Chemical Biology
Edited by Stuart L. Schreiber, Tarun M. Kupoor,and Gunther Wess

4
Controlling Protein–Protein Interactions

4.1
Chemical Complementation: Bringing the Power of Genetics to Chemistry

Pamela Peralta-Yahya and Virginia W. Cornish

Outlook

Genetics in many ways is the underpinning of modern cell biology, having provided a straightforward experimental approach to identify the proteins involved in a given biological pathway. As practised, however, genetics leaves us with a picture of the cell composed largely of proteins. The roles of other molecules, such as phosphoinositides or siRNAs, have long been overlooked. With growing interest in developing a complete description of a living cell and with the backdrop of the genome sequencing projects, the question would seem to be how to extend the ease of genetics to these other classes of molecules. With a complete palette, it would then be possible to fully harness the powerful synthetic and functional capabilities of the cell for chemistry beyond that naturally carried out by the cell (Fig. 4.1-1). Here we consider a particular genetic assay, the yeast two-hybrid assay, in light of these challenges.

4.1.1
Introduction

The two-hybrid assay, which detects protein–protein interactions as reconstitution of a transcriptional activator, provides a general, high-throughput assay for cloning any protein on the basis of its interaction with another protein. Introduced only in 1989, the two-hybrid assay has proven so robust that today roughly half of the known protein–protein interactions are determined in part using the two-hybrid assay. In this, chapter we look at more recent efforts to extend this powerful genetic assay to read-out the other important molecules in

Chemical Biology. From Small Molecules to System Biology and Drug Design.
Edited by Stuart L. Schreiber, Tarun M. Kapoor, and Günther Wess

ISBN: 978-3-527-31150-7

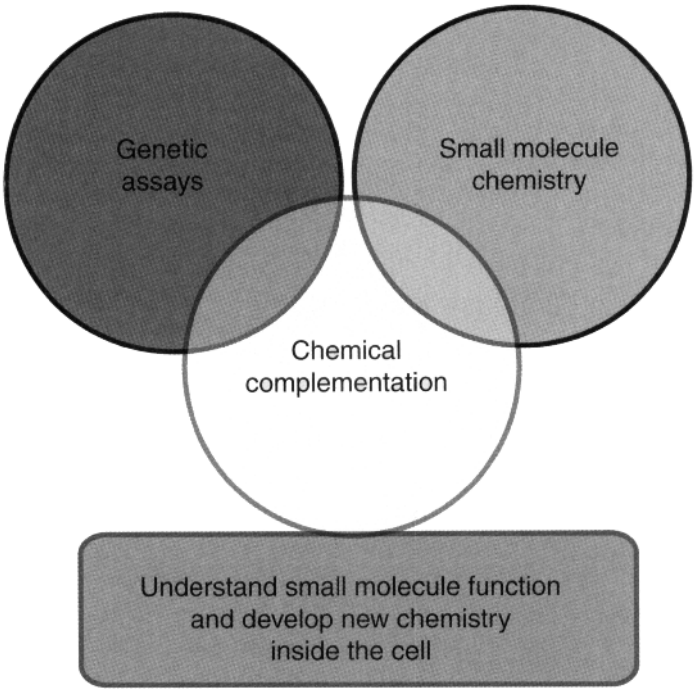


Fig. 4.1-1 Chemical Complementation combines the power of genetic assays and small molecule chemistry to understand small molecule function and develop new chemistry inside the cell.

the cell, such as nucleic acids and small molecules. We also consider the possibilities for exploiting the two-hybrid assay for chemical discovery–extending the power of genetics to chemistry not naturally carried out in the cell.

The two-hybrid assay works by detecting protein–protein interactions as reconstitution of a transcriptional activator, a natural eukaryotic transcription factor, and as activation of a reporter gene. One protein is fused to the DNA-binding domain (DBD) of the transcriptional activator, and the other protein is fused to the activation domain (AD). If the two proteins bind to one another, they effectively dimerize and hence reconstitute the transcriptional activator (Fig. 4.1-2). In practice, this assay is used not just to test a single protein–protein interaction, but to test all of the proteins expressed in a given organism or cell line for binding to the protein of interest. A library of AD-fusion proteins, encoding all ca 10^4 different proteins, is transformed *en masse* into an appropriate two-hybrid selection strain containing the DBD-protein fusion of interest. Only cells expressing an AD-protein fusion that binds to the DBD-protein fusion will then survive under the appropriate reporter gene selection conditions. The assay is general because the transcription-based selection works for any protein–protein interaction. Therefore, while

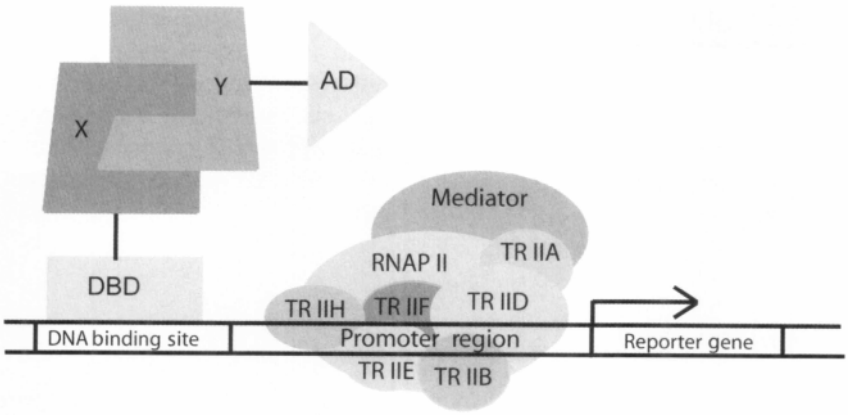


Fig. 4.1-2 In the yeast two-hybrid system, dimerization of fusion proteins X-DNA-binding domain and Y-activation domain reconstitutes the transcriptional activator. The reconstituted transcriptional activator recruits the transcriptional machinery to the promoter region of the reporter gene, initiating its transcriptional activation.

traditional genetic assays rely on pathway-specific cell survival selections or phenotypic screens, to which not all pathways or proteins in a pathway are amenable, the two-hybrid assay can be applied to any given protein–protein interaction, since the transcription-based read-out is independent of the particular pathway being studied. The assay is high-throughput because standard molecular biology techniques allow large libraries (ca 10^5–10^7 in yeast) to be tested simultaneously, where only the cells expressing an interacting protein pair survive.

The other strength of the two-hybrid assay is the ease with which it can be carried out using modern methods in molecular biology. At the end of a two-hybrid selection, the interacting proteins can be read-out simply by extracting the DNA encoding the AD-fusion proteins from the surviving cells and by sequencing the DNA. As a proof of the power of this approach, the two-hybrid assay is now essential to any effort to clone proteins along a given biological pathway. Moreover, the fortuitous development of the two-hybrid assay concurrent with genome sequence projects, enables the construction of exact cDNA-AD libraries based on this data, thus facilitating protein identity to be readily extracted from a random DNA library. The high-throughput nature of the two-hybrid assay even allows protein interaction studies to be carried out on a genome-wide scale. For example, analyzing all ca 6000 proteins expressed in yeast for binding to one another by testing all 6000 DNA-binding protein fusions to their 6000 AD counterparts.

As with the field of genetics as a whole, the two-hybrid assay is biased toward proteins. As variations of this assay, which can detect DNA, RNA, and small molecule binding, are now developed, it is exciting to imagine

the potential for basic science discovery for the roles of these molecules in the cell. Furthermore, these so-called *n*-hybrid assays extend these powerful transcription-based genetic assays to chemistry not naturally carried out in the cell. This extension should allow these genetic assays to be used not only for the discovery of biological pathways but also for new chemistry, including drug discovery and the directed evolution of molecules with new functional properties.

4.1.2 History/Development

Since the conception of the two-hybrid assay to detect protein–protein interactions *in vivo* at the end of the 1980s, key modifications to this assay have expanded its scope to detect DNA–, RNA–, and small molecule–protein interactions in so-called *n*-hybrid assays. More recently, "*n*-hybrid" assays have also been used to detect enzyme catalysis, where enzyme activity is linked to cell survival via transcription of a reporter gene. Here we look at the initial publications that moved the two-hybrid assay into each of these new directions.

4.1.2.1 Protein–Protein Interactions

In 1989 Fields and Song introduced the "Yeast Two-Hybrid Assay" which provides a straightforward method for detecting protein–protein interactions *in vivo* [1]. Until the development of the two-hybrid methodology, protein-binding interactions had been detected using traditional biochemical techniques such as coimmunoprecipitation, affinity chromatography, and photoaffinity labeling [2]. There are three significant advantages to this *in vivo* assay that led almost immediately to its widespread use: first, it is technically straightforward and can be carried out rapidly; second, the sequence of the two interacting proteins can be read off directly from the DNA sequence of the plasmids encoding them; and third, it does not depend on the identity of the interacting proteins and so is general.

The two-hybrid assay was based on the observation that eukaryotic transcriptional activators can be dissected into two functionally independent domains, a DBD and a transcription AD, and that hybrid transcriptional activators can be generated by mixing and matching these two domains [3]. It appears that the DBD only needs to bring the AD into the proximity of the transcription start site, suggesting that the linkage between the DNA-binding and the AD can be manipulated without disrupting activity. Thus, the linkage in the two-hybrid assay is the noncovalent bond between the two interacting proteins.

