Identifying Proteins from Two-Dimensional Gels by Molecular Mass Searching of Peptide Fragments in Protein Sequence Databases

WJ Henzel, TM Billeci, JT Stults, SC Wong, C Grimley, and C Watanabe

PNAS 1993;90;5011-5015
doi:10.1073/pnas.90.11.5011

Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 5011–5015, June 1993
Biochemistry

Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases

(mass spectrometry/*in situ* digestion/two-dimensional electrophoresis)

WILLIAM J. HENZEL*, TODD M. BILLECI, JOHN T. STULTS, SUSAN C. WONG, CHRISTOPHER GRIMLEY, AND COLIN WATANABE

Department of Protein Chemistry, Genentech, Inc., South San Francisco, CA 94080-4990

Communicated by David Botstein, February 23, 1993

ABSTRACT **A rapid method for the identification of known proteins separated by two-dimensional gel electrophoresis is described in which molecular masses of peptide fragments are used to search a protein sequence database. The peptides are generated by *in situ* reduction, alkylation, and tryptic digestion of proteins electroblotted from two-dimensional gels. Masses are determined at the subpicomole level by matrix-assisted laser desorption/ionization mass spectrometry of the unfractionated digest. A computer program has been developed that searches the protein sequence database for multiple peptides of individual proteins that match the measured masses. To ensure that the most recent database updates are included, a theoretical digest of the entire database is generated each time the program is executed. This method facilitates simultaneous processing of a large number of two-dimensional gel spots. The method was applied to a two-dimensional gel of a crude *Escherichia coli* extract that was electroblotted onto poly(vinylidene difluoride) membrane. Ten randomly chosen spots were analyzed. With as few as three peptide masses, each protein was uniquely identified from over 91,000 protein sequences. All identifications were verified by concurrent N-terminal sequencing of identical spots from a second blot. One of the spots contained an N-terminally blocked protein that required enzymatic cleavage, peptide separation, and Edman degradation for confirmation of its identity.**

The identification of a purified protein is necessary in many areas of biochemical research. As the resolution and sensitivity of purification tools increase, the demand for protein sequencing increases. For example, a single high-resolution two-dimensional polyacrylamide gel can separate hundreds of proteins (1, 2). Identification of all the resolvable proteins on a two-dimensional gel by conventional protein sequencing is a daunting task. The correlation of DNA from large-scale sequencing projects with their protein products will continue to place increasing demands upon protein sequencing.

Proteins that are N-terminally blocked present an additional challenge since they cannot be directly sequenced by Edman degradation. Blockage may occur by posttranslational modification during protein synthesis or during purification. Many intracellular proteins have been reported to be N-terminally acetylated (3). In order to obtain internal sequence on a blocked protein, 50–100 pmol of material is usually required. The blocked protein is chemically or enzymatically cleaved. The peptides are then separated by HPLC and sequenced, a process which can take 3–4 days. In addition, proteins initially thought to be novel may, after purification and sequencing, be found already to exist in the protein sequence database. As a result, a significant fraction of sequencer time is spent simply identifying known proteins.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

An alternative method of identifying known proteins was proposed in 1989.† This method utilized a computer program (FRAGFIT) that identified a protein by matching two or more molecular masses of peptide fragments obtained from chemical or enzymatic cleavages with all fragment masses in a protein sequence database. At that time, fast-atom bombardment and plasma desorption mass spectrometry were the most common ionization methods for analysis of peptide mixtures (4). Subsequently, matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (5–8) has become a more useful method for the analysis of complex mixtures due to its high mass range, high sensitivity, and relative tolerance to common buffer components.

Several recent studies have described *in situ* digestion of proteins electroblotted onto poly(vinylidene difluoride) (PVDF) membranes (9–13). These techniques utilize PVDF membranes as a support for protein immobilization. This allows reactions such as reduction and alkylation to be performed at the low picomole level without significant protein loss. By utilizing *in situ* digestion on a PVDF membrane with MALDI analysis, we have developed a rapid method to identify known proteins from two-dimensional gels.

