A Mass Spectrometric Technique for Detecting and Identifying By-Products in the Synthesis of Peptides[1]

Swapan K. Chowdhury and Brian T. Chait

The Rockefeller University, New York, New York 10021

The utility of a new mass spectrometric technique for detecting and identifying peptide by-products produced in the synthesis of peptides is demonstrated. The technique involves three sequential steps: (1) practically nondestructive ^{252}Cf plasma desorption mass spectrometric analysis of monolayer amounts of the peptide(s) of interest bound to a thin layer of nitrocellulose; (2) enzyme-catalyzed microscale chemical reaction of the surface-bound peptide(s) to produce structurally informative hydrolysis products; (3) plasma desorption mass spectrometric analysis of these hydrolysis products. The first step determines the presence and the molecular weights of unwanted by-products resulting from errors or incomplete reactions during synthesis. The subsequent two steps provide information on the precise location in the peptides where errors have occurred. In the present paper, the technique is applied to an investigation of unwanted peptide by-products associated with the use of tryptophan during stepwise solid-phase peptide synthesis. Synthetic preparations of melittin and [Bpa-8]dynorphin A (1–17) were each found to contain a major impurity with molecular weight 28 Da higher than that of the desired product. The impurity in the melittin preparation, in which the final deprotection step involved the high–low HF procedure, was shown to result from incomplete removal of the formyl group from Trp-19. On the other hand, the impurity in the [Bpa-8]dynorphin A (1–17) preparation, where the removal of the formyl group from Trp-14 was carried out using piperidine, was shown to result from migration of the formyl group to Lys-11 or Lys-13. Findings concerning the relative rates of hydrolysis at N^i-formylated tryptophan and Trp by chymotrypsin and N^ϵ-formylated lysine and Lys by carboxypeptidase B are also presented. © 1989 Academic Press, Inc.

There is a large and a rapidly growing need among members of the biological community for high purity synthetic peptides and small proteins. The most widely used methods for producing these materials are based on the stepwise solid-phase synthetic procedure devised by Merrifield (1) (for reviews see Refs. (2–4)). As the size and complexity of the target peptide increases, the opportunity for synthetic errors, unwanted modifications, and cumulative effects arising from incomplete reactions also increases. It is therefore imperative to have available effective means for rapidly verifying the correctness of the covalent structure of these complex materials, establishing their purity, and detecting and identifying undesired peptide by-products. The numerous methods that have been devised for these purposes have been reviewed (2–4). The most useful have involved subjecting the products to high resolution HPLC separation (4,5), amino acid analysis (6), sequence analysis (7), spectrophotometric analysis (8), NMR analysis (9), and mass spectrometry (10–14).

Perhaps the most widely utilized technique at present for assessing the homogeneity of the purified peptide product and for detecting the presence of by-products is reverse-phase HPLC. Since this same technique is commonly used on a preparative scale for purifying the desired peptide from the crude synthetic product, it is not surprising that undesired materials which copurify with the compound of interest are also often not resolved in the analytical HPLC analysis. It is thus desirable to assess homogeneity with an analytical technique which separates compounds by a different principle from that used in the purification. We have found (14) that ^{252}Cf plasma desorption mass spectrometry (PDMS)[2] (15) serves this function well. We base this assessment on
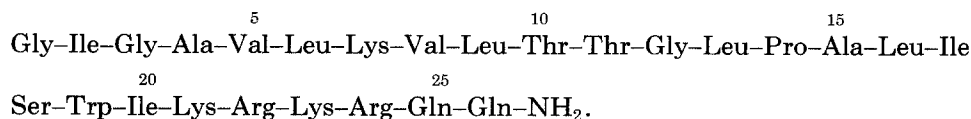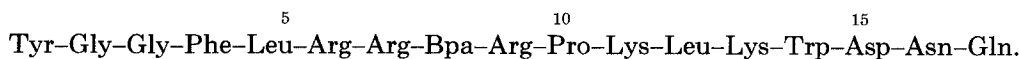
our detailed analysis of more than 800 synthetic peptides and proteins submitted to The Rockefeller University Mass Spectrometric Research Resource by 15 different laboratories in the United States (14). The resolution of this mass spectrometric technique is sufficiently high so that peptide by-products differing only very slightly from the target peptide can usually be clearly discerned and their molecular weights readily determined. While such simple molecular weight information can provide useful clues as to the origin and nature of by-products, it does not provide direct information on the site of the error(s). Recently, tandem mass spectrometry (16) has been demonstrated to be a powerful tool for pinpointing such sites of error or modification (13). We have developed a complementary mass spectrometric approach to tandem mass spectrometry for extracting information of this type (17–19). The technique involves three sequential steps: (1) practically nondestructive ^{252}Cf PDMS analysis of monolayer amounts of the peptide(s) of interest bound to a thin layer of nitrocellulose (NC); (2) enzyme-catalyzed microscale chemical reaction of the surface-bound peptide(s) to produce structurally informative hydrolysis products; (3) PDMS analysis of these hydrolysis products. The first step determines the presence and the molecular weights of unwanted by-products resulting from errors or incomplete reactions during synthesis. The subsequent two steps provide information on the location in the peptides where errors have occurred.

In the present paper we demonstrate the utility of our method for elucidating the structure of unwanted peptide by-products which we have frequently observed in tryptophan-containing peptides (14). The NH group of the indole nucleus in the tryptophan side chain can suffer alkylation if not protected (3). To protect against this side reaction, use of the formyl-masking group (21) has found general acceptance (22). Although several deprotection schemes have been used (3,4,22), the complete deprotection of N-formyl tryptophan has proved nontrivial ((14) and results presented herein). In order to assist in the elucidation of the problems associated with syntheses involving tryptophan we undertook a detailed mass spectrometric investigation of two tryptophan-containing synthetic peptides in which large amounts of the formyl group containing impurities were observed. Both materials were synthesized using the Merrifield stepwise solid-phase procedure. The first peptide, melittin, containing a Trp residue in position 19, is the principal component of the venom from the honey bee, Apis Mellifera (23). The sequence is

$$\overset{5}{}\qquad\qquad\overset{10}{}\qquad\qquad\overset{15}{}$$
Gly–Ile–Gly–Ala–Val–Leu–Lys–Val–Leu–Thr–Thr–Gly–Leu–Pro–Ala–Leu–Ile
$$\overset{20}{}\qquad\qquad\overset{25}{}$$
Ser–Trp–Ile–Lys–Arg–Lys–Arg–Gln–Gln–NH$_2$.

