

**Michal Lebl**  
**Viktor Krchňák**  
**Nikolai F. Sepetov**  
**Bruce Seligmann**  
**Peter Štrop**  
**Stephen Felder**  
*Selectide Corporation*  
*1580 E. Hanley Blvd.*  
*Tucson, AZ 85737*

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

---

## One-Bead–One-Structure Combinatorial Libraries

*Combinatorial libraries employing the one-bead–one-compound technique are reviewed. Two distinguishing features characterize this technique. First, each compound is identified with a unique solid support, enabling facile segregation of active compounds. Second, the identity of a compound on a positively reacting bead is elucidated only after its biological relevance is established. Direct methods of structure identification (Edman degradation and mass spectroscopy) as well as indirect “coding” methods facilitating the synthesis and screening of nonpeptide libraries are discussed. Nonpeptide and “scaffold” libraries, together with a new approach for the discovery of a peptide binding motif using a “library of libraries,” are also discussed. In addition, the ability to use combinatorial libraries to optimize initially discovered leads is illustrated with examples using peptide libraries. © 1995 John Wiley & Sons, Inc.*

---

### INTRODUCTION

Synthetic combinatorial libraries are fast becoming an important method for lead discovery in the pharmaceutical industry. The ability to synthesize and screen in a high throughput assay format a number of compounds that will greatly exceed that available from historical libraries or natural products is likely to make the lead discovery process more efficient. The integration of these techniques with rational drug design and traditional medicinal chemistry techniques may also facilitate the rapid optimization of compounds discovered in initial screens. Indeed, combinatorial libraries may them-

selves enhance the ability to rapidly map space through structure–activity studies, creating a data base from which to derive design ideas, or perhaps to select a drug candidate. The power and scope of combinatorial libraries has yet to be fully explored; however, advances continue to enhance the synthetic diversity that can be achieved and the speed and flexibility with which assays can be performed.

### LIBRARY TECHNIQUES

Techniques that generate arrays of structures of known identity and provide for testing on an indi-

vidual basis<sup>1-8</sup> share an advantage over the one-bead-one-compound approach, in that there is no need for structure determination. However, these are not library techniques in the truest sense. Techniques that generate families of structures<sup>9-13</sup>—so-called iterative techniques—can be used to generate structural multiplicity; however, in using this method it is difficult to identify multiple independent structural motifs. As a result, iterative techniques are not random screening techniques, since libraries are prepared and screened in a highly systematic and ordered manner<sup>14</sup> with the intent to identify a single lead compound. The principal difference between the iterative approach and the one-bead-one-compound approach<sup>15</sup> is that the iterative approach relies on multiple sequential synthetic steps and screening of complex mixtures of compounds to arrive at one of the possible (perhaps several) motifs (sequential approach). The one-bead-one-compound approach is a one-step process where all structures to be tested are screened concurrently—often resulting in the discovery of different leads (parallel approach). At the completion of the iterative screening approach (or after screening positional scanning libraries<sup>16,17</sup>), the structure of one positive ligand is known (derived from the synthetic algorithm, or deduced from residues identified in positional scanning); whereas the structure of the positive beads isolated in a one-bead-one-structure screen remains unknown until identified using a direct technique or decoded by one of a variety of described techniques. The advantage, however, is that one can possibly obtain several independent lead structures, as well as multiple analogues of the leads.

Combinatorial library techniques yielding unique but unidentified structures can be divided into three categories—biological systems, such as filamentous phage<sup>18-21</sup> or plasmids,<sup>22</sup> synthetic methods for generating compound mixtures free in solution,<sup>13,23-29</sup> or attached to the solid surface,<sup>10,25,30,31</sup> and synthetic methods for generating one compound on each bead.<sup>15,32-40</sup> Biological systems can generate up to  $10^{12}$  peptides of virtually unlimited size; however, library components are limited to only the genetically encoded amino acids. There is also a genetic bias in the creation of these libraries, causing them not to be truly random. The synthesis of solution phase mixtures is facile, but the subsequent isolation and characterization of active components has proven difficult, except in instances in which simple mixtures were tested (e.g., Ref. 29), or in the case of nucleotide libraries in which the ligand can be successively

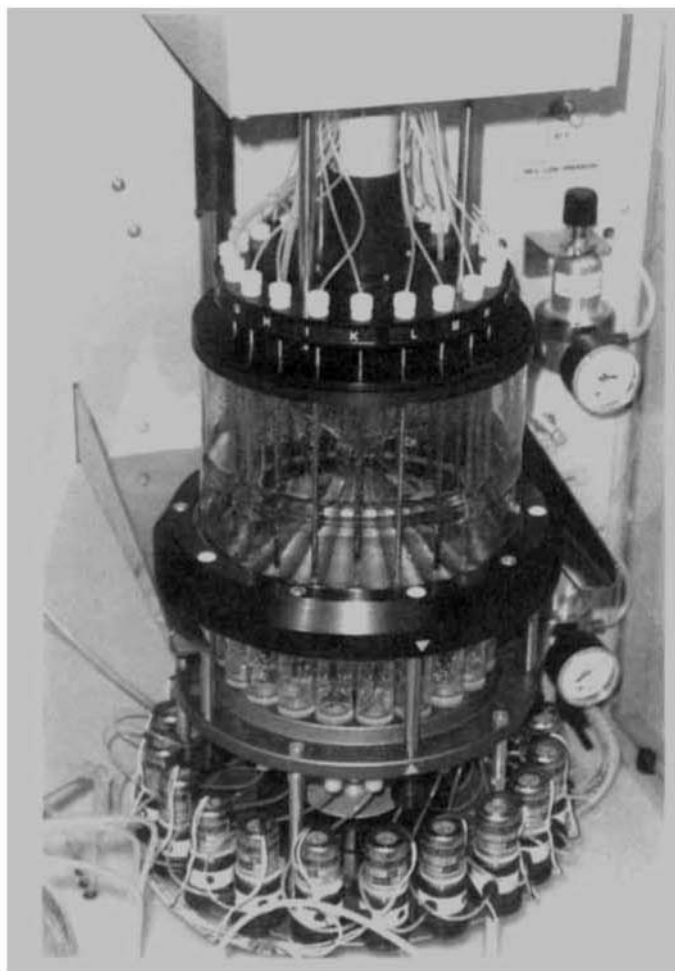
amplified.<sup>41-43</sup> This method has consequently not been widely adopted outside of nucleotides. Several reviews of the application of combinatorial libraries to drug discovery were recently published<sup>44-49</sup> as well as elsewhere in this issue. The scope of this review is limited to the potential and limitations of the synthetic one-bead-one-structure libraries described initially by Lam.<sup>15</sup>

## ONE-BEAD-ONE-STRUCTURE LIBRARIES

The Lam method, also referred to as the Selectide Process, is a random synthetic library approach based on the one-bead-one-structure concept. The split synthesis method for generating libraries of this type was first described by Furka *et al.*,<sup>50-52</sup> who applied this method for synthesis of equimolar peptide mixtures.<sup>53-55</sup> This synthetic method was later used to generate iterative libraries,<sup>9</sup> or one-bead-one-peptide libraries.<sup>15</sup> It was the pivotal recognition of Lam *et al.* that since each solid support contained a unique compound, the entire bead-bound peptide library could be screened against an acceptor molecule (e.g., natural or artificial receptor, enzyme, antibody, or even small molecules) using an ELISA-type assay (ELISA: enzyme-linked immunosorbant assay).<sup>15,32,40,56</sup> This approach was subsequently expanded by many to include a fluorescence based assay, using, for example, a fluorescence activated cell sorter (FACS),<sup>38</sup> fluorescence microscopy,<sup>34,37</sup> measuring fluorescence in solution,<sup>57</sup> radio ligand binding,<sup>36,58</sup> or magnetic bead binding.<sup>59</sup> Labeled beads could then be recovered and the identity of the compound on the bead determined directly by sequencing (in the case of sequencable peptides), by mass spectroscopy (in the case of small peptides or small organic molecules), or indirectly through a sequencable component (code) on the bead corresponding to any nonsequencable components in the test compound. Alternatively, using an orthogonal two-stage release process and tracking the relationship between the parent bead and its releasate, solution phase assays could be performed and the identity of positive compounds determined from the parent bead<sup>33,35,60-62</sup> (see below).

## SYNTHESIS OF ONE-BEAD-ONE-STRUCTURE LIBRARIES

During the last three years we have synthesized more than 400 different peptide and nonpeptide,

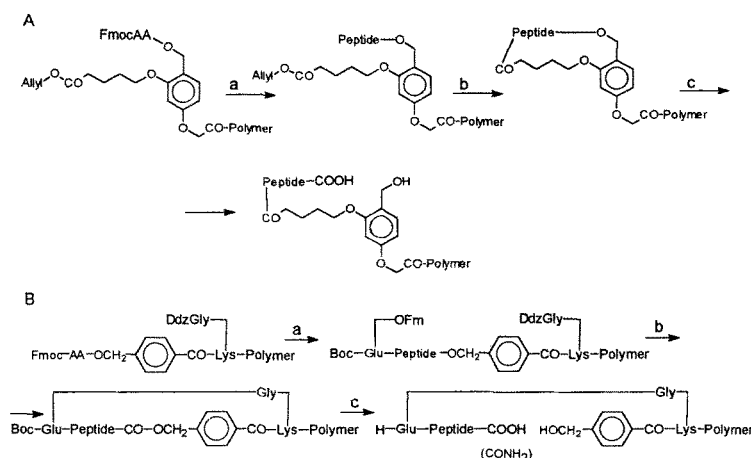


**FIGURE 1** Reactor assembly of automated library synthesizer.

cyclic and linear libraries. We realized very early that library synthesis is demanding and labor intensive, although adaptable to automation. Zuckermann et al.<sup>63</sup> constructed a robotic apparatus for library synthesis. In this instrument the resin is portioned by a robotic arm operated syringe using an isopycnic slurry of the beads. Another library synthesizer using a similar principle was described.<sup>64,65</sup> We use<sup>66</sup> a synthesizer, the main components of which are reaction vessels connected to a mixing chamber, as illustrated in Figure 1. The lower part of the apparatus consists of 20 individual reaction chambers in which couplings are performed. Randomization begins by filling both the lower reaction vessels and the mixing chamber with randomization solvent, blowing nitrogen through the frits at the bottom of the reaction chambers, and stirring. Sedimentation then redistributes the polymeric carrier for the next coupling step.

In the case of sequenceable peptides or non-

sequenceable test compounds encoded by sequenceable structures (see below), individual positive beads are loaded into a microsequencer for structure determination. Peptides attached to a polymeric support displaying a free amino terminus can be directly sequenced by Edman degradation. In some cases, however, it is required that the carboxy terminus be free for binding to occur. Since the synthesis of peptides from N- to C-terminus is not well developed, it was necessary to apply a technique to “reverse” the peptide on the polymeric carrier. The peptide was synthesized using a linker that was included in the cyclic structure. After cyclization the linker was selectively cleaved and the carboxy terminus of the peptide was displayed (Figure 2B). This approach also provides a free amino terminus, enabling N-terminal Edman degradation.<sup>67</sup> An alternative approach uses a template that serves as both the linker and cyclization point<sup>68</sup> (Figure 2A).



**FIGURE 2** Scheme of peptide reversal on solid support. (A) Using multifunctional linker<sup>68</sup>; (B) using trifunctional amino acids<sup>67</sup> (the same approach was published later by others<sup>122</sup>). (a) Synthesis of the peptide (library); (b) cyclization; (c) cleavage/reversal.

