

A method for high-sensitivity peptide sequencing using postsource decay matrix-assisted laser desorption ionization mass spectrometry

T. KEOUGH*, R. S. YOUNGQUIST, AND M. P. LACEY

The Procter and Gamble Company, Miami Valley Laboratories, P.O. Box 538707, Cincinnati, OH 45253-8707

Edited by Jesse Lee Beauchamp, California Institute of Technology, Pasadena, CA, and approved May 7, 1999 (received for review January 29, 1999)

ABSTRACT A method has been developed for *de novo* peptide sequencing using matrix-assisted laser desorption ionization mass spectrometry. This method will facilitate biological studies that require rapid determination of peptide or protein sequences, e.g., determination of posttranslational modifications, identification of active compounds isolated from combinatorial peptide libraries, and the selective identification of proteins as part of proteome studies. The method involves fast, one-step addition of a sulfonic acid group to the N terminus of tryptic peptides followed by acquisition of postsource decay (PSD) fragment ion spectra. The derivatives are designed to promote efficient charge site-initiated fragmentation of the backbone amide bonds and to selectively enhance the detection of a single fragment ion series that contains the C terminus of the molecule (γ -ions). The overall method has been applied to pmol quantities of peptides. The resulting PSD fragment ion spectra often exhibit uninterrupted sequences of 20 or more amino acid residues. However, fragmentation efficiency decreases considerably at amide bonds on the C-terminal side of Pro. The spectra are simple enough that *de novo* sequence tagging is routine. The technique has been successfully applied to peptide mixtures, to high-mass peptides (up to 3,600 Da) and to the unambiguous identification of proteins isolated from two-dimensional gel electrophoresis. The PSD spectra of these derivatized peptides often allow far more selective protein sequence database searches than those obtained from the spectra of native peptides.

Knowledge of protein sequences is fundamentally important for understanding many physiological processes at the molecular level (1). Tandem mass spectrometry has become an increasingly essential tool for protein and peptide sequencing because of its speed, sensitivity, and applicability to complex mixtures (2). Recently, postsource decay (PSD) matrix-assisted laser desorption ionization (MALDI) mass spectrometry was developed for high-sensitivity peptide sequencing applications (3–8). Subpicomole limits of analysis were reported as a result of the high yield of fragment ions and the high ion transmission inherent with time-of-flight mass spectrometry (5). Kaufmann *et al.* (5) also noted several problems associated with PSD MALDI sequencing of peptides, including the complexity of the resulting fragmentation patterns and the lack of computer algorithms capable of interpreting the complex spectra. The recent incorporation of delayed extraction (DE) (9, 10), a technique designed to improve precursor-ion mass resolution and mass-measurement accuracy, reduced the rate of PSD fragmentation by at least an order of magnitude. As a result, many low-intensity precursor ions obtained by DE MALDI do

not produce enough PSD fragmentation to allow derivation of even short sequence tags (7).

Positively charged derivatives have been used in desorption mass spectrometry for many years to improve sensitivity by enhancing ionization efficiencies (11, 12). Localization of a positive charge at a particular site on a peptide (13, 14) also profoundly affects the nature of the fragment ions observed in high-collision energy tandem mass spectrometry (15–26). A protonated Arg or Lys, or a cationic derivative located near the N terminus of a peptide, promotes the formation of fragment ions containing the N terminus (a, b, c ions; the nomenclature for peptide fragment ions has been discussed in ref. 27). A protonated Arg or Lys or a fixed-charge group near the C terminus enhances fragment ions containing the C terminus of the peptide (x, y, z ions). It is not surprising, therefore, that several researchers have used N-terminal fixed-charge derivatives (quaternary ammonium and phosphonium groups) to improve MALDI sensitivity and to try to direct and simplify peptide fragmentation patterns observed in PSD mass spectra (28–35). In our hands, generally poor-quality MALDI PSD mass spectra have been obtained with fixed positive-charge derivatives. An alternative strategy to increase the yield of sequence-specific fragment ions by PSD was developed by Dikler and coworkers (36), who modified the guanidine side chain of Arg-containing peptides. That modification was done to minimize charge localization on the basic Arg residue and to facilitate protonation of the backbone amide groups.

Several previous studies (7, 30, 33, 37) raise questions about the internal energy content of precursor ions derived from DE MALDI and whether there is enough energy to optimally produce sequence-specific products by the charge-remote fragmentation mechanisms that are required with fixed positive-charge derivatives. Wysocki and coworkers (37) showed that addition of a fixed positive charge at the N terminus of a peptide increased the internal energy needed to fragment the peptide. Those results are consistent with the electrospray ionization findings of Gaskell and coworkers (38), who showed that the formation of precharged derivatives was detrimental to the yield of sequence-specific fragment ions after low-energy collisional activation of singly charged species. On the other hand, experimental (37) and theoretical (39) studies reveal that protonation of backbone amide nitrogen atoms enhance peptide fragmentation because amide bond orders are lowered dramatically by protonation.

Questions about the internal energy content of precursor ions derived from DE MALDI (7, 30, 33) and the compelling arguments for the mobile proton model of peptide fragmentation (37, 39) led us to investigate an alternative approach to

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.

PNAS is available online at www.pnas.org.

This paper was submitted directly (Track II) to the *Proceedings* office. Abbreviations: PSD, postsource decay; MALDI, matrix-assisted laser desorption ionization; DE, delayed extraction; 2D, two-dimensional; THF, tetrahydrofuran.

