SCIENCE CHINA
Chemistry

• ARTICLES •
• SPECIAL TOPIC • Mass Spectrometry Analysis

May 2014 Vol.57 No.5: 708–717
doi: 10.1007/s11426-013-5049-8

A novel method for identification and relative quantification of N-terminal peptides using metal-element-chelated tags coupled with mass spectrometry

YAN Hui[1,2], HAO FeiRan[2], CAO QiChen[3,2], LI JiaBin[2], LI NanNan[2], TIAN Fang[2], BAI HaiHong[2], REN XiaoJun[2], LI XianYu[2], ZHANG YangJun[1,2*] & QIAN XiaoHong[2]

[1]*Anhui Medical University, Hefei 230032, China*
[2]*State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing 102206, China*
[3]*Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang 110032, China*

Received September 1, 2013; accepted November 6, 2013; published online February 24, 2014

By means of the reaction between a DOTA-NHS-ester bifunctional reagent and N-terminal peptides of proteins, and then chelation of lanthanide metal ions as tags, we established a novel method for the identification of N-terminal peptides of proteins and their relative quantification using metal-element-chelated tags coupled with mass spectrometry. The experimental results indicate that metal elements are able to completely label N-terminal peptides at the protein level. The N-terminal peptides are enriched as the peptides digested with trypsin are selectively eliminated by isothiocyanate-coupled silica beads. We successfully identified the N-terminal peptides of 158 proteins of *Thermoanaerobacter tengcongensis* incubated at 55 and 75 °C, among which N-terminal peptides of 24 proteins are partially acetylated. Moreover, metal-element tags with high molecule weights make it convenient for N-terminal peptides consisting of less than 6 amino acids to be identified; these make up 55 percent of the identified proteins. Finally, we developed a general approach for the relative quantification of proteins based on N-terminal peptides. We adopted lysozyme and ribonuclease B as model proteins; the correlation coefficients (R^2) of the standard curves for the quantitative method were 0.9994 and 0.9997, respectively, with each concentration ratio ranging from 0.1 to 10 and both relative standard derivations (RSD) measured at less than 5%. In *T. tengcongensis* at two incubation temperatures, 80 proteins possess quantitative information. In addition, compared with the proteins of *T. tengcongensis* incubated at 55 °C, in *T. tengcongensis* incubated at 75 °C, 7 proteins upregulate whereas 16 proteins downregulate, and most differential proteins are related to protein synthesis.

metal-element-chelated tags, N-terminal peptides, proteomics, mass spectrometry

1 Introduction

Shotgun strategy, which is based on two-dimensional capillary liquid chromatography and mass spectrometry, is the most common strategy in proteomic research. According to this strategy, protein mixture is first digested into peptide mixture with an enzyme, and then analyzed by high-performance liquid chromatography coupled with tandem mass spectrometric (HPLC-MS/MS) technology [1]. The proteolytic digestion increases the complexities of samples, however, which makes it difficult to identify all of the proteins even if the most advanced technologies are applied [2]. A method involving tagged peptides can be used to solve this problem because the specific enrichment remarkably reduces the complexity. The N-terminal peptides of proteins are not only the most representative tagged peptides [3, 4] but also important structural and functional parts that are

*Corresponding author (email: 13683167093@163.com)

vital both to the identification of proteins and their research of biological functions. The half-life period of a protein is significantly affected by the N-terminal amino acids (the N-end rule) [5, 6]. In addition, modifications of terminal peptides of proteins have great influence upon their functions [7, 8]. A revision and supplement of a genome could be made by the identification of terminal peptides, which is a means of finding transcription initiation sites and termination sites of protein-coding genes. Terminal sequences, whether N-terminal or C-terminal, are highly specified. A few amino acid residues could be used to reliably identify most proteins [9]. For example, Wilkins *et al*. [10] analyzed the specificity of theoretical amino-acid terminal sequences of 15519 kinds of proteins from the SWISS-PROT database for organisms *Mycoplasma genitalium*, *Bacillus subtilis*, *Escherichia coli*, *Saccharomyces cerevisiae*, and *homo sapiens*. The results showed that 43% to 83% of proteins (varying by species) could be identified by N-terminal peptides consisting of four amino acid residues, and 74% to 97% of proteins (varying by species) could be identified by C-terminal peptides consisting of four amino acid residues. If the number of amino-acid residues in terminal peptides identified increases to five, a higher specificity will be reached: 78% to 97% of proteins are identified by N-terminal peptides and 81% to 99.6% of proteins by C-terminal peptides. In summary, terminal amino-acid analysis, including identification and quantitation, is of significant importance.

Currently, two methods are in use to study N-terminal peptides, according to whether or not they are involved in chemical derivatization. The first is called the chemical derivatization method. The Christopher M. Overall group was the first to adopt the TAILS (terminal amine isotopic labeling of substrates) method to study terminal peptides of proteins. In their study of N-terminal peptides, they utilized dendritic polyglycerol aldehyde polymers for terminal peptide enrichment and identified 731 acetylated N-terminal peptides, 132 cyclized N-terminal peptides, and 288 MMP-2 sites in mouse fibroblast secretomes. It was pointed out that the methionine in the initiation sites can be easily eliminated when the second amino acid of N-terminal peptides is one of the four amino acids (valine, glycine, alanine, and serine). After the elimination, alanine, serine, and methionine at the second site are more likely to be acetylated [11]. Several optimizations of the TAILS method have been made and the application to N-terminal peptides has obtained remarkable results [12–17]. Gevaert *et al*. [18, 19] utilized the COFRADIC (combined fractional diagonal chromatography) method to reduce the complexity of samples analyzed by mass spectrometry. According to this method, proteins are first acetylated at unbounded amino groups, including side-chain amino groups and N-terminal amino groups; after digestion, the peptide mixture partly underwent RP-HPLC-MS. Then the peptide mixture was treated with 2,4,6-trinitrobenzenesulfonicacid (TNBS) and underwent RP-HPLC-MS a second time. The differential chromatography behaviors indicated all non-N-terminal peptides. McDonald *et al*. [20] used a similar method that they modified by replacing TNBS with acylated biotin; in addition, a protein mixture was fractionated by immobilized streptavidin solid phase column, which could combine biotin specifically. Thus N-terminal peptides are enriched. These researchers used trypsin to digest mouse-liver samples and confirmed that more than 85% of all proteins yielded informative N-terminal peptides. The detection rate increased to nearly 90% when two endopeptidases (trypsin and endopeptidase Glu-C) were used. Athena A. compared acetic anhydride with propionic anhydride in blocking N-terminal peptides and ε-amine of lysine, and showed that productivity could reach 95% and that propionic anhydride is more suitable for the identification of naturally acylated blocked N-terminal peptides [21]. Zhao *et al*. [22] utilized anhydride to modify proteins and enriched N-terminal peptides by isothiocyanate-coupled magnetic nanoparticles. In their study, 215 N-terminal peptides were identified in HepG2 cells, among which 129 were acetylated, 52 were cyclized, and 32 were non-modified. Mommen *et al*. [23] labeled digested peptides with phosphorylation tags after the protein was dimethylated. The N-terminal peptides were enriched by titanium dioxide and hydrophilic interaction liquid chromatography (TiO_2-HILIC). This method has been applied in many sample analyses, with remarkable results. In addition, due to their unique characteristics, the modified N-terminal peptides are used for the quantitation of proteins containing free N-terminal peptides. Hsu *et al*. [24] remarkably improved the signal of a_1 ions of modified N-terminal peptides by dimethylation in the MS/MS spectra. These a_1 ions, which were invulnerable to the detection, were used as mass tags in de novo and database searches to increase reliability. These researchers also used hemoglobin to test this method for feasibility as it is applied in relative quantitation. In Overall's approach, relative quantification methods for N-terminal peptides based on TAILS combining dimethylation, iTRAQ, and SILAC technology were developed. The same applies to qualification and quantification of C-terminal peptides [25, 26]. Tsiotis *et al*. [27] labeled a Q fever pathogen with ^{12}C4-butyryl and ^{13}C4-butyryl. Eventually they used the COFRADIC method to identify the N-terminal peptides of 322 proteins in the Q fever pathogen. Among these, differential expressions of 44 proteins were found between the two different bacterial strains and 10 of them are related to some unknown function of housekeeping enzymes.

