Pergamon

Tetrahedron Vol. 51, No. 30, pp. 8135-8173, 1995

0040-4020/95 $29.00+0.00

0040-4020(95)00467-X

TETRAHEDRON REPORT NUMBER 377

Combinatorial Synthesis – The Design of Compound Libraries and their Application to Drug Discovery

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

Discovery Chemistry Department, Pfizer Central Research, Sandwich, Kent CT13 9NJ, United Kingdom

Contents

1. Introduction

The synthesis of compounds using combinatorial library methodology is beginning to make a significant impact on the drug discovery process. This fact is a consequence of the continual identification of new pharmacological targets, and the need for new ligands upon which a drug discovery programme can be based. This need has sparked exploration of techniques that have the potential to generate large numbers of organic compounds. Thus, whilst rational design has been a key objective within the pharmaceutical industry for many years, an 'empirical' approach to lead discovery that relies instead on screening large numbers of diverse drug-like compounds has been wholeheartedly embraced as another valuable tool for finding new drug molecules.

Combinatorial chemistry is a technique by which large numbers of structurally distinct molecules may be synthesised in a time and resource-effective manner, and then be submitted for pharmacological assay in a variety of forms. The key feature of combinatorial chemistry is that compound synthesis is designed such that a range of analogues can be produced using similar reaction conditions, either in the same reaction vessel, or individually in parallel using semi-automated synthesis. In this way the bench chemist can single-handedly prepare many hundreds or thousands of compounds in the time usually taken to prepare only a few by orthodox methodology.

This review will define the key characteristics of combinatorial chemistry, consider issues involved in synthesising and screening libraries, focus on the need to develop new solid-phase chemistry, and will highlight the aspects that make it especially appealing to the pharmaceutical industry. In particular, the review aims to give a comprehensive overview of library methodology, describing published examples of combinatorial synthesis applied to drug discovery.[1]

There are many diverse methodologies described under the 'combinatorial library' banner that fall outside the scope of this review. In particular, any methods that do not rely on organic synthesis have been excluded. Thus, for example, filamentous phage display libraries are not covered as the library comprises peptide sequences on the surface of a bacteriophage coat protein. Phage libraries have played a valuable role in the definition of key peptide 'epitopes' that can bind to antibodies[2] or displace biotin from streptavidin.[3]

2 The Application of Compound Libraries to Drug Discovery

2.1 The Process of Drug Discovery

Finding novel biologically active leads is a challenging process. Historically, the main source of biologically active compounds for use in drug discovery programmes has been natural products, isolated from plant, animal or fermentation sources, and it is unnecessary in this journal to repeat the past success stories of the discovery of ß-lactams, tetracyclines, avermectins or taxol. However, despite the broad range of natural product structural diversity, finding activity for a specific biological target is frequently difficult. Pharmaceutical companies have for many years used their extensive compound collections as the source of leads with novel biological activity, and automated methods for rapid screening have allowed the testing of thousands of compounds per week. A flaw in this approach is that the structural diversity of company compound collections is biased by the range of structures previously explored by that company. It is a moot point whether the frequent discovery of benzodiazepine leads is a consequence of their structural attributes, or whether their occurrence reflects the disproportionate number of these structures held in company vaults. The same argument could be applied to dihydropyridines or ß-lactams.

Thus, we have reached a situation where novel receptors and enzymes are increasingly being identified as therapeutic targets, but drug discoverers are restricted to screening natural product extracts or moribund compound files in order to find the glimmer of affinity that could launch new synthetic programmes. Into this frustrating situation the appeal of a compound library designed to meet the chemists' specification is self-evident.

Lead discovery from compound libraries is primarily an empirical process. Initially there is no attempt to design an active compound. Rather, it is the number and variety of structures that libraries can offer that are their attraction. Often, any structural preconceptions of what might, or should have, affinity for the target protein are usually ignored in favour of the serendipitous discovery of a novel lead. That is not to say however, that rational design has no place in combinatorial chemistry. Indeed, a knowledge of the receptor or enzyme structure, however limited, can be used to design a library of compounds of a specific structural type. At its most extreme, the knowledge of the structure of a lead compound can be used to design a library of directly related analogues. In this case combinatorial synthesis is being used for the process of *lead optimisation* rather than *lead discovery* (Figure 1).

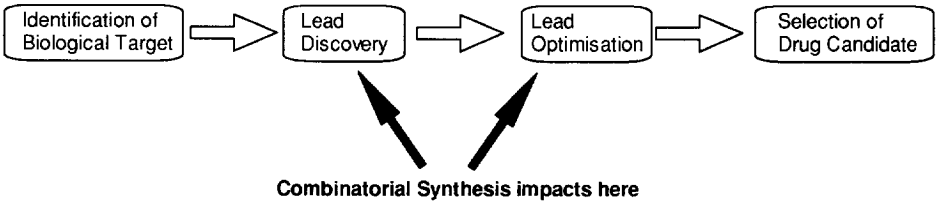


Figure 1. The drug discovery process and the impact of combinatorial libraries.

2.2 Combinatorial Chemistry

For over a hundred years the main task of the synthetic chemist has been the directed synthesis of one specific product using solution chemistry. In achieving this, chemists have usually worked on one reaction, on one substrate in one reaction vessel at a time – with each reaction targeting just one product. Combinatorial chemistry, has focused on technologies that have the *potential* to make large numbers of products, whether this is through preparing many single compounds in parallel, or many compounds simultaneously in mixtures. For example, if coupling monomer A with monomer B gives the product A-B, combinatorial synthesis can take a range of monomers A_1-A_n, and react those with B_1-B_n, and make any product combination (Figure 2). This process may be carried out using solution or solid-phase chemistry, but for reasons of reaction yield and purity, there has been a preferential focus on the use of solid-phase chemistry to give compound mixtures.

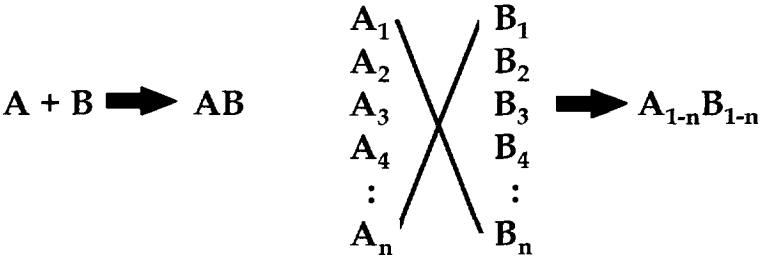


Figure 2. Contrasting orthodox and combinatorial synthesis, the former generally produces only one compound at a time, but combinatorial methods provide the potential to produce a range of products in parallel.

The universal advantage of combinatorial chemistry is that it is faster, and thus more efficient and cheaper than orthodox chemistry, and can give rise to hundreds, thousands or even millions of products. Consequently, it lends itself ideally to the search for novel biologically active compounds, especially where there is no prior information about the types of structures that are likely to be active. In this respect a parallel may be drawn with the biological assay of natural product extracts where many compounds are screened simultaneously to discover active components. Combinatorial libraries may be considered to be superior however, in that having been assembled synthetically, there is detailed information about the likely contents of the test mixtures, structural identification of active compounds should be accelerated, and lead optimisation through analogous structures should be simplified.

Performing many reactions in the same vessel imposes specific requirements. Solution chemistry is feasible, but only if conditions are chosen that minimise by-products (see Section 3). However, the rise of combinatorial chemistry has been primarily facilitated by the application of solid-phase synthesis (SPS) (Figure 3). If each substrate is linked to a solid phase, such as a resin bead, it is possible to generate products that are physically separate, and thus reagents and by-products not bound to the resin may be

removed simply by washing. Furthermore, the varying reactivity of a range of substrates may be readily accommodated by using several equivalents of reagents to drive every reaction to completion, and the excess may be removed by washing the solid phase with an inert, volatile solvent.

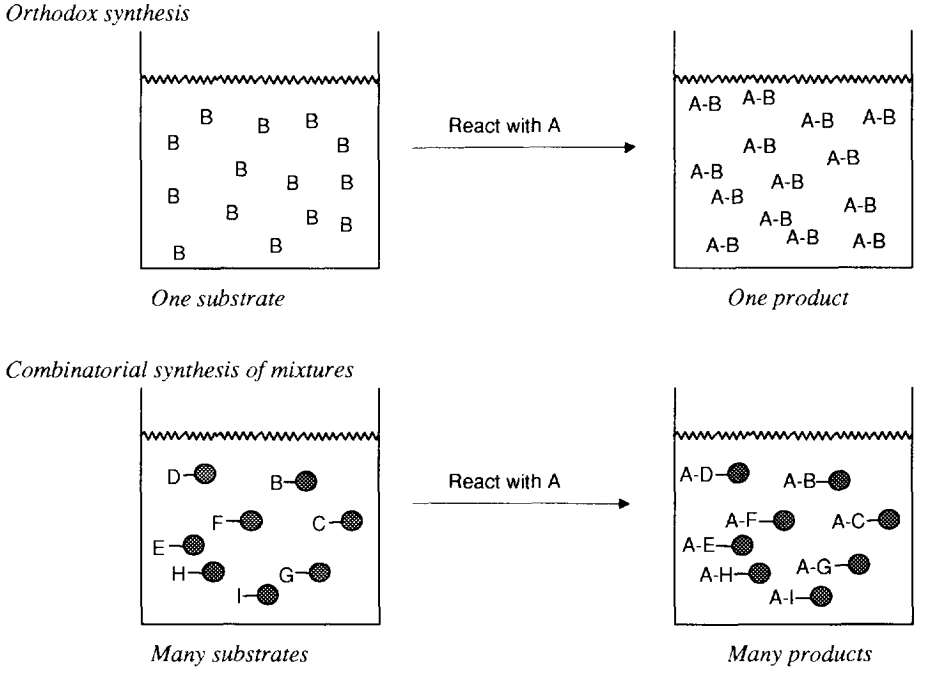


Figure 3. Synthesising mixtures of compounds has been enhanced by the use of solid-phase techniques.

2.3 Screening Issues

The sets of compounds produced by combinatorial chemistry are generally referred to as libraries, which, depending on how the solid-phase is handled, may be either mixtures or individual compounds. When individual compounds are prepared, the screening issues are generally the same as for orthodox drug discovery. If one considers the formation of a number of compounds on resin beads in a mixture, there are a range of options for testing these compounds in a biological assay:

(i) Test mixture in solution: All the compounds are cleaved from the beads and tested in solution (Figure 4). If the resin beads were intimately mixed, it is not possible to test the products separately, but rather as a mixture. If activity in a pharmacological screen is observed it is not possible to say which compound or compounds are active. In order to identify the most active component, it is necessary to resynthesise the compounds individually and thereby find the most potent. This iterative process of resynthesis and screening is one of the most simple and successful methods for identifying active compounds from libraries, and will be illustrated later by several examples of the successful discovery of pharmacologically active compounds.

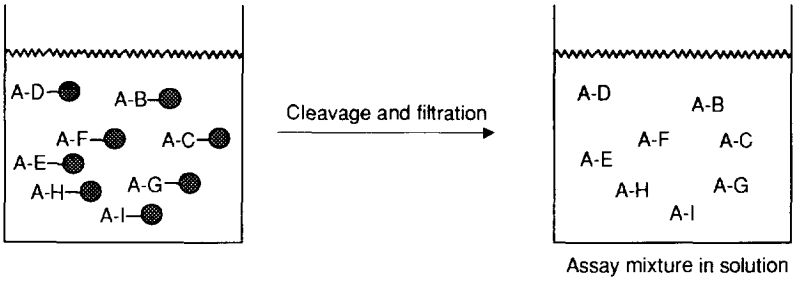


Figure 4. The products of combinatorial synthesis may be cleaved form the beads and tested in solution as a mixture of compounds.

(ii) Test individual compounds in solution: A second method is to separate the beads manually into individual wells and cleave the compounds from the solid-phase (Figure 5). These compounds can now be tested as individual entities. Note, however, that without any label on the beads, we have no idea of which compound is in which well, although we shall be able to ascertain the number of active compounds. Furthermore, the size of the bead will limit the amount of compound available for test, although standard bead sizes provide adequate material for at least one biological assay. The determination of structure is clearly more difficult and may depend on sensitive analytical methods, although for peptide libraries sequencing of the lead structure is possible. The advent of non-peptide libraries however, has encouraged the invention of novel *tagging* methods that allow the rapid identification of compound structure through the use of an encoding molecule, synthesised on the resin bead in parallel with the ligand molecule (see Section 4.3.2 below). Methods vary, but amongst the most useful of the tags are peptides (Selectide[4]), oligonucleotides (Affymax,[5] Lerner and Brenner[6]) and halogenated aromatic molecules that encode a 'binary' sequence deciphered by electron-capture capillary gas chromatography (Still, Pharmacopeia[7]). In these cases, the use of tagging methods has necessitated the design of coding chemistry that can be assembled on the bead without interfering with the chemistry of library synthesis.

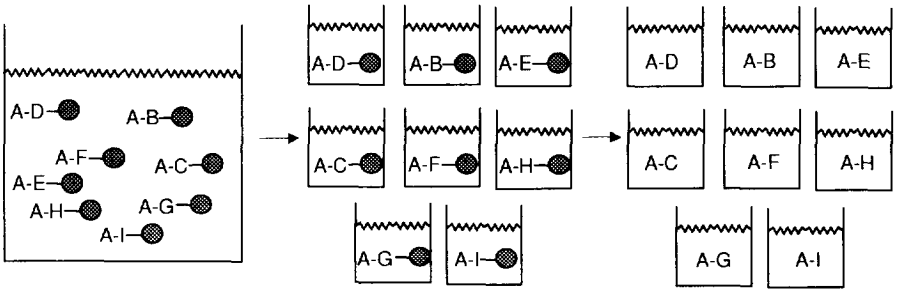


Figure 5. Using the handle provided by resin beads, individual compounds can be separated prior to cleavage from the solid-phase and screening.