*To whom reprint requests should be addressed. e-mail: keough.tw@pg.com.

promote and direct fragmentation in PSD mass spectrometry. This method involves addition of a strong acid group to the N terminus of tryptic peptides before MALDI analyses. These derivatives facilitate low-energy charge site-initiated fragmentation, and they suppress the detection of N-terminal fragment ions in the positive ion mode of analysis. Dramatic increases in fragmentation efficiencies and fragment ion signal-to-noise ratios have been observed for derivatized peptides relative to nonderivatized peptides having the same sequences. This approach provides simple, easy-to-interpret fragment ion mass spectra that can be routinely interpreted *de novo*.

In this paper we demonstrate fast, one-step derivatization chemistry that can be applied generally to tryptic peptides. The role of the mobile proton in causing cleavage of backbone amide bonds was demonstrated by comparing the fragmentation behavior of MH^+ , MNa^+ , and $(M-H)^-$ molecular ions, and by comparing differences in the fragmentation patterns of tryptic peptides containing either one or two Arg residues. The acidity of the derivatizing group was shown to have a profound effect on the observed PSD fragment ion mass spectra. Peptides derivatized with sulfonic acid groups at the N terminus show γ -type fragment ions almost exclusively, whereas peptides derivatized with carboxylic acids also show complementary N-terminal fragment ions (a - and b -type). The overall sensitivity of the method has been demonstrated by applying it to pmol quantities of peptides isolated from HPLC and by using it for the unambiguous identification of proteins separated by two-dimensional (2D) gel electrophoresis. The utility of the method for identification of phosphorylation sites has been established by analysis of a model phosphorylated peptide. Reliable sequence tags have been obtained from large peptides (up to 3,600 Da) having masses exceeding the range of conventional triple quadrupole instruments. The PSD spectra often exhibit uninterrupted sequences of 20 or more aa, particularly when the peptides lack Pro residues. This capability will increase the utility of PSD MALDI data for the design of specific oligonucleotide probes for gene cloning experiments and for protein sequence database searches used in proteome studies.

MATERIALS AND METHODS

Chemicals. The model peptides used in this study include: CDPGYIGSR (C-0668), ASHLGLAR (A-8651), YGGFLR (E-8757), YGGFLRR (D-4524), PDVDHVFRLRF-NH₂ (P-2178), CLNRQLpSSGVSEIR (H-2654), and oxidized insulin B-chain (I-6383). They were obtained from Sigma, and they were used without further purification. Dithiodiglycolic acid (D-5392) was obtained from Sigma. DTSSP (3,3'-dithiobis[sulfosuccinimidylpropionate], 21578) was obtained from Pierce. Traut's reagent (33,056-6), S-acetylmecaptosuccinic anhydride (19,732-7), 2-sulfobenzoic acid cyclic anhydride (19,169-8), chlorosulfonylacetyl chloride (25,977-2), phthalic anhydride (32,006-4), 3-nitrophthalic anhydride (15,688-4), 4-nitrophthalic anhydride (22,820-1), and succinic anhydride (23,969-0) were obtained from Aldrich.

N-Terminal Derivatization Procedures. The two derivatization procedures most often used in this work involved coupling either 2-sulfobenzoic acid cyclic anhydride or chlorosulfonylacetyl chloride to the N terminus of peptides. 2-Sulfobenzoic acid cyclic anhydride was prepared at a concentration of 0.1 M in dry tetrahydrofuran (THF) just before use. High-level reactions involving 1 nM of peptide were carried out as follows: 1 nM of peptide was diluted into 20 μ l of 0.05 M trimethylamine. Two microliters of the fresh 2-sulfobenzoic acid cyclic anhydride solution was added, and the reaction mixture was vortexed for 30 sec. The reaction proceeded for 1-2 min at room temperature before dilution and MALDI analyses. The concentration of the coupling reagent was decreased by a factor of as much as 100 when derivatizing smaller quantities

of peptides. Reactions with chlorosulfonylacetyl chloride were carried out under nonaqueous conditions. Peptide (1 nM) was mixed with 2 μ l of 0.02 M sulfoacetic acid, which had been formed by mixing 2 μ l of neat chlorosulfonylacetyl chloride with 500 μ l of H₂O. The mixture was dried and then reconstituted in 20 μ l of THF/diisopropylethyl amine 4:1. Two microliters of 0.1 M chlorosulfonylacetyl chloride in dry THF was added, and the mixture was vortexed for 30 sec. The coupling reaction proceeded for 1-2 min at room temperature. The sample was then dried, reconstituted in 20 μ l H₂O, and further diluted before MALDI analyses. Derivatization of 2D gel isolates was accomplished in the same way except the concentrations of sulfoacetic acid and chlorosulfonylacetyl chloride were reduced 10- to 100-fold. These isolates were obtained from single 2D gels (not pooled samples), and they contained <10 pM per component.

Reactions involving phthalic anhydride, 3- and 4-nitrophthalic anhydride, and succinic anhydride were carried out in 0.05 M trimethylamine as described above for 2-sulfobenzoic acid cyclic anhydride.

Formic acid oxidation of CDPGYIGSR and some of the derivatized peptide intermediates was done by mixing 1-5 nM of peptide (in 5-20 μ l H₂O) with 10 μ l of fresh formic acid (88%)/H₂O₂ (30%) prepared at a ratio of 19:1 (vol/vol). Oxidation was allowed to proceed for 30 min at room temperature, and the samples were dried before analysis.