In contrast to the biological libraries mentioned above, synthetic methods have a practical limit on the size of the library as well as length of the peptides in the libraries. The synthetic approach, however, offers the ability to incorporate into libraries unnatural amino acids, disulfide and nondisulfide cyclic structures, reduced peptide bonds, nonpeptide moieties, and specialized linkers (scaffolds) on which to attach the subunits, departing from the sequential arrangement of subunits.<sup>5,8,67,69–76</sup>

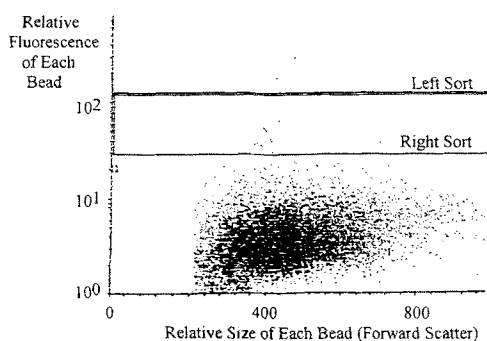
Quality control of libraries can also be performed prior to screening. This can be accomplished by sequencing randomly chosen beads or by multiple sequencing of a sample.<sup>77</sup> Amino acid analysis can be performed as well as mass spectroscopic evaluation of the sample of the cleaved library.<sup>78</sup> In all cases a statistically relevant sample must be evaluated. However, the best way to assure library quality is analytical control of each synthetic step, either by following the reaction by a noninvasive method<sup>79</sup> at the level of individual beads (inspection of indicator color disappearance under the microscope), or the application of classical analytical methods<sup>80</sup> with a statistically significant sample of the carrier.

## SCREENING OF BEAD-BOUND LIBRARIES

Compatibility of biological assays is largely dependent on the synthetic approach chosen. The one-bead-one-structure approach offers various options for screening. The most widely adopted method of

screening is the “on-bead” binding assay. Ligands that bind to monoclonal antibodies,<sup>15,36,38–40,56,67,81,82</sup> protein G, MHC class I molecules,<sup>83</sup> the platelet-derived gpIIb/IIIa receptor,<sup>33</sup> the SH3 domain of phosphatidylinositol 3-kinase,<sup>34</sup> thrombin, factor Xa, cytokine receptors, streptavidin,<sup>15,32,84</sup> avidin,<sup>32</sup> and even small molecules,<sup>85,86</sup> as well as ligands for artificial receptors,<sup>87,88</sup> have been identified. This bead-binding assay has several attributes: (a) it is extremely rapid, taking only a few hours to screen  $10^7$ – $10^8$  beads, (b) the color intensity or fluorescence of the bead is generally proportional to the binding affinity of the ligand,<sup>38</sup> and (c) the library may be reused multiple times for different probes.

Enzyme substrates have been identified using the on-bead assay of a library of compounds tagged with a fluorogen at one end and a molecule that quenches fluorescence at the other end.<sup>37</sup> Cleavage results in loss of quenching and produces a fluorescence signal from beads containing cleavable compounds. Sequencing then reveals the place of enzymatic cleavage and the structure of the substrate. Likewise, enzyme activity such as phosphorylation can be measured directly by overlaying beads with agarose containing the target enzyme and performing a colorimetric or radiolabel based enzyme assay. A specific example of a radiolabel-based assay is a screen designed to identify substrates for cAMP-dependent protein kinase. The library was covalently radiolabeled with  $\gamma$ [<sup>32</sup>P]ATP, and a substrate motif known from the literature was identified.<sup>58</sup> The use of a radiolabeled macromolecular probe for screening a one-bead-



**FIGURE 3** Result of FACS sorting of a pentapeptide L library after incubation with streptavidin labeled with a fluorescent tag. Vertical axis: fluorescence; horizontal axis: particle scatter.

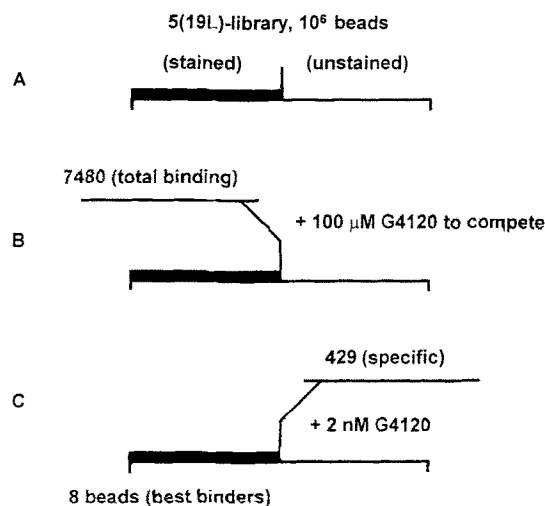
one-peptide library has also been reported.<sup>36</sup> However, from our experience, the method is often slow and insensitive, and it offers no advantage over that of the immunohistochemical method.

Fluorescence based screening methods using FACS are fast<sup>38</sup> and allow for the selection of ligands to targets with high binding off-rates (Felder et al., unpublished). Critical to this technique is the proper selection of bead material and synthetic protocol, as well as the fluorescent marker, since the autofluorescence of the library beads can render the library unsuitable for this type of screening. Figure 3 shows the distribution of library beads after addition of a model macromolecular target, streptavidin, labeled with a fluorescent tag. The most fluorescent beads were sorted by deflection to vessels at the left and right of the sheath flow, were tested for specificity of binding, and were sequenced. Both beads so selected contained the expected HPQ motif: HMHPQ and HPQQ?

Regardless of the assay format, false positive beads, i.e., beads bound nonspecifically or bound to areas of the target outside of the active site, must be eliminated. The ability to eliminate false positives and select only true positive compounds before determining their identity and carrying out re-synthesis is a major factor in the success of the screening process. Common methods such as the use of nonionic detergents, blocking agents such as bovine serum albumin or gelatin, basic or acidic conditions, chaotropic agents, etc., are useful to reduce much of the nonspecific binding seen in libraries. However, it is also necessary to design the assay to assess specificity of binding before determining the identity of positive beads. This can be achieved by the sequential screening of the library in the presence or absence of a specific blocker

(e.g., a known ligand), or with active and inactive preparations of the target molecule. These sequential steps are illustrated in Figure 4, which displays data from a histochemical based screen for ligands to the gpIIb/IIIa receptor. The use of a dual color substrate system can also help differentiate specific binding from nonspecific binding. (Lam et al., manuscript in preparation). The first dye is used to identify all compounds that bind, the second for staining in the presence of a competitor. Beads expressing the combination of both dyes are qualified as nonspecific binders; those colored only by the first dye are taken for the next treatment (destaining and restaining in the presence of various concentrations of the receptor and/or competitor). This technique eliminates the demanding step of physically removing a large number of positive beads identified in the first step for restaining and competition experiments.

The potential for false positives that originate from an interaction of the test compound with the tag used to label the target molecule can be eliminated by selecting positive beads with a target molecule labeled with one tag, and then reassaying and selecting positive beads with the same target molecule labeled with a different tag (antibody or streptavidin, for instance). All these extensive manipu-



**FIGURE 4** Scheme of screening for gpIIb/IIIa receptor ligands, eliminating nonspecific and weak ligands. (A) Incubation of the library with labeled targets; (B) reincubation of the stripped beads selected in step A with a specific inhibitor to eliminate nonspecific interactions—negative beads are selected for step C; (C) reincubation of stripped beads selected in step B with a lower concentration of inhibitor to select the most active ligands—positive beads are selected for sequencing.

lations are possible using the Selectide Process because the test compound is relatively stable on the bead, permitting positive beads to be repeatedly recycled and restained under various conditions for characterizing specificity. The optimal way for elimination of false positives is the use of a so-called hybrid assay, in the first step of which the bead is selected using the solid phase binding assay and in the second step the compound is released from the bead and tested for the activity in solution.

It is not unusual to obtain too many true positives (e.g., > 1000). To select the best binders it is possible to rescreen the library under more stringent conditions. There are several methods of increasing the stringency of screening, but we believe the best method is simply to use a lower concentration of probe. Care must be taken, however, to avoid consumption of probe and to allow sufficient time for binding to each bead to reach equilibrium. Alternatively stringency can be increased by increasing wash time—to select only those compounds with slow dissociation rates. The free probe concentration can be lowered by adding a modest concentration of competing ligand or substrate. Alternatively, one can change the conditions of the assay in order to decrease the affinity of all ligands—for example, by incorporating chaotropic agents or by altering ionic strength or pH. However, this approach involves the risk of identifying compounds that may not bind well under physiological conditions.

### SOLUTION ASSAY SCREENING OF ONE-BEAD-ONE-STRUCTURE LIBRARIES

The one-bead-one-structure technology described earlier in this article is not limited to the screening of soluble targets.<sup>15,89,90</sup> We have described the release of test compounds from library beads in such a manner that after testing for activity in solution on cellular or membrane targets active compounds could be backtracked to their bead of origin, and the identity of the positive test compound could be determined from the parent bead.<sup>33,91</sup> While it is certainly possible to divide beads one by one into wells and then release compound for testing, higher throughput can be achieved using a multistep release format. We have described several linkers suitable for this “multiple release” application.<sup>61,62</sup> A schematic of the protocol is shown in Figure 5. The library beads are distributed into the wells of microtiter filtration plates and the release of the

first portion of the compound attached to each solid support is performed under very mild conditions. In this step several hundred beads are placed in each well. A mixture of a substantial number of unrelated compounds is generated in approximately equimolar quantities. The components of the mixture are structurally unrelated, since the distribution of the beads into individual wells is based only on the statistics. Similarity of the compounds is defined only by the chemical composition of the library. The small aliquots of soluble compound mixtures (sublibraries) are filtered to the test microtiter plate and biologically assayed directly in this replicate plate or using subaliquotes from the replicate plate. Active mixtures are identified, and all beads from each well responsible for the activity are recovered from the corresponding well of the master filtration plate and redistributed at one bead per well into a second set of master plates. The second portion of each compound is then released, and the biological test is repeated. In the second step filtration may not be necessary, provided that the bead can be retrieved from the well after the test assay is complete. The bead(s) in the second masterplate corresponding to any (now individual) positive compound(s) are recovered and the identity of the test compound determined from material (test compound or code) remaining on the bead. The potential for “masking” the activity of a single compound in a mixture of other components has been studied, and it was determined that even though the activity of a specific peptide contained in a mixture may be modulated by other components of the mixture, it was still clearly detectable.<sup>92</sup>

Another approach to testing peptides released from the beads was reported recently.<sup>35</sup> The test compounds were partially released from beads that had been layered on top of live amphibian pigment skin cells transfected with specific receptors. The change in color of these cells was measured as a biological readout of stimulatory and inhibitory activity. This method of local release maintains an absolute correlation of released compound to its parent bead, and could be readily adapted to enzyme assays. Since a partial release can be repeated several times, it is possible to start with a relatively dense population of beads and identify the single bead containing the biologically active compound in several iterations using a lower density of beads.

