Chemical Research in Toxicology®

Article

Identification of Protein Thiazolidination as a Novel Molecular Signature for Oxidative Stress and Formaldehyde Exposure

Jingjing Liu, K. K. Jason Chan, and Wan Chan

Chem. Res. Toxicol., **Just Accepted Manuscript** • DOI: 10.1021/acs.chemrestox.6b00271 • Publication Date (Web): 02 Nov 2016

Just Accepted

ACS Publications

Chemical Research in Toxicology is published by the American Chemical Society.
1155 Sixteenth Street N.W., Washington, DC 20036

Identification of Protein Thiazolidination as a Novel Molecular Signature for Oxidative Stress and Formaldehyde Exposure

Jingjing Liu,[†] K. K. Jason Chan,[‡] and Wan Chan*[,†,‡]

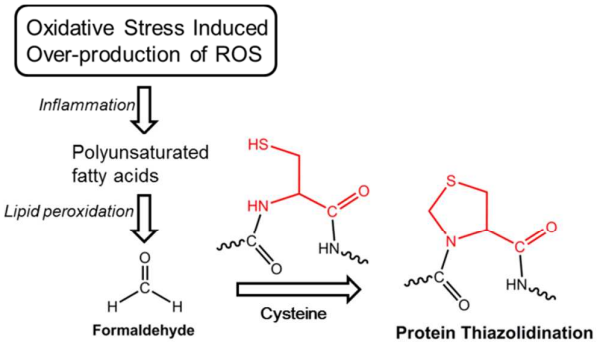
[†] Environmental Science Programs and [‡] Department of Chemistry, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

* Correspondence author

Address: Room 4520, Department of Chemistry, Academic Building, Clear Water Bay, Kowloon, Hong Kong

E-mail: chanwan@ust.hk, Phone: +852 2358-7370, Fax: +852-2358-1594.

Table of Contents Graphic



ABSTRACT

Chemical modifications of proteins have been well-documented to play important roles in normal cell physiology such as cell signaling and protein functions. They have also been demonstrated to be one of the milestones in the pathophysiology of many human diseases such as cancer, age-related pathology, and neurodegenerative disorders. Here we report the initial identification of a novel protein modification, cysteine thiazolidination, through reaction with endogenous and exogenous formaldehyde with cysteine residues in proteins. Using an isotope-dilution liquid chromatography-tandem mass spectrometric (LC−MS^3) method, we initiated the study by quantitating thioproline in formaldehyde-treated *Escherichia coli* (*E. coli*) protein. The study was then extended to quantitate thioproline in protein obtained from formaldehyde- and oxidant-exposed *E. coli*. Furthermore, N^6-formyllysine, a well-defined formylation product between formaldehyde and lysine, was exploited in a comparative study to evaluate the relative reactivity and amount of cysteine thiazolidination in the reaction of formaldehyde with proteins. It is anticipated that cysteine thiazolidination may serve as a novel biomarker for oxidative stress and formaldehyde exposure.

Keywords: Protein modification, thiazolidination, thioproline, oxidative stress, LC-MS/MS

INTRODUCTION

Proteins are molecular vehicles that drive a myriad of cellular processes, and the diverse biological functions they exhibit are attributed to their specific substrate-fitting shapes as well as unique arrangements of functional groups. Many cellular proteins contain some form of non-canonical post-translational modifications (PTMs) which can carry physiological significance and their formations are normally mediated through enzymes. However, the constant presence of reactive chemical species inside the cell, arising both endogenously and exogenously, can also lead to the proteins being modified through unintended chemical reactions. These PTMs could alter protein structures and their bioactivities, which may lead to adverse effects on cell physiology.[1,2] PTMs commonly involve electrophilic species, which form covalent adducts with nucleophilic amino acid residues through alkylation, carbonylation, glycation, or phosphorylation. Correlations between these modifications and their physiological significance in cell signaling,[3,4] functions,[5-8] and pathophysiology[9,10] are well-documented.

Among the many reactive electrophiles known to modify proteins, formaldehyde that is ubiquitous in both indoor and outdoor air, present in smoked and barbequed foods and even generated endogenouly from processes such as lipid peroxidation[11,

[12] has attracted a lot of attention as a PTM agent, especially since it is also a known human carcinogen.[13] Formaldehyde is known to condense swiftly with lysine and arginine residues in peptides to produce *N*-formylated products.[14-17] For example, the formation of N^6-formyllysine has been used as a biomarker to assess the extent of exposure to oxidative stress (Figure 1).[14] The addition of formyl groups to peptide residues has also been implicated in the misfolding, loss of function, and hydrolysis of proteins.[18-20] It was also discovered that formylation could result in cytotoxic protein aggregation products.[21, 22]

Recently, we reported on the dose-dependent formation of thiazolidine-4-carboxylic acid (thioproline) when formaldehyde condenses with cysteine, present as free amino acid in the extra-cellular fluid of *Escherichia coli* (*E. coli*) exposed to oxidative stress.[23] Since cysteine is an essential building block of proteins, the aim of this study is to test our hypothesis that formaldehyde would also react with cysteine residues within a polypeptide to form stable thiazolidine adducts.

Using a liquid chromatography-coupled tandem mass spectrometric ($LC-MS^3$) method with high sensitivity and selectivity, we demonstrated in this study for the first time that cysteine residues in proteins indeed react with formaldehyde *in vitro* through a novel thiazolidination process to form a cyclic thioproline residue in a

dose-dependent manner (Figure 1). This newly identified protein modification was also observed to occur in bacterial proteins isolated from *E. coli* whole cells that were treated with formaldehyde, along with other oxidative stress generators (Fe^{2+}-EDTA, hydrogen peroxide (H_2O_2), and sodium hypochlorite (NaOCl)).To establish the relative frequency of this newly identified cysteine thiazolidination event against well-established formylation events during the reaction between formaldehyde and proteins, we also performed a comparative study in parallel on the formation of N^6-formyllysine.[24,25]

EXPERIMENTAL SECTION

Chemicals and Reagents. All chemicals and reagents of the highest purity available were used without further purification unless otherwise stated. L-cysteine, N^6-formyllysine, L-thioproline, fluorenylmethyloxycarbonyl chloride (Fmoc-Cl), formaldehyde, formaldehyde-d_2, and *Streptomyces griseus* protease were obtained from Sigma (St. Louis, MO). Isotope labeled L-thioproline and L-thioproline-d_2, were synthesized by reacting formaldehyde-d_2 with L-cysteine and purified by HPLC as described previously.[23] Fmoc-L-thioproline and a thioproline-containing peptide (Thioproline-Arg-Glu-Gln-Arg-Leu-Gly-Arg-Gln-Trp-Ala-Val-Gly-His-Met) were acquired from GL Biochem (Shanghai, China). Protein for *in vitro* experiments was isolated in-house from cultured *E. coli* using the method described below. LC−MS grade acetonitrile and methanol were purchased from J.T. Baker (Philisburg, NJ). Deionized water was further purified by a Milli-Q Ultrapure Water System (Billerica, MA) and was used in all the experiments.