Mass Spectrometry. All mass spectrometry experiments were carried out on a PerSeptive Biosystems (Framingham, MA) Voyager DE-RP or Voyager DE-STR equipped with a N₂ laser (337 nm, 3-nsec pulse width, 20-Hz repetition rate). The mass spectra were acquired in the reflectron mode with DE. External mass calibration was performed with low-mass peptide standards, and mass-measurement accuracy was typically ± 0.3 Da. High levels of peptides and derivatized peptides were diluted to about 10 pM/ μ l in 0.1% trifluoroacetic acid (TFA). The samples then were diluted 5- to 10-fold further in α -cyano-4-hydroxycinnamic acid, which had been prepared by dissolving 10 mg in 1 ml of aqueous 50% acetonitrile containing 0.1% TFA (40). The 2D gel isolates were analyzed from thin film surfaces of α -cyano-4-hydroxycinnamic acid/nitrocellulose prepared by the fast evaporation method (41).

PSD fragment ion spectra were acquired for several peptides and peptide derivatives after isolation of the appropriate precursor ion by using timed ion selection. Fragment ions were refocused onto the final detector by stepping the voltage applied to the reflectron in the following ratios: 1.0000 (precursor ion segment), 0.9126, 0.6049, 0.4125, 0.2738, 0.1975, and 0.1213 (fragment segments). The individual segments were stitched together by using software provided by PerSeptive Biosystems. All precursor ion segments were acquired at low laser power (variable attenuator = 1,450) for <256 laser pulses to avoid saturating the detector. The laser power was increased (variable attenuator = 1,650) for all of the remaining segments of the PSD acquisitions. Typically, 256 laser pulses were acquired for each fragment-ion segment. The PSD data were acquired at a digitization rate of 20 MHz. Mass calibration was done with peptide standards only about once a month. Meta-stable decompositions were measured in all PSD mass spectrometry experiments.

Identification of Proteins Isolated From 2D Gel Electrophoresis. 2D-gel-separated protein spots were in-gel digested with trypsin following the method of Shevchenko and coworkers (42). Peptides were analyzed directly by MALDI after loading 5-10% aliquots of the digestion solutions. The measured tryptic peptide masses were used as inputs to search the NCBIInr5.18.98 database containing 299,924 protein sequences. All searches were carried out by using the PROTEIN PROSPECTOR software developed at the University of California, San Francisco (see P. R. Baker and K. R. Clauser, <http://prospector.ucsf.edu>). No restrictions were placed on the

species of origin of the protein. The allowed protein molecular mass range was 1,000 to 150,000 Da. Isoelectric points were allowed to range from 3.0 to 10.0, and oxidation of Met was included as a side reaction. Up to one missed tryptic cleavage was considered, and a conservative mass accuracy of ± 0.6 Da was used for all tryptic-mass searches. The number of tryptic peptides required for identification varied as a function of the quality of the MALDI mass spectra. The PROTEIN PROSPECTOR software also was used to search PSD fragment ion spectra either against the entire protein sequence database or against the saved set of proteins identified from the appropriate tryptic mass searches. Only y-type or (y-NH₃) ions were allowed fragment-ion possibilities when searching the PSD spectra of derivatized peptides.

RESULTS AND DISCUSSION

Method to Promote Charge-Site-Initiated Fragmentation in PSD MALDI of Tryptic Peptides. Gaskell and coworkers (43, 44) previously demonstrated that the oxidation of cysteine to cysteic acid in peptides containing a C-terminal Arg strongly increases the yield of y-type fragment ions observed by low-energy collisional activation of protonated molecules formed by electrospray ionization. These studies formed the starting point of our efforts to develop a general procedure for high-sensitivity tryptic peptide sequencing using PSD MALDI. We have developed several convenient routes to produce tryptic peptides containing an N-terminal sulfonic acid (a),



where A_i are amino acid residues. This strategy recognizes that the basic C terminus (X = Lys or Arg) will be protonated after MALDI. The strong acid group at the N terminus will be deprotonated, exactly counterbalancing the C-terminal positive charge. One additional proton then is required to ionize the derivative for mass spectrometry analyses. That proton will be more or less free to randomly ionize the backbone amide groups because the most basic residue already is occupied. The resulting PSD spectra should show increased fragmentation by direct cleavage of protonated amide bonds. The major products of these fragmentation reactions have structures like **b** and **c**,



which would be produced by fragmentation of a derivatized tryptic dipeptide (R₂ is the basic side chain of a Lys or Arg residue). Fragments containing the N terminus will be suppressed in the positive ion mode because the negative charge on the sulfonate group neutralizes the C-terminal positive charge in **b**. Internal cleavage ions will not be enhanced because they lack the positively charged C-terminal Lys or Arg, and a single series of y-type ions (**c**) should be enhanced because they contain the protonated C terminus.

A commercially available peptide, CDPGYIGSR, was analyzed to determine whether derivatives like **a** exhibit the predicted effect on PSD MALDI fragmentation behavior. Fig. 1, *Upper* shows the PSD fragment ion spectrum of the native peptide, while Fig. 1, *Lower* shows the corresponding spectrum after performic acid oxidation of the N-terminal cysteine to cysteic acid. The spectrum of the native peptide shows an incomplete series of y-ions (y₁, y₂, y₃, y₇, and y₈) dominated by the low-energy cleavage of the Asp-Pro amide bond. Prominent b₄ and weak b₅ ions are also evident. The spectrum was much improved after oxidation, Fig. 1, *Lower*. The N-terminal fragment ions were suppressed and several y-ions (y₄, y₅, y₆, and y₈) were enhanced considerably. The complete sequence of the peptide could be determined by PSD