As outlined in Fig. 4.1-3(a), the yeast two-hybrid system consists of two protein chimeras, and a reporter gene downstream from the binding site for

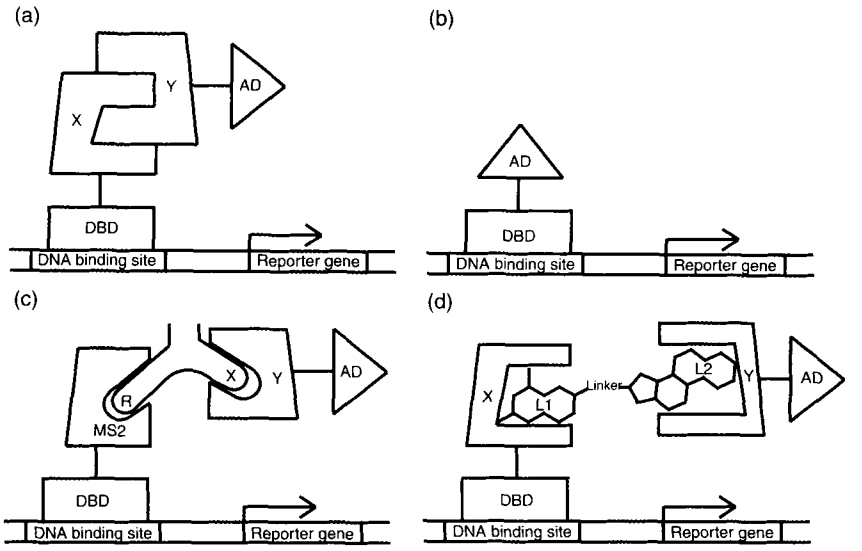


Fig. 4.1-3 Different yeast *n*-hybrid systems that have been developed to study *protein–protein*, *protein–DNA*, protein–RNA, and protein–small molecule interactions. (a) In the original version of the yeast two-hybrid system, *transcriptional* activation of the reporter gene is reconstituted by recruitment of the activation domain (AD) to the promoter region through direct interaction of protein X and Y, since protein X is fused to a DNA-binding domain (DBD) and protein Y is fused to the AD. (b) In the one-hybrid system, the AD is fused directly to the DBD. This system can be used to assay either DBDs that can bind to a specific DNA sequence or the *in vivo* binding site for a given DBD. (c) The three-hybrid system that can detect RNA–protein interactions has one more component than the yeast two-hybrid system: a hybrid RNA molecule. One half of the hybrid RNA is a known RNA (R) that binds to the MS2 coat protein (MS2) with high affinity and serves as an anchor. The other half is RNA X, whose interaction with protein Y is being tested. (d) Another version of the yeast three-hybrid system can be used to detect small *molecule–protein interactions*. Ligand L1 that interacts with protein X is covalently linked to ligand L2. Thus, if L2 interacts with Y, transcriptional activation of the reporter gene will be reconstituted.

the transcriptional activator. If the two proteins of interest (X and Y) interact, they effectively dimerize the DNA-binding protein chimera (DBD-X) and the transcription activation protein chimera (AD-Y). Dimerization of the DBD and the transcription AD helps to recruit the transcription machinery to a promoter adjacent to the binding site for the transcriptional activator, thereby activating transcription of the reporter gene.

The assay was demonstrated initially by using two yeast proteins known to be physically associated *in vivo* [1]. The yeast SNF1 protein, a serine-threonine protein kinase, was fused to the GAL4 DBD, and the SNF1 activator protein SNF4 was fused to the GAL4 transcription AD. A GAL4 binding sequence was placed upstream of a β-galactosidase reporter gene (*lacZ*). Plasmids encoding

the protein fusions and the reporter gene were introduced into the yeast. Positive protein–protein interactions lead to the increase in β-galactosidase activity inside the cell, which can be tested in a colorimetric assay using 5-bromo-4-chloro-3-indolyl β-D-galactosidase (X-gal) that turns the cells blue, or by direct measurement of enzyme activity using chlorophenol red β-D-galactopyranoside as a substrate. Control experiments established that neither the DBD and AD domains on their own nor the individual protein chimeras induced β-galactosidase synthesis above background levels. β-Galactosidase synthesis levels were increased 200-fold when the DBD-SNF1 and SNF4-AD fusion proteins were introduced together. By comparison, the direct DBD-AD fusion protein activated β-galactosidase synthesis levels 4000-fold.

It was quickly realized that the strength of the two-hybrid assay would lie not in its ability to detect a single protein–protein interaction but rather to screen an entire genome to detect novel protein–protein interactions [4–9]. For example, Murray and coworkers, as a first step toward testing their hypothesis that the cyclin-dependent kinase (CDK) Cdc20 is involved in the spindle assembly checkpoint in budding yeast, used the yeast two-hybrid assay to determine if any of the proteins known to be involved in the spindle checkpoint physically interact with Cdc20 [10]. In this experiment, haploid strains containing DBD-MAD (mitotic arrest defective) fusions were crossed with haploid strains containing AD-Cdc20 fusions. Protein–protein interactions in the resulting diploids lead to transcription activation of the *lacZ* reporter gene. As controls, haploid strains containing SNF1-AD and SNF4-DBD fusions were also mated and tested for β-galactosidase activity. The yeast two-hybrid system detected three new protein partners for Cdc20: MAD1, MAD2, and MAD3. In this experiment, the yeast two-hybrid assay was the key in rapidly and effectively identifying the new protein–protein interactions. Identification of these interactions using more traditional biochemical methods, such as coimmunoprecipitation, would have been cumbersome and time consuming since those methods require prior isolation of large quantities of all possible interacting proteins before running the assays. By facilitating the discovery of cascades of interacting proteins – in this case, the spindle assembly checkpoint – the yeast two-hybrid assay helps researchers put together entire biochemical pathways and to begin understanding how these proteins function together inside a cell.

4.1.2.2 DNA–Protein Interactions

Early on it was appreciated that, just as the yeast two-hybrid assay could be used to detect protein–protein interactions, transcriptional activators could be used directly, in a "one-hybrid" assay, to detect DNA–protein interactions (Fig. 4.1-3(b)) [11, 12]. DNA-binding proteins that bind to a given target DNA sequence could be isolated from cDNA libraries encoding all the proteins expressed in a given organism or specific cell type. Alternatively, the optimal or naturally occurring recognition sequences for a given regulatory protein

could be determined. With such an approach, Wang and Reed isolated a complementary DNA for the transcriptional activator, Olf-1, believed to be the critical switch for the coordinated expression of olfactory-specific genes [13]. To achieve this, they fused an olfactory cDNA library, consisting of 3.6 million clones, to the GAL4 transcription AD. The reporter plasmid consisted of three tandem Olf-1 binding sites upstream of a low activity promoter directing the transcriptional activation of the *HIS3* gene. The reporter plasmid requires the AD-cDNA fusion protein to bind to the Olf-1 sites and activate the transcription of the *HIS3* gene. Therefore, only cells expressing the AD-cDNA fusion are able to grow on medium lacking histidine.

4.1.2.3 **RNA–Protein Interactions**

Selecting for RNA–protein interactions is less straightforward because RNA-protein fusions cannot be generated directly *in vivo* and because routine biochemical assays that turn RNA-binding events into an amplifiable signal are not available. These difficulties were circumvented by adding a third component to the two-hybrid system to generate a "three-hybrid" assay (Fig. 4.1-3(c)) [14, 15]. The third component is a hybrid RNA molecule, in which one half is a well-studied RNA molecule that binds to a known protein with high affinity and the other half is the RNA molecule of interest whose protein-binding partner is in question. In total, the three-hybrid system consists of two protein chimeras, one RNA chimera, and a reporter gene. The hybrid RNA molecule bridges the DNA-binding and AD-fusion proteins and activates transcription of a reporter gene.

In a proof of principle experiment, Wickens and coworkers showed that the RNA three-hybrid system could detect the interactions between two well-studied protein-RNA pairs: the iron regulatory protein (IRP1) to the iron response element (IRE) RNA sequence, and the HIV transactivator (TAT) protein to the HIV transactivation response (TAR) element RNA sequence [16]. First, they constructed a bifunctional RNA containing a RNA sequence known to bind the coat protein MS2 and the RNA sequence of either IRE or TAR. Next, they fused the DNA-binding domain to the coat protein MS2, and the AD to either the IRP1 or TAT proteins. The two protein fusions and the bifunctional RNA were introduced in a yeast strain containing a reporter construct that directs activation of both a *lacZ* reporter gene and a *HIS3* reporter gene upon RNA–protein interaction. These reporter genes allow the authors to carry the assay as a colorimetric screen using the *lacZ* reporter gene and as a selection where only cells containing an interacting RNA-protein pair survive on medium lacking histidine. Furthermore, using 3-amino-1,2,3-triazole (3-AT), a competitive growth inhibitor of the enzyme encoded by the *HIS3* gene, Wickens and coworkers were able to select only cells with elevated expression levels of the *HIS3* gene, reducing the number of false positives in the HIS3 growth selection.

4.1.2.4 Small molecule–Protein Interactions

Just as a dimeric RNA molecule can be introduced to mediate the interaction between the DNA-binding and ADs, so can a dimeric small molecule [17]. In fact, well before their use in a small molecule three-hybrid assay, dimeric small molecules were used as "chemical inducers of dimerization" (CIDs) to artificially oligomerize fusion proteins *in vivo* [18]. In the yeast three-hybrid system, the union of two protein fusions and a CID reconstitute the transcription of a reporter gene (Fig. 4.1-3(d)). In 1996, Licitra and Liu built what they called a *yeast three-hybrid assay* [19]. This assay consists of two fusion proteins and a heterodimeric small molecule CID that brings these fusion proteins together to activate the transcription of a reporter gene (Fig. 4.1-3(d)).

Licitra and Liu employed two fusion proteins: the glucocorticoid receptor (GR) fused to the DBD LexA, and FK 506-binding protein (FKBP12) fused to the transcription AD B42 [19]. A heterodimeric dexamethasone (Dex)-FK506 molecule that binds to GR and FKBP12, respectively, bridges the two fusion proteins and activates the transcription of a *lacZ* reporter gene. Further, using the GR-LexA fusion protein and the Dex-FK506 molecule in their yeast three-hybrid assay, Licitra and Liu were able to isolate the FKBP isoform with the highest affinity for FK506 (FKBP12) from a Jurkat cDNA library. This experiment opened the yeast three-hybrid system as a tool for drug discovery.

