3

SOLID-PHASE PEPTIDE SYNTHESIS, LEAD GENERATION, AND OPTIMIZATION

BRUCE SELIGMANN
SIDDCO, Tucson, Arizona

MICHAEL LEBL
Trega Biosciences, San Diego, California

KIT S. LAM
Arizona Cancer Center, Tucson, Arizona

Drug discovery has been undergoing revolutionary changes for the last 15 years. The first change was in the shift from whole animal testing of compounds, to mechanistic based in vitro screening. This was based on new knowledge about molecular mechanisms brought about by advances in biochemistry. Some companies began to adapt in vitro mechanistic assays for "high throughput" screening, on the order of 10,000 to 20,000 compounds per year. At the same time, medicinal chemists began to recognize the advances being made in the understanding of the structure of enzymes and proteins, as well as to use x-ray to define the structure of the small molecules they had synthesized which exhibited desired biological activity. This approach enabled chemists to understand how structure might relate to activity in order to design even more active molecules. Computational chemistry came of age in the pharmaceutical industry, focused initially on methods of obtaining structure activity relationships (SAR) between series of active molecules, and then

evolving into de novo rational design efforts using structural information about the target molecule. The explosion of molecular biology went hand-in-hand with this evolution of rational design and mechanistic approaches by providing the means to obtain sufficient quantities of target proteins to both enable high volume assay as well as structural studies on otherwise inaccessible target molecules.

The extraordinary growth of combinatorial chemistry beginning some five years ago has led to yet another revolution, focused directly on the approach used by the medicinal chemist to generate lead compounds and to optimize their activity. However, the impact of combinatorial chemistry does not end with the successful discovery of therapeutic agents. It is now entering another phase even before its impact on lead generation and optimization has been fully developed or implemented by most companies, and even though the solid phase chemistry for non-peptide synthesis remains largely undeveloped. This new role of combinatorial chemistry is being spawned by the rush of genomics information which will produce an incredible number of new potential targets each year, more than overwhelming the capacity of the pharmaceutical industry. What is crucial is no longer the sequencing of the human and other genomes, but instead the identification of targets and their rapid evaluation to enable the industry to identify those with the greatest potential for impact on disease. To-date the use of transgenic experiments, and to a lesser extent, the use of antibodies, have served as the mainstay for determining the role of new, or even somewhat older target proteins in disease. However, combinatorial chemistry in the form of antibody display libraries, peptide display libraries, and synthetic peptide combinatorial libraries offer state-of-the-art tools to the pharmaceutical industry which can be exploited to identify antagonists/inhibitors and agonists of new target molecules which can in turn be used in vivo to validate the potential of exploiting a new target to address specific diseases. With time, non-peptide combinatorial chemistry may similarly be used to address this need, but currently such libraries are limited by the state of synthetic chemistry knowledge. While the potential for target validation is just now beginning to be exploited, results reviewed here clearly indicate that screening synthetic peptide libraries can produce leads for a large variety of targets of therapeutic interest, provide extensive databases of SAR to the medicinal chemist, and when libraries are designed based on the initial leads or other information, provide a valuable tool to rapidly optimize the activity of molecules. The envelope of peptide chemistry has been pushed to enable the reliable, high-purity synthesis of large numbers of peptides containing available proteinogenic and nonproteinogenic amino acids, amines, and carboxylic acids as linear, cyclic, and branched structures in order to produce the greatest diversity among library components and between libraries.

This chapter focuses on synthetic peptide/peptidomimetic combinatorial library synthesis, lead generation, and optimization. Numerous reviews[1-6] have discussed the various formats used to synthesize combinatorial libraries: the

synthesis of medium to large numbers of compounds by fixed array combinatorial synthesis, as described by Geysen[7]; the iterative synthesis, assay, resynthesis, and reassay of mixtures described in 1991 by Houghten[8]; the split-and-mix, one-compound–one-bead approach to synthesizing large numbers of single compounds described the same year by Lam et al.[9] (1991); and the current automated parallel synthesis of large numbers of compounds and rational limited deconvolution of relatively small mixtures.[10] Types of assays by which combinatorial libraries have been screened have also been described and reviewed extensively, with the conclusion that combinatorial libraries are amenable to screening with every high throughput assay currently known.

This overview describes specific novel peptide chemistry and the assay results that reveal the potential brought to medicinal chemistry through peptide combinatorial library synthesis and assay. The review of results is divided according to type of target, rather than type of combinatorial approach. The reader is thus encouraged to focus on what has been achieved as it applies to drug discovery regardless of the specific combinatorial chemistry synthetic/screening format. Combinatorial chemistry may reverse the dogma that has developed over the years regarding the therapeutic utility of peptides. Even the authors are divided over whether the ability to synthesize and assay large numbers of peptide and peptidomimetic analogs will produce useful therapeutic agents with acceptable pharmacological properties for a large number of targets.

COMBINATORIAL APPROACHES TO LIBRARY DESIGN, SYNTHESIS, AND SCREENING

There is no universally accepted definition of combinatorial chemistry. For the purpose of this chapter it is defined to encompass the synthesis of compounds from sets of subunit and chemical reactions used in one or more reaction steps. This definition includes both solution and solid-phase synthetic reactions, the synthesis of compounds through several sequential reaction steps in which the same or different sets of subunits and chemical reactions are used, as well as the reaction of multiple subunits in one reaction step to form multicomponent compounds. The compatibility of each building block with each chemical reaction, and the product formed must be known and validated for the process to give useful libraries of compounds. While a distinction has been made in the past by some investigators between oligomer compounds and other multicomponent compounds, there is no difference between a linear or a branching synthetic scheme. Thus, just as in polymer chemistry where there may or may not be cross-linking and branching, so in oligomer chemistry where the subunits or monomers making up the final compound are different rather than the same (thus distinguishing the compounds from polymers of the same subunits) the oligomeric compounds may be linear or branched.

There are several basic strategies to the design, synthesis, and screening of combinatorial libraries. To simplify this review the different strategies are explained below, and the terminology that is used throughout this review is established. Multiple synthesis covers any method'in which discrete compounds are synthesized simultaneously to create a library of isolated compounds whose identity is known from the synthesis scheme. Iterative synthesis/ screening involves the synthesis of compounds in such a manner that a mixture results that is not directly resolvable to determine the identity of discrete active compounds, but that instead is resolvable to determine the identity of a specific residue(s) in any mixture that shows activity when assayed. A new set of mixtures is then synthesized based on this information and assayed, and the identity of the next specific residue(s) determined. The iterative process is continued until the identity of a complete, active molecule is determined. This process has several characteristics. First, the pool size diminishes as the iterations proceed, ending at the last step with the synthesis and testing of individual compounds. Second, the relationship of the last residues to be defined depends on the selection of the residues in the initial iterative steps, so there is an algorithmic relationship between the first and last components of the compound to be identified. Methods of deconvolution fall within the iteration definition and will be considered as a form of iterative synthesis/ screening.

Positional scanning is a method of synthesizing several mixtures of compounds such that in each specific residues can be defined. The whole set of mixtures enables the investigator to identify every active residue of a virtual compound without actually having synthesized and tested that compound as a pure entity, only presumably as a component of each mixture. This process does not involve any iterative synthesis and assay steps. Furthermore, unlike the iterative process, there is no dependence of the identification of any one residue on the identity of any other residue within the active molecule. The actual synthesis of the derived active compound(s) must be performed separately, ideally as a multiple synthesis effort.

One-compound–one-bead postassay identification design (PAID) approaches (sequencable/encoded[5,9]) covers a spectrum of approaches that share a synthesis scheme that assures that each resin particle contains only one structural compound. This is efficiently and practically achieved by applying the split-and-mix method of synthesis introduced by Furka[11] and independently by others[8,9] to assure that every amino acid couples in the intended equimolar ratio. This approach has the unique requirement for investigators to maintain the relationship between assay result and the bead on which the compound was synthesized so that they can utilize information on the solid-phase particle to identify the active compound or components of the active compound. It is principally in the method of identification that variations of this approach have been introduced. In some instances the compound itself can be directly sequenced or otherwise identified by analytical methods. In other cases the synthesis is carried out in such a manner that a coding molecule is also

synthesized on the solid-phase particle, and it is the coding molecule that is used to identify the active compound. The furthest afield approach introduced to date is the use of freely mixable capsules containing both resin and a microchip which can be read by radio-frequency methods. The basis for split-and-mix libraries lies in the freely mixable character of the support, whether it is a single resin bead or a collection of resin beads in some capsule or a tube of gel. While the one-compound–one-bead postassay identification approach to library synthesis absolutely depends on use of the split-and-mix synthesis process, iterative approaches can also use the split-and-mix method of synthesis.

The concept of the iterative synthesis of mixtures, positional scanning, and efficient screening in solution of one-compound–one-bead libraries raises the issue of synthesizing and testing mixtures of compounds rather than pure, highly characterized compounds. Scientists are accustomed to testing known compounds individually so as to minimize the possibility that some unknown factor or interfering substance may influence the observations being made. In screening large numbers of compounds, and particularly in combinatorial chemistry for the synthesis of large numbers of compounds, there is frequently a great advantage to synthesizing and assaying mixtures or pools of compounds. The immediate question is what risk does this entail, and what pool sizes can be used. The absolute answer must be judged on a case-by-case, experimental basis. However, as seen from all the examples given below, there does not seem to be an adverse effect of employing mixtures of compounds. Direct examination showed that pooling of hundreds of compounds did not mask positive compounds.

Analogs of the RGD peptide sequence were synthesized, pooled, and screened for binding to the gpIIb/IIIa platelet receptor.[12] The library synthesized was of the form $YGRGY_1X_2X_3$, where Y_1 consisted of the randomization of S, D, R, H, E, while 19 L-amino acids were randomized in the X_2 position (omitting isoleucine) and 20 L-amino acids were randomized in the X_3 position, creating a library of 1900 species. The library compounds were cleaved from the resin, pooled according to the residue incorporated at Y_1, and assayed. The $YGRGD\ X_2X_3$ pool inhibited the binding of soluble gpIIb/IIIa receptor to fibrinogen-coated plates greater than any of the other pools, exhibiting similar activity assayed as a pool of 360 compounds as did control peptide, RGDS. The identification of $YGRGD\ X_2X_3$ was predictable from previously described SAR, and the authors did not define X_2X_3; the focus of this study was to examine the effect of mixtures on assay results. In this study, assaying compounds as mixtures (of 360 closely related compounds) did not negatively influence the binding data and the ability to detect an active compound. The examples cited below all suggest that screening mixtures of compounds is a powerful approach for generating important information about the structure of active compounds, and that to avoid the use of pooling, when otherwise appropriate, is to unnecessarily lessen the amount of information that can be obtained for a particular system.

COMBINATORIAL SYNTHESIS OF PEPTIDES

Choice of Solid Support

Beaded polymer used for library synthesis has to fulfill certain criteria depending on the synthetic and screening strategy. For libraries based on the one-bead–one-compound approach, the size and substitution homogeneity is very important: To be able to evaluate the biological signal created by peptide released from a single bead one must be sure that the amount of compound released from each bead is uniform. Also important is the resin resistance to the formation of clusters (resin stickiness): Clusters will prevent the statistical free redistribution of resin beads and thus substantially lower the number of structures created. The ability of the resin to swell in both organic and aqueous media is especially important when target binding to the beads is used as the criteria for positive bead identification.

Polyacrylamide beads fulfill most of the above-mentioned criteria and was actually the resin used for evaluating technology feasibility.[9] However, polyacrylamide has been replaced by TentaGel, polyoxyethylene-grafted polystyrene,[13] which has become the resin of choice for both peptide and organic solid-phase library synthesis. The PEG-PS resin[14] of Millipore (Perseptive Biosystems today has similar composition and properties, but differs in the placement of the chemically reactive group (amino group) in relation to the polystyrene matrix. TentaGel has the functionalizable group at the end of the polyoxyethylene chains, far from the hydrophobic polystyrene chain, a feature that is especially important for compound display on the bead surface for the bead-binding assay. PEG-PS has the functional group next to the polystyrene chain such that the polyoxyethylene chain does not serve as a linker connecting the synthetic compound with the polymer, but rather as modifier of polymer properties.

Alternative carriers to classical resin beads have been tested. Polyacrylamide-grafted polypropylene pins were used for the synthesis of the first library.[15] This type of support was shown to be very useful in multiple-peptide synthesis.[16] It was adapted to larger scale by the application of crowns attached to the pins,[17] and is also useful for nonpeptide synthesis.[18]

Paper is a good support for multiple (SPOT) synthesis of thousands of peptides[19] or for the synthesis of libraries.[19,20] Synthesis on paper is performed by spotting the solution of protected amino acids onto the functionalized cellulose paper in the presence of activating reagent. In this case, the reaction vessel is the carrier itself, liquid manipulation during the synthesis (usual shaking in the case of solid-phase synthesis) is eliminated, and the reaction is driven by the diffusion of liquid in the carrier. This principle of internal volume synthesis was tested using polymeric carriers on a multiple synthesizer utilizing centrifugation for liquid elimination[21] and was found comparable, if not better, than the classical arrangement of solid-phase peptide synthesis.[22] Cotton, the purest form of cellulose, was found to be a convenient solid-phase support, especially for multiple synthesis[23] or library generation.[24]

One specific feature of membrane, sheet, or thread-like carriers is their divisibility. This feature can be used for the synthesis of libraries with a nonstatistical or forced distribution of library members.[25] The synthesis of this library starts with n pieces of the carrier, which are coupled with n different building blocks (1st randomization). Each of the n pieces is then divided into m parts and these smaller parts are distributed into m reaction vessels in which m reactions are performed (second randomization). The process can be repeated as many times as permitted by the handleability of polymer particles. The result of this process is a library of $(n)(m) \cdots = X$ compounds on X polymeric particles, where no compound is missing and none is represented more than once.

Synthetic Issues

Coupling in peptide library synthesis is performed in the usual way—in a battery of bubblers, closed plastic vials, polypropylene syringes,[26] or in tea bags[27]—and standard coupling reagents are used.[28-30] Detailed descriptions of library synthesis procedures have been published.[31,32] Difficult couplings can be forced to completion by the application of more reactive coupling reagents, such as TFFA.[33] However, due to the ease of handling and simplicity in deprotection performed in multiple vessels in parallel, the Fmoc protecting group is favored over the Boc group.

The techniques used to monitor coupling reactions required adaptation. In split-and-mix synthesis,[8,9,11] beads within each reaction mixture contain different peptide sequences, and therefore coupling kinetics for each bead may be different. Since we are interested in knowing not whether the coupling is complete on average, but rather whether coupling is complete on each bead in the sample, it is advisable to follow the reaction at the level of individual beads. This can be achieved by nondestructive methods such as bromphenol blue monitoring.[34] An example of monitoring individual beads is shown in Figure 3.1. Coupling was complete for the majority of beads (clear beads), but coupling on several beads is not complete (colored beads). The classical ninhydrin test[35] did not reveal any problems at this coupling stage.

Equimolarity of incorporation of amino acids into the growing peptides of a library can be achieved in several ways. The most reliable is split synthesis,[8,9,11] which was designed for the purpose of generating equimolar peptide mixtures.[11] Equimolarity, however, can be achieved only in the cases when the number of particles used for the synthesis is substantially larger than the number of synthesized peptides (1,000,000 compounds cannot be synthesized on 100,000 beads). Split synthesis results in a collection of polymeric particles, each containing individual sequences, and was the basis for the development of the one-bead–one-compound library technique.[9] The split method, however, is inconvenient in the case of iterative libraries containing fixed positions inside of the peptide sequence. In this case the synthesis of a library with 20 amino acids in one fixed position and 20 amino acids in positions requiring equimolar
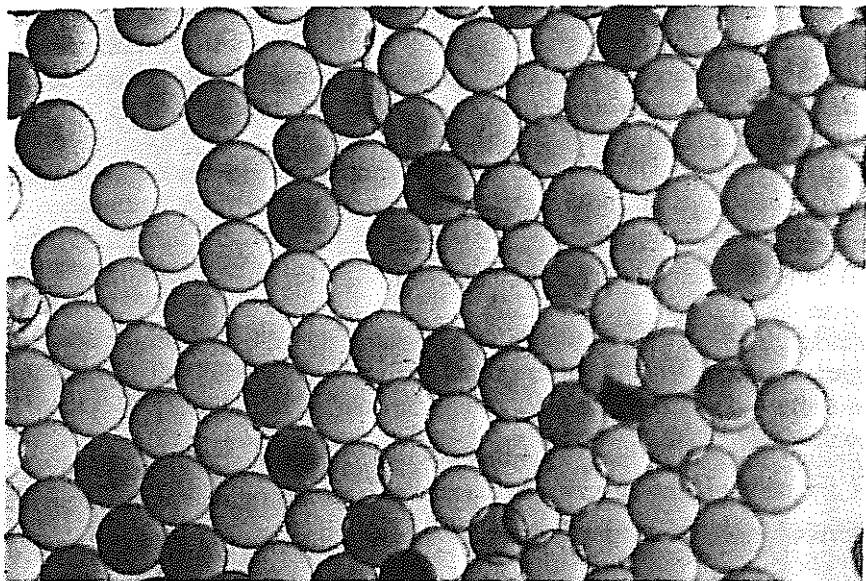
Figure 3.1. Monitoring coupling reactions on individual beads: Use of bromphenol blue to monitor the coupling reaction at the level of individual beads. At the time this photo was taken coupling was complete for the majority of resin beads, indicated by the lack of remaining color, but coupling can be seen to be incomplete of several (colored) beads.

incorporation of amino acids would require the use of 400 reaction vessels. An alternative is the use of a mixture of amino acids in which the ratio is adjusted according to the reactivity of the particular protected building block;[15,36-38] or double or triple coupling of subequimolar (0.8 molar) amount of equimolar mixtures of amino acids.[39-41] With either approach, the library would contain an equimolar representation of each peptide, though each bead would have a collection of peptides synthesized onto it (again with equimolar representation).

Synthesis of libraries can be performed manually or by automation.[42-44] The distribution of the resin is achieved by volume distribution of either a homogeneous (nonsedimenting) suspension of beads in isopycnic solution[42] (Zuckermann et al., 1992b) or a stirred suspension[43] (Saneii et al., 1993). The third design achieves distribution through sedimentation of the suspension in a symmetrical distribution vessel (Figure 3.2).[44]

Libraries immobilized on glass slides have been synthesized by the application of photolithographical techniques.[45,46] The identity of positively reacting molecules is known from the position on the glass surface (see Chapter 4). The solid-phase synthetic principle is also used in the case of liquid-phase combinatorial synthesis.[47] Synthesis is performed on soluble polymer (functionalized polyoxyethylene), which is then precipitated for removal of excess reagents. Biological screening can be performed with compounds directly attached to the soluble polymer.
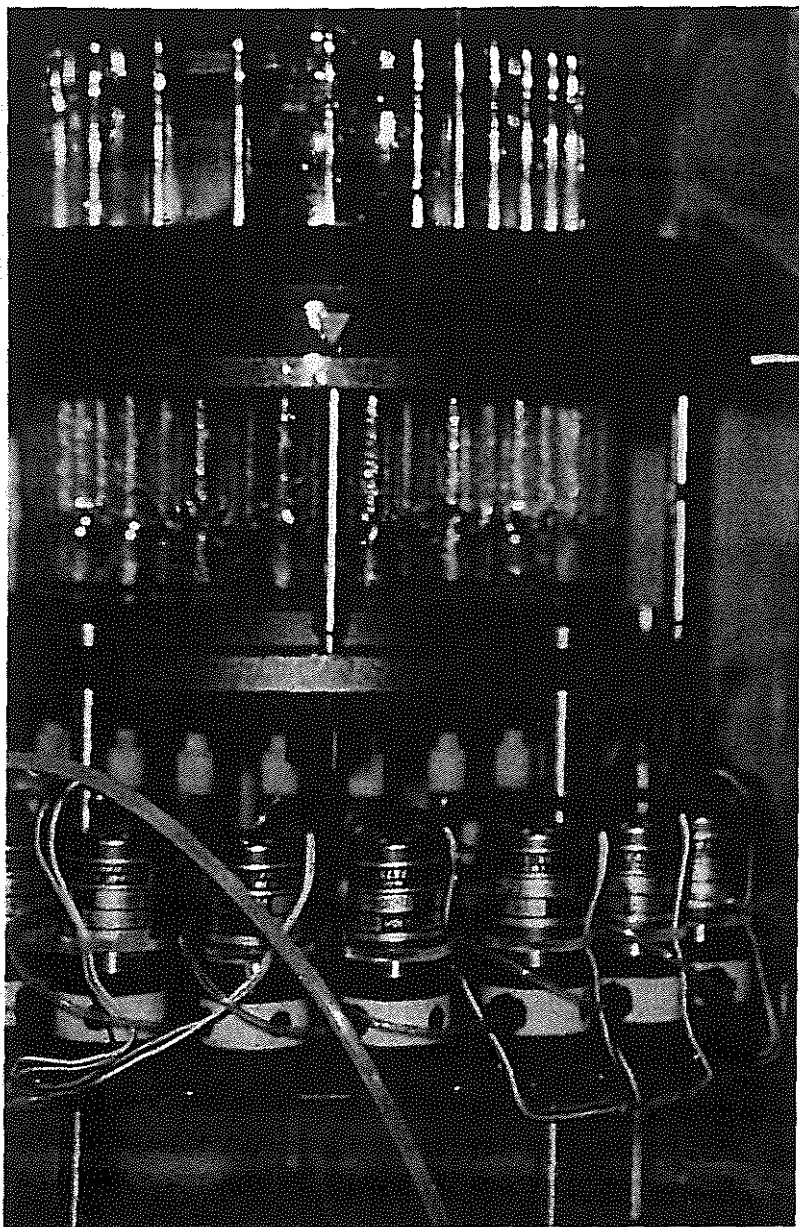
Figure 3.2. Redistribution of resin: The redistribution of resin into multiple vessels through the uniform settling of resin suspended in liquid media. The resin is thoroughly mixed and then allowed to settle. The design of the settling apparatus direct equal portions of the resin into the reaction vessels for the next step of synthesis.

Structure Determination by Nonencoded Means

Structure determination of the active component of a mixture is not an issue in the techniques utilizing iterative synthesis or positional scanning, the technique in which a certain feature of the library molecules is kept constant and the rest of the molecule structure is randomized. Determining the activity of variously defined mixtures defines the synthesis of the next generation library (or individual compound array in the case of positional scanning),[48,49] and the active structure is defined by a synthetic algorithm. Part of the material used in each step of the synthesis of an iterative library can be saved and used for the synthesis of subsequent library generations. This principle, called recursive deconvolution,[10] is attractive for laboratories evaluating one type of combinatorial library. It is not practical when hundreds of different libraries containing up to 100 building blocks in each randomized position are used, since one library composed of 4 consecutively connected building blocks, each position being randomized by 100 different blocks, would require storage of 1,010,100 resin aliquots.

