Forcing interactions as a genetic screen to identify proteins that exert a defined activity

Michael DeVit,1,3 Paul J. Cullen,2,4 Margaret Branson,1 George F. Sprague Jr.,2 and Stanley Fields1,5

1Howard Hughes Medical Institute, Departments of Genome Sciences and Medicine, University of Washington, Seattle, Washington 98195, USA; 2Institute of Molecular Biology, University of Oregon, Eugene, Oregon 97403, USA

The interaction of proteins to form macromolecular complexes is the basis for most biological processes. Approaches have been described that employ artificial constructs to promote such complexes and assess the consequences. For example, a protein interaction scheme has been described that examines the effects of a specific phosphorylation event catalyzed by a protein kinase via the provision of an artificial protein binding interface between a modified version of the kinase and a single substrate. We have generalized this type of approach to form the basis for a genetic selection to identify proteins that exert an activity when recruited to a target protein. The assay uses the leucine zipper domains from the mammalian transcription factors Fos and Jun to force the interaction of two proteins. With a target protein fused to the Jun zipper and a library of open reading frames fused to the Fos zipper, we demonstrate this approach in yeast with both a selection to identify membrane-associated proteins and a selection to identify candidate components of the filamentous growth MAP kinase pathway.

Results

Leucine zippers result in protein colocalization

Our approach is described in Figure 1A, using as an example a signaling pathway in which the presence of a signal results in the formation of a protein complex and a cellular response. By placing complementary zippers on two proteins in this pathway, we hypothesize that in some cases a response will be observed in the absence of the relevant signal. The Fos and Jun zippers preferentially heterodimerize with a dissociation constant on the order of 1 nM (Heuer et al. 1996). We first identified the optimal leucine zipper domains and their arrangement in protein fusions by linking several versions of the Fos and Jun zippers to the Gal4 DNA-binding and activation domains, and using the two-hybrid assay as a measure of the relative strength of the interaction (data not shown). These experiments led us to place the Fos leucine zipper near the amino terminus followed by a short flexible linker (MRGS – Fos residues 159–202 – SGTGQFQLTM – ORF) and the Jun leucine zipper at the carboxy terminus (ORF – GGATGG – Jun residues 277–334). To determine what fraction of the Fos and Jun fusion proteins associate within individual cells, we coexpressed a Fos–green fluorescent protein fusion (Fos-GFP) with proteins of known subcellular localization fused to Jun, with both fusion genes under the control of a tetracycline-regulated promoter. When the Fos-GFP was coexpressed with an empty Jun vector, GFP fluorescence was distributed throughout the cytoplasm and nucleus (data not shown). However, when the Fos-GFP was coexpressed with a protein fused to Jun, a large fraction of the GFP fluorescence was properly localized, as observed for Ole1, at the endoplasmic reticulum; Tom20, at the mitochondrial outer membrane; Hof1, at the septin ring at the bud neck; and Mig1, in the nucleus (Fig. 1B). Mig1 is regulated transcription factor that rapidly moves from the cytoplasm to the nucleus upon addition of glucose (DeVit et al. 1997), as did the Fos-GFP when coexpressed with Jun-Mig1 (Fig. 1B). Coassociation was typically
visible in 70%–90% of the cells in a culture. We constructed a library of all yeast ORFs fused to Fos (see Supplemental data) under tetracycline control and observed a variety of subcellular localizations in a large fraction of randomly isolated transformants that coexpressed a GFP-Jun fusion with a member of the library (data not shown). These results confirm that Fos and Jun leucine zippers can cause a strong coassociation of two proteins and can do so in a variety of subcellular localizations in a yeast cell.

