Chemical genetic strategies to delineate MAP kinase signaling pathways using protein-fragment complementation assays (PCA)

Stephen W. Michnick a,*, Marnie L. MacDonald b, John K. Westwick b

a Université de Montréal, Canada Research Chair, Integrative Genomics, CP 6128, Succ. Centre-ville, Montreal, Canada H3C 3J7
b Odyssey Thera Inc., 4550 Norris Canyon Rd., Suite 140, San Ramon, California 94583, USA

Accepted 10 July 2006

Abstract

Signal transduction pathways mediated by MAP kinases are among the most studied. Direct analysis of MAP kinase pathways has been difficult because some details of MAP kinase signaling cannot be studied in vitro. Here, we describe a strategy for directly analyzing MAP kinase signaling pathways in living cells using protein-fragment complementation assays (PCA) based on intensely fluorescent proteins. The assays allow for spatial and temporal analysis of protein complexes including those that form upstream and downstream from MAPKs as well as complexes of MAPKs with regulator and effector proteins. We describe high-content assays, high-throughput quantitative microscopic methods to follow temporal changes in complex subcellular location and quantity. Spatial and temporal changes in response to perturbations (chemical, siRNA, and hormones) allow for delineation of MAPK signaling networks and a general and high-throughput approach to identify small molecules that act directly or indirectly on MAPK pathways.

Keywords: PCA; MAP kinase; Microscopy; High-context; High throughput; Yellow fluorescent protein; Signal transduction pathways; Subcellular localization

1. Introduction

MAP kinases are among the most studied signaling proteins. Despite the high level of interest in these proteins, direct analysis of most aspects of MAPK biology in living cells has not been possible. Much of our knowledge, therefore, has been derived from in vitro analyses, or from functional assays reporting on a downstream effect of MAPK activity (such as cellular transformation or gene expression). Several dynamic events control MAPK signaling and among the most commonly studied are those of Ras protein activity and activation of the MAPK Raf family. Regulatory signals controlling Ras activity include, first, prenylation, proteolytic processing and methylation at the carboxy terminus of Ras-family proteins regulate localization and activity (reviewed in [1,2]). Second, adaptor proteins and guanine nucleotide exchange factors interact with upstream receptors, and which in turn interact with Ras proteins, regulating their activity via the regulation of localization, formation of specific protein complexes, and GDP/GTP exchange. Third, effector proteins interact with Ras in a nucleotide dependent manner, leading to regulation of downstream signaling events (such as MAP kinase activation) [3]. The number, composition, and sub-cellular localization of Ras-containing protein complexes changes depending on the level of pathway activity. Therefore, the ability to monitor the localization of Ras and Ras-mediated MAPK components is clearly desirable. Given the combinatorial complexity of Ras interactions with cellular structures, activators, and effectors, the added ability to quantify specific complexes would significantly enhance our understanding of these pathways. We describe here a strategy for directly analyzing these events using live cell, high-content protein-fragment complementation assays (PCA). These assays are also amenable to analysis of various events upstream and downstream from
MAPKs, and are not limited to a particular protein class. Therefore, this strategy may be broadly applicable to analysis of MAPK signaling networks.

2. Methods

2.1. Background and general considerations

Principles and basic methods for PCA have been extensively reviewed elsewhere [4–7]. In brief, PCA involves the use of a rationally dissected reporter protein (including enzymes and fluorescent or luminescent proteins). For the purposes of this review, we will focus on assays utilizing fragments of inherently fluorescent proteins. The ends of each reporter cDNA fragment are separately fused, in-frame and with a short flexible linker, to two test proteins that are known (or suspected) to interact. The resultant two cDNA fusion expression cassettes are co-introduced into cells. Following expression, if the two test proteins interact, the fragments of the reporter protein are brought within close proximity and spontaneously re-fold to generate a measurable signal. With fluorescence PCAs, at least three types of events can be observed and quantified—an increase or a decrease in protein complex formation, and/or a change in sub-cellular localization or concentration of the signal. As with full-length fluorescent proteins, one of the key advantages of this strategy is that the location of signals (and thus their cognate protein complexes) can be mapped with high precision. For example, signal localization to organelles, sub-organelle structures such as nucleoli, and structures such as clathrin-coated pits, can be easily resolved and quantified.

