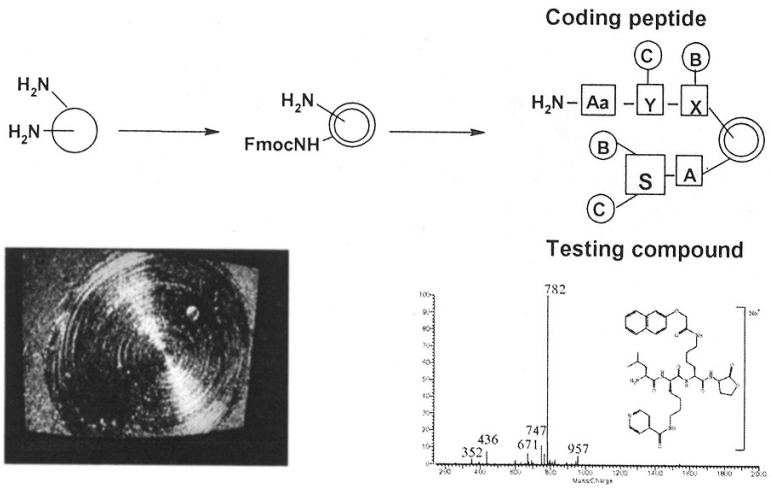
JOURNAL OF combinatorial CHEMISTRY

Article

High-Throughput One-Bead-One-Compound Approach to Peptide-Encoded Combinatorial Libraries: MALDI-MS Analysis of Single TentaGel Beads

Andreas H. Franz, Ruiwu Liu, Aimin Song, Kit S. Lam, and Carlito B. Lebrilla

J. Comb. Chem., **2003**, 5 (2), 125-137• DOI: 10.1021/cc020083a • Publication Date (Web): 18 January 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

Journal of Combinatorial Chemistry is published by the American Chemical Society.
1155 Sixteenth Street N.W., Washington, DC 20036

High-Throughput One-Bead-One-Compound Approach to Peptide-Encoded Combinatorial Libraries: MALDI-MS Analysis of Single TentaGel Beads

Andreas H. Franz,[†] Ruiwu Liu,[‡] Aimin Song,[‡] Kit S. Lam,*[,‡] and Carlito B. Lebrilla*[,†]

Department of Chemistry, University of California Davis, One Shields Avenue, Davis, California 95616, and Department of Internal Medicine, University of California Davis Cancer Center, 4501 X Street, Sacramento, California 95817

Received September 19, 2002

The identification of pharmacologically promising compounds (lead compounds) from combinatorial libraries is frequently limited by the throughput of the analytical technique employed. Fourier transform mass spectrometry (FTMS) offers high sensitivity, mass accuracy ($m/\Delta m$ > 500 000), and sequencing capabilities. A rapid and efficient method for high-throughput analysis of single beads from peptide-encoded combinatorial libraries with matrix-assisted laser desorption/ionization (MALDI) mass spectrometry is presented. Encoding peptides on single beads are identified and structurally characterized by MALDI time-of-flight (TOF) and ultrahigh-resolution MALDI Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometry. A strategy of on-probe sample preparation is developed to minimize handling of the beads.

Introduction

The advent of combinatorial chemistry in the field of drug development in the late 1980s has flourished into an independent field of research.[1−7] One particular combinatorial approach is the one-bead-one-compound (OBOC) method that was introduced more than 10 years ago.[8] Using a split−mix synthesis technique,[8−10] a library of compound beads can be rapidly generated such that each single polymeric bead carries one particular compound. The bead library is then screened, and the positive compound beads are physically isolated for structural elucidation. The synthesis and screening for the OBOC combinatorial library are fast, and tens of thousands to millions of compounds can be synthesized and screened in a week or two. Successful applications of the OBOC approach to the identification of protein kinase inhibitors, integrin-specific peptides for human prostate cancer, idiotype-specific peptides for murine lymphoma cells, MHC molecule anchor residues, and antigenic epitopes for monoclonal antibodies have been reported and reviewed.[11,12]

Structure determination of the positive compound beads is often the bottleneck of the OBOC combinatorial method. If the compound is a peptide and the N terminus is not blocked, Edman chemistry[13,14] can be employed using an automatic protein sequencer to microsequence the peptide. About four peptides can be microsequenced with one automatic protein sequencer each day. However, if the compound beads contain a small molecule, there is no simple way to directly determine its chemical structure. Multiple encoding methods have been developed to solve this problem, and the subject has been reviewed recently.[12,15−18] The coding tag must meet the following criteria to be useful: The tag has to be chemically variable to encode a sufficiently large number of library compounds, and the tag has to be inert toward the chemistries employed for the synthesis of the library. Furthermore, it is sometimes desirable that the tag can be cleaved off the bead without cleaving the library compound, and the tag's chemical structure can be determined with common analytical techniques. Finally, the tag should not interfere with biological assays that screen the library compounds. The last requirement has been met by topologically segregated bifunctional beads, as described elsewhere.[19−21] Topologically segregated bifunctional beads carry the encoding tag on the inside and expose the library compound on the outside exclusively. Therefore, interference of the tag with biological screening of compounds on the outside of the beads is minimized. In current encoding methods, chemical and nonchemical tags[22,23] have found applications. Chemical tags can be attached to the bead covalently or noncovalently. Covalently bound tags may be removed from individual beads by chemical means at the end of the screening process of the library. If a photolabile anchor is incorporated between the tag and the bead, irradiation with light of appropriate wavelength can be used to break the bond and release the tag from the bead.

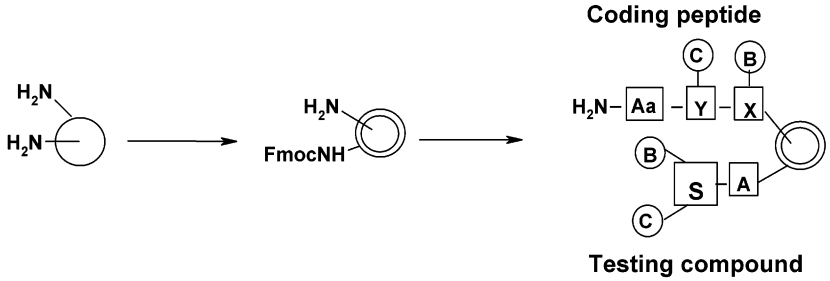
After release of the encoding tag from the polymer support, the tag has to be structurally characterized. The techniques employed in the structural determination of encoding tags have to fulfill three requirements. First, they have to be accurate and reliable. Second, the techniques have to be highly efficient in that a large number of compounds can be analyzed within a short period of time. This principle is referred to as high throughput. Third, the techniques have to be sensitive so that for example a few picomoles of material is sufficient for complete structural elucidation.

We have recently reported on the development of a novel chemical encoding method for OBOC peptidomimetic or small molecule libraries using a topologically segregated

[†] Department of Chemistry, University of California Davis.
[‡] University of California Davis Cancer Center.

10.1021/cc020083a CCC: $25.00
Published on Web 01/18/2003

Scheme 1. Encoding Scheme for Topologically Segregated Bifunctional Beads[a]



[a] The testing compound is displayed on the outer layer of the bead while the coding peptide resides in the interior of the bead. (A represents a bifunctional building block that links the scaffold to the bead, coded by amino acid Aa; S represents the scaffold of the testing arm; X and Y represent trifunctional amino acids whose side chains are derivatized with building blocks B and C, respectively.)

bifunctional bead.[21] In this encoding method, the peptidomimetic or small molecule testing compound is presented on the outer layer of the bead, and the peptide coding tag resides on the interior of the bead (Scheme 1).

The peptide tag consists of α-amino acid derivatives and, therefore, can be readily microsequenced by an automatic protein sequencer without the need to release the coding tag prior to microsequencing. Approximately eight compound beads can be decoded with one protein sequencer using this approach; therefore, the decoding process becomes the rate-limiting step. There is a need for the development of a faster method to decode the compound beads.

The sensitivity and accuracy of mass spectrometry has made this technique the best choice for analysis of compounds for which the amount of sample is a major concern. Mass spectrometry has been used for the structural elucidation of peptides[24,25] and their properties in noncovalent gas-phase complexes.[26,27] Currently, structural information from collision-induced dissociation experiments (tandem MS) can be conveniently obtained from a few picomoles of material; however, attomole amounts of ions have been reported as the absolute lower detection limit.[28−30]

Mass spectrometry has been applied to the analysis of combinatorial libraries.[31−35] A technique that meets the requirements of accuracy and sensitivity is matrix-assisted laser desorption/ionization Fourier transform ion cyclotron mass spectrometry (MALDI FT-ICR MS). In FTMS instruments, a mass resolution of $m/\Delta m$ 500 000 or greater can be achieved. High mass resolution is invaluable in discriminating compounds with a mass difference of $\ll$1 u. The analytical value of ultrahigh resolution to the discrimination of closely spaced peptide signals and isotope signals in FT mass spectra has been successfully demonstrated.[36] On the basis of mass accuracy, large pools of compounds from a combinatorial library can possibly be reduced to a few combinations of building blocks. Subsequently, the actual structure can be further elucidated by collision-induced dissociation (CID) experiments. In this article, MALDI FT-ICR MS was used to quickly and reliably identify the coding peptides from individual beads.