A technique allowing the rapid determination of the active compound from a mixture is based on an orthogonal mixture principle. In this case two (or more) sets of mixtures are generated in such a way that individual compounds are present in two (or more) different mixtures in identical concentration. The biological activity of two orthogonal mixtures defines the active compound common for both mixtures.[50] The disadvantage of this method is the complexity of the synthesis of the orthogonal libraries (Figure 3.3), but robotic synthesis of the library addresses this problem. Conceptually similar was a library design utilizing enriched and depleted mixtures of amino acids in the construction of library mixtures.[51]

Library Formats

Challenging syntheses are the norm in combinatorial peptide chemistry, not the exception. For example, some biological targets require a free carboxy terminus (or both N- and C-termini free) to be displayed from amino to carboxy terminus. This is technically problematic due to significant problems with racemization in every step. This problem was solved by variously synthesizing a cyclic peptide containing a cleavable linker in its cyclic structure such that cleavage of the intramolecular linker exposed the free terminus of the molecule (Figure 3.4).[52-54]

Screening one-bead–one-compound libraries for activity in solution requires multiple release of an equimolar quantity of compound from individual beads (vide infra).[32,35] This can be achieved by attaching the compound to the bead via a linker, allowing release in two independent steps,[56,57] or by attaching the compound onto a mixture of linkers cleavable under different conditions. The first approach has the advantage of not decreasing the amount of released compounds, since the amount of compound on each bead is multiplied by the
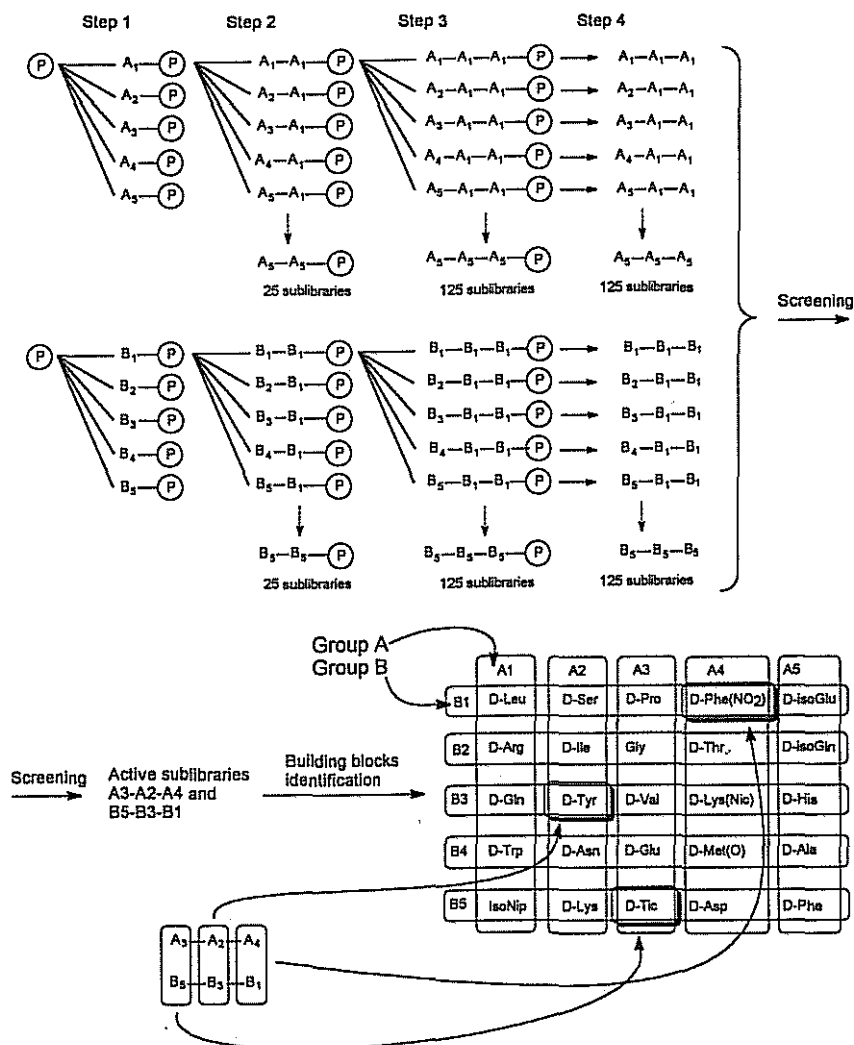
Figure 3.3. Orthogonol Deconvolution of Mixtures: Synthesis of two sets of libraries using orthogonol overlapping building block subsets permits the direct identification of an active compound from a the assay of both sets without resynthesis. Only one compound is common between the active sets. In this example mixtures contained 125 compounds each. Each set contained 125 sublibraries (as depicted for set A and Set B) which were assayed (total of 250 assays) in order to screen and identify the active compound from among the total 15,625 compounds.

Figure 3.4. Synthetic method for the solid phase synthesis and display of a free carboxy terminus of peptides remaining bound to solid phase particle during assay. The approach depends upon the synthesis of a peptide on a branched l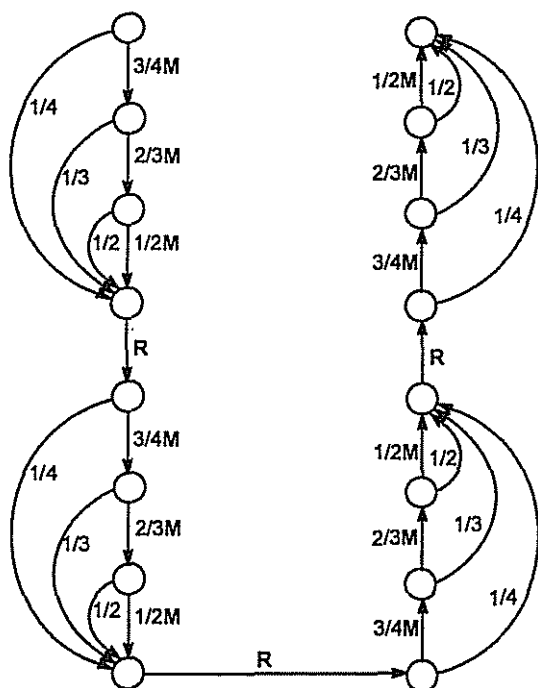inker containing a protected arm, followed by deprotection of the linker arm an cyclization back onto the linker, and then cleavage of the original linker arm on which the linear form of the peptide had been synthesized. The peptide remains bound to the resin (indicated by the "P") through the cyclizing peptide-linker bond.

branching of the linker to accommodate each step of release (the release mechanism is illustrated in Figure 3.5). An alternative method of tiered release is the kinetic release realized by either timed exposure of a relatively stable (benzhydrylamine) attachment to mild cleaving conditions (vapors of trifloro-acetic acid),[58] or exposure of photolytically cleavable attachment to light.[59] Various approaches to stepwise release of the compounds from polymeric carrier have been reviewed.[60]

The principle of a one-bead–one-compound library can be combined with the multiple-defined positional scanning concept[61] to address the problem of generation and screening incomplete libraries. A library of hexapeptides composed from 20 amino acids would contain 64,000 compounds. However, it is probable that not all six amino acids in any one peptide with significant affinity for the biological target are essential for binding. Some of the amino acids in the sequence can be easily replaced without significant loss of activity, whereas several critical amino acids cannot usually be replaced by a substitute. The task is therefore to identify the critical amino acids, or the motif required for binding. If we define the motif as an arrangement of three amino acids, not necessarily contiguous, we can create 20 different positional arrangements

Figure 3.5. Chemistry and method of release of identical copies of a compound from a doubly cleavable branched linker. Branching of the linker enables two equivalents of peptide and one equivalent of coding molecule to be synthesized on the resin (indicated by the "TG") for every equivalent of resin loading capacity. The linker used permits one copy of peptide to be released by acid, removed and tested, and then the second copy of peptide can be released at neutral pH, removed and tested, and finally the identity of peptide can be determined by the coding molecule left on the resin bead. The released peptides have the identical chemical memory (hydroxyl rather than a c-terminal carboxyl group) of the

51

Figure 3.6. Method of synthesis of a "library of libraries": Each solid phase support contain a mixture of compounds representing a library of compounds ranging in length from 3 to 15 residues to evaluate all 3-residue pharmacophores. While it is not necessary for the 3-residue pharmacophore to be composed of contiguous amino acid residues, in peptides greater than six residues in length pharmacophores containing residues separated by more than three residue uni positions (e.g. the pharmacophore of amino acid 1, 5, and 9 of a nonapeptide) will not be represented.

of those three amino acids in the framework of the hexapeptide. In the case of 20 amino acids, each of these positional arrangements contain 8000 individual compositional motifs. Therefore, the complete library of tripeptide motifs in a hexapeptide framework, a "library of libraries," is composed of 160,000 species of motifs. This is a substantial reduction in complexity compared to the 64,000,000 individual compounds in a hexapeptide library.

The synthesis of a library of libraries requires splitting the resin into several aliquots—in the case of the above-mentioned libraries of hexapeptides, into up to six aliquots. In three aliquots a mixture of amino acids is coupled and in three aliquots the position is randomized between the 20 different amino acids during parallel coupling as in the standard one-bead–one-compound synthetic method. In total, 252 couplings are performed in the synthesis of this library (for the scheme see reference 61).

The scheme of synthesis of an alternative library with variable length is depicted in Figure 3.6. At the beginning of synthesis, and after each randomiza-

tion step, one-quarter of the resin is separated and the mixture of amino acids is coupled to the remaining part. After this coupling, one-third of the resin is separated and the remainder undergoes coupling with the mixture of amino acids. The next coupling is performed with half of the resin from the previous coupling. All portions of the resin are then combined and a randomization is performed. Synthesis of a library of libraries with a three-amino acid motif, by this method, consists of three randomization steps and four stages of multiple couplings of amino acid mixtures. As a result, each solid-phase particle of the library is subjected to three mandatory randomization steps and as many as 12 acylations with the mixture of amino acids. This library, containing peptides of lengths from 3 to 15 residues, consists of 256 positional motif sublibraries. Among sublibraries of up to hexapeptides, all positional motifs are presented. However, because this synthetic scheme does not allow more then three successive acylations with the amino acid mixture, motifs in which pharmacophore positions are separated by more than three adjacent structural unit positions are not represented.

A significant increase in the diversity of peptide libraries can be achieved by using not only alpha but also other amino acids in the construction of the peptidic chain.[60] A peptide backbone can serve as the scaffold onto which a variety of building blocks can be attached via coupling to trifunctional amino acids (aminoglycine, diaminopropionic acid, diaminobutyric acid, ornithine, lysine, iminodiacetic acid, aspartic acid, glutamic acid, serine, hydroxyproline, cysteine, etc.).[54,62,63] Iminodiacetic acid is a convenient structural unit allowing construction of peptide-like libraries.[64] The scheme for its application to library construction is given in Figure 3.7.

Attachment of carboxylic acids onto the free amino groups was the basis for the construction of an alpha, beta, gamma library in which both the backbone and the side-chain arrangement was randomized.[65] On the border of peptidic and nonpeptidic libraries are the peptoids, or NSGs (N-substituted glycine peptides).[66,67] Libraries of peptoids were constructed either by coupling individual N-substituted glycines or by a submonomer approach, in which bromoacetic acid was coupled to the amino acid group of the growing peptoid chain, and primary amines were used to displace bromine in nucleophilic substitution, forming thus another N-substituted glycine in the chain[68] (see Chapter 6).

The linear arrangement of building blocks in peptide-like fashion was the basis for several library designs.[2,6,69,70] Because these designs produce nonpeptidic structures, they are not discussed here in detail.

Cyclization is a generally accepted method for decreasing the conformational flexibility of peptides. This restriction is expected to provide more potent ligands, as in the case of disulfide cyclic libraries containing ligand for the IIbIIIa receptor.[55] Cyclization can change the preference of the biological receptor for ligands, as shown in the study of streptavidin binders from cyclic libraries of various sizes.[71] A cyclic peptide composed of lysines and glutamic acid was used as a template onto which various carboxylic acids were
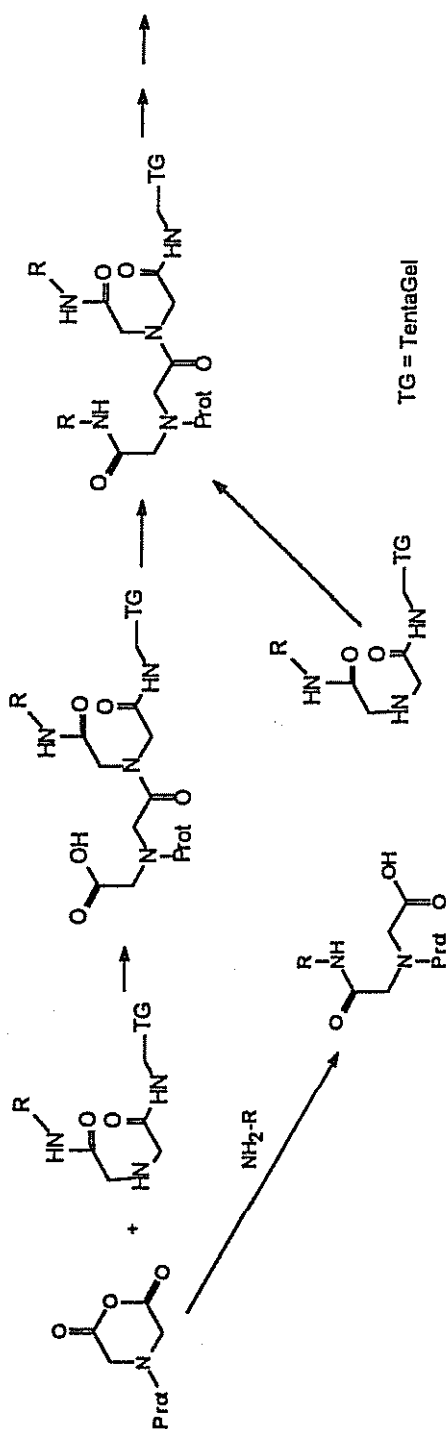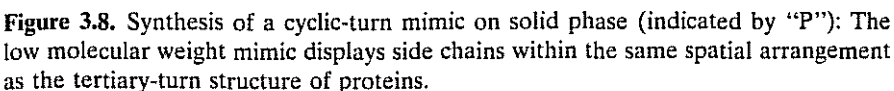
Figure 3.7. Introduction of side chain diversity: Use of a trifunctional amino acid, such as the depicted iminodiacetic acid, as a scaffold on which to synthesize diverse side chains while extending the peptide backbone the desired number of residues. Solid phase is represented by the "TG".

TG = TentaGel

Figure 3.8. Synthesis of a cyclic-turn mimic on solid phase (indicated by "P"): The low molecular weight mimic displays side chains within the same spatial arrangement as the tertiary-turn structure of proteins.

attached.[72] Cyclic peptide libraries were studied by Spatola et al., and an optimal strategy for their synthesis was devised.[73] Ellman et al., synthesized a β-turn mimetic library with the same goal (Figure 3.8).[74]

Conformational restriction has also been achieved by constructing libraries with a bias toward α-helical conformation. One such library was synthesized by randomizing four positions in the sequence of the amphipathic helix (YKLLKKLLKKLKKLLKKL) on either the hydrophobic or hydrophilic sides of the molecule, from which peptides with increased antimicrobial activity were identified.[75]

The combination of peptide structure with nonpeptidic elements has been used for library design. Potent and specific zinc endopeptidase inhibitors were identified in a library of peptides with the amino terminus modified by the Z-Phe(PO$_2$CH$_2$) peptidomimetic group.[76] A similar approach was used in a Pfizer study of the endothelin antagonist developed by Fugisawa.[77] N-terminal substitution was kept intact and all amino acids were randomized by an array of natural and unnatural α- and non-α-amino acids. Promising leads were generated from a library constructed using acylation of the peptide chain by an array of three building blocks containing amino and carboxyl functionality and capped by a set of carboxylic acids.[78] The peptide chain in this case was used as the biasing element targeting the binding pocket of the Src SH3 domain. Vinylogous sulfonyl peptide synthesis was developed,[79] and libraries were used for the studies of synthetic receptors.[80] Boc protected vinylogous sulfonyl chlorides coupling in dichloromethane catalyzed by dimethylamino-pyridine monitored by bromophenol blue (for proper base[DBU] excess during the reaction) was used in the steps of library synthesis. Monitoring of the
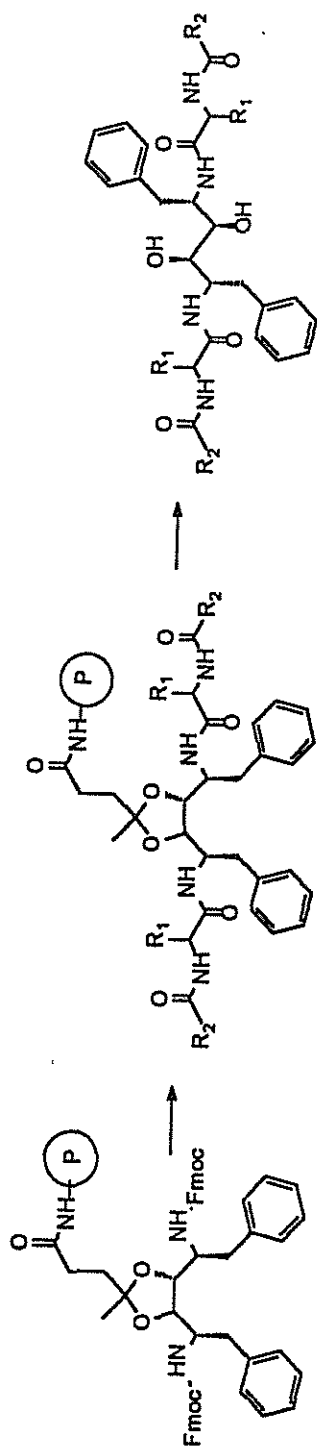
Figure 3.9. Incorporation of a target-specific (e.g. HIV protease) pharmacophore (diol motif) as the basic scaffold of a peptide library synthesized on solid phase (indicated by "P") followed by cleavage and testing of the free peptides.

reaction was found to be critical. Using an excess of the base destroyed the sulfonyl chloride, and even an excess of sulfonyl chloride did not result in complete amino group modification. Libraries of synthetic receptors were generated by the combinatorial synthesis of peptides on scaffold molecules such as a macrocyclic tetramine[81] or steroid molecule.[82]

Palladium-mediated macrocyclization was used in the synthesis of cyclic libraries. The carboxy terminal lysine side chain was acylated by acrylic acid, and the amino terminal group of the linear peptide was acylated by iodobenzoic acid. Pd(O)-mediated cyclization provided clean product in high yield.[83] The synthetic scheme of C2 symmetric inhibitors of HIV protease (Figure 3.9) serves as an example of solid-phase synthesis utilizing a specific feature of the target molecule (in this case the diol structure) to simultaneously protect the functional group critical to the function of the constructed molecules and at the same time use this protection as the attachment to the solid support.[84]

LEAD GENERATION AND OPTIMIZATION

Drug discovery is composed of two major elements: lead generation and lead optimization. The research reviewed in this section demonstrates that combinatorial peptide chemistry is a powerful tool for generating leads using generic libraries designed on general rather than target-specific principles. Combinatorial peptide chemistry is also a powerful tool for the optimization of leads through the synthesis and screening of optimization libraries designed based on a known lead or SAR. It is particularly powerful when combined with traditional medicinal chemistry and molecular modeling, though published examples of this are just now beginning to appear. Table 3.1 lists the targets for which leads have been identified and/or optimized from synthetic combinatorial peptide libraries. This list covers a variety of classes of targets. In some cases, investigators have pursued these targets as a means of validating a new method or the design of a combinatorial library. In other cases, information has been sought about the target and its chemical selectivity. In still other studies, the objective has been to identify or optimize a lead that can be used to modify the activity mediated by a specific target. Details of how each class of target has been pursued are discussed in the following sections.

Identification of Antibody Binding Domains and Putative Peptide Antigens

The identification of peptides that bind specifically to antibodies can be used to define the immunizing epitope within the native antigen. In addition, compounds can be identified that bind and elicit the same function as the native antigen, but that are structurally different, constituting a mimeotope—a mimic of the native antigen epitope. In either case there is the potential for identifying useful vaccines from synthetic combinatorial peptide libraries. There is also the potential to identify peptides that bind a portion of the eCDR region but

TABLE 3.1.

Target Molecule	Scientific Accomplishment	Method	References[a]
Multiple Synthesis			
Anti-VPI mAb	ID epitope, new method	Pin syn/solution assay, biased library[b]	7
Anti-Hemaglutin mAb	ID epitope, new method	Tea bag syn/solution assay, biased library	27
Anti-PLP mAb	ID epitope, new method	Spot syn/screen on polypropylene, biased library	85
Anti-CMV human sera	Map epitope, new method	Spot syn/screen on cellulose, biased library	86
Anti-opiate mAb	ID epitope, new method	Photochem syn & screen on glass, biased library	87
Anti-gp120 mAb	ID epitope	Microplate syn., biased library	88
Anti-gp160 mAb	ID epitope/vaccine	Pin syn/solution assay, biased library	89
Anti-TGFα mAb	Epitope ID	Spot syn/assay	90
Streptavidin	Novel binder	Spot syn/assay, surface presentation	91
DNA Binding	Novel binder	Spot syn/assay, double coupling	39
Vancomycin	Novel binder	Affinity capillary electrophoresis selection	92
c-AMP dep./c-Abl kinases	Kemtide analog	Autophosphorylation of library cpds, LCMS detection, biased library	93
Acetylcholinesterase	Inhibitor SAR	Syn & spot on TLC plate, novel detection, biased library	94
Chymase/ chymotrypsin	Inhibitor SAR, optimize specificity/activity	Pins, biased library	95
NK1 receptor	Antagonist SAR	Pin syn/solution assay, biased library	96,97
NK3 receptor	Novel antagonist, SAR, optimize activity	Microplate syn, novel information rich library	98
Split and Mix, One-Compound–One-Bead, Postassay Identification Format Libraries			
Anti-β-endorphin mAb	ID epitope	New method/on bead	9
Anti-β-endorphin mAb	ID epitope	Double release	55

TABLE 3.1. (*Continued*)

Target Molecule	Scientific Accomplishment	Method	References[a]
Split and Mix, One-Compound–One-Bead, Postassay Identification Format Libraries			
Anti-β-endorphin mAb	ID epitope	On-bead/radiolabel	99
Anti-β-endorphin mAb	ID epitope	On-bead/coded	100
Anti-cMYC mAb	ID etipope	On-bead/coded	101
Anti-gp120 mAb	ID epitope	On-bead/MS sequence	102, 103
Anti-measles virus mAb	ID epitope/vaccine	On-bead	104
B-cell surface idiotype	ID epitope	On-bead	105, 106
MHC Class I (HLA-A2 * B7)	ID binding motif	On-bead	107
Streptavidin/avidin	Novel/specificity	On-bead	9, 108, 109
Neural neurophysin	Oxytocin binding, methods validation	On-bead, magnetic bead selection, biased library	110
Src, Pl-3-kinase SH3	Novel/putative substrates from database search & ligand/SH3 NMR structure determined	On-bead	111 112
c-AMP dep. kinase	Specificity of substrates	On-bead phosphorylation/biased library	113
Src kinase	Novel/putative substrates	On-bead phosphorylation	114, 115
Hepatitis A 3C endoproteinase	Specificity of substrates	On-bead/novel sequencing method, biased library	116
Stromeylisin/collagenase	Novel/specificity of substrates	On-bead/novel fluorescence detection method	117 118, 119
Endopeptidase subtilisin Carlsberg	Novel/substrate pH specificity	On-bead/novel direct fluorescence detection method, biased library	120
Edopeptidase subtilisin Carlsberg	Novel inhibitors	On-bead/novel indirect fluorescence detection method	121
Protein tyrosine phosphatase	Inhibitors, optimize specificity/activity	Solution, rf chip coding, biased library	122
Trypanosomal gPG-kinase	Trypanosom-specific inhibitors	On-bead, fluorescent & Magnetic bead selection	123

TABLE 3.1. *(Continued)*

Target Molecule	Scientific Accomplishment	Method	References[a]
Split and Mix, One-Compound–One-Bead, Postassay Identification Format Libraries			
Leukocyte elastase	Inhibitors, optimize specificity, activity	On-bead, biased library	124
Thrombin	Inhibitor SAR, novel lead/optimize activity	On-bead, biased library	125
Factor Xa	Novel inhibitor lead, SAR/optimize activity	On-bead (lead from generic library, optimize with biased library)	126
Elastase, subtilisin, trypsin, etc.	Inhibitor SAR	Solution/novel affinity column-sequencing method, biased cyclic scaffold library	127
gpllb/llla receptor	Antagonist	On-bead & double release	55
C5a receptor	Agonist–antagonist switch, SAR	Double release, biased library	128
G-protein coupled receptors	Novel agonists & antagonists, SAR	Double release, novel melanocyte functional assay, biased library	129 130
Antimicrobial	Growth inhibitors	Single release, agarose plaque forming assay	131
Indigo carmine	Small molecule	On-bead, direct color detection	132
Iterative Synthesis and Screening Library Methods			
Anti-peptide mAb	ID epitope	New method/solution	8
Anti-peptide mAb	ID epitope	Solution/novel pooling	51
Anti-β-endorphin mAb	ID epitope	Solution	133
Anti-β-endorphin mAb	ID epitope	On-bead	10
Anti-β-endorphin mAb	ID epitope	Liquid syn/screening	47
Anti-TGFa mAb	ID epitope	Solution, spot, double couple	39
Anti-gp120 mAb	ID epitope	Robotics/affinity selection	134
Anti-LPS mAb	Discover novel epitope	NMR friendly scaffold library	135

TABLE 3.1. (*Continued*)

Target Molecule	Scientific Accomplishment	Method	References[a]
Iterative Synthesis and Screening Library Methods			
Anti-lysozyme mAb	Discontinuous epitope presentation	Rigid scaffold library	136
cAMP/cGMP kinases	Novel/substrate specificity	Solution	20
HIV-1 protease	Novel inhibitor/SAR, optimize activity	Solution, biased statin library	137
Trypsin	Novel inhibitor/SAR, optimize activity	Solution, lead generation/ optimization	24, 138
Zinc endopeptidases	Optimize inhibitor selectivity/activity	Solution, biased library	116
Endothelin receptor	Antagonist SAR, optimize activity	Solution, biased substance P library	139 140
Endothelin receptor	Antagonist SAR	Solution, biased antagonist FR-139317 library	77
μ-Opiate receptor	Novel agonists, SAR, optimize analgesic activity & selectivity	Solution	141–3
Antimicrobial	Growth inhibitors	Solution	144 145
Positional Scanning Format Library Methods			
Anti-peptide mAb	ID epitope	New method	146, 147
ACE and other enzymes	Substrate	Solution/novel seq. method	148
Chymotrypsin	Novel inhibitors	Tea bag, cyclic template library	72
IL-6 receptor	Novel antagonists, SAR	Solution, branched "polydentate" library	149
MHC class 1	Novel ligands binding MHC-1, &/or sensitize target cells for cytolysis by T-cell	Solution	150 151
RBC lysis	Inhibition of melittin mediated lysis	Solution	152

[a] Unless noted, libraries used were generic in design. Biased libraries include overlapping, deletion, truncation, or biased amino acid selection based on sequence of native antigen or ligand.

that do not elicit an antibody response, particularly when the native antigen binds through discontinuous epitopes or a conformational epitope.[106] In addition, due to inherent tight binding and a wide variety of commercially available reagents and assays, antibody binding has been used as a method for validating new approaches in the application of synthetic combinatorial chemistry.