The second method is called selective enrichment. For example, Gorman *et al*. [28] enriched terminal peptides by making use of the property that blocked N-terminal peptides and non-modified C-terminal peptides are more vulnerable than other peptides to elution in SCX. Dormeyer *et al*. [29] adopted the SCX method to fractionate blocked N-terminal peptides and non-modified C-terminal peptides, searched all

terminal peptide sequences in a forward-reversed database, and identified 263 known and 83 unknown N-terminal sequences in an IPI protein database as well as 168 known and 193 unknown C-terminal sequences. Zhang *et al*. [30] adopted a CNBr-activated sepharose resin to enrich acetylated N-terminal peptides selectively at pH 6.0. Lysine-ending N-terminal peptides would not be missed in this method, but blocked N-terminal peptides that contain histone are vulnerable to being missed in the enrichment.

All of the above methods help to identify the N-terminal peptides of proteins. However, the reagent with small molecules used in modifying N-terminal peptides makes it inappropriate for mass spectrometry to analyze the short N-terminal peptides that make up most of the N-terminal peptides. For instance, 57.03% of N-terminal peptides in the tryptic-digested proteins of *T. tengcongensis* have fewer than 6 amino acids from Swissprot in silico. Advantages such as high efficiency, stable labeling, and multiple labeling make the metal-element-chelated tags play an important role in proteomic quantitation research. Our approach is focused on the application of metal-element-chelated tags to the identification and relative quantitation of N-terminal peptides [31–33]. The N-terminal peptides react with DOTA-NHS ester, which then forms a complex with lanthanide metal ions. These peptides are finally enriched by the removal, using isothiocyanate-coupled silica beads, of newly generated peptides digested by trypsin. All of the N-terminal peptides, including the ones in short sequences, can be retained on a chromatographic column and detected by mass spectrometry because the metal-element-chelated tags add more than 500 Da to each of the N-terminal peptides and increase their hydrophobicity.

2 Experimental

2.1 Apparatus and reagents

We used a 4800 Proteomics Analyzer Matrix Assisted Laser Desorption Ionization Time of Flight Mass Spectrometry (MALDI-TOF-MS, AB SCIEX, USA). The Capillary High Performance Liquid Chromatography-Eletrospray Ionization-Linear Trap Quadruple-Fourier Transform Mass Spectrometry (HPLC-ESI-LTQ-FT MS) was from Thermo Fisher Scientific, USA. The Easy Nano LC was from Thermo Fisher Scientific, USA. The LTQ-Orbitrap Velos was from Thermo Fisher Scientific, USA. The Thermo Orion MODEL 818 pH meter was from Thermo Fisher Scientific, USA. The C18 zip tip column and 3 kD Amicon Ultra-0.5 mL Centrifugal Filters were from Millipore, USA. The Sartorius BP211d Analytical Balance was from Sartorius, Switzerland. The Super-Speed Centrifuge and Vacuum Concentrator SC100ASpeedvac Plus were from Thermo Fisher Scientific, USA. The Milli-Q Water Purification System Milli2QA10 was from Millipore, USA.

Synthetic peptides were obtained from GL Biochem Shanghai Co., Ltd., China; myoglobin, lysozyme, ribonuclease B (RNase B), α-Cyano-4-hydroxycinnamic acid (CHCA, ⩾ 95%), holmium chloride ($HoCl_3$), terbium chloride ($TbCl_3$), thulium chloride ($TmCl_3$), triethylammonium bicarbonate (TEAB), ammonium bicarbonate, and ammonium acetate were from Sigma-Aldrich, USA. The sequencing-grade trypsin was from Promega, USA. Mono-*N*-hydroxysuccinimide ester 1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid (DOTA-NHS ester) was from Macrocyclics, USA. The *o*-methylisourea hemisulfate, 1,4-phenlylene diisothiocyanate (DITC), dimethyl formamide (DMF), and dimethyl sulfoxide (DMSO) were from Acros, Belgium. The micrometer-grade amino-functionalized silica beads were from Agela, China. *T. tengcongensis* was provided by the Beijing Genomics Institute. The water we used was ultra-purified with a MilliQA10 purification system from Millipore, USA. Other reagents were of analytical grade and from Sinopharm Chemical Reagent Co., Ltd., China.

2.2 Methods

Protein guanidination

We adopted a modified method to achieve guanidination modification on the side-chains of lysines in proteins [34]. A certain amount of protein, at a final concentration of 1 μg/uL, was dissolved in a buffer solution made up of 8 mol/L urea, 0.1% SDS, and 400 mmol/L NH_4HCO_3. The DTT solution was added subsequently at a final concentration of 10 mmol/L. The whole solution was placed in a water bath at 37 °C for 4 h. Then IAA solution was added at a final concentration of 50 mmol/L and the whole solution was placed in dark for 1 h. The solution was desalted with 3 kDa Amicon Ultra-0.5 mL Centrifugal Filters, after which *o*-methylisourea in ammonia solution (pH 11) was added at a final concentration 1.5 mmol/L. The whole solution was desalted once more and dried after a microwave-assisted reaction.

Metal-element-chelated tags and proteolytic digestion

Guanidition modified protein samples were dissolved in 0.2 mol/L TEAB buffer solution (pH 8.5). DOTA-NHS-ester dissolved in acetonitrile was added to the solution at a ratio of aqueous phase/acetonitrile phase of 1/3 (*v*/*v*), and then reacted at 20 °C for 1 h. The whole product was dried before 0.1 mol/L of ammonium acetate solution was added to a final pH of 5.6. Metal salt ($TbCl_3$, $HoCl_3$) dissolved in ammonium acetate solution, according to the ratio of DOTA-NHS-ester/MCl_3 (1/2, mol/mol), was added prior to an incubation at 37 °C in a water bath for 1 h. The labeled samples were desalted before being dissolved in 50 mmol/L of ammonium bicarbonate solution. Trypsin was added according to the mass ratio (1∶50, trypsin/protein). The mixture was vortexed well prior to being incubated at 37 °C for 16 h.

Synthesis of DITC-bonded silica beads and their application

First, 1 mg of NH_2-silica beads were added to 500 μL of DITC dissolved in DMF at a concentration of 50 mg/mL. Then this solution was shaken for 1 h at room temperature before being washed by 200 μL of DMF/DMSO (1:1) to remove excessive DITC. Digested peptides were dissolved in 0.1 mol/L of borate buffer (pH 8.0) and fully mixed with DITC-silica beads. The whole mixture was shaken and then centrifuged at 14000 G at 50 °C for 5 h before the supernatant was identified by MALDI-TOF MS.

Extraction of total proteins from T. tengcongensis

An appropriate amount of lysis buffer (8 mol/L urea, 5 mg DTT, 10 μL protease inhibitor) was added to *T. tengcongensis* cells prior to ultrasonication in an ice bath. The ultrasonication was run for 8 cycles of 3 min' run and 3 min' pause. The lysis buffer was subsequently centrifuged at 20000 round/min at 4 °C for 30 min before the supernatant was harvested for a concentration measurement. The extract was split into fractions and saved at –80 °C.

Metal-element-chelated tags on N-terminal peptides of total proteins from T. tengcongensis

The DTT solution, at a final concentration of 10 mmol/L, was added to *T. tengcongensis* protein samples. The solution reacted in a water bath at 37 °C for 4 h before a final concentration of 50 mmol/L IAA solution was added. The whole solution was placed in the dark at room temperature for 1 h. The solution was then desalted before a final concentration of 50 mmol/L of O-Methylisourea dissolved in a solution of ammonia and water (pH 11) was added. The reaction was microwaved for 1 min. Subsequently, the samples were labeled by metal elements and digested before being fully mixed with synthesized DITC-silica beads and shaken at 50 °C for 5 h. The samples were centrifuged at 15000 G and the supernatant was harvested before a C18 zip-tip desalting process. The desalted solution was dried prior to being dissolved again in mobile phase A (2% ACN, 0.1% FA). The redissolved samples were centrifuged at 15000 G for 10 min before an HPLC-ESI-LTQ-FT MS analysis.

