Automation of Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Using Fuzzy Logic Feedback Control

ARTICLE *in* ANALYTICAL CHEMISTRY · JUNE 1997

Impact Factor: 5.64 · DOI: 10.1021/ac961189t · Source: PubMed

4 AUTHORS, INCLUDING:

Ole N Jensen
University of Southern Denmark
298 PUBLICATIONS **19,034** CITATIONS
SEE PROFILE

Ole Vorm
Bruker Corporation
25 PUBLICATIONS **9,846** CITATIONS
SEE PROFILE

Anal. Chem. **1997,** *69,* 1706−1714

Automation of Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry Using Fuzzy Logic Feedback Control

Ole N. Jensen, Peter Mortensen, Ole Vorm,† and Matthias Mann*

Protein & Peptide Group, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Matrix-assisted laser desorption/ionization (MALDI) mass spectrometry is a sensitive and versatile method for biomolecular analysis which has potential for high-throughput screening in many applications. To obtain mass spectra of optimal quality, however, laser fluence is continuously adjusted during data acquisition to be close to the threshold level of ion production, requiring a skilled operator and several minutes of acquisition time per sample. Using real-time fuzzy logic control of the laser fluence, we here demonstrate that the acquisition of MALDI spectra can be automated without reduction of data quality. The control algorithm evaluates signal intensity and mass resolution of the base peak. It then regulates the laser fluence to keep the ion signal intensity within the dynamic range of the data acquisition hardware while maintaining high mass resolution. This fuzzy logic control system allows unattended data acquisition using either static ion extraction or delayed ion extraction MALDI. Even for difficult samples such as femtomole-level peptide mixtures, no significant reduction in data quality is observed, as compared to manually obtained spectra. Automated analysis of 78 chromatographic fractions with high mass accuracy demonstrates the utility of the method. The control algorithm has been combined with other software modules to completely automate database identification of proteins by their peptide mass maps. The success of fuzzy logic in MALDI automation suggests wider uses of this technique in mass spectrometry.

Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry has proven to be a fast, sensitive, and versatile technique for the mass analysis of biomolecules.[1,2] Recent advances in MALDI (reviewed in ref 3), particularly the introduction of delayed extraction,[4−7] have transformed MALDI-TOFMS into a high-performance mass spectrometric technique with parts-per-million mass accuracy and subpicomole to subfemtomole detection limits. Presently, MALDI-TOFMS is mainly employed for the mass analysis of biopolymers such as proteins and peptides as well as oligonucleotides and carbohydrates, even though limited structural information can also be gained by taking advantage of the metastable decay of MALDI-generated ions.[8]

A mass measurement by MALDI-TOFMS is relatively simple, involving only the laser irradiation of a coprecipitate of UV-absorbing matrix and analyte, followed by the time-resolved detection of analyte ions. This simple nature, combined with other desirable features, has made MALDI the method of choice in many protein chemistry laboratories, particularly for routine tasks, such as quality control of synthetic peptides, screening of HPLC fractions prior to Edman degradation, and analysis of peptide mixtures derived from proteins (peptide mass mapping). These measurements are, in most cases, routine because standard sample preparation methods and standard operating conditions of the mass spectrometer are used.[9,10] It would be desirable to obtain MALDI-TOF spectra unattended in these applications, freeing up operator time and allowing the analysis of large numbers of samples. However, unlike automation of electrospray mass spectrometry,[11] with which hundreds of samples per day can be analyzed in specialized applications, the automation of MALDI has proven difficult. The reason lies in the variability of the ion signal as a function of laser fluence (or irradiance) and as a function of the local irradiated environment. "Shot-to-shot" variability of the ion signal can be several-fold as a spot on the matrix surface is depleted or as a new spot is irradiated. Optimal resolution is obtained close to the threshold of ion production, leading to a balance between optimal signal-to-noise ratio and resolution. Too high a laser fluence can also result in "overshoot" of the ion signal, saturating the detection electronics and leading to a drastic deterioration of the resolution and mass accuracy for the peak in question. For these reasons, a skilled operator is necessary to keep the laser fluence in an optimal range during the accumulation of the 50−200 single-shot spectra usually added to produce the MALDI spectrum.

Attempts to automate data acquisition in MALDI-TOFMS have been pursued by most instrument manufacturers. Typically, a

* Corresponding author. Phone: +49 6221 387 560. Fax: +49 6221 387 306. E-mail: Mann@EMBL-Heidelberg.de.URL:http://www.mann.embl-heidelberg.de.

† Current address: The Protein Analysis Co., International Science Park Odense, Forskerparken 10, DK-5230 Odense M, Denmark.

(1) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988,** *60,* 2299−2301.
(2) Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1991,** *63,* 1193A−1202A.
(3) Mann, M.; Talbo, G. *Curr. Opin. Biotechnol.* **1996,** *7,* 11−19.
(4) Brown, R. S.; Lennon, J. J. *Anal. Chem.* **1995,** *67,* 1998−2003.
(5) King, T. B.; Colby, S. M.; Reilly, J. P. *Int. J. Mass Spectrom. Ion Processes* **1995,** *145,* L1−L7.
(6) Whittal, R. M.; Li, L. *Anal. Chem.* **1995,** *67,* 1950−1954.
(7) Vestal, M. L.; Juhasz, P.; Martin, S. A. *Rapid Commun. Mass Spectrom.* **1995,** *9,* 1044−1050.
(8) Spengler, B.; Kirsch, D.; Kaufmann, R.; Jaeger, E. *Rapid Commun. Mass Spectrom.* **1992,** *6,* 105−108.
(9) Beavis, R. C.; Chait, B. T. In *Methods in Enzymology*; Karger, B. L., Hancock, W. S., Eds.; Academic Press: San Diego, 1996; Vol. 270A, pp 519−551.
(10) Jensen, O. N.; Shevchenko, A.; Mann, M. In *PROTEIN STRUCTURE—A PRACTICAL APPROACH*, 2nd ed.; Creighton, T. E., Ed.; IRL Press, Oxford University Press: Cambridge, UK, 1997; pp 29−57.
(11) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989,** *246,* 64−71.

S0003-2700(96)01189-4 CCC: $14.00

preset laser irradiance is chosen which leads to good mass spectral quality for ideal samples, such as synthetic peptides.[12] We have previously attempted to automate MALDI data acquisition by ad hoc feedback control algorithms which regulated the irradiance as a function of the ion signal intensity in "real-time", i.e., *during* MALDI data acquisition and while moving the probe relative to the laser beam. However, these algorithms were not stable in practice due to the variability of the ion signal and the consequent tendency of the control algorithm to overregulate in a particular direction. We noted from the literature that "fuzzy logic" control had been employed with success in similarly difficult control applications, and we therefore investigated this approach.