4.1.2.5 Catalysis

In all the previous applications, the *n*-hybrid assay is used to detect a binding event, whether it is protein, DNA, RNA, or small molecule binding. Our laboratory and others have been interested in the idea that this powerful genetic assay could be brought to bear on a broader variety of questions. Several different approaches have now been devised for linking enzyme catalysis to reporter gene transcription using the *n*-hybrid assay. Our laboratory introduced "Chemical Complementation", which detects enzyme catalysis of bond formation or cleavage reactions on the basis of covalent coupling of two small molecule ligands *in vivo* (Fig. 4.1-4) [20]. In this assay, the enzyme is introduced as a fourth component to the small molecule yeast three-hybrid system, and the linker in the small molecule CID acts as the substrate for the enzyme. Bond formation is detected as synthesis of the CID and hence the activation of an essential reporter gene; bond cleavage is detected as cleavage of the CID and hence the repression of a toxic reporter gene. In theory, this approach should be readily extended to new chemistry, simply by synthesizing small molecule heterodimers with different chemical linkers as the enzyme substrates. Inspired by traditional genetics, our hope is to make a general complementation assay that would link enzyme catalysis of a broad range of chemical reactions to cell survival-extending genetic selections to chemistry beyond that naturally carried out in the cell.

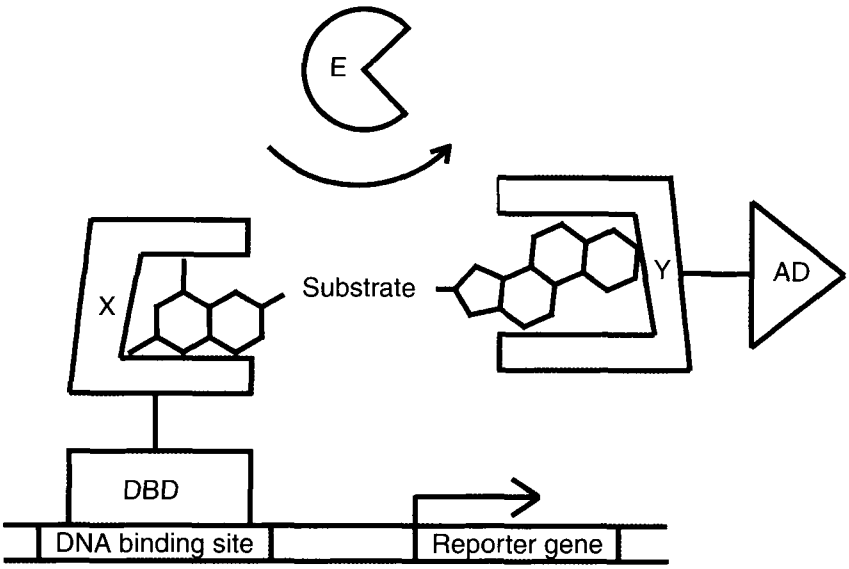


Fig. 4.1-4 Chemical Complementation. A reaction-independent complementation assay for enzyme catalysis based on the yeast three-hybrid assay. A heterodimeric small molecule bridges a DNA-binding domain–receptor fusion protein and an activation domain–receptor fusion protein, activating transcription of a downstream reporter gene *in vivo*. Enzyme catalysis of either cleavage or formation of the bond between the two small molecules can be detected as a change in transcription of the reporter gene. The assay can be applied to new chemical reactions simply by synthesizing small molecules with different substrates as linkers and adding an enzyme as a fourth component to the system.

In our initial report, we chose cephalosporin hydrolysis by the Enterobacter cloacae P99 β-lactamase (P99) as a well-studied enzyme catalyzed cleavage reaction around which to develop Chemical Complementation [20]. Cephalosporins are β-lactam antibiotics, and β-lactamases are the bacterial resistance enzymes that hydrolyze and inactivate these antibiotics. The P99 β-lactamase is well-characterized biochemically and structurally, and the synthesis of cephalosporins is well established. First, we designed a small molecule CID cephalosporin substrate, incorporating the CID ligands at the C 3′ and C7 positions of the cephem core. Using a *lacZ* reporter gene, we showed that Chemical Complementation could be used to detect β-lactamase activity using this dexamethasone-methotrexate (Dex-Mtx) heterodimer with a cephem linker (Dex-Cephem-Mtx). In the absence of enzyme, the Dex-Cephem-Mtx CID dimerizes the appropriate DBD- and AD-fusion protein activating transcription of a *lacZ* reporter gene. Expression of the P99 β-lactamase then presumably leads to cleavage of the Dex-Cephem-Mtx CID, disrupting transcription activation. We also showed that the system could distinguish the wild-type (wt) enzyme from the inactive P99:S64A variant, in which the critical

active site serine nucleophile has been mutated to an alanine, via a lacZ screen. These experiments established the feasibility of detecting enzyme catalysis using the yeast *n*-hybrid assay.

Benkovic and coworkers took a related approach in an assay they called *Quest* (*QU*erying for *E*nzyme*S* using the *T*hree-hybrid system), which detects catalysis by coupling substrate turnover to transcription of a reporter gene [21]. Here, the CID that dimerizes the transcriptional activator is a homodimer of the substrate. Enzyme catalysis of free substrate to product is detected as displacement of homodimeric CID substrate from the transcriptional activator fusion proteins. Although this approach has the advantage of using unmodified substrate, a new CID-protein pair has to be developed for each new reaction. In a more biological approach, Peterson and coworkers have developed a two-hybrid-based system to detect protein tyrosine kinase (PTK) activity [22]. This assay relies on the PTK-dependent phosphorylation of a tyrosine residue present in a peptide that has been fused to the DBD. The phosphorylated tyrosine is then bound by the phosphotyrosine-binding protein fused to the AD, leading to transcriptional activation of the reporter gene. While limited to peptide substrates, this approach has the advantage that it does not require chemical synthesis, making it more accessible to biologists.

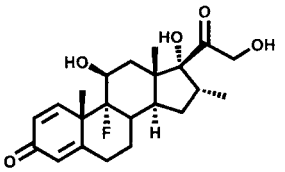
4.1.3 General Considerations

Whether being applied as in the original two-hybrid assay to detect protein–protein interactions or in the related *n*-hybrid assays to detect protein–DNA, RNA, or small molecule interactions, the basic components of the *n*-hybrid assay remain the same. Thus, while we focus in this section on the small molecule three-hybrid assay because it is in this that our laboratory specializes, this section could also be used as a technical introduction to any of the other *n*-hybrid systems. The real strength of the *n*-hybrid assays lies in how straightforward they are to implement in the laboratory with basic knowledge of *Escherichia coli* and *Saccharomyces cerevisiae* molecular biology. Moreover, the commercial availability of the components of the two-hybrid system permits any laboratory to rapidly implement the system. Finally, laboratories without prior experience working with *S. cerevisiae* should not be deterred from carrying out *n*-hybrid assays, as molecular biology techniques for this organism are similar to those for *E. coli*.

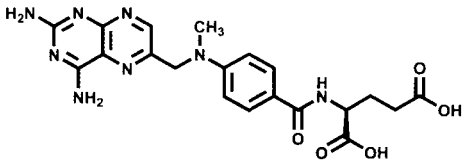
4.1.3.1 The Chemical Inducer of Dimerization (CID)

The effectiveness of any three-hybrid system depends critically on the CID used to dimerize the transcriptional activator *in vivo* [23, 24]. The subject of CIDs has been considered fully in the previous chapter by Clackson, so here we focus on the issues we have found particularly important for the use of CIDs in the

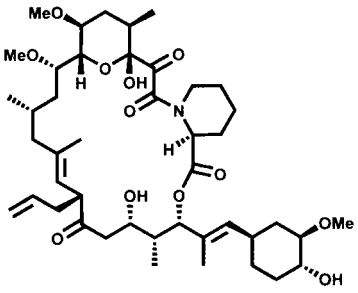
three-hybrid assay. Our presentation of these considerations is based largely on our own work with the yeast three-hybrid system and the CID ligand/receptor pairs Dex/GR, FK506/FK506 binding protein 12 (FK506/FKBP12), a synthetic analog of FK506 SLF/FK506 binding protein 12 (SLF/FKBP12), methotrexate/dihydrofolate reductase (Mtx/DHFR), O^6-benzylguanine/O^6-alkylated guanine-DNA alkyltransferase (BG/AGT), estrone/estrogen receptor (ES/ER), and biotin/streptavidin (biotin/SA) (Fig. 4.1-5) [19, 23–28].



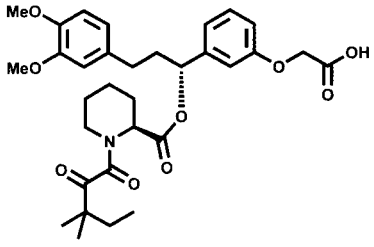
Dexamethasone



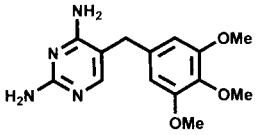
Methotrexate



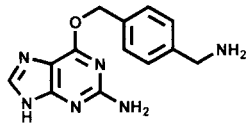
FK506



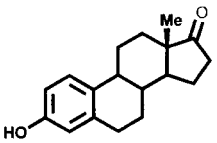
SLF



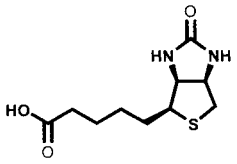
Trimethoprim



O^6-Benzylguanine



Estrone



Biotin

Fig. 4.1-5 Small molecules used to create chemical inducers of dimerization (CIDs) for the yeast three-hybrid system.