The second synthetic peptide is a modified dynorphin, [Bpa-8]Dyn A (1–17) containing a Trp residue in position 14 and a p-benzoylphenylalanine in position 8. The sequence is

$$\overset{5}{}\qquad\qquad\overset{10}{}\qquad\qquad\overset{15}{}$$
Tyr–Gly–Gly–Phe–Leu–Arg–Arg–Bpa–Arg–Pro–Lys–Leu–Lys–Trp–Asp–Asn–Gln.

MATERIALS AND METHODS

Materials. Trypsin (DPCC treated), carboxypeptidase B, proteinase K, α-chymotrypsin (TLCK treated), and NH$_4$HCO$_3$ were obtained from Sigma Chemical Co. (St. Louis, MO). Water and acetone were high purity, distilled in glass grade obtained from Burdick and Jackson Lab., Inc. (Muskegon, MI). α,α,α-Trifluoroacetic acid was HPLC spectrograde obtained from Pierce Chemical Co. (Rockford, IL). Nitrocellulose was obtained in the form of membranes (BA84 0.45-μm pore size) from Schleicher & Schuell, Inc. (Keene, NH). All chemicals were used without further purification.

Melittin. This peptide, submitted to us for mass spectrometric analysis, was synthesized on an Applied Biosystems Model 430A synthesizer by the Merrifield stepwise solid-phase procedure (1,8) on p-methylbenzhydrylamine resin (24). The protecting group strategy and synthetic protocol follows closely that described by Christensen *et al.* (25). The deprotection scheme was the low–high two-step HF cleavage procedure of Tam *et al.*

(22,26). In the low HF stage, a mixture of HF, dimethyl sulfide (DMS), p-cresol, p-thiocresol (25/65/7.5/2.5%, v/v, respectively) was used. The high HF step was carried out by evaporating all the HF and DMS *in vacuo* and recharging the vessel with anhydrous HF. The peptide product was then desalted by gel filtration and purified by preparative HPLC. This sample was submitted for mass spectrometric analysis. A high resolution HPLC analysis of this sample using a C-18 reverse-phase column showed a single symmetric peak (Fig. 1a). Amino acid analysis provided the expected composition of the peptide.

[Bpa-8]dynorphin A (1–17). This peptide, also submitted to us for mass spectrometric analysis, was synthesized by the same Merrifield solid-phase synthesis protocol as used by Yamashiro and Li (27) on benzhydrylamine resin. In order to avoid possible side reactions of Bpa with the sulfide nucleophiles generally used in low–high HF deprotection, a different deprotection scheme was adopted. First a high HF treatment was carried out with HF and p-cresol (11/1, v/v) to remove all
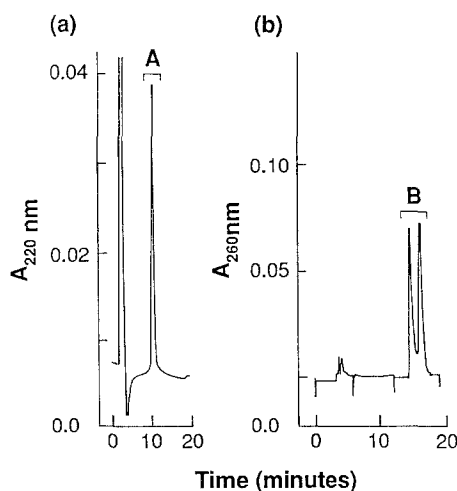
FIG. 1. Reverse-phase HPLC traces of (a) the mellitin products and (b) the [Bpa-8]dynorphin (1–17) products. These samples were previously purified by preparative HPLC (see Materials and Methods). In (a), 4 μg of the purified mellitin sample in 10% acetic acid was injected onto a Vydac C-18 reverse-phase column equilibrated in 0.04% trifluoroacetic acid. After a 10-min wash, the column was eluted with 38% CH₃CN and 0.04% trifluoroacetic acid. In (b), 50 μg of the purified [Bpa-8]dynorphin (1–17) sample in distilled water was injected onto a Vydac C-18 reverse-phase column equilibrated in 25% acetonitrile. After a 10-min wash, the column was eluted with a 30-min linear gradient of 25.0–27.5% acetonitrile in 0.1% trifluoroacetic acid. The peaks labeled A and B correspond to the mellitin products and [Bpa-8]-dynorphin (1–17) products, respectively.

the protecting groups with the exception of Trp(For). The peptide was then purified from the crude product by preparative reverse-phase HPLC. Finally, to remove the formyl group from the purified [Bpa-8][Trp(For)-14]-Dyn A (1–17), the peptide was treated with 0.2 M piperidine (distilled) for 15 min. A high resolution HPLC analysis of this material using a C-18 reverse-phase column showed two clearly separated peaks with approximately equal absorptions (λ = 260 nm, the absorption maximum of the benzophenone group) (Fig. 1b). Further treatment with piperidine for 1 h produced no significant change in the HPLC trace. The sample used in the present investigation was obtained after this second deprotection treatment with piperidine and contained both products.

[Bpa-8][Trp(For)-14]dynorphin A (1–17). This compound was obtained as an intermediate during the synthesis of [Bpa-8]Dyn A (1–17) as described above.

Methods. The detailed procedure for the mass spectrometric investigation of microscale reaction products of surface bound samples has been described previously (18,19). A brief description of this method and the additional steps taken in the present investigation are given below.

Sample preparation. Samples were prepared for mass spectrometric analysis by adsorption of the peptide(s) from solution onto a thin NC film (28,18), coated onto the surface of the mass spectrometer sample inser-

tion probe. The NC film was produced by electrospraying 50 μg of NC (1 mg/ml in acetone) onto a flat, thin (2 μm) aluminized polyester support with a surface area of 1 cm². One nanomole of peptide dissolved in 3 μl of 0.1% TFA was spread on the NC layer. Following adsorption of the peptide to the NC surface and evaporation of the solvent, the sample foil was inserted into the vacuum lock of the mass spectrometer where the film was thoroughly dried by evacuation. The resulting bare layer of peptide molecules bound to the surface of the NC was then inserted into the mass spectrometer for analysis.

