ELSEVIER

Protein fragment complementation strategies for biochemical network mapping

Stephen W Michnick

The organization of biochemical networks that make up the living cell can be defined by studying the dynamics of protein–protein interactions. To this end, experimental strategies based on protein fragment complementation assays (PCAs) have been used to map biochemical networks and to identify novel components of these networks. Pharmacological perturbations of the interactions can be observed, and the resulting pharmacological profiles and subcellular locations of interactions allow each gene product to be 'placed' at its relevant point in a network. Network mapping by PCA could be used with, or instead of, traditional target-based drug discovery strategies to increase the quantity and quality of information about the actions of small molecules on living cells and the intricate networks that make up their chemical machinery.

Addresses
Département de Biochimie, Université de Montréal, CP 6128, Succursale Centre-Ville, Montréal, Québec H3C 3J7, Canada
e-mail: stephen.michnick@umontreal.ca

Current Opinion in Biotechnology 2003, **14**:610–617

This review comes from a themed issue on
Chemical biotechnology
Edited by Frances H Arnold and Anton Glieder

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2003.10.014

Abbreviations

DHFR	dihydrofolate reductase
GFP	green fluorescent protein
PCA	protein fragment complementation assay

Introduction

The emerging offspring of the genomics revolution, variously called proteomics, functional genomics, chemical genetics or systems biology, can be attributed an overall aim. As only a fraction of gene functions can be inferred from primary gene sequences, we need to develop strategies to define gene function on a large scale. What approaches can we use to meaningfully ascribe function to genes? Systems biology and chemical genetics researchers seek a deeper appreciation of the biochemical organization of living cells and the molecular schemes that all living things share, as well as those things that make individual cells and organisms unique. What we need are conceptual and experimental approaches that will tell us how, when, where and under what circumstances, proteins are processed, become modified to be activated or inactivated, how they are destroyed and, central to all of these, what other proteins they interact with. We have devised an approach that is aimed at addressing most of these issues, a general strategy to probe the dynamics of protein–protein interactions in intact living cells, based on a technology we call protein fragment complementation assays (PCA). PCA strategies can be used to identify novel protein–protein interactions and can also be used to study perturbations of interactions by chemical or physical means. Such perturbation studies can be used to interpret the organization of biochemical networks. In an inverse sense, chemical perturbations of biochemical networks can be monitored directly or indirectly by PCA, allowing one to ask how small organic molecules effect specific cellular processes. This puts the PCA strategy into the class of chemical genetics tools that allow one to determine which networks small molecules act on.

As a tool to map sites of action of small molecules, the PCA strategy has considerable potential in the discovery of therapeutic agents. A problem with target-directed pharmaceutical discovery is that exquisitely selective small-molecule inhibitors that bind to a potential therapeutic target may, nevertheless, have completely unexpected effects when applied in a living cell. How can we quickly determine if this is the case? To isolate the mode of action of a small molecule to a specific point in the complex biochemical networks that make up a living cell, one needs a way to directly probe the networks and subnetworks themselves. A logical way to do this is to monitor the interactions of network components and to assess how these interactions are effected by perturbations of drugs that act upstream of the interactions. The PCA strategy can be applied to this problem to both validate that a small molecule acts on the network of interest, while simultaneously determining whether such molecules may also act on other, unsuspected networks. This latter application is an example of what is called in pharmaceutical research a 'fail-fast strategy': one that allows a compound whose actions are more complicated or more deleterious than first thought to be excluded from further study.

In this review I describe the many uses of PCA strategies and highlight their potential as a drug-discovery tool.

Protein fragment complementation strategies: a cellular 'spectroscopy'

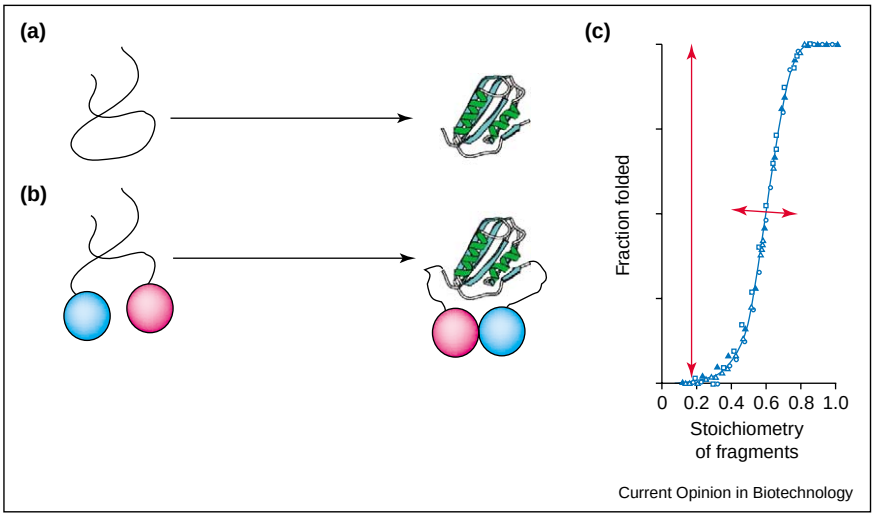
What does it mean to 'map' a biochemical network? Molecular spectroscopy can serve as a useful working