A screen for membrane-associated proteins

To demonstrate that this approach can be used as a genetic screening method, we carried out a screen to identify proteins that associate with the plasma membrane. We made use of the Sos-Ras recruitment system, which was originally designed as a method to identify protein–protein interactions (Aronheim et al. 1997). The yeast strain contains a temperature-sensitive mutant of Cdc25, an essential guanine–nucleotide exchange factor for Ras. The human homolog of Cdc25, Sos, can complement this mutant and permit growth at the restrictive temperature, but only if it is localized to the plasma membrane to be in close proximity with Ras. We sought to identify membrane-associated proteins in the Fos-ORF library by their rescue of the high-temperature growth defect of this strain in the presence of a Sos-Jun fusion (Fig. 2A). From 594 colonies able to grow at the restrictive temperature, we identified 215 Fos-fused genes (provided in Supplemental Table S1). Of these, 119 encode proteins known to be membrane-associated, as assessed by a gene ontology (GO) molecular component annotation in the Yeast Proteome Database (http://proteome.incyte.com) as plasma membrane, endoplasmic reticulum, Golgi, vesicles, vacuole, or nuclear membrane associated (Table 1). That we obtained proteins associated with various components of the endomembrane system is not surprising, given that they are continuously mixing as proteins are sorted, and that the Fos and Jun fusions are overexpressed. The rest of the ORFs identified are classified as having a nonmembrane localization, have an apparent membrane localization (examples shown in Fig. 2B), and another eight have a punctate appearance that suggests a potential association with vesicles (Table 1). Closer examination of the YPD annotations revealed that for an additional 17 ORFs classified as either uncharacterized or not membrane associated, there is some evidence of membrane association (e.g., they have roles in lipid biosynthesis or interact with known membrane proteins). Our results differ in some cases from published localization screens (Kumar et al. 2002; Huh et al. 2003) because we may be able to detect transient or weak membrane association that fulfills the genetic selection but is not sufficient to assign localization when GFP fluorescence is examined microscopically. In total, 169 (79%) of the ORFs isolated in the screen are known or likely membrane-associated proteins, and only 38 (17%) are likely to be false positives (Table 1). These results demonstrate this method works as a genomic screening method.

A screen for proteins implicated in the filamentous growth pathway

As a second and more specific test, we screened for proteins that influence the filamentous growth MAP kinase pathway, which is required for the dimorphic transition to filamentation in nutrient-limiting conditions (Gimeno et al. 1992). The signal transduction cascade mediating this response includes an integral membrane protein, Msb2 (Cullen et al. 2004), at the head of a MAP kinase cascade (Liu et al. 1993). We used strains of the filamentous background, Σ1278b, and sought proteins that could activate the pathway when forced to associate with Msb2 (Fig. 3A). Yeast carrying Jun fused to the cytoplasmic tail of Msb2 were transformed with the Fos-ORF library and selected for growth by activation of an Msb2-dependent reporter (Cullen et al. 2004).

Of 220 His′ colonies obtained from screening 230,000 transformants, only nine passed additional tests that showed that activation of the pathway was due to expression of a Fos-ORF protein. Sequencing revealed six proteins: Kss1, the MAP kinase for the filamentous growth pathway; Tpk2, a cAMP-dependent kinase subunit; Gis2, an uncharacterized protein involved in glucose sensing; Ypc1, a ceramidase enzyme; and two uncharacterized proteins, Yer158c and Ybr062c. To determine whether the
We have shown that a method based on the forced association of two proteins can be used to identify a general class of proteins (membrane associated), as well as a specific component of a signaling process (YBR062c). This method may be useful to identify proteins in any process or pathway that is activated when two or more proteins come into close physical proximity, provided that a conditional phenotype can be established. However, because these screens rely on protein tagging with leucine zipper domains, overexpression of these constructs, and forced association, any candidate proteins that emerge should be viewed with caution. The biological relevance of such proteins must be further established by additional experimentation.

Signal transduction pathways may be particularly amenable to this approach, as ligand binding can cause proteins to interact at the plasma membrane with the cytoplasmic tail of a receptor, and a screen can be performed in the absence of ligand. Proteins that are modified by other proteins, e.g., substrates of a protein kinase, might be identified if the forced association increases the rate of phosphorylation sufficiently to activate a downstream response. Similarly, a protein kinase could be identified given a substrate. However, promiscuous specificities that result from forced association may occur that would need to be ruled out, and protein kinase-substrate pairs may not be detected if kinase activation is required. For other processes, a screen could be performed in a yeast strain modified such that normal protein interactions are disrupted. For example, a protein interaction domain could be deleted or a protein that acts as a bridging component could be deleted. Other screens to identify a class of proteins might also be possible. For example, we observed that no fluorescence is visible when GFP-Jun is coexpressed with several proteins that interact.

