Application of protein-fragment complementation assays in cell biology

Ingrid Remy and Stephen W. Michnick

Département de Biochimie, Université de Montréal, Succursale Center-Ville, Montréal, QC, Canada

We have developed a general experimental strategy that enables the quantitative detection of dynamic protein-protein interactions in intact living cells, based on protein-fragment complementation assays (PCAs). In this method, protein-protein interactions are coupled to refolding of enzymes from cognate fragments where reconstitution of enzyme activity acts as the detector of a protein interaction. Here we discuss the application of PCA to different aspects of cell biology.

Introduction

Biochemical pathways are really systems of dynamically assembling and disassembling protein complexes, and thus, much of modern biological research is concerned with how, when, and where proteins interact with other proteins involved in biochemical processes. The demand for simple approaches to study protein-protein interactions, particularly on a large-scale, has grown recently with the progress in genome projects, since the associating of unknown with known gene products provides one crucial way of establishing the function of a gene. It was with this challenge in mind that our laboratory developed protein-fragment complementation assays (PCAs). In this strategy, two proteins of interest (proteins A and B) are fused to complementary fragments of a reporter protein (an enzyme, fluorescent protein, etc.). If the proteins A and B interact, the reporter fragments are brought together, fold into the native structure of the reporter and reconstitute its activity (Figure 1). PCA reporter proteins have been chosen as those producing a variety of detectable activities, including fluorescent, luminescent, and colorimetric signals, as well as simple survival selection assays (1–14). We have demonstrated that the PCA strategy has the following capabilities: (*i*) it allows detection of protein-protein interactions in vivo and in vitro in any cell type; (*ii*) it allows detection of protein-protein interactions in appropriate subcellular compartments or organelles; (*iii*) it allows detection of interactions that are specifically induced in response to developmental, nutritional, environmental, or hormone-induced signals; (*iv*) it allows monitoring of kinetic and equilibrium aspects of protein assembly in cells; and (*v*) it allows screening for novel protein-protein interactions in any cell type (2,3,6,9,15–19).

Principle

We demonstrated the principle of PCA starting with the enzyme dihydrofolate reductase (DHFR) as a reporter (1). It was obvious that if the folding of the enzyme from its fragments (as detected by reconstitution of activity) was absolutely dependent on the binding together of the interacting proteins, then the system described is, in fact, a detector of the interactions. We and others have since demonstrated that this principle can be generalized to a number of enzymes including *Gaussia* and *Renilla* luciferases, TEM β-lactamase, as well as green fluorescent protein (GFP) and its variants (1–14). A crucial feature of PCA fragments is that they are designed not to fold spontaneously without being brought into close proximity by the interaction of the proteins to which they are fused (1,20). If spontaneous folding occurred, PCA simply would not work. Spontaneous folding would lead to a false positive signal, a situation that would hopelessly confound the interpretation