First and foremost, a successful three-hybrid system seems to require a high-affinity (low nanomolar K_D) CID pair [29]. Using the most sensitive reporter genes commercially available for the Brent LexA yeast three-hybrid system, we found that FK506-Dex, Mtx-Dex, Mtx-Mtx, and Mtx-SLF could all activate transcription, but Dex-Dex and Dex-SLF could not [25]. Second, the directionality of the system is important for a strong transcription read-out. We reported that the Dex-Mtx yeast three-hybrid system showed higher levels of transcription activation when DHFR was fused to the DBD than when fused to the AD [30]. Third, as with any CID application, the ligand/receptor pair must be considered in the context of the host cell line. For example, the Dex/GR interaction is dependent on associated heat shock proteins. Thus, the K_D of this interaction is significantly higher in *S. cerevisiae*, in which there are only homologous heat shock proteins, than in the native mammalian background. Also, this CID pair cannot be used in *E. coli*, in which there are no such homologous heat shock proteins. Finally, there are also more subtle effects. For example, for reasons we do not understand, only the *E. coli* DHFR, not the murine homolog, is functional in the Dex-Mtx yeast three-hybrid system [30].

4.1.3.2 The Genetic Assay

For a laboratory new to the three-hybrid assay, we recommend beginning with the yeast two-hybrid system, which is based on reconstitution of a eukaryotic transcriptional activator protein. Not only is this assay straightforward to practice but also all the necessary strains and plasmids are commercially available. As discussed below, however, there are potential advantages to working in *E. coli* or using a nontranscription-based assay. Several *E. coli*-based transcription assays and general protein complementation assays (PCA) have now been developed as two-hybrid assays. Notably, while the *E. coli* transcription assays have proven amenable to the introduction of small molecule CIDs, the PCAs have not.

4.1.3.2.1 The Yeast *n*-Hybrid System

There are two key versions of the yeast two-hybrid system. The GAL4 system originally introduced by Fields and Song uses the DBD and the AD of the yeast *GAL4* gene [1]. The LexA system introduced by Brent and coworkers uses the *E. coli* DBD LexA and the *E. coli* B42 AD [31]. Over time, these two systems have benefited from a number of improvements. Convenient DBD and AD vectors were developed to carry diverse bacterial drug-resistance markers, yeast origins of replication, and yeast auxotrophic markers. These technical improvements facilitate the testing of large pools of protein variants (ca 10^6) using growth selections. In addition to the basic activator system, reverse and split-hybrid systems were developed to detect the disruption of protein–protein interactions, and a transcriptional repressor-based system has been reported [32, 33]. Today components for these systems are commercially

available, including Stratagene and Clontech, which market the Gal4 system, Origene, for the LexA system, and Invitrogen, which offers versions of both systems. All of the basic features of the two-hybrid system have been covered already in several excellent reviews and the chapters on methods.

In our laboratory we have used the Brent two-hybrid system to build our Dex-Mtx yeast three-hybrid system. We favor the Brent system, which uses LexA, an *E. coli* transcription factor, and B42, an artificial activator isolated from *E. coli* genomic DNA. Both LexA and B42 are orthogonal to standard yeast genetic tools and nontoxic to the yeast cell, yet the artificial LexA–B42 transcriptional activator is on par with the strongest transcriptional activators endogenous to *S. cerevisiae* [31]. Moreover, the LexA system permits the use of the tightly regulated GAL1 promoter to drive the expression of the LexA DBD and B42 AD-protein fusions by varying the ratio of galactose and glucose in the growth medium. As reported by Lin et al., we use pMW103, a multicopy 2μ plasmid with a HIS3 maker, to encode the LexA DBD fusions and pMW102, a multicopy 2μ plasmid with a TRP1 marker, to encode the B42 AD fusions. Rather than the original EGY48 LEU2 selection strain, we chose the FY251 strain (*MATa trp*1Δ63 his3Δ200 ura3-52 leu2Δ1Gal+), which provides an additional selective marker for greater flexibility. The LEU2 or URA3 markers can then be used either for the transcription activation growth selection or introduction of additional plasmids. In this initial publication, we then used the lacZ reporter plasmid pMW112, which encodes the *lacZ* gene under control of eight tandem LexA operators. Thus, small molecule CID-induced transcription activation could be detected using standard lacZ transcription assays either on plates or in liquid culture [25]. Further optimization of the yeast three-hybrid system in our lab led us to conclude that integration of either the AD or DBD into the yeast chromosome stabilizes the transcription read-out of the reporter gene without loosing transcriptional strength, effectively reducing the number of false positives in the detection of novel ligand–receptor interactions [34].

4.1.3.2.2 E. coli Transcription Activation Assays

Widespread use of the yeast two-hybrid system led several groups to develop alternate transcription-based assays. While the yeast two-hybrid assay is quite powerful, a bacterial equivalent would increase by several orders of magnitude the number of proteins that could be tested, as the transformation efficiency and doubling rate of *E. coli* are significantly greater than those of *S. cerevisiae*. There may also be applications where it is advantageous to test a eukaryotic protein in a prokaryotic environment, in which many pathways are not conserved. The yeast two-hybrid assay cannot, however, be transferred directly to bacteria since the components of the transcription machinery and the mechanism of transcriptional activation differ significantly between bacteria and yeast.

The first bacterial repressor assay was developed in 1990 by Sauer and coworkers, who adapted a bacterial λ transcriptional repressor system to

read-out the GCN4-leucine zipper fusion [35]. The transcriptional repressor λcI controls the lytic/lysogenic pathway in bacteriophage λ. As a dimer, λcI is bound to the λ operator and prevents the expression of genes involved in the lytic pathway, allowing integration of the λ DNA into the bacterial chromosome. Taking advantage of the λcI dimerization requirement, Sauer and coworkers fused the DNA-binding domain of two λcI to a GCN4 leucine zipper dimerization motive to restore a functional hybrid repressor.

Seven years later, Hochschild and coworkers designed a bacterial two-hybrid activation system based on the transcription mechanism of *E. coli* RNA polymerase (RNAP) [36]. This assay is based on their observation that binding of the C-terminus of the α subunit of the RNAP (α-CTD) to an upstream element leads to transcription activation of a downstream gene. To create a bacterial two-hybrid system, the authors replaced the α-CTD with the C-terminus of the transcriptional repressor λcI (λcI-CTD), generating a αλcI chimera. Binding of the transcriptional repressor λcI to the λ operon, leads to recruitment of RNAP via the αλcI chimera, which in turn directs transcription activation of a reporter gene downstream of the λ operon. By simply replacing the αλcI chimera with arbitrary protein–protein interactions, they created a bacterial two-hybrid activation system. This technology was successfully applied to detect two interacting yeast proteins, Gal4 and Gal11, fused to λcI and α-NTD (N-terminus of the alpha subunit of the RNAP) respectively (Fig. 4.1-6).

Our development of a successful yeast three-hybrid system and the advantages promised by an analogous system in bacteria, led us to construct a bacterial three-hybrid system from the RNAP two-hybrid system developed by Hochschild and coworkers [37]. We chose to adapt this assay because it is a transcriptional activation system, and reconstitution of transcriptional activation should be largely conformation independent. The key to converting this two-hybrid assay into a three-hybrid system was the design of a dimeric ligand that could bridge λcI and α-NTD through the receptors of the ligand. For the bridging small molecule, we chose to prepare a heterodimer of Mtx and

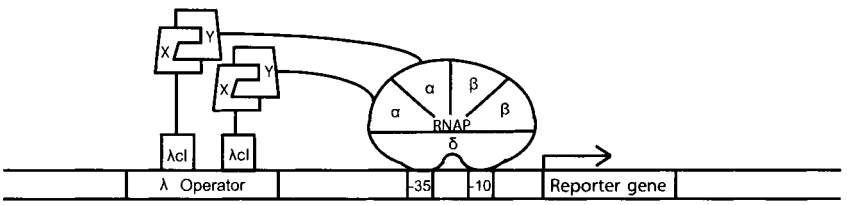


Fig. 4.1-6 The bacterial two-hybrid system developed by Hochschild and coworkers. The λcI repressor and the α-subunit of RNAP are fused to two arbitrary proteins, X and Y. Binding of the λcI repressor to the λ operon followed by dimerization of X and Y recruits RNAP leading to transcription activation of a downstream reporter gene.

a synthetic analogue of FK506 (SLF). We call this heterodimer Mtx-SLF. We did not pursue building a bacterial three-hybrid system based on the Mtx-Dex heterodimer previously used in our yeast three-hybrid system because the Dex/GR interactions require heat shock proteins that are absent in *E. coli*. The heterodimer Mtx-SLF gives a strong transcription read-out in the *E. coli* RNAP three-hybrid system, providing a robust platform for high-throughput assays based on protein–small molecule interactions.

4.1.3.3 Protein Complementation Assay

All of the above assays are based on transcription of a reporter gene. A different method for studying protein–protein interactions is the use of a PCA. Here an enzyme with a phenotype detectable via either a screen or a selection is divided into two nonfunctional fragments that are fused to proteins to be tested for dimerization. If the tested proteins dimerize, the two enzyme fragments are brought into close proximity leading to reconstitution of enzyme activity (Fig. 4.1-7) [38, 39]. Since PCAs are independent of the cell's transcription machinery, they can be used to detect protein interactions in any cell type or cell compartment *in vivo* or *in vitro*. Furthermore, PCAs can potentially quantify protein–protein interactions since there is a simple relationship between protein dimerization and reconstituted enzyme activity. PCAs have been developed using a variety of proteins including β-galactosidase, β-lactamase, DHFR, GFP (green fluorescent protein), and YFP (yellow fluorescent protein) [40–42].