Multiple-Synthesis Libraries. In the first study published describing the multiple synthesis of peptides on fixed solid supports in a 96-well microplate array, the target used to validate the approach and demonstrate utility was an antibody.[7] The pins, which had been dipped in reaction solutions to synthesize a library of known compounds, were dipped in assay medium containing a monoclonal antibody raised against the immunologically important coat protein (VPI) of foot-and-mouth disease virus to identify peptides that bound the antibody. The authors first synthesized 108 different overlapping hexapeptides required to cover the complete 213-amino acid sequence of VPI. From the results they identified GDLQVL as the epitope recognized by the antibody. They then synthesized a library of 120 different peptides in which all possible single-residue replacements were made using the 20 proteinogenic amino acids. This identified the two leucine residues as critical for antibody binding.

Another early study defining antibody–peptide recognition was reported by Houghten.[27] He described the multiple synthesis of peptides by segregating resin in individual permeable containers ("tea bags"), then simultaneously coupling amino acids to the segregated resin of a number of such tea bags suspended in a variety of reaction vessels. Each tea bag was tracked through the synthetic scheme so that at the end of the process there was a library of compounds synthesized simultaneously in a one-compound–one-bag format, ready for screening after cleavage from the resin. Using this method 238 different 13-mer peptides were synthesized, reflecting single amino acid variations of a sequence from hemagglutinin protein (residues 98–110, YPYDVPDYASLRS), which has been used to raise a monoclonal antibody. The individual peptides cleaved from the bags of resin were preadsorbed to microtiter plate wells to constitute an ELISA screening assay, and the antibody binding to these peptides was monitored. Houghten demonstrated that with this approach the residues critical for binding could be identified as, D_{101}, D_{104}, and A_{106}, validating the tea-bag multiple-synthesis concept and utility.

Antibody binding was used to validate a multiple-synthesis approach in which peptides are synthesized as spots on commercially available peptide synthesis polypropylene membranes (Millipore, Bedford, MA) and then assayed either while attached to the membranes by an ELISA-type assay or following release from the membranes.[85] A related synthesis method had been described using cellulose membranes.[41,86] The membranes were mounted on a 96-well synthesis device. The library consisted of all possible octamers (1–8, 2–9, . . . , 269–276) derived from bovine myelin proteolipid protein (PLP, 276 amino acid peptide), which had been used as the antigen for producing

the target molecule anti-PLP antibody. Synthesis was complete in 16 h. The polypropylene sheet was then incubated with the anti-PLP antibody, and bound antibody was detected with protein A conjugated to alkaline phosphatase via a calorimetric enzyme assay. The dominant epitopes mapped by this approach corresponded to peptides spanning residues 38–46 and residues 195–205. The authors thus not only validated a novel method for multiple synthesis, but also identified the binding epitopes for this anti-PLP antibody.

Frank et al. used serum from humans that was immunopositive for a 58-amino acid sequence from human cytomegalovirus 36/40 K protein (anti-CMV) as a test system.[86] They synthesized a library of 49 overlapping decapeptides on a sheet of cellulose paper, spotting the reaction mixtures as separate points of a grid to form an array of compounds. With the peptides still bound, the cellulose paper was incubated with the CMV positive serum. Positive spots were detected by identifying those to which antibody from the serum bound (using an anti-human β-galactosidase-conjugated second antibody). In this way the authors could map the binding epitope of individual human sera. This method was also used to characterize the epitope recognized by an anti-TGFa antibody, Tab2, by synthesizing a hexapeptide library spotted on cellulose of 1728 mixtures in the format XOOOOX, were X represents the positions where residues were randomized, defined by the spot location, and O represents positions where one of six amino acid mixtures was incorporated (A,P,G; D,E; H,K,R; N,Q,S,T; F,Y,W; I,L,V,M).[90]

A similar library was generated by phage display. The synthetic combinatorial library identified sequences that were native to TGFa (SHFNDC, K_d = 200 nM, and VSHFND, K_d = 80 nM) as well as a large number of other sequences with comparable affinity. The phage library identified the native HFND sequences (ASHFND, K_d = 60 nM; HFNDYL, K_d = 900 nM; and HFNDCI, K_d = 8000 nM), but, unlike the synthetic library, not the complete native sequence VSHFND or any novel sequences.

Stigler et al. characterized the binding epitope of the Fab fragment of an antibody (3D6) raised against the transmembrane protein gp41 of HIV-1 using a library of overlapping peptides, a random library, and molecular modeling.[153] The overlapping peptide library identified CSKLICTTAVPW, which was used to synthesize a secondary library on cellulose sheets to generate a library of all possible L and D conformers (494 in all). This identified critical residues and this information, together with low-resolution X-ray of the peptide bound to the Fab fragment, enabled the authors to construct a model of the bound conformation of the peptide and interactive residues. This study provides a strategy for constructing models of protein–peptide complexes using peptide library results to generate empirical experimental data to augment computer modeling and low resolution crystallographic approaches.

A useful method for determining antibody binding epitopes was demonstrated using photochemistry and light-directed synthesis on a sheet of glass to create an array of overlapping epitopes against which to screen an antibody.[87] The epitope of an antibody, C 32.39 anti-dynorphin-β-antibody, raised

against the opioid sequence YGGFLRRQFKVVT was identified by synthesizing a library of 1024 compounds consisting of the single amino acids as well as all di-, tri-, (etc.) peptides to cover all possible deletions (single, double, etc.) and truncations of, while not scrambling, the original sequence. This represents a very different strategy than simply the synthesis of overlapping sequences within the immunizing peptide. Fluorescently labeled antibody bound to the individual compounds displayed as an array on the glass surface was monitored, with the location on the glass giving the identity of each compound. Deletion sequences containing RQFKVV (four in all) exhibited high fluorescence. From the truncation results, loss of T, especially T plus V, from the C-terminus resulted in a loss of activity, while binding increased with truncation of the N-terminus to the second R, again identifying the RQFKVV(T) sequence as the epitope recognized by antibody. H-YGGFLRRQFKVVT-OH bound with an IC_{50} of 0.58 nM, while Ac-RQFKVVT-OH bound with an IC_{50} of 1.7 nM, and Ac-RQFKVVT-NH$_2$ bound with an IC_{50} of 0.49 nM.

Numerous other studies have demonstrated the utility of synthetic peptide libraries in the mapping of antibody binding domains and the production of potent antigens. Pinilla et al. recently demonstrated that calcium independent antigens could be identified by screening a library against a calcium-dependent monoclonal antibody; the anti-FLAG antibody M1, which binds DYKDDDDK-NH$_2$ in the presence of calcium and is used for protein purification.[49] The IC_{50} of DYKDDDDK-NH$_2$ determined with and without calcium was 14 and 274 nM, respectively, while a peptide identified from a hexapeptide library, DYKAKE-NH$_2$, exhibited an IC_{50} of 38 versus 3 nM in the presence and absence of calcium, respectively, and one identified from a decapeptide library, DYKEKELEDD-NH$_2$, exhibited an IC_{50} of 19 versus 2 nM with and without calcium, respectively.

Experiments with an anti-gp120 antibody identified an equipotent sequence containing nonnatural residues from a highly biased decamer library based on the native sequence Ac-RAFHTTGRII-NH$_2$.[88] The library, Ac-RA(X)HT-TG(X)I(X)-NH$_2$, produced an equipotent submillimolar affinity compound, Ac-Ra(L-napthylalanine)HTTG(R)I(L-norvaline)-NH$_2$, and related napthyla-lanine containing analogs due to the pooling method used in screening.

The multipin scanning approach of combinatorial peptide synthesis[154] was used to identify the minimal and optimal length of an endogenous sequence from HIV-1 gp160, which serves as the recognition domain for binding to the MHC class I D^d complex of murine cytotoxic T lymphocytes (CTL) to elicit a gp160 specific CTL response (lysis of peptide-binding, D^d-expressing target cells).[89] The approach consisted of synthesizing an array of either N-terminal and C-terminal truncations of the 15-mer sequence from the native protein (residues 315–329). The minimal sequence required for recognition based on this experiment consisted of an 8-mer, PGRAFVTI, residues 320–327. In the second set of experiments the C-terminus was either extended or truncated, and for each C-terminus (331, 330, 329, 328), a set of six different N-terminally

truncated peptides (residues 318 to 323) was synthesized to create, for example, six 331 C-terminal peptides with an N-terminus of residue 318, 317, 316, 315, 314, or 313, respectively. Assaying this set of compounds determined that truncation beyond residue 320 resulted in loss of activity regardless of the C-terminal extension, confirming the minimum sequence, and identifying the most active peptides as the 9-mers PGRAFVTIG and GPGRAFVTI or the 10-mer RGPGRAFVTI (residues 318–327).

A corroborating study using a conventional approach also identified the 10-mer (residues 318–327) as more active than the native 15-mer.[155,156] When the 10-mer peptide was incorporated into VAC recombinants and used to immunize mice, CTL lines obtained from the immunized mice exhibited specific priming by synthetic peptide comparable to that induced by wild-type, full-length gp160, demonstrating that peptides identified by synthetic combinatorial chemistry efforts can be used effectively as the basis for vaccine production.

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. Antibody binding was used to validate the use of the split-and-mix approach to synthesizing and screening libraries in a one-compound–one-bead, postassay, directly sequenceable identification format.[9] A library of peptides was synthesized on beads using the split-and-mix strategy[11] to assure that each amino acid coupled regardless of its inherent reactivity and to assure that there would be only one species of compound on each resin particle (Figure 3.10). With the peptides still attached to the beads, the library was incubated with mouse monoclonal anti-β-endorphin. This is an antibody that recognizes a sequence in the native ligand (binding affinity, 17 nM). Antibody binding to beads was detected using a secondary antibody conjugated to an enzyme reporter (e.g., alkaline phosphatase), the positive beads were collected, and the peptide on each was sequenced to determine the identity of the individual compounds. The ligands identified from peptide libraries in which 19 of the proteinogenic amino acids were coupled at each position included YGGF(X) motifs, among them YGGFQ (15 nM affinity). The authors demonstrated that a positive bead could be treated to remove the target molecule and reassayed to confirm binding or to assess specificity. Furthermore, the bulk of the on-bead library remained after screening (minus any positive beads) and could be reused in additional screens.

A significant problem with the use of streptavidin to identify the bound biotinylated antibody is that streptavidin itself binds to many peptides, and hence can bind directly to beads (examples given below). To avoid this, two independent methods can be used to recognize the bound biotinylated antibody: streptavidin and an anti-mouse antibody. (Lam et al, 1995). Beads were first labeled with the streptavidin/alkaline phosphatase conjugate, colorized (turquoise) with the substrate BICP, and selected. The selected beads were incubated with anti-mouse antibody conjugated with alkaline phosphatase, and colorized with BCIP plus NBT, turning positive beads dark purple.

Split and Mix, One-Bead-One-Compound Synthesis
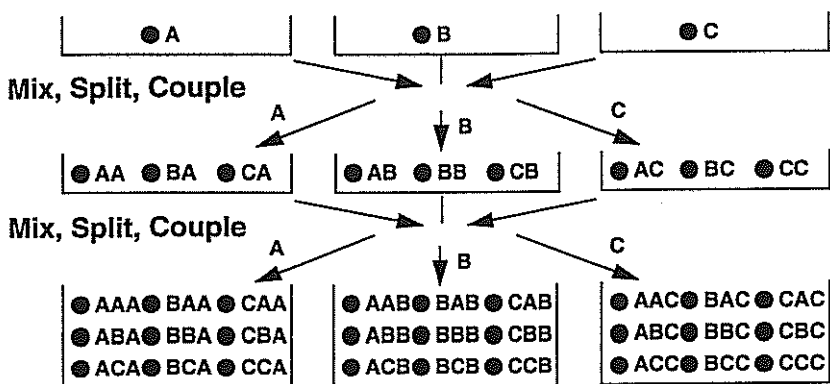
One-Bead-One-Compound Screening
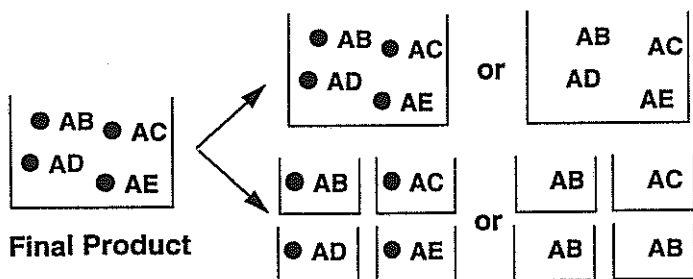
Figure 3.10. Split and mix synthesis and screening to produce one-bead-one-compound libraries for assay as mixtures of compounds or single compounds attached to the solid phase (indicated by solid circles) or as mixtures or single compounds released into solution from the solid phase. Release is carried out in a relational manner (to the solid phase bead of origin), such that once activity is detected of release compound it is possible to directly determine compound identity from chemical material remaining attached to the solid phase bead.

Those beads remaining turquoise (not binding the anti-mouse antibody) were sequenced, and it was shown that the peptides were streptavidin binders. Those that turned dark purple proved to be YGG motif peptides, true anti-β-endorphin binders.

In a follow-up study the approach was further modified to permit the controlled double release and assay of the library peptides in solution.[35] A pH-sensitive linker was employed that was stable in aqueous solution at acidic pH (4.5) but underwent a ring closure and cleavage to release compound (one-third of the total) from each bead at pH 7 in the form of

XXXXX-NH-(CH$_2$)$_3$-OH.[56] Released in randomly generated pools of 500 beads per well, the mixtures were assayed in an ELISA-type assay to identify active pools. The beads corresponding to these active pools were then recovered and redistributed, 1–2 per well, and subjected to basic pH (0.2% NaOH) to cleave another third of the compound attached through an ester linkage, again releasing molecules with the structure XXXXX-NH-(CH$_2$)$_3$-OH. After assay and identification of the positive wells, the corresponding beads were recovered and sequenced to determine identity (using the remaining third of compound which had been synthesized onto a protected linker). The identified compounds were synthesized and their activity was confirmed. This follow-up study validated the double-release assay through the identification of the spiked control YGGFL (30 nM) as well as YGGFG (200 nM), YGVFG (1000 nM), and YGAFG (700 nM) from the assay of 500,000 library beads. The double-release assay is a powerful technique, now widely used by many laboratories.

In a related study the on-bead assay was conducted by incubating beads with radiolabeled target molecule, immobilizing the beads in a thin layer of agarose, identifying positive beads by autoradiography, and manually selecting beads for sequencing.[99] The target used was again an anti-β-endorphin monoclonal antibody. From a pentapeptide library the authors identified YGSFE, YLWFQ, and YGAFE, clearly related to the known YGGFL sequence, and in the case of YGAFE, nearly identical to a peptide found using the double-release assay described above. Like Lam et al.,[9] the authors demonstrated that the recovered beads could be treated to dissociate the target and reassayed several times.

An on-bead method using a sequenceable coding molecule, rather than relying on sequencing of the active compound itself, was validated using the same antibody.[100] The assay was carried out using fluorescently labeled antibody and a fluorescence activated cell sorter to select the (positive) fluorescent beads. A heptapeptide library was synthesized on 10-mm macroporous beads, which permitted the target protein to permeate the bead and reach compound synthesized inside. The method was first validated using the known epitope RQFKVVT (K_d = 0.5 nM) synthesized onto beads. Two percent of the library beads stained above background and the bound fluorescent antibody could be completed by excess RQFKVVT, demonstrating specificity of binding. Peptides were identified from the library that, when synthesized with the threonine linker (T), ranged in binding affinity from a K_d of 0.3 to 1400 nM. The sequence of the most potent (K_d = 0.29 nM) was TFRQFKV(T), a close analog of the native epitope RQFKVV(T) (K_d = 0.51 nM).

An alternative coding method was validated using a monoclonal anti-c-MYC antibody as the target protein.[101] In this case the coding molecules consisted of a binary digital code arrangement of electrophores that could be separated and detected by electron capture gas chromatography (see Chapter 14). In this initial work the codes were attached to a fraction of the test compound, but the coding molecules can also be synthesized directly onto

the polystyrene bead. In digital coding the position and identity of each residue is represented by a three-"residue" bit code, where no residue is considered one bit of information, just as would be a specific coding residue. The library synthesized was a heptapeptide library of the format H_2N-XXXXXXEEDLGGGG-bead, containing D, E, I, K, L, Q, and S randomized at each position. It was known that the anti-c-MYC antibody 9E10 bound EQKLISEEDL, hence the selection of amino acids for the library. Using the approach described above (mixing library beads with the anti-c-MYC antibody and detecting those beads to which antibody bound using a secondary alkaline phosphatase conjugated antibody) the authors identified and sequenced the codes for 12 beads. Upon resynthesis and testing in solution the three most potent ($IC_{50} \sim 1$ mM) were EQKLIS(EEDL), LQKLIS(EEDL), and QQKLIS(EEDL), all closely related to the known epitope, validating this coding methodology.

Screening using the anti-gp120 antibody was used to validate a laser desorption mass spectroscopic method of sequencing to determine the identity of compounds on positive beads.[102,103] A synthetic hexapeptide library was synthesized in such a manner that at the end of each synthetic step a portion of the sequences were capped to prevent their complete synthesis in the remaining scheme. Thus, each bead contained the complete test peptide as well as capped truncations. The library beads were incubated with the anti-gp120 antibody, and positive beads were identified using a secondary antibody conjugated to alkaline phosphatase. Each positive bead was selected and its peptide contents cleaved and subjected to laser desorption mass spectroscopy. The spectrum contained the molecular ions for the complete compound plus all the possible capped analogs. The differences in molecular weight could be used to identify the composition of the parent peptide. This method can be applied to any molecular library, not just peptide libraries.

From the first study a family of peptides were identified, which provided a consensus binding epitope of (F/P)**GRAF**(Q/X)(F/X). This corresponded well with the peptide from the HIV gp120 protein used to produce the antibody originally, RIQRGP**GRAF**VTIGK ($IC_{50} = 11$ nM), compared to an $IC_{50} = 50$ nM for PGRAFQF. In the second study a pentapeptide library was synthesized, with and without acetylation, and screened, resulting in the identification of a XGRAF motif and an Ac-PG(R/X)AF(X/R)F motif. Ac-PGRAFQF exhibited an IC_{50} of 525 nm. The authors also screened against streptavidin (see below), identifying the HPQ motif with and without acetylation.

Another example of identifying a vaccine from a synthetic combinatorial library has been published.[104] Steward et al. synthesized an octamer library, screened on-bead with a monoclonal antibody raised against the F protein of measles virus (anti-MVF antibody F7-21). Six peptide mimetopes were identified, and when used to immunize mice, one (NIIRTKKQ), which cross-reacted with the antibody, inhibited measles virus plaque formation and acted as a vaccine to protect mice against fatal encephalitis induced by infection with measles.

Iterative Synthesis and Screening Library Methods. Hexapeptides were identified that bound a monoclonal antibody that recognized a 13-residue peptide (Ac-YPYDVPDYASLRS-NH$_2$) using an ELISA-type assay with the library peptides free in solution to identify antibody binding epitopes.[8] A total of 324 different pools were synthesized with the format Ac-O$_1$O$_2$XXXX-NH$_2$ in which the first two positions of each were defined for each pool. Screening identified one pool containing Ac-DVXXXX-NH$_2$. Subsequently, 20 pools were synthesized with the format Ac-DVO$_3$XXX-NH$_2$ and screened to define O$_3$; the process iterated until each position had been defined. In the process of defining the last position a set of individual peptides rather than mixtures was synthesized and assayed, having the same O$_{1-5}$ residues, but variable O$_6$ residues. Among them was the native sequence Ac-DVPDYA-NH$_2$ (IC$_{50}$ = 30 nM), which was also the most active. Corroborating studies using traditional approaches confirm this to be the antigenic determinant of the 13-residue peptide.[157,158] Using this approach the same group defined the binding epitope (STTS) of a monoclonal antibody raised against a surface antigen of hepatitis B, identifying the peptides Ac-STTSLI-NH$_2$ and Ac-STTSLM-NH$_2$ (IC$_{50}$ ~ 1.8 and 2.6 mM, respectively).[146] Similar results were found using a positional scanning library approach (see below).

The success of the iterative approach, together with a novel pooling strategy, was also demonstrated using both an antiserum to a tetrapeptide as well as a monoclonal antibody to a 28-amino acid peptide.[51] The antiserum used was raised against the peptide FMRF amide, but because the authors wished to exclude methionine, cysteine, tryptophane, threonine, and isoleucine from their library (reducing the number of amino acids used to 15) the ligand bound to microplates forming the basis of their antibody capture ELISA assay was FLRF. This strategy was somewhat unusual. The authors divided their amino acid mixtures into three groups, a(L,A,V,F,Y), b(G,S,P,D,E), and g(K,R,H,N,Q), with the individual amino acids present at different molar ratios designed to compensate for differences in reactivity. One strategy would have been to synthesize the library as 81 pools representing all possible positions within the tetramer and each of the three groups of amino acids. The authors instead synthesized a pool of compounds containing equal proportions of all 15 amino acids at each position (a 1:1:1 mixture of amino acid groups a:b:g), measured the activity of this pool (after the compounds were cleaved from the resin), and then compared this to at first four pools, which consisted of

AA1	AA2	AA3	AA4	
	Activity/Interpretation			
1:1:1 a:b:g	1:1:1 a:b:g	1:1:1 a:b:g	1:1:1 a:b:g	null control
2:1 b:g	1:1:1 a:b:g	1:1:1 a:b:g	1:1:1 a:b:g	<;a
1:1:1 a:b:g	2:1 b:g	1:1:1 a:b:g	1:1:1 a:b:g	<;a
1:1:1 a:b:g	1:1:1 a:b:g	2:1 b:g	1:1:1 a:b:g	=;g
1:1:1 a:b:g	1:1:1 a:b:g	1:1:1 a:b:g	2:1 b:g	<;a

Proportionate loss of activity was used to identify critical residues. The first round of screening results indicated a group residues were required at positions 1, 2, and 4 for activity, and g group residues were necessary at position 3. The next round was designed to identify triplets of amino acids from among these pools, using as the null control a, a, g, a, and four libraries:

AA1	AA2	AA3	AA4	
	Activity/Interpretation			
a	a	g	a	null control
1:1:3 LAV	a	g	a	<;F/Y
a	1:1:3 LAV	g	a	=;A/L
a	a	1:1:3 KRH	a	=;K/R
a	a	g	1:1:3 LAV	<;F/Y

The last round used the defined dipeptide mixtures as null control and the following four pools:

AA1	AA2	AA3	AA4	
	Activity/Interpretation			
F/Y	A/L	K/R	F/Y	null control
F	A/L	K/R	F/Y	=;F
F/Y	A	K/R	F/Y	<;L
F/Y	A/L	K	F/Y	<;R
F/Y	A/L	K/R	F	=;F

Thus with three rounds of iteration, synthesizing 12 pools of peptides representing 50,625 peptides, the authors had defined the peptide FLRF.

