CHIMIA 51 (1997) Nr. 11 (November)

Chimia 51 (1997) 832–837

ISSN 0009–4293

Solution-Phase Combinatorial Chemistry in Lead Discovery

Nick Bailey, Anthony W.J. Cooper, Martyn J. Deal, Anthony W. Dean, Andrea L. Gore, Michael C. Hawes, Duncan B. Judd, Andrew T. Merritt*, Richard Storer, Stuart Travers, and **Stephen P. Watson**

Abstract. Solution-phase approaches in combinatorial chemistry complement solid-phase approaches and each can be used to advantage in particular circumstances. Solution-phase synthesis of pools of compounds, whilst allowing successful identification of a selection of good lead structures for medicinal chemistry programmes, also reinforced a number of the disadvantages of such an approach. Solution-phase parallel synthesis of discrete compounds has, however, proved to be a very useful and popular approach both for lead generation and in lead optimisation work. The range of chemistry suitable for use in such approaches is expanding rapidly and some of these chemistries are discussed. The current focus is on enhancing the quality of compounds prepared in array formats, and we describe a number of useful approaches which are being developed to that end.

Introduction

High-throughput screening is now firmly established as a key component of lead-generation activities within pharmaceutical research. Initially, the emphasis was placed on screening of compound collections and natural products. However, development of assay technologies, in particular the increasing use of automation, had generated excess screening capacity, such that by 1992 it was apparent that a radically different approach to chemical synthesis was required to provide a greater number and selection of compounds. Combinatorial chemistry approaches, up to that time focusing largely on amide-bond chemistry leading to peptidic compounds, were being established externally by a wide variety of new specialist companies. An appreciation of the potential of these was evidenced both by acquisition by and the establishment of strategic alliances with major pharmaceutical companies. Establishment in-house by the latter of dedicated groups to import and develop such new technologies also became commonplace.

**Correspondence*: Dr. A.T. Merritt
Combinatorial Medicinal Chemistry
GlaxoWellcome Research and Development
Medicines Research Centre
Gunnels Wood Road
Stevenage, Herts., UK SG1 2NY

In *GlaxoWellcome*, a number of such specialist groups were established at the different major research sites with a variety of different objectives. These included pioneering new technologies for synthesis and in particular introduction of a degree of automation into the synthetic process as had been done to revolutionise the screening process. Two major objectives were envisaged: a significant increase in numbers of suitable compounds available for screening from primary library generation by specialist groups and a change in the practices employed throughout medicinal chemistry to make use of the new technologies in lead optimisation work.

GlaxoWellcome Solution-Phase Library Approaches

The emphasis of our primary library approaches was on the generation of compounds which would provide suitable lead structures for medicinal chemistry programmes. In our initial work in the UK group, we decided to focus on solution-phase work; solid-phase initiatives were being established elsewhere in our company and we wished to cover all options. Solution-phase work has a number of attractive features. A wide range of chemistry is readily applicable and no additional steps are required to attach to or detach from supports. For clean robust reactions this approach has much to recommend it. Significant numbers and appropriate quantities of compounds are generated in a form directly compatible with the requirements of many screens. It is ideal for short, high-yielding reaction sequences but is less suitable for longer, more complex reaction sequences where solid-phase work currently allows much better control of purity.

Decisions regarding library size and whether to construct it as mixtures (pools) or as single entities are determined by the major purpose for which the library is designed. In principle, each of these options has advantages in particular situations. Generation of pools requires more development work and the products obtained are more difficult to analyse. The screening data obtained is less precise and reliable than that obtained from discretes although throughput is increased. Where a library is constructed as pools, however, a viable process for 'decoding' to identify individual compounds of interest needs to be devised as part of the planning process.

Diamide-Trimer Library

An early solution-phase approach in *GlaxoWellcome* was the preparation of a diamide-trimer library using a design which would allow for a chemical decoding procedure (*Scheme 1*). We felt that a trimer library offered the best balance between numbers of building blocks and products, utilised an acceptable number of steps and gave products in the right molecular-weight range. Amide-bond chemistry was used to generate the library by coupling of sets of carboxylic acids and amines to central amino acids leading to products represented by structure **1**. From 50 amino acids, 80 carboxylic acids and 40 amines, we sought to generate all 160000 possible permutations in the first instance as 4000 pools of 40 compounds using the mixture of the 40 amines as the origin of the pool. The synthesis was car-

CHIMIA 51 (1997) Nr. 11 (November)

ried out in part manually and in part using custom-made robotic equipment. Screening of the 4000 samples would allow identification of active pools, and reverse direction synthesis would then be applied to generate the individual components of pools of interest for rescreening.