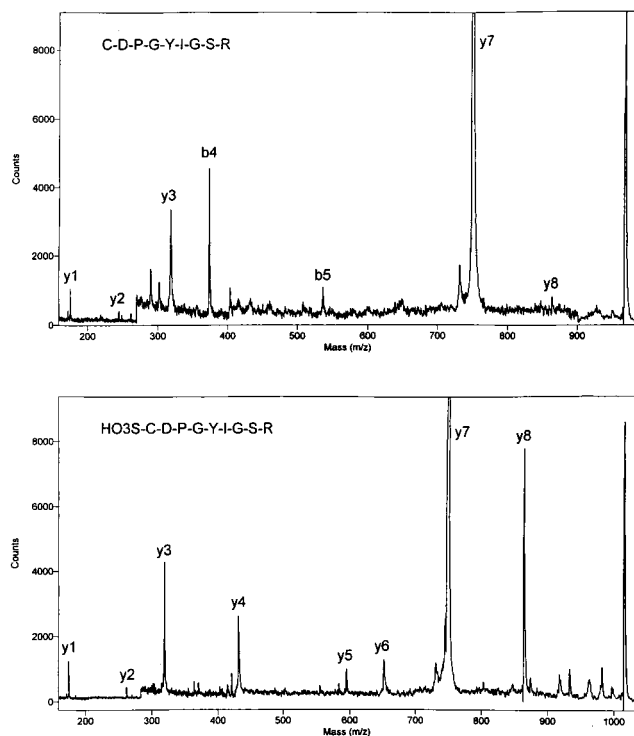


Fig. 1. PSD fragment ion mass spectra of CDPGYIGSR before (*Upper*) and after (*Lower*) oxidation with performic acid.

MALDI after oxidation. These promising results led us to investigate general methods to derivatize tryptic peptides.

N-Terminal Derivatization Methods. A number of approaches for the introduction of sulfonic acid groups to the N terminus of peptides were evaluated in this study. The first method involved introduction of an N-terminal sulfur that subsequently could be oxidized. Dithiodiglycolic anhydride (**I**) and DTSSP (3,3'-dithiobis[sulfosuccinimidylpropionate]) (**II**), a commonly used protein crosslinking agent (45), enable the introduction of a disulfide bond at the N terminus of peptides. 2-Iminothiolane **III** (Traut's reagent) (46) provides an N-terminal sulfhydryl group, whereas S-acetyl mercaptosuccinic anhydride (**IV**) provides an acetylated sulfhydryl group. However, all of these derivatives require performic acid oxidation to yield the desired sulfonic acid. Oxidation of labile amino acid residues made the resulting mass spectra more difficult to interpret. The second approach, which involved direct introduction of the sulfonic acid, was a more effective method. Sulfobenzoic acid cyclic anhydride (**V**) couples readily to free amines in aqueous solution. This reaction introduces an aromatic sulfonic acid group directly at the N terminus of the peptide. On the other hand, chlorosulfonylacetyl chloride (**VI**) did not provide high yields of derivatized peptides when the coupling reaction was carried out in aqueous solution. Presumably, it is too reactive with water. Coupling proceeded well in dry THF and subsequent addition of water converted the chlorosulfonyl group into the desired sulfonic acid.

Mass Spectrometry of Derivatized Peptides. The MALDI mass spectra of the derivatized tryptic peptides often exhibit MH⁺ ions for the starting peptide as well as for the desired final product. The ratio of these two ions does not provide a good estimate of the overall product yield. The positive-ion response of the derivatized peptide is expected to be considerably lower than that of the unreacted starting material because of the added sulfonic acid group. In-source fragmentation of the derivative to regenerate starting material also alters the ratio of these ions. To investigate this effect, the ratio of starting peptide to derivatized product was examined as a

function of laser power for the mixture formed by adding **VI** to ASHLGLAR in dry THF. The ratio of starting peptide to derivatized product increased from 0.25 at the lowest laser power used (variable attenuator set to 1,150) to 1.09 at the highest power (attenuator set to 1,650). Negative-ion MALDI is a more convenient method to monitor these reactions because the molecular anions are more stable and the MALDI spectra are simpler.

The PSD fragment ion spectra of the derivatives formed by reacting **V** and **VI** with various tryptic peptides are similar. This is true for all of the tryptic peptides we examined, whether they contain a C-terminal Arg or Lys. The spectra of both derivatives exhibit prominent *y*-ions, with few other ions observed. The elements of $O_3S-C_6H_4-CO$ are readily lost from products formed with **V**, and the most abundant product ion in those PSD spectra generally corresponds to the mass of the unreacted starting tryptic peptide. The relative abundance of the *y*-ions decreases toward lower mass. Fragmentation at amide bonds on the C-terminal side of Pro is less favorable than it is for other amino acid residues (for all derivatives examined). Derivatives formed with **VI** show losses of SO_3 and O_3S-CH_2-CO to a small extent, in addition to the abundant *y*-series ions. The product ion formed by loss of the elements of O_3S-CH_2-CO generally is not the most abundant ion in the PSD spectra of these derivatives. Lower-mass *y*-series ions dominate, which suggests that peptide derivatives formed with **VI** are more stable than those formed with **V** under MALDI conditions. A larger fraction of the PSD product ion current will be directed into structurally significant *y*-ions with chlorosulfonyl chloride derivatives than with derivatives formed with 2-sulfobenzoic acid cyclic anhydride.

Dramatic improvements often are observed in PSD MALDI fragment ion spectra after derivatization. Results obtained from a 24-aa peptide, formed by tryptic digestion of an enzyme isolated from a commercial laundry product, are given in Fig. 2. The peptide was isolated by reverse-phase HPLC. The PSD spectrum obtained from the native peptide (Fig. 2, *Upper*) demonstrates little structurally significant fragmentation. The

spectrum of the derivative formed with **VI** was obtained under very similar experimental conditions. It shows a complete series of *y*-ions, enabling rapid verification of the amino acid sequence (NTATSLGSTNLYGSGLVNAEAATR) of this portion of the enzyme.

