

Synthesis and Applications of Small Molecule Libraries

Lorin A. Thompson and Jonathan A. Ellman*

Department of Chemistry, University of California, Berkeley, California 94720

Received August 17, 1995 (Revised Manuscript Received October 18, 1995)

Contents

I. Introduction	555
II. Libraries Synthesized on a Solid Support	556
A. Library Synthesis and Evaluation Strategies	556
1. Discrete Compounds	556
2. Split Synthesis	557
3. Deconvolution of Soluble Libraries	558
4. Structural Determination by Analytical Methods	559
5. Encoding Strategies	560
B. Synthesis of Organic Compound Libraries	561
1. Introduction to Solid Supports	561
2. Post-Synthesis Peptide Modification	562
3. Biopolymer-Mimetic Libraries	562
4. Nonoligomeric Compound Libraries	566
5. Molecular Recognition in Designed Receptor Systems	590
6. Analytical Techniques	591
III. Libraries Synthesized in Solution	592
A. Spatially Separate Synthesis	592
B. Synthesis in Pools	593
1. A Library of Amides and Esters	593
2. Acetylcholinesterase Inhibitors	594
3. Amides Displayed from a Core Molecule	595
4. Oligosaccharide Libraries	595
IV. Future Directions	596
V. Acknowledgments	597
VI. Glossary	597
VII. Bibliography	597



Lorin A. Thompson was born in Lexington, KY, in 1970. He received the Bachelor of Science degree from the University of North Carolina, Chapel Hill, in 1992 where he worked under the guidance of Joseph Desimone. He is currently pursuing his doctorate in the laboratory of Jonathan Ellman at UC Berkeley where he is the 1994 Glaxo-Wellcome fellow. His research interests include the development of synthetic methodology for organic library construction.



Jonathan Ellman was born in California in 1962. He received his S.B. degree from M.I.T. in 1984 where he worked in the laboratory of K. Barry Sharpless. He received his Ph.D. degree with David A. Evans at Harvard University in 1989, where he worked on the synthesis of enantiomerically pure nonproteinogenic amino acids, cyclopeptide alkaloids, and vancomycin. After an NSF postdoctoral fellowship with Peter G. Schultz at the University of California at Berkeley on the incorporation of unnatural amino acids into proteins, he joined the faculty at the University of California at Berkeley in 1992 as an assistant professor. His laboratory is engaged in the development of new chemistry for the synthesis of organic compound libraries, and the application of organic compound libraries to different research problems in chemistry and biology.

I. Introduction

One of the initial steps in the development of therapeutic agents is the identification of lead compounds that bind to the receptor or enzyme target of interest. Many analogs of these lead compounds are then synthesized to define the key recognition elements for maximal activity. In general, many compounds must be evaluated in both the lead identification and optimization steps. Increasing burdens have been placed on these efforts due to the large number of new therapeutic targets that continue to be identified through modern molecular biology methods.¹

To address this demand, very powerful chemical and biological methods have been developed for the generation of large combinatorial libraries of peptides² and oligonucleotides³ that are then screened against a receptor or enzyme to identify high-affinity ligands or potent inhibitors, respectively. While these studies have clearly demonstrated the power of library synthesis and screening strategies, peptides

and oligonucleotides generally have poor oral activities and rapid *in vivo* clearance;⁴ therefore their utility as bioavailable therapeutic agents is often limited. Due to the favorable pharmacokinetic properties of many small organic molecules (<600–700 molecular weight),⁵ the design, synthesis, and evaluation of libraries of these compounds⁶ has rapidly become a major frontier in organic chemistry.

In addition to the development of therapeutic agents, many other applications of organic compound libraries are currently being pursued. These include important advances in molecular recognition, as well as recent efforts in materials and catalysis.

This review focuses on efforts toward the synthesis and evaluation of small organic molecule libraries. Peptide and oligonucleotide libraries will not be described, since a number of very thorough articles have recently reviewed this subject.^{2,3} In addition, greater emphasis will be placed upon library synthesis methods than upon assay results because this *Chemical Reviews* issue is devoted to frontiers in organic synthesis.

