1
Yeast-Based Chemical Genomic Approaches

Katja Hübel

1.1
Introduction

In this chapter a yeast-based chemical genomic approach to elucidate the mode of action of small molecules is described. For this approach a genome-wide screen using the *Saccharomyces cerevisiae* yeast deletion strain collection has been developed. The yeast-based screen is explained in detail including case studies for its successful use. Besides the yeast-based chemical genomic approach the reader will learn about the basic principle of chemical genomics, the advantages of employing small molecules to study biological questions and about advantages and disadvantages of cell-based assays and the use of yeast in cell-based screens.

1.2
The Biological Problem

1.2.1
Interplay between Organic Chemistry and Biology

A potential starting point in the cycle between organic synthesis and biology is the analysis of a biological system or phenomenon of interest (Box 1.1). For this purpose a given biological system (for example a certain signaling pathway) is typically perturbed by adding a small molecule and then analyzed [1].

There are two ways to perform these so called primary screens of large compound libraries (see Chapter 3): biochemical *in vitro* assays (reverse chemical genetics) and cell-based screening (forward chemical genetics) (Box 1.2). Interesting and potential small molecules or whole compound classes identified thereby serve as starting points for identification and validation of protein targets [2].

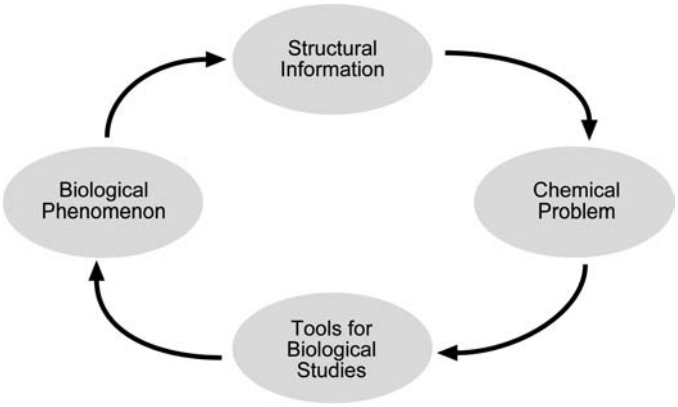
Chemical Biology: Learning through Case Studies. Edited by H. Waldmann and P. Janning

ISBN: 978-3-527-32330-2

Box 1.1

Interplay between Organic Synthesis and Biology in Chemical Biology Research: Chemical biology often starts by analyzing a biological system or phenomenon of interest. In this analysis structural information is deduced concerning the structure of biomacromolecules (for example a protein) involved in a particular biological phenomenon or the structure of small molecules that interact with these macromolecules. This structural information is then employed to define unsolved chemical problems, for example the development of new methods for the synthesis of low molecular weight compounds such as natural products and analogs thereof, and of biomacromolecules (for example semi-synthetic proteins; see Chapters 12 and 13). The design and synthesis of inhibitors that can be used to perturb and probe biological systems is also of major interest. Once methods of accessing the desired compounds have been devised and developed, the newly prepared compounds are employed in appropriately designed biological and/or biochemical experiments. The results gleaned may then give rise to a better understanding of the biological problem. They may also highlight new structural features, thus forming the basis for a new round of investigation [1].

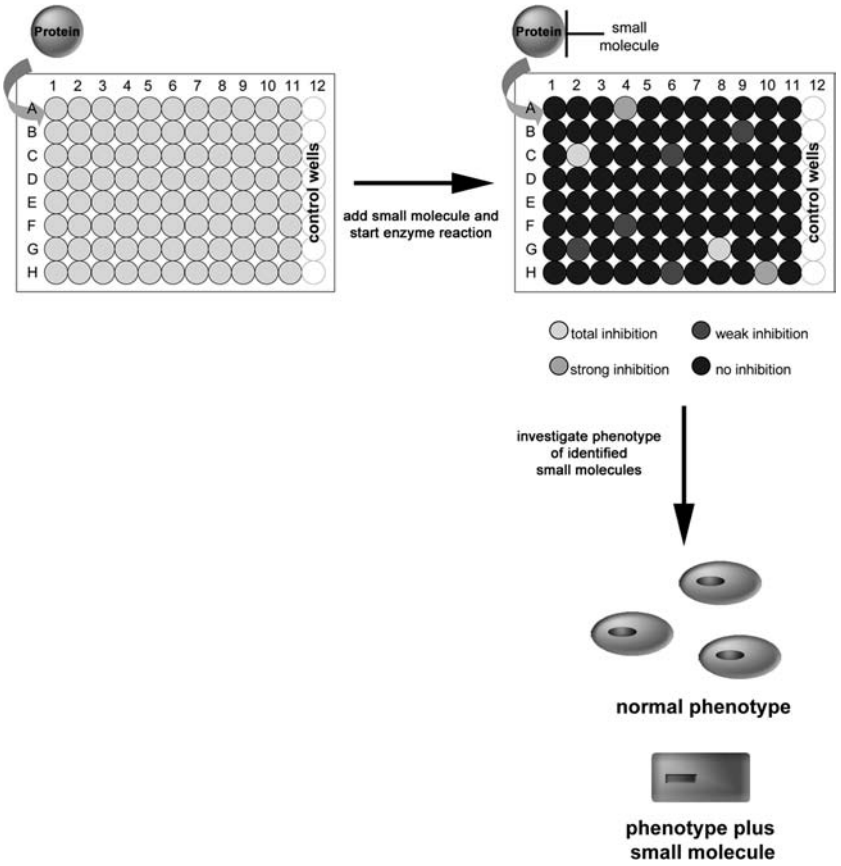
The reader should be aware that not always all criteria of the outlined cycle can be applied and that the cycle can be accessed at any starting point (Reprinted with permission from [1]).



