Front. Chem. China 2007, 2(4): 447–453
DOI 10.1007/s11458-007-0062-0

RESEARCH ARTICLE

MENG Qingfang, ZHANG Yangjun, CAI Yun, QIAN Xiaohong

Affinity labeling coupled with matrix assistant laser desorption tandem time of flight mass spectrometry for quantitative proteomics research

© Higher Education Press and Springer-Verlag 2007

Abstract A relative quantitative method for differential proteomics by cleavable isotope-coded affinity tag (cICAT) and matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry (MALDI-TOF-MS) was established. The accuracy and reproducibility of the method were evaluated by bovine serum albumin (BSA) digest as having a relative standard deviation of less than 30% and good reproducibility. The dynamic range was also evaluated by analyzing two mixtures of several standard proteins with different concentration. The experimental results showed that in the dynamic range of 1:30, the quantitation error of the method was less than 30%. Although the quantitation error becomes very large when used beyond this range, it does not affect the derivation of information on the differential proteins. All the work provides an alternative method for differential proteomics analysis in biological samples from different origins.

Keywords isotope coded affinity tag, matrix assistant laser desorption ionization tandem time of flight mass spectrometry, capillary reversed-phase liquid chromatography

1 Introduction

Quantitative proteomics has become the new hot spot in proteomics research recently. The strategy of quantitative proteomics mainly includes two methods: one is Fluorescence 2-D Diffrence Gel Electrophoresis (DIGE) method based on two-dimensional gel electrophoresis (2-DE) and mass spectrometry (MS), and the other is the isotope-coded affinity tag (ICAT) method [1], which combines the chemical label with multi-dimensional liquid chromatography and tandem MS analysis. The principle of the ICAT technology is to label different original proteins using reagents with identical chemical properties but different molecular weight (usually the isotope tags including light and heavy isotope atoms in two different reagents). After the digestion and multi-dimensional liquid chromatograph separation, the peak intensity of the labeled isotope tag peptide acquired from the MS is compared, and then the abundance of each protein producing the corresponding peptides is gained. Finally, every labeled isotope tag peptide with higher intensity in each pair of peaks is further analyzed by tandem MS/MS to validate the corresponding proteins. The current research is mainly focused on the strategy of multi-dimensional liquid separation interfaced to electrospray ionization mass spectrometry (ESI-MS) [2–4]. Because the method based on the matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is concerned with sample separation, the crystallization on the sample target and the different peptides ionization and the poor reproducibility of the experiment results, the application of the method is limited. However, the combination of the liquid chromatography and matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry (MALDI-TOF-TOF) has many advantages. First, the high sensitivity and resolution of MALDI-TOF-TOF MS can improve the reliability of the experimental data. Second, no analytical time limitation of the liquid chromatographic (LC) separation coupled with the mass spectrometry analysis exists. Third, the sample crystallized with matrix on a sample plate can be analyzed repeatedly. In addition, the result-dependent analysis mode (the selection of the precursor for MSMS dependent on the search result of the MS spectra) can be carried out on MALDI-TOF-TOF MS which enhances the analytical efficiency. Therefore, we try to establish a method based on the cleavable isotope-coded affinity tag (cICAT) (the second generation ICAT reagents) and MALDI-TOF-TOF-MS to do this research on differential proteomics. The key point of the strategy is couples with the use of the LC and MALDI-MS.

Translated from *Chinese Journal of Analytical Chemistry*, 2006, 34(7): 899–904 [译自: 分析化学]

MENG Qingfang, ZHANG Yangjun, CAI Yun, QIAN Xiaohong (✉)
Beijing Institute of Radiation Medicine, 27 Taiping Road, Beijing 100850, China
Beijing Proteome Research Center, 33 Life Park Road, Zhongguancun, Beijing 102206, China
E-mail: qianxhi@yahoo.com.cn

An automatic collector, Probot, couple was used to automatically collect LC fractions and put it directly onto a sample target in terms of time interval. Compared with the conventionally used MDLC-ESI-MS method, another off-line LC-MALDI MS method was established. According to the ABI customer instruction for the use of cICAT, the technical route for protein labelling, separation, qualitative and quantitative analyses by MS was established with bovine serum albumin (BSA). Because the reliability of the data is a serious problem in proteomics [5], we have conducted the three parallel experiments to test the accuracy and reproducibility of the technical route, and used five different standard proteins to test the dynamic range of the differential proteins and standard deviation. This formed the foundation for the analysis of real biological samples.