(iii) Test compounds on the beads: A third method for screening is testing on the beads, using a colourimetric or fluorescent assay technique (Figure 6). If there are active compounds, the appropriate

beads can be selected by colour or fluorescence, 'picked' out by micromanipulation and the product structure, if a peptide, determined by sequencing on the bead. Non-peptide structures would need to be identified by one of the tagging methods highlighted above. Screening on the bead may be an inappropriate method for drug discovery, as the bead and linker present conformational restrictions that may prevent binding to the receptor. Furthermore, for pharmaceutical applications compounds will invariably need to act, and thus ideally need to be tested, in solution.

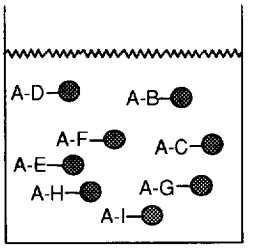


Figure 6. Compounds synthesised on solid-phase can be tested whilst still attached to the resin bead.

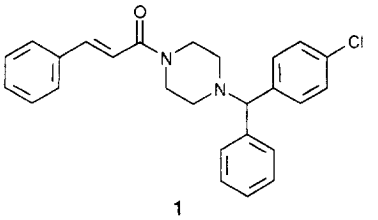
3 Combinatorial synthesis in solution

Despite the focus on the use of solid-phase techniques for the synthesis of combinatorial libraries, there have been a few examples where libraries have successfully been made and screened in solution. Indeed some groups have expressed a preference for solution libraries because there is no prior requirement to develop workable solid-phase coupling and linking techniques. Panlabs have recently disclosed an interest in making large numbers of compounds as individual components using parallel, reliable solution chemistry.[8] Reactions are pushed to completion by the use of excess quantities of the reactive reagent, and are isolated by solvent-solvent extraction. There is no further purification, and thus they prefer to describe these samples as 'reaction products'.

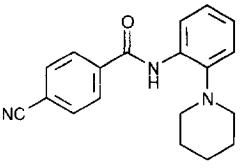
Two groups have recently disclosed solution libraries prepared in mixtures. In each case the groups from Glaxo[9] and Pirrung[10] have synthesised dimeric compounds using amide, ester or carbamate bond-forming reactions. Every library compound was prepared twice in mixtures of different composition. Testing all of these mixtures allows identification of likely active compounds without the need to resynthesise every compound in an active mixture.

In the Glaxo example 40 acid chlorides were reacted with 40 amines or alcohols to give amides or esters respectively in two sets. In the first set, each acid chloride (A) was reacted with a stoichiometric amount of an equimolar mixture of all 40 nucleophiles ($N_{1\text{-}40}$). In the second set each amine or alcohol (N) was reacted with an equimolar mixture of the acid chlorides ($A_{1\text{-}40}$). The 80 mixtures of 40 components each were screened against a wide variety of pharmacological targets, and a positive result from any sample identified

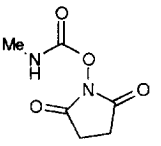
half of the structure of a likely active dimeric compound. Weak leads against the neurokinin-3 receptor (**1**, 60μM) and matrix metalloproteinase-1 (**2**, 55μM) were detected.



1



2



3

Pirrung's group have used an identical approach for the solution synthesis of a library of 54 carbamates from nine alcohols and six isocyanates. One library was constructed from each alcohol reacting with an equimolar mixture of isocyanates, and the second from each isocyanate reacting with an equimolar mixture of alcohols. The mixtures were subsequently tested against electric eel acetyl cholinesterase, and the results were used as 'indices' to the rows and columns of a two-dimensional matrix reflecting the activities of individual carbamates within the library. The most active compound (**3**) in the library was correctly identified from the structures of the most potent mixture from each sub-library.

4 Combinatorial synthesis on solid-phase

4.1 Solid phases employed for library synthesis

The techniques for solid-phase synthesis (SPS) are based extensively on the pioneering work of Merrifield,[11] who was the first to utilise substituted resins as the solid phase for the synthesis of peptides. Solid-phase synthesis naturally lends itself to the production of peptides, because of the limited range of synthetic transformations that are required for synthesis, and each of the key reactions has been optimised to allow the production of peptides of sizeable length in high overall yield.

The use of solid supports for organic synthesis relies on three interconnected requirements:

a) a cross-linked, insoluble, polymeric material that is inert to the conditions of synthesis,
b) some means of linking the substrate to this solid phase that permits selective cleavage of some or all of the product from the solid support during synthesis for analysis of the extent of reaction(s), and ultimately to give the final product of interest,
c) a chemical protection strategy to allow selective orthogonal protection and deprotection of reactive groups in the monomers (Figure 7).

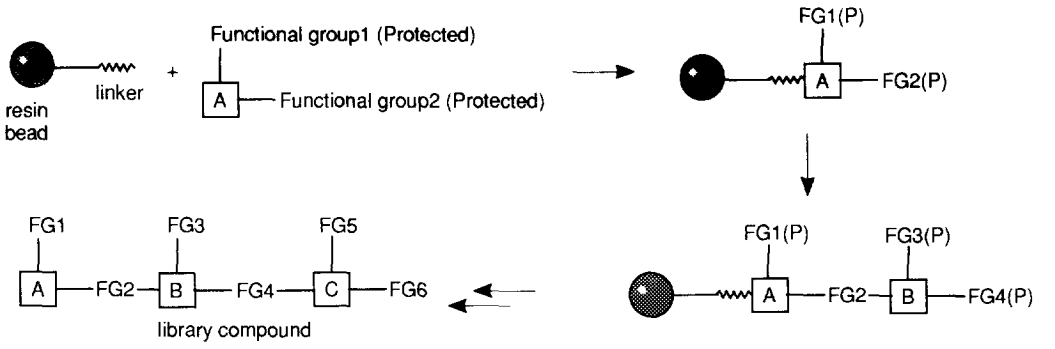


Figure 7. Essential to the success of solid-phase synthesis is judicious choice of synthetically compatible solid-phase, linker chemistry and protecting group strategy. In this example, functional groups (FG1, 3, 5 and 6) on the monomers A, B and C are protected until the end of the synthesis. Reactive groups on the monomers (FG2 and 4) involved in the coupling chemistry are also protected until required.

4.1.1 Solid Supports

The earliest forms of resin used were partially crosslinked polystyrene beads (the styrene is crosslinked with 1% divinylbenzene to give mechanical strength and insolubility whilst still permitting the flexibility apparent during solvent swelling) in a wide variety of sizes, prepared by light- or radical-catalysed polymerisation in an aqueous/organic mixture. Polymerisation takes place in micro droplets giving beads of approximately spherical shape. The consistency of size is ensured by sieving.

The earliest form of polystyrene resin (Merrifield resin) used for peptide synthesis was derivatised with a chloromethyl group to which amino acids could be coupled by nucleophilic displacement (Figure 8). The resulting ester bond was stable to the conditions of peptide synthesis and was cleaved to give carboxylic acid products under vigorous acidic conditions (hydrogen fluoride).

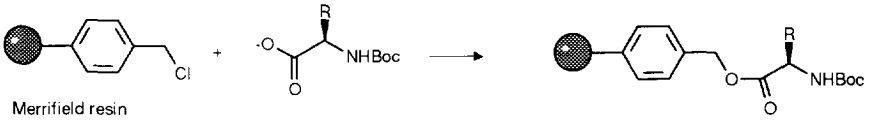


Figure 8. Loading of Merrifield chloromethyl resin with N-protected amino acids.

Many alternative linking chains on polystyrene have been described that permit the release of carboxylic acids under more mild conditions (see section 4.1.2 below). These polystyrene resins were used extensively for many years, but there was a growing realisation that the nature of the local environment around the growing molecular chain had a significant effect on the rate and extent of reaction. Polystyrene is completely hydrophobic in nature, whereas the growing peptide chain is much more hydrophilic, and this difference induces a chain-folding effect in which the peptide satisfies its own hydrogen-bond requirements rather than being solvated. As this can severely limit synthetic access to the exposed end of a growing chain, alternative

conditions have been investigated to obviate these effects. Different solid-phases such as Sheppard's polyamide resin,[12] have been developed as these polymers are hydrophilic like the growing peptide chain itself, and both can be readily solvated by dipolar, aprotic solvents (e.g. DMF, or N-methyl pyrrolidinone).

However, much of the focus of combinatorial libraries is on the synthesis of molecules other than peptides. For this purpose, there has been much use of Tentagel resin[13] (Rapp Polymere Gmbh) which consists of about 80% polyethylene glycol grafted to cross-linked polystyrene. It is generally considered that the reaction milieu within this resin is more closely related to ether and tetrahydrofuran, and consequently it has the potential for compatibility with the large range of reactions that are currently being investigated for compound library synthesis.[14] For example, work on the synthesis of β-turn mimics on solid phase could only be completed on Tentagel, as polystyrene beads did not permit solvation by aqueous solvent co-mixtures.[15]

Resins and other solid materials used for SPS have a wide variety of physical shapes. Although the most common is the spherical bead, a range of polyacrylic-grafted polyethylene extrusions called '*pins*' have been prepared with shapes of maximal surface area designed to optimise the product capacity.[16] The nature of the solid phase has a profound effect on the speed and nature of the reactions taking place in it, or on its surface. Many of the problems originally encountered with the use of this form of solid material have now been overcome by (i) increasing sizes and optimising shapes, and hence the amount of product available is much increased, and (ii) the use of selectively cleavable linker groups to permit the testing both on and off the 'pins'.[17]

Other materials have been used for solid-phase synthesis including cellulose in the form of 'Perloza' beads,[18] paper,[19] and cotton.[20] Advantages have been attributed to synthesis on paper materials, and it may well be that the more protic environment causes less of the chain-folding seen with hydrophobic materials and permits 'difficult' sequences to be prepared with greater facility. Functionalisation of glass surfaces and subsequent use in SPS has also been possible, especially for oligonucleotide and light-directed spatially-addressable parallel synthesis (see Section 4.3.1 below).

4.1.2 Linker groups

The group that joins the substrate to the resin bead is an essential part of solid-phase synthesis and thus compound library technology, as this group dictates both the method used for cleavage from the solid-phase and the terminal functionality that is revealed. With regard to combinatorial library design, using the same linker for an entire library will dictate that every compound will have the same terminal functional group, although there is no reason to stop the use of several linkers in parallel yielding different functionality in the same library.

As a consequence of the legacy of solid-phase peptide synthesis, the release of carboxylic acids and carboxamides is well served by a range of linking groups.[21] Whilst this is not a problem for peptide libraries,

as all compounds will contain a carboxylic acid or a carboxamide, there has been a pressing need for a greater range of functionality to be revealed by the cleavage step.

The preferred methods for the formation of carboxylic acids in combinatorial libraries is through the use of 'Wang' resin (Figure 9). Substrates are bound to this benzyl alcohol linker through formation of an ester, stable to most conditions except strong protic acid. This is adequate for peptide and some non-peptide synthesis, especially if a base-labile protecting group strategy is chosen. For example, the corresponding amino acid protection strategy employed is to protect the amine with fluorenylmethyloxycarbonyl (Fmoc) which can be removed by base (usually piperidine).[22] Conditions for the cleavage of products at the end of the synthesis are generally strong protic acid (95% down to 1% TFA).[23] This step also permits cleavage of *tert*-butyl protecting groups on side-chain functionality. A preferred route to primary amides is the use of Rink resin, which like Wang requires TFA for cleavage (Figure 10).

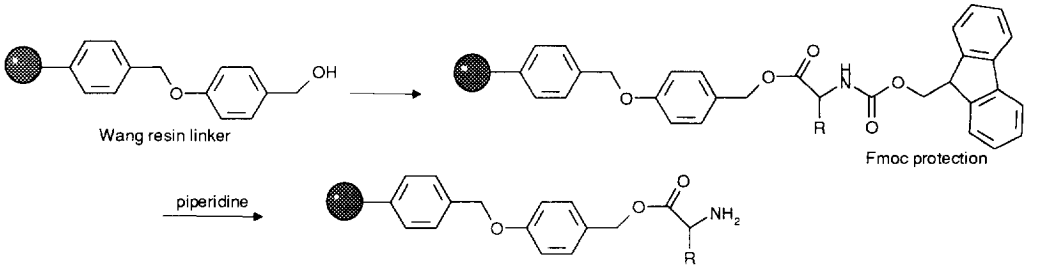


Figure 9. The use of Wang resin to couple carboxylic acids.

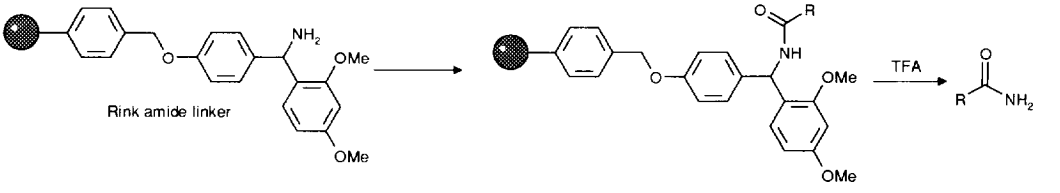


Figure 10. The use of Rink resin to produce carboxamides following TFA-catalysed cleavage.

Many other linking groups have been designed to permit the release of carboxylic acids through light-induced cleavage,[24] palladium catalysed cleavage of an allyl ester,[25] or cleavage of silicon substituted esters.[26]

Other functionality has been linked to and released from solid-phase. Ellman has described[27] a useful method for the coupling of hydroxyl groups to solid-phase through a dihydropyran-functionalised resin (Figure 11). Primary and secondary alcohols may be linked to the resin and cleaved in good yield using mildly acidic conditions.

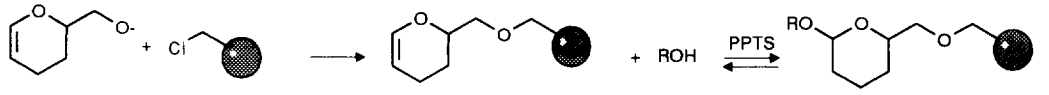


Figure 11. The linking and cleavage of alcohols from a dihydropyran-derivatised resin.

