HPLC/NMR in Combinatorial Chemistry

Jefferson Chin,† Jay B. Fell,‡ Mark Jarosinski,§ Michael J. Shapiro,*,† and James R. Wareing‡

Department of Central Technologies, Oncology Research Program, and Department of Metabolic and Cardiovascular Diseases, Preclinical Research, Novartis Pharmaceuticals Corporation, 59 Rt. 10, East Hanover, New Jersey 07936

Received July 24, 1997

When sufficient material is available, NMR spectroscopy is unsurpassed for detailed structure determination in solution. However, the application of NMR to the determination of structure in multicomponent organic chemical systems prior to physical separation of the components has not been widely reported.[1] This is due to the complexity of the resulting NMR spectrum that contains severely overlapped signals.

Mass spectrometry (MS) is a powerful tool for the analysis of compound mixtures produced by combinatorial chemistry.[2] The use of mild ionization methods has made this technique suitable for both traditional organic molecules as well as for peptides. Superb sensitivity and speed of analysis are the two advantages MS has over other analytical techniques. A potential problem with the MS methodology involves a situation when isomolecular weight compounds are present. This can occur with compounds that are stereoisomers or positional isomers or simply by chance have the same molecular weight. When this occurs, analysis by MS can be ambiguous. LC/MS and MS/MS may be helpful, but their utility is achieved at the expense of the speed of analysis.

The recent redevelopment of HPLC/NMR allows for complete identification of individual components in complex mixtures.[3] The majority of reports using HPLC/NMR have been for drug metabolites.[4] The utility of HPLC/NMR methods for the analysis of organic chemical reactions and for peptides, however, has not been as widely reported.[5] It is presently possible to obtain routine high-quality HPLC/NMR 1D NMR data with as little as 5 μg of compound in the chromatogram peak. The detection limit using specially designed probes is presently on the order of 25−100 ng.[6] The full range of NMR experiments can be applied to the structure determination of compounds if necessary.[7]

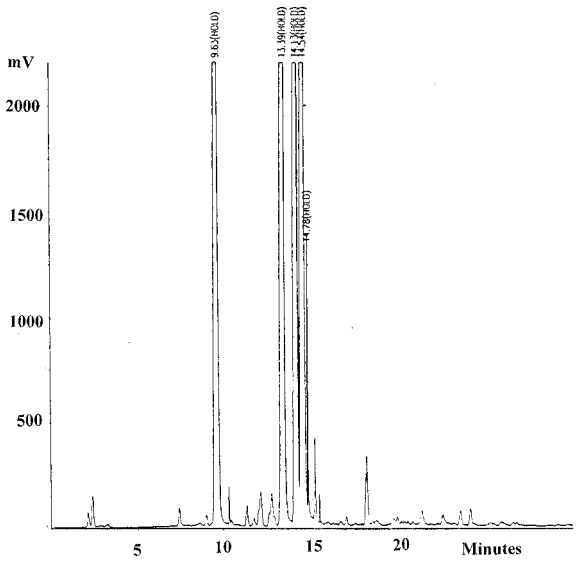

Figure 1. UV trace of the chromatographic separation for the four positional isomers **1**−**4**.

The utility of HPLC/NMR is demonstrated here for the special case of isomolecular weight compounds using two HPLC/NMR techniques: stopped-flow and on-flow NMR. The two examples investigated were (1) a four-component mixture of aromatic ring positional isomers and (2) a five-component pentapeptide mixture containing two isomolecular weight peptides.

Results

Positional Isomers. Using a split-mix synthesis, a mixture of four (dimethoxybenzoyl)glycines was prepared.[8] Samples obtained from this mixture were dissolved in 50 μL of acetonitrile−D_2O; the mixture containing approximately 100 μg of total compound was injected onto the HPLC Nova-Pak C18 column, and separation of the four positional isomers was accomplished using a gradient LC separation. The chromatographic separation, detected by UV, is shown in Figure 1. No attempt was made to further optimize the chromatography conditions to obtain better resolution.