Experimental Section

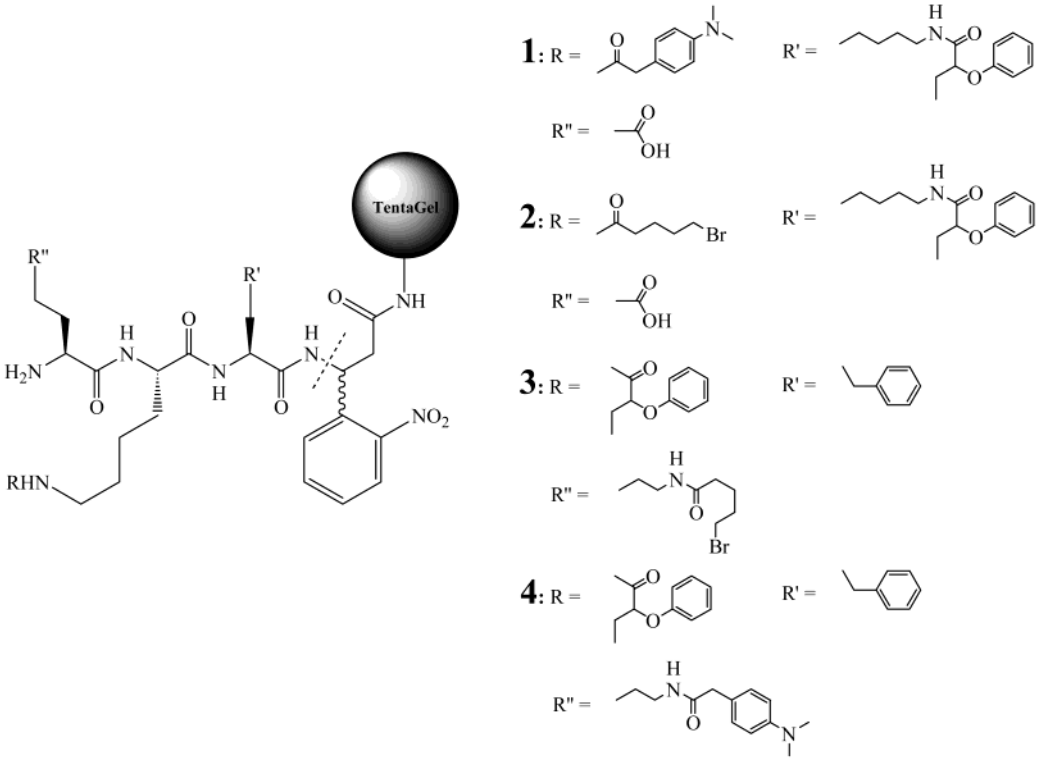
Unless stated otherwise, all solvents and general chemicals were purchased from Aldrich (Milwaukee, WI) and were used without further purification. Solvents were of HPLC grade. All peptides were synthesized on TentaGel S NH_2 resin beads (90-μm diameter, loading 0.26 mmol/g). TentaGel S NH_2 resin was purchased from Rapp Polymere GmbH (Tübingen, Germany). Fmoc-protected amino acids and 9-fluorenylmethyloxycarbonyl-*N*-hydroxysuccinimide (FmocOSu) were purchased from Chem-Impex International, Inc (Wood Dale, IL) and Advanced ChemTech (Louisville, KY). Natural amino acids were designated by the standard three-letter codes.

For the synthesis of each peptide, 10 mg TentaGel resin was used. A typical hydroxybenzotriazole (HOBt)/diisopropylcarbodiimide (DIC) coupling method[37] was employed in the amino acid coupling. A ratio of three fold of amino acids to the amount of resin was used. The average coupling time was 2 h.

Photocleavable Peptides Synthesis. Fmoc-(*R*,*S*)-3-amino-3-(2-nitrophenyl)propionic acid was first coupled to the beads as a photolabile linker by the HOBt/DIC coupling method.[37] The Fmoc was removed with 20% piperidine in DMF (first 5 min, second 15 min). For peptides **1** and **2**, Fmoc-Lys-(Boc)-OH was then coupled to the beads, followed by deprotection of Boc with 50% trifluoroacetic acid (TFA)/2.5% thioanisole in dichloromethane (DCM). The beads were washed with DCM, 2% *N*,*N*-diisopropylethylamine (DIEA) in DMF, DMF, methanol, and DMF, respectively. Acylation of the side chain of lysine with 2-phenoxybutyric acid (10 equiv) was achieved in the presence of HOBt (10 equiv) and DIC (10 equiv) in DMF at room temperature overnight. The removal of Fmoc, coupling of Fmoc-Lys(Boc)-OH, removal of Boc, and acylation of the side chain of lysine with 4-(dimethylamino) phenylacetic acid and 5-bromovaleric acid, respectively, were carried out sequentially in the same manner as above. The Fmoc deprotection was achieved with 20% piperidine in the same manner. The fourth amino acid, Fmoc−Glu(tBu), was coupled to the beads the same way as above. After removal of Fmoc, side chain deprotection was achieved by treating with a mixture of TFA, phenol, thioanisole, water, and triisopropylsilane (82.5:5:5:5:2.5, v/v) at room temperature for 2 h. After neutralization with 2% DIEA/DMF twice, the resin was washed sequentially with DMF, MeOH, and DCM. After being dried under vacuum for 2 h, the peptide beads were then ready for mass analysis.

Peptides **3** and **4** were synthesized in a similar manner as described above. Instead of using lysine as the second and glutamic acid as the fourth amino acid in peptides **1** and **2**, phenylalanine as the second and lysine as the fourth amino acid were used in peptides **3** and **4**.

Scheme 2. Chemical Structures of the Encoding Peptides **1**−**4**



Chemically Cleavable Peptide Synthesis. Chemically cleavable encoding peptides were linked by methionine to the interior of the topologically segregated TentaGel S NH_2 beads (with 50% substitution of the bead for the coding peptide), which were easily prepared by a biphasic solvent approach.[21] In brief, the beads were preswollen in water for 2 days, and the outer layer of the beads was derivatized with FmocOSu (0.5 equiv to the total resin) in DCM/diethyl ether (55/45, v/v) at room temperature for 30 min. Fmoc-Met-OH was coupled to the internal part of the beads in the presence of HOBt and DIC. After removal of Fmoc in both the interior and exterior of beads with 20% peperidine in DMF and the coding peptides, Aa-Lys(R_2)-Lys(R_1), Aa-Phe(4-$NHCOR_2$)-Lys(R_1), and Aa-Glu(R_2)-Lys(R_1) were synthesized in a similar manner as mentioned above. Acylation of the side chains of lysine and Phe(4-NH_2) with carboxylic acids (10 equiv) was achieved in the presence of HOBt (10 equiv) and DIC (10 equiv) at room temperature overnight. Amide formation of the side chain of glutamic acid with *N*-phenylpiperazine (10 equiv) was conducted in the presence of DIEA (10 equiv) and benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphoniumhexafluorophosphate (10 equiv) at room temperature overnight.

Fourier Transform Ion Cyclotron Resonance (FT-ICR) Mass Spectrometry. Mass spectra were recorded on an external source HiResMALDI (IonSpec Corporation, Irvine, CA) equipped with a 4.7-T magnet and a Nd:YAG laser. The emission at 355 nm was used for all experiments. For all CID experiments, the appropriate isolation pulses were programmed starting at 3 s after the initial ionization and followed by SORI excitation at 6 s (1 s, 5 V base to peak, +1000 Hz off-resonance). At a background pressure of 10^{-10} Torr, argon gas was administered through a pulsed valve at 6 and 6.5 s (peak pressure 5×10^{-5} Torr). Final excitation for detection was performed 12 s after the initial laser pulse.

Time-of-Flight (TOF) Mass Spectrometry. The data were recorded using a Proflex III MALDI-TOF instrument (Bruker-Daltonics, Billerica, MA) equipped with a 337-nm nitrogen laser. 2,5-Dihydroxybenzoic acid (DHB, 5 mg/100 μL in ethanol) was used as the matrix in all experiments. A saturated solution of NaCl in methanol was used as dopant. An ammonium resin was used to produce the $[M + H]^+$ quasimolecular ion of the peptides.

Preparation of the Ammonium Resin. The cation-exchange resin (ammonium form) was prepared by a previously published procedure.[38] A H^+-cation-exchange resin (100−200 mesh, Bio-Rad, NY) was stirred in a 1 M ammonium acetate solution for 12 h. The product was filtered and washed with 1 M ammonium acetate solution, deionized water, acetone, and hexane. The resin was dried and stored for future use.

Handling of Beads. The beads were manipulated by means of a glass micropipet tip. The diameter of an open-ended melting point capillary was reduced in the flame of a Bunsen burner. Under a microscope a single bead was picked up with the pipet tip. The tip was deposited into a 0.2-mL plastic microcentrifuge tube, and the beads were rinsed into the tube by 20 μL of analytical grade ethanol. The solvent was evaporated in vacuo. In the case of the modified single-well MALDI probe, the solvent volume was reduced to 2 μL.