MATERIALS AND METHODS

Two-Dimensional Gel Electrophoresis. *Escherichia coli* cells expressing human growth hormone were lysed with sonication in 8 M urea/2% (wt/vol) 3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate (CHAPS)/2% (vol/vol) 2-mercaptoethanol/8 mM phenylmethanesulfonyl fluoride, pH ≈7. The solubilized proteins were centrifuged at 12,000 × *g* in an Eppendorf microcentrifuge and the supernatant was diluted with sample buffer [8 M urea/2% 2-mercaptoethanol/2% (vol/vol) Pharmalyte pH 3–10 (Pharmacia)/2% CHAPS/0.003% bromophenol blue, pH 7]. Two-dimensional gel electrophoresis was performed on a Pharmacia Multiphor II electrophoresis apparatus, using precast pH 4–6 immobilized gradient strips for the first dimension and SDS/8–18% polyacrylamide gradient gels for the second dimension. After electrophoresis in the second dimension, the gels were equilibrated for 5 min prior to electroblotting in 10 mM Caps buffer, pH 11.0/20% methanol. Electroblotting onto PVDF membranes (Immobilon-PSQ, Millipore) was carried out for 45 min at 250-mA constant current in a Bio-Rad Trans-Blot transfer cell (14). The PVDF membrane was stained with 0.1% Coomassie blue R-250 in 40% methanol/0.1% acetic acid for 1 min and

Abbreviations: MALDI, matrix-assisted laser desorption/ionization; PVDF, poly(vinylidene difluoride).

*To whom reprint requests should be addressed.

†Henzel, W. J., Stults, J. T. & Watanabe, C. (1989) Poster presentation, Third Symposium Protein Society, July 29–August 2, Seattle, WA.

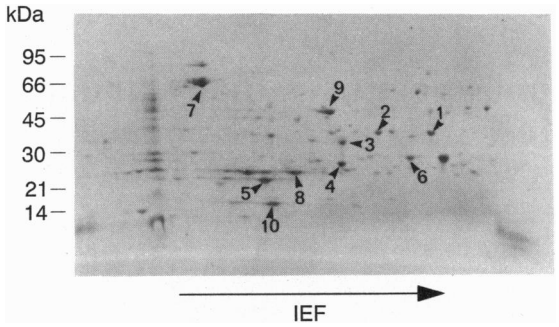

FIG. 1. *E. coli* proteins separated by isoelectric focusing (IEF) in a pH 4–6 Pharmacia precast gel in the first dimension and by electrophoresis in an SDS/8–18% polyacrylamide gel in the second dimension. The proteins were blotted onto PVDF membranes (Immobilon-PSQ, Millipore) and stained with Coomassie brilliant blue.

destained for 2–3 min with 10% acetic acid/50% methanol. The membrane was thoroughly washed with water and allowed to dry before storage at −20°C.

***In Situ* Reduction, Alkylation, and Digestion.** Single protein spots were excised from the PVDF membrane and wetted with 1 μl of methanol. The blots were reduced with 100 μl of 0.5 M Tris·HCl, pH 8.5/10% acetonitrile/5 mM EDTA/7 mM dithiothreitol for 1 hr at 45°C. The solution was then cooled to room temperature and 10 μl of 200 mM iodoacetic acid/0.5 M NaOH was added. The alkylation reaction was allowed to continue in the dark for 20 min at room temperature and the blots were then immediately rinsed with water. Blots were then incubated with 200 μl of 0.25% polyvinylpyrrolidone 40 in 0.5 M acetic acid on a shaker at room temperature for 20 min to prevent protease adsorption to the PVDF membrane. Residual polyvinylpyrrolidone 40 was removed by rinsing the blots with water and then with 20% acetonitrile. Blotted proteins were digested in 50 μl of 0.1 M ammonium bicarbonate/10% acetonitrile with 0.2 μg of Promega modified trypsin at 37°C for 17 hr (10). The supernatant was removed and used for mass spectrometric analysis and capillary HPLC peptide mapping.

Capillary HPLC Peptide Mapping. Peptides generated from *in situ* digestion of spots were separated on a C_{18} capillary column (0.32 mm × 150 mm; LC Packing, San Francisco) as described (15).

Protein Sequencing. Automated protein sequencing was performed on 470A and 477A Applied Biosystems sequencers equipped with on-line phenylthiohydantoin analyzers. Electroblotted proteins were sequenced in the Blott cartridge. Peaks were integrated with Justice Innovation software using Nelson Analytical 760 interfaces. Sequence interpretation was performed on a VAX 8650 (16).