As mentioned above, in our approach equimolar amounts of test compound are released at each step. As a model to demonstrate this, we constructed a molecule on which five copies of the

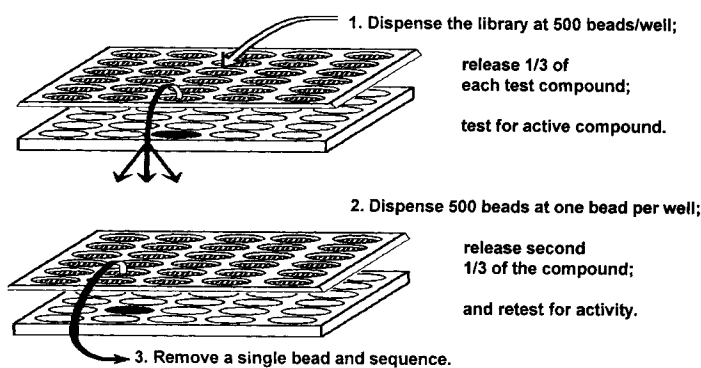


FIGURE 5 Scheme illustrating library screening using a double-cleavable linker.

same peptide were prepared<sup>62</sup> (see Figure 6). The model peptide was released in four distinct steps by taking advantage of the differential stability of the protecting groups to cleavage, followed by intramolecular cyclization at neutral pH for the first- and second-stage release, hydrolysis at higher pH for the third-stage release, and photolysis for the fourth-stage release. After the release of the fourth copy of the peptide we determined the sequence from the peptide remaining attached to the bead through a noncleaved but in principle cleavable linker.<sup>93</sup> Due to the construction of the scaffold holding the releasable molecule, the peptide content released in every step was equivalent to that which would be obtained using the classical synthesis protocol for this type of carrier. Since we did not require four discreet stages of release for functional screening, we selected the two release methods yielding the best results, intramolecular cyclization and hydrolysis, for construction of releasable libraries. Intramolecular cyclization is a very mild method for release. This has been used

for release in the case of multiple peptide synthesis on polyethylene pins.<sup>94–96</sup> A disadvantage of the method originally described was that the diketopiperazine (DKP) structure remained on the released peptide molecule. We overcame this limitation by redesigning the cleavable linker to be more convenient and substantially less expensive. A two-stage releasable linker<sup>61</sup> composed of the tripeptide motif *Ida-Ida* (*Ida*, iminodiacetic acid) was found particularly suitable for designing double-cleavable linkers. The *Ida-Ida* dipeptide is prone to DKP formation and provides three carboxyl groups, one on the amino terminal *Ida* and two on the carboxy-terminal. The structure of linkers and mechanism of the staged release are shown in Figure 7.

To test the performance of multiply cleavable linkers, we prepared multiply releasable libraries and screened them against the anti- $\beta$ -endorphin monoclonal antibody (linear tetrapeptide library) and platelet glycoprotein IIb/IIIa receptor (cyclic pentapeptide library), the ligands of which are well known.<sup>33</sup> The known ligands for the anti- $\beta$ -endorphin antibody were found—YGGF, YGVF, and YGAF. We have also combined the release assay with the on-bead binding assay and achieved the same result. The results from the screening against the gpIIb/IIIa receptor are illustrated in Figure 8. An ELISA-based assay that detected inhibitors of the binding of IIb/IIIa to immobilized fibrinogen was used. A library consisting of 100,000 beads was plated into individual wells of microtiter plates at approximately 200 beads per well and the first third of the peptide was released. Two wells were found positive and beads from these wells were replated at one bead per well (406 wells). Assay results for the second round of release are shown in the figure. Three beads were positive, two of them containing

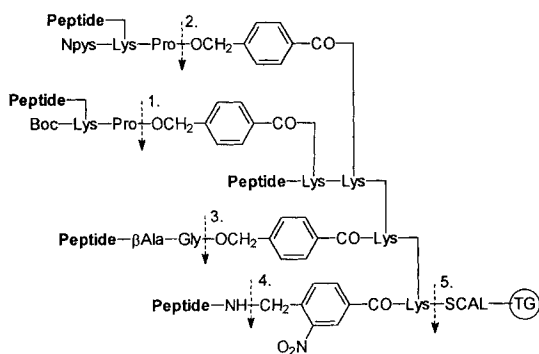
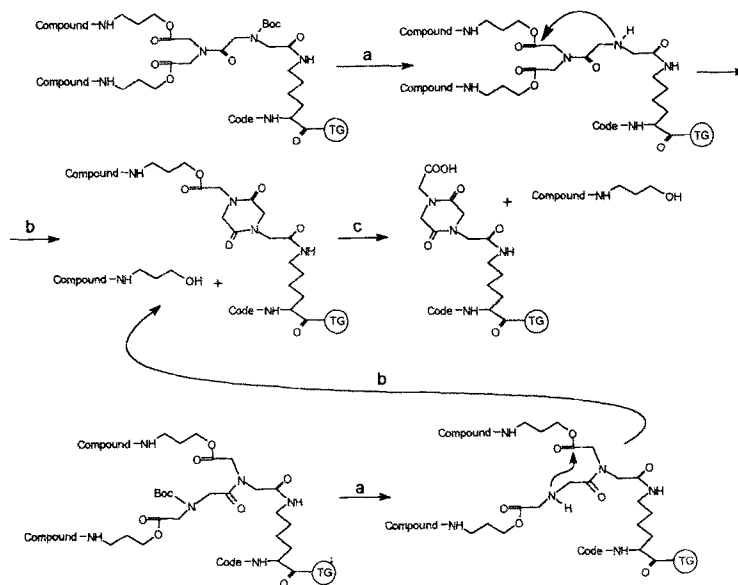


FIGURE 6 Structure of quadruply cleavable linker used for concept validation.

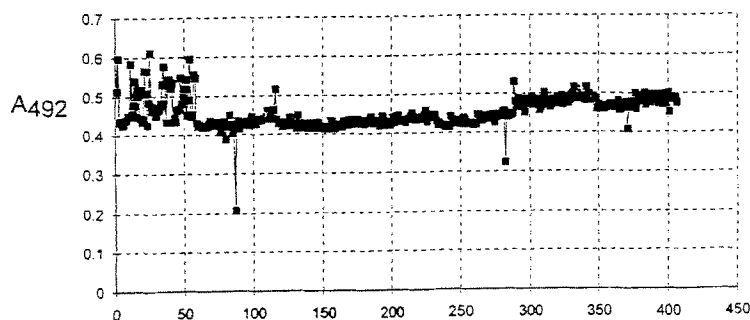


**FIGURE 7** Scheme of double release from two types of iminodiacetic acid-based doubly cleavable linkers. Reagents: (a) trifluoroacetic acid; (b) aqueous buffer pH 8.5; (c) 0.1% NaOH or gaseous ammonia.

the expected sequence CRGDC. The third sequence—CIRYC—was found to be a false positive.

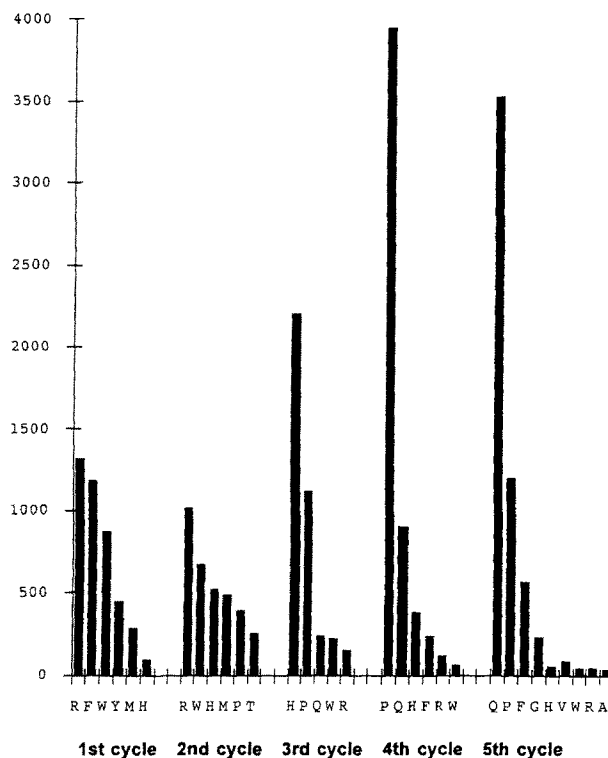
A critical issue in the solution assay is the amount of compound available from one bead. We have been using a commercially available resin, TentaGel, with 130  $\mu\text{m}$  bead size. Typical substitution is approximately 0.2 mmol of  $\text{NH}_2$  groups per one gram of resin and one gram contains approximately 1 million beads. Theoretical release in this case should be around 200 pmol of compound per

bead per arm. However, actual release in our libraries varied from 100 to 150 pmol, as calculated from the control release. Since a realistic assay volume is about 100  $\mu\text{L}$ , 100 pmol of compound yields a 1  $\mu\text{M}$  solution. This concentration is critical to the assay's sensitivity, since the lowest affinity detectable is approximately 1  $\mu\text{M}$ . To increase assay sensitivity, volume can be decreased. However, this may increase sensitivity no more than 3- to 4-fold. The more promising approach is to increase the re-



**FIGURE 8** Results of screening double-cleavable library for gpIIb/IIIa receptor ligand. Vertical axis: Absorbance of solution in individual wells; horizontal axis: well number. Individual beads taken from the two positive wells identified in the testing of the mixtures generated by the first release ( $\sim 200$  beads per well) were plated and released in the 60 inner wells of 7 microtiter plates. Note the unusually high noise in the first plate. Active compounds were identified in wells 88, 282, and 371. The activity of the peptide identified in well 371, CIRYC, was not confirmed after resynthesis. Peptides in wells 88 and 282 had the same sequence CRGDC.





**FIGURE 9** Results of multiple sequencing of 99 beads identified as ligands for streptavidin. Vertical axis: concentration (pmol) of a particular amino acid in each sequencing cycle.

lease from each bead. This can be achieved in principle in three different ways: (a) increase of substitution of the resin, (b) increase bead size, or (c) use an alternative support, e.g., cotton thread or a plastic membrane, that can be cut to pieces of almost any size and modulate the releasable quantity by length or area of support. We have tested all three possibilities. It is possible to increase substitution approximately 5 times with the same size of beads; however, further increase leads to extremely slow reaction rates and incomplete coupling. We have tested custom-made 220  $\mu\text{m}$  beads based on polystyrene grafted with PEG that yield a release in the nmol range. Synthesis on cotton was also shown to be feasible, although handling cotton thread is difficult.

### STRUCTURE ELUCIDATION OF POSITIVELY REACTING COMPOUNDS

Edman degradation is a standard method for the determination of peptide sequences. Bead-bound peptides, however, are attached to very small particles, requiring modifications to the standard sequencing protocol. This protocol may also differ

substantially for different types of polymeric supports. Another complication of the standard procedure arises from the chirality of components included in library synthesis. For these situations special procedures must be developed. However, even the most optimized conditions will not allow separation of all amino acid derivatives, and in some cases, coding is necessary. An average degradation cycle may then take up to 40 min, making structure elucidation one of the slowest steps in the screening process. However, not all positive compounds need to be sequenced individually. It may be advantageous to sequence multiple positive beads simultaneously to obtain an initial idea about structure-activity relationships.<sup>60</sup> Multiple sequencing was originally described for the characterization of peptides bound to MHC molecules.<sup>97</sup> We have applied this method to libraries, demonstrating that in the case where only one motif was present, valuable information could be derived rapidly by combining up to 100 positive beads for concurrent microsequencing. This short cut approach has proven useful for the identification of the anchor residues for MHC class I molecules,<sup>83</sup> the HPQ motif for streptavidin, the YG\_F motif for the anti- $\beta$ -endorphin monoclonal antibody,<sup>60</sup> the anti-insulin

monoclonal antibody, thrombin inhibitors, and other targets. Figure 9 shows the result of multiple sequencing of streptavidin binders as an example of motif identification in the situation where the motif could be "frame shifted," i.e., the motif was active when displayed at several overlapping positions within the compounds synthesized in the library. Two HPQ motifs are clearly visible overlapping in positions 2, 3, and 4, and positions 3, 4, and 5, with a strong preference toward the carboxy terminal.