The abundance of functional dyes for various sub-cellular structures, coupled with image analysis and de-convolution of multiple wavelength signals, enables co-localization studies.

As with any technique involving expression of exogenous proteins, it is important to consider the level of expression. Expression at levels significantly greater than the endogenous protein can abrogate normal regulatory mechanisms within a pathway, and for some proteins may be toxic. Some strategies for protein complex analysis, such as energy transfer between fluorophores and/or luminophores (FRET, BRET) require high levels of exogenous protein expression. We have found, using intensely fluorescent variants of YFP for PCA engineering, that expressing proteins at or below the endogenous level can still yield readily quantifiable signals [8].

PCA has been employed in a number of common cell lines, including CHO, HEK293, HeLa, Cos, U2OS, HepG2, and Jurkat, as well as transgenic insects and plant and bacterial cells [8–15]. Recent extension of these assays into viral vector systems suggests that any transduceable cell type, including non-dividing cells and stem cells, can be analyzed with this strategy. We describe below the method for transient PCA expression in HEK cells, with concomitant small molecule or siRNA treatment.

2.2. Fragment synthesis and construct preparation

Fusion constructs using cDNAs coding for full-length proteins are generated as described previously [12,16,17] with YFP reporter fragments with the following additional mutations: YFP[1]-(F46L, F64L, M153T) and YFP[2]-(V163A, S175G). These mutations have been shown to enhance chromophore maturation and increase the fluorescence intensity of the intact YFP protein [18]. Each test cDNA can be fused to the reporter fragment at either the amino or carboxy terminus. In addition, either fragment of the reporter can be used for each test protein, yielding four possible fusion constructs for each protein and eight possible combinations between two interacting proteins. In practice, all possible combinations should be tested for a pair of interacting proteins. Pairs yielding sufficient signal and correct sub-cellular localization are chosen for further analysis.

2.3. Cells and transfections

HEK293 cells are maintained in MEM alpha medium (Invitrogen) supplemented with 10% FBS (Gemini Bio-Products), 1% penicillin, and 1% streptomycin, and grown in a 37°C humidified incubator equilibrated to 5% CO₂. Cells are seeded at 7500 cells per well in 96-well plates 20 h prior to transfection, and co-transfected with up to 100 ng total of complementary fusion vectors using Fugene 6 (Roche) according to the manufacturer’s protocol. The amount of each vector used needs to be empirically determined to identify the lowest DNA concentration at which a quantifiable signal is obtained. We have found that transient transfection of many PCAs generates quantifiable signals in the appropriate cellular compartments that respond appropriately to pathway stimulation or inhibition. Predictably, many PCAs generate a more homogeneous signal when engineered as stably transfected cell lines. For high-throughput screening campaigns, engineering of a clonal cell line is therefore advisable.

2.4. Image acquisition and analysis

These assays are amenable to imaging on any platform capable of acquiring fluorescence signals, including microscopes, flow cytometers, and simple plate readers. We focus here on the use of confocal or epifluorescence microscopy systems and recently developed high-throughput derivatives. We have used the Discovery-1 automated fluorescence imager (Molecular Devices, Inc.) equipped with a robotic arm (CRS Catalyst Express; Thermo Electron Corp., Waltham, Mass) as well as the Opera high-throughput confocal fluorescence imaging platform (Evotec Technologies, Hamburg). For the Discovery-1, the following filter sets are used to obtain at least four images of non-overlapping populations of cells per well: excitation filter 480 ± 40 nm, emission filter 535 ± 50 nm (YFP); excitation filter 360 ± 40 nm, emission filter 465 ± 30 nm (Hoechst); excitation filter 560 ± 50 nm, emission filter 650 ± 40 nm.
(Texas Red). A constant exposure time for each wavelength is used to acquire all images for a given assay. For higher throughput analyses, cells are fixed and stained 48 h after transfection as described above. However, live cell imaging is possible, enabling real-time tracking and ‘movies’ of signalling dynamics.