2 Materials and methods

2.1 Instruments and reagents

The micro- high performance liquid chromatographic (HPLC) (Dionex, US) includes an auto-injecter (Famos module), a switcher module (Switchers) that includes a pump and two ten-port valves, a detector module (Ultimate) that includes a pump with a low pressure mixer for gradient elution, a micro splitter, a UV detector and an automatic fraction collecting and spotting system (Probot module). The chromatographic column used is 180 μm i.d × 15 cm, C18, 5 μm, 300 Å (Vydac), the precolumn is 320 μm i.d × 5 mm, C18, 5 μm, 100 Å (Pepmap). All the experiments were conducted in MALDI-TOF-TOF MS (Applied Biosystems, US). The laser is Nd:Yag generator, and the software GPS 2.0. is used for data analysis.

The standard proteins, such as BSA, Apo-transferrin, beta-lactoglobulin, alcohol dehydrogenase I, casein, and asialofetuin A were purchased from Sigma Corporation (US). The cICAT reagent, the sequenced trypsin, denatureing buffer, reducing buffer, cleaving buffer, strong cation-exchange (SCX) cartridge, and AV cartridge were purchased from Applied Biosystems Corporation (US); SARTORIUS BP211D analytical balance is the product of Sartorius Company (Germany); the super-pure water was prepared by the Milli-Q system (Millipore Company, US). The acetonitrile (HPLC grade) was purchased from J. T. Baker Company (US) and the other reagents are HPLC grade.

2.2 BSA experimental method

2.2.1 The protein labeling and digestion

Weigh two aliquots of 25 μg BSA and dissolve them in an 80 μL denaturing buffer (0.1% SDS-Tris buffer), then add 2 μL 50 mM tris (2-carboxyethyl) phosphine (TCEP) and incubate them in a boiling bath for 10 min to break the disulfide bonds. Use 20 μL acetonitrile to dissolve two vials of cICAT reagents, one is for the labeling of the light cICAT reagent and the other is for labeling the heavy cICAT reagent. Incubate the two vials at 37°C for 2 h. Mix the two labeled BSA solutions, then add 10 ug trypsin and incubate at 37°C overnight.

2.2.2 Purification of BSA digest by strong cation exchange, avidin affinity chromatograph and biotin cleavage

The solution of 10 mM potassium dihydrogen phosphate/25% acetonitrile (pH 3.0) is first used to equilibrate the strong cation exchange column and the same solution was also used to adjust the BSA solution to pH 3.0. Then, the sample was slowly injected into the equilibrated column and washed using the same solution. Finally, the solution of 10 mM potassium dihydrogen phosphate/25% acetonitrile/350 mM potassium chloride (pH 3.0) was employed to elute the column. The elution from the strong cation exchange column was adjusted to about pH 8.0 by 2 × phosphate-buffered saline (PBS) buffer (pH 7.2). The solution of 30% acetonitril/0.4% TFA was used to activate the avidin affinity column and the 2 × PBS buffer (pH 7.2) to equilibrate the column. The sample was slowly injected into the column and washed by using the 1 × PBS to desalt. Then the solution of 50 mM ammonia hydrogen carbonate/20% methonal (pH 8.3) was used to wash the unbinded peptides. Exchange of the buffer was made using water. Finally the peptides were eluted by injecting the solution of 30% acetonitril/0.4%TFA.

The elution from avidin affinity column was lyophilizated and then 90 μL of 95%TFA was added and incubated at 37°C for 2 h to cleave the biotin. Finally the sample solution was lyophilizated again.

2.2.3 Micro-scale capillary reversed-phase chromatography and online auto-collecting and -spotting

The solution A of the reversed-phase chromatography is 0.1% trifluoroacetic acid (TFA)/5% acetonitrile (ACN) and solution B is 0.1% TFA/95% ACN. The flow rate is 1 μL/min after being split. The wavelength of the UV detector is set at 214 nm. The nonlinear gradient is: 100% A, 10 min; 100% A–30% B, 6 min; 30% B–60% B, 60 min; 60% B–100% B, 10 min; 100% B, 10 min; 100% B–100% A, 10 min; 100% A, 20 min. When the absorption peaks appear (approximately 30 min), the auto-targeting system begins. Along with the flow of elution, another flow of 0.1% TFA/50% ACN was added consisting of 0.3 μL/min. Probot automatically collected and spotted the eluted fractions on the target plate for 1 min on each well. After air drying, the matrix (5 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA)/0.3% TFA/50% ACN) was spotted on each dried spot and the sample plate was waited to dry to be analyzed by MS.