analogy. Molecular spectroscopies are aimed at determining or mapping the covalent and three-dimensional structures of molecules. All molecular spectroscopic methods rely on two components: a general energy transition with distinct and distinguishable resonance conditions for different atoms in different environments and methods to probe the resonance behavior when subject to specific perturbations. It is a bit of a stretch, but not unreasonable to say that cell biologists do the same thing to understand biochemical networks. For example, the advent of DNA microarray technologies has changed the manner in which we view biochemical networks [1–6]. The practical monitoring of changes in expression of complete genomes or large subsets of genes has allowed researchers to begin to scrutinize in some detail the evolution of genetic programs and sometimes, by inference, the underlying biochemical networks that control these programs. The most well-described analyses of gene expression have been performed for *Saccharomyces cerevisiae*, for which the entire genome has been sequenced and microarrays representing all predicted genes have been available for some time [7]. Complete analyses of the cell cycle and responses of the organism to several perturbations have been explored. It is possible to conceive of mapping the output of a genetic program (gene expression changes) back to the organization of the biochemical networks underlying these changes. However, this 'transcriptocentric' approach to mapping networks does not help us to side-step an obvious problem. If you do not know the details of the underlying machinery leading to specific outputs of a system, then any new insights from analyzing outputs remain inferences that must still be tested directly. It was with these challenges in mind that our laboratory developed the PCA strategy.

In the general PCA strategy, an enzyme or fluorescent protein (the reporter) is rationally dissected into two fragments and the fragments fused to two proteins that are thought to bind to each other. Folding of the reporter enzyme from its fragments is catalyzed by the binding of the test proteins to each other, and is detected as reconstitution of enzyme activity (Figure 1a,b). The PCA strategy takes advantage of the spontaneous all-or-none nature of protein folding and the design of fragments follows from basic concepts of protein engineering. The all-or-none folding of reporter protein from fragments means that the generation of signal by complementation has an enormous dynamic range over very narrow conditions, unlike fluorescence resonance energy transfer techniques (Figure 1c) [8]. Since first demonstrating the principle, this approach has been extended to several enzymes, including dihydrofolate reductase, glycinamide ribonucleotide transformylase, aminoglycoside kinase, hygromycin B kinase, TEM β-lactamase, green fluorescent protein (GFP) and firefly and renilla luciferases [9•,10••,11,12,13•,14••,15–20,21•].

A crucial feature of PCA fragments is that they cannot fold spontaneously [22]. If this occurred, PCA simply would not work. Spontaneous folding would lead to a false-positive

Figure 1



General principles of PCA. **(a)** Proteins fold from a disordered polypeptide into a unique three-dimensional structure encoded by the amino acid sequence. **(b)** The unimolecular protein folding process is mimicked by dissecting the polypeptide into two complementary fragments fused, respectively, to two interacting proteins (blue and pink balls). The interaction of the proteins to which the fragments are fused recreates the unimolecular folding conditions and, thus, the unique three-dimensional structure of the protein 'reporter' is created. If the folded polypeptide is an enzyme or fluorescent probe binding protein, a reconstitution of catalytic or binding activity reports the interaction of the proteins to which the complementary fragments are fused. **(c)** Protein folding from fragments is, like the unimolecular process, a highly cooperative all-or-none process. An increase in the quantity of one PCA pair relative to another (horizontal axis) results in a huge change in the fraction of reconstituted reporter (vertical axis, dynamic range indicated by arrow) over a very narrow range of concentrations (horizontal arrow).

signal; a situation that would hopelessly confound the interpretation of library screens *in vivo*, where a more or less positive result could be due to differences in expression levels, solubility, or proteolytic stability rather than a protein–protein interaction [11,15]. In contrast to PCAs, there are assay systems based on β-galactosidase and split inteins that resemble PCA in a superficial way, but which are conceptually and practically different, being based on spontaneously associating subunits of the enzymes fused to interacting proteins [23,24]. The central problem here is that subunits, even if weakly associating, are always capable of doing so to some extent, meaning that there is a constant background of spontaneous assembly.