For peptide and nonpeptide molecules other than oligonucleotides, the only viable alternative for direct structure elucidation is mass spectrometry. This method can use the exchange of labile hydrogens in the peptide molecule to significantly improve the speed and precision of structure determination.<sup>98</sup> Diminishing the number of possible compositions allows the application of an algorithm beginning from the composition of a full-length peptide<sup>99</sup> rather than from small fragments (see, e.g., Ref. 100). Using time-of-flight secondary ion mass spectrometry improves significantly the sensitivity of this method<sup>101</sup> and it can be applied to structures other than peptides. Knowledge of the fragmentation pathways is important to the success of mass spectrometry in structure elucidation; however, even without this knowledge it has been possible to determine the structure of hits from a small nonpeptidic library.<sup>102</sup>

## CODING

There are three main reasons for the introduction of coding techniques: (a) to enable the structure identification of nonsequenceable compounds, (b) to improve sensitivity and (c) to increase the speed and throughput of compound identification. The idea of encoding library compounds was derived from biological systems, in which the composition of peptides/proteins is encoded by DNA. The phage technology, in which the structure of the peptide displayed on the phage surface is decoded by sequencing the phage DNA, takes advantage of this biological coding.<sup>18-21</sup> Coding of peptidic libraries by nucleic acids was suggested by Brenner and Lerner,<sup>39,103</sup> and further developed by others.<sup>38</sup> Coding by nucleic acids has several disadvantages, and also some unique advantages. Novel chemistry is necessary to enable the independent syntheses of peptide and nucleotide sequences. However, this does not address the problem that nucleotide sequences are sensitive to many chemical reagents

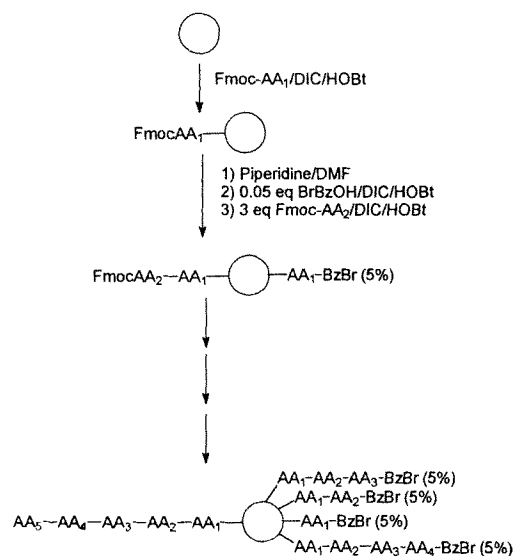


FIGURE 10 Scheme of synthesis of a library with a tag representing the compound's synthetic history.

useful for the synthesis of libraries of nonpeptidic compounds. Furthermore, long coding nucleotide sequences must be synthesized, adding to the time it takes to synthesize a library. The advantage of using nucleic acid coding is the ability to amplify the code by PCR. In this method, the code of all positive beads is amplified in individual PCR reactions, and the amplified nucleotide code sequences are digested and run in multiple lanes of a sequencing gel, each lane corresponding to a different bead. Thus, many compounds can be identified simultaneously by batchwise processing.

Recently, coding by halogenated derivatives of carboxylic acids followed by gas chromatographical analysis of photolytically cleaved and silylated mixtures of tag molecules was published.<sup>40,104</sup> This technique employs a "digital" code comprised of separate sets of coding structures for each position of the library, with coding subunits within each set defining the identity of the specific compound subunit. A single step of analysis is needed to identify the test compound. A similar technique for structural elucidation was described using mass spectrometry.<sup>82,99</sup> In these cases, in each step of synthesis, a fraction of available amino groups (in the case of a peptide) is capped, and the "history" of synthesis is documented. Once a positive bead is isolated, the test compound is cleaved together with all capped intermediates. The resulting mixture is then injected into a mass spectrometer and the sequence of the test compound is identified by reading the mass differences of individual molecules

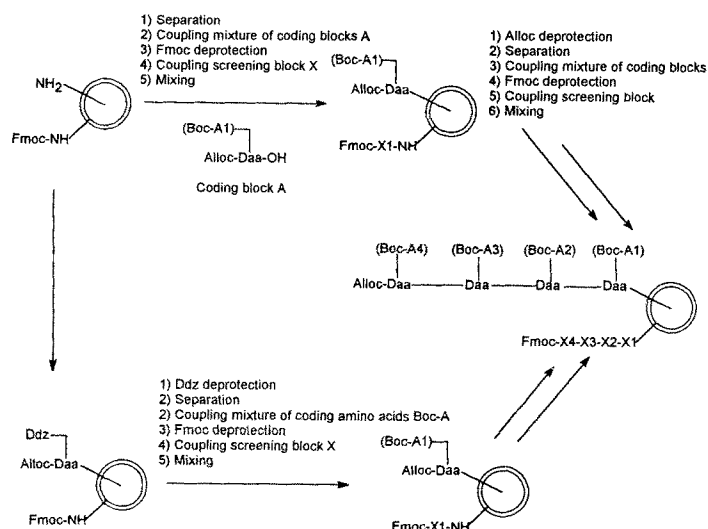


FIGURE 11 Two alternative syntheses of a library with an amino acid "bar code."

represented in the spectrum. If the capping is performed by a bromine containing carboxylic acid<sup>99</sup> (Figure 10), characteristic doublets of bromine containing molecules can simplify the identification of relevant masses in the spectrum. The principle of this method is similar to the "ladder" sequencing designed by Kent et al.,<sup>105-107</sup> which instead of building the ladder during synthesis builds it during Edman-like degradation, and is thus applicable only to peptidic compounds.

Encoding by peptide molecules was described for the identification of test compounds composed from unnatural amino acids<sup>71</sup> and nonpeptide chemical libraries or small organic libraries.<sup>69</sup> In the approach described initially, each step of synthesis of the nonsequencable structure is accompa-

nied by the attachment of one or several amino acids onto an independent attachment point on the solid phase particle—constructing a linear coding molecule. This "peptidic tag" could then be sequenced by the application of standard techniques (Edman degradation or mass spectroscopy—see Ref. 98 and references quoted therein), and the structure of the nonpeptidic molecule recalled. Alternatively, the peptidic tag can be built in such a way that it can be cleaved in one step of Edman degradation, followed by high performance liquid chromatography (HPLC) analysis as a "bar code" (Sepetov et al., unpublished). (For discussion of "digital" or "bar" coding see Ref. 40). Figure 11 shows alternative schemes of synthesis of the coding structure and Figure 12 illustrates the decoding

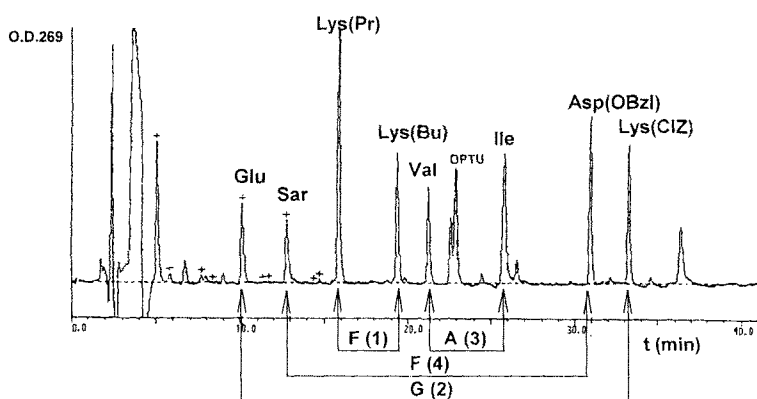
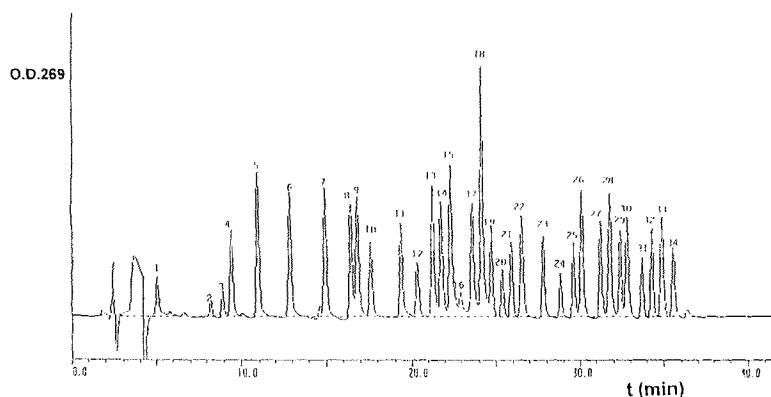


FIGURE 12 HPLC trace of bar code. Note that building block F is coded by a different amino acid doublet in position 1 [Lys (Propionyl) and Lys (Butyryl)] and 4 [Sar and Asp (OBzl)].



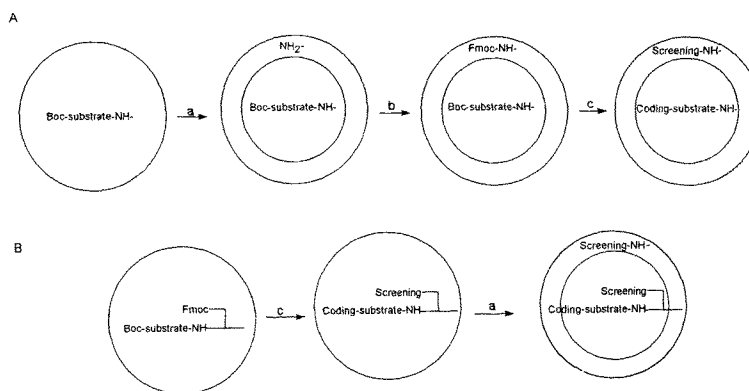
**FIGURE 13** HPLC trace of 33 phenylisothiohydantoin derivatives of coding amino acids. 1 Thr, 2 Gln, 3 Gly, 4 Dab (Acetyl), 5 Orn (Acetyl), 6 Ala, 7 Dap (Butyryl), 8 Dab (Isobutyryl), 9 Aib, 10 Orn (Butyryl), 11 Lys (Butyryl), 12 Dab (H), 13 Val, 14 Orn (H), 15 Norvaline, 16 DPTU, 17 Lys (H), 18 Dap (Caproyl), 19 Lys (Phenylacetyl), 20 Phe, 21 Orn (Caproyl), 22 Leu, 23 Orn (4-Chlorobenzoyl), 24 Lys (4-Chlorobenzoyl), 25 Orn (2-Naphthylacetyl), 26 Phe (4-Cl), 27 Dap (1-Adamantylcarbonyl), 28 2-Naphthylalanine, 29 Dab (4-Biphenylcarbonyl), 30 Dab (2,2-Diphenylacetyl), 31 Lys (1-Adamantylcarbonyl), 32 Orn (1-Adamantylacetyl), 33 Cyclohexylalanine, 34 Lys (1-Adamantylacetyl). Dab: diaminobutyric acid; Dap: diaminopropionic acid. Applied Biosystems ABI 120A Analyzer with a PTH-222 Brownlee column (PTH C-18 5 micron, 220 × 2.1 mm). HPLC buffers: (A) 0.01 M NaOAc, (B) acetonitrile; gradient: 0.0–0.4 min—8% B, 0.4–38.0 min—8–60% B, 38.0–40.0 min—60–90% B; flow rate 230  $\mu$ L/min.

of the structure of a test compound from a library in which four randomizations were performed, each coded by a doublet of amino acid derivatives.  $N$  amino acids used as a doublet can code for  $N \times (N - 1)/2$  structures (10 amino acids can code for 45 building blocks). To increase the number of amino acids available for coding, we have prepared a set of side-chain acylated diaminocarboxylic acids (Dap, Dab, Orn, Lys), the HPLC retention times of which (in the form of phenylisothiohydantoin) do not overlap with the derivatives of natural amino acids (as an example, see Figure 13). The presently available set (80 derivatives) can code for up to 130 million compounds for a library in which four randomizations are performed.