Having shown that the approach worked, the authors then applied the method to the case of the monoclonal antibody, raised against the 28-amino acid peptide Ac-RTPALG**PQAGID**TNEIAPLEPDAPPDAC amide, and for which the epitope had not been defined. The difference was that a hexapeptide format was pursued, so that there were 6 pools in each round of iteration (18 pools synthesized, representing 16,777,216 peptides), but otherwise the strategy was identical. The result was the definition of RQVGHD amide. The authors used this to identify the consensus region of the 28-mer to be PQAGID (bolded above), and confirmed this by synthesizing seven peptides and assessing their relative activity (shown in brackets) compared to the "native" peptide PQAGID [2.8]: RQVGHD [1.2], QVGHD [>750], RQVGH [128], EQVGHD [14], PQVGHD [0.08], RQAGHD [4], and RQVGID [17]. The loss of activity in the N- and C-terminal deletion analogs demonstrated specificity, and the ~35-fold improvement in activity of PQVGHD over the native peptide PQAGID confirmed specificity and demonstrates the power of the approach for optimization.
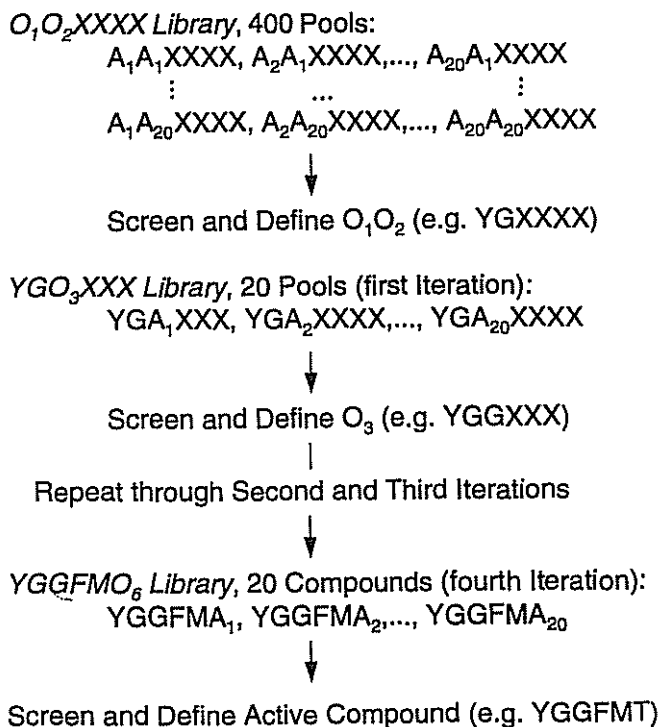
O_1O_2XXXX Library, 400 Pools:
A_1A_1XXXX, A_2A_1XXXX,..., $A_{20}A_1XXXX$
⋮ ... ⋮
$A_1A_{20}XXXX$, $A_2A_{20}XXXX$,..., $A_{20}A_{20}XXXX$
↓
Screen and Define O_1O_2 (e.g. YGXXXX)

YGO_3XXX Library, 20 Pools (first Iteration):
YGA_1XXX, YGA_2XXXX,..., $YGA_{20}XXXX$
↓
Screen and Define O_3 (e.g. YGGXXX)
|
Repeat through Second and Third Iterations
↓
$YGGFMO_6$ Library, 20 Compounds (fourth Iteration):
$YGGFMA_1$, $YGGFMA_2$,..., $YGGFMA_{20}$
↓
Screen and Define Active Compound (e.g. YGGFMT)

Figure 3.11. Iterative Deconvolution of Mixtures: A library composed of 6-mers requires five steps of synthesis and assay of the resulting mixtures, with compounds free in solution. The results of such an approach may be biased by the synthetic sequence, since the first two residues (O_1O_2) defined in the first set of mixtures have considerably more impact on binding than the residue (O_6) identified in the last set of mixtures.

In another study, peptides that bind anti-β-endorphin (3E7) were elucidated using iterative synthesis and screening of a series of hexamer libraries (Figure 3.11). Peptides were cleaved from resin as mixtures and assayed free in solution without separation in an ELISA-type format.[133] The antigen peptide (β-endorphin) was adsorbed onto the surface of microplate wells, the peptide mixture plus antibody was added, incubated, and washed, and then bound antibody was detected. In the first stage 400 libraries were synthesized, each containing specified amino acids (O_1O_2) in the first two positions, with the remainder (X) synthesized by splitting and mixing and cleaved as a complete mixture as regards the X positions (format O_1O_2XXXX). Once the first two positions (A_2A_2) were defined, 20 libraries were synthesized, with the third position (O_3) identified in each (format $A_1A_2O_3XXX$). This process continued until the complete peptide was defined, YGGFMT-NH$_2$ (IC$_{50}$ = 3.2 nM), a peptide identical to the first six residues of β-endorphin.

The most recent variation on the iterative approach is single-step deconvo-
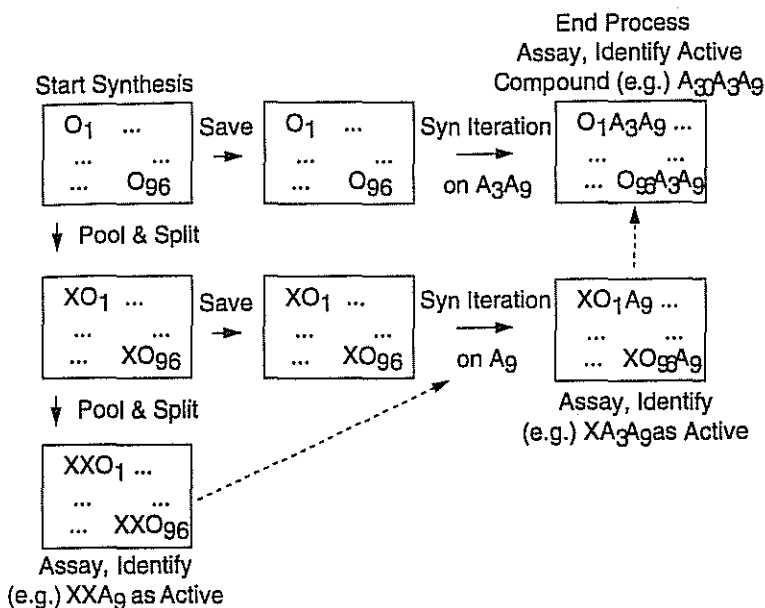
Figure 3.12. Recursive Deconvolution of Mixtures: Compounds composed of 3-mers require three assays and two additional steps of synthesis performed on starting material split and saved from each step of the original synthesis, with the advantage that results are not biased by synthetic sequence. In each case the last building block is added to mixtures in discrete wells of an array, such that the identity of each building block is known from the spatial location in the array of the active mixture. Working backwards (recursively) identifies the active compounds.

lution or recursive deconvolution where automation is utilized to simplify the iterative process (Figure 3.12). In recursive deconvolution the library is split after each synthetic step into samples that are saved for subsequent synthesis and samples that are pooled and split in a many-compounds–one-well format (rather than the one-compound–one-bead format) for the next round of synthesis.[10] Thus, for instance, at the end of the synthesis of a tripeptide library there would be multiple wells containing pools of compounds with the identity of the last amino acid unknown. Screening would identify the last amino acid, and then the investigators would take the previous saved plate of dipeptide mixtures, in which the second residue was known because of well location, synthesize onto these mixtures the identified third residue, and then screen to identify the second residue. Once identified, the investigators would go to the first round of saved single amino acid plates, synthesize the second and third residues onto each, and identify the active tripeptide. The authors synthesized a pentapeptide library, but with only four amino acids (G, L, F, Y) used at each position to create a library of 1024 members, including a peptide to bind the anti-β-endorphin antibody. This approach eliminates the need for

tags, dependence on resin bead size to provide sufficient assay material, and some of the synthesis steps required in the iterative approach. The authors screened the library compounds in the anti-β-endorphin antibody ELISA assay with compounds attached to the beads to assess whether compounds in the mixture prevented the antibody from binding to the [Leu$_5$]enkephalin–BSA-coated surface. The authors identified the known peptide NH$_2$-YGGFL as well as two other peptides, NH$_2$-YGGFF and NH$_2$-YGGLL, validating their recursive deconvolution method.

An important issue in all the approaches where a mixture of amino acids is reacted in the same reaction vessel is the fact that some will couple at a faster rate than others, producing an unequal representation within the pool. The rates of reaction are somewhat inherent in the amino acid itself, but are also dependent on the sequence to which the amino acid is being coupled. Many investigators have tried to manage this problem by changing the ratio of amino acids within the mixture inversely according to their general reactivity, using a larger number of equivalents of the slower reacting amino acids to create a "kinetically adjusted" acylating mixture. However, another method has also been described. It had long been accepted in peptide chemistry that for some difficult couplings a process of double coupling was necessary to completely react all available sites on the solid phase. Double coupling means that after the first reaction is complete, the same reaction is repeated again to force coupling to all available reactive sites.

This approach was modified by Kramer et al. to assure that in a mixture there would be essentially an equimolar representation of all amino acids, regardless of the rate at which they reacted with the solid phase.[39] They used a less than saturating mixture in the first coupling reaction, for example, 0.8 equivalents (relative to the number of reactive sites in the solid phase) of an equimolar mixture of amino acids. Thus in the first step the slowest reacting amino acid would be coupled to the solid support with close to the same efficiency as the fastest reacting amino acids. The result was an equimolar representation occupying 80% of the available coupling sites on the solid support at the end of the first reaction. The remaining reactive sites were coupled in the second reaction of the double-coupling protocol, so the fastest reacting amino acids coupled to a <20% excess over the slowest.

The authors validated this approach through the synthesis of libraries on solid-phase resin and on a cellulose sheet, using a variety of targets, among them the epitope definition of an anti-TGFa monoclonal antibody, Tab2. The iterative library design was XXO_1O_2XX, where in this case the double coupling of a mixture of 19 amino acids was used at the X positions, and O_1 and O_2 were defined, creating pools of peptides in solution following cleavage from the solid support. A standard competitive ELISA was performed using TGFa bound to microplates and identifying those wells to which horseradish peroxidase labeled anti-TGFa did not bind in the presence of the particular peptide pool. The known epitope was HFND. From the library (containing 47,045,881 compounds in 361 pools) the authors identified XXHFXX, XXFNXX, and

XXNDXX, clearly peptide pools related to, and containing, the known epitope. When these pools were synthesized as spots in a 19×19, O_1 by O_2 array on a sheet of cellulose paper, the same families and conservative substitutions were identified: FN, HF, HY, ND, YF, YN, YY. Using two additional 19×19-array cellulose-bound libraries the authors identified the amino acids permitted on either side of the O_1O_2 motif, synthesizing libraries of the structure X O_1HF O_2X, X O_1FN O_2X, and X O_1ND O_2X.

Positional Scanning Format Library Methods. A decapeptide positional scanning library approach was evaluated using the same anti-Ac-YPYDVPDYASLRS-NH$_2$ antibody.[147] Ten different positional libraries were synthesized, each containing 20 pools with a defined amino acid (O) at the specific position being investigated by that library (O_1XXXXXXXXX, XO_2XXXXXXXX, etc.). From this study 15 peptides were identified with comparable or better activity than the antigenic 13-mer (IC$_{50}$ = 6 nM), and all but one contained a consensus DYA sequence at positions 8, 9, and 10.

A derivative positional scanning method published by Wong et al. was also validated using antibody binding.[159,160] The authors synthesized three libraries of the format Ac-O_1O_2XXXX-NH$_2$, Ac-XXO_3O_4XX-NH$_2$, and Ac-XXXXO_5O_6-NH$_2$, using only 10 "physicochemical representative" proteinogenic amino acids so that each library consisted of 100 dipeptide defined pools of 10,000 compounds each. Screened against a monoclonal antibody against the pili of *Pseudomonas aeruginosa,* the authors identified the best dipeptides, EQ, FI, and PK, corresponding to the known epitope DEQFIPK. They synthesized EQFIPK, and determined an IC$_{50}$ of 1.3 mM, validating the approach. However, when the 8 peptides containing single-amino acid substitutions of the residues representing amino acid classes (D for E; N for Q; W, Y for F; V,M for I; H,R for K) were synthesized and tested, only four exhibited activity similar to that of the lead sequence (IC$_{50}$ = 3.4–0.9 mM). This meant that while the representative algorithm was effective in enriching the active hit rate upon secondary synthesis, if the wrong representative residues had been selected no lead in the case of H or R for K or a less informative lead in the case of N for Q, or Y for F would have been discovered. Therefore, the authors propose that this approach is best used where it is necessary to simplify the synthetic process, such as synthesis and screening libraries of longer peptides (heptamer, octamer, etc.), trading off the risk of missing leads with the ability to actually synthesize and screen a complete library.

Using this approach Pinilla et al, defined the binding epitope (STTS) of a monoclonal antibody raised against a surface antigen of hepatitis B, identifying the peptide consensus Ac-S(S/T)T(P/S)(A/M)(H/M)-NH$_2$, from which the three most active peptides synthesized were Ac-STTSMM-NH$_2$ (IC$_{50}$ = 0.39 mM), Ac-SSTSMM-NH$_2$ (IC$_{50}$ = 0.87 mM), and Ac-STTSAM-NH$_2$ (IC$_{50}$ = 2.8 mM).[146] These results were essentially the same as those obtained by the iterative synthesis library described above.

Liquid-Phase Combinatorial Libraries. As combinatorial chemistry is extended from peptides to nonpeptides, synthesis issues become the major limitation. A method termed liquid-phase synthesis was demonstrated in which the compounds are synthesized onto a chemical phase (polyethylene glycol, MeO-PEG 5000) that can be conveniently crystallized or precipitated by solvent changes. Liquid-phase synthesis permits filtration purification and affords all the advantages of solid-phase synthesis when it comes to the separation of product and reactants, yet permits the reactions to be carried out in solution.[47] This method was validated using anti-β-endorphin as the target and the recursive deconvolution strategy described above. From a pentapeptide library containing G, L, F, and Y, the authors identified YGGFL-MeO-PEG,   YGGFF-MeO-PEG,   YGGFY-MeO-PEG,   and YGGFG-MeO-PEG using the same ELISA assay as described above, though in this case the pools of library compounds were in solution, though not cleaved from the PEG.

Validation of Synthesis, Assay, and Compound Identification Methods. The first description of a fully automated, 96-well, robotic, solid-phase library synthesizer was validated using an anti-gp 120 antibody.[134] This system has been a model for many of the systems now in use and was based on the use of filtration microtiter plates, enabling the removal of solutions without loss of the solid-phase support beads and the addition of solutions to the wells using a single- or multiple-head pipetting device. Individual compounds can be synthesized in each well or, by transferring the beads from wells into a mixing device and then redistributing them among the reaction wells, a library could be synthesized using either the split-and-mix one-compound–one-bead approach or any of the iterative, deconvolution, or positional scanning approaches.

In the example 12 individual peptides were synthesized and tested to validate the quality of synthesis. A library of format $RAX_1HTTGRIX_2$ (19 amino acids randomized at each X position) was synthesized and assayed in pools of peptides released from the solid-phase supports using an ELISA assay. After addition of the X_1 amino acids the 19 pools (containing 19 peptides differing in X_2) were kept separate so that the residue(s) in this position would be defined by the initial assay. The anti-gp120 antibody used had been raised against recombinant gp120, and the binding epitope mapped to a 10-mer sequence (RAFHTTGRII). The residues in boldface type had been shown to be important using alanine scans. Screening identified X_1 as F = W = Y > H > A = N. Each mixture was separated by affinity chromatography using the anti-gp120 antibody as the affinity support, and the retained fractions were sequenced to identify X_2, with the finding I > L > P. It would have also been possible to resynthesize the 19 peptides corresponding to each X_1 pool in order to identify X_2, but affinity selection as demonstrated in this work can be a powerful approach in certain circumstances.

Scaffold Library. The use of a scaffold library to produce ligands with a known structure was validated by identifying peptides that bind to an anti-lipopolysaccharide (LPS) monoclonal antibody.[135] A library was synthesized into a highly structured peptide that binds zinc in its three-dimensional conformation, PYKPECFKSFSQK(X/S)(X/D)L(S/V)(X/K)HQ(X/R)THTG, presenting the randomized (X) residues to the target molecule on a defined backbone. Because of the peptides zinc binding, the conformation of the peptide mixtures and individual peptides could be confirmed by Co(II)-complex fluorescence analysis, circular dichroism measurements, and transfer NOE NMR. A positional scanning approach was employed, synthesizing five libraries of the format $O_1X_2X_3X_4X_5$, $X_1O_2X_3X_4X_5$, $X_1X_2O_3X_4X_5$, $X_1X_2X_3O_4X_5$, $X_1X_2X_3X_4O_5$, ~26 pools each (17 proteinogenic, 9 nonproteinogenic amino acids), where the O positions were defined according to the pool, and X represented mixtures. Substituting hydrophobic residues at X_1 caused a major change in conformation, and thus only 14 of the amino acids were used at this position. Screening with the anti-LPS antibody produced $H_1F_2V_3Q_4H_5$ as the best binder, and the backbone conformation of these residues was confirmed by physical studies. The authors propose this as a first step toward synthesis of a peptidomimetic and represents an approach that can be generalized for any target molecule.

In another example of the use of a template library, FmocK-P-G-BocK-A-AlocK-P-G-BocK-A was synthesized, cleaved from the resin, and cyclized to serve as a backbone to which amino acids could be attached selectively to the orthogonally protected lysine side chains.[136] A single amino acid was attached to the two Boc-protected lysines, while eight different amino acids were attached, in separate sets of reactions, to each of the other orthogonally protected lysine residue side chains. The limited library that resulted was screened against an anti-lysozyme antibody (HyHEC-5) which recognized a defined discontinuous epitope of lysozyme. Library screening identified a tyrosine that was not present in the native epitope. A similar concept was employed to display four amino acids from a Kemp's triacid backbone.[57]

Identification of Protein and DNA Binding Peptides

The binding to streptavidin has been used to validate new synthetic combinatorial methods. In general, either the direct binding of (alkaline phosphatase-conjugated) streptavidin/avidin to beads is monitored or biotin is bound to microtiter plates, test peptides are added free in solution, and the inhibition of streptavidin/avidin binding is monitored. While not of therapeutic interest, these studies are of scientific note, since they explore the binding of peptides to a protein that binds a nonpeptide (biotin) with very high affinity. Both L- and D-amino acid-containing peptides are bound by streptavidin and avidin. Other more biologically relevant protein–protein interacting systems have also been studied.

Multiple-Synthesis Libraries. Binding to streptavidin was used to validate a surface display array-based library method called Pilot.[91] The effect of the surface on the binding properties of the compounds assayed while remaining attached to the solid phase on which they were synthesized is an issue. Cass et al. attempted to address the surface presentation issue by exploring the binding of [125]-I-streptavidin to the sequence YGHPQGG synthesized onto various surfaces. They derivatized polyethylene sheets with a functionalized linker (diamino-substituted polyethylene glycol, Jeffamine ED-600) to which they covalently bonded carboxymethyl dextran (Pharmacia T500) to produce a surface coating that was amino functionalized by coupling tBOC-NH-CH$_2$)$_3$ NH$_2$ followed by cleavage of the protecting tBOC. Arrays of compounds were synthesized onto these thin-film sheets in defined positions to create reusable assay plates.

The strategy used for array design is interesting. A standard mixture of 16 selected amino acids was used (W, amino acids Y,F,D,E,K,R,Nle,V,G, DAla,A,S,H,Q,P,DNap) together with limited positions in which a mixture of two defined amino acids was incorporated. Thus, for instance, an 8 × 8 array hexapeptide library was generated by creating all possible two-amino acid combinations of the amino acids used in the W mixture in positions 1 and 2 [(Y/F, D/E, K/R, Nle/V, G/DAla, A/S, H/Q, P/DNap) by (Y/F, D/E, K/R, Nle/V, G/DAla, A/S, H/Q, P/DNap)], while the W mixture was used in positions 3, 4, 5, and 6. Probing with streptavidin produced VY. Generating the subarray to define positions 3 and 4 (same format) produced VYGF and VYHP. Generating the final sub-array to define positions 5 and 6 produced VYGFRQ and VYHPQF. Speculating that these two might overlap the authors combined the two motifs to produce HPQVYGFRQ, which was a strong binder compared to the individual peptides.

The identification of DNA binding peptides from a library was demonstrated by Kramer et al.[39] As described above, they synthesized a XXO$_1$O$_2$XX library on a cellulose membrane sheet in the form of pools in a 19 × 19, O$_1$ by O$_2$, array using the double coupling of 0.8 equimolar (less than saturating) mixtures of amino acids. The cellulose sheet was incubated with a ^{32}P-labeled, double-stranded DNA 15-mer, and analyzed by a Phosphoimager (Molecular Dynamics, Sunnyvale, CA). This led to the identification of a KK, KR, RK, RR consensus for the O$_1$O$_2$ amino acid pair. Considering the crucial role of protein recognition of DNA sequences for transcription, and the current attempts to influence disease states through the selective control of transcription, these results demonstrate the potential for applying peptide combinatorial libraries to these targets.

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. Lam et al. synthesized a pentapeptide library and assayed it directly for streptavidin binding to the compounds displayed on the beads, sequencing to identify the positive compounds, which were then resynthesized and tested to confirm that binding was competitive with biotin. The consensus motif

identified was HPQ, with a lesser number of HPM motif compounds. Furthermore, it did not matter if HPQ was N-terminal (e.g., HPQFV), internal (e.g., LHPQF), or C-terminal (e.g., MYHPQ); of course, the C-terminal position in the library was internal to the linker. The HPM motif containing peptides was displayed either as an internal motif (e.g., MHPMA) or C-terminal motif.

As a follow-up, Lam and Lebl explored the specificity of streptavidin and avidin using a library that lacked histidine.[108] A higher concentration of the streptavidin/alkaline phosphatase conjugate had to be used, but the result was the discovery of four motifs, WPAX or WXAX, RMDLY, and YMEXW, and two additional sequences, PPWPY and QYWQS. Representative compounds that were resynthesized demonstrated competitive binding with biotin, indicating they bound in the biotin binding pocket. This demonstrates the power of synthetic combinatorial libraries in discovering novel leads. Furthermore, when the apparently closely related avidin molecule was used as a target the authors found apparent overlap with streptavidin motif (HPYPX, HPFXX, and HPXPX motifs) as well as a different peptide specificity (HKXXX). Despite the apparent overlap, streptavidin did not bind to HPYPP, and avidin did not bind to the LHPQF, indicating that the HPYP motif is specific for avidin and the HPQ motif is specific for streptavidin, though both bind in the respective biotin binding pockets.

Following this study additional L- and D-amino acid peptides were discovered[109] The results show that the majority of the L-amino acid-containing peptides that bind streptavidin contain the HPQ motif, while avidin binds D-amino acid containing peptides that do not have this motif, but instead exhibit a vynxx, ltxsx, or vqsxw motif, again demonstrating the binding pocket specificity differences that existed between the two biotin binding proteins.

Protein–protein and protein–DNA binding interactions regulate signal transduction, transcription, and translation. SH3- and SH2-mediated protein–protein interaction is an important component in signal transduction, serving to interface selective ligand/receptor interactions into common cellular signal transduction pathways. Specific inhibition of SH3-dependent binding is therefore an attractive therapeutic target. Studies were undertaken utilizing a hexapeptide library and a cyclic heptapeptide library assayed on-bead using the fluoresceinated SH3 domains of phosphatidylinositol 3-kinase (P13KSH3) and Src presence of at first all 32 peptides. The set was divided into two pools of 16 and the active pool was identified, then this pool was split in half, and so on. In the end one high-affinity ligand was identified, Ac₂-L-Lys-D-Ala-D-ALa. The authors demonstrated the utility of this approach, but also evaluated the pool size that could be employed, concluding that the method was useful for pool sizes of 100–1000 compounds. Many methods can identify activity from a pool of compounds, but capillary electrophoresis gives a direct determination of binding affinity and does not require labeling of compound or a specific biological activity.

