An Overview of the Method of Positional Scanning Synthetic Combinatorial Libraries

G. MARIAPPAN*, NIHAR BHUYAN, J. P. MOHANTY, SUBARNA GANGULI AND D. DHACHINAMOORTHI
Himalayan Pharmacy Institute, Majhitar, Rangpo, East Sikkim-737 132, India.

Combinatorial chemistry is a novel synthetic strategy which leads to produce a large number of chemical libraries with predetermined structures. The advances in the research of combinatorial library synthesis and screening methods have enabled the medicinal chemists to identify highly active compounds rapidly. This method has revolutionized basic research and drug discovery. A number of combinatorial methods have been developed to boost the morale of medicinal chemists. Out of those, positional scanning synthesis plays a significant role to generate libraries from libraries. This approach is capable of producing infinite libraries. Our aim is to explore this method to understand the principle of synthesis and deconvolution screening methods to identify individual active compounds.

Combinatorial chemistry represents the product of matrix chemistry which represents the product of linear chemistry. It has caused a great cultural change by redefining the scientific methods. For more than a decade, medicinal chemists execute experiments, one at a time, with very careful control of parameters. But combinatorial chemistry changed all the steps of orthodox chemistry. Let us imagine an orthodox chemistry laboratory. Here the medicinal chemist produces one compound at a time by the standard reaction A+B→C. But in combinatorial chemistry laboratory, the combinatorial chemist produces infinite compounds by the reaction A_1 to A_n and reacts with B_1 to B_n. The combinatorial libraries are A_1 B_1, A_2 B_2, A_3 B_3,……. A_n B_n. Combinatorial chemistry started its life with amino acids. Mario Geysen, in the mid-1980, began synthesizing peptides by the hundreds, first in parallel fashion and latter as mixtures. The classic technique, i.e., solid phase organic synthesis (SPOS), developed in the 1960 by R. Bruce Merrifield at Rockefeller University, quickly became the key technique for production of entire range of proteins and enzymes during the past three decades. This technique has become the basis for the exploration of solid phase organic synthesis, whereby molecular diversity can be introduced by producing a nearly infinite variety of heterocycles, steroids and carbohydrates.

***For correspondence**
E-mail: gmariappanhpi@yahoo.co.in

Combinatorial chemistry has many parallels in nature. Our immune system is able to produce hundreds of millions of different antibodies by recombining segments of a variable region and primary structure. Even deadly South Pacific cone snails appear to have been making mixtures for the past 50 million years or so. Bal do mero Olivera of Utah has found that these creatures have the ability to make the mixture of 100 or more deadly venoms, essentially produced by combinatorial scrambling of amino acids.

Characteristics of mixture-based synthetic combinatorial libraries:

Synthetic combinatorial libraries (SCLs) represent systematically arranged mixtures of a large number of synthetic compounds[1,2]. The multiple solid-phase synthesis, also called 'tea-bag approach,' generates the synthetic combinatorial libraries. In this approach[3], compartmentalized resin-bound compounds are synthesized. SCLs are made up of individual defined building blocks at certain positions of the compound scaffold, while the remaining diverse positions are mixtures of building blocks. SCLs are cleaved from a solid support and are assayed in solution. The assay procedure allows each compound within each mixture to freely interact with the given receptor. The first SCLs were composed of peptides of various lengths and amino acids (L, D and unnatural). An approach known as 'libraries from libraries' was used in the development of

SCLs, which is made up of peptidomimetics[4], polyamines[5] and heterocycles[6-9]. The existing peptide SCLs can be transformed chemically to produce new SCLs having different physical, chemical and biological properties with respect to the original peptide SCLs used as starting material. There are two different deconvolution methods that are used to identify individual compounds from mixture-based SCLs. The first method involves identification of active mixtures having defined positions; the remaining mixture positions are then defined, one diversity position at a time, through a synthesis and selection process until individual compounds are identified[10, 11].

The second deconvolution method is known as positional scanning[12-14]. In this case, separate sub-libraries individually address each diversity position. The building blocks of the active mixtures at each diversity position are combined and the resulting individual compounds are synthesized and tested to determine their activities. Iterative deconvolution is not required in case of positional scanning format. In most cases, positional scanning format needs a single synthesis to obtain active individual compounds. This positional scanning method is used not only to prepare peptide libraries but also heterocyclic libraries using dipeptide SCLs as starting materials; examples include hydantoin, cyclic urea, thiourea, indole-pyrido-imidazole and bicyclic guanidine SCLs. At first, this method was successfully adopted to generate heterocyclic libraries, e.g., benzodiazepine library synthesis[15]. Then to generate libraries belonging to groups like hydantoins, antibacterials, antihypertensives, thrombin inhibitors and metalloprotease inhibitors[16], combinatorial approach has been successfully adopted. Recently, this technique was used to design some peptidomimetic inhibitors for Hepatitis C virus NS3 protease[17].

