Available online at www.sciencedirect.com
SCIENCE DIRECT®
Methods 32 (2004) 381–388
METHODS
www.elsevier.com/locate/ymeth

A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways

Ingrid Remy and Stephen W. Michnick*

Département de Biochimie, Université de Montréal, C.P. 6128, Succursale centre-ville, Montréal, Que., Canada H3C 3J7

Accepted 6 October 2003

Abstract

Progress towards a deeper understanding of cellular biochemical networks demands the development of methods to both identify and validate component proteins of these networks. Here, we describe a cDNA library screening strategy that achieves these aims, based on a protein-fragment complementation assay (PCA) using green fluorescent protein (GFP) as a reporter. The strategy combines a simple cell-based cDNA-screening approach (interactions of a ''bait'' protein of interest with ''prey'' cDNA products) with specific functional assays that use the same system and provide initial validation of the cDNA products as being biologically relevant. We applied this strategy to identify novel interacting partners of the protein kinase PKB/Akt. This method provides very general means of identifying and validating genes involved in any cellular process and is particularly designed for identifying enzyme substrates or regulatory proteins for which the enzyme specificity can only be defined by their interactions with other proteins in cells in which the proteins are normally expressed.

Keywords: Protein-fragment complementation assay; cDNA library screening; Protein–protein interactions; Green fluorescent protein

1. Introduction

Rapid progress in genome projects is leading to the identification or prediction of a huge number of genes, but only a fraction of gene functions can be inferred from primary gene sequences. To cope with the increasing flood of genome information, we need to develop strategies aimed at characterizing the totality of genes or large subsets thereof. In the past, many ingenious strategies have been devised to simultaneously screen cDNA libraries using a protein- or enzyme-specific assay that allow for both selection of clones and validation of their biological relevance with the same assay [1–4]. However, there are many examples of classes of genes for which there is no obvious and specific functional assay that can be combined with cDNA library screening. Particularly difficult are classes of important enzymes such as kinases, phosphatases, and proteases, that have very broad substrate specificity and bind to many proteins or protein domains when studied out of their appropriate context in intact living cells [5,6]. In the absence of simple and specific assays, researchers have turned to strategies that use as readout some general functional properties of proteins. A first step in defining the function of a novel gene is to determine its interactions with other gene products; that is, since proteins make specific interactions with other proteins as part of functional assemblies, an appropriate way to examine the function of the product of a novel gene is to determine its physical relationships with the products of other genes. This is the basis of the highly successful yeast two-hybrid expression cloning strategies [7–13]. However, a purely protein interaction-based screening approach is limited in that the assays themselves do not provide any immediate information that would allow one to decide whether a cDNA gene product is likely to be involved in a specific cellular function. A particularly powerful experimental

* Corresponding author. Fax: +1-514-343-2015.
E-mail address: stephen.michnick@umontreal.ca (S.W. Michnick).

1046-2023/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2003.10.011

approach to understand gene function would provide for the ability to establish both how proteins and other biological molecules interact in living cells and simultaneously, to validate the biological relevance of these interactions using the same assay system.

In this paper, we describe a cDNA library screening strategy that achieves these aims, using a green fluorescent protein (GFP)-based protein-fragment complementation assay (PCA). The strategy combines a simple cell-based cDNA-screening approach (interactions of a "bait" protein of interest with "prey" cDNA products) with specific functional assays that provide initial validation of the biological relevance of the prey protein. We applied this strategy to the identification of novel substrates or regulators of the serine/threonine protein kinase PKB/Akt. We also present an example of how the GFP PCA is used to achieve initial functional validation of prey cDNAs identified in the screen.

2. Materials and methods

2.1. DNA constructs

For the construction of the GFP[1]-cDNA library fusions, a human brain cDNA library was excised from the vector pEXP1 (ClonCapture cDNA library, Clontech, Palo Alto, CA) using SfiI restriction sites, fractionated by agarose gel (1.2%) electrophoresis from which four pools were isolated according the size of the cDNAs. Each pool of cDNAs was inserted into the eukaryotic expression vector pMT3 [14], 3′ of the GFP[1] fragment of GFP (amino acids 1–158), and transformed into DH5α competent cells. The GFP[1]-cDNA library fusion expression vectors were amplified at 30 °C by inoculating 2.5×10^9 cells per pool into 100 ml Terrific broth (1 L:12 g bacto-tryptone, 24 g yeast extract, 2.3 g KH_2PO_4, 12.5 g K_2HPO_4, and 4 ml glycerol) containing 100 μg/ml ampicillin in 500 ml flasks. The full-length cDNA encoding PKB was amplified by PCR and subcloned 5′ of the F[2] fragment of GFP (amino acids 159–239) into the pMT3 vector where the ampicillin resistance gene has been replaced by a chloramphenicol resistance gene, resulting in the PKB-GFP[2] fusion expressing vector. GFP[1] and GFP[2] were originally amplified by PCR from pCMS-EGFP (Clontech). In all the fusion constructs, a 10 amino acid flexible linker consisting of $(Gly.Gly.Gly.Gly.Ser)_2$ was inserted between the fused protein (or library) and the GFP fragments, to assure that the orientation/arrangement of the fusions in space is optimal to bring the GFP fragments into close proximity. The GFP[1]-GCN4 and GCN4-GFP[2] constructs consist of fusions with GCN4 leucine zipper-forming sequences and are used as controls. GFP PCA expression vectors and the cDNA library are available on request from the authors.