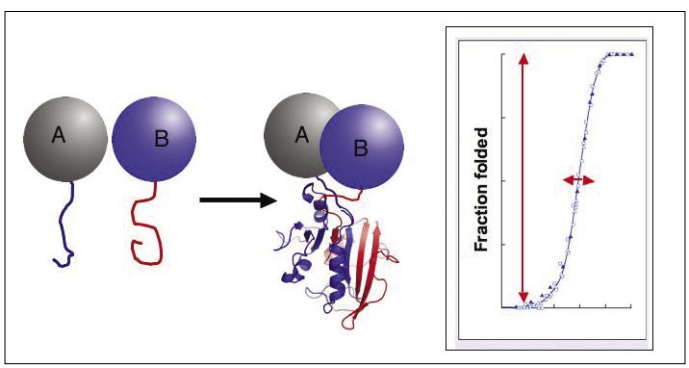

Figure 1. Protein complex dynamics can be studied by fusing proteins of interest (proteins A and B represented by two spheres) to complementary fragments of a reporter protein. If the two proteins interact, the reporter fragments are brought together, fold into the native structure of the reporter protein, and its activity is reconstituted (left). These protein-fragment complementation assays (PCAs) have physical characteristics that make them particularly useful as reporters of dynamic protein complexes. On the right is a protein folding curve where the x-axis is some variable parameter (e.g., concentration of one fragment relative to another). The high cooperativity of this process (extremely sharp increase in fraction of folded species over a very narrow range) means that the assays have an enormous dynamic range, making detection of a complex a virtual all-or-none phenomenon. This contrasts with methods such as fluorescence resonance energy transfer (FRET), which has very low dynamic range and requires careful optimization of a number of parameters. In contrast, measuring protein complex formation by PCA is no more difficult than measuring activity of the intact reporter enzyme.

of library screens in vivo (anticipated to be an important application). In contrast to PCA, there are assay systems based on β-galactosidase and split inteins that resemble PCA, but which are conceptually and practically different (21,22). In both cases, well-known naturally occurring and spontaneously associating subunits of the enzymes are fused to interacting proteins. 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.

Limitations

The PCA strategy is general, in the sense that it is not restricted to a single enzyme reporter, and it has been devised in several different forms, each of which is best suited to address a specific question. For instance, simple survival-selection PCAs, such as those based on DHFR, are most useful for library selection, while luminescence or fluorescence readout PCAs are best for studies of the spatial and temporal dynamics of protein complexes. Because the fusion proteins can be expressed in cells that are relevant models for studying a specific biochemical pathway, they are likely in their native biological state including the correct posttranslational modifications (obviously the PCA fragments themselves must not interfere with targeting or modification of the proteins, and this must be tested).

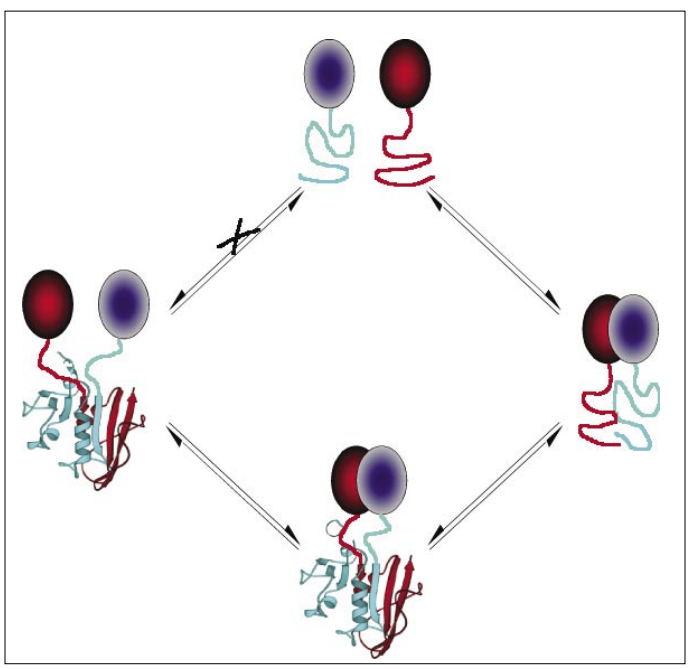

Figure 2. The protein-fragment complementation assay (PCA) strategy requires that unnatural peptide fragments be chosen that are unfolded (ribbons) prior to association of fused interacting proteins (spheres). This prevents spontaneous association of the fragments (pathway X) that can lead to a false signal. Equally, fragments are selected for which spontaneous unfolding of fragments should occur when the protein complex is disrupted (left side).

Among the simplest and therefore most popular PCAs are those based on fluorescent proteins (such as GFP and variants), because signal is provided by the intrinsic fluorophore (7–9,14,15,17,23). However, fluorescent proteins must be expressed at high levels to assure that signal is above background cellular fluorescence, and fluorescent protein PCAs have been demonstrated to be irreversible, which can be useful (trapping and visualizing rare complexes) but could also lead to misinterpretation of turnover or localization of interacting proteins (8,23,24). On the other hand, PCAs based on DHFR and β-lactamase as reporters have been demonstrated, based on indirect evidence, to be reversible following disruption of interactions, while a PCA based on *Gaussia* luciferase has been directly shown to be reversible (2,3,6). Reversibility of PCA thus allows for the detection of kinetic and equilibrium aspects of protein complex assembly and disassembly in living cells (Figure 2).