Photolytic Cleavage. Several beads of one type were placed in a plastic micro centrifuge tube. Ethanol (50 μL) was added, and the open tube was irradiated with UV light (wavelength 350 nm) for 6 h. The solvent was evaporated

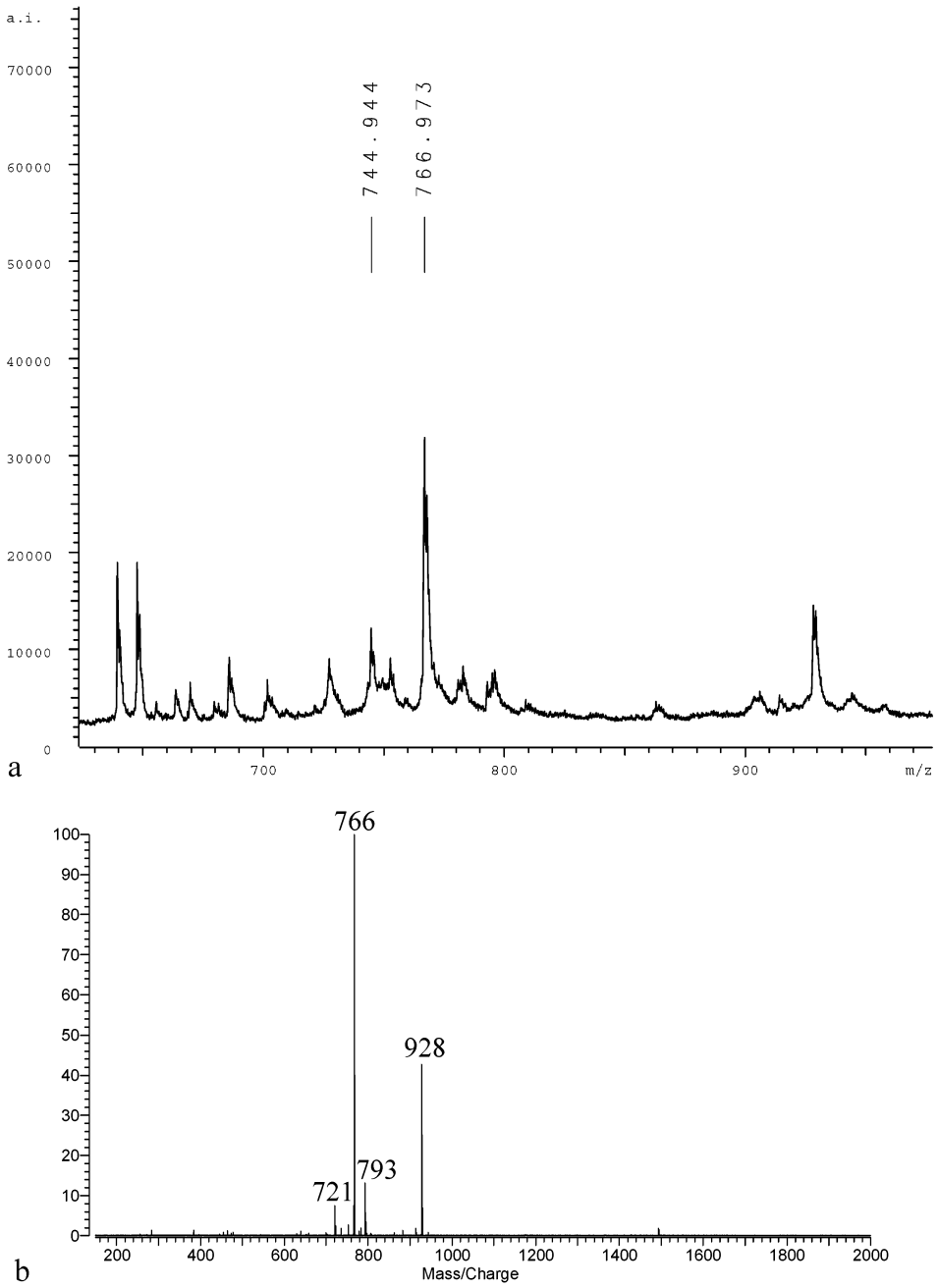


Figure 1. (a) MALDI-TOF spectrum of compound **4** $[M + Na]^{+}$ and (b) MALDI-FTMS spectrum of compound **4** $[M + Na]^{+}$. The signal at *m*/*z* 928 was due to an unknown impurity. Both spectra were obtained with material from several beads.

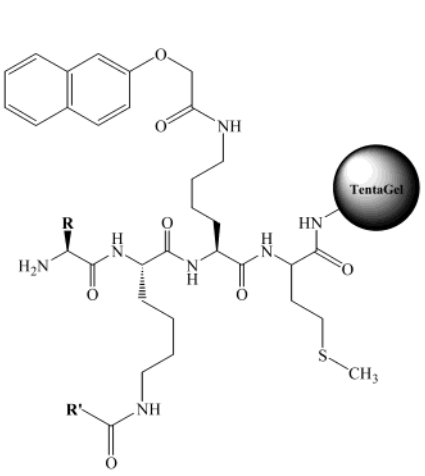
and replaced with acetonitrile/water (80:20, 10 μL). The mixture was allowed to stand at room temperature for 0.5 h to extract the peptides from the interior of the bead. Aliquots were used for MALDI-MS analysis.

Chemical Cleavage.[39] At room temperature, a solution of 0.25 M cyanogen bromide (CNBr) in 70% formic acid was added to the single bead. The mixture was allowed to sit at room temperature overnight (12 h). The solvents were evaporated, and the remainder was dissolved in 50 μL of Nanopure water and evaporated to remove traces of formic acid. The dry sample was redissolved in acetonitrile/water (8:2, 10 μL). This solvent combination gave the best results during extraction of the peptides from the interior of the bead.

Formylation of the terminal amino group in the coding peptides was achieved by warming the reaction mixture to 47 °C for 4 h in the presence of CNBr and formic acid. Shorter reaction times resulted in incomplete formylation.

MALDI Analysis. For MALDI-TOF analysis, 0.5 μL of the sample solution of interest was applied to the stainless steel 10-well sample plate. Sodium dopant (saturated NaCl solution in methanol, 0.5 μL) was added, followed by matrix solution (2,5-dihydroxybenzoic acid, $c = 5$ mg/100 μL in ethanol, 0.5 μL). Flash evaporation of the solvent in a stream of hot air gave a finely crystalline sample that was subjected to mass spectrometric analysis. The ProFLEX III MALDI-TOF instrument (Bruker-Daltonics, Billerica, MD) was equipped with a standard LSI N_2 laser (337 nm) and was operated in linear mode at a laser frequency of 2 Hz for 300 shots. The first extraction lens was charged to 20 kV, and the second extraction lens was varied between 18.5 and 18.9

Scheme 3. Chemical Structures of the Encoding Peptides **5**−**8**



5: R = i-Pr, R' =

6: R = i-Pr, R' =

7: R = , R' =

8: R = , R' =

kV to optimize the energy distribution of the ions and the resolution of the instrument. Data were collected in the shortest time window (1.0 ns) corresponding to an *m*/*z* range of 0−5000.

Results

Analysis of Encoding Peptides (Photolabile Linker). To illustrate the concept of MS analysis of peptides released from single TentaGel beads, four encoding peptides were synthesized on separate beads. Topologically segregated bifunctional beads were not used in this study. The beads were fully derivatized on the outside and inside layers for simplicity. The chemical structures of the peptides are shown in Scheme 2. The *o*-nitrophenyl moiety was cleaved (dashed line) by irradiation with UV light for several hours, as described in the Experimental Section.

Peptides **1**−**4** were chemically cleaved by irradiation with UV light for 6 h, as described in the Experimental Section. The compounds obtained were subjected to MS analysis. Figure 1 shows the MALDI-TOF and the MALDI-FTMS spectrum of compound **4**.

The quasimolecular ion $[M + Na]^+$ was detected at *m*/*z* 766.9 with the TOF analyzer and at *m*/*z* 766.470 with the FT analyzer, respectively. The theoretical mass of the sodiated species was calculated to be 766.427 u, in excellent agreement with the observed value. In the TOF spectrum (Figure 1a), the protonated molecule $[M + H]^+$ was observed at *m*/*z* 744.9, which was completely absent in the FTMS spectrum (Figure 1b). TOF analyzers have significantly shorter detection periods (∼ nanoseconds) than FT analyzers (∼ seconds). Frequently, ions that have a metastable character decompose before they can be observed with an FT analyzer. Apparently, the protonated molecule of compound **4** at *m*/*z* 744.9 had significant metastable character and decomposed during extended detection times in the ICR cell.

All other peptides were investigated analogously. However, only compound **4** was found to be fully consistent with the proposed structure. Compounds **1**, **2**, and **3** yielded inconsistent or partly inconsistent data (spectra not shown). The exact reason for the inconsistencies was not determined. However, it can be assumed that unspecific side reactions during irradiation occurred. For example, the spectrum of compound **3** did not reveal any bromine isotope pattern in the mass spectrum (spectrum not shown). Presumably, the bromine was eliminated from the compound during irradiation or cyclization with the adjacent amide occurred. In addition to the unspecific side reactions observed with several beads, the analyses run with single beads did not yield interpretable spectra for any of the compounds. Variation of the irradiation times and extraction solvents did not improve the results. Therefore, a different approach was investigated.