Mass spectrometry analysis

MALDI-TOF-MS analysis: Proteomics Analyzer Matrix-Assisted Laser Desorption Ionization Time of Flight Mass Spectrometry was used to analyze the samples. First, 1 μL of sample to be analyzed was deposited on the sample plate; next, 1 μL of CHCA solution at a concentration of 5 mg/mL was deposited after the prior drop had dried. The sample was analyzed after the second drop had dried at room temperature. The peptides from myoglobin digested with trypsin were used to calibrate the instrument, with a max outlier error of 10 ppm. The acquisition method of MS mode was MS-2 kV reflective mode (acceleration voltage 20 kV, mass range 600 to 4000 *m*/*z*, laser intensity 3800, sub-spectra accumulated in each spectrum 1500 times). The acquisition method of MS/MS mode was MS/MS 2 kV reflective mode (laser intensity 5000, sub-spectra accumulated in each spectrum 2000 times). The peaks were recorded at an *S*/*N* of more than 20 in MS mode. The spectra were analyzed by GPS software from AB SCIEX (USA).

HPLC-ESI-LTQ-FT MS analysis: The high-performance liquid chromatography was Agilent 1100 capillary liquid chromatography, loaded by autosampler. The samples were fractionated by this chromatograph before entering the ESI source and being analyzed by mass spectrometry. Liquid chromatography conditions were as follows: mobile phase A: 2% ACN, 0.1% FA; mobile phase B: 80% ACN, 0.1% FA. Elution gradients were as follows: 0 to 90 min, 6% to 40% mobile phase B; 90 to 100 min, 40% to 100%, mobile phase B; 100 to 110 min, 100%, mobile phase B; 110 to 120 min, 100% mobile phase B to 100% mobile phase A, and equilibration for 15 min. Loading volume: 20 μL. Flow rate: 300 nL/min. Mass spectrometry conditions: acquisition mode: positive ion mode; mass scan range: 375.0 to 1500.0 *m*/*z*; acquisition time: 110 min. The 10 ions with strongest intensity were selected for MS/MS analysis, and the collision energy of ions was 35 V. Dynamic exclusion was adopted, and the exclusion time was 30 s.

HPLC-ESI-LTQ-Orbitrap Velos analysis: Liquid chromatography was Easy Nano Capillary LC. Reverse phase chromatography column was Magic TM C18 column (10 cm × 75 μm i.d., 3 μm). Mobile phase A: 2% ACN, 0.1% FA, mobile phase B: 80% ACN, 0.1% FA. Flow rate: 350 μL/min. Loading volume: 5 μL. Elution gradient: 0 to 90 min, 6% to 40% mobile phase B; 90 to 100 min, 40% to 100% mobile phase B; 100 to 110 min, 100% mobile phase B; 110 to 120 min, 100% mobile phase B to 100% mobile phase A, and equilibrated for 15 min. Mass spectrometry conditions: Nano-ESI, positive ion mode. Ionization voltage: 2000 V. Temperature at capillary entrance: 350 °C. Mass range: 350 to 1200 *m*/*z*. Exclusion time: 30 s Mass spectrometry data was acquired by XCalibur.

Mass spectrometry data explorer: the raw files acquired were searched by PEAKS 6 software against the protein database of *T. tengcongensis* which contains 2588 proteins. The parameters and settings: precursor tolerance, 20 ppm; fragment tolerance, 0.5 Da; enzyme used, trypsin; max missed cleavages, 6. Fixed modification: cysteine alkylation (C+57.02 Da). Variable modifications: methionine oxidation (M+15.99 Da), N-terminal amino group formylation (+28.03 Da), acetylation (+42.01 Da), DOTA-Tb modification (+542.08 Da), DOTA-Ho modification (+548.09 Da), DOTA-Tm modification (+552.09 Da), and guanidination on ε-amino group of lysine side-chains (+42.02). The confidence level was 95%.

3 Results and discussion

3.1 Experiment principles

The workflow of metal element chelated tags applied to the identification and relative quantitation of N-terminal peptides is shown as Figure 1. Proteins were denaturated, reduced, and alkylated so that the peptide chains were fully unfolded. The ε-amine on side chains of lysine were subsequently guanidyl-modified and thus protected from reacting. DOTA-NHS-ester reacts with N-terminal amines prior to chelating with lanthanide metal ions. The isothiocyanate-coupled silica beads were utilized to remove non-N-terminal peptides from the digested proteins, leaving only N-terminal peptides. This application remarkably reduced the signal suppression caused by excessive non-N-terminal peptides in mass spectrometry analysis. The final identification and relative quantitation was achieved through data analysis.

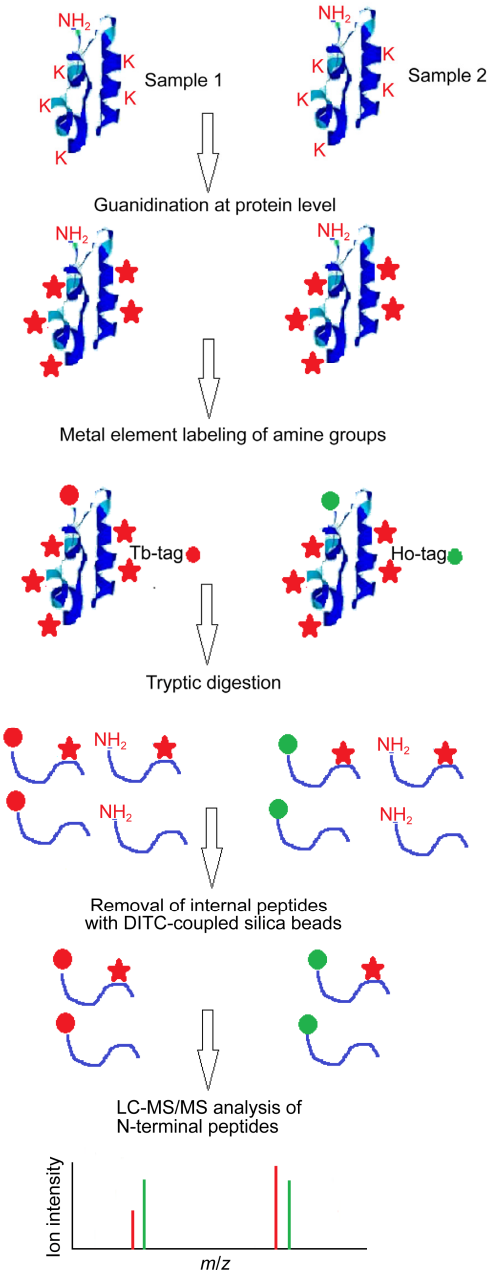

Figure 1 The workflow of metal-element-chelated tags applied to the identification and relative quantitation of N-terminal peptides.

3.2 Efficiency of guanidination of ε-amines of lysines

All free amino groups could be labeled with metal-element tags. The molecule weight of N-terminal peptides increases when tagged, which makes short N-terminal peptides to be identified, whereas the peptides' mass increases enormously when ε-amines of lysine side-chains are also tagged, if the lysine is on this N-terminal peptide. This metal tag would add more than 500 Da to the peptide and this could generally decrease the sensitivity of mass spectrometry. Guanidination could specifically transform lysine into homoarginine and thus prevent ε-amines being labeled. To avoid this side reaction, we selected protein guanidination as the first step. The N-terminal peptide sequence of ribonuclease B is KETAAAKFER, which contains two lysine residues, according to the SWISS-PROT database (http://au.expasy.org/uniprot/P61823). The two lysine ε-amines in the N-terminal peptide were successfully guanidyl-modified under microwave-assisted guanidination reaction in 1 min, as shown in Figure 2(a). After tagging Tm, most N-terminal peptides carried only one metal-element tag; in addition, the efficiency of guanidyl modification on lysine residues is 97.62%, which suggests that N-terminal peptides' guanidination on lysine ε-amines is highly efficient, as shown in Figure 2(b). Moreover, ionization efficiency of peptides that contain lysine in MALDI-source mass spectrometry increases when the peptides are guanidyl-modified [35, 36].

