29

SYNTHETIC COMBINATORIAL LIBRARIES: A NEW TOOL FOR DRUG DESIGN

Methods for Identifying the Composition of Compounds from Peptide and/or Nonpeptide Libraries

Michal Lebl,[1] Viktor Krchňák,[1] Nikolai F. Sepetov,[1] Victor Nikolaev,[1] Magda Staňková,[1] Petr Kočiš,[1] Marcel Pátek,[1] Zuzana Flegelová,[1] Ronald Ferguson[1] and Kit S. Lam[2]

[1] Selectide Corporation
1580 E. Hanley Blvd., Oro Valley, Arizona 85737
[2] Arizona Cancer Center and
Department of Medicine, University of Arizona College of Medicine
Tucson, Arizona 85724

INTRODUCTION

Development of new leads for drug design and structure/function relationship studies were revolutionized by the introduction of combinatorial or "library" techniques (for review see e.g. (Moos et al., 1993; Gallop et al., 1994; Gordon et al., 1994)). These techniques allow for the generation and screening of millions of potentially active structures. Due to the well developed and finely tuned synthetic methodology, peptides were the first group of compounds evaluated by this new approach. However, the next logical challenge is to synthesize libraries of nonpeptidic structures. The combinatorial library approach applied at Selectide consists of three basic steps: (i) chemical synthesis based on the split synthesis method yielding a library with one test compound structure per one bead; (ii) screening of the library either using an on-bead binding assay or a multiple step release assay; and (iii) recovery of positive beads and determination of the structure of the test compound (Lam et al., 1991).

CHEMICAL LIBRARY TYPES

Each chemically synthesized combinatorial library represents a certain structural diversity and multiplicity. Libraries containing sequential repetition of amino acids (peptide libraries) are easy to synthesize and the structure of compound of interest can be easily determined by sequencing. However, such libraries do not contain very high structural diversity, since the only variable parameter is the type of side-chain connected to the C-alpha

Methods in Protein Structure Analysis, Edited by M. Z. Atassi and E. Appella
Plenum Press, New York, 1995

carbon of the peptide backbone, and those side-chains occupy only limited conformational space. Combining natural L amino acids with D amino acids brings more diversity, nevertheless, it is still quite limited. Over the last three years we have synthesized and screened approximately 400 peptide libraries. These libraries ranged from linear (with exposed N- or C-terminus), to cyclic (homo or heterodetic), to libraries with a high probability of regular structural features (alpha helix, beta turn), covering most of the conformational space which can be explored by a peptide structure with molecular weight below 1000.

The advent of non-peptide libraries increased the diversity of conformational space filled by the test compound subunits, as well as increased chemical diversity due to the nature of the subunits (see e.g. Simon et al., 1992; Cho et al., 1993; DeWitt et al., 1993; Nikolaiev et al., 1993; Bunin et al., 1994; Chen et al., 1994; Gordon et al., 1994; Lebl et al., 1994; Staňková et al., 1994). Combinatorial libraries of chemically synthesized compounds can be classified into several distinct groups in which libraries from individual groups represent certain structural types: (i) Libraries of small, compact, and relatively rigid structures (e.g. N-acyl-N-alkyl amino acids); (ii) Libraries based on a more or less rigid scaffold structure (usually multifunctional cyclic scaffold, e.g. derivatized cyclopentane or cyclohexane ring, functionalized steroid skeleton, tricarboxybenzene, diaminobenzoic acid); (iii) Libraries based on a flexible scaffold that is built during the synthesis of the library and can be randomized (branched scaffold based on diamino acids, α, β, γ, δ-library); (iv) Libraries of linear, sequential compounds (typical example is peptide library, including also N-substituted glycines — peptoids, or α, β, and γ amino acids containing library); (v) Libraries of small organic molecules (e.g. benzodiazepine type). Library types which we have explored are shown in figure 1.

