High-Throughput Screening in Drug Discovery

Edited by
Jörg Hüser

WILEY-VCH Verlag GmbH & Co. KGaA

High-Throughput Screening in Drug Discovery

Edited by
Jörg Hüser

Methods and Principles in Medicinal Chemistry

Edited by R. Mannhold, H. Kubinyi, G. Folkers
Editorial Board
H.-D. Höltje, H. Timmerman, J. Vacca, H. van de Waterbeemd, T. Wieland

Related Titles

G. Cruciani (ed.)
Molecular Interaction Fields
Vol. 27

2005
ISBN 3-527-31087-8

M. Hamacher, K. Marcus, K. Stühler, A. van Hall, B. Warscheid, H. E. Meyer (eds.)
Proteomics in Drug Design
Vol. 28

2005
ISBN 3-527-31226-9

D. Triggle, M. Gopalakrishnan, D. Rampe, W. Zheng (eds.)
Voltage-Gated Ion Channels as Drug Targets
Vol. 29

2006
ISBN 3-527-31258-7

D. Rognan (ed.)
GPCR Modelling and Ligand Design
Vol. 30

2006
ISBN 3-527-31284-6

D. A. Smith, H. van de Waterbeemd, D. K. Walker
Pharmacokinetics and Metabolism in Drug Research, 2nd Ed.
Vol. 31

2006
ISBN 3-527-31368-0

T. Langer, R. D. Hoffmann (eds.)
Pharmacophore and Pharmacophore Searches
Vol. 32

2006
ISBN 3-527-31250-1

E. Francotte, W. Lindner (eds.)
Chirality in Drug Research
Vol. 33

2006
ISBN 3-527-31076-2

W. Jahnke, D. A. Erlanson (eds.)
Fragment-based Approaches in Drug Discovery
Vol. 34

2006
ISBN 3-527-31291-9

High-Throughput Screening in Drug Discovery

Edited by
Jörg Hüser

WILEY-VCH Verlag GmbH & Co. KGaA

Series Editors

Prof. Dr. Raimund Mannhold
Biomedical Research Center
Molecular Drug Research Group
Heinrich-Heine-Universität
Universitätsstrasse 1
40225 Düsseldorf
mannhold@uni-duesseldorf.de

Prof. Dr. Hugo Kubinyi
Donnersbergstrasse 9
67256 Weinheim am Sand
Germany
kubinyi@t-online.de

Prof. Dr. Gerd Folkers
Collegium Helveticum
STW/ETH Zürich
8092 Zürich
Switzerland
folkers@collegium.ethz.ch

Volume Editor

Dr. Jörg Hüser
Molecular Screening
Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany
Joerg.hueser@bayerhealthcare.com

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

Typesetting Dörr + Schiller GmbH, Stuttgart
Printing betz-druck GmbH, Darmstadt
Binding Littges & Dopf
Cover Design Grafik-Design Schulz, Fußgönheim

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN-13 978-3-527-31283-2
ISBN 10 3-527-31283-8

Foreword

Random screening of comprehensive compound collections constitutes a major source of novel lead structures reflected by industry's ongoing commitment to invest in extensive compound libraries and screening technologies. During the last decade, High-Throughput Screening (HTS) has evolved to become an innovative multidisciplinary branch in biological chemistry combining aspects of medicinal chemistry, biology, and laboratory automation. While basic medicinal chemistry techniques and views are largely conserved throughout industry, HTS strategies differ to a great extend. Each strategy can be justified by scientific rationale. However, it is also the result of different scientific backgrounds, different therapeutic areas, different technical expertise within a group, and different ways HTS is integrated within the overall discovery process in a given organization. For most strategies, the close interrelation between HTS and the molecular target approach to drug discovery renders the validity of a disease link for a selected biomolecular target an essential prerequisite for success. As a consequence, a critical assessment of HTS has to incorporate also reflections on the discovery process from target selection to appropriate screening cascades. A different approach employing phenotypic readouts, e.g. cell proliferation, has a long tradition in screening, particularly for chemotherapeutic principles in cancer and antiinfectives research. Similarly, chemical genetics makes use of small molecule perturbation of specific cellular responses to unravel the underlying gene and pathways function. Within the later paradigm, High-Throughput Screening techniques have gained increasing relevance also in academic research.