Protein Isolation. Whole protein for *in vitro* experiments and in toxicants exposed *E. coli* was isolated using the method described previously.[26] In brief, *E.coli* cells after being collected by centrifugation were washed three times with potassium phosphate buffer (100 mM, pH 7.4) and then resupended in 1M Tris buffer containing 20% glycerol (v/v). After ultrasounication and centrifugation (18,000 g, 50 min), the

supernatant containing cellular protein was collected and added with saturated ammonium sulfate to precipitate the proteins. Cellular protein was then collected by centrifugation (18,000 g, 30 min).

Instrumental Analyses. HPLC analysis and purification were performed on an 1100 series HPLC system equipped with a diode array detector (Agilent, Palo Alto, CA). High-resolution mass spectrometry (MS) and product ion scan (MS/MS) experiments were performed on a Xevo G2 QTOF mass spectrometer with a standard electrospray ionization interface (Waters, Milford, MA). LC–MS/MS and LC–MS^3 analyses were performed on an Agilent 1100 HPLC-coupled with an API 4000 QTRAP tandem mass spectrometer with a TurboV ion source (AB Sciex, Foster City, CA). UV spectrometric-based quantitation of protein was performed on a Varian UV–vis absorption spectrophotometer (Cary 50, Walnut Creek, CA).

Optimization of Reaction Conditions for the Derivatization of Thioproline with Fmoc-Cl. To determine the amount of Fmoc-Cl that is needed for the derivatization reaction of thioproline, we varied the reagent-to-thioproline molar ratios by adding different amounts of Fmoc-Cl to thioproline-fortified (1 nmole) hydrolysates of purified protein (100 μg) (50:1, 100:1, 200:1, 500:1, 1000:1, 2000:1, and 5000:1, Fmoc-Cl:thioproline) at sodium borate buffer (150 mM, pH 8.0).[27] The mixtures were left to stand in room temperature for 30 min with occasional

vortex-mixing before being analyzed using LC-MS^3 as described below to evaluate the extent of conversion of thioproline to **1**.

Reaction of Formaldehyde with Purified Protein. Solutions of purified *E. coli* protein (100 μL, 1000 μg/mL) in potassium phosphate buffer (100 mM, pH 7.4) were incubated with formaldehyde with final concentrations ranging from 0.6 to 2.5 mM at 37 ^{o}C for 1 hr. The formaldehyde-treated protein samples (100 μg) were then hydrolyzed with protease (10 μg, 37 ^{o}C for >16 hr), as described elsewhere.[28] Following the enzymatic digestion of the protein, the enzymes were removed by ultrafiltration using a 10,000 Da molecular weight cut-off Nanosep Centrifugal Device (Pall Life Sciences). The filtrates containing both naturally occurring and formaldehyde-modified amino acids (e.g. thioproline and N^6-formyllysine), were derivatized with Fmoc-Cl (Figure 2), with an added deuterium-labeled internal standard, before they were analyzed by LC-MS^3 as described below.

Quantitation of Thioproline in Formaldehyde-Treated and Oxidized *E. coli* Protein Extract. *E.coli* (DH5α, ATCC) cultured to mid-log phase were harvested by centrifugation. The cell pellet was washed twice with PBS, and then re-suspended in PBS. The resuspended cells were treated with formaldehyde (5-17 mM final concentration) and the following oxidants: Fe^{2+}-EDTA (0.5-2.5 mM), NaOCl (0.1-1.0 μM), and H_2O_2 (1.25 -8.5 μM). After 1 hr of exposure, the cells were collected by

centrifugation, the cellular protein isolated as described above, quantitated by UV spectrometry (280 nm), and digested with protease for LC–MS^3 analysis.

LC–MS^3 Analysis. A 10 μL aliquot of the Fmoc-derivatized protein hydrolysate was loaded into a GraceSmart C18 column (150 × 2.1 mm, 5 μm), and eluted at 0.3 mL/min with a gradient of acetonitrile (A) in 0.1% formic acid (v/v) in water (B) to separate the Fmoc-tagged amino acids. The gradient elution program started from 20% B (v/v), programmed linearly to 100% B in 20 min, and held for another 5 min before reconditioning to 20% B for 5 min.

After diverting the first 8 min of the LC eluate to waste, the effluent containing **1** was directed to a QTRAP mass spectrometer for MS^3 analysis. The electrospray ionization source was operated in positive ion mode with the following optimized parameters for voltages and source gas: ion spray voltage, 5500 V, collision energy, 15V, declustering potential, 60 V, and excitation energy, 150 V. The ion source gas I, gas II, curtain gas, collision gas, and temperature of gas II were set to 30, 40, 25, 5, and 400 °C, respectively. The mass spectrometer was set to monitor the fragments of the native (*m/z* 356→134→88) and isotope labeled (*m/z* 358→136→90) thioproline. Using a similar LC-MS/MS method, the Fmoc-tagged lysine-formaldehyde adduct, **2**, was analyzed in another analysis with the mass spectrometer set to monitor the fragments of *m/z* 397→175.

Calibration. A stock solution of thioproline at 500 μg/mL was prepared in methanol/water (1:1) and stored at -20 ^{o}C before use. The calibration standard solutions of thioproline (10, 25, 50, 100, 500,1000 ng/mL), were prepared by serial dilution of the stock solution, to which was added the isotope-labeled internal standard (Fmoc-l-thioproline-d_2), derivatized with Fmoc-Cl, and LC–MS^3 analyzed as described above. Calibration curves were established by plotting the peak area ratios of **1** to the isotope-labeled internal standard against the concentrations of thioproline in the calibration standard solutions. Using a similar approach, the calibration curve for N^6-formyllysine was also established.