3. Cell lines

COS-1 and HEK293T cells were grown in DMEM (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum (FBS; Hyclone, Logan, Utah).

3.1. GFP PCA-based cDNA library screen

COS-1 cells were plated in 150-mm dishes 24 h before transfection. Cells were transfected (10 μg DNA total/dish) using Lipofectamine reagent (Invitrogen, Carlsbad, CA) at around 60% confluence, with pMT3 vector harboring the human brain cDNA library fused to the F[1] fragment of GFP (GFP[1]-cDNA library) and pMT3-chloramphenicol vector containing the full-length PKB fused to the F[2] fragment of GFP (PKB-GFP[2]). The GFP[1]-cDNA library fusions were transfected in several pools, according to their size. The cotransfection efficiency was optimized with the GCN4 leucine zipper PCA control (see DNA constructs) by varying the concentrations of DNA relative to the number of cells to obtain the maximum number of fluorescent cells as measured by FACS. Forty-eight hours after transfection, positive clones were collected on a fluorescence-activated cell sorter (FACS) analyzer (FACScalibur, Becton–Dickinson, Franklin Lakes, NJ), with stimulation with an argon laser tuned to 488 nm with emission recorded through a 525 nm band width filter. The total DNA from each pool of positive cells was extracted (DNeasy tissue kit, Qiagen, Chatsworth, CA), transformed in DH5α bacterial cells, and plated on LB-agar containing 100 μg/ml ampicillin (hence, there is no propagation of the chloramphenicol-resistant vector harboring the PKB-GFP[2] fusion). DNA plasmids containing the GFP[1]-cDNA fusions were extracted from individual clones and retransfected separately with PKB-GFP[2] or F[2] alone (negative control) to discard negative clones that enter the pool during the cell sorting. After this second round of selection, the DNA plasmids corresponding to the positive clones were submitted to sequence analysis, using a sequencing primer corresponding to the 3′ end of the GFP[1] fragment (5′-CAAGGAGGACGGCAACATCCTG-3′).

3.2. GFP PCA fluorometric analysis

HEK293T cells were split in 12-well plates 24 h before transfection. Cells were cotransfected, at around 60% confluence, with the GFP[1]-hFt1 and PKB-GFP[2] or GFP[1]-GCN4 and GCN4-GFP[2] expressing vectors (1 μg DNA total/well), using Lipofectamine reagent (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. Twenty-four hours after transfection, cells were washed with PBS and serum-starved overnight in medium containing 0.5% FBS. The next day, cells were untreated or treated with 300 nM wortmannin or 50 μM

LY294002 (Calbiochem, San Diego, CA) for the last hour. Afterwards, cells were stimulated for 30 min with 10% serum or 20 μg/ml insulin (Roche Diagnostics, Indianapolis, IN), washed with PBS, gently trypsinized, and resuspended in 200 μl PBS. The total cell suspensions were transferred to 96-well black microtiter plates (Dynex; VWR Scientific, Mississauga, Ontario) and subjected to fluorometric analysis (Spectra MAX GEMINI XS, Molecular Devices, Sunnyvale, CA), using an excitation wavelength of 488 nm, emission of 525 nm, and a cut-off filter of 495 nm. Afterwards, the data were normalized to total protein concentration in cell lysates (Bio-Rad protein assay, Hercules, CA). The background fluorescence intensity corresponding to nontransfected cells was subtracted from the fluorescence intensities of all of the samples. For FACS analysis, 48 h after transfection, COS-1 cells were gently trypsinized, resuspended in 500 μl PBS, and analyzed on a fluorescence-activated cell sorter analyzer (FACScalibur, Becton–Dickinson). For fluorescence microscopy, HEK293T cells were grown on 18-mm glass coverslips prior to transfection. Cells were washed two times with PBS and mounted on glass slides. Fluorescence microscopy was performed on live cells (Nikon TE2000U, FITC filter, objective lens 100×).