Standard Controls for a PCA Study

The PCA strategy requires the fragments of the reporter protein to assemble and fold after the proteins of interest have formed a complex. The assembly and correct folding of the reporter is dependent on the recovery of both the structural geometry intrinsic to the reporter proteins and of the complex formed by the interacting proteins. This is one of the major distinctions of the PCA assays compared with fluorescence resonance energy transfer (FRET) or bioluminescence resonance energy transfer (BRET) or yeast two-hybrid assays, and this feature allowed us to perform a structure-based study of the erythropoietin receptor (19). We typically insert a 10-amino acid flexible polypeptide linker consisting of $(Gly.Gly.Gly.Gly.Ser)_2$ between the protein of interest and the PCA reporter fragment (for both fusions). This linker was chosen because it is the most flexible possible, and we have empirically observed that linkers of this length are sufficiently long to allow for fragments to find each other and fold, regardless of the size of the interacting proteins to which the fragments are fused (16).

To ensure that nonspecific responses do not occur, a set of controls should be performed. These controls could include the following, although the first is the most important: (*i*) Non-interacting proteins. A PCA response should not be observed if non-interacting proteins are used as PCA partners; nor should overexpression of a non-interacting protein alone compete for the known interaction. (*ii*) Partner protein interface mutations. A point or deletion mutation of a partner that is known to disrupt an interaction should also prevent a PCA response. (*iii*) Competition. A PCA response should be diminished by the simultaneous overexpression of one or the other of the interacting proteins that is not fused to a complementary PCA fragment. (*iv*) Fragment swapping. An observed interaction between two proteins should occur even if the proteins are swapped with the respective reporter fragments.

Application of PCA in Protein Design: Library vs. Library Screening for Optimally Interacting Proteins

Among the first applications of a PCA was to a protein design problem. The DHFR PCA assay was used in *Escherichia coli* to screen two libraries of complementarily designed leucine zipper-forming sequences with 10^{10} potential interacting pairs, of which we could practically cover 10^6. We demonstrated that the PCA screen selected for both optimal binding specificity, as well as solubility and expression of interacting zippers (18,25). The most important feature of this approach is that it was possible to simultaneously screen two libraries against each other, a process not easily achieved with comparable yeast two-hybrid screens. The simplicity of this approach and specific nature of the information obtained about the design strategy suggest broad utility of the DHFR PCA in protein design and directed evolution experiments. It also shows that PCA complements phage display strategies, since the entire selection, optimization, and stringency tests are done in vivo, making this approach easily executed.

Application of PCA to cDNA Library Screening in Mammalian Cells

A first step in defining the function of a novel gene product is to determine its interactions with other gene products. However, a purely protein interaction-based screening approach (such as yeast two-hybrid) is limited, because it only tells you that two proteins interact, while not providing any other information that could link a protein to its function. Therefore, we have shown that PCA can be used in a cDNA library screening strategy that combines a simple cell-based protein interaction screen with specific functional assays that provide initial validation of the biological relevance of the interaction (9). The first step consists of screening for physical interactions between the bait and a library of cDNA-encoded prey proteins, by monitoring the reconstitution of the PCA reporter in intact living cells. An important feature of this first step is that interactions can be detected directly and between full-length proteins in cells in which the bait protein normally functions, thus assuring that necessary subcellular targeting, posttranslational modifications, and interactions with other proteins can occur. Obviously, for experimental validity, the PCA fragments must be demonstrated not to interfere with targeting or modification of the proteins. In the second step, the protein interaction can be functionally validated, as follows: first, the protein interaction, detected by PCA, must be perturbed by agents, such as hormones or specific inhibitors, that are known to modulate the specific biochemical pathway in which the proteins participate. We have demonstrated this for the DHFR PCA and used this property to map signaling pathways in living mammalian cells (16). Second, the subcellular localization of the protein interaction, again detected by PCA, might be altered by agents that modulate the pathway. Thus, the PCA-based screening strategy combines a simple screening step with direct functional assays. We and others have applied this strategy to the identification of novel substrates or regulators of the serine/threonine protein kinase, PKB/Akt (9,15,26,27).