Method Validation. The overall efficiency of the analytical method, which comprised of the protein isolation, enzymatic digestion and chemical derivatization steps, was determined using a synthetic thioproline-containing peptide. The validation experiment entailed spiking 100 μg of purified protein with different amounts of the thioproline-containing peptide (300-3000 pmol), followed by protein isolation, protease digestion to release thioproline, addition of isotope-labeled internal standard, derivatization with Fmoc-Cl, and LC–MS^3 analysis, as described above. The efficiency of the method was evaluated by plotting the measured quantities of thioproline against the quantity of added thioproline-containing peptide.

RESULTS AND DISCUSSION

Optimization of the Derivatization Condition. Fmoc-Cl is a reagent originally developed for amino group protection and is increasingly being used as a derivatization reagent for biogenic amines.[29,30] It introduces an aromatic Fmoc carbamate group to amines and increases their hydrophobicity for reversed-phase HPLC (RP-HPLC) analysis.[31] Previous studies have demonstrated that Fmoc-Cl can react efficiently with amine-containing amino acids, food contaminants, and neurotransmitters, in a slightly alkaline medium to produce stable Fmoc-tagged derivatives for their analysis by RP-HPLC.[32]

As the basis for our method, we have previously used Fmoc-Cl and ethyl chloroformate as derivatization agents to derivatize thiazolidines prior to their LC–MS/MS analysis.[23,27] Because the Fmoc-thioproline derivative formed by reacting thioproline with Fmoc-Cl is more hydrophobic and can be efficiently extracted using ethyl acetate for LC−MS/MS analysis,[27] Fmoc-Cl was therefore used as the reagent for optimizing the chromatographic behavior of thioproline for LC–MS^3 analysis (Figure 2). We have previously demonstrated that Fmoc-Cl reacts rapidly with thioproline at room temperature,[23] and in this study, we have further optimized the amount of reagent needed for the derivatization reaction. By using protein hydrolysate from hydrolyzing 100 μg of control protein as the reaction medium, we determined

that a molar ratio of 2000:1 for Fmoc-Cl to thioproline was necessary for the efficient formation of the conjugate (Figure S1), and further increases in the molar ratio of Fmoc-Cl did not improve the yield of the derivative. Thus, the molar ratio of 2000:1 was used for the entire study.

It is not surprising that such a high molar ratio of Fmoc-Cl:thioproline (2000:1) is needed for the derivatization reaction, because the amount of free thioproline produced during protein hydrolysis would be around a thousand- fold less than all the other amino acids also liberated. All these free amino acids would also react with Fmoc-Cl via their amino terminus and compete with thioproline for the derivatization reagent. Thus, a high molar ratio (2000:1) of Fmoc-Cl to thioproline is needed for efficient conversion of thioproline in a protein hydrolysate to its Fmoc derivative, **1**.

Comparative Study of LC–MS/MS and LC–MS^3 Methods for Proteinic Thioproline Determination. LC–MS^3 analysis has previously been demonstrated to be highly specific and being more sensitive than LC–MS/MS for bioanalysis.[33,34] To address the challenge of analyzing **1** in the presence of high concentrations of interfering amino acids in the protein digest, we compared the performance of the newly developed LC–MS^3 method with that of our previously developed LC–MS/MS method for analyzing thioproline in protein digests.[27] Our analyses revealed that LC–MS^3 analysis provided a similar sensitivity as that of the previously developed LC–

MS/MS method in which the QTRAP MS was operated in multiple-reaction monitoring (MRM) mode. Specifically, we were able to detect one thioproline in the presence of 10^6 amino acids (from 100 μg of protein) using both methods.

Nevertheless, the analysis revealed the LC–MS^3 method provided higher selectivity for identifying compound **1** in the complex matrix of protein digest than LC–MS/MS analysis in MRM mode (Figure 3 and S2), where we observed significantly lower interference in the LC–MS^3 analysis of a protein hydrolysate from 100 μg of hydrolyzed protein (Figure S2), which was obtained from *E. coli* cells exposed to formaldehyde (8.5 mM). It is also believed that the added transition in the MS^3 analysis would increase the specificity of identifying thioproline in the study. A similar observation of higher confidence of quantitation was also observed in previous studies when LC–MS^3 was used to analyze oxidatively damaged DNA lesions.[33,34] Considering the high selectivity and confidence of LC-MS^3 for analyzing thioproline in protein hydrolysate, the LC–MS^3 method was adopted to quantify the amount of thioproline formed in cellular protein.

Method Validation. A thioproline-containing peptide was custom synthesized and used to validate the method. The method validation entails spiking 100 μg of *E. coli* protein with 300-5000 pmol of the peptide followed by protein isolation, protease hydrolysis, derivatization with Fmoc-Cl, and LC–MS^3 analysis of the digests, as

described in the Experimental Section. The yield of thioproline, which corrected for potential loss of the adduct during the protein isolation and in the digestion process, together with the loss of signal for incomplete derivatization, was found to be 72.2% of the theoretical value (Figure S3), which represents the overall efficiency for the method. Measurement of thioproline was thus corrected by a factor of 1.4 to arrive at the quantity of thioproline in the isolated protein samples.

Quantitation of Thioproline in Purified Protein and in Protein in *E. coli* Treated with Formaldehyde. The validated method was applied to quantitate thioproline as a novel modification to cysteine both in cell-free protein exposed to formaldehyde and in protein isolated from *E. coli* whole cell after exposure to formaldehyde (Figure S4). In both cases, the studies revealed a dose-dependent formation of thioproline. As shown in Figure 4, cell-free protein exposed to formaldehyde produced 3.0 (±0.1) thioproline per 10^4 amino acids/for every mM concentration of formaldehyde used. Given that the occurrence of cysteine in *E. coli* protein was reported to be 1.1%,[35] this result translated to 2.7 % of cysteine residues modified per mM of formaldehyde exposure.

The amount of thioproline observed in protein isolated from *E. coli* treated with formaldehyde was found to be 2.8 (±0.3) thioproline per 10^4 amino acids/mM formaldehyde, which is comparable to the result of the cell-free exposure.The

excellent agreement of thioproline content in both cases indicated that the thioproline formation in *E.coli* is mostly resulted from non-enzymatic chemical reactions and unaffected by the cellular environment.

