Grabbing Golden

Medicinal chemists may soon reap the benefits of solid-phase split-and-pool combichem techniques.

BY VIKTOR KRCHNAK AND COLIN DALTON

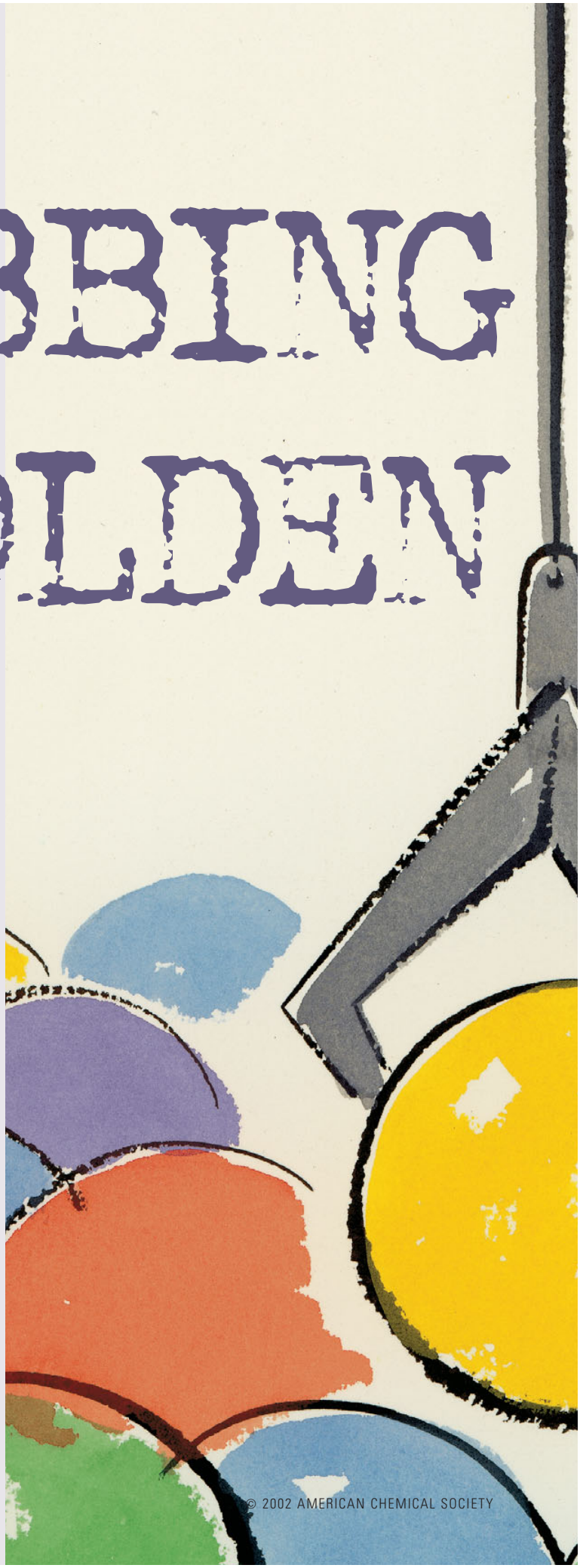
The changing role of combinatorial chemistry in drug discovery is reflected in the way library synthesis is conducted. Combinatorial chemistry has moved from centrally located laboratories, often equipped with multimillion-dollar robots, to the benches of individual medicinal chemists. Because of this change in chemistry requirements, simple instruments operated by the individual chemists are needed as well as a change in direction from inefficient parallel synthesis to the more efficient split-and-pool synthesis. The directed split-and-pool approach for solid-phase synthesis of combinatorial libraries is gaining popularity because of the availability of personal chemist tools that are simple to operate and reasonably priced.

