Medications Development: Drug Discovery, Databases, and Computer-Aided Drug Design

Editors:

Rao S. Rapaka, Ph.D.
Richard L. Hawks, Ph.D.

NIDA Research Monograph 134
1993

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
National Institutes of Health

National Institute on Drug Abuse
5600 Fishers Lane
Rockville, MD 20857

Preparation of Large Peptide Libraries With One Peptide per Bead and Their Use for the Discovery of Peptides That Bind to Acceptors

Victor J. Hruby, Kit S. Lam, Michal Lebl, Wieslaw Kazmierski, Evan M. Hersh, and Sydney E. Salmon

INTRODUCTION

Most biological processes are controlled and modulated by intermolecular interactions. These interactions can occur with a change in the covalent structure of one or more of the participants (e.g., an enzyme substrate interaction) or with no change in covalent structure (e.g., an antibody-antigen interaction). In the latter case, in general, the three-dimensional structures of both the ligand and the acceptor molecule(s) change as a consequence of the interaction, which in turn can lead to a change in the physical and chemical properties of the acceptor molecule(s) in a way that modulates cellular structure and function. Thus, these interactions are critical requirements for the conformational, dynamic, and stereoelectronic changes that are central to the "chemistry of life" in biological systems. In this regard, peptides and proteins and their posttranslationally modified derivatives constitute the most widely used "molecules of life" and can serve many functional roles (e.g., as antibodies, enzymes, growth factors, immunomodulators, hormones, neurotransmitters, and receptors) and the many other functions that are critical to cellular function and to intercellular communication in complex living systems. It is no accident that nature uses proteins as the major functional molecules in living systems, which undoubtedly is related to the enormous physical and chemical diversity that is possible for these compounds. For example, if one utilizes only the 20 standard amino acid residues found in proteins and examines the number of distinct chemical species available, millions of molecules already are possible at the pentapeptide stage (3.2×10^6); that number becomes 1.28×10^9 at the heptapeptide stage. Thus, for even a small protein of 100 amino acid residues, there are an astronomical number of possibilities. It is not clear how many of these possibilities nature has explored, but it is clear that only a small fraction of those possible are used.

There is a great opportunity in the possibility for utilizing enormous chemic diversity to obtain a better understanding of how nature works and for the discovery of novel lead molecules that can serve as starting points for the development of agonist or antagonist drugs with specific biological effects. It also is a challenge to develop a simple method to prepare such large an diverse chemical libraries in a cost-effective manner that will allow examini of the binding properties of a particular species in the library and at the sa time have each chemical species present in sufficient quantities for structu determination.

The authors recently proposed and developed a methodology for preparin large, statistically diverse peptide libraries (10^4 to 10^8 individual species) in a short time and for examining their ligand-binding activities with accept macromolecules such as antibodies and enzymes (Lam et al. 1991, 1992) This chapter briefly reviews the basic approach in this development and discusses some of the possible applications for the future.

RESULTS AND DISCUSSION

The creation of a large peptide library creates several problems. The most important is that of scale. The basic question posed was how much of each peptide is needed so that, on the one hand, sufficient concentratic are available to screen for binding and, on the other hand, sufficient amou are available for structure determination. After examining several approac the authors devised a method that maximizes chances for success with th minimum amount of peptide. This method exploits the idea of one peptide for one bead. The synthetic methodology used is the Merrifield method (Merrifield 1963) in which the peptide of interest is synthesized on a solid support. This method has revolutionized peptide chemistry (and synthetic chemistry in general) in that it greatly simplifies the multistep synthetic processes of preparing a peptide. The growing peptide chain is assemble on a resin bead, and as a consequence, the solvents and reagents requir for the synthesis can be removed by the simple expedient of washing. Th methodology provides two other previously unexploited advantages. First depending on the substitution level and size of the bead (in standard Merrifield syntheses, the substitution level is generally 0.2 to 1.0 mmol/g resin, and the bead size is generally from 50 to 200 microns), a high local concentration of peptide can be present on each bead. Calculations sugg that the local concentration of peptide on a single bead is in the millimolar to micromolar range and that the amount of peptide present would be fron 50 to 300 pmol, more than sufficient to determine the sequence, because modern sequencers can sequence in the 2 to 5 pmol range. Second, a bead provides a unique surface (like a cell) for interaction of the ligand present on the surface of the bead with a macromolecule in the surroundi

medium. Such a surface should be excellent for large-scale screening in a minimum volume but with a large surface area.

