Subfemtomole level protein sequencing by Edman degradation carried out in a microfluidic chip†

Wenzhang Chen,[a] Xuefeng Yin,*[a] Jinxia Mu[a] and Yan Yin[b]

Received (in Cambridge, UK) 8th January 2007, Accepted 8th March 2007
First published as an Advance Article on the web 27th March 2007
DOI: 10.1039/b700200a

A novel microfluidic chip based Edman degradation system is developed, in which Edman degradation can be carried out with peptide at a subfemtomole level; combined with MALDI-TOF-MS detection, the identification specificity of gel-separated protein in low abundance has been drastically improved.

The sequence analysis of proteins and peptides is important in biological research, especially for those in low abundance, *e.g.* disease-associated proteins.[1] Recently, mass analysis of tryptic peptides has become a powerful tool in protein identification due to the advancements in both mass spectrometry instrumentation and bioinformatics.[2] However, a problem of this method is its limitations in the accuracy and specificity of protein identification.[3–6] *De novo* peptide sequencing by MS/MS is not widely used because of the data complexity as well as the absence of certain key ions. Furthermore, it is normally impossible to obtain the information of N-terminal residues of a peptide by using MS/MS spectra.[7,8] Perkins *et al.* have demonstrated that the specificity of protein identification by mass spectrometry can be substantially improved by simply knowing the N-terminal residue of several peptides in a peptide mass fingerprint.[3]

Edman automated isothiocyanate degradation in 1967[9] and his method has been the standard for peptide sequencing. Edman degradation is an accurate way to easily interpret long amino acid sequences. Currently, it is still the only method that allows direct sequencing the N-terminal of an unknown protein.[10] However, the application of Edman degradation for the identification of low abundant proteins is limited due to its low sensitivity.[11] Considerable effort has been expended to improve the sensitivity of Edman degradation. Miyashita *et al.* have described an increase in sequencing sensitivity by coupling Edman degradation with accelerator mass spectrometry, however, this technique is restricted to analyzing ^{14}C-labeled proteins.[11] Powell and Tempst have described a microflow-based instrument which is capable of analyzing 100 fmol of sample.[12] The bottleneck of the current Edman degradation method is that the final elution volume formed in the sequencer is too large.

[a]*Institute of Microanalytical Systems, Department of Chemistry, Zhejiang University, Hangzhou, 310027, China. E-mail: yinxf@zju.edu.cn; Fax: +86-571-87952070; Tel: +86-571-87952070*
[b]*Chemistry Department, University of New Hampshire, 23 College Road, Durham, NH, 03824, USA*

† Electronic supplementary information (ESI) available: Experimental details. Table S1: The peptide sequences matched to the tryptic peptide mass fingerprint of Enolase 1. Table S2: The peptide sequences matched to the tryptic peptide mass fingerprint of T-cell receptor beta chain. See DOI: 10.1039/b700200a

Here we report a novel microfluidic chip based Edman degradation system that employs a nanoliter-scale C18 column as a reaction cartridge. This system is capable of analyzing as little as 0.2 fmol of sample. The C18 column with high aspect ratio and large specific interfacial area strongly retains protein/peptide, which can significantly minimize the washout during the sequencing process. The truncated peptides after Edman degradation are eluted effectively from the nanoliter-scale reaction cartridge with 50 nL eluent and are directly spotted onto a MALDI target without any sample loss. Combined with MALDI-TOF-MS detection, the sensitivity of Edman degradation and the identification specificity of gel-separated protein in low abundance have been drastically improved.

The chip-based nanoliter-scale reaction cartridge packed with C18 particles was fabricated on a glass substrate using photolithographic and wet chemical etching procedures which has been described elsewhere.[13] The volume of the reaction cartridge was 3.5 nl. The schematic of the microfluidic chip based Edman degradation system is shown in Fig. 1.