2.2.4 Experimental conditions for MS analysis

The MS data was acquired by the 1 kv-MS positive reflective mode, accelerative voltage was 20 kV, mass range was 600–4000 m/z and laser energy was 4000. Each mass spectrum

was accumulated 1000 times. Before the analysis, Myoglobin digestion was used to calibrate the mass accuracy to 0.1 Da tolerance, relative standard deviation was 50 ppm. After the MS scan, the GPS software was used to find the peak pair and select the peak with higher abundance to do the MSMS analysis. The identified peptides were selected for comparison. The detail results were given by the GPS software. The MSMS data was acquired by the 1 kV-MS positive reflective mode. Relative resolution was set at 50 when the parent ions were selected. Laser energy was 5000, and each MSMS spectrum was accumulated 3000 times.

2.3 Preparation of the mixture of five standard proteins

Weigh the standard proteins (BSA, apo-transferrin, beta-lactoglobulin, Alcohol dehydrogenase I, casein, or asialofetuinA) in six PE tubes, respectively, and add super-pure water to the concentration of 1 mg/mL when preparing protein solution. According to the composition of the test sample, the protein mixture was prepared and lyphollizated until the volume was reduced to 20 μL. Then 60 μL denaturing buffer (0.1% SDS-Tris buffer) was added, and the rest of the steps are the same as in the previous BSA experiment. In the MS analysis, with regard to peak pair, 16 peaks were selected to go into CAD from high abundance to low. In addition, six single peaks were selected for tandem MS/MS analysis.

3 Results and discussion

3.1 Labeling efficiency of protein and the side reaction

In this experiment, the relationship between the required amount of the cICAT reagent and the label efficiency with the same amount of protein used was first investigated (Fig. 1).

From Fig. 1, we can see that with the 25 μg BSA, 1/2,1/4, 1/8 vial of cICAT reagent is not enough. To obtain the high labeling efficiency, at least one vial of cICAT regent is needed. So the amount of cICAT regent was adopted in the following experiments. Because BSA is a cysteine-rich protein (containing 35 cysteines) and most proteins only have six cysteines, one vial of cICAT reagent can label 100 ug ordinary proteins, but it can only label 25 μg BSA.

In the labeling reaction, some side reactions related to cICAT reagent such as the reaction between the cICAT reagent and TCEP or Tris were found. As shown in Fig. 2, peak 1937.18 is the Tris labeled with two light cICAT reagent, peak 1955.24 is the Tris labeled with two heavy cICAT reagent. Peak 2066.19 is the TCEP labeled with two light cICAT reagent and peak 2084.26 is the TCEP labeled with two heavy cICAT reagent. The reason for these side reactions is that the stronger electron absorption induced effect of iodine exists in the cICAT reagent (iodine has higher electron gravity than carbon), which causes the rest of the reagent to be charged positively. The sulfydryls in the cysteines are the strong nucleophilic group so the labeling reaction is the alkylation reaction between the sulfydryl in the reduced cysteine with the cICAT reagent. The chemical composition of TCEP and Tris also have the strong nucleophilic group-hydroxy group, so they can react easily with cICAT reagent like the sulfydryls in cysteines. From Fig. 2, we can see that the weak peaks 1946.21 and 2075.22 are the peaks of Tris labeled with one light and one heavy cICAT reagent and possibly TCEP labeled with one light and one heavy cICAT reagent. This phenomenon is possible because of the non-sufficiently labeled Tris and TCEP continued to be labeled by the L or H cICAT reagent. However, we did not find the three hydroxy groups labeled peaks at the same time possibly because of the stereospecific blockade. For the side reaction in the labeling process, the following purification and the desalting is necessary.

