Bioorganic & Medicinal Chemistry Letters, Vol.3, No.3, pp. 387-396, 1993
Printed in Great Britain

0960-894X/93 $6.00 + .00

INTRODUCTION

THE GENERATION OF MOLECULAR DIVERSITY

Michael R. Pavia,*@ Tomi K. Sawyer,# Walter H. Moos+

@ Genesis Pharmaceuticals, Inc., 840 Memorial Drive, Cambridge, MA, 02139
# Parke-Davis Pharmaceutical Research Division, Warner-Lambert Co., 2800 Plymouth Rd., Ann Arbor, MI, 48105
+ Chiron Corp., 4560 Horton St., Emeryville, CA 94608

Ligands that bind to a given ligate (antibody, receptor, enzyme, transcription factor, or other host molecule) can be identified either by rational design or by screening large numbers of structurally diverse molecules. To define the chemical basis of ligand/ligate interactions it is often desirable to prepare and evaluate a large number of peptide analogues. The traditional serial process of synthesizing and testing native peptide analogues or protein cleavage fragments and subsequent evaluation of their biological activity is being replaced by a variety of highly sophisticated methods in which large libraries of peptide ligands are synthesized in parallel and screened against ligates of interest.[1] These techniques add tremendous power to the drug discovery process and herein we review the progress made in this arena.

Chemical Methods

Multi-Pin Method: Geysen[2] first showed that peptides could be synthesized in a reusable form in numbers three orders of magnitude greater than by conventional means. The peptides were synthesized on polyacrylic acid grafted polyethylene rods arranged in a microtiter plate format allowing ninety-six separate peptides to be simultaneously synthesized at the tips of the rods. The technique utilizes conventional multiple peptide synthesis methods, readily available materials and instrumentation, and could be used, for example, to synthesize more than 2000 hexapeptides in 10 days. The peptides can be used in assays bound to the rods, allowing their reuse for 30-60 tests, or cleaved and tested in solution.[3,4] The first cleavage method reported yielded products that contain a diketopiperazine moiety at the C-terminus which apparently does not affect biological activity in many cases. It is possible to have free acid or amide carboxy termini containing a C-terminal beta-alanine or beta-alanylmethyl amide by use of new linkers or handles for attachment to the solid support.[5] Alternatively, the support bound peptides may be cleaved with ammonia vapor to afford the corresponding C-terminal primary amides.[6]

Epitope Mapping Strategies: With the development of a method to rapidly prepare large numbers of peptides, strategies were developed to map peptide epitopes. These

epitopes can be classified as continuous, consisting of a linear sequence of amino acids homologous with the ligand or ligate of interest or discontinuous where the binding pharmacaphore consists of non-adjacent amino acids brought into proximity by secondary/tertiary structure folding. Identification of sequential epitopes within the Geysen approach is accomplished by preparing a complete set of all possible overlapping synthetic peptides of a given length homologous with the sequence of the peptide of interest, providing that the length of the peptides is equal to or greater than the longest sequential epitope. Current thinking suggests that sequential epitopes range from 5-8 amino acid residues. Scanning a protein of n residues therefore requires the synthesis and testing of n-4 (for 5mers) to n-7 (for 8mers) peptides. Once the set of linear fragments has been synthesized and tested it is then desirable to synthesize a replacement set where each of the amino acids is replaced sequentially with natural amino acids or unnatural amino acid analogues. The replacement analysis provides insight into the fine structure, or specificity of the epitope.

The majority of ligates, however, are expected to recognize discontinuous epitopes. An approach for identifying discontinuous epitopes is referred to as the mimotope approach.[7] In order to make an entire mimotope library of a peptide of n-residues incorporating only the natural L-amino acids would require 20^n peptides, an unmanageable number. As an alternative, Geysen proposed to first identify an optimum dipeptide unit with the rest of the sequence formed from randomly incorporated amino acids.[8, 9] Once the optimum dipeptide sequence was identified one could sequentially optimize the remaining amino acids. This combinatorial approach allows for the rapid synthesis and biological evaluation of millions of individual peptides for screening. The screening of these heterogeneous libraries, along with an iterative selection and synthesis procedure permits the systematic identification of optimal peptide ligands. Starting with a library of octapeptides Geysen was able to rapidly and precisely identify high affinity ligands for a monoclonal antibody shown to recognize a discontinuous epitope.