Box 1.2 Reverse Chemical Genetics and Forward Chemical Genetics

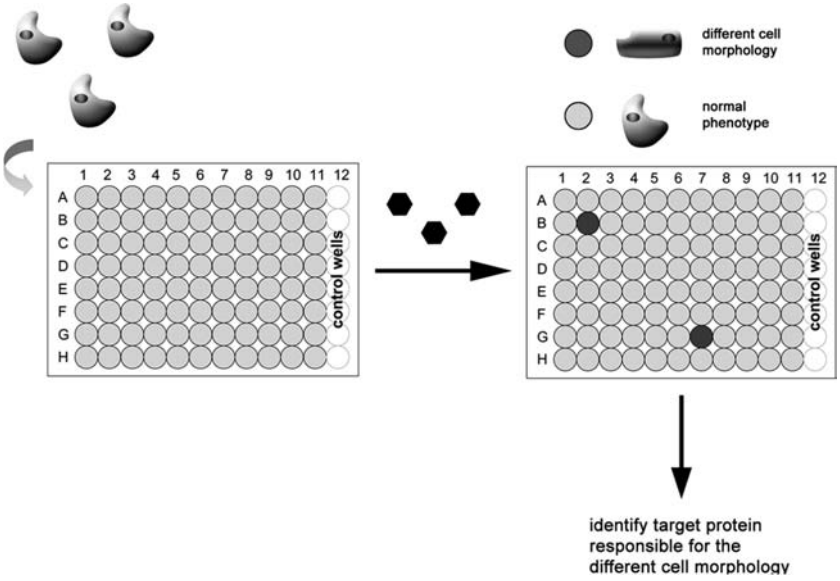
Reverse Chemical Genetics (Biochemical *in vitro* Assays) Compound collection screening can be used to efficiently investigate large numbers of compounds in biochemical *in vitro* assays to identify small molecules that are capable of modulating the biological target of interest. In this reverse chemical genetic approach a chosen target protein is screened against a compound collection. The screens are

performed for instance in the 384 well-microtiterplate format. As the target protein is a gene product the gene sequence of interest is first cloned and then the target protein is expressed. Small molecule ligands able to bind to the target and to activate or inhibit its function are then identified. In a next step phenotypic effects of adding the small molecule to a cellular system are studied. The reverse chemical genetic approach starts at the target (protein) and progresses to the phenotype [2, 3] (Reprinted with permission from [2]).



Forward Chemical Genetics (Cell-Based Screening) High-content cell-based assays are a complementary approach to the biochemical *in vitro* identification of compounds by binding for example to an isolated protein [4]. In this forward chemical genetic approach small molecules are used to screen for a desired phenotypic effect in the biological system under investigation. Once the screen reveals a suitable compound, the gene product which is modulated by the small molecule must be identified [3]. The biological system under investigation can be a prokaryotic and eukaryotic single cell organism (bacteria, fungi),

physiological or pathological cells from complex multicellular vertebrate or mammalian organisms, or a whole higher organism, such as fly, worm, zebra fish or mouse. Thus, the forward chemical genetic approach identifies potential compounds on the basis of their conditional phenotypic effect on a whole biological system and not on the basis of their inhibition of a specific protein target. The approach goes from phenotype to protein (target) [2, 3]. Reprinted with permission from [2].



1.2.2 Chemical Genomics

All biological phenomena can be traced back to chemical processes, in principle, biology is molecular. The sequencing of the genomes of model species such as *Caenorhabditis elegans* (worm), *Mus musculus* (mouse), *Homo sapiens, Arabidopsis thaliana* (plant) and *Saccharomyces cerevisiae* (yeast) is the basis to get a complete understanding of cells and organisms. For this goal, the genetic information is not sufficient. Biological research has to investigate the function, modification, and interaction of proteins and modulators of their activity. In the last few years a new approach, in which small molecules are used as modulating ligands that enable cell biological studies and an understanding of the function of certain gene-products, has received increasing attention [1]. This approach is named *Chemical Genomics* (Box 1.3).

To use genomic approaches in chemical biology loss-of-function mutants are generated by deletion, transposon insertion or RNA interference (RNAi) in a genetic model organism [4]. This global approach is often conducted in large

Box 1.3

Chemical Genomics: In chemical genomics chemical ligands are used in genomic approaches to understand the function of a protein, to elucidate pathways and to examine their effect on gene expression. The approach requires the identification of compounds that act as positive or negative regulators of individual gene products, pathways or cell types. Furthermore, the potential therapeutic use of small molecules can be evaluated [2] (see e.g. Chapters 4, 5, 7–9).

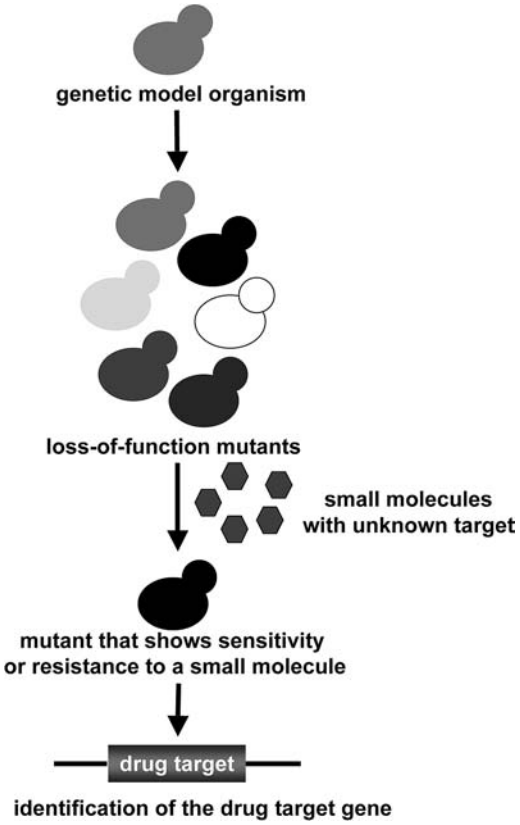


Figure 1.1 Genomic approaches to target identification. Loss-of-function-mutants are generated in a genetic model organism (yeast, worm, fly or mammalian cells). Mutants that exhibit sensitivity or resistance towards a small molecule are further investigated to identify the gene target responsible for this alteration [2]. (Reprinted with permission from [2].)

throughput (96 well- or 384 well-microtiterplate format) and assists in elucidating the mode of action of small molecules with an unknown target or can help to reveal the mode of action of compounds whose direct targets are known but whose cellular consequences are not fully understood (Figure 1.1).

1.2.3 Small Molecules in Chemical Genomics

Compared to traditional genetics (Box 1.4) the use of small molecules has several advantages. A protein can accomplish more than one function. By gene deletion a multifunctional protein will lose all its functions whereas a selective small molecule can perturb specific protein features. In some cases gene deletion may even lead to death of the organism under investigation. Other benefits are temporal control and reversibility of inhibiting protein function by using small molecule modulators of protein function. Effects can be induced by adding a small molecule and subsequent washing off of the modulating molecule results in return to the non-perturbed state. Most of those genomic approaches do not require any chemical modification. Apart from that many of the methods require minute quantities of material.

Disadvantages of small molecule modulators of protein function may be lack of specificity, cytotoxicity and solubility [2, 5]. A lack of specificity can frustrate drug development and can lead to misinterpretation of observed effects [2, 6]. Small molecules may be toxic to the organism under investigation, often in a dose-dependent manner.