Quantitation of Proteinic Thioproline in Oxidant-Exposed *E. coli* Protein Extract. The feasibility of detecting thiazolidination as a protein modification in cellular protein was extended to quantitate thioproline in protein isolated from *E. coli* cells exposed to hydroxyl radical generator of H_2O_2, Fe^{2+}-EDTA, and NaOCl to simulate an oxidative environment for the cells. It is well-known that the hydroxyl radical generators will produce hydroxyl radicals which will react with lipids and proteins to generate formaldehyde.[12, 36] The results of this analysis are shown in Figure 5, which again showed a dose-dependent formation of thioproline in protein samples isolated from the cells treated in the oxidative conditions. The thiazolidination frequencies observed was 0.38 (±0.01) per 10^4 amino acids per μM NaOCl, 0.6 (±0.05) per 10^4 amino acids per mM Fe^{2+}-EDTA, and 0.022 (±0.004) per 10^4 amino acids per μM H_2O_2 in proteins isolated from toxicant-exposed *E. coli* cells. A similar relative reactivity in the order of NaOCl > H_2O_2 > Fe^{2+}-EDTA was also observed in our previous study.[23]

Comparison with N^6-formyllysine. The results of post-translational thiazolidination to cysteine were compared with that of the well-studied lysine

formylation process.[14,37] To this end, we have also quantitated using LC–MS/MS, the amount of N^6-formyllysine in formaldehyde-treated protein (Figure 4) and in the protein samples from the *E. coli* cells treated with RONS generators (Figure S5). The study revealed the relative quantities of thioproline were 9-25% that of the N^6-formyllysine levels (Table 1).

This observed discrepancy in the levels of formylation of lysine and thiazolidination of cysteine could have been attributed by to two main factors. Firstly, the amino acid composition in the protein should be taken into consideration. It was reported that the level of lysine is 2.2 times higher than that of cysteine in protein.[38-40] In addition, cysteines are often tied up in disulfide bonds in proteins and thus not all cysteine residues would be reactive.[41,42] Therefore, it is expected that the chances of N^6-formyllysine formation would be higher than for thioproline (Figure 1). Secondly, , as indicated in the studies using formaldehyde-treated protein and *E. coli* (Table 1), N^6-formylation of lysine is both chemically driven as well as and enzymatically catalyzed, while there is currently no evidence that an enzyme would catalyse enzymatic pathway for the thiazolidination reaction. Therefore, it is reasonable to expect a larger amount of N^6-formyllysine than thioproline to be produced in these experiments.

The study revealed that the thiazolidination product, though detected to be formed at a lower level than N^6-formyllysine, is nevertheless present in a significant amount in both the *in vitro* and *E. coli* experiments. Despite the unknown biological implications, it is feasible that cysteine thiazolidination could serve as a biomarker for oxidative stress exposure to formaldehyde and it also has the potential of filling one of the gaps in mining for unknown/novel protein modifications in proteins. However, it is possible that thioproline modification may be present in proteins analyzed from tissue fixed with formaldehyde. Thus, caution has to be taken if this biomarker was chosen analyze in formaldehyde-fixed tissue samples.

CONCLUSION

We report in this study the initial identification of a novel protein modification, cysteine thiazolidination, caused by formaldehyde. Using an isotope dilution LC–MS^3 method of high sensitivity and selectivity, we quantitated the absolute and relative quantity of thiazolidination in formaldehyde-treated protein and in proteins isolated from *E. coli* whole cells that were exposed to formaldehyde and oxidative stress factors. The results from our study suggest the possibility of using cysteine thiazolidination in proteins as a potential biomarker for oxidative stress and formaldehyde exposure. It is believed that thioproline may be a better biomarker candidate for oxidative stress exposure than N^6-formyllysine, because the latter can be confused with *N*,*N*-dimethyllysine, a rather common physiological PTM that has the same nominal mass.

Corresponding Author

* E-mail: chanwan@ust.hk

Notes

The authors declare no competing financial interest.

Acknowledgements

We thank Miss Yao Zhao (Summer research student from Kuang Yaming Honors School, Nanjing University) for her assistance with the protein isolation and digestion. We also thank AB Sciex for providing the LC-MS system for the study.

Funding Sources

This research was supported by the Research Grant Council of Hong Kong (grant 16313916). W. Chan thanks the Hong Kong University of Science and Technology for Startup Funding (grant R9310).

Author Contributions

W. Chan and J. L. designed research; J. L. performed research; W. Chan, J. L., and K. K. J. C. analyzed data; and W. Chan wrote the paper.

Abbreviations

PTMs, Post-translational modifications; LC−MS^3, liquid chromatography-tandem mass spectrometry; Fmoc-Cl, fluorenylmethyloxycarbonyl chloride; MS, High-resolution mass spectrometry; MS/MS, product ion scan mass spectrometry; RP-HPLC, reversed-phase high performance liquid chromatography; MRM, multiple-reaction monitoring

Supporting Information Available

Optimization of the molar ratios of Fmoc-Cl over thioproline for the derivatization. Method validation with custom-synthesized thioproline-containing peptide. Comparative analysis of thioproline in protein from formaldehyde-treated *E. coli* by LC-MS/MS and LC-MS^3. LC-MS^3 analyses of authentic Fmoc-thioproline, authentic thioproline after derivatizing with Fmoc-Cl, thioproline in protein treated with formaldehyde, and in protein in E. coli treated with formaldehyde. Dose-dependent formation of N^6-formyllysine in protein in E. coli treated with Fe^{2+}-EDTA, NaOCl, and H_2O_2. This information is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

(1) Jiang, Y., Valente, A. J., Williamson, M. J., Zhang, L., Graves, D. T. (1990) Post-translational modification of a monocyte-specific chemoattractant synthesized by glioma, osteosarcoma, and vascular smooth muscle cells. *J. Biol. Chem. 265*, 18318-18321.

(2) Pennuto, M., Palazzolo, I., Poletti, A. (2009) Post-translational modifications of expanded polyglutamine proteins: impact on neurotoxicity. *Hum. Mol Genet. 18*, 40-47.

