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

INTERNATIONAL CENTRE
FOR SCIENCE AND HIGH TECHNOLOGY

Training Manual

Protocols in Combinatorial Chemistry and Combinatorial Technologies

Version 1.0, August 2000

Scientific editors

Giorgio Fassina and Stanislav Miertus

UNITED NATIONS
INDUSTRIAL DEVELOPMENT ORGANIZATION

Vienna, 2000

Opinions expressed in the present publication do not imply the endorsement of the United Nations Industrial Development Organization (UNIDO) or the International Centre for Science and High Technology (ICS). The mention of company names or commercial products does not imply that any such company or product is endorsed or recommended by UNIDO or ICS.

Numbers appearing in square brackets in the text refer to notes at the end of each chapter, under "references".

ID/ICS.4

UNIDO PUBLICATION
Sales No. E.00.II.B.7
ISBN 92-1-106411-2

Version 1.0, August 2000

Earth, Environmental and Marine Sciences and Technologies
ICS-UNIDO, AREA Science Park, Building L2
Padriciano 99, 34012 Trieste, Italy
Tel: +39-040-9228108, Fax: +39-040-9228136

Contributors

The following persons contributed chapters to *Protocols in Combinatorial Chemistry and Combinatorial Technologies:*

Nicola Arena of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is co-author of chapter 4 "Semi-automatic apparatus for the solid-phase synthesis of peptide libraries".

Nicola Campanile of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is co-author of chapter 4 "Semi-automatic apparatus for the solid-phase synthesis of peptide libraries".

Maria Dani of Tecnogen S.C.p.A., Italy is author of chapter 9 "Growth of phages and bacteria for biopanning experiments".

George Fassina of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is one of the scientific editors of *Protocols in Combinatorial Chemistry and Combinatorial Technologies.*

Valerie J. Gillet of the Department of Information Studies, University of Sheffield, United Kingdom; and of the Lead Discovery Unit, GlaxoWellcome Medicines Research Centre, United Kingdom is co-author of chapter 8 "Designing combinatorial libraries".*

Darren V. S. Green of the Department of Information Studies, University of Sheffield, United Kingdom; and of the Lead Discovery Unit, GlaxoWellcome Medicines Research Centre, United Kingdom is co-author of chapter 8 "Designing combinatorial libraries".*

Maria Marino of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is co-author of chapter 5 "Deconvolution of synthetic peptide libraries by non-covalent adsorption to solid supports".

Stanislav Miertus of the International Centre for Science and High Technology, United Nations Industrial Development Organization, Trieste, Italy is one of the scientific editors of *Protocols in Combinatorial Chemistry and Combinatorial Technologies.*

Andrea Missio of the Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Italy is co-author of chapter 7 "Solution-phase synthesis of a discrete library of heterocycles".

Alfredo Paio of the Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Italy is co-author of chapter 6 "Solid-phase synthesis of substituted amine libraries".

Giovanna Palombo of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is co-author of chapter 10 "Phage selection for biopanning".

Menotti Ruvo of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is author of chapter 2 "Manual synthesis of peptide libraries" and chapter 3 "Peptide library characterization by amino acid analysis and TOF-MALDI mass spectrometry", and co-author of chapter 4 "Semi-automatic apparatus for the solid-phase synthesis of peptide libraries" and chapter 5 "Deconvolution of synthetic peptide libraries by non-covalent adsorption to solid supports".

Angela Scarallo of Biopharmaceuticals, Tecnogen S.C.p.A., Italy is co-author of chapter 4 "Semi-automatic apparatus for the solid-phase synthesis of peptide libraries".

Pierfausto Seneci of the Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Italy is co-author of chapter 6 "Solid-phase synthesis of substituted amine libraries" and chapter 7 "Solution-phase synthesis of a discrete library of heterocycles".

* The authors of chapter 8, Valerie Gillet and Darren Green, wish to acknowledge the contribution of Andrew Leach for his expertise in developing the ADEPT computer program.

Foreword

The United Nations Industrial Development Organization (UNIDO) 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 that deals 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 in Trieste, Italy. The mandate of the International Centre for Science and High Technology (ICS) 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 that utilizes new and environmentally friendly technologies.

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

In the present work programme, the activities of ICS focus on specific sectors in the fields of chemistry, environment, new materials and high technology. In selecting specific subprogrammes and 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 on the harmonization of economic growth and environment conservation and protection, the Area of Pure and Applied Chemistry of ICS has identified priorities in its work programme. The themes described below are of key relevance to economic and industrial development as well as environmental protection.

Catalysis. This is an important scientific and technological area for the development of environmentally friendly chemical processes. In turn, such chemical processes form the basis for cleaner industrial development and are also the key elements in an

industrial prevention approach. New, less pollutant processes, together with the optimization of existing processes, depend to a great extent on the improvement of catalyst performance in the heavy and fine chemical production lines. Such processes, together with catalyst performance, have a direct impact on the quality and quantity of by-products or waste generated.

Environmentally degradable plastics. 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. In addition to recycling, reuse, incineration and composting, new technological developments in environmentally degradable plastics contribute dramatically to tackling environmental issues with regard to specific sectors of plastics use.

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

Combinatorial chemistry, combinatorial technologies and molecular design. Taken together, these have a strong impact on the development of new chemicals (pharma industries, agro-chemicals, new materials). Developing countries need to become familiar with and gain expertise in combinatorial technologies and molecular design in order to help local enterprises remain competitive and economically viable in coming decades. Combinatorial chemistry and combinatorial technologies have a potential influence on industrial growth and on environmental protection. 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 also fall 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 has in its entire history.

Combinatorial methods are not restricted to pharmaceutical applications. Whenever a large number of compounds need to be prepared for testing, this technique can be used. An additional field of application is agricultural research. However, the main emphasis at present is on pharmaceutical research, and most major pharmaceutical companies are active in the field. It is generally accepted that combinatorial methods offer great potential for the lead finding and drug discovery process. The technologies are expected to contribute to reductions in time and cost.

Owing to the 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 promotion and publishing activities. The training courses, workshops and expert group meetings organized by ICS include the following:

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

A training course on "Methodologies, Applications and Economics of Combinatorial Chemistry and Combinatorial Technologies", in Piana di Monte Verna, Italy, 8-19 September 1997, hosted and co-organized by Tecnogen S.C.p.A.

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

An expert group meeting on "Combinatorial Chemistry, Combinatorial Technolo-

gies and Molecular Design", in Trieste, Italy, 15-17 June 1998.

A workshop on "Combinatorial Technologies: Awareness and Familiarization for Decision Makers" and a training workshop on "Methods, Applications and Economics of Combinatorial Chemistry and Combinatorial Technology", in Hyderabad, India, 26-31 October 1998, hosted and co-organized by the Council of Scientific and Industrial Research, Industrial Institute of Chemical Technology.

A workshop on "Combinatorial Chemistry and Combinatorial Technologies", in Buenos Aires, 7-12 December 1998, hosted and co-organized by the Institute of Molecular Biology and Genetic Engineering (INGEBI), University of Quilmes.

A workshop on "Applications of CC/CT in Industry and Exploitation of Naturally Occurring Compounds", in Manila, 19-23 April 1999.

A training course on "Molecular Modelling and Computer Assisted Combinatorial Chemistry", in Trieste, Italy, 5-9 July 1999.

A training course on "Molecular Modelling and Computer-Assisted Combinatorial Chemistry", in Trieste, Italy, 29 November-3 December 1999.

More workshops and training courses are scheduled for the year 2000 and will be held in Italy, Mexico and Slovenia.

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

During the implementation of the programme, we became aware of a lack of suitable literature on combinatorial chemistry, especially with regard to training in the acquisition of laboratory practice and basic skills that are relevant to combinatorial chemistry and combinatorial technologies. This manual has, therefore, been prepared by a group of recognized field experts. It focuses on protocols in various applications of combinatorial chemistry and combinatorial technologies and thus can be used as a training package.

I would like to thank Giorgio Fassina, ICS-UNIDO Scientific Adviser, for his help in the scientific editing of *Protocols in Combinatorial Chemistry and Combinatorial Technologies*. The assistance of Susan Biggin, Paola Volpi and Emanuela Corazzi in the preparation of the publication and in the implementation of the ICS-UNIDO work programme is highly appreciated.

Thanks are also due to several staff members of the Editorial and Composition Section of the United Nations Office at Vienna for their contribution to the publication.

Trieste, April 2000
Stanislav Miertus
ICS Programme Officer
Pure and Applied Chemistry

PLEASE BE AWARE THAT
ALL OF THE MISSING PAGES IN THIS DOCUMENT
WERE ORIGINALLY BLANK

Contents

Page

List of tables

List of schemes

List of figures

Abbreviations

2D	two-dimensional
3D	three-dimensional
AcH	acetic acid
Acm	acetamidomethyl
Ala	alanine
AM-PS	aminomethylpolystyrene
Amu	atomic mass unit
Arg	arginine
Asn	asparagine
Asp	aspartic acid
Atm	atmosphere
BF_3	boron trifluoride
Boc	t-butyloxycarbonyl
BSA	bovine serum albumin
CE	capillary electrophoresis
CH_3CN	acetonitrile
CH_3OH	methanol
Cys	cysteine
DCC	dicyclohexylcarbodiimide
DCM	dichloromethane
DCR	divide-couple-recombine
DCU	dicyclohexylurea
DIBAL	diisobutylaluminium hydride
DIC	diisopropyl carbodimide
DIEA	di-isopropyl-ethylamine
DMAP	N,N-dimethylaminopyridine
DMF	N,N-dimethylformamide
DMSO	di-methylsulphoxide
ELISA	enzyme linked immunosorbent assay
ESI	*electrospray ionization*
Et_2O	ethyl ether
FAB	fast atom bombardment
FDA	Federal Drug Administration
Fmoc	9-fluorenylmethoxycarbonyl
Fmoc-Cl	9-fluorenylchloroformate
FT-IR	fourier-transform infrared spectroscopy
Gln	glutamine
Glu	glutamic acid
Gly	glycerine
h	hour
HBTU	O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluorophosphate
His	histidine
HMP	4-hydroxymethylphenoxyacetic
^{1}H-MAS-NMR	solid-phase proton nuclear magnetic resonance
HOBt	N-hydroxybenzotriazole

HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
ID	inner diameter
IgG	immunoglobulin
Ile	isoleucine
LB	luria broth
LC	liquid chromatography
Leu	leucine
LHDMS	lithium hexamethyldisilazane
Lys	lysine
MBPP	meltblown polypropylene
Met	methionine
min	minute
MS	mass spectrometry
NMP	N-methylpyrrolidone
Nu	nucleophil
OD	outer diameter
OPD	O-phenylenediamine dihydrochloride
PAM	4-hydroxymethyl-phenylacetamidomethyl
Pbf	pentamethylchromane
PBS	phosphate buffered saline
PEG	polyethylene-glycol
Phe	phenylalanine
PM	portioning-mixing
Pro	proline
PS	polystyrene
PTFE	polytetrafluoroethylene
PVC	polyvinylchloride
QSAR	quantitative structure activity relationship
RIA	radioimmunological assay
RP	reverse phase
rt	room temperature
s	second
SAR	structure activity relationship
SCX-SPE	strong cation exchange solid-phase extraction
Ser	serine
SER	sequestration enabling reagents
SOM	small organic molecules
SP	solid phase
SPE	solid-phase extraction
TBME	t-butyl-methyl ether
TBS	tris buffered saline
t-Bu	t-butyl
TFA	trifluoracetic acid
THF	tetrahydrofuran
Thr	threonine
TIS	tri-isopropylsilane
TLC	thin layer chromatography
TMOF	trimethylorthoformate
TMS-Cl	trimethylsilyl chloride
TOF-MALDI	time-of-flight matrix assisted laser desorption ionization
Tris	tris [hydroxymethyl]aminomethane
Trp	triptophan
Trt	trityl
Tyr	tyrosine
Val	valine
WDI	World Drugs Index

1

Introduction

G. Fassina and S. Miertus

1 Introduction

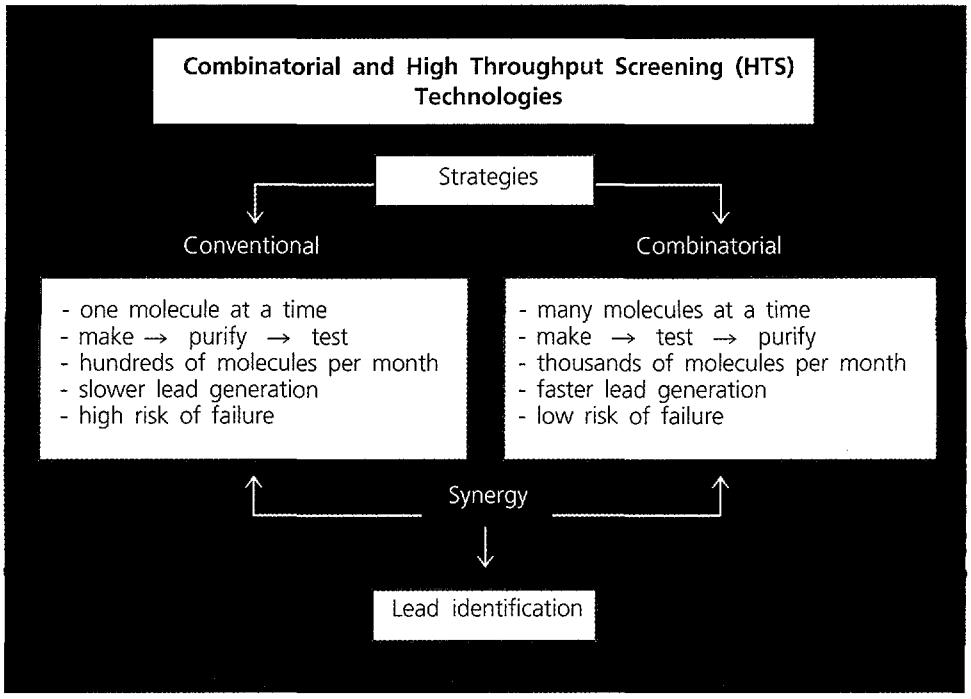
Giorgio Fassina and Stanislav Miertus

1.1. Combinatorial technologies

In its earliest expression, the primary objective of combinatorial chemistry focused on the generation of large numbers of molecules and the simultaneous screening of their activity. Originally, combinatorial approaches were based on the premise that the probability of finding a molecule in a random screening process was proportional to the number of molecules subjected to the screening process. Following such an approach, the success rate of identifying new leads has become greatly enhanced and the time required considerably reduced.

The development of new processes for the generation of collections of structurally related compounds (libraries) with the introduction of combinatorial approaches has revitalized random screening as a paradigm for drug discovery and has stimulated interest about the possibility of finding new and valuable drugs in a short time and at a reasonable cost (see figure 1.1).

Figure 1.1. Principal characteristics of conventional versus combinatorial strategies of drug discovery



The advent of this new field in drug discovery did not obscure the importance of "classical" approaches in medicinal chemistry, such as computer-aided rational drug design and QSAR. Instead, it catalysed their evolution to complement and integrate with combinatorial technologies.

The term "combinatorial" appeared in scientific literature at the beginning of the 1990s, but the first generation of combinatorial libraries dates back to the beginning of the 1980s. The first reports dealt with the simultaneous production of collections of chemically synthesized peptides, produced by solid-phase methods on solid supports [1-5]. Peptides are mentioned later in the chapter.

As mentioned, traditional methods of discovering drugs have been based on the random screening of collections of chemically synthesized compounds or extracts derived from natural sources, such as micro-organisms, bacteria, fungi, plants of terrestrial or marine origin, or by modifications of chemicals with known physiological activities.

Sources of molecular diversity

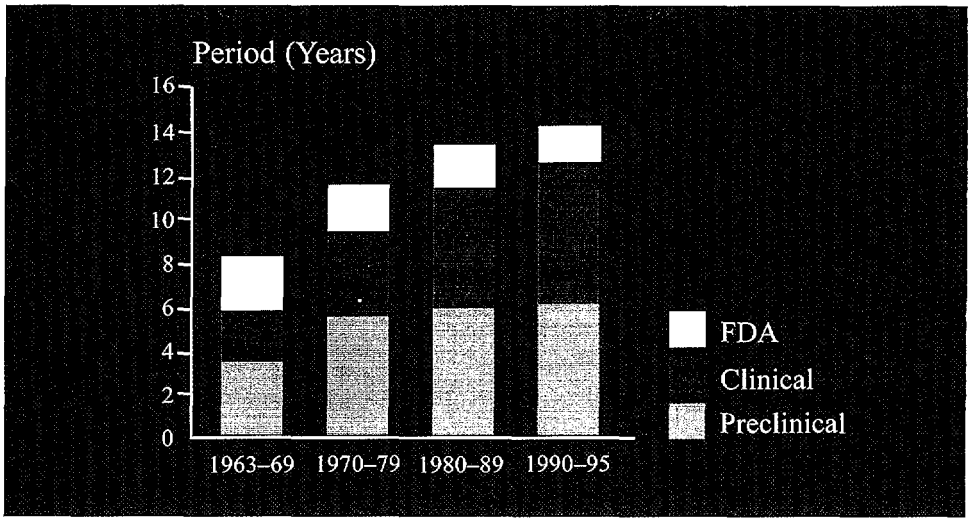
- Plant extracts
- Microbial extracts
- Collection of chemical compounds (synthetic)
- Oligonucleotide libraries (biological or synthetic)
- Oligosaccharide libraries
- Chemical compounds libraries (synthetic)
- Peptide libraries (biological or synthetic)

Libraries

Collection of structurally related compounds (peptides, oligonucleotides, oligosaccharides, organic molecules) obtainable by chemical or biological means simultaneously as a mixture and screened for activity as a mixture of compounds, without any isolation step. Identification of active compounds derives from the synthesis/production protocol used to generate the library. Faster identification of leads, since millions of different compounds can be screened simultaneously.

The traditional approach referred to has resulted in the discovery of many important drugs, however, the ratio of newly discovered to previously discovered compounds has diminished with time. In addition, the process is 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 on the number of samples tested. Chemical synthesis of new chemical entities is often a 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 $US 7,500 per compound [6]. The generation of natural extracts often provides interesting new molecular structures endowed with biological properties and leads to mixtures of different compounds in different concentrations, thus making activity comparison very difficult. In addition, once activity is found on a specific assay, the extract needs to be fractionated in order for the active component to be identified. Often, the chemical synthesis of natural compounds is extremely difficult, thus making lead development into a new drug a very complex task. The time and cost needed for the development of new drugs have increased steadily over the past three decades (see figure 1.2).

Figure 1.2. Time needed for the development of new drugs since 1963, from synthesis to approval



Estimated costs for introducing a new drug on to the market now reach between $200 million and $300 million and the process takes between 10 and 12 years after discovery. The increase in time and cost is in large part due to the extensive clinical studies of new chemical entities required by competent regulatory agencies such as the Food and Drug Administration. To a lesser extent, the increase is also due to the increased costs associated with research. The time and cost required for clinical and preclinical evaluations of new drugs is not likely to decrease in the near future. Consequently, an important strategy for pharmaceutical companies wishing to stay in the market has been to increase the number of new drugs being developed. While the pharmaceutical industry was demanding more rapid and cost-effective approaches to lead discovery, further complications arose, owing to several coinciding factors. They included the advent of new methodologies in molecular biology, biochemistry and genetics that led to the identification and production of an ever-increasing number of enzymes, proteins and receptors involved in biological processes of pharmacological relevance and good candidates for the development of screening assay. The introduction of combinatorial technologies provided an unlimited source of new compounds capable of satisfying all such needs.

Factors affecting strategy changes in drug discovery

Biotechnology (Genomics): provides molecular targets of therapeutic relevance (receptors, hormones, proteins).

Combinatorial technology: provides the possibility of generating huge collections of molecules that are produced simultaneously with a built-in decoding capability.

High throughput screening (HTS): provides the possibility of handling many assays at the same time.

As mentioned, peptides were particularly suited to 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 chemical diversity of peptides is demonstrated in figure 1. 3 below.

Figure 1.3. The number of compounds (peptides) generated by the combinatorial approach

Given a linear amino acid sequence of n residues

X_1-X_2-X_3-X_4- .....-X_n

The total number of different peptides obtainable equals:

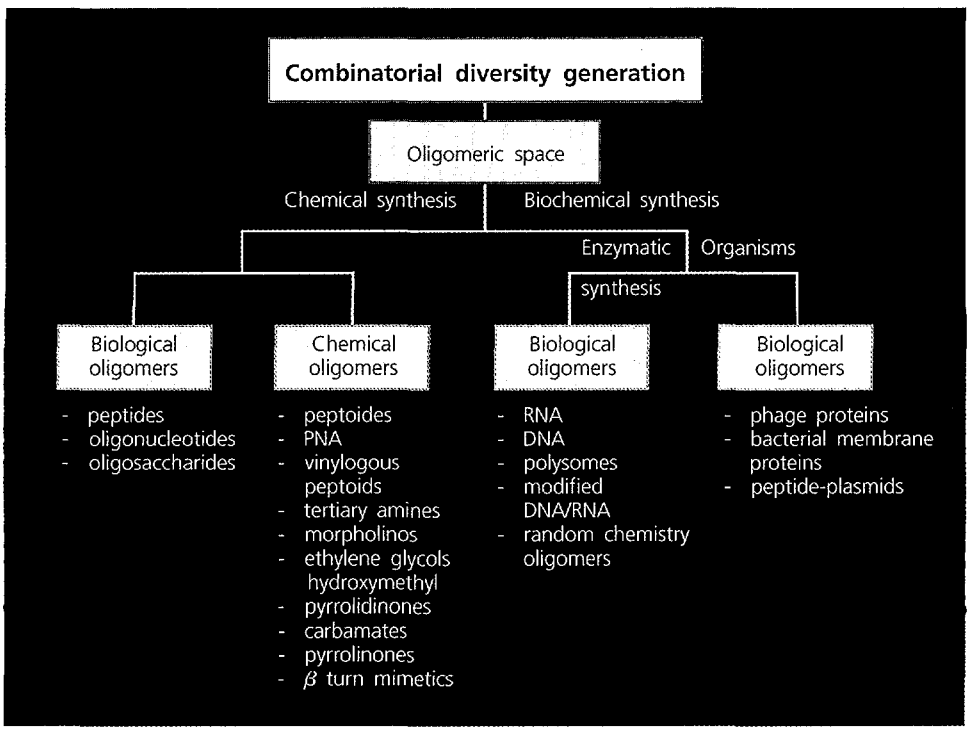
y^n

n = peptide length
y = number of different amino acids used in the synthesis (usually 18)

n = 3	5 832 peptides
n = 4	104 976
n = 5	1 889 568
n = 6	34 012 224

The use of peptide libraries was greatly accelerated by the introduction of biological methods for library preparation and by the use of phage display technology which provided interesting advantages over its synthetic counterpart [7, 8]. At the same time, the first papers on the generation of oligonucleotide libraries began to appear in the literature [9, 10]. This suggested the possibility of extending the applicability of combinatorial approaches to other classes of synthetic or natural oligomeric compounds, such as carbohydrates. There are many important, biologically active glyco-conjugate drugs whose carbohydrate constituents are associated with the molecular mechanism by which these drugs exhibit their effect. With such drugs, exploration of carbohydrate molecular diversity has the potential to identify novel agents with enhanced potency (see figure 4).

Figure 1.4. Diversity of compounds generated by a combinatorial approach



1.2. Applicability of combinatorial technologies

Many active compounds have been selected to date following combinatorial methodologies. A considerable number of those compounds have progressed to clinical trials. Combinatorial chemistry and other technologies related to producing and screening large numbers of molecules also find useful application in industrial sectors not necessarily related to the pharmaceutical industry. Emerging fields of application of combinatorial technologies include the following sectors: diagnostic, downstream processing, catalysis and new material. In the diagnostic sector, combinatorial chemistry can be applied successfully to the identification of previously unknown epitopes recognized by antibodies in biological fluids associated with pathological conditions. The selected epitopes can then be used in the development of diagnostic kits useful for the identification and quantification of the antibody of interest. In the downstream processing field, combinatorial chemistry finds application in the selection of ligands able specifically to recognize macromolecules of biotechnological interest such as proteins, antibodies or nucleic acids. This is relevant to industry, since the major costs associated with the production of recombinant molecules for therapy involves the purification of the desired target molecule from crude feedstocks. 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 also been applied to the identification of new macromolecules endowed with catalytic activity for reactions where natural enzymes are inactive. Such an application, although still at an early stage, is drawing considerable attention from the industrial sector, since the availability of new enzymes may reduce the production costs of many chemicals.

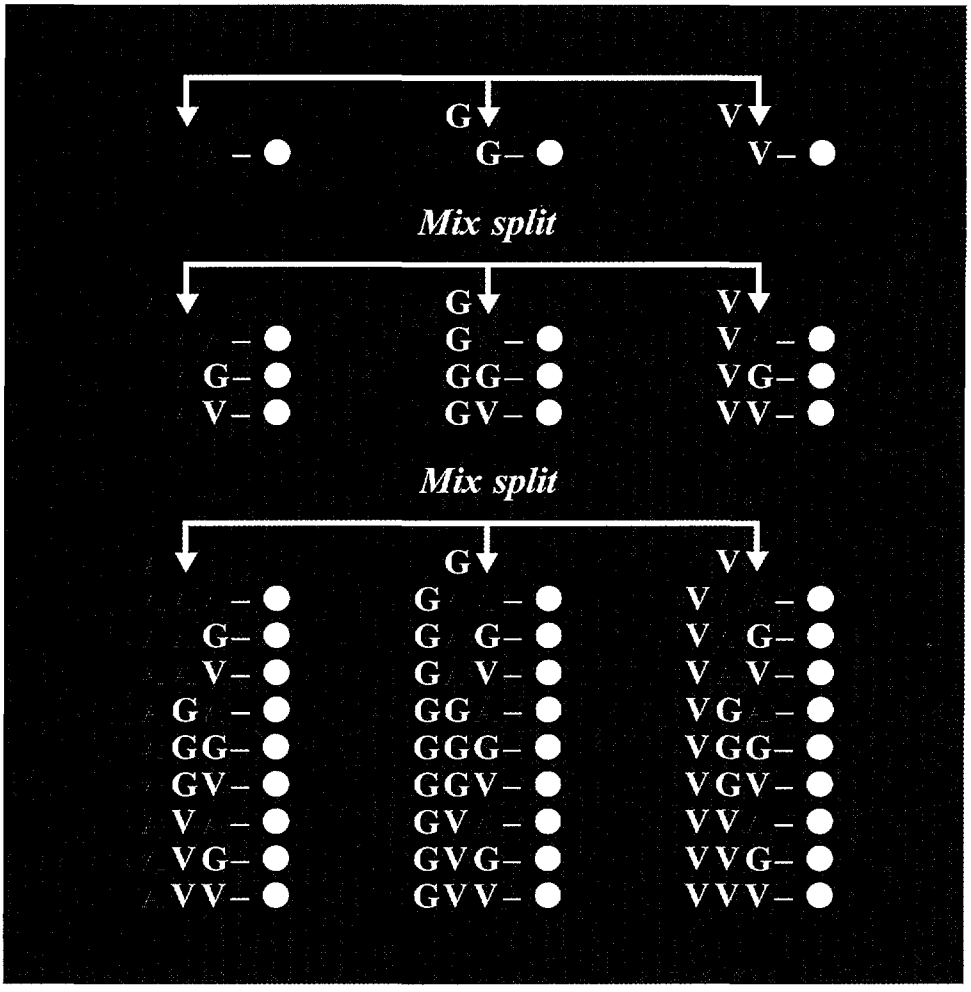
1.3. Combinatorial tools

A broad variety of new synthesis and screening methods are currently grouped under the term combinatorial. Such methods include parallel chemical synthesis and testing of multiple individual compounds or compound 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. Fully automated instruments for the synthesis and screening of libraries of compounds are integrated tools in combinatorial technologies as well as in computer-assisted approaches for library design. A significant class of molecular libraries is represented by peptide libraries. Peptides are particularly suitable for the construction of libraries since a high degree of structural diversification can be achieved easily by varying the peptide sequence length or by introducing amino acids that do not occur naturally. The number of different peptides obtainable by a combinatorial approach is governed by the simple formula:

$$N = b^x$$

where N is the total number of molecules obtainable, b is the number of residues used in the construction of the library and x is the sequence length. Generations of synthetic peptide libraries generally follow the divide-couple-recombine process (DCR) [2, 5]. The DCR process involves treating separately aliquots of resin for solid-phase synthesis with solutions that contain different activated amino acids. The amino acids are firstly coupled, then recombined, mixed and divided again in different aliquots (see figure 1.5 below). The process is repeated several times until the desired length of the library is accomplished.

Figure 1.5. The principle of the "Mix-Split" combinatorial approach in the preparation of peptide libraries



After resin cleavage and deprotection, peptides can be tested directly in biological assays. Alternatively, by using resins where the peptide is not cleaved after deprotection, peptide libraries can be tested still attached to the resin beads using target molecules tagged with appropriate labels for detection [12, 13]. Several procedures have been reported for the multiple synthesis of peptides [14, 15] or peptide libraries [16], but the majority of them require automated instruments or tailored laboratory equipment which are not readily available in many laboratories. However, simple procedures for the multiple synthesis of peptides or for the preparation of peptide libraries in the micromole scale have been developed. Such procedures require standard laboratory equipment, such as a vortex equipped with a sample holder for 25 Eppendorf tubes and a small centrifuge for polypropylene test tubes [17]. The procedure can be applied readily to the synthesis of peptide libraries as well as to the simultaneous small-scale manual synthesis of at least 30 different peptides.

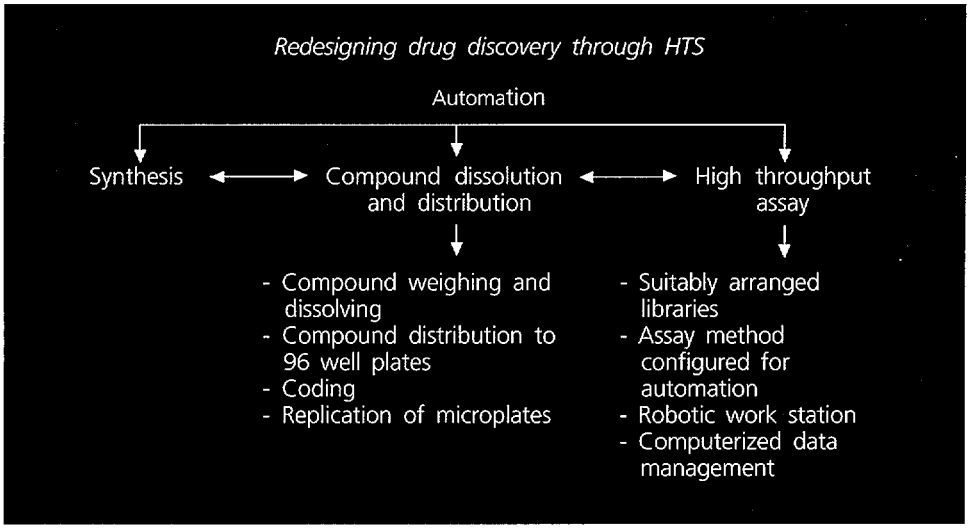
The synthesis starts suspending 0.1 mmol of resin for solid-phase synthesis in 9 ml of a DCM:NMP (6:4 v/v) mixture. Such a solvent composition allows a homogeneous dispersion of the resin, making the aliquoting for the preparation of the libraries simple and convenient by pipetting the desired amount of suspended resin in the polypropylene test tubes. Subsequently, 1 ml of NMP is added to each aliquot, which is centrifuged after vortexing. In the solvent mixture (DCM:NMP 1:4 v/v), the resin separates easily from the solvent. Excess solvent is then removed by vacuum aspiration, using a needle connected to a water vacuum pump. Following this, each resin aliquot

is treated with a solution containing the appropriate activated amino acid. Such a method provides a low-cost and easy approach to peptide libraries synthesis for laboratories whose synthesis requirements do not justify investment in an automated peptide synthesizer. All reagents and laboratory instruments are commercially available at very low cost, so there is no need for custom-made equipment. In addition, the method has a more general application, since it can be used whenever a solid-phase synthesis needs to be performed either to prepare single peptides or libraries (linear, multimeric, cyclic or peptoidic). The small synthesis scale employed allows the preparation of only a few mmols of peptides, but the amount is more than adequate for a vast array of biological assays. A protocol for the manual synthesis of peptide libraries is described in chapter 2.1.

An important aspect of combinatorial chemistry is the analytical characterization of molecular libraries. Since a considerable number of different molecules are tested separately or in combination, analytical data should indicate that all the expected components are occurring with a comparable degree of purity. In chapter 3, protocols are described on the application of amino acid analysis and TOF-MALDI mass spectrometry for the quality control of peptide libraries. The amino acid analysis is useful mainly for characterizing amino acid-based libraries (peptides, benzodiazepines, hydantoins). The analysis on its own is not sufficient to give conclusive information on the actual composition of a library, but it represents a rapid and versatile approach to evaluating the distribution of components in a mixture or to checking the presence or absence of a given peptide family. One of its most important advantages is that such an analysis 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, such as paper, cotton or glass. Using data obtained from amino acid analysis to judge the quality of a given library has one significant disadvantage, namely, that by-products deriving from incomplete side-chain deprotection or side-chain modifications cannot be detected owing to the integrity of amino acids often being restored during hydrolysis. The presence of by-products is best investigated using mass spectrometry methods (tandem mass spectrometry). Mass spectrometry methods are powerful tools for the analysis of mixtures of compounds from any source. Different techniques such as Electrospray (ES), Matrix Assisted Laser Desorption Ionization (MALDI), Fast Atom Bombardment (FAB) and tandem mass spectrometry have been used successfully to evaluate the composition and purity of synthetic peptide mixtures, but there are no limitations for their use with purely organic libraries. When interfaced to HPLC or capillary electrophoresis, the ES becomes the most powerful method for the characterization of even very complex mixtures, the reason being that the combination of the two techniques allows the identification of compounds with very similar chemical properties. MALDI is an ideal method to use when very small amounts of sample are available.