Identification of Enzyme Substrates

Peptide substrates have proven to be good starting points for developing substrate-based inhibitors of enzymes, and synthetic combinatorial chemistry peptide approaches should produce leads that can then be used for rational drug design. The identification of enzyme substrates from libraries represents a method to improve screening assays and should prove useful in creating screening assays for putative orphan enzymes, which have no known substrates. Such a tool should be instrumental in characterizing and validating new therapeutic targets through identification for substrates for the classification of orphan enzymes, identification of putative substrates from protein sequence database searches based on the sequences identified from library screening, and identification of putative substrates for kinases involved in critical functions such as signal transduction.

Multiple Synthesis Libraries. A multiple-synthesis mixture strategy was used to synthesize a library of 19 Kemptide-related peptides (format L(R/X)RASL, R is the native Kemptide residue) and a library of 19 analogs based on the v-Src autophosphorylation site (RRLIEDA(E/X)WAARG, E is the native v-Src residue). The cleaved peptides were incubated with [g-^{32}P] ATP plus either the cAMP-dependent kinase or the v-Abl protein kinase, and the phosphorylated compounds were identified by phosphopeptide-selective liquid LCMS.[93] When synthesis reached the point of randomization the resin was divided into four pools accounting for 16, 10.5, 31.5, and 42% of the total. Each pool was reacted with a different mixture of amino acids (I,V,T; Q,H; Y,D,E,K,R,P; F,G,A,L,M,S,H,W) so that the same proportion of reactive sites were occupied by equimolar amounts of each amino acid. The compounds were cleaved and the pools were recombined into one mixture of 19 compounds. Quality control was performed by MS, and the libraries were screened against their respective target kinases using [g-^{32}P] ATP. After removal of the enzyme by molecular weight centrifugal filtration (Microcon-10 unit), the peptide mixture was subjected to liquid chromatography. The recovered peaks were split for analysis of ^{32}P incorporation by g counting and specific MS detection of phosphopeptides.

Kemptide was the principle peptide phosphorylated by the cAMP-dependent kinase, but each analog was phosphorylated to some degree. The intensities of the negative ion (phosphorylated peptide, detecting the PO$_3^-$ and PO$_2^-$ ions) versus positive ion (peptide itself) permitted the authors to compute the percent conversion to phosphorylated peptide and rank them (R > H > G > Q = K > E > Y > H > A > P > T > T > S > (xle) > F > (xle) > W > V > M > A) according to suitability as substrates (MS cannot distinguish Ile and Leu, shown as xle). This ranking was consistent with that reported from a traditional study (R > H > K > A)[161] but contained much more information. The study with v-Abl protein kinase identified the Leu or Ile analog as the major substrate, with the Val analog as a minor substrate.

Resynthesis and testing of the Leu and Ile analog peptides compared with the Src substrate Glu peptide demonstrated a preference for the Ile analog as substrate (Glu: $K_m = 1.56$ mM, $V_{max} = 27.8$ nmol/min mg^{-1}; Leu: $K_m = 0.727$ mM, $V_{max} = 20$ nmol/min mg^{-1}; Ile: $K_m = 0.67$ mM, $V_{max} = 73$ nmol/min mg^{-1}). Thus libraries were shown to be effective means of defining substrates for both a receptor-dependent protein kinase (cSMP dependent kinase) and a receptor-independent protein kinase (v-Abl).

In a specialized variation of multiple synthesis, Songyang et al. synthesized a biased library MAXXXXYXXXXAKKK, where X represents 15 proteinogenic amino acids (excluding C, W, Y, S, and T), and then cleaved this library into solution.[162] The mixture was incubated with p60^{c-src} protein tyrosine kinase and [g-^{32}P]ATP and the phosphorylated peptides were isolated from the mixture by ferric chelation chromatography and identified by sequencing. The authors identified as EEIYGEFF motif.

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. Wu et al. identified peptide substrates serin/threonin cAMP-dependent kinase, known to exhibit specificity for peptides containing an RR(X)S motif.[113] A library of peptides displayed on the surface of beads were incubated with the enzyme in the presence [g-^{32}P]ATP. Substrate peptides were phosphorylated, all the beads were spread in agarose to immobilize them, and autoradiography was performed to locate the labeled beads. The ^{32}P-labeled beads were recovered and identified. Substrates that identified by this procedure had higher affinity than the commercial synthetic cAMP-dependent kinase substrate, Kemptide (LRRASLG, $K_a = 1.33$ mM). From approximately 500,000 compounds assayed, 55 positive compounds were selected from a pentapeptide library. Sequencing of two yielded the sequence RRYSV ($K_a = 0.67$ mM). Sequencing three compounds out of 60 positive beads from a hepta-peptide library produced SQRRFST ($K_a = 0.83$ mM), YRRTSLV ($K_a = 1.43$ mM), and IIRRKSE (K_a not determined). Two of these exhibited higher affinity for the enzyme than the commercial peptide.

The same approach was used subsequently to identify substrates for the p60^{c-src} kinase.[114] The sequence YIYGSFK was identified from the library. The phosphorylation of YIYGSFK and several peptides derived from cellular proteins by Src and several other kinases was studied. YIYGSF was found to be a highly specific Scr-family kinase (Scr, Lyn, Fyn) substrate, exhibiting a K_m of 55 mM for Src, compared to a K_m of 353 mM for a peptide derived from cdc2 (residues 6–20, KVEKIGEGTYGVVYK), a native substrate of Src, but not a good substrate for non-Src family protein kinases (EGFR, Fes) compared to the cdc2 peptide. Thus both a higher affinity and more specific substrate was identified.

The next step would be to search sequence data bases to identify whether this motif appears in cellular proteins that may be putative substrates for Src. Information of this type can be used to design substrate based inhibitors.

More recently, by screening a secondary library (XIYXXXX) an even more potent and specific peptide substrate, GIYWHHY, was identified.[115]

In another study the substrate specificity of an endoproteinase was explored using an acetylated library of peptides on beads.[116] Two libraries were prepared, Ac-NleELRTQ(P$_1'$)(P$_2'$)SHR-NH$_2$, with P$_1'$ fixed as phenylalanine and P$_2'$ randomized in the first library, and P$_1'$ randomized with P$_2'$ fixed as serine in the second library. Incubation of the library beads with the 3C endoproteinase from heptatis A virus resulted in removal of the acetyl group from those peptides that were cleaved by the proteinase. The absence of the acetyl group permitted direct identification of the substrate sequence by Edman degradation. Thus there was no selection step, only incubation, assessment of the extent of reaction by measuring free amino groups using fluorescamine, and then sequencing of the entire batch of beads. The native P$_1'$/P$_2'$ residues inferred from other studies are serine/phenylalanine, which demonstrated a K_{cat}/K_m value of 0.004. The library results indicated a P$_1'$ preference for alanine, serine, and glycine, and a less selective P$_2'$ preference, but peptides containing a P$_2'$ arginine or proline were not cleaved. All the peptides that were cleaved were less active than the putative substrate domain, exhibiting K_{cat}/K_m values of 0.1 to 1.0. Thus this study confirmed the putative cleavage specificity to be serine/phenylalanine for P$_1'$/P$_2'$.

The substrate specificity of human fibroblast collagenase and stromelysin was evaluated using an on-bead library approach in which a fluorescent reporter was cleaved into solution, resulting in solution fluorescence in wells containing positive substrate beads.[117–119] Positive beads were recovered and the remaining material was microsequenced to determine the cleavage site. The first assumption was that the library should encompass a tripeptide window displayed in an alanine frame, AAXXXAA. To design the optimal method of covering this window an alanine scan was performed on the peptide G$_4$P$_3$L$_2$A$_1$L$_1$,F$_2$, cleaved by both collagenase and stromelysin. In terms of conventional nomenclature shown as subscripts to the amino acid residues, the P$_3$ proline, P$_2$ leucine, and P$_1'$ leucine, and to a lesser extent the P$_2'$ phenylalanine were important for either activity or selectivity. The authors concluded that PXXL was crucial for collagenase recognition as a substrate, and designed a library of three subforms, all synthesized on controlled pore glass (CPG):

Cop-A$_1$A$_2$A$_3$X$_4$X$_5$B$_6$A$_7$-(ε-aminocaproyl)$_5$-BAla-(3-aminopropyl)-CPG
Cop-A$_1$A$_2$X$_3$A$_4$X$_5$B$_6$A$_7$-(ε-aminocaproyl)$_5$-BAla-(3-aminopropyl)-CPG
Cop-A$_1$A$_2$X$_3$X$_4$A$_5$B$_6$A$_7$-(ε-aminocaproyl)$_5$-BAla-(3-aminopropyl)-CPG

Cop is a fluorophore (7-hydroxycoumarine-4-propanoyl) developed specifically for capping peptides synthesized on solid phase.[163] A represents fixed alanine residues, X represents positions varied one at a time between 20 amino acids such that the identity of the residue was known from the position in the synthetic/assay array, and B represents a degenerate position containing 20 different amino acids. Thus each subformat library consisted of 400 pools in

which the identity of the X residues was known. The 20 proteinogenic amino acids were used, substituting S-methylcysteine for cysteine.

While a more powerful library design/identification strategy would have made it possible to simply synthesize and screen the complete tetrapeptide library, the results obtained with this approach were quite interesting. By defining cassettes of related compounds based on the shared identity of the N-terminal X residue, the relative importance of the C-terminal X residue could be compared. The cassettes defined as L, M, N, and P exhibited collagenase substrate activity. The L, M, and N cassettes exhibited an identical C-terminal X profile of U > L > M > I > Q > Y, while the P cassette had a different profile, N > U > A > D = M > P. Sequencing was used to identify the B residue identity and cleavage site. For instance, sequencing of the positive AAPANBA pool produced an abundance of and preference for B of U > L > I > M in the first round of sequencing, indicating that cleavage occurred N-terminal to these residues, and identifying AAPANUA as a good substrate for collagenase. Additional analysis indicated the abilty to detect frame shifts in the point of cleavage. The authors also confirmed that there was a good correlation between cleavage of the peptides on the bead compared to cleavage of resynthesized peptides in solution. This approach should prove to be quite useful in the characterization of putative proteases identified from gene sequencing efforts and in the optimization of screening assays.

Another strategy used to identify protease substrates relied on an internally quenched fluorescent molecule where the fluorochrome (o-aminobenzamide, ABz) was separated from the quenching moiety (3-nitrotyrosine, $Y(NO_2)$) by the point of cleavage.[120] By synthesizing the library on macroporous beads and employing on-bead screening, the beads that displayed substrate peptides developed fluorescence and could be readily identified and selected. The format of the library was $Y(NO_2)X_5X_4PX_3X_2X_1K(ABz)$ bead, designed for the endopeptidase subtilisin Carlsberg. By inserting proline at the P-2 position the site of cleavage by this enzyme was directed between X_3 and X_2. By performing the assay at pH 8.5 and 5.2 the authors demonstrated that they could not only identify substrates, but they could optimize substrates for different reaction conditions. The substrates with the highest rates of cleavage on beads at pH 8.5 were $Y(NO_2)LQPFNEK(ABz)$, $Y(NO_2)FSPLQ$-$FK(ABz)$, and $Y(NO_2)IEPFFEK(ABz)$, while those with the highest rates of cleavage at pH 5.2 were $Y(NO_2)FQPLDVK(ABz)$, $Y(NO_2)LYPLDVK(ABz)$, and $Y(NO_2)ASPMGFK(ABz)$. However, the authors noted that while reflecting the rate of cleavage, this rank order did not hold in solution where accurate kinetic constants (K_{cat}/K_m) could be determined. Thus, final optimization had to be performed using compounds assayed in solution, producing $Y(NO_2)FQPLDEK(ABz)$ and $Y(NO_2)FQPLAEK(ABz)$ as the most active structures under pH 8.5 conditions.

Iterative Synthesis and Screening Library Methods. Phosphorylation of library peptides was also used to identify substrates of cAMP- and cGMP-

dependent protein kinase.[20] The known consensus for cAMP-protein kinase substrates and inhibitors was RRX(S/A)X, but that for cGMP-protein kinase was less well defined. An iterative approach was used to identify octapeptides, incubating first a library Ac-Ac-XXXO$_1$O$_2$XXX-NH$_2$ with the kinase(s) plus [g-^{32}P]ATP, identifying Ac-XXXRRXXX for cAMP kinase and AcXXXRKXXX-NH$_2$ for cGMP kinase, respectively. Subsequent dependent sets of libraries (Ac-XXXRRO$_1$O$_2$X-NH$_2$, etc.) were synthesized until they had defined Ac-KRAEKASIY-NH$_2$ as the substrate for cAMP kinase (similar to the previously identified consensus motif RRX(S/A)X) and Ac-TQKARK-KSNA-NH$_2$ for cGMP kinase, a novel substrate series.

Validation of Library Design Concepts. An early study in which enzyme substrates were identified was based on the synthesis of several mixtures of peptides. Each mixture contained compounds of identical structure except at one position, which was randomized (e.g., Ac-GNSXR, where X = F, G, A, Y, L, V).[148] The mixtures were then exposed to an enzyme (angiotensin-I converting enzyme, atrial dipeptidyl carboxyhydrolase, bacterial dipeptidyl carboxyhydrolase, and meprin), and the products were analyzed by sequencing to determine which residues were important for binding and where the point of cleavage occurred. One cycle of sequencing revealed the preferred P$_1'$ residues, two cycles the preferred P$_2'$ residue, and so on. The authors made several new observations regarding substrate specificity, but the most interesting was perhaps the observation that meprin exhibited specificity for the Ala-Pro bond, while also cleaving Phe-Ala bonds. The authors speculate that bradykinin may be a physiological substrate of meprin.

Identification of Enzyme Inhibitors

Multiple Synthesis Libraries. Inhibitors of acetylcholinesterase activity were identified using a library of compounds synthesized and then spotted on a silica thin-layer chromatography plate and assayed.[94] Fifty-two compounds were synthesized and spotted in an array. The substrate acetylthiocholine iodide plus 5,5-dithiobis(2-nitrobenzoic acid) and then enzyme were evenly distributed over the plate by spraying. Inhibitors were detected by the clear zones that appeared because of the lack of enzymatic reaction (which produced a yellow background). Cellulose paper could also have been used as the support, in which case synthesis could have been carried out directly on the support. Positive control and several other inhibitors were identified, validating the method.

Synthesis of a library based on a noncleavable peptide inhibitor on pins was used to map the active site of and optimize inhibitor selectivity for human heart chymase, a chymotrypsin-like protease that converts angiotensin I to angiotensin II.[95] Two libraries were synthesized and the compounds were tested for inhibition of either chymotrypsin and chymase. One was based on use of 3-fluorobenzylpyruvamide (F-Phe-CO-) to bind in the P$_1$ pocket of the

enzyme. It consisted of (F-Phe-CO)-X_1X_2-RG-pin, where X_1X_2 residues were each varied among the 20 different proteinogenic amino acids to produce 400 different peptides used to determine the specificity for the P_1' and P_2' residues. This library identified the X_1X_2 residue pair ED as selective for chymase. Synthesis and testing of (F-Phe-CO)-ED-RG(OMe) resulted in the determination of a K_i for chymase of 1 mM and for chymotrypsin of 100 mM. The other library was based on the use of an α-ketoamide (Phe-CO) in the P_1 position, producing the library Z-Ile-X_3X_2-(Phe-CO)-GG-pin, where X_3X_2 residues were each varied among 18 of the different proteinogenic amino acids (Met and Cys were omitted) to produce 324 different peptides used to determine the specificity for the P_3 and P_2 residues.

The results were the identification of chymase P_3P_2 preference for the amino acid pairs E_3P_2 and I_3Q_2. E_3D_2 was the most potent inhibitor, but it was not selective. The authors selected E_3P_2 for its acceptable potency and selectivity, synthesized the peptide Z-IEP-(Phe-CO_2Me) and established a K_i of 1 nM for chymase and 10 nM for chymotrypsin. Clearly, the potency of these two series were quite different, but selectivity of either 100- or 10-fold could be achieved. The authors then combined these series into one molecule, Z-IEP-(Phe-CO)-EDR-OMe, which exhibited a K_i for chymase of 100 nM and a K_i for chymotrypsin of 40 mM, a 400-fold selectivity for chymase with reasonable inhibitory potency. This demonstrates the power of rationally designed combinatorial peptide libraries to quickly optimize the selectivity of enzyme inhibitors.

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. Meldal and Svendsen described a novel approach for identifying enzyme inhibitors.[121] The authors synthesized an intramolecularly quenched (by Ac-Y(NO_2)] fluorescent (Abz) substrate for subtilisin Carlsberg (Ac-Y(NO_2)FQPLAVK(Abz)-PEGA, cleaved at the L-A bond) inside macroporous beads composed of polyethylene glycol cross-linked polyamide (PEGA) together with a library of heptapeptides (XXXxXXVF, composed of the L (X) or D (x) stereoisomers). Enzyme inside the beads thus was in the presence of both a substrate and a potential inhibitor. Cleavage of the substrate present in each and every bead produced fluorescence. However, in those beads containing an inhibitory peptide the development of fluorescence was delayed due to the localized inhibitor. Novel inhibitors were identified, the most potent being AMMc(*tert*-butyl)MIVF (IC50 ~ 3 mM).

The identification of inhibitors of protein tyrosine phosphatase PTP1B (implicated in breast and ovarian cancer) was used to validate a novel tagging method employing a radio-frequency (rf) encodable microchip to encode the synthetic steps to which a compound carrier system was exposed in a split-and-mix, one-compound–one-carrier approach to library synthesis (See Chapter 14).[122] A tripeptide library acetylated with *p*-carboxycinnamic acid (a nonspecific inhibitor) was synthesized on solid-phase resin contained in capsules that also held an rf microchip emitting a unique binary code. The objective

was to confer specificity and additional binding affinity through the peptide interaction with enzyme residues surrounding the p-carboxycinnamic acid binding pocket. Each reaction pool was scanned and the identity of the microchips in that pool was recorded. The capsules from the various pools could then be mixed, split into the next set of reaction pools, and scanned again. Thus, at the end of the synthesis there was a record of which reactions each microchip-encoded capsule had been subjected to. In the actual example, a limited library of 125 compounds was synthesized, so the split-and-pooling strategy was also manually tracked to verify the ability of the rf tracking to accurately report the reaction history of each capsule.

Inhibitors of the glucosomal phosphoglycerate kinase (gPGK) from *Trypanosoma brucei* were identified from a pentapeptide library in an on-bead assay using fluorescently tagged or biotinylated gPGK and either manual selection of the fluorescent beads or selection onto steptavidin-coated magnetic beads.[123] Inhibitors of this gPGK are potential antiparasitic agents, because of the key role this enzyme has in the glycolytic pathway of the form of *Trypanosome brucei* found in the bloodstream, and because it is known that the *Trypanosome* is sensitive to compounds that interfere with glycolysis. Using this approach the authors found that NWMMF was a selective inhibitor of the parasite gPGK, IC_{50} of 80 mM, and was without effect on rabbit muscle gPGK at 500 mM. The next step toward a therapeutic agent is to optimize activity of this parasite-selective peptide and demonstrate anti-parasite activity.

Lowe and Quarrell started with a portion of sequence from a known 29-amino acid residue inhibitor of the serine proteases trypsin and chymotrypsin, CMT-1 (isolated from squash) to design an optimization library for leukocyte elastase.[124] The portion of sequence they used as a template defined a reactive loop in this conformationally constrained inhibitor. The authors synthesized a library containing mutants at positions P_1' to P_4' (C-terminal) to the presumed reactive site occurring between Arg-5 and Ile-6 relative to the N-terminus. Once cleaved this inhibitor remains bound to the enzyme acive site. Other investigators had shown that changing Arg-5 to Val-5 increased binding to human elastase. This led the authors to design two libraries, differing in the fifth residue, RVCP(V_5/A_5)XXXXC$_0$KKDSDCLAECVCLEHGYCG, and randomizing the P_1', P_2', P_3', and P_4' positions with the 19 proteinogenic amino acids. The residues that induced the tertiary conformation of the peptide through disulfide bond formation were all retained. Using this template library and fluorescently labeled human leukocyte elastase, they screened a therapeutically relevant serine protease in an on-bead assay using fluorescence microscopy to select positive beads. Disappointingly, the authors did not actually report on the success of screening or identify the peptides found using this library.

Inhibitors of thrombin are known to bind in both the active catalytic site as well as an exosite. An example where this was used to advantage is the inhibitor Hirulog, fPRPGGGG(NorLeu)GDFEEIPEEYL ($K_i = 2$ nM). Miru-

log was designed on a rational basis from a known peptide that constitutes a pharmacophore that binds in the active site of the enzyme (fPRPG, K_i = 20 mM), a linker, and a sequence derived from a large molecular weight inhibitor, Hirudin, comprising a pharmacophore that binds in an exosite (DFEEIPEEYL, K_i ~ 2 mM). Pentapeptide libraries were assessed using on-bead screening to identify active site inhibitors.[125] From a library of format DLDLD, where a set of D- or L-amino acid isomers were randomized at every other position, several novel compounds were identified. Among them was wFrPf, which exhibited a K_i for thrombin of 3.5 mM, and dYaEw (K_i = 6 mM), compared to a K_i of 21 mM for the known pentapeptide inhibitor fPRPG (the n-terminal sequence of Hirulog, a thrombin inhibitor in clinical trials). From a library of the format LDLDL compounds such as TrFfP (K_i = 117 mM), LiFrP (K_i = 20 mM), and GfFfK (K_i = 23 mM) were identified, while from a library of the format DDDDD compounds such as ffrfh (K_i = 23 mM), ffrpt (K_i = 43 mM), ffr(norleu)(norleu) (K_i = 53 mM), and rlmrf (K_i = 83 mM) were identified.).

One benefit of combinatorial libraries in exploring binding domains around an initial starting point is to identify a novel series, as well as to improve activity of the original. With this in mind an on-bead library was designed with the format fPRPXXXXX-(linker)-bead, randomizing 19 proteinogenic amino acids at each of the X positions.[125] Screening this library identified SEL 1478, FPRPFGYRV-NH$_2$, which exhibited a K_i of 25 nM. Thus, using a single optimization library, the affinity of the peptide lead was optimized from 20 to 25 nM and a novel binding domain was identified. Multiple synthesis was employed in a combined alanine scan and truncation study to rapidly explore the SAR of SEL 1478, with the result that the compound FPRPFG(X)R-NH$_2$ was identified as necessary for maximal activity. Optimization using an expanded library of nonproteinogenic amino acids and peptide bond modifications would be the next logical step in optimization.

Factor Xa is another enzyme in the coagulation cascade. Drug discovery teams have searched for a specific low molecular weight, orally available inhibitor of this enzyme for the last 5–10 years without success. However, by screening an octapeptide library in an on-bead assay a novel lead was identified, YIRLAAFT-NH$_2$ (SEL 1691, K_i = 15 mM).[126,164] Many additional analogs were identified from other peptide libraries, including the consensus motif YIRLA, structurally different from any known peptide serine protease inhibitors. Furthermore, the series was specific for factor Xa. This factor Xa lead was optimized in a combined combinatorial chemistry and traditional medicinal chemistry one-at-a-time synthetic approach.

The first effort was the multiple synthesis of truncation and alanine substitution analogs, resulting in the identification YIRLA, as the structure necessary for activity. Libraries which allowed multiple substitutions within this motif were designed, synthesized, and tested. This provided valuable SAR and, as the most active compound, SEL 2060 (K_i = 150 nM). Traditional medicinal chemistry approaches based on SAR and molecular modeling ultimately pro-

duced a series of very active and specific inhibitors, among them the pentapeptide SEL 2711, Ac-*p*-amidonophenylalanylcyclohexylglycyl-3(methyl pyridinium)alanyl-L-P-NH2. This molecule exhibited a K_i of 3 nM for factor Xa, was >3000-fold selective compared to thrombin and a variety of other enzymes, was not cleaved by any serum enzymes, and exhibited antithrombotic activity in animals after intravenous, subcutaneous, and oral administration.