Fuzzy logic (or continuous set theory) is based on the premise of continuous membership in sets. The theory is explained in a body of literature[13,14] and popular works[15,16] and has been applied in analytical chemistry[17−19] but will, nevertheless, be outlined briefly here. As an example, the set of "tall" people is usually mathematically segregated from the set of other people by a given boundary—such as a height of 1.85 m. This crisp boundary creates artifacts due to the discrete nature of the subdivision, i.e., a person of 1.849 m is not "tall" but a person of 1.8501 m is. In fuzzy set theory, every person would be a member in the set of tall people to a degree—according to a membership function—and the degree of membership in the above two cases would be almost the same. One of the advantages of fuzzy logic is a more rational and natural mapping of everyday objects into mathematical abstractions. This is expressed by the plain English terms which can be used in fuzzy logic modeling, such as "very" tall or "rather" tall, which are mapped into degrees of membership in sets.

Fuzzy logic control is the application of fuzzy logic concepts in the field of systems control. It allows the specification of very few rules, which are then evaluated in parallel by a fuzzy logic interpreter. Fuzzy logic control has been found to be very robust and simple to implement. A control algorithm can be specified in a few simple IF−THEN rules, which are then translated into software code by a fuzzy logic compiler. A fuzzy logic rule for nonlinear regulation of the laser irradiance in MALDI could be as simple as several statements of the following form: IF ion intensity is very low THEN change laser irradiance much. In classical linear control algorithms, in contrast, a mathematical model would have had to be defined, specifying exactly by what degree to change the laser irradiance as a response to ion signal. A further advantage of fuzzy logic control is that variables and rules can be added with little effect on the stability of the algorithm, such as IF resolution is low THEN decrease laser intensity.

In this article, we investigate whether fuzzy logic control can be applied to MALDI mass spectrometry and whether it can be used as a building block to completely automate data acquisition and interpretation. Routine automation of MALDI data acquisition would be a major step forward in making MALDI applicable to high-throughput applications.

EXPERIMENTAL SECTION

Sample Preparation. Matrix surfaces were made by fast evaporation on the MALDI probe tip of 0.2−0.3 μL of a saturated solution of 4-hydroxy-α-cyanocinnamic acid (4HCCA, Sigma, St. Louis, MO) in acetone.[20,21] Some samples were analyzed using nitrocellulose as an added component of the thin matrix films.[22,23] Dried peptide samples were redissolved in 20−30 μL of 1% formic acid/10% acetonitrile, and 0.5 μL of the peptide solution was deposited on top of the fast-evaporation matrix film. The liquid was left to evaporate at ambient temperature, and the dried sample deposit was then washed with pure water or 1% formic acid prior to mass analysis. Dried-droplet sample deposits were prepared by mixing 0.5 μL of peptide solution with 4.5 μL of matrix solution (4HCCA in 5% formic acid/acetonitrile, 2:1 v/v) and depositing 0.5 μL of the mixture on the MALDI probe tip. Dried sample deposits were rinsed with pure water or 1% formic acid prior to mass analysis.

MALDI-TOF Mass Spectrometry. MALDI mass spectra were acquired on a modified Bruker REFLEX time-of flight mass spectrometer (Bruker-Franzen Analytik, Bremen, Germany) equipped with the SCOUT multiprobe ion source.

Under static extraction conditions, ion acceleration voltage was 30 kV, and the reflector (ion mirror) voltage was 32 kV. Under delayed extraction conditions, the acceleration voltage was 22 kV. A 6 kV potential difference between the probe and the extraction lens was applied with a time delay in the range of 150−200 ns after each laser pulse using a high-voltage switch (Behlke, Frankfurt, Germany). The reflector voltage was set to 23.5 kV. A Lecroy 9350AM 1×10^9 samples/s digital storage oscilloscope was used for data acquisition (LeCroy Corp., Chestnut Ridge, NY). The laser fluence was attenuated by a circular attenuation filter (Newport, RI) mounted on a stepper motor controlled from the computer via a digital-to-analog converter (Bruker). Mass spectra were acquired as the sum of ion signals generated by irradiation of the target with 60−200 laser pulses (337 nm N_2 laser, Model VSL-337ND, Laser Science, Inc., Boston, MA).

Fuzzy Logic Software. The fuzzy logic interpreter provided in the MatLab Fuzzy Logic Toolbox (The MathWorks, Inc., Natick, MA) was converted to a stand-alone Macintosh application (FuzzyEngine) using a C++ compiler (Metrowerks, Austin, TX). The fuzzy rules and membership functions for determining the required change of laser attenuation on the basis of ion peak intensity and resolution were implemented using the MatLab graphical editor. Fuzzy logic feedback control software was used in combination with AppleScripts and LaserOne software for real-time control of the laser fluence (see below).

Data Acquisition Software. Instrument control, data acquisition, and data manipulation were performed by the LaserOne software developed in our group and running on an Apple Power Macintosh 7100/80 (Apple Computer, Inc., Cupertino, CA). The LaserOne program is scriptable by AppleScript, a scripting/macro

(12) Waddel, K.; Lessard, J.; Teller, D. Perseptive Biosystems Application Note, 1996.
(13) Zadeh, L. A. *Inf. Control* **1965,** *8*, 338−353.
(14) Kosko, B. *Fuzzy Engineering*; Simon & Schuster: Upper Saddle River, NJ, 1997.
(15) Kosko, B.; Isaka, S. *Sci. Am.* **1993**, No. 7, 62−67.
(16) McNeill, D.; Freiberger, P. *Fuzzy Logic*; Simon & Schuster: New York, 1993.
(17) Pop, H. F.; Sarbu, C. *Anal. Chem.* **1996,** *68*, 771−778.
(18) Yamakawa, T. *J. Biotechnol.* **1992,** *24*, 1−32.
(19) Xu, J.; Straus, S. K.; Sanctuary, B. C.; Trible, L. *J. Chem. Inf. Comput. Sci.* **1993,** *33*, 668−682.
(20) Vorm, O.; Roepstorff, P.; Mann, M. *Anal. Chem.* **1994,** *66*, 3281−3287.
(21) Vorm, O.; Mann, M. *J. Am. Soc. Mass Spectrom.* **1994,** *5*, 955−958.
(22) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. *Anal. Chem.* **1996,** *68*, 850−858.
(23) Jensen, O. N.; Podtelejnikov, A.; Mann, M. *Rapid Commun. Mass Spectrom.* **1996,** *10*, 1371−1378.

language built into the Macintosh operating system. This feature allows the user to write custom scripts for control of the MALDI mass spectrometer. LaserOne lets the user decide if the most recently acquired five-shot spectrum should be added to a "sum spectrum" or discarded. The final spectrum is, therefore, an average of up to 50 selected five-shot spectra, for a total of up to 250 shots/spectrum.

AppleScripts were written for coordinating the flow of data between the various applications used for automating MALDI-MS. The following applications were linked via scripts: database software (Filemaker Pro, Claris Corp., Cupertino, CA), LaserOne, FuzzyEngine, PowerPeaks (peak labeling software developed in-house), and PeptideSearch.[24,25] The scripts control probe movement, the peripheral electronics (power supplies and oscilloscope), and the laser attenuation filter (using the fuzzy logic control system). They also included routines for calibration of mass spectra, peak labeling, and saving and reporting the results.

