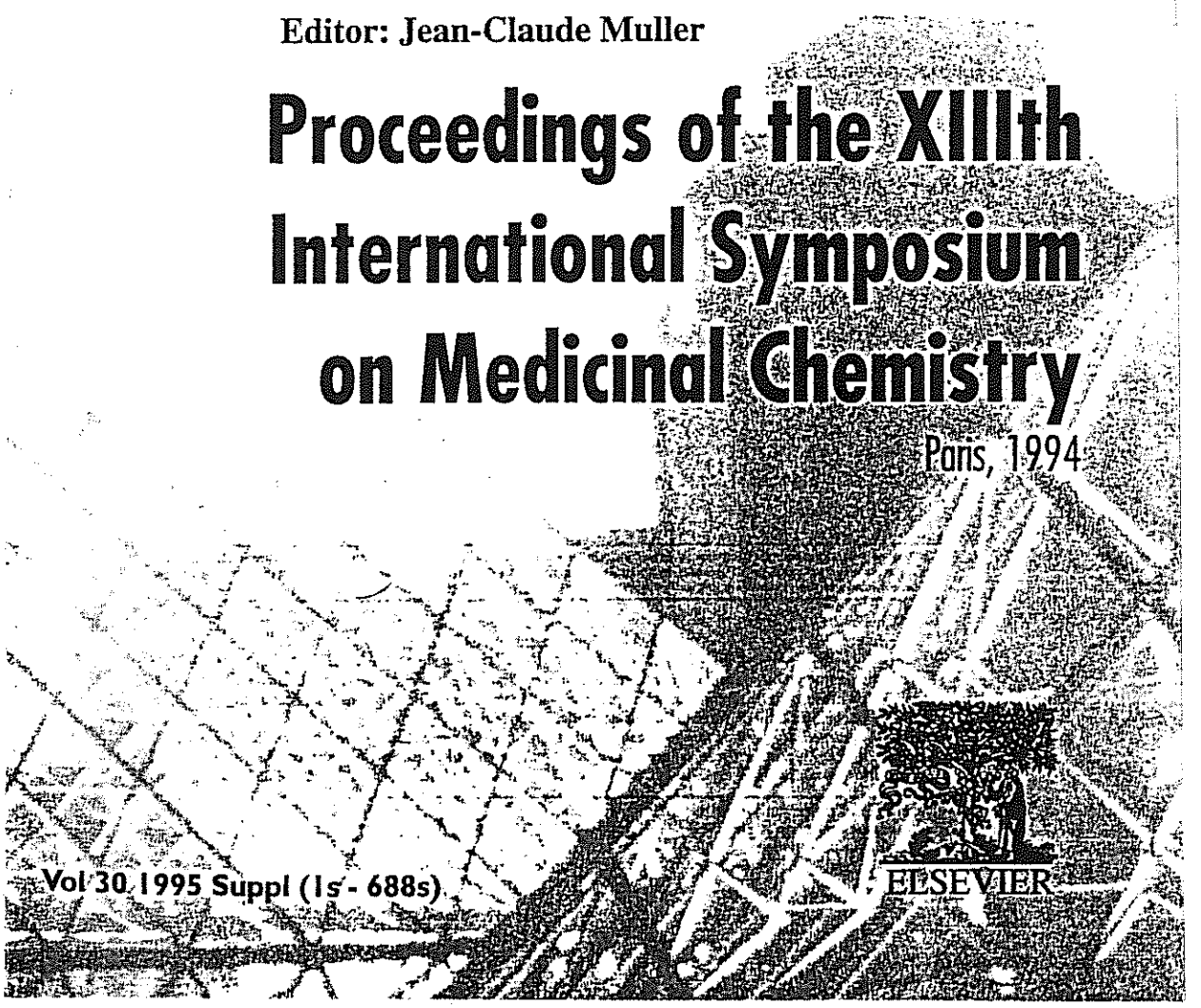
EUROPEAN JOURNAL OF

medicinal chemistry

Editor: Jean-Claude Muller

Proceedings of the XIIIth International Symposium on Medicinal Chemistry

Paris, 1994



Vol 30, 1995 Suppl (1s - 688s)