In the beginning

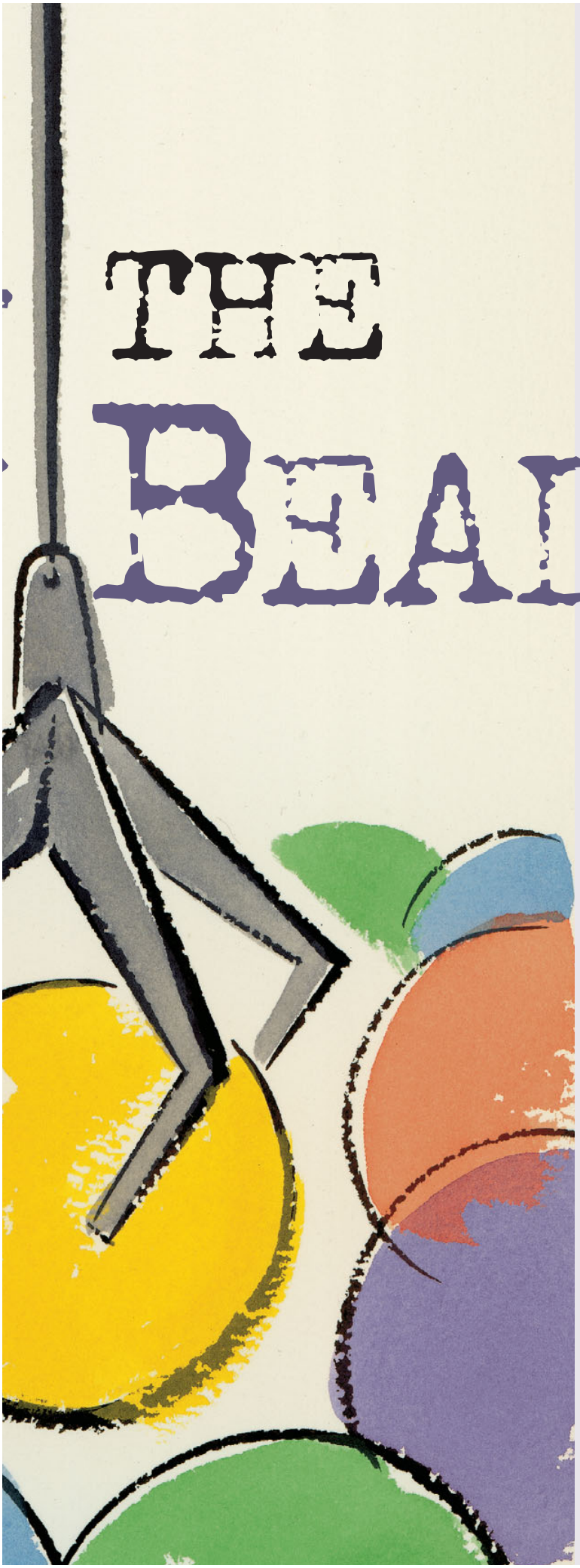
In 1963, Nobel laureate R. Bruce Merrifield introduced the concept of solid-phase synthesis and described the synthesis of a tetrapeptide. Merrifield covalently linked the growing peptide chain to an insoluble resin bead support and elongated the peptide stepwise.

Ronald Frank at the German National Research Center for Biotechnology (Braunschweig) was the first scientist to recognize the potential of combining several different solid-phase-

ILLUSTRATION: NISHAN AKGULIAN



THE BEAD



bound substrates for a reaction with a single reagent (*1*). Any number of solid-support-bound substrates can be combined (pooled) together into one reaction vessel. The individual substrates are physically separated by covalent attachment to the carrier. Frank used paper disks as solid supports for the synthesis of oligonucleotides, where each disk contained only one nucleotide substrate used in building the chain. Four reaction vessels were required for the synthesis of various oligonucleotides, and during each reaction step, the paper disks that received the same reagent were manually distributed into one reaction vessel.

Later, Richard Houghten of the Scripps Research Institute (La Jolla, CA) applied the method of combining different support-bound substrates to solid-phase peptide synthesis on resin beads. Because one bead is too small (typically 100 µm) to yield the desired micromolar amounts of material, Houghten placed resin beads into meshed polypropylene packets similar in appearance to tea (T-) bags (*2*). Before the addition of the next amino acid to the resin-bound growing peptide chain, the T-bags were distributed into different reaction vessels, each vessel containing intermediates receiving the same amino acid. This synthetic concept has the following characteristics:

- Several reaction vessels are required. The number of reaction vessels is the same as the number of building blocks in any particular reaction step. The number of compounds synthesized exponentially exceeds the number of vessels.
- The amount of compound synthesized depends on the yield from each solid-phase particle or container.
- The chemical history of the particles is recorded (i.e., each paper disk or T-bag is labeled by an alphanumeric code).
- The chemist controls the distribution of compounds in a library. Any combination of building blocks can be made as many times as required, and any combination can also be excluded from the synthesis.

Random split-and-pool method

The concept of simplifying combinatorial chemical synthesis (by reducing the number of reaction vessels) was further developed by using the random split-and-pool methods. These processes were independently invented by Arpad Furka at Eötvös Loránd University (Budapest) (*3*) and Kit Lam at the Arizona Cancer Center (Tucson) for one-bead–one-compound peptide libraries (*4*) and by Richard Houghten for iterative libraries (*5*). In the latter case, resin beads receiving the same amino acid are contained in one reaction vessel, as in Frank's method; however, the beads are pooled and split randomly before each combinatorial step.

