Pergamon

Bioorganic & Medicinal Chemistry Letters, Vol. 7, No. 7, pp. 837-842, 1997

0960-894X/97 $17.00 + 0.00

PII: S0960-894X(97)00113-3

SPATIALLY ARRAYED MIXTURE (SPAM) TECHNOLOGY: SYNTHESIS OF TWO-DIMENSIONALLY INDEXED ORTHOGONAL COMBINATORIAL LIBRARIES

Scott C. Berk* and Kevin T. Chapman
Department of Molecular Design and Diversity, Merck Research Laboratories, Rahway, NJ 07065

Abstract: A combinatorial strategy is reported that seeks to maximize information content while maintaining synthesis and screening efficiencies, and that furnishes active compounds with minimal need for synthesis or screening follow-up. The strategy was used to identify a known active from a library of 9,216 compounds.

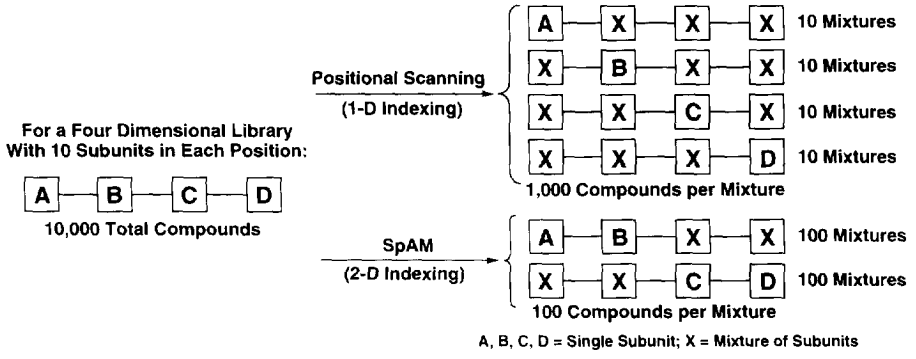
In recent years, combinatorial chemistry has emerged as an extremely useful tool for the discovery and development of new drugs.[1] When utilizing combinatorial techniques, however, there are several important issues to consider, which must be applied to the design of the library. These issues are: (1) the ease and efficiency of the synthetic strategy (the number of compounds in the library versus the number of steps required to produce them); (2) the information content of each testable entity produced by the synthesis strategy; (3) the ease and efficiency of screening the library in one or more biological assays (the number of compounds in the library versus the number of entities to assay). These three factors are intimately linked to one another, and the optimization of one usually leads to a degradation of another. For example, increasing the efficiency of the synthesis by generating mixtures of compounds results in a loss of information because for any particular mixture, it is impossible to know a specific compound's contribution to the overall activity of that mixture. Table 1 compares efficiency and information content of several popular combinatorial strategies.

Table 1. Comparison of Combinatorial Strategies

Strategy	Synthesis Ease/Efficiency	Information Content	Screening Ease/Efficiency
Matrix Synthesis of Single Compounds	Inefficient, Although RF Tracking[2] is Helpful	Complete Information	Inefficient, Especially for Low-Throughput Assays
Mix and Split w/ Encoded Beads	Efficient, But Tagging Chemistry Must be Compatible with Library Synthesis	Information on Decoded Compounds Only	Efficient, But On-Bead Screening Requires Specialized Assays
Mix and Split w/ Iterative Resynthesis	Initially Very Efficient, But Requires Follow-up to Identify Active	Information on Partial Structures, Additional Info. Only from Deconvolutions	Very Efficient, But Requires Follow-Up
Positional Scanning (1-D Indexing)	Efficient, But Requires Isokinetic Mixtures or Multiple Synthetic Routes	Instant Identification of Probable Actives, But Misses Synergistic Combinations of Subunits	Moderately Efficient

We now report an alternative combinatorial strategy that seeks to generate the maximum amount of useful information while still maintaining synthesis and screening efficiencies, and that requires minimal follow-up syntheses and screening to furnish active compounds. The method, which we have named "Spatially Arrayed Mixture" (SpAM) technology, instantly identifies probable actives after an initial screening of the library, much like other indexed methods.[3] However, instead of each mixture addressing one specific subunit in a single dimension, SpAM technology addresses pairs of neighboring subunits (see Figure 1). This two-dimensional mixture indexing provides information about potentially synergistic subunit combinations, where the activity contributed to a molecule by a specific subunit pair is greater than the sum of the activity contributed by each subunit alone.[4] The trade-off for this increased information content is a higher number of testable entities required for screening, albeit at a lower level of mixture complexity.

