OCCASION

This publication has been made available to the public on the occasion of the 50th anniversary of the United Nations Industrial Development Organisation.

DISCLAIMER

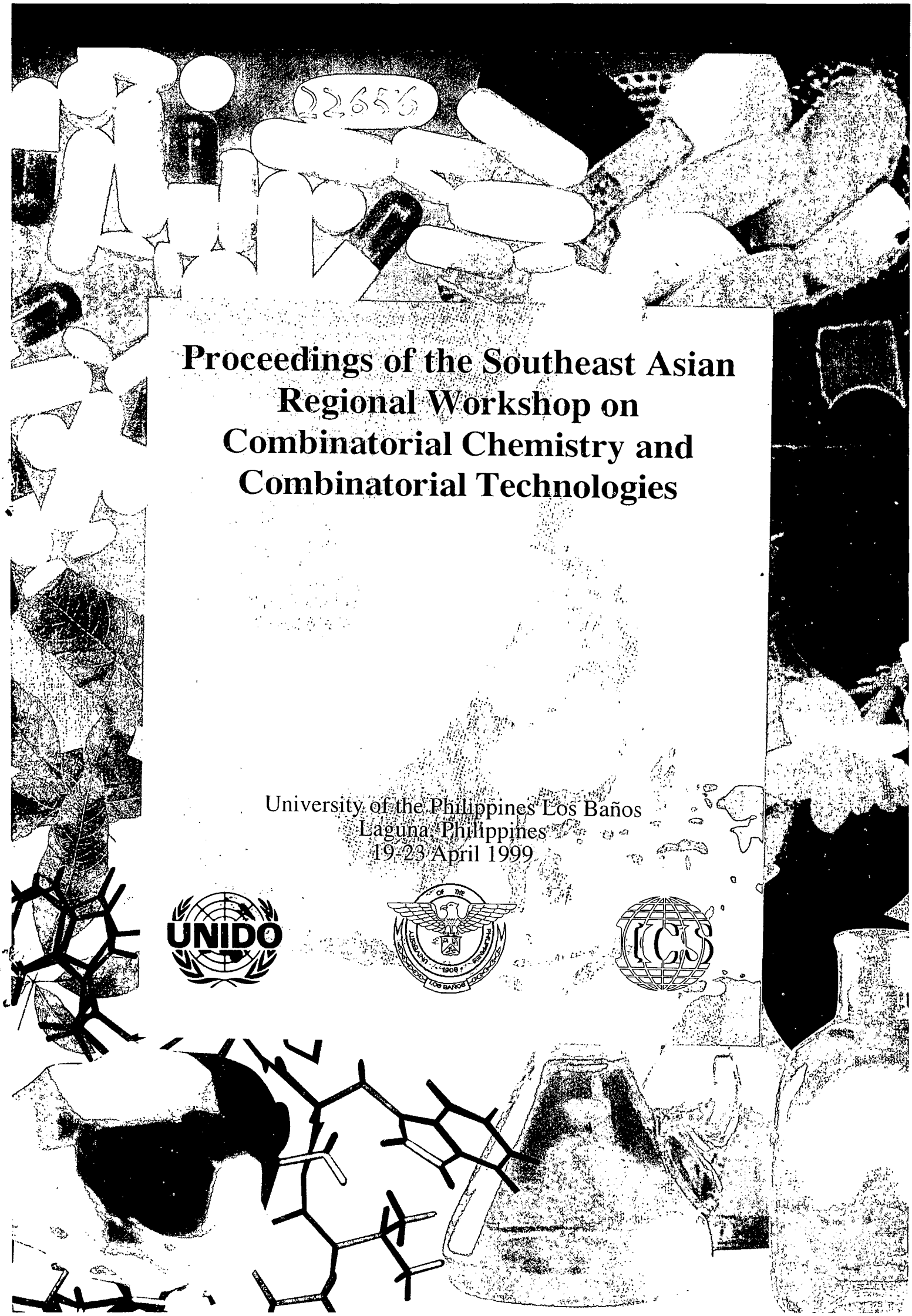
FAIR USE POLICY

CONTACT

Please contact publications@unido.org for further information concerning UNIDO publications.

For more information about UNIDO, please visit us at www.unido.org

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION
Vienna International Centre, P.O. Box 300, 1400 Vienna, Austria
Tel: (+43-1) 26026-0 · www.unido.org · unido@unido.org



Proceedings of the Southeast Asian Regional Workshop on Combinatorial Chemistry and Combinatorial Technologies

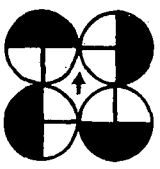
University of the Philippines Los Baños
Laguna, Philippines
19-23 April 1999

Organized by:

International Centre for Science and High Technology (ICS)
United Nations Industrial Development Organization (UNIDO)
Institute of Chemistry, University of the Philippines Los Baños (IC, UPLB)

Proceedings of the Southeast Asian Regional Workshop on Combinatorial Chemistry and Combinatorial Technologies

Editors:
Pierfausto Seneci, Stanislav Miertus, Giorgio Fassina and Cleofe A. Calanasan



This proceedings is published with support from the Philippine Council for Advanced Science and Technology Research and Development (PCASTRD) of the Department of Science and Technology (DOST), Philippines.

2000 ISBN 971-547-180-3

Preface

Combinatorial chemistry and combinatorial technologies (CC/CT) are new interdisciplinary fields that are radically changing the way the search for new drugs is being pursued. CC/CT offer rich possibilities not only for the pharmaceutical industry but also for fields like agricultural chemistry, biotechnology and new materials research and development, to name a few. Developing economies with vast natural resources, like those found in the Southeast Asian region, stand to gain from the application of these novel technologies in their drive to attain industrial competitiveness. And so to build awareness among researchers and technologies in the region about CC/CT, the Southeast Asian Regional Workshop on Combinatorial Chemistry and Combinatorial Technology was conceived during the Expert Group Meeting that I attended in Italy in June 1998.

In the summer month of April 1999, participants from the Southeast Asian region, Australia and Europe converged on the University of the Philippines Los Baños campus for the one-week workshop. The lectures delivered by top experts in CC/CT were followed by lively discussions among the participants. This proceedings documents the contributions of the participants and the other activities that transpired during the workshop.

I wish to thank the International Centre for Science and High Technology (ICS) and the United Nations Industrial Development Organization (UNIDO) for giving the Institute of Chemistry of the University of the Philippines at Los Baños the chance to organize and host this activity. I also wish to express my sincerest gratitude to the Department of Science and Technology (DOST) for the financial support that made possible the publication of this proceedings. I hope the participants also took away beautiful memories of my country and its people. Lastly, my deepest gratitude goes to the staff of the Institute of Chemistry for putting together an activity that was intellectually rewarding and gave everyone a most enjoyable time.

Norma N. Fajardo
Institute of Chemistry

December 1999

CONTENTS

Southeast Asian Regional Workshop on Combinatorial Chemistry and Combinatorial Technologies, University of the Philippines Los Baños, Laguna, Philippines, 19-23 April 1999

An Introduction to ICS-UNIDO Programme

The United Nations Industrial Development Organization is a specialized agency of the United Nations dedicated to promoting sustainable industrial development in developing countries and countries with economies in transition. It harnesses the joint forces of government and the private sector to foster competitive industrial production, develop international industrial production, develop international industrial partnerships and promote socially equitable and environmentally friendly industrial development.

UNIDO is the only worldwide organization dealing exclusively with industry from a development perspective. UNIDO's services are non-profit, neutral and specialized. UNIDO acts as a catalyst to help generate national economic wealth and raise industrial capacity through its role as a worldwide forum for industrial development and as a provider of technical cooperation services. UNIDO's ultimate goal is to create a better life for people by laying industrial foundations for long term prosperity and economic strength.

The International Centre for Science and High Technology is an Institution within the legal framework of UNIDO with headquarters located in Trieste, Italy. The Centre's mandate relates to the transfer of know-how and technology in favour of developing countries, and is justified by the perception that a competitive industrial technological capability cannot be built-up without adequate scientific knowledge and commitment to a sustainable development approach utilizing new and environment friendly technologies.

The activities of ICS follow an integrated pragmatic approach which include action-oriented research, short-term exchange between research and technologists in industry, dissemination of scientific and technology information through the creation and management of centres of excellence (focal points), consultancy and advisory services, training courses, scientific workshops, high level seminars, study tours, fellowships, promotion of training arrangements, publication and editing of frontier issues.

In the present work programme the ICS's activities focus to specific sectors within the area of chemistry, environment, new materials and high technology. In selecting the specific subprogrammes and their related activities special consideration was given to their relevance in relation to the scientific and technological development of developing countries.

Considering that sustainable development depends upon the harmonization of economic growth and environment conservation and protection, the ICS Area of Pure and Applied Chemistry has identified as priority fields in its work programme the following themes, which are of key relevance to economic and industrial development as well as environmental protection:

Catalysis, which is an important scientific and technological area for the development of environmentally friendly chemical processes, which in turn form the basis for cleaner industrial development and are also the key elements for an industrial prevention approach. New, less pollutant processes together with the optimization of existing processes depend to a great extent upon the improvement of catalyst performance in the heavy and fine chemical production lines with a direct impact on the quality and quantity of by-products or waste generated.

Environmentally Degradable Plastics, where the expanding global production and consumption of polymeric materials coupled with increasing public awareness of environmental issues have created serious concern about the problems related to the disposal of plastic waste generated by various sectors of human activity. Besides recycling, reuse, incineration and composting, new technological

developments of environmental degradable plastics contribute dramatically to the tackling of the environmental issue in specific sectors of plastics use.

Remediation Technologies, which are becoming an important and economical way to solve the problem of contaminated and polluted sites, especially in developing countries and economies in transition where the environmental issue has been until recently neglected. New technologies, methodologies and solutions are emerging from various applications and are becoming day by day more economically viable and feasible.

Combinatorial Chemistry, Combinatorial Technologies and Molecular Design which have a strong impact on the development of new chemicals (pharma industries, agro-chemicals, new materials). Developing countries need to get acquainted with and gain expertise in combinatorial technologies and molecular design to help local enterprises remain competitive and economically viable in the coming decades. Combinatorial chemistry and combinatorial technologies have a potential influence not only on industrial growth, but also on environment protection. In fact, by optimizing industrial processes and production, with the lowering of relevant costs, smaller amounts of waste and by-products are created.

Combinatorial chemistry and combinatorial technologies fall also into a new interdisciplinary field joining combinatorial informatics with automated synthesis of chemical "libraries" followed by automated screening, with the main output in medicinal chemistry and drug discovery. This nascent technology has produced more new compounds in a few years than the pharmaceutical industry did in its entire previous history.

Combinatorial methods are not restricted to pharmaceutical applications. Whenever a large number of compounds have to be prepared for testing, this technique can be used. Additional fields of application include agricultural research. However, for the time being, the main emphasis is on pharmaceutical research, and most major pharmaceutical companies are active in the field. It is generally accepted that the methods offer great potential for the lead finding and drug discovery process, and the technologies are expected to contribute to the reduction of time and costs.

Because of this important role of combinatorial chemistry, combinatorial technologies and molecular design, ICS-UNIDO has developed a programme in this field that includes the organization of awareness building events, projects promotions, publication activities, etc. The training courses, workshops and Expert Group meetings organized by ICS in this field are the following:

A TC on "Industrial Applications of Structure-Based Molecular Design and Combinatorial Chemistry", held on 28 July – 8 August 1997 in Kuala Lumpur, hosted and co-organized by the University of Malaya.

A TC on "Methodologies, Applications and Economics of Combinatorial Chemistry and Combinatorial Technologies", held in Piana di Monte Verna (CE) on 8-19 September 1997, hosted and co-organized by TECNOGEN;

A Workshop on "Applications of Molecular Design and Computer-Assisted Combinatorial Chemistry" held in Cape Town on 29 March-4 April 1998, hosted and co-organized by the University of Cape Town, Department of Chemistry.

An EGM on "Combinatorial Chemistry, Combinatorial Technologies and Molecular Design", held in Trieste, Italy on 15-17 July 1998.

A Workshop on "Combinatorial Technologies: Awareness and Familiarization for Decision Makers" and a training Workshop "Methods, Applications and Economics of Combinatorial Chemistry and

Combinatorial Technology", held in Hyderabad, India 23-30 October 1998, hosted and co-organized by the Council of Scientific and Industrial Research (CSIR), Industrial Institute of Chemical Technology (ICT).

A Workshop on "Combinatorial Chemistry and Combinatorial Technologies", held in Buenos Aires, Argentina on 7-11 December 1998, hosted and co-organized by INGEBI, University of Quilmes.

Several internationally recognized experts in the field of combinatorial chemistry and combinatorial technologies have been involved in these events and they also participated in the launching of ICS-UNIDO projects on the implementation of combinatorial chemistry and combinatorial technologies in industries of various developing countries.

The first activity organized in 1999, within the ICS-UNIDO Subprogramme of Combinatorial Chemistry was the "Southeast Asian Regional Workshop on Combinatorial Chemistry and Combinatorial Technologies" held in the Philippines, co-organized and hosted by the University of the Philippines Los Baños; the workshop specifically focused on potential applications of combinatorial technologies in the field of exploitation of naturally occurring compounds, and its objectives were the following:

- To build awareness among researchers, technologists, and academicians on the development of combinatorial chemistry and combinatorial technologies (CC/CT)
- To update participants on the principles and industrial applications of CC/CT
- To evaluate possible initiatives (as follow-up projects and feasibility studies) of industry and universities in Southeastern Asia, regarding CC/CT development and industrial implementation with the focus on the application of CC/CT to natural products research and exploitation.
- To set up a regional ICS-UNIDO network on combinatorial chemistry and combinatorial technologies.

27 selected participants, representing academia, industry and governmental institutions from Hong Kong, India, Malaysia, Philippines, Taiwan, Thailand, Vietnam, took part in the event, together with 5 recognized experts in this field.

The present publication is a collection of the contributions of the main lecturers and of selected country/institution reports presented by participants.

I would like to thank the Organizing Committee, which, under the chairmanship of Professor Norma Fajardo, did an excellent work. My thanks are also due to all the lecturers and authors of the papers in this volume, for their valuable contributions.

The efforts of Dr. P.F. Seneci, ICS-UNIDO, Scientific Adviser, for the preparation of the scientific programme of the workshop are highly appreciated, as well as the precious contribution of Dr. C.A. Calanasan for the preparation of this publicaition.

Trieste, June 1999

Stanislav Miertus
ICS Area Coordinator
Pure and Applied Chemistry

Combinatorial Chemistry: an Overview

Dr. Pierfausto Seneci
Head of Lead Discovery Department
GlaxoWellcome Medicines Research Centre
Via Fleming 4, 37100 Verona, ITALY

Introduction

Among the new technologies appeared recently in Chemical Sciences a prominent position is surely occupied by Combinatorial Chemistry. While many excellent reviews covered this subject [1], here we will briefly provide the reader with the basic principles which constitute the foundation of this new technology. We will track the history of Combinatorial Chemistry and the reasons which made it more and more popular among scientists. We will define the most important concepts and techniques used in Combinatorial Chemistry. We will highlight the usefulness and the properties of primary and secondary libraries, of solid phase and solution phase Combinatorial Chemistry, of discrete and pool libraries, with a particular emphasis on the applications for each library format.

History and Definitions

The concept of Combinatorial Chemistry appeared in literature in the late 80s, even though there are instrumental papers [2], related to peptide chemistry, that allowed the exploitation of synthetic techniques from early 80s. The immediate application of these techniques was mostly related to oligomeric chemistries (peptides, oligonucleotides) and to pharmaceutical purposes, while the first report of Combinatorial Chemistry applied to small organic molecules appeared later [3]. From then on, Combinatorial Chemistry expanded greatly its synthetic applications *via* the exploitation of many chemical reactions for library production [4] and *via* its growing popularity in non-pharma applications such as materials sciences [5], coordination chemistry [6] and catalysis [7].

A few definitions given up front will be very useful to understand the basic principles of Combinatorial Chemistry:

- a **combinatorial library** is a single entity composed of many individuals (hundreds to millions) which can be prepared in many formats and using different synthetic techniques. The library is tested for a specific activity and its active individuals, or *positives*, are identified.

- a **synthetic organic library** is prepared using standard organic chemistry, either in solution or on solid phase (SP).

- a **biological library** is usually composed of natural oligomers and is prepared by natural means (e.g. phage display libraries) or through the use of biological or biochemical reagents (e.g. enzymes and oligonucleotide amplification systems).

- a **materials science library** is made from inorganic compounds and is prepared using techniques peculiar to this field such as sputtering deposition of thin films, electron beam evaporation and moving masks techniques.

- a **scaffold** is the common structural element contained in all the library individuals: it may be an iterated chemical bond, *e.g.* in oligomeric libraries such as peptides it is a repeating amide backbone; it could be a functional group *e.g.* in libraries of substituted guanidines; it could be a ring

motif, *e.g.* in libraries of substituted benzodiazepines as in Fig. 1. A **building block** is a reagent used during the synthesis of a library (e.g. the lithium salt of acetanilide, Fig. 1).

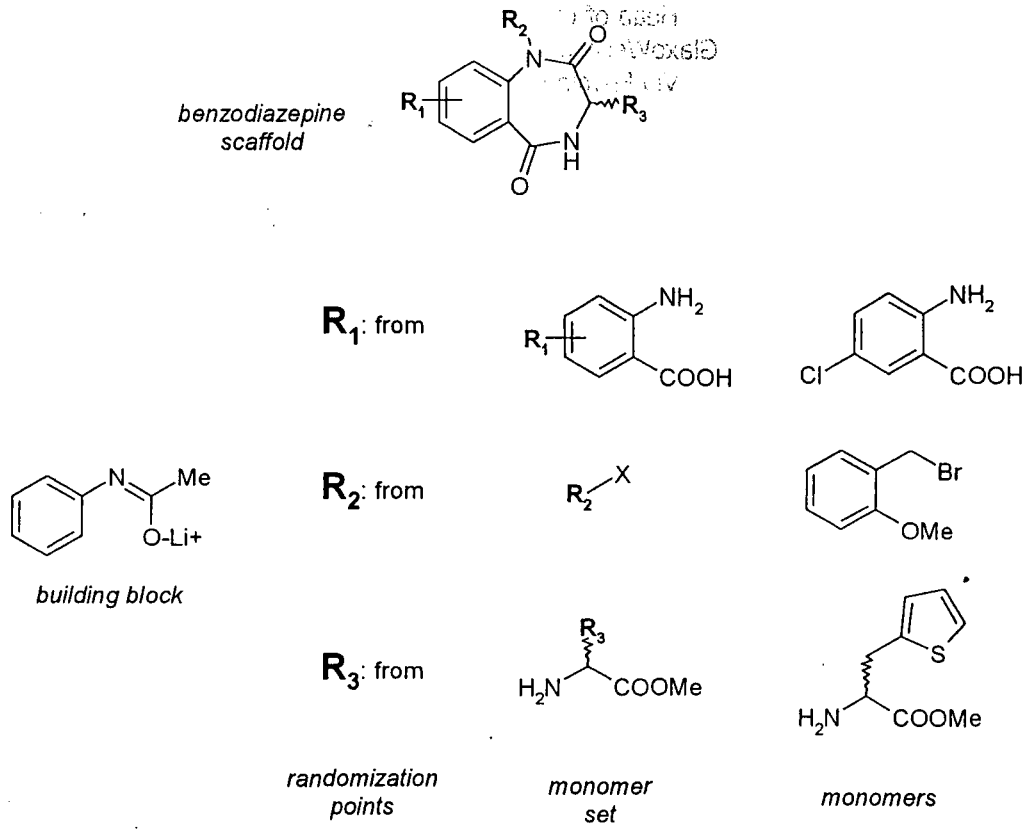


Figure 1

- a **monomer set** is a class of reagents with a common functional group which is either used to produce substituted scaffolds or to decorate pre-existing scaffolds (e.g. substituted anthranilic acids, alkyl halides and α-amino acids, Fig. 1). A **monomer** is a reagent which is part of a monomer set (e.g. 5-chloroanthranilic acid, 2-methoxybenzyl bromide and 2-thienylalanine methyl ester, Fig. 1).

- a **randomization point** ($\mathbf{R_1}$, $\mathbf{R_2}$ and $\mathbf{R_3}$, Fig. 1) is a position where a monomer can be inserted into a library during the construction of the scaffold or where it can be coupled to a pre-existing scaffold in order to produce library components containing all of the possible combinations of the selected monomers.

- **diversity** is a concept unrelated to the library size which attempts to evaluate the representation of chemical space by a chemical library using computational methods. If this space is sampled evenly by the components of a library then this library is considered to be diversity-based (Fig. 2, left). **Similarity** is a concept unrelated to the library size which is opposite to diversity: if the library components are clustered around the model structure **A**, the library is similarity-based (see Fig. 2, right).

- **chemical assessment** is the process where a known reaction in classical organic chemistry is combinatorialized. This may include the transfer of the reaction on solid phase and/or the adaptation of the reaction conditions to the use of many monomers with different reactivities and stabilities.

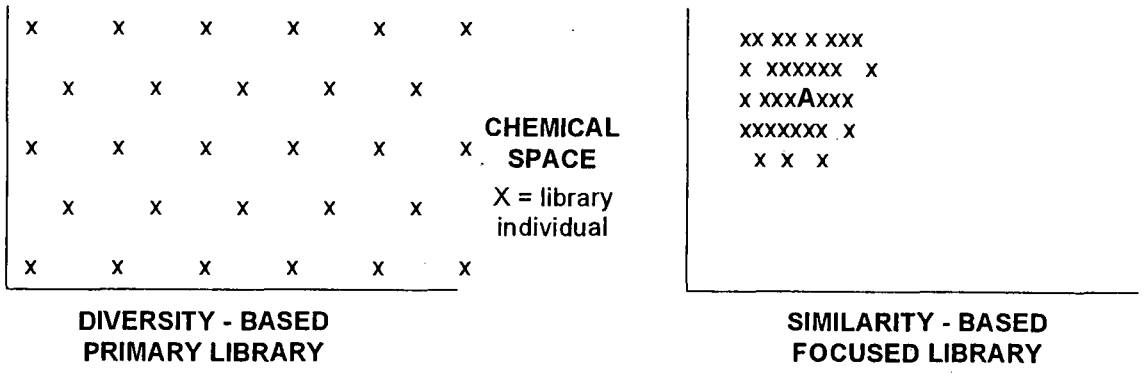


Figure 2

- **monomer rehearsal** is an accurate check of the reactivity of a monomer set in the synthetic scheme for the build up of the library so that the unreactive/difficult monomers are removed from the set.

- a **model library** is a small set of discretes, or a pool, which is prepared using the planned synthetic route for the library and is fully characterized by the appropriate analytical tools; only if the results are satisfactory the library synthesis is carried out.

- **quality control** is the process where the analytical profile of a library is determined. The analytical characterization is made on the library as a single entity, but data regarding each individual are acquired; a library with 80% confirmed pure compounds is a good quality library, but the structures of the non confirmed, or impure, 20% of compounds are known and the whole library is tested for positives' identification.

- **structure determination** is the process in which an observed activity is related to a library component through **library deconvolution** [8] (deciphering the library complexity via resynthesis of smaller pools containing "positives") or through **decoding** [9] (structure determination of a "tag" linked to the solid support which carried a positive and encoding for a single library component structure).

Library Formats: Primary (Diversity-Based) and Secondary (Focussed) Libraries

Combinatorial libraries can be made for many different purposes, and a careful selection of the best available library format for a specific application is often the key to the successful identification of positives from the library.

Libraries can be divided into **primary**, or diversity-based, or unbiased, and **secondary**, or focussed, or biased libraries. The main features of, and differences among the two categories are reported in Fig. 3.

◆ PRIMARY: hit	◆ SECONDARY: lead
◆ random	◆ focused
◆ solid phase	◆ solid/solution phase
◆ mix and split	◆ parallel synthesis
◆ >10000 cpds	◆ <10000 cpds
◆ <1mg/cpd	◆ >1mg/cpd
◆ biological assay on many targets	◆ biological assay on one target

Figure 3

Primary libraries are typically prepared to identify one, or more, positives on an unknown target; the purpose of the library components, then, is to be as "diverse" as possible and to cover a wide structural space with scattered structures. This will eventually lead to a mild positive, or **hit**, out of the large library population. This structure is now the starting point to design a secondary, or biased, library where the purpose will be to prepare analogs of the hit so to understand the structure-activity relationships (SAR) of the class of compounds. From this library a **lead**, that is a valuable molecule with a desired activity profile, will come out and in pharmaceutical research will be the basis for a project potentially leading to a new marketable drug.

Primary libraries are typically large (>10000 compounds) to span more diversity and prepared on solid phase as mixtures using the "mix and split" technique [10] (see the next paragraph) to increase productivity and decrease the number of assays. Focussed libraries are smaller (typically hundreds to few thousands), they are prepared both on solid phase and in solution and are finally obtained as discretes to have more reliable activity results on the analogs of the original hit. Primary libraries are normally tested on many assays, so to capitalize the synthetic efforts by finding more hits on various targets, while the focussed ones are tested only on the assay where their parent hit showed activity in the first place.

In general, primary libraries are used in pharmaceutical industry for the so-called exploratory phase, to reduce the time needed for the identification of a positive on a poorly known target. Focussed libraries are used to speed up the SAR data acquisition via parallel synthesis of a large number of hit analogs. This distinction is valid also for all the non-pharma applications of Combinatorial Chemistry.

Discrete Libraries (Parallel Synthesis) and Pool Libraries (Mix and Split)

Another important distinction is among libraries where the compounds are prepared as **mixtures** (sublibrary sets called pools) or as **individuals** (discrete libraries).

The latter libraries are made "simply" by automating the classical organic chemistry so to be able to handle many reaction vessels in parallel (up to tens of thousands); the name of parallel synthesis comes from this concept. While the related automation aspects and other features of these libraries are of great importance, we will focus from now on pool libraries, introducing the so called "mix and split" [10] solid-phase technique which is illustrated in Fig. 4.

A hypothetical 3-step solid phase library of 3 x 3 x 3=27 members is prepared by first splitting the resin in three equivalent portions, then adding a different monomer of the subset **A** to each portion. The three portions are then mixed and split again in three portions, now containing similar amounts of $\mathbf{A_1}$-$\mathbf{A_3}$ monomers but with a unique monomer on each resin bead. Repeating this procedure with

monomers of the subset **B**, three pools where **B** is determined and **A** is randomized are obtained. Another "mix and split" step followed by addition of the subset **C** produces three pools, each containing nine compounds; monomer **C** is determined while the other subsets are fully represented (Fig. 4). This technique allows the preparation of mixtures of compounds with two important features: 1) *all the library components are represented in similar amounts* and 2) *each resin bead carries only one library component* because each coupling is performed separately on an equivalent amount of resin.

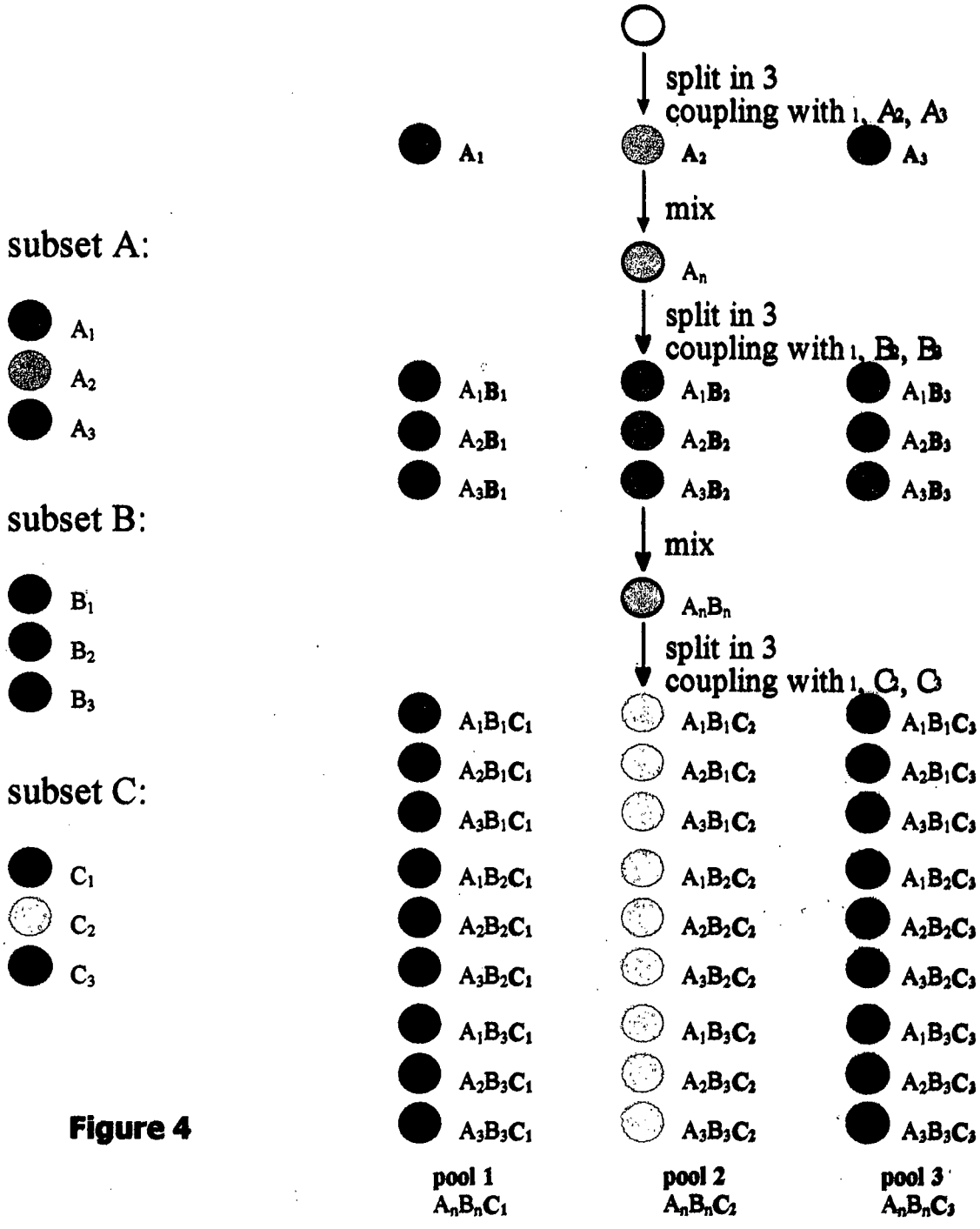


Figure 4

The main features of both discrete and pool libraries are reported in Fig. 5.

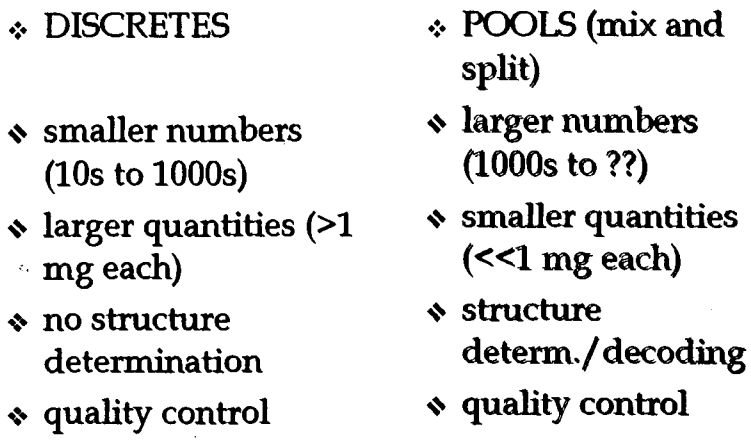


Figure 5

Discrete libraries require the parallel handling of many reaction vessels, thus limiting even with a high degree of automation the number of individuals to a maximum of few thousands; the quantity of each library component can easily be in the few mg-range. On the other side, pool libraries are theoretically unlimited in terms of numerosity (a library where only one bead carries a library component and made with 100g of resin would contain around 100.000.000 individuals!), but the amount of each individual is much smaller, typically much less than one mg of final compound.

Discrete libraries do not require structure determination, because each library well contains a single compound and its structure is known *via* the position of the reaction vessel (spatial encoding). This does not require any effort to link the activity observed for a library well to the positive structure. Pool libraries, though, require the determination of positives into a library pool *via* various processes either related to the deconvolution of the library complexity or to the decoding of suitable tags anchored to the solid support (see above). Finally, both library formats require a good quality control to produce reliable results in terms of identification of positives.

Solid Phase and Solution Phase Libraries

While Combinatorial Chemistry was for a long time intended mostly as a solid-phase related technique, as of today both solid phase [11] and solution phase [12] libraries are widely used for various purposes. The main features and advantages of these formats are highlighted in Fig. 6.

Synthesis of libraries on solid phase allows the use of the "mix and split" technique to make pool libraries with all the advantages mentioned in the previous paragraph. The synthesis of pool libraries in solution can be made [13], but with more difficulties and with by far lower flexibility. The use of solid phase anchors each reaction intermediate and the final product on the support, thus allowing the use of large excess of reagents in solution which are then simply washed away. This simplifies the work-up after each step, allowing automated washing procedures and eventually the creation of solid phase multi-step combinatorial libraries with a large population of individuals.

The synthesis of solution phase libraries, though, has other advantages. First, all the organic reactions developed in solution are immediately transferable to a solution phase library synthesis, while the transfer from solution chemistry to solid phase requires a chemical assessment which is often long and effort-requiring. Many reactions which are assessed in solution cannot be moved onto solid phase for incompatibility either with the polystyrenic supports, or with the linkers used, or for other reasons.

Solution phase libraries do not require additional complexity, such as the linker hooking the library intermediates to the resin and the cleavage to free the library components at the end of the synthesis. Finally, synthesis in solution can more easily tailor the final quantities of library products to the chemist's needs.

ON SOLID PHASE	IN SOLUTION
large excess of reagents allowed	all organic reactions can be used
multistep synthesis allowed	no chemistry assessment
easy workup-isolation	no linker/cleavage chemistry
mix and split possible	unlimited product quantities

Figure 6

In conclusion, the perfect library format does not and will never exist. Rather, each specific purpose should point the combinatorial chemist towards the better library format (solution or solid phase, primary or secondary, discrete or pool, and so on) and this choice should prove crucial in the final outcome of library testing.

References

[1] a) F. Balkenhohl, C. von den Bussche-Hannefeld, A. Lansky, C. Zechel, *Angew. Chem. Int. Ed. Eng.*, 1996, **35**, 2289-2337; b) L.A. Thompson, J.A. Ellman, *Chem. Rev.*, 1996, **96**, 555-600.

[2] a) H.M. Geysen, R.H. Meloen, S.J. Barteling, *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 3998-4002; b) R.A. Houghten, *Proc. Natl. Acad. Sci. USA.*, 1985, **82**, 5131-5135.

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

[4] a) P.H.H. Hermkens, H.C.J Ottenheijm, D. Rees, D. *Tetrahedron*, 1996, **52**, 4527-4554; b) P.H.H. Hermkens, H.C.J Ottenheijm, D. Rees, D. *Tetrahedron*, 1997, **53**, 5643-5678.

[5] W.H. Weinberg, B. Jandeleit, K. Self, H. Turner, *Curr. Opin. Solid State Mater. Sci.*, 1998, **3**, 104-110.

[6] M.T. Burger, W.C. Still, *J. Org. Chem.*, 1995, **60**, 7382-7383.

[7] a) K. Burgess, A.M. Porte, *Adv. Catal. Processes*, 1997, **2**, 69-82; b) S.J. Taylor, J.P. Morken, *Science*, 1998, **280**, 267-270.

[8] P. Seneci, in **Combinatorial Chemistry and Combinatorial Technologies: Methods and Applications** (S. Miertus, G. Fassina Eds.), 91-125, Marcel Dekker, Inc., New York, 1998.

[9] P. Seneci, in **Combinatorial Chemistry and Combinatorial Technologies: Methods and Applications** (S. Miertus, G. Fassina Eds.), 127-167, Marcel Dekker, Inc., New York, 1998.

[10] a) A. Furka, F. Sebestyen, M. Asgedom, G. Dibo, *Int. J. Pept. Protein Res.*, 1991, **37**, 487-493; b) K.S. Lam, M. Lebl, V. Krchnak, *Chem. Rev.*, 1997, **97**, 411-448.

[11] I. Sucholeiki, in **Combinatorial Chemistry: Synthesis and Application** (S.R. Wilson, A.W. Czarnik Eds.), 119-133. John Wiley & Sons, New York, 1997.

[12] a) R. Storer, *Drug Disc. Today*, 1996, **1**, 248-254; b) D.M. Coe, R. Storer, in **Annual Reports in Combinatorial Chemistry and Molecular Diversity** (W.H. Moos, M.R. Pavia, B.K. Kay, A.D. Ellington Eds.), 50-58. ESCOM Science publishers, Leiden, 1997.

[13] H.Y. An, L.L. Cummins, R.H. Griffey, R. Bharadwaj, B.D. Haly, A.S. Fraser, L. Wilson-Lingardo, L.M. Risen, J.R. Wyatt, P.D. Cook, *J. Am. Chem. Soc.*, 1997, **119**, 3696-3708.

Principles, Applications, and Economics of Combinatorial Technologies

Dr. Giorgio Fassina
TECNOGEN S.C.P.A., Parco Scientifico
81015 Piana di Monte Verna (CE), ITALY

Introduction

The time and cost needed for the development of new drugs have increased steadily during the past three decades. Estimated costs for introducing a new drug in the market now reach around 200-300 millions USD, and this process takes around 10-12 years after discovery. This increase in time and cost is due mainly to the extensive clinical studies of new chemical entities required by competent regulatory agencies, such as the FDA, and to a lesser extend to the increased costs associated to research. The time and cost required for clinical and preclinical evaluation of new drugs is not likely to decrease in the near future, and as a consequence, a key issue for pharmaceutical companies to stay in the market has been to increase the number of new drugs in the development pipeline.

Drug discovery in the past has been based traditionally on the random screening of collection of chemically synthesized compounds or extracts derived from natural sources, such as microorganisms, bacteria, fungi, plants, of terrestrial or marine origin or by modifications of chemicals with known physiological activities. This approach has resulted in many important drugs, however the ratio of novel to previously discovered compounds has diminished with time. In addition, this process is very time consuming and expensive. A limiting factor was linked to the restricted number of molecules available or extract samples to be screened, since the success rate in obtaining useful lead candidates depends directly from the number of samples tested.

Chemical synthesis of new chemical entities often is a very laborious task, and additional time is required for purification and chemical characterization. The average cost of creating a new molecular entity in a pharmaceutical company is around 7500 USD/compound [1]. Generation of natural extracts, while very often providing interesting new molecular structures endowed with biological properties, leads to mixtures of different compounds at different concentrations, thus making activity comparisons very difficult. In addition, once activity is found on a specific assay, the extract needs to be fractionated in order to identify the active component. Quite often, the chemical synthesis of natural compounds is extremely difficult, thus making the lead development in to a new drug a very complex task. While the pharmaceutical industry was demanding more rapid and cost effective approaches to lead discovery, the advent of new methodologies in molecular biology, biochemistry, and genetics, leading to the identification and production of an ever increasing number enzymes, proteins, receptors, involved in biological processes of pharmacological relevance, and good candidates for the development of screening assay, complicated even more this scenario.

The introduction of combinatorial technologies provided an unlimited source of new compounds, capable to satisfy all these needs. This approach was so appealing and full of promises that many small companies started to flourish financed by capitals raised from private investors.

Principles

Combinatorial approaches were originally based on the premise that the probability of finding a molecule in a random screening process is proportional to the number of molecules subjected to the screening process. In its earliest expression, the primary objective of combinatorial chemistry focused

on the simultaneous generation of large numbers of molecules and on the simultaneous screening of their activity. Following this approach, the success rate to identify new leads is greatly enhanced, while the time required is considerably reduced.

The development of new processes for the generation of collection of structurally related compounds (libraries) with the introduction of combinatorial approaches has revitalized random screening as a paradigm for drug discovery and has raised enormous excitement about the possibility of finding new and valuable drugs in short times and at reasonable costs. However the advent of this new field in drug discovery did not obscure the importance of "classical" medicinal chemistry approaches, such as computer-aided rational drug design and QSAR for example, but catalyzed instead their evolution to complement and integrate with combinatorial technologies.

The word "**combinatorial**" appeared in the scientific literature at the beginning of the '90s, but the generation of the first combinatorial libraries can be dated back to the beginning of the '80s. The first reports dealt with the simultaneous production of collection of chemically synthesized peptides, produced by solid phase methods on solid supports [2-6]. Peptides were particularly suited for combinatorial synthesis given the well-established synthetic protocols available, the great number of different molecules attainable, and the potential to generate leads of biological and pharmaceutical value. The use of peptide libraries was greatly accelerated by the introduction of biological methods for library preparation, by the use of the phage display technology, which provided interesting advantages over the synthetic counterpart [7,8]. At the same time, the first papers on the generation of oligonucleotide libraries appeared in the literature [9,10], thus suggesting the possibility to extend the applicability of combinatorial approaches even to other classes of synthetic or natural oligomeric compounds, such as carbohydrates. There are many important biologically active glycoconiugate drugs whose carbohydrate constituents are associated with the molecular mechanism by which these drugs exhibit their effect. With these drugs exploration of carbohydrate molecular diversity has the potential for identifying novel agents with enhanced potency. As a conformationally rigid and functionally rich system, carbohydrates also provide valuable molecular scaffold systems around which to generate primary screening libraries.

A broad variety of new synthesis and screening methods are currently grouped under the term combinatorial. These methods include parallel chemical synthesis and testing of multiple individual compounds or compounds mixtures in solution, synthesis, and testing of compounds on solid supports, and biochemical or organism-based synthesis of biological oligomers coupled to selection and amplification strategies. Combinatorial technologies merged different disciplines such as solid phase and solution phase chemistry, analytical chemistry, molecular biology, molecular design, automation and miniaturization in an integrated platform technology. The philosophy of combinatorial technologies is to make rational the random approach to drug discovery. The main advantage of using combinatorial technologies is the speed in finding and optimizing useful leads. The disadvantage is that it is impossible to explore the entire chemical space in the combinatorial format, i.e. not all chemical structures can be produced by combinatorial approaches.

Applications

Many active compounds have been selected to date following combinatorial methodologies, and a considerable number of those have progresses in to clinical trials. However, combinatorial chemistry (CC) and related technologies for producing and screening large number of molecules find useful applications also in other industrial sectors not necessarily related to **pharmaceutical industry**. Emerging fields of application of combinatorial technologies are the **diagnostic**, the **down-stream processing**, the **catalysis**, and the **new material** sectors.

In **diagnostics**, CC can be successfully applied to the identification of previously unknown epitopes recognized by antibodies in biological fluids associated to pathological conditions. The selected epitopes can be used then to for the development of diagnostic kits useful for the identification and quantification of the antibody of interest. In the **down stream processing field**, combinatorial chemistry finds application in the selection of ligands able to specifically recognize macromolecules of biotechnological interest, such as proteins, antibodies, or nucleic acids. This is of relevant industrial importance, since the major costs associated to the production of recombinant molecules for therapy is associated to the purification of the desired target molecule from crude feed-stocks. The availability of specific and selective ligands to be used in affinity chromatography for the capture and concentration of the target from crude samples will reduce considerably production costs of biopharmaceuticals, such as monoclonal antibodies [11]. Combinatorial technologies have been applied also to the identification of new macromolecules endowed with catalytic activity for reactions where natural enzymes are inactive. This application even if still at the early stage, is calling considerable attention from the industrial sector, since the availability of new enzymes may reduce the production costs of many chemicals.

The different technologies and strategies used in the production of combinatorial libraries are now so well developed that is easy to plan synthetic schemes for the generation of a huge number of compounds. Since the rate at which compounds can be screened does constitute a limitation to the use of combinatorial technologies, it is important to be selective about the compounds which are synthesized. Computational methods are very valuable from this point of view to assist in the design of combinatorial libraries. The main requirement for lead generation is often to maximize the range of structural types within the library with the expectation that a broad range of activities will result. As a consequence, diversity analysis is an important aspect of library design. The diversity of libraries may be measured by the use of similarity or dissimilarity indexes which make intermolecular comparisons possible. Measures of chemical similarity have been developed for similarity searching in chemical databases. The calculation of the similarity between two molecules involves the characterization of the molecules by using chemical/structural descriptors, and then the application of similarity coefficients to quantify the similarity.

In combinatorial chemistry, due to the high number of chemical manipulations required to synthesize libraries of compounds automation is unavoidable. Many research groups, both in academia and industrial settings are developing automated instruments specifically tailored to these needs, and this technology field is acquiring an extremely important role for the development of combinatorial technologies for the next millennium. On the other hand, the huge number of compounds produced simultaneously with these technologies requires automation also in purification protocols, quality assessment, sample dispensing and testing. In addition, the ever increasing number of compounds generated by combinatorial technologies pushes towards miniaturization of screening assays, in order to handle an increasing number of tests at the same time with little consumption of reagents.

The rapidity of new chemical entity generation and screening allows validation of molecular targets associated to diseases in short time. This is a very important emerging trend in combinatorial technologies, since the advent of new methodologies in molecular biology, biochemistry, and genetics, leads to the identification of many factors which should be screened quickly in order to define their relevance to biomedical processes. With the increased speed at which new drug entities are now synthesized and evaluated for pharmacological activity, a need has arisen to provide fundamental metabolism data at the early stages of drug discovery. Strategies are being developed to permit drug metabolism data to be an important part of early drug discovery. Many important properties of drugs related to metabolism could be the deciding factor in whether or not a compound is selected for clinical development, and application of combinatorial approaches to such assessments is emerging as a new trend of application.

