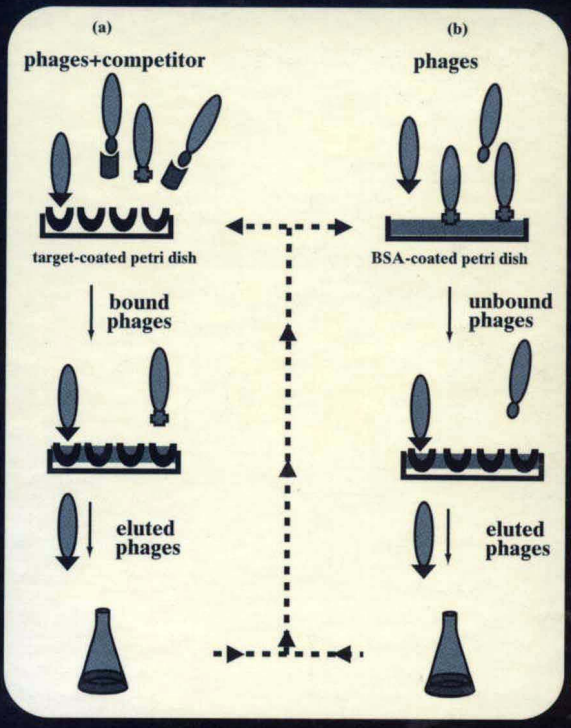
COMBINATORIAL CHEMISTRY AND TECHNOLOGY

Principles, Methods, and Applications



edited by
STANISLAV MIERTUS
GIORGIO FASSINA

COMBINATORIAL CHEMISTRY AND TECHNOLOGY

Principles, Methods, and Applications

edited by

STANISLAV MILER TUS
International Centre for Science and High Technology of UNIDO
Trieste, Italy

GIORGIO FASSINA
TECNOGEN SCpA
Piana di Monte Verna (CE), Italy

MARCEL DEKKER

MARCEL DEKKER, INC. NEW YORK · BASEL

Library of Congress Cataloging-in-Publication Data

Combinatorial chemistry and technology: principles, methods, and applications/edited by Stanislav Miertus, Giorgio Fassina.
p. cm.
Includes index.
ISBN: 0-8247-1960-3 (alk. paper)
1. Combinatorial chemistry. I. Miertus, Stanislav. II. Fassina, Giorgio
RS419.C59 1999 99-15461
615'.19—dc21 CIP
This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

Eastern Hemisphere Distribution
Marcel Dekker AG
Hutgasse 4, Postfach 812, CH-4001 Basel, Switzerland
tel: 41-61-261-8482; fax: 41-61-261-8896

World Wide Web
http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

Preface

Combinatorial methodologies have dramatically changed the drug discovery process in the pharmaceutical industry, offering an unlimited source of new molecular entities to be screened for activity. Innovative chemistries, software, hardware, and advanced molecular biology protocols have been developed in the past few years to generate numerically complex and structurally diverse libraries of compounds, as well as new automated approaches for simultaneous screening. This field not only has integrated, complemented, and revitalized existing research activities, it also has tremendously stimulated the evolution of other disciplines and technologies, such as solid-phase chemistry, molecular modeling, in vitro screening of biological activities, miniaturization, and automation, to satisfy the demand to design, produce in high yield and purity, and rapidly screen huge numbers of molecules. The rapidly growing interest in combinatorial technologies all over the world has been due not only to the possibility of identifying new drugs to treat human diseases, but also to the broad applicability to other fields, such as the diagnostic, new materials, and catalysis sectors. Many books have been published on combinatorial techniques, but most offer only a limited perspective of the field, focusing on selected aspects without covering all the different approaches and integrated technologies involved.

This book provides comprehensive coverage of the current methodologies employed for the design, synthesis, and screening of molecular libraries. Major topics include generation of molecular libraries by chemical methods using solution and solid-phase chemistries; biological approaches for the production and screening of peptide, antibody, and oligonucleotide libraries; the application of computer-assisted approaches to guide library synthesis; the use of high-throughput screening methodologies to accelerate lead discovery; development of auto-

mation and robotics; and economic and legal issues. The book is intended for scientists, technologists, and doctoral fellows who require an introduction to the combinatorial world. All the basic approaches and methodologies are reviewed in sufficient detail to provide university-level teachers with a useful textbook for a course in combinatorial technologies.