²⁵²*Cf plasma desorption mass spectrometry.* The mass spectra were obtained with The Rockefeller University ²⁵²Cf fission fragment ionization time-of-flight mass spectrometer which has been described previously (29,11). Spectral accumulation times ranged between 1 and 10 h depending on the yield of ions obtained from the sample under investigation. It should be noted that the fission fragment flux through the sample (5000 fission fragments/s) is sufficiently low to cause negligible damage (<1 part in 10³) to the total sample over the time scale of these measurements. Thus the effect of increasing the spectrum accumulation time is simply to increase the counting statistics and the signal-to-noise ratio. In addition, the practically nondestructive nature of this mass spectrometric technique allows us to carry out microscale chemical reactions of the surface-bound peptides with various enzymes subsequent to the initial mass analysis. The resulting products from the enzymatic reaction(s) can then be analyzed mass spectrometrically to obtain structural details of the peptide(s). The accuracy of the mass determinations was generally better than 300 ppm.

Enzyme-catalyzed reactions and subsequent mass spectrometric analysis of the reaction products. The sample insertion probe containing the NC-bound peptide was removed from the mass spectrometer after obtaining the mass spectrum of the unreacted peptide. Between 10^{-11} and 10^{-10} mol of the desired enzyme was applied to the surface of the NC-bound peptide in a 5-μl drop. The enzyme solution was spread and maintained in contact with the sample using a microscope coverslip. The reactions were carried out at 37°C for periods ranging from 5 to 20 min. The coverslip was then removed and the reacted surface was inserted into the mass spectrometer where the volatile reagents were evaporated and the spectrum of the remaining reaction products was obtained. In some cases further reactions were carried out on the same surface and were in turn reinvestigated mass spectrometrically.

Solutions of proteinase K (pH 8.5–9.0), α-chymotrypsin (pH 8.5), trypsin (pH 8.0), and carboxypeptidase B (pH 8.0) were prepared in 0.05 M NH₄HCO₃.

RESULTS AND DISCUSSION

Melittin. High resolution reverse-phase HPLC analysis (λ = 220 nm) of the purified synthetic melittin sam-
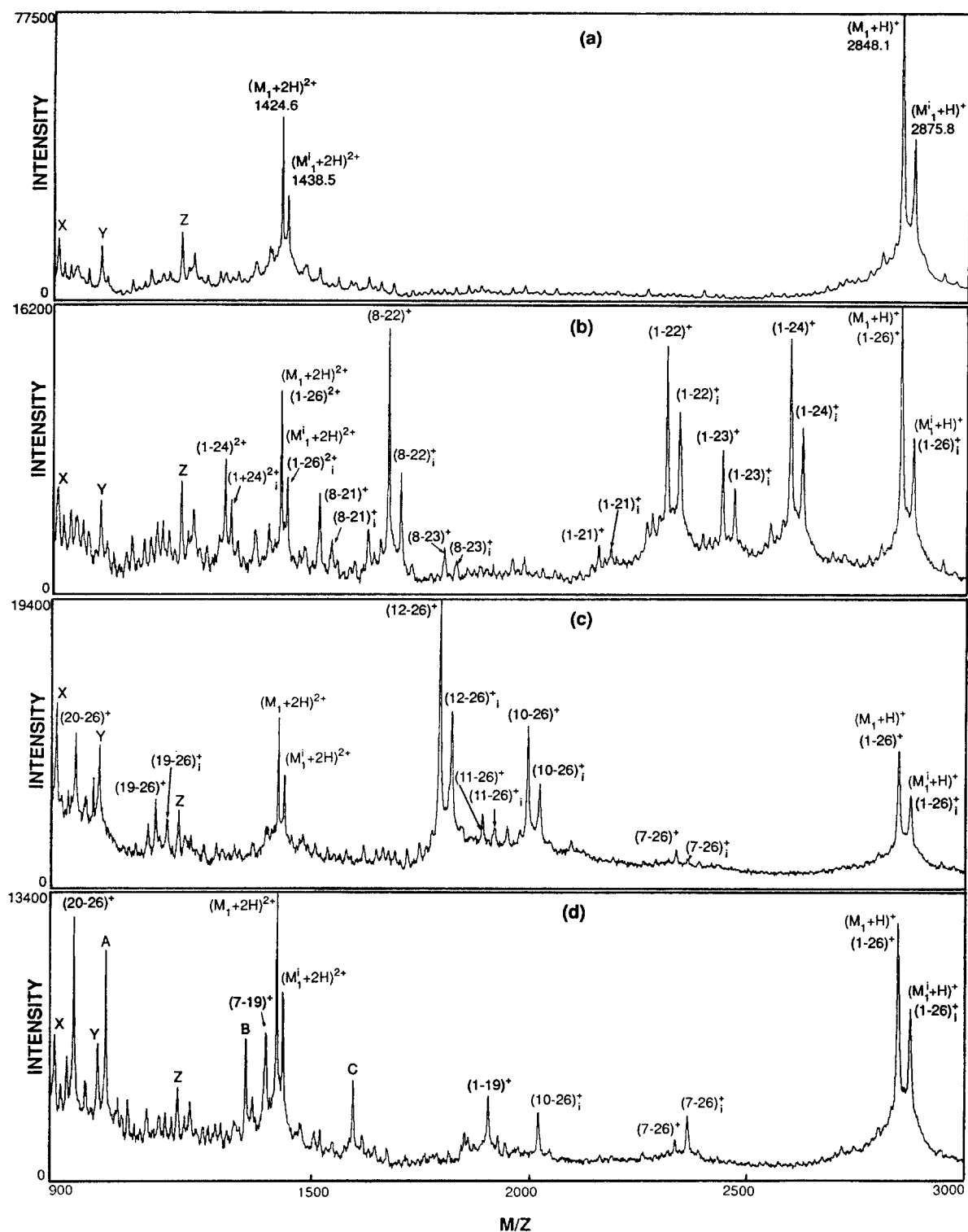
FIG. 2. ^{252}Cf plasma desorption time-of-flight mass spectra of the purified synthetic mellitin product (10^{-9} mol) bound to a thin film of nitrocellulose. (a) Prior to any enzymatic treatment. $(M_1+H)^+$ and $(M_1+2H)^{2+}$ are singly and doubly protonated ions of authentic mellitin. $(M_1^i+H)^+$ and $(M_1^{i+}2H)^{2+}$ are similar ions of an impurity with molecular mass ca. 28 Da higher than that of mellitin. X, Y, and Z result from unidentified impurities. (b) Following a 10-min digestion with trypsin (10^{-11} mol, pH 8.0) at 37°C. The residues contained in each hydrolysis product ion are indicated on top of the corresponding peak. For example, $(1-22)^+$ represents the protonated tryptic fragment of mellitin, M_1, which contains residues 1 to 22 and $(1-22)_i^+$ represents the protonated ion of the corresponding tryptic fragment from the mellitin impurity, M_1^i. The mass separation between the members of each such pair of ions is 28 Da. The observed and calculated m/z values are given in Table 1. (c) Following a 10-min reaction with proteinase K (10^{-10} mol, pH 8.5) at 37°C. (d) Following a 10-min reaction with α-chymotrypsin (10^{-10} mol, pH 8.5) at 37°C. Ions A, B, and C were not identified.