Unique features of the PCA strategy

The PCA strategy exhibits several unique features of key importance. First, molecular interactions, including complete cellular networks, are detected directly — not through secondary events such as transcription activation. Second, genes are expressed in the relevant cellular context, reflecting the native state of the protein with the correct post-translational modifications. Third, events induced by hormones or growth factors can be detected, providing target validation by linking specific interactions to specific networks. For example, using insulin-sensitive cells, molecular interactions stimulated by insulin can be detected and quantitated. Finally, the subcellular location of protein interactions can be determined, whether in the membrane, cytoplasm or nucleus, and the movement of protein complexes can be visualized.

In addition to the specific capabilities of PCA described above, some special features of this approach make it appropriate for genomic screening of molecular interactions. PCAs are not a single assay but a series of assays; an assay can be chosen because it works in a specific cell type appropriate for studying interactions of some class of proteins. PCAs are also inexpensive, requiring no specialized reagents beyond those necessary for a particular assay and off-the-shelf materials and technology. In addition, they can be automated and high-throughput screening employed. Because these assays are designed at the level of the atomic structure of the enzymes used, there is additional flexibility in designing the probe fragments to control their sensitivity and stringency. Lastly, PCAs can be based on enzymes for which the detection of protein–protein interactions can be determined differently; for example, using dominant selection or production of a fluorescent or colored product.

Mapping the biochemical machinery of cells

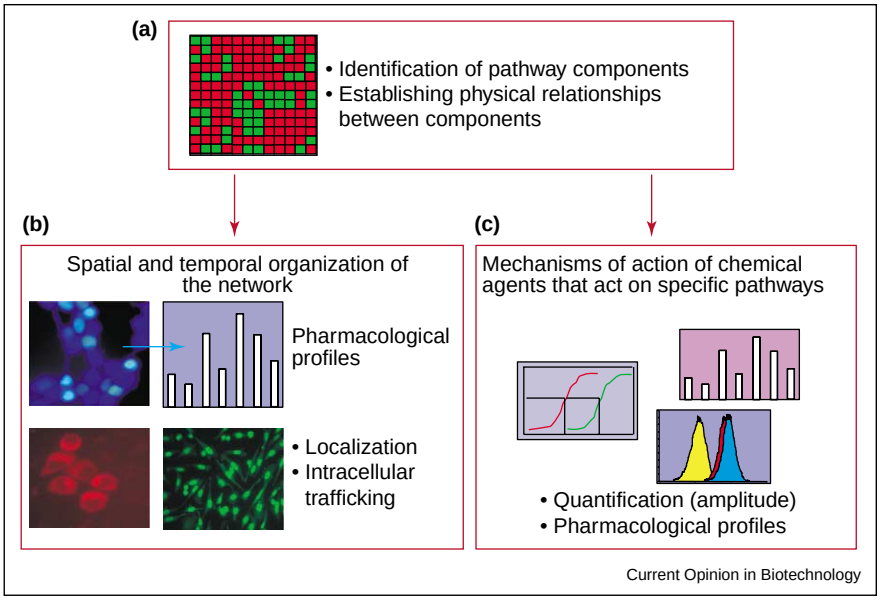
What is it about PCA that makes it ideal for mapping biochemical networks? The reason is a fundamental observation about the nature of biochemical processes. Biochemical processes are mediated by dynamic, noncovalently associated multienzyme complexes that provide the most efficient chemical machinery. The substrates and products of a series of reaction steps are transferred from one active site to another over minimal distance with minimal diffusional loss of intermediates and in chemical environments suited to stabilizing reactive intermediates [25•,26,27]. Biochemical 'networks' are networks of dynamically assembling and disassembling protein complexes. A meaningful representation of a biochemical network in a living cell would therefore result from a step-by-step analysis of the dynamics of individual protein–protein interactions in response to perturbations that impinge upon the network under study and the time and spatial distribution of these interactions. How then can PCA help us to study these dynamic assembly events? Returning to the molecular spectroscopy analogy, in a PCA-based biochemical network mapping project specific pairs of proteins that interact at strategic points in a network serve as 'sentinels' for the state of the network under different conditions. Cells containing PCA sentinels are treated with agents (chemical inhibitors small interfering RNAs, hormones, etc) that would be thought to perturb the biochemical network under study. A change (or lack of change) in the PCA sentinels reporter signal would then suggest the relationship between the point of action of the perturbing agent (say some enzyme in the network) and the sentinels. For instance, if an enzyme were inhibited with a small molecule and the sentinel signal decreased, we could hypothesize that this enzyme must be somehow positively coupled to the function of the sentinel proteins. A series of perturbations at different points in the network would result in a pattern of responses or 'pharmacological profile', as detected by PCA, which should be consistent with the response of the network under study. We have demonstrated the feasibility of the approach to measure the direct induction of protein–protein interactions [12,28] and the perturbations of protein interactions by drugs or hormones acting at steps remote from the interactions studied [14••]. Interactions of protein components of a network should also take place in specific subcellular compartments or locations consistent with the function of the network. The combined pharmacological profiles and subcellular interaction patterns serves then to describe a biochemical network (Figure 2).