(3) Barrios-Rodiles, M., Brown, K. R., Ozdamar, B., Bose, R., Liu, Z., Donovan, R. S., Shinjo, F., Liu, Y., Dembowy, J., Taylor, I. W. (2005) High-throughput mapping of a dynamic signaling network in mammalian cells. *Science 307*, 1621-1625.

(4) Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., Mann, M. (2006) Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*, *127*, 635-648.

(5) Seet, B. T., Dikic, I., Zhou, M. M., Pawson, T. (2006) Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Bio. 7*, 473-483.

(6) Hunter, T. (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond. *Mol.Cell 28*, 730-738.

(7) Choudhary, C., Kumar, C., Gnad, F., Nielsen, M. L., Rehman, M., Walther, T. C., Olsen, J. V., Mann, M. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science 325*, 834-840.

(8) Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E. b., Murray, B., Latham, V., Sullivan, M. (2011) PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res. 40*, D261–D270..

(9) Janke, C., Kneussel, M. (2010) Tubulin post-translational modifications: encoding functions on the neuronal microtubule cytoskeleton. *Trends Neurosci. 33*, 362-372.

(10) Cloos, P. A., Christgau, S. (2004) Post-translational modifications of proteins: implications for aging, antigen recognition, and autoimmunity. *Biogerontology. 5*, 139-158.

(11) Shibamoto, T. (2006) Analytical methods for trace levels of reactive carbonyl compounds formed in lipid peroxidation systems. *J. Pharmaceut Biomed. 41*, 12-25.

(12) O'Brien, P. J., Siraki, A. G., Shangari, N. (2005) Aldehyde sources, metabolism, molecular toxicity mechanisms, and possible effects on human health. *Crit. Rev. Toxicol. 35*, 609-662.

(13) Kerns, W.D., Pavkov, K.L., Donofrio, D.J., Gralla, E.J., Swenberg, J.A. (1983) Carcinogenicity of formaldehyde in rats and mice after long-term inhalation exposure. *Cancer Res. 43*, 4382-4393.

(14) Jiang, T., Zhou, X., Taghizadeh, K., Dong, M., Dedon, P. C. (2007) N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. *Proc. Natl. Acad. Sci. U. S. A. 104*, 60-65.

(15) Metz, B., Kersten, G. F., Hoogerhout, P., Brugghe, H. F., Timmermans, H. A., De Jong, A., Meiring, H., ten Hove, J., Hennink, W. E., Crommelin, D. J. (2004) Identification of formaldehyde-induced modifications in proteins reactions with model peptides. *J. Biol. Chem. 279*, 6235-6243.

(16) Boissel, J. P., Kasper, T. J., Shah, S. C., Malone, J. I., Bunn, H. F. (1985) Amino-terminal processing of proteins: hemoglobin South Florida, a variant with retention of initiator methionine and N alpha-acetylation . *Proc. Natl. Acad. Sci. U. S. A. 82*, 8448-8452.

(17) Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev. 47*, 1.

(18) Yang, X. J., Seto, E. (2008) The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat. Rev. Mol. Cell Bio. 9*, 206-218.

(19) Tokikawa, R., Loffredo, C., Uemi, M., Machini, M., Bechara, E. (2014) Radical acylation of L-lysine derivatives and L-lysine-containing peptides by peroxynitrite-treated diacetyl and methylglyoxal. *Free Radical Res. 48*, 357-370.

(20) Wang, Z. M., Lashuel, H. A. (2013) Discovery of a novel aggregation domain in the huntingtin protein: implications for the mechanisms of Htt aggregation and toxicity. *Angew. Chem. Int. Edit. 52*, 562-567.

(21) Soreghan, B., Kosmoski, J., Glabe, C. (1994) Surfactant properties of Alzheimer's A beta peptides and the mechanism of amyloid aggregation. *J. Biol. Chem. 269*, 28551-28554.

(22) Lorton, D., Schaller, J., Lala, A., De Nardin, E. (2000) Chemotactic-like receptors and Aβ peptide induced responses in Alzheimer's disease. *Neurobiol. Aging 21*, 463-473.

(23) Liu, J., Chan, W. (2014) Quantification of thiazolidine-4-carboxylic acid in toxicant-exposed cells by isotope-dilution liquid chromatography-mass spectrometry

reveals an intrinsic antagonistic response to oxidative stress-induced toxicity. *Chem. Res. Toxicol. 28*, 394-400.

(24) Edrissi, B., Taghizadeh, K., Moeller, B. C., Kracko, D., Doyle-Eisele, M., Swenberg, J. A., Dedon, P. C. (2013) Dosimetry of N^6-formyllysine adducts following [$^{13}C_2H_2$]-formaldehyde exposures in rats. *Chem. Res. Toxicol. 26*, 1421-1423.

(25) Jiang, T., Zhou, X., Dong, M., Dedon, P. C. (2005) Deoxyribose oxidation in DNA leads to the formation of N6-formyllysine residues in histone proteins. *Cancer Res. 65*, 86-87.

(26) Duong-Ly, K. C., Gabelli, S.B., Jon. L. (2014) Chapter seven-salting out of proteins using ammonium sulfate precipitation. *Methods Enzymol. 541*, 85-94.

(27) Liu, J., Meng, X., Chan, W. (2016) Quantitation of thioprolines in grape wine by isotope-dilution liquid chromatography-tandem mass spectrometry. *J. Agric. Food Chem. 64*, 1361-1366.

(28) Capelo, J., Ximenez-Embun, P., Madrid-Albarran, Y., Camara, C. (2014) Enzymatic probe sonication: enhancement of protease-catalyzed hydrolysis of selenium bound to proteins in yeast. *Anal. Chem. 76*, 233-237.

(29) Sureshbabu, V. V., Venkataramanarao, R., Naik, S. A., Chennakrishnareddy, G. (2007) Synthesis of tetrazole analogues of amino acids using Fmoc chemistry: isolation of amino free tetrazoles and their incorporation into peptides. *Tetrahedron Lett. 48*, 7038-7041.

(30) López-Cervantes, J., Sánchez-Machado, D.I., Rosas-Rodríguez, J. A. (2006) Analysis of free amino acids in fermented shrimp waste by high-performance liquid chromatography. *J. Chromatogr. A. 1105*, 106-110.