ELSEVIER

THE CONSTRUCTION AND USE OF PEPTIDE AND NON-PEPTIDIC COMBINATORIAL LIBRARIES TO DISCOVER ENZYME INHIBITORS

Bruce Seligmann, Farid Abdul-Latif, Fahad Al-Obeidi, Zuzana Flegelova, Olga Issakova, Petr Kocis, Viktor Krchnak, Kit Lam, Michal Lebl, James Ostrem, Pavel Safar, Nikolai Sepetov, Alena Stierandova, Peter Strop and Peter Wildgoose

Selectide Corp., Tucson, AZ 85737, USA (602-575-8040)

Key Words: combinatorial chemistry, non-peptide, peptide, coding, solid phase synthesis

1. INTRODUCTION

The basis of the Selectide Process, one of the drug discovery techniques based on combinatorial chemistry (for the review of various combinatorial library techniques see e.g. [1,2]), was first published by Lam in 1991[3]. This process includes the rapid synthesis of libraries of compounds and their screening either with the compounds remaining attached to solid phase particles, or following their release into solution. The identity of compounds is not tracked during synthesis, but instead the composition of positive compounds is determined after screening, using chemical information on the solid phase support. The medicinal chemist can custom design libraries to improve the probability of identifying lead compounds, or generic libraries of compounds can be synthesized and screened without any preconceived bias. Once initial leads are discovered the Selectide Process can be used to optimize the activity from μM to nM by synthesizing and screening secondary libraries based on the structure of the initial lead. Libraries have been synthesized and screened against both receptor and enzyme targets of pharmacologic interest. Selected leads identified from these libraries have been rapidly optimized using secondary libraries to produce compounds with double digit nM activity. The one-peptide-one-bead approach has been extended to non-peptide libraries, making it a one-compound-one-bead method. Methods for determining the identity of positive compounds have been developed based both on coding and mass spectroscopy. The coding method developed at Selectide requires a single step of Edman degradation and HPLC analysis, and is termed "bar" coding. Coding molecules can be sequestered in the interior of the solid phase particle away from the test compound in order to prevent the acceptor molecule from interacting with the coding molecules. This physical sequestration is termed "shaving", and is a method which can be used to chemically

differentiate the interior of the solid phase particle from its surface.

The Selectide Process, and combinatorial chemistry in general, has been validated by the successful identification of thrombin inhibitors and Factor Xa inhibitors from libraries, and the optimization of these inhibitors to produce, in the case of Factor Xa, compounds with sub-nanomolar activity *in vitro*.

2. RESULTS

2.1 ASSAYS WITH THROUGHPUT MATCHING SYNTHETIC CAPACITY

The synthetic capacity of the split and mix method [4] which is utilized in the Selectide Process can produce millions of compounds per library. This capacity is useless unless screening assays can be developed that have a matching throughput of millions of compounds per week. We have developed three types of screening formats which have high throughput rates. If the target molecule is available free in solution, screening can be carried out with the compounds remaining attached to the solid phase particles in an "on-bead" assay format

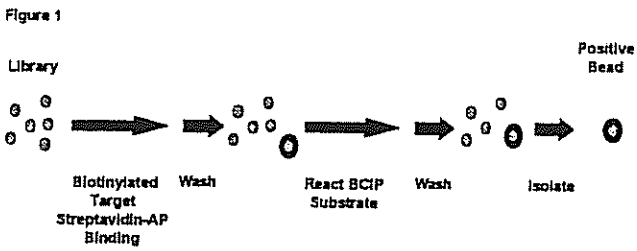


Figure 1: Library beads (grey circles) are incubated with biotinylated target plus streptavidin conjugated with alkaline phosphatase. This complex, bound to positive beads (dark ringed circle), precipitates the alkaline phosphatase substrate BCIP, (grey ringed circle), enabling positive beads to be selected for compound identification.

