METHODS IN MOLECULAR BIOLOGY™

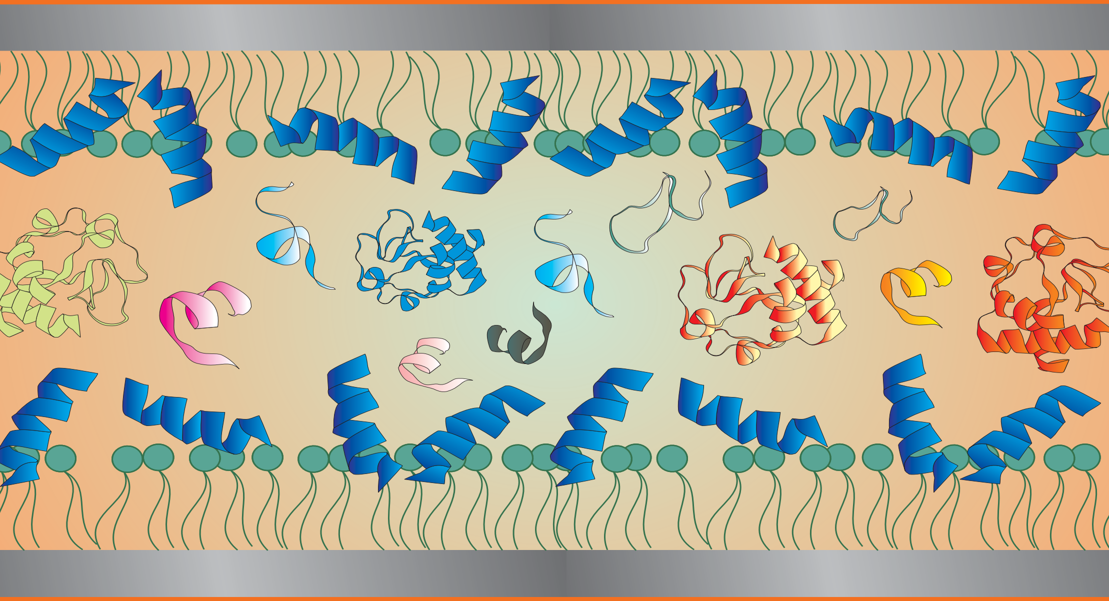
Volume 251

HPLC of Peptides and Proteins

Methods and Protocols

Edited by

Marie-Isabel Aguilar



HUMANA PRESS

26

Automated vs Manual Profiling of Peptide Libraries by Mass Spectrometry

Jean-Louis Aubagnac, Christine Enjalbal, Jean Martinez,
Pierre Sanchez, and Gilles Subra

1. Introduction

Combinatorial chemistry ***(1–7)*** has evolved from the synthesis of very large mixtures to the preparation of collections of isolated compounds by parallel syntheses. Peptides constitute ideal materials because their preparation is completely automated using solid-phase methodology allowing high throughput syntheses.

Library profiling requires both compound identification and quantification. In addition, the analysis must be as fast as possible to keep pace with the synthesis and screening. Such requirements impart considerable constraints on the analytical method. Specificity, sensitivity, and rapidity are prerequisite for efficient control of large collections of molecules.

Among the various spectroscopic methods reported in the literature on that topic ***(8–10)***, simultaneous ultraviolet (UV) and mass spectrometric detections were found the most convenient. Indeed, the two techniques are complementary. UV detection is sensitive, but not specific. Furthermore, liquid chromatography (LC)/UV experiment allows compound quantification if a suitable wavelength has been chosen. On the contrary, mass spectrometry (MS) exhibits both specificity and sensitivity, but quantification is not directly achievable because of possible discrimination during compound ionization. The recourse to hyphenated LC–MS technique allows gathering all the aforementioned data (UV chromatogram and mass spectrum) in a single experiment, thus reducing the analysis time.