Many computational approaches are now being developed to assess and describe diversity in sets of compounds, but for this initial exercise, the sets of monomers were selected as being 'diverse' by a group of medicinal chemists. The amino-acid set contained both natural L-amino acids and some of the corresponding D-enantiomers as well as unnatural α- and non-α-amino acids. The acid set contained additional masked substitution which could be uncovered during a final deprotection step to increase diversity in the products. The amine set did not contain anilines in view of their lack of reactivity, and the final selection of the constituents of the set was made in the light of observed reactivities during the development phase. The chemistry development began with discrete synthesis of representative examples in the traditional way with full characterisation and analytical evaluation. Pool synthesis was initially conducted to give pools of three which we were able to analyse by HPLC and NMR. More complex mixtures of 10 compounds were monitored by HPLC and mass spectroscopy. Mixtures of the final pool size of 40 were monitored only in a qualitative way by HPLC, emphasising the critical importance of adequate reaction development in this approach.

The robotics required for this initial work were relatively simple and based on a modified liquid-handling robot (*Fig. 1*). Customisation allowed a capability to evaporate solvent and other volatile components by a combination of heating the reaction block and use of nitrogen gas. Reactions were conducted in glass vials in 80 wells per plate to parallel the screening format. Reactants and coupling agents which were likely to remain in the final products were prescreened to ensure they had no effect on the assays. Compounds of known biological activity were prepared by the automatic procedure and the appropriate wells identified on screening. Such controls are an important feature in helping to ensure that screening data are as reliable as possible.

The 50 amino acids were each Cbz-protected and the free acid group was coupled to an equimolar mixture of the 40 amines. Removal of the Cbz group from each of the resulting mixtures was effected by hydrogenolysis. All the work indicated above was carried out manually. The final stages, coupling of the 80-acid set to each of the 50 mixtures and subsequent acid-catalysed removal of protecting groups were carried out robotically. Removal of volatile components and solvents by blowdown with final evaporation *in vacuo* furnished the 4000 products for test. The nature of the design resulted in each of the 50 individual plates of 80

Scheme 1. *Amide-Trimer Synthesis*

Carboxylic acid (Ac)-------Amino Acid (AA)--------Amine (Am)
80 50 40

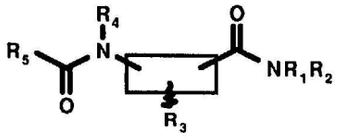
Manual/Automated Approach

Manual Preparation of feedstock dimers
Automated Synthesis of final trimers

$Am_{mix} + AA \xrightarrow[\text{manual}]{} Am_{mix}\text{-}AA \xrightarrow[\text{automated}]{} Am_{mix}\text{-}AA\text{-}Ac_1$, $Am_{mix}\text{-}AA\text{-}Ac_2$, ⋮

Docoding - Reverse direction synthesis
Maintain automated synthesis of final trimers

$Ac_x + AA \xrightarrow[\text{manual}]{} Ac_x\text{-}AA \xrightarrow[\text{automated}]{} Ac_x\text{-}AA\text{-}Am_1$, $Ac_x\text{-}AA\text{-}Am_2$, ⋮



(1)

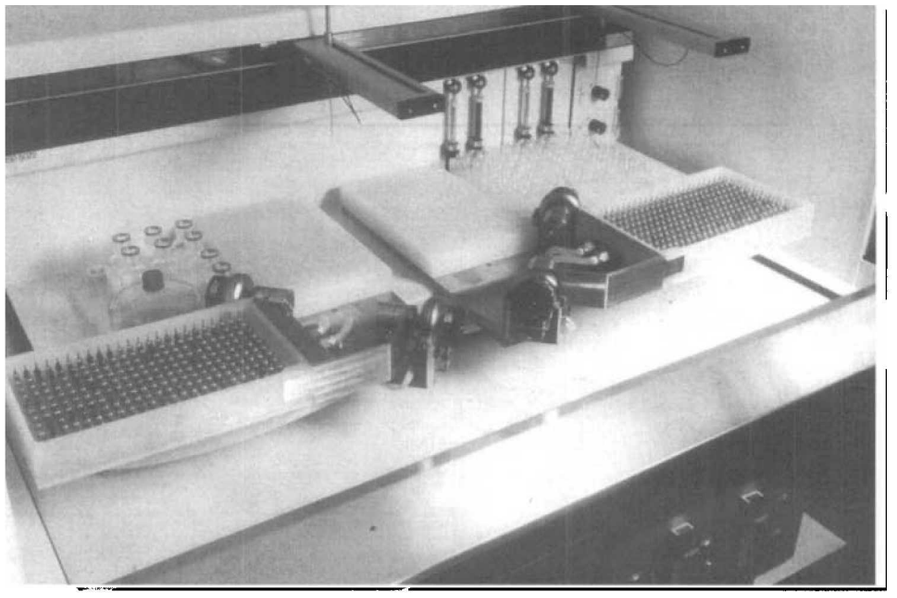


Fig. 1. *First-generation solution robotics*

CHIMIA 51 (1997) Nr. 11 (November)

