Bioorganic & Medicinal Chemistry Letters, Vol.3, No.3, pp. 413-418, 1993
Printed in Great Britain

0960-894X/93 $6.00 + .00

CHEMICAL SYNTHESIS OF PEPTIDE LIBRARIES

Ferenc Sebestyén, Gábor Dibó, Attila Kovács[¶] and Árpád Furka

Department of Organic Chemistry and [¶]Department of General Zoology,
Eötvös Loránd University, P.O. Box 32, H-1518 Budapest 112, Hungary

Abstract: The efficiency of the "portioning-mixing" method is demonstrated by the synthesis of more than one million hexapeptides in 5 days. An example is described to illustrate the possibility of completion of partial peptide libraries. Our method is also shown to be suitable to carry out binary peptide synthesis. In addition, a new screening method is outlined based on specific interaction of peptides with living cells.

Conventional peptide research often begins with the isolation of a new biologically active peptide which is then sequenced and synthesized. This is usually followed by the synthesis and screening of its fragments and analogs. The synthetic work as well as the biological screening essentially deals with individual peptides. This - even if the advantages offered by the solid phase method[1] are considered - is a slow process. A significant improvement in the efficiency of the synthetic methods has been observed in the last two decades as a result of the development of different simultaneous synthetic procedures[2-7] initiated by Rietschoten *et al.*[2] In addition, the inspiring experiments of Geysen *et al.*[5] opened a completely new possibility: screening without removal of the peptides from the solid support. The really dramatic changes in the efficiency, however, were the result of the last 4 years. Three independent - and radically different - methods were introduced enabling the user to synthesize millions of peptides assuring enormous savings in labor, time and expenses.

Our synthetic methodology introduced in 1988[8-10] is based on the Merrifield solid phase approach[1] and can be called a "portioning-mixing" procedure since the two key operations are: portioning (before coupling) and mixing (after coupling). Another important method ploneered by Smith and his colleagues appeared in 1990[11]. The peptide sequences in this method are expressed on the surface of phage after inserting into their genomes random oligonucleotide sequences. In the third method published by Fodor *et al.* in 1991[12] the peptides were synthesized in different but predetermined locations on the surface of a glass microscope slide.

Two other procedures were published in 1991: the "split synthesis"[13] and the "divide, couple and recombine (DCR) process"[14]. The essential operations in these procedures, however, were those described in our method: portioning and mixing[18].

Efficiency of the "portioning-mixing" synthesis

Our previous paper[10] was devoted to describing the technique of the "portioning-mixing" synthesis and to show through the synthesis of simpler peptide mixtures that the expected components were really formed. The efficiency of the method is now demonstrated by a brief description of the synthesis of a

mixture of resin-bound hexapeptides. All peptides were designed to carry *G* (one letter codes are used) at the C-terminus and the same set of 17 amino acids *(A,D,E,F,G,H,I,K,L,M,N,P,S,T,V,W,Y)* were varied in each of the remaining five positions. Thus - before starting the synthesis of the mixture - aminomethyl resin (200-400 mesh) was acylated by γ-aminobutyric acid to form a spacer between the resin and the peptides then coupled with glycine representing the C-terminal residue of all peptides. The synthesis was completed in 5 consecutive days by two technicians using solely the manual method without any particular modification[10]. Every day a complete set (17) of coupling cycles was executed: the resin was divided into 17 portions, every portion was coupled with one of the 17 varied amino acids, then mixed. On the last day all of the protecting groups were removed. The efficiency of the method can be characterized by comparing the number of synthesized peptides to the number of executed coupling cycles:

Number of peptides: $17 \times 17 \times 17 \times 17 \times 17 = 1,419,857$

Number of coupling cycles: $17 + 17 + 17 + 17 + 17 = 85$

The number of coupling cycles as well as the number of days needed for the synthesis of complete oligopeptide libraries are summarized in Table 1. The data clearly demonstrate the exceptional efficiency of our synthesis. Millions or even billions of peptides can be synthesized in a week or so. This means that the demand of labor is no longer a limiting factor in synthesis.

Table 1. *Number of coupling cycles in the synthesis of complete oligopeptide libraries*

residues	peptides	coupling cycles	days
2	4.00×10^2	40	2
3	8.00×10^3	60	3
4	1.60×10^5	80	4
5	3.20×10^6	100	5
6	6.40×10^7	120	6
7	1.28×10^9	140	7

In the pentapeptide stage of synthesis described above, equal samples were removed from every portion before mixing, and treated separately. Removal of the protecting groups resulted in 17 pentapeptide mixtures having 4913 components each. All of the peptides in each sample had the same C-terminus *(G)* and, in addition, they carried the same N-terminal residue (one of the 17 amino acids). These samples were used in the binding tests with cells.