4.2 The Synthesis of a Combinatorial Library

Combinatorial synthesis on solid-phase, can generate very large numbers of products, using a method described as *mix and split* synthesis. This technique was pioneered by Furka[28] and has been enthusiastically exploited by many others since its first disclosure. For example, Houghten[29] has used mix and split on a macro scale in a 'tea bag' approach[30] for the generation of large libraries of peptides.

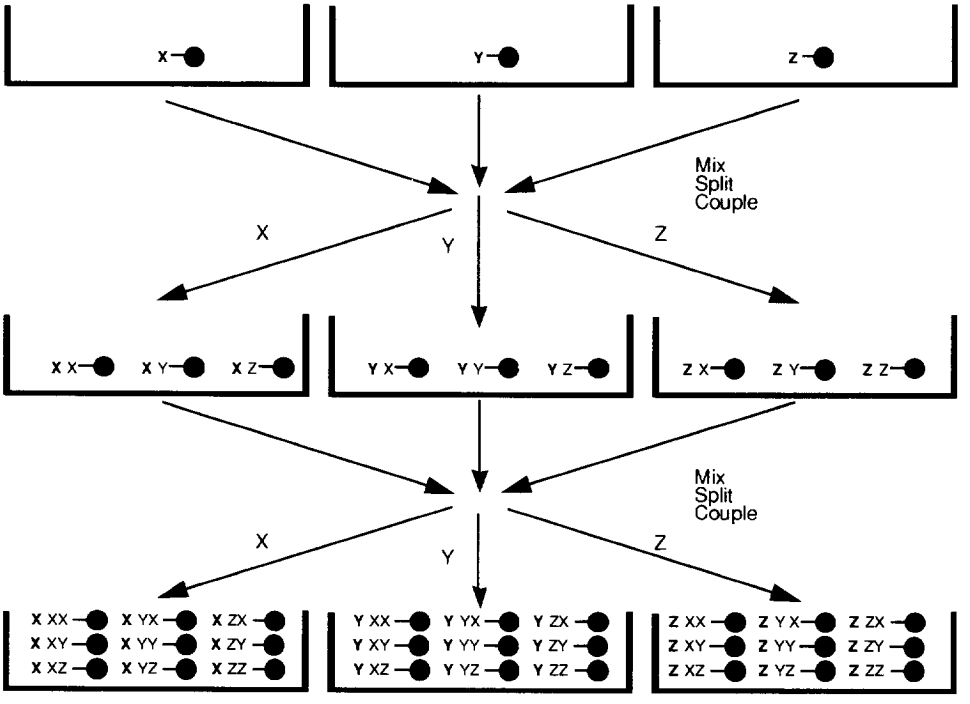


Figure 12. The application of the mix and split library procedure for the solid-phase synthesis of a 27-component trimer library.

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 reacted 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 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.

We illustrate the process schematically for the simple example of a 3 x 3 x 3 library, which gives all 27 possible combinations of trimeric products (Figure 12). X, Y and Z could be amino acids, in which case the final products would be tripeptides, but more generally they could be any type of monomeric unit or chemical precursor. It can be seen that the mix and split procedure finally gives three mixtures each consisting of nine compounds each, and as mentioned in section 2.3, there are several ways of progressing these compounds to biological screening. Although the compounds can be tested whilst still attached to the bead, a favoured method is to test the compounds as a mixture following cleavage from the solid phase. Activity in any given mixture reveals the partial structure of active compounds within the library, as the residue coupled last (usually the N-terminal residue) is unique to each mixture. Identification of the most active compound relies on deconvoluting the active mixtures in the library through further synthesis and screening. Whether the most active mixture actually contains the most potent compound in the library has been a focus of discussion for some time, but an elegant theoretical study by scientists from Isis has revealed that for an oligonucleotide library at least, the iterative process is generally successful.[31]

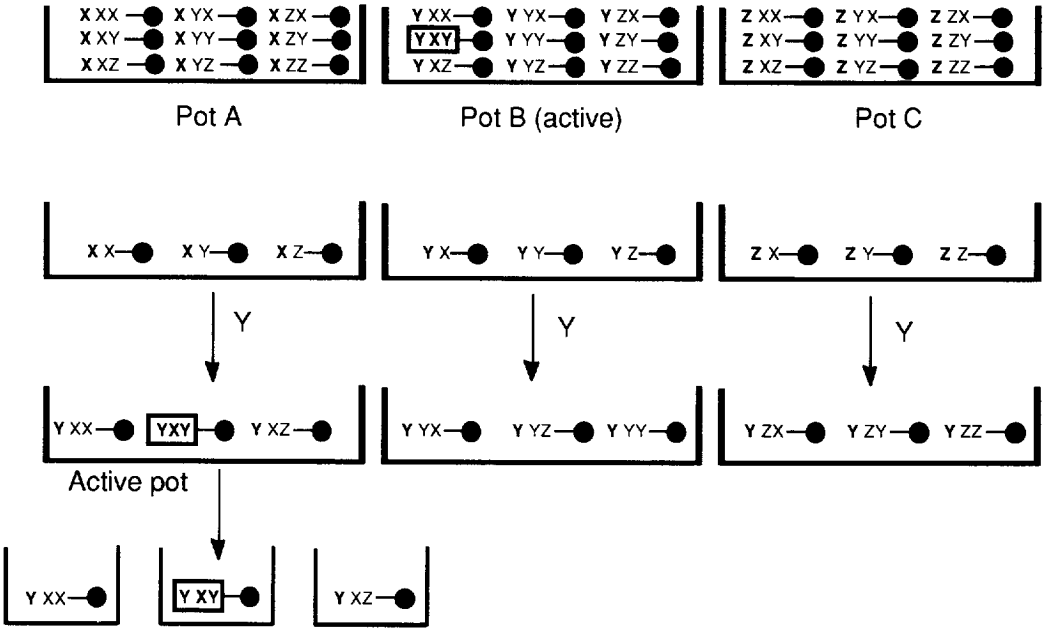


Figure 13. The identification of the most active sequence (YXY in this case) in a library through iterative resynthesis and rescreening.

In the example where the active structure is YXY, the mixture with Y at the terminal position will appear as the most active (Figure 13). Having retained samples of the intermediate dimers on resin (so-called

'recursive' deconvolution[32]) addition of Y to each of the three mixtures will give all nine compounds with Y at the terminal position, and the second position defined by the mixture. The most active mixture here defines the middle position of the most active trimer to be residue X. Finally, the three individual compounds can be independently resynthesised and tested to reveal both the most potent compound and also some structure-activity relationship data.

The ability to test compound libraries prepared on solid-phase either in solution or on the bead is exemplified by seminal publications from Houghten and Lam. Houghten and co-workers exploited the mix and split approach to prepare a library of just over 34 million N-acetylated hexapeptides.[33] Unlike the trimer example above, each final mixture had two N-terminal residues defined, and thus 324 mixtures (that is 18 x 18 as only 18 amino acids were used) were prepared on resin beads, and were cleaved before assay for binding to a monoclonal antibody. Having identified preferred N-terminal dipeptide sequences, iterative resynthesis and screening eventually revealed a number of peptides with sub-micromolar binding affinity.

In contrast, Lam *et al.*[34] tested a family of peptides whilst still attached to the resin bead solid-phase. Nineteen amino acids were incorporated into pentapeptides to generate a library of almost two and a half million compounds. By using a colourimetric assay, beads bearing peptide sequences that bound tightly to the protein streptavidin or to an antibody raised against β-endorphin were revealed by visual inspection. Bead picking using micromanipulation isolated the beads, and the active peptide structures were determined by microsequencing.

A modification of this method has allowed screening of such libraries in solution.[35] Linkers have been devised that allow several copies of the library compounds to be released sequentially. Using this method it is possible to identify an active mixture using a solution assay, and then return to the beads that produced these compounds, and redistribute them into smaller mixtures for retest. By repeatedly reducing the mixture size, ultimately to single compounds, the bead containing the most potent sequence may be identified and the peptide product sequenced (Figure 14).

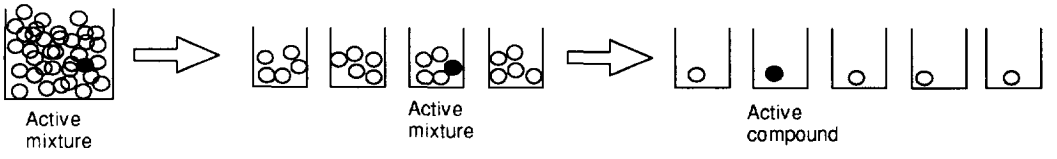


Figure 14. The identification of a single bead bearing an active component through iterative screening of successively smaller mixtures.

Houghten has also used a variation of the mix and split technique for the production of *positional scanning* libraries.[36] Positional scanning is a technique by which the same library compounds are prepared several times, and in each library a different residue in the sequence is held constant. For example, a hexapeptide library of 34 million (18^6) compounds has been produced in six sets of mixtures, each mixture containing

1,889,568 (18^5) peptides. The first set of mixtures will have the N-terminal amino acid held constant within a mixture (O below) but will have a random mixture of the 18 amino acids in each of the other positions (X below). The identity of the fixed residue in each mixture is known as a consequence of the method of assembling the library and thus the most active mixture following screening will define the preferred N-terminal residue. The second library holds the second residue constant in each of the mixtures, and from screening these mixtures a preferred second position amino acid may be identified. The same approach is used to construct the remaining four libraries. Testing all the mixtures will reveal the preferred residues at each of the six residue positions (Figure 15).

$O_1XXXXX\text{-}NH_2$
$XO_2XXXX\text{-}NH_2$
$XXO_3XXX\text{-}NH_2$
$XXXO_4XX\text{-}NH_2$
$XXXXO_5X\text{-}NH_2$
$XXXXXO_6\text{-}NH_2$

Figure 15. The six hexapeptide positional scanning libraries. 'O' is a residue that is known within the mixture as one of the monomers used to synthesise the library. 'X' is an equimolar mixture of all monomers used.

Two main issues need to be addressed with this library approach. Firstly, apart from the final monomer to be added to the mixture, the routine mix and split synthetic approach does not easily permit other positions to be held constant in a mixture. Houghten overcomes this problem by reacting the resin-bound substrate with a *mixture* of activated amino acids. Although successful in finding active compounds, this approach could potentially fail to make all of the target compounds, as some amino acids may acylate more readily than others, and the outcome could be unequal quantities, or even omissions of some products. This variable reactivity has been addressed by Houghten by using concentrations of amino acids that are dictated by their relative reactivity.

The second issue with positional scanning is common to testing all libraries in mixtures. The most active mixture does not necessarily contain the most potent compound in the library (an issue recently addressed for both iterative and positional scanning libraries by Isis[31]). Houghten addresses this problem by identifying not just one, but up to four amino acids for each position that may be in the structure of the most potent peptide. Having identified a number of preferred residues in each position (O1 = Y; O2 = G; O3 = G, or F; O4 = F; O5 = F, Y, M or L; O6 = F, Y or R), twenty-four individual peptides were synthesised to discover the structure of potent μ-opioid receptor ligands (the most potent being YGGFMY). In fact the hexamer based on the most preferred single residue in each position was only weakly active.

The libraries prepared by Houghten's group have demonstrated the power of combinatorial synthesis for generating large numbers of compounds, and successfully identifying potent peptides. However, there remains a pressing need to prepare non-peptide libraries. Houghten has shown[37] that it is possible chemically

to transform the peptide libraries whilst they are still attached to the solid-phase to give mixtures of novel non-peptidic compounds. For example, a 10-fold excess of sodium hydride in dimethyl sulphoxide followed by a 30-fold excess of methyl iodide was successful in converting peptides to their fully methylated derivatives (Figure 16).

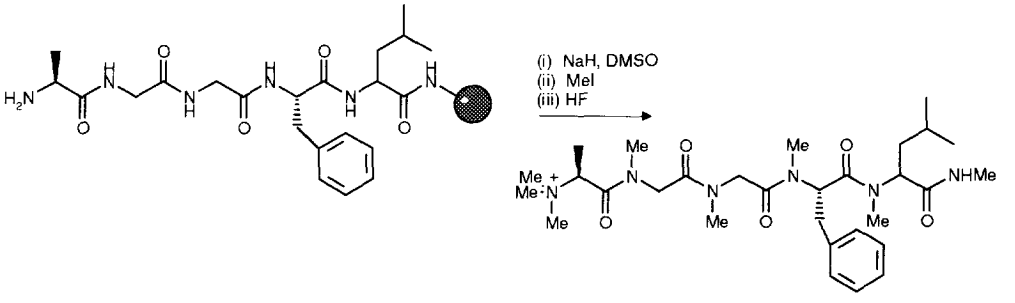


Figure 16. The derivatisation of peptides on solid-phase to give non-peptide products.

4.3 Determination of Product Structures

In the above examples the structural determination of active constituents in the library is achieved either by (i) a systematic iterative resynthesis and rescreening of specific mixtures and compounds, or (ii) microsequencing of a peptide sequence whilst still attached to a resin bead. However, many investigators may not wish to embark on the resynthesis of library components, and if the library constituents are not peptides, microsequencing can no longer be applied. Thus, a major area of investigation has been the development of other ways of determining the structures of active compounds within libraries. In particular, structures have been derived either by product location or by 'tagging'.

4.3.1 Structure of product by location

Libraries synthesised in a format in which the solid support is physically constrained (on paper, cotton, polypropylene pins or resin beads in tea-bags), have the structures of the compounds defined by position (usually supplemented by a label). An elegant extension of this is work by Affymax on light-directed, spatially addressable parallel chemical synthesis (also known as VLSIPS - very large scale immobilised polymer synthesis).[38]

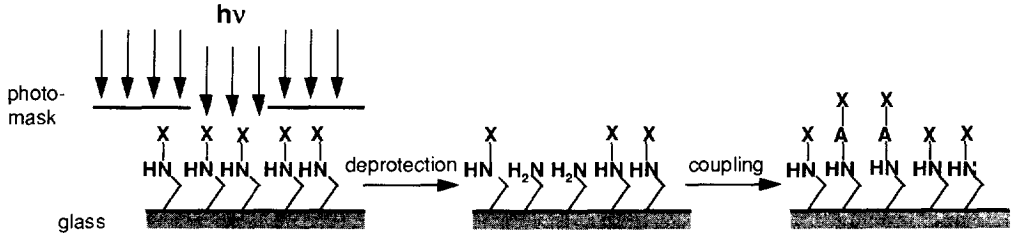


Figure 17. The use of light to deprotect solid-phase bound substrates is controlled by the use of specific masking.

