] Nuclease was purchased from Novagen. Bicinchoninic acid assay (BCA) kit was acquired from Pierce Thermo Fisher Scientific, Schwerte, Germany. All chemicals for ultra-pure HPLC solvents such as formic acid (FA), trifluoroacetic acid (TFA) and acetonitrile (ACN) were obtained from Biosolve, Valkenswaard, the Netherlands.

Methods

Preparation of stock peptide mixture

Stock protein solutions were prepared in ultra-pure water at a concentration of 1 mg/mL. 100 µg of each protein solution (100 µL) was diluted 10-fold with 50 mM AmBic (pH 7.8). All five proteins were processed independently. Disulfide bonds were reduced by addition of 10 mM DTT at 56°C for 30 min, followed by alkylation of free thiol groups with 30 mM IAA at room temperature (RT) in the dark for 30 min. Trypsin was added in a 1:50 (w/w) enzyme to substrate ratio and proteins were digested at 37°C overnight. Generated tryptic peptides (800 pmol) of each protein were acidified with TFA to a final concentration of 0.2%. SPE was conducted with 100 µL C18 tips (OMIX, Varian), according to the manufacturer's protocol. Peptides were eluted with 100 µL of 60% ACN/0.1% TFA (v/v). To prepare the stock peptide mixture, eluates (100 µL) of each digested protein (800 pmol) were pooled. This combined peptide mixture was then divided into 50 pmol aliquots, which were completely dried under vacuum.

Processing of urea-treated samples for LC-MS/MS analysis

After completion of each incubation condition, the urea-treated samples were diluted with ultra-pure water to achieve a final concentration of 0.1 pmol/ μ L. The dilutions of control samples (i.e., untreated) were made in 0.1% TFA (v/v) to attain similar final concentrations as above. For the LC-MS/MS analysis, 10 μ L (1 pmol) of each sample was used.

Two-step digest of fibroblast cells

Lyophilized RapiGest SF powder was reconstituted in 50 mM AmBic (pH 7.8) to get a final concentration of 0.1% (w/v). Approximately, 250 μ g of fibroblast cell pellets were lysed with 100 μ L of RapiGest solution. Subsequently, 1 μ L of benzonase (25 U/ μ L) and 2 mM $MgCl_2$ were added and incubated at 37°C for 30 min. Samples were centrifuged at 4°C, 16,000 *g* for 15 min. Protein concentration of the supernatant was determined by BCA assay according to the manufacturer's protocol. 20 μ g protein aliquots were taken for urea-based and urea-free processing, each. In case of urea-based, samples were diluted 10-fold with 8.0 M urea/100 mM Tris-HCl (pH 7.9), whereas urea-free samples were diluted 10-fold with RapiGest SF (0.1% w/v). Reduction of cystines was conducted with 5 mM DTT (56°C, 30 min) followed by alkylation with 15 mM IAA in the dark (RT, 30 min). Lys-C was added in 1:100 (w/w) enzyme to substrate ratio and incubated at 20°C for 18 h. Afterwards, urea concentration was lowered from 8.0 M to 2.0 M with 100 mM Tris-HCl (pH 7.9) in both sample sets. Proteolysis was continued with trypsin (1:100 w/w) at 20°C for 4 h. RapiGest was removed according to the manufacturer's instructions. In brief, protein digests were acidified to a final concentration of 0.4% TFA, followed by incubation at 37°C for 45 min. Acidified samples were then centrifuged at 13,000 *g* for 10 min and finally, the clear supernatant was collected and used for LC-MS/MS analysis.