Economics

Many companies were founded at the end of the 80's using venture capitals, such as Affymax (Palo Alto, CA), Aptein Corp. (Seattle, WA), Darwin Molecular (Seattle, WA), Gilead Sciences (Foster City, CA), Houghten Pharmaceuticals (San Diego, CA, now Trega), Isis Pharmaceuticals (Carlsbad, CA), Ixsys Inc. (San Diego, CA), NeXstar Pharmaceuticals (originally NeXagen, Boulder, CO) Protein Engineering Corp. (Cambridge, MA, now Dyax), Selectide Corp. (Tucson, AZ), and Sphinx Pharmaceuticals (Durham, NC). The creation of small and medium sized companies centered around newly introduced technologies is a well known and recurrent phenomena in the biotech world, but differently than in other sectors in this case the majority of such companies were acquired by big pharmaceutical companies.

In 1994, Eli Lilly & Co. (Indianapolis, IN) bought Sphinx for $80 million. This was followed by Marion Merrell Dow (Kansas City, MO) spending $58 million for Selectide in 1995. The last, and also one of the biggest in all of biotech's history, major acquisition was Glaxo's (London, UK) investment of $533 million in 1995 for Affymax. The venture capital raised during the 90's to set up companies focused on combinatorial chemistry has seen to date an overall increase of about 65-fold. The reasons for such growing confidence in the field's overall development are based on several peculiar aspects of combinatorial technologies. First of all, combinatorial technologies can produce many different products and also many different kinds of products. The products are all valuable and useful and can reach the marketplace relatively quickly, because they are geared towards specific applications, and address vast markets in the human, veterinary and also agro sectors.

The additional parameter indicating the health of the combinatorial business is the increase in the market capitalization over time of participating companies in the public capital markets. This is the reason why so many companies, in excess of 180, are involved in the field. These companies can be divided into four major categories, depending on their use of combinatorial technology: **library makers**, **library value-adders**, **library users**, and finally **hardware/software developers**. The industry is therefore fragmented, with no clear leadership position enjoyed by any single company. This is also probably the reason why the field is so rich in alliances and collaborations, whose value exceeds a few billion dollars, and which occur among at least 130 of the 180 companies involved in combinatorial chemistry.

An important reason for the rapid development of this field is its unique patent history. There are no strategic patents in these fields, but many companies own patents that enable them to pursue unique chemistries, thus also new entrants have the peculiar opportunity to get access to this field with no major limitations.

Combinatorial technologies have been developed at the beginning by biotech companies, mainly because the availability of recombinant macromolecules such as proteins, receptors, antibodies, and enzymes, in a purified form, endowed of important biological functions, constituted the most appropriate targets for screening libraries of compounds. Nowadays there is no shortage of therapeutic targets, given the ever-growing number of new biological targets emerging from biotechnology and the human genome project. But to gain access to combinatorial technologies, pharmaceutical companies basically have three choices. They can seek collaborative agreements, which are typically limited to screening one target or family of targets against a set of libraries. They can develop combinatorial chemistry capability in house. Or they can acquire a firm with established expertise in combinatorial chemistry, or even a part of it.

Part of the reason for the steady increase in the capital raised during the 90's to fund combinatorial companies is, of course, the arrival of new entrants and the growing confidence in the field's overall development. Combinatorial chemistry is well received in the public and private capital markets, which do not hesitate to invest. Many different reasons contribute to sustain interest in investors.

Combinatorial chemistry can produce many different products, all valuable and useful, because they are geared towards specific applications. The products of combinatorial chemistry address vast markets, and many of them can reach the marketplace relatively quickly. Combinatorial chemistry is a platform technology that is integrating very well with other technologies, such as functional genomics and proteomics, which are used to focus new lead generation. In addition, this field has an unusual patent history. No single company or individual has control of any one strategic patent, and there is no single strategic patent that defines the field, with many companies with patents that enable them to pursue unique chemistries, thus providing investors with many choices.

Combinatorial chemistry is a very dynamic and evolving technology that will always be changing and improving. Its value will increase with time because of this feature.

References

1] Chabala, J.C. *Pharmaceutical Manufacturers Association, Drug Discovery Management Subsection,* September 19-21, 1993, Philadelphia, PA.

2] Furka, A., Sebestyen, F., Asgedom, M., and Dibo, G. 1988, Abstr. *14th Int. Congr. Biochem.*, Prague, Czechoslovakia, Vol. 5, p. 47.

3] Houghten, R.A. 1985. General method for the rapid solid-phase synthesis of large number of peptides: Specificty of antigen-antibody interaction at the level of individual aminoacids. *Proc. Natl. Acad. Sci. USA*, 82,5131-5135.

4] Fassina, G., Lebl, M., and Chaiken, I.M. 1988. Screening the recognition properties of peptide hormone sequence mutants by analytical high performance liquid affinity chromatography on immobilized neurophysin. *Coll. Czech. Chem. Comm.* 53:2627-2636 .

5] Lam, K.S., Salmon, S.E., Hersh, E.M. Hruby, V.J., Kazmierski, W.M., and Knapp, R.J. 1991. A new type of synthetic peptide library for identifying Ligand binding activity. *Nature* 354, 82-84.

6] Hougthen, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T., and Cuervo, J.H. 1991. Generation and use of synthetic combinatorial libraries for basic research and drug discovery. *Nature* 354, 84-86

7] Cwirla, S., Peters, E.A., Barret, R.W. and Dower, W.J. 1990. Peptides on phage: a vast library of peptides for identifying ligands. *Proc. Natl. Acad. Sci. USA.* 87:6378-6382.

8] Scott, J.K. and Smith G.P. 1990. Searching for peptide ligands with an epitope library. *Science* 249:386-390.

9] Tuerk, C. and Gold, L. 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase *Science* 249: 505-510.

10] Ellington, A.D. and Szostak, J.W. 1990. In vitro selection of RNA molecules that bind specific ligands. *Nature* 346: 818-822.

11] Fassina, G., Verdoliva, A., Odierna , MR., Ruvo, M., and Cassani, G. 1996. A Protein A mimetic peptide for the affinity purification of antibodies. *J. Mol. Recogn.* 9:564-560.

Solution-Phase Libraries

Dr. Andrea Missio
Combinatorial Technologies
GlaxoWellcome, ITALY

In the early days of combinatorial chemistry, solution-phase synthesis evolved as the preferred method by traditional medicinal chemists involved in library synthesis. This was mainly due to the fact that organic chemists found it easier to use the methodologies they had practised for years rather than learning how to use solid supports.

The complexity of these first libraries was limited but set the scene for the introduction of the "combinatorial" concept in their research organizations.

One of these first examples is illustrated in the paper by Smith *et al.* [1]. This paper doesn't come from a combichem group, rather they are a group of chemists and biologists who intended to challenge the efficacy of combinatorial methods in lead identification. The library was prepared starting from a set of 40 acid chlorides and a set of 40 nucleophiles (amines and alcohols) to produce all the 1600 possible combinations between the two sets. The chemistry was kept simple; 0.5M solutions of each reactant were prepared in DCM. Half of each solution was mixed with an equal volume of all the other 39 monomers of the same type to give a single pool of the reactants. Each single monomer was treated with the pool solution containing the coupling partners and allowed to stand for 48h to produce the desired products. Any unreacted acid chloride was then destroyed by adding a small amount of methanol and the solutions were allowed to evaporate in the air for an additional 24-48h. No further attempts were made to purify the crude products.

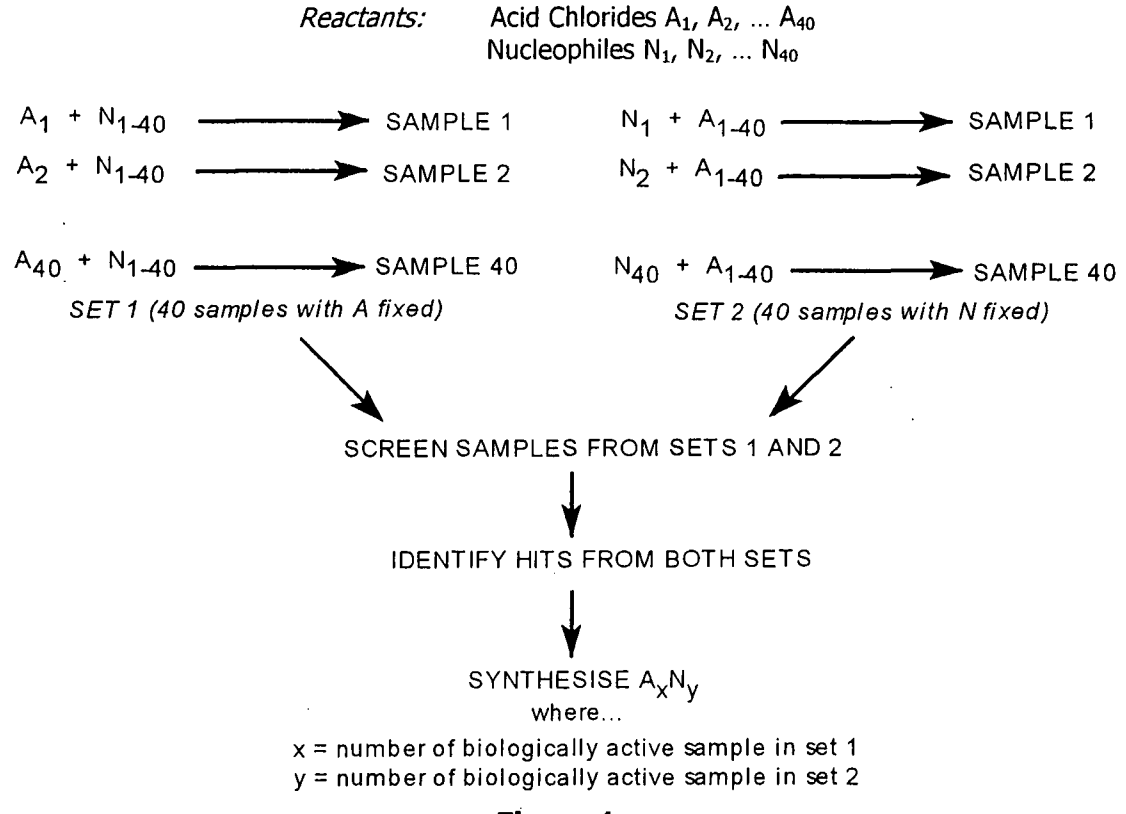
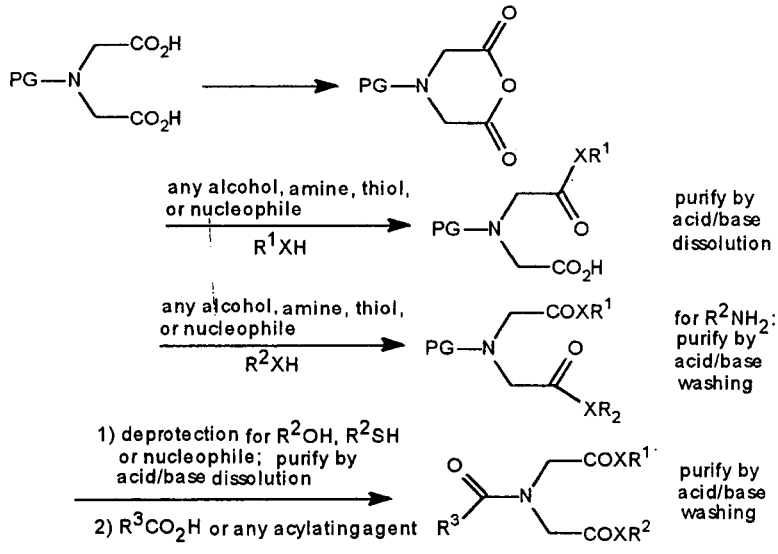


Figure 1

With only 2 x 40 = 80 reactions the authors were able to present to the screening an indexed library consisting of two sublibraries: one contained the 1600 final products indexed according to the acid chloride component (sublibrary A), the other one containing the same 1600 final compounds indexed according to the nucleophile component (sublibrary B). If an active component was found in sublibrary A, the acid chloride component was known. Activity was due to the combination of that particular acid chloride and one of the forty nucleophiles. Similarly if an active component was found in sublibrary B, the nucleophile component was known. Activity was due to the combination of that particular nucleophile and one of the forty acid chlorides. By combining positives from both sets of sample mixtures the whole structure of active components could be quickly determined (**Fig. 1**).

Purification is usually an essential part of any preparation in organic chemistry and it was soon implemented in the synthesis of libraries using solution-phase methods.

In the following example by Boger *et al.* [2] a novel solution-phase strategy for the preparation of small molecule libraries is disclosed. According to the authors a universal template useful in the synthesis of general purpose libraries is a viable concept. Such a template should be easily accessible and should allow the synthesis of library members in multi-milligram quantities. Simple liquid/liquid extractions are used to remove reactants, unreacted starting materials, reagents and their by-products, thus increasing the purity of library components irrespective of the reaction yields and without lengthy chemistry assessment. One such general template is the N-protected iminodiacetic acid anhydride which, upon reaction with a first set of nucleophiles (alcohol, amine, thiol or carbon nucleophiles) gives a monoamide/monocarboxilic acid. Acid/base washings allowed the recovery of a pure intermediate which was coupled to a second set of nucleophiles to give esters or amides. Again aqueous extractions yielded a pure intermediate. Deprotection of the secondary amine on the core scaffold provided the third point of diversity saturated with a set of acylating agents. Acid/base washings provided the final compounds in >90% purity and in 5 to 60 mg without prior optimization (**Scheme 1**).

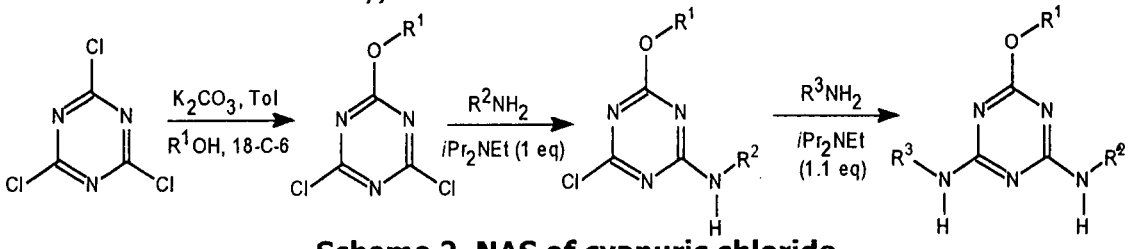


Scheme 1

One of the largest libraries addressed by manual manipulation using these chemical transformations is a 960-member library assembled in a 6 x 8 x 20 matrix. Each library member was obtained in 10 to 148 mg with purities ranging from 10 to 71%.

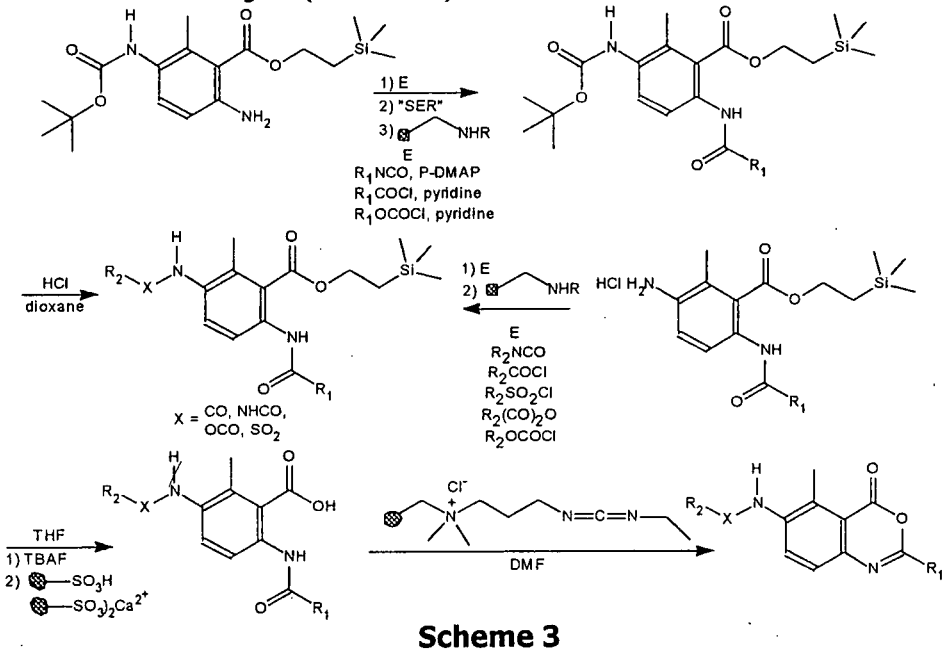
As the complexity of the chemistry increases more sophisticated methods of purification need to be used in solution-phase library synthesis.

Johnson *et al.* [3] described the synthesis of heterocyclic derivatives using Nucleophilic Aromatic Substitution (NAS) as the general reaction to assemble the library and reported the use of **Solid-supported Liquid Extraction** (SLE) for the purification of the library members. SLE was found to be an excellent method for purification, amenable to automation and allowing high-throughput removal of reaction contaminants. Inexpensive, coarse, hydrophilic diatomaceous earth was chosen as the support. It was packed in microtiter plates or in syringes but now is also commercially available from Varian. Columns were pretreated with a small quantity of aqueous phase which formed a thin film coating the hydrophilic support, then the reaction mixture was applied and eluted with a water immiscible solvent. The organic phase was recovered by gravity, free of water soluble contaminants. With this purification method in place this group set out for the synthesis of a variety of heterocycles, using multiple step synthetic sequences. The selected templates included cyanuric chloride (see **Scheme 2**); 4,6-dichloropyrimidine; 2,4-dichloro-6-methylpyrimidine; 2,3-dichloroquinoxaline and 6,7-dimethoxy,2,4-dichloroquinazoline. The nucleophiles included alcohols; phenols; phenoxides and amines. Careful chemistry assessment was performed to control the regiochemistry of the NAS reactions. These libraries included up to 1920 different compounds at > 85% relative ion abundance (as observed by MS analysis in at least 74% of the whole library).



Scheme 2. NAS of cyanuric chloride

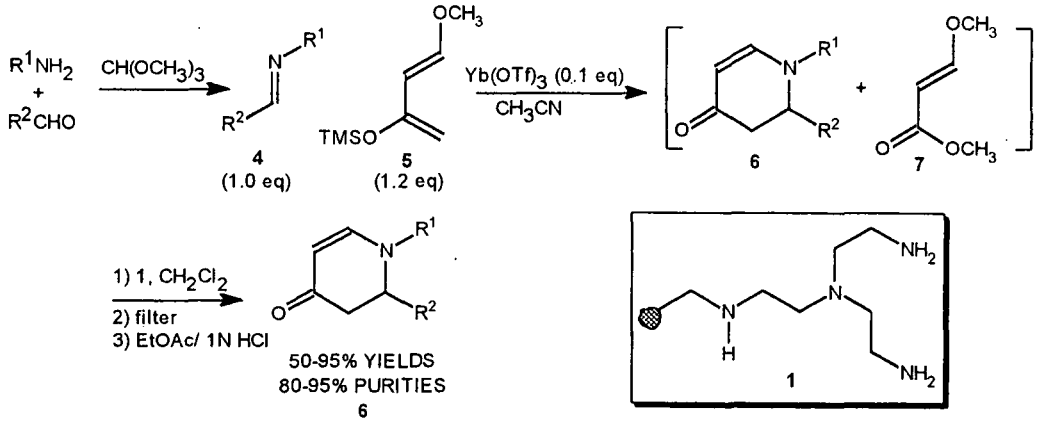
The careful choice and use of a variety of supported reagents allowed Parlow and Flynn [4] to complete the solution synthesis of a benzoxazinone library. Their purification method was called **Complementary Molecular Reactivity and Molecular Recognition** (CMR/R). The general synthetic scheme elaborated a mono-BOC protected diaminobenzoic trimethylsilylethyl (SEM) ester into the final benzoxazinone by reaction with a set of electrophiles (isocyanates, acyl chlorides, and chloroformates), BOC deprotection with HCl/dioxane, reaction with a more diverse set of electrophiles (including also anhydrides and sulphonyl chlorides), SEM deprotection using TBAF in THF and ring closure with a carbodiimide reagent (**Scheme 3**).



Scheme 3

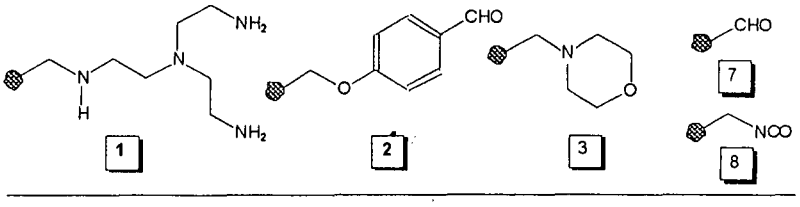
The HCl liberated in the first reaction was trapped by pyridine or polymer-supported DMAP. Excess electrophiles were scavenged by adding a high loading polyamine resin. When isocyanates were used a 10% starting benzoate was always present; this was transformed into a more reactive compound by treatment with tetrafluorophthalic anhydride to form a carboxylic acid derivative which was also scavenged by the polyamine resin. Typical purities at this stage were 97-99%. The high loading polyamine resin was successfully employed also with the second set of electrophiles. Typical purities were in the range 78-99% with the exception of products coming from the reactions with isopropyl and α-phenyl-methylisocyanate in the first step. These products gave lower purities due to undesired benzoxazinone formation. SEM protection was removed with TBAF. The free carboxylate (as its tetrabutylammonium salt) and excess TBAF were scavenged with a mixture of anion and cation exchange resins to give the desired free carboxylic acid in solution along with the polymer-supported Bu_4N^+ salt and insoluble calcium fluoride. Finally ring closure took place using polymer-supported EDC to yield pure benzoxazinones. The purity of the final library members were in the range 82-99% with the exceptions described before and yields were good to excellent (on average, 59% overall yield for the five step synthesis).

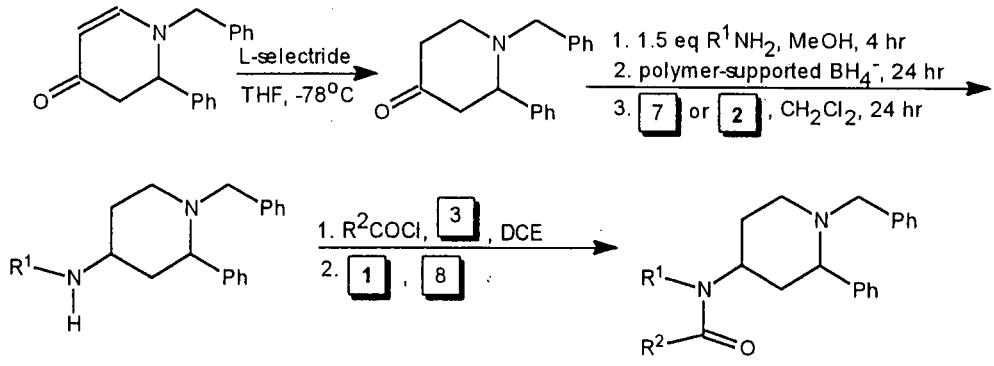
Another nice application of polymer-supported quench reagents in aid of solution-phase library synthesis came from Creswell *et al.* [5], who prepared a number of derivatives around a combinatorially assembled dihydropyridone core scaffold. This heterocycle was prepared by Diels-Alder reaction of a series of imines (from ten amines and four aldehydes) with Danishefsky's diene. 4-methoxy-3-buten-2-one (from the hydrolysis of Danishefsky's diene) was removed by reaction with a polyamine resin, while the Lewis acid catalyst was extracted with an aqueous work-up. Purities and yields were in the 80-95% and 50-95% range, respectively (**Scheme 4**).



Scheme 4

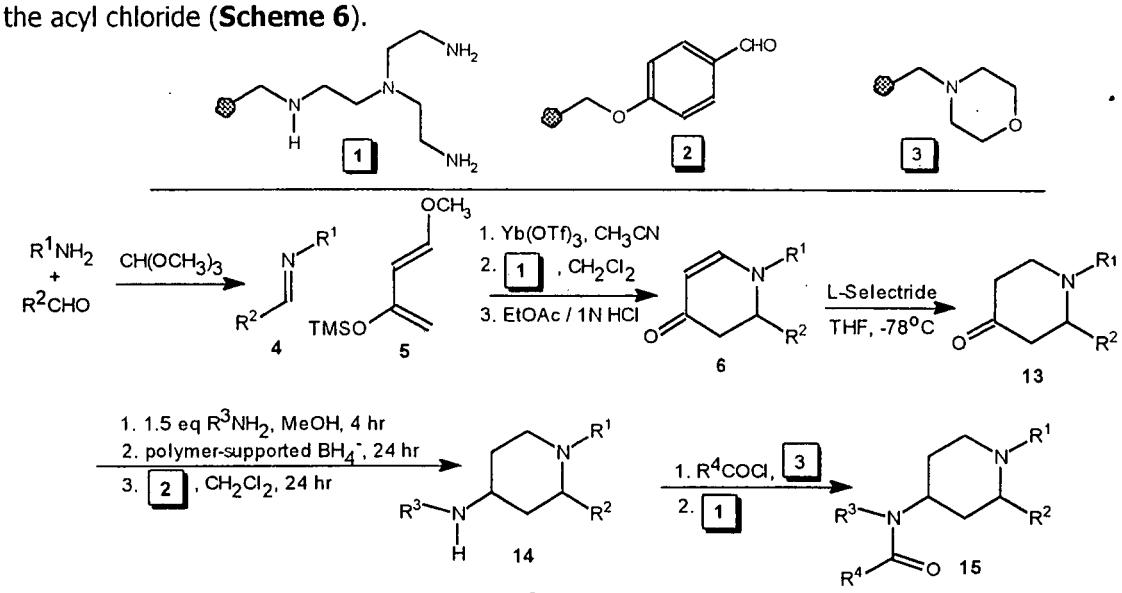
These products were further elaborated through chemoselective reduction of the double-bond with L-Selectride, which exposed the 4-oxo group to a reductive amination. A set of ten amines was used and polymer-supported borohydride was the reducing agent. Excess amines were quenched with two commercially available carboxaldehyde resins. The secondary amines from the reductive amination were reacted with a set of eight acyl chlorides, using solid-supported morpholine as a proton sponge. Excess electrophiles were removed by treatment with polyamine resin, while unreacted amines were trapped with an isocyanate resin. Yields were quite respectable, averaging 48%; purities were around 84% (**Scheme 5**).





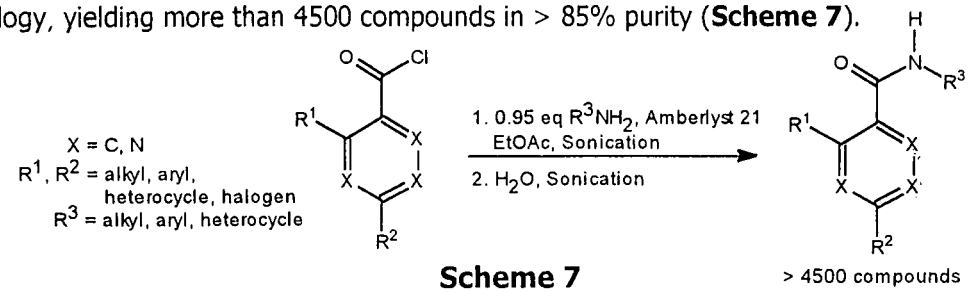
Scheme 5

To take the libraries from libraries to an extreme, a small 4 x 2 x 1 x 1 array was prepared starting from 4 amines, two aldehydes, one amine for the reduction process and one acyl chloride. Careful selection of supported reagents accounted for the final purities which were remarkable for a five-step sequence (average 80%). Polyamine resin trapped unreacted electrophiles and the hydrolysis product of Danishefsky's diene, carboxaldehyde resin reacted with excess primary amines in the presence of the secondary amine products, and morpholyne supported-reagent scavenged the HCl from the reaction with the acyl chloride (**Scheme 6**).



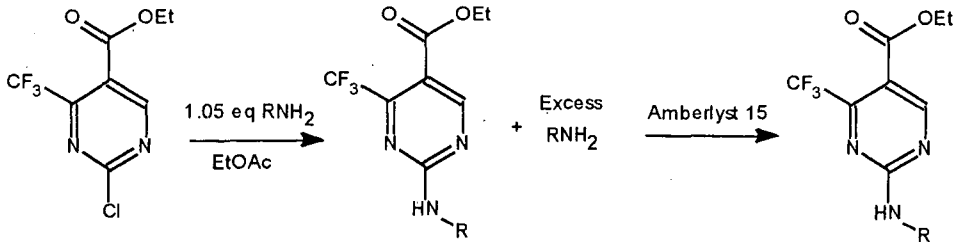
Scheme 6

A class of supported reagents chemists are familiar with are the ion-exchange resins. Suto *et al.* [6] explored their use in the preparation of a series of amides from acyl chlorides and amines. The basic ion-exchange resins soaked the HCl formed in this reaction and, upon completion of the reaction, a small amount of water was added. Any unreacted acyl chloride was converted to the corresponding free carboxylic acid which was now adsorbed by the resin. Among the many ion-exchange resins tested Amberlyst 21 performed better and was used throughout this study, while EtOAc was the solvent of choice. A series of pyrimidine, benzene and related derivatives were synthesized according to such a methodology, yielding more than 4500 compounds in > 85% purity (**Scheme 7**).



Scheme 7

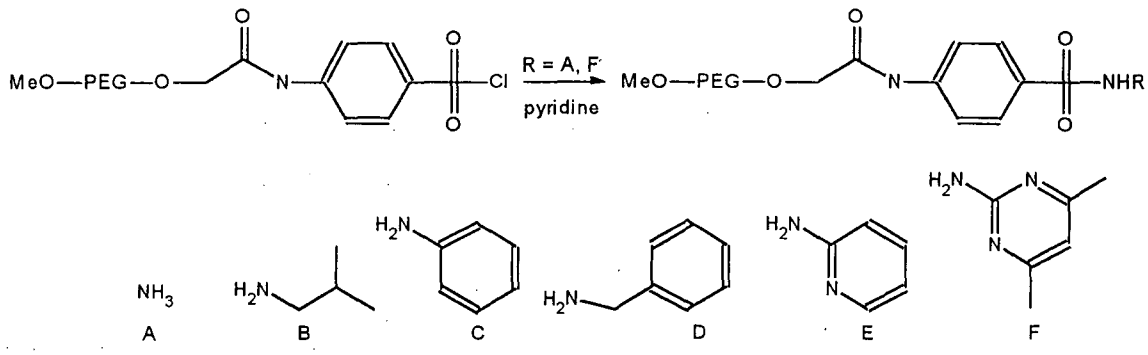
In analogy, esters were obtained by simply substituting various alcohols in place of the amines. Acidic ion-exchange resins were also evaluated and a library of ureas prepared. In this series Amberlyst 15 was used to scavenge excess amines that were reacted with substoichiometric quantities of the precious isocyanates. Purities were usually > 95%. In addition, acidic ion-exchange resins proved successful in NAS reactions on 2-chloropyrimidine derivatives. Excess amines were used to drive reactions to completion and adding Amberlyst 15 provided clean products. No scavenging action was observed on the final products, the less basic 2-aminopyrimidines (**Scheme 8**).



Scheme 8

To complete this study the authors used a mixed solid-phase/solution-phase approach. 2-aminopyrimidines from a solid-phase synthesis were polluted by excess amines and an ion-exchange resin was successfully used to purify the library members.

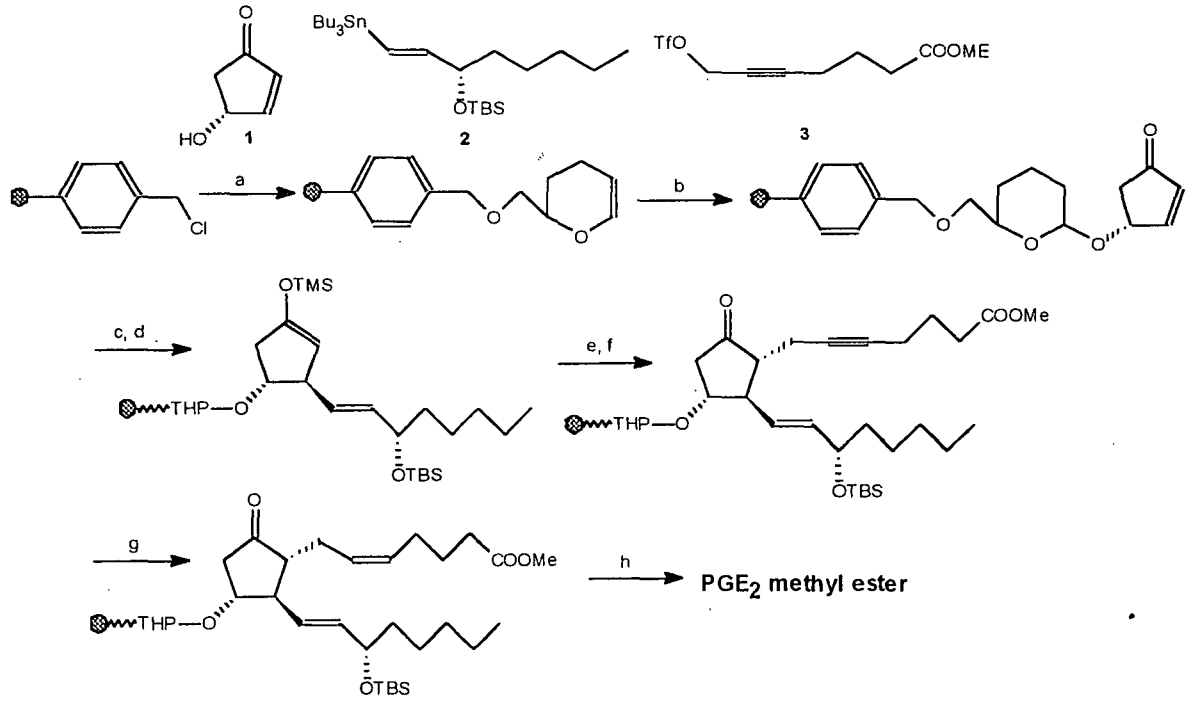
In addition to solid-phase and solution-phase methods a third method has appeared recently. Janda and Gravert [7] described the **Liquid Phase Combinatorial Synthesis** (LPCS). Their work exploited a physical property of polyethylene glycol (PEG), its solubility in most organic solvents but not in diethyl ether and cold ethanol, which allows its precipitation by addition of the latter solvents to the reaction mixture.
They reported a peptide and a sulphonamide library synthesis where, upon completion of each step, Et_2O was added to precipitate the PEG polymer with the attached monomers while the soluble excess reagents were filtered off. At the end final products were recovered by hydrolytic cleavage in analytically pure form in overall yields of 95-97% (**Scheme 9**).



Scheme 9

Continuing their search for novel soluble polymers for LPCS, Janda and Chen [8] reported the use of Non-cross-linked Chloromethylated PolyStyrene (NCPS) for the synthesis of prostaglandin E_2 methyl ester. PEG was not applicable due to its low solubility in THF at −78°C; this would have posed a problem during the aqueous work-up necessary to eliminate the inorganic and organometallic by-products. NCPS was soluble in THF, DCM and EtOAc, and insoluble in water and methanol thus enabling aqueous work-ups and precipitation with methanol. This soluble polymer was obtained by copolymerization of styrene and 4-(chloromethyl)styrene with a loading of 0.3 mmol/g. The synthesis of prostaglandin E_2 methyl ester was quite demanding since it involved the use of methyl lithium and lithium cuprate species at low temperatures, hydrogenation and cleavage with HF. The use of this

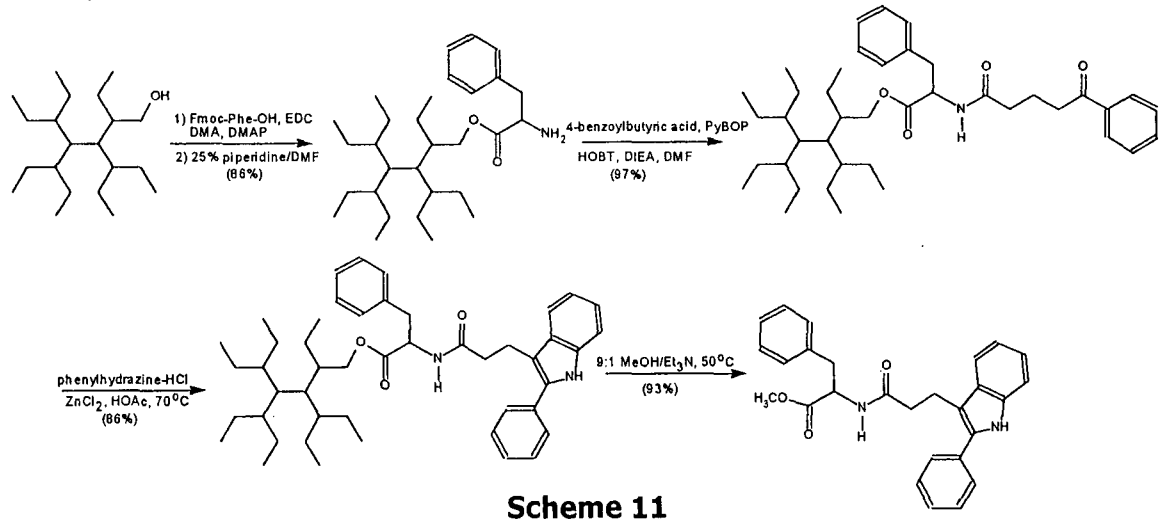
soluble polymer allowed the recovery of the pure final product, after eight steps and flash-chromatography, in 37% overall yield (**Scheme 10**).



*Reagents and conditions: (a) 3 eq of 6-(hydroxymethyl)-3,4-dihydro-2H-pyran, 3.3 eq of NaH, dimethylacetamide, rt, 24h: (b) 3 eq of **1**, 0.5 eq of PPTS, CH_2Cl_2, 40°C, 16 h; (c) 4.2 eq of **2**, 3.9 eq of $Li_2CuCNMe_2$, THF, -78°C, 15 min; (d) 15 eq of chlorotrimethylsilane, -78°C, 30 min; 30 eq of triethylamine, 0°C, 15 min; (e) 3 eq of MeLi, THF, -23°C, 20 min; (f) 6 eq of **3**, -78°C, 10 min; -23°C, 30 min; (g) H_2, 5% Pd-$BaSO_4$, qunoline, benzene/cyclohexane (1:1), 48h; (h) 48% aqueous HF/THF (3:20 v/v).*

Scheme 10

A number of soluble polymers are now available for LPCS applications and there is scope for improving either the properties and the number of these soluble supports but the wider combinatorial chemistry community is still making resistance to consider LCPS among their tools for combinatorial applications.



Scheme 11

The final application is related to dendrimer chemistry. Kim *et al.* [9] reported **Dendrimer supported Combinatorial Chemistry** (DCC) which featured solution-phase synthesis on dendrimer supports. Dendrimers are branching oligomers with a precise molecular architecture, with PAMAM being the most

common (and commercially available) dendrimer. Notably, low generation dendrimers (low molecular weight dendrimers) exist in extended form and allow great reagent accessibility. DCC is conceptually analogous to solid-phase combinatorial synthesis except that reactions are performed in solution and dendrimeric intermediates are separated by size exclusion chromatography (SEC). DCC offers a number of advantages: it is a solution-phase synthesis with no need to adapt the chemistry to solid supports; analytical methods like NMR, IR and MS can be applied to monitor reactions; extremely high loadings can be obtained on dendrimers; SEC purification does not rely on any physical property other than the dimensions of the support-bound compounds. To validate the concept of DCC this group investigated the well-known Fischer indole synthesis (**Scheme 11**).

A small array of 3 aminoacids x 3 ω-oxo-carboxylic acids x 3 phenylhydrazines was successfully assembled. The indoles were obtained in yields and purities that averaged around 90%. Like the previous soluble supports described by Janda and co-workers, also dendrimers hold the promise to become a valuable tool in combichem applications with their blend of solution- ands solid-phase methods. Further chemistry need to be done in order to prove that DCC is a valuable, general and practical way to generate libraries.

References

1. Smith et al., *Bioorg. & Med. Chem. Lett.* **1994,** *4,* 2821-2824.
2. Boger et al., *J. Am. Chem. Soc.* **1996,** *118,* 2567-2573.
3. Johnson et al., *Tetrahedron,* **1998,** *54,* 4097-4106.
4. Parlow and Flynn, *Tetrahedron,* **1998,** *54,* 4013-4031.
5. Creswell et al., *Tetrahedron,* **1998,** *54,* 3983-3998.
6. Suto et al., *Tetrahedron,* **1998,** *54,* 4141-4150.
7. Janda and Gravert, C*hem. Rev.* **1997,** *97,* 489-509.
8. Janda and Chen, *J. Am. Chem. Soc.* **1997,** *119* 8724-8725.
9. Kim et al., *Proc. Natl. Acad. Sci. USA* **1996,** *93,* 10012-10017.

Solid-Phase Libraries

Dr. Andrea Missio
Combinatorial Technologies
GlaxoWellcome, ITALY

The tool-kit of reactions adapted to solid-phase chemistry has been growing year after year in the late 90's. In particular, heterocyclic chemistry has attracted much attention from combinatorial chemists since most of the marketed drugs contain an heterocycle moiety and these represent privileged structures for many, if not all, pharma and biotech companies.

Affymax has been very active in this field and Gordon *et al.* [1] have detailed their search for a versatile synthetic intermediate that would have provided access to diverse nitrogen heterocycles. Imines were found to be both versatile and easily accessible intermediates. They are formed via condensation of amines with aldehydes or ketones. Both partners could be the supported component and the final imines demonstrated to be stable species. Gel phase ^{13}C-NMR using ^{13}C-enriched building blocks proved to be a valuable tool to monitor imine formation and confirming the intrinsic stability of these chemical entities. Substantial investigation was devoted to the study of an efficient solvent to prepare these intermediates and TMOF (TriMethyl OrthoFormate) was found to be a convenient solvent **and** dehydrating agent. THF could be used when the solubility of the monomers in TMOF was poor. The panel of heterocycles accessible starting from imines comprised: 4-thiazolidinones, β-lactams, pyrrolidines, dihydropyridines and pyridines but only β-lactams and dihydropyridines will be here discussed (**Figure 1**).

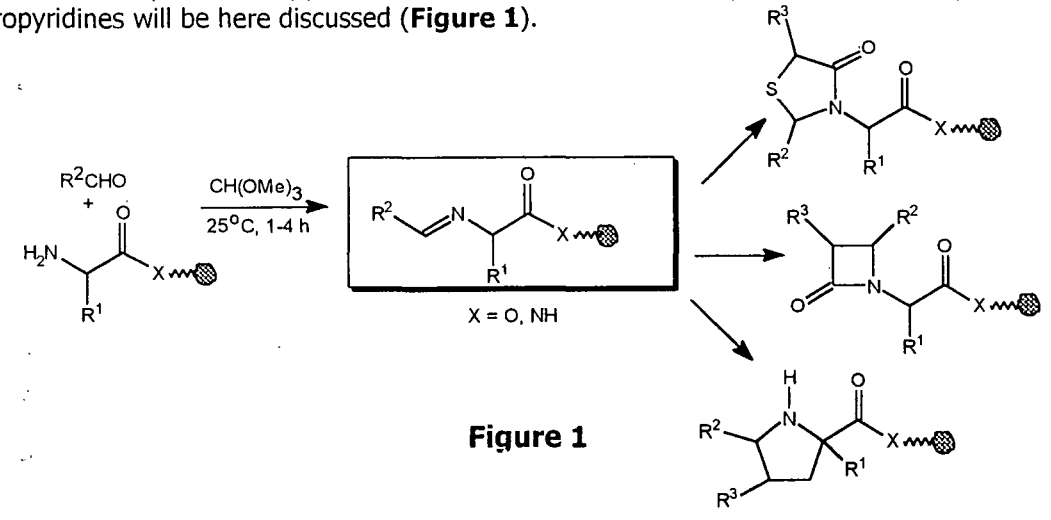


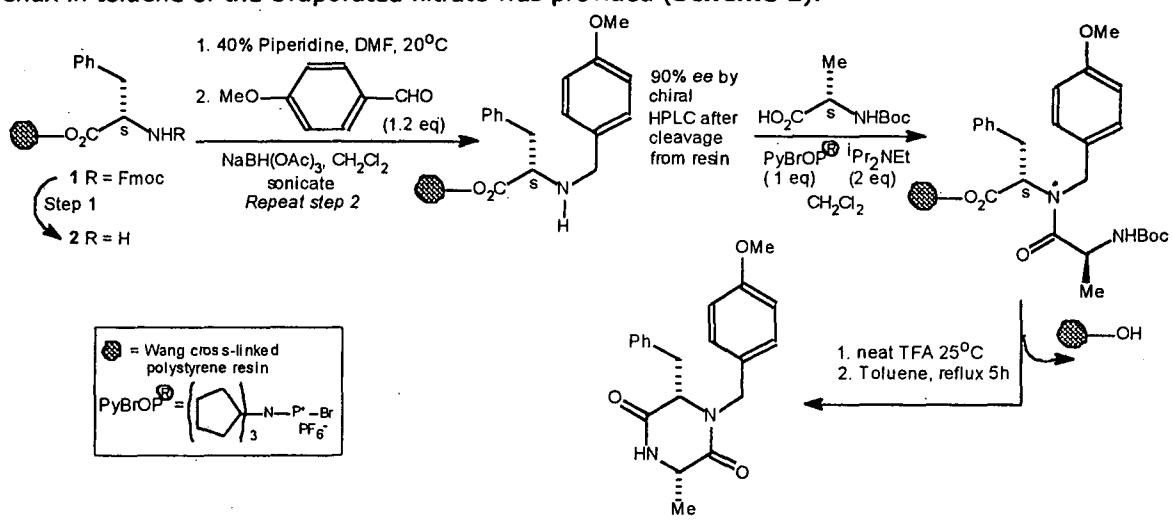
Figure 1

The Staudinger addition of ketenes to imines offered a versatile approach to the generation of β-lactams. Since only few ketenes were commercially available they were preferentially obtained *in situ* by base dehydrohalogenation of acyl halides. A [2 + 2] cycloaddition occurred and, interestingly, the quality of the crude products from solid-phase synthesis was frequently superior to that obtained by classical solution-phase synthesis. Using a ketene from commercially available phthalimidoacetyl chloride, researchers at Affymax were able to obtain, after nitrogen deprotection, the corresponding 3-amino-2-azetidinones. These, in turn, were reacted with a set of carboxylic acids to yield 3-amido-2-azetidinones in excellent yields and purities after cleavage from the resin. Imine formation was also successfully used in the synthesis of dihydropyridines, based on a two- or three-component condensation of enamino esters with 2-arylidene β-keto esters or β-keto esters and aldehydes, respectively.

