LEAD GENERATION APPROACHES IN DRUG DISCOVERY

Zoran Rankovic
Schering-Plough Corporation
Newhouse, Lanarkshire, UK

Richard Morphy
Schering-Plough Corporation
Newhouse, Lanarkshire, UK

A JOHN WILEY & SONS, INC., PUBLICATION

LEAD GENERATION APPROACHES IN DRUG DISCOVERY

LEAD GENERATION APPROACHES IN DRUG DISCOVERY

Zoran Rankovic
Schering-Plough Corporation
Newhouse, Lanarkshire, UK

Richard Morphy
Schering-Plough Corporation
Newhouse, Lanarkshire, UK

A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Library of Congress Cataloging-in-Publication Data:

Lead generation approaches in drug discovery / [edited by] Zoran Rankovic, Richard Morphy.
p. ; cm.
Includes bibliographical references and index.
ISBN 978-0-470-25761-6 (cloth)
1. Drug development. 2. Drugs—Design. 3. High throughput screening (Drug development) I. Rankovic, Zoran. II. Morphy, Richard.
[DNLM: 1. Drug Discovery—methods. 2. Lead. QV 744 L4345 2010]
RM301.25.L433 2010
615′.19—dc22

2009040188

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

Over the past decade or so, many, if not most, pharmaceutical companies have introduced a distinctive lead generation phase to their drug discovery programs. The critical importance of generating high quality starting compounds for lead optimization cannot be overstated. The decisions made during lead generation, in terms of which compounds to progress, fundamentally influence the chances of a new chemical entity progressing successfully to the market. One of the principal architects of lead generation as a separate phase within the pharmaceutical industry was Bill Michne while at AstraZeneca. In Chapter 1, he presents an overview of the lead discovery process describing its challenges and the achievements of medicinal chemists working in this area.

The main source of hits continues to be high throughput screening (HTS). In Chapter 2, the evolution of the HTS approach over the last two decades is described by Zoran Rankovic, Craig Jamieson, and Richard Morphy from Schering-Plough Corporation. Today there is greater emphasis than ever before on the quality of compound collections and the robustness and fidelity of HTS assays and equipment. Screening is followed by a hit validation phase in which the desired and undesired attributes of hits are investigated, centred on potency, selectivity, physicochemical and ADMET properties, ease of synthesis, novelty, and SAR. Sufficiently attractive hits are then selected for hit-to-lead optimization in which the goal is to build confidence in a chemical series to lower the risk that the commitment of additional resources in lead optimization will be in vain.

In Chapter 3, Dagmar Stumpfe, Hanna Geppert, and Jürgen Bajorath outline how knowledge-based *in silico* or "virtual" screening plays an increasingly important role at this early stage of the drug discovery process. Virtual screening methods can be neatly grouped into protein structure-based and ligand-based approaches, the former being used for targets for which 3D structural information is available, such as kinases and proteases, and the latter more extensively employed for membrane-bound targets such as GPCRs. Ligand-based virtual screening hinges on the concept of molecular similarity, what descriptors are used to represent molecules, and how the similarity between those representations of different molecules is assessed. HTS and *in silico* screening are complementary techniques; following HTS, readily available analog of hits can be retrieved using similarity searching thereby generating early SAR to guide the hit-to-lead process.

The ever-increasing emphasis on physicochemical properties in recent years has spawned the concept of ligand efficiency and an increased interest in

fragment-based drug discovery (FBDD). In Chapter 4, Jeffrey Albert describes the different methods of conducting FBDD, NMR spectroscopy and crystallography in particular, and some of the successes that have been achieved. The first compounds derived from FBDD are now advancing through the clinic, proving the validity of the approach, at least for enzyme targets.

Most current drug discovery projects aim to discover a drug that is highly selective for a single target. However, one of the main causes of attrition in Phase 2 proof-of-concept studies is poor efficacy. The multitarget drug discovery (MTDD) area is attracting increasing attention amongst drug discoverers, since drugs that modulate multiple targets simultaneously have the potential to provide enhanced efficacy. The challenges and opportunities presented to medicinal chemists by MTDD are described by Richard Morphy and Zoran Rankovic in Chapter 5.

Most approaches to lead generation rely upon screening of some sort, be it random HTS, FBDD, or a knowledge-based virtual screen. However, in Chapter 6, Gisbert Schneider, Markus Hartenfeller, and Ewgenij Proschak provide an alternative perspective, describing how lead compounds can be designed de novo. Algorithms that allow either structure-based or ligand-based assembly of new molecules are described. In the area of de novo design, historically one of the main challenges has been to ensure the synthetic feasibility of the designed compounds. Fortunately, new approaches to estimate the synthetic accessibility of de novo designed molecules are becoming available.

Natural products have been an important source of drugs, morphine and taxol being two prominent examples that are both ancient and modern, respectively. During the rush toward high throughput methods of lead generation in the past decade, less emphasis was placed on using the natural world as a source of starting compounds and inspiration for drug discovery. However, recently a resurgence of interest in natural products is apparent within the drug discovery community. Hugo Lachance, Stefan Wetzel, and Herbert Waldmann remind us in Chapter 7 of their historical track record and the potential of natural product-like compounds for exploring regions of biologically-relevant chemical space that are not accessible using the medicinal chemist's more usual repertoire of synthetic compounds.

In the early days of the lead generation phase, there was often an overemphasis on increasing potency during the hit-to-lead phase while neglecting physicochemical property and pharmacokinetic profiles. Today, multiple parameters are optimized in parallel to produce leads with a balanced profile of biological, physicochemical, and pharmacokinetic properties. In Chapter 8, Dennis Smith of Pfizer outlines how early screening of hits and leads for metabolic, pharmacokinetic, and toxicological liabilities can reduce attrition during the later phases of drug discovery.