Having described the network mapping process, there are immediate applications of this approach to expanding network structure by adding novel components, discussed below.

Identification of components of biochemical networks

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

Figure 2



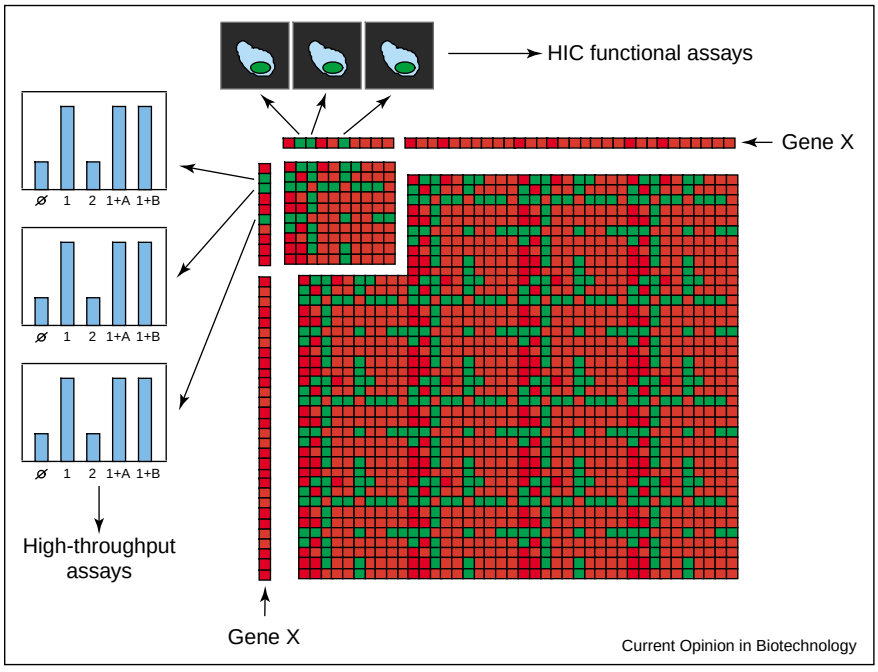
Components of pathway mapping by PCA. **(a)** Matrix representing interactions between component proteins of a biochemical network detected by a PCA screen. The matrix is symmetrical and each protein in a network is screened against every other. A green off-diagonal square means the two proteins interact, red means no interaction is detected. **(b)** Interactions observed in the screen (a) are tested for characteristic responses to pharmacological agents that act on the biochemical network and a 'pharmacological profile' for each interaction can be constructed (histogram). Further, interactions can be localized to specific cellular compartments with some PCAs (e.g. DHFR, GFP). **(c)** Quantitative pharmacological analysis of responses of some interactions to specific network-perturbing agents can be performed.

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 the selection of clones and validation of their biological relevance with the same assay [29–32]. 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, which have very broad substrate specificity and bind to many proteins or protein domains when studied out of their appropriate context in intact living cells [27,33]. In the absence of simple and specific assays, researchers have turned to strategies that use some general functional properties of proteins as readout. The best example of this is the screening for protein–protein interactions between proteins of known function against cDNA libraries as exemplified by the yeast two-hybrid expression cloning strategy [34–40]. 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 which interacts with a bait protein of known function is likely to be involved in specific cellular functions. Thus, while a two-hybrid approach achieves extraordinary generality, no information can be obtained about biological relevance of the interaction. The generality of a two-hybrid approach and the specificity of functional screening can be achieved with a PCA strategy (Figure 3). The strategy consists of two steps. The first step is a large-scale screening of 'bait' proteins (X) fused to one PCA reporter fragment against a cDNA library fused to the complementary reporter fragment. In the second step, positive hits from the screen are directly functionally 'validated' by testing for perturbations of the interaction, as measured by PCA, with agents that act on the biochemical network in which the bait protein is known to participate. A cDNA library screening strategy using a GFP-based PCA was described in which we successfully identified novel substrates and regulators of the serine/threonine protein kinase PKB/Akt [41••].

