RAPID COMMUNICATIONS IN MASS SPECTROMETRY
Rapid Commun. Mass Spectrom. 2002; **16**: 1892–1895
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rcm.798

RCM

Letter to the Editor

To the Editor-in-Chief
Sir,

An improved two-step calibration method for matrix-assisted laser desorption/ionization time-of-flight mass spectra for proteomics

Since its introduction in 1993,[1] matrix-assisted laser desorption/ionization (MALDI) peptide mass fingerprinting (PMF) has become one of the standard methods for proteome research. Among other factors, such as the size of the protein and sequence database, the reliability of a protein identification using PMF primarily depends on the number of peptide ion signals matching to the respective protein hit and the accuracy with which the *m/z* values of these signals can be determined.

In an ideal, continuous extraction linear time-of-flight instrument, a simple internal two-point calibration should yield good mass accuracy over the entire *m/z* range, since *m/z* is proportional to t^2 (t is flight time). In practice, however, good mass accuracy can only be obtained with instruments equipped with time-delayed ion extraction, low mass deflection pulsers and single- or dual-stage reflectrons that partly compensate for the space and energy distributions of ions generated by MALDI. These instruments have attained resolving power beyond 10 000 FWHM and sensitivities in the attomole range, but delayed ion extraction, low mass ion deflection, and the reflectron directly influence individual molecular ion velocities and, thus, complicate the correlation between *m/z* and flight time. Therefore, mass accuracies in mass spectra of peptides have generally not reached beyond 20 ppm.

Recently, some effort has been put into optimizing calibration procedures for increasing mass accuracy over a broad mass range.[2–5] For example, simplex optimization and quadratic equations were used to describe the correlation between systematic mass errors and *m/z*.[4,5] However, mass accuracies in the low ppm range could not be achieved. Recently, Gobom and co-workers[3] suggested a procedure using polyethylene glycol (PPG) to determine systematic mass errors and eliminate those by a 15th order correction function. Although initial velocities of PPG and peptides might be different, good mass accuracies were achieved when this method was applied to PMF spectra. However, PPG yields only one signal per mass increment of 58 Da. In order to achieve a sufficiently large data set for calculation of higher-order fitting functions, a method allowing the implementation of more data points generated from compounds with a similar chemical nature to analytes should be beneficial. Here we report a calibration method that is based on the empirical determination of systematic *m/z*-dependent mass errors in internally calibrated MALDI PMF spectra and the elimination of these errors using a higher-order *m/z* correction function. This method was found to greatly improve mass accuracy in MALDI peptide mass fingerprints and, thus, enables more reliable protein identification.

In order to determine systematic mass errors corresponding to distinct *m/z* values in MALDI-TOF spectra, a training set of peptide mass fingerprint spectra from 50 randomly picked in-gel digested (trypsin) protein samples was acquired. In addition, 25 independent measurements of 5 fmol of a solution digest of myosin with trypsin were performed. Overall, the training set consisted of more than 1800 peptide mass measurements.

All spectra were acquired using a Reflex IV™ MALDI-TOF instrument (Bruker Daltonik, Bremen, Germany) equipped with a gated detector. The instrument was run in reflector mode at a full reflector voltage of 27.5 kV. The high voltages of the instrument were always switched on between analyses in order to ensure stable instrument voltages. When sample plates were exchanged, high voltages were switched on at least 30 min prior to spectra acquisition.

Samples were spotted onto 600 µm anchor targets (Scout 384-MTP AnchorChip; Bruker Daltonik, Bremen, Germany) according to a dried-droplet protocol using α-cyano-4-hydroxycinnamic acid (CHCA) as matrix. Three different target plates were used and sample positions were evenly distributed over the target plates in order to eliminate position and plate dependent parameters. It should be noted that the outer two rows and columns of the target plates were not used because we observed far worse mass accuracies and sensitivities at these positions than at the other positions. All spectra were recorded in automatic mode. Peak labeling and internal two-point calibration on trypsin autolysis products (*m/z* 842.5100 and 2211.1046) were performed without user interference using the SNAP algorithm and Aura scripts implemented into the XTOF software. Database searches were carried out using Profound (Genomic Solutions, Ann Arbor, USA) with mass tolerance set to 70 ppm.