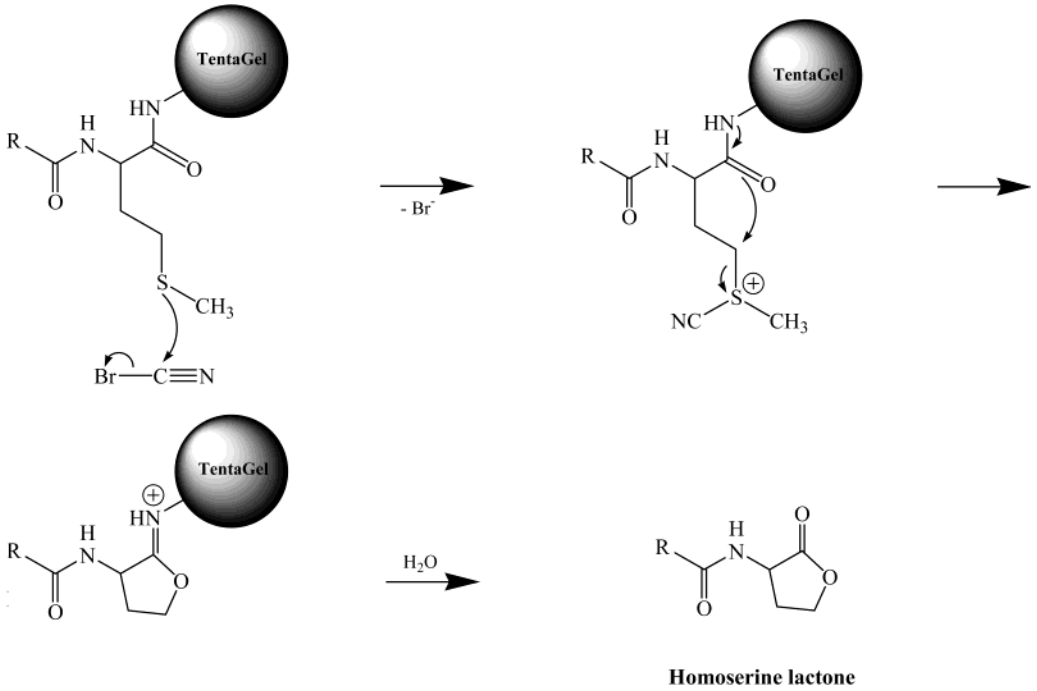
Analysis of Encoding Peptides (Methionine-Linked). In an alternative approach, four encoding peptides with *chemically* cleavable linkers were synthesized on separate topologically segregated TentaGel beads. The chemical structures of the peptides **5**−**8** are shown in Scheme 3. To simplify the peptide synthesis, the same peptides were synthesized on both the interior and exterior of the beads. The coding peptides present on the interior of the beads were linked to the polymer support by a methionine and were cleaved by exposure to CNBr, as described in the Experimental Section.[39] The peptides on the exterior were not methionine-linked and, therefore, could not be cleaved off the beads.

Treatment of the polymer with CNBr in formic acid at room temperature cleaved the peptide−polymer bond. The mechanism for the reaction is depicted by Scheme 4. Cyanation of the sulfide promotes cyclization, which is followed by hydrolysis to yield the homoserine lactone.

Several beads of each batch were treated with CNBr, and the released peptides were analyzed by MALDI-MS. Figure 2 shows the MALDI-FTMS spectrum obtained from compound **6**. The sodiated species at *m*/*z* 782.381 was consistent with the expected homoserine lactone and was the most abundant ion in the spectrum. The signal at *m*/*z* 957 was an impurity. All other peptides yielded spectra of similar quality. The experimental *m*/*z* values were within 2−10 ppm of the theoretical masses.

Subsequently, we tried to produce the protonated species by mixing the peptide with an ammonium exchange resin during MALDI sample preparation. The resulting spectrum (not shown) did indeed yield the $[M + H]^+$ quasimolecular ion; however, extensive fragmentation of the peptide was

Scheme 4. CNBr-promoted Cleavage of Coding Peptides Linked by Methionine



observed, and the protonated peptide was not the most abundant ion in the spectrum. Electrospray ionization has been reported to yield the $[M + H]^+$ species of peptides.[40,41] MALDI yielded a protonated molecule in our hands, but in lower abundance than the sodiated species.

Therefore, we used the $[M + Na]^+$ species of the peptides exclusively to investigate their fragmentation behavior during CID. A concise nomenclature for the assignment of fragment ions from peptides during CID has been published and was used in this paper.[42−46] Figure 3 shows the MALDI-CID spectra obtained from compounds **5**−**8**, respectively. The quasimolecular ions at *m*/*z* 787.328, 782.381, 821.319, and 816.352, respectively, were isolated and fragmented as described in the experimental part. Accurate masses of the observed fragment ions are compiled in Table 1.

A similarity in all spectra was the presence of a fragment ion at *m*/*z* 436. Apparently, this ion was generated from all four peptides regardless of the chemical modifications in the side chains. This was consistent with a Y_2 fragment (Figure 4). Also present in all four CID spectra were the fragment ions with *m*/*z* 446 and 252. The structural identity of the fragment ion at *m*/*z* 446 could not be determined. The fragment ion at *m*/*z* 252, common to all four peptides, resulted from a Y_2 fragmentation combined with the loss of the naphthyloxymethylcarbonyl unit from the third side chain (Lys). For compound **5**, the fragment ions at *m*/*z* 674 and 658 corresponded to a Y_3 and a X_1 cleavage, respectively. Cleavage of the naphthyloxymethylcarbonyl unit in the third side chain yielded an ion at *m*/*z* 603. The signals at *m*/*z* 357, 194, and 752 could not be assigned.

On the basis of the differences in chemical structures (e.g., compounds **5** and **6**), characteristic mass shifts were observed in the CID spectrum of compound **6**. Figure 4b clearly shows that for many fragment ions (Y_3 *m*/*z* 669, X_1 *m*/*z* 653and

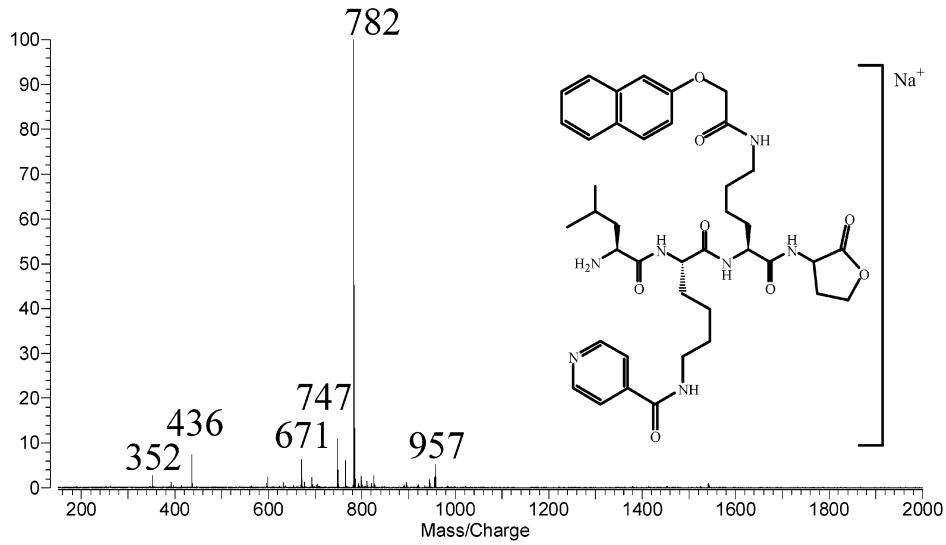


Figure 2. MALDI mass spectrum of compound **6** $[M + Na]^+$. Shown is the spectrum obtained from material cleaved off several beads.

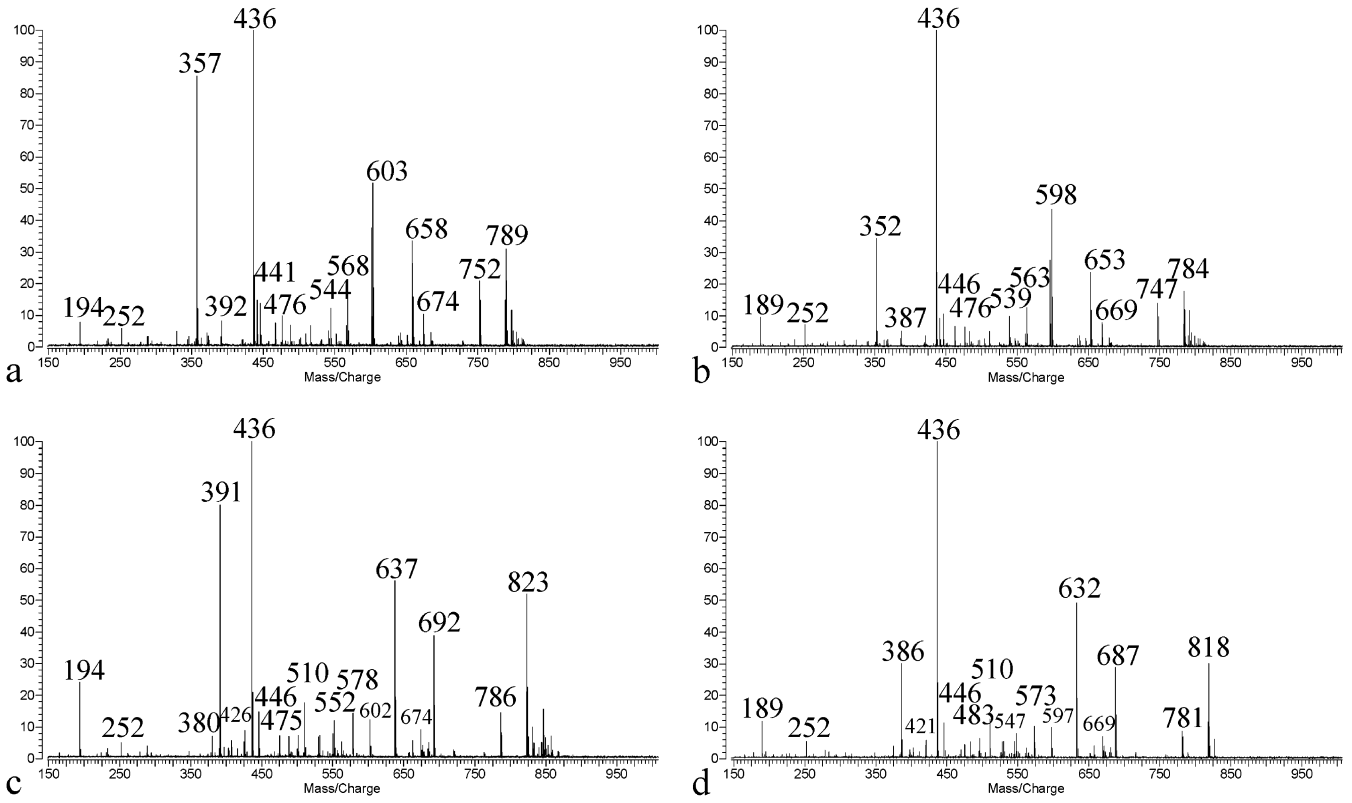


Figure 3. SORI-CID spectra obtained from the isolated quasimolecular ions $[M + Na]^+$ of (a) compound **5**, (b) compound **6**, (c) compound **7**, and (d) compound **8**. Shown are the spectra obtained from material cleaved off several beads.