The four-component mixture was first analyzed by stopped-flow HPLC/NMR. As each of the individual components was captured in the probe, the chromatography was stopped and the NMR data collection was initiated. For these data, 128 scans were more than sufficient to obtain suitable S/N. Figure 2 shows a stacked plot of the four compounds for the NMR region from $\delta = 3.5-8.5$.

The determination of structure for each compound is straightforward by consideration of the chemical shifts and the coupling patterns. The compound eluting at 9.6 min is identified as the 2,6-dimethoxy analogue **1** by the doublet at $\delta = 6.7$ for 2H, the triplet at $\delta = 7.4$, and the chemical shift equivalence of the methoxy resonances. The compound eluting at 13.4 min, **2**, shows two methoxy

† Department of Central Technologies.
‡ Department of Metabolic and Cardiovascular Diseases.
§ Oncology Research Program.

(1) Lin, M.; Shapiro, M. J. *J. Org. Chem.* **1996**, *61*, 7617−7619.
(2) Boutin, J. A.; Hennig, P.; Lambert, P.; Bertin, S.; Petit, L.; Mahieu, J.-P.; Serkiz, B.; Volland, J.-P.; Fauchere, J.-L. *Anal. Biochem.* **1996**, *234*, 126−141.
(3) (a) Spraul, M.; Hoffman, M.; Dvortsak, P.; Nicholson, J. K.; Wilson, I. D. *Anal. Chem.* **1993**, *65*, 327. (b) Strohschein, S.; Schlotterbeck, G.; Richter, J.; Pursh, M.; Tseng, Handel, H.; Albert, K. *J. Chromatogr. A* **1997**, *765*, 207−214.
(4) Shockcor, J. P.; Silver, I. S.; Wurm, R. M.; Sanderson, P. N.; Farrant, R. D.; Sweatman, B. C.; Lindon, J. C. *Xenobiotica* **1996**, *26*, 41−48.
(5) Albert, K. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 641.
(6) Olson, D. L.; Peck, T. L.; Webb, A. G.; Sweedler J. V. *Peptides: Chem., Struct. Biol.* **1996**, 730−731.
(7) Lindon, J. C.; Farrant R. D.; Sanderson, P. N.; Doyle, P. M.; Goughg, S. L.; Spraul, M.; Hoffman, M.; and Nicholson, J. K. *Magn. Reson. Chem.* **1995**, *33*, 857.
(8) Furka, A.; Sebestyan, F.; Asgedom, M.; Dibo, G. *Int. J. Peptide Protein Res.* **1991**, *37*, 487−493.

S0022-3263(97)01366-2 CCC: $15.00
Published on Web 01/23/1998

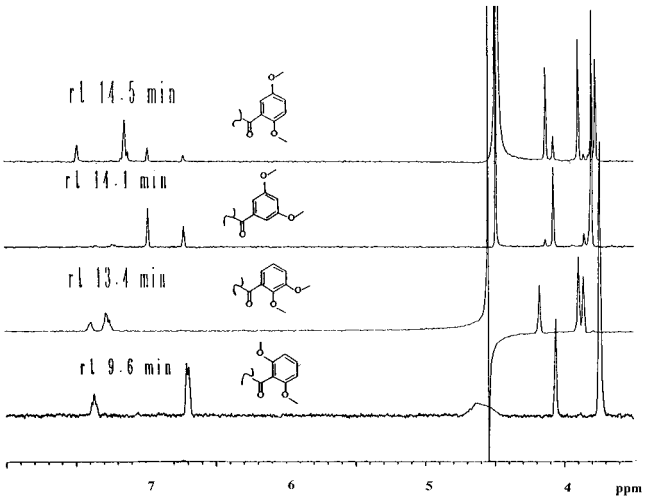

Figure 2. 500 MHz HPLC/NMR stopped-flow stacked NMR plot for compounds **1**−**4**.

resonances and a more complicated aromatic spectrum and was identified as the 2,3-dimethoxy derivative.