(Figure 1) [3, 5-7]. In the case where the target molecule is not amenable to the on-bead approach, e.g. the target is membrane bound, compounds can be released from the solid phase particles and activity determined with the compounds free in solution in a "double release" assay (Figure 2). In this approach mixing of 100 to 500 beads per well in the first release step enables the throughput of any conventional solution phase assay (membrane binding, enzyme, cytotoxic, or functional) to be increased to over 500,000 compounds per day, while permitting specificity since in the second release there is only one bead per well tested. The underlying feature of this approach is that the relationship between wells containing the beads (or bead, in the case of second release) and the assay result is tracked, so the positive beads can be recovered and the identity of the compound released from each positive bead determined from information remaining on the bead. Aspects of this assay were recently published [8-10].

Figure 2

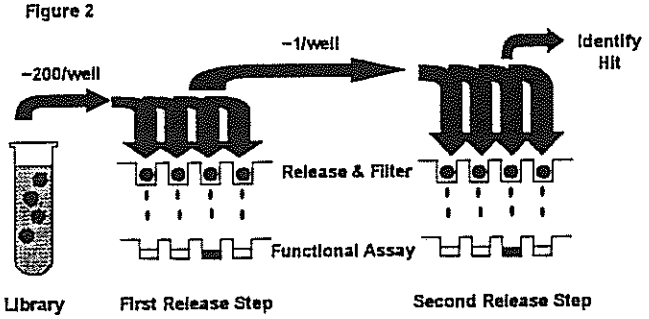


Figure 2: Library beads (circles) are pipetted into the wells of a filter plate (~200/well), a portion of compound is chemically cleaved from each bead, the resulting releasate is filtered into a replicate plate and assayed. The beads corresponding to a positive well (dark well) are redistributed into a second master plate (1/well), additional material is cleaved from the bead, the releasate is filtered into a replicate plate, and assayed. The bead corresponding to a positive assay result (dark well) is recovered and the identity of the released compound determined.

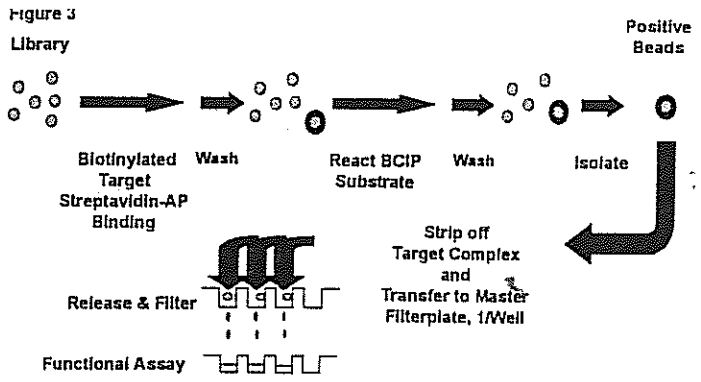


Figure 3: Library beads were screened as described in Figure 1. The positive beads were distributed into a master filter plate, a portion of compound released from each bead, releasate is filtered into a replicate plate, assayed, and analyzed.

Combining these two formats produces a "hybrid" assay where the first step of screening is conducted with compounds attached to the solid phase particles, and then the compound from positive beads is released into solution and activity accurately determined (Figure 3). The validity of this assay is shown by the following. A pentamer library was synthesized containing both D and L amino acids with the third position fixed as L-arginine (X-X-Arg-X-X), screened, and 12 hits were sequenced. Five beads had the sequence IWRVA, two IWRFA, two IWRFK, and one each of IWRVK, IWRIA, and IWRIK (single letter amino acid code). From these results a common motif of IWRVA was deduced. Since there was no way to know if the identified residues were in the D or L configuration, all 16 possible D/L isomer combinations of the IWRVA motif were synthesized on beads and screened. Two were positive. Compound was released from these two bead samples and used to determine the Ki for each positive compound. The purity of material released was assessed by HPLC

and the concentration determined by amino acid analysis, permitting precise Ki values of 31 μM for IwRvA and 21 μM for IwRva to be calculated.