samples representing one amino acid. Each unique two-dimensional well address on each plate represented a single carboxylic acid and each well on every plate had the same mixture of 40 amines represented.

The library was evaluated in the wide variety of screen types which were available at the time covering a range of target types and therapeutic areas. The chemistry to decode active pools required initial coupling of the acid and amino-acid ester defined by the well address and plate number followed by a hydrolysis step. Robotic coupling with the 40 individual amines provided the discrete products for screening. In some cases a range of activities was apparent on screening discretes which provided some tentative SAR data, but more reliably indicated only those more promising compounds which should be prepared in pure fully characterised form for rigorous evaluation. Good lead structures providing extremely acceptable starting points for medicinal chemistry programmes were identified in a variety of screens and are exemplified by those shown below (*Fig. 2*). The tryptophan derivative **2** was found to have potent activity as an NK1 antagonist with a p*Ki* value of 7.3. A variety of other samples showed activity in the NK1 assay, but in a number of other assays only one of the 4000 wells tested showed activity. One such example **3** was identified in the assay to find inhibitors of HIV proteinase. When the 40 individual constituents of the active well were generated as discretes, seven samples showed activity and were re-prepared in pure form. Compound **4** was identified as the most potent compound and provided a novel lead structure in this area. A further example where only one well of the 4000 showed activity was in the bradykinin assay. On decoding the active well, the activity was found to reside in a single compound **5** which had a p*Ki* of 5.8 and to the best of our knowledge was the first small-molecule antagonist.

A certain level of information was available by comparison of results from well to well and plate to plate, but full decoding was required to gain information about specific structures. This requirement revealed some of the limitations of this exercise. In two-thirds of cases, decoding of an active pool failed to yield a compound of any significance for reasons which were not extensively investigated. Thus, considerable time was lost following up the decodes which was in itself a rather labour-intensive exercise.

Several other extensions to the chemistry described for this library were pursued in a similar manner whilst the initial library was being screened. This included further amide couplings based not only on additional amino acids but also on diamines and diacids as central portions. Sulfonamides and oxadiazoles were also employed in these extensions to the original exercise (*Fig. 3*). In these cases also, samples were generated as pools of up to 40 compounds.