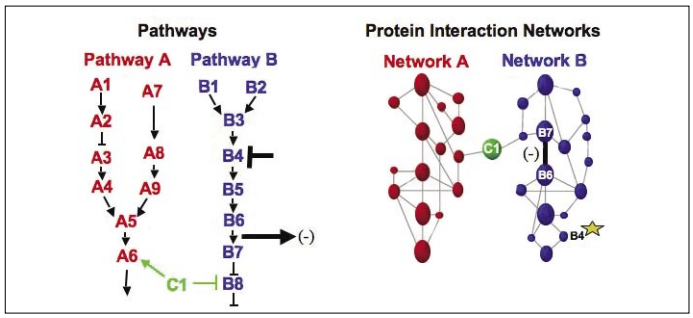

Figure 3. Using dynamic perturbations of protein-protein interactions to infer the organization of pathways. (Left) The actions of an inhibitory perturbing agent [e.g., a small molecule inhibitor, classic genetic approach, or small interfering RNA (siRNA)] that acts on protein B4 (T-bar), is detected downstream by a change in the interaction of proteins B6 and B7 with each other (arrow). In this case, the effect of the perturbation is a decrease in the number of interacting protein (-) as detected by a reporter of that interaction (output signal of the interaction as detected by PCA sentinel for instance). However, the effect could equally be positive, depending on the consequences of inhibiting the upstream protein. (Right) Within the protein interaction network for pathway B, a perturbation of protein B4 (star) somehow propagates through the network to somehow affect the link (wide bar) between proteins B6 and B7. This does not imply that protein B4 physically interacts with either B6 or B7; the propagation of an affect through the protein interaction network may be due to direct physical links or to enzymatic processes not obvious in the network.

Using PCA as a Molecular Ruler: Receptor Studies

A special feature of PCA strategies is that, if we know the three-dimensional structure of the reporter enzyme, it is possible to accurately predict how close together the fragments must be to assure that the enzyme will fold correctly and have a measurable activity. This fact was put to work to test a structural allosteric model for activation of the dimeric erythropoietin receptor (EpoR) using the DHFR PCA, and the approach could be extended to studying allosteric transitions in dimeric or multimeric protein interfaces (19). In the EpoR case, the receptor dimer transmembrane domains were shown to be separated by 73 Å, as observed in the crystal structure of unligated EpoR. It was reasoned that if this inactive state existed on the membrane of a living cell, then DHFR fragments fused to the C termini of the transmembrane domains would fold only if a ligand induced a conformation change that allow the fragments to come close enough together to assure that the precise three-dimensional structure of DHFR could be formed

(19,28). This would require that the N termini of the fragments be 8 Å apart. Insertion of flexible linker peptides between the transmembrane domain and DHFR fragments allowed us to probe the distance between the insertion points of the extracellular domain dimer and confirm that linkers long enough to span 73 Å were needed for DHFR to fold from its fragments.