The SCOUT sample probe of the MALDI mass spectrometer is computer-controlled in the $x-y$ direction by two servo motors, allowing any position on the circular probe to be irradiated by the laser beam. Samples are applied at up to 26 positions. During data acquisition at a given sample position, the probe is automatically moved in a spiral pattern to irradiate a new spot on the sample deposit for every five laser pulses. The 26 sample positions and the x,y coordinates for spiral probe movement are specified in an AppleScript. In a later version of the software, the probe movement was also controlled by the fuzzy logic system.

Mass Calibration of TOF Spectra. A straightforward and general method for mass calibration of static extraction MALDI-TOF spectra using matrix ion signals has recently been described.[26] A fast electronic switching circuit is used to reduce the potential (and thereby the sensitivity) of the detector during arrival of matrix ions, which are therefore detected as well-resolved signals that can be used for internal mass calibration. This detector bias gating method is well suited for an automated MALDI mass spectrometer because the calibration points (matrix ion signals) are inherent to every spectrum, and calibration can be performed automatically by a script as soon as a spectrum is acquired. Linear regression with the TOF and m/z values of matrix ion signals determines the mass calibration constants for the TOF mass spectrum.

In delayed extraction MALDI, not the whole mass range is time focused, and the mass scale is not linear over a wide range. Therefore, the above-mentioned approach for calibration using low-m/z matrix ion signals could not be used. External calibration with predefined values typically led to a mass accuracy of only 100−200 ppm. Instead, we added two internal peptide calibrants for analysis of synthetic peptides. Automated data acquisition and calibration with these mass standards allowed automated MALDI peptide mass determinations with an accuracy in the low ppm range, typically 5−20 ppm. Automatic mass calibration of tryptic peptide mass maps obtained by delayed extraction MALDI-MS after in-gel digestion of proteins was performed by using the high-m/z matrix ion signals (m/z 855.1, 1060.1) and peptide signals due to bovine trypsin autodigestion (m/z 2163.057, 2273.159) as internal calibrants.[23]

RESULTS AND DISCUSSION

Implementation of Fuzzy Logic Control. The important model variables monitored in MALDI are ion intensity and peak resolution. Based on these values, the operator continuously changes the laser fluence to optimize signal-to-noise ratio without compromising peak resolution. The decision process used by a human operator for determining the laser fluence during MALDI data acquisition can be simply described in everyday language by a few rules with *intensity* and *resolution* as input variables and the *fluence change* as the output variable.

Intensity rules:

1. IF Very_Low intensity THEN *fluence change* is Large_Positive
2. IF Low intensity THEN *fluence change* is Small_Positive
3. IF Average intensity THEN *fluence change* is None
4. IF High intensity THEN *fluence change* is Small_Negative
5. IF Very_High intensity THEN *fluence change* is Large_Negative

Resolution rule:

6. IF Low resolution THEN *fluence change* is Large_Negative

Using binary logic, each of the above-mentioned propositions would evaluate to a truth value of either 0 or 1. However, each of the variables can be considered a fuzzy set. Fuzzy set theory allows a proposition to be evaluated to a truth value *between* 0 and 1 by way of predefined membership functions for the variables.

In this work, we assumed that the most intense ion signal (the base peak) was in the peptide mass range from m/z 600 to 4000. The full width at half-maximum (fwhm) resolution was assumed to be in the range of 0−5000 for static ion extraction MALDI reflector TOF and in the range of 0−15 000 for delayed ion extraction MALDI reflector TOF. The variable *intensity* is the signal intensity, measured in millivolts, of the base peak in a five-shot averaged mass spectrum, whereas the variable *resolution* is the fwhm resolution of this base peak. Note that the mass resolution is easily taken into account by a single proposition in the fuzzy system (eq 6). Using a traditional feedback control algorithm, it would be significantly more complicated to define the influence of both ion intensity and resolution on the required laser attenuation change.

The allowed range (called the "universe of discourse") for the output variable *fluence change* was defined to be from −3% to +3% attenuation. On our MALDI system, a value of 1% is only a slight change in laser fluence, whereas a 3% change is a significant change.

The fuzzy qualifiers (Very_Low, Low, Average, etc.) are overlapping domains within the *intensity* membership function (Figure 1A), whereas the qualifiers Large_Positive, Small_Positive, etc. are overlapping domains within the *fluence change* membership function (Figure 1C). The *resolution* membership function is shown in Figure 1B. Note that there are no crisp boundaries between the domains of the membership functions, viz. the transition from an ion signal of low *intensity* to a signal of average *intensity* is "fuzzy". It is worth mentioning that the definitions of these membership functions are not critical—they can easily be

(24) Mann, M.; Højrup, P.; Roepstorff, P. *Biol. Mass Spectrom.* **1993,** *22*, 338−345.
(25) Mann, M. In *Microcharacterization of Proteins*; Kellner, R., Lottspeich, F., Meyer, H. E., Eds.; VCH: Weinheim, 1994; pp 223−245.
(26) Vorm, O.; Roepstorff, P. *J. Mass Spectrom.* **1996,** *31*, 351−356.