From: *Methods in Molecular Biology, vol. 251, HPLC of Peptides and Proteins: Methods and Protocols*
Edited by: M.-I. Aguilar © Humana Press Inc., Totowa, NJ

Data management is of crucial importance. Specific software has been designed to allow automatic data acquisition and interpretation. Samples can be run continuously through the use of autosamplers able to handle vials or plates. Optimization of such systems can be effected by diminishing injection-to-injection cycle times and increasing the sensitivity. LC columns dedicated to combinatorial library analysis are commercially available. Their reduced dimensions (50×2.1 mm, 3.5 µm) enable the flow rate to be decreased from 1 to 0.25 mL/min while maintaining good chromatographic separation. Reduced analysis time and better sensitivity in electrospray MS are thus achieved.

Combinatorial peptide libraries are routinely analyzed by such fast LC–MS techniques, the molecule of interest and the eventual side-products being characterized but also quantified. Above a defined purity threshold, the synthesis is considered successful whereas all experiments which failed to contain enough compound are discarded. Such pass/fail criterion is usually visualized by green and red spots in automated result reporting using commercially available softwares. The selected samples can then be screened as such or further purified prior to bioassays.

We found that such automated data interpretation must be considered with caution. Provided that all ions related to the compound including adducts and multiply charged species are monitored, a negative response should indicate an unsuccessful synthesis. Manual inspection of such presumably failed syntheses showed that the desired peptides were present but in mixtures with various byproducts. The systematic recourse to double coupling avoids contamination by deleted chains. However, the final deprotection step, which is conducted rapidly to avoid oxidation of methionine residues, is not sufficient to cleave simultaneously all side-chain protecting groups. Whereas the presence of truncated peptides irreversibly affects the product purity, partially deprotected chains can be easily converted into the desired sequences by carrying out a second deprotection reaction. This chapter outlines a method for analysis of a 96-member library by LC–MS techniques.

2. Materials

2.1. Chemicals

1. High-pressure liquid chromatography (HPLC) grade trifluoroacetic acid (TFA) is available from Merck (Schuchardt, Germany).
2. HPLC grade acetonitrile is available from Panreac Chimica (Barcelona, Spain) or from Riedel-de-Haën (Seelze, Germany).
3. Water was glass-distilled in-house and further purified using a Milli-Q water purification system (Millipore, Bedford, MA).

4. Fmoc-protected amino acids and polystyrene Rink amide resin are available from Advanced Chemtech (Louisville, KY).
5. HBTU is available from SENN Chemicals (Dielsdorf, Switzerland).

2.2. Solid-Phase Peptide Synthesis

1. Solid-phase synthesis is performed on an ACT 496 Ω multiple organic synthesizer in Fmoc strategy using PS Rink amide resin (0.8 mmol/g).
2. Coupling solution (120 mmol/L Fmoc-amino acid, 120 mmol/L HBTU, 240 mmol/L DIEA) is prepared in dimethyl formamide (DMF).
3. Two successive coupling steps of 2 h are used at each stage of the synthesis to avoid deletion products.
4. Fmoc removal is performed in 20 min with a 80/20 DMF/piperidine solution (v/v).
5. A TFA/anisole/thioanisole/ethanedithiol/water (90/2.5/2.5/2.5/2.5, v/v) solution is used for the final resin cleavage and side-chain deprotection step.
6. After cleavage, peptides are precipitated several times in diethyl ether and lyophilized.

2.3. LC/ESI-MS Instrumentation

1. Identification of the peptides is performed by high performance liquid chromatography coupled to an electrospray mass spectrometer (LC/ESI-MS).
2. The mass spectrometer (Platform II from Waters-Micromass, Milford, MA) fitted with a quadrupole is calibrated in the positive ion mode using a mixture of *NaI* and *CsI*.
3. Voltages are set at +3.5 kV for the capillary and +0.5 kV for the skimmer lens.
4. The source is heated at 120°C. Nitrogen is used as nebulizing and drying gas at 15 and 250 L/h, respectively.
5. Mass spectrometric data are acquired in the positive mode according to one of the following methods (*see* **Note 2**):
   Method A: scan mode from *m/z* 200 to *m/z* 1200 in 3 s.
   Method B: scan mode from *m/z* 400 to *m/z* 1900 in 4 s.
6. Three to five scans are summed to produce the final spectrum at the selected retention time.
7. Voltage of the sampling cone is adjusted to 30 V.
8. Data are processed by Micromass MassLynx system.
9. An Alliance 2690 (Waters, Milford, MA) LC system equipped with an autosampler is used to deliver the mobile phase, continuously degassed, at a flow rate of 250 µL/min.
10. The eluent is split after the LC column to reach only 35 µL/min of mobile phase infused into the mass spectrometer source and 215 µL/min into the UV cell.
11. UV detection is set up at 214 nm.
12. The Symmetry Shield RP18 column (50 × 2.1 mm, 3.5 µm) is available from Waters-Micromass.