For example, to identify the optimal binding sequence of a pentapeptide requires preparation of 400 peptide mixtures of known composition at the first two amino acid positions while the remainder of the peptide is made up of a random mixture of amino acids. An assay, based on binding or some other activity, is then utilized to identify the library with the most active peptide. The initial dipeptide sequence is then used as a starting point to prepare another twenty mixtures defining the third amino acid, while positions 4 and 5 are a mixture of random amino acids. The biological screening is again carried out and the most potent ligand identified. This newly identified amino acid becomes the third optimized position in the pentapeptide. This procedure is repeated two more times until the entire sequence is optimized. For delineation of a pentapeptide sequence, this method requires the preparation of 460 peptide mixtures (400 + 20 + 20 + 20). For a hexapeptide library the number increases only slightly to 480 mixtures, and represents rapid screening of 64,000,000 hexapeptides (assuming the sequence is

limited to the 20 naturally-occurring amino acids) in a rapid fashion. If amino acids other than the naturally-occurring ones are used the number of peptides screened increases rapidly.

Tea-Bag Approach: An alternative method of multiple peptide synthesis was reported by Houghten,[10] and is commonly referred to as the "tea-bag" approach, where small amounts of resins representing individual peptides are enclosed in labeled, porous polypropylene containers. The tea-bags are immersed in individual solutions of the appropriate amino acids and standard Merrifield chemistry is utilized. Amino deprotection and washing steps are carried out by mixing all the tea-bags together in one reaction vessel followed by their subsequent separation for the next specific coupling. The peptides are cleaved in a multiple vessel HF apparatus.[11] Houghten's method affords fully-characterizable, non-solid support bound peptides which may afford more realistic interaction results than solid-support bound peptides. Neither analysis nor separation is necessary for the identification of a biologically-active peptide since the biological assay identifies the tea-bag which contains the most active peptide. The process has been adapted for semi-automation to allow the simultaneous synthesis of up to 120 different peptides.[12] Houghten's method has been further elaborated[13]into a combinatorial approach which allows for the rapid synthesis and biological evaluation of millions of individual, unmodified, free peptides in quantities useful for screening in solution. From a library of more than 34 million hexapeptides Houghten has been able to rapidly identify high affinity ligands.

Split Synthesis Method: Combinatorial libraries of peptides can also be synthesized on resin beads using a split synthesis method where each bead contains a single peptide.[14,15] In this approach the polymeric support is divided into equal portions for coupling to each of the individual twenty amino acids. This affords uniform coupling since competition between amino acids is eliminated. The individual polymers are combined in a single vessel for washing and deprotection and then divided again into twenty individual portions for the next coupling. The resulting peptides exhibit a purely statistical distribution of sequences. Through this combinatorial approach the complete universe of possible random peptides is rapidly prepared in approximately equimolar amounts. The resultant peptide library is incubated with a ligate attached to a reporter molecule and bound acceptors are identified by visual inspection, and then physically removed with tiny forceps coupled to a micromanipulator. The peptide/bead is separated from the ligate and sequenced by microsequencing techniques. This reusable library of several million beads can be screened in an afternoon. The approach is currently being expanded to incorporate cleavable linkers on each bead. After exposure to cleaving reagent the beads release a portion of their peptides into solution for biological assay and still retain sufficient peptide on the bead for structure determination.

VLSIPS: A technology combining solid-phase synthesis with photolithography and miniaturization has recently been described.[16] These techniques are brought together by the use of photolabile amine protecting groups resulting in light-directed, spatially addressable, parallel chemical synthesis of thousands of oligopeptides at defined sites on a micromolar scale. The synthetic substrate is prepared for amino acid coupling through covalent attachment of photolabile protected amine linkers. Light is used to selectively remove protecting groups and thereby activate the area for further synthesis. The pattern of exposure to light through a photolithographic mask determines the regions of the surface that are activated for chemical coupling. After activation the entire surface is exposed to the first amino acid, each of them bearing a photoactivatible protecting group on the amino terminus. Amino acid coupling occurs only in regions that were exposed to light in the previous step. The procedure is repeated until the entire peptide is prepared. The pattern of masks and the sequence of reactants define the products and their locations and combinatorial masking strategies are used to form large numbers of compounds in a small number of chemical steps. The use of photolithographic resolution allows for the preparation of tens-of-thousands of peptides on a one square centimeter solid surface. Furthermore, the identity of the peptide sequence at each location is known. Detection of binding is carried out by incubating the array with a fluorescently tagged receptor. The fluorescence intensity at each site will depend on the affinity of the receptor for the compound, the concentration of the receptor, and the number and density of interacting sites on the array. This technique allows the use of unnatural amino acids, nucleotides and even completely unique building blocks.