Reaction of an amino resin with a β-keto ester to give the corresponding stable enamino ester (via imine formation and subsequent rearrangement to enamine), followed by condensation with 2-

benzylidene β-keto ester or a mixture of β-keto ester and aldehyde and final treatment with TFA, under inert atmosphere to prevent oxidation to pyridines, gave the desired dihydropyridines in 65-75% yield. Both β-lactams and dihydropyridines are well represented on the market as antiinfectives, antihypertensives, antiatherosclerotics and antidiabetics. Libraries of privileged structures like these constitute an excellent opportunity for pharma companies to fish out useful drugs for the future market.

Gordon and Steele [2] described the synthesis of a library of diketopiperazines (DKPs), selected as versatile, heterocyclic scaffolds on which to arrange pendant functionalities. Chemistry assessment involved the optimisation of reductive amination of solid-supported FMOC-aminoacids, using ultrasound-promoted reactions and a double-coupling protocol to achieve consistently high yields (>90%). The secondary amine was acylated with a set of BOC-aminoacids, using PyBrOP as the activating agent and, again, a double-coupling protocol (yields >90%). TFA deprotected the BOC group and removed the product from the resin but no cyclisation to DKPs occurred unless a short reflux in toluene of the evaporated filtrate was provided (**Scheme 1**).



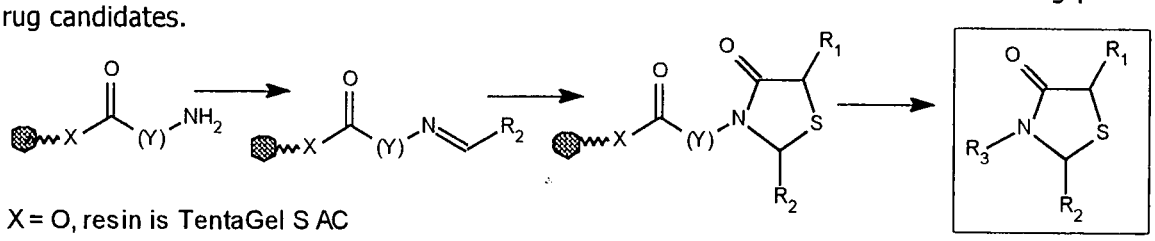
Scheme 1

Having optimised the reaction conditions the authors set out for the split/mix synthesis of 1000 DKPs from a pool of ten FMOC-aminoacids, ten aldehydes and ten BOC-aminoacids. The final library was presented as ten pools of 100 DKPs each. To assure that the quality was good the intermediate mixture of ten secondary amines was analyzed. TFA cleavage followed by HPLC-MS combined with MS-MS was used to unambiguously confirm the presence of 96 out of 100 expected products.

In an interesting paper, Look *et al.* [3] took advantage of their experience in the solid-phase synthesis of 4-thiazolidinones to prepare a series of focused libraries to be screened against the enzyme cyclooxygenase-1 (COX-1), prompted by a patent claiming a group of nearly 100 4-thiazolidinones as COX-1 inhibitors. The three libraries were prepared on amino or hydroxy resins by coupling a FMOC-aminoacid, which would yield an amide or a carboxylic acid in the final products. (**Scheme 2**).

Using split/mix synthesis the thiazolidinones were prepared by imine formation on the free amino group of the five supported aminoacids, using five (hetero)aromatic aldehydes and condensation with five different mercaptoacids. Three 125 member libraries were assembled but the overall numbers, including the stereoisomers, was actually 3 x 540. Libraries were screened for activity against COX-1: only one of the libraries showed interesting activity and its active pools were resolved by iterative deconvolution. With three rounds of resynthesis a compound active at μM

concentrations was identified, thus proving the value of combinatorial libraries in finding promising drug candidates.

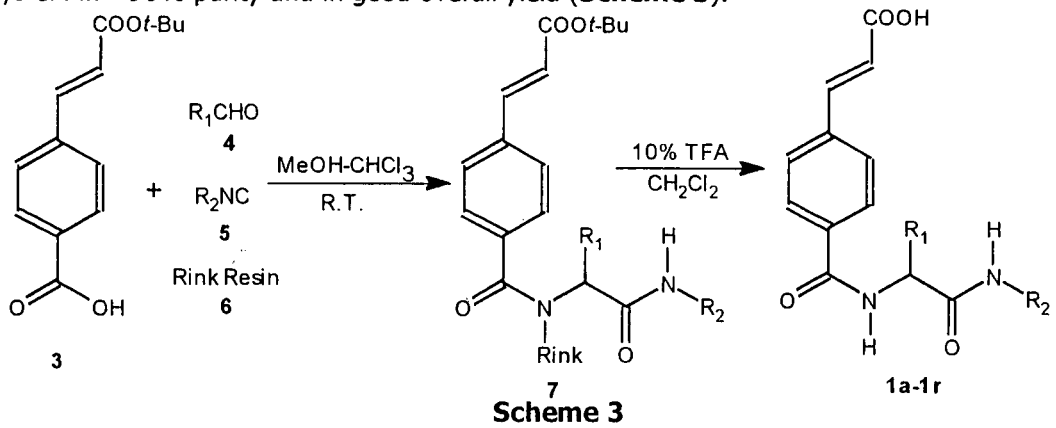


X = O, resin is TentaGel S AC

X = NH, resin is TentaGel S RAM

Scheme 2

Multiple Component Condensation (MCC) reactions are very powerful reactions in combinatorial chemistry since they add the diverse monomers all at once and produce a highly efficient combination of products in just one step. Cao et al. [4] presented the synthesis of a series of cinnamic acid derivatives in a Ugi four-component condensation on Rink resin as the amine component, t-butyl 4-carboxy-cinnamate, a set of six aldehydes and a set of seven isonitriles. t-Butyl ester group deprotection and product cleavage were accomplished simultaneously using 10% TFA/DCM in >90% purity and in good overall yield (**Scheme 3**).

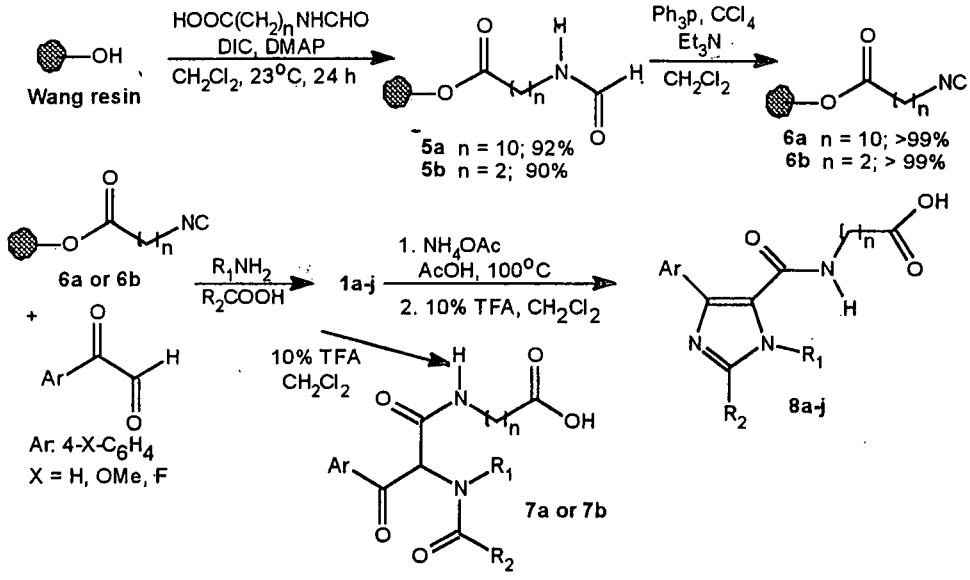


Scheme 3

Several advantages over the solution-phase were realized when Rink resin was employed as a reaction component. Not only were resonably complex products readily synthesized using the Ugi MCC reaction but purification was greatly simplified relative to the solution-phase synthesis. The resin served as a solid-support ammonia equivalent and this provided unsubstituted NH-acyl-α-aminoamides as the final products. These compounds were then tested as novel non-phosphorous based inhibitors of hematopoietic protein tyrosine phosphatase (HEPTP). Compounds active at micromolar level were detected and basic structure-activity relationships were established.

Following these encouraging results the same group described the synthesis of imidazoles from their previously disclosed NH-acyl-α-aminoamides, obtained through Ugi 4-component condensation (U-4CC) (Zhang et al. [5]). Wang resin was used in this study and N-formyl aminoacids were attached to it, dehydrated to the corresponding isonitriles and submitted to the U-4CC. The other reaction partners were 3 arylglyoxals, 4 amines (two aliphatic, two aromatic) and 4 carboxylic acids (two aliphatic, two aromatic). The intermediate NH-acyl-α-aminoamides were converted to imidazoles upon treatment with NH_4OAc in AcOH at 100°C for 20h. Yields after preparative TLC were in agreement with those reported for the solution-phase synthesis (ca. 45%) (**Scheme 4**).

The diverse set of coupling partners used in this example allowed the authors to draw some basic conclusions on the reactivity of the monomers from which it appeared that out of the examined monomers only anilines adversely affected this reaction.



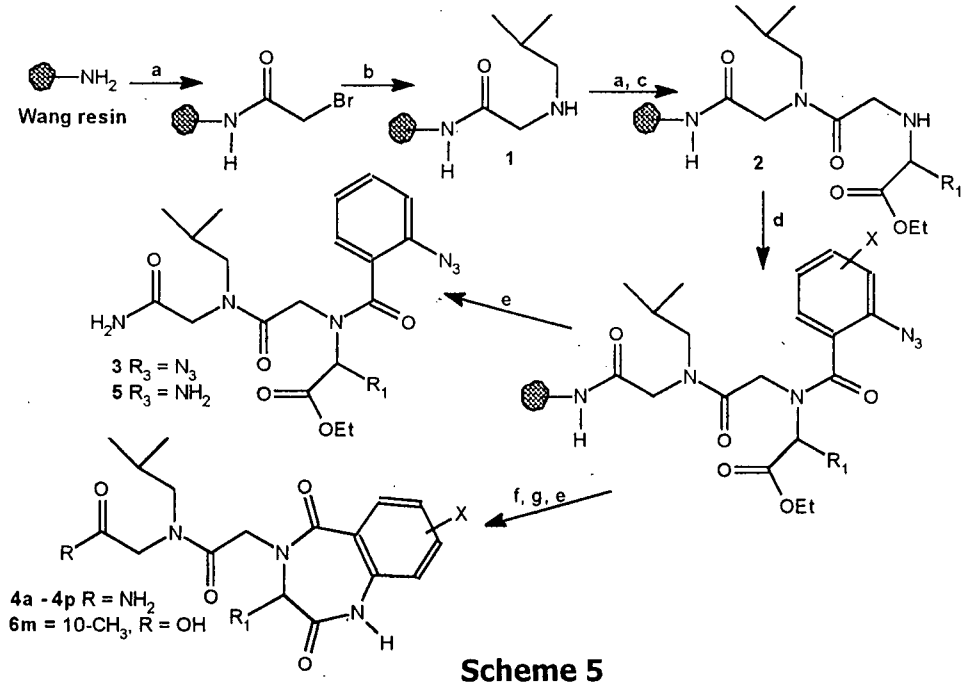
Scheme 4

One particular reaction which is quite demanding using classical methods is the macrocyclization. Hiroshige *et al.* [6] proved that the "pseudodilution" effect observed in peptide chemistry macrocyclizations could be translated also to less traditional palladium chemistry on solid-phase. Tentagel PHB resin was converted to an acid labile aminopropyl carbamate which was reacted with orthogonally protected FMOC-Lys(DDE). This aminoacid constituted the branching point from which a terminal alkene and a 3-iodophenyl residue were attached. These two groups were coupled under Heck conditions to give a small array of 20 macrocycles in good yields (>70%). Site isolation provided by the low loading polymer (0.23mmol/g) accounted for the "pseudodilution" effect and minimized the intermolecular cyclization. The authors did not comment on the yields expected for a similar reaction conducted in solution but probably, even under high dilution, that would not compare favorably with those obtained using solid-phase chemistry.

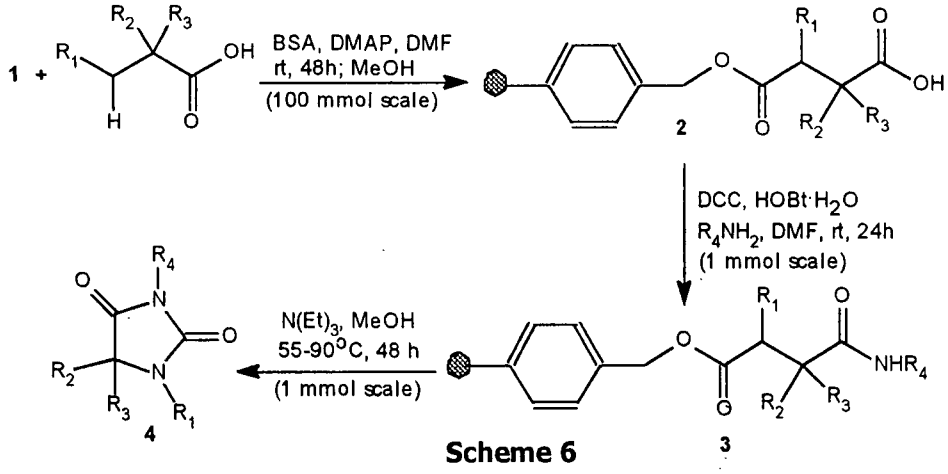
A hybrid peptoid-1,4-benzodiazepine-2,5-dione was disclosed by Goff and Zuckermann [7]. The authors had already prepared a number of N-substituted glycine peptoid libraries and now they implemented this concept to include a hybrid structure containing a small-molecule motif. The key reaction was an intramolecular aza-Wittig reaction on an *o*-azidobenzoic acid derivative. Tributylphosphine was selected as the phosphorous component and yielded iminophosphorane which was heated at 130°C to give the benzodiazepinones. The final structures were cleaved from the resin with TFA to provide the hybrid peptoid molecules in modest to good yields (**Scheme 5**).

It was noteworthy that purities of crudes averaged 80%. This chemistry has been adapted to library synthesis and a small 8-member library was assembled. HPLC showed eight major peaks and ESI-MS showed all the eight parent ions along with peaks ascribable to uncyclized material.

Hydantoins were among the first small molecules synthesized on resin and, since they are well represented drugs in the market, constituted an attractive scaffold for pharma companies. Dressman *et al.* [8] devised a simple synthesis which relied on largely represented, commercially available monomers: aminoacids and amines. The amino group of the aminoacid was hooked on the resin via a carbamate linkage, then the carboxy terminus was converted to an amide upon reaction with an amine. The amidocarbamate was subsequently cleaved to hydantoin by treatment with excess triethylamine in methanol for 48h at different temperatures (typically 55-90°C) (**Scheme 6**).



Scheme 5



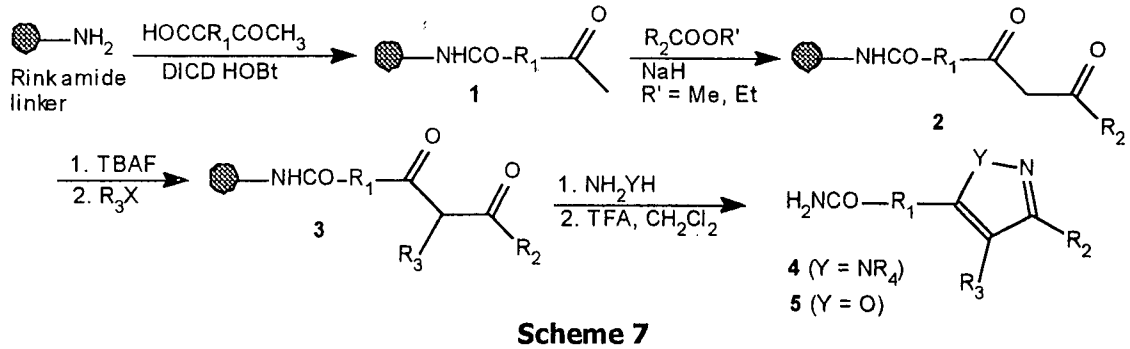
Scheme 6

Cyclative cleavage is an elegant way to overcome the drawback inherent in many solid-supports which, upon cleavage, release linker-related impurities together with the final molecule. Furthermore only the correct structures can cyclize. Failed syntheses are unable to deliver the cyclized product, thus improving the purity of the final released library members. Using such a protocol the authors were able to prepare 800 products from 20 aminoacids and over 80 amines in purities usually well above 90% and submitted them to screening for biological activity.

Marzinzik and Felder [9] were interested in the solid-phase synthesis of pyrazoles and isoxazoles. A four-step reaction sequence was devised, involving a Claisen condensation, an α-alkylation and a cyclization of a β-diketone to pyrazoles (using monosubstituted hydrazines) or to isoxazoles (using hydroxylamine) (**Scheme 7**).

Such a sequence introduced four elements of diversity into the scaffold, allowing a high degree of chemical diversity to be explored. A complete chemistry assessment was performed and valuable details on reactive and unreactive monomers were obtained. For example Claisen condensation failed to work with carboxylic esters bearing α-hydrogens and with weakly acidic heteroaromatic compounds. Ring closure yielded regioisomers with equal efficiency, unless steric or electronic

properties could account for a different ratio. This case was explored with hydralazine, which after four days showed only 20% conversion (no data were added relative to the regioisomer distribution). The collected data provided an information basis sufficient for planning combinatorial libraries and helped making informed decision on which building blocks to include and whether to skip a reaction step in favour of simplified reaction conditions.



Scheme 7

Terrett *et al.* [10] after demonstrating the ability of split/mix synthesis to deliver equimolar mixtures of compounds to be tested for biological activity became interested in a more productive way to synthesize compounds as discrete entities. The new methodology shared with mix/split the advantage that the same chemical transformation could be applied to many substrates in the same reaction mixture, but the solid-phase came in a form that allowed ready separation of the products. Furthermore the solid phase was in a format that permitted to track the synthetic history and thus the attached chemical structure to be decoded by examination of a printed code sequence. Libraries were made on a new laminar form of the solid phase material, which consisted of two woven sheets of inert polypropylene between which were sandwiched quantities of resin beads. The sheets were fused together in such a way that the beads were fixed immobile even when the sheets were put through chemical transformations or divided. To test this concept a trial library of 27 tripeptides (3 x 3 x 3) was synthesized (**Figure 2**).

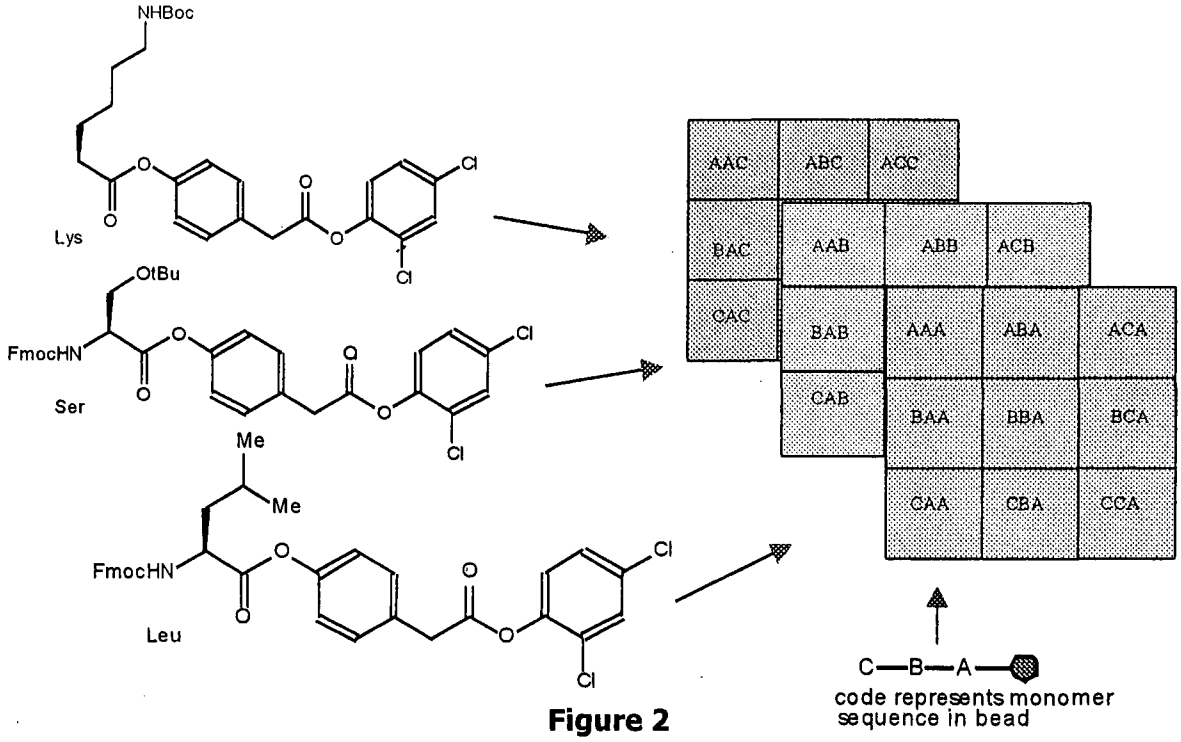
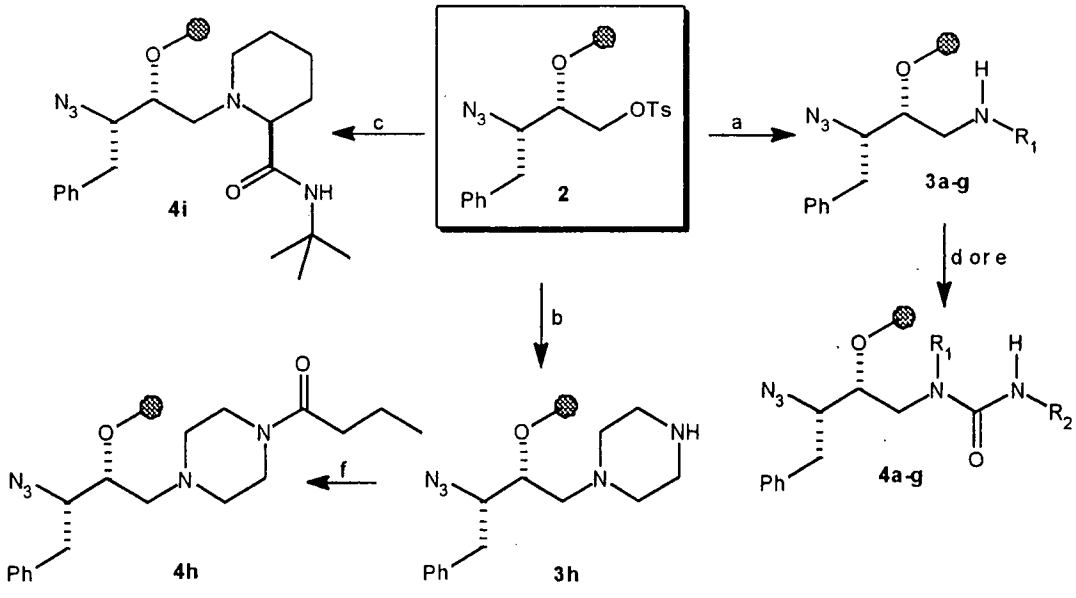


Figure 2

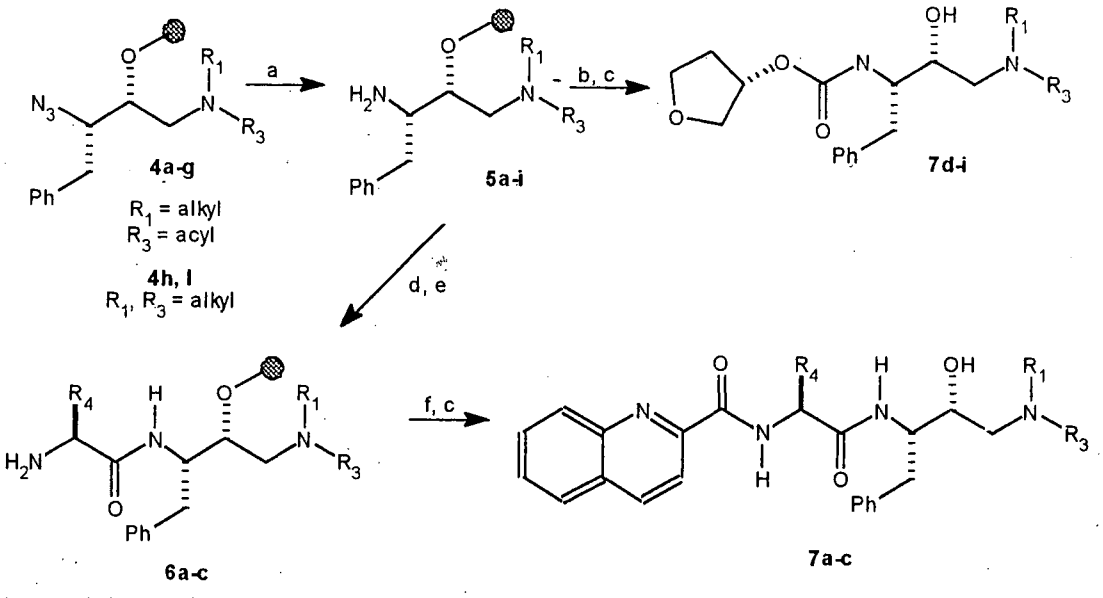
Three laminar sheets, each containing nine separate quantities of solid support, were prepared. Each portion of resin was printed with a three-letter code unique to the reagents that visited that particular area of the laminar sheet, then the synthesis could commence. Three FMOC-aminoacids were applied, one to each sheet, then the sheets were cut in three columns and the second set of three FMOC-aminoacids was applied, one to each column. The third set of three FMOC-aminoacids was applied to the sheets now cut into three rows, one aminoacid to each row. Finally TFA cleavage yielded the deprotected tripeptides which were analysed, and confirmed, by MS analysis. This novel laminar methodology allowed the rapid synthesis of large numbers of individual compounds in quantity, and in a form where their identity could be quickly determined by reading the unique code sequence printed on the sheet.

Aspartic acid proteases are a well-represented family of enzymes, characterised by two aspartic acid residues in the active site. They catalyze the peptide bond hydrolysis at the level of large hydrophobic residues. Potent inhibitors of these enzymes usually contain an isostere mimicking the tetrahedral intermediate which prevents the peptide hydrolysis. Kick and Ellman [11] presented the solid-phase synthesis of aspartic acid protease inhibitors based on (hydroxyethyl)amine or (hydroxyethyl)urea isosteres. A 3-benzyl-3-azido-propanediol, monoprotected at position 1 as tosyl derivative, was selected as the central scaffold on which to add further diverse elements. The hydroxyl function provided the point of attachment to the resin and the tosyl group was displaced by a series of primary and secondary amines. These, in turn, were derivatised to the corresponding ureas by means of a set of isocyanates. Azide reduction with $SnCl_2$ to give the free amine, followed by reaction with amino acids or succinimidyl carbonate derivatives yielded amides and carbamates, respectively. The amino group from aminoacid residues could be further elaborated to give amides.

TFA cleavage gave products in good yields (47-86%) which were subsequently sent to test for the evaluation of their activity as aspartic acid protease inhibitors (**Scheme 8**).



(a) R_1NH_2, NMP, 80°C; (b) piperazine, NMP, 80°C; (c) N-tert-butyl-L-pipecolinamide, NMP, 95°C; (d) R_2NCO, $ClCH_2CH_2Cl$; (e) (i)$OC(OCCl_3)_2$, Et_3N, cat DMAP, THF, (ii) 4-(3-aminopropyl)morpholine, THF; (f) butyryl chloride, i-Pr_2EtN, CH_2Cl_2.



(a) $SnCl_2$:HSPh:Et_3N (1:4:5), THF; (b) 3(S)-tetrahydrofuranylsuccinimidyl, i-Pr_2EtN, CH_2Cl_2; (c) 95:5 TFA/H_2O; (d) Fmoc amino acid, PyBOP, HOBt, i-Pr_2EtN (3 eq) DMF; (e) 20% piperidine in DMF; (f) pentafluorophenyl ester of quinaldic acid, HOBt, Et_3N, DMF.

Scheme 8

References

1. Gordon et al., *Accounts Chem. Res.* **1996,** *29,* 144-154.
2. Gordon and Steele, *Bioorg. Med. Chem. Lett.* **1995,** *5,* 47-50.
3. Look et al., *Bioorg. Med. Chem. Lett.* **1996,** *6,* 707-712.
4. Cao et al., *Bioorg. Med. Chem. Lett.* **1995,** *5,* 2953-2958.
5. Zhang et al., *Tetrahedron Lett.* **1996,** *37,* 751-754.
6. Hiroshige et al., *J. Am. Chem. Soc.* **1995,** *117,* 11590-11591.
7. Goff and Zuckermann, *J. Org. Chem.* **1995,** *60,* 5744-5745.
8. Dressman et al., *Tetrahedron Lett.* **1996,** *37,* 937-940.
9. Marzinzik and Felder, *Tetrahedron Lett.* **1996,** *37,* 1003-1006.
10. Terrett et al., *Chem. Eur. J.,* **1997,** *3,* 1917-1920.
11. Kick and Ellman, *J. Med. Chem.* **1995,** *38,* 1427-1430.

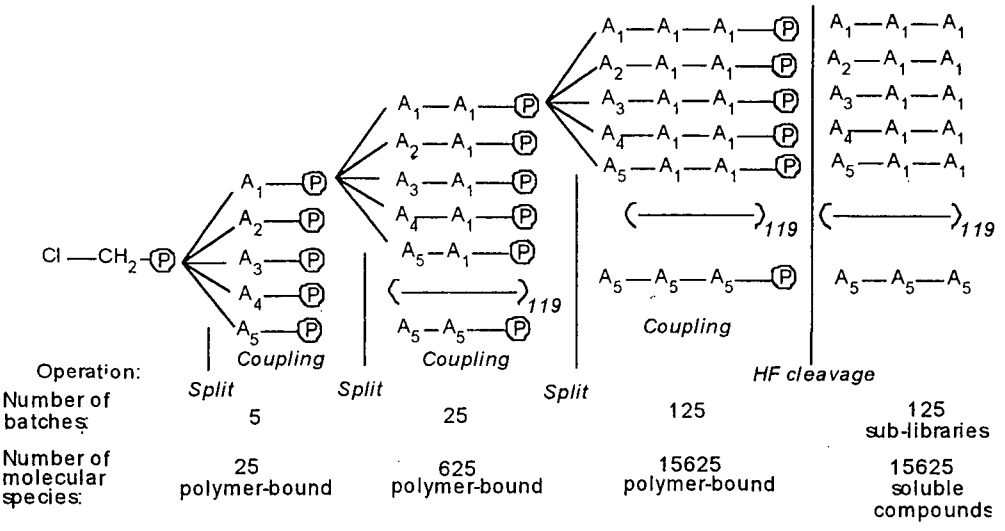
STRUCTURE DETERMINATION OF POSITIVES

Dr. Andrea Missio
Combinatorial Technologies
GlaxoWellcome, ITALY

A truly combinatorial synthesis involves at least a mix/split step which, in the end, will result in a "*x*" number of pools of compounds (where *x* is the number of monomers used in the last step). After biological test one is faced with the problem of identifying which structure(s) was(were) responsible for the observed activity. A number of solutions were proposed, the simplest being the resynthesis of the pool in individual reaction vessels and retesting the single compounds. If the pool is numerous one may desire to resolve it to smaller numbers employing split/mix synthesis in all but the last step. If "*n*" split/mix steps are applied "*n*" runs of deconvolution will be required. Other ways to solve the issue of structure determination of positives have been proposed and will be here discussed.

An elegant and conceptually simple way was proposed by Deprez *et al.* [1] with the use of "**orthogonal libraries**". The authors prepared two libraries of tripeptides using a set of 25 aminoacids (23 D-aminoacids and 2 achiral ones) at each position, resulting in 15625 trimers in both collections. Each library was partitioned in 125 sublibraries, each containing 125 tripeptides (**Figure 1**).

(a)



(b)	A1	A2	A3	A4	A5
B1	D-Leucine	D-Proline	D-Serine	(4-nitro)-D-Phenylalanine	D-Isoglutamic acid
B2	D-arginine	D-Isoleucine	Glycine	D-Threonine	D-Histidine
B3	D-Glutamine	D-Tyrosine	D-Valine	D-ε-Nicotinoyl lysine	D-Alanine
B4	D-Tryptophan	D-Asparaginine	D-Glutamic acid	D-Methionine sulfoxide	D-Phenylalanine
B5	Isonipecotic acid	D-Lysine	D-tetrahydro-isoquinoleic acid (D-Tic)	D-Aspartic acid	

Figure 1

The libraries were synthesized in the following way: library (A) contained 5 subsets (A_1 to A_5) of the 25 aminoacids. Each sublibrary was the result of the incorporation of these subsets at each of the three positions of the tripeptide. The second library (B) was prepared in the same way but the 5 subsets (B_1 to B_5) were "orthogonal" to the A subsets. The interesting feature of these two orthogonal libraries is

that any sublibrary A and any sublibrary B share only one trimer. A positive in a screen will show activity from both sublibraries. Since these are self-deciphering the structure of the active compound is easily deduced. The 25 aminoacids were selected so that there is the largest structural diversity possible to avoid the synthesis of sublibraries containing structurally correlated motifs, which could generate a cumulative effect during the screening. This structural diversity, along with the limited complexity of the pool (125 members), accounts for a low background and an easy detection of biological activity. Several advantages are realized screening libraries prepared in an orthogonal manner: compounds are screened in a soluble form, further chemical steps to add a code are not required, the method is not restricted to combinatorial libraries but can be used for the screening of any set of compounds, provided they are orthogonally arranged.

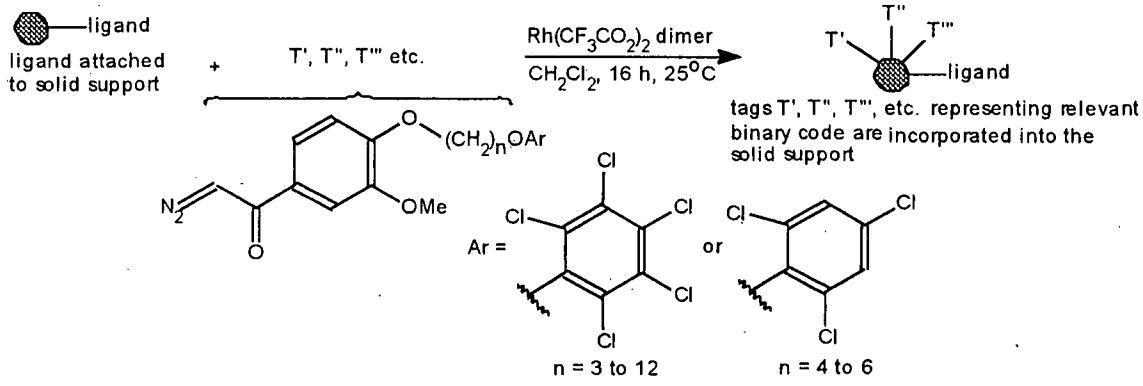
The issue of resolving complicated mixtures from combinatorial libraries has been addressed by Ohlmeyer *et al.* [2] who proposed the use of "**molecular tags**" to help in the identification of the structure of positives within a complex pool. This is a co-synthesis method, since the product and its tags are assembled on the same bead. Both oligonucleotides and oligopeptides have been proposed as chemical tags but they suffer some restrictions (especially the oligonucleotides) on the kind of chemistry that can be applied to prepare the ligand. The authors proposed the use of 20 different tags which consisted of a photolinker (to obtain orthogonal protection and allow a wider choice of chemistry to be used for the ligand synthesis) to attach the code to the bead, ten different length hydrocarbon tethers and three different haloaromatic electrophores. These tags were well resolved in Gas Chromatography (GC) and detected using Electron Capture (EC) detection at low levels (<1 pmol). 20 such tags would code for only 20 monomers thus reducing the applicability of the method. By implementing a binary code scheme the authors could keep track of 2^{20} = 1,048,576 monomers.

A binary code reflects the presence or absence of a product (0 = absent, 1 = present) and a code would be, for example, in the form 001, 010, 011, ..., 111. The haloaromatic electrophores were chosen for their chemical inertness, which could allow any harsh condition to be used in the assembly of the library without affecting the codes. A peptide library was prepared to test the feasibility of this encoding scheme, including also a sequence well recognized by a monoclonal antibody (mAb 9E10). 3-bit codes were assigned: 001 to Serine, 010 to Isoleucine, 011 to Lysine, 100 to Leucine, 101 to Glutamine, 110 to Glutamate and 111 to Aspartate. Tags were attached using the carboxy terminus of the linker, preparing an amount corresponding to 1% of the free amino groups on the bead and were cleaved by UV irradiation and unambiguously detected with EC/GC. The library was tested against mAb 9E10 and positive beads detected with colorimetric methods. Stained beads were picked up and codes were straightforwardly read. The decamer epitope was correctly identified along with a number of peptides differing for one or two substitutions at the N terminus thus proving the validity of these chemically inert molecular tags, used according to a binary code.

Baldwin *et al.* [3] continued to develop the concept of **electrophoric tags** in the synthesis of small molecule libraries. A library from 7 FMOC-aminoacids or aminoalcohols, 31 FMOC-aminoacids, and 31 different sulfonyl chlorides, isocyanates, carboxylic acids, and chloroformates was synthesized to yield 6727 different products, encoded with the electrophoric tags. Encoding of the library was accomplished by rhodium-catalyzed carbene insertion of diazomethyl ketone derivatives of the electrophoric tags (**Scheme 1**).

These were attached to an oxidatively labile linker and carbene insertion occurred directly on the polymeric bead, without affecting the synthesis sites on the bead. Three distinct sets of tags (three in step one, five in steps two and three) were used and recorded the complete history of the synthesis. Ligands were released from the library beads with UV irradiation and biological test was performed in solution. Arraying the library in microtiter plates allowed to keep track of each bead or group of beads and to link them to the corresponding codes. These were released upon Cerium Ammonium Nitrate (CAN) oxidation, derivatized with the silylating agent BSA, and analyzed by EC/GC. Approximately 0.1

pmol of each code was released from the bead but the sensitivity of EC detection allowed the correct reading of the codes.



Scheme 1

The concept of inert molecular tags was developed also by Ni *et al.* [4]. The chemically inert species chosen by this group was a set of secondary amines, which were termed "**hard tags**" because they could withstand a wide range of reaction conditions and the process of code release involved a hydrolysis at 130°C in 6N HCl. With respect to the electrophoric tags, this method employs a differentiated resin to grow the ligand on one side and the code on the other. To produce the resin differentiation, FMOC- and BOC-glycine (in a 9:1 ratio) were reacted with the starting amine resin. By carefully selecting the linker and the protection strategy the library synthesis could proceed on one side or another. Typically a ligand was synthesized starting from a photolinker (which guarantees orthogonality to FMOC- and BOC- protecting groups), while codes were attached after BOC-deprotection (orthogonal to the other protecting groups). The amine hard tags were attached to the resin as iminodiacetic acid derivatives, with the nitrogen protected by the BOC- group. A carboxylic acid reacted with the amine code to provide an amide and the second carboxylic acid was available for coupling to the resin (**Figure 2**).

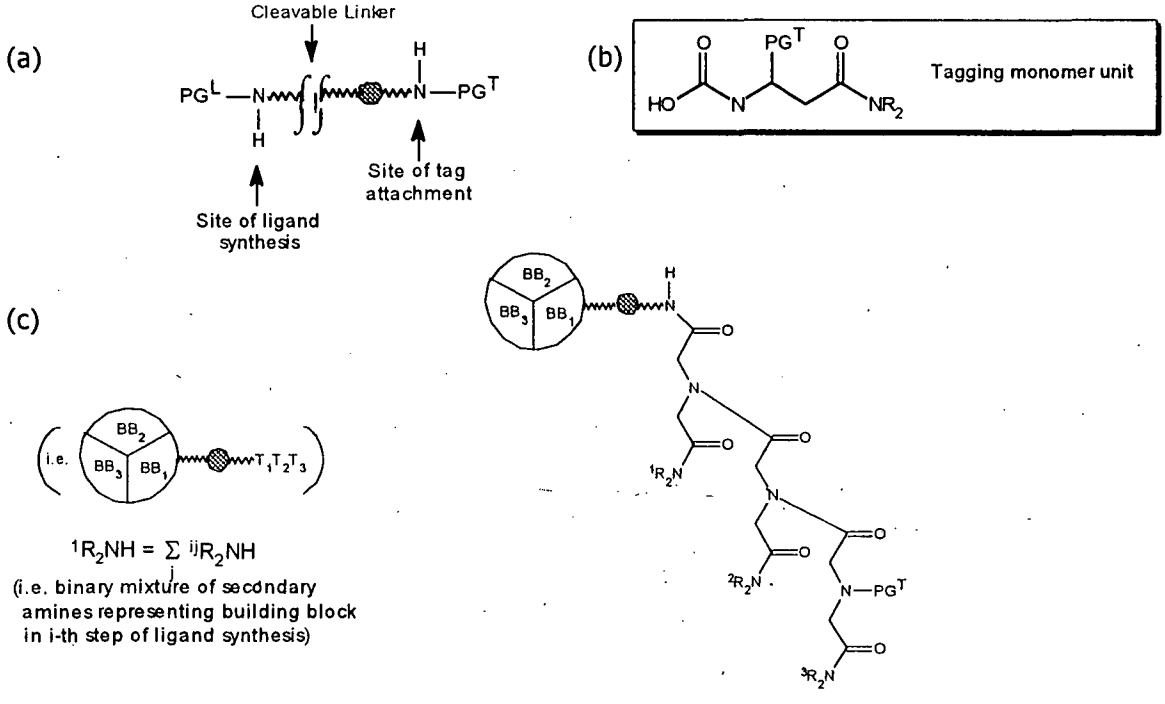
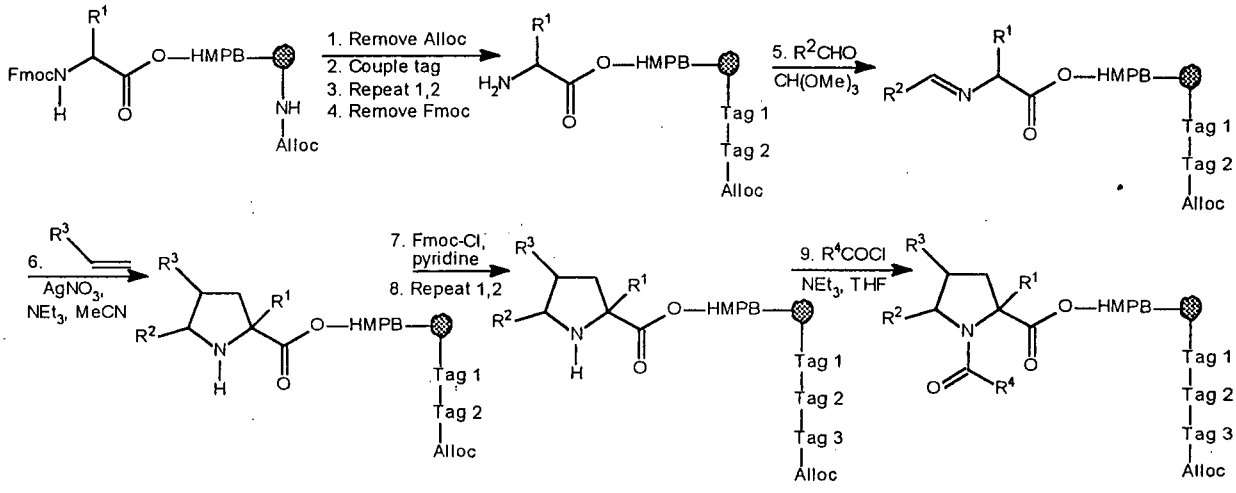


Figure 2

This construct was then totally hydrolyzed under severe hydrolysis conditions (6N HCl; 130°C) and the free amine reconstituted in solution by treatment with lithium carbonate. The solution containing the codes was derivatized with dansyl chloride and these compounds were then eluted on an HPLC column at different retention times, using a fluorimeter as the detector. This instrument allowed the recognition of the dansylated amines to picomolar levels. Also this group applied a binary code to increase the number of encoded permutations. Eighteen different amines with different molecular weight were selected to allow the encoding of a very large number of permutations. A trial library of thiazolidinones and β-lactams was prepared, screened and positives were detected. Beads from positive pools were taken and structure successfully elucidated according to the above process. To compare the ability of library encoding to elucidate the structure of active compounds, with respect to more traditional and slower deconvolution procedures two libraries of pyrrolidines (one encoded and one non-encoded) were prepared by the same group [5]. Three sets of tags (one for the four aminoacids, one for the four aldehydes and one for the five olefins) were used to encode the 240 final members produced in the tagged library (**Scheme 2**).