II. Libraries Synthesized on a Solid Support

The majority of the compound libraries that have been synthesized to date have been synthesized on a solid support (a solid support is a insoluble material to which compounds are covalently attached during a synthesis sequence). There are two advantages to solid-phase synthesis strategies. First, isolation of support-bound reaction products is accomplished simply by washing away reagents from the support-bound material, and therefore reactions can be driven to completion by the use of excess reagents. Second, innovative methods are available for the manipulation of discrete compounds and for "tracking" the identity of compounds when compounds are attached to a solid support.

A number of general strategies have been developed for the synthesis and evaluation of compound libraries synthesized on solid supports. Although most of these strategies were initially demonstrated with peptide libraries, many of these approaches have now been applied to other compound classes. Many of these strategies have been reviewed recently,² and therefore only an overview of these methods will be provided with an emphasis on recent advances.

A. Library Synthesis and Evaluation Strategies

1. Discrete Compounds

a. Spatially Separate Synthesis. Conceptually, the most straightforward method for the preparation of a compound library is to synthesize many compounds in parallel and to keep each compound in a separate reaction vessel. When the final compounds are spatially separate, the identity of a compound at a particular location is known and can be confirmed by analytical methods. In addition, biological evaluation of the library can provide specific information about each compound in the library. Of course, the compounds can still be pooled (a pool refers to a mixture of compounds) when the assay of interest is not sufficiently high-throughput. Recently reported pooling strategies may be particularly useful for this purpose.⁷

A number of approaches for the parallel synthesis of organic compounds have been reported.⁸ The first method was originally developed by Geysen for peptide epitope mapping.⁹ Since that time, several improvements have been introduced.¹⁰ In this method,

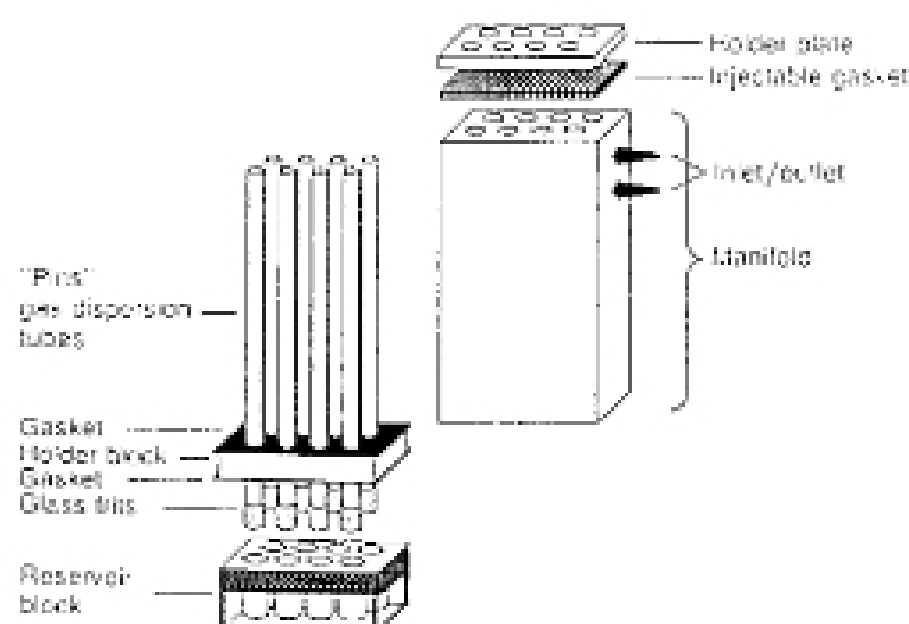


Figure 1. Diversomer™ apparatus. (Reprinted from ref 12a. Copyright the National Academy of Sciences of the United States of America.)

96 polyethylene pins are placed into a supporting block so that each pin fits into a separate well of a 96-well microtiter plate. Each pin is coated with a polymeric material that is amenable to solid-phase synthesis such as poly *N,N*-dimethylacrylamide. The polymeric material is derivatized with aminoalkyl groups to provide sites for substrate attachment. During a synthesis sequence each pin is placed in a separate well of the microtiter plate so that each well serves as a distinct reaction vessel. With this approach, on the order of 10 000 spatially separate compounds have been prepared¹¹ in parallel by employing inexpensive labware, instrumentation, and automation developed predominantly for high-throughput microtiter-based screening efforts. Currently, pin loading levels range from 100 nmol to 50 μ mol of material per pin. Even 100 nmol of material is sufficient for multiple biological assays, as well as for analytical evaluation of the purity and chemical integrity of the individual compounds.