For example, in a proof of principle paper, Michnick and coworkers showed that mDHFR can be split into two fragments that show no detectable

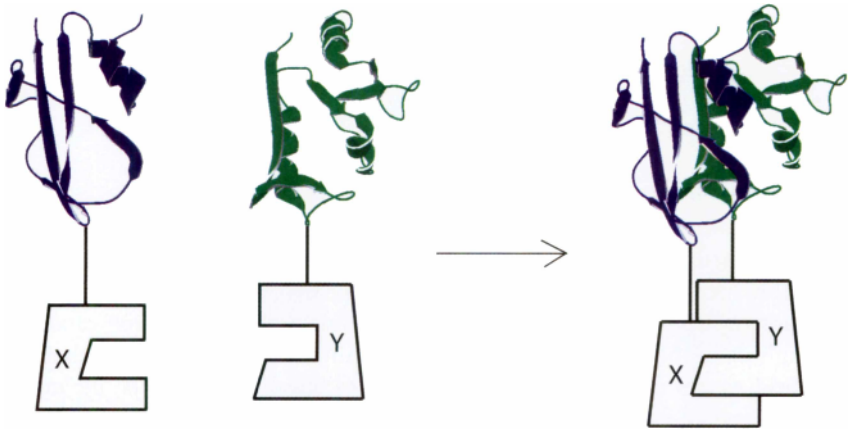


Fig. 4.1-7 Protein complementation assays. A protein that carries out a detectable function is separated into two fragments that show no detectable reconstituted enzyme activity on their own (blue and green), but can effectively reconstitute enzyme activity when fused to two interacting proteins, X and Y.

reconstituted enzyme activity on their own but can effectively reconstitute enzyme activity when fused to two interacting proteins. Bacteria expressing a functionally reassembled mDHFR can easily be selected since mDHFR activity is essential for growth of *E. coli* in the presence of trimethoprim, which selectively inhibits bacterial DHFR but not its eukaryotic counterpart mDHFR. Further, the mDHFR PCA works as a selection system in eukaryotic cells deficient in endogenous DHFR activity [43]. In a remarkable application of this system, Michnick and coworkers were able to detect a protein–protein interaction, locate the interaction to a specific cell compartment, and place the interaction in a signal transduction pathway by doing a single assay based on the DHFR PCA in mammalian cells deficient of DHFR [44]. Specifically, they examined protein interactions in the well-studied signal transduction pathway of receptor tyrosine kinase, which mediates control of initiation of translation in eukaryotes. From 35 interactions tested, the DHFR PCA selection identified 14 interacting partners that were localized to specific intracellular compartments using fluorescein-Mtx, a fluorophore in which the Mtx portion binds to the reconstituted DHFR with nanomolar affinity. The position of the protein interaction in the signal transduction pathway was determined by using three small molecule inhibitors known to act at key points of the pathway.

In view of the advantages PCAs would bring to the detection of protein–small molecule interactions, our laboratory has made some efforts to develop a small molecule PCA three-hybrid assay, though without success [45]. Specifically, we tested both the Mtx-SLF adenylate cyclase PCA and the Mtx-SLF β-lactamase PCA in *E. coli* (E. Althoff, V. Cornish, unpublished results). In addition, we tested a Dex-Mtx GFP PCA also in *E. coli* in collaboration with Regan and coworkers (E. Althoff, V. Cornish, T. Magliery, L. Regan, unpublished results). From both, a simple thermodynamic consideration and these results, we hypothesize that without the high degree of cooperativity found in the transcription-based assays, the PCAs cannot detect a three-component interaction.

4.1.3.4 **Problem Choice**

The two-hybrid assay was originally used simply for cloning proteins based on their interaction with other proteins in a given biological pathway. However, the more recent development of one- and three-hybrid assays opens the door to studying DNA, RNA, and small molecule interactions, and even catalysis. Though developed as a genetic assay for cloning, there is no reason that the *n*-hybrid assays cannot be used for a broad range of applications, including drug discovery, directed evolution, and enzymology.

It is interesting to consider how well suited the two-hybrid assay is for its original conception – the discovery of new proteins on the basis of their binding to other known proteins – particularly as this assay begins to be carried out on a genome-wide scale. An important paper that bears on this question,

in our opinion, comes from Golemis and Brent, in which they estimated that the K_D cutoff for the yeast two-hybrid assay is ca 1 μM [46]. Assuming that the proteins are being expressed at ca 1 μM concentrations, the two-hybrid assay can only detect relatively high-affinity interactions (ca $K_D = 1$ μM). Thus, while the two-hybrid assay is quite successful at identifying new interactions, it is probably not appropriate to assume that a high-throughput two-hybrid assay gives a snapshot of all interactions. In fairness, however, it should be pointed out that traditional affinity chromatography approaches are even further impaired because they rely on the natural abundance of any given protein in the cell. Extending this analysis to drug discovery using the small molecule three-hybrid assay, it is our opinion that the three-hybrid assay was long underutilized because the original systems had low sensitivity owing to the CID anchor. Recently, we have shown that our Mtx three-hybrid system has a K_D cutoff of ca 100 nM [29]. Consistent with this idea, GPC Biotech reported last year the use of the Mtx three-hybrid system for identification of protein targets of CDK inhibitors [47]. Interestingly, Hochschild and coworkers have shown that they can build additional sensitivity into their bacterial two-hybrid assay by adding cooperative interactions [48].

The *n*-hybrid assay can also be used for directed evolution. For example, Pabo and coworkers have adapted a bacterial one-hybrid assay to evolve zinc-finger variants with defined DNA-binding specificities [49]. Starting with a three zinc-finger protein that has nanomolar affinity for its DNA-binding site, the authors replaced the binding site for the third zinc finger with a new DNA sequence and then randomized the third finger to evolve a zinc-finger variant with increased affinity for the target sequence. Impressively, the evolved zinc finger showed DNA affinity within 10-fold of the wt protein, $K_D = 0.01$ nM, and a 10- to 100-fold preference for the modified over the wt DNA sequence. Given the low K_D cutoff and the fact that the *n*-hybrid assay is governed by equilibrium binding, there are two likely limitations to using this assay for directed evolution. First, the assay cannot effectively detect initial, weak binders. Second, the assay is limited in its ability to distinguish evolved variants on the basis of improvements in K_D since energy differences of only a few kilocalories per mole determine whether a molecule is bound at equilibrium. In theory, however, these limitations could be overcome by varying the concentration of the *n*-hybrid components or, again, by building in a series of tunable, cooperative interactions. Pabo and coworkers, then, choose their problem well. They began with a zinc-finger protein with two out of three zinc fingers intact. This initial binding affinity enabled them to select good binders in a single round of selection, rather than trying to improve binding affinity through multiple rounds of selection. A similar analysis suggests that the *n*-hybrid assays may be ideally suited to catalysis applications since large differences in catalytic activity are needed to significantly affect the half-life of product formation.

4.1.4 Applications

Although introduced only in 1989, the yeast two-hybrid assay has emerged as an integral tool for biology research. Two-hybrid screens now appear regularly in the biology literature. Genome-wide two-hybrid screens are even the focus of major research publications. Somewhat surprisingly then, there have been few applications of the related *n*-hybrid technologies to detect protein interactions with DNA, RNA, and small molecules, or applications beyond cloning. Here we look at more recent applications of *n*-hybrid assays with an eye for asking whether this discrepancy results from the relative power of these different *n*-hybrid assays or rather the biases of current research.

4.1.4.1 Protein–Protein Interactions

Traditional genetic assays and more recently the yeast two-hybrid assay have been primarily used to identify natural protein–protein interactions. Two-hybrid screens are now fully integrated into the biologist's toolbox and appear routinely in the published literature. Almost half of the published protein–protein interactions to date have been detected, at least in part, using the yeast two-hybrid assay [50]. Beyond these simple cloning applications, the two-hybrid assay would seem perfectly suited for genomics. For example, automation techniques were used to identify all possible protein–protein interactions in *S. cerevisiae* [51]. Every open-reading frame encoding a protein, ca 6000 in *S. cerevisiae*, was fused both to a DNA-binding domain and an AD, and the two fusion libraries were screened against one another. The major challenge in this project was how to transform all combinations of the 6000 DBD and 6000 AD fusions into yeast and then how to assay so many cells. Since a library of 10^7 is at the limit of the transformation efficiency of yeast, it is in theory achievable. Uetz and coworkers compared two approaches. In the first approach, they explicitly mated haploid mating type (MATa) cells containing 192 DBD fusions with haploid MATα cells containing the 6000 AD fusions in a spatially addressable format, such as microtiter plate, and assayed each well using a HIS3 growth selection. In the second one, MATa cells containing the 6000 DBD fusions were mated with MATα cells containing the 6000 AD fusions, and only diploids that survived in a LEU2 growth selection were arrayed and analyzed individually. Interestingly, there were significantly more "hits" in the first spatially addressable format, underscoring the importance of parameterizing new methods for high-throughput screening and the problem of distinguishing false positives and negatives in genomics. This example highlights how well suited the *n*-hybrid assays are for extracting some of the information provided by recent genome sequencing efforts.

