Drug Discovery by Combinatorial Chemistry—The Development of a Novel Method for the Rapid Synthesis of Single Compounds**

Nicholas K. Terrett,* Mark Gardner, David W. Gordon, Ryszard J. Kobylecki and John Steele

Abstract: Through the advantages conferred in the speed of synthesis, combinatorial chemistry is making a significant impact on the process of drug discovery. The mix-and-split paradigm has been an effective method for the production of compound mixtures, although there is now a need for new, fast library approaches to generate well-characterised single compounds. Having already demonstrated the successful preparation and application of library mixtures, we have now developed a novel combinatorial method for the production of single compounds.

Keywords: combinatorial chemistry · compound libraries · drug research · solid-phase synthesis

The final outcome of the drug candidate discovery process is the identification of a chemical compound that has both the desired potency and selectivity for a specific biological target, and suitable efficacy and bioavailability in appropriate test systems. Drug discovery is a lengthy and expensive business, and it frequently takes up to five years from initiating a project to the point where a drug candidate can be nominated for development and clinical trials.

Early in the lifetime of a drug discovery project, medicinal chemists require a lead—a compound with an appreciable degree of affinity for the enzyme or receptor biological target. With a lead in hand, they can proceed to the second phase, namely, the identification of a drug development candidate by the stepwise, incremental improvement of the lead's biological profile. To address the first phase, the drug discoverer can use a known literature or competitor agent as a lead, or use a knowledge of structure or mechanism of the target to design a new lead. But often, neither of these routes are possible, and empirical screening of compound files is the only approach to lead discovery. The average pharmaceutical company archive contains in the region of 200 000 compounds, so automated methods for rapid screening have been developed in order to test thousands of compounds in a period of weeks.

However, even 200 000 compounds may be too small a number to guarantee the discovery of novel biologically active lead compounds. Thus, a technique that permits the rapid synthesis of thousands or even millions of new compounds for lead discovery has quickly captured the imagination of many pharmaceutical chemists. This technique is combinatorial chemistry, an approach that in its power to generate huge numbers of new compounds is already changing the face of drug discovery.[1]

Combinatorial chemistry embraces a wide diversity of different high-speed and parallel synthetic techniques. The key to rapid production of novel molecules is the ability to use separation techniques to effect an easy purification of intermediates and final products. To achieve this, solid-phase synthesis in particular has been the main influence on library methodology. The ability to synthesise compounds on an inert polymeric resin bead, to force a reaction to completion by the addition of excess reagents and monomers, and to then remove all the unwanted material by a simple filtration and wash is at the heart of most library synthesis (Figure 1).

A second key component of combinatorial chemistry is the use of the mix-and-split paradigm, first described by Furka,[2] which can be used to generate mixtures of huge numbers of compounds. The method works as follows: a sample of resin support material is divided into a number of equal portions (x) and each of these are individually treated with a single different reagent. After completion of the reactions, and subsequent washing to remove excess reagents, the individual portions are recombined, the whole is thoroughly mixed, and may then be

[*] Dr. N. K. Terrett, M. Gardner, D. W. Gordon, R. J. Kobylecki, J. Steele
Discovery Chemistry, Pfizer Central Research
Sandwich, Kent CT13 9NJ (UK)
e-mail: nick_terrett@sandwich.pfizer.com
Fax: Int. code +(1304) 618422

[**] From a lecture delivered at the meeting of the British Association for the Advancement of Science at the University of Birmingham, September 9–13, 1996.

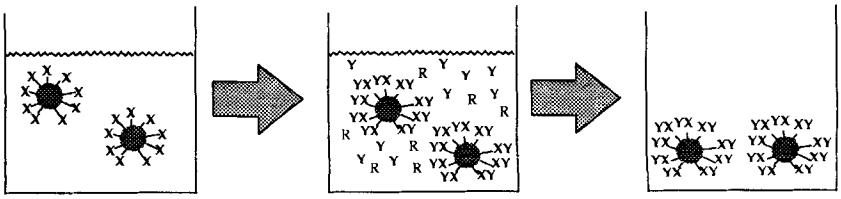

Figure 1. The use of a solid phase for the synthesis of X Y allows excess monomer (Y) and reagents (R) to be removed by a simple filtration.

  0947-6539/97/0312-1917 $ 17.50+.50/0

divided again into portions. Reaction with a further set of activated reagents gives the complete set of possible dimeric units as mixtures, and this whole process may then be repeated as necessary (for a total of n times). The number of compounds obtained arises from the geometric increase in potential products; in this case x to the power of n.