2.2 APPLICATION OF PRIMARY PEPTIDE LIBRARIES TO DISCOVER INHIBITORS OF THROMBIN

A thrombin on-bead screen was established and validated using a published peptide inhibitor, fPRPG. A variety of pentapeptide libraries were synthesized containing both D and L amino acids, and screened for binding to thrombin. Hits were identified which inhibited thrombin over a range of 1 to 300 μM. Table I lists representative hits identified from a DLDLD pentapeptide library, together with the Ki values determined for the hits after resynthesis and testing in solution. The pentapeptide fPRPG was tested in parallel as a control.

Table I.
Examples of Inhibitors of Thrombin Discovered From a DLDLD Pentapeptide Primary Library Covering Whole Range of Activities Obtained

Compound*	Ki (μM)
fPRPG (control)	21
fPrPG	700
fFrNr	200
fIrNq	78
fFrSi	32
dYaRw	6
wFrPf	3.5

* Structure of compounds indicated using single letter codes for amino acids, using uppercase for L and lower case for D isomers. 50 sequences obtained for this library, 46 contained arginine.

2.3 APPLICATION OF SECONDARY OPTIMIZATION LIBRARIES TO IMPROVE THE ACTIVITY OF THROMBIN INHIBITORS

It is known that thrombin has both an active site and an exosite to which peptides can bind.

Hirulog, a 21 residue peptide, consists of a sequence which binds in the active site of the enzyme (and is a substrate) connected with a series of glycine residues to a sequence which binds in the exosite. Alone, each of these binding sequences exhibit low affinity (20 μM and 2 μM, respectively) but Hirulog exhibits a binding affinity of 2 nM. A secondary optimization nonapeptide library was designed in which the first four positions were fixed as fPRP and the last five positions were randomized. The idea was to explore potential binding regions close to the active site, but possibly different from the previously identified exosite. The result was a 25 nM compound, truncated analogs of which had progressively lower activity (Table II). The importance of specific residues was determined from an alanine scan. Thus from a single "secondary" library the activity of a lead was improved 1000-fold, producing a peptide which was less than half the size and only 10-fold less active than Hirulog, a compound which is currently in clinical trials as a drug to prevent coagulation.

Table II.
Thrombin Inhibitors Discovered From Secondary Optimization Library

Compound*	Ki (nM)
Library Lead	
fPRP56789b	25
Analogs	
fPRP56789	45
fPRP5678	62
fPRP567	167
fPRP56	180
fPRP5	400
fPRP5678Ab	152
fPRP567A9b	490
fPRP56A89b	46
fPRP5A789b	139
fPRPA6789b	2195

* Structure of compounds indicated using single letter codes for amino acids, using uppercase for L and lower case for D isomers. Certain residues are coded with numbers. The linker contained the residue β-alanine (b), included in the resynthesized lead and analogs of the alanine scan.

2.4 APPLICATION OF FIRST GENERATION NON-PEPTIDIC PRIMARY LIBRARIES TO IDENTIFY INHIBITORS OF THROMBIN

Peptides represented a good system to develop both the combinatorial synthetic process and assay technology. However, peptides may not represent the best starting point for drug development, and libraries of peptides offer limited structural diversity, particularly conformational diversity. We began synthesizing "first-generation" non-peptidic structures composed of building blocks other than amino acids linked together through amide bonds, or linked to a multifunctional central "scaffold" subunit through amide bonds [11,12]. Both the chemical diversity as well as the structural diversity is greater in these libraries than in peptide libraries. The number of potential hydrogen bonds which can be formed with the linking chemical groups can be controlled through the use of different linking chemistry or the generation of tertiary rather than secondary amines. This attribute, and the absence of peptidic structure, make these first-generation libraries attractive. Examples are shown in Figure 4. A common feature of many of these first-generation libraries is the use of a multifunctional subunit as a "scaffold" onto which additional building blocks, amino acids or other organic molecules, can be attached [13-21]. The scaffold and building blocks together determine the structural diversity represented by each library.

In the case of peptides the test compound itself can be directly sequenced using Edman degradation or mass spectroscopy [3,10,22], and its identity determined. In the case of compounds (e.g. non-peptidic or N-substituted peptides) where the identity cannot be determined from the amount of material available, it is necessary to incorporate a different but corresponding molecule on the bead from which the identity of the test compound can be inferred. This technique is referred to as coding, methods for which have been described

[11, 13, 23-25]. We utilize an amino acid "bar-coding" method combined with bead "shaving" [26] to sequester a coding sequence in the interior of the library beads. The test compound is exhibited on the surface where it (but not the code) is accessible to macromolecular targets.