4. Results and discussion

4.1. GFP PCA-based functional cDNA library screen

The principle of the PCA strategy is that cells simultaneously expressing two proteins fused to complementary fragments (F[1] and F[2]) of a reporter protein will produce a fluorescent signal, only if the fused proteins physically interact and then bring the complementary fragments of the reporter protein into proximity where they can fold and reassemble into an active form (Fig. 1). We and others have described several PCAs that use different reporters, including GFP, that allows for detection of interactions by a variety of measurement techniques [15–18]. In the GFP PCA-based functional screening strategy presented here, the first step of screening the cDNA library consists of the detection of physical interactions between the bait and cDNA library-encoded prey proteins, by monitoring the reconstitution of GFP in intact cells by FACS (Fig. 1). An important feature of this first step is that interactions can be detected directly and between full-length proteins expressed in cells in which the bait protein normally functions, assuring that subcellular targeting, post-translational modifications, and interactions with other proteins needed for correct functioning of the bait (and prey) can occur (obviously the PCA fragments themselves must not interfere with targeting or modification of the proteins and this must be tested). These features allow for an initial functional validation as follows: first, functional assays can be based on the detection by PCA of perturbations of protein–protein interactions caused by agents, such as hormones or specific inhibitors, that are known to modulate the specific biochemical pathway in which the proteins participate [19]. We have demonstrated this strategy with a PCA based on the enzyme dihydrofolate reductase (DHFR) to map signaling pathways in living mammalian cells [19–21]. Second, observations of subcellular location and induced translocation of complexes can serve as yet another functional validation criterion [19,22]. Thus, the PCA screening strategy combines a simple screening step with direct functional assays.

4.2. cDNA library screening for PKB binding partners

PKB has been demonstrated to play a central role in a number of cellular responses to growth factors and insulin, which include growth, protein synthesis, and anti-

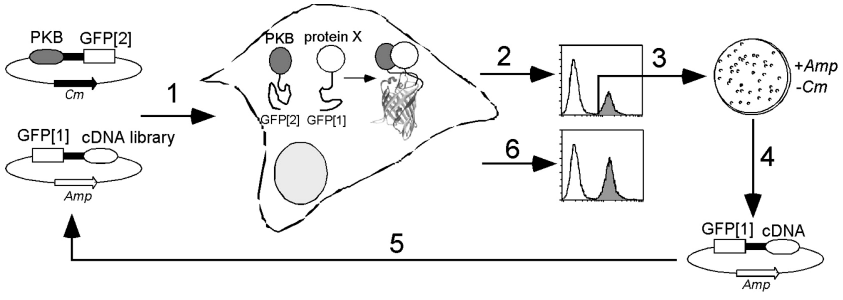

Fig. 1. The GFP PCA-based library screening strategy. A human brain cDNA library was fused to fragment 1 of GFP (GFP[1]-cDNA library) and full-length PKB cDNA to fragment 2 (PKB-GFP[2]), in mammalian expression vectors harboring *E. coli* selection markers Ampicillin (Amp) and Chloramphenicol (Cm), respectively. In the first step (1) COS-1 cells are cotransfected with PKB "bait" and cDNA library "prey" fusions and a physical interaction between the bait and a prey protein induces the folding and reconstitution of GFP from its fragments, generating fluorescence. Positive clones are collected by fluorescence-activated cell sorting (FACS) (step 2) and DNA was extracted from the pools and transformed into *E. coli* grown on Amp plates to select only for plasmids harboring cDNA (step 3). Clones are picked, plasmids are extracted, and interaction of individual proteins with PKB is reconfirmed by cotransfecting COS-1 cells with the PKB fusion and individual cDNA fusions (step 5) and detection by FACS (step 6).

apoptotic/survival signals [23–27]. As with many cellular protein kinases, considerable efforts have been made to map out the links between many cell surface receptors, PKB, and effectors of PKB that are implicated in these processes [27]. However, key to future advances in understanding the multiple functions of PKB rests in identifying and validating novel substrates and regulators. In the studies described here, we sought to identify new PKB-interacting proteins that could be involved in modulating or integrating signaling pathways that impinge upon PKB activation and/or mediate PKB substrate specificity. To achieve this aim, we screened a human brain cDNA library containing between 10^7 and 10^8 independent clones, using PKB as bait. For the screening, we chose a strategy aimed at maximizing clone recovery, while minimizing the chance of isolating false-positive clones. The whole strategy can be schematically divided into five steps (Fig. 1). In the first step, a plasmid expressing full-length PKB fused to the C-terminal fragment of GFP (PKB-GFP[2]) is transiently cotransfected in COS-1 cells with plasmids expressing the cDNA library prey proteins fused to the N-terminal fragment of GFP (GFP[1]-cDNAlibrary) (Fig. 1, step 1). To maximize the incorporation of large cDNAs into cells, the GFP[1]-cDNA library expression vectors were divided into four pools (fractions 1–4), according to the size of the inserted cDNAs. A flexible 10 amino acid linker was also inserted between the fused protein and the GFP fragments, to assure that the orientation/arrangement of the fusions in space is optimal to bring the GFP fragments into close proximity [21]. A physical interaction between a cDNA expressed protein and the bait induces the reconstitution of GFP from its fragments and positive clones can be collected by fluorescence-activated cell sorting (FACS) (Fig. 1, step 2). Plasmids are then extracted from sorted cells and transformed into DH5α bacterial cells, grown in the presence of ampicillin to select for only those plasmids harboring the cDNA constructs (*Amp* marker in the plasmid) and eliminate the PKB bait plasmid containing a chloramphenicol marker (Fig. 1, step 3). *Amp*-positive clones are picked, plasmids are extracted (Fig. 1, step 4), and interaction of individual proteins with PKB is reconfirmed by cotransfecting COS-1 cells with the PKB fusion and individual cDNA fusions (Fig. 1, step 5) and detection by FACS (Fig. 1, step 6). This last step is necessary, since an important source of contaminants in cell sorting are negative cells that are inadvertently sorted to the positive pool, even at slow sorting rates.

