Pergamon

Bioorganic & Medicinal Chemistry Letters 8 (1998) 2357–2362

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

DECONVOLUTION BY OMISSION LIBRARIES

Eugen Câmpian, Mark L. Peterson Hossain H. Saneii, and Árpád Furka*

Advanced ChemTech Inc., 5609 Fern Valley Road, Louisville, KY 40228, U.S.A.

Received 30 March 1998; accepted 24 June 1998

Abstract: Omission libraries, synthesized by omitting one amino acid in all coupling positions, are very efficient tools for the rapid identification of the amino acid components of bioactive peptides. Based on the determined amino acids, an occurrence library can be defined and prepared which is much less complex than the full one while still comprising the bioactive peptide.

Nowadays, combinatorial chemistry is an accepted and powerful approach in pharmaceutical research. It is, however, a very young field of science. One of the earliest and most popular synthetic methods, the portioning-mixing (PM) procedure, often called split-mix or split-combine method, was introduced only ten years ago.[1–3] This method has several inherent features that are responsible for its widespread application:

- It is easy.
- It is so efficient that millions of new compounds can be prepared by its use in less than a week.
- The substances are formed in nearly equimolar quantities.
- A single compound forms on each bead of the solid support.
- All possible structure combinations are formed that can be deduced from the monomers.
- It can be equally well applied to the synthesis of peptide and organic libraries.

A considerable repertoire of methods have already been developed for screening. In some of them, the one bead one compound feature is exploited by screening the library in tethered form[4] following the approach originally pioneered by Smith et al.[5] with individual peptides, or by releasing the synthesized compounds from individual beads and testing them in solution.[6,7] Recently, radiofrequency and laser encoding methods[8–10] have been introduced which have succeeded in combining the high efficiency of the PM procedure while offering all the advantages of the parallel synthesis. Since, by using this approach, individual compounds are produced in tens of milligram quantities, no specific deconvolution is needed in screening. The PM method can also be applied to produce soluble compound mixtures. In this case, in order to identify the bioactive components, special deconvolution procedure have to be followed like the iteration method[11,12] or the positional scanning.[13] Both methods are based on partial libraries. The authors introduce in this paper a new type of partial libraries: omission peptide libraries. As already shown by Carell et al.,[14] partial organic libraries prepared by omission of groups of amino acids can successfully be used in deconvolution.

0960-894X/98/$ - see front matter
PII: S0960-894X(98)00426-0

A peptide omission library can be prepared by omitting the same amino acid from all coupling positions. When an alanine omission library is synthesized, for example, no alanine is used in couplings, consequently no alanine containing peptides are formed. Using a very simple example Figure 1 shows how the composition of an omission library can be derived from that of the full one. A short symbol can be used to denote the omission libraries: the one letter symbol of the omitted amino acid preceded by the minus sign. The symbol of a glycine omission library, for example, is -G. An optional figure can appended to the symbol indicating the number of the amino acid building blocks in the peptides (length). Accordingly, the symbol of a glycine omission tripeptide library is: -G3. The number of peptides in full and omission libraries is listed in Table 1.

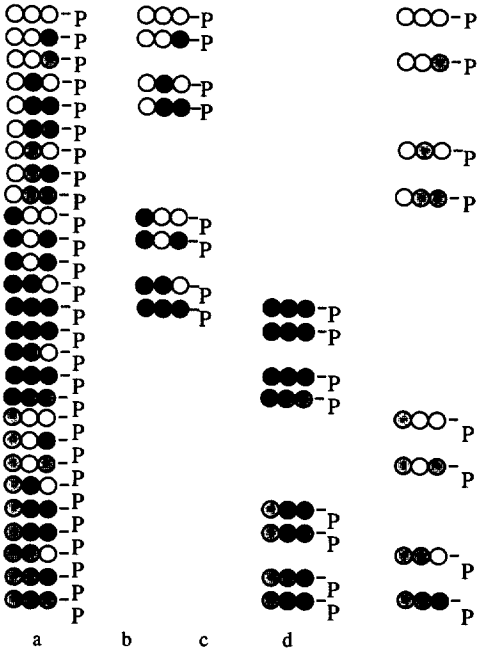

Figure 1. Deriving the three possible omission libraries (b, c, and d) from a full tripeptide library (a) composed from three (white, gray, and black) amino acids. In the synthesis of b, c and d libraries the "gray", "white", and "black" amino acids are omitted, respectively.

An omission library can be used to determine whether or not an amino acid is present as an essential (nonreplaceable) amino acid residue in the bioactive component of a peptide library. If glycine, for example, is an essential amino acid in a bioactive peptide, -G is expected to show no activity since all glycine containing peptides - including the active peptide - are missing. If a full set of omission libraries is available, it can be used to determine the amino acid composition of the bioactive peptide.