The next question posed was, given the above considerations, how to prepare a large, chemically diverse peptide mixture in which each peptide in the mixture would be localized to a single resin bead and each bead would possess a single peptide and in which each peptide in the library would be present in approximately the same concentration. The answer to the latter part of the question is to have each bead in the mixture approximately the same size and substituted to the same extent. To ensure that each bead has a unique peptide and only that peptide, it is necessary that each bead be exposed to only a single coupling reaction at a time and that the reaction be driven to completion. The approach is similar to that presented by Furka and colleagues (1988, 1991), who, however, did not recognize the potential of having a unique peptide sequence on every bead. Finally, for diversity, one must turn to statistics. For example, if the 20 standard amino acids are used for the construction of a library, by dividing the resin into 20 equal parts, adding each amino acid to only one of these parts, then thoroughly mixing all the parts, by dividing into 20 equal parts again, adding each amino acid to only one of these parts, and so forth, one can rapidly develop a library of millions of peptides on millions of beads. Because a gram of standard-size solid-phase resin contains about 1 million 100-micron beads, very large libraries are feasible using Merrifield solid-phase methodology.

It is important to recognize two important synthetic issues in the construction of such a large, diverse peptide library. First, different amino acid residues have very different chemical reactivities. For example, activated N^{α}-Fmoc-alanine is much more reactive than N^{α}-Fmoc-isoleucine. Second, and even more important, when a large peptide library is constructed, the peptide products become very diverse chemically and hence differ greatly in their reactivity (nucleophilicity). Therefore, every reaction must be driven to completion. As a practical matter, we found that using a threefold to fourfold excess of the amino acid, a 1- to 2-hour reaction time in a minimum volume, and optimized Merrifield synthetic procedures generally leads to completion of the reaction even for difficult sequences such as Val to Val-peptide coupling. To check on the fidelity of the methodology for producing single peptide moieties on a bead, we examined a large number of single resin beads chosen both randomly and from our binding assay for their structure and for the synthetic fidelity of the synthetic peptide on the bead using a Pulsed Liquid Automatic Peptide Sequencer to determine the sequence and quantitate the amount of peptide on a single bead and then used modern preview analysis to determine the fidelity of the peptide on a single bead. These studies showed that each bead contains between 50 and 300 pmols of peptide as predicted, and preview analysis demonstrated that the peptides on each bead were homogeneous, with the percentage of a single peptide species being between

97 to 99.9 percent in a pentapeptide library of more than 2 million peptides (Lam et al., in press). The results of such an analysis for several pentapepti beads that interact with streptavidin (Lam et al. 1991, 1992) are given in tabl These results demonstrate that the synthetic methods used, optimized to maximize completion of reactions on solid phase resins, can provide large peptide libraries in which each peptide produced is present in enough yield and each bead present has on its surface and throughout the bead a unique peptide sequence. The realization that a single solid-phase bead could cont enough peptide in sufficient purity for use in a screening process was a critic element in the development of this methodology.

Next, these libraries were utilized to screen for binding to acceptor macromolecules. For this purpose, the authors prepared the libraries to leave the peptide attached to the bead, but with all the protecting groups (amino terminal and side-chain moieties) removed. Thus, it was necessary to develop a synthetic strategy so that all side-chain protecting groups could be removed following construction of the peptide library while the peptide remained on the resin and, in addition, to ensure that the peptide on the surface of the bead was accessible for binding to the macromolecular acceptor by construction of a spacer of suitable length that is "biocompatible. One approach is to use polyamide resins that have been modified with a spacer arm that consisted of a β-alanine, an ε-amino caproic acid moiety, and an ethylenediamine moiety attached to the polyamide polymer. In this strategy, Fmoc protection is used for N^{α}-amino groups, and tert-

TABLE 1. *Examination of the amount of peptide and preview analysis for individual pentapeptide beads that interact with streptavidin*

Sequence Determined	Peptide Recovered (pmol)	Percent Preview for Cycle 5
HPQGP	60	0.35
HPQAG	53	1.5
FHPQG	72	0.23
QHPQG	60	2.3
GHPQG	250	0.44
REHPQ	112	0.56
IQHPQ	192	1.8
GNHPQ	222	0
WMHPQ	257	2.7
TPHPQ	158	0
WNHPM	59	2.5
MHPMA	140	0.31

butyloxycarbonyl, tert-butyl esters, and tert-butyl ether type side-chain protections are used (N^{α}-Boc protection in conjunction with Fmoc-based side-chain protecting groups also can be used). The N^{α}-Fmoc groups were removed with piperidine, and following construction of the peptide library, the side-chain protecting groups were removed by 90 percent trifluoroacetic acid in N,N-dimethylformamide (v/v) containing 1 percent anisole and 0.9 percent ethanedithiol. The peptide-resin beads were then neutralized with 10 percent diisopropylethylamine in N,N-dimethylformamide and thoroughly washed with several solvents (including water) before use in screening. Utilizing this methodology, we obtained a number of peptide libraries suitable for rapid screening of the type described below.