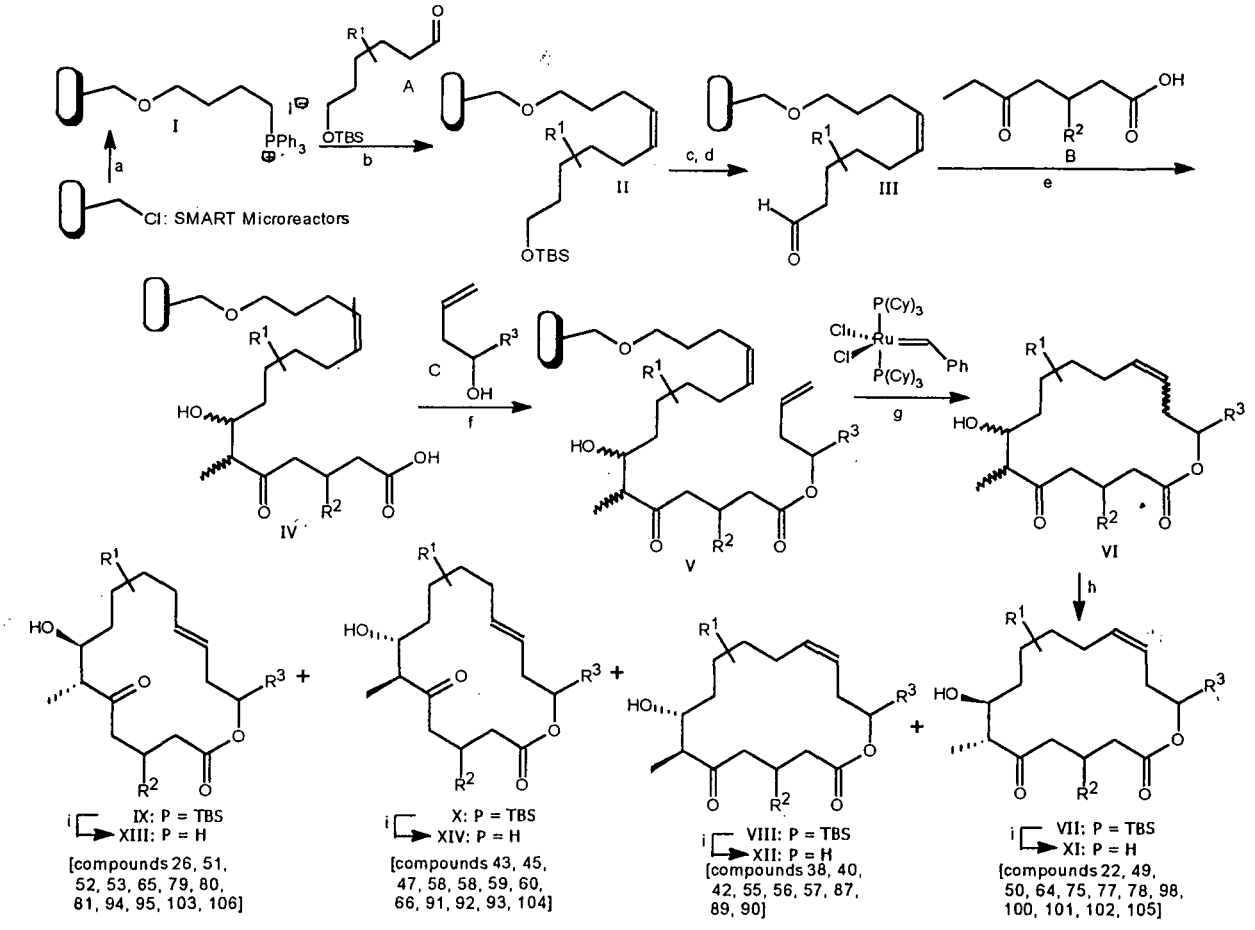


Scheme 2

A total of nine amines was sufficient to cover the whole synthesis, due to the encoding power of the binary strategy adopted. The split/mix synthesis produced 80 pyrrolidines which were pooled and split into three portions, each reacted with a single acyl chloride in spatially segregated vessels, so there was no need for a further encoding step. The non-encoded library required four rounds of deconvolution to identify the most active library components while the encoded library was screened only once and this was sufficient to determine biological activity and further structure determination by reading the associated codes. A remarkable difference between the two methods is observed during the screening: **non-encoded library is typically screened as a mixture, with the risk of false positive and false negative results while encoded library is screened on single beads, thus avoiding those risks.** Although one method is more time-intensive at the level of screening (non-encoded library) and the other is more time-intensive at the synthetic level (encoded library) both successfully identified the same active components. However the ability of an encoded library to deliver substantial SAR information at the level of a high-throughput screening can be considered a remarkable achievement for combinatorial chemistry.

Nicolaou *et al.* [6] have identified epothilones as structures amenable to combinatorial modification by changing the configuration of stereocenters, the geometry of double bonds, ring size and the nature of their substituents. The aim was to quickly establish structure-activity relationships and to start a second, more refined round of synthesis which would ultimately yield a potential drug candidate active against tumor cells. **Radiofrequency Encoded Combinatorial (REC) chemistry** was used to assemble the epothilone library. According to REC chemistry each single monomer is associated with a

unique radiofrequency which comes from a glass encapsulated microchip producing a signal recognized by a reader. The overall process is governed by a dedicated software which allows chemists to sort each radiofrequency tag into special vessels, named Kans, containing the resin. Since the Kans are recognized by their unique signal, all common steps (*i.e.*: washings, deprotections) can be performed in a single flask, saving time and reducing the associated labor.



Scheme 3

***Reagents and Conditions**: (a)1. 1,4-butanediol (5.0 eq), NaH (5.0 eq, nBu_4NI (0.1 eq), DMF, 25°C, 12h; 2. Ph_3P (4 eq), I_2 (4 eq), imidazole (4 eq), CH_2Cl_2, 25°C, 3h; 3. Ph_3P (10 eq), 90°C, 12h; (b).1, sort SMART microreactors (with an Accutag-100) apparatus); 2. NaHMDS (3 eq), THF:DMSO (1:1), 25°C, 3h, **A** (2 eq), THF, 0°C, 3h (~75% from chloromethylpolystyrene loading based on recovered aldehyde upon ozonolysis);4. Pool; (c) 0.2M HCl in THF, 25°C, 12h; (d) $(COCl)_2$ (4 eq), DMSO (8 3q), Et_3N (12.5 eq), -78→25°C (≥95% for two steps, the reactions were monitored by IR analysis of polymer-bound material and by TLC analysis of the products obtained by ozonolysis); (e)1. Sort; 2. **B** (2 eq), LDA (2.2 eq), THF, -78→ -40°C, 1h; then add resulting enolate to the resin suspended in a $ZnCl_2$ (2 eq) solution in THF, -78→ -40°C, 2 h, (≥90% based on recovered aldehyde upon ozonolysis); 3. Pool; (f) 1. Sort, 2. **C** (5 eq), DCC (5 eq), 4-DMAP (5 eq), 25°C, 15h (≥85% yield as determined by recovered heterocycle fragments obtained upon treatment with NaOMe); (g) 1. Separation of individual SMART Microreactors; 2. $[RuCl_2(=CHPh)(PCy_3)_2]$ (0.2 eq), CH_2Cl_2, 25°C, 48h; (h) automated HPLC (SiO_2, EtOAc/hexane or C18, H_2O/THF) or preparative thin-layer chromatography, EtOAc/hexanes. The stereochemistry at C-6 and C-7 as well as geometry of the olefin was tentatively assigned by ^1H NMR spectroscopy; (i) 20% TFA in CH_2Cl_2 (v/v), 25°C, 2-12h. The purity of individual compounds was established by ^1H NMR spectroscopy and HPLC. P= protecting group: TBS= tert-butyldimethylsilyl; DMSO- dimethyl sulfoxide, TFA= trifluoroacetic acid, LDA= lithium diisopropylamide, DCC= dicyclohexylcarbodiimide, 4-DMAP= 4-dimethylaminopyridine.*

The identity of each product can be easily reconstituted by sorting the Radiofrequency tags on the reader and redistributing them in a format convenient to perform the next synthetic step. This system is commercially available from IRORI Quantum Microchemistry and, along with Kans, REC synthesis can be performed also on tubes, with a properly derivatized surface to enable the synthesis, which contain

the radiotag inside. Merrifield resin was weighed into the Kans and, through a spacer diol, transformed into a phosphonium salt. The microreactors were sorted and Wittig reaction performed with a TBS-protected ω-hydroxyalkylcarboxaldehyde. The pooled vessels were jointly deprotected and oxidized to polymer-bound aldehyde. Further sorting and treatment with the dianion of a set of ketoacids yielded a group of carboxylic acids. Mix and split according to the radiotags and ester formation with ω–alkenylhydroxy components provided the key dienes which were submitted to olefin methatesis reaction to give the final library products in solution, as mixtures of four 12,13-desoxyepothilones (**Scheme 3**).

Pure compounds were obtained through preparative TLC, TFA deprotected and epoxidized. Biological tests soon delivered SAR for in vitro tubulin polymerization and in vitro cytotoxicity which are here summarized. Olefin metathesis precursor products were inactive, thus confirming the importance of the macrocyclic structure. Inversion of configuration at C3 resulted in reduced tubulin polymerization potency. Interestingly, α,β-unsaturated lactones retained significant tubulin assembly properties. Substitution of 4-*gem*-dimethyl with a 4,4-ethano group always resulted in loss of activity, pointing to the crucial importance of a proper conformation of epothilones for biological activity. The 6*R*,7*S* configuration also proved to be a key feature to maintain the activity. The C8 methyl group was also a delicate fragment: inversion of configuration, introduction of a *gem*-dimethyl group or removal of the C8 methyl group all resulted in loss of activity. Lower activity was shown by the unnatural epoxides with configuration 12*S*,13*R*. Interestingly, both *cis* and *trans* olefins were active in the tubulin assembly assays and were significantly less cytotoxic with respect to the naturally occurring epoxides. C16 methyl group, when replaced by an ethyl group, was less active. A number of oxazoles replaced the original thiazole, while exhibiting comparable activity. Finally, replacement of the methyl group in position 23 with a phenyl ring gave inactive products. In conclusion, the authors were able to demonstrate that combinatorial chemistry (REC chemistry, in particular) can quickly deliver SAR and facilitate the process of drug discovery in the anticancer research.

Guiles *et al.* [7] recently reported a simple process for mix and sort combinatorial chemistry based on a **visual tagging**. The process utilized two sets of simple color codes: color-coded glass beads and color-coded container caps. The color codes were used after the mixing step to sort into the matrix of a 96-well plate. The colored glass beads were chemically inert, whereas the colored polypropylene caps and the porous reaction vessels were compatible with a wide variety of synthetic reagents. The overall tagging method is economical and allows the recovery of 10-15 mg of compound. This technique allows for two mix and split steps followed by sorting into a parallel array. The library was typically in the X-Y-Z format, where X and Y were each assigned a set of eight bead colors and a set of twelve cap colors, while Z was added throughout a plate. The library was arrayed as 8 X monomers, 12 Y monomers and one Z monomer and was tested with the synthesis of a peptoid library. After the first two sets of monomers were introduced the 96 different products, segregated in 96 porous reaction vessels, were pooled together and treated with reagent Z. Finally the 96 samples from each container were individually sorted into a 96-well plate according to their bead and cap color. The cleavage cocktail was applied and, upon completion of the reaction, the peptoid library members were obtained in 83.5% overall average yield. This is a simple, cost-effective method, does not rely on robotics, computer monitoring controls or spectroscopic techniques and its versatility should make it useful for the encoded synthesis of small to medium size combinatorial libraries.

Xiao *et al.* [6] described a laser optical encoding technique using **Laser Optical Synthesis Chips** (**LOSC**s). LOSC technology combined the most advanced development in laser bar code etching and identification as well as organic synthesis on novel supports. It was made of a two-dimensional 16-digit bar code and a polymeric support for chemical synthesis. A chemically inert alumina ceramic plate carried the 2D-bar code. This was surrounded by the synthesis platform which was made of chemically resistant polypropylene or of a fluoropolymer. These polymers were grafted with polystyrene which, in turn, was derivatized with a proper functional group or linker to serve as the synthesis support. Typical loadings were in the 5-8 μmol range. The overall dimensions were very small (10 x 10 x 2mm) and a huge amount of data could be compressed in the densely arrayed bar code area. Directed sorting

(same as the one used in REC chemistry) was then used to prepare combinatorial libraries. A small camera and a dedicated software were used to read the codes and drive the library synthesis. To prove the utility of such a technique a small array of 27 oligonucleotides in the format of three randomized and one fixed positions (X-X-X-T) was prepared using 27 LOSC devices and the directed sorting strategy. After cleavage the 27 oligos were recovered in 2 to 5 mg quantities, with 67 to 97% purities. MS, ^{1}H-NMR and sequence analysis confirmed the quality of the library members. Potential advantages of this method include: low manufacturing cost, noninvasive encoding, high encoding reliability and capacity, total chemistry flexibility, excellent chemokinetics, easy and clean washing after each reaction, utilization of highly efficient directed sorting strategy, delivery of pure, discrete compounds in multimilligram scale, and amenability to full automation.

References

1. Deprez et al., *J. Am. Chem. Soc.* **1995,** *117,* 5405-5406.
2. Ohlmeyer et al., *Proc. Nat. Acad. Sci. USA* **1993,** *90,* 10922-10926.
3. Baldwin et al., *J. Am. Chem. Soc.* **1995,** *117,* 5588-5589.
4. Ni et al., *J. Med. Chem.* **1996,** *39,* 1601-1608.
5. Maclean et al., *Proc. Nat. Acad. Sci. USA* **1997,** *94,* 2805-2810.
6. Nicolaou et al., *Angew. Chem. Int. Ed. Engl.* **1997,** *36,* 2097-2103.
7. Guiles et al., *Angew. Chem. Int. Ed. Engl.* **1998,** *37,* 926-928.
8. Xiao et al., *Angew. Chem. Int. Ed. Engl.* **1997,** *36,* 780-782.

The Role of Computation in Combinatorial Chemistry

Dr. Anneliese F. Appleton
Molecular Simulations Pty Ltd
Sydney, AUSTRALIA

The development of a new drug involves three key phases, Discovery, Preclinical & Clinical Trials (see **Figure 1**). Over the last twenty years computation has played an increasing role in the discovery phase of the drug development cycle. The development of combinatorial synthetic techniques enabling the production of thousands of new compounds per week has opened up new avenues for the use of computation chemistry in the drug discovery phase.

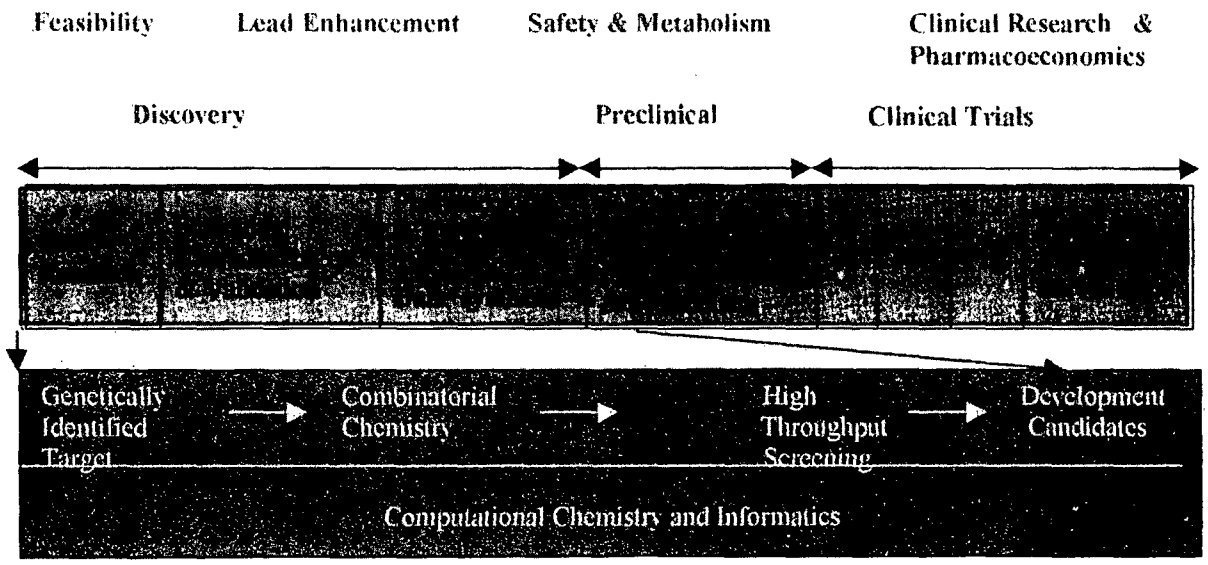


Figure 1: Stages of new drug discovery

The basic research stage of the drug discovery phase can be further broken down into two stages - Lead Discovery and Lead Optimisation. Lead discovery involves the identification of novel or existing compounds that have potential as the basis of a drug. Identification of these compounds have traditionally been obtained by screening libraries of compounds either from natural or synthetic sources. Lead optimisation involves the development of models such as pharmacophores or quantitative structure activity relationships (QSAR) to aid in the understanding of the chemical features that are necessary for activity. These models are then used to guide the modification of the previously discovered leads in order to enhance their efficacy.

Although new combinatorial chemistry techniques have enabled chemists to make many more compounds in a shorter space of time the question of which compounds to make is still key to producing useful compounds in an efficient manner. The combinatorial chemistry software packages developed over the last four years play a key role in assisting the chemist make the decision of which compounds to make.

Combinatorial libraries generally fall into two categories - **discovery libraries** which are synthesized in order to try and identify new lead compounds, and **focussed libraries** which are synthesized in order to optimize the activity of a previously identified lead compound. Discovery libraries need to be as diverse as possible in order to provide a wide variety of chemical characteristics some of which may be suitable for interaction with a specific target enzyme or receptor. Focussed libraries, on the other hand,

are designed to be similar, exploring a variety of compounds which exhibit similar chemical characteristics.

Designing a Diverse Library

So what is chemical diversity and how can it be evaluated? The diversity of a series of compounds is determined by using descriptors which have been selected to represent the chemical properties of molecules. A descriptor is basically just a value that tells you something about a molecule. For example molecular weight or the number of rotatable bonds a molecule has. The initial descriptors used in combinatorial chemistry software were based on those used for the development of QSAR models. For a useful review on the use of QSAR in drug design see the paper by Hugo Kubinyi[i].

A key consideration which had to be taken into account when choosing descriptors for combinatorial chemistry applications is the length of time it takes to calculate a particular descriptor. Unlike QSAR training sets which generally contain only 20-30 molecules, combinatorial libraries can contain hundreds or thousands of molecules, thus the length of time a particular descriptor takes to be calculated is critical. The development of suitable combinatorial chemistry descriptor sets also had to incorporate descriptors which could be used to calculate properties of R groups as well as properties of enumerated compound libraries. Over the last few years much research been focused on the development of new descriptors which can be calculated quickly and provide a large amount of information about a particular molecule.

Molecular Descriptors

Three classes of molecular descriptors are commonly used to describe the molecular properties of a combinatorial library:

- 2D descriptors (which are quick to calculate)
- 3D descriptors
- Fingerprint descriptors.

2D Descriptors include fragment constant descriptors (*e.g.* Hydrogen bond donor or acceptors), electronic descriptors (*e.g.* Dipole moment), topological descriptors (*e.g.* Molecular flexibility indexes), spatial descriptors (*e.g.* Molecular surface area) and thermodynamic descriptors (*e.g.* Log of the partition coefficient).

3D Descriptors include conformational descriptors (*e.g.* Conformational energy), receptor descriptors (*e.g.* Interaction energy between a theoretical receptor surface and a specific compound), quantum mechanical descriptors (*e.g.* Heat of formation) and structural descriptors (*e.g.* Molecular weight).

Fingerprint descriptors such as ISIS and Daylight keys are hexadecimal representations of molecular features, which convey a lot of information is a simple integer string. For more details about the use of "fingerprint" descriptors see the paper by Brown and Martin[ii].

Which descriptors you use to is really dependant on how big your library is and what types of chemistry is represented in the library. Most combinatorial chemistry software packages, like the Cerius2 Diversity module[iii], have a set of predefined set of descriptors which have been optimized for with large combinatorial libraries.

Library Specification

Once you have chosen a set of descriptors you would normally build a virtual library of compounds (i.e. a set of compounds built in the computer and not actually synthesized in the lab). Due to the sheer size of combinatorial libraries, (i.e. 1000's of molecules), traditional molecular building tools which enable the input of molecules into a computer system by sketching the molecule in 2D and then converting it to 3D, are not suitable for specifying a large number of compounds, so the development of combinatorial software tools also lead to the development of more efficient ways of entering molecular structures into a computer system.

Molecular Simulations Incorporated (MSI[iv]) combinatorial chemistry tools provide two methods for library specification, core-based and reaction based approaches. The core-based definition, such as provided in the Cerius2 Analog builder, is suitable when all molecules contain the same core structure with different R groups.
(see http://www.msi.com/solutions/products/cerius2/modules/analogbuilder.html for more details). Reaction based approaches as implemented in WebLab Diversity can provide a more general way of defining the chemical library and are more directly related to the chemical steps undertaken during the synthesis of the library. (see http://www.msi.com/weblab/diversity/index.html)

Data Reduction and Visualization

Once you have built your virtual library of compounds you then calculate descriptors for each molecule in the library. Taking an example where you used 50 descriptors and 1000 molecules you would end up with a table with 50,000 (50 columns and 1000 rows). So how can you possibly tell which molecules are more diverse than others. This is where the techniques of data reduction and visualization come into play.

The most commonly used data reduction technique used in both QSAR and combinatorial chemistry applications is Principal Component Analysis or PCA. PCA searches for relationships between the descriptors which are defined as the independent (X) variables and then creates new variables (called the Principal components) which represent most of the information contained in the independent variables. PCA does this by attempting to discover the true dimensionality of a data set by using linear combinations of the original variables that are orthogonal to each other. Generally PCA is used to reduce the information provided by the 50 or more descriptors into three principal components which can then be visualized as a 3-dimensional plot. For more information on PCA see the reference by Everitt and Dunn[v].

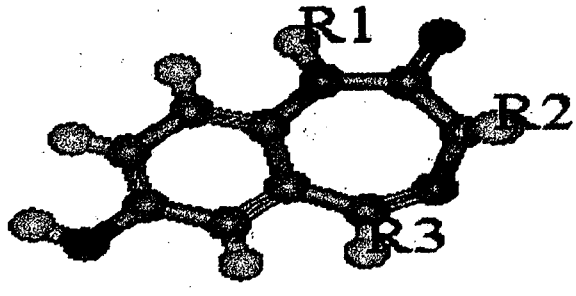
Another data reduction technique which can be used when high dimensional descriptors such as fingerprints are used to describe the molecules is Multidimensional Scaling or MDS. MDS proceeds from a complete distance matrix of the samples, which is either supplied (pairwise molecular similarities or distances are input) or computed from the descriptors. The samples can then be assigned relative position in a cartesian coordinate space where inter-compound distance are respected.

The main objective of either of these data reduction techniques is to condense the information about the molecules in the library represented by the descriptors into three components which can be plotted. Reducing the information into three components means that the molecules can then be plotted as shown in **Figure 2**. This ability to visualize molecules in 3D space makes the analysis and selection of subsets of molecules in the virtual library possible.

Figure 2: PCA plot of a virtual library.

Compound Selection Procedures

In most instances the "virtual library" you have entered into the computer to analyse will either be to big to make or contain a lot of molecules which are very similar to each other and thus synthesizing them would be unnecessary as they would not provide any additional chemical information. A good example of a virtual library is the benzodiazepine designed by Ellman[vi]. In this example (illustrated in **figure 3**) the researcher started with a virtual library of 5,760 compounds, resulting from the addition of 16 alkyl-halides in position R1, 18 amino acids in position R2 and 20 acid chlorides in position R3. Diversity analysis of this virtual library revealed that the chemical functionality of the compounds present in the 5,670 virtual library could be represented by just 672 compounds made by using just 7 of the possible 16 alkyl halides in R1, 8 of the possible amino acids in R2 and 12 of the acid chlorides in position R3.



R1: Alkyl-halides (16)

R2: Amino-acids (18)

R3: Acid-chlorides (20)

Figure 3: The Ellman Library

So how do you go about choosing which molecules in your virtual library to keep and which ones to through away? During the last four years a variety of compound selection procedures have been developed and implemented in combinatorial software tools. MSI's approach to the design and implementation of combinatorial software tools has been guided by members of a consortium set up specifically to address the needs of researchers interested in the application of combinatorial to drug design. For more information about this consortium see http://www.msi.com/solutions/consortia/cchem.html.

MSI's Cerius2 combinatorial chemistry software provides compound selection procedures which include techniques for the identification of outliers, selection of diverse or similar compounds as well as tools for comparing and augmenting combinatorial libraries. All of these compound selection procedures rely on the analysis of the molecule libraries in 3D space.

"**Outliers**" in a virtual library are identified as those molecules whose distance to the centroid, defined by the full set of molecules plotted in 3D space, is greater than the average distance to the centroid. Whether or not you chose to remove "outliers" is very dependant on the purpose for which the library is being synthesized and in some cases these "outliers" may provide key pieces of information which could result in the discovery of a new lead.

A variety of methods are available for the selection of diverse compounds, including cluster-based, distance-based and cell-based selection procedures. **Cluster-based selection** is based on the partitioning of a compound library into classes or categories consisting of elements of comparable similarity. A diverse set of molecules are then derived by selecting a representative, normally the centroid, of each cluster. Once again visualization plays a key roles in this type of analysis with the molecular clusters being visualized as dendograms (see **Figure 4**). Clustering algorithms commonly used for this type of diversity classification include Jarvis-Patrick clustering, relocation clustering and hierarchical cluster analysis.

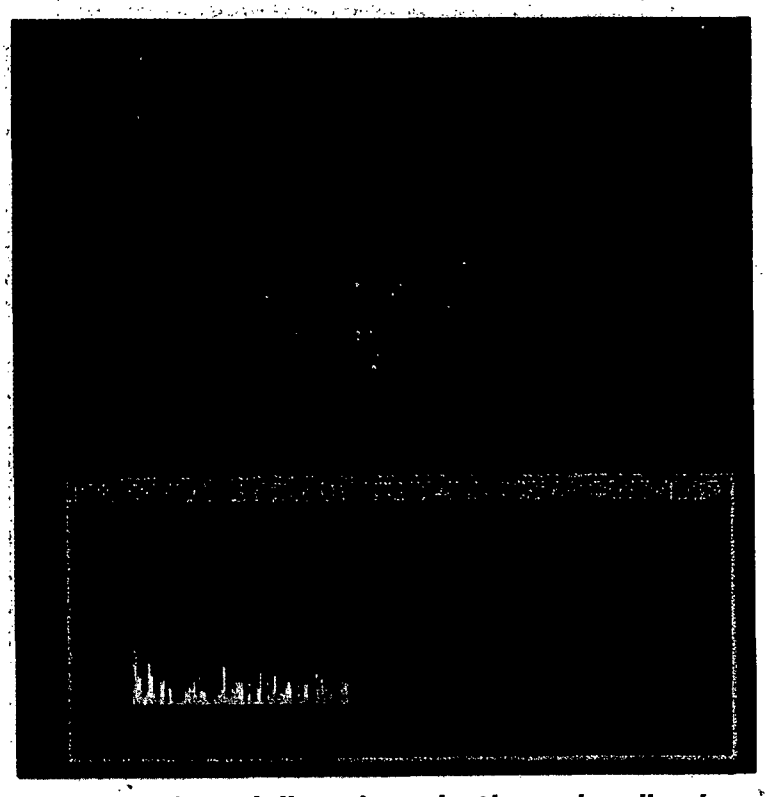


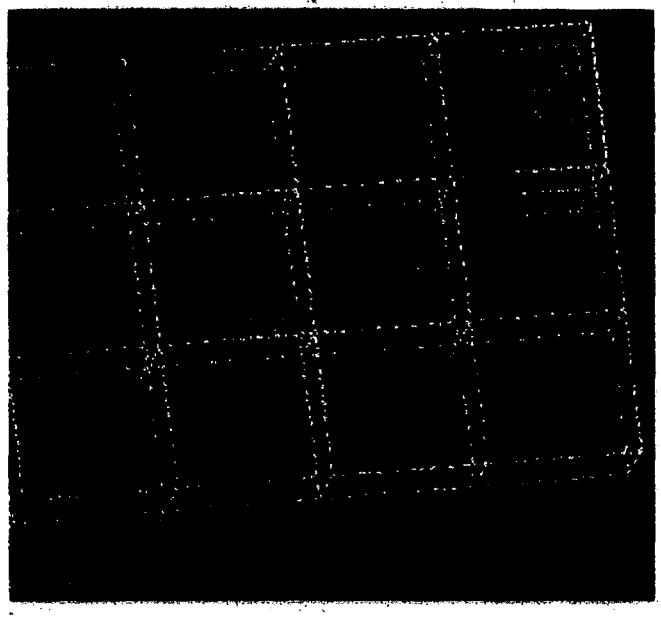
Figure 4: Cluster-based diversity selections visualized as a dendogram and then coloured on the PCA plot.

An alternative diversity selection procedure is distance-based selection. In **distance-based** approach, selection of highly diverse subsets of molecules from the virtual library is based on the stochastic optimization of diversity functions using a single-point-mutation Monte Carlo technique. The diversity target functions for distance-based selections are evaluated from inter-compound distance information. Diversity metrics used in Cerius2 include the MaxMin function[vii] which maximizes the minimum squared distance from each point to all other points in the selected subset of molecules, the PowerSum function which maximized the inverse of the sum of the reciprocals of the squares of all intermolecular distances between selected points, the Product function which maximizes the product of the squares of the intermolecular distances, and the MaxMinSpan Tree function which is based on calculating the minimum spanning tree for each subset of selected point in the stochastic optimization process, which is then used to compute an error function based on the length of each edge in the tree (see **Figure 5**).

Figure 5: MaxMinSpan Tree representation for a virtual library.

In the **Cell-based diversity selection** approach, the descriptors of the molecules in the library defined as independent variables are "binned" or grouped, to divide the property space into cells (see Figure 6). Diverse molecules are then obtained by selecting models from different cells. For each filled cell (i.e. cells containing molecules), the model closest to the cell centre is selected.

Figure 6. Diversity selection of compounds using a cell-based approach



Selecting Similar Compounds

After lead molecules with moderate activities have been identified, you may then want to design a combinatorial library which focuses on compounds which exhibit similar molecular properties to that of the lead molecule. This is often referred to as "library focussing". Searching for similar molecules is especially important during the optimization phase of library design. The selection of similar compounds basically works that same way as the selection of diverse compounds except in reverse.

Library Comparison and Augmentation

Rather than synthesizing specific combinatorial libraries many companies now purchase commercially available compound libraries. Cerius2 also contains a tools to assist the chemist decide what libraries would be best to purchase. This is done by 3D comparison of a virtual representation of the commercially available libraries where the coverage of molecular property space as represented via a PCA plot can be compared both qualitatively and quantitatively. These tools also help the chemist to identify holes in property space which could be filled by purchasing selected molecules from commercially available libraries (see **Figure 7**).

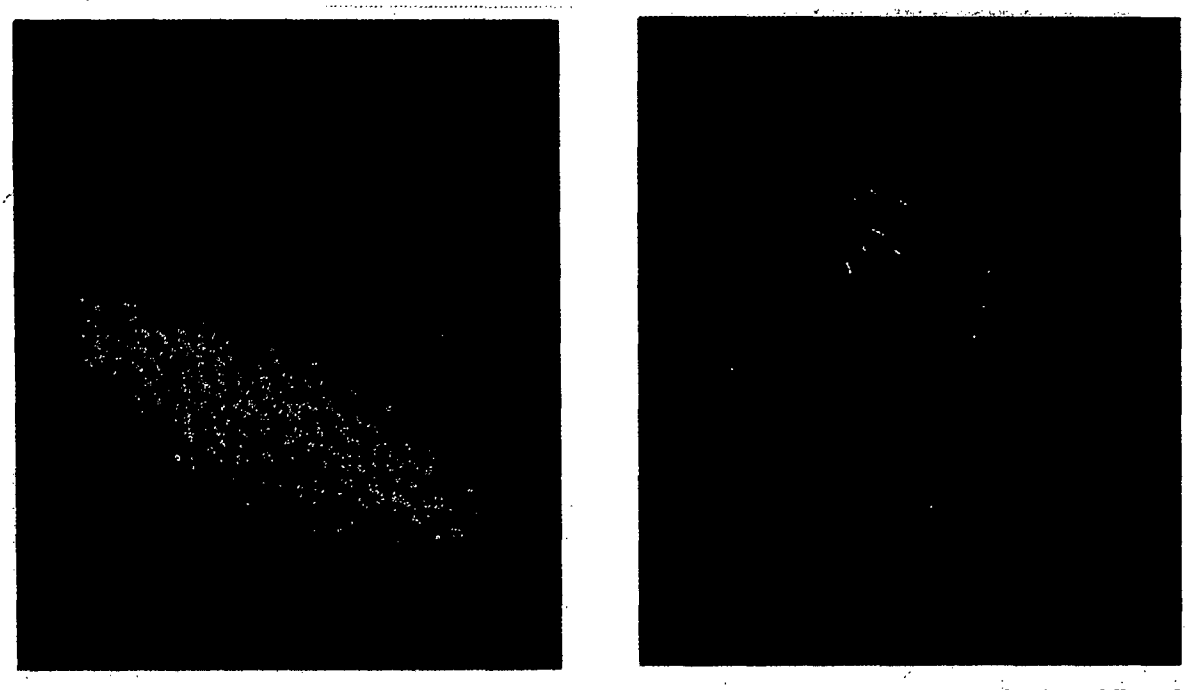


Figure 7: Comparison of three different compound libraries (left) and identification of compounds from the red library to augment the green library (right).

Lead Optimization

Once a lead molecule or class of molecules has been identified then the traditional rational drug design tools can assist in the design of focussed libraries aimed at optimizing the compounds activity (see **Figure 8**). When the target receptor structure for which the library is being designed is unknown QSAR and Pharmacophore modelling approaches can be used. For more details about these drug design techniques see 3D QSAR in Drug Design by Kubinyi[viii]. When the target of the receptor structure is known then information about the size and chemistry of the active site can also be used to design focussed libraries.

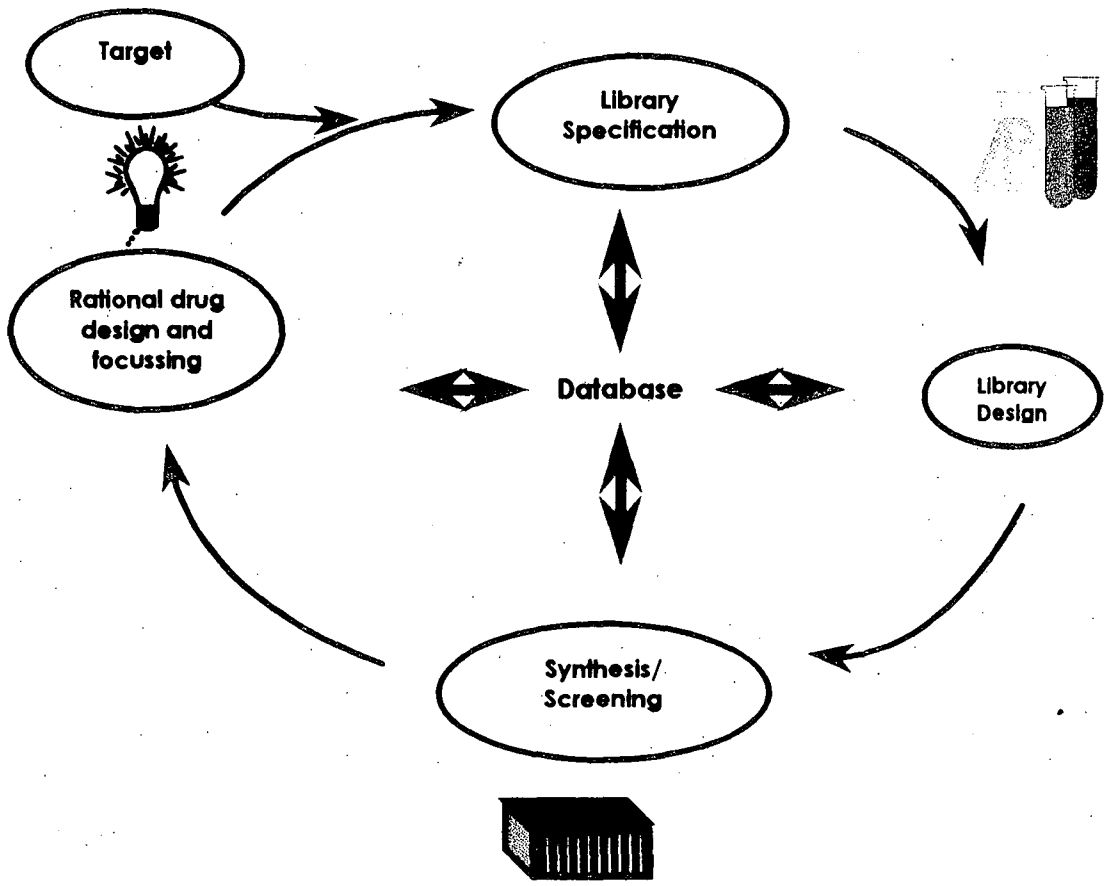


Figure 8: The combinatorial chemistry design cycle.

A new approach has recently been developed by MSI called "Structure-based focussing". **Structure Based Focusing** is a method which uses the known or suspected active site of a protein to select compounds which are likely to bind within the defined active site. This approach is often referred to as Virtual High Throughput Screening (VTHS). The defined active site is first analyzed to generate an interaction map for the active site consisting of a list of features (such as lipophilic, hydrogen donor, hydrogen acceptor) that a ligand is expected to satisfy for a reasonable interaction with the protein. A set of three dimensional queries is then derived from the interaction map. For a given library of compounds, a conformationally flexible database is constructed. This database is searched with the set of queries. The resulting hits consist of various conformers of a subset of compounds that satisfy one or more queries and therefore are expected to fit the active site reasonably well. These hits are then scored using the scoring algorithm developed by Bohm[ix] and implemented in the receptor based design program, Ludi.

Conclusion

The development of combinatorial chemistry has catalyzed the design and implementation of a completely new set of software tools which are capable of handling thousands of compounds. The software used for combinatorial chemistry applications has brought the field of computation chemistry into the bioinformatics arena - facilitating integration of databases used to catalogue libraries with the output from drug design tools.

[i] H. Kubinyi. *QSAR and 3D QSAR in drug design - Part 1*, Drug Discovery Today, 2, 457-467, **1997** & *QSAR and 3D QSAR in drug design - Part 2*, Drug Discovery Today, 2, 538-546, **1997**

[ii] ISIS (SSKEYS) and Daylight fingerprints combined with Ward's clustering the method of choice. Brown, R.D.: Martin, Y.C. J. Chem. Inf. Comput. Inf. Comput. Sci., 36, 572-584, **1996**.

[iii] For more details about Cerius2 Diversity, see the MSI website: **http://www.msi.com/products/solutions/cerius2/modules/diversity.html**

[iv] Molecular Simulations Inc. 9685 Scranton Road, San Diego, CA 92121-2777, **http://www.msi.com**

[v] Everitt, B; Dunn, G. **Applied Multivariate Data Analysis**, Oxford University Press: New York, **1992**.

[vi] Ellman, J.A. *Design, Synthesis and Evaluation of Small Molecule Libraries*, Acc. Chem. Res. 29, 132-143, **1996**.

[vii] M. Hassan, J.P. Bielawski, J.C. Hempel, and M. Waldman, *Optimization and Visualization of Molecular Diversity and Combinatorial Libraries*, Molecular Diversity, **2**, 64-74, 1996

[viii] Kubinyi, H, *3D QSAR in Drug Design: Theory Methods and Applications*, Leiden, ESCOM, **1993**.

[ix] Bohm, H.J. *The computer program LUDI: A new simple method for the de-novo design of enzyme inhibitors*, J. Comput.-Aided Mol. Design, 6, 61-78, **1992**; Bohm, H. J. LUDI: *Rule based automatic design of new substituents for enzyme inhibitor leads*, J. Comput.-Aided Mol. Design, 6, 593-606, **1992**.

Combinatorial Chemistry: Patenting Issues

Dr. Pierfausto Seneci
GlaxoWellcome Medicines Research Centre
Via Fleming 4, 37100 Verona, ITALY

Patenting a class of chemical entities is, here as in many other fields, the key that eventually leads to the return of investments and prevents the insurgence of competition. The change in the pharmaceutical market summarized in **Figs. 1 and 2**, though, has a strong impact also on patenting policies. Given that a patent lasts 20 years, the increased time to the market has reduced the profitability time window for companies; if a patent is filed and 15 years are used to reach the market, only five years of sales without generic competition are granted. Thus filing a patent in late phases of Drug Discovery looks appealing; only more assessed and promising drug candidates are patented, reducing the significant patent costs, and a larger profitability window is available.

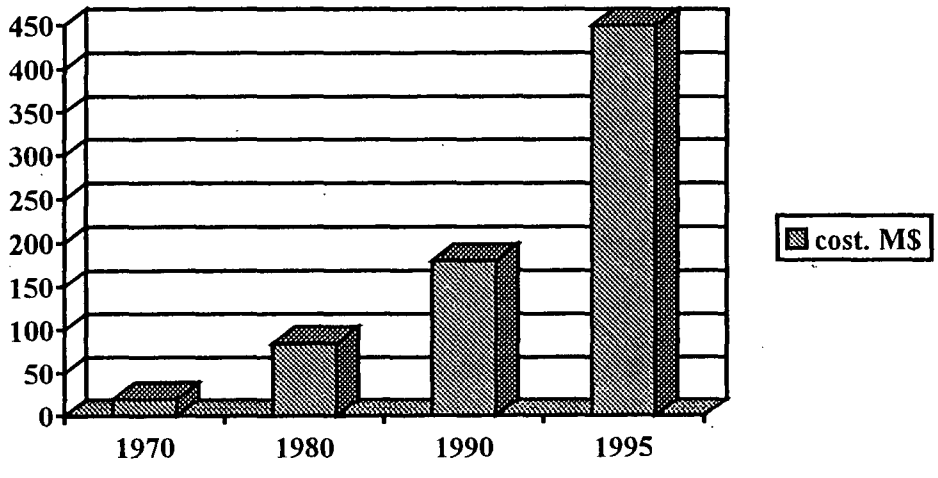


Figure 1

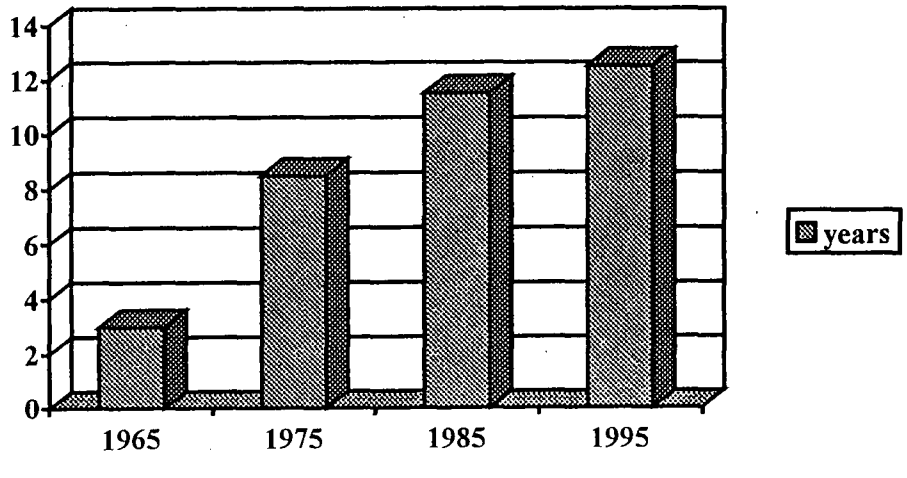
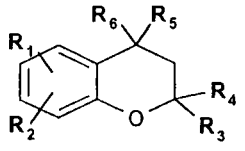


Figure 2

High throughput chemistry and biology have introduced an additional variable to the patent protection equation in Drug Discovery: chemical libraries, i.e. large collections of chemical compounds, can be patented and may either represent prior art to hinder competitive research on the same structural class or even a way to claim large collections/libraries of compounds for specific applications. A conservative patenting approach can be severely damaged if competitors are actively exploiting the same biological target using high-throughput chemical and biological strategies and patent early their results. A careful evaluation of the risks versus benefits of awaiting to file patents during the Drug Discovery process should be made, and the competition activity should be monitored regularly to apply the best patenting strategy for a specific project.