3.3 Optimization of metal-element tagging on N-terminal peptides

Highly efficient N-terminal peptide labeling is crucial to relative quantitation that is based on N-terminal peptides. We took lysozyme and ribonuclease B as models and optimized metal- element-labeling conditions. Lysozyme was reacted with different concentrations of DOTA-NHS-ester dissolved in TEAB buffer at pH 8.5 for 1 h, and then chelated with Tm. As shown in Table 1, when the ratio of lysozyme (μg) to DOTA-NHS-ester was 1∶10, the labeling efficiency of lysozyme N-terminal peptides reached 99.50%. When we added excessive DOTA-NHS-ester, it was hydrolyzed under alkaline conditions and the acid it produced decreased the pH of the solution. Thus the N-terminal primary amines were blocked as these primary amines turned into $-NH_3^+$. Further experiments indicated that when the ratio of ribonucluease B to DOTA-NHS-ester was 1∶10, the

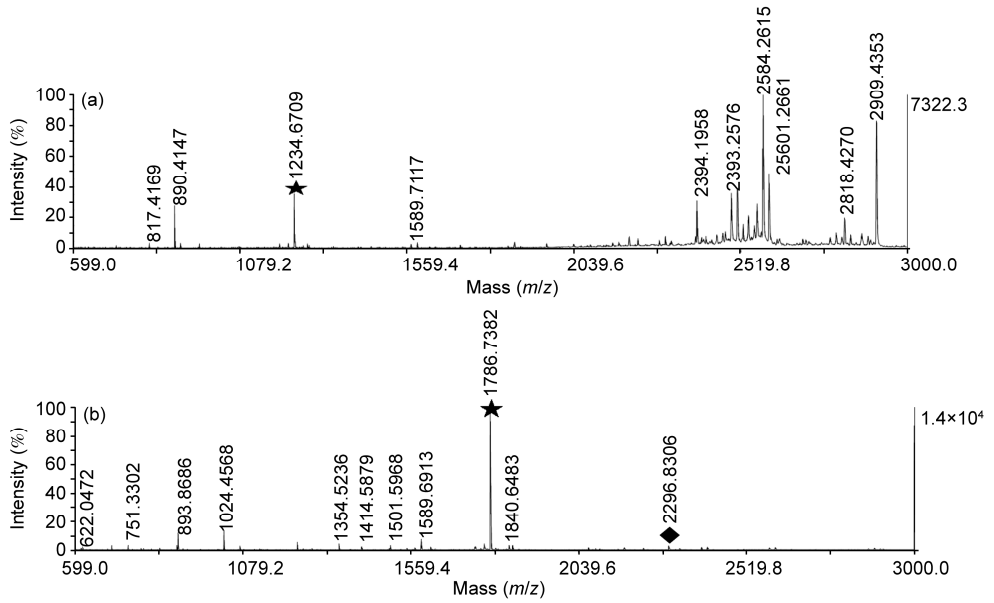

Figure 2 MALDI-TOF spectrum of trypsin digested peptides of guanidyl modified ribonuclease B (a) and guanidyl modified ribonuclease B with Tm tag (b). (a) ★ represents N-terminal peptide peak with both lysine ε-amines guanidyl modified; (b) ★ represents N-terminal peptides with one metal element tag, and ◆ represents N-terminal peptides with two metal element tag.

Table 1 Labeling efficiency of different ratios of lysozyme to DOTA-NHS-ester when reacting with N-terminal peptides

Lysozyme (μg)	DOTA-NHS-ester (nmol)	Labeling efficiency (%)
1	50	69.66
1	30	92.57
1	10	99.50

labeling efficiency reached 99.99%. Thus metal-element-chelated tags could have a high efficiency in labeling N-terminal peptides.

3.4 Enrichment and identification of N-terminal peptides by using DITC-silica beads to remove non-N-terminal peptides

Removal capability of DITC-silica beads

Synthesized peptides (sequence: AFAAGVDLGR) and Tm-labeled synthesized peptides were mixed at a ratio of 10∶1 and were added into 1 mg of DITC-silica beads at room temperature for 1 h. The experiments were repeated with different moles of synthesized peptides. We took 1 μL of supernatant to be analyzed by MALDI-TOF-MS. The experiment was repeated 3 times and the averaged results showed that 91.05% of the unblocked peptides were removed when 2 nmoles of synthesized peptides were added (Figure 3). The removed amount of peptides increased remarkably as the added amount decreased. When 0.25 nmoles of peptides were added, 99.43% of the unblocked peptides were removed. The combination of unblocked peptides with DITC-silica beads was considered to be complete

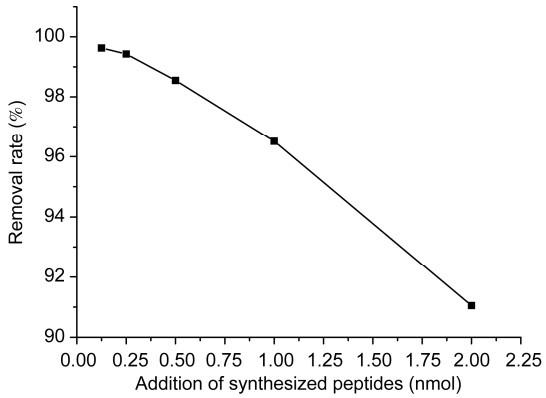

Figure 3 Capability of DITC-silica bead.

if more than 95% of the peptides were removed. The removal capability of 1 mg of DITC-silica beads was approximately 1 nmol when 1 mg of DITC-silica beads was used.

Recovery of N-terminal peptides enriched by DITC-silica beads

In order to measure the recovery of DITC-silica beads, we utilized Tb and Tm to label the synthesized peptides. We took an equal amount of differently labeled peptides to carry out mass spectrometry analysis after the labeling was complete. We first obtained the pre-combined peak area ratio (Tm/Tb). Then we incubated the Tm-labeled peptides with 1 mg of DITC-silica beads; fractionated, dried, and redissolved them; and mixed them with an equal amount of untreated Tb-labeled peptides. The mixture was analyzed by mass spectrometry and the post-combined peak area ratio

(Tm/Tb) was obtained. The ratio (the ratio of post-combined/the ratio of pre-combined) was the recovery. As shown in Table 2, three parallel experiments were conducted and the averaged result was 96.44%. The loss may come from absorption of Eppendorf tubes or unspecific absorption of the materials.

Enrichment and identification of blocked N-terminal peptides in peptide mixture

In order to investigate the specificity of the method, we tested the removal ability of DITC-silica beads added to a peptide mixture made up of 1 μg of trypsin-digested ribonuclease B that was blocked at N-terminal by metal-element tags, 20 ng of synthesized peptides that were blocked at N-terminal by metal-element tags, and 100 μg peptides of trypsin-digested myoglobin that had free amines at N-terminal peptide. We took 1 μg of this peptide mixture to incubate with 1 mg of DITC-silica beads prior to a MALDI-TOF-MS analysis. As shown in Figure 4(a), for the peptide mixture of digested ribonuclease B, digested myoglobin, and synthesized peptides, the N-terminal peptides blocked by metal-element tags had very low signal intensity; in addition, as shown in Figure 4(b), after incubation with DITC-silica beads, free-amine N-terminal peptides had been eliminated, leaving only blocked N-terminal peptides. We can conclude from these results that DITC-silica beads have very high specificity in enriching blocked N-terminal peptides. The Tm-labeled N-terminal peptides of ribonuclease B were further analyzed by tandem mass spectrometry and were matched by using Mascot. As shown in Figure 4(c), these N-terminal peptides were successfully identified.