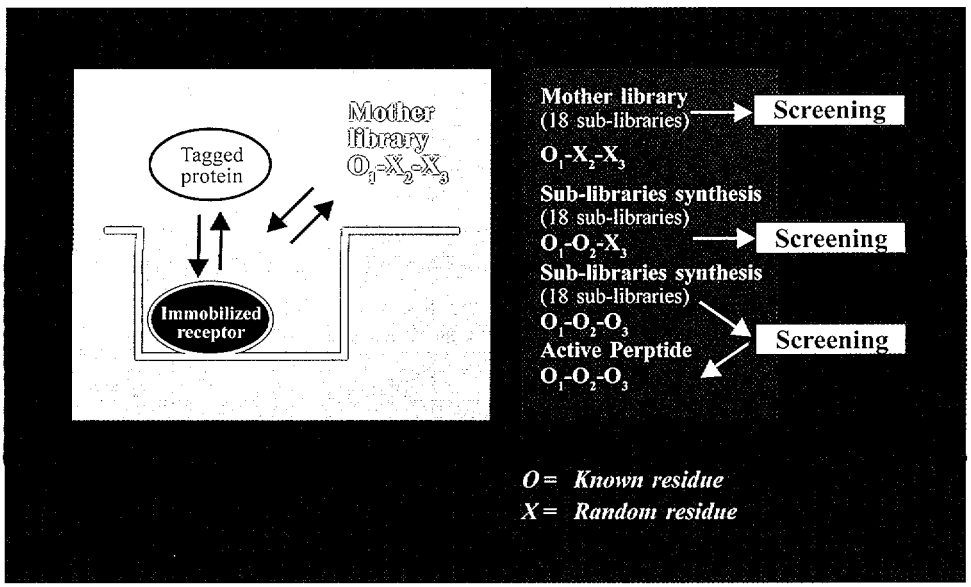
In combinatorial chemistry, owing to the high number of chemical manipulations required to synthesize libraries of compounds and the high number of screening steps, automation is unavoidable (see figure 1.6 below). Many research groups, both in academia and in industry, are developing automated instruments specifically tailored to such needs. It is a field of technology that is acquiring an extremely important role in the development of combinatorial technologies in the next 20 years. However, semi-automated instruments requiring little investment may be constructed in research laboratories on a low budget. The design and development of a small item of equipment for the semi-automatic synthesis of peptide libraries is discussed in chapter 4.

Figure 1.6. The role of automation in combinatorial chemistry and combinatorial technology



The screening steps required to decipher the active sequence from a molecular library are strictly related to the type of library used, to the synthesis or preparation cycles needed and to the kind of activity wanted. Molecular libraries can be prepared following chemical or biological approaches. For the former, libraries can be prepared free in solution or anchored to solid supports; the two situations require different screening procedures. Resin-released libraries can be conveniently used in searching for molecules that are 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 performed by 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. Figure 1.7 shows the conventional uses of screening sub-libraries.

Figure 1.7. Screening soluble peptide libraries



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, labelled 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 iterative cycles of synthesis and screening as before, the structure of the active compounds is elucidated. A protocol for the screening of peptide libraries immobilized on solid supports is the focus of chapter 5.

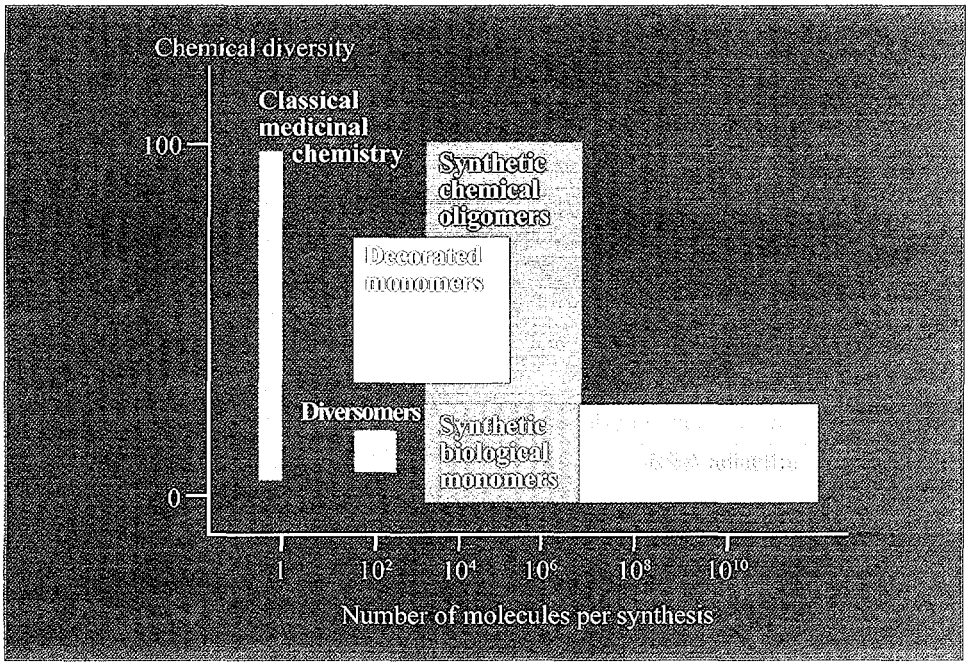
In combinatorial chemistry many different types of libraries can be produced by using solid-phase or solution-phase methods. Table 1.1 shows the characteristics of solid-phase and solution-phase combinatorial chemistry. A protocol for the solid-phase synthesis of highly functionalized amine libraries is discussed in chapter 6; the subsequent chapter provides a protocol for the solution-phase synthesis of a discrete library of heterocycles.

Table 1.1. Characteristics of solid-phase and solution-phase combinatorial chemistry

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

The different technologies and strategies used in the production of combinatorial libraries are now so well developed that it is easy to plan synthetic schemes for the generation of a huge number of compounds. Since the rate at which compounds can be screened limits the use of combinatorial technologies, it is important to be selective about the compounds that are synthesized (see figure 1.8).

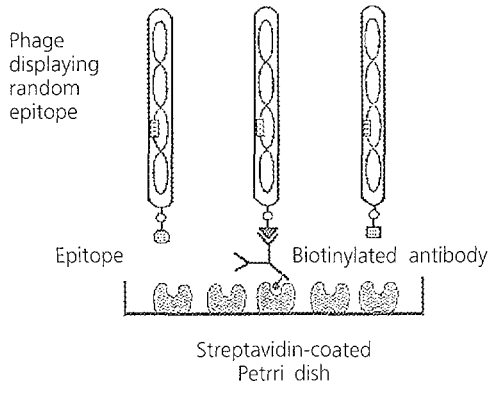
Figure 1.8. Chemical diversity and the number of molecules produced by various concepts of synthesis



From this point of view, computational methods are very valuable in assisting 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 that 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. Applying similarity coefficients then quantifies the similarity. A detailed protocol on how to design combinatorial libraries is provided in chapter 8.

Most biological methods for library preparation are 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. Figure 1.9 shows the principle governing the preparation of phage libraries.

Figure 1.9. Phage libraries



The peptides are fused to proteins normally expressed on the surface of the micro-organism used. Phage display libraries are the most commonly used. Screening is accomplished by incubation of the target molecule adsorbed on 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 conditions 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 a few active sequences. Active phages may then be subjected to DNA sequencing in order to decode the active peptide sequence.

The use of biological display libraries for the isolation of peptide ligands is a viable alternative to chemical libraries. The technique was first published in 1985 [18]. Since then, it has been used successfully in many fields of research. Biological display libraries are made up of large pools of micro-organisms (from 10^9 to 10^{10}), each one expressing a different polypeptide on its surface. Such libraries can be propagated easily 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 micro-organisms are eluted, grown and selected over the target between two and four more times. The resulting single clones are easily isolated and analysed.

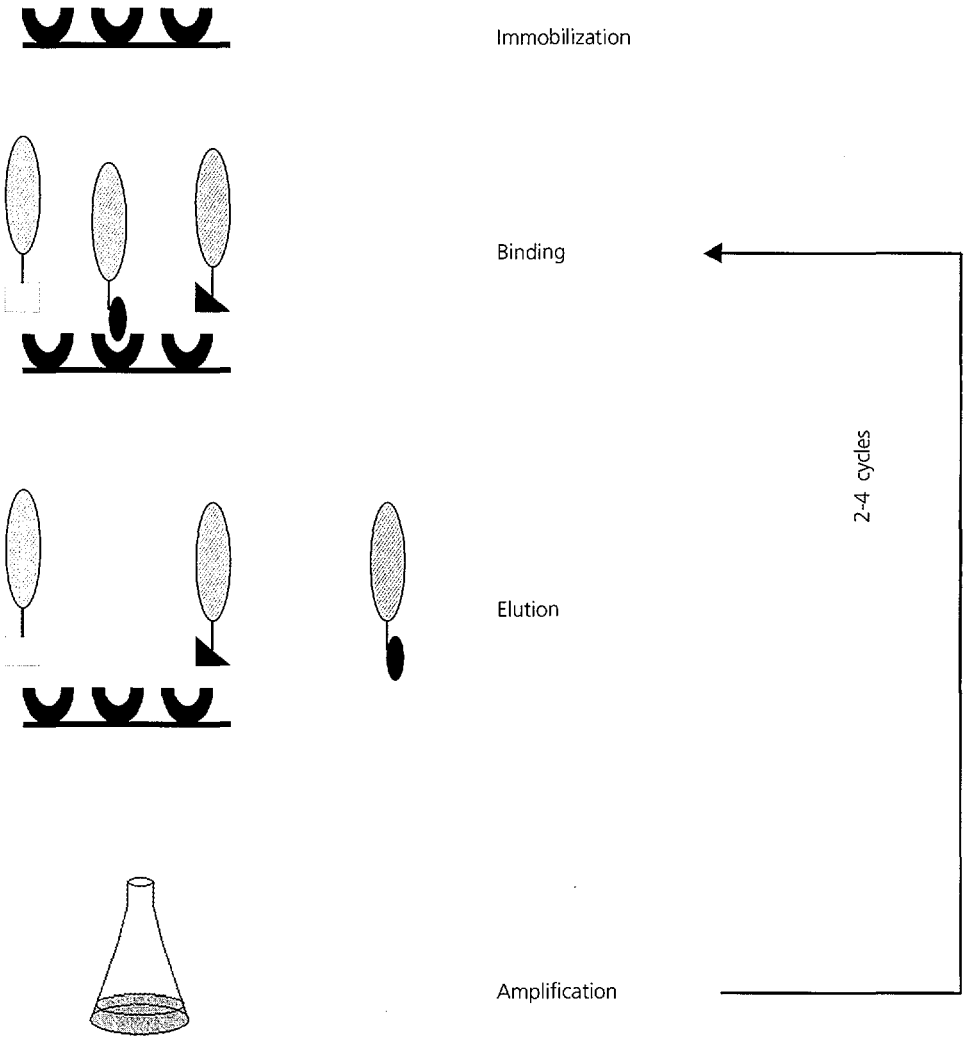
The construction of biological display libraries requires the introduction into a micro-organism of the genetic information necessary for peptide synthesis. For the construction of a random peptide display library it is necessary to synthesize pools of DNA fragments that are then inserted into specific vectors. The DNA fragments are chemically synthesized as a mixture of single-stranded degenerated oligonucleotides containing constant regions plus one degenerated stretch (or more) of DNA. DNA consists of sequences of four different nucleotides and each trinucleotide code for a corresponding amino acid. Because of the codon degeneracy, most of the amino acids are coded by more than one triplet. Since it is possible to introduce stop codons that will interrupt protein synthesis in fully degenerated oligonucleotides, the oligonucleotides are synthesized using different mixtures of nucleotides especially in the third position of each triplet [19]. 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 micro-organism.

The most common micro-organism used for peptide display is the *E. coli* filamentous bacteriophage [19]. Bacteriophages are viruses that infect bacteria by injecting their single-stranded DNA genome into bacterial cells. Once inside the cell, they start to replicate their DNA using the host protein synthesis machinery. Their coat proteins are synthesized and the DNA is packaged into phage particles across the bacterial membrane and secreted into the medium from which they are easily recovered by precipitation. Protocols for growing phages and bacteria in order to carry on screening experiments are discussed in chapter 9.

The ligand selection process from biological libraries is called biopanning. The principles of biopanning are shown in figure 1.10. The target molecule must be bound to a solid support, usually a microtiter plate or a small Petri dish. Less common alternative supports include a column with solid matrices, magnetic particles, cells and mammalian organs. In a typical experiment, the number of phages that are incubated with the target corresponds to about 100 to 1,000 times the complexity of the library. After the unbound clones are washed away, the bound ones are eluted by different methods, such as low pH, high concentration of free target, direct infection of bacteria

cells. The eluted phages are grown, purified and submitted to a new cycle of selection. Usually three to four rounds of selection are sufficient and the entire process can be completed in about one 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 micro-organism system. Several different approaches for biopanning experiments are illustrated in chapter 10.

Figure 1.10. Biopanning procedures



- **Immobilization** of the target on the support
- **Binding** of the phage library with the target
- **Washing** to remove aspecific interactions
- **Elution** of bound phages
- **Amplification** of eluted phages in *E. Coli*

References

1. Furka, A., Sebestyen, F., Asgedom, M. and Dibo, G. 1988. *Abstr. 14th Int. Congr. Biochem.*, Prague, vol. 5, p. 47.
2. Houghten, R. A. 1985. *Proc. Natl. Acad. Sci. USA*, vol. 82, pp. 5131-5135.
3. Fassina, G., Lebl, M. and Chaiken, I. M. 1988. *Coll. Czech. Chem. Comm.*, vol. 53, pp. 2627-2636.
4. Furka, A., Sebestyen, M., Asgedom, M. and Dibo, G. 1991. *Int. J. Pept. Protein. Res.*, vol. 37, p. 487.
5. Hougthen, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C.T. and Cuervo, J. H. 1991. *Nature*, vol. 354, pp. 84-86.
6. Chabala, J. C. Pharmaceutical Manufacturers Association, Drug Discovery Management Subsection, Philadelphia, 19-21 September 1993, United States.
7. Cwirla, S., Peters, E. A., Barret, R. W. and Dower, W. J. 1990. *Proc. Natl. Acad. Sci. USA*, vol. 87, pp. 6378-6382.
8. Scott, J. K. and Smith, G. P. 1990. *Science*, vol. 249, pp. 386-390.
9. Tuerk, C. and Gold, L. 1990. *Science*, vol. 249, pp. 505-510.
10. Ellington, A. D. and Szostak, J. W. 1990. *Nature*, vol. 346, pp. 818-822.
11. Fassina, G., Verdoliva, A., Odierna, M. R., Ruvo, M. and Cassani, G. 1996. *J. Mol. Recogn.*, vol. 9, p. 564.
12. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J. 1991. *Nature*, vol. 354, p. 82.
13. Lam, K. S., Hruby, V. J., Lebl, M., Knapp, R. J., Kazmierski, W. M., Hersh, E. M. and Salmon S. E. 1993. *Bioorg. Med. Chem. Lett.*, vol. 3, p. 419.
14. Holm, A. and Meldal, M. 1988. *Peptides 1988*, ESCOM, Leiden, p. 208.
15. Gausepohl, H., Kraft, M., Boulin, C. and Frank, R.W. 1991. *Peptides 1990*, ESCOM, Leiden, p. 206.
16. Knapp, D. R., Oatis, J. E. and Papac, D. I. 1993. *Int. J. Pept. Protein. Res.*, vol. 42, p. 259.
17. Ruvo, M., Scardino, P., Cassani, G. and Fassina, G. 1994. *Prot. Pept. Lett.*, vol. 1, p. 187.
18. Smith, G. P. 1985. *Science*, vol. 228, p. 1315.
19. Smith, G. P. and Scott, J. K. 1993. *Methods Enzymol.*, vol. 217, p. 228.

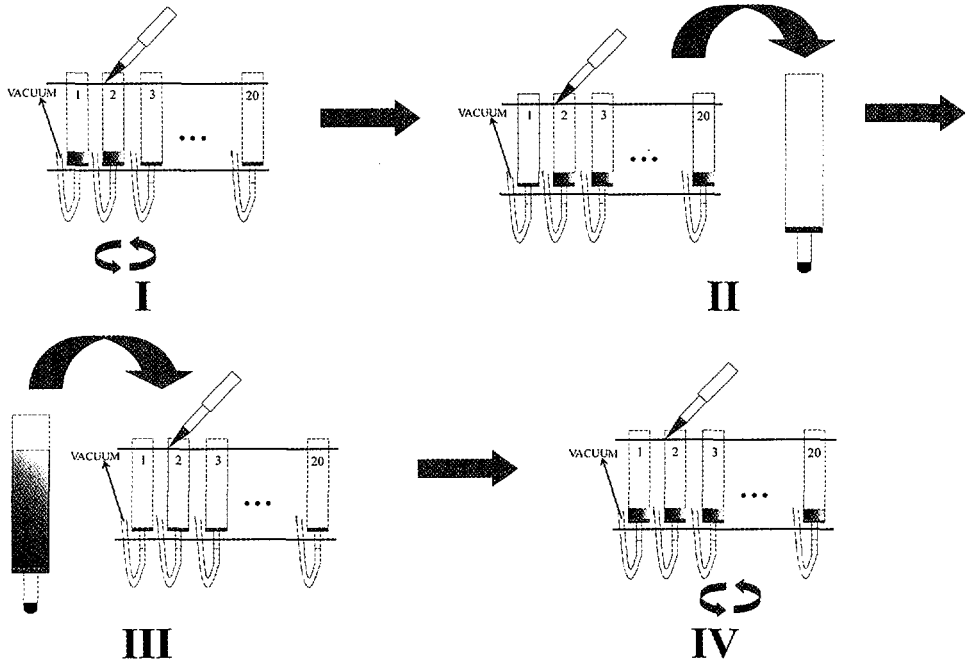
2

Manual synthesis of peptide libraries

M. Ruvo

2 Manual synthesis of peptide libraries

Menotti Ruvo



- Production of collections of peptides in equimolar amounts
- Compatible with solid-phase and solution-phase screening assays
- Low cost, non-automated procedure
- Possibility to increase chemical diversity by changing the building blocks

2.1. Introduction

In the search for new active compounds, the most rapid approach utilized today involves applying combinatorial techniques, that is, exploring the chemical diversity represented by large collections of rationally generated synthetic molecules. The larger such collections, the higher the probability of finding components able to fulfil a given biological role. Peptides are very useful molecules to be utilized for the synthesis of such collections or libraries, since by modulating the length and the number of different amino acids used to build up the library, very large and diverse arrays of compounds can be achieved [1, 2, 3]. Further, using the solid-phase method [4], peptide synthesis is accomplished by performing a few well-established reactions that lead to high pure compounds. Generally, peptide libraries are classified into two main categories: peptide libraries bound to solid supports (solid-phase peptide libraries) [3, 5] and soluble peptide libraries [2]. Both types are, in most cases, synthesized by applying the above-mentioned solid-phase method, but they differ both in physical status and in the way they are screened after preparation. The solid-phase libraries are screened while still attached to the solid support on which they have been synthesized (after removal of the side-chain protecting groups), while the soluble libraries are screened in solution after release from the resin beads used for synthesis. Solid-phase libraries are useful only in binding experiments, since the support-bound peptides can be seen as highly concentrated artificial receptors (on the surface of the resin beads) to which the target molecule can bind. They are usually prepared on polymeric resin beads [3] applying the portioning-mixing (PM) method (one-bead-one-peptide concept) [1, 2, 3], but many useful applications of libraries on plastic pins [6], cotton [7], paper [8] or microchip surfaces [9] have also been reported. The soluble libraries are much more versatile and can be used in any kind of binding or inhibition assay [10, 11]. They are prepared by applying both the portioning-mixing and the pre-mix method [10, 11, 12].

2.1.1. Portioning-mixing method

The portioning-mixing method is applied equally for the synthesis of solid and soluble peptide libraries; the difference is made by using the linker to anchor the molecules to the polymeric chain composing the resin. When an acid-stable linker (such as the PAM linker [13]) is used, the peptides are freed from the side-chain protecting groups but remain anchored to the resin beads. When an acid-labile linker (such as the HMP linker [14]) is used, the peptides are also removed from the support, becoming soluble entities that can be treated as single compounds. The portioning-mixing method is accomplished in four main steps (see chapter 1 title page):

- Coupling a single activated monomer to separated aliquots of resin;
- Mixing the separated aliquots;
- Portioning the resin into a number of new aliquots that equals the number of starting monomers;
- Re-coupling monomers to the resin aliquots.

The above four steps are repeated until the desired library length is reached leading to a collection of n^x of different peptides in equimolar amounts, where n is the number of monomers used and x is the length of the library.

The main drawback of libraries of this kind is that the number of resin beads used in the synthesis limits the number of different molecules that can be obtained. In fact, unless working with kilograms of resin, when all 20 proteinogenic amino acids are used for the construction of the library, only libraries containing up to six random residues can be prepared [10].

2.1.2. *Pre-mix method*

The pre-mix method is amenable only to the synthesis of soluble peptide libraries since randomization, achieved in this case by coupling mixtures of activated monomers to the solid support, leads to a distribution of sequences on the single bead. The product distribution is influenced strictly 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 achieved accurately. Peptide libraries up to a size of 15 residues have been generated using this method (with and without concentration corrections), and successfully screened for biological activity.

The following sections represent the detailed protocols for the preparation of a simple tripeptide library, applying the methods that have been outlined here, starting from the initial resin washes and concluding with storage conditions.

2.2. Materials and methods

2.2.1. *Materials*

- 9-fluorenylmethoxycarbonyl (Fmoc)-derivatized amino acids (purity > 99%). 0.25 M stock solutions of protected amino acids are prepared, weighing 2.5 mmol of each amino acid in 20 graduated tubes and adding DMF up to 10 ml. The solutions are then stored at −20° C until required. Equimolar mixtures of the 20 amino acids are conveniently prepared by mixing equal volumes of each 0.25 M stock solution just before use

Table 2.1. Protected amino acid derivatives used to synthesize the library

Fmoc-Ala-OH
Fmoc-Leu-OH
Fmoc-Cys(Acm)-OH
Fmoc-Asp(tBu)-OH
Fmoc-Lys(Boc)-OH
Fmoc-Phe-OH
Fmoc-Tyr(tBut)-OH
Fmoc-Pro-OH
Fmoc-Ile-OH
Fmoc-Val-OH
Fmoc-Ser(tBu)-OH
Fmoc-Thr(tBu)-OH
Fmoc-Met-OH
Fmoc-Glu(tBu)-OH
Fmoc-Gln(Trt)-OH
Fmoc-Asn(Trt)-OH
Fmoc-Gly-OH
Fmoc-Trp(Boc)-OH
Fmoc-Arg(Pbf)-OH
Fmoc-His(Trt)-OH

- 4-hydroxymethyl-phenylacetamidomethyl and 4-hydroxymethylphenoxy-acetic polystyrene resins (PS-HMP and PS-PAM resins)
- O-benzotriazole-N,N,N',N'-tetramethyl-uronium hexafluorophosphate (HBTU, > 99%) dichloromethane (DCM)
- N-methylpyrrolidone (NMP)
- N,N-dimethylformamide (DMF)
- methanol
- trifluoroacetic acid (TFA)
- water (HPLC grade)
- acetonitrile (HPLC grade)
- phenol
- thioanisol
- tri-isopropylsilane (TIS)
- ethanedithiol
- N-hydroxybenzotriazole (HOBt)
- diciclohexylcarbodiimide (DCC)
- di-isopropyl-ethylamine (DIEA)

2.2.2. *Methods*

A. Portioning-mixing method in the synthesis of soluble libraries

1. *Loading of the Gly residue*[1,2]: 100-110 mg of PS-HMP resin (subst.: 0.99 mmol/g)[3] are placed in a filter-equipped plastic vessel and washed with 3 ml of DMF, 3 ml of DCM and finally with 3 ml of DMF:DCM (2:3).

297 mg (1.0 mmols) of Fmoc-Gly-OH are placed in a 15 ml plastic tube and dissolved in 4 ml of DMF:DCM (1:4). Added to this is 0.5 ml of 1 M DCC in DMF.[4] A white precipitate of dicyclohexylurea (DCU) is rapidly formed after a few minutes. The tube is vortexed for 30 min and the precipitate removed by filtration on glass wool. Added to this is 0.5 ml of a 0.1 M solution of N,N-dimethylaminopyridine (DMAP) in DMF (a catalyst), and the activated amino acid is transferred into the resin-containing vessel. The suspension is gently stirred for 1 h at rt. The resin is vacuum drained from the bottom and washed as follows:

---

[1] This section describes the synthesis of a simple tripeptide library with the general formula O_3-X_2-X_1-G. It uses all 20 proteinogenic amino acides. (Cysteine is protected with the Acm group to prevent disulphide bridge formation and consequently polymerization.) Such a library is composed of 20^3 = 8,000 different sequences. A glycine residue is introduced at the C-terminus to overcome racemization problems connected to the anchoring of the first residue on to the resin (see also footnote 2) and, in the case of resin bound libraries, also as a spacer.

[2] Amino acids such as histidine and cysteine undergo a partial racemization during DCC-mediated esterification. Other residues such as arginine, asparagine and glutamine form unstable symmetrical anhydrides and have to be loaded on to the resin using other approaches.

[3] Assuming that 1 mg of resin corresponds to around 3,000 beads (typical values for such resins), 100 mg corresponds to 300,000 beads. In the final library, around 37 beads will represent each peptide which correspond to around 12 nmols (substitution 0.99 mmol/g) or, assuming an average molecular weight of 350 amu for all tripeptides, 4.6 μg of each peptide. If a 1 mg/ml solution is prepared from each sub-library (400 molecules), we obtain a 7.1 μM solution of each peptide. To obtain similar quantities of a tetrapeptide library (160,000 sequences), the synthesis should start from around 2 g of the same resin. When resin-bound peptide libraries are prepared following the one-bead-one-peptide approach, a minimal theoretical distribution of 10 beads per sequence should be pursued, since, due to inaccurate resin splitting and re-mixing, some sequences could be missed.

3 x 5 ml of DMF:DCM (2:3) (removal of the residual activated amino acid);
2 x 5 ml of CH_3OH (removal of the residual DCU);
3 x 5 ml of DMF.

Then 2 ml of a 0.5 M solution of acetic anhydride in DMF and 0.5 ml of 0.1 M DMAP in DMF are added in order to block the unreacted amino groups.

The Fmoc deprotection is accomplished treating the resin with 3 ml of 20% piperidine in DMF for 15 min. The deprotection solution is removed and collected along with the subsequent DMF washes (3 x 3 ml). In order to establish the actual amino acid loading on to the resin, the Fmoc-piperidine adduct concentration is spectrophotometrically determined (Σ = 7.8 mM^{-1} x cm^{-1}). The solution is 100-fold diluted and the absorbance at 301 nm determined (blank DMF). Assuming a loading yield of 99%, the absorbance should be in the range of 0.63-0.70 Absorbance (pathlength 1 cm) according to the amount of resin.[4]

2. *Resin splitting into 20 equal aliquots:* The resin is transferred into a 15 ml polypropylene graduated vessel and 12 ml of DMF:DCM (2:3) are added.[5] The suspension is thoroughly mixed and fractions of 0.5 ml are sampled in 20 polypropylene syringes (3 ml) endowed with filtration septa at the bottom.[6] The remaining volume of suspension is diluted up to 12 ml with DMF:DCM (2:3) and 0.5 ml fractions are again dispensed into the tubes. The graduated vessel is once again filled up to 10 ml and 0.5 ml aliquots distributed into the tubes. The tubes are then vacuum drained from the bottom and the resins washed once with 1 ml of DMF. Each tube contains an equal fraction of resin corresponding to around 5 μmol of NH_2 groups.

3. *Coupling of the first random residue (position 3, see figure on first page of the present chapter):* 200 μl (50 μmol) of 0.25 M stock solutions in DMF of the 20 amino acids (listed in table 2)[7] are placed in 20 different 2 ml polypropylene tubes. Added to each tube are 100 μl of 0.5 M HBTU/HOBt in DMF and 20 μl of neat DIEA. After 3 min mixing, the solutions are transferred on to the corresponding resins and the tubes vortexed for 30 min. The resins are drained and washed 3 times with 1 ml of DMF.

4. *Deprotection of the first random residue:* To each tube 0.5 ml of 20% piperidine in DMF are applied for 15 min under vortexing. The resins are then drained and washed 3 times with 1 ml of DMF.

---

[4] Determination of reaction yields measuring Fmoc-piperiding adduct release is based on the assumption that the deprotection reaction occurs 100%. In fact, the reaction rate is sequence-dependent and complete deprotection cannot be assured after only 15 min. If a lower amount of Fmoc-piperidine is detected, the piperidine treatment is repeated, thereby increasing the concentration of piperidine in the deprotection solution. When a discrepancy higher than around 10% is still obtained, the library should be discarded. Capping with acetic anhydride after each coupling step would be a good measure to prevent the formation of large populations of deleted sequences. Capping is not accomplished routinely, since typical reaction yields for peptide synthesis, mainly using large amino acid excesses, are always around 99%, and coupling problems can arise only in those cases where wrong operations or damaged chemicals are employed. Assessing the amount of Fmoc-piperidine adduct released from the resin-anchored glycine is an important step since the true starting synthesis scale is determined. The deprotection can also be performed on the separated resin aliquots, after splitting in order to check that a good resin distribution has been achieved.

[5] 60% DCM in DMF is used to suspend the resin beads. This mixture works perfectly with polystyrene resins (200-400 mesh) giving a highly homogeneous suspension. Working with other resins, the DCM/DMF ratio should be adjusted to obtain the optimal density that can assure the best dispersion.

[6] 3 ml polypropylene tubes equipped with filtration septa at the bottom are available commercially or can be assembled in-house, using disposable 3 ml syringes.

[7] The acetamidomethyl (Acm) group is a TFA-stable protecting group for cysteine, hence the cysteines in such libraries should be considered a non-proteinogenic amino acid. If free sulphidril groups are desired, Acm removal can be achieved by treatment with Hg(II)/AcH.

The deprotection solution and the subsequent washes from each single tube are collected and pooled. These solutions are 20-fold diluted and the Fmoc-piperidine adduct concentration spectrophotometrically determined by reading the absorbance at 301 nm.[4]

5. *Mixing and re-splitting:* 500 μl of DMF:DCM (2:3) is added to each tube and the resins suspended by gentle swirling. The suspensions are removed from the tubes, pipetted in a unique vessel and thoroughly mixed by vigorous shaking. After the addition of DMF:DCM (2:3) up to a final volume of around 12 ml, 0.5 ml aliquots of suspension are again pipetted into the empty tubes, repeating the operations described in the second step above. At the end, the resins are washed with 1 ml of DMF.

6. *Coupling of the second random residue (position 2, see figure on first page of the present chapter):* 200 μl (50 μmoles) of 0.25 M stock solutions in DMF of the 20 amino acids are activated with HBTU/HOBt and DIEA as described in step 4. The solutions are then transferred into the corresponding tubes for coupling. After 30 min vortexing, they are drained off from the resins. The resins are washed 3 times with 1 ml of DMF.

7. *Deprotection of the second random residue:* Added to each tube is 0.5 ml of 20% piperidine in DMF for 15 min under vortexing. The resins are then drained and washed 3 times with 1 ml of DMF. The yields of the previous couplings are determined as described in step 4.

8. *Mixing and re-splitting:* As for step 5 above.

9. *Coupling of the third known residue (position 1, see figure on first page of the present chapter):* The 20 amino acids are activated with HBTU/HOBt and DIEA as described in step 3. The solutions are then transferred into the corresponding tubes for coupling and after 30 min vortexing they are drained off from the resins. 3 ml washes with 1 ml of DMF are performed to remove the excess of amino acids.

10. *Deprotection of the third random residue:* As for step 4 above.

11. *Final washes and drying of the resins:* The resins are washed 3 times with 1 ml of DCM, 3 times with 1 ml of CH_3OH, and 2 times with 1 ml of ethyl ether. The resins are then vacuum-dried and weighed.

12. *Cleavage:* A TFA-H_2O-thioanisol-phenol-TIS (86:3:3:4.5:1.5, w/w) mixture is freshly prepared and 800 μl added to each sample. After vortexing for 5 h, the resin residues are removed by filtration through 1 ml pipette tips plugged with glass wool. Adding 1.5 ml of cold diethyl ether to each tube, the peptide mixtures are precipitated.[8] The precipitates are centrifuged and the ether removed. The crude materials recovered after centrifugation are once again washed with 1 ml of cold diethyl ether, dissolved in 0.5 ml of H_2O-CH_3CN-TFA (50:50:0.1) and, after lyophilization, separately weighed.[9] 10 mg/ml stock solutions of the peptide libraries are usually prepared in DMSO and stored in sealed vials frozen at −80° C until required.

[8] If the precipitation is slow or if the libraries do not precipitate at all, an explanation could be that the TFA used for the cleavage is too much. In such cases, the cleavage mixtures should be concentrated after resin removal or a larger volume of ethyl ether should be used for precipitation. Decreasing the temperature (to as low as - 80° C) is also an alternative. Another precipitating agent such as TBME (t-butyl-methyl ether) could also be used.

[9] An estimate of the expected weight of each sub-library could be obtained assuming an average molecular weight of 350 amu for all peptides and a theoretical final yield of 95% (4.75 μmol corresponding to 1.7 mg of product). It was possible to obtain an estimate of the expected weight of each solid sub-library by adding to the intial weight of resin divided by 20 the theoretical weight of 4.75 μmol of a tripeptide with an average molecular weight of 350 amu.

B. Portioning-mixing method in the synthesis of resin-bound libraries

1. Loading of the Gly residue: 100-110 mg of PS-PAM resin[10] (subst.: 0.80 mmol/g) are loaded with the C-terminal glycine as described in step A1.

All subsequent steps for library assembling are performed exactly as described in the previous section, steps A2-A11.

2. Cleavage: A TFA-H_2O-thioanisol-phenol-TIS (86:3:3:4.5:1.5, w/w) mixture is freshly prepared and 800 μl added to each sample. After vortexing for 5 h, the cleavage mixture is drained and discarded. The resins are then washed with 0.5 ml of neat TFA, 3 times with 1 ml of DCM, 3 times with 1 ml of ethyl ether and vacuum-dried. The resins are weighed[9] and then stored in sealed vials at −80° C until required.

C. Pre-mix method for the preparation of soluble peptide libraries

1. Loading of the Gly residue: The anchoring and deprotection of Fmoc-Gly is accomplished as described in step A1, starting from 100-110 mg of PS-HMP resin (subst.: 0.99 mmol/g).

2. Resin splitting into 20 equal aliquots: This operation is carried out as described in step A2. After splitting, each tube contains an equal fraction of resin corresponding to around 5 μmol of NH_2 groups.

3. Coupling of the first random residue: 220 μl from each 0.25 M stock solution of the 20 Fmoc-protected amino acids (see table 2) are taken and thoroughly mixed. The resulting mixture is activated for 3 min with 2.2 ml of 0.5 M HBTU/HOBt in DMF and 440 μl of DIEA. 200 μl aliquots are then pipetted on to the resins and allowed to react with the amino groups for 30 min. The solution of activated amino acids is drained by suction and the resins washed 3 times with 1 ml of DMF.