Combinatorial Technologies-related patents have appeared since early 90s, and their number is steadily growing: we can divide them mostly into structure-based patents and technology-based patents. The first are broad patents claiming chemical classes of compounds and/or their screening on large families of targets (enzymes, receptors, whole cells and so on). Some claimed generic or specific structures are reported in **Fig. 3** together with the patent number and the claimed biological activities. Technology-based patents span a wide range of applications including methods for library synthesis, tagging methods, synthetic and analytical combinatorial instrumentation. A sampling of these patents is reported in **Fig. 4** together with their main claims and contents.

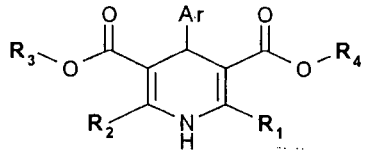


WO9530642 (1995)
Pharmacopeia

Generic dihydrobenzopyran libraries for lead generation

Activity on carbonic anhydrase

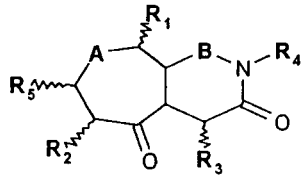
Treatment of glaucomas



WO9633972 (1996)
Affymax/Glaxo

Generic dihydropyridine libraries for lead generation

Broad pharmaceutical activity claimed



WO9715577 (1997)
Molecumetics Ltd.

Generic scaffold for peptide reverse turn mimics claimed

Libraries based on claimed scaffolds were claimed

Biological activity of claimed libraries were claimed

Figure 3

A few additional comments can be made. Many combinatorial patents have been filed by small biotechnology companies to let the scientific world appraise their libraries (or technologies), rather than really to protect them. Major pharmaceutical companies have not patented that much in the field up to now; the more conservative approach of patenting well characterized, more downstream compounds is still prevalent. The time lag between first filing and patent publication may disprove such sentence, though, in the near future. Most of the accessible patent applications have prior art which was discovered during the International Examiners' Search Report, and their relevance should be strongly affected. No major litigations concerning claimed libraries or technologies have appeared yet, but both subjects should cause such events to appear in the near future (see for example structures and claims of Figs. 3 and 4).

WO9309668 (1993)
Affymax

Light-directed masking/unmasking strategies for biooligomer library synthesis

Automated instrumentation for reagent delivery

WO9320242 (1993)
The Scripps Institute

Oligonucleotide encoded SP pool libraries

PCA deconvolution

WO9408051 (1994)
University of Columbia

Chemically encoded, bead-based SP pool libraries

Several chemical tagging methods exemplified

Several tagged SP libraries exemplified

WO9512608 (1995)
Affymax

Automated instrumentation (vessels, manifolds, flow-lines, agitation, etc.) and software to perform and control mix and split encoded SP library synthesis

Methods to transfer resin slurries using the above mentioned instrumentation

US5463564 (1995)
3-D Pharmaceutical, Inc.

Computer-based processes to define chemical libraries with selected properties

Pharmaceutical exploitation of the rational design

WO9624061 (1996)
Ontogen Corp.

Radio frequency encoded microchips

Application to SP combinatorial synthesis

Figure 4

Some recent reviews (1-7) cover the field of Combinatorial Technologies' patenting: any combinatorial scientist must know that current and future intellectual property is going to be influenced by technological breakthroughs and by the increase of popularity of Combinatorial Technologies in many other application fields.

References

1. Newton, C. G., *Exp. Opin. Ther. Patents* **6**, 827 (1996).
2. Bozicevic, K., *Annu. Rep. Combi. Chem. Mol. Diversity* **1**, 298 (1997).
3. Newton, C. G., *Exp. Opin. Ther. Patents* **7**, 1183 (1997).
4. Green, G. D., *Mol. Diversity* **3**, 247 (1998).
5. Jakobsen, P. H. et al., *Exp. Opin. Ther. Patents* **8**, 1157 (1998).
6. Mulligan, L. and Steele, P., Exp. Opin. Ther. Patents **9**, 507 (1999).
7. Harper, M. F. and Newton, C. G., *Exp. Opin. Ther. Patents* **9**, 583 (1999).

Chemical and Biological Methods for Peptide Libraries Generation

Dr. Giorgio Fassina
Tecnogen SCpA
Parco Scientifico, 81015
Piana di Monte Verna (CE), ITALY

Introduction

In spite of the poor biostability and bioavailability of peptides and consequently their scarce attitude to become good drug candidates, libraries of these kind of compounds were the first to be produced by chemical approaches, given the availability of efficient and high repetitive yield synthetic protocols for peptide synthesis [1-4]. Peptides are particularly suitable for the construction of libraries since a high degree of structural diversification can be easily achieved simply by varying the peptide sequence length or by introducing different amino acids other than those naturally occurring. The number of different peptides which can be obtained by a combinatorial approach is governed by the simple formula:

$$N = b^x$$

where **N is the total number of molecules, b is the number of residues used in the construction of the library and x is the library sequence length.** In Table I are reported some examples of the number of distinct molecules attainable varying the sequence length and using different numbers of building blocks for library synthesis.

TABLE I
No. of molecules

X	b = 20	b = 25	b = 30
2	400	625	900
3	8.000	15.625	27.000
4	160.000	390.625	810.000
5	3.200.000	9.765.625	24.300.000
6	64.000.000	244.140.625	729.000.000

Types of synthetic peptide libraries

All synthetic peptide libraries can be classified in two main categories: **peptide libraries bound to a solid support (solid-support peptide libraries)** and **soluble peptide libraries.** Both types are, in most cases, synthesized applying the solid phase method, but they differ by the way they are screened after preparation. The **solid-support libraries** are screened while still attached to the solid matrix on which they have been synthesized, after removal of the side chain protecting groups. They are usually prepared on the polymeric resin used for peptide synthesis, but many other supports such as plastic pins (5), cotton (6), paper (7), or microchip surfaces (8) have also been used. The resin-bound peptide libraries are based on the **one-bead-one-peptide concept** (3, 9) and are prepared using the **Portioning-Mixing** (PM) method firstly described by Furka in 1988 (1,10) and subsequently adopted by many others.

The **Portioning-Mixing** involves the splitting of the resin support into *n* equal aliquots, coupling a single activated monomer or performing separate reactions on each individual fraction, and then mixing the resin aliquots together. The repetition of this protocol for a total of *x* cycles produces a collection of n^x different peptides in equimolar amounts. Since each resin bead is subjected to individual reactions, the final result is a population of beads each one displaying a different peptide on its surface. The number of resin beads used in the synthesis should exceed the number of peptides in the library, in order to get an adequate representation of all sequences. Generally a ratio of ten beads-one sequence is used. At the end of the synthesis cycles, the library is composed of a set of sub-libraries where the nature of the last building block added or reaction performed is known.

The decoding step for libraries released from the solid support is based on the activity screening of such sub-libraries. Resin-bound peptides are submitted to a one-step screening process where all components are concurrently tested in a colorimetric assay and the active compounds are identified by sequencing the isolated bead. A great limitation of solid phase peptide libraries is that only binding experiments with target molecules can be carried out, and consequently positive beads may contain just ligands and not inhibitors of the target molecule used for screening. Another disadvantage of working with these libraries is that the number of resin beads required to maintain the representation of all the sequences restricts the number of different molecules obtainable, and, unless working with kg of resin, only libraries containing up to five random residues can be prepared (2,3,9).

The **soluble peptide libraries** are much more versatile, since, by their nature, can be used virtually in any screening assay. They are prepared either applying the **Portioning-Mixing method** by using solid phase (3) or solution phase (11) approaches, or using the **Mixture Synthesis** method (12). The Mixture Synthesis method involves the couplings of mixtures of activated monomers at each cycle of the synthesis. The product distribution, in this case, is strictly influenced by the relative kinetics of the competing reaction, and, unless corrections in the relative concentrations of activated amino acids are introduced, the required equal representation of components in the mixture is not easily achieved. However, very complex peptide mixtures have been generated using this method, with and without concentration corrections, and successfully screened for biological activity (13,14).

Approaches for the synthesis of peptide libraries

Several procedures have been reported for the multiple synthesis of peptides [15,16] or peptide libraries [17], and not all of them require automated instruments or tailored laboratory equipment, which are not of easily found in most laboratories. Simple procedures for the multiple synthesis of peptides or for the preparation of peptide libraries in the micromolar scale have been developed, requiring only very common laboratory equipment such as a vortex equipped with a sample holder for 25 Eppendorf tubes and a small centrifuge for polypropylene test tubes [18]. In addition to peptide libraries synthesis, this procedure can be applied very conveniently also to the simultaneous small scale manual synthesis of at least 30 different peptides.

The synthesis can start suspending 0.1 mmol of resin for solid phase synthesis in a DCM: NMP (6:4 v/v) mixture. This solvent composition allows a homogeneous dispersion of the resin, making the aliquoting for libraries preparation very simple and convenient by simple pipetting the desired amount of suspended resin in the polypropylene test tubes. Subsequently NMP is added to each aliquot, which after vortexing is centrifuged. The resin separation from solvent proceeds very easily in another solvent mixture, DCM:NMP 1:4 v/v. Excess solvent is then removed by vacuum aspiration, using a needle connected to a water vacuum pump. Each different resin aliquot is then treated with a solution containing the appropriate activated amino acid.

This method provides a low-cost, easy approach to peptide library synthesis for laboratories whose needs do not justify the costly investment in an automated peptide synthesizer. All the reagents and

laboratory instruments are commercially available at a very low cost, and there is no need for a custom-made apparatus. In addition, this method has a more general applicability, since it can be used whenever a solid phase synthesis has to be performed, either to prepare libraries (linear, multimeric, cyclic, or peptoidic) or single peptides. The small synthesis scale employed allows the preparation of only few micromoles of peptides, but this amount is more than adequate for a vast array of biological assays.

Screening of peptide libraries

Different methodologies have been developed for the screening of resin-bound peptide libraries. A wide variety of these methods rely on the presence of a target molecule tagged with a detectable moiety such as biotin (3), or a radioactive isotope (19), which is incubated with the resin beads population. The active beads interacting with the labeled target molecule will be identified, isolated from the rest of unreactive beads, washed to remove the bound target molecule, and subjected to microsequencing to determine the sequence of the active peptide. This method is straightforward, since all the beads are simultaneously tested for activity and no decoding steps are required. However, this approach lacks of rapidity and simplicity, which are very important factors, since in the search for active peptides several sub-libraries have to be tested, often under different experimental conditions.

Detection of peptide-beads complexes by means of secondary chromogenic reactions requires additional bead cleaning procedures prior to peptide characterization by micro-sequencing. In addition, the target molecule binding peptide-bead has to be first identified and then removed from the entire peptide library, which is often composed of several millions of beads. This is often difficult and tedious. The methodology for peptide-beads selection and identification may be improved by coupling the target molecules to magnetic beads and isolating the complexes resin beads-target simply by exposure to a magnetic field (20).

Resin-released libraries can be conveniently used for the search of molecules able to interfere in solution with a specific biochemical recognition event, such as in the case of hormone-receptor, antigen-antibody, or inhibitor-enzyme interactions. Screening can be conveniently performed evaluating the inhibitory activity of sub-libraries, where the nature of at least one functional group of the library is known in a predetermined position, on the assay under consideration. This allows the identification of the first functional group in the library responsible for activity. The complete deciphering of the active structure must then follow iterative cycles of synthesis and screening steps, where other sub-libraries are prepared, all of them with the previously identified functional group in the predetermined position (n) in the scaffold, and for all of them with the functional groups in the n+1 position known. The sub-libraries are screened again for activity, leading to the identification of the n+1 functional group responsible for activity.

The number of iterative synthesis-screening cycles depends consequently on the number of different functional groups on the library. Alternatively soluble libraries can be immobilized, again in the sub-library format, on solid supports such as microtiter plates for ELISA determination. The target molecule, labeled with a reporter compound such as chromophores, radioactive isotopes, biotin, or enzymes, is incubated on the plates. The sub-library with the highest activity for the target will be easily detected, and repeating as before iterative cycles of synthesis and screening the structure of the active compounds will be elucidated. Soluble libraries can be screened also directly on biological assays, including those with living cells, or isolated organs, or bacteria. As before, sub-libraries are tested individually in the biological assay, and by activity evaluation of all the different samples the active structure is decoded.

Quality control of synthetic peptide libraries

Reaction yields in peptide chemistry are high enough to assure a good quality of final products, but when mixtures are synthesized the sequence-dependent difficulties in peptide bond formation or N-terminus deprotection cumulates and an exponential number of side products could be generated. While the presence of this large population of side products in the libraries could represent a further contribution to the diversity of the library and consequently a higher probability to find active molecules, for converse, the formation of so many impurities could result in a library difficult to be decoded or libraries lacking one or more of the desired components. In most cases the quality control is entrusted to the screening assay. If positive results are obtained after the screening step, this is enough to validate the entire process.

A problem with using biological data from libraries without or with only minimal analytical data is the interpretation of negative results. Without analytical proofs of the existence of the expected components, it could be dangerous to assume that the library is inactive. Analytical studies must be performed on the newly synthesized libraries, at least to retrieve information on the presence of a given molecule or family of molecules, and on the distribution of the components in order to establish if they are equimolarly represented. Techniques like **RP-HPLC**, **amino acid analysis**, **sequencing**, **MALDI-TOF** have been already successfully applied to the analysis of peptide libraries (18, 21-25). Indirect methods, such as monitoring of synthesis progression, parallel synthesis of single test compound or test mixtures with fewer components, may also be used to evaluate the quality of synthetic peptide libraries.

RP-HPLC is the most powerful tool for the characterization and purification of polypeptides from any source. Nevertheless its use for the analysis of synthetic peptide libraries finds useful applications only in those cases where the number of mixture components does not exceed twenty. The components of libraries containing a higher number of molecules are very difficult to be separated, since they are usually synthesized in such a way that families of analogs are generated with very similar physico-chemical properties. This happens for example when the libraries are organized in sub-libraries where many residues are identical and only few others are randomized. From the RP-HPLC analysis of small peptide libraries, useful information on the number and the relative concentration of the components can be deduced (from the number of peaks in the chromatograms and the corresponding integrations), but for a more detailed understanding of the actual composition of the mixtures other studies must complement these investigations.

RP-HPLC could provide much more useful indications for the characterization of even complex mixtures when interfaced directly to mass spectrometers (**LC-MS**), since in this case an additional, orthogonal method of separation is introduced (21-22).

Capillary electrophoresis extends and complements analytical information derived by HPLC. The separation principle of CE is based on the charge differences of the compounds to be separated, and when working with peptide libraries these differences can be pronounced, since a library of length n composed of all natural amino acids shows a charge distribution between $-n$ and $+n$. The presence of charges depends on the nature of the side chains and on the pH value of the buffer in which the electrophoresis is carried out, therefore the technique could result particularly useful for the identification of problems occurring during the final cleavage step to remove the library from the resin used for synthesis. If the number of component is not too high, they can be completely resolved choosing the appropriate pH, voltage and ionic strength conditions. When more complex mixtures are analyzed, broad bands are observed in the electropherograms, corresponding to peptides families having the same approximate charge (21). Integration of the area under these bands, taking into account the effects of the charge differences on the migration rate, can be used to deduce a quantitative estimation of the number of compounds in each family. From this point of view, data from capillary electrophoresis are more satisfactory than those obtained from the RP-HPLC technique, since it

is much more difficult to predict the behavior of a given group of molecules. Nevertheless, also information retrieved from this technique must be complemented by other thorough investigations.

Amino acid analysis is not a powerful method for the full characterization of peptide libraries since it is not sufficient alone to give conclusive information on its actual composition, however it represents a rapid and versatile approach to evaluate the distribution of the components in that mixture or to check the presence or absence of some peptide families. It can be applied with both soluble (lyophilized) or support-bound libraries, since the conditions of hydrolysis are strong enough to remove the peptides from any kind of resin or other solid surfaces like paper, cotton or glass. Even in the case of soluble libraries, the analysis can be carried out directly on resin in order to retrieve the desired information on amino acid composition before the cleavage step. When strong deviations between the theoretical amino acid compositions and experimental values are found (after correction for the differences in the hydrolysis and oxidation rates of the different residues), the library can be discarded without performing any other investigation. If the found values fit very well the expected ones, there is a high probability that all molecules are represented in the mixture. A main drawback with using only data obtained from amino acid analysis to judge the quality of a given library is that side products deriving from incomplete side chain deprotection or side chain modifications can not be detected, since during hydrolysis the integrity of the amino acids could be often restored. The amino acid composition of the library is deduced by comparison of the peak areas with those of a standard pre-calibrated mixture of amino acids.

Mass spectrometry methods are powerful tools for the characterization of compounds from any source. The determination of molecular masses of peptides can not be achieved using conventional methods of ionization (like electron impact) since peptides are labile molecules which decompose if heated to temperatures beyond the melting point. The two new techniques of soft ionization, Matrix Assisted Laser Desorption Ionization (MALDI) and Electrospray ionization (ESI), may be used to characterize peptide libraries. These techniques can be employed to measure the mass of biomolecules having molecular weight higher than 200.000 Dalton and can provide information on their structural properties. Given the very high sensitivity (picomoles to femtomoles) and accuracy (0.01% or more, depending on the analyzer), they result particularly suitable for the analysis of complex mixtures like synthetic peptide libraries, where the concentration of the mixture components can be very low with molecular weights very similar to each other. Both ESI and MALDI, but also the older Fast Atom Bombardment (FAB) technique (21), have been successfully used to evaluate the composition and purity of more or less complex synthetic peptide mixtures (22). MALDI is unique in its capacity to analyze very complex mixtures, since spectra are not complicated by the presence of molecule fragments or multi-charge ions as in the ESI (the dominant ion produced by MALDI is the singly charged ion).

When interfaced to HPLC or capillary electrophoresis ES becomes the most powerful method for the characterization of even very complex mixtures [22], since the combination of the two techniques allows the identification of compounds having very similar chemical properties,

Peptide libraries by biological routes

Biological methods for library preparation are mainly limited to peptide or oligonucleotide libraries. For peptide libraries, methods are based on the construction of a pool of clones each one expressing a different peptide on its surface. The peptides are fused to proteins normally expressed on the surface of the microorganism used. Phage display libraries are the most commonly used. Screening is accomplished by incubation of the target molecule, adsorbed to a solid support, with the phage population. Active phages will bind the target even after extensive washing steps. Target-bound phages are isolated and propagated by infection of E. coli and subjected to an additional round of adsorption to the immobilized target. This procedure increases both the number of active phages and the stringency

of selection, since harsher condition may be employed in the washing steps to reduce the number of non-specifically bound phages. As for the case of synthetic libraries, iterative cycles of adsorption, washing, elution and propagation in E. coli are performed to enrich the phage population in the active or in few active sequences. Active phages may then be subjected to DNA sequencing in order to decode the active peptide sequence. In a very similar way, also oligonucleotide libraries can be screened for immobilized targets using the polymerase chain reaction (PCR) methodology to expand the number of active sequences after each selection cycles.

The use of biological display libraries for the isolation of peptide ligands is an interesting alternative to chemical libraries. Since 1985 [26], when this technique was first published, many fields of research have benefited from its use. Biological display libraries are constituted by large pools of microorganisms (up to 10^9-10^{10}), each one expressing a different polypeptide on its surface. These libraries can be easily propagated and used in repeated cycles of selection over the target molecule. In a typical experiment, the library is incubated with the target bound to a solid support; the bound microorganisms are eluted, grown and selected over the target 2 to 4 more times. At the end, single clones are easily isolated and analysed.

The construction of biological display libraries requires the introduction into a microorganism of the genetic information necessary for the peptide synthesis. For the construction of a random peptide display library it is necessary to synthesize pools of DNA fragments that are then inserted into specific vectors. The DNA fragments are chemically synthesized as a mixture of single-stranded degenerated oligonucleotides containing constant regions and one or more degenerated stretches of DNA. DNA consists of sequences of 4 different nucleotides and each trinucleotide codes for a corresponding amino acid. Because of the codon degeneracy, most of the amino acids are coded by more than one triplet. Since in fully degenerated oligonucleotides there is the possibility to introduce stop codons that will interrupt protein synthesis, the oligonucleotides are synthesized using different mixtures of nucleotides especially in the third position of each triplet [27]. The DNA fragments to be cloned must be in a double-stranded form, at least at the end of each fragment. This is normally done by annealing short oligonucleotides to a complementary constant region inserted during the synthesis and by enzymatically completing the complementary DNA strand. After compatible ends are prepared by restriction enzyme digestion, the fragments are ligated into an appropriate vector and then introduced into the microorganism.

The most common microorganism used for peptide display is the *E. coli* filamentous bacteriophage [27]. Bacteriophages are viruses that infect bacteria by injecting their single-stranded DNA genome into the bacterial cells. Once inside the cell, they start to replicate their DNA. By using the host protein synthesis machinery, their coat proteins are synthesized and the DNA packaged into phage particles across the bacterial membrane and secreted into the medium from which they are easily recovered by precipitation. Random peptides have also been displayed in *E. coli* bacterium [28]. But bacteria, as other viruses and yeast, have been most used for protein display rather than peptides.

In order to be accessible to the target molecule, the peptides must be exposed to the medium and anchored to the viral coat or bacterial external membrane. For this reason, the DNA fragments coding for the peptides are inserted within a coat or membrane protein gene. The insertion is usually at one end of the fusion protein, in a region that does not change the conformation of important domains and without disrupting the protein coding sequence. For filamentous phage display, the most commonly used proteins are pIII and pVIII. The minor coat protein pIII is present at 3-5 copies per virion and is responsible for the binding to the F pilus and infection of male bacteria. pVIII is the major coat protein present at about 2700 copies and aggregates around the phage DNA.

The entire phage genome is usually used as a vector, after specific modifications have been introduced. First, it is necessary to genetically engineer specific restriction sites at the point of insertion of the DNA fragments. Sometimes new genes or regulatory regions are introduced, or existing genes are mutated

[29]. When the entire phage genome is used as a vector, the library is constituted of viral particles displaying a number of peptide molecules equal to the number of fusion coat protein molecules. Alternatively, a defective phage vector, called phagemid [30], can be created. A phagemid is a vector that contains the genetic information for packaging into virions, but does not encode viral genes, which must be supplied by a helper phage for the production of viral particles. In this case, phages will display a mixture of peptide-coat protein fusion and the corresponding wild-type coat protein. This system was created mainly for the display of larger protein fragments and for fusion proteins that do tolerate very short peptides only. In the phagemid system, the library can be propagated as bacterial colonies, but, in the presence of a helper phage, it will be constituted of phage particles. Several phage and phagemid vectors have been engineered for the display of random peptides [31-38].

The ligand selection process is called **Biopanning**. The target molecule must be bound to a solid support, usually a microtiter plate or a small Petri dish. Less common alternative supports are magnetic particles, columns with solid matrices, cells, mammalian organs. In a typical experiment, the number of phages that are incubated with the target corresponds to about 100 to 1000 times the complexity of the library. After the unbound clones are washed away, the bound ones are eluted by different methods, like low pH, high concentration of free target and direct infection of bacteria cells. The eluted phages are grown, purified and submitted to a new cycle of selection. Usually 3 to 4 rounds of selection are sufficient, and the entire process can be completed in about a week. At the end, several clones are isolated and their DNA extracted and sequenced. The DNA portions coding for the peptides are translated into amino acids and the sequences compared. If a consensus sequence can be identified, the screening may have been successful. One or more peptides are chosen and chemically synthesized in order to verify their binding affinity, outside of the microorganism system.

If it is necessary to improve a peptide ligand affinity, it is possible to build a new library by introducing random or targeted mutations on the ligand DNA, cloning the resulting fragments and using this library for a new selection.

Many peptide ligands have been identified by using biological display libraries. Some examples are peptide ligands for enzymes [29-44], receptors [45-47], DNA [48-49], toxins [50-51], lectins [52-53], hormones [54], and even organs [55].

Compared to chemical libraries, biological display libraries have several advantages and disadvantages. Some of the major advantages are the possibility to use a library for many different selection processes (even 100s), the easy propagation of the library and of the selected clones. The possibility to build larger size libraries is another advantage together with simple selection and sequencing procedures. On the contrary, a disadvantage is the fusion of peptides to a microorganism protein, and, therefore, the binding site can be extended to the fusion protein or the fusion protein may influence the peptide conformation. Furthermore, the selection process may enrich also for those microorganisms that stick to the solid support used, and L-peptides only can be displayed.

References

1. Furka, A., Sebestyen, M., Asgedom, M., Dibo, G., (1991), *Int. J. Pept. Protein. Res.* 37, 487.
2. Houghten, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T., Cuervo, J.H. (1991). *Nature,* 354, 84.
3. Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmierski, W.M., Knapp, R.J. (1991). *Nature,* 354, 82
4. Fassina, G., Verdoliva, A., Odierna, M.R., Ruvo, M., and Cassani, G. (1996). *J. Mol. Recog.* 9, 564-569.
5. Geysen HM, Rodda SJ, Mason TJ, (1986). *Mol. Immunol.* 23 709-715.
6. Stankova, M., Wade, S., Lam, K. S., Lebl, M., (1994). *Pept. Res.* 7(6), 292-298.

7. Kramer A, Volkmer-Engert R, Malin R, Reineke U, Schneider-Mergener J, (1993). *Pept. Res.* 6 314-319.
8. Fodor SPA, Read JR, Pirrung MC, Stryer L, Lu AT, Solas D, (1991). *Science* 251 767-773.
9. Lam, K.S., Hruby, V.J., Lebl, M., Knapp, R.J., Kazmierski, W.M., Hersh, E.M., Salmon, S.E. (1993). *Bioorg. Med. Chem. Lett.* 3, 419-424.
10. Furka A, Sebestyen F, Asgedom M, Dibo G, (1988). *14th Int. Congr. Biochem.*, Prague, Czechoslovakia. 5, 47.
11. Han H, Wolfe MM, Brenner S, Janda KD, (1995). *Proc. Nat. Acad. Sci.*, 92, 6419-6423.
12. Eichler, J., Houghten, R.A., (1993). *Biochemistry*, 32, 11035-11041.
13. Pinilla C, Appel JR, Blondelle SE, Dooley CT, Doerner B, Eichler J, Ostresh J, Houghten RA, (1995). *Biopolymers (Peptide Science)*, 37 221-240.
14. Pinilla C, Appel JR, Blanc P, Houghten RA, (1992). *Biotechniques*, 13(3), 412-421.
15. Frische, K., Meldal, M., Werdelin, O., Mouritsen, S., Jensen, T., Galli-Stampino, L., Bock, K., J. (1996). *Pept. Sci.* 2(4), 212-222.
16. Gausepohl, H., Kraft, M., Boulin, C., Frank, R.W. *Peptides 1990*, ESCOM, Leiden (1991) 206-207.
17. Knapp, D.R., Oatis, J.E., Papac, D.I. (1993*)., Int. J. Pept. Protein. Res.* 42, 259.
18. Ruvo, M., Scardino, P., Cassani, G. and Fassina, G., (1994). *Peptide and Protein Letters*, 1, No.3, 187-192.
19. Kassarjian, A., Schellemberger, V., Turck, C.W., (1993). *Prot. Sci.* 2,123 (391-T).
20. Fassina, G., Bellitti, M.R., and Cassani, G. (1994). *Prot. Pept. Lett* 1,15.
21. Boutin, J. A., Hennig, P., Lambert, P.-H., Bertin, S., Petit, L., Mahieu, J.-P., Serkiz, B., Vollånd, J.-P., Fauchère, J.-L., (1996). *Anal. Biochem.* 234, 126-141.
22. Metzger, J. W., Kempter C., Wiesmueller, K.-H., Jung, G., (1994). *Anal. Biochem.*, 219, 261-277.
23. Winger, B. E., Campana, J. E., (1996). *Rapid Commun. Mass Spectrom.*, 10, 1811-1813.
24. Stevanovic, S., Jung, G., (1993). *Anal. Biochem.*, 212, 212-220.
25. Fitch, W. L., (1997). Analytical methods for the quality control of combinatorial libraries, *Annual Reports in Combinatorial Chemistry and Molecular Diversity*, 1, 59-68.
26. Smith, G.P. (1985) *Science* 228, 1315
27. Smith, G.P., Scott, J.K. (1993) *Methods Enzymol.* 217, 228.
28. Lu, Z., Murray, K.S., Van Cleave, V., laVallie, E.R., Stahl, M.L., McCoy, J.M. (1995) *Biotechnology* 13, 366.
29. Scott, J.K., Smith, G.P. (1990) *Science* 249, 386.
30. Smith, G.P. (1991) *Curr. Opin. in Biotecnol.* 2, 668.
31. Parmley, S.F., Smith, G.P. (1988) *Gene* 73, 305.
32. Cwirla, S.E., Peters, E.A., Barrett, R.W., Dower, W.J. (1990) *Proc. Natl. Acad. Sci.* USA 87, 6378.
33. McCafferty, J., Griffiths, A.D., Winter, G. Chiswell, D.J. (1990) *Nature* 348, 552.
34. Markland, W. Roberts, B.L., Saxena, M.J. Guterman, S.K., Ladner, R.C. (1991) *Gene* 109, 13.
35. Felici, F., Castagnoli L., Mustacchio, A., Jappelli, R., Cesareni, G. (1991) *J. Mol. Biol.* 222, 301.
36. Breitling, F., Dubel, S., Seehaus, T., Klewinghaus, I., Little (1991) *Gene* 104, 147.
37. Jespers, L.S., Messens, J.H.,De Keyser, A., Eeckhout, D., Van Den Brande, I., Gansemans, Y.G., Lauwereys, M.J., Vlasuk, G.P., Stanssens, P.E. (1995) *BioTechnology* 13, 378.
38. Smith, G.P., Schultz, D.A., and Ladbury, J.E. (1993) *Gene* 128, 37.
39. Bunnel, S.C., Henry, P.A., Kolluri, R., Kirchhausen, T., Rickies, RJ., Berg, L.J. (1996) *J. Biol. Chem.* 271, 25646.
40. Markland, W., Ley, A.C., and Ladner R.C. (1996) *Biochemistry* 35, 8058.
41. Markland, W., Ley, A.C., Lee, S.W., and Ladner, R.C. (1996) *Biochemistry* 35, 8045.
42. Zhao, S., Lee, E.Y. (1997) *J. Biol. Chem.* 272, 28368.
43. Mukhija, S., Erni, B. (1997*) Mol. Microbiol.* 25, 1159.
44. Li, M. (1997) *Nat. Biotechnol.* 15, 559.
45. Liu, G., Bryant, R.T., and Hilderman, R.H. (1996) *Biochemistry* 35, 197.
46. Wrighton, N.C., Farrel, F.X., Chang, R., Kashyap, A.X., Barbone, F.P., Mulcahy, L.S., Johson, D.L., Barret, R.W., Jolliffe, L.X., and Dower, W.J. (1996) *Science* 273, 458.
47. Cheng, X., Juliano, R.L. (1996) *Gene* 171, 1.

48. Krook, M., Mosbach, K, and Lindbladh, C. (1994) *Biochem. Biophys. Res. Com.* 204 (2) 849.
49. Sato, A., Ida, N., Fukuyama, M., Miwa, K., Kazami, J, Nakamura, H. (1996) *Biochemistry.* 35, 10441.
50. Balass, M., Heldman, Y, Cabilly, S., Givol, D., Katchalski-Katzir, E., Fuchs, S. (1993*)* *Proc. Natl. Acad. Sci.* USA 90, 10638.
51. Scott, J.K, Loganathan, D., Easley, R.B., Gong, X., Goldstein, I.J. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5398.
52. Oldemburg, K.R., Loganathan, D., Goldstein, I.J., Schultz, P.G., and Gallop, M.A. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5393.
53. Nord, K, Gunneriusson, E., Ringdahi, J., Stahl, S., Uhlen, M., and Nygren, P. (1997*)* *Nat. Biotechnol.* 15, 772.
54. Folgori A, Tafi R, Meola A, Felici F, Galfre G, Cortese R, Monaci P, Nicosia A (1994) *EMBO J.* 13, 2236.

Application of Combinatorial Technologies in Affinity Chromatography: Ligands for Antibodies Purification

Dr. Giorgio Fassina
TECNOGEN S.C.p.A., Parco Scientifico
81015 Piana di Monte Verna (CE), ITALY

One of the first products deriving from combinatorial technologies to reach the market has been developed in our laboratory. Under the trade name KAPTIV™, TECNOGEN s.c.p.a. produces and sells a product line designed for the purification of monoclonal and polyclonal antibodies from different sources, such as cell culture supernatants, ascitic fluid, or sera (www.tecnogen.it). Specific products have been developed for the purification of immunoglobulin M (KAPTIV-M™), immunoglobulin A or E (KAPTIV-AE™), and immunoglobulin G and Y (KAPTIV-GY™). These products are sold worldwide through a network of local distributors.

Monoclonal antibodies (MAb's) are becoming an important class of therapeutic agents useful for the treatment of a vast array of diseases. Many monoclonals are waiting for FDA approval, and they represent almost 30 % of biotechnology derived drugs under development. Production of MAb's by hybridoma technology or transgenic animals can be easily scaled up, but still immunoglobulins purification from crude feedstocks poses several problems. Main difficulties are due to the low antibody concentration in cell culture supernatants or milk of transgenic animals and the high amounts of contaminating proteins.

Purification by affinity chromatography of monoclonal antibodies of the G class for therapy is based on the use of protein A or protein G immobilized on appropriate supports [1], as a first step to capture and concentrate the immunoglobulin from diluted feedstocks. These two proteins bind to the constant portion of the immunoglobulins, and so can be used to purify the majority of antibodies. They are obtained from microorganisms or genetically modified bacteria, through complex and expensive procedures, requiring in addition time consuming analytical controls to check for the presence of contaminants such as viruses, pyrogens, or DNA fragments, which may affect the safety of the purified MAb for clinical purposes. Many monoclonal antibodies with potential therapeutic applications are immunoglobulins of the A class. For example, protection against oral challenge with the invasive pathogen *Salmonella typhimurium* has been observed in mice treated with a monoclonal IgA generated by mucosal immunization of BALB/cmice with attenuated strains of *S. typhimurium* and subsequent fusion of Peyer's patch lymphoblasts with myeloma cells [2]. Other studies showed that monoclonal immunoglobulin A specific for *Shigella flexneri* 5a polysaccharide can protect mice against the disease [3]. Protective capacity of IgA's was also found for cholera toxin-induced chloride secretory response, in a dose dependent manner [4]. IgA's have been used for intranasal passive immunization of mice, providing protection against respiratory syncytial virus infection [5].

Monoclonal IgA may be produced by hybridoma technology in serum free media, devoid of IgG and IgM. Purification of this class of immunoglobulins can be accomplished by classical chromatographic approaches with an acceptable degree of purity, but several steps, such as ammonium sulfate precipitation, ion-exchange chromatography, and gel filtration, are usually required [6,7]. The lectin jacalin, isolated from jackfruit seeds [8], binds to IgA and can be used conveniently for the affinity purification of IgA from colostrum or serum [9]. However, several aspects limit the use of this lectin for large-scale purification of monoclonal IgA from cell culture supernatants. First jacalin is a biologically active lectin, a potent T cell mitogen and a strong B cell polyclonal activator [10], thus requiring a careful control for ligand leakage in purified preparation. Second, jacalin binds to the carbohydrate moiety of IgA, and D-galactose is required to elute IgA from affinity columns; the procedure is costly and unpractical for large-scale operations.

Monoclonal antibodies of the E class have been also developed for therapeutic applications [11], and large amounts of highly pure material are required to extend the clinical investigations. In these cases, IgE's are produced by hybridoma technology in media with low IgG and IgM content. Purification of monoclonal and polyclonal IgE requires the use of conventional, often complex and laborious, fractionation procedures, since specific and cost effective ligands for this class of immunoglobulin usable for affinity chromatography applications are not yet available.

Antibodies of the E class are purified mainly by immunoaffinity chromatography using anti-IgE antibodies immobilized on solid supports [12,13]. Even if selective enough for research application, scaling up immunochromatography for preparative applications is very expensive and not easily accomplished. Other approaches for IgE purification include classical chromatographic protocols based on the combination of different sequential procedures such as salting out, affinity chromatography on lysine-Sepharose, ion-exchange, gel filtration and immunoaffinity chromatography to remove interfering proteins [14]. Studies with immobilized protein A show that this protein, known to recognize the immunoglobulins Fc region, does not bind to monoclonal IgE, but binds only 12-14 % of serum polyclonal IgE. Protein G binds to neither polyclonal nor monoclonal IgE.

Even if the large majority of MAb's currently under development for clinical treatments are immunoglobulins of the G class, IgMs are also finding application for the diagnosis and cure of certain important diseases such as cancer [15]. Conventional approaches for the purification of IgMs have been based on the combination of different fractionation techniques, such as precipitation, gel permeation chromatography and ion exchange chromatography or electrophoresis. Protein A, protein G, or protein A/G , widely used for the affinity purification of antibodies from sera or cell culture supernatants even at the industrial scale, do not recognize well immunoglobulins of the M class and are not used to capture and purify IgM from crude sources. Alternative ligands for the affinity purification of IgM include the mannan binding protein (MBP), which after immobilization on solid supports provides affinity media useful for IgM isolation based on a temperature dependent interaction of the ligand with the immunoglobulin [16]. The use of immobilized MBP for the purification of IgM is based on the adsorption in the presence of calcium at a temperature of 4°C, and the room temperature dependent elution of adsorbed immunoglobulins in the presence of ethylenediaminotetraacetic acid (EDTA). This ligand shows low binding affinity for IgG, but binds to bovine and human IgM with reduced affinity than murine IgM. However, in addition to the complexity of MBP isolation, functional binding capacities of MBP-columns are limited to 1 or 2 mg of IgM per ml of support. All these procedures are time consuming, labor intensive, may alter IgM functionality and are not compatible with industrial scaling up.

Given the importance of the application of MAb's for therapy, and given the role of the purification process in assuring the quality, consistency and safety of the products, it is clear that the availability of synthetic ligands able to recognize antibodies independently from their class is of remarkable industrial importance, since it may lead to less expensive production costs and reduced risks of contamination.

In order to circumvent these problems, in our laboratory the search for a synthetic ligand able to bind to immunoglobulins, and usable for affinity chromatography applications, has been undertaken through the synthesis and screening of combinatorial peptide libraries [17]. **Peptide libraries** are collections of peptide molecules encompassing all the possible sequences for a given peptide length, which are synthesized and screened for activity simultaneously [18,19]. The great number of molecular combinations attainable makes the identification of useful leads very convenient and time-effective. Previous studies have shown that peptide ligand multimerization enhances retention of recognition properties after immobilization on solid support for the preparation of affinity columns [20-22]. Based on these considerations, a tetrameric peptide library has been designed where four identical peptide chains are assembled starting from a tetradentate lysine core, similar to that used for the production of multimeric antigenic peptide [23].

The multimeric library, composed of randomized tripeptide tetramers, has been produced by solid phase peptide synthesis following a manual procedure developed by our group [24]. Screening of the activity of the multimeric library in terms of antibody recognition has been carried out using a very simple assay, measuring the library ability to interfere with the interaction between Protein A and biotinylated immunoglobulins, monitored on the solid phase using an ELISA format. The first screening cycle identified the sub-library with arginine at the N terminus as the most active sub-library (60 % inhibition). Subsequently a second multimeric sub-library was synthesized, with arginine at the N-terminus, the amino acid in position 2 defined, and the third unknown. This second sub-library was screened in the same assay at the same concentrations, and the best inhibitory activity was found in the sub-library with threonine in the second position. A third sub-library, composed of 18 tripeptide tetramers of the general formula $(Arg\text{-}Thr\text{-}X)_4\text{-}K_2\text{-}K\text{-}G$ was then prepared and tested in binding inhibition experiments. This allowed the final identification of the most active multimer as $(Arg\text{-}Thr\text{-}Tyr)_4\text{-}K_2\text{-}K\text{-}G$ [TG19318]. The ligand was then prepared in larger amount by automated solid phase peptide synthesis, purified by semipreparative RP-HPLC, and characterized in terms of amino acid composition and by determination of molecular weight by laser desorption mass spectrometry which confirmed the predicted chemical nature.

Ligand immobilization on solid supports indicated recognition selectivity for antibodies much broader than protein A, since immunoglobulins of different classes and subtypes from different sources could be conveniently purified from crude samples in a single step [25]. Affinity columns with immobilized TG 19318 were not affected by the presence of denaturants, detergents, or other sanitation reagents commonly used for pyrogen removal, proved stable to repeated use, and adsorption and desorption of immunoglobulins was accomplished using the same eluents used for protein A columns. The tetrameric ligand TG19318 could be easily immobilized on preactivated solid supports, since the presence of the symmetric central core and the four peptide chains departing from it lead to an oriented immobilization where the resin bound chains act as self-built spacer to optimize interaction. All the different supports tested so far maintained the ligand recognition properties for immunoglobulins, even if with different functional capacities.

Ligand denaturation did not constitute a problem, such as in the case of protein A, and TG19318 columns could withstand a large array of harsh sanitizing agents with no capacity losses. In addition, the low toxicity of TG19318 and the low molecular weight of the resulting fragments reduced considerably the problems of contamination by leaked ligand, as is the case for protein A. Preliminary experiments suggested that the ligand is more stable to proteolytic digestion when coupled to solid supports, and the enzymatic activity normally found in crude feedstock's derived from cell culture supernatants did not lead to noticeable losses of capacity. Adsorption of antibodies on TG19318 affinity columns occurred with neutral buffers at low ionic strength, conditions fully compatible with the use of crude feedstock's deriving from cell culture supernatants. Elution of adsorbed immunoglobulins could be achieved simply by changing the buffer pH to acid or alkaline conditions, with acetic acid (pH 3) or sodium bicarbonate (pH 9.0). Increasing the ionic strength of the dissociation buffer favored a more efficient elution of adsorbed antibodies.

Interaction affinity was strong enough to allow purification of antibodies directly from diluted supernatants where the immunoglobulin concentration is very low, from 10 to 50 μg/ml. The main contaminant, albumin, was always efficiently removed in the purification step with any type of support tested for TG19318 immobilization. The ligand was useful not only for IgG purification from different sources, but also for IgY, IgM [26], IgA [27], and IgE [28] isolation from sera or crude cell supernatants.

Among the classical techniques for protein purification, affinity chromatography deserves particular attention in down stream processing of antibodies since it allows several-fold purification in a single step, usually with high recovery, and is compatible with the use of large amounts of very dilute preparations, as is often the case with monoclonal antibodies produced by hybridoma technology.

Despite its potential, the use of affinity chromatography for purification of biotechnological products is limited by the availability of specific or class-specific ligands for each desired target.

The number of polypeptides produced by recombinant DNA or hydridoma technology with potential therapeutic or diagnostic applications is growing at an impressive rate, and consequently the demand of effective and compatible purification strategies is becoming more urgent. Synthesis and screening of peptide libraries might constitute an interesting approach to identify novel ligands tailored to specific biotechnological needs, and multimeric libraries are particularly suited for the identification of novel ligands for affinity chromatography applications. Compared to conventional linear peptide libraries, multimeric libraries offer the additional advantage of a more extended chemical and structural diversity, since the peptide chains linked to the central polylysine core are located in an asymmetric chemical environment, and display an increased molecular surface which may lead to more favourable contacts with the target biomolecules. In addition, multimeric peptides are already engineered for a more appropriate immobilization on solid support with retention of recognition properties, since after immobilization only a limited number of peptide chains are linked to the solid phase, leaving the others fully available for interaction.

References

1] Fuglistaller, P. (1989) Comparison of immunoglobulin binding capacities and ligand leakage using eight different protein A affinity chromatography matrices. *J. Immunol. Meth.* 124,171-177.

2] Michetti, P., Mahan, M.J., Slauch, J.M., Mekelanos, J.J., and Neutra, M.R. (1992) Monoclonal secretory immunoglobulin A protects mice against oral challenge with the invasive pathogen *Salmonella typhimurium. Infect. Immun* 60,1786-1792.

3] Phalipon, A., Kaufmann, M., Michetti, P., Cavaillon, J.M., Huerre, M., Sansonetti, P., and Kraehenbuhl, J.P. (1995) Monoclonal immunoglobulin A antibody directed against serotype-specific epitope of *Shigella flexneri* lipopolysaccharide protects against murine experimental shigellosis. *J. Exp. Med.* 182:769-778.

4] Apter, F.M., Lencer, W.I., Finkelstein, R.A., Mekalanos, J.J. and Neutra, M.R. (1993) Monoclonal immunoglobulin A antibodies directed against cholera toxin prevent the toxin-induced chloride secretory response and block toxin binding to intestinal epithelial cells in vitro. *Infect Immun* 61:5271-5278.

5] Weltzin, R., Hsu, S.A., Mittler, E.S., Georgakopoulos, K., and Monath, T.P. (1994) Intranasal monoclonal immunoglobulin A against respiratory syncytial virus protects against upper and lower respiratory tract infections in mice. *Antimicrob. Agents Chemother.* 38:2785-2791.

6] Waldam, R.H., Mach, J.P., Stella, M.M., and Rowe, D.S. (1970) Secretory IgA in human serum. *J. Immunol* 105:43-47.