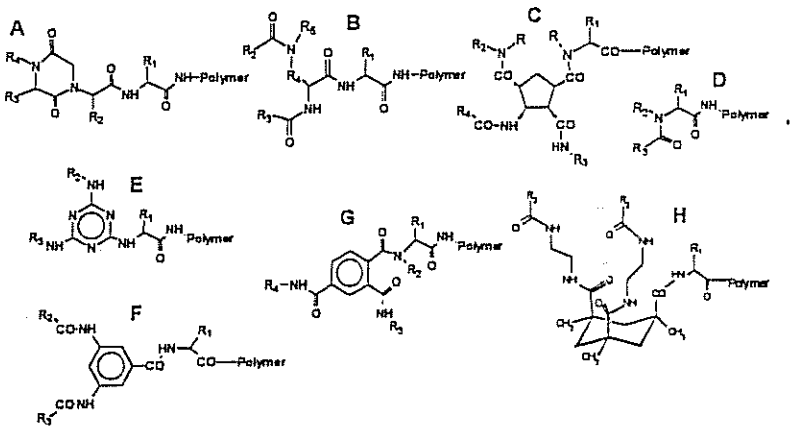


Figure 4: Structures shown with attachment points for building blocks (Rx) indicated.

In the specific case reported here a Kemp's triacid library (Figure 4H) was screened to identify first-generation non-peptidic compounds which bound to thrombin and inhibited its enzymatic activity. The identity of the Kemp's triacid compounds was determined by mass spectroscopy. Figure 5 presents an example of the fragmentation spectra of the compound from this library (upper trace), and spectrum of resynthesized compound, the structure of which was deduced from the upper spectrum (lower trace). Figure 6 depicts inhibition for several of the Kemp's triacid hits, compared to fPRPG. The best of these, SEL 2800 (structure shown), exhibited a Ki of 4 μM, 5-fold more potent than the standard peptide fPRPG (Ki ~ 20 μM) which is of similar size.

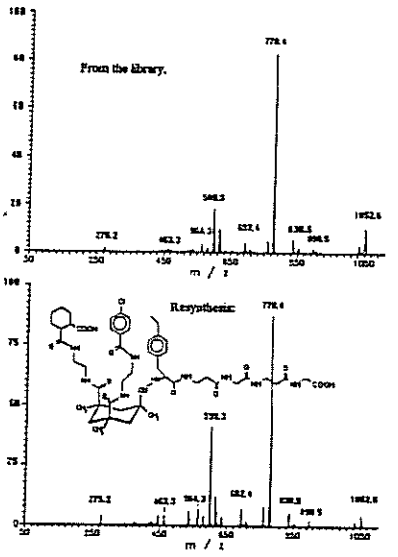


Figure 5 Legend: Panel A - the fragmentation pattern for material recovered from a single positive bead from a thrombin assay from which an identity was assigned. Panel B - the mass spectra for the resynthesized compound, confirming the compound had been properly identified.

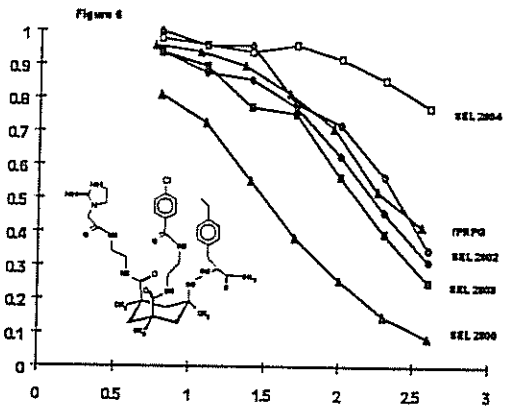


Figure 6: Inhibition curves for Kemp's triacid compounds identified by their Selectide number. The structure for the most active (IC_{50} 4 μM) is shown in the insert without the linker which was the same as shown in Figure 5. The peptide fPRPG was tested as a control (IC_{50} 20 μM).

