Construction and Screening of Libraries of Peptide and Non-Peptide Structures

Michal Lebl
Viktor Krchňák
Pavel Šafář
Alena Stierandová
Nikolai F. Sepetov
Petr Kočiš
Selectide Corporation, Tucson, Arizona 85737

Kit S. Lam
University of Arizona College of Medicine,
Arizona Cancer Center and Department of Medicine,
Tucson, Arizona 85724

I. Introduction

The generation and screening of large numbers of structurally diverse chemical compounds in the quest for pharmacologically active structures has become an important component of drug discovery efforts at pharmaceutical companies and academic institutions. (For review see e.g. (1).) This synthetic approach differs in several important aspects from traditional approaches in which the diversity of compounds is achieved from plant or microbial organism extracts or through accumulation of synthetic compounds over years of synthesis. To begin with, the diversity of structures which can be generated far exceeds that available in the compound "libraries" of all major pharmaceutical companies. Perhaps more importantly, however, is that despite the great diversity achieved, all active structures can be identified and easily synthesized in pure form.

The first generation of diversity available through these synthetic methods consisted primarily of peptides and peptide-like structures, incorporating natural and unnatural amino acids (2-5). These library techniques take advantage of over 40 years of solid phase peptide chemistry to enable the relatively easy and sequential synthesis of linear, branched, or cyclic compounds in a library format. Despite their prevalence in nature, however, peptides do not comprise a preponderance of pharmaceutical products. In large part this may be due to the constraints imposed on drug candidates, such as bioavailability and pharmacokinetics. While it has been well established that desirable characteristics in these categories can be designed into peptides through structure-activity analysis and subsequent modifications, this process is time consuming and less than certain.

TECHNIQUES IN PROTEIN CHEMISTRY V

In addition to peptides, there are several classes of chemical compounds which make up the bulk of drugs on the market today. The reasons for the dominance of certain classes of compounds are many, but a principal reason is the presence of desirable pharmacokinetic and bioavailability characteristics. In addition, it is becoming clear that certain classes of receptors, for instance, bind preferentially to a specific chemical class of compounds – what some are now calling a "preferred platform." The ability to produce in a synthetic library format a diversity of compounds based on preferred platforms or other structures known to have favorable drug characteristics should greatly enhance the applicability of the library techniques by speeding the drug discovery and optimization process and increasing the probability of successfully identifying promising compounds.

The Selectide Process offers a powerful tool to rapidly generate and screen a diversity of chemical compounds, both peptides and non-peptides. The Selectide Process is based on (i) synthesis of a multiplicity of diverse compounds on polymeric beads, with a single structure on each particle (3,6), (ii) screening this library in either a binding assay based on an enzyme-linked or fluorescent tag (3,6,7), or in standard solution-phase assays (after partial release of a defined portion of the compound from each bead into solution (8-11)), and (iii) determination of the structure on the "active" bead. Compound composition is determined either by sequencing peptides by Edman degradation or using a mass spectroscopic technique (12).

II. Materials and Methods

Sequencing by Edman degradation was performed on an ABI 4778 protein sequencer (Applied Biosystems, Foster City, CA) and Porton PI 3010 instrument (Porton Instruments, Tarzana, CA). Solid phase syntheses of single peptides were performed manually in polypropylene syringes as described by Krchňák and Vágner (13). Syntheses of libraries were performed as described earlier (3,4,7,14). Syntheses were performed on TentaGel S NH_2 (TG) resin (Rapp Polymere, Tubingen, Germany, 130 or 80μm, 0.23mmol/g) modified with an appropriate linker. Fmoc protecting groups were cleaved with 50% piperidine/DMF for 1x10min, Tfa groups by repeated treatment (3 x 1 min + 90 min) with 20% piperidine/water. Npys groups were removed by 0.3M solution of HCl in dioxane for 5 + 30 min, Boc groups were cleaved with 30% TFA/DCM containing 3% of anisole for 20min. A mixture of BOP/HOBt/DIEA (1:1:2eq) in DMF was used for the activation of both protected amino acids and carboxylic acids used for randomization. The peptide and non-peptide libraries were screened according to the published procedure (7).