13. The best chromatographic separation is achieved at a flow rate of 250 µL/min using a gradient from 0 to 50% acetonitrile in water containing 0.1% TFA in 10 min followed by a second rapid gradient from 50 to 80% acetonitrile in 0.5 min.
14. Isocratic elution is carried out for 1.5 min at 80% acetonitrile to flush lipophilic compounds.
15. The initial conditions are then reached in 0.5 min and the column is equilibrated for 4.5 min before injection of the next sample.
16. The autosampler requires 1 min to select the vial and load the injection loop. So the overall run cycle per sample is 18 min (*see* **Note 3**).
17. All samples are dissolved in acetonitrile/water (50/50, v/v) at a concentration of 30 mmol/L.
18. Each solution is placed in one autosampler vial fitted with a 200-µL restrictor.
19. In all experiments, 1 µL of sample is injected at 25°C.

3. Methods

1. Place the vials in the autosampler.
2. Equilibrate the LC column under initial conditions.
3. Load the elution conditions file (LC file).
4. Choose the MS scan method for each sample according to the molecular weight of the peptide: Method A for molecules possessing a molecular weight inferior to 1000 atomic mass unit (amu) and Method B for molecules possessing a molecular weight superior to 1000 amu.
5. Fill the sample list (file name, LC file that is identical for all samples, MS file that varies according to the chosen scan mode).
6. Start the run overnight (electrospray probe into the source, gas supply on, whole sample list selected and automatic shutdown activated) because analysis of the whole library roughly lasts 29 h.
7. Effect manual inspection of all samples in the morning (quantification from UV integration and identification from regenerated spectra).

In this example, a 96-member library designed to contain fully deprotected peptides of molecular weight ranging from 700 to 3000 was analyzed by LC–MS. Automated data interpretation was effected by targeting the protonated and sodiated molecular ions as well as the doubly and triply charged molecule. Approximately one-quarter of the samples were correct, i.e., the main UV signal that was integrated corresponded to more than 80% (*see* **Fig. 1B**) was the expected peptide. The mass spectrum regenerated at 8.87 min (*see* **Fig. 1C**) exhibited two intense signal at *m/z* 970 and 647 attributed to the doubly and triply protonated peptide, respectively. The protonated molecule was scarcely detected and always with a very low abundance. The other samples exhibited multiple signals in UV and TIC chromatograms as displayed in