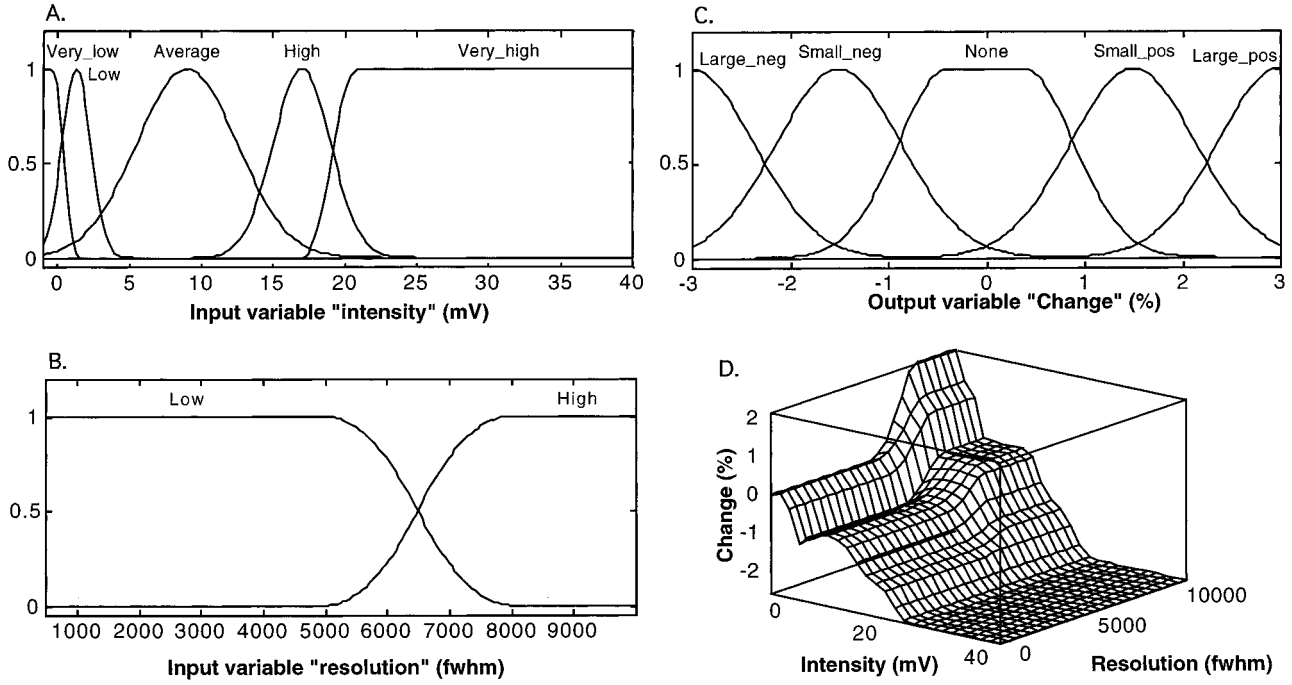

Figure 1. Definitions of fuzzy membership functions for input and output variables and the input−output space as defined for delayed extraction MALDI-TOFMS. (A) The membership function for the *intensity* input variable contains five domains (fuzzy qualifiers), i.e., Very Low, Low, Average, High, and Very High. (B) The membership function for the *resolution* input variable contains only two domains, Low and High. (C) The membership function for the output variable *fluence change*. (D) Surface plot of the input−output space defined by the fuzzy rules (eqs 1−6) and the three membership functions. The input variables *intensity* and *resolution* are mapped to the output variable, *laser fluence*, as defined by this surface.

modified to include a larger mass range or more domains. They can also be implemented by using triangular functions rather than Gaussian distributions. This provides a great deal of flexibility and error tolerance when implementing a fuzzy control system. The fuzzy membership functions are defined on the basis of the experience acquired by the operator during acquisition of many MALDI mass spectra. They basically define the acceptable range of ion intensity and peak resolution and the appropriate change in laser fluence. The membership functions can easily be modified and tuned by using the graphical editor in the Matlab Fuzzy Logic Toolbox.

The propositions are evaluated in parallel by the fuzzy inference engine provided in ANSI C source code with the MatLab Fuzzy Logic Toolbox using the above-mentioned membership functions. The set of outputs is aggregated according to the truth value and weight of each fuzzy IF−THEN rule and "defuzzified" by the Mamdani method to produce a single-number *fluence change*. The input−output space defined by the fuzzy logic propositions and membership functions is shown in Figure 1D.

The fuzzy logic control system was used not only for determining the required laser fluence change from one five-shot spectrum to the next but also for deciding if a five-shot spectrum should be discarded or added to the sum spectrum. An acceptable five-shot spectrum was defined as one for which the laser fluence change was evaluated to be small, i.e., in the range from −1.1% to +1.1%, indicating that the signal intensity was within the dynamic range of the data acquisition system and that the peak resolution was acceptable. Thus, if a five-shot spectrum was found to be acceptable by means of the laser fluence being changed only marginally, then this five-shot spectrum was added to the sum spectrum. If the laser fluence had to be changed by more than 1.1% in either direction, then the five-shot spectrum was discarded. In the current system, the total number of laser shots accumulated for each sample deposit was predefined in a sample setup sheet. For example, 60 shots were required for synthetic peptide samples. An alternative would be to implement a software routine that terminates acquisition when a certain ion intensity or signal-to-noise ratio has been reached.

Since a sample deposit becomes depleted after irradiation with a number of laser pulses, a script was written to continuously move the sample probe in a spiral pattern relative to the laser beam. Thus, a new spot of the sample deposit was irradiated for each five-shot spectrum. For any one sample deposit, a significant area is scanned to ensure that the resulting MALDI mass spectrum is representative and reproducible. This approach works very well for fast-evaporation matrix surfaces but also for homogeneous layers of matrix crystallites made by the dried-droplet method.

Performance of Fuzzy Logic Control of Static Extraction MALDI-TOFMS. The fuzzy logic control system for MALDI-MS was tested on a variety of peptide samples. Initial testing was performed by static extraction MALDI-MS analysis of a synthetic peptide (55 fmol) deposited on a fast-evaporation matrix surface of 4HCCA.

During automatic data acquisition, individual mass spectra are transferred from the storage oscilloscope to the LaserOne software. A script determines the ion intensity (range of 0−40 mV) and mass resolution of the base peak in the m/z range 600−4000 of an averaged five-shot spectrum. This information is transferred to the fuzzy logic engine, which quantitatively determines the required change of laser fluence (−3% to +3%) and forwards this information to the stepper motor controlling the laser attenuation filter. As mentioned above, a five-shot spectrum is added to the sum spectrum if the attenuation is small (−1.1% to +1.1%). The data evaluation process takes less than 1 s per five-shot spectrum.

Table 1. Results Obtained by Analysis of 12 Identical Samples by Automated and Manual MALDI-TOF[a]

automatic				manual			
no.	*m/z*	resolution	S/N	no.	*m/z*	resolution	S/N
15	1278.63	5000	27	15	1278.77	3300	30
16	1278.77	4000	30	16	1278.65	3300	28
17	1278.85	3300	20	17	1278.79	3300	15
18	1278.75	5000	15	18	1278.74	5000	11
19	1278.84	3300	20	19	1278.76	6600	17
20	1278.71	4000	20	20	1278.76	3300	13
21	1278.74	5000	17	21	1278.75	2900	13
22	1278.83	4000	20	22	1278.68	4000	23
23	1278.74	3300	22	23	1278.68	3300	11
24	1278.73	2500	21	24	1278.73	3300	25
25	1278.71	4000	35	25	1278.72	4000	25
26	1278.73	3300	25	26	1278.69	6600	30
av	1278.75	3900	23		1278.73	4100	20
SD	0.06	800	6		0.04	1300	7

[a] Aliquots (approximately 55 fmol) of synthetic peptide (RGITVAGKTYGR, calculated *m/z* 1278.73) were deposited on 12 matrix surfaces made by the fast evaporation method. The samples were analyzed in the automated mode and in the manual mode by static extraction MALDI-TOF. Spectra were mass calibrated using low *m/z* matrix ion signals as internal standards (see Experimental Section).

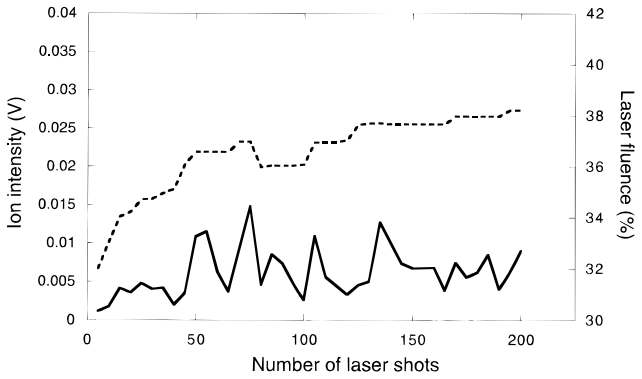

Figure 2. Performance of fuzzy logic feedback control system for automated MALDI analysis of a synthetic peptide (RGITVAGKTYGR, M_r = 1277.72; 55 fmol) deposited onto a fast-evaporation matrix surface. The ion intensity (—) and laser fluence setting (- - -) are plotted as functions of the number of laser shots.