Mass Spectrometry. A dried aliquot of the tryptic digest, reconstituted in 0.5 μl of 10% acetonitrile/0.1% trifluoroacetic acid, was mixed on the sample probe tip with 0.5 μl of 50 mM fucose and 0.5 μl of a saturated solution of 2,5-dihydroxybenzoic acid (19). Mass spectra were obtained with a Vestec (Houston) LaserTec ResearcH laser desorption linear time-of-flight mass spectrometer with a 337-nm VSL-337 ND nitrogen laser (Laser Science, Cambridge, MA). Each spectrum was the sum of the ion intensities from 32 laser shots. To achieve sufficient mass accuracy, internal mass standards were added subsequently to the sample by dissolving the crystals on the probe tip with 0.5 μl of a solution containing 200 fmol of Ala-Gly-[Arg^8]vasopressin and 200 fmol of bovine insulin in 10% acetonitrile/0.1% trifluoroacetic acid. The sample was then dried and reanalyzed. At these low sample levels, some suppression of ion signal was noted upon addition of the internal standards. Therefore, following calibration of the mass axis with internal standards, the measured masses were transferred to the original spectrum, and it was recalibrated to achieve more accurate determination of all peptide masses.

FRAGFIT Algorithm. Input for the FRAGFIT program consists of the following parameters: a list of peptide masses, the protease or cleavage reagent, a mass tolerance, a protein molecular weight range, and the number of allowed mismatches. The program‡ scans the database, generates sequence fragments based on the specified protease, and computes the molecular masses of the fragments. The terminal methionine of a CNBr fragment is assumed to be a homoserine lactone, and all cysteine residues are assumed to be carboxymethylated. If a fragment mass lies within the specified tolerance of a mass on the user's list of masses, the match is recorded. If the number of matching fragments for a protein exceeds the minimum number of matches, that protein is added to the list of matching proteins. A list of matching proteins sorted by the number of matches is output to a file.

The protein database used by the program consists of a combination of several widely available databases, supplemented with ≈2000 sequences from the primary literature. The database includes Swiss-Prot (Release 23, August 1992), the Protein Identification Resource (Release 33, June 30, 1992), and a translation of GenBank (Release 73.1, October 1, 1992, plus daily updates). The databases are merged and the accession numbers are used to remove duplicate entries. The resulting database contains 18.8 million residues, representing over 91,000 entries. Approximately 20% of the entries

‡The C-language source code for FRAGFIT can be obtained via E-mail from ckw@gene.com. The program has been used on a DEC 8650 running the Reno version of Berkeley Unix. Other systems may require slight modifications of the source.

Table 1. N-terminal sequence analysis of spots 1–10

Spot no.	Initial yield, pmol	Protein	N-terminal sequence
1	2.5	Cysteine synthase A	SKIFEDNS
2	1.1	Malate dehydrogenase	MKVAVLGAAGGIGQAL
3	1.2	Cytidine deaminase	MHPRFQTAFAQLADNLQ
4	2.8	Purine nucleoside phosphorylase	ATPHINAEMGDFADVVLMPG
5	3.5	Growth hormone	FPTIPL
6	1.0	Uridine phosphorylase	SKSDVFHLGLTKNDLQG
7	2.0	60-kDa chaperonin	AAKDVKFGN
8	4.2	Growth hormone	FPTIPL
9	0.5	60-kDa elongation factor	GITINTSHVEYDTXT*
10	11.8	10-kDa chaperonin	MNIRPLHDRVIVKRKEVE
	4.6	Universal stress factor	AYKHILIAVDLSPESKVE

*Obtained from a tryptic digest separated by HPLC. The protein was N-terminally blocked.

in the merged database are duplicates which could not be removed because of a lack of accession numbers.