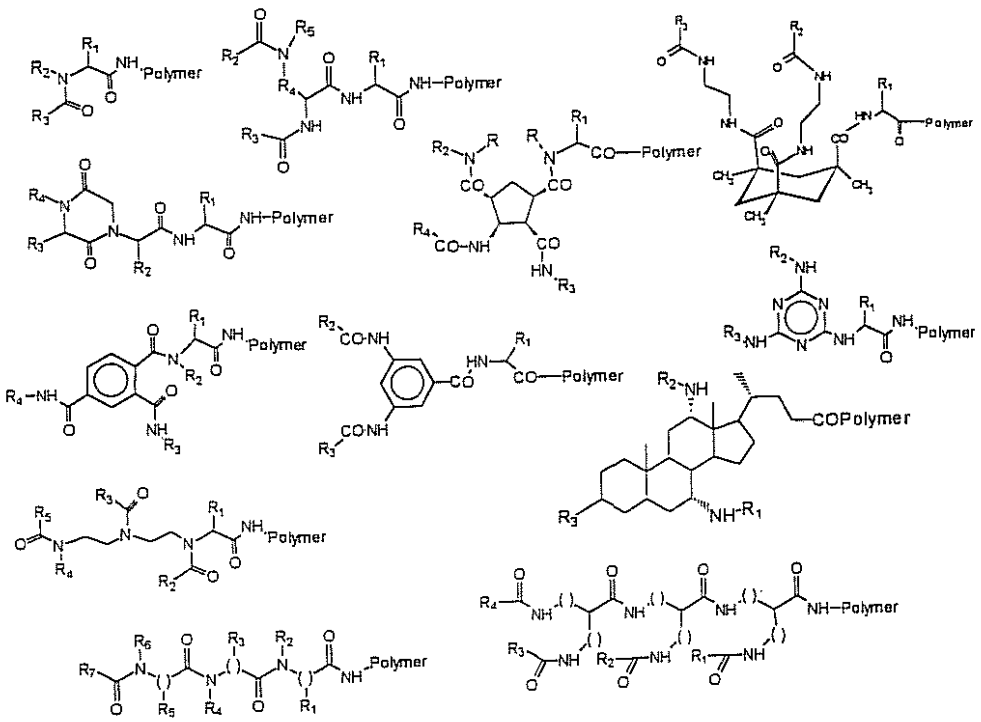

Figure 1. Structure of studied nonpeptidic library types.

DETERMINATION OF POSITIVELY REACTING STRUCTURES

Once the bead of interest is selected by the screening protocol, it is neccessary to determine the structure of the test compound responsible for the observed effect. Peptide structures can be easily determined by sequencing using automatic microsequencers. The structure determination of hits from nonpeptide libraries is complicated by the fact that the amount of positively reacting compound is limited. Standard bead of 100 μm diameter carries approximately 100 pmoles of the functional group onto which the library can be built. This amount of organic structure does not allow application of modern analytical methods for structure elucidation. The only exception is mass spectroscopy, which can be applied in cases when the library is composed of a limited number of structures or in cases where the fragmentation patterns are known and predictable. An example of mass spectroscopical structure determination is shown in figure 2. Beads expressing binding to streptavidin were selected from the small library of N-acyl-N-alkyl amino acids. The compound was cleaved from the bead and all components of the generated mixture were analyzed by MS/MS experiment. The deduced structures were resynthesized and their mass spectra matched those obtained for components cleaved from the beads. Binding to streptavidin was verified by solution assay (Stańková et al., 1994). The structures from a library based on the attachment of carboxylic acids to a modified Kemp's triacid scaffold were also analyzed by the MS/MS technique (figure 3).