4. Deprotection and washes: Applied to each tube is 0.5 ml of 20% piperidine in DMF for 15 min vortexing. The resins are then drained and washed 3 times with 1 ml of DMF.

The yield of the previous coupling is determined as described in step A4.

5. Coupling of the second random residue: As for step C3.

6. Deprotection and washes: As for step C4.

7. Coupling of the third random residue: As for step C3.

8. Deprotection and washes: As for step C4.

9. Final washes and drying of the resins: The resins are washed 3 times with 1 ml of DCM, 3 times with 1 ml of CH_3OH and 2 times with 1 ml of ethyl ether. The resins are then vacuum-dried.

10. Cleavage: A TFA-H_2O-thioanisol-phenol-TIS (86:3:3:4.5:1.5, w/w) mixture is freshly prepared and 800 μl added to each sample. After vortexing for 5 h, the resin residues are removed by filtration through 1 ml pipette tips plugged with glass wool. Adding 1.5 ml of cold diethyl ether to each tube, the peptide mixtures are precipitated.

---

[10] The resin-bound peptide libraries are obtained as described in section A, but the acid-stable linker PAM is used to prevent the cleavage of the libraries from the resin during the final TFA treatment.

The precipitates are centrifuged and the ether removed. The crude materials recoveredafter centrifugation are once again washed with 1 ml of cold diethyl ether, dissolved in 0.5 ml of H_2O-CH_3CN-TFA (50:50:0.1) and lyophilized.

10 mg/ml stock solutions of the peptide libraries are usually prepared in DMSO and stored in sealed vials frozen at −80° C until required.

References

1. Furka, A., Sebestyen, M., Asgedom, M. and Dibo, G. 1991. *Int. J. Pept. Protein. Res.*, vol. 37, p. 487.
2. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. and Cuervo, J. H. 1991. *Nature*, vol. 354, p. 84.
3. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J. 1991. *Nature*, vol. 354, p. 82.
4. Merrifield, R. B. 1963. *JACS*, vol. 85, pp. 2149-2154.
5. Fassina, G., Verdoliva, A., Odierna, M. R., Ruvo, M. and Cassani, G. 1996, *J. Mol. Recog.*, vol. 9, pp. 564-569.
6. Geysen, H. M., Rodda, S. J. and Mason, T. J. 1986. *Mol. Immunol.*, vol. 23, pp. 709-715.
7. Eichler, J., Bienert, M., Stierandova, A. and Lebl, M. 1991. *Pept. Res.*, vol. 4(5), pp. 296-307.
8. Kramer, A., Volkmer-Engert, R., Malin, R., Reineke, U. and Schneider-Mergener, J. 1993. *Pept. Res.*, vol. 6, pp. 314-319.
9. Fodor, S. P. A., Read, J. R., Pirrung, M. C., Stryer, L., Lu, A. T. and Solas, D. 1991. *Science*, vol. 251, pp. 767-773.
10. Eichler, J. and Houghten, R. A. 1993. *Biochemistry*, vol. 32, pp. 11035-11041.
11. Pinilla, C., Appel, J. R., Blondelle, S. E., Dooley, C. T., Doerner, B., Eichler, J., Ostresh, J. and Houghten, R. A. 1995. *Biopolymers (Peptide Science)*, vol. 37, pp. 221-240.
12. Pinilla, C., Appel, J. R. Blanc, P. and Houghten R. A. 1992. *Biotechniques.*
13. Gutte, B. and Merrifield, R. B. 1971. *J.Biol. Chem.*, vol. 246(6), pp. 1922-1941.
14. Wang, S. S. 1973. *J. Am. Chem. Soc.*, vol. 95(4), pp. 1328-1333.
15. Veber, D. F., Milkowski, J. D., Varga, S. L., Denkewalter, R. G. and Hirschmann, R. 1972. *J. Am. Chem. Soc.*, vol. 94(15), pp. 5456-61.

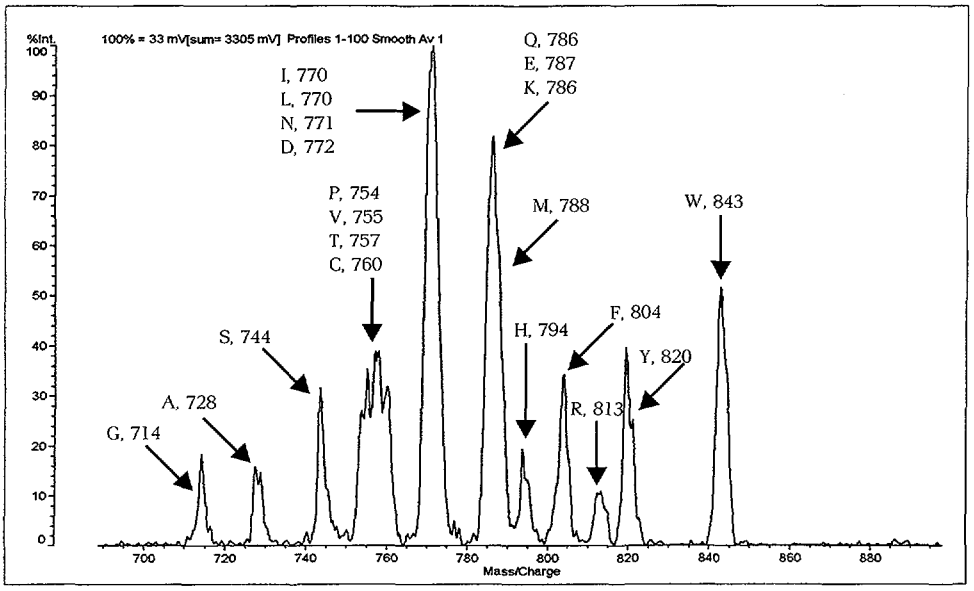
3

Peptide library characterization by amino acid analysis and TOF-MALDI mass spectrometry

M. Ruvo

3 Peptide library characterization by amino acid analysis and TOF-MALDI mass spectrometry

Menotti Ruvo



- Amino acid analysis and TOF-MALDI mass spectroscopy are versatile tools to assess amino acid distribution in a peptide library
- The methods are rapid, efficient and require a very small amount of libraries
- Amino acid analysis can be applied to both soluble and support-bound libraries

3.1. Introduction

The introduction of the combinatorial libraries has accelerated the drug discovery process, making available large equimolar collections of synthetic compounds relatively quickly and cheaply. New leads able to recognize biological targets are identified exploring the molecular space occupied by these collections of molecules [1]. From the beginning, peptide libraries have played a significant role in the development of the combinatorial technique. From a synthetic point of view, peptides represent one of the best-studied classes of compounds. In fact, applying the solid-phase method [2], peptides of very different sizes and structures, and in good yields and purity, can be obtained by carrying out a few simple reaction steps. Some methods for the production of peptide libraries have been developed on the basis of the solid-phase synthesis principle. The best known is the portioning-mixing method first described by Furka [3] and developed by many others [4, 5]. The method involves performing cycles of coupling, mixing and re-splitting of the resin. Repetition of the procedure for x times leads to peptide sequences having x randomized positions. If n is the number of monomers (amino acids) used to build the library, the total number of different peptides obtained (N_T) is:

$$N_T = n^x$$

The application of this method is mandatory when a one-bead-one-peptide [5] format is required. However, unless working with kilograms of resin, this method is limited to the preparation of peptide libraries containing up to six random residues. A more versatile and rapid method for the production of peptide libraries is the pre-mix method [6, 7]. In this case, mixtures of activated amino acids are concurrently coupled to a single batch of resin, with no splitting or re-mixing. Libraries obtained by applying this technique are screened in solution after detachment from the resin beads. Since the number of resin particles does not limit the number of different sequences, highly diversified libraries can be produced even working on small synthesis scale [8]. Fundamental prerequisites for screening a good peptide library are the purity of the compounds produced and an equimolar representation of the compounds in the final mixtures.

Both the portioning-mixing and pre-mix methods fulfil this need, even though a correct application of the latter often requires a correction in the relative concentration of the activated monomers used during couplings.

Although using the term "purity" in conjunction with mixtures suggests a contradiction in terms, in this context, purity refers to the absence of undesired side-products (truncated sequences, deletions and by-products derived from the cleavage reaction) which, to some extent, are always generated in the preparation of libraries.

While the presence of a large population of side-products could represent a further contribution to the diversity of a library and consequently a higher probability to find active molecules, conversely, the formation of so many impurities could result in a library difficult to decode or libraries lacking one or more of the desired components.

A well-characterized library is a good starting point for successful screening. Analytical investigations should be carried out in order to identify, when present, side-products or to assess the equimolarity of components. Techniques such as amino acid analysis [9], mass spectrometry (MALDI and ESI) [10, 11, 12], sequencing [13], HPLC [9] and CE [14] are useful tools for the characterization of peptides. Their application to libraries is easily extended and does not require additional skills. Amino acid analysis is a versatile and inexpensive method for peptide library characterization. Although not sufficient on its own to give conclusive information on the actual composition, this analysis represents a rapid approach to evaluate the distribution of components in a

peptide mixture. It can be applied with both soluble (lyophilized) and support-bound libraries, since the conditions of hydrolysis are strong enough to remove the peptides from any kind of resin or other solid surfaces such as 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. HPLC and CE alone can give only limited information on the composition of a complex mixture (only mixtures containing up to a few dozen different compounds can be efficiently separated). Nevertheless, when they are coupled to mass spectrometry (LC-MS and CE-MS) the separation power increases significantly, thus allowing the characterization of even very complex mixtures [15]. Pool sequencing (or multiple sequence analysis) has been successfully used to characterize peptide libraries and has proved also very promising in the screening step. The main drawback with this technique is that it requires expensive instruments and expertise that is not commonly found in peptide synthesis laboratories. Mass spectrometric methods like ESI and MALDI are also currently applied to structure elucidation in complex peptide mixtures, mainly when coupled to tandem mass analysers for the identification of side-products derived by the cleavage step. Mass spectrometry finds the best application when coupled with chromatographic and electrophoretic methods. In this instance, the ESI technique is particularly useful since it can be directly interfaced to HPLC and CE (allowing the identification of completely overlapped peaks) or used in conjunction with an autosampler for the analysis of multiple samples. MALDI has to be used off-line and analysis cannot be automated, but has the advantage of being more sensitive than ESI. The following sections describe the application of amino acid analysis and TOF-MALDI mass spectrometry to the analysis of peptide mixtures. An additional section is dedicated to a discussion of the problems more frequently encountered when applying these methodologies to peptide library characterization.

3.2. Materials and methods

3.2.1. Materials

- HPLC system (pumps, gradient-makers, injectors and UV-VIS variable wavelength detectors)
- autosampler-HPLC system
- data recorder (D-2500)
- fraction collector
- columns for amino acid analysis
- columns used for RP-HPLC analysis of libraries
- TOF-MALDI mass spectrometer Kompact MALDI III
- RF-551 spectrofluorimeter (used for Fmoc-amino acid detection)
- 9-fluorenylchloroformate (Fmoc-Cl)
- caffeic acid
- sinapinic acid
- trifluoroacetic acid (TFA)
- constant boiling 6N HCl
- 1.5 ml glass vials for amino analysis
- HPLC-grade acetonitrile
- amino acid analysis-grade pentane
- amino acid analysis-grade acetone

3.2.2. *Methods*

A. Amino acid analysis

1. Hydrolysis: Around 100 mg of all dimeric peptide sub-libraries (18 sub-libraries)[11] or amounts of resin corresponding to the same net peptide quantity, are placed in glass vials suitable for amino acid analysis.

2. 100 μl of 6N HCl are added to each sample.

3. The vials are sealed under vacuum using a membrane pump and a gas burner.

4. The vials[12] are left for 1 h at 150° C.

5. Fmoc-derivatization: 10 μl aliquots are vacuum dried and dissolved in 250 μl of 0.5 M borate buffer pH 7.7.

6. The samples are placed in an autosampler-HPLC system that automatically performs the Fmoc-derivatization[13] with Fmoc-Cl as follows:

250 μl of a 15 mM solution in acetone of Fmoc-Cl are added to sample and mixed. After 45 s reaction, the reagent excess is extracted twice with 900 μl of pentane.

7. RP-HPLC-analysis: After pentane removal, 50 μl of the resulting solutions are automatically injected on to a 125 x 4.6 RP-8 LiChroCart (Superspher 60,5 μm) HPLC column and analysed in the following conditions:

flow: 1.25 ml/min (typical operating pressure 100 atm)

temperature: 45° C

gradient (see table 3 below)

Table 3.1. Gradient

	Time (min)				
Eluents (%)	0	30	35	45	50
A	100	50	0	0	100
B	0	50	0	0	0
C	0	0	100	100	0

[11] This amount of peptide library, assuming that all amino acids are equimolarly represented and have a common average MW of 110 amu, corresponds to around 50 nmol each component. After derivatization, 5 nmol are injected and analysed. Such a quantity is more than enough for an amino acid analysis performed using the described method even when detection is carried out using UV at 263 nm.

[12] Usually hydrolysis is carried out for 24 h at 110° C. 150° C allows hydrolysis time to reduce to 1 h. If shorter times are applied, the presence of di- or tri-peptides is observed. A vacuum is required to suppress or reduce oxidation phenomena caused by oxygen on the side-chain of tyrosine, histidine and, when present, triptophan.

[13] The method is based on a pre-column Fmoc derivatization of the amino acid mixtures. Introduction of the Fmoc moiety allows the separation by reverse-phase analysis instead of the classical two steps ionic exchange (cationic and anionic) method. The eluted derivatized amino acids are conveniently detected at 263 nm, since at this wavelength a maximum in the absorption spectrum of the Fmoc group occurs. Assuming a similar reaction rate of all amino acids toward Fmoc-Cl, peaks of the same area are obtained in the chromatogram. In fact, differences are observed in peak areas due to slightly different reaction rates and solubility in pentane. These differences are annulled, introducing a correction factor that can be derived by the relative areas in the chromatogram of the standard. Using Fmoc derivatization, detection can be fulfilled also by fluorimetry. The Fmoc group has a main band of absorption at 263 nm. The energy absorbed at this wavelength is then emitted as fluorescence at 313 nm. Since all amino acids bear the same absorbing or fluorescent group, no errors due to differences in the absorption coefficient can be introduced. Tyrosine and triptophan are also fluorescent molecules but they have absorption and emission maxima at 274/303 nm and 280/348 nm respectively.

Eluents:

A = H_2O/CH_3CN 80:20, 40 mM CH_3COONa pH 4.6

B = H_2O/CH_3CN 20:80, 10 mM CH_3COONa pH 7.0

C = CH_3CN

Monitoring using either UV at λ = 263 nm (cell path length 11 mm) or fluorescence with λ_{ex} = 263 nm, λ_{em} = 313 nm (gain set to 1, sensitivity on the low position and time constant 1.5 s. Instrument equipped with a quartz flow cell having a path length of 12 mm). A 10 μl sample of a 2.5 mM standard solution[14] of amino acids in 0.1 M HCl (standard H) is dissolved in 250 $\mu\mu$l of borate buffer 0.5 M, pH 7.7 and processed in the same way. Data are collected and stored on a D-2500 Merck-Hitachi data recorder.

8. Data handling: Amino acids are identified by comparing retention times. Compositions are obtained by comparing peak areas of the different Fmoc-amino acids. A correction factor[13] is introduced in order to take into account the reaction rates of different amino acids with Fmoc-Cl. No corrections are introduced to take into account the incomplete hydrolysis of some resistant amide bonds, such as the Ile-Ile or Ile-Val bonds.

B. TOF-MALDI mass spectrometry

1. Analysis of crude libraries: 0.5 μl matrix samples[15] (sinapinic acid, 200 mg in 20 ml of H_2O/CH_3CN 2:3) are spotted on the micro-wells of the steel sample holder.

2. The solvent is allowed to dry in a forced-ventilation warm oven.

3. 0.5 μl of a 0.5 mg/ml solution in H_2O/CH_3CN 1:1, 0.1% TFA of the 20 crude sub-libraries are added on the same micro-wells.

4. The solvent is dried and 0.5 μl of matrix solution is again spotted and let dry.

5. The sample holder is placed in the instrument and left until a good vacuum is reached.

6. 200 acquisitions are recorded and automatically averaged for each sample. For libraries of this size, spectra can be conveniently acquired in the reflectron mode (0.04% accuracy). The laser power is tuned in such a way that a good signal for all components is obtained at the lowest possible energy level.[16,17]

---

[14] The standard solution of amino acids is an equimolar mixture of Ala, Asp, Glu, Phe, Gly, His, Ile, Leu, Met, Pro, Arg, Ser, Thr, Val, Lys, Tyr at 2.5 mM and cystine 1.25 mM). After Fmoc derivatization, 2.5 nmol of each amino acid are injected and analysed.

[15] A different matrix (caffeic acid or hydroxycinammic acid) can be chosen if a poor signal is obtained. Matrices are conveniently dissolved in a mixture of water and a water-miscible, volatile organic solvent.

[16] No linear correlation exists between observed peaks in the mass spectrum and the relative concentration of the mixture components. It has been observed that highly (positively) charged peptides (arginine and lysine containing peptides, for example) are more rapidly and efficiently ionized and produce higher peaks in the spectrum. In our case, all molecules have a common arginine residue and a good correlation between the relative concentration and peak heights is observed.

[17] Working with a laser power that is too low does not allow the detection of some poorly ionizable components. Too-high laser power, conversely, produces a noisy spectrum. If the laser power exceeds a certain limit, fragmentation can be induced. This results in a population of daughter peaks that makes the interpretation of results difficult.

7. Analysis of purified libraries: 40 μl samples of a 10 mg/ml solution in DMSO (400 μg of total peptide, 20 μg each) of one sub-library (containing 20 peptides) are injected on to a RP-18 250 x 4.5 mm ID Vydac RP-18 HPLC column equilibrated at a flow rate of 1.0 ml/min with 3% CH_3CN in H_2O (0.1% TFA).

8. After injection, a linear gradient from 3% to 95% of CH_3CN, 0.1% TFA in 50 min is applied to elute the peptides. The column eluate is monitored using an UV detector (cell path length: 11 mm) set at 214 nm.

9. Using a flow splitter and a fraction collector, 96 fractions (0.5 min, 50 μl) are collected in an ELISA plate, starting from 5 min.

10. 1 μl samples are taken from fractions corresponding to chromatogram regions where peaks are present. They are then spotted on the mass spectrometer sample holder as described in steps B1-B6.

References

1. Fassina, G., Verdoliva, A., Odierna, M. R., Ruvo, M. and Cassani, G. 1996. *J. Mol. Recog.*, vol. 9, pp. 564-569.
2. Merrifield, R. B. 1963. *J. Am. Chem. Soc.*, vol. 85, pp. 2149-2154.
3. Furka, A., Sebestyen, M., Asgedom, M. and Dibo, G. 1991. *Int. J. Pept. Protein. Res.*,vol. 37, p. 487.
4. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. and Cuervo, J. H. 1991. *Nature*, vol. 354, p. 84.
5. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J. 1991. *Nature*, vol. 354, p. 82.
6. Pinilla, C., Appel, J. R., Blanc, P. and Houghten, R. A. 1992. *Biotechniques*, vol. 13(3), pp. 412-421.
7. Ivanetich, K. M. and Santi, D. 1996. *Methods in Enzymology*, vol. 267, pp. 247-260.
8. Wiesmueller, K.-H., Feiertag, S., Fleckenstein, B., Kienle, S., Stoll, D., Hermann, M. and Jung, G. 1996. *Combinatorial Peptide and Nonpeptide Libraries: A Handbook* (G. Jung; ed.), VCH Weiheim, New York and Basel, pp. 203-246.
9. Ruvo, M., Scardino, P. , Cassani, G. and Fassina, G. 1994. *Peptide and Protein Letters*, vol. 1, No. 3, pp. 187-192.
10. Metzger, J. W., Kempter, C., Wiesmueller, K.-H. and Jung, G. 1994. *Anal. Biochem.*, vol. 219, pp. 261-277.
11. Winger, B. E. and Campana, J. E. 1996. *Rapid Commun. Mass Spectrom.*, vol. 10, pp. 1811-1813.
12. Davis, M. T., Stahl, D. C., Hefta, S. A. and Lee, T. D. 1995. *Anal. Chem.*, vol. 67, pp. 4549-4556.
13. Stevanovic, S. and Jung, G. 1993. *Anal. Biochem.*, vol. 212, pp. 212-220.
14. Boutin, J. A., Hennig, P., Lambert, P.-H., Bertin, S., Petit, L., Mahieu, J.-P., Serkiz, B., Volland, J.-P. and Fauchère, J.-L. 1996. *Anal. Biochem.*, vol. 234, pp. 126-141.
15. Metzger, J. W., Wiesmueller, K.-H., Kienle, S., Brunjes and Jung, G. 1996. *Combinatorial Peptide and Nonpeptide Libraries: A Handbook* (G. Jung; ed.), VCH Weiheim, New York and Basel, pp. 247-285.

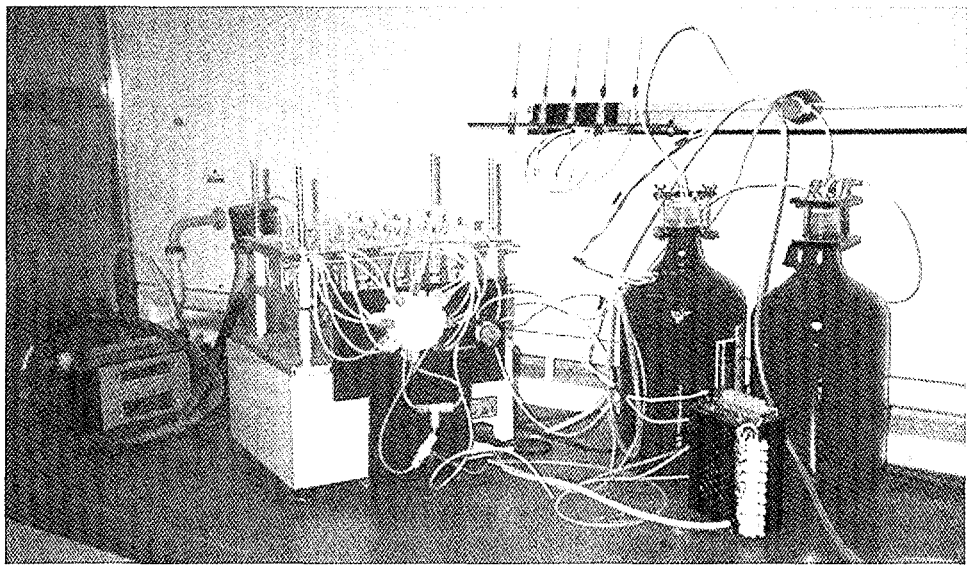
4

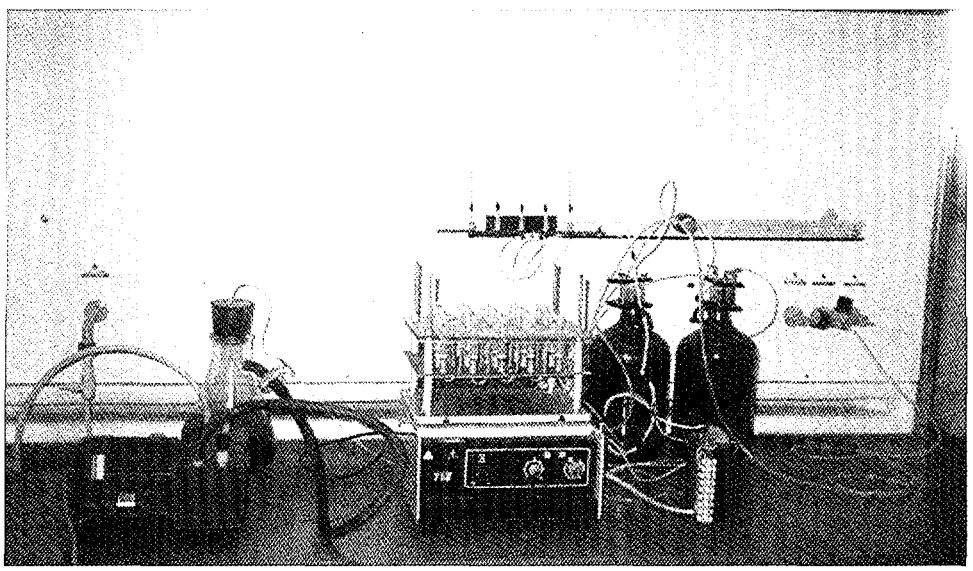
Semi-automatc apparatus for the solid-phase synthesis of peptide libraries

M. Ruvo, N. Campanile, A. Scarallo and N. Arena

4 Semi-automatic apparatus for the solid-phase synthesis of peptide libraries

Menotti Ruvo, Nicola Campanile, Angela Scarallo and Nicola Arena





Photographs courtesy of M. Ruvo

- Semi-automatic apparatus for the parallel synthesis of peptides
- Composed by a shaking module, a sucking module and a dispensing module
- Modules can be assembled starting from relatively inexpensive commercially available materials, such as a membrane pump, polypropylene cartridges, Teflon tubing and connectors

4.1. Introduction

4.1.1. The solid-phase method

Solid-phase synthesis [1] is one of the most reliable and rapid techniques for preparing molecules in high yields and purity. The method is based on the rapid and efficient removal of reagent excess from the reaction mixture once the synthesizing processes are complete. This is achieved by carrying out the chemical transformations needed to obtain the desired product on an insoluble support. Since the growing molecule remains attached to the insoluble polymer, soluble reagents are simply filtered off and the solid phase washed using a suitable solvent. This simple principle allows the use of large excesses of reagents that help to increase reaction rates and yields. The method has been successfully applied to the synthesis of peptides, oligonucleotides, sugars and, more recently, small organic molecules. For the synthesis of peptides, the method requires three basic steps:

- Deprotection (removal of a temporary N-terminal protecting group on the amino acid anchored to the solid support);
- Coupling (addition of the next protected residue);
- Washing (removal of reagents excesses by washing the product-bearing support with a solvent).

These operations are repeated for each residue to be added into the growing molecule and the final product is then detached from the resin and made free of all other protecting groups on the side-chains. The combination of the solid-phase method with a variety of efficient protecting groups and activating agents allows the preparation of even very large peptide fragments in a short time and with very good results [2]. The synthesis of peptides is in most cases performed using fully automated synthesizers where all operations are optimized in terms of time and solvent consumption. The manual method, however, is still a valuable approach in the preparation of short sequences; it has also been applied in the development of techniques of parallel peptide synthesis [3, 4]. The techniques were introduced to satisfy the increasing demand for synthetic peptides mainly for use in immunological investigations. This later became automated and a number of multiple peptide synthesizers are now available commercially.

4.1.2. Peptide libraries

Due to the simplicity and reliability of the solid-phase method and to the robustness of the chemistry supporting the entire process, peptide synthesis has been used as a source of collections of new compounds (peptide libraries) to be then screened in the search for new ligands of biologically relevant targets. Peptide libraries can be prepared either by parallel synthesis (small libraries of hundreds of peptides) or by combinatorial technologies (libraries of up to millions of different peptides). The combinatorial chemistry technique is, in fact, an extension of parallel synthesis, since it is achieved by simply performing additional steps of mixing and splitting of the resin (mix-and-split method [5, 6]) or mixing solutions of different amino acids (pre-mix method [7]). Nearly all the instrumentation devised for parallel synthesis can be used in the mix-and-split or pre-mix methods. Such techniques require further manual intervention as well as a number of approaches that exploit manual methods [7] (or the fully automatic multiple peptide synthesizers described). Whichever adaptations are necessary, the steps of mixing and splitting the resin or mixing the amino acid solutions must

be performed manually, since no machines available commercially are able to achieve these operations. Very good results can be obtained by manual methods [8], but the disadvantage is that they are time-consuming and require many repetitive, tiresome and error-inducing operations such as washing and amino acid activation. Multiple synthesizers are very expensive machines that perform automatically these and many other steps, but weighing and making solutions of dozens of amino acids for a combinatorial synthesis, arranging them on to the synthesizer, entering locations of each into a software (other than mixing resins or amino acids), remain time-consuming manual operations.

4.1.3. *Self-made semi-automatic apparatus*

There is a solution for those laboratories interested in the preparation of peptide libraries that do not want to invest money in automatic instruments but, at the same time, want to speed up the process. A self-made semi-automatic apparatus can present a compromise that requires only a small investment in terms of time and money. Such an apparatus, intended for the Fmoc-based chemistry (the TFA employed in t-Boc chemistry prevent the use of open vessels), can be assembled from inexpensive components that can be adapted to fulfil the basic operations for parallel peptide synthesis such as shaking, liquid dispensing and liquid removal. Thus the minimum requirements for a semi-automated multiple peptide synthesizer include a shaking device, a dispensing device and a sucking device. The shaking device is needed for holding and shaking reaction vessels, the dispensing device is needed for fluid dispensing and the sucking device is needed for fluid removal.

The shaking device must ensure that shaking occurs in such a way that all resin particles are submerged in the reaction and washing solutions. This can be achieved by tilting the reaction vessels or simply by vortexing them. The latter is desirable in order to avoid the need to use a leak-proof stopper for the reaction vessel after adding solvent or reagent. Since vortexing can be vigorous, reaction vessels are tightly immobilized on the shaking device. The device also works as a holder when reagents are added. Shaking can be achieved by using a vortexing system on which the reaction vessels (usually cartridges endowed with filtration septa) are assembled in rows and columns. The dispensing device must ensure a correct delivery of all reagents and solvents (the piperidine/DMF solution for deprotection and DMF for washing) in the desired amounts and in such a way that the operation is rapid and clean (cross-contamination must not occur while fluid is being dispensed). A multi-needle apparatus connected to solvent and reagent reservoirs, and equipped with valves to switch between one solvent and the others, can be used as a dispensing device. The transfer of solvent can be achieved by applying slight gas pressure (such as nitrogen or other inert gases) on the liquid surfaces. The sucking device, that could be simply a membrane pump, must ensure that fluids are thoroughly and rapidly removed from the resin. The sucking capacity can be regulated using a vacuum trap and a valve to avoid breaking the cartridge filters. The sucking device can be connected to the reaction vessels by a network of tubes through which the solvents and reagents in excess are collected (the same vacuum trap can be used for this purpose).

4.2. Materials and methods

4.2.1. Materials

- vortex with a 30 x 20 cm aluminium plate, a shaking power regulator and a timer
- membrane pump with vacuum rubber tubes
- 8 cm stainless steel needles (Reacti-VAP™), bearing a female luer lock
- glass vacuum trap with rubber stopper
- Teflon tubing (1/8 in. OD, 2.4 mm ID)
- Spe-ed™ polypropylene cartridges (3 and 6 μl) with 5 m PTFE frits and female luer lock gauge
- peek connectors with female luer locks on one side (into which the cartridge tips can be lodged) and female screws on the other side (able to lodge a peek nut)
- peek nuts and suitable ferrules fitting the 1/8 in. Teflon tubes to achieve all connections between reaction vessel, valves and needles
- polypropylene caps with two Teflon tubes (1/8 in. OD, 2.4 mm ID, the first one for gas pressure the second one for solvent delivery) that fit on to one gallon DMF brown bottles
- 4-, 6- and 8-way Teflon connectors
- 3-way rotary electric Teflon valves to achieve the switch between the piperidine and DMF
- membrane ON/OFF electric valves to control fluid dispensing or to close the gas connection on to the piperidine bottle

4.2.2. Methods

A. Assembling the shaking device[18]

1. Attach four 10 cm stainless steel rods with screw ends vertically to the aluminium plate corners. Insert the rods into 5 mm ID holes and fix tightly using threaded nuts.

2. Fix two additional plates that are the same size (30 cm x 20 cm) as the vortex plate (using threaded nuts) at 3 cm and 8 cm, respectively, from the surface of the original plate. On to the lower additional plate, arrange 20 equally distributed 4 mm ID holes (previously made), in a matrix-like disposition of 4 rows and 5 columns.

3. Tightly fit on each hole a peek connector by screwing a peek nut from the bottom (the ferrules are in polypropylene) bearing a 1/8 in. Teflon tube (about 20 cm long).[19]

[18] The automatic apparatus is built in such a way that 20 parallel synthetic steps can be performed.
[19] This tube serves as a draining line for the solvents removed by suction. Each tube is the same length to avoid back-pressure differences in the lines leading to different draining rates.

4. On the upper aluminium plate, centre 20 1 cm holes on the corresponding smaller holes on the lower plate. Place 20 polypropylene cartridges (6.5 cm high, 1 cm OD) into these holes, fitting their luer endings into the luer-shaped holes of the corresponding peek connectors fixed on to the lower plate.[20]

B. Assembling the sucking device

1. Combine the 20 Teflon tubes from the bottom of the cartridges (see step A.4) in a single outlet line (connected to the vacuum trap) with three 8-way Teflon connectors symmetrically joining each other using a 4-way Teflon connector.[21, 22, 23]

2. Connect the outlet line to an ON/OFF electric valve using screwed peek nuts and polypropylene ferrules.[24]

3. The outlet line from the ON/OFF electric valve is then joined to the vacuum trap by inserting it into a hole made in the rubber stopper.[5]

4. Connect the vacuum trap to the membrane pump using a vacuum rubber tube.

C. Assembling the solvent delivery device[26]

1. Make 5 equidistant 4 mm ID holes in an aluminium bar (30 cm x 2 cm x 0.2 cm) in such a way that they fit the holes of a line of cartridges on the shaking device.

2. Fix on each hole a peek connector as described in step A3.

3. On the female luer-shaped side of each peek connector a needle is joined using a male luer to male luer peak connector.

4. On the screw female side of the first peek connectors, connect Teflon tubes with screwed peek nuts and polypropylene ferrules. Bring the other ends of the Teflon tubes into a single incoming line using a symmetrical 6-way Teflon connector.

---

[20] So assembled, the cartridges remain tightly fixed into the made-to-measure lodgings. (The latter ensures the cartridges are not pushed off during vortexing.)

[21] Connections between connectors are achieved using Teflon tubes (1/8 in. OD), screwed peek nuts and polypropylene ferrules.

[22] Assembly must be symmetrical to avoid a non-uniform removal of solvents due to back-pressure differences in the different lines. The only way to achieve this is to join two 8-way connectors (each bearing 7 draining lines) to two opposite holes of the 4-way connector. The third 8-way connector, on which 6 draining lines are connected, is mounted on to the hole opposite the one connected to the vacuum trap. On the last connector, one way remains closed, since there are only 20 draining lines.