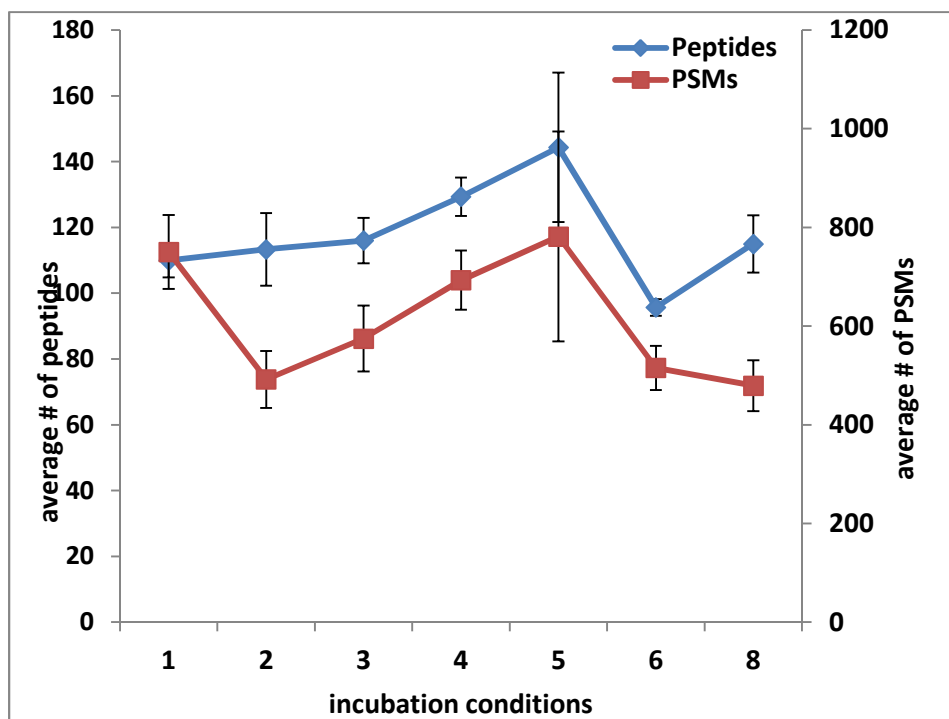
Mass Spectrometric (MS) analysis of peptide mixtures

MS data were acquired on an LTQ-Orbitrap XL mass spectrometer (Thermo Scientific). MS survey scans were acquired in the Orbitrap from *m/z* 300 to 2000 at a resolution of 60,000 using the polysiloxane ion at *m/z* 371.101236 as lock mass [1]. The five most intense signals were subjected to collision induced dissociation (CID) in the ion trap taking into account a dynamic exclusion of 10 s. CID spectra were acquired with a normalized CE of 35%, a default charge state of 2 and an activation time of 30 ms. AGC target values were set to 10^6 for Orbitrap MS and 10^4 for ion trap MSⁿ scans.

LC-MS/MS analysis and Data interpretation of two-step digested fibroblast samples

Peptides were separated on an Ultimate 3000 system (Thermo Scientific) with self-packed columns filled with reverse phase Kinetix C18 material (2.6 μm , 100 $\text{\AA}$). The concentration of tryptic peptides was estimated based on the amount of protein used prior to enzymatic digestion (i.e., 20 μg). Peptide solutions (approximately 1 μg) were preconcentrated on the trapping column 100 μm x 2 cm for 20 min using 0.1% TFA (v/v) at a flow rate of 5 $\mu\text{L}/\text{min}$ followed by separation on the main column 75 μm x 30 cm with a 120 min LC gradient from 3-38% ACN in 0.1% FA (v/v) at a flow rate of 230 nL/min. The LC was coupled to an LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific) as described above. The twenty most intense signals were subjected to collision induced dissociation (CID) in the ion trap as described above.

MS data interpretation was conducted using Proteome Discoverer 1.3 (Thermo Scientific) with two search algorithms Mascot 2.4 (Matrix Science) and SEQUEST, using a human UniProt database (downloaded on 30th of July 2012, containing 20,424 target entries). Search settings were: trypsin as enzyme, two missed cleavage sites allowed, carbamidomethylation of cysteine as fixed and carbamylation of N-terminus, lysine and arginine as dynamic modifications. MS and MS/MS tolerances were set to 10 ppm and 0.5 Da respectively, only PSMs with an FDR < 1% (high confidence PeptideValidator setting) were considered. Moreover, identified peptides were filtered for (a) maximum rank 1, (b) mass deviation of ≤ 4 ppm and (c) proteins with at least one unique peptide.



Supplementary Figure 1: Number of peptides and peptide spectrum matches (PSMs) identified under different experimental conditions, average values and standard deviations from triplicate measurements.

Supplementary table 1: Summary of two-step digest condition (7) performed with Lys-C and trypsin using fibroblast lysates.

condition	# of proteins identified with ≥ 1 unique peptide	# of peptides identified	# of PSMs identified	% of carbamylated peptides (N-termini, Lys and Arg)
urea-based	1578 $\pm$ 17	6465 $\pm$ 330	12982 $\pm$ 720	0.3% $\pm$ 0.1%
urea-free (Rapigest)	1547 $\pm$ 13	5806 $\pm$ 252	12321 $\pm$ 460	0.2% $\pm$ 0.1%

[1] Olsen, J. V., de Godoy, L. M. F., Li, G., Macek, B., *et al.*, Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap. *Molecular & Cellular Proteomics* 2005, 4, 2010-2021.