(31) Bank, R.A., Jansen, E.J., Beekman, B., te Koppele, J.M. (1996) Amino acid analysis by reverse-phase high-performance liquid chromatography: improved derivatization and detection conditions with 9-fluorenylmethyl chloroformate. *Anal. Biochem. 240*, 167-176.

(32) Pérez-Míguez, R., Marina, M.L., Castro-Puyana, M. (2016) Capillary electrophoresis determination of non-protein amino acids as quality markers in foods. *J. Chromatogr. A. 1428*, 97-114.

(33) Wang, J., Yuan, B., Guerrero, C., Bahde, R., Gupta, S, Wang, Y. (2011) Quantification of oxidative DNA lesions in tissues of Long-Evans Cinnamon rats by capillary high-performance liquid chromatography−tandem mass spectrometry coupled with stable isotope-dilution method. *Anal. Chem*. *83*, 2201-2209.

(34) Hong, H., Wang, Y. (2007) Derivatization with Girard reagent T combined with LC-MS/MS for the sensitive detection of 5-formyl-2'-deoxyuridine in cellular DNA. *Anal. Chem. 79*, 322-326.

(35) Miseta, A., Csutora, P. (2000) Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Mol. Biol. Evol. 17*, 1232-1239.

(36) Onyango, A. N. (2012) Small reactive carbonyl compounds as tissue lipid oxidation products; and the mechanisms of their formation thereby. *Chem. Phys. Lipids 165*, 777-786.

(37) Zheng, S., Doucette, A.A. (2016) Preventing N- and O- formylation of proteins when incubated in concentrated formic acid. *Proteomics 16*, 1059-1068.

(38) Wilson, R. P. (2006) Fish nutrition Academic Press. *Amsterdam*.

(39) Jukes, T. H., Holmquist, R., Moise, H. (1975) Amino acid composition of proteins: selection against the genetic code. *Science 189*, 50-51.

(40) Cedano, J., Aloy, P., Perez-Pons, J. A., Querol, E. (1997) Relation between amino acid composition and cellular location of proteins. *J. Mol. Biol. 266*, 594-600.

(41) Matsumura, M., Becktel, W.J., Levitt, M., Matthews, B.W. (1989) Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc. Natl. Acad. Sci. U. S. A. 86*, 6562-6566.

(43) Hogg, P. J. (2003) Disulfide bonds as switches for protein function. *Trends Biochem. Sci. 28*, 210-214.

Table 1

Formation of N^6-Formyllysine and Thioproline in Purified Protein and in Protein Isolated from *E. coli* Whole Cell Treated with Formaldehyde and Oxidants.

	Adducts/10^4 amino acids/mM toxicant		Ratio of Thioproline: N^6-formyllysine
	Thioproline	N^6-formyllysine	
In protein treated with:			
Formaldehyde	3.0±0.1 (2.7%)	12.9±1.3	0.20
In protein from E. coli treated with:			
Formaldehyde	2.8±0.3 (2.5%)	30.8±2.1	0.09
NaOCl[a]	0.38±0.01 (0.35%)	1.5±0.02	0.25
Fe^{2+}-EDTA	0.6±0.05 (0.55)	2.5±0.2	0.24
H_2O_2[a]	0.022±0.004 (0.02%)	0.17±0.005	0.13

[a] Adducts/10^4 amino acids/μM toxicant**.** Showed in the parenthesis is the % of cysteine modified per unit of toxicant exposed.

FIGURE LEGEND

Figure 1. Formation of N^6-formyllysine and thioproline in protein by reacting formaldehyde with the lysine and cysteine residues, respectively.

Figure 2. Reaction of thioproline and N^6-formyllysine with fluorenylmethyloxycarbonyl chloride, Fmoc-Cl, produce stable Fmoc carbamate derivatives of enhanced chromatographic performance for reversed-phase HPLC analysis.

Figure 3. Representative chromatograms for LC–MS^3 monitoring the *m/z* 356→134→88 for unlabeled (A) and *m/z* 358→136→90 for isotopic-labeled (B) thioproline after reacting with Fmoc-Cl. Together with the positive-ion MS/MS and MS^3 spectra for the unlabeled (C) and labeled (D) thioproline.

Figure 4. Formation of thioproline in protein (A) and in protein in *E. coli* (B) treated with formaldehyde. Protein and cells were exposed to formaldehyde at 37 ^oC and processed for LC–MS^3 analysis as described in the Experimental section.

Figure 5. Dose-dependent formation of thioproline in protein in *E. coli* treated with Fe^{2+}-EDTA (A), NaOCl (B), and H_2O_2 (C). Cells were exposed to oxidants at 37 ^oC and processed for LC–MS^3 analysis as described in the Experimental section.

Figure 1

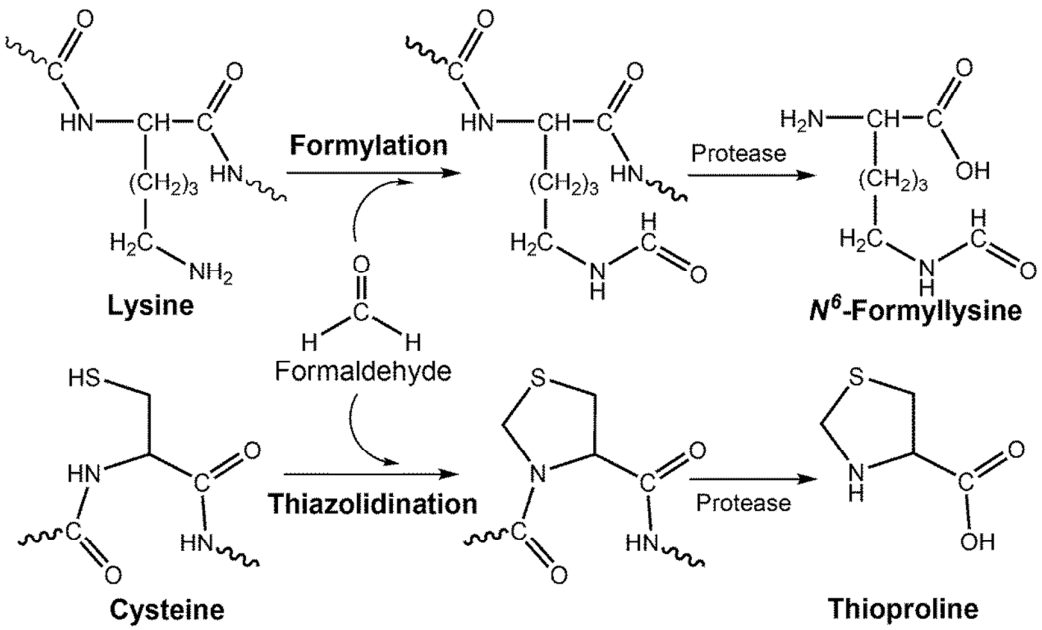

Formylation
Protease
Lysine
Formaldehyde
N^6-Formyllysine
Thiazolidination
Protease
Cysteine
Thioproline