Chemical genetics and PCA

As much as we want to know the components of biochemical networks and how they interact with each other, we also want to know how agents that perturb these networks exert their effects. Even if we know that a compound acts very specifically on some target protein *in vitro*, this in no way assures that the consequences of the actions of the molecule on the target in the cell can be predicted. Figure 4 shows examples of how this problem can be addressed by PCA. In the first example, a small molecule has been shown to act on a specific target, but we want to know if the predicted outcome on the network in which the target protein participates is perturbed in a

Figure 3



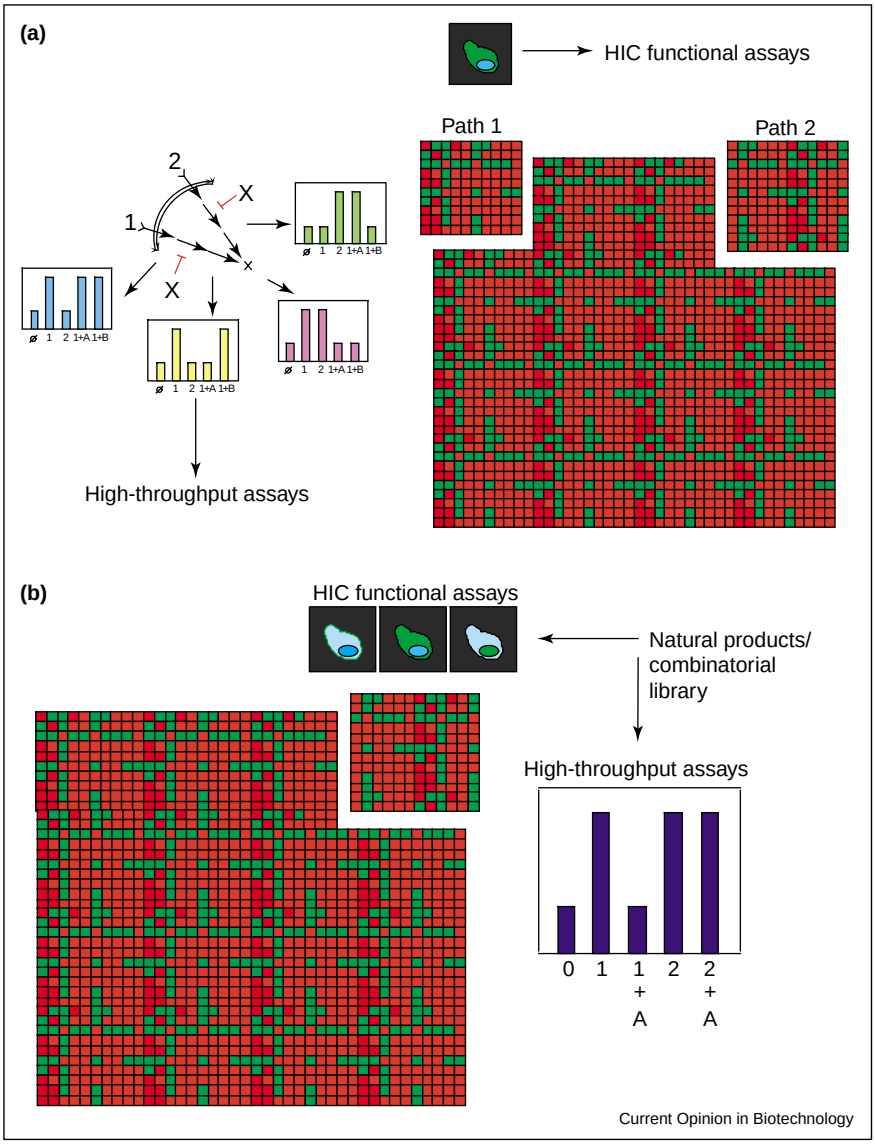
Applications of PCA in expression cloning. Screening of a bait (gene X) against a matrix of potential interactions (large matrix; could represent a cDNA library of interacting proteins or some defined set of genes that have already been screened as in Figure 2a) reveals three hits in the submatrix that represents a known biochemical network. Testing for perturbations (left-hand side) with known activators (1) and potential inhibitors of the network establishes whether interactions are perturbed in a manner that would be predicted for the subnetwork in which hits are observed for gene X. For example, in the biochemical network in which gene X interacts, the network is known to respond to an activator (1) over background (Ø), but not to activator 2 or inhibitors A and B. The physical location of complexes (top) provides a high information content (HIC) assay for whether an interaction occurs in a valid cellular compartment (i.e. that in which the rest of the component proteins of the subnetwork are expressed).

predicted way (Figure 4a). We thus test the compound against a series of PCA sentinels that report on the state of the network we predict will be affected. The compound is also tested on a series of control networks. If the molecule acts as predicted, only the target network is affected; however, if sentinels in other networks report perturbations, then the specificity of action of the molecule must be re-evaluated. In a second example, a novel compound from a natural product or synthetic library has been shown to cause some profound phenotype in a cell and we want to quickly link the observed phenotype to effects on known networks that are likely to be affected. The same strategy is used as above, but in this case we hypothesize that to obtain the observed phenotype one of several networks must be affected (Figure 4b). A third example, is one in which the specific target of a compound is known, but paradoxically, the predicted effects on the cell do not occur because some other branch in the biochemical network compensates for the inhibition of a target. There are numerous examples of such paradoxes (e.g. [42]) and we described one such paradox in the actions of the natural product rapamycin [14••]. These also represent cases in which a 'two hit' strategy of knocking out both the first target and a second target in the compensating branch might produce a profound affect, while hitting either target individually would do nothing to the cell.