Table 1. Fragment Ions and Their Respective *m*/*z* Values From the SORI-CID Spectra of Compounds **5**−**8**[a,b]

compound			
5 (787.328)	**6** (782.381)	**7** (821.319)	**8** (816.352)
752.333	747.370	786.308	781.342
674.282 (Y_3)[e]	669.319 (Y_3)[f]	674.273 (Y_3)[e]	669.309 (Y_3)[f]
658.322 (X_1)	653.360 (X_1)	692.299 (X_1)	687.335 (X_1)
603.309	598.347	637.288	632.324
601.257	596.295		
568.271	563.309	602.251	597.287
544.233	539.271	578.213	573.251
		552.197	547.234
		532.228	527.265
516.273	511.312	550.253	545.291
		510.239[e]	510.237[e]
488.169	483.207		
476.259[e]	476.257[e]		
		475.202[e]	475.201[e]
467.219	462.257	501.200	496.238
446.178[g]	446.178[g]	446.174[g]	446.173[g]
441.221[e]	441.220[e]		
436.192 (Y_2)[g]	436.192 (Y_2)[g]	436.189 (Y_2)[g]	436.188 (Y_2)[g]
392.169	387.208	426.150	421.191
357.130	352.169	391.113	386.151
346.162	341.200	380.145	375.183
329.110			
252.136 (Y_2)[g]	252.136 (Y_2)[g]	252.134 (Y_2)[g]	252.134 (Y_2)[g]
194.066[e]	189.105[f]	194.065[e]	189.103[f]

[a] All fragment ions were obtained from the $[M + Na]^+$ species. The *m*/*z* value of the quasimolecular ion is given in parentheses behind the compound number. The assignment of the fragment ions is given in parentheses behind the *m*/*z* values. [b] Similar superscripts indicate mutual differences and similarities of the four CID spectra recorded.

352), a mass shift of 5 u resulted. This reflected the difference in mass between the pyridine (78 u) and the thiophene (82 u) substituents in the side chains. The fragment ions at *m*/*z* 189 and 747 could not be assigned.

Similarly, coding peptides **7** and **8** were analyzed (Figure 3c and d). The CID spectrum for **7** yielded ions at *m*/*z* 692 (X_1), 637 (loss of naphthyloxy methyl carbonyl), and 391. Analogously, compound **8** resulted in fragment ions at *m*/*z* 687, 632, and 386, respectively, reflecting the difference in substitution on the side chains.

It was interesting that certain diagnostic ions were present only in certain spectra, but not in others. On the basis of the characteristic fragment ions observed, all four peptides could be unambiguously assigned to their chemical structures. The establishment of a CID data bank for encoding peptides could be useful in developing general search algorithms. Acquired CID spectra could then be searched against the data bank with *m*/*z* values of fingerprint fragment ions and with relative abundances of these ions for rapid identification of the coding peptide.

To increase the throughput of the chemical cleavage reaction, we performed the release of the peptides from the beads at elevated temperature (47 °C), which reduced the reaction time from overnight at room temperature to 2 h. However, upon examination of the resulting MALDI spectra (Figure 5), a side product was observed with a mass shift of 28 u higher than the homoserine lactone. The elevated temperature apparently promoted formylation of the terminal free amino group, as shown in Scheme 5. Interestingly, there was a difference in the extent of formylation after identical reaction times for compounds **1**−**4**. Whereas compounds **1** and **3** showed little to no modification, the formylation was significantly more pronounced in compounds **2** and **4**. We

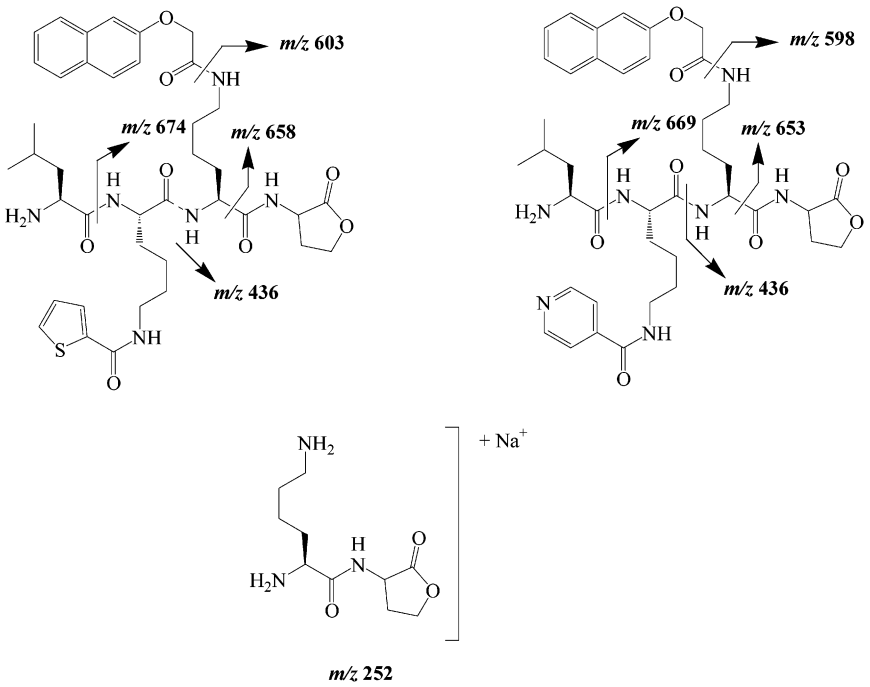


Figure 4. Fragment ions observed in the CID spectra of compounds **5** and **6**. The fragment ions were obtained from the $[M + Na]^+$ species.

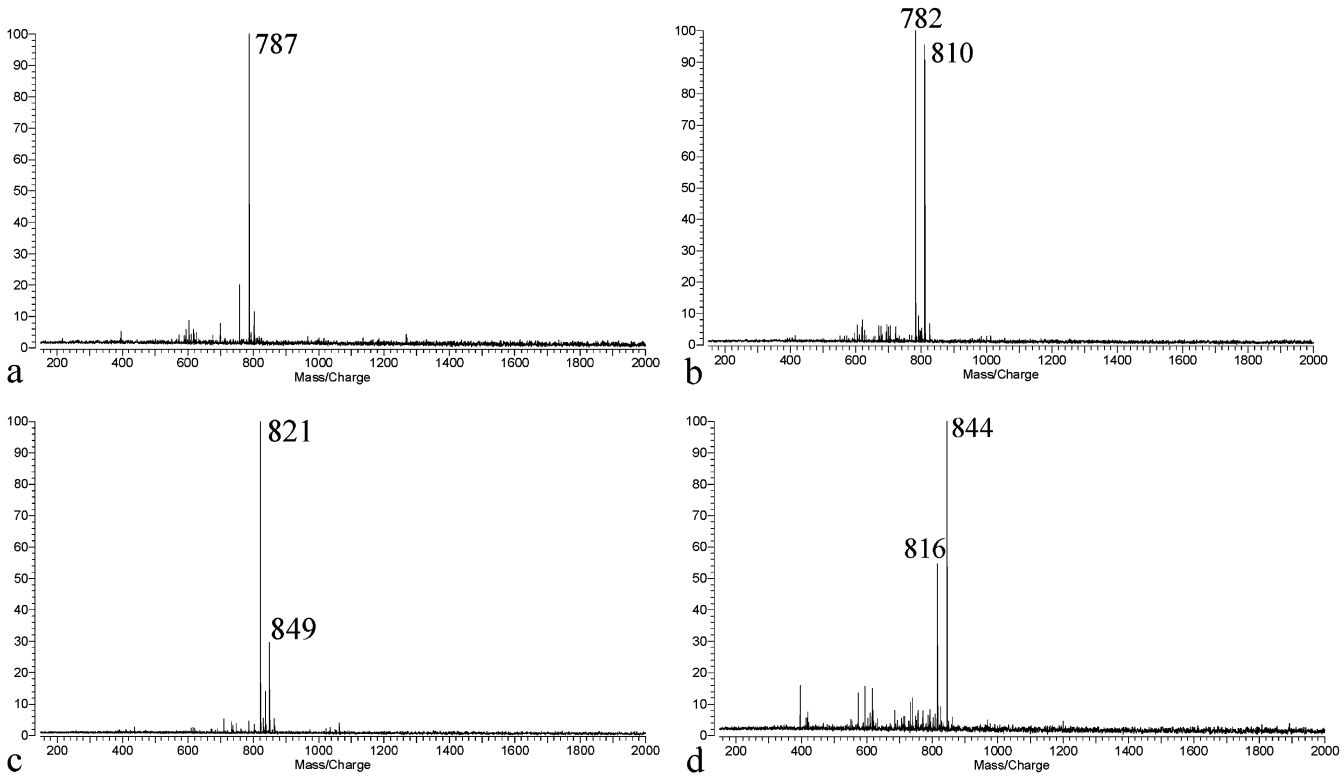


Figure 5. Formylation of the terminal amino group introduced a *m/z* shift of 28 u. This modification was observed only when the chemical cleavage with CNBr was performed at elevated temperature (47 °C). Compound **5** showed no formylation when compared to compounds **6**, **7**, and **8** under identical reaction conditions. Shown are the spectra of 10-bead samples.

are currently investigating strategies to avoid this side reaction and to enhance throughput at elevated temperatures and shorter reaction times.