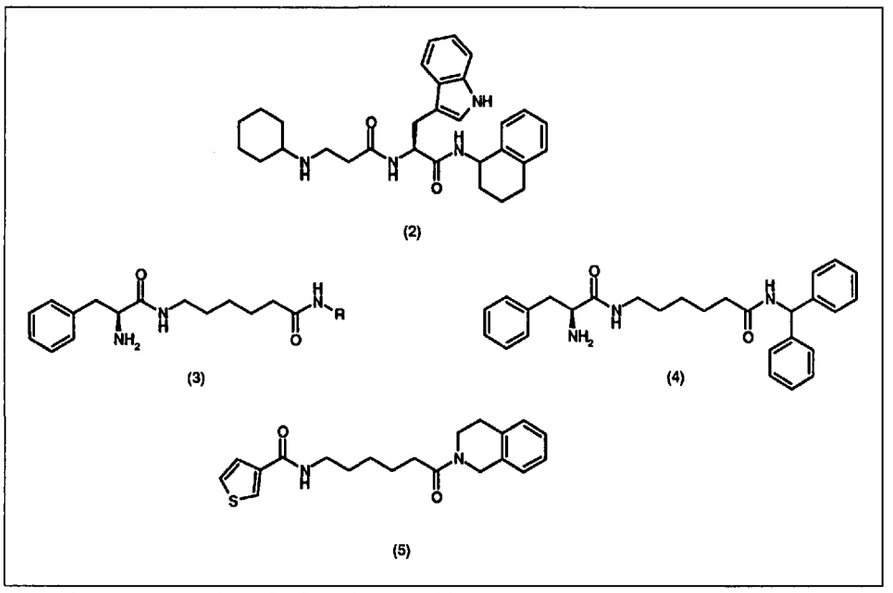


Fig. 2. *Lead structures from the diamide-trimer library*

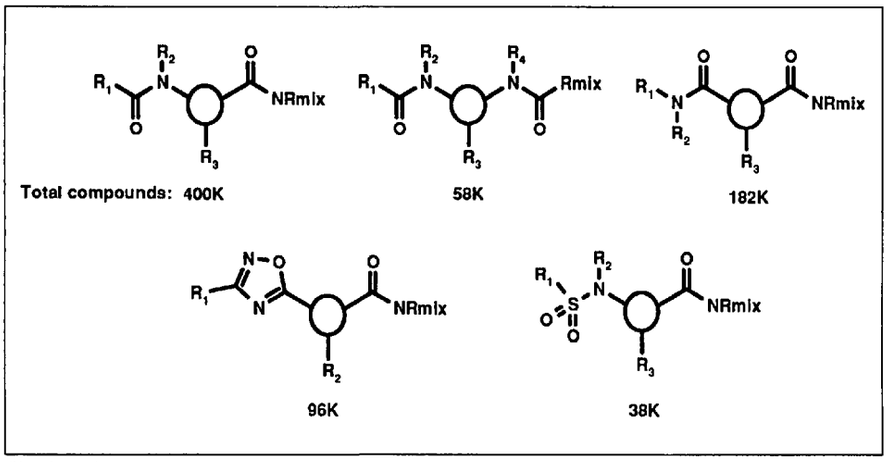


Fig. 3. *Other templates*

Extensions to the Range of Developed Chemistry

Looking to more general future requirements both for primary and focused libraries, we sought to expand our range of developed chemistries. Initially, we planned to do this by formation of libraries of dimers to validate the bond-forming reactions and subsequently to use combinations of appropriate reactions in the construction of trimer libraries to exploit fully the combinatorial potential. From our experiences with the decoding exercise from earlier libraries, we decided to focus our efforts on the synthesis of discrete compounds. This would allow subsequent robotic pooling if appropriate but would allow rapid rescreening of the individual components of active pools. This approach would also allow generation of pools of dissimilar compounds which we hoped would reduce the incidence of false hits obtained with the pools of similar compounds. We also wanted to extend the robotic capability to include some additional facilities and in particular to move to synthesis in standard microtitre footprint blocks.

On the chemical front, we looked to identify clean, high-yielding reactions with one major pathway which generated few, or ideally no, by-products. The chemistry needed to be versatile and robust to tolerate a range of substrates and to proceed

under the relatively mild conditions suitable for current robotics. The reactions had to be combinatorial (*i.e.* A + B to give C) to allow maximum benefit from the available building blocks. Also, we were interested only in those reactions which would give rise to products having a good pedigree or potential as lead structures for medicinal chemistry programmes. We are continually expanding our range of developed reactions beyond acylation and sulfonylation. Alkylations of both sulfur and nitrogen, reductive amination, cycloadditions, 1,4-additions, oxime and hydrazone formations amongst many others have proved to be of significant interest.

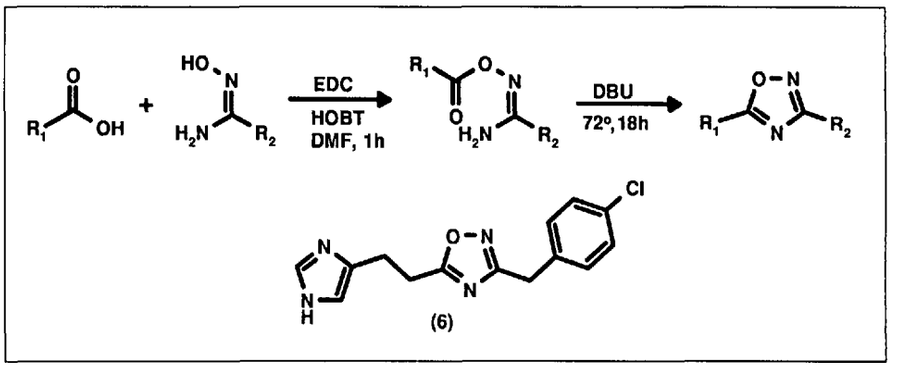
One particular area in which we have concentrated our efforts heavily has been in synthesis of a wide range of heterocyclic systems. Heterocyclic ring formation is often chemically efficient. Such structures provide good templates and are ubiquitous in drug molecules, and our current range includes a wide variety of 5-, 6- and 7-membered rings. Oxadiazoles were the first heterocyclic system we exemplified with the general approach shown in *Scheme* 2. In addition, we produced a library of compounds having potential to explore an area of H3 antagonists based on the lead structure **6** [1]. Easy access to two sites of diversity and good yielding reactions have made this a popular system which has since been included in a number of other libraries.