#### ELIMINATION OF POSSIBLE INTERACTION OF TARGET MACROMOLECULE WITH CODING STRUCTURE

The screening of coded libraries in on-bead binding assays brings about the risk that the target molecule may interact with the coding structure rather than with the test compound, or that the coding mole-

cules may perturb the test compound in a manner favoring (or disfavoring) interaction with the target molecule. This problem can be addressed in at least four ways. One approach is to resynthesize both code and test structures from positive beads independently and determine their activity. Another is to use a method where any single coding species represents only a small fraction of the material on the bead ( $\sim 1$ –5%). A mixture of molecules for coding can serve a similar purpose. This mixture, however, must provide unambiguous sequencing data. The coupling of three amino acid mixtures in five steps will produce 243 individual pentapeptides, the simultaneous sequencing of which will give only three phenylisothiohydantoin in every cycle. The fourth approach is to arrange the coding and screening structures in such a way that only the test compound is available for interaction with the target molecule. We exploited the properties of a polymeric solid phase resin, polyoxyethylene grafted polystyrene (TentaGel), to accomplish the physical segregation of compounds. The interior of TentaGel beads is not available to macromolecular targets due to the microporous structure of this polymer. This carrier was first modified with a protected peptide sequence that could be cleaved by a high molecular weight rea-



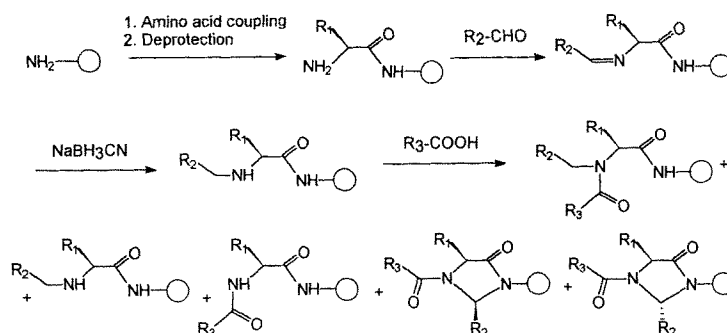
**FIGURE 14** Scheme of bead shaving. (A) Shaving performed before library synthesis; (B) “shaving after.” (a) Enzyme incubation; (b) introduction of orthogonal protecting group; (c) library synthesis using alternating Boc and Fmoc strategy in each randomization step.

gent (reactive polymer or enzyme) that could only access the molecules on the “surface” of the solid phase support, thus exposing a free amino group on the surface. Free amino group was subsequently protected with an orthogonally cleavable protecting group. In this way it was possible to synthesize the test compound exclusively on the bead surface and the coding structure exclusively in the bead interior. We have used enzymatic “shaving” to produce this selective surface modification, using chymotrypsin as the shaving agent and the sequence Boc-Ala-Gly-Val-Phe-Gly- $\beta$ -Ala-Gly-TG as the substrate as diagrammed in Figure 14.<sup>108</sup> Synthesis of model ligands D-Phe-Pro-Arg-Pro-Gly (thrombin ligand) and Leu-His-Pro-Gln-Phe (streptavidin ligand) on the surface and hidden inside of the beads showed complete separation of accessible surface from interior part of the bead. The bead was identified positive when tested with streptavidin only when Leu-His-Pro-Gln-Phe sequence was displayed on its surface and negative when it was

synthesized inside of the bead. The same was true for the thrombin specific sequence.

## RESYNTHESIS OF POSITIVE COMPOUNDS

Using any of these structure-elucidation approaches, a system can be designed where the bottleneck is no longer identification of positive test compounds, but rather their resynthesis. Resynthesis is a necessary step in the overall process, since only after confirmation of activity in solution can one be certain that a positive compound has been identified. It is important to use the identical protocol during resynthesis as used in library synthesis since it is always possible that the active compound is not the primary synthetic product, but a side product. If this is the case, the side product will be present in the resynthesized preparation and can be isolated, characterized functionally, and its struc-



**FIGURE 15** Scheme of the synthesis of a small organic library based on the alkylation and acylation of amino acids and structure of the resultant side products.

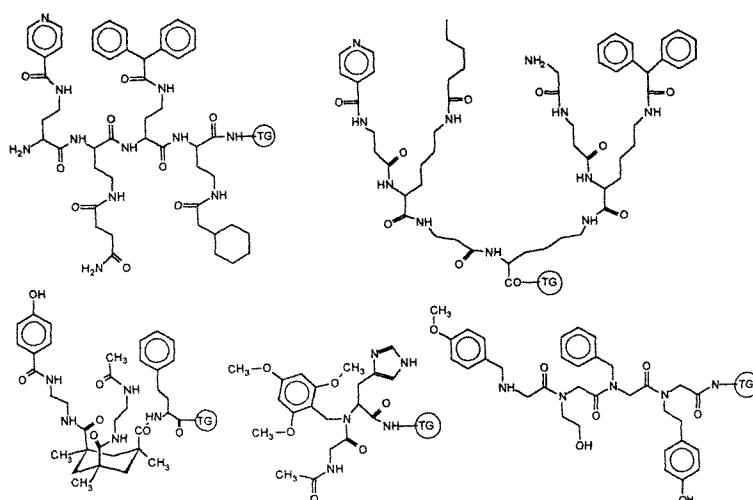


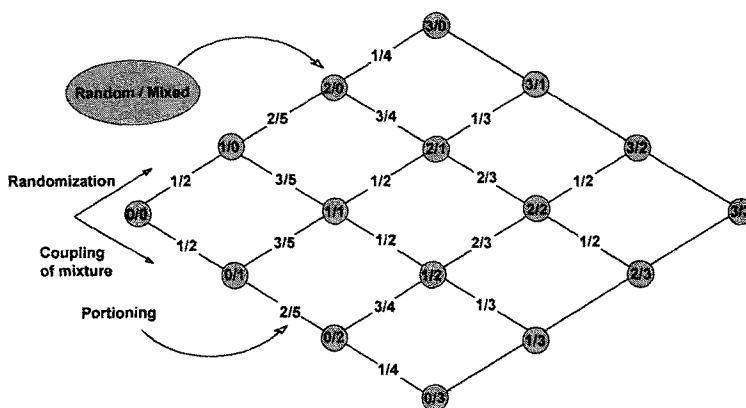
FIGURE 16 Nonpeptidic ligands of streptavidin.

ture elucidated. The less complete each step of the library synthesis, the more critical it is to duplicate the synthetic protocol during resynthesis. Since an identical resynthesis protocol is followed as for synthesis of the library, the ease of resynthesis is the same as that of library synthesis. An example of this reasoning is the result of the screening of library of alkylated and acylated amino acids, synthesized by using reactions that were not optimized for solid phase synthesis.<sup>102</sup> Positively reacting beads in a streptavidin assay were isolated. A mixture of products arising from reductive amination of the  $\alpha$ -

amino group of the amino acid attached to the resin was analyzed by mass spectroscopy and the composition of building blocks used for the synthesis of these beads was identified. Synthesis of the same mixture was performed in larger scale and all products were purified, their structure was established (including the product of unexpected cyclization—Figure 15), and binding affinities determined. The structural motif for streptavidin binding found in this library is given in Figure 16,<sup>102</sup> together with structures identified as streptavidin ligands in other libraries.

Table I Numbers of Beads (and Weight of Resin—130  $\mu\text{m}$  Bead Size) Required for Testing Full Representation of Peptides of Various Length Using (A) One Bead–One Peptide Approach and (B) Library of Libraries (Three Amino Acid Motif) Approach

Length	One-Bead–One-Peptide (A)		Library of Libraries (B)	
	Number of Beads	Resin Amount	Number of Beads	Resin Amount
3	8,000	8 mg	8,000	8 mg
4	160,000	160 mg	32,000	32 mg
5	3,200,000	3.2 g	80,000	80 mg
6	64,000,000	64 g	160,000	160 mg
7	1,280,000,000	1.28 kg	280,000	280 mg
8	25,600,000,000	25.6 kg	448,000	448 mg
9	512,000,000,000	512 kg	672,000	672 mg
10	10,240,000,000,000	10.2 t	960,000	960 mg
11	204,800,000,000,000	204 t	1,320,000	1.32 g
12	4,096,000,000,000,000		1,760,000	1.76 g
13	81,920,000,000,000,000		2,288,000	2.29 g
14	1,638,000,000,000,000,000		2,912,000	2.91 g
15	32,770,000,000,000,000,000		3,640,000	3.64 g



**FIGURE 17** Scheme of the synthesis of “library of libraries” with a tripeptide motif in a hexapeptide frame. The polymeric carrier is split in the ratio  $n_{Rr}:n_{Mr}$  (numbers located on the lines connecting circles denote the fraction of carrier undergoing the specified operation) ( $n_{Rr}$ : number of remaining randomization steps;  $n_{Mr}$ : number of remaining steps in which a mixture of amino acids is to be coupled) based on the status  $n_R/n_M$  (number located in the shaded circles;  $n_R$ : number of randomizations performed on the carrier;  $n_M$ : number mixtures to be coupled). At the beginning of the synthesis  $n_R/n_M = 0/0$ , at the end  $n_R/n_M = 3/3$ . The lines pointing up designate a randomization step, the lines pointing down designate a mixture step. This scheme can be applied to any length of library and motif.

### ONE-BEAD-ONE-MOTIF LIBRARIES (“LIBRARY OF LIBRARIES”)

The issue of incomplete representation of all compound permutations in libraries of compounds containing more than 5 residues (see Table I) may arise. This is not as limiting as it might at first appear. The number of expected positive beads depends on the number of “critical residues” in the peptide sequence (or critical pharmacophores in a nonpeptidic structure), i.e., residues required for minimal observable binding. This number can be calculated according to the formula

$$n = (x)(P_f)[S/(A_n)^{n_{crit}}]$$

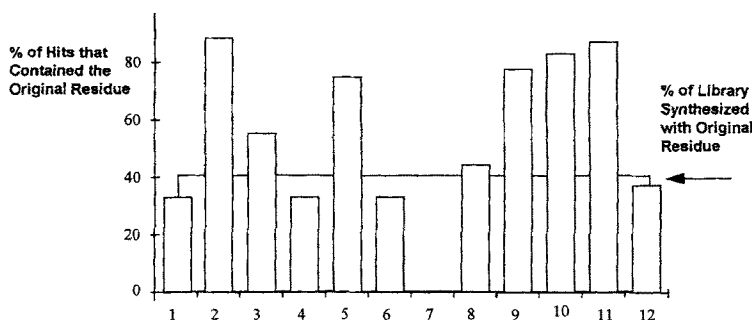
In this equation  $n$  is the number of expected positive hits,  $x$  is the number of different binding motifs,  $P_f$  is the “placement” factor, i.e., number of possible placements of each motif in the peptidic chain,  $S$  is the number of beads screened,  $A_n$  is number of amino acids (subunits) used for randomization, and  $n_{crit}$  is the number of critical residues. The number of positive hits depends on the number of beads tested, but it does not depend on the length of the library. Therefore, screening of even a very incomplete library, e.g., a fraction of library of decapeptides, can provide a reasonable number of positive beads if only 3–4 residues in the

sought-after peptide are critical for binding under the screening conditions used.