Iterative Synthesis and Screening Library Methods.　Inhibitors of the HIV-1 protease were identified from a tetrapeptide library.[137] The peptides were screened free in solution in a microplate assay in which the cleavage of a fluorescent substrate by the HIV-1 protease was monitored. The library contained 22 amino acids at each position plus the aspartic acid protease transition-state analog Statine (Sta) at the second position. Twenty-two pools were synthesized to define the first position as Ac-(P/D)-XXX-NH2. Successive rounds of synthesis defined the rest of the peptide to be Ac-PI(Sta)D-Leu/D-Phe)-NH_2, the aspartic acid having lost activity/specificity on the second round of synthesis and screening. Thus, of the 23 residues incorporated at the second position, only the Statine-containing peptides were active, the activity (IC_{50}) of the most active peptide (Ac-PI(Sta)D-Leu)-NH_2) was 1.4 mM. Subsequent subtitution of the D-leucine with amino acids not in the original library yielded three leucine-substituted peptides (I, L, V) with 10-fold greater potency; substitution of isoleucine with valine produced a 2-fold enhancement of activity. Reassessment of the valine series indicated that tryptophan was more active than phenylalanine at the N-terminus, producing the tetrapeptide Ac-WV(Sta)D-Leu)-NH_2, IC_{50} 200 nM. Further optimization from the initial lead produced a pentapeptide with a defined composition Ac-WV(Sta)VX (X not disclosed) with an IC_{50} of 5 nM.

Eichler and Houghten identified inhibitors of trypsin from an interative set of 10 hexapeptide libraries with the P_1 position defined as K or R, synthesized in an array on cotton and screened in solution (Ac-KOXXXX-NH_2, Ac-XKOXXX-NH_2, Ac-XXKOXX-NH_2, Ac-XXXKOX-NH_2, Ac-XXXXKO-NH_2, Ac-ROXXXX-NH_2, Ac-XROXXX-NH_2, Ac-XXROXX-NH_2, Ac-XXXROX-NH_2, Ac-XXXXRO-NH_2).[24] This strategy, incorporating 19 of the proteinogenic amino acids, resulting in 190 pools containing 2,606,420 peptides, produced the sequence Ac-AKIYRP-NH_2, IC_{50} = 46 mM, which was slowly cleaved by enzyme, but was a better inhibitor than the most active hexapeptide sequence derived from naturally occurring trypsin inhibitors (Ac-TKIYNP-NH_2, IC_{50} = 102 mM). The authors went on to design a secondary library to optimize the activity of this novel lead.

Lead optimization was pursued using the iterative process described above[8] of synthesizing 400 pools of composition Ac-O_1O_2XXXX-NH_2, screening, and then synthesizing and screening subsequent iterative sets of 20 libraries each (Ac-$Id_1Id_2O_3$XXX-NH_2, Ac-$Id_1Id_2Id_3O_4$XX-NH_2, etc.) and proved unsuccessful in identifying any inhibitors of trypsin.[24,138] Eichler and Houghten synthesized an iterative set of 10 hexapeptide libraries with the P_1 position defined

(biased) as K or R, as described above, and identified a novel hexamer sequence, Ac-AKIYRP-NH$_2$, IC$_{50}$ = 46 mM. The authors incorporated this into the design of a secondary dodecamer library to optimize the activity of this novel lead. Through another series of iterative synthesis/screening steps the most active dodecapeptide was identified. An example of one library is Ac-XXXAKIYRPOXX-NH$_2$. Altogether 19 libraries were synthesized in which the randomized position and the AKIYRPO motif was varied positionally throughout the template as the KO motif had been positionally varied. The final result was Ac-YYGAKIYRPDKM-NH$_2$, with an IC$_{50}$ of 10 mM.

Petithory et al. had demonstrated that endopeptidase substrates could be identified from peptide libraries. The identification of selective inhibitors of metalloproteases, the rat brain zinc endopeptidases 24-15 and 24-16, which metabolize neurotensin, was achieved using a phosphinic peptide library of the format Z-(L,D)PheY(PO$_2$CH$_2$)-(L,D) (G/A)$_1$-X$_2$X$_3$, where X was substituted by 20 amino acids (See Chapter 9). Phosphinic peptides (with in vivo activity) act as mixed inhibitors of the two peptidases, with nanomolar K_i, without affect on several other zinc peptidases, such as endopeptidase 24-11, angiotensin-converting enzyme, aminopeptidase M, leucine aminopeptidase, and carboxy-peptidases A and B.[26] However by varying the P$_1'$ position between G and A, and randomizing the P$_2'$ (X$_2$) and P$_3'$ (X$_3$) residues the authors identified selective inhibitors from their optimization library.

A single-step deconvolution strategy was used to synthesize the peptides. X$_3$ was coupled to resin in 20 different reactions, the resin was mixed and split, and then X$_2$ was coupled in 20 different reactions. Each of these mixtures was split and then either Z-(L,D)PheY(PO$_2$CH$_2$)-(L,D)G-OH or Z-(L,D)PheY (PO$_2$CH$_2$)-(L,D)A-OH was coupled. This scheme produced 20 pools containing 40 Z-(L,D)PheY(PO$_2$CH$_2$)-G . . . (X$_3$) G/A-X$_2$ each, with G and X$_2$ defined, and 20 pools containing 80 Z-(L,D)PheY(PO$_2$CH$_2$)-(L,D)A . . . (X$_3$) peptides each, with A and X$_2$ defined. The peptides were cleaved from the resin and assayed. The X$_2$ selectivity determined for the Z-(L,D)PheY(PO$_2$CH$_2$)-G-X$_2$X$_3$ sublibrary was R/K for 24-15 and P for 24-16. Twenty Z-(L,D)PheY(PO$_2$CH$_2$)-GRX$_3$ peptides were synthesized, purified by HPLC, and tested, identifying X$_3$ selectivity for 24-15 to be F/M (~400-fold selective, K_i = 2.7/0.35 nM, respectively). Pursuing the Z-(L,D)PheY(PO$_2$CH$_2$)(L,D)A-X$_2$X$_3$ series the authors identified Pro, Nle as unselective, and Arg, Lys as 24-15-selective X$_2$ residues. From the deconvolution synthesis of X$_3$ peptides a nearly 2000-fold selective inhibitor was identified, Z-(L,D)PheY(PO$_2$CH$_2$)-(L,D)A-K-M, exhibiting a 24-15 K_i of 0.12 nM and a 1300-fold selective inhibitor was identified, Z-(L,D)PheY(PO$_2$CH$_2$)-(L,D)A-R-M, exhibiting a 24-15 K_i of 0.07 nM. This is a striking achievement of both selectivity and potency from an optimization library strategy.

Positional Scanning Format Library Methods. Inhibitors of chymotrypsin were discovered using a cyclic peptide template library.[72] A peptide backbone (KKKEG-resin) was synthesized using the tea-bag method of synthesis. Three

different positional libraries were synthesized, requiring three different backbone structures using orthogonally cleavable protecting groups on the lysine residues, depending on whether the residue being coupled was to be defined (Dde) or consisted of a mixture (Boc). The peptide backbone was cyclized between the N-terminal lysine and the glutamate residue before coupling of 19 proteinogenic and 10 carboxylic acids to the side chains of the three lysine residues to generate three positional libraries, cyclo[K(X)K(X)K(O$_1$)E]G-OH, cyclo[K(X)K(O$_2$)K(X)E]G-OH, and cyclo[K(O$_3$)K(X)K(X)E]G-OH. In these libraries O represents positions defined for each pool, X represents positions in which the mixture of side-chain acylating residues were attached, and the square brackets represent the cyclization between the N-terminal K and C-terminal E residues. A novel series of compounds was identified exhibiting side chains of piperonylic acid and 2-thiophenecarboxylic acid, with ~50 mM IC$_{50}$ values (compared to 1.6 mM for chymostatin and 35 mM for Ac-ygyyyr-NH$_2$).

Validation of Library Design Concepts. Using a more constrained and well-defined cyclic peptide scaffold, Domingo et al.[127] investigated the P$_1$ specificity for a series of different proteases. Their approach was based on the synthesis of an 11-mer peptide α-chymotrypsin inhibitor[165] mimicking the larger 6- to 9-kD Bowman–Birk inhibitors. The Bowman–Birk inhibitors present a nine-residue binding loop to proteases spanning the P$_3$ to P$_6'$ binding domain, while the peptide mimic presents a seven-residue loop to the enzyme. The authors synthesized a library (SCTX$_1$SIPPQCY) wherein each member had an identical sequence, randomized only at the P$_1$ position (X$_1$) by the 21 proteinogenic amino acids to generate a mixture of compounds. They screened this mixture for binding to porcine pancreatic elastase, subtilisin BPN', trypsin, α-chymotrypsin, and enzymatically inactive anhydro-α-chymotrypsin and anhydrotrypsin. Positive peptides were selected by passing the mixture over affinity columns composed of immobilized enzyme, and then the retained peptides were sequenced to identify the cleavage site and P$_1$ residues. In each case two sequences were obtained, one for the intact, bound peptide, and one for cleaved peptide. Trypsin exhibited a preference for K,R; α-chymotrypsin for F,Y,M,L,Nle; anhydro-α-chymotrypsin for F,Y; and elastase for L,M. K_i values ranged from 9000 to 10 nM.

Identification of Receptor Antagonists

Multiple-Synthesis Libraries. The method of synthesizing individual peptides on an array of pins and then either directly assaying for activity while the peptides are still bound to the pins or after their release into solution, lends itself to exploring structure activity relationships (SAR) of analogs based on the sequence of a lead molecule. The approach was applied to exploration of substance P analog binding to the NK1 receptor.[96] A radiolabeled binding assay was employed using a rat brain synaptosome tissue preparation. All the

peptides synthesized had been reported previously, the purpose of the experiments described in the initial publication being to demonstrate that the crude material released from the pins could be used to establish SAR without further purification. The same group continued with the approach, applying it to study the SAR of substance P binding to the NK1 receptor through the simultaneous synthesis of 512 analogs to cover all possible D- and L-amino acid combinations of 9 of the 10 native residues of substance P (the C-terminal Met residue was not varied because it was directly incorporated into the support.[97]

This exhaustive SAR produced new insights and a rich database. Most interesting was that single changes were not predictive of dual changes. This is a fundamental difference in traditional medicinal chemistry and combinatorial medicinal chemistry approaches. In the traditional approach the parent molecule is frequently divided into domains and each domain is varied while the rest of the molecule is kept constant. Later on the best modifications of each domain are combined and the molecule is further refined. In the combinatorial approach, multiple changes can be introduced, and optimized simultaneously.

Much of the effort in synthetic combinatorial chemistry is now focused either on the synthesis of nonpeptide libraries or on the use of rational approaches to improve the design and utility of chemical libraries such as the design of a dipeptide library based on physicochemical properties of select amino acids.[166] After observing that the successful conversion of several peptides into nonpeptide peptidomimetics occurred through the discovery of millimolar dipeptides (angiotensin converting enzyme inhibitor CCK-B and CCK-A selective receptor antagonists), the authors considered that just using the 20 naturally occurring amino acids and four N-terminal protecting groups would generate 1600 possible peptides. They designed and synthesized an information-rich library of 256 N-protected dipeptides, encoding the greatest possible diversity of physicochemical information. The minimum analog peptide set was used,[167] plus various N-terminal blocking groups (the des-amino derivatives of three amino acids, A, G, F), and the C-terminus was amidated, providing structures with at least three potential pharmacophores, which were described by the z scales calculated using the physical properties of 55 amino acids. The set of 256 compounds was composed of dipeptides formed from of eight amino acids (D, K, S, F, L, W, T, and V).

This library was used to identify NK3 selective antagonists, followed by more traditional optimization incorporating multiple synthesis where appropriate.[98] The lead from the library was Boc(S)F(S)F-NH$_2$ (IC$_{50}$ = 1550 nM), where the S indicates the chirality of the Phe residues. This compound was somewhat specific for the NK3 receptor, exhibiting an IC$_{50}$ > 10 mM for NK1 and NK2 receptors. Of methylated analogs, only the C-terminal derivative retained activity, producing Boc(S)F(R,S)αMeF-NH$_2$ (IC$_{50}$ = 1520 nM). Exploration of the optimal chirality for the α-MeF and attempts to functionalize the N- and C-terminal groups produced Boc(S)F(R)αMeF-(CH$_2$)$_7$NHCONH$_2$ (IC$_{50}$ = 16 nM), exhibiting >700-fold selectivity for NK3 over NK1 and NK2, and greater potency than the reference compound; SuccDFNMeFGLM-NH$_2$.

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. This method has been successfully applied to discover peptides that bind to the gpIIb/IIIa platelet integrin receptor.[55] A cyclic pentapeptide library containing both D- and L-amino acids was synthesized and assayed by competition of soluble gpIIb/IIIa binding to fibrogen-coated microplates. The peptide antagonists identified from the assay of only 50,000 compounds (a subportion of the library) were CRCDC and CRGdC.

The C5a receptor has been the target of drug discovery efforts for many years. For nearly 5 years a traditional peptide effort was conducted at Abbott Laboratories to discover receptor antagonists using a rational approach focused on the C-terminus of the C5a ligand. A series of peptides containing nonproteinogenic amino acids were identified, but all were agonists, typified by the structure FKA(D-cha)(L-Cha)r-OH (binding $IC_{50} = 70$ nM).[168–170] A subsequent effort carried out at another company did succeed in identifying an antagonist from this series after a similarly extensive effort.[171] This target thus represented a relevant model to compare the advantages of a combinatorial peptide library to traditional peptide approaches.

A library was designed based on an analog of the Abbott agonist (FKP(D-cha)(L-Cha)r-OH) in which the parent residue was coupled to 20% of the resin at each step of the synthesis, while the remaining 80% was split among the other residues (proteinogenic and nonproteinogenic amino acids) used for randomization and the C-terminal argenine was not varied. A library of 1,200,000 compounds was created on a double releasable linker designed to release the peptides as a free C-terminal OH group. Screened to identify only those compounds with a binding IC_{50} better than 50 nM, 37 compounds were identified (e.g., FL(Tic)(D-cha)(L-Cha)r-OH, (NMe-F)LY(D-cha)(L-Cha)-r-OH, FLY(D-cha)(L-Cha)r-OH), among them were close analogs to the antagonist described by Kawai et al.[169] In this case a single library not only provided new SAR, but was successful in converting an agonist into an antagonist. The library approach identified the same lead in a matter of a few weeks of synthesis and screening, which had taken years by a traditional approach.

Iterative Synthesis and Screening Library Methods. Three publications have addressed the discovery and optimization of endothelin receptor antagonists using combinatorial chemistry approaches. Exploring SAR, all possible (64) L- and D-amino acid combinations were assessed for a potent ETR_A hexapeptide ligand (Ac-f(Orn)DIIW-OH, $IC_{50} = 230$ nM), resulting in the identification of three analogs with millimolar activity, two containing 5 D-amino acid replacements, and with Ac-f(D-orn)Diiw-OH exhibiting a comparable IC_{50} as the parent of 250 nM.[139] The technique used was the split-and-mix synthesis of a 30,752-member tripeptide library (containing 31 or 32 residues consisting of proteinogenic or nonproteinogenic amino acids at each position) designed around a known ET antagonist, FR-139317.[77] Thirty-one pools (of mixtures of 992 compounds each) defined by the identity of the N-terminal residue, were screened. The authors then resynthesized and screened successive libraries to

define the remaining two residues. Not only was FR-139317 identified, but extensive SAR around this molecule was obtained. Deconvolution of mixtures from a 19 × 19 combinatorial array, based on a similar but more potent hexapeptide ETR receptor antagonist (Ac-(D-BHG)LDIIW, IC_{50} = 8 nM) investigated by Spellmeyer et al.[139] not only produced extensive SAR, but produced several analogs with comparable affinity as the parent and one analog with higher affinity (Ac-(D-BHG)QDVIW-OH, IC_{50} = 1.7 nM).[140] Thus, all three independent studies demonstrated the power of combinatorial chemistry optimization libraries to identify potent compounds and to provide extensive SAR.

Positional Scanning Format Library Methods. To discover antagonists of the binding of IL-6 to its soluble receptor (sIL-6R), Wallace et al.[149] synthesized a positional scanning pentapeptide library. The ligand, IL-6, was bound to microtiter plates and then the library peptides were added with sIL-6R. Bound sIL-6R was detected with a specific anti-receptor antibody. No antagonists were discovered, so the investigators proceeded to synthesize the same library in a branched, polydentate configuration using lysine residues as the branching points to produce an octameric library. The branched lysine core is shown in Figure 3.13. The library was synthesized onto the eight free amino groups, such that each branch was identical. Thus, a multimeric or polydentate library was created in which each compound consisted of eight identical pentamers displayed off the branched lysine core. Again positional scanning was employed, but this time antagonists were identified, which produced a positional consensus (E/D/Y)(F/I/E)(Y/L/E/F/W)(F/I/W/Y)(W/Y/F) from which a possible 360 compounds could be derived. To simplify follow-up the degree of inhibition caused by certain pools was used to limit the residue selection, reducing the number of compounds synthesized in the follow-up to 36 peptides.
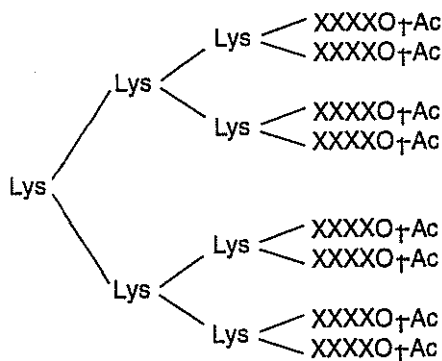

Figure 3.13. Polydentate library: Use of the solid phase and a highly branched linker complex of lysine residues to present multiple copies of the same peptide to the target, enhancing weak binding interactions. Shown is a multimeric "octodentate" library of compounds.

At a concentration of 45 mM, 20 of these produced greater than 50% inhibition, and dose–response curves were obtained for the 18 most active. These results indicated a consensus sequence of (E/D)(F/I)(L/Y)(I/F)W, with an IC_{50} of 75 nM for the six most active peptides (EFLIW, EFYIW, EILFW, EIYFW, EIYIW, DIYFW). Finally, the authors synthesized these as well as negative controls as tetramers and dimers. The tetramers exhibited similar activity as the octamers, but the dimers were inactive.

Ligands that bind the T-cell receptor were defined in order to assess the structural requirements for recognition of the MHC class 1/antigen complexes.[150,151] The mouse MHC class I amorphs H-2K and H-2L were chosen for study because both prefer octapeptides. The authors designed an octapeptide library, synthesizing eight libraries, each containing 19 pools with the position being defined occupied by 19 proteinogenic amino acids (OXXXXXXX, XOXXXXXX, XXOXXXXX, etc.). The ability to bind these peptides was assessed by the restoration of MHC expression. All the libraries contained peptides that bound and stabilized the two MHC class I molecules. In addition, the library pools sensitized target cells for cytolysis by T-lymphocytes, and the authors were able to define the specificity for several T-lymphocyte clones. In this functional assay the individual peptides identified were active in the pico- to femtomolar range. For instance, for one T-cell clone (4G3) that recognized the chicken ovalbumin-derived SIINFEKL epitope, they identified a novel epitope (DIKVGIEF), which was just as active. Even of greater interest was the identification of T-cell epitopes that induced cellular responses but that were inefficient in binding the MHC molecules. These latter molecules would appear to bypass the need for MHC interaction. This work demonstrates the potential of using synthetic combinatorial peptide libraries to understand the immune response and autoimmunity, and represents another approach to the development of synthetic vaccines.

Identification of Receptor Agonists

Iterative Synthesis and Screening Library Methods. A novel agonist of the m-opioid receptor, Ac-rfwink-NH$_2$, was identified from a hexapeptide library by testing the peptides in solution in a traditional radiolabel binding assay employing [^{3}H][D-Ala^2,MePhe^4,Gly^5-ol]enkephalin and rat brain tissue preparations.[142] The selectivity of this all-D-amino acid-containing peptide was high (K_i for $m_1 = 16$ nM and for $m_2 = 41$ nM), compared to the s receptor ($K_i > 1500$ nM) and the k_1 ($K_i > 2000$ nM), k_2 ($K_i > 5000$ nM), and k_3 ($K_i = 288$ nM) receptor subtypes. Functional m-opioid receptor agonist activity (reversed by naloxone) was demonstrated on guinea pig ileum where Ac-rfwink-NH$_2$ exhibited an IC_{50} of 433 nM, compared to leu-enkephalin, which exhibited IC_{50} of 246 nM. In contrast, Ac-rfwink-NH$_2$ was functionally 100-fold less potent at the s receptor than leu-enkephalin, measured as the inhibition of electrically evoked contractions or mouse vas deferens tissue. Finally, intracerebroventricular injection of Ac-rfwink-NH$_2$ (3 nmol) showed analgesic prop-

erties in a nociception tail-flick model, exhibiting a long duration of action (120 min). Noteworthy is that the analgesic properties of this compound are more potent and longer lasting than those of morphine. The analgesic effect of Ac-rfwink-NH$_2$ was reversed by the irreversible m-selective antagonist β-funaltrexamine. Given intraperitoneal, Ac-rfwink-NH$_2$ also exhibited analgesic effects that were equipotent with morphine and reversed by intracerebroventricular naloxone. This observation suggests that the compound has bioavailability comparable to morphine. Due to its novel structure, bioavailability, potent activity, and good duration of effect, this compound or close analogs are being pursued as analgesics in an ongoing preclinical program.

L-Amino acid hexapeptides were also identified using a similar strategy.[141] While very potent at the m receptor, (Ac-RFMWMT-NH$_2$ K_i 0.8 nM, Ac-RFMWMK-NH$_2$ K_i 0.4 nM, Ac-RFMWMR-NH$_2$ K_i 0.5 nM), these compounds were less selective at the s receptor (K_i of 0.9, 5.6, and 7.4 nM, respectively) and the k$_3$ receptor (K_i of 0.6, 0.4, and 1.4 nM, respectively) than Ac-rfwink-NH$_2$. Subsequently, Dooley et al. identified several unrelated m-opioid agonists series, YPFGFX-NH$_2$ (X = R, IC$_{50}$ = 13 nM), WWPKHX-NH$_2$ (X = G, IC$_{50}$ = 9 nM), and Ac-FRWWYX-NH$_2$ (X = M, IC$_{50}$ = 33 nM), as well as antagonist, Ac-RWIGWR-NH$_2$ (IC$_{50}$ = 5 nM).[143] Interestingly, the truncated sequence WWPK-NH$_2$ exhibited an IC$_{50}$ of 17 nM, WWPR-NH$_2$ exhibited an IC$_{50}$ of 10 nM, and both retained >2000-fold selectivity for m- over s- and k-opioid receptors. These activities compare quite impressively with DAMGO (IC$_{50}$ = 3 nM; 253- and 844-fold selectivity over s and k, respectively) and YGGFL-OH (IC$_{50}$ = 5 nM; 0.7- and 257-fold selectivity over s and k, respectively. Relevant to the use of the iterative approach, the authors observed that at early stages in the iteration the YGG motif was more active, but because they pursued less active pools as well, they were able to define the YPFGFX-NH$_2$ and WWPKHX-NH$_2$ series, which were much more active and selective that the best YGGFMX-NH$_2$ molecules (IC$_{50}$ = 28 nM).

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. A novel double-release assay was described for the identification of agonists and antagonists of G-protein coupled receptors.[129,130] After library synthesis, dry beads attached by surface effect to polyethylene film were exposed to gaseous TFA in a controlled manner to cleave ~4% of compound from each, and then neutralized with gaseous ammonium hydroxide. The released compound remained trapped in the microporous beads. For the cellular receptor assay, tissue culture dishes of confluent frog (*Xenopus laevis*) melanocytes, transiently expressing the bombesin receptor introduced by electroporation or expressing either the frog α-MSH receptor or the transfected human α-MSH receptor, were treated with 1 nM melatonin. They were then covered with a thin layer of agarose containing melatonin and, in the case of the α-MSH receptor studies, α-MSH. The polyethylene sheet displaying the library beads was inverted over this.

With the beads now in (agarose) solution the released compounds were

free to diffuse out, contacting the underlying melanocytes. In this cellular system the pigmentation of the melanocytes is controlled through variations in the level of cAMP, which in turn is regulated by the G-coupled (bombesin or α-MSH) receptor. Agonists increase pigment darkening, whereas antagonists decrease pigmentation. In the bombesin studies, the plates with positive responses were frozen and the polyethylene film was removed (leaving the beads imbedded in the agarose). Image analysis was used to locate the positive beads, which were then collected by aspiration, washed, and subjected to a second round of release and screening.