III. Results and Discussion

A necessary requirement for incorporating building blocks other than amino acids or nucleotides into a library in order to increase diversity is the ability to identify the composition of positively reacting compounds. This is straightforward for library technologies in which the stepwise tracking of synthesis (iteration) leads

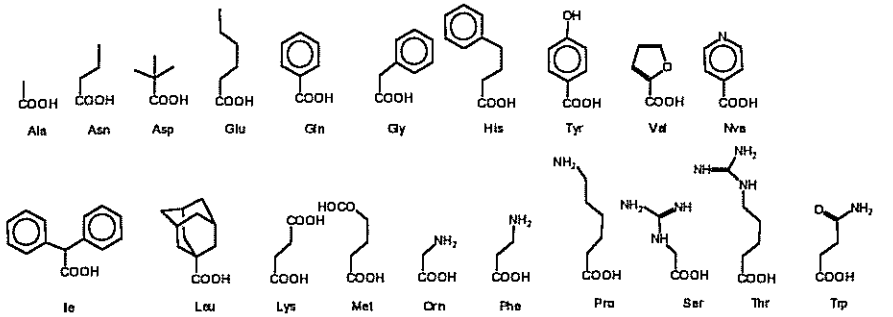

Figure 1. Building blocks (and their coding) used for the construction of nonpeptide libraries on nonpeptide (Kemp's triacid based) and branched peptide scaffolding.

to the eventual identification of an interacting molecule (4,15). However, the need to start from a given set of "structural families," having one or two positions in the structure fixed, limits the ability of these methods to identify all possible motifs with affinity for the given acceptor. If the "critical" residues for a given motif were not chosen for the fixation in the beginning of library construction, an important motif can be easily overlooked.

This potential liability is overcome by library synthesis and assay techniques where all possible combinations are synthesized, screened, and then the identity of positively reacting compounds are determined by direct or indirect analysis. The need to determine the structure found on positively reacting beads is the challenge for this approach. The principle of encoding can be used for structure elucidation, based on a parallel synthesis of peptide (16,17) or nucleotide (18) structures together with the nonsequenceable structures.

We decided to combine the simplicity of peptide structures with the diversity available in alternative building blocks. Building blocks (which we have used in two of our examples) for library construction are given in Figure 1. Trifunctional amino acids are suitable for construction of a library template (scaffold) and modification of their side chains with assorted building blocks can be used to construct a library of great diversity. Amino acids such as diaminobutyric acid, aspartic acid, cysteine and/or iminodiacetic acid are the smallest building blocks onto the side chains of which the universe (>3,000 commercially available) of carboxylic acids, amines or halides (aliphatic, aromatic, heterocyclic) can be attached. This approach is simpler than the use of unnatural amino acids for library construction (7,19), where the special amino acids have to be synthesized.

To achieve reasonable binding to a receptor, antibody, enzyme, nucleic acid, the appropriate spatial arrangement of the interacting structures must be realized. Linear presentation of building blocks in libraries may not be an optimal format for the selection of the best binding structures. One strategy for displaying the interacting structures would be their placement onto a scaffold which would map the appropriate segments of space.