Table 2 Recovery of DITC-silica beads

Experiments	Recovery (%)
Run 1	93.63
Run 2	94.37
Run 3	98.06
Averaged result	96.44

3.5 Relative quantitation based on N-terminal peptides

We took lysozyme and ribonuclease B as models to develop the relative quantitative method. Lysozyme was guanidyl-modified, labeled by Tb and Ho, respectively, and mixed at the ratio of 1∶1 before trypsin digestion. The mixture was incubated with DITC-silica beads prior to an HPLC-ESI-LTQ-FT MS analysis. The same method was applied to ribonuclease B. The results are shown in Figure 5. N-terminal peptides with both Tb and Ho labeled have almost the same retention time; in addition, the peak area ratio was 0.9 to 1, which was approximately the theoretical ratios. Proteins labeled with Ho and Tb were mixed at the ratios of 10, 5, 2, 1, 0.5, 0.2, 0.1, respectively, and incubated with DITC-

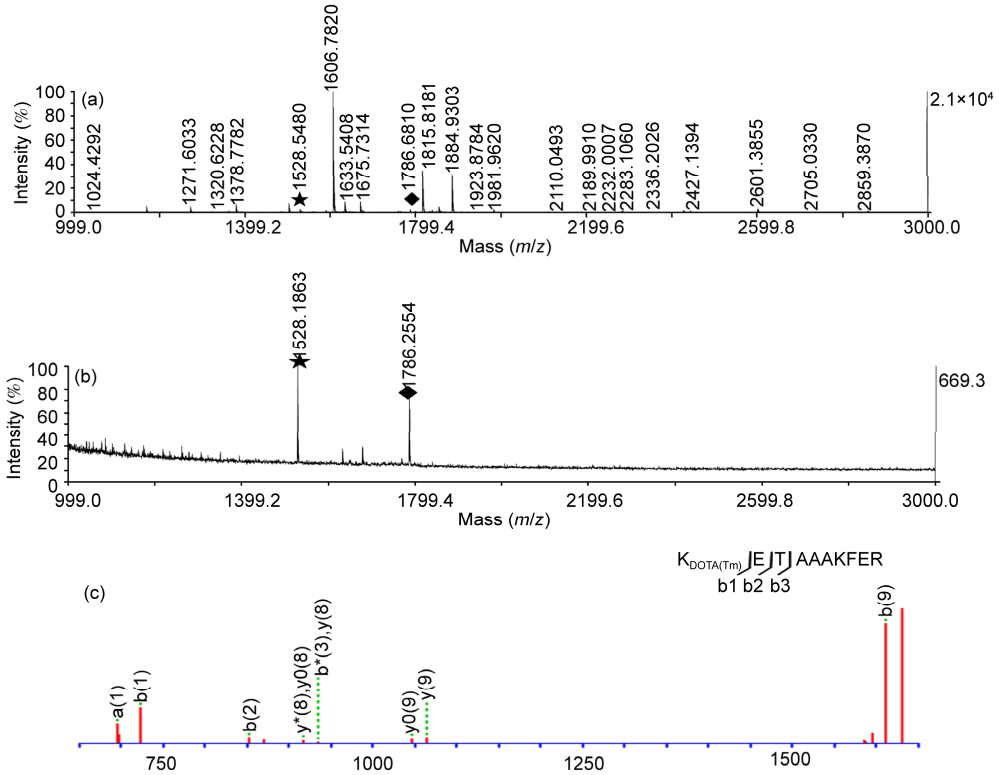

Figure 4 MALDI-TOF MS spectrum of peptide mixture uncombined with DITC-silica beads (a), combined with DITC-silica beads (b), and Mascot matching of Tm-labeled ribonuclease B N-terminal peptides analyzed by MS/MS (c). Note: ★ represents synthesized peptides blocked at N-terminal; ◆ represents Tm-labeled N-terminal peptides of ribonuclease B.

silica beads before mass spectrometry analysis. These experiments were repeated twice. As shown in Figure 6, in the dynamic range of 0.1 to 10, the correlation coefficients (R^2) of the standard curves of lysozyme and ribonuclease B were 0.9994 and 0.9997, respectively. Both of the relative standard deviations (RSD) were less than 5%. These results suggest that the relative quantitation method based on metal tags and N-terminal peptides was of excellent accuracy and reproducibility, with good linearity.

3.6 Identification and relative quantitation of N-terminal peptides of proteins from *T. tengcongensis* incubated at different temperatures

The total protein extract of *T. tengcongensis* incubated at 55 °C was labeled with Tb and the extract of *T. tengcongensis* incubated at 75 °C was labeled with Ho. Both metal-labeled proteins were 1:1 mixed before trypsin digestion. After being digested with trypsin, 15 μg of each of the two peptide mixtures were mixed, and then incubated with DITC-silica beads before an HPLC-ESI-LTQ-Orbitrap-Velos analysis. The data were analyzed qualitatively and quantitatively by PEAKS 6. The identification results are shown in Table S1 and Figure 7(a). A total of 158 N-terminal peptides were identified, among which 86 proteins overlapped between the *T. tengcongensis* proteins that had been incubated at two different temperatures. Further analysis indicated that among the 158 N-terminal peptides, partial acetylations in existed in 24. As shown in Figure 7(b), the N-terminal peptides were identified as retained initiator methionine (66%), methionine cleavage (26%), and proteins with signal peptide cleavage (8%). As shown in Figure 7(c), further examination of the length of the N-terminal peptides indicated that 55% had short sequences (fewer than 6 amino acids). Finally, we analyzed the identified proteins quantitatively: a total of 80 proteins had gotten quantitation information. Compared with proteins of *T. tengcongensis* incubated at 55 °C, in the extract of *T. tengcongensis* incubated at 75 °C, 7 proteins upregulated their expression whereas 16 proteins downregulated according to whether the peak area ratios were more than 2 or less than 0.5, respectively. The function of these 23

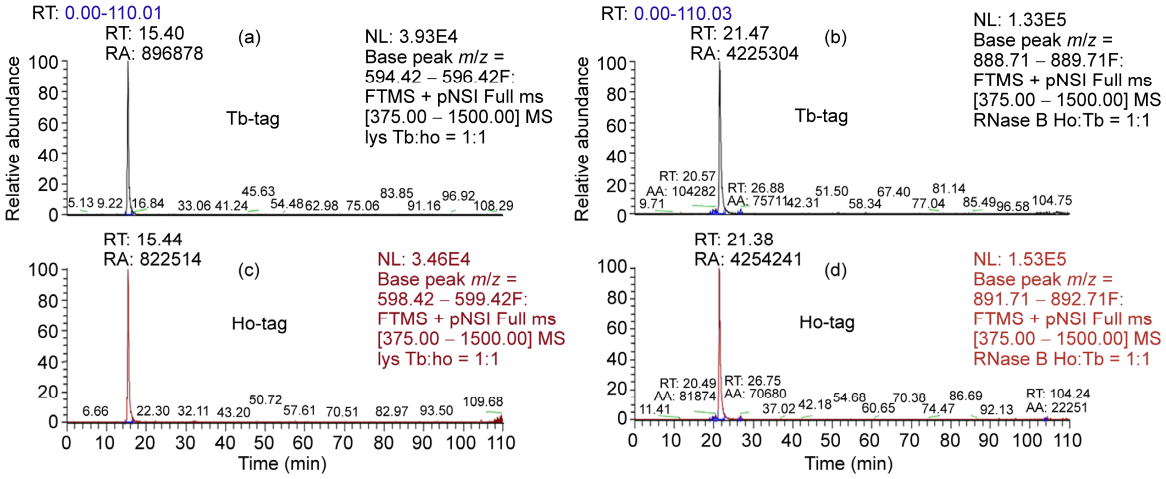

Figure 5 XIC chromatogram of lysozyme ((a), (c)) and ribonuclease B ((b), (d)) N-terminal peptides labeled with Tb and Ho, respectively.