No matter whether hits are derived from random screening or knowledge-based approaches, medicinal chemists will naturally assess and prioritize hits on the basis of their synthetic feasibility. Those that are amenable to analog

synthesis in a high throughput parallel fashion will be favored. The skills for making libraries of compounds are now well developed amongst lead generation chemists, and the historical development of this area is described by Roland Dolle and Karin Worm in Chapter 9.

One approach to lead generation that is not covered explicitly in this book, but which is as relevant today as it ever was, is starting from an existing drug. As Nobel Laureate James Black once remarked, "The most fruitful basis for the discovery of a new drug is to start with an old drug.". If there are no suitable drugs with which to start, preclinical molecules that are described in the primary literature or in a patent can also be considered. The advantage of starting with an existing agent is that much knowledge concerning its behavior in *in vitro* and *in vivo* assays, and perhaps also in the clinic, is already in the public domain. This can facilitate rapid progress toward a second-generation "best-in-class" molecule that has an improved profile, which could be, for example, improved pharmacokinetics, allowing an easier dosing schedule for patients or an improved safety profile by virtue of being more selective for the target of interest.

It is hoped that this book succinctly captures the essence of lead generation toward the end of the first decade of the twenty-first century and, looking further ahead, this critical phase of the drug discovery process is unlikely to diminish in its relevance and importance, and thus many new developments can be eagerly anticipated.

CONTRIBUTORS

JEFFREY S. ALBERT, AstraZeneca Pharmaceuticals, Wilmington, Delaware, USA

JÜRGEN BAJORATH, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

ROLAND E. DOLLE, Adolor Corporation, Exton, Pennsylvania, USA

HANNA GEPPERT, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

MARKUS HARTENFELLER, Eidgenössische Technische Hochschule (ETH), Institute of Pharmaceutical Sciences, Wolfgang-Pauli-Str. 10, 8093 Zürich, Switzerland.

CRAIG JAMIESON, Schering-Plough Corporation, Newhouse, Lanarkshire, UK

HUGO LACHANCE, Max-Planck Institute for Molecular Physiology, Dortmund, Germany

WILLIAM F. MICHNE, MedOrion, Inc., Oriental, North Carolina, USA

RICHARD MORPHY, Schering-Plough Corporation, Newhouse, Lanarkshire, UK

EWGENIJ PROSCHAK, Johann Wolfgang Goethe-University, Frankfurt am Main, Germany

ZORAN RANKOVIC, Schering-Plough Corporation, Newhouse, Lanarkshire, UK

GISBERT SCHNEIDER, Eidgenössische Technische Hochschule (ETH), Institute of Pharmaceutical Sciences, Wolfgang-Pauli-Str. 10, 8093 Zürich, Switzerland.

DENNIS A. SMITH, Pfizer Global R&D, Sandwich, Kent, UK

DAGMAR STUMPFE, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany

HERBERT WALDMANN, Max-Planck Institute for Molecular Physiology, Dortmund, Germany

STEFAN WETZEL, Max-Planck Institute for Molecular Physiology, Dortmund, Germany

KARIN WORM, Adolor Corporation, Exton, Pennsylvania, USA

1 Lead Discovery: The Process

WILLIAM F. MICHNE
MedOrion, Inc., 6027 Dolphin Road, Oriental, NC 28571, USA.
billmichne@embarqmail.com

CONTENTS

1.1. INTRODUCTION

The discovery of a lead compound is arguably the first significant chemistry milestone along the path to the discovery of a new small molecule drug. Without a solid lead a new drug discovery project must be put on hold or abandoned. The process by which this milestone is achieved has been and is one of continuous evolution. Indeed, even the definition of what constitutes a lead has been refined over the past two decades such that lead criteria are much more stringent than ever before. This is a necessary consequence of high

Lead Generation Approaches in Drug Discovery,
Edited by Zoran Rankovic and Richard Morphy.

attrition rates in the clinic, the escalation of drug development costs beyond the billion dollar mark, and the ever higher expectations of safety and efficacy on the part of regulatory agencies, physicians, and patient populations at large. It is incumbent upon medicinal chemists to find the strongest possible leads so as to minimize the risk of failure as the compounds move along the increasingly expensive path to the pharmacists' shelves. This chapter is intended to provide an introduction to the volume by presenting a historical overview of lead discovery, along with an assessment of where we stand today and the major barriers that still remain. The coverage cannot be exhaustive, considering the volume of literature on the subject. Rather, it will be selective, focusing on key contributions that form the body of basic principles in current use, and providing indications of directions for further research likely to improve the efficiency and cost effectiveness of lead discovery.

1.2. HISTORICAL OVERVIEW

The use of exogenous substances to treat human diseases can be traced to very ancient times. The use of opium, for example, was described by the Sumerians several thousand years ago. In 1805 Serturner isolated from opium a single pure substance that he called morphine. This allowed the study of the effects of morphine without the interfering effects of other constituents of the opium mixture, and marks the beginning of the science of pharmacology. There followed rapidly the isolation of additional alkaloids from opium, and their comparative biological evaluations were the beginnings of modern medicinal chemistry, notwithstanding the fact that the structures of these substances would not be known for another century or more. Biological studies were purely descriptive. Compounds were administered to whole animals and effects were observed. If a compound's effect was considered desirable as it may have related to human disease, the compound might have been evaluated for toxicity, and then tried in humans. In the early twentieth century the theory of drug effect resulting through specific interaction with a receptor was advanced by Paul Ehrlich. His basic concepts of specificity, reversibility, and development of tachyphylaxis are still largely valid today. In combination with the development of molecular structure and its elucidation, this gave rise to the concept of structure–activity relationships (SAR) as a means of improving drug action. Wohler's synthesis of urea in 1828 established the science of synthetic organic chemistry, which allowed chemists to explore non-natural analogs of natural product structures for biological activity. Thus, the process of discovering new drugs relied almost entirely on the discovery of analogs of natural products, prepared one at a time, with biological activity in whole animal systems, through about two-thirds of the twentieth century. About that time, the pace of discovery and understanding in the field of molecular biology began to accelerate. We quickly learned that genes were expressed as proteins, and that some of these proteins, either because they were overexpressed, or