Box 1.4

Traditional Genetics: Traditional genetics employs a variety of methods. In *forward genetics* the genome of an organism is randomly mutated and thus the mutants produce a change in phenotype. In *reverse genetics* an already identified gene is mutated or deleted and the resultant phenotype is studied. Other genetic methods are the use of *conditional alleles*, the *Cre-loxP system* or *RNAi*. For further information the reader is referred to a review by Zheng and Chan (and references therein) [4] and to "*Genes VI*" by Lewin [7]. A short introduction to RNAi is also given in Chapter 7.

1.2.4 Cell-based Genomic Approaches

Genomic approaches are cell-based which brings several advantages. First, the target is in a cellular environment. Thus, its native conformation and its association with cellular cofactors are kept. Furthermore, a cell-based screen will select for compounds that are cell permeable, considering only compounds that pass into a cell and act before they are inactivated or exported [9]. In comparison, *in vitro* screens (Box 1.2) are limited by the fact that the small molecule selected may be impermeable to the cell. *In vitro* effects might not mimic the natural state of events in a cell. A compound could be rapidly metabolized in the cell, leading to differences in results [10].

1.2.5 Yeast-Based Chemical Genomic Approaches

For several reasons the baker's yeast *Saccharomyces cerevisiae* is a valuable system for identification of new drug targets, target-based and non-target-based compound screening, and detailed analysis of the cellular effects of compounds. Yeast-based assays have most of the advantages of cell-based assays, even if the target is not itself a yeast protein. Many human proteins will function or can be made to function in yeast. This means that conformation or protein–protein associations required for the function can occur. Most compounds that pass through the yeast cell wall and plasma membrane can also be expected to pass into mammalian cells, and the yeast Pleiotropic Drug Resistance exporter proteins are structurally similar to the mammalian Multiple Drug Resistance efflux pumps [9]. In general, 31% of proteins encoded by yeast have human homologs. The budding yeast *Saccharomyces cerevisiae* has additional benefits. This eukaryotic model system has long been used to study cellular processes, mammalian diseases and pathways due to its ease of manipulation and genetic tractability. It is inexpensive to maintain and grow, has a short life cycle of 90 min and is stable in both the diploid and haploid state (Box 1.5). Information on the function of yeast genes and their corresponding protein products is available through several databases, including the *Saccharomyces* Genome Database (SGD; www.yeastgenome.org), the Yeast Protein Database (YPD; www.proteome.com), the Munich Information Center for Protein Sequences (MIPS), Comprehensive Yeast Genome Database (CYGD; http://mips.gsf.de/genre/proj/yeast/index.jsp), and the Yeast Resource Center (http://depts.washington.edu/~yeastrc) [8].

Box 1.5

Diploid, Haploid: Diploid cells contain two copies of each chromosome and haploid cells one copy. A chromosome is a discrete unit of the genome carrying many genes.

Homozygote, Heterozygote: A homozygote is an individual with the same allele at a corresponding locus on the homolog's chromosome. A heterozygote is an individual with different alleles at particular loci. An allele is one of several alternative forms of a gene occupying a given locus on a chromosome. A locus is the position on a chromosome at which the gene for a particular trait resides; a locus may be occupied by any one of the alleles for the gene [7].

MATa, MATα: Yeast has a locus responsible for mating. This mating type locus contains two wild type alleles MATa and MATα. These are able to mate with one another to yield a MATa/MATα. MATa/MATα cells cannot mate with cells of either mating type. Generally, MATa and MATα strains will be haploid and MATa/MATα strains will be diploid [11].

However, yeast-based chemical genomic approaches have certain limitations. For example, some drug targets may not be encoded in the relatively small yeast genome and target gene phenotypes may be masked by genetic and functional redundancy [12]. Many oncogenes and tumor-suppressor genes that constitute potential drug targets have no clear orthologs in yeast [8]. In addition, results may not translate to more complex eukaryotes, hence additional validation steps are required [12]. It is difficult to study the effect of drugs on certain processes, like cell–cell signalling, that occur only in multicellular organisms. Thus, drug screening in yeast may serve as a primary step, with subsequent complementary approaches implemented in mammalian systems [13].

The most important yeast-based functional genomic and proteomic technologies include the yeast deletion strain collection, drug-induced haploinsufficiency screens, synthetic lethal screens and the yeast two-hybrid and the yeast three-hybrid system.

The yeast deletion strain collection as well as the drug-induced haploinsufficiency screen are explained in detail in Section 1.4.

In synthetic lethal screens drug-induced inhibition of growth is observed. The underlying principle is that, without drug, inactivation of two genes in redundant pathways leads to a loss of viability, whereas inactivation of either gene has no effect. Adding a drug, a protein product of a gene is inactivated. The protein product of a second gene is by itself not essential but prevents loss of viability in the presence of the drug. Deletion of the second gene leads to hypersensitivity to a dose of the drug that is not lethal in the wild-type cell (Figure 1.2) [8, 10].

Originally described by Fields and Song [14] the yeast two-hybrid system is based on the yeast Gal4 transcription factor and uses the *Saccharomyces cerevisiae* transcriptional machinery to discover new protein interactions (Box 1.6). A known protein of interest, the bait, is expressed as a fusion to the DNA binding domain of GAL4. The potentially interacting protein of interest, the prey, is expressed as a fusion to the GAL4 transactivation domain. Neither the bait nor the prey activates the reporter genes integrated into the yeast genome when expressed alone. Co-expression of bait and prey leads to interaction and thus reconstitutes a functional transcription factor that is situated upstream of the reporter genes. This two-hybrid transcription factor activates the reporter genes, whose output is measured, for instance, as growth of yeast clones on a selective medium or as blue coloration in a β-galactosidase assay (Figure 1.3) [8]. Current development using the yeast two-hybrid system is reviewed by Suter, Auerbach and Stagljar [8] and by Lentze and Auerbach [16].