With optimized reaction conditions for release of peptides from several beads, we turned our attention to the single-bead scale. Single beads were manipulated as described in the experimental part. All four batches were successfully analyzed. Figure 6a shows the representative MALDI-MS spectrum of compound **1**. The amount of material analyzed corresponded to ~0.5 bead with full substitution or to one bead with 50% substitution. This resembled the scenario of topologically segregated bifunctional beads with the encoding

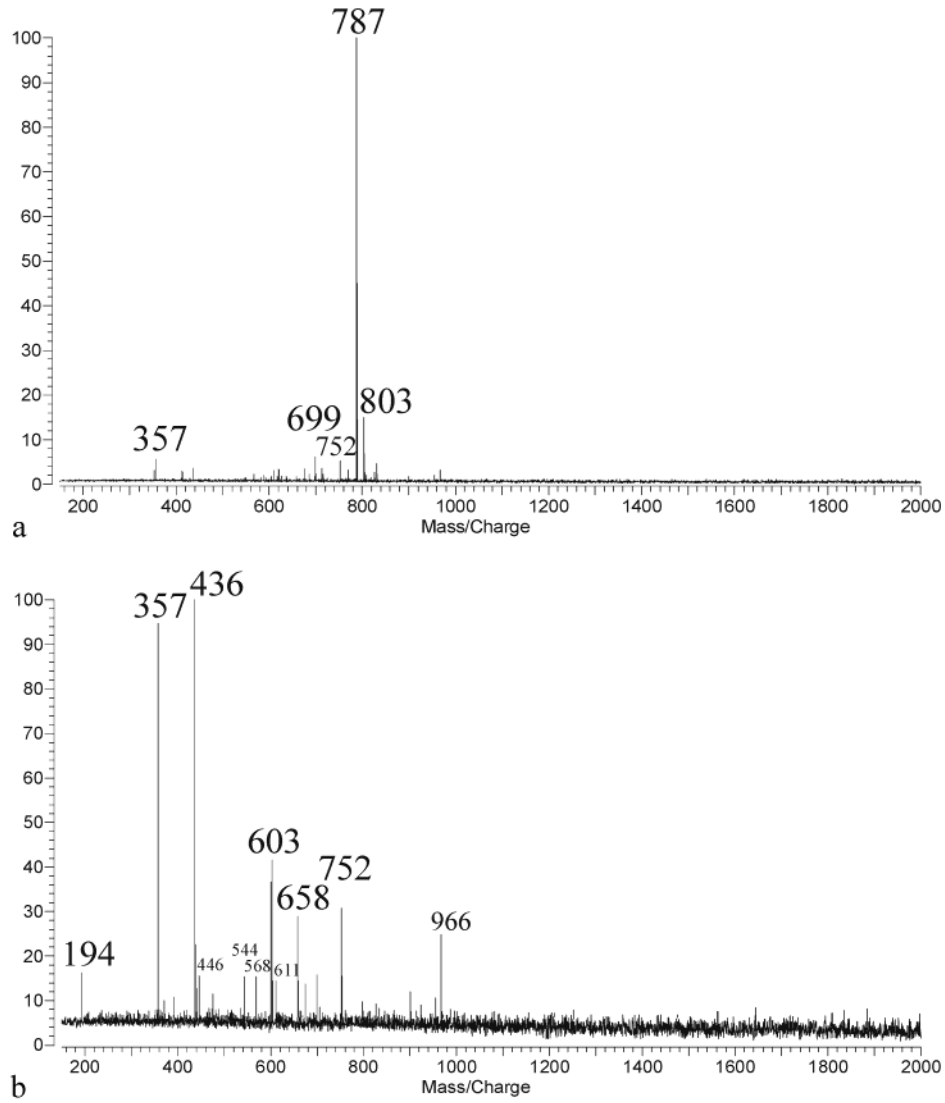
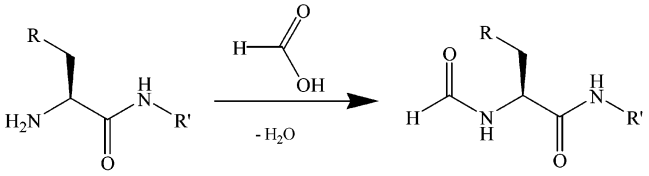


Figure 6. (a) MALDI-MS spectrum of compound **5** from a single-bead analysis $[M + Na]^+$ and (b) SORI-CID spectrum of compound **5** from a single-bead analysis $[M + Na]^+$.

Scheme 5. Formylation of the Terminal Amino Group of Coding Peptides at Elevated Temperatures during Cleavage



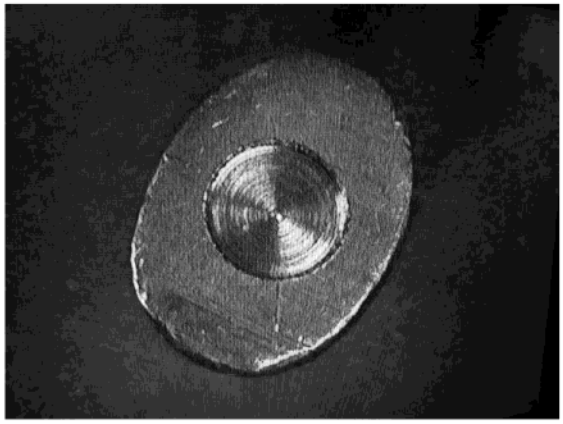
Δm = 28 u

peptide present only on the inside of the bead (~50% substitution). The sodiated molecule at *m*/*z* 787 yielded an abundant signal. The quasimolecular ion was isolated and subjected to CID. The resulting spectrum (Figure 6b) compared well with the CID spectrum obtained with material from several beads (Figure 3a).

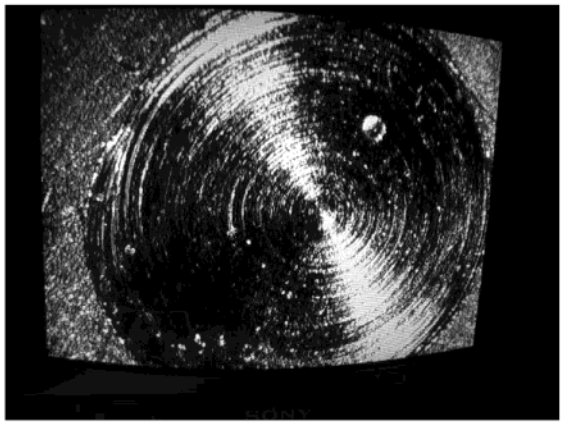
In an attempt to develop a high-throughput system for structural elucidation of the encoding peptides, we modified the surface of the MALDI sample probe to perform on-probe cleavages. A small reservoir (2-mm diameter, 100-μm depth) was drilled into the sample probe (Figure 7a). Carrying out the chemical cleavage, the extraction, and the MALDI analysis on the same probe surface would significantly reduce the time needed for sample preparation and would essentially eliminate sample handling. The release of the peptide from a single bead was then carried out according to the procedure used before. It was clearly demonstrated that the on-probe single-bead analysis was successful (spectra not shown).

The most significant problem encountered during the on-probe digest was the small volume of CNBr solution and the inherent rapid evaporation. Therefore, the release of the peptide was not quantitative, and the resulting abundance of the quasimolecular ion at *m*/*z* 782 was insufficient for CID experiments. However, an appropriate design for a MALDI probe cover is currently being developed to minimize solvent evaporation overnight. In this paper, we replenished the reagent mixture three times over the course of 4 h and once after the release of the peptide had progressed overnight at room temperature.

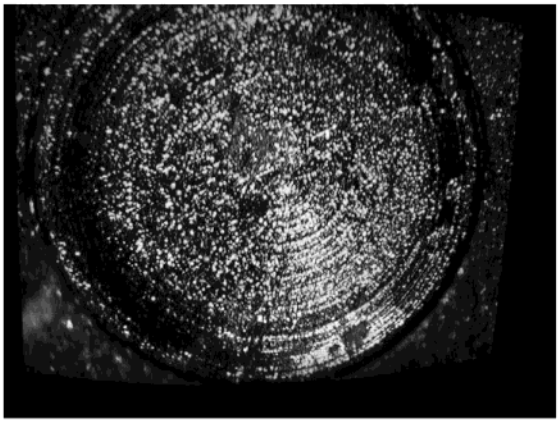
Analysis of Unknown Encoding Peptides (Methionine-Linked). To verify and test the developed method for



a



b

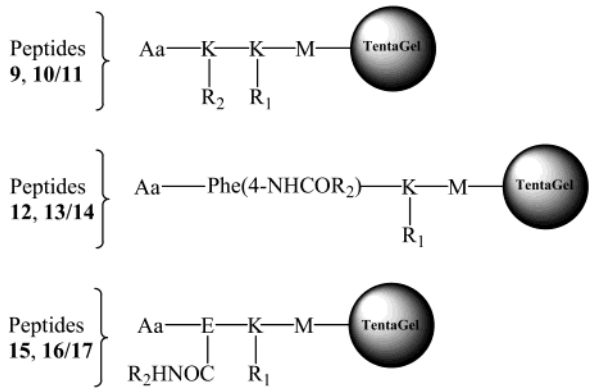


c

Figure 7. (a) Modified surface of a MALDI sample probe. The micro well has a diameter of 2 mm and a depth of 100 μm. Coding peptides on a positive bead from a biological assay can be cleaved, extracted, and prepared for MALDI-MS analysis in the same well. This reduces sample handling. (b) A single bead inside the microwell for analysis. (c) DHB matrix-covered MALDI sample probe (the bead can still be seen at the very right of the well). Some areas are matrix-depleted where the laser hit during analysis.

structural elucidation of encoding peptides from single beads, TentaGel beads encoded with nine different unknown peptides were analyzed on a single-bead scale with 50% substitution. Three batches of beads (**9**, **12**, and **15**) were encoded with a single peptide; the other three batches (**10**/**11**, **13**/**14**, and **16**/**17**) carried two peptides each per single bead. The exact structural identity of the peptides and, therefore, their expected *m*/*z* values were unknown prior to MS analysis. Only the building blocks Aa, R_1, and R_2 used in the synthesis of the library were known (Table 2). Three terminal amino acids (Aa = alanine, phenylalanine, and norvaline) and three compounds used as substituents (R_1, R_2 = 2,5-dimethoxybenzoic acid, 2-pyrazine carboxylic acid, and *N*-phenyl piperazine) were used. Scheme 6 shows the general building block structures of the peptides **9**−**17**.