incorrectly expressed because of genetic mutations, were largely responsible for many human diseases. We also learned that these proteins were in many cases the targets of drugs shown to be effective in treating the disease. We learned how to manipulate genetic sequences, how to transfect modified sequences into cells that would then produce the modified protein, how to isolate those proteins, and how to assay for protein function in the presence of a potential drug. Advances in robotics allowed us to conduct tens of thousands of assays in a matter of weeks or days, considered a trivial exercise by today's standards. By the early 1990s, high throughput screening (HTS) emerged as the ultimate solution to drug discovery, if only we could provide the numbers of compounds to stoke the robots. A company's legacy compound collection became its screening library. These collections, typically in the range of 50–100,000 compounds, reflected the past chemical interests of the company, and as a result, were of a low order of chemical diversity, consisting rather of a relatively small number of groups of similar compounds. This quickly gave rise to compound acquisition strategies. New positions and new departments with new budgets were created with the purpose of buying compounds from whomever was willing to sell them, such as university professors and small chemical companies that had closed their doors for whatever reason. Eastern European chemical operations became a rich source of compounds with the collapse of the Soviet Union. Still, screening libraries grew slowly during the early 1990s, while screening robots and assay development moved forward at a similarly slow pace. The situation is captured in this quote from the first paper published on the hit-to-lead (HTL) process [1]:

> Consider the fairly typical situation of a screening program consisting of sixteen assays and a library of 150,000 compounds to be completed in one year. If each compound is screened at a single concentration, then 2.4 million data points will be generated. If the hit rate is only 0.01%, then 20 active compounds per month can be expected.

At that time the question of how to sort out which compounds to work on was a fairly daunting task.

1.3. THE HIT-TO-LEAD PROCESS

Several companies worked on the problem, but Sterling-Winthrop was the first to publish a method [1]. They began by defining *hit-to-lead chemistry* as "a process wherein the likelihood of failure or long term success for a given chemical series can be evaluated in a systematic way." They further defined an *active* as "any substance that shows reproducible activity in any screen." They defined a *hit* as having the following attributes: activity in a relevant bioassay is reproducible; structure and high purity has been confirmed; selectivity data has been generated; available analogs have been tested; the structure is chemically

tractable; and there is potential for novelty. Finally, a *lead* was defined as a hit with genuine SAR, which suggests that more potent and selective compounds can be found. The process that was developed had five basic objectives that are still valid today. The first is the validation of structure and purity. Many of the screening samples were decades old, prepared and purified prior to the widespread use of high field NMR spectroscopy for structure determination and chromatographic techniques for purity assessment. Further, many samples simply decomposed during long-term storage. It was found that 10–30 percent of the samples in some collections either had incorrect structures or were less than 90 percent pure. A second objective is to find the minimum active fragment (pharmacophore) in a complex active molecule, as such information impacts the strategy for SAR development of the series. A third objective is to enhance selectivity in order to minimize activity at closely related molecular targets. A fourth objective is to exclude compounds with inappropriate modes of action, such as nonspecific binding to the molecular target. The fifth objective is to increase potency. Actives often have activity in the micromolar range. Such compounds can be problematic later in development with respect to achieving effective concentrations in the target tissue or organ.

The overall objective of HTL chemistry is to find the best possible leads that can be progressed to lead optimization (LO) in the shortest period of time. Project leaders must be ruthlessly Darwinian in their willingness to drop projects for which one or more of the above objectives cannot be achieved in about six months. Beyond that, time-precious resources are increasingly expended on projects with decreasing probabilities of success. The effectiveness of an HTL chemistry group is determined solely by the rate and quality of leads generated; the timely termination of hit series for cause must be considered to play a vital role in that effectiveness.

In the ensuing years since the formulation of the HTL objectives, most if not all lead discovery organizations established an HTL function. Beyond the five fundamental objectives most had differences in degree, rather than kind, as to the exact process to be followed. Nevertheless, essentially every drug discovery organization in the world had incorporated, by the late 1990s, an HTL function, by whatever name they chose, into their overall drug discovery pathway. One such typical pathway is illustrated in Figure 1.1.

In general, such pathways captured the key activities and timelines for progression of compounds, but often lacked specific criteria to be met. As experience was gained, and successes and failures were evaluated, new attributes of the process were developed. Many organizations became more rigid in their milestone criteria for hits and leads. Perhaps one of the most complete yet succinct descriptions of criteria in use was published by Steele and coworkers [2] at AstraZeneca. While these criteria may not be universally applicable because of the nature of a particular molecular or disease target, they were developed as a result of their internal experience with the HTS–HTL–LO history across many disease targets and therapeutic areas. Hit criteria are aimed at providing post-HTS evidence that the activity is not associated with

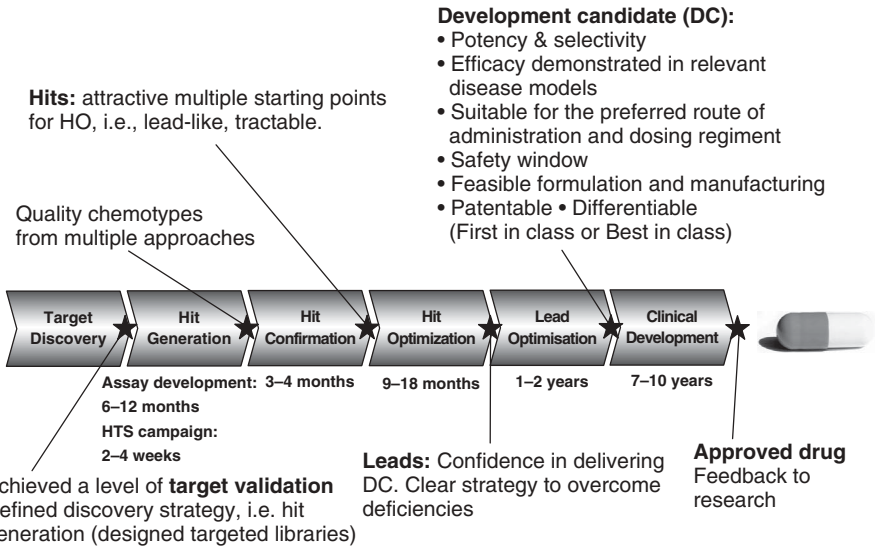


Figure 1.1. The drug discovery pathway from the target to the clinic. The figure was kindly provided by Dr. Zoran Rankovic.