A combination of photolithography and solid phase synthesis has been used for the generation of an array of peptides on a glass microscope slide. Initially all the free amino groups on the glass surface were protected with the photolabile NVOC (6-nitroveratryloxycarbonyl) group. A mask was used to control the irradiation of predefined regions of the surface to reveal free amino groups that could then be functionalised (Figure 17). Using a range of patterned masks during the synthetic procedure allowed the creation of a checkerboard array of 1024 different peptides in elements measuring 50μm by 50μm. Testing took place on the 'chip', and the structure of active(s) read by their position on the surface.

4.3.2 Structural identification by tagging

The resin bead mix and split method can be used to generate hundreds, thousands or even millions of different products. Although synthesis is rapid, the power of combinatorial libraries is only evident if structural information on active components may be easily obtained. The iterative resynthesis and rescreening offers a solution, but as it can be slow and requires a further dedication of synthetic and screening resource, there have been a number of new methods devised where information concerning the active compound may be carried on the bead in the form of a 'tag'. Three major methods of chemical tagging have been described, and all involve co-synthesis of a chemical tag or identifier on the same resin beads.

Selectide have described a method that allows the determination of structure for non-peptides or other non-sequenceable library products.[4] The method involves the simultaneous synthesis *on the same bead* of a test compound and a peptide that codes for this non-sequenceable compound (Figure 18). To permit this parallel synthesis, Selectide have devised a multiple linker that includes a lysine branch-point to attach both sequences.

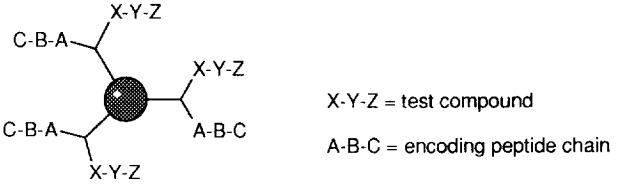


Figure 18. Resin bead with branched linkers that permit the synthesis of both a library component and an encoding peptide chain.

The non-peptide sequence is constructed by initially attaching an alpha-bromo carboxylic acid to a tryptophan residue on the linker. This was reacted with an amine or N-protected aminothiol to give either an amine or thio-ether, and the compounds were finally capped with an acylating group. Cleavage from the resin gave products with a C-terminal tryptophan carboxamide (Figure 19)

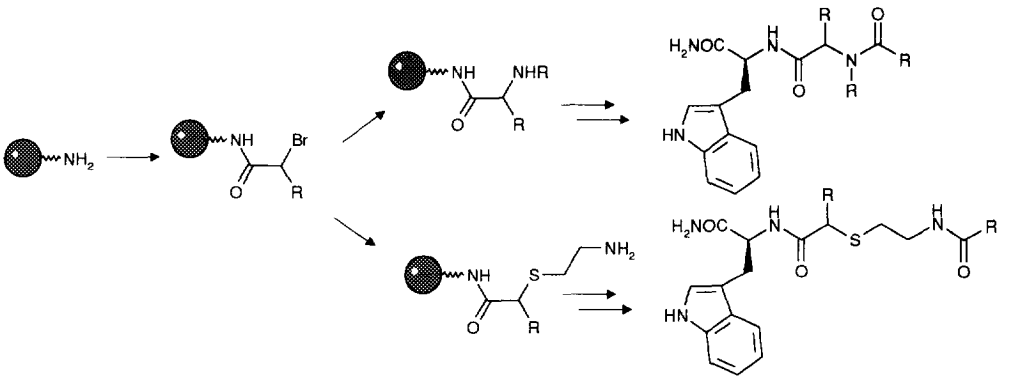


Figure 19. Selectide non-peptide compound sequences.

The encoding chain is a tripeptide, where each amino-acid represents one component in the non-sequenceable chain. Thus as each building block is attached to the growing compound there follows the coupling of an amino acid to the encoding arm of the linker (Figure 20). The library compound is released by cleavage of a safety-catch linker (SCAL)[39] that will only cleave under acidic conditions following reduction of two sulphoxide groups (using 20% $(EtO)_2P(S)SH$ in DMPU). The structure of any of the library compounds can be ascertained by Edman degradation of the peptide tag whilst attached to the bead.

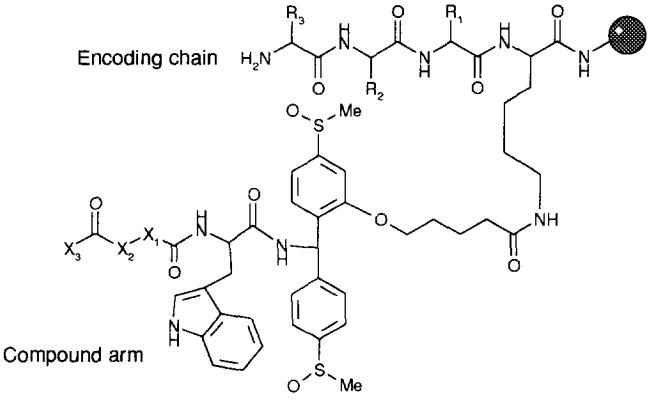


Figure 20. A branched linker for assembly of non-peptide compounds on a safety-catch linker and an encoding peptide chain.

In this example there is a one-to-one correlation between the amino-acids and the non-peptide components. Clearly this will only permit the use of 20 building blocks, but with the use of dipeptide sequences to encode for each building block, the number could be increased to 400.

A second method of tagging is by use of oligonucleotides to encode the structure of the library components.[6] Thus, a process of alternating steps is performed in which each chemical step in the synthesis of the library

members is succeeded by a parallel step to encode with a unique oligonucleotide sequence. The genetic code attached to beads bearing active compounds may then be amplified by replication, and the code read by sequencing the DNA. Brenner has demonstrated this approach by the synthesis of oligonucleotide-encoded peptides on derivatised controlled pore glass (CPG).[40] Several linkers are described that each depend on the use of serine branching residues to permit the synthesis of both the peptide and oligonucleotide chains (Figure 21). Orthogonal reaction conditions were established for the synthesis of each of the chains. Oligonucleotide synthesis using standard 2-cyanoethyl-protected nucleoside phosphoramidite chemistry did not affect the Fmoc-amino acid derivatives used in peptide synthesis. Likewise, the conditions used for removing Fmoc protection did not interfere with the standard protecting groups on the nucleotides. Although not used for the synthesis of a library, trial experiments demonstrated the preparation of standard peptide and oligonucleotide chains, the identities of which were confirmed by Edman degradation and PCR amplification respectively. In a binding experiment, a bead-attached peptide chain had affinity for the 3-E7 antibody despite the presence of the adjacent oligonucleotide.

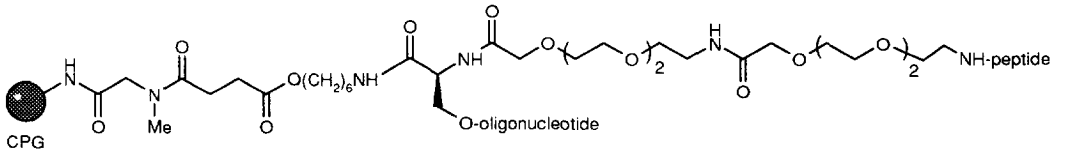


Figure 21. The Brenner branched linker for parallel synthesis of a peptide and an oligonucleotide encoding chain.

Oligonucleotide tagging has been used by scientists at Affymax for the deconvolution of a million heptapeptide library, and these have been screened against an antidynorphin B monoclonal antibody.[5] Using a cross-linked polystyrene bead derivatised with a 1,12-diaminododecane linker, the peptide was built on a threonine residue whilst the oligonucleotide tag was constructed on a 4-hydroxybutyric acid group (Figure 22). The library was assayed whilst the compounds were still attached to the beads, and those that bound the antibody were detected by acquired fluorescence. PCR amplification and sequencing of the oligonucleotide tag attached to the most brightly stained beads revealed the structure of the active peptides.

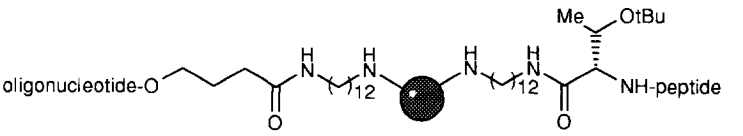


Figure 22. The Affymax branched linker for parallel synthesis of a peptide and an oligonucleotide encoding chain.

A third method of tagging uses a similar method of parallel synthesis, but in this instance the tag is one of a series of inert haloaromatic substituted alkanols.[7] During each step of a peptide library synthesis, tagging molecules are attached to the beads to encode for both the step number and the chemical building block used

in that step. Unlike the peptide or oligonucleotide tags described above, the tags are not built into a parallel chain, but instead are added onto the free amino terminus of the growing peptides (Figure 23). As a very small amount of the tagging molecule is used on each occasion, only about 0.5% of the growing peptide is terminated by each tagging step. Thus there is no need for a special branched linking group, and the total amount of tag represents only a few per cent of material on the bead. More recently diazomethane derivatives have been used to add tags directly to the resin backbone by the generation of acylcarbenes.[41] Screening of a library of 117,649 peptides was accomplished whilst still attached to the bead, and active sequences were picked by staining of the bead through a colourimetric assay. The identities of active peptides are determined by analysis of the tags after cleavage by UV irradiation of a labile linker, and 'read' by means of electron capture gas chromatography (ECGC). The tags are designed in such a way that the GC trace could be interpreted as a 'binary code' revealing both the amino acids used and their position in the synthesis; the presence or absence of each of the tag molecules defining precisely the amino acids used in each position of the hexameric sequence. The most potent peptides in this library had micromolar affinity for the anti-c-MYV monoclonal antibody.

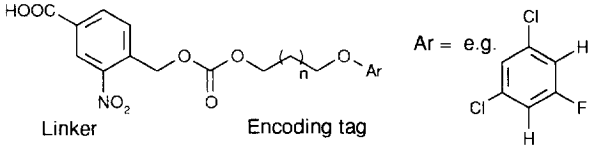


Figure 23. The reactive tagging molecules used by Pharmacopeia to encode for non-peptide sequences.

It can be seen that in each of the three cases above, the use of a tagging group allows the synthesis of any type of compound within the library. The tagging molecules can encode for any building block and any synthetic transformation. Furthermore, given the uncertainties of much synthetic chemistry, the tag may be looked upon as not so much encoding a specific compound structure, but encoding instead a synthetic procedure. Thus, even if the intended compound was not made but biological activity was detected, the tagging system facilitates a replication of the synthetic steps employed in producing the active compound, and thus aids structure determination.

5. Range and evolution of library solid phase chemistry

5.1 Peptides and Oligonucleotides

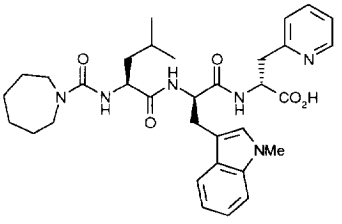
In the effort to identify novel diagnostics and therapeutics, the majority of combinatorial libraries prepared prior to 1994 have focused on oligonucleotide[42] and, particularly, peptide structures. Consequently, the polymer-supported organic chemistry required for their synthesis has evolved from the peptide-bond forming procedures originally devised by Merrifield.[11] Synthetic peptide combinatorial libraries using amide chemistry have been successfully applied, for example by Owens and co-workers at Lilly,[43] who have identified HIV protease inhibitors through the synthesis and screening of peptidic mixtures consisting of precursors including L- and D-α-amino acids, and statine. In parallel, Houghten and his group as described

in section 4.2 have used peptide libraries to identify a range of bioactive products including antigenic determinants,[33 44 45] antimicrobials,[29] enzyme inhibitors[46] and ligands for opioid receptors,[36 47] and have also described a combinatorial library of *cyclic* peptides[48] that offer reduced flexibility over analogous linear peptides as well as increasing the potential for high binding affinities and enzymatic stability.

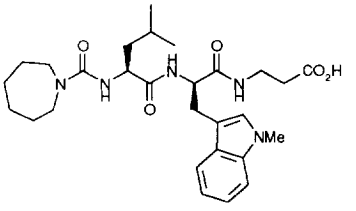
However, despite compound populations exceeding 10^9 in some libraries, the components are still peptides and as such suffer from major deficiencies as new leads - predominantly poor oral availability and facile metabolism due to extensive proteolysis. The emphasis of combinatorial library synthesis is now shifting away from this reliance on peptide bond formation toward the synthesis of so-called 'small organic' components that are more attractive as pharmaceutical leads. To achieve this, new solid-phase organic chemistry is appearing at an increasing rate - the primary advantage of the solid support being the ease of purification (even of mixtures) at each stage of a synthetic sequence. Furthermore, a key component of any library is the relative diversity of the constituents, and a corollary of this is that the syntheses used should permit of as wide a monomer base as possible. The remainder of Section 5 aims to highlight solid-phase chemistry that has been applied to library synthesis *and* new chemistry that has obvious or untapped application in library synthesis.

5.2 Peptoids and other amides

A first small step away from peptides and their inherent metabolic limitations is to exploit the potentially limitless supply of non-α-amino acids whilst still using amide bond coupling methods. This approach was used by our group[49] in the synthesis of a 31,000-component library of triamides prepared to identify potent endothelin antagonists and derive limited SAR around the known Fujisawa endothelin inhibitor, FR-139,317. A pool of 32 Fmoc-protected amino acids including only seven natural α-amino acids was assembled into 32 mixtures of 992 triamides using a standard solid-phase coupling methodology (TBTU-HOBT-DIPEA). Screening and iterative deconvolution identified FR-139,317 as the most active component and highlighted several potent analogues including (**4**).