The current book presents a collection of review-style papers written by experts in the field intended to provide insights into selected aspects of the experimental lead discovery process in High-Throughput Screening. It is by no means claimed to comprehensively cover the entire field. A number of aspects have been discussed in previous volumes within this series on "Methods and Principles in Medicinal Chemistry". It complements this book series by illustrating HTS as one of the technologies of great relevance to the medicinal chemist and molecular pharmacologist working in pharmaceutical or academic research.

I am personally thankful to the Series Editors not only for providing the opportunity to present High-Throughput Screening within a single dedicated

volume, but also for their patience during the preparation of this volume. In addition, the continuous support of my colleagues, Stefan Mundt, Nils Griebenow and Peter Nell, is gratefully acknowledged.

Wuppertal, July 2006
Jörg Hüser

List of Contents

Preface

Biological "trial and error" testing of large collections of small molecules for a specific pharmacological effect is the classical route to discover novel lead compounds which subsequently serve as templates for further optimization in medicinal chemistry programs. During the past two decades, the advent of recombinant DNA technologies together with improved assay techniques and high-performance laboratory automation dramatically changed the pharmacological screening process. Nowadays, high-throughput screening (HTS) has gained attention also in the academic environment. Using HTS of compound collections in combination with cell-based tests, small molecule modulators of relevant biochemical and signal transduction pathways should be identified. Following identification of the biological target, the small molecule modulator (inhibitor or activator) substitutes mutational analysis to unravel the target protein function. The vision of this approach, referred to as "Chemical Genetics", is to identify a small molecule partner for every gene product. In the future, pharmaceutical drug research might be stimulated by Chemical Genetics by revealing novel drug targets and initial lead structures.

The present volume provides fascinating insights into this important part of the early drug discovery process. Four most important issues relevant to HTS are covered: a) concepts of pathway/phenotypic versus target-based screening, b) automation technologies, c) assay technologies, and d) data analysis.

Part I contrasts the two approaches of "Chemical Genetics" using pathway assays dependent on a initially not defined set of possible drug receptor sites and the target-based lead finding process commonly used in pharmaceutical research. Caroline Shamu uses case studies to introduce the basic concept, the assay technologies and selected results. Jörg Hüser and colleagues summarize the concepts of target-directed screening for lead discovery and discuss strengths and weaknesses of random/diversity screening when compared to knowledge-based *in silico* methodologies.

In **Part II**, John Comley provides a general overview of laboratory automation technologies covering assay carriers, liquid handling automats, signal detection instrumentation, and robotic integration.

Part III focuses on HTS assay technologies. For the discovery of novel lead candidates the choice of the appropiate assay technology and its technical realiza-

tion will determine the overall quality of the screening experiment. There are two general appoaches for pharmacological assays: Assays measuring the binding of a candidate molecule to the target receptor ("binding tests") and assays monitoring the function of a target (or pathway) to visualize a possible modulation by small molecules. In chapter 4, Jörg Hüser and colleagues focus on principles of functional cell-based test systems. Improved readout techniques together with the rich molecular biology toolbox have rendered cell-based assays an important methodology for the targeted discovery of pharmacological lead compounds in HTS. Designed cell-based HTS assays ideally combine high specificity for and superior sensitivity towards the targeted receptor. In addition, measuring receptor function rather than binding allows one to monitor all possible drug-receptor interactions including allosteric modulation and reveals additional information on ligand efficacy, i.e. agonism or antagonism. Chapter 5 by William Mallender and colleagues gives an overview on functional biochemical tests, providing examples for the most important drug target classes approached by enzyme assays (proteases, kinases and others). In chapter 6, Peter Lipp and Lars Kästner introduce and critically discuss "Image-based High-content Screening", a recently emerging technology using subcellular imaging for target-based and pathway assays.

The last set of papers (**Part IV**) covers various aspects of HTS data analysis. Chapter 7, by Hanspeter Gubler, introduces concepts of HTS data management, assay quality assurance, and analysis. It touches on some fundamental statistical consideration relevant to handling large sets of compound activity data. Chapter 8 (Peter Nell and Stefan Mundt) illustrates the use of chemoinformatic tools, e.g. structural clustering of active compound sets, aiming to discriminate between specific hits and compounds acting through unspecific mechanisms and providing clues for preliminary SARs or pharmacophoric elements (i.e. molecular fragments contributing to activity). In chapter 9, Roger Smith and Nils Griebenow discuss pro's and con's of focused library screening, including methodologies and strategies to design such subsets of the available compound file and its use for lead discovery. The data analysis section is concluded by a paper by Jeremy Caldwell describing the consequent exploitation of large activity databases derived from functional cell-based screening by data mining technologies to reveal molecules acting through unexpected mechanisms.