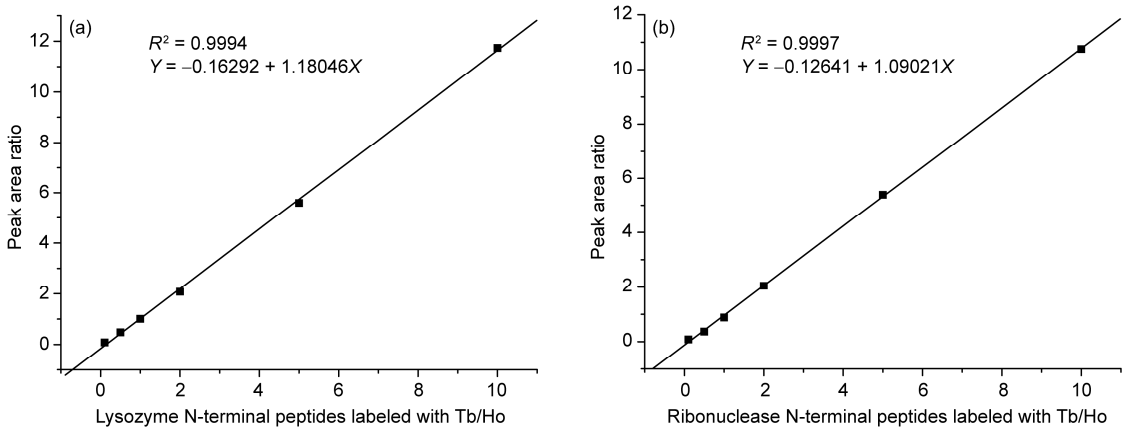

Figure 6 Linear dynamic range curve of lysozyme (a) and ribonuclease (b) N-terminal peptides labeled with Tb and Ho, respectively.

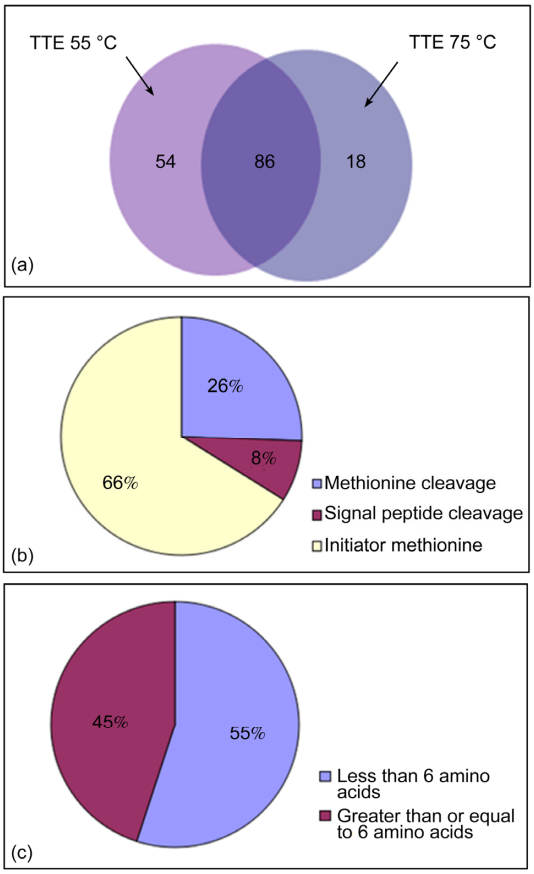

Figure 7 Identified N-terminal peptides of *T. tengcongensis* at two different temperatures (a); modification distribution of N-terminal peptide initiation sites (b); and length distribution of N-terminal peptides (c).

differentially expressed proteins is analyzed at http://string-db.org. Interactions exist among 20 differentially expressed proteins, most of which were involved in the synthesis of *T. tengcongensis*. Three of these 20 proteins were ribosomal proteins and were in charge of the synthesis of proteins. Six were related to energy metabolism, which could provide energy to the synthesis of proteins. Two participated in the synthesis of amino acids for the proteins. These results indicate that metal element tags are competent for the identification and relative quantitation of N-terminal peptides in complex samples.

4 Conclusions

We established a novel method for the identification of N-terminal peptides of proteins and their relative quantification using metal-element-chelated tags coupled with mass spectrometry. The results indicate that metal elements are able to label N-terminal peptides completely at protein level, and that sample complexity can decrease remarkably by selective elimination of proteolytic digested peptides using isothiocyanate-coupled silica beads. This method was applied in the identification and relative quantitation of N-terminal peptides of proteins from *T. tengcongensis* incubated at two different temperatures. We successfully identified the N-terminal peptides of 158 proteins of *T. tengcongensis* incubated at 55 and 75 °C, among which the N-terminal peptides of 24 proteins were partially acetylated. Moreover, metal-element tags with high molecule weight make it convenient for N-terminal peptides consisting of fewer than 6 amino acids to be identified. N-terminal peptides consisting of fewer than 6 amino acids make up 55% of the identified proteins. Finally, we developed a general approach for the relative quantification of proteins based on N-terminal peptides. As we adopted lysozyme and ribonuclease B as model proteins, the correlation coefficients (R^2) of the standard curves for the quantitative method were 0.9994 and 0.9997, respectively, with each concentration ratio ranging from 0.1 to 10, and both relative standard derivations (RSD) were less than 5%. In *T. tengcongensis* at two incubation temperatures, 80 proteins possessed quantitative information. In *T. tengcongensis* incubated at 75 °C, 7 proteins upregulated whereas 16 proteins downregulated, compared with the proteins of *T. tengcongensis* incubated at 55 °C.

This work was supported by the National Basic Research Program of China (2012CB910603, 2010CB912701), the National Key Scientific Instrument Development Program of China (2011YQ09000504, 2012YQ-12004407, 2011YQ06008408), the High-Technology Research and Development Programme of China (2012AA020202), and the National Natural Science Foundation of China (20735005, 21275159, 31100591).

1 Resing KA, Meyer-Arendt K, Mendoza AM. Improving reproducibility and sensitivity in identifying human proteins by shotgun proteomics. Anal Chem, 2004, 76: 3556–3568
2 Cox J, Mann M. Quantitative, high-resolution proteomics for data-driven systems biology. Annu Rev Biochem, 2011, 80: 273–299
3 Mahrus S, Trinidad JC, Barkan DT, Sali A, Burlingame AL, Wells JA. Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell*, 2008, 134: 866–876
4 Van Damme P, Maurer-Stroh S, Plasman K, Van Durme J, Colaert N, Timmerman E, De Bock PJ, Goethals M, Rousseau F, Schymkowitz J, Vandekerckhove J, Gevaert K. Analysis of protein processing by N-terminal proteomics reveals novel species-specific substrate determinants of granzyme B orthologs. *Mol Cell Proteomics*, 2009, 8: 258–272
5 Bachmair A, Finley D, Varshavsky A. *In vivo* half-life of a protein is a function of its amino-terminal residue. *Science*, 1986, 234: 179–186
6 Bonissone S, Gupta N, Romine M, Bradshaw RA, Pevzner PA. N-terminal protein processing: a comparative proteogenomic analysis. *Mol Cell Proteomics*, 2013, 12: 14–28
7 Mann M, Wilm M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem*, 1994, 66: 4390–4399
8 Foyn H, Van Damme P, Stove SI, Glomnes N, Evjenth R, Gevaert K, Arnesen T. Protein N-terminal acetyltransferases act as N-terminal propionyltransferases *in vitro* and *in vivo*. *Mol Cell Proteomics*, 2013, 12: 42–54
9 Wilkins MR, Ou K, Appel RD. Rapid protein identification using