FR-139,317



4

An alternative approach that retains the tried and tested solid-phase amide bond chemistry whilst avoiding many of the drawbacks of peptides has been devised by Zuckerman and colleagues at Chiron.[50] This consists of libraries of N-substituted oligomers such as poly-(N-substituted glycines) (NSGs). These so-called '*peptoids*' bear a close resemblance to the corresponding 'natural' peptides but are achiral and have retro amide groups and side-chains attached to N rather than the α-carbon atom (Figure 24). Peptoids have recently been shown[51] to have less susceptibility to proteolysis than native peptides and have already led to

the discovery[52] of nanomolar ligands for seven-transmembrane, G-protein-coupled receptors. These include CHIR 2279 (α_1-adrenergic K_i 5nM) and CHIR 4531 (μ-opiate K_i 6nM).

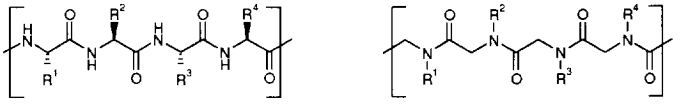
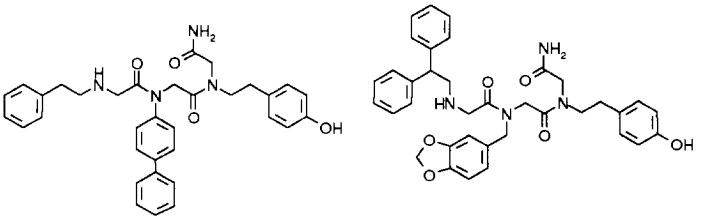


Figure 24. Comparison of the structures of peptides and peptoids.



Recent peptoid syntheses have utilised an automated cycle of Rink amide resin haloacetylation followed by displacement with a range of primary amines and further acylation of the new secondary amine (Figure 25).

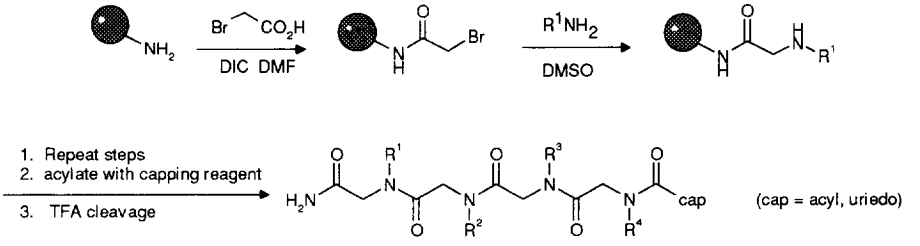


Figure 25. The synthesis of peptoids.

A library of oligocarbamates has been prepared from optically active aminocarbonates, using the Affymax VLSIPS chip technology (see section 4.3.1 above).[53] A total of 256 oligocarbamates that express every deletion combination of the sequence, $AcY^cF^cA^cS^cK^cI^cF^cL^c$ (where X^c is the carbamate equivalent of the amino acid X) were synthesised using photolabile protecting groups (NVOC) and a binary masking strategy (Figure 26). The library was screened against an antibody prepared using a conjugate of the sequence $AcY^cK^cF^cL^cG^c$ and several products were observed to bind.

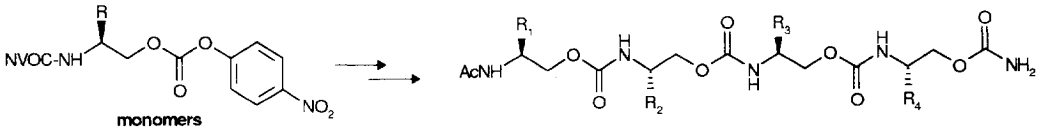


Figure 26. The preparation of oligocarbamate structures.

5.3 Heterocyclic synthesis

As a large number of drugs feature a heterocyclic component, methods are rapidly appearing for the assembly of heterocyclic compound libraries. In the first key example of solid-phase heterocyclic synthesis, Bunin and Ellman[54] illustrated a general and expedient route to 1,4-benzodiazepine derivatives (Figure 27) that is ideally suited to library design as three discrete monomeric components are combined to form each benzodiazepine. In the synthesis, a range of independently synthesised aminobenzophenones (**5**) is linked to Tentagel resin through a phenol or acid residue and acylated with a set of Fmoc-protected α-amino acids. Deprotection and acid-catalysed cyclisation gives representative benzodiazepines (**6**) on resin that are further functionalised by amide N-alkylation. Acid catalysed cleavage liberates the final benzodiazepine (**7**) in a relatively pure state for screening. The use of this chemistry in library synthesis is the subject of a US Patent application.[55]

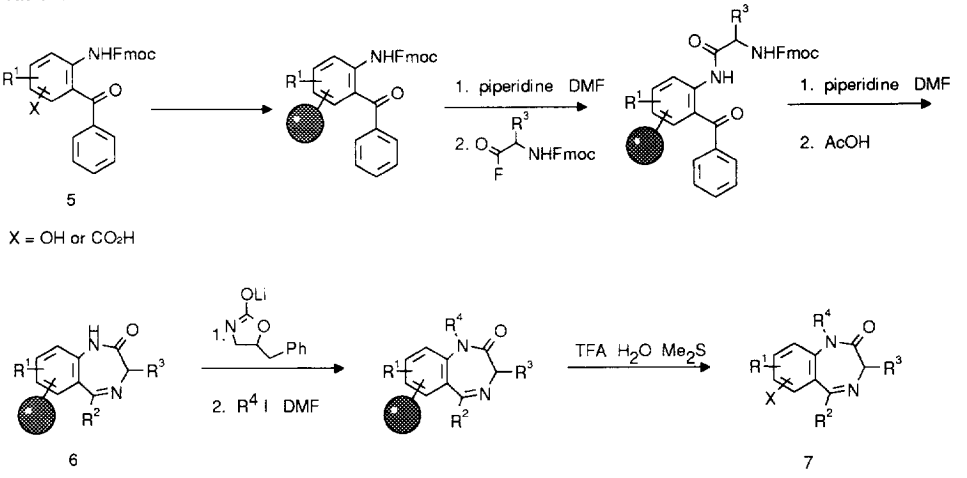


Figure 27. Ellman's solid-phase synthesis of benzodiazepines.

A further approach to benzodiazepine synthesis due to DeWitt *et al.* inverts the foregoing methodology by transimination of presynthesised aminobenzophenone imines (**8**) using commercially available α-amino acids on Wang resin (Figure 28). The resulting imine is cleaved and cyclised by TFA treatment to give (**9**).[56] The Parke-Davis chemists have used this chemistry for the parallel assembly of forty discrete benzodiazepines in their Diversomer™ pin apparatus. A similar approach has been devised by the same group for the synthesis of trisubstituted hydantoins (**10**), relying on the reaction of individual α-amino acids with isocyanates

followed by a protic acid hydrolysis. In the quoted example, 39 out of an expected 40 single hydantoins were isolated and characterised.

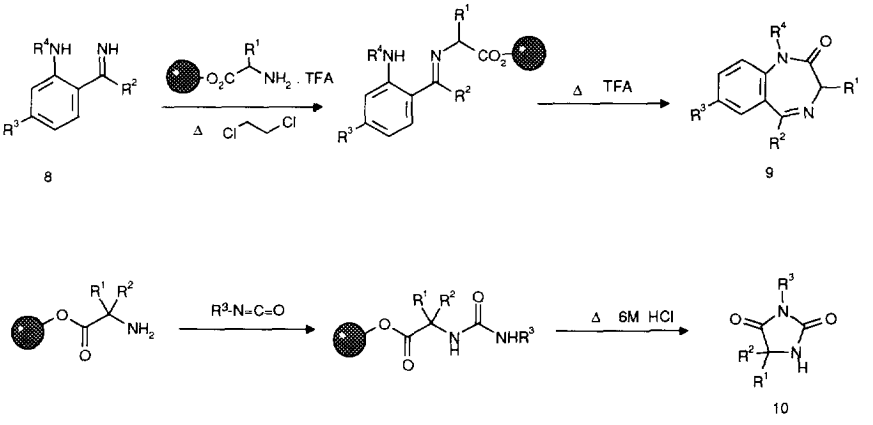


Figure 28.

Gordon and Steele[57] have recently reported the synthesis of a prototype library of 1,000 trisubstituted piperazinediones (**13**) (diketopiperazines, DKPs) by a sequence of three key steps: a novel solid-phase reductive alkylation, acylation by a second amino acid and cyclisation to the DKP products (Figure 29). Amino acids on Wang resin were reductively alkylated with alkyl or aryl aldehydes using sodium triacetoxyborohydride as the reductant, to give resin-bound secondary amines (**11**). The acylation of secondary benzylic amines is very difficult and required a double coupling procedure using PyBrOP to achieve good conversions to (**12**). TFA-mediated cleavage of (**12**) from the support did not cause spontaneous cyclisation and so a brief reflux in toluene effected closure to (**13**). Iterative screening and resynthesis of this library has identified several DKPs with significant biological activity including (**14**), which has a high affinity for the neurokinin-2 receptor (IC_{50}=313nM).[58]

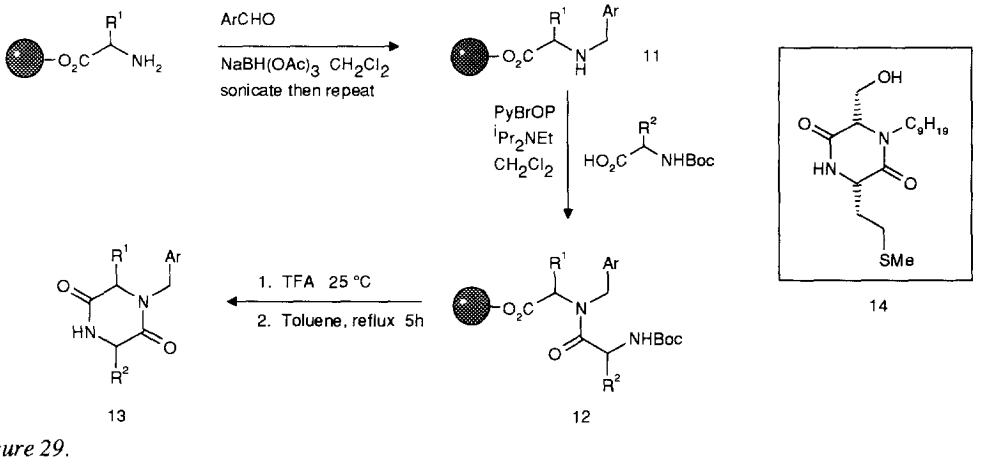


Figure 29.

A polymer-supported synthesis of 2,5-disubstituted tetrahydrofurans such as (**17**) has been described by Kurth and co-workers (Figure 30).[59] Dehydration of the nitroalkyl polystyrene resin (**15**) (available in two steps from Merrifield resin) by phenyl isocyanate generates a transient, resin-bound nitrile oxide that is readily intercepted by 1,5-hexadiene, giving (**16**). Oxidative cleavage and cyclisation generated the tetrahydrofuran (**17**) as a 1:2 mixture of diastereomers in 11-17% yield and returned a re-useable resin. In a related synthesis of γ-butyrolactones, Kurth has reported a diastereoselective alkylation of the resin-linked auxiliary (**18**) (Figure 30).[60] Iodolactonisation of the product (**19**), liberates the γ-butyrolactone (**20**) with 87%ee. The polymer-supported auxiliary can be recycled for further syntheses by simple filtration.

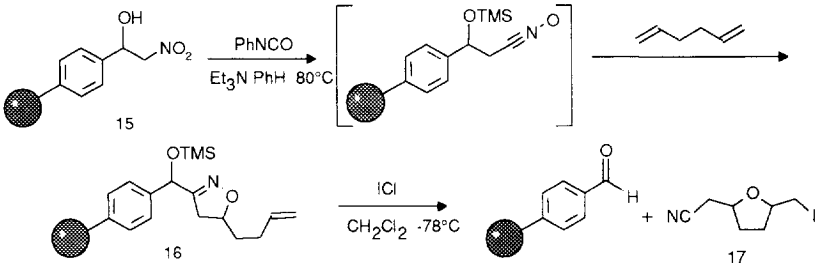


Figure 30.

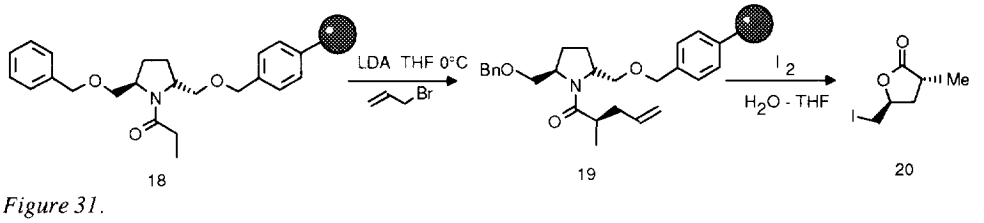


Figure 31.

A library of discrete, macrocyclic β-turn mimetics generically represented by (**24**) has been assembled by Virgilio and Ellman using three monomer sets: the ubiquitous α-amino acids, α-bromoacids and ω-aminodisulfides.[15] The key steps are as follows; reaction of the α-bromoacetyl Phe derivative (**21**) with an ω-aminodisulphide followed by amide coupling to an Fmoc-protected amino acid produces a *pseudo*-tripeptide (**22**). Removal of the Fmoc group and acylation by an α-bromoanhydride gives the cyclisation precursor (**23**). Reductive deprotection of the thiol was best achieved by tributylphosphine in PrOH/DMF/H_2O and only proceeded cleanly when the solid support was a polyethylene glycol - polystyrene copolymer. The cyclisation itself to give (**24**) (9 or 10-membered ring) was accomplished using tetramethylguanidine followed by TFA-mediated removal from the resin (Figure 32).