The yeast three hybrid system is an adaptation of the yeast two-hybrid system used to detect small molecule–protein interactions instead of protein–protein interactions. For that purpose a protein with a known affinity for a defined small molecule is fused to the DNA-binding domain of a transcription factor. In most cases, the bait is dihydrofolate reductase (DHFR) which binds to the small molecule methotrexate. Methotrexate is covalently bound via a linker to the small molecule of interest. On the other hand, a cDNA library is fused to

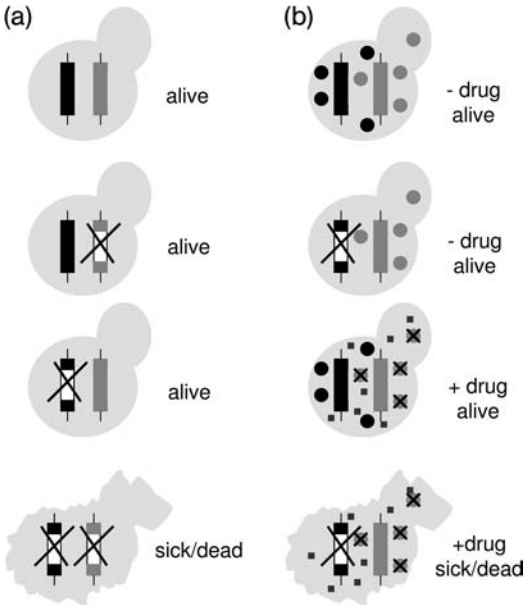


Figure 1.2 Synthetic lethal screen. (a) Inactivation (deletion) of two genes (black and dark gray) in redundant pathways leads to loss of viability, whereas inactivation of either one gene has no effect; (b) a protein product of a gene (dark gray circles) is inactivated by treatment with a drug (dark squares). The protein product of a second gene (black circles) is by itself not essential but prevents loss of viability in the presence of the drug. Deletion of the second gene leads therefore to hypersensitivity to a dose of the drug that is not lethal in a wild-type cell [8]. (Reprinted with permission from [8].)

Box 1.6

Transcription Factors: Transcription factors are proteins supporting the binding of RNA-polymerase to initiate transcription. In order to produce their effects, transcription factors require the ability to bind to DNA and then to influence transcription positively or negatively. For that purpose a prototypic transcription factor has a DNA binding domain (DBD), a signal sensing domain (SSD) and a transactivation domain (TAD).

Transcription is RNA synthesis based on a DNA matrix. Transcription together with translation are the key steps in getting a protein product from a gene [7, 17, 18]. For mechanistic details the reader is referred to the textbooks *Genes VI* by B. Lewin [7] and *Biology* by N. A. Campbell and J. B. Reece [17].

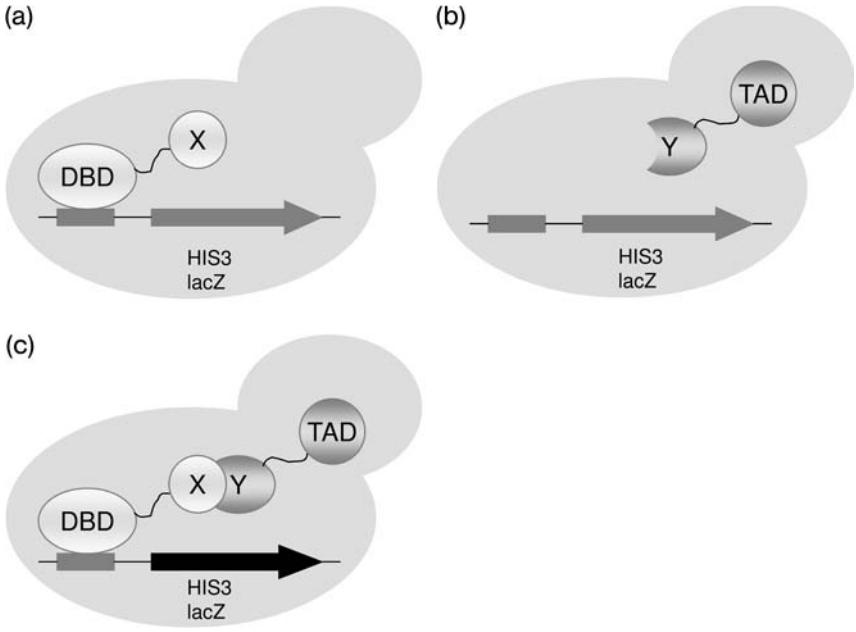


Figure 1.3 The yeast two-hybrid system. (a) To construct a bait in the yeast two-hybrid system, a protein of interest X is fused to the DNA binding domain (DBD) of a transcription factor. When expressed on its own in yeast, the bait will not activate transcription since it lacks a transcriptional activation domain (TAD). (b) Likewise, a prey is constructed by fusing a second protein of interest Y to the TAD of a transcription factor. The TAD–Y fusion is unable to activate transcription on its own, since it is not situated near a promoter. (c) Co-expression of the interacting DBD–X and TAD–Y fusion proteins reconstitutes a functional transcription factor situated at a promoter. Consequently, the reporter gene located downstream of the reporter is activated, and the protein–protein interaction between the proteins X and Y is measured using the product of the reporter gene. Common reporter genes in yeast two-hybrid systems include auxotrophic growth markers, such as the *HIS3* or *ADE2* genes, or a color marker, such as *lacZ* [8]. (Reprinted with permission from [8].)

the activation domain of the transcription factor. Only when the methotrexate–small molecule hybrid binds with the methotrexate part to the dihydrofolate reductase and with the small molecule to its binding partner protein (expressed via the cDNA library) is expression of the reporter gene achieved (Figure 1.4) [8].

For summaries of yeast-based technologies including recent research activities the reader is referred to the reviews by Suter *et al.* [8] and Luesch [13]. All in all these global approaches are only useful in indirectly determining small molecule targets but at least they give the molecular basis for bioactivity [13].

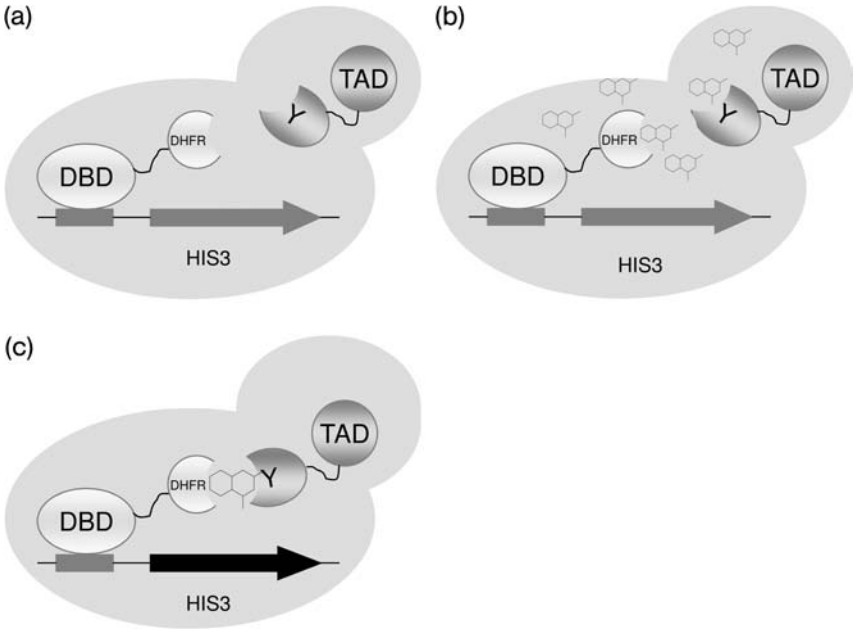


Figure 1.4 The yeast three-hybrid system. (a) The bait is constructed by fusing a DNA binding domain (DBD) to a small molecule binding protein such as dihyrofolate reductase (DHFR). Simultaneously, each yeast cell expresses a particular activation domain (TAD)-fused prey from a cDNA library. (b) A hybrid compound consisting of a small molecule covalently linked to methotrexate is added, which crosses the yeast cell membrane and binds to the DBD-DHFR bait via its methotrexate part. In this way, the other part of the small molecule is displayed by the bait. (c) If the TAD-prey binds to the small molecule displayed from the scaffold bait, a functional transcription factor is reconstituted via the small molecule–protein interaction, resulting in activation of the downstream reporter gene [8]. (Reprinted with permission from [8].)