TABLE 1

Comparison of Observed and Calculated Molecular Weights for Tryptic Fragments
of NC-Bound Mellitin and Mellitin Impurity

Tryptic product ion species[a]	Observed molecular weight	Calculated molecular weight	Δ[b]	Tryptic product ion species[a,c]	Observed molecular weight	Calculated molecular weight	Δ[b]
$(1-7)^+$	656.7	656.7	0.0	$(1-7)_i^+$	—	684.7	—
$(8-21)^+$	1511.0	1511.9	−0.9	$(8-21)_i^+$	1538.9	1539.9	−1.0
$(8-22)^+$	1668.1	1668.1	0.0	$(8-22)_i^+$	1696.2	1696.1	+0.1
$(8-23)^+$	1796.3	1796.3	0.0	$(8-23)_i^+$	1824.0	1824.3	−0.3
$(1-21)^+$	2150.6	2150.6	0.0	$(1-21)_i^+$	2178.2	2178.6	−0.4
$(1-22)^+$	2306.8	2306.8	0.0	$(1-22)_i^+$	2335.2	2334.8	+0.4
$(1-23)^+$	2434.8	2435.0	−0.2	$(1-23)_i^+$	2462.7	2463.0	−0.3
$(1-24)^+$	2591.3	2591.2	+0.1	$(1-24)_i^+$	2619.2	2619.2	0.0
$(1-24)^{2+}$	2591.2^d	2591.2	0.0	$(1-24)_i^{2+}$	2619.0	2619.2	−0.2
$(1-26)^+$	Calibration mass[e]	2846.4	—	$(1-26)_i^+$	2874.5	2874.4	+0.1
$(1-26)^{2+}$	2846.8	2846.4	+0.4	$(1-26)_i^{2+}$	2875.0	2874.4	+0.6

[a] The ion species is designated $(j-k)^{n+}$. j denotes the amino-terminal residue and k the carboxyl-terminal residue of the tryptic fragment of mellitin. n denotes the number of charges (protons) attached to the fragment.

[b] Δ, observed molecular weight − calculated molecular weight.

[c] i denotes the tryptic fragments obtained from the impurity mellitin species.

[d] The molecular weights of doubly protonated species are obtained by multiplying the m/z values by 2 and then subtracting 2.0 Da.

[e] To obtain the highest possible mass accuracy for the tryptic fragments, a calibration was made using the Na ion at m/z 23 and the protonated mellitin ion at m/z 2846.4 as described previously (29).

ple yielded a single component peak suggesting that the sample was pure. Inspection of the ^{252}Cf plasma desorption mass spectrum obtained from 10^{-9} mol of this mellitin sample bound to NC (Fig. 2a), however, indicated the presence of two distinct peptide components (designated M_1 and M_1^i). Each of these components is observed in the form of a singly and a doubly protonated ion species. The measured m/z values of the ions at 2848.1 and 1424.6 correspond closely to those expected for singly and doubly protonated melittin (calculated values are 2847.4 and 1424.2, respectively). The m/z values of the ions at 2875.8 and 1438.5 correspond to a singly and doubly protonated impurity with molecular mass 28 Da higher than that measured for mellitin. These m/z values were obtained using a calibration based on the known m/z values of the H^+ and Na^+ ions which are always present in plasma desorption mass spectra (11). Because of the limited resolution of the plasma desorption time-of-flight mass spectrometer, individual isotopic peaks are not resolved above m/z 1000. Molecular weight measurements are therefore made by determining the centroid of the envelope of the isotope distribution. The resulting average molecular weights are then compared with the average molecular weights calculated using the natural isotopic abundances (11). The $(M_1+H)^+$ and $(M_1^i+H)^+$ ion peaks (Fig. 2a) are observed with relative intensities of 70:30. Since these two peptides are almost identical (see later), their mass spectrometric responses are expected to be very similar. Thus the observed relative ion intensities are expected to closely reflect the relative amounts of the two compo-

nents present in the sample. Since Trp(For) was used in the synthesis, the observed impurity is likely the result of either an incomplete deprotection of Trp(For) or a transformylation modification reaction during deprotection (3,20,22). The simple molecular weight of the impurity does not, however, provide direct information on the actual location of the formyl group. In order to extract such information, we carried out a series of enzyme-catalyzed hydrolysis reactions on the sample mixture containing both the authentic melittin and the impurity. The procedure is described below.