The Role of the Mobile Ionizing Proton. The role of the mobile ionizing proton was assessed by comparing the PSD fragment ion spectra of the MH^+ , MNa^+ , and $(M-H)^-$ ions of the derivative formed between ASHLGLAR and **V**. The spectrum of the protonated derivative showed a complete series of *y*-ions. It also showed an MH^+ molecular ion that was of about half the abundance to the most prominent product ion, which corresponded to loss of 184 Da ($O_3S-C_6H_4-CO$) as discussed above. Changing the ionizing agent from a proton to a sodium cation greatly reduced PSD fragmentation. The *y*-ion series was virtually absent from the spectrum of the MNa^+ precursor ion. The most intense ion in the spectrum was the parent MNa^+ ion. The main product ion corresponded to loss of 201 Da (the elements of $HO_3S-C_6H_4-CO-NH_2$). Weak signals also were observed for loss of 184 Da ($O_3S-C_6H_4-CO$) and loss of (184 + 156 Da), where the 156-Da fragment comes from the Arg residue at the C terminus. Similarly poor results were observed for the $(M-H)^-$ ion, which also lacks the mobile ionizing proton. No *y*-ions were observed and the undissociated molecular ion was the largest ion observed in the PSD spectrum. The only other prominent fragment ions correspond to loss of the C-terminal R (-156 Da) and loss of 184 Da. It would not be possible to sequence this peptide by using the fragment ion mass spectra obtained from either the MNa^+ or $(M-H)^-$ precursor ions.

The importance of the mobile proton in promoting fragmentation in these systems was explored further by obtaining the PSD spectra of the derivatives formed between **V** and YGGFLR and YGGFLRR. Wysocki's group (37) previously examined the doubly charged ions formed from some bradykinin-related peptides that contained two Arg residues. Their experiments demonstrated that both protons were preferentially localized at the basic Arg residues and neither proton was mobile across the peptide backbone. We expected similar charge localization with these peptides and a corresponding reduction in PSD fragmentation. Derivatized YGGFLR showed typical behavior, i.e., a complete and abundant *y*-ion series dominated by y_5 and $MH^+-(O_3S-C_6H_4-CO)$. The non-dissociated MH^+ ion had a relative abundance of about 30% compared with that of y_5 . Derivatized YGGFLRR showed quite different PSD behavior when run under the same conditions as for YGGFLR. The MH^+ precursor ion was the most abundant ion in this PSD spectrum and the fragment ions were reduced in relative abundance by at least a factor of 2 or 3. The *y*-ion series was evident in the spectrum, but all *y*-ions except y_1 were accompanied with a (y_i-17) ion that was at least twice as large as the corresponding y_i ion. Tang and Boyd (47) observed similar ($y-17$) ions in the electrospray tandem mass spectra of singly charged peptides containing a C-terminal Arg.

It should be noted that YGGFLRR is an interesting model for peptides formed by incomplete tryptic digestion of proteins. The differences in fragmentation behavior between complete and incomplete tryptic peptides must be considered when searching protein sequence databases against the PSD spectra of these derivatives (see below).

Acidity of the Derivatizing Group. A number of model peptides (ASHLGLAR, YGGFLRR, and PDVDHVFLRF-NH₂) were derivatized with various anhydrides like phthalic anhydride, 3- and 4-nitrophthalic anhydride and succinic anhydride to test the effect of acid strength on the PSD MALDI spectra of derivatized tryptic-like peptides. The spectra of the peptides derivatized with the anhydrides showed extensive fragmentation, but they were much more complex than those obtained with any of the sulfonic acid derivatives. They showed

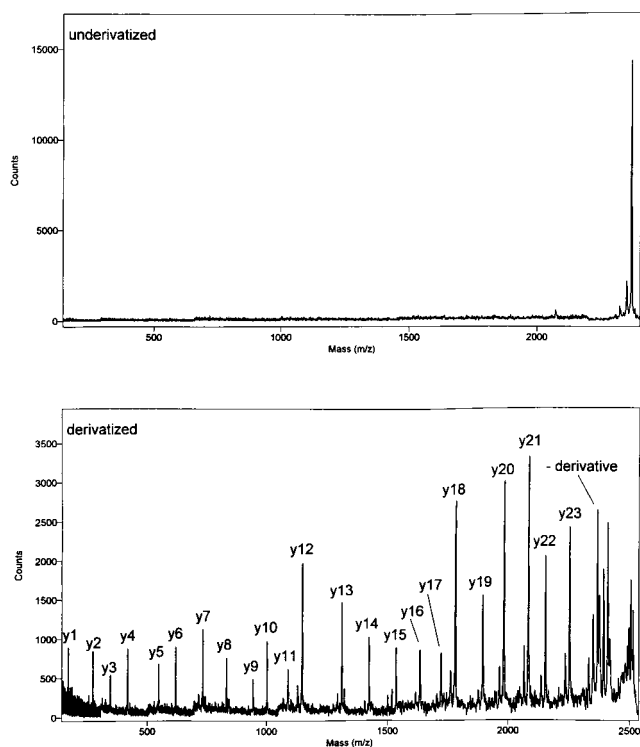


Fig. 2. PSD MALDI mass spectra of a tryptic peptide before (*Upper*) and after (*Lower*) derivatization.

a- and b-ions, both with and without the added anhydride, in addition to y-ions. None of the anhydride derivatives we investigated gave simple, single-series spectra like the ones obtained after derivatization with **V** or **VI**. They were unable to suppress detection of the N-terminal fragment ions, probably because the resulting acids are more basic (pKas ranging from 1.8 to 4.2, ref. 48) than protonated amide groups [pKa of protonated acetamide = -1 (nonaqueous)]. Strong acids, like sulfonic acids [pKas about -7 (nonaqueous)], appear to be required to suppress detection of the N-terminal fragment ions and produce easy-to-interpret spectra.