The random split-and-pool synthesis is illustrated in Figure 1, which shows an example of the combinatorial synthesis of peptides from amino acids. A sufficient quantity of resin beads is split into 20 equal portions in separate reaction vessels (only five portions are shown). One amino acid is coupled to each resin portion, and then the resin beads from all 20 reaction vessels are combined, thoroughly mixed, and split again into 20 portions. At this stage, each portion contains a mixture of 20 resin-bound amino acids. One amino acid is then coupled to each resin portion, creating 20 different resin-bound dipeptides in each reaction vessel (the second combinatorial step). The resin beads are combined, creating a mixture of 400 dipeptides. The same process is then repeated *x* times, where *x* is the length of the target peptides (only three combinatorial steps are shown).

Random split-and-pool is differentiated from Frank's directed split-and-pool method by two major features:

- The chemical history of the particles is lost. After each combinatorial step, the resin beads from all reaction vessels are pooled and randomly split for the next combinatorial step.
- The distribution of compounds in a library is driven by statistics. Depending on the number of beads and the number of compounds, each compound is synthesized numerous times when the number of beads exceeds several times the number of compounds, or only subsets of compounds are produced when the number of beads is lower than the number of possible combinations of building blocks.

To determine the chemical structure of the synthesized compounds, a coding principle has been introduced because the amount of material on one 100-µm bead—some hundreds of picomoles—is insufficient to allow structure determination using classical analytical methods.

A novel solution to this problem was borrowed from the genetic code, in which the sequence of amino acids in proteins is encoded by a sequence of bases in DNA. The idea of nucleic acid coding was applied to random split-and-pool combinatorial synthesis such that each combinatorial chemical transformation was followed by an independent step that introduced an identification code. Thus, two compounds were synthesized on each bead (Figure 2), one containing the target structure, the other the coding structure (coding tag).

However, the application of nucleic acids for coding organic libraries is limited by the incompatibility of nucleic acids with other organic reactions. Other coding strategies using different molecular tags were then developed, including the use of peptide tags, electrophoric tags (halocarbon molecules determined after silylation by gas chromatography), amine tags that can be decoded using high-resolution mass spectrometry, fluorine-containing tags for detection by ^{19}F NMR, and resins that are distinguished by infrared or Raman spectroscopic bar codes.