7] Khayam-Bashi, H., Blanken, R.M., and Schwartz, C.L. (1977) Chromatographic separation and purification of secretory IgA from human milk. *Prep. Biochem* 7:225-229.

8] Roque-Barreira, M.R. and Campos-Nieto, A. (1985) Jacalin: An IgA-binding lectin *J. Biol. Chem.* 134:1740-1743.

9] Kondoh, H., Kobayashi, K., and Hagiwara, K. (1987) A simple procedure for the isolation of human secretory IgA of IgA1 and IgA2 subclass by a jackfruit lectin, jacalin, affinity chromatography. *Molec. Immunol.* 24:1219-1222..

10] Bunn-Moreno, M.M., and Campos-Neto, A. (1981) Lectin(s) extracted from seeds of artocarpus integrifolia (jackfruit): potent and selective stimulator(s) of distinct human T and B cell functions. *J. Immunol* 127:427-430.

11] Nagy, E., Berczi, I., and Sehon, A.H. (1991) Growth inhibition of murine mammary carcinoma by monoclonal IgE antibodies specific for the mammary tumor virus. *Cancer Immunol. Immunother.* 34:63-69.

12] Phillips, T.M., More, N.S., Queen, W.D., and Thompson, A.M. (1985). Isolation and quantitation of serum IgE levels by high-performance immunoaffinity chromatography. *J. Chromatogr.* 327: 205-211.

13] Lehrer, S.B. (1979). Isolation of IgE from normal mouse serum. *Immunology* 36:103-109.

14] Ikeyama, S., Nakagawa, S., Arakawa, M., Sugino, H., and Kakinuma, A. (1986). Purification and characterization of IgE produced by human myeloma cell line, U266. *Mol. Immunol.* 23:159-167.

15] S.M. Quadri, A.B. Malik, X.Z. Tang, R. Patenia, R.S. Freedman, and H.M. Vriesendorp (1995) Preclinical analysis of intraperitoneal administration of 111In-labeled human tumor reactive monoclonal IgM AC63-2B12. *Cancer Res.* 55:5736-5742.

16] Nevens J.R., Mallia A. K., Wendt M. W., and Smith P. K., (1992) Affinity chromatographic purification of immunoglobulin M antibodies utilizing immobilized mannan binding protein. *J. Chromatogr.* 597: 247-253.

17] Fassina, G., Verdoliva, A., Odierna, M.R., Ruvo, M., and Cassani, G. (1996) Protein A mimetic peptide ligand for affinity purification of antibodies. *J. Mol. Recogn.* 9: 564-569.

18] Houghten, R.A., Pinilla, C., Blondelle, S.E., Appel, J.R., Dooley, C.T., and Cuervo, J.H. (1991) Generation and use of synthetic peptide combinatorial libraries for basic research and discovery. *Nature* 35484-86.

19] Lam, K.S., Salmon, S.E., Hersh, E.M., Hruby, V.J., Kazmiersky, W.M., and Knapp, R.J. (1991) A new type of synthetic peptide library for identifying ligand-binding activity. *Nature* 354:82-84.

20] Fassina, G. (1992). Oriented immobilization of peptide ligands on solid supports *J. Chromatog.* 591:99-106.

21] Fournier, A., Couvineau, A., and Laburthe, M. (1993) Synthesis of a hydrophilic affinity matrix for the purification of the vasoactive intestinal peptide receptor. *Anal Biochem.* 211:305-310.

22] Butz, S., Rawer, S., Rapp, W., Birsner, U. (1994) Immunization and affinity purification of antibodies using resin-immobilized lysine-branched synthetic peptides. *Pept. Res.* 7:20-23.

23] Tam, J.P. (1988). Synthetic peptide vaccine design: Synthesis and properties of a high-density multiple antigenic peptide system. *Proc. Natl. Acad. Sci. USA* 85:5409-5413

24] Ruvo, M., Scardino, P., Cassani, G., and Fassina, G. (1994) Facile manual synthesis of peptide libraries. *Prot. & Pept. Lett.* 1:187-192.

25] Fassina, G.,Verdoliva, A., Palombo, G., Ruvo, M., and Cassani, G., (1998) Immunoglobulin specificity of TG 19318: A novel synthetic ligand for antibody affinity purification. *J. Mol. Recogn.* (1998) 11: 128-133.

26] Palombo, G., Verdoliva, A., and Fassina, G. (1998) Affinity purification of IgM using a novel synthetic ligand. *J. Chromatogr. Biom. Appl.* (1998) 715:137-145.

27] Palombo, G., De Falco, S., Tortora, M., Cassani, G., and Fassina, G. (1998) A synthetic ligand for IgA affinity purification . *J. Molec. Recogn.* (1998) 11:243-246.

28] Palombo, G., Rossi, M., Cassani, G., and Fassina,G. (1998) Affinity purification of mouse monoclonal IgE using a protein A mimetic ligand [TG 19318] immobilized on solid supports. *J. Molec. Recogn.*, (1998) 11:247-249.

Bioassay Design, High Throughput Compound Screening and the Impact of Combinatorial Technology

Dr. Frank T. M. van Amsterdam
GlaxoWellcome Medicines Research Centre,
Via A. Fleming 4, Verona, ITALY

Introduction

In recent years, the use of large compound stores in the pharmaceutical industry and the synthesis of ever-larger compound libraries have had a profound effect on the way these compounds are being assayed. The development of new technologies in chemistry, such as parallel and combinatorial design and synthesis of compound libraries, have led to an enormous growth in the availability of chemical entities for screening biological targets. Moreover, the developments in molecular biology and the advent of the knowledge of the complete human genome will make (and have already made it) possible to use completely new technologies for screening a number of targets that will be much higher than ever before. In particular, the development of versatile, more or less large, fully automated (robotic) screening stations that are able to take over the routine work from highly skilled laboratory scientists and technicians make it now possible to routinely screen tens of thousands of assay points per day.

Today the drug discovery process is typically subdivided in a number of phases, comprising target identification, target validation, assay development, high throughput screening (hit identification), hit optimisation and lead optimisation. Most of these phases are directly or indirectly related to the areas of bioassay design and high throughput screening, the main areas presented in this contribution.

The present paper concentrates on the early phases of drug discovery, starting with a description of the issues in target identification and validation and continuing via the assay design phase to a description of high throughput screening technologies. Further detail of parallel and combinatorial chemistry in the design of libraries to be used in the process of hit and lead optimisation is amply described in other accompanying papers.

Bioassay Design

Bioassay Design is preceded by two other essential phases in the drug discovery process, the identification of a target and its validation. Only when a valid target has been identified we start to discuss the design and testing of the most appropriate assay format and a detailed test protocol.

Target Identification

Not too long ago most targets were found by looking at the action of hormones and other natural ligands that could be identified in or outside the human body and were easily linked to known physiological processes. Today new targets are often linked to elusive components of metabolic pathways or to diseases of which the multi-factorial (patho)-physiology is not yet known at all, or only at the rim of being understood.

The unravelling of the human genome, which is continuing at ever a faster pace, will lead to the identification of all human genes and implicitly will allow the search for new drugs starting from a complete blue-print of the human body. However, many diseases are not linked to single genic effects and therefore a simple relation between a gene and a disease most often does not exist.

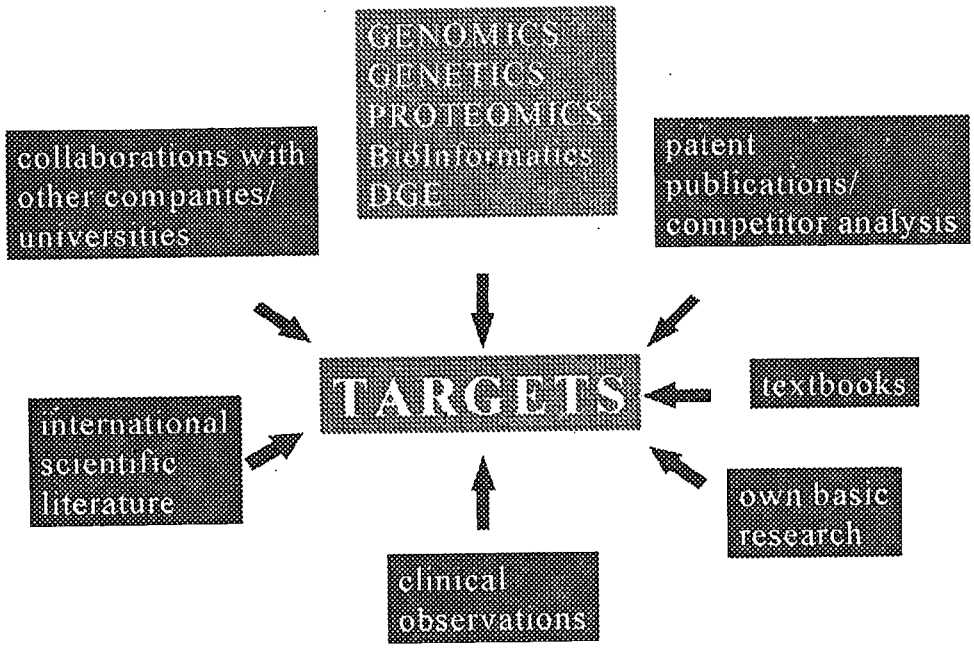


Fig 1. *Sources for new targets to be exploited for Drug Discovery. The role of bioinformatics as a tool of using genetic and genomic information is becoming more and more important.*

So-called susceptibility genes are those genes which contribute to the risk and onset of the disease and have been identified for many rare, inherited diseases seen in families. A particularly well know case is the gene coding for the ApoE enzyme, which has been shown to be directly involved in the development of Alzheimer disease. Studying human diseases therefore means finding out which (poly)-genic effects are involved in a particular disease. At that point we may start looking at a way of understanding the function of the gene, discovering which metabolic pathways are involved and converting one of the steps into a reliable bioassay.

The role of bio-informatics as a support science is particularly important as it permits the development of tools and databases to store, organise, manipulate and analyse genetic data. Linking this information electronically to the scientific literature provides ways for obtaining information that could have been impossible to reach only few years ago (Beeley and Duckworth, 1996). In this process, the use of the internet as a means of connecting widely varying information sources has become one of the most important bio-informatics tools. As we deal with enormous amounts of information that is dispersed over a huge number of sources, it would be unthinkable to obtain the necessary basic data from a single source.

While current drug targets represent approximately 400 human gene products like enzymes, receptors and ion channels, it is estimated that there may be 2,500 to 5,000 molecular targets that can somehow be related to the recovery of lost functions related to human polygenic diseases. The central question remains how to identify these gene products.

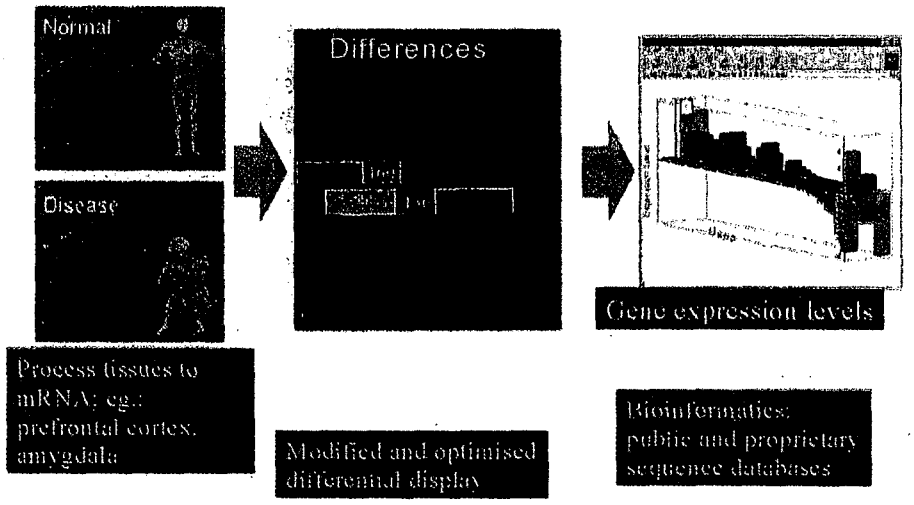


Fig. 2. Systematic identification of disease-related genes, which represent potential drug targets or surrogate markers/diagnostics

A widely used technique of identifying which genes are involved in a disease is the so-called differential gene expression (DGE). In short, healthy and sick patients are studied, isolating samples of the tissues most probably involved in the disease and directly comparing the expression level of all genes in these tissues. It is supposed that the difference between the two expression patterns somehow will be related to the disease. This assumption becomes more close to reality when the basic genome of the people involved is closer and therefore it is preferable to study complete families or closely contained populations (like the Iceland people). DGE is used in many variations, going from Serial Analysis Gene Expression (SAGE; Genzyme), via the Affymetrix Gene Chips, Sequencing by hybridisation (Hyseq inc.) and oligo fingerprinting (DGE-barcode; Genome Pharmaceuticals Corp) to Differential Display (Curagen, Digital Gene Tech., GeneLogic).

Of the about 100,000 human genes only about 10,000 are expressed in a given cell. As any disease phenotype will show differences in its expressed genes with respect to the healthy counterpart, this means that the differentially expressed genes and their gene products represent potential drug targets. Bio-informatics is the basic tool to relate the information of the differentially expressed genes with information available to everyone and with proprietary sequence databases. As differences and not absolute levels are studied, the technique is very sensitive and also applicable to the identification of rare genes.

Essentially the same approach can be used in proteomics. Comparing the protein content of diseased against control tissue by 2D-electrophoresis, the peptide sequence of differentially expressed proteins can be obtained after mass spectrometric analysis. From here the DNA sequence can be obtained which gives essentially the same starting point for the bio-informaticist as the gene-based DGE technology described above. Proteomics has particular usefulness in the study of protein-protein interactions such as with multi-protein complexes, signalling complexes and the creation of metabolic pathways or spatial reference maps of cellular proteins.

Functional Analysis / Target Validation

Once a gene or a series of genes has been identified as related to a particular disease it is imperative to understand what the function of the gene(s) is, in which pathways the gene products are involved and whether or not the disease association is the cause of the disease. Should the gene product be a

consequence of the disease rather than the cause, this may exclude a way of finding a marketable drug. Nevertheless, the information may still be used to develop a diagnostic kit or lead to the development of a surrogate marker i.e. a marker that can be easily measured and indicates the predisposition to, the presence, or the severity of a particular disease. These markers are of particular use in early clinical trials where they can make a massive difference in the cost of the trial.

Functional analysis as a discipline of research can make use of a series of supportive technologies ranging from antisense technology, knock-out experiments in transgenic animals (usually in mice) or by introducing point mutations in defined genes. Also Yeast Two Hybrid assays (pathway elucidation) and appropriate use of proteomics (protein inhibition control) can be helpful in elucidating the functions of a gene.

The unravelling of the function of a gene or a gene product and the verification of its involvement in the generation of a certain patho-physiological condition constitutes de facto the target validation and justifies the development of a screening assay for finding developable lead compounds.

In selecting the best targets for developing an assay for high throughput screening, several criteria must be taken into consideration: i) the target must be validated to greatest extent for disease relevance; ii) affecting the target should cure the disease or at least alleviate the symptoms; iii) the target should preferably be present only in the cells or tissues involved in the disease process; iv) the target must be tractable i.e. it must be possible to develop an assay while a therapeutic intervention must also be possible. Two more objectives must be met: the target must be in line with the institution objectives/portfolio and Intellectual Property rights must be sufficiently secure to allow development of a medicine.

Assay Design

Targets for bioassay design are usually proteinergic gene products like membrane receptors, ion channels, enzymes, structural proteins, transduction intermediates or nuclear receptors. Protein-protein interactions can be measured to study specific steps in signal transduction pathways. These targets can often be isolated and assayed independently. Alternatively, whole cells can be used for viability or toxicity studies, or for studying mechanisms that need the presence of other cellular machinery for activity. Transfected cells (cells in which foreign genes from one organism have been introduced into the genome of another organism) allow the use of specific cellular machinery for the construction of functional assays even when this normally, in a non-modified cell, would not be possible. Often these assays are developed in the so-called reporter gene format (*vide infra*).

In order to allow a high sample throughput it is necessary to generate extremely simple assays that can be executed in microtiter plates or in other high-density formats. Filtration, centrifugation or other complex multiplate steps must be avoided as much as possible to increase throughput. It is further necessary to decrease the volume per sample, not only to reduce the cost of reagents and the efforts to manipulate and store solutions, but also to increase speed and to allow the synthesis of compounds to be done on a smaller scale.

Assay Technologies

Spectroscopic detection techniques:

Colourimetric detection of enzymatic reactions is probably one of the oldest and simplest techniques used for screening. This technique is usually based on the direct detection of specific metabolic species (NAD/NADH) or on complex formation of specific reaction products forming a coloured species (Pi /

vanadium heptamolybdate). This detection technique is very widely used as it is usually cheap and easy, with readers that usually can read microtiter-plates in a matter of seconds.

Chemo-luminescence detection is the second easiest way of measuring a series of reaction products such as the very popular measurement of firefly luciferase activity. Often, as is the case of the luciferase reaction, chemiluminescence is a very fast phenomenon (flash luminescence) which requires direct injection of reagents in the wells of the microtiter plates immediately before reading the luminescence yield. Recent developments in this field have produced signal stabilisation kits (e.g. Packards LucLite) that allow measurement of the signal also up to an hour after the addition, thus allowing to read the plates as if it were a glow luminescence assay.

Fluorescence detection has become a popular way of measuring metabolic reactions and molecular interactions, especially now that many specific reagents have become commercially available. These reagents are very specific and may allow signal multiplexing (performing two or more assays to be carried out in the same assay well) and even the substitution of radioisotopes. However, their use has not become as widespread as should be the case as direct fluorescence is often hindered by cellular/aspecific background (auto) fluorescence, while the technique is also dependent on many secondary reaction conditions.

Time Resolved Fluorescence is a more recent technology, introduced by Wallac and Packard. This technology uses the fact that certain fluorescence phenomena last longer (millisecond range) than the classical fluorescence peaks, allowing measurement of the signal in a specific time window that starts after the main signal of background fluorescence has already extinguished to zero values. This gives an important advantage, especially in functional, cellular assays.

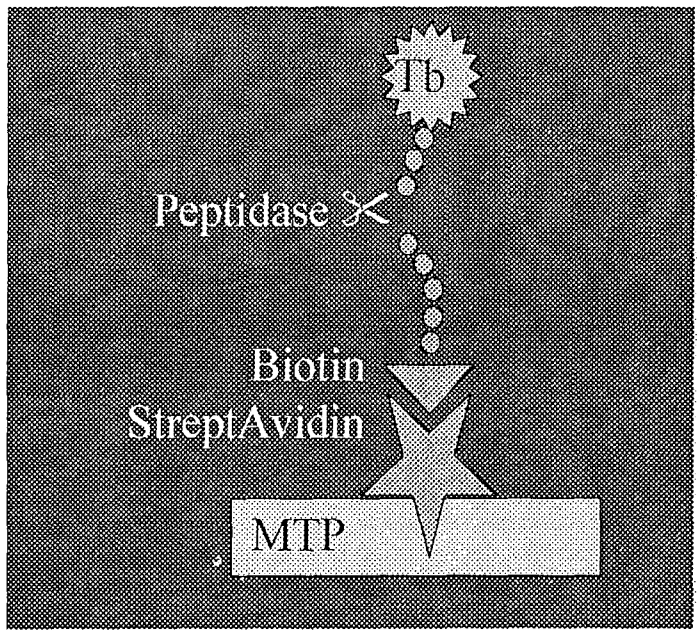


Fig 3. *Example of a time-resolved fluorescence assay using Terbium as the fluorescence probe. Washing the plates after the peptidase activity measures the amount of Terbium still bound to the plates.*

The Wallac technology uses the fluorescence characteristics of Lanthanides like Terbium and Europium that can be incorporated in many different ways in macromolecular structures. Using biotinylated macromolecules that can be linked to streptavidin-coated plates or by plating cells on poly-D-Lys-coated plates, it is possible to study interactions and molecular reactions by washing away that

part of a lanthanide containing label not linked to the plate at the end of the reaction. For a review see Hemmila and Web (1997).

Homogeneous TRF is based on **Fluorescence Resonance Energy Transer (FRET) technology.** Packard developed this technique as a specific product, the Discovery analyzer. In this case a proprietory chelate of Europium (Eu-cryptate) acts as fluorescence donor while a second reactant (XL665) serves as acceptor and secondary emitter. The secondary event is much slower than the cryptate, but has a very high energy yield of 50% at 9 nm distance. This allows a different kind of proximity assay entirely based on fluorescence signals and the possibility to perform the assay in homogeneous conditions (all reagents in solution). In addition this format allows kinetic measurements, while colour quench can be compensated by dual wavelength measurements.

Fluorescence Polarisation has become popular recently as it allows receptor or enzyme binding / activation experiments to be performed without radioisotopes. The technology is based on measurement of the increase of polarisation of emissions from fluorescent labels irradiated with polarised light when a small ligand is bound to a macromolecule. High polarisation is related to low rotational movements of molecules during their excited-state lifetime. When the molecule is unbound, it can freely rotate and emitted light is much less polarised. The main advantage of this technology is that there is no separation required, all reactions take place in solution phase and data can be collected at equilibrium, and there is a relative insensitivity to compound colour quench. FP certainly is a technology that will further develop into the HTS area.

Radioisotope measurements:

A basic assay format to measure affinities of chemical entities for membrane receptors or other cellular proteins is the **radioligand binding format**. The use of a radioactive tracer (usually 3H or 125I) allows the measurement of the binding of minute quantities of a specific ligand to its receptor and the subsequent displacement of that ligand through the interaction with the chemical entities to be tested. Major disadvantages are the use of radionuclides and the fact that bound and free radioligand need to be separated before the bound ligand can be measured after dissolving the sample in a liquid organic scintillant. The technique has been successfully used for many years, especially during the '80s.

New developments in this area have come from Amersham, who have developed the **Scintillation proximity assays (SPA),** a technique already used in the late seventies by Hart and Greenwald (1979). In this format small beads are used that contain a quantity of scintillant. The receptor or the enzyme is bound to the beads with an appropriate binding medium (for membrane receptors this may be wheat germ agglutinin) which gets the receptors in close proximity to the scintillant in the beads. In this case, when the radioligand binds to the receptor, it is in proximity of the scintillant and will induce the emission of light, while those radioligands that are not bound to the receptor will be too far from the scintillant to induce the same effect. The result is that there is no need to separate bound from free radioligand and the assay is usually simplified to a simple add, mix, incubate and measure protocol.

The latest developments in this field are present in the **Leadseeker program** from Amersham, where smaller beads are used with different spectral characteristics. There are two major advantages: firstly, it is possible to use the telecentric lens of the Leadseeker system to record an image of the complete microtiterplate, instead of collecting the scintillation events of a limited number of wells at a time. Secondly, the light yield of these beads are so much higher that the system can be used for 384, 864 and even 1536 well plates, which can not be used anymore with the previous generation of SPA beads. An additional advantage is the reduced colour quench by chemical entities which is very rare in the spectral range of these beads.

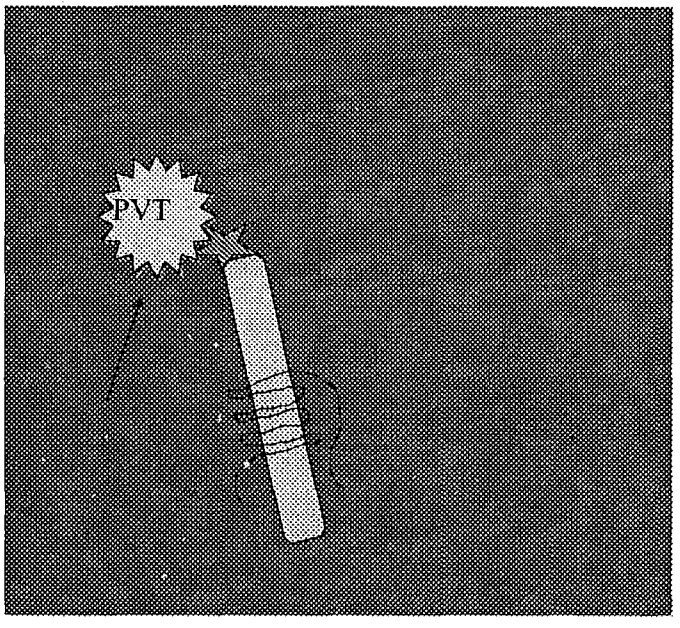


Fig 4. *Example of a Scintillation Proximity Assay. The SPA principle renders washing or filtering of the plate no longer necessary, which makes the assay method one of the simplest imaginable.*

Altogether the newest technologies allow faster and easier screening of high numbers of chemical entities, while again reducing the cost per well which was relatively high for the first generation SPA beads.

Other Assay Formats

Reporter Gene Technology

A very important technology used more and more with the general spreading of molecular biology tools is the so called reporter gene assay (Alam and Cook, 1990; Suto and Ignar, 1997). In general, a reporter gene construct consists of an inducible transcriptional control element driving the expression of a reporter gene. This technology allows the measurement of quite complex metabolic processes with simple detection technologies, without the need of complex biochemical assay procedures. Often the reporter gene is introduced by transfection using a now wide series of different vectors to work in a number of appropriate reciepeint cells. Second messengers or cellular metabolites serve as specific activators for DNA elements that are linked to a promoter which activates the reporter gene and consequently the production of a reporter molecule. The most widely used reporter systems include chloramphenicol acetyltransferase (CAT), firefly luciferase, renilla luciferase, secreted alkaline phosphatase (SEAP), beta-galactosidase and Green Fluorescent Protein (GFP).

A specific example is the measurement of second messengers like cAMP (George et al, 1997), where until recently a two-phase protocol had to be used. This consisted of a first phase, which represented the actual experiment, and a second phase which was usually an antibody detection of the cAMP generated during the experiment. With the advent of 6CRE-Luc constructs transfected in the cell, the detection of cAMP can be measured very simply by measuring a luciferase signal that develops as a consequence of the intrracellular cAMP production and the interaction with the cAMP responsive element. The luminescence signal can be measured as a glow reaction on a scintillation counter using the LucLight system from Packard.

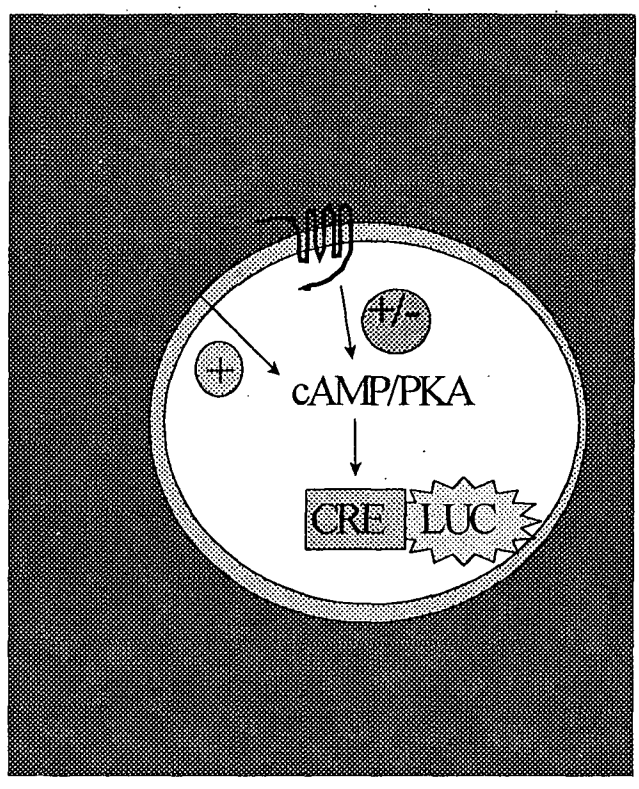


Fig 5. *Reporter Gene Assays: measurement of cAMP accumulation through the activation of a cAMP Responsive Element and the consequent production of Luciferase.*

Lawn Assay Format

Microtiter plates are becoming available with ever-lower volumes. Now the 1536 well format has become the highest density standard, although even higher density formats are being experimented widely.

An alternative way of approaching the problem of volumes and density of sample distribution is the free format. Working on a sheet of material or gel without physical boundaries between the areas available for each sample allows screening of very high sample densities. This is particularly interesting in the case where the samples can be identified without knowing their position, as is the case when encoded bead-based chemical libraries are screened.

Using encoded libraries, one may distribute the beads on the gel, cleave (some of) the compound by acid or photolytic treatment so that they can diffuse into the gel and react with the reagents. Positive reactions will give a a coloured or fluorescent halo which can be easily detected. The positive bead is taken from the gel and the tag analysed and identified with available technology.

Also from the point of view of the assay this format may be favourable as it is very well possible to grow a monolayer of cells or inoculate a germ on the gel without the need of doing that in each individual well.

High Throughput Screening Technology

High Throughput Screening (HTS) is a logical development of recent advances in chemistry, biology and technology.

The advent of parallel and combinatorial synthetic techniques has opened the way to fast synthesis of potentially very large compound libraries. Now that (tens of) thousands of compounds can be prepared in relatively short time intervals, classical ways of manual screening are no longer sufficient. Also the presence of large compound collection in the pharmaceutical industry and the clear intent to use those collections to screen as many targets as desirable has increased the availability of large compound numbers for screening.

As indicated before, the human genome initiative will soon make a much larger amount of targets available for screening than ever has been possible. Screens to find tools for these targets or to help select the better ones need a high screening capacity. The same is true for the larger number of selected targets linked to various diseases that have become new ways of finding high quality leads. In addition, recent developments in cellular and molecular biology have made a whole range of new assay technologies available that allow screening of many targets in an easy way with high throughput, where some years ago no screening would have been possible.

A third component and not the least in importance, is the development and availability of modern technological and data management tools that allow to deal with high amounts of samples and enormous amounts of raw and elaborated screening data. Although this is often not realised, the amount of resource that can be gained by using automated data analysis is probably even higher than the amount of resources gained by automated assay execution. Technological progress will also allow further assay miniaturisation and therefore cost containment, both in screening consumables and in the use of the main asset, the compounds.

From these consideration it may be derived that HTS is a combination of a number of disciplines that lead to a set of conditions that allow screening of tens of thousands of samples per week or even per day (Babiak, 1997). The necessary conditions comprise the availability of a robust, miniaturised assay, a flexible (robotic) station that can deal with the multi-well plates, a powerful IS infrastructure that can deal with the handling of all the compounds as well as analytical and biological data and a way of scheduling all activities. Besides that there is the need of appropriate support for storing and managing all consumables and for preparing and maintaining all biological materials necessary for the screens.

The Screening Process

A typical screening process comprises the following distinct phases:

- Selection of the compound sets/libraries to be screened. Essentially there are two different approaches, the 'diversity' approach or the 'rational' approach. The diversity approach aims at screening as high as possible a chemical diversity, including central file compounds, acquired compounds, combinatorial libraries or natural products. In contrast, the rational approach starts from already available knowledge on the target and tries to model as good as possible the required chemical characteristics of the compounds to test to this pharmacophore model.

- Collecting the compounds/libraries selected. Usually a request to the central compound store will have to be made, checking for availability and the compounds must be plated or copied for the screen requested. Then they must be transported to the screening site.

- Set up of the data structure. When the compounds or plates get to the screening site they need to be plated in the assay plates and bar-coded. The accompanying plate map needs to be stored in the database against the new bar code

- Perform dependent on the phase of the process a compound chemical quality check. This check should preferably be done by LC/MS, but also nmr may be sufficient information. Data are stored in a central database.

- The automated screening assay is then performed according to an optimised protocol. Data are collected and saved for elaboration together with the plate bar-code. In the standard protocol used in the GlaxoWellcome laboratories in Italy the primary screening result is confirmed and subsequently a quantitative dose response curve is obtained with a fresh sample of the positive compounds that is obtained from the solid store and dissolved only shortly before the experiment.

- These data can then be elaborated and linked to the plate map information. In that way it is possible to have an immediate result file or report comprising all registration numbers of the compounds that were positive in the screen. The data, either positive/negative or an activity or affinity value must also be saved in a central database.

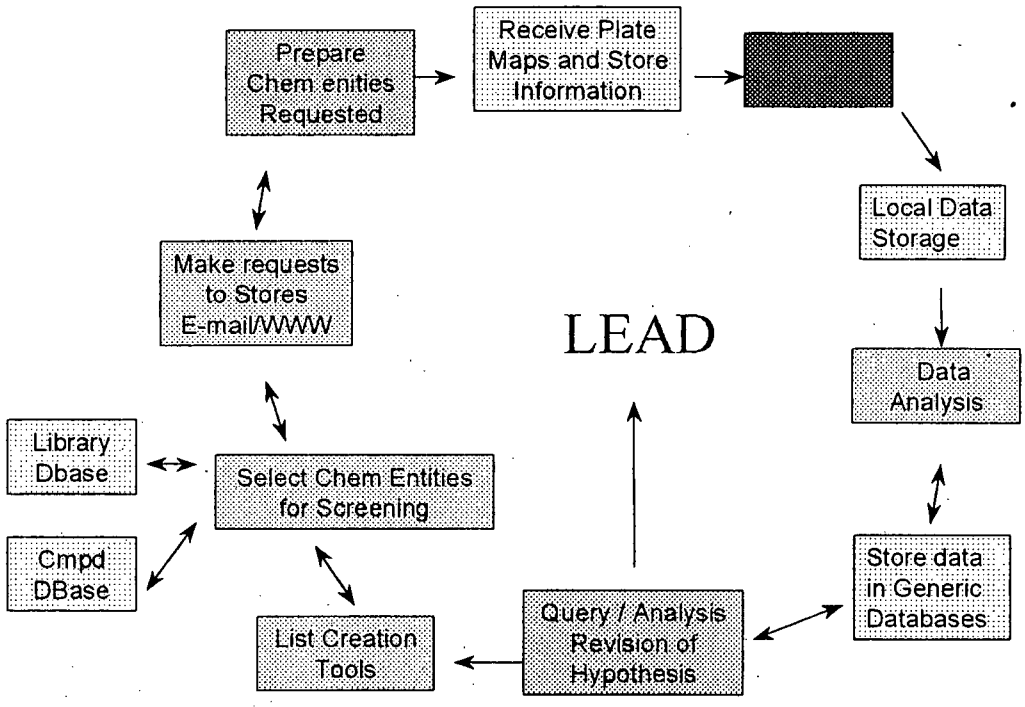


Fig 6. *Compound and Data Flow during the Screening and Hit Validation Process*

- The last phase of the screening process is the analysis of the data using computational modelling tools and the review of the model that had been hypothesised to describe the interaction of the compounds with the active site. On the basis of the results a new set of compounds will be selected or designed and the process is restarted from scratch. As these data are available to everybody in the company, they can be used as such by project biologists or chemists, but also as secondary data for other projects in the same or other research centres of the company.

Hardware and Software considerations for Robotic Screening

Hardware

Running a robotic, *i.e.* a fully automated screening station requires a set of instruments that can be controlled by a computer and serviced by a robotic arm or a gripper.

Typically a screening station processes microtiter plates and comprises some form of a robotic arm, liquid management instruments (dispensers, aspirators, either with fixed tubing or with disposable pipette tips), an incubator and one or more plate readers. In addition a series of accessory instruments is used that may comprise plate storage devices, shakers, plate sealers, centrifuges, turntables, delidding devices and not in the least place a barcode reader (Harding et al, 1997). All these components can be commercially acquired and integrated by specialised engineering companies. The choice of the instruments depends clearly on the function of the station. If a single assay type will be performed, usually relatively small stations will be produced with a fixed or only minimally moving robotic arm. When a large flexibility is required, multiple readers, incubators and various accessories may be needed at the station and often the robotic arm (one or more) may be moved over a 4-6 m track to allow the servicing of a relatively large surface area with a variety of instruments.

Small systems may have advantages, as they can be provided by several suppliers almost off the shelf, cost relatively little and can easily be put in special conditions to allow particularly hazardous operations (radioisotopes, use of cells or bacteria). The larger systems, however, are very flexible, have a large hands off capacity, and may still be easily adapted to new needs.

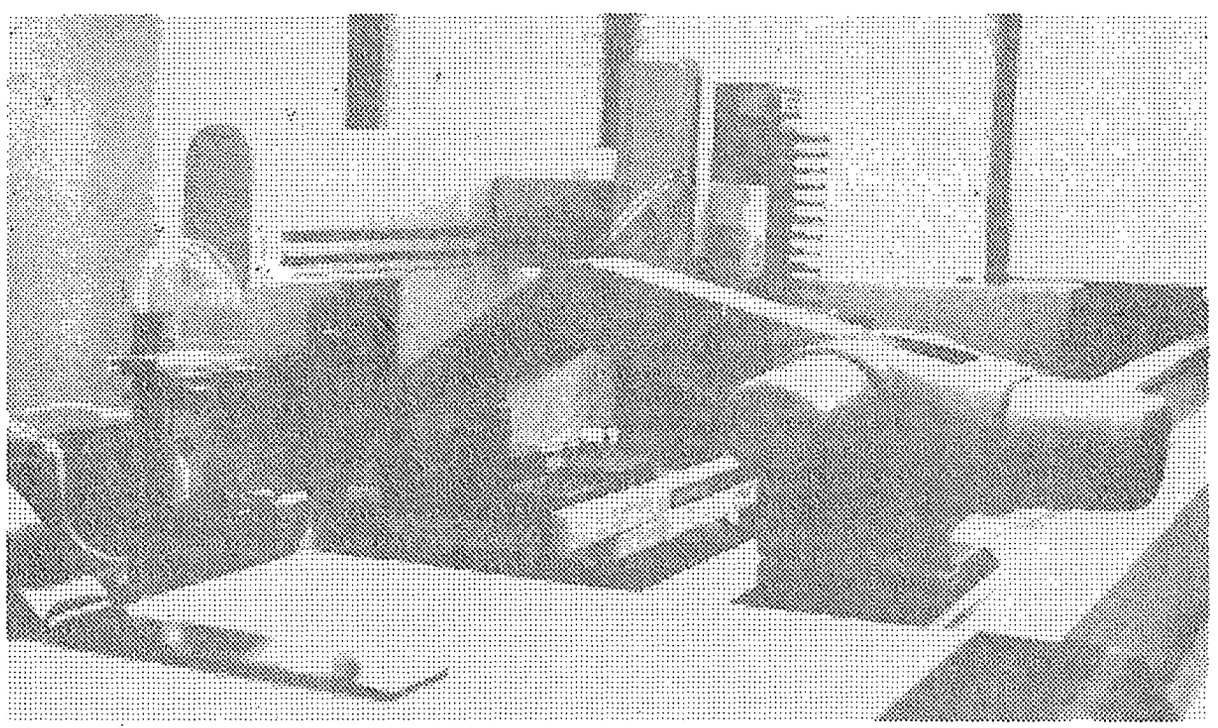


Fig. 7. *A view on a robotic screening station*

Software

Several pieces of software are needed to support robotic screening and to render it fully hands-off:

- Plate map management: there is a need to identify the plates that are being screened and to have a database from which the content of each well in each plate can be retrieved. In this respect, you will also need a utility that permits you to create plate maps and to save them in your database. This activity can also be performed by the chemists that create the plate or by a compound control function in the company.

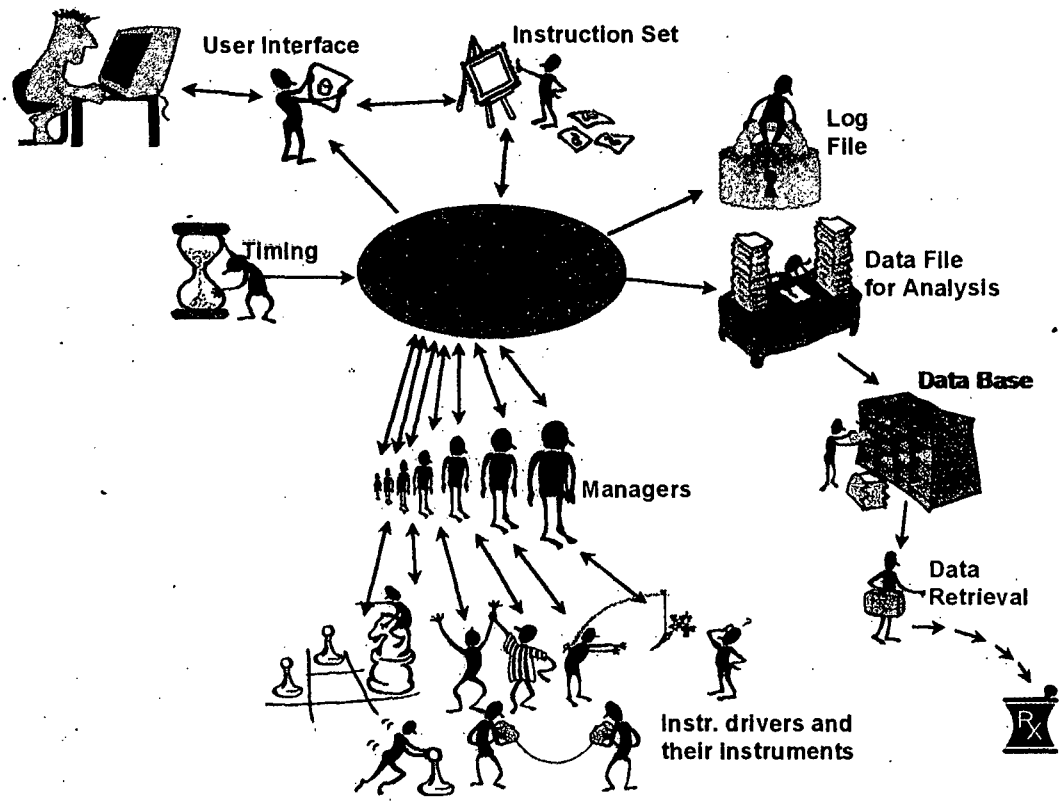


Fig 8: *The dynamic scheduler is instructed by the experimenter. The instructions are converted into machine language and passed to the kernel that takes full control of the system and informs the user on its progress.*

- Scheduling software: to screen hundreds of microtiter plates in a continuous fashion you need a scheduling software that can decide which plate has to execute a given step of the protocol at a certain moment. In this respect the better schedulers are so called dynamic schedulers, which are able to define the protocol priorities while the experiment is running. In this way the system becomes reactive to unexpected situations in the laboratory (refill of reagents; an error condition of any of the instruments of the station; queuing of plates in a particular stage of the protocol as the time expected for that particular step may be longer than expected). A non-dynamic scheduler often leaves no other choice but quickly eliminating plates in case of an unexpected situation. This may, however, be sufficient in small stations with relatively simple protocols and an easy direct control. It must be recalled that the scheduling software should be able to translate the assay conditions, which will be inserted in a simple way by the experimenter, in an optimised activity

map. The scheduler should allow to define incubation times quite precisely in such a way that all plates are treated in an almost identical way.

- Data collecting software: the data produced during the robotic experiment need to be captured and saved together with the barcode of each plate on the controlling computer, preferably in a way which is independent of disturbances of networks.

- Data analysis software: The collected files need to be analysed against the original plate maps, which allows the identification of positive compounds. Usually two different ways of analysis are requested in HTS: the first existing of the identification of those compounds that have a signal with a value below or above a predefined cut-off value, the second should allow some form of quantification of the affinity/potency/activity of the compounds studied, preferably using non-linear least squares fitting algorithms.

- Finally, a centralised database is needed where all results of the tests can be stored for each compound on each plate. These data can then be used for more sophisticated chemometrics analysis in any laboratory that can access the same centralised database.

Future developments:

Large, flexible, multifunctional screening stations are particularly useful in situations where the station must be used for all possible targets and its activities are concentrated on the high throughput part of the screening campaign. Alternatively, especially in large companies, several more specialised systems may be recommended that concentrate on different areas of screening such as non radioactive in vitro assays, radioisotope assays, cellular assays or others.

An additional development is the introduction of smaller, dedicated screening stations that are particularly useful in the assay preparation phase and in the hit/lead optimisation phase. These small stations can easily be incorporated in any biological lab where the project scientists themselves can perform the many activities in the course of a project that need processing of typically tens of plates each time.

Assay formats used in these stations will gradually shift from radioisotope formats and their problems of waste disposal to non-radioisotope formats ([tr]-fluorescence and polarised fluorescence). However, the ease of SPA technology and the miniaturisation that can be introduced with the Leadseeker technology will certainly keep the use of radioisotope assays alive for quite some time. Also more complex, cellular screening formats may be used more and more, particularly using smaller, specialised screening stations.

Future developments also include the use of ultra large libraries in ultra-high density formats, not anymore confined to the classical plate format where each entity tested is placed in a physically separated container, the plate well. This assay format is particularly indicated for assays using bead-based chemistry. As a consequence, also a different way of data collection, mostly centered around the collection of ar gel image of a plate rather than measuring the signal of each well must be used as analysing each position separately takes too much time. Image analysis software converts the collected image in data related to each compound, which can then be analysed as before. This way of analysis may then also be applied to the lower density screening formats and in the end may render classical plate readers redundant. In those cases where the bead itself remains on the assay matrix, analysis of the assay happens also simply by visual inspection, as the place where the positive reaction has taken place can be easily identified by eye.

It will be an interesting experiment to compare the success rate of screens performed with compound collections and with combinatorial and/or bead-based chemistry, measured with the only measure possible, the number of high quality leads that can be derived from a screening campaign.
Paradoxically, in this particular ultra high throughput screening format, when the beads have been properly tagged, the 'positive' bead may be picked up manually for analysis of the compounds involved, while the assay can often be performed on the bench without the need of large and expensive robotic screening stations.