[23] All connections between tubes and connectors are achieved using peek nuts and polypropylene ferrules that fit the tube diameter.

[24] This valve serves to prevent draining of solvents by gravity from the reaction vessels while reagent is added and vortexing.

[25] The vacuum trap should be able to contain several litres of waste. It should not be filled up to its maximum capacity as it is important to avoid solvents being sucked directly into the pump.

[26] The solvent delivery device is made from 5 needles assembled in parallel and spaced in such a way that they fit the holes of a line of cartridges on the shaking device.

5. Connect this unique tube to an ON/OFF 2-way electric valve using the same nuts and ferrules described previously.[27]

6. Connect the tube coming out from this valve to the outlet of a 3-way rotary electric valve.[28]

7. To the two inlet holes of the 3-way electric valve, connect the delivery tube from the polypropylene caps fixed on to the DMF and piperidine solution bottles.[29,30]

8. Connect the pressure line from the cap to the piperidine bottle to an ON/OFF 2-way electric valve.[31]

9. Connect the two pressure lines from this last valve and that from the DMF cap to a nitrogen cylinder using a Teflon 3-way connector.[32]

---

[27] The controlling button of the valve is placed on the same bar holding the needles. The button must be accessible and easily operated in order for the valve to be opened and closed rapidly during liquid dispensing.

[28] This valve is used to switch between DMF and the piperidine solution.

[29] The caps are equipped with two Teflon tubes. One is used for solvent delivery and is the one connected to the 3-way electric valve, while the second is used to apply a small gas pressure on the liquid surface.

[30] The Teflon sealed caps should fit perfectly on to the top of the bottles.

[31] This valve is needed only when the 20% (or higher) piperidine solution is used as a deprotection reagent since the piperidine vapours can reach the bottle of DMF and dissolve in it. The presence of piperidine in the DMF must be absolutely avoided to prevent polymerization and other undesired side reactions during coupling. This valve can be omitted if a solution of 2% piperidine and 2% DBU is used instead of 20% piperidine.

[32] The cylinder must have a pressure reducer able to work in the range of 1 to 3 bars.

References

1. Merrifield, R. B. 1963. *J. Am. Chem. Soc.*, vol. 85, pp. 2149-2154.
2. Ruvo, M., Villain, M., De Falco, S., Verdoliva, A. and Fassina, G. 1998. *Peptides 1996*, (R. Ramage and R. Epton, eds.), pp. 769-770.
3. Geysen, H. M., Rodda, S. J. and Mason, T. J. 1986. *Mol. Immunol.*, vol. 23, pp. 709-715.
4. Houghten, R. 1985. *Proc. Natl. Acad. Sci.*, vol. 82(15), pp. 5131-5135.
5. Furka, A., Sebestyen, M., Asgedom, M. and Dibo, G. 1991. *Int. J. Pept. Protein. Res.*, vol. 37, p. 487.
6. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J. 1991. *Nature*, vol. 354, p. 82.
7. Ruvo, M., Scardino, P. , Cassani, G. and Fassina, G. 1994. *Peptide and Protein Letters*, vol. 1, No. 3, pp. 187-192.

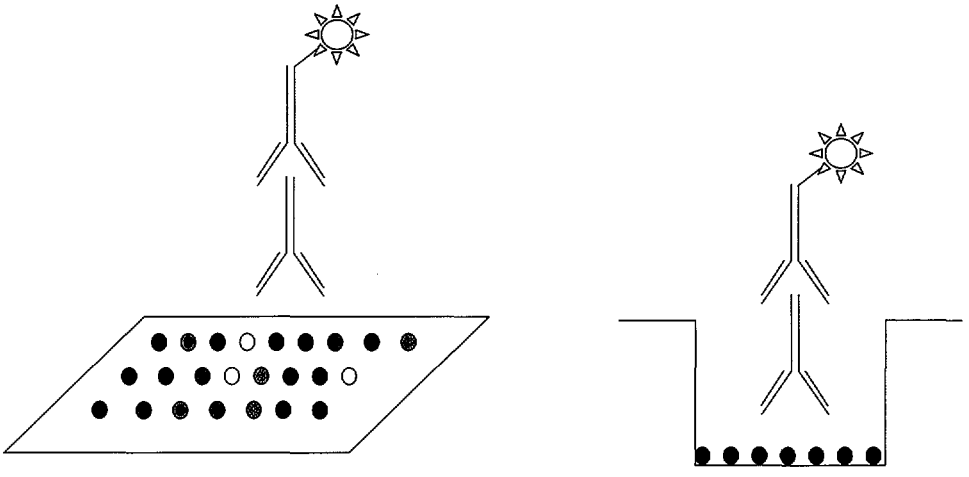
5

Deconvolution of synthetic peptide libraries by non-convalent adsorption to solid supports

M. Ruvo and M. Marino

5 Deconvolution of synthetic peptide libraries by non-covalent adsorption to solid supports

Menotti Ruvo and Maria Marino



- ❑ Useful for selecting ligands
- ❑ Requires low amounts of library
- ❑ Low-cost procedures
- ❑ Amenable to handling many libraries simultaneously

5.1. Introduction

Exploring the molecular space described by large families of new chemical entities that have a synthetic origin is a challenging new way of identifying structural motifs able to recognize biological targets and interfere with their function. Such a process is conceptually straightforward and involves three basic, interrelated steps. The first is the identification of targets represented by a biochemical entity (such as proteins, nucleic acids but also whole cells) that have a demonstrated role in the disease or pathology under study. The second involves the preparation of collections of compounds (libraries) whose chemical structure is known, as in the case of synthetic libraries. In the third step, the library components are concurrently tested in order to select one or more of those that have the desired activity. The choice of the screening assay is dictated by the type of target under study, the type of ligand searched for (agonist, antagonist or simply an affinity chromatography ligand) and also by the type of library available.

Synthetic libraries are obtained by solid-phase methods and represent collections of single compounds or mixtures of compounds (usually having a common structural motif) where all the components are equimolarly represented. Peptide libraries are a classic example of synthetic libraries where huge numbers of different molecular entities are achieved by varying the sequence length and the number of building blocks (amino acids). The complexity of such libraries is around 10^6 (six randomized residues with the 20 natural amino acids) when prepared by the split and mix method [1, 2] but can reach a value considerably higher when obtained by the pre-mix method [3-5]. The first method is essentially adopted for libraries based on the one-bead-one-peptide concept [1, 2, 6], while the second one is more versatile and has a more general application. Depending on the chemistry adopted, peptide libraries can be obtained as a solid or soluble material. The identification of active components in synthetic peptide libraries is achieved in different ways that are linked both to their physical status (soluble or solid) and to the format resulting from the synthetic process adopted for their preparation. Since peptides have a built-in code (the amino acid sequence itself), in both soluble and solid libraries, active molecules can be identified by direct sequencing after isolation [2, 6]. A different method is based on the preparation of sub-libraries where one residue (the N-terminal) is fixed and known while the others are randomized. Cycles of screening and re-synthesis of the most active sub-libraries lead to the decoding of the active structure (one residue per cycle) [2, 7]. A variation of this method is the positional scanning method introduced by Houghten et al. [4] where positions other than the N-terminal can be fixed. Other methods, used essentially with resin-bound libraries, are based on tagging techniques where molecular tags are introduced during library synthesis. Such tags act as a label for the structures to which they are associated. After the selection process, the tags can be isolated easily and the corresponding active molecule identified by analytical techniques [8, 9]. Screening assays on solid libraries prepared on paper or glass allow the identification of active components by localizing their position on the surface on which they have been synthesized [10, 11].

When working with soluble peptide libraries, the activity evaluation is usually accomplished in a biological assay where the ability of the library components to interact with the target is measured using ELISA- or RIA-type assays [7]. Since soluble libraries are amenable to both binding and competition experiments, a broad spectrum of different assay formats can be utilized to carry out the screening. In the most common competition assay format, the target (for example, a soluble protein or a cell-bound receptor) is adsorbed on to the wells of the microtiter plate and the peptide library components are tested as possible competitors of a natural or known ligand [7].

In the present chapter, two simple procedures for the screening of peptide libraries are outlined. The number of screening steps required to completely elucidate the sequence of the active peptide depends on the number of randomized residues. The screening procedure may be applied to searching ligands for monoclonal antibodies or

to defining the antigenic sequences recognized by the antibody. Both procedures exploit a simple binding assay where the peptide pools (sub-libraries) are spotted on a nitrocellulose sheet or adsorbed on the 96-well microtiter plates commonly used for ELISA tests. The library-bound antibody is then detected by a second specific antibody conjugated to an enzyme (a peroxidase) able to process chromogenic substrates, whose concentration is then determined in a multi-wavelength ELISA plate reader.

5.2. Materials and methods

5.2.1. Materials

- ❑ vacuum manifold apparatus (Bio-Dot apparatus Bio-Rad, Milan Italy)
- ❑ shaking-table (Delchimica, Naples, Italy)
- ❑ nitrocellulose membrane pieces cut to size (usually 9 cm x 12 cm) from larger sheets
- ❑ second antibody: alkaline phosphatase labelled anti-mouse IgG (Sigma, Milan, Italy)[33]
- ❑ blocking reagent (Boehringer Mannheim, Monza, Italy)
- ❑ buffer 1: 100mM Tris, 150 mM NaCl pH 7.5 (add 12.1 g Tris to 8.76 g NaCl and make up to 1 litre with deionized water; adjust to pH 7.5 with HCl)
- ❑ buffer 2: blocking solution (add 5 g blocking reagent to 500 ml buffer 1; stir at about 65° C to dissolve blocking reagent)
- ❑ buffer 3: antibody solution (add 6 ml normal goat serum to 0.2 g bovine serum albumin and make up to 200 ml with buffer 2)
- ❑ buffer 4: 100 mM Tris, 100 mM NaCl, 50 mM $MgCl_2$ pH 9.5 (add 2.42 g Tris to 1.17 g NaCl and make up to 150 ml with deionized water; adjust to pH 9.5 with HCl; add 0.95 g $MgCl_2$ to the solution and make up to 200 ml with deionized water)
- ❑ NBT solution (Boehringer Mannheim, Monza, Italy)
- ❑ X-phosphate solution (Boehringer Mannheim, Monza, Italy)
- ❑ PVC microtitre plates (Falcon, cat. no. 3912)
- ❑ ELISA plate reader (Merck, Darmstadt, Germany)
- ❑ bovine serum albumin (BSA) (Sigma, Milan, Italy)
- ❑ Tween 20 (Sigma, Milan, Italy)
- ❑ 30% hydrogen peroxide (H_2O_2) (Sigma, Milan, Italy)
- ❑ O-phenylenediamine dihydrochloride (OPD) (Sigma, Milan, Italy)
- ❑ second antibody: HRP-labelled anti-mouse IgG (Sigma, Milan, Italy)[34]

---

[33] In this assay, a mouse IgG antibody is used as the molecular target. Depending on the origin and isotype of the target antibody, use appropriate detection antibodies.

[34] Tripeptide libraries synthesized using only 18 out of 20 natural amino acids (cysteine and triptophan omitted) contain a total of 5,832 molecules (18^3). The libraries are arranged in 18 mother sub-libraries where the N-terminal residue is known, each containing 324 peptides. In the second screening, the sub-library (324 peptides) with the highest activity in the first step is re-synthesized in such a way that 18 sub-sub-libraries are obtained (each containing 18 peptides), where the N-terminal residue is known. In the third screening, the 18 single peptides belonging to the most active sub-sub-library in the previous step are separately synthesized and tested. In each mixture of peptides all components are equimolarly distributed. Since 10 μg of each pool are applied on to the filter in the first screening (assuming that the library components are all equally and completely adsorbed on to the membrane), 0.5 nmol of each peptide are tested assuming an average molecular weight of 1,100 amu. To keep the concentration constant in the second screening step, 0.55 μg of each sub-sub-library should be spotted on nitrocellulose. In the third screening step, a total of 0.03 μg of each single peptide should be applied to use the same amounts of sub-sub-libraries.

- ❑ dibasic sodium phosphate (Sigma, Milan, Italy): to make a 0.2 M solution add 5.68 g of powder to 200 ml deionized water
- ❑ citric acid (Sigma, Milan, Italy): to make a 0.1 M solution dissolve 4.2 g of powder in 200 ml deionized water
- ❑ substrate buffer: 0.05 M phosphate-citrate buffer pH 5.0 (add 25.7 ml 0.2 dibasic sodium phosphate to 24.3 ml 0.1 M citric acid and bring to 100 ml with deionized water; adjust to pH 5.0, if necessary)
- ❑ OPD solution: dissolve one OPD tablet (Sigma, Milan, Italy) to 11 ml substrate buffer and add 6 μl of fresh 30% H_2O_2 immediately prior to use[33]
- ❑ phosphate buffer: PBS (add 120 g NaH_2PO_4 to 32 g NaOH and make up to 1 litre; adjust to pH 7.2 with a 1 M NaOH solution)
- ❑ coating buffer: 0.1 M carbonate buffer pH 8.5 (add 3.2 g Na_2CO_3 to 5.8 g $NaHCO_3$ and make up to 1 litre; adjust to pH 8.5)
- ❑ blocking solution: PBS-3% BSA (add 0.75 g BSA to 25 ml PBS)
- ❑ washing solution: PBS-T (add 1 ml Tween 20 to 2 litres of PBS)
- ❑ antibody dilution buffer: PBS-0.5% BSA (add 0.25 g BSA to 50 ml PBS)

5.2.2. Methods

A. Screening a synthesized tripeptide library by adsorption to nitrocellulose sheets (blotting method)[34]

1. Clean and dry the Bio-Dot apparatus and gasket prior to assembly.

2. Wet the membrane by slowly sliding it at a 45° angle into buffer 1 and allowing it to soak for 10 min.[35]

3. Assemble the Bio-Dot apparatus as described by the manufacturer.[36]

4. Adjust the flow valve so that the vacuum manifold is open to the atmospheric pressure. Apply 100 μl of buffer 1 to all 96 sample wells using a multi-channel pipette.[37]

5. Fill the appropriate wells with aliquots of library solutions (100 μg for each library) using a minimal sample volume. Allow the entire sample to filter through the membrane by gravity flow. Make sure that the flow valve is positioned at a level below the sample wells. Generally, it takes 30 to 40 min for 100 μl of the antigen solution to filter through the membrane.[38]

---

[35] It is important to completely wet the membrane. This ensures proper drainage of the solutions during operations. The membrane should not be exposed to the atmosphere by extending beyond the edge of the gasket after the Bio-Dot apparatus is assembled. Always use forceps or wear gloves when handling membranes.

[36] Visually inspect the gasket to make sure that the holes are properly aligned. If the gasket is not centred, pull slightly at the corners until it is aligned. When removing the wet membrane sheet from the buffer, let the excess of buffer drain from the membrane. Blotting the membrane on a sheet of filter paper is a simple method of removing excess buffer.

[37] Addition of buffer is necessary to re-hydrate the membrane following the vacuum procedure in step 3. If this step is not performed prior to applying samples, assay results will show haloes or a weak detection signal.

[38] Spot peptide samples in duplicate. Set up negative controls to assess whether any reagent binds unexpectedly to any reagent and to nitrocellulose membrane. As negative control, buffer 1 may be used.

6. Remove the membrane from the apparatus. Cut the membrane to obtain only the portion containing the samples (usually 3 cm x 8 cm), and mark sample positions lightly with a pencil.

7. Incubate the membrane in blocking buffer for 1 h.

8. Wash the membrane in binding buffer for 1 min.

9. Incubate the membrane with the antibody target molecule diluted in buffer 3 at a final concentration of 30 μg/ml for 2 h.

10. Wash 3 times, 5 min per wash in buffer 3.

11. While waiting for step 10, prepare the second antibody solution making a dilution of 1:5000 in buffer 3.

12. Incubate the membrane with the second antibody solution for 2 h.

13. Wash the membrane three times in buffer 3.

14. While waiting for step 13, mix 45 μl NBT solution and 35 μl X-Phosphate solution in 10 ml of buffer 4. This freshly prepared colour substrate solution will be used in the next step.

15. Incubate the filter with the colour substrate solution and allow colour development to proceed in the dark.[39]

16. When the desired intensity spots have developed, stop the reaction by washing the membrane several times with re-distilled water, then dry the filter on paper towels.[40]

17. To obtain a numerical evaluation (expressed in arbitrary units) of the colour intensity in each spot, the spots themselves are scanned with a scanning densitometer and the values calculated by the software associated with the densitometer (Image-Pro Plus, Media Cybernetics) reported as bar plots.

18. The sub-library giving the highest value is re-synthesized, as briefly described in step A1. The screening of the second generation of sub-libraries (sub-sub-libraries) is as described in steps A1-A17.

19. The most active sub-sub-library in the latter screening step is in turn re-synthesized and screened as described in steps A1-A17.

B. Screening synthesized peptide libraries adsorbed on microtiter plates (ELISA method)

1. Dilute the peptide library in the coating buffer in order to obtain a final concentration of 50 μg/ml; add 100 μl of this solution to the microtiter plate and leave overnight at 4° C.[41]

---

[39] The colour precipitate starts to form within a few minutes and the reaction is usually complete within 16 h. Do not shake or mix while the colour is developing.

[40] At this point the filter can be photographed and kept for documentation.

[41] As for the blotting method, each sub-library contains an equimolar distribution of all components. If 50 μg of each pool are used in each data point and, assuming an average molecular weight of 1,100 amu, the library is tested at a final concentration of 2.5 μM per each peptide. To keep the peptide concentration constant in following assays, the concentrations should be 2.7 μg/ml (1.4 μM) and 0.15 μg/ml (7.7 nM) respectively for the second and third screening step. Peptide library samples are always tested at least as duplicate data points (two wells for the same sample). Negative controls can be set up to assess whether any reagent binds unexpectedly and aspecifically to any other reagent and to the plastic surface of plates. As negative control, antibody dilution buffer may be used.

2. Wash the plate three times with PBS.

3. Add 200 μl/well of blocking solution and incubate for 2 h at 37° C.

4. Wash the plate three times with PBS.

5. Dilute the target molecule in PBS-0.5% BSA in order to obtain a final concentration of 20 μg/ml; add to the microtiter plate 100 μl/well of the resulting solution and incubate for 1 h at 37° C.

6. Wash the plate six times with PBS-T.

7. Prepare the second antibody solution performing a 1,000-fold dilution in a PBS-0.5% BSA solution.

8. Add 100 μl/well of freshly prepared second antibody solution and incubate for 1 h at 37° C.[33]

9. Wash the plate 6 times with PBS-T.

10. Add to the microtiter plate a freshly prepared OPD solution (110 μl/well) and let the colour develop for 15 min in the dark.[42]

11. Read the absorbance of the developed colour on a microplate reader at 450 nm.

12. The reaction may be stopped with 25 μl/well of 3N HCl or 3M H_2SO_4. In this case, read the absorbance at 492 nm.

13. The data are obtained already as optical density (OD) values and after subtraction of the corresponding blank lines are reported as a bar plot.

14. The sub-library giving the highest value of OD is re-synthesized. The screening of this second generation of sub-libraries (sub-sub-libraries) is as described in steps B1-B13.

15. The most active sub-sub-library in the latter screening is in turn re-synthesized as briefly described in step A1 and screened as described in steps C1-C12.

[42] Usually 11 ml of OPD solution are sufficient for one plate.

References

1. Furka, A., Sebestyen, M., Asgedom, M. and Dibo, G. 1991. *Int. J. Pept. Protein. Res.*, vol. 37, p. 487.
2. Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M. and Knapp, R. J. 1991. *Nature*, vol. 354, p. 82.
3. Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T. and Cuervo, J. H. 1991. *Nature*, vol. 354, p. 84.
4. Pinilla C., Appel, J. R., Blondelle, S. E., Dooley, C. T., Doerner, B., Eichler, J., Ostresh, J. and Houghten, R. A. 1995. *Biopolymers (Peptide Science)*, vol. 37, pp. 221-240.
5. Geysen, H. M., Rodda, S. J. and Mason, T. J. 1986. *Mol. Immunol.*, vol. 23, pp. 709-715.
6. Lebl, M., Krchnak, V., Sepetov, N. F., Seligman, B., Strop, P. , Felder, S. and Lam, K. S. 1995. *Biopolymers (Peptide Science)*, vol. 37, pp. 177-198.
7. Fassina, G., Verdoliva, A., Odierna, M. R., Ruvo, M. and Cassani, G. 1996. *J. Mol. Recog.*, vol. 9, pp. 564-569.
8. Ohlmeyer, M. H. J., Swanson, R. N., Dillard, L. W., Reader, J. C., Asouline, G., Kobayashi, R., Wigler, M. and Still, W. C. 1992. *Prot. Natl. Acad. Sci. USA*, vol. 89, pp. 5381-5383.
9. Nielsen, J., Brenner, S. and Janda, K. D. 1993. *J. Am. Chem. Soc.*, vol. 115, pp. 9812-9813.
10. Kramer, A., Volkmer-Engert, R., Malin, R., Reineke, U. and Schneider-Mergener, J. 1993. *Pept. Res.*, vol. 6, pp. 314-319.
11. Fodor, S. P. A., Read, J. R., Pirrung, M. C., Stryer, L., Lu, A. T. and Solas, D. 1991. *Science*, vol. 251, pp. 767-773.

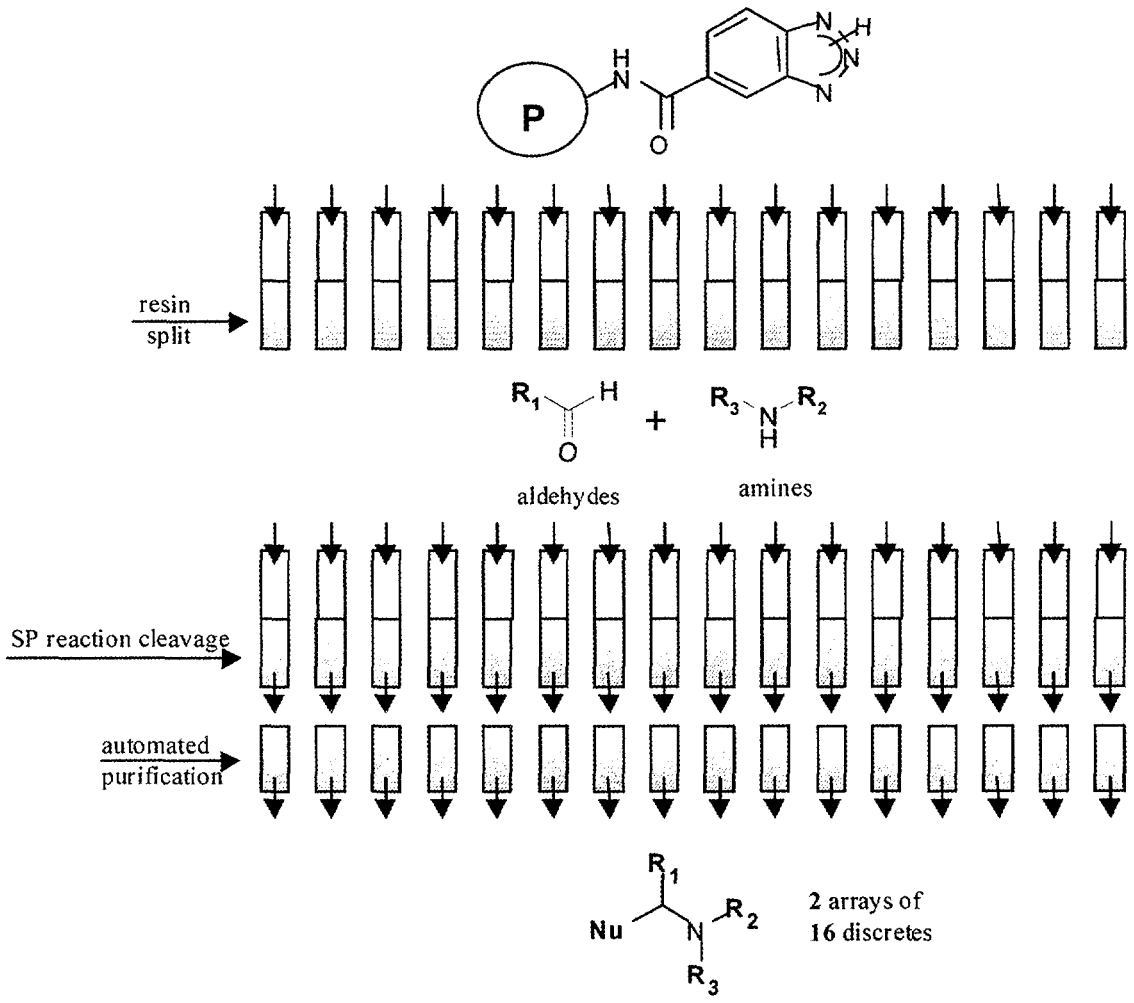
6

Solid-phase synthesis of substituted amine libraries

P. Seneci and A. Paio

6 Solid-phase synthesis of substituted amine libraries

Pierfausto Seneci and Alfredo Paio



- Parallel synthesis of arrays of amines using assessed organic chemistry on solid phase
- Access to remarkable chemical diversity
- Implementation of novel chemistry on solid phase
- Reliable and parallel purification procedures
- Cheap techniques and reagents used

6.1. Introduction

In the present chapter, the synthesis of a small discrete array library on solid phase is described. A growing interest in developing solid-phase synthetic strategies [1-5] for preparing highly differentiated small drug-like molecules in a combinatorial library format [6-10] has been witnessed during the past few years. The development of new linkers [3, 11-17] has increased the compound diversity obtainable from solid-phase chemistry, especially regarding small organic molecules: the so-called traceless linkers [18-21] that leave no residual tethering on cleaved compounds and are for this reason particularly appealing to the combinatorial chemist. Recent advances in the use of solid-supported reagents [22-25] have opened new routes leading to library synthesis, both in solution and on solid phase.

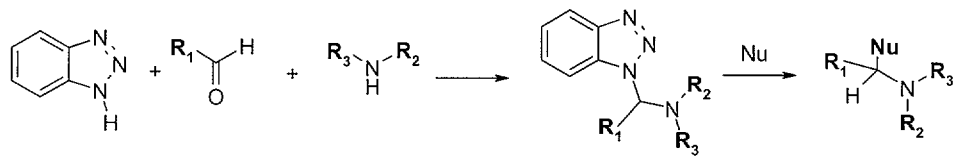
The use of automated or semi-automated purification procedures of final library compounds coming from parallel synthesis of discretes has become very popular [26-29] for producing high purity samples. In this chapter, an example of automated purification via aqueous extraction and ionic exchange chromatography is described.

6.2. Chemical assessment and library design

6.2.1. Reaction scheme

The reaction scheme for the production of functionalized amines in solution using known benzotriazole chemistry [30-32] is illustrated in scheme 1 below.

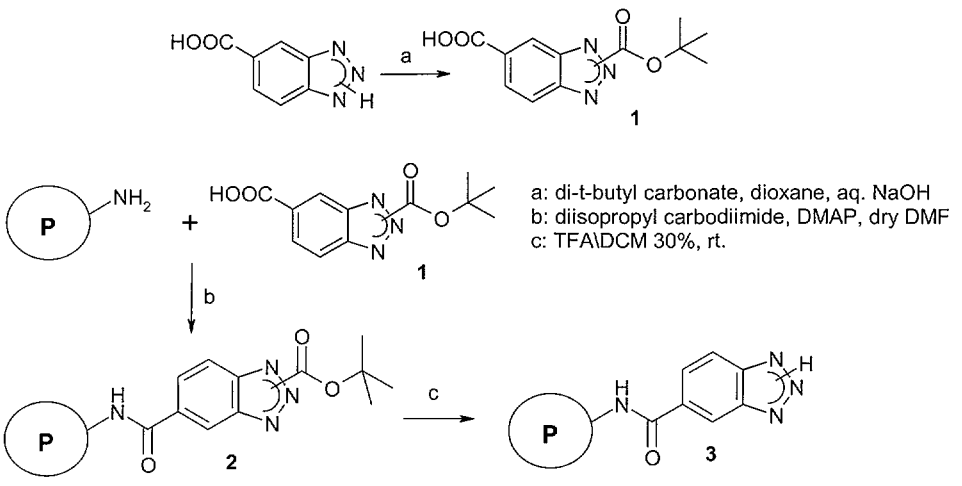
Scheme 1



The possibility of anchoring the benzotriazole on solid phase would allow a faster work-up and purification of the libraries produced using such chemistry and would confer more stability to the intermediate adducts.

6.2.2. Synthesis of the supported benzotriazole

Scheme 2



The synthesis of the polymer-supported key synthon 3 is illustrated in scheme 2. The commercially available 5-carboxybenzotriazole was N-protected to give the synthon 1 which was loaded on to the solid phase using standard solid-phase peptide coupling conditions. The intermediate 3 was obtained by Boc deprotection in acidic conditions.

6.2.3. *Chemical assessment on solid phase*

The significant reactions carried out on solid phase to optimize the reaction conditions to be used for the library synthesis are reported in scheme 3. The supported benzotriazole 3 was reacted with a few aldehydes and amines and the adducts were cleaved with sodium borohydride (Nu=H), allyl (Nu=Allyl) or phenyl magnesium halide (Nu=Ph) in different experimental conditions. Products 7A-E were obtained and the best reaction conditions for adduct formation (b and c, scheme 3) and for cleavage in solution (d, Nu=H; e, Nu=Allyl; f, Nu=Ph) were determined.

6.2.4. *Library structure and monomer selection*

The 30-member array library of discretes is shown in figure 6.1. It is composed of two sub-libraries: in the first, a common amine is reacted with 8 aldehydes, a standard aldehyde is reacted with 8 amines and the 16 adducts (two are identical, see figure 6.1) are cleaved with sodium borohydride (sub-library 1, figure 6.1, as in condition d, scheme 3).[43] The second sub-library is identical except for the cleavage performed with allyl magnesium chloride (sub-library 2, figure 6.1, as in condition e, scheme 3). The analytical quality control for the library is provided in a later section in the chapter.

[43] This library must be intended as a sort of model library, and the exploitation of the assessed SP method to larger libraries is under evaluation.

Scheme 3

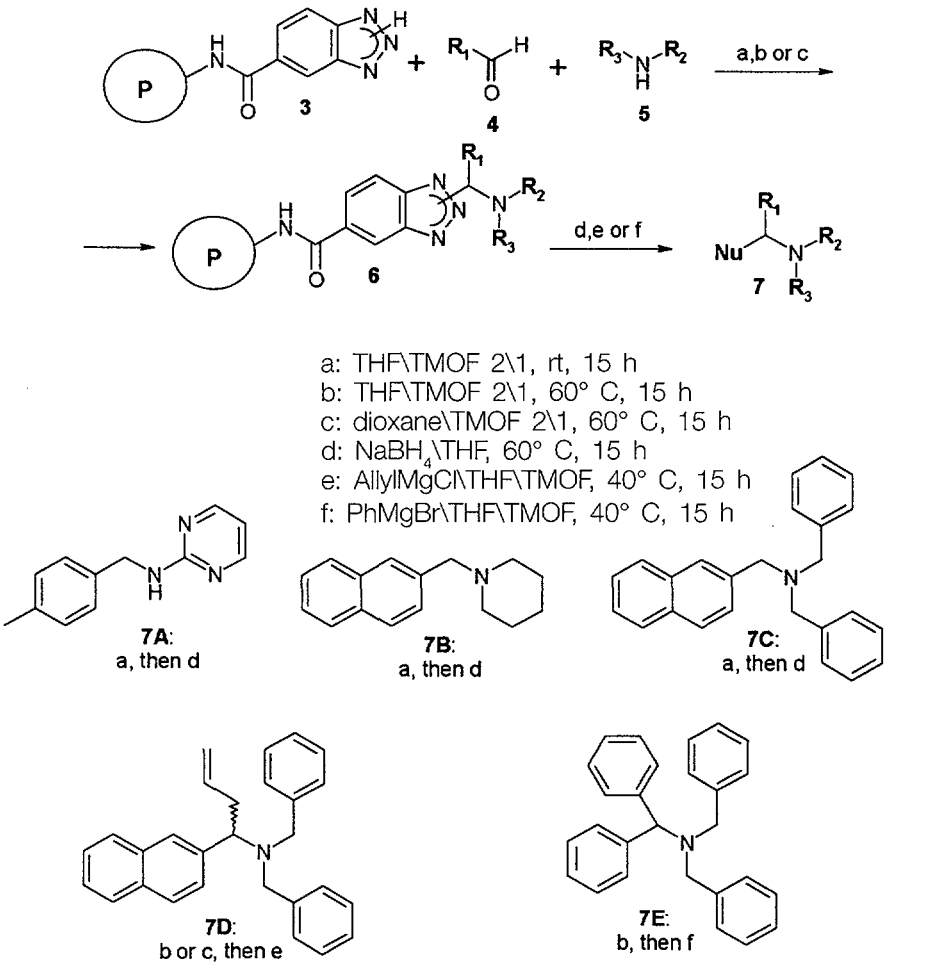
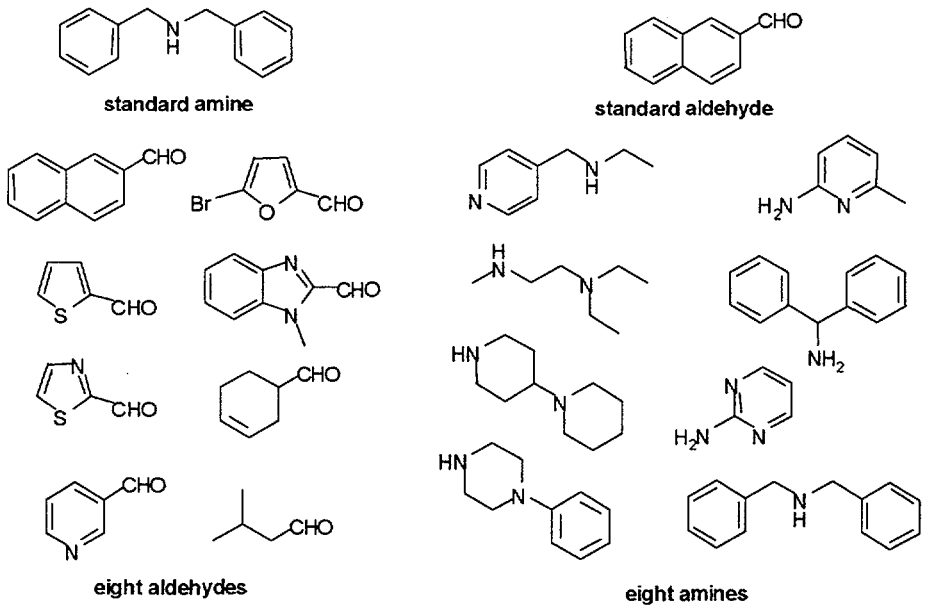


Figure 6.1. A 30-member array library of discretes



sub-library 1: adducts cleaved with sodium borohydride (d, scheme 3)
sub-library 2: adducts cleaved with allyl magnesium chloride (e, scheme 3)

6.3. Materials and methods

6.3.1. Materials

A. Equipment for chemical assessment, library synthesis and purification[44]

- ❑ Wheaton sample vials (4 ml, 8 ml, 20 ml, from Aldrich Chemicals)
- ❑ 40 ml glass vials for sample storage with screw caps and Teflon liners, used as reaction vessels for up to 4 g of resin (Aldrich Chemicals)
- ❑ extract clean tubes (1.5 ml, 4 ml and 8 ml from Alltech)
- ❑ polypropylene fritted syringes (15 ml, 25 ml, 70 ml from International Sorbent Technology, IST) used for resin filtration and washing operations
- ❑ vacuum manifold with luer fittings (VacMaster available from IST) for multiple sample resin washing and SPE extraction procedures
- ❑ phase-separation syringes (6 and 12 ml from Whatman) for water-alogenated solvents separation
- ❑ Elut Bond strong cation exchange SCX-SPE columns (Varian)
- ❑ adjustable pipettors (0.02-10 ml) for reagents dispensing and resin transferring (Gilson)
- ❑ multichannels pipettors (0.1 ml-1.2 ml) for solvents and reagent solutions dispensing (Matrix or Eppendorf)
- ❑ 96 deep-well polypropylene fritted microtiter plate for parallel solid-phase synthesis (GF/D unifilter plates available from Polyfiltronic/Whatman)
- ❑ 96 deep-well polypropylene microtiter plates for solution collection after cleavage (uniplate available from Polyfiltronic/Whatman)
- ❑ 96 deep-well polypropylene microtiter plate for parallel liquid-phase separation (MBPP unifilter plates from Polyfiltronic/Whatman)
- ❑ 96 deep-well polypropylene plates for solid-phase extraction (bioplate SCX from Polyfiltronic/Whatman)
- ❑ nitrogen gas supply
- ❑ heating blocks for reactions carried out at high temperature (VWR Scientific)
- ❑ orbital shaker (KS 250, IKA-BDH)
- ❑ rotative shaker (Heidolph Reax 2)
- ❑ HiTOPS kit for parallel solid-phase organic synthesis (Polyfiltronic Whatman)
- ❑ vacuum centrifuge for multiple sample concentration (Speed Vac SC210A, Savant)[45]
- ❑ oven for reactions at high temperature (vacuum oven 65, Tecnovetro)

---

[44] All the materials and reagents used for SP synthesis are available commercially and not particularly expensive; the analytical characterization also used sophisticated equipment but a simpler characterization with more conventional methods (TLC, IR, NMR in solution) could have been possible.