Following the seminal publications from Houghten[3] and Lam,[4] we at Pfizer considered whether we could use the methods of combinatorial chemistry for our in-house lead discovery programmes. We initially embarked on a trial synthesis of a library of 30752 trimeric compounds, to test whether the mix-and-split technique would permit the identification of a known endothelin antagonist, FR-139,317.[5]

The endothelins[6] are a family of peptides with potent and sustained vasoconstrictor activity. Raised plasma levels have been observed in conditions related to systemic vasoconstriction, such as hypertension, heart failure and angina, and endothelin is thought to play a key role in acute and chronic renal failure. Consequently, there has been a concerted search for antagonists of the endothelin receptor subtypes. Many different antagonists have been described in the literature, and one of the first described, with selectivity for the ET_A receptor subtype, was the modified tripeptide, FR-139,317 from Fujisawa[7] (Figure 2).

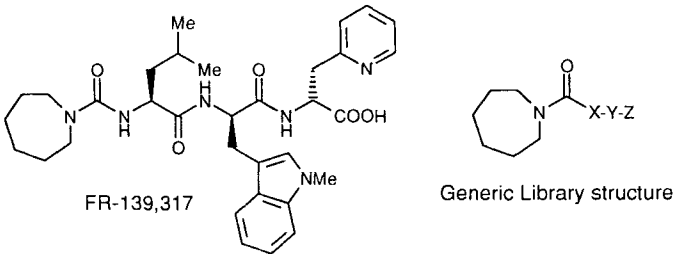

Figure 2. The endothelin receptor antagonist FR-139,317, and the generic structure of the combinatorial library designed to include this ligand.

The synthesis employed a set of 32 monomeric amino acids containing both L and D natural amino acids and a number of unnatural amino acids. Three monomers in particular (L-Leu, *N*-Me-D-Trp and D-2-PyrAla) were included to ensure that FR-139,317 was synthesised as a positive control within the library, and all trimers were *N*-acylated to generate the homopiperidine urea (see generic library structure in Figure 2). The library synthesis followed a standard procedure on Wang polystyrene resin with the use of monomers protected by the 9-fluorenylmethoxycarbonyl (Fmoc) group, to allow the formation of C-terminal carboxylic acids following trifluoroacetic acid (TFA) catalysed cleavage from the resin beads. As it was observed that D-2-PyrAla would only couple in very low yield under the library synthesis conditions, this amino acid was omitted from the X and Y positions, and thus the total library size was 30752 compounds ($32 \times 31 \times 31$).

The library synthesis and subsequent deconvolution of active mixtures proceeded uneventfully, and gratifyingly FR-139,317 was "rediscovered" as one of the most potent constituents of the library. A number of other endothelin antagonists were identified and their structures confirmed by mass spectrometry. One unexpected observation during the first round of screening was that the mixture with X = *p*-(aminomethyl)benzoic acid appeared to possess the greatest activity. This mixture was followed up by further synthesis, and it was found that no mixtures (position Y fixed) containing X = *p*-(aminomethyl)benzoic acid had any significant activity. The analogue of FR-139,317 that contains *p*-(aminomethyl)benzoic acid in the place of L-Leu was synthesised independently and found to have an IC_{50} for the ET_A receptor of 1.1 μM. It is frequently observed that mixtures can give rise to biological activity that cannot be confirmed in deconvolution. The activity of the *p*-(aminomethyl)benzoic acid mixture in this case may be explained by the presence of a large number of weakly active components that additively give rise to the observed receptor affinity. However, in this library experiment we successfully demonstrated that it is possible to synthesise thousands of compounds in very few individual chemical steps, and furthermore that we could discover a known biologically active receptor antagonist and a number of structurally related analogues.

We have since applied the mix-and-split synthesis to the production of many other libraries, and in particular have synthesised libraries of non-peptides, which offer more suitable starting points for drug discovery. Although there are many peptide drugs, as a class peptides tend to suffer from poor oral bioavailability and rapid clearance from the body. Thus, the applications of combinatorial chemistry to drug discovery are most powerful when used to prepare libraries of drug-like compounds employing coupling chemistry other than peptide bond formation. The range of solid-phase chemistry now available is substantial and continues to grow with the weekly publication of papers describing new synthetic methodology.[8]

Following the synthesis of the library targeted at the endothelin receptor described above we synthesised other combinatorial libraries with the specific intention of discovering unprecedented lead compounds. A number of trimeric structures (Figure 3)

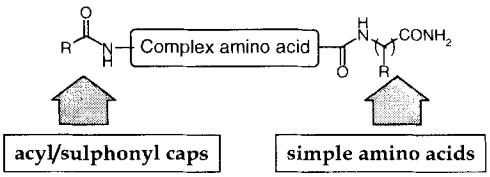

Figure 3. The generic structure for the library tested against the neurokinin-2 receptor. The compounds were based on a central complex amino acid capped on the C-terminus with a simple amino acid and on the N-terminus with an acyl or sulphonyl group.

were generated which consisted of a complex amino acid, capped with an acyl group on the amino terminus and a "simple" amino carboxamide on the C-terminal. Each central amino acid was used in the generation of 10000 component library, prepared as mixtures of 100 compounds, and screened against a range of biologically important targets. In particular, we were seeking novel antagonists of the neurokinin-2 receptor.