RESULTS

Proteins extracted from a human growth hormone-producing *E. coli* fermentation were separated by two-dimensional electrophoresis, blotted onto a PVDF membrane, and stained with Coomassie blue (Fig. 1). Ten spots were arbitrarily chosen to reflect a range of molecular weights and isoelectric points. Two identical gels were run. One gel was used for N-terminal sequence analysis. The initial yield obtained from sequence analysis was used to estimate the amount of protein in each spot and the amount used for mass spectrometry (Table 1). The second gel was used for *in situ* tryptic digestion followed by mass spectrometry. Aliquots of the tryptic digests were analyzed by MALDI mass measurement. Amounts ranged from 840 fmol (20% of the total) in spot 8 (the largest amount analyzed) to 75 fmol (3%) in spot 1. Fig. 2 shows selected spectra for the tryptic digest of spots 1, 8, and 9 and a trypsin control. Some autolysis fragments from trypsin were present in the spectra of the spots (masses: 842, 2218, 2301, and 3347 Da). The peak corresponding to a mass of 1798 Da was observed in most of the spectra. This mass may be associated with material from the PVDF membrane, since it is absent from the trypsin solution control and was present in most of the digests. Masses larger than 900 and less than 2500 Da obtained from MALDI mass spectrometry were input into the FRAGFIT program. A molecular mass range based on estimates from the second gel dimension (Fig. 1) was used to limit the size of the proteins searched in the database by the computer program. This value was obtained by increasing the observed molecular mass by 10–20% to allow for potential posttranslational modifications. The program default value of 500 Da was used for the lower mass limit. A summary of the output of the FRAGFIT program obtained from analysis of spots 1–10 is shown in Table 2. This table lists all observed masses that were used by the program and the peptide sequences that matched each mass value. Masses that are listed as "not found" did not match with any of the predicted masses but did not prevent the program from identifying the protein.

Although most peptides in Table 2 contain masses more accurate than ±4 Da, we chose this value to reflect the largest error observed. Larger mass tolerances increase the possibility of obtaining proteins that are unrelated to the sample. Decreasing the mass tolerance results in higher confidence in the matches but also increases the chance of missing a match.

By using this search strategy to analyze the data from the protein spots by FRAGFIT, a single protein was usually identified from the >91,000 proteins in the database. Spots 7, 8, and 9 resulted in proteins from more than one species of protein being identified, since the masses used result from sequences which are invariant between species. Spot 7 was identified as 60-kDa chaperonin with both *E. coli* and human sequences reported. The program identified spot 8 as growth hormones from human and monkey. Two species of elongation factor Tu were also found for spot 9.

Sequence analysis of spot 10 identified a mixture of two proteins: a 10-kDa chaperonin and a universal stress protein (Table 1). The FRAGFIT program identified chaperonin from the masses used but was unable to identify the stress protein. Examination of the stress-protein sequence, which was submitted as an unpublished entry in the Dayhoff database, revealed a total of 12 residues labeled X. All but one peptide longer than 6 residues contained unidentified residues.

N-terminal analysis was able to confirm the identity of all spots, except for spot 9, which was N-terminally blocked. To make a positive identification of spot 9, the tryptic digest of that spot was separated by HPLC on a capillary C_{18} column.