The series editors are grateful to Jörg Hüser for his enthusiasm to organize this volume and to work with such a fine selection of authors. We believe that this book adds a fascinating new facet to our series on "Methods and Principles in Medicinal Chemistry". Last, but not least we thank the publisher Wiley-VCH, in particular Renate Dötzer and Dr. Frank Weinreich, for their valuable contributions to this project and to the entire series.

July 2006

Raimund Mannhold, Düsseldorf
Hugo Kubinyi, Weisenheim am Sand
Gerd Folkers, Zürich

List of Contributors

Martin Bechem
Molecular Screening Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Michael Bembenek
Department of High Throughput Biochemistry
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Ulf Brüggemeier
PH-R&D Molecular Screening Technology
Bayer HealthCare AG
Research Center Wuppertal
42096 Wuppertal
Germany

Nils Burkhardt
Molecular Screening Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Jeremy S. Caldwell
10675 John Jay Hopkins Drive
San Diego, CA 92121
USA

Sa L. Chiang
Seeley Madd 504, Harvard Medical School
250 Longwood Avenue
Boston, Ma 02115
USA

John Comley
HTStec Limited
12 Cockle Close, Newton
Cambridge, CB2 5TW
UK

Larry Dick
Department of Oncology Biochemistry
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Nils Griebenow
Medicinal Chemistry
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Hanspeter Gubler
Discovery Technologies
Novartis Institutes for BioMedical Research
WSJ-350.P.15
4002 Basel
Switzerland

Jörg Hüser
Molecular Screening Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Jeff Janes
10675 John Jay Hopkins Drive
San Diego, CA 92121
USA

Lars Kaestner
Molecular Cell Biology
University Hospital
Medical Faculty
Saarland University
66421 Homburg/Saar
Germany

Bernd Kalthof
Molecular Screening Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Michael Kuranda
Department of High Throughput Biochemistry
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Ping Li
Department of Protein Sciences
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Peter Lipp
Molecular Cell Biology
University Hospital
Medical Faculty
Saarland University
66421 Homburg/Saar
Germany

Emanuel Lohrmann
Molecular Screening Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

William D. Mallender
Department of Oncology Biochemistry
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Saurabh Menon
Department of High Throughput Biochemistry
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Stefan M. Mundt
Molecular Screening Technology
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Peter G. Nell
Medicinal Chemistry
Bayer HealthCare AG
Pharma Discovery Research Europe
42096 Wuppertal
Germany

Eneida Pardo
Department of High Throughput Biochemistry
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Tom Parsons
Department of Protein Sciences
Millennium Pharmaceuticals, Inc.
Cambridge, MA 02139
USA

Roger A. Smith
Department of Chemistry Research
Pharmaceuticals Division
Bayer HealthCare
400 Morgan Lane
West Haven, CT 06516
USA

Jochen Strayle
PH-R&D Molecular Screening Technology
Bayer HealthCare
400 Morgan Lane
West Haven, CT 06516
USA

Part I
Concept of Screening

1
Chemical Genetics: Use of High-throughput Screening to Identify Small-molecule Modulators of Proteins Involved in Cellular Pathways with the Aim of Uncovering Protein Function

Sa L. Chiang

1.1
Introduction

Understanding cellular pathways and their molecular mechanisms is one of the longstanding goals of scientific research. However, the tools that researchers utilize to study biological processes become progressively more sophisticated as technology and our knowledge of biology advance. This chapter discusses one of the most recent and exciting developments in this area: chemical genetics and the use of small, bioactive molecules to characterize the components of cellular pathways and their functions.

1.2
Classical and Chemical Genetics

All genetic approaches depend on the ability to perturb gene function and to correlate phenotypic changes with specific changes in gene function. Classical genetics relies on physical perturbation of a gene, through methods such as irradiation, chemical mutagenesis or insertional mutagenesis (e.g. via the use of transposons). There are also numerous molecular biology techniques for creating directed mutations that allow highly specific modifications at the level of the gene or even the nucleotide in a variety of experimental systems. These methods are well established and readers should consult a standard genetics text for the relevant descriptions.