Scheme 6. Building Block Structures of Encoding Peptides **9**−**17** Used as Mass Spectral "Unknowns"[a]



[a] The residues Aa, R_1, and R_2 were variable.

Scheme 7. Structures of Encoding Peptides **9**−**17** as Synthesized

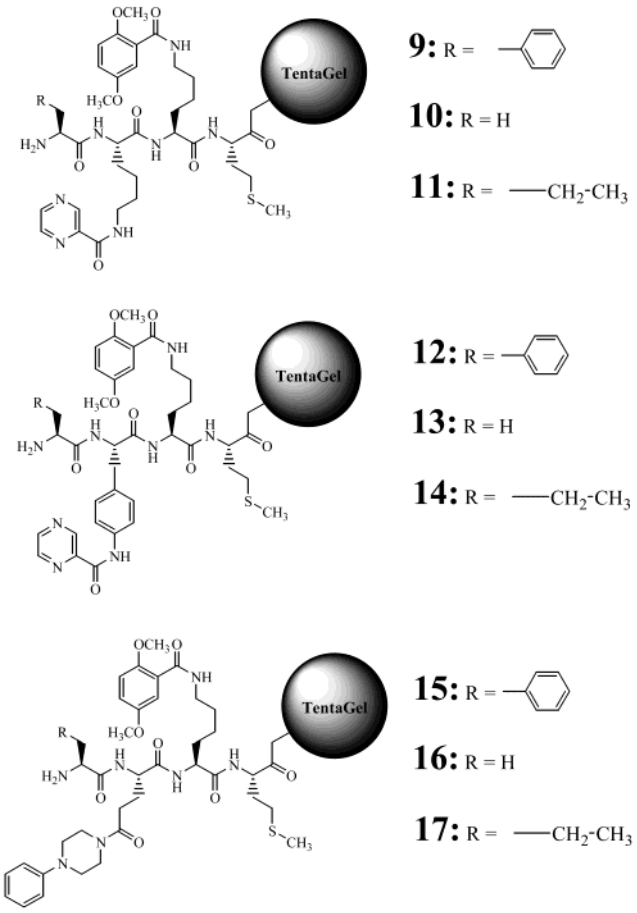


Table 2. Building Blocks Used In the Synthesis of Peptides **9**−**17**

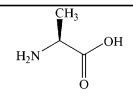
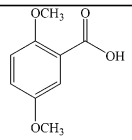
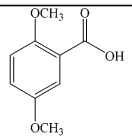
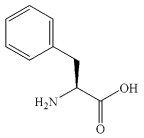
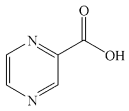
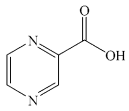
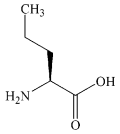
Aa	R_1	R_2
 ($C_3H_7O_2N$, 89.048 u)	 ($C_9H_{10}O_4$, 182.058 u)	 ($C_9H_{10}O_4$, 182.058 u)
 ($C_9H_{11}O_2N$, 165.079 u)	 ($C_5H_4O_2N_2$, 124.027 u)	 ($C_5H_4O_2N_2$, 124.027 u)
 ($C_5H_{11}O_2N$, 117.079 u)		

Table 3. Theoretical and Experimental *m*/*z* Values for $[M + Na]^+$ of Sample Peptides **9**−**17**

entry	*m*/*z* (theor)	*m*/*z* (exptl)	error (ppm)
9	797.360	797.362	2.5
10	721.329	721.322	10
11	749.360	749.354	8.0
12	831.344	831.332	14
13	755.313	755.303	13
14	783.344	783.334	13
15	836.396	836.392	4.8
16	760.365	760.361	5.3
17	788.396	788.390	7.6

All peptides were cleaved off in a single-bead experiment and subjected to mass spectrometric analysis. Figure 8a and b shows two representative MALDI-FTMS spectra of compounds **9** and **12**, respectively. The sodiated molecules were observed with satisfactory abundance. Given the exact masses of the building blocks Aa, R_1, and R_2, all peptides were successfully assigned to specific chemical compositions by calculating the possible permutations of masses. Their experimental *m*/*z* values corresponded to the expected values with excellent accuracy of 2−10 ppm. Compounds **10**/**11** and **13**−**17** gave similar results, and their respective *m*/*z* values are summarized in Table 3.

For *structural* elucidation, the quasimolecular ions were isolated and subjected to CID. Figure 8c and d shows representative CID spectra of compounds **9** and **12**, respectively. Compound **9** yielded fragment ions at *m*/*z* 691 (loss of the pyrazine), 668 (X_1), and 416 (Y_2) as shown in Figure 8e. The signals at *m*/*z* 762 and 661 could not be assigned. The CID spectrum of compound **12** contained ions at *m*/*z* 702 (X_1) and 416 (Y_2). The encoding peptides **10**/**11** and **13**−**17** were analyzed analogously. The fragment ions and their respective *m*/*z* values are summarized in Table 4 for all nine peptides. The exact masses of the fragment ions were sufficient to fully assign the structure of the peptides (Scheme 7).

Because of the small amounts of sample used in the analysis, the CID spectra did not show all the fragment ions observed with more concentrated samples. However, the most abundant fragment ions were detected. The exact masses of the quasimolecular ion and its fragment ions will be used in future studies to fully derive the structures of compounds from combinatorial libraries.

Table 4. Fragment Ions and Their Respective *m*/*z* Values from the SORI-CID Spectra of Compounds **9**−**17**[a,b]

compound								
9	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
797.362	721.355	749.388	831.368	755.303	783.334	836.426	760.361	788.390
762.342	686.309	714.351	796.332	720.291	748.320	801.354	725.341	753.366
691.359	615.325			619.239	677.327	674.275		
668.332	592.300	620.337	702.317	626.268	654.316		631.331	659.359
(X_1)	(X_1)	(X_1)	(X_1)	(X_1)	(X_1)		(X_1)	(X_1)
661.290	585.259	613.297	584.238			656.265	580.247	608.275
							563.220	591.248
						527.209[e]	527.219[e]	527.217[e]
		426.175[e]	426.170[e]				426.169[e]	
416.186	416.186	416.189	416.185	416.185	416.185	416.184	416.184	416.183
(Y_2)[f]	(Y_2)[f]	(Y_2)[f]	(Y_2)[f]	(Y_2)[f]	(Y_2)[f]	(Y_2)[f]	(Y_2)[f]	(Y_2)[f]
					373.131			

[a] All fragment ions were obtained from the $[M + Na]^+$ species. The *m*/*z* value of the quasimolecular ion is given behind the compound number [b] Similar superscripts indicate mutual differences and similarities of the CID spectra.