impurities or false positives, and that the issues to be addressed in HTL have at least been evaluated. Lead criteria build on the hit criteria to demonstrate sufficient scope and robustness to sustain a major LO project. Both sets of criteria are subdivided into three categories as they relate to the compound series; pharmacology; and absorption, distribution, metabolism, and excretion (ADME), including *in vivo* data for leads. The authors admit that this is a bewildering array of information to capture during a supposedly fast and low resource plan. However, they state that, ''this set of criteria has grown from lessons learned by the unwitting progression of avoidable but time-consuming issues into LO, leading to more attrition in a phase where it is less acceptable.'' Subsequent publications from this company have included an abbreviated version of the criteria to illustrate the HTL process (*vide infra*).

1.4. THE MAGNITUDE OF THE PROBLEM

Over the last two decades there have been tremendous advances both in synthetic organic chemistry and in screening technology. The former, through methods in combinatorial chemistry, and more recently in multiple parallel syntheses, has increased the available numbers and diversity of screening libraries, and the latter has increased the rate at which assays can be run. As these technologies advanced, libraries of billions of compounds, and the robots to screen them, were envisaged. A logical extension of these technologies would

be to synthesize and test all possible drug-like molecules against all possible disease targets. This turns out to be utterly and completely impossible. Defining as drug-like molecules that consist of no more than 30 nonhydrogen atoms, limited to C, N, O, P, S, F, Cl, and Br, with molecular weight ≤ 500 Da, and making an allowance for those unstable to moisture and oxygen at room temperature, the total number of possible combinations has been estimated to be of the order of 3×10^{62} [3]. Assuming an average molecular weight of 350 Da, if one were to prepare only a single molecule of each of these possible structures and place them in a sphere at a density of 1 g/cm^3, the diameter of the sphere would be approximately 40 billion miles. Note, however, that the best small molecule drug for every disease is contained within this sphere of chemistry space. Note also that we can never know whether any particular drug is the best possible one for its use.

1.5. SCREENING LIBRARIES

Early thinking in the design of screening libraries was that biological activity was to be found uniformly in diversity space. Accordingly, vast combinatorial libraries, based solely on available synthetic methodologies, and the desire to be as broadly diverse as possible, were prepared and screened against a broad range of molecular targets. Hit rates turned out to be disappointingly low, and attempts to progress many of these hits to leads were unsuccessful. It rapidly became clear that within drug chemistry space biological activity occurred only in certain regions, which in all likelihood occupied only a small portion of the total space. Thus, the random approach to building productive screening libraries had only a correspondingly small probability of success. This gave rise to new thinking about how to describe biologically relevant chemistry space, and how then to design libraries that expand and further explore that space. Many approaches were taken, but all are, for the most part, variations on one of the following three themes: physical properties, chemical structures, and biological behavior. Most of the effort has been concentrated on the first two of these, as structure and properties enjoy a high level of intellectual comfort among chemists. The relationship of biological behavior to library design is less intuitive, but may ultimately prove to be the most effective approach.

Through about the early to mid-1990s medicinal chemistry strategy for drug discovery was to make compounds of similar structure to some lead compound with the hope of finding increasingly potent compounds. The most potent compounds were then advanced. When a compound failed in development it was replaced with a related potent compound and the process repeated. The relationship between physical properties and successful drug development was not fully appreciated until publication of the "rule of five" by Lipinski and coworkers [4]. The importance of this paper is apparent from its well over 1500 citations. The rule states that poor absorption or permeability are more

likely when the calculated 1-octanol/water partition coefficient is >5, the molecular mass is >500 Da, the number of hydrogen bond donors (OH plus NH count) is >5, and the number of hydrogen bond acceptors (O plus N atoms) is >10. Somewhat later the importance of the number of rotatable bonds [5] and polar surface area [6] to oral bioavailability was established. These six easily calculated physical properties have become important considerations in the design of screening libraries, as well as HTL and lead optimization programs. Just how important these properties are was assessed by Wenlock and coworkers [7]. This work compared profiles of compounds in various stages of development (Phase I, II, III, Preregistration, and Marketed oral drugs), and included data on compounds that were discontinued from the first three phases in order to evaluate properties trends. The properties studied did not include polar surface area, but did include $\log D_{7.4}$, a measure of the lipid water distribution of ionizable compounds at pH 7.4. The data was reported as the mean value and standard deviation of each property for the set of compounds in each phase, both active and discontinued. The mean molecular weights for drugs in all phases of development are higher than that of marketed oral drugs. The mean $\log P$ of compounds in the active phases are very similar to that of marketed oral drugs, whereas the values for discontinued compounds are approximately 0.6 units higher. The situation is exactly the same for $\log D_{7.4}$. For H-bond donors, only the Phase I group was significantly different at the 95 percent level from the marketed drugs. For H-bond acceptors, four groups differed from the mean value for marketed drugs. For rotatable bonds, five groups differed significantly from marketed drugs. One of the clearest findings of this work is that mean molecular weight in development decreases on passing through each of the phases. The same appears to be true for $\log P$ and $\log D_{7.4}$. Thus, molecular weight and lipophilicity are two properties that show the clearest influence on the successful passage of a candidate drug through the different stages of development.