Binary synthesis by the "portioning-mixing" method

The very remarkable synthetic process called binary synthesis introduced recently by Fodor *et al.*[12] seems to be even more efficient than that described above since the number the products generated is expected to be 2^n, where n is the number of coupling cycles. Thus, as few as 10 coupling cycles yield 1024 peptides! In the original ten step process, a glass microscope slide was applied as the solid support and a

directed beam of light was used to remove the protecting groups. Although our "portioning-mixing" method applies a substantially different technique, it is as good - as shown below - in realization of the binary synthesis as the original light directed method.

A four step binary synthesis was realized to yield a mixture simple enough to assure the easy identification of the products. An alanyl-resin (prepared from chloromethyl resin) was used as starting material. Four coupling cycles were performed in the following order: *K, W, E,* and *L*. In every coupling cycle three operations were executed:

(i) dividing the resin into 2 equal portions

(ii) coupling an amino acid to one of the portions

(iii) mixing the two (coupled and uncoupled) portions

After the last *(L)* coupling step the two portions of the resin were not mixed. The peptides were removed from the resin separately and submitted to two-dimensional paper electrophoresis (at pH 6.5 and 2.0). All of the 16 expected products could be identified using our computer-aided identification procedure[15]. The following components were found:

portion uncoupled with *L:*

A, KA, WA, EA, WKA, EKA, EWA, EWKA

portion coupled with *L:*

LA, LKA, LWA, LEA, LWKA, LEKA, LEWA, LEWKA

The binary synthesis is a very efficient process indeed. As shown above, synthesis of the mixture of 1,419,857 hexapeptides by our original method needs 85 coupling cycles. The binary synthesis would yield as many as 2,097,152 components after only 21 coupling steps. Despite this fact, if one decides to systematically check possible peptide sequences for biologically active components, the use of a binary synthesis strategy is not advised because of an inherent feature of the method. The product of the synthesis is always a mixture of peptides of different length. Thus, a complete library can not be synthesized in one run. Furthermore, a mixture - once synthesized - can not be transformed to a complete library by additional runs.

Completion of partial libraries synthesized by the "portioning-mixing" method

Partial peptide libraries are formed in the "portioning-mixing" process if less than 20 amino acids are varied in one or more positions. It can be shown - as contrasted with the binary synthesis - that any such library can be completed. This is demonstrated by a simple example.

Let's suppose that a partial pentapeptide library is synthesized by varying the following 17 amino acids in every position:

A,D,E,F,G,H,I,K,L,M,N,P,R,S,T,V,W,Y

that is, three amino acids *(C, Q* and *R)* are omitted from variations in each position. This library can be synthesized - as shown above - in 85 coupling cycles and the number of components is 1,419,857. As Table 1 shows, the complete pentapeptide library has 3.2 million components, and the synthesis needs 100 coupling cycles. The question is: how can one synthesize the missing 1,780,143 peptides to complete the library?

Table 2. *Synthesis of missing peptides in five different runs*

Position	Run 1	Run 2	Run 3	Run 4	Run 5
1	3	17	17	17	17
2	20	3	17	17	17
3	20	20	3	17	17
4	20	20	20	3	17
5	20	20	20	20	3
Coupling cycles	83	80	77	74	71
Number of peptides	480 000	408 000	346 800	294 780	250 563

Synthesis of the missing peptides can be carried out in 5 parallel runs, as indicated in Table 2, by varying in the different positions 3 *(C, Q* and *R)*, 17 *(A,D,E,F,G,H,I,K,L,M,N,P,S,T,V,W,Y)* or all 20 amino acids (positions numbered from the C-terminus). The number of coupling cycles or the number of peptides synthesized in any of the runs can be calculated by summing or multiplying the figures within the columns, respectively. The total number of peptides formed in the five runs is 1,780,143 (exactly what was expected) and that of the additional coupling cycles is 385. This latter number - which seems to be too high if one considers that only 100 coupling cycles are needed in the synthesis of the complete library - can significantly be reduced by executing most of the coupling cycles in combined form according to Table 3. Thus, for example, in position 1, runs 2-5, and in position 4, runs 1-3 can be executed in combined form. The resin is always divided between runs according to the number of expected peptides. As a consequence, for example, couplings in position 1 are preceded by dividing the resin between run 1 and runs 2-5 in a ratio of 480,000:1,300,143 (*ca.* 27:73).

Table 3. *Coupling cycles in combined form*

Position	Run 1	Run 2	Run 3	Run 4	Run 5
1	3				17
2	20	3			17
3	20		3		17
4	20			3	17
5	20				3

The minimum number of coupling cycles (by summing the figures of Table 3) is 163. This means that completing such a partial library needs more coupling cycles than the synthesis of the complete library. It deserves consideration, however, that the synthesis of the rest of the components can be accomplished using a smaller amount of reagents (56% in this case, since only 56% of the 3.2 million peptides have to be synthesized). So depending on the relative price of the labor and that of the reagents one may decide to resynthesize the whole library or only the missing components.