Results for the screening of fractions 2, 3, and 4 of the cDNA library are presented in Fig. 2. Fractions 2, 3, and 4 correspond to cDNAs between 0.5–2 kb, 2–3 kb, and 3–4.5 kb, respectively. Fraction 1 (cDNAs <0.5 kb) was not screened because it was likely to contain a large proportion of truncated cDNAs. We compared two negative controls to establish whether false-positive

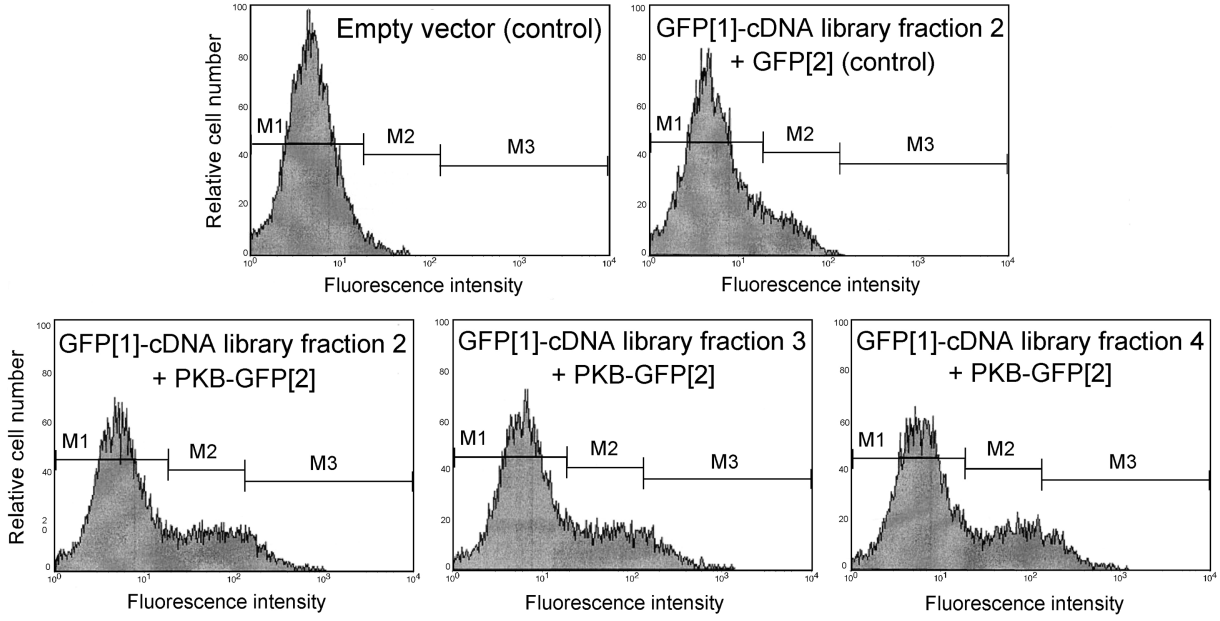

Fig. 2. First-pass screening of size pooled cDNA clones by FACS. COS-1 cells were cotransfected with the cDNA library fusions (GFP[1]-cDNA library) and the full-length PKB fusion (PKB-GFP[2]) expressing vectors. The F[1]-cDNA library fusions were transfected in several pools, according to their size. Fraction 1 (cDNAs <0.5 kb) was not used because of the high probability of truncated cDNAs. Fractions 2, 3, and 4 correspond to cDNAs between 0.5–2 kb, 2–3 kb, and 3–4.5 kb, respectively. Controls included: (1) transfection with an empty vector and (2) cotransfection with the GFP[1]-cDNA library fusions and the GFP[2] fragment alone (without fusion) expressing vectors, to exclude cases in which a cDNA library member can nonspecifically induce the folding and reconstitution of GFP from its fragments (identified as gate window M2). A physical interaction between PKB and a prey protein (cDNA encoded protein) induces the folding and reconstitution of GFP from its fragments, generating fluorescence. Positive clones (gate window M3) were collected by fluorescence-activated cell sorting (FACS).

