ProtTech™

Protein identification in the study of protein-protein interactions

With the complete sequencing of human genome and genomes of many other organisms, biological research is entering the post-genomic era. Enormous attention is now increasingly being drawn towards the study of the functions of proteins in the context of cellular molecular machines and functional networks. Study of protein-protein interactions is becoming a key step towards the understanding of the functions of proteins, the cellular pathways, and molecular mechanism of human diseases.

Two different approaches are now widely used in the study of protein-protein interactions: the yeast two-hybrid system based screening and the biochemical approach, which is combination of affinity protein purifications and MS based protein identifications. In the basic version of the yeast two-hybrid system, a bait protein is fused with the DNA binding domain of the yeast GCN4 protein, and the interacting partners are fused with GCN4 activation domain. The interaction is revealed by the turn-on expression of a report gene by alpha complementation. In a biochemical approach, ether a bait protein immobilized in an affinity column is used to directly fish out its interaction proteins, or antibody against the bait protein is used to direct precipitate the whole protein complex.

It is often surprising to find out that interacting proteins identified by yeast two-hybrid system based genetic approach are quite different from that identified by affinity purification –MS protein identification based biochemical approach. For example, in systematic genome-wide yeast two-hybrid screenings, seventeen proteins were identified that interact with Skp1 and Cdc53, both are components of SCF ubiquitin ligase complex, belongs to so-called F-box protein family (Uetz, P. et al. A comprehensive Analysis of Protein-Protein Interactions in Saccharomyces cerevisiae. Nature 430, 623-627. 2000). In a study carried by Raymond Deshaies and collogues using affinity purification and MS based protein identification, sixteen proteins were identified as Skp1 and Cdc53 interacting proteins. However, only three proteins are overlap in both study (Seol, JH. et al. Skp1 forms multiple protein complexes, including RAVE, a regulator of V-ATPase assembly. Nature Cell. Biol. 3, 384-391. 2001).

The above examples clearly demonstrate the need to apply multiple methods in studying protein-protein interactions, because each method has some limitations. Yeast two-hybrid designed mainly to detect the interactions between two proteins or two protein domains, is prone to false positives and false negatives in detecting protein-protein interactions involving multiple components--- which is often the rule than exception in mammalian cells. The force that stabilizes a multiple component protein complex is from cooperative interacting of many domains of multiple proteins in the complex, while each pair wise interaction between two proteins or domains is only at basal level. Such cooperative action is essential since it makes the formation and destruction of cellular protein machines a dynamic process in concert with cell cycle and in response to change of physiological conditions: e.g. phosphorylation at one sites in a protein component of a complex may lead to destabilizing and destruction of whole protein complex. Therefore, yeast two-hybrid system will miss out identifying some interacting proteins that its associations with bait proteins depending on the cooperativity of whole protein complex. On the other hand, if detection threshold is set low, many false positives will show up in such a screening.

There are some other factors that also affect the outcome from a yeast two-hybrid screening. First, protein-protein interaction may depend on certain cellular conditions that do not exist inside a yeast cell. For example, interacting between two proteins requires phosphorylation of Tyr site in one partner, but such protein cannot be phosphorylated due to lack of similar kinase in yeast cells. Second, the conditions used in two-hybrid screening: assembling of protein-protein interaction at the GAL1 DNA binding site, and protein fusions with GAL4p DNA binding domain and activation domain, may also contribute to

some artifacts.

Biochemical approach based on affinity protein purification and MS protein identifications also has some limitations. The most challenging problem is how to purify a protein complex. Unfortunately, protein complex often exist in a cell at a level that requires 10-5 to 10-7 folds of purification. Even a good affinity column may only yield 10-3 folds purifications. If only one affinity purification step is used and combing effort of other purification steps does not give enough folds of purification, there is a real danger that the real interacting proteins are buried among dozens or even hundreds of contaminant proteins. Although it is still possible to use LC-MS/MS method to identify interacting proteins among contaminating proteins, such protein identification is almost impossible for MALDI-TOF based fingerprinting method. Furthermore, following up functional study will also become very difficult.

In summary, in the study of protein-protein interaction, yeast two-hybrid system and affinity protein purification –MS protein identification base biochemical approach can complement each other to give a more complete picture about the functions and regulations of complex and dynamic cellular protein machines and pathways.