While the two-hybrid method has been extensively used to detect natural protein–protein interactions, it should also be well suited for protein evolution. Brent and coworkers demonstrated that the two-hybrid assay can be used to

Table 4.1-1 The sequences and binding affinities of 14 different aptamers for binding to Cdk2 isolated in a yeast two-hybrid system

Aptamer	K_D(nM)	Amino acid sequence
pep1	ND[a]	ELRHRLGRAL SEDMVRGLAW GPTSHCATVP GRSDLWRVIR
pep2	64 ± 16	LVCKSYRLDW EAGALFRSLF
pep3	112 ± 17	YRWQQGVVPS NMASCSFRQ
pep4	ND	SSFSLWLLMV KSIKRAAWEL GPSSAWNTSG WASLSDFY
pep5	52 ± 3	SVRMRYGIDA FFDLGGLLHG
pep6	ND	RVKLGYSFWA QSLLRCISVG
pep7	ND	QLYAGCYLGV VIASSLSIRV
pep8	38 ± 5	YSFVHHGFFN FRVSWREMLA
pep9	ND	QQRFVFSPSW FTCAGTSDFW GPEPLFDWTR D
pep10	105 ± 10	QVWSLWALGW RWLRRYGWNM
pep11	87 ± 7	WRRMELDAEI RWVKPISPLE
pep12	ND	RPLTGRWVVW GRRHEECGLT
pep13	ND	PVCCMMYGHR TAPHSVFNVD
pep14	ND	WSPELLRAMV AFRWLLERRP

a ND – not determined.

identify peptide aptamers that inhibit Cdk2 from a library of random peptide sequences (Table 4.1-1) [52]. The 20-residue peptide library was displayed in the active site loop of *E. coli* thioredoxin (TrxA). The TrxA loop library was fused to the AD, and Cdk2 was fused to the DBD. In a single round of assay, 6×10^6 TrxA-AD transformants, a very small percentage of the 10^{27} 20mers possible, were tested for binding to LexA–Cdk2. From this assay, they isolated 66 colonies that activated transcription of both a *LEU2* and a *lacZ* reporter gene. Remarkably, these colonies converged on 14 different peptide sequences that bound Cdk2 with high affinity. Using surface plasmon resonance, the peptide aptamers were shown to bind Cdk2 with K_Ds of 30–120 nM. In kinase inhibition assays, the peptide aptamers had IC_{50}s for the Cdk2/cyclin E kinase complex of 1–100 nM. What is particularly impressive about this experiment is that nanomolar affinity ligands are being isolated in a single round of selection from a library only on the order of 10^6–10^8. Similar results have been obtained using peptide aptamers in a traditional genetic selection [53].

Given the success of this and related "aptamer" selections, it is somewhat surprising that these "aptamer" scaffolds are not more widely used. There are several potential advantages to directed evolution over traditional monoclonal antibody technology for generating selective binding proteins. Optimistically, six months are required from the start of immunization, through immortalization, and finally screening to generate a monoclonal antibody. On the other hand, if several peptide aptamer libraries were maintained for routine use, the libraries could be screened against a new target, false positives could be sorted out, and biochemical assays could validate a target in less than a month and at considerably less expense. Moreover, protein

scaffolds other than antibodies may prove more robust for use as reagents and therapeutic applications. Perhaps because monoclonal antibody technology has become so robust over the years, the momentum does not seem to be there to seriously explore replacing this technology with directed evolution. It is also interesting to compare these "aptamer" scaffolds to chemical genetic approaches for generating inhibitors for a broad array of biological targets.

4.1.4.2 DNA–Protein Interactions

Just as the yeast two-hybrid assay can be used to detect protein–protein interactions, transcriptional activators can be used directly to detect protein–DNA interactions. In truth, this type of experiment was done before the one-hybrid assay was conceptualized as such. For example, as early as 1983 a His6 → Pro Mnt variant was generated that preferentially binds a mutant Mnt operator using a transcription-based selection [54]. A plasmid encoding Mnt was mutagenized both by irradiation with UV light and by passage through a mutator strain. The mutant plasmids were then introduced into *E. coli* and selected against binding to the wt operator and for binding to the mutant operator. Because there are a variety of convenient reporter genes, the *E. coli* was engineered to link DNA recognition to cell survival in both the negative (selection against binding to the wt operator) and the positive (selection for binding to the mutant operator) directions. Binding to the wt Mnt operator was selected against by placing a tet resistance (tet^R) gene under negative control of the wt Mnt operator. If a Mnt mutant bound the wt operator, it would block synthesis of the tet^R gene, and the *E. coli* cells would die in the presence of tetracycline. Then Mnt variants with altered DNA-binding specificity were selected for on the basis of immunity to infection by a P22 phage containing a mutant Mnt operator. The mutant Mnt operator controlled synthesis of the proteins responsible for lysing the bacterial host. If a Mnt variant could bind to this mutant operator, it would turn off the lytic machinery, and the bacteria would survive phage infection. Four independent colonies were isolated from the two selections. Again, only a single round of selection was required for each step. All four colonies encoded the same His6 → Pro mutation, two by a CAC → CCC and two by a CAC → CCT mutation. Not only did these mutants bind to the mutant operator but they also did not bind efficiently to the wt operator.

More recently, Pabo and coworkers adapted a bacterial two-hybrid assay into a bacterial one-hybrid system to evolve zinc-finger variants with defined DNA-binding specificities [49]. In this assay, three tandem zinc fingers function as the DBD of this one-hybrid system and are fused to Gal11 protein, known to dimerize with Gal4, which is fused to the RNA polymerase. Binding of the three tandem zinc fingers to a specific DNA sequence upstream of the reporter gene, mobilizes the RNAP to the promoter region of the reporter gene and initiates transcription thereof (Fig. 4.1-8). This assay allows testing $\pm 10^8$ protein variants per round of selection. However, if all three zinc fingers were to be randomized simultaneously it would create 8×10^{24} protein variants (using

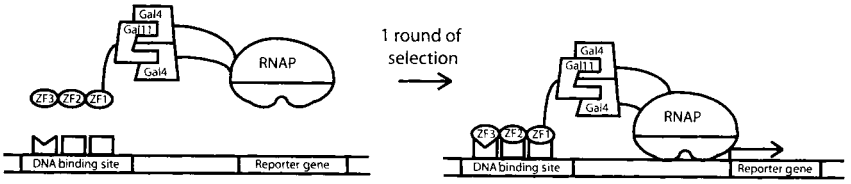


Fig. 4.1-8 Development of zinc fingers specific for a specific DNA sequence using a one-hybrid assay adapted from a bacterial two-hybrid system. Zinc fingers (ZF) 1, 2, and 3 from the Zif268 protein were fused to the Gal11 protein. The Gal4 protein, which binds Gal11 with high affinity, was fused to the α-subunit of RNAP. If ZF3 bound to the first site with high affinity, the RNAP complex would be recruited, activating transcription of a *HIS3* reporter gene. Significantly, in just one round of assay, several proteins were identified that bound specifically to the target DNA sequence.

24 codons at six amino acids per three zinc finger = $(24^6)^3$), which cannot be covered by this high-throughput method. Thus, the authors are limited to randomizing one finger at a time, while keeping the other two unchanged. We believe that conserving the high affinity of two zinc fingers for the DNA may be important for the success of Pabo and coworkers' directed evolution, because starting a directed evolution with a high-affinity protein for DNA ensures the evolution of proteins within the dynamic range of the *n*-hybrid system. For this zinc-finger evolution, they created a library of ca 10^9 variants, and identified a total of nine sequences that bound specifically to three target DNAs with a preference of 10- to100-fold for the modified over the wt DNA.

Comparing their results for the zinc-finger evolution using the bacterial hybrid system with earlier results obtained in a similar zinc-finger evolution study using phage display, Pabo and coworkers conclude that the affinity and specificity of the selected zinc fingers is superior to those obtained in earlier phage display studies. Moreover, the bacterial hybrid system is a more rapid alternative to phage display because it permits isolation of functional fingers in a single selection step instead of using multiple rounds of enrichments. Speaking to the power of this approach, Sangamo uses a modified one-hybrid assay for its selection of artificial DNA-binding proteins for commercial applications [55, 56]. The success found here raises the question of other binding interactions. One could speculate that the success here depends on starting with two known zinc fingers with high affinity for their DNA target, except that the protein "aptamer" scaffold selections described in the previous section have begun with scaffolds with no measurable affinity for their protein target.

4.1.4.3 RNA–Protein Interactions

Before the development of the RNA three-hybrid system, identification of protein–RNA interactions was limited to *in vitro* methods such as pull-down assays using radiolabeled RNA. The introduction of the RNA three-hybrid system has allowed not only the detection of well-studied protein–RNA

pairs, but also the identification of novel protein–RNA interactions. An impressive application of this system is the cloning of a regulatory protein from *Caenorhabditis elegans* that binds to the 3′ untranslated region of the FEM-3 (fem-33′UTR) and mediates the sperm/oocyte switch in hermaphrodites [57]. In this assay, a bifunctional RNA plasmid possessing fem-33′UTR and the RNA ligand for the MS2 coat protein was introduced into a yeast strain expressing a DBD-MS2 upstream of the *HIS3* and *lacZ* reporter genes. Into this strain, a complementary DNA-AD library was introduced. Cells containing a positive protein–RNA interaction were selected first for HIS3 and lacZ activation followed by screening for the presence of the bifunctional RNA plasmid. The RNA plasmid from successful candidates was lost by reverse selection and the cells were tested again for lacZ activation to reduce the number of false positives. Cells that failed to activate lacZ after plasmid loss were tested for fem-33′UTR binding specificity by reintroduction of the bifunctional RNA plasmids. The protein encoded in the only cDNA-AD that satisfied all selection and screening criteria was found to have 93% homology at the nucleotide level with two genes encoded in the *C. elegans* genome. Further testings confirmed these genes to be regulators of the sperm/oocyte switch in hermaphrodite *C. elegans*. The specificity with which the RNA three-hybrid assay selected just one protein from thousands for the selected protein–RNA interaction illustrates the power of this assay for finding novel protein–RNA interactions [16]. The recent discovery, for example, of RNAi highlights the need not to forget about molecules other than proteins when carrying genetic assays [58, 59].