Figure 1. Comparison of Simple Indexing vs. SpAM (Multi-Dimensional) Indexing



In one typical SpAM experiment, shown schematically in Figure 2, for a four component molecule **ABCD**, a basis set of eight **A**, twelve **B**, eight **C**, and twelve **D** subunits are chosen for a total compound count of 9,216. Each compound can be referred to by a sequence of its constituent subunits in the form $\mathbf{A}_i\mathbf{B}_j\mathbf{C}_k\mathbf{D}_l$. A robotic synthesizer with a 96-well reactor block arranged in an eight by twelve array is then used to add the **A** and **B** portions of the molecule to resin beads such that each unique **AB** pair is associated with a particular reaction vessel. Thus, the reaction vessel in row 2, column 8 contains the molecule $\mathbf{A}_8\mathbf{B}_2$ attached to the resin. The resins in all reaction wells are then combined, thoroughly mixed, and redistributed to the reaction vessels. Each vessel now contains all possible **AB** combinations attached to the resin, but each bead still contains a unique **AB** pair. Next, the **C** and **D** subunits are sequentially attached to the molecules on the resin with the **C** subunits addressed by column and the **D** subunits addressed by row. These 96 mixtures of 96 compounds per mixture make up the **CD** Plate of the library. In a separate run, the **AB** pairs are synthesized as before. Instead of mixing and redistributing the resin, however, each **AB** pair is kept spatially arrayed, and a mixture of all **C** subunits is added to the molecules in each reaction vessel in such a way that each **C** component is equally represented in the resulting mixture. Finally, a mixture of **D** subunits is similarly added to each molecule to complete the solid-phase synthesis. These 96 mixtures of 96 compounds constitute the **AB** Plate of the library.[5] Both the **AB** Plate and the **CD** Plate contain the same 9,216 compounds, but the compounds are distributed in different reactor positions. The compound $\mathbf{A}_8\mathbf{B}_5\mathbf{C}_8\mathbf{D}_1$ appears in row 1, column 8 of the **CD** Plate and row 5, column 8 of the **AB** Plate. Note that no other compound appears in *both* of these positions. Cleaving the compounds from the resin and screening the resulting 192 mixtures identifies the most active **AB** pairs and the most active **CD** pairs. The combination of these pairs produces a set of compounds which encompass the most active compounds in the library. At the same time, the mixture activity data provides useful structure–activity relationship information about each subunit pair.

To validate the SpAM strategy, a library was prepared which contains a known α1-adrenergic receptor agonist.[6] This molecule, which binds to the receptor with a K_i of 5 nM is a tri(N-substituted-glycine), or "peptoid."[7] The relative ease of synthesis and availability of starting materials made peptoid chemistry an ideal candidate for this proof of concept experiment (see Figure 3). Figure 4 shows the **A**, **B**, **C**, and **D** subunits chosen for the library. The **A** subunits are all differentially functionalized bis-electrophiles, and offer a structurally diverse replacement for the bromoacetic acid linkage unit used in standard peptoid chemistry. The **B**, **C**, and **D** subunits are all amines. The **C** subunits were chosen to be similarly reactive in the nucleophilic

displacement of an alkyl bromide to simplify the process of determining the composition of an isokinetic mixture[8] for this set of fragments (see below). Thus, each **C** subunit except for the fragment found in the known active is a primary amine with no α-carbon substituents. The **D** subunits were also selected to react at similar rates. Before inclusion in the basis set, the subunits were "proofed" to ensure compatibility with the synthesis. For each reagent, a single-compound solid-phase synthesis was run under standard conditions in which the test subunit was incorporated into a model compound. If the synthesis was unsuccessful, the subunit was dropped from the library.

Once the basis set was chosen, the remaining challenge of the library development was to determine how to install the **C** and **D** subunits for the **AB** plate. Each **C** and **D** subunit had to be added to the spatially arrayed **AB** pairs such that each molecule in the library would be equally represented. For a linear peptoid synthesis, adding the subunits to the molecule by creating isokinetic **C** and **D** mixtures for the alkylation reactions seemed to be the best choice.[9] To optimize such mixtures, an equimolar mixture of the subunits was reacted with a representative dipeptoid and, following cleavage from the resin, the resulting tripeptoid mixture was subjected to electrospray mass spectrometric analysis. The concentrations of each component of the subunit mixture were then adjusted based on the relative ion intensity of their corresponding product. By reacting this new mixture with the original dipeptoid, it was shown that each component of the product mixture was indeed present at roughly the same concentration. Unfortunately, the subunits required to include the known active did not have similar reactivity to the rest of the subunits ($\mathbf{C_8}$ and $\mathbf{D_{12}}$ reacted much more slowly than the other **C** and **D** fragments, respectively). In order to insure that the known active was present in the library, it was necessary to install these subunits separately, after first determining the optimal concentrations and reaction times required for partial alkylation. By combining the use of isokinetic mixtures for some subunits with this partial alkylation strategy for others, a large diversity of subunits could theoretically be included in SpAM libraries. With these solutions and conditions in hand, the rest of the synthesis was relatively straightforward, and is presented in Scheme 1.