Stanislav Miertus
Giorgio Fassina

Contributors

Valery V. Antonenko High Throughput Parallel Organic Synthesis, Affymax Research Institute, Santa Clara, California

Andrew Bradbury Department of Biophysics, International School for Advanced Studies, Trieste, Italy

Lucia Carrano Microbial Technologies Department, Biosearch Italia SpA, Gerenzano, Italy

Maria Dani TECNOGEN SCpA, Piana di Monte Verna (CE), Italy

Elisabetta de Magistris Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Verona, Italy

Stefano Donadio Microbial Technologies Department, Biosearch Italia SpA, Gerenzano, Italy

Giorgio Fassina Department of Biopharmaceuticals, TECNOGEN SCpA, Piana di Monte Verna (CE), Italy

Eduard R. Felder Core Technology Area, Novartis Pharma AG, Basel, Switzerland

Rafael Ferritto Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Verona, Italy

Valerie J. Gillet Department of Information Studies, University of Sheffield, Sheffield, United Kingdom

Larry Gold NeXstar Pharmaceuticals, Boulder, Colorado

Hennie R. Hoogenboom CESAME at Department of Pathology, University Hospital Maastricht and Target Quest B.V., Maastricht, The Netherlands

Simon Hufton CESAME at Department of Pathology, University Hospital Maastricht and Target Quest B.V., Maastricht, The Netherlands

Olga Issakova Department of Chemistry, Selectide, a subsidiary of Hoechst Marion Roussel, Tucson, Arizona

Stanislav Miertus International Centre for Science and High Technology of UNIDO, Trieste, Italy

Andrea Missio Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Verona, Italy

Alfredo Paio Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Verona, Italy

Giovanna Palombo Department of Biopharmaceuticals, TECNOGEN SCpA, Piana di Monte Verna (CE), Italy

Aris Persidis Business Development, Argonex, Inc., Charlottesville, Virginia

Menotti Ruvo Department of Biopharmaceuticals, TECNOGEN SCpA, Piana di Monte Verna (CE), Italy

Daniele Sblattero Department of Biophysics, International School for Advanced Studies, Trieste, Italy

Pierfausto Seneci Lead Discovery Department, GlaxoWellcome Medicines Research Centre, Verona, Italy

Nikolai Sepetov Nanoscale Combinatorial Synthesis, Inc., Tucson, Arizona

Introduction of the ICS-UNIDO Program in Combinatorial Chemistry and Technology

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 and partnerships, and promote socially equitable and environmentally friendly industrial development.

UNIDO is the only worldwide organization dealing exclusively with industry from a development perspective. Its services are nonprofit, 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. The agency'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 (ICS) is an institution within the legal framework of UNIDO with headquarters in Trieste, Italy. The Centre's mandate relates to the transfer of know-how and technology in favor of developing countries and is justified by the perception that a competitive industrial technological capability cannot be built without adequate scientific knowledge and commitment to a sustainable development approach utilizing new and environmentally friendly technologies.

The activities of the ICS follow an integrated pragmatic approach that includes action-oriented research, short-term exchange between researchers and technologists in industry, dissemination of scientific and technological information through the creation and management of centers of excellence (focal points),

consultancy and advisory services, training courses, scientific workshops, high-level seminars, study tours, fellowships, promotion of training arrangements, and publication and editing of frontier issues.

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

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

Catalysis, an important scientific and technological area for the development of environmentally friendly chemical processes, which in turn form the basis for cleaner industrial development and are also the key elements for an industrial prevention approach. New, less polluting 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 with a direct impact on the quality and quantity of by-products or waste generated.

Environmentally degradable plastics, an area in which the expanding global production and consumption of polymeric materials coupled with increasing public awareness of environmental issues have created serious concern about the problems related to the disposal of plastic waste generated by various sectors of human activity. Besides recycling, reuse, incineration, and composting, new technological developments of environmental degradable plastics contribute dramatically to the tackling of the environmental issue in specific sectors of plastics use.

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

Combinatorial chemistry and combinatorial technologies, which have a strong impact on the development of new chemicals (pharma industries, agrochemicals, new materials). Developing countries need to become acquainted with and gain expertise in combinatorial technologies to help local enterprises remain competitive and economically viable in the coming decades. Combinatorial chemistry and combinatorial technologies

have a potential influence not only on industrial growth but also on environment protection. In fact, by optimizing industrial processes and production, with the lowering of relevant costs, smaller amounts of waste and by-products are created.

Combinatorial chemistry and combinatorial technologies fall 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 just a few years than the pharmaceutical industry did in its entire previous history.

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

Because of this important role of combinatorial chemistry and combinatorial technologies, ICS-UNIDO has developed a program in this field that includes the following.