The strategies described above refocus chemical genetics and therapeutic discovery from protein targeting to 'network-based targeting'. There are some considerable advantages of network targeting that are summarized in Box 1.

The biochemical network-based strategy could achieve enormous economies of scale in the drug discovery process. By judiciously creating PCA sentinel probes for specific networks, with a few assays one can achieve the monitoring of many other potential protein targets. The rest of the proteins that are involved in the network are simultaneously monitored, as the sentinels detect what is happening to all members of the network to which each sentinel is linked, directly or indirectly. Thus, for example, a screen for 10 networks with 20 proteins per network would capture any of 200 potential drugable targets. The upshot of this remarkable economy of scale is that, in principle, we can do what an entire pharmaceutical company effort would accomplish in a single year, but in fact yield even more validated hits from fewer screens, in an effort in which prior knowledge of

Figure 4



Applications of PCA in chemical genetics and reverse chemical genetics. **(a)** A compound (X) that is a known inhibitor of a specific target in one pathway (Path 1) turns out to also act on a target in a second pathway (Path 2) when screened against several networks (large matrix), as shown by pharmacological profiles in which both pathways are affected. **(b)** Screening of natural product or synthetic libraries produces an interesting effect on cells, which can then be mapped to a specific biochemical network. A screen of all networks probed by PCA sentinels reveals that one subnetwork is responsive at a specific point in the network (upper right submatrix of complete matrix of network sentinel reporters), as revealed by a pharmacological profile in which the compound (A) inhibits subnetwork (1), but not a different subnetwork (2). HIC, high information content.

Box 1 Advantages of network targeting.

1. Optimal target(s) for a network of interest can be identified in a simple cell-based assay.
2. The evaluation of off-pathway, non-specific or potential toxic effects can be performed with the same assays, thereby circumventing expensive mid-stage to late-stage clinical failures.
3. The development of PCAs are simple and necessitate neither the production of recombinant proteins nor protein purification, as is required for the development of large-scale *in vitro* based assays of drug targets; often an intractable problem.
4. PCAs enable the analysis of combinations of drugs that may impact multiple targets simultaneously. The 'drug cocktail' is becoming particularly relevant in the oncology arena, where recent data demonstrated highly synergistic effects of kinase inhibitors (e.g. in combination with traditional cytotoxic chemotherapeutic agents).

the most drugable step in any specific network is not necessary. Furthermore, because the primary screens are in living cells, small molecule hits are shown to be membrane permeable and therefore are more likely to be bioavailable. As described above, our approach enables an understanding of the mechanism of action of the compound in the context of the complex networks of the living cell.

Finally, as important as the successful identification of positive hits with promising pharmacological effects, is the rapid exclusion of compounds that could have deleterious effects. Success in this effort is called a 'fail fast' strategy, which is built into the PCA process. Here, we fail fast by screening a positive hit against networks in which significant effects could result in serious secondary off-pathway effects. Thus, by pinpointing off-pathway effects early in the discovery process, compounds can be flagged for careful re-evaluation, preventing a potentially expensive rush to further pre-clinical, or worse, clinical studies.