The validity of this formula was demonstrated using the example of an octapeptide library synthesized from L-amino acids and assays designed to identify ligands for the anti- $\beta$ -endorphin antibody, streptavidin and anti-insulin antibody.<sup>89</sup> The number of expected positive beads was very close to the number actually identified in each of the given assays. However, in certain assays, a smaller than expected number of positive beads was identified

**Table II** Sequences Identified as Positive and Specific in the Thrombin Screening (Sequences are Listed in the Order of Staining Intensity) (x Denotes Mixture of Amino Acids)

I	R	x	W	x	x
I	x	F	x	Y	x
I	F	x	W	x	x
I	x	F	W	x	x
I	x	F	W	x	x
I	R	W	x	x	x
L	R	x	W	x	x
I	R	W	x	x	x
L	R	Y	x	x	x
L	R	Y	x	x	x
L	R	Y	x	x	x
I	x	F	R	x	x
I	R	Y	x	x	x



**FIGURE 18** Results of sequencing 28 beads selected from the screening of a homolog library. Horizontal axis: Position of amino acid in the peptide chain.

(Štřop and Ostrem, unpublished). This phenomenon might be explained by irregularities in the mechanical properties of individual beads, different presentation of peptide on the bead surface, different surface density of the ligand, etc.

As mentioned above, the individual sequence identified in the peptide library with incomplete representation of all permutations is less valuable than information describing the motif responsible for binding or other biological activity. In order to verify this concept, we have designed a library that contains all possible motifs, (e.g., all tripeptide motifs in a hexapeptide library) in the framework of a library of given length.<sup>109</sup> In this library, each individual bead contains a multiplicity of peptides (each with the same motif). Synthesis of this library follows a simple scheme, illustrated in Figure 17. This library combines the principles of iterative library techniques<sup>9-11,110</sup> with those of the one-bead-one-compound approach<sup>15</sup>; however, it enables the screening of all possible motifs at the same time, rather than consecutive dipeptide motifs—which may be too short<sup>111</sup> or not appropriately localized in the peptide chain to identify the ligands. Each bead of the library of libraries represents a unique motif, the structure of which is determined only after its biological relevance is established. In this arrangement all tripeptide motifs in up to a 15-meric peptide can be synthesized and screened in a manageable library (see Table I).

Results from screening this type of library are illustrated in Table II. The library of tripeptide motifs in the framework of a linear hexapeptide composed of 19 natural amino acids (Cys excluded) was screened against thrombin. Peptides with the motif I/LRFW were detected. When tested in solution after resynthesis, the peptide IRFWA (sequence derived from motifs identified in screening) demonstrated an  $IC_{50}$  of  $2.3 \mu M$ . This primary motif can be used as a template for sec-

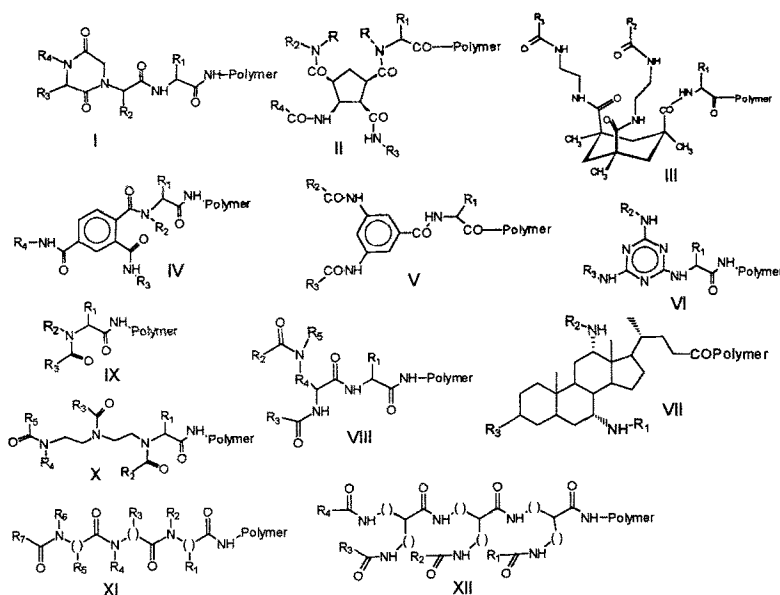
ondary peptide libraries (see below) or for the design of scaffold or other nonpeptidic libraries. It should be noted that only a small sample of the library (300,000 beads) was screened—demonstrating substantial savings in the consumption of reagents needed for both synthesis and screening, compared with a complete library of individual hexapeptides, which would be represented on 64 millions beads.

## STRATEGIES FOR OPTIMIZATION OF THE PRIMARY LIGAND

Ligands isolated from the primary screen often have low to moderate activity. The one-bead-one-structure process enables optimization to be pursued by designing and synthesizing secondary “optimization” libraries based on structural information from the initial hit. It is possible to synthesize extension libraries, keeping the lead intact but randomizing residues on both N- and C-terminus. It is also possible to synthesize analogue libraries, substituting a different set of building blocks at each position based on the subunits identified in the initial lead. When multiple pharmacophores or “motifs” are identified from the primary library, it is possible to link these with a variable length/variable composition linker library. Finally, in the case where cyclization is pursued as an optimization strategy, spacer subunits can be incorporated and the points of cyclization can be randomized around the active motif in a library format.

The anti-insulin monoclonal antibody system serves as an example to illustrate this technique. This antibody recognizes discontinuous epitopes of insulin. Several motifs were identified from primary libraries of various lengths. A secondary “optimization” library was constructed using one of the motifs as a template.<sup>60,81</sup> The secondary library





**FIGURE 19** Scaffolds used for the synthesis of libraries incorporating a multiplicity of amines, carboxylic acids, aldehydes, and halogen derivatives.

was screened under more stringent conditions to select ligands of higher affinity. Tertiary and quaternary libraries were synthesized and screened, the result of which was the identification of compounds containing the originally identified motifs with affinities comparable to that of native insulin<sup>67</sup> ( $IC_{50}$  35 nM for QSSVNH<sub>1</sub>PGWKYGF vs 14 nM for insulin). The identified peptides have no sequence homology with native insulin, indicating that the peptide sequence can mimic the discontinuous epitope.

Another example of this approach is the thrombin ligand selected from the biased (not completely randomized in all positions) primary library of pentapeptides. Ligands of the highest affinity had a  $K_i$  in the range of 20  $\mu$ M. The known motif D-Phe-Pro-Arg-Pro was used as a template for a secondary library (extension at C-terminus) and a nonapeptide ligand with  $K_i$  25 nM was identified.<sup>112</sup> This ligand is one order of magnitude weaker than hirulog, a molecule that is more than twice the size.<sup>113</sup>

We have also used a so-called homolog library to optimize a hit from an incomplete peptidic library. In each step of library synthesis, a fixed percentage (in our case, 40%) of the solid support is coupled to the residue corresponding to that in the initially discovered hit, while the remainder of the carrier is randomized using the standard procedure. After completion of coupling the carrier is recombined and the next, similarly biased coupling is performed. The portion of resin to which the original residue is coupled is a function of the

length of peptide and number of amino acids used in each randomization step. This is done to assure that the sample to be screened contains a reasonable number of compounds with the original sequence (e.g., 10 in 1 million). The likelihood of finding a bead having the original sequence is thus increased, as is the probability of finding sequences with one or more substitutions. Positive compounds are sequenced, and distribution of amino acids in each position is evaluated. Positions in which the frequency of a given amino acid is that at which the same amino acid was introduced during the synthesis (40%) can be considered unimportant for binding. Positions in which the frequency is significantly higher can be considered critical. More interestingly, positions in which the frequency of the original amino acid is significantly lower than 40% indicate that the original hit contained a nonoptimal residue in that position. An example of this approach is shown in Figure 18. A primary hit from a 12-mer peptidic library was used as the basis for the synthesis of a homolog library with 40% enrichment in each position. Twenty-eight positive beads from this library were sequenced and the frequency of the original amino acid found in each position was plotted. A clearly identifiable motif of amino acids originally presented in positions 2, 5, 9, 10, and 11 was detected. Positions 1, 4, 6, 8, and 12 can be replaced by many alternative amino acids. An additional important residue was found for position 7, in which the original hit contained a clearly nonoptimal amino acid.

**Table III Results from Screening for Several Targets in Selectide Corp**

Project	Primary Lead	Optimized Lead
gpIIb/IIIa antagonist	1–320 $\mu M$	na
Thrombin inhibitor	0.8–700 $\mu M$	25 nM
Factor Xa inhibitor	15 $\mu M$	300 pM
HER-2 ligand	40 $\mu M$	70 nM
IL-8 ligand	5 $\mu M$	na
NADPH oxidase inhibitor	5 $\mu M$	na
HIV-1 RNase inhibitor	0.3–200 $\mu M$	150 nM

In the illustrated case the binding of peptide synthesized based on this structural information improved from  $IC_{50}$  4 to 0.03  $\mu M$ .

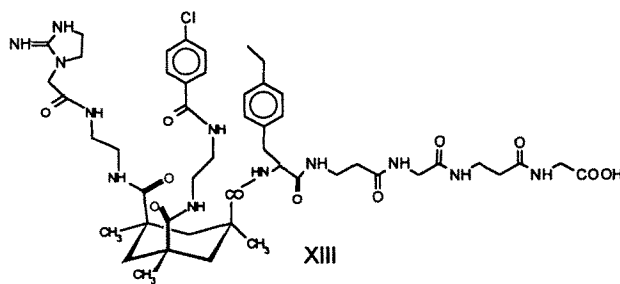
## DIVERSITY AND COMPLEXITY

Each combinatorial library represents a certain complexity and diversity. The complexity of a given library reflects the number of individual components regardless of their similarity. Conversely, diversity indicates the structural dissimilarity of compounds. It is obvious that if the structural and functional requirements for a small molecule to bind to a target molecule are not known, the greater the diversity explored the greater the chance to find a molecule that binds specifically. Both the conformational space mapped by the members of a library and the functional groups displayed in this space contribute to the overall diversity represented by this library. The conformational space is determined by the size and flexibility of library members, whereas functional groups accommodated in this space should represent all major interactions known to play a significant role in the interaction between two molecules. Flexible molecules will cover more conformational space; however, their flexibility will make them less likely to have high affinity because of entropy factors. To

compensate for lost entropy, the fit must be more optimal; nevertheless, there is no method available to quantitatively compare energy factors connected with reducing the freedom of conformational change with the energy loss when a certain conformation is not optimally fixed. One solution to this dilemma is intuition and experience of the experimenter. Future experimental data will shed more light on the appropriate balance. We have taken a pragmatic approach in this respect, in that we have designed and synthesized libraries of small, rigid, and compact structures, as well as those that exhibit intermediate flexibility, to those that are based on the linear presentation of building blocks and are very flexible. Examples of these libraries are shown in Figure 19.