[45] The use of sophisticated equipment for the work-up, such as the centrifuges, could be substituted with normal rotavapours.

B. Equipment for analytical characterization

- HP1100 liquid chromatography system (Hewlett Packard)
- mass spectrometer Platform II (Micromass)
- Gilson XL233 autosampler (Gilson)
- diode array detector (Hewlett Packard)
- ^{1}H-MAS-NMR for solid-phase proton NMR analyses (400 MHz Nano-probe™, Varian)
- FT-IR Magna 760 (Nicolet)
- elemental analysis instrument EA 1108 (Carlo Erba)
- Supelcosil ABZ+ Plus (Supelco) column, 3.3 cm x 0.46 cm, 5 μm particle size
- HPLC buffers:
  ammonium acetate 10 mM, pH=6.4 buffer (A)
  acetonitrile (B)

C. Reagents

- The following reagents were purchased from Aldrich Chemicals and used without further purification: 1,2-diamino-4-methoxybenzene, sodium nitrite, boron tribromide, di-t-butyl carbonate, 5-carboxybenzotriazole, diisopropyl carbodiimide, dimethylaminopyridine, triethyl ortoformate, trifluoroacetic acid, sodium borohydride, sodium carbonate, ammonium chloride
- 0.87 M aldehyde (Aldrich, Sigma, Fluka, Acros or Janssen) stock solutions in dry THF/TMOF 2/1 or dioxane/TMOF 2/1 (Aldrich Chemicals)
- 0.5 M amine (Aldrich, Sigma, Fluka, Acros or Janssen) stock solutions in dry THF or dioxane (Aldrich Chemicals)
- aminomethylpolystyrene resin was purchased from Polymer Laboratories
- 0.75 M stock solution of 5-carboxy-N-tertbutyloxycarbonyl benzotriazole in dry DMF (3.55 g of benzotriazole in 18 ml of DMF) for the amide coupling
- 0.37 M stock solution of diisopropyl carbodiimide (DIC) in dry DMF for the amide coupling
- ninhidrine kaiser test kit solutions (available from Perseptive Biosystem) for the monitoring of the amide coupling
- 30% v/v TFA/DCM stock solutions for Boc deprotection
- 1 M stock mixtures of $NaBH_4$ in THF for the hydride ion cleavage methodology
- saturated aqueous Na_2CO_3 stock solutions for the extraction procedures
- 2 M allylmagnesium chloride solution in THF for the Grignard's cleavage (Aldrich Chemicals)
- 2 M phenyl magnesium bromide solution in THF for the Grignard's cleavage (Aldrich Chemicals)
- saturated aqueous NH_4Cl stock solutions for the extraction procedures
- 2 M stock solutions of ammonia in methanol for the final amines elution from the SCX column

6.3.2. *Methods*

A. Resin handling

Stir resins on orbital or rotative shaker for reactions carried out at rt; no stirring is required for the reaction carried out at high temperatures. Wash resin by adding a solvent while stirring (usually the reaction solvent followed by DMF, DCM and MeOH), then drain under vacuum. Repeat the operation for a number of cycles (typically 5 for each solvent). For the washing operations it is useful to have dispensers with adjustable volumes. The resin is then dried under vacuum usually at rt. The intermediates supported on the resins are usually stored in a dessicator under P_2O_5.

B. Synthesis of polymer-supported benzotriazole 3

1. Add a solution of di-t-butylcarbonate (18 g, 80 mmol) in dioxane (40 ml) dropwise over 30 min to a solution of commercially available 5-carboxybenzotriazole (10 g, 60 mmol) in dioxane/1M NaOH 4/6 (100 ml). After stirring 2 h at rt, dilute the solution with water (100 ml) and acidify to pH 3 with HCl 2N, then extract with ethyl acetate (2 x 250 ml). Wash the organic phase with brine and dry ($MgSO_4$), then concentrate to a small volume (60 ml). Adding cyclohexane (350 ml) while stirring vigorously causes precipitation of the product which is then filtered, washed with cyclohexane and dried to give pure 1 (12.2 g, 46.4 mmol, yield 80%) characterized by ^{1}H-NMR and MS.

2. Add diisopropyl carbodiimide (1.06 ml, 6.75 mmol) and catalytic DMAP (0.066 g, 0.54 mmol) to a mixture of AM-PS resin (3 g, 2.7 mmol) and benzotriazole 1 (3.55 g, 13.5 mmol) in dry DMF (18 ml) while stirring at rt. Shake the mixture gently for 15 h at rt, monitoring its completion via the colorimetric Kaiser test [32]. After filtration, wash the resin with DMF (15 ml), MeOH (15 ml) and DCM (15 ml) for 5 repeating cycles. Then vacuum dry the resin; characterize the supported benzotriazole 2 by IR (1645 cm^{-1}, C=O amide stretching; 1715 cm^{-1}, C=O carbamate stretching).

3. Treat the supported 2 (3 g, 2.7 mmol) with 30% TFA/DCM (30 ml) and stir at rt for 3 h. Monitor the reaction by IR (disappearance of the C=O carbamate stretching). Wash the resin after filtration with DCM (30 ml), MeOH (30 ml) and DCM (30 ml) for 5 repeating cycles. Then vacuum dry the resin; characterize by ^{1}H-MAS-NMR the supported benzotriazole 3 and by elemental analysis (N: 4.46%, found 4.37%, yield 98%).[46]

C. 2-(4-methylbenzylamino)pyrimidine 7A

1. Heat a mixture of supported 3 (0.1 g, 0.087 mmol), 4-methylbenzaldehyde (0.104 g, 0.87 mmol) and 2-aminopyrimidine (0.1 g, 1.044 mmol) in THF/TMOF 2/1 (1 ml) at 60°C over 15 h. After cooling at rt, filter the resin and wash it with THF (2 ml) and DCM (2 ml) for 5 repeating cycles; vacuum dry.

2. Heat a mixture of the dried adduct and sodium borohydride (0.033 g, 0.87 mmol) in dry THF at 60°C for 15 h. Add methanol (1 ml) and continue heating for a further 60 min (this breaks amine-borane complexes). Wash the exhausted resin with a mixture 1/1 MeOH/DCM (2 x 2 ml) and filter off. Then concentrate the organic solution to dryness, yielding a crude mixture (0.12 g) that contains the desired compound.

---

[46] Other solid-supported benzotriazoles were prepared but 3 resulted in the best construct for SP synthesis.

3. Take up the residue with DCM (2 ml) and extract with saturated aqueous Na_2CO_3, separate the two layers with a phase-separation polystyrene syringe and evaporate the recovered DCM layer to dryness. Take up the residue with methanol (2 ml) and purify it through an SCX column by elution with NH_3/MeOH. After concentration at reduced pressure, recover the final amine 7A (0.074 g, 43% yield) as a brownish solid, m.p. 112-114° C (given 122-123°C) [9]; ^{1}H-NMR ($CDCl_3$) δ 8.3 (d, 2H), 7.26 (m, 2H), 7.15 (m, 2H), 6.56 (m, 1H), 4.60 (s, 2H), 2.35 (s, 3H); MS *m/z* 200 $(M + H)^+$; LC-MS (DAD) 77% purity.

D. Dibenzyl-(1-naphthalen-2-yl-but-3-enyl) amine 7D

1. Adopt the same procedures described for adduct formation as seen for 7A, using 2-naphthaldehyde (0.135 g, 0.87 mmol) and N,N-dibenzylamine (0.2 g, 1.04 mmol).

2. Heat at 40° C for 15 h, a mixture of dried adduct (0.1 g, 0.087 mmol) and allyl magnesium chloride (2M THF sol., 1.3 ml, 2.61 mmol) in dry TMOF (0.7 ml). Then wash the exhausted resin with DCM (2 x 2 ml) and filter off.

3. Extract the DCM solution with saturated aqueous NH_4Cl, separate the two layers with a phase-separation polystyrene syringe and recover the DCM layer when it has completely evaporated. Take up the residue with methanol (2 ml) and purify it through an SCX column by elution with NH_3/MeOH. After concentration at reduced pressure, recover the final amine 7D (0.016 g, 48% yield) as a solid, 66-68°C; ^{1}H-NMR ($CDCl_3$) δ 7.85 (m, 2H), 7.64 (bs, 2H), 7.49 - 7.4 (m, 3H), 7.41 (d, 4H), 7.33 (t, 4H), 7.24 (m, 2H), 5.83 (m, 1H), 5.07 (m, 2H), 3.99 (m, 1H), 3.84 (d, 2H), 3.30 (d, 2H), 2.96 (m, 1H), 2.73 (m, 1H); MS *m/z* 378 $(M + H)^+$; LC-MS (DAD) 100% purity.

E. Synthesis of sub-library 1

1. Adopt the same procedures described for the preparation of compound 7A, for the parallel preparation of the discrete array (25 mg of AM-PS resin for each library individual).

2. Take up the 16 residues with DCM (0.7 ml each) and extract with saturated aqueous Na_2CO_3 (0.7 ml each); separate the two layers in parallel using a MBPP (meltblown polypropylene) microtiter plate bearing an hydrophobic polypropylene membrane with 10 μm pore size. Recover the DCM layers into a 96 deep well-plate and evaporate to dryness in a vacuum centrifuge. Take up the residues with methanol (1 ml) and purify in parallel using a Bioplate™-SPE-SCX 96 well-plate. Collect the samples into a 96 deep well-plate by elution with NH_3/MeOH. After concentration at reduced pressure in a vacuum centrifuge, control the purity of the recovered amines by LC-MS (on the entire set) and by ^{1}H-NMR on a few samples.

F. Synthesis of sub-library 2

1. Adopt the same procedures described for the preparation of compound 7D, for the parallel preparation of the discrete array (25 mg of AM-PS resin for each library individual).

2. Extract the DCM solutions (0.7 ml each) with saturated aqueous NH_4Cl (0.7 ml) and separate the two layers in parallel using a MBPP microtiter plate as described previously. Recover the DCM layers into a 96 deep well-plate and evaporate to dryness in a vacuum centrifuge. Take up the residues with methanol (1 ml) and purify in parallel using a Bioplate™-SPE-SCX 96 deep well-plate as described previously. Collect

the samples into a 96 deep well-plate by elution with NH_3/MeOH and concentrate at reduced pressure in a vacuum centrifuge. Control the purity of recovered amines by LC-MS (on the entire set) and by ^{1}H-NMR on a few samples.

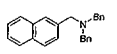
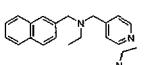
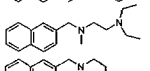
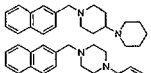
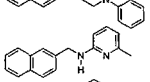
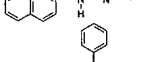
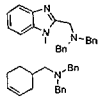
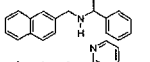
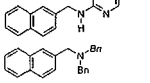
G. Analytical characterization

1. The analytical characterization of supported intermediates is normally carried out using spectroscopic techniques:
   (a) FT-IR where spectra are acquired in nujol or KBr pellets as sample vehicles;
   (b) ^{1}H-MAS-NMR carried out in deutero-CH_2Cl_2;
   (c) elemental analysis determinations.
2. The compounds after cleavage are characterized by:
   (a) NMR, usually in $CDCl_3$ as solvent unless otherwise stated;
   (b) HPLC/MS. Data are obtained using the HP1100 liquid chromatography system equipped with diode array detector (Hewlett Packard, Germany) coupled with the mass spectrometer Platform II (Micromass, United Kingdom). The mobile phase is water (A) and acetonitrile (B) from 20% to 90% of B in 8 min, and then 5 min with 90% of B. The re-equilibration time between two injections is 3 min. The flow rate is 0.8 ml/min. All samples are injected using a Gilson XL233 autosampler (Gilson, France). The injection volume is 20 μl. Diode array chromatograms are collected using a large bandwidth (220 nm-350 nm). The mass spectrometer work in positive electrospray ionization mode (ES+); the scan range is between 170 amu and 800 amu (centroid mode). All data acquired are processed using the MassLynx software with the OpenLynx Diversity tools (Micromass).

Quality control of the library

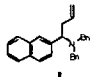
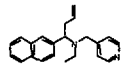
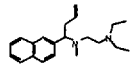
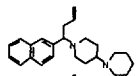
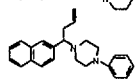
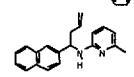
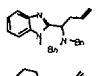
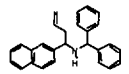
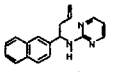
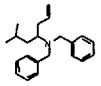
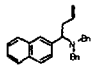
Tables 6.1 and 6.2 present the analytical results of sub-libraries 1 and 2 respectively.

Table 6.1. Sub-library 1

Product	*Yield*[a]	*LC-MS purity*[b]	*Product*	*Yield*[a]	*LC-MS purity*[b]
	43	90		60	82
	26	61		-	-
	22	82		-	-
	41	79		65	91
	-	-		25	27
	24	60		-	-
	47	90		55	86
	30	75		50	86

a. calculated as w/w yield after SCX purification
b. calculated as area/area percentage

Table 6.2. Sub-library 2

Product	Yield[a]	LC-MS purity[b]	Product	Yield[a]	LC-MS purity[b]
	33	100		-	-
	48	100		-	-
	45	100		-	-
	41	55		26	94
	-	-		22	13
	36	81		-	-
	20	100		-	-
	11	75		40	100

a. calculated as w/w yield after SCX purification
b. calculated as area/area percentage

A good reactivity can be observed in sub-library 1 (see table 6.1 above) with aromatic, hetero-aromatic and aliphatic aldehydes giving the desired products with comparable overall yields (20-50%) and purities (60-90% a/a by HPLC/MS, table 6.1). The only exception is 5-bromo-2-furancarboxaldehyde which does not produce the expected amine due to aldehyde polymerization during the adduct formation. The reactivity profile of amines is worse: some aliphatic amines fail to react while four others, including low-reacting 2-amino pyridine, give the expected products varying from acceptable to good overall yields (25-65%) and purities (80-90% a/a by HPLC/MS, table 6.1).

Once again, a good reactivity for the various aldehydes in sub-library 2 (see table 6.2 above) can be observed obtaining the desired products with moderate overall yields (20-50%) and high purities (55-100% a/a by HPLC/MS, table 6.2). Amines show the same behaviour, as seen for the previous array, with additional interference caused by primary amines which, producing a free NH after the adduct formation, does not react properly with the Grignard's reagent.[47] Only two compounds are obtained with moderate yields (25-40%) and high purities (95-100% a/a by HPLC/MS, table 6.2).

[47] The SP chemistry proved to be extremely robust throughout the whole synthetesis scheme and we did not take any particular care for anhydrous conditions or any other sophisticated reaction condition apart from the Grignard cleavage.

References

1. Thompson, M. A. and Ellman, J. A. 1996. *Chem. Rev.*, vol. 96, pp. 555-600.
2. Fruchtel, J. S. and Jung, G. 1996. *Angew. Chem., Int. Ed. Engl.*, vol. 35, pp. 17-42.
3. Balkenhohl, F., von dem Bussche-Hunnefeld, C., Lansky, A. and Zechel, C. 1996. *Angew. Chem., Int. Ed. Engl.*, vol. 35, pp. 2288-2337.
4. Hermkens, P. H. H., Ottenheijm, H. C. J. and Rees, D. C. 1997. *Tetrahedron*, vol. 53, pp. 5643-5678.
5. Brown, A. R., Hermkens, P. H. H., Ottenheijm, H. C. J. and Rees, D. C. 1998. *Synlett*, pp. 817-827.
6. Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A. and Gordon, E. M. 1994. *J. Med. Chem.*, vol. 37, pp. 1233-1251.
7. Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A. and Gallop, M. A. 1994 *J. Med. Chem.*, vol. 37, pp. 1384-1401.
8. Brown, D. 1997. *Mol. Diversity*, vol. 2, pp. 217-222.
9. Lam, K. S., Lebl, M. and Krchnak, V. 1997. *Chem. Rev.*, vol. 97, pp. 411-448.
10. Kauvar, L. M. and Laborde, E. 1998. *Curr. Opin. Drug. Disc. Develop.*, vol. 1, pp. 66-70.
11. Han, Y., Walker, S. D. and Young, R. N. 1996. *Tetrahedron Lett.*, vol. 37, pp. 2703-2706.
12. Beaver, K. A., Seigmund, A. C. and Spear, K. L. 1996. *Tetrahedron Lett.*, vol. 37, pp. 1145-1148.
13. Boehm, T. L. and Showalter, H. D. 1996. *J. Org. Chem.*, vol. 61, pp. 6498-6499.
14. Ngu, K. and Patel, D. V. 1997. *Tetrahedron Lett.*, vol. 38, pp. 973-976.
15. Routledge, A., Abell, C. and Balasubramanian, S. 1997. *Tetrahedron Lett.*, vol. 38, pp. 1227-1230.
16. Dressman, B. A., Singh, U. and Kaldor, S. W. 1998. *Tetrahedron Lett.*, vol. 39, pp. 3631-3634.
17. Bohm, G., Dowden, J., Rice, D. C., Burgess, I., Pilard, J., Guilbert, B., Haxton, A., Hunter, R. C., Turner, N. J. and Flitsch, S. L. 1998. *Tetrahedron Lett.*, vol. 39, pp. 3819-3822.
18. Plunkett, M. J. and Ellman, J. A. 1997. *J. Org. Chem.*, vol. 62, pp. 2885-2893.
19. Woolard, F. X., Paetsch, J. and Ellman, J. A. 1997. *J. Org. Chem.*, vol. 62, pp. 6102-6103.
20. Gayo, L. M. and Suto, M. J. 1997. *Tetrahedron Lett.*, vol. 38, pp. 211-214.
21. Kroll, F. E. K., Morphy, R., Rees, D. and Gani, D. 1997. *Tetrahedron Lett.*, vol. 38, pp. 8573-8576.

22. Kaldor, S. W. and Siegel, M. G. 1997. *Curr. Opin. Chem. Biol.*, vol. 1, pp. 101-106.

23. Ley, S. V., Bolli, M. H., Hinzen, B., Gervois, A.-G. and Hall, B. J. 1998. *J. Chem. Soc. Perkin Trans. I*, pp. 2239-2242.

24. Booth, R. J. and Hodges, J. C. 1998. *Acc. Chem. Res.*

25. Flynn, D. L., Devraj, R. V. and Parlow, J. J. 1998. *Curr. Opin. Drug. Disc. Develop.*, vol. 1, pp. 41-50.

26. Curran, D. P. 1998. *Angew. Chem., Int. Ed. Engl.*, vol. 37, pp. 1174-1196.

27. Shuker, A. J., Siegel, M. G., Matthews, D. P. and Weigel, L. O. 1997. *Tetrahedron Lett.*, vol. 38, pp. 6149-6152.

28. Siegel, M. G., Hahn, P. J., Dressman, B. A., Fritz, J. E., Grunwell, J. R. and Kaldor, S. W. 1997. *Tetrahedron Lett.*, vol. 38, pp. 3357-3360.

29. Lawrence, M. R., Biller, S. A., Fryszman, O. A. and Poss, M. A. 1997. *Synthesis*, pp. 553-558.

30. Katritzky, A. R., Rachwal, S. and Hitchings G. J. 1991. *Tetrahedron*, vol. 47, pp. 2683-2732.

31. Katritzky, A. R., Lan, X., Yang, J. Z. and Denisko, O. V. 1998. *Chem. Rev.*, vol. 98, pp. 409-548.

32. Katritzky, A. R., Rachwal, S. and Rachwal, B. 1987. *J. Chem. Soc. Perkin Trans. I*, pp. 799-804.

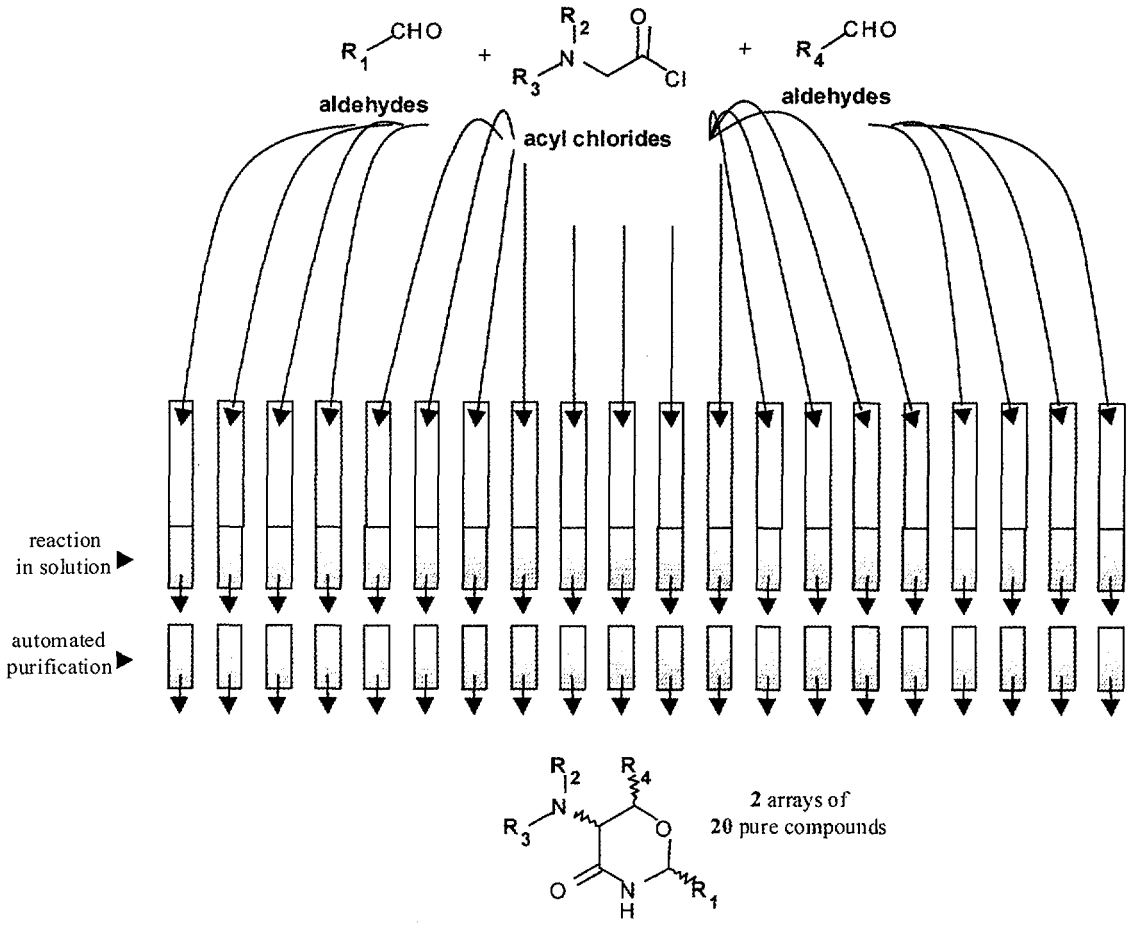
7

Solution-phase synthesis of a discrete library of heterocycles

P. Seneci and A. Missio

7 Solution-phase synthesis of a discrete library of heterocycles

Pierfausto Seneci and Andrea Missio



- Fast parallel synthesis of arrays of heterocycles using classical organic chemistry
- Access to remarkable chemical diversity
- Use of solid supported scavengers to purify the solution products
- Reliable and parallel purification procedures
- Cheap techniques and reagents used

7.1. Introduction

The focus of the present chapter is the presentation and analysis of a typical method used for generating a combinatorial library of small organic molecules (SOM) in solution. The use of combinatorial libraries for identifying novel chemical leads or for optimizing promising lead candidates has emerged as a powerful method in accelerating the drug discovery process [1-3]. Until recently, such libraries were usually prepared by solid-phase techniques, but solution-phase combinatorial chemistry [4-8] is becoming more commonly used, and a sensible option for short reaction sequences and for small to medium-sized libraries. The increased requirements for reaction purification procedures and for their handling is balanced by the easier adaptation of classical organic chemistry reaction conditions to combinatorial protocols in solution where all the arsenal of organic reactions can be used. The chapter illustrates an example of complex reaction procedures adapted to the production of a small array of biologically relevant heterocycles.

The use of automated purification procedures [9-12] for parallel synthesis in solution has enabled the combinatorialization of many synthesis procedures requiring extractions or chromatographic separations. The use of solid supported reagents as scavengers [13-15], or sequestration enabling reagents (SER), to purify solution-phase libraries has also emerged as a new, powerful technique and is also discussed in the chapter.

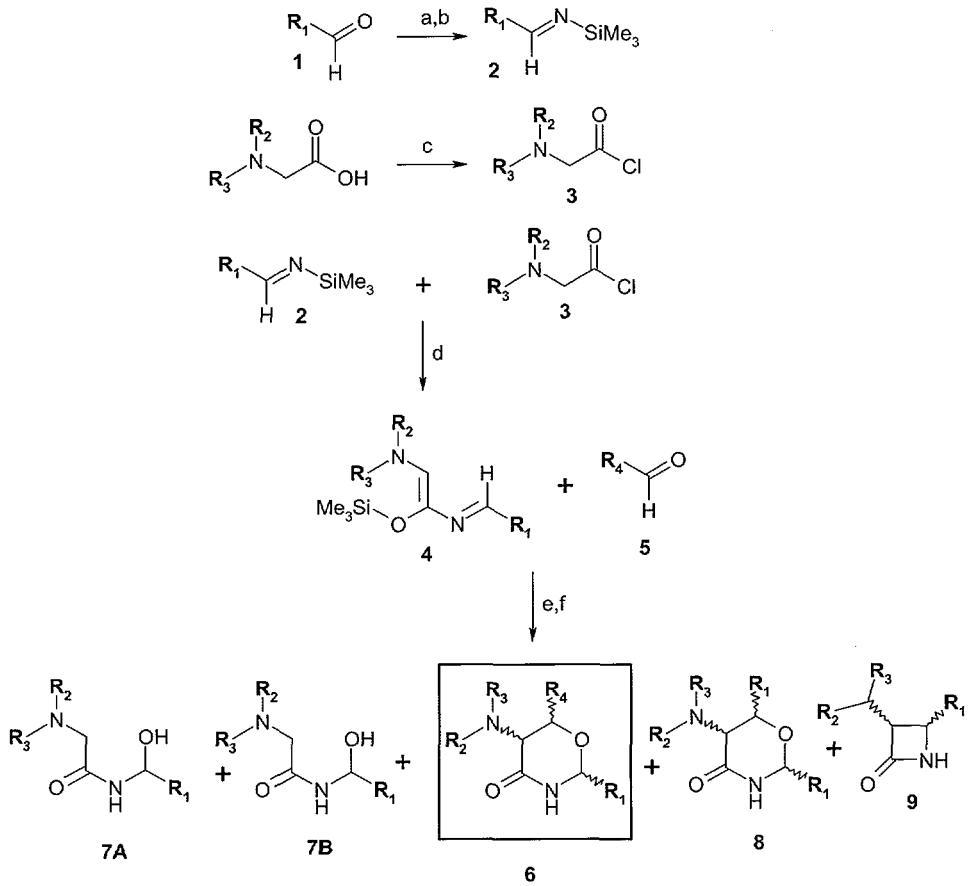
7.2. Synthetic scheme and library design

7.2.1. The reaction scheme

The organic reaction scheme used to prepare the key intermediates 2, 3 and 4 and to design the library synthesis is reported in scheme 4 below.[48]

The synthesis of the silylimines 2, of the acyl chlorides 3 and of the azadienes 4 were performed as already reported [16-18]. The azadienes 4 were reacted with the second aldehyde monomer set to give as a major product the expected oxazines 6 and varying amounts of side-products such as open amides 7A, B, bis adducts 8 and β-lactams 9. The experimental procedure to prepare and purify (via SER reagents) the 40-membered discrete library is reported in the following section.

Scheme 4



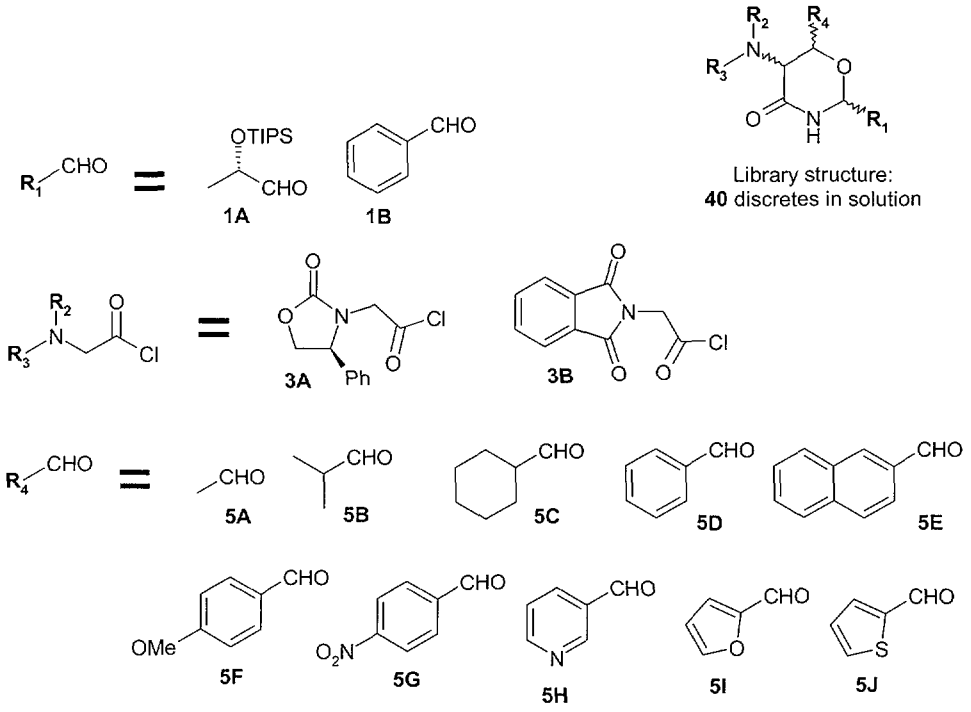
a. lithium hexamethyldisilazane, heptane, 0° C, 15', then rt, 1 h
b. TMS-Cl, rt, 1h
c. oxalyl chloride, dry toluene, 60° C, 3 h
d. TEA, 0° C, 30', then rt, 1 h
e dry DCM, - 78° C, then $BF_3.Et_2O$, -78° C, 3 h, then warming to rt
f. aminomethylated polystyrene resin

[48] The library could easily be expanded by using more monomers. Such a procedure is currently under evaluation in our laboratories.

7.2.2. *Library design*

The generic structure of the library and the monomers used are illustrated in Figure 12. The syntheses of aldehyde 1A and acyl chloride 3A are reported in the section that follows. The use of 2 aldehydes as first monomer set (2 R_1), 2 acyl chlorides as second monomer set (2 R_2,R_3) and 10 aldehydes as third monomer set (10 R_4) produced a 40-member discrete library.[49] The quality of results of the library obtained by HPLC\MS are described in the section on "Library results".

Figure 7.1. Library structure and monomers used



7.3. Materials and methods

7.3.1. *Materials*

A. Equipment for library synthesis and purification

- ❑ screw-top V-vials equipped with Teflon Mininert syringe valves and magnetic spin vanes (all available from Aldrich Chemicals) to be used as reaction vessels
- ❑ aluminium test tube racks built in-house to hold reaction vessels
- ❑ nitrogen gas supply

[49] The synthetic scheme involved many reactions where complex experimental conditions and handling of extremely sensitive intermediates were necessary. As the library synthesis was performed in solution, classical glassware (found in any synthesis laboratory or which can be bought at a reasonable cost) was used in a parallel fashion.