N-terminal "sequence tag" and amino acid analysis. *Biochem Biophys Res Commun*, 1996, 21: 609–613
10 Wilkins MR, Gasteiger E, Tonella L. Protein identification with N and C-terminal sequence tags in proteome projects. *J Mol Biol*, 1998, 278: 599–608
11 Kleifeld O, Doucet A, auf dem Keller U, Prudova A, Schilling O, Kainthan RK, Starr AE, Foster LJ, Kizhakkedathu JN, Overall CM. Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nat Biotechnol*, 2010, 28: 281–288
12 Lange PF, Overall CM. Protein TAILS: when termini tell tales of proteolysis and function. *Curr Opin Chem Biol*, 2013, 17: 73–82
13 Prudova A, auf dem Keller U, Butler GS, Overall CM. Multiplex N-terminome analysis of MMP-2 and MMP-9 substrate degradomes by iTRAQ-TAILS quantitative proteomics. *Mol Cell Proteomics*, 2010, 9: 894–911
14 Jefferson T, Auf dem Keller U, Bellac C, Metz VV, Broder C, Hedrich J, Ohler A, Maier W, Magdolen V, Sterchi E, Bond JS, Jayakumar A, Traupe H, Chalaris A,Rose-John S, Pietrzik CU, Postina R, Overall CM, Becker-Pauly C. The substrate degradome of meprin metalloproteases reveals an unexpected proteolytic link between meprin β and ADAM10. *Cell Mol Life Sci*, 2013, 70: 309–333
15 Auf dem Keller U, Prudova A, Gioia M, Butler GS, Overall CM. A statistics based platform for quantitative N- terminome analysis and identification of protease cleavage products. *Mol Cell Proteomics*, 2010, 9: 912–927
16 Kovanich D, Cappadona S, Raijmakers R, Mohammed S, Scholten A, Heck AJ. Applications of stable isotope dimethyl labeling in quantitative proteomics. *Anal Bioanal Chem,* 2012, 404: 991–1009
17 Guryca V, Lamerz J, Ducret A, Cutler P. Qualitative improvement and quantitative assessment of N-terminomics. *Proteomics,* 2012, 12: 1207–1216
18 Gevaert K, Goethals M, Martens L, Van Damme J, Staes A, Thomas GR, Vandekerckhove J. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nat Biotechnol*, 2003, 21: 566–569
19 Staes A, Impens F, Van Damme P, RuT. tengcongensisns B, Goethals M, Demol H, Timmerman E, Vandekerckhove J, Gevaert K. Selecting protein N-terminal peptides by combined fractional diagonal chromatography. *Nat Protoc*, 2011, 6: 1130–1141
20 McDonald L, Robertson DH, Hurst JL, Beynon RJ. Positional proteomics: Selective recovery and analysis of N-terminal proteolytic peptides. *Nat Methods*, 2005, 2: 955–957
21 Schepmoes AA, Zhang Q, Petritis BO, Qian WJ, Smith RD. N-terminal enrichment: Developing a protocol to detect specific proteolytic fragments. *J Biomol Tech*, 2009, 20: 263–265
22 Zhao L, Zhang Y, Wei J, Cao D, Liu K, Qian X. A rapid isolation and identification method for blocked N-terminal peptides by isothiocyanate-coupled magneticnanoparticles and MS. *Proteomics*, 2009, 9: 4416–4420
23 Mommen GP, van de Waterbeemd B, Meiring HD, Kersten G, Heck AJ, de Jong AP. Unbiased selective isolation of protein N-terminal peptides from complex proteome samples using phosphotagging (PTAG) and TiO(2)-based depletion. *Mol Cell Proteomics*, 2012, 11: 832–842
24 Hsu JL, Chen SH, Li DT. Enhanced a1 fragmentation for dimethylated proteins and its applications for n-terminal identification and comparative protein quantitation. *J Proteome Res*, 2007, 6: 2376–2383
25 Schilling O, Barré O, Huesgen PF, Overall CM. Proteome-wide analysis of protein carboxy termini: C terminomics. *Nat Methods*, 2010, 7: 508–511
26 Kleifeld O, Doucet A, Prudova A, auf dem Keller U, Gioia M, Kizhakkedathu JN, Overall CM. Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. *Nat Protoc*, 2011, 6: 1578–1611
27 Papadioti A, De Bock PJ, Vranakis I, Tselentis Y, Gevaert K, Psaroulaki A, Tsiotis G. Study of the whole cell lysate of two *Coxiella burnetii* strains using N-terminomics. *J Proteome Res*, 2012, 11: 3150–3159
28 Gorman JJ, Shiell BJ. Isolation of carboxyl-termini and blocked amino-termini of viral proteins by high-performance cation exchange chromatography. *J Chromatogr*, 1993, 646 : 193–205
29 Dormeyer W, Mohammed S, Breukelen Bv, Krijgsveld J, Heck AJ. Targeted analysis of protein termini. *J Proteome Res*, 2007, 6: 4634–4645
30 Zhang X, Ye J, Engholm-Keller K, Højrup P. A proteome-scale study on *in vivo* protein Na-acetylation using an optimized method. *Proteomics*, 2011, 11: 81–93
31 Liu H, Zhang Y, Wang J, Wang D, Zhou C, Cai Y, Qian X. Method for quantitative proteomics research by using metal element chelated tags coupled with mass spectrometry. *Anal Chem*, 2006, 78: 6614–6621
32 Wang X, Zhang YJ, Wang XY, Lin HJ, Qin WJ, Qian XH. Multiplex relative quantitation of peptides and proteins using amine-reactive metal element chelated tags coupled with mass spectrometry. *Anal Methods*, 2012, 4: 1629–1632
33 Wang XY, Wang X, Qin WJ, Lin HJ, Wang JF, Wei JY, Zhang YJ, Qian XH. Metal-tag labeling coupled with multiple reaction monitoring-mass spectrometry for absolute quantitation ofproteins. *Analyst*, 2013, 138: 5309–5317
34 Wang XY, Wang X, Zhang YJ, Qian XH. A novel method of fast guanidination in protein analysis. *J Chin Mass Spectr Soc*, 2012, 32: 258–270
35 Brancia FL, Oliver SG, Gaskell SJ. Improved matrix-assisted laser desorption/ionization mass spectrometric analysis of tryptic hydrolysates of proteins following guanidination of lysine-containing peptides. *Rapid Commun Mass Spectrom*, 2000, 14: 2070–2073
36 Beardsley RL, Reilly JP. Optimization of guanidination procedures for MALDI massmapping. *Anal Chem*, 2002, 74: 1884–1890

SCIENCE CHINA
Chemistry

• SUPPORTING INFORMATION •
• SPECIAL TOPIC • Mass Spectrometry Analysis

doi: 10.1007/s11426-013-5049-8

A novel method for identification and relative quantification of N-terminal peptides using metal-element-chelated tags coupled with mass spectrometry

YAN Hui[1,2], HAO FeiRan[2], CAO QiChen[3,2], LI JiaBin[2], LI NanNan[2], TIAN Fang[2], BAI HaiHong[2], REN XiaoJun[2], LI XianYu[2], ZHANG YangJun[1,2*] & QIAN XiaoHong[2]

[1]*Anhui Medical University, Hefei 230032, China*
[2]*State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing 102206, China*
[3]*Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang 110032, China*

Received September 1, 2013; accepted November 6, 2013; published online February 24, 2014

Table S1 Identified N-terminal peptides of *T. tengcongensis* at two different temperatures

Number	Proteins	N-terminal peptides
1	gi|20808157	AEKGYGFIER
2	gi|20808207	AIAFITGATDGIGR
3	gi|20806821	AIETFKAVSR
4	gi|20807101	AISKSEVEYIAKLAR
5	gi|20807076	AKQIKYGEEAR
6	gi|20808405	AKYVMALDQGTTSSR
7	gi|20806820	ANLDMKDLER
8	gi|20808968	ANSTSSYLWR
9	gi|20806899	AQKIEMKVPIVEMDGDEMTR
10	gi|20808635	ARIAGVDLPR
11	gi|20808879	ARNIVEFR
12	gi|20807807	ASSKAAINLQDIFLNQVR
13	gi|20808625	ATIQYYGTGR
14	gi|20806551	ATYEDFLKLDIR
15	gi|20806841	AYSVEIDR
16	gi|20806985	GATIKDVAR
17	gi|20808966	GEDGKLYIVR
18	gi|20807683	GEQLGGLKR
19	gi|20808657	GIKAYKPTSPGR
20	gi|20807128	GIRPEEISSVLAEK
21	gi|20808783	GITVTKAEEAVDYFLNR
22	gi|20808654	GQKVHPYGLR