The assignment of compounds **3** and **4** was also readily accomplished, even though the peaks eluting at 14.1 and 14.5 min for compounds **3** and **4**, respectively, are overlapped in the chromatography. It is clear from Figure 2, by intensity and chemical shift considerations, that the extra resonances in each NMR spectrum are due to the incomplete separation. The compound eluting at 14.1 min is assigned to the 3,5-dimethoxy component from symmetry considerations while the remaining substance is the 2,5-dimethoxy compound based upon chemical shift considerations.

A second method of HPLC/NMR involves collecting the NMR data continuously as the sample passes through the probe. On-flow HPLC/NMR analysis, when sufficient material is available, is an efficient method for structure evaluation. In our on-flow experiment, approximately 200 μg of the crude four-component mixture was injected into the column in 50 μL of solvent. The data are represented in a two-dimensional plot where the *X* direction contains chemical shift information and the *Y* direction is the reconstructed LC retention time. The on-flow HPLC/NMR characterization of the four components is shown in Figure 3.

Four distinct sets of resonances can be observed in the 2D data set. The individual spectra can be extracted from the 1D slices along the *X* axis if so desired. The resolution in the individual spectra is of somewhat lower quality than in the stopped-flow method. However, the introduction of the second dimension allows for easy NMR signal assignment and thus structure evaluation, even of overlapping peaks in the LC separation.

Peptides. We investigated three different pentapeptide mixtures, ENXEF−OH, FNXEF−OH, and VNXEF−OH, where the variable amino acid X = D, Q, I, K, or T, to determine the structure of each component. In each of these five component mixtures are two compounds having identical molecular weights that could be difficult to identify by the simple MS methods used for combinatorial chemistry analysis, which generally yields unit mass resolution. Due to the small size of the sample, estimated to be less than 50 μg per peak in the chromatography separation, and the need to possibly perform 2D NMR on these samples, we chose to use only the stopped-flow methodology for structure determination.

The format used to assign the structures of FNXEF is illustrative of the assignment procedure and was also used in the other peptide mixtures. The chromatogram for FNXEF, as detected by UV, is shown in Figure 4.

All five components can be observed, although baseline resolution was not achieved and extensive overlap was present for chromatographic peaks 3 and 4. Nevertheless, stopped flow HPLC/NMR was performed, and all the compounds could be unambiguously assigned using 2D TOCSY data (Figure 5).

Since the solvent is D_2O-CH_3CN, the proton NMR shift correlations established for amino acids, which were obtained for aqueous solutions, cannot be relied on to make the assignments.[9] Fraction 1, retention time of 10.8 min, showed proton resonances in the TOCSY spectrum at δ = 1.3, 1.6, 2.9, and 4.1 as indicated in Figure 5. This is indicative of a lysine residue. The additional resonances are consistent with the amino acids phenylalanine δ = 4.2 and 3.1, 4.6, 3.18, 2.98; asparagine δ = 4.7 2.7, 2.55; and glutamate δ = 4.27, 2.34−2.05. Therefore, the sequence was assigned FNKEF.

The proton resonances for fraction 2, retention time of 11.4 min, at δ = 4.15 and 2.35 overlapped with the signals from glutamate as indicated in Figure 6. These data are consistent with the presence of glutamine, and therefore, the peptide is FNQEF.

The presence of aspartate in fraction 3 (retention time 12.4 min) was deduced from the 1D NMR spectrum, where additional signals in the δ = 2.6−2.8 region were observed. Due to severe overlap with both asparagine and phenylalanine resonances, the 2D TOCSY spectrum could not be used effectively for complete spectral assignment for this peptide. The use of the 1D NMR spectrum, however, still allowed the assignment as FNDEF.

Fraction 4 (retention time of 12.7 min) showed peaks from fraction 3 as well as new resonances arising from the presence of threonine at δ = 4.1, 4.05, and 1.07. A new phenylalanine peak is also observed at δ = 4.68, 2.7, and 2.6. This identifies the peptide as FNTEF. The final fraction, retention time of 16.7 min, contains isoleucine on the basis of its unique NMR spectrum, shown in Figure 7, consisting of a methyl doublet and triplet at δ = 0.78 and 1.04, respectively. This gives peptide FNIEF.