Figure 2. SpAM Technology Schematic

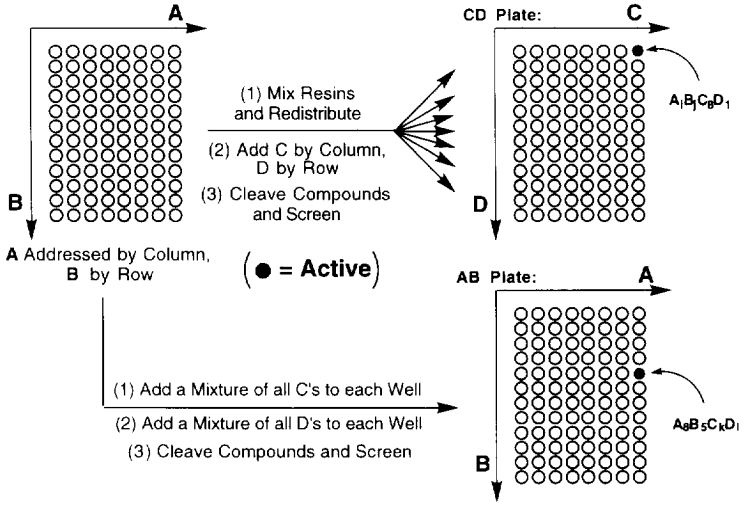


Figure 3. Library Generic Structure and Known Active Compound

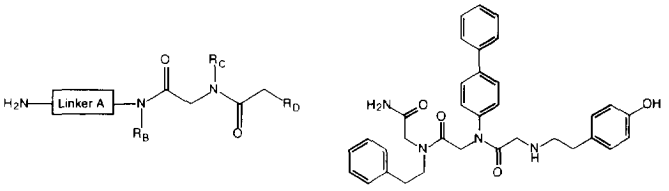
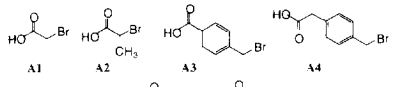
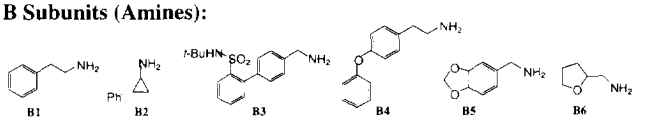


Figure 4. Library Basis Set

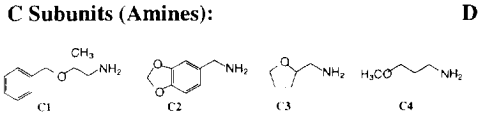
A Subunits (Linkers):



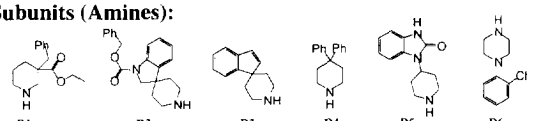
B Subunits (Amines):



C Subunits (Amines):



D Subunits (Amines):