*Corresponding author (email: 13683167093@163.com)

(*Table S1 Continued*)

Number	Proteins	N-terminal peptides
23	gi|20808573	GRLFGTDGVR
24	gi|20808607	GSEVLKLLR
25	gi|20807315	GVNVAIVGATGMVGR
26	gi|20808661	ILDQSAQKIVETAKR
27	gi|20809104	KAGEVVNVSDGYGR
28	gi|20808832	MDAKEIVLK
29	gi|20808638	MDDLQLGQVVR
30	gi|20808768	MDGLMTAVVK
31	gi|20808520	MDIEIIR
32	gi|20807131	MDKTFHLEVLTPYR
33	gi|20807904	MDLIKAVESEQLKKDITPFNVGDTVR
34	gi|20808165	MDLNSYIDSMRDEIIK
35	gi|20808292	MDNIVDAR
36	gi|20807090	MDSRPIGVFDSGVGGLTVLKR
37	gi|20807591	MDWDYVEDAMKR
38	gi|20808087	MEEEKMVVLR
39	gi|20808186	MEEITALLK
40	gi|20806659	MEEIVIIR
41	gi|20807740	MEENVKNFENQELGSVK
42	gi|20806574	MEFMNVSLPEEIAR
43	gi|20807143	MEIELINLKR
44	gi|20807744	MEKEFNFEEDLKR
45	gi|20808447	MEKEMAGKVEVEAK
46	gi|20809028	MEKFVIR
47	gi|20807697	MEKLGMNEIR
48	gi|20808271	MEKMGVLEGR
49	gi|20807781	MELWFTEHQDENLR
50	gi|20806617	MENLHTR
51	gi|20807127	MEQIIAKR
52	gi|20807759	MEQKEPTVSTR
53	gi|20806793	MEVKELVLK
54	gi|20807817	MEVLKVSAK
55	gi|20807914	MEVLKVSAK
56	gi|20807731	MEYLNER
57	gi|20806970	MEYRDYEKLIEIAKEAR
58	gi|20808552	MFDEIRPTR
59	gi|20808632	MFEIIEPKIEIVEQSEDGR
60	gi|20807181	MFKILEKR
61	gi|20808519	MFVDFHCDTLYR
62	gi|20807976	MFYQPVLETSR
63	gi|20807978	MFYQPVLETSR
64	gi|20808435	MHGTIIQLEEKPLR
65	gi|20807730	MIAAGDFR
66	gi|20807821	MIEKQKALEMAISQIER
67	gi|20807650	MIEKVINR
68	gi|20807742	MIIDGKEISNR
69	gi|20808644	MIIKPSR
70	gi|20807955	MIKSMTGYGR
71	gi|20808650	MIRPQTR
72	gi|20807855	MISAQAVKELR

(*Table S1 Continued*)

Number	Proteins	N-terminal peptides
73	gi|20807944	MIVAALTDIGNFR
74	gi|20807409	MKDIAIALLK(+542.08)ER
75	gi|20808019	MKEIKTMEFR
76	gi|20808237	MKEVTIEIK
77	gi|20808922	MKITDVR
78	gi|20807130	MKKGYITQVIGPVVDIR
79	gi|20808306	MKKLQEIAEEISR
80	gi|20806668	MKPGIHPTYYHDAVVK
81	gi|20808755	MKSDIEIAQEAKMLHIR
82	gi|20809107	MKSDLEIAKNIR
83	gi|20808626	MKSYMAKPSDVER
84	gi|20808233	MKTIGILTSGGDAPGMNAAIR
85	gi|20808032	MKTIINTDAAPK
86	gi|20807373	MKTLEELER
87	gi|20808301	MKVAVSSQGKTLESHVDTR
88	gi|20807088	MKVLVVGGGGR
89	gi|20807502	MKVMIVR
90	gi|20806566	MLDIKFIR
91	gi|20807747	MLEQINSPYDLK
92	gi|20806635	MLKSTGIVR
93	gi|20806951	MLLKELTELLGASGDEKEVR
94	gi|20808483	MLNLEKSATFR
95	gi|20808694	MLQYDKDVVLSSR
96	gi|20808025	MLTPMDIHNKEFR.
97	gi|20806666	MLVTGIEILEK
98	gi|20808603	MNALEFAIK
99	gi|20806553	MNEEMDKVIPVDIEDEMR
100	gi|20807175	MNIPKETLMR
101	gi|20808810	MNKAELVAK
102	gi|20807452	MNVLIDNR
103	gi|20808100	MNYDVIVVGGGPGGYTAAIR
104	gi|20808386	MQAITLEAVKR
105	gi|20807515	MQEPQNINELIR
106	gi|20808226	MQKAVEFTYNR
107	gi|20808479	MQMVDINLLKEHPR
108	gi|20808082	MQQVYDEKVR
109	gi|20808571	MQSVYINTPLR
110	gi|20807050	MREAVIVSAVR
111	gi|20806992	MREMTKKNLLDAYAGESQAHMR
112	gi|20807814	MREVYETPLVTR
113	gi|20806687	MRIGVPKEIKTAESR
114	gi|20808268	MRIYFER
115	gi|20808642	MRLHDLKPAEGSTK
116	gi|20807075	MRLKPLGDR
117	gi|20807218	MRLNENMFR
118	gi|20809110	MRSYETMYVLSPDLNEEER
119	gi|20807972	MRVKAEIMDEKAIDR
120	gi|20808669	MSKEEILQAIK
121	gi|20808452	MSLNKVMLVG
122	gi|20808178	MSSIIDIYAR
123	gi|20808875	MTDDFVPEITSPLR

(*Table S1 Continued*)

Number	Proteins	N-terminal peptides
124	gi|20808441	MTEAWQDIYASR
125	gi|20807649	MTLDDIKEMIR
126	gi|20808477	MTLKEELAFLR
127	gi|20807893	MTLLEKTR
128	gi|20807994	MTQLEYALSGIITK
129	gi|20807457	MVDEQLEKIDMIVER
130	gi|20808382	MVDYNLLR
131	gi|20806763	MVENEVINLLELR
132	gi|20808327	MVESSGAKGVLR
133	gi|20808004	MVEVGKKAPDFVLPDADGR
134	gi|20807394	MYAIIETGGKQYMVR
135	gi|20807684	MYEESREDLLR
136	gi|20806676	MYENVFVIDHPLIQHKISLIR
137	gi|20808842	MYKVVLLR
138	gi|20807031	MYLLKGGR
139	gi|20807012	MYQDKVLVCKDCGR
140	gi|20808302	PRGDGTGPLGLGPR
141	gi|20808665	PTVNQLVR
142	gi|20807776	SDHPIDALMK
143	gi|20807853	SDYLKDSEER
144	gi|20808171	SIHIGAKEGDIAETVLLPGDPLR
145	gi|20808440	SKALNDRLDDIIYEKVK
146	gi|20807875	SKILIVDDAAFMR
147	gi|20807119	SLVPIVVEQTNR
148	gi|20806612	SNDVKPTYTLDER
149	gi|20807787	SRVLMQGNEAVVEGAIR
150	gi|20807791	STESLNPLVIAQK
151	gi|20808219	SVEAVRDFFR
152	gi|20809023	SVLEVLKER
153	gi|20807926	SVMDNIIQK
154	gi|20807095	SWGATSSALAK
155	gi|20807461	TKKYVYFFNEGDASMR
156	gi|20809108	TTANTNETAAAAAAKNRR
157	gi|20808646	VMTDPIADMLTR
158	gi|20807103	YEAVIGLEVHAELLTDSK