Leeson and Springthorpe [8] extended this work to evaluate trends in physical properties of launched drugs from 1983 to 2006. The number of drugs trended steadily downward, while the molecular weights trended steadily upward. Indeed, over this time period, there are increases in three of the four rule-of-five properties, with lipophilicity changing less appreciably. Polar surface area and rotatable bond count also increased. The authors suggest that factors driving the increases in physical properties include a greater number of apparently less druggable new targets, for which larger and more lipophilic molecules appear to be necessary for high affinity binding to active sites. But the fact that drug lipophilicity is changing less over time than other physical properties suggests that this is an especially important drug-like property. This property essentially reflects the key event of molecular desolvation in transfer from aqueous phases to cell membranes and to protein binding sites, which are mostly hydrophobic in nature. The authors' thesis is that if lipophilicity is too high, there is an increased likelihood of binding to multiple targets. They used the Cerep BioPrint database of drugs and reference compounds to examine the

role of physical properties in influencing drug promiscuity and side effects. They found that promiscuity is predominantly controlled by lipophilicity and ionization state. Thus, they concluded that the most important challenge today is delivering candidate drugs that will not eventually fail owing to properties-based toxicity. The implications of working increasingly closer to the extremities of drug-like chemical space appear serious for overall productivity and promiscuity, leading to increased risks of pharmacologically based toxicity.

The assessment and characterization of the structures of all known biologically active molecules would be a daunting task. A simpler and arguably still representative approach would be to study the structural attributes of known drugs. Bemis and Murcko [9] used some 5000 compounds from the Comprehensive Medicinal Chemistry database. They analyzed only the graph representation of the molecule, that is, the representation of the molecule by considering each atom to be a vertex and each bond to be an edge on a graph, without regard to atom type or hybridization. Ring systems were defined as cycles within the graph representation and cycles sharing an edge. Linker atoms are on the direct path connecting two ring systems. A framework was defined as the union of ring systems and linkers. Interestingly, only 6 percent of the molecules were acyclic, and only 32 frameworks were found to occur in 20 or more compounds, accounting for 50 percent of the total compounds. By separating rings from linkers it was found that only 14 rings and 8 linkers (chains of 0–7 atoms) were represented. While it is tempting to speculate about reasons for finding so few frameworks among so many drugs, the authors stated that the results do not necessarily reveal some fundamental truths about drugs, receptors, metabolism, or toxicity. Instead, it may reflect the constraints imposed by the scientists who have produced the drugs. Synthetic and patent considerations, and a tendency to make new compounds that are similar to known compounds, may all be reflected in these findings.

In a second publication by these authors [10] a similar analysis was carried out for the side chains, defined as nonring, nonlinker atoms. They found that 73 percent of side-chain occurrences were accounted for by only 20 side chains. This suggests that the diversity that side chains provide to drug molecules is quite low, and may explain the poor performance of combinatorial screening libraries in which diversity elements are added as side chains. The authors suggest that one way their findings could be used is to generate a large virtual library by attaching side chains to frameworks at various points. The virtual library could then be computationally filtered to optimize a particular property of interest prior to the library actually being synthesized. They have illustrated this application by designing a library optimized for blood brain barrier penetration [11].

Most companies began to analyze in detail the composition and HTS hit rates of their legacy libraries, combinatorial libraries, and natural products libraries. Jacoby and coworkers published a detailed analysis of their experience at Novartis [12]. They found that the validated hit rate of the combinatorial collection was 0.02 percent. The hit rate of the historic medicinal

chemistry archive was higher at 0.03 percent, but both were outperformed by the natural products collection that had a hit rate of 0.12 percent. However, when the percentage of assays with hits from each compound source is considered, a very different picture emerges. The combinatorial and natural products collections were successful in about 60 percent of the assays, whereas the historic archive collection always produced hits. They suggested that the higher performance of the latter collection is due in part to its consisting of individually synthesized compounds with a resulting overall higher scaffold diversity, and in part to the compounds having been driven toward a historic target focus, and therefore have more similarity to active compounds of similar classical druggable targets. Using this and additional analyses based on cheminformatics, as well as the molecular framework approach of Bemis and Murcko (*vide supra*), these authors developed a compound collection enhancement program that also included the extraction of privileged scaffolds from HTS data.

The literature soon saw a proliferation of computational methods to extract the best compounds and the most information from screening libraries. Many of these methods, while interesting in their own right, were unavailable to the broad scientific community because of issues of hardware, software support, and the proprietary nature of the method. The most valuable of these methods, then, were the empirical, practical, hands-on strategies. In these cases, chemists developed sets of rules to define the acceptability of compounds. This approach is not without its pitfalls. For example, Lajiness and coworkers [13] conducted a study in which 13 experienced medicinal chemists reviewed a total of about 22,000 compounds with regard to their acceptability for purchase for use in screening. The compounds had already passed a number of standard compound filters designed to eliminate chemically and/or pharmaceutically undesirable compounds. The compounds were divided into 11 lists of about 2000 compounds each. Most of the chemists reviewed two lists. Unknown to the chemists, each list contained a subset of 250 compounds previously rejected by a very experienced senior medicinal chemist. The results indicated that there is very little consistency among medicinal chemists in deciding which compounds to reject. It was also shown that when medicinal chemists review the same compounds a second time but embedded within a different 2000 compound set, they reject the same compounds only about 50 percent of the time. The authors were careful to point out that the results make no judgments of right or wrong, but only that opinions often differ. There is probably better agreement among chemists when considering substructures and functional groups. Steele and coworkers [2] at AstraZeneca developed a set of such chemistry filters for eliminating undesirable compounds. Their resultant rules are quite simple. The first rule states that compounds must have at least one N or O atom and at least one noncyclic bond. The second rule states that compounds must not have any of the substructures or functional groups listed in a table of chemistry filters with 46 entries. In contrast to the above discussion of a low level of consistency among individual chemists considering specific compounds, these filters were derived from a consensus of opinions of a

panel of experienced medicinal chemists considering only substructures and functional groups. Similar tables have certainly been developed at other companies, and it would be interesting to compare them if they were to become available. The AstraZeneca group went on to partition their collection into the Good Set (largely lead-like compounds passing all chemical filters, molecular weight < 400, $\log P < 5.0$); the Bad Set (pass all chemical filters, molecular weight < 550, $\log P < 7.0$); and the Ugly Set (molecular weight > 550, or $\log P > 7.0$, or fail any chemical filter). The Ugly Set was excluded from HTS.