Conclusions and perspectives

I have attempted here to summarize PCA-based strategies and their application to mapping biochemical networks and identifying novel components of such networks. Pharmacological perturbations of interactions can be observed, even if the site of action of the perturbant is distant from the interaction being studied. Resulting pharmacological profiles and subcellular locations of interactions allow us to 'place' each gene product at its relevant point in a network. Although specific chemical inhibitors such as those described may not be available in many cases, other perturbants could be used to generate a functional profile, including dominant-negative forms of enzymes, receptor- or enzyme-specific peptides or siRNAs. The ability to monitor the network in living cells containing all of the components of the network studied can reveal hidden connections. This approach will allow for a full investigation of the global organization of biochemical networks in cells and will enable fundamental questions about how biochemical machineries are organized to be addressed [43]. Finally, network mapping by PCA could be used alone or in conjunction with traditional target-based drug discovery strategies to increase the quantity and quality of information about the actions of small molecules on living cells and to reveal the intricate networks that make up their chemical machinery.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO, Herskowitz I: **The transcriptional program of sporulation in budding yeast**. *Science* 1998, **282**:699-705.
2. Holstege FC, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA: **Dissecting the regulatory circuitry of a eukaryotic genome**. *Cell* 1998, **95**:717-728.
3. Hughes TR, Marton MJ, Jones AR, Roberts CJ, Stoughton R, Armour CD, Bennett HA, Coffey E, Dai H, He YD *et al.*: **Functional discovery via a compendium of expression profiles**. *Cell* 2000, **102**:109-126.
4. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L: **Integrated genomic and proteomic analyses of a systematically perturbed metabolic network**. *Science* 2001, **292**:929-934.
5. Roberts CJ, Nelson B, Marton MJ, Stoughton R, Meyer MR, Bennett HA, He YD, Dai H, Walker WL, Hughes TR *et al.*: **Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles**. *Science* 2000, **287**:873-880.
6. Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D, Futcher B: **Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization**. *Mol Biol Cell* 1998, **9**:3273-3297.
7. Lashkari DA, DeRisi JL, McCusker JH, Namath AF, Gentile C, Hwang SY, Brown PO, Davis RW: **Yeast microarrays for genome wide parallel genetic and gene expression analysis**. *Proc Natl Acad Sci USA* 1997, **94**:13057-13062.
8. Zhang J, Campbell RE, Ting AY, Tsien RY: **Creating new fluorescent probes for cell biology**. *Nat Rev Mol Cell Biol* 2002, **3**:906-918.
9. • Galarneau A, Primeau M, Trudeau LE, Michnick SW: **β-Lactamase protein fragment complementation assays as *in vivo* and *in vivo***. *Nat Biotechnol* 2002, **20**:619-622.

Describes a particularly sensitive and versatile PCA based on an enzyme reporter.

10. •• Michnick SW, Remy I, C-Valois F-X, V-Belisle A, Pelletier JN: **Detection of protein–protein interactions by protein fragment complementation strategies.** In *Methods in Enzymology*. Edited by Abelson JN, Emr SD, Thorner J: London, UK: Academic Press; 2000, vol 328:208-230.

Everything you ever wanted to know about how PCAs are designed, optimized and used in a variety of applications to studies of protein–protein interactions and library screening.

11. Pelletier JN, Arndt KM, Plückthun A, Michnick SW: **An *in vivo* competition strategy for the selection of optimized protein–protein interactions**. *Nat Biotechnol* 1999, **17**:683-690.
12. Remy I, Wilson IA, Michnick SW: **Erythropoietin receptor activation by a ligand-induced conformation change**. *Science* 1999, **283**:990-993.
13. • Remy I, Pelletier JN, Galarneau A, Michnick SW: **Protein interactions and library screening with protein fragment complementation strategies.** In *Protein–Protein interactions: A Molecular Cloning Manual*. Edited by Golemis EA: New York: Cold Spring Harbor Laboratory Press; 2001:449-475.

A practical manual for implementing several PCAs with references to materials, products and instruments needed for measurements.

14. •• Remy I, Michnick SW: **Visualization of biochemical networks in living cells**. *Proc Natl Acad Sci USA* 2001, **98**:7678-7683.

A key article describing a proof-of-principle PCA pathway mapping strategy.

15. Pelletier JN, Remy I, Michnick SW: **Protein-fragment complementation assays: a general strategy for the *in vivo* detection of protein–protein interactions.** *J Biomol Tech* 1998, **accession number S0012**.
16. Ghosh I, Hamilton AD, Regan L: **Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein**. *J Am Chem Soc* 2000, **122**:5658-5659.
17. Paulmurugan R, Umezawa Y, Gambhir SS: **Noninvasive imaging of protein–protein interactions in living subjects by using reporter protein complementation and reconstitution strategies**. *Proc Natl Acad Sci USA* 2002, **99**:15608-15613.

18. Paulmurugan R, Gambhir SS: **Monitoring protein–protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation**. *Anal Chem* 2003, **75**:1584-1589.

19. Hu CD, Chinenov Y, Kerppola TK: **Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation**. *Mol Cell* 2002, **9**:789-798.

20. Wehrman T, Kleaveland B, Her JH, Balint RF, Blau HM: **Protein–protein interactions monitored in mammalian cells via complementation of β-lactamase enzyme fragments**. *Proc Natl Acad Sci USA* 2002, **99**:3469-3474.

21. • Spotts JM, Dolmetsch RE, Greenberg ME: **Time-lapse imaging of a dynamic phosphorylation-dependent protein–protein interaction in mammalian cells**. *Proc Natl Acad Sci USA* 2002, **99**:15142-15147.
Presents results showing that the β-lactamase PCA can be used to monitor protein–protein interactions that are activated by signal transduction pathways and to follow these events in real time.