Positional scanning concept:

Positional scanning synthetic combinatorial libraries (PS-SCLs) are composed of positional SCLs or sub-libraries. Here each diversity position is defined with a single building block. The remaining positions are composed of mixtures of building blocks. Each positional sub-library represents the same collection of individual compounds. The positional sub-libraries are assayed and the assay screening data yield information about the most important building block for every diversity position of the PS-SCL. Table 1 represents the tripeptide combinatorial libraries. Here four different amino acids are incorporated at each of the three diversity positions, resulting in a diversity of 64 (4^3) individual peptides. Only 12-peptide mixture (4 amino acids and 3 positions) needs to be synthesized when the same diversity is arranged as a PS-SCL. Each of the three positional sub-libraries, namely, OXX, XOX and XXO, contain the same diversity of peptides; they differ only in the location of the position defined with a single amino acid (Table 2).

Any of the four amino acids is represented by O position, while each of the remaining two positions is a mixture (X) of the same four amino acids. It is shown below that each mixture is the 16 peptides (4^2) that make up that mixture. Here it is assumed that ART is the only tripeptide in these libraries that is recognized by the receptor. It is found that each positional sub-library contains the same diversity of polypeptides; the ART tripeptide is present in all three positional sub-libraries (Table 3). The ART tripeptide is present only in the mixture with activity. The tripeptide ART is obtained by combination of these amino acids in their respective positions. This tripeptide is to be synthesized and tested for activity against the receptor. It is to be noted that the activity observed for each of the three mixtures (AXX, XRX and XXT) is due to the presence of the tripeptide ART within each mixture and not due to the individual amino acids (A, R and T) that occupy the defined position.

Tripeptide PS-SCL synthesis:

A predetermined ratio of amino acids is coupled to the resin-bound amino acid. It is necessary to establish a set

TABLE 1: TRI-PEPTIDE COMBINATORIAL LIBRARY – 64 INDIVIDUAL PEPTIDES

AAA	RAA	TAA	WAA
AAR	RAR	TAR	WAR
AAT	RAT	TAT	WAT
AAW	RAW	TAW	WAW
ARA	RRA	TRA	WRA
ARR	RRR	TRR	WRR
ART	RRT	TRT	WRT
ARW	RRW	TRW	WRW
ATA	RTA	TTA	WTA
ATR	RTR	TTR	WTR
ATT	RTT	TTT	WTT
ATW	RTW	TTW	WTW
AWA	RWA	TWA	WWA
AWR	RWR	TWR	WWR
AWT	RWT	TWT	WWT
AWW	RWW	TWW	WWW

A = Alanine, R = Arginine, T = Tyrosine, W = Tryptophan

TABLE 2: POSITIONAL SCANNING – SYNTHETIC COMBINATORIAL POSSIBLE SUB-LIBRARIES

OXX	XOX	XXO
AXX	XAX	XXA
RXX	XRX	XXR
TXX	XTX	XXT
WXX	XWX	XXW

X = [A – Alanine, R – Arginine, T – Tyrosine, W – Tryptophan], O = Position

TABLE 3: POSITIONAL SCANNING – SYNTHETIC COMBINATORIAL LIBRARIES

OXX				XOX				XXO			
AXX	RXX	TXX	WXX	XAX	XRX	XTX	XWX	XXA	XXR	XXT	XXW
AAA	RAA	TAA	WAA	AAA	ARA	ATA	AWA	AAA	AAR	AAT	AAW
AAR	RAR	TAR	WAR	AAR	ARR	ATR	AWR	ARA	ARR	ART	ARW
AAT	RAT	TAT	WAT	AAT	ART	ATT	AWT	ATA	ATR	ATT	ATW
AAW	RAW	TAW	WAW	AAW	ARW	ATW	AWW	AWA	AWR	AWT	AWW
ARA	RRA	TRA	WRA	RAA	RRA	RTA	RWA	RAA	RAR	RAT	RAW
ARR	RRR	TRR	WRR	RAR	RRR	RTR	RWR	RRA	RRR	RRT	RRW
ART	RRT	TRT	WRT	RAT	RRT	RTT	RWT	RTA	RTR	RTT	RTW
ARW	RRW	TRW	WRW	RAW	RRW	RTW	RWW	RWA	RWR	RWT	RWW
ATA	RTA	TTA	WTA	TAA	TRA	TTA	TWA	TAA	TAR	TAT	TAW
ATR	RTR	TTR	WTR	TAR	TRR	TTR	TWR	TRA	TRR	TRT	TRW
ATT	RTT	TTT	WTT	TAT	TRT	TTT	TWT	TTA	TTR	TTT	TTW
ATW	RTW	TTW	WTW	TAW	TRW	TTW	TWW	TWA	TWR	TWT	TWW
AWA	RWA	TWA	WWA	WAA	WRA	WTA	WWA	WAA	WAR	WAT	WAW
AWR	RWR	TWR	WWR	WAR	WRR	WTR	WWR	WRA	WRR	WRT	WRW
AWT	RWT	TWT	WWT	WAT	WRT	WTT	WWT	WTA	WTR	WTT	WTW
AWW	RWW	TWW	WWW	WAW	WRW	WTW	WWW	WWA	WWR	WWT	WWW