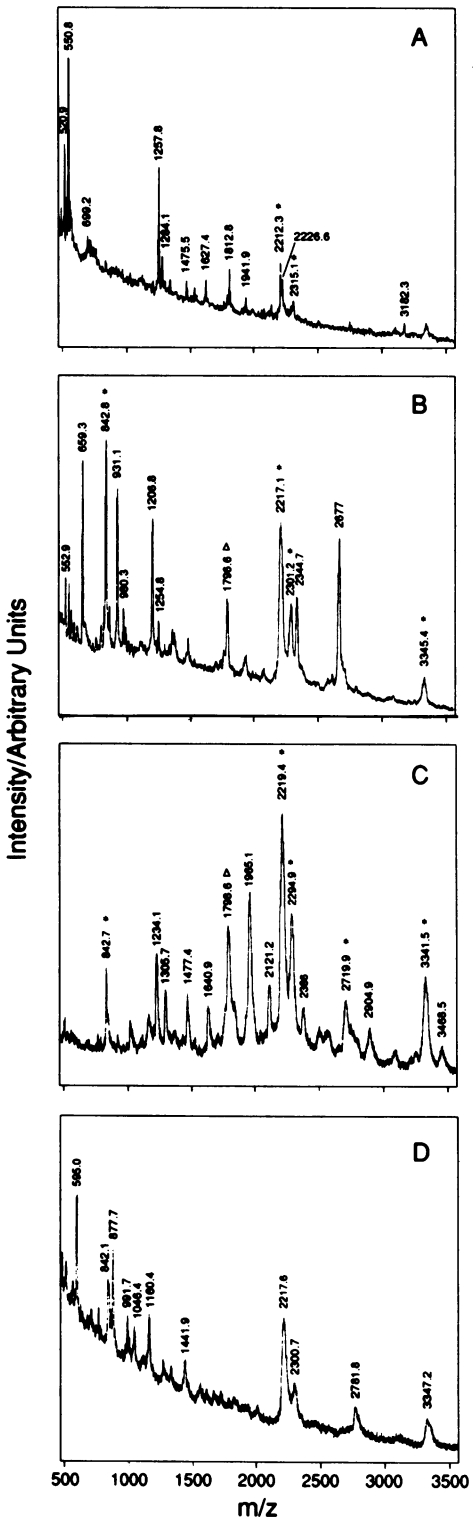

FIG. 2. MALDI mass spectra of *in situ* tryptic digestion of spot 1, cysteine synthase A (*A*); spot 8, human growth hormone (*B*); spot 9, 60-kDa elongation factor (*C*); and a trypsin control (*D*). Asterisks indicate trypsin autolysis fragments. Peaks marked with Δ are unknown contaminants.

Individual fractions were sequenced which matched the sequence of elongation factor Tu (Table 1).

DISCUSSION

Molecular mass searching of peptide fragments provides a rapid method of identifying known proteins separated by

Table 2. Summary of FRAGFIT output of 2-D gel spots 1-10

Spot no.	Percent analyzed (fmol)	Misses allowed	Mass range, kDa	Protein identified	Mass observed, Da	Δ*, Da	Sequences from database
1	3 (75)	1	0.5–45	Cysteine synthase A (34,358.1 Da†)	1257.8	−0.3	ALGANLVLTEGAK
					1284.1	−0.4	NIVVILPSSGER
					1475.5	+0.9	VIGITNEEAISTAR
					1627.4	−0.5	IQGIGAGFIPANLDLK
					1812.8	−0.3	IFEDNSLTIGHTPLVR
					1941.9		Not found
2	20 (220)	0	0.5–40	Malate dehydrogenase (32,387.2 Da†)	1146.5	−3.9	FFSQPLLLGK
					1274.7	−2.8	SDLFNVNAGIVK
					2398.3	−3.3	ALQGEQGVVECAYVEGDGQYAR
3	20 (240)	2	0.5–40	Cytidine deaminase (31,539.4 Da†)	1123.8	+0.6	DYLPDAFGPK
					1123.8	+3.4	GYDYDIQR
					1361.9	+1.3	TPLSNFNVGAIAR
					1538.8	+1.1	QFMNELNSGLDLR
					1701.9	+2.0	ADAPLIQWDATSATLK
4	5 (140)	2	0.5–30	Purine nucleoside phosphorylase (25,949.7 Da†)	1022.5	+0.3	ELITDFGVK
					1096.4	−1.9	QTTFNDMIK
					1145.2	+1.0	THEQTTAAER
					1225.6	+0.1	VGSCGAVLPHVK
					1330.1	+1.6	YIAETFLEDAR
					1388.2	+0.6	ALTICTVSDHIR
					1046.2		Not found
					1479.1		Not found
5	20 (640)	1	0.5–25	Growth hormone (22,127.8 Da‡)	931.2	+0.1	FPTIPLSR
					1207.8	+0.4	NYGLLYCFR
					2266.4	+2.9	SVFANSLVYGASNSDVYDLLK§
					2347.5	+3.9	LHQLAFDTYQEFEEAYIPK
					2680.8		Not found
6	20 (200)	2	0.5–35	Uridine phosphorylase (27,158.9 Da†)	1755.6	+2.7	NDLQGATLAIVPGPDR
					2386.2	−3.6	SIGATTHVGVTASSDTFYPGQER
					3189.7	−3.9	LDGASLHFAPLEFPAVADFECTTALVEAAK
					3338.7	−2.1	AELDGKPVIVCSTGIGGPSTSIAVEELAQLGIR
					1942.8		Not found
7	20 (400)	3	0.5–65	Chaperonin (57,137.1 Da†; 60,512.1 Da‡)	1199.0	−1.5	EMLPVLEAVAK
					1202.3	−3.3	GQNEDQNVGIK
					1568.6	−0.2	AAVEEGVVAGGGVALIR
					1760.6	+1.3	AIAQVGTISANSDETVGK
					1846.1	−0.9	DTTTIIDGVGEEAAIQGR
					1957.7	−1.5	QIVLNCGEEPSVVANTVK
					2404.6	+1.0	ANDAAGDGTTTATVLAQAIITEGLK
					2856.3	+3.2	EGVITVEDGTGLQDELDVVEGMQFDR
					1437.8		Not found
					1942.2		Not found
8	20 (840)	2	0.5–30	Growth hormone (22,127.8 Da‡; 22,210.9 Da¶)	931.1	0	FPTIPLSR
					980.3	+0.1	LFDNAMLR
					1206.8	−0.6	NYGLLYCFR
					2344.7	+1.1	LHQLAFDTYQEFEEAYIPK
					2677.0	+1.1	YSFLQNPQTSLCFSESIPTPSNR
					1254.8		Not found
9	20 (N terminus blocked)	2	0.5–45	Elongation factor (43,182.1 and 41,511.4 Da†; 43,152.1 Da‖)	1234.4	−0.3	GYRPQFYFR
					1305.7	+1.1	TTLTAAITTVLAK
					1965.1	+1.9	ILELAGFLDSYIPEPER
					1965.1	−1.0	ELLSQYDFPGDDTPIVR
					2121.2	+2.7	AIDKPFLLPIEDVFSISGR
					1477.4		Not found
					1640.9		Not found
10	0.5 (60)	2	0.5–30	Chaperonin (10,386.9 Da†)	974.0	+1.9	GEVLAVGNGR
					1154.1	+1.7	MNIRPLHDR
					1201.5	−0.9	SAGGIVLTGSAAAK
					1454.1	−0.6	ILENGEVKPLDVK
					1494.7	−2.0	VGDIVIFNDGYGVK
					1046.6		Not found
					1476.8		Not found