Figure 2

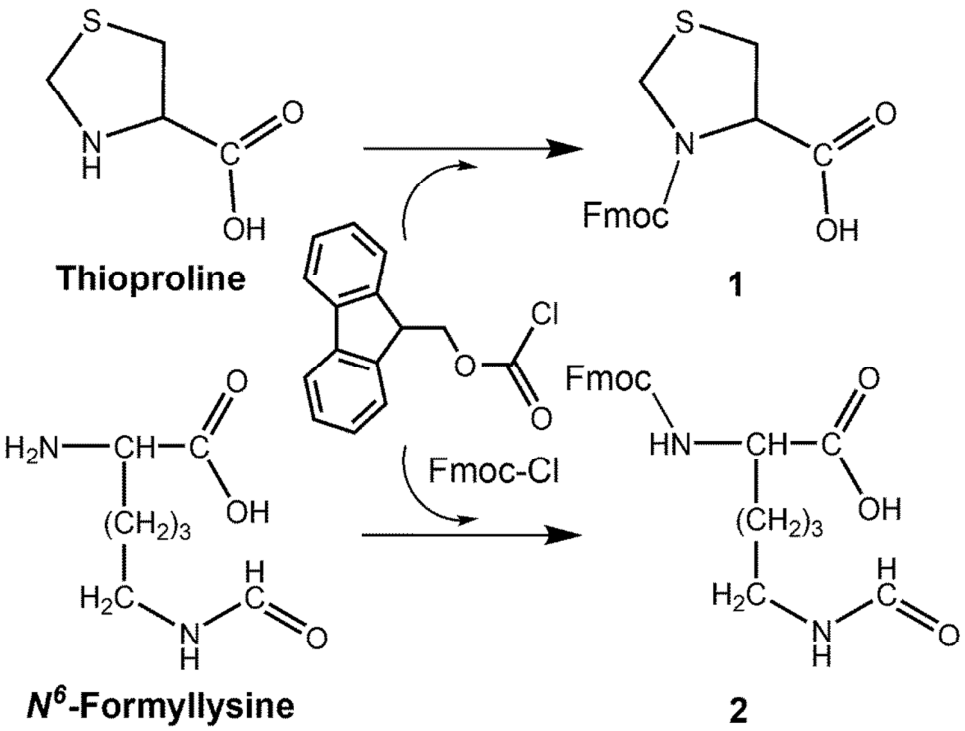


Figure 3

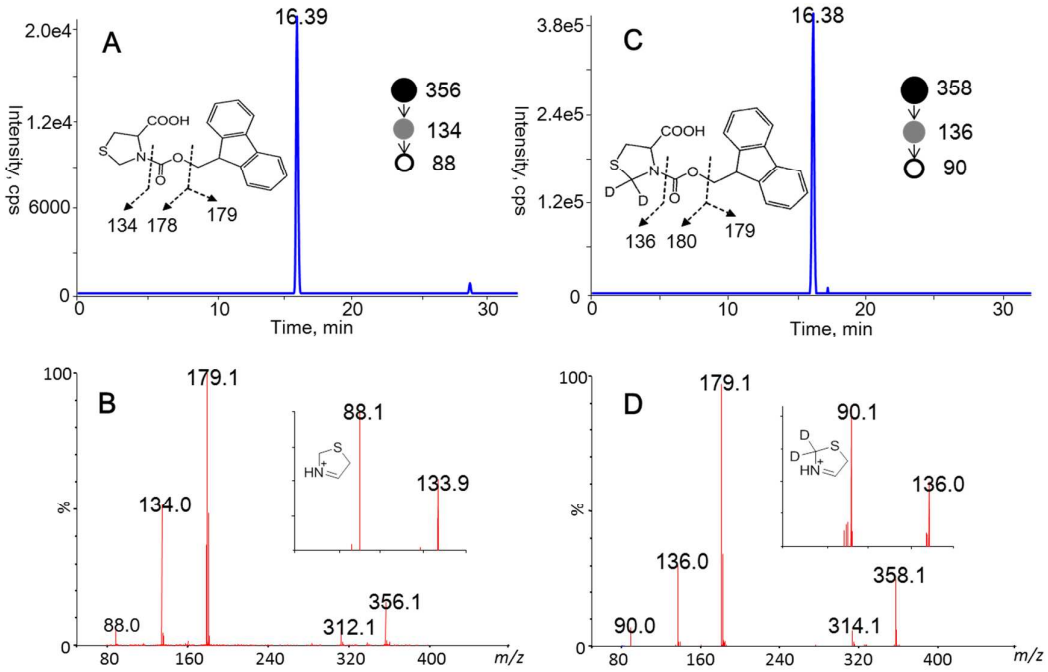

A
16.39
2.0e4
1.2e4
6000
0
Intensity, cps
356
134
88
COOH
134 178
179
0 10 20 30
Time, min
C
16.38
3.8e5
2.4e5
1.2e5
0
Intensity, cps
358
136
90
COOH
D
D
136 180
179
0 10 20 30
Time, min
B
179.1
100
%
88.1
133.9
134.0
88.0
356.1
312.1
0
80 160 240 320 400
m/z
D
179.1
100
%
90.1
136.0
136.0
358.1
90.0
314.1
0
80 160 240 320 400
m/z