2.5 DEVELOPMENT OF A HIGHLY SPECIFIC FACTOR Xa INHIBITOR SERIES FROM A LIBRARY LEAD

The identification of hits from a library does not in itself validate the concept that this approach is useful in the development of drugs. The question remains as to whether "quality" leads can be identified which can readily be optimized into clinical candidates with the required potency, selectivity, and pharmacokinetic properties, and whether the process of optimization can be completed in a timely manner without undue expenditure of resources. Table III summarizes the activity of a lead series of Factor Xa inhibitors which were discovered from a library and subsequently optimized. SEL 1691 was the original hit obtained from the library. While not particularly potent as an inhibitor (Ki ~20 μM), this lead was selective, as evidenced by the lack of inhibition of thrombin when tested at a concentration as high as 200 μM. The activities of four compounds optimized from this lead are depicted in Table III.

Table III.
Summary Table of Factor Xa Inhibitors

Compound[a]	Mol. Wt. (g/mole)	Factor Xa Inhib.[b] (Ki, nM)	Thrombin Inhib.[c] (Ki, nM)	Dilute PT Inhib.[d] (IC_{50}, nM)
rTAP[e]	6977	0.4	>200,000	5
SEL 1691	952	20,000	NT	NT
SEL 2489	726	58	>1,000,000	200
SEL 2452	761	31	25,000	140
SEL 2711	760	3	52,000	220
SEL 2392	753	0.9	10,000	55
SEL 2684	810	0.3	1,000	114

[a]Compounds identified by Selectide number.
[b]Inhibition of purified human Factor Xa [30]
[c]Inhibition of purified human thrombin
[d]Inhibition of coagulation, measured using a dilute PT assay
[e]rTAP is recombinant tick anticoagulant peptide [27]

The most potent compound inhibited Factor Xa with a Ki of ~300 pM. Selectivity was maintained. Furthermore, we have shown that these compounds inhibit coagulation. For reference the activity of recombinant tick anticoagulant peptide (rTAP) is shown [27]. This small protein (6985 MW) inhibitor has proven efficacious in animal models of thrombosis. Studies which are not shown indicate that analogs of the compounds whose activities are shown in Table III have a suitable potency and half life in animals and are active in models of thrombosis, making it likely that this series will produce an attractive clinical candidate for testing in humans. Factor Xa is a particularly attractive target in an area (anticoagulation) where there is a large unmet clinical need for an improved parenteral agent and an effective oral drug. Despite years of research by the pharmaceutical industry, only one other low molecular weight inhibitor has been described [28]. What is additionally of interest is the time line from initiation of screening, through identification and confirmation of the initial lead, and initiation of *in vivo* studies. Once the on-bead bioassay had been validated (using Tenstop, [29] attached to the polymeric beads) and screening was initiated, SEL 1691 was discovered, resynthesized, and characterized as a lead within two months. Testing in animals began four months later.

3. DISCUSSION

The schemes presented in Figures 1 and 2 summarize how the throughput of the bioassay screens can approach 500,000 to several million compounds per day, a necessary feature in order to take advantage of the synthetic capacity of solid phase combinatorial chemistry which can be realized by adapting the split synthesis method for synthesizing mixtures of compounds [4] to the synthesis and use of one-compound-one-bead [3]. Screening a secondary optimization library based on the structure of an active molecule can produce numerous

positive compounds, but the investigator is only interested in those which are most potent, and must be able to distinguish between compounds of similar potency. The hybrid assay summarized in Figure 3 combines the ability of the on-bead assay to efficiently identify positive compounds and the release assay to reliably determine the activity of compounds in solution before determining their identity.

Armed with this battery of assays, inhibitors of thrombin were discovered (Table I). A single optimization library provided a novel structure 1000-fold more active than the lead on which the optimization library was based (Table II) and less than half the size and having nearly the potency of Hiruolg, a compound which is in human trials. Furthermore, this discovery defined an additional binding region on thrombin that could be exploited to develop novel low molecular weight (oral) inhibitors of thrombin. The significance of this example to the drug discovery process and the medicinal chemist is that it demonstrates that screening of primary libraries can produce a variety of leads from which to select those with the best potential for optimization, and secondary libraries based on these initial leads can be easily and quickly synthesized to explore the potential for optimization and to provide compounds of greater activity.