Raw images in 16-bit grayscale TIFF format are analyzed using modules from the ImageJ API/library (http://rsb.info.nih.gov/ij/, NIH, MD). First, images from each fluorescence channel (Hoechst, YFP and Texas Red) are normalized using the ImageJ built-in rolling-ball algorithm [19]. Since each PCA generates signal in a specific subcellular compartment or organelle, and treatment with a small molecules or siRNA may affect a change in complex localization or signal intensity, different algorithms are required to accurately quantitate fluorescent signals localized to the membrane, nucleus or cytosol. Each assay was categorized according to the subcellular localization of the fluorescent signal, and changes in signal intensity across each sample population were quantified using one of multiple automated image analysis algorithms. A detailed description of specific algorithms is beyond the scope of this review, but it should be noted that, as with most cell-based assays, the signals across the population do not form a normal distribution. We and others have found that the optimal statistical approach to these data utilizes non-parametric [20].

3. Applications and examples

3.1. Mapping signaling complexes, pathways, and networks

PCA has been used extensively to identify novel interacting partners for known signaling proteins and to map signaling networks. For example, we have devised a PCA-based system to screen cDNA libraries to identify novel proteins implicated in signaling by the protein kinase PKB [12,13,16]. We have also recently reported a screening of 107 common drugs or small molecules against several PCA reporters for MAK pathways [8]. It should be noted that these studies map protein complexes, not just binary protein interactions, as a positive signal is not absolutely dependent on a direct interaction between two proteins. If proteins are in sufficiently close proximity in the context of a larger protein complex a positive signal can result. The absolute proximity of two proteins required to generate a signal is a function of their structure and the length of the polypeptide linkers between proteins and the PCA reporter fragments (10 amino acids in these studies) [21], but it should be noted that a positive PCA signal is essentially an “all or none” phenomenon. That is, PCA results from folding, not association of the reporter protein fragments. Protein folding is a highly co-operative process in which, once conditions are created for a polypeptide to fold (in the case of PCA, when the fragments are brought together by the interacting proteins) folding will proceed to completion spontaneously. A further consequence of this all or none behavior is the extremely high dynamic range of PCA compared to FRET-based interaction assays, regardless of the relative expression levels of the proteins of interest [22,23].

Fig. 1 provides examples of the wide range of small MAP kinase-related signaling events that can be probed with PCA. Diverse processes and target classes can be visualized and quantified, including upstream activators of Ras protein regulation (integrin dimerization; ITGα5/ITGβ1), GTPase/scaffold interactions (Cdc42/WASP), guanine nucleotide exchange factor/GTPase complexes (Vav/Cdc42), GTPase/effectors complexes (Cdc42/Pak4), and downstream signaling events (Raf, MEK, Elk, and SRF-containing protein complexes; Fig. 1). Also notable is the fact that multiple interactions involving the same protein can be probed, for example complexes of Raf with modulatory proteins or chaperones (14-3-3 or Cdc37; Fig. 1) or downstream effectors (Raf/Mek; Fig. 3).

These examples span multiple steps in MAPK and related signaling pathways and probe signaling events specific to distinct sub-cellular compartments. For example, a major fraction of ERK protein kinase pools reside in the cytoplasm of these cells in a complex with its activators MEK1 and MEK2 (Fig. 3 and data not shown), but by using an assay such as ERK/ELK or ELK/SRF, events specific for nuclear ERK can be visualized and quantified (Fig. 1). Importantly, the response of the common protein to small molecules or siRNAs is often dependent on the specific context in which it is probed.

3.2. Probing pathway architecture and signaling dynamics

A unique feature of PCA is the ability to capture the dynamics of complexes in their native context. In conjunction with high-content cell imaging, detailed information on the levels and localization of signaling complexes is obtained. Thus the tools can be used not just to identify the components of signaling networks, but also the activity of pathways within these networks. The examples described below demonstrate that it is possible to probe discreet signaling nodes for agents that act directly on the signaling proteins of which it is composed, as well as for targets and agents that act “upstream” of the node.

Because PCA assays are generally performed as plasmid-based transfections, the strategy works well in conjunction with other genetic probes requiring transfection, such as dominant negative/active mutants and RNAi. For example, the JNK2/e-Jun PCA was co-transfected with a constitutively active signaling protein, pDCR.RasV12 [24] or the corresponding empty vector (pDCR) (Fig. 2). A dramatically higher level of fluorescence was seen in cells co-transfected with only 1ng of Ras cDNA (activated G12V mutant). The JNK-Jun pathway is known to respond to Ras activation [25], and this response was robust, occurring over a range (30–100ng) of PCA vector transfection levels (Fig. 2).