Thiazole ring formation provides another attractive option in the synthesis of combinatorial libraries. Reaction of α-halo ketones with thioureas and thioamides has been employed to give dimer libraries of thiazole and 2-aminothiazole derivatives (*Scheme 3*) [2]. The option to use monomer sets in different reactions to produce libraries is a very attractive one which allows maximum benefit to be derived. The same α-halo ketones were also used in the preparation of a thiazole-imine trimer library which demonstrates clearly the potential to make significant numbers of compounds from modest numbers of building blocks (*Scheme 4*). Reaction of 12 isothiocyanates with 80 different amines and subsequent reaction of the products with 12 α-halo ketones gave a library consisting of 11 520 discrete products from only 104 monomers. We attempted to generate a range of diversity in the products by exemplifying lipophilic, H-bonding, acidic and basic substituents in each component set.

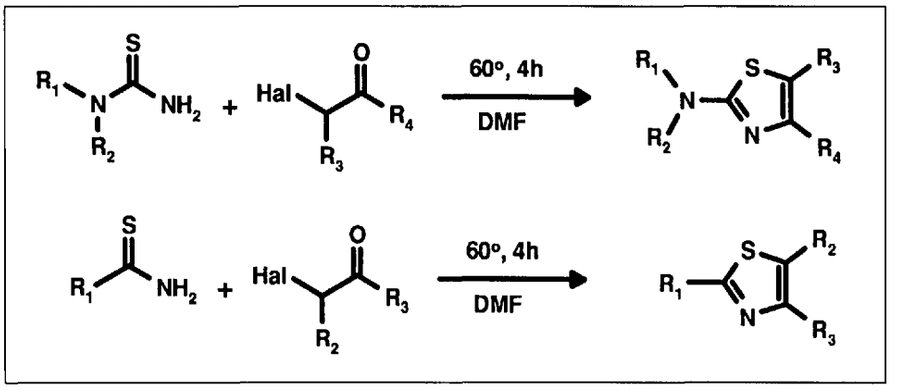
Trimer libraries of a variety of other heterocyclic systems have been prepared, some examples of which are discussed below. β-Keto esters have provided a set of very versatile building blocks. Reaction with hydrazines allowed synthesis of a dimer library of pyrazolones (*Scheme 5*). Initial reactions were conducted under microwave radiation but it was also possible to use simple elevated temperatures. Further reaction of the activated methylene component of the pyrazolone with an aldehyde set afforded the trimers. The use of β-keto esters in the preparation of a library of pyrimidines (*Scheme 6*) further demonstrates the ready access in two sequential steps to simple templates having three points of diversity. In a similar fashion, reaction of acylhydrazines with isothiocyanates and simple base-catalysed cyclisation of the product afforded 3-mercapto-1,2,4-triazoles allowing introduction of a third diversity element by subsequent S-alkylation (*Scheme 7*).