signals could arise due to nonspecific reconstitution of GFP from its fragments. The two controls consisted of cells transfected with empty vector (Fig. 2, top left) versus cells cotransfected with GFP[1]-cDNA library fractions and the GFP C-terminal fragment alone (GFP[2]) (Fig. 2, top right; fraction 2 shown). Cells transfected with empty vector showed a homogeneous population (Fig. 2, top left) whereas a clear population of cells with enhanced fluorescence was detected in the case of GFP[1]-cDNA library fractions and GFP[2] alone (Fig. 2, top right), suggesting that some cDNA library members nonspecifically induce folding and reconstitution of GFP from its fragments (there is no significant *spontaneous* reconstitution of GFP from its fragments, as shown in Fig. 3A, left panel). Separate populations of cells were then defined based on comparison to those with higher fluorescence than untransfected cells (gate window M2) and to those with higher fluorescence than cells coexpressing the GFP[1]-cDNA library fusions with the GFP[2] fragment alone (gate window M3) (Fig. 2). Only 2% of the population of untransfected cells is found in M2 and none are found in M3. For the control cells coexpressing the GFP[1]-cDNA library fusions with the GFP[2] fragment alone, around 10% of the cell population is found in M2 and none are found in M3. The gate window M3 was then defined as the positive population of cells for the sorting (M3 corresponded to less than 5% of the cell population for all the three cDNA fractions screened; Fig. 2, lower panels). Positive cells were sorted by FACS and several thousand clones were recovered. The total DNAs from the M3 pools of positive cells for fractions 2, 3, and 4 were separately extracted and transformed into DH5α bacterial cells, grown on LB-agar/ampicillin plates to select only for plasmids harboring cDNA. We obtained a total of approximately 2500 colonies. Three hundred clones were picked, plasmids were extracted, and interaction of individual proteins with the bait (PKB) was confirmed by cotransfecting individual positive GFP[1]-cDNA fusions with PKB-GFP[2] in COS-1 cells (or with GFP[2] alone, as a negative control) and analysis by FACS. After this second round of selection, plasmid cDNAs for 100 of the positive clones were sequenced.

4.3. Analysis of individual cDNA clones

Among the 100 clones sequenced, 54 yielded readable sequences. Among these, 22 sequences corresponded to genes of potential interest, 5 sequences were identified as contaminants (coding for genomic sequence from macaque and adenovirus), while the rest were determined to have no sequence homology to a gene of known function (20 clones) or encoded potential false-positives (7 clones). Five of the 20 clones of unknown function have human ESTs but no known homologues. We used conservative criteria to flag the potential false-positives based on direct or anecdotal evidence from two-hybrid screens showing that genes for certain classes of housekeeping enzymes and proteins often yield false-positives. These seven potential false-positive clones included a gene related to S10 (ribosomal protein), profoldin I (chaperone), carbohydrate metabolism enzymes including pyruvate kinase, malate dehydrogenase, and aldolase A, an endoplasmic reticulum α-mannosidase, and the ion pump Na–K ATPase subunit β. Most of the 22 promising hits (corresponding to 17 individual genes) could be linked to PKB function on the basis of evidence supporting their role in cellular functions in which PKB is implicated. These include genes involved in cytoskeletal organization, chemotaxis, differentiation, and apoptosis. Below we focus on validation of a cDNA encoded protein that was identified three times in the clones sequenced. This cDNA corresponds to a human gene encoding a protein highly similar to the mouse protein Ft1 (Accession No. Z67963), which we called hFt1 (for human Ft1). The cDNA encoding hFt1 was complete in the three clones isolated. The corresponding mouse *Ft1* gene was found to be deleted in a mouse mutant characterized by developmental abnormalities including fused toes on the forelimbs and thymic hyperplasia in heterozygous animals [28,29].

4.4. First pass functional validation of hFt1

The GFP PCA strategy described here uses direct functional assays to establish the likely biological validity of a bait–prey interaction. As we previously demonstrated, protein–protein interactions that occur within a specific biochemical pathway can be modulated in predicted ways by conditions or molecules that activate or inhibit the pathway and we have demonstrated this for pathways linked to PKB [19]. As a "first-pass" functional validation, we applied this strategy to the newly discovered PKB/hFt1 interaction (Fig. 3). First, the PKB/hFt1 interaction was confirmed by transiently cotransfecting the GFP[1]-hFt1 and PKB-GFP[2] fusions (or with GFP[2] alone, as a negative control) in COS-1 cells and analysis by FACS (Fig. 3A). We next attempted to determine if the PKB/hFt1 interaction is modulated by agents that activate or inhibit PI3K-mediated signal transduction pathways, of which PKB is a downstream effector (Fig. 3B). Specifically, serum and insulin, which activate PI3K-mediated signal transduction pathways, and two specific inhibitors of PI3K activities, wortmannin and LY294002, were tested. As we have previously shown for known PKB interactions [19], the PKB/hFt1 interaction was enhanced 4- to 5-fold after treatment of HEK293T cells with serum or insulin, while both wortmannin and LY294002 inhibited insulin-induced stimulation of the interaction (Fig. 3B, left). These patterns of stimulation/inhibition are consistent with protein–protein interactions activated through PI3K-associated