Phosphopeptide Analysis. The utility of this method for determination of phosphorylation sites was demonstrated by analysis of a model phosphopeptide CLNRQLpSSGVSEIR obtained commercially. The PSD MALDI mass spectrum obtained from the native peptide showed poor sequence coverage. Only weak y_1 and y_7 ions were observed above background noise. PSD MALDI of the peptide derivatized with **V** showed much better fragment ion sensitivity and sequence coverage. A complete series of unmodified y-ions was observed from y_1 to y_7 confirming that Ser residues 8 and 11 (numbered from the N terminus) were not phosphorylated. The phosphate group was located on Ser-7 by the observation of an abundant (y_8 -H₃PO₄) ion. Facile loss of neutral phosphate is a characteristic of serine- and threonine-phosphorylated peptides (49). The y_9 , y_{10} , y_{11} , and y_{12} fragment ions also showed quantitative loss of H₃PO₄. The y_{11} and y_{12} ions showed further loss of NH₃, as discussed above for tryptic peptides containing a second basic residue. No N-terminal fragment ions were observed in this spectrum even though the peptide contains an Arg at residue 4.

Variant Proteins. The MALDI mass spectrum of a tryptic digest obtained from about 1 nmol of an enzyme isolated from a commercial product showed many masses identical to those of savinase, a widely used commercial protease. One of the expected tryptic peptides GVLVVAASGNSGAGSISYPAR (MH⁺ = 1,933 Da) was missing from the spectrum, and an unknown ion was observed at m/z 1963. This finding suggests that the isolate is a savinase variant. Eleven single amino acid changes could account for the observed +30-Da mass shift (four G → S, four A → T, and three V → E). Two or more amino acid changes are also possible, and a complete sequence is required to identify the variant. The PSD MALDI spectrum obtained after derivatization of the isolate (with **VI**) is given in Fig. 3. A complete series of y-ions is observed for this 21-aa peptide. The spectrum verifies that only a single amino acid has been changed. The Gly at residue 14 was converted to a Ser. Incidentally, the spectrum was automatically interpreted by using a sequencing program previously written by PerSeptive Biosystems. That program had been developed for automated

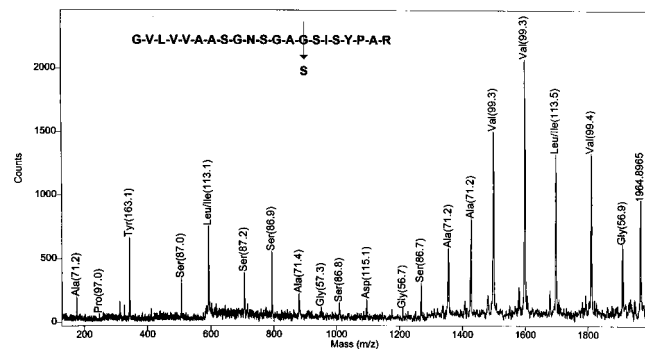


Fig. 3. PSD MALDI mass spectrum of a derivatized tryptic peptide from a variant protein. The y_1 ion is observed at m/z 175 establishing the C-terminal residue as R (not labeled by the software). Mass differences between adjacent y-ions establish the identities of the remaining amino acid residues (indicated on the figure).

interpretation of MALDI mass spectra of peptide ladders generated chemically (50) or by carboxypeptidase Y digestion (51). The computer-assigned amino acid residues are indicated in Fig. 3. One error was made in the automated interpretation of this spectrum, i.e., N (114 Da) was assigned as D (115 Da). Careful mass calibration should minimize this problem. The result illustrates that automated interpretation of these PSD fragment ion mass spectra is feasible.

High-Mass Peptides. PSD mass spectrometry of high-mass peptides was demonstrated with oxidized insulin B-chain (MM_{calc} = 3,495.9 Da). The mass of this peptide far exceeds the upper limit of most conventional triple quadrupole instruments. The fragment ion mass spectrum of the native peptide showed a relatively abundant product-ion series consisting mainly of y-ions. All of the y-ions between y_9 and y_{24} were clearly observed. This abundant fragment ion series probably results because the native peptide contains a basic Arg and Lys residue near the C terminus and it has sulfonic acid groups attached to the two Cys residues. The sequence-specific fragments defining the N terminus of the molecule, y_{25} to y_{29} , were entirely absent from this spectrum. The PSD spectrum of this compound was improved by derivatization with **V**. The derivatized peptide showed significant enhancement of the y_{25} , y_{26} , y_{27} , y_{28} , and y_{29} ions. A complete series of y-ions was observed from y_8 to y_{29} after derivatization and no prominent b-ions were detected. The spectrum is less complex than that obtained on a MALDI ion trap mass spectrometer (52), which showed significant b-type ions interspersed with the y-ions.