Alternative Solid Supports: The use of cellulose discs[17,18] which can be tightly stacked together into a reactor column with as many as 100 discs per column has been reported. By optimal utilization of identical amino acid additions the number of coupling cycles is kept to a minimum. Advantages of the cellulose approach include low cost, simple labelling, and easily variable quantities. Furthermore, the rigid structure and low swelling properties of the discs allow multiple couplings to be carried out under low pressures and continuous flow conditions. An apparatus can be built from inexpensive chromatography equipment[19] which synthesizes multiple peptides on cellulose sheets. The first amino acid is spotted and coupled to the sheet, then the entire sheet is immersed into appropriate solutions for washing, acetylation, and deprotection. The sheet is then dried, the next amino acid spotted, and the chemistry cycle repeated. Many compounds can be generated on one sheet and the process can be automated.

More recently, the use of cotton fabric as a solid support for multiple peptide synthesis has been reported and it appears that sequences which are difficult to prepare on polystyrene support work well on cotton.[20] In addition, the substantial amount of solvent held by cotton makes the use of reaction vessels unnecessary. Cotton segments placed on the perimeter of a disk-like rotor of a centrifuge release the solvent upon application of a centrifugal force.

An interesting variant utilizing two polymeric supports (Kel-F-g-styrene and polystyrene-1% divinylbenzene) to simultaneously synthesize two related peptides with most of the steps being carried out in the same reaction vessel under essentially identical conditions has been reported.[21] The polymers have different densities and hence can be separated by flotation at different stages of the synthesis. The method works well in the 4-19 amino acid residue range.

Finally, long-chain polystyrene-grafted polyethylene film matrices can be handled as individually labeled and readily separable sheets to afford an alternative solid support.[22]

Automated Methods: A fully automated peptide synthesizer capable of the simultaneous synthesis of up to 36 individual peptides or equimolar peptide mixtures (resin bound or solution phase) has been described and used to simultaneously prepare 361 peptides (19 pools of 19 peptides). This instrument can automate the resin-splitting approach described earlier.[23,24] In addition, several instruments based on traditional solid phase methods have been designed which allow for multiple peptide synthesis,[19,25-27] and robotics systems for these synthesizers or for use with the multi-pin technique have been described.[28]

Biological Methods

Phage Biopanning: Powerful biological methods have recently been described for the generation of molecular diversity in which libraries are presented on the surface of a bacteriophage.[29,30] These libraries comprise tens of millions of filamentous phage clones, each displaying a unique peptide sequence on the phage surface, and within the phage genome, the DNA sequence encoding for the peptide. The library is prepared by synthesizing a repertoire of random oligonucleotides to generate all combinations, followed by their insertion into the phage gene III and displayed as part of the gene III protein (pIII) at one tip of the virion. Alternatively, they can be cloned into the pVIII coat protein which has several thousand copies on the phage body.[31] Both proteins are synthesized with amino terminal signal peptides which are cleaved off concomitantly with translocation of the polypeptides through the inner membrane of the envelope. A ligate of interest (antibody, receptor, or other binding molecule) is used to affinity purify phage that display binding peptides by reacting the biotinylated ligate with the library and then "panning" the ligate/phage mixture on streptavidin coated plates. The unbound phage is removed by washing, the bound phage separated from the ligate, the phage propagated in E. coli, and the amino acid sequences of the peptides displayed on the phage are identified by sequencing the corresponding coding region of the viral DNA. Tens of millions of

peptides can be rapidly surveyed for binding and ideally, peptide ligands would be found that recognize all available sites in a binding pocket.

The initial phage peptide libraries afforded peptides of six amino acids. While on several occasions specific peptides for protein-binding ligates have been found, affinity is usually relatively weak and consensus peptides for protein-binding ligates, especially receptors, are not always found. It appears that a hexapeptide motif may not be suitable for identifying tight-binding peptides for many ligates, especially those that recognize folded determinants on proteins. Longer epitope regions and/or constrained epitopes may be required.