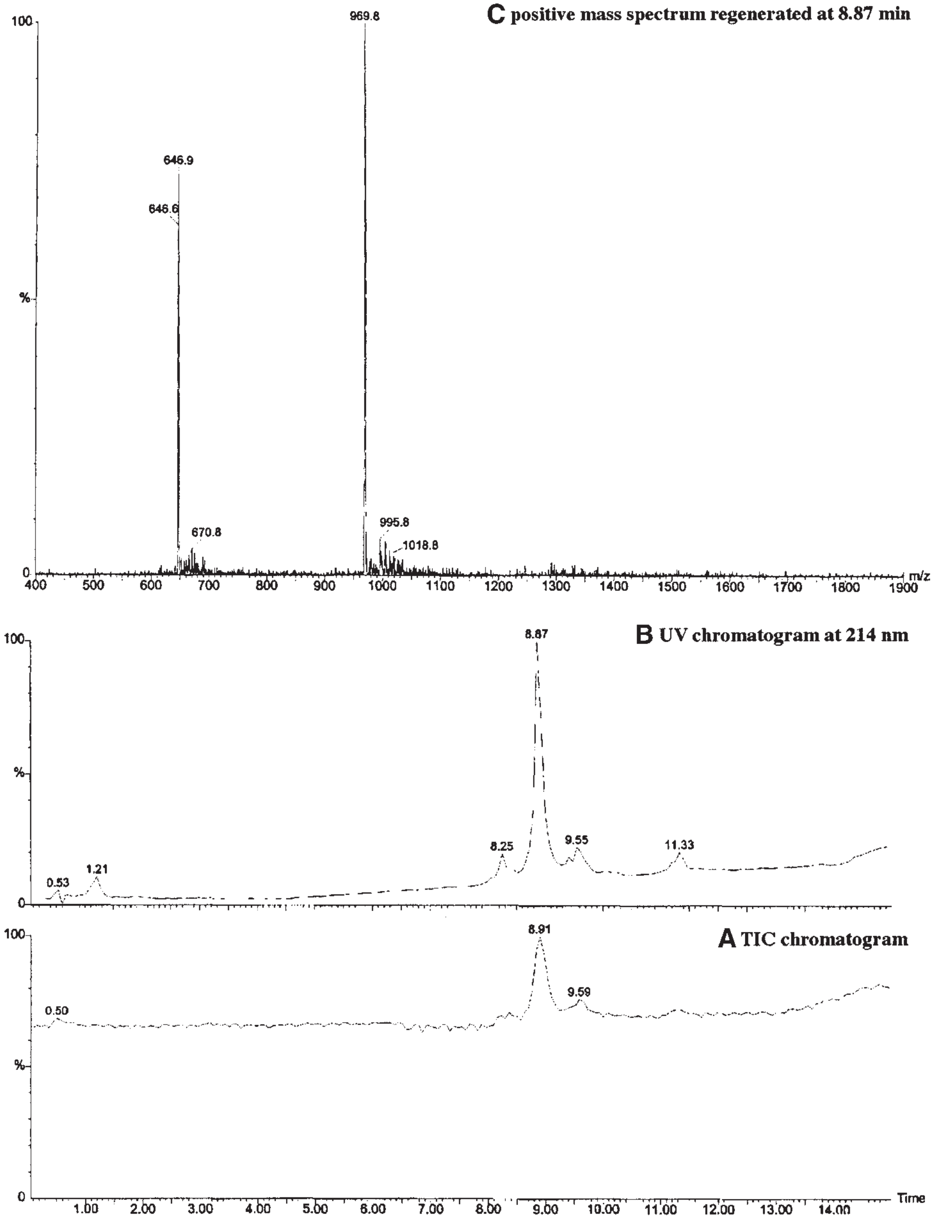


Fig. 1. **(A)** TIC chromatogram in the positive mode. **(B)** UV chromatogram at 214 nm. **(C)** Positive mass spectrum regenerated at 8.87 min showing the expected peptide: $(M+2H)^{2+} = 969.8$ and $(M+3H)^{3+} = 646.9$.

Fig. 2. The main peak at 8.98 min corresponded to the expected peptide [*see* **Fig. 3A**: $(M+H)^+ = 1281$ and $(M+2H)^{2+} = 641$] but three major byproducts were detected at 9.66, 11.85, and 12.33 min. The first eluted byproduct at 9.66 min [*see* **Fig. 3B**: $(M+2H)^{2+} = 669$] possessed one t-Bu protected side chain because the mass increment on the doubly charged species equaled 28 Th. Accordingly, the compound at 11.85 min corresponded to a sequence containing a side-chain-protected arginine residue [*see* **Fig. 3C**: $(M+2H)^{2+} = 774$, mass increment of 133 Th on the doubly charged species]. The more lipophilic molecule eluted at 12.33 min contained a t-Bu and a Pmc protecting group [*see* **Fig. 3D**: $(M+2H)^{2+} = 802$]. Most of the unsuccessful syntheses were a result of partial removal of the Pmc protection of the guanidinium moiety, especially in the case of arginine rich sequences (*see* **Note 1**). In some extreme cases, the desired fully deprotected peptide was hardly detected and more lipophilic protected structures were eluted at very high concentrations of acetonitrile. In several samples, partially deprotected sequences were detected at different retention times implying that regioisomers were present. Thus, the whole library was submitted again to a deprotection step and analyzed by LC–MS a second time.