Identification of Proteins Isolated from 2D Gel Electrophoresis. We use MALDI and PSD mass spectrometry extensively for the identification of proteins separated from 2D gel electrophoresis, and these derivatization methods have proven useful even with the low-pmol levels of tryptic peptides generally available from these experiments. In a recent study, an excellent MALDI mass map was obtained from a protein spot after in-gel tryptic digestion (42, 53). It showed several abundant MH⁺ signals including ions at m/z 1060.8, 1090.8, 1271.0, 1299.0, 1312.0, 1344.0, 1399.1, 1450.1, 1460.1, and 1794.4. Database searching against the 299,924 entries in the NCBI nr5.18.98 protein sequence library produced a list of 41 candidate proteins that matched five or more of the input tryptic masses. The top molecular weight search (MOWSE) scores were quite low, only in the 560 range. The entire peptide mixture was then reacted with **VI**, and PSD mass spectrometry was carried out on the derivatized peptide weighing about 1,572 Da (1450 + 122 Da), Fig. 4. A sequence tag was obtained and it consisted of y-ions at m/z 574.5, 661.7, 875.9, 1003.8, 1103.9, 1204.6, 1304.1, and the (MH⁺ derivative) ion at m/z 1450.3. That spectrum was searched against the entire protein sequence database, and six candidate proteins (all mitochondrial aspartate aminotransferases) were returned. This level of specificity was obtained because the database search was

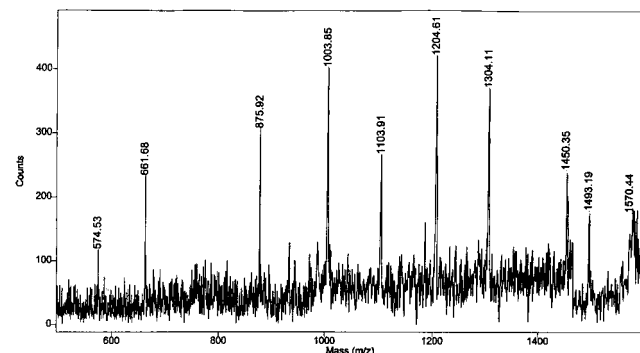


Fig. 4. PSD MALDI fragment ion mass spectrum of a derivatized peptide obtained from an in-gel tryptic digestion of a protein separated by 2D gel electrophoresis.

constrained to consider γ -ions as the only allowed fragments. This search constraint is possible because N-terminal and internal cleavage ions are not prominent in the PSD mass spectra of these derivatives. Much less specificity was obtained (239 candidate proteins returned) when this same fragment ion mass spectrum was searched against the full library and all fragment ion types [a, b, (b + H₂O), (b-NH₃), (b-H₂O), internal, and γ] were allowed.

Confirmation of the protein identification was sought by PSD MALDI of the derivative weighing 1,916 Da (1,794 + 122 Da). The PSD spectrum contained seven fragment ions at m/z 724.6, 1154.2, 1299.3, 1439.0, 1551.9, 1665.5, and 1779.2 in addition to the protonated precursor ion. This spectrum was searched against the full database assuming that the fragments were γ -ions. The search did not return mitochondrial aspartate aminotransferase or any other candidate protein. Researching the database allowing both γ and (γ -NH₃) ions returned 16 candidate proteins, 10 of which were mitochondrial aspartate aminotransferases. This latter search confirmed the protein identification. Inclusion of (γ -NH₃) ions as possible fragments was required for this particular search because the peptide turned out to be an incomplete tryptic product, ILIRPLYS-NPPLNGAR, that contained a second Arg residue. In this case, the PSD fragment ions were exclusively (γ -NH₃) ions and not γ -ions.

We thank Dr. Raymond A. Grant, Karen B. Begley, and Angela M. Fieno for the separations and sample preparations associated with the 2D gel electrophoresis experiments reported in this paper.