While success with peptide libraries is illustrative, the true potential of combinatorial chemistry lies in its application to libraries of structurally and chemically diverse small molecules. The methods to determine the structure of such compounds indirectly using coding [11,13,23-25], or directly by mass spectroscopy have been developed [22]. In the illustrated case, mass spectroscopy was used to directly identify the structure of positive compounds from a Kemps' triacid library screened in a thrombin assay (Figure 5). We have focused on a mass spectroscopy example rather than a coding example in this review because

mass spectroscopy represents an elegant, though not universally applicable, approach to structure determination of compounds from libraries. The results prove that this method can be applied successfully to identify a compound using material from a single bead containing 10-100 pmoles of organic compound.

The Kemps' triacid library represents what we refer to as a first-generation non-peptidic library, designed to explore non-linear structural diversity. It is noteworthy that one of the initial leads identified was of similar size but more potent (Figure 6) than the original small peptide lead (fPRPG) that was found years ago and has served as the basis for most of the published work on small peptide thrombin inhibitors.

Data from our Factor Xa drug discovery project was presented as illustrative of success with a much more specific enzyme in the coagulation cascade and a target which has proven difficult using classical drug discovery approaches. Factor Xa screening produced a 20 μM lead within two months of assay validation, which was of sufficient quality to begin optimization. Overall success is evidenced by the data shown in Table III. Optimization resulted in 40,000-fold increase in inhibitory activity while retaining the ~10,000-fold selectivity of these inhibitors, and also in the successful development of the *in vivo* activity required of a potential therapeutic agent. The project time line of 6 months from establishing the screen to *in vivo* testing (4 months from lead discovery to *in vivo* testing) highlights the speed of both lead identification and optimization using our combinatorial chemistry approach. Factor Xa is a challenging target. The success we report therefore validates the concept that synthetic combinatorial chemistry represents a valuable new approach to drug development.

4. CONCLUSION

We have demonstrated that using the Selectide Process combinatorial chemistry can be used to discover novel peptide and first-generation non-peptidic inhibitors. Furthermore, combinatorial chemistry can be used to optimize hits and shorten the drug discovery and lead optimization processes. Not only can high quality leads be obtained in short order, but we believe our initial successes present both an opportunity and a challenge. The medicinal chemists have before them the power of numbers in a functional process of synthesis and analysis of millions of compounds. The challenge will be to develop the synthetic reactions and design the next generations of libraries to increase molecular diversity and direct pharmaceutical potential of the compounds being synthesized and screened. Combinatorial chemistry may well represent a drug discovery tool for the medicinal chemist akin to the tool molecular biology has become for our industry.

5. REFERENCES

1. Gallop, M.A., Barrett, R.V., Dower, W.J., Fodor, S.P.A. and Gordon, E.M., *J. Med. Chem.*, 37 (1994) 1233-1251.

2. Gordon, E.M., Barrett, R.V., Dower, W.J., Fodor, S.P.A. and Gallop, M.A., *J. Med. Chem.*, 37 (1994) 1385-1401.

3 Lam KS, Salmon SE, Hersh EM, Hruby VJ, Kazmierski WM, Knapp RJ (1991) *Nature 354*, 82-84

4 Furka A, Sebestyen F, Asgedom M, Dibo G (1991) *Int J Pept Protein Res 37*, 487-493

5 Lam KS, Lebl M (1992) *Immunomethods 1*, 11-15

6 Lam KS, Hruby VJ, Lebl M, Knapp RJ, Kazmierski WM, Hersh EM, Salmon SE (1993) *Bioorg. Med. Chem. Lett 3*, 419-424

7 Salmon SE, Lam KS, Felder S, Yeoman H, Schlessinger J, Ullrich A, Krchnak V, Lebl M (1994) *Acta Oncologica 33*, 127-131