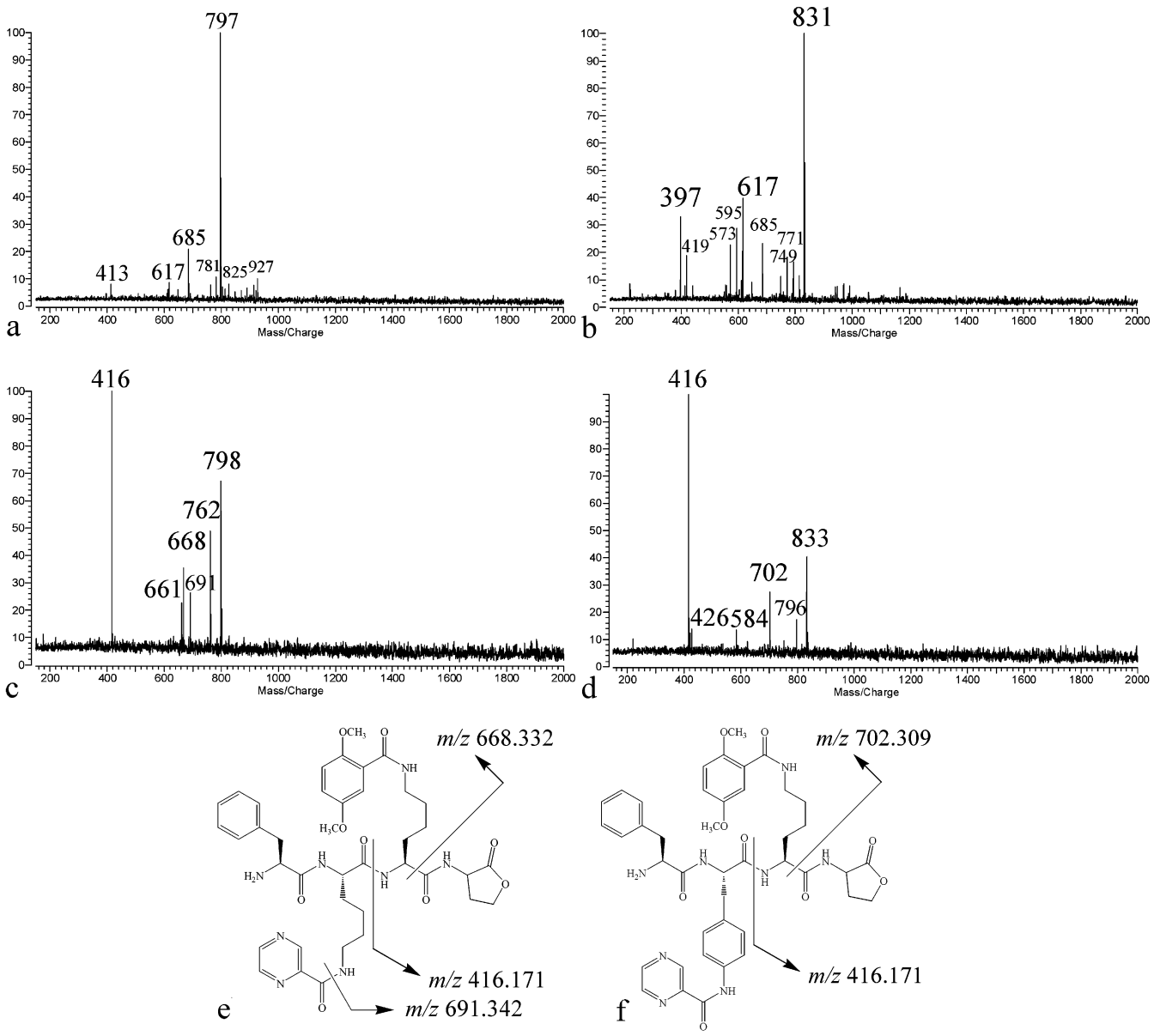


Figure 8. Representative MALDI-FTMS spectra of (a) compound **9** and (b) compound **12** and MALDI-FTMS-CID spectra of (c) compound **9** and (d) compound **12**. All spectra were recorded on a one-bead scale with 50% substitution. The $[M + Na]^+$ quasimolecular ion was chosen for the CID experiments. (e) Observed fragment ions from compound **9**. (f) Observed fragment ions from compound **12**. On the basis of the exact masses of the fragment ions, the correct structure of the peptides was derived.

Conclusion

Encoding peptides can be cleaved off single TentaGel beads and can be fully characterized by ultrahigh-resolution FT-ICR mass spectrometry. Peptides linked to the bead by methionine and cleaved by CNBr showed more consistent results than peptides linked by a UV-cleavable linker. The amounts of peptides obtained from a single bead (with 50% substitution) were sufficient to yield abundant quasimolecular ions in MALDI-TOF and MALDI-FTMS spectra. CID experiments were performed and yielded satisfactory spectra. We have successfully demonstrated that peptides from single beads can be cleaved directly on the MALDI probe to reduce sample preparation times and handling. In future studies, the ultrahigh mass accuracy of FTMS will be used to serve two purposes. First, the exact mass of the quasimolecular ion will be determined and screened against a data bank of combinatorial masses on the basis of the building blocks of the library. FTMS mass measurements with mass resolution of $m/\Delta m > 500\,000$ will be sufficient to reduce the number of chemical compositions from the library to a few possibilities. Second, the CID fragment ions and their exact masses will be used to elucidate the chemical structure of the encoding peptide.

Combinatorial OBOC chemistry on topologically segregated beads in combination with concurrent biological screening of hundreds of thousands of small-molecule beads and ultrahigh-resolution FT mass spectrometry on a single-bead scale will offer a fast high-throughput system for drug discovery.

Acknowledgment. This work was supported by the National Institute of Health., NIH R33CA-86364. R. L. is supported in part by the University of California Systemwide Biotechnology Research Program, Grant no. 2001-07.

References and Notes

(1) Boyle, N. A.; Janda, K. D. *Curr. Opin. Chem. Biol.* **2002**, *6*, 339−346.
(2) Nielsen, J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 297−305.
(3) Marcaurelle, L. A.; Seeberger, P. H. *Curr. Opin. Chem. Biol.* **2002**, *6*, 289−296.
(4) Lebl, M. *J. Comb. Chem.* **1999**, *1*, 3−24.
(5) Drews, J. *Science* **2000**, *287*, 1960−1963.
(6) Dolle, R. E. *Mol. Diversity* **1998**, *3*, 199−233.
(7) Chabala, J. C. *Comb. Chem. Mol. Diversity Drug Discov.* **1998**, 3−15.
(8) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−84.
(9) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Apple, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84−86.
(10) Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 487−493.
(11) Lam, K. S.; Sroka, T.; Chen, M.-L.; Zhao, Y.; Lou, Q.; Wu, J.; Zhao, Z.-G. *Life Sci.* **1998**, *62*, 1577−1583.
(12) Lam, K. S.; Lebl, M.; Krchnak, V. *Chem. Rev.* **1997**, *97*, 411−448.
(13) Edman, P. *Acta Chem. Scand.* **1950**, *4*, 283−293.
(14) Edman, P.; Begg, G. *Eur. J. Biochem.* **1967**, *1*, 80−91.
(15) Czarnik, A. W. *Curr. Opin. Chem. Biol.* **1997**, *1*, 60−66.
(16) Xiao, X.-Y. *Front. Biotechnol. Pharm.* **2000**, *1*, 114−149.
(17) Barnes, C.; Balasubramanian, S. *Curr. Opin. Chem. Biol.* **2000**, *4*, 346−350.
(18) Affleck, R. L. *Curr. Opin. Chem. Biol.* **2001**, *5*, 257−263.
(19) Lebl, M.; Lam, K. S.; Salmon, S. E.; Krchnak, V.; Sepetov, N.; Kocis, P. Topologically Segregated, Encoded Solid Phase Libraries, U.S. Patent, 1998; US5840485.
(20) Vagner, J.; Barany, G.; Lam, K. S.; Krchnak, V.; Sepetov, N. F.; Ostrem, J. A.; Strop, P.; Lebl, M. *Proc. Natl. Acad. Sci.* **1996**, *93*, 8194−8199.
(21) Liu, R.; Marik, J.; Lam, K. S. *J. Am. Chem. Soc.* **2002**, *124*, 7678−7680.
(22) Hinzen, B. *Methods Princ. Med. Chem.* **2000**, *9*, 239−242.
(23) Gauglitz, G. *Curr. Opin. Chem. Biol.* **2000**, *4*, 351−355.
(24) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53−71.
(25) Wu, J.; Gard, E.; Bregar, J.; Green, M. K.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1995**, *117*, 9900−9905.
(26) He, F.; Ramirez, J.; Garcia, B. A.; Lebrilla, C. B. *Int. J. Mass Spectrom.* **1999**, *182*, 261−273.
(27) Penn, S. G.; He, F.; Lebrilla, C. B. *J. Phys. Chem. B* **1998**, *102*, 9119−9126.
(28) Haskins, W. E.; Wang, Z.; Watson, C. J.; Rostand, R. R.; Witowski, S. R.; Powell, D. H.; Kennedy, R. T. *Anal. Chem.* **2001**, *73*, 5005−5014.
(29) Nedelkov, D.; Nelson, R. W. *Anal. Chim. Acta* **2000**, *423*, 1−7.
(30) Chatman, K.; Hollenbeck, T.; Hagey, L.; Vallee, M.; Purdy, R.; Weiss, F.; Siuzdak, G. *Anal. Chem.* **1999**, *71*, 2358−2363.
(31) Cheng, X.; Hochlowski, J. *Anal. Chem.* **2002**, *74*, 2679−2690.
(32) Triolo, A.; Altamura, M.; Cardinali, F.; Sisto, A.; Maggi, C. A. *J. Mass Spectrom.* **2001**, *36*, 1249−1259.
(33) Kassel, D. B. *Chem. Rev.* **2001**, *101*, 255−267.
(34) Schmid, D. G.; Grosche, P.; Bandel, H.; Jung, G. *Biotechnol. Bioeng.* **2001**, *71*, 149−161.
(35) Enjalbal, C.; Martinez, J.; Aubagnac, J.-L. *Mass Spectrom. Rev.* **2000**, *19*, 139−161.
(36) Stults, J. T. *Anal. Chem.* **1997**, *69*, 1815−1819.
(37) Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* **1990**, *35*, 161−214.
(38) Wang, B. H.; Biemann, K. *Anal. Chem.* **1994**, *66*, 1918−1924.
(39) Crimmins, D. L.; Mische, S. M.; Denslow, N. D. *Curr. Protoc. Protein Sci.* **2000**, 11.5.1−11.5.13.
(40) Griffiths, W. J.; Jonsson, A. P.; Liu, S.; Rai, D. K.; Wang, Y. *Biochem. J.* **2001**, *355*, 545−561.
(41) Jonsson, A. P. *Cell. Mol. Life Sci.* **2001**, *58*, 868−884.
(42) Roepstorff, P. *Biomed. Mass Spectrom.* **1984**, *11*, 601.
(43) Roepstorff, P.; Hoejrup, P.; Moeller, J. *Biomed. Mass Spectrom.* **1985**, *12*, 181−189.
(44) Craig, A. G.; Koerber, S. C.; Porter, J.; Hoeger, C.; Rivier, J. E. *Biol. Mass Spectrom.* **1993**, *22*, 31−44.
(45) Biemann, K. *Methods Enzymol.* **1990**, *193*, 886−887.
(46) Tuinman, A. A.; Pettit, G. R. *Int. J. Pept. Protein Res.* **1990**, *36*, 331−334.

CC020083A