Of the samples that contained partially deprotected chains after the first synthesis, only 25% of them were pure enough to undergo the screening process. Resubmitting the collection of samples to a second deprotection procedure raised this figure to 99%. So, the manual inspection of the data saved time by preventing the time consumed by resynthesis of the library.

4. Notes

1. Since the Pmc protection of the arginine side chain was not completely removed after two TFA treatments in some peptides, the Pbf moiety was preferred to mask the guanidinium functionality due to its faster deprotection rate ***(11)***.
2. Two MS scan modes were required due to the heterogeneity of the studied library in terms of peptide length. The low-molecular-weight sequences (<1000 amu) exhibited mostly singly protonated molecular ion, whereas longer chains (>1000 amu) produced to multiple ions at different charge state. A wider mass range was thus necessary for the former samples to detect all ions. The upper limit of 1900 amu was determined according to preliminary studies which showed that peptides possess sufficient basic sites to promote multiple protonations thus decreasing the recorded *m/z* values.
3. More efficient library profiling was achieved by using a shorter LC column. The XTerra RP18 column (30 × 2.1 mm, 3.5 µm) was purchased from Waters and showed comparable chromatographic separation as the Symmetry Shield stationary phase. The flow rate was kept at 250 µL/min but the elution was performed faster using a gradient from 0 to 50% acetonitrile in water containing 0.1% TFA in 6.5 min (instead of 10 min) followed by a second rapid gradient from 50 to 80%