Edman degradation was carried out by a gas phase sequencing procedure on a hot plate at 45 °C for the coupling and cleavage reactions according the following scheme: (1) the peptide(s) was immobilized on the reaction cartridge and the salts and buffers present in the sample were washed away by using 0.3 μL 10% ethanol; (2) the reaction cartridge was completely dried by N_2 for 3 min; (3) TMA vapor was purged through the reaction cartridge for 3 min; (4) 50 nl 5% PITC was delivered to the reaction cartridge; (5) TMA vapor was purged for 30 min to keep the reaction cartridge at the desired basic environment for the coupling reaction; (6) excess TMA and PITC were washed away by using 0.3 μL 10% ethanol and heptane, respectively, and the column was dried completely by N_2 for 3 min; (7) TFA vapor was purged for 3 min for the cleavage reaction; (8) nitrogen was purged to remove the residual TFA vapor and the reaction cartridge was washed by using 0.3 μL 10% ethanol to remove phenylthiohydantoin amino acid as well as the other byproducts of the cleavage reaction; (9) truncated peptides were eluted from the reaction cartridge with 50 nL MALDI matrix solution, and the eluent was directly spotted onto the target for MALDI-TOF-MS. The delivery of gas reagents was achieved by aspirating the gas reagents into a plastic injector, which was followed by placing the plastic injector firmly on the access hole and gently pressing the plunger. The reagents flowed through the reaction cartridge slowly. The mass of the truncated peptide after the Edman degradation cycle was determined by MALDI-TOF-MS. The N-terminal sequence of each peptide was identified by using the subtractive Edman degradation.[14]

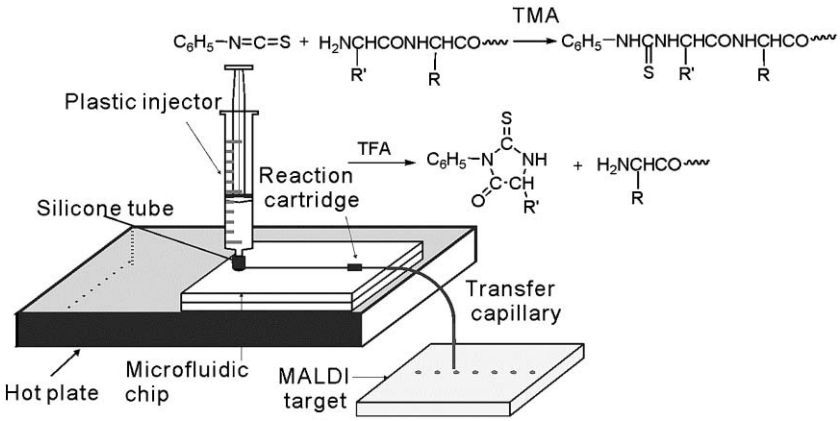

Fig. 1 Schematic of the microfluidic chip based Edman degradation system. The dimensions of the chip are 2.5 cm × 4.5 cm.

A major challenge to improve the sensitivity of the Edman degradation system was how to accurately deliver very small volumes of reagent. In order to reproducibly deliver 50 nL PITC and eluent through the nanoliter-scale reaction cartridge, a manual delivery procedure was developed. According to the calculation, the cross-sectional area of the micro channel was about 2.25×10^{-3} mm^2. After adding the reagent into the access hole, the reagent was pushed to flow slowly within the channel by pressing the plastic injector gently. 50 nL reagent, corresponding to 2.2 cm liquid column, was filled in the micro channel, and the residual reagent in the access hole was aspirated away by using a pipette. Then, the 2.2 cm liquid column moved slowly through the column under the air pressure generated by pressing the plastic injector. In this way, both 50 nL eluent and 5% PITC solution could be transferred quantitatively through the column.

The sensitivity of this method was evaluated by using 0.5 fmol and 0.2 fmol of Angiotensin II as the starting materials, whose mass spectra obtained after one cycle of Edman degradation are shown in Fig. 2(a) and (b), respectively. For the 0.5 fmol sample (Fig. 2(a)), the truncated peptide peak (931.5) was easily recognized (S/N = 8), which indicated that the first amino acid residue was aspartic acid. Even for the 0.2 fmol sample of Angiotensin II, the degraded peptide peak could be distinguished from the noise (S/N = 3, Fig. 2(b)). Our results demonstrated that the sensitivity of the microfluidic chip based Edman degradation system was increased by several hundred-fold compared to the current Edman degradation method. The results also demonstrated that the aspartic acid was completely cleaved from Angiotensin II, which indicated that a very high yield of recovery was achieved. However, the sequencing accuracy would not be affected by the recovery yield. It had been reported that proline residue underwent a slower reaction than other amino acids in the Edman chemistry, which resulted in the significant sequencing lag.[11] Therefore, the performance of the microfluidic chip based Edman degradation system was evaluated by sequencing 2 fmol of synthetic peptide P14R (PPPPPPPPPPPPPR). Fig. 2(c) and (d) correspond to the mass spectra of P14R acquired after five or two consecutive degradation cycles, respectively, significant sequencing lag with a mass difference of 97 could be seen, which was consistent with the adjacent proline residues in P14R. The MALDI mass data revealed that about one-third of P14R was converted to P9R after five degradation cycles and about 70% of P14R was converted to