22. Gegg CV, Bowers KE, Matthews CR: **Probing minimal independent folding units in dihydrofolate reductase by molecular dissection**. *Protein Sci* 1997, **6**:1885-1892.

23. Ozawa T, Nogami S, Sato M, Ohya Y, Umezawa Y: **A fluorescent indicator for detecting protein–protein interactions *in vivo* based on protein splicing**. *Anal Chem* 2000, **72**:5151-5157.

24. Rossi F, Charlton CA, Blau HM: **Monitoring protein–protein interactions in intact eukaryotic cells by β-galactosidase complementation**. *Proc Natl Acad Sci USA* 1997, **94**:8405-8410.

25. • Vidal M: **A biological atlas of functional maps**. *Cell* 2001, **104**:333-339.
A simple and thoughtful perspective on pathway mapping and functional annotations based on protein–protein interactions.

26. Perham RN: **Self-assembly of biological macromolecules**. *Philos Trans R Soc Lond B Biol Sci* 1975, **272**:123-136.

27. Pawson T, Nash P: **Protein-protein interactions define specificity in signal transduction**. *Genes Dev* 2000, **14**:1027-1047.

28. Remy I, Michnick SW: **Clonal selection and *in vivo* quantitation of protein interactions with protein fragment complementation assays**. *Proc Natl Acad Sci USA* 1999, **96**:5394-5399.

29. D'Andrea AD, Lodish HF, Wong GG: **Expression cloning of the murine erythropoietin receptor**. *Cell* 1989, **57**:277-285.

30. Lin HY, Wang XF, Ng-Eaton E, Weinberg RA, Lodish HF: **Expression cloning of the TGF-β type II receptor, a functional transmembrane serine/threonine kinase**. *Cell* 1992, **68**:775-785.

31. Aruffo A, Seed B: **Molecular cloning of a CD28 cDNA by a high-efficiency COS cell expression system**. *Proc Natl Acad Sci USA* 1987, **84**:8573-8577.

32. Sako D, Chang XJ, Barone KM, Vachino G, White HM, Shaw G, Veldman GM, Bean KM, Ahern TJ, Furie B *et al.*: **Expression cloning of a functional glycoprotein ligand for P-selectin**. *Cell* 1993, **75**:1179-1186.

33. Weston CR, Davis RJ: **Signal transduction: signaling specificity — a complex affair**. *Science* 2001, **292**:2439-2440.

34. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y, Zhang J, Ma Y, Taylor SS, Tsien RY: **A comprehensive two-hybrid analysis to explore the yeast protein interactome**. *Proc Natl Acad Sci USA* 2001, **98**:4569-4574.

35. Ito T, Tashiro K, Muta S, Ozawa R, Chiba T, Nishizawa M, Yamamoto K, Kuhara S, Sakaki Y, Giaever G *et al.*: **Toward a protein–protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins**. *Proc Natl Acad Sci USA* 2000, **97**:1143-1147.

36. Drees BL: **Progress and variations in two-hybrid and three-hybrid technologies**. *Curr Opin Chem Biol* 1999, **3**:64-70.

37. Walhout AJ, Sordella R, Lu X, Hartley JL, Temple GF, Brasch MA, Thierry-Mieg N, Vidal M: **Protein interaction mapping in *C. elegans* using proteins involved in vulval development**. *Science* 2000, **287**:116-122.

38. Vidal M, Legrain P: **Yeast forward and reverse 'n'-hybrid systems**. *Nucleic Acids Res* 1999, **27**:919-929.

39. Fields S, Song O: **A novel genetic system to detect protein–protein interactions**. *Nature* 1989, **340**:245-246.

40. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P *et al.*: **A comprehensive analysis of protein–protein interactions in *Saccharomyces cerevisiae***. *Nature* 2000, **403**:623-627.

41. •• Remy I, Michnick S: **A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways.** *Methods* 2003, in press.
A description of how to screen cDNA libraries by PCA, discussing both details of a fluorescence-activated cell sorting based screening strategy, followed by functional validation studies using pharmacological profile and subcellular localization analyses.

42. Hall-Jackson CA, Eyers PA, Cohen P, Goedert M, Boyle FT, Hewitt N, Plant H, Hedge P: **Paradoxical activation of Raf by a novel Raf inhibitor**. *Chem Biol* 1999, **6**:559-568.

43. Hartwell LH, Hopfield JJ, Leibler S, Murray AW: **From molecular to modular cell biology**. *Nature* 1999, **402**:C47-C52.