PCA and siRNA strategies are highly complementary. The co-transfection frequency—determined with fluorescently labeled siRNAs co-transfected with fluorescence
PCAs—approaches 100% (data not shown). For assessment of siRNA-mediated target knock-down, siRNAs are transiently co-transfected with PCA plasmid vectors pairs. Lipofectamine 2000 is the optimal transfection reagent for siRNA/PCA co-transfections. siRNA SMART pools designed to target human genes and two ‘GC-matched’ non-specific siRNA pools were designed and obtained from Dharmacon (Boulder, CO). siRNA pools directly targeting one of the components of a PCA used in the study serve as a control for siRNA efficacy (such as H-Ras siRNA and the H-Ras/Raf PCA; Fig. 3). siRNA pools are generally designed to target endoge-

Fig. 1. Examples of MAP kinase signaling complexes probed with PCA. HEK293 cells were transfected with the indicated pairs of PCA vectors. Forty-eight hours after transfection, cells were fixed and stained with Hoechst, and images were captured on the Discovery-1 as described.

Fig. 2. Formation of a JNK2/c-Jun signaling complexes following co-transfection of Ras(G12V). HEK 293 cells were transfected with JNK2 and c-Jun PCA fusion vectors (30, 60, or 100 ng total PCA vector DNA, as indicated) along with 1 ng of empty vector (pDCR) or 1 ng of PDCR Ras G12V. Cells were fixed, stained, and imaged as described in the text.
nous proteins, allowing analyses of the effects of endogenous protein knockdown on pathway activity. To determine the optimal siRNA concentration at which targets are modulated but non-specific effects are minimized, we evaluated the effects of siGFP (Dharmacon) and the non-specific siRNA controls on four different PCAs. Under the conditions listed here, 40 nM siRNA was found to be optimal. While higher concentrations (100 nM) are typically reported, and lead to higher levels of target protein knock-down, we found evidence of widespread non-specific activity at these concentrations (data not shown). siRNA effects on PCA activity are quantified as described above, and compared to the pooled mean fluorescence of the corresponding non-specific siRNAs. Each 96-well plate should contain five internal controls: mock (no PCA), no siRNA, non-specific siRNA controls (such as control IX and control XI; 47% and 36% GC content, respectively, Dharmacon, Boulder CO), and a PCA-specific control (to confirm degree of stimulation for assays treated with agonists).

Fig. 3 illustrates several PCAs that were co-transfected with a siRNA pool targeting H-Ras or a GC-matched siRNA control. As expected, H-Ras siRNA knocked down complexes of H-Ras with Raf, Raf with MEK1, and MEK1 with ERK2. It is notable that the Ras/Raf complexes are distinctly localized at the plasma membrane (Fig. 3, top left image) whereas the Raf/MEK and MEK/ERK complexes are localized in the cytoplasm. Raf proteins associate with the effector domain of active Ras proteins, an event known to occur at the plasma membrane [26]. Thus, by probing the Ras/Raf complex (as opposed to, for example, fluorescently labeled Raf protein alone), this assay focuses exclusively on active signaling complexes. In addition, while MEK/ERK complexes were seen exclusively in the cytoplasm, complexes of ERK with transcription factor substrates or complexes of these substrates with other proteins (such as ELK/SRF complexes) are seen exclusively in the nucleus (Fig. 1, and data not shown). Thus the combination of complex-based assays and high-content image analysis can yield information on specific steps in a pathway, not just the endpoint.

Other events known to be downstream of Ras activation are inhibited following cell treatment with Ras siRNA. For example, c-Jun is phosphorylated and activated (and hence complexes with the prolyl isomerase Pin1) following activation of the Ras pathway [27]. In logarithmically growing cells, an appreciable level of Pin1/Jun complexes is evident and exclusively localized in the nucleus, as expected (Fig. 3). If an activated allele of H-Ras is co-transfected with the PCA pair, the level of nuclear Pin1/Jun complexes increases and also adopts a distinct pattern suggesting enhanced interaction with chromatin. In each case, co-transfection of siRNA targeting Ras clearly diminishes the levels of Pin1/Jun complexes, indicating that endogenous H-Ras is involved in the pathway leading to formation or stability of these complexes. Several PCAs in the cell cycle control/DNA damage response paths were also regulated by treatment with H-Ras siRNA, indicating that Ras controls pathways leading to these protein complexes (p53/Mdm2, Chk1/Cdc25C, and p53/p53; Fig. 3). Thus, both expected and unexpected connections between pathways can be visualized and quantified with this strategy.
3.3. HTS and probing novel targets