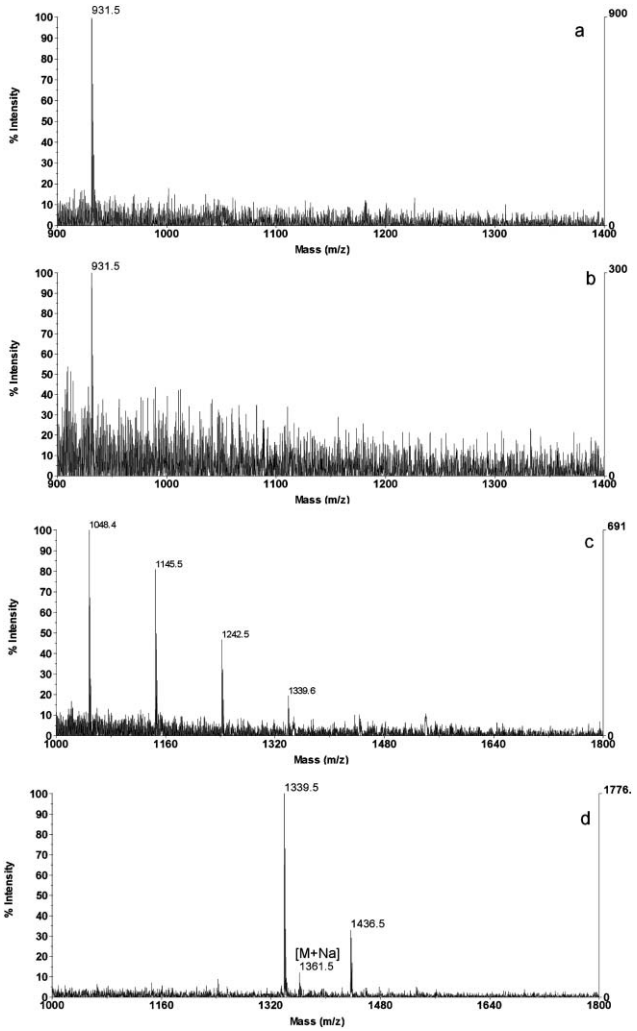

Fig. 2 Mass spectra of synthetic peptides after Edman degradation: (a) 0.5 fmol of Angiotensin II after one cycle of Edman degradation, (b) 0.2 fmol of Angiotensin II after one cycle of Edman degradation, (c) 2 fmol of P14R after five cycles of Edman degradation and (d) 2 fmol of P14R after two cycles of Edman degradation. The signals at *m/z* 1436, 1339, 1242, 1145 and 1048 correspond to $[M + H]^+$ for P13R, P12R, P11R, P10R and P9R, respectively.

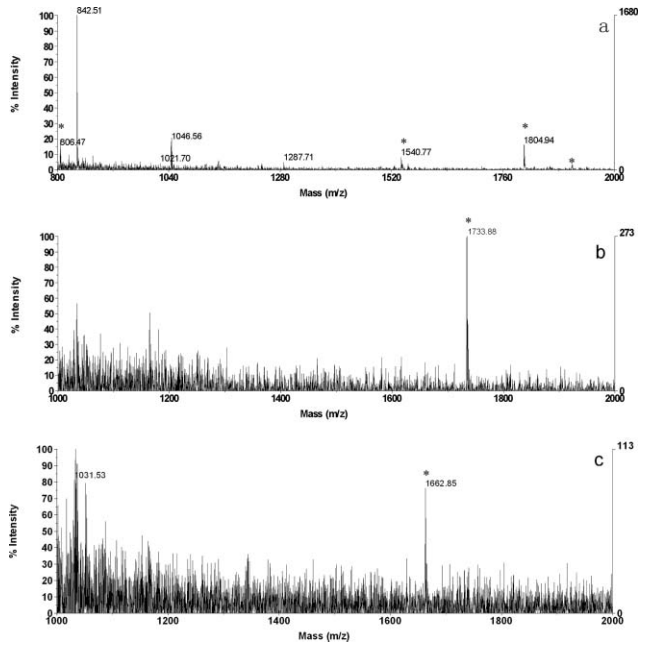

Fig. 3 Mass spectra of the tryptic peptides derived from in-gel digested protein after one and two cycles of Edman degradation: (a) native peptide, (b) one cycle degradation, and (c) two cycle degradation; *identified peaks.

