Library of libraries: A novel approach in synthetic combinatorial libraries

Nikolai F. Sepetov[a], Viktor Krchňák[a], Magda Staňková[a], Kit S. Lam[b], Shelly Wade[a] and Michal Lebl[a]

[a]*Selectide Corporation, 1580 E. Hanley Blvd., Tucson, AZ 85737, U.S.A.*

[b]*Arizona Cancer Center and Department of Medicine, University of Arizona, College of Medicine, Tucson, AZ 85724, U.S.A.*

Introduction

The goal of screening peptide libraries for pharmacological purposes is not necessarily to find the most active peptide, but to identify structural features of the peptides critical for biological activity i.e. to find a motif which can be used subsequently for developing a drug. Therefore a *library of motifs* can give as much or more useful information as a *library of peptides*.

Results and Discussion

There are two types of library-based screening for new pharmacologically important leads: parallel and iterative (for the review of library techniques see e.g. [1,2]). The parallel approach is based on the principle of generating the complete, or as complete as possible, multiplicity of unique structures in a single library, which is then screened to identify the solid phase supports on which active compounds were synthesized. The compositions of the positive test compounds are then determined from information on the solid phase support. Since the number of unique compounds which must be synthesized to obtain a complete library increases exponentially, synthesis of complete libraries longer than pentapeptide is becoming impractical. This is the main disadvantage of the parallel approach. On the other hand it is accepted that in a long peptide not all amino acid residues are equally important for activity and only certain residues at certain positions (motif) are critical. Consequently it is not necessary to synthesize all possible structures, but instead to synthesize only compounds containing all possible motifs but not all structures sharing the same motif.

For a particular target it is impossible to predict *a priori* how many residues are important for activity and how many residues must be defined in order to observe activity in a particular screening assay. In the iterative approach to the combinatorial libraries a number of libraries are synthesized, each initially containing one or two sequencial positions which are defined, with the remaining positions completely randomized. The most active library in a particular biological test is selected, beginning an iterative process of synthesis and screening of additional libraries until all portions of the active molecule are defined. In other words, the first step of the iterative process consists of screening libraries with very short motifs containing only one or two sequencial residues. The "Library of Libraries" (library of motifs) approach is a powerful symbiosis of both approaches: parallel and iterative. To generate a single

library of motifs we need to randomize only a few residues in a long peptide and also to randomize the positions of these randomized residues. In all nonrandomized positions we need to have some "average" amino acid, side chains of which may not participate in binding with the macromolecular target, but which may properly orient the critical residues. In our experiments we used a mixture of 19 natural amino acids to fill the "average" amino acid positions. We believe that it is reasonable to synthesize libraries with motifs containing three or four residues. This allows us to reduce dramatically the size of complete libraries and yet still expect sufficiently high activity for detection in screening.

We have synthesized several "Libraries of Libraries" containing three residue size motifs. The libraries were screened using anti-β-endorphin antibodies, streptavidin and thrombin. For the synthesis of the first library we used 19 proteinogenic (without Cys) amino acids in randomized positions and a mixture of the same amino acids in nonrandomized positions. In another library 78 amino acids were incorporated at randomized positions and a mixture of 19 proteinogenic amino acids of either L or D configuration were used to fill nonrandomized positions.

Motifs found in the library constructed from 78 natural and unnatural amino acids in the screening for streptavidin binders are given in Fig. 1. Similarity to the natural motif His-Pro-Gln is clearly visible. The imidazole ring of His is indispensable for binding as well as the secondary amine in the next position. Glutamine can be replaced by another residue with hydrogen bonding potential. Aromatic residue probably finds an additional beneficial interaction in the binding site.

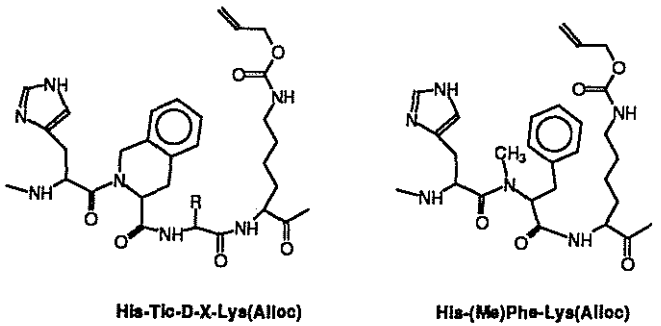

Fig. 1. Structure of streptavidin ligand motifs.

References

1. Gallop, M.A., J. Med. Chem., 37(1994)1233.
2. Gordon, E.M., J. Med. Chem., 37(1994)1385.