The other two peptide mixtures were also evaluated, and each component was identified. The retention order, under the same chromatographic conditions, for the variable amino acids (K, Q, D, T, I) for each component in the other two mixtures was identical with the one described. The isomolecular weight complication of lysine and glutamine is readily resolved in every case. All of the mixtures were also evaluated by LC/MS. As expected, fractions 2 and 3 gave the same molecular weight and could not be readily assigned. The other fractions could be assigned by MS and confirmed the results obtained by the HPLC/ NMR.

Conclusions

Although HPLC/NMR cannot be favorably compared to MS techniques in terms of sensitivity, it can neverthe-

(9) Wuthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.

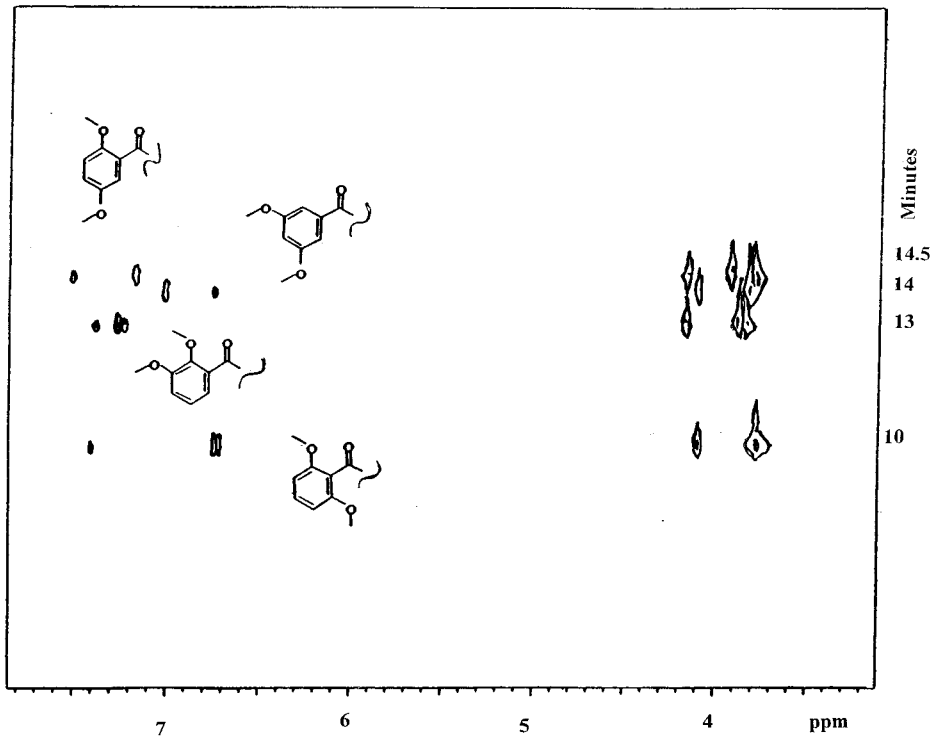

Figure 3. 500 MHz HPLC/NMR on-flow 2D plot for compounds **1**−**4**.

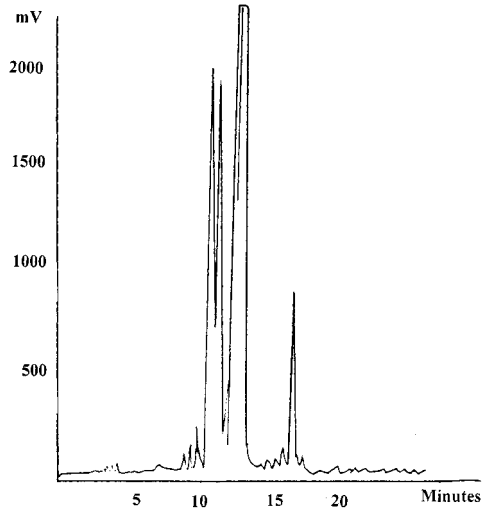

Figure 4. UV trace of the chromatographic separation for FNXEF peptide mixture.