DeWitt and co-workers have reported the Diversomer™ apparatus, which is one of the first reaction apparatus designed expressly for the parallel synthesis of small organic molecules (Figure 1).¹² This apparatus is based on the use of porous gas dispersion tubes that serve as containers for resin beads. Reagents and solvents are placed in up to 40 vials that are located in a reservoir block, and the ends of the gas dispersion tubes containing resin are placed into the vials allowing the reagents to diffuse into the tubes and contact the support. The temperature of the reaction solutions can be controlled by heating and cooling the block. The apparatus is also enclosed in a manifold with an injectable gasket employed for reagent and solvent additions so that reactions can be maintained under an inert atmosphere.

Meyers and co-workers have recently reported a conceptually related approach (Figure 2).¹³ The reaction apparatus is prepared by drilling a hole into each well of a deep-well polypropylene microtiter plate. Porous frits are then placed at the bottom of each well. The solid support and reaction solutions are placed into the wells of the apparatus. During a reaction step, the plate is clamped against a Viton gasket in order to seal the hole at the bottom of the well. In between steps in the synthesis sequence, the reaction solution can be drained and the resin can

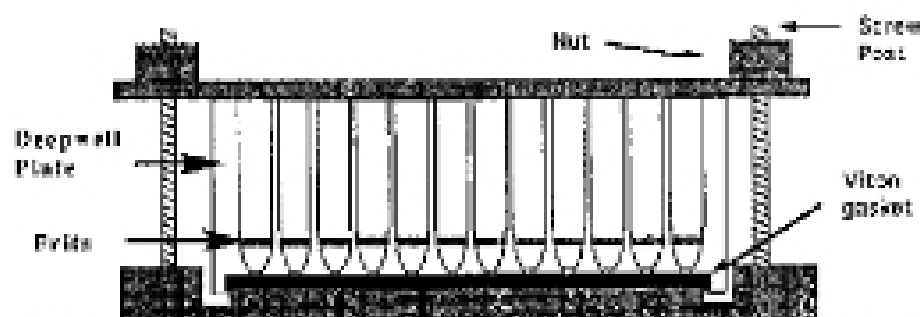


Figure 2. Multiple synthesis apparatus of Meyers and co-workers. (Reprinted from ref 13. Copyright ESCOM Scientific publishers.)

then be rinsed by removal of the Viton gasket.

The reaction apparatus employed in the three strategies described above are relatively inexpensive. This can be an important consideration, since the parallel synthesis of hundreds to thousands of spatially separate compounds will generally require that multiple reaction apparatus be used simultaneously. Clearly, when compounds are synthesized in pools, a single parallel synthesis apparatus is usually sufficient for the synthesis of large numbers of compounds (*vide infra*).

Commercial apparatus for parallel organic synthesis have also recently been developed. As one example, Advanced Chemtech has developed the model 496 multiple organic synthesizer instrument for automated chemical synthesis.¹⁴ The instrument is designed to produce 96 different compounds. A range of temperatures, mix times, and speeds can be employed. The instrument is also compatible with a wide range of reaction conditions. Finally, reactions can be performed under an inert atmosphere.

b. Light-Directed, Spatially Addressable Parallel Chemical Synthesis. While libraries of greater than 10 000 compounds based upon a single structural type have been prepared by parallel synthesis using separate reaction vessels, the synthesis of libraries of hundreds of thousands to millions of compounds per structural type would not be practical by this approach. Fodor and co-workers have developed a strategy based upon photolithographic methods that can be used to synthesize libraries containing more than 100 000 spatially separate compounds.¹⁵ In this method a silica wafer (borosilicate glass microscope slide) serves as the solid support (Figure 3). Aminoalkyl groups or other reactive functionalities that are attached to the surface of the wafer are blocked with photolabile protecting groups. The photolabile protecting groups are cleaved at specific regions on the silica wafer by site-specific illumination using masks and instrumentation initially developed for computer microchip construction. The silica wafer is then exposed to a reaction solution with reactions occurring only at the regions that were deprotected by illumination. Only a single photolabile group is necessary for the preparation of oligomeric compounds, since a monomer that is protected with the same photolabile group is introduced at each step in oligomer synthesis. However, orthogonal photolabile protecting groups that are cleaved at different wavelengths of light would likely be required in order to differentiate between different sites in nonoligomeric structures. The structure of a final compound at a specific location on the silica wafer is dependent upon the masking scheme and