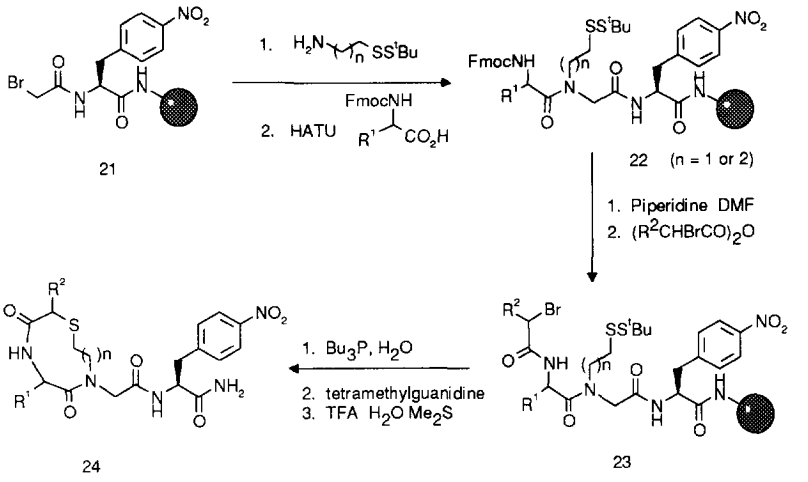


Figure 32.

5.4 Other Solid-Phase chemistry amenable to Library Synthesis

An increasing number of synthetic transformations that are routine in traditional solution phase organic chemistry are now being re-established on various solid supports. The resource required to achieve this represents no small effort mainly due to the different requirements and properties of the two phases. Several of the reactions mentioned below have yet to be reported as component steps in a library synthesis.

Kurth and co-workers have recently described the synthesis of a library of 1,3-diols with potential antioxidant activity by use of a two-step aldol-reduction sequence (Figure 33).[61] The common precursor, (**25**) is readily prepared from Merrifield resin and can readily be converted to a zinc enolate that reacts smoothly with a range of aldehydes. DIBAL reduction of the intermediate aldol products effects concomitant removal from the resin giving diols such as (**26**) (26% overall as a 7:5 *threo:erythro* mixture) . In all, 27 discrete compounds were prepared in this manner.

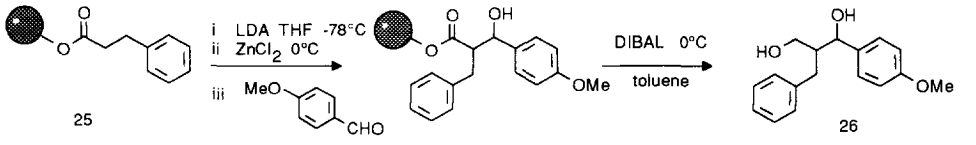


Figure 33.

In a further prototype (3x3) library synthesis of β-mercapto ketones (**30**), Kurth has highlighted a range of novel transformations on solid phase (Figure 34).[62] The five-step synthesis starts with a S_N1 reaction between 'trityl chloride' resin (**27**) and butanediol followed by oxidation of the remaining hydroxyl group using sulphur trioxide-pyridine complex to give an aldehyde (**28**). Wittig reaction of (**28**) with any of a range of stabilised phosphoranes gives an enone (**29**) that will readily undergo 1,4-addition of thiophenols, liberating the requisite β-mercapto ketones (**30**) after formic acid catalysed removal from the support. This

prototype library is a clear indicator that complex multistep syntheses can be used in a combinatorial sequence to build very large arrays.

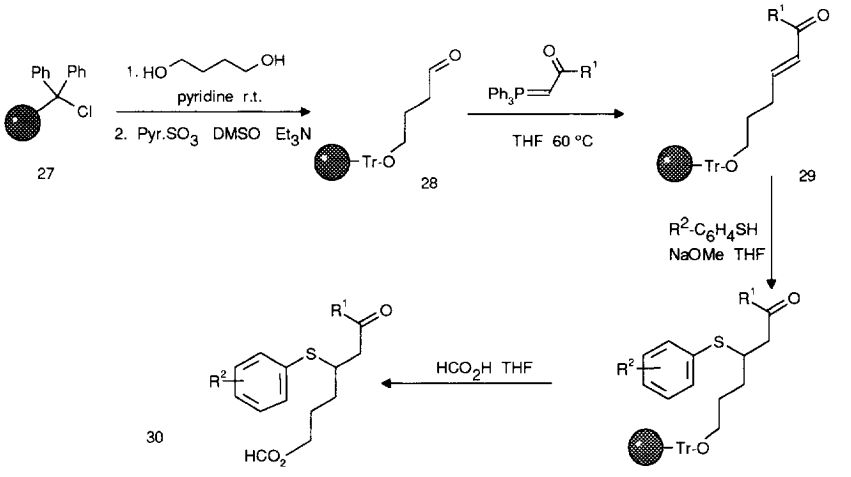


Figure 34.

The foregoing work of Kurth and the following examples serve to dispel a commonly held view that moisture sensitive or organometallic reactions are not feasible on solid phase. Deshpande and colleagues at Bristol-Myers-Squibb have now shown that both Heck[63] and Stille[64] coupling reactions proceed well on Wang polystyrene resin. Solid-phase Heck reactions proceed under conditions similar to those documented for solution phase equivalents - thus the immobilised styrene (**31**) reacts with 2-bromothiophene under palladium catalysis to give (**32**) in 76% yield (Figure 35).

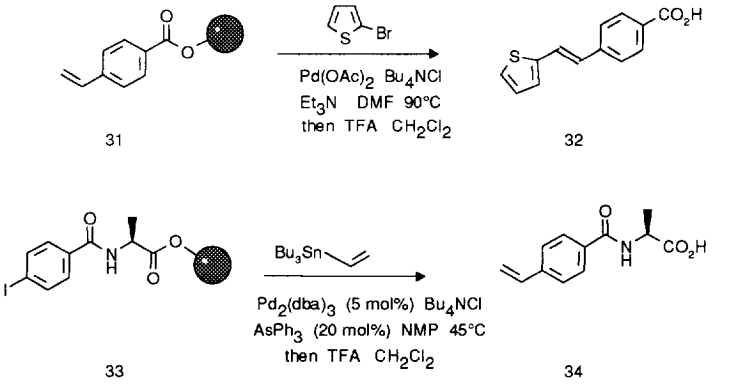


Figure 35.

Yields generally varied from 64-91%. Similarly, vinyl- or arylstannanes can be coupled efficiently to polymer-bound aryl iodides, representative yields after resin cleavage being 85-92%.[64] In a specific example,

the iodobenzamide (**33**) is coupled to vinyltributylstannane to give the styrene (**34**) in 92% yield after removal from the resin. It should be re-emphasised that in these and most related cases, only the desired product remains bound to polymer, giving ample opportunity to wash all reagent impurities out of the system before separating product(s) and resin.

5.5 Analysis

The analysis of combinatorial libraries presents particular difficulties which escalate proportionally as the mixture size increases. The process of analysis and validation is best handled in two discrete phases:

(i) Thorough validation of all chemical steps by model syntheses of single compounds
(ii) Spectroscopic analysis of representative library mixtures and surrogate chemical markers

As library synthesis moves away from amide coupling to other reactions less well-tailored to solid-phase synthesis, the validation of all chemical steps to ensure good yields will be vital. A useful guideline adopted by several groups including ours[49] is to identify several likely 'worst case' coupling examples (due to steric or electronic deactivation) and optimise reaction conditions based on these components. Alternatively, a full assessment of the relative coupling affinities of every monomer in a proposed library can be completed to weed out unreactive monomers.[52] Product analysis can be achieved by cleavage from the support followed by traditional characterisation, although techniques have emerged now for quite sophisticated, non-destructive product analysis 'on the bead' by IR, or NMR spectroscopy. Resin bound substrates are ideal for KBr pellet IR and also for *photoacoustic* IR, a surface analysis technique requiring minimal sample preparation that has only rarely been applied to polymer substrates.[65] In a recent example,[66] our group monitored progress of an esterification reaction using the increasing carbonyl absorptions as indicators (Figure 36).

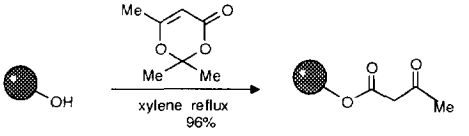


Figure 36.

Gel-phase carbon NMR spectroscopy[67] can now give adequate resolution to monitor reaction progress in small molecule substrates whereas, in the case of proton spectra, the line widths are often too broad to give any meaningful coupling information. Vastly improved resolution solid-phase ^{1}H NMR spectra have recently been achieved by using Magic-angle spinning (MAS) to minimise the line broadening observed in a conventional probe.[68] This technique offers near-solution spectral quality, although specialist MAS probes require a major capital outlay.

Analysis of multicomponent mixtures has generally been tackled by HPLC, MS or HPLC-MS combinations. Small mixtures (<20 components) may be teased apart by HPLC only with difficulty to give meaningful

yields, although as mixture size increases the main aim of any analysis is to ensure that the vast majority of expected components are *represented* in the mixture. The latter is generally best achieved by mass spectrometry and all of the familiar MS techniques have been used for mixture analysis, although electrospray appears to be the most reliable method to date.[69] Matrix-Assisted Laser Desorption Ionisation - time-of-flight (MALDI-TOF) mass spectrometry is a tool of increasing importance in multicomponent analysis because of its reduced tendency to preferential ionisation, therefore giving an increased likelihood of observing ions from all components of a given mixture.[70] To date, its main application has been in characterisation of multipeptide mixtures,[71] but a logical extension of this is to combinatorial libraries of small nonpeptides.

In general, the parallel synthesis of at least one single-compound surrogate marker should be undertaken for each library. The successful synthesis of this compound (analysed by normal methods) ensures that each library mixture was exposed to the appropriate sequence of chemical steps.

6 Future developments in combinatorial chemistry

The technique of combinatorial chemistry has developed primarily as a method of generating increasingly large numbers of compounds. However, making large numbers of compounds is not the only objective: after all it only takes one compound to make a drug. The most likely developments in the future therefore are those which exploit the potential of combinatorial chemistry in smarter and more creative ways. Quality rather than quantity will become the new goal.

Many of the aspects of combinatorial chemistry discussed in this review have described efficient ways of using libraries. In the future, we will see developments in compound selection strategies to ensure maximum value from the synthetic and screening effort employed. Furthermore, there will be a drive to increase the quality (purity and characterisation) of the compounds synthesised as well as an increase in the range of chemistry used in compound synthesis.

6.1 Compound Selection Strategies in Lead Discovery and Lead Optimisation

Combinatorial chemistry has increased the number of compounds that can be synthesised by orders of magnitude over other methods but it is still impossible to synthesise and screen millions of compounds in a format that allows useful information about their individual potential as drugs to be discovered. Thus, some method that selects which library compounds to make is still attractive. As more structural information about the binding of compound to receptor becomes available, the number of methods that can be used for rational selection of appropriate library compounds increases, and this has a profound effect on the library design strategy.

6.1.1 No knowledge of target

In some situations there may be no knowledge of the molecular target at all. For example many successful anti-infective agents have been discovered and used for years without any mechanistic understanding. However, even in these circumstances there are some 'design principles' that can be applied to compound selection. For example:

Diversity: Concern is frequently expressed that the diversity of a corporate chemical database (or combinatorial library) is too limited (particularly if screening has failed to deliver a lead for a target of biological interest!). Armed with molecular spreadsheets capable of multivariate statistics and large numbers of structure-based parameters such as shape, lipophilicity, dipole moment *etc.*, it is relatively easy to group molecules together and generate some diversity measurements. In addition, it is certainly true that most pharmaceutical companies' catalogues of synthesised compounds contain clusters of compounds made for specific projects. However, this apparent lack of diversity primarily reflects the chemist's perception of diversity. There are countless examples where a small change (to the chemist's eye) in structure dramatically affects biological activity. Clearly, these changes are perceived by the receptor as being considerable, and thus the challenge in the use of diversity measurements for library design and compound selection is to decide which parameters are relevant. Some attempts have been made to generate compound libraries with diversity measured in this way,[72] but it is difficult to assess the impact on success these considerations have made.

In general, it may be said that although everyone seems to agree that compound diversity in libraries is important, there is uncertainty that the methods available really generate a meaningful expression of such diversity. At a more fundamental level, it may be desirable and sufficient to use available diversity measurements to provide a spread of compounds that at least look like good starting points for medicinal chemistry programs.

Design with optimisation in mind: Lead discovery screening is mostly carried out *in vitro*, often with mixtures of compounds. The discovery of a lead is only the first stage in drug discovery, and thus it seems sensible to select compound libraries that will facilitate the next steps. Consequently, the lead should have some drug-like qualities. The selection of library compounds should thus consider such parameters as a molecular weight limit, to avoid high clearance associated with many large molecules. Assuming one knows the intended route of delivery (e.g. oral or injection) one may also wish to choose molecules which fall within specified lipophilicity ranges and avoid known metabolic vulnerability or functional groups associated with toxicity or carcinogenicity. In addition it may be argued that some types of molecule might be intrinsically easier to optimise than others. For example, the problems of down-sizing large molecules are well documented. Conversely, adding lipophilic side chains is frequently a successful route to enhancing the potency of small polar molecules. Conformational restraint is often a good way of increasing potency or selectivity, but it may be that some flexibility would be advantageous for lead discovery because of the greater number of conformations available to such molecules. Thus the optimal library for lead discovery may include molecules which retain some flexibility but are relatively small, to simplify further optimisation.

6.1.2 Design based on structural information

In many situations in lead discovery a certain amount of information about the target may be available, and this can be used for intelligent design of a compound library.

Knowledge of target class: It is possible to make use of even a limited amount of information about the intended target as an aid to lead discovery library design. An obvious example is the design of inhibitors of protease enzymes, as these usually have relatively well defined active site residues. For example, if the intended target was a zinc-containing protease (such as angiotensin-converting enzyme) then it might be appropriate to design a library of carboxylic acids or thiols.

Limited structural information: The target of many medicinal chemistry programs is frequently a membrane-bound receptor. Although not so well defined as many enzymes, techniques such as site-directed mutagenesis experiments often generate information useful in library design. It may be possible to identify key residues in the receptor, or in a known protein agonist/antagonist, that can give clues to some of the functionality that a small molecule ligand should contain.