8 Salmon SE, Lam KS, Lebl M, Kandola A, Khattri PS, Wade S, Patek M, Kocis P, Krchnak V, Thorpe DE, Felder S (1993) *Proc Natl Acad Sci USA 90*, 11708-11712

9 Lebl M, Patek M, Kocis P, Krchnak V, Hruby VJ, Salmon SE, Lam KS (1993) *J Pept Protein Res 41*, 201-203

10 Kocis P, Krchnak V, Lebl M (1993) *Tetrahedron Letters 34*, 7251-7252

11 Lebl M., Krchňák V., Šafář P., Stierandová A., Sepetov N.F., Kočiš P., Lam K.S. (1994)In *Techniques in Protein Chemistry (Crabb J., Ed.)*, Vol.5, pp. 541-547. Academic Press, Orlando

12 Lebl M., Krchňák V., Sepetov N.F., Nikolaev V., Stierandová A., Šafář P., Seligmann B., Štrop P., Lam K.S., Salmon S.E. (1994) In *Innovation and Perspectives in Solid Phase Synthesis (R. Epton, Ed.)*, pp. 233-238, Mayflower Worldwide Ltd., Birmingham

13 Nikolaiev V, Stierandova A, Krchnak V, Seligmann B, Lam KS, Salmon SE, Lebl M (1993) *Peptide Res 6*, 161-170

14 Pavia MR, Sawyer TK, Moos WH (1993) *Bioorg Med Chem Lett 3*, 387

15 Bunin BA, Plunkett MJ, Ellman JA (1994) *Proc Natl Acad Sci USA 91*, 4708-4712

16 Deshpande MS (1994) *Tetrahedron Letters 35*, 5613-5614

17 Simon RJ, Kania RS, Zuckerman RN, Huebner VD, Jewell DA, Banville S, Ng S, Wang L, Rosenberg S, Marlowe CK, Spellmeyer DC, Tan R, Frankel AD, Santi DV, Cohen FE, Bartlett PA (1992) *Proc Natl Acad Sci USA 89*, 9367-9371

18 Zuckerman 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) *J Med Chem 37*, 2678-2685

19 DeWitt SH, Kiely JS, Stankovic CJ, Schroeder MC, Cody DMR, Pavia MR (1993) *Proc Natl Acad Sci USA 90*, 6909-6913

20 Zuckerman RN, Kerr JM, Kent SBH, Moos WH (1992) *J Am Chem Soc 114*, 10646- 10647

21 Chen C, Randall ALA, Miller BR, Jones DA, Kurth MJ (1994) *J Am Chem Soc 116*, 2661-2662

22 Stankova M., Issakova O., Sepetov N.F., Krchnak V., Lam K.S., Lebl M (1994) *Drug Development Research* 33, 146-156

23 Needles MC, Jones DG, Tate EH, Heinkel GL, Kochersperger LM, Dower WJ, Barret RW, Gallop MA (1993) *Proc Natl Acad Sci USA 90*, 10700-10704

24 Ohlmeyer MHJ, Swanson RN, Dillard LW, Reader JC, Asouline G, Kobayashi R, Wigler M, Still WC (1993) *Proc Natl Acad Sci USA 90*, 10922-10926

25 Kerr JM, Banville SC, Zuckermann RN (1993) *J Am Chem Soc 115*, 2529-2531

26 Vágner J., Sepetov N.F., Krchňák V., Lam K.S., Lebl M., Barany G. (1994) In *Innovation and Perspectives in Solid Phase Synthesis (R. Epton, Ed.)*, pp.347-352. Mayflower Worldwide, Birmingham

27 Vlasuk GP (1993) *Thrombosis and Haemostasis 70*, 212-216

28 Katakura S, Nagahara T, Hai T, Iwamoto M (1993) *Biochem Biophys Res Comm 197*, 965-972

29 Sturzebecher J, Sturzebecher U, Vieweg H, Wagner G, Hauptmann J, Markwardt F (1989) *Thrombosis Res 54*, 245-252

30 Lottenberg R, Christensen, Jackson CM, Coleman RL (1981) *In: Methods in Enzymology* Vol 80 (Academic Press) 341-361