Mapping Biochemical Networks

Cellular biochemical machineries for metabolism, signaling cascades, and cell cycle are examples of dynamically assembling and disassembling macromolecular complexes. These are defined by grouping interacting proteins according to their similar responses to a set of perturbations (hormones, metabolites, enzyme inhibitors, etc.). Protein-protein interactions can be used to link a protein of unknown function to proteins that are known to be involved in a known biochemical process. We have demonstrated that pharmacological profiling (monitoring effects of pathway-specific drugs and protein hormones on protein-protein interactions) and determining the cellular location of protein-protein interactions can be achieved using PCAs (9,15–17,26). Analysis of these results allow for a representation of how biochemical networks evolve in time and space and in response to specific stimuli. As a proof of principle, we reported the application of this strategy to the mapping of a signal transduction pathway mediated by receptor tyrosine kinases (RTKs) (16). The pharmacological profiles and cellular location of interactions we observed allowed us to place each gene product at its relevant point in the pathways (Figure 3). From the results of our analysis, a map of the organization of the RTK network emerged that was consistent with existing models, but that also included several novel interactions. The ability to monitor a network of protein interactions in living cells containing all of the components of the underlying pathway studied revealed hidden connections, not observed before, in spite of intense scrutiny of this network. The results presented demonstrate that the PCA strategy has the features necessary for a general gene function validation and pathway mapping strategy. A recent application of a larger set of PCAs allowed for development of a general approach to link the actions of drugs on specific signaling pathways and to detect unforeseen activities of drugs (17).

Conclusion

The development and application of PCA are still in progress. For instance, in addition to the limited, though informative, sets of applications described here, the strategy is being applied to large-scale screening of entire genomes. More sophisticated problems of protein design and protein folding are being explored, including studies of the factors that control the selection of sequences for optimal interactions between proteins, proteins and nucleic acids, and proteins and small organic molecules. PCA is a very general and flexible experimental approach, and thus we should expect to see a growing number of novel applications of this basic tool to molecular and cell biology in the near future.

Acknowledgments

Stephen Michnick holds the Canada Research Chair in Integrative Genomics. Research cited from our lab was funded by the Canadian Institutes of Health Research.

References

1. **Pelletier, J.N., F.X. Campbell-Valois, and S.W. Michnick.** 1998. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. Proc. Natl. Acad. Sci. USA *95*:12141-12146.
2. **Remy, I. and S.W. Michnick.** 1999. Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays. Proc. Natl. Acad. Sci. USA *96*:5394-5399.
3. **Galarneau, A., M. Primeau, L.E. Trudeau, and S.W. Michnick.** 2002. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. Nat. Biotechnol. *20*:619-622.
4. **Wehrman, T., B. Kleaveland, J.H. Her, R.F. Balint, and H.M. Blau.** 2002. Protein-protein interactions monitored in mammalian cells via complementation of beta-lactamase enzyme fragments. Proc. Natl. Acad. Sci. USA *99*:3469-3474.
5. **Spotts, J.M., R.E. Dolmetsch, and M.E. Greenberg.** 2002. Time-lapse imaging of a dynamic phosphorylation-dependent protein-protein interaction in mammalian cells. Proc. Natl. Acad. Sci. USA *99*:15142-15147.
6. **Remy, I. and S.W. Michnick.** 2006. A highly sensitive protein-protein interaction assay based on *Gaussia* luciferase. Nat. Methods *3*:977-979.
7. **Ghosh, I., A.D. Hamilton, and L. Regan.** 2000. Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. J. Am. Chem. Soc. *122*:5658-5659.
8. **Hu, C.D., Y. Chinenov, and T.K. Kerppola.** 2002. Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation. Mol. Cell *9*:789-798.
9. **Remy, I. and S.W. Michnick.** 2004. A cDNA library functional screening strategy based on fluorescent protein complementation assays to identify novel components of signaling pathways. Methods *32*:381-388.
10. **Remy, I., F.X. Campbell-Valois, G. Ghaddar, S. Aquin, and S.W. Michnick.** 2005. Detection of protein interactions and library screening with protein-fragment complementation assays, p. 637-672. *In* Protein-Protein Interactions: A Molecular Cloning Manual, 2nd ed. CSH Laboratory Press, Cold Spring Harbor, NY.
11. **Paulmurugan, R. and S.S. Gambhir.** 2003. Monitoring protein-protein interactions using split synthetic renilla luciferase protein-fragment-assisted complementation. Anal. Chem. *75*:1584-1589.
12. **Paulmurugan, R., Y. Umezawa, and S.S. Gambhir.** 2002. Noninvasive imaging of protein-protein interactions in living subjects by using reporter protein complementation

and reconstitution strategies. Proc. Natl. Acad. Sci. USA *99*:15608-15613.