1.3 The Chemical Approach

Hitherto, most of the yeast-based approaches have been carried out using commercial compound collections with the aim being to explain the mode of action of small molecules or known drugs and to find new modulators of bioactivity. Among the commercial compound collections used are substances from ChemBridge Cooperation (www.chembridge.com) [20], chemical libraries from the National Cancer Institute (NCI: www.cancer.gov) [21] as well as from MicroSource Discovery Inc. (www.msdiscovery.com) and FDA approved drugs (U. S. Food and Drug Administration: http://www.accessdata.fda.gov/scripts/cder/drugsatfda/) [22]. These compound collections contain diverse substances of general interest amongst other anticancer, antifungal or psychotic therapeutics as well as natural products and bioactive compounds.

1.4
Chemical Biological Research/Evaluation – Chemogenomic Profiling: Elucidating the Mode of Action of Small Molecules

1.4.1
Assay Principle

The yeast-based chemical genomic approach described below makes use of the barcoded *Saccharomyces cerevisiae* yeast deletion strain collection (see Section 1.4.2). In this forward genetic screen (Box 1.2) pools of yeast deletion strains are grown in the presence or absence of a small molecule. In the deletion strain collection each gene deletion is flanked by two sequences that contain unique barcodes. After DNA-extraction the barcodes are amplified by polymerase chain reaction (PCR) (Box 1.7) and then hybridized to a barcode microarray (Box 1.8). Absence of a hybridization signal in the drug-treated sample reveals sensitivity of the corresponding deletion strain to the drug (Figure 1.5) [8, 15].

The steps involved are:

1. Yeast deletion strain pool construction,
2. Pooled growth in absence or presence of the drug,
3. Purification of genomic DNA,
4. PCR amplification of barcodes,
5. Array hybridization and scanning,
6. Analysis of the microarray results.

A detailed protocol of all necessary steps is published by Giaever *et al*. in *Nature Protocols* [15].

Box 1.7

Polymerase Chain Reaction: PCR describes a technique in which cycles of denaturation, annealing with primer, and extension with DNA polymerase, are used to amplify the number of copies of target DNA sequence by $> 10^6$ times [7].

Box 1.8

DNA Microarray: A DNA microarray is a collection of microscopic DNA spots (defined as features), commonly representing single genes or transcripts (in the present case: oligonucleotides), arrayed on a solid surface by covalent attachment to chemically suitable matrices or directly synthesized on them. DNA microarrays use DNA as part of their detection system. Qualitative or quantitative measurements with DNA microarrays use the selective nature of DNA–DNA or DNA–RNA hybridization under high-stringency conditions and fluorophore-based detection [24]. For additional information about DNA microarrays see Chapter 2.

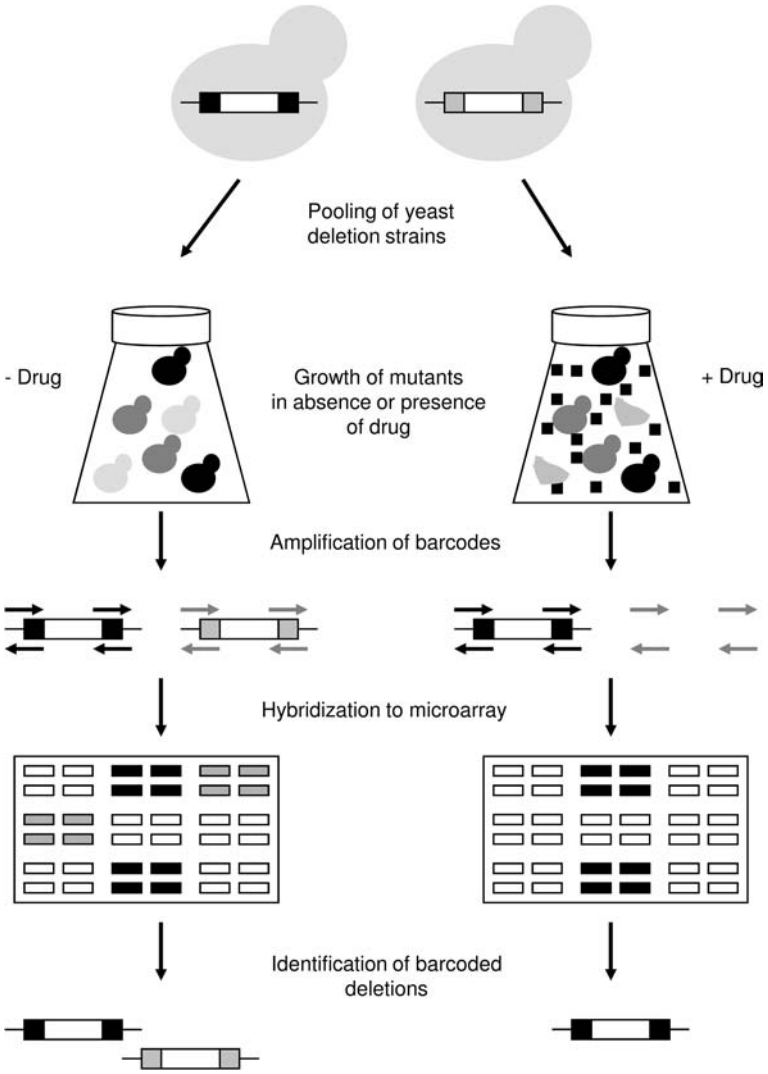


Figure 1.5 Chemogenomic profiling using the yeast deletion strain collection. The heterozygous or homozygous yeast deletion strains are grown in pools in the presence or absence of a drug (black squares). Unique barcodes enclose each gene deletion (black and gray). After DNA-extraction the barcodes are amplified by PCR and then hybridized to a barcode microarray. If the hybridization signal in the drug-treated sample is absent the corresponding deletion strain reveals sensitivity to the drug. The scheme shows the hybridization of treated and untreated samples to different arrays [8]. (Reprinted with permission from [8].)