In cases when mass spectroscopy cannot be used, a coding principle is applied (Brenner & Lerner, 1992; Kerr et al., 1993; Needels et al., 1993; Nielsen et al., 1993; Nikolaiev et al., 1993; Ohlmeyer et al., 1993). The various formats of coded libraries are given in figure 4. Linear coding is based on parallel synthesis of the screening and coding structure. Fractional coding is realized in two ways: (i) Simultaneous coupling of a tag together with tagged building block - e.g. coupling 0.05 equivalents of norleucine together with a D-amino acid to identify the configuration of the amino acid during sequencing, or (ii) Capping part of the growing chain by the tag which can be later cleaved and identified as such (Ohlmeyer et al., 1993), or as a tagged molecule (Sepetov, 1992; Youngquist et al., 1994). This last possibility is illustrated in figure 5, showing the tagging of a growing peptide chain by bromobenzoylation. After cleaving the mixture of full length peptide and truncated bromobenzoylated fragments, mass spectroscopic evaluation allows the elucidation of the peptide sequence (Sepetov, 1992). Binary coding utilizes a mixture of several blocks instead of a single coding block for coding building block of screening structure. Using a different set of coding blocks for coding different positions in the library allows for the construction of a coding structure in such a way that the coding blocks are cleaved and analyzed in a single step (Ohlmeyer et al., 1993).

Nature has coded proteins by nucleic acids for ages. However, coding nonpeptidic compounds by peptide structures is robust and reliable (Kerr et al., 1993; Nikolaiev et al., 1993). Each chemical individuum in the synthetic library is independently coded by a peptide whose composition can be easily resolved using an established technique (Edman degradation). The synthesis scheme of a coded library is shown in figure 6. Peptidic tags can be constructed in such a way that one cycle of Edman degradation will cleave all coding amino acids and a single HPLC run will reveal all components. The structure of a coding molecule (more appropriately a mixture of coding molecules) is shown in figure 7 together with the HPLC trace of the product of one Edman degradation cycle of this molecule.

The coding principle bears one inherent complication. If the screening process is being performed on the bead, the coding structure can interact with the target molecule. Three possibilities exist to prevent the interaction of the coding structure with the target: (i) The coding structure can be present in a very low concentration so that the interaction with the target molecule will not be seen under the conditions of the experiment; (ii) The coding and test structures can be physically separated; (iii) The test structure can be coded by a