Dipolar cycloadditions are another class of reaction that can provide a range

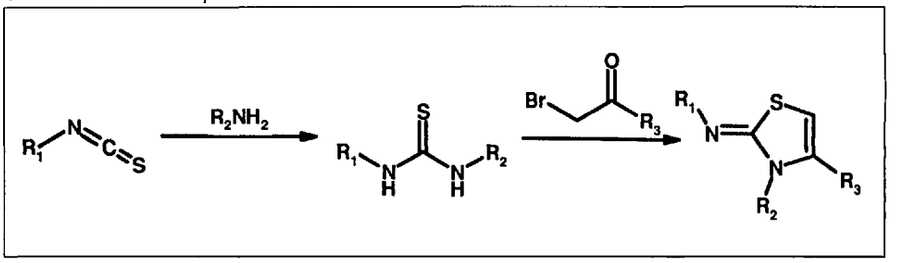
Scheme 2. *Oxadiazole Libraries*



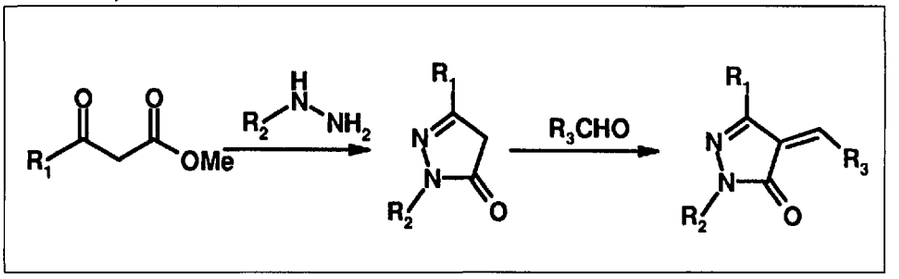
Scheme 3. *Thiazole Libraries*



Scheme 4. *Three-Component Libraries*



Scheme 5. *Pyrazolone Condensation Libraries*



of diverse product structures. For example, the three-component reaction using *in situ* generation of the active iminium species and subsequent cycloaddition with the maleimide gives interesting bicyclic products (*Scheme 8*). Diversity at four sites in the products is accessed easily from the readily available starting materials. Additionally, the products benefit from a carboxylic ester group which can be further elaborated. In addition to the examples described here, many other systems may be readily accessed by an appropriate choice of suitable chemical steps, and chemists will continue to develop further examples. Such simple reaction sequences demonstrate the ease with which solution-phase methods can be used to advantage when applied in appropriate circumstances.

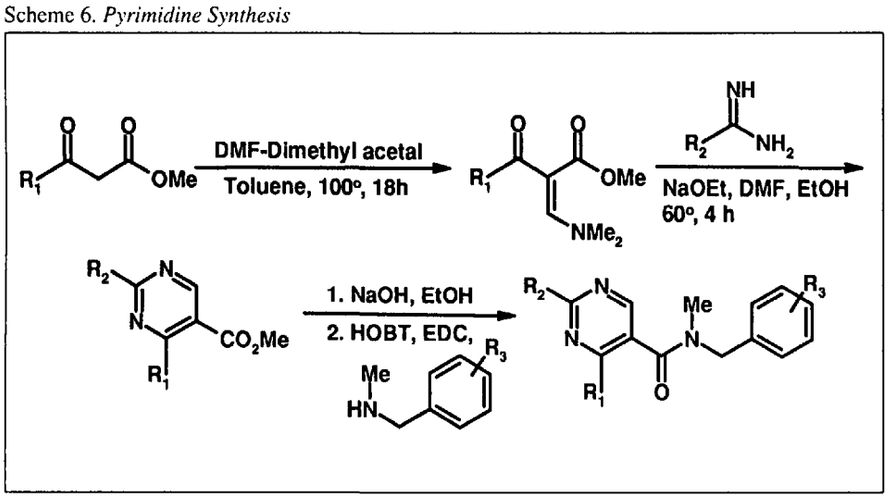
Future Enhancements of Solution-Phase Work

For less clean reactions and for longer sequences, solution phase is currently a less attractive option than solid phase. Developments in technologies for isolation and purification which can be employed to enhance the quality of solution-phase products will help to provide an important future for solution-phase synthesis. A number of such procedures are receiving attention by a variety of groups including ourselves. Use of supported reagents has the potential to combine the versatility of solution-phase synthesis with some of the advantage of solid phase. This is an area which is likely to develop rapidly with additional reagents becoming available on supports. The potential to use combinations of supported reagents where they would otherwise be incompatible is particularly attractive. Complementary approaches employ solid-supported scavenging agents to remove excess reagents or to isolate the product of interest from a mixture.

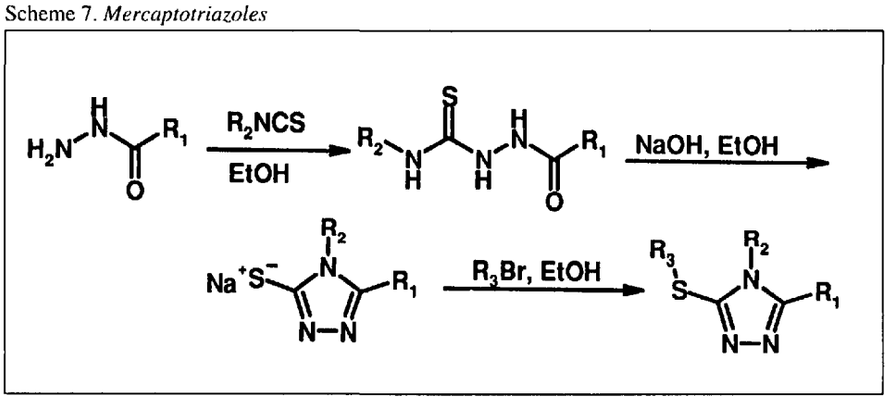
Whilst methods for parallel synthesis are now fairly well-advanced, methods for parallel purification are yet at a relatively early stage and serial purification, *e.g.* by preparative HPLC is currently the norm. There is a real need for methods which would allow genuine parallel purification to be employed with array synthesis. We now use routinely solid-phase extraction (SPE) cartridges for purification in array format. At the same time, we have been actively working to develop methods which would allow us to perform phase separations which would permit facile aqueous washing which could significantly enhance the purity of products produced. The initial approaches to aqueous separation make use of a commercially available hydrophobic membrane in a polypropylene cartridge to separate a $CHCl_3$ or CH_2Cl_2 phase from an aqueous phase (*Fig. 4*). When the biphasic mixture is applied to the tube, the membrane allows passage of the organic phase under gravity. The aqueous phase requires application of vacuum to allow passage through membrane and separation is thus easily achieved. The method has the further advantage that it works well with emulsions. A modification involving an absorbent packing which would allow similar phase separations where the organic phase is lighter than water has also been developed. Other groups have also reported work on similar principles to this for phase separation [3].