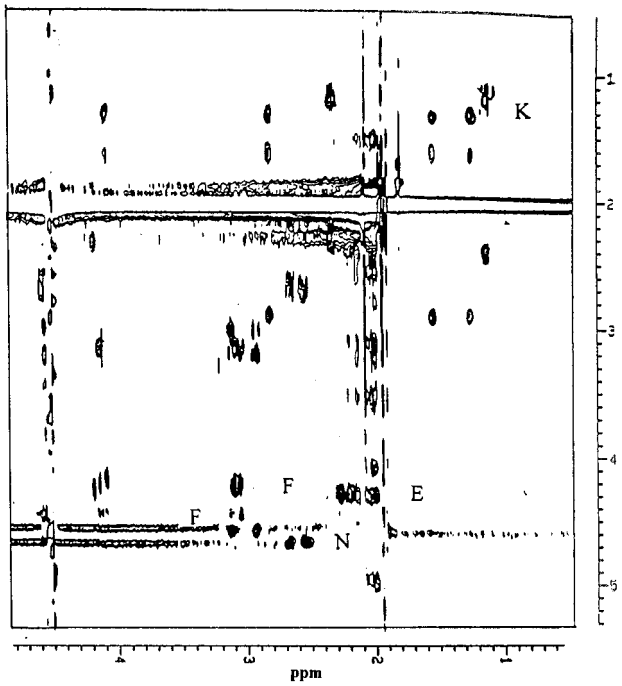

Figure 5. 500 MHz 2D TOCSY NMR spectrum for FNKEF obtained in the HPLC/NMR stopped-flow mode. Assignments of the amino acids are indicated on the spectrum.

less play a role in the structure determination of both peptides and organic molecules in mixtures generated by solid-phase synthesis. HPLC/NMR can be particularly useful in the identification of isomolecular weight compounds complimenting MS techniques. Furthermore, the HPLC/NMR method "sees" the totality of the sample and is not subject to either ionizing or chromophoric problems. In addition, HPLC/NMR also has the ability to provide detailed structure information in the absence of complete chromatographic resolution.

Experimental Section

(Dimethoxybenzoyl)amides. HMPB-Gly-FMOC resin (400 mg, 0.69 mmol/g substitution) (Novobiochem) was placed in a 25 mL capacity glass reaction vessel and swollen with dimethylformamide (DMF) (8 mL). The solvent was drained, and the resin was mixed for 20 min with 20% v/v piperidine in DMF (8 mL) via nitrogen bubbling. The solvent was drained, and the resin was washed with DMF (2 × 8 mL × 5 min), methanol (MeOH) (2 × 8 mL × 5 min), and DMF (2 × 8 mL × 5 min). 2,3-Dimethoxybenzoic acid (125 mg, 0.69 mmol) was placed into a reaction flask with DMF (2 mL). To this solution was added disopropylcarbodimide (55 μL, 0.35 mmol) and the mixture stirred at rt for 10 min. This anhydride formation procedure was repeated with the 2,5-dimethoxybenzoic acid, 2,6-dimethoxybenzoic acid, and 3,5-dimethoxybenzoic acid, all in separate