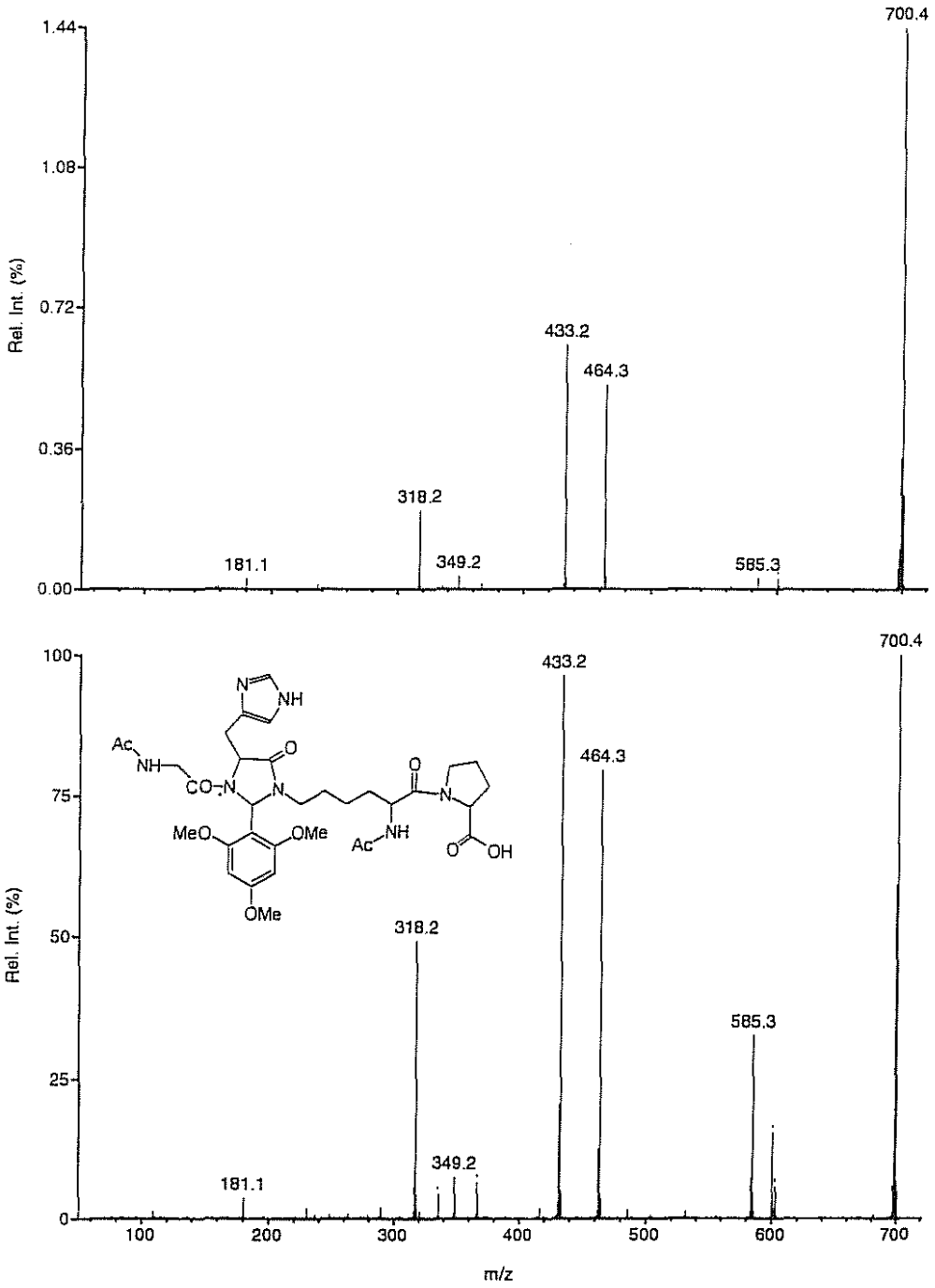

Figure 2. MS/MS spectrum of a compound released from a positively reacting bead (upper trace), and spectrum of the resynthesized compound, the structure of which was deduced from the upper spectrum (lower trace).