Atom-level structural information: Design principles can be most comprehensively applied when the structure of the intended target (or a ligand) is known at the atomic level. One of the growth areas in molecular modelling recently has been the use of *de novo* ligand design to fit cavities in proteins.[73] However, a drawback of many examples of this work is that the algorithms frequently suggest molecules which are hard to synthesise. In contrast, combinatorial chemistry deals with large numbers of molecules that can be made relatively easily. The term '*Virtual Library*' has been used to denote a set of compounds which are accessible using parallel synthesis methods. If the cavity-filling algorithms can be restricted to examine only these virtual libraries, the predictions made can be tested by synthesis and biological assay without difficulty. Furthermore as *de novo* design is in its infancy and still a very difficult challenge, it is a distinct advantage that combinatorial chemistry is able to synthesise many compounds predicted by an algorithm to have some affinity, rather than just one 'best shot' molecule.

Clearly virtual libraries can be used in any of the situations in which 3D database searching is currently employed. One might, for example, have a pharmacophore model constructed from competitor compounds. This model could be used in a flexible search of the virtual library to pick examples for synthesis and screening.

Ready access to any of the compounds in a virtual library by rapid synthesis makes such a library virtually indistinguishable from a set of compounds that already exist from a structure-based lead discovery perspective. Given that it is not difficult to generate combinatorial library plans to make millions or even billions of drug-like molecules, it seems reasonable to expect that most structure-based lead discovery in the near future will generate leads from combinatorial synthesis rather than screening databases of existing compounds.

6.2 Purification and analysis requirements

At the moment, combinatorial chemistry is almost universally carried out with no purification and minimal analysis, as these steps for large numbers of compounds would be much too time-consuming to be worthwhile. However, yield and confirmation of structure is certainly desirable if they can be achieved easily, but this depends largely on whether we have prepared mixtures or single compounds.

The analytical separation of compounds within mixtures has been demonstrated.[50 74] This information is very useful in providing reassurance that many compounds have been made in a combinatorial chemistry experiment but is difficult to reproduce on a preparative scale. Also, biological information obtained on a mixture is usually employed to decide which mixture should be deconvoluted to identify the compound(s) responsible for activity. It will be interesting to see whether methods will be developed that make use of data obtained on mixtures to construct structure-activity relationships, rather than at present merely to guide isolation of single compounds.

The purification and analysis of single compounds could put the products of combinatorial chemistry on an equal footing with those of 'traditional' synthesis, allowing quantitative use of the biological data obtained. One of the biggest problems with using biological data obtained on compounds made with minimal analytical data is in the interpretation of negative results. Without analytical proof of the presence of a designed compound, it is dangerous to assume that the compound is inactive, as it may not have been present in the assay at all.

6.3 Numbers: mixtures *vs.* single compounds

The concept of deliberately generating and screening compound mixtures has caused some alarm to biologists and chemists alike. However, in the natural biological environments, enzymes and receptors interact with their target substrates from amongst a mixture of other low molecular weight compounds. What then are the pros and cons of mixture screening, and what does this mean for the future of mixtures in combinatorial chemistry?

For the sake of simplicity it is usually assumed that the components of a mixture of compounds don't interact in any way with each other. The observed biological activity can then be regarded as a sum of the individual effects of all of the compounds present in the mixture. Two extremes are possible: i) all the compounds have equal biological activity and ii) one compound gives rise to all of the activity and the others are completely inactive. The truth will almost certainly lie somewhere between these extremes.

For drug discovery purposes the end objective is a single potent compound, rather than an active mixture, and thus the mixture is a means to an end rather than an end in itself. For simplicity's sake, when deconvoluting a mixture, it is assumed that the mixture demonstrating greatest activity also contains the most active single compound. Theoretically this need not be the case and indeed there are several examples where the most active mixture did not contain the most active single component. For example, the 'all-D' hexapeptide Ac-RFWINK-NH_2 was identified by a deconvolution strategy from a peptide library. The most

active mixtures at each stage in the deconvolution were: Ac-RYXXXX-NH_2, Ac-RFWXXX-NH_2, Ac-RFWWXX-NH_2 and Ac-RFWINX-NH_2.[75] In other words, the most active peptide found was in the most active mixture on only two out of four occasions. In other published examples of iterative deconvolution the most active compound discovered has been present in nine out of 13 of the most active mixtures at each stage and a total of one out of four occasions for the first round of iterative deconvolution.[76] This may not be a surprising result, but it does demonstrate that it is frequently necessary to deconvolute several active mixtures in order to find the most potent single compound.

Mixture Size: Another key question connected with mixture screening is what is the optimal size of a mixture? A solution to this question can usually be arrived at following a consideration of several key parameters. For example, what is the total concentration of material that the biological assay will tolerate, or can be assembled without precipitation from an assay-compatible solution? Secondly, what level of activity is being sought from an active compound in the library? Is there a requirement for a nanomolar lead, or would 50 micromolar provide a sufficient starting point for lead optimisation?

To illustrate these considerations, we could put some hypothetical figures on these parameters. Suppose that the goal is to identify a ligand with an IC_{50} of 1μM and that solubility considerations and an apparent background level of activity prevent the screening of compounds at a total assay concentration above 100μM To detect activity, it would be necessary to screen each component of the mixture at 1μM or greater leading to a maximum mixture size of 100. However, because of the combinatorial 'mix and split' methodology generally used to synthesise mixtures, each mixture usually consists of a family of related structures. Thus, for an active mixture, many of the compounds in the mixture will have some activity, and so mixture sizes of over 100 (and concentrations of each component below 1 μM) are still likely to allow the detection of single compounds with an IC_{50} of 1 μM. 100 is therefore a *minimum* value for the optimal mixture size in this example.

There is experimental evidence to suggest that mixtures contain compounds with related activity. The table below shows one deconvolution experiment[75] and contrasts the observed mixture activity with the activity that would be expected if only one active compound was present in the mixture (Ac-RFWINK-NH_2, 18 nM).

Mixture definition	*Observed activity of mixture (nM)*	*Calculated activity assuming only AcRFWINK-NH2 is active (nM)*	*Increase in activity due to presence of other active compounds in the mixture (ratio)**
Ac-RFXXXX-NH_2	12,000	2,880,000	240
Ac-RFWXXX-NH_2	900	144,000	160
Ac-RFWIXX-NH_2	500	7,200	14
Ac-RFWINX-NH_2	200	360	2

* *This number has been defined as a sub-optimal binding factor (SBF) and has been calculated for several examples.*[31]

6.4 Development of new chemistry on solid phase.

Solid phase synthesis is highly suited to the synthesis of biopolymers such as DNA, RNA and peptides, as the chemistry required for chain extension is consistent for each step. It has therefore been worthwhile to put considerable effort into the optimisation of the coupling conditions to give highly efficient syntheses. However, the history of drug discovery suggests that no single class of compound will provide all the drugs of the future, and thus for combinatorial chemistry to have maximum impact, a large range of bond-forming reactions need to be developed on solid-phase. As discussed in section 5 above, there are many new methods emerging. However, much work remains to be done in this area and this is clearly an area of massive growth for the future.

6.5 Integration of combinatorial chemistry with medicinal chemistry

Combinatorial chemistry represents a broad spectrum of techniques that are rapidly becoming a standard part of the medicinal chemist's tool kit. But how will this technology develop in the future? Will it become a routine method of lead discovery used by all medicinal chemists or will it remain in the hands of specialists?

At one extreme, combinatorial chemistry can be considered to be purely a lead discovery tool. This might be most accurate where the techniques used require novel screening or synthetic strategies specifically designed to address the difficulties of handling large numbers of single compounds. Thus sophisticated approaches using labelling strategies, reporter gene constructs and compounds tending towards bio-polymers may drive combinatorial chemistry into the preserve of specialists. The medicinal chemist would then remain a customer of combinatorial chemistry, picking up the most useful discoveries and advances made using the technique.

The other extreme is to place combinatorial chemistry in the hands of existing drug discovery projects. This is much easier to achieve when a 'low-tech' library method is employed, particularly if synthesis is in a format that can use existing screens and makes use of techniques familiar to most chemists. Adoption of the technology is also more likely where the potential for using combinatorial chemistry for lead optimisation is seen as particularly important. The benefit of this is the seamless integration of the technology into the hands of those who will use the results.

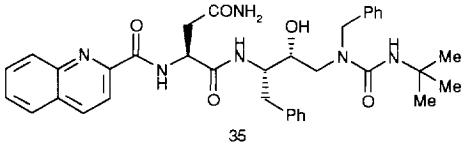
Whatever, the degree of integration into the medicinal chemist's laboratory, one thing is certain. Combinatorial chemistry as a technique for the rapid synthesis of drug-like compounds will continue to make a major impact on the way drug molecules are discovered.

Acknowledgements: We are grateful to Charlotte Allerton for checking the manuscript

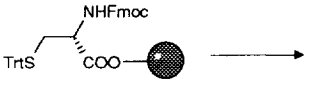
Note added in proof: The area of combinatorial chemistry research has continued to generate large numbers of publications. The following are a selection of some of the most significant recent papers.

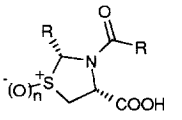
The concept of an 'orthogonal' combinatorial library has been described.[77] The same 15,625 trimers were synthesised separately in two libraries in such a way that any mixture from the first library and any mixture from the second library share only one trimer in common. This library method, similar in concept to 'indexed' libraries'[10] revealed a vasopressin-2 ligand of nanomolar affinity without any need for deconvolution of the library mixtures.

A number of new structural types have been synthesised on solid-phase and could now be prepared in large numbers through combinatorial synthesis. Aspartic acid protease inhibitors such as (**35**) have been prepared on solid-phase using a hydroxyl group as the tethering functionality.[78] The reaction of beta-mercaptoalkylamines such as cysteine derivatives with aldehydes on solid-phase has given a range of thiazolidines (**36**).[79] Reactions recently demonstrated on solid-phase include biaryl formation *via* Stille coupling,[80] and aryl ether synthesis using the Mitsunobu reaction.[81] Valuable linking groups for attaching organic molecules to solid-phase include a photolabile *ortho*-nitrobenzyl group (**37**) that yields either carboxamides following cleavage in aqueous solution,[82] and a solid-phase imidazolide carbamate (**38**) that can be used as a Cbz chloride equivalent.[83]

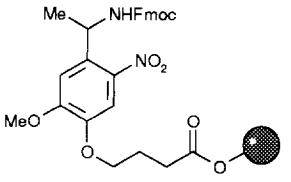


35

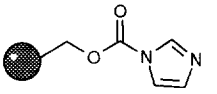




36



37



38

Further examples of analytical techniques being applied to combinatorial library samples have appeared, including affinity electrophoresis/mass spectrometry (ACE/MS)[84] and matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF MS).[85]

[1] This area has already been reviewed. See: Gallop, M.A.; Barrett, R.W.; Dower, W.J.; Fodor, S.P.A.; Gordon, E.M. *J. Med. Chem.* **1994,** *37*, 1233-1251 and 1385-1401. Janda, K.D. *Proc. Natl. Acad. Sci. USA* **1994,** *91*, 10779-10785. DeWitt, S.H. *Pharmaceutical News* **1994**, *1*, 11-14. Nielsen, J. *Chem. & Ind.* **1994**, 902-905.

[2] Smith, G.P.; Scott, J.K. *Science* **1990,** *249*, 386-390; Cwirla, S.; Peters, E.A.; Barrett, R.W.; Dower, W.J. *Proc. Natl. Acads. Sci. U.S.A.* **1990**, *87*, 6378-6382.

[3] Devlin, J.J.; Panganiban, L.C.; Devlin, P.E. *Science* **1990**, *249*, 404-406.

[4] Nikolaiev, V.; Stierandova, A.; Krchnak, V.; Seligmann, B.; Lam, K.S.; Salmon, S.E.; Lebl, M. *Peptide Res.* **1993**, *6*, 161-170. Salmon, S.E.; Lam, K.S.; Lebl, M.; Kandola, A.; Khattri, P.S.; Healy, S.; Wade, S.; Patek, M.; Kocis, P.; Krchnak, V.; Thorpe, D.; Felder, S. in *Peptides: Chemistry, Structure and Biology* ed. Hodges, R.S.; Smith, J.A. **1994,** Leiden: ESCOM-I11, 1001-1002.

[5] Needels, M.C.; Jones, D.G.; Tate, E.H.; Heinkel, G.L.; Kochersperger, L.M.; Dower, W.J.; Barrett, R.W.; Gallop, M.A. *Proc. Natl. Acad.. Sci. USA* **1993**, *90*, 10700-10704.

[6] Brenner, S.; Lerner, R.A. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5381-5383.

[7] Ohlmeyer, M.H.J.; Swanson, R.N.; Dillard, L.W.; Reader, J.C.; Asouline, G.; Kobayashi, R.; Wigler, M.; Still, W.C. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10922-10926.

[8] Peterson, J.R. *A Simple Approach to Small Molecule Synthetic Libraries for Drug Discovery and Directed Chemical Analoguing* presentation at *Exploiting Molecular Diversity: Small Molecule Libraries for Drug Discovery* January 23-25, 1995, La Jolla, California.

[9] Smith, P.W.; Lai, J.Y.Q.; Whittington, A.R.; Cox, B.; Houston, J.G.; Stylli, C.H.; Banks, M.N.; Tiller, P.R. *Biomed. Chem. Lett.* **1994**, *4*, 2821-2824.

[10] Pirrung, M.C.; Chen. J. *J. Am. Chem. Soc.* **1995**, *117*, 1240-1245.

[11] Merrifield, R.B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.

[12] Atherton, E; Sheppard, R.C.; *Solid phase peptide synthesis: a practical approach.* **1989,** IRL Press at Oxford University Press.

[13] Grubler, G.; Stoeva, S.; Echner, H.; Voelter, W. in *Peptides: Chemistry, Structure and Biology* ed. Hodges, R.A.; Smith, J.A. **1994**, Leiden: ESCOM-I11, 51-53.

[14] Rapp, W.; Zhang, L.; Bayer, E. in *Innovation and Perspectives in Solid Phase Synthesis; 1st Intl. Symposium*, ed. Epton, R., SPCC(UK) Ltd, Birmingham **1990**, 205-210. Rapp, W.E.; Bayer, E. in *Peptides: Chemistry, Structure and Biology* ed. Hodges, R.S.; Smith, J.A. **1994,** Leiden: ESCOM-I11, 40-43.