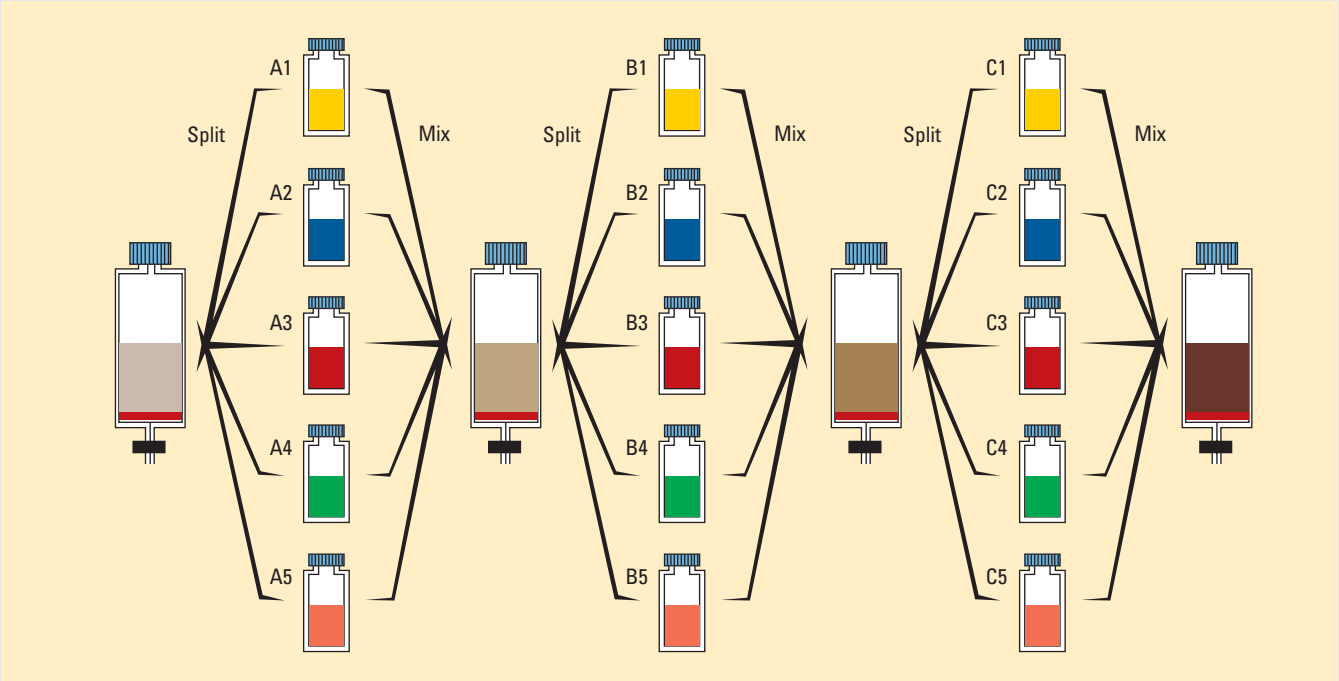


Figure 1. Split-and-pool synthesis. The combinatorial method can be performed to generate products in either a random or directed mode. In the random mode, pooled solid-phase particles are randomly split into reaction vessels. In the directed mode, each solid-phase particle is handled on an individual basis and directed into its corresponding reaction vessel.

Directed split-and-pool

For the directed split-and-pool method to attain more widespread acceptance, several technical issues must be solved. The first is recording the chemical history of individual compounds. This problem has been addressed by several methods, including labeling paper disks and T-bags with an alphanumeric code readable by a chemist. Reaction containers have also been color-coded, and the cut-and-combine method uses a planar solid-phase support that is divided into smaller pieces after each combinatorial step and labeled by a laser printer. Coding by shape of solid-phase particles has also been reported.

Radio-frequency tagging and optical encoding of containers have enabled computer-assisted reading of the tag and automation of the directed split-and-pool process. A radio-frequency tag is inserted into each container with resin beads, and before a combinatorial step, the individual tags are read and the containers distributed into corresponding reaction vessels. The process is called directed sorting and was fully automated and commercialized by IRORI (*www.irori.com*), a subsidiary of Discovery Partners International (San Diego). The Mimotopes SynPhase Lanterns (Clayton, Australia, *www.mimotopes.com*) have been used similarly as radio-frequency-tagged containers.

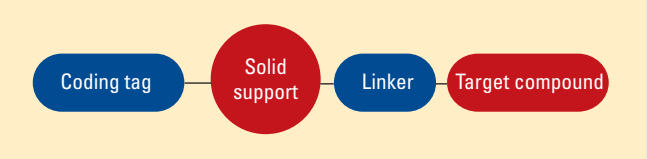


Figure 2. Chemical encoding. To identify target compounds upon combinatorial completion, researchers often resort to chemical tags

Viktor Krchnak has developed a spatially addressable directed split-and-pool procedure for tracking the chemical history of a synthesis carried out on SynPhase crowns, referred to as necklace coding (*6*). Individual solid-phase supports (crowns or lanterns, or resin plugs) are strung on a Teflon thread, and the position of a particle on the necklace encodes the previous chemical history.

Another challenge to acceptance of the directed split-and-pool method is that the solid-phase support must be formulated so that one unit provides the required amount of compound. This problem has been addressed in two ways. Resin beads have been packaged into containers made of meshed polypropylene (T-bags), discussed above, which IRORI changed into a can (MicroKan, NanoKan) to facilitate robotic handling. Recently, Mark Bradley at the University of Southampton (U.K.) introduced resin plugs, which have been commercialized by Polymer Laboratories (Amherst, MA, *www.polymerlabs.com*). To form the plugs, resin beads are mixed with fine-powdered high-density polyethylene, which is then melted, and after cooling, the plugs are used as a solid-phase support (*7*).

The second solution is to produce one particle that yields the required amount of material, as with SynPhase Lanterns, a polypropylene mold with a grafted layer of polystyrene. Grafted macroporous polymer monolithic disks have also been introduced, and paper as a modular solid support has already been mentioned.

However, the chemical incompatibility of paper with many chemical transformations has prompted the development of inert planar membranes.