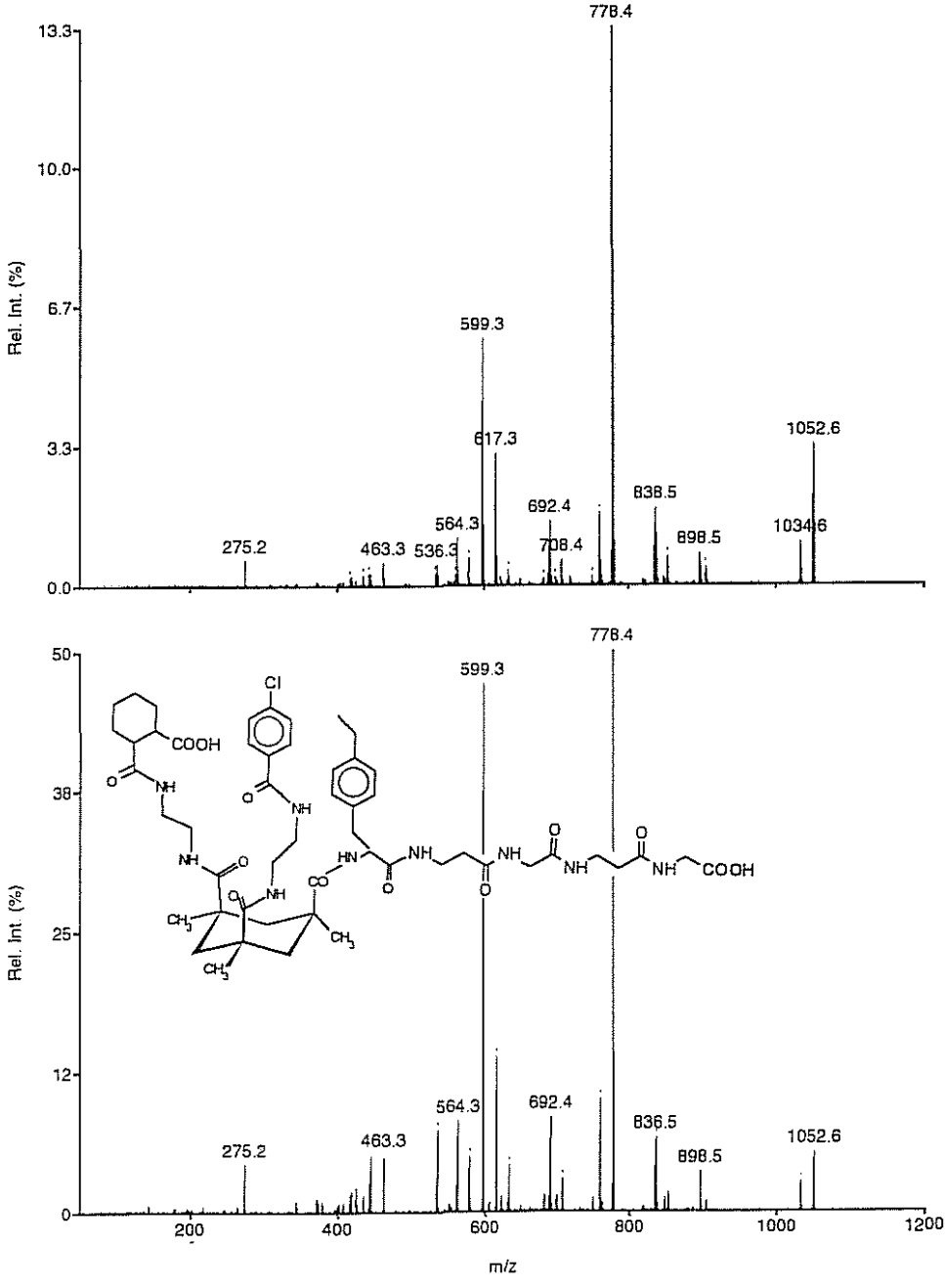

Figure 3. MS/MS spectrum of a compound from library constructed on Kemp's triacid. Spectrum of bead bound compound (upper trace) and of resynthesized compound.

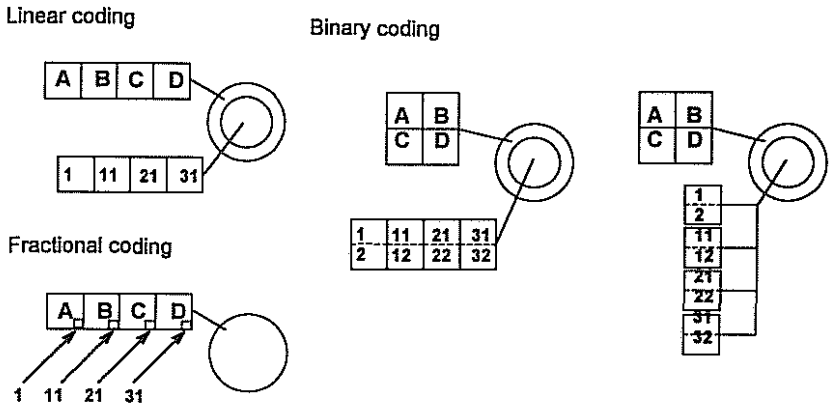

Figure 4. Coded library formats.