Acknowledgment

Dr G. Terstappen is thanked for careful reading of the manuscript and for his contributions to the bio-technological part of the paper.

References

Alam J and Cook JL, Reporter Genes: Application to the study of mammalian gene transcription, *Anal. Biochem.* 1990; 188:245-249.

Babiak J, Transforming your robotics into an infrastructure for the future. *J Biomol Screening* 1997; 2(3):139-143.

Beeley LJ and Duckworth DM, The impact of genomics on drug design. *Drug Discovery Today* 1996; 1(11): 474-480.

George SE, Bungay PJ and Naylor LH, Evaluation of a CRE-Directed Luciferase Reporter Gene Assay as an Alternative to Measuring cAMP Accumulation. *J Biomol Screening* 1997;2(4):235-240.

Harding D, Banks M, Fogarty S and Binnie A, Development of an automated high-throughput screening system: a case history. *Drug Discovery Today* 1997;2 (9): 385-390.

Hart HE and Greenwald EB, Scintillation Proximity Assay (SPA)-A new method of immunoassay. *Mol Immunol* 1979; 16: 265-267.

Hemmila I and Webb, S, *Drug Discovery Today,* 1997;2(9):373-381.

Suto CM and Ignar DM, Selection of an Optimal Reporter Gene for Cell-Based High Throughput Screening. *J Biomol Screening* 1997;2 (1):7-9.

Natural Products and Drug Discovery

Dr. Lucia Carrano
Microbial Technologies
Biosearch, ITALY

Introduction

Natural products were largely used in all the ancient and folk medicine. The first documented use of natural materials and plants was in 3000 BC. Officially, the screening of natural compounds from soil microorganisms was started in 1940 by Wakesman. Since then, thousands of molecules with antitumour, antiinfective, antiparasitary properties or able to act as immunomodulators, herbicides, fungicides, pesticides have been discovered. Pharmaceuticals of microbial, plant or animal origin account for more than 30% of the current worldwide human therapeutic sales; four out of the twenty best-selling drugs are obtained from natural sources. It is important to note, however, that penicillin was discovered in 1928 by Fleming and rediscovered by Chain *et al* in 1940. This represents the biggest issue in the natural products study: to discover a molecule not yet described by someone else.

The term **natural product** is used as synonymous of **secondary metabolite**. These are molecules not essential to the growth and life of the producing organism that often confer selective advantage on it. These molecules are synthesised by a wide variety of biochemical pathways. The term secondary metabolite is used often as synonymous of **antibiotic**, where this latter term includes all substances that possess a biological activity.

Antibiotics are rather small molecules, whose MW range from 200 to 5000 Da, have a complex elemental composition and a complex molecular structure. Generally microbes start antibiotic production in the presence of nutritional limitations or other form of stress. It is also important to note that the same metabolite may be produced by different microorganisms.

Antibiotics belong to a definite number of chemical classes. It is important to underline that compounds belonging to the same chemical class do not necessarily have the same biological properties. One of the most important characteristic of natural compounds is that they have much higher affinity for the target enzyme or receptor and more unpredictable chemical structures than synthetic drugs.

The search for new bioactive secondary metabolites

New strategies for searching new natural compounds involve all the steps of the discovery process: identification of sources, samples preparation, target and assay identification, screening, compound purification and structure determination. The natural compounds sources include: microbes, plants, insects, sponges. Many efforts are directed at identifying new sources to be exploited. Therefore, in order to increase the possibility of isolating new molecules from plants, botanists are seeking plants from unusual regions of the world and are carefully examining previously studied plants.

Tropical rain forests represent one of the most investigated environments. The usual output in the search for new drugs from plants is information on the chemical structure of the molecule active in the biological system under investigation. Once the structure is known the amount necessary for further studies must be available. Extraction of the active compound from the specific plants cultivated in assigned lots usually implies a long lag period. Often, chemical synthesis is necessary to reduce the ecological impact of extracting the compounds from the plant's natural environment. The most clear example is represented by taxol: this molecule has a very good activity against ovarian cancer, but up of 3000 trees were necessary to afford sufficient compound for phase 1 clinical studies. The *Taxus*

brevifolia is a very slow growing tree and several studies were undertaken to obtain the molecule from different sources without causing an ecological damage. Some investigated whether the taxol was produced by more easily cultivated plant species, but unsuccessfully. Total synthesis of the molecule was attempted and realized through four different strategies, all too complicate to be economically viable. Attempts to engineer plants to produce taxol were also unsuccessful. Hopefully, semisynthesis of taxol starting from the natural, easily obtainable precursor 10-deacylbacchatine opened up the way to an easy supply.

Microorganisms represent the easiest scalable producers of molecules. They are virtually ubiquitous and can be isolated from many different sources including soils, water, plants, etc. In order to increase the possibility to discover new molecules efforts are devoted to the isolation of "unusual" strains. The most exploited microbial source is represented by *Actinomycetes* (82% of the microbial producers, among them *Streptomyces* the best producers) these microorganisms are also fast and easy growing species. For the isolation of unusual species key factors are source, treatment and cultivation. New strain isolation procedures can be divided in two types: sample source pretreatment and selective culture conditions. A typical example of sample pretreatment is the extraction with an organic solvent of the sample source followed by the plating of the two phases separately; spores are extracted by the organic phase and fungi are eliminated from the sample source. Another typical pretreatment consists in heating the sample source at high temperature, e.g. 100°C for several minutes. In this way for example *Actinomadura* species will be isolated more easily. As an example of selective culture conditions, we can consider plating on medium containing chitin as the sole carbon source only those strains able to metabolize this polymer wil be isolated.

An important contribution to the isolation of new species from sample source is given by **molecular biology techniques**. It is possible to know if desired species are present in a given source. It is well known that 16S RNA are very conserved and characteristic of a genus/species. So, specific oligoprobes for a given 16S RNA sequence can reveal the presence of the desired bacterium in the source. Once detected, appropriate conditions can be devised to isolate it.

It is well known that antibiotic production is strain specific and temporarily regulated so once microbes have been isolated, it is possible to intervene on the fermentation conditions. Secondary metabolite production is affected by physicochemical factors such as medium aeration, pH and temperature; nutritional factors such as glucose, ammonium, phosphate, metal ions; the age of the culture and many other factors. Many screening programs include fermentation of the isolated strains in 3 or 4 different media to increase chances of discovery new metabolites.

Screening process

In many instances, a microbial culture is not suitable for screening. Therefore, sample preparation may be necessary and several methods have been established. They range from solvent/solvent extraction (*e.g.* culture broth is extracted with water immiscible solvents), solid/solvent (*e.g* lyophilized plant material is extracted with organic solvents), chromatographic procedure such as solid-phase extraction (SPE). These procedures are then followed by concentration and preparation of the sample in a biocompatible organic solvent. These procedures are normally executed by robots in industrial labs to achieve speed and consistency.

The identification, purification and structural characterization of active individual compounds from complex mixtures, that constitute plant or microbial extract, are costly and time consuming. Two methods are used for compound identification: chemical- or activity based screening.

An example of the chemical screening is the identification of a peculiar functional group in the molecule typical of a given chemical class, believed to possess interesting biological properties. For this purpose

several simple colorimetric methods have been developed (*e.g.,* alkaloids can be purified by using Dragendorff's reagent).

Activity based identification is preferred because it does not put any structural bias on the active compound. An assay must be designed and validated to allow identification of interesting activities. In the pharmaceutical world the assay is a simplification of the pathology to fight; e.g. for finding a new drug able to cure AIDS several assays are conceivable that are representative of the therapeutic target and that can be run in microplate format. A possibility would be to look for inhibitors of the HIV-protease another for the viral titre in an infected cell line. In any case a simple HTS assay is designed, whose output results in the detection and quantification of a measurable (colorimatric, fluorometric etc.) signal.

The assay has to be representative of the target and obey to simple rules. It must be selective and specific; it must allow the identification of an active molecule in a mixture, while sensitive enough to detect faint activities, since the active compound may be present in small traces or have poor activity; it must be robust, *i.e.* not affected by interferences and with good reproducibility. The assay will be run in automated and miniaturized HTS format (to 100 µl or less). The ideal assay should also be simple. The most intriguing step in the search for new natural products is probably how to transform a target in a simple assay whose output is clearly measurable. Attractive systems are reporter gene assays, which look at the expression of a reporter gene (*e.g.* coding for luciferase) under the control of a promoter whose activity modulates the presence and amount of luciferase; they can be easily quantitated by measuring a chemiluminescent signal. Modulation of the promoter will result in modulation of the luciferase gene expression. A big impact on target identification is coming from genomics. Several genes involved in a pathology have been identified (*e.g* through gene knock out experiments). These genes represent good targets and modulators of the activity of these genes or of their products may have good chance of becoming drugs.

The advent of automation has affected the way of looking for activity. For example, once robots started running assays, the most common antimicrobial test (the presence of an inhibition halo, on an agar plate inoculated with the test microorganism) was substituted by measuring the turbidometry of a microbial suspension grown in the presence of the sample.

The speed of the discovery process has been increased not only by automating assays but also by the construction of large databases containing information available on known natural products. These databases allow simple queries for the rapid identification of molecules with similar properties (*e.g.* molecular weight, solubility , biological activity) making it possible assess if the compound under analysis has been already described. NAPRALERT is one of such databases available on the web. Large agrochemical and pharmaceutical industries have their proprietary database. Two imperatives in discovering new compounds are miniaturization and speed. A drug discovery program from natural sources is time consuming. Significant progress in analytical technologies makes it possible to collect information on MW, fragmentation profile and UV-spectrum of active compounds present in the mixture.

Unnatural natural products

Despite the fact that natural products are being found continually by pharmaceuticals companies through targeted screening programs, fewer new type of drugs have been found in the last decade.

Biological modification of antibiotics with the aim of obtaining products with better therapeutic properties have been carried out since the beginning of antibiotic era, when several penicillins were obtained by adding specific precursors to *Penicillium* cultures. Cephalosporins are important antibiotic obtained by semisynthesis using as aminocephalosporanic acid (7ACA) or

aminodeacetoxycephalosporanic acid (7-ADCA) as precursors. They are produced by *Cephalosporium*. *Penicillium* produce large amounts of penicillin N. Introduction in *Penicillium* of the gene coding for the expandase (which changes the 5 atom ring of penicillin into the 6 atoms ring of cephalosporin) and feeding the culture with adipic acid results in the production of 7-ACA. At Merck researchers managed to produce of 7-ACA or 7-ADCA with high yield.

Then combinatorial chemistry approaches together with combinatorial biosynthesis help to remedy the lack of new molecules. The goal is to make new antibiotics by exploiting the microbial genes and enzymes that make these substances and produce unnatural natural products. Combinatorial biosynthesis is being successfully applied to bacterial polyketide synthases (PKSs).

Polyketides constitute represent a large family of natural products some of them are successful drugs. Their biosynthesis is mechanistically equivalent to formation of long fatty acids chain, but many polyketides are made by modular systems. By exchanging some modules with other ones from different sources it has been possible to produce by genetic manipulation polyketide strutures that were not described previously.

Conclusion

Biologically active compounds isolated from natural sources are vital to the discovery and development of new active principles. New technological development in many disciplines from genetics to analytical chemistry, from automation to informatics, have opened the way to more effective discovery programs for natural products. Furthermore, combinatorial biosynthesis has increased the possibility of modifying existing natural products. Further developments will continue to be required to ensure success in the discovery of new biologically active natural compounds.

References

Omura, S. ed., **The search for bioactive compounds from microorganisms**. Brock/Springer series in contemporary Bioscience.

Rehm, H.J. Reed, G.,Puhler, A. and Stadler., P. eds., **Biotechnology** vol 7, VCH.

Grabley, S. and Thiericke, R. eds., **Drug discovery from Nature** , Springer-Verlag.

Hutchinson C., *Microbial polyketide synthases: more and more prolific.* 1999 Proc. Natl. Acad. Sci. 96,3336-3338.

Cordell, G. 1995. *Changing strategies in natural product chemistry. Phytochemistry.*40, 1585-1612.

Carrano, L. and Donadio, S., **Key ingredient for high throughput screening. in Combinatorial chemistry and combinatorial technologies: method and applications**. S. Miertus and G. Fassina eds., Marcel Dekker inc (in press).

COUNTRY

REPORTS

Development of Combinatorial chemistry and Combinatorial Technologies in Hong Kong

Dr. P.K. Wong
Department of Biology
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong, SAR, CHINA

Introduction

To the turn of the millennium, there is a rapid change in Hong Kong to cope with the new challenges ahead and to restore herself to be a prospective city after the Asian economic crisis. The success of Hong Kong in being a well-known, most competitive and attractive city for investment lies in the manufacturing of high-quality and comparatively low-cost products in the early 60's-80's. The rapid economic development of Hong Kong in 80's-90's is mainly derived from the finance and property sectors, while the people employed by manufacturing industry dropped from 46% in 1980 to 12.3% in 1997 (Fig. 1) and the GDP contributed by manufacturing industries dropped from 23.7% in 1980 to 7.3% in 1996 (Fig. 2) (Industry Department, 1999).

The major manufacturing industries in Hong Kong are:
(1) Textile and clothing industry (Textile: First world exporter; Clothing: Second world exporter)
(2) Electronic industry
(3) Printing and publishing industry
(4) Food and beverage industry
(5) Metal products industry
(6) Watches and clocks industry (Watches: Second world exporter; Clocks: First world exporter)
(7) Chemical industry
(8) Jewelry industry (Imitation jewelry: First world exporter; Precious metal jewelry: Fourth world exporter)
(9) Industrial machinery industry
(10) Toy industry (First world exporter)

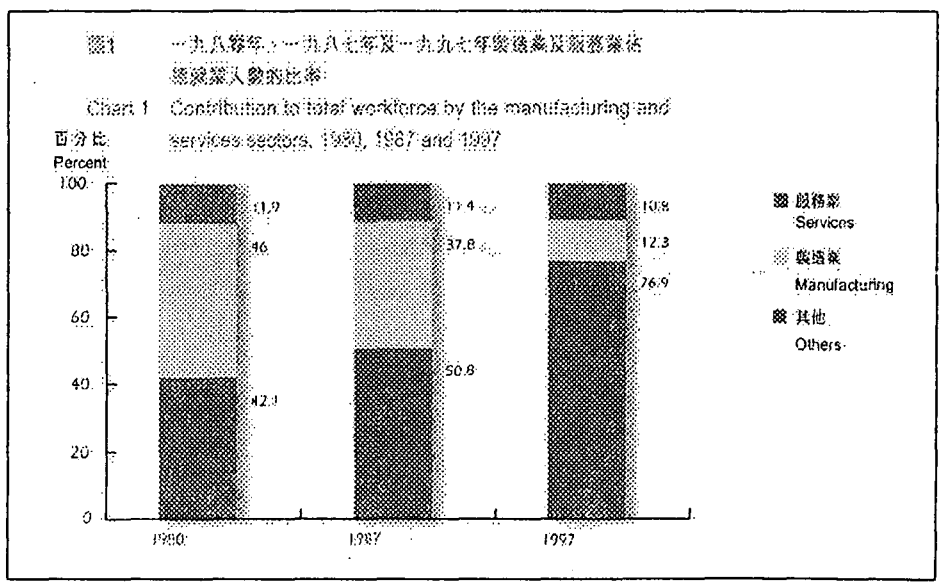


Fig. 1. Contribution to total workforce by manufacturing and services sectors, 1980, 1987, 1997.

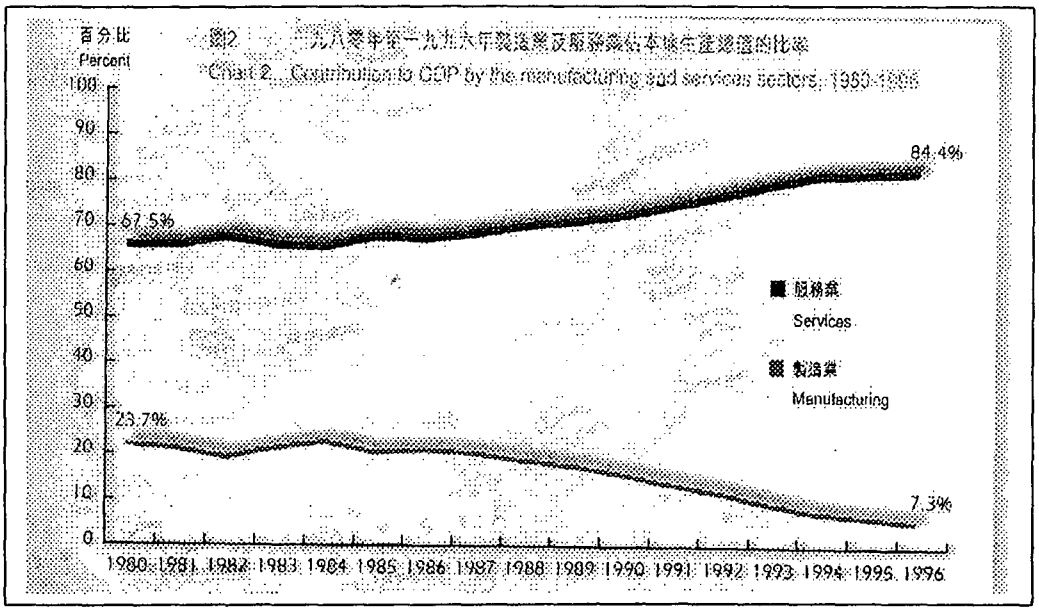


Fig. 2. Contribution to GDP by the manufacturing and services sectors, 1980-1996.

Electronic, textile and clothing are the mainstays of Hong Kong's manufacturing sectors. Although there is chemical industry activity in Hong Kong, the major activities are packing, mixing and blending of chemicals into marketable products. There is no production of chemicals in Hong Kong.

After the Asian economic crisis, Hong Kong government like all her counter-parts in Asian countries refocused the importance of building strong industrial foundation for the continuous growth of the national or regional economies. Since 1993, Hong Kong government started to provide comparatively large sums to support the development and application of new technologies and to enhance the competitiveness of the industries in the region. However, there is still no sufficient funding to support industries (data available up to 1998). The government has planned two main projects that will directly increase the quantity and quality of the Hong Kong industries: they are the "Science Park" and the "Cyber-Port" projects. The former will emphasize technology development and manufacturing, while the latter will provide excellent infra-structures for marketing and promoting the information technologies invented locally or overseas into the region.

The Science Park will be completed and operative by the year 2001. It will accommodate many ventures for research and development (R & D) in the high technology fields such as information technology, material science, biotechnology and traditional Chinese medicine (TCM), etc. Biologically active compounds and drugs will be the major products of the R & D in biotechnology and TCM industries. Besides using molecular technologies to develop new drugs in biotechnology and to routinely screen for medicinal compounds from Chinese herbs, the use of combinatorial chemistry/combinatorial technologies (CC/CT) should be an important area in the development biotechnology and TCM industries.

Although the use of combinatorial chemistry was introduced in early 80s, extensive application of the technology did not appear until the last few years. In Hong Kong, as in most countries in the region, not enough efforts have been allocated in the development and applications of CC/CT. Few studies in this area have been supported mainly by the funding from the Research Grant Council (RGC) and Industrial Support Fund (ISF) of the government. Most of these studies focused on the development of new drugs in medicinal applications. However, due to the lack of development/studies in the other areas of CC/CT, such as automation and high throughput screening, the advance of the CC/CT in Hong Kong is comparatively slower than that in Japan, Taiwan and Singapore. The introduction of CC/CT in the R & D for biotechnology and TCM industries will greatly enhance the potential and success of the

development of these industries in Hong Kong. The provision of a research centre of CC/CT by the Hong Kong government is the most productive initiative to develop and promote the CC/CT in the city/region.

Funding agencies in Hong Kong

There are several funding sources in Hong Kong to support research in various disciplines, and most of the researches are conducted in universities in Hong Kong. Since there is no biotechnology and pharmaceutical manufacturing industry in Hong Kong, the major funding support for the research in universities is Research Grant Council (RGC, a governmental organization to fund research projects in all universities in Hong Kong). There was no central funding agency in Hong Kong until the establishment of RGC in 1988. In 1988, there was only HK$ 20 millions per year allocated to support research projects in 2 universities, 2 polytechnics and two colleges. In 1998, the total amount of research grants allocated to projects in 9 government-funded institutions through the RGC was increased to 200 million per year (a 10 folds increase). Most funded projects receive an average of about 0.25 to 0.5 million per year. Most projects received funding for two years, some projects were funded for 3 years and few projects were funded for 1 year.

The Croucher Foundation is a non-governmental funding agency established in early 80's. It was a foundation of a British businessman who had his business in Hong Kong until early 80's. The foundation has supported and will support various academic activities including research projects. The allocated sums and the duration for each project were very similar to those of RGC. After the turn-over of Hong Kong to China in 1997, the foundation refocused his mission and started to supports academic activities other than research projects.

In 1994, Hong Kong government established an Industrial Support Fund (ISF) which supports research projects with the goal of increasing the productivity and quality of Hong Kong industry. There are two categories of funding, one to support projects under HK$ 1 million and another to support projects over 1 million. The latter can even be used to support establishment of infrastructures. The amount of funding of the second category ranged from more than HK$ 1 million to 38 millions, and most projects received funding of HK$ 1 to 10 millions for 2-4 years.

Research centres of biotechnology or traditional Chinese medicine in Hong Kong

Hong Kong Institute of Biotechnology (HKIB)

It was established in 1991 by a donation of HK$ 170 millions from the Hong Kong Jockey Club. The major portion of the funding was used to build a main building for research laboratories, offices and equipment rooms and a building for staff and visiting researchers' accommodation. A Food and Drug Administration (FDA)-standard laboratory was built. The remaining funding was used to support projects in various areas. As the establishment of ISF in 1994, most projects of the institute have been supported by research grants of ISF.

Biotechnology Research Institute (BRI)

The institute is located at Hong Kong University of Science and Technology, and was established by a donation of HK$ 130 millions from Hong Kong Jockey Club in 1991. The entire amount of donation has been used to support research projects in biotechnology. The duration and amount of funding to support each project are 1-3 years and HK$ 0.2 to 0.5 million per year.

Institute of Molecular Biology (IMB)

The institute was established by private donations at The University of Hong Kong. Similar to Hong Kong Institute of Biotechnology, a major portion of the donations was used to build new premise for research laboratories and offices, and the remaining portion of the donations has been used to support research projects.

Chinese Medicinal Materials Research Centre (CMMRC)

The Centre was established in early 80's with the support of private donations to The Chinese University of Hong Kong. The Centre has its research laboratories and offices and an accommodation facility for staff and visiting researchers. The Centre attracts donations and research contracts from private sectors engaging in businesses related to Chinese medicine. Starting from 1995, the Centre also received research grants from the ISF to support research projects in traditional Chinese medicine (TCM).

There are many ongoing research projects, involving application of CC/CT, that can be categorized into 2 parts - on-going and completed projects which were funded by funding organizations in 1997 and before, and new projects which have been funded by various funding organizations in 1998 and after. Almost all projects were/are conducted in six government-funded universities. They are:

- City University of Hong Kong (CiUHK)
- Hong Kong Baptist University (HKBU)
- Hong Kong Polytechnic University (HKPU)
- Hong Kong University of Science and Technology (HKUST)
- The Chinese University of Hong Kong (CUHK)
- The University of Hong Kong (UHK)

On-going and completed projects on CC/CT

List 1 compiles all funded on-going and completed research projects on CC/CT in Hong Kong.

List 1: On-going and completed research projects on CC/CT (extracted from the reports of University Grant Committee and Industry Department, Hong Kong SAR Government)

- Discovery of traditional Chinese medicine (TCM)-derived saponins with anti-tumour and cardiotonic activities for use as health food supplements or therapeutics (ISF, 1997/98, K.P. Fung, CUHK, HK$ 1,968,000)

- Protein engineered enzymes for efficient manufacturing of β-lactam antibiotics (ISF, 1997/98, M. Wu, HKUST, HK$ 3,100,000)

- Oral DNA vaccines (ISF, 1997/98, K.Y. Yuen, UHK, HK$ 4,945,000

- Novel chiral catalysts containing spirocyclic ligands and their application in asymmetric catalyst (RGC, 1997/98, .S.C. Chan, HKPU, HK$ 446,000)

- Design, synthesis, and application of novel bimetallic and multimetallic chiral homogeneous transition metal catalysts (RGC, 1996/97, S.C. Chan, KHPU, HK$ 1,162,000)

- Novel polymer-supported chiral catalysts for the production of fine chemicals and pharmaceuticals (RGC, 1995/96, S.C. Chan, HKPU, HK$ 531,000)

- High level production of biologically active recombinant protein in bacteria (ISF, 1996/97, H.C. Wang, HKUST, HK$ 2,329,000)
- Development of high throughput DNA-base diagnostic microchips for the international health-care market (ISF, 1996/97, M. Yang, CiUHK, HK$ 996,000)
- Investigation of anti-rheumatic drugs: triptolide, bucillamine and their analogues (RGC, 1996/97, UHK, HK$ 692,000)
- Antihormonal substances and inhibitors of 5-reductase and aromatase from natural source (RGC, 1995/96, C.T. Che, HKUST, HK$ 431,000)
- Optically active polyacetylenes containing heteroatom chiral centers: Synthesis, characterization, properties, and application (RGC, 1995/96, B.Z. Tang, HKUST, HK$ 431,000)
- Development of rapid diagnostic tests for concomitant detection and enumeration of pathogens by multiplex PCR (ISF, 1995/96, Y.C. Kong, CiUHK, HK$ 2,231,000)
- Trichosanthin derivatives with reduced immunogenicity and nephrotoxicity and long plasma half-life (RGC, 1995/96, P.C. Shaw, CUHK, HK$ 581,000)
- Novel transition-metal phosphine complexes and their potential industrial applications (RGC, 1994/95, S.C. Chan, HKPU, HK$ 1,157,000)
- Structure/function relationships among avain immunoglobulins (RGC, 1994/95, D.A. Higgins, UHK, HK$ 619,000)
- Synthesis and structural studies of metal complexes containing polyfunctional carboxylate-like ligands (RGC, 1993/94, C.W. Mak, CUHK, HK$ 1,000,000)
- DNA as affinity probe for the detection and isolation of bioactive molecules (RGC, 1993/94, C.T. Che, HKUST, HK$ 150,000)
- Production of autoantigens by gene fusion and peptide synthesis for use in the immunodiagnosis of autoimmune disease (RGC, 1993/94, L.P. Leong, CUHK, HK$ 551,000)
- Design and synthesis of potent DNA-cleaving and anticancer agents by *in-situ* formation of 10-membered ring enediyne structures (RGC, 1993/94, W.M. Da, HKUST, HK$ 462,000)
- Synthetic, structural, mechanistic and catalytic studies of metal-organic compounds (RGC, 1991/92, C.W. Mak, CUHK, HK$ 840,000)

New Projects on CC/CT

In addition, projects in CC/CT funded in 1998 (denoted as new projects) are given in List 2:

List 2: Newly funded research projects on CC/CT (extracted from the reports of University Grant Committee and Industry Department, Hong Kong SAR Government)

- Synthesis and applications of novel chiral catalysts containing a partially hydrogenated binaphthyl backbone (RGC, 1998/99, S.C. Chan, HKPU, HK$ 485,000)

- Rapid high throughput DNA analysis technology: Integration of DNA amplification and detection on microchip (ISF, 1998/99, M. Yang, CiUHK, HK$ 3,166,200)

According to a new journal "Combinatorial Chemistry" (Fig. 3) which was first published in winter 1998 as an affiliated section of a well-known biotechnology journal - "Biotechnology and Bioengineering", the major areas of CC/CT are: Solid phase synthesis, solution phase synthesis, automation, computational, analytical, screening and patents. The categories are matched perfectly with the thematic topics of the ICS-UNIDO-UPLB Workshop held in the Philippines except studies in "natural products" should be added in the lists. The new, on-going and completed research projects mentioned above can be categorized into these 8 areas (Table 1).

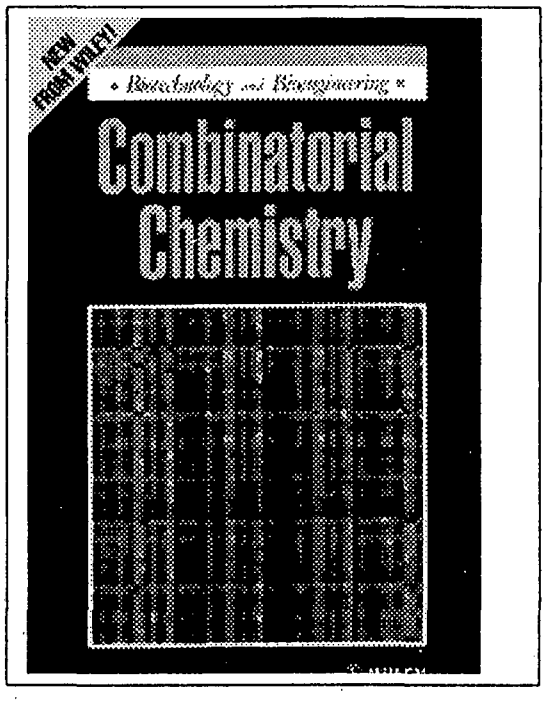


Fig. 3. *Combinatorial Chemistry* - a new journal/an affiliate section of *Biotechnology and Bioengineering.*

Table 1: Classification of research project in various areas of CC/CT

Area	Frequency
(1) Solid phase synthesis	0
(2) Solution phase synthesis	0
(3) Automation	
(a) synthesis	0
(b) screening	+
(4) Computational	+
(5) Analytical	+++
(6) Screening	+++
(7) Patents	+++
(8) Natural products	+++++

Most of the new, on-going and completed projects focus on the production or improvement of natural products, some involve "analytical", "screening" and "patenting" aspects in CC/CT. Few projects are on "automation" and "computational". There is no project involves in "solid phase synthesis" and "solution phase synthesis".

Further development of CC/CT in Hong Kong

Biotechnology and TCM will be two major developing industries that can use CC/CT in the "Science Park" project. Efforts should be concentrated in the application of CC/CT especially in these two disciplines. The Chief Executive of Hong Kong Special Administrative Region (SAR) has stressed on the importance of the development of TCM to Hong Kong economic in his "Executive Report" in 1999. In addition, China is the major producer of TCM and Hong Kong has been the major exporter of TCM in the region. From List 3, there is a trend of the funding agencies in Hong Kong to increase their supports to the research projects on TCM. I believe Hong Kong has most of advantages to develop TCM as one of her major industries in the coming decade.

List 3: Funded research projects on traditional Chinese medicine (extracted from the reports of University Grant Committee and Industry Department, Hong Kong SAR Government)

1998

- Screening and characterisation of proteins that interact with trichosanthin (RGC, 1998/99, P.C. Shaw, CUHK, HK$555,000)
- The application of verbascoside, a purified extract of Chinese medicine on recovery of muscle injury after immobilization (RGC, 1998/99, M.H. Yip, CUHK, HK$ 630,000)
- Internalization and routing of trichosanthin (RGC, 1998/99, S.C. Tam, CUHK, HK$ 645,000)
- Role of Chinese medicinal compounds in the regulation of stress-activated protein kinase in ischaemia/reperfusion (RGC, 1998/99, Y.L. Siow, UHK, HK$ 735,000)
- International ginseng conference '99 - Biotechnology, medicinal uses and marketing of ginseng (ISF, 1998/99, P.H. But, CUHK, HK$ 560,000)
- Workshop on "Intellectual Properties in Chinese Medicine, Health Food, Pharmaceuticals, and Biotechnology Industries" (ISF, 1998/99, H.W. Yeung, HKBU, HK$ 288,000)
- A process development facility for the extraction and concentration of TCM and herbs (ISF 1998/99, C.W. Kwong, HKIB, **HK$ 14,793,000**)

1997

- Discovery of traditional Chinese medicine-derived saponins with anti-tumor and cardiotonic activities for use as health food supplements or therapeutics (ISF, 1997/98, K.P. Fung, CUHK, HK$ 1,968,000)
- The Development of Chinese medicine and dietary supplements for treating gastrointestinal disorders and coughing (ISF, 1997/98, P.H. But, CUHK, **HK$ 8,799,000**)
- Development of over-the-counter pharmaceutical products based on local medicinal plants (ISF, 1997/98, Kong, Y.C., CUHK, **HK$ 8,128,000**)

- Chinese material standardization: Chemical studies of herbs for the United State's dietary supplement/nutraceutical market (ISF, 1997/98, H.W. Leung, HKBU, **HK$ 8,611,776**)

- International symposium on "The Worldwide Herbal Industry: Present and Future" (ISF, 1997/98, D. Chia, Federation of Hong Kong Industries, HK$ 608,300)

- Biomedical and traditional Chinese medicine database of China and Hong Kong (ISF, 1997/98, W.M. Chung, HKUST, HK$ 3,285,000)

1995

- Trichosanthin derivatives with reduced immunogenicity and nephrotoxicity and long plasma half-life (RGC, 1995/96, P.C. Shaw, CUHK, HK$ 581,000)

- Laboratory for quality control of Chinese medicines and health foods based on chemical components (ISF, 1995/96, P.H. But, CUHK, **HK$ 7,103,000**)

1993

- Identification of Chinese medicinal herbs by combined molecular and chemical approaches (RGC, 1993/94, P.C. Shaw, CUHK, HK$ 766,000)

1991

- Protein engineering of trichosanthin (TCS) and β-momorcharin (MMC): Modification and antigenic determinants and structure-function studies (RGC, 1991/92, P.C. Shaw, CUHK, HK$ 2,290,000)

The development of TCM industry in Hong Kong can be approached in following areas:

1) The policy to certify and license qualified practitioners in order to control the use Chinese medicine,
2) The policy and guidelines to control the quality (i.e. QA/QC system) of TCM,
3) Purification, identification and production of biologically active component(s) of TCM, and
4) Synthesis of purified component of TCM and produce new biologically active compounds by improvement of purified components of TCM.

The CC/CT can be applied to 2), 3) and 4). There are a number of centres which have been established or supported by ISF (List 4); most likely to build a CC/CT centre in Hong Kong will be crucial to introduce the vast and potential applications of CC/CT in TCM industry. Only a rapid and focused development of TCM industry can make it to be a competitive and economically viable industry in the region.

List 4: Some centres established or supported by ISF (extracted from the reports of Industry Department, Hong Kong Government).

- The establishment of a BRI traditional Chinese medicine centre: Drug development, safety, standardization and reformulation (ISF, 1997/98, Y. Ip, HKUST, HK$ 13,570,000)

- Establishment of Hong Kong Bioinformatics Centre (ISF, 1997/98, K.P. Fung, CUHK, HK$ 4,983,000)

- Manufacturing technology centre for human vaccines and pharmaceuticals (ISF, 1996/97, C.W. Kwong, HKIB, **HK$ 38,285,000**)

- Food packing technology centre (ISF, 1995/96, H.H. Liang, HKPU, HK$ 3,125,000)

- Drug delivery technology centre (ISF, 1994/95, T.F. Wang, HKUST, HK$ 2,529,000)

References:

Abelson, J.N. 1996. *Methods in Enzymology - Combinatorial Chemistry.* volume 267, Academic Press, New York.

Doyle, P.M. 1995. Combinatorial chemistry in the discovery and development of drugs. *Journal of Chemical Technology and Biotechnology* 64:317-324.

Ecker, D.J. 1998. *Combinatorial Chemistry.* John Wiley & Sons, Inc., New York.

Hall, S.E. 1997. The future of combinatorial chemistry as drug discovery paradigm. *Pharmaceutical Research* 14:1104-1105.

Industry Department. 1999. *Hong Kong Industry.* Printing Department, The Government of The Hong Kong Special Administrative Region, Hong Kong.

Lyttle, M.H. 1995. Combinatorial chemistry: A conservative perspective. *Drug Development Research* 35:230-236.

Terrett, N. K. 1998. *Combinatorial Chemistry.* Oxford University Press, Oxford.

Wang, G.P. and Xie, M.H. 1996. Combinatorial chemistry and its application to drug discovery. *Zhongguo Yiyao Gongye Zazhi* 27:552-559.

Wilson, S.R. and Crarnik, A.W. 1997. *Combinatorial Chemistry: Synthesis and Application.* John Wiley & Sons, Inc., New York.

Status of Research & Development on Combinatorial Chemistry and Combinatorial Technologies in India

Dr. Madugula Sree Rama Murty
Organic Division-I
Indian Institute of Chemical Technology
Hyderabad, INDIA

Combinatorial Chemistry and Combinatorial Technologies (CC/CT) are the important interdisciplinary subjects of current interest because of their fruitful output results in many fields of application. In India the application of CC and CT are mainly focussed on the pharmaceutical and catalysis applications and will be extended to other fields e.g., natural products and agrochemicals. The main focus of pharmaceutical researchers is to design and to synthesize a large number of organic compounds and to subject them for screening. This will be done by preparing single, pure compounds as much as possible and testing them for biological activity. Generally, thousands or even millions of new compounds have to be synthesized to find a single new drug molecule and this is practically a gambling game where probability and statistics will play a relevant role. In addition to this, the synthesis of many new compounds followed by the testing of each molecule is an extremely expensive, time consuming process.

Combinatorial chemistry has emerged as a result of the recent technology developments and aims at improving the speed and the efficiency of pharmaceutical research and of other disciplines. In order to discover a new lead drug structure, the philosophy of combinatorial chemistry is aimed at make discretes, or mixtures of large number of compounds and then submit them for activity testing.

The work on combinatorial chemistry in India was focussed on the above points and such type of work was initiated a year back at Indian Institute of Chemical Technology (I.I.C.T.), Hyderabad.

To our knowledge only few Industries have shown interest on CC/CT and no collaborative work has been already assessed between industry and CSIR institute on CC/CT in India. Mainly basic research has been initiated and no applied research was started.

Some of the funding organizations in India to promote research are the Department of Science and Technology (DST), the Department of Biotechnology (DBT), the Department Ocean Development (DOD) and CSIR. I.I.C.T. has international collaborative programs with SKB, Dupont and Cyanamid among the major pharmaceutical companies.

The widespread application of combinatorial Chemistry to drug discovery has led to an explosion of interest in solid phase organic synthesis. The Combinatorial Chemistry approach for the generation of libraries leading to useful drug-like principles is a very fascinating area. I.I.C.T. in collaboration with SmithKlineBeecham, U.K. has initiated the work on Solid Phase Organic Chemistry (SPOC). I.I.C.T. on its part has taken the development of several reactions on solid phase. The main aim is to evaluate the Solid-Phase reaction schemes, to assess them in solution, to transfer them on the support and finally to cleave from the resin the final target compounds for biological evaluation of the New Chemical Entities (NCE) synthesized. After the standardization of the reaction on solid phase the next target is the synthesis of chemical libraries. I.I.C.T. has an automatic synthesizer with well-trained scientists who will be engaged in the development of libraries based on the reactions that are developed at I.I.C.T.

I.I.C.T. has well established research areas both in the catalysis and material sciences. With the knowledge gained through SPOC at I.I.C.T. efforts are being initiated to work in both the above areas,

where development of good CC would lead to good CT. The research groups that are working in the area of Natural Products at I.I.C.T. have slowly but steadily initiated the wok in this area.

The ICS and UNIDO organizations are doing extremely great job by organizing workshops/Symposia on CC/CT to bring awareness among the researchers, technologists and academicians. The recent Workshop held in Hyderabad is an excellent example of ICS-UNIDO activity.

Combinatorial Chemistry in Malaysia*

The introduction of combinatorial chemistry in Malaysia is very recent. Researches in combinatorial chemistry and combinatorial technology (CC/CT) have not been really started yet but there is a group of researchers who is working on the computational aspects of CC/CT. This discussion will focus first on the different activities in my own institution and university, and present our plans to get involved in CC/CT. The government programs and initiatives to attract workers from various disciplines to conduct researches in CC/CT will then be discussed.

The University of Malaya

My institution (Faculty of Medicine) is part of the University of Malaya and is involved in the teaching and training of undergraduate and postgraduate students in medicine, pharmacy and other biomedical sciences. Aside from teaching, we are involved in training of hospital personnel. The university is attached to a hospital and this gives us the advantage of having access to patients when we conduct studies. We are also involved in training professionals in analytical methods and techniques from different industries, including pharmaceuticals.

The Faculty of Medicine also serves a lot of industries in the vicinities- chemical, pharmaceutical as well as industries involved in environmental control. Some of the work done involves bioequivalence and toxicological studies for large pharmaceutical companies as well as small local companies investing in the production of generic drugs.

Natural products research in our institute is not similar to what conventional chemists do in other environments. There is a clinic for emergency and accident cases as well as for normal diseases. In this clinic, cases of patients who take traditional medicine are studied. When a patient takes traditional medicine, he either recovers or goes into toxic state. When the patient recovers, then we analyze what is the active ingredient (if they can be found in the plasma) from the plant that was used. Toxicology cases involving the use of traditional medicine are also studied. On the premise that toxicology is pharmacology in small doses, the pharmacological properties of the toxic compounds are also studied. If results are considered interesting, the phytochemistry of the medicinal plant is then considered. There are a few cases wherein interesting biologically active compounds have been found so far using this strategy. There are now plans on pursuing combinatorial chemistry along these lines.

Other phytochemical researches that are carried out in the faculty are:

- Analysis for Forensic Medicine (screening and confirmation for drug abuse);
- Detection of adulterants such as steroids/toxic compounds in traditional medicinal preparations (adulterants are sometimes added even to traditional preparations to make them more potent; this is usually done for "Chinese" locally-made drugs, which are imitations of Chinese herbal preparations);
- Analysis of toxic compounds in food and food products, especially those which can act as endocrine disruptors.

In the university, there is CC/CT group based in the Chemistry Department, Faculty of Science. There is a computational group and a regional center for Computer-Aided Molecular Design is being established in coordination with ICS-UNIDO. They are at present doing more theoretical rather than practical work involving CC/CT.

Government Programs and Initiatives

In Malaysia, major efforts have been focussed on natural products research. This is because there are many natural sources and most of the work has been carried out through conventional phytochemical

---

* Text prepared by Dr. Cleofe A. Calanasan based on transcription of the presentation of *Dr. Mustafa Ali Mohd,* Department of Pharmacology, Faculty of Medicine, University of Malaya, 59100 Kuala Lumpur, MALAYSIA.

studies. Not much has been achieved so far because there is lack of coordination among workers in different disciplines. However, things are beginning to shape up and the government is more serious now in funding multi-disciplinary types of research projects.

At present, there is **IRPA** (Intensified Research on Priority Areas grant), and natural products research has been identified as one of the priority areas of the government. A big grant was given to researchers and these are assessed and revised every 5 years. There are also other overseas grants on natural products research from Japan, Korea and other UN-linked agencies.

The government has also formed a group called **MIGHT** (Malaysia Industry-Government High Technology group) under the supervision of the Prime Minister Department which focuses on national products research. The Prime Minister is very serious in developing the drug industry in the country and he wants a lot of efforts to be focussed on these because of the natural resources in the country. The main objective of MIGHT is to establish technology and processes related to the discovery of natural products. A big group of different disciplines from different universities, industries and research institutions is involved in MIGHT. A lot of funds are available since our R&D funds were not reduced with the Asian economic crisis.

The policies for MIGHT are as follows:

- Ensure that from 1999 onwards, national development of pharmaceuticals will be industry-led. In connection with this, the research directions of government institutions and universities are programmed together with the industries;
- Institute bioequivalence/bioavailability research for generic products;
- Develop a strong and comprehensive R&D program based on natural bioactive compounds;
- Select and review research priorities as an ongoing activity.

The main focus or directive of the MIGHT group is the High Throughput Screening of natural products for pharmaceutical, cosmetics, food and industrial products. The methodologies requested by MIGHT include combinatorial technology, conventional screening and crude product development.

With funds available from the government through MIGHT, several institutions have started conducting or have indicated interest in conducting CC/CT work. These are the following:

Teaching Institutions:

- University of Malaya
- National University of Malaysia
- University Putra Malaysia (an agricultural university)
- Technology University Malaysia

Research Institutes:

- Institute of Medical Research
- Institute of Forest Research
- Institute of Palm Oil Research
- Institute of Technology Research

At present, what Malaysia needs is help from initiatives/groups coordinated by ICS-UNIDO involving developed countries. If these sources can help Malaysia in a more aggressive way, then maybe the development can be sped up. In this workshop, there were apparently very good experts who can help motivate us. The problem in our country at this point is that many are still not aware of what combinatorial chemistry and combinatorial technology are all about. The technology is not fully transferred and not accessible to many researchers. Preliminary work must be done first to establish a CC/CT team. With support from the government and a willingness to go into combinatorial chemistry, the country will be ready to go into CC/CT work.

Current State of Combinatorial Chemistry and Molecular Design in the Philippines: Possibilities for Industrial Applications

Dr. Norma N. Fajardo
Institute of Chemistry, University of the Philippines Los Baños
4031 Los Baños, Laguna, PHILIPPINES

The rapid advances in science and technology have revolutionized the way research is being conducted in many fields of science, chemistry included. Combinatorial chemistry, combinatorial technologies and molecular design (CC/CT/MD) are among the computer-assisted technologies that have emerged in recent years. CC/CT/MD have become powerful tools in the search for new compounds whose activities are desired for varied purposes (Terrett, *et. al.*, 1995). They have had the greatest impact on the drug and biotechnology industries. However, many countries have yet to adopt these novel technologies in their research programs. The Philippines is among those countries.