1.4.2 The Yeast Deletion Strain Collection

The yeast deletion strain collection is a near-complete (96%) collection of gene-deletion mutants in *Saccharomyces cerevisiae* facilitating the systematic analysis of gene function in yeast. It contains only one strain per gene and allows individual deletion mutants to be rapidly identified. To easily identify individual deletion strains unique DNA barcodes or tags are included in each strain. The yeast deletion strain collection can be obtained in 96-well microtiter plates via OpenBiosystems (part nos. YSC1056 and YSC1055) [15].

1.4.2.1 Homozygous Deletion Strains

The homozygous deletion strains (Box 1.5) can be screened in a condition of interest to deduce functional information about the deleted genes or to identify the mechanism of action of small molecules. The gene deletion in the homozygous strain leads to total omission of the corresponding protein (gene product). As depicted in Figure 1.6, the homozygous profiling works as follows: In the wild type strain the gene of interest is not deleted and the yeast is growing. Gene deletion leads to total omission of the gene product but the yeast is still alive. When adding a small molecule to the wild type strain, the strain is still growing whereas the deletion strain displays sensitivity towards the small molecule (reduced growth). If the protein function of the gene product is known this screen will provide information about the mode of action of the small molecule [15].

1.4.2.2 Heterozygous Deletion Strains

The heterozygous deletion strains (Box 1.5) are often used for genomic profiling of sensitivity to small molecules via induced haploinsufficiency [8, 19]. A cell is haploinsufficient when the dosage of a gene encoding a small molecule target protein is lowered from two copies to one copy. Thus, the amount of target protein is reduced conferring hypersensitivity to the applied compound (Figure 1.6).

The readouts from haploinsufficiency screens are distinct from chemical genetic screens with complete homozygous null mutations. Screens with homozygous deletions identify genes that are important for survival in the presence of a compound, the haploinsufficiency profiling is based on a dosage effect and is expected to reveal the primary target for the small molecule. Another advantage of the heterozygous profiling is that fitness defects in essential genes can be detected, which is not possible with homozygous diploid or haploid strains. When the drug target is not a protein (for example DNA damage) or when the direct effects of the drug are masked by a redundant protein function the haploinsufficiency approach is limited. Furthermore, when the potential drug target is in excess (even after reduction to one gene copy) haploinsufficiency cannot be observed [8].

The yeast deletion collection consists of four different sets of strains: homozygous diploids, heterozygous diploids, MATa haploids and MATα haploids (Box 1.5). Each set of strains uses the same set of barcodes. For example, the deletion of gene X will use the same two barcodes in all four strain backgrounds. The

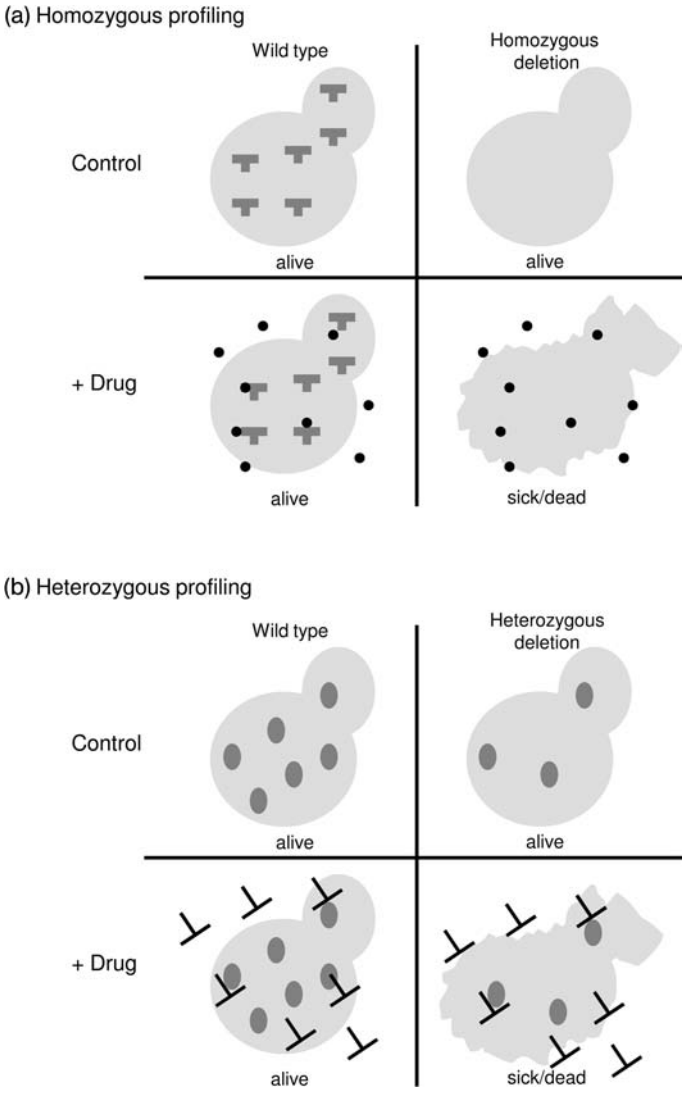


Figure 1.6 Homozygous and heterozygous profiling. (a) Screening the homozygous deletion collection in a condition of interest will identify genes that are needed for growth in that condition and can provide information about a cytotoxic compound's mode of action. Here a drug is shown in black, and a protein needed to protect against that drug is shown in gray. Deletion of the gene encoding this protein results in sensitivity to the drug. In cases where the drugs effects are known, this will give functional information about the deleted gene. In cases where the protein's function is known, this will give information about the drugs mode of action.
(b) Screening the heterozygous deletion collection can identify drug targets for genes whose homozygous deletion causes a growth defect or death. This method exploits the fact that the copy number of a gene is related to its expression level. Here, a drug with a specific protein target is shown in black, and the corresponding protein target is shown in gray. Adding an intermediate concentration of a drug that deactivates the protein target will inactivate most target function in the heterozygote but not in strains with two functional gene copies. This will result in a growth defect for the strain that carries a heterozygous deletion of the gene encoding the drugs target. This phenotype mimics the growth defect caused by the corresponding homozygous deletion [15]. (Reprinted with permission from [15].)

only case where overlap does not occur is for the heterozygous deletion of an essential gene, which will be present only in the heterozygous collection. Because of this barcode overlap, the homozygous deletion collection and the heterozygous deletion collection must be pooled and grown separately to ensure that each barcode corresponds to a unique strain within the pool [15].