To screen a library composed of millions of different chemical species requires careful consideration. Among the most common problems for consideration are the signal-to-noise (background signal) ratio, nonspecific chemical reactions, unexpected chemical reactions, and the usual problems of sensitivity and specificity. To overcome these problems, the authors developed a rapid screening method that utilizes an enzyme-linked immunoassay method in which the acceptor molecule (e.g., a monoclonal antibody) is coupled to the enzyme alkaline phosphatase, which can then catalyze a reaction in which those beads that interact with the monoclonal antibody become colored as a result of the enzyme-catalyzed reaction. These beads can be easily distinguished from those that do not react because they are colorless. The reacting beads (those that are colored) can be removed easily from the incubation medium by a micromanipulator under a low-power dissecting microscope as previously discussed (Lam et al. 1991, 1992).

The colored beads are then washed, the bound monoclonal antibody-enzyme complex removed by treatment with denaturing reagents that can solubilize proteins (e.g., 6M guanidine hydrochloride), and the single bead subjected to sequence analysis. In this way, unique new structures can be discovered that bind with high affinity (micromolar to nanomolar) to acceptor molecules. Ordinarily, we separately synthesize the binding peptides discovered by these procedures and examine their binding to the acceptor molecules by standard radioreceptor binding assays. Generally, there is a good correspondence between the binding seen on the bead and that obtained in the solution binding experiments (Lam et al. 1991, 1992, in press).

SUMMARY AND OVERVIEW

The authors developed a simple, highly reproducible, and synthetically clean method for preparing large, statistically diverse peptide libraries that can be used to discover ligands that will bind to acceptor molecules. Several other

methods were developed using both biological and chemical methods to provide polypeptide diversity for a variety of applications (Scott and Smith 1990; Geysen et al. 1984; Fodor et al. 1991; Cwirla et al. 1990; Houghten et al. 1991; Oldenburg et al. 1992; Devlin et al. 1990; Scott et al. 1992; Cull et al. 1992; Brenner and Lerner 1992). These methods, and others to be discovered and developed, offer a number of opportunities to examir in new ways the fundamental issues of molecular recognition and the diversity of ways in which molecular recognition problems can be solved. With the availability of large, diverse peptide libraries, it is likely that, for the same acceptor site on a macromolecule, several different peptide structures will bind specifically to the same site with an affinity within a few orders of magnitude of each other.

From these and other fundamental studies, it will be possible to address issues that could not be addressed before the availability of these libraries For example, there are the basic questions of whether a small peptide of 5 to 15 residues can mimic a protein in its activities, whether a small linear peptide can serve as a mimotope for a discontinuous epitope on an antibo or how big an area an inhibitor needs to recognize to be able to block the binding of the native ligand. The last question is an especially critical ques for protein-protein interactions, where fairly large surface areas appear to l involved in molecular recognition. Can a small peptide serve as a competi inhibitor for one of the sites of interaction to block the entire interaction? The answer is not clear.

Another major issue is the difference between a peptide agonist and a pep antagonist. It was suggested many years ago that peptide agonists and antagonists, even competitive ones, bind differently to the acceptor molecu (Meraldi et al. 1977). In this case, a peptide antagonist had greatly reduce flexibility relative to the agonist, and the specific induced rigidity was critica for antagonist bioactivity (Meraldi et al. 1977). Perhaps this effect of rigidit explains why small, relatively rigid peptidomimetics generally are antagoni of ligands for macromolecular acceptors.

Another issue is how many different ways a particular molecular recognitio problem can be solved. For example, it already has been noted that a particular recognition favors a particular primary sequence, called a consel sequence, and that this sequence is conserved throughout evolution. The question is whether this sequence is truly "essential" or whether some othe sequence can serve the same purpose. This question now can be explore by constructing a large library of peptides (millions or more) that do not possess this consensus to determine whether one or more new sequences can accomplish the same binding interactions and whether a new consens sequence" will be developed. Another extension of these methods is mole

recognition processes involving nonpeptides, such as sugars, polysaccharides, nucleotides, and lipids. Again, some very critical issues can be addressed.

Many other possibilities exist for exploiting molecular diversity, and acceleration can be expected in the development of assay methods, chemical methods, and physical methods to examine complex diverse mixtures. In this regard, it is likely that the development of new synthetic methods to prepare large, diverse chemical structures will challenge the current methods of structural elucidation.

REFERENCES

Brenner, S., and Lerner, R.A. Encoded combinatorial chemistry. *Proc Natl Acad Sci U S A* 89:5381-5383, 1992.

Cull, M.G.; Miller, J.F.; and Schatz, P.J. Screening for receptor ligands using large libraries of peptides linked to the C terminus of the lac repressor. *Proc Natl Acad Sci U S A* 89:1865-1869, 1992.