1.6. MOLECULAR COMPLEXITY

Molecular complexity is probably impossible to define precisely, but like art, most chemists know it when they see it. Such molecular features as multiple chiral centers, polycyclic structures, and multiple functionality, symmetry, heteroatoms, etc., all contribute to one's perception of complexity. Notwithstanding our inability to define it, there have been numerous attempts to quantitate it [14–16]. All of these attempts have been structured from an organic synthesis perspective. For example, using these methods the complexity index of each intermediate of a synthesis can be calculated and the change in complexity plotted as a function of step number in the synthesis. In this way, several syntheses of the same molecule can be compared with respect to progression of complexity as the syntheses proceed. Of course, a complex molecule may require a complex synthesis today, but new developments in synthetic methodology may result in a dramatically simplified synthesis. Does this mean the molecule is no longer considered complex? From a biological perspective, however, the notion of molecular complexity has a rather different meaning. The extent to which two molecules bind, or not, is dependent on the complementarity of their various surface features, such as charge, dipole effects, hydrogen binding, local hydrophobicity, etc. If complementarity is thought of in terms of complexity, it seems intuitive that the more complex a molecule is, the less likely it is to bind to multiple protein binding sites. In other words, the more complex the molecule, the more likely it is to be selective for a single, or small number of protein targets. This notion appears to have been first expressed by Bentley and Hardy in 1967 [17]. They were trying to discover analgesics superior to morphine (**1**, Fig. 1.2) by separating desirable from undesirable effects. All efforts to that point had been in the direction of making structures simpler than morphine, such as meperidine (**2**). They reasoned that compounds of simpler structure would be able to fit all of the receptor surfaces associated with the different effects. These considerations led them to examine bases more complex and more rigid than morphine, in the hope that the reduced flexibility and the differences in peripheral shape between such compounds and other known analgesics would result in the new bases being unacceptable at some of the receptor surfaces, and thus to a separation of the various effects. Among the molecules that have emerged from this work is

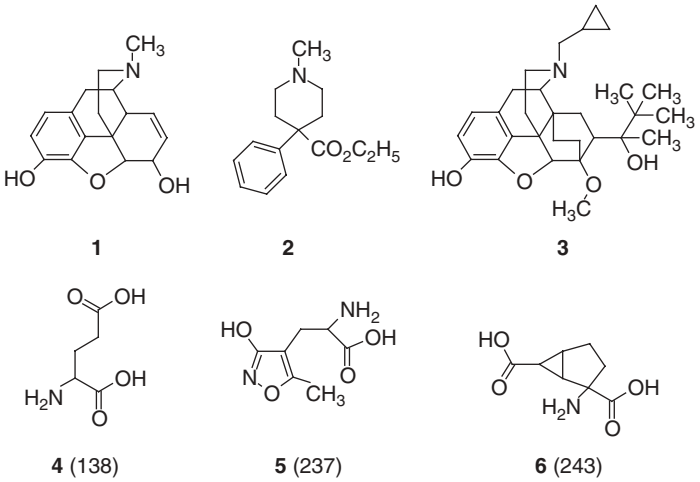


Figure 1.2. The structures of the opiates morphine (**1**), the simpler analog meperidine (**2**), and the more complex buprenorphine (**3**). The glutamate ligands glutamic acid (**4**), (S)-AMPA (**5**), and Ly-354740 (**6**) are shown with calculated Barone–Chanon complexity indices in parentheses.

buprenorphine (**3**), clearly more complex than morphine by the addition of a bridged bicylic system, a hydroxyl group, two new chiral centers, and peripheral hydrophobicity. A strict comparison of the promiscuities of **2** and **3** apparently has not been done. However, using the Cerep BioPrint database (*vide supra*) it was found that meperidine produced 24 hits, and buprenorphine produced only 8 hits, defined as >30 percent inhibition at 30 μM.[1] It thus appears that higher orders of complexity are consistent with the Bentley and Hardy hypothesis. A second example using structures more closely related to each other is from the glutamate field. For each structure, the Barone–Chanon complexity index [16] has been calculated and is shown in Figure 1.2.

Glutamic acid itself (**4**) has an index of 138, and is obviously a ligand for all 20-plus glutamate receptors. (S)-AMPA (**5**) has an index of 237 and is highly selective for the ionotropic glutamate receptors, and LY-354740 (**6**), with an index of 243, is highly selective for a subset of metabotropic glutamate receptors. Again, there does seem to be a relationship between molecular complexity and receptor selectivity.[2] This suggests that for an LO program, increasing complexity while holding the values of other properties within drug-like boundaries might be a viable strategy. But what about screening libraries? The above observations suggest that libraries of very complex molecules might produce relatively few hits because the compounds are too complex to fit the

[1] We thank Dr. Paul D. Leeson for providing this information.
[2] Michne, W. F., unpublished data.

binding sites on the protein targets. Indeed, Tan et al. [18] assembled a library of over two million compounds based on a very complex natural product core. Screened in three assays, the library produced no hits in two of them, and only one hit in the third, with an EC_{50} of 50 μM. Hann and coworkers [19] developed a very simple model to represent probabilistic aspects of molecular recognition. The model does not incorporate flexibility in either the ligand or the binding site. With this type of model it is possible to explore how the complexity of a ligand affects its chance of matching a binding site of given complexity. They found that the probability of finding any type of match decays exponentially as the complexity of the ligand increases, and that the chance of finding a ligand that matches only one way, that is, a unique binding mode, in a random choice of ligand peaks at a very low ligand complexity. The results further suggested that libraries containing very complex molecules have a low chance of individual molecules binding. In their view it is better to start with less complex molecules, and to increase potency by increasing complexity. They suggest that there is an optimal complexity of molecules that should be considered when screening collections of molecules, but did not speculate as to what that complexity should be, or how to measure it. Schuffenhauer and coworkers [20] used historical IC_{50}/EC_{50} summary data of 160 assays run at Novartis covering a diverse range of targets, and compared this to the background of inactive compounds. As complexity measures they used the number of structural features present in various molecular fingerprints and descriptors. They found that with increasing activity of the ligands, their average complexity also increased. They could therefore establish a minimum number of structural features in each descriptor needed for biological activity. They did not, however, address the issue of complexity versus selectivity.