The sample used to obtain the spectrum shown in Fig. 2a was removed from the mass spectrometer and treated with 10^{-11} mol of trypsin for 10 min at 37°C. The trypsin-
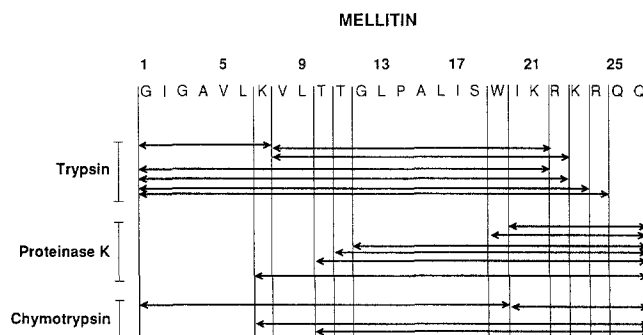

FIG. 3. A summary of all the mass spectrometrically observed enzymatically generated fragments of mellitin. These results are obtained from the data shown in Fig. 2 for the digestion of mellitin by trypsin, proteinase K, and α-chymotrypsin, respectively.
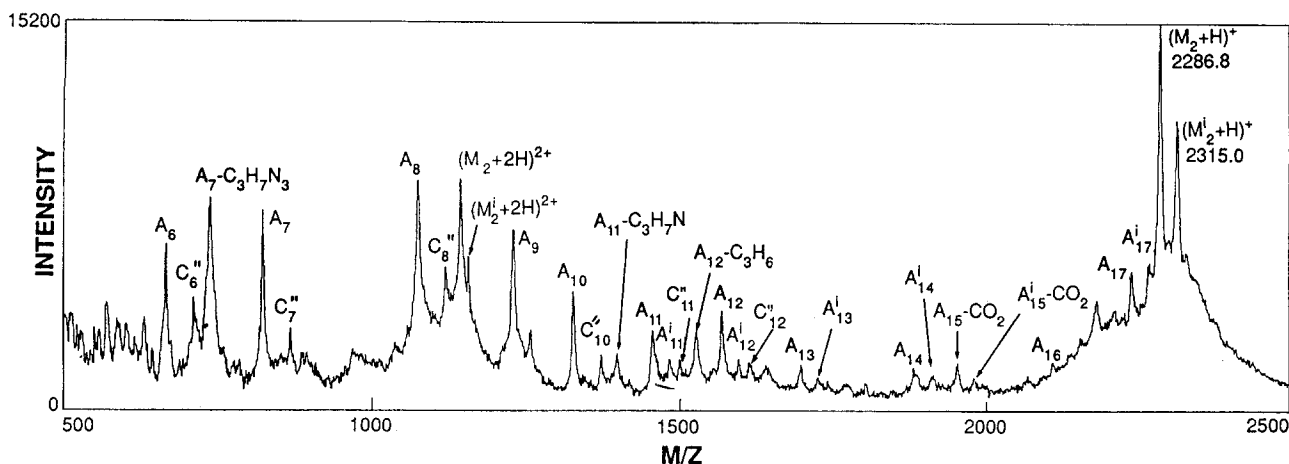
FIG. 4. Fission fragment mass spectrum of 10^{-9} mol of [Bpa-8]Dyn A sample bound to a NC surface. $(M_2+H)^+$ and $(M_2^i+H)^+$ are protonated ions of respectively intact [Bpa-8]Dyn A and an impurity present in the sample with molecular mass 28 Da higher than that of M_2. Three different series of fragment ions are present, denoted Aj, C_j'', and Aj-R. These labels refer to the fragmentation notation shown in Scheme I and described in the text. R denotes portions of the side groups of various amino acid residues. Fragment ions formed from the impurity M_2^i are designated with a superscript i. Since A series ions from M_2^i occur at the same masses as B series ions from M_2, the peaks labeled A_j^i cannot be unambiguously assigned.

treated sample was then reinserted into the mass spectrometer to give the spectrum shown in Fig. 2b. A series of new peaks corresponding to products of hydrolysis at the carboxytermini of the lysine and arginine residues present in the melittin and the melittin impurity were clearly observed in the mass spectrum. The identities of these products, as deduced from the measured m/z values given in Table 1, are indicated above the peaks in Fig. 2b. The series of observed tryptic fragments are summarized diagrammatically in Fig. 3 together with the results from other enzymatic treatments (see later). The tryptic fragments observed in Fig. 2b occur as paired peaks, 28 Da apart. The lower mass component of each pair of peaks (e.g., $(1-22)^+$ representing protonated melittin $(1-22)$) originates from the authentic melittin, while the upper component (e.g., $(1-22)_i^+$) originates from the formylated melittin impurity. Since all the fragments are observed as doublets we can deduce that the formyl group in the impurity, M_1^i, is located between residues 8 and 21 inclusive.

The position of the formyl group in the impurity melittin (M_1^i) was more tightly established by treatment with proteinase K of a freshly deposited sample of the mixture on NC. The mass spectrum of the resulting hydrolysis products (Fig. 3) is shown in Fig. 2c. All of the enzyme-generated fragment ions in the mass spectrum occur as pairs 28 Da apart with the exception of the ion at the m/z 957.3. We attribute this ion to the fragment $(20-26)^+$ having a calculated m/z 956.1. The absence of a companion ion peak 28 Da higher indicates that the formyl group in the impurity melittin is not located between residues 20 and 26 inclusive. On the other hand, the presence of the pair $(19-26)^+$ (observed m/z 1143.2, calculated m/z 1143.3) and $(19-26)_i^+$ (observed m/z 1170.8, calculated m/z 1171.3) which are ca. 28 Da apart

indicates that the formyl group is present between residues 19 and 26 inclusive of the impurity melittin. Therefore, the formyl group is located on tryptophan-19 of the impurity melittin species.

Supporting evidence for the location of the formyl group on tryptophan-19 was deduced from the mass spectrum (Fig. 2d) obtained after α-chymotrypsin treatment of the melittin sample (Fig. 3) bound to NC for 10 min at 37°C. Here too a single peak $(20-26)^+$ with m/z 956.6 (calculated m/z 956.1) was observed with no accompanying $(20-26)_i^+$ peak, again indicating the absence of the formyl group in residues 20-26. Interestingly, the peak which we attribute to $(1-19)^+$ (observed m/z 1910.2, calculated m/z 1910.3) also has no corresponding $(1-19)_i^+$ peak. One would deduce from this finding that the formyl group is also absent from residues 1-19. However, this is not the case because the presence of the formyl group on the indole nitrogen of the Trp greatly reduces the rate of hydrolysis at Trp by α-chymotrypsin. This result was confirmed in a separate experiment by comparing the rates of hydrolysis by α-chymotrypsin of [Bpa-8][Trp(For)-14]Dyn A (1-17) with [Bpa-8]Dyn A (1-17). We found (data not shown) that the rate of hydrolysis at Trp(For) was more than an order of magnitude slower than that at unprotected Trp. Thus the nonobservation of $(1-19)_i^+$ provides additional evidence that the formyl group is present on Trp-19 in the impurity. Another manifestation of the reduced rate of hydrolysis at Trp(For)-19 is the enhanced intensities observed in Fig. 2d of $(10-26)_i^+$ and $(7-26)_i^+$ compared to those of $(10-26)^+$ and $(7-26)^+$.