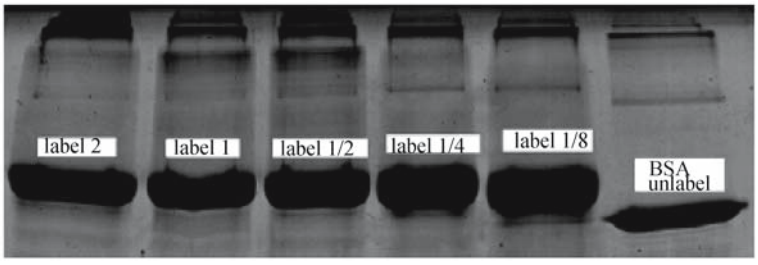

Fig. 1 The relationship between cICAT consumption and labeling efficiency

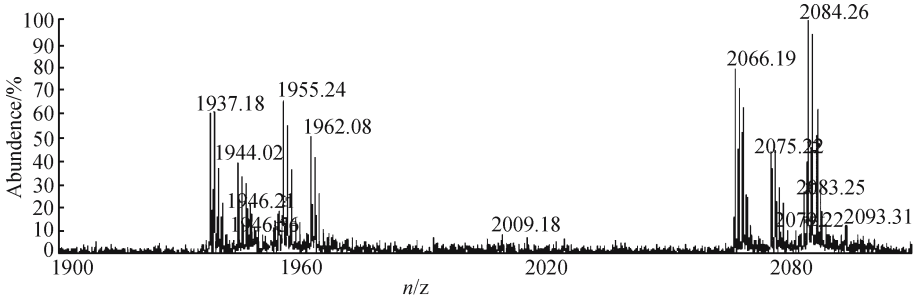

Fig. 2 The nonspecific reaction of cICAT reagent: TCEP and Tris labeled by cICAT

In addition, in the general digestion of proteins, the ratio of the protein amount versus the enzyme is about 20–50:1, in this experiment, because of the presence of the denaturing reagent SDS, the amount of the trypsin was increased. Even so, a few missed cleavages were still found possibly because the high stereospecific blockade of the labeled protein influences the binding of the trypsin with the protein.

3.2 Separation of the digested peptides by reversed-phase liquid chromatography

Pump B was first used to apply the sample in the precolumn, then the ten-port valve was switched to pump A to back flush the precolumn in order to make the sample elute into the analytical column which can realize online desalting. In the linear gradient elution, severe gradient elution delay was found if solution B started to run below 30% and the chromatographic peaks appeared at 30–50 min. Thus, solution B was set at 30% as the gradient elution started. The slope of the linear gradient elution was controlled below 1%/min, which enabled acquiring of a good resolution. Three parallel experiments were performed and Fig. 3 shows the first chromatogram. In Fig. 3, through MS analysis, the chromatographic peaks after 60 min are principally biotin and it was found that peptide peaks were rare. This indicates that the biotin has stronger hydrophobicity than peptides.

Experimental conditions: column, 180 μm I.d.×15 cm, C18, 5 μm, 300 Å; flow rate, 1 μL/min; wave length, 214 nm; gradient elution: 100%A, 10 min: 100% A–30% B, 6 min; 30% B–60% B, 60 min, 60% B–100% B, 10 min; 100% B, 10 min; 100% B–100% A, 10 min; 100% A, 20 min. fraction collected every 1.0 min on MALDI plate.

3.3 MALDI-TOF-TOF MS analysis

Under the conditions for MS analysis mentioned previously, the MS data were first acquired. Then GPS software was used with the setting of search parameters, and finally several precursor ions were selected. Figure 4 shows one MS spectrum of the fraction at 54 min of reversed-phase chromatography. Because of the purification, by using affinity chromatograph, most peptides containing no cysteines were removed. Figure 4 only shows the cysteine-containing peptides labeled with light or heavy cICAT reagent. As shown in Fig. 4, the peak pair of 1589.9 and 1598.9 are the peptides labeled with light and heavy cICAT reagent, respectively. Figure 5 is the MSMS spectrum of the parent ion of 1598.9 in Fig. 4. Figure 5 shows that the b and y serial ions are continuous and complete. With the assistance of Mascot searching against the Swissprot 5.0 database, the matched peptide is SLHTLFGDELCK labeled by a light cICAT reagent, and this peptide is from the protein BSA. The protein identification is based on the criterion of the confidence level of more than 95% from Mascot. In the tryptic digest of BSA, there are 25 cysteine-containing peptides and 17 of them were obtained, including 11 peptides containing one cysteine, four peptides