Table 1. Classification of ORFs identified in membrane screen (215 total)

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Membrane associated</td>
<td>119</td>
</tr>
<tr>
<td>By Jun-GFP colocalization</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>144</td>
</tr>
<tr>
<td>Possible membrane associated</td>
<td>8</td>
</tr>
<tr>
<td>By Jun-GFP colocalization</td>
<td>8</td>
</tr>
<tr>
<td>By other evidence</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>25</td>
</tr>
<tr>
<td>Not membrane associated</td>
<td>31</td>
</tr>
<tr>
<td>By GO annotation</td>
<td>31</td>
</tr>
<tr>
<td>By Jun-GFP colocalization</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>38</td>
</tr>
<tr>
<td>Not enough information to assess</td>
<td>8</td>
</tr>
</tbody>
</table>
eral Fos-fused proteins. It seems likely that the association of GFP-Jun to the protein fused to Fos caused GFP degradation, which if so, suggests the possibility for a screen to identify proteins involved in protein degradation.

Although these proof of principle experiments have been carried out in yeast, it should be possible to carry out such screens in other organisms. As Sharma et al. (2003) have shown, a signal transduction cascade in mammalian cells can be activated by forcing substrates to associate with tyrosine kinase receptors. In fact, our approach may be particularly useful for mammalian cells, as it uses ectopic expression of proteins and dominant gain-of-function phenotypes, although the presence of endogenous Fos and Jun interactions may require the use of different dimerization domains. Screens can be carried out without the need to create specially engineered mutant strains or to later identify mutated genes. The most limiting factor is the construction of a library of genes fused to Fos (or other small protein–protein interaction domain). However, the availability of complete genome sequences, which is leading to the construction of extensive gene sets, makes this method a powerful new tool for protein discovery studies.

Methods

Strains

Yeast were grown at 30°C in rich (YPD), or minimal (SD) medium lacking nutrients required to maintain selection of plasmids (Rose et al. 1990). Strain 839 used to observe GFP fluorescence was derived from strain DBY746 (MATa, ura3-52, his3Δ1, leu2-3,112, trp1-28). The TTA transactivator for tetracycline regulated promoters was integrated at the leu2 locus with plasmid pCM149 (Gari et al. 1997). The LEU2 marker of pCM149 was then removed by PCR-mediated gene replacement with a kanMX cassette (Guldener et al. 1996). Strain 843 used to screen for genes activating the filamentous growth pathway was created from PC1199 (MATa, ura3-52, leu2::NatMX, ste4, Fus1-His3, Msb2-HA). TRP1 was deleted by using plasmid pNYK1009 (Alani et al. 1987), which inserts the URA3 gene. After TRP1 disruption, the URA3 sequence was removed as the result of homologous recombination between repeated sequences that flank the URA3. Overexpression of the full-length YBR062c gene was performed by using primers 5’-AGGTCTATTCTCAAATCTTGGTAA CACCGATGGTGAGGTAGTGGCATTAGCGTGTTTAAAC-3’ and 5’-TATCCATTCTTGGCGAACCATCTCATTGTAAT GATCTCAGGAGGCTTGGAGATCGCAGGTTTT-3’ to amplify the GAL1 promoter by PCR from the pFA6a-kanMX6-PGAL1 template provided by John Pringle (Longtine et al. 1998). Integration was targeted to the YBR062c locus by homologous recombination (Baudin et al. 1993) in cells of the 1278b background in strain S38 (MATa ste4 Fus1-His3 Fus1-lacZ) to create strain 1492 (MATa ste4 Fus1-His3 Fus1-lacZ pGAL1-YBR062c::KanMX6).