*The difference between the mass found and the calculated mass. A tolerance of 4 Da was used for the analysis of all spots. All masses represent average isotopic masses.

†*E. coli.*

‡Human.

§This sequence obtained from the protein database is incorrect. The order of the sequence NSD should be reversed.

¶Monkey.

‖*Salmonella.*

two-dimensional electrophoresis. This can be achieved by simultaneous digestion of a large number of protein spots. The analysis of 5–10 *in situ* digests by mass spectrometry and computer searching can be completed in <1 hr. A theoretical digest of the protein sequence database is generated each time the FRAGFIT program is run, in order to access the latest version of the database. A typical search using peptide masses is usually complete in <5 min on a VAX 8650. This is considerably faster than that which can be accomplished by conventional protein sequencing. No separation is needed, since the unfractionated peptide digest is placed directly in the ion source of the mass spectrometer. While this may result in suppression of some peptide fragment ions, only three to five fragments with masses between 1 and 2 kDa are needed to correctly identify a protein. The program requires a minimum of two masses in order to perform a search.

MALDI provides high sensitivity for peptide mixtures when operated under the conditions described here, requiring only a fraction of a 1-pmol spot (10–20%) to produce a useful spectrum. If necessary the remainder of the digest can then be separated by capillary HPLC and the peaks can be sequenced by Edman degradation. The addition of carbohydrate to the UV-absorbing matrix (17, 19) is essential to mass analysis of peptide mixtures at this sample level. Other mass spectrometry ionization techniques, such as electrospray, could be utilized (18). However, the spectrum of a complex mixture delivered to the electrospray source by direct infusion may be difficult to interpret. HPLC could be used to separate the components of the mixture, but the analysis time would be significantly longer than that required by MALDI analysis.

Proteins which contain posttranslational modifications present additional challenges. Modified peptides from these proteins cannot be identified by the FRAGFIT program, since current commercial databases do not link posttranslational modifications with protein sequence. In addition, many protein databases are either poorly documented or lack information on these modifications. To identify proteins containing posttranslational modifications, the number of mismatches allowed can be increased to exclude masses corresponding to the modified peptides. If more than several mismatches are necessary for a match, the protein should be positively identified by protein sequencing. Another challenge is presented by proteins which are not completely resolved during the two-dimensional gel separation. Spots that contain more than one protein may also be identified by increasing the number of allowed mismatches. Although spot 10 contained a mixture of two proteins, the program was able to identify one of the proteins in the mixture.