Observed masses were compared with calculated masses of matching peptides of the identified proteins and relative mass errors (in ppm) were calculated for each peptide ion signal. Systematic mass errors corresponding to distinct *m/z* values were determined by averaging the measured mass errors of individual peptides corresponding to the same *m/z* value. Plotting of the relative mass error over *m/z* indicated a continuous distribution of data points (Fig. 1(a)) that was subsequently subjected to a polynomial fitting procedure using standard Excel worksheet extensions. A 7th order polynomial function was found to approximate the correlation between *m/z* and relative mass errors (Fig. 1(a)). Application of this polynomial function to the training set drastically reduced relative mass errors. All data points were now found within ±20 ppm (Fig. 1(b)). Using internal calibration only, the average relative mass error of the training set

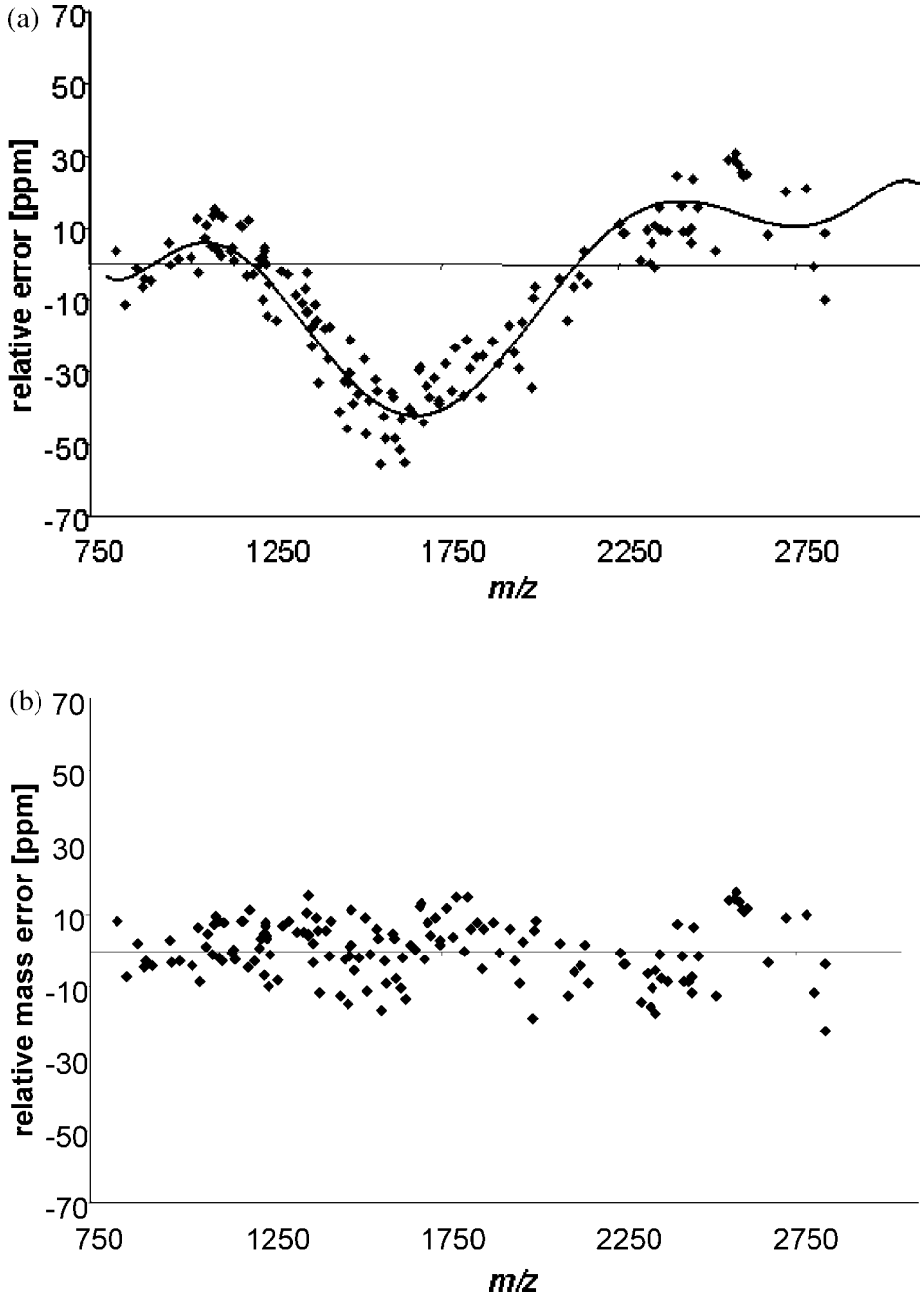

Figure 1. (a) Distribution of average relative mass errors over *m/z* using internal two-point calibration. Squares indicate average mass errors corresponding to distinct *m/z* values as calculated from the training set of peptide mass fingerprints. The 7th order polynomial function that was fitted to this data set is indicated by a black line. The function used was ppm = −1.50688 E-19 * (*m*/*z*)^7 + 2.08088 E-15 * (*m*/*z*)^6 − 1.18717 E-11 * (*m*/*z*)^5 + 3.60873 E-8*(*m*/*z*)^4 − 6.27943 E-5* (*m*/*z*)^3 + 0.062285023* (*m*/*z*)^2 − 32.56037441 * (*m*/*z*) + 6927.728364. (b) Squares indicate average mass deviations corresponding to distinct *m/z* values from the training set of peptide mass fingerprints after the higher order correction function was applied.

was −10.3 ppm with a standard deviation of 21.1 ppm. The average absolute mass error was 18.5 ppm with a standard deviation of 14.4 ppm. After recalculation of observed masses using the polynomial correction function, the average relative mass error was reduced to 0.5 ppm with a standard deviation of 8.2 ppm and the average absolute mass error was reduced to 6.8 ppm with a standard deviation of 4.6 ppm.

Having implemented this mass correction procedure into routine analysis using a simple visual basic script that automatically corrects the peak lists of a given set of samples according to the correction function, we could reduce the mass tolerance allowed for protein identification using PMF from 70 to 20 ppm. Based on 50