Plasmids

Standard procedures for the manipulation of plasmid DNA and growth and transformation of Escherichia coli were followed (Sambrook et al. 1990). All plasmid constructs were made by recombinational cloning in yeast (Hudson Jr. et al. 1997) by linearization with a restriction enzyme that cuts at the position where the new DNA sequence was to be inserted. DNA sequences were amplified by PCR with primers that add flanking sequence such that homologous recombination targets the insert to the desired position in the recipient plasmid. The insert and linearized recipient plasmid DNAs were cotransformed into yeast.
(Gietz et al. 1992) followed by selection for the auxotrophic marker on the plasmid. Plasmids were recovered from extracts of yeast colonies by transformation into E. coli. Correct plasmids were identified by restriction enzyme digestion mapping and/or sequencing of inserts and fusion points and introduced back into yeast.

Fos and Jun two hybrid constructs were derived by removal or addition of sequences to a Gal4 binding-domain Fos fusion (pPC76) and a Gal4 activation-domain Jun fusion (pPC79) (Chevray and Nathans 1992). Fos-BD and Jun-AD fusions were transformed into two-hybrid strain, PJ69-4a, and β-galactosidase assays were carried as previously described (Hudson Jr. et al. 1997). The following fusions were carried on a series of vectors created for tetracycline-regulated expression in yeast (Gari et al. 1997): Fos-GFP, GFP-Jun, Mig1-Jun, Tom20-Jun, Hof1-Jun, Ole1-Jun, Msb2-Jun, and Rgt2Tad-Jun. Fos fusions were constructed in plasmid pCM181 that was modified so that the following amino acid sequence was added to the amino terminus of the desired protein: MRGS-Fos leucine zipper (amino acids 159–202)–SGTSQQLTTM-ORF. Jun fusions were constructed in plasmid pCM187, so that the following sequence was added to the carboxy terminus of the desired protein: ORF-GGGTG-Jun leucine zipper (amino acids 277–334). A version of Ybr062c lacking the Fos leucine zipper was created by inserting a DNA segment containing the YBR062c ORF amplified from yeast genomic DNA into pCM187. Both wild-type and Fos-fused versions used here were based on an early ORF prediction in the Saccharomyces Genome Database (http://www.yeastgenome.org/). Since these experiments were completed, the sequence of the region upstream of YBR062c has been revised, resulting in a larger predicted ORF of 180 amino acids (see http://db.yeastgenome.org/cgi-bin/SGD/locusHistory.pl?sgdid=S0000266).

Library construction

A plasmid library of yeast proteins fused to Fos (Fos-ORF library) was created by inserting the sequence encoding the Fos leucine zipper in place of that encoding the Gal4 activation domain from a two-hybrid plasmid (GAD) library (Ito et al. 2000). This was carried out en masse by homologous recombination in yeast; the pooled GAD library, which was linearized by restriction enzyme digestion with enzymes that recognize sites in the Gal4 activation domain encoding sequence, was cotransformed with a DNA cassette containing the Fos encoding sequence, and as well as the tet promoter that replaces the ADH1 promoter in the GAD plasmid and the kanMX gene as a marker to select for insertion of the cassette into the plasmid. The cassette also contains flanking sequence homologous to sequence in the pGAD plasmids so that homologous recombination causes the cassette to be inserted in the proper position and orientation: The 5′ homology targets the 5′ end of the cassette to be inserted upstream of the ADH1 promoter, and the 3′ homology targets the 3′ end of the cassette to be inserted just upstream of the sequence encoding for a yeast ORF. The homology was designed to cause the Fos coding sequence to be in the same reading frame as the yeast ORF and include an intervening linker. The resulting proteins have the following sequence: MRGS-Fos leucine zipper (amino acids 159–202)–ELGSGAPEFQLT-ORF. For the recombinant, the GAD library was linearized in separate reactions with three different restriction enzymes. The reactions were then combined and fractionated on 0.8% agarose gels to separate the linearized plasmids based on the size of the ORF. Pools of two size ranges were isolated by cutting the gel into two sections and electroeluting the DNA. Each size pool was cotransformed with the Fos cassette in separate reactions for the recombination. The transformations were plated on media lacking leucine and containing G418 to select for plasmids that have undergone a successful homologous recombination. The transformants were pooled by scraping colonies from all plates into one batch. Plasmids were recovered by transformation of purified extracts of aliquots of the pooled yeast into E. coli. The bacterial colonies were pooled by scraping the colonies into a single batch. Plasmids were purified from these bacteria. Sequencing of randomly picked bacterial colonies before pooling confirmed that the recombination successfully produced a library of Fos-fused yeast ORFs.