Cwirla, S.E.; Peter, E.A.; Barrett, R.W.; and Dower, W.J. Peptides on phage: A vast library of peptides for identifying ligands. *Proc Natl Acad Sci U S A* 87:6378-6382, 1990.

Devlin, J.J.; Pangamban, L.C.; and Devlin, P.E. Random peptide libraries: A source of specific protein binding molecules. *Science* 249:404-406, 1990.

Fodor, S.P.A.; Reed, J.L.; Pirrung, M.C.; Stryer, L.; Lee, A.T.; and Solas, D. Light-directed, spacially addressable parallel chemical synthesis. *Science* 251:767-773, 1991.

Furka, A.; Sebestyén, F.; Asgedom, M.; and Dibó, G. Cornucopia of peptides by synthesis. Abstracts, 14th International Congress of Biochemistry, Vol. 5. Prague, 1988. p. 47.

Furka, A.; Sebestyén, F.; Asgedom, M.; and Dibó, G. General method for rapid synthesis of multicomponent peptide mixtures. *Int J Pept Protein Res* 37:487-493, 1991.

Geysen, H.M.; Meloen, R.H.; and Barteling, S.J. Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid. *Proc Natl Acad Sci U S A* 81:3998-4002, 1984.

Houghten, R.A.; Pinilla, C.; Blondelle, S.E.; Appel, J.R.; Dooley, C.T.; and Cuervo, J.H. Generation and use of synthetic combinatorial libraries for basic research and drug discovery. *Nature* 354:84-86, 1991.

Lam, K.S.; Hruby, V.J.; Lebl, M.; Kazmierski, W.M.; Hersh, E.M.; and Salmon, S.E. The chemical synthesis of large random peptide libraries and their use for the discovery of ligands for macromolecular acceptors. *Bioorg Med Chem Lett*, in press.

Lam, K.S.; Salmon, S.; Hersh, E.M.; Hruby, V.J.; Al-Obeidi, F.; Kazmierski, W.M.; and Knapp, R.J. The selectide process: Rapid generation of large synthetic peptide libraries linked to identification and structure determination of acceptor-binding ligands. In: Smith, J.E., and Rivier, J.E., eds. *Peptides: Chemistry and Biology*. Leiden, The Netherlands: ESCOM Publishers, 1992. pp. 492-495.

Lam, K.S.; Salmon, S.E.; Hersh, E.M.; Hruby, V.J.; Kazmierski, W.M.; and Knapp, R.J. A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 354:82-84, 1991.

Meraldi, J.P.; Hruby, V.J.; and Brewster, A.I.R. Relative conformational rigidity in oxytocin and [1-penicillamine]oxytocin: A proposal for the relationship of conformational flexibility to peptide hormone agonism and antagonism. *Proc Natl Acad Sci U S A* 74:1373-1377, 1977.

Merrifield, R.B. Solid phase peptide synthesis. I. Synthesis of a tetrapeptide. *J Am Chem Soc* 85:2149-2154, 1963.

Oldenburg, K.R.; Loganathan, D.; Goldstein, I.J.; Schultz, P.G.; and Gallop, M.A. Peptide ligands for a sugar-binding protein isolated from a random peptide library. *Proc Natl Acad Sci U S A* 89:5393-5397, 1992.

Scott, J.K.; Loganathan, D.; Easley, R.B.; Gong, X.; and Goldstein, I.J. A family of concanavalin A-binding peptides from a hexapeptide epitope library. *Proc Natl Acad Sci U S A* 89:5398-5402, 1992.

Scott, J.K., and Smith, G.P. Searching for peptide ligands with an epitope library. *Science* 249:386-390, 1990.

ACKNOWLEDGMENTS

Fahad Al-Obeidi and Ralph Hirschmann made helpful suggestions and offered constructive criticisms; Farid Abdul Latif provided assistance in sequencing; Ronald Ferguson assisted in the synthesis of peptide libraries and ligands; and Richard Knapp helped in the determination of binding affinities.

This work was supported in part by the Arizona Disease Control Commission and the Selectide Corporation. Kit S. Lam is a recipient of a Leukemia Specia Fellow Award (1989-1992) and a Leukemia Scholar Award (1992-1997).

AUTHORS

Victor J. Hruby, Ph.D.
Regents Professor
Department of Chemistry

Kit S. Lam, M.D., Ph.D.
Assistant Professor
Arizona Cancer Center

Wieslaw Kazmierski, Ph.D.
Research Scientist
Department of Chemistry

Evan M. Hersh, M.D.
Professor
Arizona Cancer Center

Sydney E. Salmon, M.D.
Regents Professor
Arizona Cancer Center

Department of Chemistry and Arizona Cancer Center
University of Arizona
Tucson, AZ 85721

Michal Lebl, Ph.D.
Director of Chemistry
Selectide Corporation
10900 Stallard Place
Suite 122
Tucson, AZ 85737