X = [A – Alanine, R – Arginine, T – Tyrosine, W – Tryptophan], O = Position

of ratio for the amino acids that will be used in the coupling step[18]. The method of simultaneous multiple peptide synthesis (SMPS)[19], also known as the 'tea-bag approach,' is advantageous as it provides all wash and deprotection step in a common vessel. It is possible as the resin used for each peptide is enclosed in separate polypropylene mesh bags (tea bags). Individual tea bags are simply placed in the appropriate amino acids solution for the coupling reaction. It is necessary that sufficient solvent covers the tea bags and vigorous shaking be used for each step. For the method described here, methybenzhydrylamine (MBHA) polystyrene resin is used in conjunction with t-Boc chemistry[20].

Protocol for the synthesis of a tripeptide PS-SCL:
Polypropylene tea bags are used to prepare PS-SC libraries. These tea bags are numbered from 1 to 60. To each bag, 100 mg of MBHA polypropylene resin is added and sealed. The first coupling step involves coupling of the 20-proteinogenic amino acids individually to resin-filled bags 41-60. To bags 1-40, 19 amino acid mixtures are coupled by using a predetermined ratio of amino acids. The bags are neutralized before the next process of coupling. Cysteine is excluded to avoid disulfide aggregates in 19 amino acid mixtures. The second coupling step involves coupling of 20 amino acids individually to bags 21-40. Then 19 amino acid mixtures are coupled to bags 1-20 and 41-60. The N-terminal BOC is removed and bags are neutralized for the next coupling process. In the third coupling step, 20 amino acids are coupled individually to bags 1-20. The N-terminal BOC is removed to have an N-terminal amine. Further, the N-terminal amino groups of the peptide mixture can be acetylated if needed.

PS-SCL deconvolution:
In order to identify the most active individual compound from the PS-SCLs, the synthesis and testing of the individual compound that corresponds to the combination of building blocks defined in the most active mixtures at each position are needed. The minimum number of building blocks is to be selected from each position that will be used to synthesize individual compounds. For example, if two amino acids were selected from each position of a hexapeptide PS-SCL, one needs to synthesize 64 peptides; and if three amino acids were selected at each position, 729 would be required. It should be noted that different number of building blocks could be chosen from each position, depending upon the activity of the mixtures.

Since PS-SCLs are composed of separate positional SCLs, each one can be considered independent of the others. Thus each positional SCL can be independently screened and pursued using an iterative synthesis and selection process[21]. Deconvolution of active individual compounds from mixture-based libraries depends on reproducible screening data and clear dose response activities of the most active mixtures. Dose response curve is to be determined for the most active mixtures and the activities based on IC_{50} values are used to select the building blocks that will be included in the synthesis of individual compounds. It has been noted that the building blocks of similar chemical character will yield similar activities at a given position.

One aspect of screening PS-SCL is that the activity of a mixture with a given building block in the defined position can be due to one or more families of

compounds having that same building block at its respective position. When the combination of the most active building blocks are synthesized and tested as individual compounds, it becomes clear whether the activities of the mixtures between positions are connected. Several strategies can be used for deconvolution. The activities of the mixtures of the library are due to the activities of individual compounds and they can be deconvoluted. In order to identify the highly active mixture, the connectivity of the active mixture at each position of diversity is the most important step.

Assay optimization:

PS-SCLs have been used in a number of biological assays. The knowledge of the various assay parameters, such as signal to noise ratio, variability and sensitivity, is required for successful identification of active compounds from libraries. Table 4 illustrates various assays in which PS-SCLs have been used and the signal ranges obtained.

All these parameters are inherent to the assay and are not influenced with respect to whether complex mixtures are being tested as individual compounds. The parameter used as a control for an assay system is the variability. While screening a complex mixture, it is critical that the assay variability is known. For an assay with low variability, one has confidence that a 5-10-fold difference in observed activity between mixtures is significant. For an assay with high variability, the variation between replicates for a given mixture may obscure real differences in activities from other mixtures. The use of repeated experiments and averaged data ensures accurate deconvolution (selective of truly active mixtures), which results in the identification of individual compounds having significant activity.