13. **Luker, K.E., M.C. Smith, G.D. Luker, S.T. Gammon, H. Piwnica-Worms, and D. Piwnica-Worms.** 2004. Kinetics of regulated protein-protein interactions revealed with firefly luciferase complementation imaging in cells and living animals. Proc. Natl. Acad. Sci. USA *101*:12288-12293.
14. **Jach, G., M. Pesch, K. Richter, S. Frings, and J.F. Uhrig.** 2006. An improved mRFP1 adds red to bimolecular fluorescence complementation. Nat. Methods *3*:597-600.
15. **Remy, I., A. Montmarquette, and S.W. Michnick.** 2004. PKB/Akt modulates TGF-beta signalling through a direct interaction with Smad3. Nat. Cell Biol. *6*:358-365.
16. **Remy, I. and S.W. Michnick.** 2001. Visualization of biochemical networks in living cells. Proc. Natl. Acad. Sci. USA *98*:7678-7683.
17. **Macdonald, M.L., J. Lamerdin, S. Owens, B.H. Keon, G.K. Bilter, Z. Shang, Z. Huang, H. Yu, et al.** 2006. Identifying off-target effects and hidden phenotypes of drugs in human cells. Nat. Chem. Biol. *2*:329-337.
18. **Pelletier, J.N., K.M. Arndt, A. Pluckthun, and S.W. Michnick.** 1999. An in vivo library-versus-library selection of optimized protein-protein interactions. Nat. Biotechnol. *17*:683-690.
19. **Remy, I., I.A. Wilson, and S.W. Michnick.** 1999. Erythropoietin receptor activation by a ligand-induced conformation change. Science *283*:990-993.
20. **Gegg, C.V., K.E. Bowers, and C.R. Matthews.** 1997. Probing minimal independent folding units in dihydrofolate reductase by molecular dissection. Protein Sci. *6*:1885-1892.
21. **Rossi, F., C.A. Charlton, and H.M. Blau.** 1997. Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. Proc. Natl. Acad. Sci. USA *94*:8405-8410.
22. **Ozawa, T., S. Nogami, M. Sato, Y. Ohya, and Y. Umezawa.** 2000. A fluorescent indicator for detecting protein-protein interactions in vivo based on protein splicing. Anal. Chem. *72*:5151-5157.
23. **Magliery, T.J., C.G. Wilson, W. Pan, D. Mishler, I. Ghosh, A.D. Hamilton, and L. Regan.** 2005. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. J. Am. Chem. Soc. *127*:146-157.
24. **Nyfeler, B., S.W. Michnick, and H.P. Hauri.** 2005. Capturing protein interactions in the secretory pathway of living cells. Proc. Natl. Acad. Sci. USA *102*:6350-6355.
25. **Arndt, K.M., J.N. Pelletier, K.M. Muller, T. Alber, S.W. Michnick, and A. Pluckthun.** 2000. A heterodimeric coiled-coil peptide pair selected in vivo from a designed library-versus-library ensemble. J. Mol. Biol. *295*:627-639.
26. **Remy, I. and S.W. Michnick.** 2004. Regulation of apoptosis by the Ft1 protein, a new modulator of protein kinase B/Akt. Mol. Cell. Biol. *24*:1493-1504.
27. **Ding, Z., J. Liang, Y. Lu, Q. Yu, Z. Songyang, S.Y. Lin, and G.B. Mills.** 2006. A retrovirus-based protein complementation assay screen reveals functional AKT1-binding partners. Proc. Natl. Acad. Sci. USA *103*:15014-15019.
28. **Cody, V., J.R. Luft, E. Ciszak, T.I. Kalman, and J.H. Freisheim.** 1992. Crystal structure determination at 2.3 A of recombinant human dihydrofolate reductase ternary complex with NADPH and methotrexate-gamma-tetrazole. Anticancer Drug Des. *7*:483-491.