- A training course, "Methodologies, Applications and Economics of Combinatorial Chemistry and Combinatorial Technologies," held in Piana di Monte Verna, Italy, on September 8–19, 1997, hosted and co-organized by TECNOGEN.
- A workshop, "Applications of Molecular Design and Computer-Assisted Combinatorial Chemistry," held in Cape Town, South Africa, March 29–April 4, 1998, hosted and co-organized by the University of Cape Town, Department of Chemistry
- A workshop, "Combinatorial Technologies—Awareness and Familiarization for Decision Makers" and a training workshop "Methods, Applications and Economics of Combinatorial Chemistry and Combinatorial Technology" held in Hyderabad, India October 23–31, 1998, hosted and co-organized by the Council of Scientific and Industrial Research (CSIR), Indian Institute of Chemical Technology (IICT).
- A workshop "Combinatorial Chemistry and Combinatorial Technologies" held in Buenos Aires, Argentina, December 7–11, 1998, hosted and co-organized by INGEBI, University of Quilmes.
- A workshop "Combinatorial Chemistry and Combinatorial Technologies" held in Laguna, Philippines, April 19–23, 1999, hosted and co-organized by the University of the Philippines Los Banos.

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

When implementing the program, the lack of suitable literature introducing and overviewing the various aspects of combinatorial chemistry and combinatorial technologies (CC/CT) was denounced. It was therefore decided to prepare, beside the ICS-UNIDO publications covering various aspects of specific issues related to the implementation of CC/CT in industries, a general book on CC/CT overviewing methods and applications. This project was then promoted by Marcel Dekker, Inc., with the participation of more than 20 experts working in various branches of CC and CT.

As indicated more specifically in the Preface, this book should appeal to a broad spectrum of readers from both industrialized and developing countries.

I would like to thank all the authors who contributed to this book, for their expertise and high-quality collaboration. My thanks are due to Dr. G. Fassina, ICS Scientific Consultant in this field, for his collaboration in the coordination of the subprogram on Combinatorial Chemistry and Combinatorial Technologies.

I am grateful to Marcel Dekker, Inc., for their collaborative effort, and especially to Ms. Anita Lekhwani and Mr. Joseph Stubenrauch for their assistance.

The assistance and secretarial support of Ms. Paola Volpi and Ms. Marina Varvesi in the implementation of the ICS program on combinatorial chemistry and combinatorial technology are highly appreciated.

Stanislav Miertus

Contents

1

Combinatorial Chemistry and Combinatorial Technologies: Principles and Applications

Giorgio Fassina
TECNOGEN SCpA, Piana di Monte Verna (CE), Italy

Stanislav Miertus
International Centre for Science and High Technology of UNIDO, Trieste, Italy

The time and cost needed for the development of new drugs have increased steadily during the past three decades. Estimated costs for introducing a new drug in the market now reach around $200–300 million U.S., and this process takes around 10–12 years after discovery. This increase in time and cost is due mainly to the extensive clinical studies of new chemical entities required by competent regulatory agencies, such as the U.S. Food and Drug Administration (FDA), and to a lesser extent to the increased costs associated to research. The time and cost required for clinical and preclinical evaluation of new drugs is not likely to decrease in the near future, and as a consequence, a key issue for pharmaceutical companies to stay in the market has been to increase the number of new drugs in the development pipeline. Drug discovery in the past has been based traditionally on the random screening of collections of chemically synthesized compounds or extracts derived from natural sources, such as microorganisms, bacteria, fungi, and plants, of terrestrial or marine origin, or by modifications of chemicals with known physiological activities. This approach has resulted in many important drugs, but the ratio of novel to previously discovered compounds has diminished with time. In addition, this process is very time consuming and expensive. A limiting factor was 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 synthe-