At the same time, we have developed an alternative novel phase separation procedure which also works well where the organic phase is lighter than the aqueous phase. This method, which we have termed the 'lollipop' method, involves cooling the biphasic mixture in a bath at –20° in the presence of an array of pins which extends through the organic phase and down into the lower aqueous phase (*Fig. 5*). After the freezing process, the solidified aqueous phase is removed attached to the pins as

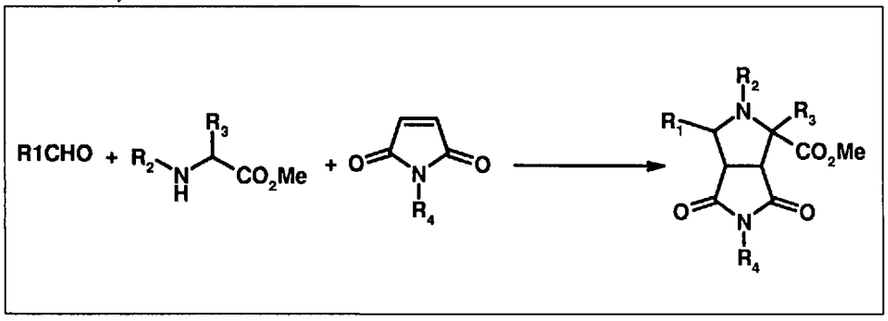
Scheme 6. *Pyrimidine Synthesis*



Scheme 7. *Mercaptotriazoles*



Scheme 8. *Cycloadditions*



CHIMIA 51 (1997) Nr. 11 (November)