- ❑ 1-ml polypropylene disposable syringes equipped with rigid polyethylene plungers and stainless steel disposable needles (all available from Aldrich Chemicals) to add reagents
- ❑ 15-ml polypropylene fritted syringes fitted with a stopcock and a stopper to be used in the aqueous work-up and in the resin handling
- ❑ vacuum manifold with luer fittings (available from IST)
- ❑ 20-ml glass scintillation vials (Wheaton) to be used in the vacuum centrifuge
- ❑ vacuum centrifuge SpeedVac Plus (Savant)[50]
- ❑ magnetic stirring plates MR 3003 (Heidolph)
- ❑ orbital shaker K5250 (IKA)

B. Equipment for analytical characterization

- ❑ HP1100 liquid chromatography system (Hewlett Packard)
- ❑ mass spectrometer Platform II (Micromass)
- ❑ Gilson XL233 Autosampler (Gilson)
- ❑ diode array detector (Hewlett Packard)
- ❑ evaporative light scattering detector Sedex 55 (S.E.D.E.R.E.)
- ❑ supelcosil ABZ+ Plus (Supelco) column, 3.3 cm x 0.46 cm, 5 μm particle size
- ❑ HPLC buffers:

ammonium acetate 10 mM, pH=6.4 buffer (A)
acetonitrile (B)

C. Reagents

- ❑ 0.2 M R_1 aldehyde (Aldrich Chemicals) stock solutions in n-hexanes (Aldrich Chemicals)
- ❑ 1 M lithium hexamethyldisilazane (LHDMS) solution in dry THF (Aldrich Chemicals)
- ❑ trimethylsilyl chloride (Aldrich Chemicals)
- ❑ triethylamine (Aldrich Chemicals)
- ❑ Evan's oxazolidinone-derived carboxylic acid (Aldrich Chemicals)
- ❑ 0.66 M acyl chloride (Aldrich Chemicals) stock solutions in dry toluene (Aldrich Chemicals)
- ❑ 1 M R_4 aldehyde (Aldrich Chemicals) stock solutions in dry DCM (Aldrich Chemicals)
- ❑ 0.2 M $BF_3.Et_2O$ solution in dry DCM (Aldrich Chemicals)
- ❑ saturated $NaHCO_3$ aqueous solution
- ❑ saturated NaCl aqueous solution

[50] The use of sophisticated equipment for the work-up, such as the centrifuges, could be substituted with normal rotavapours.

- ❑ aminomethylated polystyrene resin (NovaBioChem, 200-400 mesh; 1.2 mmol/g) for excess aldehyde trapping
- ❑ dry THF (Aldrich Chemicals)
- ❑ trimethylorthoformate (TMOF, Aldrich Chemicals).

7.3.2. Methods

A. General considerations

Perform all reactions inert gas (nitrogen). Add nitrogen and reagents through a mini-nert valve. Perform reactions at 0° C or at −78° C by placing a test tube rack containing the vials in an ice bath or in a dry ice/acetone bath, respectively, with the bath mounted on a magnetic stirrer. Remove solvents using a vacuum centrifuge.

B. Synthesis of key intermediates

1. O-TIPS protected L-lactic aldehyde (1A of figure 7.1):

(a) Add triisopropylsilyl chloride (30 mmol) to a solution of ethyl-(S)-(-)-lactate (30 mmol) and imidazole (60 mmol) in dry DMF cooled with an ice/water bath. Remove from bath after stirring 10 min, then stir the reaction at rt for 3 h. Pour the mixture into ice/water and extract with hexane (3 x 50 ml), wash with brine (2 x 50 ml), dry (Na_2SO_4) and concentrate *in vacuo*;

(b) Add DIBAL (1 M in hexane, 30 ml, 30 mmol) dropwise to the TIPS-protected ester (20 mmol) in dry Et_2O (80 ml) at −78° C. After stirring for 2 h at −78° C, pour the mixture into ice/water and extract with ethyl ether. Wash the solvent twice with diluted cold HCl, brine, then dry and concentrate *in vacuo*. The residue is chromatographed on silica gel to give pure 1A (15.7 mmol, yield 78.5%).

2. O-TIPS protected L-lactic aldehyde trimethylsilylimine (2A of figure 7.1:

(a) Add a solution of 1A (2.3 g, 10 mmol) in heptane (50 ml) dropwise to an ice-cooled solution of lithium hexamethyldisilyl amide (LHMDSA, 1 M in THF, 10 ml). After the complete addition of 1A, stir the reaction mixture at 0° C for 15 min and at rt for 1 h. Add trimethylsilyl chloride (1.3 ml, 10 mmol) to the solution in one portion and stir this mixture for 1 h at rt;

(b) Confirm the formation of 2A by an infrared spectrum of the reaction mixture (imine, 1653 cm^{-1}).

3. Evan's oxazolidinone-derived acyl chloride (3A at figure 7.1):

(a) Add one portion of oxalyl chloride (1.6 m, 18 mmol) to a solution of the corresponding, commercially available Evan's oxazolidinone-derived acid (2.6 g, 12 mmol) in dry toluene (50 ml), then warm the mixture to 60° C for 3 h;

(b) Remove the solvent and excess oxalyl chloride *in vacuo* and re-dissolve the resulting crude acid chloride 3A in toluene (50 ml).

4. O-TMS protected Evan's lactic azadiene (4A of figure 7.1):

(a) Cool the mixture containing 2A to 0° C and add TEA (3 ml, 20 mmol);

(b) After stirring the mixture for 5 min at 0° C, add the solution of 3A slowly (10 min). Continue stirring for 30 min at 0° C for 1 h at rt. Then filter this yellow mixture through celite and remove the solvent *in vacuo*;

(c) Remove a small sample from the reaction mixture and run a ^{1}H-NMR of this sample to check for the presence of 4A.

C. Library synthesis

1. Dissolve the four crude azadienes 4A-D (theoretically 2.4 mmol) into dry DCM (12 ml for each azadiene). Distribute an aliquot of each solution (1 ml, theoretically 0.2 mmol) in each of the 40 vials (5 ml capacity) equipped with a rubber septum and a stirring bar and kept in an inert atmosphere (nitrogen).

2. Cool the vials at −78° C (dry ice/acetone bath) and withdraw aliquots from stock solutions of the third monomer set (0.22 mmol of aldehydes 5A-J in 0.2 ml of dry DCM) and add dropwise to the 40 vials. Each of the four aliquots for the same aldehyde should be added to each of the four vials containing four azadienes 4A-D. An inert atmosphere (nitrogen) should be maintained.

3. Immediately after the aldehyde addition, treat each vial with aliquots from a stock solution of 0.2 M $BF_3.Et_2O$ in dry DCM (1 ml). Stir the mixtures at −78° C for 3 h, then slowly warm to rt and stir overnight.

D. Parallel purification and use of a sequestration enabling reagent (SER)

1. Dilute the crudes from the library synthesis to 4 ml with DCM, then pour into four 10-manifold supported polypropylene syringes (15 ml) equipped each with a bottom frit and a stopcock.

2. Extract each solution with saturated aqueous sodium bicarbonate solution (2 x 4 ml) and wash with brine (6 ml). Remove the 40 organic phases from the syringes and dry ($MgSO_4$) for 5 min while stirring vigorously. After filtration, remove solvents and other volatile components using a vacuum centrifuge.

3. Dissolve the crudes in dry THF\TMOF 1\1 (1 ml) and treat with aminomethylated polystyrene resin (200-400 mesh; 1.2 mmol/g, 167 mg for each solution, 0.2 mmol). After shaking overnight, wash thoroughly the resins loaded with the imines (resulting from reaction with the excess aldehydes) with THF and DCM (2 x 1 ml for each solvent and vial). Collect the organic solutions in 20 ml glass scintillation vials and evaporate to dryness using a vacuum centrifuge.

E. Library compounds: analytical characterization

1. Obtain HPLC/MS data using the HP1100 liquid chromatography system equipped with diode array detector (Hewlett Packard) and coupled with the mass spectrometer Platform II (Micromass) and the evaporative light scattering detector Sedex 55 (S.E.D.E.R.E.).

2. Obtain the chromatographic separations using a Supelcosil column ABZ+ Plus (Supelco) 3.3 x 0.46 cm, 5 μm particle size. The mobile phase should be ammonium acetate 10 mM, pH=6.4 buffer (A) and acetonitrile (B) from 60% to 90% of B in 5 min, and then 10 min with 90% of B. The re-equilibration time between two injections should be 3 min and the flow rate 0.6 ml/min.

3. Inject all samples using a Gilson XL233 Autosampler (Gilson). The injection volume is 10 μL. Collect the diode array chromatograms using a large bandwidth (from 220 nm to 350 nm) and work the ELS detector using nitrogen as nebulizer gas (P = 2 bar) at an inner temperature of 50° C.

4. Operate the mass spectrometer in positive electrospray ionization mode (ES+); the scan range should be between 200 amu and 700 amu (centroid mode). Process all data acquired using the MassLynx software with the OpenLynx Diversity tools (Micromass).

F. Library results

Table 7.1 sets out the library results in terms of yields and purity of oxazin-4-ones 6 and quantities of side-products 7A, B, 8 and 9.

Table 7.1. Quality control of the library

Compound[a]	*Oxazin-4-one 6*[b]	*Open amides 7A, B*[b]	*Bis adducts 8*[b]	*β-lactams 9*[b]	*Purity class*
A-A-A	80	traces	ND[c]	traces	G[d]
A-A-B	65	5	ND		G
A-A-C	55	15	ND	5	S[e]
A-A-D	40	30	ND	10	U[f]
A-A-E	65	traces	ND	30	G
A-A-F	50	30	ND	10	S
A-A-G	55	25	ND	10	S
A-A-H	30	15	ND	5	U
A-A-I	70	15	ND	5	G
A-A-J	60	25	ND	10	G
A-B-A	75	traces	ND	traces	G
A-B-B	50	15	ND	10	S
A-B-C	60	5	ND	traces	G
A-B-D	55	10	ND	traces	S
A-B-E	40	25	ND	10	S
A-B-F	35	35	ND	15	U
A-B-G	55	15	ND	10	S
A-B-H	20	20	ND	10	U
A-B-I	65	5	ND	traces	G
A-B-J	50	10	ND	5	S
B-A-A	35	10	ND	15	U
B-A-B	60	5	15	10	G
B-A-C	35	20	10	5	U
B-A-D	70	10	ND	15	G
B-A-E	45	25	10	20	S
B-A-F	50	5	20	20	S
B-A-G	60	20	traces	15	G
B-A-H	80	traces	15	traces	G
B-A-I	45	35	15	traces	S
B-A-J	30	30	15	10	U
B-B-A	70	10	traces	10	G
B-B-B	55	15	5	10	S
B-B-C	45	15	5	10	S
B-B-D	60	10	ND	5	G
B-B-E	25	20	10	25	U
B-B-F	10	25	20	25	U
B-B-G	15	30	10	30	U
B-B-H	35	25	5	10	U
B-B-I	50	20	traces	25	S
B-B-J	65	15	ND	5	G

a. compounds identified as monomer R_1 (A,B) - monomer $R_{2,3}$ (A,B) - monomer R_4 (A-J)
b. yield
c. ND, not detected
d. G, good purity, 60%-80%
e. S, satisfactory, 40%-60%
f. U, unsatisfactory, 0-40%

References

1. Gallop, M. A., Barrett, R. W., Dower, W. J., Fodor, S. P. A. and Gordon, E. M. 1994. *J. Med. Chem.*, vol. 37, pp. 1233-1251.
2. Gordon, E. M., Barrett, R. W., Dower, W. J., Fodor, S. P. A. and Gallop, M. A. 1994. *J. Med. Chem.*, vol. 37, p. 1385.
3. Terrett, N. K., Gardner, M., Gordon, D. W., Kobylecki, R. J. and Steele, J. 1995. *Tetrahedron*, vol. 51, p. 8135.
4. Storer, R. 1996. *Drug Discovery Today*, vol. 1, p. 248.
5. Bailey, N., Cooper, A. W. J., Deal, M. J., Dean, A. W., Gore, A. L., Hawes, M. C., Judd, D. B., Merritt, A. T., Storer, R., Travers, S. and Watson, S. P. 1997. *Chimia*, vol. 51, p. 832.
6. Boger, D. L. and Chai, W. 1998. *Tetrahedron*, vol. 54, p. 3955.
7. An, H., Wang, T., Mohan, V., Griffey, R. H. and Cook, D. P. 1998. *Tetrahedron*, vol. 54, p. 4067.
8. Boger, D. L., Chai, W. and Jin, Q. 1998. *J. Am. Chem. Soc.*, vol. 120, p. 7220.
9. Zeng, L. and Kassel, D. B. 1998. *Anal. Chem.*, vol. 70, p. 4380.
10. Zeng, L., Wang, X., Wang, T. and Kassel, D. B. 1998. *Combi. Chem. High Throughput Screening*, vol. 1, p. 101.
11. Griffey, R. H., An, H., Cummins, L. L., Gaus, H. J., Haly, B., Herrmann, R. and Cook, D. P. 1998. *Tetrahedron*, vol. 54, p. 4067.
12. Zeng, L., Burton, L., Yung, K., Shushan, B. and Kassel, D. B. 1998. *J. Chromatogr., A*, vol. 794, p. 3.
13. Suto, M. J., Gayo, L. M., Palanki, M. S. S. and Sullivan, R. 1998. *Tetrahedron*, vol. 54, p. 4141.
14. Ley, S. V., Bolli, M. H., Hinzen, B., Gervois, A.-G. and Hall, B. J. 1998. *J. Chem Soc. Perkin Trans. I*, p. 2239.
15. Flynn, D. L., Devraj, R. V. and Parlow, J. J. 1998. *Curr. Opin. Drug Discovery Dev.*, 1, p. 41.
16. Bongini, A., Panunzio, M., Bandini, E., Martelli, G. and Spunta, G. 1997. *J. Org. Chem.*, vol. 62, p. 8911.
17. Cainelli, G., Panunzio, M., Bandini, E., Martelli, G. and Spunta, G. 1996. *Tetrahedron*, vol. 52, p. 1685.
18. Evans, D. A., Sjogren, E. B., Weber, A. E. and Conn, R. E. 1987. *Tetrahedron Lett.*, vol. 28, p. 39.

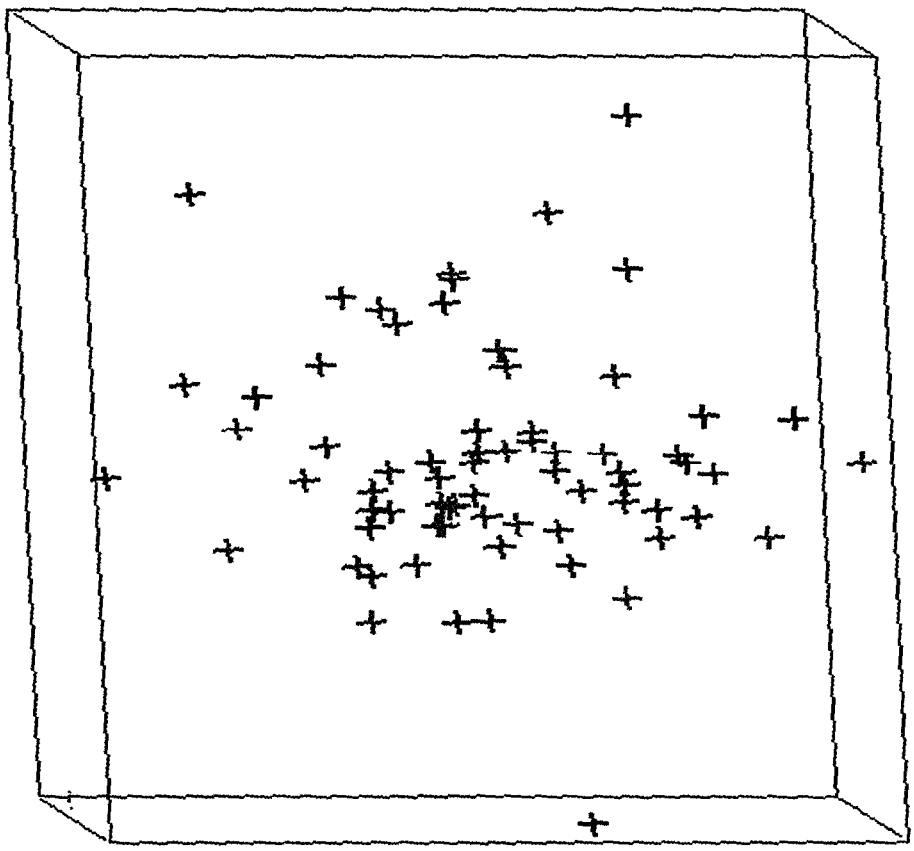
8

Designing combinatorial libraries

J. V. Gillet and D. V. S. Green

8 Designing combinatorial libraries

Valerie J. Gillet and Darren V. S. Green



- ☐ Identify potential reactants
- ☐ Filter out undesirable compounds
- ☐ Choose library design strategy
- ☐ Calculate molecular descriptors
- ☐ Apply compound selection technique

8.1. Introduction

Combinatorial chemistry techniques are now used routinely in the search for novel active compounds. They allow a much more rapid throughput of structures in terms of synthesis and testing than traditional synthesis methods. A typical combinatorial experiment may involve three or four substitution positions with hundreds or even thousands of possible reactants at each position so, theoretically, massive numbers of structures are possible. However, the number of potential compounds can easily exceed the capacity of synthesis and screening apparatus, thus there is a need to be selective about the compounds that are actually made. Many computational tools have been developed to assist in the process of selecting which compounds to include within a combinatorial experiment. In the present chapter, a brief overview is given of the many alternative techniques that are available. The discussion that follows focuses on applying some of the techniques involved in the design of a two-component combinatorial library experiment.

The number of compounds that can be potentially synthesized in a combinatorial experiment by using all possible reactants at each substitution position defines what is referred to as a virtual library. A virtual library can be constructed by searching for all readily available reactants, for example, in in-house databases as well as in databases of commercially available compounds. In library design, the aim is to select a library of a size that is within the capabilities of the synthesis or screening systems available. In most cases, this will be a subset of the virtual library associated with a given set of chemistries. The end result of library design is a real combinatorial library that can be synthesized

There are different strategies that can be employed to reduce the size of a virtual library and the strategy for a particular library will depend on the purpose for which the library is being designed, for example, the methods used for lead generation tend to be different from those used for lead optimization.

Some techniques are common to all drug discovery programmes and can be used as a first step in the design process. The aim is to reduce the size of the virtual library by including only those compounds viable as drug compounds. While it is difficult to clearly define the concept of "drug-likeness" [1, 2, 3], there are some general criteria that can be applied, for example, flexibility, molecular weight and the presence of reactive groups that may interfere with the intended reaction or lead to toxic or unstable products. Such techniques can be thought of as filters whereby undesirable molecules are eliminated before design factors are brought into play. Normally, such filters are applied to the reactant pools prior to enumerating a virtual library.

Once the undesirable molecules have been eliminated, the next step in library design is to select reactants of interest from the compounds that remain in the virtual library. When searching for new leads, usually the aim is to select a structurally diverse range of compounds based on the assumption that if similar compounds are likely to have similar biological activity then dissimilar or diverse compounds are likely to have different biological activity. Diverse libraries can be designed to be screened either against a range of different biological targets or against a specific target. Conversely, when a library is being designed to follow up a lead compound in a lead optimization experiment, the criteria for selecting compounds are normally based on similarity to the lead in order to design focused or targeted libraries.

The diversity of a library of compounds reflects the degree of heterogeneity, structural range or dissimilarity within the set of compounds. A number of different diversity metrics have been suggested and all are based, either directly or indirectly, on the concept of intermolecular similarity or distance. Determining the similarity or dissimilarity between two molecules requires, firstly, that the molecules be represented by appropriate numerical descriptors and, secondly, that a quantitative method be used to determine the degree of resemblance between the two sets of descriptors. Many differ-

ent descriptors have been used in diversity studies [4, 5, 6] including: 1D descriptors such as topological indices and physicochemical properties [7]; 2D descriptors such as bitstrings [8] and fingerprints [9, 10]; and 3D descriptors such as pharmacophore keys [11]. The choice of descriptor depends on factors such as the size of the virtual and real combinatorial libraries, the commercial software that is available and the method that will be used to select compounds.

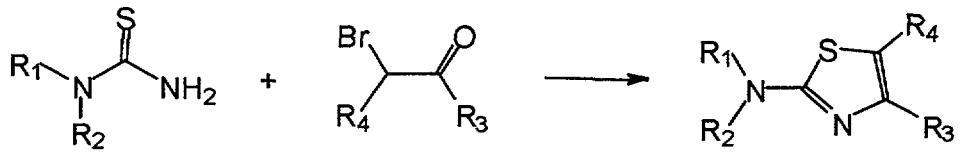
A number of different algorithms have been developed to select compounds. The main methods are dissimilarity-based compound selection [12], clustering [13], partitioning [14] and optimization techniques [15]. Clustering and partitioning methods can also be used to select compounds based on similarity rather than diversity. In some cases, the choice of selection algorithm limits the descriptors that can be used, for example, partitioning methods require that the compounds be represented by low dimensionality descriptors, whereas dissimilarity-based methods can be used with high dimensionality descriptors such as fingerprints. Some methods are more costly in terms of computational time: clustering is an example of a computationally expensive process. The expense may restrict the size of libraries that can be handled.

Finally, selection can be performed at the product or the reactant level. If selection is performed at the reactant level, there is no need for library enumeration. However, there is evidence to suggest that library design is more effective when the analysis is done in product space [16]. In such a case, library enumeration is required so that the selection algorithm can be applied to the product molecules themselves. When selecting molecules in product space, consideration must be given to the combinatorial constraint, that is, for a two-component reaction, every reactant from one pool must be combined combinatorially with all the reactants in the second pool.

The theoretical aspects concerning the use of computational approaches for the design of combinatorial libraries are reviewed by Gillet, Willott et al. [17]. The present chapter is concerned with more practical aspects. There is no single scheme for designing combinatorial libraries since the appropriate procedure depends on a number of different factors such as the characteristics of the library itself, the information available (for example, the crystal structure of an enzyme target) and the software that is at hand. (See Warr [18] for a recent review of commercially available software for diversity analysis.) However, it is possible to identify a general scheme where alternative tools can be used to perform each of the steps. The initial step is to identify the set of available reactants of interest, for example, by performing substructure searches on in-house databases and databases of commercially available compounds. The next step is to filter out compounds that are known to be undesirable (such as high molecular weight compounds and compounds containing reactive groups). Compound selection is then required. This involves calculating molecular descriptors for the compounds followed by the application of a selection algorithm in order that, for example, a diverse set of compounds or a targeted set of compounds be selected. Such a step may also involve library enumeration if the selection is done through an analysis of product space.

In the next section, such steps are described for the design of a library of 2-aminothiazoles, see figure 8.1 below. Libraries of this type have been generated using solution-phase techniques and their preparation has been shown to be simple, practical and effective [19].

Figure 8.1. A library of 2-aminothiazoles



The combinatorial reaction involves the coupling of α-bromoketones and thioureas to form 2-aminothiazoles. Theoretically, it is possible to generate many tens of thousands of 2-aminothiazoles using readily available reactants, however, it is preferable to limit a library to 2,000 diverse products constructed from 40 α-bromoketones and 50 thioureas. In the section discussing methods, a procedure is described for selecting a diverse library through an analysis of reactant space. A method for selecting a diverse library through an analysis of product space that results in improved diversity is also described. Following this is an approach to selecting a library in product space that optimizes diversity and physical property profiles. Finally, a method of selecting a library focusing on a particular compound of interest is described.

8.2. Materials and methods

8.2.1. Materials

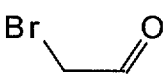
- ❑ sources of available reactants
- ❑ substructure searching software
- ❑ software for enumerating the virtual library of products if selection is to be performed in product space
- ❑ software for computing molecular descriptors
- ❑ compound selection software

8.2.2. Methods

A. Identifying candidate reactants

Generate the 2-aminothiazole product compounds by reacting α-bromoketones with thioureas. Identify candidate reactants for each of the α-bromoketone and thiourea reactant pools by performing substructure searches on the Available Chemicals Directory (ACD) [20]. The Daylight substructure searching software [10][51] can be used to search for α-bromoketones that contain the substructure shown in figure 8.2, and a second search can be conducted for thioureas that contain the substructure shown in figure 8.3.

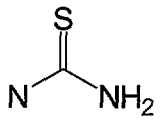
Figure 8.2. The substructure used to search for α-bromoketones



Note that the substructure should not be embedded within a ring system.

[51] The substructure searches for available reactions were performed using Merlin, a Daylight software product [10], however, other substructure searching packages are equally appropriate, such as MDL's ISIS [26].

Figure 8.3. The substructure used to search for thioureas



Note that substitution on the S atom should not be allowed and that explicit hydrogens should be attached to one of the N atoms.

At this stage, it can be useful to examine the subsets of structures retrieved in order to identify and remove structures that are inappropriate, for example, the search for α-bromoketones may find some structures where the substructure itself is embedded within a ring and which should not therefore take part in the reaction.

B. the reactant pools

The reactant sets identified in step A may still contain reactants that have undesirable characteristics, for example, highly reactive functional groups that could interfere with the intended reaction. In addition, the reactants may be present as salts. The ADEPT software [21] can be used to clean up each reactant set:[52] the parent structures should be extracted and retained where compounds are present as salts; reactants having a molecular weight greater than 300 and more than 8 rotatable bonds were removed; and a series of substructure searches should be performed to remove reactants that contain undesirable substructural fragments. An example of the functionality available in ADEPT is shown in figure 8.4 below. In one trial, after the filtering and clean-up operations had been performed on each of the reactant sets, there were 175 thioureas and 74 α-bromoketones remaining, respectively. This represents a virtual library of 12,950 potential products.

[52] The ADEPT [21] software is a GlaxoWellcome in-house program that has been developed to assist in the design of combinatorial libraries. ADEPT provides a user-friendly interface that allows non-experts to perform various functions such as substructure searches, used to filter reactant pools and library enumeration tasks. The functions performed by ADEPT can also be performed using commercially available software such as MDL's ISIS [26]. The authors of chapter 8, Valerie Gillet and Darren Green, wish to acknowledge the contribution of Andrew Leach for his expertise in developing the ADEPT computer program.

Figure 8.4. ADEPT filtering of undesirable compounds

Use this section to define the functionality that you require to be present in your compound list. Note that each query is treated *independently* by the search engine. Thus to specify primary or secondary amines it is only necessary to click the "Primary or secondary amine (aliphatic or aromatic)" button; it is not also necessary to click the aliphatic and aromatic amines on as well.

reactive groups			
carboxylic acid	Don't care	beta keto ester	Must NOT be present
alpha halo ketone	One only	nitrile	Don't care
isocyanate	Must NOT be present	primary alkyl amine	Don't care
secondary alkyl amine	Don't care	primary aryl amine	Don't care
alkyl alcohol	Don't care	phenol	Don't care
alkyl aldehyde	Must NOT be present	aryl aldehyde	Don't care
sulphonyl chloride or bromide	Must NOT be present	epoxide	Must NOT be present
anhydride	Must NOT be present	boronic acid or ester	Must NOT be present

C=O, C=S groups			
carboxylic acid	Don't care	amide	Don't care
ester	Don't care	anhydride	Don't care
thiocarbonyl	Don't care	thioamide	One only

substructures			
beta lactam	Must NOT be present	steroid	Must NOT be present

Generate 3D file? No

Format of 3D file? sdf/mol

Molecular weight range (leave blank if not important)

Lower value Upper value

Flexibility (rotatable bonds). Leave blank if not important.

Lower value Upper value

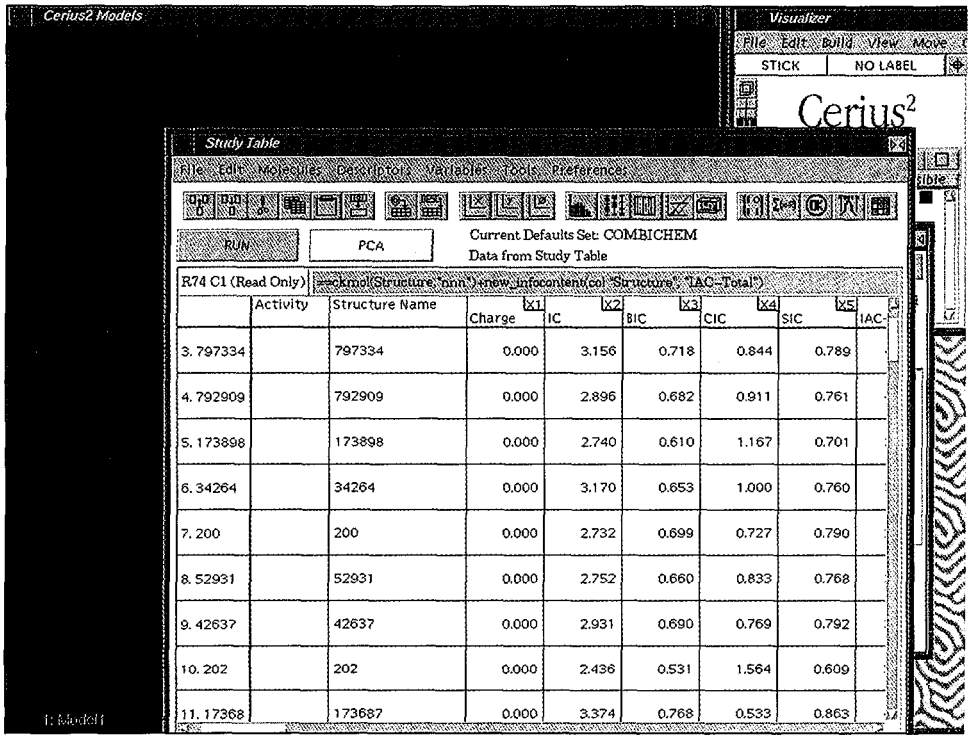
C. Selecting diverse subsets of reactants

The simplest library design method is to analyse each set of reactants independently and to produce diverse subsets of reactants that can then be used to generate a library of products. In this method, the assumption is made that a diverse combinatorial library will result from enumerating the diverse subsets of reactants. Reactant selection can be performed using the "Cerius^2" software [22].[53]

[53] "Cerius^2" [22] is commercially available software from Molecular Simulations Inc. Other commercial vendors providing software for designing combinatorial libraries include Tripos [27], Barnard Chemical Information Systems [28], Synopsys Scientific Systems [29], Chemical Design Ltd [11], Daylight Chemical Information Systems [10] and MDL Information Systems Inc. [26].

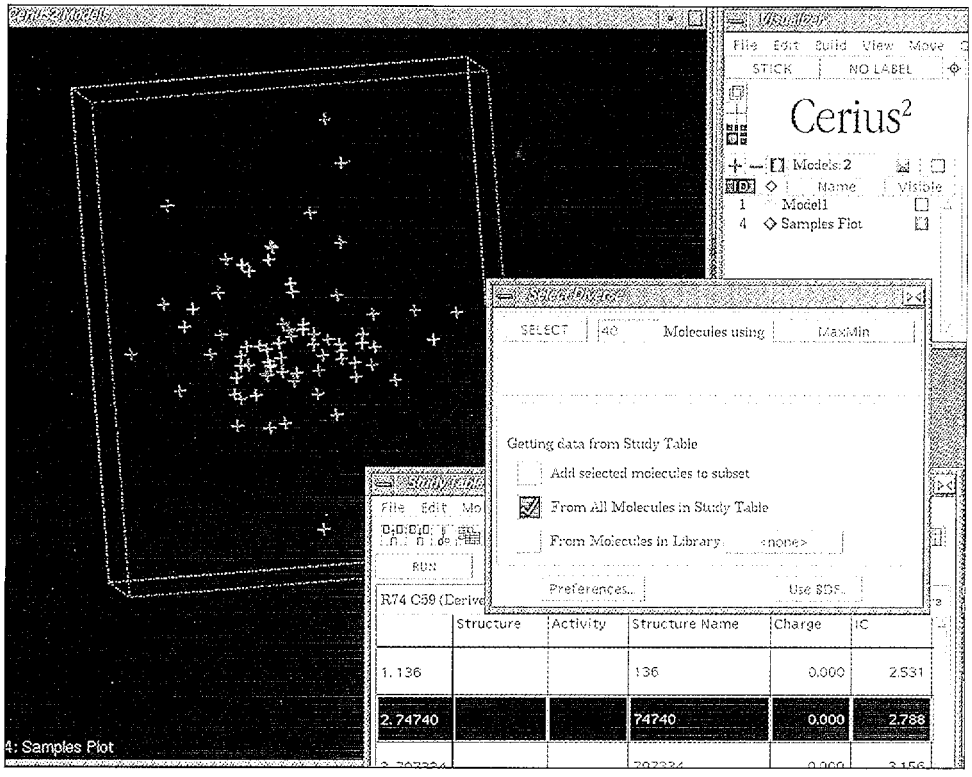
Analyse the α-bromoketones reactant set first with the aim of selecting the 40 most diverse reactants from the pool of 74 reactants. The 74 α-bromoketones can then be read into "Cerius^2" and the 50 default descriptors calculated (mainly topological indices but also AlogP, molecular weight, counts of hydrogen bond donors and acceptors), see figure 8.5 below. Perform a principal components analysis and plot the structures in the 3D space constructed from the first three principal components, as shown in figure 8.6 below. Several diversity metrics are available in "Cerius^2" for choosing a diverse subset of compounds. Select the 40 most diverse reactants according to the MaxMin distance-based target function.[54] Figure 8.6 shows the positions of the subset of reactants in the 3D space by red crosses. Repeat the process for the thioureas: the 174 thioureas read into "Cerius^2"; calculate the descriptors; perform the principal components analysis and select the 50 most diverse reactants using the MaxMin target function.

Figure 8.5. PCA analysis performed on the 74 α-bromoketones using the "Cerius^2" topological descriptors



[54] Distance-based compound selection algorithms have a tendency to favour outliers, however, they have the advantage of being rapid to calculate. Alternative compound selection techniques and molecular descriptors were discussed briefly early in the chapter. To date there is no clear picture of the relative effectiveness of the various methods.