multiplicity of coding structures. The first possibility is not realistic in the case of peptide coding due to the limited sensitivity of peptide sequencing. However, it can be used advantageously in cases of coding by nucleic acids, where the coding structure can be conveniently amplified (Needels et al., 1993). The second option was explored by us recently (Vágner et al., 1994). Separation of the "surface" of the bead, which is available for interaction with the macromolecular target, from the "interior" of the bead, was achieved by enzymatic "shaving". To this target inaccessible "interior" was coupled the coding structure. The last possibility is based on the idea of coding using a different set of structures rather than one unique structure. This set of structures must provide unambiguous information about the chemistry performed on screening arm.

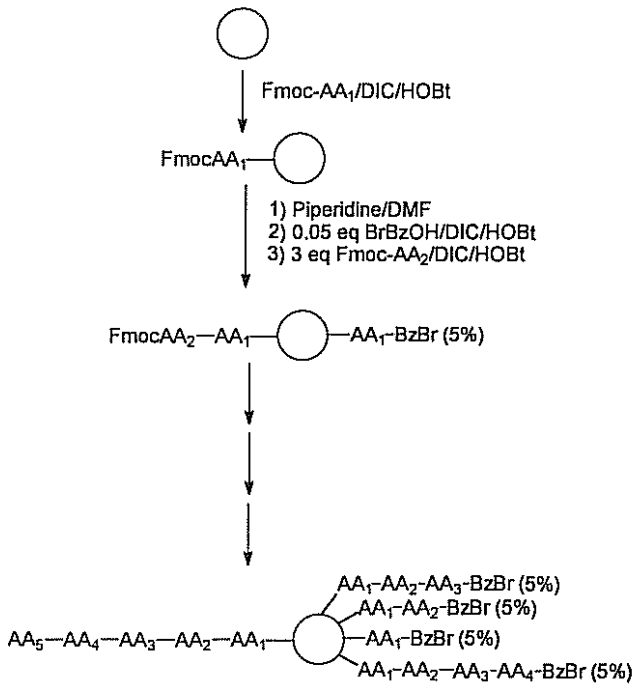

Figure 5. Coding by bromobenzoyl cap.

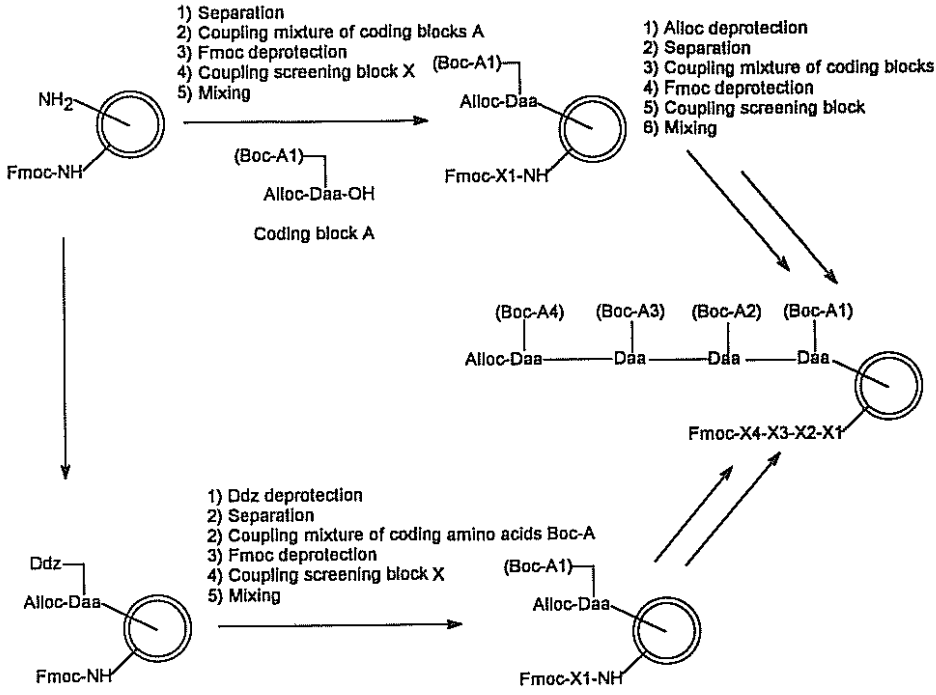

Figure 6. Synthesis scheme of a coded library.