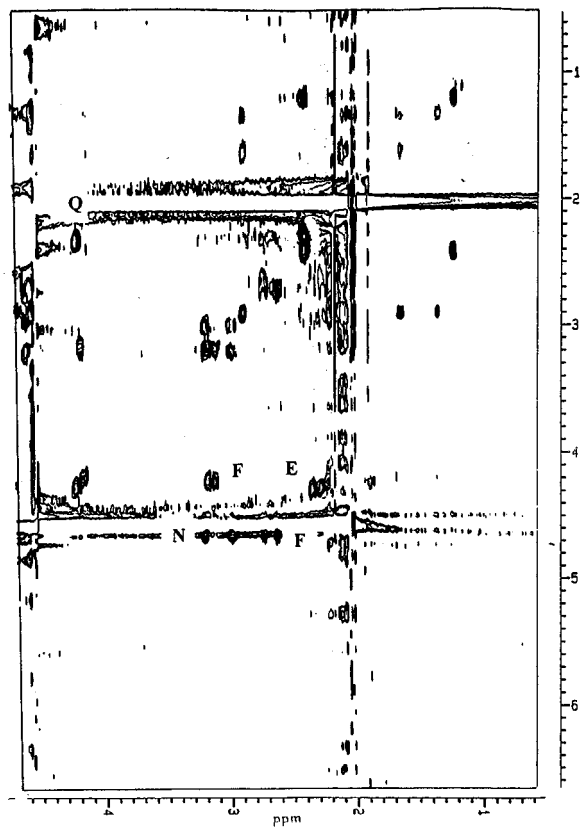

Figure 6. 500 MHz 2D TOCSY NMR spectrum for FNQEF obtained in the HPLC/NMR stopped-flow mode.

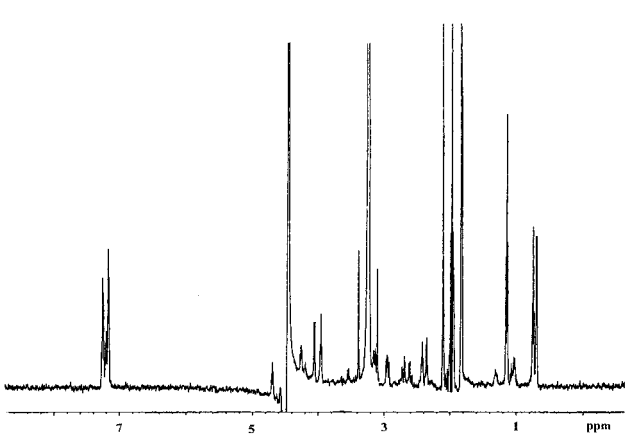

Figure 7. 500 MHz 1D NMR spectrum for FNIEF obtained in the HPLC/NMR stopped-flow mode.

flasks. To each of the preformed anhydrides was added the deprotected resin (100 mg) followed by hydroxybenzotriazole (HOBT) (65 mg, 0.43 mmol). After the solution was mixed on an orbital shaker at rt for 18 h, the solvent was drained and the resin was washed with DMF (2 × 8 mL × 5 min), MeOH (2 × 8 mL × 5 min), and dichloromethane (DCM) (2 × 8 mL × 5 min). The four pools of resin were mixed and cleaved with 10% v/v trifluoroacetic acid (TFA) in DCM (6 mL) for 1 h. The cleavage solution was collected, and the resin was washed with 10% v/v TFA in DCM (6 mL). The wash solution was combined with the cleavage solution, and the solvent was removed under a stream of nitrogen to give a yellow oil, which was further dried under vacuum (yield 53 mg, 80%).

Peptide Library. FmocGlu(OtBu)-Phe-Wang Resin (#). Fmoc-Phe-Wang resin (5.0 g, 0.540 mmol/g) (Novobiochem) was swollen, washed with *N*-methylpyrrolinone (NMP) (4 × 50 mL, 5 min per wash), and treated with 20% piperidine in NMP (1 × 50 mL for 5 min, then 1 × 50 mL for 20 min). Following successive NMP-DCM-NMP washes (2 × 30 mL for 2 min each solvent), a preactivated 5 degree solution of *N*-FmocGlu(OtBu)-OH (8.10 mmol, 3.45 g), HOBT (8.10 mmol, 27 mL of 0.3 M solution in NMP), PyBOP (8.1 mmol, 4.21 g), DIEA (24.1 mmol, 12 mL of 2 M solution in NMP) in NMP (25 mL) were added to the deprotected Phe-Wang resin and mixed by dry nitrogen sparge for 3 h. The reaction was complete by ninhydrin test, washed successively with NMP (2 × 25 mL for 2 min) and DCM (3 × 25 mL for 2 min), and dried in vacuo and resin loading determined by weight gain on resin and quantitative Fmoc titration (found 0.477 mmol/g, 98% yield).[10]