A common feature of all classical genetic methods is that they cause a permanent change in the structure of a gene. Therefore, except for the few situations noted below, the phenotypes arising from classical genetic mutations are irreversible. A notable disadvantage of this irreversibility is that it hinders the study of genes that are essential for viability. Irreversibility also makes it difficult to study the effects of

temporal variations in gene expression or protein function. Placing genes under the control of inducible promoters has gone some way towards solving these problems, but inducible promoters, of course, act at the level of transcription and the researcher may not obtain sufficient control over the activity of the encoded protein. Temperature-sensitive mutations may provide some control over protein activity, but they are not easy to construct and generally cannot be used in animal models. Pleiotropic effects caused by temperature shifts may also complicate analysis.

In the last decade, researchers have increasingly explored the use of low molecular weight chemical entities to modulate and characterize protein function. These methods are generally analogous to classical genetic approaches and have accordingly been termed *chemical genetic* strategies [1, 2]. Whereas classical genetics uses physical modification of a gene to perturb protein function, chemical genetics employs specific, biologically active small-molecule modulators to perturb protein function.

It is relatively simple to imagine mechanisms by which small molecules may modulate protein function. Enzymatic activity can be affected by the binding of small molecules to active or allosteric sites and protein–ligand interactions may be disrupted by small molecules that interfere with binding between interaction partners. Alternatively, interaction partners could be brought together more effectively by ligands that bind to both. Interactions between small molecules and their targets may be reversible or irreversible and protein function can be either diminished or enhanced, depending on the situation.

One advantage of small molecules is that they can be used in biological systems where there is little or no ability for classical genetic manipulation. Moreover, since many small molecules will not interact irreversibly with their targets, chemical genetics is expected to provide enhanced opportunities to create conditional phenotypes. Theoretically, a high degree of control over protein function could be afforded by simply adding or washing away a small-molecule modulator. Indeed, several widely used inducible promoter systems (lac, arabinose, tetracycline) employ small-molecule inducers. In addition, conditional phenotypes induced by small molecules can potentially be studied in animal models, where temperature-dependent phenotypes generally cannot be used.

In actuality, the ability of small molecules to induce and reverse phenotypes will depend on factors such as binding kinetics and the physical accessibility of the target. More important, most protein targets have no known small-molecule modulators of their activity. Obtaining a specific and potent small-molecule modulator for a chosen target often requires structure-based drug design or full-scale high-throughput screening (HTS) and medicinal chemistry optimization may also be necessary. Owing to this "front-end" effort, directed perturbations in many systems are currently more difficult to achieve with chemical genetics than with classical genetics.

1.2.1 Forward and Reverse Screens

Many classical genetic screens begin with mutagenesis of organisms or populations of cells, followed by attempts to associate the resulting phenotypic changes with specific genes. These approaches are termed *forward genetic* strategies and can be generally characterized as starting with phenotypes and progressing towards the identification of the genes that are responsible for those phenotypes. With the advent of molecular biology, directed mutagenesis techniques became available and these advances allowed *reverse genetic* studies, which begin with the introduction of programmed or directed mutations into a known target gene, followed by analysis of the resulting phenotypes to obtain information about the function of that gene.

Both forward and reverse genetic strategies have also been employed for chemical screens. In a forward chemical genetic screen, chemical libraries are screened for compounds that produce a phenotype of interest, typically in a cell-based assay. An example of forward chemical genetics is the strategy that was employed to identify inhibitors of mitotic spindle bipolarity [3]. Monastrol was identified through a chemical genetic screen employing a whole-cell immunodetection assay to screen a library of 16 320 compounds for those that increase nucleolin phosphorylation, a phenotype predicted for cells experiencing mitotic arrest. A total of 139 compounds found to increase nucleolin phosphorylation were subjected to further analysis, resulting in the identification of five compounds that affect the structure of the mitotic spindle. One of these induced the formation of a monoastral microtubule array and was accordingly named monastrol. Monastrol has since been employed in multiple studies as a tool for investigating the process of cell division.