1.7. SIMILARITY

Generations of medicinal chemists have developed SAR by synthesizing molecules similar to an active molecule. This is generally considered successful, although any experienced chemist will agree that the approach is full of surprises. When considering compounds for purchase to augment an existing library, it is important to consider whether structurally similar molecules have similar biological activity. This question was specifically addressed by Martin and coworkers [21]. They defined similar as a compound with ≥ 0.85 Daylight Tanimoto similarity to an active compound. They showed that for IC_{50} values determined as a follow-up to 115 HTS assays, there is only a 30 percent chance that a compound similar to an active is itself active.

They also considered whether biologically similar compounds have similar structures. They compared the nicotinic agonists acetylcholine (**7**) and nicotine (**8**), and the dopaminergic agonists dopamine (**9**) and pergolide (**10**) (Fig. 1.3). They found that the highest Daylight Tanimoto similarity within this group of four compounds (0.32) was between nicotine and pergolide, and the second

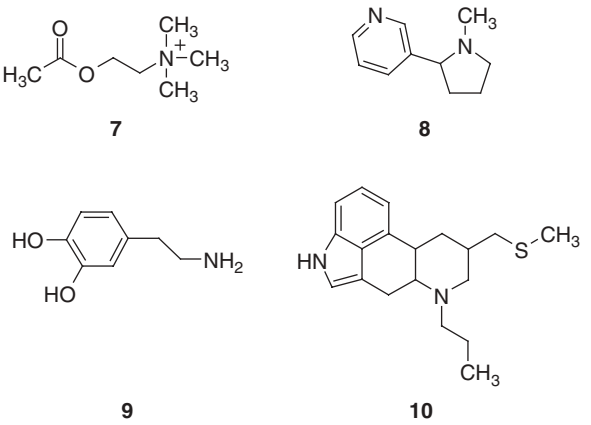


Figure 1.3. Biologically similar molecule pairs: acetylcholine (**7**) and nicotine (**8**), and dopamine (**9**) and pergolide (**10**).

highest (0.22) was between nicotine and dopamine. These results were likely in part due to deficiencies in the similarity calculations, and in part because similar compounds may not necessarily interact with the target macromolecule in similar ways. The latter explanation was addressed elegantly by Boström et al. [22]. They exhaustively analyzed experimental data from the protein database (PDB). The complete PDB was searched for pairs of structurally similar ligands binding to the same biological target. The binding sites of the pairs of proteins complexing structurally similar ligands were found to differ in 83 percent of the cases. The most recurrent structural change among the pairs involves different water molecule architecture. Side-chain movements were observed in half of the pairs, whereas backbone movements rarely occurred.

1.8. BIOLOGY-BASED LIBRARIES

Consideration of the foregoing discussions of complexity, structural similarity, and binding mode dissimilarity suggests that the vast majority of the possible compounds in drug-like chemistry space are simply not biologically active. However, the fact that a compound similar to another biologically active compound has a 70 percent probability of being inactive at that molecular target does not preclude its being active at a different unrelated molecular target. This suggested to us that compounds that have exhibited biological activity at any target, as well as compounds similar to biologically active compounds, might have activity at other targets, and would therefore serve as the basis for building efficient screening libraries, which we dubbed biology-based libraries. We tested this hypothesis[2] by selecting all the compounds in the AstraZeneca compound collection that had shown confirmed activity in any

screen, including compounds that had been prepared for SAR development in LO projects. At the time of this work the set consisted of approximately 72,000 compounds. The compounds were filtered to retain only those that had lead-like physicochemical properties and were devoid of objectionable functionality. We then selected against frequent hitters by choosing only those compounds that had been tested in a minimum of 10 assays, and found active in at least 2, but in no more than 5 assays. This resulted in 25,536 compounds. We assumed that if this set consisted largely of frequent hitters, the distribution of compounds would be weighted toward those that were active in five assays, with correspondingly fewer actives in four, three, and two assays. However, the distribution was exactly the opposite: 15,351 were active in two, 5,849 were active in three, 2,863 were active in four, and only 1,473 were active in five assays. We then compared the performance of these compounds against the performance of the general screening collection across 40 new assays of diverse molecular targets and found that hit rates for the selected set exceeded those of the general collection in 38 of the assays, in some cases dramatically so. The 25,536 compounds consisted of about 9,000 clusters, of which about 5000 were singletons. We concluded that these singletons would serve as good starting points for expanding the screening collection. The advantages of this approach is that hits could be quickly found in time-consuming and expensive low throughput assays, and that new privileged cores will be found. The disadvantage is that new chemotypes are not likely to be found very often. Nevertheless, we believe that the approach can add value to an overall screening program.

Another interesting approach to new hit and lead discovery has been developed at Telik, and described in detail with examples in a paper by Beroza et al. [23]. Briefly, target-related affinity profiling (TRAP) is based on the principle that almost all drugs produce their effects by binding to proteins. This fact suggests that protein-binding patterns might be a particularly relevant way to classify molecules. In TRAP, the characterization of a molecule's protein-binding preferences is captured by assaying it against a panel of proteins, the "reference panel." The panel is fixed, and each molecule in a set is assayed against it. The vector of each compound's binding affinities to the panel is its affinity fingerprint. A library of compounds can therefore be considered a set of vectors in "affinity fingerprint space," an alternative space to those defined by molecular structure. Molecules that interact with the same affinity toward the same proteins would be considered biologically similar, regardless of their chemical structures. For example, met-enkephalin and morphine are obviously very different structurally, but exhibit very similar opioid pharmacology. They also have very similar affinity fingerprints, as shown in Figure 1.4.