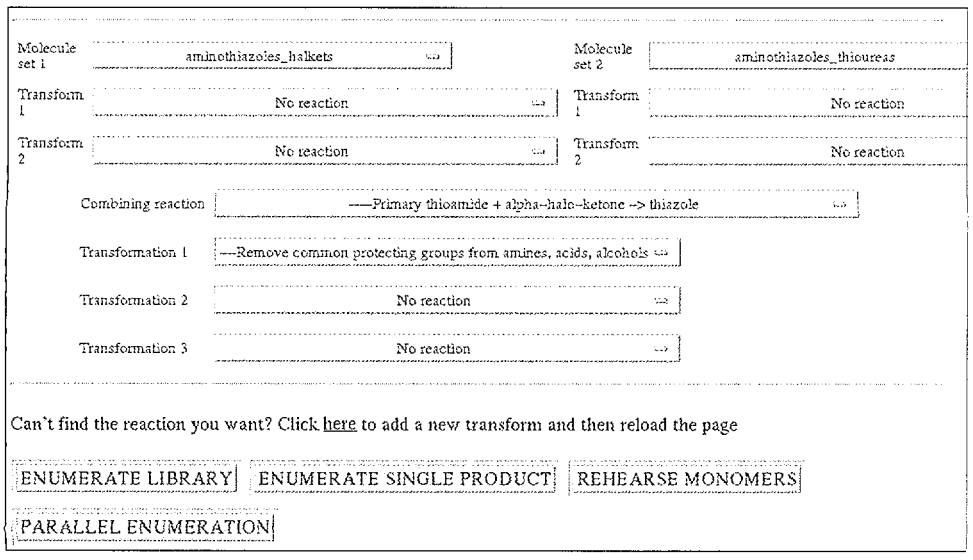
Figure 8.6. A subset of 40 diverse α-bromoketones



The subject is shown by small crosses, selected using the MaxMin target function. The compounds are plotted in the 3D space defined by the first three principal components.

Enumerate the library of 2,000 product molecules from the 40 and 50 subsets, remove any protecting groups and profile the physical property characteristics of the library. Library enumeration and profiling can be performed using ADEPT.[55] The functions available for library enumeration in ADEPT are shown in figure 8.7 below.

Figure 8.7. Functions available for library enumeration in ADEPT



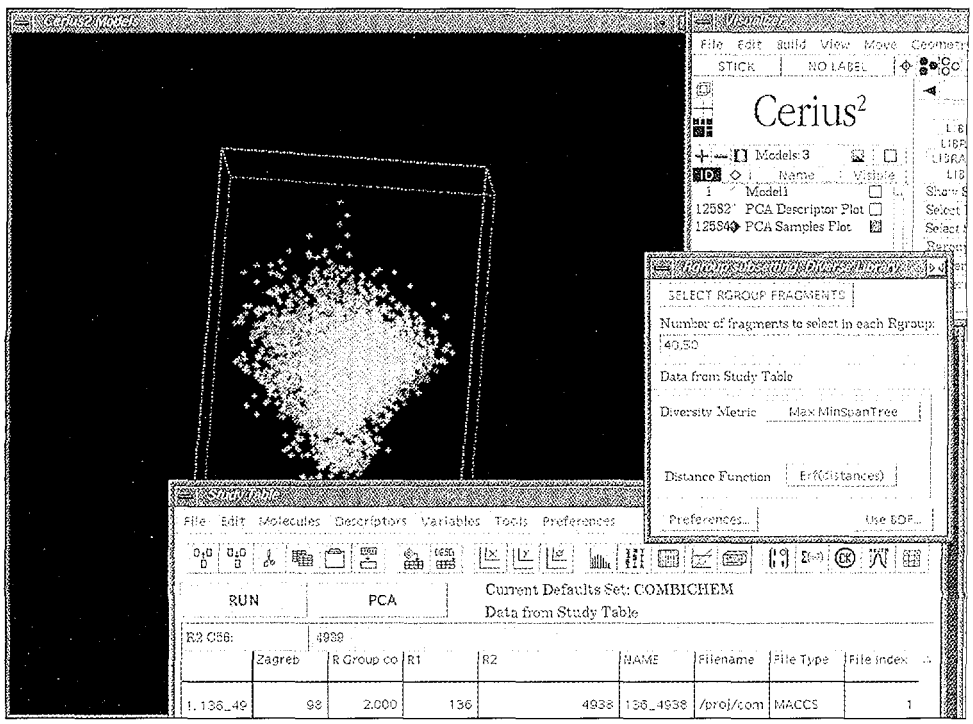
[55] The ADEPT library enumeration functionality is based on Daylight transformations [10]. There are a number of commercially available tools for library enumeration including products supplied by MSI [22], MDL [26], Tripos [27] and Synopsys Scientific Systems [29].

ADEPT is used to enumerate the combinatorial and remove protecting groups from the product molecules.

D. Selecting a combinatorial library in product space

As described earlier, there is evidence to suggest that greater diversity can be achieved through an analysis of product space. In addition, the pitfalls inherent in the application of, say, molecular weight filters to protected reactants is avoided. Product-based selection is a computationally more complicated and time-consuming process, however, the benefits resulting from a more optimized library can outweigh the extra computational cost. It is also possible to compare different configurations of libraries. For example, libraries of configuration 20 x 100 and 25 x 80 could be examined as well as 40 x 50, to assess the appropriateness of such a choice. In order to select a library through an analysis of product space, it is necessary to enumerate the virtual library of products. Library enumeration can be performed using ADEPT as described in step C, using the full reactant pools of 74 α-bromoketones and 174 thioureas. This should result in a virtual library of 12,950 product molecules. Following enumeration, commonly occurring protecting groups can be removed from the product molecules. The Rgroup subsetting facility in "Cerius^2" can then be used to select a diverse set of 2,000 product molecules that preserve the combinatorial constraint (that is, a combinatorial library constructed from 40 α-bromoketones and 50 thioureas): the virtual library can be read into "Cerius^2"; the same descriptors calculated for the product molecules as used for the reactant-based selection described in step C; principal components analysis performed and the molecules plotted in 3D space according to the first three principal components as shown in figure 8.8 below; the Rgroup subsetting feature of "Cerius^2" can then be used to select a diverse 40 x 50 library using the minimum spanning tree distance-based target function [23]. Figure 8.9 below shows the combinatorial library of 2,000 compounds by the small crosses. The characteristics of the library were calculated using ADEPT.

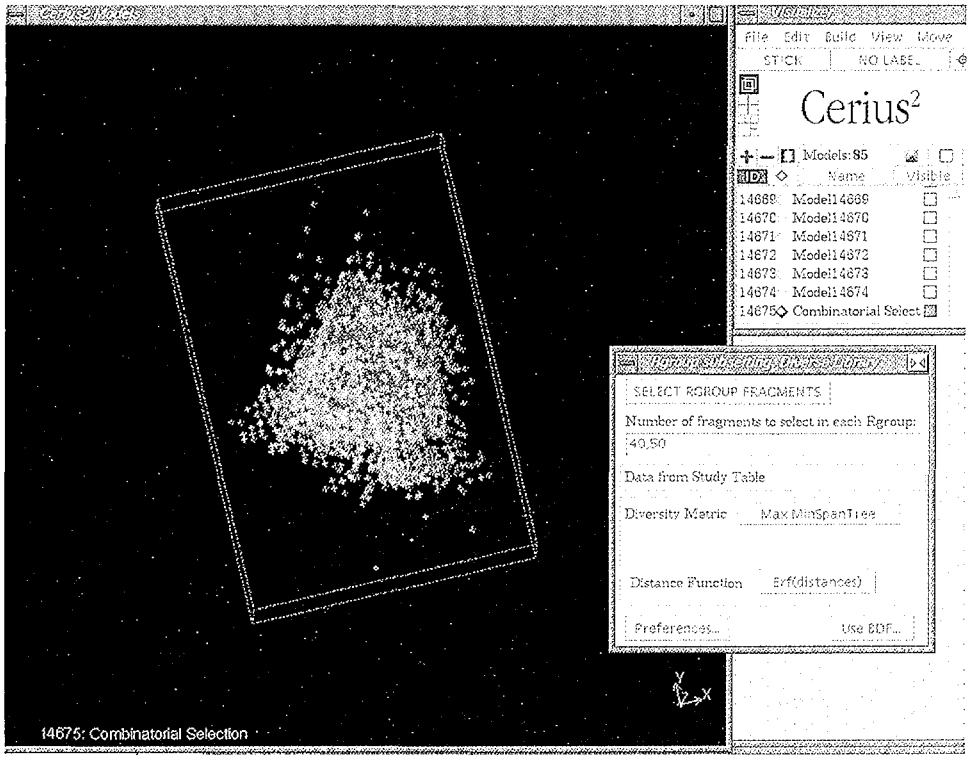
Figure 8.8. The Rgroup subsetting feature of "Cerius^2"



The Rgroup subsetting feature of "Cerius²" is used to select a diverse (40 x 50) library from the full virtual library of 12,950 products. The product molecules are plotted in the 3D space defined by the first three principal components derived from the topological descriptors.

The subset of 2,000 diverse 2-aminothiazoles selected by analysing the product space of the virtual library of 12,950 products is shown by small crosses. This subset represents a combinatorial library.

Figure 8.9. The subset of 2,000 diverse 2-aminothiazoles



The subset of 2,000 diverse 2-aminothiazoles selected by analysing the product space of the virtual library of 12,950 products is shown by small crosses. This subset represents a combinatorial library.

E. Selecting a combinatorial library optimized on diversity and physicochemical properties

The library selected in step D should be optimized for diversity, however, many of the compounds in the final library will have higher ClogP than is considered desirable for bioactive compounds. The average ClogP for the compounds selected in step D, for example, is 5.2 (with standard deviation 2.0). This compares with an average ClogP of 2.3 for compounds in the World Drugs Index (WDI) [24]. It would be ideal to be able to select a diverse subset within a "drug-like" space. In a trial, SELECT [15] was the programme used to design a 40 x 50 library that is optimized both for diversity and ClogP.[56] SELECT optimized the ClogP profile of the library by minimizing the RMS

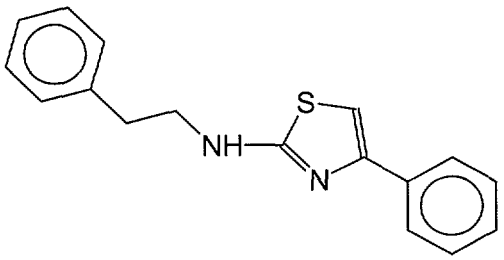
[56] SELECT [15] can be tailored to design libraries that are optimized for any physicochemical property that can be calculated for the compounds in the library, for example, libraries can be designed to have a distribution of molecular weights or rotatable bonds that are similar to the distributions found in the WDI (or some other collection of interest). Multiple properties, including diversity, can be optimized simultaneously.

difference between the library's profile and the ClogP profile of the WDI (which represents a collection of "drug-like" molecules). The average ClogP for the SELECT designed library is 4.1 (standard deviation 2.1), which represents a significant improvement in terms of the characteristics of the compounds contained within the library as compared with a library selected on diversity alone. To lower this even more, it would be necessary either to search for more hydrophilic reactants to incorporate in the library, to reduce the library size to limit the number of lipophilic compounds, or to vary the library configuration, should one of the reactant sets contain a high percentage of lipophilic compounds.

F. Selecting a focused combinatorial library

The final library design procedure is used to design a combinatorial library that is focused around a particular compound of interest. Assume one of the previous diverse libraries was synthesized and tested in a high throughput assay; now assume that one active 2-aminothiazole is found (see figure 8.9 above). The structure is fanetizole (see figure 8.10 below), a known anti-inflammatory agent. It would then be normal practice to make a small number of analogues of this "hit" to study SAR, since according to the similar property principal [25] it is likely that compounds close to fanetizole in structure space are also biologically active compounds. Thus, in the final example, a combinatorial library of 80 compounds (1 "plate" of compounds) can be designed in product space around fanetizole.

Figure 8.10. Fanetizole



The design is carried out using the Rgroup subsetting procedure in "Cerius²". Read in the virtual library of 12,950 compounds and calculate the descriptors as before. In this way the target compound, fanetizole, is identified. Figure 8.11 below shows the 8 x 10 focused combinatorial library by the small crosses. A subset of this focused library is shown in figure 8.12.

Figure 8.11. An 8 x 10 combinatorial library focused around fanetizole (shown by small crosses)

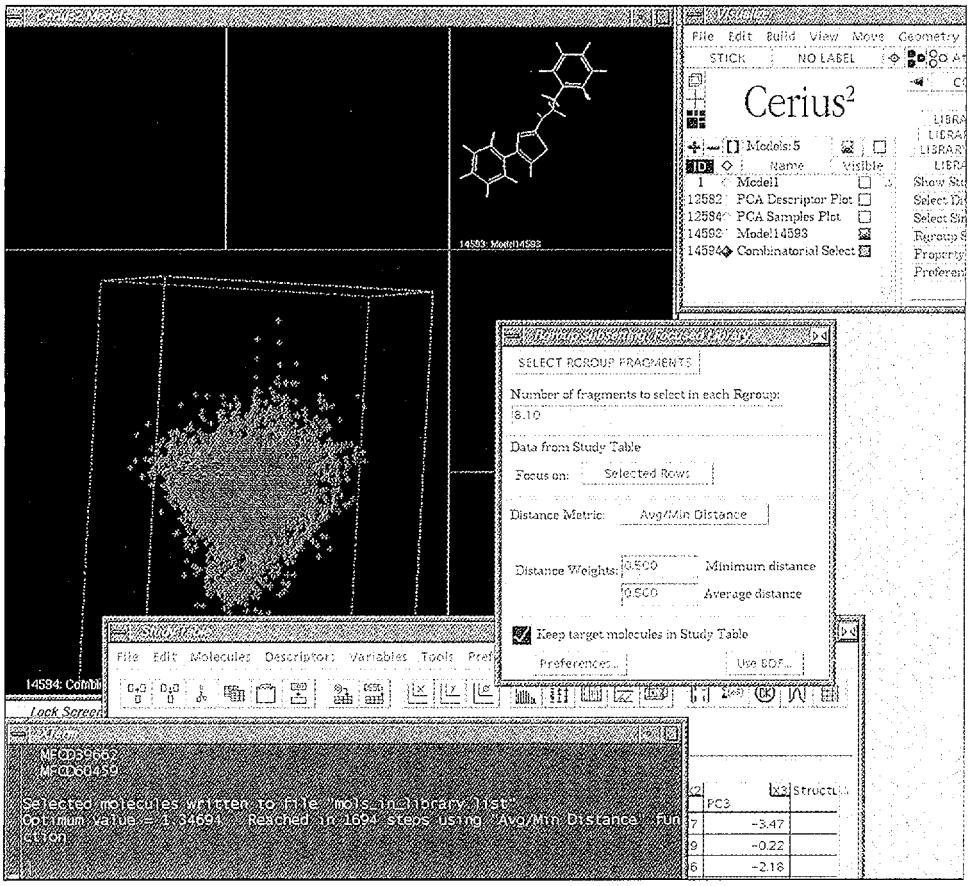
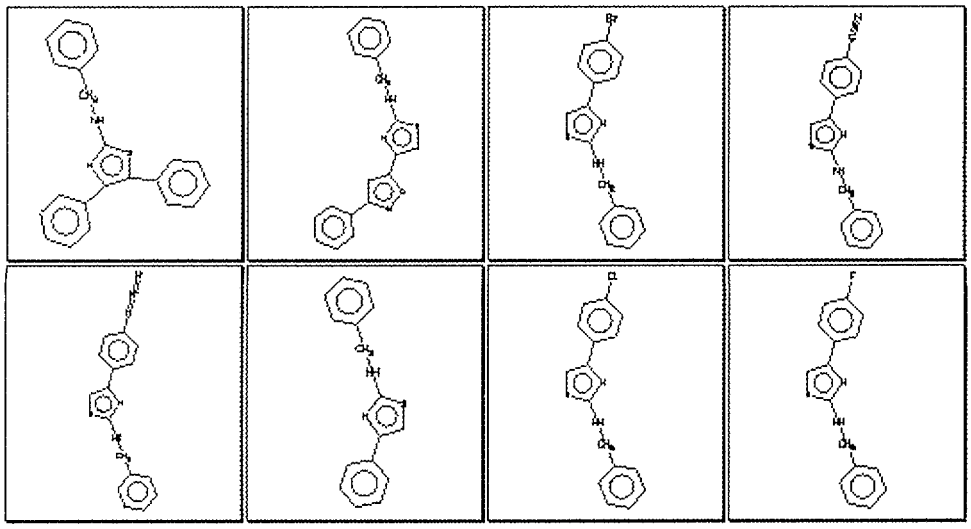


Figure 8.12. A sample of the 8 x 10 library focused around fanetizole



References

1. Gillet, V. J., Willett, P. and Bradshaw, J. 1998. "Identification of Biological Activity Profiles Using Substructural Analysis and Genetic Algorithms". *J. Chem. Inf. Comput. Sci.*, vol. 38, pp. 165-179.
2. Ajay, Walters, W. P. and Murcko, M. 1998. "Can We Learn to Distinguish between 'Drug-like' and 'Nondrug-like' Molecules?". *J. Med. Chem.*, vol. 41, pp. 3314-3324.
3. Sadowski, J. and Kubinyi, H. 1998. "A Scoring Scheme for Discriminating between Drugs and Nondrugs". *J. Med. Chem.*, vol. 41, pp. 3325-3329.
4. Brown, R.D. and Martin, Y. C. 1996. "Use of Structure-Activity Data to Compare Structure-Based Clustering Methods and Descriptors for Use in Compound Selection". *J. Chem. Inf. Comput. Sci.*, vol. 36, pp. 572-584.
5. Brown, R. D. "Descriptors for Diversity Analysis". 1997. *Perspect. Drug Disc. Des.*, vol. 7/8, pp. 31-49.
6. Agrafiotis, D. K. 1998. "Molecular Diversity". *Encyclopedia of Computational Chemistry*, (ed. P. v. R. Schleyer), Wiley, New York.
7. Molconn-X, 1993, available from Hall Associates, Massachusetts, United States.
8. MACCS II, Molecular Design Ltd., San Leandro, California, United States.
9. UNITY Chemical Information Software, Tripos Inc., Missouri, United States.
10. Daylight Chemical Information Systems, Inc., Mission Viejo, California, United States.
11. Chem-X, Chemical Design Ltd., Oxford Molecular Group, The Medawar Centre, Oxford Science Park, Oxford, United Kingdom.
12. Lajiness, M. S. 1997. "Dissimilarity-Based Compound Selection Techniques". *Perspect. Drug Disc. Des.*, vol. 7/8, pp. 65-84.
13. Dunbar Jr., J. B. 1997. "Cluster-Based Selection". *Perspect. Drug Discov. Design*, vol. 7/8, pp. 51-63.
14. Mason, J. S. and Pickett, S. D. 1997. "Partition-Based Selection". *Perspect. Drug Disc. Des.*, vol. 7/8, pp. 85-114.
15. Gillet, V. J., Willett, P. , Bradshaw, J. and Green, D. V. S. 1999. "Selecting Combinatorial Libraries to Optimize Diversity and Physical Properties". *J. Chem. Inf. Comput. Sci.*, vol. 39, pp. 169-177.
16. Gillet, V. J., Willett, P. and Bradshaw, J. 1997. "The Effectiveness of Reactant Pools for Generating Structurally-Diverse Combinatorial Libraries". *J. Chem. Inf. Comput. Sci.*, vol. 37, pp. 731-740.
17. Gillet, V. J. 1999. "Computational Aspects of Combinatorial Chemistry". *Combinatorial Chemistry and Combinatorial Technologies: Methods and Applications* (S.

Miertus and G. Fassina, eds.), Marcel Dekker Inc., New York.

18. Warr, W. A. 1997. "Commercial software systems for diversity analysis". *Perspect. Drug Disc. Des.*, vol. 7-8, pp. 115-130.
19. Bailey, N., Dean, A. W., Judd, D. B. and Middlemiss, D. 1996. "A convenient procedure for the solution-phase preparation of 2-aminothiazole combinatorial libraries". *Bioorg. Med. Chem. Letts.*, vol. 6, p. 1409.
20. The Available Chemicals Directory, distributed by MDL Information Systems Inc., California, United States.
21. Leach, A. 1998. "An ADEPT approach to library enumeration and design", http://www.daylight.com/meetings/emug98/Leach/index.html
22. Cerius^2, Molecular Simulations Inc., California, United States.
23. Waldman, M. 1998. "Novel Metrics for the Optimization of Molecular Diversity of Combinatorial Libraries", presented at the Computational Approaches to the Design and Analysis of Combinatorial Libraries, Molecular Graphics and Modelling Society (MGMS) and the Chemical Structure Association (CSA), Sheffield, United Kingdom.
24. The World Drug Index (WDI), maintained by Derwent Publications Ltd., London, United Kingdom.
25. Johnson, M. A. and Maggiora, G. M. 1990. *Concepts and Applications of Molecular Similarity*, John Wiley, New York.
26. MDL Information Systems Inc., California, United States.
27. Tripos Inc., Missouri, United States.
28. Barnard Chemical Information Systems, Sheffield, United Kingdom.
29. Synopsys Scientific Systems, Leeds, United Kingdom.

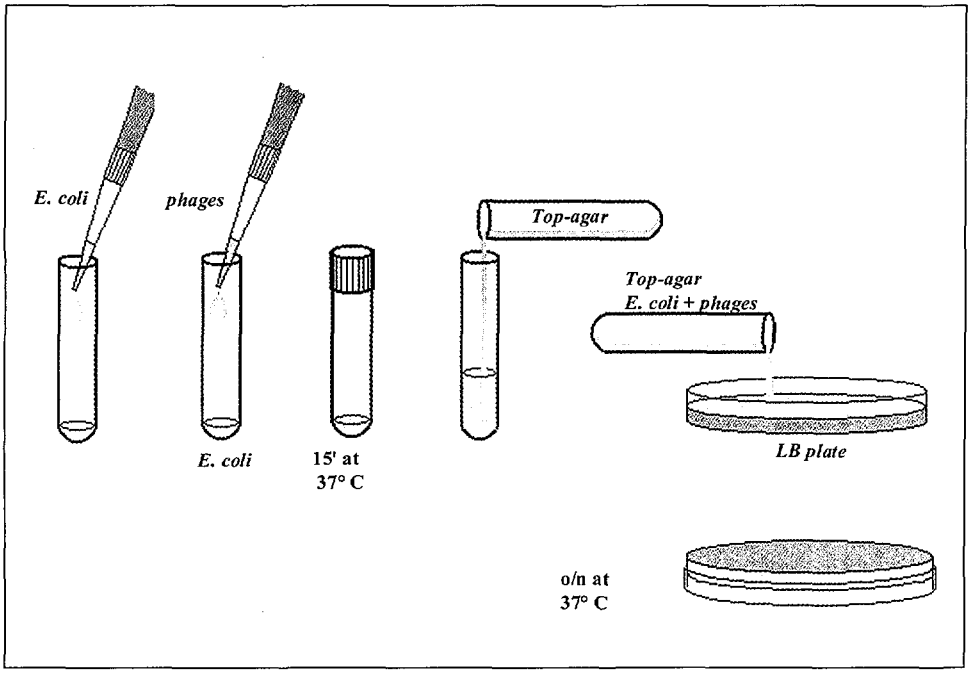
9

Growth of phages and bacteria for biopanning experiments

M. Dani

9 Growth of phages and bacteria for biopanning experiments

Maria Dani



- ❑ Appropriate bacterial strain growth
- ❑ Phage growth on solid medium
- ❑ Phage growth on liquid medium
- ❑ Single phage clones isolation

9.1. Introduction

Large pools of peptides can be displayed on the surface of phages [1, 2]. The peptides are fused to a viral coat protein. The protein has the function of exposing the peptides to the medium and consequently allows their interaction with specific target molecules.

Phage display libraries can be used for the selection of peptide ligands by a simple *in vitro* procedure called biopanning. In the most simple protocol, the phages displaying the peptides are incubated with the immobilized target. The bound phages are then eluted by disrupting the binding with the target molecules and amplified. This selection cycle is usually repeated three times in order to enrich the phages with the tightest binding peptides. At the end of the selection process, individual clones are isolated, their DNA extracted and sequenced.

The above steps involve the growth and preparation of viable viral particles. This is achieved through basic microbiology techniques, such as growing and storing bacterial strains, growing phages in liquid and solid medium, phage titering and isolating single plaques isolation [3, 4].

Since the most commonly used phage for display libraries is the filamentous M13 type, the focus of the following discussion will be on such bacteriophage methodologies.

The filamentous phages (M13, f1, fd) [5] are viruses that infect male *E. coli* strains that contain an *F pilus*. They are rod-shaped and their single-stranded DNA genome is a circular DNA molecule that is 6,407 nucleotides long, which codes for ten proteins and contains the origin of replication, a transcription terminator and the signal for packaging the (+) DNA strand into phage particles [6]. The viral coat is composed primarily of about 2,700 copies of pVIII protein and several minor coat proteins at the ends. One of the minor coat proteins (pIII) is essential for the *E. coli* infection because pIII attaches to the receptor at the tip of the *F pilus*, allowing the phage to inject its DNA inside the bacterium membrane [7]. When the DNA is inside the cell, it starts to replicate. The viral coat proteins are then synthesized by using *E. coli* protein synthesis machinery. New viral particles are formed at the membrane levels and exported out of the cell [8]. Ordinarily, filamentous phages do not lyse the host cells, although they slow the rate of growth by about two-fold.

In order to be infected by filamentous phages, the bacterial strain must carry the F episome (F^+, F' or Hfr). F^+ strains containing the Sup E suppressor mutation (such as XL1-Blue, DH5aF') can be used for phage propagation. However, if a commercial display library or vector is purchased, it is better to use the supplied strain which is certainly tested in that particular system.

E. coli is a rod-shaped bacterium that can grow on solid and liquid medium. In liquid culture, a small number of cells is inoculated into sterile medium. The medium can be minimal and contain salts (such as nitrogen, phosphorus and trace metals) and glucose, which provides a source of carbon and energy. Alternatively, a rich medium may contain other metabolites, such as amino acids, vitamins and nucleotide precursors. For M13 growth or for liquid cultures, it is better to use a rich medium such as LB. After a lag period, the cells start to divide and grow exponentially (log phase), doubling, in a rich medium, the number of cells every 20 to 30 minutes. In laboratory conditions, when the culture reaches a density of 1 to 2 x 10^9 cells/ml, the cells stop dividing owing to both the lack of nutrients and oxygen, and the accumulation of waste products. Growing bacteria on solid medium is usually used for single-colony isolation and titering. Bacterial cells are either spread or streaked on top of plates containing solid medium and grown for 12 to 24 hours. Well separated cells grow as round colonies. The size is usually 1 mm in diameter but can vary depending on colony density and time. Because the F episome is maintained in minimal medium, it is strongly recommended that single colonies from minimal plates only be picked. When plating bac-

teria on minimal medium, the growth is slow and it takes from 18 to 24 hours at 37° C before colonies are visible.

During biopanning experiments, the phages must be grown, purified and titrated. Phages can be grown in the presence of *E. coli* both in liquid and solid medium. For the growth in liquid medium an appropriate amount of phages is added to a bacteria culture in the early log phase or to a diluted overnight bacteria culture. The culture is then incubated for 4.5 hours with vigorous agitation at 37° C. Phage particles are purified from the medium by centrifugation of bacteria cells and precipitation of supernatant viral particles with PEG (polyetilene-glycol).

For the growth on solid medium, an appropriate amount of phages is added to a small amount of mid-log bacteria (usually 200 μl). After incubating for about 15 minutes incubation at 37° C (necessary for bacteria infection), a melted top agar is added and the mixture poured on top of LB plates. When the top agar is solidified, the plates are incubated overnight at 37° C or until plaques are visible. Since the M13-type phage does not lyse the cell, the plaques will not appear as completely clear areas but will be visible when plates are held against a light. The size of the plaques varies depending on the phage concentration. If the number of phages per plate is too high (for example, more than about 10,000), the plaques will be confluent and not visible.

The entire procedure must be carried out in sterile conditions. It is also important to separate bacterial culture from phage cultures in order to avoid contamination with phages and to use aerosol-resistant tips.

9.2. Materials and methods

9.2.1. Materials

- ❑ F' *E. coli* strain
- ❑ filamentous phages
- ❑ sterile disposable Petri dishes for bacterial culture (95 mm)
- ❑ sterile Pyrex or disposable Erlenmeyer flasks of appropriate size (e.g. 100 ml, 250 ml, 500 ml)
- ❑ sterile inoculating loops, aerosol-resistant tips, Eppendorf tubes, pipettes
- ❑ 10 ml sterile disposable tubes (snap caps, round bottom)
- ❑ sterile low-speed and medium-speed centrifuge tubes
- ❑ 3-5 ml screw-capped vials

LB medium

- ❑ per litre: 10 g bacto-tryptone, 5 g yeast extract, 5 g NaCl. Dissolve, autoclave and store at rt.

LB agar plates

- ❑ per litre: 10 g bacto-tryptone, 5 g yeast extract, 5 g NaCl, 15 g agar. Autoclave, cool to < 70° C, pour on sterile disposable Petri dishes and let the medium solidify on a flat surface.

Minimal plates

- Prepare separately:

  2 x M9 salts: per litre: 12 g Na_2HPO_4, 6 g KH_2PO_4, 1 g NaCl, 2 g NH_4Cl; autoclave;

  3% agar: 15 g agar in 500 ml H_2O; autoclave;

  1 M $MgSO_4$; autoclave;

  1 M $CaCl_2$; autoclave;

  thiamine 10 mg/ml; filter sterilize;

  20% glucose: 20 g glucose in H_2O; dissolve and filter sterilize.;

Autoclave or sterilize as indicated all components separately and cool to < 70° C before combining as follows:

500 ml of 2 x M9 salts, 500 ml 3% agar, 20 ml 20% glucose, 2 ml $MgSO_4$, 0.1 ml 1 M $CaCl_2$, 1 ml thiamine (10 mg/ml). Pour immediately on sterile disposable Petri dishes and let the medium solidify on a flat surface.

Top agar

per litre: 10 g bacto-tryptone, 5 g yeast extract, 5 g NaCl, 1 g $MgCl_2$ x $6H_2O$, 7 g agar. Autoclave and dispense into 50 or 100 ml aliquots; store at rt and melt in the microwave oven when needed.

PEG/NaCl

20% (w/v) polyethylene-glycol 8000, 2.5 M NaCl; autoclave and store at rt.

SM

per litre: 5 g NaCl, 2 g $MgSO_4$, $7H_2O$, 50 ml 1 M Tris-Cl, pH 7.5, 0.01% gelatin; autoclave and store at rt.

TBS

50 mM Tris-HCl, pH 7.5, 150 mM NaCl; autoclave and store at rt.

9.2.2. Methods

A. Bacteria culture on solid medium[57]

1. Prepare minimal agar plates as indicated in the section on materials, above, by adding agar at the final concentration of 15 g/l, to 2 x M9 salts and other ingredients after the autoclaving temperature has cooled off to between 50° C and 60° C. Pour

[57] For the selection of male *E. coli* cells, bacteria are grown on minimal agar plates.

the medium into plates and let solidify at rt. Since freshly poured plates are wet, it is better to dry the plates by leaving them with lids off for about 30 min in a laminar flow hood. Plates can be stored up to several weeks in a plastic bag in the refrigerator.

2. Streak an inoculum of bacteria, either from a colony or from a liquid culture across a minimal agar plate several times using an inoculating loop. Incubate the plate inverted overnight at 37° C in a static incubator.[58]

B. Bacteria culture in liquid medium[59]

1. Inoculate a single colony into an Erlenmeyer flask containing 10 ml-50 ml LB medium. For large cultures (more than 500 ml), it is necessary to inoculate first a smaller volume, grow for 8 h to overnight at 37° C and then dilute the culture 1:100 in fresh medium.

2. Incubate in an orbital shaker incubator for a few hours to overnight at 37° C, 250 rpm.

3. Store at 4° C for up to a few weeks.

C. Storage of bacteria[60]

1. Preparation of stabs: Pour 2 ml-3 ml of LB agar medium into small, sterile screw-capped vials. Inoculate bacteria with a sterile inoculating needle by picking a single colony from a plate and inserting it into the agar until it reaches the bottom of the vial. The cap should be kept loosened. Incubate the vials overnight at the appropriate temperature (usually 37° C), then tightly cap and store in the dark at rt.

2. Preparation of glycerol stocks: Inoculate 10 ml of LB liquid medium with a single colony from a minimal plate and grow the cells overnight or to a late log phase. Add sterile glycerol to a final concentration of 15%, mix well and store in aliquots at −80° C.

---

[58] Do not use air flow incubators, since the plates will dry out and the bacteria will not be able to grow. Store the plate sealed at 4° C for up to a few weeks. This plate constitutes the stock from which to isolate single colonies for bacterial culture.

[59] A liquid medium is prepared by dissolving in water and autoclaving several compounds like salt and nutrient mixture (e.g. yeast extract). There are different media and they must be chosen based on the bacterial strain and application of the culture. The most commonly used medium is LB. In order to give a good aeration to the culture, Erlenmeyer flasks are used and the amount of liquid medium should be no more than 1/5 of the flask capacity. Vigorous agitation (250 rpm-300 rpm) is also applied to ensure proper aeration. The time needed for the culture to grow varies. The growth can be divided into 3 phases: a short lag phase in which the bacteria become acclimated and start to divide, a log phase in which they grow rapidly (usually once every 20-30 min) and a stationary phase characterized by a slow growth and eventually by lack of cell division. The stationary phase corresponds to an *E. coli* concentration of about 1 x 10^9 cells/ml and can be reached in 6 h to 8 h. If a culture in the early log-phase is needed, bacteria are grown during the day for up to a few hours.

[60] There are several bacterial storage methods in use. For a period of a few weeks, bacteria can be stored on agar plates, tightly wrapped in parafilm and stored inverted at 4° C. For longer storage (a year or two), bacteria can be maintained in stab cultures. Alternatively, they can be stored for many years in media containing 15% glycerol at low temperature without significant loss of viability. Bacteria can be recovered from glycerol stocks by scratching the frozen surface of the stock with an inoculating loop and returning immediately the frozen suspension back to the freezer.

D. Storage of phages

1. Short-term storage:[61] Grow phages in liquid culture for 4.5 h at 37° C, 250 rpm as described below in the protocol "E phage amplification in liquid medium". Spin cells by centrifuging for 10 min at 10,000 rpm and transfer the supernatant to a fresh tube. Store at 4° C for up to a few weeks. Alternatively, grow phages on LB agar plates overnight at 37° C as described below. Store plates at 4° C for up to a few weeks.

2. Long-term storage: To store for a few years, add glycerol to the phage suspension at a final concentration of 50%. Store in aliquots in tightly capped tubes at −80° C.

E. Phage amplification in liquid medium[62]

1. Pick a single bacterial colony from a minimal plate and inoculate 10 ml-50 ml of LB in a 100 ml-250 ml Erlenmeyer flask. Grow overnight at 37° C, 250 rpm. The day after, make a 1:100 dilution of the overnight bacteria culture in a final volume of 20 ml LB medium and grow until it reaches early log phase (< 0.01 OD_{600}).