Two ways to construct a molecular scaffold are illustrated in Figure 2. The

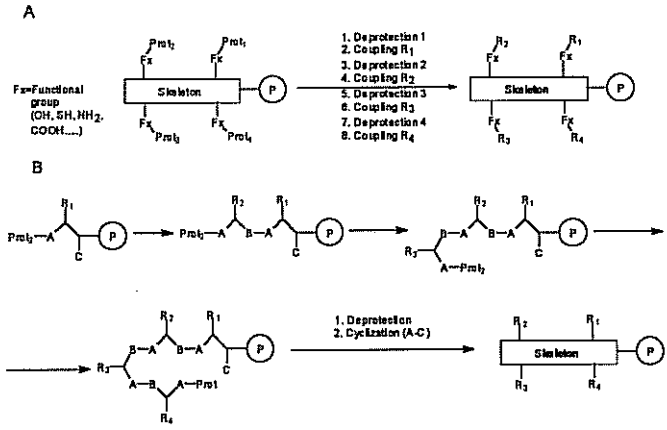

Figure 2. Two strategies for building scaffolding (template, skeleton) libraries. A: "Glucose approach"; B: "Benzodiazepine approach".

first approach, "glucose approach," utilizes a selectively protected scaffold on which the randomizations are performed (Fig. 2A). In the second approach, "benzodiazepine approach," the randomizations are performed during scaffold construction (Fig. 2B). We have used both approaches in our examples.

An example of a non-peptidic scaffold is based on the use of Kemp's triacid (20) (Fig. 3). In this structure the three carboxyl groups are constrained to the triaxial conformation. Structure I with two orthogonal protecting groups attached via an ethylenediamine linker and one free carboxyl group was synthesized. For generating the library, the first randomization was performed on the side chain of lysine attached to the resin. In the next step, scaffold I was bound to the free amino group of lysine, and randomization was performed in two steps after removal of Fmoc in the first step and Boc group in the second step. The set of carboxylic acids used for randomization in all three steps (without coding) is shown in Fig. 1.

Libraries were also designed based on an attachment of building blocks to a scaffold constructed by the consecutive coupling of diamino carboxylic acids.

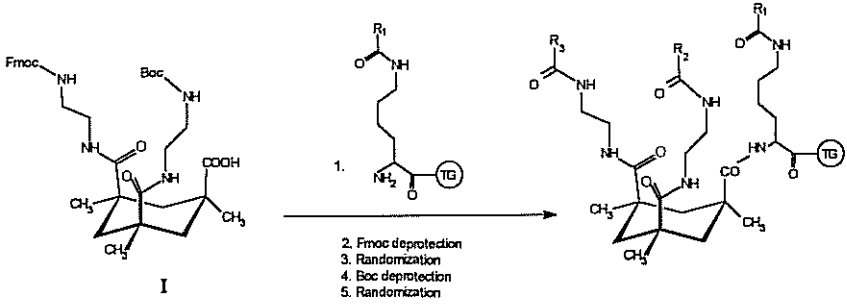

Figure 3. Building of scaffolding library based on Kemp's triacid. R_1-R_4, carboxylic acids given in Figure 1.

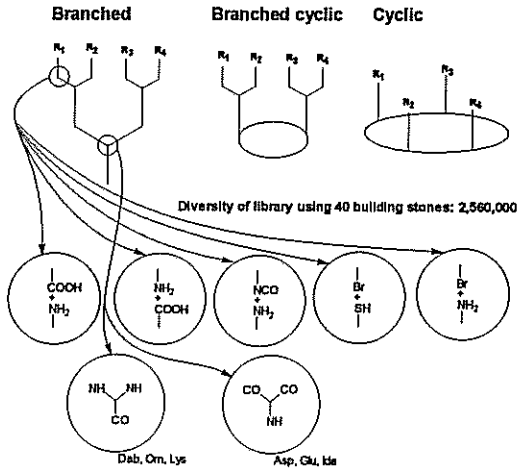

Figure 4. Basic types of peptidic scaffoldings and chemistries readily available for their construction.