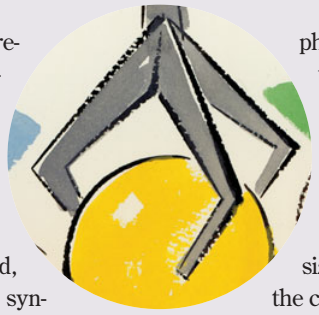
Finally, the process of distributing solid-phase particles, particularly for sizable libraries, needs to be integrated. During the early years of the directed split-and-pool method, the entire process was done manually. Although this is still useful for small libraries, the manual process is tedious and error-prone. Highly automated and sophisticated robots have been designed to execute the directed split-and-pool combinatorial synthesis. IRORI's fully automated directed-sorting robot is based on optical encoding of NanoKans. Similarly, scientists at Selectide (Tucson, AZ, *www.selectide.com*), a subsidiary of Aventis Pharmaceuticals, designed a robot system based on a sophisticated, somewhat complex, two- and three-dimensional spatial-encoded directed split-and-pool method. An intermediate level of sophistication is achieved by radio-frequency tagging of containers or lanterns coupled with a reader connected to a computer that informs the operator where to put the MicroKan. A sorter is also available that automatically distributes MicroKans into reaction vessels.

The distribution of compounds in a library is driven by statistics.

A user-friendly instrument that enables individual chemists to perform directed split-and-pool combinatorial chemistry is the Encore Synthesizer from Encore International Corp. (Tucson, AZ, *www.eico.cc*). The synthesizer is designed as a personal chemistry tool that integrates the process of one-dimensional spatial necklace encoding. Necklace coding with separated individual sequences is the basis of the Encore technique, which stands for *E*ncoding by a *N*ecklace, *Co*lor, and *Re*action vessel. The technique combines three different coding methods: sequential position on a necklace for the first combinatorial step, color coding of individual necklaces for the second combinatorial step, and reaction vessel coding for identifying the last building block. The Encore system was developed to use SynPhase Lanterns, but it can be adapted to other solid supports, including the resin plugs.

Split-and-pool synthesis, as shown in Figure 1, not only simplifies the combinatorial synthetic process, but also offers other important benefits. To undertake a full range of solid-phase chemical reactions, elaborate reaction conditions are needed for some chemical transformations, including, but not limited to, low temperature and inert atmosphere. Parallel synthesis of 1000 compounds requires handling 1000 reaction vessels. To add sensitive reagents (e.g., butyl lithium) in a timely manner at low temperature (–78 °C) under an inert atmosphere during parallel synthesis is not

trivial and requires automated synthesizers. It is unrealistic to perform this synthesis in a manually operated reactor unless the synthesis is split into small batches of compounds. However, such a synthesis can be performed easily using a split-and-pool method that requires only 10 reaction vessels and manual handling.



Now that the technical issues have been solved, we think it is fair to say that directed split-and-pool synthesis might be the preferred approach for constructing compound libraries containing compounds with highly desirable biological properties.

The next step?

Harvard University's Stuart Schreiber recently introduced the new trend of diversity-oriented synthesis of small complex molecules in designing combinatorial libraries for drug discovery (*8*, *9*). Diversity-oriented synthesis, in contrast to target-oriented synthesis, is not focused on one target but rather on disrupting or interfering with complex biochemical pathways. According to Schreiber, the diversity and complexity of libraries will play a critical role in future drug discovery.

Collecting diverse compounds is more likely to be successful in phenotypic screens involving cells or microorganisms than in a collection of related compounds. Compound structural complexity is important because many biological processes depend critically on protein–protein interactions, and many small molecules known to disrupt the interaction are complex natural products. Because of the potentially larger size of diversity-oriented combinatorial libraries and the complexity of chemical transformations, both random and directed split-and-pool approaches will be required.

References

(1) Frank, R.; et al. *Nucleic Acids Res.* **1983,** *11,* 4365–4377.
(2) Houghten, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1985,** *82,* 5131–5135.
(3) Furka, A.; et al. *Int. J. Pept. Protein Res.* **1991,** *37,* 487–493.
(4) Lam, K. S.; et al. *Nature* **1991,** *354,* 82–84.
(5) Houghten, R. A.; et al. *Nature* **1991,** *354,* 84–86.
(6) Smith, J.; et al. *J. Comb. Chem.* **1999,** *1,* 368–370.
(7) Atrash, B.; et al. *Angew. Chem., Int. Ed.* **2001,** *40,* 938–941.
(8) Schreiber, S. L. *Science* **2000,** *287,* 1964–1969.
(9) Schreiber, S. L.; Nicolaou, K. C.; Davies, K. *Chem. Biol.* **2002,** *9,* 1–2.

Viktor Krchnak and **Colin Dalton** have been associated with many of the developments in combinatorial chemistry and are the founders of Encore International Corp. Send your comments or questions about this article to mdd@acs.org or to the Editorial Office address on page 3. ■

Advertisers