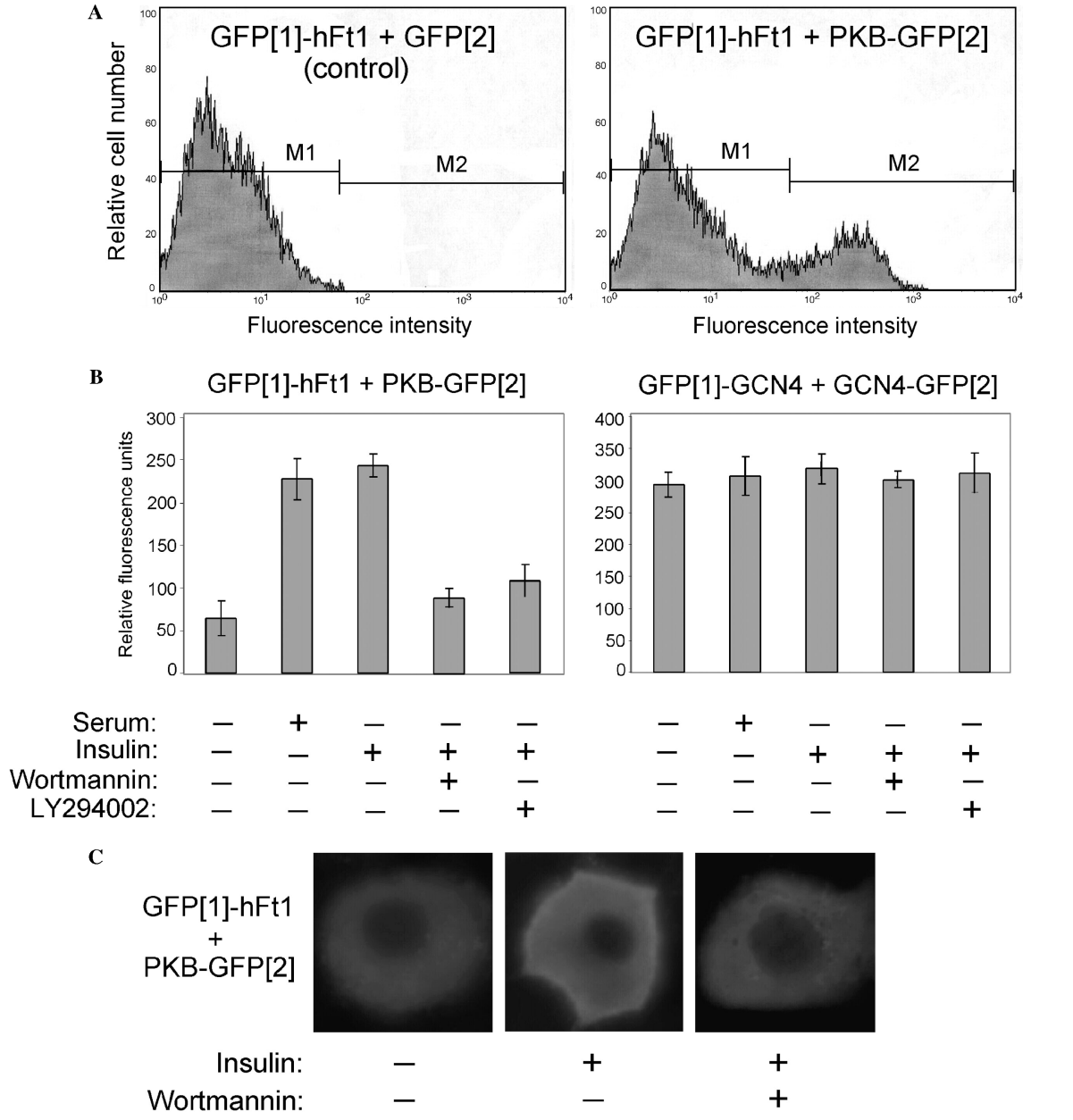

Fig. 3. Biological validation of a newly identified protein–protein interaction with the GFP PCA. (A) Interaction of hFt1 with PKB was confirmed by cotransfecting COS-1 cells with the individual GFP[1]-cDNA fusions coding for full-length hFt1 (GFP[1]-hFt1) and the PKB-GFP[2] fusion, followed by FACS analysis. The physical interaction between PKB and hFt1 induces the folding and reconstitution of GFP from its fragments, generating a fluorescent signal (gate window M2). Cotransfection of cells with GFP[1]-hFt1 fusion and free GFP[2] expressing vectors was used as a negative control. (B) Pharmacological modulation of the PKB/hFt1 interaction. HEK293T cells expressing the GFP[1]-hFt1 and PKB-GFP[2] fusions were serum-starved and untreated or treated with 300 nM wortmannin or 50 μM LY294002 for 60 min. Afterwards, cells were stimulated for 30 min with 10% serum or 20 μg/ml insulin. The relative amount of reconstituted GFP, a measure of the interaction between the fused protein partners, was detected by fluorometric analysis in intact cells. The dimerization of GCN4 leucine zipper was used as a control to assure that cell treatments do not alter protein–protein interactions in a nonspecific way. Fluorescence intensity is given in relative fluorescence units (*y* axis). Error bars represent standard errors of the mean calculated from three independent samples. (C) To determine the cellular location of the PKB/hFt1 protein complex, HEK293T cells were cotransfected with GFP[1]-hFt1 and PKB-GFP[2] fusions and treated with insulin or wortmannin as in (B). Fluorescence microscopy was performed on live cells.

signaling pathways. The dimerization of GCN4 leucine zipper was used as a control to assure that cell treatments do not alter protein–protein interactions in a nonspecific way (Fig. 3B, right). A first step in activation of PKB is its recruitment to the plasma membrane via interaction of its N-terminal PH domain with PIP_3 lipids that themselves are synthesized by PI3K activities [30–32]. Thus, our results also suggest that the hFt1 interaction with PKB may occur at the membrane or be recruited to the membrane with PKB as part of an PKB activation complex.