The principles of these libraries are illustrated in Fig. 4. We explored the synthesis of the branched library. The synthesis required the use of four orthogonal protecting groups (Tfa, Npys, Fmoc, and Ddz). We found that the Tfa group was not cleaved during Fmoc deprotection using 50% piperidine in dimethylformamide, but was completely cleaved by 1 - 2 h exposure to piperidine solution (20%) in water, which, however, also cleaved the Fmoc group. This limitation dictated the

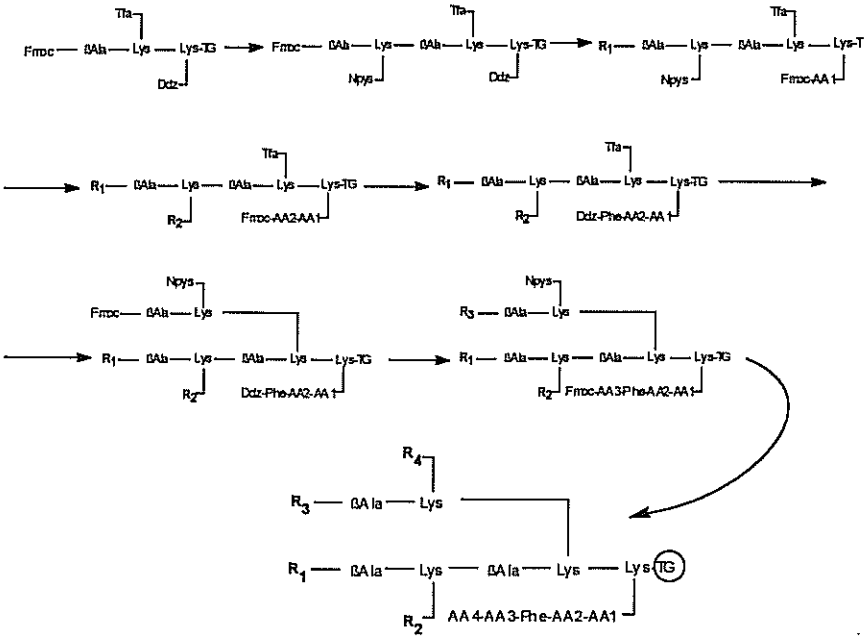

Figure 5. Synthetic scheme for the nonpeptide library on branched peptidic scaffolding. R_1-R_4, carboxylic acids given in Figure 1; AA1-AA4, amino acids used for coding.