P12R after two degradation cycles, which implied that the yield of the recovery of the peptide after each cycle of Edman degradation was about 80%.

A major issue in the use of database searching for peptide mass fingerprinting is false identifications.[6] Therefore, the performance of this method for proteomic applications was further evaluated by analyzing the tryptic peptides of some proteins, which was extracted and digested from the silver stained 2-D gels. Mass spectra of the peptides in the absence and presence of Edman degradation are shown in Fig. 3. As the control, Fig. 3(a) is the mass spectrum of 2 μL of in-gel digested peptides concentrated and desalted by the micro SPE column.[13] Only a few peaks with low intensities were detected by MALDI-TOF MS, indicating the low abundance sample. It should be mentioned that the sample was identified as T-cell receptor beta chain and Enolase in the absence of Edman degradation. Both results received a score of 51 by the database search with the mass mapping data,[13] which was below the significance threshold, 64, for a database of this size. Therefore, it was difficult to give a clear identification. Fig. 3(b) and (c) are the mass spectra of 1 μL sample acquired after one and two degradation cycles, respectively. In Fig. 3(b), the peak at *m*/*z* 1733.8 matched the tryptic peptides (AAVPSGASTGIYEALELR) of Enolase missing one N-terminal residue. As seen in Fig. 3(c), the peak at *m*/*z* 1662.8 matched Enolase tryptic peptides losing two N-terminal residues. Thus, it was confirmed that the source protein was Enolase (see ESI†). A score of 76 was achieved by the database searching with the N-terminal sequence information.

In summary, the sensitivity of the Edman degradation has been greatly enhanced by using the microfluidic chip based Edman degradation system, which leads to an extremely small eluent volume required for detection. This new method has the feasibility of direct coupling with MALDI-MS to minimize the sample loss. It has been successfully applied to determine the N-terminal sequence of tryptic peptides derived from in-gel digested proteins excised from 2-D silver stained gels. In combination with the peptide mass fingerprint, the confidence level of low abundance protein identification can be drastically increased.

This work was supported by the National Natural Science Foundation of China under project No. 20475049.

Notes and references

1 G. Cagney and A. Emili, *Nat. Biotechnol.*, 2002, **20**, 163–170.
2 H. Steen and M. Mann, *Nat. Rev. Mol. Cell. Biol.*, 2004, **5**, 699–711.
3 D. N. Perkins, D. J. C. Pappin, D. M. Creasy and J. S. Cottrell, *Electrophoresis*, 1999, **20**, 3551–3567.
4 L. A. Huber, *Nat. Rev. Mol. Cell. Biol.*, 2003, **4**, 74–80.
5 D. Chamrad and H. E. Meyer, *Nat. Methods*, 2005, **2**, 647–648.
6 D. J. States, G. S. Omenn, T. W. Blackwell, D. Fermin, J. Eng, D. W. Speicher and S. M. Hanash, *Nat. Biotechnol.*, 2006, **24**, 333–338.
7 Q. Fu and L. Li, *Anal. Chem.*, 2005, **77**, 7783–7795.
8 E. S. Boja, E. A. Sokoloski and H. M. Fales, *Anal. Chem.*, 2004, **76**, 3958–3970.
9 P. Edman and G. Begg, *Eur. J. Biochem.*, 1967, **1**, 80–91.
10 B. Samyn, K. Sergeant, P. Castanheira, C. Faro and J. V. Beeumen, *Nat. Methods*, 2005, **2**, 193–200.
11 M. Miyashita, J. M. Presley, B. A. Buchholz, K. S. Lam, Y. M. Lee, J. S. Vogel and B. D. Hammock, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 4403–4408.
12 M. Powell and P. Tempst, *Anal. Chem.*, 2001, **73**, 776–786.
13 W. Z. Chen, J. Shen, X. F. Yin and Y. N. Yu, *Rapid Commun. Mass Spectrom.*, 2007, **21**, 35–43.
14 C. V. Bradley, D. H. Williams and M. R. Hanley, *Biochem. Biophys. Res. Commun.*, 1982, **104**, 1223–1230.