In Devlin's[32] library, the phage display contains 15-amino acid peptides. This increases the effective size of the library since each phage represents several smaller peptide determinants in various contexts of flanking residues. The ability to rapidly prepare a library of this magnitude means that an "all purpose library" is possible, meaning that no prior knowledge of the ligate is required. In the phage library constructed by Cwirla,[33] the peptide epitope begins immediately after the signal peptide cleavage site, making this library particularly useful for finding free peptide ligands since the free and phage bound peptides have the same free amino terminus.

Expression of bovine pancreatic trypsin inhibitor[34]and growth hormone fused to the pIII has been accomplished,[35]demonstrating that protein domains can fold properly on the surface of phage. These methods allow the use of phage to clone and select genes by the binding properties of the expressed proteins as well as the ability to rapidly explore structure-activity relationships on proteins. Proteins involved in DNA recognition might also be modified by random mutagenesis and selected for binding to novel DNA sequences.

Phage technology can also be applied to antibodies. Huse[36]first described a technique for the generation of recombinant libraries encoding the entire human repertoire using PCR and subsequent expression in E. coli. Heavy and light chains were expressed in different vectors and combined artificially in vitro to afford monovalent Fab fragments. Random combinatorial libraries of rearranged heavy and light chains isolated by PCR could be much more easily screened with the phage system than with the original Huse method because selection and mutagenesis is facilitated and antibodies reactive against defined epitopes may be directly selected from phage libraries. Filamentous phage displaying functional antibody domains on the surface have been constructed. Single chain Fv domains, comprising the H- and L-chain variable regions have been prepared by linking these two regions with a short nondescript peptide followed by phage expression,[37]or by simple non-covalent heavy and light chain association[38,39]where one chain is attached to the phage surface and the other is secreted into the bacterial periplasm where it associates with the phage bound chain. Two-chain Fab domains have also been

expressed in which the H-chain that comprises the variable region and the first subdomain of the constant region is fused to pVIII. The L-chain is again synthesized separately and associates with the phage attached H-chain.[40] These phage-attached antibodies retain the antigen binding specificities of the parent antibodies.

The question of how one obtains a naive library for the generation of antibody combinatorial libraries is important since a library obtained from an organism already has considerable bias. One way to solve this problem would be to chemically synthesize the six complementarity-determining regions (CDRs) in a random way and present them on a varied framework region. The first step toward this goal has been taken by the synthesis of random heavy chain (CDR3) which significantly contributes to antigen-binding contacts and diversity. A large library of Fab fragments differing only in HCDR3 was therefore displayed on phage surface and selected for antigen specificity.[41] This semisynthetic approach to gene synthesis allows for the production of human antibodies without immunization and without the biases imposed by constructing libraries from animals.

Alternative Expression Systems: An expression system that relies on a DNA binding protein, the lac repressor protein, to establish the connection between the peptide and the genetic material has been described.[42] This technique is referred to as peptides-on-plasmids. Random peptides are fused to the C-terminus of the repressor by cloning degenerate sequences at the end of the repressor gene present on a plasmid. The plasmid also contains lac repressor binding sites so the fusions bind the same plasmid that encodes them. In theory, many other DNA binding proteins can be used as well.

Non-Peptide Building Blocks

Several of the methods described above have also been used for the simultaneous synthesis of multiple oligonucleotide fragments. Frank[43] has described a segmental support approach utilizing paper discs as the synthetic support. The discs are stacked into four different reaction vessels for the addition of either A, C, G, or T. After each coupling cycle the filters are reshuffled into appropriate vessels for the next chain elongation. In addition, silica gel, controlled-pore glass (CPG), or teflon wafers containing CPG as the solid support[44] have been used to synthesize up to 100 oligonucleotides simultaneously.[45] Lastly, forty unique glycopeptides have been synthesized in a manual multiple column peptide synthesizer.[46]

Biological methods can also be used to prepare libraries of synthetic oligonucleotides that are selected for binding to an acceptor and then amplified by the polymerase chain reaction (PCR).[47-49] In this case, however, the repertoire is

limited to nucleotides or nucleotide analogues that preserve specific Watson-Crick pairing and can be copied by a polymerase.