Microscopy

Bud site–selection phenotypes were scored by direct visualization of cells scraped from the plate at 1000× using DIC and were confirmed by staining with 1 µg/mL Calcofluor and analysis by using a UV filter at 1000× using established protocols (Chant and Pringle 1995; Pringle et al. 1995; Cullen and Sprague Jr. 2002). More than 200 cells were examined for budding pattern by each method. To visualize GFP fluorescence, yeast were grown to early log phase, concentrated, and imaged by fluorescence microscopy on a Zeiss Axiocent microscope with a 63× 1.4 NA objective.

Membrane screen

A calc25 temperature-sensitive mutant yeast strain (MATa, calc25, ura3, lys2, leu2, trp1, his3&200, ade101) carrying two plasmids, one expressing the tTA transactivator (pCM188) (Gari et al. 1997) and another expressing a Sos-Jun fusion (pSON8-Jun) (Aronheim et al. 1997), was transformed with the Fos-ORF library. The yeast were plated on SD media lacking uracil, tryptophan, and leucine to select for all three plasmids and incubated at 36°C. Aliquots were plated on the same media and incubated at the permissive temperature of 25°C to determine the total number of transformants. Plasmids were recovered from the yeast colonies that formed at 36°C, and the identity of the ORF fused to Fos was determined by sequencing with a primer that anneals in the Fos coding sequence.

Filamentous growth pathway screen

The Fos-ORF library was transformed into strain 843, which carried MSB2-Jun on pCM187. Cells were plated on SD medium lacking uracil, tryptophan, leucine, and histidine and containing 3 mM 3-aminotriazole (3AT). His+ colonies were restreaked on the same media containing 25 µg/mL doxycycline, which represses expression of the Msb2-Jun and Fos-ORF fusions. Only candidates that did not grow, or grew much more slowly, on doxycycline were analyzed further. The candidates were recovered into bacteria, and the identity of the ORF was determined by sequencing with a primer complementary to the Fos coding sequence. Each unique ORF was retransformed into the test strain along with plasmids expressing Jun-Msb2, GFP-Jun, or Jun-Rgt2A tail and tested for its ability to confer growth on media lacking histidine and containing 3 mM 3AT. The single cell invasive growth assay was performed as described (Cullen and Sprague Jr. 2000). Cells containing fusion protein-bearing plasmids and control plasmids were also examined from colonies grown to saturation on SD minus uracil, leucine, and tryptophan solid-agar medium by microscopy at 100×. Protein overexpression was induced by serial passages of strains in YEP media supplemented with 2% galactose (YEPGal) followed by 24-h incubation at 30°C for bud site–selection assays and 72 h at 30°C to examine colony ruffling phenotypes. Colony ruffling was ob-
served by direct visualization of patches on YEPGal semisolid agar medium.

Acknowledgments

We thank Takashi Ito for providing the GAD library, and Enrique Hererro and John Pringle for providing plasmids. S.F. is an investigator of the Howard Hughes Medical Institute.

References


Web site references

http://proteome.incyte.com; Yeast Protein Database.

http://www.yeastgenome.org; Saccharomyces Genome Database.

http://db.yeastgenome.org/cgi-bin/SGD/histLoc?locus=YBR062c; description of YBR062c in the Saccharomyces Genome Database.

Received September 15, 2004; accepted in revised form January 13, 2005.