One mixture (**1**) from a particular subset of 10000 compounds (100 mixtures of 100 compounds each) demonstrated potent affinity for the NK-2 receptor (Figure 4). We already knew that this 10000 component library was based on 2-(phenyl)-piperidinyloxyacetic acid and that the active mixture contained the adamantylacetyl group. As the mixture complexity was rela-

0947-6539/97/0312-1918 $ 17.50+.50/0 *Chem. Eur. J.* **1997**, *3*, No. 12

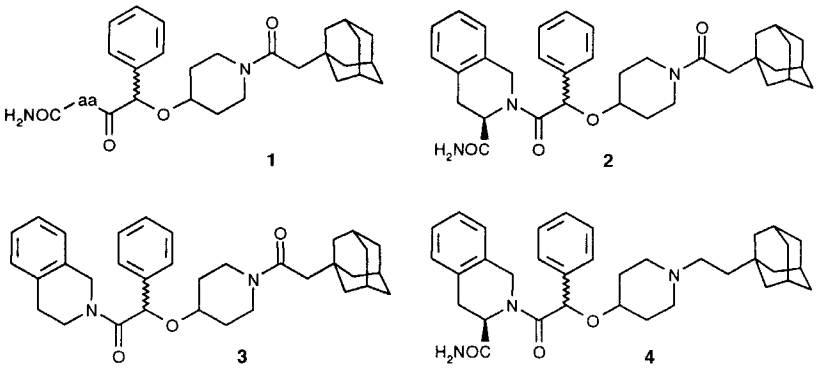

Figure 4. The most active mixture in the library tested against the neurokinin-2 receptor had the generic structure **1**. Deconvolution identified the simple amino acid on the carboxy terminus, and defined the most active ligand to have the structure **2**. Further orthodox follow-up synthesis investigated the analogues **3** and **4**.

tively low in this library, we used only one stage of deconvolution. All 100 individual compounds within the mixture were resynthesised and assayed, and the identity of the preferred "simple" amino carboxamide determined. From this deconvolution, the diastereomeric mixture **2** was identified as a potent NK-2 receptor binder ($IC_{50} = 59$ nM).

This lead compound has been followed up through both orthodox and combinatorial chemistry. For example, orthodox synthesis was used to separate the diastereoisomers and to prepare the derivatives lacking the carboxamide (**3**, $IC_{50} = 100$ nM) and the amide carbonyl (**4**, $IC_{50} = 40$ nM). Moreover, rapid parallel synthesis was used to replace the adamantylacetyl group with other acyl groups. Thus, both techniques have a place in drug discovery, their use being dictated by the type of chemistry being attempted.

The mix-and-split approach to library synthesis is clearly a successful method for the discovery of novel lead compounds. However, the necessity to test mixtures does present several problems. As revealed in the endothelin library experiment, it is possible to be misled by false positives. In particular, if a mixture is composed of a number of weakly active compounds, the measured biological activity is indistinguishable from a mixture containing one potent compound amongst inactive partners. Furthermore, the only way to determine whether the mixture contains a real active lead is to resynthesise the constituent compounds and resubmit them for test. This procedure, although relatively straightforward, as the chemist retains samples of key resin-bound intermediates and has the synthetic methodology ready to hand, still relies on additional chemical synthesis. The deconvolution process would be much easier if the constituent single compounds were already available to screen without needing to resort to resynthesis.

So far, chemists have been unable to synthesise libraries easily where all the compounds are available in quantity as separate components. It is possible to synthesise individual compounds in parallel using 96-well microtitre plates or commercially available automated synthesisers, but the throughput is limited to around one hundred compounds in each run. To overcome these limitations, we have developed a new technique whereby hundreds or even thousands of individual compounds can be prepared using a modification of the mix-and-split synthesis.