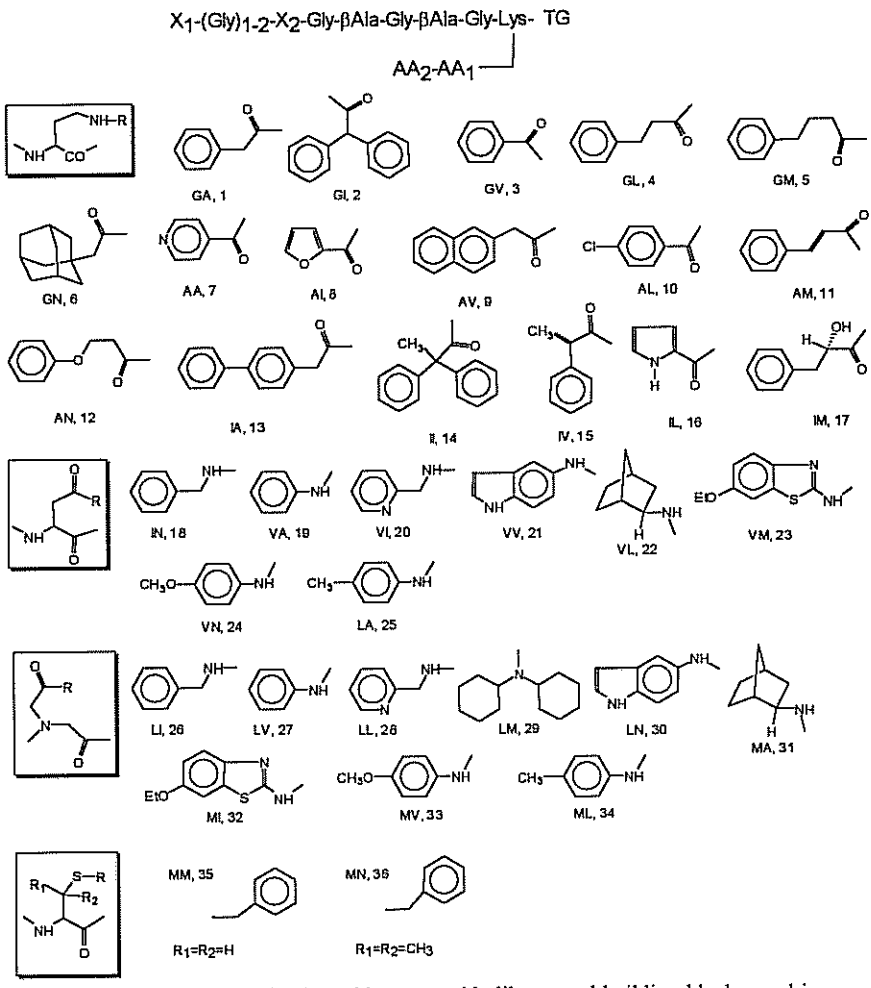

Figure 6. Structure of the mixed peptide-nonpeptide library and building blocks used in position 4. Doublets of amino acids used for the coding are given in one-letter code.

strategy used in the construction of this library (Fig. 5).

Unnatural building blocks can be combined with standard amino acids. We have shown that this approach can yield structures with reasonable affinity by constructing a library of 936 members, having selected natural L-amino acids randomized in position 1 (X_1, Fig. 6), one or two glycines in position 2 and 3 and a set of aromatic amines coupled to the β-carboxyl group of aspartic acid or side chain modified iminodicarboxylic acid, or aromatic acids coupled to the side chain of diaminobutyric acid, or benzylhalides coupled to the side chain of sulfur containing amino acids (cysteine and penicillamine) in position 4 (X_2, Fig. 6). Since these substitutes were not necessarily sequenceable we coded for this position, using a doublet of amino acids to code for each building block since

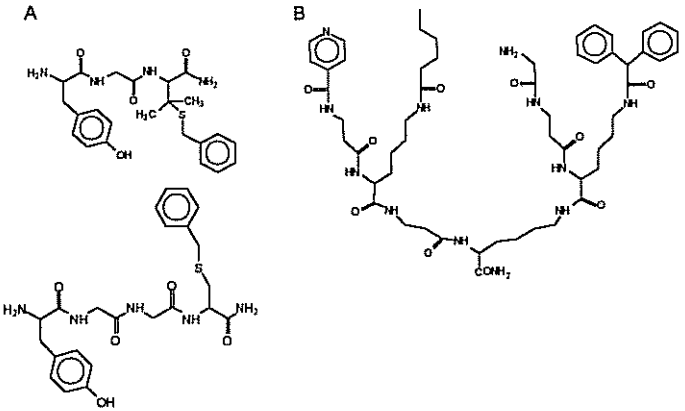

Figure 7. Structures identified in screening for ligand to (A) anti-β-endorphin monoclonal antibody, and (B) streptavidin.

more than 20 building blocks were used in the randomization. Using a doublet of 6 different amino acids we could code for up to 36 different building blocks. The structure of the library and the coding (AA_2-AA_1) is given in Figure 6.

The above described libraries were screened in our model systems to find ligands to anti-β-endorphin monoclonal antibody and streptavidin. Positively reacting beads were subjected to Edman degradation, and the interacting structures were deduced (Fig. 7). After resynthesis on beads to confirm binding, these compounds were synthesized in free form and their binding affinities were determined. Bead bound sequences exhibited specific binding (competable by leucine enkephalin and biotin, respectively). The results indicated that binding to the antibody requires two aromatic groups at the appropriate distance. The structure connecting the two aromatic groups also contribute to binding affinity. The non-peptide structure binding to streptavidin is quite different from both the natural ligand biotin and from the peptide found from screening peptide libraries.