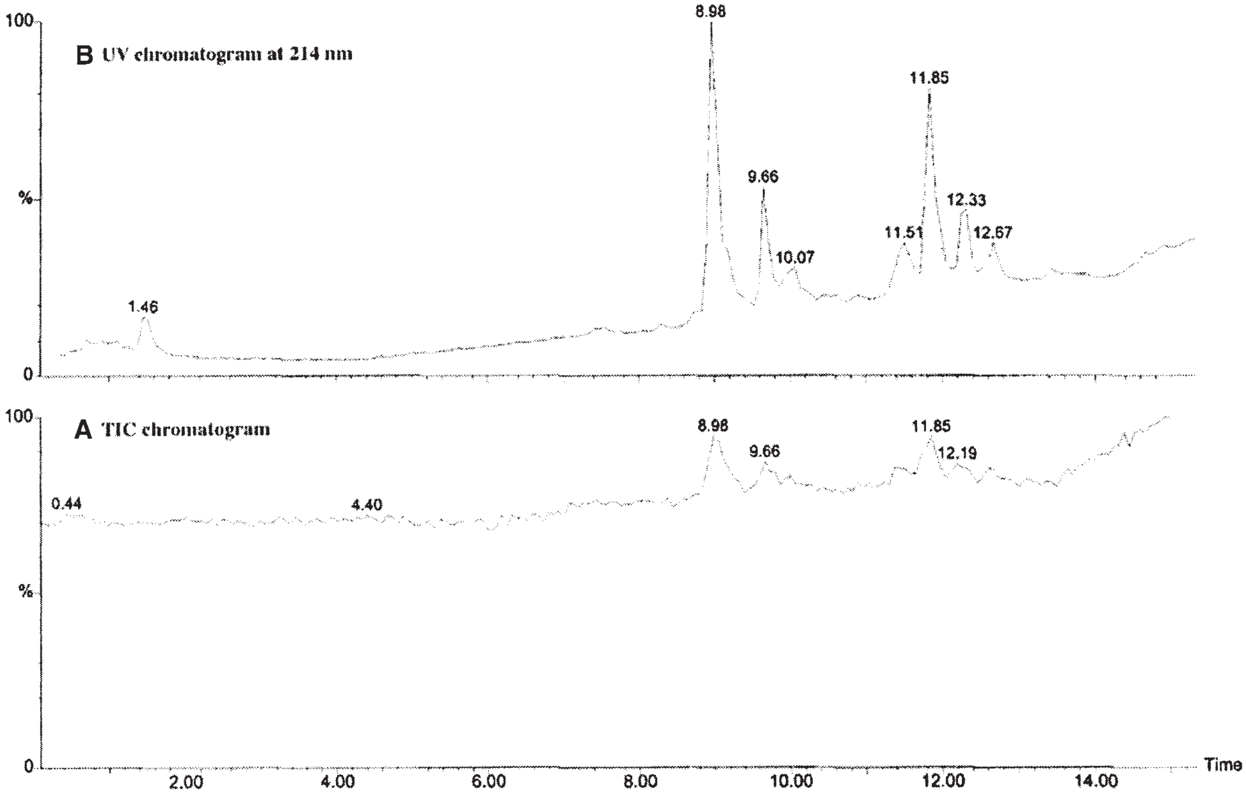
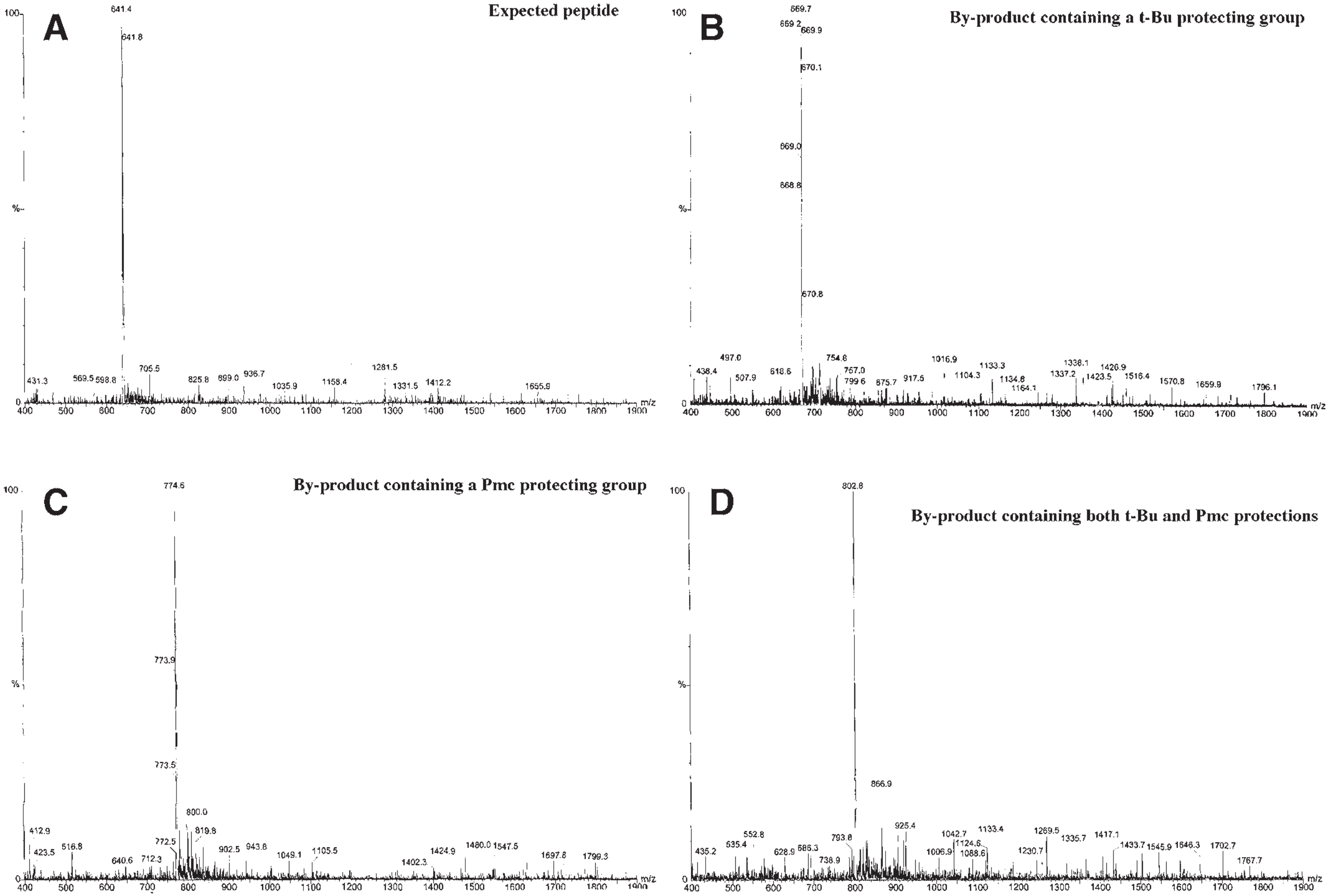