The synthesis of hexapeptide sublibraries was carried out on the Advanced ChemTech (ACT357) multiple peptide synthesizer using (#) FmocGlu(OtBu)-Phe-Wang resin (0.487 mmol/g) as starting material. The syntheses were scaled to provide approximately 0.3 g of the libraries. Coupling was affected with 5-fold excess of activated amino acid using DIC/HOBT in NMP for 1 h. Fmoc deprotection used 25% piperidine in DMF (1 × 3.5 mL per reaction vessel for 19 min). Resin splitting and pooling was carried out using the slurry method using DMF/DCM (1:1) as a suspending solvent. The split and pooling scheme provided libraries with the variation at position 3, and at that position Asp(tBu), Gln(Trt), Ile, Lys(Boc), and Thr(OtBu) were used.[11] The robotic programming was defined using the ACT357 synthesis managing software, and the synthesis run time was 2.5 days.

Fmoc-RNXEF-OH Fmoc-NNXEF-OH
Fmoc-ENXEF-OH

Fmoc-SNXEF-OH Fmoc-FNXEF-OH
Fmoc-VNXEF-OH

where X = Asp(tBu), Gln(Trt), Ile, Lys(Boc), and Thr(OtBu).

Each library was ninhydrin negative at the end of the synthesis, loading determined by quantitative Fmoc titration, and manually deprotected from the solid support using TFA−water−thioanisole−ethanedithiol−phenol 90:5:5:2.5:7.5 for the Arg-containing library and TFA−water−ethanedithiol 90:2.5:2.5, 60 min, for the remaining libraries. After deprotection, anhydrous ether (2 × 35 mL) was added to the filtered mixture and the precipitated peptides were collected by centrifugation. The crude libraries were vacuum-dried and were characterized by analytical reversed-phase HPLC (C-18 Vydac), electrospray mass spectroscopy, and HPLC NMR.

HPLC/NMR. The HPLC system consists of a Bruker LC22C pump, Bischoff UV detector, and Bruker's peak sampling unit (BPSU). Separation of the (dimethoxybenzyl)glycine derivatives was accomplished using a Waters 4.6 × 150 mm Nova-Pak C18 column. The mobile phases used consisted of 0.1% H_3PO_4/D_2O (A) and 0.1% H_3PO_4/ CH_3CN (B). The UV detector was set at 210 nm and the flow rate set for 0.8 mL/min. For on-flow analysis, elution of the compounds was achieved in under 10 min with an isocratic program of 25% B. Stopped-flow analysis required an LC gradient of 5−75% B in 35 min.

The pentamers were separated with a Vydac 4.6 × 250 mm C18 column. The mobile phases used were 0.1% TFA/D_2O (A) and 0.1% TFA/CD_3CN (B). The UV detector was set at 210 nm, and the flow rate was 0.8 mL/min. Stopped-flow analysis was performed using an LC gradient of 10−70% B in 30 min.

NMR analysis of the compounds used the DMX500 equipped with a 4 mm inverse triple (^1H, ^{13}C, ^{31}P) flow probe with a shielded *Z* axis gradient coil. Proton spectra were acquired using a standard solvent suppression 1D NOESY program with the irradiation of both the acetonitrile and HDO. For the on-flow

(10) Meinhofer, J.; Waki, M.; Heimer, E. P.; Lambross, T. J.; Makofske, R. C.; Chang, C. D. *Int. J. Peptide Protein Res.* **1979**, *13*, 35−42.

(11) (a) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−84. (b) Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature* **1991**, *354*, 84−86.

data, a data matrix of 2K in F2 by 64W in F1 with each increment having 32 scans was used. The spectral width was set for 12000 Hz and the ^{1}H 90° was 7.5 μs at a power level of 9 dB.

TOCSY data was collected with the irradiation of both acetonitrile and HDO. The spectral width for the experiments was set for 7500 Hz and a 70 ms spin lock was used. The TOCSY data was collected using 2K data in F2 and with 128 increments of 128 scans in F1. Data were processed using shifted sine squared function in both dimensions.

Acknowledgment. We would like to acknowledge Somchai Suwanrumpha for the MS data evaluation of the peptide mixture.

JO9713666