The performance of the fuzzy logic feedback control system for obtaining a peptide mass spectrum as a function of the number of laser pulses is shown in Figure 2. The peptide ion signal intensity and the laser fluence setting are plotted as a function of the number of five-shot spectra. The laser fluence is initially being increased because of low ion signal intensity, but it rapidly reaches a stable level of approximately 37%, around which it fluctuates only slightly in order to keep the ion signal intensity of the base peak in the range of 5−25 mV and the resolution above 1500 as specified in the fuzzy membership functions for static extraction MALDI. Note that an increase in ion intensity above the "threshold level" immediately results in a down-regulation of the laser fluence. This demonstrates that the control system is able to keep the laser fluence at the optimal level without drifting off and without major fluctuations. The software automatically discarded five-shot spectra which required too large a change in laser fluence, as described above.

The data obtained by manual and automated MALDI analysis of 12 identical samples are compared in Table 1. The spectra were acquired as the sum of signals generated by 60 shots, i.e., the sum of 12 selected five-shot spectra. The fwhm resolution of all these spectra was always better than 2500, typically around 4000, which is also the case for manually obtained static extraction MALDI mass spectra of the same samples (Table 1). This shows that the automatic system successfully keeps the laser fluence close to the threshold level, where ion intensity is relatively low and peak resolution is high. On average, 70−75 shots were fired to obtain the 60-shot spectra whose results are listed in the table. The signal-to-noise ratio of the analyte ion signal fluctuated 25%−35% from spectrum to spectrum for both automatically and manually obtained spectra. This test on 12 identical samples thus showed that the automated system reproducibly produced high-quality spectra comparable to those obtained by a human operator.

The performance of the automatic mass calibration script in the case of static extraction MALDI is documented in Table 1 as well. Low-*m/z* matrix ion signals were used for internal calibration. In no case was the peptide mass deviation from the mean larger than 0.12 Da, and the standard deviation was 0.06 Da. This mass accuracy is only slightly inferior to that obtained by manual MALDI analysis of the same samples (largest deviation of 0.07 Da and a standard deviation of 0.04 Da, Table 1).

We next tested how the fuzzy logic feedback control would perform for MALDI-MS of challenging samples. Peptide mass mapping of unseparated peptide mixtures is an obvious application of MALDI but requires careful control of data acquisition because of the large dynamic range and the low signal strength of some of the mixture components. This applies particularly to the analysis of contaminated mixtures such as those obtained after in-gel digestion of proteins. Therefore, the system performance was tested on an aliquot of a peptide mixture derived from in-gel tryptic digestion of 1 pmol of bovine serum albumin. Three percent of the sample solution, i.e., a maximum of 30 fmol/peptide, was applied to a fast-evaporation matrix surface.

The peptide mass map (200 laser shots) was generated by automated MALDI as the sum of selected five-shot spectra obtained from different spots on the sample deposit, using the intensity and resolution of the base peak as the input for fuzzy logic control of the laser attenuation. Note that this base peak may differ from one five-shot spectrum to the next due to heterogeneities in the sample deposit or ion suppression effects.

The MALDI peptide mass maps obtained in automatic and manual data acquisition modes are compared in Figure 3. Little systematic difference in mass accuracy or resolution is evident

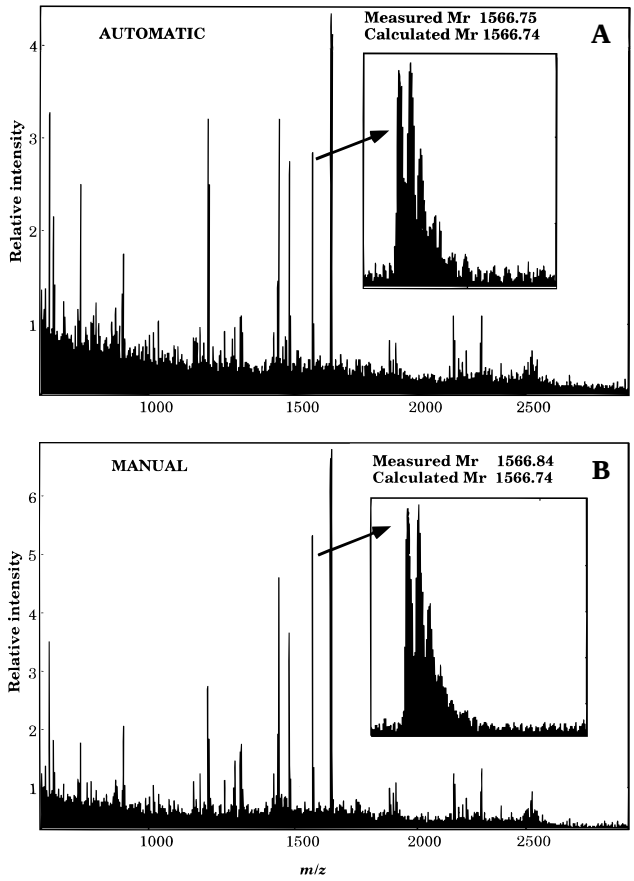

Figure 3. Peptide mass mapping by automatic and manual MALDI-TOFMS. An aliquot of an in-gel tryptic digest of 1 pmol of BSA was loaded onto a fast-evaporation matrix surface. Peptide mass maps were acquired by static extraction MALDI-TOFMS in (A) automatic mode and (B) manual mode. The insets show the mass resolution and mass accuracy of the M_r = 1566.7 peptide ion signal.

from the two panels. The signal-to-noise ratio was only slightly lower in automatic mode compared to that in manual mode. These and other experiments demonstrated that automatic peptide mass mapping by MALDI is routinely possible, even at very low sample levels.

Automatic MALDI Mass Analysis of HPLC Fractions. Reversed-phase HPLC with UV detection is widely used for isolation of peptides prior to Edman sequencing and for comparative peptide mapping of proteins. Mass spectrometric analysis of eluting peptides can be performed on-line by direct interfacing to electrospray MS (LC-MS) or off-line by MALDI-MS analysis of collected fractions. Mass spectrometric analysis often reveals heterogeneities or impurities in peptide fractions that were judged to be pure on the basis of the UV-trace. However, MALDI-MS analysis of many HPLC fractions is a time-consuming and repetitive task that would gain considerably more popularity by being automated. To test the reliability and throughput of our automated MALDI system, a batch of HPLC fractions was analyzed.