As a further validation step, we examined the cellular location of the PKB/hFt1 complex and changes in location following perturbation of the PI3K pathway, in intact living cells (Fig. 3C). The interaction between PKB and hFt1 predominantly occurs at the plasma membrane in insulin-stimulated cells, but complexes were also observed in the cytosol (Fig. 3C). The interaction appears to be both disrupted and to dissociate from the membrane after treatment of cells with wortmannin (Fig. 3C). All these results strongly suggest that formation of the PKB/hFt1 complex is induced by stimulation of PI3K-associated signaling pathways. These pathway perturbation and visualization studies serve to support hFt1 as a *bona fide* partner of PKB. For our purposes here, these results illustrate the complete cycle of functional screening by a GFP PCA strategy, from FACS to first-pass functional validation. A detailed study of the role of hFt1 in PKB-mediated processes will be reported in detail elsewhere (I. Remy and S.W. Michnick, submitted for publication).

5. Conclusions

We have presented a general approach to screen expressed cDNAs using PCA. The approach consists of a genome-wide screening of a directional cDNA library for protein–protein interactions, combined with simple functional validation experiments that use the same system. The biological relevance of a newly discovered protein–protein interaction is then tested immediately and preliminary clues about the function of the prey protein are also obtained using this first-pass validation strategy (identification of the corresponding signal transduction pathway). However, a library protein could constitutively interact with the bait and thus not respond to pathway perturbations. Such an interaction would be considered to be a false-positive according to our validation strategy, but could in fact be a biologically relevant bait-interacting protein. Judgment, based on available functional annotations for the library member or homologues and lists of frequent false-positive interacting proteins, should be taken into account in deciding whether to do further validation studies. Obviously, further validation through studies of the endogenous proteins is necessary to fully validate a newly discovered protein–protein interaction. In the present study, in spite of the relatively small sampling of the library presented here, 17 potentially novel proteins that interact with PKB were identified. The screening steps are easily scalable and in principle, tens of thousands of clones could be isolated in a few hours if automation were applied to the follow-up steps of cloning, colony picking, DNA isolation, and sequencing.

The PCA strategy is a completely general approach to cDNA library screening, both in that it can be used in any cell type or organism that can be transfected or transformed and express the fusion proteins and that it can be used to study interactions in any subcellular compartment. These features are due to the fact that unlike other protein–protein interaction-based screening strategies, the fusion proteins themselves represent the entire assay system, requiring no other endogenous cellular machinery [11,33–35]. As a result, cDNAs can be expressed in cells that are relevant models for studying any biochemical pathway and expressed proteins are likely in their native biological state including the correct post-translational modifications. The consequences of these features, combined with the ability to detect pathway-specific perturbations of putatively biologically relevant protein–protein interactions, provide for the first-pass functional validation that we have demonstrated here. For these reasons, the cDNA library screening strategies that we have presented could have wide applications in quantitative analysis and mapping of biochemical pathways and in genome-wide functional annotation efforts.

Acknowledgments

We gratefully acknowledge Serge Sénéchal for his technical assistance in FACS analysis. We also thank Jing Jin and Jim Woodgett for the gift of PKB cDNA and Monique Davies for the pMT3 vector. This work was supported by a grant from the Canadian Institutes of Health Research (CIHR, MOP-42477) to S.W.M.

References

[1] A.D. D'Andrea, H.F. Lodish, G.G. Wong, Cell 57 (1989) 277–285.
[2] H.Y. Lin, X.F. Wang, E. Ng-Eaton, R.A. Weinberg, H.F. Lodish, Cell 68 (1992) 775–785.
[3] A. Aruffo, B. Seed, Proc. Natl. Acad. Sci. USA 84 (1987) 8573–8577.
[4] D. Sako, X.J. Chang, K.M. Barone, G. Vachino, H.M. White, G. Shaw, G.M. Veldman, K.M. Bean, T.J. Ahern, B. Furie, Cell 75 (1993) 1179–1186.
[5] T. Pawson, P. Nash, Genes. Dev. 14 (2000) 1027–1047.
[6] C.R. Weston, R.J. Davis, Science 292 (2001) 2439–2440.
[7] B.L. Drees, Curr. Opin. Chem. Biol. 3 (1999) 64–70.
[8] M. Vidal, P. Legrain, Nucleic Acids Res. 27 (1999) 919–929.
[9] A.J. Walhout, R. Sordella, X. Lu, J.L. Hartley, G.F. Temple, M.A. Brasch, N. Thierry-Mieg, M. Vidal, Science 287 (2000) 116–122.
[10] 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, Nature 403 (2000) 623–627.
[11] S. Fields, O. Song, Nature 340 (1989) 245–246.
[12] T. Ito, K. Tashiro, S. Muta, R. Ozawa, T. Chiba, M. Nishizawa, K. Yamamoto, S. Kuhara, Y. Sakaki, G. Giaever, A.M. Chu, L. Ni, C. Connelly, L. Riles, S. Veronneau, S. Dow, A. Lucau-Danila,