[Bpa-8]dynorphin A (1-17). High resolution reverse-phase HPLC analysis (λ = 260 nm) of the synthetic dynorphin product yielded two closely spaced peaks with nearly equal intensities (Fig. 1b). Inspection
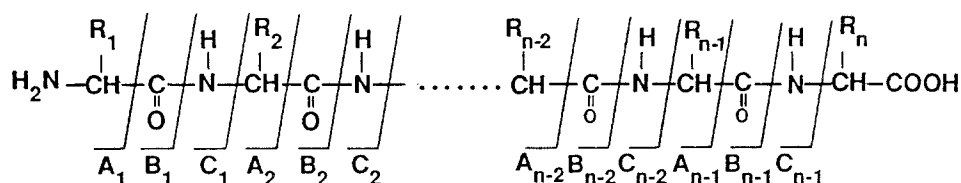
SCHEME 1

of the mass spectrum of 10^{-9} mol of this product bound to NC (Fig. 4) also showed the presence of two peptide components, M_2 and M_2^i, with observed relative intensities (peak areas) of 58:42. The most intense ion $(M_2+H)^+$ had an observed m/z 2286.8 in close agreement with the value of 2286.5 calculated for [Bpa-8]Dyn A (1–17). The ion arising from the second component, observed at m/z 2315.0 is an impurity with m/z ca. 28 Da higher than that measured for authentic [Bpa-8]Dyn A (1–17). The 28 Da mass difference implies the presence of a formyl group in this dynorphin impurity. Doubly protonated ions of the two components are also observed in the mass spectrum. The other ions in Fig. 4 result from the unimolecular fragmentation of the $(M_2+H)^+$ and $(M_2^i+H)^+$ during the ^{252}Cf plasma desorption ionization process. The identities of these fragment ions are indicated by the peak labels which refer to the notation (30) in Scheme I. The double prime on the C-series notation used in Fig. 4 indicates the transfer of a hydrogen atom to the protonated fragment. It should be noted that the A series fragment ions from the Dyn impurity (M_2^i) occur at the same masses as the B series fragment ions from [Bpa-8]Dyn A (1–17) (M_2). Because of this unfortunate coincidence, the site of attachment of the formyl group on the dynorphin impurity cannot be unambiguously deduced from the unimolecular fragmentation data. However, since no fragment ions are observed 28 Da higher

than A_6, A_7, A_8, and A_{10}, we can conclude that the formyl group is absent from residues 1–10 of M_2^i and therefore must be present in residues 11–17. In order to obtain more specific information regarding the site of formylation we subjected the NC-bound Dyn sample mixture to enzymatic degradation.

The mass spectrum obtained following a 10-min digestion of 10^{-9} mol of the NC-bound Dyn sample mixture with 10^{-10} mol of α-chymotrypsin (pH 8.5) at 37°C is given in Fig. 5. Intense pairs of chymotryptic fragment ions containing residues 5–14, 6–14, and 1–14 are observed. The members of each pair are separated by 28 Da. It should be noted that, in contrast to the results obtained with the melittin sample mixture (Fig. 2d), the impurity Dyn (M_2^i) is cleaved at the Trp-14 by α-chymotrypsin. Furthermore, the relative intensities of the fragment ions in each pair are similar to the relative intensities of $(M_2+H)^+$ and $(M_2^i+H)^+$. This implies that the rates of the cleavage reaction at Trp-14 in both the authentic and the impurity Dyn A are similar. By contrast, in a separate experiment we determined that the rate of cleavage of formylated Trp in [Bpa-8][Trp(For)-14]Dyn A (1–17) by α-chymotrypsin was more than an order of magnitude slower than that of unformylated Trp. From the observation of the pair of chymotryptic fragment ions $(6–14)^+$ and $(6–14)_i^+$ we thus deduce that the formyl group has migrated to the region between residues 6 and
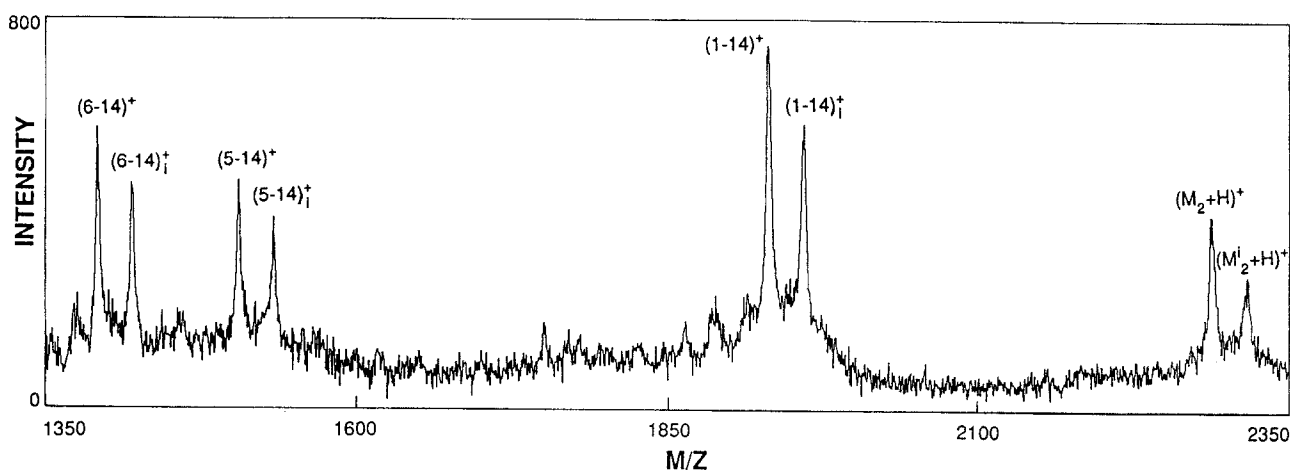