Limitations:

The major limitations include the following: 1. the chemistry for SPOS is limited; 2. the laboratory should be completely sophisticated to carry out SPOS; 3. side chain de-protection and cleavage is the mandatory last stage; 4. molecules should have polar groups, such as NH_2, OH, to bind with resins; 5. The resins introduce other problems, such as poor homogeneity; swelling, low loading capacity and these factors may hamper reaction rate and site accessibility; 6. High Throughput Positional Scanning (HTPS) is needed to screen these libraries; 7. analytical difficulties are inevitable, i.e., 100 000 NMR is needed to study 100 000 libraries; and finally, 8. physical requirement is inevitable, i.e., to make 10 000 compounds, 10 000 vials or wells are needed.

TABLE 4: BIOLOGICAL ASSAYS WITH VARIOUS PARAMETERS

Assay	Readout	Range	Ratio
Radio receptor	Counts/min	200-1000	5
T cell proliferation	Counts/min	100-500	5
Micro dilution	Optical density	0.1-0.5	5
Enzyme inhibition	Optical density	0.1-0.8	8
ELISA	Optical density	0.1-2.0	20

CONCLUSION

We have dealt with an elegant method, that is, PS-SCL, to generate libraries from libraries. This method is amenable to the production of a huge number of libraries from commercially available building blocks. Nowadays, combinatorial chemistry has become popular in academia and industry, but it is still a relatively young technology whose value remains unproven. The PS-SCL will certainly enrich the efforts of medicinal chemists towards discovery of effective drug candidates for the benefits of mankind.

REFERENCES

1. Houghten, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T. and Cuerro, J.H., **Nature**, 1991, 84, 354.
2. Pinialla, C., Appel, J., Blondelle, S.E., Dooley, C.T., Dorner, B. and Eichler, J., **Biopolymers Peptide Sci.,** 1995, 37, 221.
3. Houghten, R.A., **Proc. Natl. Acad. Sci.**, USA, 1985, 82, 5131.
4. Ostresh, J.M., Husar, G.M., Blondelle, S.E., Dorner, B., Weber, P.A. and Houghten, R.A., **Proc. Natl. Acad**. **Sci.,** USA, 1994, 91, 11138.
5. Abell, A., Dorner, B., Ostresh, J.M., Blondelle, S.E., Dolley, C.T. and Houghten, R.A., Advances in Amino Acid Mimetics and Peptidomimetics, JAI Press, Greenwich, 1998, 109.
6. Nefzi, A., Ostresh, J.M., and Houghten, R.A., **Chem. Rev.,** 1997, 97, 449.
7. Nefzi, A., Ostresh, J.M., Meyer, J.P. and Houghten, R.A., **Tetrahedron Lett.,** 1997, 38, 931.
8. Nefzi, A., Ostresh, J.M. and Houghten, R.A., **Tetrahedron Lett.,** 1997, 38, 4943.
9. Ostresh, J.M., Schoner, C.C., Hamashin, V.T., Meyer, J.P. and Houghten, R.A., **J. Org. Chem.,** 1998, 63, 8622.
10. Dooley, C.T., Chung, N.N., Schiller, P.W. and Houghten, R.A., **Proc. Natl. Acad. Sci.,** USA, 1993, 90, 10811.
11. Dooley, C.T., Chung, N.N., Wilkes, B.C., Schiller, P.W., Bidlack J.M.and Pasternak, G.W., **Sci.,** 1994, 266, 2019.
12. Pinilla, C., Appel, J.R., Blanc, P. and Houghten, R.A., **Biotechniques,** 1992, 13, 901.
13. Dooley, C.T. and Houghten, R.A., **Life Sci.,** 1993, 52, 1509.
14. Pinilla, C., Appel, J.R., Blondelle, S.E., Dooley, C.T., Eichler, J.and Ostresh. J.M., **Drug Dev. Res.,** 1994, 33, 133.
15. Sastry, B.S. and Muralidhar, P., **Indian J. Pharm. Sci.,** 2001, 63, 279.
16. Galande, A.K., **Indian J. Pharm. Sci.,** 2000, 62, 84.
17. Merrifield, R.B., **J. Amer. Chem. Soc.,** 1963, 85, 2149.
18. Ostresh, J.M., Winkle, J.H., Hamashin, V.T. and Houghten, R.A., **Biopolymers,** 1994, 34, 1681.

19. Houghten, R.A., Bray, M.K., De Graw, S.T. and Kirby, C.J., **Int. J. Peptide Protein Res.,** 1986, 27, 673.
20. Stewart, J.M. and Young, J.D., **Solid Phase Peptide Synthesis**, Pierce Chemical Company, Rockford, Illinois, 1984, 212.
21. Bothara, K.G., Gaikwad, A.D. and Saxena, A., **Indian J. Pharm. Sci.**, 1999, 61, 255.

Accepted 2 July 2006
Revised 8 March 2006
Received 18 February 2005
Indian J. Pharm. Sci., 2006, 68 (4): 420-424