The new methodology shares with mix-and-split the advantage that the same chemical transformation can be applied to many substrates in the same reaction mixture, but the solid phase comes in a form that allows ready separation of the products. Furthermore the solid phase is in a format that allows the synthetic history and thus the attached chemical structure to be decoded by examination of a printed code sequence. Libraries are made on a new laminar form of the solid-phase material, which consists of two woven sheets of inert polypropylene between which are sandwiched quantities of resin beads. The sheets are fused together in such a way that the beads are fixed immobile even when the sheets are put through chemical transformations or divided.

The process was exemplified by the synthesis of a trial 27-component tripeptide library, but it is applicable to larger numbers of compounds and a variety of structural types. The synthesis commenced with the supply of three sheets of the laminar material approximately 6 cm by 6 cm in size, containing 1% cross-linker polystyrene resin (approximately 1.3 mg cm^{-2}) functionalised with an aminomethyl residue. Each sheet was printed with a three-by-three grid enclosing three-letter code sequences unique to the positions on the sheets. The printing was done prior to the chemistry by using a UV-curable ink insoluble in commonly used organic solvents. In this example, the first synthetic step was the derivatisation of each sheet with the activated ester of the carboxymethylphenyl ester of a unique Fmoc-protected amino acid (Figure 5). This step derivatised the aminomethyl resin with an amino acid, attached through a linker that could be cleaved under acidic conditions at a later point in the synthesis.

Following complete solid-phase derivatisation, the sheets were washed, dried and then treated with piperidine to remove the Fmoc protecting group. Following reassembly of the three

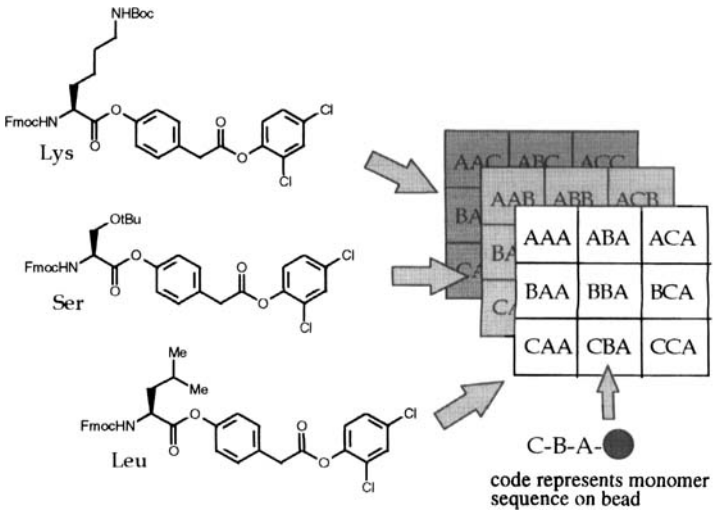

Figure 5. In the synthesis of the 27-component combinatorial library on laminar solid phase, the first step was the derivatisation of the individual sheets with the first amino acid derivative. The code sequence printed on the laminar sheets defined the synthetic history and thus the compound attached to each piece of solid phase at the end of the synthesis.

  0947-6539/97/0312-1919 $ 17.50+.50/0

sheets into a stack, they were cut along the gridlines to give three sets of columns. Each set was individually treated with an activated, Fmoc-amino acid to generate a unique dipeptide on each strip of laminar material (Figure 6).

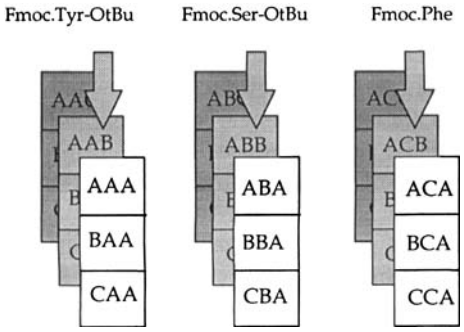

Figure 6. Following the vertical cutting of the sheets, each stack of columns was individually derivatised with the second amino acid, represented by the middle letter of the printed code.

Following Fmoc removal, the laminar pieces were washed, dried and reassembled into the original stacks to facilitate cutting into three rows. The pieces from each row were combined and finally treated with individual activated amino acids to generate the final tripeptides (Figure 7).

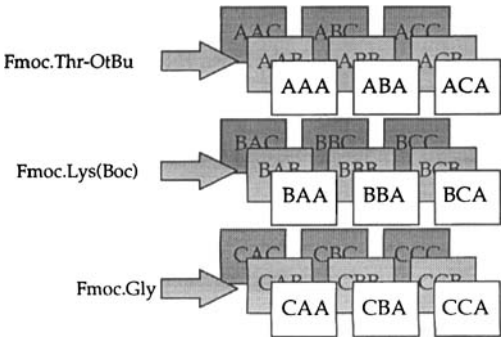

Figure 7. The laminar stack was reassembled, and cut horizontally. The last coupling reaction grouped the rows of laminar pieces from the stack and attached the third amino acid represented by the first letter in the code sequence.