IV. Conclusions

We have shown that the principle of peptide encoding allows for the generation of libraries constructed from very simple and readily available non-amino acid building blocks. This opens completely new ways to the area of generating structural multiplicity through the synthesis of libraries of compounds. The screening of nonsequenceable/non-peptide libraries was shown to produce compounds which bound specifically to the given targets.

References

1. Pavia, M.R., Sawyer, T.K., Moos, W.H. (1993). Bioorg. Med. Chem. Lett. **3**:387.
2. Geysen, M.H., Rodda, H.M., Mason, T.J. (1986). Molecular Immunology **23**:709.
3. Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M., Knapp, R.J. (1991). Nature **354**:82.
4. Houghten, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T., Cuervo, J.H. (1991). Nature **357**:84.
5. Fodor, S.P.A., Read, J.L., Pirrung, M.C., Stryer, L., Lu, A.T., Solas, D. (1991). Science **251**:763.
6. Lam, K.S., Hruby, V.J., Lebl, M., Knapp, R.J., Kazmierski,W.M., Hersh, E.M., Salmon, S.E. (1993). Bioorg. Med. Chem. Lett. **3**:419.
7. Lam, K.S., Lebl, M. (1992). Immunomethods **1**:11.
8. Lebl, M., Pátek, M., Kočiš, P., Krchňák, V., Hruby, V.J., Salmon, S.E., Lam, K.S. (1993). Int. J. Peptide Protein Res. **41**:201.
9. Lebl, M., Lam, K.S., Kočiš, P., Krchňák, V., Pátek, M., Salmon, S.E., and Hruby, V.J. (1993). In "Peptides 1992", Proc. 22nd Eur.Pept.Symp. (Schneider, K.C. and Eberle, A.M., eds.) 66.
10. Kočiš, P., Krchňák, V., Lebl, M. (1993). Tetrahedron Lett. Submitted.
11. Salmon, S.E., Lam, K.S., Lebl, M., Kandola, A., Khattri, P.S., Healy, S., Wade, S., Pátek, M., Kočiš, P., Krchňák, V., Thorpe, D., Felder, S. (1993). Proc. Nat. Acad. Sci. U.S.A., in press.
12. Sepetov, N.F., Issakova, O.L., Lebl, M., Swiderek, K., Stahl, D.C., and Lee T.D. (1993). Rapid Comm. Mass Spectrom. **7**, 58.
13. Krchňák, V., Vágner, J. (1990). Peptide Research **3**:182.
14. Furka, A., Sebestyen, F., Asgedom, M., Dibo, G. (1991). Int. J. Peptide Protein Res. **37**:487.
15. Pinilla, C., Appel, J.R., Blanc, P., Houghten, R.A. (1992). BioTechniques **13**:901.
16. Nikolaiev, V., Stierandová, A., Krchňák, V., Seligmann, B., Lam. K.S., Salmon, S.E., Lebl, M. (1993). Peptide Res. **6**:161.
17. Kerr, J.M., Banville, S.C., Zuckermann, R.N. (1993). J. Am. Chem. Soc. **115**:2529.
18. Brenner, S., Lerner, R.A. (1992). Proc. Natl. Acad. Sci. USA **89**:5381.
19. Simon, R.J., Kaina, R.S., Zuckermann, R.N., Hueber, V.D., Jewell, D.A., Banville, S., Simon, N.G., Wang, L., Rosenberg, S., Marlowe, C.K., Spellmeyer, D.C., Tan, R., Frankel, A.D., Santi, D.V., Cohen, F.E., Bartlett, P.A. (1992). Proc. Natl. Acad. Sci. USA **89**:9367.
20. Kemp, D.S., Petrakis, K.S. (1981). J. Org. Chem. **46**:5140.