The strategy described here is not limited to traditional measures of protein level or post-translational modification, and therefore adds extensively to the list of potential assays for pathway mapping. In addition to enabling direct probing of signaling complexes (Fig. 3) we demonstrate a diverse range of protein interactions that can be probed with this approach. Assays representing a wide range of protein classes have been constructed, including kinase/kinase complexes (e.g., Raf/MEK and MEK/ERK; Fig. 3), GTPase/scaffold complexes (e.g., Cdc42/WASP; Fig. 1), complexes with modulatory proteins (e.g., Raf/14-3-3) and complexes with chaperones and co-chaperones (e.g., HSP90/MEK and Raf/Cdc37; Fig. 1).

In addition to traditional signaling protein complexes classes, novel interactions can be probed. For example, the direct ubiquitination or sumoylation of specific target proteins can be visualized and quantified (e.g., SRF/SUMO; Fig. 1, and data not shown). Other complexes related to proteasome-mediated regulation can also be probed, such as the interaction of E3 ubiquitin ligases with their client proteins (Mdm2/p53; Fig. 3). Other enzyme/substrate interactions, such as the prolyl isomerase interaction with its substrate c-Jun, provide a unique assessment of this pathway (Fig. 3). Finally, receptor mediated events, including GPCR, receptor tyrosine kinase, and integrin signaling, are amenable to analysis (e.g., ITGα5/ITGβ1; Fig. 1).

For small molecule studies in transiently transfected cells, compounds or vehicle controls are generally added 24 or 48 h after transfection, and cells are incubated for various periods of time before imaging. Small molecule-induced changes in protein complexes have been visualized and quantified within seconds of treatment, and can be followed over the course of several days. Following small molecule or other treatments, cells can be stained with 33 μg/ml Hoechst 33342 (Molecular Probes) and 15 μg/ml TexasRed-conjugated Wheat Germ Agglutinin (WGA; Molecular Probes) to localize nuclei and plasma membrane, respectively. Cells can be imaged live, or fixed with 2% formaldehyde (Ted Pella) for 10 min. When fixed, cells are subsequently rinsed with HBSS (Invitrogen) and maintained in the same buffer during image acquisition. With the instrumentation and liquid handling capabilities described here, PCAs have been used to screen 20,000 compounds per day.

3.4. Pharmacological profiling

The diversity of the assays described here, and the richness of data provided by high-content imaging, provides a unique opportunity for gaining a better understanding of signaling pathways in general. We have found that the combination of a broad panel of assays and equally broad panel of siRNAs and small molecules can yield novel information regarding pathway architecture [8]. Challenges include the need for expensive imaging platforms, high-throughput liquid handling robotics and extensive computational capabilities. For example, profiling studies involving hundreds of assays and dozens of small molecules routinely require terabytes of image storage capacity. Due to these instrumentation and computational requirements, and the necessity for engineering a broad panel of diverse assays, broad target, and small molecule profiling is currently beyond the scope of most academic laboratories. However, focused use of these tools is achievable using generally available laboratory equipment (such as fluorescence microscopes and plate readers). In addition, the existence of high-throughput, high-content imaging platforms appears likely to become widespread in larger university core laboratories.

4. Conclusions

There is a growing appreciation for the fact that signaling events occur in the context of large, multi-protein complexes. Changes in protein complex levels and localization in response to target modulation or small molecule treatment reflect the activity of the pathways in which these complexes reside. Despite this knowledge of signal transduction mechanics, most of our understanding of pathway architecture and activity is based on measurable transcriptional or post-transcriptional changes, or from proteomic approaches using preparations from disrupted cells. Application of strategies such as those described here should improve our knowledge of how these pathways operate in living cells, and may provide novel approaches for the identification of therapeutic agents.

Acknowledgments

The authors thank Jane Lamerdin and Jennifer Dias for generating the data used in this review. We thank Anastasia Khvorova, Dharmacon, for the siRNA pools used in this study.

References