4.1.4.4 Small Molecule–Protein Interactions

While several small molecule three-hybrid systems have now been reported, it was only in 2004 that such a system was used successfully for drug discovery research. Specifically, Becker and coworkers reported that the Mtx yeast three-hybrid system developed in our laboratory could be used to clone novel protein targets of CDK inhibitors (Table 4.1-2) [47]. The CIDs used in this study took advantage of the low picomolar affinity of Mtx for DHFR [25]. Three known CDK inhibitors, roscovitine, purvalanol B, and indenopyrazole, were linked to Mtx and introduced into a yeast strain expressing a DBD-DHFR protein fusion upstream of the *HIS3* reporter gene and a library of kinase cDNAs linked to a transcription AD. With this system they isolated, besides the known CDK targets, 29 new kinase targets, 22 of which were either confirmed by *in vitro* binding or enzyme inhibition assays. We speculate that the success here was from the use of the high-affinity Mtx/DHFR anchor, which, as we recently showed, gives a K_D cutoff of ca 100 nM in the yeast three-hybrid assay.

4.1.4.5 Catalysis

The widespread utility and robust transcription read-out of the *n*-hybrid system motivated several laboratories to develop general methods to detect enzyme

Table 4.1-2 Summary of biochemical analysis of purvalanol B–Protein interactions. Binding of proteins to immobilized purvalanol B but not to CDK-inactive-N6-methylated purvalanol B was evaluated by immunoblotting or liquid chromatography-mass spectrometry (for endogenous Jurkat proteins). Enzyme assays were performed with purified enzymes and percentage inhibition of kinase activity observed with 1 μM purvalanol B

	Name	Group	Y3H array anaysis	Secondary validation assays				Y3H targets confirmed in at least one secondary assay
				Affinity chromatography			Enzyme assay	
			Purvalanol B-MFC	MS\Endo	WB\Endo	WB\Exo	%Inhibition at 1 μM compound	
Targets Identified in cDNA screens with Purvalanol B-MFC MFC = Methotrexate fusion compound	CDK2	CMGC	3	+	+	+	99	
	CSNK1D	CK	3	+	+	+	nd	
	CSNK1E	CK	3	+	nd	nd	nd	
	CSNK1G	CK	3	–	nd	+	nd	
	CLK1	CMGC	2	–	nd	–	nd	
	CLK2	CMGC	3	–	nd	nd	nd	
	CLK3	CMGC	3	–	nd	+	nd	
	PAK4	CMGC	3	+	nd	nd	nd	
	PCTK1	CMGC	3	–	nd	+	nd	
	PCTK2	CMGC	1[H]	+	nd	nd	nd	
	RSK-3	AGC	3	+	nd	nd	61	
	EPHB2	TK	3[H]	–	nd	+	nd	
	EPHB4	TK	3	–	–	–	nd	
	FLT4	TK	3[H]	–	nd	+	nd	
	TRK-B	TK	3	–	nd	nd	nd	
	FYN	TK	3	+	nd	nd	98	
	YES	TK		+	nd	nd	98	

catalysis *in vivo* around the small molecule three-hybrid system. Several proofs of principle papers have been published in the last few years, and now the key test of these systems is whether they can be readily applied to new chemistry. Toward that end, our laboratory recently demonstrated that Chemical Complementation could be used to detect glycosidic bond formation using a glycosynthase [60].

We chose glycosidic bond formation because despite the fundamental role of carbohydrates in biological processes and their potential use as therapeutics, carbohydrates still remain difficult to synthesize. Specifically, this system was developed using the E197A mutant of Cel7B from *Humicola insolens*, which

had previously been shown to be an efficient "glycosynthase" using an α-fluoro donor substrate. Here, enzymatic activity is detected as formation of a bond between a Mtx-disaccharide-fluoride donor (Mtx-Lac-F) and a dexamethasone-disaccharide acceptor (Dex-Cel), which dimerize DBD-eDHFR and AD-GR activating transcription of a *LEU2* reporter gene that permits survival under appropriate selective conditions. The growth advantage conferred by the glycosynthase activity was used to select the Cel7B:E197A glycosynthase from a pool of inactive variants (Cel7B). A mock library containing 100 : 1 inactive variants to glycosynthase underwent 400-fold enrichment in glycosynthase after a single round of selection. Encouraged by this result, we carry out the directed evolution of the glycosidase Cel7B to improve its glycosynthase activity using a Glu197 saturation library. From a library of 10^5 mutants, Cel7BE197S was selected, which showed a fivefold improvement glycosynthase activity over the known Cel7B:E197A glycosynthase (Table 4.1-3).

As intended, no further modifications to Chemical Complementation were needed to extend this assay to detect glycosynthase activity. All that was required to detect glycosynthase activity was to add the Dex and Mtx saccharide substrates. This result shows the generality of Chemical Complementation, and the ease with which it can be applied to new chemical reactions. Moreover, it shows that Chemical Complementation can detect not only bond cleavage but also bond formation reactions. Although, the size of the Glu197 saturation library selected here was quite small, with only 32 members at the DNA level, the transformation efficiency of *S. cerevisiae*, however, allows much larger libraries, in the order of 10^5–10^7.

4.1.5 Future Development

The yeast two-hybrid assay no doubt will continue to be a mainstay technique for the discovery of new protein–protein interactions. As biological pathways

Table 4.1-3 Glycosynthase activities and protein purification yields for Cel7B variants

	E197A	E197S	N196D/E197A
Specific activity (mol $[F]$)/(min^{-1} mol $[E_0]$)	8 ± 2	40 ± 5	7 ± 1
Protein purification yield [nmol l^{-1}]	6.1	4.6	7.3

Glycosynthase activity for tetrasaccharide synthesis from α-lactosyl fluoride and *p*-nitrophenyl β-cellobioside (PNPC) was measured for the *Humicola insolens* Cel7B variants in sodium phosphate buffer, pH 7.0, at room temperature. Specific activities were determined by measuring the fluoride ion release rate by a fluoride ion selective electrode. The protein purification yields are the yield of purified protein as determined by western analysis from total cell culture.

are being studied increasingly at the systems level, the two-hybrid assay has the potential to be quite useful for analyzing total protein dynamics in living cells. As seen in the PCA work by Michnick and coworkers, it is here that technical improvements will prove important for the two-hybrid assay.

But it is the *n*-hybrid assays that have the potential to extend the power of genetics to molecules other than proteins, such as nucleic acids and small molecules. Despite this enormous potential, use of these other *n*-hybrid assays pales in comparison to that of the two-hybrid assay. As we argue in this chapter, a consideration of the published literature suggests that this discrepancy is not the result of some inherent technical limitation to the *n*-hybrid assays, but rather likely reflects the bias of current practice. Thus, it is here that we believe there is most potential for the future development of the *n*-hybrid assay and indeed genetics as a whole. Technically, the *n*-hybrid assays probably still can be further developed for different classes of molecules or posttranslational modifications. But already in their present form these assays seem to have tremendous potential for biological discovery, uncovering new functions for the many classes of molecules that make up the cell.

These advances also expand our ability to engineer the cell to harness its synthetic and functional capabilities for chemical discovery. Just as protein engineering impacted both basic research and the biotechnology and pharmaceutical industries in the last 25 years, so should cell engineering in this century. Such systems engineering likely will require a much more quantitative understanding of cellular processes, and accordingly the *n*-hybrid assays will have to be characterized and rebuilt on this level, allowing, for example, the K_D cutoff of the assay to be dialed-in. Using this genetic assay in entirely new ways should then open the door for new chemistry, with the potential to match the complexity of cell function.

References

1. S. Fields, O. Song, A novel genetic system to detect protein-protein interactions, *Nature* **1989**, *340*, 245–246.
2. E.M. Phizicky, S. Fields, Protein-protein interactions: methods for detection and analysis, *Microbiol. Rev.* **1995**, *59*, 94–123.
3. L. Keegan, G. Gill, M. Ptashne, Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein, *Science* **1986**, *231*, 699–704.
4. E.A. Golemis, *Protein-Protein Interactions: a Molecular Cloning Manual*, 1st ed., Cold Spring Harbor Lab Press, New York, 2002.
5. B.T. Carter, H. Lin, V.W. Cornish, in *Directed Molecular Evolution of Proteins*, (Eds.: S. Brakmann, K. Johnsson), Wiley-VCH Verlag, Weinheim, 2002.
6. E. Phizicky, P.I. Bastiaens, H. Zhu, M. Snyder, S. Fields, Protein analysis on a proteomic scale, *Nature* **2003**, *422*, 208–215.
7. C.R. Geyer, R. Brent, Selection of genetic agents from random peptide aptamer expression libraries, *Methods Enzymol.* **2000**, *328*, 171–208.