Figure 4

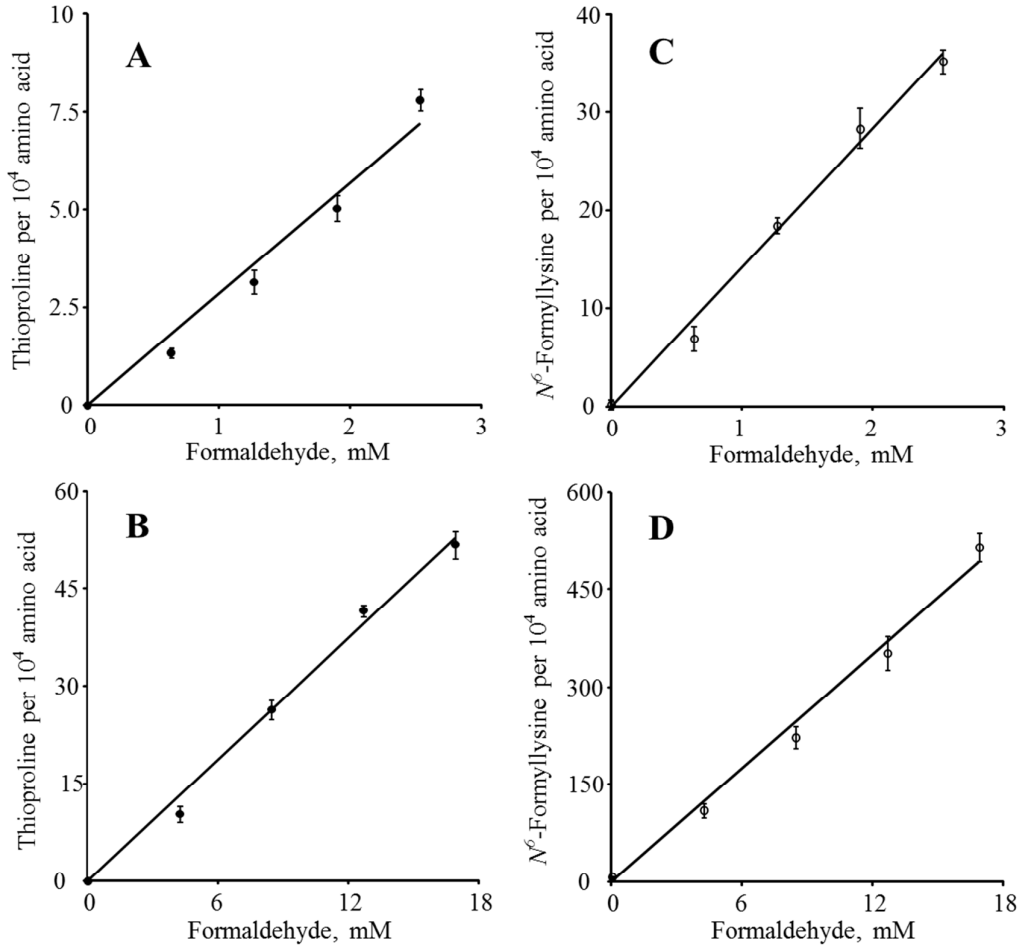
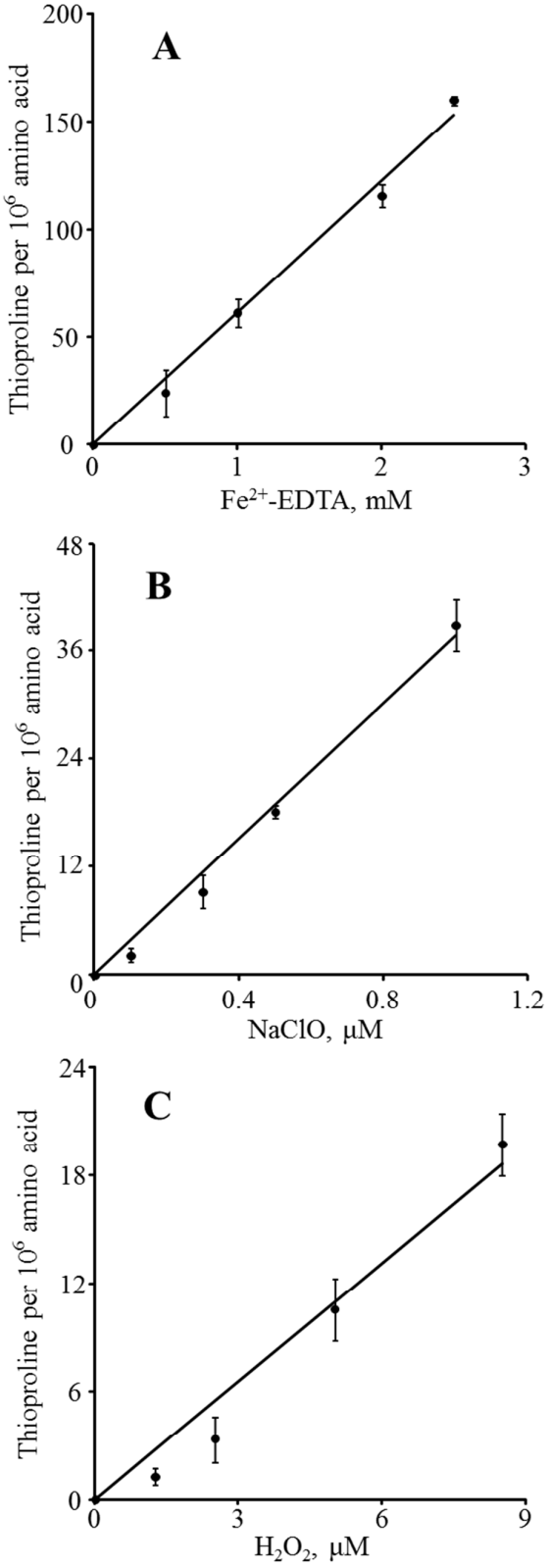

A
Thioproline per 10^4 amino acid
Formaldehyde, mM
B
Thioproline per 10^4 amino acid
Formaldehyde, mM
C
N^6-Formyllysine per 10^4 amino acid
Formaldehyde, mM
D
N^6-Formyllysine per 10^4 amino acid
Formaldehyde, mM

Figure 5


A
Thioproline per 10^6 amino acid
200
150
100
50
0
0
1
2
3
Fe^{2+}-EDTA, mM
B
Thioproline per 10^6 amino acid
48
36
24
12
0
0
0.4
0.8
1.2
NaClO, μM
C
Thioproline per 10^6 amino acid
24
18
12
6
0
0
3
6
9
H_2O_2, μM