Fig. 2. **(A)** TIC chromatogram in the positive mode. **(B)** UV chromatogram at 214 nm.


Expected peptide

By-product containing a t-Bu protecting group

By-product containing a Pmc protecting group

By-product containing both t-Bu and Pmc protections acetonitrile in 0.5 min. Isocratic elution was carried out for 1.0 min at 80% acetonitrile to flush lipophilic compounds. The initial conditions were then reached in 0.5 min and the column was equilibrated for 4.0 min before injection of the next sample (1 min delay to load the injection loop). The injection-to-injection cycle time was thus reduced from 18 to 13.5 min. The same 96-member's library was analyzed in less than 22 h.

References

1. Czarnik, A. W. and Dewitt, S. H. (1997) *A Practical Guide to Combinatorial Chemistry*. American Chemical Society, Washington, DC.
2. Wilson, S. R. and Czarnik, A. W. (1997) *Combinatorial Chemistry—Synthesis and Application*. Wiley, New York.
3. Bunin, B. A. (1998) *The Combinatorial Index*. Academic, London.
4. Terrett, N. K. (1998) *Combinatorial Chemistry*. Oxford University Press, Oxford, UK.
5. Gordon, E. M. and Kervin, J. F. (1998) *Combinatorial Chemistry and Molecular Diversity in Drug Discovery*. Wiley, New York.
6. Obrecht, D. and Villalgordo, J. M. (1998) *Solid-Supported Combinatorial and Parallel Synthesis of Small Molecular Weight Compound Libraries*. Tetrahedron Organic Chemistry Series, Vol. 17, Pergamon, Elsevier, Oxford, UK.
7. Jung, G. (1999) *Combinatorial Chemistry—Synthesis, Analysis, Screening*. Wiley-VCH, Weiheim, RFA.
8. Shapiro, M. J. and Wareing, J. R. (1998) NMR methods in combinatorial chemistry. *Curr. Opin. Chem. Biol.* **2,** 372–375.
9. Fitch, W. L. (1999) Analytical methods for quality control of combinatorial libraries. *Mol. Diversity* **4,** 39–45.
10. Enjalbal, C., Martinez, J., and Aubagnac, J-L. (2000) Mass spectrometry in combinatorial chemistry. *Mass Spectrom. Rev.* **19,** 139–161.
11. Bray, A. M., Jhingran, A. G., Valerio, R. M., and Maeji, N. J. (1994) Simultaneous multiple synthesis of peptide amides by the multipin method—Application of vapor-phase ammonolysis. *J. Org. Chem.* **59,** 2197–2203.

Fig. 3. *(see opposite page)* Positive mass spectra regenerated from the TIC chromatogram of Fig. 2 at different retention times: **(A)** 8.98 min: expected peptide, $(M+2H)^{2+}$ = 641.4. **(B)** 9.66 min: byproduct containing a t-Bu protecting group, $(M+2H)^{2+}$ = 669.7. **(C)** 11.85 min: byproduct containing a Pmc protecting group, $(M+2H)^{2+}$ = 774.6. **(D)** 12.33 min: byproduct containing both t-Bu and Pmc protecting groups, $(M+2H)^{2+}$ = 802.8.