In order to digest an electroblotted protein at the picomole level, high enzyme-to-substrate ratios (1:1) are necessary. This may result in a significant number of autolysis fragments, which can be identified by utilizing a trypsin control. We find that using reductively methylated trypsin significantly reduces autolysis.

Although we have used trypsin to digest the proteins, any specific chemical or enzymatic cleavage could be used. Specific cleavages that result in peptides with higher mass values (e.g., CNBr) enable FRAGFIT to identify a protein with fewer molecular ions. We typically find that two or three molecular ions are sufficient for protein identification by FRAGFIT for CNBr-generated peptides of ≥2 kDa. In general, the longer the peptide, the more likely only one specific match for a molecular ion will be found in the database.

Our best strategy for searching is to start with a mass tolerance of 4 Da and one or two mismatches. The number of mismatches is then increased or decreased to obtain more or fewer matches. All molecular ions that have a good signal-to-noise ratio and are absent from the enzyme control should be utilized as input. A tighter mass tolerance (1–2 Da) permits identification with fewer masses. However, this mass accuracy is difficult to achieve on a routine basis with our current instrumentation. Continued improvements in instrumentation and sample matrices for MALDI should make this mass accuracy possible.

The method described here should significantly increase the speed of identifying known proteins. This method is not a complete substitute for protein sequencing. When the confidence of the FRAGFIT output is low, or conflicting data exist, protein sequencing should be used to confirm the identity of the protein. Other methods of protein cleavage need to be tested for compatibility with MALDI analysis.

We thank R. Vandlen for valuable comments on the manuscript and L. Tamayo for preparation of the figures.

1. Harrington, M. G., Gudeman, D., Zewert, T., Yun, M. & Hood, L. (1991) *Methods: Companion Methods Enzymol.* **3,** 98–108.
2. Strahler, J. R. & Hanash, S. (1991) *Methods: Companion Methods Enzymol.* **3,** 109–114.
3. Tsunasawa, S. & Sakiyama, F. (1984) *Methods Enzymol.* **106,** 165–170.
4. Desiderio, D., ed. (1991) in *Mass Spectrometry of Peptides* (CRC, Boca Raton, FL).
5. Karas, M., Bachmann, D., Bahr, U. & Hillenkamp, F. (1987) *Int. J. Mass Spectrom. Ion Processes* **78,** 53–68.
6. Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y. & Yoshida, T. (1988) *Rapid Commun. Mass Spectrom.* **2,** 151–153.
7. Beavis, R. C. & Chait, B. T. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 6873–6877.
8. Hillenkamp, F., Karas, M., Beavis, R. C. & Chait, B. T. (1991) *Anal. Chem.* **63,** 1193A–1203A.
9. Iwamatsu, A. (1992) *Electrophoresis* **13,** 142–147.
10. Wong, S., Grimley, C., Padua, A., Bourell, J. H. & Henzel, W. J. (1993) in *Techniques in Protein Chemistry IV,* ed. Angeletti, R. H. (Academic, San Diego).
11. Bauw, G. W., Van Damme, S., Puype, M., Vandekerchove, B. G., Ratz, G. P., Lauridsen, J. B. & Celis, J. E. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7701–7705.
12. Fernandez, J., DeMott, M., Atherton, D. & Mische, S. (1992) *Anal. Biochem.* **201,** 255–264.
13. Yuen, S. W., Chui, A. H., Wilson, K. J. & Yuan, P. M. (1989) *BioTechniques* **7,** 74–83.
14. Matsudaira, P. (1987) *J. Biol. Chem.* **262,** 10035–10038.
15. Henzel, W. J., Bourell, J. H. & Stults, J. T. (1990) *Anal. Biochem.* **187,** 228–233.
16. Henzel, W. J., Rodriguez, H. & Watanabe, C. (1987) *J. Chromatogr.* **404,** 41–52.
17. Köster, C., Castoro, J. A. & Wilkins, C. L. (1992) *J. Am. Chem. Soc.* **114,** 7572–7574.
18. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989) *Science* **246,** 64–71.
19. Billeci, T. M. & Stults, J. T. (1993) *Anal. Chem.* **65,** in press.