In order to prove the utility of omission libraries, we first prepared a full tripeptide amide library, a full set (20) of omission tripeptide amide libraries, and tripeptide acid library (omission of the amide group).

Table 1. Number of peptides in full and omission libraries (constructed from 19 amino acids + pyroglutamic acid in N-terminal position)

Length	Full	Omission
2	380	342
3	7,220	6,156
4	137,180	110,808
5	2,606,420	1,994,544
6	49,521,980	35,901,792

In construction of the full libraries, 19 amino acids were used in the first and second coupling position (cysteine was omitted from all libraries) and pyroglutamic acid was added to this set for the N-terminal position.

The libraries were prepared according to the PM method[1-3] using the Advanced ChemTech Model 357 Flexible Biomolecular Synthesizer[15] following the Fmoc protection strategy. The synthesized libraries were submitted to model screening experiments which examined, by radioimmunoassay, the competitive inhibition of binding of LH-RH[16] to its polyclonal antibody.[17] The LH-RH radioimmunoassay kit used in determinations was also obtained from Advanced ChemTech.

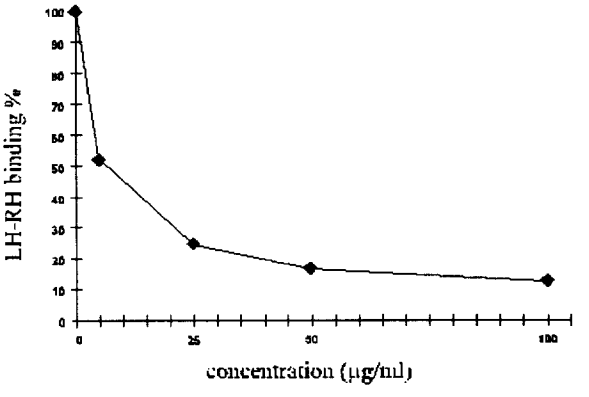

Figure 2. Effect of the concentration of the full tripeptide amide library on binding of radioactively labeled LH-RH to its polyclonal antibody.

The first experiment was carried out to determine the dependence of inhibition of binding on the concentration of the full library. The results presented in Figure 2 show that the full tripeptide amide library competes with the radioactively labeled LH-RH in binding to the antibody since raising the concentration of the

library reduces the quantity of LH-RH bound to the antibody. This result also suggests that the optimal concentration for further experiments should be around 50 μg/mL.

Next, the 20 omission libraries were submitted to binding experiments. Figure 3 demonstrates the result. Omission of pyroglutamic acid and the C-terminal amide group is marked by p and -a, respectively. It can be clearly seen that binding of the radio-labeled LH-RH is reduced less by the following omission libraries:

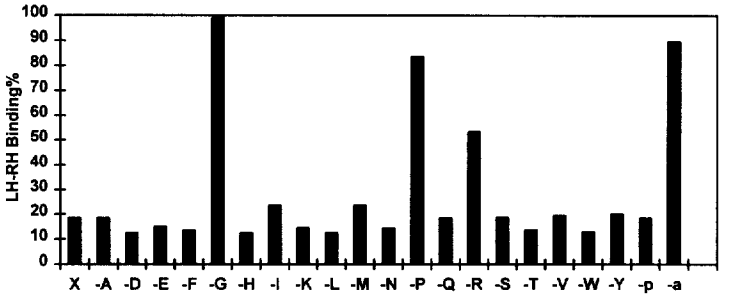

Figure 3. Effect of omission libraries. X and -a mean full tripeptide amide and full tripeptide acid libraries, respectively, while -p denotes pyroglutamic acid omission library. Concentration in these tests was 50 μg/mL.

-G, -P, and **-R**. This means that the amino acids essential for binding are: **glycine, proline, and arginine**. The **terminal amide group** also proved to be essential since the full tripeptide acid library (-a) does not signifficantly reduce the LH-RH binding.

The results of the experiments carried out with omission libraries gave no indication about the position of the amino acids within the sequence of the active peptide. Despite this, the information gained by only 21 screening experiments was very valuable. They defined a library of very low complexity. If an "amino acid occurrence library" is synthesized by varying only three amino acids, Gly, Pro, and Arg in all of the three positions, the active tripeptide is expected to be present among the 27 components of this library. In other words, by screening with omission libraries, the complexity of the library in which the active peptide is found can be reduced from the original 7220 to only 27.

Positions of the identified amino acids could be determined by using one of the following three possibilities:

- Preparation by parallel synthesis and screening of the 27 components of the occurrence library.
- Application of positional scanning to the occurrence library (preparation and screening of nine sublibraries).
- Positional scanning with nine sublibraries of the full library.

Since we previously prepared a full set (58) of sublibraries for using them in other experiments, we chose to realize the latter possibility.