RCM

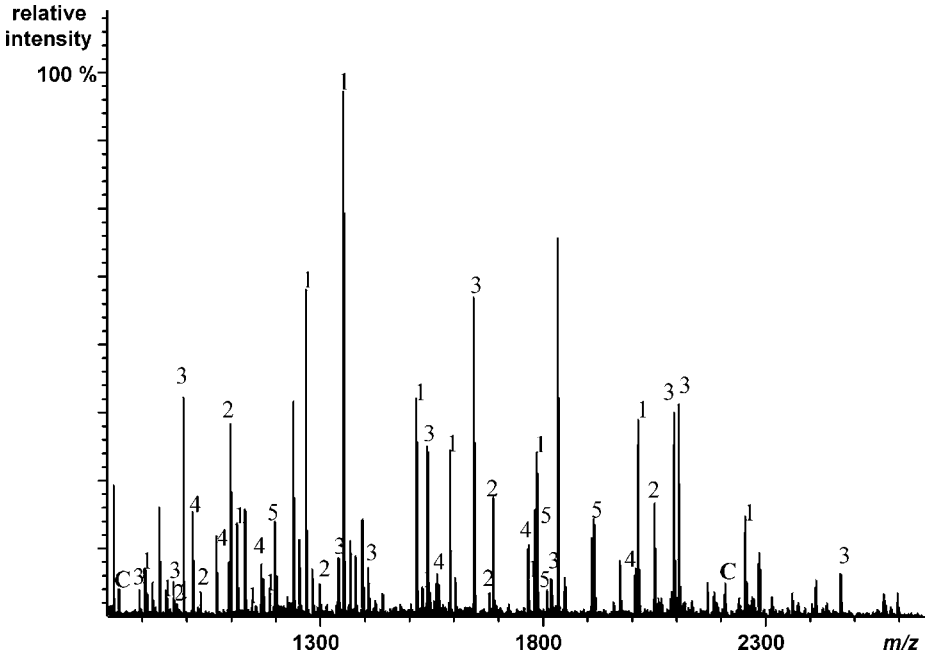

Figure 2. MALDI-TOF spectrum of a mixture consisting of five in-gel digested proteins. 1: HSP90 alpha subunit was identified with an average absolute mass error of (4.3 ± 3.8) ppm; 2: mitogen-activated protein kinase kinase kinase 14 (5.9 ± 4.9) ppm; 3: inhibitor of NFκB kinase alpha subunit (6.7 ± 3.4) ppm; 4: inhibitor of NFkB kinase beta subunit (5.5 ± 4.3) ppm; and 5: HSP90 beta subunit (4.3 ± 3.8) ppm.

randomly picked protein identifications acquired over a period of 3 months, we found an average relative mass error of 0.1 ppm with a standard deviation of 7.8 ppm and an average absolute mass error of 6.5 ppm with a standard deviation of 4.8 ppm. Within the observed time frame mass accuracies achieved with this procedure were stable. A further refinement of the fitted function was therefore not necessary.

Owing to the increased mass accuracy, false positive protein identifications were drastically reduced and identifications of multiple proteins present in one band of gel-separated protein mixtures became more reliable. For example, Fig. 2 shows the identification of five different proteins present in a mixture using the described calibration technique (20 ppm mass tolerance allowed). For all identified proteins, average absolute mass errors between 4.6 and 6.7 ppm with standard deviations between 3.6 and 5.0 ppm were achieved. Even for weak ion signals, a significant increase in mass error could not be observed. When only internal calibration was applied (70 ppm mass tolerance allowed), merely the first two proteins could be unambiguously identified.

The method presented here provides a simple means to improve mass accuracy in MALDI-TOF spectra using a two-step calibration method. In a first step, spectra are internally calibrated using trypsin autolysis products and in a second step peak lists are corrected for higher-order calibration coefficients using a polynomial function. Calculation of the appropriate function is rather straightforward. Spectra acquisition of a sufficient number of peptide mass fingerprints, calculation of mass errors and polynomial fitting of the mass error dependency on *m/z* can be done in one day. Although, for high-throughput analyses, an automated recalculation of peak lists would be appropriate, the simplest implementation of the correction function can take the form of an Excel spreadsheet.

It seems important to mention that the correlation between *m/z* and flight time and, thus, the appropriate polynomial function to use, strongly depends on instrument design and the instrumental parameters applied. Delayed ion extraction time and voltages, ion acceleration voltage, reflector voltage, and time settings of low-mass ion deflection pulsers strongly influence the mass error distribution over *m/z*. Therefore, a new correction function must be obtained for every set of instrumental parameters. Since usually only one optimized set of instrumental parameters is used for peptide mass fingerprinting, this should not negatively influence the usability of this method in everyday laboratory work.

Acknowledgements

We would like to thank Dr. Volker Schnaible (Stiftung CAESAR, Bonn, Germany) for the kind provision of a Visual Basic script for automated peak list processing.

Marcus Bantscheff, Birgit Duempelfeld, Bernhard Kuster*
Cellzome AG, Meyerhofstrasse 1, 69117 Heidelberg, Germany

**Correspondence to: B. Kuster, Cellzome AG, Meyerhofstrasse 1, 69117 Heidelberg, Germany*
E-mail: bernhard.kuster@cellzome.com

REFERENCES

1. Henzel WJ, Billeci TM, Stults JT,

Wong SC, Grimley C, Watanabe C. *Proc. Natl. Acad. Sci. USA* 1993; **90**: 5011.
2. Whittal RM, Russon LM, Weinberger SR, Li L. *Anal. Chem.* 1997; **69**: 2147.
3. Gobom J, Mueller M, Engelhofer V, Theiss D, Lehrach H, Nordhoff E. *Anal. Chem.* 2002; ASAP article, Web release Date June 22 2002.
4. Christian NP, Arnold RJ, Reilly JP. *Anal. Chem.* 2000; **72**: 3327.
5. Hack CA, Benner WH. *Rapid. Commun. Mass. Spectrom.* 2002; **16**: 1304.

Received 18 July 2002
Revised 24 July 2002
Accepted 24 July 2002