After cleavage from the resin and sample preparation, the 192 library mixtures were assayed for binding to the α1a receptor. The results from each plate of the library are shown in Figure 5. Each bar on the graph corresponds to the relative activity (estimated $1/IC_{50}$ calculated from the % binding to the α1a receptor at 1.25 μM) of a particular library mixture, indexed as shown. Note that the most active mixture in each plate, $\mathbf{A_1B_1}$ and $\mathbf{C_8D_{12}}$, contains the known active. This confirms that the SpAM strategy is capable of immediately identifying the most active compounds in a relatively complex (9,216 compound) library. At the same time, the data provides other information. Comparison of the data for the two plates shows that the receptor seems much more tolerant of changes to the **A** and **B** positions than to the **C** and **D** positions of the molecule. In addition to the $\mathbf{A_1B_1}$ mixture, which contains the known active, several other **AB** mixtures showed similar activity. From the **B** rows in the graph of the **AB** plate, the greatest number of active mixtures contain subunit $\mathbf{B_1}$, phenylethylamine. Several mixtures with $\mathbf{B_8}$, 3,5-dimethoxyphenylethylamine, are also active. This can be contrasted with mixtures containing $\mathbf{B_{10}}$ and $\mathbf{B_{12}}$, both of which are dimethoxyphenylethylamines, but which lead to relatively inactive compounds. Thus, the data indicate that *ortho* methoxy substitution of this subunit is not well tolerated, but *meta* substitution is. Interpretation of the **CD** plate is much more straightforward. The mixture containing the $\mathbf{C_8D_{12}}$ pair is about an order of magnitude more active than any other mixture on the plate. While a one-dimensional indexing method would have clearly identified $\mathbf{C_8}$ and $\mathbf{D_{12}}$ as the most desirable fragments, only the two-dimensional SpAM indexing confirms that it is the $\mathbf{C_8D_{12}}$ pair that is exclusively responsible for the enhanced activity. In a more ambiguous instance, a small subset of potentially active library compounds (formed from all combinations of the most active **AB** and **CD** pairs) would be identified. These compounds could then be synthesized and screened individually to find the actual actives.

Scheme 1

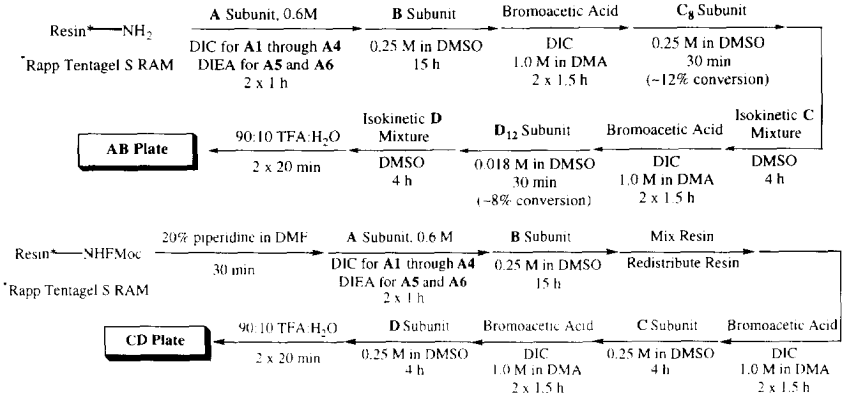
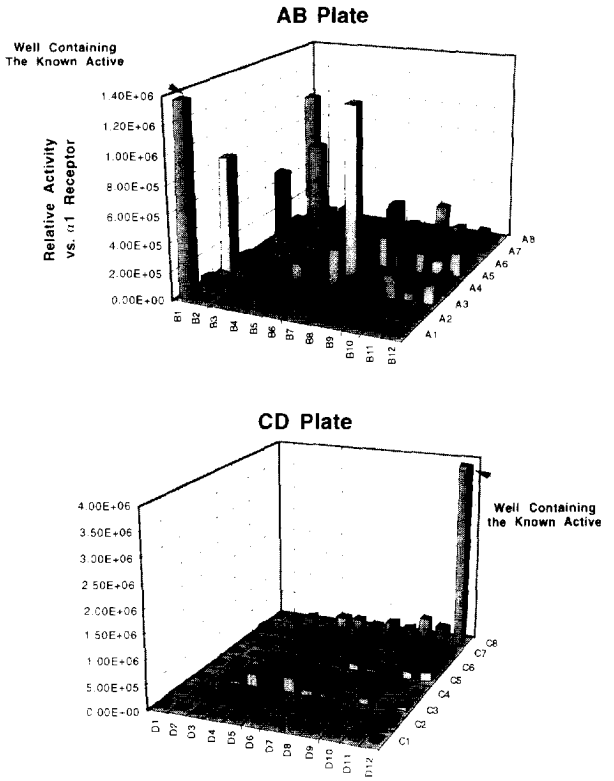


Figure 5. Biological Data for the Library Mixtures



In conclusion, we have developed a new combinatorial strategy which produces two orthogonal sets of mixtures in solution, each indexed in two dimensions. Screening both sets of mixtures identifies a small subset of the library which contains the most active compounds without the need for decoding or iterative resynthesis techniques, and also provides some potentially useful structure–activity information. While SpAM

technology produces more testable entities than one dimensional indexing, the mixtures are made up of fewer compounds. Furthermore, the resulting information not only identifies the contributions of individual subunits towards biological activity, but the interactions of adjacent pairs of subunits in each molecule.

