40. Michal Lebl,[1] Viktor Krchňák,[1] Nikolai F. Sepetov,[1] Petr Kočš,[1] Marcel Pátek,[1] Zuzana Flegelová,[1] Ronald Ferguson,[1] and Kit S. Lam.[2] **Synthetic Combinatorial Libraries—A New Tool for Drug Design: Methods for Identifying the Composition of Compounds from Peptide and/or Nonpeptide Libraries.** ([1]Selectide Corporation, Oro Valley, Arizona 85737; [2]Arizona Cancer Center and Department of Medicine, University of Arizona College of Medicine, Tucson, Arizona 85724)

J. Protein Chem. 13, 489 (94)

MPSA Short Communications

Development of new leads for drug design and structure/function relationship studies was revolutionized by the introduction of combinatorial or "library" techniques (for review see, e.g., Moos *et al.*, 1993). These techniques allow the generation and screening of millions of potentially active structures. Due to the well-developed and fine-tuned synthetic methodology, peptides were the first group of compounds evaluated by this new approach. However, the challenge is to synthesize libraries of nonpeptidic structures. The combinatorial library approach applied at Selectide consists of three basic steps: (i) chemical synthesis based on the split synthesis method yielding a library with one test compound structure per bead; (ii) screening of the library either using an on-bead binding assay or a multiple-step release assay; and (iii) recovery of positive beads and the determination of the structure of the test compound (Lam *et al.*, 1991).

Each chemically synthesized combinatorial library represents a certain structural diversity and multiplicity. Libraries containing sequential repetition of amino acids (peptide libraries) are easy to synthesize and the structure of compounds of interest can be easily determined by sequencing. However, such libraries do not contain very high structural diversity, since the only variable parameter is the type of side chain connected to the C-alpha carbon of the peptide backbone, and those side cahins occupy only limited conformational space. Combining natural L amino acids with D amino acids brings more diversity; nevertheless, it is still quite limited.

The advent of nonpeptide libraries increased the diversity of conformational space filled by the test compound subunits, as well as increased chemical diversity due to the nature of the subunits (Simon *et al.*, 1992; Bunin and Ellman, 1992; Nikolaiev *et al.*, 1993; Cho *et al.* 1993; Lebl *et al.*, 1994). Combinatorial libraries of chemically synthesized compounds can be classified into several distinct groups. In the classification used here libraries from individual groups represent certain structural types: (i) libraries of small, compact, and relatively rigid structures (e.g., N-acyl-N-alkyl amino acids); (ii) libraries based on a more or less rigid scaffold structure (usually multifunctional cyclic scaffold, e.g., derivatized cyclopentane or cyclohexane ring, functionalized steroid skeleton, tricarboxybenzene, diaminobenzoic acid); (iii) libraries based on a flexible scaffold that is built during the synthesis of the library and can be randomized

(branched scaffold based on diamino acids, α, β, γ, δ-library); (iv) libraries of linear, sequential compounds (a typical example is a peptide library, including also N-substituted glycines—peptoids, or α-, β-, and γ-amino acid-containing library); (v) libraries of organic molecules (e.g., benzodiazepine type).

Once the bead of interest is selected by the screening protocol, it is necessary to determine the structure of the test compound responsible for the desired effect. We apply, similarly as others, a coding principle (Brenner and Lerner, 1992; Nikolaiev *et al.*, 1993; Kerr *et al.*, 1993; Ohlmeyer *et al.*, 1993). Nature has used nucleic acid to code for amino acids for ages. Each chemical individuum in the synthetic library is independently coded by another structure (peptide) whose composition can be easily resolved using an established technique (Edman degradation). We have developed a so-called binary (bar) coding technique in which each chemical reaction performed to synthesize the test compound is coded by two amino acids. If a different set of amino acids is used for coding each reaction step, then the coding arm can be constructed in such a way that one cycle of Edman degradation will cleave all coding amino acids and a single HPLC run will reveal all components.

The coding principle bears one inherent complication. If the screening process is being performed on the bead, the coding structure can interact with the target molecule. There are three possibilities to prevent the interaction of the coding structure with the target: (i) The coding structure can be present in a very low concentration so that the interaction with the target molecule will not be seen under the conditions of the experiment; (ii) the coding and test structures can be physically separated; (iii) the test structure can be coded by a multiplicity of coding structures. The first possibility is not realistic in the case of peptide coding due to the limited sensitivity of peptide sequencing. However, it can be used advantageously in the cases of coding by nucleic acids, where the coding structure can be conveniently amplified (Needels *et al.*, 1993). The second option was explored by us recently (Vagner *et al.*, 1994). Separation of the "surface" available for interaction of the macromolecule target with the test structure and the "interior" of the bead inaccessible to the target can be achieved by enzymatic "shaving." The last possibility is based on the idea of coding using a set of different structures rather than one unique structure. This set of structures must provide unambiguous information about the chemistry performed on screening arm.

References

Brenner, S., and Lerner, R. A. (1992). *Proc. Natl. Acad. Sci. USA* **89** 5381–5383.

Bunin, B. A., and Ellman, J. A. (1992). *J. Am. Chem. Soc.* **114**, 10997–10998.

Cho, C. Y., *et al.* (1993). *Science* **261**, 1303–1305.

Kerr, J. M., Banville, S. C. and Zuckermann, R. N. (1993). *J. Am. Chem. Soc.* **115**, 2529–2531.

Lam, K., *et al.* (1991). *Nature* **354**, 82–84.

Lebl, M., *et al.* (1994). In Crabb, J. W. (ed.), *Techniques in Protein Chemistry* (Academic Press, Orlando, Florida), pp. 541–548.

Moos, W. H., *et al.* (1993). *Annu. Rep. Med. Chem.* **28**, 315–324.

Needels, M. C., *et al.* (1993). *Proc. Natl. Acad. Sci. USA* **90**, 10700–10704.

Nikolaiev, V., *et al.* (1993). *Peptide Res.* **6**, 161–170.

Ohlmeyer, M. H., *et al.* (1993). *Proc. Natl. Acad. Sci. USA* **90**, 10922–10926.

Simon, R., *et al.* (1992). *Proc. Natl. Acad. Sci. USA* **89**, 9367–9371.

Vagner, J., *et al.* (1994). In Epton, R. (ed.), *Innovation and Perspectives in Solid Phase Synthesis*, (in press).