The library synthesized for bombesin was derived from the sequence of bombesin, residues 8–14 (YAVGHLM-NH$_2$), and had the composition X_1X_2 VGHLX$_3$-NH$_2$. X_1 and X_2 were randomized with 19 L-amino acids, and X_3 was randomized with L, M, P, W to produce a 1444-member library. In the first round of screening 5000 beads were assayed and 500 were selected and reassayed. From this second round of screening 52 beads were selected and reassayed to produce four positive compounds. The beads producing these compounds were recovered and sequenced to determine their identity. The activities of agonists identified ranged from ~100 nM (WFVGHLM and WAVGHLM) to 10 mM (AWVGHLM), compared to bombesin itself, which exhibited an EC$_{50}$ of ~0.1 nM.

This study demonstrated an elegant method for identifying receptor agonists. The α-MSH study was directed toward the identification of antagonists. α-MSH is a 13-mer, while the library synthesized for screening was a tripeptide library released as 48 pools of Ac-X$_1$X$_2$X$_3$-NH$_2$ and 48 pools of X$_1$X$_2$X$_3$-NH$_2$ in which X_1 was defined for each pool. Iterative synthesis and screening of subpools was used to define X_2 and X_3, though it would also have been possible to recover the beads corresponding to underlying pigmentation changes and determine the identity of the compounds from information on the beads had a coding method been employed as well as either a multiple release pooling strategy or more sparse seeding of beads over the cell layer. From these studies the authors identified numerous antagonists, among them the most active was wRL-NH$_2$ ($K_d = 63$ nM). The potential for this approach is that any G-coupled receptor transfected into the melanocytes will be automatically coupled to the pigmentation response through stimulus-dependent effects on cAMP levels. Thus, this system represents a universal assay for a wide variety of therapeutically relevant receptors that is capable of identifying both agonists and antagonists.

Expanding on this work the same group synthesized three libraries based on the α-MSH-[5–13] sequence: a 7-mer (minux X_1X_2), 8-mer (minus X_1), and 9-mer library, with the format X_1X_2(D,L-F)$_3$(D,L-R)$_4$(D,L-W)$_5$X$_6$K$_7$P$_8$V$_9$-NH$_2$.[129] The approach used was split and mix, one compound–one bead, but split into 8 subpools depending on the chirality of the residues 3, 4, and 5 (e.g., LLL, DLL, LDL, etc.). Thus, by knowing the pool and sequencing the individual active peptide the authors could identify the compound including the chirality. Because their assay enabled them to assess the activity of the

compounds from individual beads, as well as to identify both agonist and antagonist activity, they could assess individual responses within pools and not worry that compounds would interfere or cancel due to the simultaneous presence of an antagonist or agonist within the same pool. They identified an antagonist MP(D-F)(R)(D-W)FKPV-NH$_2$ (IC$_{50}$ = 11 nM), a partial agonist MP(D-F)(R)(D-W)WKPV-NH$_2$ (EC$_{50}$ = 690 nM), and a full agonist FH (D-F)(R)(L-W)QKPV-NH$_2$ (EC$_{50}$ = 0.4 nM); the agonist activity of α-MSH-[5-13] itself was EC$_{50}$ = 2 mM.

Identification of Functional Inhibitors

Iterative Synthesis and Screening Library Methods. Synthetic combinatorial peptide libraries have been screened to identify antimicrobial agents. In several reports of this application an iterative approach was used to identify L-amino acid-containing hexapeptides that exhibited antimicrobial activity,[8] as well as D-amino acid- and nonproteinogenic amino acid-containing peptides with or without acetylation of the N-terminus or other postsynthetic modification.[145] A liquid microbial growth assay was employed, and from these studies the peptide Ac-RRWWCR-NH$_2$ was identified as the most active among the L-amino acid-containing peptides, exhibiting a minimal inhibitory concentration (MIC) of 3.4 mg/mL against *Staphylococcus aureus*. From the libraries containing 56 different L, D, and nonproteinogenic amino acids at each position, four tetramers were identified: (aFmoc-elys)WfR-NH$_2$, (aFmoc-elys)Wfl-NH$_2$, (aFmoc-elys)Wf(aAIB)-NH$_2$, and (aFmoc-elys)Wfi-NH$_2$. The latter exhibited specific bacteriostatic activity for gram-positive bacteria.

Positional Scanning Format Library Methods. A hexapeptide positional scanning library was also used to identify peptides that inhibited the ability of melittin to be lyse red blood cells.[152] From a library containing 52,128,400 peptides, this study identified Ac-IVIFDC-NH2, with an IC$_{50}$ of 11 mg/mL. This compound had no similarity to melittin and inhibited hemolysis in a competitive manner, which was overcome by raising the concentration of melittin. The authors suggested that the effect of the peptide is through a specific interaction, perhaps with melittin itself, rather than through a nonspecific membrane effect.

Split-and-Mix, One-Compound–One-Bead, Postassay Identification Format Libraries. A novel antimicrobial assay has been described based on pH-sensitive single release.[131] The method was validated by use of analogs of the antimicrobial compound Ac-RRWWCR-NH$_2$, described originally by Houghten et al.[8,144] Multiple synthesis of an alanine scan series and N- and C-terminal truncations of the peptide rrwwcrc identified the importance of a C-terminal crc motif. There was only a marginal effect of deleting the N-terminal rr residues, and that the single replacements of the ww residues had the greatest effect on activity of any of the alanine replacements (the C-terminal

cysteine was not replaced). Single replacements of the other residues had no affect on activity.

Once validated, the method was used to screen a pentapeptide library of D-proteinogenic and nonproteinogenic amino acids. The library peptides were synthesized on a linker that cleaved in aqueous solution at pH 7. The library beads were mixed in agarose with bacteria, poured into tissue culture plates (where the agar hardened), and incubated a sufficient period of time (12 h) for a lawn of bacteria to grow. Compounds were cleaved by the pH 7 medium and diffused out of the beads. Where the released peptides inhibited bacterial growth a clear ring formed around the bead of origin. The bead in the center of each clear zone was recovered and the identity of the compound was determined (either by directly sequencing or through sequencing of an encoding molecule). A series of compounds were identified, all analogs of the Houghten compound.

DISCUSSION

Peptide combinatorial chemistry has confronted the medicinal chemist with the problem of how to carry out challenging reactions in a routinely predictable, automatable manner. A variety of solid supports have been explored and are generally available for the solid-phase synthesis of the various types of library formats. Similarly, much work has been devoted to overcoming the synthetic issues that are specific to library synthesis. Good methods exist to assure the equimolar incorporation of multiple subunits and to monitor reactions at the level of single beads, a necessity in cases of library synthesis, with the exception of multiple synthesis. Newer methods of structure determination have also been described, and with the advent of automation and deconvolution methods, this is often more an iterative or deductive process rather than an analytical challenge.

There are growing bodies of literature addressing cyclization of peptides, the scaffolds created to confer structural rigidity onto library members, and the functionalization of side chains; N- and C-termini, and the peptide backbone. One successful approach has been the use of an N-substituted glycine backbone, referred to commonly as peptoids (see Chapter 6).

The three most common synthetic and screening combinatorial approaches being employed are multiple synthesis, iterative/deconvolution synthesis/identification, and one-compound–one-bead synthesis and postassay identification. All three general approaches work, though in specific circumstances one approach may be the most appropriate. Multiple synthesis plays an important role in interfacing lead generation with lead optimization, but it can also serve a valuable lead generation function in situations where there is basis for good design, just as the potentially larger iterative or one-compound–one-bead libraries can play a crucial role in lead optimization and SAR generation. The advantage of a sophisticated one-compound–one-bead/encoding/sequencing

approach must be weighed against the investment that must be made in analytics for compound identification and in developing compatible chemical reactions and the expense of the linkers. Similarly, the elegance of a single synthesis must be compared with the need for iterative/recursive synthesis and screening, and the advantages of identifying an active compound versus elucidating the structure of a compound that should be active from the iterative algorithm. The synthesis and screening of mixtures is not just acceptable, but provides sound data and is in many instances crucial to success.

The research results that have been reviewed in this chapter demonstrate that combinatorial peptide approaches can identify peptide leads for many classes of targets that are of current therapeutic interest. For targets that have been extensively studied by traditional means, peptide combinatorial chemistry has frequently discovered novel leads and provided valuable new SAR. This in large part may be due to the ability to synthesize and screen so many compounds, and to incorporate many unnatural amino acids. It is also likely that such success is due to a difference between traditional and combinatorial approaches. In the former, medicinal chemists usually make one change at a time and assesss the affects, while in the latter multiple changes are made at one time, permitting the chemist for the first time to efficiently explore the role of and synergistic intramolecular interactions.

There are fewer reports demonstrating the use of combinatorial peptide chemistry to optimize lead structures, but those that are reviewed here indicate the power of the approach. Identifying more leads is not the solution to all problems of drug discovery; it simply moves the bottleneck upward in the process to optimization. Recognition that combinatorial chemistry approaches can assist in lead optimization is fundamentally important to the combinatorial chemistry/drug discovery revolution. To realize the real benefits of combinatorial chemistry, the drug discovery team must think and apply the combinatorial paradigm through all phases of optimization, including animal testing.

REFERENCES

1. Pavia MR, Sawyer TK, Moos WH (1993): The generation of molecular diversity. *Bioorg Med Chem Lett* 3:387–396.
2. Gallop MA, Barrett RW, Dower WJ, Fodor SPA, Gordon EM (1994): Applications of combinatorial technologies to drug discovery, 1: background and peptide combinatorial libraries. *J Med Chem* 37:1233–1251.
3. Gordon EM, Barrett RW, Dower WJ, Fodor SPA, Gallop MA (1994): Applications of combinatorial technologies to drug discovery, 2: combinatorial organic synthesis, library screening strategies, and future directions. *J Med Chem* 37:1385–1401.
4. Janda KD (1994): Tagged versus untagged libraries; methods for the generation and screening of combinatorial chemical libraries. *Proc Natl Acad Sci USA* 91:10779–10785.

5. Lebl M, Krchnak V, Sepetov NF, Seligmann B, Strop P, Felder S, Lam KS (1995): One-bead–one-structure combinatorial libraries. *Biopolymers (Pept Sci)* 37: 177–198.

6. Terrett NK, Gardner M, Fordon DW, Kobylecki RJ, Steele J (1995): Combinatorial synthesis: the design of compound libraries and their application to drug discovery. *Tetrahedron* 51:8135–8173.

7. Geysen HM, Meloen RH, Barteling SJ (1984): Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc Natl Acad Sci USA* 81:3998–4002.

8. Houghten RA, Pinilla C, Blondelle SE, Appel JR, Dooley CT, Cuervo JH (1991): Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature* 354:84–86.

9. Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kasmierski WM, Knapp RJ (1991): A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 354:82–84.

10. Erb E, Janda KD, Brenner S (1994): Recursive deconvolution of combinatorial chemical libraries. *Proc Natl Acad Sci USA* 91:11422–11426.

11. Furka A, Sebestyen F, Asgedom M, Dibo G (1991): General method for rapid synthesis of multicomponent peptide mixtures. *Int J Pept Protein Res* 37:487–493.

12. Hortin, GL, Stantz WD, Santoro SA (1992): Preparation of soluble peptide libraries: application to studies of platelet adhesion sequences. *Biochem Int* 26:731–738.

13. Rapp W, Zhang L, Habich R, Bayer E (1989): Polystyrene–polyoxyethylene graft copolymer for high-speed peptide synthesis. In Jung G, Bayer E, eds. *Pept 88, Proc 20th Eur Pept Soc.* Berlin: de Gruyter, pp 199–201.

14. Barany G, Albericio F, Biancalana S, Bontems SL, Chang JL, Eritja R, Ferrer M, Fields CG, Fields GB, Lyttle MH, Sole NA, Tian Z, Van Abel RJ, Wright PB, Salipsky S, Hudson D (1992): Biopolymer synthesis on novel polyethylene glycol–polystyrene (PEG-PS) graft supports, In Smith JA, Rivier JE, eds. *Peptides: Chemistry and Biology.* Leiden: ESCOM Sci, pp 603–604.

15. Geysen HM, Meloen RH, Mason TJ (1986): A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol Immunol* 23:709–715.

16. Bray AM, Valerio RM, Dipasquale AJ, Greig J, Maeji NJ (1995): Multiple synthesis by the multipin method as a methodological tool. *J Pept Sci* 1:80–87.

17. Maeji NJU, Bray AM, Valerio RM, Wang W (1995): Larger-scale multipin peptide synthesis. *Pept Res* 8:33–38.

18. Bray AM, Chierfari DS, Valerio RM, Maeji NJ (1995): Rapid optimization of organic reactions on solid phase using multipin approach: synthesis of 4-aminoproline analogues by reductive amination. *Tetrahedron Lett* 36:5081–5084.

19. Frank R (1995): Simultaneous and combinatorial chemical synthesis techniques for the generation and screening of molecular diversity. *J Biotechnol* 41:259–272.

20. Tegge W, Frank R, Hofmann F, Dostmann WRG (1995): Determination of cyclic nucleotide-dependent protein kinase substrate specificity by the use of peptide libraries on cellulose paper. *Biochemistry* 34:10569–10577.

21. Pokorny V, Mudra P, Jehnicka J, Zenisek K, Pavlik M, Voburka Z, Rinnova M, Stierandova, A, Lucka AW, Eichler J, Houghten RA, Lebl M (1994): Compas 242: new type of multiple peptide synthesizer utilizing cotton and tea bag technology. In

Epton R, ed. *Innovative Perspect Solid Phase Syn Collects Pept Int Symp, 3d* Oxford, UK: Mayflower Worldwide, pp 643–648.

22. Eichler J, Houghten RA, Lebl M (1996): Inclusion volume solid-phase peptide synthesis. *J Pept Sci,* 2:240–244.

23. Eichler J, Bienert M, Stierandova A, Lebl M (1991): Evaluation of cotton as a carrier for solid phase peptide synthesis. *Pept Res* 4:296–307.

24. Eichler J, Houghten RA (1993): Identification of substrate-analog trypsin inhibitors through the screening of synthetic peptide combinatorial libraries. *Biochemistry* 32:11033–11041.

25. Stankova M, Wade S, Lam KS, Lebl M (1994): Synthesis of combinatorial libraries with only one representation of each structure. *Pept Res* 7:292–298.

26. Krchnak V, Vagner J (1990): Color-monitored solid-phase multiple peptide synthesis under low-pressure continuous flow conditions. *Pept Res* 3:182–193.

27. Houghten RA (1985): General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen–antibody interaction at the level of individual amino acids. *Proc Natl Acad Sci USA* 82:5131–5135.

28. Steward JM, Young CM (1984): *Solid Phase Peptide Synthesis.* Rockford, IL: Pierce Chemical.

29. Grant GA, Ed (1992): *Synthetic Peptides: A User's Guide.* New York: Freeman, pp 1–382.

30. Merrifield B (1995): Solid-phase peptide synthesis. In Gutte B, ed. *Pept: Synthesis Structures and Applications.* San Diego, CA: Academic, pp 94–169.

31. Lam KS, Lebl M (1994): Selectide technology: bead-binding screening. *Methods: A Companion to Methods in Enzymology* 6:372–380.

32. Lebl M, Krchnak V, Salmon SE, Lam KS (1994): Screening of completely random one-bead–one-peptide libraries for activities in solution. *Methods: A Companion to Methods in Enzymology* 6:381–387.

33. Carpino LA, El-Faham A (1995): Tetramethylfluoroformamidium hexafluorophosphate: a rapid-acting peptide coupling reagent for solution and solid phase peptide synthesis. *J Am Chem Soc* 117:5401–5402.

34. Krchnak V, Vagner J, Safar P, Lebl M (1988): Noninvasive continuous monitoring of solid-phase peptide synthesis by acid–base indicator. *Collect Czech Chem Commun* 53:2542–2548.

35. Kaiser E, Colescott RL, Bossinger CD, Cook PI (1969): Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal Biochem* 34:595–598.

36. Rutter WJ, Santi DV (1991): General method for producing and selecting peptides with specific properties. *US Patent 5,010,175.*

37. Pinilla C, Appel JR, Blanc P, Houghten RA (1992): Rapid identification of high-affinity peptide ligands using positional scanning synthetic peptide combinatorial libraries. *Bio Techniques* 13:901–905.

38. Pinilla C, Appel J, Blondelle SE, Dooley CT, Dorner B, Eichler J, Ostresh J, Houghten RA (1995): A review of the utility of soluble peptide combinatorial libraries. *Biopolymers (Pept Sci)* 37:221–240.

39. Kramer A, Volkmer–Engert R, Malin R, Reineke U, Schneider–Mergener J (1993): Simultaneous synthesis of peptide libraries on single resin and continuous

cellulose membrane supports: examples for the identification of protein, metal, and DNA binding peptide mixtures. *Pept Res* 6:314–319.

40. Andrews PC, Boyd J, Loo OR, Zhao R, Zhu CQ, Grant K, Williams S (1994): Synthesis of uniform peptide libraries and methods for physico-chemical analysis, In Crabb JW, ed. *Techniques in Protein Chemistry V*. San Diego: Academic, pp 485–492.

41. Frank R (1994): Spot synthesis: an easy and flexible tool to study molecular recognition. In Epton R, ed. *Innovative Perspect Solid Phase Syn Collect Pept Int Symp, 3d.* Birmingham, UK: Mayflower Worldwide, pp 509–512.

42. Zuckermann RN, Kerr JM, Siani MA, Banville SC (1992): Design, construction and application of a fully automated equimolar peptide mixture synthesizer. *Int J Pept Protein Res* 40:497–506.

43. Saneii HH, Shannon JD, Miceli RM, Fischer HD, Smith CW (1993): Comparison of a phage-derived peptide library with one synthesized chemically on the peptide librarian. *Pept Chem* 31:117–120.

44. Bartak Z, Bolf J, Kalousek J, Mudra P, Pavlik M, Pokorny V, Rinnova M, Voburka, Z, Zenisek K, Krchnak V, Lebl M, Salmon SE, Lam KS. (1994): Design and construction of the automatic peptide library synthesizer. *Methods: A Companion to Methods in Enzymology* 6:432–437.

45a. Fodor SPA, Read JL, Pierrung MC, Stryer L, Lu AT, Solas D (1991): Light-directed, spatially addressable parallel chemical synthesis. *Science* 251:767–773.

45b. Fodor SPA, Read JL, Pierrung MC, Stryer L, Lu AT, Solas D. (1991b): Light-directed, spatially addressable parallel chemical synthesis. *Science* 251:767–773.

46. Jacobs JW, Fodor SPA (1994): Combinatorial chemistry: applications of light-directed chemical synthesis. *Trends Biotechol* 12:19–26.

47. Han H, Wolfe MM, Brenner S, Janda KD (1995): Liquid-phase combinatorial chemistry. *Proc Natl Acad Sci USA* 92:6419–6423.

48. Dooley CT, Houghten RA (1995): A comparison of combinatorial library approaches for the study of opioid compounds. *Perspect Drug Disc Design* 2:287–304.

49. Pinilla C, Buencamino J, Appel JR, Hopp TP, Houghten RA (1995): Mapping the detailed specificity of a calcium-dependent monoclonal antibody through the use of soluble positional scanning combinatorial libraries: identification of potent calcium-independent antigens. *Mol Diversity* 1:21–28.

50. Deprez B, Williard X, Bourel L, Coste H, Hyafil F, Tartar A (1995): Orthogonal combinatorial chemical libraries. *J Am Chem Soc* 117:5405–5406.

51. Blake J, Litzi–Davis L (1992): Evaluation of peptide libraries: an interative strategy to analyze the reactivity of peptide mixtures with antibodies. *Bioconjugate Chem* 3:510–513.

52. Holmes CP, Rybak CM (1994): Peptide reversal on solid supports: a technique for the generation of C-terminal exposed peptide libraries. In Hodges RS, Smith JA, eds. *Pept: Proc 13 Am Pept Symp*. Leiden: ESCOM Sci, pp 992–994.

53. Kania RS, Zuckermann RN, Marlowe CK, (1994): Free C-terminal resin-bound peptides: reversal of peptide orientation via a cyclization/cleavage protocol. *J Am Chem Soc* 116:8835–8836.

54. Lebl M, Krchnak V, Sepetov NF, Nikolaev V, Stierandova A, Safar P, Seligmann

B, Strop P, Thorpe D, Felder S, Lake DF, Lam KS, Salmon SEL (1994): One-bead–one-structure libraries. In Epton R, ed. *Innovative Perspect Solid Phase Syn Collect Pept Int Symp, 3d.* Birmingham, UK: Mayflower Worldwide, pp 233–238.

55. Salmon SE, Lam KS, Lebl M, Kandola A, Khattri PS, Wade S, Patek M, Kocis P, Krchnak V, Thorpe D, Felder S (1993): Discovery of biologically active peptides in random libraries: solution-phase testing after staged orthogonal release from resin beads. *Proc Natl Acad Sci USA* 90:11708–11712.

56. Lebl M, Patek M, Kocis P, Krchnak V, Hruby VJ, Salmono SE, Lam KS (1993): Multiple release of equimolar amounts of peptides from a polymeric carrier using orthogonal linkage-cleavage chemistry. *Int J Pept Protein Res* 41:201–203.

57. Kocis P, Krchnak V, Lebl M (1993): Symmetrical structure allowing the selective multiple release of a defined quantity of peptide from a single bead of polymeric support. *Tetrahedron Lett* 34:7251–7252.

58. Jayawickreme CK, Graminski GF, Quillan JM, Lerner MR (1994): Creation and functional screening of a multi-use peptide library. *Proc Natl Acad Sci USA* 91:1614–1618.

59. Burbaum JJ, Ohlmeyer MHJ. Reader JC, Henderson I, Dillard LW, Li G, Randle TL, Sigal NH, Chelsky D, Balwin JJ (1995): A paradigm for drug discovery employing encoded combinatorial libraries. *Proc Natl Acad Sci USA* 92:6027–6031.

60. Madden D, Krchnak V, Lebl M (1995): Synthetic combinatorial libraries: views on techniques and their application. *Perspect Drug Disc Design* 2:269–286.

61. Sepetov N, Krchnak V, Stankova M, Wade S, Lam KS, Lebl M (1995): Library of libraries: approach to synthetic combinatorial library design and screening of "pharmacophore" motifs. *Proc Natl Acad Sci USA* 92:5426–5430.

62. Lebl M, Krchnak V, Stierandova A, Safar P, Kocis P, Nikolaev V, Sepetov NF, Ferguson R, Seligmann B, Lam KS, Salmon SE (1994): Nonsequencable and/or nonpeptide libraries. In Hodges RS, Smith JA, eds. *Pept: Proc 13 Am Pept Symp.* Leiden: ESCOM Sci, pp 1007–1008.

63. Lebl M, Krchnak V, Safar P, Stierandova A, Sepetov NF, Kocis P, Lam KS. (1994): Construction and screening of libraries of peptide and nonpeptide structures. In Crabb JW, ed. *Techniques in Protein Chem V.* San Diego, CA: Academic, pp 541–548.

64. Safar P, Stierandova A, Lebl M (1995): Amino acid like subunits based on iminodiacetic acid and their application in linear and DKP libraries. In Maia HLS, ed. *Pept 94, Proc 23d Eur Pept Soc.* Leider: ESCOM Sci pp 471–472.

65. Krchnak V, Weichsel AS, Cabel D, Lebl M (1995): Linear presentation of variable side-chain spacing in a highly diverse combinatorial library. *Pept Res* 8:198–205.

66. Simon RJ, Kaina RS, Zuckermann RN, Huebner VD, Jewell DA, Banville S, Ng S, Wang L, Hosenberg S, Marlowe CK, Spellmeyer DC, Tan R, Frnkel AD, Santi DV, Cohen FE, Bartlett PA (1992): Peptoids: a modular approach to drug discovery. *Proc Natl Acad Sci USA* 89:9367–9371.

67. Zuckermann RN, Martin EJ, Spellmeyer DC, Stauber GB, Shoemaker KR, Kerr JM, Figliozzi GM, Goff DA, Siani MA, Simon RJ, Banville SC, Brown EG, Wang L, Richter LS, Moos WH (1994): Discovery of nanomolar ligands for 7-transmembrane G-protein-coupled receptors form a diverse *N*-(substituted)gly-cine peptoid library. *J Med Chem* 37:2678–2685.

68. Zuckerman RN, Kerr JM, Kent SB:H, Moos WH (1992): Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis. J Am Chem Soc 114:10646–10647.