## NONPEPTIDIC LIBRARIES

The discussion of complexity and diversity relates directly to the synthesis and use of nonpeptidic, or low molecular weight organic compound, libraries. The potential for small organic libraries is enormous. Thousands of building blocks are now available, eliminating the limitations imposed by the 50 or so commercially available amino acids. For instance, a small organic trimer with 100 possible subunits in each coupling step will generate a di-



**FIGURE 20** Structure of a thrombin inhibitor selected from nonpeptide (scaffold) library.

versity of  $10^6$  and the synthesis of such libraries is feasible using current synthetic methods. Ligands isolated from such libraries may have a better chance of having oral activity as well as intracellular activity, resulting in a lead with characteristics closer to that of a therapeutic agent. An example of a simple nonpeptidic library constructed using subunits selected from readily available amino acids, aldehydes, and carboxylic acids was already mentioned.<sup>102,114</sup> We have used alkylation and acylation reactions for the construction of model libraries<sup>69</sup> as well as libraries screened for pharmaceutically relevant targets. Construction of heterocyclic molecules in solid phase is very attractive way of library preparation.<sup>5,8,115</sup> The addition of amines to isocyanates<sup>116</sup> and Wittig reaction followed by Michael addition<sup>74</sup> were also used for nonpeptidic library construction. Several laboratories focused recently on the development of synthetic methods for formation of carbon–carbon or carbon–heteroatom bonds in solid phase, which can be used for the library construction.<sup>8,13,117–119</sup>

An alternative to the use of various organic reactions for linking building blocks containing different functional groups is the application of well-developed techniques for amide bond formation and use of vast diversity available with amines and carboxylic acids.<sup>67,75,76</sup> These building blocks can be attached to a great variety of “scaffolds” (see, e.g., Refs. 120 and 121), ranging from linear and flexible structures, allowing the exploration of a wide conformational space (Figure 19, structures VIII–XII), to cyclic and rigid structures, mapping a smaller space, but potentially providing higher affinity ligands (Figure 19 structures I–VII). We have prepared and tested libraries based on all of the illustrated scaffolds. Figure 20 shows structures inhibiting enzymatic activity of thrombin as an example of ligands found in these libraries (Kociš et al., in preparation). Inhibitory activity of compound XIII was at the level of  $4 \mu\text{M}$ , which is equivalent to that found in the primary peptide libraries.

## CONCLUSION

One-bead–one-structure libraries can be viewed as a real-life situation of a library to which a person (receptor) goes with the wish to find (affinity) a certain book (ligand). By browsing through the titles or searching in the catalogue (incubation), he finds the book (binding) and carries it home (retrieval), not knowing its content. Then he reads

it (sequencing), or he reads only the critique (decoding).

One-bead–one-structure, or one-bead–one-motif libraries represent a powerful tool in the drug discovery process. The applicability of these techniques has been expanded broadly from the initial application to peptide and other oligomeric libraries, to libraries of small organic structures. The applicability will continue to expand in scope, with the reduction to practice in solid phase of many standard organic reactions—potentially leading to a point at which virtually any compound that can be synthesized in solution could become the basis for a library of compounds. We have reviewed here some techniques used for building one-bead–one-structure libraries, or one-bead-one-motif libraries, and some results obtained from their use. Initial data indicate the utility of these libraries, and Table III shows several results of primary screening and optimization of ligands from some projects run in our laboratories. Binding experiments were followed by functional assays and the most successful compounds from the factor Xa inhibitor project have entered preclinical evaluation, demonstrating the potential of the one-bead-one-structure approach to lead generation.

The authors are indebted to all scientists in Selectide Corp., who made the realization of all ideas mentioned in this review possible. This work was partially supported by the NCDDG Cooperative Agreement (CA57723).

## REFERENCES

1. Geysen, M. H., Meloen, R. H. & Barteling, S. J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 3998–4002.
2. Houghten, R. A. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 5131–5135.
3. Frank, R. & Döring, R. (1988) *Tetrahed. Lett.* **44**, 6031–6040.
4. Fodor, S. P. A., Leighton, R. J., Pirrung, M. C., Stryer, L., Tsai, L. A. & Solas D. (1991) *Science* **251**, 767–773.
5. DeWitt, S. H., Kiely, J. S., Stankovic, C. J., Schroeder, M. C., Reynolds Cody, D. M. & Pavia, M. R. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 6909–6913.
6. Frank, R. (1992) *Tetrahedron* **48**, 9217–9232.
7. Krchňák, V. & Vágner, J. (1990) *Peptide Res.* **3**, 182–193.
8. Bunin, B. A., Plunkett, M. J. & Ellman, J. A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4708–4712.
9. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. & Cuervo, J. H. (1991) *Nature* **354**, 84–86.

10. Geysen, M. H., Rodda, H. M. & Mason, T. J. (1986) *Mol. Immunol.* **23**, 709–715.
11. Owens, R. A., Gesellchen, R. D., Houchins, B. J. & DiMarchi, R. D. (1991) *Biochem. Biophys. Res. Commun.* **181**, 402–408.
12. Blake, J. & Litz-Davis, L. (1992) *Bioconj. Chem.* **3**, 510–513.
13. Zuckermann, R. N., Martin, E. J., Spellmeyer, D. C., Stauber, G. R., Simon, R. J., Banville, S. C., Brown, E. G., Wang, L., Richter, L. S. & Moos, W. H. (1994) *J. Med. Chem.* **37**, 2678–2685.
14. Houghten, R. A., Appel, J. R., Blondelle, S. E., Cuervo, J. H., Dooley, C. T. & Pinilla, C. (1992) *BioTechniques* **13**, 412–421.
15. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. & Knapp, R. J. (1991) *Nature* **354**, 82–84.
16. Pinilla, C., Appel, J. R., Blanc, P. & Houghten, R. A. (1992) *BioTechniques* **13**, 901–905.
17. Pinilla, C., Appel, J. R. & Houghten, R. A. (1994) *Biochem. J.* **301**, 847–853.
18. Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
19. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
20. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990) *Science* **249**, 404–406.
21. Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. & Cesareni, G. (1991) *J. Mol. Biol.* **222**, 301–310.
22. Cull, M. G., Miller, J. F. & Schatz, P. J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 1865–1869.
23. Fassina, G., Lebl, M. & Chaiken, I. (1988) *Collect. Czech. Chem. Commun.* **53**, 2627–2634.
24. Tjoeng, F. S., Towery, D. S., Bullock, J. W., Whipple, D. E., Fok, K. F., Williams, M. H., Zupec, M. E. & Adams, S. P. (1990) *Int. J. Peptide Protein Res.* **35**, 141–146.
25. Kramer, A., Volkmer-Engert, R., Malin, R., Reincke, U. & Schneider-Mergener, J. (1993) *Peptide Res.* **6**, 314–319.
26. Rutter, W. J. & Santi, D. V. (1991) U.S. Patent 5,010,175.
27. Birkett, A. J., Soler, D. F., Wolz, R. L., Bond, J. S., Wiseman, J., Berman, J. & Harris, R. B. (1991) *Biochemistry* **196**, 137–143.
28. Petithory, J. R., Masiarz, E. R., Kirsch, J. E., Santi, D. V. & Malcolm, B. A. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 11510–11514.
29. Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C. & Santi, D. V. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4505–4509.
30. Frank, R. (1994) *Innovation and Perspectives in Solid Phase Synthesis, Vol. 3*, Epton, R., Ed., Mayflower Worldwide Ltd., Birmingham, pp. 509–512.
31. Cass, R., Dreyer, M. L., Giebel, L. R., Hudson, D., Johnson, C. R., Ross, M. J., Schaeck, J. & Shoemaker, K. R. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds. ESCOM, Leiden, pp. 975–977.
32. Lam, K. S. & Lebl, M. (1992) *Immunomethods* **1**, 11–15.
33. Salmon, S. E., Lam, K. S., Lebl, M., Kandola, A., Khattri, P. S., Wade, S., Pátek, M., Kočiš, P., Krchňák, V., Thrope, D. & Felder, S. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 11708–11712.
34. Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A. & Schreiber, S. L. (1993) *J. Am. Chem. Soc.* **115**, 12591–12592.
35. Jayawickreme, C. K., Graminski, G. E., Quillan, J. M. & Lerner, M. R. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 1614–1618.
36. Kassarijan, A., Schellenberger, V. & Turck, C. W. (1993) *Peptide Res.* **6**, 129–133.
37. Meldal, M., Svendsen, I., Breddam, K. & Auzanneau France-I. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3314–3318.
38. Needels, M. C., Jones, D. G., Tate, E. H., Heinkel, G. L., Kochersperger, L. M., Dower, W. J., Barrett, R. W. & Gallop, M. A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10700–10704.
39. Nielsen, J., Brenner, R. A. & Janda, K. D. (1993) *J. Am. Chem. Soc.* **115**, 9812–9813.
40. Ohlmeyer, M. H., Swanson, R. N., Dillard, L. W., Reader, J. C., Asouline, G., Kobayashi, R., Wigler, M. & Still, W. C. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 10922–10926.
41. Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. (1992) *Nature* **355**, 564–566.
42. Tuerk, C. & Gold, L. (1990) *Science* **249**, 505–510.
43. Ellington, A. D. & Szostak, J. W. (1992) *Nature* **355**, 850–852.
44. Moos, W. H., Green, G. D. & Pavia, M. R. (1993) *Annual Reports in Medicinal Chemistry*, Bristol, J. A. Ed., Academic Press, San Diego, CA, pp. 315–324.
45. Scott, J. K. (1992) *TIBS* **17**, 241–245.
46. Houghten, R. A. (1993) *Trends Genet.* **9**, 235–239.
47. Pavia, M. R., Sawyer, T. K. & Moos, W. H. (1993) *Bioorg. Med. Chem. Lett.* **3**, 387–396.
48. Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A. & Gordon, E. M. (1994) *J. Med. Chem.* **37**, 1233–1251.
49. Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A. & Gallop, M. A. (1994) *J. Med. Chem.* **37**, 1385–1401.
50. Furka, A., Sebastyen, F., Asgedom, M. & Dibo, G. (1988) poster presented at 14th International Congress of Biochemistry, Prague.
51. Furka, A., Sebastyen, F., Asgedom, A. & Dibo, G. (1988) poster presented at Xth International Symposium on Medicinal Chemistry, Budapest.
52. Furka, A., Sebastyen, F., Asgedom, M. & Dibo, G. (1991) *Int. J. Peptide Protein Res.* **37**, 487–493.