Conversely, in a reverse chemical genetic screen, small-molecule libraries might be screened for compounds that bind to a purified protein target, modulate the activity of the target or affect the target's ability to interact with other proteins. Such compounds could then be used in cell-based assays to characterize the function of the target protein in cellular pathways. For instance, a luminescence-based reverse chemical genetic screen was employed to identify inhibitors of rabbit muscle myosin II subfragment (S1) actin-stimulated ATPase activity [4]. The most potent compound identified (*N*-benzyl-*p*-toluenesulfonamide; BTS) inhibited S1 ATPase activity with an IC_{50} of ~5 μM and BTS also inhibited the activity of skeletal muscle myosin in a gliding-filament assay. Subsequent studies demonstrated that BTS inhibits the binding of myosin–ADP to actin and also affects various properties of rabbit and frog muscle preparations.

A similar reverse chemical genetic screen identified an inhibitor (blebbistatin) of nonmuscle myosin II [5]. Blebbistatin was found to inhibit a variety of activities in whole vertebrate cells, including directed cell migration and cytokinesis. The use of blebbistatin and additional drugs (including monastrol) to manipulate the mitotic process in whole cells led to the discovery that ubiquitin-dependent proteolysis is required for exit from the cytokinetic phase of the cell cycle. Blebbistatin is thus

another example of how small molecules identified via HTS can be used to study the function of both individual proteins and major cellular processes.

The above points are summarized in Table 1.1.

Table 1.1 Attributes of classical genetic, chemical genetic, and RNAi approaches.

	Classical genetics	Chemical genetics	RNAi
Nature of perturbation	Permanent genetic change; true null is usually possible; heritable	Transient or irreversible, depending on situation; not heritable	Transient; generally not heritable
Directed perturbation	Possible in many systems, can be accomplished at the single-nucleotide level	Requires identification of a specific effector; may require synthetic chemistry effort to optimize	Yes; some occasional off-target effects
Conditional perturbation	Possible in some situations (temperature-sensitive alleles, inducible expression)	Possible, depending on specific situation	Yes, although difficult to control temporally
Target identification	Often simple, depending on situation	Often difficult	Target is known

1.3 Identifying Bioactive Molecules

Biologically active small molecules have often been discovered by testing a single compound at a time, but such an approach is obviously highly inefficient and cost-intensive in terms of both reagents and personnel time. As a result, HTS technologies have been developed to screen large numbers of compounds simultaneously, typically through the miniaturization and automation of assay protocols. What constitutes high throughput will vary depending on technical considerations and the screener's economic situation, but generally, the number of compounds involved may range from tens of thousands to several million. With current technologies, throughput is high enough that screening this number of compounds can be accomplished within several weeks, if not within several days.

A large number of screening technologies are available today for identifying bioactive small molecules; in fact, there are too many methods to be discussed adequately in this chapter. A very general discussion is provided here, but the reader is referred to other chapters in this volume for greater detail.

High-throughput screening assays can be divided into two main classes: "pure protein" and "cell-based". Pure protein screens generally have optical assay readouts that monitor enzymatic or binding activity. For instance, fluorescence polar-

ization or FRET techniques are commonly used to screen for compounds that affect binding between protein partners. In pure protein assays, every compound screened should have equal access to the target. However, the membrane permeability characteristics of any active molecules that are identified may subsequently pose a major concern if the target is intracellular.

Readouts for cell-based screens may rely on reporter gene systems (e.g. luciferase, β-lactamase), cell density, cell viability or cell morphology. Screen readouts can also be divided into two broad classes: uniform well readout acquired via plate readers and images acquired via automated microscopy. Screens involving image-based readouts are usually technically more difficult and their computational analysis more challenging than plate-reader screens. However, image-based readouts can provide a vast amount of information and microscopy screens are therefore often referred to as *high-content screens* [6–8]. The development of equipment and data analysis techniques for automated imaging screens is an area of active research, but application of early technologies has already yielded promising results [9–15].

1.4 Target Identification

When biologically active small molecules are identified through chemical screens, particularly forward genetic screens, a substantial amount of work is often required to identify the molecular target. Ignorance of the target does not preclude clinical or research use of the molecule; indeed, many clinical agents have been used effectively even when their targets were not known (e.g. aspirin, nitroglycerin, fumagillin, epoxomycin). Nevertheless, defining the molecular mechanism of action is vital to understanding the biological principles involved and also for potentially creating more potent or specific molecules through synthetic chemistry approaches.