sis of new chemical entities often is a very laborious task, and additional time is required for purification and chemical characterization. The average cost of creating a new molecular entity in a pharmaceutical company is around $7500 U.S./compound [1]. Generation of natural extracts, while very often providing interesting new molecular structures endowed with biological properties, leads to mixtures of different compounds at different concentrations, thus making activity comparisons very difficult. In addition, once activity is found in a specific assay, the extract needs to be fractionated in order to identify the active component. Quite often, the chemical synthesis of natural compounds is extremely difficult, thus making the lead development into a new drug a very complex task. While the pharmaceutical industry was demanding more rapid and cost-effective approaches to lead discovery, the advent of new methodologies in molecular biology, biochemistry, and genetics, leading to the identification and production of an ever-increasing number of enzymes, proteins, and receptors involved in biological processes of pharmacological relevance, and good candidates for the development of screening assays, complicated this scenario even more. The introduction of combinatorial technologies provided an unlimited source of new compounds, capable of satisfying all these needs. This approach was so appealing and full of promise that many small companies started to flourish, financed by capital raised from private investors. Once combinatorial technologies clearly demonstrated the potential to identify new leads with a previously unknown speed, the majority of these companies were purchased by big pharmaceutical companies. Combinatorial approaches were originally based on the premise that the probability of finding a molecule in a random screening process is proportional to the number of molecules subjected to the screening process. In its earliest expression, the primary objective of combinatorial chemistry focused on the simultaneous generation of large numbers of molecules and on the simultaneous screening of their activity. Following this approach, the success rate of identifying new leads is greatly enhanced, while the time required is 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 raised enormous excitement about the possibility of finding new and valuable drugs in short times and at reasonable costs. However, the advent of this new field in drug discovery did not obscure the importance of "classical" medicinal chemistry approaches, such as computer-aided rational drug design and QSAR, for example, but instead catalyzed their evolution to complement and integrate with combinatorial technologies.

The word "combinatorial" appeared in the scientific literature at the beginning of the 1990s, but the generation of the first combinatorial libraries can be dated back to the beginning of the 1980s. The first reports dealt with the simul-

taneous production of collections of chemically synthesized peptides, produced by solid-phase methods on solid supports [2–6]. Peptides were particularly suited for combinatorial synthesis given the well-established synthetic protocols available, the great number of different molecules attainable, and the potential to generate leads of biological and pharmaceutical value. The use of peptide libraries was greatly accelerated by the introduction of biological methods for library preparation, and by the use of phage display technology, which provided interesting advantages over the synthetic counterpart [7,8]. At the same time, the first papers on the generation of oligonucleotide libraries appeared in the literature [9,10], thus suggesting the possibility 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 glycoconjugate drugs whose carbohydrate constituents are associated with the molecular mechanism by which these drugs exhibit their effects. With these drugs, exploration of carbohydrate molecular diversity has the potential for identifying novel agents with enhanced potency. As a conformationally rigid and functionally rich system, carbohydrates also provide valuable molecular scaffold systems around which to generate primary screening libraries.

A broad variety of new synthesis and screening methods are currently grouped under the term "combinatorial." These methods include parallel chemical synthesis and testing of multiple individual compounds or 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. All these different methods have expanded rapidly, each with its putative advantages, disadvantages, and proponents, and a broad coverage of all the diverse approaches used for library generation and screening is provided in the following chapters.

Many active compounds have been selected to date following combinatorial methodologies, and a considerable number of those have progressed to clinical trials. However, combinatorial chemistry (CC) and related technologies for producing and screening large number of molecules also find useful applications in other industrial sectors not necessarily related to pharmaceutical industry. Emerging fields of application of combinatorial technologies are the diagnostic, the downstream, processing, the catalysis, and the new-material sectors. In the first case, CC can be successfully applied to the identification of previously unknown epitopes recognized by antibodies in biological fluids associated with pathological conditions. The selected epitopes can then be used for 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 to recognize specific macromolecules of biotechnological interest, such as proteins, antibodies, or nucleic acids. This is of great importance to industry, since the major costs associated with

the production of recombinant molecules for therapy are associated with the purification of the desired target molecule from crude feedstocks. The availability of specific and selective ligands, such as monoclonal antibodies, to be used in affinity chromatography for the capture and concentration of the target from crude samples, will reduce considerably the costs of producing biopharmaceuticals [11]. Combinatorial technologies have also been applied to the identification of new macromolecules endowed with catalytic activity for reactions where natural enzymes are inactive. This application, even if 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.

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

With the increased speed at which new drug entities are now synthesized and evaluated for pharmacological activity, a need has arisen to provide fundamental metabolism data at the early stages of drug discovery. Strategies are being developed to permit drug metabolism data to be an important part of early drug discovery. Many important properties of drugs related to metabolism can be the deciding factor in whether or not a compound is selected for clinical development. Some of these include the pharmacokinetic properties. Other related factors that can help discovery teams make decisions about which structures to pursue include measurements of metabolic stability, protein binding, P450 inhibition and absorbtion.

In combinatorial chemistry, due to the high number of chemical manipulations required to synthesize libraries of compounds, automation is unavoidable. Many research groups, in both academia and industrial settings, are developing automated instruments tailored specifically to these needs, and this technology field is acquiring an extremely important role for the development of combinatorial technologies for the next millennium.