Following the synthesis, the final library consisted of 27 smaller pieces of laminar solid phase, each 2 cm by 2 cm, containing unique tripeptide sequences encoded by the three letter sequence printed on each piece. To demonstrate that we had indeed made the compounds, cleavage with trifluoroacetic acid gave the final deprotected peptides in solution, and mass spectrometry gave the expected accurate molecular ion weights (Figure 8 with some example mass spectral data).

Thus, we have demonstrated that with the use of a novel laminar form of solid phase, it is possible to generate libraries of

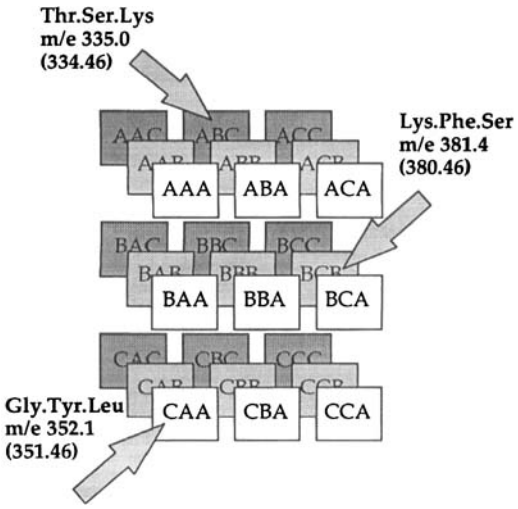

Figure 8. The compounds in the 27-member tripeptide library were cleaved from the laminar sheets under acidic conditions. Mass spectrometry confirmed the presence of the expected molecular ions for every compound (the observed molecular masses for three examples are quoted above with the expected masses in parenthesis).

compounds by means of a modified mix-and-split synthetic procedure. We are now investigating the use of this laminar library approach for the synthesis of larger libraries of non-peptidic compounds to assist our drug discovery programmes.

Overall, combinatorial chemistry is increasingly being used for the rapid synthesis of individual high-quality compounds. The novel laminar methodology we have developed allows the rapid synthesis of large numbers of individual compounds in quantity, and in a form whereby their identity can be quickly ascertained by reference to the unique code sequence printed on the solid phase. This library approach is now being used in the synthesis of libraries for our drug discovery programmes.

Received: April 8, 1997 [C664]

[1] a) N. K. Terrett, M. Gardner, D. W. Gordon, R. J. Kobylecki, J. Steele, *Tetrahedron* **1995**, *51*, 8135; b) F. Balkenhohl, C. von dem Bussche-Hünnefeld, A. Lansky, C. Zechel, *Angew. Chem.* **1996**, *108*, 2436; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2288; c) *Peptide and Nonpeptide Libraries* (Ed.: G. Jung), VCH, Weinheim, **1996**.
[2] a) A. Furka, F. Sebestyen, M. Asgedom, G. Dibo, *Abstr. 14th Int. Congr. Biochem.*, Prague, Czechoslovakia, **1988**, *5*, 47; b) A. Furka, F. Sebestyen, M. Asgedom, G. Dibo, *Int. J. Peptide Protein Res.* **1991**, *37*, 487.
[3] R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* (London) **1991**, *354*, 84.
[4] K. S. Lam, S. E. Salmon, E. M. Hersch, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* (London) **1991**, *354*, 82.
[5] N. K. Terrett, D. Bojanic, D. Brown, P. J. Bungay, M. Gardner, D. W. Gordon, C. J. Mayers, J. Steele, *Bioorg. Med. Chem. Lett.* **1995**, *5*, 917.
[6] A. M. Doherty, *J. Med. Chem.* **1992**, *35*, 411.
[7] a) K. Hemmi, N. Fukami, M. Hashimoto, H. Tanaka, N. Kayakiri (Fujisawa), *World Patent* WO93/10144, **1991**. b) K. Sogabe, H. Nirei, M. Shoubo, A. Nomoto, S. Ao, Y. Notsu, T Ono, *J. Pharmacol. Exp. Ther.* **1993**, *264*, 1040.
[8] a) L. A. Thompson, J. A. Ellman, *Chem. Rev.* **1996**, *96*, 555. b) P. H. H. Hermkens, H. C. J. Ottenheijm, D. Rees. *Tetrahedron* **1996**, *52*, 4527. c) J. S. Früchtel, G. Jung, *Angew. Chem.* **1996**, *108*, 19; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 17.

0947-6539/97/0312-1920 $ 17.50+.50/0