[15] Virgilio, A.A.; Ellman, J.A. *J.Am. Chem. Soc.* **1994**, *116*, 11580-11581.

[16] Geysen, H.M.; Meloen, R.H.; Barteling, S.J. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3998-4002. Geysen, H.M.; Barteling, S.J.; Meloen, R.H. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 178-182.

[17] Valerio, R.M.; Bray, A.M.; Campbell, R.A.; Dipasquale, A.; Margellis, C.; Rodda, S.J.; Geysen, H.M.; Maeji, N.J. *Int. J. Peptide Protein Res.* **1993**, *42*, 1-9.

[18] Englebretsen, D.R.; Harding, D.R.K. in *Peptides: Chemistry and Biology* eds. Smith, J.A.; Rivier, J.E. **1992** Leiden: ESCOM, 597-598. Englebretsen, D.R.; Harding, D.R.K. *Int. J. Peptide Protein Res.* **1992,** *40*, 487-496. Englebretsen, D.R.; Harding, D.R.K. *Int. J. Peptide Protein Res.* **1994**, *43* 546-554.

[19] Frank, R.; Doring, R. *Tetrahedron* **1988**, *44*, 6031-6040. Blankemeyer-Menge, B.; Nimtz, M.; Frank, R. *Tetrahedron Letters* **1990**, *31*, 1701-1704. Bartl, R.; Hoffmann, S.; Tegge, W.; Frank, R. in *Innovation and Perspectives in Solid Phase Synthesis; 2nd Intl. Symposium* ed. Epton, R., Intercept Ltd Andover **1992**, 281-287. Frank, R. *Tetrahedron* **1992**, *48*, 9217-9232.

[20] Lebl, M.; Stierandova, A.; Eichler, J.; Patek, M.; Pokorny, V.; Jehnicka, J.; Mudra, P.; Zenisek, K.; Kalousek, J. in *Innovation and Perspectives in Solid Phase Synthesis; 2nd Intl. Symposium* ed. Epton, R., Intercept Ltd Andover **1992**, 251-257.

[21] Lloyd-Williams, P.; Albericio, F.; Giralt, E. *Tetrahedron* **1993**, *49*, 11065-11133.

[22] Atherton, E.; Clive, D.L.J.; Sheppard, R.C. *J. Am. Chem. Soc.* **1975**, *97*, 6584-6585. Fields, G.B.; Noble, R.L. *Int. J. Peptide Protein Res.* **1990**, *35*, 161-214.

[23] Shao, J.; Li, Y.H.; Voelter, W. *Int. J. Peptide Protein Res.* **1990**, *36*, 182-187.

[24] Hammer, R.P.; Albericio, F.; Gera, L.; Barany, G. *Int. J. Peptide Protein Res.* **1990**, *36*, 31-45.

[25] Guibe, F.; Dangles, O.; Balavoine, G.; Loffet, A. *Tetrahedron Letters* **1989**, *30*, 2641-2644. Blankemeyer-Menge, B.; Frank, F. *Tetrahedron Letters* **1988**, *29*, 5871-5874.

[26] Chao, H-G.; Bernatowicz, M.S.; Reiss, P.D.; Klimas, C.E.; Matsueda, G.R. *J.Am.Chem.Soc.* **1994**, *116*, 1746-1752.

[27] Thompson, L.A.; Ellman, J.A. *Tetrahedron Letters* **1994**, *35*, 9333-9336.

[28] Furka. A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Abstr. 14th Int. Congr. Biochem., Prague, Czechoslovakia.* **1988**, *5*, 47. Furka. A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Peptide Protein Res.*, **1991**, *37*, 487-493.

[29] Houghten, R.A.; Dooley, C.T. *BioMed. Chem. Lett.* **1993**, *3*, 405-412.

[30] Houghten, R.A. *Proc. Natl. Acad. Sci. USA.* **1985**, *82*, 5131-5135.

[31] Freier, S.M.; Konings, D.A.M.; Wyatt, J.R.; Ecker, D.J. *J. Med. Chem.* **1995**, *38*, 344-352.

[32] Erb, E.; Janda, K.D.; Brenner, S. *Proc. Natl. Acad. Sci.* **1994**, *91*, 11422-11426.

[33] Houghten, R.A.; Pinilla, C.; Blondelle, S.E.; Appel, J.R.; Dooley, C.T.; Cuervo, J.H. *Nature* **1991**, *354*, 84-86.

[34] Lam, K.S.; Salmon, S.E.; Hersch, E.M.; Hruby, V.J.; Kazmierski, W.M.; Knapp, R.J. *Nature* (London) **1991**, *354*, 82-84. Lam, K.S.; Hruby, V.J.; Lebl, M.; Knapp, R.J.; Kazmierski, W.M.; Hersch, E.M.; Salmon, S.E. *BioMed. Chem. Lett.* **1993**, *3*, 419-424.

[35] Lebl, M.; Patek, M.; Kocis, P.; Krchnak, V.; Hruby, V.J.; Salmon, S.E.; Lam, K.S. *Int. J. Peptide Protein Res.* **1993**, *41*, 201-203.

[36] Dooley, C.T.; Houghten, R.A. *Life Sci.* **1993**, *52*, 1509-1517.

[37] Ostresh, J.M.; Husar, G.M.; Blondelle, S.E.; Dorner, B.; Weber, P.A.; Houghten, R.A. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11138-11142.

[38] Fodor, S.P.A.; Read, J.L.; Pirrung, M.C.; Stryer, L.; Lu, A.T.; Solas, D. *Science* **1991**, *251*, 767-773.

[39] Patek, M.; Lebl, M. *Tetrahedron Letters* **1991**, *32*, 3891-3894.

[40] Nielsen, J.; Brenner, S.; Janda, K.D. *J. Am. Chem. Soc.* **1993**, *115*, 9812-9813.

[41] Nestler, H.P.; Bartlett, P.A.; Still, W.C. *J.Org.Chem.* **1994**, *59*, 4723-4724.

[42] Wyatt, J.R.; Vickers, T.A.; Roberson, J.L.; Buckheit Jr, R.W.; Klimkait, T.; DeBaets, E.; Davis, P.W.; Rayner, B; Imbach, J.L.; Ecker, D.J. *Proc. Natl. Acad. Sci.* **1994**, *91*, 1356-1360 and references therein.

[43] Owens, R.A.; Gesellchen, P.D.; Houchins, B.J.; DiMarchi, R.D. *Biochem. Biophys. Res.Commun.* **1991**, *181*, 402-408.

[44] Pinilla, C.; Appel, J.R.; Houghten, R.A. *Gene* **1993**, *128*, 71-76.

[45] Houghten, R.A.; Appel, J.R.; Blondelle, S.E.; Cuervo, J.H.; Dooley, C.T.; Pinilla, C. *Biotechniques* **1992**, *13*, 412-421.

[46] Eichler, J.; Houghten, R.A. *Biochemistry* **1993**, *32*, 11035-11041.

[47] Houghten, R.; Dooley, C. *BioMed. Chem. Lett.* **1993**, *3*, 405. Dooley, C.T.; Chung, N.N.; Schiller, P.W.; Houghten, R.A. *Proc. Natl. Acad. Sci.* **1993**, *90*, 10811-10815.

[48] Eichler, J.; Lucka, A.W.; Houghten, R.A. *Proceedings of the 23rd European Peptide Symposium,* Braga, September 1994, Poster 198.

[49] Terrett, N.K.; Bojanic, D.; Brown, D; Bungay, P.J., Gardner, M.; Gordon, D.W.; Mayers, C.J.; Steele, J. *BioMed. Chem. Lett.* **1995**, *5*, 917-922.

[50] 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.* **1992**, *89*, 9367-9371. Zuckermann, R.N.; Kerr, J.M.; Kent, S.B.H.; Moos, W.H. *J.Am.Chem.Soc.* **1992**, *114*, 10646-10649.

[51] Miller, S.M.; Simon, R.J. Ng, S.; Zuckermann, R.N.; Kerr, J.M; Moos, W.H. *BioMed.Chem.Lett.* **1994**, *4*, 2657-2662.

[52] Zuckermann, R.N.; Martin, E.J.; Spellmeyer, D.C.; Stauber, G.B.; Shoemaker, K.R.; Kerr, J.M; Figliozzi, G.M.; Goff, D.A.; Siani, M.A.; Simon, R.J.; Banville, S.C.; Brown, E.G.; Wang, L.; Richter, L.S.; Moos, W.H. *J.Med.Chem.* **1994**, *37*, 2678-2685.

[53] Cho, C.Y.; Moran, E.J.; Cherry, S.R.; Stephans, J.C.; Fodor,S.P.A.; Adams, C.L.; Sundaram, A.; Jacobs, J.W.; Schultz, P.G. *Science* **1993**, *261*, 1303-1305.

[54] Bunin, B.A.; Ellman, J.A. *J.Am.Chem.Soc.* **1992**, *114*, 10997-10998.

[55] Ellman, J.A.; Virgilio, A.A. *U.S.Patent 5,288,514* **1994**.

[56] DeWitt, S.H.; Kiely, J.S.; Stankovic, C.J.; Schroeder, M.C.; Reynolds Cody, D.M.; Pavia, M.R. *Proc. Natl. Acad. Sci.* **1993**, *90*, 6909-6913.

[57] Gordon, D.W.; Steele, J. *BioMed. Chem. Lett.* **1995**, *5*, 47-50.

[58] Gordon, D.W.; Steele, J. unpublished results.

[59] Beebe, X.; Schore, N.E.; Kurth, M.J. *J.Am.Chem.Soc.* **1992**, *114*, 10061-10062.

[60] Moon, H.-S.; Schore, N.E.; Kurth, M.J. *Tetrahedron Lett.* **1994**, *35*, 8915-8919.

[61] Kurth, M.J.; Ahlberg Randall, L.A.; Chen, C.; Melander, C.; Miller, R.B.; McAlister, K.; Reitz, G.; Kang, R.; Nakatssu, T.; Green, C. *J.Org.Chem.* **1994**, *59*, 5862-5864.

[62] Chen, C.; Ahlberg Randall, L.A..; Miller, R.B.; Jones, A.D.; Kurth, M.J. *J.Am.Chem.Soc.* **1994**, *116*, 2661-2662.

[63] Yu, K.-L.; Deshpande, M.S.; Vyas, D.M. *Tetrahedron Letters* **1994**, *35*, 8919-8922.

[64] Deshpande, M.S. *Tetrahedron Letters* **1994**, *35*, 5613-5614.

[65] Vidrine, D.W.; *Photoacoustic Fourier Transform Infrared Spectroscopy of Solids and Liquids*, Chapter 4 of *Fourier Transform Infrared Spectroscopy*, Academic Press **1982**, *3*, 125-148.

[66] Parrott, H.M.; Clark, D.A. (Physical Sciences, Pfizer, Sandwich) unpublished results.

[67] Look, G.C.; Holmes, C.P.; Chinn, J.P; Gallop, M.A. *J. Org. Chem.* **1994**, *59*, 7588-7590.

[68] Fitch, W.L.; Detre, G; Holmes, C.P.; Shoolery, J.N.; Keifer, P.A. *J. Org. Chem.* **1994**, *59*, 7955-7956.

[69] Metzger, J.W.; Wiesmuller, K.-H.; Gnau, V.; Brunjes, J.; Jung, G. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 894-896.

[70] Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proc.* **1987**, *78*, 53-68. Karas, M; Bahr, U.; Hillenkamp, F. **1989**, *92*, 231-242.

[71] Youngquist, R.S.; Fuentes, G.R.; Lacey, M.P.; Keough, T. *Rapid Communications in Mass Spectrometry* **1994**, *8*, 77-81 and references therein.

[72] Martin, E.J.; *Measuring Chemical Diversity: Random Screening or Rational Design?* presentation at *Exploiting Molecular Diversity: Small Molecule Libraries for Drug Discovery* January 23-25, 1995, La Jolla, California.

[73] Leach, A.R.; Kilvington, S.R. *J. Computer-Aided Molecular Design* 1994, *8*, 283-298.

[74] Kerr, J.M.; Banville, S.C.; Zuckermann, R.N. *BioMed. Chem. Lett.* **1993**, *3*, 463-468. Carell, T.; Wintner, E.A.; Bashir-Hashem, A.; Rebek Jr, J. *Angew. Chem. Int. Ed. Engl.* **1994**, 33, 2059-2061.

[75] Dooley, C.T.; Chung, N.N.; Wilkes, B.C.; Schiller, P.W.; Bidlack, J.M.; Pasternak, G.W.; Houghten, R.A. *Science* **1994**, *266*, 2019-2022.

[76] See deconvolution experiments described in refs. 36, 42 and 48.

[77] Déprez, B.; Williard, X.; Bourel, L.; Coste, H.; Hyafil, F.; Tartar, A. *J.Am.Chem.Soc.* **1995**, *117*, 5405-5406.

[78] Kick, E.K.; Ellman, J.A. *J.Med.Chem.* **1995**, *38*, 1427-1430.

[79] Pátek, M.; Drake, B.; Lebl, M. *Tetrahedron Letters* **1995**, *36*, 2227-2230.

[80] Forman, F.W.; Sucholeiki, I. *J. Org. Chem.* **1995**, *60*, 523-528.

[81] Rano, T.A.; Chapman, K.T. *Tetrahedron Letters* **1995**, *36*, 3789-3792.

[82] Holmes, C.P.; Jones, D.G. *J. Org. Chem.* **1995**, *60*, 2318-2319.

[83] Hauske, J.R.; Dorff, P. *Tetrahedron Letters* **1995**, *36*, 1589-1592.

[84] Chu, Y-H.; Kirby, D.P.; Karger, B.L. *J.Am.Chem.Soc.* **1995**, *117*, 5419-5420.

[85] Egner, B.J.; Langley, G.J.; Bradley, M. *J. Org. Chem.* **1995**, *60*, 2652-2653.

(*Received* 13 *June* 1995)