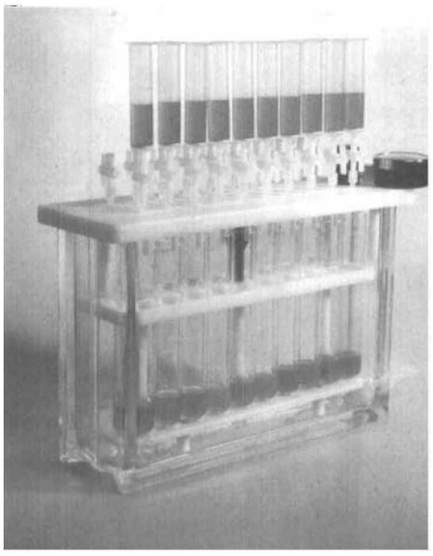


Fig. 4. *Liquid-phase separation*

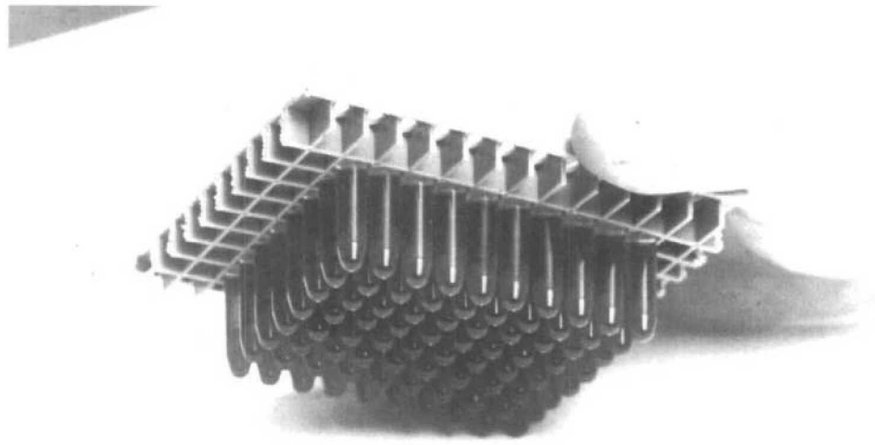


Fig. 5. *Lollipops – Pins in mixture after cooling*

the block bearing the pins is raised allowing complete separation of the phase in array format very easily and reproducibly (*Fig. 6*). This method is also applicable to phase separations involving water and chlorinated solvent mixtures. It is important that the tubes holding the mixtures are tapered to allow facile removal of the frozen aqueous phase. These methods are applicable for array formats in a variety of sample sizes including microtitre format. The ability to include simple workup procedures such as an aqueous wash in a solution-phase protocol adds significant value to the complexity of reagents that can be utilised and to the purity of the products.

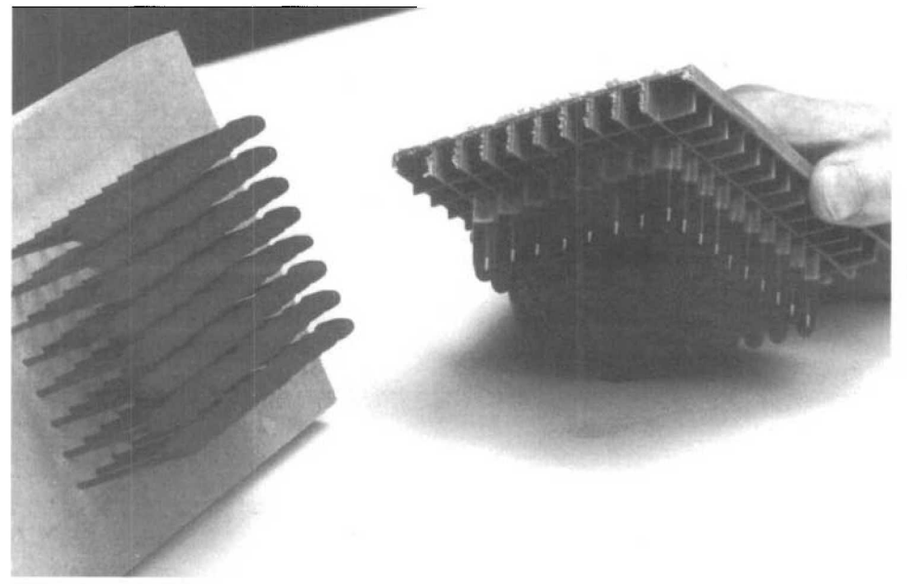


Fig. 6. *Lollipops – Frozen aqueous removed*

Concluding Remarks

Medicinal chemists are currently having to acquire a range of new skills to allow them to work as efficiently as possible in all aspects of the lead discovery and optimisation processes. There is currently emphasis on developing expertise in solid-phase synthesis to complement more traditional solution-phase skills. In future it will be imperative to be able to apply the most appropriate synthetic methods to address each particular requirement. Use of array formats both for lead generation and lead optimisation is rapidly becoming standard procedure and both highly automated and essentially manual methods find application. The introduction of combinatorial chemistry was initially driven by requirement for synthesis of larger numbers of compounds which is now established. Currently, the emphasis is being directed to improving the quality of materials produced using the new techniques and in the development of analytical methods.

It is also important to remember that synthesis is only one part of the drug discovery process. We also need to strive for the ability to generate smarter rather than larger libraries. Much effort is currently being applied to deriving methods for helping the design of potentially smarter libraries and to allow us to learn from the vast amount of data that is being generated from screening larger numbers of compounds. The value of any or all of these will only become apparent with time. We need to work to improve efficiency throughout all parts of the overall process and to have in place methods for evaluating those appropriate numbers of compounds which we will progress through the various stages of both *in vitro* and *in vivo* evaluation. Current developments in those areas are as fundamental and wide-ranging as the changes being introduced in chemistry.

Received: June 23, 1997

[1] J.W. Clitherow, P. Beswick, W.J. Irving, D.I.C. Scopes, J.C. Barnes, J. Clapham, J.D. Brown, D.J. Evans, A.G. Hayes, *Bio. Med. Chem. Lett.* **1996**, *6*, 833.

[2] N. Bailey, A.W. Dean, D.B. Judd, D. Middlemiss, R. Storer, S.P. Watson, *Bio. Med. Chem. Lett.* **1996**, *6*, 1409.

[3] M. Bush, T.J. Hamilton, J.C. Robins, S.H. DeWitt, CHI Conference on Exploiting Molecular Diversity, Coronado, California, February 1997.