Binding test with cells

Screening tests with both resin bound and free mixtures of peptides synthesized by the "portioning-mixing" method have already been described[13,14]. The binding experiments were carried out with proteins. Our preliminary results suggest, however, that a new kind of screening test can be introduced which is based on specific interaction between peptides and cells. The principle of the method was inspired by the well known fact that cells can be separated by affinity chromatography, in which, specific binding is involved. Screening with cells has a real advantage: no staining is needed since binding of cells to the beads can be simply observed by a microscope.

In a typical experiment, 0.5 mg of pentapeptide mixture bound to resin beads (N-terminal residue: *Y*, see above) and 0.5 mL of red blood cell suspension[16] was kept at 0ºC for 15 min, shaking gently several times, then pipetted into 10 mL Seglen's suspension buffer (Ssb)[17]. The diluted suspension was gently centrifuged (200 rpm, 3 min) to facilitate sedimentation of the beads. The excess of cells were removed by carefully pipetting out the supernatant. The sediment was diluted with 3 mL Ssb, pipetted into a Petri-dish then surveyed in an Opton inverted microscope. The great majority of the beads were completely bare. The fully packed ones could easily be observed and removed by a capillary tube attached to a pipette. The cells, although proved to be strongly bound to the beads, could be removed by adding a drop of 8 M urea. After washing, the beads were ready for sequencing as described by Lam *et al.*[13].

Comparison of the "portioning-mixing" synthesis to other methods

The "portioning-mixing" method (PMM) has two important advantages as compared to the biosynthetic method (BSM)[11] and the light directed one (LDM)[12]: (i) it can be used to prepare both resin-bound and free peptide mixtures, thus assuring a greater versatility (and perhaps relevance) in screening[13,14] and (ii) the mixtures can be synthesized in large quantities sufficient for many screening tests. Only PMM and LDM allow monomers like D-, and unnatural amino acids, sugars *etc.* to be used as building blocks in the synthesis of the oligomers. Furthermore, it can be shown, that both methods need the same number of coupling cycles if complete peptide libraries are synthesized (Table 1), but the process is expected to take a longer time for LDM since the couplings can be carried out only one by one. Parallel couplings (like coupling of amino acids to the portions of resins in PMM) can not be carried out. It is, however, a unique and enviable feature of LDM that its screening test does not need sequencing. It is an additional advantage of LDM that the glass slides covered by peptides can repeatedly be used in screening tests since the attached proteins can probably be completely removed. In PMM only the non-binding part of

the resin can be reused since the beads showing specific binding are removed and destroyed during sequencing. BSM and PMM have already been shown to be suitable for preparing millions of peptides (although only PMM assures the formation of complete libraries with a 1:1 molar ratio of the components), which in the case of LDM awaits experimental realization.

References and Notes

1. Merrifield, R. B. *J. Amer. Chem. Soc.* **1963**, *85*, 2149.
2. van Rietschoten, J.; Tregear, G. W.; Leeman, S.; Powell, D.; Niall, H.; Potts, J. T. *Peptides 1974*; Wolman, Y., Ed.; Wiley: New York, 1975; pp.113-115.
3. Gorman, J. J. *Anal. Biochem.* **1984**, *136*, 397.
4. Krchnak, V.; Vagner, J.; Mach, O. *Int. J. Peptide Prot. Res.* **1989**, *33*, 209.
5. Geysen, H. M.; Meloen, R. H.; Barteling, S. J. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3998.
6. Houghten, R. A. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5131.
7. Frank, R.; Döring, R. *Tetrahedron* **1988**, *44*, 6031.
8. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Abstr. 14th Int. Congr. Biochem.*, Prague, Czechoslovakia, **1988**; Vol. 5., p. 47.
9. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Abstr. 10th Int. Symp. Med. Chem.*, Budapest, Hungary, **1988**; p. 288.
10. Furka, Á.; Sebestyén, F.; Asgedom, M.; Dibó, G. *Int. J. Peptide Prot. Res.*, **1991**, *37*, 487.
11. Scott, J. K.; Smith, G. P. *Science*, **1990**, *249*, 386.
12. Fodor, S. P. A.; Read, J. L.; Pirrung, M. C.; Stryer, L.; Tsai Lu, A.; Solas, D. *Science*, **1991**, *251*, 767.
13. Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature*, **1991**, *354*, 82.
14. Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature*, **1991**, *354*, 84.
15. Furka, Á.; Sebestyén, F.; Gulyás, J. *Proc. 2nd Int. Conf. Biochem. Separations*, Keszthely, Hungary, **1988**; p. 35.
16. A suspension of red blood cells was prepared from heparinized rat blood. The cells were washed and centrifuged three times before use to remove serum proteins.
17. Seglen, P. O. *Meth. Cell Biol.* **1976**, *13*, 29.
18. Note added in proof: Lam et al. published a correction (Nature, **1992**, *358*, 434) and included one of our publications [8] into the list of references of their paper[13].