53. Sebestyén, F., Dibó, G. & Furka, A. (1993) in *Peptides 1992, Proceedings of the 22nd European Peptide Symposium*, Schneider, C. H. & Eberle, A. N. Eds., ESCOM, Leiden, pp. 63–64.
54. Sebestyén, F., Dibó, G., Kovács, A. & Furka, A. (1993) *Bioorg. Med. Chem. Lett.* **3**, 413–418.
55. Furka, A., Sebestyén, F. & Campian, E. (1994) in *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds., ESCOM, Leiden, pp. 986–988.
56. 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**, 419–424.
57. Ator, M. A., Beigel, S., Dankanich, T. C., Echols, M., Gainor, J. A., Gilliam, C. L., Gordon, T. D., Koch, D., Koch, J. F., Kruse, L. I., Morgan, B. A., Krupinski-Olsen, R., Siahaan, T. J., Singh, J. & Whipple, D. A. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds., ESCOM, Leiden, pp. 1012–1015.
58. Wu, J., Ma, Q. N. & Lam, K. S. (1994) *Biochemistry* **33**, 14825–14833.
59. Fassina, G., Bellitti, M. R. & Cassani, G. (1994) *Protein Peptide Lett.* **1**, 15–18.
60. Lebl, M., Lam, K. S., Kociš, P., Krchňák, V., Pátek, M., Salmon, S. E. & Hruby, V. J. (1993) *Peptides 1992, Proceedings of the 22nd European Peptide Symposium*, Schneider, C. H. & Eberle, A. N. Eds., ESCOM, Leiden, pp. 67–69.
61. Kociš, P., Krchňák, V. & Lebl, M. (1993) *Tetrahed. Lett.* **34**, 7251–7252.
62. Lebl, M., Pátek, M., Kociš, P., Krchňák, V., Hruby, V. J., Salmon, S. E. & Lam, K. S. (1993) *Int. J. Peptide Protein Res.* **41**, 201–203.
63. Zuckermann, R. N., Kerr, J. M., Siani, M. A. & Banville, S. C. (1992) *Int. J. Peptide Protein Res.* **40**, 497–506.
64. Saneii, H. H. & Shannon, J. D. (1994) *Innovation and Perspectives in Solid Phase Synthesis*, Epton, R., Ed., Mayflower Worldwide Ltd., Birmingham, pp. 335–338.
65. Saneii, H. H., Shannon, J. D., Miceli, R. M., Fischer, H. D. & Smith, C. W. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A. Eds., ESCOM, Leiden, pp. 1018–1020.
66. Barták, Z., Bolf, J., Kalousek, J., Mudra, P., Pavlík, M., Pokorný, V., Rinnová, M., Vobúrka, Z., Ženišek, K., Krchňák, V., Lebl, M., Salmon, S. E. & Lam, K. S. (1994) *Methods, A Companion to Methods in Enzymology*, **6**, 432–437.
67. Lebl, M., Krchňák, V., Sepetov, N. F., Nikolaev, V., Stierandová, A., Šafář, P., Seligman, B., Štrop, P., Lam, K. S. & Salmon, S. E. (1994) *Innovation and Perspectives in Solid Phase Synthesis*, Vol. 3, Epton, R., Ed., Mayflower Worldwide Ltd., Birmingham.
68. Holmes, C. P. & Rybak, C. M. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds., ESCOM, Leiden, pp. 992–994.
69. Nikolaiev, V., Stierandová, A., Krchňák, V., Seligman, B., Lam, K. S., Salmon, S. E. & Lebl, M. (1993) *Peptide Res.* **6**, 161–170.
70. Simon, R. J., Kaina, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Simon, N. G., Wang, L., Rosenberg, S., Marlowe, O. 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**, 9367–9371.
71. Kerr, J. M., Banville, S. C. & Zuckermann, R. N. (1993) *J. Am. Chem. Soc.* **115**, 2529–2531.
72. Zuckermann, R. N. (1993) *Curr. Opin. Struct. Biol.* **3**, 580–584.
73. Bunin, B. A. & Ellman, J. A. (1992) *J. Am. Chem. Soc.* **114**, 10997–10998.
74. Chen, C., Randall, A. L. A., Miller, B. R., Jones, D. A. & Kurth, M. J. (1994) *J. Am. Chem. Soc.* **116**, 2661–2662.
75. Lebl, M., Krchňák, V., Šafář, P., Stierandová, A., Sepetov, N. F., Kociš, P. & Lam, K. S. (1994) *Techniques in Protein Chemistry*, Vol. V, Crabb, J. W., Ed., Academic Press, San Diego, CA, pp. 541–548.
76. Lebl, M., Krchňák, V., Stierandová, A., Šafář, P., Kociš, P., Nikolaev, V., Sepetov, N. F., Ferguson, R., Seligman, B., Lam, K. S. & Salmon, S. E. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds., ESCOM, Leiden, pp. 1007–1008.
77. Stevanovic, S. & Jung, G. (1993) *Anal. Biochem.* **212**, 212–220.
78. Andrews, P. C., Boyd, J., Ogorzalek, Loo, R., Zhao, R., Zhu, C. O., Grant, K. & Williams, S. (1994) *Techniques in Protein Chemistry*, Vol. V, Crabb, J. W., Ed., Academic Press, San Diego, CA, pp. 485–492.
79. Krchňák, V., Vágner, J., Šafář, P. & Lebl, M. (1988) *Collect. Czech. Chem. Commun.* **53**, 2542–2548.
80. Kaiser, E., Colescott, R. L., Bossinger, C. D. & Cook, P. I. (1969) *Anal. Biochem.* **34**, 595–598.
81. Lam, K. S., Lebl, M., Krchňák, V., Lake, D. F., Smith, J., Wade, S., Ferguson, R., Ackerman-Berrier, M. & Wertman, K. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A. Eds., ESCOM, Leiden, pp. 1003–1004.
82. Youngquist, S. R., Fuentes, G. R., Lacey, M. P. & Keough, T. (1994) *Rapid Commun. Mass Spectrom.* **8**, 77–81.
83. Smith, M. H., Lam, K. S., Hersh, E. M., Lebl, M. & Grimes, W. J. (1994) *Mol. Immunol.*, **31**, 1431–1437.
84. Lam, K. S., Lebl, M., Wade, S., Stierandová, A.,

- Khattari, P. S., Collins, N. & Hruba, V. J. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds., ESCOM, Leiden, pp. 1005–1006.
85. Lam, K. S., Zhao, Z.-G., Wade, S., Krchňák, V. & Lebl, M. (1994) *Drug Develop. Res.* **33**, 157–160.
  86. Wennemers, H. & Still, W. C. (1994) *Tetrahed. Lett.* **35**, 6413–6416.
  87. Borchardt, A. & Still, C. W. (1994) *J. Am. Chem. Soc.* **116**, 373–374.
  88. Boyce, R., Li, G., Nestler, H. P., Suenaga, T. & Still, W. C. (1994) *J. Am. Chem. Soc.* **116**, 7955–7956.
  89. Lam, K. S. & Lebl, M. (1994) *Methods, A Companion to Methods in Enzymology*, **6**, 372–380.
  90. Lebl, M., Krchňák, V., Salmon, S. E. & Lam, K. (1994) *Methods, A companion to Methods in Enzymology*, **6**, 381–387.
  91. Lam, K. S. & Salmon, S. E. (1991) U.S.A. Patent Application **5**, 382–513.
  92. Wiesmuller, K. H., Beck-Sickinger, A. G., Ihlenfeldt, H. G., Wieland, H. A., Udaka, K., Walden, P. & Jung, G. (1994) *Peptides, Proceedings of the 13th American Peptide Symposium*, Hodges, R. S. & Smith, J. A., Eds., ESCOM, Leiden, pp. 998–1000.
  93. Patek, M. & Lebl, M. (1991) *Tetrahed. Lett.* **32**, 3891–3894.
  94. Maeji, N. J., Bray, A. M. & Geysen, H. M. (1990) *J. Immunol. Methods* **134**, 23–33.
  95. Bray, A. M., Maeji, J. N., Valerio, R. M., Campbell, R. A. & Geysen, M. H. (1991) *J. Org. Chem.* **56**, 6659–6666.
  96. Bray, A. M., Maeji, J. N., Jhingran, A. G. & Valerio, R. M. (1991) *Tetrahed. Lett.* **32**, 6163–6166.
  97. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. (1991) *Nature* **351**, 290–296.
  98. Sepetov, N. F., Issakova, O. L., Lebl, M., Swiderek, K., Stahl, D. C. & Lee, T. D. (1993) *Rapid Commun. Mass Spectrom.* **7**, 58–62.
  99. Sepetov, N. F., Issakova, O. L., Krchňák, V. & Lebl, M. (1992) U.S. Patent Application 07/939,811.
  100. Biemann, K. (1990) *Methods Enzymol.* **193**, 455–479.
  101. Brummel, C. L., Lee, I. N. W., Zhou, Y., Benkovic, S. J. & Winograd, N. (1994) *Science* **264**, 399–402.
  102. Staňková, M., Issakova, O., Sepetov, N. F., Krchňák, V., Lam, K. S. & Lebl, M. (1994) *Drug Develop. Res.* **33**, 146–156.
  103. Brenner, S., Lerner, R. A. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5381–5383.
  104. Nestler, H. P., Bartlett, P. A. & Still, W. C. (1994) *J. Org. Chem.* **59**, 4723–4724.
  105. Wang, R., Chait, B. T. & Kent, S. B. H. (1994) *Techniques in Protein Chemistry V*, Vol. V, Crabb, J. V., Ed., Academic Press, San Diego, CA, pp. 19–26.
  106. Chait, B. T., Wang, R., Beavis, R. C. & Kent, S. B. H. (1993) *Science* **262**, 89–92.
  107. Wang, R., Chait, B. T. & Kent, S. B. H. (1993) *Techniques in Protein Chemistry*, Vol. IV, Angelletti, R. H., Ed., Academic Press, San Diego, CA, pp. 471–478.
  108. Vágner, J., Krchnak, V., Sepetov, N. F., Strop, P., Lam, K. S., Barany, G. & Lebl, M. (1994) *Innovation and Perspectives in Solid Phase Synthesis*, Vol. 3, Epton, R., Ed., Mayflower Worldwide Ltd., Birmingham, pp. 347–352.
  109. Sepetov, N. F., Krchňák, V., Staňková, M., Wade, S., Lam, K. S. & Lebl, M. (1994) *Proc. Natl. Acad. Sci. USA*, in press.
  110. Geysen, M. H. (1993) *Peptide Chemistry 1992, Proceedings of the 2nd Japanese Symposium on Peptide Chemistry*, Yanaihara, N., Ed., ESCOM, Leiden, pp. 3–6.
  111. Eichler, J. & Houghten, R. A. (1993) *Biochemistry* **32**, 11035–11041.
  112. Strop, P., Chen, C., Haney, K., Spoonamore, J., Ostrem, J., Stierandová, A., Šafář, P., Krchňák, V., Kociš, P., Sepetov, N. F., Cabel, D., Latif, F. A. & Lebl, M. (1993) poster at EMBO/IRBM Workshop on Molecular Repertoires and Methods of Selection, Gubio, Italy, September 26–October 1.
  113. Maraganore, J. M., Bourdon, P., Jablonski, J., Ramachandran, K. L. & Fenton, J. W., II (1990) *Biochemistry* **29**, 7095–7101.
  114. Flegelová, Z., Krchňák, V., Sepetov, N. F., Staňková, M., Issakova, O., Cabel, D., Lam, K. S. & Lebl, M. (1994) *Peptides 94, Proceedings of the 23rd European Peptide Symposium*, Maia, H. S., Ed., in press.
  115. Pátek, M., Drake, B. & Lebl, M. (1994) *Tetrahed. Lett.*, submitted.
  116. Cho, C. Y., Moran, E. J., Cherry, S. R., Stephans, J. C., Fodor, S. P. A., Adams, C. L., Sundaram, A., Jacobs, J. W. & Schultz, P. G. (1993) *Science* **261**, 1303–1305.
  117. Deshpande, M. S. (1994) *Tetrahed. Lett.*, **35**, 5613–5614.
  118. Kuhn, H. & Neumann, W. P. (1994) *Synlett* **2**, 123–124.
  119. Pei, Y. & Moos, W. H. (1994) *Tetrahed. Lett.* **35**, 5825–5828.
  120. Pátek, M., Drake, B., Lebl, M. (1994) *Tetrahed. Lett.*, **35**, 9169–9172.
  121. Šafář, P., Stierandová, A. & Lebl, M. (1994) in Maia H. S. (ed) *Peptides 94, Proceedings of the 23rd European Peptide Symposium*, in press.
  122. Kania, R. S., Zuckermann, R. N. & Marlowe, C. K. (1994) *J. Am. Chem. Soc.* **116**, 8835–8836.