FIG. 5. Mass spectrum of the [Bpa-8]Dyn sample mixture bound to a NC surface after digestion with 10^{-10} mol α-chymotrypsin (pH 8.5 for 10 min at 37°C). $(M_2+H)^+$ and $(M_2^i+H)^+$ are respectively protonated [Bpa-8]Dyn A and an impurity of mass 28 Da higher than that of M_2. Chymotryptic fragment ions formed from the hydrolysis of the impurity M_2^i are designated with a subscript i.
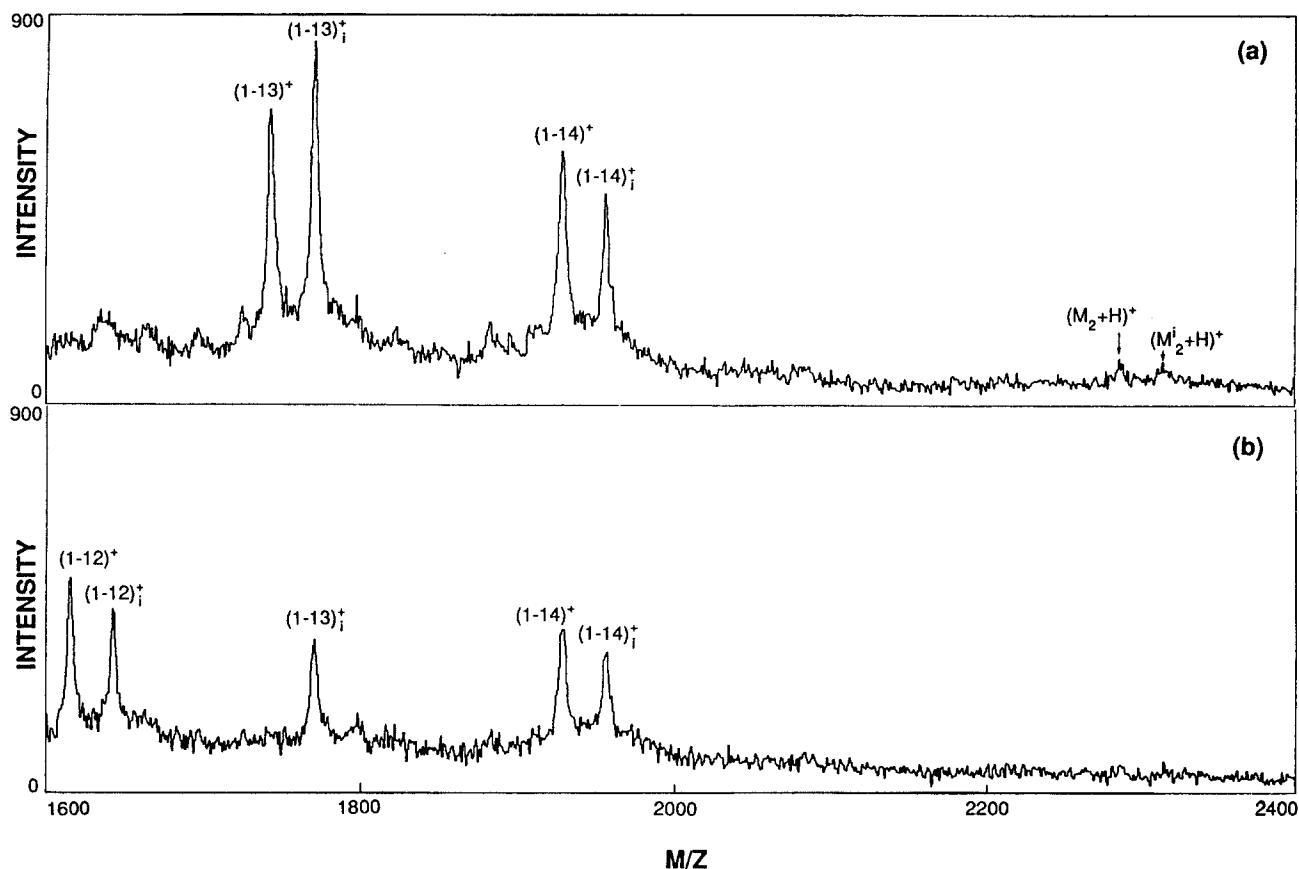
FIG. 6. Mass spectra of 10^{-9} mol of the [Bpa-8]Dyn A sample mixture following (a) a 10-min digestion with proteinase K (10^{-10} mol, pH 9.0) and (b) a successive treatment of the proteinase K-digested sample shown above with carboxypeptidase B (10^{-11} mol, pH 8) for 10 min at 37°C.

13 inclusive. We previously deduced from the mass spectrometric fragmentation data that the formyl group was present in residues 11–17. Combining these two results we conclude that the formyl group has migrated to the region 11–13.

In order to determine which of residues 11–13 is/are formylated in M_2^i, a freshly deposited sample of the Dyn mixture was treated with proteinase K (10^{-10} mol for 10 min at 37°C, pH 9.0) on NC. A portion of the mass spectrum (m/z 1600–2400) of the resulting hydrolysis products is shown in Fig. 6a. Peaks corresponding to M_2 and M_2^i are almost absent indicating virtually complete enzymatic fragmentation of the starting materials. The observation of the pair of peaks $(1-13)_i^+$ again shows that the formyl group has migrated to residues 1–13. A number of other structurally informative fragment ions were observed (data not shown) including the pair corresponding to $(6-11)^+$ and $(6-11)_i^+$. Taken together with the previously described results, the observation of these latter peaks demonstrate that at least some of the formyl groups have migrated to Lys-11.

After the mass spectrum of the proteinase K-treated Dyn sample mixture was obtained (Fig. 6a), the NC-bound digestion products were further treated with carboxypeptidase B (CPB) (10^{-11} mol for 10 min at 37°C,

pH 8.0). The spectrum of the resulting reaction products is shown in Fig. 6b. Any C-terminal Lys or Arg residues which are present on the proteinase K-generated fragments are specifically removed by CPB. Thus the doublet comprised of (1–13) and (1–13)$_i$ was cleaved to produce fragments (1–12) and (1–12)$_i$. Of special interest is the complete disappearance of (1–13) and the survival of a portion of (1–13)$_i$. These results can be readily understood if we assume that CPB cleaves C-terminal Lys(For) at a significantly slower rate than Lys. The observation of residual $(1-13)_i^+$ (Fig. 6b) then confirms the presence of some formylation on Lys-13 of M_2^i. The observation of the $(1-12)^+$ and $(1-12)_i^+$ doublet demonstrates that Lys-11 of M_2^i is also partially formylated since it is unlikely that Leu-12 will be thus modified. Additional confirmation of this finding was evident from the presence of $(6-11)^+$ and $(6-11)_i^+$ in the lower portion of the mass spectrum of the proteinase K-treated samples (data not shown). Thus we conclude that the formyl group is no longer located on Trp-14 of M_2^i and has migrated to the neighboring lysine residues.