69. Ellman JA, Thompson LA (1996): Synthesis and application of small molecule libraries. *Chem Rev* 96:555–600.

70. Rinnova M, Lebl M (1996): Molecular diversity and libraries of structures: synthesis and screening. *Collect Czech Chem Commun* 61:171–231.

71. Lam KS, Lebl M, Wade S, Stierandova A, Khattri PS, Collins N, Hruby VJ (1994): Streptavidin–peptide interaction as a model system for molecular recognition. In Hodges RS, Smith JA, eds. *Pept: Proc 13 Am Pept Sym* Leiden: ESCOM Sci, pp 1005–1006.

72. Eichler J, Lucka AW, Houghten RA (1994): Cyclic peptide template combinatorial libraries: synthesis and identification of chymotrypsin inhibitors. *Pept Res* 7:300–307.

73. Spatola AF, Darlak K, Romanovskis P (1996): An approch to cyclic peptide libraries: reducing epimerization in medium sized rings during solid phase synthesis. *Tetrahedron Lett* 37:591–594.

74. Virgilio AA, Ellmann JA (1994): Simultaneous solid-phase synthesis of β-turn mimetics incorporating side-chain functionality. *J Am Chem Soc* 116:11580–11581.

75. Blondelle SE, Takahashi E, Houghten RA, Perez–Paya E (1996): Rapid identification of compounds with enhanced antimicrobial activity by using conformationally defined combinatorial libraries. *Biochem J* 313:141–147.

76. Jiracek J, Yiotakis A, Vincet B, Lecoq A, Nicolaou A, Checler F, Dive V (1995) Development of highly potent and selective phosphinic peptide inhibitors of zinc endopeptidase 24–15 using combinatorial chemistry. *J Biol Chem* 270:21701–21706.

77. Terrett NK, Bojanic D, Brown D, Bungay PJ, Gardner M, Gordon DW, Mayers CJ, Steele J (1995): The combinatorial synthesis of a 30,752-compound library: discovery of SAR around the endothelin antagonist, FR-139,317. *Bioorg Med Chem Lett* 5:917–922.

78. Combs AP, Kapoor TM, Feng S, Chen JK, Daude–Snow LF, Schreiber SL (1996): Protein structure-based combinatorial chemistry: discovery of non-peptide binding elements to Src SH3 domain. *J Am Chem Soc* 118:287–288.

79. Gennari C, Nestler HP, Salmon B, Still WC (1995): Solid-phase synthesis of vinylogous sulphonyl peptides. *Angew Chem Int Ed* 34:1763–1765.

80. Gennari C, Nestler HP, Salmon B, Still WC (1995): Synthetic receptors based on vinylogous sulfonyl peptides. *Angew Chem Int Ed* 34:1765–1768.

81. Burger MT, Still WC (1995): Synthetic ionophores: encoded combinatorial libraries of cyclen-based receptors for Cu^{2+} and Co^{2+}. *J Org Chem* 60:7382–7383.

82. Boyce R, Li G, Nestler HP, Suenaga T, Still WC (1994): Peptidosteroidal receptors for opioid peptides: sequence-selective binding using a synthetic receptor library. *J Am Chem Soc* 116:7955–7956.

83. Hiroshige M, Hauske JR, Zhou P (1995): Palladium-mediated macrocyclization on solid support and its applications to combinatorial synthesis. *J Am Chem Soc* 117:11590–11591.

84. Wang GT, Li S, Wideburg N, Krafft GA, Kempf DJ (1996): Synthetic chemical

diversity: solid phase synthesis of libraries of C(2) symmetric inhibitors of HIV protease containing diamino diol and diamino alcohol cores. *J Med Chem* 38:2995–3002.

85. Wang Z, Laursen RA (1992): Multiple peptide synthesis on polypropylene membranes for rapid screening of bioactive peptides. *Peptide Res* 5:275–280.

86. Frank R, Guler S, Krause S, Lindenmaier W (1990): Facile and rapid "spot-synthesis" of large numbers of peptides on membrane sheets. In Giralt E, Andreu D, eds. *Peptides: Proc 21st Eur Pept Symp.* Leiden: ESCOM Sci, pp 151–152.

87. Holmes CP, Adams CL, Kochersperger LM, Mortensen RB, Aldwin LA (1995): The use of light-directed combinatorial peptide synthesis in epitope mapping. *Biopolymers (Pep Sci)* 37:199–211.

88. Kerr JM, Banville SC, Zuckermann RN (1993): Encoded combinatorial peptide libraries containing nonnatural amino acids. *J Am Chem Soc* 115:2529–2531.

89. Bergmann C, Stohlmann SA, McMillan M (1993): An endogenously synthesized decamer peptide efficiently primes cytotoxic T cells specific for the HIV-1 envelope glycoprotein. *Eur J Immunol* 23:2777–2781.

90. Kramer A, Vakalopoulou E, Schleuning W-D, Schneider–Mergener J (1995): A general route to fingerprint analysis of peptide-antibody interactions using a clustered amino acid peptide library: comparison with a phage display library. *Mol Immunol* 32:459–465.

91. Cass R, Dreyer ML, Giebel LB, Hudson D, Johnson CR, Ross MJ, Schaeck J, Shoemaker KR (1994): Pilot, a new peptide lead optimization technique and its application as a general library method. In Hodges RE, Smith JA, eds. *Pep: Chem, Struct, Biol,; Proc Am Pept Symp 13.* Leiden: ESCOM Sci, pp 975–977.

92. Chu Y-H, Avila LZ, Biebuyck HA, Whitesides GM (1993): Using affinity capillary electrophoresis to identify the peptide in a peptide library that binds most tightly to vancomycin. *J Org Chem* 58:649–652.

93. Till JH, Annan RS, Carr SA, Miller WT (1994): Use of synthetic peptide libraries and phosphopeptide-selective mass spectrometry to probe protein kinase substrate specificity. *J Biol Chem* 269:7423–7428.

94. Kiely JS, Moos, WH, Pavia MR, Schwarz RD, Woodard GL (1991): A silica gel plate-based qualitative assay for acetylcholinesterase activity: a mass method to screen for potential inhibitors. *Anal Biochem* 196:438–442.

95. Bastos M, Marji NJ, Abeles RH (1995): Inhibitors of human heart chymase based on a peptide library. *Proc Natl Acad Sci USA* 92:6738–6742.

96. Wang J-X, Bray AM, Dipasquale AJ, Maeji NJ, Geysen HM (1993): Systemic study of substance P analogs, I: evaluation of peptides synthesized by the multipin method for quantitative receptor binding assay. *Int J Pept Protein Res* 42:384–391.

97. Wang J-X, DiPasquale AJ, Bray AM, Maeji NJ, Geysen HM (1993): Study of stereo-requirements of substance P binding to NK1 receptors using analogues with systematic D-amino acid replacements. *Bioorg Med Chem Lett* 3:451–456.

98. Boden P, Eden JM, Hodgson J, Horwell DC, Howson W, Hughes J, McKnight AT, Meecham K, Pritchard MC, Raphy J, Ratcliffe GS, Suman–Chauhan N, Woodruff GN (1994): The rational development of small molecule tachykinin NK3 receptor selective antagonists: the utilisation of a dipeptide chemical library in drug design. *Bioorg Med Chem Lett* 4:1679–1684.

99. Kassarjian A, Schellenberger V, Turck CW (1993): Screening of synthetic peptide libraries with radiolabeled acceptor molecules. *Pept Res* 6:129–133.

100. Needels MC, Jones DG, Tate EH, Heinkel GL, Kochersperger LM, Dower WJ, Barrett RW, Gallop MA (1993): Generation and screening of an oligonucleotide-encoded synthetic peptide library. *Proc Natl Acad Sci USA* 90:10700–10704.

101. Ohlmeyer MHJ, Swanson RN, Dillard LW, Reader JC, Asouline G, Kobayashi R, Wigler M, Still WC (1993): Complex synthetic chemical libraries indexed with molecular tags. *Proc Natl Acad Sci USA* 90:10922–10926.

102. Youngquist RS, Fuentes GR, Lacey MP, Keough T (1994): Matrix-assisted laser desorption ionization for rapid determination of the sequences of biologically active peptides isolated from support-bound combinatorial peptide libraries. *Rapid Commun Mass Spectrosc* 8:77–81.

103. Youngquist RS, Fuentes GR, Lacey MP, Keough T (1995): Generation and screening of combinatorial peptide libraries designed for rapid sequencing by mass spectroscopy. *J Am Chem Soc* 117:3900–3906.

104. Steward MW, Stanley CM, Obeid OE (1995): A mimotope from a solid-phase peptide library induces a measles virus-neutralizing and protective antibody response. *J Virol* 69:7668–7673.

105. Lam KS, Lou Q, Zhao Z-G, Smith J, Chen M-L, Pleshko E, Salmon SE (1995): Idiotype specific peptides bind to the surface immunoglobulins of two murine B-cell lymphoma lines, inducing signal transduction. In *Biomedical Pept, Protein & Nucleic Acids, I.* Birmingham, UK: Mayflower Worldwide, pp 205–210.

106. Lam KS, Lake D, Salmon SE, Smith J, Chen M-L, Wade S, Abdul-Latif F, Leblova Z, Ferguson RD, Krchnak V, Sepetov NF, Lebl M (1996): Application of a one-bead–one-peptide combinatorial library method for B-cell epitope mapping. *Immunomethods,* in press.

107. Smith MH, Lam KS, Hersh EM, Lebl M, Grimes WJ (1994): Peptide sequences binding to MHC class I proteins. *Mol Immunol* 31:1431–1437.

108. Lam KS, Lebl M (1992): Streptavidin and avidin recognize peptide ligands with different motifs. *Immunomethods* 1:11–15.

109. Gissel B, Jensen MR, Gregorius K, Elsner HI, Svendsen I, Mouritsen S (1995): L- and D-peptide ligands to streptavidin and avidin found in a synthetic peptide library. In Maia HLS, ed. *Pept 94; Proc 23 Eur Pept Soc* Leiden: ESCOM Sci, pp 1–6.

110. Fassina G, Bellitti MR, Cassani G (1994): Screening synthetic peptide libraries with targets bound to magnetic beads. *Protein Pept Lett* 1:15–18.

111. Chen JK, Lane WS, Brauer AW, Tanaka A, Schreiber SL (1993): Biased combinatorial libraries: novel ligands for the SH3 domain of phosphatidylinositol 3-kinase. *J Am Chem Soc* 115:12591–12592.

112. Yu H, Chen JK, Feng S, Dalgarno DC, Brauer AW, Schreiber SL (1994): Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* 76:933–945.

113. Wu J, Ma QN, Lam KS (1994): Identifying substrate motifs of protein kinases by a random library approach. *Biochemistry* 33:14825–14833.

114. Lam KS, Wu J, Lou Q (1995): Identification and characterization of a novel

synthetic peptide substrate specific for Src-family protein tyrosine kinases. *Int J Pept Protein Res* 45:587–592.

115a. Lou Q, Wu J, Salmon SE, Lam KS (1995): Structure–activity relationship of a novel peptide substrate for p60^{c-src} protein tyrosine kinase. *Lett Pept Sci* 2:289–296.

115b. Lou Q, Leftwich ME, Lam KS (1996): Identification of GIYWHHY as a novel peptide substrate for human p60^{c-src} protein tyrosine kinase. *Bioorg Med Chem, Bioorg Med Chem* 4:677–682.

116. Petithory JR, Masiarz FR, Kirsch JF, Santi DV, Malcolm BA (1991): A rapid method for determination of endoproteinase substrate specificity: specificity of the 3C proteinase from hepatitis A virus. *Proc Natl Acad Sci USA* 88:11510–11514.

117. Ator MA, Beigel S, Dankanich TC, Echols M, Gainor JA, Gilliam CL, Gordon TD, Koch D, Koch JF, Kruse LI, Morgan BA, Krupinski–Olsen R, Siahaan TJ, Singh J, Whipple DA (1994): Immobilized peptide arrays: a new technology for the characterization of protease function In Hodges RE, Smith JA, eds. *Pep: Chem, Struct, Biol, Proc Am Pept Symp 13*. Leiden: ESCOM Sci, pp 1012–1015.

118. Singh J, Allen M, Solowiej, Ator M, Olsen R, Whipple D, Gainor J, Gilliam C, Treasurywala A, Echols M, Petersen J, Kruse L, Upson D (1995): Screening of "unbiased" immobilized peptide library versus proteases: an approach for rapid determination of substrate specificity and selectivity. In Maia HLS, ed. *Pept 94: Proc 23rd Eur Pept Symp*. Leiden: ESCOM Sci.

119. Singh J, Allen M, Ator M, Gainor J, Whipple D, Solowiej JE, Treasurywala A, Morgan BA, Gordon TD, Upson D (1995): Validation of screening immobilized peptide libraries for discovery of protease substrates. *J Med Chem* 38:217–219.

120. Meldal M, Svendsen I, Breddam K, Auzanneau F-I (1994): Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. *Proc Natl Acad Sci USA* 91:3314–3318.

121. Meldal M, Svendsen I (1995): Direct visualization of enzyme inhibitors using a portion mixing inhibitor library containing a quenched fluorogenic peptide substrate, I: inhibitors for subtilisin Carlsberg. *J Chem Soc Perkin Trans* 1:1591–1596.

122. Moran EJ, Sarshar S, Cargill JF, Shahbaz MM, Lio A, Mjalli AMM, Armstrong RW (1995): Radio frequency tag encoded combinatorial library method for the discovery of tripeptide-substituted cinnamic acid inhibitors of the protein tyrosine phosphatase PTP1B. *J Am Chem Soc* 117:10787–10788.

123. Samson I, Kerremans L, Rozenski J, Samyn B, Van Beeumen J, Van Aerschot A, Herdewijn P (1995): Identification of a peptide inhibitor against glycosomal phosphoglycerate kinase of *Trypanosoma brucei* by a synthetic peptide library approach. *Bioorg Med Chem* 3:257–265.

124. Lowe G, Quarrell R (1994): Screening peptide bead libraries for enzyme inhibitors. *Methods: A Comparison to Methods in Enzymology* 6:411–416.

125. Seligmann, B, in press.

126. Seligmann, B, in press.

127. Domingo GJ, Leatherbarrow RJ, Freeman N, Patel S, Weir M (1995): Synthesis of a mixture of cyclic peptides based on the Bowman–Birk reactive site loop to screen for serine protease inhibitors. *Int J Pept Res* 46:79–87.

128. Vlattas I, Sytwn II, Dellureficio J, Stanton J, Braunwalder AF, Galakatos N, Kramer R, Seligmann B, Sills MA, Wasvary J (1994): Identification of a receptor-

binding region in the core segment of the human anaphylatoxin C_{5a}. *J Med Chem* 37:2783–2790.

129. Jayawickreme CK, Quillan JM, Graminski GF, Lerner MR (1994): Discovery and structure–function analysis of *a*-melanocyte-stimulating hormone antagonists. *J Biol Chem* 47:29846–29854.

130. Quillan JM, Jayawickreme CK, Lerner MR (1995): Combinatorial diffusion assay used to identify topically active melanocyte-stimulating hormone receptor antagonists. *Proc Natl Acad Sci USA* 92:2894–2898.

131. Seligmann, B. in press.

132. Lam KS, Zhao Z-C, Wade S, Krchnak V, Lebl M (1994): Identification of small peptides that interact specifically with a small organic dye. *Drug Dev Res* 33:157–160.

133. Pinilla C, Appel JR, Houghten RA (1993): Synthetic peptide combinatorial libraries (SPLCs): identification of the antigenic determinant of β-endorphin recognized by monoclonal antibody 3E7. *Gene* 128:71–76.

134. Zuckermann RN, Kerr JM, Siani MA, Banville SC, Santi DV (1992): Identification of highest-affinity ligands by affinity selection from equimolar peptide mixtures generated by robotic synthesis. *Proc Natl Acad Sci USA* 89: 4505–4509.

135. Bianchi E, Folgori A, Wallace A, Nicotra M, Acali S, Phalipon A, Barbato G, Bazzo R, Cortese R, Felici F, Pessi A (1995): A conformationally homogenous combinatorial peptide library. *J Mol Biol* 247:154–160.

136. Sila U, Mutter M (1995): Topical templates as a tool in molecular recognition and peptide mimicry: synthesis of a TASK library. *J Mol Recog* 8:29–34.

137. Owens RA, Gesellchen PD, Houchins BJ, DiMarchi RD (1991): The rapid identification of HIV protease inhibitors through the synthesis and screening of defined peptide mixtures. *Biomed Biophys Res Commun* 181:402–408.

138. Eichler J, Houghten RA (1993): Identification of substrate-analog trypsin inhibitors through the screening of synthetic peptide combinatorial libraries. *Biochemistry* 32: 11035–11041.

139. Spellmeyer DC, Brown S, Stauber GB (1993): Endothelin receptor ligands. multiple D-amino acid replacement net approach. *Bioorg Med Chem Lett* 3:1253–1256.

140. Neustadt B, Wu A, Smith EM, Nechuta T, Fawzi A, Zhang H, Ganguly AK (1995): A case study of combinatorial libraries: endothelin receptor antagonist hexapeptides. *Bioorg Med Chem Lett* 5:2041–2044.

141. Dooley CT, Chung NN, Schiller PW, Houghten RA (1993): Acetalins: opioid receptor antagonists determined through the use of synthetic peptide combinatorial libraries. *Proc Natl Acad Sci USA* 90:10811–10815.

142. Dooley CT, Chung NN, Wilkes BC, Schiller PW, Bidlack JM, Pasternak GW, Houghten RA (1994): An all D-amino acid opioid peptide with central analgesic activity from a combinatorial library. *Science* 266:2019–2022.

143. Dooley CT, Kaplan RA, Chung NN, Schiller PW, Bidlack JM, Houghten RA (1995): Six highly active mu-selective opioid peptides identified from two synthetic combinatorial libraries. *Pept Res* 8:124–137.

144. Houghten RA, Appel JR, Blondelle SE, Cuervo JH, Dooley CT, Pinilla C (1992): The use of synthetic peptide combinatorial libraries for the identification of bioactive peptides. *Biotechniques* 13:412–421.

145. Blondelle SE, Takahashi E, Weber PA, Houghten RA (1994): Identification of antimicrobial peptides by using combinatorial libraries made up of unnatural amino acids. *Antimic Agents Chemother* 38:2280–2286.

146. Pinilla C, Appel JR, Milich D, Houghten RA (1994): Identification of an antigenic determinant for the surface antigen of hepatitis B virus using peptide libraries. In Hodges RE, Smith JA, eds. *Pep: Chem, Struct, Biol, Proc Am Pept Symp 13.* Leiden: ESCOM Sci, pp 1016–1017.

147. Pinilla C, Appel JR, Houghten RA (1994): Investigation of antigen–antibody interactions using a soluble, non-support-bound synthetic decapeptide library composed of four trillion (4×10^{12}) sequences. *Biochem J* 301:847–853.

148. Birkett AJ, Soler DF, Wolz RL, Bond JS, Wiseman J, Berman J, Harris RB (1991): Determination of enzyme specificity in a complex mixture of peptide substrates by N-terminal sequence analysis. *Anal Biochem* 196:137–143.

149. Wallace A, Altamura S, Toniatti C, Vitelli A, Bianchi E, Delmastro P, Ciliberto G, Pessi A (1994): A multimeric synthetic peptide combinatorial library. *Peptide Res* 7:27–31.

150. Walden P, Wiesmuller KH, Jung G (1995): Elucidation of T-cell epitopes: a synthetic approach with random peptide libraries. *Biochem Soc Trans* 23:678–681.

151. Udaka K, Wiesmuller K-H, Kienle S, Jung G, Walden P (1995): Decrypting the structure of major histocompatability complex class I-restricted cytotoxic T lymphocyte epitopes with complex peptide libraries. *J Exp Med* 181:2097–2108.

152. Blondelle SE, Simpkins LR, Houghten RA (1992): Inhibition of melittin's hemolytic activity using synthetic peptide combinatorial libraries. In Schneider CH, Eberle AN, eds. *Peptides: Proc 22nd Eur Pept Symp 1.* Leiden: ESCOM Sci, pp 761–762.

153. Stigler R-D, Ruker F, Katinger D, Elliott G, Hohne W, Henklein P, Ho, JX, Keeling K, Carter DC, Nugel E, 'Kramer A, Psrstmann T, Schneider–Mergener J (1995): Interaction between a Fab fragment against gp41 of human immunodeficiency virus 1 and its peptide epitope: characterization using a peptide epitope library and molecular modeling., *Protein Eng* 8:471–479.

154. Maeji NJ, Bray AM, Geysen HM (1990): Multipin peptide synthesis strategy for T cell determinant analysis. *J Immunol Methods* 134:23–33.

155. Kozlowski S, Corr M, Takeshita T, Boyd LF, Pendleton CD, Germain RN, Berzofsky JA, Margulies DH. (1992): Serum angiotensin-1 converting enzyme activity processes a human immunodeficiency virus 1 gp160 peptide for presentation for presentation by Major Histocompatability Complex Class I molecules. *J Exp Med* 175:1417–1422.

156. Shirai M, Pendleton D, Berzofsky JA (1992): Broad recognition of cytotoxic T cell epitopes form the HIV-1 envelope protein with multiple class I histocompatability molecules. *J Immunol* 148:1657–1667.

157. Wilson IA, Niman HL, Houghten RA, Cherenson ML (1984): The structure of an antigenic determinant in a protein. *Cell* 37:767–778.

158. Appel JR, Pinilla C, Niman H, Houghten RA (1990): Elucidation of discontinuous linear determinants in peptides. *J Immunol* 144:976–983.

159. Wong WY, Sheth HB, Holm A, Irvin RT, Hodges RS (1994): Investigation of antibody binding diversity utilizing representative combinatorial peptide libraries

In Hodges RE, Smith JA, eds. *Pep: Chem, Struct, Biol, Proc Am Pept Symp 13.* Leiden: ESCOM Sci, pp 995–997.

160. Wong WY, Sheth HB, Holm A, Hodges RS, Irvin RT (1994): Representative combinatorial peptide libraries: an approach to reduce both synthesis and screening efforts. *Methods: A Companion to Methods in Enzymology* 6:404–410.

161. Kemp BE, Graves DJ, Benjamini E, Krebs EG (1977): Role of multiple basic residues in determining the substrate specificity of cyclic AMP-dependent protein kinase. *J Biol Chem* 252:4888–4894.

162. Songyang Z, Carraway KL, Eck MJ, Harrison SC, Feldman RA, Mohammadi M, Schlessinger J, Hubbard SR, Smith DP, Eng C, Lorenso MJ, Ponder BAJ, Mayer BJ, Cantley LC (1995): Catalytic specificity of protein–tyrosine kinase is critical for selective signalling. *Nature* 373:591–594.

163. Gainor JA, Gordon TD, Morgan BA (1994): The synthesis and coupling efficiency of 7-hydroxycoumarin-4-propionic acid, a fluorescent marker useful in immobilized substrate libraries. In Hodges RE, Smith JA, eds. *Peptides: Chemistry Structure and Biology: Proceedings of the 13th American Peptide Symposium.* Leiden: ESCOM Sci, pp 989–991.

164. Al-Obeidi F, Lebl M, Ostrem JA, Safar P, Stierandova A, Strop P, Walser A (1995): Factor Xa Inhibitors. *WO PCT 95/29189.*

165. Maeder DL, Sunde M, Botes DP (1992): Design and inhibitory properties of synthetic Bowman–Birk loops. *Int J Protein Res* 40:97–102.

166. Horwell DC, Howson W, Ratcliffe GS, Rees DC (1993): The design of a dipeptide library for screening at peptide receptor sites. *Bioorg Med Chem Lett* 3:799–802.

167. Horwell DC, Hughes J, Hunter JC, Pritchard MC, Richardson RS, Roberts E, Woodruff GN (1991): Rationally designed "dipeptide" analogs of CCK: a-methyl-tryptophan derivatives as highly selective and orally active gastrin and CCK-B antagonists with potent anxiolytic properties. *J Med Chem* 34:404–414.

168. Drapeau G, Brochu S, Godin D, Levesque L, Rioux R, Marceau F (1993): Synthetic Cra receptor agonists: pharmacology, metabolism and in vivo cardiovascular and hematologic effects. *Biochem Pharm* 45:1289–1299.

169. Kawai M, Wiedeman PE, Luly JR, Or YS (1992): *Publication WO92/12168-A1.*

170. Luly JR, Megumi K, Wiedeman PE (1994): *US Patent 005 190922A.*