Reversed-phase (RP) HPLC was used to separate tryptic peptides derived from reduced and *S*-alkylated bovine G_{a0} subunit. To be able to retrieve high molecular weight proteolytic fragments, a 1 mm i.d. alkyl-C_4 reversed-phase column was used. Eluting peptides were manually collected in 78 fractions, aliquots of which were deposited on premade fast-evaporation matrix surfaces. The samples were subsequently mass analyzed in batches of 26 by automated static extraction MALDI-MS (100 laser shots/spectrum). The mass spectra were automatically mass calibrated using the matrix ion signals as internal mass standards. No problems with the fuzzy control of the laser fluence during the analysis were encountered, and automated MALDI mass analysis of all of the 78 fractions was finished in less than 4 h. We estimate that the same analysis performed by a skilled operator would take significantly longer because of the time required for manual acquisition and mass calibration.

The measured peptide masses were assigned to the protein sequence on the basis of the predicted (calculated) tryptic peptide masses. A representative subset of the acquired data for the 78 HPLC fractions is presented in Table 2. The MALDI spectrum quality (signal-to-noise ratio and resolution) was excellent for all peptide-containing HPLC fractions. The automatic mass calibration routine also performed well: mass accuracy was better than 0.03% for all peptides where matrix ions could be assigned. In a few spectra (data not shown), the intense peptide signal suppressed the matrix ions, and for these the default (external) mass calibration was used (this problem could be solved by a variable detector bias gating).

The experiment had been designed to preferentially collect hydrophobic peptides by the use of a C_4 column. Even so, the tryptic peptides that were assigned to the bovine G_{a0} polypeptide constituted more than 65% of the amino acid sequence, including the late-eluting myristoylated N-terminal peptide (fraction 69) and some of the larger tryptic peptides from the hydrophobic C-terminal domain (fractions 57, 58, and 70 in Table 2).

Performance of Fuzzy Logic Control of Delayed Extraction MALDI. During the course of this study, we installed a delayed extraction (DE) ion source on our MALDI instrument. Since delayed extraction MALDI produces "cooler" ions than static extraction MALDI, laser fluence control should become less critical. However, heterogeneities in the analyte/matrix crystallites still generate varying ion signal intensities when rastering the laser beam across a sample deposit. This is particularly the case for subpicomole amounts of sample. Real-time control of laser fluence was, therefore, also found to be necessary for DE MALDI. To take into account the improved mass resolution obtained with delayed extraction MALDI, the fuzzy control system was slightly modified by changing the membership function associated with rule no. 6 (*resolution*, see above and Figure 1). With this minor change, the system performed as robustly for delayed extraction MALDI as it did for static extraction MALDI.

For automated MALDI analysis, homogeneous analyte/matrix deposits, such as those made by the fast-evaporation technique, have the advantage of generating analyte ion signals from almost any position on the deposit. The classical dried-droplet method leads to more heterogeneity due to the less uniform distribution and larger size of crystallites, and it was therefore of interest to test automated DE MALDI for the analysis of peptide samples prepared by this method. A solution of three peptides was mixed with matrix solution, deposited on the probe, and allowed to crystallize. The sample was then analyzed by automated delayed extraction MALDI reflector TOF mass spectrometry (Figure 4). As expected, ion signals generated by moving the laser beam across this preparation fluctuated more than ion signals from samples prepared by the fast evaporation method (compare Figure 2 and Figure 4A). Nevertheless, the fuzzy logic control system was able to rapidly and efficiently correct the laser fluence so that

Table 2. Automatic MALDI-TOF Mass Analysis of Fractions Collected from a C_4 RP-HPLC Separation of Tryptic Peptides Derived from a Bovine G_{a0} Protein[a]

fraction no.	M_r measd	M_r calcd	ΔM	residues	sequence[b]
5	814.55	814.49	0.06	17−23	AIEKNLK
7	1066.47	1066.49	−0.02	145−153	EYQLNDSAK
19	707.57	707.35	0.22	272−277	DLFGEK
22, 23	890.44	890.46	−0.02	198−205	LFDVGGQR
23	1585.81	1585.87	−0.06	17−31	AIEKNLKEDGISAAK
33	1077.61 and 1391.69	1077.63	−0.02	271−279	KDLFGEKIK or DLFGEKIKK
	1391.69	1391.74	−0.05	198−209	LFDVGGQRSERK
42	1838.74	1838.82	−0.082	243−257	MHESLMLFDSICNNK
49	2150.30	2150.20	0.10	67−85	QYKPVVYSNTIQSLAAIVR
51−54	1984.11	1983.89	0.22	113−129	MEDTEPFSPELLSAMMR
55	3104.18	3104.64 (av)	−0.46	24−53	EDGISAAKDVKLLLLGAGESGKSTIVKQM*K
56	3531.78	3531.88 (av)	−0.10	280−310	KSPLTICFPEYTGSNTYEDAAAYIQAQFESK
57, 58	3403.94	3403.70 (av)	0.24	281−310	SPLTICFPEYTGSNTYEDAAAYIQAQFESK
65	3520.68	3519.86	0.82	179−208	VKTTGIVETHFTFKNLHFRLFDVGGQRSER
69	1231.70	1231.63	0.07	1−9	myristoyl-GCTLSAEER
70	3727.30	3727.23 (av)	0.06	317−348	EIYCHMTCATDTNNIQVVFDAVTDIIIANNLR(G)

[a] A representative subset of the data obtained by analysis of 78 fractions is shown. Total MALDI analysis time was 4 h. The spectra were automatically mass calibrated using matrix ion signals, and monoisotopic masses were determined unless otherwise stated. Peptide masses were assigned to the polypeptide sequence on the basis of calculated molecular weights of predicted tryptic peptides.

it was kept close to the threshold for peptide ion production throughout the experiment (Figure 4A). No drifting or uncontrolled fluctuation was observed.

The automatically acquired mass spectrum of the peptide mixture is shown in Figure 4B. The signal-to-noise ratio was about 50:1, and the mass resolution on the peptide peak was better than 13 000 (Figure 4B, inset). Automatic calibration using the *m/z* 1278.728 and 2374.300 peptide ion signals as internal standards allowed the *m/z* of the middle peptide ion to be determined as 1714.941, which is only 0.016 Da less than the calculated *m/z* 1714.957, corresponding to a deviation of 10 ppm. Averaging the masses determined from four different sample deposits improved the mass accuracy to 3 ppm. Thus, very high mass accuracy is also achievable by automatic data acquisition and automatic internal calibration, even when employing the dried-droplet sample preparation method.

Protein Identification by Automated MALDI Peptide Mass Mapping with On-Line Database Searching. Protein identification by mass spectrometry is currently the most sensitive and versatile method for linking expressed proteins (the proteome) of an organism to the corresponding genome.[27] The proteins are usually separated and visualized on one- or two-dimensional polyacrylamide gels. A main interest of our laboratory is to develop integrated methods for the handling and analysis of such samples. Delayed extraction MALDI-MS is the preferred method for initial screening of tryptic digests of gel-isolated proteins in our laboratory.[23] This is because MALDI-MS is simpler and faster than other analytical methods such as amino acid analysis, Edman sequencing, LC-MS, and MS/MS and consumes only a small percentage of the sample. The automation of the key components of MALDI-MS, i.e., data acquisition and mass calibration, should make it possible to automate the whole process of protein identification—from data acquisition to retrieval of the matching entry from a sequence database.