By contrast, morphine, an agonist, and naloxone, an antagonist, have very similar structures, but very different biological activities. The use of affinity fingerprints for lead discovery has a limitation because there is no *in silico* method to generate an affinity fingerprint for a compound. A physical sample of the compound must be available, which carries with it the costs associated with compound acquisition, chemical inventory, and assaying the compounds

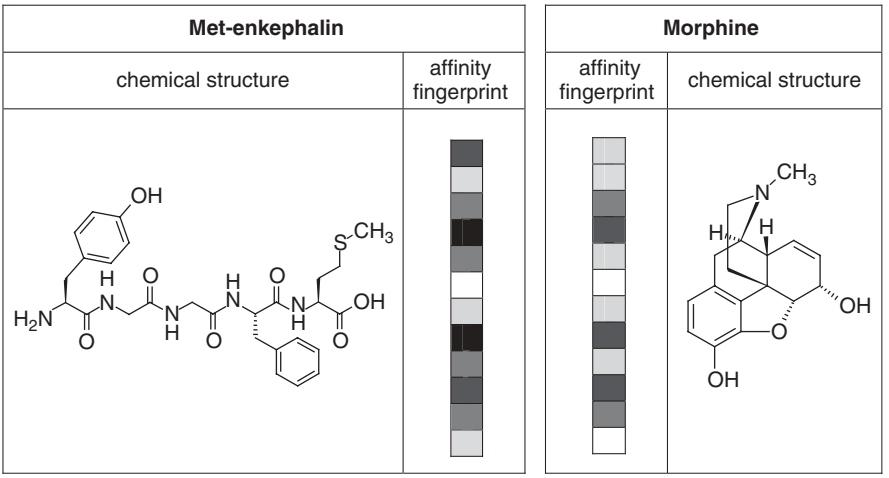


Figure 1.4. Chemical structures and affinity fingerprints for met-enkephalin and morphine. The protein panel is able to recognize the biological similarities between these two structurally distinct molecules. The two molecules have the same affinity for over half the proteins in the 12-member panel.

against the protein panel. Despite this limitation, affinity fingerprints have been shown to be an effective descriptor for identifying drug leads. The screening process that uses them for low throughput, high information assays reliably identifies molecules with activities in the low micromolar range.

1.9. A HIT-TO-LEAD EXAMPLE

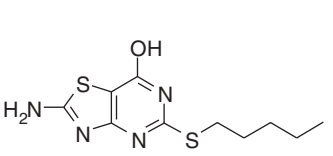
Baxter and coworkers at AstraZeneca have published a number of papers [24–26] that specifically illustrate the HTL process applied to various molecular targets following HTS campaigns. These works are distinguished by their emphasis on meeting a set of generic lead criteria. These criteria are an abbreviated version of those listed in Reference 2, and are shown in Table 1.1. Lead compounds should fulfill the majority of the lead criteria, but it is understood that very few compounds will be found that meet them all. By way of example, Reference 26 describes the process as applied to a series of thiazolopyrimidine CXCR2 receptor antagonists. The profile of the screening hit **11** is shown in Table 1.2.

The compound is not particularly potent (10 μM) in the binding assay, but did show comparable potency (2 μM) in a functional FLIPR assay, suggesting that the observed activity was real. The HTL strategy was to examine in turn the three thiazolopyrimidine substituents looking for SAR leading to increasing potency. Replacement of the 2-amino group with hydrogen retained activity, but other changes eliminated activity. The 2-amino group was therefore held

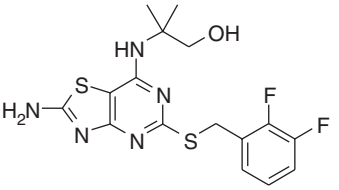
Table 1.1. Generic lead target profile

Generic assay	Generic criteria
Potency	$IC_{50} < 0.1$ μM
Rat hepatocytes	Clearance < 14 μL/min/10^6 cells
Human liver microsomes	Clearance < 23 μL/min/mg
Rat iv pharmacokinetics	Clearance < 35 mL/min/kg, Vss > 0.5 L/kg, $T_{1/2} > 0.5$ h
Oral bioavailability	F > 10%, PPB < 99.5%
Physical chemical	MW < 450, *c*log*P* < 3.0, log*D* < 3.0
Additionally	Clear SAR, appropriate selectivity data, and patentable multiple series

Table 1.2. Lead criteria profiles of screening active 11 and resulting lead 12



11



12

Generic lead criteria[a]	Compound **11**	Compound **12**
Binding $IC_{50} < 0.1$ μM	10	0.014
Ca flux $IC_{50} < 0.1$ μM	2.0	0.04
Rat hepatocyte clearance <14	49	4
Human microsome clearance <23	18	31
Rat iv clearance <35	–	25
Rat iv Vss > 0.5 L/kg	–	1.9
Rat iv $T_{1/2} > 0.5$ h	–	1.2
Rat po bioavailability >10%	–	15
Plasma protein binding >99.5%	–	98.4
Molecular weight <450	270	397
Solubility >10 μg/mL	27	0.5
*c*log*P* < 3.0	2.0	3.1
Log*D* < 3.0	2.9	3.4

[a] Units as in Table 1.1 where not stated.

constant. The 7-hydroxy group could be replaced by hydroxyethylamines, but not by the slightly larger hydroxypropylamine. The 5-alkylthio group could be replaced by a benzylthio group with a significant increase in potency. The chemistry lent itself readily to combinatorial expansions at positions 5 and 7,