1.4.3 Advantages and Disadvantages

The yeast-based chemical genomic profiling features several benefits. The DNA barcodes enable a fast and simultaneous screening of multiple yeast strains in pooled cultures. This means that the amount of small molecule is advantageously reduced. The heterozygous deletion strain collection allows the study of ~1000 essential genes. Moreover this deletion strain collection can be used for haploinsufficiency screens including all the advantages of this method (see Section 1.4.2).

A disadvantage of the described method may be that the expensive microarray technology has to be available.

General instruction for working with yeast is given in *Methods in Yeast Genetics* [11].

1.4.4 Case Studies

In an initial study Giaever and coworkers [19] screened a pooled culture of 233 heterozygous deletion strains in the presence of 0.5 μg ml^{-1} tunicamycin. Three of the 233 strains grew less abundant in the pool. Those three strains were termed as drug-induced haploinsufficient. The three strains were alg7/ALG7 (Figure 1.7), ymr007w/YMR007w (both of which diminished in the pool at a rate approximately five times that of any other strain) and ymr266w/YMR266w (which diminished in the pool at an intermediate rate). The ALG7 locus (required for glycosylation) encodes the known target of tunicamycin which is GlcNAc phosphotransferase (GPT); YMR007W encodes a 126 amino acid protein of unknown function and YMR266W encodes a 953 amino acid protein with homology to the multifacilitator superfamily (MFS). The newly identified YMR007w and YMR266w genes were further characterized by studying the individual growth of the heterozygous and homozygous strains (Box 1.5 and Section 1.4.2). The heterozygous strains showed drug-induced haploinsufficiency at growth rates consistent with their behavior in the pool. The homozygous strains ymr007w/ymr007w and mr266w/ymr266w exhibited increased sensitivity to tunicamycin, with the ymr007w/ymr007w strain being the more sensitive of the two.

As the two homozygous strains were also drug-sensitive, these loci were ruled out as possible drug targets of tunicamycin because the proteins they encode are absent in these homozygotes. As a control experiment all strains were tested for sensitivity towards the unrelated drugs hygromycin B and fluconazole. The

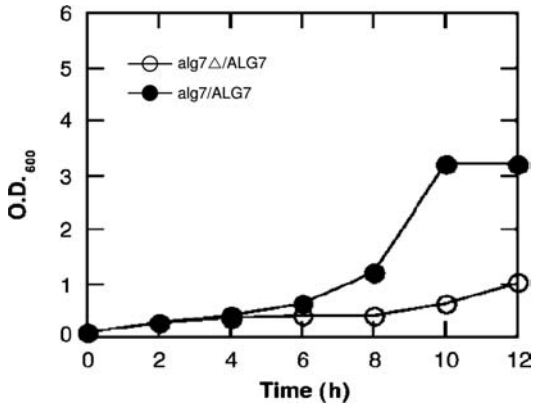


Figure 1.7 Tunicamycin sensitivity of the alg7/ALG7. Growth of alg7Δ/ALG7 deletion strain and alg7/ALG7 wild-type strain (OD_{600}) as a function of time. Drug-induced haploinsufficiency is seen because a concentration of 0.5 μg ml^{-1} tunicamycin reveals a difference in drug response between the two strains [19]. (Reprinted with permission from [19].)

reader has to be aware that a validation of the two newly identified genes and the corresponding gene products has to follow.

The success of this initial approach using the 233 heterozygous deletion strains inspired efforts to scale to a genome-wide level [23, 25]. In this continuative approach Giaever and coworkers screened 10 diverse compounds against the complete collection of ~6000 heterozygous yeast deletion strains (haploinsufficiency screen) [23]. The compounds include anticancer (methotrexate, 5-fluorouracil (5-FU), cisplatin) and antifungal agents (miconazole, itraconazole, fluconazole, fenpropimorph), statins (atorvastatin, lovastatin) and dyclonine. The complete collection of bar-coded heterozygous deletion strains pooled in a single culture allowed screening of all strains in parallel for each of the 10 compounds. Relative sensitivities were quantified using high-density oligonucleotide arrays (Box 1.5 and Box 1.8) carrying the barcode complements. Statistical treatment of the resulting signal intensity data resulted in a ranking in order of sensitivity on a gene-by-gene basis. Below, the results of this genome-wide screen will be exemplified for methotrexate.

The known target of methotrexate is dihydrofolate reductase which is encoded by the gene *DFR1*. The corresponding heterozygous strain was identified as a highly sensitive strain at the optimal concentration of 250 μM methotrexate (Figure 1.8).

Four other strains were identified as significantly sensitive in eight of nine replicate experiments: The *FOL1*- and *FOL2*- as well as the *YBT1*- and *YOR072w*-heterozygous deletion strains.

The *FOL1* and *FOL2* genes, act upstream of *DFR1* and are required for biosynthesis of folic acid in yeast. As the readout of a haploinsufficiency screen is based

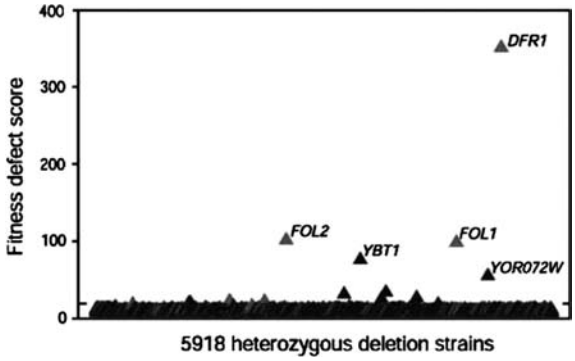


Figure 1.8 A genome-wide readout of heterozygous strain sensitivity profiled at 250 μM methotrexate. The FD (fitness defect) score is plotted along the *y* axis as a function of the 5918 heterozygous yeast deletion strains ordered by gene name. The greater the FD score, the more sensitive the strain. Strains above the dashed line are considered significantly sensitive. Only those genes that scored significantly sensitive in eight out of nine replicate experiments are labeled [23]. (Reprinted with permission from [23].)

on growth inhibition, identified genes (and the corresponding gene products) are expected to interact directly with the applied small molecule (in this case methotrexate).

The present example states that an exception to this may be gene products that are rate-limiting in the drug target pathway. The product of the *FOL2* gene catalyzes the known rate-limiting step in the biosynthesis of a variety of pterins (pterins are needed for the biosynthesis of folic acid). Although the *FOL1* product is not known to be rate-limiting in this pathway, it is possible that under these conditions it may be. Because the haploinsufficiency screen does not distinguish between gene products that directly interact with a compound and those that become rate-limiting in the presence of a compound, however, it is also possible that *FOL1* and *FOL2* gene products bind directly to methotrexate. The methotrexate results also revealed that the screen can identify genes involved in compound availability. The *YBT1* and *YOR072w* heterozygous deletion strains are highly sensitive to methotrexate, and both nonessential gene products may be involved in small molecule transport.