Scripts were written to connect the various software components, including a database file for sample specification and reporting of search results, the automated MALDI control software, and the protein database search software. The overall

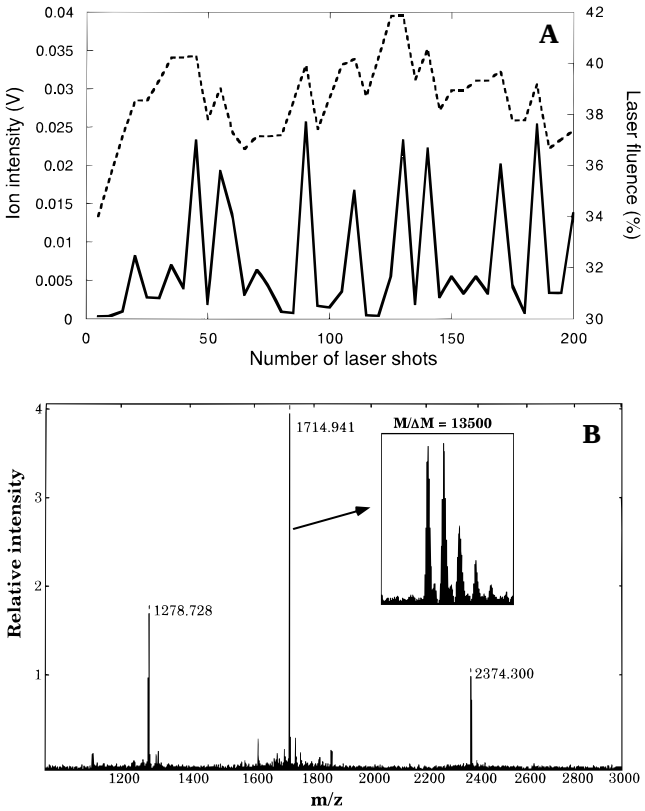

Figure 4. Automated delayed extraction MALDI analysis of a peptide mixture prepared with the dried-droplet method. (A) Performance of the fuzzy logic feedback control system for data acquisition. The ion intensity (solid line) and laser fluence setting (broken line) are plotted as a function of the number of laser shots. Major changes in ion intensity result in significant adjustment of the laser fluence. (B) The mass spectrum acquired during the experiment documented in (A). The signal-to-noise ratio and mass resolution (inset) are comparable to those obtained by manual acquisition. Automatic mass calibration using the *m/z* 1278.728 and 2374.300 peptides determined the *m/z* of the middle peptide to be 1714.941 (calculated *m/z* 1714.957).

(27) Shevchenko, A.; Jensen, O. N.; Podtelejnikov, A. V.; Sagliocco, F.; Wilm, M.; Vorm, O.; Mortensen, P.; Shevchenko, A.; Boucherie, H.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996,** *93,* 14440−14445.

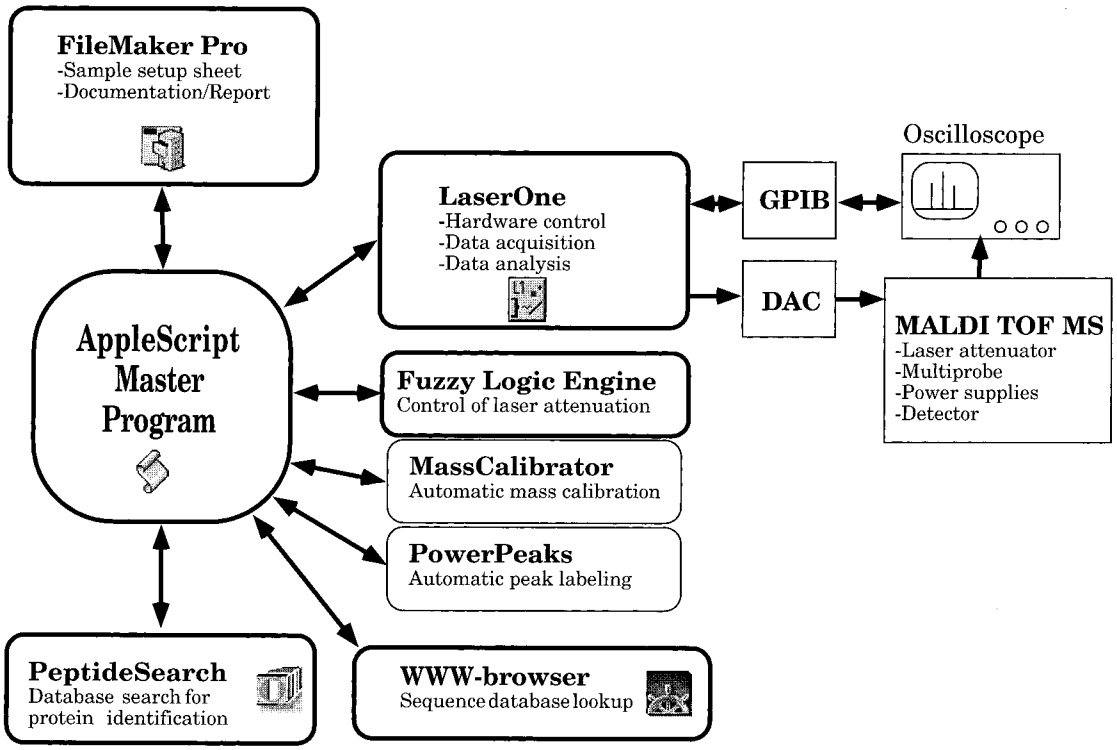

Figure 5. Software components used in the complete automation of MALDI-TOFMS for protein identification. An Applescript master program controls the data flow between the individual applications used for sample specification, data acquisition, mass calibration, peak labeling, database searching, and reporting. See main text for further details.

setup is shown in Figure 5. A master script controls the data flow between the various software components as follows: a MALDI mass spectrum was automatically acquired by the fuzzy logic control system, and the resulting mass spectrum was automatically mass calibrated. The list of measured tryptic peptide masses was transferred to the PeptideSearch program, and a database search was started. Searching a nonredundant database with more than 200 000 protein entries with a set of tryptic peptide masses took less than 15 s on the computer used. The highest scoring protein matches were then returned to the database program for subsequent inspection and evaluation. Direct linking to protein sequence databases, such as SwissProt, for further evaluation was implemented via a World Wide Web (www) interface.