K. Anderson, B. Andre, A.P. Arkin, A. Astromoff, M. El-Bakkoury, R. Bangham, R. Benito, S. Brachat, S. Campanaro, M. Curtiss, K. Davis, A. Deutschbauer, K.D. Entian, P. Flaherty, F. Foury, D.J. Garfinkel, M. Gerstein, D. Gotte, U. Guldener, J.H. Hegemann, S. Hempel, Z. Herman, D.F. Jaramillo, D.E. Kelly, S.L. Kelly, P. Kotter, D. LaBonte, D.C. Lamb, N. Lan, H. Liang, H. Liao, L. Liu, C. Luo, M. Lussier, R. Mao, P. Menard, S.L. Ooi, J.L. Revuelta, C.J. Roberts, M. Rose, P. Ross-Macdonald, B. Scherens, G. Schimmack, B. Shafer, D.D. Shoemaker, S. Sookhai-Mahadeo, R.K. Storms, J.N. Strathern, G. Valle, M. Voet, G. Volckaert, C.Y. Wang, T.R. Ward, J. Wilhelmy, E.A. Winzeler, Y. Yang, G. Yen, E. Youngman, K. Yu, H. Bussey, J.D. Boeke, M. Snyder, P. Philippsen, R.W. Davis, M. Johnston, Proc. Natl. Acad. Sci. USA 97 (2000) 1143–1147.

[13] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, J. Zhang, Y. Ma, S.S. Taylor, R.Y. Tsien, Proc. Natl. Acad. Sci. USA 98 (2001) 4569–4574.

[14] R.J. Kaufman, M.V. Davies, V.K. Pathak, J.W. Hershey, Mol. Cell. Biol. 9 (1989) 946–958.

[15] S.W. Michnick, I. Remy, F.X.C. Valois, A.V. Belisle, J.N. Pelletier, in: J.N. Abelson, S.D. Emr, J. Thorner (Eds.), Methods in Enzymology, vol. 328, Academic Press, New York, 2000, pp. 208–230.

[16] I. Remy, J.N. Pelletier, A. Galarneau, S.W. Michnick, in: E.A. Golemis (Ed.), Protein–Protein Interactions: A Molecular Cloning Manual, Cold Spring Harbor Laboratory Press, 2001, pp. 449–475.

[17] I. Ghosh, A.D. Hamilton, L. Regan, J. Am. Chem. Soc. (2000) 5658–5659.

[18] C.D. Hu, Y. Chinenov, T.K. Kerppola, Mol. Cell 9 (2002) 789–798.

[19] I. Remy, S.W. Michnick, Proc. Natl. Acad. Sci. USA 98 (2001) 7678–7683.

[20] I. Remy, S.W. Michnick, Proc. Natl. Acad. Sci. USA 96 (1999) 5394–5399.

[21] I. Remy, I.A. Wilson, S.W. Michnick, Science 283 (1999) 990–993.

[22] R. Subramaniam, D. Desveaux, C. Spickler, S.W. Michnick, N. Brisson, Nat. Biotechnol. 19 (2001) 769–772.

[23] D.P. Brazil, B.A. Hemmings, Trends Biochem. Sci. 26 (2001) 657–664.

[24] S.R. Datta, A. Brunet, M.E. Greenberg, Genes Dev. 13 (1999) 2905–2927.

[25] E.S. Kandel, N. Hay, Exp. Cell Res. 253 (1999) 210–229.

[26] M.A. Lawlor, D.R. Alessi, J. Cell Sci. 114 (2001) 2903–2910.

[27] D.P. Brazil, J. Park, B.A. Hemmings, Cell 111 (2002) 293–303.

[28] F. van der Hoeven, T. Schimmang, A. Volkmann, M.G. Mattei, B. Kyewski, U. Ruther, Development 120 (1994) 2601–2607.

[29] R. Lesche, A. Peetz, F. van der Hoeven, U. Ruther, Mamm. Genome 8 (1997) 879–883.

[30] C. Belham, S. Wu, J. Avruch, Curr. Biol. 9 (1999) R93–R96.

[31] R.T. Peterson, S.L. Schreiber, Curr. Biol. 9 (1999) R521–R524.

[32] B. Vanhaesebroeck, D.R. Alessi, Biochem. J. 3 (2000) 561–576.

[33] N. Johnsson, A. Varshavsky, Proc. Natl. Acad. Sci. USA 91 (1994) 10340–10344.

[34] H. Laser, C. Bongards, J. Schuller, S. Heck, N. Johnsson, N. Lehming, Proc. Natl. Acad. Sci. USA 97 (2000) 13732–13737.

[35] J. Zhang, S. Lautar, Anal. Biochem. 242 (1996) 68–72.