The human homolog of *YBT1* encodes the known methotrexate transporter and up-regulation of this gene in human cancer cells causes methotrexate resistance. Although the function of the nonessential gene *YOR072w* is unknown, the gene encodes a predicted transmembrane domain, indicating that it may play a role in methotrexate transport. As in the first example, validation of the four newly identified genes and the corresponding gene products has to follow.

At the same time Lum, Armour and coworkers [25] published a similar approach using a genome-wide pool of tagged heterozygous yeast deletion strains to assess the cellular effects of 78 compounds in *Saccharomyces cerevisiae*. They came to the

same conclusion as Giaever and associates [23], that the use of barcoded deletion strains is a powerful approach for analyzing the mode of action of small molecules.

Recent examples for the successful use of the yeast deletion strain collection are given by Chung, Yim, Lee and coworkers [21] and by the group of Langston and associates [26].

1.5 Conclusions

One general chemical biology approach is to perturb a given biological system by adding a small molecule and then analyze the observed effect.

In yeast-based genomic approaches this chemical biology approach is realized by using the advantages of the model organism *Saccharomyces cerevisiae*.

In the presented case studies the yeast-based genomic approach makes use of the barcoded *Saccharomyces cerevisiae* yeast deletion strain collection. In this forward genetic screen, pools of yeast deletion strains are grown in the presence or absence of a small molecule. In the deletion strain collection each gene deletion is flanked by two sequences that contain unique barcodes. After DNA-extraction the barcodes are amplified by PCR and then hybridized to a barcode microarray. Absence of a hybridization signal in the drug-treated sample reveals sensitivity of the corresponding deletion strain to the drug.

The studies by Giaever and coworkers [19, 23], Lum, Armour and associates [25], and the recent work by Chung, Yim, Lee and coworkers [21] and the group of Langston and associates [26] demonstrate the efficacy of this genome-wide protocol in yeast that allows the identification of those gene products that functionally interact with small molecules and result in the inhibition of cellular proliferation. They could identify previously known interactions for already known drugs and the analysis revealed novel cellular interactions, even when the relationship between a compound and its cellular target had been well established.

References

1 Hinterding, K., Alonso-Diaz, D. and Waldmann, H. (1998) *Angew. Chem. Int. Ed.*, **37**, 688–749.

2 Hübel, K., Leßmann, T. and Waldmann, H. (2008) *Chem. Soc. Rev.*, **37**, 1361–74.

3 Spring, D.R. (2005) *Chem. Soc. Rev.*, **34**, 472–82.

4 Zheng, X.S., Chan, T.-F. and Thou, H.H. (2004) *J. Chembiol.*, **11**, 609–18.

5 Stockwell, B.R. (2004) *Nature*, **432**, 846–54.

6 Guiffant, D., Tribouillard, D., Gug, F., Galons, H., Meijer, L., Blondel, M. and Bach, S. (2007) *Biotechnol. J.*, **2**, 68–75.

7 Lewin, B. (1997) *Genes VI*, Oxford University Press, Oxford, United Kingdom.

8 Suter, B., Auerbach, D., Stagljar, I. (2006) *BioTechniques*, **40**, 625–44.

9 Hughes, T.R. (2002) *Funct. Integr. Genomics*, **2**, 199–211.

10 Bharucha, N. and Kumar, A. (2007) *Comb. Chem. High Throughput Screen.*, **10**, 618–34.
11 Amberg, D.C., Burke, D.J. and Strathern, J.N. (2005) *Methods in Yeast Genetics*, Cold Spring Harbor Laboratory Press, USA.
12 Armour, C.D. and Lum, P.Y. (2005) *Curr. Opin. Chem. Biol.*, **9**, 20–4.
13 Luesch, H. (2006) *Mol. BioSyst.*, **2**, 609–20.
14 Fields, S. and Song, O. (1989) *Nature*, **340**, 245–6.
15 Pierce, S.E., Davis, R.W., Nislow, C. and Giaever, G. (2007) *Nat. Protoc.*, **2**, 2958–74.
16 Lentze, N. and Auerbach, D. (2008) *Expert Opin. Ther. Targets*, **12**, 505–15.
17 Campbell, N.A. and Reece, J.B. (2004) *Biology*, Benjamin Cummings, Pearson Education, USA.
18 Latchman, D.S. (1997) *Int. J. Biochem. Cell Biol.*, **29**, 1305–12.
19 Giaever, G., Shoemaker, D.D., Jones, T.W., Liang, H., Winzeler, E.A., Astromoff, A. and Davis, R.W. (1999) *Nat. Genet.*, **21**, 278–83.
20 Duncan, M., Ho, D.G., Huang, J., Jung, M.E. and Payne, G.S. (2007) *Proc. Nat. Acad. Sci.*, **104**, 6235–40.
21 Chung, K.S., Yim, N.H., Lee, S.H., Choi, S.J., Hur, K.S., Hoe, K.L., Kim, D.U., Goehle, S., Kim, H.B., Song, K.B., Yoo, H.S., Bae, K.H., Simon, J. and Won, M. (2008) *Invest. New Drug*, **26**, 299–307.
22 Perlstein, E.O., Ruderfer, D.M., Roberts, D.C., Schreiber, S.L. and Kruglyak, L. (2007) *Nat. Genet.*, **39**, 496–502.
23 Giaever, G., Flaherty, P., Kumm, J., Proctor, M., Nislow, C., Jaramillo, D.F., Chu, A.M., Jordan, M.I., Arkin, A.P. and Davis, R.W. (2004) *Proc. Nat. Acad. Sci.*, **101**, 793–8.
24 Marchionni, L., Wilson, R.F., Wolff, A.C., Marinopoulos, S., Parmigiani, G., Bass, E.B. and Goodman, S.N. (2008) *Ann. Intern. Med.*, **148**, 358–69.
25 Lum, P.Y., Armour, C.D., Stepaniants, S.B., Cavet, G., Wolf, M.K., Butler, J.S., Hinshaw, J.C., Garnier, P., Prestwich, G.D., Leonardson, A., Garrett-Engele, P., Rush, C.M., Bard, M., Schimmack, G., Phillips, J.W., Roberts, C.J. and Shoemaker, D.D. (2004) *Cell*, **116**, 121–37.
26 Doostzadeh, J., Davis, R.W., Giaever, G.N., Nislow, C. and Langston, J.W. (2007) *Toxicol. Sci.*, **95**, 182–7.