CONCLUSION

The presence of formylated impurities in two synthetic tryptophan-containing peptide samples was de-

tected by ^{252}Cf plasma desorption mass spectrometry. The sites of formylation were determined using a selection of enzymatic reactions on monolayer amounts of peptide samples bound to nitrocellulose followed by mass spectrometric analysis of the reaction products. In the case of the melittin sample, we found that the impurity was due to residual formyl moieties on Trp-19 which were not efficiently removed during the low HF deprotection procedure (25,22). By contrast, in the case of the [Bpa-8]Dyn A (1–17) sample, we found that the impurity originated from the migration of the Trp formyl group to the neighboring Lys-11 or Lys-13 residues during piperidine deprotection.

We have found through the analysis of the peptides discussed here as well as a number of other Trp-containing synthetic peptides (e.g., see Ref. (14)) that complete removal of the Trp formyl group is highly problematic. Since reverse-phase HPLC does not always resolve the formylated from the deformylated peptide, the presence of these impurities frequently goes undetected. Our findings also indicate that the spectrophotometric method often used to assess the degree of formyl deprotection of Trp may not always be adequate. Thus the present technique provides valuable information to the synthetic peptide chemist regarding the efficiency of the chosen deprotection method and the presence and identity of unwanted by-products.

The use of the present technique is by no means limited to providing information involving formylated impurities. We believe that many other peptide by-products can be studied in a manner similar to that described in this paper and that the technique will provide valuable complementary information to that provided by the more established methods.

ACKNOWLEDGMENTS

REFERENCES

1. Merrifield, R. B. (1963) *J. Amer. Chem. Soc.* **85**, 2149–2154.

2. Erickson, B. W., and Merrifield, R. B. (1976) *in* The Proteins (Neurath, H., and Hill, R. L., Eds.), 3rd ed., Vol. 2, pp. 255–527, Academic Press, New York.

3. Barany, G., and Merrifield, R. B. (1980) *in* The Peptide (Gross, E. and Meienhofer, J., Eds.), Vol. 2, Part A, pp. 1–284, Academic Press, New York.

4. Kent, S. B. H. (1988) *Annu. Rev. Biochem.* **57**, 957–990.

5. Larsen, B., Fox, L. B., Burke, M. F., and Hruby, V. J. (1979) *Int. J. Pept. Protein Res.* **13**, 12–21.

6. Hare, P. E. (1975) *in* Protein Sequence Determination (Needleman, S. B., Ed.), Vol. 8, pp. 204–231, Springer-Verlag, New York.

7. Edman, P., and Henischen, A. (1975) *in* Protein Sequence Determination (Needleman, S. B., Ed.), Vol. 8, pp. 232–279, Springer-Verlag, New York.

8. Merrifield, R. B., Vizioli, L. D., and Boman, H. G. (1982) *Biochemistry* **21**, 5020–5031.

9. Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids, pp. 1–292, Wiley-Interscience, New York.

10. Bayer, E., Eckstein, H., Hagele, K., Konig, W. A., Biuring, W., Hagenmaier, H., and Parr, W. (1970) *J. Amer. Chem. Soc.* **92**, 1735–1738.

11. Chait, B. T., Gisin, B. F., and Field, F. H. (1982) *J. Amer. Chem. Soc.* **104**, 5157–5162.

12. Biemann, K., and Martin, S. A. (1987) *Mass Spectrom. Rev.* **6**, 1–76.

13. Biemann, K., and Scoble, H. A. (1987) *Science* **237**, 992–998.

14. Chait, B. T. (1988) *in* The Analysis of Peptides and Proteins by Mass Spectrometry (McNeal, C. J., Ed.), pp. 21–39, Wiley, New York.

15. McFarlane, R. D., and Torgerson, D. F. (1976) *Science* **191**, 920–925.

16. McLafferty, F. W. (Ed.) (1983) Tandem Mass Spectrometry, Wiley, New York.

17. Chait, B. T., and Field, F. H. (1985) *J. Amer. Chem. Soc.* **107**, 6743–6744.

18. Chait, B. T., and Field, F. H. (1986) *Biochem. Biophys. Res. Commun.* **134**, 420–426.

19. Chait, B. T., Chaudhary, T., and Field, F. H. (1987) *in* Methods in Protein Sequence Analysis (Walsh, K. A., Ed.), pp. 483–491, The Humana Press, Clifton, NJ.

20. Yamashiro, D., and Li, C. H. (1973) *J. Org. Chem.* **38**, 2594–2597.

21. Ohno, M., Tsukamoto, S., Makisumi, S., and Izumiya, N. (1972) *Bull. Chem. Soc. Japan* **45**, 2852–2855.

22. Tam, J. P. (1988) *in* Macromolecular Sequencing and Synthesis: Selected Methods and Applications, pp. 153–184, Alan R. Liss, New York.

23. Habermann, E., and Jentsch, J. (1967) *Hoppe Seyler's Z. Physiol. Chem.* **348**, 37–50.

24. Matsueda, G. R., and Stewart, J. M. (1981) *Peptides* **2**, 45–50.

25. Christensen, B., Fink, J., Merrifield, R. B., and Mauzerall, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 5072–5076.

26. Tam, J. P., Heath, F. W., and Merrifield, R. B. (1983) *J. Amer. Chem. Soc.* **105**, 6442–6455.

27. Yamashiro, D., and Li, H. (1978) *J. Amer. Chem. Soc.* **100**, 5174–5179.

28. Jonsson, G. P., Hedin, A. B., Hakansson, P. L., Sundqvist, U. R., Save, B. G., Neilsen, P. F., Roepstorff, P., Johansson, K., Kamensky, I., and Lindburg, M. S. L. (1986) *Anal. Chem.* **58**, 1084–1087.

29. Chait, B. T., Agosta, W. C., and Field, F. H. (1981) *Int. J. Mass Spectrom. Ion Phys.* **39**, 339–366.

30. Roepstroff, P., and Fohlman, J. (1984) *Biomed. Mass Spectrom.* **11**, 601.