Target identification is currently a major area of research in chemical genetics and Tochtrop and King have recently provided an excellent discussion of this topic [16]. This section will therefore emphasize HTS-related examples and also discuss some recent work not included in that review.

1.4.1 Hypothesis-driven Target Identification

For well-characterized cellular pathways, it is sometimes possible to deduce the target of a small molecule by comparing data from across the field. For instance, characteristic phenotypes induced by the small molecule may permit the assignment of the target to a previously identified complementation group. Subsequent hypothesis-driven testing of potential targets can then be undertaken.

This deductive approach was used successfully to define the target of monastrol, a small-molecule inhibitor of mitotic spindle bipolarity [3]. As noted above, mon-

astrol was identified in a forward chemical genetic screen for compounds that affect the structure of the mitotic spindle. Monastrol induces the formation of a monoastral microtubule array and, since previous studies had noted that antibody inhibition of the Eg5 kinesin also causes monoaster formation, it was postulated that Eg5 might be a target of monastrol. Subsequent work demonstrated that monastrol reversibly inhibits Eg5-driven microtubule motility *in vitro* but is not a general inhibitor of motor proteins.

Hypothesis-driven target identification was also used in a chemical screen for inhibitors of SARS coronavirus (SARS-CoV) replication [17]. In this study, a library of 50 240 compounds was screened by imaging for compounds that inhibit the cytopathic effect (CPE) of SARS-CoV towards Vero cells. A total of 104 compounds demonstrating effective inhibition of CPE and viral plaque formation were then tested *in vitro* against two protein targets known to affect SARS-CoV replication (M^{pro} protease and the NTPase/helicase), and also in a pseudotype virus assay for inhibition of S protein–ACE2-mediated entry of SARS-CoV into 293T cells. Two inhibitors of M^{pro}, seven inhibitors of Hel and 18 inhibitors of viral entry were identified and each of these three classes contained at least one inhibitor active in the low micromolar range. The authors subsequently assayed the 104 compounds against other common RNA viruses and found that most were specifically active against SARS-CoV, and approximately 3% were active against all viruses tested.

1.4.2 Affinity-based Target Identification

Affinity-based methods such as affinity labeling, affinity chromatography and crosslinking are also commonly used target identification strategies. Since these approaches often require synthetic modification of the molecule (e.g. addition of linkers or immunoreactive epitopes), care must be taken that the modifications do not interfere with the molecule–target interaction. Nonspecific interactions may also complicate the analysis and controls must be designed to address this issue. Despite such limitations, however, these strategies are of tremendous utility in target identification.

With affinity labeling, molecules may be radioactively or chemically labeled; they may be synthetically modified with reactive groups to promote covalent attachment to the target or tagged with specific moieties to facilitate detection. Standard protein fractionation and detection techniques can then be used to identify proteins from crude extracts that are specifically labeled by the molecule. Examples of drugs whose cellular targets were determined by such methods include acetylcholine, the anti-angiogenic agent fumagillin [18] and the antifungal lipopetide echinocandin [19].

Affinity purification of putative targets from cellular extracts is accomplished with molecules immobilized on a solid support used either as a column matrix or as a bead slurry. Small molecules for which cellular targets were identified through affinity purification include the immunosuppressant FK506 [20], the kinase inhibitor purvalanol [21] and the anti-inflammatory compound SB 203580 [22].

Additional studies involving affinity chromatography were instrumental in further characterizing the biochemistry of FK506 and purifying the target of the structurally related compound rapamycin [23].

A natural technological development of affinity-based target identification involves probing proteome chips with small molecules [24]. A recent application of this approach [25] is discussed in Section 1.4.4. Conversely, it is also possible to probe small-molecule microarrays with a known protein to identify specific interactions [26].

1.4.3 Genomic Methods of Target Identification

Yeast three-hybrid system The yeast two-hybrid system has been widely used to study protein–protein interactions and the approach was adapted to create a yeast three-hybrid system for identifying protein–ligand interactions [27]. The method relies on the use of a hybrid "bait" ligand consisting of the query molecule linked to a known ligand. A protein fusion between the known ligand's receptor and a DNA-binding domain serves as the "hook", while the "fish" is a protein fusion between a transactivation domain and the target protein. If the target protein interacts with the query molecule, the "fish" and "hook" will be brought together by the "bait" and transcription from a reporter gene is activated. Hence it should be possible to identify the target of the query molecule by cloning a library into the "fish" domain and screening or selecting for activity of the reporter.