With this setup, a protein identification experiment can be finished in only 5−6 min. With a multiprobe sample inlet, a throughput of 10−12 samples/h is thus currently possible. An example of automated MALDI-MS protein identification is shown in Figure 6. The peptide mass map shown was generated from an aliquot of a tryptic in-gel digest of a yeast protein spot excised from a single 2D gel. Automated MALDI analysis with database searching retrieved *S*-adenosylmethionine synthase 2 (SwissProt accession number P19358) as the highest scoring candidate, with 14 matching peptides within 50 ppm mass deviation. The next protein matched with only five peptides and was furthermore from a different species. Subsequent evaluation of the peptide mass data by a "second pass search" [23] revealed that the automatic mass calibration using a matrix ion signal (m/z 855.1) and a trypsin autodigestion peptide signal (m/z 2163.057) resulted in a peptide mass accuracy better than 30 ppm throughout the m/z range 1000−3500 (Figure 6). Four additional ion signals of low intensity could be assigned to the sequence resulting in a total of 18 peptide matches. These peptides covered 48% of the protein sequence and also contain two terminal peptide sequence tags[28] which further confirm this identification. This example demonstrates that the performance of the automated MALDI system is sufficient to identify gel-isolated proteins on the basis of their peptide mass maps without human intervention (the most demanding experiments, however, are still repeated by a skilled operator if the results are not conclusive). In large-scale applications, the automated MALDI system is now routinely used to screen all peptide mixtures after in-gel digestion of proteins spots excised from Coomassie- and silver-stained 1D and 2D gels.

CONCLUSION AND PROSPECTS

We have demonstrated that automation of MALDI-MS for the acquisition of mass spectra of biomolecules is feasible and relatively simple when using a fuzzy logic-based feedback control system. The performance of the automated MALDI-TOF mass spectrometer is comparable to that of a skilled human operator. During analysis of hundreds of samples, including synthetic peptides, oligonucleotides, HPLC fractionated peptides, and peptide mixtures produced by in-gel digestion of proteins, no problems with instability in the form of major fluctuations or drifting of the fuzzy logic control system have been encountered. The fuzzy logic control system works equally well for static extraction and delayed extraction MALDI and for samples prepared with the dried-droplet method and the fast-evaporation method of sample deposition. Automatic mass calibration and peak labeling routines performed well and allowed peptide masses to be assigned to their calculated counterparts with high confidence. Complete automation of protein identification from acquisition of a MALDI peptide mass map to database searching with the peptide masses was demonstrated.

(28) Jensen, O. N.; Vorm, O.; Mann, M. *Electrophoresis* **1996**, *17*, 938−944.

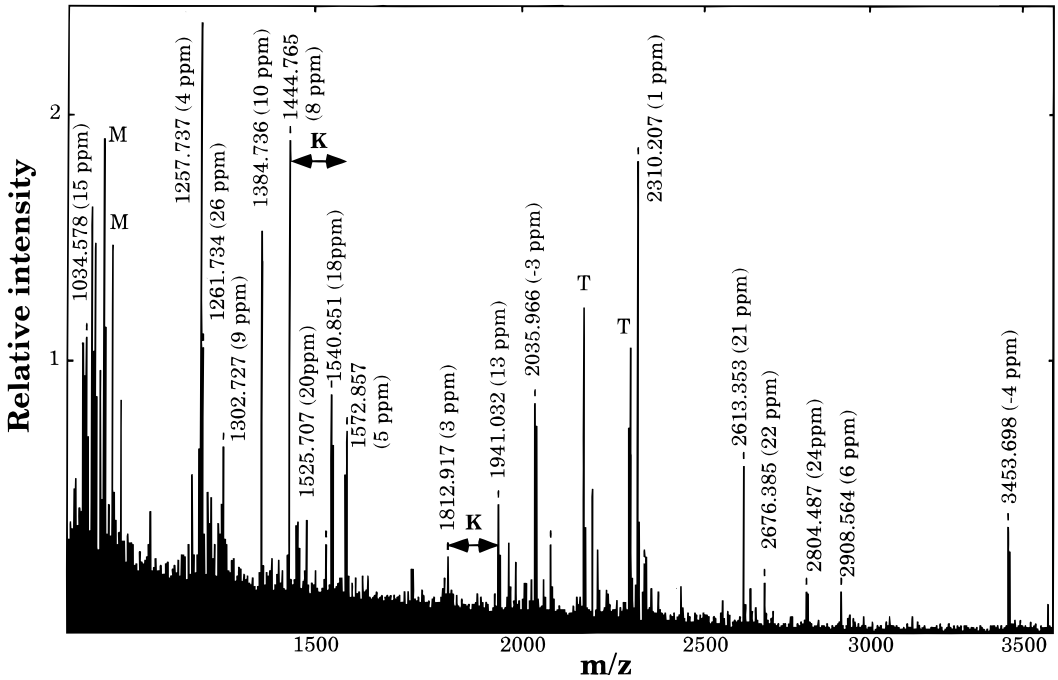

Figure 6. Protein identification by automated MALDI-TOF. An aliquot (0.3 μL) of supernatant from in-gel digestion of a gel-spot excised from a silver-stained 2D gel was deposited on a fast-evaporation matrix surface containing nitrocellulose. The peptide mixture was automatically mass analyzed, and the list of peptides was automatically searched in a protein database. Yeast *S*-adenosylmethionine synthetase 2 (SwissProt P19358) was retrieved. The 18 matching peptides and their mass deviation from the calculated masses are labeled in the spectrum. Two terminal sequence tags due to adjacent tryptic cleavage sites are labeled with arrows. Matrix-related ion signals (M) and trypsin autodigestion peptides (T) are indicated. Based on the intensity of staining and the ion intensity in the mass spectrum, the protein amount in the gel was estimated to be 1−2 pmol.

Automated MALDI-MS will prove useful in high-throughput applications such as screening and quality control of synthetic products and libraries and HPLC fractions. Automated MALDI peptide mass mapping in combination with on-line protein sequence database searches facilitates large-scale protein identification and characterization with a throughput of 10 or more samples per hour.

The stability, flexibility, and ease of implementation of a fuzzy logic system for real-time control of MALDI suggests other uses of this technique in the field of mass spectrometry. For example, we have now expanded the fuzzy logic system to also control the MALDI probe movement as a function of ion intensity. A human operator decides to move the irradiated spot little when a good signal is obtained but will move it by a larger distance when no or few ions are generated. This strategy was easily implemented in the fuzzy logic system by adding a few rules of the type "IF Low *intensity* THEN *move* Much" and by adding the corresponding membership function for probe movement. Other uses in mass spectrometry do not need to be confined to the area of control. For example, a measured mass is typically said to match a calculated mass if it is within a certain "crisp" mass tolerance. It would be more appropriate to "soften" this criterion according to fuzzy logic precepts. Thus, a measured mass would match the calculated one according to a membership function. Work is underway in our laboratory to further exploit the potential of fuzzy logic for instrument control and for data analysis in MALDI and electrospray mass spectrometry.

ACKNOWLEDGMENT

We thank the other members of the Peptide & Protein group for fruitful discussions. Dr. Bernd Nürnberg and Torsten Exner (Freie Universität Berlin) provided the bovine G_{a0} protein. O.N.J. is supported by a fellowship from the EU Biotechnology program.

Received for review November 22, 1996. Accepted February 13, 1997.[⊗]

AC961189T

[⊗] Abstract published in *Advance ACS Abstracts,* April 1, 1997.