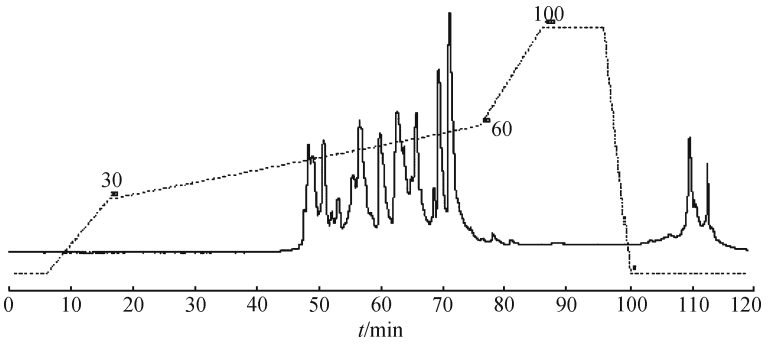

Fig. 3 The capillary reversed-phase liquid chromatography for BSA digests eluted from avidin column

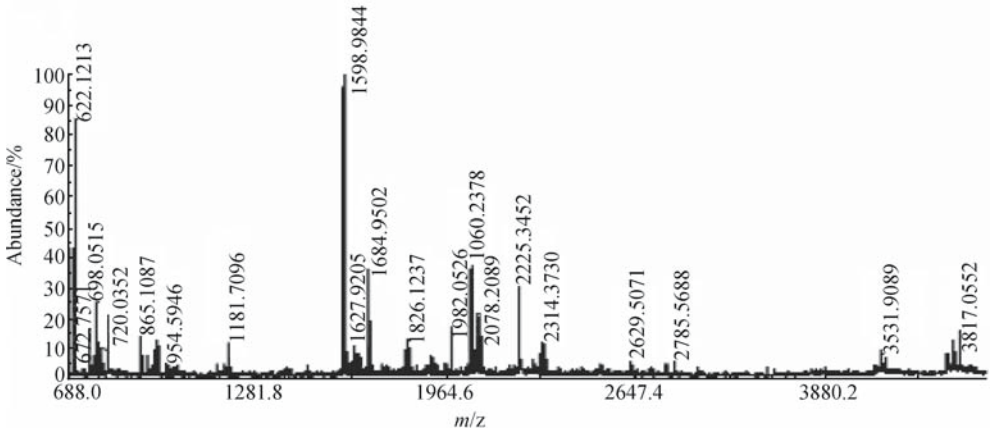

Fig. 4 Mass spectrum of the fraction at 54 min by capillary reversed-phase liquid chromatography

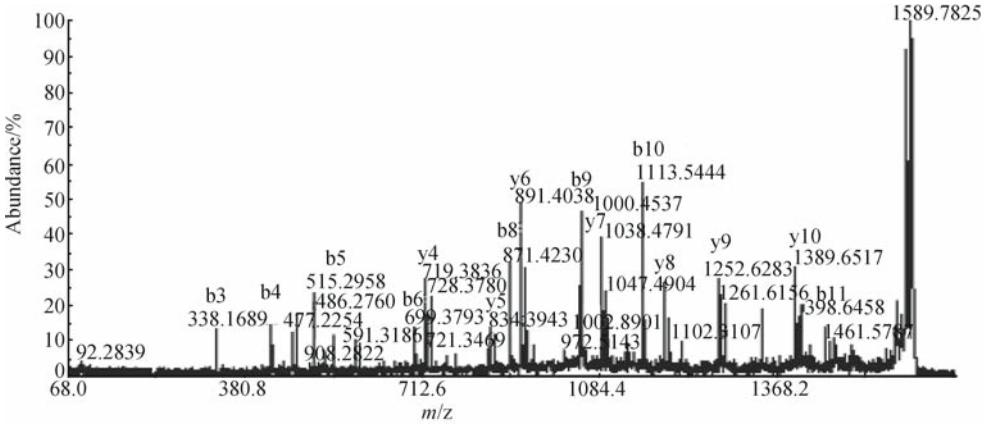

Fig. 5 MSMS spectrum for the peak 1598.98 in Fig. 4

containing two cysteines, and two peptides containing three cysteines.