Concluding Remarks

Currently available libraries are largely restricted to peptides and oligonucleotides. While the use of novel chemical building blocks could afford unlimited diversity, they lack the capacity for serial enrichment and there are potential difficulties in identifying the structures of active ligands. A method has recently been proposed which couples the chemical and biological methods through the construction of encoded combinatorial chemical libraries, in which each chemical sequence is labeled by an appended genetic tag, itself constructed by chemical synthesis.[50] In effect, a retrogenetic way of specifying each chemical structure is accomplished. Two alternating parallel combinatorial syntheses are carried out so that when a monomeric chemical unit is added it is followed by addition of an oligonucleotide sequence, which is defined as encoding that chemical building block. The library is built up by repetition of the process after pooling and division. Non-amino acid building blocks can be used as long as the synthetic techniques are compatible with the oligonucleotide chemistry.

What will the future hold? We can safely predict that the methods described herein will continue to be improved and we can expect new methods for the simultaneous synthesis of carbohydrates, peptide mimetics,[51] as well as for the synthesis of multiple structurally-unique compounds utilizing well-known synthetic organic reactions.

References:

1. Jung, G., Beck-Sickinger, A. G., *Angew. Chem. Int. Ed. Engl.* **1992**, 31, 367.

2. Geysen, H. M., Meloen, R., Barteling, S., *Proc. Natl. Acad. Sci. ,USA* **1984**, 81, 3998.

3. Bray, A. M., Maeji, N. J., Geysen, H. M., *Tet. Lett.* **1990**, 31, 5811.

4. Bray, A. M., Geysen, H. M., *J. Org. Chem.* **1991**, 56, 6659.

5. Valerio, R. M., Benstead, M., Bray, A. M., Campbell, R. A., Maeji, N. J., *Anal. Biochem.* **1991**, 171, 168.

6. Bray, A. M., Maeji, N. J., Jhingran, A. G., Valerio, R. M., *Tet. Lett.* **1991**, 32, 6163.

7. Geysen, H. M., Barteling, S., Meloen, R., *Proc. Natl. Acad. Sci. USA* **1985**, 82, 178.

8. Geysen, H. M., Rodda, S., Mason, T., Tribbick, G., Schoofs, P., *J. Immunol. Meth.* **1987**, 102, 259.

9. Geysen, H. M., Rodda, S. J., Mason, T. J., *Mol. Immunol.* **1986**, 23, 709.

10. Houghten, R. A., *Proc. Natl. Acad. Sci. ,USA* **1985**, 82, 5131.

11. Houghten, R. A., Bray, M. K., Degraw, S. T., Kirby, C. J., *Int. J. Pept. Prot. Res.* **1986**, 27, 673.

12. Beck-Sickinger, A. G., Duerr, H., Jung, G., *Pept. Res.* **1991**, 4, 88.

13. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T., Cuervo, J. H., *Nature* **1991**, 354, 64.

14. Lam, K., Salmon, S., Hersh, E., Hruby, V., Kazmierski, W., Knapp, R., *Nature* **1991**, 354, 82.

15. Furka, A., Sebestyen, F., Asgedom, M., Dibo, G., *Int. J. Pept. Prot. Res.* **1991**, 37, 487.

16. Fodor, S.P.A., Read, J. L., Pirrung, M.C., Stryer, L., Lu, A.T., Solas, D., *Science* **1991**, 251, 767.

17. Frank, R., Doering, R., *Tet. Lett.* **1988**, 44, 6031.

18. Eichler, J., Beyermann, M., Bienert, M., *Collect. Czech. Chem. Commun.* **1989**, 54, 1746.

19. Gausepohl, H., Kraft, M., Boulin, C., Frank, R., *Peptides Proc. Eur. Pept. Symp.* **1990**, 21, 206.

20. Eichler, J., Furkert, J., Bienert, M., Rohde, W., Lebl, M., *Pept. 1990, Proc. Eur. Pept. Symp.* **1991**, 21, 156.

21. Albericio, F., Ruiz-Gayo, M., Pedroso, E., Giralt, E., *React. Polym.* **1989**, 10, 259.

22. Berg, R. H., Almdal, K., Pedersen, W. B., Holm, A., Tam, J. P., Merrifield, R. B., *Peptides: Chem. Struct. Biol., Proc. Am. Pept. Symp.* **1990**, 11, 1036.

23. Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C., *Proc. Natl. Acad. Sci., USA* **1992**, 89, 4505.

24. Zuckermann, R. N., Kerr, J. M., Siani, M. A., Banville, S. C., *Int. J. Pept. Prot. Res.* **1992**, in press.

25. Fujii, N., Funakoshi, S., Otaka, A., Morimoto, H., Tamamura, H., Carpino, L., Yajima, H., *Pept. Chem.* **1989**, 26, 147.

26. Schnorrenberg, G., Gerhardt, H., *Tetrahedron* **1989**, 45, 7759.

27. Schnorrenberg, G., Wiesmueller, K. H., Beck-Sickinger, A. G., Drechsel, H., Guenther, J., *Peptides 1990, Proc. Eur. Pept. Symp.* **1990**, 21, 202.

28. Gausepohl, H., Kraft, M., Boulin, C., Frank, R., *Peptides: Chem. Struct. Biol., Proc. Am. Pept. Symp.* **1990**, 11, 1003.

29. Smith, G. P., *Science* **1985**, 228, 1315.

30. Scott, J. K., Smith, G. P., *Science* **1990**, 249, 386.

31. Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R., Casareni, G., *J. Mol. Biol.* **1991**, 222, 301.

32. Devlin, J. J., Panganiban, L. C., Devlin, P. E., *Science* **1990**, 249, 404.

33. Cwirla, S. E., Barrett, R. W., Peters, E. A., Dower, W. J., *Proc. Natl. Acad. Sci., USA* **1990**, 87, 6378.

34. Markland, W., Roberts, B., Saxena, M., Guterman, S., Ladner, R., *Gene* **1991**, 109, 13.

35. Bass, S. H., Wells, J. A., *Prot. Struct. Funct. Gen.* **1990**, 8, 309.

36. Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting, M., Burton, D. R., *Science* **1989**, 246, 1275.

37. McCafferty, J., Griffiths, A. D., Winter, G., Chiswell, D., *Nature* **1990**, 248, 552.

38. Hoogenboom, H. R., Griffiths, A. D., Johnson, K. S., Chiswell, D. J., Husdon, P., Winter, G., *Nucleic Acids Res.* **1991**, 19, 4133.

39. Barbas, C. F., Kang, A. S., R.A. Lerner, Benkovic, S. J., *Proc. Natl. Acad. Sci., USA* **1991**, 88, 7978.

40. Kang, A. S., Barbas, C. F., Janda, K. D., Benkovic, S. J., Lerner, R. A., *Proc. Natl. Acad. Sci., USA* **1991**, 88, 4363.

41. Barbas, C. F., Bain, J. D., Hoekstra, D. M., Lerner, R. A., *Proc. Natl. Acad. Sci., USA* **1992**, 89, 4457.

42. Cull, M. G., Miller, J. F., Schatz, P. J., *Proc. Natl. Acad. Sci., USA* **1992**, 89, 1865.

43. Frank, R., Meyerhans, A., Schwellnus, K., Blocker, H., *Meth. Enzymol.* **1987**, 154, 221.

44. Beattie, K., Logsdon, N. J., Anderson, R. S., Espinosa-Lara, J.M., Maldonado-Rodriquez, R.,Frost, J.D., *Biotechnol. Appl. Biochem.*, **1988**, 10, 510.

45. Seliger, H., Herold, A., Kotschi, U., Lyons, J., Schmidt, G., Eisenbeiss, F., *Nucleosides Nucleotides*, **1987**, 6, 137.

46. Peters, S., Biefeldt, T., Medal, M., Bock, K., Paulsen, H., *Tet. Lett.* **1991**, 32, 5067.

47. Tuerk, C., Gold, L., *Science* **1990**, 24, 505.

48. Ellington, A. D., Szostak, J. W., *Nature* **1990**, 346, 818.

49. Bock, L. C., Griffin, L. C., Latham, J. A., Vermas, E. H., Toole, J. T., *Nature* **1992**, 355, 564.

50. Brenner, S., Lerner, R. A., *Proc. Natl. Acad. Sci., USA* **1992**, 89, 5381.

51. Simon, R. J., Kania, R. S., Zuckermann, R. N., Huebner, V. D., Jewell, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frankel, A. D., Santi, D. V., Cohen, F. E., Bartlett, P. A., *Proc. Natl. Acad. Sci., USA* **1992**, in press.