References and Notes

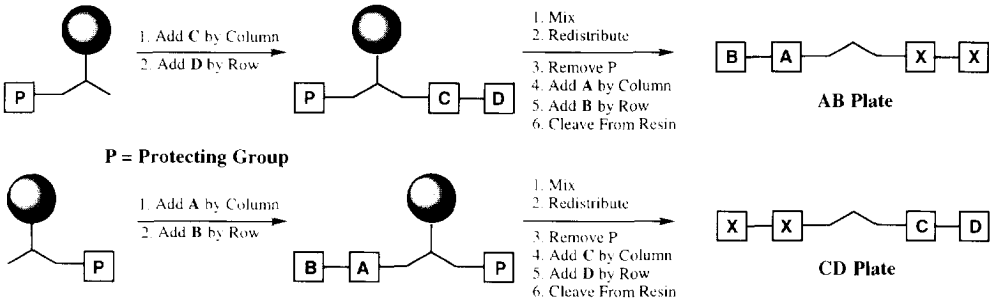
1. For several excellent reviews, see: (a) Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gordon, E. M. *J. Med. Chem.* **1994**, *37*, 1233. (b) Gallop, M. A.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gordon, E.M. *Ibid.* 1385. (c) Thompson, L. A.; Ellman, J. *Chem. Rev.* **1996**, *96*, 555.

2. (a) 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. (b) Nicolaou, K. C.; Xiao, X-Y.; Parandoosh, Z.; Senyei, A.; Nova, M. P. *Angew. Chem., Int. Ed. Engl.* **1995**, *34*, 2289.

3. (a) Dooley, C. T.; Houghton, R. A. *Life Sci.* **1993**, *52*, 1509. (b) Smith, P. W.; Lai, J. Y. Q.; Whittington, A. R.; Cox, B.; Houston, J. G.; Stylli, C. H.; Banks, M. N.; Tiller, P. R. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2821. (c) Pirrung, M. C.; Chen, J. *J. Am. Chem. Soc.* **1995**, *117*, 1240. (c) Pirrung, M. C.; Chau, J. H-L.; Chen, J. *Chemistry & Biology* **1995**, *2*, 621. (d) Déprez, B.; Williard, X.; Bourel, L.; Coste, H.; Hyafil, F.; Tarter, A. *J. Am. Chem. Soc.*, **1995**, *117*, 5405.

4. Two-dimensional indexing has been used to create the first set of mixtures for six-dimensional, iterative resynthesis-based peptide libraries. See: (a) Geysen, H. M.; Mason, T. J. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 397. (b) Houghton, R. A.; Dooley, C. T. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 405.

5. Alternatively, the **AB** and **CD** plates could be produced by two *different* synthetic pathways, such as those shown below, where the order of the installation of each portion of the molecule is exploited to avoid using mixtures of reagents or partial reaction protocols. While this requires considerably more development time to work out the synthetic routes, there are no restrictions on the relative reactivity of the subunits. With the appropriate synthetic sequence, this strategy could also be applied to the synthesis of **AC**, **AD**, **BC** or **BD** plates.



6. Zuckermann, R. N.; Martin, E. J.; Spellmeyer, D. C.; Stauber, G. B.; Shoemaker, K. R.; Kerr, J. M.; Figliozzi, G. M.; Goff, D. A.; Siani, M. A.; Simon, R. J.; Banville, S. C.; Brown, E. G.; Wang, L.; Richter, L. S.; Moos, W. H. *J. Med. Chem.* **1994**, *37*, 2678.

7. (a) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tan, R.; Frankel, A. D.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9367. (b) Zuckermann, R. N.; Kerr, J. M.; Kent, S. B.; Moos, W. H. *J. Am. Chem. Soc.* **1992** *114*, 10646.

8. See, for example: Ostresh, J. M.; Winkle, J. H.; Hamashin, V. T.; Houghten, R. A. *Biopolymers* **1994**, *34*, 1681.

9. As an alternative to isokinetic mixtures, Ellman has successfully employed a limiting amount (0.5 equiv) of an equimolar mixture of reagents: Backes, B. J.; Virgilio, A. A.; Ellman, J. A. *J. Am. Chem. Soc.* **1996**, *118*, 3055. A series of inidividual reactions with limiting amounts of each building block may also be employed.

(*Received in USA* 3 *February* 1997; *accepted* 20 *February* 1997)