ACKNOWLEDGMENTS

This work was supported in part by the NCDDG Cooperative Agreement (CA57723).

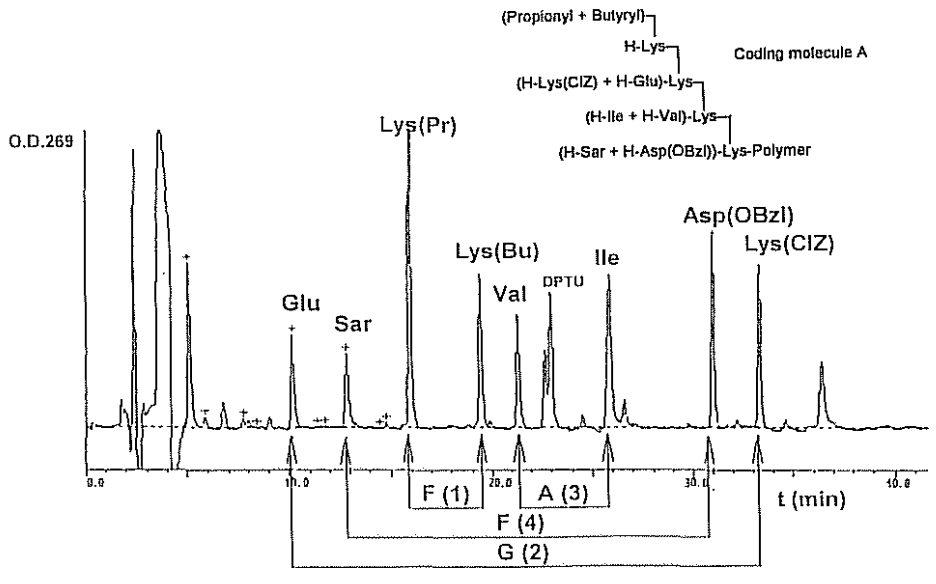

Figure 7. HPLC trace of the digital code generated by Edman degradation of the coding molecule A.

REFERENCES

Brenner, S., Lerner, R. A., 1992 Encoded combinatorial chemistry, *Proc.Natl.Acad.Sci.USA* 89(June):5381.

Bunin, B. A., Plunkett, M. J. and Ellman, J. A., 1994 The combinatorial synthesis and chemical and biological evaluation of 1,4-benzodiazepine library, *Proc.Natl.Acad.Sci.USA* 91(May):4708.

Chen, C., Randall, A. L. A., Miller, B. R., Jones, D. A. and Kurth, M. J., 1994 Analogous Organic Synthesis of Small-Compound Libraries: Validation of Combinatorial Chemistry in Small-Molecule Synthesis, *J.Am.Chem.Soc.* 116(6):2661.

Cho, C. Y., Moran, E. J., Cherry, S. R., Stephans, J. C., Fodor, S. P. A., Adams, C. L., Sundaram, A., Jacobs, J. W. and Schultz, P. G., 1993 An unnatural biopolymer, *Science* 261(3 Sep):1303.

DeWitt, S. H., Kiely, J. S., Stankovic, C. J., Schroeder, M. C., Reynolds Cody, D. M. and Pavia, M. R., 1993 "Diversomers": An approach to nonpeptide, nonoligomeric chemical diversity, *Proc.Natl.Acad.Sci.USA* 90:6909.

Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A. and Gordon, E. M., 1994 Applications of Combinatorial Technologies to Drug Discovery .1. Background and Peptide Combinatorial Libraries, *J. Med. Chem.* 37(9):1233.

Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A. and Gallop, M. A., 1994 Applications of combinatorial technologies to drug discovery.2. Combinatorial organic synthesis, library screening strategies, and future directions, *J. Med. Chem.* 37:1385.

Kerr, J. M., Banville, S. C. and Zuckermann, R. N., 1993 Encoded combinatorial peptide libraries containing non-natural amino acids, *J.Am.Chem.Soc.* 115:2529.

Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J., 1991 A new type of synthetic peptide library for identifying ligand-binding activity, *Nature* 354(7 November):82.

Lebl, M., Krchňák, V., Šafář, P., Stierandová, A., Sepetov, N. F., Kočiš, P. and Lam, K. S., 1994 Construction and screening of libraries of peptide and nonpeptide structures, in: Techniques in Protein Chemistry, Crabb, J. W., ed., Academic Press, San Diego, pp. 541-548.

Moos, W. H., Green, G. D. and Pavia, M. R., 1993 Recent advances in the generation of molecular diversity, in: Annual Reports in Medicinal Chemistry, Bristol, J. A., ed., Academic press. Inc., San Diego, CA, pp. 315-324.

Needels, M. C., Jones, D. G., Tate, E. H., Heinkel, G. L., Kochersperger, L. M., Dower, W. J., Barrett, R. W. and Gallop, M. A., 1993 Generation and screening of an oligonucleotide-encoded synthetic peptide library, *Proc.Natl.Acad.Sci.USA* 90(Nov):10700.

Nielsen, J., Brenner, R. A. and Janda, K. D., 1993 Synthetic methods for the implementation of encoded combinatorial chemistry, *J.Am.Chem.Soc.* 115:9812.

Nikolaiev, V., Stierandová, A., Krchňák, V., Seligman, B., Lam, K. S., Salmon, S. E. and Lebl, M., 1993 Peptide-encoding for structure determination of nonsequenceable polymers within libraries synthesized and tested on solid-phase supports, *Pept. Res.* 6:161.

Ohlmeyer, M. H., Swanson, R. N., Dillard, L. W., Reader, J. C., Asouline, G., Kobayashi, R., Wigler, M. and Still, W. C., 1993 Complex synthetic chemical libraries indexed with molecular tags, *Proc.Natl.Acad.Sci.USA* 90:10922.

Sepetov, N. F., Issakova, O. L., Krchňák, V. and Lebl, M., 1992 Peptide sequencing using mass spectrometry, U.S.A., 07/939,811.

Simon, R. J., Kaina, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Simon, N. G., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frankel, A. D., Santi, D. V., Cohen, F. E. and Bartlett, P. A., 1992 Peptoids: A modular approach to drug discovery, *Proc.Natl.Acad.Sci.USA* 89(October):9367.

Staňková, M., Issakova, O., Sepetov, N. F., Krchňák, V., Lam, K. S. and Lebl, M., 1994 Application of one-bead-one-structure approach to identification of nonpeptidic ligands, *Drug Development Research*, 33:146.

Vágner, J., Krchňák, V., Sepetov, N. F., Štrop, P., Lam, K. S., Barany, G. and Lebl, M., 1994 Novel methodology for differentiation of "surface" and "interior" areas of poly(oxyethylene)-polystyrene (POE-PS) supports: Application to library screening procedures, in: Innovation and Perspectives in Solid Phase Synthesis, Epton, R., ed., Mayflower Worldwide Ltd., Birmingham, pp. 347-352.

Youngquist, S., R, Fuentes, G., R, Lacey, M., P and Keough, T., 1994 Matrix-assisted Laser Desorption Ionization for Rapid Determination of the Sequences of Biologically Active Peptides Isolated from Support-bound Combinatorial Peptide Libraries, *Rapid Comm.Mass Spectrom.* 8:77.