The State of the Chemical Sciences in the Philippines

Cognizant of the role of science and technology in the development of the country, the National Academy of Science and Technology (NAST) embarked on a project to assess the state of the various sciences in the Philippines. One study in which the author was personally involved focused on the state of the chemical sciences in the Philippines (NAST, 1998). This was used as a major material for this report.

Active areas of research in the country were determined on the basis of:

- Survey of the papers presented over a three-year period (1995-1997) during the National Chemistry Congress (Table 1), and
- Survey of research activities in the leading academic institutions in the country (Table 2)

Both surveys show that the most active areas of research are analytical/environmental chemistry and the search for bioactive compounds for therapeutic purposes, followed by agricultural chemistry/biochemistry/molecular biology and materials science/polymer chemistry. The study also takes note of the fact that a good number of the papers surveyed are based on dissertations conducted outside the Philippines.

The NAST study also shows that no work is currently being done in the Philippines involving CC/CT/MD. The premier biotechnology institution in the country, the National Institute of Molecular Biology and Biotechnology (BIOTECH), does not have any project involving these new technologies (personal communications). Neither does the Department of Science and Technology (DOST) nor any of its agencies (personal communications).

Table 1. Research topics of paper presented during the Chemistry Congress (1995-1997)

Field in Chemistry	Number of Papers		
	1995	1996	1997
Analytical	16 (21)*	10 (26)	21 (20)
Environmental	20 (26)	6 (15)	6 (6)
Inorganic			5 (5)
Physical	2 (3)		3 (3)
Organic	5 (7)	2 (5)	6 (6)
Materials/Polymer	11 (14)	6 (15)	13 (13)
Natural Products	12 (16)	12 (31)	31 (30)
Agricultural/Biochemistry/Molecular Biology	10 (13)	3(8)	18 (17)

** Percent of total number of papers within a column*

Table 2. Topics of Local Researches in the Chemical and Allied Sciences

Topic	Institutions Involved*
Analytical/Environmental	
Analysis of metals, other substances in environmental waters	UPD, UPLB, AdMU
Development of biosensors	UST, UPD
Electrochemistry	UPD, UPLB
Organic (Synthesis, Cocochemicals)	UPD, UPLB
Materials Science/Polymer Chemistry	
Liquid crystals	UPD
Polymeric membranes for ultrafiltration	UPLB
Biodegradable polystyrene	AdMU
Polyurethane production from coconut oil	UPD
Natural Products Chemistry	UST, UPD, UPLB, AdMU, DLSU
Agricultural Chemistry/Biochem/Molecular Biology	
Towards increased food production	UPLB, UPD, PCA
Utilization of bioresources and waste products	UPLB, DOST, UST, AdMU, DLSU

**UPLB- University of the Philippines Los Baños*
UPD- University of the Philippines Diliman
UST- University of Santo Tomas
DLSU- De La Salle University
AdMU- Ateneo de Manila University
PCA- Philippine Coconut Authority

CC/CT/MD are relatively new technologies. For a developing country like the Philippines, adopting novel strategies could be a very slow process. Funding is a very real problem. It is especially true for projects requiring high initial inputs in terms of facilities and equipment. In the Philippines, research activities are largely funded by the government, followed by international funding agencies. The contribution of the private sector is very minimal (Table 3).

Table 3. Sources of Funding for Researches in the Chemical Sciences in the Philippines

1. Government
   Part of Core Budget
   DOST and its agencies
   - Philippine Council for Advanced Science and Technology Research and Development (PCASTRD)
   - Philippine Council for Industry and Energy Research and Development (PCIERD)
   - Philippine Council for Agriculture, Forestry and Natural Resources Research and Development (PCARRD)
   - Philippine Council for Health Research and Development (PCHRD)
   - Philippine Council for Aquatic and Marine Research and Development (PCAMRD)
2. Private
   United Laboratories (UNILAB)
3. International
   USAID
   International Foundation for Science
   International Atomic Energy Commission
   Japan Society for the Promotion of Science
   Japan International Cooperation Agency

Another possible reason for the scarcity of efforts to adopt CC/CT/MD is the poor linkage between the academe and industry. In spite of the recognized need to develop greater collaborations between the

two, such a status has not been attained in the Philippine setting. Rivera (1997) estimates that on a scale of 1-10 (1- negligible; 10-high), academe-industry links average 4.75 (Table 4).

Possibilities for Industrial Applications

The possibilities for industrial applications of CC/CT/MD in the Philippines are very promising. Current research activities (Tables 1 and 2) can very easily incorporate these nascent technologies in their programs with the industry as the direct beneficiary of their outputs. For example, the search for bioactive substances of various origins (plant, animal, bacterial, etc.) stand to benefit from these novel approaches. The resulting technology can then be passed on to industry for possible commercialization.

Table 4. Academe-Industry Links (Philippines)

	Estimated Level (10 pt scale)*
Manpower Source	
Graduates (employees)	3-6
Professionals (faculty)	3
Internship	3
Company-sponsored scholarships	4
Conferences, workshops and seminars	6
Laboratory/analytical services	2
Consultancy	2
Industry directed research	1
University grants	1
Sponsorship to special events	3

*1-negligible; 10-high
From Rivera (1997)

In the last few years, selected universities in the country namely UPD, UPLB, AdMU, DLSU and the Industrial technology Development Institute (ITDI) of DOST have received massive support for the acquisition of facilities and equipment from the DOST's World bank-Engineering and Science Education Project (WB-ESEP). In addition, WB-ESEP has also granted scholarships and a number of its fellows are in laboratories outside the country working in high technology areas. This should enable these institutions to adopt new technologies in their research programs. The first phase of this project has just come to a close. Negotiations are under way for a second phase. If ever this materializes, plans are for this second phase to focus on activities that will foster greater academe-industry linkage.

The Commission on Higher Education (CHED), the government agency that oversees tertiary education in the Philippines, in a resolution which took effect in June 1998, has designated as centers of excellence (COE) in chemistry seven academic institutions in the country. This is another development that could help accelerate the adoption of novel approaches in scientific research. To carry out this mandate, said institutions were given financial assistance to further strengthen their capabilities in research and instruction, particularly at the graduate level. The Institute of Chemistry at UPLB, for instance, plans to set up new facilities in computational chemistry as part of its work plan as a COE.

Better linkage between academe and industry are starting to emerge. For example, United Laboratories (UNILAB), UP Diliman and UP Los Baños have joined up in a drug development project which uses asymmetric synthesis and kinetic resolution in the preparation of racemic switches, the single enantiomer formulation of traditional racemic formulations (personal communications).

In the NAST study, special note is taken of the central role of the chemical industry to the manufacturing sector and the extensive upstream linkages existing. A few of those linkages are

- agriculture (fertilizers and pesticides)
- electronics (plastics, industrial gases)

- construction (paints, coatings, adhesives, cement additives, plastics and sealants)
- metals and engineering (foundry resins, industrial gases)
- packaging (plastics, adhesives, surface coatings, inks)
- health care (drugs and pharmaceuticals)

It is envisioned that with these linkages, advances in chemistry and the chemical industry will inevitably have a great impact on the growth and development of the manufacturing sector.

In conclusion, while it is a fact that CC/CT/MD have not made their impact on chemical research in the Philippines, conditions for their adoption in the near future are very favorable.

References:

Terrett, N.K., Gardner, M., Gordon, D.W., Kobylecki, R.J. and Steele, J., *Combinatorial synthesis- the design of compound libraries and their application to drug discovery. Tetrahedron*, 51, 30, 8135-8173, 1995.

National Academy of Science and Technology. 1998. Assessment of the State of the Chemical Sciences in the Philippines.

Rivera, A.D., 1997. Strengthening linkages between industry, government and academe with strategic technology management. Paper presented during the 13th Philippine Chemistry Congress, Puerto Princesa City, Palawan.

The Status of Research on Combinatorial Chemistry/ Combinatorial Technology in Thailand*

This paper will outline the research status and priorities in Thailand, the different organizations that play important roles in research in the country, and the possible role of our government institution (the Department of Medical Sciences) in the development of combinatorial chemistry/combinatorial technology in Thailand. Some results from the medicinal plants research institute which can be a possible take off for combinatorial approaches will also be mentioned.

Status of research in Thailand

The Thai Government policy on the development of science and technology are the following:

- Promote the use of technology as a tool to boost the competitiveness of the manufacturing sector by emphasizing the development of appropriate technologies that are consistent with the workers' capacity to learn, while bearing in mind the economic returns from adoption of new technology. Moreover, the government shall promote the methodical development of technology to reduce imports of and dependency on foreign technology.
- Promote private sector participation in the development of science and technology and technology transfer, by granting tax privileges and promoting investment in research and development, so that research is consistent with the demands of business establishments.
- Continuously expedite the production and development of personnel in science and technology, as well as encourage the private sector competitiveness in international markets.

There are three organizations that play important roles in scientific research, namely: The Science Society of Thailand, Thailand's National Science and Technology Development (NSTDA) and the National Research Council of Thailand (NRCT).

The Science Society of Thailand

The Science Society of Thailand was founded on 27 January 1948 under the Patronage of His Majesty the King in order to facilitate the exchange of knowledge among the scientists. The Society plays the roles of initiator, disseminator and supporter of several scientific activities through the academic sections in order to promote science education and scientific research under the policy "Scientists and Technologists are the Brains of the Country".

The Society has launched several programs and activities such as

- Outstanding Scientist Award.
- Senior Scientist Award.
- Best Science Teacher Award.
- National Science Week Fair (yearly, 18-24 August).
- Organization of seminars, workshops and meetings in collaboration with other local and foreign agencies/associations,
- Publication of the Journal of the Science Society of Thailand.
- Organization of the Annual Congress on Science and Technology of Thailand.
- Scientific services to several national agencies, such as the Thai Industrial Standards Institute, National Youth Commission, Council of Scientific and Technological Associations of Thailand, National Forum of the Deans of Faculties of Science, and various universities.

---

* Text prepared by Dr. Cleofé A. Calanasan based on transcription of the presentations of *Dr. Duangchan Suprasert* (Division of Food) and *Dr. Angkana Herunsalee* (Medicinal Plant Research Institute), Department of Medical Sciences, Ministry of Public Health, Nonthaburi 11000, THAILAND

Thailand's National Science and Technology Development Agency (NSTDA)

NSTDA offers funding to support RD&E in areas covered by its three national research centers:
- National Center for Genetic Engineering and Biotechnolgy (BIOTEC)
- National Metal and Materials Technology Center (MTEC)
- National Electronics and Computer Technology Center (NECTEC)

In 1995, funds were awarded to 129 projects for 130.8 million baht. At present, there are 69 research projects for 130.8 million baht. The agency also conducts policy research and provides recommendations to the government on scientific and technological issues impacting Thai society. NSTDA is proposing Thailand as the site of APEC's Center for Technology Foresight in collaboration with agencies in Japan, Australia and other countries.

National Research Council of Thailand (NRCT)

NRCT was inaugurated by the National Research Council Act B.E. 2502 (1959), designated to be the government's technical adviser or the National Council of Science. The Council's function is to recommend the Prime Minister and the Cabinet research policy and development in assigned areas. Since Thailand is an agricultural country, the technologies needed for national agriculture are as follows:
- Production technology.
- Management technology.
- Technology for better utilization of raw materials.
- Technology for natural resource conservation.

The research in the field of chemistry, biology, biochemistry and pharmacology is usually done in the university and government institutes according to one's interest and the budget supply. At present, cooperation between research institutes is still lacking, and research is limited due to insufficient number of qualified researchers, modern equipment and technology.

The Department of Medical Sciences: Organization responsibilities

The department where we belong, the Department of Medical Sciences (DMSc) was established on 10 March 1942 when the Ministry of Public Health was originally founded. It is comprised of 4 Departments, namely; University Bureau, Department of Medicine, Department of Health, and Department of Medical Sciences. Initially, DMSc was responsible for controlling the quality of drugs, production of vaccines, medical diagnosis and analysis of food. For the past 40 years the structure has been re-organised several times to improve the efficiency as well as to expand its responsibilities in serving the public nationwide. The last improvement was in February 1990.

The responsibilities of DMSc are as following:

1. Services
- Laboratory examination and diagnosis in bacteriology, mycology, virology, immunology, parasitology, hematology, clinical chemistry and AIDS.
- Analysis of pharmaceutical products, materials intended for medicinal use and containers for sterile products as well as all kinds of narcotics and psychotropic substances. Supervising the local drug manufacturers on good laboratory practices.
- Identification of toxic substances and analysis of cosmetics, medical equipment, toxic household products and environment hazardous.
- Analysis of food, beverages, water, food containers and food for export to ensure the consumer safety.
- Radiation protection in clinical and industrial uses of radiological equipment.
- Conducting quality assurance of health laboratories in hospitals.

2. Research

- Research in the field of Medical Sciences such as virology, bacteriology, biotechnology, mycology, parasitology, immunology, hematology, clinical chemistry, medical entomology, pharmacology, pharmaceutical chemistry, toxicology and radiology.
- Research on medicinal plants, for example, synthesis of principle substances, structure elucidation of active components from medicinal plants and clinical study as drug.
- Research and production of biological products used for diagnosis and immunization.
- Quality control of biological products produced locally and imported.
- Support research activities by providing experimental animals (both normal and specific pathogen free).

3. Training

- Organizing training courses in different technical areas for both inside officers and outsiders (from other government offices and private sectors).
- Conducting on the job training for certain analysis technique areas for both domestic and international levels.

4. Others

- Acts as WHO National Influenza Center, WHO National Institution for Viral Hepatites, WHO Collaborating Center For Rabies Diagnosis, WHO Collaborating Center for Salmonella and Shigella, WHO National Phage Typing Center, WHO collaborating Center for Evaluation of Pesticides and Applicators, and WHO Collaborating Center for Essential Drug, WHO Regional Reference Laboratory for Polio Diagnosis in Southeast Asia.
- Expansion and development of botanical garden for collecting, cultivating and breeding of medicinal plants and serving as information center and distribution of medicinal plant seeds.
- Performing the duties according to the Act for Pathogenic Microorganisms and Animal Toxins B.E. 2525
- Cooperation in technical aspects with other government authorities and international agencies.

The organization of DMSc are as follows:
DMSc is composed of 8 divisions, 2 institutes, 2 offices and 12 regional Medical Sciences Centers namely:

- Office of the Secretary.
- Div. of Plan and Technical.
- Div. of Biological Products.
- Div. of Cosmetics and Hazardous substances.
- Div. of Radiation and Medical Instrument.
- Div. of Drug Analysis.
- Div. of Narcotics.
- Div. of Food.
- Div. of Food-for-Export.
- Medicinal Plants Research Institute.
- National Institute of Health.
- Office of DMSc Quality Assurance.
- Regional Med.Sciences Center (Chonburi, Nakhon Ratchasima, Phitsanulok, Ubonratchathani, Khonkaen, Chiang Mai, Songkhla, Trang, Chiang Rai, Samut Songkhram, Suratthani, Udonthani).

The Medicinal Plants Research Institute (MPRI)

One of the institutes that can play an important role on research in combinatorial chemistry/combinatorial technology in Thailand is the medicinal plants research institute (MPRI). The role of combinatorial chemistry and technology, especially in the field of natural products, in the development of the country is well recognized. The fact that research work in this field might be applied in public health, agriculture and industry would enhance the value of natural products which are abundant in Thailand.

MPRI is involved in several research areas as shown in figure 1. Before a medicinal plant or plant products can be used commercially in hospitals, its botany, phytochemistry and pharmacological

properties are first determined. Organic synthesis of interesting biologically active compounds are carried out, and their pharmacokinetic and toxicological properties determined. Clinical trials are then implemented in coordination with the Ethics Committee before they are approved for use. Once a drug formulation is developed, the institute advertises the results to the government pharmaceutical organization as well as to local pharmaceutical companies who might be interested in marketing the product.

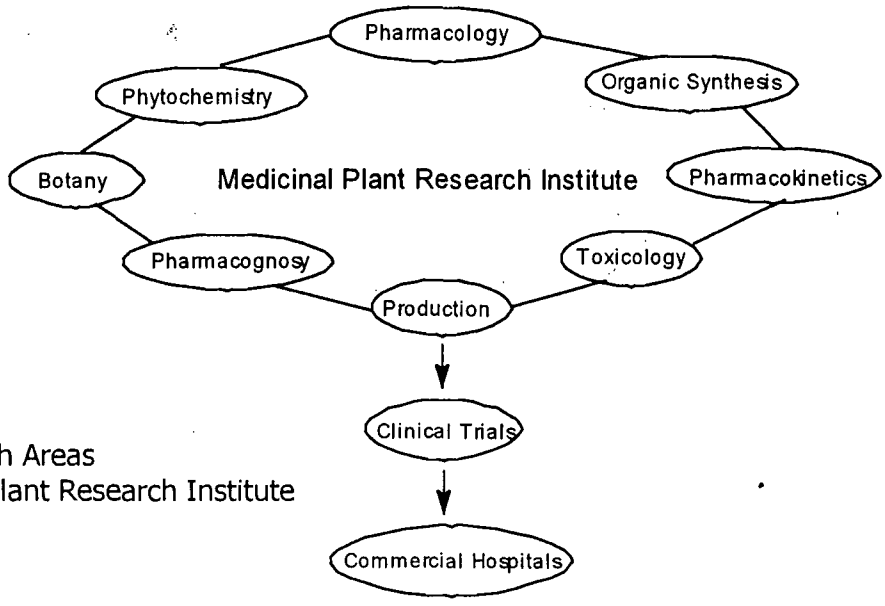


Figure 1. Research Areas
at the Medicinal Plant Research Institute

One of the plants that had been studied in the institute is *Andrographis paniculata* (Burm.f.) *Nees*. The plant has been used as antipyretic, in the treatment of sore throat, treatment for pyrogenic infection and ulcerous diseases in oral cavity, as well as antidiarrheal and antidysentery. The biologically active compounds isolated are the andrographolide and its derivatives (figure 2).

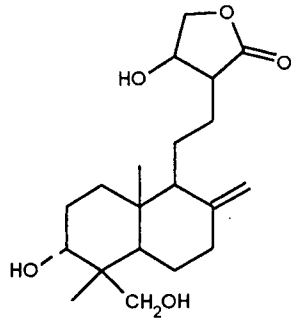


Figure 2. Andrographolide

Other biologically important compounds which have also been studied in collaboration with other research groups in universities are those which act as insect repellants like (+) citronellol (from *Cymbopogon* sp.), azadirachtin (from *Azadirachta indica A. Jussa*) and eugenol. These different compounds and other compounds which have novel skeletal structures can probably be used as scaffolds for combinatorial synthesis.

In Thailand, it will be better to start combinatorial chemistry/combinatorial technology in research institutes and universities since it is difficult for local industries to start combinatorial research. Once the research institutes have done some CC/CT work, they can then collaborate with the private industries. At present, only one government agency is doing CC/CT work in collaboration with different private companies like Aji-no-moto and that is BIOTEC (National Center for Genetic Engineering and Biotechnolgy). However, with more researchers becoming more aware with CC/CT, and with such diverse natural products to start with, the development of CC/CT in Thailand will be forthcoming.

FINAL REPORT

SOUTHEAST ASIAN REGIONAL WORKSHOP ON "COMBINATORIAL CHEMISTRY AND COMBINATORIAL TECHNOLOGIES"

19-23 APRIL 1999
LOS BAÑOS, LAGUNA, PHILIPPINES
UNIDO PROJECT No. TF/GLO/96/105
CONTRACT No. 99/086

Objectives

The major objectives of the workshop are as follows:

- to build awareness among researchers, technologists and academicians on the development of combinatorial chemistry and combinatorial technologies (CC/CT).
- to update the participants on the principles and industrial applications of CC/CT.
- to evaluate possible initiatives (as follow-up projects and feasibility studies) of industry and academe in Southeast Asia regarding CC/CT development and industrial implementation with the focus on the application of CC/CT to natural products research and exploitation.
- to set up a regional ICS-UNIDO network on Combinatorial Chemistry and Combinatorial Technologies.

Organization

This workshop was jointly organized by the International Centre for Science and High Technology (ICS-UNIDO) and the Institute of Chemistry, University of the Philippines, Los Baños (IC-UPLB). The local members of the organizing committee were:

Dr. Norma N. Fajardo (Chair)
Dr. Cleofe A. Calanasan (Head, Technical Committee)
Mr. Roy Roberto L. Gerona (Head, Secretariat)
Dr. Maxima E. Flavier (Head, Ways and Means Committee)
Ms. Maria Cecilia D. de Mesa (Head, Publicity and Physical Arrangements Committee)

The scientific committee was composed of local and international scientists who advised organizers about the content of the scientific programme. This committee included the following:

Prof. Stanislav Miertus (ICS-UNIDO, Italy)
Dr. Pierfausto Seneci (GlaxoWellcome, Italy)
Dr. Giorgio Fassina (Tecnogen, Italy)
Dr. Cleofe A. Calanasan (IC-UPLB, Philippines)

Dates

The workshop was held at the Continuing Education Center, University of the Philippines Los Baños (UPLB), Laguna, Philippines 19-23 April 1999.

Funding

The total contribution from UNIDO was US$ 30,000. The Institute of Chemistry, UPLB provided the manpower support, conference facilities, documentation services and communication expenses.

Lecturers

A total of 6 lecturers from different institutions and countries participated in the workshop. Another lecturer, Dr Andrea Missio from GlaxoWellcome Italy, was unable to attend due to ill health, so the papers he prepared were presented by Dr. P.F. Seneci.

The lecturers who presented the state-of-the-art in combinatorial chemistry and technologies were mainly from industrial institutions. The lectures were supported by discussions on different case

studies highlighting the applications of CC/CT. The economics and patenting issues regarding CC/CT were also covered.

Lecturer	*Institution*	*Country*
Prof. Stanislav Miertus	ICS-UNIDO	Italy
Dr. Pierfausto Seneci	GlaxoWellcome	Italy
Dr. Giorgio Fassina	TECNOGEN	Italy
Dr. Anneliese Appleton	Molecular Simulations P/L	Australia
Dr. Lucia Carrano	Biosearch	Italy
Dr. Franciscus T. M. van Amsterdam	GlaxoWellcome	Italy

The lecturers' accommodations, living expenses and airfare (except Prof. S. Miertus) were fully sponsored with the workshop budget.

Participants

The workshop was intended for participants from universities, research institutions, government institutions, and private companies engaged in the petrochemical, agrochemical and pharmaceutical industries. Participants involved in R and D and industrial processes were given preference. Applications were solicited from institutions throughout the Southeast Asian region via mail, e-mail, fax and telephone calls. The local committee with the assistance of ICS selected the participants.

Twenty seven (27) participants out of the 28 selected attended the workshop. The lone invitee from Indonesia failed to attend. Details of the geographical and institutional distributions of the participants are presented below.

Country	*Number of Participants*
Hongkong	**1** (academe)
India	**1** (research institute)
Malaysia	**1** (academe)
Philippines	**18** (7 from industry, 2 from govt institutes, 1 from research institute, 8 from academe)
Taiwan	**1** (industry)
Thailand	**4** (1 from academe, 3 from govt institutes)
Vietnam	**1** (research institute)

A complete list of participants is given on pages 117-119.

Materials Distributed

Lecturers and participants were provided with a workshop bag containing the souvenir program, note pads, pens, directory, Philippine map and computer diskettes. Copies of lecture materials were also given to each participant.

Opening Ceremony

The opening ceremony was attended by Prof. Stanislav Miertus (ICS-UNIDO Area Coordinator), the lecturers and participants. Invited guests to this ceremony included Dr. Francisco Peñalba (Vice Chancellor for Community Affairs, UPLB), Dr. Ma. Cristina D. Padolina (Chancellor, UP Open University), Dr. Carol Yorobe (Assistant Secretary, Departmentt of Science and Technology), Dr. Pacifico C. Payawal (Dean, College of Arts and Sciences, UPLB), Dr. Ernesto J. del Rosario (Director, Institute of Chemistry, UPLB) and Engr. Jun Montes (Representative, Chemical Industries Association of the Philippines, SPIK).

Meetings

On 21 April 1999, a meeting was held among Mr. Roberto di Virgilio (Commercial Attache of the Italian embassy), Prof. Stanislav Miertus, Dr. Lucia Carrano, Dr. Giorgio Fassina and Dr. Pierfausto

Seneci. Mr. Di Virgilio showed interest in the joint programme between ICS-UNIDO and the Institute of Chemistry, UPLB, and expressed his support for the plan of a CC/CT network in the Southeast Asian Region. Another meeting was held on 23 April 1999 between Prof. S. Miertus and Ms. Betty Koreh (UNIDO Program Officer in the Philippines). The latter expressed her strong interest and support for any follow-up activities to the workshop.

Workshop Programme

The workshop was organized as lectures and discussions from 9.00 AM to 18.30 PM. The programme is given on pages 114-115.

Social, Educational and Cultural Activities

The following social events were attended by the lecturers, participants and organizers during the workshop:

A visit to the flower and garden show in the university after the lectures on 19 April.

Swimming in a local hot spring resort in the evening of 19 April.

A banquet at the ANEST Tower on 21 April 1999 which featured Philippine folk dances.

A trip to nearby towns of Los Baños, Laguna in the afternoon of 23 April 1999.

Prof. Miertus, Dr. van Amsterdam and Dr Fassina took a boat ride in Laguna lake, the biggest body of freshwater in the Philippines on 20 April.

On the 24th of April, the lecturers and international participants also visited Tagaytay which overlooks Taal volcano, the smallest active volcano in the world.

Closing Ceremony

The round table discussion on future actions and closing ceremony were attended by Ms. Betty Koreh (UNIDO Program Officer), Prof. S. Miertus (ICS-UNIDO Area Coordinator), Dr. Ernesto J. del Rosario (IC-UPLB director) and Dr. Alan Feranil (Philippine Council for Health Research and Development, DOST representative).

Assessment

The workshop was assessed by the participants using the questionnaire provided by ICS-UNIDO. The participants' evaluation of the overall programme organization is from very good to excellent. All the participants said that they benefited from the workshop and will recommend others to attend a similar activity in the future.

The participants opined that the lectures ranged from very good to excellent and the number of days spent was just right. The evaluation of the training facilities/accommodation and the food ranged from good to excellent.

Results and Follow-up

At the start of the workshop, most of the participants had little or no knowledge of combinatorial chemistry/combinatorial technologies. Based on the results of the evaluation, the participants felt that the workshop has enlightened them on what CC/CT is all about. They found all the topics useful and are very enthusiastic for a follow-up to the workshop which would have more hands-on activities. A number of the participants have even discussed going into collaborative work particularly in the search for bioactive natural products which will be good scaffolds for combinatorial chemistry.

At the last round table discussion, the setting up of a regional network for combinatorial chemistry in Southeast Asia was proposed. This will facilitate regular interactions among the different institutions that will be carrying out CC/CT in the region. A program proposal for the establishment of a CC/CT network for Southeast Asia is now in preparation.

Since most of the objectives of the workshop were met, it is deemed that the workshop has been a success.

PROGRAMME

Monday, 19 April

9:00 AM	**Opening Ceremonies**	
	Philippine National Anthem	Raymond B. Monterey
	Welcome Remarks	Dr. Francisco Peñalba *Vice Chancellor for Community Affairs, UPLB*
	Overview of ICS-UNIDO Programmes and Activities	Prof. Stanislav Miertus *Area Coordinator, ICS-UNIDO*
	Introduction of Workshop Speakers	Dr. Cleofe A. Calanasan, *Head, Technical Committee*
	Introduction of Participants	Roy Roberto L. Gerona *Head, Secretariat*
	Closing Remarks	Dr. Ernesto J. del Rosario, *Director, IC, CAS, UPLB*

Dr. Evamarie Capareda
Master of Ceremonies

10:30 AM	C O F F E E B R E A K	
11:00 AM	Combinatorial Chemistry and Combinatorial Technologies: An Overview	*Prof. S. Miertus, Dr. G. Fassina*
12:00 NN	Solid Phase Synthesis	*Dr. P. F. Seneci*
12:45 PM	L U N C H	
2:00 PM	Case Study 1: Solid-Phase Synthesis	*Dr. P. F. Seneci*
2:30 PM	Round Table Discussion on the Morning Session	
3:00 PM	Combinatorial Synthetic Libraries: Design and Formats	*Dr. P. F. Seneci*
4:00 PM	C O F F E E B R E A K	
4:30 PM	Solid Phase Synthetic Libraries	*Dr. P. F. Seneci*
5:30 PM	Case Study 2: Solid Phase Synthetic Libraries	*Dr. P. F. Seneci*
6:00 PM	Round Table Discussion on the Afternoon Session Moderator: Dr. Evelyn B. Rodriguez	

Tuesday, 20 April

9:00 AM	Structure Determination of Positives from Solid Phase Libraries	*Dr. P. F. Seneci*
10:00 AM	Case Study 3: Structure Determination of Positives from Solid Phase Libraries	*Dr. P. F. Seneci*
10:30 AM	C O F F E E B R E A K	
11:00 AM	Solution Phase Synthetic Libraries	*Dr. P. F. Seneci*
12:00 NN	Case Study 4: Solution Phase Synthetic Libraries	*Dr. P. F. Seneci*
12:30 PM	Round Table Discussion on the Morning Session Moderator: Dr. Evamarie P. Capareda	
1:00 PM	L U N C H	
2:30 PM	Purification and Quality Control of Libraries	*Dr. P. F. Seneci*
3:30 PM	Solid Phase Methodologies, Peptide Library Generation and Screening	*Dr. G. Fassina*
4:30 PM	C O F F E E B R E A K	
5:00 PM	Case Study 5: Identification of a Synthetic Ligand for the Affinity Purification of Antibodies	*Dr. G. Fassina*
5:30 PM	Biological Libraries	*Dr. G. Fassina*
6:30 PM	Round Table Discussion on the Afternoon Session Moderator: Mr. Roy Roberto L. Gerona	

Wednesday, 21 April

9:00 AM	Computation Methods in Library Design	*Dr. A. Appleton*
10:00 AM	Software for Combinatorial Technologies	*Dr. A. Appleton*
11:00 AM	C O F F E E B R E A K	
11:30 AM	Case Study 6: Computational Methods in Library Design	*Dr. A. Appleton*
12:00 NN	Round Table Discussion on the Morning Session Moderator: Dr. Evelyn B. Rodriguez	
12:30 PM	L U N C H	
2:00 PM	Bioassay Design	*Dr. F. V. Amsterdam*
3:00 PM	LTS, MTS, and HTS: Implications for Drug Discovery	*Dr. F. V. Amsterdam*

4:00 PM	C O F F E E B R E A K	
4:30 PM	Case Study 7: Low Throughput Screening	*Dr. F. V. Amsterdam*
5:00 PM	Case Study 8: High Throughput Screening	*Dr. F. V. Amsterdam*
5:30 PM	Development of Combinatorial Chemistry and Combinatorial Technologies in Hongkong	*Dr. P. K. Wong*
6:00 PM	Round Table Discussion on the Afternoon Session Moderator: Dr. Evelyn B. Rodriguez	
7:30 PM	B A N Q U E T	ANEST Towers

Thursday, 22 April

9:00 AM	Natural Products: An Overview	*Dr. L. Carrano*
10:00 AM	Natural Products as Sources of Relevant Drugs	*Dr. L. Carrano*
11:00 AM	C O F F E E B R E A K	
11:30 AM	Case Study 9: Natural Products and Drug Discovery	*Dr. L. Carrano*
12:00 NN	L U N C H	
2:00 PM	Case Study 10: Combinatorialization of Natural Products	*Dr. P. F. Seneci*
3:00 PM	Round Table Discussion on Natural Products and Combinatorial Technologies	
3:30 PM	Patenting Issues in Combinatorial Technologies	*Dr. P. F. Seneci*
4:30 PM	C O F F E E B R E A K	
5:00 PM	Economics of Combinatorial Chemistry and Technology	*Dr. G. Fassina*
6:00 PM	Round Table Discussion on Patenting Issues and Economics of Combinatorial Technologies Moderator: Dr. Cleofe A. Calanasan	

Friday, 23 April

8:30 AM	**Country Reports**:	
	• Combinatorial Chemistry in Malaysia	*Dr. M. Ali Mohd*
	• The Status of Research on CC/CT in Thailand	*Dr. A. Herunsalee*
	• Plastic Industries Degradable Polymers in Taiwan	*Mr. P. C. Lin*
	• An Overview of Water Quality of Some Industrializing Craft Villages in the Delta of Northern Vietnam and their Impact on Natural Waterbodies	*Dr. Ho Than Hai*
	• Combinatorial Chemistry and Technologies in India	*Dr. S. Murty*
	• Status of CC/CT in the Philippines	*Dr. N. N. Fajardo*
11:00 AM	C O F F E E B R E A K	
11:30 AM	Round Table Discussion: Future Actions Moderator: Dr. Cleofe A. Calanasan	
12:00 NN	**Closing Ceremonies**	
	Workshop Assessment and Summary	*Dr. P. F. Seneci* *Dr. C. A. Calanasan*
	Response from Participants	*Dr. M. Ali Mohd* *Ms. R. Nolasco* *Ms. E. Suplido*
	Distribution of Certificates of Attendance	*Prof. S. Miertus*
	Closing Remarks	*Ms. B. Koreh* *Dr. N. N. Fajardo*
1:00 PM	**E N D O F W O R K S H O P** L U N C H	

DIRECTORY

RESOURCE PERSONS

Dr. ANNELIESE APPLETON
Scientific Consultant
Molecular Simulations P/L
Level 1, 65 York Street
Sydney NSW 2000
AUSTRALIA
Phone: 61-2-9299-9533
Fax: 61-2-9262-2860
E-mail: liese@mpx.com.au

Dr. FRANK Th.M. VAN AMSTERDAM
Head
Scientific Computing & Lead Generation
Glaxo Wellcome Medicines Research Centre, Verona
ITALY
Phone: 39-045-9218837
Fax: 39-045-9218153
E-mail:fa8935@glaxowellcome.co.uk

Dr. GIORGIO FASSINA
Director
Biopharmaceuticals
Tecnogen S.C.P.A.
Parco Scientifico
81015 Piana di Monte Verna
ITALY
ICS-UNIDO Scientific Consultant on CC/CT
Phone: 39-0823-612-214
Fax: 39-0823-612-230
E-mail: fassina@tecnogen.it

Dr. LUCIA CARRANO
Senior Scientist
Microbial Technologies
Biosearch Italia spa
ITALY
Phone: 39-029-647-4432
Fax: 39-029-647-4238
E-mail: lcarrano@biosearch.it

Dr. PIERFAUSTO SENECI
Head
Lead Discovery Department
GlaxoWellcome Medicines Research Centre
Via Fleming 4
37100 Verona, ITALY
ICS-UNIDO Scientific Consultant on CC/CT
Phone: 39-045-9218884
Fax: 39-045-9218196
Email:pfs37082@glaxowellcome.co.uk

Dr. ANDREA MISSIO
Combinatorial Technologies
GlaxoWellcome, Verona
ITALY
Present address:
Hoechst Roussel Vet GmbH
Research Pharmaceuticals: Chemistry
Althanstrasse 14
A-1090 Vienna, AUSTRIA
Phone: +43 (0) 1 3195082-22
Fax: +43 (0) 1 3195082-40
E-mail: amissio@hrvet.com

Prof. STANISLAV MIERTUS
ICS Area Coordinator
Pure and Applied Chemistry
AREA Science Park, Building L2
Padriciano, 99
34012 Trieste
ITALY
Phone: 39-040-9228116
Fax: 39-040-9228115
E-mail: miertus@ics.trieste.it

PARTICIPANTS

DANILO N. AÑOS
General Manager
Biogenerics Pharmaceuticals Corp.
1052 Metro Properties Bldg., EDSA,
Magallanes Village,
Makati City
PHILIPPINES
Phone: 63-2-551-0974
Fax: 63-2-834-2375
E-mail: bgen@iconn.com.ph

CHANTANA AROMDEE, PhD
Department Head (FPS)
Department of Pharmaceutical Chemistry, Faculty of Pharmaceutical Sciences
Khon Kaen University,
Khon Kaen 40002
THAILAND
Phone: 66-043-236907
66-043-241243
E-mail: chaaro@kkul.kku.ac.th

EDWIN ALCANTARA, PhD
University Researcher
National Institute of Biotechnology and Molecular Biology
Univ of the Philippines Los Baños
College, Laguna
PHILIPPINES
Phone: 536 2721

EVAMARIE P. CAPAREDA, PhD
Associate Professor
Institute of Chemistry
Univ of the Philippines Los Baños
4031 College, Laguna
PHILIPPINES
Phone: 63-49-536-2220
Fax: 63-49-536-2241
E-mail: epc@chem.uplb.edu.ph

SHARON ROSE P. GARCIA
Science Research Specialist
Phil Council for Health Res and Dev
Department of Science and Technology
Gen. Santos Ave., Bicutan, Taguig,
Metro Manila
PHILIPPINES
Phone: 63-2-837-7535 to 36
Fax: 63-2-837-2924
63-2-837-2942
E-mail: sharon@pchrd.dost.gov.ph

ROMULO D. CONDE
Technical Supervisor
Philippine Coconut R&D Foundation
3/F PCRDF Bldg., Pearl Dr. cor.
Lourdes St.,
Ortigas, Pasig
PHILIPPINES
Phone:63-2-633-8488
63-2-632-9053
Fax: 63-2-633-8032
E-mail:info@pcrdf.org.ph

CHRISTIAN JOSEPH R. CUMAGUN, MSc
Instructor
Department of Plant Pathology
Univ of the Philippines Los Baños
College, Laguna
PHILIPPINES
Phone:63-49-536-4765
Fax: 63-49-536-3534
E-mail:cjrc@mudspring.uplb.edu.ph

RAMON S. DEL FIERRO, PhD
Department Chair
Department of Chemistry
University of San Carlos
Talamban, Cebu City
PHILIPPINES
Phone:63-32-346-1128
Fax: 63-32-344-6715
E-mail:chem@mangga.usc.edu.ph

IVAN-MARCELO A. DUKA, MSc
Assistant Professor
Genetics and Molecular Biology Division
Institute of Biological Sciences
Univ of the Philippines Los Baños
College, Laguna
PHILIPPINES
Phone: 63-49-536-2893

PI CHOU LIN, MSc
Technical Director
USI Far East Corporation
Customer Service Laboratory
3F, No. 1, Lane 147, Ming Sheng E. Rd.
Sec. 2, Taipei 10447
TAIWAN
Phone:886-2-2609-4269
886-2-2609-5270
E-mail:PCLIN@CSL.USIFE.COM.TW

ANGKANA HERUNSALEE, PhD
Medical Scientist
Institute of Medicinal Plant Research,
Department of Medical Science,
Tivanoth Road, Nonthaburi 11000
THAILAND
Ph:66-2-5899850 to 58 ext 9044
66-2-5899866
E-mail:
angkanah@health.moph.go.th

HAI THANH HO, PhD
Department Head
Institute of Ecology and Bioresources
Nghiado-Cua Giay, Hanoi
VIETNAM
Phone: 84-4-843-4288
84-4-836-1196
E-mail: canh@iebr.ac.vn

JOSE O JULIANO, PhD
Vice Chairman and Consultant
Interphil Laboratories, Inc.
3/F Zuelig Bldg., Sen. Gil Puyat Ave.
1265 Makati City
PHILIPPINES
Phone: 63-2-845-7587
Fax: 63-2-893-8036
E-mail:jjuliano@interpharma.com.ph

MOHD ALI MUSTAFA, PhD
Lecturer
Department of Pharmacology,
Faculty of Medicine,
University of Malaya,
Kuala Lumpur 59100
MALAYSIA
Phone: 603-759-4709
603-759-4791
E-mail:
mustafa@medicine.med.um.edu.my

EDNA DC. NACIANCENO
Science Research Specialist
Philippine Council for the Advancement of Science and Technology Research & Dev
Department of Science and Technology
Gen. Santos Ave., Bicutan, Taguig,
Metro Manila
PHILIPPINES
Phone: 63-2-837-7516
Fax: 63-2-837-3468
E-mail: edna@agham.dost.gov.ph

SREE RAMA M. MAGUDULA, PhD
Scientist
Organic Chemistry Division I
Indian Institute of Chemical Technology
Hyderabad
INDIA
Phone:91-40-7170512
91-40-7173378
E-mail:madugula51@hotmail.com

ELMER RICO E. MOJICA
Instructor
Physical Sciences Department
Cavite State University
Indang, Cavite
PHILIPPINES
Phone:63-46-415-0010
Fax: 63-46-415-0012

RAYMOND B. MONTEREY, ChE
Instructor
Institute of Chemistry
Univ of the Philippines Los Baños
4031 College, Laguna
PHILIPPINES
Phone:63-49-536-2220
Fax: 63-49-536-2241
E-mail:arein@chem.uplb.edu.ph

ROSALINA D. NOLASCO
Senior Vice-President
Swiss-Pharma Research Lab
Bo. Diezmo, Cabuyao, Laguna
PHILIPPINES
Phone: 63-49-549-2376
Fax: 63-49-549-7210

ANG LONG BING M. PAO
Supervising Chemist
Quality Assurance / Quality Control Department
Swiss-Pharma Research Lab
Bo. Diezmo, Cabuyao, Laguna
PHILIPPINES
Phone: 63-49-549-2376
Fax: 63-49-549-7210

GEORGE A. SALVILLA, MBA
Plant Manager
YSS Laboratories
1133 United Nations Avenue
Paco, Manila
PHILIPPINES
Phone:63-2-524-0061 to 63
Fax: 63-2-523-6385
E-mail:yss@info.com.ph

JOSE H. SANTOS, PhD
Assistant Professor
Institute of Chemistry
Univ of the Philippines Los Baños
4031 College, Laguna
PHILIPPINES
Phone:63-49-536-2220
Fax: 63-49-536-2241
E-mail:joey@chem.uplb.edu.ph

SUWUNNA SOONTRANON, PhD
Biological and Safety Testing Section
Division of Cosmetics and Hazardous Substances
Ministry of Public Health
Nonthaburi 11000
THAILAND
Phone:66-2-591-5436
66-2-591-5436
E-mail:suwunna@dmsc.moph.go.th

EMMA R. SUPLIDO, MSc
Dean
College of Science
Palawan State University
Tiniguiban Heights
Puerto Princesa City, Palawan
PHILIPPINES
Phone:63-48-434-3405

DOUNGCHAN SUPRASERT, PhD
Division of Food
Department of Medical Sciences
Ministry of Public Health
Nonthaburi 11000
THAILAND
Phone: 66-2-951-0000 ext 9566
66-2-951-1023
E-mail: dchan@dmsc.moph.go.th

ELIZABETH T. TAN, PhD
Department Manager
United Laboratories, Inc.
Mandaluyong City
PHILIPPINES
Phone: 63-2-631-8501 loc 461, 463 or 836
Fax: 63-2-631-9694
E-mail: bethtan@ibm.net

PO-KEUNG WONG, PhD
Professor
Department of Biology
The Chinese University of Hong Kong
Shatin, N.T.
HONG KONG
Phone:852-2609-6383
852-2603-5767
E-mail:pkwong@cuhk.edu.hk

SCIENTIFIC COMMITTEE

PROF. STANISLAV MIERTUS	ICS-UNIDO, Trieste, Italy
DR. PIERFAUSTO SENECI	GlaxoWellcome, Verona, Italy
DR. GIORGIO FASSINA	Tecnogen, Naples, Italy
DR. CLEOFE A. CALANASAN	Univ. of the Phil. Los Baños, Philippines

ORGANIZING COMMITTEE

Institute of Chemistry
University of the Philippines Los Baños
4031 College, Laguna
PHILIPPINES
Phone: 63-49-536-2220
Fax: 63-49-536-2241
E-mail: postmaster@chem.uplb.edu.ph

Chair :	Norma N. Fajardo, PhD, Associate Professor	nnf@chem.uplb.edu.ph
Technical Committee		
Head:	Cleofe A. Calanasan, PhD, Assistant Professor	cac@chem.uplb.edu.ph
Members:	Evelyn B. Rodriguez, PhD, Associate Professor	ebr@chem.uplb.edu.ph
	Jose H. Santos,PhD, Assistant Professor	joey@chem.uplb.edu.ph
Ways and Means Committee		
Head:	Maxima E. Flavier, PhD, Associate Professor	mef@chem.uplb.edu.ph
Memberr:	Ma. Flor A. Navera, University Research Assoc	marif@chem.uplb.edu.ph
Travel and Physical Arrangements		
Head:	Ma. Cecilia D. De Mesa, MSc, Assistant Prof	mcddm@chem.uplb.edu.ph
Members:	Gladys Cherisse J. Completo, MSc, Instructor	gladys@chem.uplb.edu.ph
	Alvin M. A. Altamirano III, ChE, Instructor	mac@chem.uplb.edu.ph
	Maria Elena C. Viernes, MSc, Instructor	mecv@chem.uplb.edu.ph
Secretariat		
Head:	Roy Roberto L. Gerona, MSc, Instructor	rrlg@chem.uplb.edu.ph
	Evamarie P. Capareda, Associate Professor	epc@chem.uplb.edu.ph
	Raymond B. Monterey, Instructor	arein@chem.uplb.edu.ph
	Jose Rene L. Micor, Instructor & Sys Ad	micor@chem.uplb.edu.ph
	Aurea N. Miranda, Instructor	h37anm@chem.uplb.edu.ph
	Nestor U. Soriano, Instructor	jon@chem.uplb.edu.ph