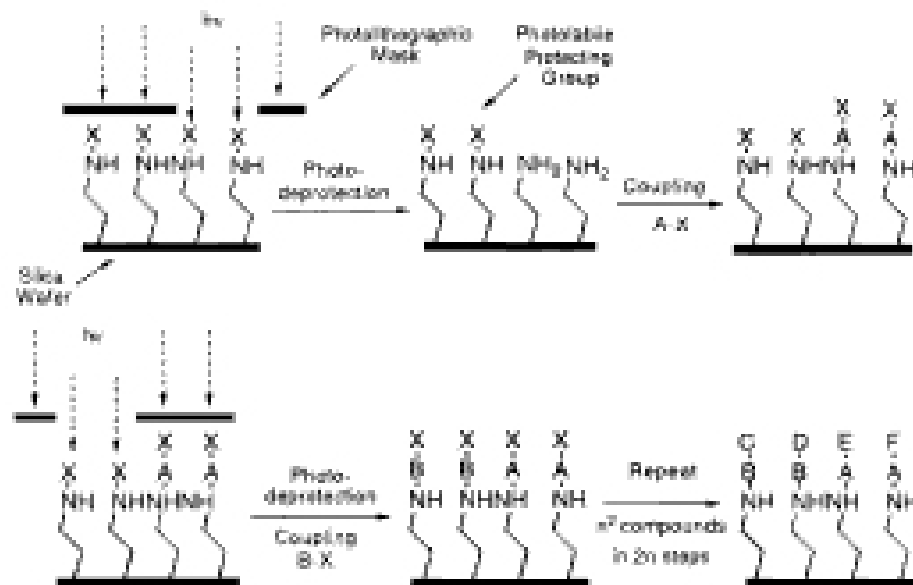


Figure 3. Light-directed, spatially addressable chemical synthesis.

upon the order of addition of the reagents. Large numbers of spatially addressable compounds can be prepared due to the exceptional spatial resolution of photolithography techniques. For example, Fodor and co-workers have assembled peptide libraries with each peptide localized to a $50 \mu\text{m}$ square.^{15c} This density allows the synthesis of 40 000 oligomers per centimeter squared, and the current limits of photolithography should allow synthesis to be performed at much higher densities.¹⁶

In this spatially addressable array, the location of the compound on the silica wafer provides the structure of the compound. The compounds in a library therefore must be assayed while still tethered to the wafer. Although this requirement imposes constraints on the range of feasible biological assays, successful strategies have been developed for this purpose. One strategy is to measure the percent binding of a soluble fluorescently labeled receptor to different locations on the silica wafer as monitored by epifluorescence microscopy. The silica support or the linker that serves to attach the compound to the support can, however, have a pronounced effect on the binding affinity.^{17,18a}

2 Split Synthesis

In general, large libraries of compounds are synthesized by employing pooling strategies. The most direct method is to employ equimolar mixtures of reactants in each synthesis step. In fact, Geysen employed this strategy in one of the early peptide library synthesis efforts.¹⁸ One problem with this approach is that equimolar quantities of the final compounds in the library will only be obtained if all of the reactants in the mixture have comparable reactivity, and for most reaction classes, reactivity is highly dependent upon the structure of the reactants. For the purpose of synthesizing peptide libraries, several researchers have demonstrated that modest differences in reactivity can be corrected by adjusting the relative concentrations of the activated amino acids in the mixture.¹⁹ It is also possible to synthesize approximately equimolar mixtures of products by using a total of 1 equiv of a mixture of reagents in a coupling step.²⁰ This strategy, however, requires that the concentration of each reagent in the mixture is fairly low and therefore is restricted to reactions that are very efficient.