- Allen, G. (1981) *Sequencing of Proteins and Peptides* (Elsevier/North-Holland, New York).
- Carr, S. A., Hemling, M. E., Bean, M. F. & Roberts, G. D. (1991) *Anal. Chem.* **63**, 2802–2824.
- Spengler, B., Kirsch, D., Kaufmann, R. & Jaeger, E. (1992) *Rapid Commun. Mass Spectrom.* **6**, 105–108.
- Spengler, B., Kirsch, D. & Kaufmann, R. (1992) *J. Phys. Chem.* **96**, 9678–9684.
- Kaufmann, R., Spengler, B. & Lutzenkirchen, F. (1993) *Rapid Commun. Mass Spectrom.* **7**, 902–910.
- Kaufmann, R., Kirsch, D. & Spengler, B. (1994) *Int. J. Mass Spectrom. Ion Processes* **131**, 355–385.
- Kaufmann, R., Chaurand, P., Kirsch, D. & Spengler, B. (1996) *Rapid Commun. Mass Spectrom.* **10**, 1199–1208.
- Spengler, B. (1997) *J. Mass Spectrom.* **32**, 1019–1036.
- Brown, R. S. & Lennon, J. J. (1995) *Anal. Chem.* **67**, 1998–2003.
- Vestal, M. & Juhasz, P. (1998) *J. Am. Soc. Mass Spectrom.* **9**, 892–911.
- Keough, T. & DeStefano, A. J. (1981) *Anal. Chem.* **53**, 25–29.
- Busch, K. L., Unger, S. E., Vincze, A., Cooks, R. G. & Keough, T. (1982) *J. Am. Chem. Soc.* **104**, 1507–1511.
- Johnson, R. S., Martin, S. A. & Biemann, K. (1988) *Int. J. Mass Spectrom. Ion Processes* **86**, 137–154.
- Vath, J. E. & Biemann, K. (1990) *Int. J. Mass Spectrom. Ion Processes* **100**, 287–299.
- Kidwell, D. A., Ross, M. M. & Colton, R. J. (1984) *J. Am. Chem. Soc.* **106**, 2219–2220.
- Renner, D. & Spiteller, G. (1986) *Biomed. Environ. Mass Spectrom.* **13**, 405–410.
- Wagner, D. S., Salari, A., Gage, D. A., Leykam, J., Fetter, J., Hollingsworth, R. & Watson, J. T. (1991) *Biol. Mass Spectrom.* **20**, 419–425.
- Zhao, Y.-F., Zhang, D.-Q. & Xue, C.-B. (1991) *Int. J. Peptide Protein Res.* **37**, 457–461.
- Watson, J. T., Wagner, D. S., Chang, Y.-S., Strahler, J. R., Hanash, S. M. & Gage, D. A. (1991) *Int. J. Mass Spectrom. Ion Processes* **111**, 191–209.
- Yang, H.-J., He, M.-Y., Ye, Y.-H. & Zhao, Y.-F. (1992) *Org. Mass Spectrom.* **27**, 746–749.
- Kiplinger, J. P., Contillo, L., Hendrick, W. L. & Grodski, A. (1992) *Rapid Commun. Mass Spectrom.* **6**, 747–752.
- Chang, Y.-S., Gage, D. A. & Watson, J. T. (1993) *Biol. Mass Spectrom.* **22**, 176–180.
- Bunk, D. M. & Macfarlane, R. D. (1993) *Int. J. Mass Spectrom. Ion Processes* **126**, 123–136.
- Stults, J. T., Shannon, J. L., McCune, S. & Wetzel, R. (1993) *Anal. Chem.* **65**, 1703–1708.
- Downard, K. M. & Biemann, K. (1994) *J. Am. Soc. Mass Spectrom.* **5**, 966–975.
- Zaia, J. & Biemann, K. (1995) *J. Am. Soc. Mass Spectrom.* **6**, 428–436.
- Biemann, K. (1988) *Biomed. Environ. Mass Spectrom.* **16**, 99–111.
- Bartlett-Jones, M., Jeffery, W. A., Hansen, H. F. & Pappin, D. J. C. (1994) *Rapid Commun. Mass Spectrom.* **8**, 737–742.
- Liao, P.-C. & Allison, J. (1995) *J. Mass Spectrom.* **30**, 511–512.
- Liao, P.-C., Huang, Z.-H. & Allison, J. (1997) *J. Am. Soc. Mass Spectrom.* **8**, 501–509.
- Schurch, S., Scott, J. R. & Wilkins, C. L. (1997) *Int. J. Mass Spectrom. Ion Processes* **169/170**, 141–152.
- Stimson, E., Truong, O., Richter, W. J., Waterfield, M. D. & Burlingame, A. L. (1997) *Int. J. Mass Spectrom. Ion Processes* **169/170**, 231–240.
- Spengler, B., Luetzenkirchen, F., Metzger, S., Chaurand, P., Kaufmann, R., Jeffery, W., Bartlett-Jones, M. & Pappin, D. J. C. (1997) *Int. J. Mass Spectrom. Ion Processes* **169/170**, 127–140.
- Huang, Z.-H., Wu, J., Roth, K. D. W., Yang, Y., Gage, D. A. & Watson, J. T. (1997) *Anal. Chem.* **69**, 137–144.
- Strahler, J. R., Smelyanskiy, Y., Lavine, G. & Allison, J. (1997) *Int. J. Mass Spectrom. Ion Processes* **169/170**, 111–126.
- Dikler, S., Kelly, J. W. & Russell, D. H. (1997) *J. Mass Spectrom.* **32**, 1337–1349.
- Dongre, A. R., Jones, J. L., Somogyi, A. & Wysocki, V. H. (1996) *J. Am. Chem. Soc.* **118**, 8365–8374.
- Burlet, O., Orkiszewski, R. S., Ballard, K. D. & Gaskell, S. J. (1992) *Rapid Commun. Mass Spectrom.* **6**, 658–662.
- McCormack, A. L., Somogyi, A., Dongre, A. R. & Wysocki, V. H. (1993) *Anal. Chem.* **65**, 2859–2872.
- Beavis, R. C., Chaudhary, T. & Chait, B. T. (1992) *Org. Mass Spectrom.* **27**, 156–158.
- Arnott, D., O'Connell, K. L., King, K. L. & Stults, J. T. (1998) *Anal. Biochem.* **258**, 1–18.
- Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. (1996) *Anal. Chem.* **68**, 850–858.
- Burlet, O., Yang, C.-Y. & Gaskell, S. J. (1992) *J. Am. Soc. Mass Spectrom.* **3**, 337–344.
- Cox, K. A., Gaskell, S. J., Morris, M. & Whiting, A. (1996) *J. Am. Soc. Mass Spectrom.* **7**, 522–531.
- Lomant, A. J. & Fairbanks, G. (1976) *J. Mol. Biol.* **104**, 243–261.
- Traut, R. J., Bollen, A., Sun, T. T., Hershey, J. W. B., Sunberg, J. & Pierce, L. R. (1973) *Biochemistry* **12**, 3266–3273.
- Tang, X.-J. & Boyd, R. K. (1992) *Rapid Commun. Mass Spectrom.* **6**, 651–657.
- Dean, J. A. (1992) *Lange's Handbook of Chemistry* (McGraw Hill, New York), 14th Ed.
- Annan, R. S. & Carr, S. A. (1996) *Anal. Chem.* **68**, 3413–3421.
- Chait, B. T., Wang, R., Beavis, R. C. & Kent, S. B. H. (1993) *Science* **262**, 89–92.
- Patterson, D. H., Tarr, G. E., Regnier, F. E. & Martin, S. A. (1995) *Anal. Chem.* **67**, 3971–3978.
- Quin, J., Steenvoorden, R. J. J. M. & Chait, B. T. (1996) *Anal. Chem.* **68**, 1784–1791.
- Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., Wilm, M., Vorm, O., Mortensen, P., Shevchenko, A., Boucherie, H. & Mann, M. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 14440–14445.