3.4 Investigation of the accuracy and reproducibility

Due to the presence of the "sweet cake" effect, the challenge of the proteomics quantitative method by using MALDI-TOF-TOF MS is data reproducibility. Therefore, BSA was used as the model protein and three parallel experiments were performed. Three sets of repeated data from the same experiments were acquired, and the accuracy and reproducibility of the data were analyzed. The accuracy is measured by the deviation of the measured peak area ratios of peptides labeled with heavy versus the light cICAT reagent from the theoretically designed ratios.

The reproducibility is measured by that of the final results of the three experiments, the relative standard deviation of H/L value of the same peptides in the different wells on one sample plate, the relative standard deviation of H/L value of the different peptides originated from one protein in one well on one sample plate and the relative standard deviation of H/L value of the different peptides that originated from one protein in different wells on one sample plate.

All the designed ratios for the three parallel experiments are 1. The measured ratios are 0.874, 0.921 and 0.970 with the relative errors being 12.6%, 7.9% and 3.0% respectively. The relative standard deviation of the accuracy of the three experiments is 25.5%.

When data reproducibility was investigated, the data with biggest relative error from one of the three experiments was chosen. In order to measure the reproducibility of the same peptides in the different wells, peak integral areas for peptides labeled with heavy or light cICAT reagent were measured and H/L values were obtained. Relative standard deviation for all the ratios was calculated, with 30.60% as the greatest value. The relative standard deviations for the different peptides that originated from one protein in each well on one sample plate could be calculated with the same method above. The relative deviation was 31.99%. The relative standard deviation for the different peptides that originated from one protein in different wells on one sample plate obtained was 29.99%. The theoretical basis for the data comparison is that the moles of the peptides digested from the same protein are identical, so although these peptides labeled with heavy or light cICAT reagent were spotted on either one well or different wells, they should be equal in terms of their amount. From the above results, the relative standard deviation of the method is 30%, which shows that when comparing differential abundance of proteins that originated from different physiological states, the evaluation criterion should at least be greater than 30%.

3.5 Investigation of the dynamic range

Five proteins (BSA, Apo-transferrin, beta-lactglobulin, alcohol dehydrogenase I, casein) with significant differences in physical and chemical properties (isoelectric point molecular weight and et al., refer to (http://au.expasy.org) were selected. The designed ratio of H/L in the small range was 1:10. The sample composition and the results are shown in Table 1. With BSA, Apo-transferrin, beta- lactglobulin, and alcohol dehydrogenase I, the values of the measured H/L were close to the designed ones, and there was room for in the application of the method. Moreover, the relative standard deviation for the different peptides from each protein is less than 30%. This suggests that when the dynamic range of the different proteins in a sample is between 1:1 and 1:10, the method can be used to measure the proteins' relative amounts accurately.

In order to investigate the dynamic range of the method, we extend the dynamic range of the different proteins from 1:10 to 1:100. The sample composition and the results are shown in the Table 2. In protein identification, the five cysteine-containing proteins are identified by their scores in the database against all species. Among them, BSA was identified with 22 cysteine-containing peptides that accounts for 88% of the all theoretically cysteine-containing peptides from BSA. Transferrin was identified with 18 cysteine-containing peptides that account for 60% of the all theoretically cysteine-containing peptides. This shows that the large dynamic range

Table 1 Sample preparation of small dynamic range and the experimental results

Protein name	Sample 1 (labeled L) consumption /μg	Sample 2 (labeled H) consumption /μg	Dynamic range (H/L)	Designed H/L ratio	Observed H/L ratio	Relative standard deviation of the identified peptides /%	Error /%
BSA	5	40	8:1	8	8.699	30	8.74
Apo-transferrin	35	5	1:7	0.143	0.149	15.9	5.59
Beta-lactoglobulin	10	35	3.5:1	3.5	3.181	24.9	9.11
Alcohol dehydrogenase I	10	10	1:1	1	0.895	0	10.5
Casein	40	10	1:4	0.25	0	/	