In order to determine the positions of glycine, proline and arginine in the tripeptide responsible for the competitive binding to the LH-RH antibody, the following sublibraries were tested:

1G, 2G, 3G

1P, 2P, 3P

1R, 2R, 3R

G, P, and R are amino acids occupying nonrandomized (1, 2, or 3) coupling positions.

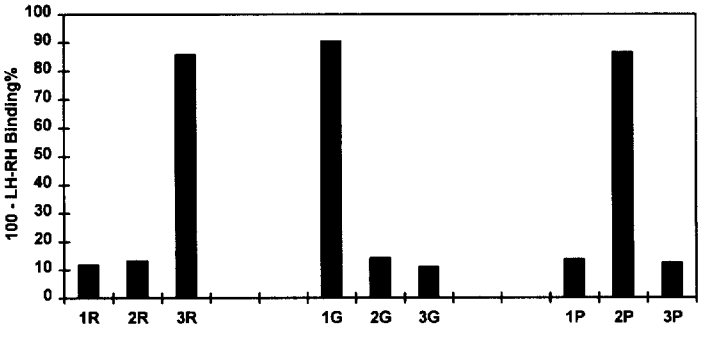

Figure 4. Determination of the position of the previously identified amino acids by positional scanning.

The result of the experiments is shown in Figure 4. It can be seen that 3R, 1G, and 2P considerably reduce the binding of the labeled LH-RH to the antibody. This means that glycine, proline and arginine occupy the first, second and third coupling position in the active tripeptide. Therefore, the sequence of the tripeptide exhibiting specific binding is **Arg-Pro-Gly-NH_2**, which happens to be identical with the C-terminal sequence of LH-RH.

The omission libraries can be easily synthesized and can be used very effectively in screening of peptide libraries. The amino acid sequence of the bioactive peptides can be determined, of course, without omission libraries, for example, by positional scanning. In that case, however, signifficantly more screening experiments would have to be performed (in the above example 58 instead of 29). Applicability of omission libraries is by no means limited to the screening of peptide libraries. They are expected to work equally well with nonpeptide libraries as well since the methods of construction and screening of these libraries is essentially the same as those of the peptide libraries.

Acknowledgement: The authors thank Dr. Joanne Chou, Advanced ChemTech Inc., for the binding experiments.

References and Notes

1. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. In *Highlights of Modern Biochemistry, Proceedings of the 14th International Congress of Biochemistry;* VSP: Utrecht, 1988; Vol. 5, p 47.
2. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Abstrracts of Papers,* 10th International Symposium of Medicinal Chemistry; Budapest, 1988; p 288.
3. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Int. J. Peptide Protein Res.* **1991**, *37*, 487.
4. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature (London)* **1991**, *354*, 82. and its correction: *Nature (London)* **1992**, *360*, 768.
5. Smith, J. A.; Hurrel, J. G. R.; Leach, S. J. *Immunochemistry* **1977**, *14*, 565.
6. Burbaum, J. J.; Ohlmeyer, M. H. J.; Reader, J. C.; Henderson, I.; Dillard, L. W.; Li, G.; Randle, T. L.; Sigal, N. H.; Chelsky, D.; Baldwin, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6027.
7. Jayawickreme C. K.; Geraminski G. F.; Quillan J. M.; Lerner M. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91,* 1614.
8. Moran E. J.; Sarshar, S.; Cargill, J. F.; Shahbaz, M. M.; Lio, A.; Mjalli, A. M. M.; Armstrong, R. W. *J. Am. Chem. Soc.* **1995**, *117*, 10787.
9. Nicolaou, K. C.; Xiao, X. Y.; Parandoosh, Z.; Senyei, A.; Nova, M. P. *Angew.Chem.* **1995**, *107*, 2476.
10. Xiao, X.; Zhao, C.; Potash, H.; Nova, M. P. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 780.
11. Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature (London)* **1991**, *354*, 84.
12. Erb, E.; Janda, K. D.; Brenner, S. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11422.
13. Pinilla, C.; Appel, J. R.; Blanc, P.; Houghten, R. A. *Biotechniques* **1992**, *13*, 901.
14. Carell, T.; Winter, E. A.; Rebek, J. Jr. *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2061.
15. Saneii, H. H.; Shannon, J. D.; Miceli, R. M.; Fischer, H. D.; Smith, C. W. In *Peptide Chemistry 1993;* Okada, Y., Ed.; Protein Research Foundation: Osaka, 1994, pp 117–119.
16. Matsuo, H.; Baba, Y.; Nair, R. M. G.; Arimura, A.; Schally, A. V. *Biochem. Biophys. Res. Commun.* **1971**, *43*, 1334.
17. Berson, S. A.; Yalow, R. S. J. *Clinical Investigations* **1957**, *36*, 873.
18. Furka, Á. *Drug Development Research* **1994**, *33,* 90.