2. Amplify the eluted phages by mixing them with this exponentially growing bacteria and incubate the culture at 37° C, 250 rpm, for 4.5 h to 5 h.[63]

3. Transfer the liquid culture to a sterile centrifuge tube and spin for 10 min at 4° C, 10,000 rpm (Sorvall SS-34 or equivalent).

4. Collect 4/5 of the supernatant and discard the cells containing pellet.

5. If phages are used for the biopanning experiment, precipitate the phages twice with PEG/NaCl in order to obtain purified phage particles. Add 1/6 volume of PEG/NaCl and allow phages to precipitate at 4° C for at least 60 min, preferably overnight. Spin the PEG precipitate 10 min at 10,000 rpm, 4° C. Decant supernatant and re-spin briefly in order to remove residual supernatant. Suspend the phage pellet in 1 ml TBS, transfer the suspension to a microcentrifuge tube and spin for 5 min at 10,000 rpm to pellet residual cells. Transfer the supernatant to a fresh tube and precipitate again with 1/6 of PEG/NaCl. After incubating 30 to 60 min on ice, spin 10 min at 10,000 rpm, 4° C in a microcentrifuge and discard supernatant completely (re-spin briefly). Re-suspend the phage pellet in a small volume (200 μl) of TBS.

6. Titer the phage culture as described in protocol "G phage titration".

F. Phage amplification on solid medium[64]

1. Inoculate 5 ml LB with a single bacterial colony from a minimal plate or, alternatively, dilute an overnight culture 1:100 in LB. Grow the cells at 37° C, 250 rpm, for about 2 h, until early to mid-log phase.

2. Mix 2 ml of the bacteria with the eluted phages, incubate at 37° C for 15 min.

---

[61] For better storage, precipitate phages with PEG and re-suspend them in a small volume of SM or TBS. More concentrated phage suspensions (e.g. more than 109-1010 pfu/ml) will have a slower decrease in viable particles with time, compared to less concentrated suspensions. For solid medium storage, be sure to keep plates inverted and check for water condensation on the cover lid and contamination before using them.

[62] When working with phages, care must be taken to avoid cross contamination. Use aerosol-resistant pipette tips for all protocols. Avoid spilling and clean the hood often.

[63] Longer incubations may result in DNA insert deletions and are not recommended.

[64] If phage amplification in liquid culture results in a low titer, it is possible to amplify the phages on solid medium.

3. Add 30 ml melted top agar (50° C) and pour 3 ml into each LB plate. Let the top agar harden and incubate overnight at 37° C.

4. After incubation, put 3 ml of SM on top of each plate and incubate at 4° C for 2 h to 3 h. Collect the SM phage containing solution. Centrifuge the SM phage suspension for 10 min at 4° C and 10,000 rpm (Sorvall SS-34 or equivalent). Collect 4/5 of the supernatant and discard the cells containing pellet.

5. If phages are used for the biopanning experiment, precipitate the phages twice with PEG/NaCl in order to obtain purified phage particles. For this purpose, add 1/6 volume of PEG/NaCl and follow the same procedure described for phage amplification in liquid medium. At the end, re-suspend the phage pellet in a small volume (200 μl) of TBS. Titer the phage culture as described below.

G. Phage titration

1. Streak a suitable F^+ strain on a minimal plate. Grow at 37° C for 18 h to 24 h. Store sealed at 4° C for up to a few weeks.

2. Inoculate 5 ml LB with a single bacterial colony from a minimal plate or, alternatively, dilute an overnight culture 1:100 in LB. Grow the cells at 37° C, 250 rpm, for 2 h to 3 h, until early to mid-log phase.

3. Melt top agar in a microwave oven and equilibrate the temperature at 45° C to 50° C in a water bath. Warm up the necessary LB plates in the incubator at 37° C for about 1 h. Add 0.2 ml of the grown bacteria culture to each 10 ml of disposable tissue-culture tube.

4. Make serial dilution of phages in SM (e.g. 10^{-3}, 10^{-5}, 10^{-7}, 10^{-9}). Add 10 μl-100 μl of phage dilution to each tube containing bacteria, according to the amount of pfu (plaque-forming units) required in each plate. Label the tubes. If a high titer is expected (e.g. after phage amplification), make serial dilution of phages in SM from 10^{-8} to 10^{-11}. If a low titer is expected (e.g. after panning), place 1 μl directly and a 10^{-2} dilution.[65]

5. Incubate bacteria and phages for 15 min at 37° C, then add 3 ml of 45° C top agar and pour immediately on top of pre-warmed (37° C) LB plates.[66] Let the top agar harden on a flat surface, then incubate the plates upside down overnight at 37° C in a static incubator.

6. After incubation, count the plaques and multiply each number by the dilution factor for the plate and then calculate the phage titer in pfu/ml.[67]

---

[65] It is very helpful to include a control plate that contains *E. coli* only. If the plate has plaques, this means that the bacterial culture has been contaminated with some airborne phages and it is necessary to repeat the experiment with fresh cells.

[66] Work quickly with one tube at a time and keep the top agar warm. If the top starts to solidify before it is poured on top of plates, the plaques will not be visible.

[67] M13 is a non-lytic phage and does not produce clear plaques, only areas of diminished cell growth. In order to count the plaques, hold the plate up to a light. If plaques are confluent, they will not be visible. If plaques are present in the plate that should contain bacteria only, then a contamination has occurred. Counting the same number of plaques on different phage dilution is another sign of contamination. If this happens, discard everything and repeat the experiment with freshly prepared bacteria directly from a minimum plate.

H. Single plaque isolation and amplification[68]

1. Inoculate 5 ml LB with a single bacterial colony from a minimal plate or, alternatively, dilute an overnight culture 1:100 in LB. Grow the cells at 37° C, 250 rpm, for about 2 h to 3 h, until early to mid-log phase.

2. Melt top agar in a microwave oven and equilibrate the temperature at 45° C to 50° C in a water bath. Warm up the necessary LB plates in the incubator at 37° C for about 1 h. Add 0.2 ml of the grown bacteria culture to each 10 ml disposable tissue-culture tube. Make the appropriate dilution of phages in SM. Add 100-200 phages to each tube containing bacteria. Incubate 15 min at 37° C, then add 3 ml top agar to each tube and pour the suspension on top of LB plates. Grow overnight at 37° C.

3. Prepare a fresh bacteria culture by inoculating a single colony into 5 ml-10 ml of LB medium. Grow overnight at 37° C, 250 rpm.

4. The day after, sterilize several Pasteur pipettes and use them to pick a small top agar plug containing a well-isolated plaque. Transfer the plaque into a 10 ml conical tube containing 2 ml of a 1:100 dilution in LB medium of an overnight bacteria culture. Grow at 37° C, 250 rpm for 4.5 to 5 h.

5. Transfer the liquid culture to 2 ml Eppendorf tubes and centrifuge for 10 min at 4° C, 10,000 rpm (Eppendorf microcentrifuge). Collect 4/5 of the supernatant and discard the cells containing pellet.[69]

---

[68] The best way to isolate single plaques is by plating serial dilution of a phage suspension. If the phage concentration is known, simply plate the amount of phages needed. Be sure to obtain plates with no more than 100 pfu-200 pfu, in order to have well-isolated plaques. Keep in mind that the titer can vary slightly, depending on the growth of the plating bacteria and the dilution error.

[69] The supernatant containing the viral particles can be used for DNA isolation and sequencing or as a phage stock.

References

1. Smith, G. P. 1985. "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface". *Science,* vol. 228, pp. 1315-1317.
2. Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L. and Curtiss, R. 1997. "Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines". *Nat. Biotechnol.,* vol. 15, pp. 29-34.
3. Miller, J. 1972. *Methods in molecular genetics.* Cold Spring Harbor Laboratory, New York.
4. Ausubel, F. M., Brent, R, Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. *Current protocols in molecular biology.* Wiley, New York, 1.1.1-1.3.5.
5. Marvin, D. A., Hale, R. D., Nave, C. and Citterich, M. H. 1994. "Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe". *J. Mol. Bio.,* vol. 235(1), pp. 260-286.
6. Denhardt, D. J., Dressler, D., and Ray, D. S. 1978. *The single-stranded DNA phages.* Cold Spring Harbor Laboratory, New York.
7. Deng, L. W., Malik, P. and Perham, R. N. 1999. "Interaction of the globular domains of pIII protein of filamentous bacteriophage fd with the F-pilus of *Escherichia coli*". *Virology,* vol. 253(2), pp. 271-277.
8. Day, L. A., Marzec, C. J., Reisberg, S. A. and Casadevall, A. 1988. "DNA packing in filamentous bacteriophages". *Annu. Rev. Biophys. Biophys. Chem.,* vol. 17, pp. 509-539.

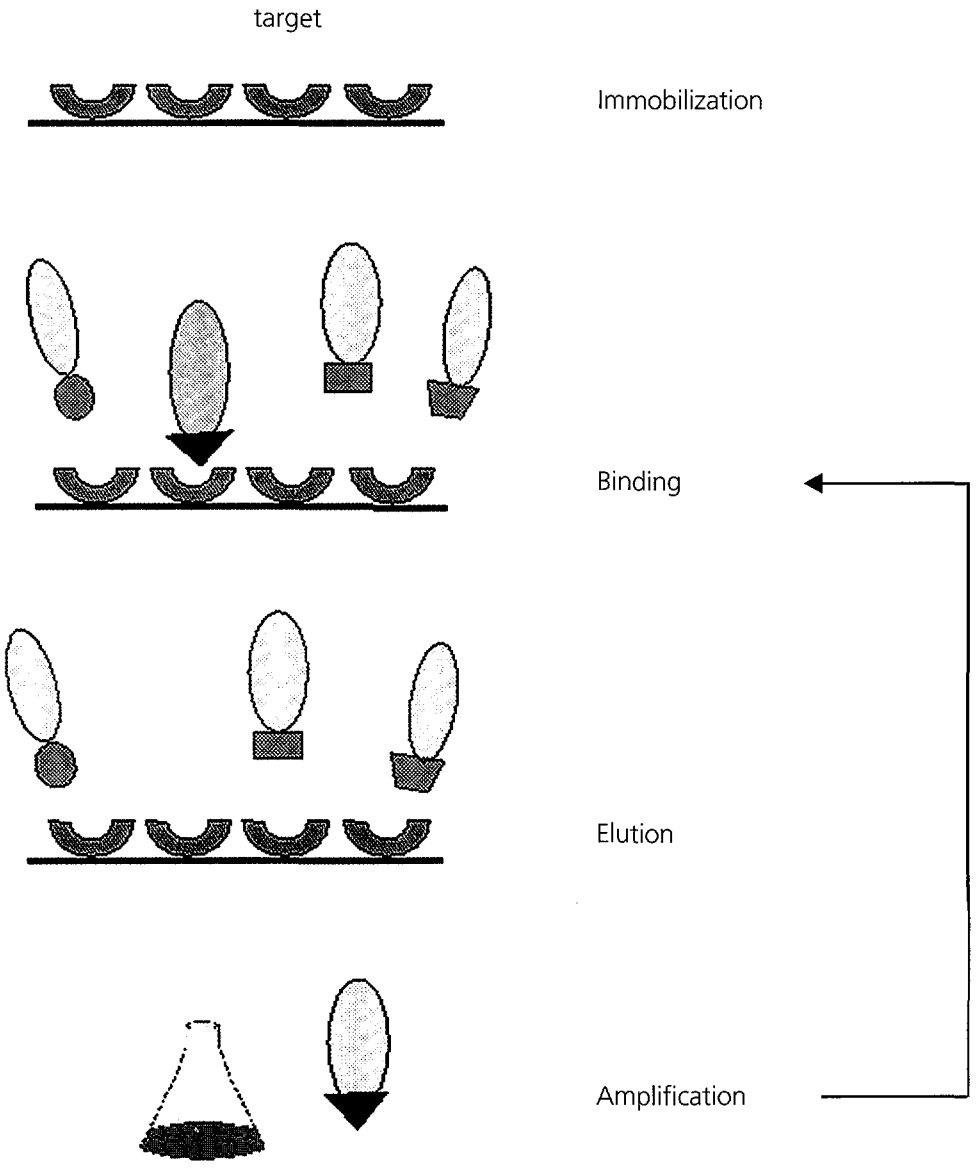
10

Phage selection for biopanning

G. Palombo

10 Phage selection for biopanning

Giovanna Palombo



- Select target binding peptides from libraries of phage-displaying peptides
- Enrich phages with the active sequence by repetitive cycles
- Modulate stringency of selection
- Sequence active phages for sequence determination

10.1. Introduction

Repertoires of peptides and antibodies have been displayed on the surface of filamentous bacteriophages [1-3] or expressed on the surface of bacteria [4]. The great power of biological libraries lies in the possibility of making use of biopanning [5-7]. Biopanning represents an *in vitro* selection method that has been employed successfully to isolate target-binding sequences. Since its introduction [8], such a technique has become increasingly important in combinatorial methodologies, finding a wide range of applications in both research and diagnostic fields. In its simpler form, biopanning is carried out by incubating a library of phage-displayed peptides with a support coated with the target, washing away the unbound phages and eluting the retained phages. The desorbed particles are then propagated in *E. coli* and taken through additional cycles of biopanning and amplification to successively enrich the pool of phages in favour of the tightest binding sequences (see figure on page 10.1). After two or three selection rounds, individual clones are picked up and characterized either by DNA sequencing and/or ELISA assay.

Attempts to optimize the selection process have resulted in the setting up of new biopanning procedures that can be selectively applied, depending on the chosen selection strategy. Such methods differ principally in the kind of support chosen for the coating procedure. The choice of support depends on the chemical and physical properties of the target in question.

The most frequently used *in vitro* selection method is performed on plates, usually microtiter plates or Petri dishes, coated directly with the target (using the direct coating method [9-13]). One of the most common artefacts occurring with this biopanning procedure is the selection of matrix-binding clones that are difficult to remove even through extensive washing. To avoid such aspecific interactions, it is often necessary to increase the stringency of the selection. Generally, non-stringent conditions are used early in a selection process to maximize the yield of bound phages even if selection of them will be less rigorous. In subsequent rounds, higher stringency of selection can be reached in several ways by:

- Increasing the number and length of the washing steps [14] as well as the number of biopanning cycles [15];
- Changing the ionic strength of the washing buffer to remove weakly bound phages;
- Decreasing the protein coating concentration to select for higher affinity ligand [16];
- Recovering the bound phages through a competitive elution rather than by an acid elution [17, 18].

In addition, aspecific phage interactions can be avoided by introducing subtractive selection cycles, in which the phage pool is allowed first to react with the plate alone (no target), then the non-adherent phages are added to the target-coated plate [9, 12].

The biopanning on plate, as described, requires the availability of a purified target and its consequent immobilization on the solid support. In some instances, however, the coating procedure leads to the partial destruction of the 3D structure of the protein. Such a consideration is more critical if the target in question is an antibody or an antigen, since the loss of its native structure may allow the isolation of sequences against other irrelevant epitopes. Those problems can be avoided by using an alternative biopanning selection, based on the strong biotin-streptavidin interaction to isolate target-binding molecules. The library is reacted with the target in solution, previously biotinylated, then the phage-target complexes are captured on a streptavidin-coated dish, recovered by elution, amplified and submitted to a successive round of biopanning [5, 19-21]. A similar selec-

tion process utilizes magnetic particles as support for isolating active-binding molecules. The steps, including the binding of the library to a biotinylated target and the capture on the streptavidin-coated beads, are identical to those described above but the recovery of bound phages is performed using a magnetic field [22].

A different strategy of selection is performed on columns prepared by covalently immobilizing the target protein on a chromatographic support [23, 24]. The phage library is then applied to the column, equilibrated with buffer at a defined pH and ionic strength. The nature of the target-phage interaction is not yet clear; presumably the phage is thin enough to partially penetrate the pore where it became "caged" in the immediate vicinity of the target. After extensive washing steps (required to remove aspecifically bound molecules), adsorbed phages are dissociated by low pH buffer or by competitive elution using a high concentration of the same target or of the real ligand, if available.

Successful biopanning experiments have also been performed using intact cells [25, 26], as well as whole tissues [27] as targets for phage selection. More recently, there has been an exciting development in the use of such methodology. An *in vivo* selection method has been carried out by injecting a phage population into living animals and analysing phages rescued from specific organs in order to identify new peptide motifs responsible for organ-specificity [28].

Among the large number of selection procedures mentioned above, the present chapter analyses the biopanning on plate and explains in detail each step required to perform such a technique.

10.2. Materials and methods

10.2.1. Materials

- ❑ target (enough to perform three panning cycles: 80 μg-800 μg)
- ❑ phage library (~2 x 10^{12} pfu/ml)
- ❑ *E. coli* strain cells suitable for the library
- ❑ Petri dish (65 mm x 15 mm) or 96 well microtiter plates
- ❑ 0.1 M sodium bicarbonate pH 8.5
- ❑ PBS stock solution:
  50 mM phosphate (pH 7.5), 150 mM NaCl; autoclave and store at 4° C
- ❑ PBS-T:
  PBS containing 0.5% Tween-20; adjust the Tween concentration for desired stringency
- ❑ blocking buffer:
  PBS containing 5% non-fat dry milk or 5% BSA or 1% gelatin; use immediately
- ❑ 20% Tween-20 stock solution; autoclave and store at rt
- ❑ 0.2 M glycine-HCl, pH 2.2
- ❑ 1 M Tris-HCl, pH 9.0
- ❑ 10% Triton X-100 stock solution; filter sterilize and store at rt
- ❑ PEG/NaCl
  20% (w/v) polyethylene glycol-8000, 2.5 M NaCl; autoclave and store at rt

10.2.2. Methods

1. Dissolve the target at a final concentration of 100 μg/ml[70,71] in a 0.1 M sodium bicarbonate solution, pH 8.5 and dispose 1.5 ml in a sterile 60 mm x 15 mm Petri dish.[72]

2. Place the dish in a humidified container (to prevent evaporation of the solution) and incubate overnight at 4° C or at least 2 h at rt.[73]

3. Pour off the coating solution and firmly place the plate face down on to a clean paper towel. Fill the plate up to the top with blocking solution[74] and incubate at least 1 h at 4° C.

4. Discard the blocking solution and wash the dish with PBS buffer six times, emptying it completely after each wash.

5. Add to the plate about 2 x 10^{11} virions of a phage library diluted in 1 ml of PBS[75,76,77] and incubate for 45 min at rt.[78,79]

---

[70] If a pure target is not available, use a monoclonal or polyclonal antibody-coated dish to capture the target from a crude extract [29].

[71] The coating concentration of the target must be determined empirically and depends on both the nature of the target and on the desired affinity of the selected peptide. In general, concentrations between 1 μg/ml and 100 μg/ml are used but those values can be changed during the same selection process. In particular, a decrease of the target concentration in the last cycles of selection may allow the isolation of higher binding affinity peptides.

[72] Alternatively, use a 96 well microtiter plate and add 150 μl of target solution to each well.

[73] The coating is a critical step in the biopanning selection. In some instances, such a procedure may lead to the partial destruction of the 3D structure of the target protein. An associated risk is that sequences are selected against useless epitopes. More frequently, the immobilization of the protein fails because of the chemical and physical properties either of the target or of the support utilized. Generally, polystyrene plates are indicated for protein immobilization, while polyvinyl chloride dishes are recommended for smaller target coating, such as peptides or small antigens. However, before starting the selection, it is strongly recommended that an ELISA assay to verify the immobilization of the target on the chosen support be performed. Such an assay could be carried out by adding an antibody solution directed against the target to the protein-coated dish and detecting the target-antibody complex through the use of a secondary, labelled antibody.

[74] BSA, dried milk and gelatin are commonly used as blocking reagents to saturate the uncoated plastic surface. The more suitable for the selection may be determined empirically and depends on the nature of both the target and the library.

[75] Add a 100 μl aliquot if a 96 microtiter plate is used. If the library is not concentrated enough, a precipitation step is required. Add 1/6 fraction volume of PEG/NaCl to the library solution and incubate at least 3 h at 4° C. Centrifuge 15 min at 10,000 rpm, pour off the supernatant and re-suspend the library in a smaller volume.

[76] Aspecific, hydrophobic interactions may occur between capsid phage proteins and the target during the panning. Such an interaction can be reduced if Triton X-100 is added to the phage library to a final concentration of 1%.

[77] As a negative control, incubate the same amount of phage library in a well or Petri dish saturated with the blocking reagent (no target) and carry out a parallel selection. The analysis of phages recovered from the two different selections will distinguish, at the end of the process, aspecifically bound molecules from those selectively binding the target.

[78] The conditions for binding (i.e. buffer, incubation time and temperature) determine the stringency of the selection and may be modified during the panning procedure. Generally, non-stringent conditions (high-binding temperature and long incubation time) are used early in the first selection cycles to ensure that even molecules with low-binding activity are not going to be lost. In subsequent rounds, the stringency may be increased to select for higher-affinity peptide ligands.

[79] One of the most frequent artefacts that occurs during a biopanning procedure is the selection of molecules binding aspecifically either the blocking reagent, the plastic surface or other undesirable contaminants. In order to minimize such a risk, alternative selection methods can be performed, such as: (a) subtractive biopanning [9, 12] in which the library is first incubated in a dish, coated only with the blocking solution. Non-adherent phages are removed and added to the target-coated dish. Such a procedure is useful for removing molecules from the entire population that bind the blocking reagent and select for those binding specifically the target; (b) competitive biopanning [10, 13], in which the library is sorted against an immobilized target in the presence of a competitor in solution (usually a protein with similar chemical and physical properties as the target). In such a process, a phage particle having a higher affinity toward the competitor will bind to it, resulting in no more being available for the interaction with the target. Such an approach, if successful, allows the selection of phages bearing epitopes with high selectivity and specificity to the target of interest.

6. Pour off the unbound phages and wash 10 times with PBS-T, placing firmly face down on to the plate each time.[80]

7. Elute the bound phage by incubating the plate for 10 min (maximum) in a 0.2 M glycine-HCl, pH 2.2 solution while gently rocking. Neutralize it immediately with 1 M Tris-HCl, pH 9.0.[81]

8. Titer 1 μl of the eluate and amplify the remaining by adding it to a 20 ml of a 1:100 dilution of *E. coli* culture that has grown overnight[82] and incubate 4.5 h at 37° C, while shaking at 250 rpm.[83]

9. Spin the culture 10 min at 10,000 rpm at 4° C. Transfer the supernatant and re-spin to pellet residual cells.

10. Recover phages from the supernatant by adding 1/6 volume of PEG/NaCl and incubating overnight at 4° C (or for at least 3 h).

11. Spin PEG precipitation 20 min at 10,000 rpm in a refrigerated centrifuge, pour off the supernatant, re-spin and remove any remaining supernatant.

12. Thoroughly suspend the pellet in 1 ml of TBS, spin out residual cells and re-precipitate phage with PEG/NaCl solution on ice for 60 min. Pellet phages by centrifugation for 15 min at 4° C at 10,000 rpm.

13. Finally, re-suspend the pellet which contains the amplified library in 200 μl-300 μl of PBS containing 0.02% NaN_3.

14. Titer the amplified phage[84] and store it at 4° C.[85]

15. Carry out the successive round of biopanning, repeating all the steps described above and using 2 x 10^{11} phages of the amplified pool as input phage.[86]

16. Titer the eluate from the last panning (it is not necessary to amplify it) and plate several-fold dilutions in order to obtain isolated plaques. Store the remaining eluate at 4° C.

17. Pick up well-isolated plaques[87] using sterile Pasteur pipettes or pipet tips and transfer them in separate tubes containing 2 ml of a 1:100 dilution of an overnight culture of *E. coli* cells.

---

[80] The presence of detergent as well as a higher ionic strength of the washing buffer reduces aspecific interaction between the phage and the target and/or the blocking reagent. Lower Tween concentration in early rounds will result in higher eluate titer and the stringency can be gradually increased during each round by raising the Tween concentration up to a maximum of 0.5%.

[81] Alternatively, a competitive elution can be performed by incubating the plate with a solution of known ligand (if available), or free target (concentration range 10 μg/ml-100 mg/ml) in order to take the bound phage away from the immobilized target on the plate. Such an elution method is more specific and is recommended to increase the stringency of the selection.

[82] If the library has been built in M13 coliphage, the *E. coli* strain must carry the F episome, which is responsible for the M13 infection. F^+ *E. coli* strain such as ER2537, XL1-Blu and DH5aF are available commercially. However, all commercial libraries are supplied with the appropriate *E. coli* strain which has been tested for that phage display system.

[83] Carry on the amplification in a well-aerated environment (use a 250 ml flask for 20 ml of culture) and for no more than 5 h at 37° C to avoid undesirable deletions.

[84] If no plaques are visible after 16 h, check that the dilution range used to plate was appropriate ($1:10^8$-$1:10^{11}$ for amplified phage, $1:10^1$-$1:10^4$ for early rounds, $1:10^4$-$1:10^7$ for later rounds). If the phage suspension is not diluted sufficiently, the resulting plaques will be confluent so that the plate will look as if there are no plaques at all. If the phage titer is lower than that expected, check that the amplification conditions are those recommended above (see footnote 82 above). Furthermore, if an acid elution has been performed, check that the eluted phage solution has been neutralized before the amplification.

[85] If the selection is not continued at this point, the phage preparation can be stored up to 4 weeks at 4° C. However, it is recommended that the successive biopaning cycles be performed as soon as possible, since some phages may be unstable and could be lost if stored for longer periods.

[86] By maintaining a constant input of phages in each round of biopanning (input phage), a stepwise increase of the number of eluted phages (output phage) should be observed. Depending on the interaction under examination and on the stringency with which the selection has been carried out, usually 2 or 3 rounds are enough to reach a considerable amount of active binding sequences against the target of interest.

[87] It is very important that plates have no more than 100 plaques. This will ensure that a single plaque is picked and, successively, to analyse a single DNA sequence. Usually 10 to 20 clones are sufficient to detect a consensus binding sequence.

18. Carry out the amplification procedure as described before.

19. After the centrifugation step, transfer the supernatant of each clone in a fresh tube and store at 4° C.[88]

20. Analyse each clone for sequence consensus[89] and for binding properties.[90]

[88] This is the amplified phage stock and can be stored at 4° C for several weeks. For long-term storage, however, diluting it 1:1 with sterile glycerol and storing it at −20° C is recommended.

[89] After the last round of panning, the pool of eluted phages should be enriched significantly with selected binding peptides that share a similar or identical amino acid sequence. If a significant consensus can be identified, the screening may have been successful. One or more peptides are chosen, chemically synthesized and tested for their ability to bind the target outside the context of a phage.

[90] An ELISA assay can be performed by incubating the phage clones on a target-coated dish and using an anti-M13 antibody HRP-conjugate to detect the phage-target complex. Such a detection system presents the main limitation to allow only the detection of higher affinity ligands because of the very low phage concentration than can be added to the coated-target well (10^{12} virions/100 μl well correspond to a phage concentration of 16 nM). Alternatively, a sandwich ELISA can be carried out by immobilizing the phage clones on plate and adding an excess of target in solution. The availability of an antibody against the target will allow the detection of positive clones. Sometimes, however, there is no correspondence between consensus sequence and ELISA signal. The only definite evidence of sequence-binding affinity will result from binding experiments carried out with the peptide deduced from the screening, once it has been synthesized.

References

1. Smith, G. P. 1991. "Surface presentation of protein epitopes using bacteriophage expression system". *Curr. Opin. Biotech.*, vol. 2, pp. 668-673.
2. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. and Winter, G. 1991. "Making antibodies fragment using phage display library", *Nature*, vol. 352, pp. 624-628.
3. McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J. 1990. "Phage antibodies: filamentous phage displaying antibody variable domains", *Nature*, vol. 348, pp. 552-554.
4. Lu, Z., Murray, K. S., Van Cleave, V., La Vallie, E. R., Stahl, M. L. and McCoy, J. M. 1995. "Expression of thioredoxin random peptide libraries on the *Echerichia coli* cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interaction", *Biotechnology*, vol. 13, pp. 366-372.
5. Scott, J. K. and Smith, G. P. 1990. "Searching for peptide ligands with an epitope library", *Science*, vol. 249, pp. 386-390.
6. Parmley, S. F. and Smith, G. P. 1988. "Antibody-selectable filamentous fd phage vectors: affinity purification of target genes", *Gene*, vol. 73(2), pp. 305-318.
7. Hoogenboom, H. R. 1997. "Designing and optimizing library selection strategies for generating high-affinity antibodies", *Trends in Biotech.*, vol. 15, pp. 62-70.
8. Smith, G. P. 1985. "Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface", *Science*, vol. 228, pp. 1315-1317. Smith, G. P. 1991. "Surface presentation of protein epitopes using bacteriophage expression system", *Curr. Opin. Biotech.*, vol. 2, pp. 668-673.
9. Koivunen, E., Gay, D.A. and Rouslahti, E. 1993. "Selection of peptides binding to the $\alpha_5\beta_1$ integrin from phage display library", *J. Biol. Chem.*, vol. 268(27), pp. 20205-20210.
10. Smith, J. W., Hu, D., Satterthwait, A., Pinz-Sweeney, S. and Barbas III, C. F. 1994. "Building synthetic antibodies as adhesive ligands for integrins", *J. Biol. Chem.*, vol. 269(52), pp. 32788-32795.
11. Sparks, A. B., Quilliam, L. A., Thorn, J. M., Der, C. J. and Kay, B. K. 1994. "Identification and characterization of Src SH3 ligands from phage-displayed random peptide libraries", *J. Biol. Chem.*, vol. 269(39), pp. 23853-23856.
12. Sato, A., Ida, N., Fukuyama, M., Miwa, K., Kazami, J. and Nakamura, H. 1996. "Identification from a phage display library of peptides that bind to toxic shock syndrome toxin-1 that inhibits its binding to major histocompatibility complex (MHC) class II molecules", *Biochemistry*, vol. 35, pp. 10441-10447.
13. Ames, R. S., Tornetta, M. A., Jones, C. S. and Tsui, P. 1994. "Isolation of neutralizing anti-C5a monoclonal antibodies from a filamentous phage monovalent Fab display library", *J. Immunol.*, vol. 152, pp. 4572-4581.
14. Daniels, D. A. and Lane, D. P. 1994. "The characterization of p53 binding phage isolated from phage peptide display libraries", *J. Mol. Biol.*, vol. 243, pp. 639-652.

15. Baca, M., Presta, L. G., O'Connors, S. J. and Wells, J. A. 1997. "Antibody hymanization using phage display", *J. Biol. Chem.*, vol. 272(16), pp. 10678-10684.

16. Fu, Y., Shearing, L. N., Haynes, S., Crewther, P., Tilley, L., Anders, R. F. and Foley, M. 1997. "Isolation from phage display libraries of single chain variable fragment antibodies that recognize conformational epitopes in the malaria vaccine candidate, apical membrane antigen-1", *J. Biol. Chem.*, vol. 272(41), pp. 25678-25684.

17. Choi, S. J., Ahn, M., Lee, J. S. and Jung, W. J. 1997. "Selection of high affinity angiogenin-binding peptide from a peptide library displayed on phage coat protein", *Mol. Cells*, vol. 7 (5), pp. 575-581.

18. Healy, J. M., Murayama, O., Maeda, T., Yoshino, K., Sekiguchi, K. and Kikuchi, M. 1995. "Peptide ligands for integrin $\alpha_v\beta_3$ selected from random phage display libraries", *Biochemistry*, vol. 34, pp. 3948-3955.

19. Balass, M., Heldman, Y., Cabilly, S., Givol, D., Katchalski-Katzir, E. and Fuchs, S. 1993. "Identification of a hexapeptide that mimics a conformation-dependent binding site of acetylcholine receptor by use of a phage-epitope library", *Proc. Natl. Acad. Sci.*, vol. 90, pp. 10638-10642.

20. Barchan, D., Balass, M., Souroujon, M. C., Katchalski-Katzir, E. and Fuchs, S. 1995. "Identification of epitopes within a highly immunogenic region of acetylcholine receptor by phage epitope library", *J. Immunol.*, vol. 155, pp. 4264-4269.

21. Scott, J. K., Loganathan, D., Easley, R. B., Gong, X. and Goldstein, I. J. 1992. "A family of concavalin A-binding peptides from a hexapeptide epitope library", *Proc. Natl. Acad. Sci.*, vol. 89, pp. 5398-5402.

22. Hawkins, R. E., Russell, S. J. and Winter, G. 1992. "Selection of phage antibodies by binding: affinity mimicking affinity maturation", *J. Mol. Biol.*, vol. 226, pp. 889-896.

23. Wang, B., Dickinson, L. A., Koivunen, E., Rouslahti, E. and Kohwi-Shigematsu, T. 1995. "A novel matrix attachment region DNA binding motif identified using a random phage peptide library", *J. Biol. Chem.*, vol. 270(40), pp. 23239-23242.

24. DeLeo, F. R., Yu, L., Burritt, J. B., Loetterle, L. R., Bond, C. W., Jesaitis, A. J. and Quinn, M. T. 1995. "Mapping sites of interaction of p47-phox and flavocytochrome b with random-sequence peptide phage display libraries", *Proc. Natl. Acad. Sci.*, vol. 92, pp. 7110-7114.

25. Marks, J. D., Ouwehand, W. H., Bye, J. M., Finnern, R., Gorick, B. D., Voak, D., Thorpe, S. J., Hughes-Jones, N. C. and Winter, G. 1993. "Human antibody fragments specific for human blood group antigens from a phage display library", *Biotechnol.*, vol. 11, pp. 1145-1149.

26. Watters, J. M., Telleman, P. and Junghans, R. P. "An optimized method for cell-based phage display panning", *Immunotechnol.*, vol. 3(1), pp. 21-29.

27. van Ewijk, W., de Kruif, J., Germeraad, W. T. V., Berendes, P. , Röpke, C., Platenburg, P. P. and Logtenberg, T. 1997. "Subtractive isolation of phage-displayed single-chain antibodies to thymic stromal cells by using intact thymic fragments", *Proc. Natl. Acad. Sci.*, vol. 94, pp. 3903-3908.

28. Pasqualini, R. and Rouslahti, E. 1996. "Organ targeting *in vivo* using phage display peptide libraries", *Nature*, vol. 380, pp. 364-366.

29. Sanna, P. P. , Williamson, R. A. and De Logu, A. 1995. "Directed selection of recombinant human monoclonal antibodies to herpes simplex virus glycoproteins from phage display libraries", *Proc. Natl. Acad. Sci.*, vol. 92, pp. 6439-6443.

Printed in Austria
V.99-85600—September 2000—1,000

UNIDO publication
Sales No. E.00.II.B.7
ISBN 92-1-106411-2

ID/ICS.4