8. H. Lin, V.W. Cornish, In vivo protein-protein interaction assays: beyond proteins we would like to thank Tony Siu, Dr. Charles Cho, and the members of our lab for their helpful comments as we were preparing this manuscript, *Angew. Chem., Int. Ed. Engl.* **2001**, *40*, 871–875.
9. H. Lin, V.W. Cornish, Screening and selection methods for large-scale analysis of protein function, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 4402–4425.
10. L.H. Hwang, L.F. Lau, D.L. Smith, C.A. Mistrot, K.G. Hardwick, E.S. Hwang, A. Amon, A.W. Murray, Budding yeast Cdc20: a target of the spindle checkpoint, *Science* **1998**, *279*, 1041–1044.
11. J.A. Chong, G. Mandel, in *The Yeast Two-Hybrid System*, (Eds.: B. P.L., S. Fields), Oxford University Press, New York, 1997, pp. 289–297.
12. M.K. Alexander, D. Bourns, V.A. Zakian, in *Two-Hybrid Systems, Methods and Protocols*, Vol. 177 (Ed.: P.N. MacDonald), Humana Press, New Jersey, 2001, pp. 241–260.
13. M.M. Wang, R.R. Reed, Molecular cloning of the olfactory neuronal transcription factor Olf-1 by genetic selection in yeast, *Nature* **1993**, *364*, 121–126.
14. S. Jaeger, G. Eriani, F. Martin, Results and prospects of the yeast three-hybrid system, *FEBS Lett.* **2004**, *556*, 7–12.
15. B. Zhang, B. Kraemer, D. SenGupta, S. Fields, M. Wickens, Yeast three-hybrid system to detect and analyze interactions between RNA and protein, *Methods Enzymol.* **1999**, *306*, 93–113.
16. D.J. SenGupta, B. Zhang, B. Kraemer, P. Pochart, S. Fields, M. Wickens, A three-hybrid system to detect RNA-protein interactions in vivo, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8496–8501.
17. N. Kley, Chemical dimerizers and three-hybrid systems: scanning the proteome for targets of organic small molecules, *Chem. Biol.* **2004**, *11*, 599–608.
18. S.L. Schreiber, Chemistry and biology of the immunophilins and their immunosuppressive ligands, *Science* **1991**, *251*, 283–287.
19. E.J. Licitra, J.O. Liu, A three-hybrid system for detecting small ligand-protein receptor interactions, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12817–12821.
20. K. Baker, C. Bleczinski, H. Lin, G. Salazar-Jimenez, D. Sengupta, S. Krane, V.W. Cornish, Chemical complementation: a reaction-independent genetic assay for enzyme catalysis, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 16537–16542.
21. S.M. Firestine, F. Salinas, A.E. Nixon, S.J. Baker, S.J. Benkovic, Using an AraC-based three-hybrid system to detect biocatalysts in vivo, *Nat Biotechnol* **2000**, *18*, 544–547.
22. D.D. Clark, B.R. Peterson, Rapid detection of protein tyrosine kinase activity in recombinant yeast expressing a universal substrate, *J. Proteome Res.* **2002**, *1*, 207–209.
23. D.M. Spencer, T.J. Wandless, S.L. Schreiber, G.R. Crabtree, Controlling signal transduction with synthetic ligands, *Science* **1993**, *262*, 1019–1024.
24. J.F. Amara, T. Clackson, V.M. Rivera, T. Guo, T. Keenan, S. Natesan, R. Pollock, W. Yang, N.L. Courage, D.A. Holt, M. Gilman, A versatile synthetic dimerizer for the regulation of protein-protein interactions, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 10618–10623.
25. H. Lin, W. Abida, R. Sauer, W.V. Cornish, Dexamethasone-methotrexate: an efficient chemical inducer of protein dimerization in vivo, *J. Am. Chem. Soc.* **2000**, *122*, 4247–4248.
26. S.J. Kopytek, R.F. Standaert, J.C. Dyer, J.C. Hu, Chemically induced dimerization of dihydrofolate reductase by a homobifunctional dimer of methotrexate, *Chem. Biol.* **2000**, *7*, 313–321.
27. S. Gendreizig, M. Kindermann, K. Johnsson, Induced protein dimerization in vivo through covalent

labeling, *J. Am. Chem. Soc.* **2003**, *125*, 14970–14971.

28. S.S. Muddana, B.R. Peterson, Facile synthesis of cids: biotinylated estrone oximes efficiently heterodimerize estrogen receptor and streptavidin proteins in yeast three hybrid systems, *Org. Lett.* **2004**, *6*, 1409–1412.
29. K.S. de Felipe, B.T. Carter, E.A. Althoff, V.W. Cornish, Correlation between ligand-receptor affinity and the transcription readout in a yeast three-hybrid system, *Biochemistry* **2004**, *43*, 10353–10363.
30. W.M. Abida, B.T. Carter, E.A. Althoff, H. Lin, V.W. Cornish, Receptor-dependence of the transcription read-out in a small-molecule three-hybrid system, *Chembiochem* **2002**, *3*, 887–895.
31. J. Gyuris, E. Golemis, H. Chertkov, R. Brent, Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2, *Cell* **1993**, *75*, 791–803.
32. M. Vidal, R.K. Brachmann, A. Fattaey, E. Harlow, J.D. Boeke, Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 10315–10320.
33. H.M. Shih, P.S. Goldman, A.J. DeMaggio, S.M. Hollenberg, R.H. Goodman, M.F. Hoekstra, A positive genetic selection for disrupting protein-protein interactions: identification of CREB mutations that prevent association with the coactivator CBP, *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13896–13901.
34. K. Baker, D. Sengupta, G. Salazar-Jimenez, V.W. Cornish, An optimized dexamethasone-methotrexate yeast 3-hybrid system for high-throughput screening of small molecule-protein interactions, *Anal. Biochem.* **2003**, *315*, 134–137.
35. J.C. Hu, E.K. O'Shea, P.S. Kim, R.T. Sauer, Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions, *Science* **1990**, *250*, 1400–1403.
36. S.L. Dove, J.K. Joung, A. Hochschild, Activation of prokaryotic transcription through arbitrary protein-protein contacts, *Nature* **1997**, *386*, 627–630.
37. E.A. Althoff, V.W. Cornish, A bacterial small-molecule three-hybrid system, *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 2327–2330.
38. S.W. Michnick, I. Remy, F.X. Campbell-Valois, A. Vallee-Belisle, J.N. Pelletier, Detection of protein-protein interactions by protein fragment complementation strategies, *Methods Enzymol.* **2000**, *328*, 208–230.
39. I. Remy, J.N. Pelletier, A. Galarneau, in *Protein-Protein Interactions*, (Ed.: E. Golemis), Cold Spring Harbor Laboratory Press, New York, 2001, pp. 449–475.
40. S.W. Michnick, I. Remy, F. Valois, in *Methods in Enzymology*, Vol. 14, (Eds.: J. Abelson, S. Emr, J. Thorner), Academic Press, London, 2000, pp. 208–230.
41. F. Rossi, C.A. Charlton, H.M. Blau, Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation, *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8405–8410.
42. T. Wehrman, B. Kleaveland, J.H. Her, R.F. Balint, H.M. Blau, Protein-protein interactions monitored in mammalian cells via complementation of beta-lactamase enzyme fragments, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3469–3474.
43. I. Remy, S.W. Michnick, Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays, *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5394–5399.
44. I. Remy, S.W. Michnick, Visualization of biochemical networks in living cells, *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7678–7683.
45. E.A. Althoff, *Engineering Ligand-Receptor Interactions Using a Bacterial Three-Hybrid System*, Columbia University, New York, 2004.
46. J. Estojak, R. Brent, E.A. Golemis, Correlation of two-hybrid affinity data

with in vitro measurements, *Mol. Cell. Biol.* **1995**, *15*, 5820–5829.

47. F. Becker, K. Murthi, C. Smith, J. Come, N. Costa-Roldan, C. Kaufmann, A. Hanke, S. Dedier, S. Dill, D. Kinsman, N. Hediger, N. Bockovich, S. Meier-Ewert, A three-hybrid approach to scanning the proteome for targets of small molecule kinase inhibitors, *Chem. Biol.* **2004**, *11*, 211–223.
48. A. Hochschild, M. Ptashne, Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix, *Cell* **1986**, *44*, 681–687.
49. K. Joung, E. Ramm, C. Pabo, A bacterial two-hybrid selection system to study protein-DNA and protein-protein interactions, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7382–7387.
50. I. Xenarios, L. Salwinski, X.J. Duan, P. Higney, S.M. Kim, D. Eisenberg, DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions, *Nucleic Acids Res.* **2002**, *30*, 303–305.
51. P. Uetz, L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, J.M. Rothberg, A comprehensive analysis of protein-protein interactions in saccharomyces cerevisiae, *Nature* **2000**, *403*, 623–627.
52. P. Colas, B. Cohen, T. Jessen, I. Grishina, J. McCoy, R. Brent, Genetic selection of peptide aptamers that recognize and inhibit cyclin-dependent kinase 2, *Nature* **1996**, *380*, 548–550.
53. M. Yang, Z. Wu, S. Fields, Protein-peptide interactions analyzed with the yeast two-hybrid system, *Nucleic Acids Res.* **1995**, *23*, 1152–1156.
54. P. Youderian, A. Vershon, S. Bouvier, R.T. Sauer, M.M. Susskind, Changing the DNA-binding specificity of a repressor, *Cell* **1983**, *35*, 777–783.
55. S. Tan, D. Guschin, A. Davalos, Y.L. Lee, A.W. Snowden, Y. Jouvenot, H.S. Zhang, K. Howes, A.R. McNamara, A. Lai, C. Ullman, L. Reynolds, M. Moore, M. Isalan, L.P. Berg, B. Campos, H. Qi, S.K. Spratt, C.C. Case, C.O. Pabo, J. Campisi, P.D. Gregory, Zinc-finger protein-targeted gene regulation: genomewide single-gene specificity, *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11997–12002.
56. Sangamo, Biosciences, *Vol. 2005*, 2005, pp. Sangamo Bio Science Inc, www.sangamo.com; Biotechnology company focused on the research and development of novel transcription factors for regulating human, plant, and microbial genes.
57. B. Zhang, M. Gallegos, A. Puoti, E. Durkin, S. Fields, J. Kimble, M.P. Wickens, A conserved RNA-binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line, *Nature* **1997**, *390*, 477–484.
58. G.J. Hannon, RNA interference, *Nature* **2002**, *418*, 244–251.
59. D.R. Engelke, J.J. Rossi, *RNA Interference, Methods Enzymology* Vol. 392, 2005, 1–454.
60. H. Lin, H. Tao, V.W. Cornish, Directed evolution of a glycosynthase via chemical complementation, *J. Am. Chem. Soc.* **2004**, *126*, 15051–15059.