Table 2 Sample preparation of large dynamic range and the experimental results

Protein name	Sample 1 (labeled L) consumption /μg	Sample 2 (labeled H) consumption /μg	Dynamic range (H/L)	Designed H/L ratio	Observed H/L ratio	Relative standard deviation of the identified peptides /%	Error /%
BSA	0.5	25	50:1	50	24.892	77	50.2
Apo-transferrin	50	0.5	1:100	0.01	0.024	76	140
Beta-lactoglobulin	5	50	10:1	10	8.819	23	11.8
Alcohol dehydrogenase I	30	1	1:30	0.033	0.039	0	18.2
Asialofetuin A	10	20	2:1	2	2.192	0	9.6

of different proteins cannot affect protein identification if their amounts are above the MS detection limit. As far as the quantitation of the protein is concerned, as shown in Table 2, the small dynamic range of asialofetuin A (1:2) and beta-lactglobulin (1:10) and even that of the alcohol dehydrogenase I (1:30) have been also quantified accurately and the relative error is less than 10%. And the relative standard deviation of H/L values for different peptides from the same protein are all below the value of 30%. However, with BSA and the transferring with large dynamic range, the relative errors of the measured versus the theoretical H/L values are 50% and 140%, respectively. And the relative standard deviations of H/L values for different peptides from the same protein are 70%. The accuracy of the quantification method is not satisfactory. The relationship of the quantification error and the dynamic range of the different proteins were plotted as shown in Fig. 6. From Fig. 6, when the dynamic range is between 1:1 and 1:30, the method can be used to quantitate differential proteins accurately and has a relatively small error. When the dynamic range increased, the relative error also increased sharply. That shows that the dynamic range of the method is from 1:1 to 1:30. When the dynamic range is out of this range, and although the relative error also becomes bigger, the identification of differential proteins is not hindered. Therefore, this method is applicable in the field of comparable proteomics.

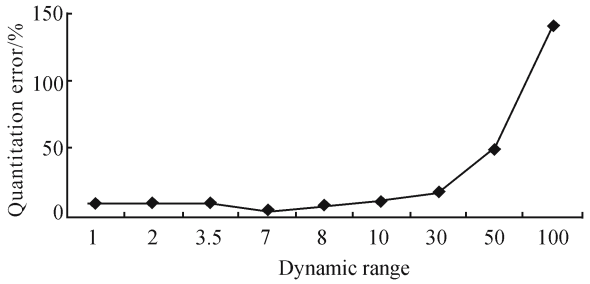

Fig. 6 The relationship between the quantitative error and dynamic range for differential proteins

4 Conclusions

A relative quantitative method for differential proteomics was established based on cICAT and MALDI-TOF-TOF MS. The experimental result shows that the accuracy and reproducibility of the method is approximately 30%. Five standard proteins were used as the model proteins for the investigation of the dynamic range of the method. The result shows that the dynamic range of the method is from 1:1 to1:30. When the dynamic range is out of this range, and although the relative error also becomes bigger, the identification of differential proteins is not hindered. Therefore, this method is applicable in field of comparable proteomics.

Acknowledgements This work was supported by the National High Technologies R&D Program of China (2004CB518707, 2006CB0D1601), the National Natural Science Foundation of China (Grant Nos. 30621063, 20405017, 20505018, 20505019) and the Science and Technology Foundation of Beijing Municipality (H030230280190).

References

1. Gygi S P, Rist B, Gerber S A, Turecek F, Gelb M H, Aebersold R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology, 1999, 17(10): 994−999
2. Han D K, Eng J, Zhou H, Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded

affinity tags and mass spectrometry. Nature Biotechnology, 2001, 19(10): 946−951
3. Gygi S P, Rist B, Griffin T J, Eng J, Aebersold R. Proteome analysis of low-abundance proteins using multidimensional chromatography and isotope-coded affinity tags. Jounal of Proteome Research, 2002, 1(1): 47−54
4. Yi E C, Li X J, Cooke K, Lee H, Raught B, Page A, Aneliunas V, Hieter P, Goodlett D R, Aebersold R. Increased quantitative proteome coverage with $^{13}C/^{12}C$-based, acid-cleavable isotope-coded affinity tag reagent and modified data acquisition scheme. Proteomics, 2005, 5(2): 380−387
5. Molloy M P, Donohoe S, Brzezinski E E, Kilby G W, Stevenson T I, Baker J D, Goodlett D R, Gage D A. Large-scale evaluation of quantitative reproducibility and proteome coverage using acid cleavable isotope coded affinity tag mass spectrometry for proteomic profiling. Proteomics, 2005, 5(5): 1204−1208