The three-hybrid proof-of-principle study was conducted with a dexamethasone–FK506 hybrid ligand and a Jurkat cDNA library. This experiment successfully identified two variants of human FKBP12 as targets of FK506. Another proof-of-principle study employed a dexamethasone–methotrexate hybrid ligand to screen a mouse cDNA library and identified dihydrofolate reductase as a target of methotrexate [28].

A three-hybrid approach was recently used to identify targets of various kinase inhibitors [29]. For each inhibitor, a hybrid ligand was synthesized by attachment to methotrexate and the hook protein was a LexA–DHFR fusion. A three-hybrid screen of a human cDNA library with a purvalanol B–methotrexate ligand identified several (but not all) known purvalanol targets and also several new candidate targets, and all identified targets were kinases. Affinity chromatography and enzymatic assays confirmed 12 of 16 novel candidate targets identified in the cDNA screens.

Induced haploinsufficiency One genomic approach to target identification relies on the premise that altering the gene dosage of the target will affect sensitivity to the small molecule. For example, inactivating one copy of the target gene in a diploid organism would in many cases be expected to increase sensitivity. This method of identifying drug targets through *induced haploinsufficiency* was established by Giaever et al. [30], who constructed and screened a set of 233 heterozygous yeast deletion mutants for those demonstrating increased sensitivity to

known drugs. Each mutant was chromosomally tagged with a unique oligonucleotide and the mutants were pooled and grown in the presence of tunicamycin, at a level of drug that is sublethal for wild type. The relative number of each mutant in the pool was monitored at various time points by polymerase chain reaction (PCR) amplification and fluorescence labeling of all tags in the pool, followed by hybridization of the PCR-generated probes to an oligonucleotide microarray. The fluorescence intensity generated at each spot on the array permitted quantitation of the relative abundance of each corresponding heterozygote in the pool. Mutants unaffected by tunicamycin showed no reduction in signal over time, whereas tunicamycin-sensitive heterozygotes showed varying decreases in signal. In this study, one known and two new loci were identified and confirmed as involved in tunicamycin resistance.

A subsequent study extended this approach to test 78 compounds against a pool of 3503 yeast heterozygotes, representing over half the yeast genome [31]. Most of the compounds tested had known activities and many also had known targets. Of 20 compounds with known targets, in most cases this method correctly identified the target or members of the target complex. Targets were also identified for a number of compounds with previously unknown targets in yeast.

Recently, a complementary approach of using gene overexpression to identify small-molecule targets has been pioneered [32]. A plasmid-based yeast genomic library was introduced into yeast and 7296 individual transformants were arrayed in 384-well plates. The arrayed library was then replicated on to solid agar containing an inhibitor. The plasmid inserts from resistant strains were then recovered and sequenced. Pkc1 was identified via this method as a target for the inhibitor and this was subsequently confirmed by affinity chromatography and genetic and biochemical assays.

Expression profiling Expression profiling was recently employed in identifying the targets of a class of small-molecule antagonists of FK506 in yeast cells subjected to salt stress [33]. These molecules (termed SFKs for suppressors of FK506) were identified in a chemical genetic screen for molecules that rescue yeast growth in the presence of high salt and FK506 [34]. Expression profiling results obtained with SFK2-treated yeast suggested that the Ald6 p pathway was a target of SFK2 and haploinsufficency screening supported this hypothesis. Overexpression of ALD6 was found to suppress the effects of SFK2–SFK4 on growth and the ability of SFKs to inhibit Ald6 p *in vitro* was subsequently demonstrated.

In addition to using DNA microarray technologies in conjunction with haploinsufficiency studies, it is likely that gene expression profiling will be used increasingly as a primary means of identifying the targets of bioactive small molecules. Many research groups have used expression profiling to identify characteristic patterns of gene expression ("fingerprints") that are associated with certain disease states or biological pathways. If the patterns are sufficiently unique, then it is sometimes possible to assign previously uncharacterized mutants to specific cellular pathways or complementation groups based on their expression profiles. Analogously, by profiling cells grown with and without a small-molecule modu-